



**ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ  
ΤΜΗΜΑ ΨΗΦΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ**

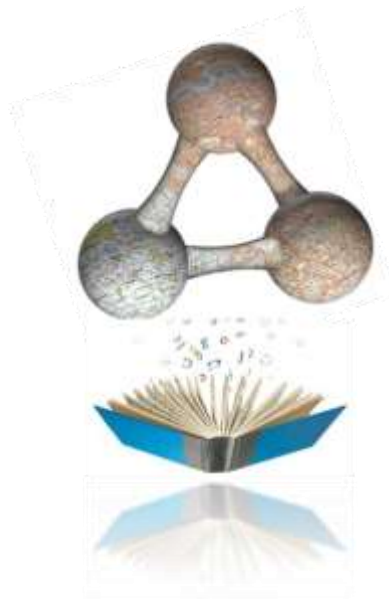
*Πρόγραμμα Μεταπτυχιακών Σπουδών:*

**«Διδακτική της Τεχνολογίας και Ψηφιακά Συστήματα»**

*Κατεύθυνση: «Ηλεκτρονική Μάθηση»*

**ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ**

*«Ανοιχτά Διασυνδεδεμένα Δεδομένα και Βιβλιοθήκες: η διάθεση και η  
διασύνδεση των βιβλιογραφικών δεδομένων της Δημοτικής Βιβλιοθήκης  
Ευόσμου»*



**Ζαχαρούλα Σαμοθρακίτου**

**(Α.Μ.: ΜΕ11039)**

**Επιβλέπων:**

**Γεώργιος Βούρος, Καθηγητής**

**Πειραιάς, Σεπτέμβριος 2015**



# Ευχαριστίες

Αισθάνομαι τη βαθύτατη ανάγκη να ευχαριστήσω τον κ. Γεώργιο Βούρο, Καθηγητή του Τμήματος Ψηφιακών Συστημάτων για την αποτελεσματική και άκρως εποικοδομητική συνεργασία μας, για την απόλυτα καθοριστική καθοδήγησή του καθώς επίσης και για την υπομονή που υπέδειξε ώστε να ολοκληρωθεί με απόλυτη επιτυχία η εν λόγω διπλωματική εργασία.

Ακόμα, είναι αναγκαίο να ευχαριστήσω τον Σωτήρη Καραμπατάκη, κάτοχο του μεταπτυχιακού διπλώματος «Επιστήμη του Διαδικτύου» του τμήματος Μαθηματικών του Αριστοτελείου Πανεπιστημίου Θεσσαλονίκης, για την παραχώρηση της εφαρμογής που χρησιμοποιήθηκε και την απεριόριστη βοήθεια του για την ολοκλήρωση της διπλωματικής αυτής αλλά και τον Γεώργιο Σαντιπαντάκη, μεταδιδάκτορα του τμήματος Ψηφιακών Συστημάτων, για την πολύτιμη βοήθεια και την προμηθεία του να προσφέρει τις γνώσεις τους για την επίλυση οποιονδήποτε προβλημάτων προέκυπταν στα πλαίσια της διπλωματικής εργασίας.

Επιπρόσθετα, δε θα μπορούσα να παραλείψω την έκφραση της ευγνωμοσύνης μου στους δικούς μου ανθρώπους, ιδιαίτερα στην οικογένεια μου και την καλύτερη μου φίλη και συμφοιτήτρια στο συγκεκριμένο μεταπτυχιακό πρόγραμμα σπουδών, Αικατερίνη Τσαρτσάλη, για την αμέριστη συμβολή και στήριξή τους όλο το διάστημα της φοίτησής μου, αφού χωρίς την αγάπη, την εμπιστοσύνη, την κατανόηση και την αμέριστη υποστήριξη τους, δε θα ήμουν σε θέση να επιτύχω τους στόχους μου.

*Στον πατέρα μου, Γιάννη..*

# Περίληψη

Η εμφάνιση του Παγκόσμιου Ιστού επηρέασε σημαντικά τις βιβλιοθήκες, καθώς έως τότε αποτελούσαν την κύρια πηγή γνώσης και πληροφόρησης. Η μετάβαση από τον Παγκόσμιο Ιστό στον Σημασιολογικό Ιστό επέλυσε σημαντικά προβλήματα που σχετίζονταν με την παροχή εμπλουτισμένης πληροφορίας στον τελικό χρήστη. Το άνοιγμα και η διασύνδεση των δεδομένων που προσφέρει η χρήση των τεχνολογιών του Σημασιολογικού Ιστού επιτρέπει την άμεση αξιοποίηση των πληροφοριών των δεδομένων που προσφέρονται από τις βιβλιοθήκες. Στην παρούσα διπλωματική εξετάζονται μέθοδοι σημασιολογικής περιγραφής βιβλιογραφικών δεδομένων και παρουσίασής τους στον τελικό χρήστη ως Διασυνδεδεμένα Δεδομένα. Πιο συγκεκριμένα, μετατράπηκαν τα βιβλιογραφικά δεδομένα της Δημοτικής Βιβλιοθήκης του Ευόσμου από τη μορφή MARC σε μορφή RDF και διασυνδέθηκαν με τα σύνολα δεδομένων της Δημόσιας Βιβλιοθήκης Βέροιας και της Dbpedia, εμπλουτίζοντας με αυτόν τον τρόπο τις υπηρεσίες που μπορεί να προσφέρει η βιβλιοθήκη αυτή. Τέλος, η εν λόγω διπλωματική αποτελεί μία προτεινόμενη πρακτική για τους βιβλιοθηκονόμους που ενδιαφέρονται να ανοίξουν και να διασυνδέσουν τα βιβλιογραφικά δεδομένα των βιβλιοθηκών που εργάζονται.

Λέξεις-κλειδιά: Σημασιολογικός Ιστός, Διασυνδεδεμένα Δεδομένα, Βιβλιοθήκες, Βιβλιοθήκες και Ανοιχτά Διασυνδεδεμένα Δεδομένα, Βιβλιογραφικά Δεδομένα, MARC

# Abstract

The emergence of the World Wide Web strongly influenced the Libraries, which until then were the main source of knowledge and information. The transition from the Web to the Semantic Web has solved many significant problems related to the supply of enhanced information to the end user. The opening and interconnection of data provided by the use of Semantic Web technologies allows the direct use of the data information offered by the libraries. This thesis examines methods that semantically describe bibliographic data and present these to the end user as Linked Data. In particular, the bibliographic data of the Public Library of Evosmos were transformed from the MARC format to the RDF format and were interconnected with the datasets of the Public Library of Veria and of DBpedia. Finally, this thesis can be a recommended practice for librarians who are interested in opening and linking the bibliographic data of the libraries they work in.

Keywords: Semantic Web, Linked Data, Libraries, Libraries and Linked Data, Bibliographic Data, MARC

# Περιεχόμενα

Ευχαριστίες.....	3
Περίληψη .....	5
Abstract.....	6
Περιεχόμενα.....	7
1.1. Αντικείμενο εργασίας.....	13
1.2. Δομή εργασίας.....	14
2. Θεωρητικό υπόβαθρο.....	16
2.1. Καταλογογράφηση και πρότυπα.....	16
2.1.1. Καταλογογράφηση.....	16
2.1.2. Πρότυπα Βιβλιοθηκών.....	16
2.1.2.1. Τα πρότυπα MARC ( <i>Machine Readable Catalogue</i> ).....	17
2.2. Σημασιολογικός Ιστός.....	19
2.3. Αρχιτεκτονική Σημασιολογικού Ιστού.....	19
2.4. Τεχνολογίες Σημασιολογικού Ιστού.....	26
2.4.1. URI γ.....	26
2.4.2. Συντάξεις RDF.....	27
2.4.3. Οντολογίες ( <i>Ontologies</i> ).....	30
2.4.4. Αποθετήρια τριπλετών ( <i>triplestores</i> ).....	32
2.5. Ανοιχτά Διασυνδεδεμένα Δεδομένα.....	34
2.5.1. Διασυνδεδεμένα Δεδομένα και Βιβλιοθήκες.....	35
2.5.1.1. <i>LOD Cloud</i> και Βιβλιοθήκες.....	36
3. Πρακτικές Ανοιχτών Διασυνδεδεμένων Δεδομένων σε Βιβλιοθήκες.....	40
3.1. BNB ( <i>British National Bibliography</i> ).....	40
3.2. <i>Bibliothèque Nationale de France</i> .....	44

3.3. Deutsche Nationalbibliothek (DNB).....	47
3.4. Biblioteca Nacional de España (BNE).....	49
3.5. Δημόσια Βιβλιοθήκη Βέροιας.....	51
4. Δημοτική Βιβλιοθήκη Ευόσμου και Ανοιχτά Διασυνδεδεμένα Δεδομένα: Προτεινόμενη Πρακτική.....	54
4.1. Εξαγωγή Δεδομένων Δημοτικής Βιβλιοθήκης Ευόσμου.....	55
4.1.1. Κατάλογος Βιβλιοθήκης (ΑΒΕΚΤ 5.5).....	55
4.1.2. Μορφή εξαγωγίμου αρχείου δεδομένων.....	56
4.2. Σημασιολογική Διαχείριση Δεδομένων.....	61
4.2.1. Λεξιλόγια εφαρμογής.....	63
4.2.2. Αρχικοποίηση Μοντέλου Δεδομένων.....	64
4.2.3. Ανάγνωση Εγγραφής.....	65
4.2.4. Μετατροπή σε RDF.....	66
4.3. “Ανοιγμα” Δεδομένων.....	68
4.3.1. Εικονικό Μηχάνημα (Virtual Machine).....	68
4.3.2. Εξυπηρετητής OpenLink Virtuoso.....	68
4.3.2.1. Φόρτωση αρχείου τριπλετών.....	69
4.3.2.2. Αποσαφήνιση URI.....	70
4.4. Μεταφόρτωση RDF/XML αρχείου στο Datahub.....	74
4.5. Διασύνδεση Δεδομένων.....	76
4.5.1. Εύρεση κι επιλογή άλλων ανοιχτών δεδομένων.....	77
4.5.2. Επιλογή κατάλληλης εφαρμογής για τη διασύνδεση.....	77
4.5.3. Silk Workbench: παράδειγμα διασύνδεσης συνόλων δεδομένων.....	78
4.6. Παρουσίαση Δεδομένων.....	83
Συμπεράσματα.....	88
Βιβλιογραφία.....	89



# Πίνακας εικόνων

Εικόνα 1: Παράδειγμα βιβλιογραφικής εγγραφής Παντείου Πανεπιστημίου.....	18
Εικόνα 2: Παράδειγμα MARC βιβλιογραφικής εγγραφής Παντείου Πανεπιστημίου .....	18
Εικόνα 3: Αρχιτεκτονική του Σηματολογικού Ιστού .....	20
Εικόνα 4: Παράδειγμα 303 URIs .....	27
Εικόνα 5: Παράδειγμα Hash URIs .....	27
Εικόνα 6: Παράδειγμα σε μορφή αρχείου N3.....	28
Εικόνα 7: Παράδειγμα σε μορφή αρχείου CSV .....	29
Εικόνα 8: Παράδειγμα μορφής αρχείου Turtle.....	29
Εικόνα 9: Παράδειγμα σε μορφή αρχείου JSON .....	30
Εικόνα 10: Μονόδρομα κι αμφίδρομα τόξα στο LOD Cloud.....	38
Εικόνα 11: Κατηγορία "Publications" στο LOD Cloud .....	38
Εικόνα 12: Sparql Endpoint της BNB.....	43
Εικόνα 13: Παράδειγμα ερωτήματος SPARQL στο Endpoint της BNB.....	44
Εικόνα 14: Δείγμα αποτελεσμάτων από το SPARQL Endpoint της BNB.....	44
Εικόνα 15: Παράδειγμα ερωτήματος SPARQL στο Endpoint της BNF .....	46
Εικόνα 16: Δείγμα αποτελεσμάτων από το SPARQL Endpoint της BNF .....	47
Εικόνα 17: Παράδειγμα Ερωτήματος SPARQL στο Endpoint της Δημόσιας Βιβλιοθήκης Βέροιας.....	52
Εικόνα 18: Δείγμα αποτελεσμάτων από το SPARQL Endpoint της Δημόσιας Βιβλιοθήκης Βέροιας.....	53
Εικόνα 19: Βιβλιοθήκες που διαθέτουν το ABEKT .....	56
Εικόνα 20: Παράδειγμα δεδομένων Βιβλιοθήκης σε μορφή ISO 2709 .....	58
Εικόνα 21: Ανάλυση δομής record.....	65
Εικόνα 22: Ανάλυση δομής marc_field.....	66
Εικόνα 23: Ανάλυση δομής tags .....	66
Εικόνα 24: Παράδειγμα εγγραφής στο αρχείο n3 της Δημοτικής Βιβλιοθήκης Ευόσμου .....	67
Εικόνα 25: Εγκατάσταση OpenLink Virtuoso 6.1. ....	69
Εικόνα 26: Φόρτωσης τριπλετών από το μενού του Virtuoso .....	70
Εικόνα 27: Τελικό Σημείο SPARQL Δημοτικής Βιβλιοθήκης Ευόσμου .....	70
Εικόνα 28: Σχηματική απεικόνιση 303 ανακατεύθυνσης URI .....	71
Εικόνα 29: Μορφή 303 ανακατεύθυνσης - Ιστοσελίδα HTML.....	72
Εικόνα 30: Μορφή 303 ανακατεύθυνσης - RDF αρχείο.....	73
Εικόνα 31: Δημιουργία λογαριασμού στο Datahub .....	75
Εικόνα 32: Βιβλιογραφικά δεδομένα Δημοτικής Βιβλιοθήκης Ευόσμου στο Datahub.....	76
Εικόνα 33: Επιλογές σύνδεσης στο Silk.....	78
Εικόνα 34: Διεπαφή χρήστη Silk Workbench .....	79
Εικόνα 35: Ρυθμίσεις αρχικών δηλώσεων του Silk.....	80
Εικόνα 36: Επιλογή πεδίων από τα σύνολα δεδομένων στο Silk.....	80
Εικόνα 37: Επιλογή μετασχηματισμών για τα σύνολα δεδομένων στο Silk.....	81
Εικόνα 38: Επιλογή τελεστή σύγκρισης δεδομένων στο Silk.....	81

Εικόνα 39: Επιλογή τελεστών συνάθροισης στο Silk.....	82
Εικόνα 40: Σχηματική απεικόνιση κανόνων στον Editor του Silk.....	82
Εικόνα 41: Σύνδεση owl:sameAs στο SPARQL ENdpoint.....	82
Εικόνα 42: Κεντρικό Μενού Drupal.....	84
Εικόνα 43: Αρχική σελίδα Drupal.....	84
Εικόνα 44: Μενού "Διασυνδεδεμένα Δεδομένα" στο Drupal.....	85
Εικόνα 45: Μενού "Σημασιολογική Αναζήτηση" στο Drupal.....	86
Εικόνα 46: Μενού "SPARQL Endpoint" στο Drupal.....	87

# Εισαγωγή

Με την εμφάνιση του Παγκόσμιου Ιστού, οι Βιβλιοθήκες ήταν από τους φορείς που επηρεάστηκαν δραματικά κι αυτό γίνεται κατανοητό αν αναλογιστεί κανείς πως ως τότε ήταν η κύρια και βασική πηγή γνώσης και πληροφόρησης.

Οι Βιβλιοθήκες προσπάθησαν να προσαρμοστούν και να εναρμονίσουν τις υπηρεσίες τους όλα αυτά τα χρόνια με τα νέα δεδομένα των εποχών (Park, 2015). Πιο ειδικά, οι έντυποι δελτιοκατάλογοι αντικαταστάθηκαν με την ηλεκτρονική καταλογογράφηση βάσει προτύπων (όπως MARC και UNIMARC) και γενικά πραγματοποιήθηκε η αυτοματοποίηση των τεκμηρίων των Βιβλιοθηκών.

Η επιρροή αυτή έγινε μεγαλύτερη αφού τα τελευταία χρόνια η μορφή του διαδικτύου αλλάζει και παρατηρείται η μετάβαση από τον Παγκόσμιο Ιστό που είναι γεμάτος με υπερσυνδέσμους, σ' έναν «έξυπνο» παγκόσμιο ιστό, τον Σημαιολογικό Ιστό. Σκοπός δεν είναι απλά η δημοσίευση εγγράφων και δεδομένων, αλλά αντιθέτως, η δημιουργία δομημένων δεδομένων και η σύνδεσή τους. Η ωρίμανση του Σημαιολογικού Ιστού οδήγησε στο Δίκτυο των Δεδομένων (Web of data), το οποίο αποτελείται από Διασυνδεδεμένα Δεδομένα (Bizer, Heath, Idehen & Berners-Lee, 2008).

Το πέρασμα από τον Ιστό 2.0 στον Ιστό 3.0 δε μπορούσε να αφήσει ανεπηρέαστες τις Βιβλιοθήκες, λύνοντας ταυτόχρονα και πολλά προβλήματα στην παροχή εμπλουτισμένης πληροφορίας στον τελικό χρήστη αλλά και δημιούργησε νέες δυνατότητες στην επικοινωνία μεταξύ των Βιβλιοθηκών. Αυτό γίνεται απόλυτα νοητό καθώς οι κατάλογοι των Βιβλιοθηκών κατέχουν τέτοιο πλούτο πληροφοριών, ο οποίος έμενε αναξιοποίητος με τις τεχνολογίες του Ιστού 2.0. Το Άνοιγμα και η Διασύνδεση των Δεδομένων ήρθε να δώσει λύση στην αξιοποίηση των πληροφοριών των δεδομένων που παρέχονται από τους καταλόγους των Βιβλιοθηκών.

Πιο ειδικά, η διαδικασία της τεκμηρίωσης είναι από τις πλέον σημαντικές λειτουργίες ενός πολιτιστικού/ εκπαιδευτικού φορέα και εξυπηρετεί στον εντοπισμό και τη διαχείριση τόσο των εγγραφών όσο και των ψηφιακών τους υποκατάστατων, στην ερμηνεία και την παροχή πρόσβασης σε αυτά. Η τεκμηριωτική πολιτιστική πληροφορία αντιπροσωπεύει έναν τεράστιο όγκο γνώσης γύρω από την προέλευση και την ερμηνεία των πολιτιστικών εγγραφών και συγκροτεί ένα από τα πλέον κατάλληλα σύνολα δεδομένων για δημοσίευση ως Διασυνδεδεμένα Δεδομένα, καθώς παραδοσιακά οι πολιτιστικοί φορείς χρησιμοποιούν πρότυπα για την οργάνωση και την ταξινόμηση της πληροφορίας, και άλλες δομημένες γλωσσικές πηγές όπως λεξιλόγια, θησαυρούς, κ.λπ., για την περιγραφή του υλικού και άυλου πολιτισμικού τους πλούτου. Με το άνοιγμά της πληροφορίας αυτής, καθίσταται δυνατή η διασύνδεσή της με άλλες πηγές δομημένης πληροφορίας όπως ημερομηνίες γεννήσεως και θανάτου δημιουργών, γεωγραφικές συντεταγμένες τοπωνυμίων, πληροφορίες από τη Βικιπαίδεια και άλλες έγκριτες πηγές πληροφόρησης, μετατρέποντας πολύ πιο πλούσια την εμπειρία του χρήστη (Ανοιχτά Δεδομένα: Η πρώτη ύλη για την Κοινωνία της Γνώσης, 2014).

Πολύ γρήγορα, πολλά αποθετήρια γνώσης και ειδικότερα οι Βιβλιοθήκες συνειδητοποίησαν τα ωφέλη από την υιοθέτηση τεχνολογιών βασισμένων σε συνδέσεις μεταξύ των διαφορετικών πηγών δεδομένων (Paercke, Chang, Winograd, & Garcia-Molina, 1998), ή αλλιώς των Διασυνδεδεμένων Δεδομένων.

Οι υπηρεσίες αυτές δίνουν τη δυνατότητα στις βιβλιοθήκες να ενσωματώσουν τη σημασιολογική πληροφορία που υπάρχει στις παραδοσιακές τους υπηρεσίες, όπως είναι ο κατάλογος ανοιχτής πρόσβασης (OPAC – Open Public Access Catalog) με πληροφορίες που προέρχονται από τρίτους, όπως είναι τα αρχεία βίντεο και εικόνων, ή βάσεις γνώσης (π.χ. DBpedia) (Κυπριανός & Παπαδάκης, 2012).

Στο σημείο αυτό είναι αναγκαίο να ειπωθεί πως το άνοιγμα και η διασύνδεση των δεδομένων των βιβλιοθηκών έχουν αξία τόσο για το απευθυνόμενο κοινό τους, όσο και για την κοινωνία στο σύνολο της. Ένας τομέας στον οποίο παρατηρείται προστιθέμενη αξία είναι κι η εκπαίδευση, η οποία είναι άμεσα συνυφασμένη με τις βιβλιοθήκες.

Πιο συγκεκριμένα, η δια βίου μάθηση είναι ένας από τους στόχους των Ιδρυμάτων της Πρωτοβάθμιας, Δευτεροβάθμιας και Τριτοβάθμιας Εκπαίδευσης. Δια βίου μάθηση μπορεί να οριστεί κάθε σκόπιμη μαθησιακή δραστηριότητα που διενεργείται σε συνεχή βάση με σκοπό τη βελτίωση των γνώσεων, των δεξιοτήτων και των ικανοτήτων του ατόμου. Εμπεριέχει όλες τις μορφές εκπαίδευσης και κατάρτισης, τυπικής, μη-τυπικής και άτυπης. Όλες αυτές οι μορφές εκπαίδευσης και κατάρτισης βασίζονται σε μεθόδους ανάπτυξης της ικανότητας του ατόμου να αναζητεί πληροφορίες και να οικοδομεί τις γνώσεις του ανεξάρτητα. Οι βιβλιοθήκες έχουν τη δυνατότητα να κάνουν τη διαφορά ανάμεσα σε ένα παραδοσιακό σύστημα τυπικής εκπαίδευσης και ένα ευρύτερο σύστημα μάθησης. Το ευρύτερο αυτό σύστημα μάθησης δεν θα περιορίζεται στα στενά όρια των βιβλίων ή των συγγραμμάτων, αλλά θα χαρακτηρίζεται από την πρόσβαση του δια βίου εκπαιδευόμενου σε αμέτρητες ποιοτικά εγγυημένες πηγές του εκάστοτε ενδιαφερόμενου πεδίου. Μέσα από αυτό δημιουργείται ένα εξατομικευμένο περιβάλλον μάθησης για τον εκάστοτε εκπαιδευόμενο που οδηγεί στην προσωπική οικοδόμηση της γνώσης. Για τους παραπάνω λόγους θεωρείται πως οι βιβλιοθήκες είναι αναγκαίο συμπλήρωμα των Ιδρυμάτων Εκπαίδευσης (Häggström, 2004).

Με βάση τη σημαντικότητα των βιβλιοθηκών στον τομέα της εκπαίδευσης, στο σημείο αυτό είναι αναγκαίο να ειπωθεί πως το Άνοιγμα και η Διασύνδεση των βιβλιογραφικών εγγραφών που εμπεριέχονται στους καταλόγους των βιβλιοθηκών (OPACs) μπορούν να προσφέρουν πληροφορίες που είναι πιο ολοκληρωμένες κι εμπλουτισμένες και να ενισχύσουν ακόμα περισσότερο την μαθησιακή εμπειρία ή να διευκολύνουν την οικοδόμηση της γνώσης από το ίδιο το άτομο. Για να γίνει αυτό πιο κατανοητό είναι απαραίτητο να δοθεί ένα παράδειγμα στο πεδίο.

Ας υποθέσουμε πως ένας φοιτητής της Φιλοσοφικής Σχολής έχει αναλάβει μία εργασία σχετικά με τον ποιητή «Οδυσσέα Ελύτη» και το έργο του. Με μία απλή αναζήτηση σε έναν κατάλογο βιβλιοθήκης (OPAC) με λέξεις κλειδιά το ονοματεπώνυμο του ποιητή, θα επιστραφούν αποτελέσματα που απλώς πληροφορούν τον φοιτητή για τους τίτλους των βιβλίων που είναι διαθέσιμοι στην εν λόγω βιβλιοθήκη που αναζητά αλλά κι επιπρόσθετες

πληροφορίες, όπως η τοποθεσία των βιβλίων στα ράφια και η διαθεσιμότητα τους για δανεισμό.

Από την άλλη μεριά, αν τα δεδομένα του καταλόγου αυτού έχουν «ανοιχτεί» και έχουν «διασυνδεθεί» με δεδομένα άλλων καταλόγων ή δεδομένα άλλων πηγών όπως είναι η Βικιπαίδεια κι η Europeana, θα του επιστραφούν αποτελέσματα που θα δίνουν συνδυαστικές και συγκριτικές πληροφορίες για τους τίτλους των βιβλίων που προσφέρονται από διάφορες βιβλιοθήκες, όπως και για τις διάφορες εκδόσεις κ.α.. Ενώ αντίστοιχα στη διασύνδεση με δεδομένα άλλων πηγών, θα επιστρέφονται βιβλιογραφικές πληροφορίες για τον ίδιο τον ποιητή από την Βικιπαίδεια αλλά και ορισμένα ψηφιοποιημένα έργα του, που βρίσκονται διαθέσιμα στο πολιτισμικό αποθετήριο της Europeana.

Με τον τρόπο αυτό η εμπειρία του τελικού χρήστη διέπεται από την πρόσβαση σε πληροφορίες με μεγαλύτερη εκπαιδευτική αξία και μπορεί να χαρακτηριστεί ως συνδυαστική πρόσβαση στη γνώση για ένα συγκεκριμένο θέμα με πιο εύκολο και γρήγορο τρόπο.

## 1.1. Αντικείμενο εργασίας

Η παρούσα διπλωματική κινείται στο πεδίο του Σημασιολογικού Ιστού και πιο ειδικά, τα Ανοιχτά και Διασυνδεδεμένα Δεδομένα και την εφαρμογή τους στις Βιβλιοθήκες. Η εργασία αυτή διακυβεύεται την πρόταση μίας πρακτικής για το άνοιγμα και τη διασύνδεση των δεδομένων βιβλιοθηκών και απευθύνεται σε βιβλιοθήκες που έχουν τα εξής χαρακτηριστικά:

- ✓ Πραγματοποιούν καταλογογράφηση στο υλικό τους
- ✓ Χρησιμοποιούν το Αυτοματοποιημένο Σύστημα Βιβλιοθηκών ΑΒΕΚΤ
- ✓ Διαθέτουν βιβλιοθηκονόμους στο προσωπικό τους
- ✓ Οι βιβλιοθηκονόμοι διαθέτουν βασικές προγραμματιστικές γνώσεις
- ✓ Επιθυμούν να συνεισφέρουν στο Σύννεφο των Ανοιχτών και Διασυνδεδεμένων Δεδομένων δημοσιοποιώντας τα δεδομένα τους

Εν συνεχεία, για την ανάπτυξη της πρακτικής αυτής τέθηκαν ο σκοπός και οι επιμέρους στόχοι της. Πιο ειδικά, ως σκοπός της διπλωματικής αυτής εργασίας ορίζεται η παράθεση της βήμα – βήμα διαδικασίας του «άνοιγματος» και της «διασύνδεσης» των βιβλιογραφικών δεδομένων μιας βιβλιοθήκης, που αποτελεί αντιπροσωπευτικό δείγμα στην Ελλάδα και υλοποιείται από έναν βιβλιοθηκονόμο με βασικές προγραμματιστικές γνώσεις.

Ακόμα, τέθηκαν επιπλέον επιμέρους στόχοι που συνοψίζονται στα εξής:

1. Η συμβολή στο Σύννεφο των Διασυνδεδεμένων Δεδομένων
2. Οι δυνατότητες που προσφέρουν τα Ανοιχτά και Διασυνδεδεμένα Δεδομένα στον κόσμο των βιβλιοθηκών
3. Η παρακίνηση των υπόλοιπων ελληνικών βιβλιοθηκών για το «άνοιγμα» και τη «διασύνδεση» των δεδομένων τους.

Η μεθοδολογία που ακολουθήθηκε για την επιτυχή ολοκλήρωση της εργασίας αυτής παρουσιάζεται στα εξής βήματα:

**1° Βήμα:** Μελέτη της βιβλιογραφίας για παρόμοιες πρακτικές υλοποίησης

**2° Βήμα:** Εύρεση αντιπροσωπευτικής βιβλιοθήκης

**3° Βήμα:** Εξαγωγή βιβλιογραφικών εγγραφών από το αυτοματοποιημένο σύστημα της βιβλιοθήκης

**4° Βήμα:** Μετατροπή βιβλιογραφικών δεδομένων σε ανοιχτά δεδομένα

**5° Βήμα:** Επίτευξη διασύνδεσης βιβλιογραφικών δεδομένων με άλλα δεδομένα

**6° Βήμα:** Παρουσίαση Ανοιχτών και Διασυνδεδεμένων Δεδομένων σε ένα περιβάλλον φιλικό προς τον χρήστη

## 1.2. Δομή εργασίας

Έχοντας κατά νου τον σκοπό και τους στόχους της εν λόγω διπλωματικής και επιδιώκοντας την - από το γενικό προς το ειδικό- παρουσίαση της, τα κεφάλαια δομήθηκαν ως εξής:

*2ο Κεφάλαιο: Θεωρητικό Υπόβαθρο*

Στο κεφάλαιο αυτό παρουσιάζονται κι αποσαφηνίζονται οι βασικές έννοιες που πραγματεύεται η εργασία αυτή. Πιο ειδικά, παρουσιάζονται τα πρότυπα των Βιβλιοθηκών, οι έννοιες του Σημαιολογικού Ιστού, η αρχιτεκτονική και οι τεχνολογίες που αξιοποιήθηκαν κατά την υλοποίηση της εργασίας αυτής. Ακόμα, δε θα μπορούσαν να λείπουν από το κεφάλαιο αυτό οι έννοιες των Ανοιχτών και Διασυνδεδεμένων Δεδομένων και μία σύνδεση αυτών με τις Βιβλιοθήκες.

*3ο Κεφάλαιο: Πρακτικές Ανοιχτών Διασυνδεδεμένων Δεδομένων σε Βιβλιοθήκες*

Το τρίτο κεφάλαιο είναι μία αναλυτική παρουσίαση των πρακτικών που έχουν υλοποιηθεί στο εξωτερικό και σχετίζονται με το θέμα της εργασίας αυτής. Στην ουσία, πρόκειται για μία βιβλιογραφική ανασκόπηση που τονίζει τα δυνατά και αδύναμα σημεία της κάθε πρακτικής και οδηγεί στην υιοθέτηση ή μη των εκάστοτε τεχνολογιών ή εφαρμογών για την ολοκλήρωση της διπλωματικής αυτής.

*4ο Κεφάλαιο: Δημοτική Βιβλιοθήκη Ευόσμου και Ανοιχτά Διασυνδεδεμένα Δεδομένα: Προτεινόμενη Πρακτική*

Το κεφάλαιο 4 είναι η αναλυτική και βήμα – βήμα παρουσίαση της διαδικασίας εύρεσης των βιβλιογραφικών εγγραφών της Δημοτικής Βιβλιοθήκης Ευόσμου, στη συνέχεια την διαδικασία ανοίγματος αυτών, τη διασύνδεση τους με τα βιβλιογραφικά δεδομένα της Δημόσιας Βιβλιοθήκης Βέροιας και τα δεδομένα των συγγραφέων με τη Dbpedia και τέλος,

την παρουσίαση αυτών των Ανοιχτών και Διασυνδεδεμένων Δεδομένων σε μία φιλική για τον χρήστη διεπαφή.

### *Συμπεράσματα*

Τέλος παρουσιάζονται τα συμπεράσματα που εξήχθησαν από την υλοποίηση της εν λόγω διπλωματικής και οι προτάσεις για μελλοντική εργασία.

### *Βιβλιογραφία*

Είναι αναγκαίο να αναφερθεί πως στο τελευταίο μη αριθμημένο κεφάλαιο παρατίθεται η βιβλιογραφία, η οποία έχει συνταχθεί με το διεθνές αναγνωρισμένο πρότυπο βιβλιογραφικής σύνταξης APA (The American Psychological Association).

## 2. Θεωρητικό υπόβαθρο

### 2.1. Καταλογογράφηση και πρότυπα

#### 2.1.1. Καταλογογράφηση

Οι βιβλιοθήκες παρέχουν υπηρεσίες και πρόσβαση σε έναν τεράστιο όγκο πληροφοριών. Αλλά για να είναι χρήσιμες αυτές οι πληροφορίες, το υλικό πρέπει να είναι οργανωμένο και προσβάσιμο. Ειδικότερα, πρέπει να περιγραφεί και να καταλογογραφηθεί αποτελεσματικά αλλά και πρέπει να ταξινομηθεί και να τοποθετηθεί με συνέπεια στα ράφια. Η εύκολη και γρήγορη ανάκτηση είναι απαραίτητη για τη προσφορά ενός φιλικού περιβάλλοντος στον χρήστη βιβλιοθηκών.

Για την κατανόηση των παραπάνω κρίνεται αναγκαίο να δοθεί ο ορισμός της καταλογογράφησης. Η καταλογογράφηση, επομένως, είναι η διαδικασία της δημιουργίας βιβλιογραφικών εγγραφών για τα τεκμήρια της βιβλιοθήκης, σύμφωνα με συγκεκριμένους κανόνες ή πρότυπα. Κατά την καταλογογράφηση γίνεται κωδικοποίηση των πληροφοριών που αφορούν ένα βιβλιογραφικό τεκμήριο (βιβλία, αλλά και οποιοδήποτε άλλο τεκμήριο, ακόμα και ηλεκτρονικά) και δημιουργούνται εγγραφές σε μία βάση δεδομένων, τον κατάλογο της βιβλιοθήκης (Sales, 2005).

Η εγγραφή καταλογογράφησης είναι μια βιβλιογραφική εγγραφή, ή αλλιώς οι πληροφορίες που περιλαμβάνονται σε έναν κατάλογο βιβλιοθήκης για ένα τεκμήριο (What is a MARC record, and why it is important?, 2009).

#### 2.1.2. Πρότυπα Βιβλιοθηκών

Οι κατάλογοι των βιβλιοθηκών έχουν την ανάγκη εφαρμογής ενός προτύπου για την αυτοματοποιημένη ανάγνωση των δεδομένων. Τη δυνατότητα αυτή δίνουν τα MARC formats. Η δημιουργία τους ξεκίνησε από τη Βιβλιοθήκη του Κογκρέσσου των ΗΠΑ κατά τη διάρκεια της δεκαετίας του 1960 με τη δημιουργία του LC-MARC (USMARC) format. Η μορφή αυτή μαζί με το ομοειδές CAN-MARC εναρμόνισαν τις διαφορές τους, με αποτέλεσμα στις αρχές του 1999 να εκδοθεί ένας τόμος με την ονομασία MARC21. Εν τω μεταξύ είχαν αναπτυχθεί από άλλες βιβλιογραφικές υπηρεσίες πάνω από 20 διαφορετικά MARC πρότυπα όπως το UKMARC. Το πρόβλημα των πολλών και διαφορετικών προτύπων, της ανταλλαγής τους και της συμβατότητάς τους προσπάθησε να λύσει το 1977 η IFLA με την έκδοση του UNIMARC (Γερόλιμος, Σίτας & Κώνστα, 2011).



### 2.1.2.1. Τα πρότυπα MARC (Machine Readable Catalogue)

Οι πληροφορίες από τον κατάλογο δεν μπορούν απλά να είναι δακτυλογραφημένες σε έναν υπολογιστή για να παράγουν έναν αυτοματοποιημένο κατάλογο. Ο υπολογιστής χρειάζεται ένα μέσο για την ερμηνεία των πληροφοριών που βρίσκονται στο αρχείο της καταλογογράφησης. Για το λόγο αυτό εμφανίστηκαν τα πρότυπα MARC που περιέχουν έναν οδηγό για τα δεδομένα των βιβλιογραφικών εγγραφών και προσθέτουν "ταμπέλες" πριν από κάθε κομμάτι των βιβλιογραφικών πληροφοριών τεκμήριο (What is a MARC record, and why it is important?, 2009).

Το ακρωνύμιο MARC σημαίνει "καταλογογράφηση αναγνώσιμη από μηχανή" και μ' αυτό νοείται ότι ένας συγκεκριμένος τύπος μηχανής, δηλ. ένας υπολογιστής, μπορεί να διαβάσει και να ερμηνεύσει τα δεδομένα της εκάστοτε βιβλιογραφικής εγγραφής που έχει καταλογογραφηθεί.

Η αναζήτηση στοιχείων με εγγραφές MARC βασίζεται σε αναγνωριστικές ετικέτες και κωδικούς υπο-πεδίων που εκχωρούνται σε κάθε πεδίο μιας εγγραφής (Furrie, 2003). Αυτό που πιο πάνω αναφέρθηκε γενικά ως "ταμπέλες". Συγκεκριμένα, μια εγγραφή MARC περιλαμβάνει:

- *Περιγραφή του αντικειμένου:* Συνήθως είναι ο τίτλος, η έκδοση, πληροφορίες δημοσίευσης κτλ.
- *Κύρια είσοδο και επιπρόσθετες εισόδους:* Πρόκειται για σημεία πρόσβασης, δηλαδή σημεία ανάκτησης της πληροφορίας του αντικειμένου στον κατάλογο της βιβλιοθήκης (για παράδειγμα, κάποιο βιβλίο είναι δυνατόν να υπάρχει στον κατάλογο με περισσότερους του ενός τίτλους).
- *Επικεφαλίδες θεμάτων:* Υπάρχουν διάφορες πρότυπες λίστες για επικεφαλίδες θεμάτων που μπορούν να χρησιμοποιηθούν προκειμένου να επιλεγούν τα θέματα στα οποία θα συμπεριληφθεί ένα αντικείμενο.
- *Ταξινόμηση:* Ο αριθμός αυτός χρησιμεύει στην τοποθέτηση των αντικειμένων του ίδιου θέματος στο ίδιο ράφι στη βιβλιοθήκη.

Βέβαια, στην πλειονότητα των βιβλιοθηκών, οι εγγραφές MARC περιλαμβάνουν περισσότερες πληροφορίες. Κάθε εγγραφή MARC αποτελείται από πεδία, κάθε ένα από τα οποία περιλαμβάνει μια πληροφορία (τίτλο, συγγραφέα κτλ). Οι εγγραφές έχουν σταθερό αριθμό πεδίων, κάθε ένα από τα οποία περιλαμβάνει σταθερό αριθμό χαρακτήρων. Ωστόσο, για τη δημιουργία ενός ορθού καταλόγου, η καλύτερη δομή αρχείου είναι αυτή που επιτρέπει απεριόριστο αριθμό πεδίων και χαρακτήρων μέσα στα πεδία.

Κάθε πεδίο περιλαμβάνει ετικέτες που χρησιμοποιούνται ως αναγνωριστικά των πεδίων, για χάριν ευκολίας, αντί των ονομάτων των πεδίων. Μια ετικέτα προσδιορίζει την πληροφορία που ακολουθεί στο πεδίο. Για παράδειγμα, η ετικέτα 100 περιέχει το όνομα του συγγραφέα. Οι ετικέτες αυτές έχουν κοινή μορφή για όλες τις εγγραφές MARC και περιλαμβάνουν τις εξής πληροφορίες:

Ετικέτα	Πληροφορίες
---------	-------------

1XX	Κύριες πληροφορίες
2XX	Τίτλος, έκδοση, πληροφορίες δημοσίευσης
3XX	Περιγραφή υλικής υπόστασης εγγράφου
4XX	Πληροφορίες σειράς
5XX	Σημειώσεις
6XX	Πληροφορίες θεμάτων
7XX	Άλλες πληροφορίες
8XX	Επιπρόσθετες πληροφορίες σειράς
9XX	Είναι ελεύθερες ετικέτες

Παρακάτω ακολουθεί ένα παράδειγμα βιβλιογραφικής εγγραφής του καταλόγου της Βιβλιοθήκης του Παντείου Πανεπιστημίου σε html μορφή και σε MARC:

**Κύρια Φυσικά πρόσωπα :** [Davies, J. \( N. John \)](#) W  
**:** [Studer, Rudi](#) W  
**Άλλα Φυσικά πρόσωπα :** [Warren, Paul](#) W  
**Τίτλος :** Semantic web technologies : trends and research in ontology-based systems / John Davies, Rudi Studer, Paul Warren  
**Δημοσίευση/Διάθεση :** Chichester, West Sussex, England : John Wiley & Sons , 2006, 2007 print  
**Φυσική Περιγραφή :** xiii, 312 σ. , 24 εκ.  
**ISBN :** 0-470-02596-4  
**Βιβλιογραφία, ευρετήρια :** Περιέχει ευρετήριο  
**Γλώσσα :** Αγγλική  
**Θέματα :** [Σημασιολογικός Ιστός](#)  
[Οντολογίες \(Ανάκτηση πληροφοριών\)](#)  
[Σημασιολογικά δίκτυα \(Θεωρία πληροφοριών\)](#)  
**Αριθμός Dewey :** 025.04

**Εικόνα 1: Παράδειγμα βιβλιογραφικής εγγραφής Παντείου Πανεπιστημίου**

Η παραπάνω εγγραφή εμφανίζεται στο χρήστη με την παραπάνω μορφή, ωστόσο, για να επιτευχθεί αυτό έχει καταλογογραφηθεί με το πρότυπο MARC που παρουσιάζεται παρακάτω για την ίδια ακριβώς εγγραφή:

Πεδία MARC

```
01059nam 2200301 i 450
001 GR-AtPPVb10069400
003 GR-AtPPV
005 20080821125231.0
010 $a 0-470-02596-4
029 $a ΕΠΕΑΕΚ
100 $a 20080821e20072006 0grey0105 ba
101 0 $a eng
200 1 $a Semantic web technologies $e trends and research in ontology-based systems $f John Davies, Rudi Studer, Paul Warren
210 $a Chichester, West Sussex, England $c John Wiley & Sons $d 2006, 2007 print
215 $a xiii, 312 σ. $d 24 εκ.
320 $a Περιέχει ευρετήριο
606 0 $a Σημασιολογικός Ιστός
606 0 $a Οντολογίες (Ανάκτηση πληροφοριών)
606 0 $a Σημασιολογικά δίκτυα (Θεωρία πληροφοριών)
676 $a 025.04 $v 22η
700 1 $a Davies $b J. $g N. John
701 1 $a Studer $b Rudi
702 1 $a Warren $b Paul
709 $a John Wiley & Sons
801 0 $a GR $b Πάντειο Πανεπιστήμιο $g AACR2
960 $a 2008-BIBLIO-2
970 $a Κρητικού $b Ελένη $z 2008
852 $a GR-AtPPV $b PANTIO $b LIBRARY $h 025.04 DAV $7 10070875 $p 002000081421 $t 1 $y Στο ραφι
```

**Εικόνα 2: Παράδειγμα MARC βιβλιογραφικής εγγραφής Παντείου Πανεπιστημίου**

Όπως μπορεί να παρατηρήσει κανείς στο αριστερό μέρος εμφανίζονται τα πεδία MARC και στο δεξί μέρος τα υποπεδία και τα δεδομένα της εγγραφής αυτής.

## 2.2. Σημασιολογικός Ιστός

Ο Σημασιολογικός Ιστός (Semantic Web) είναι ένας ιστός δεδομένων, η ανάπτυξη του οποίου σηματοδοτεί την εξέλιξη του Παγκόσμιου Ιστού (World Wide Web) που αποτελείται σε μεγάλο βαθμό από υλικό προορισμένο για ανθρώπινη κατανάλωση, σε έναν ιστό που θα περιλαμβάνει δεδομένα και πληροφορίες, η διαχείριση των οποίων μπορεί να γίνει κι από υπολογιστές.

Πρόκειται για έναν ιστό πληροφορίας που απαρτίζεται από δεδομένα κι από διαδικασίες σημασιολογικής μετάφρασης συμβόλων. Μέσα από τις διαδικασίες αυτές, παρέχεται ένα είδος νοήματος κατά τη λογική σύνδεση όρων μέσω λεξιλογίων, ώστε να επιτευχθεί η διαλειτουργικότητα μεταξύ διαφορετικών συστημάτων (Shadbolt, Hall & Berners – Lee, 2006).

Ο εμπνευστής του Παγκόσμιου Ιστού, Tim Berners - Lee αλλά και άλλοι επιστήμονες του πεδίου, περιγράφουν ως αναγκαία τη μετάβαση της σύγχρονης εποχής από τον «Ιστό των Εγγράφων» προς τον «Ιστό των Δεδομένων», στον οποίο πολλά από τα δεδομένα επιμελούνται ώστε να διατίθενται μέσω των υπολογιστών, διαμορφωμένα με την απλή αρχιτεκτονική και την τεχνολογία του υπάρχοντος Ιστού (Miller, 2010). Η αναγκαιότητα αυτή προκύπτει από την έλλειψη λογικής στον σημερινό ιστό.

Όμως σε τι διαφέρει ο Σημασιολογικός Ιστός από τον Παγκόσμιο Ιστό; Ο τελευταίος δίνει τη δυνατότητα καλύτερης ολοκλήρωσης των δεδομένων, επιτρέποντας στον καθένα, ο οποίος τοποθετεί μεμονωμένα στοιχεία δεδομένων στον Ιστό, να τα συνδέσει με άλλα κομμάτια δεδομένων, χρησιμοποιώντας πρότυπες μορφοποιήσεις. Το αποτέλεσμα της χρήσης των τεχνολογιών του Σημασιολογικού Ιστού είναι να (Berners – Lee, 2007):

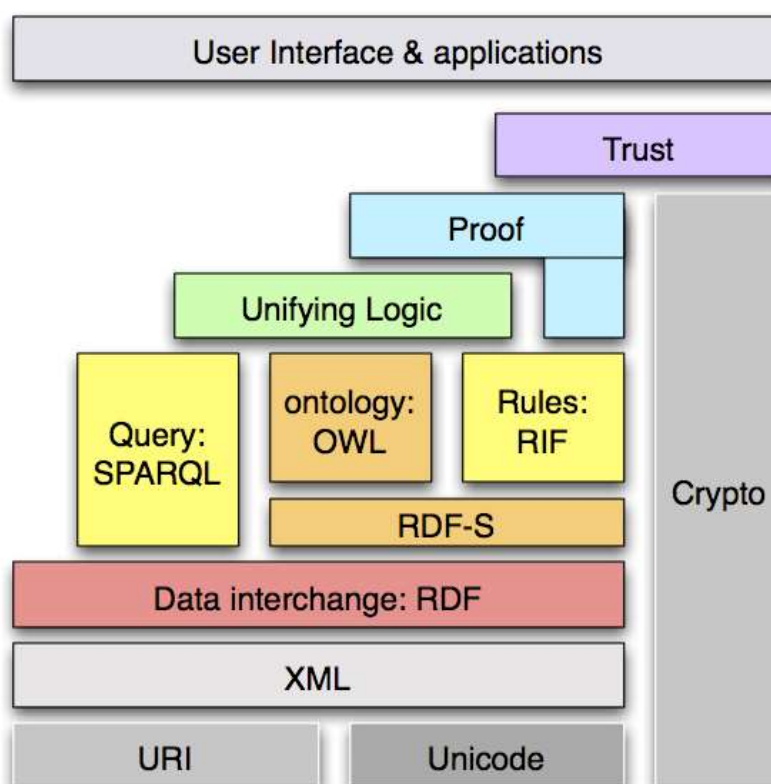
- ✚ δημοσιευθούν δομημένα δεδομένα στον Ιστό
- ✚ τεθούν συνδέσεις μεταξύ δεδομένων από μια πηγή δεδομένων προς δεδομένα άλλων πηγών (Συνδεδεμένα Δεδομένα - Linked Data).

## 2.3. Αρχιτεκτονική Σημασιολογικού Ιστού

Στο σημείο αυτό κρίνεται αναγκαία η επιγραμματική αναφορά στη δομή ή αλλιώς αρχιτεκτονική του σημασιολογικού ιστού, ώστε να γίνει αντιληπτή η εφαρμογή των εκάστοτε τεχνολογιών στην συγκεκριμένη πρακτική που αναλύεται στην εν λόγω διπλωματική εργασία καθώς οι περισσότερες από τις τεχνολογίες έχουν αξιοποιηθεί για την υλοποίηση της.

Ο Σημασιολογικός Ιστός αποτελείται από στρώματα (layers) και κάθε στρώμα υλοποιεί μια λειτουργικότητα (functionality), χρησιμοποιώντας και επεκτείνοντας τη λειτουργικότητα και τις τεχνολογίες που παρέχονται από τα χαμηλότερα στρώματα προς τα υψηλότερα. Έτσι λοιπόν στα χαμηλά επίπεδα υλοποιούνται λειτουργίες οι οποίες είναι πολύ κοντά στον Παγκόσμιο Ιστό και στις μηχανές, όπως είναι οι τεχνολογίες που ασχολούνται με τον καθορισμό και την αναγνώριση των πόρων (URIs). Με την έννοια πόρος αναφερόμαστε σε οποιαδήποτε οντότητα του Παγκόσμιου Ιστού, όπως είναι μια ιστοσελίδα, ένα τμήμα ή ένα σύνολο από ιστοσελίδες, ηλεκτρονικά αρχεία ή ακόμα και πόροι, οι οποίοι δεν είναι άμεσα διαθέσιμα στο διαδίκτυο, όπως είναι για παράδειγμα ένα βιβλίο (Στοιλος, χ.χ., σ. 1-2) (βλ. εικ. 3). Στην παρούσα εργασία πόροι θεωρούνται όλα τα βιβλία, τα περιοδικά, οι εγκυκλοπαίδειες, τα λεξικά και το οπτικοακουστικό υλικό που εμπεριέχεται στον κατάλογο της Δημοτικής Βιβλιοθήκης Ευόσμου, η οποία είναι η βιβλιοθήκη που επιλέχθηκε για την υλοποίηση της προτεινόμενης πρακτικής.

Όσο πραγματοποιείται άνοδος στην ιεραρχία των επιπέδων συναντώνται επίπεδα τα οποία υλοποιούν λειτουργικότητες αναπαράστασης γνώσης, πολύπλοκης συλλογιστικής κι εμπιστοσύνης που πλησιάζουν την ανθρώπινη σκέψη (Στοιλος, χ.χ., σ. 1) (βλ. εικ. 3).



**Εικόνα 3: Αρχιτεκτονική του Σημασιολογικού Ιστού**

Παρακάτω παρουσιάζονται πιο αναλυτικά οι διάφορες τεχνολογίες που εμπεριέχονται στα στρώματα του Σημασιολογικού Ιστού (βλ. εικ. 3). Τα κατώτερα επίπεδα της αρχιτεκτονικής του Σημασιολογικού Ιστού αφορούν πρότυπα που είναι κοινά με τον Παγκόσμιο Ιστό. Πιο ειδικά (Hogan, 2014):

- **Unicode (Χαρακτήρες)**

Όπως και ο Ιστός 2.0, έτσι κι ο Σημασιολογικός Ιστός απαιτεί κάποια πρότυπα για την αντιστοίχιση των δυαδικών συνόλων σε μορφή κειμένου. Για το σκοπό αυτό, ο Σημασιολογικός Ιστός βασίζεται στο πρότυπο Unicode, που είναι ένα πλήρες σύνολο χαρακτήρων διεθνώς αποδεκτό.

- **URIs (Αναγνωριστικά)**

Ο Σημασιολογικός Ιστός είναι απόλυτα συναφής με την περιγραφή των πόρων. Όσον αφορά την αναγνώριση αυτών των αντικειμένων ή αλλιώς πόρων υπάρχουν συγκεκριμένα αναγνωριστικά. Για τον λόγο αυτό χρησιμοποιούνται τα Uniform Resource Identifiers (URIs) (βλ. εικ. 3), τα οποία χρησιμοποιούνται ήδη στο διαδίκτυο για τον εντοπισμό των εγγράφων. Ένα νεότερο αναγνωριστικό πρότυπο του Σημασιολογικού Ιστού είναι τα Internationalized Resource Identifiers (IRIs), το οποίο είναι μια γενίκευση των URIs για την υποστήριξη του προτύπου Unicode, που αναφέρθηκε παραπάνω.

- **Σύνταξη**

Ο τρόπος μέσω του οποίου θα επιτραπεί στις μηχανές να αναλύσουν αυτόματα το περιεχόμενο των πόρων στα στοιχειώδη συστατικά τους είναι η χρήση σύνταξης με επίσημα καθορισμένες γραμματικές. Για το σκοπό αυτό, στον Σημασιολογικό Ιστό χρησιμοποιούνται υπάρχοντα πρότυπα/ γλώσσες, όπως XML και JSON (βλ. εικ. 3). Εκτός από τη χρήση των ήδη υπάρχοντων προτύπων/ γλωσσών έχουν δημιουργηθεί επίσης προσαρμοσμένα πρότυπα/ γλώσσες για να κωδικοποιήσουν τα δεδομένα του Σημασιολογικού Ιστού. Αυτά τα νέα πρότυπα είναι όλα παράγωγα της γλώσσας Terse RDF Triple Language (Turtle).

- **XML (Extensible Markup Language)**

Τι είναι όμως η XML που αποτελεί τη βάση του συγκεκριμένου στρώματος; Σύμφωνα με το W3C, η XML (Extensible Markup Language) είναι μία γλώσσα σήμανσης, που περιέχει ένα σύνολο κανόνων για την ηλεκτρονική κωδικοποίηση κειμένων. Η XML σχεδιάστηκε δίνοντας έμφαση στην απλότητα, τη γενικότητα και τη χρησιμότητα στο Διαδίκτυο. Τα XML έγγραφα αποτελούνται από μονάδες αποθήκευσης που καλούνται entities (στοιχεία), τα οποία περιέχουν πληροφορίες αναλυμένες ή μη. Οι αναλυμένες πληροφορίες αποτελούνται από χαρακτήρες (characters), οι οποίοι συνθέτουν χαρακτήρες δεδομένων (character data) και άλλοι οι οποίοι συνθέτουν μία επισήμανση (markup). Η μορφή markup κωδικοποιεί την περιγραφή της τελικής αποθήκευσης του εγγράφου καθώς και τη λογική δομή του.

Πάνω από το στρώμα της Σύνταξης βρίσκονται τα πιο βασικά δομικά συστατικά του Σημασιολογικού Ιστού (Hogan, 2014) (βλ. εικ. 3):

- **Μοντέλο Δεδομένων (Data Models)**

Προκειμένου οι μηχανές να ανταλλάξουν δεδομένα που είναι αναγνώσιμα από μηχανές, κρίνεται αναγκαία η χρήση ενός μοντέλου δεδομένων σύμφωνα με το οποίο θα δομηθεί το περιεχόμενο των πόρων. Το μοντέλο δεδομένων που χρησιμοποιείται στο Σημασιολογικό Ιστό είναι η RDF (Resource Description Framework) (βλ. εικ. 3).

- **RDF (Resource Description Framework)**

Στο σημείο αυτό δε θα μπορούσε να παραληφθεί η συνοπτική παρουσίαση της RDF, η οποία είναι μια γλώσσα που χρησιμοποιείται για την απλή περιγραφή πόρων του διαδικτύου. Η περιγραφή αυτή εστιάζεται κυρίως στην απόδοση μετα-πληροφορίας για τους πόρους αυτούς, όπως είναι η περιγραφή του τίτλου, του ονόματος, της ημερομηνίας δημιουργίας και άλλων χαρακτηριστικών κάποιου πόρου του διαδικτύου (Στοϊλος, χ.χ., σ.2).

Για την αναπαράσταση των δεδομένων που περιγράφουν πόρους στην RDF κάθε δήλωση (statement) είναι καταμεμημένη σε μορφή τριπλετών (triples) - υποκείμενο, κατηγορημα, αντικείμενο (subject, predicate, object). Οι τριπλέτες (triples) παρίστανται ως ένας κατευθυνόμενος γράφος, όπου το υποκείμενο και το αντικείμενο της εκάστοτε τριπλέτας (triple) διαμορφώνονται ως κόμβοι και το κατηγορημα ως κατευθυνόμενη σύνδεση που περιγράφει τη σχέση μεταξύ των κόμβων. Είναι αναγκαίο να σημειωθεί πως η κατεύθυνση του συνδέσμου δείχνει πάντα προς το αντικείμενο (Alexander & Ravada, χ.χ.).

Αντίστοιχα, οι πόροι που περιγράφονται από την RDF έχουν ιδιότητες. Πιο ειδικά, η RDF ορίζει έναν πόρο ως κάθε αντικείμενο που είναι άμεσα αναγνωρίσιμο από Uniform Resource Identifier (URI). Οι ιδιότητες (properties) που σχετίζονται με τους πόρους προσδιορίζονται από τύπους ιδιοτήτων (property types) και οι τύποι ιδιοτήτων έχουν αντίστοιχες τιμές (values). Οι τύποι ιδιοτήτων εκφράζουν τις σχέσεις των τιμών που συνδέονται με τους πόρους (Miller, 1998).

Στην RDF, οι τιμές μπορεί να έχουν ατομικό χαρακτήρα (μορφή απλή κειμένου, αριθμούς, κλπ) ή άλλους πόρους, που με τη σειρά τους μπορεί να έχουν τις δικές τους ιδιότητες. Μια συλλογή από αυτές τις ιδιότητες που αναφέρονται στον ίδιο πόρο ονομάζεται περιγραφή (description) (Miller, 1998).

- **Σχήματα & Οντολογίες (Schemas & Ontologies)**

Ενώ το μοντέλο δεδομένων της RDF φέρνει μια καθολική δομή στο περιεχόμενο, δεν φέρει σημασιολογία ή αλλιώς νόημα σε αυτό. Έτσι, ο Σημασιολογικός Ιστός απαιτεί μία επίσημη γλώσσα με την οποία να προβάλλονται αξιώματα και το νόημα των πραγμάτων που περιγράφονται στο περιεχόμενο της RDF.

Αυτές οι επίσημες γλώσσες προσφέρουν ένα μετα-λεξιλόγιο με σαφώς καθορισμένη σημασιολογία που μπορεί να χρησιμοποιηθεί σε συνδυασμό με την RDF για τον ορισμό σχημάτων και οντολογιών με σκοπό την αναπαράσταση και την περιγραφή των δεδομένων της RDF. Οι βασικές γλώσσες που προσφέρονται ως μέρος των ισχυόντων κανόνων του Σημασιολογικού Ιστού είναι η RDF Schema (RDFS) και η Web Ontology Language (γνωστή κι ως OWL) (βλ. εικ. 3). Παρακάτω ακολουθεί η συνοπτική περιγραφή τους:

- **RDFS (Resource Description Framework Schema)**

Η RDFS (RDF Schema Specification) παρέχει ένα τυπικό μηχανισμό για τον ορισμό κλάσεων (classes) και ιδιοτήτων (properties), καθώς και για τον καθορισμό των σχέσεων μεταξύ τους, χρησιμοποιώντας τη μοντέλο RDF (Pan & Horrocks, 2003). Η RDF Schema

ορίζει περαιτέρω βασικές έννοιες μοντελοποίησης στην RDF. Δηλαδή, η RDF Schema επεκτείνει (ή εμπλουτίζει) την RDF δίνοντας μια καθορισμένη σημασιολογία σε συγκεκριμένους πόρους (Broekstra, κ. α., 2001).

Πιο αναλυτικά, η RDF Schema είναι ένα σύνολο κλάσεων με καθορισμένες ιδιότητες που χρησιμοποιούν το μοντέλο δεδομένων της RDF, παρέχοντας βασικά στοιχεία για την περιγραφή των οντολογιών με σκοπό τη δόμηση των RDF πόρων (Brickley & Guha, 2014).

- **OWL (Web Ontology Language)**

Η OWL (Web Ontology Language), όπως κι οι RDF, RDFs, σχεδιάστηκε για τις εφαρμογές που είναι αναγκαίο να επεξεργαστούν το περιεχόμενο των πληροφοριών αντί να παρουσιάσουν μόνο τις πληροφορίες στους ανθρώπους. Ο σκοπός της OWL, όπως έχει διατυπωθεί από το W3C, είναι να παραχθεί ένα τυποποιημένο σχήμα που είναι συμβατό με την αρχιτεκτονική του Παγκόσμιου Ιστού αλλά και του Σημασιολογικού Ιστού, το οποίο σχεδιάστηκε για να αναπαραστήσει την πλούσια και πολύπλοκη γνώση των πραγμάτων, τις ομάδες των πραγμάτων και τις σχέσεις μεταξύ τους (Seeley, 2003). Αναλυτικότερα, ο σχεδιασμός της γλώσσας αυτής είναι απόλυτα συνυφασμένος με την ανάγκη επεξεργασίας του περιεχομένου των πληροφοριών αντί της απλής παρουσίασής τους.

Η OWL υπερβαίνει τις γνωστές και ευρέως διαδεδομένες γλώσσες, όπως η XML, XMLS (XML Schema), RDF, RDFS (RDF Schema) στην υποστήριξη μεγαλύτερης αποσαφήνισης του περιεχομένου του Ιστού από τις μηχανές, παρέχοντας επιπλέον λεξιλόγιο μαζί με μία επίσημη σημασιολογία. Πιο ειδικά, προσθέτει λεξιλόγιο για την περιγραφή ιδιοτήτων και κλάσεων (π.χ. cardinality, disjointness κ.α.).

Το πρότυπο της OWL καθορίζει ουσιαστικά τρεις υπογλώσσες αυξανόμενης εκφραστικής δυνατότητας, οι οποίες η μία είναι υποσύνολο της άλλης. Οι γλώσσες αυτές είναι οι ακόλουθες (Στοϊλος, χ.χ., σ. 13):

*OWL Lite:* Η γλώσσα αυτή απευθύνεται σε χρήστες, οι οποίοι επιθυμούν να χρησιμοποιήσουν την OWL για την περιγραφή γνώσης σε εφαρμογές που δεν έχουν μεγάλες απαιτήσεις σε εκφραστικές δυνατότητες. Έτσι δίνεται η δυνατότητα ανάπτυξης εξειδικευμένων εργαλείων και μηχανισμών εξαγωγής συμπερασμάτων τα οποία αναμένεται να λειτουργούν ταχύτερα από εργαλεία τα οποία υλοποιούν περισσότερο εκφραστικές γλώσσες.

*OWL DL:* Η γλώσσα αυτή δίνει τη μέγιστη εκφραστική δυνατότητα που προσφέρεται από τη γλώσσα OWL χωρίς όμως να χάνονται οι καλές υπολογιστικές ιδιότητές της. Αυτό σημαίνει ότι η γλώσσα αυτή, σε αντίθεση με την τελευταία υπογλώσσα της OWL, είναι αποφασίσιμη (decidable).

*OWL Full:* Η γλώσσα αυτή προσφέρει το ίδιο λεξιλόγιο με τη γλώσσα OWL DL. Επιπρόσθετα όμως παρέχει τη συντακτική ελευθερία και τα χαρακτηριστικά της γλώσσας RDF και πιο συγκεκριμένα τη δυνατότητα μετα- μοντελοποίησης. Η γλώσσα αυτή είναι εμφανώς μη-αποφασίσιμη (undecidable) (Boris, 2005).

Είναι αναγκαίο να αναφερθεί πως η OWL επεκτάθηκε στην OWL 2 διότι αναδύθηκαν τρία βασικά προβλήματα (Patel – Schneider, 2012):

1. Οι περισσότερες εφαρμογές χρησιμοποιούσαν μεγάλες οντολογίες με έναν τεράστιο αριθμό κλάσεων και συσχετίσεων κι αυτό είχε ως αποτέλεσμα να προκαλούνται προβλήματα εξελξιμότητας αλλά και απόδοσης συλλογισμών.
2. Η έλλειψη διαλειτουργικότητας με τεχνολογίες και εργαλεία βάσεων δεδομένων καθώς οι υπογλώσσες της OWL χρησιμοποιούσαν ελαφριές οντολογίες συνδυαστικά με άλλες γλώσσες όπως SQL για τη βελτιστοποίηση των ερωτήσεων σε πολύ μεγάλα σύνολα αποθηκευμένων δεδομένων.
3. Η έλλειψη διαλειτουργικότητας με τις γλώσσες και τις μηχανές κανόνων.

Η OWL 2 επεκτάθηκε σε τρεις υπογλώσσες, την EL, QL και RL. Πιο ειδικά:

*OWL 2 EL:* Η γλώσσα αυτή είναι ιδιαίτερα κατάλληλη για εφαρμογές όπου απαιτούνται πολύ μεγάλες οντολογίες και μεγάλη εκφραστική δύναμη για εγγύηση στη απόδοση λόγω του μεγάλου αριθμού κλάσεων και ιδιοτήτων.

*OWL 2 QL:* Η εν λόγω γλώσσα είναι ιδιαίτερα κατάλληλη για εφαρμογές που διαχειρίζονται πολύ μεγάλο όγκο δεδομένων και η απάντηση ερωτημάτων είναι η πιο συχνή εργασία.

*OWL 2 RL* Η γλώσσα αυτή είναι ιδιαίτερα κατάλληλη για εφαρμογές που απαιτούν η συλλογιστική να έχει καλές ιδιότητες κλιμάκωσης χωρίς να θυσιάζεται η εκφραστική δύναμη.

#### ▪ **Επερωτήσεις & Κανόνες (Query & Rules)**

Για την παροχή κανόνων πέρα από τις δομές που διατίθενται από τις γλώσσες του Σημαιολογικού Ιστού, δομήθηκαν γλώσσες κανόνων. Επίσης, για την ανάκτηση των RDF δεδομένων, καθώς και της RDFS και της OWL δομήθηκαν πρότυπα επερωτήσεων.

Το τρέχον πρότυπο επερωτήσεων για το Σημαιολογικό Ιστό είναι το πρωτόκολλο της SPARQL, η οποία είναι μία ώριμη, πλούσια σε χαρακτηριστικά γλώσσα επερωτήσεων για το περιεχόμενο της RDF και το αντίστοιχο πρότυπο για τους κανόνες σχετικά με το Σημαιολογικό Ιστό είναι το Rule Interchange Format (RIF) και η SWRL (Semantic Web Rule Language) (βλ. εικ. 3), που αποθηκεύει την εκφραστικότητα των διαφόρων υφιστάμενων γλωσσών που βασίζονται σε κανόνες.

#### ○ **SPARQL (Simple Protocol and RDF Query Language)**

Τι προσφέρει όμως η SPARQL; Το πρωτόκολλο SPARQL (Simple Protocol and RDF Query Language) καθορίζει το συντακτικό αλλά και τη σημασιολογία της γλώσσας επερωτήσεων ενός RDF γράφου και μπορεί να ανακτήσει αλλά ταυτόχρονα και να επεξεργαστεί δεδομένα που είναι συμβατά με το μοντέλο της RDF (Prud'hommeaux & Seaborne, 2008). Η SPARQL είναι μία τεχνολογία του Σημαιολογικού Ιστού που αξιοποιήθηκε για την υλοποίηση της προτεινόμενης πρακτικής που παρουσιάζεται στην εν λόγω διπλωματική και θα αναλυθεί στα παρακάτω κεφάλαια.



- **SPARQL Endpoint (Τελικά Σημεία SPARQL)**

Τα ερωτήματα σε SPARQL υλοποιούνται στα τελικά σημεία (Endpoints), που δεν είναι άλλο από μία υπηρεσία διαδικτύου που υποστηρίζει απομακρυσμένα ερωτήματα σε Διασυνδεδεμένα Δεδομένα (Acosta, κ. α., 2011). Τα τελικά σημεία (Endpoints) μπορούν να αναζητηθούν χρησιμοποιώντας οποιοδήποτε εργαλείο που υποστηρίζει το πρωτόκολλο HTTP ή μέσω μιας HTML φόρμας που υποστηρίζει την εισαγωγή κειμένου (Maali, 2014). Μερικά γνωστά εργαλεία που προσφέρουν τελικά σημεία (Endpoints) αναζήτησης σε SPARQL είναι ο Virtuoso, το Twinkle, το ViziQuer, κ.α.

- **Λογική (Unifying Logic)**

Το στρώμα της Λογικής (βλ. εικ. 3) παρέχει το υπόβαθρο για τη δυνατότητα αυτοματοποιημένου συλλογισμού και συμπερασμών βάσει των πληροφοριών που δομούνται σε μία οντολογία. Ακόμα, το επίπεδο αυτό καθιστά δυνατή τη χρήση τυπικών κανόνων, βάσει των οποίων γίνεται εφικτή η ψευδο-νοήμων διαδικασία λήψης αποφάσεων από τις υπολογιστικές μηχανές.

Μία Μηχανή Συμπερασμού (reasoner) ή αλλιώς μία μηχανή που εξαγεί σημασιολογικά συμπεράσματα, θα μπορούσε να χρησιμοποιηθεί για να ελέγξει και να επιλύσει τα προβλήματα της συνέπειας και της συνάφειας των εννοιών (Bassiliades, 2005).

- **Απόδειξη (Proof)**

Αυτό το στρώμα (βλ. εικ. 3) έχει ως στόχο να εξηγήσει τις απαντήσεις που λαμβάνονται από τους αυτοματοποιημένους πράκτορες που αξιοποιούν τις παρεχόμενες πληροφορίες (Sizon, 2007).

- **Εμπιστοσύνη (Trust)**

Η εμπιστοσύνη (βλ. εικ. 3) είναι το τελικό στρώμα του Σημασιολογικού Ιστού κι αφορά την αξιοπιστία των πληροφοριών στο Διαδίκτυο με σκοπό την παροχή εξασφάλισης της ποιότητας τους (Taye, 2010, σ. 184).

Συνοπτικά, το Unicode και το URI που βρίσκονται στο κατώτερο στρώμα βεβαιώνουν ότι χρησιμοποιείται το διεθνές σύνολο χαρακτήρων και παρέχονται τα μέσα για τον εντοπισμό των πόρων στο Σημασιολογικό Ιστό αντίστοιχα. Το στρώμα της XML με τους ορισμούς των χώρων ονομάτων (Namespaces) και του σχήματος (Schema) διαβεβαιώνουν ότι είναι εφικτό να ενσωματωθούν οι ορισμοί του Σημασιολογικού Ιστού με τα άλλα πρότυπα που βασίζονται στην XML. Με την RDF και το RDF Schema είναι δυνατόν να πραγματοποιηθούν δηλώσεις σχετικά με τους πόρους που έχουν URI και να καθοριστούν λεξιλόγια στα οποία μπορεί να γίνει αναφορά μέσω URIs. Αυτό είναι το στρώμα όπου μπορούν να δοθούν τύποι στους πόρους και τις συνδέσεις. Το στρώμα της Οντολογίας υποστηρίζει την εξέλιξη των λεξιλογίων, καθώς μπορεί να καθορίσει τις σχέσεις μεταξύ των διαφόρων εννοιών. Το στρώμα της Λογικής επιτρέπει τη συγγραφή των κανόνων, ενώ το στρώμα της Απόδειξης εκτελεί τους κανόνες και αξιολογεί μαζί με το στρώμα της Εμπιστοσύνη εάν πρέπει να εμπιστευθεί κανείς την απόδειξη ή όχι (Koivunen & Miller, 2001).

## 2.4. Τεχνολογίες Σημασιολογικού Ιστού

Οι κύριες τεχνολογίες του σημασιολογικού ιστού που εξυπηρετούν το όραμα του και μπορούν να χρησιμοποιηθούν και στον τομέα της εκπαίδευσης και των βιβλιοθηκών, ώστε να παραχθούν εφαρμογές σημασιολογικού ιστού με εκπαιδευτική χρησιμότητα (Basiliades, 2005). Παρακάτω παρουσιάζονται οι τεχνολογίες του Σημασιολογικού Ιστού που αξιοποιήθηκαν για την υλοποίηση της προτεινόμενης πρακτικής που παρουσιάζεται στην εν λόγω διπλωματική.

### 2.4.1. URI s

Ο Ιστός είναι ένας χώρος πληροφοριών που μπορεί να χρησιμοποιηθεί τόσο από τον άνθρωπο, όσο κι από τις μηχανές. Επομένως, και οι δύο θα πρέπει να είναι σε θέση να ανακτήσουν αναπαραστάσεις των πόρων σε μια κατάλληλη μορφή γι' αυτούς. Για παράδειγμα, HTML μορφή για τον άνθρωπο και RDF μορφή για τις μηχανές. Αυτό υλοποιείται χρησιμοποιώντας ένα μηχανισμό που ονομάζεται HTTP διαπραγμάτευση περιεχομένου (content negotiation).

Σύμφωνα με τις αρχές των Διασυνδεδεμένων Δεδομένων (βλ. υποκεφ. 2.4.) τα HTTP URIs πρέπει να είναι dereferenceable, δηλαδή οι HTTP clients να είναι σε θέση κοιτώντας το URI, χρησιμοποιώντας το πρωτόκολλο HTTP, να ανακτήσουν μια περιγραφή του προσδιοριζόμενου πόρου.

Για να πραγματοποιηθούν τα παραπάνω έχουν υιοθετηθεί οι εξής δύο στρατηγικές για τα URIs στο Σημασιολογικό Ιστό:

- 303 URIs: χρησιμοποιούνται για να περιγράψουν πόρους που αποτελούν μέρος των πολύ μεγάλων συνόλων δεδομένων, όπως είναι η περιγραφή ενός ατόμου από την Dbpedia. Τα URI στον Σημασιολογικό Ιστό δεν προσδιορίζουν μόνο έγγραφα του Ιστού αλλά και αντικείμενα του πραγματικού κόσμου (ανθρώπους, σπίτια κ.α.), η προσέγγιση αυτή χρησιμοποιεί έναν κωδικό HTTP, τον 303 See Other, ώστε ο εκάστοτε πελάτης (άνθρωπος ή μηχανή) να αντιληφθεί ότι δεν πρόκειται για ένα συνηθισμένο έγγραφο του Ιστού. Η στρατηγική αυτή ανακατευθύνει τον πελάτη σε έγγραφα Ιστού που περιέχουν πληροφορίες για το αντικείμενο αυτό σε άλλη διεύθυνση, έτσι ώστε να επιτευχθεί σαφής διαχωρισμός μεταξύ URI που προσδιορίζουν αντικείμενα του πραγματικού κόσμου και URI που προσδιορίζουν έγγραφα Ιστού (Health & Bizer, 2011). Η μορφή των 303 URIs είναι η εξής:

```
http://www.example.com/id/exampleinc
Example Inc., the company
http://www.example.com/id/bob
Bob, the person
http://www.example.com/id/alice
Alice, the person
```

**Εικόνα 4: Παράδειγμα 303 URIs**

- Hash URIs: χρησιμοποιούνται για να περιγράψουν πόρους που ανήκουν σε μικρά σύνολα δεδομένων. Η στρατηγική Hash URI βασίζεται στο χαρακτηριστικό ότι τα URIs μπορεί να περιέχουν ένα ειδικό τμήμα που διαχωρίζεται από το τμήμα βάσης του URI με ένα σύμβολο δέσης (#). Αυτό το ειδικό τμήμα ονομάζεται αναγνωριστικό κομμάτι. Πιο ειδικά, όταν ένας πελάτης θέλει να ανακτήσει ένα Hash URI, το πρωτόκολλο HTTP αποκόβει το τμήμα που βρίσκεται αμέσως μετά το «#» και επιστρέφει ως απάντηση το υπόλοιπο τμήμα, το οποίο είναι συνήθως ένα στατικό αρχείο RDF (Health & Bizer, 2011). Η μορφή των Hash URIs είναι η εξής:

```
http://www.example.com/about#exampleinc
Example Inc., the company
http://www.example.com/about#bob
Bob, the person
http://www.example.com/about#alice
Alice, the person
```

**Εικόνα 5: Παράδειγμα Hash URIs**

Και οι δύο στρατηγικές διασφαλίζουν ότι τα αντικείμενα και τα έγγραφα που περιγράφονται δεν είναι συγκεχυμένα κι ότι οι άνθρωποι και τα μηχανήματα μπορούν να τα ανακτήσουν στην κατάλληλη μορφή.

## 2.4.2. Συντάξεις RDF

Για την ανταλλαγή δεδομένων μεταξύ των εφαρμογών στο Σημασιολογικό Ιστό, πέρα από την XML έχουν προταθεί και κάποιες άλλες μορφοποιήσεις του μοντέλου RDF. Η προτεινόμενη μορφή από το W3C είναι το βασισμένο στην XML συντακτικό της γλώσσας RDF/XML. Ωστόσο, η μορφή αυτή δεν είναι φιλική προς τον άνθρωπο κι έτσι δημιουργήθηκε η ανάγκη ύπαρξης άλλων μορφών σύνταξης. Βάσει του παραπάνω δημιουργήθηκαν επιπλέον συντάξεις RDF, όπως η Terse RDF Triple Language (Turtle), η Notation-3 (n3) και η JavaScript Object Notation (JSON), οι οποίες είναι συντάξεις που αξιοποιήθηκαν στα πλαίσια της εν λόγω διπλωματικής και βοήθησαν στην επιτυχή ολοκλήρωση της.

- Notation-3<sup>1</sup> (.n3)

Η μορφή αρχείου Notation3 ή αλλιώς N3, όπως είναι ευρύτερα γνωστή, είναι μια συντομογραφία των μη XML αρχείων των μοντέλων δεδομένων της RDF. Η μορφή αυτή είναι σχεδιασμένη βάσει της δυνατότητας ανάγνωσης από τους ανθρώπους (Berners - Lee, κ.α., 2007). Η N3 επεκτείνει το RDF μοντέλο δεδομένων με την προσθήκη τύπων, μεταβλητών, λογικής συνεπαγωγής και λειτουργικών κατηγορημάτων (Berners – Lee & Connolly, 2011).

Παρακάτω ακολουθεί ένα παράδειγμα της μορφής .n3:

```
@prefix library: <http://83.212.169.101/resource/> .
@prefix rdfs: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfrs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix bibo: <http://url.org/ontology/bibo/> .
@prefix rdvocab: <http://RDVocab.info/elements/> .
@prefix dcterm: <http://url.org/dc/terms/> .
@prefix dc: <http://url.org/dc/elements/1.1/> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix schema: <http://schema.org/> .
@prefix gnd: <http://d-nb.info/standards/elementset/gnd#> .
@prefix gn: <http://www.geonames.org/ontology#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
library:record9294 rdfs:type bibo:Book .
library:record9294 bibo:record "9294" .
library:record9294 bibo:isbn "960-225-177-8" .
library:record9294 dcterm:created "2003-01-27"^^xsd:date .
library:record9294 bibo:title "ΚΑΠΟΙΑ ΚΑΛΙΚΑΝΤΖΑΡΟ ΧΡΙΣΤΟΥΓΕΝΝΑ" .
library:record9294 bibo:authorList "ΚΑΤΕΡΙΝΑ ΤΖΑΜΤΖΗ" .
library:record9294 rdfs:label "ΚΑΠΟΙΑ ΚΑΛΙΚΑΝΤΖΑΡΟ ΧΡΙΣΤΟΥΓΕΝΝΑ-ΚΑΤΕΡΙΝΑ ΤΖΑΜΤΖΗ-ΚΑΤΕΡΙΝΑ ΤΖΑΜΤΖΗ-ΘΕΟΔΩΡΑ ΜΟΛΕ" .
library:place3 rdfs:type schema:Place .
library:place3 rdfs:label "ΑΘΗΝΑ" .
library:record9294 rdvocab:placeOfPublication library:place3 .
library:publisher1173 rdfs:type foaf:Organization .
library:publisher1173 rdfs:label "ΡΩΣΣΗ" .
library:record9294 dcterm:publisher library:publisher1173 .
library:record9294 dcterm:issued "1999"^^xsd:gYear .
library:record9294 dcterm:extent "20σ." .
```

**Εικόνα 6: Παράδειγμα σε μορφή αρχείου N3**

- Comma-separated values<sup>2</sup> (.csv)

Η μορφή αρχείου CSV (Comma-separated values) χρησιμοποιείται συχνά για την ανταλλαγή δεδομένων μεταξύ των διαφορετικών εφαρμογών. Το csv είναι τα αρχικά των λέξεων Comma Separated Values που υποδηλώνει στην ουσία μία σειρά μεταβλητών με τιμές μέσα τους, χωρισμένες απλά με ένα κόμμα (Edoceo: Comma Separated Values (CSV) Standard File, χ.χ.).

Τα CSV αρχεία σώζονται ως κανονικό κείμενο σε μορφή κωδικοποίησης ANSI, διαχωρίζοντας τα δεδομένα των πεδίων απλά με ένα κόμμα και μπορεί κανείς να τα επεξεργαστεί ακόμα και με έναν απλό editor κειμένου. Η μορφή του αρχείου CSV είναι επεξεργάσιμη από το KSpread, το OpenOffice Calc και το Microsoft Excel. Πολλές άλλες εφαρμογές υποστηρίζουν τη μορφή αρχείου CSV με κάποιον από τους εξής τρόπους, είτε

<sup>1</sup> <http://www.w3.org/DesignIssues/Notation3>

<sup>2</sup> <http://edoceo.com/utilitas/csv-file-format>

να εισάγουν ή να εξάγουν τα δεδομένα τους σε αυτή τη μορφή (Edoceo: Comma Separated Values (CSV) Standard File, χ.χ.).

Παρακάτω παρατίθεται ένα παράδειγμα σε μορφή CSV:

```
Title, Author, ISBN13, Pages
1984, George Orwell, 978-0451524935, 268
Animal Farm, George Orwell, 978-0451526342, 144
Brave New World, Aldous Huxley, 978-0060929879, 288
Fahrenheit 451, Ray Bradbury, 978-0345342966, 208
Jane Eyre, Charlotte Brontë, 978-0142437209, 532
Wuthering Heights, Emily Brontë, 978-0141439556, 416
Agnes Grey, Anne Brontë, 978-1593083236, 256
Walden, Henry David Thoreau, 978-1420922615, 156
Walden Two, B. F. Skinner, 978-0872207783, 301
"Eats, Shoots & Leaves", Lynne Truss, 978-1592400874, 209
```

**Εικόνα 7: Παράδειγμα σε μορφή αρχείου CSV**

- Turtle<sup>3</sup>

Η μορφή αρχείου Turtle (.ttl) είναι μια σύνταξη κειμένου για το μοντέλο δεδομένων της RDF που επιτρέπει σε έναν γράφο RDF να αποδοθεί σε μία πλήρη και συμπαγή μορφή κειμένου, με συντομογραφίες για κοινά πρότυπα χρήσης και τύπους δεδομένων. Η Turtle είναι συμβατή με τη μορφή N-Triples, καθώς την SPARQL (Beckett, Berners-Lee, Eric Prud'hommeaux & Carothers, 2014).

Ακολουθεί ένα παράδειγμα της μορφής αρχείου Turtle (.ttl):

```
@prefix xs: <http://www.w3.org/2001/XMLSchema#> .
<http://dbpedia.org/resource/Abraham_Lincoln> <http://dbpedia.org/ontology/birthDate> "1809-02-12"^^xs:date ;
<http://dbpedia.org/ontology/deathDate> "1865-04-15"^^xs:date ;
<http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://xmlns.com/foaf/0.1/Person> ;
<http://xmlns.com/foaf/0.1/name> "Abraham Lincoln"@en ;
<http://dbpedia.org/ontology/birthPlace> <http://dbpedia.org/resource/Hardin_County,_Kentucky> ;
<http://purl.org/dc/elements/1.1/description> "16th President of the United States"@en .
<http://dbpedia.org/resource/Alexander_Mackenzie> <http://dbpedia.org/ontology/birthDate> "1822-01-28"^^xs:date ;
<http://dbpedia.org/ontology/deathDate> "1892-04-17"^^xs:date ;
<http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://xmlns.com/foaf/0.1/Person> ;
<http://xmlns.com/foaf/0.1/surname> "Mackenzie"@en ;
<http://dbpedia.org/ontology/deathPlace> <http://dbpedia.org/resource/Toronto> ;
<http://xmlns.com/foaf/0.1/name> "Alexander Mackenzie"@en ;
<http://dbpedia.org/ontology/birthPlace> <http://dbpedia.org/resource/Logierait> ,
<http://dbpedia.org/resource/Perthshire> ,
<http://dbpedia.org/resource/Scotland> ;
<http://xmlns.com/foaf/0.1/givenName> "Alexander"@en ;
<http://purl.org/dc/elements/1.1/description> "2nd Prime Minister of Canada (1873-1878)"@en .
<http://dbpedia.org/resource/Alan_Turing> <http://dbpedia.org/ontology/birthDate> "1912-06-23"^^xs:date ;
<http://dbpedia.org/ontology/deathDate> "1954-06-07"^^xs:date ;
<http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://xmlns.com/foaf/0.1/Person> ;
<http://xmlns.com/foaf/0.1/surname> "Turing"@en ;
<http://dbpedia.org/ontology/deathPlace> <http://dbpedia.org/resource/Manchester> ;
<http://xmlns.com/foaf/0.1/name> "Alan Mathison Turing"@en ;
<http://dbpedia.org/ontology/birthPlace> <http://dbpedia.org/resource/Paddington> ;
<http://xmlns.com/foaf/0.1/givenName> "Alan Mathison"@en ;
<http://purl.org/dc/elements/1.1/description> "Computer scientist, mathematician, and cryptographer"@en .
```

**Εικόνα 8: Παράδειγμα μορφής αρχείου Turtle**

- JavaScript Object Notation<sup>4</sup> (.json)

<sup>3</sup> <http://www.w3.org/TR/turtle/>

Η μορφή JSON (JavaScript Object Notation) είναι μία ελαφριά μορφή ανταλλαγής δεδομένων. Είναι εύκολη για τους ανθρώπους να διαβάζουν και να γράφουν και αντίστοιχα για τις μηχανές να την αναλύσουν και να τη δημιουργήσουν. Όσον αφορά τη δομή της, βασίζεται σε ένα υποσύνολο της γλώσσας προγραμματισμού JavaScript. Η JSON είναι μια μορφή κειμένου που είναι εντελώς ανεξάρτητη γλώσσα, αλλά χρησιμοποιεί συμβάσεις που είναι γνωστές στους προγραμματιστές από την οικογένεια των γλωσσών της C, συμπεριλαμβανομένων της C, C++, C#, Java, JavaScript, Perl, Python αλλά και πολλές άλλες (Introducing JSON, χ.χ.).

Η JSON έχει δύο δομές:

1. Μια συλλογή από ζεύγη ονόματος / τιμής. Σε διάφορες γλώσσες, αυτή αντιστοιχεί σ' ένα αντικείμενο, μία εγγραφή, μία δομή, ένα λεξικό, έναν πίνακα ή μία λίστα.
2. Μια ταξινομημένη λίστα τιμών. Στις περισσότερες γλώσσες, αυτή αντιστοιχεί σε μια σειρά, διάνυσμα, λίστα, ή συχνότητα.

Ακολουθεί ένα παράδειγμα στη μορφή JSON:

```
{
  "glossary": {
    "title": "example glossary",
    "GlossDiv": {
      "title": "S",
      "GlossList": {
        "GlossEntry": {
          "ID": "SGML",
          "SortAs": "SGML",
          "GlossTerm": "Standard Generalized Markup Language",
          "Acronym": "SGML",
          "Abbrev": "ISO 8879:1986",
          "GlossDef": {
            "para": "A meta-markup language, used to create markup languages such as DocBook.",
            "GlossSeeAlso": ["GML", "XML"]
          },
          "GlossSee": "markup"
        }
      }
    }
  }
}
```

Εικόνα 9: Παράδειγμα σε μορφή αρχείου JSON

### 2.4.3. Οντολογίες (Ontologies)

Η λέξη «οντολογία» χρησιμοποιείται με διαφορετικές σημασίες σε διαφορετικές επιστημονικές κοινότητες. Κατά τη χρήση του όρου πραγματοποιείται διάκριση μεταξύ του μη μετρήσιμου ουσιαστικού, δηλαδή του ουσιαστικού που αναφέρεται σε μια φιλοσοφική έννοια, δηλαδή τον κλάδο της φιλοσοφίας που ασχολείται με τη φύση και τη δομή της "πραγματικότητας" (Basiliades, 2005) και του μετρήσιμου ουσιαστικού δηλαδή αυτού που αναφέρεται στην επιστήμη των Υπολογιστών.

---

<sup>4</sup> <http://json.org/>

Πιο ειδικά, στην Επιστήμη των Υπολογιστών, μια οντολογία αντικατοπτρίζει τη περιγραφή της σημασιολογίας των δεδομένων, παρέχοντας έναν τρόπο μεταγωγής δεδομένων μεταξύ των διάφορων μερών, ώστε να μπορεί να καταλάβει το ένα - το νήμα των δεδομένων που του στέλνει το άλλο με ρητό και σαφή τρόπο (Tayе, 2010, σ. 184).

Οι οντολογίες παρέχουν μια κοινή κατανόηση ενός γνωστικού πεδίου. Μια τέτοια προσέγγιση μπορεί να καταργήσει τις διαφορές στην ορολογία ανάμεσα στους φορείς στο ίδιο γνωστικό πεδίο. Επίσης μπορεί να λύσει το πρόβλημα που υπάρχει όταν δύο φορείς έχουν την ίδια ορολογία για διαφορετικές έννοιες. Οι σημαντικότερες γλώσσες οντολογιών για τον Ιστό είναι οι παρακάτω (Antonίου & Harmelen, 2008):

- Η XML παρέχει το βασικό συντακτικό για δομημένα έγγραφα αλλά δεν βάζει σημασιολογικούς περιορισμούς στους όρους του εγγράφου.
- Η XML Schema είναι μια γλώσσα, η οποία επιβάλλει περιορισμούς στα έγγραφα XML.
- Το RDF είναι ένα μοντέλο δεδομένων για πόρους και σχέσεις μεταξύ των πόρων αυτών, που παρέχει βασική σημασιολογική λειτουργία και το μοντέλο αναπαριστάται σε σύνταξη XML.

Τα τελευταία χρόνια έχει πραγματοποιηθεί μία προσπάθεια να αναπτυχθούν οντολογίες που μπορούν να περιγράψουν με σαφήνεια τις βιβλιογραφικές εγγραφές, ωστόσο, προς το παρόν δεν έχει αναπτυχθεί μία καθολική οντολογία γι' αυτόν τον σκοπό. Έτσι, οι βιβλιοθήκες χρησιμοποιούν ένα σύνολο οντολογιών ώστε να καλυφθούν οι ανάγκες σημασιολογικής αναπαράστασης των βιβλιογραφικών δεδομένων. Πιο ειδικά, για την περιγραφή βιβλίων κι άλλων σχετικών πόρων χρησιμοποιείται η BIBO, για την περιγραφή προσώπων, η FOAF και για την περιγραφή θεματικών επικεφαλίδων η SKOS (Καραμπατάκης, 2013).

Πιο αναλυτικά:

- ***BIBO (Bibliographic Ontology)***

Η BIBO παρέχει βασικές έννοιες και ιδιότητες για την περιγραφή βιβλιογραφικών εγγραφών στο Σημασιολογικό Ιστό σε RDF. Αυτή η οντολογία μπορεί να χρησιμοποιηθεί ως μια οντολογία παραπομπών, ως οντολογία κατάταξης εγγράφων, ή απλά ως ένας τρόπος περιγραφής κάθε είδους εγγράφων RDF (D'Arcus & Giasson, 2009).

- ***FOAF (Friend of a Friend)***

Η οντολογία FOAF περιγράφει ανθρώπους, τους δεσμούς μεταξύ τους και τα πράγματα που έχουν κάνει, τις εργασίες του κι ό,τι έχουν δημιουργήσει (Brickley & Miller, χ.χ.). Η οντολογία αυτή περιγράφει δύο τομείς των ψηφιακών δεδομένων: τα βιογραφικά δεδομένα και τα δεδομένα των κοινωνικών δικτύων (Al-Mukhtar & Al-Assafy, 2014).

- ***SKOS (Simple Knowledge Organization System)***

Το SKOS αναπτύχθηκε για την υποστήριξη της χρήσης των συστημάτων οργάνωσης γνώσης, όπως θησαυροί, σχήματα ταξινόμησης, συστήματα θεματικών επικεφαλίδων και

ταξονομίες στα πλαίσια του Σημασιολογικού Ιστού (SKOS Simple Knowledge Organization System - Home Page, χ.χ.).

### *2.4.4. Αποθετήρια τριπλετών (triplestores)*

Ένα αποθετήριο τριπλετών (triplestore) ή αλλιώς ένα RDF κατάστημα είναι μια βάση δεδομένων ειδικά διαμορφωμένη για την αποθήκευση και ανάκτηση τριπλετών μέσω σημασιολογικών ερωτημάτων. Τα αποθετήρια τριπλετών παρουσιάζουν κάποια συγκεκριμένα κοινά λειτουργικά χαρακτηριστικά αφού χρησιμοποιούνται για τους εξής λόγους:

1. Δημιουργία βάσης δεδομένων
2. Εισαγωγή δεδομένων
3. Επερωτήματα
4. Διακομιστή SPARQL
5. Backup / Restore
6. Βιβλιοθήκες πελατών

Τα αποθετήρια τριπλετών παρουσιάζουν τα εξής κοινά πλεονεκτήματα κατά την χρήση τους:

- ✓ Προσφέρουν καλύτερη απόδοση ερωτημάτων σε SPARQL απ' ό,τι οι σχεσιακές βάσεις δεδομένων στις οποίες πραγματοποιούνται ερωτήματα στη γλώσσα ερωτημάτων της SQL.
- ✓ Προσφέρουν επιπλέον διεπαφές καθώς τα αποθετήρια τριπλετών υποστηρίζουν τη γλώσσα SPARQL και προσφέρουν τη δυνατότητα και σε άλλες εφαρμογές να πραγματοποιήσουν ερωτήματα SPARQL για την ανάκτηση των δεδομένων που έχουν αποθηκευμένα.
- ✓ Προσφέρουν τη δυνατότητα αποθήκευσης επιπλέον δεδομένων, τα οποία μπορούν να συνδεθούν και να επαναχρησιμοποιηθούν, λειτουργώντας έτσι ως πλατφόρμες για την ενοποίηση δεδομένων και την επαναχρησιμοποίηση οντολογιών.

Για την απόδοση μιας ολοκληρωμένης εικόνας παρακάτω παρουσιάζονται τα πιο γνωστά αποθετήρια τριπλετών καθώς και τα χαρακτηριστικά που κάνουν το καθένα να διαφέρει από τα υπόλοιπα, έτσι όπως συναντήθηκαν στη βιβλιογραφία:

- 4store<sup>5</sup>

Το 4store σχεδιάστηκε από τον Steve Harris και αναπτύχθηκε στο Garlik για τη στήριξη των εφαρμογών του Σημασιολογικού Ιστού. Το 4store τε προορίζεται κυρίως για να

---

<sup>5</sup> <http://4store.org/>



λειτουργεί σε πλατφόρμες Linux, αλλά είναι γνωστό ότι λειτουργεί και σε Mac OS X. Τα κύρια πλεονεκτήματα του είναι η απόδοση, η επεκτασιμότητα και η σταθερότητα που προσφέρει. Ωστόσο, δεν παρέχει πολλές δυνατότητες πέρα από την αποθήκευση RDF και τη δυνατότητα υλοποίησης SPARQL ερωτημάτων.

- Stardog<sup>6</sup>

Το Stardog είναι μια βάση δεδομένων για γράφους σημασιολογίας, η οποία υποστηρίζει το μοντέλο δεδομένων της RDF, τις επερωτήσεις SPARQL, τα HTTP και SNAPL πρωτόκολλα για απομακρυσμένη πρόσβαση και έλεγχο, την οντολογία OWL 2 και την προγραμματιστική αλληλεπίδραση μέσω της υποστήριξης πολλών γλωσσών προγραμματισμού.

- AllegroGraph<sup>7</sup>

Είναι μία σύγχρονη, υψηλής απόδοσης, γραφική βάση δεδομένων. Το AllegroGraph υποστηρίζει SPARQL, RDFS ++ και συλλογισμούς Prolog.

- OpenLink Virtuoso<sup>8</sup>

Ο Virtuoso είναι μια σύγχρονη λύση για την πρόσβαση σε δεδομένα, την ένταξη και διαχείριση σχεσιακών βάσεων δεδομένων (SQL πίνακες ή / και RDF βασιζόμενα σε γράφους). Η μοναδική υβριδική αρχιτεκτονική του Virtuoso παρέχει τη δυνατότητα να προσφερθούν οι εξής διακριτές λειτουργίες:

1. Διαχείριση σχεσιακών πινάκων δεδομένων (Columnar or Column-Store SQL RDBMS)
2. Διαχείριση γράφων δεδομένων (SPARQL RDF based Quad Store)
3. Διαχείριση περιεχομένου (HTML, TEXT, TURTLE, RDF/XML, JSON, JSON-LD, XML)
4. Υπηρεσίες Ιστού και άλλων αρχείων (Web Document or File Server)
5. Ανάπτυξη Ανοιχτών και Διασυνδεδεμένων Δεδομένων 5 αστέρων (RDF-based Linked Data Server)
6. Διακομιστής εφαρμογής Ιστού (SOAP or RESTful interaction modes).

- Jena TDB<sup>9</sup>

Το αποθετήριο τριπλετών Jena TDB έχει αναπτυχθεί από τη Jena για την αποθήκευση και τις επερωτήσεις στη γλώσσα RDF. Το TDB μπορεί να χρησιμοποιηθεί ως αποθήκη RDF υψηλής απόδοσης σε ένα μόνο μηχάνημα. Για την υλοποίηση των SPARQL ερωτημάτων η Jena προσφέρει ένα component, που ονομάζεται Fuseki και μπορεί να χρησιμοποιηθεί σε συνδυασμό με το TDB.

---

<sup>6</sup> <http://stardog.com/>

<sup>7</sup> <http://franz.com/agraph/allegrograph/>

<sup>8</sup> <http://virtuoso.openlinksw.com/>

<sup>9</sup> <https://jena.apache.org/documentation/tdb/>

## 2.5. Ανοιχτά Διασυνδεδεμένα Δεδομένα

Όπως αναφέρθηκε και παραπάνω, ο σημασιολογικός ιστός έρχεται να δώσει λύση στην «έξυπνη» διασύνδεση και την αναζήτηση των δομημένων δεδομένων που μπορούν να προκύψουν είτε από κάποια διεπαφή εφαρμογών, είτε από την πρωτογενή περιγραφή δεδομένων. Τα Διασυνδεδεμένα Δεδομένα (linked data) φέρονται ως η λύση στο πρόβλημα της αναζήτησης και της σύνδεσης των δεδομένων του παγκόσμιου ιστού μέσα από τη δομημένη περιγραφή τους. Αυτή η νέα λύση που ως ιδέα υπάρχει αρκετά χρόνια, αρχίζει να υλοποιείται τα τελευταία χρόνια και παρατηρείται η έντονη εφαρμογή της και στον τομέα της Εκπαίδευσης (Κωνσταντινίδης, κ. συν., 2011).

Τι είναι όμως τα Διασυνδεδεμένα Δεδομένα και ποιες είναι οι προϋποθέσεις για να χαρακτηρίσουμε συγκεκριμένα δεδομένα ως διασυνδεδεμένα; Για να γίνει η μετάβαση από τη σύλληψη της ιδέας του Σημασιολογικού Ιστού στην πραγματοποίησή της απαιτείται να υπάρχει μεγάλος όγκος δεδομένων στον Παγκόσμιο Ιστό σε μορφή καθορισμένη, προσπελάσιμη κι εύκολα διαχειρίσιμη. Επιπλέον, θα πρέπει να είναι γνωστές και καθορισμένες οι σχέσεις αυτών των δεδομένων μεταξύ τους. Τα δεδομένα με τα συγκεκριμένα χαρακτηριστικά εμφανίζονται με τον όρο Διασυνδεδεμένα Δεδομένα (Linked Data) (Berners – Lee, 2006).

Τα Διασυνδεδεμένα Δεδομένα ως όρος ύπαρξης και ανάπτυξης του Σημασιολογικού Ιστού απαιτείται να είναι ενσωματωμένα στον Παγκόσμιο Ιστό ανεξάρτητα από την πηγή από την οποία προέρχονται (Heath & Bizer, 2011) κι αυτό διότι ο στόχος των Διασυνδεδεμένων Δεδομένων είναι να επιτρέψει στους ανθρώπους να μοιράζονται δομημένα δεδομένα στον Ιστό όσο εύκολα μπορούν να μοιράζονται και απλά έγγραφα (Bizer, Cyganiak & Heath, 2008).

Τεχνικά, ο όρος Διασυνδεδεμένα Δεδομένα αναφέρεται σε δεδομένα που δημοσιεύονται στο διαδίκτυο με τέτοιο τρόπο, ώστε να είναι αναγνώσιμα από τις μηχανές, οι έννοιες τους να ορίζονται ρητά βάσει συγκεκριμένων λεξιλογίων και να συνδέονται με άλλα ανοιχτά σύνολα δεδομένων.

Ο εμπνευστής του σημασιολογικού ιστού, Tim Berners – Lee (2006), διατύπωσε τέσσερις βασικές αρχές για τα διασυνδεδεμένα δεδομένα και παρατίθενται παρακάτω:

- Χρήση URIs για την αναγνώριση οντοτήτων, έτσι ώστε να εξασφαλίζεται η μοναδικότητα των αναγνωριστικών
- Χρήση HTTP URIs, έτσι ώστε να είναι σε θέση κανείς να προσπελάσει την πληροφορία
- Στην προσπέλαση ενός URI, να παρέχονται χρήσιμες πληροφορίες, χρησιμοποιώντας τα πρότυπα RDF και SPARQL
- Διασύνδεση των αναγνωριστικών URIs με άλλα που έχουν δημιουργήσει τρίτοι, έτσι ώστε να μπορεί κανείς να λάβει περισσότερα αποτελέσματα κι από συμπληρωματικές πηγές πληροφόρησης

Επιπρόσθετα, για την αξιολόγηση των Διασυνδεδεμένων Δεδομένων και το κατά πόσο συγκεκριμένα δεδομένα μπορούν να χαρακτηριστούν ως διασυνδεδεμένα, ο ίδιος ο Berners – Lee, διατύπωσε ένα σύστημα αξιολόγησης πέντε (5) αστέρων που αναλύεται παρακάτω (Janowicz, κ.α., 2014):

★	<b>Διαθέσιμα στο διαδίκτυο σε οποιαδήποτε μορφή με ανοιχτή άδεια</b>
★★	<b>Διαθέσιμα ως δομημένα δεδομένα που μπορούν να διαβαστούν από τις μηχανές</b>
★★★	<b>Δομημένα χωρίς αποκλειστική μορφή (π.χ. CSV αντί για Excel)</b>
★★★★	<b>Προσδιορισμένα με διευθύνσεις URL, έτσι ώστε οι άνθρωποι και οι μηχανές να έχουν πρόσβαση σε αυτά</b>
★★★★★	<b>Συνδεδεμένα με άλλα δεδομένα για ενίσχυση περιεχομένου</b>

Από την άλλη, συναντάται και ο όρος «Ανοιχτά Διασυνδεδεμένα Δεδομένα». Τα Ανοιχτά Διασυνδεδεμένα Δεδομένα (Open Linked Data) σχετίζονται με τη χρήση του διαδικτύου για τη δημιουργία τυπικών δεσμών μεταξύ δεδομένων από διαφορετικές πηγές. Αυτά μπορεί να είναι τόσο διαφορετικά όσο οι βάσεις δεδομένων που διατηρούνται από δύο οργανισμούς ή συγκεκριμένα ετερογενή συστήματα ενός οργανισμού που, ιστορικά, δεν έχουν εύκολη δυνατότητα διαλειτουργικότητας σε επίπεδο δεδομένων (Bizer, Heath & Berners – Lee, 2009).

Η σχέση μεταξύ των Διασυνδεδεμένων Δεδομένων και των Ανοιχτών δεδομένων είναι σήμερα διφορούμενη. Το JISC Linked Data Horizon Scan αναφέρει πως τα Διασυνδεδεμένα Δεδομένα μπορεί να είναι ανοικτά και τα ανοιχτά δεδομένα μπορούν να συνδέονται μεταξύ τους, αλλά είναι εξίσου πιθανό τα Διασυνδεδεμένα Δεδομένα να έχουν συγκεκριμένη αδειοδότηση ή άλλους περιορισμούς που τα εμποδίζουν να θεωρηθούν ανοιχτά. Τα δεδομένα πρέπει να διατίθενται με τρόπο, ώστε να σέβονται το σύνολο των Κανόνων του Berners-Lee για τη σύνδεση (Miller, 2010).

### *2.5.1. Διασυνδεδεμένα Δεδομένα και Βιβλιοθήκες*

Ο Σημαιολογικός Ιστός και ειδικότερα η σύνδεση των ανοιχτών διασυνδεδεμένων δεδομένων παροτρύνουν τα ιδρύματα να δημοσιεύουν, να μοιράζονται και να συνδέουν με πολλαπλό τρόπο τα δεδομένα τους χρησιμοποιώντας το διαδίκτυο (Hanneman & Kett, 2010).

Οι κύριοι λόγοι για την υιοθέτηση των Ανοιχτών Διασυνδεδεμένων Δεδομένων από τις βιβλιοθήκες, τα αρχεία και τα μουσεία είναι οι εξής (Armstrong, Steward & Ward, 2013):

- Ως ιδρύματα πολιτιστικής κληρονομιάς, οι βιβλιοθήκες, τα αρχεία και τα μουσεία είναι αναγκαίο να ενσωματώνουν στη δραστηριότητα τους, την παροχή πρόσβασης σε πόρους που να διατίθενται στο κοινό.
- Το άνοιγμα των (μετα)δεδομένων επιτρέπει την επαναχρησιμοποίηση τους από τους επαγγελματίες της πληροφόρησης.
- Η διασύνδεση των ανοικτών (μετα)δεδομένων επιτρέπει την απόλυτη σύνδεση των πόρων που δεν περιορίζεται πλέον στο φάσμα των ιδρυμάτων αλλά επεκτείνεται και πέρα από τα στενά όρια αυτών (π.χ. διασύνδεση με άλλα σύνολα δεδομένων από πολλά διαφορετικά ιδρύματα).
- Τα διασυνδεδεμένα (μετα)δεδομένα παρέχουν τα μέσα στις βιβλιοθήκες, τα αρχεία και τα μουσεία να προβάλλουν τους πόρους τους σε ένα πλαίσιο που ξεπερνά την δομή και το θέμα (θεματική κατηγοριοποίηση) των πόρων αυτών καθώς μπορούν να εμπλουτιστούν με επιπλέον πληροφορίες (π.χ. παροχή βιογραφίας ενός συγγραφέα κ.α.).

Τα συνδεδεμένα δεδομένα στις βιβλιοθήκες εστιάζονται στην ανταλλαγή των δεδομένων των βιβλιοθηκών και στη δυναμική που διαθέτουν για τη δημιουργία Διασυνδεδεμένων Δεδομένων Βιβλιοθηκών σε παγκόσμιο επίπεδο. Με άλλα λόγια, τα δεδομένα διαφορετικών βιβλιοθηκών μπορούν να συνδυαστούν ώστε να παρέχουν εμπλουτισμένη πληροφορία στον εκάστοτε ενδιαφερόμενο χρήστη. Ακόμα, εστιάζονται στην ανταλλαγή και την από κοινού αξιοποίηση των δεδομένων πέρα από τα στενά όρια των βιβλιοθηκών, παρέχοντας έτσι πρόσβαση σε ένα μεγάλο πλήθος δεδομένων που προέρχονται από διαφορετικές πηγές. Τα διασυνδεδεμένα δεδομένα στις βιβλιοθήκες εστιάζονται, επίσης, στην μεγαλύτερη εμπιστοσύνη στον αναπτυσσόμενο σημασιολογικό ιστό αλλά και στη διατήρηση ενός παγκόσμιου «γραφήματος» πολιτισμικών πληροφοριών που είναι τόσο αξιόπιστες, όσο και ανθεκτικές (Bauer & Kalternbock, 2012).

Τα προγράμματα που αφορούν τα ανοιχτά συνδεδεμένα δεδομένα και πραγματοποιούνται σε όλο τον κόσμο, επεκτείνουν τον τρόπο που αποκτά κανείς πρόσβαση στην πολιτιστική κληρονομιά. Οι βιβλιοθήκες, τα μουσεία, και τα αρχεία έχουν εφεύρει νέους τρόπους για να εξάγουν τα δεδομένα τους σε τριπλέτες, για να ενσωματώνουν εξωτερικούς συνδέσμους στα σύνολα δεδομένων των συλλογών τους και να αναπτύσσουν νέες διεπαφές για τους χρήστες ώστε να αποκτήσουν εμπειρίες σχετικά με την πολιτιστική κληρονομιά (Edelstein κ. α., 2013). Μια τέτοια προσέγγιση πραγματεύεται και η εν λόγω διπλωματική εργασία, εστιασμένη στα δεδομένα μίας Δημοτικής Βιβλιοθήκης για την εφαρμογή μίας προτεινόμενης πρακτικής από την οπτική της βιβλιοθηκονομικής κοινότητας.

### *2.5.1.1. LOD Cloud και Βιβλιοθήκες*

Το πιο χαρακτηριστικό παράδειγμα υιοθέτησης και εφαρμογής των αρχών των Διασυνδεδεμένων Δεδομένων (Linked Data) είναι το Σύννεφο των Ανοιχτών Διασυνδεδεμένων Δεδομένων (Linked Open Data Cloud). Το Σύννεφο των Ανοιχτών και Διασυνδεδεμένων Δεδομένων αναπτύχθηκε στα πλαίσια του προγράμματος Linked Open Data Project που ξεκίνησε την υλοποίησή του από τον Ιανουάριο του 2007. Η προσπάθεια

αυτή πραγματοποιήθηκε από την ευρύτερη κοινότητα και υποστηρίζεται από το W3C Semantic Web Education and Outreach Group. Ο αρχικός και συνεχής στόχος του έργου αυτού ήταν η αρχή της δημιουργίας του Ιστού των Δεδομένων μέσω του εντοπισμού συνόλων δεδομένων που είναι διαθέσιμα κάτω από ανοικτές άδειες, τα οποία μετατρέπονται σε RDF σύνολα, σύμφωνα πάντα με τις αρχές των Διασυνδεδεμένων Δεδομένων και στη συνέχεια δημοσιεύονται στο Διαδίκτυο (Bizer, Heath & Berners – Lee, 2009).

Οι συμμετέχοντες στα πρώτα στάδια του έργου ήταν κυρίως ερευνητές και προγραμματιστές σε πανεπιστημιακά ερευνητικά εργαστήρια και μικρές επιχειρήσεις. Στη συνέχεια, το έργο αυτό έλαβε μεγάλες διαστάσεις και συμπεριλήφθηκαν σ' αυτό μεγάλοι οργανισμοί όπως το BBC<sup>10</sup>, η Thomson Reuters<sup>11</sup> και η Βιβλιοθήκη του Κογκρέσου<sup>12</sup>. Η αύξηση αυτή των συμμετοχών είναι απόλυτα κατανοητή αν αναλογιστεί κανείς ότι πρόκειται για ένα έργο ανοικτού χαρακτήρα, όπου ο καθένας μπορεί να δημοσιεύσει ένα σύνολο δεδομένων αρκεί να πληρούν τις αρχές των Διασυνδεδεμένων Δεδομένων και να είναι διασυνδεδεμένα με ήδη υπάρχοντα σύνολα δεδομένων (Bizer, Heath & Berners – Lee, 2009).

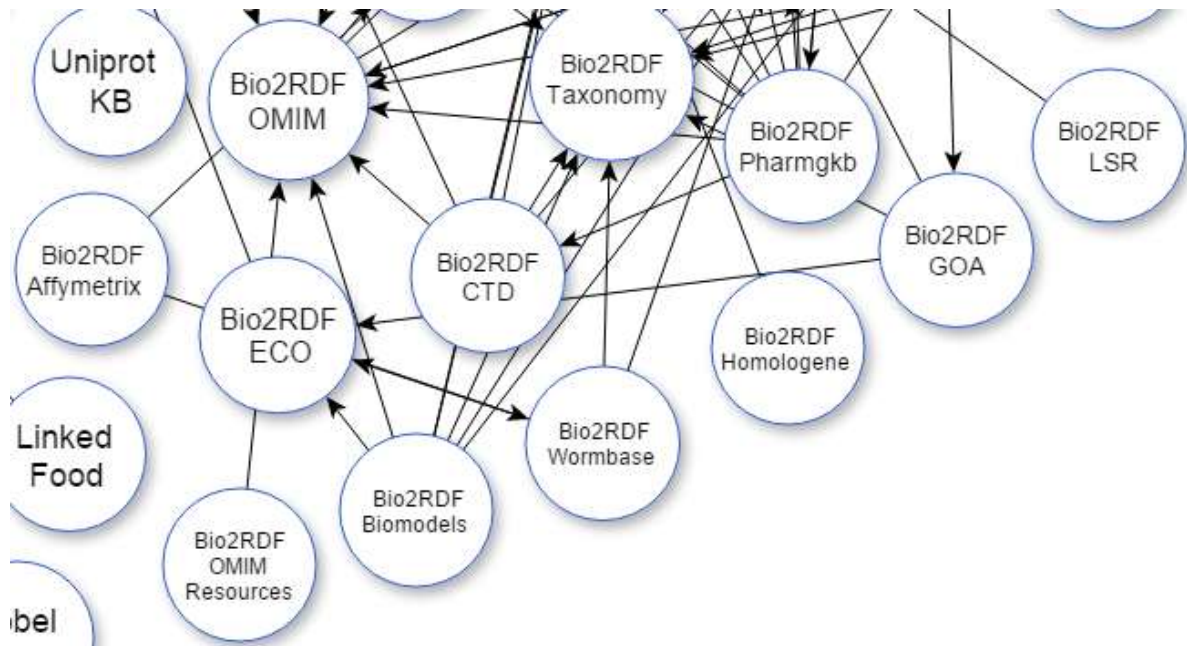
Τα τόξα του Σύννεφου δείχνουν τους δεσμούς μεταξύ των δύο συνδεδεμένων συνόλων δεδομένων. Τα πιο έντονα τόξα αντιστοιχούν σε μεγαλύτερο αριθμό συνδέσεων μεταξύ δύο συνόλων δεδομένων, ενώ τα αμφίδρομα τόξα δείχνουν τις εξωτερικές συνδέσεις από και προς τα δυο σύνολα δεδομένων (βλ. εικ. 10) (Bizer, Heath & Berners – Lee, 2009). Το περιεχόμενο του Σύννεφου είναι διαφορετικό ως προς τη φύση των δεδομένων, τα οποία σχετίζονται με γεωγραφικές περιοχές, ανθρώπους, επιχειρήσεις, βιβλία (Bizer, Cyganiak & Gauss, 2007), επιστημονικές εκδόσεις (Sompel κ. α., 2009), ταινίες (Hassanzadeh, κ. α., 2009), μουσική, τηλεοπτικά και ραδιοφωνικά προγράμματα (Kobilaron κ. α., 2009), γονίδια, πρωτεΐνες, φάρμακα και κλινικές δοκιμές (Belleau κ. α., 2008, Jentzsch κ. α., 2009), διαδικτυακές κοινότητες, στατιστικά δεδομένα, αποτελέσματα απογραφών και σχόλια (Heath & Motta, 2008).

---

<sup>10</sup> <http://www.bbc.com/>

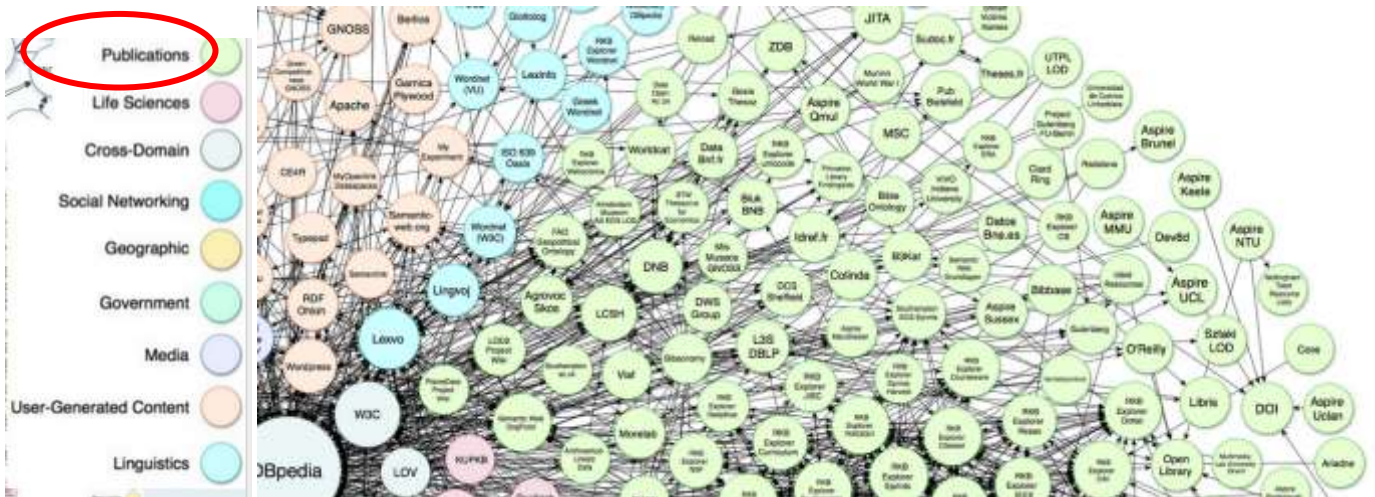
<sup>11</sup> <http://thomsonreuters.com/en.html>

<sup>12</sup> <https://www.loc.gov/>



**Εικόνα 10: Μονόδρομα κι αμφίδρομα τόξα στο LOD Cloud**

Το Σύννεφο των Ανοιχτών Διασυνδεδεμένων Δεδομένων αποτέλεσε ένα σημαντικό εργαλείο αναζήτησης δεδομένων για την υλοποίηση της παρούσας διπλωματικής και πιο ειδικά, τα δεδομένα των βιβλιοθηκών και των φορέων πολιτισμικής κληρονομιάς που εμπεριέχονται στην κατηγορία «Publications» του LOD Cloud προβαλλόμενα με πράσινο χρώμα (βλ. εικ. 11):



**Εικόνα 11: Κατηγορία "Publications" στο LOD Cloud**

Χαρακτηριστικά παραδείγματα βιβλιοθηκών και Φορέων Πολιτισμικής Κληρονομιάς που εμπεριέχονται στο Σύννεφο των Ανοιχτών Διασυνδεδεμένων Δεδομένων είναι τα εξής:

- Open Library<sup>13</sup>
- Library of Congress<sup>14</sup>
- British National Bibliography<sup>15</sup>

<sup>13</sup> <https://openlibrary.org/>

<sup>14</sup> <https://www.loc.gov/>

- Europeana<sup>16</sup> Linked Open Data
- National Library of Scotland<sup>17</sup>
- Biblioteca Nacional de España<sup>18</sup>
- The European Library<sup>19</sup>
- Public Library of Veroia<sup>20</sup>

---

<sup>15</sup> <http://bnb.bl.uk/>

<sup>16</sup> <http://www.europeana.eu/portal/>

<sup>17</sup> <http://www.nls.uk/>

<sup>18</sup> <http://www.bne.es/es/Inicio/index.html>

<sup>19</sup> <http://www.theeuropeanlibrary.org/tel4/>

<sup>20</sup> <http://www.libver.gr/>

# 3. Πρακτικές Ανοιχτών Διασυνδεδεμένων Δεδομένων σε Βιβλιοθήκες

Στο κεφάλαιο αυτό παρουσιάζονται ορισμένες πρακτικές Ανοιχτών και Διασυνδεδεμένων Δεδομένων που προέρχονται από τον χώρο των βιβλιοθηκών τόσο στο εξωτερικό, όσο και στην Ελλάδα, και επηρέασαν την υλοποίηση της προτεινόμενης πρακτικής που υλοποιήθηκε στη διπλωματική αυτή. Η δομή των παρουσιάσεων αυτών είναι πολύ συγκεκριμένη αφού εκτός από την αναφορά στη γενική σύνοψη, εξετάζονται συγκεκριμένα κριτήρια που εξυπηρετούν την εξαγωγή των δυνατών και αδύναμων σημείων των πρακτικών αυτών. Τα κριτήρια αυτά είναι τα εξής:

- Χρήση URIs
- Μετατροπή Δεδομένων σε RDF
- Χρήση οντολογιών
- Πλατφόρμα φιλοξενίας δεδομένων
- Διασύνδεση με άλλα Δεδομένα
- SPARQL Endpoint

Με τον τρόπο αυτόν, πραγματοποιείται μία ανασκόπηση της βιβλιογραφίας, ώστε να ενισχυθεί η επιλογή και η χρήση συγκεκριμένων τεχνολογιών για την υλοποίηση της πρακτικής που πραγματοποιήθηκε στην Δημοτική Βιβλιοθήκη Ευόσμου και παρουσιάζεται ως μία προτεινόμενη πρακτική «ανοίγματος» και «διασύνδεσης» των δεδομένων των βιβλιοθηκών από τους ίδιους τους βιβλιοθηκονόμους.

## 3.1. BNB (British National Bibliography)

Η Βρετανική Βιβλιοθήκη διέθεσε τη Εθνική Βρετανική Βιβλιογραφία ως Ανοιχτά Διασυνδεδεμένα Δεδομένα τον Ιούλιο 2011. Τα κίνητρα της ενέργειας αυτής ήταν αφενός να ανοίξει τα δημόσια δεδομένα για την ευρύτερη επαναχρησιμοποίηση και αφετέρου να δοθεί στα δεδομένα της Βιβλιοθήκης μία μορφή που δεν περιορίζεται στα στενά βιβλιοθηκονομικά πρότυπα αλλά σε πρότυπα που βασίζονται στην XML (Deliot, 2014).

Για την υλοποίηση του έργου αξιοποιήθηκαν 3 εκατομμύρια δεδομένα σε διάφορες γλώσσες, γεγονός που καθιστά το άνοιγμα των δεδομένων αυτών αξιοποιήσιμα σε παγκόσμιο επίπεδο. Το προσωπικό των Υπηρεσιών Μεταδεδομένων που συμμετείχαν στο



έργο είχαν αναπτύξει τεχνογνωσία σε βιβλιογραφικά πρότυπα (κυρίως σε MARC) και είχαν ορισμένες βασικές γνώσεις γύρω από την HTML, XML και XSLT, ωστόσο όμως η ομάδα δεν περιλάμβανε προγραμματιστές. Για την επιτυχή ολοκλήρωση του έργου, η κατάρτιση στην RDF και τις αρχές μοντελοποίησης των δεδομένων πραγματοποιήθηκαν από το Talis (Deliot, 2014). Ωστόσο, η κατάρτιση που προσφέρθηκε από το Talis, δεν θα ήταν απαραίτητη εάν υπήρχε μία πρακτική, όπως αυτή που προτείνεται, η οποία προέρχεται από έναν βιβλιοθηκονόμο που διαθέτει τις βασικές προγραμματιστικές γνώσεις και θα αποτελούσε τον οδηγό υλοποίησης του έργου. Η βήμα – βήμα καθοδήγηση αυτής της πρακτικής θα καθιστούσε την υλοποίηση του έργου μία πιο εύκολη διαδικασία για τους βιβλιοθηκονόμους που επιθυμούν να επιχειρήσουν το άνοιγμα και τη διασύνδεση των δεδομένων της βιβλιοθήκης όπου εργάζονται.

Ωστόσο, είναι θεμιτό ν' αναφερθεί πως το έργο αυτό επηρέασε την υλοποίηση της προτεινόμενης αυτής πρακτικής και υιοθετήθηκαν συγκεκριμένες τεχνολογίες. Τα δεδομένα της Βιβλιοθήκης περιελάμβαναν έννοιες αλλά και υλικά αντικείμενα, για παράδειγμα βιβλιογραφικές πηγές, πρόσωπα, οργανώσεις, χώρους, αντικείμενα, κ.λπ. Για να εντοπιστούν όλοι αυτοί οι πόροι θα έπρεπε να σηματοδοτηθούν με URIs. Αυτό είναι πιο περίπλοκο από ό, τι ακούγεται και περιλαμβάνει μια σειρά από αποφάσεις που πρέπει να ληφθούν από αυτόν ή αυτούς που υλοποιούν το έργο. Στη Βρετανική Βιβλιοθήκη αποφασίστηκε να χρησιμοποιηθούν τα δικά τους URIs για την πλειονότητα των φορέων, αντί να στηριχθούν σε εξωτερικές πηγές (Deliot, 2014). Για παράδειγμα, έχουν δημιουργήσει το δικό τους αναγνωριστικό για τον William Shakespeare, αντί να βασίζεται στην ταυτότητα VIAF (The Virtual International Authority File) που είναι ένα διεθνές αρχείο καθιερωμένων όρων (OCLC: VIAF, 2015).

Το επόμενο βήμα στη διαδικασία υλοποίησης είναι η περιγραφή οντοτήτων και το πώς σχετίζονται μεταξύ τους. Η προσέγγισή που υλοποιήθηκε ήταν η χρήση των κλάσεων και των ιδιοτήτων από τα υπάρχοντα λεξιλόγια όσο το δυνατόν περισσότερο. Για τον λόγο αυτόν, ερευνήθηκαν τα άλλα έργα που σχετίζονται με Βιβλιοθήκες του LOD Cloud και προέκυψε ένας συνδυασμός του Dublin Core, του Bibliographic Ontology, της FOAF, της Event Ontology, κλπ (Deliot, 2014). Άξιο αναφοράς είναι το γεγονός πως δε χρησιμοποιήθηκε η οντολογία FRBR (Functional Requirements for Bibliographic Records) που αφορά εξ ολοκλήρου τις βιβλιογραφικές εγγραφές κι αυτό διότι ο σκοπός του έργου ήταν να δημιουργηθούν Διασυνδεδεμένα Δεδομένα που δε θα περιορίζονται από βιβλιοθηκονομικά πρότυπα. Το παραπάνω αποτέλεσε στόχο και της πρακτικής που αναλύεται στην εν λόγω διπλωματική με τη σκέψη πως τα δεδομένα των βιβλιοθηκών θα ήταν χρήσιμο να χρησιμοποιηθούν από την ευρύτερη κοινότητα για τη δημιουργία γενικών εφαρμογών κι όχι μόνο από τους βιβλιοθηκονόμους και τους επιστήμονες της πληροφόρησης.

Επιπλέον, η μετατροπή του MARC αρχείου σε RDF / XML πραγματοποιήθηκε με το XSLT, που είναι μία γλώσσα που μετατρέπει έγγραφα XML σε άλλα έγγραφα XML (Clark, 1999). Μετά τον έλεγχο της ποιότητας του RDF / XML με ένα εργαλείο ανοικτού κώδικα, το Jena Eyeball, το προκύπτον αρχείο του καταλόγου της BNB μετατράπηκε σε N-Triple αρχείο για φόρτωση στην πλατφόρμα φιλοξενίας. Η πλατφόρμα που επιλέχθηκε είναι η

TSO<sup>21</sup> (OpenUp® Platform: What is the OpenUp® Platform?, χ.χ.) πλατφόρμα, μια συλλογή ολοκληρωμένων υπηρεσιών που βασίζονται στο cloud διαθέσιμο ως Software as a Service (SaaS), επιτρέποντας ταχεία ανάπτυξη σε λογικό κόστος. Οι υπηρεσίες παρέχουν μια εξαιρετικά επεκτάσιμη και ανθεκτική πλατφόρμα που επιτρέπει τη συγκομιδή, τον εμπλουτισμό, την αποθήκευση και τη διάθεση δεδομένων για την αναζήτηση (OpenUp® Platform: What is the OpenUp® Platform?, χ.χ.). Ωστόσο, στην εν λόγω πρακτική που εμπεριέχεται στην διπλωματική αυτή δε χρησιμοποιήθηκε η πλατφόρμα αυτή λόγω του οικονομικού κόστους, παρόλο που οι δυνατότητες και οι υπηρεσίες της θα εξυπηρετούσαν τον σκοπό και τους στόχους της.

Επιπλέον, το σύνολο των δεδομένων BNB περιλαμβάνει συνδέσμους ή αναφορές σε άλλα σύνολα δεδομένων που περιέχουν χρήσιμα συγκριτικά στοιχεία και ειδικότερα είναι συνδεδεμένο με τα εξής σύνολα δεδομένων:

- ISNI<sup>22</sup>, το Διεθνές Αναγνωριστικό Πρότυπο Ονομάτων (International Standard Name Identifier)
- VIAF<sup>23</sup>, το Virtual International Authority File
- LCSH<sup>24</sup>, το Library of Congress subject headings δηλωμένες σε RDF;
- Lexvo<sup>25</sup>, τυποποιημένοι κωδικοί γλωσσών
- GeoNames<sup>26</sup>, για τις χώρες δημοσίευσης
- MARC<sup>27</sup>, πεδία χωρών και γλώσσας δηλωμένα σε RDF σύμφωνα με τη Βιβλιοθήκη του Κονγκρέσου
- Dewey.info, για την πρόσβαση στα πρώτα τρία επίπεδα του ταξινομικού συστήματος
- RDF Book Mashup, για πληροφορίες βιβλίων, συγγραφέων, κριτικών και online βιβλιοπωλείων διαθέσιμων στο Σημασιολογικό Ιστό
- UK Government Interval Set, για τη χρονιά δημοσίευσης.

Μέσω των παραπάνω συνδέσεων, τα δεδομένα της Βρετανικής Βιβλιοθήκης λαμβάνουν μία πιο εμπλουτισμένη διάσταση και προσφέρουν πιο εμπεριστατωμένες και ολοκληρωμένες πληροφορίες στον τελικό χρήστη.

Τέλος, τα δεδομένα της BNB μπορούν να είναι προσβάσιμα μέσω του SPARQL Endpoint που παρέχεται στη διεύθυνση: <http://bnb.data.bl.uk/flint-sparql>. Το SPARQL Endpoint προσφέρεται μέσω του Flint SPARQL Editor (βλ. εικ. 12), ενός συντάκτη με ανοιχτή άδεια, τον οποίο μπορεί να ενσωματώσει ο καθένας στην ιστοσελίδα του. Ο συντάκτης αυτός είναι περίπλοκος για κάποιον που δεν έχει άμεση επαφή με τις γλώσσες του Σημασιολογικού Ιστού και τη γλώσσα SPARQL ενώ δεν είναι εύχρηστος. Τα ερωτήματα και τα αποτελέσματα εμφανίζονται στο ίδιο παράθυρο και ο ερευνητής πρέπει να μετακινηθεί πολύ προς τα κάτω (scrolling) μέχρι να αναγνώσει όλα τα αποτελέσματα. Τα παραπάνω

<sup>21</sup> <http://openup.tso.co.uk/openup-platform>

<sup>22</sup> <http://www.isni.org/>

<sup>23</sup> <https://viaf.org/>

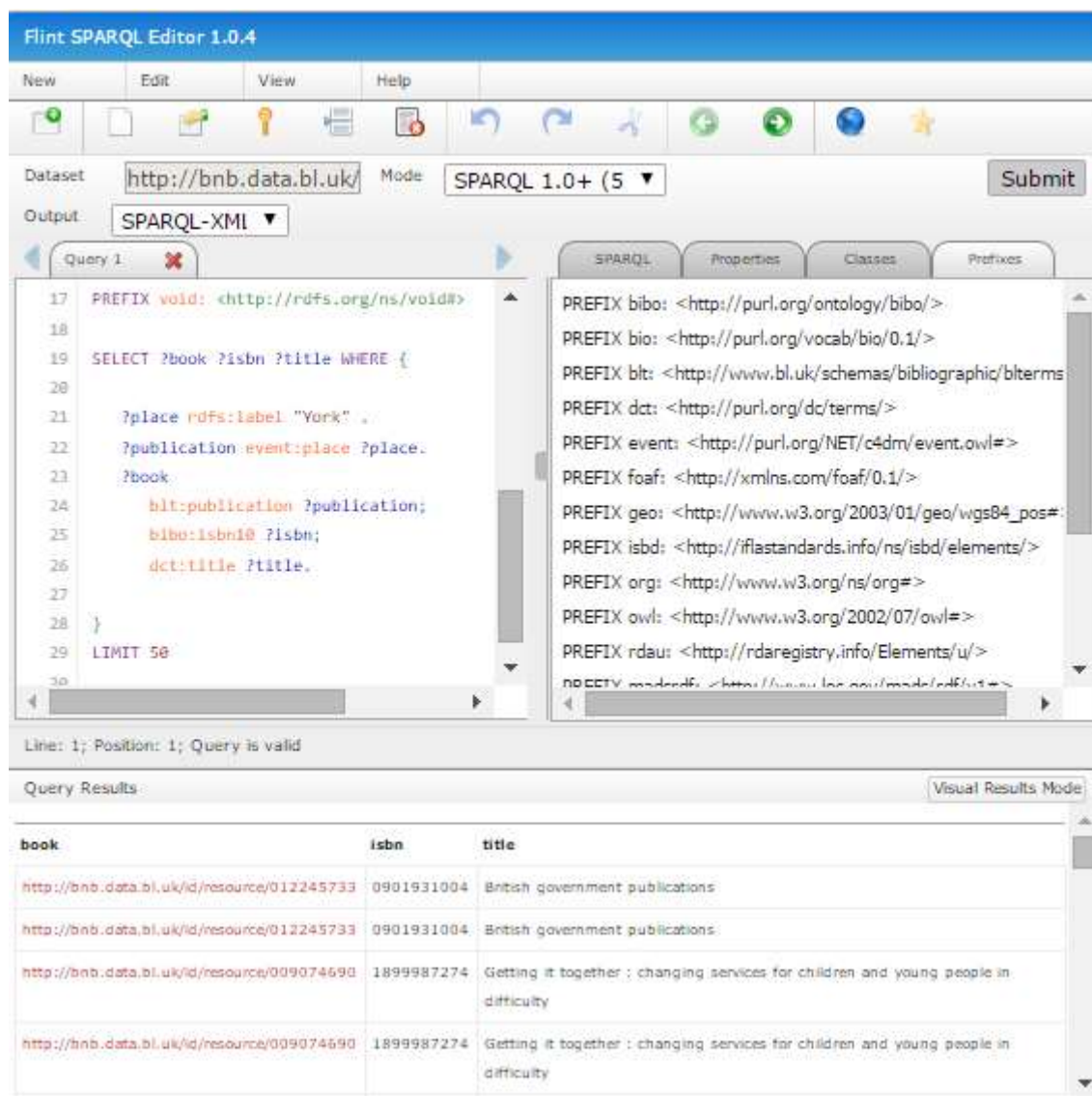
<sup>24</sup> <http://id.loc.gov/authorities/subjects.html>

<sup>25</sup> <http://www.lexvo.org/>

<sup>26</sup> <http://www.geonames.org/>

<sup>27</sup> <http://www.loc.gov/marc/>

είναι οι βασικοί λόγοι που ο συντάκτης αυτός δεν επιλέχθηκε για την προτεινόμενη πρακτική που παρουσιάζεται στην εργασία αυτή.



**Εικόνα 12: Sparql Endpoint της BNB**

Παρακάτω μπορεί να δει κανείς ένα παράδειγμα ερωτήματος (βλ. εικ. 13) και τα αποτελέσματα του στο συγκεκριμένο τελικό σημείο (βλ. εικ. 14). Στο ερώτημα αυτό ζητούνται 50 τίτλοι βιβλίων που έχουν δημοσιευθεί στη Νέα Υόρκη.

```

19 SELECT ?book ?isbn ?title WHERE {
20
21   ?place rdfs:label "York" .
22   ?publication event:place ?place.
23   ?book
24     blt:publication ?publication;
25     bibo:isbn10 ?isbn;
26     dct:title ?title.
27
28 }
29 LIMIT 50

```

**Εικόνα 13: Παράδειγμα ερωτήματος SPARQL στο Endpoint της BNB**

<a href="http://bnb.data.bl.uk/id/resource/012245733">http://bnb.data.bl.uk/id/resource/012245733</a>	0901931004	British government publications
<a href="http://bnb.data.bl.uk/id/resource/012245733">http://bnb.data.bl.uk/id/resource/012245733</a>	0901931004	British government publications
<a href="http://bnb.data.bl.uk/id/resource/009074690">http://bnb.data.bl.uk/id/resource/009074690</a>	1899987274	Getting it together : changing services for children and young people in difficulty
<a href="http://bnb.data.bl.uk/id/resource/009074690">http://bnb.data.bl.uk/id/resource/009074690</a>	1899987274	Getting it together : changing services for children and young people in difficulty
<a href="http://bnb.data.bl.uk/id/resource/011601687">http://bnb.data.bl.uk/id/resource/011601687</a>	0900657073	Yorkshire flooding - some effects on man and nature

**Εικόνα 14: Δείγμα αποτελεσμάτων από το SPARQL Endpoint της BNB**

## 3.2. *Bibliothèque Nationale de France*

Η Εθνική Βιβλιοθήκη της Γαλλίας (BNF) διέθεσε τα δεδομένα της ως Ανοιχτά και Διασυνδεδεμένα Δεδομένα τον Ιούνιο του 2011. Οι βασικοί στόχοι του έργου αυτού ήταν οι εξής (About data.bnf.fr, χ.χ.):

- Να γίνουν τα δεδομένα που παράγονται από την BNF περισσότερο ορατά μέσω του Παγκόσμιου Ιστού
- Να συνδεθούν τα δεδομένα που παράγονται από την BNF, τόσο εντός όσο και εκτός των καταλόγων της
- Να συμβάλλουν στη συνεργασία και την ανταλλαγή μεταδεδομένων μέσω της δημιουργίας συνδέσεων μεταξύ αξιόπιστων πηγών δεδομένων,
- Να διευκολυνθεί η επαναχρησιμοποίηση των μεταδεδομένων (υπό Ανοιχτές Άδειες) από τρίτους.

Τον Ιούλιο του 2014, το έργο διασύνδεσης των δεδομένων της Εθνικής Βιβλιοθήκης της Γαλλίας που ονομάζεται «data.bnf.fr»<sup>28</sup> αποτελείτο από 400.000 συγγραφείς, που συνδέονταν με πάνω από 7 εκατομμύρια έγγραφα.

<sup>28</sup> <http://data.bnf.fr/>

Για την οικοδόμηση των URIs, η Εθνική Βιβλιοθήκη της Γαλλίας (BNF) χρησιμοποίησε τις προδιαγραφές ARK (Archival Resource Key), τα οποία είναι αναγνωριστικά URLs (Uniform Resource Locator) πολλαπλών χρήσεων για πληροφοριακά αντικείμενα οποιουδήποτε τύπου, ώστε να εκχωρήσει αυτόματα ένα αναγνωριστικό στα έγγραφα, με βάση τον αριθμό των αρχείων. Ο στόχος ήταν να οικοδομηθούν URIs που είναι και σταθερά και αναγνώσιμα από τον άνθρωπο (έτσι ώστε ένα URI μπορεί να σχετίζονται σαφώς με το περιεχόμενο) (Simon, κ. α., 2013).

Η δραστηριότητα μοντελοποίησης των δεδομένων πραγματοποιήθηκε με το μοντέλο βιβλιογραφικών εγγραφών της FRBR. Πιο ειδικά, η μοντελοποίηση πραγματοποιήθηκε μετατρέποντας τα προϋπάρχοντα αρχεία σε μορφή MARC και τους καθιερωμένους όρους, σε RDF τριπλέτες. Στη συνέχεια, εφαρμόστηκε το πρότυπο FRBR κι έτσι δημιουργήθηκαν νέα σύνολα δεδομένων που ήταν πιο κοντά στις αρχές των Διασυνδεδεμένων Δεδομένων αλλά ακόμα δεν μπορούσαν να θεωρηθούν ως Διασυνδεδεμένα Δεδομένα. Στο σημείο αυτό πρέπει να αναφερθεί πως αν και η χρήση του παραπάνω μοντέλου θεωρείται ως μία πρακτική για το πεδίο της βιβλιοθηκονομίας, έρχεται σε αναντιστοιχία με τον στόχο του συγκεκριμένου έργου να διατεθούν τα δεδομένα της Βιβλιοθήκης στην ευρύτερη κοινότητα. Ίσως να ήταν πιο χρήσιμο να χρησιμοποιηθούν ευρέως γνωστές οντολογίες του Σημασιολογικού Ιστού για την εύκολη αξιοποίηση των δεδομένων σε εφαρμογές παντός τύπου και με πολλαπλή χρησιμότητα.

Τα επαναχρησιμοποιήσιμα δεδομένα της Εθνικής Βιβλιοθήκης της Γαλλίας (BNF) φιλοξενήθηκαν αρχικά στο αποθετήριο RAMEAU, το οποίο χρησιμοποιήθηκε για την προβολή και τη διαχείριση των βιβλιογραφικών εγγραφών της BNF. Ωστόσο, το αποθετήριο αυτό αντικαταστάθηκε το 2012 με το αποθετήριο CATCH (About data.bnf.fr, χ.χ).

Η Εθνική Βιβλιοθήκη της Γαλλίας επέλεξε το CubicWeb<sup>29</sup> μεταξύ άλλων λύσεων λογισμικού για να φιλοξενήσει τα δεδομένα της επειδή είχε καλές συστάσεις, δεν είχε οικονομικό κόστος καθώς επίσης και γιατί έδινε τη δυνατότητα να δημοσιεύονται δεδομένα με τα πρότυπα του Σημασιολογικού Ιστού (RDF, SPARQL, HTML5, CSS3) αλλά και να είναι διαχειρίσιμα υπό πολλές μορφές (Simon, κ. α., 2013).

Η Βιβλιοθήκη που παρουσιάζεται είναι διασυνδεδεμένη με διάφορα άλλα σύνολα δεδομένων, ώστε να προσφέρει στους χρήστες της μία πιο ολοκληρωμένη εμπειρία αναζήτησης. Πιο συγκεκριμένα, υπάρχουν συνδέσεις της Εθνικής Βιβλιοθήκης της Γαλλίας και των συνόλων δεδομένων που προέρχονται από τα εξής:

- θεματικές επικεφαλίδες της Βιβλιοθήκης του Κογκρέσου (Library of Congress<sup>30</sup>)
- θεματικές επικεφαλίδες της Εθνικής Γερμανικής Βιβλιοθήκης (Deutsche Nationalbibliothek, DNB<sup>31</sup>)
- τις θεματικές επικεφαλίδες του θησαυρού των Εθνικών Αρχείων της Γαλλίας (Thesaurus W)
- γεωγραφικές επικεφαλίδες Geonames<sup>32</sup>

<sup>29</sup> <https://www.cubicweb.org/>

<sup>30</sup> <http://id.loc.gov/authorities/subjects.html>

<sup>31</sup> [http://www.dnb.de/DE/Home/home\\_node.html](http://www.dnb.de/DE/Home/home_node.html)

- θησαυρός Agrovoc<sup>33</sup>
- θεματικές επικεφαλίδες από το σύστημα ταξινόμησης Dewey
- id.loc.gov για τις γλώσσες και τους κωδικούς χωρών

*Τα δεδομένα σχετικά με τους συγγραφείς συνδέονται επίσης με:*

- VIAF<sup>34</sup>, Virtual International Authority File
- Το Διεθνές Αναγνωριστικό Πρότυπο Ονομάτων (ISNI<sup>35</sup> - International Standard Name Identifier) για τα πρόσωπα και τους οργανισμούς
- IDREF<sup>36</sup>, το αποθετήριο των Βιβλιοθηκών της Τριτοβάθμιας Εκπαίδευσης της Γαλλίας
- DBpedia<sup>37</sup>

Όσον αφορά τα δεδομένα της Βιβλιοθήκης αυτής είναι προσπελάσιμα μέσω ενός τελικού σημείου αναζήτησης (Endpoint), που βρίσκεται στην εξής διεύθυνση: <http://data.bnf.fr/sparql/> και το οποίο παρουσιάζει απόλυτη λειτουργικότητα. Το τελικό αυτό σημείο προσφέρεται από τον OpenLink Virtuoso<sup>38</sup>, που είναι ένας Server Δεδομένων πολλαπλών χρήσεων και πρωτοκόλλων. Στην παρακάτω εικόνα παρουσιάζεται ένα συγκεκριμένο ερώτημα εκφρασμένο σε γλώσσα ερωτήσεων SPARQL στα γαλλικά, που αφορά όλες τις ημερομηνίες που σχετίζονται με τους συγγραφείς αλλά και το ονοματεπώνυμο αυτών (βλ. εικ. 15).

**Εικόνα 15: Παράδειγμα ερωτήματος SPARQL στο Endpoint της BNF**

<sup>32</sup> <http://www.geonames.org/>

<sup>33</sup> <http://www4.fao.org/faobib/kwocinana.html>

<sup>34</sup> <https://viaf.org/>

<sup>35</sup> <http://www.isni.org/>

<sup>36</sup> <http://www.idref.fr/autorites/autorites.html>

<sup>37</sup> <http://wiki.dbpedia.org/>

<sup>38</sup> <http://virtuoso.openlinksw.com/>

Ένα δείγμα των αποτελεσμάτων του ερωτήματος αυτού παρουσιάζεται παρακάτω (βλ. εικ. 16). Όπως μπορεί να αντιληφθεί κανείς κι από την εικόνα αυτή τα αποτελέσματα εμφανίζονται με τη μορφή πινάκων σύμφωνα με τις μεταβλητές που ζητήθηκαν από το ερώτημα. Πιο ειδικά, στην πρώτη στήλη εμφανίζονται τα αναγνωριστικά URI των «Συγγραφέων» (auteur), στη δεύτερη η «Ημερομηνία Γέννησης» (jour), στην τρίτη η «Έτος Γέννησης» (date1), στην τέταρτη η «Έτος Θανάτου» (date2) και στην Πέμπτη το «Όνοματεπώνυμο» (nom) του συγγραφέα:

auteur	jour	date1	date2	nom
<a href="http://data.bnf.fr/ark:/12148/cb124191432#foaf:Person">http://data.bnf.fr/ark:/12148/cb124191432#foaf:Person</a>	"..."	"1475"	"1541"	"Francesco Bassano"
<a href="http://data.bnf.fr/ark:/12148/cb12448580#foaf:Person">http://data.bnf.fr/ark:/12148/cb12448580#foaf:Person</a>	"..."	"1440"	"1518"	"Rodrigo de Osona"
<a href="http://data.bnf.fr/ark:/12148/cb12448582#foaf:Person">http://data.bnf.fr/ark:/12148/cb12448582#foaf:Person</a>	"..."	"1465"	"1514"	"Francesc d' Osona"
<a href="http://data.bnf.fr/ark:/12148/cb12449390#foaf:Person">http://data.bnf.fr/ark:/12148/cb12449390#foaf:Person</a>	"..."	"1623"	"1683"	"Giovanni Ghisolfi"
<a href="http://data.bnf.fr/ark:/12148/cb124566697#foaf:Person">http://data.bnf.fr/ark:/12148/cb124566697#foaf:Person</a>	"..."	"1748"	"1832"	"Asensio Julia"
<a href="http://data.bnf.fr/ark:/12148/cb12460109#foaf:Person">http://data.bnf.fr/ark:/12148/cb12460109#foaf:Person</a>	"..."	"1525"	"1593"	"Rafael de León"
<a href="http://data.bnf.fr/ark:/12148/cb12507229#foaf:Person">http://data.bnf.fr/ark:/12148/cb12507229#foaf:Person</a>	"..."	"1706"	"176."	"Diego Camporredondo"
<a href="http://data.bnf.fr/ark:/12148/cb125149390#foaf:Person">http://data.bnf.fr/ark:/12148/cb125149390#foaf:Person</a>	"..."	"19."	"19."	"Wolfgang Schreiner"
<a href="http://data.bnf.fr/ark:/12148/cb12516278#foaf:Person">http://data.bnf.fr/ark:/12148/cb12516278#foaf:Person</a>	"..."	"189."	"197."	"William Appleton Coolidge"
<a href="http://data.bnf.fr/ark:/12148/cb13167998q#foaf:Person">http://data.bnf.fr/ark:/12148/cb13167998q#foaf:Person</a>	"..."	"1716"	"1796"	"Carl Fredrik Adelcrantz"
<a href="http://data.bnf.fr/ark:/12148/cb13330348j#foaf:Person">http://data.bnf.fr/ark:/12148/cb13330348j#foaf:Person</a>	"..."	"1745"	"1824"	"Anne-Louise Boivin d'Hardencourt Brillon de Jouy"
<a href="http://data.bnf.fr/ark:/12148/cb135069231#foaf:Person">http://data.bnf.fr/ark:/12148/cb135069231#foaf:Person</a>	"..."	"715"	"794"	
<a href="http://data.bnf.fr/ark:/12148/cb13510549c#foaf:Person">http://data.bnf.fr/ark:/12148/cb13510549c#foaf:Person</a>	"..."	"1015"	"1076-04-18"	
<a href="http://data.bnf.fr/ark:/12148/cb13534923q#foaf:Person">http://data.bnf.fr/ark:/12148/cb13534923q#foaf:Person</a>	"..."	"1760"	"1793-03-31"	"Louis Bruneau"
<a href="http://data.bnf.fr/ark:/12148/cb13618819c#foaf:Person">http://data.bnf.fr/ark:/12148/cb13618819c#foaf:Person</a>	"..."	"1667"	"1693"	"George Douglas"
<a href="http://data.bnf.fr/ark:/12148/cb13736306#foaf:Person">http://data.bnf.fr/ark:/12148/cb13736306#foaf:Person</a>	"..."	"1435"	"1498"	"Michael Pachet"
<a href="http://data.bnf.fr/ark:/12148/cb13747299#foaf:Person">http://data.bnf.fr/ark:/12148/cb13747299#foaf:Person</a>	"..."	"1767"	"1822"	"Denmark Vesey"
<a href="http://data.bnf.fr/ark:/12148/cb137566378#foaf:Person">http://data.bnf.fr/ark:/12148/cb137566378#foaf:Person</a>	"..."	"1450"	"1494-02-04"	
<a href="http://data.bnf.fr/ark:/12148/cb139524941g#foaf:Person">http://data.bnf.fr/ark:/12148/cb139524941g#foaf:Person</a>	"..."	"1867"	"1958"	"Sydney Fremantle"
<a href="http://data.bnf.fr/ark:/12148/cb15072680#foaf:Person">http://data.bnf.fr/ark:/12148/cb15072680#foaf:Person</a>	"..."	"18."	"18."	"Aimé Carpiet"
<a href="http://data.bnf.fr/ark:/12148/cb14522616h#foaf:Person">http://data.bnf.fr/ark:/12148/cb14522616h#foaf:Person</a>	"..."	"1764"	"1813-11-15"	"Jean-Pierre Simon"
<a href="http://data.bnf.fr/ark:/12148/cb15121890#foaf:Person">http://data.bnf.fr/ark:/12148/cb15121890#foaf:Person</a>	"..."	"1792"	"1858-11-28"	"Giuseppe Nizzoli"
<a href="http://data.bnf.fr/ark:/12148/cb155216690#foaf:Person">http://data.bnf.fr/ark:/12148/cb155216690#foaf:Person</a>	"..."	"1875"	"1939"	"Federico Alvarez de Toledo"

**Εικόνα 16: Δείγμα αποτελεσμάτων από το SPARQL Endpoint της BNF**

Τέλος, είναι αναγκαίο να αναφερθεί πως ο συγκεκριμένος Server (δηλ. ο OpenLink Virtuoso), όπως και το τελικό σημείο αναζήτησης (Endpoint) που προσφέρει αξιοποιήθηκε και για την υλοποίηση της πρακτικής που παρουσιάζεται στη συγκεκριμένη διπλωματική. Οι λόγοι επιλογής και υπηρεσίες που προσφέρονται από τον OpenLink Virtuoso θα αναλυθούν στα επόμενα κεφάλαια.

### 3.3. Deutsche Nationalbibliothek (DNB)

Η Γερμανική Εθνική Βιβλιοθήκη (Deutsche Nationalbibliothek - DNB) το 2010 αποφάσισε να κάνει την συμβολή της στο Σύννεφο των Ανοιχτών και Διασυνδεδεμένων Δεδομένων με την παροχή δεδομένων υψηλής ποιότητας. Με την προσφορά των δεδομένων της σε RDF μορφή επιτρέπει στους χρήστες σαν μονάδες αλλά και σε ομάδες χρηστών να επαναχρησιμοποιήσουν τα δεδομένα της κατά τρόπο που δεν απαιτεί καμία γνώση των βιβλιογραφικών μορφών. Η υπηρεσία αυτή υπόκειται έως και σήμερα σε συνεχή περαιτέρω ανάπτυξη και βελτιστοποίηση όσον αφορά το τεχνικό κομμάτι, το περιεχόμενο και το

οργανωτικό επίπεδο (Case study: German National Library publishes its authority and bibliographic records as linked data, χ.χ.).

Τα βιβλιογραφικά δεδομένα της βιβλιοθήκης ήταν σε μορφή MARC21, όπως και οι περισσότερες βιβλιοθήκες που έχουν αναλυθεί στο παρόν κεφάλαιο, σε αντίθεση με τα δεδομένα της Δημοτικής Βιβλιοθήκης Ευόσμου που αποτελεί το παράδειγμα για την παρουσίαση της προτεινόμενης πρακτικής στην εν λόγω διπλωματική εργασία. Τα MARC21 δεδομένα μετατράπηκαν σε RDF μέσω του Metafacture<sup>39</sup>, που αναπτύχθηκε από την [culturegraph.org](http://culturegraph.org) και είναι ένα εργαλείο που μετατρέπει δεδομένα σε RDF. Ακόμα, για την περιγραφή των δεδομένων της Βιβλιοθήκης δημιουργήθηκε μία οντολογία που συναντάται με το όνομα "GND ontology" και αξιοποιήθηκε για τη μετατροπή των δεδομένων από το πρότυπο MARC21 στο μοντέλο RDF (Haffner, 2015).

Η πληροφόρηση για τα URIs των δεδομένων της Βιβλιοθήκης αυτής είναι περιορισμένη. Στο σημείο αυτό είναι αξιοσημείωτο να ειπωθεί πως το αγγλικό κείμενο τεκμηρίωσης του έργου δεν έχει ενημερωθεί και διατίθεται μόνο στην γερμανική γλώσσα. Το γεγονός αυτό δυσκολεύει και περιορίζει την έρευνα.

Τα δεδομένα της Βιβλιοθήκης δε φιλοξενούνται σε κάποια πλατφόρμα, ωστόσο μπορεί να ανακτήσει κανείς το RDF αρχείο που περιέχει τα δεδομένα του καταλόγου από το Datahub, την ελεύθερη πλατφόρμα διαχείρισης δεδομένων που προσφέρεται από το Ίδρυμα Ανοιχτής Γνώσης (Open Knowledge Foundation<sup>40</sup>) και βασίζεται στο σύστημα διαχείρισης δεδομένων CKAN<sup>41</sup> (Datahub, 2015). Στο Datahub διατίθενται τα αρχεία RDF ή κομμάτια αυτών από όλα τα σύνολα δεδομένων που υπάρχουν στο Σύννεφο Ανοιχτών Διασυνδεδεμένων Δεδομένων (Open Linked Data Cloud).

Η Γερμανική Εθνική Βιβλιοθήκη (DNB) δεν παρέχει καμία διασύνδεση με άλλες πηγές δεδομένων, εκτός από το θησαυρό οικονομικών STW<sup>42</sup> (Papadakis, Kyprianos & Stefanidakis, 2015).

Τα δεδομένα της Βιβλιοθήκης αυτής δεν προσφέρονται μέσω κάποιου τελικού σημείου (Endpoint) για ερωτήματα σε SPARQL. Το παραπάνω θεωρείται ένα μειονέκτημα του έργου αυτού καθώς τα δεδομένα δεν μπορούν να ανακτηθούν από τους ενδιαφερόμενους χρήστες, παρά μόνο να χρησιμοποιηθούν σε πιθανές εφαρμογές.

Εν κατακλείδι, αξίζει να σημειωθεί πως το έργο της Βιβλιοθήκης αυτής για το «άνοιγμα» και τη «διασύνδεση» των δεδομένων χαρακτηρίζεται ως ελλιπές και στην παρούσα διπλωματική αξιοποιήθηκε ως μία πρακτική προς αποφυγή για τη δημοσιοποίηση Ανοιχτών Διασυνδεδεμένων Δεδομένων σε βιβλιοθήκες.

---

<sup>39</sup> <https://github.com/culturegraph/metafacture-core/wiki>

<sup>40</sup> <https://okfn.org/>

<sup>41</sup> <http://ckan.org/>

<sup>42</sup> <http://zbw.eu/stw/version/latest/about>



### 3.4. *Biblioteca Nacional de España* (BNE)

Η Εθνική Βιβλιοθήκη της Ισπανίας (Biblioteca Nacional de España - BNE), πραγματοποίησε ένα έργο με όνομα «Datos.bne.es» που ως σκοπό είχε τον εμπλουτισμό του σημασιολογικού Ιστού με τα δεδομένα της Βιβλιοθήκης με τη μορφή των Ανοιχτών και Διασυνδεδεμένων Δεδομένων. Η πρώτη συνάντηση πραγματοποιήθηκε το Σεπτέμβριο του 2010, ενώ η συμφωνία συνεργασίας υπεγράφη τον Οκτώβριο του 2010. Η πρώτη σειρά δεδομένων μεταφορτώθηκε και διασυνδέθηκε τον Απρίλιο του 2011, αλλά η μεταφόρτωση και διασύνδεση του σημαντικού συνόλου των δεδομένων υλοποιήθηκε τον Δεκέμβριο του 2011.

Τα δεδομένα της Βιβλιοθήκης αυτής ήταν περιγεγραμμένα με το πρότυπο MARC21 (Papadakis, Kyprianos & Stefanidakis, 2015) και εξήχθησαν πάνω από 4.000.000 εγγραφές των βιβλιογραφικών δεδομένων, συμπεριλαμβανομένων των σύγχρονων και των αρχαίων μονογραφιών, των ηλεκτρονικών πηγών, των χειρόγραφων, των περιοδικών, των χαρτών, των φωτογραφιών, των παρτιτούρων και των οπτικοακουστικών έργων. Ακόμα, έχουν εξαχθεί πάνω από 4.000.000 δεδομένα καθιερωμένων όρων για ανθρώπους, οντότητες, συνέδρια, έργα, εκφράσεις, αντικείμενα, γεωγραφικά ονόματα και είδη (Biblioteca Nacional de España: Data Source, χ.χ.).

Όσον αφορά τις οντολογίες που χρησιμοποιήθηκαν για την αναπαράσταση των πόρων σε RDF, το έργο «datos.bne.es» συμπεριέλαβε τις οντολογίες που αναπτύχθηκαν από την IFLA<sup>43</sup> (International Federation of Library Associations) και αφορούν τα δεδομένα των Βιβλιοθηκών, δηλαδή την FRBR (Functional Requirements for Bibliographic Records), την FRAD (Functional Requirements for Authority Data), την FRSAD (Functional Requirements for Subject Authority Data) και το ISBD (International Standard for Bibliographic Description) (Vila - Suero, Villazón - Terrazas & Gómez - Pérez, 2012).

Για τη διαδικασία μετασχηματισμού των δεδομένων από το πρότυπο MARC21 σε μορφή RDF χρησιμοποιήθηκε το εργαλείο MARiMba, που αναπτύχθηκε από το Ontology Engineering Group<sup>44</sup> και είναι ένα εργαλείο σχεδιασμένο για τους βιβλιοθηκονόμους και υποστηρίζει το σύνολο της διαδικασίας παραγωγής RDF αρχείων χρησιμοποιώντας δεδομένα που είναι περιγεγραμμένα με το πρότυπο MARC21 αλλά και της σύνδεσης αυτών με τα σύνολα δεδομένων όπως το VIAF<sup>44</sup> και τη DBpedia<sup>45</sup> (Biblioteca Nacional de España: Technologies Used, χ.χ.). Για την κατασκευή των URIs των RDF πόρων, το MARiMba παρείχε τη λειτουργικότητα για να δημιουργηθούν σύνδεσμοι εκμεταλλευόμενοι τη σύνδεση με το VIAF (Vila-Suero, Villazón-Terrazas & Gómez-Pérez, 2012). Στο σημείο αυτό είναι αναγκαίο να σημειωθεί πως ενώ το εργαλείο αυτό θεωρείται ως μία καλή επιλογή μετατροπής των δεδομένων, δε διατίθεται στο διαδίκτυο για την ευρύτερη χρήση του από το ενδιαφερόμενο κοινό κι αυτός είναι ο λόγος που δεν αξιοποιήθηκε στην εν

<sup>43</sup> <http://www.ifla.org/>

<sup>44</sup> <https://viaf.org/>

<sup>45</sup> <http://wiki.dbpedia.org/>

λόγω διπλωματική εργασία για τη μετατροπή των δεδομένων της Δημοτικής Βιβλιοθήκης Ευόσμου.

Μέσω της διαδικασίας αυτής προέκυψαν περίπου 30.000.000 τριπλέτες στα ισπανικά, αυξάνοντας σημαντικά την παρουσία της ισπανικής γλώσσας στο Σύννεφο Ανοιχτών και Διασυνδεδεμένων Δεδομένων (Biblioteca Nacional de España: Technologies Used, χ.χ.).

Τα δεδομένα της Βιβλιοθήκης εμπλουτίστηκαν με συνδέσεις προς τις αντίστοιχες πηγές δεδομένων και πιο ειδικά, για τους συγγραφείς παρέχονται συνδέσεις προς (Vila - Suero, Villazón - Terrazas & Gómez - Pérez, 2012):

- τη Βιβλιοθήκη του Κογκρέσου<sup>46</sup> (Library of Congress)
- τη Γερμανική Εθνική Βιβλιοθήκη<sup>47</sup> (Deutsche Nationalbibliothek - DNB)
- την Εθνική Βιβλιοθήκη της Γαλλίας<sup>48</sup> (Bibliothèque nationale de France - BnF)
- Sudoc<sup>49</sup>, τον γαλλικό συλλογικό κατάλογο των Ανώτατων Εκπαιδευτικών Ιδρυμάτων και των Ερευνητικών Βιβλιοθηκών (Sudoc Catalogue, 2015).
- την Εθνική Βιβλιοθήκη της Σουηδίας<sup>50</sup> (Kungliga biblioteket)
- Lexvo<sup>51</sup>
- VIAF<sup>52</sup>, το Virtual International Authority File
- ISNI<sup>53</sup>, το Διεθνές Αναγνωριστικό Πρότυπο Ονομάτων (International Standard Name Identifier)

Τα δεδομένα που σχετίζονται με τα θέματα, τις γεωγραφικές περιοχές και τα είδη συνδέονται με τη Βιβλιοθήκη του Κογκρέσου.

Τα δεδομένα της Εθνικής Βιβλιοθήκης της Ισπανίας (Biblioteca Nacional de España - BNE) φιλοξενούνται στον διακομιστή OpenLink Virtuoso<sup>54</sup> και ταυτόχρονα προσφέρονται από το τελικό σημείο (Endpoint) για ερωτήματα SPARQL στους χρήστες χρησιμοποιώντας το δωρεάν εργαλείο Pubby<sup>55</sup>, που είναι ένα εργαλείο διεπαφής για τελικά σημεία (Endpoints) SPARQL. Αν και υπάρχει το τελικό σημείο (Endpoint) και διατίθεται στον σύνδεσμο: <http://datos.bne.es/sparql>, ωστόσο, δε λειτουργεί κατά την υλοποίηση την τεκμηρίωσης της εν λόγω διπλωματικής κι έτσι είναι ανέφικτο να παρουσιαστούν εικόνες παραδειγμάτων σε γλώσσα SPARQL. Η παραπάνω δυσλειτουργία είναι λογική αν αναλογιστεί κανείς πως τα δεδομένα της Βιβλιοθήκης αυτής έχουν να ανανεωθούν από το 2013.

---

<sup>46</sup> <http://id.loc.gov/authorities/subjects.html>

<sup>47</sup> [http://www.dnb.de/DE/Home/home\\_node.html](http://www.dnb.de/DE/Home/home_node.html)

<sup>48</sup> <http://www.bnf.fr/fr/acc/x.accueil.html>

<sup>49</sup> <http://www.sudoc.abes.fr/?COOKIE=U10178,Klecteurweb,D2.1,E88a9c5f0-25b,I250,B341720009+,SY,A%5C9008+1,,J,H2-26,,29,,34,,39,,44,,49-50,,53-78,,80-87,NLECTEUR+PSI,R109.242.205.79,FN>

<sup>50</sup> <http://www.kb.se/english/>

<sup>51</sup> <http://www.lexvo.org/>

<sup>52</sup> <https://viaf.org/>

<sup>53</sup> <http://www.isni.org/>

<sup>54</sup> <http://virtuoso.openlinksw.com/>

<sup>55</sup> <http://wifo5-03.informatik.uni-mannheim.de/pubby/>

### 3.5. Δημόσια Βιβλιοθήκη Βέροιας

Στα πλαίσια Διπλωματικής Εργασίας του ΠΜΣ «Επιστήμη του Διαδικτύου» του Αριστοτέλειου Πανεπιστημίου Θεσσαλονίκης στο Τμήμα Μαθηματικών, το 2013 πραγματοποιήθηκε η προσπάθεια δημοσιοποίησης των βιβλιογραφικών δεδομένων της Δημόσιας Βιβλιοθήκης Βέροιας υπό τις Αρχές των Διασυνδεδεμένων Δεδομένων, προσφέροντας με τον τρόπο αυτό υπηρεσίες Σηματολογικού Ιστού στους χρήστες της Βιβλιοθήκης (Καραμπατάκης, 2013). Η προσπάθεια αυτή ήταν η πρώτη στο χώρο των ελληνικών βιβλιοθηκών.

Τα δεδομένα της Βιβλιοθήκης ήταν σε μορφή .iso2709, περιέχοντας 70.000 εγγραφές κωδικοποιημένες σύμφωνα με το πρότυπο MARC21. Για την μετατροπή τους χρησιμοποιώντας "γνωστές" οντολογίες όπως Dublin Core, FOAF, BIBO και RDVOAB, ώστε τα δεδομένα να είναι κατανοητά ευανάγνωστα είτε από ανθρώπους είτε από μηχανές και οι προγραμματιστές να μπορούν να επαχρησιμοποιήσουν εύκολα τα δεδομένα αυτά (Καραμπατάκης, 2013).

Ο ερευνητής αδυνατώντας να βρει μία εφαρμογή μετατροπής αρχείων από το πρότυπο MARC21 σε RDF, ανέπτυξε μία νέα εφαρμογή σε γλώσσα C++, που αναλαμβάνει τη μετατροπή ή αντιστοίχιση των δεδομένων από το πρότυπο MARC σε RDF τριπλέτες. Αναλυτική αναφορά για τη συγκεκριμένη εφαρμογή θα πραγματοποιηθεί παρακάτω, καθώς ήταν αυτή που χρησιμοποιήθηκε και για την υλοποίηση της εν λόγω πρακτικής που περιγράφεται στην εν λόγω διπλωματική.

Με την εφαρμογή αυτή παράχθηκαν 1.300.000 τριπλέτες και στη συνέχεια χρησιμοποιήθηκε το Silk Workbench, μια διαδικτυακή εφαρμογή η οποία καθοδηγεί το χρήστη για δημιουργία συνδέσμων μεταξύ δύο πηγών δεδομένων. Τα δεδομένα της εν λόγω Βιβλιοθήκης διασυνδέθηκαν με την DBpedia.

Τα URI του εκάστοτε πληροφοριακού πόρου δημιουργήθηκαν με τέτοιο τρόπο, ώστε να είναι dereferencable, δηλαδή αν αναζητήσει κανείς ένα συγκεκριμένο πόρο και το αναγνωριστικό του (URI), να λαμβάνει ως απάντηση την περιγραφή του πόρου αυτού.

Για τη φιλοξενία των δεδομένων της εν λόγω Βιβλιοθήκης χρησιμοποιήθηκε, όπως και στις περισσότερες Βιβλιοθήκες που παρουσιάζονται στο κεφάλαιο αυτό, ο OpenLink Virtuoso<sup>56</sup> ως αποθετήριο τριπλετών. Εντούτοις, όπως παρουσιάστηκε και στις πιο πάνω μελέτες περιπτώσεις, ο Virtuoso προσφέρει τελικό σημείο (Endpoint) για ερωτήματα σε SPARQL, το οποίο αξιοποιήθηκε και από τον συγκεκριμένο ερευνητή.

Το SPARQL Endpoint της Δημόσιας Βιβλιοθήκης της Βέροιας, που προσφέρεται από τον σύνδεσμο: <http://libver.math.auth.gr:8890/sparql/>, χαρακτηρίζεται από απόλυτη λειτουργικότητα και αυτό μπορεί να γίνει απόλυτα αντιληπτό κι από το παράδειγμα ερωτήματος SPARQL που παρατίθεται παρακάτω (βλ. εικ. 17).

<sup>56</sup> <http://virtuoso.openlinksw.com/>

Στο παράδειγμα αυτό ζητούνται ο «Συγγραφέας», το «Όνομα» του συγγραφέα, ο «Τίτλος» του έργου που έχει συγγράψει και η «Ημερομηνία» (βλ. εικ. 17).

d)'. At the bottom are 'Run Query' and 'Reset' buttons."/>

Virtuoso SPARQL Query Editor

Default Data Set Name (Graph IRI)

Query Text

```
PREFIX bibo: <http://purl.org/ontology/bibo/>
PREFIX dcterms: <http://purl.org/dc/terms/>
PREFIX dc: <http://purl.org/dc/elements/1.1/>
SELECT *
WHERE
{
  ?author foaf:name ?name .
  ?x dc:contributor ?author .
  ?x dcterms:issued ?date .
  ?x bibo:title ?title .
}
```

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

Results Format:

Execution timeout:  milliseconds (values less than 0 are not allowed)

Options:  Strict checking of void variables

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

**Εικόνα 17: Παράδειγμα Ερωτήματος SPARQL στο Endpoint της Δημόσιας Βιβλιοθήκης Βέροιας**

Παρακάτω παρουσιάζονται στην εικόνα 10 (βλ. εικ. 18) τα αποτελέσματα του ερωτήματος αυτού:

author	name	x	date	title
http://libver.math.auth.gr/resource/agent20161	Juan Antonio Samaranch	http://libver.math.auth.gr/resource/record1_2752	2003	21 χρόνια στην προεδρία της Διεθνούς Ολυμπιακής Επιτροπής;
http://libver.math.auth.gr/resource/agent35517	Joe Turner	http://libver.math.auth.gr/resource/record1_45458	1988	Κοκτίδι;
http://libver.math.auth.gr/resource/agent1549	Edgar Wallace	http://libver.math.auth.gr/resource/record1_10918	1983	Ο άγγελος του τρόμου
http://libver.math.auth.gr/resource/agent1549	Edgar Wallace	http://libver.math.auth.gr/resource/record1_34210	1985 ή 1986	Οι τρεις δόκτοι
http://libver.math.auth.gr/resource/agent1549	Edgar Wallace	http://libver.math.auth.gr/resource/record1_45651	1988 ή 1989	Το πετρόμυο σμαράγδι
http://libver.math.auth.gr/resource/agent1549	Edgar Wallace	http://libver.math.auth.gr/resource/record1_51965	1991 ή 1992	Ο καταδότης;
http://libver.math.auth.gr/resource/agent1549	Edgar Wallace	http://libver.math.auth.gr/resource/record1_5529	2008	Το παιχνίδι της παλάμης;
http://libver.math.auth.gr/resource/agent1549	Edgar Wallace	http://libver.math.auth.gr/resource/record1_64637	2008 ή 2009	Το παιχνίδι της παλάμης;
http://libver.math.auth.gr/resource/agent1549	Edgar Wallace	http://libver.math.auth.gr/resource/record1_64676	2008 ή 2009	Η πόρτα με τις επτά κλειδαριές;
http://libver.math.auth.gr/resource/agent1549	Edgar Wallace	http://libver.math.auth.gr/resource/record1_64677	1982	Μια μερμή στο παράθυρο
http://libver.math.auth.gr/resource/agent1549	Edgar Wallace	http://libver.math.auth.gr/resource/record1_64680	2005	Και ο ύμνος χριστιανισμός
http://libver.math.auth.gr/resource/agent26185	Douglas Leon	http://libver.math.auth.gr/resource/record1_3410	2005	Θάνατος σε ζένη χάρη
http://libver.math.auth.gr/resource/agent21541	David Byrne	http://libver.math.auth.gr/resource/record1_29066	1981	My life in the bush of ghosts

**Εικόνα 18: Δείγμα αποτελεσμάτων από το SPARQL Endpoint της Δημόσιας Βιβλιοθήκης Βέροιας**

# 4. Δημοτική Βιβλιοθήκη Ευόσμου και Ανοιχτά Διασυνδεδεμένα Δεδομένα: Προτεινόμενη Πρακτική

Αφού μελετήθηκαν προσεκτικά οι πρακτικές που έχουν εφαρμοστεί τα προηγούμενα χρόνια στην Ελλάδα αλλά και στο εξωτερικό στον τομέα των βιβλιοθηκών και των Ανοιχτών και Διασυνδεδεμένων Δεδομένων, στο κεφάλαιο αυτό παρουσιάζεται αναλυτικά η διαδικασία υλοποίησης της προτεινόμενης πρακτικής που προτείνεται στην παρούσα διπλωματική. Τα δεδομένα που αξιοποιήθηκαν για την επιτυχή ολοκλήρωση της προτεινόμενης αυτής πρακτικής συλλέχθηκαν από τον κατάλογο της Δημοτικής Βιβλιοθήκης Ευόσμου Θεσσαλονίκης.

Η επιλογή της Βιβλιοθήκης αυτής κάθε άλλο παρά τυχαία μπορεί να θεωρηθεί κι αυτό διότι πρόκειται για μία τοπική βιβλιοθήκη μικρού μεγέθους που διαθέτει μία σχετικά μικρή συλλογή και η διαδικασία της καταλογογράφησης παρουσιάζει πολλά λάθη στις εγγραφές, αφού πραγματοποιείται από ευκαιριακό προσωπικό (δηλ. συμβασιούχους και φοιτητές πρακτικής άσκησης). Η παραπάνω εικόνα μπορεί να αντικατοπτρίσει σχεδόν το σύνολο των Δημοτικών Βιβλιοθηκών στην Ελλάδα.

Με τη χρήση του καταλόγου της εν λόγω Βιβλιοθήκης επιδιώκεται η απόδοση μίας απλής πρακτικής - συμβολής στο Σύννεφο των Ανοιχτών και Διασυνδεδεμένων Δεδομένων που απευθύνεται σε βιβλιοθηκονόμους κι επιστήμονες της πληροφόρησης χωρίς ιδιαίτερη τεχνογνωσία στις τεχνολογίες του Σημασιολογικού Ιστού. Τα παραπάνω καθιστούν την πρακτική αυτή ευέλικτη και εν δυνάμει εφαρμόσιμη σε οποιαδήποτε Δημοτική Βιβλιοθήκη της Ελλάδας.

## 4.1. Εξαγωγή Δεδομένων Δημοτικής Βιβλιοθήκης Ευόσμου

### 4.1.1. Κατάλογος Βιβλιοθήκης (ΑΒΕΚΤ 5.5)

Αρχικά, στη Δημοτική Βιβλιοθήκη Ευόσμου οι υπηρεσίες του δανεισμού, των κρατήσεων υλικού, οι βιβλιογραφικές αναζητήσεις, η διαδικασία οργάνωσης και διαχείρισης του υλικού, η καταλογογράφηση, η ταξινόμηση, οι εγγραφές νέων μελών καθώς και η εξαγωγή στατιστικών στοιχείων πραγματοποιούνται μέσω του αυτοματοποιημένου συστήματος ΑΒΕΚΤ 5.5 του Εθνικού Κέντρου Τεκμηρίωσης (ΕΚΤ).

Το σύστημα αυτοματισμού βιβλιοθηκών ΑΒΕΚΤ (Αυτοματισμός Βιβλιοθηκών Εθνικού Κέντρου Τεκμηρίωσης) είναι μια εφαρμογή λογισμικού, η οποία έχει αναπτυχθεί εξ' ολοκλήρου από το ΕΚΤ, για την όσο το δυνατόν πιο ολοκληρωμένη αυτοματοποίηση των βιβλιοθηκών και τη διάθεση των βιβλιογραφικών δεδομένων (ΑΒΕΚΤ: Γενικά για το ΑΒΕΚΤ, χ.χ.).

Η ανάπτυξη του ΑΒΕΚΤ σηματοδοτείται το 1986, ενώ έως τις μέρες μας έχουν αναπτυχθεί 5 εκδόσεις του (συγκεκριμένα έως και το ΑΒΕΚΤ 5.6). Το ΑΒΕΚΤ χρησιμοποιείται για την αυτοματοποίηση περισσότερων από 2.500 δημόσιων και ιδιωτικών βιβλιοθηκών στην Ελλάδα, την Κύπρο και το εξωτερικό. Οι βιβλιοθήκες αυτές ανήκουν σε φορείς όπως (ΑΒΕΚΤ: Γενικά για το ΑΒΕΚΤ, χ.χ.):

- πρωτοβάθμια, δευτεροβάθμια και τριτοβάθμια εκπαίδευση,
- δημόσιες / δημοτικές βιβλιοθήκες, μουσεία, αρχαία,
- διάφορες σχολές, ωδεία, επιμορφωτικά κέντρα, κέντρα εκπαίδευσης (ΙΕΚ),
- ερευνητικά ιδρύματα και ινστιτούτα, κέντρα έρευνας, ακαδημίες,
- οργανισμούς κοινωνικής ωφέλειας και άλλους οργανισμούς,
- νοσοκομεία, κέντρα υγείας,
- φιλανθρωπικά ιδρύματα,
- θρησκευτικά ιδρύματα,
- υπουργεία, τοπική αυτοδιοίκηση, πνευματικά κέντρα, συλλόγους, κόμματα,
- ενώσεις, επιμελητήρια, εταιρείες,

Παρακάτω (βλ. εικ. 19) μπορεί να διαπιστώσει κανείς το πλήθος των φορέων που χρησιμοποιούν το ΑΒΕΚΤ ως σύστημα αυτοματοποίησης της συλλογής τους και ταυτόχρονα μπορεί να υποθέσει κανείς πως η πρακτική που παρουσιάζεται στην εν λόγω διπλωματική εργασία είναι δυνατό να εφαρμοστεί στο σύνολο των φορέων αυτών με απλό τρόπο από το ίδιο το προσωπικό που αυτοί απασχολούν. Με την εφαρμογή του παραπάνω θα μπορεί να γίνει λόγος για μία μεγάλη συμβολή των Ελληνικών Βιβλιοθηκών στον κόσμο των Ανοιχτών και Διασυνδεδεμένων Δεδομένων.

Είδος Βιβλιοθήκης*	Πλήθος Βιβλιοθηκών	Ποσοστό
Ακαδημαϊκή	35	5,22%
Δημόσια	45	6,72%
Δημοτική	128	19,10%
Ειδική	143	21,34%
Ειδική. Θρησκευτικό Ίδρυμα	13	1,94%
Ειδική. Μουσείο	6	0,90%
Ειδική. Νομική	14	2,09%
Ειδική. Νοσηλευτικό Ίδρυμα	9	1,34%
Ειδική. Στρατιωτικός Οργανισμός	10	1,50%
Ειδική. Τραπεζικός Οργανισμός	7	1,04%
Σχολική	260	38,81%
Σύνολο	670	100,00%

**Εικόνα 19: Βιβλιοθήκες που διαθέτουν το ABEKT**

#### *4.1.2. Μορφή εξαχώριμου αρχείου δεδομένων*

Το ABEKT υποστηρίζει την καταχώρηση και ενημέρωση των βιβλιογραφικών εγγραφών, βάσει του προτύπου ISO 2709 και των διατάξεων UNIMARC Bibliographic και UNIMARC Authorities (πρότυπα MARC, βλ. υποκεφάλαιο 2.3.1.) (ABEKT: ABEKT 5.6, χ.χ.).

Η εξαγωγή των δεδομένων από το αυτοματοποιημένο σύστημα της Βιβλιοθήκης πραγματοποιήθηκε με μία απλή διαδικασία (με την εντολή «Export») που διαθέτει το ίδιο το ABEKT και εξήχθη ένα αρχείο σε μορφή ISO 2709 UNIMARC, το οποίο είναι ένα πρότυπο ISO<sup>57</sup> (International Organization for Standardization) που αφορά την ανταλλαγή πληροφοριών μεταξύ των βιβλιογραφικών περιγραφών σε UNIMARC (ISO 2709: 2008, 2011).

<sup>57</sup> <http://www.iso.org/iso/home.html>



Η μορφή ISO 2709 αποτελείται από τρία μέρη για κάθε εγγραφή (Lam, 2001). Πιο ειδικά:

- **Ετικέτα Εγγραφής:** Είναι οι πρώτοι 24 χαρακτήρες του αρχείου. Κάθε θέση έχει συγκεκριμένη σημασία, αλλά το μεγαλύτερο μέρος των πληροφοριών της ετικέτας εγγραφής είναι για χρήση από τον υπολογιστή. Οι παραπάνω 24 χαρακτήρες της ετικέτας εγγραφής οποιασδήποτε εγγραφής αντιστοιχούν στις εξής επεξηγήσεις:

Χαρακτήρες	Επεξήγηση	
0-4	Μήκος εγγραφής	
5	Κατάσταση εγγραφής	
6-9	Κώδικας υλοποίησης	
10	Μήκος δείκτη	
11	Μήκος συμβόλων υποπεδίων	
12-16	Βάση αρχής πεδίων δεδομένων	
17-19	Πρόσθετες πληροφορίες	
20	Μήκος του μήκους πεδίου δεδομένων σε κάθε εγγραφή	<b>Χάρτης Ευρετηρίου</b>
21	Μήκος της αρχικής θέσης χαρακτήρων σε κάθε εγγραφή	
22	Μήκος του τμήματος υλοποίησης σε κάθε εγγραφή	
23	Για μελλοντική χρήση	

- **Ευρετήριο Διευθύνσεων:** Αμέσως μετά την ετικέτα εγγραφής είναι ένα σύνολο αριθμών που ονομάζεται ευρετήριο. Αυτό το ευρετήριο παρουσιάζει τις ετικέτες που βρίσκονται στο αρχείο με τη θέση που εμφανίζονται. Το ευρετήριο διευθύνσεων περιλαμβάνει τρία επιμέρους μέρη:

1. Ετικέτα πεδίου (Tag): Αποτελείται από 3 χαρακτήρες και εμπεριέχει το αναγνωριστικό του πεδίου
2. Αριθμός χαρακτήρων πεδίου δεδομένων: Αποτελείται από 4 χαρακτήρες και παρουσιάζει το πλήθος των χαρακτήρων που υπάρχουν σε ένα συγκεκριμένο

πεδίο συνυπολογίζοντας διαχωριστές πεδίων, υποπεδίων, δείκτες αλλά και τα ίδια τα δεδομένα

3. Αρχική θέση χαρακτήρων: Αποτελείται από 5 χαρακτήρες και παρουσιάζει τη θέση που ξεκινά ο πρώτος χαρακτήρας του συγκεκριμένου πεδίου

- **Πεδία Δεδομένων:** Μετά το ευρετήριο διευθύνσεων είναι τα πραγματικά δεδομένα της εγγραφής MARC. Τα πεδία δεδομένων μιας εγγραφής UNIMARC σε μορφή ISO 2709 αποτελείται από δείκτες κι από υποπεδία που χωρίζονται μεταξύ τους μέσω ενός διαχωριστή υποπεδίων ή αλλιώς US (Unit Separator). Αντίστοιχα, το τέλος ενός πεδίου δεδομένων σημαίνεται μέσω ενός διαχωριστή πεδίου ή RS (Record Separator). Ένα πεδίο δεδομένων UNIMARC έχει την εξής γενική δομή στην μορφή ISO 2709:

Δείκτης 1	Δείκτης 2	US	Αναγνωριστικά Υποπεδίου	Δεδομένα	US	Αναγνωριστικά Υποπεδίου	Δεδομένα	RS
-----------	-----------	----	-------------------------	----------	----	-------------------------	----------	----

Η κάθε εγγραφή τερματίζεται από ένα μοναδικό χαρακτήρα GS (Group Separator).

Παρακάτω παρουσιάζεται ένα παράδειγμα της μορφής ISO 2709 του αρχείου της Δημοτικής Βιβλιοθήκης Ευόσμου (βλ. εικ. 20):

Εικόνα 20: Παράδειγμα δεδομένων Βιβλιοθήκης σε μορφή ISO 2709

### Ετικέτα εγγραφής παραδείγματος

Η ετικέτα εγγραφής του παραπάνω παραδείγματος του αρχείου ISO 2709 της Δημοτικής Βιβλιοθήκης Ευόσμου είναι η εξής (βλ. εικ. 20):

00641nam \_ 2200169 \_ \_ \_ 450 \_

Η ετικέτα εγγραφής του παραδείγματος μεταφράζεται ως εξής:

Θέση χαρακτήρα	Τιμή	Επεξήγηση
0	0	Το πλήθος των χαρακτήρων της εγγραφής είναι 641
1	0	
2	6	
3	4	
4	1	
5	n	Πρόκειται για μία νέα εγγραφή
6	a	Γλωσσικό υλικό, έντυπο
7	m	Μονογραφία, βιβλίο
8		Δεν έχει προσδιοριστεί καμία ιεραρχική δομή
9		Απροσδιόριστο
10	2	Μήκος δείκτη
11	2	Μήκος συμβόλων υποπεδίων
12	0	Το πλήθος των χαρακτήρων στην ετικέτα εγγραφής και το ευρετήριο είναι ακριβώς 169 χαρακτήρες. Το πρώτα πεδίο δεδομένων συναντάται στη θέση 169.
13	0	
14	1	
15	6	
16	9	
17		Απροσδιόριστο
18		Απροσδιόριστο
19		Απροσδιόριστο
20	4	Το μήκος πεδίου δεδομένων της εγγραφής είναι 4 χαρακτήρες
21	5	Το μήκος της αρχικής θέσης χαρακτήρων της εγγραφής είναι 5

22	0	Το μήκος του τμήματος υλοποίησης είναι 0 χαρακτήρες
23		Απροσδιόριστο

### Ευρετήριο διεθύνσεων παραδείγματος

Το ευρετήριο διεθύνσεων του πρώτου πεδίου δεδομένων του παραπάνω παραδείγματος του αρχείου ISO 2709 της Δημοτικής Βιβλιοθήκης Ευόσμου είναι η εξής (βλ. εικ. 20):

001000200000

Αναλυτικά, πρόκειται για το πεδίο 001, που είναι το πεδίο της "Ταυτότητας Εγγραφής" στο UNIMARC και περιέχει 2 χαρακτήρες ενώ ξεκινάει από τη θέση 0.

### Πεδίο δεδομένων παραδείγματος

1 USaΤα επιά θαύματα του κόσμουRS

Το παραπάνω μέρος του παραδείγματος αντιστοιχεί στο πεδίο 200 (βλ. εικ. 20) (βλ. υποκεφ. 2.1.2.1.) μπορεί να παρατηρήσει κανείς πως στον δείκτη 1 η εγγραφή αυτή έχει την τιμή «1» που σημαίνει πως πρόκειται για τον κύριο τίτλο της εγγραφής. Ο δείκτης δύο είναι μηδενικός και έχει ένα υποπεδίο το \$a που εμπεριέχει τον τίτλο του βιβλίου σε ρέοντα λόγο.

Το σύνολο των δεδομένων της Δημοτικής Βιβλιοθήκης Ευόσμου αντιστοιχούν ακριβώς σε 18.907 βιβλιογραφικές εγγραφές έως τον Αύγουστο του 2014 που εξήλθαν από τον κατάλογο της Βιβλιοθήκης, γεγονός που επιβεβαιώνει την άποψη που ειπώθηκε παραπάνω πως η Βιβλιοθήκη έχει μία σχετικά μικρή συλλογή. Ωστόσο, αυτό όπως θα παρουσιαστεί και στις επόμενες ενότητες του κεφαλαίου, δεν είναι ανασταλτικός παράγοντας διασύνδεσης των δεδομένων της με άλλα σύνολα δεδομένων. Αντιθέτως, το παραπάνω μπορεί να θεωρηθεί θετικός παράγοντας από την άποψη πως οι περισσότερες Δημοτικές Βιβλιοθήκες της Ελλάδας εμπεριέχουν έναν μικρότερο αριθμό δεδομένων στους καταλόγους τους σε σύγκριση με τις αντίστοιχες βιβλιοθήκες του εξωτερικού κι έτσι καθιστούν το δείγμα αυτό αντιπροσωπευτικό και την προτεινόμενη πρακτική εύκολα υλοποιήσιμη από τις αντίστοιχες βιβλιοθήκες τις Ελλάδος.

## 4.2. Σημασιολογική Διαχείριση Δεδομένων

Η εξαγωγή των δεδομένων από τον κατάλογο της Βιβλιοθήκης αποτέλεσε το πρώτο βήμα για το «Άνοιγμα» και τη «Διασύνδεση» των δεδομένων αυτών, ωστόσο, η μορφή ISO 2907 UNIMARC δεν είναι εύκολα επεξεργάσιμη από συστήματα εκτός των Βιβλιοθηκών κι επομένως, καταστήθηκε αναγκαία η μετατροπή τους σε μορφή RDF.

Όπως αναλύθηκε και στο κεφάλαιο 3 (βλ. κεφ. 3), οι βιβλιοθήκες του εξωτερικού χρησιμοποίησαν διάφορες τεχνικές για τη μετατροπή των δεδομένων που εξήλθαν από τους καταλόγους τους από τα πρότυπα MARC (κυρίως από το MARC21) σε RDF μορφή. Πιο ειδικά, η Βιβλιοθήκη της Βρετανίας χρησιμοποίησε τη γλώσσα XSLT (βλ. υποκεφ. 3.1.), η Εθνική Βιβλιοθήκη της Γερμανίας το Metafacture<sup>58</sup> (βλ. υποκεφ. 3.2.) και η Εθνική Βιβλιοθήκη της Ισπανίας την εφαρμογή MARiMba<sup>59</sup> (βλ. υποκεφ. 3.3.).

Στην προσπάθεια εύρεσης της καλύτερης πρακτικής εφαρμογής για τη μετατροπή των δεδομένων της Δημοτικής Βιβλιοθήκης Ευόσμου αξιολογήθηκαν όλες οι παραπάνω εφαρμογές βάσει συγκεκριμένων κριτηρίων. Τα κριτήρια αυτά είναι:

1. Η ευκολία χρήσης τους από βιβλιοθηκονόμους με ελάχιστη τεχνογνωσία
2. Οι μορφές των αρχείων εισόδου
3. Η ανοιχτή και δωρεάν διάθεση τους
4. Η υποστήριξη της ελληνικής γλώσσας

Τα συμπεράσματα της αξιολόγησης συνοψίζονται στα εξής:

- Η Βιβλιοθήκη της Βρετανίας δε χρησιμοποίησε κάποιο εργαλείο ή εφαρμογή, καθώς ακολούθησε την πρακτική που χρησιμοποίησε η Εθνική Βιβλιοθήκη της Σουηδίας (Kungliga Biblioteket) για να μετατρέψει τα δεδομένα του καταλόγου της, LIBRIS<sup>60</sup> σε RDF μορφή. Οι δύο αυτές Βιβλιοθήκες αξιοποίησαν την γλώσσα XSLT για να μετατρέψουν τα βιβλιογραφικά δεδομένα σε RDF/ XML μορφή, ωστόσο, η διαδικασία μετατροπής δεν τεκμηριώνεται βήμα – βήμα από τις βιβλιοθήκες και κρίνεται δύσκολο οι βιβλιοθηκονόμοι, χωρίς ιδιαίτερη τεχνογνωσία, να καταφέρουν με ευκολία να υλοποιήσουν τη διαδικασία αυτή. Ίσως η δημιουργία ενός εγχειριδίου μετατροπής των δεδομένων να βοηθούσε το βιβλιοθηκονομικό κοινό να εφαρμόσει την πρακτική αυτή.
- Το εργαλείο Metafacture<sup>61</sup> που χρησιμοποίησε η Εθνική Βιβλιοθήκη της Γερμανίας είναι ένα ευέλικτο, εύχρηστο, ανοιχτού κώδικα εργαλείο βασισμένο σε Java, το οποίο μπορεί να χρησιμοποιηθεί για την μετατροπή εύκολα από

<sup>58</sup> <https://github.com/culturegraph/metafacture-core>

<sup>59</sup> <http://mayor2.dia.fi.upm.es/oeg-upm/index.php/en/technologies/228-marimba>

<sup>60</sup> <http://libris.kb.se/>

<sup>61</sup> <https://github.com/culturegraph/metafacture-core>

βιβλιοθηκονόμους κι επιστήμονες της πληροφόρησης. Ωστόσο, κατά τη δοκιμή του εν λόγω εργαλείου προέκυψαν προβλήματα με την ανάγνωση της ελληνικής γλώσσας και οι τριπλέτες που δημιουργήθηκαν δεν ήταν αναγνώσιμες. Πραγματοποιήθηκε προσπάθεια αποφυγής του εν λόγω προβλήματος με την αλλαγή της κωδικοποίησης του αρχείου σε UTF-8, ωστόσο η εμφάνιση των ελληνικών χαρακτήρων με τη μορφή συμβόλων δεν επιλύθηκε.

- Το MARiMba<sup>62</sup> είναι ένα εργαλείο που δε χρησιμοποιήθηκε εφόσον δε διατίθεται από τους φορείς δημιουργίας του προς ελεύθερη χρήση. Έπειτα από επικοινωνία με τους υπεύθυνους του εργαλείου διαπιστώθηκε πως η χρήση του μπορεί να πραγματοποιηθεί μόνο στα στενά πλαίσια του Πανεπιστημίου της Ισπανίας με τη συνεργασία των δημιουργών του. Βάσει των παραπάνω δε μπορεί να αποδοθεί μία ξεκάθαρη εικόνα για το αν το εργαλείο αυτό θα εξυπηρετούσε ή όχι με επιτυχία τη μετατροπή των δεδομένων της Δημοτικής Βιβλιοθήκης του Ευόσμου. Ωστόσο, από την αντίστοιχη βιβλιογραφία προκύπτει πως το εργαλείο αυτό είναι εύκολο στη χρήση και δέχεται αρχεία σε μορφή ISO 2709, γεγονός που θα εξυπηρετούσε τις ανάγκες της πρακτικής αυτής που αναλύεται στην παρούσα διπλωματική.

Συμπερασματικά, η εφαρμογή που χρησιμοποιήθηκε για τη μετατροπή των βιβλιογραφικών δεδομένων της Δημοτικής Βιβλιοθήκης του Ευόσμου είναι εκείνη που δημιουργήθηκε από τον Έλληνα μεταπτυχιακό φοιτητή του Αριστοτελείου Πανεπιστημίου, Σωτήρη Καραμπατάκη (Καραμπατάκης, 2013) για τη μετατροπή των δεδομένων της Δημόσιας Βιβλιοθήκης της Βέροιας. Πέραν των παραπάνω μειονεκτημάτων που προέκυψαν από την έρευνα που αφορούσε τις άλλες εφαρμογές που χρησιμοποιήθηκαν από τις βιβλιοθήκες του εξωτερικού, το συγκεκριμένο εργαλείο που επιλέχθηκε τελικά παρουσίασε κι επιπλέον θετικά στοιχεία.

Παρακάτω παρουσιάζονται επιγραμματικά τα πλεονεκτήματα που παρατηρήθηκαν στην εν λόγω εφαρμογή:

- Αναγνωρίζει τους ελληνικούς χαρακτήρες
- Είναι ανοιχτού κώδικα (open source), που σημαίνει ότι μπορεί να παραμετροποιηθεί με τέτοιο τρόπο ώστε να προσαρμοστεί για την επεξεργασία οποιωνδήποτε δεδομένων ή λάθος δεδομένων (error data) προκύψουν κατά τη διαδικασία μετατροπής
- Τα αρχεία εισόδου δύναται να είναι σε μορφή ISO 2709, δηλαδή την εξαγωγή μορφή του αρχείου που προήλθε από τον κατάλογο της Δημοτικής Βιβλιοθήκης Ευόσμου
- Είναι σχεδιασμένο για να επεξεργάζεται δεδομένα που είναι περιγεγραμμένα από το πρότυπο UNIMARC, όπως τα δεδομένα της Δημόσιας Βιβλιοθήκης Βέροιας αλλά και της Δημοτικής Βιβλιοθήκης Ευόσμου

---

<sup>62</sup> <http://mayor2.dia.fi.upm.es/oeg-upm/index.php/en/technologies/228-marimba>

- Εξάγει αρχεία n3 (notation3), μία γλώσσα η οποία είναι η πιο συμπαγής και ευανάγνωστη εναλλακτική λύση της RDF/ XML σύνταξης, αλλά ταυτόχρονα διακατέχεται κι από μεγαλύτερη εκφραστικότητα από μορφές εξαγωγής αρχείων όπως η Turtle και η JSON (Berners - Lee & Connolly, 2011).
- Έχει δημιουργηθεί για αρχεία που προέρχονται από καταλόγους Βιβλιοθηκών που χρησιμοποιούν το αυτοματοποιημένο σύστημα ABEKT (βλ. υποκεφ. 4.1.1.)

Εντούτοις, δε μπορούν παρά να αναφερθούν και τα μειονεκτήματα που διαπιστώθηκαν κατά τη χρήση της εφαρμογής αυτής. Πιο ειδικά, πρόκειται για μία εφαρμογή που δεν έχει διεπαφή χρήστη αλλά λειτουργεί μέσω κονσόλας. Ωστόσο οι οδηγίες «τρεξιματος» είναι απόλυτα σαφείς και με τον τρόπο αυτό μπορεί να χρησιμοποιηθεί κι από βιβλιοθηκονόμους που δεν έχουν τεχνογνωσία στις τεχνολογίες του Σημασιολογικού Ιστού

Η δομή του κώδικα μπορεί να διαχωριστεί σε τρία βασικά στάδια όπως έχουν οριστεί από τον ίδιο το δημιουργό και παρουσιάζονται παρακάτω (Καραμπατάκης, 2013):

- Αρχικοποίηση Μοντέλου Δεδομένων: Όπου πραγματοποιείται η εισαγωγή στη μνήμη του μοντέλου δεδομένων που θα χρησιμοποιηθεί
- Ανάγνωση Εγγραφής: Ανάγνωση εγγραφής MARC από το αρχείο που έχει εισαχθεί σε μορφή ISO 2709
- Μετατροπή σε RDF: Δημιουργία RDF τριπλετών και εγγραφή στο εξαγώμενο αρχείο υπό την μορφή .n3

Η επεξήγηση των βασικών σταδίων για τα δεδομένα της Δημοτικής Βιβλιοθήκης Ευόσμου παρουσιάζονται στα επόμενα υποκεφάλαια.

### 4.2.1. Λεξιλόγια εφαρμογής

Η εφαρμογή που χρησιμοποιήθηκε για τη μετατροπή των δεδομένων της Δημοτικής Βιβλιοθήκης Ευόσμου χρησιμοποιεί τις εξής οντολογίες για την περιγραφή των οντοτήτων:

- Η οντολογία BIBO για την περιγραφή των βιβλιογραφικών δεδομένων. Ανάλογα με τον τύπο της, η εκάστοτε εγγραφή αντιστοιχίζεται σε κάποια κλάση μέσω RDF συνδέσμου του τύπου `rdf:type`. Πιο ειδικά, οι αντιστοιχήσεις αυτές είναι οι εξής:

Τύπος	Κλάση
Βιβλία	bibo: Book
Έγγραφα	bibo: Document
Άρθρα	bibo: Article
Ταινίες	bibo: Film
Περιοδικά	bibo: Magazine
Ηχογραφήσεις και CDs	bibo: AudioDocument

Οπτικοακουστικό υλικό	bibo:AudioVisualDocument
Γραφικά	bibo:Image
Τεχνολογικός εξοπλισμός	Owl:Thing
Εφημερίδες	bibo: Newspaper

- Η οντολογία FOAF για την περιγραφή των πνευματικά υπεύθυνων ή των συντελεστών της εγγραφής. Πιο συγκεκριμένα, κάθε πνευματικά υπεύθυνος για την εγγραφή (π.χ. συγγραφέας, εκδότης, συνθέτης, σκηνοθέτης κ.α.) αντιστοιχίζεται στην κλάση foaf:Person αν είναι φυσικό πρόσωπο και foaf:Organization αν είναι συλλογικό όργανο.
- Η οντολογία SKOS για τις θεματικές επικεφαλίδες ή αλλιώς τις θεματικές κατηγορίες που ανήκει η κάθε εγγραφή. Σαν κλάση για το θέμα αποδίδεται η κλάση skos:Concept ενώ υπάρχουν κι εσωτερικοί σύνδεσμοι προς άλλες σχετικές κατηγορίες με σχέσεις του τύπου skos:narrower. Αν κάποια θεματική κατηγορία σχετίζεται με κάποιο άτομο, δημιουργείται σχέση μεταξύ του θέματος και του ατόμου με τη μορφή συνδέσμου rdfs:seeAlso.

#### 4.2.2. Αρχικοποίηση Μοντέλου Δεδομένων

Αρχικά, η εφαρμογή αυτή πραγματοποιεί μία πρώτη επεξεργασία στο αρχείο εισόδου και καταγράφονται όλοι οι προσδιορισμοί ή αλλιώς τύποι των εγγραφών όπως έχουν οριστεί στο πεδίο 200\_\$b της MARC εγγραφής (βλ. υποκ. 2.1.1.). Αυτό το υποπεδίο UNIMARC είναι ο Γενικός Προσδιορισμός Υλικού του κειμενικού υλικού και περιέχει έναν τυποποιημένο όρο που υποδεικνύει την ευρεία κατηγορία, στην οποία ανήκει το στοιχείο. Ο τυποποιημένος αυτός όρος προέρχεται από το πρότυπο MARC (IFLA: 200 Title and statement of responsibility, χ.χ.).

Στη συνέχεια παράγονται τα αρχεία αρχικοποίησης του μοντέλου δεδομένων classes.txt και mappings.txt. Ειδικότερα, το μοντέλο δεδομένων ολοκληρώνεται με τη δημιουργία τριών ξεχωριστών αρχείων (Καραμπατάκης, 2013):

- **classes.txt:** περιέχει τις αντιστοιχήσεις των κλάσεων των βιβλιογραφικών εγγραφών. Για την περιγραφή των βιβλιογραφικών εγγραφών σύμφωνα με το μοντέλο της RDF χρησιμοποιήθηκαν κλάσεις από την οντολογία BIBO για να προσδιορίσουν τους τυποποιημένους όρους του προτύπου MARC που αφορούν τον γενικό προσδιοριστικό όρο του κειμενικού υλικού και ανάλογα με τον τύπο της, η κάθε εγγραφή αντιστοιχίζεται σε κάποια κλάση μέσω RDF συνδέσμου του τύπου rdf:type (βλ. υποκ. 4.2.1.).
- **mappings.txt:** περιέχει τις αντιστοιχήσεις των πεδίων και των υποπεδίων του UNIMARC προς κάποια ιδιότητα από τις οντολογίες που χρησιμοποιούνται στην εφαρμογή (BIBO, FOAF κ.α.) (βλ. υποκ. 4.2.1.).



- **prefixes.txt:** περιέχει τις συντμήσεις των οντολογιών που χρησιμοποιούνται (βλ. υποκ. 4.2.1.).

### 4.2.3. Ανάγνωση Εγγραφής

Η εφαρμογή «διαβάζει» κάθε φορά μία εγγραφή του αρχείου εισόδου μέχρι να φτάσει και στην τελευταία εγγραφή, δηλαδή στο τέλος του αρχείου. Σε περίπτωση που η ανάγνωση της εκάστοτε εγγραφής είναι έγκυρη, προχωρά στη μετατροπή σε RDF κι αυτή η διαδικασία επαναλαμβάνεται έως ότου τελειώσουν οι εγγραφές. Πώς λειτουργεί όμως αυτή η διαδικασία; Η εκάστοτε εγγραφή ελέγχεται ως προς τα τρία μέρη της μορφής ISO 2709 που αναφέρθηκαν παραπάνω (βλ. 4.1.2.), δηλαδή την ετικέτα εγγραφής, του ευρετηρίου διευθύνσεων και του πεδίου δεδομένων.

Ειδικότερα, η Ετικέτα Εγγραφής ελέγχεται ως προς το αν το δηλωθέν πλήθος χαρακτήρων έρχεται σε αντιστοιχία με τον πραγματικό αριθμό χαρακτήρων της εγγραφής. Σε περίπτωση σφάλματος καταγράφεται η εγγραφή σε ένα συγκεκριμένο αρχείο καταγραφής σφαλμάτων και πραγματοποιείται μετάβαση στην επόμενη εγγραφή. Στη συνέχεια, η εφαρμογή πραγματοποιεί ανάγνωση στο Ευρετήριο Διευθύνσεων, το οποίο χρησιμοποιείται για την εξαγωγή των πληροφοριών για την ταυτότητα του πεδίου, την θέση έναρξης υποπεδίου και το πλήθος των χαρακτήρων που δηλώνονται. Στο μέρος αυτό ελέγχεται επίσης η εγκυρότητα της εγγραφής και σε περίπτωση σφάλματος, υλοποιείται η καταγραφή στο αρχείο σφαλμάτων και η μετάβαση στην επόμενη εγγραφή. Στο τέλος, η εφαρμογή «διαβάζει» το πεδίο δεδομένων κάθε εγγραφής, που σημαίνει πως πραγματοποιεί ανάγνωση στα δεδομένα που εμπεριέχονται στα πεδία και στα υποπεδία κάθε εγγραφής. Για κάθε εγγραφή δημιουργείται μία δομή, η οποία αποτελείται από επιμέρους δομές (Καραμπατάκης, 2013):

- **Δομή record:** περιέχει όλα τα δεδομένα που έχουν εξαχθεί από την εκάστοτε εγγραφή.

```
struct record {
    marc_field field[1000]; //Τα πεδία της εγγραφής ξεχωριστά ανά Πεδίο
    string leader; //Η Έκταση Εγγραφής
    string directory; //Το Ευρετήριο
    string field_container; //Ολόκληρο το πεδίο δεδομένων
    int no_fields; //Ο συνολικός αριθμός πεδίων που υπάρχουν στην Εγγραφή
    string type; //Η κλάση της εγγραφής όπως εξάγεται από το leader[7] ή το υποπεδίο 200$b
    int recordID; //Το αναγνωριστικό της εγγραφής, η τιμή του πεδίου 001
};
```

Εικόνα 21: Ανάλυση δομής record

- **Δομή marc\_field:** αποθηκεύεται κάθε πεδίο UNIMARC της εκάστοτε εγγραφής.

```

struct marc_field {
    string tag; //Περιέχει την Ετικία Πεδίου
    int start; //Η θέση αρχής του πεδίου στο Πεδίο Δεδομένων
    int length; //Το πλήθος των χαρακτήρων του πεδίου
    char ind1; //Ο Δείκτης 1 του πεδίου
    char ind2; //Ο Δείκτης 2 του πεδίου
    string data; //Το περιεχόμενο του πεδίου χωρίς επεξεργασία
    tags subfield; //Συλλογή δομών υποπεδίων που υπάρχουν στο πεδίο
    int subfn; //Συνολικός αριθμός υποπεδίων στον πεδίο
    bool processed; //TRUE αν το πεδίο έχει επεξεργαστεί αλλιώς FALSE
};

```

Εικόνα 22: Ανάλυση δομής marc\_field

- **Δομή tags:** αποθηκεύεται κάθε υποπεδίο UNIMARC της εκάστοτε εγγραφής.

```

struct tags {
    string a[10];
    int an;
    string b[50];
    int bn;
    string c[10];
    int cn;
    .
    .
    .
    string nosub;
    int count;
};

```

Εικόνα 23: Ανάλυση δομής tags

#### 4.2.4. Μετατροπή σε RDF

Τη διαδικασία ανάγνωσης της εγγραφής ακολουθεί η μετατροπή των εγγραφών από το πρότυπο MARC σε μορφή ISO 2709. Συγκεκριμένα, η διαδικασία αυτή δέχεται ως ορίσματα τη δομή record (βλ. υποκεφ. 4.2.2.), το URI της εγγραφής και τους πίνακες με τις αντιστοιχίσεις που έχουν ορισθεί από το δημιουργό του εργαλείου.

Το URI της εκάστοτε εγγραφής δημιουργείται από το prefix που έχει δηλωθεί και από την τιμή του πεδίου 001 της κάθε MARC εγγραφής. Το πεδίο 001 είναι ο αριθμός ελέγχου που αποδίδεται από τον οργανισμό που δημιουργεί, χρησιμοποιεί ή διαθέτει την εγγραφή και δημιουργείται αυτόματα από το ίδιο το αυτοματοποιημένο σύστημα, στην εν λόγω περίπτωση από το ΑΒΕΚΤ 5.1. (Τσουκαλά & Κοντού, 2015, σ. 15). Με τον τρόπο αυτό η εφαρμογή αυτή διασφαλίζει τη μοναδικότητα των URI.

Η τιμή του prefix library: καθορίζεται από την τιμή που έχει δηλωθεί στο αρχείο αρχικοποίησης prefixes.txt (Καραμπατάκης, 2013, σ. 72). Η Δημοτική Βιβλιοθήκη Ευόσμου αντιστοιχεί στην τιμή <http://83.212.169.101/resource/>. Για παράδειγμα, το URI μιας εγγραφής με τιμή πεδίου 001: «library:record1», παίρνει τη μοναδική τιμή <http://83.212.169.101/resource/record1>.

Μετά το τέλος της διαδικασίας, παράγονται τριπλέτες που εγγράφονται στο αρχείο εξόδου που είναι σε μορφή n3 (notation3). Η εφαρμογή αυτή δημιουργεί τις τριπλέτες για κάθε εγγραφή και συνεχίζει στην επόμενη έως την εξάντληση τους. Τα τρία συστατικά της εκάστοτε τριπλέτας αντιστοιχούν για κάθε εγγραφή στο υποκείμενο, το κατηγορήμα και το αντικείμενο. Για παράδειγμα, η εγγραφή record5 (library:record5) (το υποκείμενο) έχει τίτλο (bibo:title) (το κατηγορήμα) «Ο Γιος του Ήλιου» (η τιμή σε string) (το αντικείμενο) (βλ. εικ. 24). Παρακάτω μπορεί να δει κανείς ένα παράδειγμα του αρχείου τριπλετών για μία εγγραφή και πιο ειδικά για την εγγραφή «library:record5»:

```

library:record5 rdf:type      bibo:Document .
library:record5 bibo:record  "5" .
library:record5 dcterms:created "2002-10-18"^^xsd:date .
library:record5 bibo:title    "Ο ΓΙΟΣ ΤΟΥ ΗΛΙΟΥ" .
library:record5 rdfs:label    "Ο ΓΙΟΣ ΤΟΥ ΗΛΙΟΥ" .
library:place3  rdf:type      schema:Place .
library:place3  rdfs:label    "ΑΘΗΝΑ" .
library:record5 rdvocab:placeOfPublication library:place3 .
library:publisher5 rdf:type    foaf:Organization .
library:publisher5 rdfs:label  "ΚΟΛΛΑΡΟΣ" .
library:record5 dcterms:publisher library:publisher5 .
library:record5 dcterms:extent "431 σ." .
library:subject5  rdf:type      skos:Concept .
library:subject5  skos:prefLabel "ΕΛΛΗΝΙΚΗ ΛΟΓΟΤΕΧΝΙΑ" .
library:subject5  rdfs:label    "ΕΛΛΗΝΙΚΗ ΛΟΓΟΤΕΧΝΙΑ" .
library:record5 dc:subject library:subject5 .
library:subject7  rdf:type      skos:Concept .
library:subject7  skos:prefLabel " ΕΛΛΗΝΙΚΗ ΛΟΓΟΤΕΧΝΙΑ" .
library:subject7  rdfs:label    " ΕΛΛΗΝΙΚΗ ΛΟΓΟΤΕΧΝΙΑ" .
library:subject7  skos:narrower library:subject5 .
library:record5 dc:contributor library:agent7 .
library:agent7    rdf:type      foaf:Person .
library:agent7    rdfs:label    "ΑΘΑΝΑΣΙΑΔΗΣ, ΤΑΣΟΣ" .
library:agent7    foaf:name     "ΑΘΑΝΑΣΙΑΔΗΣ ΤΑΣΟΣ" .
library:agent7    foaf:givenName "ΤΑΣΟΣ" .
library:agent7    foaf:surname  "ΑΘΑΝΑΣΙΑΔΗΣ" .
library:record5 dc:creator library:library1 .
library:record5 dcterms:created "20021018" .
library:library1  rdf:type      foaf:Organization .
library:library1  rdfs:label    "Δημοτική Βιβλιοθήκη Ευόσμου." .
library:record5 dc:creator library:library1 .
library:record5 dcterms:created "20021018" .
library:library1  rdf:type      foaf:Organization .
library:library1  rdfs:label    "Δημοτική Βιβλιοθήκη Ευόσμου." .

```

**Εικόνα 24: Παράδειγμα εγγραφής στο αρχείο n3 της Δημοτικής Βιβλιοθήκης Ευόσμου**

Το αρχείο του καταλόγου της Δημοτικής Βιβλιοθήκης του Ευόσμου μετατράπηκε σε ένα αρχείο που εμπεριέχει συνολικά 354.525 τριπλέτες που περιγράφουν 9.293 εγγραφές.

## 4.3. “Άνοιγμα” Δεδομένων

Μετά τη μετατροπή του αρχείου από μορφή MARC σε μορφή n3, οι τριπλέτες που εμπεριέχονταν σε αυτό δύνανται να δημοσιοποιηθούν ώστε να προκύψει το «άνοιγμα» των δεδομένων της βιβλιοθήκης. Για την δημοσιοποίηση των τριπλετών ήταν απαραίτητο ένα αποθετήριο φιλοξενίας τριπλετών. Στο υποκεφάλαιο αυτό παρουσιάζονται τα χαρακτηριστικά του εικονικού μηχανήματος (Virtual Machine) που φιλοξενεί το αποθετήριο και η διαδικασία επιλογής του κατάλληλου αποθετηρίου τριπλετών (triplestore) που εξυπηρετούσε τους σκοπούς και στόχους της εν λόγω διπλωματικής.

### 4.3.1. Εικονικό Μηχάνημα (Virtual Machine)

Το Εικονικό Μηχάνημα (Virtual Machine) στο οποίο υλοποιήθηκε η δημοσιοποίηση των τριπλετών της Δημοτικής Βιβλιοθήκης Ευόσμου έχει τα εξής χαρακτηριστικά, τα οποία θεωρήθηκαν κατάλληλα για την υλοποίηση της προτεινόμενης πρακτικής:

<b>Cluster (Ομάδα Ιδρυμάτων)</b>	Υπουργείο Παιδείας
<b>Λειτουργικό Σύστημα</b>	Ubuntu 14.04 LTS
<b>Μνήμη</b>	4.0 GB
<b>Επεξεργαστές</b>	2
<b>Thread(s) per core</b>	1
<b>ID Προμηθευτή</b>	Genuine Intel
<b>Σκληρός Δίσκος</b>	50.0 GB
<b>Διεύθυνση</b>	<a href="http://83.212.169.101">http://83.212.169.101</a>

### 4.3.2. Εξυπηρετητής OpenLink Virtuoso

Για την επιλογή του αποθετηρίου των τριπλετών (triplestore) πραγματοποιήθηκε μία έρευνα που κατέληξε σε σύγκριση πολύ γνωστών triplestores (αποθετήρια τριπλετών) (βλ. υποενοτ. 2.4.2.) και τελικά στην επιλογή του OpenLink Virtuoso για τους εξής λόγους:

- ✓ Είναι ανοιχτού κώδικα κι επιδέχεται οποιασδήποτε παραμετροποίησης

- ✓ Χρησιμοποιείται από γνωστούς φορείς ως αποθετήριο τριπλετών (Dbpedia, BNF, Δημόσια Βιβλιοθήκη Βέροιας), γεγονός που το καθιστά το πιο διαδεδομένο από όλα τα υπόλοιπα
- ✓ Διαθέτει SPARQL Endpoint
- ✓ Η διαχείριση του είναι εύκολη και δεν απαιτεί γνώσεις προγραμματισμού, με τον τρόπο αυτόν καθιστά τη διαχείριση του από βιβλιοθηκονόμους απόλυτα δυνατή
- ✓ Διατίθεται δωρεάν
- ✓ Υποστηρίζει την ελληνική γλώσσα

Η έκδοση που εγκαταστάθηκε στο εικονικό μηχάνημα ήταν η 6.1. (βλ. εικ. 25) και η μοναδική αλλαγή που χρειάστηκε να πραγματοποιηθεί ήταν η αλλαγή του κωδικού του διαχειριστή:



**Εικόνα 25: Εγκατάσταση OpenLink Virtuoso 6.1.**

#### *4.3.2.1. Φόρτωση αρχείου τριπλετών*

Αρχικά πραγματοποιήθηκε προσπάθεια φόρτωσης του αρχείου τριπλετών στη μορφή .n3, που είχε το εξαγωγίμο αρχείο των δεδομένων της Δημοτικής Βιβλιοθήκης Ευόσμου. Για την επιτυχή ολοκλήρωση της φόρτωσης χρησιμοποιήθηκε το WinSCP<sup>63</sup>. Μετά τη μεταφορά του αρχείου .n3 από τον τοπικό στον απομακρυσμένο υπολογιστή, φορτώθηκαν οι τριπλέτες στον OpenLink Virtuoso από την επιλογή Quad Store Upload που διαθέτει στο μενού του, αφού πρώτα διαχωρίστηκε σε επιμέρους δύο αρχεία (λόγω του μεγάλου όγκου) κι αφού μετατράπηκε η κωδικοποίηση του αρχείου από ANSI σε UTF-8 (για αναγνώριση των ελληνικών χαρακτήρων). Η φόρτωση τριπλετών πραγματοποιήθηκε σ' έναν γράφο που του δόθηκε το όνομα <http://www.libevo.gr>.

<sup>63</sup> <https://winscp.net/eng/download.php>



**Εικόνα 26: Φόρτωση τριπλετών από το μενού του Virtuoso**

Έπειτα από τη διαδικασία φόρτωσης, τα δεδομένα της Δημοτικής Βιβλιοθήκης Ευόσμου είναι διαθέσιμα για υλοποίηση ερωτημάτων από το τελικό σημείο ερωτημάτων SPARQL που διατίθεται στη διεύθυνση <http://83.212.169.101:8890/sparql> (βλ. εικ. 27). Επομένως, όπως είναι αντιληπτό, τα δεδομένα της Δημοτικής Βιβλιοθήκης Ευόσμου ανήκουν πλέον στα Ανοιχτά Δεδομένα, ωστόσο, δεν ανήκουν ακόμα στα Ανοιχτά και Διασυνδεδεμένα Δεδομένα.



**Εικόνα 27: Τελικό Σημείο SPARQL Δημοτικής Βιβλιοθήκης Ευόσμου**

#### 4.3.2.2. Αποαφήνιση URI

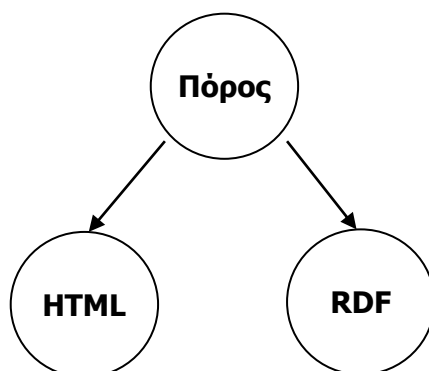
Για την προτεινόμενη πρακτική υιοθετήθηκαν τα 303 URIs (βλ. υποκ. 2.4.1.) για τους εξής λόγους:

- Τα δεδομένα της Δημοτικής Βιβλιοθήκης Ευόσμου αποτελούν ένα μεγάλο σύνολο δεδομένων καθώς εμπεριέχουν 354.525 τριπλέτες (βλ. υποκ. 4.3.2.).
- Είναι πολύ ευέλικτα, διότι ο στόχος ανακατεύθυνσης μπορεί να ρυθμιστεί ξεχωριστά για κάθε πόρο. Θα μπορούσε να υπάρξει ένα περιγραφικό έγγραφο για κάθε πόρο, ή ένα μεγάλο έγγραφο για όλους αυτούς, ή ακόμα κι ένας συνδυασμός των παραπάνω.
- Η επιλεγμένη στρατηγική URI είναι δυνατό να αλλάξει με εύκολο τρόπο ακόμα και μετά την εφαρμογή (από 303 URI σε Hash URI).

Για να επιτευχθεί η υιοθέτηση των 303 URIs πραγματοποιήθηκαν οι κατάλληλες ρυθμίσεις στον Apache και πιο ειδικά, ενεργοποιήθηκε το module «mod\_rewrite» και η παραμετροποίηση των σχετικών εντολών στο αρχείο «.ht access» που βρίσκεται στον κεντρικό φάκελο του εξυπηρετητή.

Για να γίνουν όλα τα παραπάνω πιο κατανοητά. Έστω πως ένας client ζητάει τον πόρο «<http://83.212.169.101/resource/agent353>» που αντιστοιχεί στον συγγραφέα «Μενέλαο Λουντέμη». Τη στιγμή εκείνη ο Apache αναλύει το GET Request που αποστέλλει ο client, που στην προκειμένη περίπτωση είναι χρήστης και διαβάζει το Accept Header. Εφόσον, πρόκειται για χρήστη, πραγματοποιείται 303 ανακατεύθυνση στην εξής διεύθυνση «<http://83.212.169.101/page/agent353>». Η σελίδα αυτή που επιστρέφεται στον χρήστη δεν είναι τίποτα άλλο από html ιστοσελίδα ικανή να αναγνωστεί από έναν χρήστη.

Με τον ίδιο τρόπο, αν ο client είναι μηχανή και ζητήσει περιεχόμενο σε μορφή RDF/XML, ο εξυπηρετητής θα πραγματοποιήσει 303 ανακατεύθυνση στη διεύθυνση «<http://83.212.169.101/data/agent353.rdf>» και θα προσφέρει το περιεχόμενο σε RDF/XML μορφή. Σχηματικά, η 303 ανακατεύθυνση παρουσιάζεται ως εξής:



**Εικόνα 28: Σχηματική απεικόνιση 303 ανακατεύθυνσης URI**

Έπειτα από τη σχηματική απεικόνιση παρουσιάζονται εικόνες για τις δύο αυτές μορφές ανακατεύθυνσης:

- HTML ιστοσελίδα

← → ↻ 🏠 83.212.169.101/page/agent353

## Σχετικά με:agent353

Μία Οντότητα του Τύπου : [Person](#) Από τον Σημασμένο Γράφο : <http://www.libevo.gr> στο Πεδίο Ορισμού : <http://83.212.169.101>

Ιδιότητα	Τιμή
rdf:type	<ul style="list-style-type: none"> <li>foaf:Person</li> </ul>
rdfs:label	<ul style="list-style-type: none"> <li>ΛΟΥΝΤΕΜΗΣ ΜΕΝΕΛΑΟΣ</li> </ul>
foaf:name	<ul style="list-style-type: none"> <li>ΛΟΥΝΤΕΜΗΣ ΜΕΝΕΛΑΟΣ</li> </ul>
foaf:givenName	<ul style="list-style-type: none"> <li>ΜΕΝΕΛΑΟΣ</li> </ul>
foaf:surname	<ul style="list-style-type: none"> <li>ΛΟΥΝΤΕΜΗΣ</li> </ul>
είναι dc:contributor του	<ul style="list-style-type: none"> <li><a href="http://83.212.169.101/resource/record10565">http://83.212.169.101/resource/record10565</a></li> <li><a href="http://83.212.169.101/resource/record10696">http://83.212.169.101/resource/record10696</a></li> <li><a href="http://83.212.169.101/resource/record10722">http://83.212.169.101/resource/record10722</a></li> <li><a href="http://83.212.169.101/resource/record10723">http://83.212.169.101/resource/record10723</a></li> <li><a href="http://83.212.169.101/resource/record10731">http://83.212.169.101/resource/record10731</a></li> <li><a href="http://83.212.169.101/resource/record10734">http://83.212.169.101/resource/record10734</a></li> <li><a href="http://83.212.169.101/resource/record16269">http://83.212.169.101/resource/record16269</a></li> <li><a href="http://83.212.169.101/resource/record17630">http://83.212.169.101/resource/record17630</a></li> <li><a href="http://83.212.169.101/resource/record17631">http://83.212.169.101/resource/record17631</a></li> <li><a href="http://83.212.169.101/resource/record17632">http://83.212.169.101/resource/record17632</a></li> <li><a href="http://83.212.169.101/resource/record17633">http://83.212.169.101/resource/record17633</a></li> <li><a href="http://83.212.169.101/resource/record334">http://83.212.169.101/resource/record334</a></li> <li><a href="http://83.212.169.101/resource/record335">http://83.212.169.101/resource/record335</a></li> <li><a href="http://83.212.169.101/resource/record339">http://83.212.169.101/resource/record339</a></li> <li><a href="http://83.212.169.101/resource/record340">http://83.212.169.101/resource/record340</a></li> <li><a href="http://83.212.169.101/resource/record341">http://83.212.169.101/resource/record341</a></li> <li><a href="http://83.212.169.101/resource/record342">http://83.212.169.101/resource/record342</a></li> <li><a href="http://83.212.169.101/resource/record343">http://83.212.169.101/resource/record343</a></li> <li><a href="http://83.212.169.101/resource/record344">http://83.212.169.101/resource/record344</a></li> <li><a href="http://83.212.169.101/resource/record345">http://83.212.169.101/resource/record345</a></li> </ul>

**Εικόνα 29: Μορφή 303 ανακατεύθυνσης - Ιστοσελίδα HTML**

- RDF/XML αρχείο



```

<?xml version="1.0" encoding="utf-8" ?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:foaf="http://xmlns.com/foaf/0.1/" >
  <rdf:Description rdf:about="http://83.212.169.101/resource/record11947">
    <dc:contributor rdf:resource="http://83.212.169.101/resource/agent11669" />
  </rdf:Description>
  <rdf:Description rdf:about="http://83.212.169.101/resource/agent11669">
    <rdf:type rdf:resource="http://xmlns.com/foaf/0.1/Person" />
    <rdfs:label>ΜΑΝΔΗΛΑΡΑΣ, ΦΙΛΙΠΠΟΣ</rdfs:label>
    <foaf:name>ΜΑΝΔΗΛΑΡΑΣ ΦΙΛΙΠΠΟΣ</foaf:name>
    <foaf:givenName>ΦΙΛΙΠΠΟΣ</foaf:givenName>
    <foaf:surname>ΜΑΝΔΗΛΑΡΑΣ</foaf:surname>
  </rdf:Description>
  <rdf:Description rdf:about="http://83.212.169.101/resource/record11948">
    <dc:contributor rdf:resource="http://83.212.169.101/resource/agent11669" />
  </rdf:Description>
  <rdf:Description rdf:about="http://83.212.169.101/resource/record11949">
    <dc:contributor rdf:resource="http://83.212.169.101/resource/agent11669" />
  </rdf:Description>
  <rdf:Description rdf:about="http://83.212.169.101/resource/record11950">
    <dc:contributor rdf:resource="http://83.212.169.101/resource/agent11669" />
  </rdf:Description>
  <rdf:Description rdf:about="http://83.212.169.101/resource/record12197">
    <dc:contributor rdf:resource="http://83.212.169.101/resource/agent11669" />
  </rdf:Description>
  <rdf:Description rdf:about="http://83.212.169.101/resource/record12654">
    <dc:contributor rdf:resource="http://83.212.169.101/resource/agent11669" />
  </rdf:Description>
</rdf:RDF>

```

**Εικόνα 30: Μορφή 303 ανακατεύθυνσης - RDF αρχείο**

Η HTML ιστοσελίδα παράγεται δυναμικά μέσω της εφαρμογής Babel που χρησιμοποιεί και η Dbpedia. Η εφαρμογή αυτή είναι ανοιχτού κώδικα και παραμετροποιήθηκε με τέτοιο τρόπο ώστε να προσαρμοστεί στις ανάγκες της συγκεκριμένης διπλωματικής.

Ανάλογα με τον τύπο του πόρου που περιγράφεται, υπάρχουν οι εξής μορφές URI για τα δεδομένα της Δημοτικής Βιβλιοθήκης Ευόσμου. Συγκεκριμένα:

Τυποποιημένο URI	Τύπος πόρου
<a href="http://83.212.169.101/resource/entityID">http://83.212.169.101/resource/entityID</a>	Μοναδικό αναγνωριστικό πόρου
<a href="http://83.212.169.101/page/entityID">http://83.212.169.101/page/entityID</a>	Html απεικόνιση πόρου
<a href="http://83.212.169.101/data/entityID.rdf">http://83.212.169.101/data/entityID.rdf</a>	Δεδομένα πόρου σε μορφή αρχείου RDF/XML
<a href="http://83.212.169.101/data/entityID.n3">http://83.212.169.101/data/entityID.n3</a>	Δεδομένα πόρου σε μορφή αρχείου

	Notation3
<a href="http://83.212.169.101/data/entityID.csv">http://83.212.169.101/data/entityID.csv</a>	Δεδομένα πόρου σε μορφή αρχείου Comma-separated values
<a href="http://83.212.169.101/data/entityID.ttl">http://83.212.169.101/data/entityID.ttl</a>	Δεδομένα πόρου σε μορφή αρχείου Turtle
<a href="http://83.212.169.101/data/entityID.json">http://83.212.169.101/data/entityID.json</a>	Δεδομένα πόρου σε μορφή αρχείου JavaScript Object Notation

## 4.4. Μεταφόρτωση RDF/XML αρχείου στο Datahub

Έπειτα από τη φόρτωση του αρχείου των βιβλιογραφικών δεδομένων της Δημοτικής Βιβλιοθήκης Ευόσμου στο αποθετήριο τριπλετών (triplestore), κρίθηκε αναγκαίο να διατεθεί στο Datahub<sup>64</sup>, το αποθετήριο των συνόλων δεδομένων του Linked Data Cloud.

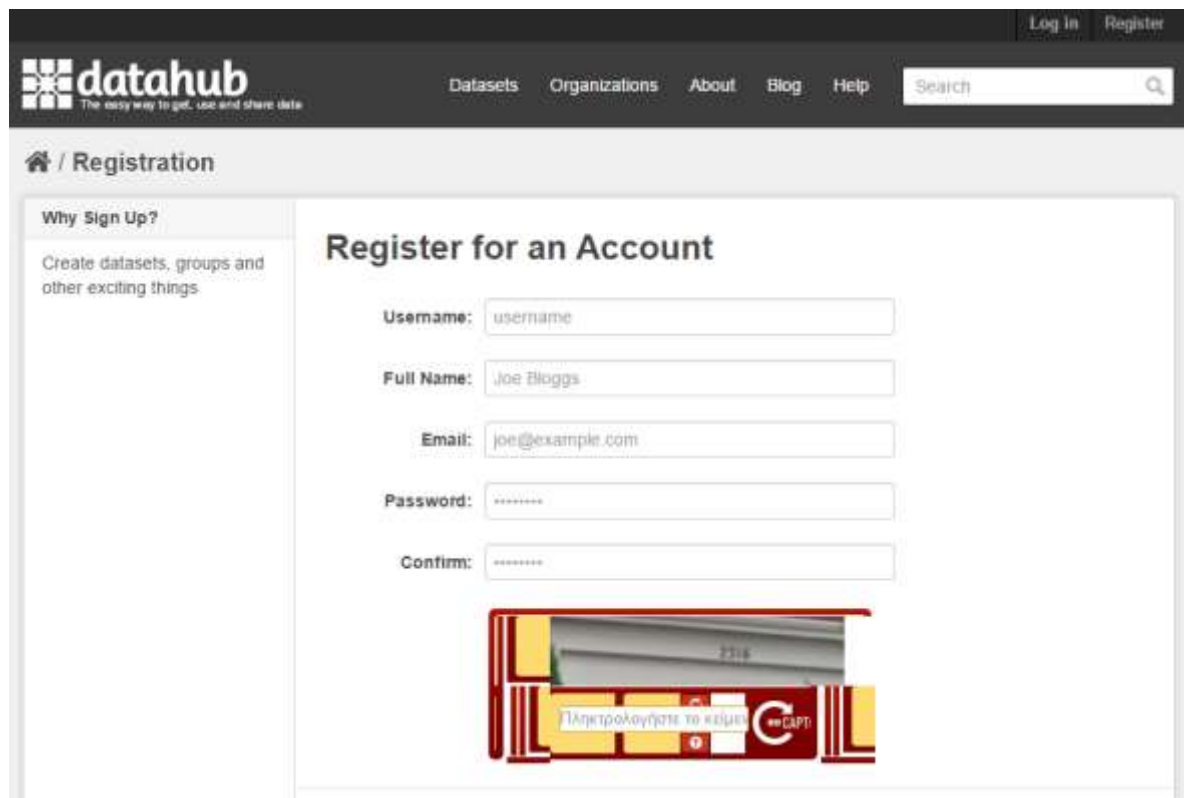
Για να επιτευχθεί αυτό ακολουθήθηκε η εξής διαδικασία:

1. Δημιουργία λογαριασμού στο Datahub
2. Δημιουργία ενός Οργανισμού
3. Διαμόρφωση χαρακτηριστικών για το σύνολο δεδομένων
4. Μεταφόρτωση αρχείου και τελικού σημείου SPARQL

Πιο ειδικά, δημιουργήθηκε ένας απλός λογαριασμός στο Datahub, όπως μπορεί να δει κανείς στην παρακάτω εικόνα:

---

<sup>64</sup> <http://datahub.io/>



**Εικόνα 31: Δημιουργία λογαριασμού στο Datahub**

Στη συνέχεια, δημιουργήθηκε ο οργανισμός «Public Library of Evosmos» και μέσω αυτού του οργανισμού μεταφορτώθηκαν το RDF/XML αρχείο της βιβλιοθήκης και το τελικό σημείο SPARQL (SPARQL Endpoint) απ' όπου προσφέρεται. Το αρχείο μεταφορτώθηκε με την ανοιχτή άδεια Creative Commons Not Commercial Use, η οποία προσφέρει ελεύθερα το αρχείο και το τελικό σημείο SPARQL για ανάκτηση των δεδομένων για οποιαδήποτε χρήση, ωστόσο, δεν μπορεί κανείς να εκμεταλλευτεί τα παραπάνω για εμπορική χρήση. Το σύνολο δεδομένων της Δημοτικής Βιβλιοθήκης Ευόσμου και το τελικό σημείο SPARQL φιλοξενούνται στη διεύθυνση: <http://datahub.io/dataset/library-of-evosmos-bibliographic-data> (βλ. εικ. 32).

# Library of Evosmos Bibliographic Data

The emergence of the World Wide Web strongly influenced the Libraries, which until then were the main source of knowledge and information. The transition from the Web to the Semantic Web has solved many significant problems related to the supply of enhanced information to the end user. The opening and interconnection of data provided by the use of Semantic Web technologies allows the direct use of the data information offered by the libraries. This thesis examines methods that semantically describe bibliographic data and present these to the end user as Linked Data. In particular, the bibliographic data of the Public Library of Evosmos were transformed from the MARC format to the RDF format and were interconnected with the datasets of the Public Library of Veria and of DBpedia. Finally, this thesis can be a recommended practice for librarians who are interested in opening and linking the bibliographic data of the libraries they work in.

## Data and Resources

**The RDF/XML dataset**  
The RDF/XML dataset

**SPARQL Endpoint**  
A SPARQL Endpoint to retrieve the data of Public Library

bibliographic data > greek dataset > library > public libraries

## Additional Info

Field	Value
Author	Chara Samothrakitou
Maintainer	Chara Samothrakitou
State	active

Εικόνα 32: Βιβλιογραφικά δεδομένα Δημοτικής Βιβλιοθήκης Ευόσμου στο Datahub

## 4.5. Διασύνδεση Δεδομένων

Εφόσον έχουν δημοσιευθεί οι τριπλέτες της Δημοτικής Βιβλιοθήκης Ευόσμου και αυτές δύναται να ανακτηθούν μέσω του τελικού σημείου SPARQL (SPARQL Endpoint), μπορεί να γίνει λόγος για τα ανοιχτά δεδομένα της. Ωστόσο, τα δεδομένα αυτά δεν είναι ακόμα διασυνδεδεμένα. Για τη δημιουργία Ανοιχτών και Διασυνδεδεμένων Δεδομένων κρίνεται αναγκαία η διασύνδεση των δεδομένων με άλλα ανοιχτά δεδομένα. Για την υλοποίηση της διασύνδεσης ακολουθήθηκαν δύο βήματα:

1. Εύρεση κι επιλογή άλλων ανοιχτών δεδομένων που θα εμπλούτιζαν τα δεδομένα της Δημοτικής Βιβλιοθήκης Ευόσμου
2. Επιλογή κατάλληλης εφαρμογής για τη διασύνδεση

Τα παραπάνω βήματα αναλύονται στα επόμενα υποκεφάλαια δίνοντας μία ολοκληρωμένη εικόνα για τη διαδικασία σύνδεσης των δεδομένων.

#### *4.5.1. Εύρεση κι επιλογή άλλων ανοιχτών δεδομένων*

Η έρευνα για την εύρεση των άλλων ανοιχτών δεδομένων που θα εμπλούτιζαν τα βιβλιογραφικά δεδομένα της Δημοτικής Βιβλιοθήκης Ευόσμου αξιοποιήθηκε το Open Linked Data Cloud και πιο ειδικά το Datahub<sup>65</sup>, το οποίο είναι το αποθετήριο των συνόλων δεδομένων του cloud, που επιτρέπει την αναζήτηση σε δεδομένα, τη δημοσίευση δεδομένων και τη διαχείριση των συνόλων δεδομένων.

Για την εύρεση συνόλων δεδομένων πραγματοποιήθηκε αναζήτηση με την εξής λέξη – κλειδί: “bibliographic records”. Η αναζήτηση άλλων βιβλιογραφικών δεδομένων ανέκτυσε 40 αποτελέσματα, στα οποία πραγματοποιήθηκε εκτενής έρευνα για την εύρεση ολόκληρων αρχείων του συνόλου δεδομένων και όχι παραδείγματα αυτών ή ενός εν λειτουργία τελικού σημείου SPARQL. Μέσω της διαδικασίας αυτής επιλέχθηκαν τα δεδομένα της Δημόσιας Βιβλιοθήκης Βέροιας με τη λογική της ύπαρξης περισσότερων συνδέσεων βάσει της γλώσσας τους. Ακόμα, για τον εμπλουτισμό των ονομάτων των συγγραφέων επιλέχθηκαν τα δεδομένα της Dbpedia.

#### *4.5.2. Επιλογή κατάλληλης εφαρμογής για τη διασύνδεση*

Στη βιβλιογραφία, όπως αναλύθηκε και στο κεφάλαιο 3 (βλ. κεφ. 3), στα περισσότερα σύνολα δεδομένων δεν παρουσιάζεται λεπτομερώς η διαδικασία διασύνδεσης με άλλα δεδομένα, ωστόσο σε ορισμένα από αυτά αναφέρεται ονομαστικά η εφαρμογή που χρησιμοποιήθηκε για τη διασύνδεση αυτή. Αρχικά, η έρευνα ξεκίνησε από τις εφαρμογές αυτές με κριτήρια την ευκολία της χρήσης, το επίπεδο τεχνογνωσίας που έπρεπε να διαθέτει ο χρήστης για να τη χρησιμοποιήσει και τον τρόπο διασύνδεσης των δεδομένων. Οι εφαρμογές που χρησιμοποιήθηκαν στις καλές πρακτικές που αναφέρθηκαν στο κεφάλαιο 3 είναι οι εξής:

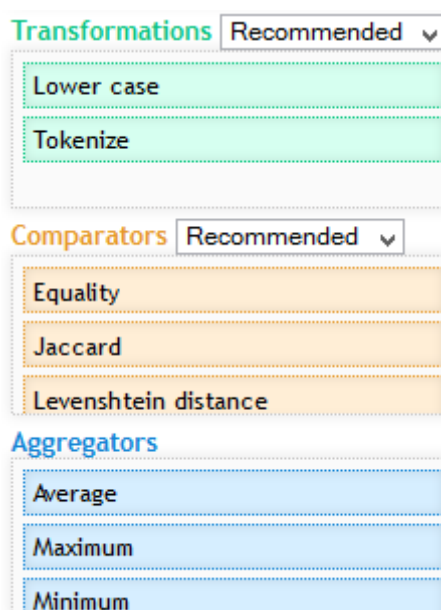
- nazca<sup>66</sup> - Εθνική Βιβλιοθήκη της Γαλλίας (BNF)
- MaRiMba<sup>67</sup> - Εθνική Βιβλιοθήκη Ισπανίας (BNE)
- Silk Workbench<sup>68</sup> - Δημόσια Βιβλιοθήκη Βέροιας

<sup>65</sup> <http://datahub.io/>

<sup>66</sup> <http://www.logilab.org/project/nazca>

<sup>67</sup> <http://mayor2.dia.fi.upm.es/oeg-upm/index.php/en/technologies/228-marimba>

Οι παραπάνω εφαρμογές αξιολογήθηκαν βάσει των κριτηρίων που τέθηκαν και επιλέχθηκε το Silk Workbench για τη διασύνδεση των δεδομένων της Δημοτικής Βιβλιοθήκης Ευόσμου. Οι λόγοι που απορρίφθηκαν οι υπόλοιπες εφαρμογές ήταν οι υψηλές απαιτήσεις τεχνογνωσίας στην περίπτωση της nazca και η απουσία διάθεσης της εφαρμογής από τους δημιουργούς στην περίπτωση του MaRiMba. Ωστόσο, το Silk Workbench δεν ήταν μία ευκαιριακή επιλογή. Αντιθέτως, είναι μια διαδικτυακή εφαρμογή ανοιχτού κώδικα, η οποία προσφέρει στο χρήστη μία φιλική διεπαφή και η διασύνδεση δύο πηγών δεδομένων πραγματοποιείται μέσω απλής drag and drop διαδικασίας, γεγονός που διασφαλίζει την ευκολία χρήσης και από ανθρώπους που δεν έχουν ιδιαίτερες προγραμματιστικές γνώσεις. Ακόμα, για τη διασύνδεση των δεδομένων προσφέρονται πολλές επιλογές όσον αφορά τους μετασχηματισμούς (transformations), τους συγκριτές (comparators) και τις συναθροίσεις (aggregators) (βλ. εικ. 33).



**Εικόνα 33: Επιλογές σύνδεσης στο Silk**

#### *4.5.3. Silk Workbench: παράδειγμα διασύνδεσης συνόλων δεδομένων*

Στο σημείο αυτό κρίνεται αναγκαία η αναλυτική παρουσίαση της διασύνδεσης δύο συγκεκριμένων συνόλων δεδομένων με το Silk Workbench, έτσι ώστε να αποδοθεί μία ολοκληρωμένη εικόνα και να μπορεί να υλοποιηθεί από οποιονδήποτε ενδιαφερόμενο.

Το Silk Workbench εκτελείται μέσω του command-line (cmd) του υπολογιστή (τρέχοντας το «silk-workbench.bat»), ωστόσο, διαθέτει μία γραφική διεπαφή για τον χρήστη μέσω ενός κοινού εξυπηρετητή στη διεύθυνση «localhost:9000» (βλ. εικ. 34). Στο σημείο αυτό είναι αναγκαίο να ειπωθεί πως η εφαρμογή δεν ήταν απόλυτα λειτουργική σε

---

<sup>68</sup> [https://www.assembla.com/spaces/silk/wiki/Silk\\_Workbench](https://www.assembla.com/spaces/silk/wiki/Silk_Workbench)

όλους τους εξυπηρετητές και έπειτα από πολλές δοκιμές διαπιστώθηκε πως η εφαρμογή προσφέρει όλες τις δυνατότητες της μόνο μέσω του Mozilla Firefox.



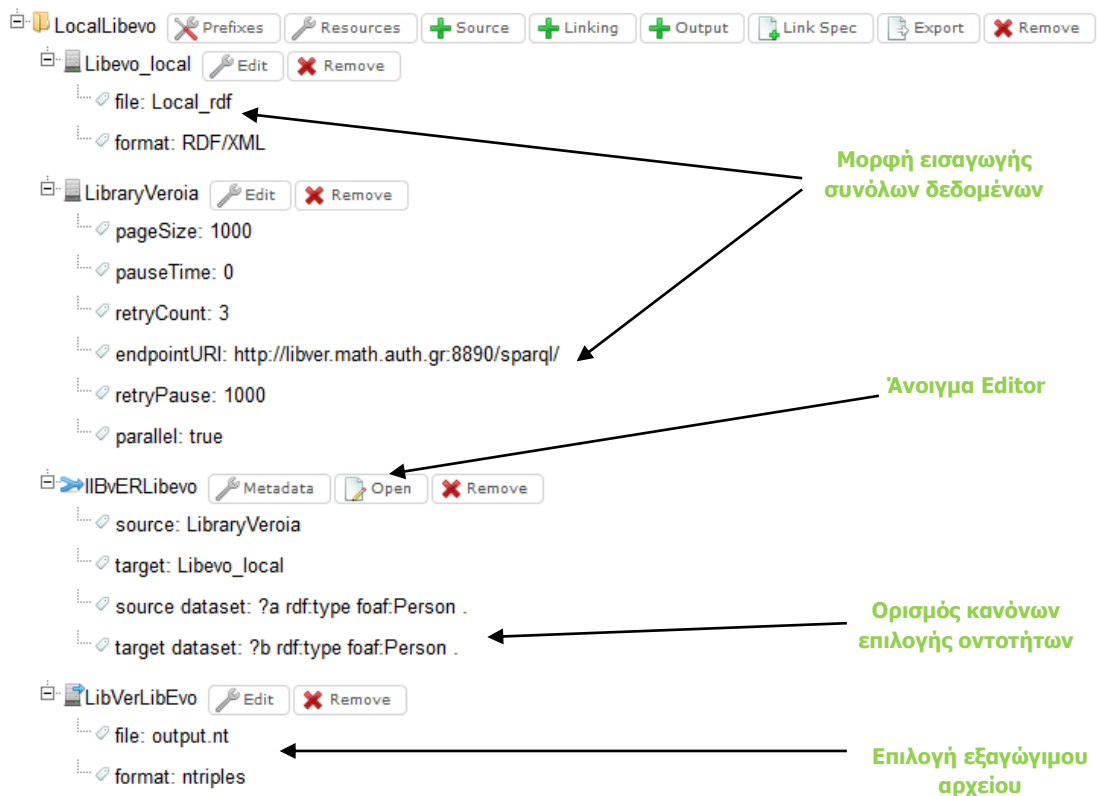
**Εικόνα 34: Διεπαφή χρήστη Silk Workbench**

Αφού λοιπόν ανοίξει κανείς έναν καινούργιο χώρο εργασίας (Open Workspace) πρέπει να δημιουργήσει ένα καινούργιο έργο (Project) και να του δώσει το όνομα που επιθυμεί. Αφού δημιουργηθεί ένα καινούργιο έργο (Project) προσθέτει τις πηγές των δεδομένων. Το Silk προσφέρει στους χρήστες του δύο επιλογές εισαγωγής πηγών των συνόλων δεδομένων που επιθυμείται να συνδεθούν. Ο πρώτος είναι μέσω εισαγωγής ενός αρχείου σε μία από τις εξής μορφές: RDF/XML, N-Triples, N-Quads, Turtle και ο δεύτερος μέσω της εισαγωγής του URI του τελικού σημείου SPARQL (SPARQL Endpoint) των δεδομένων.

Στην παρουσίαση του παραδείγματος αυτού επιθυμείται η διασύνδεση των συνόλων δεδομένων της Δημοτικής Βιβλιοθήκης Ευόσμου και της Δημόσιας Βιβλιοθήκης Βέροιας. Η εισαγωγή των συνόλων δεδομένων πραγματοποιήθηκε και με τους δύο παραπάνω τρόπους και πιο ειδικά, τα δεδομένα της Δημοτικής Βιβλιοθήκης Ευόσμου εισήχθησαν με τη μορφή RDF/XML και τα δεδομένα της Δημόσιας Βιβλιοθήκης Βέροιας μέσω του τελικού σημείου SPARQL (βλ. εικ. 35).

Στη συνέχεια δηλώνονται τα σύνολα δεδομένων που θα διασυνδεθούν και ορίζονται κανόνες για την επιλογή των οντοτήτων που επιθυμείται να διασυνδεθούν μέσω της γλώσσας επερωτήσεων SPARQL αφού πρώτα αποφασιστεί τι οντότητες επιδιώκεται να συνδεθούν ανάμεσα στα δύο σύνολα δεδομένων. Στο εν λόγω παράδειγμα επιθυμείται η διασύνδεση των πνευματικά υπεύθυνων φυσικών προσώπων (βλ. υποκ. 4.2.1.) (Πηγή: `?a rdf:type foaf:Person, Στόχος: ?b rdf:type foaf:Person`) (βλ. εικ. 35).

Τέλος, δηλώθηκε το αρχείο που θα εξαχθούν οι νέες διασυνδεδεμένες τριπλέτες (output.nt) (βλ. εικ.35).



**Εικόνα 35: Ρυθμίσεις αρχικών δηλώσεων του Silk**

Το επόμενο στάδιο της διασύνδεσης των συνόλων δεδομένων μέσω της εφαρμογής Silk ήταν ο ορισμός κανόνων διασύνδεσης αφού πρώτα επιλέχθηκε το άνοιγμα του Editor (Open) (βλ. εικ. 35).

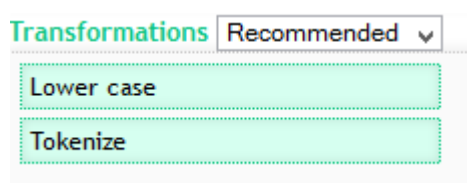
Αρχικά, στον Editor επιλέχθηκαν τα πεδία που επιθυμείται να διασυνδεθούν τόσο από την πηγή (Δημόσια Βιβλιοθήκη Βέροιας), όσο κι από το στόχο (Δημοτική Βιβλιοθήκη Ευόσμου) (βλ. εικ. 36). Στην προκειμένη περίπτωση, όπως έχει αναφερθεί και παραπάνω έχει επιλεγθεί το πεδίο των πνευματικά υπεύθυνων και για τα δύο σύνολα δεδομένων (βλ. εικ. 36).



**Εικόνα 36: Επιλογή πεδίων από τα σύνολα δεδομένων στο Silk**

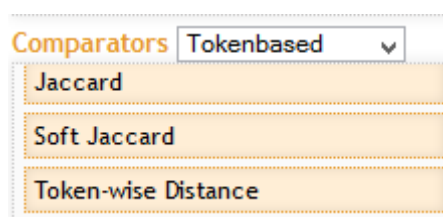


Στη συνέχεια επιλέγονται τα φίλτρα μετατροπής των τιμών των πεδίων που επιθυμούνται καθώς τα διαφορετικά σύνολα δεδομένων συνήθως χρησιμοποιούν διαφορετικές μορφές, ένας μετασχηματισμός μπορεί να χρησιμοποιηθεί για την ομαλοποίηση των τιμών πριν τη σύγκρισή τους. Το Silk προσφέρει συνολικά 37 φίλτρα μετατροπής των τιμών (βλ. εικ. 37) μεταξύ αυτών ένα φίλτρο για τη μετατροπή όλων των δεδομένων σε πεζά γράμματα (Lower Case), ένα φίλτρο για την αφαίρεση παρενθέσεων (Remove blanks), ένα φίλτρο για την αφαίρεση κενών (Remove Parentheses), κ.α. Τα φίλτρα που χρησιμοποιήθηκαν για το παράδειγμα αυτό ήταν το Lower Case και το Tokenize. Με το πρώτο φίλτρο επιδιώχθηκε η μετατροπή όλων των δεδομένων και της πηγής και του στόχου σε πεζά γράμματα, έτσι ώστε να μπορούν να συγκριθούν με μεγαλύτερη ευκολία. Το δεύτερο φίλτρο κατακερμάτισε όλα τα δεδομένα της πηγής και του στόχου σε ξεχωριστές λέξεις, έτσι ώστε να μπορούν να συγκριθούν ανεξάρτητα τη σειρά με την οποία εμφανίζονται (βλ. εικ. 37).



**Εικόνα 37: Επιλογή μετασχηματισμών για τα σύνολα δεδομένων στο Silk**

Έπειτα, επιλέγονται τα κριτήρια σύγκρισης που θα εφαρμοστούν στα δεδομένα μέσα από το πλήθος των κριτηρίων που διαθέτει το Silk, τα οποία είναι 22 στο σύνολο τους (βλ. εικ. 38). Ένας τελεστής σύγκρισης αξιολογεί δύο εισόδους και υπολογίζει την ομοιότητα τους που βασίζεται σε μία μετρική καθορισμένη από το χρήστη. Στο συγκεκριμένο παράδειγμα χρησιμοποιήθηκε ο τελεστής σύγκρισης Soft Jaccard, ο οποίος είναι μία στατιστική μετρική, η οποία παίρνει τιμές από 0 μέχρι και 1, όπου το 0 σημαίνει καμία ομοιότητα και το 1 πλήρη ταύτιση. Ο τελεστής αυτός ορίζεται ως το σύνολο των όμοιων δεδομένων προς το σύνολο όλων των δεδομένων που συγκρίνονται. Όταν τα δεδομένα είναι ταυτόσημα τότε το Silk πραγματοποιεί σύνδεση μεταξύ των δύο συνόλων δεδομένων.



**Εικόνα 38: Επιλογή τελεστή σύγκρισης δεδομένων στο Silk**

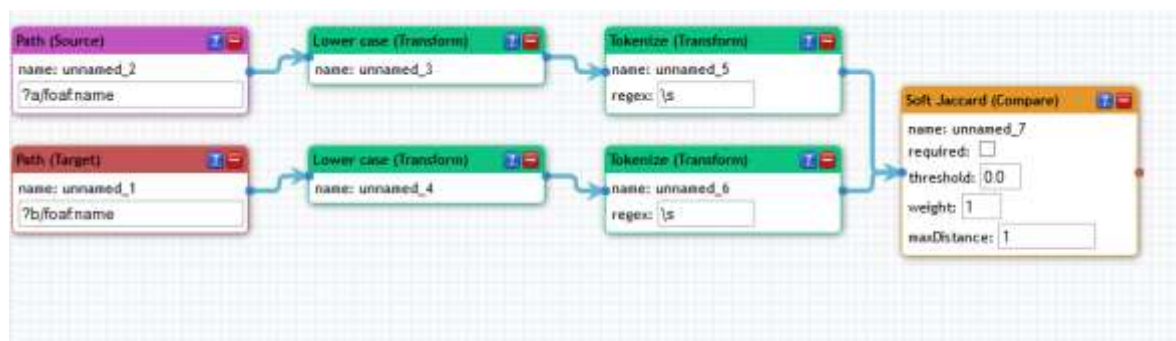
Τέλος, το Silk δίνει τη δυνατότητα συνάθροισης των δεδομένων μέσω συγκεκριμένων τελεστών (βλ. εικ. 39), οι οποίοι είναι συνολικά 5 και συγκεντρώνουν πολλαπλές τιμές σε μια ενιαία τιμή. Στο συγκεκριμένο παράδειγμα δε χρησιμοποιήθηκαν τελεστές συνάθροισης.

## Aggregators



Εικόνα 39: Επιλογή τελεστών συνάθροισης στο Silk

Αφού καθορίστηκαν όλα τα παραπάνω για τα σύνολα δεδομένων με drag and drop στον Editor (βλ. εικ. 40), στη συνέχεια επιλέχθηκε το «Generate Links» για την έναρξη της διαδικασίας ταυτοποίησης.



Εικόνα 40: Σχηματική απεικόνιση κανόνων στον Editor του Silk

Με την επιλογή Start ξεκινάει η διαδικασία, αφού πρώτα επιβεβαιωθεί το output που θα εξαχθούν οι νέες τριπλέτες (output.nt).

Το αρχείο αυτό των τριπλετών μεταφορτώθηκε στον Virtuoso, στον ίδιο γράφο (<http://www.libevo.gr>) που είχε δημιουργηθεί για το σύνολο των δεδομένων της Δημοτικής Βιβλιοθήκης Ευόσμου (βλ. 4.3.2.1.). Από τη διασύνδεση με τη Δημόσια Βιβλιοθήκη της Βέροιας, προέκυψαν 150 νέες τριπλέτες συνδεδεμένες με owl:sameAs για τους πνευματικά υπεύθυνους (βλ. εικ. 41).

## Σχετικά με: agent2987

Μία Οντότητα του Τύπου: [Person](#) Από τον Σημειωμένο Γράφο: <http://www.libevo.gr> στο Πεδίο Ορισμού: <http://83.212.169.101>

Ιδιότητα	Τιμή
<a href="#">rdf:type</a>	<ul style="list-style-type: none"><li>foaf:Person</li></ul>
<a href="#">rdfs:label</a>	<ul style="list-style-type: none"><li>BENOITE, GROULT</li></ul>
<a href="#">foaf:name</a>	<ul style="list-style-type: none"><li>BENOITE GROULT</li></ul>
<a href="#">foaf:givenName</a>	<ul style="list-style-type: none"><li>GROULT</li></ul>
<a href="#">foaf:surname</a>	<ul style="list-style-type: none"><li>BENOITE</li></ul>
<a href="#">owl:sameAs</a> του	<ul style="list-style-type: none"><li><a href="http://libev.math.auth.gr/resource/agent8417">http://libev.math.auth.gr/resource/agent8417</a></li></ul>
<a href="#">owl:contributor</a> του	<ul style="list-style-type: none"><li><a href="http://83.212.169.101/resource/record3079">http://83.212.169.101/resource/record3079</a></li></ul>

Δεδομένα σε μορφή: [CSV](#) [RDF](#) ([N-Triples](#) [N3](#) [Turtle](#) [JSON](#) [XML](#))

This content was extracted as a part of master thesis of [M.Sc. e-Learning Program, Department of Digital Systems of University of Piraeus](#) student [Samothraikou Zacharoula](#) under supervisor [Dr. George Vouros](#)

WebSite generated by [ReteSite 1160](#) (L3 Developer version) © ReteSite  
Developed by [Tactica Engineering](#)

Εικόνα 41: Σύνδεση owl:sameAs στο SPARQL ENDpoint

## 4.6. Παρουσίαση Δεδομένων

Έπειτα από το άνοιγμα και τη διασύνδεση του συνόλου των δεδομένων της Δημοτικής Βιβλιοθήκης Ευόσμου κι αφού τα δεδομένα πληρούν όλες τις βασικές αρχές των Διασυνδεδεμένων Δεδομένων (βλ. υποκεφ. 2.5.) κρίθηκε αναγκαία η δημιουργία ενός ενιαίου σημείου πρόσβασης στα δεδομένα διευκολύνοντας την περιήγηση του εκάστοτε ενδιαφερόμενου χρήστη. Για την υλοποίηση του ενιαίου αυτού σημείου πρόσβασης επιλέχθηκε η διαμόρφωση ενός CMS (Content Management Systems), του Drupal<sup>69</sup>.

Η επιλογή του συγκεκριμένου CMS (Content Management Systems) δεν μπορεί να θεωρηθεί τυχαία. Κατά την αξιολόγηση των τριών δημοφιλέστερων CMS (Joomla<sup>70</sup>, WordPress<sup>71</sup> και Drupal<sup>72</sup>) ως προς τον προσανατολισμό, την επάρκεια και την ωριμότητα υποστήριξης του Σημαιολογικού Ιστού. Μέσω αυτής της αξιολόγησης το Drupal αναδείχθηκε ως το πρώτο CMS με εγγενή υποστήριξη RDF και τεχνολογιών Σημαιολογικού Ιστού. Επιπλέον, το Drupal επιλέχθηκε για την ανάπτυξη ιστοχώρου σχετικού με τον Σημαιολογικό Ιστό λόγω του ευέλικτου συστήματος διαμόρφωσης εμφάνισης, της προηγμένης και φιλικής διεπαφής χρήστη και της δόμησης περιεχομένου σε κόμβους. Είναι αξιοσημείωτο να ειπωθεί πως οι δυνατότητες αξιοποίησης των τεχνολογιών του Σημαιολογικού Ιστού που προσφέρονται από το συγκεκριμένο CMS, δεν αξιοποιήθηκαν. Ωστόσο, η σχέση και η συνάφεια του με τον Σημαιολογικό Ιστό ήταν καθοριστικοί λόγοι για την επιλογή του.

Η ιστοσελίδα που δημιουργήθηκε βρίσκεται στη διεύθυνση: <http://83.212.169.101/samothrakitou/> και διαθέτει τέσσερα κεντρικά Μενού (βλ. εικ. 42):

---

<sup>69</sup> <https://www.drupal.org/>

<sup>70</sup> <https://www.joomla.org/>

<sup>71</sup> <https://wordpress.org/>

<sup>72</sup> <https://www.drupal.org/>



**Εικόνα 42: Κεντρικό Μενού Drupal**

- **Αρχική**

Η Αρχική σελίδα προσφέρει πληροφορίες για τη διπλωματική που υλοποιήθηκε, τους σκοπούς της και για τις προσφερόμενες υπηρεσίες της ιστοσελίδας (βλ. εικ. 43).



**Καλωσήρθατε στη Σημασιολογική Αναζήτηση της Δημοτικής Βιβλιοθήκης Ευόσμου**

Posted By admin On Monday, 21 September 2015

Η ιστοσελίδα αυτή δημιουργήθηκε στα πλαίσια της Διπλωματικής Εργασίας με τίτλο "Ανοιχτά Διασυνδεδεμένα Δεδομένα και Βιβλιοθήκες: η διάθεση και η διασύνδεση των βιβλιογραφικών δεδομένων της Δημοτικής Βιβλιοθήκης Ευόσμου" που πραγματοποιήθηκε στα πλαίσια του ΜΠΣ "Ηλεκτρονική Μάθηση" του Πανεπιστημίου Πειραιώς. Στην παρούσα διπλωματική εξετάζονται μέθοδοι σημασιολογικής περιγραφής βιβλιογραφικών δεδομένων και παρουσιάστές τους στον τελικό χρήστη ως Διασυνδεδεμένα Δεδομένα.

[Read more](#)

**Εικόνα 43: Αρχική σελίδα Drupal**

## ▪ Διασυνδεδεμένα Δεδομένα

Στο μενού «Διασυνδεδεμένα Δεδομένα» μπορεί κανείς να βρει πληροφορίες για τα Διασυνδεδεμένα Δεδομένα και τη σχέση τους με τις Βιβλιοθήκες (βλ. εικ. 44).



The screenshot shows the website header for the National Library of Greece (Δημοτική Βιβλιοθήκη Ευόσμου). The main navigation menu includes 'Αρχική', 'Διασυνδεδεμένα Δεδομένα' (highlighted), 'Σημαιολογική Αναζήτηση', and 'SPARQL Endpoint'. Below the menu, the page title is 'Home » Διασυνδεδεμένα Δεδομένα'. The main heading is 'Διασυνδεδεμένα Δεδομένα'. The text explains that semantic web standards facilitate the connection of open data from various institutions, allowing for a more comprehensive view of the cultural heritage. It lists several benefits: 1) It enables the reuse of data from various sources (libraries, archives, museums) to create new services. 2) It allows for the integration of data from different institutions, providing a more complete picture. 3) It provides a structured framework for data, making it easier to search and use. The text concludes by stating that semantic web standards are essential for the development of a digital cultural heritage infrastructure.

**Εικόνα 44: Μενού "Διασυνδεδεμένα Δεδομένα" στο Drupal**

## ▪ Σημαιολογική Αναζήτηση

Στο μενού «Σημαιολογική Αναζήτηση» προσφέρονται κάποιες έτοιμες επερωτήσεις SPARQL για την ανάκτηση των δεδομένων μαζί με τις επεξηγήσεις τους για τη διευκόλυνση των ενδιαφερόμενων χρηστών καθώς επίσης και για την απεικόνιση των περισσότερων λεξιλογίων που αξιοποιήθηκαν για το άνοιγμα των δεδομένων (βλ. εικ. 45).



## Σημασιολογική Αναζήτηση

Για τη σημασιολογική αναζήτηση στα δεδομένα της Δημοτικής Βιβλιοθήκης Εύόσμου, προσφέρεται ένα τελικό σημείο SPARQL μέσω του Open Link Virtuoso που μπορεί να το βρει κανείς στο μενού "SPARQL Endpoint". Για τη διευκόλυνση των ενδιαφερόμενων χρηστών καθώς επίσης και την απεικόνιση των περισσότερων λεξιλογίων που αξιοποιήθηκαν για το άνοιγμα των δεδομένων, παρακάτω προσφέρονται κάποιες έτοιμες επερωτήσεις SPARQL για την ανάκτηση των δεδομένων μαζί με τις επεξηγήσεις τους:

### SPARQL Example #1

```
PREFIX bibo: <http://purl.org/ontology/bibo/>
PREFIX dctems: <http://purl.org/dc/terms/>
PREFIX dc: <http://purl.org/dc/elements/1.1/>
```

```
SELECT ?name
WHERE
{ ?x foaf:name ?name
.
}
```

(Στο παραπάνω παράδειγμα επιθυμείται η ανάκτηση όλων των ονομάτων των συγγραφέων που υπάρχουν στο σύνολο των δεδομένων).

### SPARQL Example #2

```
PREFIX bibo: <http://purl.org/ontology/bibo/>
PREFIX dctems: <http://purl.org/dc/terms/>
```

**Εικόνα 45: Μενού "Σημασιολογική Αναζήτηση" στο Drupal**

### ▪ **SPARQL Endpoint**

Στο συγκεκριμένο μενού παρέχονται πληροφορίες για τα δεδομένα που βρίσκονται στο Datahub, έτσι ώστε να ανατρέξει όποιος επιθυμεί να τα χρησιμοποιήσει και έχει ενσωματωθεί το τελικό σημείο SPARQL, ώστε ο χρήστης να μην μεταβαίνει στη διεύθυνση όπου αυτό φιλοξενείται αλλά να έχει όλες τις πληροφορίες συγκεντρωμένες σε ένα ενιαίο σημείο πρόσβασης (βλ. εικ. 46).



Αρχική

Διασυνδεδεμένα Δεδομένα

Σημασιολογική Αναζήτηση

**SPARQL Endpoint**

Home » SPARQL Endpoint

## SPARQL Endpoint

Posted By admin On Wednesday, 11 February 2015

Παρακάτω προσφέρεται το τελικό σημείο SPARQL του συνόλου δεδομένων της Δημοτικής Βιβλιοθήκης Εύοςμου, το οποίο βρίσκεται και στη διεύθυνση: <http://83.212.169.101:8890/sparql>. Ο εκάστοτε ενδιαφερόμενος χρήστης μπορεί να περιηγηθεί στα δεδομένα της βιβλιοθήκης μέσω επερωτήσεων SPARQL. Για περισσότερες πληροφορίες για τη γλώσσα επερωτήσεων SPARQL μπορεί κανείς να επισκεφθεί τη σελίδα του W3C: <https://www.w3.org/TR/rdf-sparql-query/>.

### Virtuoso SPARQL Query Editor

[About](#) | [Namespace Prefixes](#) | [Inference rules](#)

Default Data Set Name (Graph IRI)

Query Text

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

**Εικόνα 46: Μενού "SPARQL Endpoint" στο Drupal**

# Συμπεράσματα

Στην παρούσα διπλωματική εργασία παρουσιάστηκε ένας τρόπος για τη διάθεση και τη διασύνδεση των βιβλιογραφικών δεδομένων μίας Δημοτικής Βιβλιοθήκης. Η επιλογή της Δημοτικής Βιβλιοθήκης πραγματοποιήθηκε με το σκεπτικό πως η πρακτική που προτείνεται θα μπορούσε να εφαρμοστεί στο μεγαλύτερο μέρος των βιβλιοθηκών που διαθέτει η χώρα. Μέσω της επιλογής αυτής προτείνεται μία πρακτική που μπορεί να εφαρμοστεί καθολικά από όλες τις Δημοτικές και Δημόσιες Βιβλιοθήκες της χώρας, ώστε να πραγματοποιηθεί μεγάλη συνεισφορά στο Σύννεφο των Ανοιχτών και Διασυνδεδεμένων Δεδομένων στο πεδίο των Βιβλιοθηκών από την Ελλάδα.

Ωστόσο, η ολοκλήρωση της εργασίας αυτής δεν ήταν μία εύκολη διαδικασία. Αυτό διαπιστώθηκε από τα πρώτα κίονα στάδια υλοποίησης της καθώς πολλές από τις προσπάθειες που έχουν πραγματοποιηθεί στο εξωτερικό δεν διαθέτουν αναλυτική τεκμηρίωση με τη διαδικασία υλοποίησης. Επίσης, πολλά από τα σύνολα δεδομένων που εντοπίστηκαν εξ αρχής δεν διατίθενται ελεύθερα για χρήση, παρά διατίθενται μέρη αυτών. Το ίδιο συμβαίνει και με τα τελικά σημεία τους, των οποίων τα περισσότερα δεν ήταν ανανεωμένα κι άλλα δε λειτουργούσαν καθόλου. Σε σύνοψη των παραπάνω, δε μπορούσε να διαπιστωθεί με ακρίβεια ο λεπτομερής τρόπος που πραγματοποιήθηκε η διαδικασία και πιο ειδικά, οι εφαρμογές και τα λεξιλόγια που χρησιμοποιήθηκαν.

Ακόμα κι όμως αν συνδύαζε κανείς τις καλές πρακτικές θα διαπίστωνε πως απαιτούνται ιδιαίτερες προγραμματιστικές γνώσεις, τις οποίες δεν κατέχουν οι βιβλιοθηκονόμοι και οι επιστήμονες της πληροφόρησης. Για να λυθούν όλα τα παραπάνω πραγματοποιήθηκε προσπάθεια απλούστευσης της διαδικασίας με έναν τρόπο ώστε ακόμα και οι βιβλιοθηκονόμοι να είναι σε θέση να διαθέσουν και να διασυνδέσουν τα δεδομένα της βιβλιοθήκης όπου εργάζονται. Το παραπάνω αποτελεί μία καινοτομία στην επιστήμη της Βιβλιοθηκονομίας και ανοίγει τους ορίζοντες σε όσους βιβλιοθηκονόμους ενδιαφέρονται να αξιοποιήσουν τις τεχνολογίες του Σημαιολογικού Ιστού για να εμπλουτίσουν τις προσφερόμενες υπηρεσίες τους.

Εν κατακλείδι, το πεδίο του Σημαιολογικού Ιστού έχει πολλά να προσφέρει στην επιστήμη της βιβλιοθηκονομίας. Το μόνο που αρκεί να συμβεί είναι να προσφερθούν στο βιβλιοθηκονομικό κοινό τα κίνητρα και η πληροφόρηση για το τι μπορεί να τους προσφέρει. Η παρούσα εργασία θα ήταν δυνατό να είναι η αρχή της συνεισφοράς όλων των βιβλιοθηκών της Ελλάδας στο Σύννεφο των Ανοιχτών και Διασυνδεδεμένων Δεδομένων κι αυτό είναι που προτείνεται ως μελλοντική εργασία. Για να πραγματοποιηθεί αυτό θα πρέπει οι βιβλιοθήκες της χώρας να κινητοποιηθούν για να διαθέσουν τα βιβλιογραφικά τους δεδομένα ως ανοιχτά και στη συνέχεια να πραγματοποιήσουν συνδέσεις με όλα τα διαθέσιμα σύνολα δεδομένων για τον εμπλουτισμό της βιβλιογραφικής πληροφορίας. Μόνο έτσι τα πολύτιμα δεδομένα των βιβλιοθηκών θα ξεφύγουν από τα στενά πλαίσια αυτών και θα μπορούν να χρησιμοποιηθούν για τη δημιουργία καινοτόμων εφαρμογών που θα εμπλουτίσουν στο μέγιστο την εμπειρία των χρηστών τους.



# Βιβλιογραφία

*About data.bnf.fr* (χ.χ.). Διαθέσιμο στο: <http://data.bnf.fr/about>

Acosta, M., Vidal, M. E., Lampo, T., Castillo, J. & Ruckhaus, E. (2011). *ANAPSID: An adaptive query processing engine for SPARQL Endpoints*. Διαθέσιμο στο: [http://iswc2011.semanticweb.org/fileadmin/iswc/Papers/Research\\_Paper/03/70310017.pdf](http://iswc2011.semanticweb.org/fileadmin/iswc/Papers/Research_Paper/03/70310017.pdf)

Al – Mukhtar, M. M. & Al - Assafy, A. T. A. (2014). The implementation of FOAF Ontology for an academic social network, *International Journal of Computer Science Engineering and Technology (IJCSET)*, 4 (1), σ. 10 – 14. Διαθέσιμο στο: <http://www.ijcset.net/docs/Volumes/volume4issue1/ijcset2014040103.pdf>

Alexander, N. & Ravada, S. (χ.χ.). *RDF object type and reification in Oracle*. USA: Oracle Corporation. Διαθέσιμο στο: [http://download.oracle.com/otndocs/tech/semantic\\_technologies/pdf/rdf\\_reification.pdf](http://download.oracle.com/otndocs/tech/semantic_technologies/pdf/rdf_reification.pdf)

Antoniou, G. & Harmelen, F. (2008). *A Semantic Web Primer*. USA: The MIT Press. Διαθέσιμο στο: <http://www.dcc.fc.up.pt/~zp/aulas/1213/pde/geral/bibliografia/MIT.Press.A.Semantic.Web.Primer.eBook-TLFeBOOK.pdf>

Armstrong, P., Steward, M. & Ward, A. (2013). *Case study: using linked data to integrate resources from cultural heritage institutions across Canada*. Διαθέσιμο στο: <http://or2013.net/sites/or2013.net/files/Integrating%20Resources%20with%20Linked%20Data/index.pdf>

Beckett, D., Berners-Lee, T., Prud'hommeaux, E. & Carothers, G. (2014). *RDF 1.1 Turtle: Terse RDF Triple Language, W3C*. Διαθέσιμο στο: <http://www.w3.org/TR/turtle/>

Bassiliades, N. (2005). *Semantic web: vision and technologies*. Διαθέσιμο στο: <http://lpis.csd.auth.gr/publications/conf-cs-halkidiki-bassiliades.pdf>

Bauer, F. & Kaltenbock, M. (2012). *Linked open data: the essentials*. A quick start guide for decision makers. Διαθέσιμο στο: <http://www.reeep.org/LOD-the-Essentials.pdf>

Belleau, F., Nolin, M. A., Tourigny, N., Rigault, P. & Morissette, J. (2008). *Bio2RDF: Towards a mashup to build bioinformatics knowledge system*, *Journal of Biomedical Informatics*, 41 (5), σ. 706 - 716. Διαθέσιμο στο: [http://www2007.org/workshops/paper\\_143.pdf](http://www2007.org/workshops/paper_143.pdf)

Berners – Lee, T. (2006). *Linked data: design issues, W3C*. Διαθέσιμο στο: <http://www.w3.org/DesignIssues/LinkedData>

- Berners – Lee, T. (2007). *Digital Future of the United States: Part I -- The Future of the World Wide Web*. Διαθέσιμο στο: <http://dig.csail.mit.edu/2007/03/01-ushouse-future-of-the-web.pdf>
- Berners – Lee, T. & Connolly, D. (2011) Notation3 (N3): A readable RDF syntax, *W3C*. Διαθέσιμο στο: <http://www.w3.org/TeamSubmission/n3/>
- Berners - Lee, T., Connolly, D., Kagal, L., Scharf, Y. & Hendler, J. (2007). N3Logic: A logical framework for the World Wide Web, *Theory and Practice of Logic Programming (TPLP)*, 8 (3). Διαθέσιμο στο: <http://arxiv.org/abs/0711.1533>
- Biblioteca Nacional de España: Data Source* (χ.χ.). Διαθέσιμο στο: <http://www.bne.es/en/Inicio/Perfiles/Bibliotecarios/DatosEnlazados/FuentesDatos/>
- Biblioteca Nacional de España: Technologies Used* (χ.χ.). Διαθέσιμο στο: <http://www.bne.es/en/Inicio/Perfiles/Bibliotecarios/DatosEnlazados/Tecnologia/>
- Bizer, C., Cyganiak, R. & Gauss, T. (2007): *The RDF Book Mashup: From Web APIs to a Web of Data*. Διαθέσιμο στο: <http://ceur-ws.org/Vol-248/paper4.pdf>
- Bizer, C., Heath, T. & Berners-Lee, T. (2009). *Linked data – The story so far*. Διαθέσιμο στο: <http://tomheath.com/papers/bizer-heath-berners-lee-ijswis-linked-data.pdf>
- Bizer, C., Heath, T., Idehen, K. & Berners – Lee, T. (2008). *Linked data on the web. Πρακτικά 17<sup>ου</sup> Διεθνούς Συνεδρίου WWW*. Διαθέσιμο στο: <http://dl.acm.org/citation.cfm?doid=1367497.1367760>
- Boris, M., (2005) *On the Properties of Metamodeling in OWL. Πρακτικά 4<sup>ου</sup> Διεθνούς Συνεδρίου Semantic Web (ISWC 05)*, σ. 548 – 562. Galway, Ireland. Διαθέσιμο στο: <https://www.cs.ox.ac.uk/boris.motik/pubs/motik07metamodeling-journal.pdf>
- Brickley, D. & Guha, R. V. (2014). *RDF Schema 1.1, W3C*. Διαθέσιμο στο: <http://www.w3.org/TR/rdf-schema/>
- Brickley, D. & Miller, L. (χ.χ.). *FOAF (2000 – 2015+)*. Διαθέσιμο στο: <http://www.foaf-project.org/>
- Broekstra, J., Klein, M., Decker, S., Fensel, D., Harmelen, F. & Horrocks, I. (2001). *Enabling knowledge representation on the Web by extending RDF Schema*. Διαθέσιμο στο: <http://www.cs.ox.ac.uk/ian.horrocks/Publications/download/2001/extending-RDF.pdf>
- Case study: German National Library publishes its authority and bibliographic records as linked data* (χ.χ.). Διαθέσιμο στο: <http://www.theeuropeanlibrary.org/tel4/newsitem/5802>
- Clark, J. (1999). *XSL Transformations (XSLT): Version 1.0, W3C*. Διαθέσιμο στο: <http://www.w3.org/TR/xslt>

- D'Arcus, B. & Giasson, F. (2009). Bibliographic Ontology specification, *The Bibliographic Ontology*. Διαθέσιμο στο: <http://bibliontology.com/specification>
- Datahub (2015). Διαθέσιμο στο: <http://datahub.io/el/about>
- Deliot, C. (2014) Publishing the British National Bibliography as Linked Open Data, *Catalogue & Index*, (174), σ. 13 – 18. Διαθέσιμο στο: [http://www.bl.uk/bibliographic/pdfs/publishing\\_bnb\\_as\\_lod.pdf](http://www.bl.uk/bibliographic/pdfs/publishing_bnb_as_lod.pdf)
- Dublin Core Metadata Element Set, Version 1.1* (2012) διαθέσιμο στο: <http://dublincore.org/documents/dces/>
- Edelstein, J., Galla, L. Li – Madeo, C., Marden, J., Rhonemus, A., Whysel, N. (2013). *Linked open data for cultural heritage: evolution of an information technology*. Διαθέσιμο στο: <http://www.whysel.com/papers/LIS670-Linked-Open-Data-for-Cultural-Heritage.pdf>
- Edoceo: Comma Separated Values (CSV) Standard File* (χ.χ.). Διαθέσιμο στο: <http://edoceo.com/utilitas/csv-file-format>
- Furrie, B. (2003). *Understanding MARC Bibliographic: Machine-Readable Cataloging*. USA: Library of Congress. Διαθέσιμο στο: <http://www.loc.gov/marc/umb/>
- Haffner, A. (2015) *GND ONTOLOGY: Namespace document. German: Deutsche Nationalbibliothek*. Διαθέσιμο στο: <http://d-nb.info/standards/elementset/gnd#>
- Hägström, M. (2004). *The role of libraries in lifelong learning: Final report of the IFLA project under the section for Public Libraries*. USA: IFLA. Διαθέσιμο στο: <http://archive.ifla.org/VII/s8/proj/Lifelong-LearningReport.pdf>
- Hannemann, J. & Kett, J. (2010). Linked data for libraries, *76ο Γενικό Συνέδριο της IFLA*, Γκέτεμποργκ, Σουηδία, 10 - 15 Αυγούστου. Διαθέσιμο στο: <http://conference.ifla.org/past-wlic/2010/149-hannemann-en.pdf>
- Hassanzadeh, O., Lim, L., Kementsietsidis, A. & Wang, M. (2009). *A Declarative Framework for Semantic Link Discovery over Relational Data*. Διαθέσιμο στο: <http://ra.ethz.ch/CDstore/www2009/proc/docs/p1101.pdf>
- Heath, T. & Bizer, C. (2011) Linked Data: Evolving the Web into a Global Data Space, *Synthesis Lectures on the Semantic Web: Theory and Technology*, 1(1), σ. 1 - 136. USA: Morgan & Claypool. Διαθέσιμο στο: <http://info.slis.indiana.edu/~dingying/Teaching/S604/LODBook.pdf>
- Heath, T. & Motta, E. (2008). Revyu.com: a reviewing and rating site for the web of data, *Journal of Web Semantics*, 6(4), σ. 266 - 273. Διαθέσιμο στο: [http://data.semanticweb.org/pdfs/iswc-aswc/2007/ISWC2007\\_SWC\\_Heath.pdf](http://data.semanticweb.org/pdfs/iswc-aswc/2007/ISWC2007_SWC_Heath.pdf)
- Hogan, A. (2014). Linked data & the semantic web standards. Διαθέσιμο στο: [http://aidanhogan.com/docs/ldmgmt\\_local\\_chapter.pdf](http://aidanhogan.com/docs/ldmgmt_local_chapter.pdf)

- IFLA: 200 Title and statement of responsibility (χ.χ.) Διαθέσιμο στο: [http://www.ifla.org/files/assets/uca/unimarc\\_updates/BIBLIOGRAPHIC/ub\\_200\\_update.pdf](http://www.ifla.org/files/assets/uca/unimarc_updates/BIBLIOGRAPHIC/ub_200_update.pdf)
- Introducing JSON (χ.χ.). Διαθέσιμο στο: <http://json.org/>
- ISO 2709: 2008 (2011). Διαθέσιμο στο: [http://www.iso.org/iso/iso\\_catalogue/catalogue\\_tc/catalogue\\_detail.htm?csnumber=41319](http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=41319)
- Janowicz , K., Hitzler, P., Adams, B., Kolas, D. & Vardeman II, C. (2014). Five stars of linked data vocabulary use, *Semantic Web Journal*, 1. Διαθέσιμο στο: <http://geog.ucsb.edu/~jano/swj653.pdf>
- Jentzsch, A., Andersson, B., Hassanzadeh, O., Stephens, S. & Bizer, C. (2009). *Enabling Tailored Therapeutics with Linked Data*. Πρακτικά 2<sup>ου</sup> Συνεδρίου Linked Data on the Web (LDOW2009). Διαθέσιμο στο: [http://events.linkedata.org/ldow2009/papers/ldow2009\\_paper9.pdf](http://events.linkedata.org/ldow2009/papers/ldow2009_paper9.pdf)
- Kobilarov, G., Scott, T., Raimond, Y., Oliver, S., Sizemore, C., Smethurst, >, Bizer, C. & Lee, E. (2009). Media Meets Semantic Web - How the BBC Uses DBpedia and Linked Data to Make Conections, *The Semantic Web: Research and Applications*. 5554, σ. 723 -737.
- Koivunen, M. R. & Miller, E. (2001). *W3C Semantic Web activity*. Διαθέσιμο στο: <http://www.w3.org/2001/12/semweb-fin/w3csw>
- Lam, K. T. (2002). *Lecture notes: Metadata for library uses*. Διαθέσιμο στο: <http://ihome.ust.hk/~lblkt/diploma/metadata/metadata.html>
- Maali, F. (2014) *SPARQLture: A more welcoming entry to SPARQL Endpoints*. Διαθέσιμο στο: [http://ceur-ws.org/Vol-1279/iesd14\\_9.pdf](http://ceur-ws.org/Vol-1279/iesd14_9.pdf)
- Marc Standards: Library of Congress (χ.χ.). Διαθέσιμο στο: <http://www.loc.gov/marc/>
- Miller, E. (1998). An introduction to the Resource Description Framework, *D-Lib Magazine*. Διαθέσιμο στο: <http://www.dlib.org/dlib/may98/miller/05miller.html#URI1>
- Miller, E. (2011). The Serials Librarian: From the Printed Page to the Digital Age, *Linked Data and Libraries*, 60 (1-4), σ. 17 -22. Διαθέσιμο στο: <http://dx.doi.org/10.1080/0361526X.2011.556427>
- Miller, P. (2010). Linked data and government, *European PSI Platform*. Διαθέσιμο στο: <http://www.epsplatform.eu/sites/default/files/ePSPlatform%20Topic%20Report%20Nο.%207%20%20Linked%20Data%20and%20Gov.pdf>
- OCLC: VIAF (2015). Διαθέσιμο στο: <http://www.oclc.org/viaf.en.html>
- OpenUp® Platform: What is the OpenUp® Platform? (χ.χ.). Διαθέσιμο στο: <http://www.tso.co.uk/our-expertise/technology/openup-platform>

- Paepcke, A., Chang, C. K., Winograd, T. & Garcia-Molina, H. (1998): Interoperability for digital libraries worldwide. *Communications of the ACM*, 41 (4), σ. 33-42. Διαθέσιμο στο: <http://dl.acm.org/citation.cfm?id=273044>
- Pan, J. Z. & Horrocks, I. (2003). *RDFS(FA) and RDF MT: Two semantics for RDFS*. Διαθέσιμο στο: <http://www.cs.ox.ac.uk/ian.horrocks/Publications/download/2003/PaHo03b.pdf>
- Papadakis, I., Kyprianos, K. & Stefanidakis, M. (2015). Linked data and libraries: The story so far, *D-Lib Magazine*, 21 (5 - 6). Διαθέσιμο στο: <http://www.dlib.org/dlib/may15/papadakis/05papadakis.html>
- Park, O. N. (2015). Development of linked data for archives in Korea, *D-Lib Magazine*, 21 (3/4). Διαθέσιμο στο: <http://www.dlib.org/dlib/march15/park/03park.html#8>
- Patel – Schneider, P. F. (2012). *OWL 2 Web Ontology Language: New features and rationale (Second Edition)*. Διαθέσιμο στο: <http://www.w3.org/TR/owl2-new-features/>
- Prud'hommeaux, E. & Seaborne, A. (2008). *SPARQL query language for RDF*, W3C. Διαθέσιμο στο: <http://www.w3.org/TR/rdf-sparql-query/>
- Sales, M. H. (2005). Cataloging and classification, *Journal of Adventist Education*. Διαθέσιμο στο: <http://circle.adventist.org/files/jae/en/jae200467021804.pdf>
- Seeley, R. (2003). The Semantic Web: The OWL has landed, *ADT Magazine*. Διαθέσιμο στο: <https://adtmag.com/article.asp?id=8144>
- Shadbolt, N., Hall, W. & Berners – Lee, T. (2006). The semantic web revisited, *IEEE Intelligent Systems*. Διαθέσιμο στο: [http://eprints.soton.ac.uk/262614/1/Semantic\\_Web\\_Revisited.pdf](http://eprints.soton.ac.uk/262614/1/Semantic_Web_Revisited.pdf)
- Simon, A., Wenz, R., Michel, V & Mascio, A. (2013). *Publishing bibliographic records on the Web of Data: opportunities for the BnF (French National Library)*. Διαθέσιμο στο: <http://eswc-conferences.org/sites/default/files/papers2013/simon.pdf>
- Sizov, S. (2007). What makes you think that? The semantic web's proof layer, *IEEE Intelligent Systems*. Διαθέσιμο στο: <http://www.ualberta.ca/~reformat/ece720w2012/papers/IEEEExplore-8.pdf>
- SKOS Simple Knowledge Organization System - Home Page* (χ.χ.) Διαθέσιμο στο: <http://www.w3.org/2004/02/skos/>
- Sompel, V. H., Lagoze, C., Nelson, M. L., Warner, S., Sanderson, R., Johnston, P. (2009). Adding eScience Assets to the Data Web. *Πρακτικά 2<sup>ου</sup> Συνεδρίου Linked Data on the Web* (LDOW2009). Διαθέσιμο στο: [http://public.lanl.gov/herbertv/presentations/ORE\\_LDOW2009\\_hvds.pdf](http://public.lanl.gov/herbertv/presentations/ORE_LDOW2009_hvds.pdf)
- Sudoc Catalogue* (2015). Διαθέσιμο στο: <http://www.sudoc.abes.fr//DB=2.1/LNG=EN/?COOKIE=U10178,Klecteurweb,D2.1,E88>

[a9c5f0-25b,I250,B341720009+,SY,A%5C9008+1,,J,H2-26,,29,,34,,39,,44,,49-50,,53-78,,80-87,NLECTEUR+PSI,R109.242.205.79,FN](http://arxiv.org/ftp/arxiv/papers/1006/1006.4567.pdf)

Taye, M. M. (2010). Understanding semantic web and ontologies: theory and applications, *Journal of Computing*, 2 (6). Διαθέσιμο στο: <http://arxiv.org/ftp/arxiv/papers/1006/1006.4567.pdf>

Vila - Suero, D., Villazón-Terrazas, B. & Gómez-Pérez, A. (2013), datos.bne.es: A library linked data dataset, *Library Hi Tech*, 31 (4), σ. 575 - 601. Διαθέσιμο στο: [http://www.semantic-web-journal.net/system/files/swj290\\_1.pdf](http://www.semantic-web-journal.net/system/files/swj290_1.pdf)

*What is a MARC record, and why it is important?* (2009). USA: Library of Congress. Διαθέσιμο στο: <http://www.loc.gov/marc/umb/um01to06.html>

Wooldridge, M. J. (2002). *An introduction to multiagent systems*. Sussex: John Wiley & Sons Ltd.

ABEKT: ABEKT 5.6 (χ.χ.). Διαθέσιμο στο: [http://abekt.ekt.gr/abekt/abekt\\_56](http://abekt.ekt.gr/abekt/abekt_56)

ABEKT: Γενικά για το ABEKT (χ.χ.). Διαθέσιμο στο: <http://abekt.ekt.gr/abekt>

*Ανοιχτά Δεδομένα: Η πρώτη ύλη για την Κοινωνία της Γνώσης* (2014). Διαθέσιμο στο: [http://www.ekt.gr/content/img/product/89312/20\\_27.pdf](http://www.ekt.gr/content/img/product/89312/20_27.pdf)

Γερόλιμος, Μ., Σίτας, Α. & Κώνστα, Ρ. (2011). *Διαχείριση τεκμηρίων στο μεταβαλλόμενο περιβάλλον πληροφόρησης*. Διαθέσιμο στο: <http://repository.edulll.gr/edulll/retrieve/2201/684.pdf>

Καραμπατάκης, Σ. (2013). Σημασιολογική αναπαράσταση βιβλιογραφικών δεδομένων στον Παγκόσμιο Ιστό: Η Βιβλιοθήκη της Βέροιας ως παράδειγμα Διασυνδεδεμένων Δεδομένων. *Διπλωματική Εργασία*. Θεσσαλονίκη: Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης.

Κυπριανός, Κ. & Παπαδάκης, Ι. (2012). *Καθιστώντας μια υπηρεσία θεματικής πλοήγησης στο διαδίκτυο συμβατή με τις τεχνολογίες των συνδεδεμένων δεδομένων*. Διαθέσιμο στο: [http://leo.hua.gr/palc2012/sites/default/files/pictures/Kyprianos\\_papadakis\\_palc2012.pdf](http://leo.hua.gr/palc2012/sites/default/files/pictures/Kyprianos_papadakis_palc2012.pdf)

Κωνσταντινίδης, Σ. Θ., Λάκκα, Χ., Μπράτσας, Χ., Παππάς, Κ., Μπαμίδης, Π. Δ. (2011). Σημασιολογική περιγραφή ψηφιακών εκπαιδευτικών αντικειμένων σε ηλεκτρονικό περιβάλλον εκπαίδευσης. *2ο Πανελλήνιο Συνέδριο Ένταξη και χρήση των ΤΠΕ στην εκπαιδευτική διαδικασία*. Διαθέσιμο στο: <http://www.etpe.gr/custom/pdf/etpe1723.pdf>

Στοΐλος, Γ. (χ.χ.). Γλώσσες αναπαράστασης γνώσης στο Σημασιολογικό Ιστό. Διαθέσιμο στο: <http://www.image.ece.ntua.gr/~gstoil/RDF-OWL.pdf>

Τσουκαλά, Α. & Κοντού, Π. (2015). *Εγχειρίδιο MARC 21: Βιβλιογραφικά δεδομένα*. Σύρος: Βιβλιοθήκη Πανεπιστημίου Αιγαίου. Διαθέσιμο στο:

<http://portal.lib.aegean.gr/cs/UserFiles/File/MARC%2021%20Bibliographic%20Greek.pdf>