

Θεματική ενοποίηση δημοσίων καταλόγων και ψηφιακών βιβλιοθηκών

Δημήτρης Γαβρίλης, Κωνσταντία Κακάλη, Άννα Διακάκη,
{gavrilis, nkakal, annad }@panteion.gr

Περίληψη: Στην παρούσα εργασία παρουσιάζεται η νέα ψηφιακή βιβλιοθήκη του Παντείου Πανεπιστημίου – ΠΑΝΔΗΜΟΣ (ΠΑΝτειακές ΔΗΜΟΣιεύσεις). Παρουσιάζεται η προτεινόμενη αρχιτεκτονική της βιβλιοθήκης, η υλοποίησή της και τα προβλήματα που παρουσιάστηκαν, καθώς και οι πολιτικές πρόσβασης και διαχείρισης. Πιο συγκεκριμένα, αναλύεται ο προβληματισμός της επιλογής σχήματος μεταδεδομένων σε συνδυασμό με τη χρήση καθιερωμένων όρων στην ψηφιακή βιβλιοθήκη. Επιπλέον, παρουσιάζεται ενσωμάτωση της ψηφιακής βιβλιοθήκης στο σύστημα ενοποιημένης αναζήτησης της Βιβλιοθήκης, η προβληματική που δημιουργήθηκε από τα ζητήματα αντιστοίχισης μεταξύ προτύπων μεταδεδομένων και η τελική απόφαση.

Λέξεις κλειδιά: Ακαδημαϊκά αποθετήρια, ψηφιακές βιβλιοθήκες, μεταδεδομένα, σημασιολογική διαλειτουργικότητα, θεματικές επικεφαλίδες

1. Εισαγωγή

Η ανάπτυξη του Διαδικτύου και η ανάγκη διαχείρισης της παρεχόμενης από αυτό πληροφορίας δεν μπορούν να μας αφήσουν αδιάφορους. Η ιδέα της δημιουργίας του σημασιολογικού ιστού (semantic web), με κύριο σκοπό την οργάνωση της πληροφορίας με την χρήση τόσο των παραδοσιακών βιβλιοθηκονομικών εργαλείων (θησαυρών, ελεγχόμενων λεξιλογίων, ταξινομικών συστημάτων), όσο και των νέων τα οποία βασίζονται σε αυτά (ταξονομίες, οντολογίες) εμπλέκει σε μεγάλο βαθμό την βιβλιοθηκονομική κοινότητα.

Η ανάγκη για διαλειτουργικότητα και υιοθέτηση διεθνών βιβλιοθηκονομικών εργαλείων, προτύπων και κανόνων ήταν ένα ζήτημα που ανέκαθεν απασχολούσε τους βιβλιοθηκονόμους, ενώ σήμερα τους απασχολεί ακόμα περισσότερο, ενόψει της αναγκαιότητας για ενοποιημένη αναζήτηση και ανάκτηση της πληροφορίας, προερχόμενης από διαφορετικά περιβάλλοντα. Ως εκ τούτου, η χρήση των παραδοσιακών βιβλιοθηκονομικών εργαλείων, ενσωματωμένων στα κατάλληλα πρότυπα μεταδεδομένων, στο περιβάλλον μίας ψηφιακής βιβλιοθήκης είναι κάτι το οποίο δεν μπορούμε να αγνοήσουμε.

Τα διαφορετικά εννοιολογικά εργαλεία (θεματικές επικεφαλίδες, θησαυροί, ελεγχόμενα λεξιλόγια), που χρησιμοποιούν σήμερα οι περισσότερες βιβλιοθήκες, δεν είναι συμβατά μεταξύ τους. Με την ανάπτυξη του ψηφιακού περιβάλλοντος, τη δημιουργία πολλαπλών καταλόγων και βιβλιογραφικών βάσεων δεδομένων, καθώς και ακαδημαϊκών αποθετηρίων, η εννοιολογική ανομοιογένεια μεγαλώνει και το γεγονός αυτό επηρεάζει αρνητικά τη δυνατότητα ενοποιημένης αναζήτησης και ανταλλαγής μηχαναγνώσιμων δεδομένων, στον ενιαίο χώρο της πληροφόρησης.

Κατά την διαδικασία προσδιορισμού των απαιτήσεων για την δημιουργία της ψηφιακής βιβλιοθήκης του Παντείου Πανεπιστημίου, Πάνδημος (<http://library.panteion.gr/pandemos>), τα ζητήματα αυτά αποτέλεσαν το βασικό προβληματισμό μας. Το ζήτημα της ενοποίησης ψηφιακού και ηλεκτρονικού καταλόγου με τη χρήση κοινών εργαλείων αποτελεί βασικό στόχο για την ολοκληρωμένη πρόσβαση των χρηστών στην πληροφορία.

Στόχος του Πάνδημου είναι να αποτελέσει τον χώρο ηλεκτρονικής δημοσίευσης της συγγραφικής δραστηριότητας των μελών της επιστημονικής κοινότητας του Πανεπιστημίου. Στον Πάνδημο κατατίθενται σε χωριστές συλλογές οι μεταπτυχιακές εργασίες των ΠΜΣ, οι διδακτορικές διατριβές καθώς και οι πανεπιστημιακές σημειώσεις των προπτυχιακών και μεταπτυχιακών μαθημάτων. Οι συλλογές αυτές οργανώνονται ανά κοινότητα που αποτελούν αρχικά τα Τμήματα του Πανεπιστημίου. Ιδιαίτερη κοινότητα του Πανδήμου αποτελεί αυτή

των Ελληνικών περιοδικών, που περιλαμβάνει τα ηλεκτρονικά άρθρα περιοδικών, που έχουν δοθεί στη Βιβλιοθήκη.

Η προσέγγιση που ακολουθείται και παρουσιάζεται σε αυτή την εργασία είναι αυτή της θεματικής ομογενοποίησης μέσω της ενσωμάτωσης των θεμάτων του καταλόγου της βιβλιοθήκης στα θέματα της ΨΒ. Η διαδικασία αυτή παρουσιάζει αρκετά προβλήματα τόσο σε επίπεδο σχεδιασμού, όπως το πρόβλημα της αντιστοίχισης σχημάτων μεταδεδομένων (metadata mappings), όσο και σε επίπεδο υλοποίησης, όπως το πρόβλημα της ενσωμάτωσης των καθιερωμένων όρων από το σύστημα αυτοματοποίησης της βιβλιοθήκης στην ψηφιακή βιβλιοθήκη.

2. Λειτουργικές Απαιτήσεις

Το πρώτο στάδιο, ως προς το σχεδιασμό μιας ψηφιακής βιβλιοθήκης, περιλαμβάνει τον καθορισμό των απαιτήσεων που θα πρέπει να πληρεί. Οι απαιτήσεις – προδιαγραφές της ψηφιακής βιβλιοθήκης έχουν να κάνουν με παράγοντες, όπως το σχήμα μεταδεδομένων που χρησιμοποιεί, τις διάφορες πολιτικές πρόσβασης-διαχείρισης που υλοποιεί, θέματα καθιερωμένων όρων, δυνατότητες προσβασιμότητας-αναζήτησης, δυνατότητες συμβατότητας με άλλα συστήματα. Αφού καθοριστούν οι απαιτήσεις, θα προκύψουν οι προδιαγραφές βάσει των οποίων θα υλοποιηθεί η ψηφιακή βιβλιοθήκη.

Για τον καθορισμό της πλατφόρμας ανάπτυξης του ΠΑΝΔΗΜΟΥ, εξετάστηκαν τα συστήματα ψηφιακών βιβλιοθηκών DSpace, Fedora και Greenstone. Και τα τρία συστήματα χρησιμοποιούν το Dublin Core ως σχήμα μεταδεδομένων. Από αυτά, επιλέχθηκαν τα δύο πρώτα λόγω της στιβαρότητάς τους σε σχέση με το τρίτο. Μεταξύ του DSpace και του Fedora, επιλέχθηκε το πρώτο λόγω της δυνατότητάς του (μέχρι τη στιγμή της εξέτασης) να υποστηρίζει κοινότητες χρηστών, σε αντίθεση με το Fedora. Άλλα πλεονεκτήματα του DSpace είναι ότι είναι ιδιαίτερα εύκολο στην εγκατάσταση και έχει καλή τεκμηρίωση. Ενδεικτικό είναι ότι η διαδικασία εγκατάστασής του δεν πήρε πάνω από μια ώρα, ενώ όλη η διαδικασία παραμετροποίησης και μετάφρασης της διεπαφής στα ελληνικά ολοκληρώθηκε σχεδόν σε μία εβδομάδα.

2.1. Βασικές πολιτικές

Οι πολιτικές του ΠΑΝΔΗΜΟΥ διαμορφώθηκαν ως εξής:

Πολιτική πρόσβασης: Όλα τα τεκμήρια είναι προσβάσιμα από όλους.

Πολιτική υποβολής: Η υποβολή είναι πλήρως ελεγχόμενη. Έχει κατασκευαστεί ειδική φόρμα στο πληροφοριακό σύστημα της Βιβλιοθήκης, μέσω της οποίας οι χρήστες μπορούν να υποβάλλουν τα τεκμήρια τους. Η περαιτέρω διαδικασία έχει ως εξής: α)τα μεταδεδομένα του χρήστη ελέγχονται, εμπλουτίζονται από εξειδικευμένο προσωπικό. β)οι θεματικές επικεφαλίδες αντιστοιχίζονται με αυτές του καταλόγου της Βιβλιοθήκης και γ) στα αρχεία προστίθεται εξώφυλλο για λόγους ομοιομορφίας, μετατρέπονται σε PDF και βελτιστοποιούνται ως προς το μέγεθος.

Πολιτική ενημέρωσης: Όλοι οι χρήστες μπορούν να τροποποιήσουν τα τεκμήρια τους μετά από αίτηση και ακολουθώντας διαδικασία αντίστοιχη με αυτή της υποβολής.

Πολιτική απόσυρσης: Μόνο οι σημειώσεις μαθημάτων μπορούν να αποσυρθούν από τον ΠΑΝΔΗΜΟ, αλλά κρατούνται σε αρχείο.

2.2. Διαλειτουργικότητα

Η διαλειτουργικότητα είναι η δυνατότητα ανταλλαγής και ενοποίησης (integration) μηχαναγνώσιμων δεδομένων που προέρχονται από διαφορετικά πληροφοριακά περιβάλλοντα μέσω της υιοθέτησης κοινών προτύπων.

Ένα σύστημα ψηφιακής βιβλιοθήκης θα πρέπει να μπορεί να υποστηρίζει γνωστά πρωτόκολλα διαλειτουργικότητας, όπως το OpenURL και το Open Archives Initiative for

metadata exchange (OAI). Τα δύο αυτά πρωτόκολλα θεωρήθηκαν απαραίτητα για την επιλογή του συστήματος ψηφιακής βιβλιοθήκης. Μια άλλη απαίτηση πιο εξεζητημένη είναι ένας μηχανισμός για τη σύνδεση της ψηφιακής βιβλιοθήκης με συστήματα ενοποιημένης αναζήτησης (federated search), έτσι ώστε να μπορεί ο χρήστης να αναζητά π.χ. από τον ιστότοπο της βιβλιοθήκης ταυτόχρονα σε όλες τις ηλεκτρονικές πηγές της.

Όσον αφορά τα πρωτόκολλα διαλειτουργικότητας που υποστηρίζει το DSpace, τα OpenURL και OAI υποστηρίζονται απευθείας, οπότε και δεν υπήρξε πρόβλημα. Παρόλα αυτά, το DSpace δεν παρέχει κάποιο τρόπο σύνδεσής του με το σύστημα ενοποιημένης αναζήτησης της Βιβλιοθήκης του Παντείου Πανεπιστημίου. Για αυτό το λόγο, κατασκευάστηκε ένα Application Programming Interface (API) το οποίο έδωσε τη δυνατότητα στους χρήστες του ιστότοπου ή/και της θεματικής πύλης (<http://library.panteion.gr/>) της Βιβλιοθήκης να αναζητούν τεκμήρια στον ΠΑΝΔΗΜΟ. Η υπηρεσία αυτή είναι προσβάσιμη και μέσω της υποδομής που αναπτύχθηκε για την αναζήτηση και πλοήγηση των χρηστών στις φυσικές και ψηφιακές συλλογές της Βιβλιοθήκης με τη χρήση ασύρματων φορητών συσκευών (PDAs, κινητά τηλέφωνα).

2.3. Καθιερωμένοι Όροι

Στην αρχή της ανάπτυξης των υπηρεσιών πληροφόρησης του Διαδικτύου, η συστηματική οργάνωση γνώσης, η ευρετηρίαση θεμάτων και η ταξινόμηση θεωρήθηκε πως θα αντικατασταθούν από την ευρετηρίαση πλήρους κειμένου (όπως οι μηχανές αναζήτησης) και από τη πλοήγηση μέσω των συνδέσμων των υπερκειμένων. Ωστόσο, παρατηρείται το τελευταίο καιρό μια αναγέννηση της μάλλον "παραδοσιακής" θεματικής πρόσβασης. Όλο και περισσότερο συνειδητοποιείται ότι η οργάνωση γνώσης με τα ελεγχόμενα λεξιλόγια/KOS (συστήματα οργάνωσης γνώσης) όπως οι θησαυροί, τα συστήματα ταξινόμησης, οι ταξονομίες, τα γεωγραφικά λεξικά, οι οντολογίες κ.λπ. είναι τόσο χρήσιμη για την ανακάλυψη των πόρων και τη πλοήγηση της πληροφορίας όπως ήταν και στον κόσμο των έντυπων μέσων. (Kohn, 2006). Καθώς η δημιουργία του σημασιολογικού ιστού έχει μεταθέσει αυτή την πρακτική και στο ψηφιακό περιβάλλον και πολλές προσπάθειες προσανατολίζονται στη μεταφορά των θεματικών επικεφαλίδων, των θησαυρών και άλλων συστημάτων διαχείρισης γνώσης σε ψηφιακή μορφή (Hodge, 2000).

Η χρήση καθιερωμένων όρων στο περιβάλλον μίας ψηφιακής βιβλιοθήκης είναι θεμελιώδους σημασίας, καθώς συμβάλλει στην συλλογική διαχείριση των εννοιών από τους βιβλιοθηκονόμους και στην άμεση πρόσβαση του χρήστη στην πληροφορία που επιθυμεί, αλλά ειδικά στην δημιουργία του ΠΑΝΔΗΜΟΥ, βασική απαίτηση υπήρξε η χρήση των θεματικών επικεφαλίδων, με βάση το υλικό που τεκμηριώνει. Για να γίνει σωστά η τεκμηρίωση με τους καθιερωμένους όρους και παραπέρα η ενιαία αναζήτηση, θα πρέπει οι όροι αυτοί να ταυτίζονται με αυτούς του καταλόγου της βιβλιοθήκης.

Ο ηλεκτρονικός κατάλογος της Βιβλιοθήκης διαθέτει μία βάση δεδομένων θεματικών επικεφαλίδων της Βιβλιοθήκης του Κογκρέσου, η οποία αναπτύσσεται και ενημερώνεται συνεχώς. Επίσης, σε περιπτώσεις που αυτές δεν επαρκούν για την θεματική περιγραφή, χρησιμοποιείται και ο Κατάλογος των Καθιερωμένων Επικεφαλίδων της Εθνικής Βιβλιοθήκης της Ελλάδας.

Η μαζική εισαγωγή καθιερωμένων όρων του καταλόγου της Βιβλιοθήκης υποστηρίζεται από το DSpace μέσω της καταχώρησης τους σε ένα xml αρχείο. Ο χρήστης μπορεί αντί να πληκτρολογήσει τον όρο, να επιλέξει από το δέντρο των καθιερωμένων όρων. Για τη σύνδεση του αρχείου αυτού με τους καθιερωμένους όρους του κυρίου καταλόγου (ADVANCE) κατασκευάστηκε ένας parser σε Java, ο οποίος μετατρέποντας τα θέματα του καταλόγου από UNIMARC σε UNICODE και XML έδωσε τη δυνατότητα της σύνδεσης των θεματικών επικεφαλίδων του ADVANCE με τους καθιερωμένους όρους του DSpace. Σημαντικό πρόβλημα αποτέλεσε το γεγονός ότι οι θεματικές επικεφαλίδες του καταλόγου

είναι πάρα πολλές (περίπου 48000) και το DSpace δεν μπορεί να τις απεικονίσει με ικανοποιητική ταχύτητα.



Εικόνα 1 Θεματικός κατάλογος του αποθετηρίου

Για τη θεματική ευρετηρίαση χρησιμοποιήθηκε το ήδη υπάρχον στοιχείο του DC subject.lcsh και προστέθηκε ένας νέος προσδιοριστής στο στοιχείο subject, ο nlg για την ένταξη των θεματικών επικεφαλίδων της Εθνικής Βιβλιοθήκης της Ελλάδας. Οι θεματικές επικεφαλίδες που περιγράφουν το κάθε τεκμήριο εντάσσονται στον θεματικό κατάλογο του DSpace, δίνοντας στους χρήστες την δυνατότητα πλοήγησης, βάσει θέματος.

Η χρήση όμως των προσδιοριστών στο στοιχείο Subject είχε ως αποτέλεσμα, κατά την εμφάνισή τους στο χρήστη, τα θέματα να διαχωρίζονται και να μην παρουσιάζονται ενιαία (όπως συμβαίνει στον ηλεκτρονικό κατάλογο). Πρέπει να σημειωθεί εδώ ότι το Dublin Core αδυνατεί να ενσωματώσει τόσο πλούσια πληροφορία, όπως αυτή των LSCH.

2.4. Σχήματα Μεταδεδομένων

Το σχήμα μεταδεδομένων μπορεί να θεωρηθεί ως η καρδιά μιας ψηφιακής βιβλιοθήκης. Αυτό καθορίζει τι πληροφορία θα καταγράφεται, τις δυνατότητες παρουσίασης της πληροφορίας των τεκμηρίων στο χρήστη και την αντιστοίχιση των μεταδεδομένων αυτών με άλλα σχήματα (για λόγους συμβατότητας).

Η απλότητα του DC συντελεί στο γεγονός ότι γίνεται εύκολα κατανοητό και δεν απαιτεί εξειδικευμένες γνώσεις για τη χρήση του (καθιστώντας το κατάλληλο για χρήση από «μη εξειδικευμένους» χρήστες), απεναντίας όμως δεν μπορεί να εκφράσει πιο πολύπλοκες σχέσεις (Beall, 2004). Η βασική του αδυναμία, σε αντίθεση με τα άλλα πιο πλούσια σχήματα, όπως το MARC, συνίσταται στο γεγονός ότι δεν μπορεί να κάνει χρήση και προσδιορισμό των καθιερωμένων όρων ως ονομάτων, ή θεμάτων, ή γεωγραφικών όρων.

Το σχήμα μεταδεδομένων που προτάθηκε για την ψηφιακή βιβλιοθήκη ΠΑΝΔΗΜΟΣ ήταν το Metadata Object Description Schema (MODS), καθώς και το MADS (Metadata Authority Description Schema) για σύνολα στοιχείων καθιερωμένων όρων, τα οποία αναπτύχθηκαν και υποστηρίζονται από τη Βιβλιοθήκη του Κογκρέσου, αποτελούν παράγωγα του MARC και είναι πλέον από τα καλύτερα και πιο σύγχρονα σχήματα μεταδεδομένων (MacCallum, 2004). Το MODS αποτελείται από 20 βασικά στοιχεία (elements) και 64 υποστοιχεία (subelements) το καθένα από τα οποία χαρακτηρίζεται από γνωρίσματα (attributes) τα οποία προσδίδουν μια πιο αναλυτική περιγραφή και υλοποιείται από τη γλώσσα

XML Schema. Βασικό του πλεονέκτημα είναι ότι αντικαθιστά με λεκτικούς χαρακτήρες τα αριθμητικά πεδία και υποπεδία του MARC. Είναι συμβατό με τα παραδοσιακά σχήματα καταλογογράφησης, έχει σημασιολογική συμβατότητα με το ISO 2709. Το MODS, σε αντίθεση με το MARC, παρέχει τις δυνατότητες αφομοίωσης μεγάλου όγκου δεδομένων, λεπτομερούς σήμανσης και ανάδειξης των ιεραρχικών σχέσεων μεταξύ τους. Επίσης, μπορεί να ανταποκριθεί τόσο στις ανάγκες του ψηφιακού όσο και του συμβατικού υλικού μιας βιβλιοθήκης. Οφείλουμε να σημειώσουμε ότι αποτελεί ένα ιδιαίτερα ευέλικτο πρότυπο, απλούστερο της διάταξης MARC, ενώ δίνει την δυνατότητα διαφόρων επιπέδων λεπτομέρειας κατά την περιγραφή των τεκμηρίων και ηλεκτρονικών πόρων, γεγονός που απορρέει από το ότι όλα τα πεδία του είναι επαναλαμβανόμενα, ενώ κανένα από αυτά δεν είναι υποχρεωτικό (Andresen, 2004).

3. Θέματα αντιστοίχισης μεταδεδομένων

Το DSpace χρησιμοποιεί ενδογενώς το DC για την εσωτερική αναπαράσταση των μεταδεδομένων, δίνοντας παράλληλα τη δυνατότητα στο χρήστη να ορίσει τα δικά του μεταδεδομένα. Για την ενσωμάτωση όμως άλλου πρότυπου μεταδεδομένων στο Dspace απαιτείται αντιστοίχιση. Η αντιστοίχιση (mapping) είναι η νοητική λειτουργία του καθορισμού της σημασιολογικής ισοδυναμίας ανάμεσα σε στοιχεία πρωταρχικού προτύπου μεταδεδομένων προς το επιδιωκόμενο πρότυπο μεταδεδομένων. Περιλαμβάνει την προδιαγραφή μιας αντιστοιχίας κάθε στοιχείου στο αρχικό πρότυπο μεταδεδομένων με ένα σημασιολογικά ισοδύναμο στοιχείο στο τελικό πρότυπο μεταδεδομένων.

Η υλοποίηση της αντιστοίχισης δεν ήταν ένα εύκολο εγχείρημα. Κάθε στοιχείο του MODS έπρεπε να αντιστοιχηθεί σε ένα στοιχείο του DC. Η αντιστοίχιση από ένα ειδικότερο πρότυπο (αναλυτικότερο και μεγαλύτερο σε στοιχεία μεταδεδομένων) σε ένα γενικότερο (με λιγότερα και γενικής σημασίας στοιχεία μεταδεδομένων) γίνεται ιδιαίτερα δύσκολη.

Πιο συγκεκριμένα, τα στοιχεία του MODS οργανώνονται σε υπο-στοιχεία με πολλαπλά επίπεδα ιεραρχίας υποπεδίων, κάθε στοιχείο ή υπο-στοιχείο διαθέτει ένα πλήθος των γνωρισμάτων (attributes) και κάθε γνώρισμα περιλαμβάνει μία ή περισσότερες επιτρεπτές τιμές. Αντίθετα το DC έχει ένα επίπεδο ενσωμάτωσης μόνο και δεν χρησιμοποιεί γνωρίσματα.

Θεωρητικά, για την αντιστοίχιση ενός σύνθετου σχήματος μεταδεδομένων (με πολλαπλά γνωρίσματα) σε ένα απλό σχήμα μεταδεδομένων (με ένα επίπεδο ενσωμάτωσης χωρίς γνωρίσματα) θα ισχύουν τα εξής: έστω ένα πεδίο x του σύνθετου σχήματος με N γνωρίσματα: y_1, y_2, \dots, y_N καθένα από τα οποία μπορεί να πάρει $k_i(y) = |y_i|$ τιμές όπου $i = 1, \dots, N$. Ο συνολικός αριθμός διακριτών τιμών όλων των γνωρισμάτων είναι:

$$K = \sum_{i=1}^N k_i = \sum_{i=1}^N |y_i|$$

Για να αντιστοιχηθούν αυτά στη βάση του απλού σχήματος, τα απαραίτητα πεδία:

$${}_K C_N = \binom{K}{N} = \frac{K!}{N!(K-N)!} = \frac{\left(\sum_{i=1}^N |y_i|\right)!}{N! \left(\sum_{i=1}^N |y_i| - N\right)!}, \quad K \geq N$$

Όπως φαίνεται και από την παραπάνω εξίσωση, μια τέτοια μετατροπή εξαρτάται από δύο παράγοντες: τον αριθμό των γνωρισμάτων και το συνολικό αριθμό των δυνατών τιμών όλων των γνωρισμάτων. Δηλαδή, όσο περισσότερα γνωρίσματα έχουμε και όσο περισσότερες τιμές έχουν αυτά, τόσο περισσότερα στοιχεία για το DC θα χρειαστούμε. Θα πρέπει επίσης να σημειωθεί, ότι η παραπάνω ανάλυση δεν λαμβάνει υπόψη την ενσωμάτωση των πολλαπλών επιπέδων του σύνθετου σχήματος μιας και κάτι τέτοιο δεν επηρεάζει την αντιστοίχιση.

Εφαρμόζοντας τα παραπάνω στο MODS (σύνθετο σχήμα) και το DC (απλό σχήμα), μόνο για 4 πεδία, η απαιτούμενη ανάπτυξη της δομής της βάσης δεδομένων γίνεται:

Πεδίο MODS	Γνωρίσματα	N	K	C
titleInfo.Title	type (4 τιμές)	1	4	4
originInfo.place.placeTerm	type (2 τιμές) authority (3 τιμές)	2	5	10
Name	type (3 τιμές)	1	3	3
Name.namePart	type (3 τιμές)	1	3	3
Συνολικός αριθμός πεδίων DC:				20

Σύμφωνα με το παραπάνω πίνακα, ο αριθμός των στοιχείων στο DC που θα χρειαστεί να κατασκευαστούν για τα παραπάνω 4 πεδία του MODS είναι 20. Είναι λοιπόν εμφανές πως κάτι τέτοιο είναι ιδιαίτερα δύσκολο ειδικά αν αναλογιστεί κανείς πως το MODS περιλαμβάνει δεκάδες πεδία και όλα σχεδόν έχουν τουλάχιστον ένα γνώρισμα με κατά μέσο όρο 3 τιμές. Στην περίπτωση που κάποιος θέλει να λάβει υπόψη τα πολλαπλά επίπεδα ιεραρχίας του MODS, μια σωστή αντιστοίχιση σε βάση ενός μόνο επιπέδου είναι αδύνατη χωρίς την τροποποίηση της ίδιας της βάσης της εφαρμογής.

Η επέκταση του DC για την αντιστοίχσή του με τα στοιχεία του MODS δεν είναι αποδεκτή λύση για λόγους διαλειτουργικότητας. Ωστόσο μια αντιστοίχιση του MODS στο DC- qualified χωρίς την απαιτούμενη επέκταση προκαλεί απώλεια πληροφορίας. Τελικά, η ιδέα της ενσωμάτωσης του MODS σε σύστημα όπως το DSpace εγκαταλείφθηκε. Αποφασίστηκε η χρήση του MODS να γίνει με την ανάπτυξη νέας πλατφόρμας ψηφιακής βιβλιοθήκης, που να το χρησιμοποιεί ενδογενώς, αποφεύγοντας έτσι τα προβλήματα που προκύπτουν. Παρόλα αυτά το DSpace παρέμεινε μιας και αποτελεί ένα ολοκληρωμένο και δοκιμασμένο σύστημα, το οποίο μπορεί να καλύψει τις βασικές ανάγκες των περισσότερων βιβλιοθηκών.

5. Συμπεράσματα

Ο ΠΑΝΔΗΜΟΣ αποτελεί μια ψηφιακή βιβλιοθήκη η οποία εισάγει αρκετές καινοτομίες. Η θεματική ενοποίηση με τον κατάλογο της βιβλιοθήκης έχει επιτευχθεί παρά τις δυσλειτουργίες της εργασίας αυτής λόγω των περιορισμών του DC.

Το βασικότερο όμως είναι η δυνατότητα ενιαίας αναζήτησης στους ψηφιακούς πόρους της Βιβλιοθήκης.

Για πρώτη φορά ίσως, έγινε προσπάθεια υιοθέτησης του σχήματος μεταδεδομένων MODS για την κατασκευή μιας ψηφιακής βιβλιοθήκης. Τα αποτελέσματα αυτής της προσπάθειας ήταν ιδιαίτερα σημαντικά και καθοριστικά για τις περαιτέρω αποφάσεις στη Βιβλιοθήκη του Παντείου Πανεπιστημίου. Το εγχείρημα της αντιστοίχισης του MODS με το DC και παρόλο που αυτό είναι τεχνικά δυνατό κατέδειξε αδυναμία του DSpace να ενσωματώσει πλουσιότερα πρότυπα μεταδεδομένων.

Συνολικά, γίνεται φανερή η έλλειψη συστημάτων τα οποία να υποστηρίζουν πρότυπα επεξεργασίας και διαχείρισης καθιερωμένων όρων. Σήμερα, έχουμε ήδη προχωρήσει στην υλοποίηση τέτοιων συστημάτων, ώστε να μπορέσει να ανταπεξέλθει στις ολοένα αυξανόμενες απαιτήσεις στο χώρο των Ψηφιακών Βιβλιοθηκών.

Βιβλιογραφία

Andresen, L. (2004). "After MARC - what then?". *Library Hi Tech*, vol. 22, no. 1, pp. 40-51.

Beall, J. (2004). "Dublin Core : An Obituary." *Library Hi Tech News*. no. 8. pp. 40-1.

- Carvalho, J.R.d., Cordeiro, M.I., Lopes, A. & Vieira, M. (2004), "Meta-information about MARC: an XML framework for validation, explanation and help systems". *Library Hi Tech*, vol. 22, no. 2, pp. 131-137.
- DeZelar-Tiedman, C. (2004). "Crashing the party: cataloguers as digital librarians", *OCLC Systems & Services*, vol. 20, no. 4, pp. 145-147.
- Dublin Core Metadata Initiative (2007) Διαθέσιμο στο: <http://dublincore.org/> [Τελευταία πρόσβαση 5/8/2007].
- Dspace (2007). Διαθέσιμο στο: <http://www.dspace.org/>[Τελευταία πρόσβαση 5/8/2007].
- El-Sherbini, M. & Klim, G. 2004, "Metadata and cataloging practices", *The Electronic Library*, vol. 22, no. 3, pp. 238-248.
- Fox, R. 2006, "Cataloging for the masses", *OCLC Systems and Services*, vol. 22, no. 3, pp. 166-172.
- Hodge, G. (2000), *Systems of Knowledge Organization for Digital Libraries: Beyond Traditional Authority Files*, CLIR, Washington, DC. Διαθέσιμο στο: www.clir.org/pubs/abstract/pub91abst.html [Τελευταία πρόσβαση 5/8/2007].
- Koch, Traugott (2006). Electronic thesis and dissertations services: Semantic interoperability, subject access, multilinguality. *Background paper for the E-Thesis Workshop (JISC, CURL, SURF), Amsterdam, Netherlands*. Διαθέσιμο στο: <http://www.ukoln.ac.uk/ukoln/staff/t.koch/publ/e-thesis-200601.html> [Τελευταία πρόσβαση 10/8/2007].
- Lee, H.-., Jeon, Y.-. & Han, S.-. (2006). "MARCXTM: Topic Maps Modeling of MARC Bibliographic Information". *Lecture notes in computer science*, no. 3873, pp. 241-252.
- Lopatin, L. (2006). "Library digitization projects, issues and guidelines: A survey of the literature". *Library Hi Tech*, vol. 24, no. 2, pp. 273-289.
- Mann, T., & Library of Congress. Professional Guild (2006). The changing nature of the catalog and its integration with other discovery tools, final report, March 17, 2006, prepared for the Library of Congress by Karen Calhoun a critical review [Homepage of AFSCME, Local 2910],
- Martin, K.E. (2006). "Moving into the digital age: A conceptual model for a publications repository". *Internet Reference Services Quarterly*, vol. 11, no. 2, pp. 27-47.
- Metadata Object Description Schema Official Website, July 25, 2007-last update. Διαθέσιμο στο: <http://www.loc.gov/standards/mods/> [2007, August 5]
- McCallum, S.H. (2004) "An introduction to the Metadata Object Description Schema (MODS)", *Library Hi Tech*, vol. 22, no. 1, pp. 82-88
- Open Archives Initiative. Διαθέσιμο στο: <http://www.openarchives.org/> [Τελευταία πρόσβαση 5/8/2007].
- Nicholson, D., Wake, S. and Currier, S. (2001), HILT: High Level Thesaurus Project: Final Report, Διαθέσιμο στο: <http://hilt.cdlr.strath.ac.uk/Reports/Documents/HILTFinalreport.doc>
- Shiri, A. and Molberg K., C. (2005), "Interface to knowledge organization systems in Canadian digital libraries collections". *Online Information Review*, Vol. 29 No. 6, pp. 604-620.

Todd, C. (2003). "Metadata mayhem: Cataloguing electronic resources in the national library of New Zealand". *Electronic Library*, vol. 21, no. 3, pp. 214-222.

Weng, C. & Mi, J. (2006). "Towards accessibility to digital cultural materials: An FRBRized approach", *OCLC Systems and Services*, vol. 22, no. 3, pp. 217-232.