



«DARIAH-ΑΤΤΙΚΗ Ανάπτυξη της ελληνικής ερευνητικής υποδομής για τις ανθρωπιστικές επιστήμες ΔΥΑΣ»

ΑΚΑΔΗΜΙΑ ΑΘΗΝΩΝ

ΕΚ ΑΘΗΝΑ

ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ

ΕΡΕΥΝΗΤΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΑΚΟ ΙΝΣΤΙΤΟΥΤΟ

ΣΥΣΤΗΜΑΤΩΝ ΕΠΙΚΟΙΝΩΝΙΩΝ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

ΑΝΩΤΑΤΗ ΣΧΟΛΗ ΚΑΛΩΝ ΤΕΧΝΩΝ

ΙΔΡΥΜΑ ΤΕΧΝΟΛΟΓΙΑΣ ΚΑΙ ΕΡΕΥΝΑΣ

ΥΠΟΕΡΓΟ 4

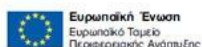
Δράσεις ανάπτυξης της εθνικής υποδομής του ΕΠΙΣΕΥ

Παραδοτέο

ΠΑ.1.4.ΕΠΙΣΕΥ.3 Σημασιολογικός εμπλουτισμός των μεταδεδομένων με χρήση πολυμεσικών πληροφοριών και συνδεδεμένων δεδομένων

Έκδοση 1.0-Τελική
27/3/2015

Έγγραφο:



η περιφέρεια στο επίκεντρο της ανάπτυξης

Με τη συγχρηματοδότηση της Ελλάδας και του Ευρωπαϊκού Ταμείου Περιφερειακής Ανάπτυξης της Ευρωπαϊκής Ένωσης στο πλαίσιο του Ε.Π. Ανταγωνιστικότητα & Επιχειρηματικότητα και των Π.Ε.Π. Αττικής, Π.Ε.Π. Μακεδονίας Θράκης, Π.Ε.Π. Κρήτης & Νήσων Αιγαίου, Π.Ε.Π. Θεσσαλίας-Στερεάς Ελλάδας-Ηπείρου

Ταυτότητα Εγγράφου

ΥΠΟΕΡΓΟ	4. Δράσεις ανάπτυξης της εθνικής υποδομής του ΕΠΙΣΕΥ
Ενότητα εργασιών	1. ΚΟΙΝΩΝΙΑ ΨΗΦΙΑΚΩΝ ΠΟΡΩΝ
ΔΡΑΣΗ	1.4. Μητρώο λογισμικών υπηρεσιών
ΠΑΡΑΔΟΤΕΟ	ΠΑ.1.4.ΕΠΙΣΕΥ.3 Σημαιολογικός εμπλουτισμός των μεταδεδομένων με χρήση πολυμεσικών πληροφοριών και συνδεδεμένων δεδομένων
Εργασία	Συγγραφή εγγράφου
Ομάδα Έργου	Στέφανος Κόλλιας Γεώργιος Μαρανδιανός Φοίβος Μυλωνάς Αλέξανδρος Χορταράς
Χαρακτηρισμός	Εσωτερικό Ομάδας Έργου

Ιστορικό Αλλαγών

Έκδ.	Ημ/μηνία	Αιτιολογία	Σύνοψη Αλλαγών	Επισπεύδοντες
0.1	31/12/2014	Αρχική	Δημιουργία και Σύνταξη	Ομάδα Έργου
1.0	27/3/2015	Τελική		

Πίνακας Περιεχομένων

ΕΙΣΑΓΩΓΗ	5
Σκοπός	5
Ομάδες Στόχοι	6
1. ΨΗΦΙΑΚΕΣ ΒΙΒΛΙΟΘΗΚΕΣ & ΑΝΑΖΗΤΗΣΗ ΠΟΛΙΤΙΣΤΙΚΟΥ ΠΕΡΙΕΧΟΜΕΝΟΥ	7
2. ΣΗΜΑΣΙΟΛΟΓΙΚΗ ΑΝΑΖΗΤΗΣΗ ΣΕ ΨΗΦΙΑΚΕΣ ΒΙΒΛΙΟΘΗΚΕΣ	11
2.1 Το σύστημα σημασιολογικής αναζήτησης	11
2.2 Ο αλγόριθμος σημασιολογικής αναζήτησης	14
3. ΕΞΑΓΩΓΗ ΟΠΤΙΚΩΝ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ ΚΑΙ ΔΗΜΙΟΥΡΓΙΑ ΟΠΤΙΚΟΥ ΛΕΞΙΚΟΥ	16
3.1 Εξαγωγή MPEG-7 χαρακτηριστικών	16
3.2 Εξαγωγή τοπικών χαρακτηριστικών και δημιουργία οπτικού λεξικού	20
4. ΣΥΝΔΥΑΖΟΝΤΑΣ ΤΗ ΣΗΜΑΣΙΟΛΟΓΙΚΗ ΑΝΑΖΗΤΗΣΗ ΜΕ ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ ΤΩΝ ΟΠΤΙΚΩΝ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ	23
5. ΜΕΛΕΤΗ ΑΞΙΟΛΟΓΗΣΗΣ: ΒΕΛΤΙΩΣΗ ΤΗΣ ΣΗΜΑΣΙΟΛΟΓΙΚΗΣ ΑΝΑΖΗΤΗΣΗΣ ΜΕΣΩ ΤΗΣ ΟΠΤΙΚΗΣ ΑΝΑΛΥΣΗΣ	27
5.1 Πειραματική μελέτη	27
5.2 Τα αποτελέσματα της απάντησης ερωτημάτων	30
5.3 Ζητήματα κλιμάκωσης: Μια οπτική μηχανή αναζήτησης για την Europeana (VIEU)	34
6. ΕΜΠΛΟΥΤΙΣΜΟΣ ΜΕΣΩ ΔΙΑΣΥΝΔΕΔΕΜΕΝΩΝ ΔΕΔΟΜΕΝΩΝ	36
6.1 Βασικές αρχές διασυνδεδεμένων δεδομένων	36
6.2 Σημασιολογική αναπαράσταση πολυμεσικού περιεχομένου	37
6.3 Σημασιολογικός εμπλουτισμός	41
6.4 Ανάπτυξη πιλοτικής εφαρμογής	42
6.4.1 Σημασιολογική βάση δεδομένων	43
6.4.2 Αποτελέσματα του εμπλουτισμού	45
7. ΕΡΓΑΛΕΙΑ ΕΜΠΛΟΥΤΙΣΜΟΥ ΜΕ ΔΙΑΣΥΝΔΕΔΕΜΕΝΑ ΔΕΔΟΜΕΝΑ	48
7.1 Karma	48
7.2 PoolParty	49
7.3 Alchemy API	50

7.4	Dexter 2.0	51
7.5	LODE	52
7.6	Silk 52	
7.7	Pundit	53
8.	ΣΥΜΠΕΡΑΣΜΑΤΑ	55
9.	ΒΙΒΛΙΟΓΡΑΦΙΑ	57

Εισαγωγή

Το παραδοτέο αυτό περιγράφει τη μεθοδολογία και τους αλγορίθμους που επιτρέπουν τον σημασιολογικό εμπλουτισμό των μεταδεδομένων πολιτιστικού περιεχομένου με ταυτόχρονη ανάλυση πολυμεσικών πληροφοριών, κυρίως οπτικού και οπτικοακουστικού υλικού, για την απάντηση επερωτημάτων των χρηστών, όπως και χρήση συνδεδεμένων δεδομένων.

Το παρόν Παραδοτέο είναι το τρίτο της Δράσης για Μελέτη και Δημιουργία «Μητρώου Λογισμικών Υπηρεσιών» που υλοποιεί το Ε.Π.Ι.Σ.Ε.Υ. στο πλαίσιο του έργου DARIAH-Αττική.

Σκοπός

Το παρόν Παραδοτέο αναφέρεται στη χρήση σημασιολογικής αναπαράστασης γνώσης για την απάντηση επερωτημάτων σε ψηφιακές βιβλιοθήκες που περιλαμβάνουν και πολυμεσικό περιεχόμενο. Σκοπός του είναι να παρουσιάσει την αιχμή της τεχνολογίας στην ανάπτυξη συστημάτων που προσφέρουν αποτελεσματική αναζήτηση και πρόσβαση στο περιεχόμενο ψηφιακών βιβλιοθηκών, κειμενικό και οπτικό, συνδυάζοντας σημασιολογική ανάλυση των μεταδεδομένων περιγραφής του περιεχομένου αυτού με τεχνικές μηχανικής μάθησης που εφαρμόζονται στο οπτικό τμήμα του περιεχομένου. Η μεθοδολογία περιλαμβάνει τα επόμενα:

- Συσσώρευση των μεταδεδομένων περιγραφής του περιεχομένου των ψηφιακών βιβλιοθηκών και δημιουργία του υποσυστήματος της τυπικής γνώσης με χρησιμοποίηση θεματικών ορολογιών και οντολογιών.
- Χρησιμοποίηση μηχανής συμπερασματολογίας για την απάντηση στα επερωτήματα των χρηστών.
- Ανάλυση πολυμεσικού περιεχομένου με εξαγωγή χαρακτηριστικών από τις εικόνες των πολιτιστικών τεκμηρίων και δημιουργία οπτικού λεξιλογίου για ταυτοποίησή τους με χρησιμοποίηση τεχνικών μηχανικής μάθησης.
- Συνδυασμό των υποσυστημάτων σημασιολογικής και οπτικής ανάλυσης για βελτίωση της παραγωγής αποτελεσματικών απαντήσεων στα ερωτήματα των χρηστών.

Για όλα τα ανωτέρω παρουσιάζονται αποτελέσματα και γίνεται αξιολόγηση της απόδοσης της μεθοδολογίας και των επιμέρους τεχνολογιών πάνω σε πραγματικά δεδομένα της Ευρωπαϊκής Ψηφιακής Βιβλιοθήκης (Europeana).

Περιγράφεται τέλος μια πιλοτική εφαρμογή για τον σημασιολογικό εμπλουτισμό των μεταδεδομένων πολυμεσικού περιεχομένου, με έμφαση στο οπτικοακουστικό περιεχόμενο, η οποία περιλαμβάνει και τη σύνδεση με την πλατφόρμα ενσωμάτωσης μεταδεδομένων MINT που χρησιμοποιείται για τη διαλειτουργική συγκέντρωσή τους. Στο πλαίσιο αυτής περιγράφεται η μετατροπή των μεταδεδομένων σε Διασυνδεδεμένα Δεδομένα (Linked Data) και ο σημασιολογικός εμπλουτισμός τους για αποτελεσματική αναζήτηση και απάντηση στα ερωτήματα των χρηστών.

Ομάδες Στόχοι

Το παραδοτέο απευθύνεται τόσο στα μέλη των φορέων του έργου DARIAH-ΑΤΤΙΚΗ, και στους ερευνητές των φορέων αυτών που απασχολούνται στο πρόγραμμα, όσο και στους φορείς-χρήστες και στους ερευνητές που θα χρησιμοποιήσουν τα αποτελέσματα του έργου.

1. Ψηφιακές βιβλιοθήκες & αναζήτηση πολιτιστικού περιεχομένου

Οι εξελίξεις στον τομέα της δημιουργίας και πρόσβασης στην Ψηφιακή Πολιτιστική Κληρονομιά είναι ραγδαίες τα τελευταία χρόνια. Σημαντικό ρόλο σε αυτό έχουν παίξει η μαζική ψηφιοποίηση και τεκμηρίωση πολιτιστικού περιεχομένου που βρίσκεται σε εξέλιξη σε όλη την Ευρώπη και τον κόσμο, που ξεκίνησε πριν από δέκα πέντε περίπου χρόνια με την δημοσίευση των "αρχών του Lund» [1]. Επιπλέον, η έντονη συμμετοχή των εταιρειών όπως η Google, σε συνδυασμό με τη θετική αντίδραση και αυξανόμενη στήριξη της Ευρωπαϊκής Ένωσης, έχουν οδηγήσει σε μια ποικιλία συγκλινουσών δράσεων που στοχεύουν στην παραγωγή και ευρεία πρόσβαση σε πολυτροπικό (multimodal) και πολυμεσικό περιεχόμενο από κάθε δυνατή πηγή (δηλαδή, πινακοθήκες, βιβλιοθήκες, αρχεία, μουσεία, οπτικοακουστικά αρχεία, που σε συντομία αναφέρονται ως GLAMs).

Η δημιουργία και η εξέλιξη της Europeana¹, ως ένα μοναδικό σημείο πρόσβασης στην Ευρωπαϊκή πολιτιστική κληρονομιά, αποτελεί ένα από τα σημαντικότερα επιτεύγματα αυτών των προσπαθειών. Αυτή τη στιγμή, πάνω από 34 εκατομμύρια αντικείμενα, που απεικονίζουν τον Ευρωπαϊκό πολιτιστικό πλούτο, είναι προσβάσιμα μέσω της πύλης Europeana, και αναμένεται ότι ο αριθμός αυτός θα αυξηθεί σημαντικά μέσα στα επόμενα τρία χρόνια. Η διαδικτυακή πύλη της Europeana παρέχει πρόσβαση σε διάφορα πολιτιστικά αντικείμενα καθώς και τις ψηφιακές αναπαραστάσεις τους, με την πλειοψηφία τους να είναι κείμενο ή εικόνες.

Οι πολιτιστικοί φορείς προσθέτουν μεταδεδομένα στα ψηφιοποιημένα τεκμήρια, προσδιορίζοντας για παράδειγμα τον δημιουργό, την προέλευση, τον χρόνο δημιουργίας, παρέχοντας σχετικές και τεχνικές πληροφορίες για τη μορφή και τα χαρακτηριστικά των αντικειμένων, ώστε να διευκολυνθεί η ανίχνευσή τους από μηχανές αναζήτησης. Μια εικόνα του αντικειμένου είναι επίσης συνδεδεμένη με την περιγραφή του διευκολύνοντας έτσι την ευρεία χρήση τους.

Οι υπηρεσίες που παρέχονται από τις ψηφιακές βιβλιοθήκες (συμπεριλαμβανομένης της Europeana), αναφέρονται είτε στους παρόχους περιεχομένου, δηλαδή στους φορείς που ψηφιοποιούν και παρέχουν το περιεχόμενο και τα μεταδεδομένα on-line, ή στους καταναλωτές περιεχομένου, π.χ. σε ερευνητές, σπουδαστές και γενικούς χρήστες,

¹ <http://www.europeana.eu>

επιτρέποντας το σχολιασμό, τη συζήτηση και τη συλλογή υλικού. Οι πιο διαδεδομένες τεχνολογίες και σχετικές υπηρεσίες που έχουν συγκεντρώσει την προσοχή στο πεδίο αυτό είναι η διαλειτουργικότητα (Interoperability) και η σημασιολογική αναζήτηση (Semantic Search) [18, 22, 23]. Η διαλειτουργικότητα μπορεί να απαιτήσει το μετασχηματισμό των μεταδεδομένων σε καθιερωμένα πρότυπα, όπως τα Europeana Semantic Element (ESE) και Europeana Data Model (EDM) της Europeana, ή τη δημιουργία αντιστοιχίσεων μεταξύ των μεταδεδομένων που χρησιμοποιούνται από τους διαφορετικούς παρόχους περιεχομένου. Η πλατφόρμα `MINT' (<http://mint.image.ece.ntua.gr/redmine/projects/mint/wiki>) έχει χρησιμοποιηθεί σε περισσότερα από 15 προγράμματα της Europeana για την συσσώρευση περισσότερων από 5 εκατομμυρίων τεκμηρίων. Η δημόσια ψηφιακή βιβλιοθήκη της Αμερικής (Digital Public Library of America (<http://blogs.law.harvard.edu/dpla/>) έχει χρησιμοποιήσει επίσης την πλατφόρμα.

Η σημασιολογική αναζήτηση αποσκοπεί στην απάντηση των ερωτημάτων των χρηστών, εκμεταλλευόμενη τόσο την ρητά (explicit) εκφρασμένη όσο και την υπονοούμενη (implicit) γνώση. Η συλλογιστική (reasoning) σε διαθέσιμες βάσεις δεδομένων [19], η οποία βασίζεται σε περιγραφικές λογικές ή στις γλώσσες RDF και OWL [4,8,9] αποτελεί τη βάση για την κατεύθυνση αυτή. Η δημιουργία των linked data stores [25,26] από πόρους της ψηφιακής πολιτισμικής κληρονομιάς επιτρέπει τη σύνδεση πολλαπλών δεδομένων, αυξάνοντας έτσι την αποδοτικότητα και επιτρέποντας συνδυασμένες ή συνδεδεμένες αναζητήσεις.

Η σημασιολογική αναζήτηση και απάντηση σε πολύπλοκα επερωτήματα των χρηστών που αναπτύσσεται στο 2^ο Παραδοτέο του ΕΠΙΣΕΥ (1.4.2) επεκτείνεται στο Παρόν Παραδοτέο, στοχεύοντας σε βελτιωμένη αναζήτηση με ταυτόχρονη χρήση του οπτικού και οπτικοακουστικού περιεχομένου που συνοδεύει τα μεταδεδομένα.

Το οπτικοακουστικό περιεχόμενο που περιλαμβάνει τόσο τις εικόνες που συνοδεύουν τα πολιτιστικά τεκμήρια, όσο και υλικό από την τηλεόραση, τις ταινιοθήκες, τα οπτικοακουστικά αρχεία, θεωρείται ιδιαίτερης σπουδαιότητας συστατικό της πολιτιστικής κληρονομιάς της Ευρώπης, επειδή αποτελεί συλλογική μνήμη και ταυτότητα για την Ευρώπη, ενώ και αποτελεί την πλέον πολύπλοκη μορφή πληροφορίας για ψηφιακή αναζήτηση και πρόσβαση.

Το πολυμεσικό περιεχόμενο που επιλέχθηκε στην παρούσα μελέτη για σημασιολογικό εμπλουτισμό και δημιουργία μιας πιλοτικής εφαρμογής προέρχεται από τα έργα EUScreen και EUScreenXL². Ο στόχος αυτών των έργων είναι η δημιουργία μιας αντιπροσωπευτικής συλλογής τηλεοπτικών προγραμμάτων, δευτερευουσών πηγών και άρθρων που επιτρέπουν την αποτελεσματική πρόσβαση στους μαθητές, στους μελετητές και στο ευρύ κοινό. Ωστόσο, η παροχή πρόσβασης σε μεγάλες ψηφιακές συλλογές αντικειμένων πολιτιστικής κληρονομιάς είναι ένα δύσκολο έργο που περιλαμβάνει την επίλυση πολλών προβλημάτων. Αρχικά πρέπει να εξεταστεί η συγκέντρωση των μεταδεδομένων σε συνδυασμό με την διαλειτουργικότητα, δηλαδή την διαδικασία εναρμόνισής τους, καθώς οι διάφοροι πάροχοι περιεχομένου χρησιμοποιούν διαφορετικά πρότυπα μεταδεδομένων. Στη συνέχεια, τα μεταδεδομένα πρέπει να τίθενται στη διάθεση του κοινού με συνεπή τρόπο, όχι μόνο προσφέροντας μια - φιλική προς το χρήστη - πλοήγηση και προεπισκόπηση σε αυτά, αλλά παράλληλα επιτρέποντας την χρήση τους από οποιαδήποτε εφαρμογή.

Στο παρόν Παραδοτέο παρουσιάζουμε τις ροές εργασίας και τα αντίστοιχα εργαλεία που χρησιμοποιούνται για τον εμπλουτισμό του περιεχομένου και την δημοσίευσή του στο πλαίσιο των Διασυνδεδεμένων Δεδομένων. Συγκεκριμένα, η συνολική ροή εργασίας αποτελείται από τρία βασικά βήματα που αναλύονται στη συνέχεια:

- Συσσώρευση των μεταδεδομένων,
- Μετατροπή τους σε ένα κοινό σχήμα αναφοράς,
- Δημοσίευσή τους ως Linked Data.

Η δυνατότητα ύπαρξης διαφορετικών παρόχων περιεχομένου, και κατά συνέπεια η χρήση διαφορετικών συστημάτων διαχείρισης και τύπων μεταδεδομένων που συνεπάγεται κάτι τέτοιο, δημιουργεί την ανάγκη για σημασιολογική διαλειτουργικότητα. Για την επίτευξη διαλειτουργικότητας με τις εξωτερικές εφαρμογές σημασιολογικού ιστού, ένα σχήμα συσσώρευσης μεταδεδομένων υλοποιήθηκε με βάση το EBUCore [2], το οποίο είναι ένα καθιερωμένο πρότυπο στον τομέα των οπτικοακουστικών μεταδεδομένων, όπως έχει αναφερθεί στο Παραδοτέο 1.4.1. Η επιλογή έγινε μετά από εκτεταμένη αξιολόγηση των εναλλακτικών προτύπων στον τομέα αυτό (MPEG7, DCMI, TV Anytime) [3].

Το EBUCore περιλαμβάνει ένα ελάχιστο κατάλογο των χαρακτηριστικών που περιγράφουν πόρους ήχου και βίντεο για ένα ευρύ φάσμα ραδιοτηλεοπτικών εφαρμογών

² <http://euscreen.eu/>

συμπεριλαμβανομένων της αρχειοθέτησης, της ανταλλαγής και της δημοσίευσής τους. Είναι βασισμένο στο πρότυπο Dublin Core για να μεγιστοποιηθεί η διαλειτουργικότητά του με την κοινότητα των χρηστών του Dublin Core. Το EBUCore επεκτείνει τον κατάλογο των στοιχείων που ορίζονταν αρχικά στο EBU Tech 3293-2001 για ραδιοφωνικά αρχεία, που επίσης βασιζόταν στο Dublin Core. Το εργαλείο ενσωμάτωσης πολιτιστικών αποθετηρίων MINT³ χρησιμοποιήθηκε και εδώ για τη λήψη και τη μετατροπή των μεταδεδομένων. Όπως έχει αναφερθεί, το MINT είναι μια web-based πλατφόρμα που έχει ως στόχο να βοηθήσει στη αντιστοίχιση των υπάρχοντων μεταδεδομένων του εκάστοτε παρόχου με βάση το προτεινόμενο μοντέλο μεταδεδομένων.

Το επόμενο βήμα, μετά τη μετατροπή των μεταδεδομένων του περιεχομένου σε ένα κοινό μοντέλο, είναι η δημοσίευσή τους στο Διαδίκτυο ως Διασυνδεδεμένα Δεδομένα (Linked Data). Τα Διασυνδεδεμένα Δεδομένα έχουν ως στόχο να καταστήσουν προσβάσιμα τα δεδομένα όχι μόνο στους ανθρώπους αλλά και στους πράκτορες λογισμικού, δημιουργώντας με αυτόν τον τρόπο ένα σημασιολογικό στρώμα για τη βελτίωση και τον εμπλουτισμό της αλληλεπίδρασης των χρηστών με τις πληροφορίες. Για την επίτευξη του στόχου αυτού απαιτείται η μετατροπή των μεταδεδομένων σε αναπαράσταση RDF χρησιμοποιώντας ένα εκφραστικό μοντέλο δεδομένων. Η οντολογία EBUCore είναι το καταλληλότερο μοντέλο για να οδηγήσει αυτό το σημασιολογικό μετασχηματισμό. Έτσι, πραγματοποιείται ο εμπλουτισμός των μεταδεδομένων με σύνδεσή τους με εσωτερικές και εξωτερικές πηγές και το αποτέλεσμα αποθηκεύεται σε μια σημασιολογική βάση δεδομένων (triplestore) που υποστηρίζει απαντήσεις σε ερωτήματα με τη μορφή SPARQL, που έχουμε αναλύσει στο Παραδοτέο 1.4.2 και στο παρόν Παραδοτέο.

³ <http://mint.image.ece.ntua.gr/>

2. Σημασιολογική αναζήτηση σε ψηφιακές βιβλιοθήκες

2.1 Το σύστημα σημασιολογικής αναζήτησης

Στο Σχήμα 2.1 απεικονίζεται ένα σύστημα, το οποίο επιτρέπει τη σημασιολογική ανάλυση των ερωτημάτων των χρηστών εκμεταλλευόμενο τόσο τα μεταδεδομένα όσο και την οπτική αναπαράσταση των αντικειμένων μιας ψηφιακής βιβλιοθήκης.

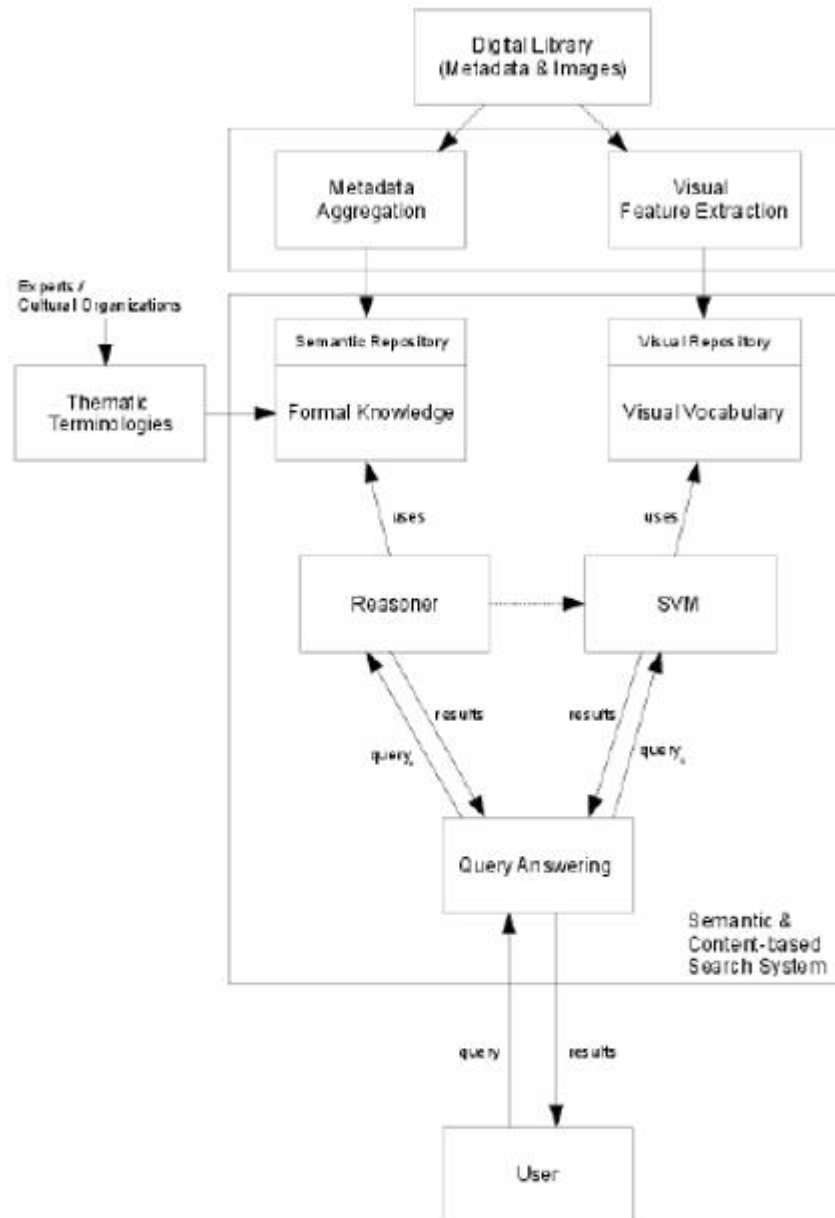
Μια ψηφιακή βιβλιοθήκη αποτελείται είτε από μια ενιαία πύλη, όπως στην περίπτωση της Europeana, είτε από ένα δίκτυο παροχής περιεχομένου, όπως πιθανόν θα συμβεί με την εξελισσόμενη ψηφιακή δημόσια βιβλιοθήκη της Αμερικής. Στην περίπτωση της Europeana, η διαλειτουργικότητα κατά τη διαδικασία των αναζητήσεων των χρηστών επιτυγχάνεται μέσω της ανάπτυξης και της χρησιμοποίησης ενός κοινού μοντέλου μεταδεδομένων (Europeana Data Model (EDM)) που είναι συμβατό με το Dublin Core, με το πρότυπο METS που χρησιμοποιείται από τις βιβλιοθήκες, με το LIDO που χρησιμοποιείται από τα μουσεία, το EAD που χρησιμοποιείται από τα αρχεία και το EBU Core που χρησιμοποιείται από τα οπτικοακουστικά αρχεία. Τα στοιχεία του μοντέλου είναι περιγραφές των αντικειμένων που παρέχουν πληροφορίες για αυτά, ξεκινώντας από τις απαντήσεις στα `ποιος;`, `τι;`, `πότε;` και `πού;`. Επιπλέον συνήθως δίνεται το url της εικόνας κάθε ψηφιακού αντικειμένου μαζί με μια εικόνα του. Όταν οι χρήστες υποβάλλουν ένα ερώτημα για ένα αντικείμενο, γίνεται αναζήτηση στα μεταδεδομένα όλων των αντικειμένων της βάσης και όποτε υπάρχει αντιστοιχία το αντικείμενο συμπεριλαμβάνεται στα αποτελέσματα. Στο σύστημά μας, τα μεταδεδομένα συσσωρεύονται και αναπαρίστανται ως τριπλέτες RDF (οι οποίες αποτελούν την τυπική γνώση) στα πλαίσια της οντολογίας EDM. Στη συνέχεια αυτές οι τριπλέτες αποθηκεύονται στη Σημασιολογική Αποθήκη (Semantic Repository).

Εάν θέλουμε να επιτρέπουμε στους χρήστες να υποβάλλουν σύνθετες ερωτήσεις και να λαμβάνουν κατάλληλες απαντήσεις, χρειαζόμαστε μια πιο λεπτομερή περιγραφή του πολιτιστικού περιεχομένου υπό τη μορφή θεματικών ορολογιών (Thematic Terminologies). Στο παρόν Παραδοτέο, δείχνουμε ότι όποτε είναι διαθέσιμη τέτοια γνώση, μπορούμε να τη χρησιμοποιήσουμε για να απαντήσουμε στα ερωτήματα του χρήστη βασιζόμενοι όχι μόνο στο 'ταίριασμα φράσης' ανάμεσα στο ερώτημα και στα ψηφιακά μεταδεδομένα των βιβλιοθηκών, αλλά και στην υπονοούμενη γνώση που μπορεί να εξαχθεί με συλλογιστική [5,6], χρησιμοποιώντας τη θεματική ορολογία (η οποία

παρέχει λεπτομέρειες σχετικά με τα είδη, τις κατηγορίες, ιδιότητες, και τις αμοιβαίες σχέσεις των αντικειμένων (π.χ. οι πόρπες κατασκευάζονται από χαλκό ή χρυσό)). Η τελευταία γνώση (οι θεματικές ορολογίες) αναπτύσσεται από εμπειρογνώμονες ή τους παρόχους περιεχομένου, και αποθηκεύεται στο Τυπικό Υποσύστημα Γνώσης (Formal Knowledge Subsystem).

Παρ'όλα αυτά η δημιουργία γενικών αξιωμάτων που ισχύουν για όλα τα αντικείμενα της ψηφιακής βιβλιοθήκης είναι δύσκολη. Μία δυνατή προσέγγιση είναι η χρήση μόνο των αξιωμάτων, τα οποία ισχύουν για όλα τα δεδομένα και όχι αυτών που ισχύουν για ένα ποσοστό των δεδομένων. Σε κάθε περίπτωση η έμφυτη ή συνεπής μη πληρότητα (incompleteness) της τυπικής γνώσης επιβάλλει περιορισμούς στη χρήση της για σημασιολογική αναζήτηση και απάντηση ερωτημάτων πολιτισμικού περιεχομένου.

Στη μέθοδο που περιγράφουμε δείχνουμε ότι αυτό το πρόβλημα μπορεί εν μέρει να ξεπεραστεί αν χρησιμοποιήσουμε επιπλέον τις εικόνες των τεκμηρίων των ψηφιακών βιβλιοθηκών. Η αναζήτηση με βάση το περιεχόμενο στις εικόνες της κάθε ψηφιακής βιβλιοθήκης βελτιώνει την λίστα αποτελεσμάτων του χρήστη και συμπληρώνει την αναζήτηση που βασίζεται μόνο στα μεταδεδομένα. Όπως φαίνεται στο Σχήμα 2.1, η διαδικασία εξαγωγής οπτικών χαρακτηριστικών και δημιουργίας της οπτικής αποθήκης (Visual Repository) υλοποιείται παράλληλα με τη διαδικασία συσσώρευσης μεταδεδομένων. Στη συνέχεια δημιουργούνται κατάλληλα οπτικά λεξικά, τα οποία τροφοδοτούν τεχνικές μηχανικής μάθησης για οπτική αναζήτηση.



Σχήμα 2.1: Η αρχιτεκτονική του συστήματος σημασιολογικής αναζήτησης

2.2 Ο αλγόριθμος σημασιολογικής αναζήτησης

Για την αναπαράσταση των θεματικών ορολογιών χρησιμοποιείται η OWL 2 (που είναι το πρότυπο W3C για οντολογική αναπαράσταση στο διαδίκτυο) [9]. Η θεωρητική βάση της OWL 2 είναι οι περιγραφικές λογικές (Description Logics-DL) [4]. Συγκεκριμένα, η OWL 2 στηρίζεται στην περιγραφική λογική SROIQ.

Όπως έχει αναφερθεί και στο Παραδοτέο 1.4.2, τα δομικά συστατικά μιας βάσης γνώσης περιγραφικής λογικής είναι οι ατομικές έννοιες, οι ατομικοί ρόλοι και τα άτομα, τα οποία είναι στοιχεία των μετρήσιμων και ξένων συνόλων **C,R,I**, αντίστοιχα. Μία βάση γνώσης περιγραφικής λογικής (KB) δηλώνεται ως $K = \langle T, A \rangle$, όπου T είναι το σώμα ορολογίας (terminological knowledge-TBox) που αναπαριστά τις οντότητες του πεδίου ενδιαφέροντος και A είναι το σώμα ισχυρισμών (assertional knowledge-ABox), η οποία περιγράφει τα αντικείμενα του κόσμου στο πλαίσιο των παραπάνω οντοτήτων. Τυπικά, το T είναι ένα σύνολο αξιωμάτων ορολογίας της μορφής $C_1 \sqsubseteq C_2$ ή $R_1 \sqsubseteq R_2$, όπου οι C_1, C_2 είναι SROIQ περιγραφές εννοιών και οι R_1, R_2 είναι SROIQ περιγραφές ρόλων. Οι σύνθετες έννοιες στη SROIQ δημιουργούνται χρησιμοποιώντας σύζευξη, ή διάζευξη, ($C_1 \sqcap C_2$), ($C_1 \sqcup C_2$)), καθολικούς και υπαρξιακούς περιορισμούς τιμής ($\forall R.C$, $\exists R.C$), περιορισμούς πληθυσμότητας ($\geq R.C$, $\leq R.C$), ονοματικές έννοιες ($\{a\}$), ενώ η εκφραστικότητα σύνθετων SROIQ ρόλων επιτρέπει τον ορισμό ανάστροφων ρόλων (R^-) και τον ορισμό της σύνθεσης ρόλων ($R_1 \circ R_2$) στην αριστερή πλευρά αξιωμάτων υπαγωγής ρόλων. Το TBox T περιγράφει τους περιορισμούς του πεδίου που αναπαριστάται. Το ABox A είναι ένα πεπερασμένο σύνολο ισχυρισμών της μορφής $A(a)$ ή $R(a,b)$, όπου $a, b \in \mathbf{I}, A \in \mathbf{C}, R \in \mathbf{R}$.

Μια διερμηνεία I αντιστοιχεί έννοιες σε υποσύνολα του πεδίου αντικειμένων, ρόλους σε ζευγάρια στοιχείων του πεδίου αντικειμένων και άτομα σε στοιχεία του πεδίου αντικειμένων. Για να θεωρηθεί μια ερμηνεία μοντέλο μιας βάσης γνώσης πρέπει να πληρούνται κάποιες συνθήκες [10]. Αν ένα αξίωμα ax ικανοποιείται για κάθε μοντέλο μιας βάσης γνώσης K , τότε η K συνεπάγεται το αξίωμα ax (που γράφεται $K \models ax$). Οι έλεγχοι συνεπαγωγής γίνονται από κατάλληλα εργαλεία, τα οποία καλούνται μηχανές συλλογιστικής (Reasoners).

Στη συνέχεια μελετούμε τα εννοιολογικά ερωτήματα (concept queries). Ένα εννοιολογικό ερώτημα q είναι της μορφής $q : Q(x) \leftarrow \bigwedge_{i=1..n} C_i(x)$, όπου όπου x είναι μία

μεταβλητή και $C_i(x)$ είναι κατηγορήματα/εννοιολογικά άτομα. Ένα παράδειγμα ενός εννοιολογικού ερωτήματος είναι το $Q(x) \leftarrow OpenVase(x) \wedge VaseWithTwoHandles(x)$, το οποίο αποτελείται από δυο εννοιολογικά άτομα. Ένα άτομο a αποτελεί απάντηση ενός εννοιολογικού ερωτήματος q που τίθεται σε μια βάση γνώσης K αν $K \models Q(a)$.

Η διαδικασία που ακολουθούμε για την εύρεση απαντήσεων σε εννοιολογικά ερωτήματα φαίνεται στον Αλγόριθμο 1.

Αλγόριθμος 1 Απάντηση ερωτημάτων με χρήση μιας βάσης γνώσης

Input: $K(T, A)$: the SROIQ knowledge base DL ontology

q : a concept query

Output: Ans : the set of answers to query q

```
1:  $Ans := \emptyset$ 
2:  $C^1, \dots, C^n := queryAtomsOf(q)$ 
3: for  $j=1, \dots, n$  do
4:   for all individual  $a \in A$  do
5:     if  $K \models C_j(a)$  then
6:        $Ans := Ans \cup a$ 
7:     else
8:       if  $a \in Ans$  then
9:          $Ans := Ans \setminus a$ 
10:      end if
11:    end if
12:  end for
13: end for
14: return  $Ans$ 
```

Ο αλγόριθμος δέχεται ως είσοδο την τυπική γνώση K και το ερώτημα του χρηστή q και επιστρέφει τα αποτελέσματα, δηλαδή τα άτομα της βάσης γνώσης που ικανοποιούν το ερώτημα. Αυτό γίνεται διατρέχοντας επαναληπτικά όλα τα εννοιολογικά άτομα C_j του ερωτήματος q και όλα τα άτομα a που εμφανίζονται στη βάση γνώσης K ελέγχοντας αν η K συνεπάγεται ότι το a είναι ένα στιγμιότυπο της έννοιας C_j . Αν η K συνεπάγεται ότι το a είναι στιγμιότυπο, τότε προσθέτουμε το άτομο a στο σύνολο των απαντήσεων Ans , αλλιώς πρέπει να ελέγξουμε αν το είναι ήδη στην Ans . Αν είναι, το αφαιρούμε από το σύνολο, αλλιώς αφήνουμε το σύνολο ανέπαφο.

3. Εξαγωγή οπτικών χαρακτηριστικών και δημιουργία οπτικού λεξικού

3.1 Εξαγωγή MPEG-7 χαρακτηριστικών

Τα MPEG-7 χαρακτηριστικά αποτελούνται από ένα σύνολο περιγραφών για ήχο, χρώμα, υφή και σχήμα, τα οποία έχουν χρησιμοποιηθεί ευρέως για κατηγοριοποίηση εικόνων, ανίχνευση εννοιών υψηλού επιπέδου και ανάκτηση εικόνων/ακολουθιών \cite{CHA01}.

Στην ανάλυσή μας χρησιμοποιούμε περιγραφείς χρώματος, σχήματος και υφής. Συγκεκριμένα, τον Dominant Color Descriptor (DCD - ένα σύνολο κυρίαρχων χρωμάτων σε μία περιοχή ή στο σύνολο της εικόνας), τον Color Structure Descriptor (CSD - ένα σύνολο συνολικών και τοπικών χαρακτηριστικών χρώματος της εικόνας), τον Color Layout Descriptor (CLD - έναν αναλλοίωτο σε ανάλυση περιγραφέα της χωρικής κατανομής στον χώρο YCbCr), τον Scalable Color Descriptor (SCD - έναν μετασχηματισμό Haar στις τιμές του ιστογράμματος χρώματος της εικόνας), τον Region-Based Shape descriptor (RSD - έναν περιγραφέα της χωρικής κατανομής των εικονοστοιχείων μέσα σε μία περιοχή βασισμένο στο περίγραμμά της, αλλά και στην εσωτερική διάταξή της), τον Homogeneous Texture Descriptor (HTD - έναν ποσοτικό περιγραφέα υφής βασισμένο σε κατευθυντικά φίλτρα) και τον Edge Histogram Descriptor (EHD - έναν περιγραφέα κατανομής των ακμών).

Στο Σχήμα 3.1 φαίνονται μερικές εικόνες της Europeana οι οποίες περιλαμβάνονται στις απαντήσεις στο ερώτημα χρήστη `Κοσμήματα της ύστερης βυζαντινής περιόδου' (`Jewelry of the Late Byzantine period') και περιέχουν σταυρούς και μωσαϊκά. Στο ίδιο Σχήμα φαίνονται και οι αντίστοιχες μάσκες τμηματοποίησης (segmentation masks), οι οποίες παράγονται μετά από κατάλληλη κατωφλίωση και εφαρμογή ενδιάμεσου φίλτρου. Ακολουθεί η εξαγωγή των MPEG-7 περιγραφών σχήματος. Οι τιμές του Region Shape descriptor για όμοιες και ανόμοιες εικόνες φαίνονται στον Πίνακα 3.1. Οι τιμές είναι μεγάλες για τις ανόμοιες εικόνες του Σχήματος 3.2 και μικρές για τις όμοιες. Εκμεταλλευόμενοι αυτόν τον οπτικό συσχετισμό εικόνων μπορούμε να ανακτήσουμε όμοιες εικόνες των οποίων όμως τα μεταδεδομένα διαφέρουν σημαντικά, όπως συμβαίνει με τους σταυρούς του Σχήματος 3.1.

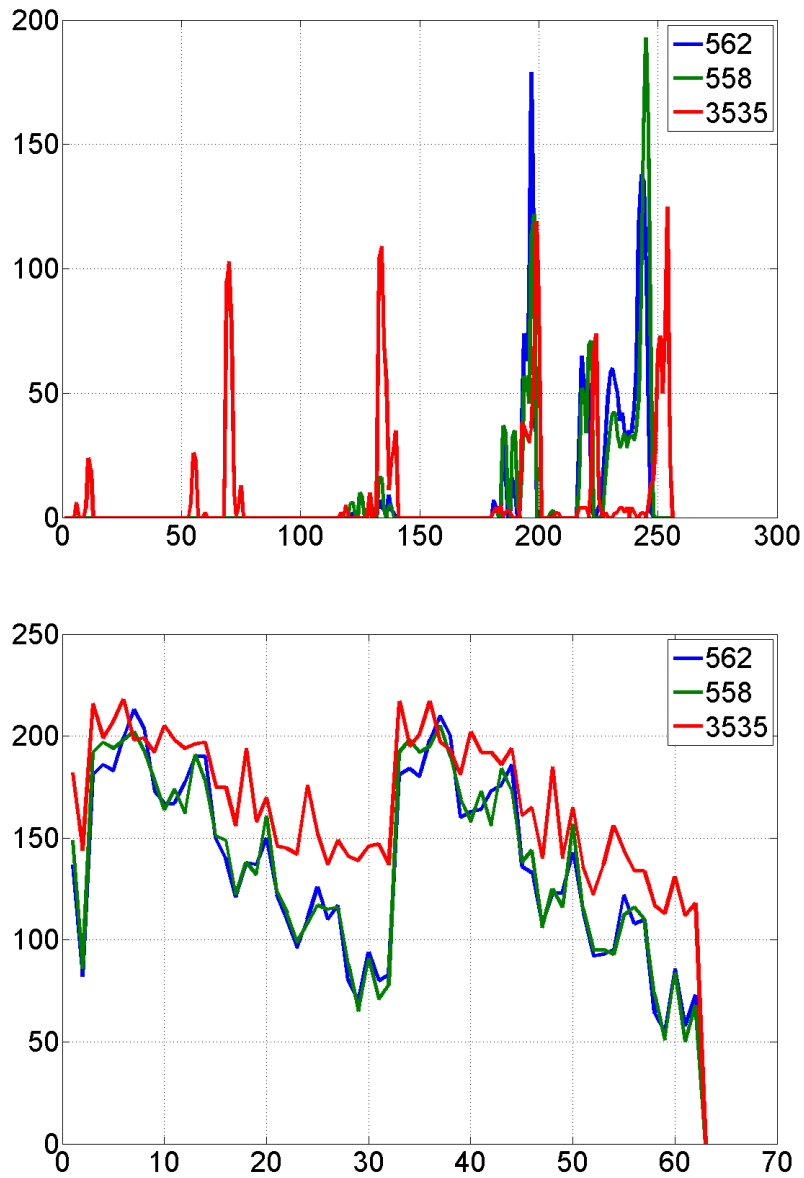
Στο Σχήμα 3.2 απεικονίζονται τα ιστογράμματα των περιγραφών Color Structure και Homogeneous Texture descriptor για τρεις από τις εικόνες του Σχήματος 3.1. Δύο από αυτές τις εικόνες είναι όμοιες (562, 558) ενώ η τρίτη είναι διαφορετική (3535). Το Σχήμα 3.3 απεικονίζει αντικείμενα που ανήκουν στην κατηγορία 'στολίδι του λαιμού' ('brooch'), τα οποία είναι φτιαγμένα είτε από χρυσό είτε από χαλκό. Οι δυαδικές εικόνες των αντικειμένων μετά την κατωφλίωση χρησιμοποιούνται για την εξαγωγή περιγραφών χρώματος, ώστε να διαφοροποιηθούν τα αντικείμενα βάσει του μετάλλου που είναι κατασκευασμένα. Οι κατανομές φαίνονται στο Σχήμα 3.4. Στα μεταδεδομένα των αντικειμένων δεν περιέχεται πάντα πληροφορία για το υλικό που χρησιμοποιήθηκε. Με την κατηγοριοποίηση επομένως μπορούμε να τα εμπλουτίσουμε με επιπλέον πληροφορία.

Πίνακας 3.1: Πίνακας σύγκρισης (απόσταση L1) του περιγραφέα Region Shape

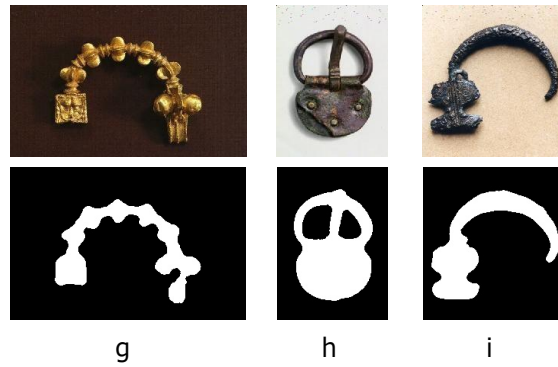
image id	(e)	(f)	(a)	(b)
(e)	0	85	181	171
(f)	85	0	154	148
(a)	181	154	0	120
(b)	171	148	120	0



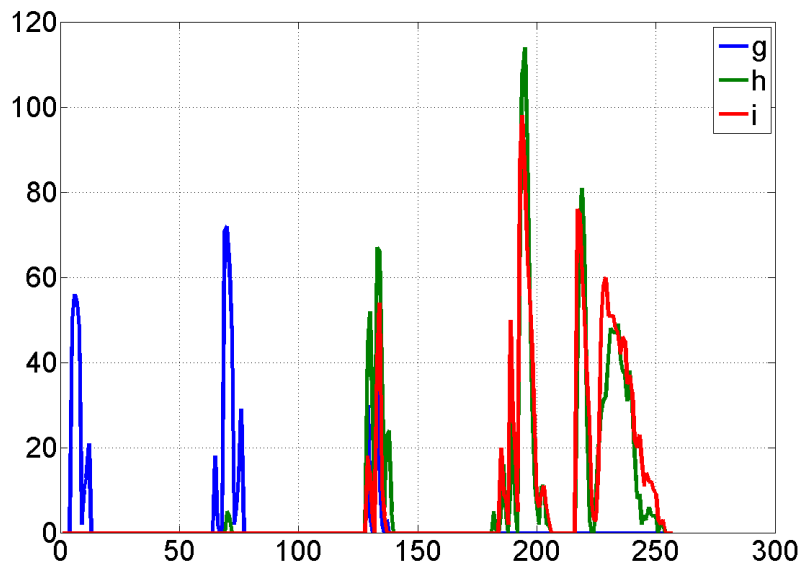
Σχήμα 3.1: Εικόνες από την βάση και οι αντίστοιχες μάσκες τμηματοποίησης.



Σχήμα 3.2: Συγκρίσεις ιστογραμμάτων των περιγραφέων MPEG-7 (CST: πάνω, HT: κάτω). Δύο ιστογράματα αντιστοιχούν σε παρόμοιες εικόνες, ενώ το τρίτο σε ανόμοιες.



Σχήμα 3.3: Εικόνες και οι αντίστοιχες μάσκες τμηματοποίησης. g: χρυσός, h-i: χαλκός.

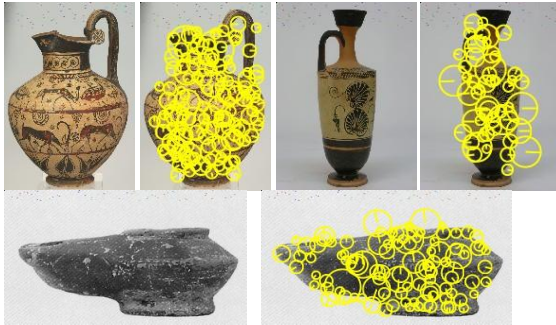


Σχήμα 3.4: Σύγκριση ιστογραμμάτων του CST περιγραφέα για τις δυαδικές μάσκες του Σχήματος 3.3.

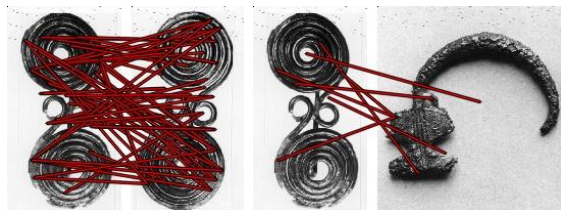
3.2 Εξαγωγή τοπικών χαρακτηριστικών και δημιουργία οπτικού λεξικού

Μεταξύ των πιο διαδεδομένων τοπικών χαρακτηριστικών και περιγραφών είναι οι αφινικά αναλλοίωτες περιοχές [51], δηλαδή οι περιοχές που `ακολουθούν' τον αφινικό μετασχηματισμό της εικόνας και είναι σχετικά αναλλοίωτες σε αλλαγή οπτικής γωνίας και επικάλυψη. Σύμφωνα με την μελέτη των Mikolajczyk et al [42], οι πιο αποδοτικοί είναι οι ανιχνευτές $\{M\}$ aximally $\{S\}$ table $\{E\}$ xtremal $\{R\}$ egion (MSER) [41] και Hessian-affine [42].

Η ανάγκη για καλύτερη ισορροπία μεταξύ υπολογιστικής πολυπλοκότητας και απόδοσης οδήγησε σε μία ποικιλία τοπικών ανιχνευτών, οι οποίοι βασίζονται στην κατανομή φωτεινότητας και της παραγώγου της. Σε μια προσπάθεια να βελτιωθεί η ακρίβεια ανίχνευσης οι Rapantzikos et al πρότειναν έναν ανιχνευτή βασισμένο σε ακμές [48], ο οποίος ανιχνεύει σύνθετες περιοχές ξεκινώντας από κατωφλιωμένες ακμές. Σε μια πιο πρόσφατη εργασία οι Avrithis et al [31] πρότειναν τον Medial Feature Detector (MFD), ο οποίος βασίζεται στον υπολογισμό της σταθμισμένης απόστασης από παραγώγους της εικόνας. Στις προσπάθειες για υπολογιστικά πιο αποδοτικούς ανιχνευτές ανήκουν οι CenSurE [53], FAST [56], BRIEF [52], με τον SURF (Speeded-Up Robust Features) [32] να παραμένει ο πιο διαδεδομένος. Ο ανιχνευτής SURF βασίζεται στα integral images για γρήγορο φιλτράρισμα στον χώρο κλίμακας-χώρου και σε μία προσέγγιση του πίνακα Hessian. Υιοθετούμε τα SURF για την εξαγωγή χαρακτηριστικών από τις εικόνες της ψηφιακής βιβλιοθήκης λόγω της καλής ισορροπίας μεταξύ πολυπλοκότητας και απόδοσης. Το Σχήμα 3.5 απεικονίζει αγγεία και τα αντίστοιχα τοπικά χαρακτηριστικά, τα οποία εστιάζονται σε περιοχές έντονης υψής (corner-like) και ομαλού σχήματος (blob-like).



Σχήμα 3.5: Ενδεικτικές εικόνες αγγείων και τα αντίστοιχα χαρακτηριστικά SURF. Η κλίμακα των χαρακτηριστικών και η κυρίαρχη κατεύθυνση της περιοχής φαίνεται με τους κίτρινους κύκλους.



Σχήμα 3.6: Αντιστοιχίσεις τοπικών χαρακτηριστικών μεταξύ (α) όμοιων και (β) ανόμοιων εικόνων. Οι όμοιες εικόνες μοιράζονται πολλές οπτικές λέξεις.

Για την αξιόπιστη αναπαράσταση των εικόνων από σύνολα τοπικών χαρακτηριστικών χρησιμοποιείται ευρέως το μοντέλο Bag-of-Words (BoW) ή -συνώνυμα- το οπτικό λεξικό (visual vocabulary) σαν ισοδύναμο του γλωσσικού λεξικού, με μία εικόνα να αντιστοιχεί σε ένα αρχείο κειμένου. Οι περιγραφείς εξάγονται από τις περιοχές των τοπικών χαρακτηριστικών και κβαντίζονται για να κατασκευαστεί το λεξικό, το οποίο τελικά χρησιμοποιείται για να περιγράψει την εικόνα με το σύνολο των οπτικών λέξεων που περιέχει [50,35].

Για την κατασκευή του οπτικού λεξικού κάνουμε συσταδοποίηση των περιγραφέων και αντιστοιχίζουμε κάθε χαρακτηριστικό στο κοντινότερο κέντρο συστάδας (οπτική λέξη). Πρέπει να σημειωθεί ότι λόγω της πολυσημίας οι οπτικές λέξεις δεν μπορούν να έχουν την ακρίβεια των λέξεων φυσικής γλώσσας. Πρακτικά, αυτό σημαίνει ότι μία οπτική λέξη

δεν μπορεί να συσχετιστεί με μία και μόνο συγκεκριμένη έννοια, αλλά με περισσότερες. Γενικά, τέτοιου είδους λάθη αντισταθμίζονται με ένα στάδιο χωρικής επαλήθευσης (spatial verification) των ταιριασμένων χαρακτηριστικών (matched features).

Τυπικά, το οπτικό λεξικό κατασκευάζεται με τον αλγόριθμο k-means [40]. Παρόλα αυτά όταν ο αριθμός των χαρακτηριστικών είναι πολύ μεγάλος, τότε η συσταδοποίηση με τον απλό k-means είναι πολύ αργή. Χρησιμοποιούμε μία γρήγορη παραλλαγή του k-means, η οποία βασίζεται σε αναζήτηση προσεγγιστικά κοντινότερου γείτονα (approximate nearest neighbor search), δηλαδή τα κέντρα κοντινότερων συστάδων προσδιορίζονται με randomized kd-trees [49]. Συγκεκριμένα, χρησιμοποιούμε την βιβλιοθήκη FLANN των Muja και Lowe [45]. Έχοντας την προηγούμενη αναπαράσταση μπορούμε να αναπαραστήσουμε κάθε εικόνα με το σύνολο των οπτικών λέξεων που περιέχει με την μορφή ενός ιστογράμματος σταθερού μήκους που περιέχει τις συχνότητες εμφάνισης κάθε λέξης. Αυτό είναι το BoW-ιστόγραμμα, ένα διάνυσμα N_v διαστάσεων $H_v(I)$ της εικόνας I :

$$H_v(I) = [tf_I(0), tf_I(1), \dots, tf_I(N_v)] \quad (3.1)$$

όπου $tf_i(i)$ δηλώνει τη συχνότητα εμφάνισης της οπτικής λέξης i στην εικόνα I .

Ένα παράδειγμα με εικόνες πολιτιστικής κληρονομιάς φαίνεται στο Σχήμα 3.6.

Το διάνυσμα $H_v(I)$ αποτελεί το πρώτο μέρος της εισόδου στον κατηγοριοποιητή SVM της αρχιτεκτονικής που περιγράφεται, ενώ το δεύτερο αποτελείται από τα MPEG-7 χαρακτηριστικά.

4. Συνδυάζοντας τη σημασιολογική αναζήτηση με μηχανική μάθηση των οπτικών χαρακτηριστικών

Στο Σχήμα 2.1 χρησιμοποιούμε τεχνικές μηχανικής μάθησης (ειδικότερα εκπαιδεύουμε μηχανές διανυσμάτων στήριξης (Support Vector Machines (SVMs))), οι οποίες λαμβάνουν υπόψιν τους τα οπτικά χαρακτηριστικά που έχουμε εξάγει, για την ταξινόμηση των αντικειμένων της ψηφιακής βιβλιοθήκης σε διάφορες έννοιες που εμφανίζονται σε ερωτήματα. Οι τεχνικές αυτές μπορούν να καθορίζουν τα άτομα της βάσης γνώσης που ικανοποιούν τα εννοιολογικά άτομα του ερωτήματος, ανεξάρτητα από το αν αυτά τα άτομα έχουν βρεθεί ως απαντήσεις ερωτημάτων από τη Μηχανή Συλλογιστικής (Reasoner) που εκμεταλλεύεται την επίσημη γνώση του συστήματος. Αυτό βοηθάει στη γεφύρωση του χάσματος μεταξύ των περιορισμών που επιβάλλονται από τις οντολογίες και των πραγματικών περιορισμών (δηλ. των περιορισμών που επιβάλλονται από τα οπτικά χαρακτηριστικά) κάθε στοιχείου της πολιτιστικής κληρονομιάς.

Οι Μηχανές Διανυσμάτων Στήριξης αποτελούν μια πολύ γνωστή μέθοδο η οποία βασίζεται σε συναρτήσεις (συναρτήσεις πυρήνα) προκειμένου να δημιουργήσει αποτελεσματικά κατάλληλους ταξινομητές. Οι ταξινομητές αυτοί αντιστοιχίζουν τα οπτικά χαρακτηριστικά, καθώς και τα αντίστοιχα αντικείμενα, σε ένα διανυσματικό χώρο μεγαλύτερης διάστασης (χώρος ενσωμάτωσης), στον οποίο τα αντικείμενα μπορούν να διακριθούν μέσω ενός γραμμικού ταξινομητή. Μπορούμε να χρησιμοποιήσουμε ένα SVM για κάθε έννοια. Το διάνυσμα εισόδου του SVM περιλαμβάνει $N_1 + N_2$ στοιχεία, τα N_1 εκ των οποίων είναι τα SURF χαρακτηριστικά του οπτικού λεξικού και τα N_2 τα επιλεγμένα MPEG-7 χαρακτηριστικά. Κάθε τύπος χαρακτηριστικών έχει κανονικοποιηθεί κατάλληλα. Ως εκ τούτου, οι MPEG-7 περιγραφείς είναι κανονικοποιημένοι χρησιμοποιώντας την L_1 ή L_2 νόρμα (ανάλογα με τον τύπο τους [34]) και τα SURF χρησιμοποιώντας τη L_1 νόρμα. Τέλος, όλα τα χαρακτηριστικά κανονικοποιούνται σύμφωνα με τη L_1 νόρμα και τροφοδοτούν το SVM.

Η οπτική αναζήτηση στοχεύει στη βελτίωση της ακρίβειας της σημασιολογικής αναζήτησης. Για το λόγο αυτό, χρησιμοποιούμε τις σωστές απαντήσεις που παρέχονται από τη μηχανή συλλογιστικής, η οποία χρησιμοποιεί την τυπική γνώση, ως δείγματα για την εκπαίδευση του κάθε SVM, το οποίο στη συνέχεια χρησιμοποιείται για τον έλεγχο της

απόδοσής του σε νέες εισόδους. Αυτή η αλληλεπίδραση μεταξύ της μηχανής συλλογιστικής και των SVMs φαίνεται στο Σχήμα 2.1 με ένα διακεκομμένο βέλος που υπάρχει από τη μηχανή συλλογιστικής (reasoner) προς το SVM.

Ο πυρήνας που χρησιμοποιείται στο SVM για την κωδικοποίηση της οπτικής γνώσης μέσω της ομοιότητας μεταξύ διαφορετικών εικόνων είναι μια κανονικοποιημένη γραμμική συνάρτηση που ορίζεται ως εξής [12]:

$$k_l(x, y) = (x^T y + c) / (\|x\| \|y\|) \quad (4.1)$$

όπου x, y είναι τα διανύσματα των χαρακτηριστικών, $\| \cdot \|$ είναι η Ευκλείδεια νόρμα και το c θεωρείται ίσο με μηδέν.

Το Υποσύστημα Απάντησης Ερωτημάτων (Query Answering Subsystem) συγχωνεύει τις εξόδους των δύο ενοτήτων, δηλαδή, της μηχανής συλλογιστικής (reasoner) και του SVM.

Είναι δυνατόν να επεκταθεί ο πυρήνας του SVM, έτσι ώστε να συμπεριλάβει τη γνώση που αφορά στα άτομα της τυπικής γνώσης [15,17] και, κατά συνέπεια, μόνο η έξοδος του SVM να τροφοδοτεί το Υποσύστημα Απάντησης Ερωτημάτων.

Η επέκταση προέρχεται από μια οικογένεια συναρτήσεων πυρήνα που ορίζονται ως

$$k_p^F = \text{Ind}(A) \times \text{Ind}(A) \rightarrow [0, 1] \quad (4.2)$$

για μια βάση γνώσης $K = \langle T, A \rangle$. Το $\text{Ind}(A)$ είναι το σύνολο των ατόμων που εμφανίζονται στο ABox A και $F = \{F_1, F_2, \dots, F_m\}$ είναι ένα σύνολο περιγραφέντων εννοιών. Αυτές οι συναρτήσεις ορίζονται ως L_p μέσος, έστω m , απλών συναρτήσεων πυρήνα εννοιών $k_i, i = 1, \dots, m$ ανάμεσα σε κάθε δύο άτομα a, b .

Η ανωτέρω συνάρτηση πυρήνα κωδικοποιεί την τυπική γνώση για το υπό εξέταση πρόβλημα μέσω της ομοιότητας ζευγών ατόμων με βάση υψηλού επιπέδου χαρακτηριστικά, δηλαδή των εννοιών της βάσης γνώσης.

Οι πυρήνες αυτοί στηρίζονται στο γεγονός ότι η ομοιότητα μεταξύ των ατόμων καθορίζεται από την ομοιότητά τους σε σχέση με κάθε έννοια F_i , δηλαδή, εάν δύο αντικείμενα είναι στιγμιότυπα της έννοιας ή της άρνησής της. Λόγω της υπόθεσης ανοιχτού κόσμου των περιγραφικών λογικών για την υποκείμενη σημασιολογία, μια πιθανή αβεβαιότητα στην συμμετοχή του ατόμου στην έννοια αντιπροσωπεύεται από μία ενδιάμεση τιμή του πυρήνα. Η τιμή $\rho = 1$ χρησιμοποιείται γενικά στις εφαρμογές.

Η επέκταση που μπορούμε να χρησιμοποιήσουμε είναι ένας συνδυασμένος πυρήνας SVM, που υπολογίζεται ως η μέση τιμή των ανωτέρω περιγραφέντων πυρήνων, δηλαδή,

$$k_c(a, b) = k_p^F(a, b) + k_l(a, b) \quad (4.3)$$

όπου $k_p^F(a, b)$ είναι ο ανωτέρω, καθοδηγούμενος από τη γνώση πυρήνας και $k_l(a, b)$ είναι ο κανονικοποιημένος γραμμικός πυρήνας.

Στη συνέχεια, θα σκιαγραφήσουμε τον τρόπο με τον οποίον απαντώνται τα ερωτήματα, χρησιμοποιώντας τα SVMs που έχουν ήδη εκπαιδευτεί για να ταξινομήσουν τα αντικείμενα της ψηφιακής βιβλιοθήκης σε έννοιες.

Ο αλγόριθμος 2 δείχνει τη διαδικασία. Ο αλγόριθμος δέχεται ως είσοδο τα αντικείμενα στα οποία θέλουμε να γίνει αναζήτηση, μαζί με τα οπτικά χαρακτηριστικά τους και χρησιμοποιεί εκπαιδευμένα SVMs για να ελέγξει ποιά άτομα/αντικείμενα ανήκουν ταυτόχρονα σε όλες τις έννοιες που εμφανίζονται στο ερώτημα. Το SVMpredict προβλέπει την ετικέτα ενός ατόμου ως προς την έννοια C_j χρησιμοποιώντας το SVM που έχει εκπαιδευτεί για να ταξινομή τα άτομα σε αυτή την έννοια.

Ο αλγόριθμος 3 δείχνει πώς μπορούμε να συνδυάσουμε τις δύο προσεγγίσεις, δηλ. τη βασισμένη στη γνώση και τη βασισμένη στον πυρήνα προσέγγιση, ούτως ώστε να εμπλουτίσουμε με τον τρόπο αυτό τα αποτελέσματα της αναζήτησης. Η μέθοδος `extractFeatures` εξάγει τα χαρακτηριστικά των εικόνων, ενώ οι μέθοδοι `evaluate_{KB}` και `evaluate_{SVM}` εκτελούν τους αλγορίθμους 1 και 2 για να εξάγουν τις απαντήσεις των ερωτημάτων από τις δύο μεθόδους. Στη συνέχεια λαμβάνουμε σαν απαντήσεις του ερωτήματος την ένωση των απαντήσεων από τις δύο μεθόδους και τις παρέχουμε στο χρήστη.

Αλγόριθμος 2 Διαδικασία Αξιολόγησης Ερωτήματος με χρήση οπτικών χαρακτηριστικών

Input: *trained SVM* : ένα διάνυσμα εκπαιδευμένων SVM

Input: *data* : τα ερωτήματα

Input: *features* : τα χαρακτηριστικά των *data*

q : ένα εννοιολογικό ερώτημα

Output: *Ans*: το σύνολο των απαντήσεων στο ερώτημα *q*

```
1: Ans := 0
2:  $C^1, \dots, C^n := queryAtomsOf(q)$ 
3: for  $j=1, \dots, n$  do
4:   for all  $a \in data$  do
5:      $output(a) := SVMpredict(features(a), trained\ SVM_j)$ 
6:     if  $output(a) = 1$  then
7:        $Ans := Ans \cup a$ 
8:     else
9:       if  $output(a) = 0$  then
10:         $Ans := Ans \setminus a$ 
11:      end if
12:    end if
13:  end for
14: end for
15: return Ans
```

Αλγόριθμος 3 Συνδυασμένη Διαδικασία Αξιολόγησης Ερωτήματος

Input: *KB* : η SROIQ βάση γνώσης

Input: *SVM* : το διάνυσμα των εκπαιδευμένων SVM

Input: *data* : τα ερωτήματα

q : ένα εννοιολογικό ερώτημα *q*

Output: *Ans*: το σύνολο των απαντήσεων στο ερώτημα *q*

```
1: Ans := 0
2:  $Ans_{KB} := Evaluate_{KB}(KB, q)$ 
3:  $features := ExtractFeatures(data)$ 
4:  $Ans_{SVM} := Evaluate_{SVM}(SVM, data, features, q)$ 
5:  $Ans := Ans_{KB} \cup Ans_{SVM}$ 
6: return Ans
```

5. Μελέτη αξιολόγησης: Βελτίωση της σημασιολογικής αναζήτησης μέσω της οπτικής ανάλυσης

5.1 Πειραματική μελέτη

Η πειραματική μελέτη που παρουσιάζεται σε αυτή την ενότητα έχει ως στόχο να παρουσιάσει τη βελτίωση που επιτυγχάνεται αξιοποιώντας τόσο τη σημασιολογική, όσο και τη με βάση το περιεχόμενο οπτική αναζήτηση των τεκμηρίων μιας ψηφιακής βιβλιοθήκης κατά την απάντηση στα ερωτήματα των χρηστών. Επικεντρωνόμαστε στην Europeana, διότι αποτελεί ένα πραγματικό περιβάλλον, εύκολα προσβάσιμο από όλους μέσω της διαδικτυακής του πύλης <http://www.europeana.eu>. Το περιεχόμενο που εξετάζουμε είναι, από τη μία πλευρά, ελληνικό περιεχόμενο, που αποτελείται από περίπου 40.000 αντικείμενα και, από την άλλη πλευρά, μουσειακό περιεχόμενο που έχει συγκεντρωθεί μέσω των προγραμμάτων ATHENA και ATHENAPLUS [30] και αποτελείται από περίπου 3.600.000 αντικείμενα της Europeana. Τα μεταδεδομένα του περιεχομένου αυτού μετατράπηκαν σε EDM OWL. Τα πειράματα που παρουσιάζονται βασίζονται στο ελληνικό περιεχόμενο, αφού για αυτό το περιεχόμενο διαθέτουμε και αντίστοιχη θεματική γνώση. Θα προβούμε σε αναφορά στο περιεχόμενο των προγραμμάτων ATHENA και ATHENAPLUS, όταν θα ασχοληθούμε με θέματα κλιμάκωσης.

Η θεματική γνώση που χρησιμοποιήσαμε για τα ελληνικά μνημεία, και ειδικότερα για τα αγγεία (για τα οποία παρέχονται τα μεταδεδομένα και οι εικόνες τους), έχει δημιουργηθεί στο πλαίσιο των έργων Polemon και Ψηφιοποίηση των Συλλογών Κινητών Μνημείων του Υπουργείου Πολιτισμού της Διεύθυνσης του Εθνικού Αρχείου Μνημείων <http://nam.culture.gr> και η οποία έχει συμπεριληφθεί στην ορολογία Πολυδεύκης στο Θησαυρό Αρχαιολογικών Συλλογών και Μνημείων [28,29].

Αυτή η γνώση περιέχει αξιώματα σχετικά με αγγεία στην αρχαία Ελλάδα, δηλαδή, αξιώματα ιεραρχίας κλάσεων που αφορούν τα διάφορα είδη των αγγείων, όπως αμφορείς, αλάβαστρο, κρατήρας, καθώς και αξιώματα σχετικά με την εμφάνιση, τη χρήση, την περίοδο δημιουργίας και τα υλικά από τα οποία κατασκευάστηκε καθένα από τα αγγεία αυτά. Περιέχει, δε, 55 κατηγορίες πολιτιστικών αγαθών (όπως

DARIAH-ΑΤΤΙΚΗ Ανάπτυξη της ελληνικής ερευνητικής υποδομής για τις ανθρωπιστικές επιστήμες ΔΥΑΣ

κεραμικά, κοσμήματα, γραμματόσημα, τοιχογραφίες, χαρακτικά, νομίσματα) και περισσότερα από 300 είδη. Ένα απόσπασμα της σχετικής ορολογίας που αφορά τους τύπους των αγγείων φαίνεται στον Πίνακα 5.1.

Με βάση αυτή τη γνώση και την ανάλυση των σχετικών εικόνων, στόχος μας είναι να απαντήσουμε σε ερωτήματα χρηστών που αναζητούν π.χ. την Κεραμικά της Μυκηναϊκής περιόδου, μινωικά αγγεία με θαλασσινή διακόσμηση, κοσμήματα της Ελληνιστικής περιόδου, μολυβδόβουλα (σφραγίδες του βασιλιά) της Μέσης και Ύστερης Βυζαντινής περιόδου, νομίσματα της Ύστερης Βυζαντινής περιόδου, ανοιχτά αγγεία, αγγεία με δύο λαβές, ανοιχτά αγγεία με δύο λαβές.

Κατά συνέπεια, το TBox της τυπικής γνώσης του συστήματός μας αποτελείται από την οντολογία EDM μαζί με τη θεματική οντολογία. Το ABox αποτελείται από στιγμιότυπα της EDM οντολογίας, καθένα από τα οποία περιγράφεται από τον τύπο του, την ημερομηνία δημιουργίας του, το υλικό του και το μουσείο στο οποίο μπορεί να βρεθεί. Δημιουργούνται περίπου 1.000.000 τριάδες RDF από τα 40.000 είδη της Europeana, οι οποίες αποθηκεύονται σε μια αποθήκη Sesame.

Εκτός από τα μεταδεδομένα, εξάγονται και τα οπτικά χαρακτηριστικά των πολιτιστικών αντικειμένων σύμφωνα με τη μεθοδολογία που παρουσιάστηκε στην ενότητα [6].

Η εφαρμογή της μεθοδολογίας για την εξαγωγή χαρακτηριστικών SURF και τη δημιουργία οπτικού λεξικού απαιτεί τη ρύθμιση κάποιων παραμέτρων που χρησιμοποιούνται σε αυτή. Συγκεκριμένα, ο κατά προσέγγιση αριθμός των χαρακτηριστικών (ρυθμισμένου κατάλληλα μέσω ενός κατωφλίου) έχει επιλεγεί να είναι 500, ενώ ο αριθμός των οπτικών λέξεων έχει οριστεί ως μια ρυθμιζόμενη παράμετρος, η οποία μεταβάλλεται στην περιοχή από 100-4000, σύμφωνα με την σχετική βιβλιογραφία. Όσον αφορά στα MPEG-7 χαρακτηριστικά, είτε χρησιμοποιούνται αμφότερα τα χαρακτηριστικά χρώματος και υφής, παρέχοντας ένα διάνυσμα με 666 στοιχεία, ή μόνο τα χαρακτηριστικά υφής (χρώματος), παρέχοντας ένα διάνυσμα με 142 (524) στοιχεία, αντίστοιχα.

DARIAH-ΑΤΤΙΚΗ Ανάπτυξη της ελληνικής ερευνητικής υποδομής για τις ανθρωπιστικές επιστήμες ΔΥΑΣ

Πίνακας 5.1: Απόσπασμα από τη θεματική οντολογία που χρησιμοποιείται, εκπεφρασμένη σε Περιγραφικές Λογικές

Amphora \sqsubseteq *BigVase* \sqcap *CloseVase*
Alabaster \sqsubseteq *VaseWithoutHandles*
Crater \sqsubseteq \exists *hasBase.NarrowBase*
Pycnometer \sqsubseteq \exists *hasBody.CylindricalBody*
Amphora \neq *Alabaster*
Bowl \sqsubseteq *OpenVase*
EnclosedProduct \sqsubseteq *Solid* \sqcup *Liquid*
Solid \neq *Liquid*
DrinkingLiquid \sqsubseteq *Liquid*
Water \sqsubseteq *DrinkingLiquid*
Wine \sqsubseteq *DrinkingLiquid*
Oil \sqsubseteq *Liquid*
Perfume \sqsubseteq *Liquid*
Cereal \sqsubseteq *Solid*
Grain \sqsubseteq *Solid*
Usage \equiv *Carrying* \sqcup *Storing* \sqcup *Drinking*
 \exists *contains*⁻ . \top \sqsubseteq *EnclosedProduct*
 \exists *isUsedFor*⁻ . \top \sqsubseteq *Usage*
Alabaster \sqsubseteq \exists *isUsedFor.Carrying* \sqcap \exists *contains*⁻*(Oil* \sqcup *Perfume)*
Amphora \sqsubseteq \exists *isUsedFor.Carrying* \sqcup \exists *isUsedFor.Storing*
Aryballos \sqsubseteq \exists *isUsedFor.Storing*
Aryballos \sqsubseteq \exists *contains.Perfume*
Cup \sqsubseteq \exists *isUsedFor.Drinking*
Lecythus \sqsubseteq \exists *isUsedFor.Storing* \sqcap \exists *contains*⁻*.(Perfume* \sqcup *Oil)*
Pithos \sqsubseteq \exists *isUsedFor.Storing* \sqcap \exists *contains*⁻*.(Oil* \sqcup *Cereal* \sqcup *Grain)*
Hydria \sqsubseteq \exists *isUsedFor.Carrying* \sqcap \exists *contains.Water*
Vase \sqcap \exists *isUsedFor.Storing* \sqsubseteq *StorageVase*
Vase \sqcap \exists *madeIn.ArchaicPeriod* \sqsubseteq *ArchaicVase*
ArchaicVase \sqcap *Amphora* \sqsubseteq *ArchaicAmphora*
 \exists *isUsedFor.Storing* \sqcap \exists *contains.Liquid* \sqsubseteq *LiquidStorageVase*

5.2 Τα αποτελέσματα της απάντησης ερωτημάτων

Στη συνέχεια εφαρμόζουμε την προσέγγιση Απάντησης Ερωτημάτων που φαίνεται στο Σχήμα 2.1, εστιάζοντας στις τεχνικές που περιγράφονται στις ενότητες 2 και 3 για σημασιολογική απάντηση ερωτημάτων και το συνδυασμό της με τις εξαγόμενες οπτικές πληροφορίες των παραπάνω αντικειμένων της Europeana.

Κατά την υλοποίηση χρησιμοποιούμε τη μηχανή συλλογιστικής Hermit [6,7] και τη βιβλιοθήκη LIBSVM (<http://www.csie.ntu.edu.tw>) για την εξαγωγή συμπερασμάτων χρησιμοποιώντας τη βάση γνώσης και για την εκμάθηση των οπτικών περιγραφών, αντίστοιχα.

Ας εξετάσουμε, τώρα, τρεις από τις αναζητήσεις των χρηστών που αναφέρονται στην προηγούμενη ενότητα και ειδικότερα αυτές που εμφανίζονται στην πρώτη στήλη του Πίνακα 5.2, δηλαδή, τα ερωτήματα για αγγεία με δύο λαβές (VaseWithTwoHandles), ή ανοικτά αγγεία (OpenVase), ή ανοικτά αγγεία με δύο λαβές (OpenVaseWithTwoHandles).

Πίνακας 5.2: Ακρίβεια (%) της Απάντησης Ερωτημάτων

Ερώτημα	Αλγόριθμος 1	Αλγόριθμος 2	Αλγόριθμος 3
$Q(x) \leftarrow OpenVase(x)$	85.5	78.4	96.9
$Q(x) \leftarrow VaseWithTwoHandles(x)$	84.6	52.7	92.1
$Q(x) \leftarrow OpenVase(x)$ $\wedge VaseWithTwoHandles(x)$	90.2	42.4	93.2

Η δεύτερη στήλη αναφέρεται στην ακρίβεια της διαδικασίας απάντησης ερωτημάτων, που ορίζεται ως ο λόγος του αριθμού των αληθώς θετικών και αληθώς αρνητικών προς όλα τα αντικείμενα ελέγχου, όταν χρησιμοποιείται ο Αλγόριθμος 1 της ενότητας 2. Η βάση γνώσης που λαμβάνει ο αλγόριθμος αυτός ως είσοδο περιλαμβάνει τόσο τη γνώση ορολογίας που περιγράφεται στην προηγούμενη ενότητα, όσο και τα αντικείμενα της Europeana. Περίπου 4.000 αντικείμενα από τα 40.000 ανήκουν στην έννοια βάζο, με περίπου 2.200 να είναι ανοικτά αγγεία, 1.800 να είναι αγγεία με δύο λαβές και με περίπου 900 να ανήκουν και στις δύο κατηγορίες. Η υπόθεση αλήθειας (ground truth) για όλα αυτά τα στοιχεία έχει δημιουργηθεί από εμπειρογνώμονες στον πολιτιστικό τομέα.

Μπορεί εύκολα να επαληθευθεί ότι και στα τρία ερωτήματα, η γνώση που χρησιμοποιείται για τον ορισμό των ανοικτών αγγείων ή/και αγγείων με δύο λαβές

DARIAH-ΑΤΤΙΚΗ Ανάπτυξη της ελληνικής ερευνητικής υποδομής για τις ανθρωπιστικές επιστήμες ΔΥΑΣ

αντιπροσωπεύει περίπου το 85% των περιπτώσεων, με αποτέλεσμα ένα σφάλμα της τάξεως του 10 έως 15%.

Ας εξετάσουμε τώρα τη χρήση της αναζήτησης με βάση το περιεχόμενο για βελτίωση της ακρίβειας της παραπάνω διαδικασίας απάντησης ερωτήματος. Ειδικότερα, πρώτα εκπαιδεύουμε τα SVMs (ένα για κάθε μία από τις δύο διαφορετικές έννοιες των ερωτημάτων), χρησιμοποιώντας τα εξαγόμενα οπτικά χαρακτηριστικά των αντικειμένων που επιστρέφονται ως απαντήσεις των ερωτημάτων από τη βάση γνώσης. Τα SVMs αυτά εκπαιδεύονται χρησιμοποιώντας τον κανονικοποιημένο γραμμικό πυρήνα που περιγράφεται στην ενότητα [\ref{sec:combination}](#) με βάση τα οπτικά χαρακτηριστικά και τις ετικέτες που έχουν δοθεί στις εικόνες για όλα τα αντικείμενα που σωστά έχουν επιστραφεί ως απαντήσεις του εκάστοτε ερωτήματος από τη μηχανή συλλογιστικής που χρησιμοποιήθηκε προηγουμένως στον Αλγόριθμο 1. Και στις τρεις περιπτώσεις, τα SVMs μαθαίνουν να ταξινομούν σωστά όλα τα δεδομένα που παρέχονται προς εκπαίδευση.

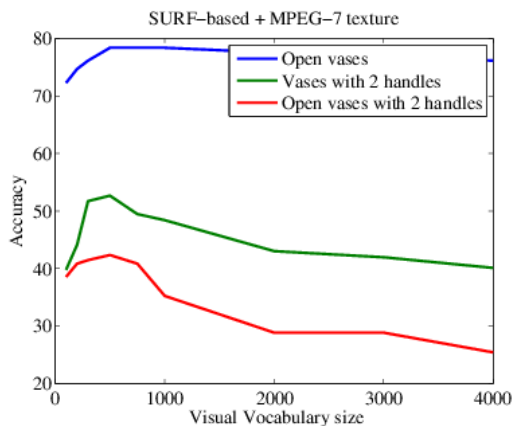
Ως εισόδοι στα SVMs μπορούν να δοθούν διάφοροι συνδυασμοί των οπτικών εισόδων που αναφέρθηκαν παραπάνω, ήτοι: α) μόνο τα MPEG-7 χαρακτηριστικά, με αποτέλεσμα ένα διάνυσμα εισόδου 666 στοιχείων, β) μόνο το με βάση τα SURF οπτικό λεξικό, με τον αριθμό των οπτικών λέξεων να κυμαίνεται μεταξύ 100 και 4000, γ) όλες οι MPEG-7 και οπτικές λέξεις, δ) ο συνδυασμός των χαρακτηριστικών υψής του MPEG-7 (με 142 στοιχεία) και οπτικών λέξεων, το οποίο και αποδείχθηκε ότι παρέχει τα καλύτερα αποτελέσματα.

Μετά την εκπαίδευση, ελέγξαμε - σύμφωνα με τον Αλγόριθμο 2 - την απόδοση των SVMs για τα υπόλοιπα δεδομένα (περίπου 10-15% και στις τρεις περιπτώσεις), τα οποία δεν έχουν επιστραφεί ως απαντήσεις του εκάστοτε ερωτήματος σύμφωνα με τη γνώση.

Πραγματοποιούμε μία ποικιλία από πειράματα με τις περιπτώσεις (α) - (δ) που αναφέρονται παραπάνω και αναλύουμε τις επιδόσεις των SVMs όσον αφορά: i) στην ακρίβεια (accuracy), ii) στην ακρίβεια (precision) (που ορίζεται ως ο λόγος των αληθώς θετικών προς το σύνολο των επιστρεφόμενων από τα SVM απαντήσεων) και iii) στην ανάκληση (recall) (που ορίζεται ως ο λόγος των αληθώς θετικών προς τη θεωρητικά σωστή απάντηση).

DARIAH-ΑΤΤΙΚΗ Ανάπτυξη της ελληνικής ερευνητικής υποδομής για τις ανθρωπιστικές επιστήμες ΔΥΑΣ

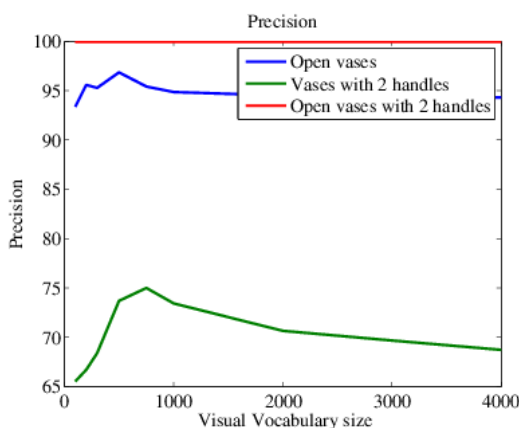
Το Σχήμα 5.1 παρουσιάζει την ακρίβεια που λαμβάνουμε και στα τρία ερωτήματα για την προαναφερθείσα περίπτωση (δ), καθώς μεταβάλλεται ο αριθμός των οπτικών λέξεων, που δείχνει ότι τα καλύτερα αποτελέσματα ελήφθησαν χρησιμοποιώντας ένα λεξικό 500 οπτικών λέξεων. Η ακρίβεια είναι υψηλότερη στην κατηγορία ανοικτών βάζων και χαμηλότερη σε αυτή των βάζων με δύο λαβές.



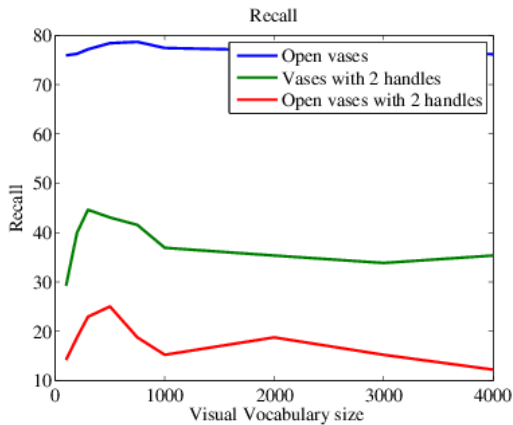
Σχήμα 5.1: Ακρίβεια της μεθόδου σε σχέση με το μέγεθος του οπτικού λεξικού

Το Σχήμα 5.2 εμφανίζει τις αντίστοιχες τιμές ακρίβειας (precision) και ανάκλησης για ένα μέγεθος λεξικού μεταξύ 100 και 4.000 λέξεων, που επιβεβαιώνει τα αποτελέσματα που προκύπτουν από το Σχήμα 5.1.

Για αυτόν τον αριθμό οπτικών λέξεων (500), ο Πίνακας 5.3 δείχνει μία σύγκριση της απόδοσης που λαμβάνεται από κάθε ένα από τα πειράματα δοκιμής (α) - (δ), επαληθεύοντας ότι ο συνδυασμός των οπτικών λέξεων και των MPEG-7 χαρακτηριστικών υφής παρέχει τα καλύτερα δυνατά αποτελέσματα, όπως αποδεικνύεται από όλα τα σχετικά κριτήρια.



DARIAH-ΑΤΤΙΚΗ Ανάπτυξη της ελληνικής ερευνητικής υποδομής για τις ανθρωπιστικές επιστήμες ΔΥΑΣ



Σχήμα 5.2: Ακρίβεια και Ανάκληση για τις τρεις κατηγορίες περιεχομένου

Η τρίτη στήλη του Πίνακα 5.2 συνοψίζει την ακρίβεια της απάντησης ερωτήματος όταν χρησιμοποιούμε τον Αλγόριθμο 2 της ενότητας 4. Με άλλα λόγια η στήλη αυτή δείχνει το ποσοστό των αντικειμένων που προβλέπουν σωστά τα SVMs ως απαντήσεις ερωτήματος μεταξύ εκείνων που δεν είχαν προβλεφθεί ως τέτοια από την γνώση.

Η τέταρτη στήλη του Πίνακα 5.2 αναφέρεται στην ακρίβεια της απάντησης ερωτημάτων όταν χρησιμοποιείται ο Αλγόριθμος 3 για να συνδυάσει τα αποτελέσματα με βάση τη γνώση και με βάση τους πυρήνες προσεγγίσεων. Βλέπουμε ότι η ακρίβεια του αλγορίθμου έχει αυξηθεί σε όλες τις περιπτώσεις. Αυτό απεικονίζει την βελτιωμένη απόδοση της σημασιολογικής αναζήτησης, όταν λαμβάνονται υπόψη οπτικές πληροφορίες με βάση τις διαδικασίες των ενοτήτων 2,3,4.

Για παράδειγμα, τα αγγεία με δύο λαβές ορίζονται στη βάση γνώσης έτσι ώστε να περιλαμβάνουν τις κατηγορίες: αμφορέας, σκαθάρι, μπουλ, κρατήρας, πελίκη. Παρόλα αυτά, υπάρχουν, επίσης, ορισμένα αγγεία σε άλλες κατηγορίες, όπως η κανάτα ή η λεκάνη που μπορεί να έχουν δύο λαβές. Αυτά τα αντικείμενα δε θα είναι μεταξύ των αποτελεσμάτων που θα επιστρέψει στο χρήστη η βασισμένη στη γνώση απάντηση ερωτήματος (Αλγόριθμος 1). Ακριβώς για αυτές τις περιπτώσεις, ο συνδυασμός της με τη μηχανική μάθηση των οπτικών χαρακτηριστικών (Αλγόριθμοι 2 και 3) βελτιώνει την απόδοση της σημασιολογικής υπηρεσίας αναζήτησης που προσφέρεται στους χρήστες. Στον Πίνακα 5.2, η πρώτη μέθοδος επιτυγχάνει απόδοση της τάξης του 85,5%, 84,6% και 90,2% στα ερωτήματα. Η

DARIAH-ΑΤΤΙΚΗ Ανάπτυξη της ελληνικής ερευνητικής υποδομής για τις ανθρωπιστικές επιστήμες ΔΥΑΣ

προσέγγιση συνδυασμού της γνώσης και της μηχανικής μάθησης αυξάνει την απόδοση στο 96,9%, 92,1% και 93,2%, αντίστοιχα.

Πίνακας 5.3: Σύγκριση της Απόδοσης των SVM για Διαφορετικά Διανύσματα Οπτικής Εισόδου

Οπτικά Χαρακτηριστικά	Accuracy			Precision			Recall		
	Q	Q	Q	Q1	Q2	Q3	Q1	Q2	Q3
μόνο SURF-based	55.6	51.6	42.3	93.48	71.79	100	54.43	43.08	26.15
μόνο MPEG-7	67.0	48.3	30.5	93.1	71.79	100	68.35	43.08	14.58
SURF-based + MPEG-7	64.7	49.4	30.5	92.86	72.5	100	65.82	41.54	14.58
SURF-based + MPEG-7 Υφή	78.4	52.6	42.3	96.88	73.68	100	78.48	43.08	25

5.3 Ζητήματα κλιμάκωσης: Μια οπτική μηχανή αναζήτησης για την Europeana (VIEU)

Ένα ζήτημα που αξίζει ιδιαίτερης προσοχής κατά την εφαρμογή της προτεινόμενης προσέγγισης είναι η επεκτασιμότητά της.

Έχει αποδειχθεί ότι η απάντηση ερωτημάτων σε OWL 2 βάσεις γνώσης είναι μια υπολογιστικά εντατική διαδικασία, που πάσχει από εξαιρετικά υψηλή πολυπλοκότητα. Αυτό είναι σίγουρα ένα σοβαρό ζήτημα, αν θέλουμε να εφαρμοστεί η μέθοδος σε μια μεγάλη ψηφιακή βιβλιοθήκη, όπως η Europeana, όπου ο αριθμός των τριπλετών RDF θα είναι της τάξης των εκατομμυρίων ή δισεκατομμυρίων. Η μεθοδολογία που μπορούμε να ακολουθήσουμε για να αντιμετωπίσουμε αυτό το πρόβλημα είναι: α) να χρησιμοποιήσουμε τις λειτουργίες του συστήματος αποθήκευσης τριπλετών για να παρέχουμε προσεγγιστικές απαντήσεις στα ερωτήματα [11] μειώνοντας έτσι το υπολογιστικό φορτίο, β) να επωφεληθούμε της εξαιρετικά αρθρωτής μορφής των μεταδεδομένων και των ορολογιών, έτσι ώστε να τις τεμαχίσουμε σε ένα σύνολο από σημαντικά μικρότερες ανεξάρτητες βάσεις γνώσης.

Ας επικεντρωθούμε, επίσης, και στην εξαγωγή των οπτικών χαρακτηριστικών, τη δημιουργία του οπτικού λεξικού και τη σύνδεση των διαφορετικών εικόνων με βάση τα οπτικά χαρακτηριστικά τους. Όπως φαίνεται στην προηγούμενη υποενότητα, η υλοποίηση της διαδικασίας για την εξαγωγή χαρακτηριστικών και οπτικού λεξικού

DARIAH-ΑΤΤΙΚΗ Ανάπτυξη της ελληνικής ερευνητικής υποδομής για τις ανθρωπιστικές επιστήμες ΔΥΑΣ

απαιτεί τη ρύθμιση κάποιων παραμέτρων που χρησιμοποιούνται σε αυτή. Συγκεκριμένα, πρέπει να καθορισθούν ο αριθμός των χαρακτηριστικών (ρυθμιζόμενος μέσω ενός κατωφλίου), καθώς και ο αριθμός των οπτικών λέξεων. Για να μπορέσουμε να χρησιμοποιήσουμε την διαδικασία σε ένα περιβάλλον μεγάλης κλίμακας μιας ψηφιακής βιβλιοθήκης, πραγματοποιήσαμε ένα μεγάλης κλίμακας πείραμα ανάκτησης, το οποίο βασίστηκε αποκλειστικά και μόνο στις εικόνες που περιλαμβάνονται στο περιεχόμενο των προγραμμάτων ATHENA και ATHENAPLUS που παρέχεται στην Europeana. Το εργαλείο που αναπτύχθηκε, δηλαδή η οπτική μηχανή αναζήτησης για την Europeana, μπορεί να αξιολογηθεί on-line στη διεύθυνση <http://vieu.image.ntua.gr>, συμπεριλαμβανομένου περίπου του 10% των εικόνων των αντικειμένων του ΑΘΗΝΑ, ήτοι περίπου 375.000 εικόνων από τη διαδικτυακή πύλη της Europeana.

Η εξαγωγή των τοπικών χαρακτηριστικών, η κατασκευή του οπτικού λεξικού και η εκχώρηση των οπτικών λέξεων έχουν πραγματοποιηθεί σε ένα offline βήμα για όλες τις εικόνες. Μετά από αυτό, η οπτική αναπαράσταση όλων των εικόνων έχει οργανωθεί σε ένα κατάλληλο ευρετήριο. Το οπτικό λεξικό έχει εξαχθεί χρησιμοποιώντας ένα σύνολο 10.000 εικόνων και το μέγεθός του έχει επιλεγεί μετά από μια διαδικασία trial-and-test, κατά την οποία έχει επιλεγεί το λεξικό με τις καλύτερες επιδόσεις. Ειδικότερα, για αυτό το μέγεθος της βάσης δεδομένων επιλέξαμε ένα οπτικό λεξικό των 30.000 συστάδων (με τον αριθμό των τοπικών χαρακτηριστικών να είναι 500 ανά εικόνα). Οι καλύτερα καταταγμένες εικόνες πέρασαν σε μια νέα φάση επανακατάταξης, όπου ελέγχθηκαν ως προς τη χωροταξική επιβεβαίωση. Τέλος, η χωρική αντιστοίχιση έγινε με το Fast Spatial Matching (FSM) [54] και/ή το Hough Pyramid Matching (HPM) [55].

Το αποτέλεσμα αυτής της διαδικασίας, που προσφέρει το σύστημα VIEU, είναι η ικανότητα να αποκτήσει κανείς δεσμούς (links) μεταξύ όσο το δυνατόν περισσότερων εικόνων με βάση την οπτική τους ομοιότητα. Αυτό μπορεί να βοηθήσει και να επιταχύνει τη διαδικασία απάντησης ερωτημάτων. Επιπλέον, το σύστημα VIEU μπορεί να επεκταθεί με τη χρήση ετικετών με άμεσο τρόπο. Όπως και στην περίπτωση του συστήματος Visual Image Retrieval and Localization (VIRaL) που βρίσκεται στη διεύθυνση <http://viral.image.ntua.gr>, ετικέτες που σχετίζονται με οπτικά όμοιες εικόνες μπορούν να μεταφερθούν στην υπό εξέταση εικόνα του ερωτήματος. Αυτές οι ετικέτες συνήθως συγκεντρώνονται απευθείας από την πηγή των εικόνων ή εμπλουτίζονται από εξωτερικούς χώρους, όπως π.χ., η Wikipedia.

6. Εμπλουτισμός μέσω διασυνδεδεμένων δεδομένων

Στην ενότητα αυτή παρουσιάζονται τα βήματα που προτείνονται και μπορούν να ακολουθηθούν για τον εμπλουτισμό του πολυμεσικού περιεχομένου και την δημοσίευσή του με τη μορφή Διασυνδεδεμένων Δεδομένων. Αρχικά κάνουμε μια σύντομη εισαγωγή στα Διασυνδεδεμένα Δεδομένα και στις αρχές που τα διέπουν. Αυτή ακολουθείται από αναλυτική περιγραφή της διαδικασίας παραγωγής RDF στιγμιοτύπων (instances) από τα μεταδεδομένα, μαζί με τον τρόπο που αυτά εμπλουτίζονται και συνδέονται με εξωτερικές πηγές.

6.1 Βασικές αρχές διασυνδεδεμένων δεδομένων

Τα τελευταία χρόνια ο Παγκόσμιος Ιστός έχει εξελιχθεί σε ένα χώρο πληροφοριών, στον οποίο συνδέονται τόσο τα έγγραφα όσο και τα δεδομένα. Η εξέλιξη αυτή είχε ως αποτέλεσμα ένα σύνολο βέλτιστων πρακτικών για τη δημοσίευση και τη σύνδεση δομημένων δεδομένων στον Παγκόσμιο Ιστό που είναι ορίζουν τη δομή των Διασυνδεδεμένων Δεδομένων. Ο όρος Διασυνδεδεμένα Δεδομένα αναφέρεται στον καθορισμό σχέσεων μεταξύ δικτυακών δεδομένων, προερχόμενων από μια ποικιλία πηγών. Τα δεδομένα αυτά μπορεί να προέρχονται από βάσεις δεδομένων που συντηρούνται από δύο οργανισμούς σε διαφορετικές γεωγραφικές περιοχές, ή από ετερογενή συστήματα εντός ενός οργανισμού που δεν παρέχουν διαλειτουργικότητα σε επίπεδο δεδομένων. Τεχνικά, τα Διασυνδεδεμένα Δεδομένα αναφέρονται σε δεδομένα που δημοσιεύονται στο Διαδίκτυο με τέτοιο τρόπο ώστε να είναι αναγνώσιμα από μηχανές (machine-readable), να διαθέτουν ρητά καθορισμένη σημασία, να μπορούν να εμπλουτίζονται μέσω της συνδέσής τους με άλλα εξωτερικά σύνολα δεδομένων και να επιτρέπουν σύνδεση σε αυτά από άλλα εξωτερικά σύνολα δεδομένων [58].

Η κύρια διαφορά του Παγκόσμιου Ιστού και των Διασυνδεδεμένων Δεδομένων είναι ότι ο πρώτος βασίζεται σε έγγραφα με αναπαράσταση HTML (HyperText Markup Language) που συνδέονται με υπερσυνδέσμους χωρίς τύπους, ενώ τα Διασυνδεδεμένα Δεδομένα βασίζονται σε έγγραφα που περιέχουν δεδομένα σε μορφή RDF (Resource Description Framework) [59]. Το αποτέλεσμα, ή όπως είναι γνωστό ευρέως, ο Σημασιολογικός Ιστός (Semantic Web), μπορεί να οριστεί με μεγαλύτερη ακρίβεια ως ένα δίκτυο πραγμάτων στον κόσμο που περιγράφονται από τα δεδομένα στο δίκτυο. Ο Berners-Lee το 2006 παρουσίασε τους παρακάτω

DARIAH-ΑΤΤΙΚΗ Ανάπτυξη της ελληνικής ερευνητικής υποδομής για τις ανθρωπιστικές επιστήμες ΔΥΑΣ

«κανόνες» για τη δημοσίευση δεδομένων στο Διαδίκτυο με τέτοιο τρόπο ώστε όλα τα δημοσιευμένα δεδομένα να γίνονται μέρος ενός ενιαίου παγκόσμιου χώρου δεδομένων:

1. Χρήση των URIs ως ονόματα για καθορισμό των πραγμάτων.
2. Χρήση των HTTP URIs έτσι ώστε οι άνθρωποι να μπορούν να αναζητήσουν αυτά τα ονόματα.
3. Όταν κάποιος αναζητεί ένα URI, να παρέχεται χρήσιμη πληροφορία σύμφωνα με το πρότυπο αναπαράστασης RDF και τη γλώσσα SPARQL.
4. Καθορισμός συνδέσμων με άλλα URIs, έτσι ώστε να μπορούν να ανακαλυφθούν περισσότερα πράγματα.

Τα παραπάνω έχουν γίνει γνωστά ως «Αρχές των Διασυνδεδεμένων Δεδομένων», και παρέχουν μια βασική συνταγή για τη δημοσίευση και τη σύνδεση δεδομένων χρησιμοποιώντας την υποδομή του Web και τηρώντας την αρχιτεκτονική και τα πρότυπά της.

6.2 Σημασιολογική αναπαράσταση πολυμεσικού περιεχομένου

Για την αναπαράσταση των μεταδεδομένων του πολυμεσικού περιεχομένου ως Διασυνδεδεμένων Δεδομένων, είναι απαραίτητη η δημιουργία μιας κατανοητής από μηχανές αναπαράστασης - σε RDF -, μια διαδικασία που συχνά αναφέρεται και ως RDFization στη κοινότητα του σημασιολογικού ιστού. Για να κατανοήσουμε πλήρως τον λόγο για τον οποίο αυτή η διαδικασία είναι σημαντική, πρέπει να εντοπίσουμε την ειδοποιό διαφορά μεταξύ των αναπαραστάσεων πολιτιστικού περιεχομένου XML και RDF. Στην προτεινόμενη προσέγγιση, η XML χρησιμοποιείται για τη συλλογή των μεταδεδομένων του περιεχομένου, ενώ η RDF χρησιμοποιείται για να ορίσει δηλώσεις σχετικά με τους πόρους (ιδίως τους πόρους Web) με τη μορφή εκφράσεων υποκειμένου-κατηγορήματος-αντικειμένου (τις λεγόμενες τριπλέτες - triples). Ως εκ τούτου, σε αντίθεση με τις διαδικασίες μετασχηματισμών XML κατά τις οποίες ένα έγγραφο XML μετατρέπεται σε ένα άλλο έγγραφο XML με διαφορετική δομή, κατά τη διάρκεια της διαδικασίας RDFization τα πράγματα που περιγράφονται στο έγγραφο XML πρέπει πρώτα να προσδιοριστούν σε συνδυασμό με τις σχετικές με αυτά δηλώσεις και στη συνέχεια να ακολουθήσει η δημιουργία του εγγράφου RDF.

Η RDF παρέχει ένα γενικό, αφηρημένο μοντέλο δεδομένων για την περιγραφή πόρων με τη χρήση της τριπλέτας: υποκείμενο, κατηγορημα, αντικείμενο. Ωστόσο, δεν παρέχει κάποιο όρο για την περιγραφή των κατηγοριών των πραγμάτων στον

DARIAH-ΑΤΤΙΚΗ Ανάπτυξη της ελληνικής ερευνητικής υποδομής για τις ανθρωπιστικές επιστήμες ΔΥΑΣ

κόσμο και πώς αυτά σχετίζονται μεταξύ τους. Αυτή η λειτουργία εξυπηρετείται από τα λεξιλόγια και τις οντολογίες, συνήθως εκφρασμένες σε Simple Knowledge Organization System (SKOS) [60], σε RDF Vocabulary Description Language, επίσης γνωστή ως RDF Schema (RDFS) [61] και σε Web Ontology Language (OWL) [62]. Στην προτεινόμενη υλοποίηση επιλέγεται η οντολογία EBUCore [63] ως το λεξιλόγιο που χρησιμοποιείται για την υποστήριξη της αναπαράστασης οπτικοακουστικών δεδομένων σε RDF. Ειδικότερα επικεντρωνόμαστε σε οπτικοακουστικό, και ειδικότερα σε τηλεοπτικό περιεχόμενο, που σε Ευρωπαϊκό επίπεδο συσσωρεύεται - για την Europeana - από το έργο EU-Screen (και τη συνέχεια του EUScreen-XL), με στόχο την εξερεύνηση της πλούσιας και ποικιλόμορφης οπτικοακουστικής κληρονομιάς της Ευρώπης. Το μέγεθος του περιεχομένου ανέρχεται σε πάνω από 50.000 προγράμματα και συνοδευτικές πληροφορίες, περιλαμβάνοντας περιεχόμενο πολύτιμο για πολιτιστικές, μορφωτικές και ψυχαγωγικές εφαρμογές. Το Ευρωπαϊκό έργο υλοποιείται από κοινοπραξία που αποτελείται από 30 παρόχους οπτικοακουστικού περιεχομένου και επιστήμονες που παρέχουν την ιστορική γνώση, τις τεχνικές προϋποθέσεις, τις νομικές εμπειρογνομοσύνες και τις εκπαιδευτικές εφαρμογές.

Σε συνεργασία με κορυφαίους ιστορικούς της τηλεόρασης και διεθνείς πρακτικές (Kaye, 2011), έγινε η επιλογή του πολυμεσικού περιεχομένου, που επικεντρώνεται στις παρακάτω θεματικές ενότητες:

1. Ιστορικά θέματα: 14 σημαντικά θέματα στην ιστορία της Ευρώπης του 20ού αιώνα (70% του περιεχομένου)
2. Συγκριτικές Εικονικές Εκθέσεις: δύο ειδικά θέματα που εξερευνούν πιο εξειδικευμένες πτυχές της ευρωπαϊκής ιστορίας με συγκριτικό τρόπο (10% του περιεχομένου - περιλαμβάνει έγγραφα, φωτογραφίες, άρθρα)
3. Εικονικές Εκθέσεις Παρόχων: Κάθε πάροχος περιεχομένου επιλέγει το περιεχόμενο που υποστηρίζεται με άλλα ψηφιακά υλικά και έγγραφες πληροφορίες σχετικά με τα ως άνω θέματα ή θέματα της επιλογής του (20% του περιεχομένου).

Η οντολογία EBUCore είναι μια αναπαράσταση RDF της EBU Class Conceptual Data Model (CCDM). Η CCDM ορίζεται ως ένα διαρθρωμένο σύνολο οπτικοακουστικών κατηγοριών, περιλαμβάνοντας ομάδες πόρων, πόρους πολυμέσων, επιμέρους τμήματα, αντικείμενα πολυμέσων, αλλά και περιοχές, γεγονότα, πρόσωπα και οργανισμούς. Η οντολογία EBUCore καθορίζει επίσης τις σημασιολογικές σχέσεις (objectProperties) μεταξύ αυτών των κατηγοριών, καθώς και τις ιδιότητες (dataProperties) που χαρακτηρίζουν αυτές τις κατηγορίες. Ένα μεγάλο μέρος της

DARIAH-ΑΤΤΙΚΗ Ανάπτυξη της ελληνικής ερευνητικής υποδομής για τις ανθρωπιστικές επιστήμες ΔΥΑΣ

γνώσης που περιλαμβάνονται στην EBU Core αποτελεί τη βάση για την ανάπτυξη της W3C οντολογίας επισημείωσης πολυμέσων (W3C MAWG).

Το επόμενο βήμα, μετά την επιλογή των κατάλληλων λεξιλογίων για την RDF αναπαράσταση του περιεχομένου, είναι η δημιουργία των προτάσεων (υποκείμενο –κατηγορημα – αντικείμενο) για τα πράγματα που περιγράφονται. Με άλλα λόγια, πρόκειται για την εκπλήρωση της πρώτης αρχής των Διασυνδεδεμένων Δεδομένων, που δηλώνει τη χρήση των URIs για τα αντικείμενα. Υπάρχουν διάφορες κατευθυντήριες γραμμές για τη δημιουργία URIs στον σημασιολογικό ιστό [64-66] και τα δύο βασικά χαρακτηριστικά που πρέπει να έχουν είναι: α) να είναι μοναδικές για κάθε αντικείμενο και β) να είναι συνεπείς. Σύμφωνα με αυτές τις κατευθυντήριες γραμμές κάθε οντότητα που εκπροσωπείται στο σύνολο των δεδομένων οδηγεί σε τουλάχιστον τρία URIs:

- Ένα URI για κάθε αντικείμενο του πραγματικού κόσμου.
- Ένα URI για κάθε σχετική πηγή πληροφοριών που περιγράφει το αντικείμενο του πραγματικού κόσμου και έχει μια HTML αναπαράσταση (dereferencable).
- Ένα URI για κάθε σχετική πηγή πληροφοριών που περιγράφει το αντικείμενο του πραγματικού κόσμου και έχει μια RDF/XML αναπαράσταση.

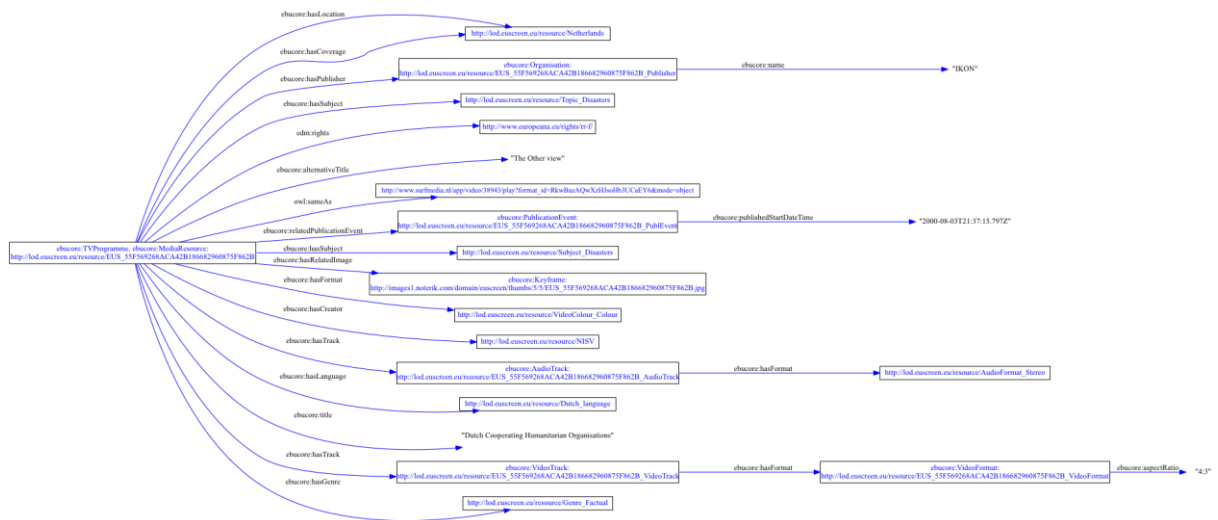
Για την εξασφάλιση της μοναδικότητας των URIs, τα υποκείμενα των πόρων (resources) εξυπηρετούνται από έναν διακομιστή - για το έργο EUScreen είναι ο (<http://lod.euscreen.eu>) - μαζί με τον αποκλειστικό αναγνωριστικό κωδικό που αποδίδεται στο αντικείμενο που αποτελεί μέρος του URI. Το αντίστοιχο σύνολο των URIs για ένα παράδειγμα πολυμεσικού περιεχομένου φαίνεται παρακάτω:

- http://lod.euscreen.eu/resource/EUS_55F569268ACA42B186682960875F862B
- http://www.euscreen.eu/play.html?id=EUS_55F569268ACA42B186682960875F862B
- http://lod.euscreen.eu/data/EUS_55F569268ACA42B186682960875F862B

Μετά τον καθορισμό του τρόπου δημιουργίας των URIs, τα πράγματα που περιγράφονται εντοπίζονται μαζί με τις κατάλληλες EBU Core κατηγορίες και τις ιδιότητες που θα χρησιμοποιηθούν για την RDF αναπαράστασή τους. Πιο συγκεκριμένα, ο τύπος ενός βίντεο καθορίζεται στα μεταδεδομένα του έγγραφο XML και μπορεί να είναι, ένα τηλεοπτικό πρόγραμμα, ένα έγγραφο, ένα

DARIAH-ΑΤΤΙΚΗ Ανάπτυξη της ελληνικής ερευνητικής υποδομής για τις ανθρωπιστικές επιστήμες ΔΥΑΣ

ραδιοφωνικό πρόγραμμα, ή μια εικόνα. Ανάλογα με αυτές τις πληροφορίες ο πόρος που δημιουργήθηκε για το βίντεο μπορεί να είναι ένα στιγμιότυπο των κατηγοριών/κλάσων TVProgramme, RadioProgramme, Document και Image αντίστοιχα του EBUCore₂. Τα επιπλέον χαρακτηριστικά των πόρων βίντεο μετατρέπονται σε RDF με τη χρήση ιδιοτήτων του EBUCore που έχουν ως εύρος, είτε τιμές κειμένου (typed literals - π.χ. αρχικός τίτλος, στο ebucore: originalTitle) ή άλλους εσωτερικούς πόρους (π.χ. για κάθε πάροχο βίντεο, ένας νέος πόρος δημιουργείται, ως στιγμιότυπο της ebucore: Agent). Στο Σχήμα 6.1 φαίνεται μια γραφική αναπαράσταση ενός τηλεοπτικού προγράμματος σε RDF⁴.



Σχήμα 6.1: Γραφική αναπαράσταση τηλεοπτικού προγράμματος σε RDF

Τέλος, μια άλλη σύσταση, η οποία είναι πολύ σημαντική και πρέπει να λαμβάνεται υπόψη κατά τη διάρκεια της δημοσίευσης Διασυνδεδεμένων Δεδομένων είναι η ιδιοκτησία και η προέλευση των πόρων των πληροφοριών. Ως εκ τούτου, για κάθε RDF αναπαράσταση ενός στοιχείου μεταδεδομένων περιλαμβάνονται πόροι προέλευσης που καθορίζουν την ημερομηνία δημοσίευσης και τον δημιουργό, έτσι ώστε οι καταναλωτές (consumers) να μπορούν να παρακολουθούν την προέλευση των συγκεκριμένων τμημάτων δεδομένων.

⁴ Το πλήρες σύνολο των ιδιοτήτων και κλάσεων που χρησιμοποιούνται για την αντιστοίχιση (mapping) του συνόλου στοιχείων όλων των σχημάτων που συλλέγονται (harvrsting) είναι διαθέσιμα στο: https://docs.google.com/spreadsheet/cc?key=0Akruw5a0_oaLdEQyMI85NVQxZ_2ImT00wcVU4ZVRJZ0E&hl=en_US#gid=3

6.3 Σημασιολογικός εμπλουτισμός

Όπως αναφέρθηκε ανωτέρω, τα Διασυνδεδεμένα Δεδομένα αφορούν στη χρήση του Web για τη δημιουργία συνδέσμων (links) μεταξύ δεδομένων από διαφορετικές πηγές, ως εκ τούτου, μετά την RDF αναπαράσταση κάποιου περιεχόμενου, δημιουργούνται συνδέσεις με άλλους πόρους. Υπάρχουν δύο είδη συνδέσμων, οι εσωτερικοί και εξωτερικοί RDF σύνδεσμοι. Οι εσωτερικοί σύνδεσμοι συνδέουν RDF πόρους μέσα σε ίδια πηγή Διασυνδεδεμένων Δεδομένων. Έτσι, τα URIs του υποκείμενου και του αντικείμενου είναι στον ίδιο χώρο ονομάτων (namespace). Οι εξωτερικοί σύνδεσμοι RDF συνδέουν πόρους που εξυπηρετούνται από διαφορετικές πηγές Διασυνδεδεμένων Δεδομένων. Τα URIs του υποκείμενου και του αντικείμενου των εξωτερικών συνδέσεων RDF βρίσκονται σε διαφορετικά namespaces. Οι εξωτερικές συνδέσεις RDF είναι ζωτικής σημασίας για τον Σημασιολογικό Ιστό καθώς με αυτό το τρόπο συνδέονται τα νησιά των δεδομένων σε ένα παγκόσμιο διασυνδεδεμένο χώρο δεδομένων παρέχοντας τη δυνατότητα εμπλουτισμού του συνόλου των δεδομένων.

Για τις εσωτερικές συνδέσεις, στην υλοποίησή μας χρησιμοποιούνται συγκεκριμένα στοιχεία του σχήματος συλλογής που συσχετίζουν αντικείμενα. Για παράδειγμα, η τιμή του στοιχείου του σχήματος συλλογής "isRelatedToItem" είναι ένα αναγνωριστικό στοιχείο. Αντίστοιχα, στην αναπαράσταση RDF η EBUCore ιδιότητα "isRelatedTo" χρησιμοποιείται έχοντας ως εύρος τον πόρο από το συγκεκριμένο στοιχείο. Επιπλέον, η πρόσθετη εσωτερική διασύνδεση χρησιμοποιείται για τις χώρες, τους φορείς και τις οργανώσεις. Η εξειδίκευση αυτών των πληροφοριών είναι το γεγονός που διαμοιράζεται μεταξύ του συνόλου δεδομένων. Με άλλα λόγια, μια χώρα μπορεί να είναι η τοποθεσία παραγωγής περισσότερων του ενός αντικειμένων βίντεο. Ως εκ τούτου, νέοι πόροι δημιουργούνται για αυτές τις τιμές στοιχείων χωρίς κανένα αναγνωριστικό κωδικό - μόνο, χρησιμοποιώντας το όνομα τους - αφού αυτοί είναι ήδη μοναδικοί. Π.χ. για την περίπτωση της Ολλανδίας, στο έργο EUScreen, δημιουργήθηκε ο κοινόχρηστος πόρος <http://lod.euscreen.eu/resource/Netherlands>. Ως εκ τούτου μπορεί να χρησιμοποιηθεί το αντικείμενο μιας τριπλέτας που έχει ως κατηγορηματή την EBUCore ιδιότητα "createdIn" και υποκείμενο τον πόρο του βίντεο (Σχήμα 6.1).

Οι πόροι που υλοποιούνται για τις χώρες συνδέονται επίσης εξωτερικά, δεδομένου ότι οι πληροφορίες σχετικά με τις χώρες εξυπηρετούνται από πολλές πηγές δεδομένων. Για τη δημιουργία των εξωτερικών συνδέσεων χρησιμοποιούμε

DARIAH-ΑΤΤΙΚΗ Ανάπτυξη της ελληνικής ερευνητικής υποδομής για τις ανθρωπιστικές επιστήμες ΔΥΑΣ

την DBpedia⁵. Τα ονόματα των χωρών των τοπικών συνόλων δεδομένων συγκρίνονται χρησιμοποιώντας τη SPARQL[57] με τα ονόματα των χωρών των πόρων που εξυπηρετούνται από τη DBpedia. Μετά τη δημιουργία μιας σύνδεσης στο DBpedia, επιπρόσθετοι πόροι Διασυνδεδεμένων Δεδομένων ανακαλύπτονται με χρήση της SPARQL. Με αυτόν τον τρόπο καταλήγουμε με εξωτερικές συνδέσεις προς όλες τις πηγές δεδομένων (Freebase, DBpedia, Eurostat, NYTimes) που συνδέεται η DBpedia.

Εκτός από αυτές τις συνδέσεις, νέες εξωτερικές συνδέσεις εξάγονται από τις περιλήψεις βίντεο με τη χρήση του DBpedia spotlight⁶, ενός εργαλείου που μπορεί να εξαγάγει πόρους από ελεύθερο κείμενο. Στη συνοπτική περιγραφή ενός βίντεο αναφέρονται, αρκετά συχνά, τα ονόματα προσώπων που είτε συμμετέχουν στο βίντεο, ή το βίντεο τους εμπλέκει κατά κάποιο τρόπο. Με τη χρήση του spotlight εξάγονται οι πόροι για τέτοιες περιπτώσεις, καθόσον αυτοί παρέχουν πολύ χρήσιμες επιπρόσθετες πληροφορίες σχετικά με το βίντεο και ως εκ τούτου για τη σημασιολογική του ανάκτηση.

6.4 Ανάπτυξη πιλοτικής εφαρμογής

Μέχρι τώρα έχουμε περιγράψει τα κύρια ζητήματα που αφορούν στη μετατροπή των συλλεγόμενων και ομογενών XML στοιχείων σε RDF και στην εσωτερική και εξωτερική τους σύνδεση. Ωστόσο, προκειμένου να εκπληρώσουμε τις 4 κύριες αρχές των Διασυνδεδεμένων Δεδομένων πρέπει να πραγματοποιηθούν κάποιες επιπλέον ενέργειες. Πρώτον, τόσο η πληροφορία που είναι κατανοητή από τους πράκτορες λογισμικού (RDF) όσο και από τον άνθρωπο (HTML) πρέπει να παρέχονται μέσω κάποιου εξυπηρετητή. Στην παρούσα πρόταση, όπου έχουν παραχθεί στατικά αρχεία RDF, η καλύτερη πρακτική είναι να φορτωθούν απευθείας τα αρχεία στο διακομιστή. Ένα άλλο σημαντικό θέμα κατά τη δημοσίευση Διασυνδεδεμένων Δεδομένων είναι να προσφέρουμε ένα μηχανισμό διαπραγμάτευσης περιεχομένου [67], η βασική ιδέα του οποίου είναι ότι οι HTTP πελάτες στέλνουν επικεφαλίδες (headers) HTTP με κάθε αίτηση για να δείξουν τι είδους έγγραφα προτιμούν. Το παρακάτω σχήμα απεικονίζει το αποτέλεσμα του VAPOUR⁷ παρουσιάζοντας τη διαδικασία dereferencing για τα δεδομένα του έργου EUscreen που αναφερόμαστε.

⁵ <http://dbpedia.org>

⁶ <http://dbpedia.org/spotlight>

⁷ Εκτιμητής Διάσυνδεδεμένων Δεδομένων <http://vapor.sourceforge.net/>

DARIAH-ΑΤΤΙΚΗ Ανάπτυξη της ελληνικής ερευνητικής υποδομής για τις ανθρωπιστικές επιστήμες ΔΥΑΣ

Vapour Report



Σχήμα 6.2: Αξιολόγηση της διαδικασίας διαπραγμάτευσης με χρήση του VAPOUR

6.4.1 Σημσιολογική βάση δεδομένων

Για την αποθήκευση των δεδομένων που παράγονται από τη διαδικασία RDFization χρησιμοποιούνται βάσεις δεδομένων ειδικά κατασκευασμένες για την αποθήκευση και την ανάκτηση των RDF μεταδεδομένων, που ονομάζονται triplestores. Στη συνέχεια παραθέτουμε μερικά από τα πιο δημοφιλή triplestores που υποστηρίζουν την αρχιτεκτονική πλέγματος (grid computing).

- Το Bigdata⁸ είναι ένα σύσταδοποιημένο RDF store για διατεταγμένα δεδομένα (B + δένδρα) και έχει σχεδιαστεί για να λειτουργεί ως διακομιστής. Είναι διαθέσιμο από την GNU General Public License (GPL) και έχει σχεδιαστεί για συστήματα UNIX (Linux) με συνδέσεις πελάτη για Java (Sesame). Επιπλέον κλιμάκωση μπορεί να επιτευχθεί, με απλή σύνδεση σε περισσότερες υπηρεσίες δεδομένων, δυναμικά κατά το χρόνο εκτέλεσης, οι οποίες θα εγγραφούν από μόνες τους στην κεντρική υπηρεσία διαχείρισης και θα ξεκινήσουν τη διαχείριση των δεδομένων αυτόματα. Υπο-κλιμάκωση επιτυγχάνεται μέσω δυναμικής κατάτμησης του φάσματος κλειδών των δεικτών των B + Δέντρων.

⁸ <http://www.systap.com/bigdata.html>

DARIAH-ΑΤΤΙΚΗ Ανάπτυξη της ελληνικής ερευνητικής υποδομής για τις ανθρωπιστικές επιστήμες ΔΥΑΣ

- Το OWLIM⁹ είναι μια οικογένεια των εμπορικών λύσεων αποθήκευσης RDF, που παρέχονται από την Ontotext. Είναι διαθέσιμο σε δύο διαφορετικές εκδόσεις: το SwiftOWLIM που είναι σχεδιασμένο για μεσαίου όγκου δεδομένα (μέχρι 100 εκατομμύρια triples) εφόσον η συλλογιστική (reasoning) και η αναζήτηση (query) εκτελούνται στην κύρια μνήμη, ενώ το BigOWLIM έχει σχεδιαστεί για μεγάλους όγκους δεδομένων και έχει τη δυνατότητα υποστήριξης δισεκατομμυρίων RDF triples υποστηρίζοντας δυναμική διαμόρφωση του συμπλέγματος. Επιπλέον, το OWLIM είναι διαθέσιμο ως SAIL (Storage and Inference Layer) για το Sesame πλαίσιο RDF. Ο SwiftOWLIM πηγαίος κώδικας παρέχεται για οποιοδήποτε σκοπό, με άδεια GNU LGPL. Το BigOWLIM είναι ελεύθερο για χρήση στην έρευνα, στην αξιολόγηση, και στον έλεγχο. Για εμπορικές εφαρμογές απαιτείται κατάλληλη άδεια. Και οι δύο εκδόσεις δεν παρέχουν ειδικό μηχανισμό επεκτασιμότητας, αλλά επιτρέπουν τον ορισμό προσαρμοζόμενων κανόνων και γλωσσών κανόνων για τη διαδικασία λογικής συμπερασματολογίας. Το BigOWLIM είναι σε χρήση από μεγάλο αριθμό εφαρμογών του Σημαιολογικού Ιστού και των Διασυνδεδεμένων Δεδομένων.
- Το 4Store¹⁰ είναι μια RDF βάση δεδομένων που αναπτύχθηκε από τη Garlik Inc. Είναι υλοποιημένο σε ANSIC99 και διαθέσιμο υπό την GNU General Public License (GPL), έκδοση 3. Είναι σχεδιασμένο για συστήματα τύπου UNIX (Linux, Mac OS) και λειτουργεί ως server σε ένα μηχάνημα ή σε σύμπλεγμα από 64bit μηχανήματα. Συνδέσεις clients είναι διαθέσιμες σε PHP, Ruby, Python, και Java. Ειδικοί μηχανισμοί επέκτασης δεν προβλέπονται.
- Το SHARD¹¹ (Scalable, High-Performance, Robust and Distributed) είναι μια open source cloud-based τεχνολογία triplestore που επιτρέπει κλιμακωτή επεξεργασία δεδομένων και ανάλυση με βάση τη διανομή Cloudera για την Hadoop υλοποίηση του MapReduce φορμαλισμού. Το SHARD αντιμετωπίζει τα δεδομένα σαν τυπικά RDF και τρέχει ερωτήματα πάνω σε αυτά χρησιμοποιώντας την τυπική γλώσσα ερωτημάτων SPARQL.
- Το Dydra¹² προσφέρεται ως υπηρεσία σε ένα cloud-based RDF store. Ο χρήστης δεν χρειάζεται να εκτίμησει το μέγεθος των δεδομένων, ή να

⁹ <http://www.ontotext.com/OWLIM>

¹⁰ <http://4store.org/>

¹¹ <http://www.dist-systems.bbn.com/people/krohloff/shard.shtml>

¹² <http://docs.dydra.com/dydra>

DARIAH-ΑΤΤΙΚΗ Ανάπτυξη της ελληνικής ερευνητικής υποδομής για τις ανθρωπιστικές επιστήμες ΔΥΑΣ

ανησυχεί για τη χρήση των συστάδων (clusters), των κόμβων, των πόρων, ή να κάνει δεσμεύσεις για μεγάλες άδειες εκ των προτέρων.

- Το Sesame ¹³ είναι ένα ανοικτού κώδικα πλαίσιο εραγοσίας σε Java για την αποθήκευση, αναζήτηση και λογική συμπερασματολογία με RDF και RDFS. Μπορεί να χρησιμοποιηθεί ως βάση δεδομένων για την RDF και RDF Schema, ή ως μια βιβλιοθήκη της Java για εφαρμογές που χρειάζονται να συνεργαστούν με RDF εσωτερικά.

Για την επιλογή ενός triplestore για την αποθήκευση των δεδομένων και την δημιουργία της πιλοτικής εφαρμογής πρέπει να πληρούνται οι ακόλουθες απαιτήσεις:

1. Το triplestore πρέπει να είναι κατανεμημένο. Αυτό σημαίνει ότι οι συσκευές αποθήκευσης δεν είναι συνδεδεμένες με την ίδια CPU, αλλά μπορούν να είναι σε πολλούς υπολογιστές που βρίσκονται στην ίδια φυσική θέση, ή μπορούν να είναι διασκορπισμένες σε ένα δίκτυο διασυνδεδεμένων υπολογιστών.
2. Το triplestore πρέπει να έχει άδεια open source. Μία open-source άδεια χρήσης είναι μια άδεια πνευματικών δικαιωμάτων για το λογισμικό υπολογιστών που καθιστά τον πηγαίο κώδικα διαθέσιμο έτσι ώστε να μπορούν όλοι να τον χρησιμοποιήσουν. Αυτό επιτρέπει στους τελικούς χρήστες να επανεξετάσουν και να τροποποιήσουν τον πηγαίο κώδικα για τη δική τους προσαρμογή και / ή την αντιμετώπιση προβλημάτων ανάλογα με τις εκάστοτε ανάγκες. Οι άδειες χρήσης ανοιχτού κώδικα είναι συνήθως δωρεάν, επιτρέποντας την τροποποίηση, αναδιανομή, και εμπορική χρήση, χωρίς να χρειάζεται να πληρωθεί ο αρχικός συγγραφέας.
3. Το triplestore πρέπει να υποστηρίζει τη SPARQL που είναι μια RDF Query Language σχεδιασμένη να καλύπτει τις περιπτώσεις χρήσης και τις απαιτήσεις που προσδιορίζονται από την RDF Data Access Ομάδα Εργασίας.

6.4.2 Αποτελέσματα του εμπλουτισμού

Εφαρμόζοντας την ανωτέρω διαδικασία στα προαναφερθέντα δεδομένα των έργων EUScreen-EUScreenXL, πετύχαμε τον σημασιολογικό εμπλουτισμό τους. Ειδικότερα, τα σημασιολογικά εμπλουτισμένα μεταδεδομένα αποθηκεύτηκαν σε

¹³ <http://www.openrdf.org/>

DARIAH-ΑΤΤΙΚΗ Ανάπτυξη της ελληνικής ερευνητικής υποδομής για τις ανθρωπιστικές επιστήμες ΔΥΑΣ

4store (<http://4store.org>) - μια ειδικά διαμορφωμένη βάση δεδομένων - προκειμένου να παρέχεται SPARQL πρόσβαση στα δεδομένα. Με αυτόν τον τρόπο τα δεδομένα μπορούν καταναλωθούν με χρήση της γλώσσας ερωτημάτων SPARQL από το <http://lod.euscreen.eu/sparql/> ή με τη χρήση του web interface του αποθετηρίου 4store <http://oreo.image.ece.ntua.gr:10999/test/>.

Ο παρακάτω πίνακας παραθέτει αποτελέσματα που αφορούν στους πόρους που δημιουργήθηκαν για την πιλοτική εφαρμογή καθώς και στον εμπλουτισμό τους.

Πόροι μεταδεδομένων	
Πόροι προγραμμάτων	41.622
Πόροι Προσώπων	19.995
Πόροι Χωρών	586
Πόροι Γλωσσών	46
<i>Συνολικοί πόροι</i>	<i>511.816</i>
Στατιστικά εμπλουτισμού	
Εξωτερικές συνδέσεις για τους πόρους χωρών	5940
Εξωτερικές συνδέσεις για τους πόρους γλωσσών	618
Προγράμματα από τις περιλήψεις των οποίων εξήχθηκαν πόροι προσώπων	905
Πόροι προσώπων που εξήχθηκαν από προγράμματα	1081
<i>Συνολικός αριθμός εξωτερικών πόρων</i>	<i>15.036</i>

Ως αποτέλεσμα του εμπλουτισμού ένας χρήστης μπορεί να πραγματοποιήσει το παρακάτω ερωτήμα για να αναζητήσει τηλεοπτικά προγράμματα του έργου EUscreen στα οποία αναφέρεται ένας ηθοποιός που έχει υποδυθεί τον ρόλο του James Bond:

DARIAH-ΑΤΤΙΚΗ Ανάπτυξη της ελληνικής ερευνητικής υποδομής για τις ανθρωπιστικές επιστήμες ΔΥΑΣ

```
PREFIX ebu:<http://www.ebu.ch/metadata/ontologies/ebucore/ebucore#>
PREFIX db:<http://dbpedia.org/resource/>
PREFIX db-on:<http://dbpedia.org/ontology/>

SELECT ?video ?actor
WHERE
{
    SERVICE <http://dbpedia.org/sparql> {
        db:James_Bond db-on:portrayer ?actor.
    }
    ?video ebu:isMentionedinSummary ?actor.
}
```

7. Εργαλεία εμπλουτισμού με διασυνδεδεμένα δεδομένα

Στην ενότητα αυτή θα παρουσιάσουμε τα διαθέσιμα εργαλεία και συστήματα που εξυπηρετούν στη δημιουργία συνδέσμων μεταξύ δεδομένων που προέρχονται από διάφορες πηγές με τα Διασυνδεδεμένα Δεδομένα του διαδικτύου.

7.1 Karma

Το Karma (<http://www.isi.edu/integration/karma/>) είναι ένα εργαλείο αντιστοίχισης μεταδεδομένων από συλλογές μουσειών σε διασυνδεδεμένα δεδομένα με βάση μια οντολογία επιλογής του χρήστη. Προηγούμενη έκδοση του προγράμματος είχε χρησιμοποιηθεί στην αντιστοίχιση δεδομένων του Smithsonian American Art Museum στην οντολογία EDM της Europeana, ενώ η τρέχουσα έκδοση έχει εφαρμοστεί με την οντολογία CIDOC CRM για δεδομένα πολιτιστικής κληρονομιάς.

Αρχικά, προκειμένου να πραγματοποιηθεί η αντιστοίχιση ορίζεται μια ιεραρχική δομή για τα URIs που θα δημιουργηθούν ανάλογα με την δομή των δεδομένων. Το Karma προσφέρει μια διεπαφή για τον χρήστη όμοια με αυτή του Microsoft Excel, όπου δηλαδή τα URIs ορίζονται σε τύποι για επεξεργασία συμβολοσειρών. Στη συνέχεια, δίνεται η δυνατότητα να επεκτασης της οντολογίας καθώς δεν είναι πάντοτε εφικτή η αντιστοίχιση των πεδίων των δεδομένων με τις κλάσεις της οντολογίας. Έπειτα, το ελεγχόμενο λεξιλόγιο του συστήματος διαχείρισης των δεδομένων ορίζεται ως θησαυρός με βάση το SKOS και πραγματοποιείται και μια δεύτερη αντιστοίχιση των εννοιών του SKOS με τις κλάσεις της οντολογίας. Αυτή αναπαριστάται με μορφή RDF όπως δηλαδή και οι αντιστοιχίσεις δεδομένων – οντολογίας. Τα δεδομένα στη συνέχεια μπορούν να καθαριστούν, όπως για παράδειγμα οι ημερομηνίες που πρέπει να έχουν την ίδια μορφοποίηση. Οι απαραίτητες μετατροπές γίνονται με συναρτήσεις που υλοποιούνται στη γλώσσα Python.

Η βασική λειτουργία του Karma είναι η διεπαφή που επιτρέπει στον χρήστη να αντιστοιχεί την οντολογία με τους πίνακες των δεδομένων. Ο χρήστης έχει στη διάθεσή του τους πίνακες με τα δεδομένα και τα URIs που παράχθηκαν, ενώ πάνω από αυτούς εμφανίζονται έννοιες της οντολογίας, που μπορούν να επιλεγθούν από λίστες, και συνδέονται μεταξύ τους με βέλη που υποδεικνύουν ρόλους της οντολογίας ή την ιεραρχική δομή τους. Διαφορετικά βέλη δείχνουν την αντιστοίχιση

DARIAH-ΑΤΤΙΚΗ Ανάπτυξη της ελληνικής ερευνητικής υποδομής για τις ανθρωπιστικές επιστήμες ΔΥΑΣ

με τις στήλες του πίνακα των δεδομένων. Ο χρήστης έχει πλήρη έλεγχο των αντιστοιχίσεων ενώ υποδεικνύονται αυτόματα συσχετίσεις. Το τελικό αρχείο της αντιστοίχισης είναι σε μορφή RDF. Η διαδικασία της εκτέλεσης διαρκεί λίγα μόνο λεπτά για δεδομένα των τάξεων των δεκάδων χιλιάδων εγγραφών, και οντολογίες με εκατοντάδες έννοιες.

7.2 PoolParty

Το PoolParty (<http://www.poolparty.biz/portfolioportfolio-4-columns/products/>) είναι ένα εργαλείο διαχείρισης θησαυρών που έχει αναπτυχθεί με βάση τα πρότυπα του W3C για το σημασιολογικό ιστό. Στον πυρήνα της η εφαρμογή χρησιμοποιεί μια αποθήκη τριάδων για να αναπαραστήσει λεξιλόγια (vocabularies) στη γλώσσα SKOS, στην οποία ο χρήστης μπορεί να προσθέσει και δικές του επεκτάσεις, με τη μορφή RDF. Επομένως έχει ένα πλεονέκτημα έναντι άλλων εργαλείων που χρησιμοποιούν σχεσιακές βάσεις δεδομένων. Η πρόσβαση στο εργαλείο μέσω διαδικτύου διευκολύνει για εργασίες που απαιτούν συνεργασία πολλών χρηστών, για θησαυρούς με αποκεντρωμένη δομή.

Ο σχεδιασμός του PoolParty στοχεύει στην ευκολία κατά τη δημιουργία και διαχείριση θησαυρών, από χρήστες χωρίς εξειδίκευση. Η ιστοσελίδα διαθέτει μια φόρμα αναζήτησης, ένα κύριο μενού και κουμπιά για λειτουργίες κειμένου, στα αριστερά το δέντρο με τις ιεραρχήσεις των εννοιών και τέλος στα δεξιά λεπτομέρειες για κάθε έννοια και το αντίστοιχο σχήμα. Οι βασικές διευκολύνσεις που προσφέρει η διεπαφή με τον χρήστη είναι λειτουργίες “drag and drop” και αυτόματης συμπλήρωσης σε σχεδόν όλες τις αλληλεπιδράσεις του χρήστη, όπως η δημιουργία, μετακίνηση και σύνδεση εννοιών ή η πρόσθεση συσχετίσεων, αλλά και η αυτόματη μετάφραση για ετικέτες. Ακόμη, η πλατφόρμα υποστηρίζει την αυτόματη επεξεργασία σωμάτων κειμένου με αλγόριθμους εξαγωγής πληροφορίας και στατιστικής επεξεργασίας, ώστε να προτείνει στο χρήστη αυτόματα πιθανές έννοιες που αυτός μπορεί να επιλέξει να συμπεριλάβει στον θησαυρό που κατασκευάζει. Τέλος, η ιεραρχία των εννοιών μπορεί να συμπληρωθεί αυτόματα με έννοιες της DBpedia με αντίστοιχη δομή και κατηγοριοποιήσεις. Οι θησαυροί που ορίζει ο χρήστης μπορούν να οριστούν στη συνέχεια ως διασυνδεδεμένα δεδομένα (LOD) με τη δημιουργία διευθύνσεων URI. Τα URIs εμφανίζουν ιστοσελίδες σε επίπεδα εξυπηρετητή, έργων, εννοιών και υποεννοιών του θησαυρού, με συνδέσμους και πληροφορίες για διάφορα μεταδεδομένα στο περιεχόμενο των σελίδων. Ακόμη, μπορούν να γίνουν διασυνδέσεις με πηγές διασυνδεδεμένων

DARIAH-ΑΤΤΙΚΗ Ανάπτυξη της ελληνικής ερευνητικής υποδομής για τις ανθρωπιστικές επιστήμες ΔΥΑΣ

δεδομένων όπως οι Dbpedia, GeoNames, LCSH, WordNet, Yago και Umbel, ή άλλες που προσθέτει ο χρήστης. Οι αντιστοιχίσεις γίνονται από τον χρήστη με την επιλογή μίας έννοιας και έπειτα με την αναζήτηση στις πηγές δυνδεδεμένων δεδομένων με μια απλή φόρμα αναζήτησης και επιλογής των κατάλληλων αποτελεσμάτων και τον τύπο αντιστοίχισης, από αυτούς που προσφέρει το SKOS. Μετά την αντιστοίχιση το πρόγραμμα μπορεί να εισάγει από αυτήν και τα αντίστοιχα μεταδεδομένα που περιέχονται.

Η εφαρμογή έχει μια πολύ καλή τεκμηρίωση με διάφορα εγχειρίδια για χρήστες, διαχειριστές και τεχνικούς ενώ λόγω του διαδικτυακού της χαρακτήρα προσφέρει και λειτουργίες διαχείρισης χρηστών, APIs, τερματικό σε SPARQL και άλλες διαδικτυακές υπηρεσίες. Οι εξυπηρετητές της εφαρμογής ανάλογα με την έκδοση μπορούν να αποθηκεύσουν έως και 500.000 έννοιες με 10 ετικέτες και 5 διασυνδέσεις κατά μέσο όρο, σε ένα ή περισσότερα έργα.

7.3 Alchemy API

Το AlchemyAPI (<http://www.alchemyapi.com/api/register.html>) είναι μία διεπαφή με σκοπό να διευκολύνει εφαρμογές που απαιτούν δυνατότητες κατανόησης φυσικής γλώσσας και εικόνων. Το API προσφέρει τερματικά REST, SDKs σε δημοφιλείς γλώσσες προγραμματισμού, υποστήριξη διάφορων σειριοποιήσεων για μεταδεδομένα όπως XML, JSON, RDF, ενώ παρέχει πλήρη και άριστη αυτοματοποίηση.

Το AlchemyAPI προσφέρει υπηρεσίες αυτόματης εξαγωγής εννοιών, σχέσεων ή λέξεων κλειδιών ή οντοτήτων σημασιολογικού περιεχομένου από πηγές κειμένου, αρχείων HTML ή άλλες πηγές διαδικτυακού περιεχομένου. Τα αποτελέσματα αυτών των διαδικασιών συνδέονται αυτόματα με πηγές διασυνδεδεμένων δεδομένων, συγκεκριμένα τις Freebase, US Census, GeoNames, UMBEL, OpenCyc, YAGO, MusicBrainz, CIA Factbook και CrunchBase. Το Alchemy API αντιμετωπίζει ενδεχόμενες αμφισημίες κατά την εξαγωγή οντοτήτων χρησιμοποιώντας διάφορους στατιστικούς αλγόριθμους που ελέγχουν το περιεχόμενο του γειτονικού κειμένου για να καταλήξουν στην σωστή σημασία της κάθε λέξης και να τη συνδέσουν σωστά με την αντίστοιχη οντότητα και τα διασυνδεδεμένα δεδομένα.

DARIAH-ΑΤΤΙΚΗ Ανάπτυξη της ελληνικής ερευνητικής υποδομής για τις ανθρωπιστικές επιστήμες ΔΥΑΣ

Η διαδικτυακή υπηρεσία απαιτεί πληρωμή σε κάθε κλήση του API. Η εταιρεία εκτιμά ότι δέχεται 3 δισεκατομμύρια κλήσεις ανά μήνα. Πέραν των υπηρεσιών εμπλουτισμού κειμένου, προσφέρει και άλλες υπηρεσίες όπως η εκτίμηση αισθήματος, αναγνώριση προσώπου σε εικόνες, και επισημείωση εικόνων με δεδομένα.

7.4 Dexter 2.0

Το Dexter 2.0 (<http://dexter.isti.cnr.it/download/>) είναι ένα εργαλείο ανοιχτού κώδικα που συνδέει αυτόματα μη δομημένα δεδομένα με οντότητες σε μια βάση γνώσης. Η πρώτη έκδοση του εργαλείου είχε σκοπό την υλοποίηση μεθόδων εμπλουτισμού και τη σύγκρισή τους σε μια κοινή πλατφόρμα. Η νέα έκδοση προσφέρει δυνατότητες επιλογής στρατηγικών εμπλουτισμού κατά την εκτέλεση καθώς και ένα καλά τεκμηριωμένο REST-API. Το εργαλείο προορίζεται κυρίως για ερευνητές και σπουδαστές, ενώ περιέχεται σε ένα αρχείο .jar και μπορεί να χρησιμοποιηθεί από το API ή από σελίδα διαδικτύου. Ο εμπλουτισμός των δεδομένων πραγματοποιείται χρησιμοποιώντας URIs του Wikipedia.

Η αρχιτεκτονική του εργαλείου στηρίζεται σε διάφορα τμήματα το καθένα από τα οποία εκτελεί συγκεκριμένες λειτουργίες. Ανάμεσα σε αυτά, αρχικά, υπάρχει ένας μετατροπέας μεταδεδομένων της Wikipedia από XML σε JSON ο οποίος αξιοποιεί την πληροφορία του MediaWiki. Αυτό έχει ως αποτέλεσμα σε μια εγγραφή JSON να περιέχονται δεδομένα όπως οι επικεφαλίδες, όλο το κείμενο του άρθρου και άλλες λεπτομέρειες που δεν εμφανίζονται στα RDF δεδομένα της Dbpedia. Το τμήμα του πυρήνα της εφαρμογής επιλέγει ανάμεσα σε τμήματα του κειμένου, τα κωδικοποιεί σε μια λίστα και στη συνέχεια τα αντιστοιχεί με έννοιες αντιμετωπίζοντας αμφισημίες. Όλα τα τμήματα μπορούν να αλλαχθούν ή να συνδυαστούν με άλλες εκδόσεις τους με αποτέλεσμα ο χρήστης που θα βελτιώσει τις τελικές δυνατότητες να μπορεί να βοηθήσει και άλλους χρήστες.

Η διεπαφή χρήστη οργανώνεται σε 4 κατηγοριοποιήσεις, που αφορούν στις επισημειώσεις του κειμένου, στην εξαγωγή των υποψήφιων σημείων για εμπλουτισμό και την οπτικοποίηση των χαρακτηριστικών τους, στη διεπαφή γράφου και κατηγοριών, και στην αναζήτηση στα δεδομένα της Wikipedia. Η συνεργασία των τμημάτων ορίζεται ως παράμετρος στη διεπαφή χρήστη επομένως διαφορετικοί συνδυασμοί είναι πιθανοί, ενώ είναι δυνατόν να τροποποιηθούν κατά

DARIAH-ΑΤΤΙΚΗ Ανάπτυξη της ελληνικής ερευνητικής υποδομής για τις ανθρωπιστικές επιστήμες ΔΥΑΣ

την εκτέλεση της εφαρμογής. Τέλος, το εργαλείο έχει τη δυνατότητα να επεξεργαστεί και ημιδομημένα κείμενα, δηλαδή να διαχειριστεί τίτλους και διαφορετικές παραγράφους χωρίς να πρέπει να παρέμβει ο χρήστης.

7.5 LODE

Το LODE (<http://www.linkedhumanities.com/>) είναι ένα εργαλείο διασύνδεσης δεδομένων με διασυνδεδεμένα δεδομένα που διατίθεται σε πλατφόρμα διαδικτυακής χρήσης. Έχει αναπτυχθεί με σκοπό τη διασύνδεση και τον εμπλουτισμό ψηφιακών συλλογών για τις ανθρωπιστικές επιστήμες.

Η εφαρμογή σκοπεύει να βοηθήσει χρήστες που θέλουν να πραγματοποιήσουν συνδέσεις με διασυνδεδεμένα δεδομένα με ένα απλό σύστημα βασισμένο σε λειτουργίες drag and drop. Οι βασικές λειτουργίες είναι η εξερεύνηση, η διασύνδεση και ο εμπλουτισμός δεδομένων. Η ιστοσελίδα του εργαλείου δίνει αρχικά στο χρήστη μια απλή φόρμα αναζήτησης. Ο χρήστης αναζητά όρους μιας τοπικής βάσης δεδομένων και ως αποτελέσματα της αναζήτησης εμφανίζονται τα URI καθώς και τα ονόματα και οι έννοιες με τις οποίες σχετίζονται. Παράλληλα, υπάρχει μια ένδειξη συσχέτισης με διασυνδεδεμένα δεδομένα που επιτρέπει στο χρήστη να μεταφερθεί στη σελίδα διασύνδεσης όπου βλέπει προτάσεις για διασύνδεση που έχουν παραχθεί αυτόματα από την εφαρμογή ενώ η διασύνδεση επιτυγχάνεται επιλέγοντας και τραβώντας με το ποντίκι. Στη συνέχεια, η εφαρμογή προτείνει στο χρήστη να εμπλουτίσει τα δεδομένα με έννοιες και ιδιότητες που προκύπτουν από την διασύνδεση που πραγματοποίησε.

Η δομή του εργαλείου είναι πολύ απλή και εύληπτη, αλλά οι αυτοματοποιήσεις περιορίζονται σε προτάσεις στον χρήστη. Όλες οι διαδικασίες του εργαλείου εκτελούνται ταχύτατα και με μεγάλη ευκολία.

7.6 Silk

Το Silk (<http://silk-framework.com/download>) αποτελεί προγραμματιστικό πλαίσιο ανοιχτού κώδικα για την ενοποίηση ετερογενών πηγών δεδομένων. Προορίζεται για λειτουργίες αντιστοίχισης ανάμεσα σε διαφορετικές πηγές διασυνδεδεμένων

DARIAH-ΑΤΤΙΚΗ Ανάπτυξη της ελληνικής ερευνητικής υποδομής για τις ανθρωπιστικές επιστήμες ΔΥΑΣ

δεδομένων, για τη δημιουργία συνδέσμων RDF και εφαρμογής μετασχηματισμών σε πηγές δομημένων δεδομένων.

Το σύστημα χρησιμοποιεί μια δική του γλώσσα δηλώσεων την Silk LSL (γλώσσα προσδιορισμού συνδέσμων) με την οποία οι χρήστες μπορούν να ορίσουν τους τύπους συνδέσμων RDF που θα εξάγονται από τις πηγές, καθώς και τα κριτήρια που θα επιτρέπουν την πραγματοποίηση της διασύνδεσης όπως είναι μέτρα ομοιότητας που συμπεριλαμβάνουν και η δομή του γράφου RDF που χαρακτηρίζει τα δεδομένα. Τα κριτήρια αυτά μπορούν να καθοριστούν σε αρχεία XML ή χρησιμοποιώντας το γραφικό περιβάλλον του Silk. Η πρόσβαση στα δεδομένα γίνεται μέσω διεπαφής SPARQL ενώ μπορούν να αντιμετωπιστούν καταστάσεις που χρησιμοποιούνται όροι από διαφορετικά λεξιλόγια και δεν υπάρχει ενιαίο σχήμα RDFS ή OWL. Επιπλέον, είναι δυνατός ο ορισμός κανόνων μετασχηματισμού των δεδομένων, όπως η απομάκρυνση ανεπιθύμητων τιμών, οι αντιστοιχίσεις διαφορετικών ιδιοτήτων ή προσθήκη νέων, καθώς και μετατροπές ανάμεσα σε διαφορετικές σειριοποιήσεις όπως RDF, CSV ή XML.

Το γραφικό περιβάλλον του συστήματος εμφανίζεται σε μια διαδικτυακή εφαρμογή που καθοδηγεί το χρήστη στις διαδικασίες διασύνδεσης δεδομένων. Οπτικοποιεί τα σύνολα των δεδομένων καθώς και τις διαδικασίες σύνδεσης και μετατροπής. Παραθέτει προτάσεις για συνδέσεις που ο χρήστης μπορεί να εγκρίνει ή να απορρίψει και τέλος του επιτρέπει να χρησιμοποιήσει ένα σύνολο συνδέσμων αναφοράς για να επιθεωρήσει την τρέχουσα διασύνδεση. Τέλος, το σύστημα παρέχει έναν επεξεργαστή φυσικής γλώσσας που μπορεί να εκπαιδευτεί με κείμενο σε φυσική γλώσσα ώστε να εξάγει ένα αντίστοιχο αρχείο RDF.

7.7 Pundit

Το Pundit (<http://thepund.it/>, <https://github.com/net7/pundit2>) είναι ένα εργαλείο ανοιχτού κώδικα για τη δημιουργία σημασιολογικών επισημειώσεων. Οι χρήστες του μπορούν να προσθέσουν σημασιολογικό περιεχόμενο σε ιστοσελίδες με τη μορφή σχολίων, αναφορών, σελιδοδεικτών και άλλων ώστε να μπορούν να συμπεριληφθούν στον σημασιολογικό ιστό. Το πλαίσιο χρησιμοποιεί τα πρότυπα της W3C για την απόδοση της δομής με σημασιολογικό περιεχόμενο και οι διασυνδέσεις των δεδομένων γίνονται με χρήση οντολογιών και λεξικών όπως το Freebase, η Europeana και η DBpedia.

DARIAH-ΑΤΤΙΚΗ Ανάπτυξη της ελληνικής ερευνητικής υποδομής για τις ανθρωπιστικές επιστήμες ΔΥΑΣ

Η ιδιαιτερότητα του Pundit εμφανίζεται στην συλλογική πλατφόρμα του. Οι χρήστες μπορούν να χρησιμοποιήσουν τη διαδικτυακή εφαρμογή η οποία εμφανίζεται με τη μορφή παραθύρων πάνω από την ιστοσελίδα που θέλουν να επεξεργαστούν. Αυτό επιτρέπει σε πολλούς χρήστες να συνεργαστούν και να μοιραστούν συλλογές επισημειώσεων που αφορούν στο σημασιολογικό εμπλουτισμό της σελίδας και που μπορούν να οριστούν ως ανοιχτές ή κλειστές στο κοινό. Οι επισημειώσεις οργανώνονται σε ταξινομίες ώστε να υπάρχει η δυνατότητα μελλοντικής αξιοποίησής τους. Στην ίδια σελίδα γίνονται όλες οι εργασίες, δηλαδή και η επισημείωση της σελίδας και η σύνδεση με διασυνδεδεμένα δεδομένα με πηγές που μπορεί να επιλέξει ο χρήστης. Οι επισημειώσεις αποθηκεύονται σε μορφή RDF σε έναν εξυπηρετητή με SPARQL διεπαφές. Τέλος, προσφέρονται διάφοροι τρόποι οπτικοποίησης των επισημειώσεων που είναι δυνατό να αξιοποιηθούν για τη διερεύνηση και επέκτασή τους.

8. Συμπεράσματα

Η τρέχουσα τεχνολογική στάθμη (state-of-the-art) στις ψηφιακές βιβλιοθήκες, όπως η Europeana, στοχεύει στην αποτελεσματική πρόσβαση στο περιεχόμενο, αξιοποιώντας τις σημασιολογικές τεχνολογίες, όπως την αναπαράσταση γνώσης, τα πρότυπα μεταδεδομένων και τις αντιστοιχίσεις, τα διασυνδεδεμένα δεδομένα (linked data) και τα μοντέλα των χρηστών. Τα μεταδεδομένα του περιεχομένου αποτελούν το κύριο χαρακτηριστικό των πολιτιστικών αντικειμένων που αναλύονται, αντιστοιχίζονται και χρησιμοποιούνται για να ερμηνεύσουν τα ερωτήματα των χρηστών, έτσι ώστε να τους παρουσιάζεται πάντοτε το πιο κατάλληλο περιεχόμενο. Σε αυτό το πλαίσιο, η έρευνα και η ανάπτυξη στο πλαίσιο του DARIAH-GR που παρουσιάζεται στο παρόν Παραδοτέο εστιάζει στη δημιουργία αποτελεσματικών και αποδοτικών προηγμένων μηχανισμών αναζήτησης.

Το παρόν Παραδοτέο παρουσιάζει τη μεθοδολογία σημασιολογικής αναζήτησης, που περιλαμβάνει μηχανισμό απάντησης ερωτημάτων με βάση την σημασιολογία τους και ταυτόχρονα να εμπλουτίζει αυτές τις απαντήσεις, κατ' αρχήν αξιοποιώντας κατάλληλα οπτικά χαρακτηριστικά, τόσο τοπικά, όσο και MPEG-7. Αυτό επιτυγχάνεται μέσω του συνδυασμού τεχνικών τυπικής γνώσης με τεχνικές μηχανικής μάθησης. Δείχνεται ότι με την οπτική εξαγωγή των χαρακτηριστικών γνωρισμάτων, τη συσταδοποίηση, τη μηχανική μάθηση και το συνδυασμό αυτών με τεχνικές της βάσης γνώσης για την απάντηση των ερωτημάτων των χρηστών, μπορεί να επιτευχθεί σημαντική βελτίωση της σημασιολογικής αναζήτησης στις ψηφιακές βιβλιοθήκες.

Στη συνέχεια παρουσιάστηκε η μεθοδολογία που ακολουθείται για τον εμπλουτισμό των μεταδεδομένων και την δημοσίευσή τους ως Διασυνδεδεμένα Δεδομένα. Η πιλοτική εφαρμογή που αναπτύχθηκε βασίζεται σε οπτικοακουστικό υλικό των έργων EUscreen/EUScreenXL, τα οποία περιέχουν καλή τεκμηρίωση. Με τη δημοσίευσή του με τη μορφή των Διασυνδεδεμένων δεδομένων, το περιεχόμενο αυτό μπορεί να αντληθεί εύκολα, καθιστώντας με αυτόν τον τρόπο την υλοποίηση των διαφόρων εφαρμογών που το χρησιμοποιούν πολύ απλούστερη. Οι εφαρμογές αυτές μπορούν να χρησιμεύσουν για εκπαιδευτικούς, ερευνητικούς, ή ιστορικούς σκοπούς, αφού η

DARIAH-ΑΤΤΙΚΗ Ανάπτυξη της ελληνικής ερευνητικής υποδομής για τις ανθρωπιστικές επιστήμες ΔΥΑΣ

πλειοψηφία του περιεχομένου καλύπτει ιστορικά θέματα. Επιπλέον, το περιεχόμενο εμπλουτίζεται με τη σύνδεσή του με εξωτερικές πηγές δεδομένων, όπως οι DBpedia, Eurostat, Freebase και NY Times, επιτρέποντας με αυτόν τον τρόπο την πιο εκφραστική αναζήτηση και ανάκτηση.

Πρέπει τέλος να αναφερθεί ότι η εφαρμογή της βελτιωμένης και εμπλουτισμένης σημασιολογικής αναζήτησης μπορεί να γίνει σε μια ποικιλία υποπεριοχών, όπως η αρχαιολογία, η έρευνα, οι σύγχρονες τέχνες, κλπ., όπου μπορεί να παραχθεί και να χρησιμοποιηθεί συγκεκριμένη και σπονδυλωτή θεματική γνώση, καθώς και ο συνδυασμός της με τον εξελισσόμενο τομέα των linked open data.

9. Βιβλιογραφία

[1] European Council, European Content in Global Networks: Coordination Mechanisms for Digitisation Programmes (Lund Principles), 2001, available:ftp://ftp.cordis.lu/pub/ist/docs/digicult/lund_principles-en.pdf

[2] Evain, Jean Pierre (ed.) EBU Core Metadata Set EBU, Consulted January 31, 2012, available: http://tech.ebu.ch/docs/tech/tech3293v1_3.pdf

[3] G. Schreiber, Metadata Models, Interoperability Gaps, and Extensions to Preservation Metadata Standards, 2010

[4] F. Baader, D. Calvanese, D. McGuinness, D. Nardi, and P. F. Patel-Schneider, editors. The Description Logic Handbook: Theory, Implementation, and Applications. Cambridge University Press (2007)

[5] Perez-Urbina, H., Motik, B., Horrocks, I.: Tractable query answering and rewriting under description logic constraints. Journal of Applied Logic 8(2), 186-209 (2010)

[6] Ilianna Kollia, Birte Glimm and Ian Horrocks, SPARQL Query Answering over OWL Ontologies. In: Proceedings of the 8th Extended Semantic Web Conference (ESWC 2011). Springer Verlag (2011)

[7] Rob Shearer, Boris Motik and Rob Shearer and Ian Horrocks, Hermit: A Highly-Efficient OWL Reasoner, Proc. of the 5th Int. Workshop on OWL: Experiences and Directions (OWLED 2008 EU) (2008)

[8] Frank Manola and Eric Miller, editors. Resource Description Framework (RDF): Primer, W3C Recommendation (2004)

[9] Boris Motik, Peter F. Patel-Schneider and Bijan Parsia, editors, OWL 2 Web Ontology Language: Structural Specification and Functional-Style Syntax, W3C Recommendation (2009)

DARIAH-ΑΤΤΙΚΗ Ανάπτυξη της ελληνικής ερευνητικής υποδομής για τις ανθρωπιστικές επιστήμες ΔΥΑΣ

[10] B. Motik, P.F. Patel-Schneider, B. Guenca Grau (eds.): OWL 2 Web Ontology Language Direct Semantics. W3C Recommendation (27 October 2009), available at <http://www.w3.org/TR/owl2-direct-semantics/>

[11] H. J. Horst. Completeness, decidability and complexity of entailment for RDF Schema and a semantic extension involving the OWL vocabulary. *Journal of Web Semantics*, 3(2-3):79-115, (2005)

[12] Bernhard Schoelkopf and Alexander. J. Smola. *Learning with Kernels*. The MIT Press, Cambridge, Massachusetts, 2002

[13] Richard Duda, Peter Hart, David Stork. *Pattern Classification*. John Wiley & Sons, Inc, 2001

[14] Simon Haykin. *Neural Networks A Comprehensive Foundation*. Prentice Hall International, Inc., 1999

[15] Ilianna Kollia, Nikolaos Simou, Andreas Stafylopatis and Stefanos Kollias. Semantic Image Analysis using a Symbolic Neural Architecture. *Journal of Image Analysis and Stereology*.2010.

[16] N. Fanizzi, C. d'Amato, F. Esposito. Learning to Rank Individuals in Description Logics Using Kernel Perceptrons. RR 2010, 4th International Conference on Web Reasoning and Rule Systems, September, 22 - 24, 2010, Bressanone/Brixen, Italy.

[17] N. Fanizzi, C. d'Amato, F. Esposito. Statistical Learning for Inductive Query Answering on OWL Ontologies. ISWC 2008, International Semantic Web Conference, October, 26-30, 2008, Karlsruhe, Germany.

[18] SIEDL: First Workshop on Semantic Interoperability in the European Digital Library, 5th European Semantic Web Conference, Tenerife, Spain, June 2, 2008.

[19] Hitzler, P., Kroetzsch, M., Rundolph, S.: *Foundations of Semantic Web Technologies*. Chapman & Hall/CRC (2009)

DARIAH-ΑΤΤΙΚΗ Ανάπτυξη της ελληνικής ερευνητικής υποδομής για τις ανθρωπιστικές επιστήμες ΔΥΑΣ

[20] Report 'Existing standards applied by European Museums', <http://www.athenaeurope.org>

[21] Report 'Converting XML files to Apenet EAD', <http://www.apenet.eu>

[22] The New Renaissance Report of the European Reflection Group on Digital Libraries (Comite des Sages), January 10, 2011, <http://www.europeana.eu>

[23] Report of European Commission Member State Expert Group on Digitisation and Digital Preservation (MSEG), <http://www.ec.europa.eu>

[24] http://ec.europa.eu/information_society/apps/projects/

[25] Christian Bizer, Tom Heath, Tim Berners-Lee, Linked Data: The Story So Far, International Journal on Semantic Web and Information Systems (IJSWIS), 2009.

[26] Keynote speech 'Linked Data and Europeana: Perspectives and issues', Europeana Plenary Conference, The Hague, The Netherlands, September 14, 2009.

[27] Report of Joint Programming Initiative in Cultural Heritage and Global Change, <http://www.cordis.europa.eu/pub/fp7/coordination>

[28] M. Doer, D. Kalomirakis, 'A Metastructure for Thesauri in Archaeology, Computing Archaeology for Understanding the Past'. In Proceedings of the 28th Conference, BAR International Series, Lublijana, 2000.

[29] D. Kalomirakis, 'Polydefkis: A Terminology Thesauri for Monuments'. In M. Tsipopoulou (ed.), Proc. of 'Digital Heritage in the New Knowledge Environment: Shared spaces and open paths to cultural content', Athens, 2008.

[30] Proceedings of the ATHENA Conference 'Cultural Institutions Online', Rome, 28 April 2011.

[31] Y.Avrithis and K.Rapantzikos, Capturing boundary structure for feature detection, International Conference on Computer Vision (ICCV), 2011.

DARIAH-ΑΤΤΙΚΗ Ανάπτυξη της ελληνικής ερευνητικής υποδομής για τις ανθρωπιστικές επιστήμες ΔΥΑΣ

[32] H.Bay, T.Tuytelaars, and L.Van Gool, SURF: Speeded up robust features, ECCV, 2006.

[33] J.L. Bentley, Multidimensional binary search trees used for associative Searching, Communications of the ACM, 18(9):509--517, 1975.

[34] Shih-Fu Chang, Thomas Sikora, and Atum Puri, Overview of the MPEG-7 Standard, IEEE Trans. on Circuits and Systems for Video Technology, 11(6):688-695, 2001.

[35] G.Csurka, C.Dance, L.Fan, J.Willamowski, and C.Bray, Visual categorization with bags of keypoints, Workshop on statistical learning in computer vision, ECCV, volume 1, page 22. Citeseer, 2004.

[36] J.H. Freidman, J.L. Bentley, and R.A. Finkel, An Algorithm for Finding Best Matches in Logarithmic Expected Time, ACM Transactions on Mathematical Software (TOMS), 3(3):209--226, 1977.

[37] C.Harris and M.Stephens, A combined corner and edge detector, Alvey Vision Conference, 1988.

[38] T.Lindeberg, Feature detection with automatic scale selection, IJCV, 30(2):79--116, 1998.

[39] D.G. Lowe, Distinctive image features from scale-invariant keypoints, IJCV, 60(2):91--110, 2004.

[40] J.MacQueen, Some methods for classification and analysis of multivariate Observations, Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, pages 1-297, 1967.

[41] J.Matas, O.Chum, M.Urban, and T.Pajdla, Robust wide-baseline stereo from maximally stable extremal regions, Image and Vision Computing, 22(10):761--767, 2004.

[42] K.Mikolajczyk and C.Schmid, Scale & affine invariant interest point detectors, IJCV, 60(1):63--86, 2004.

DARIAH-ΑΤΤΙΚΗ Ανάπτυξη της ελληνικής ερευνητικής υποδομής για τις ανθρωπιστικές επιστήμες ΔΥΑΣ

- [43] K.Mikolajczyk, T.Tuytelaars, C.Schmid, A.Zisserman, J.Matas, F.Schaffalitzky, T.Kadir, and L.V Gool, A comparison of affine region detectors, *IJCV*, 65(1):43--72, 2005.
- [44] A.W. Moore, An introductory tutorial on kd-trees, Technical report, Technical Report,
- [45] Marius Muja and David~G. Lowe, Fast approximate nearest neighbors with automatic algorithm configuration, *International Conference on Computer Vision Theory and Application VISSAPP'09*), pages 331--340. INSTICC Press, 2009.
- [46] A.Neubeck and L.Van~Gool, Efficient Non-Maximum Suppression, *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on Pattern Recognition*, volume 3, 2006.
- [47] S.M. Omohundro, Efficient algorithms with neural network behavior, *Complex Systems*, 1(2):273--347, 1987.
- [48] K.Rapantzikos, Y.Avrithis, and S.Kollias, Detecting regions from single scale edges, *International Workshop on Sign, Gesture and Activity (SGA'10)*, *European Conference on Computer Vision (ECCV)*, 2010.
- [49] C.Silpa-Anan and R.Hartley, Optimised kd-trees for fast image descriptor matching, 2008.
- [50] J.Sivic and A.Zisserman, Video google: A text retrieval approach to object matching in videos, *International Conference on Computer Vision*, volume~2, pages 1470--1477, 2003.
- [51] T.Tuytelaars and L.J Van Gool, Content-based image retrieval based on local affinity invariant regions, *Visual Information Systems*, 1999.
- [52] M.Calonder, V.Lepetit, C.Strecha, and P.Fua, Brief: Binary robust independent elementary features, *European Conference on Computer Vision*, ser. *Lecture Notes in Computer Science*, K.Daniilidis, P.Maragos and N.Paragios, Eds., vol. 6314, Springer, 2010, pp. 778--792.

DARIAH-ΑΤΤΙΚΗ Ανάπτυξη της ελληνικής ερευνητικής υποδομής για τις ανθρωπιστικές επιστήμες ΔΥΑΣ

[53] M.Agrawal, K.Konolige, and M.R. Blas, CenSurE: Center surround extremas for realtime feature detection and matching, ECCV, 2008.

[54] J.Philbin, O.Chum, M.Isard, J.Sivic, and A.Zisserman, Object retrieval with large vocabularies and fast spatial matching, CVPR}, 2007.

[55] G.Tolias and Y.Avrithis, Speeded-up, relaxed spatial matching, Proceedings of International Conference on Computer Vision (ICCV 2011), Barcelona, Spain, November 2011.

[56] E.Rosten and T.Drummond, Machine learning for high-speed corner detection, European Conference on Computer Vision, vol. 3951, 2006, p. 430.

[57] Prud'hommeaux, E. & Seaborne, A. (2008) SPARQL Query Language for RDF - W3C Recommendation, Consulted January 31, 2012. Available: <http://www.w3.org/TR/rdf-sparql-query/>

[58] Bizer, C. & Heath, T. & Berners-Lee, T. (2009), Linked Data - The Story So Far, International Journal on Semantic Web and Information Systems, Volume: 5, Issue: 3, Publisher: Elsevier, Pages: 1-22 ISSN: 15526283, DOI: 10.4018/jswis.2009081901

[59] Klyne, G. & Carroll, J. (2004), Resource Description Framework (RDF): Concepts and Abstract Syntax- W3C Recommendation. Retrieved June 14, 2009, Consulted January 31, 2012. Available: <http://www.w3.org/TR/rdfconcepts/>

[60] Miles, A. & Bechhofer, S (2009), Skos simple knowledge organization system, Consulted January 31, 2012. Available: <http://www.w3.org/TR/skos-reference/>

[61] Brickley, D. & Guha, R. V.(2004), RDF Vocabulary Description Language 1.0: RDF Schema - W3C Recommendation, Consulted January 31, 2012. Available: <http://www.w3.org/TR/rdf-schema/>

[62] McGuinness, D. L. & Harmelen, F.V (2004), OWL Web Ontology Language Overview - W3C Recommendation, Consulted January 31, 2012. Available: <http://www.w3.org/TR/2004/REC-owl-features-20040210/>

DARIAH-ΑΤΤΙΚΗ Ανάπτυξη της ελληνικής ερευνητικής υποδομής για τις ανθρωπιστικές επιστήμες ΔΥΑΣ

[63] Buerger, T. & Evain, J-P. & Champin P-A. (2011), W3C Media Annotation Working Group RDF ontology (ma-ont), Consulted January 31, 2012. Available: <http://www.w3.org/ns/ma-ont.rdf> (see also <http://www.w3.org/TR/2011/PR-mediaont-10-20111129/>)

[64] Heath, T. & Bizer, C (2011), Linked Data: Evolving the Web into a Global Data Space (1st edition). Synthesis Lectures on the Semantic Web: Theory and Technology, 1:1, 1-136.

[65] Berners-Lee, T. (1998), Cool uris don't change, Consulted January 31, 2012. Available: <http://www.w3.org/Provider/Style/URI>

[66] Berners-Lee, T. (2006), Linked Data - Design Issues, Consulted January 31, 2012. Available: <http://www.w3.org/DesignIssues/LinkedData.htm>

[67] Fielding, R. (1999), Hypertext transfer protocol http/1.1, Consulted January 31, 2012. Available: <http://www.w3.org/Protocols/rfc2616/rfc2616.html>