



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

Ανάπτυξη Διαδικτυακού Συστήματος για Σηματολογική Περιγραφή και Ανάκτηση Ιατρικών Απεικονιστικών Δεδομένων

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

των

**ΓΕΩΡΓΙΟΥ ΚΥΡΓΙΑΖΟΥ
ΗΛΙΑ ΓΕΡΟΣΤΑΘΟΠΟΥΛΟΥ**

Επιβλέπουσα : Κωνσταντίνα Νικήτα,
Καθηγήτρια Ε.Μ.Π.

Αθήνα, Οκτώβριος 2010



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ
ΥΛΙΚΩΝ

Ανάπτυξη Διαδικτυακού Συστήματος για Σηματολογική Περιγραφή και Ανάκτηση Ιατρικών Απεικονιστικών Δεδομένων

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΤΩΝ

ΓΕΩΡΓΙΟΥ ΚΥΡΓΙΑΖΟΥ
ΗΛΙΑ ΓΕΡΟΣΤΑΘΟΠΟΥΛΟΥ

Επιβλέπουσα : Κωνσταντίνα Νικήτα,
Καθηγήτρια Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 25^η Οκτωβρίου 2010.

.....
Κωνσταντίνα Νικήτα
Καθηγήτρια Ε.Μ.Π.

.....
Κώστας Κοντογιάννης
Αν. Καθηγητής Ε.Μ.Π.

.....
Γιώργος Στάμου
Λέκτορας Ε.Μ.Π.

Αθήνα, Οκτώβριος 2010

.....

ΓΕΩΡΓΙΟΣ Κ. ΚΥΡΓΙΑΖΟΣ

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

.....

ΗΛΙΑΣ Θ. ΓΕΡΟΣΤΑΘΟΠΟΥΛΟΣ

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

© 2010 – All rights reserved

Περίληψη

Ο σκοπός της παρούσας διπλωματικής εργασίας ήταν η σημασιολογική προσέγγιση της περιγραφής ιατρικών εικόνων. Η ανάγκη επισημείωσης του διαρκώς αυξανόμενου όγκου δεδομένων των ιατρικών εικόνων είναι αναγνωρισμένη από ειδικούς του χώρου για διάφορους λόγους, ανεξαρτήτως αν αυτοί είναι κλινικής φύσης, ερευνητικής ή εκπαιδευτικής. Το πλούσιο μεν, λανθάνον δε, περιεχόμενο πληροφορίας των ιατρικών εικόνων μπορεί να καταστεί ρητό και σαφές με τη χρήση καλώς ορισμένων οντολογιών.

Για το σκοπό αυτό σχεδιάστηκε και αναπτύχθηκε ολοκληρωμένο διαδικτυακό σύστημα για την σημασιολογική επισημείωση και αναζήτηση ιατρικών εικόνων, χρησιμοποιώντας τεχνολογίες Σημασιολογικού Ιστού. Η εξέλιξη του Σημασιολογικού Ιστού προσφέρει πλέον τη δυνατότητα απομακρυσμένης πρόσβασης, μέσω διαδικτυακών υπηρεσιών, σε οντολογίες και ταξινομήσεις, που έχουν καθοριστεί σε διεθνές επίπεδο για το χώρο της ιατρικής, όπως η οντολογία ανατομίας FMA (Foundational Model of Anatomy) και η ταξινόμηση ασθενειών ICD-10 (International Classification of Diseases – 10th Revision). Αξιοποιώντας κατάλληλα αυτή την υποδομή, το υλοποιημένο σύστημα αποκτά πρόσβαση στο απαραίτητο οντολογικό πλαίσιο αναφοράς, με αποτέλεσμα να επιτρέπει την αξιοποιήσιμη επισημείωση και την αποδοτική αναζήτηση των ιατρικών εικόνων με χρήση όρων ανατομίας και ασθενειών.

Συγκεκριμένα, το υλοποιημένο διαδικτυακό σύστημα προσφέρει δύο βασικές διεπιφάνειες: τη διεπιφάνεια επισημείωσης και τη διεπιφάνεια αναζήτησης. Η πρώτη δίνει τη δυνατότητα στο χρήστη να «ανοίξει» ένα αρχείο, τύπου DICOM (Digital Imaging and Communications in Medicine), και να επεξεργαστεί τα περιεχόμενα του, δηλαδή την εικόνα και τα βασικά πεδία της επικεφαλίδας. Κατά την επεξεργασία ο χρήστης μπορεί να σημειώσει τα επιθυμητά ROIs (Regions Of Interest – Περιοχές Ενδιαφέροντος) πάνω στην εικόνα, τα οποία στη συνέχεια επισημαίνει με κατάλληλους όρους ανατομίας και ασθενειών. Καθ' όλη τη διαδικασία, το σύστημα υποβοηθά το χρήστη στην επιλογή όρου επισημείωσης με συστάσεις-προτάσεις. Η δεύτερη διεπιφάνεια - διεπιφάνεια αναζήτησης και ανάκτησης εικόνων - επιτρέπει στο χρήστη την αναζήτηση με βάση όρους ανατομίας και ασθενειών, επιστρέφοντας λίστα εικόνων, οι οποίες ικανοποιούν τα σημασιολογικά διευρυμένα κριτήρια αναζήτησης. Ο χρήστης μπορεί προαιρετικά να εφαρμόσει μία σειρά φίλτρων αναζήτησης, όπως ηλικία ασθενή ή τροπικότητα εικόνας (image modality), ώστε να εξειδικεύσει την αναζήτησή του, περιορίζοντας με τον τρόπο αυτό το εύρος των αποτελεσμάτων.

Τα πλεονεκτήματα της προτεινόμενης σημασιολογικής προσέγγισης αποτυπώνονται τόσο στη διαλειτουργικότητα που προσφέρει το κοινό σημασιολογικό πλαίσιο αναπαράστασης, όσο και στον ευφυή μηχανισμό αναζήτησης.

Λέξεις Κλειδιά: Σημασιολογικός Ιστός, οντολογία, ιατρική εικόνα, επισημείωση εικόνας, σημασιολογική αναζήτηση, διαδικτυακές υπηρεσίες

Abstract

The scope of this diploma thesis was the semantic approach to medical image annotation. The need for annotating the growing amount of medical image data is recognized from domain experts for a variety of purposes, regardless if this is medical practice, research or education. The rich, though latent, information content of medical images can be made explicit and formal with the use of well-defined ontologies.

Under this objective, a complete online system of semantic annotation and search on medical images was designed and developed, using Semantic Web technologies. The evolution of Semantic Web now offers the possibility of remote access, via web services, to ontologies and classifications, which are defined in global scale for the medical domain, such as the ontology of anatomy FMA (Foundational Model of Anatomy) and the classification of disease ICD-10 (International Classification of Diseases – 10th revision). Utilizing this infrastructure, the implemented system gains access to the essential ontological framework, resulting in allowing computer-usable annotations and efficient search over medical images using anatomy and disease terms.

Specifically, the implemented online system offers two basic interfaces: the annotation interface and the search interface. In the first one, the user is able to open a file, in DICOM (Digital Imaging and Communications in Medicine) format, and process its contents, namely the image data and the basic header fields. While image processing, the user may select the preferred ROIs (Regions of Interest) on the image. Continuing, the user annotates the ROIs with relevant anatomy and disease terms. During the whole process, the system assists the user to select the annotation terms with meaningful recommendations. The second interface - the image search and retrieval interface – allows the user to perform search by anatomy and disease terms, retrieving a list of images, which satisfy the semantically expanded criteria. The user can optionally apply several search filters, such as patient age or image modality, in order to refine their search by limiting the range of results.

The advantages of the proposed semantic approach are reflected in both the interoperability, enabled by the common semantic representation layer, and the intelligent search mechanism.

Keywords: Semantic Web, ontology, medical image, image annotation, semantic search, web services

Ευχαριστίες

Ευχαριστούμε ιδιαίτερα το φίλο και συνεργάτη μας Γιάννη Στοίτση, που μας θυμίζει ότι κάποιες φορές πρέπει να σταματάς ό,τι κάνεις για να θαυμάσεις το θαύμα της – νέας – ζωής, για την καθοδήγηση και τις πολύτιμες συμβουλές του.

Ευχαριστούμε επίσης την καθηγήτριά μας κ. Κωνσταντίνα Νικήτα για τη συνεργασία, τις ευκαιρίες που μας έδωσε και την εμπιστοσύνη που μας έδειξε.

Στους γονείς μου και την πολυαγαπημένη μου γιαγιά,
Ηλίας

Στην οικογένειά μου για την εμπιστοσύνη τους σε κάθε μου επιλογή
Γ.Κ.

Πίνακας Περιεχομένων

| | | |
|----------|--|-----------|
| 1 | Εισαγωγή..... | 1 |
| 1.1 | Σημασιολογικός Ιστός..... | 1 |
| 1.1.1 | Ο Ιστός Σήμερα..... | 1 |
| 1.1.2 | Το Όραμα του Σημασιολογικού Ιστού..... | 2 |
| 1.2 | Ιατρικά Δεδομένα και Σημασιολογία..... | 3 |
| 1.2.1 | Ιατρικές εικόνες..... | 4 |
| 1.2.2 | Σημασιολογική επισημείωση ιατρικών εικόνων | 5 |
| 1.2.3 | Αναζήτηση ιατρικών εικόνων με σημασιολογικά κριτήρια | 6 |
| 1.3 | Σκοπός Διπλωματικής Εργασίας..... | 9 |
| 2 | Θεωρητικό Υπόβαθρο | 11 |
| 2.1 | Οντολογίες | 11 |
| 2.2 | Πρότυπα και Γλώσσες Σημασιολογικού Ιστού..... | 14 |
| 2.2.1 | RDF | 14 |
| 2.2.2 | SPARQL..... | 16 |
| 2.2.3 | RDF Schema..... | 17 |
| 2.2.4 | OWL..... | 17 |
| 2.3 | Λογική και Συλλογιστική..... | 19 |
| 3 | Σημασιολογικός Ιστός και Ιατρική | 21 |
| 3.1 | Πρότυπα και Κωδικοποιήσεις στην Ιατρική..... | 21 |
| 3.1.1 | DICOM | 22 |
| 3.1.2 | HL7 | 22 |
| 3.1.3 | SNOMED-CT..... | 23 |
| 3.1.4 | openEHR..... | 23 |
| 3.1.5 | ICD | 24 |
| 3.2 | Οντολογίες στην Ιατρική | 24 |
| 3.2.1 | FMA..... | 25 |
| 3.2.2 | GALEN..... | 27 |
| 3.2.3 | Gene Ontology..... | 27 |

| | | |
|----------|---|-----------|
| 3.2.4 | <i>UMLS</i> | 28 |
| 3.3 | Διαλειτουργικότητα στην Ιατρική..... | 29 |
| 4 | Συστήματα Περιγραφής και Αναζήτησης Ιατρικών Εικόνων | 32 |
| 4.1 | PACS..... | 32 |
| 4.2 | Σημασιολογικά Συστήματα..... | 34 |
| 4.2.1 | <i>RadSem</i> | 34 |
| 4.2.2 | <i>Annotation and Image Markup</i> | 38 |
| 4.2.3 | <i>Πληροφοριακό Σύστημα Ιατρικών Εικόνων των Wei και Barnaghi</i> | 39 |
| 5 | Σχεδίαση Συστήματος Σημασιολογικής Επισημείωσης και Αναζήτησης Ιατρικών Εικόνων..... | 42 |
| 5.1 | Απαιτήσεις Συστήματος..... | 42 |
| 5.1.1 | <i>Ομάδες χρηστών</i> | 42 |
| 5.1.2 | <i>Περιγραφές χρήσης</i> | 44 |
| 5.1.3 | <i>Οντολογίες Αναφοράς</i> | 50 |
| 5.2 | Αρχιτεκτονική..... | 51 |
| 5.2.1 | <i>MVC</i> | 51 |
| 5.2.2 | <i>SOA και Οντολογίες Αναφοράς</i> | 53 |
| 5.2.3 | <i>Υβριδικό Αρχιτεκτονικό Πλαίσιο</i> | 54 |
| 6 | Σύστημα Σημασιολογικής Επισημείωσης και Αναζήτησης Ιατρικών Εικόνων | 56 |
| 6.1 | Λεπτομέρειες Υλοποίησης..... | 56 |
| 6.1.1 | <i>Οντολογίες αναφοράς</i> | 56 |
| 6.1.2 | <i>Οντολογία εφαρμογής</i> | 59 |
| 6.1.3 | <i>Διάταξη Συστήματος και Διαδικτυακές Υπηρεσίες</i> | 60 |
| 6.1.4 | <i>Καθορισμός Σημασιολογικής Γειτονιάς</i> | 61 |
| 6.1.5 | <i>Διεπιφάνεια χρήση</i> | 63 |
| 6.2 | Πλατφόρμες και Προγραμματιστικά Εργαλεία | 65 |
| 6.2.1 | <i>Java servlet</i> | 65 |
| 6.2.2 | <i>Java applet</i> | 66 |
| 6.2.3 | <i>jQuery</i> | 66 |
| 6.2.4 | <i>Sesame</i> | 66 |
| 6.2.5 | <i>SwiftOWLIM</i> | 67 |

| | | |
|----------|--|-----------|
| 6.2.6 | <i>Flash</i> | 67 |
| 6.2.7 | <i>Protégé</i> | 68 |
| 6.2.8 | <i>Subversion</i> | 68 |
| 6.2.9 | <i>Πλατφόρμα ανάπτυξης</i> | 69 |
| 6.3 | Έλεγχος με βάση Σενάρια Λειτουργίας..... | 69 |
| 6.3.1 | <i>Προσθήκη νέου χρήστη</i> | 69 |
| 6.3.2 | <i>Επισημείωση εικόνας</i> | 72 |
| 6.3.3 | <i>Αναζήτηση εικόνων</i> | 79 |
| 7 | Συμπεράσματα | 88 |
| 7.1 | Αποτελέσματα..... | 88 |
| 7.2 | Περιορισμοί – Προκλήσεις..... | 90 |
| 7.3 | Μελλοντικές Επεκτάσεις..... | 91 |
| 8 | Αναφορές | 93 |
| 9 | Παραρτήματα | 96 |
| | Παράρτημα Α: Οντολογία εφαρμογής σε μορφή RDF/XML..... | 96 |

1

Εισαγωγή

Στο παρόν εισαγωγικό κεφάλαιο αναλύονται έννοιες με τις οποίες θα καταπιαστεί η παρούσα εργασία όπως αυτή του Σημασιολογικού Ιστού και γίνεται μια παρουσίαση του πεδίου των ιατρικών εικόνων, στο οποίο και θα εστιάσουμε. Στο τέλος παρουσιάζεται ο σκοπός της εργασίας.

1.1 Σημασιολογικός Ιστός

1.1.1 Ο Ιστός Σήμερα

Η επίδραση του Παγκόσμιου Ιστού στην εποχή που διανύουμε είναι εμφανής τόσο στον τρόπο που επικοινωνούμε και ενημερωνόμαστε όσο και στον τρόπο που εργαζόμαστε αλλά και ψυχαγωγούμαστε. Μεγάλη συμβολή στην επιτυχία του Παγκοσμίου Ιστού έχουν οι μηχανές αναζήτησής του, οι οποίες επιτρέπουν την πρόσβαση στις δαιδαλώδεις λεωφόρους των πληροφοριών και την άντληση της επιθυμητής πληροφορίας ανάμεσα στη διάχυτη γνώση του Ιστού. Η λειτουργία τους βασίζεται στην αναζήτηση με κριτήριο κάποιες λέξεις-κλειδιά που εισάγει ο χρήστης. Η μέθοδος αυτή κρύβει όμως κάποιους περιορισμούς, όπως ευαισθησία στο λεξιλόγιο και επιστροφή μεγάλου πλήθους αποτελεσμάτων [1].

Κυριότερο μειονέκτημα της παραπάνω μεθόδου, το οποίο φανερώνεται όλο και περισσότερο λόγω της αύξησης της ποσότητας του περιεχομένου του Ιστού, είναι ότι το νόημα του περιεχομένου αυτού δεν είναι προς το παρόν επεξεργάσιμο από υπολογιστές, αλλά κατανοητό

μόνο στους ανθρώπους. Η μη επεξεργασιμότητα καθιστά αδύνατη την εξόρυξη γνώσης και αυτό συχνά οδηγεί σε σημασιολογικά σφάλματα κατά την αναζήτηση, που ως συνέπεια έχουν τη χαμηλή ανάκληση πληροφορίας ή την ανάκληση ανακριβούς πληροφορίας.

1.1.2 Το Όραμα του Σημασιολογικού Ιστού

Μια λύση στο παραπάνω πρόβλημα είναι η αναπαράσταση του διαδικτυακού περιεχομένου σε μορφή που είναι ευκολότερα επεξεργάσιμη από υπολογιστές και η χρήση νοημών τεχνικών για την εκμετάλλευση αυτών των αναπαραστάσεων. Το όλο εγχείρημα αναφέρεται συχνά ως το όραμα του «Σημασιολογικού Ιστού». Ο Σημασιολογικός Ιστός (Semantic Web) δεν πρόκειται, ωστόσο, για μια νέα τεχνολογία πρόσβασης στη πληροφορία, που αναπτύσσεται παράλληλα με τον υπάρχοντα Ιστό, αλλά για μια νέα κατεύθυνση σταδιακής εξέλιξης του υφιστάμενου Ιστού.

Στο Σημασιολογικό Ιστό η έννοια της πληροφορίας κατέχει πιο σημαντικό ρόλο από αυτόν που διαδραματίζει στα πλαίσια του σημερινού Ιστού. Άλλωστε η λέξη «σημασιολογικός» έχει τη ρίζα της στις ελληνικές λέξεις «σημάδι», «σημαίνω» και «σημαντικός» που αναφέρονται στο νόημα – τη σημασία – του περιεχομένου. Η σαφής αναπαράσταση του νοήματος των πληροφοριών και των εγγράφων είναι το κλειδί που θα ανοίξει το δρόμο για την αυτόματη επεξεργασία των πληροφοριών προς εξαγωγή εννοιολογικών συμπερασμάτων που θα ικανοποιήσουν την ανάγκη για ακριβή και σημασιολογικά ορθή αναζήτηση. Το κοινό πλαίσιο αναπαράστασης, σε αντίθεση με την ελεύθερη μορφή κειμένου, επιτρέπει επίσης την ενοποίηση διαδικτυακών πόρων από ευφυή προγράμματα – πράκτορες και την ενοποίηση συστημάτων ώστε να συνεργάζονται διαλειτουργικά σε παγκόσμιο επίπεδο.

Εμπνευστής και καθοδηγητής στην προσπάθεια θεμελίωσης του Σημασιολογικού Ιστού είναι η Κοινοπραξία Παγκόσμιου Ιστού (World Wide Web Consortium ή W3C) και συγκεκριμένα ο ιδρυτής της Tim Berners-Lee, ο άνθρωπος που επινόησε το World Wide Web στα τέλη της δεκαετίας του '80 [2]. Τα βασικά μοντέλα και πρότυπα, όπως έχουν προταθεί από την W3C, περιλαμβάνουν το «Πλαίσιο Περιγραφής Πόρων» (Resource Description Framework ή RDF) και τις γλώσσες περιγραφής λεξιλογίων και οντολογιών «Σχήμα Περιγραφής Πόρων» (RDF Schema ή RDFS) και «Γλώσσα Οντολογιών Ιστού» (Ontology Web Language ή OWL).

Το σημείο στο οποίο πρέπει να σταθούμε είναι ότι η νέα κατεύθυνση του Ιστού προχωράει πέρα από διασύνδεση ιστοσελίδων μέσω υπερσυνδέσμων και περιγράφει σχέσεις μεταξύ αντικειμένων και ιδιότητες των αντικειμένων σε μοντέλα που είναι ταυτόχρονα κατανοητά από τους ανθρώπους και επεξεργάσιμα από τους υπολογιστές. Τα λόγια του Berners-Lee περιγράφουν το όραμα αυτό γλαφυρά [3]:

«Αν η HTML και ο Ιστός έκαναν όλα τα δικτυακά έγγραφα να φαίνονται σαν ένα μεγάλο βιβλίο, η RDF, RDFS και οι γλώσσες συνεπαγωγής θα κάνουν όλη

αυτήν την πληροφορία στον κόσμο να φαίνεται σαν μια τεράστια βάση δεδομένων.»

1.2 Ιατρικά Δεδομένα και Σημασιολογία

Πέρα από το όραμα του Ιστού Δεδομένων, το οποίο ακόμα μπορεί να φαντάζει μακρινό και ουτοπικό, ο Σημασιολογικός Ιστός έχει εμπνεύσει και οδηγήσει στη δημιουργία πλήθους καινοτόμων τεχνολογιών και εφαρμογών. Οι τεχνολογίες αυτές βρίσκουν εφαρμογή σε πολλά πεδία, μεταξύ των οποίων και στην ιατρική, πεδίο στο οποίο εστιάζει η συγκεκριμένη εργασία.

Τις τελευταίες δύο δεκαετίες καινούριες ανακαλύψεις στους χώρους της ιατρικής, της βιοϊατρικής και της βιολογίας έχουν πραγματοποιηθεί με την βοήθεια πειραμάτων μεγάλης κλίμακας με χρήση αλλά και παραγωγή μεγάλου όγκου δεδομένων. Τα παρατηρούμενα αποτελέσματα διατηρούνται σε κλιμακούμενες, επιστημονικές και ανοιχτές στο κοινό βάσεις δεδομένων ως δομημένα δεδομένα, ενώ τα επιστημονικά ευρήματα καταγράφονται στα επιστημονικά κείμενα. Πολλές από τις προσπάθειες και εξελίξεις στους τομείς της ιατρικής πληροφορικής και της υπολογιστικής βιολογίας εστιάζουν στην διαχείριση των δεδομένων που βρίσκονται σε δομημένη ή ημιδομημένη μορφή.

Με βάση της τελευταίες εξελίξεις, η τάση στην βιοϊατρική έρευνα είναι να μη βασίζεται αποκλειστικά σε τοπικά αποκτώμενα δεδομένα και ευρήματα του στενού κύκλου των επιμέρους εργαστηρίων, αλλά να ενσωματώσει και να χρησιμοποιήσει διαμοιραζόμενες και δημοσίως διαθέσιμες αποθήκες βιοϊατρικών δεδομένων. Επιπλέον, υπάρχει η ανάγκη για εκμετάλλευση των πηγών δεδομένων από διαφορετικά πεδία. Παράδειγμα τέτοιας περίπτωσης αποτελούν μοριακοί βιολόγοι που χρησιμοποιούν τοξικολογικά δεδομένα ή γιατροί που διερευνούν δεδομένα της μοριακής βιολογίας ή των φυσικών επιστημών. Η αρχιτεκτονική και οι τεχνολογίες του Σημασιολογικού Ιστού είναι μία λύση στην αναπαράσταση, διαχείριση και επιτυχή διαβίβαση των δεδομένων τόσο μεταξύ διαφορετικών ινστιτούτων και εργαστηρίων όσο και μεταξύ επιστημόνων διαφορετικών πεδίων και ειδικοτήτων, χωρίς να χάνεται το νόημα τους [4]. Ήδη η βιοϊατρική έρευνα έχει προχωρήσει σε προσεγγίσεις που κάνουν χρήση μεγάλης κλίμακας, σημασιολογικά τεκμηριωμένων βιοϊατρικών δεδομένων.

Επιπρόσθετα προς την προσπάθεια για σημασιολογική ενοποίηση των δομημένων δεδομένων, οι ερευνητές στοχεύουν στην αυτόματη ενσωμάτωση μη δομημένων δεδομένων, όπως τα επιστημονικά κείμενα, έτσι ώστε να εκμεταλλευτούν τα αποτελέσματα και ευρήματα της θεωρίας. Με τον τρόπο αυτό, θα δημιουργηθεί ένα δίκτυο πλήρως τεκμηριωμένων και διασυνδεδεμένων δεδομένων με τυπική σημασιολογία.

Η επισημοποίηση της ιατρικής γνώσης σε ρητά μοντέλα πεδίου με στόχο την επιτυχή αναπαράσταση της, που αποτελεί προϋπόθεση για τη χρήση των σημασιολογικών τεχνολογιών, έχει οδηγήσει τα τελευταία χρόνια στην παραγωγή επισημειωμένων συνόλων δεδομένων, σημασιολογικών δικτύων και οντολογιών. Βέβαια, ο τομέας της ιατρικής και βιοϊατρικής είχε από το παρελθόν παράδοση στην κατηγοριοποίηση της γνώσης μέσω θησαυρών, λεξικών ιατρικών όρων, ταξινομήσεων, κατηγοριοποιήσεων και προτύπων. Παράδειγμα ταξινόμησης αποτελεί η Διεθνής Κατηγοριοποίηση των Ασθενειών (International Classification of Diseases ή ICD). Το ζητούμενο είναι η μοντελοποίηση της ήδη υπάρχουσας γνώσης πεδίου σε τυπικά ορισμένες, ενδεδειγμένες μελετημένες και σημασιολογικά ορθές αναπαραστάσεις που καλούνται οντολογίες. Οι πιο αξιόλογες οντολογίες στο χώρο της ιατρικής μελετώνται στην Ενότητα 3.2.

1.2.1 Ιατρικές εικόνες

Καθώς οι νοσοκομειακές μονάδες προσανατολίζονται όλο και περισσότερο σε μία διαχείριση των ιατρικών δεδομένων βασισμένη στην τεχνολογία των πληροφοριών, όπως ο Ηλεκτρονικός Ιατρικός Φάκελος (Electronic Health Records ή EHR), μεγάλες ποσότητες ιατρικών δεδομένων πρέπει να αποθηκευτούν και να γίνουν προσιτά με χρήση υπολογιστικών συστημάτων [5]. Τα ιατρικά δεδομένα ποικίλουν τόσο στο περιεχόμενο όσο και στη μορφή. Ως προς το περιεχόμενο, κυμαίνονται από καταχωρίσεις σχετικές με το όνομα, την ηλικία και το φύλο του ασθενούς μέχρι αναφορές συγκεκριμένων διαγνώσεων και ιατρικών περιπτώσεων. Η μορφή από την άλλη δεν περιορίζεται σε έγγραφα κειμένου, αλλά περιλαμβάνει και ιατρικές απεικονίσεις.

Η εξέλιξη στην τεχνολογία απεικόνισης έχει αυξήσει δραματικά την ποσότητα των ιατρικών απεικονιστικών δεδομένων που παράγονται από νοσοκομεία, κλινικές, διαγνωστικά κέντρα, φαρμακευτικές εταιρίες και την ακαδημαϊκή ιατρική έρευνα. Τεχνολογίες όπως η αξονική τομογραφία (Computed Tomography ή CT) 4Δ και 64-τομών, μαγνητική τομογραφία συντονισμού (Magnetic Resonance Imaging ή MRI) ολόκληρου σώματος, υπερηχογράφημα (Ultrasound ή US) 4Δ και υβριδική τομογραφία εκπομπής ποζιτρονίου (Positron Emitted Tomography ή PET) με αξονική τομογραφία μπορούν να παράγουν απίστευτη λεπτομέρεια και πλούτο πληροφορίας σχετικά με την ανατομία και τη λειτουργία του ανθρώπινου σώματος και τις συσχετίσεις αυτών με ασθένειες. Ενώ όμως η αύξηση του όγκου των δεδομένων προώθησε την εξέλιξη τεχνικών ανάλυσης απεικονιστικών δεδομένων, όπως «έξυπνων» αλγορίθμων αναγνώρισης προτύπων και τμηματοποίησης, αυτή η πρόοδος δεν οδήγησε σε πιο ευέλικτες ή γενικευμένες τεχνικές κατανόησης του περιεχομένου της εικόνας.

1.2.2 Σημασιολογική επισημείωση ιατρικών εικόνων

Από τα παραπάνω είναι φανερό ότι η κοινότητα ιατρικών εικόνων, και συγκεκριμένα η ακτινολογία, αντιμετωπίζει παρόμοια προβλήματα και έχει παρόμοιες ανάγκες με τις υπόλοιπες κοινότητες της βιοϊατρικής και βιοπληροφορικής στη διαχείριση, αναζήτηση και χρήση των υπέρογκων ποσοτήτων δεδομένων που παράγονται με εκρηκτικό ρυθμό. Ωστόσο, η ιατρική απεικόνιση προβάλλει μοναδικές προκλήσεις οι οποίες δυσχεραίνουν την απευθείας χρήση δοκιμασμένων μεθόδων που εφαρμόζονται ήδη σε μη απεικονιστικά ιατρικά δεδομένα. Συγκεκριμένα [6] :

- Οι εικόνες περιέχουν πλούσιο περιεχόμενο το οποίο δεν είναι ρητά ορισμένο και συνεπακόλουθα προσβάσιμο από μηχανές. Η υπονοούμενη γνώση σχετικά με την ανατομία και τη δυσλειτουργία που εξάγεται από τον παρατηρητή των εικονοστοιχείων της εικόνας γενικά δεν καταγράφεται με δομημένο τρόπο ούτε συνδέεται άμεσα με την εικόνα. Συνεπακόλουθα, οι εικόνες δεν μπορούν να αναζητηθούν με βάση το σημασιολογικό περιεχόμενό τους, για παράδειγμα «βρες όλες τις εικόνες που εικονίζουν μια συγκεκριμένη ανατομία ή παθολογία».
- Η σύνταξη και ορολογία κατά την επισημείωση εικόνων ποικίλει με συνέπεια να μειώνεται η διαλειτουργικότητα. Το περιεχόμενο των ιατρικών εικόνων συνήθως περιγράφεται και αποθηκεύεται σε ελεύθερο κείμενο, καθώς δεν υπάρχουν ευρέως αποδεκτά πρότυπα για την επισημείωση των παρατηρούμενων χαρακτηριστικών, της ανατομίας και της παθολογίας.
- Η συγκεκριμένη πληροφορία που επιθυμείται να περιγραφεί συχνά εξαρτάται από ένα ειδικό πλαίσιο: διαφορετικού τύπου εικόνες αποκτώνται για διαφορετικούς σκοπούς και το είδος των επισημειώσεων (απαιτήσεις επισημείωσης) εξαρτώνται από το πλαίσιο αυτό. Για παράδειγμα, σε εικόνες της κοιλιακής χώρας ενός καρκινοπαθούς υπάρχει ανάγκη επισημειώσεων που περιγράφουν το ήπαρ, που είναι όργανο της κοιλιακής χώρας, και σε περίπτωση καρκίνου στο ήπαρ θα πρέπει να επισημειωθούν και τα όρια του καρκίνου. Ένα εργαλείο επισημείωσης θα πρέπει ιδανικά να λαμβάνει υπόψη τέτοιου είδους εξαρτήσεις έτσι ώστε να παρακινεί το χρήστη να κάνει κατάλληλες επισημειώσεις με βάση τα διαφορετικά πλαίσια.

Κρίνεται λοιπόν απαραίτητο να βρεθούν μέθοδοι που θα καταστήσουν ρητή την υπονοούμενη γνώση που μόνο η ανθρώπινη παρατήρηση μπορεί να εξάγει από τις ιατρικές εικόνες ώστε να διασυνδεθεί η γνώση αυτή με άλλες πηγές από μη απεικονιστικά δεδομένα και να είναι εύκολα προσπελάσιμη από υπολογιστές. Η επισημείωση εικόνων πρέπει όμως να γίνει με τρόπο απλό, σαφή, επίσημο και όσο το δυνατόν πιο ευέλικτο και γενικευμένο. Η χρήση δομών αναπαράστασης γνώσης όπως η RDF και OWL μπορεί να οδηγήσει στη δημιουργία ευέλικτων και τυπικών αναπαραστάσεων γνώσης για τις εν λόγω επισημειώσεις

με σημαντικά πλεονεκτήματα σε σύγκριση με παραδοσιακά μοντέλα δεδομένων όπως η XML και οι σχεσιακές βάσεις δεδομένων. Επίσης, οι οντολογίες που χρησιμοποιούνται από τις εφαρμογές του Σημασιολογικού Ιστού παρέχουν μια κοινή αντίληψη του εκάστοτε πεδίου αναφοράς. Η περιγραφή και επισημείωση που γίνεται με βάση οντολογίες παρέχει ανεξαρτήτου πλατφόρμας, διαλειτουργικές, κλιμακούμενες και επαναχρησιμοποιούμενες δομές αναπαράστασης περιεχομένου. Η διαλειτουργικότητα εμφανίζεται σε πολλά επίπεδα: σε επίπεδο τροπικότητας – κοινές επισημειώσεις τόσο για εικόνες CT όσο και για MRI – σε επίπεδο προτύπων εικόνων, σε επίπεδο νοσοκομείων και διαγνωστικών κέντρων και τέλος σε επίπεδο γλώσσας. Περισσότερα για τα οφέλη της διαλειτουργικότητας μέσω της σημασιολογικής προσέγγισης στην επισημείωση ιατρικών εικόνων αναφέρονται στην Ενότητα 3.3.

1.2.3 Αναζήτηση ιατρικών εικόνων με σημασιολογικά κριτήρια

Η δεδομενοκεντρική προσέγγιση στα μοντέλα γνώσης και περιγραφής που εφαρμόζεται με τη χρήση σημασιολογικών μεθόδων επισημείωσης παρέχει μεγάλα πλεονεκτήματα, εκτός των άλλων, και στην αναζήτηση των ιατρικών εικόνων. Σε μια οντολογική αναπαράσταση ορίζεται η σημασία των εννοιών και οι σχέσεις μεταξύ τους, επιτρέποντας σε πράκτορες λογισμικού να εξάγουν ρητή γνώση από την υπονοούμενη πληροφορία και να παράγουν επιπλέον γνώση με χρήση του λογικού συμπερασμού για τις έννοιες που αναπαριστώνται σε κάθε εικόνα. Οι οντολογίες επίσης επιτρέπουν στους χρήστες να αναζητούν έννοιες από την παρεχόμενη οντολογική δομή και να συγκροτούν βελτιωμένα ερωτήματα, ακόμη και αν δεν έχουν μια ξεκάθαρη ιδέα για το εκάστοτε πεδίο αναζήτησης [7]. Ως παράδειγμα αναφέρεται η προσπάθεια αναζήτησης για εικόνες που παριστούν καρκίνο στην κοιλιακή χώρα, στην οποία το σύστημα ανιχνεύει εικόνες με επισημείωση «καρκίνος στο ήπαρ», γίνεται ο λογικός συμπερασμός ότι το ήπαρ βρίσκεται στην κοιλιακή χώρα και τις επιστρέφει ή ρωτάει το χρήστη αν θέλει να συγκεκριμενοποιήσει το ερώτημά του αναζητώντας εικόνες με καρκίνο σε κάποιο συγκεκριμένο σημείο της κοιλιακής χώρας, όπως στο ήπαρ (βελτιωμένο ερώτημα).

Παρότι η λειτουργία των συστημάτων ανάκτησης εικόνων με βάση το περιεχόμενο (Content-Based Image Retrieval ή CBIR) με σημασιολογικά κριτήρια είναι σε γενικές γραμμές δύσκολη έως ανέφικτη, το πεδίο των ιατρικών εικόνων διαθέτει αρκετά μοναδικά χαρακτηριστικά που καθιστούν τη λειτουργία αυτή εφικτή. Τα θετικά αυτά χαρακτηριστικά είναι [8]:

- Καλύτερα ορισμένη σημασιολογία. Εξαιτίας του υψηλού κόστους και πιθανών βλαβών (π.χ. από ακτινοβολία), οι ιατρικές εικόνες συνήθως παράγονται με ιδιαίτερη φροντίδα και σχεδιασμό και με καλά ορισμένους και συγκεκριμένους σκοπούς. Συνεπώς, σπάνια υπάρχει πολυσημία ως προς το περιεχόμενό τους, όπως συμβαίνει

συχνά με τις κοινές φωτογραφίες. Αν και μία ιατρική εικόνα μπορεί να χρησιμοποιηθεί για πολλούς σκοπούς, από διάγνωση ασθένειας μέχρι σχεδιασμό και αξιολόγηση θεραπείας και αναδρομικές μελέτες με γνωστά αποτελέσματα, το περιεχόμενο της εικόνας θα συνδέεται πάντα άμεσα, μοναδικά και ξεκάθαρα με το σημασιολογικό πλαίσιο χρήσης της εικόνας.

- Καθοδηγούμενος και αντικειμενικός συμπερασμός. Ο τομέας της ιατρικής στερείται της συναισθηματικής πτυχής που περιπλέκει συχνά άλλους CBIR τομείς. Στην ιατρική ευτυχώς η εξαγόμενη σημασιολογία, δηλαδή τα σημασιολογικά χαρακτηριστικά που δεν περιλαμβάνονται απαραίτητα στα περιεχόμενα της ιατρικής εικόνας, όπως για παράδειγμα «διαλειτουργικότητα όσον αφορά την εκτομή του ήπατος», συνήθως ορίζεται σε μεγάλο βαθμό από πρωτόκολλα, συστάσεις και τεκμηριωμένες ιατρικές μελέτες. Αβεβαιότητα και ασάφειες λύνονται με κρίση βασισμένη σε αποδείξεις, εμπειρία και το ποιόν του ασθενούς και όχι με βάση τις προσωπικές προκαταλήψεις του γιατρού.
- Προγενέστερη εργασία. Παρότι υπάρχουν πολλές ιατρικές απεικονιστικές τροπικότητες (CT, MRI, US, κ.ά.), όλες αναφέρονται στην ίδια οντότητα-στόχο: το ανθρώπινο σώμα, την ανατομία και λειτουργία του. Η ανθρώπινη ανατομία και λειτουργία μελετάται εμβριθώς για αιώνες και υπολογιστικά μοντέλα για ανθρώπινα όργανα όπως ο εγκέφαλος, η καρδιά ή ο πνεύμονας είναι διαθέσιμα προς χρήση σήμερα για την καθοδήγηση της σημασιολογικής ανάλυσης και τμηματοποίησης των εικόνων. Κλειδί ωστόσο για την σημασιολογική λειτουργία ενός CBIR είναι όχι μόνο η γνώση του πεδίου της ανατομίας και λειτουργίας του ανθρώπινου σώματος, αλλά και η σύνδεσή του με τις νόσους και τη συμπεριφορά τους, καθώς τα πιο κρίσιμα ερωτήματα στην ιατρική αφορούν όλα ασθένειες – πώς να τις προλαμβάνεις, ανιχνεύεις, να κάνεις διαγνώσεις και να τις θεραπεύεις. Ευτυχώς, πολλές ασθένειες και τα εξελικτικά τους βήματα είναι, αν και σύνθετα, γνωστά και μπορούμε να κατασκευάσουμε υπολογιστικά μοντέλα για αυτά.
- Πλούσιο περιεχόμενο. Υπάρχει πλούσιο πλαίσιο που περιβάλλει κάθε ιατρική εικόνα. Οι περισσότερες ιατρικές εικόνες σήμερα κωδικοποιούνται με βάση το πρότυπο DICOM (Digital Imaging and Communication in Medicine), το οποίο αποτελεί το βασικό πρότυπο στην ψηφιακή ιατρική εικόνα τα τελευταία 25 χρόνια. Κάθε αρχείο DICOM περιλαμβάνει εκτός από τα εικονοστοιχεία της εικόνας πληροφορία σχετική με τον ασθενή (ηλικία, φύλο, βάρος, ύψος, κ.ά.) και τις απεικονιστικές παραμέτρους (ημέρα και ώρα εξέτασης, αριθμός εικονοστοιχείων, τροπικότητα κ.ά.). Επιπρόσθετα, μπορεί υπάρχουν κλινικά δεδομένα όπως εργαστηριακές μελέτες, ιστορικό ασθενούς, ακόμη και γονιδιακά δεδομένα.

- Υφιστάμενη αναπαράσταση γνώσης. Όπως προαναφέρθηκε, υπάρχει παράδοση στην κατηγοριοποίηση της ιατρικής γνώσης εδώ και δεκαετίες με την ύπαρξη θησαυρών, ταξινομιών και οντολογιών που αναπτύχθηκαν και αναπτύσσονται από τους τομείς της υγειονομικής περίθαλψης και της βιοϊατρικής πληροφορικής. Από την πληθώρα των πόρων επίσημης αναπαράστασης γνώσης αναφέρεται ενδεικτικά το Ενιαίο Ιατρικό Σύστημα Γλώσσας (Unified Medical Language System ή UMLS), που καλύπτει το ευρύ πεδίο των βιοϊατρικών επιστημών, το Θεμελιώδες Μοντέλο Ανατομίας (Foundational Model of Anatomy ή FMA) και τη Διεθνή Κατηγοριοποίηση Ασθενειών (International Classification of Diseases ή ICD), που εκδίδεται και συνεχώς ενημερώνεται από τον Διεθνή Οργανισμό Υγείας (World Health Organization ή WHO).

Οι ευκαιρίες και δυνατότητες που παρουσιάζουν τα συστήματα σημασιολογικής αναζήτησης ιατρικών εικόνων είναι μεγάλες. Ωστόσο, υπάρχουν και αρκετές ανοιχτές προκλήσεις, περιορισμοί και κενά που πρέπει να ξεπεραστούν, τα σημαντικότερα των οποίων αναλύονται παρακάτω [8].

- Σημασιολογικό κενό. Η διαφορά ανάμεσα στις χαμηλού επιπέδου περιγραφές χαρακτηριστικών, που παρέχονται από εργαλεία ανάλυσης εικόνας, και στις υψηλού επιπέδου περιγραφές περιεχομένου, που είναι αναγκαίες από τις εφαρμογές, αναφέρεται συχνά ως «Σημασιολογικό Κενό» (Semantic Gap). Για να διευκολυνθεί ο αυτόματος λογικός συμπερασμός για καλύτερη ανάκτηση πληροφορίας, επιπρόσθετη σημασιολογία πρέπει να συνδεθεί με τα δεδομένα. Το πρόβλημα είναι ότι ο μεγάλος όγκος απεικονιστικών δεδομένων καθιστά απαγορευτική την επιλογή της χειροκίνητης επισημείωσης. Για το λόγο αυτό, εργαλεία αυτόματης ή έστω ημι-αυτόματης επισημείωσης με σημασιολογικά κριτήρια πρέπει να αναπτυχθούν παράλληλα με τη βελτιστοποίηση των υφιστάμενων εργαλείων αυτόματης ανάλυσης και τμηματοποίησης.
- Διεπιστημονικό κενό. Η ιατρική εικόνα είναι συχνά τελείως ακατανόητη σε μη ειδικούς. Είναι μια πρόκληση λοιπόν για τους επιστήμονες πληροφορικής να κατανοήσουν το ιατρικό πεδίο και τη σημασιολογία του, όπως όμως και για τους επαγγελματίες υγείας να γνωρίσουν τα θετικά και αρνητικά των χρησιμοποιούμενων δομών αναπαράστασης γνώσης, τη δυναμική και τους περιορισμούς τους.
- Κενό επίδοσης. Οι ιατρικές εφαρμογές είναι πολύ απαιτητικές στον τομέα της επίδοσης καθώς αφήνουν λίγο χώρο για λάθη ή παραλείψεις. Καθώς ένα σύστημα ανάκτησης εικόνων θα γίνεται πιο σημασιολογικό θα διαδραματίζει πιο καθοριστικό ρόλο στην κλινική απόφαση. Η ευθύνη των συστημάτων θα αυξηθεί και συνεπακόλουθα θα αυξηθούν οι ανάγκες για επίδοση του συστήματος.

- Κενό χρηστικότητα. Οι εξελίξεις στην τεχνολογία έχουν ως αποτέλεσμα την βελτίωση της καθημερινής ροής εργασιών των επαγγελματιών υγείας και την εξοικονόμηση χρόνου. Ένα καινούργιο σύστημα, εάν δε συμβάλλει στην εμφανή βελτίωση των όρων εργασίας, θα δυσκολευτεί να βρει εμπορική χρήση.
- Κενό διαθέσιμης σχετικής πληροφορίας. Οι γιατροί έχουν ανάγκη χρήσης όλων των διαθέσιμων δεδομένων για τη λήψη εμπειριστατωμένων αποφάσεων. Το CBIR σύστημα πρέπει να ενσωματώσει όλη αυτήν την πληροφορία για την υποστήριξη της λήψης αποφάσεων σε σημασιολογικό επίπεδο. Το πρόβλημα είναι ότι πολλά από τα δεδομένα μπορεί να μην είναι σε μορφή δομημένη και κατάλληλη για χρήση από προγράμματα και πράκτορες ή να μην είναι καν σε ηλεκτρονική μορφή.
- Κενό δεδομένων. Μολονότι παράγονται πολλές ιατρικές απεικονίσεις, αυτές δεν εύκολα διαθέσιμες λόγω κωλυμάτων ασφάλειας, προστασίας της ιδιωτικότητας και νομικών περιορισμών που περιβάλλουν τα δεδομένα υγείας γενικά. Πρέπει να λαμβάνονται μέτρα ανωνυμίας και σε πολλές περιπτώσεις απαιτείται η συγκατάθεση του ασθενούς για τη χρήση των εικόνων, κάτι που περιορίζει τον εφοδιασμό των συστημάτων αναζήτησης και ανακάλυψης γνώσης με ιατρικές απεικονίσεις.

Ιδανικά, όταν ξεπεραστούν τα παραπάνω εμπόδια, τα σημασιολογικά ερωτήματα που θα υποστηρίζονται από τα συστήματα ανάκτησης γνώσης από ιατρικές εικόνες δεν θα περιορίζονται μόνο στο περιεχόμενο των εικόνων, αλλά θα ενσωματώνουν και εξωτερικές πηγές γνώσης μέσα σε ένα κοινό πλαίσιο αναπαράστασης. Για παράδειγμα, ένα τέτοιο σύστημα, σε περίπτωση διάγνωσης καρκίνου του εντέρου σε κάποιον ασθενή, θα επιτρέπει την αναζήτηση προηγούμενων περιπτώσεων ασθενών με καρκίνο του εντέρου με παρόμοιο μοντέλο εξάπλωσης, ώστε να καθοδηγηθεί η θεραπευτική απόφαση για τον τωρινό ασθενή.

1.3 Σκοπός Διπλωματικής Εργασίας

Στις προηγούμενες ενότητες παρουσιάστηκαν τα σημαντικότερα προβλήματα αλλά και οι ευκαιρίες που παρουσιάζουν η επισημείωση και η αναζήτηση ιατρικών εικόνων. Σκοπός της παρούσας εργασίας είναι να αναδειχτούν λύσεις στα παραπάνω προβλήματα μέσω των αναπτυσσόμενων σημασιολογικών τεχνολογιών. Στα πλαίσια της εργασίας σχεδιάστηκε και υλοποιήθηκε ολοκληρωμένο διαδικτυακό σύστημα για τη σημασιολογική περιγραφή ιατρικών εικόνων.

Το σύστημα περιλαμβάνει δύο βασικές λειτουργίες: τη λειτουργία καταχώρησης σημασιολογικών κριτηρίων για ιατρικές εικόνες και τη λειτουργία αναζήτησης απεικονιστικών δεδομένων με βάση σημασιολογικά κριτήρια. Η περιγραφή των εικόνων πραγματοποιείται μέσω κατάλληλων οντολογιών, το οποίο αποσκοπεί στη μοντελοποίηση

των ιατρικών δεδομένων και της ιατρικής γνώσης. Για το λόγο αυτό χρησιμοποιήθηκαν οντολογίες και ταξινομήσεις που έχουν καθοριστεί σε διεθνές επίπεδο για το χώρο της ιατρικής, όπως η οντολογία Foundational Model of Anatomy για την ανατομία και η ταξινόμηση ICD-10 για την κατηγοριοποίηση των ασθενειών.

Σκοπός του συστήματος που αναπτύχθηκε είναι η ανάδειξη των πλεονεκτημάτων της σημασιολογικής προσέγγισης στον τομέα της διαχείρισης απεικονιστικών δεδομένων. Η διαχείριση των δεδομένων που βασίζεται σε σημασιολογικά κριτήρια θα επιτρέψει τη μελλοντική εφαρμογή αλγορίθμων τεχνητής νοημοσύνης για την εξόρυξη της γνώσης από τα απεικονιστικά δεδομένα και τη διασύνδεση αυτών με άλλες πηγές γνώσης. Επίσης, θα επιτρέψει την αποτελεσματική πρόσβαση σε απεικονιστικά δεδομένα προκειμένου να υποστηριχθεί η εκπαίδευση ιατρικού προσωπικού, η εξατομικευμένη διάγνωση και θεραπεία καθώς και οι επιδημιολογικές μελέτες. Η κοινή αναπαράσταση της γνώσης θα διευκολύνει, τέλος, τη διαλειτουργικότητα μεταξύ των ινστιτούτων και των ερευνητικών εργαστηρίων.

Συγκεντρωτικά, η εργασία περιλαμβάνει τρία βασικά στάδια:

1. Τη μοντελοποίηση της ιατρικής γνώσης με χρήση οντολογιών που έχουν καθοριστεί στη βιβλιογραφία για την αναπαράσταση της ιατρικής γνώσης.
2. Την ανάπτυξη ολοκληρωμένου διαδικτυακού συστήματος για τη σημασιολογική περιγραφή των απεικονιστικών δεδομένων και την επακόλουθη ανάκτησή τους, με βάση σημασιολογικά κριτήρια.
3. Τη σύνδεση των σημασιολογικών δεδομένων με πρότυπα που χρησιμοποιούνται διεθνώς όπως DICOM, RDF και OWL, για να είναι εφικτή η διασύνδεση των απεικονιστικών δεδομένων μεταξύ διαφορετικών πληροφοριακών συστημάτων αλλά και του διαδικτύου.

2

Θεωρητικό Υπόβαθρο

Στο κεφάλαιο αυτό παρουσιάζονται κάποιες βασικές έννοιες, η κατανόηση των οποίων κρίθηκε απαραίτητη για τη μελέτη των κεφαλαίων που ακολουθούν. Αναλύονται οι βασικές έννοιες της οντολογίας και της συλλογιστικής και γίνεται παρουσίαση των θεμελιωδών προτύπων και γλωσσών του Σημασιολογικού Ιστού: RDF, SPARQL, RDF Schema, OWL.

2.1 Οντολογίες

Ο όρος *οντολογία* (ontology) προέρχεται από την αρχαία ελληνική φιλοσοφία, με την κυριολεκτική μετάφραση της λέξης να είναι «μελέτη της φύσης της ύπαρξης». Η ιδέα της σύλληψης της γνώσης με δομημένο τρόπο αποδίδεται στον Αριστοτέλη, ο οποίος ήταν ο πρώτος που εστίασε με συστηματικό τρόπο στο πρακτικό πρόβλημα της αναπαράστασης της δομής της πραγματικότητας. Αν και η φιλοσοφία έχει επιστρατεύσει από τότε ένα σημαντικό σώμα αναλυτικών εργαλείων για την εξέταση των οντολογικών προβλημάτων, πολλές ιδέες και όροι της οντολογικής θεωρίας, όπως οι έννοιες της *κλάσης* και της *ιεραρχίας*, μπορούν δίκαια να αποδοθούν στον Αριστοτέλη. Η φιλοσοφική οντολογία έχει πάρει πολλές μορφές και έχει προσεγγιστεί από διάφορες γωνίες, ωστόσο κεντρικός στόχος παραμένει η ρητή και εξαντλητική ταξινόμηση όλων των οντοτήτων.

Εκτός από τον τομέα της φιλοσοφίας, τα τελευταία χρόνια ο όρος *οντολογία* βρέθηκε στο προσκήνιο και στο χώρο της επιστήμης των υπολογιστών, με τον τομέα της τεχνητής νοημοσύνης να δείχνει ιδιαίτερο ενδιαφέρον. Εδώ, ο όρος *οντολογία* προσεγγίζεται

διαφορετικά και αναφέρεται ως μια ρητή και τυπική προδιαγραφή μιας επινοίας (conceptualization), έτσι δεν μιλάμε πλέον για «οντολογία» αλλά για «μία οντολογία».

Με άλλα λόγια, η οντολογία αποτελεί το μοντέλο ενός συγκεκριμένου πεδίου που δημιουργείται για ένα συγκεκριμένο σκοπό. Επομένως, δεν υπάρχει σωστή οντολογία για ένα ορισμένο πεδίο, αλλά μια οντολογία πάντα αποτελεί γενίκευση ενός συγκεκριμένου πεδίου και πάντα υπάρχουν εφικτές εναλλακτικές επιλογές. Τα στοιχεία που θα συμπεριληφθούν σε αυτή τη γενίκευση πρέπει να καθορίζονται από τη χρήση για την οποία προορίζεται η οντολογία.

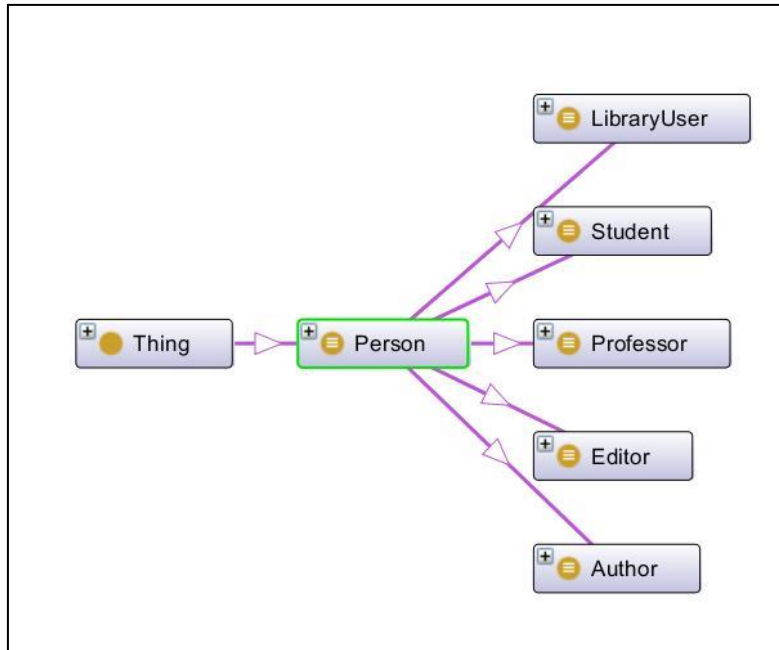
Εντούτοις, τελευταία παρατηρείται μία τάση ανάπτυξης οντολογιών γενικού περιεχόμενου και χωρίς καθορισμένη χρήση. Έτσι, λοιπόν, είναι αποδεκτός ο διαχωρισμός σε *οντολογίες αναφοράς* και *οντολογίες εφαρμογής*. Ως *οντολογία αναφοράς* [9] μπορούμε να ορίσουμε μια οντολογία όταν αποτελεί εκτεταμένη γενίκευση και δεν έχει σχεδιαστεί για κάποια συγκεκριμένη εφαρμογή, αλλά προορίζεται για επαναχρησιμοποίηση σε διάφορες εφαρμογές. Οι οντολογίες αναφοράς έχουν προταθεί ως ένας μηχανισμός για την παροχή του αναγκαίου οντολογικού πλαισίου για τη διασύνδεση οντολογιών και περιγράφουν μέρος ενός γενικευμένου πεδίου ενδιαφέροντος. Σχεδιάζονται ειδικά για να καλύπτουν ένα ευρύ φάσμα σχετικών πληροφοριών και συχνά τα μέλη της ομάδας σχεδίασης είναι ειδικοί σε υποπεδία του εκάστοτε γενικού πεδίου ενδιαφέροντος. Ως αποτέλεσμα, συχνά περιέχουν περισσότερες πληροφορίες από ότι χρειάζονται οι ειδικοί που ασχολούνται με ένα συγκεκριμένο υποτομέα. Αντίθετα όταν μια οντολογία σχεδιάζεται στοχευμένα για συγκεκριμένη εφαρμογή χαρακτηρίζεται ως *οντολογία εφαρμογής*. Μια ή περισσότερες οντολογίες αναφοράς μπορούν να χρησιμοποιηθούν για το σχεδιασμό και τη δημιουργία μιας οντολογίας εφαρμογής [10]. Για το σκοπό αυτό τελευταία αναπτύσσονται εφαρμογές διαχείρισης και τμηματοποίησης των μεγάλων οντολογιών αναφοράς, καθώς η απλή ανάκτηση «μιας φέτας» από μία οντολογία αναφοράς δεν είναι αποτελεσματική μέθοδος.

Γενικά, στην τεχνητή νοημοσύνη μια οντολογία περιγράφει ένα συγκεκριμένο πεδίο ενδιαφέροντος και απαρτίζεται από διάφορες οντότητες. Οι οντότητες αυτές μπορεί να είναι κλάσεις αντικειμένων (σημαντικές έννοιες του πεδίου), τα αντικείμενα τα ίδια και σχέσεις. Οι σχέσεις καθορίζουν συνήθως ιεραρχίες κλάσεων. Μια ιεραρχία ορίζει ότι μια κλάση A' είναι υποκλάση μιας άλλης κλάσης A , αν και μόνο αν κάθε άτομο της A' είναι ταυτόχρονα και άτομο της κλάσης A . Οι σχέσεις σε πρακτικό επίπεδο εφαρμόζονται στα αντικείμενα, καθότι όμως τα αντικείμενα συμπύσσουν κλάσεις θεωρούμε σχέσεις μεταξύ κλάσεων. Εκτός από σχέσεις κλάσης-υποκλάσης (ιεραρχία), οι σχέσεις μπορεί να καθορίζουν επίσης

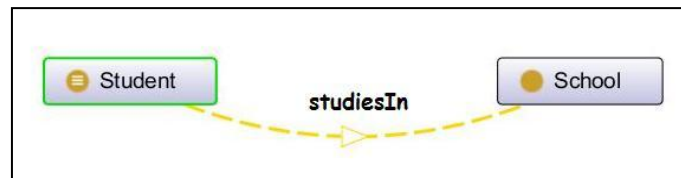
1. ιδιότητες
2. περιορισμούς τιμών
3. προτάσεις μη επικάλυψης

4. προδιαγραφές λογικών σχέσεων μεταξύ αντικειμένων

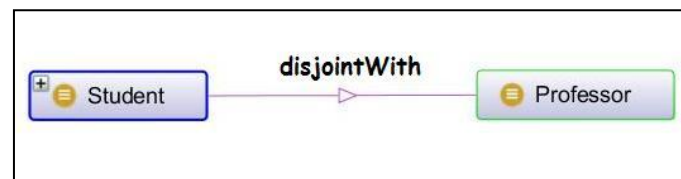
Στη συνέχεια με βάση μία οντολογία βιβλιοθήκης πανεπιστημίου (University Library Ontology) παρατίθενται γραφικά παραδείγματα.



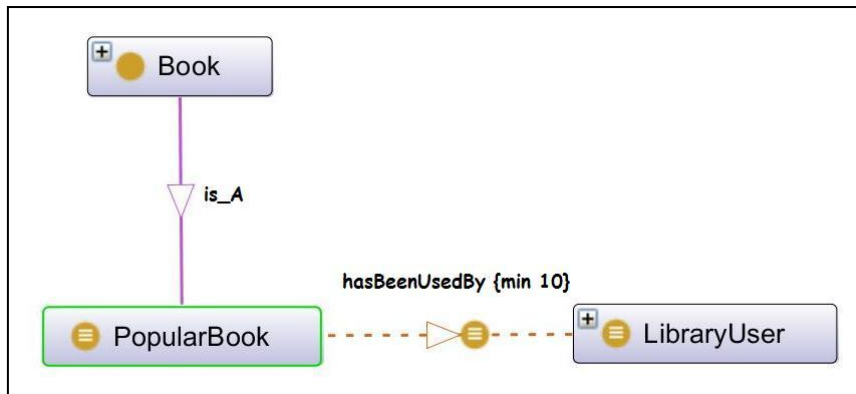
Εικόνα 2.1.1: Παράδειγμα Ιεραρχίας.



Εικόνα 2.1.2: Παράδειγμα Ιδιότητας: «some Student studiesIn some School»



Εικόνα 2.1.3: Παράδειγμα Πρότασης μη επικάλυψης. Όποιο αντικείμενο ανήκει στην κλάση Student δεν μπορεί να ανήκει και στην κλάση Professor, και αντίστροφα.



Εικόνα 2.1.4: Παράδειγμα λογικών σχέσεων. Όποιο αντικείμενο έχει χρησιμοποιηθεί (*hasBeenUsedBy*) από τουλάχιστον 10 αντικείμενα της κλάσης *LibraryUser* και ανήκει επίσης στην κλάση *Book*, τότε ανήκει και στην κλάση *PopularBook*.

2.2 Πρότυπα και Γλώσσες Σημασιολογικού Ιστού

Η παρουσίαση πληροφοριών στον Παγκόσμιο Ιστό γίνεται μέσω των ιστοσελίδων του, οι οποίες δημιουργούνται με γλώσσες σήμανσης, κυρίως HTML και XML, με την XML να παρουσιάζει καλύτερη δόμηση της πληροφορίας. Και οι δύο αυτές έχουν προέλθει από την Πρότυπη Γενικευμένη Γλώσσα Σήμανσης (Standard Generalized Markup Language ή SGML), ένα διεθνές πρότυπο (ISO 8879). Τέτοια πρότυπα είναι σημαντικά, καθώς επιτρέπουν την αποτελεσματική επικοινωνία και υποστηρίζουν έτσι την τεχνολογική πρόοδο και τη συνεργασία μεταξύ των επιχειρήσεων. Στο χώρο του Παγκόσμιου Ιστού, τα πρότυπα καθορίζονται από τον οργανισμό W3C και αποκαλούνται συστάσεις (recommendations), σε αναγνώριση του γεγονότος ότι η επιβολή προτύπων δεν είναι δυνατή σε ένα κατακεκομμένο περιβάλλον χωρίς κεντρική αρχή. Για τις ανάγκες του Σημασιολογικού Ιστού κρίθηκε απαραίτητη η σύσταση νέων προτύπων, με την W3C να πρωτοστατεί. Η νέα κατεύθυνση του Ιστού περιγράφει πλέον σχέσεις μεταξύ αντικειμένων και ιδιότητες των αντικειμένων σε μοντέλα που είναι ταυτόχρονα κατανοητά από τους ανθρώπους και επεξεργάσιμα από τους υπολογιστές.

2.2.1 RDF

Η XML είναι μια καθολική μετά-γλώσσα για τον ορισμό σήμανσης. Περιέχει ένα ενιαίο πλαίσιο, καθώς και ένα σύνολο από εργαλεία, όπως συντακτικοί αναλυτές, για την ανταλλαγή δεδομένων και μεταδεδομένων μεταξύ εφαρμογών. Παρόλα αυτά, η XML δεν παρέχει κάποιον τρόπο καθορισμού της σημασιολογίας (σημασίας) των δεδομένων.

Παρόλο που το Πλαίσιο Περιγραφής Πόρων (Resource Description Framework ή RDF) αποκαλείται καταχρηστικά «γλώσσα» στην ουσία είναι ένα μοντέλο δεδομένων [1]. Φυσικά, ένα αφηρημένο μοντέλο δεδομένων χρειάζεται μια καθορισμένη σύνταξη ώστε να αναπαρίσταται και να μεταδίδεται. Η πιο διαδεδομένη μορφή σειριακής διάταξης της RDF είναι η RDF/XML, η οποία περιλαμβάνεται στις προδιαγραφές της RDF που συντάχθηκαν από την W3C [11]. Η σύνταξη της RDF/XML είναι βασισμένη στην XML και οι αναλυτικές προδιαγραφές της παρουσιάζονται στην αντίστοιχη σύσταση της W3C [12]. Άλλες μορφές είναι η Notation 3 (N3), η Turtle και η N-Triples. Η RDF επιτρέπει την αναπαράσταση πληροφορίας σχετικής με πόρους στον Παγκόσμιο Ιστό. Το βασικό δομικό στοιχείο της είναι μία τριάδα (triple) υποκειμένου - κατηγορήματος - αντικείμενου (subject - predicate - object) και καλείται *πρόταση*. Υποκείμενο είναι ένας *πόρος*, κατηγορημα είναι μία *ιδιότητα* και αντικείμενο είναι είτε ένας *πόρος* είτε ένα *λεκτικό* (*literal*).

Πόρος θεωρείται κάθε αντικείμενο για το οποίο θέλουμε να μιλήσουμε και έχει έναν Ενιαίο Αναγνωριστικό Πόρου (Unique Resource Identifier ή URI). Σημειώνεται ότι η ύπαρξη του αναγνωριστικού συνεπάγεται τη δυνατότητα περιγραφής του πόρου, αλλά όχι απαραίτητα και τη δυνατότητα πρόσβασης σε αυτόν. Συνήθως ένα URI αντιστοιχίζεται σε μία Ενιαία Τοποθεσία Πόρου (Unique Resource Location ή URL), επιτρέποντας την πρόσβαση σε αυτόν. Οι συνδυασμοί URI δεν έχουν οριστεί όμως μόνο για τοποθεσίες Ιστού, αλλά και για αντικείμενα διαφόρων ειδών, όπως τηλεφωνικούς αριθμούς, αριθμούς ISBN και γεωγραφικές τοποθεσίες. Οι *ιδιότητες* είναι μία ειδική περίπτωση πόρων που περιγράφουν τις σχέσεις μεταξύ πόρων. Τέτοιες σχέσεις για παράδειγμα είναι οι: «τίτλος», «δημιουργός», «ημερομηνία γέννησης», κ.ά.. Οι ιδιότητες, ως πόροι και αυτές, καθορίζονται επίσης από URIs, κάτι το οποίο έχει ιδιαίτερη σημασία, καθώς προσφέρει μία μέθοδο παγκόσμιας, μοναδικής ονομασίας μειώνοντας δραστικά το πρόβλημα της ομωνυμίας που «μαστιρίζει» την αναπαράσταση των κατανεμημένων δεδομένων μέχρι σήμερα. Τα *λεκτικά*, τέλος, είναι αλφαριθμητικά (strings).

Μία πρόταση δηλώνει ότι ο πόρος-υποκείμενο συνδέεται με τη σχέση-κατηγορημα με κάποιον άλλο πόρο-αντικείμενο ή ότι ο πόρος-υποκείμενο συνδέεται με τη σχέση-κατηγορημα με κάποιο αλφαριθμητικό-αντικείμενο, το οποίο είναι η λεκτική αναπαράσταση της τιμής που έχει κάποια ιδιότητά του (το κατηγορημα). Για παράδειγμα η πρόταση «*H Alice γνωρίζει τον Bob*» μπορεί να μεταφραστεί στην παρακάτω RDF τριάδα :

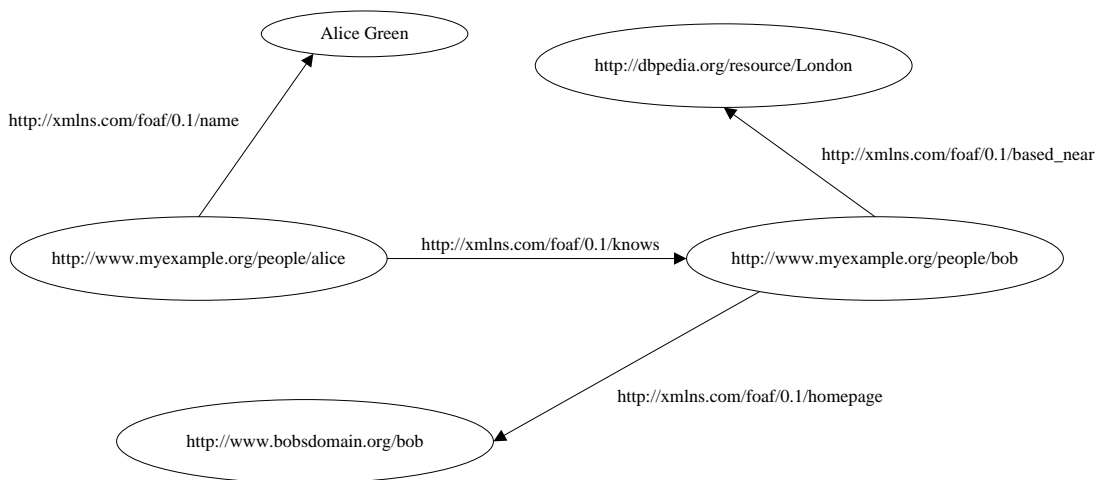
```
(http://www.myexample.org/people/alice,  
http://xmlns.com/foaf/0.1/knows, http://www.myexample.org/people/bob)
```

Μία πρόταση μπορεί να αναπαρασταθεί γραφικά από δύο κόμβους που συνδέονται με μία κατευθυνόμενη ακμή. Για παράδειγμα, η παραπάνω πρόταση αναπαρίσταται γραφικά στην Εικόνα 2.2.1.



Εικόνα 2.2.1 : Γραφική αναπαράσταση RDF τριάδας.

Αντίστοιχα, πολλές προτάσεις μπορούν να αναπαρασταθούν από έναν κατευθυνόμενο γράφο, που ονομάζεται σημασιολογικό δίκτυο, ένα παράδειγμα του οποίου φαίνεται στην Εικόνα 2.2.2.



Εικόνα 2.2.2: Ένα σημασιολογικό δίκτυο.

2.2.2 SPARQL

Η SPARQL θεωρείται η επίσημη γλώσσα ερωτημάτων για την RDF, μετά τη σχετική σύσταση της W3C [13]. Η SPARQL είναι για την RDF ότι ακριβώς είναι για τις σχεσιακές βάσεις δεδομένων η SQL ή για την XML η XQuery. Τα SPARQL ερωτήματα βασίζονται σε μοτίβα τριάδων (triple patterns). Ένα μοτίβο τριάδας είναι το ίδιο με μία RDF τριάδα, με τη μόνη διαφορά ότι ένας ή περισσότεροι από τους πόρους-συνιστώσες του είναι μεταβλητές. Μία μεταβλητή συμβολίζεται *?name*, όπου *name* είναι το όνομά της. Η μηχανή SPARQL που εκτελεί ένα ερώτημα αναζητά από όλους τους πόρους, εκείνους που επαληθεύουν τα μοτίβα τριάδων του ερωτήματος, σύμφωνα με τις RDF προτάσεις που υπάρχουν στη βάση γνώσης με την οποία η SPARQL μηχανή είναι συνδεδεμένη. Υπάρχουν τέσσερις μορφές SPARQL ερωτημάτων: SELECT, CONSTRUCT, ASK και DESCRIBE.

2.2.3 *RDF Schema*

Το Σχήμα Περιγραφής Πόρων (RDF Schema ή αλλιώς RDFS) είναι μία γλώσσα για την περιγραφή RDF λεξιλογίων [14]. Ενώ η RDF είναι ανεξάρτητη πεδίου, με την έννοια ότι δεν έχει γίνει καμία υπόθεση σχετικά με κάποιο συγκεκριμένο πεδίο χρήσης της, η RDFS δίνει τη δυνατότητα καθορισμού της σημασιολογίας κάποιου πεδίου, μέσω του ορισμού της συγκεκριμένης ορολογίας πεδίου. Συγκεκριμένα, η RDFS προσφέρει μηχανισμούς για την περιγραφή ομάδων σχετικών πόρων, που ονομάζονται κλάσεις, και για την περιγραφή των σχέσεων μεταξύ των πόρων. Η ταξινόμηση των μεμονωμένων αντικειμένων σε κλάσεις που ορίζουν τύπους αντικειμένων μας επιτρέπει να ορίσουμε σχέσεις στις κλάσεις, να περιορίσουμε τις ιδιότητες των αντικειμένων συγκεκριμένων κλάσεων, να ορίσουμε ιεραρχικές σχέσεις μεταξύ κλάσεων και γενικά να εισάγουμε δομή στα έγγραφα RDF. Για παράδειγμα, η RDFS καθορίζει τη σημασιολογία της φράσης «έχει πεδίο ορισμού» έτσι ώστε η ερμηνεία της να είναι μοναδική και να μην εξαρτάται από την εκάστοτε εφαρμογή. Με τον τρόπο αυτό, το γνωστό επιδιωκόμενο νόημα της φράσης πρέπει να χρησιμοποιηθεί από όλα τα προγράμματα επεξεργασίας RDF. Θα αντιπαραβάλλουμε τη σχέση μεταξύ RDF και RDFS με τη σχέση μεταξύ στιγμιότυπων και σχήματος μιας τυπικής σχεσιακής βάσης δεδομένων, αναλογία η οποία, παρότι υπεραπλουστευτική και ενδεχομένως ανακόλουθη, μας δίνει μια γενική εικόνα της διαφοράς μεταξύ των δύο γλωσσών περιγραφής πόρων.

Με βάση τα παραπάνω, η RDFS γίνεται μια περιορισμένη γλώσσα περιγραφής της σημασιολογίας συγκεκριμένων πεδίων ή, αλλιώς διατυπωμένο, μια στοιχειώδης γλώσσα οντολογιών. Η μελέτη χαρακτηριστικών περιπτώσεων χρήσης για το Σημασιολογικό Ιστό όμως γρήγορα οδήγησε στην ανάγκη για μεγαλύτερη εκφραστικότητα κατά την περιγραφή δεδομένων από αυτήν των γλωσσών RDF και RDFS [15].

2.2.4 *OWL*

Για να επιτευχθεί μεγαλύτερη εκφραστικότητα αναπτύχθηκε από την Ομάδα Εργασίας Οντολογιών του οργανισμού W3C η ευρέως αποδεκτή Γλώσσα Οντολογιών Ιστού (Web Ontology Language ή OWL) [16]. Η OWL επεκτείνει την εκφραστικότητα της RDFS επιτρέποντας τη δυνατότητα για

- ισότητα
- περιορισμό πληθικότητας κλάσεων (π.χ. «τουλάχιστον ένα»)
- τοπική εμβέλεια ιδιοτήτων (δηλ. πεδίο ορισμού)
- μη επικάλυψη κλάσεων
- λογικοί συνδυασμοί (ένωση, τομή κ.ά.)
- ειδικά χαρακτηριστικά ιδιοτήτων (π.χ. μεταβατική)

Επιπρόσθετα, ικανοποιεί όλες τις απαιτήσεις μιας οντολογικής γλώσσας: καλά ορισμένη σύνταξη, τυπική σημασιολογία, αποδοτική υποστήριξη συλλογισμών, επαρκής εκφραστική ισχύς και ευκολία στην έκφραση. Η υποστήριξη συλλογισμών προϋποθέτει την ύπαρξη τυπικής σημασιολογίας και οδηγεί στην αυτόματη παραγωγή γνώσης που αφορά τόσο στον έλεγχο της συνέπειας της οντολογίας και της γνώσης όσο και στην αυτόματη ταξινόμηση των στιγμιότυπων σε κλάσεις.

Αν και το ιδανικό προς τη συνέπεια και τη γενική αρχιτεκτονική του Σημασιολογικού Ιστού θα ήταν η OWL να αποτελεί επέκταση της RDF Schema, κάτι τέτοιο κρίθηκε ανέφικτο καθώς θα λειτουργούσε ενάντια στην εκφραστική ισχύ και στην αποδοτική συλλογιστική. Έτσι οι απαιτήσεις αυτές ώθησαν την Ομάδα Εργασίας Οντολογιών Ιστού της W3C να την ορίσουν ως τρεις διαφορετικές υπογλώσσες. Καθεμία από τις υπογλώσσες αυτές έχει προσαρμοστεί ώστε να ανταποκρίνεται σε συγκεκριμένες προδιαγραφές.

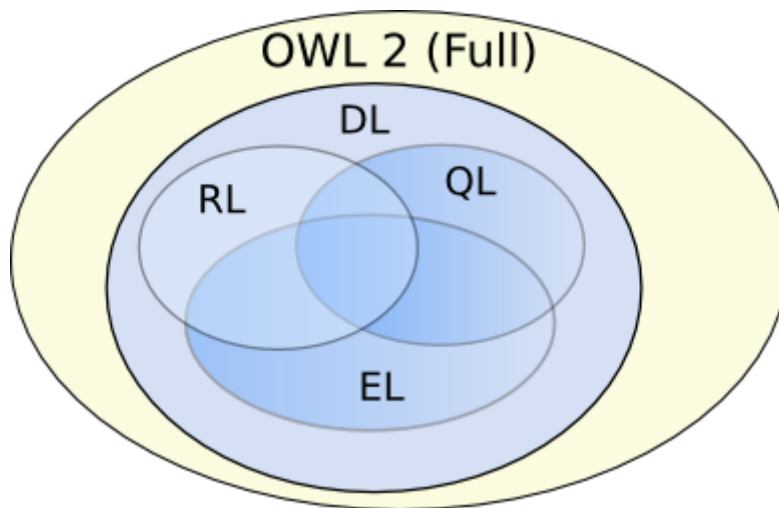
1. Η OWL Lite υποστηρίζει εκείνους τους χρήστες που χρειάζονται κυρίως μια ιεραρχία ταξινόμησης και απλούς περιορισμούς. Επίσης έχει χαμηλότερη πολυπλοκότητα από την OWL DL.
2. Η OWL DL υποστηρίζει εκείνους τους χρήστες που θέλουν τη μέγιστη εκφραστικότητα διατηρώντας την υπολογιστική πληρότητα (όλα τα συμπεράσματα είναι εγγυημένα ότι είναι επεξεργάσιμα από υπολογιστές) και αποκρισιμότητα (όλοι οι υπολογισμοί θα τελειώσουν σε πεπερασμένο χρόνο). Περιλαμβάνει όλες τις γλωσσικές δομές της OWL, αλλά μπορούν να χρησιμοποιηθούν μόνο υπό ορισμένους περιορισμούς. Η OWL DL ονομάζεται έτσι λόγω αντιστοιχία της με τις Περιγραφικές Λογικές (Description Logics) [17].
3. Η OWL Full προορίζεται για χρήστες που θέλουν μέγιστη εκφραστικότητα και την ελευθερία του RDF συντακτικού, χωρίς υπολογιστικές εγγυήσεις. Επιτρέπει σε μια οντολογία να αυξήσει την έννοια του προκαθορισμένου (RDF ή OWL) λεξιλογίου. Είναι απίθανο, όμως, οποιοδήποτε λογισμικό συλλογιστικής (reasoning) να είναι σε θέση να υποστηρίξει κάθε χαρακτηριστικό της OWL Full.

Πρόσφατα, τον Οκτώβριο του 2009, από την Ομάδα Εργασίας για την OWL της W3C παρουσιάστηκε η OWL2, η οποία προσθέτει νέες λειτουργικότητες [18]. Κάποιες είναι απλές και προσδίδουν ευχρηστία, ενώ κάποιες άλλες επεκτείνουν την εκφραστικότητα και παρουσιάζονται επιγραμματικά:

- κλειδιά
- αλυσίδες ιδιοτήτων
- πλουσιότερους τύπους δεδομένων, πεδίο ορισμού δεδομένων
- ειδικοί περιορισμοί πληθικότητας

- μη-συμμετρική, ανακλαστική και ιδιότητα μη-επικάλυψης
- επαυξημένες δυνατότητες σχολιασμού

Η OWL2 παρουσιάζεται με κάποια προφίλ, τα οποία είναι υπογλώσσες (συντακτικά υποσύνολα της OWL2) και προσφέρουν σημαντικά πλεονεκτήματα σε συγκεκριμένα σενάρια εφαρμογής. Έχουν οριστεί τρία προφίλ: OWL2 EL, OWL2 QL και OWL2 RL. Κάθε προφίλ ορίζεται ως συντακτικός περιορισμός των διαρθρωτικών προδιαγραφών της OWL2, δηλαδή, ως υποσύνολο των δομικών στοιχείων που μπορούν να χρησιμοποιηθούν σε μια οντολογία, και το καθένα είναι πιο περιοριστικό από OWL DL. Κάθε ένα από τα προφίλ αποσκοπεί σε διάφορες πτυχές της εκφραστικής δύναμης της OWL με αντάλλαγμα διαφορετικά υπολογιστικά ή / και εφαρμόσιμα οφέλη.



Εικόνα 2.2.3: Διάγραμμα Venn των συντακτικών υποσυνόλων (προφίλ) της OWL2 [19].

2.3 Λογική και Συλλογιστική

Η *λογική* (logic) είναι το επιστημονικό πεδίο που μελετά τις αρχές της *συλλογιστικής* (reasoning) και οι ρίζες της φτάνουν μέχρι τον Αριστοτέλη. Η λογική αποτελεί τη βάση της αναπαράστασης γνώσης, ειδικά με τη μορφή της κατηγορηματικής λογικής που είναι γνωστή και ως *λογική πρώτης τάξης* (first-order-logic). Η *περιγραφική λογική* (description logic), τώρα, αποτελεί ένα υποσύνολο της κατηγορηματικής λογικής, για το οποίο είναι δυνατόν να υπάρξει αποδοτική υποστήριξη συλλογισμών [20]. Σημειώνεται βέβαια ότι πρέπει να υπάρχει πάντα ένας συμβιβασμός μεταξύ εκφραστικότητας της γλώσσας και αποδοτικής υποστήριξης συλλογισμών, καθώς όσο πλουσιότερη είναι η γλώσσα, τόσο λιγότερο αποδοτική γίνεται η υποστήριξη συλλογισμών, ξεπερνώντας συχνά το όριο της μη υπολογισιμότητας.

Από τη στιγμή που αναπτύχθηκαν γλώσσες αναπαράστασης γνώσης, που μπορούν να αντιστοιχηθούν σε μια περιγραφική λογική, άνοιξε ο δρόμος για την ανάπτυξη προγραμμάτων αυτοματοποιημένης συλλογιστικής (*semantic reasoners* ή *reasoners*), προγράμματα που είναι σε θέση να συναγάγουν λογικές συνέπειες από ένα σύνολο ισχυρισμών ή αξιωμάτων [21]. Η έννοια του reasoner γενικεύει εκείνη της μηχανής συμπερασμάτων, παρέχοντας μια πιο πλούσια σειρά μηχανισμών για να εργαστεί. Οι κανόνες συμπερασμάτων συνήθως καθορίζονται μέσω μιας γλώσσας οντολογίας, και συχνά μιας γλώσσας περιγραφής. Πολλοί reasoners κάνουν χρήση λογικής πρώτης τάξης για να εκτελέσουν συλλογιστική και να εξάγουν ασφαλή συμπεράσματα, ενώ άλλοι κάνουν χρήση αλγορίθμων Tableaux. Τα συμπεράσμα αυτά συνήθως εξάγονται από εμπροσθοδρομική ή οπισθοδρομική αλύσωση (forward ή backward chaining). Από τους πιο διαδεδομένους reasoners είναι οι Pellet, Fact++, RacerPro, Hermit και SwiftOWLIM.

3

Σημασιολογικός Ιστός και Ιατρική

Οι μονάδες υγείας δημιουργούν σε καθημερινή βάση τεράστιες ποσότητες δεδομένων διαφορετικού είδους (διοικητικά, οικονομικά, ιατρικά δεδομένα κ.ά.) και διαφορετικής φύσης (ιατρικοί φάκελοι, ιατρικές εικόνες, βιοσήματα κ.ά.). Συνεπώς η καλύτερη οργάνωση των δεδομένων με σκοπό την αποτελεσματικότερη ανάκτηση πληροφορίας αποτελεί μονόδρομο. Για το λόγο αυτό ο ευρύτερος χώρος της Ιατρικής έχει στραφεί εδώ και καιρό στις τεχνολογίες του Σημασιολογικού Ιστού, συνεισφέροντας τα δέοντα στην ανάπτυξη του. Αλλά και η κοινότητα του Σημασιολογικού Ιστού έχει στραφεί προς το χώρο της Ιατρικής, καθώς βρίσκει έφορο έδαφος για την πρακτική εφαρμογή των αναπτυσσόμενων τεχνολογιών.

3.1 Πρότυπα και Κωδικοποιήσεις στην Ιατρική

Όταν τα ιατρικά δεδομένα άρχισαν να αποθηκεύονται σε ηλεκτρονική μορφή γεννήθηκε η ανάγκη για πρότυπα, κωδικοποιήσεις και πρωτόκολλα, ώστε όχι μόνο να αποθηκεύονται αλλά να τηρούνται και ανταλλάσσονται με σωστό και δομημένο τρόπο. Έτσι επιτρέπεται η συστηματική, τυποποιημένη και αξιοποιήσιμη καταγραφή πληροφοριών με σκοπό την ανάπτυξη και τήρηση ιατρικού ιστορικού ασθενών, την υποβοήθηση της διάγνωσης και την αναβάθμιση της ιατροφαρμακευτικής περίθαλψης. Επίσης, επιτρέπεται η ανάλυση, σύγκριση δεδομένων και η τήρηση βασικών στατιστικών στοιχείων για εθνικές και διεθνείς συγκρίσεις. Η χρήση προτύπων και κωδικοποιήσεων στα πληροφοριακά συστήματα διευκολύνει τη διαλειτουργικότητα μεταξύ των μονάδων υγείας από εθνικό ως και διεθνές επίπεδο.

Υπάρχουν πολλών ειδών κωδικοποιήσεις, κυριαρχούν όμως οι ταξινομήσεις (classifications) και οι ονοματολογίες (nomenclatures). Άλλες έννοιες κωδικοποίησης είναι οι θησαυροί λέξεων (thesaurus), οι ταξινομίες (taxonomies) και οι τυπικές ορολογίες (formal terminologies).

Με βάση τα παραπάνω και τις διεθνείς πρακτικές έχουν δημιουργηθεί πολλά πρότυπα και κωδικοποιήσεις για την παρουσίαση των κλινικών δεδομένων, ώστε να εκφράσουν με συστηματικό τρόπο διαγνώσεις και διαδικασίες. Ακολουθεί περιγραφή των σημαντικότερων από αυτά.

3.1.1 DICOM

Η Ψηφιακή Απεικόνιση και Επικοινωνίες στην Ιατρική (Digital Imaging and Communications in Medicine ή DICOM) είναι ένα πρότυπο που δημιουργήθηκε από την Εθνική Ένωση Ηλεκτρολόγων Κατασκευαστών (National Electrical Manufacturers Association ή NEMA) για να ενισχύσει τη διανομή και την προβολή των ιατρικών εικόνων, όπως η σάρωση αξονικής τομογραφίας (CT scan), η μαγνητική τομογραφία συντονισμού (MRI), και το υπερηχογράφημα (US). Το Μέρος 10 του προτύπου περιγράφει μια μορφή αρχείου για τη διανομή των εικόνων [22]. Αυτή η μορφή είναι μια επέκταση του παλαιού προτύπου NEMA. Οι περισσότεροι άνθρωποι αναφέρονται σε αρχεία εικόνων οι οποίες είναι σύμφωνες με το μέρος 10 του προτύπου DICOM όπως αρχεία μορφής DICOM. Ένα ενιαίο αρχείο DICOM περιλαμβάνει τόσο μια επικεφαλίδα (η οποία αποθηκεύει πληροφορίες σχετικά με το όνομα του ασθενούς, το είδος της σάρωσης, οι διαστάσεις της εικόνας, κλπ), καθώς και όλα τα δεδομένα εικόνας (που μπορεί να είναι πληροφορίες σε τρεις διαστάσεις). Αυτό είναι διαφορετικό από τη δημοφιλή μορφή Analyze, η οποία αποθηκεύει τα δεδομένα εικόνας σε ένα αρχείο (*.img) και τα στοιχεία της κεφαλίδας σε άλλο αρχείο (*.HDR). Μια άλλη διαφορά μεταξύ των DICOM και Analyze είναι ότι τα δεδομένα εικόνας DICOM μπορούν να συμπιεστούν (ενθυλακωμένα) ώστε να μειωθεί το μέγεθος της εικόνας. Τα αρχεία μπορούν να συμπιεστούν χρησιμοποιώντας παραλλαγές της μορφής JPEG (με ή χωρίς απώλειες), καθώς και χρησιμοποιώντας μια μορφή κωδικοποίησης χωρίς απώλειες ονόματι Run-Length Classification (η οποία είναι ταυτόσημη με την συμπίεση packed-bits που παρατηρείται σε ορισμένες εικόνες μορφής TIFF).

3.1.2 HL7

Το Επίπεδο Υγείας Επτά (Health Level Seven ή HL7) είναι το πιο καταξιωμένο διεθνές πρότυπο διασύνδεσης εφαρμογών στον χώρο της Υγείας, το οποίο δίνει στους φορείς (νοσοκομεία, κέντρα υγείας, ασφαλιστικοί οργανισμοί κλπ) τη δυνατότητα να θέτουν συγκεκριμένες προδιαγραφές και να προσδιορίζουν επακριβώς το τρόπο διασύνδεσης τόσο

των υφιστάμενων όσο και των νέων πληροφοριακών συστημάτων. Αυτό βοηθά στη σωστή λειτουργία κάτω από ενιαίο πλαίσιο, ένα ενιαίο πληροφοριακό σύστημα [23]. Επιπρόσθετα παρέχει σε όλους του Φορείς Υγείας, την δυνατότητα να τυποποιήσουν τις καθημερινές τους λειτουργίες και διαδικασίες, να εξασφαλίσουν τον οργανισμό έναντι των προμηθευτών - στη διαδικασία προμήθειας και εγκατάστασης πληροφοριακών συστημάτων - και να πετύχουν σημαντικά οικονομικά οφέλη μέσω της τυποποίησης και της διαλειτουργικότητας των συστημάτων.

Παρατήρηση: Στα πλαίσια του HL7 έχει προταθεί και πρότυπο για τον Ηλεκτρονικό Ιατρικό Φάκελο (Electronic Health Record ή EHR).

3.1.3 SNOMED-CT

Η Συστηματοποιημένη Ονοματολογία Ιατρικών και Κλινικών Όρων (Systematized Nomenclature of Medical-Clinical Terms ή SNOMED-CT) αποσκοπεί στη βελτίωση της φροντίδας του ασθενούς μέσω της ανάπτυξης συστημάτων για την ακριβή καταγραφή περίθαλψης [24]. Σε τελική ανάλυση, οι ασθενείς θα ωφεληθούν από τη χρήση του SNOMED-CT, με την οικοδόμηση και προώθηση της επικοινωνίας και της διαλειτουργικότητας στην ηλεκτρονική ανταλλαγή δεδομένων υγείας. Η SNOMED-CT δημιουργήθηκε αρχικά από το Κολλέγιο Αμερικανών Παθολόγων συνδυάζοντας τη SNOMED-RT και μια ονοματολογία-ταξινόμηση, γνωστή ως Κλινικοί Όροι Έκδοση 3 (Clinical Terms Version 3), παλαιότερα γνωστή ως Read Codes Version 3, η οποία δημιουργήθηκε για λογαριασμό του βρετανικού Υπουργείου Υγείας.

3.1.4 openEHR

Το openEHR είναι ένα ανοικτό πρότυπο προδιαγραφής στο χώρο της ιατρικής πληροφορικής που περιγράφει τη διαχείριση, αποθήκευση, ανάκληση και ανταλλαγή των δεδομένων υγείας των Ηλεκτρονικών Ιατρικών Φακέλων (EHRs) [25]. Στο openEHR όλα τα δεδομένα υγείας του ασθενούς αποθηκεύονται σε ένα μοναδικό και ενιαίο, ανεξαρτήτου προμηθευτών και προσωποκεντρικό Ηλεκτρονικό Φάκελο. Οι προδιαγραφές δεν αφορούν την ανταλλαγή των δεδομένων μεταξύ EHR συστημάτων, όπου εστιάζουν άλλα πρότυπα όπως το HL7, αλλά περιλαμβάνουν μοντέλα πληροφορίας και υπηρεσιών για τον Ηλεκτρονικό Φάκελο, δημογραφικά, κλινική ροή εργασίας και «αρχέτυπα». Αρχέτυπα ορίζονται τα υψηλής ποιότητας, επαναχρησιμοποιήσιμα κλινικά μοντέλα περιεχομένου και διαδικασίας. Το openEHR είναι μια πρωτοβουλία που στηρίχθηκε σε πολύχρονη έρευνα ευρωπαϊκών και αυστραλιανών πανεπιστημίων και εξελίχθηκε σε διεθνή μη-κερδοσκοπική οργάνωση που στόχο έχει την «ενσωμάτωση των προσαρμόσιμων υπολογιστικών συστημάτων υγείας και των εξατομικευμένων Ηλεκτρονικών Φακέλων Υγείας κάτω από ένα υπολογιστικό πλαίσιο

προσανατολισμένο στη γνώση, που περιλαμβάνει οντολογίες, ορολογίες και μία σημασιολογικά υποβοηθούμενη υπολογιστική πλατφόρμα υγείας, στην οποία σύνθετες έννοιες μπορούν να αναπαρασταθούν και να μοιράζονται» [26].

3.1.5 ICD

Η Διεθνής Κατηγοριοποίηση των Ασθενειών (International Classification of Diseases ή ICD) βρίσκεται ήδη στη 10η αναθεώρηση. Οι κωδικοί που την απαρτίζουν συντηρούνται από τον Παγκόσμιο Οργανισμό Υγείας (World Health Organization ή WHO) και είναι αποδεκτοί παγκοσμίως. Η ICD είναι διεθνές πρότυπο διαγνωστικής κατηγοριοποίησης με αρκετά πεδία εφαρμογής, όπως γενικά επιδημιολογικά θέματα, διαχείριση των φορέων υγείας και κλινική χρήση, ανάλυση της γενικής κατάστασης της υγείας των πληθυσμιακών ομάδων και η παρακολούθηση της εμφάνισης των ασθενειών. Η ICD χρησιμοποιείται, ως επί των πλείστον, για την ταξινόμηση των ασθενειών και άλλων προβλημάτων υγείας, που καταγράφονται σε πολλά είδη στοιχείων της υγείας, συμπεριλαμβανομένων των πιστοποιητικών θανάτου και των ιατρικών φακέλων. Εκτός από το να καταστεί δυνατή η αποθήκευση και ανάκτηση των διαγνωστικών πληροφοριών για κλινικούς και επιδημιολογικούς σκοπούς, τα αρχεία αυτά παρέχουν επίσης τη βάση για την κατάρτιση των εθνικών στατιστικών στοιχείων θνησιμότητας και νοσηρότητας των κρατών-μελών του WHO.

Παρατήρηση: Βρίσκεται σε εξέλιξη η δημιουργία της ICD-11, η οποία θα είναι εννοιολογικά εμπλουτισμένη, ώστε να συμπεριληφθούν σχέσεις σύνδεσης με ανατομία.

3.2 Οντολογίες στην Ιατρική

Στην ιατρική υπάρχει ένας μεγάλος αριθμός ορολογικών συλλογών που αναπτύσσονται για διαφορετικούς λόγους (ευρετηρίαση και ανάκτηση πληροφοριών, ηλεκτρονικά αρχεία ασθενών, στατιστικές εκθέσεις σχετικά με τη θνησιμότητα κ.ά.), σε διαφορετικά υποπεδία (ασθένειες, μικροοργανισμοί, ιατρικές συσκευές, διαδικασίες, φάρμακα) και από διαφορετικές οργανώσεις (Παγκόσμιος Οργανισμός Υγείας, κυβερνητικές αντιπροσωπείες, επαγγελματικές ενώσεις). Κάθε ορολογικό προϊόν αναπαριστάει τον κόσμο με τρόπο συμβατό και κατάλληλο για το σκοπό για τον οποίο έχει αναπτυχθεί. Οι προσπάθειες με παγκόσμιο και καθολικό χαρακτήρα για την αναπαράσταση της γνώσης στο πεδίο της ιατρικής είναι συνήθως προβληματικές, αφού, είτε είναι περιορισμένες σε εμβέλεια και καλύπτουν λίγους υποτομείς (π.χ. GALEN), είτε παρουσιάζουν προβλήματα οντολογικής οργάνωσης (π.χ. SNOMED). Έτσι, η επιθυμητή σημασιολογική διαλειτουργικότητα είναι δύσκολο να επιτευχθεί μέσω των υπάρχοντων ορολογικών πόρων.

Η αυξανόμενη διαθεσιμότητα μεγάλου όγκου ιατρικών κειμένων σε online μορφή (π.χ. άρθρα σε περιοδικά, εγκυκλοπαίδειες, αρχεία ασθενών) επιβάλλει την ανάπτυξη εφαρμογών βασισμένων στην επεξεργασία της γνώσης (knowledge processing). Παραδείγματα τέτοιων εφαρμογών περιλαμβάνουν την εννοιοκεντρική ευρετηρίαση και ανάκτηση, εφαρμογές ερωτήσεων-απαντήσεων και, γενικότερα, εφαρμογές κατανόησης κειμένου και εξαγωγής πληροφορίας από κείμενα ή βάσεις δεδομένων. Μια έγκυρη αναπαράσταση της γνώσης του ιατρικού πεδίου, κυρίως μέσω οντολογιών, απαιτείται από όλες τις προαναφερθείσες εφαρμογές.

Πολύ συχνά, αντί για τον όρο οντολογία, χρησιμοποιείται ο όρος *σημασιολογικό δίκτυο* (semantic network), που αποτελεί μια τεχνική για την αναπαράσταση της γνώσης και χρησιμοποιείται κυρίως για τη λεπτομερή ανάλυση των κειμένων. Το πιο γνωστό σημασιολογικό δίκτυο στο χώρο των βιοϊατρικής πληροφορικής είναι το UMLS. Όπως και άλλα δίκτυα, τα σημασιολογικά δίκτυα αποτελούνται από κόμβους με συνδέσεις μεταξύ τους. Οι κόμβοι αντιπροσωπεύουν τις έννοιες. Μια έννοια είναι μια αφηρημένη κατηγορία, τα μέλη της οποίας είναι αντικείμενα που ομαδοποιούνται επειδή μοιράζονται κοινά χαρακτηριστικά γνωρίσματα ή ιδιότητες. Τα σημασιολογικά δίκτυα υποστηρίζουν ιεραρχικές και συσχετικές σχέσεις μεταξύ των κόμβων, όπως είναι για παράδειγμα η σχέση μεταξύ των εννοιών δολοφονία και θάνατος, που μπορεί να περιγραφεί ως εξής: δολοφονία <είναι μια αιτία> θάνατος. Η αντίστροφη σχέση μπορεί να εκφραστεί: θάνατος <προκαλείται από> δολοφονία. Τα αντικείμενα καλούνται *πραγματώσεις* (instances) της έννοιας. Οι συνδέσεις εμφανίζονται σε ζεύγη για να αντιπροσωπεύσουν μια σχέση και την αντίστροφη σχέση της. Για παράδειγμα, η έννοια «μηρός» συσχετίζεται με «το πάνω μέρος της έννοιας πόδι» με τη σχέση <has_location>, ενώ η αντίστροφη της σχέση <location_of> συσχετίζει αντίστοιχα «το πάνω μέρος του ποδιού» με την έννοια «μηρός».

Αναφέρονται στη συνέχεια ορισμένα παραδείγματα οντολογιών και σημασιολογικών δικτύων που κυριαρχούν στο πεδίο της βιοϊατρικής.

3.2.1 FMA

Μια από τις πιο συστηματικά δομημένες οντολογίες στη βιοϊατρική είναι το Θεμελιώδες Μοντέλο Ανατομίας (Foundational Model of Anatomy ή FMA), το οποίο ξεκίνησε ως μια οντολογία πεδίου (domain ontology) με κλάσεις και σχέσεις που αναφέρονται στη δομική οργάνωση του ανθρώπινου σώματος [27]. Οι εκδότες του έχουν περιγράψει εκτενώς την αυστηρή προσέγγιση που χρησιμοποίησαν για την ανάπτυξή της, στηριζόμενοι σε ένα σύνολο δηλωμένων αρχών, σε σχήματα ύψιστης στάθμης, στους αριστοτελικούς ορισμούς και στο φορμαλισμό πλαισίων (frame-based). Ταυτόχρονα, προσπάθειες για τη μετατροπή της πλαισιοκεντρικής αναπαράστασης της FMA σε μια αναπαράσταση που βασίζεται στην

περιγραφική λογική (description logic), χρησιμοποιώντας το φορμαλισμό OWL, έχουν αποδώσει και η FMA σε μορφή OWL είναι πλέον διαθέσιμη. Αν και αναπτύχθηκε αρχικά ως επέκταση του περιεχομένου του UMLS που αφορά τις ανατομικές έννοιες, η FMA προτείνεται τώρα ως *οντολογία αναφοράς* χρήσιμη για τη σύνδεση των διαφορετικών απόψεων στον τομέα της ανατομίας, την εναρμόνιση των υπαρχουσών και αναδυόμενων οντολογιών στη βιοϊατρική πληροφορική και την παροχή ενός προτύπου αναπαράστασης των βιολογικών λειτουργιών, που βασίζεται στην έννοια του δομημένου σχεδίου τύπου. Για παράδειγμα, ο ορισμός της ανατομικής δομής στην FMA εκφράζεται με λογικούς όρους ως εξής:

Η Ανατομική Δομή

<είναι> υλική φυσική ανατομική οντότητα
(η οποία) <έχει> μια εγγενώς τρισδιάστατη μορφή
παράγεται από τη συντονισμένη έκφραση
των ίδιων των δομικών γονιδίων του οργανισμού
<αποτελείται> από μέρη που είναι ανατομικές δομές
και <σχετίζονται χωρικά> η μία με την άλλη σε σχήματα που
<καθορίζονται> από τη συντονισμένη έκφραση των γονιδίων.

Σε αυτό τον ορισμό, η υλική φυσική ανατομική οντότητα είναι το γένος, στο οποίο ανήκει η ανατομική δομή, ενώ τα υπόλοιπα μέρη της περιγραφής είναι τα διαφοροποιητικά χαρακτηριστικά, που διακρίνουν την ανατομική δομή από οποιουδήποτε άλλους τύπους που επίσης μπορούν να υπαχθούν στην κλάση «υλική φυσική ανατομική οντότητα».

Η FMA αποκαλείται *θεμελιώδης* από τους υπεύθυνους ανάπτυξης για δύο λόγους:

1. η ανατομία είναι θεμελιώδης σε όλους τους τομείς της βιοϊατρικής και
2. οι ανατομικές έννοιες και οι σχέσεις που καλύπτονται από την FMA μπορούν να γενικευτούν σε όλους αυτούς τους τομείς.

Με τον όρο «ανατομική έννοια» εννοείται εδώ μια μονάδα γνώσης που αναφέρεται σε μια ανατομική οντότητα. Η FMA περιέχει σήμερα πάνω από 80.000 ευδιάκριτες ανατομικές έννοιες, αναπαριστώντας δομές που κυμαίνονται στο μέγεθος από μακρομοριακά συστήματα και τμήματα κυττάρων σε μεγαλύτερα μέρη του σώματος. Αυτές οι έννοιες συνδέονται με περισσότερους από 110.000 όρους και συσχετίζονται μεταξύ τους με περισσότερες από 1,5 εκατομμύριο πραγματώσεις και 176 είδη σχέσεων. Έτσι η FMA, σήμερα, θεωρείται η μεγαλύτερη και πληρέστερη οντολογία αναφοράς.

3.2.2 *GALEN*

Ο στόχος του έργου GALEN, εξ αρχής, ήταν να δημιουργήσει επαναχρησιμοποιήσιμους ορολογικούς πόρους για τα κλινικά συστήματα. Στην καρδιά του GALEN βρίσκεται το Common Reference Model (CRM), μια οντολογία που διατυπώνεται με ένα εξειδικευμένο φορμαλισμό περιγραφικής λογικής, το GRAIL. Οι υπεύθυνοι ανάπτυξης έχουν περιγράψει τα οντολογικά ζητήματα που αντιμετώπισαν, καθώς επίσης και τις βασικές αρχές και τις συγκεκριμένες μεθόδους που χρησιμοποίησαν για να εξετάσουν τις διάφορες προκλήσεις κατά τη μοντελοποίηση του. Μερικά από τα πιο ενδιαφέροντα προβλήματα που κλήθηκαν να αντιμετωπίσουν είναι ο χειρισμός της αβεβαιότητας (uncertainty), η αναπαράσταση της γνώσης που αφορά τις ασθένειες και η αναπαράσταση των προκαθορισμένων επιλογών και των εξαιρέσεων. Σήμερα, δεδομένου ότι το πρόγραμμα κινείται μακριά της χρηματοδοτούμενης έρευνας από την ΕΕ, τα μέλη των κοινοπραξιών που συμμετείχαν στις προηγούμενες φάσεις του προγράμματος, έχουν σχηματίσει το OpenGALEN, ένα ίδρυμα με σκοπό να καταστήσει δυνατή την εκμετάλλευση και την περαιτέρω ανάπτυξη της τεχνολογίας GALEN [28]. Έτσι σήμερα ο στόχος του OpenGALEN είναι η προώθηση της υγείας μέσω τόνωσης της χρήσης και ανάπτυξης της τεχνολογίας GALEN. Αυτό θα επιτευχθεί έχοντας την εμπειρία, από την GALEN, ως βάση για τη διδασκαλία, την κατάρτιση και τις υπηρεσίες στον τομέα της ιατρικής ορολογίας, της γλώσσας, της γνώσης και των πληροφοριών.

3.2.3 *Gene Ontology*

Η Οντολογία Γονιδίων (Gene Ontology ή GO) δημιουργήθηκε με στόχο να καλύψει την ανάγκη για συνεπή αναπαράσταση των πληροφοριών που αφορούν τα προϊόντα των γονιδίων σε διαφορετικές βάσεις δεδομένων [29]. Το έργο ξεκίνησε ως συνεργασία μεταξύ των υπεύθυνων ανάπτυξης τριών πρότυπων βάσεων σχετικά με οργανισμούς: τη FlyBase (*Drosophila*), τη Saccharomyces Genome Database (SGD) και τη Mouse Genome Database (MGD). Έκτοτε, έχει ενσωματώσει πολλές βάσεις δεδομένων, συμπεριλαμβανομένων των σημαντικότερων αρχείων γονιδιωμάτων σε όλο τον κόσμο. Η χρήση των όρων της οντολογίας GO από διάφορες συνεργαζόμενες βάσεις δεδομένων διευκολύνει τη διατύπωση ομοιόμορφων ερωτήσεων σε αυτές. Από δομικής απόψεως, η GO διαιρείται σε τρεις οντολογίες, οι κορυφαίοι κόμβοι των οποίων είναι οι: Cellular component (Συστατικό κυττάρου), Molecular function (Μοριακή λειτουργία) και Biological process (Βιολογική διεργασία). Οι τρεις επιμέρους οντολογίες επιτρέπουν την περιγραφή των προϊόντων γονιδίων σε σχέση με αυτές τις κατηγορίες, δηλαδή επιτρέπουν τη διατύπωση απαντήσεων στους τρεις σημαντικότερους τύπους ερωτήσεων που προκύπτουν όταν ανακαλύπτεται ένα νέο γονιδιακό προϊόν:

1. Σε ποιο σημείο του κυττάρου βρίσκεται;
2. Ποιες λειτουργίες επιτελεί στο μοριακό επίπεδο;
3. Σε ποιες βιολογικές διεργασίες συμβάλλουν αυτές οι λειτουργίες;

Οι οντολογίες κατασκευάζονται με βάση τις λογικές σχέσεις υπαγωγής, δηλαδή τη σχέση γένους-είδους *is_a* και τη μεριστική σχέση *part_of*. Αυτά τα τρία δομημένα δίκτυα αντιμετωπίζονται στη GO ως χωριστές οντολογίες, ενώ καμία οντολογική σχέση δεν καθορίζεται μεταξύ τους. Ένα από τα βασικότερα προβλήματα που έχουν εντοπιστεί στην οντολογία GO είναι η ασυνεπής διαχείριση των σχέσεων μεταξύ των εννοιών και κυρίως στη σχέση *is_a*. Παρά τους περιορισμούς της, η GO έχει επιτύχει τη διαδεδομένη χρήση στη βιολογική κοινότητα και καταβάλλονται προσπάθειες να αναπαρασταθεί σε κάποιο μορμαλισμό περιγραφικής λογικής ώστε να βελτιωθεί η καταλληλότητά της για χρήση από υπολογιστές.

3.2.4 UMLS

Το Σύστημα Ενοποιημένης Ιατρικής Γλώσσας (Unified Medical Language System ή UMLS), που έχει αναπτύξει η Εθνική Βιβλιοθήκη Ιατρικής (National Library of Medicine ή NLM) των Η.Π.Α. έχει ως σκοπό να ξεπεραστούν τα προβλήματα ετερογένειας μεταξύ της ορολογίας που χρησιμοποιούν διαφορετικοί οργανισμοί στον τομέα της υγείας και να διευκολυνθεί η ανάπτυξη υπολογιστικών συστημάτων, που συμπεριφέρονται σαν να «καταλαβαίνουν» τη σημασία της γλώσσας της βιοϊατρικής και της υγείας [30]. Με αυτό το στόχο, η NLM παράγει και διανέμει τις πηγές γνώσης UMLS (βάσεις δεδομένων) και τα σχετικά εργαλεία λογισμικού (προγράμματα) προς χρήση για την ανάπτυξη συστημάτων ηλεκτρονικών πληροφοριών που δημιουργούν, επεξεργάζονται, ανακτούν και ενσωματώνουν στοιχεία βιοϊατρικής και πληροφορίες σχετικά με τον τομέα της υγείας. Οι πηγές γνώσης UMLS έχουν σχεδιαστεί για πολλές χρήσεις. Ο Μεταθησαυρός (Metathesaurus) είναι μια εκτενής, πολυχρηστική και πολύγλωσση εννοιοκεντρική βάση δεδομένων, που έχει σχεδιαστεί για πολλές εφαρμογές και χρήσεις και περιέχει έννοιες και όρους από περίπου 100 διαφορετικά λεξιλόγια, ταξινομίες και θησαυρούς. Ο Μεταθησαυρός περιέχει πληροφορίες για τις βιοϊατρικές και σχετικές με την υγεία έννοιες, τα διάφορα ονόματά τους και τις σχέσεις μεταξύ τους. Χτίζεται από τις ηλεκτρονικές εκδόσεις πολλών διαφορετικών θησαυρών, τις ταξινομήσεις και τους καταλόγους ελεγχόμενων ορολογιών που χρησιμοποιούνται στην περίθαλψη. Επιπρόσθετα η ανάπτυξή του ενισχύεται από τις υγειονομικές υπηρεσίες, τις στατιστικές υπηρεσίες δημόσιας υγείας, τη βιοϊατρική βιβλιογραφία, ή/και την έρευνα βασικών, κλινικών και υγειονομικών υπηρεσιών. Οι πηγές αυτές αναφέρονται ως "λεξιλόγια πηγής" του Metathesaurus. Το Σημασιολογικό Δίκτυο (Semantic Network) παρέχει μια κατηγοριοποίηση των εννοιών, που αντιπροσωπεύονται στο

Μεταθησαυρό του UMLS, σε σημασιολογικούς τύπους και ένα σύνολο σχέσεων που υφίστανται μεταξύ αυτών των σημασιολογικών τύπων. Η τρέχουσα έκδοση του Σημασιολογικού Δικτύου περιέχει 135 σημασιολογικούς τύπους (ως κόμβους) και 54 σχέσεις (ως συνδέσεις μεταξύ των κόμβων). Οι σημασιολογικοί τύποι καθορίζονται με βάση τις περιγραφές και με τη βοήθεια άλλων πληροφοριών, που είναι εγγενείς στις ιεραρχίες του. Βασικοί σημασιολογικοί κόμβοι στην ιεραρχία του UMLS είναι οι: Οργανισμός, Ανατομική Δομή, Βιολογική Λειτουργία, Χημική Ουσία, Συμβάν, Φυσικό αντικείμενο, Έννοια ή Ιδέα.

Τα τελευταία χρόνια έχει δημοσιευτεί πλήθος μελετών που αξιολογούν τη χρησιμότητα του UMLS ως πόρου ορολογίας και γνώσης σε ποικίλες εφαρμογές, από τη μετάφραση όρων ως την κατασκευή οντολογιών πεδίου (όπως οι Foundational Model of Anatomy και Gene Ontology, οι οποίες έχουν στηριχθεί στο Σημασιολογικό Δίκτυο του UMLS). Άλλες μελέτες έχουν εστιάσει στο ζήτημα του ρόλου του ίδιου του Σημασιολογικού Δικτύου του UMLS ως οντολογίας της βιοϊατρικής περιοχής. Μερικά από τα σημαντικότερα προβλήματα, που εντοπίστηκαν, αφορούν την ασυμβατότητα που παρατηρήθηκε ανάμεσα σε οντολογίες γενικής γλώσσας ύψιστης στάθμης (top-level ontologies), όπως η οντολογία Wordnet και CyC, και στο UMLS. Έτσι, παρατηρήθηκε ότι δύο κατηγορίες που έχουν τα ίδια χαρακτηριστικά και τον ίδιο εντατικό ορισμό στο Wordnet και στο UMLS, παρ' όλα αυτά δεν περιλαμβάνουν τις ίδιες εκτάσεις. Για παράδειγμα, μολονότι το «Σύμπτωμα» έχει ισοδύναμους ορισμούς στο WordNet και στο UMLS, το «Σύμπτωμα» στο WordNet περιλαμβάνει την «εγκεφαλίτιδα» αλλά και άλλους όρους, οι οποίοι στο UMLS είναι ταξινομημένοι ως «Ασθένεια ή Σύνδρομο». Τελευταία προτείνονται αναθεωρήσεις στο Σημασιολογικό Δίκτυο του UMLS, που αφορούν τη σαφή και ευκρινή διάκριση των ιεραρχικών σχέσεων is_a και part_of και κατά συνέπεια των σημασιολογικών τύπων που αναφέρονται σε είδη και μέρη, διάκριση που δεν εφαρμόζεται συστηματικά στις κατηγοριοποιήσεις του Σημασιολογικού Δικτύου.

3.3 Διαλειτουργικότητα στην Ιατρική

Ένα από τα πιο καίρια ζητήματα προς επίλυση αναπτύσσοντας πληροφοριακά συστήματα, ιδιαίτερα όταν σχετίζονται με τον ευρύτερο τομέα της ιατρικής, είναι η επίτευξη διαλειτουργικότητας. Διαλειτουργικότητα όχι μόνο σε πρωταρχικό επίπεδο (λεξιλόγιο κ.ά.), αλλά και σε εννοιολογικό επίπεδο. Οι απαιτήσεις αυτές υποδεικνύουν τη χρήση σημασιολογικών μεθόδων, όπως οντολογίες και συστήματα γνώσης, και την αξιοποίηση των σύγχρονων τεχνολογιών Σημασιολογικού Ιστού, όπως οι γλώσσες αναπαράστασης RDF και OWL.

Η διαλειτουργικότητα στην πληροφορική ορίζεται ως η ικανότητα δύο ή περισσότερων πληροφοριακών συστημάτων να ανταλλάσσουν πληροφορία και να χρησιμοποιούν την

πληροφορία που έχει ανταλλαχθεί. Για την επίτευξη πραγματικής διαλειτουργικότητας, τα συστήματα δεν πρέπει μόνο να είναι σε θέση να ανταλλάσσουν πληροφορία, χρησιμοποιώντας κοινές αρχιτεκτονικές, μεθόδους και πλαίσια επικοινωνίας, αλλά να την ερμηνεύουν σωστά και να τη χρησιμοποιούν κατάλληλα χρησιμοποιώντας κοινούς τύπους δεδομένων, ορολογίες και οντολογίες και σχήματα κωδικοποίησης. Γίνεται λοιπόν διάκριση μεταξύ λειτουργικής και σημασιολογικής διαλειτουργικότητας. Η δυνατότητα που δίνεται για επίτευξη διαλειτουργικότητας σε σημασιολογικό επίπεδο με χρήση των αναπτυσσόμενων τεχνολογιών του Σημασιολογικού Ιστού έχει αρκετές προοπτικές και πλεονεκτήματα:

- Ανεξαρτησία από πρότυπα και απεικονιστικές τροπικότητες. Το κοινό μοντέλο αναπαράστασης γνώσης που προσφέρει η χρήση των RDF τριάδων επιτρέπει τη διαχείριση της πληροφορίας με κοινό τρόπο, είτε αυτή έχει κωδικοποιηθεί σε αρχεία κειμένου, είτε σε αρχεία DICOM, είτε σε άλλα ιατρικά πρότυπα όπως openEHR. Ειδικά για τις ιατρικές εικόνες, η προσέγγιση αυτή προσφέρει κοινή αντιμετώπιση της ρητής γνώσης που περιέχουν οι εικόνες ανεξάρτητα από την τροπικότητα (modality) της απεικόνισης.
- Συνεργασία νοσοκομειακών κέντρων και κλινικών. Η σημασιολογικές τεχνολογίες μπορούν να διαδραματίσουν το ρόλο μίας «ομπρέλας» που αναλαμβάνει τη διασύνδεση των ενδεχομένως διαφορετικών πληροφοριακών συστημάτων διαφόρων ιατρικών μονάδων και επιτρέπει τη συνεργασία των επιστημόνων υγείας μέσω μίας κοινής πλατφόρμας. Η συνεργασία των παραπάνω κέντρων με τους φορείς υγείας και κοινωνικής ασφάλισης μπορεί επίσης να βοηθήσει στη μείωση του βάρους των – δημοφιλών στην Ελλάδα - γραφειοκρατικών διαδικασιών.
- Βελτίωση ακρίβειας διάγνωσης. Τα συστήματα που βασίζονται σε κοινά πρότυπα επιτυγχάνουν το συσχετισμό σε πραγματικό χρόνο συμπτωμάτων, αποτελεσμάτων εξετάσεων και εξατομικευμένων ιατρικών ιστορικών για τη συστηματική, διασταυρωμένη διάγνωση και την ανίχνευση τυχόν σφαλμάτων ενός απομονωμένου συστήματος.
- Στατιστικά αποτελέσματα μεγάλης κλίμακας. Η διαλειτουργικότητα μέσω των τεχνολογιών του Παγκόσμιου Ιστού δεν περιορίζεται στα στενά όρια των χωρών αλλά εκτείνεται σε παγκόσμια κλίμακα. Ήδη μεγάλες οντολογίες, όπως η FMA, περιέχουν ισοδύναμες αναπαραστάσεις εννοιών σε άλλες γλώσσες εκτός της αγγλικής, προωθώντας τη διαλειτουργικότητα σε επίπεδο γλώσσας. Ένα αισιόδοξο μελλοντικό σενάριο περιλαμβάνει, για παράδειγμα, τη συνεργασία των ιατρικών και βιοϊατρικών κέντρων ανά τον κόσμο για την αντιμετώπιση, μέσω στατιστικών μετρήσεων και πειραμάτων παγκόσμιας κλίμακας, μιας ενδεχόμενης πανδημίας.

Καθώς το κέντρο βάρους της ιατρικής και βιοϊατρικής έρευνας και πράξης μετατοπίζεται από τα μεμονωμένα συστήματα προς την ενσωμάτωση μεγάλου όγκου δεδομένων μέσω των τεχνολογιών του Σηματολογικού Ιστού, η σηματολογική διαλειτουργικότητα θα διαδραματίζει όλο και πιο αποφασιστικό ρόλο στην επιτυχία του παραπάνω εγχειρήματος, απώτερος στόχος του οποίου δεν είναι άλλος από την ουσιαστική βελτίωση των διαδικασιών ιατρικής διάγνωσης και θεραπείας.

4

Συστήματα Περιγραφής και Αναζήτησης

Ιατρικών Εικόνων

Στο παρόν κεφάλαιο παρουσιάζονται υπάρχοντα συστήματα περιγραφής και αναζήτησης ιατρικών εικόνων, τα οποία χωρίζονται σε δύο κατηγορίες. Στην πρώτη κατηγορία ανήκουν τα ευρέως χρησιμοποιούμενα συστήματα PACS, ενώ στη δεύτερη ανήκουν τα συστήματα σημασιολογικής περιγραφής και αναζήτησης. Παρότι τα τελευταία βρίσκονται ακόμα σε πειραματικό στάδιο και δεν έχουν ενσωματωθεί σε λειτουργικά πληροφοριακά συστήματα, παρουσιάζονται οι προσπάθειες που έχουν γίνει και έχουν ανοίξει το δρόμο προς μια σημασιολογική προσέγγιση.

4.1 PACS

Τα Συστήματα Αρχειοθέτησης Εικόνων και Επικοινωνίας (Picture Archiving and Communication Systems ή PACS) είναι είδος ιατρικών πληροφοριακών συστημάτων τα οποία επιτρέπουν οικονομική αποθήκευση, γρήγορη ανάκληση εικόνων, ενιαία πρόσβαση σε εικόνες διαφορετικής τροπικότητας και ταυτόχρονη πρόσβαση από πολλαπλές τοποθεσίες. Τα PACS έχουν γίνει ευρέως αποδεκτά στο χώρο της διαχείρισης ιατρικών εικόνων και πολλές εταιρίες συναγωνίζονται για την επικράτηση στην αγορά, προσφέροντας παραπλήσια προϊόντα. Είναι η ψηφιακή λύση στην αρχειοθέτηση και διαχείριση ιατρικών εικόνων, που έρχεται να αντικαταστήσει τα παλιά συστήματα που επέτρεπαν καταγραφή των

απεικονιστικών δεδομένων μόνο σε ταινίες ή ειδικό χαρτί (κλασική ακτινογραφία). Η καθολική μορφή αποθήκευσης και μεταφοράς εικόνων για τα PACS είναι το πρότυπο DICOM.

Ένα τυπικό σύστημα PACS αποτελείται από τέσσερα βασικά συστατικά μέρη [31]:

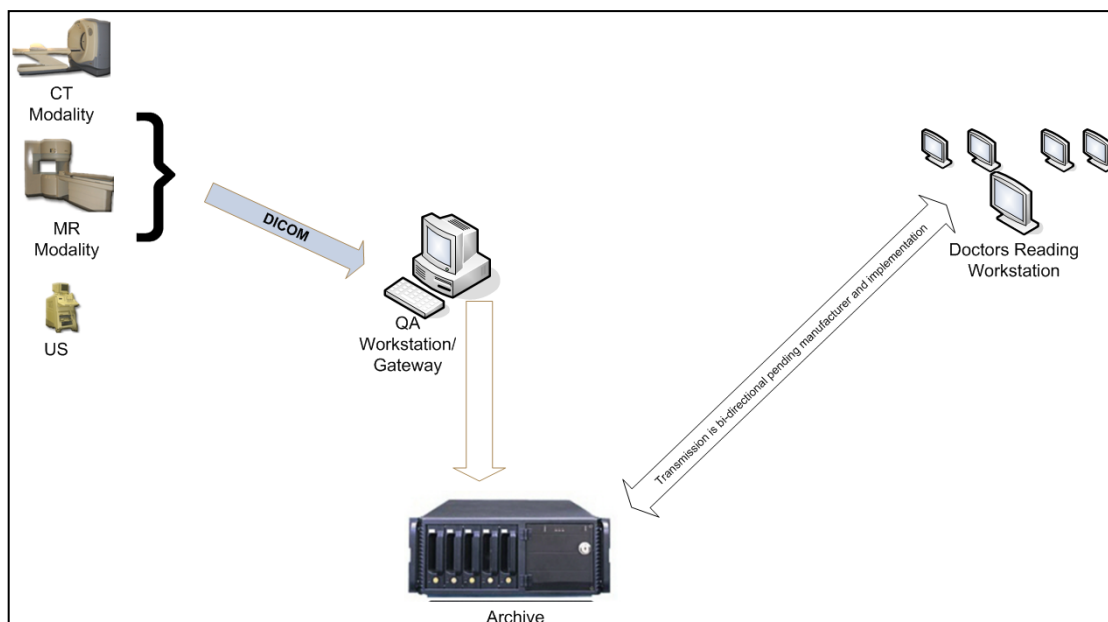
1. Απεικονιστική τροπικότητα. Αφορά το είδος της εικόνας που παράγεται από το μηχάνημα καταγραφής και μπορεί, για παράδειγμα, να είναι ακτινογραφία, υπερηχογράφημα, μαγνητική τομογραφία κ.ά.
2. Δίκτυο μετάδοσης. Αναλαμβάνει την ασφαλή μετάδοση των απεικονιστικών δεδομένων, καθώς και των δεδομένων που αφορούν το ασθενή (όπως δημογραφικές πληροφορίες) και την εικόνα (όπως περιγραφή μελέτης, παραπέμποντες ιατροί κ.ά.).
3. Αρχείο. Στην κεντρική συσκευή αποθήκευσης αποθηκεύονται οι εικόνες και σε κάποιες περιπτώσεις αναφορές και μετρήσεις. Είναι σημαντικό και σε πολλές περιπτώσεις επιβάλλεται από τον νόμο να υπάρχουν μονάδες εφεδρικής αποθήκευσης, ώστε να ανασύρονται αντίγραφα σε περίπτωση απώλειας των πρωτότυπων.
4. Σταθμοί ανάγνωσης εικόνων. Πρόκειται για σταθμούς εργασίας όπου γίνεται η ερμηνεία και επανεξέταση των εικόνων.

Στο διάγραμμα ροής εργασιών της Εικόνα 4.1.1 παρατηρούμε τα επιμέρους εξαρτήματα τα οποία απαρτίζουν ένα τυπικό σύστημα PACS. Ο σταθμός-πύλη εξασφάλισης ποιότητας (QA Workstation/Gateway) είναι ένα σημείο ελέγχου στο οποίο ελέγχονται - και εισάγονται, αν χρειάζεται - όλες οι πληροφορίες και για τον ασθενή και την εικόνα πριν από την εισαγωγή τους στις επικεφαλίδες των DICOM αρχείων και την αποθήκευση αυτών στο Αρχείο (Archive).

Η εισαγωγή δεδομένων σχετικά με τις ιατρικές εικόνες γίνεται χρησιμοποιώντας τα πεδία των DICOM επικεφαλίδων, εισάγοντας δηλαδή τιμές στα πεδία αυτά. Ενώ οι περισσότερες τιμές εισάγονται αυτόματα από τις ρυθμίσεις των μηχανημάτων καταγραφής, κάποια πεδία, όπως η περιοχή του σώματος που εξετάζεται (πεδίο BodyPartExamined) απαιτούν τη χειροκίνητη εισαγωγή τιμών. Τα περισσότερα από αυτά τα πεδία είναι ελεύθερου κειμένου, αν και προσπάθειες γίνονται για επιβολή συγκεκριμένων επιτρεπτών τιμών, ώστε να προωθείται η επαναχρησιμοποίηση και η διαλειτουργικότητα.

Τα ερωτήματα για εικόνες και η ανάκληση τους βασίζονται και πάλι στο DICOM πρότυπο. Συγκεκριμένα πραγματοποιούνται με χρήση αντικειμένων DICOM, τα οποία μοιάζουν με τις εικόνες DICOM, έχουν όμως διαφορετικές ετικέτες πεδίων. Ο χρήστης συνδέεται με τον PACS εξυπηρετητή και συμπληρώνει τις τιμές των πεδίων για τα οποία επιθυμεί να αναζητήσει και πεδία τα οποία επιθυμεί να λάβει ως απάντηση. Στη συνέχεια, το αντικείμενο του ερωτήματος αποστέλλεται στον εξυπηρετητή με την μορφή της, συμβατής με το

πρωτόκολλο DICOM, αίτησης C-FIND. Αν η αίτηση είναι επιτυχής, το σύστημα πραγματοποιεί μία C-MOVE αίτηση, με την οποία η μελέτη αποστέλλεται από το αρχείο στην καθορισμένη συσκευή απεικόνισης. Με τον τρόπο αυτό μπορούν, για παράδειγμα, να ανακληθούν επιτυχώς παλαιότερες μελέτες του υπό-εξέταση ασθενούς.



Εικόνα 4.1.1: Διάγραμμα ροής εργασιών ενός τυπικού PACS [31].

Είναι σημαντικό να τονιστεί, τέλος, ότι τα PACS σπάνια λειτουργούν ανεξάρτητα αλλά αντίθετα παρέχουν την ηλεκτρονική πλατφόρμα εικόνων που επικοινωνεί με άλλα συστήματα στα οποία ενσωματώνεται, όπως Πληροφοριακά Συστήματα Νοσοκομείων (Hospital Information Systems ή HISS) και Ακτινολογικά Πληροφοριακά Συστήματα (Radiology Information Systems ή RIS).

4.2 Σημαιολογικά Συστήματα

4.2.1 RadSem

Το RadSem είναι μία αυτόνομη (stand-alone) εφαρμογή για σημαιολογική επισημείωση και ανάκτηση ιατρικών εικόνων που αναπτύχθηκε στα πλαίσια του MEDICO project. Το συγκεκριμένο εργαλείο αξιοποιεί την οντολογία MEDICO, στην οποία έχουν ενσωματωθεί τμήματα από έγκυρες υπάρχουσες οντολογίες όπως η FMA και ταξινομήσεις όπως η ICD-10 και το Βασικό Λεξικό Ακτινολογίας (Standart Radiology Lexicon ή RadLex¹) έτσι ώστε να

¹ <http://www.rsna.org/radlex/>

είναι αξιοποιηθεί η προ-υπάρχουσα επίσημα κωδικοποιημένη γνώση. Η οντολογία MEDICO είναι αρκετά ευρεία, καλύπτει έννοιες που παρουσιάζονται σε ποικίλες φάσεις της κλινικής διαδικασίας και χρησιμοποιείται σε αρκετά βήματα της επισημείωσης και ανάκτησης εικόνων. Οι βασικές λειτουργίες – βήματα της εφαρμογής παρουσιάζονται παρακάτω [32].

- Δεδομένα, όπως όνομα ασθενούς, ηλικία, περιοχή του σώματος και παράμετροι εικόνας, εξάγονται αυτόματα από τις πληροφορίες επικεφαλίδας κάθε αρχείου εικόνας που έχει κωδικοποιηθεί με βάση το πρότυπο ψηφιακών αρχείων DICOM.
- Τα μεταδεδομένα αυτά χρησιμοποιούνται για την απλοποίηση της χειροκίνητης επισημείωσης παρέχοντας ένα προεπιλεγμένο υποσύνολο από έννοιες που μπορούν να επισημειωθούν σε κάθε εικόνα. Επίσης παρέχεται μια οπτικοποίηση μέσω του φωτισμού της περιοχής του σώματος που εμφανίζεται στην εικόνα σε μία ανθρώπινη σιλουέτα.
- Τα εξαγόμενα μεταδεδομένα συνδυάζονται με τις επισημειωμένες έννοιες των οντολογιών και κλινικά ευρήματα ώστε να παρέχουν μια ενοποιημένη όψη του ιατρικού ιστορικού του ασθενούς.
- Στον τομέα της αναζήτησης, γίνεται επέκταση των ερωτημάτων με βάση τη δομή των ιατρικών οντολογιών.
- Η οντολογία για τη διαχείριση των κλινικών δεδομένων επιτρέπει το συνδυασμό ασθενών, ιατρικών εικόνων και επισημειώσεων σε μια ενιαία και περιεκτική λίστα αποτελεσμάτων.
- Οι ιατρικές επισημειώσεις επεκτείνονται μέσω συνδέσεων με εξωτερικούς πόρους όπως η Wikipedia για την παροχή επιπρόσθετης πληροφορίας.

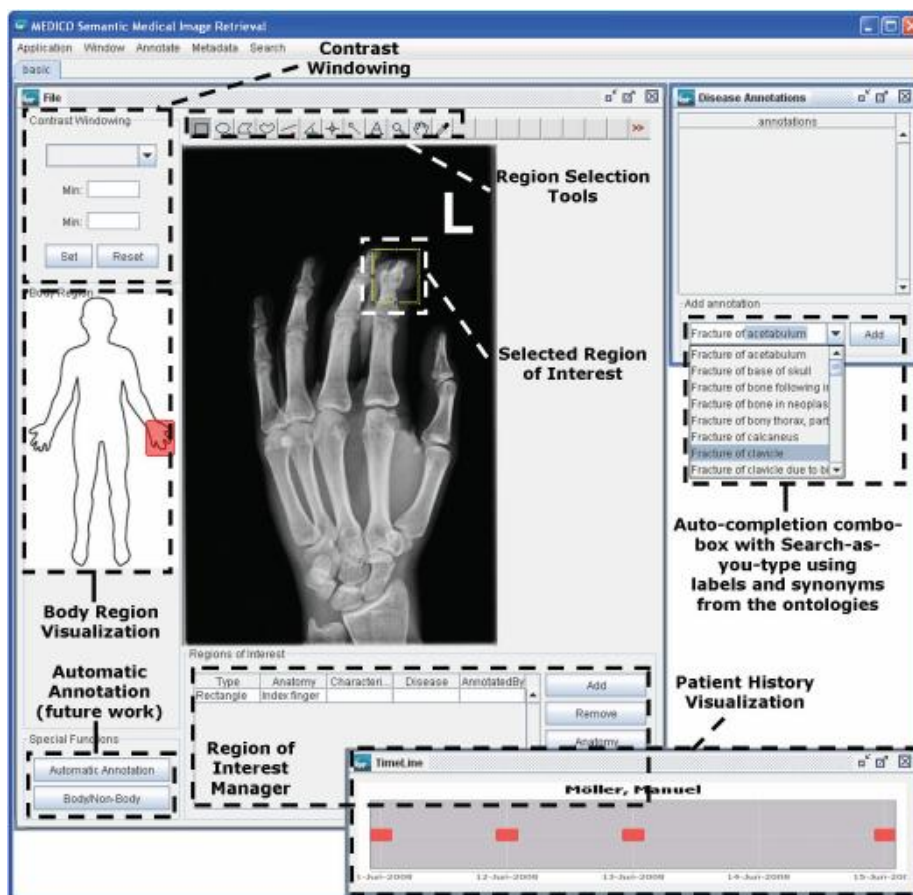
Η εφαρμογή απαρτίζεται από δύο διεπαφές χρήστη καθεμία από τις οποίες επιτελεί μια από τις δύο βασικές λειτουργίες: επισημείωση εικόνων και αναζήτηση με βάση τις επισημειώσεις και τα χαρακτηριστικά των εικόνων και των ασθενών. Η βασική οθόνη της επισημείωσης φαίνεται στην Εικόνα 4.2.1.

Οι εικόνες φορτώνονται και εμφανίζονται χρησιμοποιώντας την προγραμματιστική διεπαφή του προγράμματος επεξεργασίας εικόνας ImageJ¹. Με τη βοήθεια της συγκεκριμένης βιβλιοθήκης ο χρήστης μπορεί να σημειώσει Περιοχές Ενδιαφέροντος (Regions of Interest ή ROIs) πάνω στην εικόνα. Για κάθε ένα από τα ROIs μπορεί να επισημειωθεί ξεχωριστά η ανατομία, η ασθένεια και το παρατηρούμενο χαρακτηριστικό.

Για την επισημείωση του ανατομικού χαρακτηριστικού χρησιμοποιείται η οντολογία FMA, ενώ οι έννοιες που αφορούν τα παρατηρούμενα χαρακτηριστικά αντλούνται από τα υποδέντρα modifier και image observation characteristic της οντολογίας RadLex. Για την

¹ <http://rsbweb.nih.gov/ij/>

επισημείωση της ασθένειας χρησιμοποιείται μια ελαφριά έκδοση της ICD-10 σε OWL μορφή, η οποία αναπτύχθηκε ειδικά για τις ανάγκες της εφαρμογής, καθώς την περίοδο της ανάπτυξης της εφαρμογής δεν υπήρχε διαθέσιμη η τωρινή έκδοση της ICD-10 σε OWL¹.



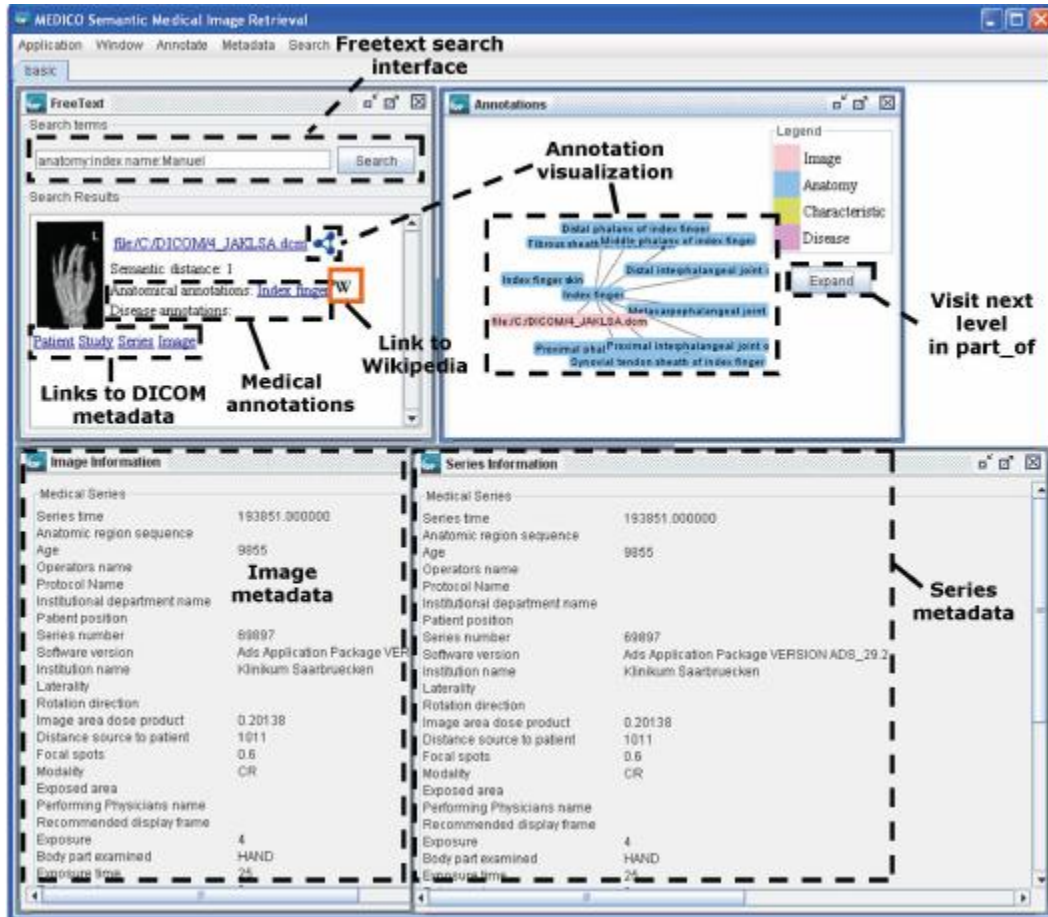
Εικόνα 4.2.1: Βασική οθόνη επισημείωσης.

Για την διευκόλυνση του χρήστη χρησιμοποιούνται λίστες που ενημερώνονται αυτόματα, καθώς ο χρήστης πληκτρολογεί τα πρώτα γράμματα του όρου, με κατάλληλες έννοιες από την εκάστοτε οντολογία, ώστε να επιλεγεί η καταλληλότερη έννοια, όπως φαίνεται στην Εικόνα 4.2.1. Ένα άλλο σημαντικό στοιχείο της διεπαφής είναι η οπτικοποίηση του ιστορικού του ασθενούς μέσω ενός γραφήματος, στο οποίο φαίνονται οι προηγούμενες εξετάσεις του ασθενούς σε χρονολογική σειρά μαζί με τις επισημειώσεις, όπου αυτές υπάρχουν.

Η διεπαφή αναζήτησης, που παρουσιάζεται στην Εικόνα 4.1.2, επιτρέπει τη δημιουργία υβριδικών ερωτημάτων σε μορφή ελεύθερου κειμένου, που συνδυάζουν την αναζήτηση για ιατρικές επισημειώσεις όσο και για συγκεκριμένες τιμές μεταδεδομένων όπως αυτά εξήχθησαν από την επικεφαλίδα των DICOM αρχείων. Για παράδειγμα, εάν επιθυμείται η

¹ <http://bioportal.bioontology.org/ontologies/44103>

αναζήτηση εικόνων που παρουσιάζουν «δείκτη» (ανατομικό χαρακτηριστικό) και ανήκουν στον ασθενή «Manuel» (μεταδεδομένο εικόνας) αρκεί να πληκτρολογήσει κανείς «anatomy:Index name:Manuel».



Εικόνα 4.2.2: Βασική οθόνη αναζήτησης.

Στα αποτελέσματα της αναζήτησης περιλαμβάνονται μικρογραφία της εικόνας που βρέθηκε, καθώς και όλες οι διαθέσιμες επισημειώσεις. Επίσης εμφανίζονται τα μεταδεδομένα της εικόνας και της σειράς εικόνων στην οποία ανήκει η συγκεκριμένη εικόνα που έχουν εξαχθεί από τις κεφαλίδες DICOM. Ενδιαφέρον χαρακτηριστικό είναι η οπτικοποίηση όλων των οντολογικών σχέσεων που διέπουν τις έννοιες με τις οποίες έχει επισημειωθεί η εικόνα σε ένα κοινό σημασιολογικό δίκτυο, κάτι που επιτρέπει την εξοικείωση του χρήστη με τη δομή των οντολογιών και την εξερεύνηση της «σημασιολογικής γειτονιάς» των επιλεγμένων εννοιών.

Η σημασιολογική αναζήτηση πραγματοποιείται χρησιμοποιώντας τη δομή των ιατρικών οντολογιών. Κατά την αναζήτηση με βάση ανατομικό χαρακτηριστικό γίνεται διατρέχεται η FMA με βάση τη μεταβατική σχέση <μέρος του> (part-of). Για παράδειγμα, εάν η αναζήτηση γίνεται με βάση τον όρο «δείκτης» όπως στο παραπάνω παράδειγμα και βρεθούν εικόνες με

επισημείωση «δείκτης του χεριού» τότε αυτές επιστρέφονται, καθώς ο δείκτης του χεριού είναι μέρος του δείκτη με βάση την κωδικοποίηση της FMA. Αντίθετα, η ICD-10 διατρέχεται με βάση τις υποκλάσεις, δηλαδή ανιχνεύονται όλες οι εικόνες που έχουν σημειωθεί με έννοια που είναι υποκλάση του όρου αναζήτησης, με βάση πάντα την ιεραρχία της ICD-10.

4.2.2 Annotation and Image Markup

Η Επισημείωση και Σήμανση Εικόνων (Annotation and Image Markup ή AIM) είναι ένα από τα έργα που έχει προωθήσει το Πλέγμα Βιοϊατρικής Πληροφορικής κατά του Καρκίνου (cancer Biomedical Informatics Grid ή caBIG) του Εθνικού Ινστιτούτου κατά του Καρκίνου των Η.Π.Α.. Σκοπός του AIM είναι η δημιουργία ενός πλαισίου επιτυχούς σημασιολογικής επισημείωσης ιατρικών εικόνων και η ανάπτυξη κατάλληλων εργαλείων ώστε να διευκολυνθεί η διαδικασία αυτή για τους ειδικούς [6].

Στο πλαίσιο αυτό, δημιουργήθηκε μία οντολογία σε OWL-DL για την αναπαράσταση των σχετικών οντοτήτων που απαιτούνται κατά την επισημείωση των ιατρικών εικόνων, με την ονομασία οντολογία AIM. Η οντολογία περιλαμβάνει ανατομικές δομές που απεικονίζονται στις εικόνες, παρατηρήσεις των ακτινολόγων για τις εικόνες (όπως «αδιαφάνεια» ή «πυκνότητα»), τις χωρικές περιοχές που μπορούν να απεικονιστούν σε εικόνες, όπως και άλλα μεταδεδομένα των εικόνων. Οι ανατομικές δομές και οι οπτικές παρατηρήσεις προμηθεύτηκαν από την οντολογία RadLex, με την ενσωμάτωση των κατάλληλων κομματιών της RadLex στη AIM οντολογία.

Επίσης, δημιουργήθηκε ένα πληροφοριακό μοντέλο – το «σχήμα AIM» – σε UML για την περιγραφή της ελάχιστης πληροφορίας που είναι απαραίτητο να καταγραφεί για την επιτυχή επισημείωση των εικόνων. Το μοντέλο αυτό στη συνέχεια μετασχηματίστηκε σε OWL αναπαράσταση για μεγαλύτερη διαλειτουργικότητα αλλά και σε σχήμα XML (αρχείο XSD) για την επικύρωση των στιγμιότυπων της οντολογίας που θα αποθηκεύονταν σε XML αρχεία.

Η επισημείωση με βάση την προσέγγιση του AIM στηρίζεται στην ιδέα ότι το πλαίσιο της εικόνας καθορίζει και τις ανάγκες επισημείωσης. Για παράδειγμα, σε ορισμένες περιοχές του σώματος, όπως ο θώρακας, μόνο ορισμένες ανατομικές δομές είναι κατάλληλες για αναφορά σε μία επισημείωση, όπως καρδιά, πνεύμονας ή πλευρά στο συγκεκριμένο παράδειγμα. Ομοίως, κάποιες παρατηρήσεις σε εικόνες μπορούν να πραγματοποιηθούν μόνο σε ιδιαίτερες ανατομικές δομές, για παράδειγμα, ένα κάταγμα μπορεί να παρατηρηθεί στα πλευρά αλλά όχι στον πνεύμονα.

Για να επιτευχθεί το παραπάνω, το εργαλείο επισημείωσης αρχικά παρακινεί τα χρήστη να ορίσει το πλαίσιο της επισημείωσης, για παράδειγμα «αξιολόγηση μορφήματος σε εικόνα αξονικής τομογραφίας στήθους». Στη συνέχεια, το σύστημα εισάγει το ορισμένο από το χρήστη πλαίσιο στην οντολογία και εκτελεί τον ταξινομητή (classifier) ώστε να εξάγει τα

πεδία δεδομένων από το AIM σχήμα που ο χρήστης πρέπει να συμπληρώσει για το συγκεκριμένο πλαίσιο επισημείωσης. Στο τέλος τα δεδομένα της επισημείωσης αποθηκεύονται σε αρχεία AIM XML, που παρέχουν μια ομοιόμορφη σύνταξη για την αναπαράσταση των μεταδεδομένων σε ένα κοινό πληροφοριακό μοντέλο.

Για την παροχή της μεγαλύτερης δυνατής διαλειτουργικότητας για τα αποθηκευμένα δεδομένα δίνεται η δυνατότητα μετατροπής τους σε μορφή OWL, έτσι ώστε να δημοσιευτούν στο διαδίκτυο και να γίνουν προσβάσιμα από διαδικτυακούς πράκτορες, αλλά και στα πρότυπα DICOM-SR και HL7-CDA, ώστε να είναι προσβάσιμα από τα νοσοκομειακά πληροφοριακά συστήματα. Ωστόσο, το AIM σχήμα διαθέτει ένα μοναδικό αναγνωριστή ο οποίος είναι διαθέσιμος σε όλες τις γλώσσες αναπαράστασης, έτσι ώστε η εικόνα συνδέεται με την επισημείωση ανεξάρτητα από τον τρόπο αναπαράστασης της τελευταίας.

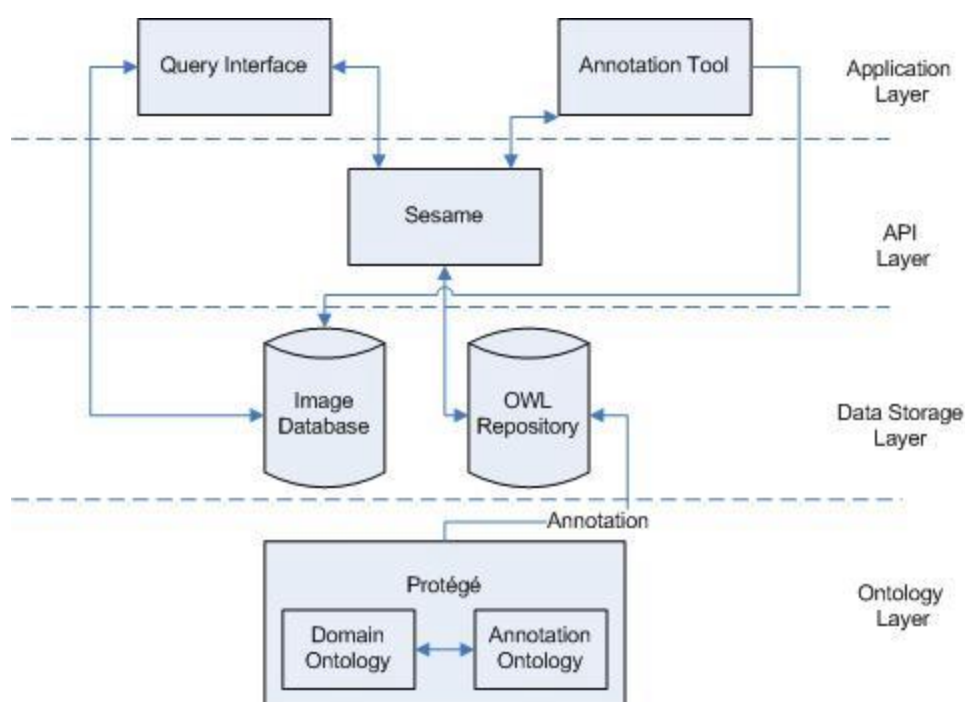
Αν και η κατεύθυνση προς την στοχευμένη επισημείωση με βάση το πλαίσιο της ιατρικής εικόνας φαντάζει σωστή, μένει να εκτιμηθεί με τη επισημείωση ενός αρκούντος μεγάλου αριθμού εικόνων αν πρακτικά βοηθάει τη διαδικασία επισημείωσης ή δυσχεραίνει την εργασία του επιστήμονα ακτινολόγου.

4.2.3 Πληροφοριακό Σύστημα Ιατρικών Εικόνων των Wei και Barnaghi

Το Πληροφοριακό Σύστημα για σημασιολογική αναζήτηση και ανάκτηση Ιατρικών Εικόνων αναπτύχθηκε στα πλαίσια της έρευνας των επιστημόνων Wang Wei και Payam Barnaghi του πανεπιστημίου του Nottingham της Μαλαισίας. Στόχος του συστήματος ήταν η βελτίωση της ακρίβειας ανάκλησης κατά την αναζήτηση ιατρικών εικόνων και η διασύνδεση των δεδομένων σχετικών με ιατρικές εικόνες, όπως δεδομένα παρακολούθησης ασθενών, παρόμοιες περιπτώσεις από καταγραφές άλλων ασθενών και σχετικά έγγραφα, σε μία κοινή, ολοκληρωμένη αναφορά που θα προσφέρει ουσιαστική βοήθεια στη διάγνωση και θεραπεία ασθενειών [7]. Αν και το σύστημα υλοποιήθηκε και δοκιμάστηκε για χρήση με ακτινογραφίες ανθρώπινου πνεύμονα και προδιαγραφές σχετικών διαταραχών, προσφέρει μια αρχιτεκτονική που μπορεί να χρησιμοποιηθεί για σημασιολογική αναζήτηση σε ένα πιο γενικευμένο ιατρικό πεδίο.

Η οντολογία πεδίου που αναπτύχθηκε σε OWL βασίστηκε σε έννοιες και σχέσεις που έχουν κωδικοποιηθεί στον τυποποιημένο και επίσημο μετα-θησαυρό του Εθνικού Ινστιτούτου κατά του Καρκίνου (National Cancer Institute ή NCI) των Η.Π.Α., ο οποίος περιέχει περιεκτικές έννοιες σχετικές με την ανθρώπινη ανατομία, ασθένειες, ευρήματα, δυσλειτουργίες και ιατρικές τεχνικές. Επίσης, αναπτύχθηκε μία οντολογία επισημείωσης για την κωδικοποίηση των δεδομένων της επισημείωσης, όπως τις Περιοχές Ενδιαφέροντος (ROIs) και τα μεταδεδομένα της εικόνας, όπως μέγεθος, ανάλυση κ.τ.λ..

Στην Εικόνα 4.2.3 παρουσιάζεται η αρχιτεκτονική του συστήματος. Το επίπεδο των οντολογιών (Ontology Layer) περιέχει τα επισημειωμένα δεδομένα και τις οντολογίες πεδίου σε μορφή RDF/XML. Το επίπεδο της αποθήκευσης δεδομένων (Data Storage Layer) παρέχει τις εγκαταστάσεις αποθήκευσης και χειρισμού των δεδομένων. Οι αναπαραστάσεις σε RDF/XML αποθηκεύονται σε μία αποθήκη RDF τριάδων, ενώ τα δυαδικά δεδομένα εικόνων αποθηκεύονται σε σχεσιακή βάση. Το επίπεδο προγραμματιστικής διεπαφής (API Layer) χρησιμοποιεί την προγραμματιστική διεπαφή της αποθήκης RDF τριάδων για να πραγματοποιήσει τα ερωτήματα στις RDF αναπαραστάσεις. Τέλος, το επίπεδο εφαρμογής (Application Layer) περιέχει μία διεπαφή για επισημείωση και μία διεπαφή για υποβολή ερωτημάτων.



Εικόνα 4.2.3: Η αρχιτεκτονική του συστήματος.

Το εργαλείο επισημείωσης φορτώνει την εικόνα και τις επισημειώσεις της, αν υπάρχουν, και επιτρέπει στο χρήστη εισάγει νέα δεδομένα ή να αλλάξει/ανανεώσει τα υπάρχοντα. Το εργαλείο επισημείωσης καθορίζει τις προδιαγραφές της επισημείωσης με βάση τη δομή της οντολογίας πεδίου και απαιτεί τον ορισμό περιγραφών με βάση τις έννοιες της οντολογίας.

Η διεπαφή υποβολής ερωτημάτων επιτρέπει στο χρήστη να περιηγηθεί στις έννοιες και τις σχέσεις της οντολογίας πεδίου και επιλέξει ανάμεσά τους. Στη συνέχεια ο μηχανισμός συλλογιστικής βρίσκει και επιστρέφει τις κατάλληλες εικόνες αλλά προχωρά και σε «σημασιολογικές συστάσεις» με βάση σημασιολογικές συσχετίσεις. Αυτό επιτρέπει στους ειδικούς της υγείας να ανιχνεύσουν για παράδειγμα παρόμοια πρότυπα στην εμφάνιση

συγκεκριμένων συμπτωμάτων. Ένα ενδιαφέρον χαρακτηριστικό είναι ότι κατά την επιλογή μιας εικόνας από το χρήστη τα χαρακτηριστικά της εικόνας εμφανίζονται στο χρήστη και δίνεται η δυνατότητα να εστιάσει σε κάποιο από αυτά για την εύρεση σχετικής πληροφορίας. Για παράδειγμα, εάν ο χρήστης επιλέξει μία εικόνα που έχει επισημειωθεί ότι παρουσιάζει έναν όγκο, τότε εικόνες που έχουν επισημειωθεί ότι παρουσιάζουν υποκλάσεις της κλάσης «όγκος», όπως «όγκος σε θωρακικό τοίχωμα» θα θεωρηθούν σχετικές από το σύστημα.

Μετά από εκτίμηση της αποδοτικότητας του συστήματος μέσω της σύγκρισής του με το σύστημα MedPix, σύστημα το οποίο επιτρέπει την αναζήτηση με βάση επισημειώσεις ελεύθερου κειμένου, φανερώθηκε ότι η οντολογική προσέγγιση βοηθάει στην εξεύρεση της επιθυμητής πληροφορίας.

5

Σχεδίαση Συστήματος Σημασιολογικής

Επισημείωσης και Αναζήτησης Ιατρικών Εικόνων

Στο κεφάλαιο αυτό αναλύονται οι λειτουργικές απαιτήσεις του υλοποιημένου συστήματος, δηλαδή οι απαιτήσεις για τις λειτουργίες που θα πρέπει να εκτελεί, και περιγράφεται η αρχιτεκτονική του.

5.1 Απαιτήσεις Συστήματος

5.1.1 Ομάδες χρηστών

Πριν την ανάπτυξη του συστήματος σημασιολογικής επισημείωσης και αναζήτησης ιατρικών εικόνων, έγινε ανάλυση των πιθανών χρηστών της εφαρμογής, των χαρακτηριστικών και των αναγκών τους. Η κατανόηση του χρήστη και του περιβάλλοντός του είναι απαραίτητη για την ανάπτυξη ενός συστήματος που θα ικανοποιεί τις ανάγκες του χρήστη, ενώ ταυτόχρονα θα είναι φιλικό προς αυτόν. Επίσης, η ανάλυση χρηστών ομαδοποιεί τους χρήστες σε ομάδες χρηστών με όμοια χαρακτηριστικά. Κάθε ομάδα έχει ενδεχομένως διαφορετικές, αν και πολλές φορές επικαλυπτόμενες, ανάγκες από τις υπόλοιπες και για την ευρεία αποδοχή του συστήματος πρέπει να ικανοποιηθούν όλες οι διαφορετικές ανάγκες.

Η ανάλυση χρηστών οδήγησε στις εξής ομάδες πιθανών χρηστών:

- *Γιατροί.* Κυριότεροι υποψήφιοι για χρήστες του συστήματος είναι το ιατρικό και νοσοκομειακό προσωπικό. Ιδιαίτερα χαρακτηριστικά της ομάδας αυτής είναι ο περιορισμένος χρόνος, η ανάγκη για ακρίβεια και αξιοπιστία και η έλλειψη εξοικείωσης με τα τεχνολογικά μέσα. Με γνώμονα λοιπόν τη συγκεκριμένη ομάδα, το σύστημα θα πρέπει να είναι ιδιαίτερα εύχρηστο και αποκρίσιμο προς το χρήστη, χωρίς να είναι κουραστικό, ενώ ταυτόχρονα θα πρέπει να είναι αρκετά αξιόπιστο ώστε να κερδίσει την εμπιστοσύνη των γιατρών και να τους βοηθήσει ουσιαστικά στις διαδικασίες διάγνωσης και θεραπείας.
- *Βοηθητικό προσωπικό.* Οι εργαζόμενοι στην γραμματειακή υποστήριξη των νοσοκομείων, διαγνωστικών κέντρων και ινστιτούτων υγείας είναι αυτοί που στην πράξη μαθαίνουν και χειρίζονται τα πληροφοριακά συστήματα για λογαριασμό των γιατρών ή τουλάχιστον βοηθητικά προς αυτούς. Το σύστημα θα πρέπει να είναι απλό και εύχρηστο έτσι ώστε να μπορεί να χρησιμοποιηθεί και από μη εξειδικευμένο προσωπικό. Επίσης, θα πρέπει να μαθαίνεται εύκολα.
- *Ερευνητές.* Το ερευνητικό προσωπικό από τον τομέα της υγείας και της βιολογίας είναι πιθανοί τελικοί χρήστες. Τα χαρακτηριστικά της ομάδας αυτής είναι μεγάλη εξοικείωση με την τεχνολογία, ανάγκη για συσχέτιση πληροφορίας από διαφορετικές πηγές και ερευνητικό ενδιαφέρον. Η ύπαρξη προχωρημένων επιλογών στην αναζήτηση και η ενσωμάτωση υλικού ενδιαφέροντος από το διαδίκτυο με χρήση εξωτερικών υπερσυνδέσμων είναι δύο παραδείγματα στοιχείων που θα ικανοποιούσαν τις ανάγκες αυτής της πιο απαιτητικής ομάδας.
- *Ακαδημαϊκό προσωπικό.* Τόσο το εκπαιδευτικό προσωπικό των πανεπιστημίων όσο και οι φοιτητές θα μπορούσαν να επωφεληθούν από τη χρήση του συστήματος για εκπαιδευτικούς σκοπούς. Για παράδειγμα, νέοι ακτινολόγοι θα μπορούσαν να εξοικειώνονται με συγκεκριμένες ανωμαλίες που εμφανίζονται συχνά σε ιατρικές εικόνες ώστε να τις αναγνωρίζουν με μεγαλύτερη ευκολία. Μια ευχάριστη και διαδραστική διεπαφή χρήστη αλλά και «έξυπνες» πρωτοβουλίες – όπως εμφάνιση όμοιων εικόνων – σίγουρα θα βοηθούσαν στη χρήση του συστήματος για εκπαιδευτικούς σκοπούς.

Από την παραπάνω ανάλυση προκύπτει ότι το σύστημα, για να ικανοποιεί όλες τις πιθανές ομάδες χρηστών, πρέπει να διαθέτει ευχρηστία, σαφήνεια, αξιοπιστία, υψηλή αποκρισμότητα και πλούσιο περιεχόμενο. Επίσης, είναι σκόπιμο να διαθέτει δύο τρόπους λειτουργίας, για παράδειγμα στην αναζήτηση, την απλή και προχωρημένη λειτουργία. Η απλή λειτουργία θα είναι γρήγορη και εύχρηστη, ενώ η προχωρημένη θα είναι πιο σύνθετη, με περισσότερες επιλογές και θα προϋποθέτει μεγαλύτερη εξοικείωση με το σύστημα. Με τη

δυνατότητα προσθήκης και αφαίρεσης λειτουργιών (παραμετροποίηση λειτουργίας) θα ικανοποιούνται ποικίλες ομάδες χρηστών.

5.1.2 Περιγραφές χρήσης

Τα σενάρια σε συνδυασμό με τα διαγράμματα χρήσης απαρτίζουν της περιγραφές χρήσης και επιτρέπουν να καθορίζουμε με ακρίβεια τις λειτουργικές απαιτήσεις που μας καθοδηγούν στην ανάπτυξη του συστήματος. Στην αρχή παρουσιάζονται κάποια σενάρια σχετικά με το υποσύστημα λογαριασμού χρηστών και στη συνέχεια αναλύονται τα δύο βασικά σενάρια του συστήματος, αυτά της επισημείωσης και αναζήτησης εικόνων.

5.1.2.1 Σενάριο εγγραφής χρήστη στο σύστημα

Το σύστημα θα πρέπει να προσφέρει κάποιον έλεγχο στο ποιος μπορεί να ανεβάσει εικόνες και να τις επισημειώσει. Για το λόγο αυτό, δημιουργούνται λογαριασμοί εγγεγραμμένων χρηστών οι οποίοι αποκτούν τις παραπάνω δυνατότητες. Το κείμενο του σεναρίου εγγραφής είναι το εξής:

1. Ο χρήστης αποστέλλει τα στοιχεία του, όπως όνομα, επίθετο, ηλεκτρονική διεύθυνση, ειδικότητα και ενδεχομένως επιλέγει κάποια διαπιστευτήρια (κωδικό πρόσβασης).
2. Το σύστημα ελέγχει αν μπορεί να γίνει η εγγραφή νέου χρήστη, για παράδειγμα ελέγχοντας αν υπάρχει ήδη χρήστης με το ίδιο όνομα στη βάση.
3. Το σύστημα δημιουργεί το νέο λογαριασμό, ενημερώνει το χρήστη και τον προτρέπει να προχωρήσει στην είσοδό του (sign in).

5.1.2.2 Σενάριο εισόδου χρήστη στο σύστημα

Για να μπορέσει κάποιος χρήστης να χρησιμοποιήσει το εργαλείο επισημείωσης εικόνων θα πρέπει πρώτα να πιστοποιηθεί από το σύστημα μέσω της σύνδεσής του με αυτό (sign in). Το κείμενο σεναρίου είναι:

1. Ο χρήστης αποστέλλει το όνομά του και ενδεχομένως κάποιο διαπιστευτήριο, όπως κωδικός πρόσβασης.
2. Το σύστημα ελέγχει αν υπάρχει εγγεγραμμένος χρήστης με το συγκεκριμένο όνομα και κωδικό.
3. Αν γίνει η ταυτοποίηση και βρεθεί χρήστης τότε τον συνδέει στο σύστημα και τον ενημερώνει κατάλληλα.
4. Αν δεν βρεθεί η αντίστοιχη εγγραφή προτρέπει το χρήστη να ολοκληρώσει πρώτα την εγγραφή του στο σύστημα (σενάριο 4.1.2.1).

Σημειώνεται ότι ανά πάσα στιγμή, δηλαδή στις περισσότερες αν όχι σε όλες, οθόνες της εφαρμογής ο χρήστης θα μπορεί να βλέπει αν είναι συνδεδεμένος ή όχι.

5.1.2.3 Σενάριο εξόδου χρήστη από το σύστημα

Το απλό αυτό σενάριο υλοποιεί τη λειτουργικότητα της αποσύνδεσης του χρήστη από το σύστημα (sign out). Κείμενο σεναρίου:

1. Ο χρήστης ενημερώνει το σύστημα ότι θέλει να αποσυνδεθεί. Προτιμάται αυτό να γίνεται με κάποιο κουμπί που βρίσκεται σε εμφανή θέση και είναι διαθέσιμο σε όλες τις οθόνες της εφαρμογής.
2. Το σύστημα αποσυνδέει το χρήστη.

5.1.2.4 Σενάριο επισημείωσης εικόνας

Μία από τις βασικές λειτουργικότητες του συστήματος είναι η επισημείωση εικόνας. Αυτή είναι μια σύνθετη διαδικασία που απαρτίζεται σε επιμέρους διαδικασίες και απαιτεί τη συνεργασία πολλών υποσυστημάτων για την πραγματοποίησή της. Μοναδική προϋπόθεση για την πραγματοποίηση του σεναρίου είναι ο χρήστης να έχει συνδεθεί στο σύστημα. Αυτή είναι μια ασφαλιστική «δικλείδα» που στόχο έχει να εξασφαλίσει την ποιότητα των επισημειώσεων καθώς αυτές είναι επώνυμες και όχι ανώνυμες. Ακολουθεί το κείμενο του σεναρίου και το αντίστοιχο διάγραμμα χρήσης:

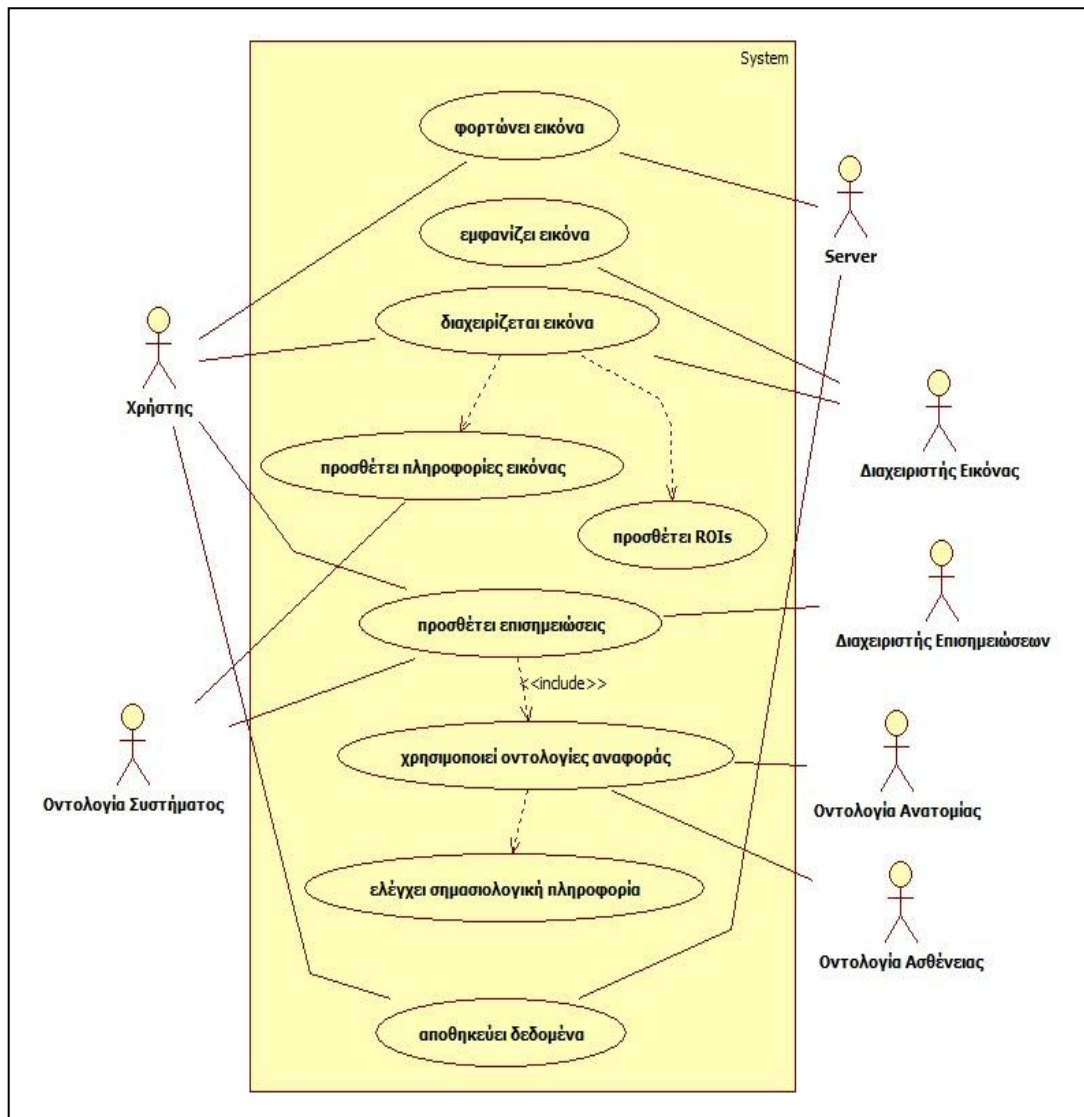
1. Ο χρήστης ανεβάζει/φορτώνει (upload) μια καινούργια εικόνα από τον υπολογιστή του στον υπολογιστή που φιλοξενεί την εφαρμογή (server).
2. Σε περίπτωση επιτυχίας το σύστημα ενημερώνει το χρήστη και του επιτρέπει να προχωρήσει στην επισημείωση της φορτωμένης εικόνας. Αλλιώς, προτρέπει το χρήστη να ξαναπροσπαθήσει.
3. Κατάλληλο πρόγραμμα εμφανίζει την εικόνα στον φυλλομετρητή ιστού (browser).
4. Το σύστημα προτρέπει τον χρήστη να συμπληρώσει πληροφορίες σχετικές με την εικόνα, όπως τροπικότητα, όνομα ινστιτούτου στο οποίο τραβήχτηκε η εικόνα, μέρος του σώματος που απεικονίζεται, και άλλα.
5. Το σύστημα προτρέπει τον χρήστη να χρησιμοποιήσει κατά βούληση φίλτρα επεξεργασίας της εικόνας, όπως βελτίωση φωτεινότητας, όξυνση γραμμών κ.τ.λ..
6. Το σύστημα προτρέπει το χρήστη να επιλέξει τις Περιοχές Ενδιαφέροντος (Regions of Interest ή ROIs) στην εικόνα με χρήση κατάλληλων επιλογέων σχήματος ορθογωνίου, έλλειψης και άλλων.
7. Ο χρήστης εφαρμόζει τα επιθυμητά φίλτρα και σημειώνει τα ROIs στην εικόνα.
8. Το σύστημα προτρέπει το χρήστη να επιλέξει την ανατομική έννοια που καλύτερα περιγράφει το κάθε ROI.

9. Ο χρήστης επιλέγει το ανατομικό χαρακτηριστικό από λίστα που ενημερώνεται αυτόματα από έννοιες της οντολογίας ανατομίας.
10. Το σύστημα προτρέπει το χρήστη να επιλέξει την έννοια της ασθένειας που παρατηρείται σε κάθε ROI.
11. Ο χρήστης επιλέγει την ασθένεια, ανωμαλία ή δυσλειτουργία από λίστα που ενημερώνεται αυτόματα από έννοιες της οντολογίας ασθένειας.
12. Το σύστημα παρουσιάζει το χρήστη συγκεντρωτικά τις επιλογές του και ζητά επιβεβαίωση ώστε αυτές να αποθηκευτούν.
13. Ο χρήστης επιβεβαιώνει τις επιλογές του. Αλλιώς, επιστρέφει στο βήμα 6 και επαναπροσδιορίζει τις επισημειώσεις του.
14. Το σύστημα αποθηκεύει την εικόνα με τα ROIs και τις επισημειώσεις στον server και παρουσιάζει στο χρήστη την επισημειωμένη εικόνα.

Παρατηρήσεις:

- Το σύστημα θα πρέπει να προτρέπει το χρήστη να ανεβάζει εικόνες της μορφής DICOM – χωρίς αυτό να είναι δεσμευτικό. Σε τέτοια περίπτωση, μαζί με την εμφάνιση της εικόνας (βλ. βήμα 3) θα πρέπει να εμφανίζονται τα δεδομένα που έχουν εξαχθεί από την επικεφαλίδα του αρχείου DICOM, ώστε να βοηθούν το χρήστη στη συμπλήρωση των πληροφοριών της εικόνας (βλ. βήμα 4).
- Σε περίπτωση που ο χρήστης επιλέξει να μη χρησιμοποιήσει τη λίστα των εννοιών της οντολογίας αναφοράς (βλ. βήματα 9 και 11) και προχωρήσει σε επισημείωση ελεύθερου κειμένου, το σύστημα θα πρέπει να ελέγχει το κείμενο που έχει εισάγει ο χρήστης. Αν το κείμενο δεν είναι όρος της εκάστοτε οντολογίας το σύστημα θα πρέπει να ενημερώνει σχετικά τον χρήστη. Με τον τρόπο αυτό το σύστημα αποτρέπει τον χρήστη από την επισημείωση σε μορφή ελεύθερου κειμένου αλλά δεν του το απαγορεύει, δίνοντάς του τη μεγαλύτερη δυνατή ελευθερία και ευελιξία.
- Κατά την επιλογή των όρων από τις οντολογίες (βλ. βήματα 9 και 11) το σύστημα ιδανικά θα πρέπει να παρέχει κατάλληλες οπτικοποιήσεις, όπως για παράδειγμα δέντρα οντολογιών, αλλά και επιπλέον πληροφορίες για τους όρους της επιλογής του, όπως εξωτερικές συνδέσεις προς ιστοτόπους με περιγραφή των όρων, ώστε ο χρήστης να μπορεί να πλοηγηθεί στη «σημασιολογική γειτονιά» των εννοιών που έχει επιλέξει και να πραγματοποιεί εξειδικευμένες και επακριβείς επισημειώσεις, αξιοποιώντας στο έπακρο το πλεονέκτημα της χρήσης σημασιολογικών μεθόδων.

Στην Εικόνα 5.1.1 παρουσιάζεται το διάγραμμα χρήσης (use case diagram) στο οποίο αποτυπώνονται οι ενέργειες που είναι αναγκαίες για την πραγματοποίηση της επισημείωσης εικόνας.



Εικόνα 5.1.1: Διάγραμμα Χρήσης της επισημείωσης εικόνας.

Στην Εικόνα 5.1.1 παρατηρούμε τους κυριότερους δράστες και τις κυριότερες περιπτώσεις χρήσης που εμπλέκονται στο σενάριο επισημείωσης εικόνας. Οι ενέργειες γίνονται κατά σειρά από πάνω προς τα κάτω, σύμφωνα με την σύμβαση των use case διαγραμμάτων: Ο *Χρήστης* φορτώνει μία εικόνα στον *Server*, στη συνέχεια με τη βοήθεια του *Διαχειριστή Εικόνας* τη διαχειρίζεται αφού πρώτα ο *Διαχειριστής Εικόνας* την εμφανίσει. Η διαχείριση της εικόνας περιλαμβάνει την προσθήκη πληροφορίας εικόνας σύμφωνα με τα πεδία που υπάρχουν στην *Οντολογία Συστήματος* και την προσθήκη των ROIs. Στη συνέχεια ο *Χρήστης* με τη βοήθεια του *Διαχειριστή Επισημειώσεων* προσθέτει επισημειώσεις χρησιμοποιώντας τις οντολογίες αναφοράς, δηλαδή την *Οντολογία Ανατομίας* και την *Οντολογία Ασθένειας*.

Ταυτόχρονα γίνεται έλεγχος της σημασιολογικής πληροφορίας των επισημειώσεων. Τέλος ο *Χρήστης* αποθηκεύει τα δεδομένα στον *Server*.

5.1.2.5 Σενάριο αναζήτησης εικόνας

Η δεύτερη βασική λειτουργικότητα της εφαρμογής είναι η αναζήτηση εικόνων. Δεν υπάρχει καμία προϋπόθεση για την χρήση της συγκεκριμένης λειτουργικότητας, που σημαίνει ότι ο χρήστης πρέπει απλά να έχει πρόσβαση στην εφαρμογή μέσω ενός φυλλομετρητή ιστού (browser), χωρίς να χρειάζεται να εισαχθεί στο σύστημα και να δώσει τα διαπιστευτήριά του. Αυτό γίνεται καθώς κατά την αναζήτηση εικόνων ο χρήστης απλά πλοηγείται στα δεδομένα του συστήματος, αλλά δεν μπορεί να τα τροποποιήσει, ούτε μπορεί να εισάγει νέα δεδομένα. Το κείμενο του σεναρίου ακολουθεί:

1. Ο χρήστης εισάγει τα δεδομένα απλής αναζήτησης, όπως το ανατομικό χαρακτηριστικό και την ασθένεια που αναζητεί. Προτιμάται αυτό να γίνεται με τη βοήθεια λιστών που ενημερώνονται καθώς ο χρήστης πληκτρολογεί με βάση τις οντολογίες αναφοράς.
2. Το σύστημα φροντίζει να παρουσιάσει στο χρήστη τις υπόλοιπες δυνατότητες της σύνθετης αναζήτησης, όπως αναζήτηση με βάση την ηλικία, το φύλο του ασθενούς, το ιστοιούτο που τραβήχτηκε η εικόνα, την τροπικότητα της εικόνας, το όνομα του χρήστη που την επισήμανε και άλλα.
3. Ο χρήστης επιλέγει όλα τα επιθυμητά κριτήρια αναζήτησης και πατάει το κουμπί της αναζήτησης.
4. Το σύστημα επιστρέφει στο χρήστη τις εικόνες που βρέθηκαν να ταιριάζουν στα δεδομένα αναζήτησης σε μία λίστα.
5. Το σύστημα εμφανίζει, εκτός από μια προεπισκόπηση των εικόνων της λίστας, συνοπτικά τις πληροφορίες κάθε εικόνας, όπως αυτές έχουν αποθηκευτεί στον server.
6. Ο χρήστης επιλέγει μία εικόνα από τη λίστα.
7. Το σύστημα εμφανίζει την εικόνα σε κανονικό πλέον μέγεθος σε καινούργια οθόνη. Οι Περιοχές Ενδιαφέροντος (ROIs) εμφανίζονται στην εικόνα.
8. Το σύστημα εμφανίζει όλες τις πληροφορίες της εικόνας και τις επισημειώσεις που έχουν γίνει στην εικόνα με απλό και κατατοπιστικό τρόπο.
9. Το σύστημα επιτρέπει στο χρήστη να εξάγει τις σημασιολογική πληροφορία (RDF τριάδες) της εικόνας σε μορφή RDF/XML.

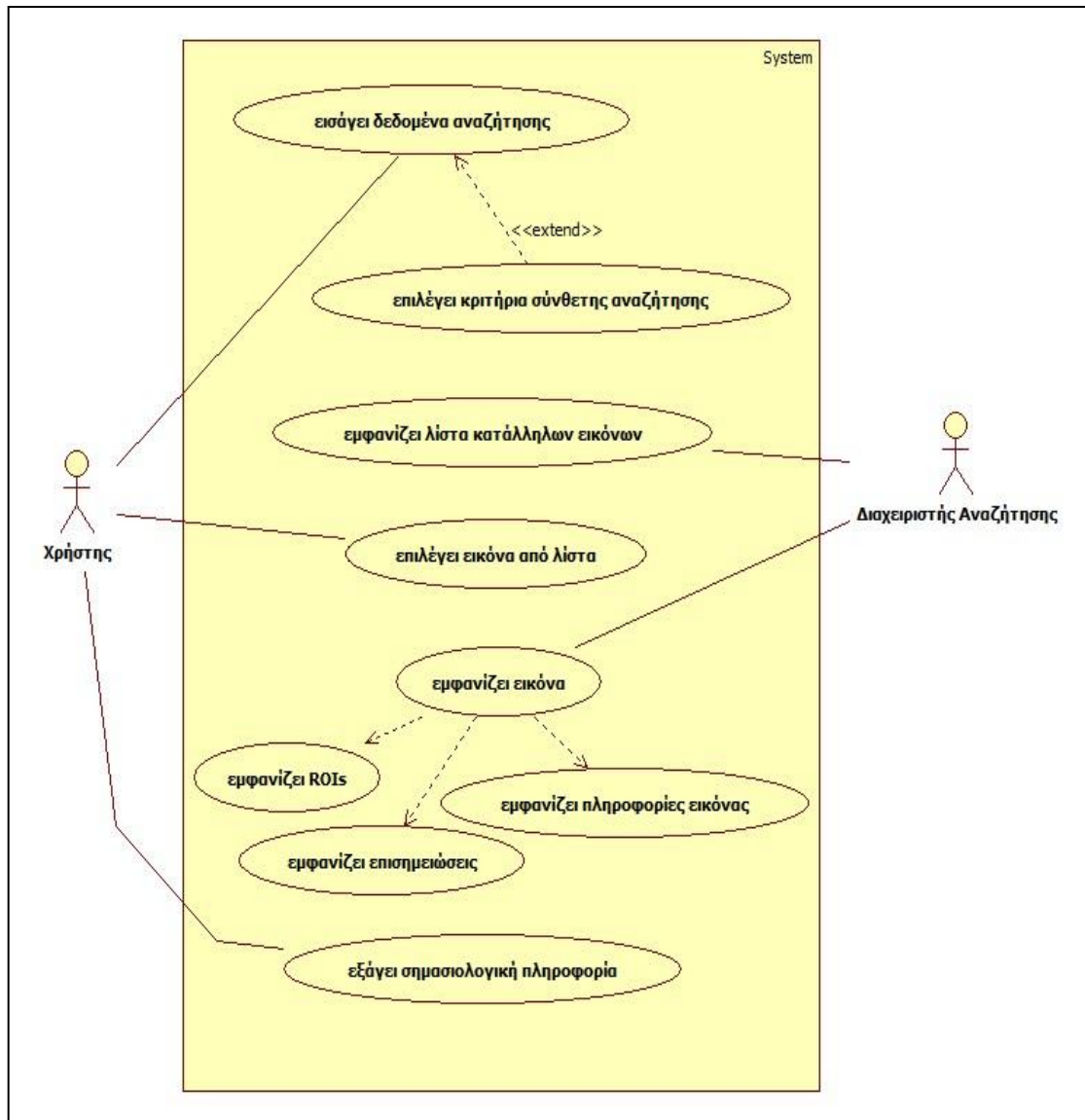
Παρατηρήσεις:

- Οι επιλογές σύνθετης αναζήτησης (βλ. βήμα 2) θα πρέπει να είναι εύκολα προσβάσιμες στον χρήστη, χωρίς όμως να τον εμποδίζουν να πραγματοποιεί γρήγορη

και απλή αναζήτηση μόνο με βάση τα χαρακτηριστικά ανατομίας και ασθένειας, εάν αυτό επιθυμεί.

- Το σύστημα θα πρέπει ενδεχομένως να δίνει τη δυνατότητα στο χρήστη να παραμετροποιεί τη μέθοδο σημασιολογικής αναζήτησης, για παράδειγμα αναζήτηση σε βάθος μίας κλάσης για τους όρους της ανατομίας ή ασθένειας ή αναζήτηση με βάση τη σχέση «οργανικό μέρος» της οντολογίας ανατομίας. Αυτό βέβαια προϋποθέτει αρκετή εξοικείωση του χρήστη με τις οντολογίες και πρέπει να γίνεται με προσοχή, καθώς μπορεί να καταστήσει το σύστημα δύσχρηστο σε αρχάριους χρήστες. Για αυτό υπάρχει πάντα ένας προεπιλεγμένος τρόπος σημασιολογικής επαύξησης των ερωτημάτων.
- Σε περίπτωση αποτυχίας εύρεσης εικόνων (βλ. βήμα 4), το σύστημα ιδανικά θα πρέπει να προχωρά σε σημασιολογικές «συστάσεις» του τύπου: «Μήπως ενδιαφέρεστε να αναζητήσετε εικόνες που παρουσιάζουν ένα συγγενές ανατομικό χαρακτηριστικό;».
- Κατά την εμφάνιση μιας εικόνας (βλ. βήμα 7) κρίνεται σκόπιμο το σύστημα να παρουσιάζει μια συνοπτική λίστα από *παρόμοιες* εικόνες, με βάση τα σημασιολογικά χαρακτηριστικά των επισημειώσεων. Επίσης, μπορεί να παρέχει εξωτερικούς συνδέσμους προς ιστοσελίδες με περιγραφή των εννοιών. Με τον τρόπο αυτό υπογραμμίζεται το πλεονέκτημα της σημασιολογικής προσέγγισης, που επιτρέπει το διαχωρισμό μεταξύ *μορφής* (η επισημείωση πάνω σε μια συγκεκριμένη εικόνα) και *περιεχομένου* (η έννοια της επισημείωσης είναι ανεξάρτητη από την εικόνα).

Στην Εικόνα 5.1.2 παρατηρούμε τους κυριότερους δράστες και τις κυριότερες περιπτώσεις χρήσης που εμπλέκονται στο σενάριο αναζήτησης εικόνας: ο *Χρήστης* εισάγει τα δεδομένα αναζήτησης και προαιρετικά επιλέγει τα κριτήρια της σύνθετης αναζήτησης. Ο *Διαχειριστής Αναζήτησης* εμφανίζει τη λίστα με τις ευρεθείσες εικόνες και ο *Χρήστης* επιλέγει μία από αυτές. Ο *Διαχειριστής Αναζήτησης* τότε εμφανίζει την εικόνα, μια ενέργεια που περιλαμβάνει επίσης την εμφάνιση των ROIs της εικόνας, των επισημειώσεων και των πληροφοριών της εικόνας. Τέλος, ο *Χρήστης* επιλέγει να εξάγει τη σημασιολογική πληροφορία σχετική με την εικόνα, αν το επιθυμεί.



Εικόνα 5.1.2: Διάγραμμα Χρήσης της αναζήτησης εικόνας.

5.1.3 Οντολογίες Αναφοράς

5.1.3.1 Οντολογία Ανατομίας

Η οντολογία ανατομίας απαιτείται κυρίως να καλύπτει ευρύ και γενικευμένο φάσμα ανατομικής γνώσης. Είναι απαραίτητο να εμπεριέχει κατηγοριοποιημένη και δομημένη σωστά όλη τη γνώση της ανατομίας του ανθρώπινου σώματος. Πρέπει, δηλαδή, οι οντότητες να περιγράφουν τα διάφορα μέρη του ανθρώπινου σώματος, όπως ζωτικά όργανα (καρδιά, πνεύμονας κ.ά.), ανθρώπινα μέλη, επιμέρους ιστούς, διαφόρων ειδών κύτταρα, ακόμη και δευτερεύουσας σημασίας χαρακτηριστικά. Ακόμη είναι σημαντικό να περιγράφονται οι σχέσεις μεταξύ διαφόρων μερών του σώματος. Τέτοιες σχέσεις θα είναι σχέσεις

- γειννίασης (π.χ. «πρόσθιο στήθος» είναι χωρικό μέρος του «θώρακα»)
- μέρος – όλου (π.χ. «διάφραγμα» είναι συστατικό μέρος του «θώρακα»)
- διασύνδεσης
- λειτουργίας
- κ.ά.

5.1.3.2 Οντολογία Ασθενειών

Η οντολογία ασθενειών πρέπει να καλύπτει όλο το φάσμα των δυσλειτουργιών, ασθενειών και αιτιών θανάτου. Είναι απαραίτητο η κατηγοριοποίηση των ασθενειών να έχει γίνει σε μορφή κλάσεων – υποκλάσεων, ώστε να επιτρέπεται η εξειδίκευση ή η γενίκευση του όρου ασθένειας ανάλογα με το ποσοστό βεβαιότητας του επιστήμονα-παρατηρητή της εικόνας. Επίσης, επιθυμητό είναι στην οντολογία να περιγράφονται σχέσεις μεταξύ των ασθενειών όπως σχέσεις:

- αλληλεξάρτησης
- αμοιβαίου αποκλεισμού
- πιθανής συσχέτισης
- κ.ά.

Ιδανικό θα ήταν επίσης να υπάρχει μία χαρτογράφηση των ασθενειών σε σχέση με το πού μπορούν να εμφανιστούν στο ανθρώπινο σώμα. Αυτή είναι ουσιαστικά μία συσχέτιση μεταξύ των δύο οντολογιών αναφοράς που θα διευκόλυνε το έργο του επιστήμονα κατά τη διάρκεια της επισημείωσης.

5.2 Αρχιτεκτονική

5.2.1 MVC

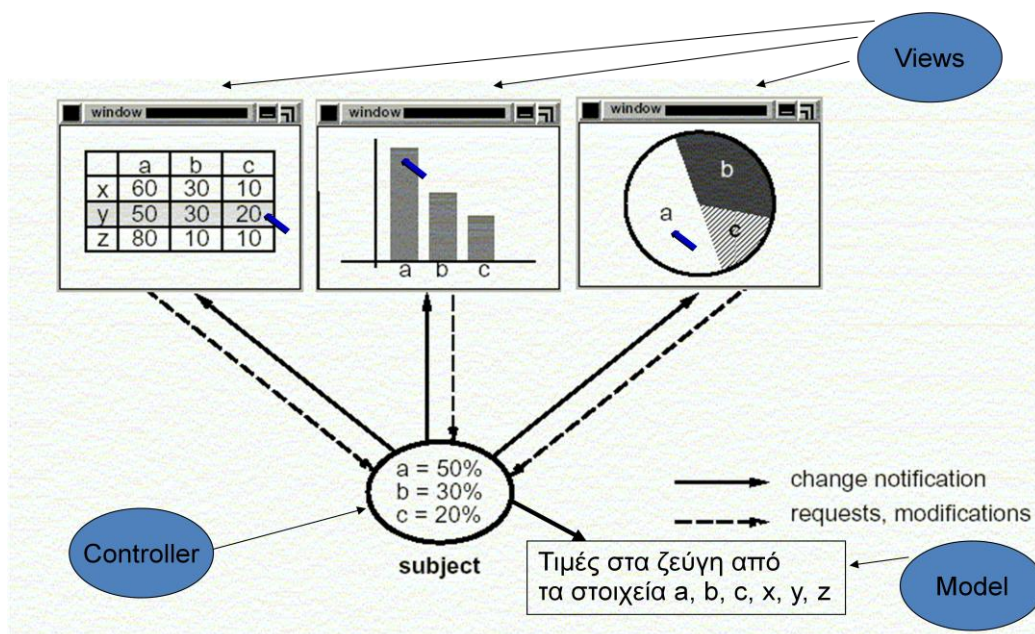
Για να είναι εύκολα προσβάσιμη η εφαρμογή κρίθηκε απαραίτητο να είναι διαθέσιμη μέσω φυλλομετρητών Ιστού (browsers). Έτσι, για το σχεδιασμό του συστήματος κρίθηκε ιδανική η αρχιτεκτονική τεχνοτροπία *Μοντέλο-Παρουσίαση-Έλεγχος* (Model-View-Controller ή MVC). Η τεχνοτροπία αυτή είναι ευρέως διαδεδομένη για ανάπτυξη εφαρμογών Ιστού, καθώς βασίζεται στη διάσπαση ενός συστήματος σε τρεις μονάδες:

- Τη μονάδα *Μοντέλο*, που περικλείει και μοντελοποιεί τις βασικές λειτουργίες και τα δεδομένα του συστήματος. Συνήθως είναι μία βάση δεδομένων, όπου είναι αποθηκευμένα τα δεδομένα της εφαρμογής, αλλά μπορεί να είναι και οποιοδήποτε άλλο σύστημα δομημένης αποθήκευσης δεδομένων, όπως XML έγγραφα ή μία αποθήκη RDF τριάδων (RDF triple store).

- Μια ή περισσότερες μονάδες *Παρουσίασης*, που ορίζουν διαφορετικούς τρόπους παρουσίασης δεδομένων και αποτελεσμάτων στο χρήστη. Οι μονάδες παρουσίασης ή αλλιώς *όψεις*, στην περίπτωση μίας web εφαρμογής, είναι κάθε html σελίδα που παρέχεται στο χρήστη. Αυτή διαμορφώνεται κάθε φορά με τη βοήθεια των ελεγκτών και αντλεί πληροφορίες απευθείας από τα μοντέλα που έχουν δημιουργηθεί, δίνοντας πάντα μια έγκυρη εικόνα των δεδομένων. Οι όψεις αποτελούν ουσιαστικά τη διεπαφή του συστήματος με το χρήστη. Για το λόγο αυτό σχεδιάζονται με προσοχή ώστε να παρέχουν την καλύτερη δυνατή εμπειρία χρήσης.
- Μια ή περισσότερες μονάδες *Ελέγχου*, ή αλλιώς ελεγκτές, που διαχειρίζονται με τον κατάλληλο τρόπο την εισαγωγή πληροφοριών και αιτημάτων από τον χρήστη. Στο κομμάτι των ελεγκτών (controllers), ουσιαστικά, ενσωματώνονται όλες οι λογικές λειτουργίες του συστήματος και αποτελούν το δομικό κρίκο που ενώνει τα μοντέλα με τις όψεις.

Με τον τρόπο αυτό, τα δομικά στοιχεία της εφαρμογής ομαδοποιούνται με βάση τη λογική λειτουργία που επιτελούν, κάτι που έχει σαν αποτέλεσμα έναν συγκροτημένο, καθαρό και κυρίως εύκολα συντηρήσιμο κώδικα.

Στην Εικόνα 5.2.1 δίνεται σχηματικά ένα παράδειγμα της MVC αρχιτεκτονικής τεχνοτροπίας, προς αποσαφήνιση των παραπάνω.



Εικόνα 5.2.1 : Παράδειγμα αρχιτεκτονικής τεχνοτροπίας MVC.

5.2.2 SOA και Οντολογίες Αναφοράς

Κρίνεται απαραίτητο οι οντολογίες αναφοράς να είναι διαθέσιμες μέσω υπηρεσιών Ιστού (web services) για να είναι προσβάσιμες σε πολλές εφαρμογές και χρήστες ταυτόχρονα, ενισχύοντας έτσι τη διαλειτουργικότητα και την επαναχρησιμοποίηση, έννοιες πρωταρχικές για το Όραμα του Σημασιολογικού Ιστού [33].

Παρατίθενται οι πιο βασικοί λόγοι:

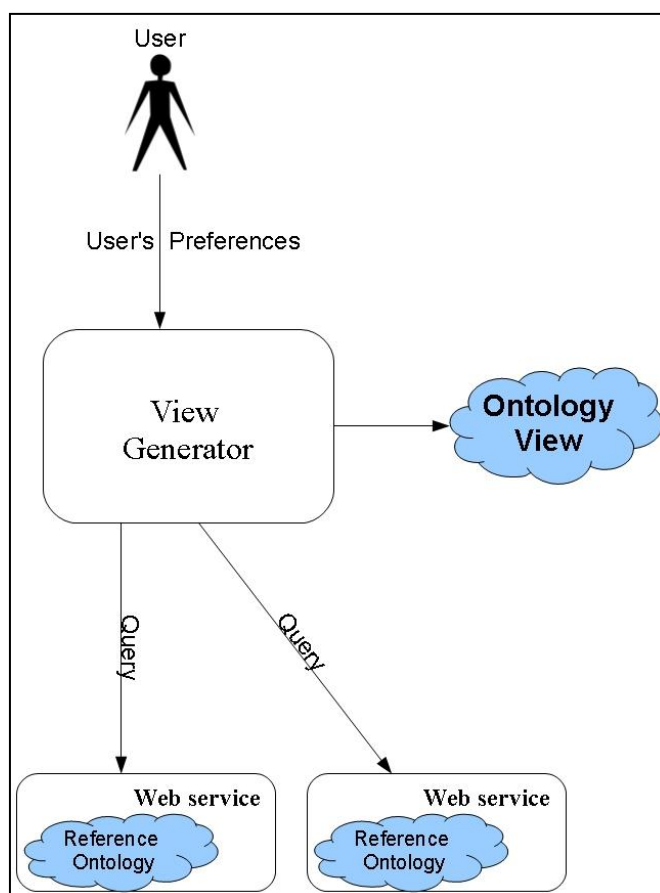
- Πρόσβαση σε οντολογίες μέσω ερωτήσεων (queries).
- Παραγωγή οντολογικών όψεων (ontology views).
- Μετάφραση των οντολογιών από μία γλώσσα σε άλλη.
- Διαχείριση πολλαπλών οντολογιών.
- Συλλογιστική σε οντολογίες.

Η λογική αυτή της διαθεσιμότητας δεδομένων και υπηρεσιών, μέσω υπηρεσιών Ιστού, είναι αρκετά διαδεδομένη στις μέρες μας και αποτελεί το κύριο χαρακτηριστικό της Προσανατολισμένης σε Υπηρεσίες Αρχιτεκτονικής τεχνοτροπίας (Service Oriented Architecture ή SOA). Στα πλαίσια αυτής της τεχνοτροπίας, εφαρμογές αναπτύσσονται, χρησιμοποιώντας ένα συνονθύλευμα υπηρεσιών Ιστού οι οποίες εκτελούν επιμέρους λειτουργίες της εφαρμογής ή διαθέτουν δεδομένα σε αυτή.

Στην περίπτωση μας, σκοπός είναι να παίρνουμε τις οντολογικές όψεις που χρειαζόμαστε από τις οντολογίες αναφοράς, ανατομίας και ασθενειών, μέσω των επιμέρους υπηρεσιών. Ως οντολογική όψη ορίζεται ένα κομμάτι της οντολογίας αναφοράς, το οποίο μπορεί να είναι όλη η οντολογία αναφοράς, μία κλάση της ή συνηθέστερα μία ομάδα κλάσεων με όλες τις σχέσεις που τις συνδέουν. Η οντολογική όψη κάθε φορά καθορίζεται από το χρήστη και αποτυπώνει την υποπεριοχή ενδιαφέροντός του. Έτσι απαραίτητη είναι η ανάπτυξη μιας υπηρεσίας του συστήματος (Ontology View Generator), που να λαμβάνει παραμετροποιημένη την υποπεριοχή ενδιαφέροντος του χρήστη και να δημιουργεί τη συνδυαστική οντολογική όψη από τις οντολογίες αναφοράς [33]. Έχοντας διαθέσιμες τις οντολογίες αναφοράς (reference ontologies) μέσω υπηρεσιών Ιστού (web services), ο Ontology View Generator επικοινωνεί με αυτές τις υπηρεσίες και, έχοντας ως παράμετρο την υποπεριοχή ενδιαφέροντος του χρήστη, δημιουργεί την *οντολογική όψη* (ontology view), όπως φαίνεται και στην Εικόνα 5.2.2. Μετέπειτα η όψη αυτή, η οποία και από μόνη της μπορεί να αποτελεί οντολογία, ενθυλακώνεται στην οντολογία εφαρμογής (application ontology), επεκτείνοντάς την ή αφήνοντάς την ανέπαφη, στην περίπτωση που προϋπάρχει η εν λόγω όψη.

Η ανάκτηση των επιμέρους οντολογικών όψεων πρέπει να γίνεται μέσω ερωτημάτων (queries) στις οντολογίες αναφοράς. Τα ερωτήματα, στην ουσία, θα τίθεται στην υπηρεσία

διαχείρισης της οντολογίας και αυτή θα αναλαμβάνει να επερωτήσει την οντολογία και να μας επιστρέψει τα αποτελέσματα στη μορφή που τα ζητήσαμε.



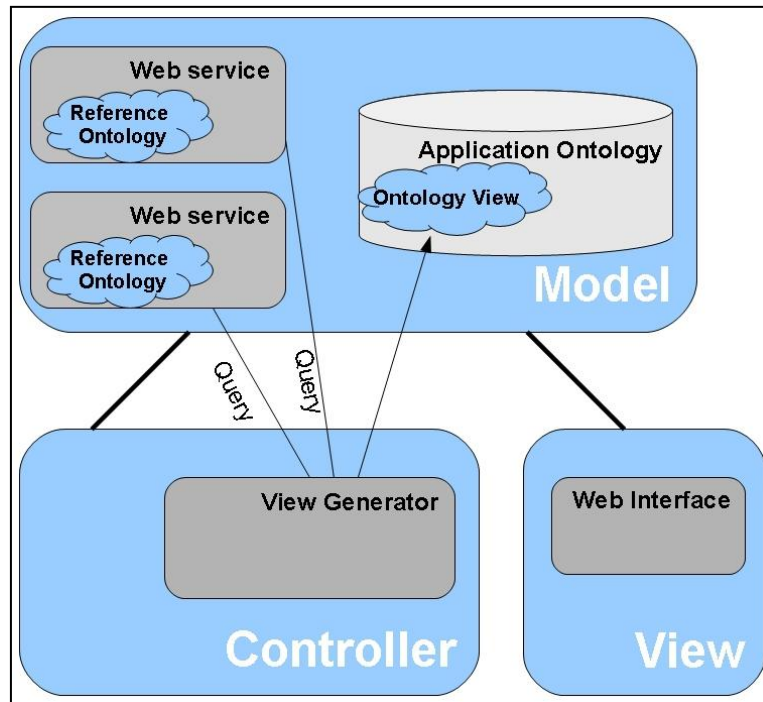
Εικόνα 5.2.2: Παραγωγή Οντολογικής Όψης

5.2.3 Υβριδικό Αρχιτεκτονικό Πλαίσιο

Οι ανάγκες, έτσι, της εφαρμογής επιβάλλουν τη χρήση ενός Υβριδικού Αρχιτεκτονικού Πλαισίου, που είναι αποτέλεσμα «ανακατέματος» (mash up) των δύο αρχιτεκτονικών τεχνοτροπιών, οι οποίες παρουσιάστηκαν στις παραγράφους 5.2.1 και 5.2.2. Συγκεκριμένα οι ανάγκες επιβάλλονται από τη δομική μονάδα του *Μοντέλου*, καθώς αυτό αποτελείται από *μοντέλο αναφοράς* και *μοντέλο εφαρμογής*, με το μοντέλο εφαρμογής να καθορίζεται και επεκτείνεται κάθε φορά από το μοντέλο αναφοράς.

Το *μοντέλο αναφοράς* αποτελείται από τις οντολογίες αναφοράς, ανατομίας και ασθενειών. Ενώ το *μοντέλο εφαρμογής* αποτελείται από την οντολογία εφαρμογής. Η οντολογία εφαρμογής επιβάλλεται να εμπεριέχει μια συνδυαστική *οντολογική όψη (ontology view)* των δύο οντολογιών αναφοράς, να εμπεριέχει δηλαδή μερική γνώση των οντολογιών αναφοράς

[9]. Στην Εικόνα 5.2.3 παρατίθεται απεικονισμένο το προαναφερθέν αρχιτεκτονικό πλαίσιο για την καλύτερη κατανόηση του.



Εικόνα 5.2.3: Υβριδικό Αρχιτεκτονικό Πλαίσιο

6

Σύστημα Σημασιολογικής Επισημείωσης και Αναζήτησης Ιατρικών Εικόνων

Στο κεφάλαιο αυτό αναλύονται κάποια θέματα σχετικά με την υλοποίηση του συστήματός μας που έχουν ιδιαίτερο ενδιαφέρον, παρουσιάζονται οι πλατφόρμες και τα προγραμματιστικά εργαλεία που χρησιμοποιήθηκαν, ενώ επίσης γίνεται έλεγχος με βάση κάποια σενάρια λειτουργίας.

6.1 Λεπτομέρειες Υλοποίησης

Στην ενότητα αυτή παρουσιάζονται λεπτομερώς συγκεκριμένα κομμάτια της υλοποίησης που παρουσιάζουν ενδιαφέρον. Συγκεκριμένα, αυτά είναι οι Οντολογίες αναφοράς που επιλέχθηκαν, η Οντολογία εφαρμογής, η Διάταξη Συστήματος και Διαδικτυακές Υπηρεσίες, ο Καθορισμός Σημασιολογικής Γειτονιάς καθώς και η Διεπιφάνεια χρήστη.

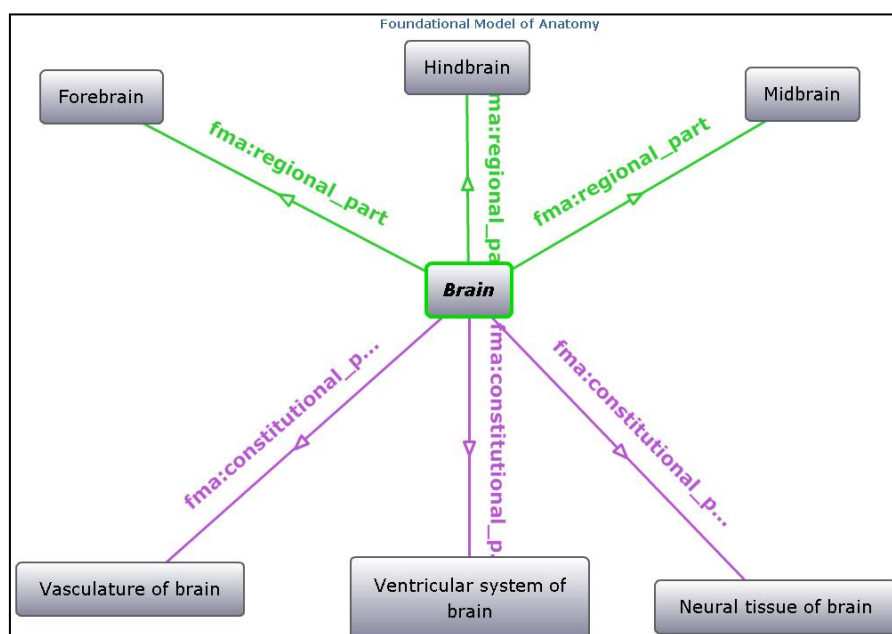
6.1.1 Οντολογίες αναφοράς

Για την υλοποίηση του συστήματος επιλέχθηκαν ως οντολογίες αναφοράς η FMA, για την αναπαράσταση της ανατομίας, και η ICD-10, για την αναπαράσταση των ασθενειών.

6.1.1.1 FMA

Η FMA (Foundational Model of Anatomy) [34], όπως παρουσιάστηκε θεωρητικά και στην παράγραφο 3.2.1, είναι μία από τις πιο συστηματικά δομημένες οντολογίες αναφοράς στο τομέα βιοϊατρικής. Επί του παρόντος, η έκδοση 3.1 είναι διαθέσιμη σε μορφή OWL. Η ανατομικές έννοιες και οι σχέσεις που καλύπτονται από την FMA μπορούν να γενικευθούν σε όλους τους τομείς της βιοϊατρικής, αυτός είναι και ο λόγος που χαρακτηρίζεται ως «Θεμελιώδης» (Foundational). Ο όρος «ανατομική έννοια» εδώ σημαίνει μια μονάδα γνώσης που αναφέρεται σε μια ανατομική οντότητα. Η FMA περιέχει σήμερα πάνω από 80.000 διαφορετικές ανατομικές έννοιες, που αναπαριστούν δομές που κυμαίνονται σε μέγεθος από μακρομοριακά συστήματα και τμήματα των κυττάρων έως τα περισσότερα μέρη του σώματος. Οι έννοιες αυτές συνδέονται με πάνω από 110.000 όρους και σχετίζονται μεταξύ τους με 176 είδη των σχέσεων. Η FMA θεωρείται σήμερα το μεγαλύτερο και πληρέστερο οντολογία αναφοράς στον τομέα της βιοϊατρικής.

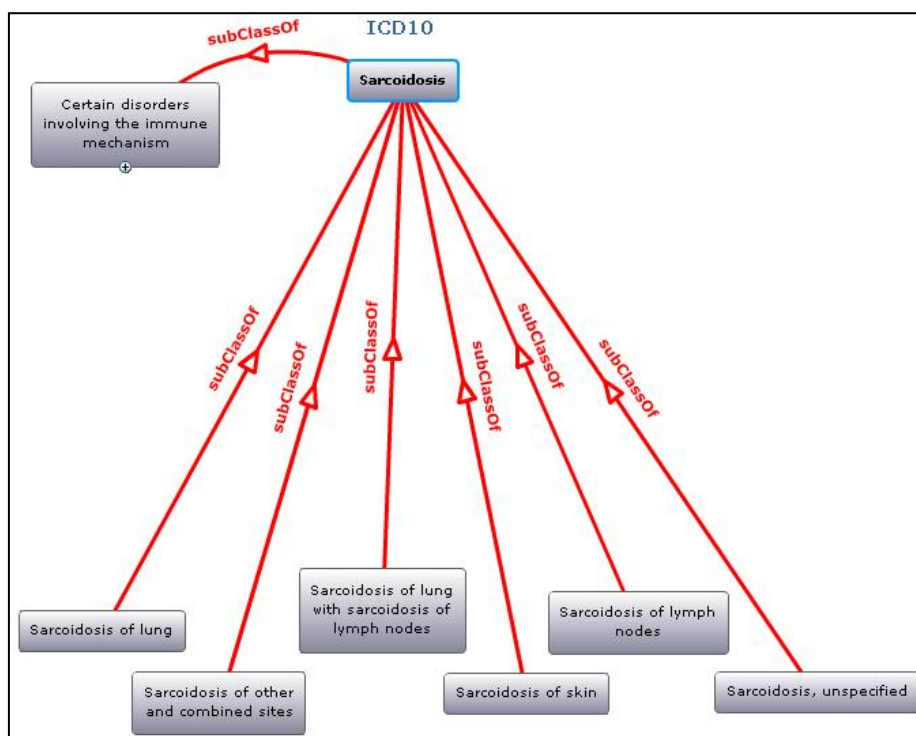
Για τις ανάγκες του έργου αυτού, το ενδιαφέρον επικεντρώνεται στην ιεραρχία (υποκλάσεις κυρίως), και τις τρεις σημαντικότερες σχέσεις: regional part (χωρικό μέρος), constitutional part (συστατικό μέρος) και systemic part (μέρος συστήματος). Παίρνοντας ως παράδειγμα την ανατομική οντότητα «Brain» (Εγκέφαλος), όπως παρουσιάζεται στην Εικόνα 6.1.1 τα χωρικά τμήματά της είναι «Forebrain», «Hindbrain», «Midbrain» και τα συστατικά της τμήματα είναι «Vasculature of brain», «Ventricular system of brain», «Neural tissue of brain». Δεν υπάρχουν υποκλάσεις ή τμήματα συστήματος της ανατομικής οντότητας «Brain», σύμφωνα με την FMA.



Εικόνα 6.1.1: «Brain» και οι κυριότερες συγγενείς κλάσεις, κατά FMA

6.1.1.2 ICD-10

Η ICD (International Classification of Diseases), όπως παρουσιάστηκε θεωρητικά και στην παράγραφο 3.1.5, είναι επί του παρόντος στην 10η αναθεώρηση (ICD-10) [35]. Η ICD-10 είναι μια κωδικοποίηση ασθενειών, συμπτωμάτων, παθολογικών ευρημάτων, κοινωνικών συνθηκών και εξωτερικών αιτιών τραυματισμών ή ασθενειών, όπως έχουν ταξινομηθεί από τον Παγκόσμιο Οργανισμό Υγείας (World Health Organization - WHO) και είναι αποδεκτοί παγκοσμίως. Το σύστημα κωδικοποίησης επιτρέπει την ύπαρξη πάνω από 14.400 διαφορετικούς κωδικούς και επιτρέπει την παρακολούθηση πολλών νέων διαγνώσεων. Χρησιμοποιώντας προαιρετική υπο-ταξινόμηση, το εύρος των κωδικών μπορεί να επεκταθεί σε πάνω από 16.000 κωδικούς. Ωστόσο, ICD-10 δεν μπορεί να θεωρηθεί ως οντολογία, είναι απλά μια ταξινόμηση. Αυτό σημαίνει ότι μόνο οι σχέσεις ιεραρχίας μπορούν να εφαρμοστούν στους όρους της. Δεν υπάρχουν άλλες σημασιολογικές σχέσεις μεταξύ των όρων της ICD-10. Παρ' όλα αυτά, είναι μία από τις πιο πλήρεις και ευρείας αποδοχής βάσεις γνώσης στον τομέα των ασθενειών. Για αυτό το λόγο κρίθηκε απαραίτητο να χρησιμοποιηθεί η ICD-10. Όπως παρουσιάζεται στην Εικόνα 6.1.2, που ακολουθεί, η «Sarcoidosis» (Σαρκοείδωση) είναι υποκλάση της κλάσης «Certain disorders involving the immune mechanism», ενώ η ίδια έχει υποκλάσεις τις ακόλουθες: «Sarcoidosis of lung», «Sarcoidosis of other and combined sites», «Sarcoidosis of lung with sarcoidosis of lymph nodes», «Sarcoidosis of skin», «Sarcoidosis of lymph nodes» και «Sarcoidosis, unspecified».



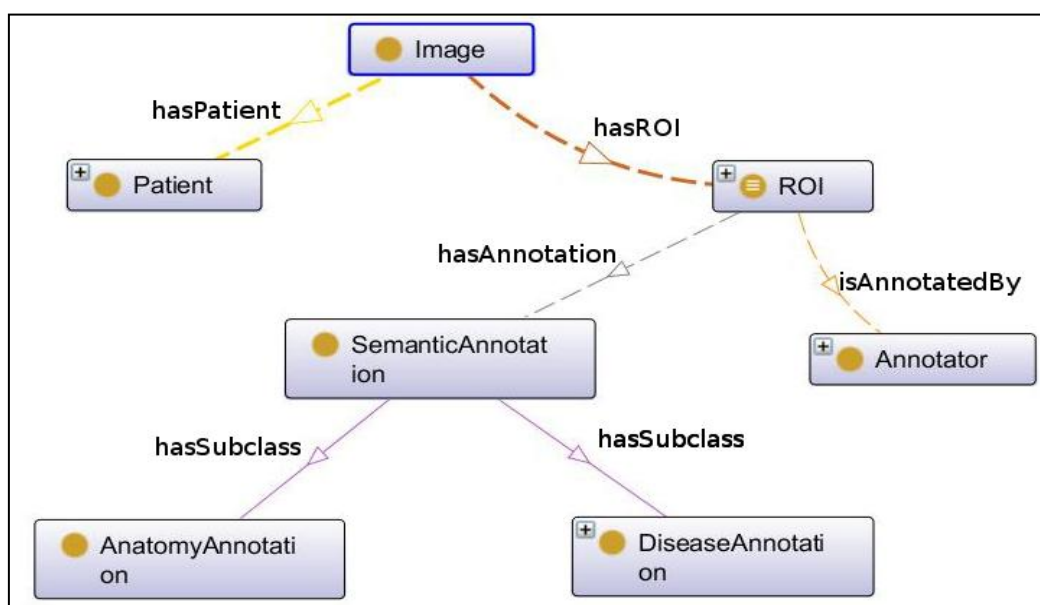
Εικόνα 6.1.2: «Sarcoidosis» και συγγενείς κλάσεις, κατά ICD-10

6.1.2 Οντολογία εφαρμογής

Για τις ανάγκες της εφαρμογής κρίθηκε αποτελεσματικό ο σχεδιασμός μιας οντολογίας εφαρμογής. Η οντολογία εφαρμογής σχεδιάστηκε με τη βοήθεια του εργαλείου Protégé 4.1 (βλ. παράγραφο 6.2.7). Ο βασικός κορμός αποτελείται, όπως παρουσιάζεται και στην Εικόνα 6.1.3, από τις ακόλουθες κλάσεις:

- Image (Εικόνα),
- ROI – Region Of Interest (Περιοχή Ενδιαφέροντος),
- SemanticAnnotation (Σημασιολογική Επισημείωση),
- AnatomyAnnotation (Επισημείωση Ανατομίας),
- DiseaseAnnotation (Επισημείωση Ασθένειας)
- Patient (Ασθενής),
- Annotator (Επισημειωτής).

Επίσης παρουσιάζονται και οι βασικές σχέσεις που τις συνδέουν. Η λογική ερμηνεία των σχέσεων είναι για παράδειγμα ότι «κάποια Εικόνα έχει Ασθενή κάποιον Ασθενή». Υπάρχουν, επίσης, μια σειρά από άλλες κλάσεις και πολλές άλλες σχέσεις. Επιπλέον, η οντολογία ενθυλακώνει κάποια αξιώματα, κανόνες και περιορισμούς για τον καθορισμό των κλάσεων και τον τρόπο που τις συνδέουν οι σχέσεις. Πληροφορίες όπως η ηλικία του ασθενούς αναπαρίστανται στην οντολογία εφαρμογής ως λεκτικό (literal) μέσω σχέσεων δεδομένων (DataProperties). Η οντολογία εφαρμογής, με όνομα «semia», παρατίθεται σε μορφή RDF/XML στο Παράρτημα Α.



Εικόνα 6.1.3: Κύριες κλάσεις και σχέσεις της Οντολογίας Εφαρμογής

Η οντολογία εφαρμογής αποτυπώνεται και οργανώνεται ως RDF τριάδες (RDF triples), ένα μοντέλο δεδομένων που συνιστάται από την W3C για τη σημασιολογική αναπαράσταση των δεδομένων [11]. Αυτές οι τριάδες αποθηκεύονται σε μία RDF αποθήκη (RDF repository). Για την υλοποίηση της εφαρμογής, χρησιμοποιήθηκε το σημασιολογικό αποθετήριο Sesame (βλ. παράγραφο 6.2.4), το οποίο είναι ανοικτού κώδικα (open-source). Επιπρόσθετα χρησιμοποιείται το ShiftOWLIM reasoner (βλ. παράγραφο 6.2.5) ως μηχανισμός συλλογιστικής και είναι ενσωματωμένος ως «Στρώμα Αποθήκευσης και Συμπερασμού» (Storage and Inference Layer) στο Sesame. Το ShiftOWLIM καλύπτει τις ανάγκες και την εκφραστικότητα της οντολογίας εφαρμογής και έτσι δεν κρίθηκε απαραίτητο να χρησιμοποιηθεί να χρησιμοποιηθεί κάποιο άλλο πρόγραμμα αυτοματοποιημένης συλλογιστικής, ικανό να διαχειριστεί γνώση υψηλότερης εκφραστικότητας, όπως το δημοφιλή Pellet.

6.1.3 Διάταξη Συστήματος και Διαδικτυακές Υπηρεσίες

Κατά τη φάση του σχεδιασμού, ο μηχανισμός των οντολογιών αναφοράς κρίθηκε απαραίτητος για την εφαρμογή, προκειμένου να επιτευχθεί υψηλό επίπεδο διαλειτουργικότητας με άλλα συστήματα. Η διαδικτυακή πρόσβαση, από την πλευρά της, σε αυτές τις οντολογίες οδήγησε στην τεχνοτροπία του Service Oriented Architecture (βλ. παράγραφο 5.2.2).

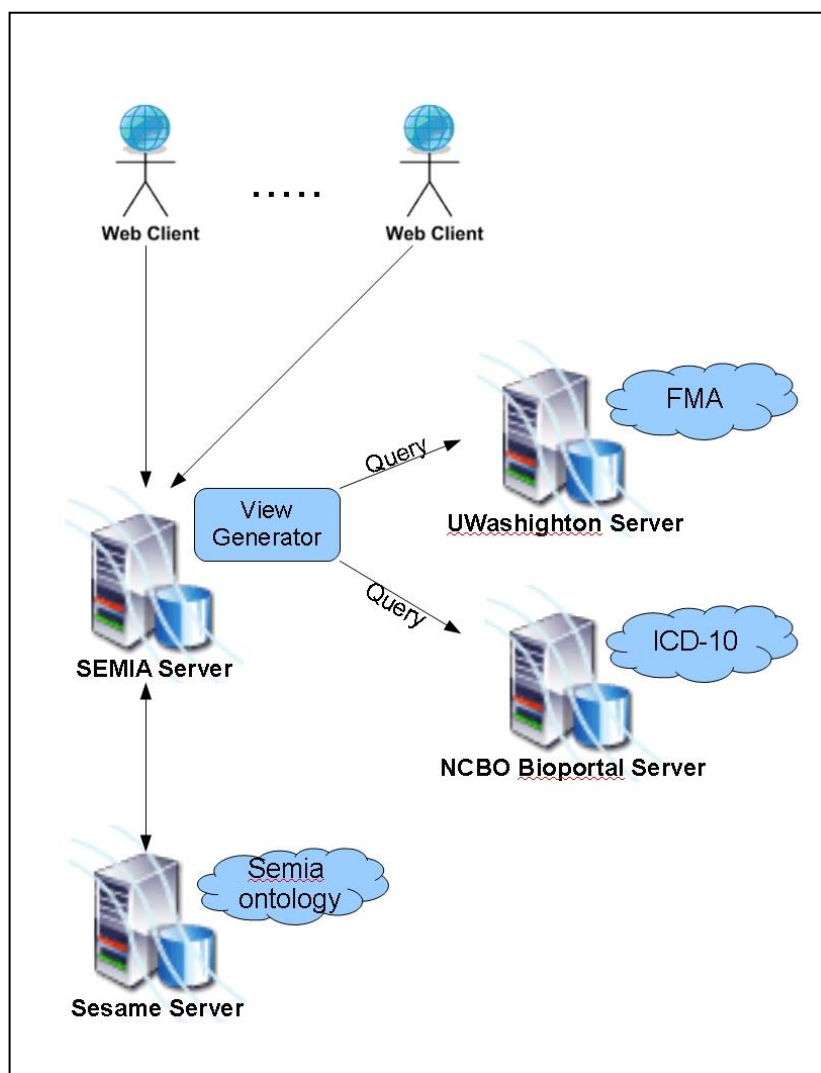
Συγκεκριμένα, η ICD-10 είναι προσβάσιμη μέσω υπηρεσιών, που είναι διαθέσιμες από τον εξυπηρετητή (server) του NCBO Bioportal¹ και είναι κυρίως τύπου REST. Το Bioportal αποτελεί το πληρέστερο αποθετήριο οντολογιών βιοϊατρικής και προσφέρει υπηρεσίες τόσο για την ανάκτηση των οντολογιών και ερωτημάτων σε αυτές (για παράδειγμα ανάκτηση των υποκλάσεων μιας συγκεκριμένης κλάσης, συγκεκριμένης οντολογίας) όσο και οπτικοποίηση αυτών.

Η FMA είναι προσβάσιμη μέσω δύο διαδικτυακών υπηρεσιών. Η πρώτη και κύρια βρίσκεται σε ένα εξυπηρετητή (server) του Πανεπιστημίου της Ουάσιγκτον, όπου η FMA σχεδιάστηκε και υλοποιήθηκε, και είναι μια υπηρεσία SOAP. Η γλώσσα ερωτημάτων που χρησιμοποιείται είναι η SPARQL, και ειδικότερα η VSPARQL που αποτελεί επέκταση της. Ο άλλος εξυπηρετητής (server) που καταστεί την FMA διαδικτυακά προσβάσιμη και διαθέσιμη είναι ο εξυπηρετητής του NCBO Bioportal.

Ακολουθώντας την τεχνοτροπία του Service Oriented Architecture, το RDF αποθετήριο του συστήματός μας είναι προσβάσιμο μέσω ερωτημάτων SRARQL και διαθέσιμο από τον SESAME server. Κατά αυτή τη παράταξη, το RDF αποθετήριο είναι ανοιχτό για χρήση από

¹ <http://bioportal.bioontology.org/>

άλλες υπηρεσίες. Έτσι, λοιπόν, επιτυγχάνεται η επιθυμητή διαλειτουργικότητα και επιπρόσθετα το σύστημα μπορεί να επεκταθεί ώστε να επικοινωνεί και με άλλα αποθετήρια. Επίσης, δεν υπάρχει περιορισμός για την επέκταση με άλλες οντολογίες αναφοράς ή την ενημέρωση των ήδη υπαρχόντων. Στην Εικόνα 6.1.4 παρουσιάζεται η Διάταξη του συστήματος.



Εικόνα 6.1.4: Διάταξη Συστήματος

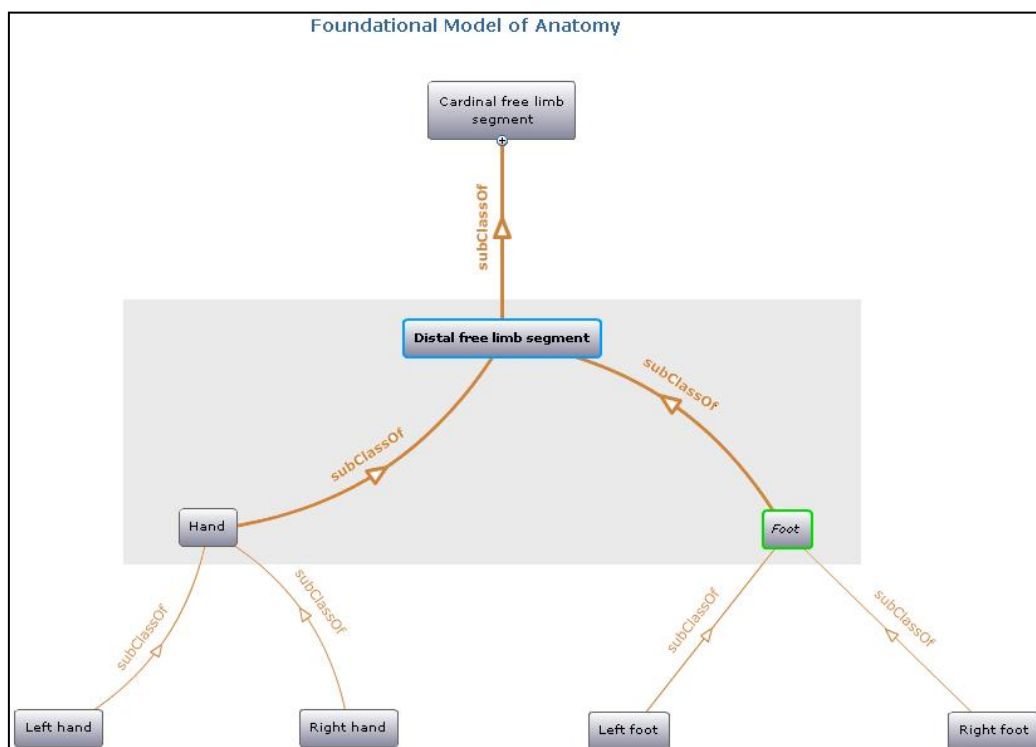
6.1.4 Καθορισμός Σημασιολογικής Γειτονιάς

Για τις ανάγκες του συστήματος θεωρήθηκε σκόπιμο να καθοριστεί η σημασιολογική γειτονιά των όρων ανατομίας και ασθένειας. Προσεγγίζοντας το ζήτημα αυτό από τη σκοπιά του τελικού χρήστη, διακρίνονται τρία σημεία ιδιαίτερης σημασίας:

1. Ο τελικός χρήστης είναι σχετικός και γνώστης των όρων ανατομίας και ασθενειών.

2. (Λόγω του 1) Η σημασιολογική γειτονιά δεν πρέπει να είναι γενικευμένη, αλλά στοχευμένη και εξειδικευμένη στον κάθε όρο.
3. (Λόγω των 1 και 2) Η σημασιολογική γειτονιά δεν πρέπει να είναι πολύ αναλυτική και εξειδικευμένη σε βάθος, καθώς αυτό θα ήταν κουραστικό και άχρηστο για το χρήστη.

Έτσι, σε ότι αφορά την FMA, η σημασιολογική γειτονιά των ανατομικών όρων καθορίζεται ως τα χωρικά μέρη (regional parts), συστατικά μέρη (constitutional parts), μέρη συστήματος (systemic parts) και οι υποκλάσεις (subclasses) της εκάστοτε κλάσης - ανατομικής οντότητας. Προσεγγίζοντας εννοιολογικά τη σημασιολογική γειτονιά, σε αυτή θα έπρεπε να συμπεριληφθούν οι υποκλάσεις των υποκλάσεων της εκάστοτε κλάσης - ανατομικής οντότητας, καθώς και μία επιπλέον πληθώρα κλάσεων ακολουθώντας την ίδια λογική και για τις άλλες σχέσεις. Εντούτοις γίνεται η παραδοχή να μη συμπεριλαμβάνονται αυτές οι κλάσεις και η εξόρυξη γνώσης να φτάνει σε βάθος μέχρι το πρώτο επίπεδο, όπως φαίνεται στην Εικόνα 6.1.5, διότι αυτές θα αποτελούσαν πλεόνασμα γνώσης για τον τελικό χρήστη (βλ. τρίτο σημείο ιδιαίτερης σημασίας). Συμπληρωματικά, ούτε η υπερκλάση, αντίστοιχα και για τις άλλες σχέσεις, της εκάστοτε κλάσης συμπεριλαμβάνεται στη σημασιολογική γειτονιά, γιατί ο τελικός χρήστης είναι γνώστης και δεν επιθυμεί γενικεύσεις (βλ. δεύτερο σημείο ιδιαίτερης σημασίας).



Εικόνα 6.1.5: Παράδειγμα Σημασιολογικής Γειτονιάς της ανατομικής κλάσης «Distal free limb segment», μόνο ως προς τη σχέση ιεραρχίας subClassOf

Ακολουθώντας την ίδια λογική και για την ICD-10, η οποία εμπεριέχει μόνο σχέσεις ιεραρχίας, γίνεται η παραδοχή να καθοριστεί η σημασιολογική γειτονιά των όρων ασθενειών ως οι υποκλάσεις της εκάστοτε κλάσης – όρου ασθένειας.

6.1.5 Διεπιφάνεια χρήστη

Τα ιδιαίτερα χαρακτηριστικά των πιθανών τελικών χρηστών του συστήματος επέβαλλαν τη δημιουργία μιας φιλικής προς το χρήστη διεπιφάνειας. Οι επαγγελματίες στο χώρο της ιατρικής και οι ερευνητές δεν είναι συχνά εξοικειωμένοι με τα πληροφοριακά συστήματα, ούτε επιθυμούν να υιοθετήσουν ένα νέο πρόγραμμα με όμορφα χαρακτηριστικά, αν αυτό δεν οδηγεί σε μια σημαντική βελτιστοποίηση της καθημερινής ροής εργασίας τους [9]. Έτσι, η ταχύτητα και η ευχρηστία αποτέλεσαν τον κεντρικό γνώμονα κατά την υλοποίηση της διεπαφής χρήστη. Το σκοπό αυτό, εξυπηρετούν και τα ακόλουθα συστατικά μέρη της διεπαφής χρήστη.

- **Αυτόματη συμπλήρωση.** Η αυτόματη συμπλήρωση προτεινόμενου κείμενου σε λίστα (auto-complete) χρησιμοποιείται σε διάφορα σημεία εισαγωγής κειμένου του συστήματος, όπου ο χρήστης καλείται να εισάγει κάποιον όρο ανατομίας ή ασθένειας. Έτσι ο χρήστης αρχίζει να πληκτρολογεί τον επιθυμητό όρο ανατομίας ή ασθένειας και το auto-complete αναλαμβάνει να του παρουσιάσει μία λίστα από όρους της FMA ή της ICD-10, ανάλογα το πεδίο, που ταιριάζουν με το ήδη υπάρχον κείμενο του χρήστη. Κατ' αυτό τον τρόπο ο χρήστης δεν χρειάζεται να πληκτρολογεί κάθε φορά ολόκληρο τον όρο που επιθυμεί και επιπλέον αποφεύγονται τυχόν ορθογραφικά λάθη. Επιπρόσθετα ο χρήστης εκπαιδεύεται στην ορολογία που χρησιμοποιείται από την FMA και την ICD-10.
- **Οδηγός Βημάτων.** Ο οδηγός βημάτων (wizard) χρησιμοποιείται στη διαδικασία επισημείωσης ιατρικών εικόνων και καθοδηγεί το χρήστη δίνοντας του οδηγίες για το τι πρέπει να κάνει στο κάθε βήμα. Ο οδηγός βημάτων είναι αρκετά διαδομένος, ως λειτουργικότητα υποβοήθησης, στη διαδικασία εκτέλεσης ενεργειών που πρέπει να διεκπεραιώσει ο χρήστης, όπως η συμπλήρωση μιας φόρμας. Υποδεικνύοντας ο οδηγός τις ενέργειες κατά βήμα, απλοποιείται και διεκπεραιώνεται ταχύτερα η διαδικασία, ενώ παράλληλα αποφεύγονται λάθη ή παρερμηνείες του χρήστη.
- **Τμήμα Υποστήριξης.** Η ιδέα του τμήματος υποστήριξης (help desk) είναι ευρέως διαδεδομένη σε νέα συστήματα που δεν παρουσιάζουν ομοιότητες με άλλα, στα οποία είναι συνηθισμένοι οι χρήστες. Σύμφωνα με αυτή τη λογική κρίθηκε σημαντική η ενσωμάτωση ενός online help desk στις ιστοσελίδες του συστήματος, που είναι απαραίτητο. Στο τμήμα υποστήριξης ενθυλακώνονται διαφορετικές λειτουργικότητες, ανάλογα με τις ενέργειες που πρόκειται να εκτελέσει ο χρήστης, με

σκοπό να παρέχουν βοήθεια στο χρήστη. Πολλές από αυτές τις λειτουργικότητες παρουσιάζονται στη συνέχεια.

- **Οπτικοποίηση Περιοχής Σώματος.** Αυτή η λειτουργικότητα αποτελεί μέρος του Help Desk και παρουσιάζεται στο χρήστη κατά τη διαδικασία της επισημείωσης μιας ιατρικής εικόνας. Η οπτικοποίηση περιοχής σώματος (body region visualization), ουσιαστικά, οπτικοποιεί την τιμή της ετικέτας «Body Part Examined», η οποία ανήκει στην επικεφαλίδα του DICOM αρχείου. Παρουσιάζει, δηλαδή, σε μια πρότυπη ανθρώπινη σιλουέτα την περιοχή σώματος που απεικονίζεται στην εικόνα του DICOM αρχείου. Στόχος της συγκεκριμένης λειτουργικότητας είναι να βοηθήσει το μη έμπειρο χρήστη να λύσει αμφιβολίες περί την απεικονισμένη περιοχή σώματος.
- **Σύννεφο Όρων Ανατομίας.** Το σύννεφο όρων ανατομίας (anatomic term cloud) αποτελεί και αυτό μέρος του Help Desk και συγκεκριμένα είναι διαθέσιμο στο χρήστη κατά την ενέργεια επισημείωσης με ανατομικούς όρους. Στο βήμα αυτό παρουσιάζονται στο σύννεφο όρων οι κλάσεις, που αποτελούν τη Σημασιολογική Γειτονιά, όπως αυτή καθορίστηκε στην παράγραφο 6.1.4, του επιθυμητού ανατομικού όρου. Στην ουσία η Σημασιολογική Γειτονιά παίζει το ρόλο συστάσεων προς το χρήστη ώστε να κάνει πιο συγκεκριμένη επισημείωση ή να επισημειώσει κάποια άλλη Περιοχή Ενδιαφέροντος (Region of Interest ή ROI). Πιθανότατα Περιοχές Ενδιαφέροντος της ίδιας εικόνας ανήκουν στην ίδια Σημασιολογική Γειτονιά.
- **Προβολή Δέντρου Οντολογίας.** Η προβολή δέντρου οντολογίας (ontology tree view) είναι χρήσιμη στο χρήστη τόσο κατά τη διαδικασία επισημείωσης όσο και κατά τη διαδικασία αναζήτησης. Ο χρήστης ανάλογα με τις ανάγκες του μπορεί να αποφασίσει ποια οντολογία προβάλλεται ανά πάσα στιγμή, η FMA ή η ICD-10. Συγκεκριμένα προβάλλεται μόνο η ιεραρχία της εκάστοτε οντολογίας. Η χρησιμότητα αυτής της λειτουργικότητας έγκειται στην ευχέρεια που δίνει στο χρήστη να πλοηγηθεί στην ιεραρχία και να κατανοήσει τη διάρθρωσή της ή να αποσαφηνίσει τη Σημασιολογική Γειτονιά των όρων, κυρίως της ICD-10, καθώς αυτή έχει μόνο σχέσεις ιεραρχίας μεταξύ των όρων της.
- **Εξωτερικοί Σύνδεσμοι.** Οι εξωτερικοί σύνδεσμοι (external links), οι οποίοι είναι εξατομικευμένοι ανά όρο, οδηγούν το χρήστη σε ιστοσελίδες του NCBO Bioportal. Οι σελίδες αυτές παρουσιάζουν τις λεπτομέρειες του εκάστου όρου ή την οπτικοποίηση του στο δίκτυο των σχέσεων του με άλλους όρους, και έτσι ολοκληρώνεται η γνώση του χρήστη για τις ιδιότητες του κάθε όρου.
- **Φίλτρα Αναζήτησης.** Το σύστημα προσφέρει διάφορα φίλτρα κατά τη διαδικασία της αναζήτησης, με σκοπό ο χρήστης να μειώσει τον αριθμό των αποτελεσμάτων,

καθορίζοντας συγκεκριμένες τιμές των μεταδεδομένων των εικόνων. Για παράδειγμα, ο χρήστης μπορεί να αναζητήσει εικόνες με μια συγκεκριμένη τροπικότητα (π.χ. CT) ή που ανήκουν σε ασθενείς εντός ενός συγκεκριμένου ηλικιακού φάσματος. Επιπλέον, ο χρήστης μπορεί να καθορίσει ρητά ποιιά σχέση της FMA να συμπεριληφθεί στο ερώτημα χρησιμοποιώντας το φίλτρο "ανατομικές σχέσεις".

- Κατηγοριοποίηση Αποτελεσμάτων κατά Κλάσεις Ανατομίας και Σύνδεσμοι Κλάσεων. Τα αποτελέσματα αναζήτησης επισυνάπτεται στην λίστα των αποτελεσμάτων, κατηγοριοποιημένα σύμφωνα με την κλάση ανατομίας, στην οποία ανήκουν. Κατ' αυτόν τον τρόπο μία εικόνα μπορεί να ανήκει ταυτόχρονα σε δύο ή και περισσότερες κλάσεις, αλλά αυτό είναι αποδεκτό και εννοιολογικά σωστό. Ακόμη για την πιο εύκολη πλοήγηση, του χρήστη, στη λίστα αποτελεσμάτων, κρίθηκε σκόπιμο στην αρχή της λίστας να υπάρχουν σύνδεσμοι μεταπήδησης στην αντίστοιχη κλάση.

6.2 Πλατφόρμες και Προγραμματιστικά Εργαλεία

Σε αυτή την ενότητα αναλύονται όλες οι τεχνολογίες και τα προγραμματιστικά εργαλεία που χρησιμοποιήθηκαν για την υλοποίηση του συστήματος.

6.2.1 Java servlet

Για τη συγγραφή του κώδικα εξυπηρετητή (server-side programming) επιλέχθηκε η προγραμματιστική γλώσσα Java και συγκεκριμένα υλοποιήθηκε ένας αριθμός από Java Servlets. Το Java Servlet είναι μία κλάση της Java που επεκτείνει τις δυνατότητες των εξυπηρετητών ιστού δημιουργώντας δυναμικό περιεχόμενο – συνήθως μορφής html – μέσω του προγραμματιστικού μοντέλου ερωτήσεων – απαντήσεων. Αναλυτικά, ο εξυπηρετητής ιστού Apache Tomcat 6.0 που χρησιμοποιήσαμε δέχεται την http αίτηση από τον πελάτη και καλεί το αντίστοιχο servlet, περνώντας του τις παραμέτρους του χρήστη. Η κλάση του servlet επιτελεί τις απαραίτητες λειτουργίες σύμφωνα με τη λογική που περιέχει και τις παραμέτρους του χρήστη και αποκρίνεται επιστρέφοντας δεδομένα σε μορφή html ή άλλη. Σε αρκετά σημεία της εφαρμογής χρησιμοποιήθηκε η τεχνολογία των Java Servlets σε συνδυασμό με την τεχνολογία AJAX (Asynchronous Javascript And XML) για την πραγματοποίηση ασύγχρονων αιτήσεων μεταξύ πελάτη και εξυπηρετητή. Τα Java Servlets εκμεταλλεύονται όλα τα πλεονεκτήματα που προσφέρει η Java, όπως αντικειμενοστρέφεια, λειτουργία ανεξάρτητου πλατφόρμας και ασφάλεια μέσω χειρισμού εξαιρέσεων και συλλέκτη σκουπιδιών. Επίσης, επιτρέπουν την ενσωμάτωση της πληθώρας έτοιμων εργαλείων και βιβλιοθηκών της java με εύκολα προσβάσιμη και πλήρη τεκμηρίωση (documentation), αλλά και υπηρεσιών τύπου REST, SOAP και άλλων.

6.2.2 *Java applet*

Η ανάγκη για επεξεργασία της ιατρικής εικόνας από την πλευρά του πελάτη (client-side) οδήγησε στη συγγραφή ενός Java Applet. Τα Java Applets είναι μικρές εφαρμογές γραμμένες σε Java, οι οποίες ενσωματώνονται στις html σελίδες και εκτελούνται στους περιηγητές ιστού. Αναλυτικά, όταν ανιχνευτεί ένα αντικείμενο τύπου Java Applet μέσα στην html σελίδα, ο περιηγητής του συστήματος του πελάτη φορτώνει από τον εξυπηρετητή το συρμό από bytes και εκτελείται στην εικονική μηχανή της java (virtual machine) του περιηγητή. Στην εφαρμογή μας, το applet που δημιουργήθηκε επιτελεί τις εξής λειτουργίες: φόρτωση εικόνας, εφαρμογή φίλτρων στην εικόνα, επιλογή των σημείων ενδιαφέροντος στην εικόνα και, τέλος, μετατροπή της τελικής εικόνας σε μορφή html και αποστολή της στον εξυπηρετητή ιστού προς αποθήκευση. Όλες οι παραπάνω λειτουργίες γίνονται από την πλευρά του πελάτη και θα ήταν πολύ δύσκολο να υλοποιηθούν χωρίς τη χρήση του applet, μέσω, για παράδειγμα, της χρήσης μίας γλώσσας συγγραφής σεναρίων (script language) όπως javascript ή php. Το υλοποιημένο applet είναι μία επέκταση του ImageJA¹, ενός λογισμικού ανοικτού κώδικα για διαχείριση εικόνων.

6.2.3 *jQuery*

Η jQuery είναι μια ανεξάρτητη περιηγητή βιβλιοθήκη Javascript που βοηθά τον προγραμματιστή στη συγγραφή κώδικα πελάτη σε σελίδες HTML. Η jQuery απλοποιεί τη διάσχιση του html εγγράφου, το χειρισμό συμβάντων και την αλληλεπίδραση μέσω AJAX κλήσεων με τον εξυπηρετητή. Η εφαρμογή που δημιουργήθηκε είναι αρκετά πλούσια σε Javascript, έτσι ώστε να αυξάνεται η εμπειρία χρήσης. Για το λόγο αυτό χρησιμοποιήθηκε αρκετά η βιβλιοθήκη jQuery ως εργαλείο για την ευκολότερη συγγραφή του κώδικα πελάτη. Επιπλέον, χρησιμοποιήθηκε η πρόσθετη βιβλιοθήκη JQuery UI η οποία παρέχει μια πληθώρα από οπτικά εφέ, βασισμένα στη JQuery.

6.2.4 *Sesame*

Όλα τα δεδομένα που αποθηκεύονται στο σύστημα είναι τα στιγμιότυπα των κλάσεων της οντολογίας εφαρμογής και είναι σε μορφή τριάδων. Για την αποθήκευση των τριάδων κρίθηκε πως απαιτούνταν η χρήση μίας αποθήκης τριάδων (triplestore), καθώς οι αποθήκες τριάδων είναι σχεδιασμένες εξ' αρχής με βάση τη φιλοσοφία των τριάδων και προσφέρουν καλύτερη διαχείριση των RDF δεδομένων σε σχέση, για παράδειγμα, με την κλασική λύση της σχεσιακής βάσης δεδομένων. Εξάλλου, οι σχεσιακές βάσεις δεδομένων δίνουν περιορισμένες δυνατότητες, καθώς όσο αυξάνει η πολυπλοκότητα των ερωτημάτων τόσο πιο

¹ <http://pacific.mpi-cbg.de/wiki/index.php/ImageJA>

δύσκολο είναι αυτά να μεταφραστούν σε ερωτήματα SQL. Ένα από τα μειονεκτήματα των αποθηκών τριάδων είναι αυτό της δυσκολίας κλιμάκωσης, δηλαδή η διαχείρισης τεράστιου όγκου δεδομένων. Στην εφαρμογή μας έγινε η παραδοχή ότι η κλιμάκωση θα παραμείνει σε επίπεδα που επιτρέπουν της χρήση μιας αμιγώς σημασιολογικής αρχιτεκτονικής για την αποθήκευση.

Φυσικά, ο τομέας των αποθηκών τριάδων βρίσκεται σε ένα σχετικά πρώιμο στάδιο και απέχει αρκετά από την ωριμότητα των σχεσιακών βάσεων δεδομένων. Παρόλα αυτά, έχουν αναπτυχθεί κάποια αρκετά αξιόλογα πλαίσια αποθηκών τριάδων, ανάμεσα στα οποία βρίσκεται το Sesame, ένα ανοιχτού κώδικα (άδεια τύπου BSD) πλαίσιο που καλύπτει όλο το φάσμα των απαιτήσεων της εφαρμογής. Το Sesame υποστηρίζει τη γλώσσα SPARQL για ερωτήματα επί των RDF δεδομένων. Παρέχει πρόσβαση είτε μέσω API είτε μέσω REST υπηρεσίας. Στην εφαρμογή που υλοποιήθηκε χρησιμοποιήθηκε η REST υπηρεσία, καθώς είναι ανεξάρτητη γλώσσας προγραμματισμού και μπορεί να οδηγήσει εύκολα σε αποκεντρωμένη πρόσβαση. Το Sesame εγκαταστάθηκε στον ίδιο εξυπηρετητή Apache Tomcat στον οποίο εγκαταστάθηκε και η εφαρμογή. Είναι σημαντικό να τονιστεί ότι η συγκεκριμένη αποθήκη τριάδων παρέχει τη δυνατότητα, μέσω της REST υπηρεσίας, σε οποιονδήποτε έχει τα κατάλληλα διαπιστευτήρια για πρόσβαση στα δεδομένα της εφαρμογής με ένα πρωτόκολλο τόσο απλό και διαδεδομένο όσο το http πρωτόκολλο.

6.2.5 *SwiftOWLIM*

Το πρόγραμμα συλλογιστικής (reasoner) που χρησιμοποιήθηκε είναι το SwiftOWLIM, το οποίο προσφέρεται ως *Στρώμα Αποθήκευσης και Συμπερασμού* για την αποθήκη τριάδων Sesame. Το SwiftOWLIM υποστηρίζει συλλογιστική για οντολογίες που αποτυπώνονται σε ένα υποσύνολο της OWL-Lite διαλέκτου, την Owl-Horst. καθώς η εκφραστικότητα της οντολογίας εφαρμογής είναι σχετικά μικρή, ο συμπερασμός με SwiftOWLIM καλύπτει τις ανάγκες της εφαρμογής μας. Ο συγκεκριμένος reasoner χρησιμοποιεί εμπροστοδρομική αλύσωση και συμπερασμό με βάση κανόνες (rule based reasoning). Για παράδειγμα, όταν μια νέα τριάδα εισαχθεί στο σύστημα, ενεργοποιούνται όλοι οι σχετικοί κανόνες, παράγονται όλες οι εξαγόμενες τριάδες – υπονοούμενη γνώση – και εισάγονται και αυτές στο σύστημα. Στα πλεονεκτήματα του SwiftOWLIM συγκαταλέγονται η ταχύτητα απόκρισης και παραγωγής γνώσης και η δυνατότητα κλιμάκωσης που προσφέρει.

6.2.6 *Flash*

Για την οπτικοποίηση των οντολογιών αναφοράς FMA και ICD μέσω δέντρων χρησιμοποιήθηκαν τα εργαλεία που προσφέρονται από το Bioportal σε μορφή Adobe Flash. Οι εφαρμογές Flash της εταιρίας Adobe είναι αντικείμενα που ενσωματώνονται σε html

σελίδες και εμπλουτίζουν τις σελίδες με βίντεο, γραφικά και κίνηση κάνοντας τις πιο διαδραστικές. Στην εφαρμογή τα δέντρα χρησιμοποιούνται για την εύρεση των σημασιολογικά κοντινών όρων στους όρους επισημείωσης και αναζήτησης, όπως υπερκλάσεις και υποκλάσεις. Για παράδειγμα, όταν ο χρήστης επιλέγει έναν όρο ασθένειας, ο όρος αυτός ανιχνεύεται στο δέντρο ασθενειών έτσι ώστε να παρουσιαστούν ασθένειες που ανήκουν σε υποκλάσεις (πιο συγκεκριμένη επισημείωση), υπερκλάσεις (πιο γενική επισημείωση) ή συγγενείς κλάσεις (διόρθωση επισημείωση).

6.2.7 Protégé

Για την ανάπτυξη της οντολογίας εφαρμογής χρησιμοποιήσαμε το πρόγραμμα Protégé. Το Protégé είναι ένας συντάκτης οντολογιών ανοικτού κώδικα του πανεπιστημίου του Stanford. Περιέχει γραφικό περιβάλλον στο οποίο ο χρήστης έχει τη δυνατότητα να ορίσει κλάσεις, σχέσεις ανάμεσα σε αυτές και στιγμιότυπα των κλάσεων. Επίσης περιέχει πρόσθετα εργαλεία, όπως ταξινομητές (classifiers) και συλλογιστές (reasoners) που επιτρέπουν τον έλεγχο ορθότητας της μοντελοποιούμενης γνώσης και την εύρεση τυχόν ασυνεπειών. Επίσης, αποτυπώνει την οντολογία σε μία περιγραφική λογική, φανερώνοντας την εκφραστικότητά της.

Η διαδικασία που ακολουθήθηκε κατά την ανάπτυξη της εφαρμογής είχε ως πρώτο βήμα την ανάπτυξη της οντολογίας της εφαρμογής με τη χρήση του Protégé, χωρίς την προσθήκη στιγμιότυπων. Στη συνέχεια, η οντολογία σε μορφή RDF/XML (RDF τριάδες) αποθηκεύτηκε σε μια αποθήκη τριάδων Sesame, στην οποία τοποθετούνται και οι τριάδες που παράγονται από τη διαδικασία της επισημείωσης ως στιγμιότυπα των κλάσεων της οντολογίας. Έτσι, στην ίδια αποθήκη τριάδων υπάρχουν και οι καταχωρήσεις (στιγμιότυπα) και η σημασιολογία αυτών (οντολογία) και καθίσταται δυνατή η αυτόματη εξαγωγή γνώσης.

6.2.8 Subversion

Ένα από τα πιο σημαντικά εργαλεία που χρησιμοποιηθήκαν για την ανάπτυξη της εφαρμογής είναι το λογισμικό subversion. Το πρόγραμμα αυτό χαρακτηρίζεται ως σύστημα ελέγχου αναθεωρήσεων (revision control system) του κώδικα που γράφεται κατά τη διαδικασία ανάπτυξης μίας εφαρμογής. Η λειτουργία που επιτελεί είναι να κρατά τις επιμέρους αναθεωρήσεις σε μία κεντρική αποθήκη και να εντοπίζει τις διαφορές από έκδοση σε έκδοση και τις κάνει εμφανείς στους προγραμματιστές όταν επιχειρούν να αποθηκεύσουν κάποια νέα αναθεώρηση του κώδικα. Όταν εντοπίζεται κάποιο σημείο σύγκρουσης ανάμεσα σε δύο εκδόσεις του ίδιου κώδικα, το πρόγραμμα δεν επιτρέπει την αποθήκευση της νέας έκδοσης έως ότου επιλυθεί η σύγκρουση. Με τον τρόπο αυτό, πολλοί προγραμματιστές μπορούν να δουλεύουν στο ίδιο έργο χωρίς να φοβούνται μήπως σβήσουν ή χαλάσουν κάποιο κομμάτι

κώδικα. Αυτό επιτυγχάνεται τόσο επειδή υπάρχουν διαθέσιμες ανά πάσα στιγμή όλες οι προηγούμενες εκδόσεις, όσο και γιατί εάν υπάρξει κάποια σύγκρουση αυτό γίνεται σαφές. Η χρήση λοιπόν του λογισμικού αυτού, έκανε αρκετά εύκολο έργο το να συντονιστούν οι εργασίες των δύο προγραμματιστών που συνεργάστηκαν για την ανάπτυξη της εφαρμογής. Όσον αφορά στην αποθήκη των εκδόσεων του κώδικα της εφαρμογής, έγινε χρήση της online υπηρεσίας *assembla*¹ που παρέχει δωρεάν online χώρο αποθήκευσης και δυνατότητες *subversioning*.

6.2.9 Πλατφόρμα ανάπτυξης

Η ανάπτυξη του κώδικα έγινε σε περιβάλλον Windows Vista και Windows XP (δοκιμάστηκε ωστόσο και σε Linux). Για τη συγγραφή του κώδικα χρησιμοποιήθηκε το ολοκληρωμένο περιβάλλον ανάπτυξης (Integrated Development Environment ή IDE) Eclipse που παρέχει τόσο έναν συντάκτη κώδικα με πλήρη υποστήριξη για αρχεία html, javascript και java, όσο και ένα γραφικό περιβάλλον για το *subversion*. Η εφαρμογή δοκιμάστηκε κατά την ανάπτυξη της τόσο μέσω των Mozilla Firefox και Chrome όσο και μέσω του Safari. Για την επίδειξη η εφαρμογή μεταφέρθηκε σε εξυπηρετητή του Ε.Μ.Π (περιβάλλον Ubuntu Linux) κάνοντάς την ευρέως διαθέσιμη μέσω διαδικτύου.

6.3 Έλεγχος με βάση Σενάρια Λειτουργίας

Στην ενότητα αυτή παρουσιάζονται τα σενάρια λειτουργίας, δηλαδή τα υποθετικά σενάρια δράσης ενός χρήστη όπου επιδεικνύονται οι απαιτούμενες ενέργειες του χρήστη και η απόκριση του συστήματος σε αυτές. Μέσα από την ανάλυση των σεναρίων λειτουργίας το σύστημα μπορεί τόσο να ελεγχθεί όσο και να αξιολογηθεί ως προς τη λειτουργικότητα και την ορθότητά του. Τα σενάρια που ελέγχθηκαν είναι τα εξής:

- Προσθήκη νέου χρήστη
- Επισημείωση εικόνας
- Αναζήτηση εικόνων

Καθένα από τα παραπάνω σενάρια παρουσιάζεται αναλυτικά στη συνέχεια.

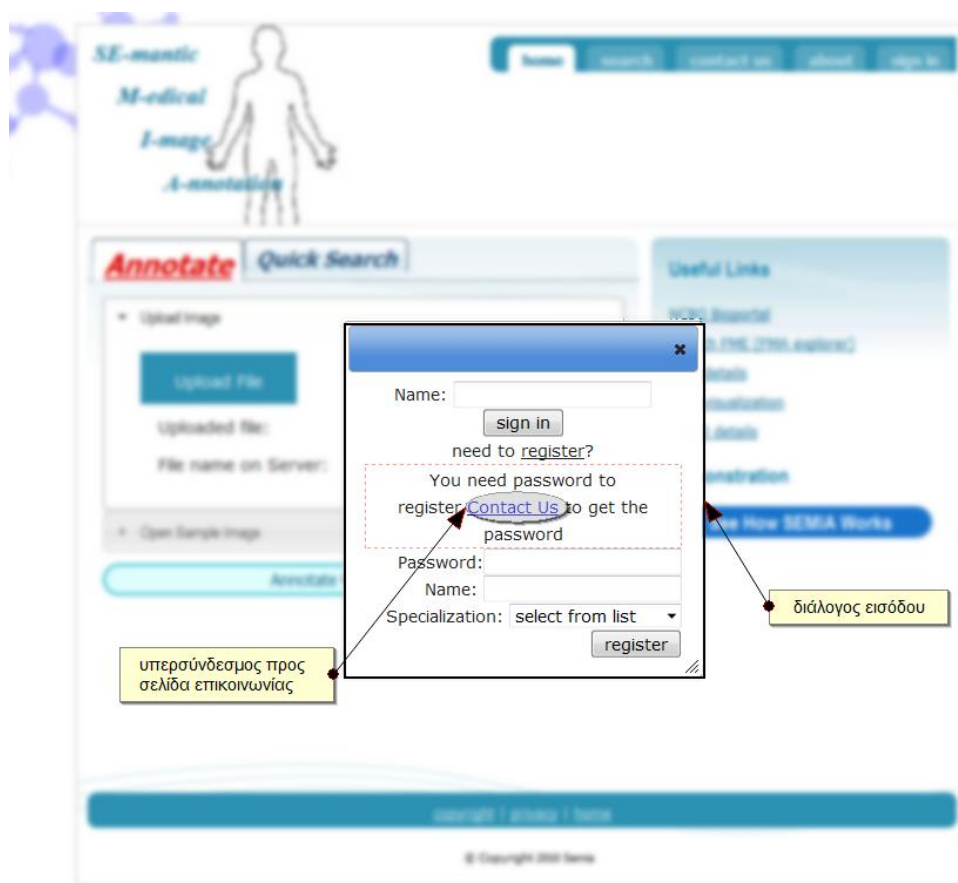
6.3.1 Προσθήκη νέου χρήστη

Για να μπορέσει ένας χρήστης να χρησιμοποιήσει το εργαλείο επισημείωσης πρέπει να εγγραφεί στο σύστημα. Για την εγγραφή απαιτείται η επικοινωνία με τους διαχειριστές του συστήματος προς παροχή του απαραίτητου αναγνωριστικού εγγραφής (password). Αυτός

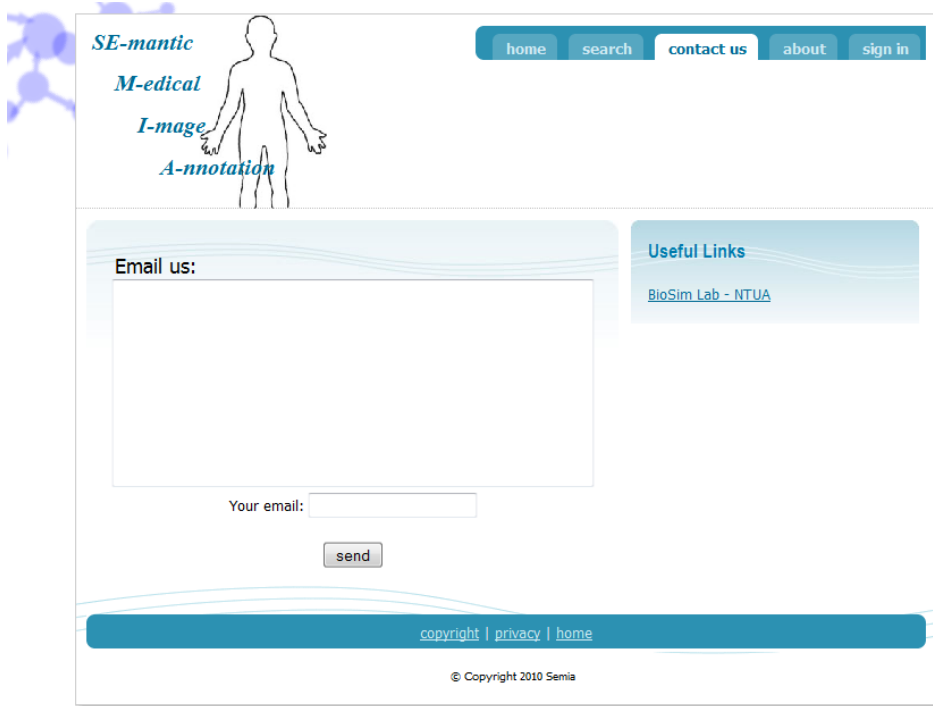
¹ <http://www.assembla.com>

είναι ένας τρόπος με τον οποίο εξασφαλίζεται ότι το σύστημα θα έχει πάντα έγκυρες, επιστημονικές επισημειώσεις.

Όταν ο χρήστης επιχειρήσει να εισέλθει στο σύστημα από το κουμπί εισόδου, ανοίγει ο διάλογος εισόδου (Εικόνα 6.3.1). Σε περίπτωση νέου χρήστη, μπορεί να χρησιμοποιηθεί ο υπερσύνδεσμος προς τη σελίδα επικοινωνίας (Εικόνα 6.3.2) στην οποία ο χρήστης αποστέλλει τα στοιχεία του που κρίνει απαραίτητα για να γίνει δεκτός από τους διαχειριστές του συστήματος ως επισημειωτής (annotator).

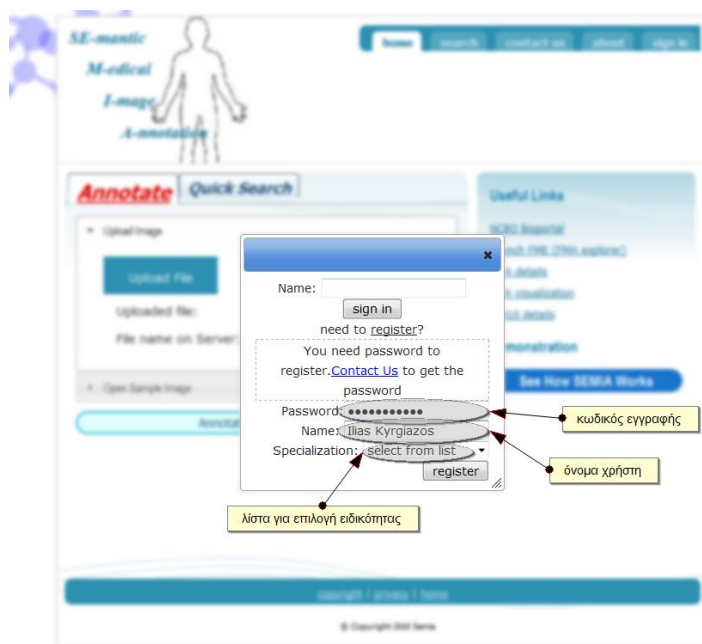


Εικόνα 6.3.1: Διάλογος Εγγραφής νέου χρήστη.

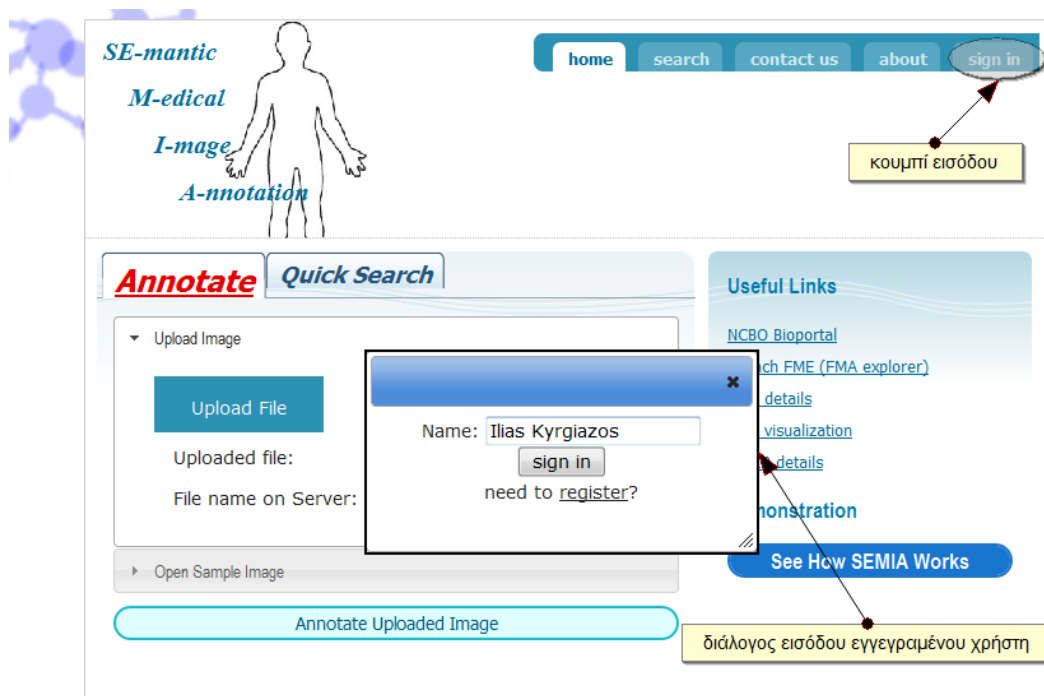


Εικόνα 6.3.2: Σελίδα επικοινωνίας.

Όταν αποκτηθεί ο κωδικός εγγραφής, ο χρήστης εισάγει τον κωδικό εγγραφής και το όνομά του, επιλέγει την ειδικότητά του και εγγράφεται στο σύστημα (Εικόνα 6.3.3). Από αυτό το σημείο μπορεί να εισέρχεται στο σύστημα χρησιμοποιώντας το κουμπί εισόδου πληκτρολογώντας κάθε φορά το όνομά του και μόνο (Εικόνα 6.3.4).



Εικόνα 6.3.3: Εγγραφή νέου χρήστη.



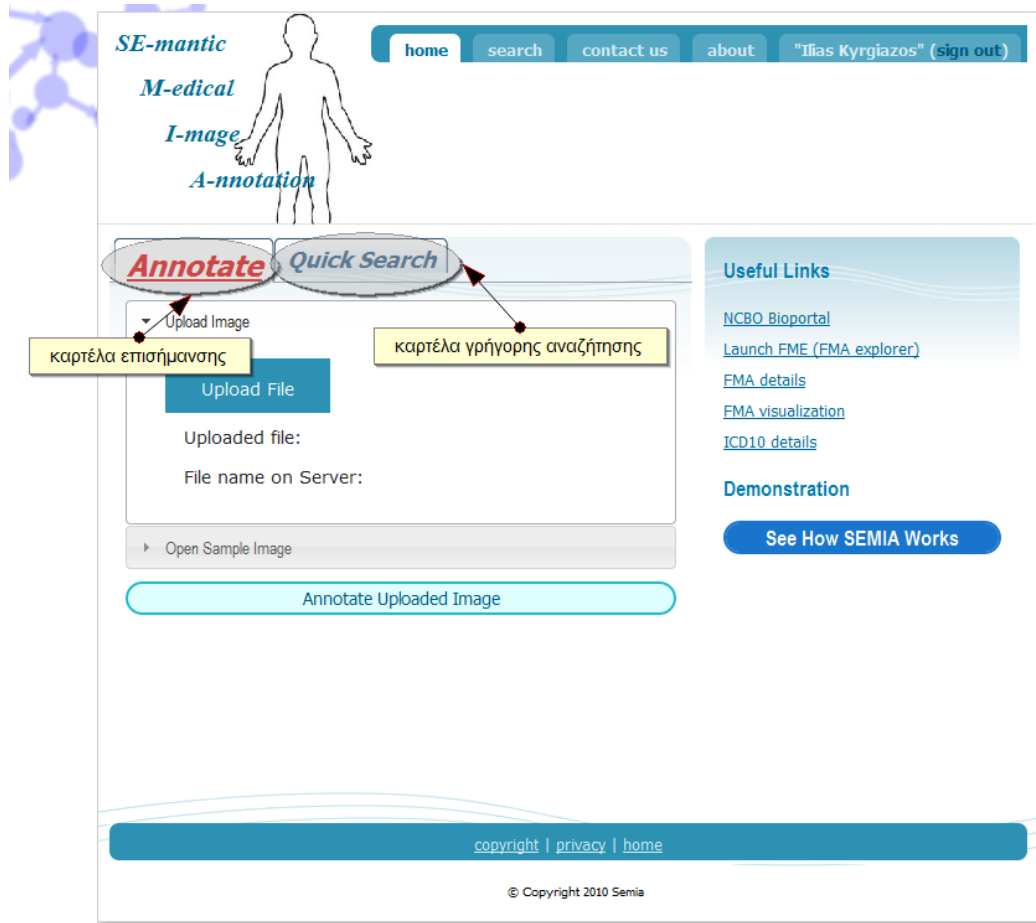
Εικόνα 6.3.4: Είσοδος χρήστη στο σύστημα.

6.3.2 Επισημείωση εικόνας

Μετά από την είσοδό του στο σύστημα ο χρήστης μπορεί να προχωρήσει στην επισημείωση εικόνων. Καθώς η διαδικασία αυτή είναι αρκετά σύνθετη και περιέχει πολλά βήματα διαιρείται εδώ σε υπο-διαδικασίες για μεγαλύτερη σαφήνεια.

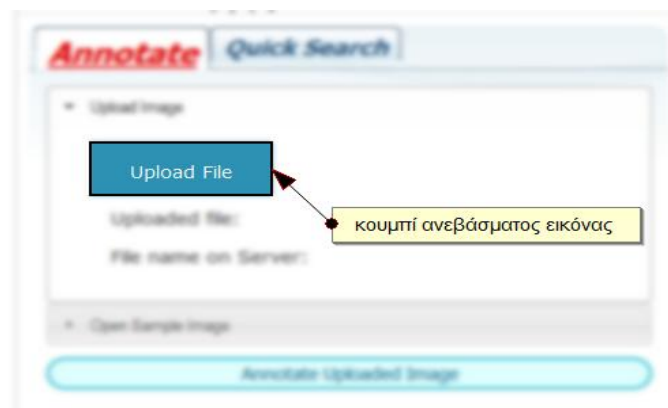
6.3.2.1 Ανέβασμα εικόνας (Upload)

Στην κεντρική οθόνη της εφαρμογής παρατηρούμε ότι υπάρχουν δύο βασικές καρτέλες: επισημείωση και αναζήτηση (Εικόνα 6.3.5).



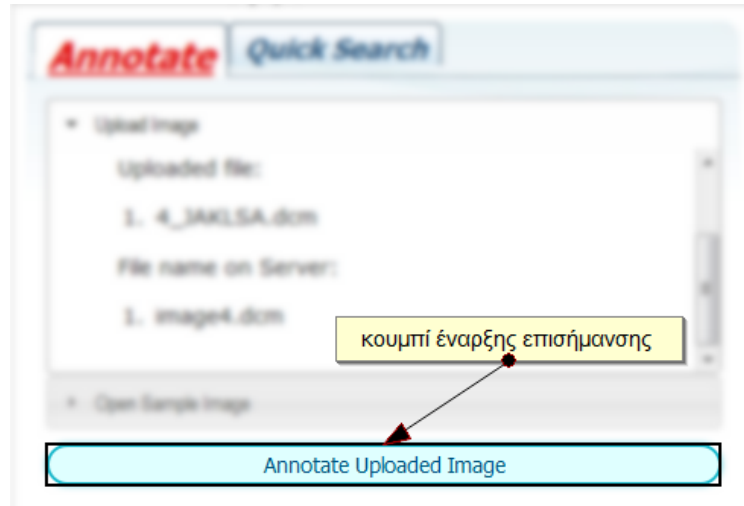
Εικόνα 6.3.5: Κεντρική οθόνη εφαρμογής - βασικές καρτέλες

Ο χρήστης επιλέγει την καρτέλα «Επισήμειωση» (Annotate) (η οποία είναι και προεπιλεγμένη) και επιλέγει μία DICOM εικόνα από το τοπικό του αρχείο πατώντας το κουμπί Upload File (Εικόνα 6.3.6). Να σημειωθεί ότι ο χρήστης έχει τη δυνατότητα να ανεβάσει και άλλες εικόνες, πέρα από DICOM, αλλά κάποιες λειτουργικότητες δεν θα είναι διαθέσιμες στην περίπτωση αυτή. Επίσης, είναι επιθυμητό τα αρχεία να έχουν την κατάληξη “.dcm”.



Εικόνα 6.3.6: Ανέβασμα εικόνας

Όταν η εικόνα φορτωθεί στον εξυπηρετητή με επιτυχία, κατάλληλο μήνυμα ενημερώνει το χρήστη και αυτός μπορεί να προχωρήσει στην επισήμανσή της πατώντας το κουμπί Annotate Uploaded Image (Εικόνα 6.3.7).

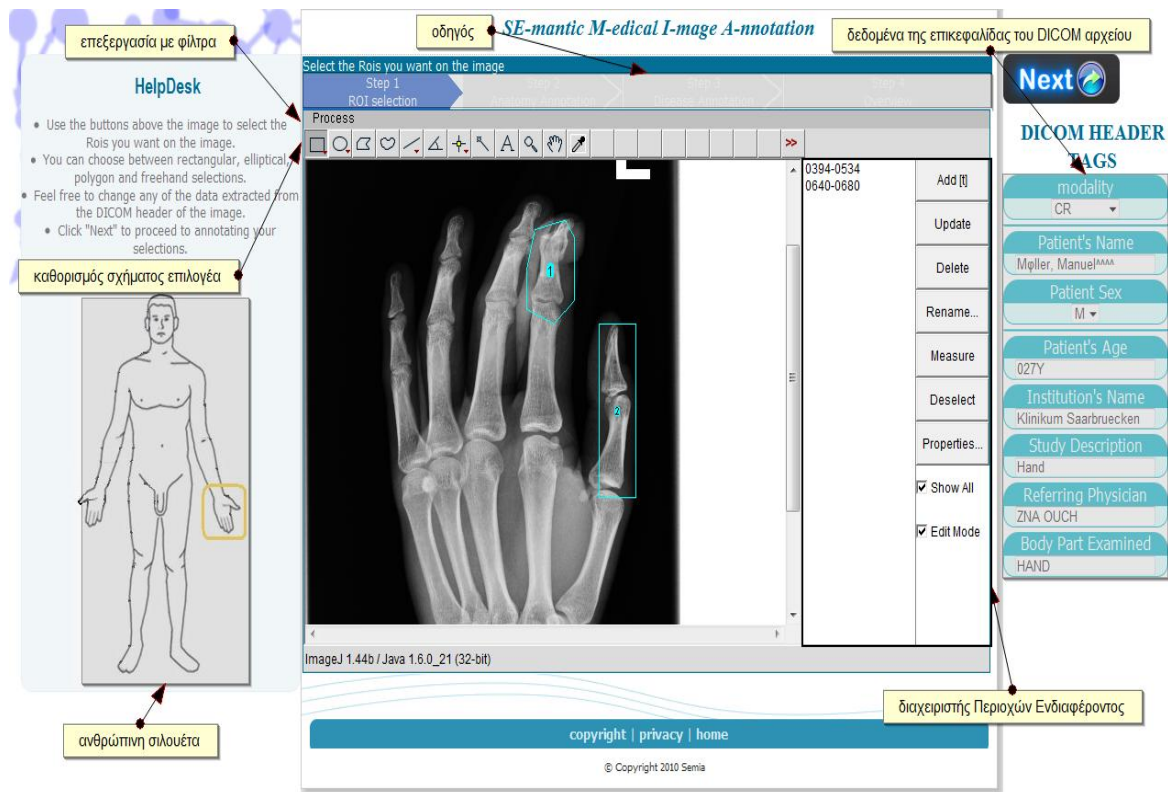


Εικόνα 6.3.7: Έναρξη επισημείωσης.

Στη συνέχεια αρχίζει η επισημείωση της εικόνας. Η επισημείωση έχει χωριστεί σε τέσσερα βήματα, τα οποία φαίνονται στο οδηγό στο πάνω μέρος της οθόνης επισημείωσης. Κάθε φορά το τρέχον βήμα όπως και τα ολοκληρωμένα βήματα είναι σημειωμένα στον οδηγό με βάση έναν χρωματικό κώδικα (Εικόνα 6.3.8).

6.3.2.2 Επισημείωση: Βήμα Πρώτο

Στο πρώτο βήμα η εικόνα εμφανίζεται στον χρήστη και τα δεδομένα που εξάγονται αυτόματα από τα επιλεγμένα πεδία της επικεφαλίδας DICOM εμφανίζονται δεξιά. Τα δεδομένα αυτά μπορούν να διορθωθούν ή να συμπληρωθούν από το χρήστη και αποθηκεύονται στο τέλος της διαδικασίας, μαζί με τα χαρακτηριστικά ανατομίας και ασθένειας, ως μεταδεδομένα της εικόνας.



Εικόνα 6.3.8: Πρώτο βήμα επισημείωσης.

Όπως βλέπουμε στην Εικόνα 6.3.8, ο χρήστης μπορεί να επιλέξει ανάμεσα σε διάφορα φίλτρα που μπορεί να εφαρμόσει στην εικόνα πατώντας το κουμπί Process αλλά και να επιλέξει το επιθυμητό σχήμα για τον επιλογέα των Περιοχών Ενδιαφέροντος (ROIs) από το μενού πάνω από την εικόνα. Στη συνέχεια ο χρήστης επιλέγει και εισάγει διαδοχικά τα κατάλληλα ROIs στο διαχειριστή των ROIs δεξιά από την εικόνα. Ο διαχειριστής επιτρέπει την εισαγωγή, ανανέωση, διαγραφή, μέτρηση χαρακτηριστικών των ROIs με τη χρήση των αντίστοιχων κουμπιών.

Τέλος, υπάρχει μια οπτικοποίηση του μέρους του ανθρώπινου σώματος το οποίο απεικονίζει η εικόνα – εδώ το χέρι - μέσω της χρωματικής επισήμανσής της σε μία ανθρώπινη σιλουέτα.

6.3.2.3 Επισημείωση: Βήμα Δεύτερο

Πατώντας το κουμπί Next η πλοήγηση συνεχίζεται και ο χρήστης προχωρά στο δεύτερο βήμα (Εικόνα 6.3.9). Εδώ μπορεί να εισάγει τα κατάλληλα ανατομικά χαρακτηριστικά στα ROIs που επιλέχθηκαν στο βήμα 1. Για κάθε ROI αρχίζει να πληκτρολογεί έναν ανατομικό όρο και επιλέγει από τη λίστα που ενημερώνεται αυτόματα με σχετικούς όρους από την οντολογία ανατομίας. Όταν επιλεγεί ο όρος, εμφανίζεται στα αριστερά της οθόνης ένα σύννεφο από σχετικούς όρους – συστάσεις από τις οποίες ο χρήστης μπορεί να

χρησιμοποιήσει, πατώντας το κουμπί με το σύμβολο της πρόσθεσης, για να διορθώσει ή να εξειδικεύσει την επισημείωση του. Επίσης, κάθε φορά που επιλέγεται ένας όρος εμφανίζεται εξωτερικός υπερσύνδεσμος προς τη σελίδα του στο Bioportal μέσω του οποίου παρουσιάζονται στο χρήστη περισσότερες λεπτομέρειες για τον όρο, όπως και μια οπτικοποίηση του όρου στην οντολογία ανατομίας.



Εικόνα 6.3.9: Δεύτερο βήμα επισημείωσης.

Στο παράδειγμα της Εικόνα 6.3.9 ο χρήστης έχει σημειώσει τα ανατομικά χαρακτηριστικά «δείκτης» και «αντίχειρας» για να περιγράψει τις Περιοχές Ενδιαφέροντος 1 και 2 αντίστοιχα. Παρατηρούμε ότι για τους παραπάνω όρους υπάρχουν εξωτερικοί σύνδεσμοι στο Bioportal, ενώ το σύννεφο όρων περιέχει όρους σχετικούς με τον όρο «αντίχειρας», ο οποίος είναι και ο τελευταίος που πληκτρολογήθηκε.

6.3.2.4 Επισημείωση: Βήμα Τρίτο

Στο τρίτο βήμα ο χρήστης έχει τη δυνατότητα να εισάγει τα χαρακτηριστικά ασθένειας ή τις σχετικές ανωμαλίες που παρατηρεί για κάθε Περιοχή Ενδιαφέροντος (Εικόνα 6.3.10). Όπως στο προηγούμενο βήμα, ο χρήστης ξεκινά να πληκτρολογεί έναν όρο και στη συνέχεια επιλέγει το κατάλληλο όρο από τους σχετικούς όρους της οντολογίας ασθένειας από λίστα που ενημερώνεται αυτόματα καθώς αυτός πληκτρολογεί. Στη συνέχεια, ο όρος αυτός

εμφανίζεται στο δέντρο της οντολογίας αριστερά. Ο χρήστης μπορεί να χρησιμοποιήσει κάποιον όρο από το δέντρο, πατώντας το κουμπί με το σύμβολο της πρόσθεσης, για να βελτιώσει την επισημείωση του. Για παράδειγμα, μπορεί να επιλέξει την κλάση-πατέρας του όρου που έχει εισάγει, ώστε να κάνει μια πιο γενική επισημείωση. Επίσης, όπως για τους ανατομικούς όρους του βήματος 2, κάθε φορά που ένας καινούργιος όρος εισάγεται ως επισημείωση, προστίθεται μία εξωτερική σύνδεση προς τη σελίδα του στο Bioportal.



Εικόνα 6.3.10: Τρίτο βήμα επισημείωσης.

Στο παράδειγμα της Εικόνα 6.3.10 ο χρήστης έχει σημειώσει τα χαρακτηριστικά «εξάρθρωση δακτύλου» και «δυσμορφία δακτύλου» για να περιγράψει τις ανωμαλίες που παρουσιάζονται στις Περιοχές Ενδιαφέροντος 1 και 2 αντίστοιχα. Στο δέντρο έχει ανιχνευτεί και επιλεγεί αυτόματα ο όρος «δυσμορφία δακτύλου», καθώς αυτός είναι ο τελευταίος που πληκτρολογήθηκε. Παρατηρούμε όλες τις υπερκλάσεις του όρου αυτού.

6.3.2.5 Επισημείωση: Βήμα Τέταρτο

Στο τέταρτο και τελευταίο βήμα παρέχεται στο χρήστη μια επισκόπηση των δεδομένων που έχει εισάγει και ζητείται επιβεβαίωση του χρήστη πριν τα δεδομένα αποθηκευτούν στο σύστημα (Εικόνα 6.3.11). Τα δεδομένα περιλαμβάνουν τόσο τα πεδία ενδιαφέροντος από την

επικεφαλίδα του DICOM αρχείου, που μπορούν να αλλαχθούν στο βήμα 1, όσο και τις επισημειώσεις των Περιοχών Ενδιαφέροντος που έχουν σημειωθεί στην εικόνα. Για την αποθήκευση των δεδομένων και τον τερματισμό της διαδικασίας επισημείωσης ο χρήστης αρκεί να πατήσει το κουμπί Save, αλλιώς μπορεί να γυρίσει στα προηγούμενα βήματα πατώντας το κουμπί Back. Επισημαίνεται ότι ο οδηγός παρέχει τη δυνατότητα επισθοδρόμησης σε όλα τα βήματα με το πάτημα του κουμπιού Back ή πατώντας το αντίστοιχο back button του περιηγητή ιστού.

Εικόνα 6.3.11: Τέταρτο βήμα επισημείωσης.

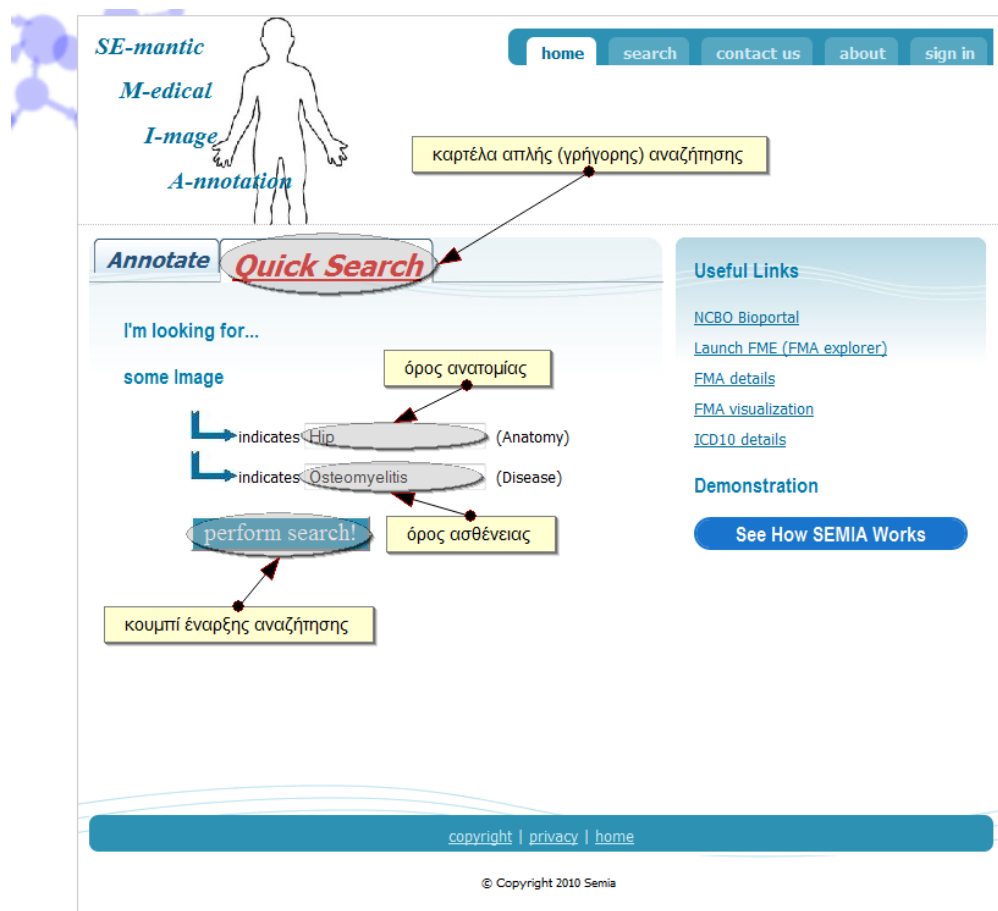
Με την αποθήκευση των δεδομένων επισημείωσης, ο χρήστης μεταφέρεται αυτόματα στη σελίδα παρουσίασης της εικόνας που μόλις έχει επισημάνει, η οποία θα παρουσιαστεί αναλυτικά στο σενάριο αναζήτησης εικόνων.

6.3.3 Αναζήτηση εικόνων

Ενώ η επισημείωση εικόνων απαιτείται η είσοδος του χρήστη στο σύστημα, η λειτουργία της αναζήτησης εικόνων είναι ελεύθερη για χρήση χωρίς πιστοποίηση. Υπάρχουν δύο τρόποι αναζήτησης εικόνων στο σύστημά μας: η απλή και η σύνθετη αναζήτηση.

6.3.3.1 Απλή αναζήτηση

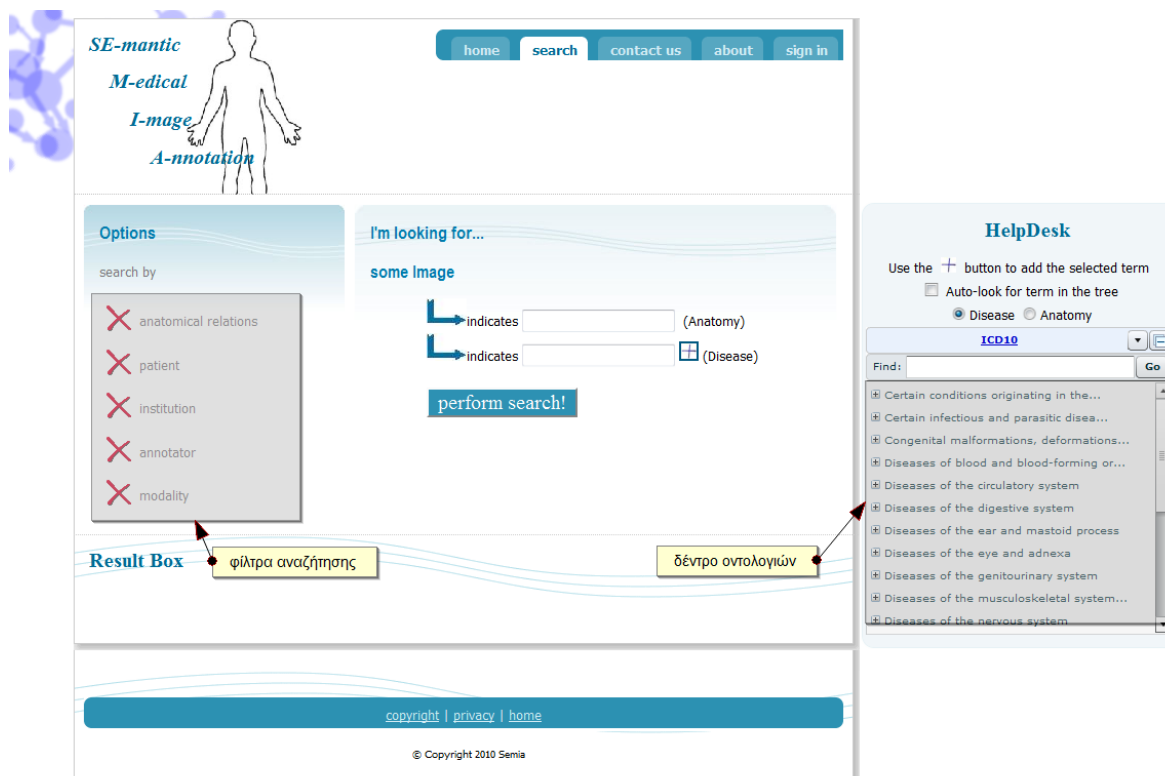
Στην κεντρική οθόνη της εφαρμογής ο χρήστης επιλέγει την καρτέλα «γρήγορη αναζήτηση» και στη συνέχεια εισάγει το ανατομικό χαρακτηριστικό και τον όρο της ασθένειας για τον οποίο επιθυμεί να αναζητήσει εικόνες. Καθώς πληκτρολογεί το όρο και στις δύο περιπτώσεις εμφανίζεται μια λίστα από όρους των οντολογιών που ενημερώνεται αυτόματα. Ο χρήστης επιλέγει τα επιθυμητά χαρακτηριστικά και πατάει το κουμπί Perform Search (Εικόνα 6.3.12). Στη συνέχεια μεταφέρεται αυτόματα στη σελίδα αναζήτησης (ή σύνθετης αναζήτησης) όπου του παρουσιάζονται τα αποτελέσματα της αναζήτησής του. Η σελίδα αυτή παρουσιάζεται στο σενάριο σύνθετης αναζήτησης. Παρατηρούμε δηλαδή ότι η απλή αναζήτηση είναι μια απλή και γρήγορη διεπαφή της σύνθετης αναζήτησης.



Εικόνα 6.3.12: Απλή αναζήτηση.

6.3.3.2 Σύνθετη αναζήτηση

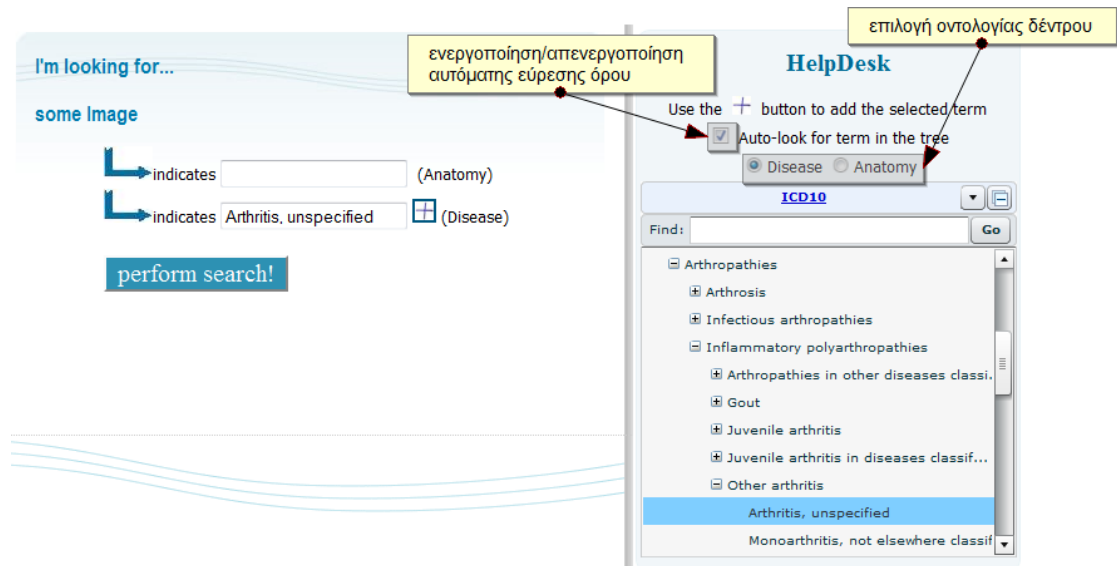
Στην οθόνη σύνθετης αναζήτησης () ο χρήστης έχει περισσότερες δυνατότητες σε σχέση με την απλή αναζήτηση, κυρίως σε ότι αφορά τα φίλτρα αναζήτησης. Πέρα από την εισαγωγή των όρων ανατομίας και ασθένειας με βάση τους οποίους αναζητά εικόνες, ο χρήστης μπορεί χρησιμοποιώντας τις επιλογές αριστερά να επιβάλλει φίλτρα στην αναζήτησή του. Δεξιά παρατηρούνται τα δέντρα των δύο οντολογιών – ασθένειας και ανατομίας.



Εικόνα 6.3.13: Σύνθετη αναζήτηση.

Στο δέντρο οντολογιών μπορεί να αλλάξει η οντολογία που απεικονίζεται από την επιλογή του radiobutton πάνω από αυτό. Επίσης μπορεί να ρυθμιστεί να θα ανιχνευτεί αυτόματα ο όρος που πληκτρολογείται από τον χρήστη στο δέντρο ή όχι με τη χρήση του checkbox (Εικόνα 6.3.14). Από προεπιλογή η λειτουργικότητα αυτή είναι απενεργοποιημένη και η οντολογία που φορτώνεται στο δέντρο είναι αυτή των ασθενειών.

Στην Εικόνα 6.3.14 παρατηρούμε ότι ο χρήστης έχει ενεργοποιημένη την αυτόματη εύρεση των όρων και έχει επιλέξει την οντολογία ασθενειών να απεικονιστεί στο δέντρο. Όταν επιλέγει τον όρο «αρθρίτιδα», αυτός εμφανίζεται στο δέντρο.



Εικόνα 6.3.14: Δέντρο οντολογιών.

Ο χρήστης μπορεί να επιλέξει να φιλτράρει τα αποτελέσματα με βάση τον ασθενή, το ινστιτούτο στο οποίο παράχθηκε η εικόνα, τον επιστήμονα που έκανε την επισημείωση, την τροπικότητα της εικόνας και τις ανατομικές σχέσεις. Αρκεί να πατήσει πάνω στο επιθυμητό φίλτρο για να του παρουσιαστούν οι επιλογές του.

Στο φίλτρο με βάση τον ασθενή, όπως φαίνεται στην Εικόνα 6.3.15, ο χρήστης μπορεί να επιλέξει το φύλο των ασθενών στους οποίους ανήκουν οι εικόνες που θα επιστραφούν, και το εύρος της ηλικίας των ασθενών. Στην περίπτωση που εικονίζεται, ο χρήστης αναζητά εικόνες ασθενών που είναι άντρες ηλικίας μεταξύ 11 και 50 ετών.



Εικόνα 6.3.15: Φίλτρο με βάση τον ασθενή.

Στο φίλτρο με βάση το ινστιτούτο, ο χρήστης μπορεί να επιλέξει να φιλτράρει τα αποτελέσματα της αναζήτησης με βάση ένα συγκεκριμένο ινστιτούτο που επιλέγει από λίστα που περιέχει όλα τα ινστιτούτα καταχωρημένα στο σύστημα (Εικόνα 6.3.16).

Εικόνα 6.3.16: Φίλτρο με βάση το ινστιτούτο.

Στο φίλτρο με βάση το άτομο που έκανε την επισημείωση, ο χρήστης μπορεί να ζητήσει εικόνες που έχουν επισημειωθεί από επιστήμονες με συγκεκριμένο όνομα ή/και ειδικότητα (Εικόνα 6.3.17).

Εικόνα 6.3.17: Φίλτρο με βάση τον επισημειωτή (annotator).

Στο φίλτρο τροπικότητας, ο χρήστης μπορεί να ζητήσει εικόνες συγκεκριμένης τροπικότητας και μόνο, για παράδειγμα μόνο εικόνες CT (Εικόνα 6.3.18).

Options

search by

- anatomical relations
- patient
- institution
- annotator
- modality

I'm looking for...

some Image

- indicates (Anatomy)
- indicates (Disease)
- hasModality

perform search!

Εικόνα 6.3.18: Φίλτρο με βάση την τροπικότητα.

Τελευταίο εξετάζεται το φίλτρο ανατομικών χαρακτηριστικών, το οποίο διαφέρει από τα υπόλοιπα. Το φίλτρο αυτό έχει νόημα μόνο όταν ο χρήστης αναζητά με βάση κάποιον ανατομικό όρο. Αν είναι ενεργοποιημένο, με την εισαγωγή του ανατομικού χαρακτηριστικού από το χρήστη το σύστημα αναζητεί ποιες από τις σχέσεις «υποκλάση», «χωρικό μέρος», «συστατικό μέρος» και «μέρος συστήματος» έχουν νόημα για τον όρο αυτόν. Δηλαδή, αν ανιχνευτεί ότι υπάρχουν όροι στην οντολογία ανατομίας που είναι υποκλάσεις του όρου ανατομίας που έχει εισαχθεί, τότε η σχέση «υποκλάση» επιστρέφεται. Οι ευρεθείσες σχέσεις τοποθετούνται σε λίστα από την οποία ο χρήστης μπορεί να επιλέξει με βάση ποια σχέση το σύστημα να αναζητήσει ανατομικά χαρακτηριστικά (Εικόνα 6.3.19). Ως προεπιλογή και αν το φίλτρο αυτό είναι απενεργοποιημένο το σύστημα ψάχνει για τον ανατομικό όρο που εισήχθη και για όλους τους όρους οι οποίοι συνδέονται μέσω κάποιας από τις παραπάνω τέσσερις σχέσεις με τον όρο αναζήτησης.

Options

search by

- anatomical relations
- patient
- institution
- annotator
- modality

I'm looking for...

some Image

- indicates (Anatomy)
 - and its
- indicates (Disease)

perform search!

Εικόνα 6.3.19: Φίλτρο με βάση τις ανατομικές σχέσεις.

Ασφαλώς επιτρέπεται η χρήση περισσότερων από ένα φίλτρα ταυτόχρονα για το καλύτερο φιλτράρισμα των αποτελεσμάτων.

Όταν ο χρήστης εισάγει τα δεδομένα αναζήτησης και επιλέξει όλα τα επιθυμητά φίλτρα πατά το κουμπί Perform Search για να ξεκινήσει η αναζήτηση. Τα αποτελέσματα της αναζήτησης εμφανίζονται στο κάτω μέρος της σελίδας σε μορφή λίστας. Για κάθε εικόνα παρουσιάζεται ένα thumbnail και όλες οι πληροφορίες τις, δηλαδή οι τιμές των DICOM πεδίων, όπως έχουν εισαχθεί στο σύστημα, και τα χαρακτηριστικά ανατομίας και ασθένειας, όπως έχουν εισαχθεί για κάθε Περιοχή Ενδιαφέροντος της εικόνας.

Εάν ο χρήστης έχει επιχειρήσει αναζήτηση με βάση ένα ανατομικό όρο, τότε τα αποτελέσματα κατηγοριοποιούνται με βάση την κλάση ανατομίας στην οποία ανήκουν (Εικόνα 6.3.20). Κατάλληλο μήνυμα εμφανίζεται ακριβώς πάνω από τη λίστα των εικόνων που ενημερώνει το χρήστη για το ποιες κλάσεις βρέθηκαν και μέσω υπερσυνδέσεων προσφέρει άμεση μετάβαση στον επιθυμητό όρο, ώστε να αποφευχθεί η χρονοβόρα κύλιση σε περίπτωση πολλών αποτελεσμάτων. Επίσης, κάτω δεξιά το σύστημα ενημερώνει το χρήστη σχετικά με τους ανατομικούς όρους με βάση τους οποίους έγινε η αναζήτηση και με έντονη γραφή σημειώνονται οι κλάσεις που βρέθηκαν να περιέχουν αποτελέσματα.

The screenshot displays a search interface with several key components:

- Options:** A sidebar on the left with search filters, all of which are currently disabled (indicated by red 'X' marks): anatomical relations, patient, institution, annotator, and modality.
- I'm looking for...:** A search input area where the user has entered 'some image'. Below the input, there are two dropdown menus: one for 'Anatomy' (with 'hip' selected) and one for 'Disease'. A 'perform search!' button is located below these menus.
- Result Box:** A section below the search area containing a message: 'Found images indicating Hip bone'. A yellow callout box labeled 'ενημερωτικό μήνυμα' points to this message.
- Hip bone:** A detailed view of the search result, including a thumbnail of a hip X-ray and metadata: 'anatomy: Hip bone', 'disease: Osteomyelitis', 'modality: OT', 'original name: OT-MONO2-8-hip.dcm', 'institution: LAIKO', 'patient's age: 50', 'patient's sex: female', and 'annotators: Ilias Kyrgiazos [Radiologist]'. A yellow callout box labeled 'λίστα αποτελεσμάτων' points to this section.
- HelpDesk:** A sidebar on the right titled 'HelpDesk' that provides a hierarchical list of related classes for the searched term 'hip'. The list includes:
 - Multiple fractures of metacarpal bones
 - Injury of blood vessels at wrist and h...
 - Injury of muscle and tendon at wrist a...
 - Injury of nerves at wrist and hand level
 - Open wound of wrist and hand
 - Other and unspecified injuries of wis...
 - Superficial injury of wrist and hand
 - Traumatic amputation of wrist and hand
 - Injuries to unspecified part of trunk,...
 - Other and unspecified effects of exter...
 - Poisoning by drugs, medicaments and bi...
- Related Classes:** A table below the HelpDesk sidebar titled 'Searched for 'hip', whose related classes are:'. The table lists related classes under different categories:

| | |
|----------------------|--|
| subclasses | <ul style="list-style-type: none"> Right hip Left hip |
| regional parts | - |
| constitutional parts | <ul style="list-style-type: none"> Vasculature of hip Skin of hip Musculature of hip Hip joint Superficial fascia of hip Hip bone |
| systemic parts | - |

Εικόνα 6.3.20: Αποτελέσματα αναζήτησης με βάση ανατομικό όρο.

Στην Εικόνα 6.3.20 παρατηρούμε ότι ο χρήστης επιχειρήσει αναζήτηση με βάση τον ανατομικό όρο «γοφόρς» και το σύστημα επέστρεψε εικόνα που περιέχει Περιοχή Ενδιαφέροντος με ανατομική επισημείωση «οστό του γοφού», αφού αυτό είναι συστατικό μέρος του γοφού.

Εάν ο χρήστης δεν επιχειρήσει αναζήτηση με βάση ανατομικό όρο, τότε τα αποτελέσματα δεν κατηγοριοποιούνται, όπως φαίνεται στην Εικόνα 6.3.21, στην οποία ο χρήστης έχει επιχειρήσει αναζήτηση με βάση την ασθένεια «πολλαπλό κάταγμα μετακαρπίου οστού».

The screenshot displays a search interface with the following components:

- Options:** A list of search criteria with red 'X' marks indicating they are not selected: anatomical relations, patient, institution, annotator, and modality.
- I'm looking for...:** A search box containing 'some Image'. Below it, two arrows point to input fields: the first is empty and labeled '(Anatomy)', the second contains 'Multiple fractures of met:' and is labeled '(Disease)'. A 'perform search!' button is located below these fields.
- Result Box:** A section titled 'Searched for images with any anatomical characteristic.' containing two results:
 - Result 3:** Includes an X-ray image of a hand, anatomy: Hinge joint - Surface of metacarpus, disease: Dislocation of finger - Multiple fractures of metacarpal bones, modality: CR, original name: 6WF0WTH8.dcm, referring physicians: ZNA OUCH, body part examined: HAND, study description: Hand, institution: Klinikum Saarbruecken, patient's age: 27, patient's sex: male, and annotators: Ilias Kyrgiazos [Radiologist].
 - Result 2:** Includes an X-ray image of a hand, anatomy: Hinge joint - Surface of metacarpus, disease: Arthritis, unspecified - Multiple fractures of metacarpal bones, modality: CR, original name: 4_JAKLSA.dcm, referring physicians: ZNA OUCH, body part examined: HAND, study description: Hand, institution: Klinikum Saarbruecken, patient's age: 27, patient's sex: male, and annotators: Ilias Kyrgiazos [Radiologist].
- HelpDesk:** A sidebar with a search bar containing 'ICD10' and a 'Go' button. Below the search bar is a list of ICD10 terms, with 'Multiple fractures of metacarpal bones' highlighted.

Εικόνα 6.3.21: Αποτελέσματα αναζήτησης με βάση όρο ασθένειας.

Ασφαλώς, ο χρήστης μπορεί να επιχειρήσει συνδυαστική αναζήτηση με βάση και ανατομικό όρο και όρο ασθένειας. Τότε, το σύστημα αναζητά εικόνες που να πληρούν και τα δύο κριτήρια.

Από τη λίστα των αποτελεσμάτων ο χρήστης μπορεί να επιλέξει την επιθυμητή εικόνα πατώντας πάνω στη μικρή εικόνα (thumbnail), όπως φαίνεται στην Εικόνα 6.3.22.



Εικόνα 6.3.22: Επιλογή εικόνας από λίστα.

Με την επιλογή της εικόνας, αυτή παρουσιάζεται στον οθόνη παρουσίασης (Εικόνα 6.3.23). Εδώ εμφανίζονται στον χρήστη η εικόνα στο φυσικό της μέγεθος με τις πληροφορίες τις στα δεξιά. Ο χρήστης μπορεί με κύλιση του ποντικιού όσο αυτό βρίσκεται μέσα στην περιοχή της εικόνας ή χρησιμοποιώντας την μπάρα που βρίσκεται ακριβώς πάνω από την εικόνα να την μεγεθύνει ή να την σμικρύνει. Επίσης μπορεί πατώντας το κουμπί Extract Content να εξάγει την σημασιολογική πληροφορία σε μορφή RDF/XML, καθώς και την ίδια την εικόνα σε μορφή jpeg, με τις Περιοχές Ενδιαφέροντος σημειωμένες πάνω σε αυτήν. Ακόμη, μπορεί πατώντας το κουμπί quick search, που βρίσκεται δίπλα σε κάθε ROI, να επιχειρήσει γρήγορη αναζήτηση με βάση τους όρους ανατομίας και ασθένειας του συγκεκριμένου ROI.

Κάτω από την εικόνα εμφανίζεται μια λίστα από παρόμοιες εικόνες, οι οποίες είναι εικόνες που περιέχουν Περιοχές Ενδιαφέροντος με ίδιες ανατομικές επισημειώσεις ή/και ίδιες επισημειώσεις ασθενειών. Ο χρήστης τοποθετώντας το ποντίκι πάνω από κάποια από τις εικόνες τις λίστας μπορεί να ενημερωθεί για τους κοινούς όρους της συγκεκριμένης με την κεντρική εικόνα, όπως φαίνεται στην Εικόνα 6.3.24. Οι εικόνες στη λίστα είναι ταξινομημένες με βάση το βαθμό ομοιότητας με τις εικόνες που έχουν περισσότερους κοινούς όρους επισημείωσης να βρίσκονται αριστερότερα.

SE-mantic
M-edical
I-mage
A-notation

home search contact us about sign in

εξαγωγή περιεχομένου

Help Desk
extract content

γρήγορη αναζήτηση

Image Info

ROI1 quick search
• Anatomy: Surface of metacarpus
• Disease: Multiple fractures of metacarpal bones

ROI2 quick search
• Anatomy: Hinge joint
• Disease: Dislocation of finger

modality
CR

original name
6WF0WTH8.dcm

referring physicians
ZNA OUCH

body part examined
HAND

study description
Hand

institution
Klinikum Saarbruecken

patient's age
27

patient's sex
male

annotators
Ilias Kyrgiazos [Radiologist]

Similar Images

λίστα παρόμοιων εικόνων

Εικόνα 6.3.23: Παρουσίαση εικόνας.

Similar Images



Hinge joint - Surface of metacarpus - Multiple fractures of metacarpal bones

Εικόνα 6.3.24: Παρόμοιες εικόνες –τοποθέτηση ποντικιού πάνω από εικόνα.

7

Συμπεράσματα

Στο κεφάλαιο αυτό παρουσιάζονται τα αποτελέσματα της παρούσας διπλωματικής εργασίας, οι περιορισμοί και οι προκλήσεις που συναντήθηκαν καθώς και προτείνονται ορισμένες μελλοντικές επεκτάσεις του συστήματος που αναπτύχθηκε, οι οποίες κρίθηκαν άξιες ενδιαφέροντος.

7.1 Αποτελέσματα

Η παρούσα διπλωματική εργασία είχε ως αντικείμενο την ανάπτυξη ενός διαδικτυακού συστήματος σημασιολογικής περιγραφής ιατρικών απεικονιστικών δεδομένων. Ένα βασικό θετικό αποτέλεσμα της χρήσης του υλοποιημένου συστήματος είναι η ενιαία και κοινή αντιμετώπιση των δεδομένων αυτών, η οποία είναι ανεξάρτητη του είδους (τροπικότητας), των υφιστάμενων προτύπων και του περιεχομένου απεικόνισης. Με τον τρόπο αυτό μπορεί να προωθηθεί η διασύνδεση και η διαλειτουργικότητα των διαφόρων φορέων υγείας που διαχειρίζονται ιατρικές εικόνες. Η υιοθέτηση της ανοικτής, διαδικτυακής λύσης που προτείνεται εκμεταλλεύεται την υποδομή του υπάρχοντος Ιστού και τις διαδικτυακές υπηρεσίες που αυτός προσφέρει. Στο σημείο αυτό είναι σημαντικό να σημειωθεί ότι, πέρα από τη χρήση των υπηρεσιών για πρόσβαση στις οντολογίες αναφοράς, το ίδιο το σύστημα που αναπτύχθηκε προσφέρει πρόσβαση μέσω ανοικτής υπηρεσίας τύπου REST, μέσω της οποίας επιτρέπεται η πραγματοποίηση ερωτημάτων και ανάκτηση γνώσης από την αποθήκη τριάδων του συστήματος.

Ένα άλλο σημαντικό αποτέλεσμα είναι η απόδειξη της δυνατότητας χρησιμοποίησης της υπάρχουσας κωδικοποιημένης γνώσης. Το σύστημα που αναπτύχθηκε χρησιμοποιεί τις οντολογίες Foundational Model of Anatomy και ICD-10, έχοντας απομακρυσμένη πρόσβαση σε αυτές. Η επαναχρησιμοποίηση είναι άλλωστε ένας από τους βασικούς σκοπούς σχεδιασμού τέτοιων μεγάλων οντολογιών αναφοράς. Η σύνδεση της τοπικά αποκτώμενης γνώσης με ένα ευρύτερο σημασιολογικό δίκτυο μας φέρνει ένα βήμα πιο κοντά στην πραγμάτωση του οράματος του Σημασιολογικού Ιστού. Η ευελιξία και η επεκτασιμότητα, τέλος, που προσφέρει η απομακρυσμένη πρόσβαση στις οντολογίες αναφοράς επιτρέπει την αυτόματη ενημέρωση του συστήματος με τις τελευταίες εκδόσεις των οντολογιών, καθώς επίσης και την αντικατάστασή τους ή την προσθήκη νέων οντολογιών στο σύστημα.

Ένα από τα κυριότερα και πιο απτά αποτελέσματα είναι η «έξυπνη» αναζήτηση που εκμεταλλεύεται τη σημασιολογική φύση των επισημειώσεων. Όταν γίνεται αναζήτηση εικόνων με βάση έναν ανατομικό όρο, το σύστημα επεκτείνει το ερώτημα, συμπεριλαμβάνοντας στους όρους αναζήτησης όλους τους ανατομικούς όρους που βρίσκονται στη «σημασιολογική γειτονιά» του αρχικού όρου αναζήτησης, με βάση την οντολογία ανατομίας. Αντίστοιχα, όταν η αναζήτηση γίνεται με βάση κάποιον ανατομικό όρο, το σύστημα αναζητεί εικόνες με βάση τον συγκεκριμένο όρο, αλλά και όλες τις υποκλάσεις του όρου αυτού, με βάση την ταξινόμηση των ασθενειών στην ορολογία ασθενειών. Με τον τρόπο αυτό, οι εικόνες συνδέονται μεταξύ τους σύμφωνα με τους όρους επισημείωσης και μπορούν να ανακτηθούν κατά σημασιολογικά συγγενείς ομάδες.

Η σημασιολογική επέκταση των ερωτημάτων αλλά και η επιβολή φίλτρων για πιο ειδική αναζήτηση φανερώνουν τα πλεονεκτήματα στην αναζήτηση που προσφέρει το προτεινόμενο σύστημα σε σχέση με υφιστάμενα λειτουργικά συστήματα αναζήτησης εικόνας όπως τα PACS. Η αναζήτηση, για παράδειγμα, για εικόνες ενός συγκεκριμένου ασθενούς που καταδεικνύουν «τενοντίτιδα του δακτύλου του δείκτη» πραγματοποιείται σε δύο βήματα από τα υπάρχοντα συστήματα. Στο πρώτο βήμα το Ακτινολογικό Πληροφοριακό Σύστημα, στο οποίο κρατούνται όλα τα ακτινολογικά ευρήματα του ασθενούς, πρέπει να ερωτηθεί και περιγραφές «τενοντίτιδας του δακτύλου του δείκτη» πρέπει να ανιχνευτούν στις σχετικές αναφορές. Ως δεύτερο βήμα, ο ακτινολόγος πρέπει να επιλέξει την ανάκτηση μόνο των επιθυμητών εικόνων από τις εικόνες που αναφέρονται στα ακτινολογικά ευρήματα. Η διαδικασία αυτή είναι όχι μόνο πολύπλοκη, αλλά έχει συχνά ανακριβή ανάκληση πληροφορίας. Στο σύστημα που αναπτύχθηκε, η συγκεκριμένη αναζήτηση γίνεται σε ένα μόνο βήμα, με τον ακτινολόγο να επιλέγει ανατομικό χαρακτηριστικό (δάκτυλο του δείκτη), χαρακτηριστικό ασθένειας (τενοντίτιδα) και φίλτρο με βάση το όνομα του ασθενούς.

Το υλοποιημένο σύστημα δεν έχει στόχο να αντικαταστήσει τα ευρέως χρησιμοποιούμενα ιατρικά απεικονιστικά συστήματα εισαγωγής και αναζήτησης (PACS), αλλά να τα βελτιώσει,

προσφέροντας μια πιο λειτουργική διεπαφή χρήστη. Μέσω της διεπαφής αυτής ο επιστήμονας-ακτινολόγος θα μπορεί να καταγράφει τα χαρακτηριστικά που παρατηρεί κατά την επανεξέταση της εικόνας και στη συνέχεια να αναζητεί εικόνες με βάση τα χαρακτηριστικά αυτά. Η ενσωμάτωση του υλοποιημένου συστήματος σε ένα ευρύτερο πληροφοριακό σύστημα υγείας παρουσιάζεται ως απώτερος στόχος. Η καταγραφή των χαρακτηριστικών με βάση πρότυπα του παγκόσμιου Ιστού, όπως το RDF μοντέλο, και η δημοσίευσή τους στο διαδίκτυο θα επιτρέψει, επίσης, τη χρησιμοποίηση των επισημειωμένων δεδομένων για ερευνητικούς σκοπούς.

Μια ακόμη δυνατότητα χρήσης του υλοποιημένου συστήματος είναι για εκπαιδευτικούς σκοπούς. Εάν χρησιμοποιηθεί ευρέως από την ιατρική κοινότητα για την περιγραφή ιατρικών εικόνων, το σύστημα θα διαθέτει πληροφορία αρκετή ώστε να λειτουργήσει ως διαδικτυακή πύλη εκπαίδευσης νέων επιστημόνων. Για παράδειγμα, νέοι ακτινολόγοι θα μπορούν να πλοηγούνται στο σύστημα και να έρχονται σε επαφή με απεικονιστικά δεδομένα που έχουν περιγραφεί από έγκριτους επιστήμονες του κλάδου τους.

7.2 Περιορισμοί – Προκλήσεις

Κατά τη φάση του σχεδιασμού και της υλοποίησης του συστήματος τέθηκαν από εξωτερικές πηγές διάφοροι περιορισμοί. Ορισμένοι από αυτούς ξεπεράστηκαν εύκολα και άλλοι επέβαλλαν διαφορετική προσέγγιση για να παρακαμφθούν και αντιμετωπίστηκαν κυρίως ως προκλήσεις. Στη συνέχεια παρατίθενται περιορισμοί - προκλήσεις που χρίζουν ιδιαίτερης προσοχής, συνοδευόμενοι από τον τρόπο που αντιμετωπίστηκαν ή τον προτεινόμενο τρόπο αντιμετώπισής τους.

Αρχική σκέψη κατά τη φάση του σχεδιασμού ήταν να ενσωματωθούν οι οντολογίες αναφοράς στο σύστημα, γιατί κατ' αυτό τον τρόπο το σύστημα είναι αυτόνομο και δεν εξαρτάται από άλλους εξωτερικούς παράγοντες. Εντούτοις, κάτι τέτοιο ήταν απαγορευτικό, εξαιτίας του μεγάλου μεγέθους των οντολογιών και ιδιαίτερα της FMA. Οι ανάγκες ενός υπολογιστικού συστήματος στο οποίο είναι αποθηκευμένη η FMA ξεπερνούν τις ανάγκες της συγκεκριμένης διπλωματικής εργασίας. Η πρόκληση αυτή οδήγησε στην χρήση των διαδικτυακών υπηρεσιών που καθιστούν τις FMA και ICD-10 διαθέσιμες και προσβάσιμες. Η τακτική αυτή είναι και η ενδεδειγμένη, σύμφωνα και με αναφορές σχεδιασμού άλλων παρόμοιων συστημάτων [36]. Έτσι, τελικά, το σύστημα σχεδιάστηκε κατά το υβριδικό αρχιτεκτονικό πλαίσιο που περιγράφεται στην παράγραφο 5.2. Το μόνο μειονέκτημα αυτής της τακτικής είναι ότι η άρτια λειτουργία του συστήματος εξαρτάται άμεσα από εξωτερικές πηγές, οι οποίες αν μη τι άλλο επιβάλλεται να είναι αξιόπιστες. Ωστόσο αυτό το μειονέκτημα λήφθηκε υπόψη κατά τη φάση της υλοποίησης, με αποτέλεσμα οι λειτουργικότητες των

διαδικτυακών αυτών υπηρεσιών να ενεργούν συμπληρωματικά στον βασικό κορμό του συστήματος.

Ένας άλλος σημαντικός περιορισμός που πρέπει να υπογραμμισθεί είναι η έλλειψη αντιστοιχίας μεταξύ ανατομίας και ασθενειών στον κόσμο των οντολογιών, με πεδίο αναφοράς την ιατρική. Η ύπαρξη μίας τέτοιας ενοποιημένης οντολογίας αναφοράς θα επέτρεπε την προσθήκη στο υλοποιημένο σύστημα μίας επιπλέον λειτουργικότητας, αυτή των πιθανών ασθενειών. Ο κατάλογος των ασθενειών που αντιστοιχούν σε κάθε ανατομικό όρο, θα βοηθούσε τους ειδικούς να πραγματοποιήσουν «διαφορική διάγνωση», με βάση τις ιατρικές εικόνες. Το συγκεκριμένο σύστημα θα μπορούσε να θεωρηθεί ως ένα σύστημα εκμάθησης της σύνδεσης των ασθενειών σε ανατομικούς όρους, καθώς επισημειώσεις ανατομίας και ασθενειών αποθηκεύονται μαζί και είναι διασυνδεδεμένοι μέσω του ROI (Περιοχή Ενδιαφέροντος). Σκόπιμα, όμως, δεν παρουσιάζεται αυτή η γνώση στο χρήστη, δεδομένου ότι η εν λόγω διασύνδεση δεν είναι προτυποποιημένη από ειδικούς του χώρου. Παρόλα αυτά, η εξαγόμενη γνώση της διασύνδεσης από την οντολογία εφαρμογής θα μπορούσε να αποτελέσει την κύρια βάση γνώσης προς επεξεργασία από ειδικούς του ευρύτερου τομέα της ιατρικής. Έτσι, θα μπου οι θεμέλιοι λίθοι για τη δημιουργία, με τη βοήθεια και καθοδήγηση ειδικών, μίας ενοποιημένης οντολογίας αναφοράς, ή εναλλακτικά, θα μπορούσε να δημιουργηθεί ένας μηχανισμός διασύνδεσης ανατομίας και ασθενειών.

Κυριότερος περιορισμός όλων, αποτέλεσε η απουσία της ιατρικής ορολογίας στα ελληνικά από τις οντολογίες αναφοράς. Κάτι τέτοιο θα βοηθούσε το σύστημα να είναι διαθέσιμο και στα ελληνικά, και να αποτελέσει, συνάμα, σύστημα εκμάθησης της ιατρικής ορολογίας στα αγγλικά αλλά και σε άλλες γλώσσες. Η ανάγκη της ύπαρξης οντολογιών αναφοράς στα ελληνικά, ειδικά στον πεδίο της βιολογικής, έχει επισημανθεί και σε άλλα αξιόλογα έργα [37]. Σε αυτό το σημείο αξίζει να σημειωθεί ότι η FMA έχει ενσωματώσει σε αρκετούς ανατομικούς όρους την ορολογία τους και σε άλλες γλώσσες, πέραν των αγγλικών. Συγκεκριμένα, η παρουσίαση της γνώσης των οντολογιών αναφοράς στα ελληνικά θα μπορούσε να γίνει με δύο κυρίως τρόπους: (1) με τη χρήση «ετικέτας» (label) για κάθε όρο των οντολογιών αναφοράς, πρακτική που είναι διαδεδομένη στις τεχνολογίες των οντολογιών και έχει ήδη συμπεριληφθεί από την W3C στο «λεξιλόγιο RDFS» (RDFS vocabulary) ως `rdfs:label` και (2) με τη χρησιμοποίηση διαδικτυακών υπηρεσιών, οι οποίες θα τηρούν σε αρχεία την αντιστοιχία όρου οντολογίας (URI) – ετικέτας, εκφρασμένη στα ελληνικά.

7.3 Μελλοντικές Επεκτάσεις

Πιθανές μελλοντικές επεκτάσεις του υλοποιημένου συστήματος στα πλαίσια της παρούσας διπλωματικής είναι οι εξής:

- Τροποποίηση των επισημειώσεων. Στο παρόν σύστημα δε δίνεται η δυνατότητα για αλλαγή ή διαγραφή των επισημειώσεων, πράγμα δεσμευτικό για το χρήστη. Είναι σκόπιμο να συμπεριληφθεί αυτή η λειτουργικότητα σε μελλοντική έκδοση του συστήματος, καθώς είναι απαραίτητη για τη διόρθωση λαθών ή ασυνεπειών κατά την επισημείωση. Θα μπορούσε επίσης να καθιερωθεί ένα σύστημα αξιολόγησης των επισημειώσεων από επιστημονική επιτροπή, το οποίο θα κατοχύρωνε την αξιοπιστία των εισαχθέντων στοιχείων.
- Αυτόματη επισημείωση, ώστε να επιταχυνθεί η διαδικασία περιγραφής των εικόνων. Ένας αλγόριθμος αναγνώρισης προτύπων και αυτόματης τμηματοποίησης της εικόνας θα χρησίμευε στην αυτόματη επιλογή των Περιοχών Ενδιαφέροντος στην εικόνα. Στη συνέχεια και με βάση κριτήρια ομοιότητας με τμήματα εικόνων που υπάρχουν ήδη στη βάση γνώσης, ο αλγόριθμος θα προτείνει κατάλληλους όρους από τις οντολογίες, ώστε να διευκολυνθεί η διαδικασία επισημείωσης.
- Σύνδεση με άλλες πηγές γνώσης, όπως η DBpedia. Οι όροι περιγραφής μπορούν να συνδεθούν με εξωτερικές σημασιολογικές πηγές γνώσεις όπως η DBpedia - μια προσπάθεια για εξαγωγή σε δομημένη μορφή της πληροφορίας που περιέχει η Wikipedia - ώστε να παρέχεται αυτόματη μετάβαση μέσω της διεπαφής του υλοποιημένου συστήματος σε τοποθεσίες σχετικές με τους όρους περιγραφής. Αυτού του είδους τα Συνδεδεμένα Δεδομένα (Linked Data) αποτελούν άλλωστε και τη γέφυρα για τη μετάβαση στο Σημασιολογικό Ιστό.
- Σύνδεση με Ηλεκτρονικό Ιατρικό Φάκελο. Μια μελλοντική ενσωμάτωση του υλοποιημένου συστήματος σε ένα Νοσοκομειακό Πληροφοριακό Σύστημα και η σύνδεσή του με τεχνολογίες Ηλεκτρονικού Ιατρικού Φακέλου θα προσφέρει τη δυνατότητα για μια προσωποκεντρική προσέγγιση στη διαχείριση των απεικονιστικών δεδομένων και της περιγραφής τους. Με τον τρόπο αυτό, όλες οι εξετάσεις ενός ασθενούς, μαζί με τη σημασιολογική περιγραφή τους και τις κλινικές αναφορές θα μεταχειρίζονται με ενιαίο τρόπο.

8

Αναφορές

- [1] G. Antoniou and F. Van Harmelen, *Εισαγωγή στον Σημασιολογικό Ιστό*, 2nd ed. Κλειδάριθμος, 2009.
- [2] “Semantic Web - Wikipedia, the free encyclopedia,” 11-Oct-2010. [Online]. Available: http://en.wikipedia.org/wiki/Semantic_Web.
- [3] T. Berners-Lee, *Weaving the Web*. HarperBusiness, 1999.
- [4] D. Rebholz-Schuhmann and G. Nenadic, “Biomedical Semantics: the Hub for Biomedical Research 2.0,” *Journal of Biomedical Semantics*, vol. 1, p. 1, 2010.
- [5] M. Moeller and M. Sintek, “A Generic Framework for Semantic Medical Image Retrieval,” in *Proceedings of the Knowledge Acquisition from Multimedia Content (KAMC) Workshop held at the 2nd international conference on Semantics And digital Media Technologies (SAMT)*, 2007.
- [6] D. Rubin, P. Mongkolwat, V. Kleper, K. Supekar, and D. Channin, “Medical Imaging on the Semantic Web: Annotation and Image Markup,” presented at the Association for the Advancement of Artificial intelligence (AAAI) Spring Symposium Series, Stanford, CA, 2008.
- [7] W. Wei and P. Barnaghi, “Semantic Support for Medical Image Search and Retrieval,” in *Proceedings of the fifth International Association of Science and Technology for Development (IASTED) International Conference: biomedical engineering*, pp. 315 -

319, 2007.

- [8] X. Zhou, S. Zillner, M. Moeller, Y. Zhan, A. Krishnan, and A. Gupta, “Semantics and CBIR: A Medical Imaging Perspective,” in *Proceedings of the 2008 international conference on Content-based image and video retrieval*, pp. 571-580, 2008.
- [9] J. Brinkley, D. Suci, L. Detwiler, J. Gennari, and C. Rosse, “A framework for using reference ontologies as a foundation for the semantic web,” in *American Medical Informatics Association (AMIA) Annual Symposium Proceedings*, pp. 96–100, 2006.
- [10] M. Shaw, L. Detwiler, J. Brinkley, and D. Suci, “Generating Application Ontologies from Reference Ontologies,” in *American Medical Informatics Association (AMIA) Annual Symposium Proceedings*, pp. 672–676, 2008.
- [11] F. Manola and E. Miller, “RDF Primer,” 10-Feb-2004. [Online]. Available: <http://www.w3.org/TR/rdf-primer/>.
- [12] D. Beckett, “RDF/XML Syntax Specification (Revised),” 10-Feb-2004. [Online]. Available: <http://www.w3.org/TR/REC-rdf-syntax/>.
- [13] E. Prub'hommeaux and A. Seaborne, “SPARQL Query Language for RDF,” 15-Jan-2008. [Online]. Available: <http://www.w3.org/TR/rdf-sparql-query/>.
- [14] D. Brickley and R. Guha, “RDF Vocabulary Description Language 1.0: RDF Schema,” 10-Feb-2004. [Online]. Available: <http://www.w3.org/TR/rdf-schema/>.
- [15] I. Herman, “W3C Semantic Web FAQ,” 12-Nov-2009. [Online]. Available: <http://www.w3.org/RDF/FAQ>.
- [16] D. McGuinness and F. Van Harmelen, “OWL Web Ontology Language Overview,” 10-Feb-2004. [Online]. Available: <http://www.w3.org/TR/owl-features/>.
- [17] C. Lutz, “Description Logics.” [Online]. Available: <http://dl.kr.org/>.
- [18] W3C OWL Working Group, “OWL 2 Web Ontology Language Document Overview,” 27-Oct-2009. [Online]. Available: <http://www.w3.org/TR/owl2-overview/>.
- [19] W3C OWL Working Group, “OWL 2 Web Ontology Language Document Overview - Draft,” 27-Mar-2009. [Online]. Available: <http://www.w3.org/TR/2009/WD-owl2-overview-20090327/>.
- [20] F. Baader, D. Calvanese, D. McGuinness, D. Nardi, and P. Patel-Schneider, *The Description Logic Handbook*. Cambridge University Press, 2003.
- [21] “Semantic reasoner - Wikipedia, the free encyclopedia,” 29-Sep-2010. [Online]. Available: http://en.wikipedia.org/wiki/Semantic_reasoner.

- [22] “The DICOM Standart.” [Online]. Available: <http://www.e-radiography.net/cr/dicom/dicomintro.htm>.
- [23] “Health Level Seven International - Homepage.” [Online]. Available: <http://www.hl7.org/>.
- [24] “IHTSDO: SNOMED CT.” [Online]. Available: <http://www.ihtsdo.org/snomed-ct/>.
- [25] “openEHR - Wikipedia, the free encyclopedia,” 21-Jun-2010. [Online]. Available: <http://en.wikipedia.org/wiki/OpenEHR>.
- [26] “openEHR : future proof and flexible EHR specifications.” [Online]. Available: <http://www.openehr.org/home.html>.
- [27] “Foundational Model of Anatomy ontology - About.” [Online]. Available: <http://sig.biostr.washington.edu/projects/fm/AboutFM.html>.
- [28] “OpenGALEN Mission Statement.” [Online]. Available: <http://www.opengalen.org/>.
- [29] “Gene Ontology Documentation.” [Online]. Available: <http://www.geneontology.org/GO.contents.doc.shtml>.
- [30] “Unified Medical Language System (UMLS) - Home.” [Online]. Available: <http://www.nlm.nih.gov/research/umls/>.
- [31] “Picture archiving and communication system - Wikipedia, the free encyclopedia,” 27-Sep-2010. [Online]. Available: http://en.wikipedia.org/wiki/Picture_archiving_and_communication_system.
- [32] M. Moeller and M. Sintek, “RadSem: Semantic Annotation and Retrieval for Medical Images,” in *Proceedings of The 6th Annual European Semantic Web Conference (ESWC2009)*, pp. 21-35, 2009.
- [33] O. Dameron, N. Noy, H. Knublauch, and M. Musen, “Accessing and Manipulating Ontologies Using Web Services,” in *Proceedings of Semantic Web Services workshop held at the Third International Semantic Web Conference (ISWC2004)*, 2004.
- [34] “Foundational Model of Anatomy ontology - Home.” [Online]. Available: <http://sig.biostr.washington.edu/projects/fm/>.
- [35] “ICD-10.” [Online]. Available: <http://apps.who.int/classifications/apps/icd/icd10online/>.
- [36] J. Mejino, J. Franklin, L. Detwiler, D. Rubin, and J. Brinkley, “Web service access to semantic web ontologies for data annotation,” in *American Medical Informatics Association (AMIA) Annual Symposium Proceedings*, p. 946, 2008.
- [37] “IATΡΟΛΕΞΗ.” [Online]. Available: <http://www.iatrolexi.gr/iatrolexi/index.html>.

9

Παραρτήματα

Παράρτημα Α: Οντολογία εφαρμογής σε μορφή RDF/XML

```
<?xml version="1.0"?>
<!DOCTYPE rdf:RDF [
  <!ENTITY owl "http://www.w3.org/2002/07/owl#" >
  <!ENTITY xsd "http://www.w3.org/2001/XMLSchema#" >
  <!ENTITY rdfs "http://www.w3.org/2000/01/rdf-schema#" >
  <!ENTITY fma3 "http://sig.biostr.washington.edu/fma3.0#" >
  <!ENTITY rdf "http://www.w3.org/1999/02/22-rdf-syntax-ns#" >
  <!ENTITY semia "http://www.semanticweb.org/ontologies/semia.owl#"
>
]>
<rdf:RDF xmlns="http://www.semanticweb.org/ontologies/semia.owl#"
  xml:base="http://www.semanticweb.org/ontologies/semia.owl"
  xmlns:semia="http://www.semanticweb.org/ontologies/semia.owl#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:fma3="http://sig.biostr.washington.edu/fma3.0#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"

```

```

xmlns:owl="http://www.w3.org/2002/07/owl#">
  <owl:Ontology
rdf:about="http://www.semanticweb.org/ontologies/semia.owl"/>
  <!--
////////////////////////////////////
////////////////////////////////////
  //
  // Datatypes
  //

////////////////////////////////////
////////////////////////////////////
  -->
  <!-- http://www.w3.org/2001/XMLSchema#ID -->

  <rdfs:Datatype rdf:about="&xsd;ID"/>
  <!--

////////////////////////////////////
////////////////////////////////////
  //
  // Object Properties
  //

////////////////////////////////////
////////////////////////////////////
  -->
  <!-- http://www.semanticweb.org/ontologies/semia.owl#annotated --
>

  <owl:ObjectProperty rdf:about="&semia;annotated">
    <rdfs:domain rdf:resource="&semia;Annotator"/>
    <rdfs:range rdf:resource="&semia;ROI"/>
    <owl:inverseOf rdf:resource="&semia;isAnnotatedBy"/>
    <rdfs:subPropertyOf rdf:resource="&owl;topObjectProperty"/>
  </owl:ObjectProperty>

```

```

    <!--
http://www.semanticweb.org/ontologies/semia.owl#hasAnatomyAnnotation
-->

    <owl:ObjectProperty rdf:about="&semia;hasAnatomyAnnotation">
        <rdfs:range rdf:resource="&fma3;Anatomical_entity"/>
        <rdfs:domain rdf:resource="&semia;ROI"/>
        <rdfs:subPropertyOf rdf:resource="&semia;hasAnnotation"/>
    </owl:ObjectProperty>

    <!-- http://www.semanticweb.org/ontologies/semia.owl#hasAnnotation -
->

    <owl:ObjectProperty rdf:about="&semia;hasAnnotation">
        <rdfs:domain rdf:resource="&semia;ROI"/>
        <rdfs:range rdf:resource="&semia;SemanticAnnotation"/>
        <rdfs:subPropertyOf rdf:resource="&owl;topObjectProperty"/>
    </owl:ObjectProperty>

    <!--
http://www.semanticweb.org/ontologies/semia.owl#hasDiseaseAnnotation
-->

    <owl:ObjectProperty rdf:about="&semia;hasDiseaseAnnotation">
        <rdfs:range rdf:resource="&semia;DiseaseAnnotation"/>
        <rdfs:domain rdf:resource="&semia;ROI"/>
        <rdfs:subPropertyOf rdf:resource="&semia;hasAnnotation"/>
    </owl:ObjectProperty>

    <!-- http://www.semanticweb.org/ontologies/semia.owl#hasPatient -
->

    <owl:ObjectProperty rdf:about="&semia;hasPatient">
        <rdfs:domain rdf:resource="&semia;Image"/>
        <rdfs:range rdf:resource="&semia;Patient"/>
        <rdfs:subPropertyOf rdf:resource="&owl;topObjectProperty"/>
    </owl:ObjectProperty>

```

```

<!-- http://www.semanticweb.org/ontologies/semia.owl#hasROI -->

<owl:ObjectProperty rdf:about="&semia;hasROI">
  <rdfs:domain rdf:resource="&semia;Image"/>
  <rdfs:range rdf:resource="&semia;ROI"/>
  <owl:inverseOf rdf:resource="&semia;isROIOf"/>
</owl:ObjectProperty>

<!-- http://www.semanticweb.org/ontologies/semia.owl#hasSex -->

<owl:ObjectProperty rdf:about="&semia;hasSex">
  <rdfs:domain rdf:resource="&semia;Person"/>
  <rdfs:range rdf:resource="&semia;Sex"/>
  <rdfs:subPropertyOf rdf:resource="&owl;topObjectProperty"/>
</owl:ObjectProperty>

<!--
http://www.semanticweb.org/ontologies/semia.owl#hasSpecialization -->

<owl:ObjectProperty rdf:about="&semia;hasSpecialization">
  <rdfs:domain rdf:resource="&semia;Annotator"/>
  <rdfs:range rdf:resource="&semia;Specialization"/>
  <rdfs:subPropertyOf rdf:resource="&owl;topObjectProperty"/>
</owl:ObjectProperty>

<!--
http://www.semanticweb.org/ontologies/semia.owl#isAnnotatedBy -->

<owl:ObjectProperty rdf:about="&semia;isAnnotatedBy">
  <rdf:type rdf:resource="&owl;FunctionalProperty"/>
  <rdfs:range rdf:resource="&semia;Annotator"/>
  <rdfs:domain rdf:resource="&semia;ROI"/>
  <rdfs:subPropertyOf rdf:resource="&owl;topObjectProperty"/>
</owl:ObjectProperty>

```



```
<!-- http://www.semanticweb.org/ontologies/semia.owl#isROIOf
-->
```

```
<owl:ObjectProperty rdf:about="&semia;isROIOf">
  <rdf:type rdf:resource="&owl;FunctionalProperty"/>
  <rdfs:range rdf:resource="&semia;Image"/>
  <rdfs:domain rdf:resource="&semia;ROI"/>
  <rdfs:subPropertyOf rdf:resource="&owl;topObjectProperty"/>
</owl:ObjectProperty>
```

```
<!-- http://www.w3.org/2002/07/owl#topObjectProperty -->
```

```
<owl:ObjectProperty rdf:about="&owl;topObjectProperty"/>
```

```
<!--
```

```
////////////////////////////////////
////////////////////////////////////
```

```
//
// Data properties
//
```

```
////////////////////////////////////
////////////////////////////////////
```

```
-->
```

```
<!-- http://www.semanticweb.org/ontologies/semia.owl#hasAge -->
```

```
<owl:DatatypeProperty rdf:about="&semia;hasAge">
  <rdfs:domain rdf:resource="&semia;Patient"/>
  <rdfs:range rdf:resource="&xsd;integer"/>
  <rdfs:subPropertyOf rdf:resource="&owl;topDataProperty"/>
</owl:DatatypeProperty>
```

```

<!--
http://www.semanticweb.org/ontologies/semia.owl#hasAnnotatorID -->

<owl:DatatypeProperty rdf:about="&semia;hasAnnotatorID">
  <rdf:type rdf:resource="&owl;FunctionalProperty"/>
  <rdfs:domain rdf:resource="&semia;Annotator"/>
  <rdfs:range rdf:resource="&xsd;integer"/>
  <rdfs:subPropertyOf rdf:resource="&owl;topDataProperty"/>
</owl:DatatypeProperty>

<!--
http://www.semanticweb.org/ontologies/semia.owl#hasBodyPartExamined -
->

<owl:DatatypeProperty rdf:about="&semia;hasBodyPartExamined">
  <rdfs:domain rdf:resource="&semia;Image"/>
  <rdfs:range rdf:resource="&xsd;string"/>
  <rdfs:subPropertyOf rdf:resource="&owl;topDataProperty"/>
</owl:DatatypeProperty>

<!-- http://www.semanticweb.org/ontologies/semia.owl#hasImageID -
->

<owl:DatatypeProperty rdf:about="&semia;hasImageID">
  <rdf:type rdf:resource="&owl;FunctionalProperty"/>
  <rdfs:domain rdf:resource="&semia;Image"/>
  <rdfs:range rdf:resource="&xsd;integer"/>
  <rdfs:subPropertyOf rdf:resource="&owl;topDataProperty"/>
</owl:DatatypeProperty>

<!-- http://www.semanticweb.org/ontologies/semia.owl#hasModality
-->

<owl:DatatypeProperty rdf:about="&semia;hasModality">
  <rdfs:domain rdf:resource="&semia;Image"/>
  <rdfs:range rdf:resource="&xsd;string"/>

```

```

    <rdfs:subPropertyOf rdf:resource="&owl;topDataProperty"/>
</owl:DatatypeProperty>

<!-- http://www.semanticweb.org/ontologies/semia.owl#hasName -->

<owl:DatatypeProperty rdf:about="&semia;hasName">
    <rdfs:domain rdf:resource="&semia;Person"/>
    <rdfs:range rdf:resource="&xsd:string"/>
    <rdfs:subPropertyOf rdf:resource="&owl;topDataProperty"/>
</owl:DatatypeProperty>

<!--
http://www.semanticweb.org/ontologies/semia.owl#hasOriginalName -->

<owl:DatatypeProperty rdf:about="&semia;hasOriginalName">
    <rdfs:domain rdf:resource="&semia;Image"/>
    <rdfs:range rdf:resource="&xsd:string"/>
    <rdfs:subPropertyOf rdf:resource="&owl;topDataProperty"/>
</owl:DatatypeProperty>

<!-- http://www.semanticweb.org/ontologies/semia.owl#hasPatientID
-->

<owl:DatatypeProperty rdf:about="&semia;hasPatientID">
    <rdfs:type rdf:resource="&owl;FunctionalProperty"/>
    <rdfs:domain rdf:resource="&semia;Patient"/>
    <rdfs:range rdf:resource="&xsd;integer"/>
    <rdfs:subPropertyOf rdf:resource="&owl;topDataProperty"/>
</owl:DatatypeProperty>

<!--
http://www.semanticweb.org/ontologies/semia.owl#hasReferringPhysicians -->

<owl:DatatypeProperty rdf:about="&semia;hasReferringPhysicians">

```

```

    <rdfs:domain rdf:resource="&semia;Image"/>
    <rdfs:range rdf:resource="&xsd:string"/>
    <rdfs:subPropertyOf rdf:resource="&owl;topDataProperty"/>
</owl:DatatypeProperty>

<!--
http://www.semanticweb.org/ontologies/semia.owl#hasStudyDescription -
->

<owl:DatatypeProperty rdf:about="&semia;hasStudyDescription">
    <rdfs:domain rdf:resource="&semia;Image"/>
    <rdfs:range rdf:resource="&xsd:string"/>
    <rdfs:subPropertyOf rdf:resource="&owl;topDataProperty"/>
</owl:DatatypeProperty>

<!--
http://www.semanticweb.org/ontologies/semia.owl#isOfInstitution -->

<owl:DatatypeProperty rdf:about="&semia;isOfInstitution">
    <rdfs:domain rdf:resource="&semia;Image"/>
    <rdfs:range rdf:resource="&xsd:string"/>
    <rdfs:subPropertyOf rdf:resource="&owl;topDataProperty"/>
</owl:DatatypeProperty>

<!-- http://www.w3.org/2002/07/owl#topDataProperty -->

<owl:DatatypeProperty rdf:about="&owl;topDataProperty"/>

<!--

////////////////////////////////////
////////////////////////////////////
//
// Classes
//

```

```

////////////////////////////////////
////////////////////////////////////
-->
<!-- http://sig.biostr.washington.edu/fma3.0#Anatomical_entity --
>

<owl:Class rdf:about="&fma3;Anatomical_entity">
  <rdfs:subClassOf rdf:resource="&semia;SemanticAnnotation"/>
</owl:Class>

<!-- http://www.semanticweb.org/ontologies/semia.owl#Annotator --
>

<owl:Class rdf:about="&semia;Annotator">
  <rdfs:subClassOf rdf:resource="&semia;Person"/>
</owl:Class>

<!--
http://www.semanticweb.org/ontologies/semia.owl#DiseaseAnnotation -->

<owl:Class rdf:about="&semia;DiseaseAnnotation">
  <rdfs:subClassOf rdf:resource="&semia;SemanticAnnotation"/>
</owl:Class>

<!-- http://www.semanticweb.org/ontologies/semia.owl#Image -->

<owl:Class rdf:about="&semia;Image"/>

<!-- http://www.semanticweb.org/ontologies/semia.owl#Patient -->

<owl:Class rdf:about="&semia;Patient">
  <rdfs:subClassOf rdf:resource="&semia;Person"/>
</owl:Class>

```

```

<!-- http://www.semanticweb.org/ontologies/semia.owl#Person -->

<owl:Class rdf:about="&semia;Person">
  <owl:equivalentClass>
    <owl:Restriction>
      <owl:onProperty rdf:resource="&semia;hasName"/>
      <owl:someValuesFrom rdf:resource="&xsd;Name"/>
    </owl:Restriction>
  </owl:equivalentClass>
</owl:Class>

<!-- http://www.semanticweb.org/ontologies/semia.owl#ROI -->

<owl:Class rdf:about="&semia;ROI">
  <owl:equivalentClass>
    <owl:Restriction>
      <owl:onProperty rdf:resource="&semia;isROIOf"/>
      <owl:someValuesFrom rdf:resource="&semia;Image"/>
    </owl:Restriction>
  </owl:equivalentClass>
  <rdfs:subClassOf rdf:resource="&owl;Thing"/>
</owl:Class>

<!--
http://www.semanticweb.org/ontologies/semia.owl#SemanticAnnotation --
>

<owl:Class rdf:about="&semia;SemanticAnnotation"/>

<!-- http://www.semanticweb.org/ontologies/semia.owl#Sex -->

<owl:Class rdf:about="&semia;Sex">
  <owl:equivalentClass>
    <owl:Class>

```

```

        <owl:oneOf rdf:parseType="Collection">
            <rdf:Description rdf:about="&semia;male"/>
            <rdf:Description rdf:about="&semia;female"/>
        </owl:oneOf>
    </owl:Class>
</owl:equivalentClass>
</owl:Class>

<!--
http://www.semanticweb.org/ontologies/semia.owl#Specialization -->

<owl:Class rdf:about="&semia;Specialization">
    <owl:equivalentClass>
        <owl:Class>
            <owl:oneOf rdf:parseType="Collection">
                <rdf:Description rdf:about="&semia;Biologist"/>
                <rdf:Description rdf:about="&semia;Doctor"/>
                <rdf:Description rdf:about="&semia;Radiologist"/>
                <rdf:Description
rdf:about="&semia;MedicineStudent"/>
                <rdf:Description rdf:about="&semia;Physician"/>
            </owl:oneOf>
        </owl:Class>
    </owl:equivalentClass>
</owl:Class>

<!-- http://www.w3.org/2002/07/owl#Thing -->

<owl:Class rdf:about="&owl;Thing"/>

<!--

////////////////////////////////////
////////////////////////////////////
//

```

```

// Individuals
//

////////////////////////////////////
////////////////////////////////////

-->

<!-- http://www.semanticweb.org/ontologies/semia.owl#Biologist --
>

<owl:NamedIndividual rdf:about="&semia;Biologist">
  <rdf:type rdf:resource="&semia;Specialization"/>
</owl:NamedIndividual>

<!-- http://www.semanticweb.org/ontologies/semia.owl#Doctor -->

<owl:NamedIndividual rdf:about="&semia;Doctor">
  <rdf:type rdf:resource="&semia;Specialization"/>
</owl:NamedIndividual>

<!--
http://www.semanticweb.org/ontologies/semia.owl#MedicineStudent -->

<owl:NamedIndividual rdf:about="&semia;MedicineStudent">
  <rdf:type rdf:resource="&semia;Specialization"/>
</owl:NamedIndividual>

<!-- http://www.semanticweb.org/ontologies/semia.owl#Physician --
>

<owl:NamedIndividual rdf:about="&semia;Physician">
  <rdf:type rdf:resource="&semia;Specialization"/>
</owl:NamedIndividual>

```



```
<!-- http://www.semanticweb.org/ontologies/semia.owl#Radiologist
-->

<owl:NamedIndividual rdf:about="&semia;Radiologist">
  <rdf:type rdf:resource="&semia;Specialization"/>
</owl:NamedIndividual>

<!-- http://www.semanticweb.org/ontologies/semia.owl#female -->

<owl:NamedIndividual rdf:about="&semia;female">
  <rdf:type rdf:resource="&semia;Sex"/>
</owl:NamedIndividual>

<!-- http://www.semanticweb.org/ontologies/semia.owl#male -->

<owl:NamedIndividual rdf:about="&semia;male">
  <rdf:type rdf:resource="&semia;Sex"/>
</owl:NamedIndividual>
</rdf:RDF>
<!-- Generated by the OWL API (version 3.0.0.1451) http://owlapi.sourceforge.net -->
```