



ΙΟΝΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ
ΤΜΗΜΑ ΑΡΧΕΙΟΝΟΜΙΑΣ, ΒΙΒΛΙΟΘΗΚΟΝΟΜΙΑΣ
ΚΑΙ ΜΟΥΣΕΙΟΛΟΓΙΑΣ

**ΣΥΝΔΕΔΕΜΕΝΑ ΔΕΔΟΜΕΝΑ ΚΑΙ Η ΑΞΙΟΠΟΙΗΣΗ ΤΟΥΣ ΑΠΟ ΤΙΣ
ΒΙΒΛΙΟΘΗΚΕΣ**

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

Μυροφόρα Καυζαλίδου

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες

Επιβλέπουσα διδάσκουσα: Σοφία Στάμου

(Υπογραφή)

.....

Μυροφόρα Καυζαλίδου

© Κέρκυρα 2019 – All rights reserved

Περίληψη

Αναμφισβήτητα ζούμε στην εποχή της πληροφόρησης. Όλοι λίγο ή πολύ είμαστε στη θέση να δεχόμαστε δεδομένα και να τα διαχειριζόμαστε καθημερινά, ως χρήστες του Ιστού. Στον Ιστό έχει συγκεντρωθεί ένας τεράστιος όγκος δεδομένων. Η ανάγκη για τη καλύτερη διαχείριση των δεδομένων οδήγησε στη δραματική αλλαγή στον τρόπο που αναζητούμε και δημοσιεύουμε πληροφορίες. Μέσω της πεποίθησης αυτής γεννήθηκε το όραμα του Σημασιολογικού Ιστού, ο οποίος έχει σκοπό να δομήσει όλη την συγκεντρωμένη πληροφορία μετατρέποντας τον σημερινό Ιστό από μια σειρά σημασιολογικά απομονωμένων δεδομένων, σε μια κολοσσιαία βάση δεδομένων αποθήκευσης και ανάκτησης πληροφοριών. Πάνω στις τεχνολογίες του Σημασιολογικού Ιστού πλαισιώνεται το κίνημα των Συνδεδεμένων Δεδομένων, διεθνώς Linked Data. Η συγκεκριμένη εργασία ασχολείται με το θέμα των Συνδεδεμένων Δεδομένων. Με τον όρο Συνδεδεμένα Δεδομένα αναφερόμαστε σε ένα σύνολο τεχνολογιών και πρακτικών που αφορούν την δημοσίευση των δομημένων δεδομένων στον Ιστό καθώς και την διασύνδεση τους με άλλα δομημένα δεδομένα. Προκύπτει πως η χρήση των Συνδεδεμένων Δεδομένων στο χώρο των βιβλιοθηκών και εκτενέστερα στην κοινότητα της βιβλιοθήκης, προσφέρει πολλά οφέλη, δημιουργώντας μια νέα πρόκληση για τον τρόπο που διαχειρίζονται οι βιβλιοθήκες το βιβλιογραφικό περιεχόμενο. Μια από τις εφαρμογές των Συνδεδεμένων Δεδομένων αποτελεί το μοντέλο δεδομένων BIBFRAME το οποίο παρουσιάζεται στο τελευταίο κεφάλαιο της παρούσας εργασίας.

Λέξεις Κλειδιά:

Συνδεδεμένα Δεδομένα, RDF, BIBFRAME, Σημασιολογικός Ιστός, Ιστός των Δεδομένων, Μεταδεδομένα, Τεχνολογίες Βιβλιοθηκών

Abstract

We are definitely living in the information age. All of us are able to accept data and deal with them daily as web users. A huge amount of data has been gathered on the Web. The need for better data management has led to a dramatic change in the way we're searching and publishing information. Through this need was born the Semantic Web vision, which aims to structure all the concentrated information on Web by transforming the current Web from a series of semantically isolated data into a colossal database of information storage and retrieval. The Linked Data Movement is framed on the Semantic Web technologies. This Thesis deals with the concept of Linked Data. Linked Data refers to a set of technologies and practices that relate to the publication of structured data on the Web as well as their interconnection with other structured data. It follows that the use of Linked Data in libraries and more extensively in the library community offers many benefits, creating a new challenge for the way libraries manage the bibliographic content. One of the Linked Data applications is the BIBFRAME data model, that is presented in the last chapter of this paper.

Keywords:

Linked Data, RDF, BIBFRAME, Semantic Web, Web of Data, Metadata, Library Technologies

Ευχαριστήριο Σημείωμα

“Η εργασία αυτή εκπονήθηκε υπό την επίβλεψη της κας.Σοφίας Στάμου, Επίκουρη Καθηγήτρια του τμήματος Αρχειονομίας, Βιβλιοθηκονομίας και Μουσειολογίας, του Ιονίου Πανεπιστημίου, στην οποία θα ήθελα να εκφράσω την εκτίμηση, τον σεβασμό και τις ευχαριστίες μου για την καθοδήγηση και την βοήθεια που μου προσέφερε. Επιπλέον, θα ήθελα να ευχαριστήσω την οικογένεια μου και τους φίλους μου για την πολύτιμη στήριξη τους και την αμέριστη συμπαράσταση.”

Περιεχόμενα

Κεφάλαιο 1. Εισαγωγή	7
1.1 Αντικείμενο της εργασίας	8
1.2 Συνεισφορά της εργασίας	9
1.3 Δομή της εργασίας	10
Κεφάλαιο 2. Σχετική έρευνα και Ιστορικό Πλαίσιο	12
2.1 Σχετική Έρευνα	12
2.2 Ιστορικό Πλαίσιο	16
2.2.1 Εξελικτική Πορεία του Παγκόσμιου Ιστού: Από το Web 1.0 στο Web 3.0	16
2.2.2 Ο Σημασιολογικός Ιστός	18
Κεφάλαιο 3 Συνδεδεμένα Δεδομένα	29
3.1 Εισαγωγή στα Συνδεδεμένα Δεδομένα.....	29
3.2 Ενιαία Αναγνωριστικά Πόρων URIs.....	30
3.3 RDF (Resource Description Framework)	31
3.3.1 Οργάνωση Δεδομένων σε Τριπλέτες	32
3.3.2 Λεξιλόγια του RDF	35
3.3.3 Πλεονεκτήματα του RDF	38
3.4 Συνδεδεμένα Δεδομένα και Βιβλιοθήκες	38
3.4.1 Η κατάσταση σήμερα και η ανάγκη για μετάβαση.....	39
3.4.2 Οφέλη των Συνδεδεμένων Δεδομένων.....	41
3.4.3 Εμπόδια στην αξιοποίηση των Συνδεδεμένων Δεδομένων από τις βιβλιοθήκες	42
Κεφάλαιο 4. Case Study, Bibliographic Framework Initiative (BIBFRAME).....	43
4.1 Ιστορικό Πλαίσιο του BIBFRAME.....	43
4.2 Το μοντέλο δεδομένων BIBFRAME	43

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες	
4.3 Δομή του μοντέλου BIBFRAME 1.0	44
4.4 BIBFRAME και Συνδεδεμένα Δεδομένα	47
Κεφάλαιο 5 Αποτίμηση του Έργου και Μελλοντικές Κατευθύνσεις	49
5.1 Αποτίμηση του Έργου.....	49
5.2 Μελλοντικές Κατευθύνσεις.....	51
Βιβλιογραφία	52

Κεφάλαιο 1. Εισαγωγή

Η σημερινή εποχή τεχνολογικά διαφέρει πάρα πολύ σε σχέση με τα προηγούμενα χρόνια. Ο Παγκόσμιος Ιστός περνάει σε μια νέα εποχή όσο αφορά τον τρόπο που δημοσιεύουμε την πληροφορία στον Ιστό και γίνεται όλο και περισσότερο η πιο διαδεδομένη πηγή δεδομένων και πλατφόρμα διασύνδεσης δεδομένων.

Μέχρι σήμερα ο τεράστιος όγκος δεδομένων που βρίσκεται στον Ιστό, έχει δημοσιευθεί με τρόπο που δεν είναι πολύ χρήσιμος για τους υπολογιστές καθώς δεν υπάρχει καμία σύνδεση μεταξύ των δεδομένων, δημιουργώντας προβλήματα στη διαχείριση τους αλλά και στην αξιοποίηση τους από τους ανθρώπους χρήστες. Τα Συνδεδεμένα Δεδομένα αποτελούν ένα σχετικά νέο τρόπο σκέψης όσο αφορά τη δημοσίευση των δεδομένων και σκοπός τους είναι μέσω της αξιοποίησης τους, να αντιμετωπιστούν τα προβλήματα που δημιουργούνται από τη συσσώρευση των δεδομένων που βρίσκονται στον Ιστό. Με την επίτευξη του στόχου των Συνδεδεμένων Δεδομένων οι υπολογιστές θα μπορούν να χρησιμοποιούν τα δεδομένα πιο εύκολα και, ως εκ τούτου, θα κάνουν τη ζωή μας καλύτερη. Το RDF είναι ένας τρόπος δημοσίευσης των Συνδεδεμένων Δεδομένων και επιλύει τόσο το πρόβλημα της συσσώρευσης δεδομένων όσο και το πρόβλημα της σύνδεσης δεδομένων.

Η αξιοποίηση των τεχνολογιών των Συνδεδεμένων δεδομένων ωφελεί τις βιβλιοθήκες, προσφέροντας τους νέες ευκαιρίες και μια νέα πρόκληση σχετικά με τον ιδανικό τρόπο διαχείρισης του περιεχομένου των βιβλιοθηκών. Τα Συνδεδεμένα Δεδομένα θα καθορίσουν ένα πιο ελκυστικό για τους χρήστες περιβάλλον βιβλιοθήκης, προσφέροντας τους καλύτερες υπηρεσίες που θα αντεπεξέρχονται στις ανάγκες των σύγχρονων χρηστών.

1.1 Αντικείμενο της εργασίας

Ο Παγκόσμιος Ιστός έχει αλλάξει την θεώρηση μας για τον τρόπο που διαμοιραζόμαστε της πληροφορίες, περνώντας σε μια νέα εποχή, την εποχή του Ιστού των Δεδομένων. Αντικείμενο της συγκεκριμένης εργασίας είναι η προσπάθεια να γίνει ορισμός της έννοιας των Συνδεδεμένων Δεδομένων, καθώς και άλλων εννοιών που σχετίζονται με αυτά όπως ο Σημασιολογικός Ιστός που αποτελεί το εννοιολογικό πλαίσιο στο οποίο είναι βασισμένα τα Συνδεδεμένα Δεδομένα. Πιο συγκεκριμένα αναλύεται η σημασία των Συνδεδεμένων Δεδομένων μέσω της παρουσίασης των τεχνολογιών που τα υλοποιούν, όπως το RDF, το RDFS, τα αναγνωριστικά URIs κ.λ.π. Στόχος της παρούσας εργασίας είναι η επισκόπηση των τεχνολογιών που βασίζονται τα Συνδεδεμένα Δεδομένα καθώς και η περιγραφή του τρόπου δημοσίευσης τους.

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες

Μέσω της συγκεκριμένης εργασίας διερευνάται η σημασία της αξιοποίησης των Συνδεδεμένων Δεδομένων από ένα σύγχρονο πρότυπο βιβλιοθήκης και τα οφέλη που αυτά προσφέρουν για την κοινότητα της βιβλιοθήκης.

Τέλος γίνεται μια προσπάθεια να παρουσιαστεί ο βαθμός που αξιοποιούνται σήμερα τα Συνδεδεμένα Δεδομένα για την περιγραφή βιβλιογραφικών πόρων, μέσω της μελέτης περίπτωσης σχετικά με το μοντέλο δεδομένων BIBFRAME.

1.2 Συνεισφορά της εργασίας

Σκοπός της συγκεκριμένης εργασίας είναι να εισάγει τον αναγνώστη στο ερευνητικό θέμα που επεξεργάζεται και συγκεκριμένα στη παρουσίαση της έννοιας των Συνδεδεμένων Δεδομένων. Επίσης η συγκεκριμένη εργασία παρουσιάζει μια εικόνα των τρεχουσών εξελίξεων πάνω στις οποίες αναπτύσσεται το όραμα του Σηματολογικού Ιστού και της μετάβασης από τον Ιστό των εγγράφων στον Ιστό των Δεδομένων, ένα ενιαίο χώρο όπου υπάρχει σύνδεση μεταξύ όλων των δεδομένων.

Επιπλέον έχει σκοπό να παρουσιάσει στον αναγνώστη την οργάνωση του συνόλου της γνώσης με δομημένο τρόπο, όπως προκύπτει μέσα από τη μελέτη της βιβλιογραφίας που έχει δημοσιευθεί για το συγκεκριμένο θέμα. Στόχος είναι η εργασία να αποτελέσει ένα πλήρες τεκμήριο βιβλιογραφικής επισκόπησης μέσα από την κριτική επεξεργασία και συνθετική θεώρηση των διαθέσιμων θεωρητικών και ερευνητικών δεδομένων. Δηλαδή στις εξελίξεις στον τρόπο που δημοσιεύονται τα Δεδομένα σήμερα, τις μεταξύ τους σχέσεις και τα εργαλεία που τα συνδέουν.

Τέλος μέσω της παρούσας εργασίας, γίνεται προσπάθεια για την κατανόηση της σημερινής άμεσης ανάγκης για μετάβαση των βιβλιοθηκών σε μια νέα εποχή, όπου ο τρόπος διαχείρισης του περιεχομένου των βιβλιοθηκών διαφέρει από τον τρόπο που διαχειριζόταν οι βιβλιοθήκες το υλικό τους τα προηγούμενα χρόνια. Με αυτές τις αλλαγές αποσκοπείτε η καλύτερη εξυπηρέτηση των χρηστών των βιβλιοθηκών, προσφέροντας τους νέες υπηρεσίες. Καθήκον όλων των υποψηφίων βιβλιοθηκονόμων είναι να κατανοήσουν και να μην αγνοήσουν την αλλαγή αυτή. Δηλαδή την αλλαγή στον τρόπο περιγραφής και διαχείρισης των βιβλιογραφικών πόρων, περνώντας από την εποχή της περιγραφής μεμονωμένων βιβλιογραφικών αντικειμένων στην εποχή της περιγραφής εννοιών.

Με τη αξιοποίηση των Συνδεδεμένων Δεδομένων και τη δημοσίευση των βιβλιογραφικών πόρων στον Ιστό, προσφέρονται νέες δυνατότητες για τις βιβλιοθήκες και δημιουργούνται νέες προκλήσεις, για την επίτευξη της εισόδου στη νέα εποχή. Στη συγκεκριμένη εργασία μελετάται η περίπτωση για τη δυνατότητα αξιοποίησης των Συνδεδεμένων Δεδομένων από τις βιβλιοθήκες και τα πιθανά εμπόδια που προκύπτουν. Τέλος μέσω της συγκεκριμένης

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες

εργασίας αποσκοπείτε η σύνθεση και η κριτική αξιολόγηση του ερευνητικού υλικού, για την διεξαγωγή συμπερασμάτων που αφορούν την μελλοντική επέκταση και ερευνά του συγκεκριμένου θέματος.

1.3 Δομή της εργασίας

Πιο αναλυτικά η εργασία δομείται στα εξής κεφάλαια:

Στο κεφάλαιο 2, στην πρώτη ενότητα γίνεται μια καταγραφή της ήδη υπάρχουσας βιβλιογραφίας σχετικά με το θέμα των Συνδεδεμένων Δεδομένων, της περίπτωσης αξιοποίησης των Συνδεδεμένων Δεδομένων από τις βιβλιοθήκες. Η συγκεκριμένη ενότητα αποτελεί τον παράγοντα βαρύτητας για τη διεξαγωγή συμπερασμάτων που παρουσιάζονται στο πέμπτο κεφάλαιο, μέσω της επισκόπησης της σχετικής βιβλιογραφίας. Στη δεύτερη ενότητα του ίδιου κεφαλαίου παρουσιάζεται η χρονολογική εξέλιξη του Ιστού, με σκοπό να δοθεί το ιστορικό πλαίσιο στο οποίο εξελίχθηκαν τα Συνδεδεμένα Δεδομένα. Επίσης στην ίδια ενότητα δίνονται κάποια στοιχεία για το Σημασιολογικό Ιστό που αποτελεί το πλαίσιο στο οποίο στηρίχθηκαν τα Συνδεδεμένα Δεδομένα, της αρχιτεκτονικής στην οποία είναι βασισμένος καθώς και βασικά στοιχεία για τις τεχνολογίες που χρησιμοποιεί. Τέλος στην ενότητα αυτή παρουσιάζονται και ερμηνεύονται διάφοροι σχετικοί όροι που θα χρησιμοποιηθούν στο σύνολο της εργασίας, ώστε να γίνει πιο κατανοητό το πλαίσιο του θέματος της εργασίας. Σε όλο το εύρος του κεφαλαίου υπάρχουν πολλές αναφορές σε διάφορα εργαλεία που χρησιμοποιούνται για τη διαδικασία της δημοσίευσης και σύνδεσης των Συνδεδεμένων Δεδομένων.

Το κεφάλαιο 3, αποτελεί τον πυρήνα του θέματος της εργασίας, τα Συνδεδεμένα Δεδομένα. Αρχικά γίνεται η πρώτη ουσιαστική επαφή με τα Συνδεδεμένα Δεδομένα όπου και περιγράφονται οι βασικές τους αρχές. Στη συνέχεια αναλύεται ο ορισμός των Αναγνωριστικών URIS και επεξηγείται ο σημαντικός ρόλος τους για την ύπαρξη των Συνδεδεμένων Δεδομένων. Στην επόμενη ενότητα του ίδιου κεφαλαίου γίνεται παρουσίαση και ανάλυση της βασικής πληροφοριακής δομής των Συνδεδεμένων Δεδομένων, η οποία είναι βασισμένη στις τριπλέτες των γράφων του προτύπου RDF καθώς και μια αναλυτική παρουσίαση του προτύπου RDF. Αφού γίνει πλήρης κατανόηση της τεχνολογίας των Συνδεδεμένων Δεδομένων, στις επόμενες ενότητες του τρίτου κεφαλαίου γίνεται αναφορά στα Ανοιχτά Συνδεδεμένα Δεδομένα και στην αξιοποίηση τους από τις βιβλιοθήκες όπως και στα οφέλη της αξιοποίησης τους για τους χρήστες και το προσωπικό των βιβλιοθηκών.

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες

Στο κεφάλαιο 4, μελετάται η περίπτωση για την αξιοποίηση των Συνδεδεμένων Δεδομένων από τις βιβλιοθήκες που εφαρμόζεται σήμερα, μέσω του μοντέλου δεδομένων BIBFRAME (Bibliographic Framework Initiative). Το BIBFRAME αναπτύχθηκε με πρωτοβουλία της βιβλιοθήκης του Κογκρέσου, σε συνεργασία με την Zepheira και σκοπός του BIBFRAME είναι η αξιοποίηση των συνδεδεμένων δεδομένων, ώστε να γίνουν τα βιβλιογραφικά δεδομένα πιο χρήσιμα εντός και εκτός της κοινότητας των βιβλιοθηκών.

Στο κεφάλαιο 5 που αποτελεί και το τελευταίο κεφάλαιο της εργασίας, παρουσιάζονται τα συμπεράσματα που προκύπτουν από την ενασχόληση με το συγκεκριμένο θέμα, αλλά και μερικές κατευθύνσεις για μελλοντική έρευνα.

Τέλος το υπόλοιπο μέρος της εργασίας είναι αφιερωμένο στις πηγές και στη βιβλιογραφία που χρησιμοποιήθηκε.

Κεφάλαιο 2. Σχετική έρευνα και Ιστορικό Πλαίσιο

2.1 Σχετική Έρευνα

Στην ενότητα αυτή πραγματοποιείται βιβλιογραφική επισκόπηση της υπάρχουσας βιβλιογραφίας που αφορά το θέμα της αξιοποίησης της τεχνολογίας των Συνδεδεμένων Δεδομένων. Μέσω της μελέτης της σχετικής βιβλιογραφίας προκύπτουν κάποιες βασικές θεματικές ενότητες οι οποίες καταγράφονται σε αυτή την ενότητα με σκοπό να κατηγοριοποιήσουν την επισκόπηση της σχετικής με το θέμα βιβλιογραφίας. Αρχικά πραγματοποιείται μια επισκόπηση της κατάστασης που επικρατεί σήμερα όσο αφορά την υιοθέτηση της τεχνολογίας των Συνδεδεμένων Δεδομένων από τις βιβλιοθήκες. Έπειτα παρουσιάζονται οι προκλήσεις και τα προβλήματα που εμποδίζουν τις βιβλιοθήκες να αξιοποιήσουν τα Συνδεδεμένα Δεδομένα για την ανάπτυξη των συλλογών τους. Παρουσιάζονται επίσης τα οφέλη από την υιοθέτηση των Συνδεδεμένων Δεδομένων για τις βιβλιοθήκες και τέλος μελλοντικές κατευθύνσεις του θέματος της σύνδεσης Συνδεδεμένων Δεδομένων και Βιβλιοθηκών.

Η κατάσταση σήμερα στη χρήση των Συνδεδεμένων Δεδομένων από τις Βιβλιοθήκες

Σύμφωνα με το άρθρο του Hannemann & Kett (Hannemann & Kett, 2010), οι βιβλιοθήκες σήμερα είναι σε μεγάλο βαθμό απομονωμένες όσον αφορά την ανταλλαγή δεδομένων, δεδομένου ότι τα δεδομένα που συλλέγονται είναι κατά κύριο λόγο από βιβλιοθήκες για βιβλιοθήκες. Στο άρθρο του Alemu, et.al. (Alemu, et al., 2012), μελετάται και παρουσιάζεται η σημερινή κατάσταση της χρήσης των Συνδεδεμένων Δεδομένων από τις βιβλιοθήκες. Συγκεκριμένα αναφέρεται πως στη τρέχουσα κατάσταση της χρήσης των Συνδεδεμένων Δεδομένων στις βιβλιοθήκες φαίνεται να υπάρχει κάποια ώθηση, ιδίως μεταξύ των οργανώσεων του δημόσιου τομέα, όπως είναι οι εθνικές βιβλιοθήκες που κάνουν ελεύθερα και ανοιχτά τα βιβλιογραφικά τους δεδομένα. Επίσης όπως αναφέρεται μέσα στο ίδιο άρθρο η Europeana και οι συνδεδεμένες εθνικές και περιφερειακές βιβλιοθήκες έχουν συναινέσει να προχωρήσουν στον τρόπο σύνδεσης δεδομένων. Η

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες
Βρετανική Βιβλιοθήκη έχει ήδη αναπτύξει ένα μοντέλο συνδεδεμένων δεδομένων. Ομοίως, η Europeana έχει αναθεωρήσει το μοντέλο μεταδεδομένων της από τα Σημασιολογικά Στοιχεία της Europeana (EDM), η οποία είναι πιο προσαρμοσμένη στα Συνδεδεμένα Δεδομένα.

Προκλήσεις και προβλήματα

Στο άρθρο του Hannemann και Kett (Hannemann & Kett, 2010), παρουσιάζεται η πρώτη εφαρμογή υπηρεσίας Συνδεδεμένων Δεδομένων που έχει πραγματοποιηθεί από την Εθνική Βιβλιοθήκη Γερμανίας. Η προσέγγιση αυτή στοχεύει στην αναζήτηση των προκλήσεων που πρέπει να αντιμετωπιστούν στην μετάβαση και αξιοποίηση των νέων αυτών τεχνολογιών. Τα ζητήματα που προκύπτουν λόγω της αδύναμης σύνδεσης που έχουν οι βιβλιοθήκες στην σημερινή εποχή παρουσιάζονται επίσης στο συγκεκριμένο άρθρο, με έμφαση στην αναποτελεσματική αξιοποίηση παλαιότερων τεχνολογιών και μεθόδων. Σε αντίθεση με τις κλασσικές μεθόδους και προσεγγίσεις η συγκεκριμένη έρευνα προσπαθεί να βοηθήσει στην ανάλυση των χαρακτηριστικών αυτών που μπορούν να ενισχύσουν το χαρακτηριστικό της διασύνδεσης. Κατά την ανάλυση αυτή, παρουσιάζονται τα βασικά προβλήματα που καλούνται να επιλύσουν οι βιβλιοθήκες με βάση τα τεχνικά, σχεδιαστικά, νομικά και κοινωνικά προβλήματα καθώς και τις προκλήσεις που προκύπτουν. Τέλος, η έρευνα εξειδικεύει την ανάλυσή της παρουσιάζοντας τις βιβλιοθήκες ως πολιτισμικά κέντρα και τις προκλήσεις που προκύπτουν με βάση αυτό το χαρακτηριστικό.

Σχετικά με τα τεχνικά προβλήματα, το άρθρο αναφέρει πως προκειμένου να δημιουργηθεί μια συνδεδεμένη υπηρεσία δεδομένων, απαιτούνται ορισμένες υποδομές. Δεδομένης της σχετικής καινοτομίας των Συνδεδεμένων Δεδομένων, οι τεχνολογικές επιλογές παραμένουν σε μεγάλο βαθμό σε ανάπτυξη και συχνά είναι ανεπαρκώς τεκμηριωμένες. Για τα σχεδιαστικά - εννοιολογικά προβλήματα ο συγγραφέας του άρθρου υποστηρίζει πως αντιμετωπίζονται ζητήματα με την καθιέρωση και επιλογή οντοτήτων, δημιουργώντας προβλήματα στη μοντελοποίηση των δεδομένων. Σχετικά με τα προβλήματα νομικής φύσης δημιουργούνται δύο μεγάλες προκλήσεις. Η πρώτη αφορά τα δικαιώματα

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες δημοσίευσης ενώ η δεύτερη έχει να κάνει με θέματα δανειοδότησης των Συνδεδεμένων Δεδομένων. Τέλος τα κοινωνικά προβλήματα που σχετίζονται με τις παραπάνω κατηγορίες προβλημάτων, πηγάζουν από την γενική έλλειψη εμπειρίας για την καθιέρωση των Συνδεδεμένων Δεδομένων.

Η έλλειψη εκπαίδευσης και ενημέρωσης βιβλιοθηκονόμων σχετικά με τη διασύνδεση δεδομένων επισημαίνεται και στο άρθρο των Byrne και Goddard, (Byrne & Goddard, 2010), ως το κύριο εμπόδιο για την υιοθέτηση των Συνδεδεμένων Δεδομένων από τις βιβλιοθήκες, καθώς απαιτούν πιο εξειδικευμένες γνώσεις και περιέχουν πολύπλοκες εργασίες. Το πρόβλημα αυτό, όπως αναφέρεται στο ίδιο άρθρο οξύνεται από την γενική έλλειψη ευαισθητοποίησης των βιβλιοθηκονόμων για την σύνδεση των δεδομένων αλλά και από την σχετική τους ανησυχία για την αλλαγή των βάσεων δεδομένων της βιβλιοθηκών.

Οφέλη αξιοποίησης των Συνδεδεμένων Δεδομένων για τις βιβλιοθήκες

Όπως ο Σημασιολογικός Ιστός, ο στόχος των Συνδεδεμένων Δεδομένων είναι να καταστήσει το διαδίκτυο "ενημερωμένο" και να μετατρέψει τον υπάρχοντα ιστό εγγράφων σε ένα δίκτυο δεδομένων. Στο άρθρο του Byrne και της Goddard, (Byrne & Goddard, 2010), επισημαίνονται τα βασικά οφέλη που έχουν οι βιβλιοθήκες από την αξιοποίηση των τεχνολογιών Συνδεδεμένων Δεδομένων. Τα οφέλη αυτά σχετίζονται με τη διαλειτουργικότητα και την διαχείριση των πληροφοριών. Τα οφέλη αυτά προέρχονται από τον διαμοιρασμό των βιβλιογραφικών δεδομένων μέσω της συνεργασίας με άλλους εκδότες δεδομένων, και της δημιουργίας μιας κοινής βάσης γνώσεων. Στο ίδιο άρθρο επίσης αναφέρεται πως μέσω των Συνδεδεμένων Δεδομένων και της χρήσης ελεγχόμενων λεξιλογίων μπορούν να αντιμετωπιστούν προβλήματα που σχετίζονται με την αναζήτηση πληροφοριών σε μηχανές αναζήτησης, οι οποίες θα έχουν τη δυνατότητα να προσφέρουν μέσω σημασιολογικής αναζήτησης, εξελιγμένες δυνατότητες αναζήτησης και ανάκτησης πληροφοριών. Στο άρθρο του Alemu, et al. (Alemu, et al., 2012), συμπεραίνεται πως η σωστή υιοθέτηση των Συνδεδεμένων Δεδομένων για τα μεταδεδομένα της βιβλιοθήκης, θα βοηθούσε τη βιβλιοθήκη να αναπτύξει νέες εφαρμογές και υπηρεσίες βιβλιοθήκης όπως

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες
πολύπλευρη περιήγηση, γρήγορη περιήγηση, μεταγλωττιστές μεταδεδομένων και
εντοπισμός αναδυόμενων τις τάσεις και τα λεξιλόγια χρηστών.

Εφαρμογές Συνδεδεμένων Δεδομένων σε Βιβλιοθήκες

Στο άρθρο του Tharani (Tharani, 2015), παρουσιάζεται μια υπάρχουσα μορφή εφαρμογής Συνδεδεμένων Δεδομένων σε συστήματα βιβλιοθηκών, το μοντέλο BIBFRAME, το οποίο μελετάται και παρουσιάζεται κατά κύριο λόγο στο συγκεκριμένο άρθρο. Το BIBFRAME είναι ένα αναδυόμενο πλαίσιο που αναπτύχθηκε από τη Βιβλιοθήκη του Κογκρέσου για βιβλιογραφική περιγραφή με βάση τα Συνδεδεμένα Δεδομένα. Στο αρχικό μέρος στο άρθρο του Tharani, πραγματοποιείται μια εισαγωγή στο BIBFRAME και στα Συνδεδεμένα Δεδομένα, ενώ το τελικό μέρος αυτού του εγγράφου περιγράφει και απεικονίζει μια βήμα προς βήμα διαδικασία εφαρμογής του BIBFRAME σε υπάρχοντα μεταδεδομένα βιβλιοθήκης. Στο συγκεκριμένο άρθρο συμπεραίνεται πως το BIBFRAME έχει τη δυνατότητα να αντιμετωπίσει πολλά από τα ζητήματα που πλήττουν τον ιστό από την οπτική των βιβλιοθηκών συμπεριλαμβανομένης της ακριβούς αναζήτησης, του Authority control, της ταξινόμησης, της φορητότητας και της αποσαφήνισης των δεδομένων.

Μελλοντικές κατευθύνσεις και ο ρόλος των βιβλιοθηκών

Ο Hannemann και Kett (Hannemann & Kett, 2010), στο άρθρο τους, παραθέτοντας, δεδομένα που πηγάζουν από την συνολική εμπειρία, παρουσιάζουν την μελλοντική προσέγγιση στην ανταλλαγή δεδομένων παραθέτοντας τις διασυνδέσεις που προκύπτουν. Επιπλέον παρουσιάζεται πως μπορούν να αναπτυχθούν οι κατάλληλες διαδικασίες και πως μπορούν να αναπτυχθούν οι νέες υπηρεσίες με βάση τον αυξανόμενο όγκο συλλογών μεταδεδομένων. Συνολικά μέσω της συγκεκριμένης έρευνας που παρουσιάζεται σε αυτό το άρθρο, αναδεικνύεται το θέμα της ανοιχτότητας των δεδομένων και του διαμοιρασμού μεταξύ των βιβλιοθηκών. Με βάση τα χαρακτηριστικά αυτά παρατίθενται η προσέγγιση ότι τα Συνδεδεμένα Δεδομένα μπορούν να αξιοποιηθούν στην ανάπτυξη των κατάλληλων και σύγχρονων υπηρεσιών με την προϋπόθεση να επιτευχθούν συγκεκριμένοι

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες μακροπρόθεσμοι και βραχυπρόθεσμοι στόχοι οι οποίοι παρατίθενται στο άρθρο, μέσω της απαραίτητης συνεργασίας με άλλους σχετικούς φορείς και ιδρύματα.

Στο άρθρο του Byrne και της Goddard (Byrne & Goddard, 2010) αναφέρεται πως με την τεχνογνωσία τους στην αναζήτηση, την παραγωγή μεταδεδομένων και την ανάπτυξη οντολογιών, οι βιβλιοθηκονόμοι θα πρέπει να είναι σε θέση να κατανοήσουν και να εφαρμόσουν τα σχετικά δεδομένα. Επίσης στο ίδιο άρθρο παρουσιάζεται ο ρόλος των βιβλιοθηκονόμων στα Συνδεδεμένα Δεδομένα μέσω τεσσάρων διαδικασιών. Οι διαδικασίες αυτές αφορούν την έκθεση συλλόγων και τη διαθεσιμότητα του περιεχομένου τους με τη χρήση Σημασιολογικού Ιστού, την έκδοση διαδικτυακών θησαυρών, χαρτογραφιών και υπηρεσιών, την κοινή χρήση γενικών πληροφοριών που σχετίζονται με την κάθε βιβλιοθήκη και τέλος την επιμονή και την συγκέντρωση στον ρόλο τους όσο αφορά τη χρήση των Συνδεδεμένων Δεδομένων και την ένταξη τους σε οποιαδήποτε εργασία τους μέσα στη βιβλιοθήκη.

2.2 Ιστορικό Πλαίσιο

Για την κατανόηση της σημασίας των Συνδεδεμένων δεδομένων είναι απαραίτητη η ιστορική αναδρομή της εξέλιξης του Παγκόσμιου Ιστού, από την δημιουργία του ώστε να γίνει κατανοητή η εξέλιξη του από την πρώτη του μορφή ως την μελλοντική επέκτασή του, τον Σημασιολογικό Ιστό και την όλο και περισσότερη αξιοποίηση των τεχνολογιών που αφορούν τα Συνδεδεμένα Δεδομένα που αναπτύσσεται σήμερα. Είναι σημαντικό να γίνει ανάλυση της έννοιας και της δομής του Σημασιολογικού Ιστού καθώς ο Σημασιολογικός Ιστός αποτελεί το εννοιολογικό πλαίσιο πάνω στο οποίο είναι στηριγμένα τα Συνδεδεμένα Δεδομένα. Επίσης θα γίνει επεξήγηση όρων οι οποίοι είναι απαραίτητοι για την κατανόηση και συνοχή της εργασίας στο σύνολο της.

2.2.1 Εξελικτική Πορεία του Παγκόσμιου Ιστού: Από το Web 1.0 στο Web 3.0

Η τεχνολογία του Ιστού εμφανίζεται από το δεύτερο μισό της δεκαετίας του 1980 μέσα στο έργο του Άγγλου επιστήμονα Tim Berners Lee, με την επίτευξή ενός παγκόσμιου συστήματος υπερσύνδεσης πληροφοριών (Enzer, 2018). Αυτό που οδήγησε τον Tim Berners Lee στην εφεύρεση του Παγκόσμιου Ιστού ήταν το όραμα για μια τεχνολογία που

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες

θα έδινε τη δυνατότητα στον καθένα να ανταλλάσσει πληροφορίες και ιδέες άμεσα οι οποίες θα ήταν την ίδια στιγμή προσβάσιμες από τους υπόλοιπους.

Το όνομα του Παγκόσμιου Ιστού το έδωσε ο ίδιος ο Tim Berners Lee στη τεχνολογία που είχε δημιουργήσει, όρος γνωστός ως "www". Η οργάνωση του Παγκοσμίου Ιστού γίνεται με ιστοσελίδες δηλαδή κόμβους, που συνδέονται μεταξύ τους μέσω των συνδέσμων. Η τεχνολογία που χρησιμοποιεί ο Ιστός, παρέχει τη δυνατότητα δημιουργίας *υπερκειμένου (hypertext)*, δηλαδή κείμενα των οποίων οι μεμονωμένες λέξεις συνδέονται μεταξύ τους με *συνδέσμους (links)* και προσφέρουν στον αναγνώστη διαφορετικές πορείες (Nelson, 1965). Το υπερκείμενο αποτελεί την οπτική απεικόνιση της πληροφορίας που βρίσκεται στα μέχρι τώρα συστήματα βάσεων δεδομένων. Οι «παραδοσιακές» βάσεις δεδομένων λειτουργούν υπό μία αυστηρή δομή, ενώ το υπερκείμενο δεν περιορίζεται από κάποια συγκεκριμένη δομή.

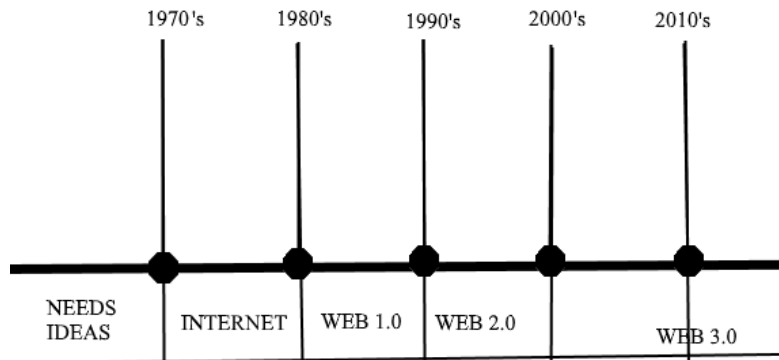
Το πρώτο στάδιο του Παγκόσμιου Ιστού αποτελεί το Web 1.0 στο οποίο χρησιμοποιήθηκε η *Hyper Text Markup Language (HTML)*, Γλώσσα Σήμανσης Υπερκειμένου. Η HTML αποτελεί τη κύρια γλώσσα σήμανσης για τις ιστοσελίδες και τα στοιχεία της οποίας είναι τα βασικά δομικά στοιχεία των ιστοσελίδων. Σκοπός του Web 1.0 αποτέλεσε η προβολή και μόνο του υλικού που βρίσκεται στις ιστοσελίδες, δηλαδή μέσω αυτού ο χρήστης είχε τη δυνατότητα της παθητικής και μόνο ανάγνωσης πληροφοριών των ιστοσελίδων (Cormode & Krishnamurthy, 2008). Αντίθετα η σημαντική μετάβαση από το Web 1.0 στο Web 2.0, δηλαδή την επόμενη μορφή του Παγκόσμιου Ιστού, πρόσφερε στο χρήστη τη δυνατότητα να συμμετέχει στο περιεχόμενο των ιστοσελίδων του Web (Berners-Lee, 2001).

Η έλευση του Web 2.0 έγινε σε μια εποχή όπου υπήρχε η δυνατότητα της ύπαρξης υπολογιστή στην οικία κάθε χρήστη. Έτσι αυξάνονται οι γνώσεις συγγραφής κώδικα και μαζί και ο αριθμός κατασκευής ιστοσελίδων. Το κύριο χαρακτηριστικό του Web 2.0 είναι η δυνατότητα που προσφέρει στο χρήστη να αλληλοεπιδρά στις σελίδες του Web. Το Web 2.0 είναι βασισμένο στην αρχιτεκτονική κατά την οποία ο χρήστης παράγει περιεχόμενο και δεν είναι απλώς ένας καταναλωτής. Η αρχιτεκτονική αυτή λειτούργησε ως ο κύριος καταλύτης της ανάπτυξης του Web 2.0 και χρησιμοποιεί τεχνολογίες οι οποίες αποτελούν μέρος των εργαλείων κοινωνικού λογισμικού/κοινωνικής δικτύωσης, *social software/networks*. Μέσω της τεχνολογίας του social software, επιτράπηκε η επικοινωνία και η συνεργασία μεταξύ δύο ή περισσότερων υπολογιστών οι οποίοι είτε βρίσκονται σε ίδια ή διαφορετική τοποθεσία (Tsetsos, 2008). Οι βασικές ιδέες που βασίζεται το Web 2.0 είναι, η παραγωγή περιεχομένου από τους χρήστες μέσω της αλληλεπίδρασής τους με τον Ιστό, η συμμετοχικότητα, η ανοικτότητα, καθώς και ο σε εξαιρετικά μεγάλος όγκος δεδομένων (Jentzsch, 2009).

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες

Με την μετάβαση από το Web 1.0 στο Web 2.0 επέρχεται η αλλαγή από τον “στατικό” ιστό στον “κοινωνικό” ιστό, ενώ με την εξέλιξη του Παγκόσμιου Ιστού επισημαίνεται η μετάβαση σε έναν “νοήμονα” ιστό το Web 3.0 ή αλλιώς Σημασιολογικό Ιστό.

Όπως έχει ήδη επισημανθεί το Web 2.0, επικεντρώνεται στη δημιουργία πληροφορίας και περιεχομένου καθώς και στον διαμοιρασμό αυτών μέσω διαφόρων εργαλείων. Από την άλλη πλευρά, η τρίτη γενιά του Παγκοσμίου Ιστού το Web 3.0, επικεντρώνεται στην οργάνωση των πληροφοριών αυτών και του περιεχομένου και στην ενσωμάτωση αυτού του δυναμικού περιεχομένου στον Παγκόσμιο Ιστό (Berners-Lee, 2008). Στο σχήμα 1 που ακολουθεί διακρίνεται η χρονολογική εξέλιξη του Ιστού και των επεκτάσεων αυτού.



Σχήμα 1: Χρονολογική Εξέλιξη του WEB

2.2.2 Ο Σημασιολογικός Ιστός

Αδιαμφισβήτητα ο Παγκόσμιος Ιστός έχει διευκολύνει τη δημοσίευση και την πρόσβαση σε έγγραφα που αποτελούν τμήμα ενός παγκοσμίου χώρου πληροφοριών. Ωστόσο σήμερα δημιουργείται η ανάγκη να αποκτήσει όλη αυτή η πληροφορία δομή και σημασιολογία ώστε να υπάρξει καλύτερη διασύνδεση των δεδομένων του Ιστού. Η ανάγκη αυτή οδήγησε στο όραμα του Σημασιολογικού Ιστού.

Ο Σημασιολογικός Ιστός (Semantic Web), αποτελεί μια επέκταση και έναν ανασχηματισμό του σημερινού Παγκόσμιου Ιστού. Κυρίως σε ότι αφορά τη δόμηση της πληροφορίας και έχει απώτερο στόχο να είναι η πληροφορία προσιτή και κατανοητή από τα υπολογιστικά συστήματα. Στο σημερινό Ιστό η αναπαράσταση των κειμένων στις σελίδες του Ιστού προορίζονται για εφαρμογή και χρήση από τους ανθρώπους. Σκοπός του Σημασιολογικού Ιστού είναι η εφαρμογή αυτή να αντικατασταθεί από σαφή αναπαράσταση και κατανόηση από τους υπολογιστές. Αυτό σημαίνει πως με τον Σημασιολογικό ιστό, οι υπολογιστές είναι σε θέση να κατανοούν τα δεδομένα τα οποία σήμερα απλώς εμφανίζουν (Berners-Lee, 2001). Η νέα αυτή επέκταση του Ιστού με τη χρήση σύγχρονων τεχνολογιών και ταχύτερων συνδέσεων έχει ως αποτέλεσμα τη προσφορά περισσότερων δυνατοτήτων και υπηρεσιών προς τους χρήστες.

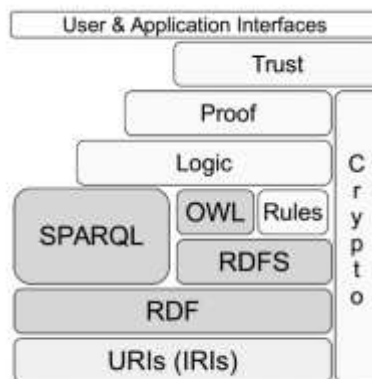
Η αξιοποίηση του Σημασιολογικού Ιστού δίνει τη δυνατότητα στους υπολογιστές να έχουν πρόσβαση σε περισσότερες πληροφορίες που μέχρι σήμερα απαιτούσαν την ανθρώπινη επέμβαση, προσοχή και χρόνο. Επιπλέον το όραμα αυτής της εξέλιξης του Ιστού στοχεύει στη μετατροπή όλων των αδόμητων εγγράφων του ιστού σε δεδομένα ιστού προκειμένου να υπάρξει καλύτερη διασύνδεση των δεδομένων του Ιστού και καλύτερη επεξεργασία των δεδομένων αυτών από ευφυείς *διαμεσολαβητές Πράκτορες (agents)* οι οποίοι θα πραγματοποιούν ανάλυση και σύγκριση και θα βγάζουν συμπέρασμα ανάλογα τη σημασία τους. Συγκεκριμένα, ο Tim Berners Lee ορίζει τον Σημασιολογικό Ιστό ως *“ένα δίκτυο δεδομένων που μπορούν να επεξεργαστούν άμεσα και έμμεσα οι μηχανές. Με την εφαρμογή του Σημασιολογικού Ιστού, οι μηχανές θα κατανοούν γρηγορότερα και θα αντιδρούν σε περίπλοκα ανθρώπινα αιτήματα που υπόκεινται στο νόημά τους.”*

Δεν είναι απαραίτητο να καταργηθεί ο σημερινός Ιστός για την υλοποίηση του Σημασιολογικού Ιστού, Αντίθετα ο Σημασιολογικός Ιστός είναι αποτέλεσμα του εμπλουτισμού του σημερινού Ιστού. *“Ο Σημασιολογικός Ιστός είναι μια επέκταση του τωρινού Ιστού, στον οποίο η πληροφορία αποκτά ορισμένο νόημα, επιτρέποντας στους υπολογιστές και στους ανθρώπους να συνεργαστούν καλύτερα μαζί”.* (Berners-Lee, 2001). Η κατασκευή του απαιτεί μια σύνδεση των τρεχουσών ιστοσελίδων σε ένα δομημένο σύνολο δεδομένων που υποδηλώνει τη σημασιολογία αυτής της συνδεδεμένης σελίδας. Γίνεται κατανοητό λοιπόν πως το κλειδί για την υλοποίηση του Σημασιολογικού Ιστού είναι η ύπαρξη δομημένου συνόλου δεδομένων, καθώς ένα δομημένο σύνολο δεδομένων σημαίνει πως τα δεδομένα που βρίσκονται σε αυτό το σύνολο είναι μηχαναγνώσιμα.

Δομή και αρχιτεκτονική του Σημασιολογικού Ιστού

Ο Σημασιολογικός Ιστός είναι δομημένος σε μια σειρά επιπέδων. Το σύστημα των επιπέδων βασίζεται σε μια “προς τα πίσω” συμβατότητα, καθώς το κάθε επίπεδο υλοποιεί τεχνολογίες που βασίζονται στο προηγούμενο επίπεδο, όπως φαίνεται και στο σχήμα 2. Τη βάση αυτής της δομής αποτελούν τα αναγνωριστικά URIs, τα οποία αποτελούν και τον πυρήνα πάνω στον οποίο οικοδομούνται τα Συνδεδεμένα Δεδομένα. Επομένως μέσω τη συγκεκριμένης αρχιτεκτονικής γίνεται κατανοητή η σύνδεση που υπάρχει μεταξύ Σημασιολογικού Ιστού και Συνδεδεμένων Δεδομένων και των κοινών τους στοιχείων.

Στο Σχήμα 2, παρουσιάζεται η διαστρωμάτωση των τεχνολογιών στις οποίες βασίζεται η εξέλιξη του Σημασιολογικού Ιστού. Επομένως η ανάπτυξη του Σημασιολογικού Ιστού συμβαδίζει με ένα ιεραρχικό σύστημα επιπέδων, στο οποίο το κάθε επίπεδο αποτελεί επέκταση του προηγούμενου (διαστρωμάτωση). Η αρχιτεκτονική του Σημασιολογικού Ιστού είναι θεμελιωμένη πάνω στην ήδη υπάρχουσα δομή του Παγκοσμίου Ιστού.



Σχήμα 2: Η αρχιτεκτονική του Σημασιολογικού Ιστού.
(Berners-Lee, T., 2000) www.w3.org/2000/Talks/1206-xml2k-tbl/slide10-0.html

Η αρχιτεκτονική του Σημασιολογικού Ιστού βασίζεται στο γνωστό μοντέλο δομής του Σημασιολογικού Ιστού που περιγράφει ο Tim Berners-Lee σε άρθρο του το 2000. Σχήμα 2.

- Στο πρωτόκολλο *HTTP* (*HyperText Transfer Protocol*), για τη μεταφορά δεδομένων ανάμεσα στον διακομιστή (server) και τον πελάτη (client).

- Στα *URIs (Uniform Resource Identifier)*, ενιαίο αναγνωριστικό των πόρων, ώστε κάθε οντότητα να ταυτοποιηθεί μοναδικά για την ονοματολογία.
- Στη κωδικοποίηση *UNICODE*, για την καθολική προσπέλαση των δεδομένων, χρησιμεύει στο να μπορούμε να γράψουμε σε πολλαπλές γλώσσες.
- Στη γλώσσα σήμανσης *XML (eXtensible Markup Language)*, για κοινή σύνταξη με καθορισμένο λεξιλόγιο από το χρήστη. Η γλώσσα XML επιτρέπει στο χρήστη να δημιουργεί τις δικές του ετικέτες για να χαρακτηρίζει τα δεδομένα του. Επίσης η γλώσσα XML αποτελεί ένα συμπλήρωμα της HTML και είναι το πρότυπο που έχει θεσπίσει η κοινότητα W3C για την αποστολή εγγράφων και αρχείων μέσα στο περιβάλλον του Διαδικτύου. Στο ίδιο επίπεδο υπάρχουν και τα XML Namespace χώροι ονομάτων και XML Schema, τα οποία χρησιμεύουν για την κατανόηση των ετικετών ενός XML εγγράφου.

Στο επόμενο επίπεδο της αρχιτεκτονικής του Σημασιολογικού Ιστού βρίσκονται καταναεμημένες οι τεχνολογίες οι τεχνολογίες που αφορούν τα μεταδεδομένα. Ορισμένα από αυτά έχουν υλοποιηθεί σε RDF και επεκτείνουν τα προηγούμενα δεδομένα σε XML. Σε επόμενη ενότητα θα γίνει ανάλυση των τεχνολογιών αυτών. Αναφορικά:

- Η τεχνολογία *RDF (Resource Description Framework)*, αποτελεί μοντέλο δεδομένων και χρησιμοποιείται ως μια γενική μέθοδος για την εννοιολογική περιγραφή των πληροφοριών, την αναπαράσταση δεδομένων και μεταδεδομένων που εφαρμόζονται στον Ιστό καθώς και για ανταλλαγή γνώσης στο Διαδίκτυο.
- Το επίπεδο όπου βρίσκονται οι γλώσσες οντολογιών (RDFS, OWL), για περιγραφή μεταδεδομένων. Το συγκεκριμένο επίπεδο αποτελεί την μεγάλη πρόκληση για την υλοποίηση του Σημασιολογικού Ιστού και είναι το σημείο που έχει πέσει το μεγαλύτερο βάρος από ερευνητική προσέγγιση αυτή τη στιγμή.
- Το επίπεδο της *Λογικής (Logic)*. Η λογική παρέχει την ικανότητά περιγραφής αυτοματοποιημένων λογικών συλλογισμών και κανόνων για την εξαγωγή αποτελεσμάτων που προκύπτουν μέσω της διαδικασίας χρήσης των πληροφοριών των δομών των οντολογιών και χρησιμοποιούνται ως επεκτάσεις αυτών ή χρησιμεύουν στη δημιουργία των δηλωτικών συστημάτων. Στο στάδιο αυτό πραγματώνεται η νοήμων διαδικασία των υπολογιστών για την λήψη αποφάσεων.
- Το επίπεδο της *Απόδειξης (Proof)*, αποτελεί τον μηχανισμό κατά τον οποίο τα αποτελέσματα που ελήφθησαν από τα δεδομένα στον Σημασιολογικό Ιστό να αποκτήσουν την ικανότητα της επιστροφής πίσω στις υποθέσεις από όπου προκλήθηκαν μέσω μηχανισμών συμπερασμάτων. Στο επίπεδο αυτό εξηγούνται οι λογικοί κανόνες.
- Το επίπεδο της *Εμπιστοσύνης (Trust)*. Το επίπεδο αυτό είναι πολύ σημαντικό για τον Σημασιολογικό Ιστό. Το επίπεδο αυτό προσφέρει την αξιοπιστία των

πληροφοριών που επεξεργάστηκαν και οδήγησαν στα αποτελέσματα και διασφαλίζει την εμπιστοσύνη των χρηστών για τους διαμεσολαβητές Πράκτορες ώστε να υπάρχει η πλήρης ασφάλεια στην εργασία της αναζήτησης και της επεξεργασίας των δεδομένων. Για την επίτευξη αυτού του επιπέδου θα πρέπει να υπάρξει συνεργασία με την τεχνολογία των *ψηφιακών υπογραφών* (*digital signatures*) και χρήση κρυπτογράφησης σε όλα τα υπόλοιπα επίπεδα της αρχιτεκτονικής του Σημασιολογικού Ιστού. Ενώ πολλοί έχουν τοποθετήσει ένα ακόμη επίπεδο στην κορυφή της αρχιτεκτονικής, η οποία θα επιτρέπει στο χρήστη να χρησιμοποιεί τις υπηρεσίες του Σημασιολογικού Ιστού, την λεγόμενη διεπαφή.

Βασικές Έννοιες του Σημασιολογικού Ιστού

Όπως γίνεται κατανοητό από την ενότητα “Αρχιτεκτονική του Σημασιολογικού Ιστού”, για να υλοποιηθεί ο Σημασιολογικός Ιστός χρειάζεται να χρησιμοποιηθούν διάφορες νέες τεχνολογίες. Στην ενότητα αυτή θα γίνει ανάλυση των μερικών τεχνολογιών του Σημασιολογικού Ιστού που αποτελούν και τις βασικές έννοιες του Σημασιολογικού Ιστού.

Διαμεσολαβητές Πράκτορες (Agents)

Με τους ταχύτατους ρυθμούς που αυξάνονται οι πληροφορίες και τα δεδομένα στον Ιστό, στο μέλλον οι άνθρωποι δεν θα μπορούν να διαχειριστούν αυτό το τεράστιο όγκο πληροφοριών που θα συγκροτεί την γνώση που βρίσκεται συγκεντρωμένη στον Ιστό. Επομένως η πραγματική δύναμη και αλλαγή που θα φέρει ο Σημασιολογικός Ιστός στον τρόπο που χρησιμοποιούμε τον Ιστό και τις πληροφορίες, θα γίνει συνειδητή όταν θα επιτευχθεί η δημιουργία προγραμμάτων και λογισμικών που θα συλλέγουν περιεχόμενο από τον Ιστό προερχόμενο από διαφορετικές πηγές και θα επεξεργάζονται τις πληροφορίες, ενώ τέλος θα ανταλλάσσουν τα αποτελέσματα με άλλα προγράμματα, απαλλάσσοντας τους χρήστες ανθρώπους από τις διαδικασίες πολύωρων αναζητήσεων στην οθόνη του υπολογιστή .

Όπως έχει ήδη αναφερθεί η υλοποίηση του Σημασιολογικού ιστού δεν θα είναι εφικτή χωρίς την ύπαρξη τέτοιων έξυπνων λογισμικών που ο Tim Berners-Lee τα έχει ονομάσει διαμεσολαβητές Πράκτορες. Οι διαμεσολαβητές Πράκτορες αποτελούν αυτό το έξυπνο λογισμικό που θα επιτρέπει τις διάφορες αυτοματοποιημένες εργασίες που απαιτούνται στον Σημασιολογικό Ιστό. Μετά την καταχώρηση εντολών από τους χρήστες οι διαμεσολαβητές Πράκτορες αυτοί θα μπορούν να χρησιμοποιήσουν τους όρους και τις

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες

σχέσεις που υπάρχουν μεταξύ τους, από τις πληροφορίες που βρίσκονται διαθέσιμες στον Ιστό ώστε να εντοπίσουν την καταλληλότερη απάντηση και να την παρουσιάσουν στους χρήστες.

Πρακτικά οι διαμεσολαβητές Πράκτορες συλλέγουν το περιεχόμενο από όλο τον Ιστό και μετά από την διαδικασία της επεξεργασίας των πληροφοριών, ανταλλάσσουν δεδομένα με άλλους διαμεσολαβητές Πράκτορες προωθώντας ένα είδος συνεργασίας. Επιπλέον οι διαμεσολαβητές Πράκτορες είναι υπεύθυνοι για τον έλεγχο της αξιοπιστίας των πηγών απ' όπου προέρχονται οι πληροφορίες που διαχειρίζονται, μέσω των ψηφιακών υπογραφών. Επιπλέον οι διαμεσολαβητές Πράκτορες χρησιμοποιούν τη τεχνολογία CRC για έλεγχο πως τα δεδομένα δεν έχουν παραποιηθεί, ενώ οι Αποδείξεις επαληθεύουν πως τα δεδομένα είναι αληθή και έγκυρα. (Kuck, 2004). Οι διαμεσολαβητές Πράκτορες θα αυξάνονται όσο το περιεχόμενο του ιστού θα γίνεται όλο και πιο μηχαναγνώσιμο. Επιπροσθέτως οι διαμεσολαβητές Πράκτορες που δεν δημιουργήθηκαν συγκεκριμένα για να συνεργάζονται, θα έχουν την ικανότητα να μεταφέρουν μεταξύ τους δεδομένα που φέρνουν σημασιολογία, προωθώντας έτσι και τον Σημασιολογικό Ιστό στο σύνολο του (Berners-Lee, et al., 2001).

Οι διαμεσολαβητές Πράκτορες διακρίνονται από τα εξής βασικά χαρακτηριστικά:

- Ορθολογισμός: Έχουν στόχους και ενεργούν με τον καλύτερο δυνατό τρόπο για την εκπλήρωσή τους.
- Αυτονομία: Μπορούν να αποφασίζουν μόνοι τους για το ποια είναι η καταλληλότερη πράξη υπό συγκεκριμένες συνθήκες, όταν υπάρχουν διαθέσιμα πολλά πιθανά σενάρια δράσης. (αντίληψη για το περιβάλλον).
- Κοινωνική Ικανότητα: Έχουν τη δυνατότητα αλληλεπίδρασης με το χρήστη, το περιβάλλον και άλλους πράκτορες όχι μόνο για μια τυπική διαδικασία ανταλλαγής δεδομένων, αλλά και για λόγους που συσχετίζουμε με την ανθρώπινη επικοινωνία, όπως για παράδειγμα η συνεργασία, η διαπραγμάτευση, ο ανταγωνισμός κ.λ.π.

Κατανοούμε λοιπόν το πόσο σημαντικός είναι ο ρόλος των Πρακτόρων για την υλοποίηση του Σημασιολογικού Ιστού.

Μεταδεδομένα

Για να γίνει κατανοητή η έννοια των *μεταδεδομένων (metadata)* στο Σημασιολογικό Ιστό, είναι σημαντικό να γίνει κατανοητή η σημασία των μεταδεδομένων και πως αυτά

χρησιμοποιούνται και περιγράφονται στο σήμερα. Τα μεταδεδομένα είναι οι πληροφορίες των δεδομένων (δεδομένα σχετικά με δεδομένα) και περιγράφουν πόρους πληροφοριών. Είναι δομημένα δεδομένα που μπορούν να αναγνωστούν από τις μηχανές υπολογιστές. (Yu, 2007)

Πιο συγκεκριμένα, τα μεταδεδομένα είναι μια συστηματική μέθοδος για την περιγραφή πόρων και με τον τρόπο αυτό βελτιώνεται η πρόσβαση σε αυτά. Είναι σημαντικό να τονιστεί η λέξη συστηματική. Στον παγκόσμιο ιστό, συστηματικό σημαίνει δομημένο και, επιπλέον, τα δομημένα δεδομένα υποδηλώνουν την αναγνωσιμότητα και την κατανόηση από τις μηχανές. Αυτό αποτελεί και τη βασική ιδέα στο όραμα του Σημασιολογικού Ιστού. Τα μεταδεδομένα είναι ακριβώς εκείνα που θα μεταδίδουν τη σημασιολογία των κανονικών δεδομένων, ενώ στην εφαρμογή τους στον Σημασιολογικό Ιστό, θα μεταδίδονται ρητά με τα άλλα δεδομένα καθώς είναι το ίδιο απαραίτητα και έχουν την ίδια αξία όσο τα άλλα δεδομένα. Τα μεταδεδομένα αποτελούν τον συνδετικό κρίκο ανάμεσα στο περιεχόμενο του πόρου και της σημασίας του περιεχομένου.

Ας εξετάσουμε ορισμένα παραδείγματα μεταδεδομένων από τον Παγκόσμιο Ιστό. Σαφώς, και ο Ιστός αποτελείται από πολλά έγγραφα ιστού. Με βάση τον ορισμό που δόθηκε για τον Παγκόσμιο Ιστό, τα μεταδεδομένα ενός εγγράφου του Ιστού είναι τα δεδομένα που χρησιμοποιήθηκαν για την περιγραφή του εγγράφου. Τα μεταδεδομένα αυτά μπορεί να περιλαμβάνουν πληροφορίες για τον τίτλο του εγγράφου, τον συντάκτη του εγγράφου και την ημερομηνία του πότε δημιουργήθηκε το παρών έγγραφο. Άλλα στοιχεία μεταδεδομένων μπορούν επίσης να προστεθούν για να περιγράψουν ένα συγκεκριμένο έγγραφο. Επίσης, η περιγραφή ενός εγγράφου του Ιστού μπορεί να γίνει με διαφορετικά στοιχεία δεδομένων από διαφορετικούς συγγραφείς ή καταλογογράφους. Το τελικό αποτέλεσμα είναι ότι τα μεταδεδομένα κάθε εγγράφου του Ιστού έχουν τη δική τους μοναδική δομή και έτσι θα γινόταν αδύνατο για έναν αυτοματοποιημένο πράκτορα να επεξεργαστεί αυτά τα μεταδεδομένα με έναν ομοιόμορφο τρόπο σε παγκόσμια κλίμακα.

Επομένως, για να διασφαλιστεί ότι τα μεταδεδομένα μπορούν να υποστούν αυτόματη επεξεργασία από μηχανήματα, απαιτείται η ύπαρξη κάποιων προτύπων μεταδεδομένων, καθώς τα σημασιολογικά δεδομένα, όπως και κάθε είδος δεδομένων χρειάζονται ένα μοντέλο οργάνωσης. Ένα τέτοιο πρότυπο μπορεί να είναι ένα σύνολο κοινών κριτηρίων και όρων για την περιγραφή των δεδομένων. Για παράδειγμα, ένα πρότυπο μεταδεδομένων μπορεί να καθορίζει ότι κάθε εγγραφή μεταδεδομένων θα πρέπει να αποτελείται από έναν αριθμό προκαθορισμένων *στοιχείων (elements)* που να αντιπροσωπεύουν κάποια συγκεκριμένα *στοιχεία ιδιότητες (attributes)* ενός πόρου (σε αυτή την περίπτωση, το έγγραφο του Ιστού) και κάθε στοιχείο (*element*) μπορεί να έχει μία ή περισσότερες τιμές. Αυτό το είδος προτύπου ονομάζεται πρότυπο μεταδεδομένων. (Yu, 2007). Ένας αριθμός χαρακτηριστικών, ονομάτων και της σημασίας τους εμπεριέχεται σε κάθε πρότυπο

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες

μεταδεδομένων ορίζονται ως στοιχεία. Τα στοιχεία αυτά πολλές φορές στο Σημασιολογικό Ιστό περιγράφουν πληροφορίες όπως, ο τύπος, η μορφή, οι φυσικές ιδιότητες και ο χώρος ύπαρξής του. Στην περίπτωση ψηφιακού υλικού και όχι συμβατικού φυσικού υλικού, συνήθως παρέχονται και άλλα μεταδεδομένα που βοηθάνε στον εντοπισμό του υλικού αυτού, όπως η ιστοσελίδα της βιβλιοθήκης κ.α. (Kasse, 2011)

Τα μεταδεδομένα χρησιμοποιούνται στην επιστήμη της πληροφόρησης για τα αρχεία των ψηφιακών βιβλιοθηκών η αναζήτηση των πηγών από τον χρήστη να γίνεται καλύτερα και γρηγορότερα. Στη συγκεκριμένη περίπτωση, για την περιγραφή των βιβλιογραφικών αρχείων, χρησιμοποιούνται τρεις κατηγορίες μεταδεδομένων: (DCMI, 2014)

- Τα περιγραφικά μεταδεδομένα είναι πληροφορίες οι οποίες περιγράφουν το περιεχόμενο ενός αρχείου, όπως για παράδειγμα οι εγγραφές καταλόγων που είναι σε μορφή MARC.
- Τα δομικά μεταδεδομένα είναι πληροφορίες οι οποίες συνδέουν ένα αρχείο με κάποια άλλα αρχεία, όπως για παράδειγμα η σύνδεση μίας εικόνας ενός βιβλίου με τις υπόλοιπες εικόνες του.
- Τα διοικητικά μεταδεδομένα είναι πληροφορίες που αφορούν τη διαχείριση ενός αρχείου, όπως για παράδειγμα τα δικαιώματα πνευματικής ιδιοκτησίας του αρχείου, η μορφή φύλαξής του, η συντήρησή ψηφιακών αρχείων κ.λ.π.
- Τα δεδομένα προέλευσης (provenance metadata) είναι μέρος των μεταδεδομένων και το Dublin Core αποτελεί το λεξιλόγιο για την αναπαράστασή τους με τη χρήση στοιχείων όπως dc:creator, dc:publisher and dc:date.

Το πρότυπο Dublin Core Metadata Initiative (DC), αποτελεί ένα παράδειγμα προτύπου μεταδεδομένων για την περιγραφή αντικειμένων και χρησιμοποιείται σήμερα σε παγκόσμια κλίμακα. Αναπτύχθηκε το Μάρτιο του 1995 στο Εργαστήριο Μεταδεδομένων που χρηματοδοτείται από το OCLC (Online Computer Library Center), Κέντρο Ηλεκτρονικών Βιβλιοθηκών και από NCSA (National Center for Supercomputing Applications), το Εθνικό Κέντρο για Εφαρμογές Υπερυπολογιστών. Η σημασιολογία των όρων του Dublin Core έχει συνδιαμορφωθεί από μια διεθνή ομάδα επαγγελματιών που προέρχεται από τον χώρο της βιβλιοθηκονομίας, της επιστήμης των υπολογιστών, των μουσείων και άλλους συναφείς τομείς.

Στόχος της δημιουργίας του Dublin Core είναι να οριστεί η σημασιολογία μέσω ενός μικρού συνόλου περιγραφικών δεικτών (core set), αποτελούμενο από καθορισμένους όρους που επιτρέπει την ανακάλυψη όλων των πόρων του Ιστού, κάνοντας είτε είναι κατανοητή από τις μηχανές είτε σε φυσική γλώσσα. Όλη η εμπειρία των βιβλιοθηκονόμων

και των αρχειοθετών έχει επενδυθεί για τη δημιουργία του ορισμού του βασικού συνόλου (core set). Ένας από τους στόχους ήταν να δημιουργηθεί ένα πιο εύχρηστο σύνολο από αυτό που καθορίστηκε για παράδειγμα στο πρότυπο *MARC (MACHine-Readable-Cataloging)*, για τις βιβλιοθήκες (Wittenburg, et al., 2008.). Στην αρχή είχε 13 στοιχεία ενώ στη συνέχεια αυξήθηκε στα 15 στοιχεία για την περιγραφή μεταδεδομένων, τα οποία ονομάζονται *Σύνολο στοιχείων μεταδεδομένων Dublin Core (DCMES)* και σκοπός τους είναι η ευκολότερη και πιο προσιτή ανακάλυψη εγγράφων στο Διαδύκτιο.

Στην συνέχεια αναφέρονται τα 15 στοιχεία του Dublin Core και η επεξήγηση τους καθώς και με ποιόν τρόπο πρέπει να χρησιμοποιούνται, όπως παρατίθενται από το Dublin Core Metadata Initiative:

1. **Τίτλος** (Ετικέτα: "Title"): Το όνομα που δίνεται στον πόρο, συνήθως από το Δημιουργό ή τον Εκδότη.
2. **Συγγραφέας ή Δημιουργός** (Ετικέτα: "Creator"): Το άτομο ή ο οργανισμός που είναι κύρια μνεία ευθύνης για την δημιουργία του πνευματικού περιεχομένου του πόρου. Για παράδειγμα, συγγραφείς σε περίπτωση γραπτών κειμένων, καλλιτέχνες, φωτογράφοι, ή εικονογράφοι σε περίπτωση οπτικών πόρων.
3. **Θέμα και Λέξεις Κλειδιά** (Ετικέτα: "Subject"): Το θέμα αυτού του πόρου. Τυπικά, το θέμα θα είναι εκφρασμένο σαν λέξεις κλειδιά ή φράσεις που περιγράφουν το θέμα ή το περιεχόμενο του πόρου. Παροτρύνεται η χρήση καθορισμένων λεξιλογίων και επίσημων σχημάτων ταξινόμησης.
4. **Περιγραφή** (Ετικέτα: "Description"): Μία αυτούσια περιγραφή του περιεχομένου του πόρου, που περιέχει περιλήψεις σε περίπτωση γραπτών κειμένων ως αντικειμένων ή περιγραφές περιεχομένων στην περίπτωση οπτικών πόρων.
5. **Εκδότης** (Ετικέτα: "Publisher"): Η υπεύθυνη οντότητα που έκανε διαθέσιμο τον πόρο στην παρούσα μορφή, όπως ένας εκδοτικός οίκος, ένα τμήμα πανεπιστημίου, ή μία συνεταιρική οντότητα.
6. **Συντελεστής** (Ετικέτα: "Contributor"): Το πρόσωπο ή ο οργανισμός που δεν προσδιορίζεται στο στοιχείο Δημιουργός που έχει συμβάλει πνευματικά σημαντικά στον πόρο αλλά η συνεισφορά του είναι δευτερεύουσα σε σχέση με οποιοδήποτε πρόσωπο ή οργανισμό που προσδιορίζεται στο στοιχείο Δημιουργός (για παράδειγμα, εκδότης, μεταφραστής και εικονογράφος).
7. **Ημερομηνία** (Ετικέτα: "Date"): Μία ημερομηνία σχετική με τη δημιουργία και τη διαθεσιμότητα του πόρου. Τέτοια ημερομηνία δεν πρέπει να συγχυστεί με αυτή που είναι στο στοιχείο Περιεχόμενο, η οποία θα συσχετιστεί με τον πόρο μόνο στο βαθμό που το πνευματικό περιεχόμενο είναι κατά κάποιον τρόπο κοντά στην ημερομηνία.

8. **Τύπος Πόρου** (Ετικέτα: "Type"): Η κατηγορία του πόρου, όπως μυθιστόρημα, ποίημα, αναφορά εργασίας, τεχνική αναφορά, έκθεση, λεξικό. Για χάρη της λειτουργικότητας, ο Τύπος πρέπει να επιλεγθεί από ένα αριθμημένο σύνολο όρων.
9. **Μορφότυπο** (Ετικέτα: "Format"): Η μορφή των δεδομένων του πόρου που χρησιμοποιείται για να πιστοποιήσει το λογισμικό *software* και πιθανόν το λειτουργικό *hardware* που μπορεί να χρειαστεί για την έκθεση ή την λειτουργία της πηγής. Για χάρη της λειτουργικότητας, η Μορφή πρέπει να επιλεγθεί από ένα αριθμημένο σύνολο όρων.
10. **Κωδικός Πόρου** (Ετικέτα: "Identifier"): Μία φράση ή αριθμός που χρησιμοποιείται για να προσδιορίσει μοναδικά τον πόρο. Παραδείγματα για πόρους δικτύου περιέχουν αναγνωριστικά URLs ή URNs (όταν χρησιμοποιείται). Άλλοι καθολικά μοναδικοί κώδικες, όπως το *International Standard Book Numbers (ISBN)* ή άλλα επίσημα ονόματα, μπορούν επίσης να χρησιμοποιηθούν για αυτό το στοιχείο.
11. **Πηγή** (Ετικέτα: "Source"): Πληροφορίες για έναν δεύτερο πόρο από τον οποίο ο παρών πόρος απορρέει. Συνιστάται περισσότερο στην πράξη να χρησιμοποιείται το στοιχείο Σχέση αντί αυτού. Για παράδειγμα είναι δυνατόν να χρησιμοποιείται μία Πηγή του 1603 στην περιγραφή μίας διασκευής του 1996 ενός έργου του Σαίξπηρ, αλλά είναι προτιμότερο να χρησιμοποιηθεί η Σχέση "IsBaseOn" με αναφορά σε έναν ξεχωριστό πόρο του οποίου η περιγραφή περιέχει την Ημερομηνία 1603.
12. **Γλώσσα** (Ετικέτα: "Language"): Η γλώσσα του πνευματικού περιεχομένου του πόρου.
13. **Σχέση** (Ετικέτα: "Relation"): Ένα αναγνωριστικό ενός δεύτερου πόρου και η σχέση του με τον παρόντα πόρο. Αυτό το στοιχείο επιτρέπει τη δήλωση συνδέσμων μεταξύ σχετικών πόρων και περιγραφών πόρων. Παράδειγμα, να περιέχει μια έκδοση μίας εργασίας (IsVersionOf), μία μετάφραση μίας εργασίας (IsBasedOn), ένα κεφάλαιο ενός βιβλίου (IsPartOf) και μια διαμόρφωση των δεδομένων σε μία εικόνα (IsFormatOf).
14. **Κάλυψη** (Ετικέτα: "Coverage"): Τα χαρακτηριστικά χώρου ή χρόνου του πνευματικού περιεχομένου του πόρου. Η Κάλυψη χώρου αναφέρεται στην υλική περιοχή με χρήση συντεταγμένων (π.χ. γεωγραφικό μήκος και πλάτος) ή θέτει ονόματα που είναι από μία καθορισμένη λίστα. Η χρονική κάλυψη αναφέρεται στο πότε ο πόρος σχετίζεται και όχι στο πότε δημιουργήθηκε ή έγινε διαθέσιμος.
15. **Δικαιώματα Χρήσης** (Ετικέτα: "Rights"): Μία δήλωση δικαιωμάτων χρήσης, έναν κωδικό που οδηγεί σε μία δήλωση δικαιωμάτων χρήσης του πόρου.

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες

Ανακεφαλαιώνοντας λοιπόν, τα μεταδεδομένα:

- Κάνουν σαφή την πληροφορία που είναι αόριστη “making explicit the implicit”
- Εκθέτουν την πληροφορία για αναζήτηση, επεξεργασία και ενοποίηση
- Τα μεταδεδομένα περιγραφών είναι τα δεδομένα που περιγράφουν το περιεχόμενο και τη σημασία των πηγών
- Οι Οντολογίες στα μεταδεδομένα: προσφέρουν την δυνατότητα να γίνονται κατανοητά σε πολλαπλές γλώσσες καθώς, διαμοιράζονται κοινά λεξιλόγια και καθιστούν διαμοιραζόμενη και κοινή τη κατανόηση ενός πεδίου ενδιαφέροντος. Οι οντολογίες στα μεταδεδομένα είναι απαραίτητες για ανταλλαγή και αναζήτηση
- Συμπερασματολογία: Εξαγωγή καινούργιας γνώσης μέσω μεταδεδομένων από την υπάρχουσα γνώση

Η αξιοποίηση του Σημασιολογικού Ιστού στη περιγραφή βιβλιογραφικών δεδομένων

Την τελευταία δεκαετία, η σημασιολογική περιγραφή και έκφραση της πληροφορίας, καθώς και οι τρόποι αλληλεπίδρασης με τα πληροφοριακά συστήματα έχουν εξελιχθεί. Οι πλατφόρμες που υποστηρίζουν την έκφραση και επικοινωνία των βιβλιογραφικών δεδομένων είναι απαραίτητο να προχωρήσουν στα νέα πρότυπα όπως το RDF (Resource Description Framework) και να ανασχεδιαστούν με βάση τα νέα δεδομένα.

Σαν ιδέα ο Σημασιολογικός Ιστός αποτελεί την επέκταση του σημερινού Ιστού με την λογική πως η δημοσιευμένη πληροφορία θα περιέχει δεδομένα τα οποία είναι κοινά προς όλους και αναγνώσιμα από χρήστες και από μηχανές. Ο σημασιολογικός ιστός βασίζεται σε παλιότερες τεχνολογίες όπως τα URIs την XML, αλλά επεκτείνεται και σε νεότερες τεχνολογίες όπως ενδεικτικά το RDF, RDF Schema και την OWL.

Προκύπτει πως η εφαρμογή του Σημασιολογικού Ιστού σκοπεύει στην αξιοποίηση μεταδεδομένων τα οποία θα είναι κοινά προς όλους και τα οποία θα παρέχονται μέσω της δημοσιευμένης πληροφορίας στον παγκόσμιο Ιστό. Τα δεδομένα που προκύπτουν είναι αναγνώσιμα και επεξεργάσιμα από χρήστες καθώς και από μηχανές. Τα συνδεδεμένα δεδομένα αποτελούν ένα σημαντικό εργαλείο για τον εντοπισμό, την χρήση και γενικά την διαδικασία της καλύτερης αξιοποίησης και επεξεργασίας της πληροφορίας στα πλαίσια ενός ενιαίου χώρου. Οι προσπάθειες της υλοποίησης του έργου του Σημασιολογικού Ιστού οδήγησαν στην δημιουργία διαφόρων εργαλείων όπως το RDF.

Κεφάλαιο 3 Συνδεδεμένα Δεδομένα

3.1 Εισαγωγή στα Συνδεδεμένα Δεδομένα

Ο Σημασιολογικός Ιστός στοχεύει στην εξέλιξη του Παγκοσμίου Ιστού σε έναν χώρο όπου η σημασία των δεδομένων θα είναι το ίδιο καλά κατανοητή και από τους ανθρώπους και από τις μηχανές (Berners-Lee, et al., 2001). Ωστόσο ο Σημασιολογικός Ιστός δεν αφορά μόνο τη τοποθέτηση δεδομένων στο Διαδίκτυο. Πρόκειται για τη δημιουργία συνδέσμων, οι οποίοι θα δίνουν τη δυνατότητα στους χρήστες και στις μηχανές να μπορούν να εξερευνήσουν τον Ιστό των δεδομένων. Όταν υπάρχουν οι σύνδεσμοι αυτοί, ο ρόλος των Συνδεδεμένων Δεδομένων είναι να εντοπίζουν περισσότερα σχετικά δεδομένα.

Όπως έχει αναφερθεί στο προηγούμενο κεφάλαιο, τα βασικά δομικά στοιχεία του Ιστού του υπερκειμένου είναι τα έγγραφα HTML (HyperText Markup Language), που συνδέονται με μη προσαρμοσμένους υπερσυνδέσμους. Τα Συνδεδεμένα Δεδομένα ωστόσο είναι βασισμένα σε έγγραφα που περιέχουν δεδομένα δομημένα σε RDF (Klyne & Carroll, 2006). Οι συνδέσεις στα Συνδεδεμένα Δεδομένα γίνονται με RDF δηλώσεις (*RDF statements*) που συνδέουν αυθαίρετα αντικείμενα στον κόσμο. Το αποτέλεσμα που προκύπτει από τις συνδέσεις αυτές είναι η δημιουργία ενός Ιστού Δεδομένων, ο οποίος μπορεί να ορισθεί πιο συγκεκριμένα ως Ιστός αντικειμένων ο οποίος περιγράφεται με δεδομένα στο Ιστό. Οι RDF δηλώσεις θα αναλυθούν στην ενότητα που είναι αφιερωμένη στο πρότυπο RDF παρακάτω.

Ο όρος Συνδεδεμένα Δεδομένα (Linked Data) αναφέρεται σε ένα σύνολο βέλτιστων πρακτικών για τη δημοσίευση και διασύνδεση δομημένων δεδομένων στον Παγκόσμιο Ιστό. Το 2006 ο Tim Berners Lee, δημοσίευσε, τέσσερις βασικούς κανόνες που αφορούν τη δημοσίευση των Συνδεδεμένων Δεδομένων:

- Κανόνας 1: Για την ταυτοποίηση οποιουδήποτε αντικειμένου (οντότητας) θα πρέπει να γίνεται χρήση αναγνωριστικών URIs. Ο συγκεκριμένος κανόνας είναι κοινός με τις αρχές του Σημασιολογικού Ιστού.
- Κανόνας 2: Πρέπει να γίνεται χρήση των HTTP URIs, ώστε να αποτελούν επισκέψιμες σελίδες για τους χρήστες, όπου ο χρήστης θα μπορούν να αναζητήσουν μια περιγραφή των αντικειμένων, που ταυτοποιείται μέσω των HTTP URIs.
- Κανόνας 3: Όταν κάποιος ανατρέχει σε διεύθυνση ενός URI, πρέπει να προσφέρεται χρήσιμη πληροφορία μέσω της χρήσης των προτύπων (RDF, SPARQL).

- Κανόνας 4: Θα πρέπει να συμπεριλαμβάνονται σύνδεσμοι προς άλλες οντότητες, οι οποίες επίσης αναγνωρίζονται με αντίστοιχα URIs, με σκοπό την ανακάλυψή όλο και περισσότερων οντοτήτων.

Οι τέσσερις κανόνες που παραθέτει ο Tim Berners-Lee, είναι γνωστοί ως οι Αρχές για τα Συνδεδεμένα Δεδομένα και εκφράζουν ακριβώς την πεποίθηση αυτή και αποτελούν την αφετηρία του κινήματος των Συνδεδεμένων Δεδομένων (*Linked Data movement*) (Stefanidakis, et al., 2015). Για την υλοποίηση των τεσσάρων Αρχών των Συνδεδεμένων Δεδομένων είναι απαραίτητο το πρωτόκολλο HTTP, ως ενιαίος μηχανισμός πρόσβασης παγκόσμιας κλίμακας καθώς και νέες τεχνολογίες όπως το πρότυπο RDF για τη δομημένη περιγραφή οντοτήτων καθώς και η γλώσσα SPARQL ώστε να γίνεται εφικτή η αναζήτηση αυτών.

Τα Συνδεδεμένα Δεδομένα δεν αποτελούν ένα διαφορετικό σύνολο από τον Σημασιολογικό Ιστό, αντίθετα ορίζονται ως ένα σύνολο βέλτιστων πρακτικών που χρησιμεύουν ώστε να γίνει δυνατή η υλοποίηση του Σημασιολογικού Ιστού πάνω στον Παγκόσμιο Ιστό. Οι αρχές των Συνδεδεμένων Δεδομένων υποστηρίζουν ακριβώς αυτή τη πεποίθηση, πως για τη δημοσίευση και τη σύνδεση δεδομένων πρέπει να χρησιμοποιηθούν οι υποδομές του τωρινού Ιστού καθώς επίσης να διατηρηθεί η αρχιτεκτονική και τα πρότυπα του. Σήμερα έχει πραγματοποιηθεί εφαρμογή αυτών των βέλτιστων πρακτικών για την δημοσίευση σημασιολογικών δεδομένων καθώς και για τη δημιουργία συνδέσμων μεταξύ των δεδομένων από διαφορετικές πηγές.

3.2 Ενιαία Αναγνωριστικά Πόρων URIs

Για τον εντοπισμό στοιχείων στον Ιστό χρησιμοποιούμε αναγνωριστικά, γνωστά ως URIs (*Universal Resource Identifiers*). Τα αναγνωριστικά URIs αποτελούν τη βάση και το πρώτο επίπεδο του Σημασιολογικού Ιστού, πάνω στο οποίο στηρίζονται και όλες οι λειτουργίες των σημασιολογικών εφαρμογών Σχήμα 2. Ο Σημασιολογικός Ιστός βασίστηκε στο προϋπάρχον πρότυπο των ενιαίων αναγνωριστικών URI και δεν ήταν αναγκαία η δημιουργία ενός καινούργιου πρότυπου. Τα URIs αποτελούν το πρώτο επίπεδο στο οποίο βασίζεται η λειτουργία των Σημασιολογικών εφαρμογών (Stefanidakis, et al., 2015) και αποτελούν ένα πιο γενικό τρόπο για τη ταυτοποίηση κάθε οντότητας που βρίσκεται στον κόσμο.

Στο Παγκόσμιο Ιστό, όπου γίνεται σύνδεση μεταξύ εγγράφων, για την αναγνώριση των εγγράφων και της τοποθεσίας τους, χρησιμοποιούνται τα *URLs* (*Uniform Resource Locator*). Στο Σημασιολογικό Ιστό όπου δεν υπάρχει μόνο σύνδεση εγγράφων αλλά και εννοιών, χρησιμοποιούνται τα Ενιαία Αναγνωριστικά Πόρων URIs. Ένα URL αποτελείται

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες

από μια ενότητα που προσδιορίζει τον υπολογιστή και τη διεύθυνση της σελίδας στο διαδίκτυο. Τα URIs βασίζονται στην ίδια λογική, αλλά δεν χρησιμοποιούνται ως μέσο εντοπισμού μιας διεύθυνσης ενός σημείου αλλά ως αναγνωριστικά οντοτήτων και εννοιών.

Τα Συνδεδεμένα Δεδομένα βασίζονται σε δύο τεχνολογίες που είναι θεμελιώδεις για τον Ιστό. Τη τεχνολογία των Ενιαίων Αναγνωριστικών Πόρων URIs (Berners-Lee et al., 2005) και το Πρωτόκολλο Μεταφοράς Υπερκειμένου HTTP (Berners-Le, et al., 2004). Ενώ οι Ενιαίοι Εντοπιστές Πόρων (URLs) ταυτοποιούνται ως διευθύνσεις για έγγραφα και άλλες οντότητες που μπορούν να εντοπιστούν στον Ιστό, τα ενιαία αναγνωριστικά πόρων URIs παρέχουν ένα πιο γενικό μέσο για την ταυτοποίηση οποιασδήποτε οντότητας που υπάρχει στο κόσμο. Είναι σημαντικό να κατανοήσουμε ότι χρησιμοποιώντας URIs, υπάρχει δυνατότητα να ταυτοποιήσουμε τόσο ένα αντικείμενο (του φυσικού κόσμου), όσο και ένα διαδικτυακό έγγραφο το οποίο περιγράφει το αντικείμενο (Sauerman & Cyganiak, 2008).

Βασισμένο στη γενική θεώρηση που χαρακτηρίζει τα URI, στο Σημασιολογικό Ιστό το RDF χρησιμοποιεί αναφορές URI (references-URIs) έτσι ώστε να πραγματοποιεί αναγνώριση με μοναδικότητα των εκφράσεων μιας *τριπλέτας* (*triples*). Μία αναφορά URI αποτελείται από ένα URI και ένα αναγνωριστικό το οποίο μπαίνει στο τέλος του URI. Τέτοιο παράδειγμα είναι το URI: <http://www.example.org/cities#Corfu> το οποίο αποτελείται από το <http://www.example.org/cities> και το αναγνωριστικό Corfu τα οποία διαχωρίζονται με τη δίεση (#). Οι αναφορές URI μπορούν να περιέχουν UNICODE ώστε τα URIs να κάνουν δυνατή τη χρήση πολλαπλών γλωσσών. Τα URIs επομένως βοηθούν το RDF ώστε να περιγράφει οποιοδήποτε δεδομένο και να εκφράζει τη μεταξύ τους σχέση.

Μετάπειτα εμφανίστηκαν και άλλες μορφές URIs, όπως για παράδειγμα τα info URIs, τα οποία υποστηρίχθηκαν από μεγάλα ιδρύματα βιβλιοθηκών που ήθελαν να μεταφέρουν την πληροφορία των τεκμηρίων τους στον παγκόσμιο ιστό (Stefanidakis, et al., 2015). Τέτοια URIs έχουμε ήδη εντοπίσει στη Βιβλιοθήκη του Κογκρέσου:

info:lc/authorities/sh2002000569

3.3 RDF (Resource Description Framework)

Το RDF αποτελεί το πρότυπο που συνιστά (recommendation) ο οργανισμός W3C για τη δημοσίευση δεδομένων στον Ιστό και αποτελεί μία από τις βασικές τεχνολογίες του Σημασιολογικού Ιστού. Το RDF είναι μια από τις κύριες τεχνολογίες των Συνδεδεμένων Δεδομένων. Το πρότυπο RDF βρίσκεται τοποθετημένο στο δεύτερο επίπεδο της αρχιτεκτονικής του Σημασιολογικού Ιστού Σχήμα 2, επομένως και χρησιμοποιεί τα αναγνωριστικά URIs, που βρίσκονται στη βάση του Σχήματος, για τη μονοσήμαντη ονομασία των οντοτήτων ενώ αποτελεί τη βάση για όλα τα ανώτερα σημασιολογικά

επίπεδα (Stefanidakis, et al., 2015). Συχνά αναφερόμαστε στο πρότυπο *RDF (Resource Description Framework)* ως τη γλώσσα που χρησιμοποιείται για τη περιγραφή των δεδομένων στο Σημασιολογικό Ιστό. Στη πραγματικότητα το RDF, είναι ένα μοντέλο οργάνωσης πληροφοριών (data model) (Stefanidakis, et al., 2015) και συγκεκριμένα χρησιμοποιείται για τη περιγραφή της δομημένης πληροφορίας που βρίσκεται στον Σημασιολογικό Ιστό, χρησιμοποιώντας ένα μοντέλο τριπλετών (triples) Σχήμα 3. Μια τριπλέτα αποτελείται από τρία στοιχεία Υποκείμενο-Κατηγορημα-Αντικείμενο (Subject-Predicate-Object) (Allemang & Hendler, 2008). Το RDF στοχεύει κατά κύριο λόγο στην προσθήκη μεταδεδομένων στα δεδομένα που δημοσιεύονται στον Ιστό, για παράδειγμα, το RDF μπορεί να χρησιμοποιηθεί για την προσθήκη δομημένων πληροφοριών, όπως ο δημιουργός, η ημερομηνία δημιουργίας και ο τίτλος.

Η σχέση RDF και Σημασιολογικού Ιστού, συνοψίζεται σε ορισμένα βασικά σημεία (Yu, 2007):

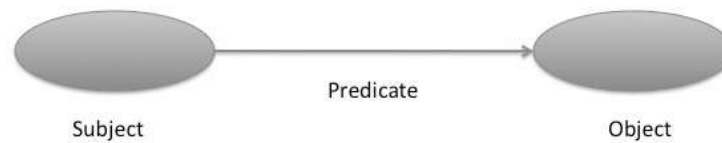
- Το RDF είναι το βασικό δομικό στοιχείο για την υποστήριξη του Σημασιολογικού Ιστού.
- Το RDF είναι στο Σημασιολογικό Ιστό ότι η HTML στον Παγκόσμιο Ιστό.
- Το RDF παρέχει μια βάση για την κωδικοποίηση, την ανταλλαγή και την επαναχρησιμοποίηση των δομημένων μεταδεδομένων.
- Το RDF είναι δομημένο. δηλ. είναι κατανοητό από μηχανές. Οι μηχανές μπορούν να κάνουν χρήσιμες πράξεις με τις γνώσεις που εκφράζονται στο RDF.
- Το RDF επιτρέπει τη διαλειτουργικότητα μεταξύ των εφαρμογών που ανταλλάσσουν μηχανικά κατανοητές πληροφορίες στον Ιστό.
- Το RDF Schema (RDFS), αποτελεί τη γλώσσα περιγραφής του λεξιλογίου του RDF

3.3.1 Οργάνωση Δεδομένων σε Τριπλέτες

Η δομή του RDF αποτελείται από *τριπλέτες (triples)*. Μια τριπλέτα αποτελείται από τρία μέρη: το υποκείμενο, το αντικείμενο και το κατηγορημα. Μια τριπλέτα εκφράζει μια σύνδεση μεταξύ δύο πόρων. Το υποκείμενο και το αντικείμενο αποτελούνται από τα URIs ενώ το κατηγορημα, που καλείται επίσης και ιδιότητα, περιγράφει τη μεταξύ τους σχέση. Το σύνολο πολλών τριπλετών καλείται γράφημα RDF. Οι συνδέσεις αυτές του RDF αποτελούν το Σημασιολογικό Ιστό (Γιατράς, 2011). Κάθε τριπλέτα αποτελεί μια δήλωση (statement) της σχέσης μεταξύ των πράγματων που υποδηλώνονται από τους κόμβους που συνδέει. Το υποκείμενο αποτελεί το θέμα της τριπλέτας και είναι ο πόρος που θέλουμε να περιγράψουμε, το αντικείμενο είναι ο πόρος που θέλουμε να συνδέσουμε με το θέμα και

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες

το κατηγορήμα παρέχει τη σχέση μεταξύ αυτών των δύο πόρων. Σχήμα 3. Ενώ κάθε γράφος RDF αποτελείται από ένα σύνολο τριάδων RDF.



Σχήμα 3 Αναπαράσταση απλού γραφήματος RDF, Υποκείμενο-Κατηγορούμενο-Αντικείμενο.

Το RDF κωδικοποιεί δεδομένα με τη χρήση τριπλετών (υποκείμενο-κατηγορήμα-αντικείμενο). Το θέμα και το αντικείμενο μιας τριπλέτας είναι και τα δύο URI που προσδιορίζουν το καθένα έναν πόρο ή ένα URI. Το κατηγορήμα καθορίζει τον τρόπο με τον οποίο σχετίζονται το θέμα και το αντικείμενο και εκπροσωπείται επίσης από ένα URI.

Για παράδειγμα, σε μια τριπλέτα RDF δηλώνεται ότι δύο άτομα, Α και Β, το καθένα που προσδιορίζεται από ένα URI, σχετίζονται από το γεγονός ότι το υποκείμενο-Α γνωρίζει το αντικείμενο-Β. Ομοίως, μια τριπλέτα RDF μπορεί να δηλώνει τη σχέση ενός ατόμου Γ σε ένα επιστημονικό άρθρο Δ που βρίσκεται σε μια βιβλιογραφική βάση δηλώνοντας ότι ο Γ είναι ο συγγραφέας του Δ. Με αυτό τον τρόπο μπορούν να συνδεθούν δυο πόροι οι οποίοι βρίσκονται σε διαφορετικά σύνολα δεδομένων στον Ιστό, επιτρέποντας τα δεδομένα που βρίσκονται στη μια πηγή δεδομένων να συνδέονται με αυτά σε μια άλλη πηγή, δημιουργώντας έτσι έναν ιστό δεδομένων. Τη συγκεκριμένη σύνδεση στοιχείων προερχόμενων από διαφορετικά σύνολα δεδομένων, που δημιουργεί μια τριπλέτα RDF, μπορούμε να την συγκρίνουμε με τις συνδέσεις υπερκειμένου που υπάρχουν στον Ιστό των εγγράφων (Heath, et al., n.d.). Στο σχήμα 4 δίνεται ένα παράδειγμα μιας δήλωσης RDF, σε μορφή υποκειμένου-κατηγορήματος-αντικειμένου (Heath, et al., n.d.).

```
Subject:http://data.linkedmdb.org/resource/film/77
Predicate:http://www.w3.org/2002/07/owl#sameAs
Object:http://dbpedia.org/resource/Pulp_Fiction_%28film%29
```

Σχήμα 4 Παράδειγμα δήλωσης RDF (Heath, et al., n.d.)

Για τη σύνταξη μιας τριπλέτας υπάρχουν διαφορετικοί τρόποι:

- RDF/XML: Αποτελεί τον αρχικό τρόπο σύνταξης της RDF. Η RDF/XML παρόλο που είναι φιλική με τις μηχανές δεν είναι με τους ανθρώπους. (Manola & Miller, 2014)
- RDFa: Μια άλλη μορφή σύνταξης RDF αποτελεί η RDFa, η οποία προσφέρει τη δυνατότητα ενσωμάτωσης του RDF σε γλώσσα HTML και XML. Επίσης επιτρέπει στις μηχανές αναζήτησης τη συλλογή δεδομένων RDF από τον Ιστό ώστε να εμπλουτίσουν τα αποτελέσματα αναζήτησης (Manola & Miller, 2014).
- Turtle: Στη συγκεκριμένη μέθοδο σύνταξης RDF, περιλαμβάνονται 4 γλώσσες (Manola & Miller 2014):
 - N-Triples: Προσφέρει τον κύριο τύπο για τη διαμόρφωση του RDF.
 - Turtle: Η οποία επεκτείνει τον κύριο τύπο της σύνταξης (N-Triples), ώστε να ώστε να γίνει αναγνώσιμο.
 - Trig: Αποτελεί μια επέκταση της Turtle.
 - N-Quads: Αποτελεί μια επέκταση της N-Triples για την ανταλλαγή RDF συνόλων δεδομένων.
- JSON-LD: Η *JSON-LD (JavaScript Object Notation for Linked Data)*, αποτελεί μια συντακτική μέθοδο για γράφους¹⁵ και σύνολα δεδομένων RDF. Χρησιμοποιείται για τη μετατροπή εγγράφων από JSON σε RDF (Manola & Miller 2014). Η δομή της είναι σχετικά απλή και έτσι μπορεί να γίνει κατανοητή και από τις μηχανές αναζήτησης. Η JSON-LD δημιουργήθηκε κυρίως για να παρουσιάζει πληροφορίες με τη χρήση λεξιλογίων συνδεδεμένων δεδομένων.
- SPARQL: Η SPARQL χρησιμοποιείται για την ανάκτηση δεδομένων RDF όπως και για την αναζήτηση δεδομένων στο Σημασιολογικό Ιστό. Η SPARQL χρησιμοποιείται επίσης για τη διατύπωση ερωτημάτων (Prud'hommeaux & Seaborne 2008).

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες

Η αναπαράσταση σε μορφή RDFa, υποστηρίζεται ενεργά από τις μεγάλες μηχανές αναζήτησης, τα on-line καταστήματα και τις εφαρμογές κοινωνικής δικτύωσης και βασίζεται στα λεξιλόγια που παρέχει το schema.org. Ως μορφή, βέβαια, μπορεί να υποστηρίξει οποιαδήποτε εφαρμογή, αρκεί να υπάρχει το κατάλληλο λεξιλόγιο (Stefanidakis, et al., 2015).

Για παράδειγμα, σε ένα βιβλιοθηκονομικό σύστημα Koha, ο χρήστης βλέπει τα διαθέσιμα αντίτυπα ενός βιβλίου στην υποθετική διεύθυνση <http://op.ac/891>, Σχήμα 5, μια εφαρμογή που μπορεί να φιλτράρει την πληροφορία RDFa («RDFa distiller») θα εξαγάγει τις παρακάτω τριπλέτες Σχήμα 6. Με άλλα λόγια, υπάρχει μια «προσφορά» ο ανώνυμος κόμβος _:1) του βιβλίου της σελίδας αυτής. Η προσφορά είναι διαθέσιμη, δανείζεται και ανήκει στην κεντρική βιβλιοθήκη με αριθμό "938.301 ΘΕΟ" (Το παράδειγμα εξαγωγής τριπλετών εφαρμόζεται μόνο για το πρώτο αντίτυπο):

Item type *	Current location *	Call number *	Copy number *	Status *
book	Κεντρικό Παράρτημα	938.301 ΘΕΟ (Browse shelf)	1	Available
book	Κεντρικό Παράρτημα	938.301 ΘΕΟ (Browse shelf)	1	Available
book	Κεντρικό Παράρτημα	938.301 ΘΕΟ (Browse shelf)	1	Available

Σχήμα 5 Εφαρμογή βιβλιοθήκης, τι είναι ορατό στο χρήστη σε ένα opac (Stefanidakis, et al., 2015).

```
@prefix schema: <http://schema.org/> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix gr: <http://purl.org/goodrelations/v1#> .

_:1 rdf:type schema:Offer;
  schema:seller "Κεντρικό Παράρτημα";
  schema:itemOffered <http://op.ac/891#record>;
  schema:businessFunction gr:LeaseOut;
  schema:sku "938.301 ΘΕΟ";
  schema:availability schema:InStock .
```

Σχήμα 6 Παράδειγμα εξαγωγής τριπλετών του αντιτύπου (Stefanidakis, et al., 2015).

3.3.2 Λεξιλόγια του RDF

Οντολογίες

Η Οντολογία εννοιολογικά σχετίζεται με το λεξιλόγιο, ενώ στην επιστήμη της πληροφόρησης μια οντολογία αποτελεί ένα ελεγχόμενο λεξιλόγιο. Το ελεγχόμενο λεξιλόγιο εκφράζεται με γλώσσες αναπαράστασης οντολογιών. Οι οντολογίες έχουν καθιερωθεί ως δομημένα πλαίσια για την οργάνωση πληροφορίας και χρησιμοποιούνται ως μια μορφή αναπαράστασης γνώσης για τον κόσμο. Αυτή η τυπική αναπαράσταση γνώσης ως ένα σύνολο εννοιών, σχέσεων και ιδιοτήτων μπορεί να

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες

χρησιμοποιηθεί για συλλογιστική (εξαγωγή συμπερασμάτων/νέας γνώσης) και για την δομημένη περιγραφή γνώσης ενός πεδίου ενδιαφέροντος (Spiliopoulos, et al., 2008)

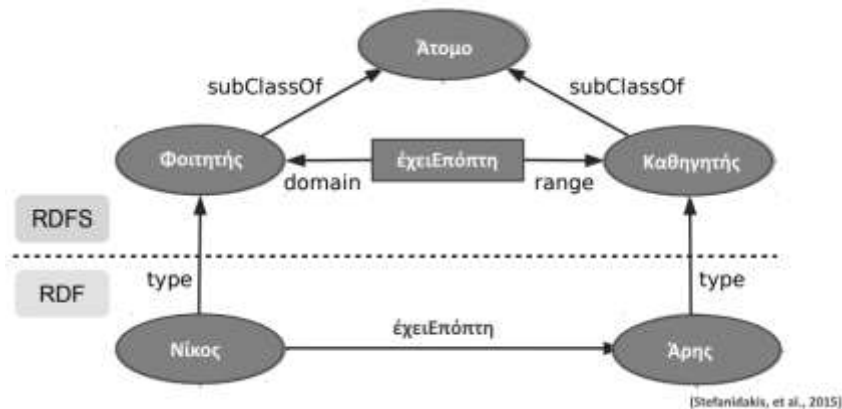
Στον ευρύτερο χώρο του Σημασιολογικού Ιστού οντολογίες (ontologies) ονομάζονται τα σύνολα στοιχείων μεταδεδωμένων. Μια οντολογία αποτελεί την αυστηρή περιγραφή μιας Θεματικής Περιοχής και, εκτός από σύνολα κλάσεων και ιδιοτήτων, μπορεί επίσης να περιέχει και ταξινομίες κλάσεων και ιδιοτήτων, καθώς και αξιώματα που πρέπει να ικανοποιούν οι εν λόγω κλάσεις και ιδιότητες. Οι οντολογίες μπορούν να εκφραστούν σε γλώσσες αναπαράστασης γνώσης, όπως η RDFS (Brickley & Guha, 2014) και η OWL (Hitzler, et al., 2012). Η χρήση όρων οντολογιών σε RDF γράφους καθιστά δυνατή την εφαρμογή συλλογισμού (reasoning) για την εξαγωγή νέων RDF προτάσεων.

Χαρακτηριστικά παραδείγματα οντολογιών που χρησιμοποιούνται από τις Βιβλιοθήκες είναι τα εξής: Dublin Core, SKOS (Simple Knowledge Organization System), BIBO (Bibliographic Ontology) και FOAF (Friend of a Friend). Η γνώση των υπαρχόντων διαθέσιμων λεξιλογίων είναι απαραίτητη για τη μετατροπή υπαρχόντων συνόλων δεδομένων σε RDF γράφους, συντελώντας στη δημιουργία Συνδεδεμένων Δεδομένων υψηλής ποιότητας, τα οποία είναι δυνατόν να ενσωματωθούν άμεσα στο παγκόσμιο Σύννεφο Ανοικτών Συνδεδεμένων Δεδομένων (Linked Open Data Cloud) (Spanos, 2015).

Άξια αναφοράς είναι τα παραδείγματα της Εθνικής Βιβλιοθήκης της Σουηδίας (Malmsten, 2008), της Γερμανικής Εθνικής Βιβλιοθήκης (<http://www.dnb.de/EN/lds>), της Εθνικής Βιβλιοθήκης της Ισπανίας (Vila-Suero & Gomez-Perez, 2013), καθώς και της Βρετανικής Βιβλιοθήκης (<http://bnb.data.bl.uk/>), οι οποίες δημοσιεύουν τους βιβλιογραφικούς καταλόγους τους και τα αρχεία καθιερωμένων όρων τους, θεματικούς όρους και πρόσωπα σε μορφή Συνδεδεμένων Δεδομένων (Spanos, 2015), έχοντας υιοθετήσει την τεχνολογία αυτή, σε γενικές γραμμές.

RDF Schema (RDFS)

Η πρώτη έκδοση του RDFS δημοσιεύθηκε το 1998, ενώ σήμερα χρησιμοποιείται η έκδοση RDFS 1.1 που ανακοινώθηκε το Φεβρουάριο του 2004. Το RDF Schema (RDFS) αποτελεί μια σημασιολογική επέκταση του προτύπου RDF. Χρησιμοποιείται για να περιγράψει σύνολα παρόμοιων πόρων καθώς και τη σχέση μεταξύ τους. Για να λειτουργήσει το μοντέλο RDF, απαιτείται μια καθορισμένη γλώσσα όπως ειδικά λεξιλόγια. Την ιδιότητα αυτή την προσφέρει το RDFS το οποίο καθορίζει βασικές σχέσεις από καθορισμένα λεξιλόγια όπως Class, SubClassOf, Domain, Range, Label και Comment (W3C, 2004). Τα λεξιλόγια μπορούν να χρησιμοποιηθούν για να οριστούν άλλα λεξιλόγια για συγκεκριμένα πεδία εφαρμογών. Η λειτουργία που προσφέρει το RDFS είναι η δημιουργία λεξιλογίων. Το RDFS προσδιορίζει τη σημασιολογία των πόρων, ενώ το RDF αποτελεί το μοντέλο για τη σύνταξη και τη περιγραφή των πόρων. Στην πράξη το πρότυπο RDFS χρησιμοποιείται ευρέως ως ένας μηχανισμός για τον ορισμό μιας βασικής δομής/σχήματος που διέπει μια ομάδα δεδομένων RDF. Πολλά από τα λεξιλόγια RDF βασίζονται στο Σχήμα RDF (Stefanidakis, et al., 2015). Στο παρακάτω σχήμα εκφράζεται η σχέση και η εφαρμογή του RDFS στο RDF, Σχήμα 7.



Σχήμα 7 Παράδειγμα εφαρμογής RDFS στο RDF (Stefanidakis, et al., 2015)

3.3.3 Πλεονεκτήματα του RDF

Το RDF, είναι πολύ σημαντικό για την εφαρμογή των Συνδεδεμένων Δεδομένων. Η χρήση του RDF προσφέρει κάποια βασικά πλεονεκτήματα για τα Συνδεδεμένα Δεδομένα, τα οποία είναι (Heath,Bizer, 2011):

- Οφελεί η χρήση των HTTP URIs, ως μοναδικά αναγνωριστικά αντικειμένων δεδομένων και όρων λεξιλογίων σε παγκόσμια κλίμακα, καθώς το RDF είναι σχεδιασμένο για χρήση σε παγκόσμια κλίμακα, όπου ο καθένας μπορεί να αναφέρει οτιδήποτε.
- Οι χρήστες μπορούν να αναζητήσουν οποιοδήποτε URI σε ένα γράφο RDF, ώστε να ανακτήσουν πρόσθετες πληροφορίες. Κάθε τριπλέτα RDF μπορεί να αποτελέσει το σημείο εκκίνησης για την εξερεύνηση αυτού του δικτύου.
- Είναι δυνατή η σύνδεση δεδομένων από διαφορετικές πηγές και σύνολα δεδομένων.
- Πληροφορίες από διαφορετικές πηγές δεδομένων μπορούν ευκολά να συνδυαστούν με την συγχώνευση των δύο συνόλων τριπλετών σε ένα κοινό γράφο.
- Το RDF επιτρέπει πληροφοριών με διαφορετικούς όρους λεξιλογίου για τον ίδιο γράφο.
- Σε συνδυασμό με γλώσσες περιγραφής όπως η RDFS και η OWL, παρέχεται η δυνατότητα επιλογής για το πόσο δομημένα θα είναι τα δεδομένα.

3.4 Συνδεδεμένα Δεδομένα και Βιβλιοθήκες

Παρόλο που ο τρόπος με το οποίο είναι διαμορφωμένα τα δεδομένα των περισσότερων βιβλιοθηκών σήμερα, ο οποίος δεν υποστηρίζει τις λειτουργίες των Ανοιχτών Συνδεδεμένων Δεδομένων, οι βιβλιοθήκες σταδιακά αντιλαμβάνονται τα οφέλη που προσφέρει η υιοθέτηση τους, όπως και τα εμπόδια που υπάρχουν κατά τη διαδικασία ενσωμάτωσής τους. (Hannemann & Kett 2010).

Πλέον έχει διαμορφωθεί η πεποίθηση, ότι τα Ανοιχτά Συνδεδεμένα Δεδομένα μπορούν να προσφέρουν πολλά πλεονεκτήματα για τις βιβλιοθήκες. Μπορούν να βοηθήσουν τις βιβλιοθήκες να βελτιώσουν την ορατότητα των δεδομένων τους και να διευκολύνουν την αναζήτηση και σύνδεσή τους μέσω Διαδικτύου και εκτός των καταλόγων στους οποίους φυλάσσονται. Το αποτέλεσμα της υιοθέτησης των Συνδεδεμένων Δεδομένων θα είναι όχι μόνο να βοηθήσει τους χρήστες να εντοπίσουν πληροφορίες διαδικτυακά, αλλά θα προσελκύσει και νέους χρήστες στο περιβάλλον της βιβλιοθήκης. Ακόμα, η χρήση υπηρεσιών των Ανοιχτών Συνδεδεμένων Δεδομένων θα βοηθούσε τις βιβλιοθήκες στην

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες

ανταλλαγή δεδομένων και πληροφοριών, την κοινή χρήση δεδομένων και τη συνεργασία όχι μόνο μεταξύ των βιβλιοθηκών αλλά και με διαφορετικούς οργανισμούς. (Illien, 2012)

3.4.1 Η κατάσταση σήμερα και η ανάγκη για μετάβαση

Στο ιστορικό πλαίσιο των ραγδαίων τεχνολογικών εξελίξεων, το MARC (Machine Readable Cataloging), ήταν και είναι ο κυρίαρχος στα πρότυπα κωδικοποίησης δεδομένων. Το MARC σχεδιάστηκε στις αρχές της δεκαετίας του '60 και πρωτοεμφανίστηκε το 1966 (Avram, 1975). Δημιουργήθηκε για να καλύψει την ανάγκη για δημιουργία αρχείων που θα μπορούσαν να διαβαστούν από ηλεκτρονικούς υπολογιστές και να μοιραστούν μεταξύ των βιβλιοθηκών. Αποτελεί ένα σύνολο ψηφιακών μορφών που περιγράφουν αντικείμενα σε έναν κατάλογο βιβλιοθήκης.

Ένα από τα μειονεκτήματα του MARC είναι ότι για την αξιοποίηση του είναι αναγκαία η χρήση ενός μεγάλου όγκου εγχειριδίων, ωστόσο το πρόβλημα αυτό δεν είναι το μεγαλύτερο πρόβλημα. Τα περισσότερα εμπόδια προκύπτουν κατά τη διαδικασία της κωδικοποίησης σύνθετων βιβλιογραφικών δεδομένων, όπως εικόνες και στη συμβατότητα με τα εκάστοτε λογισμικά. Τα προβλήματα του μετασχηματισμού δεδομένων για την επέκτασή τους στο να μπορούν να διαβάζονται από τα λογισμικά με τον βέλτιστο τρόπο αποτέλεσε και συνεχίζει να αποτελεί την βασική πρόκληση των προτύπων κωδικοποίησης δεδομένων (Tennant, 2002).

Το MARC δεν επινοήθηκε για την ομαλή αξιοποίηση των συστημάτων πληροφορικής για την ανάκτηση πληροφοριών με αυτοματοποιημένο τρόπο. Αρχικός του σκοπός ήταν να αυτοματοποιήσει τις διαδικασίες και τα καθήκοντα ενός τμήματος τεχνικών υπηρεσιών της δεκαετίας του 1950, όπως τη δημιουργία και την εκτύπωση καρτών καταλογογράφησης (Avram, 1975; Coyle, 2005; Coyle & Hillmann, 2007; Tennant, 2002). Στην σημερινή εποχή εμφανίζονται νέες απαιτήσεις για την εγκαθίδρυση σύγχρονων υποδομών μεταδεδομένων για τις βιβλιοθήκες, όπως η αξιοποίηση των Συνδεδεμένων Δεδομένων, συμπεριλαμβανομένων των ιδιοτήτων όπως η ευελιξία, η επεκτασιμότητα, η ευκρίνεια και η ανοικτότητα. Στην πορεία αυτή εντοπίζονται ορισμένες βασικές προκλήσεις που πρέπει να αντιμετωπιστούν για την εφαρμογή μιας τέτοιας αλλαγής (Tennant, 2004).

Παρόλο τον μεγαλεπήβολο τίτλο του άρθρου “MARC must die” του Tennant το 2004, ο ίδιος αναφέρει ότι αυτό που πρέπει να σταματήσει δεν είναι το MARC αλλά η αποκλειστική εξάρτηση από αυτά τα εργαλεία για την περιγραφή των βιβλιογραφικών δεδομένων. Με άλλα λόγια, πρέπει να αφομοιώσουμε το MARC σε ένα ευρύτερο και πλουσιότερο σύνολο εργαλείων, προτύπων και πρωτοκόλλων (Tennant, 2004).

Οι προκλήσεις επεκτείνονται επομένως στην ανάγκη ανάπτυξης αποδοτικών λύσεων που θα καταστήσουν τα δεδομένα που βρίσκονται στον Παγκόσμιο Ιστό ευδιάκριτα και την δυνατότητα απόδοσης τους με τρόπο που να καλύπτεται η ιδιότητα της αναγνωσιμότητας από μηχανές (Wahid, 2017). Σήμερα γίνεται διαρκώς λόγος για την αξιοποίηση των τεχνολογιών του Σημασιολογικού Ιστού και εκτενέστερα των Συνδεδεμένων Δεδομένων, καθώς και τη δημοσίευση των βιβλιογραφικών πόρων στο Διαδίκτυο. Η πλουσιότερη περιγραφή βιβλιογραφικών δεδομένων η επαναχρησιμοποίηση και η αξιοποίηση των μεταδεδομένων και η συνολικότερη αξιοποίηση της καταγεγραμμένης πληροφορίας, αποτελούν τις σημερινές προκλήσεις της εποχής στον κόσμο των μεγάλων δεδομένων και του διαδικτύου.

Ωστόσο σήμερα η τεχνολογία των Συνδεδεμένων Δεδομένων έχει υιοθετηθεί ως ένα βαθμό από βιβλιοθήκες σε διεθνές επίπεδο. Προς το παρόν υπάρχουν αρκετά παραδείγματα ευρωπαϊκών βιβλιοθηκών που έχουν υιοθετήσει τη φιλοσοφία των συνδεδεμένων δεδομένων, κάτι το οποίο μπορεί να αποτελέσει παράδειγμα και για τις Ελληνικές βιβλιοθήκες. Οι παρακάτω βιβλιοθήκες δημοσιεύουν τους βιβλιογραφικούς τους καταλόγους και τους θεματικούς όρους σε μορφή Συνδεδεμένων Δεδομένων. Παραδείγματα ευρωπαϊκών βιβλιοθηκών είναι

- Η Εθνική βιβλιοθήκη της Σουηδίας
- Η Γερμανική βιβλιοθήκη (<http://www.dnb.de/EN/lds>)
- Η Εθνική βιβλιοθήκη της Ισπανίας
- Η Βρετανική βιβλιοθήκη (<http://bnb.data.bl.uk/>)

Μια ακόμη εφαρμογή που έχει πραγματοποιηθεί ήδη και αφορά την σύνδεση βιβλιοθηκών με τα Συνδεδεμένα Δεδομένα είναι τα επιστημονικά συνέδρια MTSR (Metadata and Semantics Research) και SWIB (Semantic Web In Biblioteken). Το SWIB είναι αγγλόφωνο επιστημονικό συνέδριο και έχει ως στόχο την έρευνα σχετικά με την εφαρμογή των τεχνικών του Σημασιολογικού Ιστού στις βιβλιοθήκες. Αυτό το συνέδριο διεξάγεται μέχρι σήμερα σε διάφορες πόλεις της Γερμανίας, ενώ το 2018 έγινε στη Βόννη. Άλλα παραδείγματα εφαρμογής Συνδεδεμένων Δεδομένων σε βιβλιοθήκες αποτελούν το επιστημονικό περιοδικό SWJ (Semantic Web Journal). Το JWS καλύπτει όλες τις πτυχές του Σημασιολογικού Ιστού. Επίσης, ενθαρρύνει τη δημοσίευση πειραμάτων μεγάλης κλίμακας και την ανάλυσή τους, έτσι ώστε να αναδειχθούν σενάρια και μέθοδοι που εισάγουν τη σημασιολογία σε συστήματα που είναι ήδη διαθέσιμα στον παγκόσμιο ιστό. Ενδεικτικά το JWS καλύπτει επιστημονικές περιοχές όπως ηλεκτρονικό επιχειρείν (e-business), ηλεκτρονική κοινωνία (e-community), διαχείριση γνώσης, ηλεκτρονική μάθηση, ψηφιακές βιβλιοθήκες και ηλεκτρονικές επιστήμες (Stefanisakis, et al., 2015).

3.4.2 Οφέλη των Συνδεδεμένων Δεδομένων

Η τεχνολογία των Συνδεδεμένων Δεδομένων εμφανίζει ποικίλα πλεονεκτήματα, τόσο για τον εκδότη και διαχειριστή δεδομένων που αποφασίζει να τα δημοσιεύσει σε μορφή Συνδεδεμένων Δεδομένων όσο και για τον τελικό χρήστη που τα καταναλώνει. (Spanos, 2015). Ένα από αυτά τα πλεονεκτήματα είναι, ο οικονομικός και εύκολος τρόπος σύνδεσης δεδομένων με άλλα εξωτερικά σύνολα δομημένων ή αδόμητων δεδομένων. Όσον αφορά το χρήστη-καταναλωτή Συνδεδεμένων Δεδομένων, με τη χρήση URI για τον μονοσήμαντο προσδιορισμό οντοτήτων και την ύπαρξη συνδέσμων μεταξύ σχετικών δεδομένων, μπορεί να αποφευχθεί σε μεγάλο βαθμό τις χρονοβόρες διαδικασίες εύρεσης αντιστοιχιών μεταξύ διαφορετικών συνόλων δεδομένων και ανακάλυψης νέων συνόλων δεδομένων.

Τα Συνδεδεμένα Δεδομένα κρίνονται επωφελή και για τις Βιβλιοθήκες, γεγονός που έχει διεθνώς αναγνωριστεί τα τελευταία χρόνια (Alemu, et al., 2012; Byrne & Goddard, 2010; Dunsire & Willer, 2011). Τα οφέλη που προσφέρουν τα Συνδεδεμένα Δεδομένα αφορούν όχι μόνο στις Βιβλιοθήκες ως οργανισμούς, αλλά και σε μεμονωμένα άτομα, όπως σε βιβλιοθηκονόμους, προγραμματιστές, ερευνητές και φοιτητές και γενικώς όλης της κοινότητας των Βιβλιοθηκών (Baker, 2012).

Οφέλη για τους Χρήστες

Συγκεκριμένα, η αποτύπωση των μεταδεδομένων των Βιβλιοθηκών σε μορφή Συνδεδεμένων Δεδομένων καθιστά ευκολότερη την ανακάλυψη των τεκμηρίων της από τους χρήστες και υπόσχονται την παροχή περισσότερων υπηρεσιών ευκολότερης αναζήτησης και ανάκτησης πληροφοριών. Τα δομημένα δεδομένα μπορούν να προστεθούν σε υπάρχοντα έγγραφα Ιστού μέσω γλωσσών σημασιολογικής επισημείωσης όπως RDFa, ώστε να εντοπιστούν μέσω αλγορίθμων υπολογισμού συνάφειας των μηχανών αναζήτησης. Αποτέλεσμα αυτού είναι η βελτίωση στη θέση που εμφανίζεται ένα έγγραφο στα αποτελέσματα αναζήτησης.

Οφέλη για την Βιβλιοθήκη

Η διασύνδεση δεδομένων από διαφορετικές πηγές προσφέρει στις Βιβλιοθήκες τη δυνατότητα να βελτιώσουν σημαντικά την ποιότητα των μεταδεδομένων τους. Ένας δεδομένος πόρος μπορεί πλέον να περιγραφεί σε συνεργασία με άλλες Βιβλιοθήκες και να συνδεθεί με δεδομένα που προέρχονται από άλλες. Με αυτόν τον τρόπο, οι βιβλιοθήκες προσφέρουν πιο περιγραφικά και λεπτομερή δεδομένα στους χρήστες άρα πιθανόν καλύτερης ποιότητας.

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες

Εάν η βιβλιοθήκη χρησιμοποιεί συνδεδεμένα δεδομένα για να παρουσιάσει τα υλικά της στον ιστό, μια αναζήτηση για ένα βιβλίο θα μπορούσε να φέρει ένα παρόμοιο γράφημα, με πληροφορίες σχετικά με το βιβλίο και τη διαθεσιμότητά του στη βιβλιοθήκη κατά τη στιγμή της αναζήτησης. Οι συνδέσεις με τον συγγραφέα και τον εκδότη και άλλες πληροφορίες σχετικά με το βιβλίο θα είναι επίσης διαθέσιμες μέσω της διασύνδεσης γραφημάτων. Οι βιβλιοθήκες ήδη συλλέγουν και διαχειρίζονται αυτές τις πληροφορίες, καθιστώντας τους χρήσιμες για τους ερευνητές στο διαδίκτυο, είναι ένα προφανές επόμενο βήμα.

Οι βιβλιοθήκες που πραγματοποιούν συνδέσεις με σύνολα δεδομένων που είναι ανοιχτά και ελεύθερα χρησιμοποιήσιμα σαν αυτό θα είναι ένα όφελος για τις βιβλιοθήκες καθώς κάνουν τα δεδομένα τους να συνδέονται μέσω URI στο διαδίκτυο. Ένα παράδειγμα αυτού θα μπορούσε να είναι ένα βιβλίο (το υποκείμενο στη τριπλέτα) της συλλογής που έχει σαν κατηγορία τον εκδότη του (το αντικείμενο). Χρησιμοποιώντας το URI του φορέα που φέρει το υποκείμενο στη συλλογή του θα μπορούσε να φέρει πλούσια πηγή δεδομένων για το υλικό της βιβλιοθήκης.

3.4.3 Εμπόδια στην αξιοποίηση των Συνδεδεμένων Δεδομένων από τις βιβλιοθήκες

Τα συνδεδεμένα δεδομένα δεν χρησιμοποιούνται ακόμη ευρέως σε κανένα πεδίο, συμπεριλαμβανομένων των βιβλιοθηκών. Υπάρχουν ορισμένα εμπόδια στις βιβλιοθήκες που αποφασίζουν να μην τα χρησιμοποιήσουν, συμπεριλαμβανομένης της σχετικής πολυπλοκότητας της τεχνολογίας. Εμπόδια για τη χρήση των Συνδεδεμένων Δεδομένων από πολλές βιβλιοθήκες έχουν οικονομικούς, πολιτικούς και συστημικούς περιορισμούς. (Byrne & Goddard, 2010).

Μια βασική πρόκληση για την ανάπτυξη Συνδεδεμένων Δεδομένων στις βιβλιοθήκες είναι η έλλειψη ενημέρωσης. Παρόλο που υπάρχουν θεμέλια ανάπτυξης, όπως οι ψηφιακές βιβλιοθήκες, που κάνουν πολύπλοκες εργασίες στον τομέα, οι έννοιες των συνδεδεμένων δεδομένων και του σημασιολογικού ιστού θα πρέπει να διαμορφωθούν ακόμα αρκετά στη συνείδηση πολλών βιβλιοθηκονόμων.

Τέλος, υπάρχουν εμπόδια ειδικής σημασίας που αντιμετωπίζουν γενικά οι βιβλιοθήκες για τη χρήση Συνδεδεμένων Δεδομένων. Το απόρρητο είναι μια τεράστια ανησυχία όσο αφορά τα Συνδεδεμένα Δεδομένα. Όσο περισσότερα δεδομένα συνδέονται, τόσο περισσότερο είναι πιθανό να εκτίθενται οι προσωπικές πληροφορίες. Οι βιβλιοθηκονόμοι, με τη μακρόχρονη παράδοσή τους για την προστασία της ιδιωτικής ζωής των πολιτών, θα πρέπει να αναλάβουν ενεργό ρόλο στη διασύνδεση των δεδομένων για να εξασφαλίσουν την προστασία των δικαιωμάτων (Byrne & Goddard, 2010).

Κεφάλαιο 4. Case Study, Bibliographic Framework Initiative (BIBFRAME)

Καθώς ο αριθμός των δεδομένων αυξάνεται παγκοσμίως διαρκώς, στην σημερινή εποχή προκύπτουν νέες ανάγκες για την ανάπτυξη μοντέλων δεδομένων με δυνατότητες σύνδεσης βιβλιογραφικών δεδομένων. Το Bibliographic Framework Initiative (BIBFRAME) αποτελεί ένα μοντέλο δεδομένων το οποίο αναπτύχθηκε με πρωτοβουλία της Βιβλιοθήκης του Κογκρέσου, σε συνεργασία με την Zepheira. Αποτελεί μια ήδη εφαρμογή της τεχνολογίας των Συνδεδεμένων Δεδομένων στις βιβλιοθήκες και προάγει την υιοθέτηση και την αξιοποίηση των Συνδεδεμένων Δεδομένων από όλο και περισσότερες βιβλιοθήκες Παγκοσμίως. Σκοπός του BIBFRAME είναι η πλήρης αντικατάσταση των προτύπων του MARC (Machine Readable Cataloging) μέσω της αξιοποίησης των συνδεδεμένων δεδομένων (Linked Data). Με αυτόν τον τρόπο θα γίνουν τα βιβλιογραφικά δεδομένα πιο χρήσιμα εντός και εκτός της κοινότητας των βιβλιοθηκών.

4.1 Ιστορικό Πλαίσιο του BIBFRAME

Από το 1989, όλο και περισσότερες πληροφορίες έχουν μεταφερθεί στον Παγκόσμιο Ιστό. Από τα κοινωνικά δίκτυα, διάφορες υπηρεσίες, φορολογικά έντυπα κυβερνητικών υπηρεσιών, ιατρικά δεδομένα, εμπορικές υπηρεσίες αλλά και βιβλιοθήκες που πλέον δημοσιεύουν μέρος του υλικού τους βρίσκονται πλέον στον Παγκόσμιο Ιστό. Ωστόσο οι μεθοδολογίες καλύτερης καταγραφής της πληροφορίας των βιβλιοθηκών ξεκινούν τη δεκαετία του 1960 (Schreur, 2018).

4.2 Το μοντέλο δεδομένων BIBFRAME

Το 2011 η Βιβλιοθήκη του Κογκρέσου ανέθεσε την ανάπτυξη της επόμενης γενιάς μηχανικά αναγνώσιμων καταλόγων. Το αποτέλεσμα αυτής της προσπάθειας ήταν το μοντέλο δεδομένων *BIBFRAME (Bibliographic Framework)*. Το BIBFRAME χρησιμοποιεί Συνδεδεμένα Δεδομένα και τεχνολογίες Σημασιολογικού Ιστού με σκοπό να αποτελέσει τη βάση για το μέλλον της βιβλιογραφικής περιγραφής (Pesch & Miller, 2016). Είναι το θεμέλιο για το μέλλον της βιβλιογραφικής περιγραφής τόσο στο διαδίκτυο όσο και στον ευρύτερο δικτυωμένο κόσμο

Σκοπός της βιβλιοθήκης του Κογκρέσου ήταν να αναπτυχθεί μια επόμενη γενιά από το MARC, για την περιγραφή των βιβλιογραφικών δεδομένων, που μπορεί να λειτουργεί στο διαδίκτυο και είναι αρκετά ευέλικτη για ανάπτυξη και εφαρμογή. Το λεξιλόγιο που

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες

χρησιμοποιεί το BIBFRAME είναι κοινό με του RDF, δηλαδή τις RDF ιδιότητες, κλάσεις και τις μεταξύ τους σχέσεις. Το BIBFRAME ενσωματώνει τις αρχές, τις πρακτικές και τις τεχνολογίες των Συνδεδεμένων Δεδομένων, ώστε να επιτρέπεται η δημιουργία σύνδεσης πόρων προερχόμενων από διαφορετικές πηγές που συνδέονται και με άλλους πόρους (Pesch & Miller, 2016).

Το BIBFRAME είναι σχεδιασμένο για να ενσωματώνεται και να συμμετέχει στην ευρύτερη πληροφοριακή κοινότητα καθώς και να εξυπηρετεί τις συγκεκριμένες ανάγκες των βιβλιοθηκών. Το BIBFRAME συμβάλει με τους παρακάτω νέους τρόπους στην περιγραφή βιβλιογραφικού περιεχομένου:

- Κάνει σαφής τη διαφοροποίηση μεταξύ του εννοιολογικού περιεχομένου του βιβλιογραφικού υλικού και της φυσικής/ψηφιακής εκδήλωσής του.
- Προσδιορίζει τις οντότητες της πληροφορίας όπως για παράδειγμα τις εκδόσεις.
- Αξιοποιεί τις σχέσεις ανάμεσα σε δύο ή περισσότερες οντότητες.

Το μοντέλο BIBFRAME δημιουργήθηκε με σκοπό τον συντονισμό της καταλογογράφησης και των μεταδεδομένων που δημιουργούν οι βιβλιοθήκες και αποτελεί το σημείο για τη πρόσβαση της κοινότητας των βιβλιοθηκών σε έναν πολύ πιο μεγάλο Ιστό Δεδομένων (L.C., 2012).

Οι βιβλιοθηκονόμοι είναι ειδικοί στον εντοπισμό και τη σύλληψη βιβλιογραφικών σχέσεων και το μοντέλο BIBFRAME είναι αυτό που θέτει τη βάση ώστε να μπορεί αυτή η εξειδικευμένη γνώση να παρουσιαστεί. Οι βιβλιογραφικές σχέσεις θα συμβάλουν σημαντικά στον πλούτο του διαδικτύου και θα βοηθήσουν άτομα που αναζητούν πληροφορίες, μέσω της αναζήτησης σχετικών πόρων από τους χρήστες.

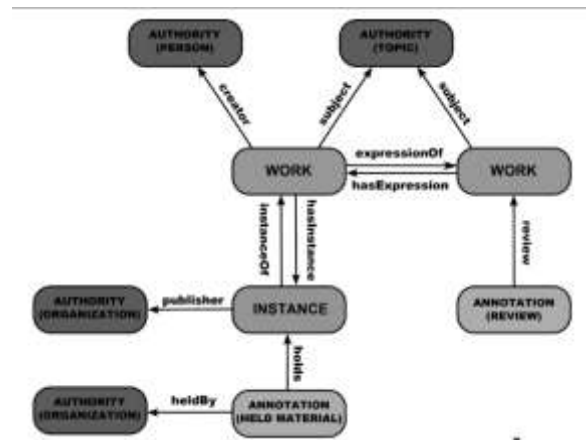
4.3 Δομή του μοντέλου BIBFRAME 1.0

Το μοντέλο BIBFRAME αποτελείται από τις ακόλουθες κύριες κατηγορίες: Σχήμα 8 (L.C., 2012)

- **Creative Work** - ένας πόρος που αντικατοπτρίζει μια εννοιολογική ουσία του στοιχείου καταλογογράφησης.
- **Instance** - ένας πόρος που αντικατοπτρίζει μια ατομική, υλική ενσωμάτωση του έργου.
- **Authority** - ένας πόρος που αντικατοπτρίζει έννοιες βασικών αρχών που έχουν ορίσει σχέσεις

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες που αντικατοπτρίζεται στο έργο και την υπόθεση. Τα παραδείγματα των πόρων της Αρχής περιλαμβάνουν άτομα, τοποθεσίες, θέματα, οργανισμοί κ.λπ.

• **Annotation** - ένας πόρος που διακοσμεί άλλους πόρους BIBFRAME με επιπλέον πληροφορίες. Παραδείγματα τέτοιων σχολιασμών περιλαμβάνουν τις πληροφορίες για την κατοχή βιβλιοθηκών, κάλυψη της τέχνης και κριτικές.

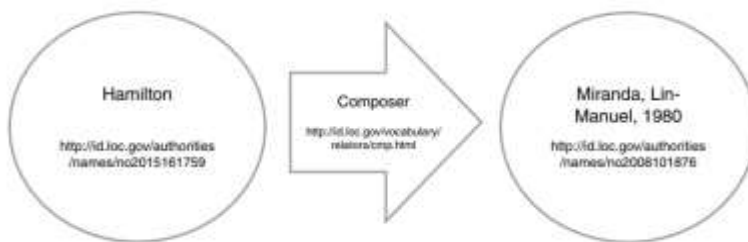


Σχήμα 8 Το μοντέλο BIBFRAME
alamw15.ala.org/sessions/handouts

Η πρωτότυπη μορφή του BIBFRAME δημοσιεύθηκε το 2012 και αποτελεί την έκδοση 1.0 ή οποία έχει αντικατασταθεί με την έκδοση BIBFRAME 2.0 από τον Απρίλιο του 2016. Ενώ το BIBFRAME 1.0 αποτελείται από τέσσερις βασικές κλάσεις: Creative Work, Instance, Authority, Annotation, το BIBFRAME 2.0 ο αριθμός των βασικών κλάσεων μειώθηκε στα τρία: Work, Instance και Item (Library of Congress, 2016).

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες

Το μεγαλύτερο επίπεδο στο BIBFRAME 2.0 είναι η κλάση Work. Ένα στοιχείο Work αντιστοιχεί στην εννοιολογική ουσία του καταλογοποιημένου πόρου όπως, συγγραφείς, γλώσσες, θέμα (Schreur, 2018). Ενώ το Work αποτελεί μια εννοιολογική κλάση, το Instance αποτελεί μια πιο φυσική. Ένα στοιχείο Instance μπορεί να αντιστοιχεί σε πληροφορίες όπως, ο εκδότης, η τοποθεσία, η ημερομηνία δημοσίευσης, η μορφή (Library of Congress, 2016). Το τελευταίο επίπεδο του BIBFRAME 2.0 είναι το Item. Μπορεί να είναι το φυσικό ή το ψηφιακό αντίγραφο ενός μεμονωμένου αντικειμένου. Σε ένα στοιχείο Item αντιστοιχούν πληροφορίες σχετικές με τη τοποθεσία (φυσική ή εικονική), κωδικός BARCODE (Library of Congress, 2016). Στο σχήμα 10 παρουσιάζεται ένα παράδειγμα μιας εγγραφής ενός σε BIBFRAME καθώς και η αναπαράσταση της σχετικής τριπλέτας RDF Σχήμα 9.



Σχήμα 9 Παράδειγμα τριπλέτας RDF (Steele, 2016)



Σχήμα 10 Παράδειγμα BIBFRAME εγγραφής (Steele, 2016)

Στο συγκεκριμένο παράδειγμα, Σχήμα 7, περιγράφεται το CD της ηχογράφησης του soundtrack του Hamilton Musical. Κάθε αντίγραφο CD που κατέχει η βιβλιοθήκη αποτελεί το Item. Στο σχήμα να αντίγραφο Item προστίθενται ακόμα τρεις βασικές έννοιες, το Subject, Agent και Event. Οι άνθρωποι, οι οργανισμοί και οι δικαιοδοσίες που σχετίζονται

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες με το Work και το Instance ορίζονται ως Agents. Οι Agents πιθανόν να είναι οι συγγραφείς, συντάκτες, εκδότες, εικονογράφοι, συνθέτες κ.α. Το Subject μπορεί να είναι μόνο ένα και επικεντρώνεται σε αυτό που αφορά το Work. Το BIBFRAME 2.0 χρησιμοποιεί και τις ιδιότητες και τις κλάσεις του RDF. *"Μια κλάση αναγνωρίζει έναν τύπο πόρων BIBFRAME. Οι ιδιότητες χρησιμεύουν ως μέσο για την περαιτέρω περιγραφή του πόρου BIBFRAME"* (Library of Congress, n.d.). "Μια πλήρης λίστα του λεξιλογίου BIBFRAME 2.0 υπάρχει διαθέσιμη στη διεύθυνση <http://id.loc.gov/ontologies/bibframe.html>.

Εργαλεία που διατίθενται από τη Βιβλιοθήκη του Κογκρέσου για μετατροπή των εγγραφών από MARC σε BIBFRAME και για δημιουργία εγγραφών BIBFRAME (Library of Congress).

- MARC-to-BIBFRAME Comparison
- MARC-to-BIBFRAME Transformation
- BIBFRAME Editor

4.4 BIBFRAME και Συνδεδεμένα Δεδομένα

Το κλειδί για την πλοήγηση στο Σημασιολογικό Ιστό είναι τα Συνδεδεμένα Δεδομένα. Οι κανόνες για την εφαρμογή των Συνδεδεμένων Δεδομένων θεσπίστηκαν από τον Tim Berners-Lee. Το μοντέλο BIBFRAME έχει αναπτυχθεί βασισμένο στον ορισμό αυτό των τεσσάρων κανόνων του Berners-Lee. Οι κανόνες για τα Συνδεδεμένα Δεδομένα απαιτούν ένα πρότυπο. Το BIBFRAME χρησιμοποιεί το σύστημα τριπλετών του RDF με σκοπό να δημοσιευτούν οι βιβλιογραφικές εγγραφές των καταλόγων στο Σημασιολογικό Ιστό (Steele, 2018).

Σήμερα έχει εντοπιστεί η ανάγκη των βιβλιοθηκών να προσαρμοστούν στις προσδοκίες των χρηστών να προσφέρουν μεγαλύτερη διασύνδεση έργων και πληροφοριών, όπως τα δημοσιευμένα έργα μιας ακαδημαϊκής κοινότητας, με ιστοσελίδες και άλλες σχετικές πληροφορίες. Επομένως ο κατάλογος των βιβλιοθηκών δεν είναι απαραίτητο να περιέχει

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες

ένα τεράστιο όγκο δεδομένων αλλά να μπορεί να διασυνδέσει πολλά δεδομένα και πληροφορίες μεταξύ τους τις οποίες οι χρήστες μπορούν να εντοπίζουν μέσω μηχανών αναζήτησης στις βιβλιοθήκες (Calhoun, 2011). Επομένως συνδέοντας αυτές τις πληροφορίες στο διαδίκτυο, μπορεί να αυξηθεί η προβολή της βιβλιοθήκης, καθώς επίσης θα μπορεί να προσφέρει υλικό ήδη υπάρχον αλλά και νέο.

Σε μια μελέτη περίπτωσης με τη χρήση του BIBFRAME που έγινε από το Χάρβαρντ, ο Karim Tharani κατέληξε στο BIBFRAME έχει τη δυνατότητα να αντιμετωπιστούν πολλά από τα ζητήματα που απομακρύνουν την ενσωμάτωση μια βιβλιοθήκης στον Ιστό, συμπεριλαμβανομένων των ζητημάτων της ακριβούς αναζήτησης, των authority controls, της ταξινόμησης και της αποσαφήνισης δεδομένων (Tharani, 2015).

Κεφάλαιο 5 Αποτίμηση του Έργου και Μελλοντικές Κατευθύνσεις

Το παρόν κεφάλαιο διακρίνεται σε δύο ενότητες. Στη πρώτη ενότητα καταγράφονται τα συμπεράσματα που διεξάχθηκαν από την ενασχόληση με το συγκεκριμένο θέμα. Στη δεύτερη ενότητα καταγράφονται μερικές προτάσεις που αποτελούν μελλοντικές κατευθύνσεις σχετικά με την αξιοποίηση των Συνδεδεμένων Δεδομένων από τις Βιβλιοθήκες.

5.1 Αποτίμηση του Έργου

Η συγκεκριμένη ενότητα αφιερώνεται στην καταγραφή των συμπερασμάτων που προκύπτουν από την ενασχόληση και τη μελέτη με το συγκεκριμένο θέμα, την αξιοποίηση των Συνδεδεμένων Δεδομένων από τις βιβλιοθήκες. Τα συμπεράσματα που προκύπτουν, αναφορικά αφορούν τις παρακάτω θεματικές ενότητες:

- Η ανάγκη που προκύπτει για μετάβαση σε ένα νέο τρόπο διαχείρισης των πόρων των βιβλιοθηκών.
- Οι εφαρμογές και οι τρόποι με τους οποίους μπορεί να επιτευχθεί η μετάβαση.
- Τα εμπόδια που δημιουργούνται
- Πιθανές λύσεις

Οι βιβλιοθήκες σήμερα είναι σε μεγάλο βαθμό απομονωμένες όσον αφορά την ανταλλαγή δεδομένων, δεδομένου ότι τα δεδομένα είναι κατά κύριο λόγο που συλλέγονται από βιβλιοθήκες για βιβλιοθήκες. Η διαδικασία ανταλλαγής και από κοινού χρήσης δεδομένων με ιδρύματα διαφορετικά των βιβλιοθηκών, βρίσκεται ακόμη σε πρώιμο στάδιο, καθώς πολλές βιβλιοθήκες που κατέχουν συλλογές χρησιμοποιούν απαρχαιωμένα συστήματα όπως το MARC. Οι συνεργασίες που υπάρχουν είναι κυρίως μεταξύ μόνο οι βιβλιοθηκών και τα δεδομένα της βιβλιοθήκης δεν αποτελούν ακόμα αναπόσπαστο μέρος του ιστού. Αυτό οφείλεται κυρίως στο το κακό επίπεδο σύνδεσης μεταξύ συνόλων δεδομένων βιβλιοθήκης και δεδομένων από άλλους τομείς, αλλά και λόγω των τρεχουσών διαδικασιών συλλογής δεδομένων και μορφών δεδομένων, οι οποίες επικεντρώνονται σε συντηρητικές μεθόδους διαχείρισης δεδομένων.

Σήμερα το MARC, το οποίο αποτελεί το κύριο σύστημα διαχείρισης δεδομένων βιβλιοθηκών παγκοσμίως, θεωρείται υπεύθυνο για το “κλείδωμα” των ψηφιακών γνώσεων των βιβλιοθηκών, εντός των τειχών των βιβλιοθηκών, καθώς εμποδίζει την δυνατότητα της κοινής χρήσης και της σύνδεσης των πόρων. Όσον αφορά τα δεδομένα, η

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες

ορατότητα σχετίζεται άμεσα με τη διασύνδεση. Χωρίς σύνδεση με τον Ιστό και σφραγισμένα στις βάσεις των βιβλιοθηκών, τα αρχεία των βιβλιοθηκών είναι δύσκολο να βρεθούν και να ενσωματωθούν σε ένα μεγαλύτερο σύνολο δεδομένων. Για να αποσφραγιστούν οι εγγραφές της βιβλιοθήκης από την απαρχαιωμένη μορφή τους, το MARC και να καταφέρουν να συνδέσουν οι βιβλιοθήκες τους χρήστες τους με το περιεχόμενο που αυτοί χρειάζονται, πρέπει να γίνει με ένα καινοτόμο τρόπο.

Για να επιτευχθεί σήμερα η είσοδος των βιβλιοθηκών στο μεγαλύτερο Ιστό των Δεδομένων, θα πρέπει η έμφαση των βιβλιοθηκών να στραφεί από την λεπτομερή καταγραφή των πόρων των βιβλιοθηκών, στον προσδιορισμό και στην καθιέρωση των σχέσεων μεταξύ δύο ή περισσότερων πόρων. Για αυτό το λόγο είναι αναγκαίο να εφαρμοστούν σύγχρονες τεχνολογίες όπως τα Συνδεδεμένα Δεδομένα. Ήδη παρατηρούνται σημαντικά βήματα προόδου, με την υιοθέτηση του μοντέλου BIBFRAME, που αποτελεί μια εφαρμογή των Συνδεδεμένων Δεδομένων, από έναν αριθμό συλλογών και βιβλιοθηκών στον κόσμο.

Ένας γενικός ορισμός του BIBFRAME είναι πως αποτελεί μια εφαρμογή των Συνδεδεμένων Δεδομένων, με σκοπό την πλήρη αντικατάσταση του MARC. Το BIBFRAME χαρακτηρίζεται ως «θεμέλιο για το μέλλον της βιβλιογραφικής περιγραφής, τόσο στον Ιστό όσο και στον ευρύτερο παγκόσμιο κόσμο» Είναι αυτό που μπορεί να επιτρέψει τη διασύνδεση δεδομένων βιβλιοθηκών και τελικά πλούσιες εμπειρίες γνώσης.

Καθώς τα Συνδεδεμένα Δεδομένα αποτελούν μια νέα τεχνολογία, βρίσκονται σε συνεχή ανάπτυξη. Επομένως ακόμα υπολείπονται αρκετά ώστε να χαρακτηρισθούν τα Συνδεδεμένα Δεδομένα πλήρως λειτουργικά αλλά και έτοιμα προς χρήση και υιοθέτηση από πολύ μεγάλο αριθμό χρηστών. Τα Συνδεδεμένα Δεδομένα είναι ακόμα πολύπλοκα ως προς τη διαχείριση τους με αποτέλεσμα να απευθύνονται σε χρήστες με εξειδικευμένη γνώση. Η μεγάλη πρόκληση επομένως για τα Συνδεδεμένα Δεδομένα είναι να γίνουν ευρέως αποδεχτά από τους χρήστες του Ιστού.

Μια πιθανή λύση ώστε να επιτευχθεί ο διαμοιρασμός των πληροφοριών και η ανάπτυξη των Συνδεδεμένων Δεδομένων είναι, οι υπεύθυνοι για την έκδοση δεδομένων να εμπιστευτούν τα Συνδεδεμένα Δεδομένα για τη δημοσίευση των Δεδομένων τους.

Η σύνδεση των δεδομένων βιβλιοθήκης είναι αυτό που μπορεί να βοηθήσει τις βιβλιοθήκες να διαχειριστούν το περιεχόμενό τους με οικονομικά αποδοτικό τρόπο και να το κάνουν ορατό για το κοινό τους. Χρησιμοποιώντας τις τεχνολογίες του Linked Data, οι

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες

βιβλιοθήκες μπορούν να παρουσιάσουν αυτό που πρέπει να προσφέρουν με έναν ελκυστικό, εμπλουτισμό και πάνω από όλα εύκολο και καινοτόμο τρόπο

5.2 Μελλοντικές Κατευθύνσεις

Θέματα που θα μπορούσαν να ερευνηθούν ως μελλοντικές κατευθύνσεις σχετικά με τα Συνδεδεμένα Δεδομένα καθώς και με την αξιοποίηση τους από τις βιβλιοθήκες είναι πολλά. Ένα από αυτά θα μπορούσε να είναι η εκτενής μελέτη για την ανάπτυξη των Συνδεδεμένων Δεδομένων ώστε να καθοριστεί ένα πλήρες και εύχρηστο για τους χρήστες του μοντέλο, ώστε να υπάρξει μεγαλύτερη αποδοχή παγκοσμίως.

Ένα άλλο θέμα για μελλοντική μελέτη θα μπορούσε να αφορά την ερευνα σχετικά με το τι βαθμό είναι έτοιμες οι βιβλιοθήκες να υιοθετήσουν τα Συνδεδεμένα Δεδομένα και να εξάγουν τα δεδομένα τους στον Ιστό, με τη χρήση ερωτηματολογίων και άλλων εργαλείων, είτε σε εθνικό είτε σε παγκόσμιο επίπεδο. Ενώ δεν θα πρέπει να υπάρχει περιορισμός μόνο στο χώρο των βιβλιοθηκών αλλά θα μπορούσε να μελετηθεί και η περίπτωση υιοθέτησης Συνδεδεμένων Δεδομένων και από διαφορετικούς φορείς διαχείρισης πόρων όπως, από Μουσειακές και Αρχειακές συλλογές αντίστοιχα.

Τέλος γίνεται κατανοητό πως υπάρχει άμεση ανάγκη να κινηθούν οι βιβλιοθήκες προς αυτή τη κατεύθυνση καθώς πλέον η διαδικασία της καταλογογράφησης αναπτύσσεται και αλλάζει. Επομένως δεν θα πρέπει οι βιβλιοθήκες να μένουν απαγκιστρωμένες σε απαρχαιωμένα συστήματα και τεχνολογίες. Θα πρέπει να δούμε τους καταλόγους των βιβλιοθηκών ως ένα δίκτυο μεταξύ αυτόνομων οντοτήτων. Με τη βοήθεια των συνδεδεμένων δεδομένων προσφέρεται αυτή η δυνατότητα να ορίζεται σχεδόν οτιδήποτε ως οντότητα-πόρος και στη συνέχεια να μπορεί να υπάρξει σύνδεση, με άλλες οντότητες-πόρους. Οι συνδέσεις αυτές δεν αποτελούν μέρος μιας μεμονωμένης βάσης δεδομένων αλλά έχουν καθολική ισχύ. Έτσι, αντί για απομονωμένες βάσεις δεδομένων, μπορεί να δημιουργηθεί ένα παγκόσμιο δίκτυο όπου υπάρχει διαλειτουργικότητα μεταξύ όλων των στοιχείων και κάθε στοιχείο αποτελεί το θεμέλιο για περαιτέρω ανακάλυψη. Έτσι θα μπορούσε μελλοντικά να γίνει λόγος για μια ενιαία βιβλιοθήκη.

Βιβλιογραφία

Alemu, Getaneh, Stevens, Brett, Ross, Penny, Chandler, Jane. (2012). Linked Data for libraries: Benefits of a conceptual shift from library-specific record structures to RDF-based data models.

Allemang, Dean & Hendler, James. (2008). Semantic Web for the Working Ontologist: Effective Modeling in RDFS and OWL.

Anderson, T., & Whitelock, D. (2004). The educational semantic web: Visioning and practicing the future of education. Journal of interactive Media in Education.

Avram, D. Henriette. (1975). Marc: its history and implications Library of Congress.

Bake, Thomas. (2012). Libraries, languages of description, and linked data: a Dublin Core perspective.

Becker, C., & Bizer, C. (2008). DBpedia Mobile: A Location-Enabled Linked Data Browser. Ldow, 369, 2008.

Becker, C., & Bizer, C. (2009). Marbles.

Berners-Lee, T. (1997). Axioms of web architecture: Metadata. Retrieved April, 8, 2006.

Berners-Lee, T., & Connolly, D. (2004). Delta: an ontology for the distribution of differences between RDF graphs. World Wide Web, <http://www.w3.org/DesignIssues/Diff>, 4(3), 4-3

Berners-Lee, T., Chen, Y., Chilton, L., Connolly, D., Dhanaraj, R., Hollenbach, J., ... & Sheets, D. (2006, November). Tabulator: Exploring and analyzing linked data on the semantic web. In Proceedings of the 3rd international semantic web user interaction workshop (Vol. 2006, p. 159).

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες

Berners-Lee, T., Connolly, D., Kagal, L., Scharf, Y., & Hendler, J. (2008). N3logic: A logical framework for the world wide web. *Theory and Practice of Logic Programming*, 8(3), 249-269.

Berners-Lee, T., Fielding, R., & Masinter, L. (2004). Uniform resource identifier (URI): Generic syntax (No. RFC 3986).

Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The semantic web. *Scientific american*, 284(5), 28-37.

Bizer, C., Heath, T., & Berners-Lee, T. (2011). Linked data: The story so far. In *Semantic services, interoperability and web applications: emerging concepts* (pp. 205-227). IGI Global.

Board, D. U. (2014). DCMi Metadata Terms. 2012 <http://dublincore.org/documents/dcmi-terms>.

Brickley, D. & Guha, R.V. (2014). RDF Schema 1.1

Byrne, Gillian & Goddard, Lisa. (2010). *The Strongest Link: Libraries and Linked Data*.

Calhoun, K. (2011). The changing nature of the catalog and its integration with other discovery tools, in McIntosh, J. (Ed.), *Cataloging and Indexing*, Apple Academic Press, Oakville, pp. 123-173.

Cardoso, J., & Sheth, A. (2006). The semantic web and its applications. In *Semantic Web Services, Processes and Applications* (pp. 3-33). Springer, Boston, MA.

Cheng, G., Wu, H., Gong, S., Ge, W., & Qu, Y. (2010). Falcons Explorer: Tabular and relational end-user programming for the Web of data. *Semantic Web Challenge*.

Cormode, G., & Krishnamurthy, B. (2008). Key differences between Web 1.0 and Web 2.0. *First Monday*, 13(6).

Coyle, Karen. (2005). *Management of RFID in Libraries*.

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες

Coyle, Karen & Diane Hillman. (2007). Resource Description and Access (RDA): Cataloging Rules for the 20th Century. DLib Magazine 13 no.1/2. <http://www.dlib.org/dlib/january07/coyle/01coyle.html>

DCMI, Available: <http://www.dublincore.org/news/2014/>

Dhuria, S., & Chawla, S. (2014). Ontologies for personalized e-learning in the semantic web. International Journal of Advanced Engineering and Nano Technology (IJAENT), 1(4), 13-18.

Dunsire, Gordon & Willer, Mirna. (2011). Standard library metadata models and structures for the Semantic Web

Enzer, Larry (August 31, 2018). The Evolution of the World Wide Web.

Feigenbaum, L., Herman, I., Hongsermeier, T., Neumann, E., & Stephens, S. (2007). The semantic web in action. Scientific American, 297(6), 90-97.

Frauenfelder, M. (2001). A smarter web. TECHNOLOGY REVIEW-MANCHESTER NH-, 104(9), 52-59.

Graham, Klyne & Jeremy, Carroll. (2006). Resource Description Framework (RDF): Concepts and Abstract Syntax.

Greenberg, J., Sutton, S., & Campbell, D. G. (2003). Metadata: a fundamental component of the semantic web. Bulletin of the American Society for Information Science and Technology, 29(4), 16-18.

Hannemann, Jan & Kett, Jürgen. (2010). Linked Data for Libraries.

Hahn, R., Bizer, C., Sahnwaldt, C., Herta, C., Robinson, S., Bürge, M., ... & Scheel, U. (2010, May). Faceted wikipedia search. In International Conference on Business Information Systems (pp. 1-11). Springer, Berlin, Heidelberg.

Heath, T. (2008a). Information-seeking on the Web with Trusted Social Networks – from Theory to Systems. PhD Thesis, The Open University.

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες

Heath, T. (2008b). How Will We Interact with the Web of Data? In IEEE Internet Computing, Vol. 12(5): 88-91.

Heath, T., Motta, E. (2008). Revyu: Linking reviews and ratings into the Web of Data. Journal of Web Semantics, 6(4):266-273.

Hendler, J., Brners-Lee, T., & Miller, E. (2002). Integrating applications on the semantic web. JOURNAL-INSTITUTE OF ELECTRICAL ENGINEERS OF JAPAN, 122(10), 676-680.

Hitzler, P., Krötzsch, M., Parsia, B., Patel-Schneider, P.F. & Rudolph, S. (2012). OWL 2 Web Ontology Language Primer.

Illien, Gildas. (2012). Are You Ready to Dive in? A Case for Open Data in National Libraries. Paper Presented at the IFLA World Library and Information Congress, Helsinki.

Jentzsch, A., Zhao, J., Hassanzadeh, O., Cheung, K. H., Samwald, M., & Andersson, B. (2009, September). Linking open drug data. In I-SEMANTICS.

Kasse, Paraskevi. (2011). Αξιοποίηση Τεχνολογιών Ανοικτού Κώδικα για την ανάπτυξη εφαρμογών Σημασιολογικού Ιστού.

Kuck, G. (2004). Tim Berners-Lee's Semantic Web. South African Journal of information management, 6(1).

Library of Congress. (2012). Bibliographic Framework as a Web of Data: Linked Data Model and Supporting Services.

Library of Congress. Available <https://www.loc.gov/bibframe/>

Malmsten, M. (2008). Making a library catalogue part of the Semantic Web. In J. Greenberg (Ed.), Proceedings of the International Conference on Dublin Core and Metadata. Metadata for semantic and social applications, 22-26 September 2008 (pp. 146-152). Berlin: Universitätsverlag Göttingen.

Manola, F., Miller, E., & McBride, B. (2014). RDF 1.1 Primer. W3C Working Group Note, February, 25.

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες

McGuinness, D. L. (2002). Ontologies come of age. Spinning the semantic web: bringing the World Wide Web to its full potential, 171-194.

Murugesan, S. (2007). "Understanding Web 2.0," in IT Professional, vol. 9, no. 4, pp. 34-41,

elson, T. H. (1965, August). Complex information processing: a file structure for the complex, the changing and the indeterminate. In Proceedings of the 1965 20th national conference (pp. 84-100). ACM.

Papadakis, I., Kyprianos, K., & Stefanidakis, M. (2015). Linked data URIs and libraries: the story so far. D-Lib Magazine, 21(5/6).

Pesch, Oliver & Miller, Eric. (2016). Using BIBFRAME and Library Linked Data to Solve Real Problems: An Interview with Eric Miller of Zepheira. The Serials Librarian. 71. 1-8. 10.1080/0361526X.2016.1183159

Sauermann, L., Cyganiak, R. (2008): Cool URIs for the Semantic Web. W3C Interest Group Note.

Schreur, Phillip. (2018). The Evolution of BIBFRAME: from MARC Surrogate to Web Conformant Data Model.

Semantic Web, Available: http://en.wikipedia.org/wiki/Semantic_Web

Spanos, D. (2015). Συνδεδεμένα Δεδομένα: Μια ευκαιρία για τις ελληνικές Βιβλιοθήκες. Κύρια άρθρα.

Spiliopoulos, V., Kotis, K., & Vouros, G. A. (2008). Semantic retrieval and ranking of Semantic Web documents using free-form queries. IJMSO, 3(2), 95-108

Stefanidakis, M. et al., 2016. Open Linked Data and Related Applications.

Stefanidakis, M., Andronikos, T., Papadakis, I., Στεφανιδάκης, Μ., Ανδρόνικος, Θ., & Παπαδάκης, Ι. (2015). Open Linked Data and Related Applications.

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες

Steele D. Thomas. (2018). "What comes next: understanding BIBFRAME", Library Hi Tech, [https:// doi.org/10.1108/LHT-06-2018-0085](https://doi.org/10.1108/LHT-06-2018-0085)

Swartz, A. (2002). The semantic web in breadth. URL: <http://logicerror.com/semanticWeb-long>

Tennant, Roy. (2002). Marc must die.

Tennant, Roy. (2004). A bibliographic metadata infrastructure for the twenty-first century.

Tharani, K. (2015), Linked data in libraries: a case study of harvesting and sharing bibliographic metadata with BIBFRAME”, Information Technology and Libraries, Vol. 34 No. 1, pp. 5-19.

Tsetsos, Vassileios & Papataxiarhis, Vassilis & Hadjiefthymiades, Stathes. (2009). Personalizing pedestrian location services through ontologies and rules. Personalization of Interactive Multimedia Services: A Research and Development Perspective. 333-349

Tummarello, G., Cyganiak, R., Catasta, M., Danielczyk, S., Delbru, R., & Decker, S. (2010). Sig. ma: Live views on the web of data. Web Semantics: Science, Services and Agents on the World Wide Web, 8(4), 355-364

Vila-Suero, D. & Gomez-Perez, A. (2013). datos.bne.es and MARiMbA: an insight into library linked data. Library Hi Tech, 31(4), 575-601. doi: 10.1108/LHT-03-2013-0031

Wahid, Nazia. (2017). Mapping the cataloguing practices in information environment: a review of linked data challenges.

Wittenburg, P., Broeder, D., Strömqvist, S., Van Veenendaal, R., & Ringersma, J. (2008). Digital Archive Federations. Talk presented at IASA. Sydney. 2008-09-19.

Yeung, C. M. A., Liccardi, I., Lu, K., Seneviratne, O., & Berners-Lee, T. (2009, January). Decentralization: The future of online social networking. In W3C Workshop on the Future of Social Networking Position Papers (Vol. 2, pp. 2-7).

Yu, L. (2007). Introduction to the semantic web and semantic web services. Chapman and Hall/CRC.

Μυροφόρα Καυζαλίδου - Συνδεδεμένα Δεδομένα και η αξιοποίηση τους από τις βιβλιοθήκες

Γιατράς, Δ., 2011. Συνδεδεμένα Δεδομένα: Επισκόπηση της τεχνολογίας και των υπαρχόντων συστημάτων.

Καλτσογιάννης, Α. (2007). WEB 2.0: Χαρακτηριστικά και επίδρασή του σε επιχειρήσεις, κεντρική διοίκηση και χρήστες. Παρατηρητήριο για την Κοινωνία της Πληροφορίας ΑΕ, Νοέμβριος.

Πεπονάκης, Μ. (2013). Οι ακαδημαϊκές βιβλιοθήκες ως πάροχοι δεδομένων στο σημασιολογικό ιστό: η μετάβαση από τη μία εγγραφή ανά πόρο στα δίκτυα σχέσεων μεταξύ αυτόνομων οντοτήτων.