

**Ιόνιο Πανεπιστήμιο
Τμήμα Αρχειονομίας - Βιβλιοθηκονομίας
Μεταπτυχιακό Πρόγραμμα Αθήνας
Έτος 2007-2008**

**Εργασία στα πλαίσια του μαθήματος
«Τεχνολογίες της Πληροφορίας»**

Διαλειτουργικότητα

Μανόλης Πεπονάκης και Παναγιώτης Στάικος

**Υπεύθυνη καθηγήτρια
Κατερίνα Τοράκη**

Ιούνιος 2008

true interoperability ([...] this decade's "greatest lie" in library automation)

Andrew Pace
Αναφέρεται στο Bahr, Ellen (2007)

Πίνακας Περιεχομένων

Γενικά περί διαλειτουργικότητας.....	4
Οι απαιτήσεις και οι ανάγκες στην εποχή της τεχνολογίας.....	4
Τι είναι διαλειτουργικότητα	5
Web Services: η διαλειτουργικότητα σε επίπεδο συστημάτων	7
Διαλειτουργικότητα στις βιβλιοθήκες	9
Δημιουργία και ανταλλαγή μεταδεδομένων	11
Πρόσβαση στα μεταδεδομένα	16
Διαλειτουργικότητα και Z39.50 – SRU/SRW	17
Διαλειτουργικότητα και ψηφιακές βιβλιοθήκες.....	20
Διαλειτουργικότητα σε επίπεδο τεκμηρίων	28
Συμπεράσματα	30
Βιβλιογραφία	34

Γενικά περί διαλειτουργικότητας

Οι απαιτήσεις και οι ανάγκες στην εποχή της τεχνολογίας

Είναι κοινός τόπος πως οι βιβλιοθήκες είναι (ή θα έπρεπε να είναι) ανοικτά συστήματα. Στο χώρο των βιβλιοθηκών είναι αρκετά διαδεδομένη η φράση «οι βιβλιοθήκες δεν έχουν τοίχους». Αυτή η «απουσία τοίχων» θα μπορούσε να ιδωθεί από δύο πλευρές. Τόσο από την άποψη ότι οι βιβλιοθήκες δεν περιορίζονται πλέον από το υλικό που διαθέτουν εντός του κτιρίου τους ή εκείνο που κατέχουν άμεσα (για παράδειγμα μια βιβλιοθήκη μπορεί να προσφέρει στους χρήστες της τα ελληνικά διδακτορικά που συγκεντρώνει και διαθέτει το ΕΚΤ) όσο και ότι η πρόσβαση στο υλικό που μπορούν να διαθέσουν δεν περιορίζεται από την επίσκεψη στο φυσικό χώρο της βιβλιοθήκης (π.χ. ψηφιοποιεί κάποια έργα και τα διαθέτει στο διαδίκτυο).

Η εποχή μας χαρακτηρίζεται από την εμμονή της αξιοποίησης προϋπάρχουσας εργασίας (να μην ξαναγίνει κάτι που έχει ήδη γίνει) και την επιβολή ενιαίας αντιμετώπισης για το σύνολο των ανθρώπων. Η προσέγγιση είναι: όλοι έχουμε τις ίδιες ανάγκες και μπορούμε να τις καλύψουμε με τους ίδιους περίπου τρόπους¹. Ένα ακόμη σημαντικό ζητούμενο της εποχής είναι η συνεργασία. Από τα διάφορα ευρωπαϊκά προγράμματα μέχρι τα consortiums που λαμβάνουν χώρα σε ένα μεγάλο εύρος του πλανήτη όλοι επιδιώκουν συνεργασίες. Η ισχύς εν τη ενώσει, λοιπόν, ώστε να επιτευχθεί ο βασικός στόχος της σύγχρονης κοσμοθεωρίας: η μείωση του κόστους².

Η τεχνολογία, από τη βιομηχανική επανάσταση και μετά, ανέλαβε το ρόλο του ηγέτη στην υπηρεσία του υπέρτατου αξιώματος της εποχής. Από τις πρώτες μηχανές στα εργοστάσια μέχρι την επανάσταση των υπολογιστών στο τέλος του 20^{ου} αιώνα πολλοί ήταν εκείνοι που κήρυτταν ότι πλησιάζει πλέον η εποχή της ευημερίας για το σύνολο της ανθρωπότητας. Η παρούσα εργασία δεν εστιάζει σε ζητήματα κοινωνικής ανάλυσης. Σε κάθε περίπτωση, ωστόσο, η διαλειτουργικότητα, η οποία και αποτελεί το θέμα της εργασίας, είναι επικοινωνία και η επικοινωνία είναι πριν και πάνω απ' όλα μια κοινωνική σχέση (ακόμη και ετυμολογικά αν ιδωθεί).

Για τους δυτικούς του 21^{ου} αιώνα θεωρείται αυτονόητο πως όλοι έχουν δικαίωμα να μπορούν να επικοινωνήσουν με όλους. Όμως η επικοινωνία προϋποθέτει κοινή βάση πάνω στην οποία θα χτιστεί. Η βαθιά πίστη ότι αναγκαία και ικανή συνθήκη για να επικοινωνήσουν οι άνθρωποι αποτελεί το καλώδιο του εκάστοτε ΟΤΕ ήταν και είναι βαθιά ριζωμένη στην κοινωνία μας. Οι τεχνολογίες υπόσχονταν πολλά. Παρά όμως τα καλώδια οπτικών ινών της εποχής μας και τις τεράστιες ταχύτητες του σήματος η επικοινωνία δεν είναι ακόμη αντάξια των προσδοκιών.

Το ζητούμενο της εποχής είναι επικοινωνία μέσω των συστημάτων αλλά ταυτόχρονα επικοινωνία πέραν των συστημάτων κάτι που σημαίνει ότι ναι μεν η επικοινωνία επιτυγχάνεται μέσα από τα συστήματα ωστόσο δεν πρέπει να καθορίζεται από τα

¹ Μπορεί οι κοινωνίες της εποχής μας να φαντάζουν συχνά πολυπολιτισμικές όμως διανύουμε την πιο ομογενοποιημένη φάση της ανθρώπινης κοινωνίας στον ιστορικό χρόνο.

² Βλέπε: Heilbroner, Robert (2000) "Οι φιλόσοφοι του οικονομικού κόσμου". Αθήνα: Κριτική

επιμέρους συστήματα. Αντίθετα πρέπει να διαθέτει ένα καθολικό χαρακτήρα εύκολα «κατανοητό» από τα συμβαλλόμενα συστήματα). Για να χρησιμοποιήσουμε το γνωστό ανέκδοτο: η πιτσαρία της γειτονιάς πρέπει να μειώνει την ποσότητα του μπέικον στην πίτσα που μας παραδίδει διότι η σύνδεση του συστήματος της με αυτό του συστήματος υγείας την πληροφορεί ότι τα επίπεδα χοληστερίνης του πελάτη έχουν ανέβει τελευταία!

Ο άξονας της παρούσας εργασίας είναι τα προβλήματα που σχετίζονται με τη διαλειτουργικότητα μεταξύ των συστημάτων των βιβλιοθηκών. Η παραπάνω εισαγωγή κρίθηκε απαραίτητη για την ένταξη του προβλήματος στις σωστές κοινωνικές του διαστάσεις διότι **δεν είναι οι υπολογιστές που επικοινωνούν αλλά οι άνθρωποι μέσω των υπολογιστών**. Το θέμα της σημασιολογικής διαλειτουργικότητας εξετάζεται ακροθιγώς στην εργασία αλλά στην ουσία αποτέλεσε το έναυσμα για την παραπάνω εισαγωγή και μέσα από αυτό το πρίσμα νοηματοδοτείται.

Τι είναι διαλειτουργικότητα

Η διαλειτουργικότητα είναι ένας όρος ο οποίος σχετίζεται καταρχάς με την επικοινωνία υπολογιστικών συστημάτων. Τυπικά η διαλειτουργικότητα ορίζεται ως η δυνατότητα δύο ή περισσότερων συστημάτων (ή συστατικών στοιχείων συστημάτων) να ανταλλάσσουν πληροφορίες και να χρησιμοποιούν τις πληροφορίες που αντάλλαξαν. Συνήθως αυτή η δυνατότητα ερμηνεύεται στο πλαίσιο συστημάτων διαφορετικής κατασκευής (φτιαγμένα από διαφορετικούς παραγωγούς). *Το σημείο κλειδί σε αυτή την προσέγγιση είναι η χρήση των πληροφοριών που ανταλλάχθηκαν και μάλιστα χωρίς κάποια ιδιαίτερη προσπάθεια από το κάθε σύστημα*³. Επομένως το επιθυμητό αποτέλεσμα δεν είναι απλά και μόνο η ανταλλαγή αλλά η “κατανόηση” των μηνυμάτων που ανταλλάσσονται. Όπως θα φανεί στη συνέχεια ο βαθμός επιτυχίας αυτής της “κατανόησης” μπορεί να οριστεί σε διάφορα επίπεδα.

Από τις πρώτες προσεγγίσεις για να επιτευχθεί διαλειτουργικότητα ήταν το ότι όλοι οι συμβαλλόμενοι πρέπει να χρησιμοποιούν τα ίδια πρότυπα. Μάλιστα από κάποιους ορίζεται ως η απόλυτη διαλειτουργικότητα.

By using the same standard, a high level of consistency can be maintained. This, of course, has been the approach in the library community for over a century. It is the ultimate solution to the interoperability problem.

(Chan, 2005).

Από την άλλη όμως μια τέτοια προσέγγιση επιβάλλει έναν απόλυτο κομορμισμό και όπως είναι αυτονόητο μειώνονται τα επίπεδα εκφραστικότητας. Ωριμάζοντας οι συνθήκες η διαλειτουργικότητα αλλάζει πρόσωπο και από εκεί που έχει ως αφετηρία τον κομορμισμό συντάσσεται περισσότερο με την ελεγχόμενη ετερογένεια. Η σύγχρονες θεωρήσεις της διαλειτουργικότητας μάλλον ευνοούν και επικροτούν την ετερογένεια. Υπάρχει εντούτοις μια ειδοποιός διαφορά σε σχέση με τις παλαιότερες

³ Πληθώρα ορισμών για τη διαλειτουργικότητα στο “Metadata Interoperability and standardization” (Chan and Zeng, 2005)

προσεγγίσεις: η ετερογένεια είναι ελεγχόμενη και επιζητείται μόνο σε εκείνες τις περιπτώσεις που αυξάνει τα νοήματα, την εκφραστικότητα και την εξειδίκευση.

Δύο είναι οι βασικοί λόγοι για τους οποίους ένα σχήμα δεν είναι η βέλτιστη λύση για την περιγραφή ή την αναπαράσταση αντικειμένων. Ο πρώτος έχει να κάνει με το ότι όλα τα προς περιγραφή αντικείμενα δεν είναι ίδια. Άλλα στοιχεία απαιτούνται για την περιγραφή ενός βιβλίου, άλλα για την περιγραφή ενός ανθρώπου, άλλα για την περιγραφή ενός καλοριφέρ κ.ο.κ. Ο δεύτερος λόγος είναι πως για τα ίδια αντικείμενα οι περιγραφές μπορεί να διαφέρουν διότι απευθύνονται σε διαφορετικές κοινότητες ανθρώπων. Για παράδειγμα δεν συντρέχει λόγος να περιγραφεί ένα βιβλίο με διαφορετικό τρόπο από τη βιβλιοθήκη Α και τη βιβλιοθήκη Β εάν η χρήση της περιγραφής έχει τους ίδιους σκοπούς: δηλαδή την ένταξη μιας “παραδοσιακής” βιβλιογραφικής εγγραφής σε έναν επίσης “παραδοσιακό” OPAC. Όμως στην περίπτωση που η περιγραφή του ίδιου βιβλίου γίνεται στο πλαίσιο ενός εκπαιδευτικού προγράμματος όπου αυτά που πρωτίστως ενδιαφέρουν είναι το σε ποιο επίπεδο σπουδών αναφέρεται, αν είναι βιβλίο δασκάλου ή μαθητή, πόσος χρόνος απαιτείται για τη διδασκαλία κτλ τότε μάλλον είναι αναγκαίο ένα άλλο πρότυπο περιγραφής για το βιβλίο.

Εάν επιβάλλονταν συγκεκριμένα μορφότυπα τότε θα υπήρχαν πολλοί περιορισμοί και θα έπρεπε να γίνονται πολλοί συμβιβασμοί -και ενδεχομένως απλουστεύσεις- ώστε να επιτευχθεί ομοιογένεια. Το αποτέλεσμα μειωμένη εκφραστικότητα και δυσκολίες στην εξειδίκευση.

Συμπεραίνεται ότι η διαλειτουργικότητα αποτελεί το εφελτήριο της λύσης διότι **επιτρέπει την ετερογένεια**, η οποία με τη σειρά της επιτρέπει υψηλούς βαθμούς εξειδίκευσης, **και ταυτόχρονα καθιστά δυνατή την επικοινωνία** μέσα από τη δημιουργία βαθμών εξειδίκευσης. Με τον τρόπο αυτό ένα σύστημα βιβλιοθήκης μπορεί να στέλνει σε ένα άλλο σύστημα (το οποίο δεν είναι κατ’ ανάγκη σύστημα βιβλιοθήκης) δεδομένα που το δεύτερο μπορεί να “καταλάβει” ανεξάρτητα από την κωδικοποίηση που έχουν στο πρώτο. Για παράδειγμα το σύστημα μιας βιβλιοθήκης μπορεί να στέλνει σε απλό κείμενο τα δεδομένα του καταλόγου σε ένα σύστημα που καταλαβαίνει μόνο το κείμενο, σε απλό Dublin Core σε ένα άλλο και σε MARC σε ένα τρίτο. Έτσι δημιουργούνται διάφορα επίπεδα επικοινωνίας και τα δεδομένα που στέλνει το πρώτο σύστημα δεν είναι κατ’ ανάγκη κωδικοποιημένα σε κανένα από τα μορφότυπα που τα στέλνει αλλά σε κάποιο άλλο μορφότυπο το οποίο μεταβάλλεται δυναμικά ανάλογα με τις απαιτήσεις αυτού που ζητάει.

Οι Ouksel and Sheth (αναφέρεται στο DELOS 2005, σ. 7) ορίζουν τέσσερις τύπους (επίπεδα) ετερογένειας και ως εκ τούτου τέσσερα σημεία στα οποία χωρίζεται αντίστοιχα η διαλειτουργικότητα:

- **Συστήματος:** ασυμβατότητες σε επίπεδο λειτουργικών συστημάτων
- **Σύνταξης:** διαφορές στην κωδικοποίηση και την αναπαράσταση
- **Δομής:** ποικιλία στα μοντέλα δεδομένων, στις δομές δεδομένων και τα σχήματα
- **Σημασιολογική:** ασυνέπειες στην ορολογία και το νόημα

Θα πρέπει να διευκρινιστεί ότι η ύπαρξη του ίδιου λογισμικού δεν είναι ούτε αναγκαία αλλά μερικές φορές ούτε και ικανή συνθήκη για την επίτευξη της διαλειτουργικότητας. Ένα παράδειγμα που καθιστά σαφή την αλήθεια της προηγούμενης πρότασης θα μπορούσε να αντληθεί από δύο ελληνικές ακαδημαϊκές βιβλιοθήκες. Τόσο η βιβλιοθήκη του Καποδιστριακού Πανεπιστημίου Αθηνών όσο και εκείνη του Αριστοτελείου Πανεπιστημίου Θεσσαλονίκης διαθέτουν το ίδιο πρόγραμμα αυτοματισμού βιβλιοθήκης. Η διαλειτουργικότητα μεταξύ των δύο βιβλιοθηκών, ωστόσο, βρίσκεται σε πολύ χαμηλά επίπεδα διότι δεν συμφωνούν ούτε σε επίπεδο διάταξης εγγραφής (η μία ακολουθεί το MARC 21 ενώ η άλλη το UNIMARC) ούτε σε επίπεδο κωδικοποίησης χαρακτήρων⁴. Αν προστεθούν και ζητήματα σημασιολογίας τότε τα πράγματα περιπλέκονται ακόμη περισσότερο.

Web Services: η διαλειτουργικότητα σε επίπεδο συστημάτων

Με δεδομένο πως η διαλειτουργικότητα είναι τεχνικός όρος είναι λογικό να απασχολούν την παρούσα εργασία οι σύγχρονες τεχνολογίες. Μια από τις πολλά υποσχόμενες νέες τεχνολογίες για τη διαλειτουργικότητα είναι τα web services⁵.

Πρόκειται για έναν τρόπο να επικοινωνούν μεταξύ τους οι εφαρμογές στο περιβάλλον του διαδικτύου. Αποτελούν πρότυπο του W3C. Επιτυγχάνουν υψηλό βαθμό διαλειτουργικότητας αφού είναι ανεξάρτητα υπολογιστικής πλατφόρμας κυρίως χάρη στην XML και τη χρήση των πρωτοκόλλων HTTP και SOAP. Αποτελούν αντιστοιχία των API⁶ στο περιβάλλον του παγκόσμιου ιστού. Πρόκειται για application components τα οποία μπορούν να χρησιμοποιηθούν και από άλλες εφαρμογές. Κάθε application component δημοσιεύει τις λειτουργίες του και τον τρόπο που πρέπει να χρησιμοποιούνται και έτσι είναι σε θέση να διαλειτουργήσει με άλλα application components και συνήθως να επιτελέσουν όλα μαζί μια πιο σύνθετη εργασία. Η υποδομή των web services στηρίζεται κυρίως σε τρία ανοικτά πρότυπα:

- SOAP (Simple Object Access Protocol): αποτελεί πρωτόκολλο επικοινωνίας μεταξύ εφαρμογών στο διαδίκτυο περιγράφοντας τη μορφή των μηνυμάτων μεταξύ των εφαρμογών με τη χρήση της γλώσσας xml.
- UDDI (Universal Description, Discovery and Integration): Αποτελεί έναν κατάλογο καταγραφής υπηρεσιών web services.
- WSDL (Web Services Description Language): είναι μια xml γλώσσα που χρησιμοποιείται για να περιγράψει τα web services και πως σε αυτά θα έχει πρόσβαση μια εφαρμογή.

Συνήθως τα web services βρίσκονται σε κάποιον SOAP Server και η καταχώρησή τους γίνεται σε ένα UDDI κατάλογο. Η εύρεσή τους λοιπόν γίνεται με αναζήτηση στους UDDI καταλόγους και στη συνέχεια καλούνται από τον αντίστοιχο SOAP

⁴ Για να μπορέσει η βιβλιοθήκη του Πανεπιστημίου Αθηνών να παίρνει έτοιμες εγγραφές από εκείνη του ΑΠΘ είχε προμηθευτεί λογισμικό από το ΕΚΤ το οποίο μετέτρεπε τόσο τη διάταξη όσο και το σύνολο χαρακτήρων.

⁵ <http://www.w3.org/2002/ws>

⁶ Application programming interface: <http://en.wikipedia.org/wiki/Api>

Server. Υπάρχουν SOAP servers με web services ελεύθερα προς χρήση ή επί πληρωμή. Το μοντέλο αυτό κάνει την κατασκευή web εφαρμογών πιο εύκολη και πιο οικονομική αφού αν υπάρχουν τα web services για την εκτέλεση κάποιας εργασίας τότε αυτό που απαιτείται είναι η χρήση τους και η κατασκευή μόνο του τρόπου παρουσίασής στον παγκόσμιο ιστό. Για παράδειγμα έστω ότι πρέπει να κατασκευαστεί ένα site για να παρέχει κάποιες πληροφορίες στους χρήστες όπως καιρός, πληροφορίες κίνησης, πληροφορίες δημόσιων υπηρεσιών κλπ. Τότε αν υπάρχουν έτοιμα αντίστοιχα web services το μόνο που θα πρέπει να κατασκευαστεί θα είναι η προσαρμογή των πληροφοριών που θα επιστρέφουν οι αντίστοιχοι SOAP servers και η εμφάνισή τους στο υπό κατασκευή site.

Τα πλεονεκτήματα από τη χρήση τους είναι προφανή. Είναι διαλειτουργικά αφού επιτρέπουν την επικοινωνία μεταξύ εφαρμογών στο διαδίκτυο που μπορεί να βρίσκονται σε ετερογενή υπολογιστικά συστήματα και κάτω από διαφορετικά λειτουργικά συστήματα. Είναι διαθέσιμα αφού δημοσιεύονται σε καταλόγους. Είναι εύκολα τροποποιήσιμα αφού είναι απλή η κατασκευή τους και επειδή στηρίζονται σε πολλά ανοικτά πρότυπα. Είναι κατάλληλα για ενσωμάτωση σε άλλες εφαρμογές και έτσι κάνουν πιο εύκολη και οικονομική την κατασκευή και διάδοση εφαρμογών στο διαδίκτυο.

Οι βιβλιοθήκες μπορούν να χρησιμοποιήσουν αυτή την τεχνολογία για να επιτελέσουν πολλές λειτουργίες. Ήδη η αναζήτηση / ανάκτηση σύμφωνα με τα πρότυπα SRU/SRW (βλέπε επόμενο κεφάλαιο) είναι ένα καλό παράδειγμα. Έτσι αν κατασκευάζονται web services για λειτουργίες των βιβλιοθηκών θα μπορούν να αποτελέσουν κίνητρο για τη δημιουργία ολοκληρωμένων υπηρεσιών στο διαδίκτυο γρήγορα και εύκολα. Άλλο χαρακτηριστικό παράδειγμα είναι η ενσωμάτωση λειτουργιών της βιβλιοθήκης σε portals εκπαιδευτικών ιδρυμάτων ή άλλων οργανισμών χωρίς να απαιτείται η μετάβαση στο portal της βιβλιοθήκης. Στο portal λοιπόν που θα απαιτείται η χρήση πληροφοριών / υπηρεσιών της βιβλιοθήκης θα γίνεται απλά η κλήση των κατάλληλων web services της βιβλιοθήκης και έτσι θα μπορούν να εμφανίζονται εμφωλευμένα τα αποτελέσματα χωρίς την έξοδο του χρήστη από το portal. Προφανές το πλεονέκτημα της ολοκλήρωσης των πληροφοριών από διαφορετικές πηγές με πρώτο κερδισμένο τον ίδιο τον χρήστη ο οποίος από ένα σημείο έχει πρόσβαση σε πολλές διαφορετικών πηγών πληροφορίες. Προφανές όμως και το πλεονέκτημα για την βιβλιοθήκη η οποία εξασφαλίζοντας την πρόσβαση στις πληροφορίες της / υπηρεσίες διαμέσου τρίτων αναβαθμίζει τη θέση της και το ρόλο της στο νέο περιβάλλον του διαδικτύου.

Διαλειτουργικότητα στις βιβλιοθήκες

Όπως ορίστηκε παραπάνω η διαλειτουργικότητα είναι η δυνατότητα ανταλλαγής πληροφοριών μεταξύ συστημάτων και η χρήση αυτών των πληροφοριών από τα συστήματα. Ποιες όμως είναι οι πληροφορίες εκείνες που ανταλλάσσουν οι βιβλιοθήκες; Όπως έχει ήδη επισημανθεί, διαλειτουργικότητα σημαίνει επικοινωνία. Αρχικά, λοιπόν, θα πρέπει να οριστούν οι υπηρεσίες εκείνες που συνήθως επικοινωνούν μεταξύ των βιβλιοθηκών.

Στις παραδοσιακές βιβλιοθήκες πληθώρα προτύπων αναπτύχθηκαν γύρω από τη συγκρότηση καταλόγων και αυτό είναι αναμενόμενο διότι πάνω στη συγκρότηση καταλόγων βασίστηκε η αιχμή του δόρατος για τις βιβλιοθήκες σε όλη τη διάρκεια του 20^{ου} αιώνα. Άρα μια από τις βασικές υπηρεσίες που παρείχε μια βιβλιοθήκη ήταν ο ενιαίος τρόπος συγκρότησης ενός καταλόγου. Συγκεκριμένοι τρόποι περιγραφής των τεκμηρίων που βρίσκονταν στη συλλογή της. Το δεύτερο σημείο είναι οι τρόποι πρόσβασης στον κατάλογο. Όσο υπάρχουν οι έντυποι δελτιοκατάλογοι το πρόβλημα έχει να κάνει μόνο με την επιλογή των κλειδιών. Υπάρχει θεματικός κατάλογος, κατά εκδότη κτλ;

Με την έλευση της υπολογιστικής τεχνολογίας εντός των τειχών της βιβλιοθήκης τόσο η περιγραφή των τεκμηρίων όσο και η αναζήτηση των περιγραφών αυτών αρχίζουν να μεταλλάσσονται. Είναι ευρέως γνωστό ότι οι πρώτοι υπολογιστές χρησιμοποιήθηκαν από τις βιβλιοθήκες ως έξυπνες γραφομηχανές που μπορούσαν να τυπώνουν μια πληθώρα δελτίων. Το MARC (ειδικά το MARC 21) κατατρύχεται ακόμη (σαράντα χρόνια μετά τη γέννησή του) από τις “αμαρτίες” εκτύπωσης των δελτίων όπως η έννοια της μη επαναλαμβανόμενης μίας και υποχρεωτικής Κύριας Αναγραφής. Οι υπολογιστές, λοιπόν, τουλάχιστον στην πρώτη φάση της εισόδου τους, δεν αλλάζουν τα θεμέλια του οικοδομήματος εντούτοις προκαλούν μετασχηματισμούς. Η ουσία, ωστόσο, παραμένει η ίδια: «ελάτε να σας δείξουμε τα βιβλία που έχουμε».

Τα πράγματα αλλάζουν άρδην με την έλευση του διαδικτύου. Πλέον οι βιβλιοθήκες μπορούν να επικοινωνούν μεταξύ τους χωρίς γεωγραφικούς περιορισμούς. Το πρώτο που θα γνωστοποιήσουν και θα διαφημίσουν είναι το υλικό τους. Δηλαδή ο κατάλογος. Εδώ είναι που τίθενται και τα πρώτα θέματα διαλειτουργικότητας με τη σημερινή έννοια του όρου. Πώς οι επιμέρους περιγραφές θα συνυπάρξουν μαζί και δεύτερον πώς θα αναζητηθούν μαζί. Χρειαζόμαστε πρότυπα. Και οι βιβλιοθήκες δεν είναι νέες στη δημιουργία προτύπων. Οι βιβλιοθήκες στις ΗΠΑ αναπτύσσουν πρότυπα ήδη από τα τέλη του 19^{ου} αιώνα. Ένα όμως πρόβλημα είναι ότι πλέον τα πρότυπα των βιβλιοθηκών δεν αφορούν μόνο τις βιβλιοθήκες αλλά βασίζονται πάνω σε άλλα πρότυπα πολύ πιο τεχνικά (π.χ. TCIP, HTTP). Αυτό σε αρκετές περιπτώσεις δημιουργεί προβλήματα στα βιβλιοθηκονομικά πρότυπα αφού ένα μέρος από την ανάπτυξη και εξέλιξη τους νομοτελειακά ανατίθεται εκτός της βιβλιοθηκονομικής κοινότητας. Όπως εύστοχα παρατηρεί η Coyle (2005) πλέον οι βιβλιοθήκες ασχολούνται και με πολλά μη βιβλιοθηκονομικά πρότυπα.

Με πρώτο δεδομένο πως η διαλειτουργικότητα ισούται με την αποδέσμευση των δεδομένων από τα προγράμματα διαχείρισής τους (δε χρειάζομαι κατ' ανάγκη το

Word για να διαβάσω την εργασία μου αλλά μπορώ και με έναν άλλο επεξεργαστή κειμένου, δεν είναι απαραίτητο το X πρόγραμμα αυτοματισμού για να αποκωδικοποιήσει τα στοιχεία των βιβλιογραφικών εγγραφών) και με δεύτερο δεδομένο πως η ανάπτυξη του καταλόγου είναι από τις πιο δαπανηρές εργασίες⁷ στο πλαίσιο λειτουργίας μιας βιβλιοθήκης είναι, επομένως, αναμενόμενη η βαρύτητα που έχει δοθεί μέχρι σήμερα στη διαλειτουργικότητα των καταλόγων.

Έτσι ενώ πολύ συχνά ένα μέλος μιας βιβλιοθήκης εγγράφεται ως μέλος και σε μια άλλη βιβλιοθήκη δεν θεωρείται αναγκαίο να μπορεί το σύστημα της δεύτερης βιβλιοθήκης να μπορεί να πάρει στοιχεία για το συγκεκριμένο μέλος από την πρώτη βιβλιοθήκη. Η τα στατιστικά των δανεισμών της X βιβλιοθήκης δεν είναι ανάγκη να μπορούν να γίνουν “αντιληπτά” από το σύστημα αυτοματισμού της Ψ βιβλιοθήκης. Η εξωστρεφής πλευρά της βιβλιοθήκης παρέμενε ο κατάλογος. Οι οποίες άλλες υπηρεσίες κρίνονταν εσωτερική υπόθεση και σπάνια εμπλέκονταν σε διαδικασίες διαλειτουργικότητας.

Εξαίρεση σε αυτό υπήρξε η διαδικασία του διαδανεισμού. Πρότυπα για το διαδανεισμό (Interlibrary Loan, z39.50 item order) έχουν αναπτυχθεί και υπάρχει μια σχετική διαλειτουργικότητα στο συγκεκριμένο θέμα. Δεν διαθέτουν βέβαια το μερίδιο αγοράς των καταλόγων και είναι απολύτως λογικό αφού για παράδειγμα το z39.50 item order πάνω στον κατάλογο βασίζεται. Άρα ακόμη και αυτό τον προϋποθέτει.

Μέσα σε αυτό το περιβάλλον τα ολοκληρωμένα συστήματα αυτοματισμού βιβλιοθηκών υπόσχονταν όλες τις υπηρεσίες διαλειτουργικότητας που παρείχαν τα πρότυπα και ακόμη περισσότερες. Και στο σημείο αυτό τίθεται ένα πολύ σοβαρό θέμα. Η βιβλιοθήκη πρέπει να ακολουθήσει κατά γράμμα τα πρότυπα ώστε να επιτύχει την όποια διαλειτουργικότητα ή πρέπει να δώσει βαρύτητα στις ανάγκες των χρηστών της;

Όμως σε ένα κόσμο ανοικτό το κριτήριο δεν είναι πόσα να κάνει η βιβλιοθήκη αλλά σε πόσα αλληλεπιδρά με τους γύρω της, με εκείνους που έχουν τους ίδιους στόχους με αυτήν. Στην εποχή μας υπάρχει μια στροφή από το να κάνω το διαφορετικό (το ξεχωριστό) στο να κάνω καταρχάς αυτό που όλοι (δυναμικά) θα μπορούν να καταλάβουν. Αυτό άλλωστε είναι και η βάση της διαλειτουργικότητας.

Τα πρότυπα είναι δεσμευτικά και απαιτούν συγκεκριμένες υλοποιήσεις. Σε καμία όμως περίπτωση δεν είναι απαγορευτικά στο να δίνει κανείς την πληροφορία και με άλλους τρόπους. Για παράδειγμα κανείς δεν απαγορεύει σε ένα σύστημα αυτοματισμού να φτιάξει ένα περιβάλλον αναζήτησης με πιο σύνθετα και αναλυτικά semantics από αυτά που δίνει το z39-50 (bib-1) αρκεί να μπορεί να παρέχει τα δεδομένα της και με z39-50. Μια βιβλιοθήκη δεν υποχρεούται να έχει εσωτερική διάταξη το MARC αρκεί να μπορεί να δίνει και σε MARC και η διαλειτουργικότητα

⁷ Για τη Βιβλιοθήκη του Κογκρέσου μία βιβλιογραφική εγγραφή κοστίζει κατά μέσο όρο πάνω από 120 δολάρια (βλέπε <http://www.loc.gov/faq/catfaq.html#9>) ενώ άλλοι ερευνητές (βλέπε Lugg, Rick (2007), *Working Group on the Future of Bibliographic Control: Economics and Organization of Bibliographic Data*, Library of Congress, Washington, διαθέσιμο στο: <http://www.loc.gov/bibliographic-future/meetings/docs/ricklugg-july9-2007.pdf>) εκτιμούν ότι το κόστος για την πλήρη καταλογογράφηση μιας αγγλικής μονογραφίας -με όλη την απαιτούμενη εργασία στα Authorities- είναι από 150-200 δολάρια η εγγραφή.

έχει επιτευχθεί. Πλέον όμως τα πρότυπα είναι αρκετά σύνθετα και απαιτητικά οπότε καλύπτουν ένα μεγάλο εύρος περιπτώσεων και δεν είναι εύκολο να διατεθούν τόσο πόροι για παράλληλη ανάπτυξη και εκτός των προτύπων. Η αυτονομία καθίσταται μύθος και η δυνατότητα της υπέρβασης των προτύπων ευσεβής αλλά ανέφικτος στόχος.

Έτσι παρότι τα πρότυπα μπορούν θεωρητικά να αποτελέσουν τροχοπέδη για την ανάπτυξη (ως ένα απαρχαιωμένο νομικό σύστημα που δεν ανταποκρίνεται στην κοινωνική πραγματικότητα) στην παρούσα φάση αποτελούν το ισχυρότερο όχημα της.

Δημιουργία και ανταλλαγή μεταδεδομένων

Από τα παραπάνω είναι φανερό η σημασία που αποδίδεται στα μεταδεδομένα. Στην ενότητα αυτή θα επιχειρηθεί μια συστηματική προσέγγιση της διαλειτουργικότητας στο επίπεδο δημιουργίας των μεταδεδομένων και παρότι σημαντικός σκοπός αυτής της δημιουργίας είναι η αναζήτηση, αυτή θα εξεταστεί στο επόμενο κεφάλαιο αφού διαθέτει επαρκή στοιχεία για να αντιμετωπιστεί ως ξεχωριστή ενότητα.

Ο μικρότερος -και ίσως ο περιεκτικότερος- ορισμός για τα μεταδεδομένα αναφέρει ότι πρόκειται για *δεδομένα για άλλα δεδομένα* (data about data). Σίγουρα η προηγούμενη διατύπωση είναι πολύ γενική αλλά και ακόμη και έτσι οι βιβλιοθήκες αποτελούν το ναό των μεταδεδομένων καθώς μια από τις κύριες εργασίες τους είναι η δημιουργία δεδομένων για άλλα δεδομένα. Οι βιβλιογραφικοί κατάλογοι μπορεί να αποτελούν το σκληρό πυρήνα αυτών αλλά μάλλον τα μεταδεδομένα δεν περιορίζονται σε αυτούς εντός των βιβλιοθηκών. Πόσες φορές δανείστηκε ένα βιβλίο ή πόσες φορές έγινε download ένα ηλεκτρονικό βιβλίο σίγουρα είναι δεδομένα για άλλα δεδομένα.

Το πρώτο σημείο στο οποίο οι βιβλιοθήκες επεδίωξαν διαλειτουργικότητα ήταν στο εκείνο των μεταδεδομένων και μάλιστα πολλά χρόνια πριν εμφανιστεί τόσο ο όρος μεταδεδομένα όσο και ο όρος διαλειτουργικότητα! Η βιβλιοθήκη του Κογκρέσου έστειλε έτοιμα δελτία στις βιβλιοθήκες ήδη από τις αρχές του προηγούμενου (20^{ου}) αιώνα.

Τα τελευταία χρόνια ο μεγάλος αριθμός των πηγών σε συνδυασμό με την πληθώρα των ειδικοτήτων που ασχολούνται με την οργάνωσή τους (βιβλιοθηκονόμοι, πληροφορικοί, domain experts κτλ⁸) δημιουργεί ιδιαίτερα αυξημένο ενδιαφέρον γύρω από την οργάνωση των πληροφοριών και την παραγωγή μεταδεδομένων. Μέσα σε αυτό το κλίμα παλιά σχήματα μετασχηματίζονται και εξελίσσονται καθώς και νέα σχήματα κάνουν την εμφάνισή τους. Από την άλλη όμως υπάρχει η ανάγκη για μια πιο αφαιρετική προσέγγιση και για δημιουργία μιας πιο συνολικής θεώρησης της κατάστασης.

⁸ Ενδιαφέρον παρουσιάζουν προσεγγίσεις από χώρους τελείως διαφορετικούς από το χώρο της πληροφόρησης-πληροφορίας. Ένα ενδιαφέρον παράδειγμα είναι ένα μικρό σε διάρκεια βίντεο που έχει δημιουργηθεί από έναν ανθρωπολόγο. Το βίντεο διατίθεται από το YouTube με τίτλο “The Machine is Us/ing Us” (http://www.youtube.com/watch?v=NLIgopyXT_g)

Ακριβώς αυτή την προσπάθεια δημιουργίας μιας συνεκτικής θεωρίας μέσω της οποίας να προσεγγιστούν όλα τα πρότυπα μεταδεδομένων για τις βιβλιοθήκες προσπάθησε να κάνει η IFLA. Εδώ και μια δεκαετία δημιούργησε ένα εννοιολογικό μοντέλο με τις «Λειτουργικές προδιαγραφές για Βιβλιογραφικές Εγγραφές»⁹. Στο εν λόγω μοντέλο θεωρείται ότι μια βιβλιογραφική εγγραφή εξυπηρετεί τέσσερις βασικούς σκοπούς:

- Χρησιμοποιείται για τον εντοπισμό υλικού που αντιστοιχεί στα κριτήρια που θέτει ο χρήστης
- Να μπορούν να χρησιμοποιηθούν τα δεδομένα που ανακτήθηκαν για να γίνει ταυτοποίηση μιας οντότητας
- Να μπορούν να χρησιμοποιηθούν τα δεδομένα για την επιλογή μιας οντότητας η οποία είναι κατάλληλη για τον χρήστη
- Να μπορούν να χρησιμοποιηθούν τα δεδομένα με σκοπό την απόκτηση του ίδιου του αντικειμένου ή της πρόσβασης σε αυτό

Συμπερασματικά για να είναι πλήρης μια βιβλιογραφική περιγραφή θα πρέπει να μπορεί να καλύψει τις παραπάνω απαιτήσεις. Άρα, ανεξάρτητα από το σχήμα μεταδεδομένων οι ανάγκες είναι αυτές και το σχήμα οφείλει να τις λαμβάνει υπόψη του. Το μοντέλο επίσης ορίζει τις επιμέρους οντότητες που αποτελούν μια βιβλιογραφική εγγραφή. Οι υλοποιήσεις του εν λόγω μοντέλου είναι ακόμη σε πειραματικό στάδιο¹⁰ αλλά μας απασχολεί εδώ διότι δείχνει προς την κατεύθυνση του μέλλοντος.

Για να αντιμετωπιστεί το πρόβλημα της διαλειτουργικότητας σε επίπεδο μεταδεδομένων έχουν διατυπωθεί διάφορες προσεγγίσεις. Είναι προφανές πως αν όλοι χρησιμοποιούσαν το ίδιο πρότυπο μεταδεδομένων θα υπήρχε υψηλό επίπεδο διαλειτουργικότητας. Οι βιβλιοθήκες το έκαναν κατά κανόνα για περισσότερα από σαράντα χρόνια με το MARC. Παρά τις όποιες διαφορές στα διάφορα MARC, συνήθως σε εθνικό επίπεδο¹¹ ακολουθούνταν μία συγκεκριμένη υλοποίησή του (USMARC, UNIMARC κτλ).

Το ερώτημα που προκύπτει είναι εάν είναι δυνατό ή/και επιθυμητό να χρησιμοποιείται το ίδιο πρότυπο από όλους; Όπως αναφέρθηκε και στην εισαγωγή η απάντηση στην συγκεκριμένη ερώτηση είναι μάλλον αρνητική και αυτό για δύο βασικούς λόγους. Ο ένας λόγος είναι πως οι πληροφοριακοί πόροι έχουν διάφορες μορφές και ως εκ τούτου υπάρχουν διαφορετικές ανάγκες περιγραφής και ανάλυσης. Ο δεύτερος αλλά όχι λιγότερο σημαντικός είναι πως υπάρχουν διαφορετικές ανάγκες περιγραφής και ανάλυσης για τον ίδιο πόρο ανάλογα με την κοινότητα η οποία τον εξετάζει. Συνεπώς μια ομογενοποιημένη περιγραφή

⁹ Functional Requirements for Bibliographic Records (διαθέσιμο στο web <http://www.ifla.org/VII/s13/frbr/frbr.pdf>)

¹⁰ Ο OCLC συμμετέχει σε διάφορες προσπάθειες για εναρμόνιση των υφιστάμενων εγγραφών με τις αρχές του μοντέλου. Άλλωστε η βάση Worldcat παρέχει τις εγγραφές όχι σε επίπεδο **Manifestation** αλλά περισσότερο σε επίπεδο **Expression**.

¹¹ Στην Ελλάδα, δυστυχώς, υπάρχει σχετική πολυφωνία αφού στις ελληνικές βιβλιοθήκες χρησιμοποιείται τόσο το UNIMARC όσο και το MARC 21.

αποτελεί μονοδιάστατη ανάλυση και αποκόπτει πλευρές της πληροφορίας οι οποίες αποτελούν “σημαίνον” σε μια κοινότητα αλλά ίσως θορυβώδη πλεονασμό σε μια άλλη. Το να γίνει αναφορά στη γραμματοσειρά που χρησιμοποιείται σε ένα βιβλίο θεωρείται άχρηστο όταν δημιουργείται μια βιβλιογραφική εγγραφή σε μια βιβλιοθήκη, αλλά ακριβώς αυτό το στοιχείο μπορεί να είναι το σημείο κλειδί στην περίπτωση μιας βιβλιοθήκης που εξυπηρετεί τις ανάγκες ενός ιστορικού της τυπογραφίας.

Είναι, λοιπόν, προφανές πως δεν διευκολύνει η χρήση ενός και μόνο προτύπου μεταδεδομένων. Η αφθονία προτύπων επιτρέπει μεγάλη εκφραστικότητα εντός του κάθε προτύπου. Ο Rogers (2005) αναφέρει πως από έρευνα που έγινε σε επτά διαδεδομένα σχήματα μεταδεδομένων εντοπίστηκε ένα σύνολο από 293 metatags. Μόνο όμως 5 από αυτά περιλαμβάνονταν και στα 7 σχήματα!

Πως είναι δυνατό να επιτευχθεί διαλειτουργικότητα μέσα σε ένα τέτοιο περιβάλλον; Ο Chan (2005) εξετάζει το πρόβλημα της διαλειτουργικότητας των μεταδεδομένων από μία γενική οπτική και θεωρεί ότι τα τελευταία χρόνια επιχειρήθηκε να λυθεί με την εφαρμογή επτά μοντέλων, χωρίς κατ’ ανάγκη η χρήση του ενός μοντέλου να αποκλείει τη χρήση των άλλων. Τα μοντέλα που διαχωρίζει είναι τα εξής:

- Uniform standard = Ενιαίο πρότυπο
Ένα και μόνο πρότυπο μεταδεδομένων
- Application profiling/adaptation/modification = Προφίλ εφαρμογών / υιοθετήσεις / τροποποιήσεις
Στοιχεία από διάφορα πρότυπα
- Derivation = Παράγωγο κάποιου άλλου
Συνήθως αναφέρεται στη διαδικασία παραγωγής ενός προτύπου από κάποιο άλλο το οποίο παράγωγο είναι και κατά κανόνα απλούστερο του αρχικού. Για παράδειγμα το MODS είναι παράγωγο του MARC
- Crosswalk/mapping = Αντιστοίχιση
Η αντιστοίχιση των στοιχείων, της σημασιολογίας και της σύνταξης από ένα σχήμα μεταδεδομένων σε ένα άλλο
- Switching schema = Σχήμα Αλλαγής
Χρήση ενός γενικού σχήματος στο οποίο μετατρέπονται τα επιμέρους. Συχνά γίνεται με χρήση του DC.
- Lingua franca
Στις περιπτώσεις που κανένα σχήμα δεν είναι κατάλληλο για «Σχήμα Αλλαγής» τότε γίνεται χρήση μιας lingua franca ως υπερδομής. Στην ουσία πρόκειται για γενικές κατευθύνσεις με βάση κοινά στοιχεία που συναντώνται σε όλους τους πόρους όπως: θέμα, ημερομηνία, όροι χρήσης, εκδότης, ονόματα που σχετίζονται με τον πόρο, γλώσσα, τύπος, δημιουργός και έκδοση.
- Metadata framework/container = Πλαίσιο μεταδεδομένων
Αυτή η προσέγγιση έχει να κάνει με τη δημιουργία ενός γενικού κελύφους μέσα στο οποίο μπορούν να συνυπάρχουν πολλαπλά σχήματα μεταδεδομένων. Ένα καλό παράδειγμα αυτής της προσέγγισης είναι το METS.

Με γνώμονα την παραπάνω κατηγοριοποίηση θα επιχειρηθεί μια αποτίμηση της κάθε μίας από τις κατηγορίες.

Η απλή περίπτωση του ενός και μόνο προτύπου είναι ανέφικτη σε μεγάλες κλίμακες όπου εμπλέκονται πολλοί και σχετικά ετερογενείς φορείς. Άλλωστε έχει ήδη γίνει στο πλαίσιο της παρούσας εργασίας εκτενής αναφορά στις αδυναμίες του.

Μια πολύ διαδεδομένη πρακτική των ημερών μας είναι τα application profiles. Στην περίπτωση αυτή αντί να χρησιμοποιείται μόνο ένα σχήμα μεταδεδομένων καλούνται στοιχεία από διάφορα σχήματα ανάλογα με τις ανάγκες. Για παράδειγμα μπορεί να περιγράφεται το Αρχείο του X σε EAD και όταν φτάσει η ώρα της περιγραφής ενός συγκεκριμένου τεκμηρίου του Αρχείου το οποίο είναι βιβλίο να χρησιμοποιούνται στοιχεία από το MODS ενώ στη συνέχεια να περιγράφεται μια φωτογραφία σε VRA κτλ. Για την υλοποίηση αυτών εκμεταλλεύονται η δυνατότητες της XML σχετικά με τους χώρους ονομάτων (namespaces). Αυτό το οποίο παράγεται τελικά είναι ένα EAD αρχείο το οποίο όμως είναι εμπλουτισμένο με στοιχεία και από άλλα πρότυπα (εν προκειμένω με το MODS και το VRA). Αυτή η προσέγγιση είναι αναμφισβήτητα προτιμότερη από την προσθήκη νέων στοιχείων που δημιουργούνται τοπικά και δεν έχουν καθολική ισχύ. Από την άλλη όμως μειώνουν τη διαλειτουργικότητα διότι δεν είναι προεξοφλημένο το πώς θα πρέπει να συμπεριφερθεί σε αυτά το σύστημα. Όπως για παράδειγμα πως θα συμπεριφερθεί stylesheet που θα δημιουργήσει την εμφάνιση των δεδομένων ή με ποιον τρόπο θα γίνει μια ενδεχόμενη ευρετηρίαση.

Η περίπτωση του παραγώγου σχήματος είναι λίγο προβληματική στον ορισμό της. Παρθενογένεση δεν υπάρχει. Όλα προέρχονται από κάπου αλλού και αυτό που αλλάζει είναι ο βαθμός μετάλλαξης από τους “προγόνους”. Άρα κάτι τέτοιο κρίνεται κατά περίπτωση με βάση τα χαρακτηριστικά που κληρονομούνται από τον πρόγονο στον απόγονο.

Η κατηγορία των crosswalks είναι αρκετά σύνθετη και παράλληλα ένας από τους πιο διαδεδομένους τρόπους για να επιτευχθεί διαλειτουργικότητα. Ακόμη και με μια σύντομη ματιά στη βιβλιογραφία φαίνεται ότι ένας μεγάλος όγκος της αναφέρεται στα crosswalks¹². Οι Zeng and Xiao (αναφέρεται στο Chan, 2006) κατηγοριοποιούν τις αντιστοιχίσεις στις εξής περιπτώσεις:

- ένα προς ένα
- ένα προς πολλά
- πολλά προς ένα
- ένα προς κανένα

Η περίπτωση του ένα προς ένα είναι η απλούστερη και κάνει εύκολη της διαδικασία της αντιστοίχισης του ενός σχήματος στο άλλο. Ένα στοιχείο του ενός σχήματος αντιστοιχεί σε ένα του άλλου. Στην περίπτωση όπου ένα στοιχείο του σχήματος ενδέχεται να αντιστοιχεί σε πολλά του άλλου πρέπει να ληφθούν υπόψη συγκεκριμένες παράμετροι που σχετίζονται με τα εν λόγω δεδομένα και δεν ισχύουν καθολικοί νόμοι. Οι Chan και Zeng (2006) ισχυρίζονται ότι τα

¹² Εκτός από τη βιβλιογραφία συχνά οι αντιστοιχίσεις είναι αποτέλεσμα διαφόρων έργων που σχετίζονται με τη διαλειτουργικότητα. Χαρακτηριστικό παράδειγμα η σελίδα της UKOLN <http://www.ukoln.ac.uk/metadata/interoperability/>

crosswalks επειδή βασίζονται σε γενικά μοντέλα έχουν ως αποτέλεσμα σοβαρές μειώσεις στην ποιότητα των μεταδεδομένων όταν δεν λαμβάνονται σοβαρά υπόψη οι επιμέρους συνθήκες παραγωγής των εκάστοτε δεδομένων. Η εμπειρία των υπογραφόντων την παρούσα εργασία από μετατροπές μεταδεδομένων από διάφορα μορφότυπα στο πλαίσιο της εργασίας τους στο EKT όχι μόνο συνηγορεί αλλά και επαυξάνει. Για παράδειγμα αν μετατρέπονται δεδομένα από DC σε UNIMARC και το στοιχείο του τίτλου στο DC επαναλαμβάνεται για να δηλώσει όλους τους προηγούμενους τίτλους μιας περιοδικής έκδοσης (χωρίς εξειδικευτές) αυτό δημιουργεί πρόβλημα σε μια πιθανή μετατροπή σε UNIMARC αφού δεν είναι δυνατό να είναι γνωστό ποιος τίτλος θα πάει στο 200 και ποιοι στο 5XX. Αν εξεταστούν οι συνθήκες παραγωγής των συγκεκριμένων δεδομένων (δηλαδή οι πολιτικές καταχώρησης) το αποτέλεσμα θα είναι σίγουρα ακριβέστερο. Στην περίπτωση του πολλά προς ένα το μόνο πρόβλημα είναι πως χάνεται μέρος της εκφραστικότητας του αρχικού σχήματος αφού το δεύτερο απλοποιεί τα δεδομένα (π.χ. το στοιχείο θέμα του DC δεν μπορεί να καθορίσει αν το θέμα είναι πρόσωπο ή γεωγραφικό όνομα κτλ. Άρα σε μια πιθανή μετατροπή από UNIMARC σε DC η αντίστοιχη εξειδίκευση θα χαθεί και θα ομαδοποιηθούν όλα κάτω από το θέμα). Η περίπτωση ένα προς τίποτα είναι η πιο επώδυνη αφού χάνονται εντελώς δεδομένα μιας και δεν υπάρχει καμία αντιστοίχιση.

Η περίπτωση του Σχήματος Αλλαγής (Switching schema) μοιάζει με αυτή των crosswalks. Αλλά αντί να γίνονται αντιστοιχίσεις από κάθε σχήμα μεταδεδομένων σε ένα άλλο να γίνονται σε ένα κεντρικό: το Σχήμα Αλλαγής. Στην περίπτωση αυτή υπάρχει το πλεονέκτημα πως απαιτούνται τόσες αντιστοιχίσεις όσα και τα σχήματα. Όλα αντιστοιχίζονται από μία φορά στο Σχήμα Αλλαγής. Αν το σχήμα αυτό είναι ένα πραγματικό σχήμα (συνήθως είναι το DC) χάνεται σημαντική πληροφορία η οποία δεν θα χανόταν αν υπήρχαν crosswalks. Για παράδειγμα αν μετατρέπεται μια εγγραφή από UNIMARC σε MARC 21 απευθείας δεν χάνεται η πληροφορία για το αν ο συγγραφέας είναι φυσικό πρόσωπο ή συλλογικό όργανο. Αν όμως μεσολαβήσει το DC ως Σχήμα Αλλαγής τότε η πληροφορία αυτή θα χαθεί. Με την ίδια ακριβώς λογική που θα γινόταν και στη φυσική γλώσσα. Αν μεταφραστεί από τη γλώσσα των Εσκιμώων προς τα Σουηδικά κάποια από τις τριάντα λέξεις που διαθέτουν οι πρώτοι για το χιόνι είναι πιθανό πως υπάρξει καλύτερη αντιστοιχία σε μία από τις δέκα που διαθέτουν οι Σουηδοί από ότι αν μεσολαβήσει η ελληνική γλώσσα όπου και οι τριάντα διαφορετικές λέξεις που περιγράφουν διαφορετικής μορφής χιόνι θα καταλήξουν στη μία και μοναδική της ελληνικής γλώσσας.

Η Lingua franca αφού στην ουσία δεν αποτελεί ένα πραγματικό σχήμα αλλά συμβάσεις πάνω σε κοινά αποδεκτές αρχές μεταξύ των εμπλεκομένων μάλλον δεν ενδείκνυται για ουσιαστική διαλειτουργικότητα.

Ιδιαίτερο ενδιαφέρον παρουσιάζει η κατηγορία του Metadata framework/container. Εδώ πλέον μιλάμε για την πρόθεση να επιτευχθεί ο ύψιστος βαθμός διαλειτουργικότητας, όπου πλέον δεν στήνονται μόνο μεταδεδομένα ή δεδομένα αλλά συλλογές ολόκληρες και οι σχέσεις μεταξύ των αντικειμένων τους. Από την άλλη ακριβώς ο υψηλός βαθμός πολυπλοκότητας σε συνδυασμό με την μεγάλη ελευθερία εκφραστικότητας καθιστούν το όλο εγχείρημα να εξαρτάται σε μεγάλο βαθμό από τις οδηγίες της εκάστοτε υλοποίησης. Ίσως δεν είναι ακόμη αρκετά ώριμες οι συνθήκες για τέτοιες προσπάθειες σε μεγάλη κλίμακα.

Αναζήτηση μεταδεδομένων

Τα όσα αναφέρθηκαν στο προηγούμενο κεφάλαιο ίσως να έχουν ενδιαφέρον ίσως όχι. Το μόνο σίγουρο είναι πως αν έχουν ενδιαφέρον έχουν μόνο για τους “ειδικούς” του χώρου. Ο τελικός χρήστης δεν θέλει να γνωρίζει όλα τα παραπάνω. “[Users] should not have to know or understand the methods used to describe and represent the contents of the digital collection” (Chan, 2006). Εντούτοις όλα τα παραπάνω έγιναν για να γίνει η ζωή του τελικού χρήστη ευκολότερη και η αναζήτηση του αποτελεσματικότερη.

Το πρόβλημα που αναδύεται είναι πως ο τελικός χρήστης θα έχει πρόσβαση όχι πλέον στον κατάλογο της κάθε βιβλιοθήκης αλλά σε πολλές βιβλιοθήκες ταυτόχρονα. Σε μία πρώτη φάση το πρόβλημα λύθηκε με τη δημιουργία συλλογικών καταλόγων. Ένας φορέας αναλάμβανε να μαζεύει τις βιβλιογραφικές εγγραφές από τις επιμέρους βιβλιοθήκες και να δημιουργεί μια κοινή βάση που αποτελούσε το συλλογικό κατάλογο. Για τις διαδικασίες ταύτισης των εγγραφών σε αυτό το επίπεδο έχουν γραφεί αρκετά¹³. Όμως οι συλλογικοί κατάλογοι παρουσιάζουν ένα σημαντικό μειονέκτημα. Η πληροφορία που παρέχουν δεν είναι πληροφορία πραγματικού χρόνου. Δεν είναι δυνατό να γίνει για παράδειγμα δανεισμός μέσω ενός συλλογικού καταλόγου διότι η πληροφορία για τη διαθεσιμότητα του αντιτύπου δεν αντιστοιχεί στον πραγματικό χρόνο αλλά στο χρόνο αποστολής της πληροφορίας προς τον δημιουργό του συλλογικού καταλόγου.

Από την άλλη πλευρά ξεκίνησαν να αναπτύσσονται προσπάθειες που εντάσσονταν σε μια κατανεμημένη λογική πραγματικού χρόνου. Αντί κάποιος κεντρικός φορέας να συγκεντρώνει τις εγγραφές η ερώτηση να αποστέλλεται μέσω κάποιου πρωτοκόλλου σε όλες τις βιβλιοθήκες και να λαμβάνεται η απάντηση από όλες σε πραγματικό χρόνο. Από αυτή την κατανεμημένη λογική ξεκίνησε και το πρωτόκολλο Z39.50 στο οποίο και θα αφιερωθεί το μεγαλύτερο μέρος της επόμενης ενότητας.

Τα τελευταία όμως χρόνια η έννοια των συλλογικών καταλόγων αναβιώνει με ελαφρώς τροποποιημένη την αρχική της μορφή. Ένα πρωτόκολλο το οποίο κερδίζει έδαφος είναι το Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH). Πρόκειται για ένα πρωτόκολλο που αναπτύσσεται από το Open Archives Initiative¹⁴ και το οποίο έχει βασικό στόχο τη συγκομιδή (harvesting) μεταδεδομένων από αποθετήρια και ψηφιακές βιβλιοθήκες. Πρόκειται για ένα “απλό” σχετικά πρωτόκολλο το οποίο χρησιμοποιεί την XML και επιτυγχάνει την επικοινωνία μέσω HTTP. Χρησιμοποιείται ευρέως από τις ψηφιακές βιβλιοθήκες (ψηφιακά αποθετήρια, ψηφιακά αρχεία κτλ) κυρίως διότι δίνει τη δυνατότητα να συλλέγονται εύκολα δεδομένα τα οποία ικανοποιούν συγκεκριμένα τεκμήρια. Η συλλογή (harvesting) έχει συνήθως ένα περιοδικό χαρακτήρα και πάνω στα δεδομένα που συλλέγονται από τις διάφορες πηγές είναι δυνατό να στηθούν υπηρεσίες προστιθέμενης αξίας. Συνήθως απαντάται σε συστήματα ψηφιακών βιβλιοθηκών και λιγότερο σε συμβατικές

¹³ Για μια εισαγωγή στο θέμα βλέπε τη διπλωματική εργασία του Ανέστη Σίτα με τίτλο “Εντοπισμός διπλών βιβλιογραφικών πόρων ή εγγραφών σε ηλεκτρονικές περιγραφές”. Διατίθεται στο http://dlib.ionio.gr/mtheses/sitas_duplication.doc (τελευταία επίσκεψη 2/6/2008)

¹⁴ <http://www.openarchives.org/>

βιβλιοθήκες αφού οι τελευταίες δίνουν ακόμη μεγαλύτερη έμφαση στα συστήματα πραγματικού χρόνου στα οποία και αφιερώνεται η επόμενη ενότητα.

Z39.50 και SRU/SRW

Ένα από τα πρότυπα που μελετήθηκε σχολαστικά στο χώρο της επιστήμης της Πληροφορίας και συγκεκριμένα από τον αμερικάνικο οργανισμό προτυποποίησης σε θέματα πληροφοριών NISO¹⁵, είναι το Z39.50¹⁶. Το Z39.50 προτάθηκε αρχικά το 1984 ως ένα πρότυπο αναζήτησης / ανάκτησης σε βιβλιογραφικές βάσεις δεδομένων. Προέκυψε μετά από μελέτη πολλών στο χώρο των βιβλιοθηκών και συντηρείται από την Βιβλιοθήκη του Κογκρέσου. Είναι NISO/ANSI πρότυπο και υιοθετήθηκε και από τον διεθνή οργανισμό προτυποποίησης ως ISO 23950. Η τελευταία έκδοση είναι η Z39.50-2003 που είναι αναθεώρηση της Z39.50-1995 – έκδοση 3. Έχει υιοθετηθεί κυρίως από βιβλιοθήκες ως ο τρόπος να κάνουν διαθέσιμο τον κατάλόγό τους προς αναζήτηση από άλλες βιβλιοθήκες διαμέσου του διαδικτύου. Το πρότυπο αποτελεί πρωτόκολλο επικοινωνίας μεταξύ δύο υπολογιστών, του ‘origin’ που αρχικοποιεί ερωτήσεις και του ‘target’ που επεξεργάζεται τις ερωτήσεις και στέλνει τις απαντήσεις, αλλά επίσης περιγράφει και τη μορφή των μηνυμάτων που ανταλλάσσονται μεταξύ ‘origin’ και ‘target’. Στηρίζεται στο client/server μοντέλο και ο ‘origin’ συνήθως καλείται z-client και ο ‘target’ z-server. Η επικοινωνία μεταξύ z-client και z-server επιτυγχάνεται αφού πρώτα ξεκινήσει μεταξύ τους μια νέα συνεδρία (session). Η συνεδρία αυτή, που ονομάζεται από το πρότυπο ‘z-association’, είναι ενεργή για όσο χρόνο ανταλλάσσουν μεταξύ τους δεδομένα ο z-client με τον z-server. Η συνεδρία τερματίζεται όταν ο χρόνος που θα μεσολαβήσει μεταξύ δύο καταστάσεων επικοινωνίας των z-client και z-server ξεπεράσει ένα προκαθορισμένο χρονικό διάστημα (time out). Το χρονικό αυτό διάστημα όπως και άλλα στοιχεία απαραίτητα για την επικοινωνία / διαλειτουργία των z-client και z-server διαπραγματεύονται κατά την αρχικοποίηση της συνεδρίας. Μετά την αρχικοποίηση της συνεδρίας και εφόσον έγινε αποδεκτή από τον z-server ξεκινά η επικοινωνία και η αίτηση εξυπηρέτησης διάφορων λειτουργιών που γνωρίζει ο συγκεκριμένος z-server. Συγκεκριμένα για την επίτευξη μιας αναζήτησης αναλυτικά θα λάβει χώρα η παρακάτω διαδικασία:

- Ο z-client στέλνει μήνυμα αρχικοποίησης στον z-server.
- Ο z-server απαντά αν δέχεται την αρχικοποίηση οπότε και ξεκινά ένα ‘z-association’ το οποίο τερματίζεται είτε από τον z-client είτε από τον z-server για διάφορους λόγους είτε τερματίζεται φυσιολογικά όταν παρέλθει ένα χρονικό διάστημα χωρίς δραστηριότητα από τον z-client.
- Στη συνέχεια ο z-client στέλνει την ερώτηση αναζήτησης την οποία έχει πάρει από τον χρήστη και αναμένει την απάντηση από τον z-server. Η ερώτηση προηγουμένως έχει μετατραπεί από τη μορφή του χρήστη σε μορφή σύμφωνη με το πρότυπο.
- Ο z-server μεταφράζει την ερώτηση σε μορφή που είναι κατανοητή από τις βάσεις δεδομένων που εξυπηρετεί και στη συνέχεια πακετάρει τα αποτελέσματα και τα στέλνει πίσω στον z-client με τη μορφή ενός συνόλου αποτελεσμάτων. Σε αυτό το

¹⁵ <http://www.niso.org/home>

¹⁶ <http://www.loc.gov/z3950/agency/>

σύνολο υπάρχουν πληροφορίες για την αναζήτηση όπως επίσης και οι εγγραφές της βάσης που ικανοποιούν τα κριτήρια της αναζήτησης αν βέβαια υπάρχουν.

Για την ανταλλαγή των βιβλιογραφικών δεδομένων το πρότυπο δίνει τη δυνατότητα χρήσεις αρκετών συντακτικών εγγραφής (record syntax) όπως MARC, SUTRS, OPAC, GRS-1 κλπ. Τη μορφή XML το πρότυπο την αντιμετωπίζει ως συντακτικό εγγραφής. Επίσης ορίζει το σύνολο ιδιοτήτων bib-1 το οποίο καλύπτει έξι κατηγορίες ιδιοτήτων που χρησιμοποιούνται για την σύνθεση της ερώτησης. Αυτές οι κατηγορίες είναι οι:

- Use: σημεία πρόσβασης όπως συγγραφέας, τίτλος κλπ
- Relation: η σχέση ανάμεσα στο σημείο πρόσβασης και του όρου αναζήτησης π.χ. Date-publication <= 1975.
- Position: η θέση του όρου αναζήτησης μέσα στο πεδίο ή υποπεδίο π.χ. στην αρχή του πεδίου.
- Structure: περιγράφει τη δομή του όρου αναζήτησης π.χ. λέξεις ή φράση κλπ.
- Truncation: καθορίζει αν μπορεί να απουσιάζουν χαρακτήρες στη διαδικασία ταιριάσματος του όρου αναζήτησης με τον όρο της βάσης π.χ. δεξιά αποκοπή σημαίνει ότι ζητείται ο όρος στη βάση να ξεκινά με αυτό που υπαγορεύει ο όρος αναζήτησης.
- Completeness: περιγράφει αν ο όρος αναζήτησης αντιπροσωπεύει για παράδειγμα την τιμή όλου του πεδίο στη βάση.

Το πρότυπο ορίζει επίσης και λογικούς τελεστές για την σύνθεση πιο πολύπλοκης ερώτησης όπως και τελεστές εγγύτητας (proximity).

Οι σύγχρονοι z-clients έχουν τη δυνατότητα να αρχικοποιούν συνεδρίες ταυτόχρονα με πολλούς z-servers με προφανές το πλεονέκτημα της ταχύτητας και της καλύτερης εξυπηρέτησης του χρήστη. Βέβαια κάθε επιχειρούμενη ομογενοποίηση δημιουργεί προβλήματα διαλειτουργικότητας. Η αναζήτηση ταυτόχρονα σε πολλές πηγές επιτάσσει την ολοκλήρωση των επιστρεφόμενων αποτελεσμάτων για την καλύτερή τους αξιοποίηση από τον χρήστη. Δεν θα ήταν και τόσο επιθυμητό ο χρήστης να βλέπει δεδομένα σε διαφορετικά σχήματα και κωδικοποιήσεις ως αποτέλεσμα μιας μόνο αναζήτησής του. Επομένως απαιτούνται διαδικασίες συγχώνευσης, ταξινόμησης, διαγραφής των ίδιων εγγραφών κ.α. δηλαδή αρχικά διαδικασίες εξασφάλισης διαλειτουργικότητας δεδομένων.

Το πρότυπο εκτός από τις υπηρεσίες αναζήτησης / ανάκτησης βιβλιογραφικών εγγραφών παρέχει και άλλες υπηρεσίες όπως είναι η παρουσίαση όρων ευρετηρίου (scan/browse), η ταξιθέτηση (sorting), ο έλεγχος πρόσβασης σε πόρους, η παρουσίαση λεπτομερειών για τις βάσεις δεδομένων τις οποίες υποστηρίζει ο z-server και οι εκτεταμένες υπηρεσίες όπως η παραγγελία αντιτύπου, η αποθήκευση ερωτημάτων, η ενημέρωση και διαγραφή εγγραφών στη βάση δεδομένων, ο διαδανεισμός. Χρειάζεται όμως εφαρμογή πολιτικών από τις βιβλιοθήκες για να επιτευχθεί η σημασιολογική διαλειτουργικότητα μεταξύ των ανταλλασσόμενων δεδομένων και υπηρεσιών. Το πρότυπο δεν μπορεί να την εξασφαλίσει. Αποτελεί μόνο το συντακτικό – τεχνικό όχημα για τη μεταφορά των δεδομένων / εκτέλεση υπηρεσιών μεταξύ των βιβλιοθηκών. Όταν για παράδειγμα γίνεται αναζήτηση σε

πολλούς z-servers ταυτόχρονα και απαιτείται ενοποίηση των αποτελεσμάτων και πιθανόν ολοκληρωμένη εμφάνισή τους τότε θα πρέπει να είναι γνωστό τι σημαίνουν τα δεδομένα προερχόμενα από διαφορετικά σχήματα. Αλλά και η δομική διαλειτουργικότητα είναι ένα θέμα. Το πρότυπο προτείνει τρόπους δόμησης / «πακεταρίσματος» των δεδομένων έτσι ώστε να μεταφερθούν και να αναγνωριστούν. Η γλώσσα XML αποτελεί το καλύτερο παράδειγμα την οποία το πρότυπο όμως αντιμετωπίζει ως record syntax αλλά τα νέα πρότυπα SRU/SRW βασισμένα στο Z39.50 τη θεωρούν ως εξ ορισμού τρόπο πακεταρίσματος.

Ένα άλλο θέμα που ανακύπτει από τις υπηρεσίες που παρέχει το πρότυπο είναι η πρόσβαση στην πληροφορία. Τεχνικά ο οποιοσδήποτε θα έχει τη δυνατότητα να αναζητήσει και να κατεβάσει τη βιβλιογραφική εγγραφή. Αυτό όμως μπορεί να έρχεται σε σύγκρουση με την πολιτική της εκάστοτε βιβλιοθήκης που δίνει τα δεδομένα της. Επομένως θα πρέπει να χρησιμοποιηθούν οι υπηρεσίες ελέγχου αυθεντικότητας του χρήστη και να εφαρμοστούν κανόνες πρόσβασης στο υλικό έτσι ώστε να υπάρχει πρόσβαση ελεγχόμενη σε επίπεδο χρήστη.

Το πρότυπο δεν υπαγορεύει πολιτικές υλοποίησης της βάσης δεδομένων ή άλλων μερών στην εκάστοτε βιβλιοθήκη απλά την συμμόρφωση στους κανόνες επικοινωνίας και στη μορφή των μηνυμάτων που ανταλλάσσονται. Παρέχει έτσι αρκετή ευελιξία τόσο στην υλοποίηση των υπηρεσιών αλλά και στη μορφή των ανταλλάξιμων δεδομένων. Αυτή όμως η ευελιξία δημιουργεί προβλήματα διαλειτουργικότητας όταν δεν έχει καθοριστεί κάποιο επίπεδο συμμόρφωσης τόσο στο είδος των υπηρεσιών που θα πρέπει να υλοποιηθούν όσο και στη μορφή των δεδομένων που θα ανταλλάσσονται. Έτσι αν και ξεκίνησε το πρότυπο ως ένα καλό εργαλείο διαλειτουργικότητας μεταξύ των βιβλιοθηκών για την αναζήτηση και ανταλλαγή βιβλιογραφικών δεδομένων στην πραγματικότητα δημιουργεί προβλήματα διαλειτουργικότητας τα οποία βέβαια δεν ανήκουν ή δεν παρουσιάζονται μόνο στην z39.50 επικοινωνία. Αυτή η έλλειψη καλύφθηκε σε κάποιο βαθμό στην έκδοση 3 του προτύπου αλλά κυρίως καλύφθηκε από αυτό που ονομάζεται προφίλ εφαρμογής (application profile). Το Bath Profile¹⁷ είναι το πιο γνωστό το οποίο θέτει τους ελάχιστους κανόνες συμμόρφωσης σε προσφερόμενες υπηρεσίες αλλά και σε μορφή δεδομένων και άλλων λεπτομερειών όπως το είδος των προσφερόμενων ιδιοτήτων αναζήτησης (attribute set).

Το πρότυπο παρόλο που είναι πολύ διαδεδομένο στη βιβλιοθηκονομική κοινότητα δεν έχει την ίδια εξάπλωση σε άλλες κοινότητες αν και δίνει τη δυνατότητα για διαχείριση όχι μόνο βιβλιογραφικής πληροφορίας. Αυτό ίσως οφείλεται γενικά στην πολυπλοκότητά του. Αρχικά η ανάγκη αρχικοποίησης συνεδρίας για τη λειτουργία δημιουργεί προβλήματα στην περίπτωση που η συνεδρία λήξει και ο χρήστης θέλει να συνεχίσει επειδή όλη η πληροφορία της συνεδρίας χάνεται (αποτελέσματα αναζητήσεων κλπ). Επίσης θεωρείται πολύπλοκη η δημιουργία και επεξεργασία των μηνυμάτων μεταξύ z-client και z-server τα οποία στηρίζονται σε κανόνες προτύπων όπως ASN και BER που αποτελούν τρόπους αναπαράστασης και κωδικοποίησης των δεδομένων. Η χρήση πιο υψηλού επιπέδου προτύπων/πρωτοκόλλων όπως η XML, τα πρωτόκολλα HTTP και SOAP θα έκανε πιθανόν πιο ελκυστική τη χρήση του Z39.50 και από άλλους χώρους εκτός των βιβλιοθηκών. Όμως το Z39.50 λόγω της μεγάλης του αποδοχής από τη βιβλιοθηκονομική κοινότητα δεν παραμένει σταθερό αλλά είτε

¹⁷ <http://www.collectionscanada.gc.ca/bath/ap-bath-e.htm>

εξελίσσεται το ίδιο είτε δημιουργούνται και άλλοι τρόποι χρήσης του που στηρίζονται στις νέες τεχνολογίες κάθε φορά. Χαρακτηριστικό παράδειγμα η τεχνολογία των Web Services που οδήγησε στην υλοποίηση του προτύπου SRU/SRW για την χρήση των υπηρεσιών του πρωτοκόλλου Z39.50 ως web services είτε με την τεχνολογία REST είτε με την τεχνολογία SOAP.

Το SRU¹⁸ είναι ένα πρότυπο για εκτέλεση αναζητήσεων στο περιβάλλον του Web, αποτελεί δηλαδή ένα web service. Τα ερωτήματα συντάσσονται σύμφωνα με τη γλώσσα ερωτήσεων CQL¹⁹ και μαζί με όλες τις άλλες πληροφορίες που απαιτούνται στέλνονται σε έναν εξυπηρετητή ερωτήσεων, στην περίπτωση βιβλιογραφικών δεδομένων σε έναν Z39.50 server. Όλο το ερώτημα πακετάρεται σε ένα HTTP URL και η επικοινωνία με τον server γίνεται διαμέσου του πρωτοκόλλου HTTP η οποία δεν δημιουργεί και κάποια συγκεκριμένη κατάσταση μεταξύ client και server όπως συμβαίνει στον παραδοσιακό τρόπο χρήσης του Z39.50. Η απάντηση είναι ένα XML document το οποίο περιέχει τις εγγραφές και/ή άλλες πληροφορίες.

Το SRW²⁰ είναι επίσης ένα web service για την εκτέλεση αναζητήσεων στο περιβάλλον του Web. Η διαφορά του με το SRU είναι ότι το ερώτημα δεν στέλνεται στον server διαμέσου ενός απλού HTTP URL αλλά πακετάρεται σύμφωνα με το πρωτόκολλο SOAP²¹ δημιουργώντας έτσι ένα xml document. Το SOAP αποτελεί πρωτόκολλο επικοινωνίας μεταξύ εφαρμογών στο διαδίκτυο περιγράφοντας τη μορφή των μηνυμάτων μεταξύ των εφαρμογών με τη χρήση της xml. Τα μηνύματα διακινούνται με τη χρήση των http και https πρωτοκόλλων και δεν αποκλείονται από τα firewalls.

Με τη χρήση των προτύπων SRU/SRW είναι δυνατή και πιο απλή πλέον η αναζήτηση / ανάκτηση βιβλιογραφικών δεδομένων στο περιβάλλον του παγκοσμίου ιστού από Z39.50 servers και έτσι οι βιβλιοθήκες εντάσσοντας αυτό τον τρόπο στο σύστημά τους παραμένουν διαλειτουργικές και άμεσα προσβάσιμες από άλλες βιβλιοθήκες. Με εύκολο τρόπο μπορούν να ανταλλάσσουν δεδομένα μεταξύ τους. Μεγάλο πλεονέκτημα αποτελεί η υποστήριξη της Unicode κωδικοποίησης χαρακτήρων και έτσι μπορεί να γίνει αναζήτηση σε οποιαδήποτε γλώσσα. Το SRU/SRW είναι ένα παράδειγμα χρήσης της τεχνολογίας των web services στο χώρο των βιβλιοθηκών.

Διαλειτουργικότητα και ψηφιακές βιβλιοθήκες

Οι ψηφιακές βιβλιοθήκες γνωρίζουν μεγάλη ανάπτυξη στις μέρες μας. Αυτό οφείλεται κυρίως στη μεγάλη παραγωγή ψηφιακού υλικού και την ανάγκη για επεξεργασία αυτού του υλικού ώστε για να γίνει προσιτό στους χρήστες. Στο περιβάλλον του διαδικτύου είναι επιβεβλημένη η συνεργασία των ψηφιακών βιβλιοθηκών για να καταφέρουν να ανταποκριθούν στο ρόλο τους. Για να επιτευχθεί αυτή η συνεργασία θα πρέπει να θεσπιστούν κανόνες τόσο στην κατασκευή των

¹⁸ Search/Retrieve via URL: <http://www.loc.gov/standards/sru/>

¹⁹ Contextual Query Language: <http://www.loc.gov/standards/sru/specs/cql.html>

²⁰ Search/Retrieve Web Service: <http://www.loc.gov/standards/sru/sru1-1archive/srw.html>

²¹ Simple Object Access Protocol: <http://www.w3.org/TR/soap/>

ψηφιακών βιβλιοθηκών όσο και στην μεταξύ τους επικοινωνία και ανταλλαγή δεδομένων και μεταδεδομένων έτσι ώστε να καταφέρουν να διαλειτουργήσουν.

Η ψηφιακή βιβλιοθήκη είναι ο χώρος όπου η διαλειτουργικότητα είναι τόσο στόχος όσο και προϋπόθεση. Στην παρούσα εργασία αντιμετωπίζονται ως χωριστή κατηγορία η οποία προϋποθέτει όλα τα προηγούμενα αλλά προσθέτει και το ίδιο το τεκμήριο. Επίσης, στις ψηφιακές βιβλιοθήκες είναι on line οι περισσότερες υπηρεσίες που παρέχονται (π.χ. το document delivery δε γίνεται από courier αλλά πιθανότατα μέσω http) άρα τίθεται πιο έντονο το ζήτημα των υπηρεσιών που προσφέρονται και το πως αυτές υλοποιούνται μέσω των υπολογιστικών συστημάτων. Έτσι στην συγκεκριμένη ενότητα θα επιχειρηθεί μια κάθετη τομή σε θέματα διαλειτουργικότητας που διαπερνά όλο το φάσμα των υπηρεσιών μιας «ψηφιακής» βιβλιοθήκης και αντίστοιχα όλο το φάσμα της διαλειτουργικότητας.

Επιλέχθηκε να αναλυθεί εδώ η σύλληψη των components που διαλειτουργούν διότι εδώ αποκαλύπτεται σε όλες της τις διαστάσεις. Γίνεται σαφές ότι πρόκειται για ξεχωριστές υπηρεσίες που επικοινωνούν μεταξύ τους και παράλληλα μπορούν να διατηρούν μια σχετική αυτονομία και αυτοτέλεια.

Διάφορες αρχιτεκτονικές υλοποίησης έχουν προταθεί και σύμφωνα με τους σκοπούς της κάθε υπό κατασκευή ψηφιακή βιβλιοθήκη προτιμά την κατάλληλη για την εξυπηρέτηση αυτών των σκοπών. Στο ολόενα και αυξανόμενο πληροφοριακό περιβάλλον υπάρχει η ανάγκη για τη δημιουργία ψηφιακών βιβλιοθηκών με παγκόσμιες διαστάσεις όχι μόνο για τεχνικούς λόγους αλλά και εξαιτίας του γεγονότος της ύπαρξης των αποθετηρίων πληροφορίας (information repositories) και των υπηρεσιών επεξεργασίας των πληροφοριών (information processing services) σε ανεξάρτητους οργανισμούς που είναι διασκορπισμένοι σε όλο τον κόσμο. Αυτό καταστεί αναγκαία την κατασκευή της ψηφιακής βιβλιοθήκης ως μια συλλογή από ανεξάρτητα κατασκευασμένα components τα οποία βασίζονται το ένα στο άλλο για να επιτύχουν μεγαλύτερες εργασίες. Ο τελικός σκοπός τέτοιων συστημάτων είναι αν και τα components αναπτύσσονται ανεξάρτητα να καλούν το ένα το άλλο αποδοτικά και εύκολα. Αυτά τα components είναι μικρά, ανεξάρτητα, προγράμματα τα οποία επιτελούν κάποια λειτουργία αλλά που συνήθως πολλά από αυτά θα πρέπει να συνεργαστούν για να ολοκληρώσουν πιο σύνθετες εργασίες που απαιτούνται στα πλαίσια ενός συστήματος βιβλιοθηκών είτε πρόκειται για το ψηφιακό είτε για το συμβατικό περιβάλλον. Πολλές τεχνολογίες υποστηρίζουν αυτό το είδος παραγωγής λογισμικού όπως οι τεχνολογίες OLE²², COM²³, DCOM²⁴, CORBA²⁵ και πιο πρόσφατα Web Services. Η κατασκευή μιας παγκόσμιας ψηφιακής βιβλιοθήκης που βασίζεται στην παραπάνω λογική της συνεργασίας των components παρουσιάζει ιδιαίτερο ενδιαφέρον από πλευράς διαλειτουργικότητας, αφού ανακύπτουν ζητήματα τα οποία είναι παρόμοια με τα θέματα διαλειτουργικότητας σε ένα ολοκληρωμένο σύστημα βιβλιοθήκης με την παραδοσιακή σημασία. Συνήθως χρησιμοποιούνται διάφοροι όροι για να περιγράψουν μια τέτοια συνεργασία όπως οι: "heterogeneous systems" και "federated systems". Υπάρχει επίσης και ο πιο γενικός όρος "distributed systems" που ίσως όμως αφήνει να εννοηθεί ότι τα components που απαρτίζουν το

²² Object Linking and Embedding: http://en.wikipedia.org/wiki/Object_linking_and_embedding

²³ Component Object Model: http://en.wikipedia.org/wiki/Component_Object_Model

²⁴ Distributed Component Object Model:

http://en.wikipedia.org/wiki/Distributed_Component_Object_Model

²⁵ Common Object Request Broker Architecture: <http://en.wikipedia.org/wiki/CORBA>

τελικό σύστημα είναι πιο προσεγγμένα κατασκευασμένα και έχουν ως σκοπό να διαλειτουργούν ολοκληρωμένα μεταξύ τους.

Τα θέματα διαλειτουργικότητας που προκύπτουν σε ένα παγκόσμιο σύστημα ψηφιακής βιβλιοθήκης που στηρίζεται στη συνεργασία ανεξάρτητων components είναι αντίστοιχα με αυτά που προκύπτουν όταν τα παραδοσιακά ολοκληρωμένα συστήματα βιβλιοθηκών θέλουν να διαλειτουργήσουν είτε μεταξύ τους είτε με άλλα προγράμματα, όπως για παράδειγμα με το portal του ιδρύματος στο οποίο υπάγεται η βιβλιοθήκη. Ο διαχωρισμός βέβαια σε παραδοσιακά ολοκληρωμένα προγράμματα βιβλιοθηκών και προγράμματα ψηφιακών βιβλιοθηκών μάλλον δεν είναι τόσο διακριτός στις μέρες μας αφού και τα δύο στηρίζονται στις ίδιες τεχνολογίες και τα πρώτα πλέον διαχειρίζονται και ψηφιακό υλικό. Επομένως αν και υπάρχουν κάποιες διαφορές μεταξύ τους, τα θέματα διαλειτουργικότητας που ανακύπτουν στο καθαρά ψηφιακό περιβάλλον μάλλον είναι κοινά και το μόνο σίγουρο είναι ότι σήμερα με την πρόκληση της μεγάλης ποσότητας της ψηφιακής πληροφορίας και της ανάγκης της διαχείρισής της ανέκυψαν περισσότερα προβλήματα διαλειτουργικότητας κάνοντας επιβεβλημένη τη λύση τους για να μπορεί να υπάρξει ψηφιακή βιβλιοθήκη και κυρίως ψηφιακή βιβλιοθήκη σε παγκόσμια διάσταση. Σύμφωνα με τους Paercke κ.α. (1998) τα θέματα διαλειτουργικότητας που προκύπτουν από μια τέτοια συνεργατική παγκόσμια βιβλιοθήκη αναδύονται αν εξετάσουμε πέντε βασικές λειτουργίες των ψηφιακών βιβλιοθηκών όπως αναφέρονται παρακάτω.

- Διαχείριση Πληροφορίας (Information Management): αναφέρεται στην αποθήκευση, την οργάνωση, και την ανάκτηση των πληροφοριών. Περιλαμβάνει το μεγαλύτερο κομμάτι δραστηριοτήτων σε μια βιβλιοθήκη και ο τελικός στόχος για την επίτευξη της διαλειτουργικότητας είναι η πλήρης ανεξαρτησία από μορφότυπα δεδομένων, μοντέλα αναπαράστασης τεκμηρίων και γλώσσες ανθρώπινης επικοινωνίας. Κάθε ένα από τα συνεργαζόμενα components/συστήματα βιβλιοθηκών να μπορεί για παράδειγμα να χρησιμοποιεί, τον δικό του τρόπο για την αναπαράσταση των τεκμηρίων, αλλά τα τεκμήρια να μπορούν να ανταλλάσσονται και να εμφανίζονται σε διαφορετικές υπολογιστικές πλατφόρμες και ίσως και τα εμπόδια που θέτουν οι ανθρώπινες γλώσσες να μπορούν να ξεπεραστούν. Τεχνολογίες ενδιάμεσων (wrappers/mediators) που επιτελούν τουλάχιστον το ρόλο του μεταφραστή μεταξύ των διαφορετικών τοπικών προσεγγίσεων στη διαχείριση της πληροφορίας είναι επιβεβλημένες για τη εξασφάλιση της διαλειτουργικότητας.
- Παρουσίαση Πληροφορίας (Information Presentation): αναφέρεται στην παρουσίαση του τεκμηρίου σε διαφορετικές υπολογιστικές πλατφόρμες. Απαιτείται η συμμόρφωση σε κοινούς κανόνες για την κατασκευή τεχνολογικών συστημάτων / προγραμμάτων που να δείχνουν την πληροφορία με τον ίδιο τρόπο. Ο σκοπός όμως είναι να μπορεί να εμφανιστεί η πληροφορία σε διαφορετικές υπολογιστικές πλατφόρμες και σε διαφορετικά υπολογιστικά συστήματα χωρίς την υιοθέτηση των ίδιων κοινών κανόνων αλλά η εξασφάλιση ενός βαθμού ετερογένειας.
- Επικοινωνία (Communication): Αναφέρεται στην επικοινωνία μεταξύ των components/συστημάτων βιβλιοθηκών. Ανεξάρτητα των διαφορετικών δικτύων των συστημάτων ο τελικός σκοπός είναι να μπορούν να επικοινωνούν μεταξύ τους. Η χρήση πάλι τεχνολογιών ενδιάμεσων αποτελούν λύση όπως

για παράδειγμα η χρήση των πυλών (gateways) για τη γεφύρωση διαφορετικής τεχνολογίας δικτύων.

- Λειτουργίες (Operations): Αναφέρεται στην έναρξη και τον έλεγχο των λειτουργιών μεταξύ των διαφορετικών components/συστήματα βιβλιοθηκών. Οι τεχνολογίες που δίνουν ευκολίες απομακρυσμένου υπολογισμού όπως οι CORBA, DCOM, Java δίνουν τη δυνατότητα ετερογένειας/αυτονομίας στα υπό επικοινωνία components αλλά ταυτόχρονα εξασφαλίζουν τη διαλειτουργικότητάς τους. Ο τελικός στόχος λοιπόν είναι τα συστήματα να λειτουργούν επιτρέποντας τα «heterogeneous components» να έρχονται on-line, να διαφημίζουν τις δυνατότητές τους και να συμμετέχουν σε peer-to-peer αλληλεπίδραση με άλλα components. Αυτό το όραμα είναι πολύ δύσκολο να πραγματοποιηθεί, επειδή δεν είναι σαφές πώς να περιγραφεί η λειτουργικότητα του κάθε component έτσι ώστε άλλα components να μπορούν να την καταλάβουν και "να αποφασίσουν" αυτόματα ότι αυτή η λειτουργία που προσφέρεται είναι κατάλληλη για μια δεδομένη εργασία. Τεχνολογίες και μοντέλα αναπαράστασης/διαχείρισης γνώσης και του σημασιολογικού ιστού κινούνται προς αυτή την κατεύθυνση όπως οι XML, RDF, OWL, Web Services κλπ.
- Προστασία (Protection): Αναφέρεται στην προστασία των χρηστών και των τεκμηρίων τους και στις πηγές πληροφοριών. Ο στόχος είναι απλά να δηλωθούν οι όροι και οι διατάξεις για μια αλληλεπίδραση και να έχει το σύστημα να φροντίζει τα υπόλοιπα. Παραδείγματος χάριν, ένα τεκμήριο να μπορεί να ταξιδέψει μεταξύ των components με συνημμένες οδηγίες για το γεγονός ότι το περιεχόμενό του μπορεί να διαβαστεί και να διαβιβαστεί προς ένα άλλο component αλλά δεν πρέπει να αντιγραφεί. Κατάλληλο watermarking για την ανίχνευση παραβιάσεων μπορεί να χρησιμοποιηθεί.

Σε ποιο βαθμό όμως απαιτείται διαλειτουργικότητα είναι ένα θέμα που πρέπει να ληφθεί υπόψη στην κατασκευή συστημάτων από αλληλοσυνεργαζόμενα components. Σε ένα διαλειτουργικό σύστημα το ζητούμενο είναι όλα τα components να είναι το ίδιο πλούσια σε λειτουργικότητα, το ίδιο εκφραστικά στη μοντελοποίηση των δεδομένων, δηλαδή περιβάλλον απόλυτης συμμόρφωσης και πλήρους ομογένειας; Στην περίπτωση αυτή όμως, αν κάποιο component υλοποιήσει επιπλέον λειτουργικότητα τότε αυτή η επιπλέον λειτουργικότητα θα χαθεί αφού δεν έχει υιοθετηθεί από άλλους. Αν όμως ζητείται χαμηλότερος βαθμός διαλειτουργικότητας τότε μπορεί να υπάρξουν κάποιες διαφορές στα components του συστήματος δίνοντας την ευελιξία της διαφορετικότητας στην υλοποίηση και στις πολιτικές. Στις διεπαφές χρήστη τι βαθμός διαλειτουργικότητας απαιτείται; Αν απαιτείται πλήρης διαλειτουργικότητα τότε όλες οι διεπαφές από τις διαφορετικές συλλογές που συμμετέχουν στο καταναμημένο σύστημα θα πρέπει να βρίσκονται πίσω από μια ενιαία διεπαφή για να μην αντιμετωπίζει ο χρήστης διαφορετικά περιβάλλοντα σε κάθε συλλογή. Αν όμως δεν ζητείται πλήρης ομοιογένεια τότε θα μπορούν να υπάρξουν και διαφοροποιήσεις στις διεπαφές. Επίσης σε ένα σύστημα ομόσπονδης αναζήτησης (federated search) ποιος είναι ο επιθυμητός αλλά και πραγματικός βαθμός διαλειτουργικότητας; Απαντούν όλα τα συστήματα τα οποία ερωτώνται από τον metasearch client στους ίδιους χρόνους; Αν δεν συμβαίνει τότε είναι δύσκολη η αναζήτηση σε πραγματικό χρόνο και άρα απαιτούνται τεχνικές προεπεξεργασίας και προαποθήκευσης των δεδομένων για να επιτευχθεί η πλήρη διαφάνεια στις

διαδικασίες σαν να ήταν ένα σύστημα, μια συλλογή που λειτουργεί σε πραγματικό χρόνο. Παράδειγμα η γνωστή μηχανή αναζήτησης Google στην οποία οι πληροφορίες συγκεντρώνονται πιο πριν και η αναζήτηση δεν γίνεται σε δεδομένα του πραγματικού χρόνου. Άλλο παράδειγμα το πρωτόκολλο OAI-PMH για τη συγκομιδή δεδομένων από ψηφιακές βιβλιοθήκες και αποθετήρια για μετέπειτα χρήση τους.

Πώς μπορεί όμως να μετρηθεί η διαλειτουργικότητα ενός συνεργατικού συστήματος με μέρη ανεξάρτητα components; Τα παρακάτω κριτήρια θεωρούνται σημαντικά κατά τους Paercke κ.α. (1998):

- Υψηλός βαθμός αυτονομίας των components
- Χαμηλό κόστος υποδομής
- Ευκολία στην παροχή/συνεισφορά νέων components
- Ευκολία στη χρήση των components
- Εύρος υποστηριζόμενων εργασιών
- Ευελιξία στον αριθμό των components

Αυτά δεν αποτελούν ποσοτικές μετρήσεις αλλά μπορούν να χρησιμοποιηθούν ως κατευθυντήριες οδηγίες για την κατανόηση των κατανεμημένων και διαλειτουργικών ψηφιακών βιβλιοθηκών. Βέβαια ορισμένες φορές όταν ικανοποιείται καλά ένα κριτήριο μπορεί να επηρεάζεται αρνητικά ένα άλλο. Για παράδειγμα έχοντας χαμηλό κόστος στην υποδομή θα κατασκευαστούν απλά components τα οποία θα μπορούν να χρησιμοποιηθούν ενδεχομένως σε απλές εργασίες. Άλλο παράδειγμα η κατασκευή όσο το δυνατό ανεξάρτητων components οδηγεί σε υψηλό κόστος υποδομής για την επίτευξη της διαλειτουργικότητας αφού πρέπει να περιγραφούν με λεπτομέρεια πολλά διαφορετικά components και να κατασκευαστούν πολλοί διαφορετικοί και ενδεχομένως πολύπλοκοι ενδιάμεσοι (wrappers/mediators) μηχανισμοί. Με τόση διαφορετικότητα γίνεται επίσης δύσκολη η χρήση αλλά και η συνεισφορά νέων components.

Διάφορες προσεγγίσεις έχουν ακολουθηθεί κατά καιρούς για την κατασκευή διαλειτουργικών συστημάτων. Πολλές φορές είναι δύσκολο να χρησιμοποιηθεί μία και μόνο προσέγγιση. Συνήθως ο συνδυασμός διαφορετικών προσεγγίσεων είναι αυτός που δίνει τα καλύτερα αποτελέσματα. Σύμφωνα με τους Paercke κ.α. (1998) οι προσεγγίσεις αυτές μπορούν να συνοψιστούν στις παρακάτω:

Χρήση γνωστών και ευρέως διαδεδομένων προτύπων

Μία από τις πιο παλιές μεθόδους για να επιτευχθεί διαλειτουργικότητα μεταξύ ετερογενών συστημάτων είναι η συμφωνία στη χρήση κάποιων προτύπων έτσι ώστε να επιτευχθεί ο βαθμός ομοιογένειας που απαιτείται. Τα πρότυπα αυτά, όπως η σειρά ISO 802 για τα δίκτυα υπολογιστών, το z39.50 για αναζήτηση/ανάκτηση πληροφοριών κλπ προκύπτουν από την ανάγκη για διαλειτουργικότητα όταν υπάρχουν πλέον πολλά διαφορετικά συστήματα τα οποία είναι αδύνατο να διαλειτουργήσουν. Αυτά τα πρότυπα προτείνονται και συντηρούνται από διεθνείς οργανισμούς. Υπάρχουν βέβαια και τα de facto πρότυπα, τα οποία είναι λύσεις κάποιων εταιριών που τυγχάνουν μεγάλης αποδοχής όπως το λειτουργικό σύστημα

DOS ή τα Windows στο χώρο των λειτουργικών συστημάτων. Επίσης ορισμένες φορές κυβερνητικοί οργανισμοί προωθούν τα πρότυπα και γίνονται παγκόσμια αποδεκτά όπως έγινε με το USMARC. Τα de facto πρότυπα προκύπτουν συνήθως σε περιπτώσεις που νέες ιδέες έρχονται το σωστό χρόνο για να καλύψουν κάποια νέα ανάγκη και γίνονται παγκοσμίως αποδεκτά όπως έγινε με τα πρωτόκολλα μεταφοράς HTTP και MIME ή με τη γλώσσα επισήμανσης HTML.

Ο βαθμός επιτυχίας ή όχι ενός προτύπου εξαρτάται πολλές φορές και από κοινωνικούς και επιχειρηματικούς παράγοντες και όχι μόνο από τις τεχνικές του δυνατότητες. Οι εταιρίες για παράδειγμα αντιδρούν στην επισημοποίηση ενός προτύπου όταν νομίζουν ότι οι ίδιες μπορούν να κατασκευάσουν ένα de facto πρότυπο που θα τις κάνει να έχουν ηγεμονικό ρόλο στο συγκεκριμένο χώρο της αγοράς.

Ιδιαίτερη αναφορά πρέπει να γίνει στην περίπτωση των ανοικτών προτύπων. Ένα ανοικτό πρότυπο πληροί τα παρακάτω κριτήρια²⁶:

- Το πρότυπο ορίζεται ή υιοθετείται και κατόπιν συντηρείται από κάποιον μη κερδοσκοπικό οργανισμό, και η ανάπτυξή του γίνεται στα πλαίσια μιας διαδικασίας αποφάσεων που είναι ανοικτή σε όλους τους ενδιαφερόμενους (συμφωνία πλειοψηφίας κλπ).
- Το πρότυπο έχει δημοσιευτεί και το έγγραφο προδιαγραφών (specification) γίνεται διαθέσιμο είτε δωρεάν είτε με συγκεκριμένη χρέωση. Πρέπει να επιτρέπεται η αντιγραφή, διανομή και χρήση δωρεάν ή με συγκεκριμένο κόστος.
- Η πνευματική ιδιοκτησία (πχ. πατέντες) μέρη ή όλου του προτύπου διατίθεται δια παντός για χρήση χωρίς χρέωση.

Χαρακτηριστικά ανοικτά πρότυπα τα οποία είναι ευρέως διαδεδομένα αποτελούν τα PDF, HTML για εμφάνιση και μορφοποίηση κειμένου, HTTP για μεταφορά υπερκειμένου στο διαδίκτυο και FTP για μεταφορά αρχείων στο διαδίκτυο. Από τα κριτήρια προκύπτει ότι η χρήση τους από τα ολοκληρωμένα προγράμματα βιβλιοθηκών (ψηφιακών ή μη) θα οδηγήσει σε χαμηλότερο κόστος αφού τα περισσότερα είναι ελεύθερα προς χρήση αλλά και σε ελευθερία χρήσης της πληροφορίας χωρίς περιορισμούς. Αν το πρότυπο είναι ιδιοκτησιακό (proprietary) τότε δεν εξασφαλίζει άμεση και πλήρη διαλειτουργικότητα ειδικά στην περίπτωση που είναι και κλειστό (δηλαδή και κάποιος έχει ιδιοκτησία και δεν έχει δημοσιεύσει την κωδικοποίηση). Ενδεχομένως δεν θα υπάρχουν για παράδειγμα μετατροπές του υλικού που είναι αποθηκευμένο σύμφωνα με αυτό το κλειστό πρότυπο σε άλλες μορφές όπως συνέβαινε παλιότερα με το de facto πρότυπο μορφοποίησης κειμένου το "doc". Σε αντίθεση το ανοικτό πρότυπο επιτρέπει της μετατροπή της πληροφορίας σε άλλες μορφές από όποιον το επιχειρήσει αφού είναι διαφανής ο τρόπος υλοποίησής του, ο πρώτος με τον οποίο αναπαρίσταται η πληροφορία. Ένα άλλο σημαντικό πλεονέκτημα του ανοικτού προτύπου ακριβώς λόγω της διαφάνειάς του είναι ότι εξασφαλίζει μεγαλύτερο χρόνο ζωής στην πληροφορία από το ίδιο το πρότυπο ή το λογισμικό χρήσης. Άρα βοηθάει και στη διατήρηση (preservation). Ένα ανοικτό πρότυπο λοιπόν είναι ένα καλό εργαλείο διαλειτουργικότητας ιδίως όταν χρησιμοποιείται από πολλούς γιατί τότε δεν χρειάζονται και πολλές μετατροπές.

²⁶ Διαθέσιμα στο <http://www.digitalrights.gr/tiki/tiki-index.php?page=OpenStandards>

Χαρακτηριστικό παράδειγμα το Z39.50 στο οποίο στηρίζονται πολλές μηχανές αναζήτησης βιβλιογραφικής πληροφορίας.

Ένα πρότυπο γενικά αν είναι κλειστό ή ανοικτό είναι λοιπόν ένα καλό εργαλείο διαλειτουργικότητας αν χρησιμοποιείται από πολλούς. Δίνει τη δυνατότητα να στηριχθεί η δημιουργία προγραμμάτων / υπηρεσιών πάνω σε αυτό και έτσι καλύπτεται το κριτήριο της ευκολίας στη χρήση των components και της συνεισφοράς νέων components και υπηρεσιών. Επίσης ένα αξιόπιστο πρότυπο ενθαρρύνει την επένδυση σε υποδομή. Μπορεί όμως να είναι πολλά τα πρότυπα που είναι κατάλληλα για την εκάστοτε εργασία, το οποίο σημαίνει δυσκολία στην απόφαση της χρήσης ενός κοινού προτύπου. Επίσης η υιοθέτηση ενός προτύπου μειώνει την αυτονομία του κάθε component και ίσως του στερεί τη δημιουργία τοπικών παρεκκλίσεων. Μία λύση φαίνεται να είναι η υιοθέτηση προαιρετικών τμημάτων στο πρότυπο έτσι ώστε να δίνεται η δυνατότητα τοπικών αποφάσεων / υλοποιήσεων αλλά αυτό μπορεί να οδηγήσει σε πολυπλοκότητα του προτύπου και ελάττωση της διαλειτουργικότητας. Χαρακτηριστικό παράδειγμα αποτελεί το Z39.50 το οποίο έχει αρκετά προαιρετικά στοιχεία στην ανταλλαγή των μηνυμάτων καθώς επίσης και ευελιξία στην υλοποίηση μορφών αναπαράστασης δεδομένων εκτός των κανόνων του προτύπου. Έτσι υπάρχουν προφανή πλεονεκτήματα αυτονομίας αλλά ταυτόχρονα ο κίνδυνος της ασυμβατότητας με άλλες υλοποιήσεις του προτύπου που δεν υιοθέτησαν τις ίδιες κοινές στρατηγικές υλοποίησης. Βέβαια στην περίπτωση του Z39.50 επικράτησε η ιδέα των προφίλ εφαρμογής για κοινή συμμόρφωση. Τελικά συνετή λύση φαντάζει η υιοθέτηση περισσότερων του ενός κοινών προτύπων.

Οικογένειες προτύπων

Στις οικογένειες προτύπων (families-of-standards) οι κατασκευαστές των components έχουν την επιλογή να υλοποιήσουν ένα ή περισσότερα από διάφορα πρότυπα. Όταν δύο components αρχίζουν να επικοινωνούν πρώτα ενημερώνονται αυτόματα ή με ανθρώπινη παρέμβαση μεταξύ τους για το τι πρότυπα έχουν υλοποιήσει και με ποια θα επικοινωνήσουν. Στα συστήματα ηλεκτρονικού εμπορίου ακολουθείται αυτή η λογική. Για παράδειγμα μπορεί να προτιμηθεί να υλοποιηθεί ως τρόπος πληρωμής σε συστήματα ηλεκτρονικού εμπορίου κάποιος από τους ήδη διαθέσιμους όπως η χρήση πιστωτικής κάρτας κάποιας εταιρίας. Στη διασύνδεση υπολογιστών το πρότυπο OSI είναι ένα παράδειγμα χρήσης οικογένειας προτύπων. Σε κάθε ένα από τα επτά επίπεδά του χρησιμοποιούνται / έχουν υλοποιηθεί μια οικογένεια από πρότυπα με αποτέλεσμα η κάθε διασύνδεση μεταξύ υπολογιστών να επιλέγει κάθε φορά κάποιο από τα πρότυπα σε κάθε επίπεδο για να πραγματοποιηθεί η επικοινωνία. Η προσέγγιση αυτή της χρήσης της οικογένειας προτύπων εξασφαλίζει ως ένα βαθμό αυτονομία στα components αλλά μπορεί όμως πολλές φορές λόγω μη υλοποίησης κάποιου προτύπου να μην είναι δυνατή η επικοινωνία.

Εξωτερική διαμεσολάβηση

Ένας τρόπος για να εξασφαλιστεί μεγάλος βαθμός αυτονομίας των components είναι να διαμεσολαβήσουν διαλειτουργικοί μηχανισμοί εκτός των components που συμμετέχουν στην επικοινωνία. Μια κύρια λειτουργία ενός τέτοιου μηχανισμού είναι η μετάφραση των μορφότυπων δεδομένων και των τρόπων αλληλεπίδρασης. Για παράδειγμα στα δίκτυα υπολογιστών οι πύλες (gateways) συνδέουν μεταξύ τους

ετερόκλητα δίκτυα. Η απεικόνιση καθολικών σχημάτων σε τοπικά σχήματα ανήκει επίσης σε αυτή την προσέγγιση.

Αυτοί οι διαλειτουργικοί μηχανισμοί, που συνήθως καλούνται 'wrappers' ή 'proxies', συνήθως επιτελούν πιο σύνθετες εργασίες από μια απλή μετάφραση. Για παράδειγμα αν υπάρχουν δύο συλλογές στις οποίες προσφέρονται υπηρεσίες αναζήτησης διαμέσου διαφορετικών ψηφιακών βιβλιοθηκών και μόνο η πρώτη παρέχει και υπηρεσία ταξινόησης αποτελεσμάτων (sorting) τότε αν ένας client αναζητήσει ταυτόχρονα και στις δύο συλλογές θα πρέπει ο εξωτερικός διαμεσολαβητής να προσφέρει την υπηρεσία ταξινόησης για τη δεύτερη συλλογή προτού τα αποτελέσματα φτάσουν στον client.

Η προσέγγιση της διαμεσολάβησης προσφέρει σημαντική διαλειτουργικότητα και ικανοποιεί τα κριτήρια της αυτονομίας, της ευκολίας της χρήσης και της κλιμάκωσης. Δεν απαιτείται ιδιαίτερη συμμόρφωση σε κοινούς κανόνες από τα components και ο χρήστης όταν χρησιμοποιεί το συνολικό σύστημα έχει της ψευδαίσθηση της χρήσης ενός ολοκληρωμένου συστήματος. Το πρόβλημα κυρίως που ανακύπτει με αυτή την προσέγγιση είναι της συνεισφοράς νέων components. Στην κατασκευή κάθε νέου διαφορετικού component πρέπει να κατασκευάζεται και ο αντίστοιχος διαμεσολαβητικός μηχανισμός. Γενικά αν πρέπει να διαλειτουργήσουν n διαφορετικά components με άλλα m διαφορετικά components τότε $n \times m$ διαφορετικοί διαμεσολαβητικοί μηχανισμοί πρέπει να ακολουθηθούν. Συνήθως η λύση που υιοθετείται είναι η χρήση ενός διαμεσολαβητή με ένα εσωτερικό πρότυπο στο οποίο θα γίνεται η μετατροπή της εισερχόμενης πληροφορίας από το ένα component και στη συνέχεια η μετατροπή της στη μορφή που καταλαβαίνει το άλλο component. Για παράδειγμα στην κατηγορία των σχημάτων κωδικοποίησης των βιβλιογραφικών δεδομένων το εσωτερικό σχήμα του διαμεσολαβητή θα μπορούσε να είναι το USMARC. Βέβαια όταν ένα πρότυπο δεν είναι αρκετό γιατί μπορεί να οδηγεί σε σημαντική απώλεια πληροφορίας κατά τη μετάφραση μπορεί να υιοθετείται συνδυαστικά η προσέγγιση της οικογένειας προτύπων και έτσι ο διαμεσολαβητής να χρησιμοποιεί περισσότερα του ενός εσωτερικά πρότυπα.

Ένα σημαντικό εργαλείο πλέον στη διαμεσολάβηση είναι η χρήση μεταδεδομένων για την περιγραφή αλλά και μετάφραση των αλληλοσυνεργαζόμενων μερών που είτε είναι υπηρεσίες, διαδικασίες, components ολόκληρα και κυρίως στοιχεία δεδομένων.

Αλληλεπίδραση βασιζόμενη στις προδιαγραφές

Σύμφωνα με αυτή την προσέγγιση τα components περιγράφουν δομικά και σημασιολογικά αρκετά αναλυτικά τις λειτουργίες και τα δεδομένα τους. Έτσι όταν θέλουν να διαλειτουργήσουν δεν χρειάζονται διαμεσολάβηση ή άλλη προηγούμενη διευθέτηση αλλά μόνα τους, διαθέτοντας μηχανισμούς ανίχνευσης της προδιαγραφής του άλλου component, θα επικοινωνούν. Αυτή η προσέγγιση δίνει μεγάλη αυτονομία στα components αφού διαχωρίζει τη λειτουργικότητα και την περιγραφή των δεδομένων από την υλοποίηση. Προσθέτει όμως πολυπλοκότητα και κάνει δύσκολη πολλές φορές την ακριβή περιγραφή των components. Τεχνολογίες και μοντέλα αναπαράστασης/διαχείρισης γνώσης και του σημασιολογικού ιστού κινούνται προς αυτή την κατεύθυνση όπως οι τεχνολογίες XML²⁷, RDF²⁸, OWL²⁹, Web Services³⁰

²⁷ <http://www.w3.org/XML/>

κλπ. Εξέχοντα ρόλο και σε αυτή την τεχνική έχουν τα μεταδεδομένα αφού είναι ο τρόπος περιγραφής δομής, διαδικασιών και δεδομένων.

Διαλειτουργικότητα σε επίπεδο τεκμηρίων

Στις παραδοσιακές βιβλιοθήκες το θέμα δεν τίθεται. Όμως στην εποχή μας οι παραδοσιακές βιβλιοθήκες αποτελούν είδος προς εξαφάνιση. Ακόμη και για τις ελληνικές δημόσιες βιβλιοθήκες “τρέχει” τη στιγμή που πληκτρολογούνται αυτές οι γραμμές έργο ψηφιοποίησης. Μέσα σε αυτό σε αυτό το πλαίσιο οι σύγχρονες βιβλιοθήκες -είτε είναι ψηφιακές είτε υβριδικές- διαθέτουν έναν διαρκώς αυξανόμενο όγκο ψηφιακών τεκμηρίων.

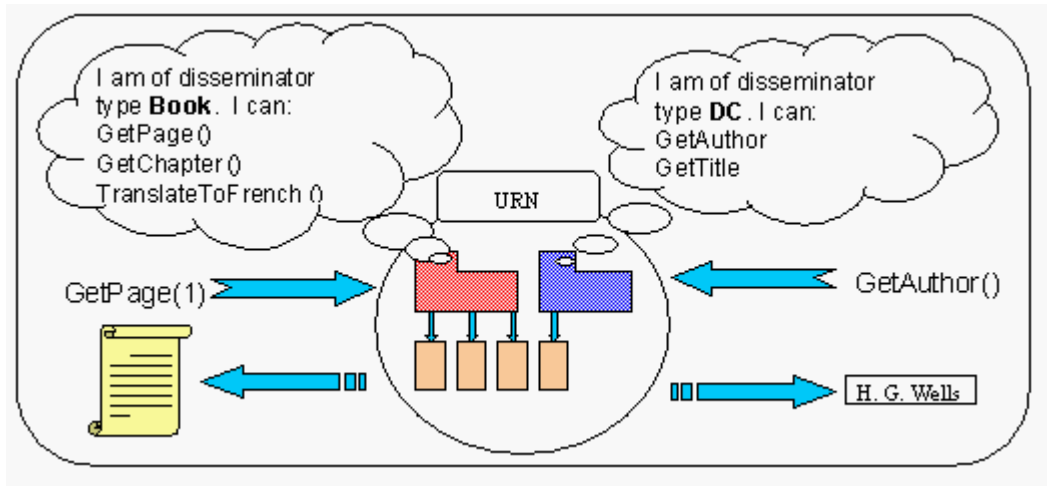
Στο σημείο αυτό τίθεται ένα πολύ σοβαρό θέμα για τη διαλειτουργικότητα αυτών των δεδομένων. Πως θα είναι δομημένα τα ψηφιακά τεκμήρια ώστε να διαλειτουργούν μεταξύ τους; Μια σειρά από οδηγίες έχουν εκδοθεί στο πλαίσιο διαφόρων έργων ψηφιοποίησης αλλά και οδηγίες για τεκμήρια που γεννιούνται ψηφιακά. Κεντρικός στόχος: να είναι σωστά δομημένα τα ψηφιακά τεκμήρια. Αυτό για να επιτευχθεί υψηλός βαθμός διαλειτουργικότητας η οποία άλλωστε συνδέεται από πολλούς με την ίδια τη διασφάλιση ποιότητας (βλέπε Kelly, 2004).

Στο χώρο των ψηφιακών βιβλιοθηκών έχουν προταθεί διάφορες αρχιτεκτονικές για την εξασφάλιση της διαλειτουργικότητας σε επίπεδο τεκμηρίου και συλλογής. Χαρακτηριστικό το παράδειγμα της συνεργασίας των Digital Library Research Group του Cornell University και του Corporation for National Research Initiatives (CNRI) για τη σχεδίαση μιας πλούσιας σε λειτουργικότητα και εύκολης σε ανάπτυξη αρχιτεκτονικής για διαλειτουργικά ψηφιακά τεκμήρια και αποθετήρια όπως περιγράφεται στο άρθρο των Payette, Sandra et all (1999). Σύμφωνα με αυτήν την αρχιτεκτονική το ψηφιακό τεκμήριο ορίζεται ως ένα αντικείμενο το οποίο φέρει την πληροφορία σε δυαδική μορφή αλλά που πάνω του μπορούν να επενεργήσουν διάφοροι μηχανισμοί για να εξορύξουν και να εμφανίσουν αυτήν την πληροφορία κατά το δοκούν ανάλογα το είδος του πραγματικού υλικού στο οποίο αντιστοιχεί (βιβλίο, περιοδικό, αναφορά, ηχητικό ντοκουμέντο κλπ). Αυτοί οι μηχανισμοί ονομάζονται ‘disseminators’ στα πλαίσια αυτής της αρχιτεκτονικής. Αυτή η τεχνική μπορεί να προσφέρει πολυπλοκότητα στο σύστημα κυρίως από πλευράς υλοποίησης αλλά κάνει το ψηφιακό τεκμήριο διαλειτουργικό αφού μπορεί να εμφανιστεί με όποια μορφή του ζητηθεί. Νέοι τρόποι εμφάνισης τους μπορούν να προστίθενται προσφέροντας επιπλέον λειτουργικότητα αλλά και διαλειτουργικότητα στο ευρύ ψηφιακό περιβάλλον αφού θα μπορεί να είναι προσβάσιμο και ορατό από διάφορους μηχανισμούς αναζήτησης / ανάκτησης. Στην παρακάτω εικόνα παρουσιάζεται σχηματικά αυτή η διαφορετικότητα στη λειτουργικότητα και συνεπώς διαλειτουργικότητα πάνω στο ίδιο ψηφιακό αντικείμενο.

²⁸ <http://www.w3.org/2001/sw/>

²⁹ <http://www.w3.org/2001/sw/>

³⁰ <http://www.w3.org/2002/ws/>



Πηγή: Payette, Sandra et all (1999)

Το ψηφιακό αντικείμενο αντιστοιχεί σε ένα βιβλίο και έχουν προστεθεί πάνω του δύο διαφορετικοί μηχανισμοί ανάγνωσης και παρουσίασης της πληροφορίας του. Ο ένας μηχανισμός (disseminator) το βλέπει ως βιβλίο και δίνει τρόπους ανάγνωσης της πληροφορίας του όπως GetPage, GetChapter ενώ ο άλλος μηχανισμός το βλέπει μεταδεδομένα και προσφέρει λειτουργίες όπως GetAuthor και GetTitle.

Συμπεράσματα

Στις 19 Μαΐου 2008 η εταιρεία Google Inc και ο οργανισμός OCLC υπέγραψαν συμφωνία για ανταλλαγή δεδομένων τα οποία θα διευκολύνουν την ανακάλυψη (discovery) των συλλογών των βιβλιοθηκών μέσα από τις υπηρεσίες αναζήτησης της Google³¹. Ήδη εδώ και αρκετό καιρό η υπηρεσία Google Books προσφέρει (μέσω ενός συνδέσμου με τίτλο «*Find this book in a library*») στο χρήστη τη δυνατότητα να εντοπίσει την πλησιέστερη σε αυτόν βιβλιοθήκη που διαθέτει το βιβλίο. Η Google ήξερε και που μένω (IP recognition) και ποια βιβλία έχει η βιβλιοθήκη δίπλα μου (ο εντοπισμός της βιβλιοθήκης και των βιβλίων της επιτυγχάνονταν μέσω του WorldCat). Τελικά μάλλον το ανέκδοτο με την πιτσαρία στην αρχή του παρόντος κειμένου δε φαντάζει και πολύ εξωπραγματικό, από τεχνικής τουλάχιστον άποψης.

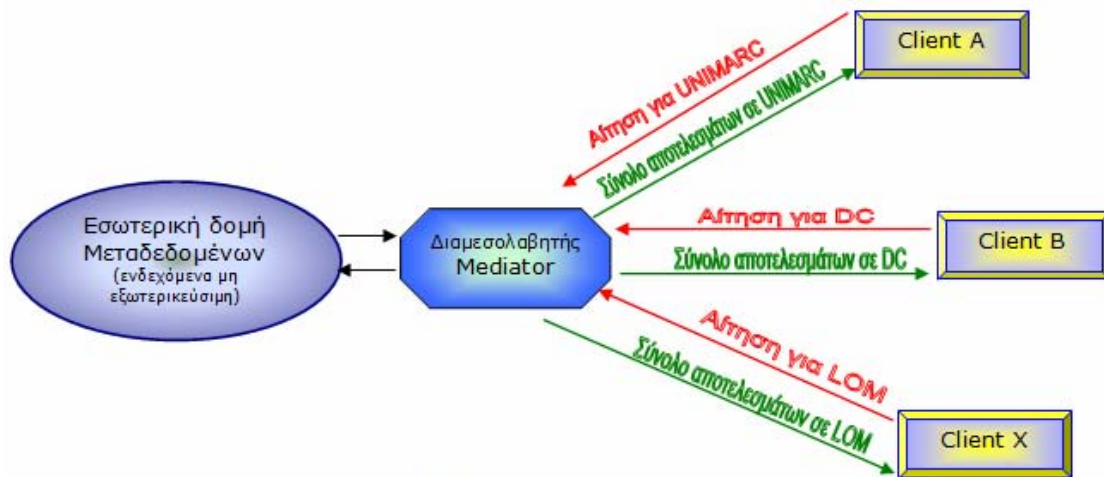
Η διαλειτουργικότητα δεν είναι διαπραγματεύσιμη στην εποχή μας. Είναι βασικός στόχος. Όσο αποτελεσματικότερα επιτυγχάνεται τόσο αποτελεσματικότερες υπηρεσίες παρέχονται. Όλα στηρίζονται στον κατακερματισμό και στη σύνθεση. Πολλές διαφορετικές οντότητες, πολλά διαφορετικά components και αμέτρητοι συνδυασμοί. Αλλιώς χρειάζεται να περιγραφούν τα βιβλία και αλλιώς τα περιοδικά για να επιτευχθεί μεγαλύτερη εκφραστικότητα στην περιγραφή, όμως χρειάζεται να αναζητηθούν και ταυτόχρονα. Αν σκοπός είναι η εξειδίκευση της αναζήτησης μόνο σε περιοδικά και η χρήση κριτηρίων που σχετίζονται μόνο με αυτά (π.χ. ποια εξακολουθούν να εκδίδονται, ή αυτά που βγαίνουν κάθε εβδομάδα) τότε σαφώς δεν έχει νόημα να τεθούν αυτά τα κριτήρια στα βιβλία. Αυτό σε καμία περίπτωση δεν σημαίνει πως αν το ζητούμενο είναι τεκμήρια που σχετίζονται με το X θέμα, θα πρέπει να ψάξει σε δύο καταλόγους διότι τα περιοδικά έχουν συχνότητα έκδοσης ενώ τα βιβλία δεν έχουν!

Ακριβώς σε αυτό τον άξονα των γενικών κατευθυντήριων οδηγιών κινείται η όλη προσπάθεια για δημιουργία εννοιολογικών μοντέλων (conceptual models). Η τάση είναι η συνεννόηση σε εννοιολογικά μοντέλα (για παράδειγμα FRBR, CIDOC) και όχι σε συγκεκριμένα σχήματα, τα σχήματα έπονται. Ο λόγος είναι ότι έτσι οριοθετείται το πλαίσιο μέσα στο οποίο νοηματοδοτούνται τα συμφραζόμενα και είναι ευκολότερη μια αποτελεσματικότερη επικοινωνία και κατ' επέκταση η διαλειτουργικότητα.

Στην παρούσα φάση και μάλλον στο άμεσο μέλλον φαίνεται ότι η βέλτιστη προσέγγιση είναι αυτή των διαμεσολαβητών. Ήδη χρησιμοποιούνται σε μεγάλο αριθμό περιπτώσεων και η χρήση τους φαίνεται αυξανόμενη³². Η προσέγγιση αυτή παρουσιάζεται σχηματικά στην επόμενη εικόνα.

³¹ Η ανακοίνωση αναρτήθηκε στον ιστότοπο του OCLC (<http://www.oclc.org/news/releases/200811.htm> τελευταία επίσκεψη 1/6/2008)

³² Για παράδειγμα το ψηφιακό αποθετήριο “**University of Southampton's Research Repository**” (<http://eprints.ecs.soton.ac.uk>) όπου δίνεται η δυνατότητα εξαγωγής των μεταδεδομένων σε: ASCII, EndNote, MODS, DC κτλ.

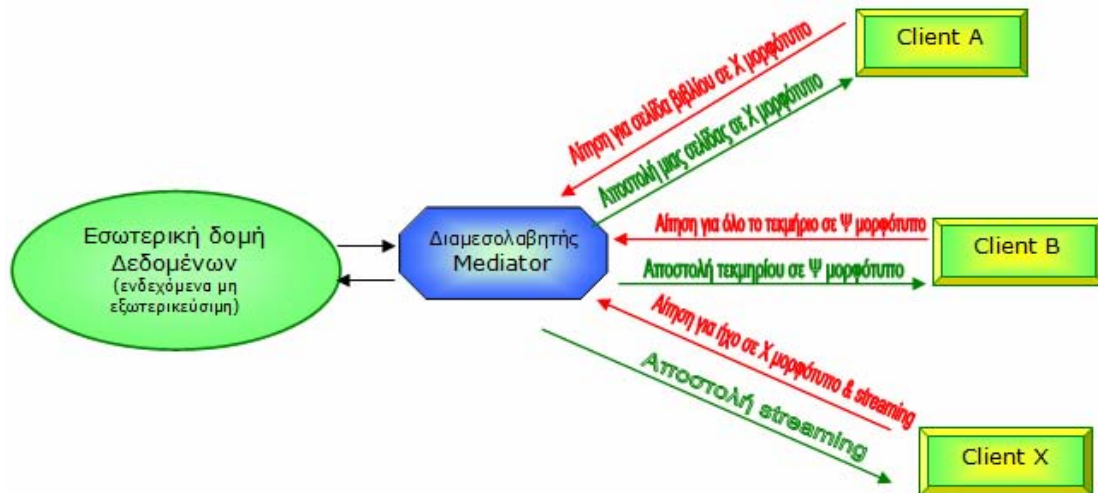


Σχήμα 1
Μεταδεδομένα

Βέβαια στο παραπάνω σχήμα ο client μπορεί να είναι ένα πρόγραμμα που χρησιμοποιεί απευθείας ο τελικός χρήστης αλλά μπορεί να είναι με τη σειρά του ένας server (agent) ο οποίος μαζεύει και επεξεργάζεται δεδομένα και από άλλες πηγές για να τα στείλει στη συνέχεια είτε σε τελικό χρήστη είτε σε άλλο client (agent) κλπ. Έτσι ο εκάστοτε χρήστης (τελικός ή ενδιάμεσος) δεν έχει καμιά εικόνα για την αρχική (ή εσωτερική) κωδικοποίηση των μεταδεδομένων αλλά έρχεται σε επαφή με την προτιμώμενη από αυτόν διάταξη-κωδικοποίηση. Όμως είναι πάντα δυνατή η μετατροπή; Προφανέστατα όχι. Πως είναι δυνατό να μετατραπεί μια πλήρης περιγραφή ενός Αρχείου (Fond) που έχει γίνει σε EAD σε ένα μια εγγραφή απλού DC; Αν γίνει θα διαθέτει ένα ελάχιστο μέρος της αρχικής πληροφορίας.

Το σχήμα θα μπορούσε να είναι πολύ πιο σύνθετο και ο διαμεσολαβητής(wrapper, mediator) να είναι μια σειρά από διαδικασίες (βλέπε και Haslhofer), να γίνεται χρήση οντολογιών για την ολοκλήρωση των μεταδεδομένων κλπ.

Ακριβώς η ίδια προσέγγιση μπορεί να ακολουθηθεί και για τα ίδια τα ψηφιακά αντικείμενα. Ο πελάτης να ζητάει μια σελίδα ενός βιβλίου ή ένα συγκεκριμένο κεφάλαιο ή την επωδό ενός τραγουδιού κτλ. Πρώτα από όλα θα πρέπει ο διαμεσολαβητής να μπορεί να καταλαβαίνει το αίτημα. Στη συνέχεια η εσωτερική δομή των δεδομένων θα πρέπει να επιτρέπει την απομόνωση της συγκεκριμένης ενότητας. Τέλος ο διαμεσολαβητής θα πρέπει να μπορεί να στείλει αυτή την ενότητα στην προτιμώμενη από τον πελάτη μορφή (μορφότυπο). Το παρακάτω σχήμα αναπαριστά αυτή τη λογική.



Σχήμα 2
Ψηφιακά Τεκμήρια

Για να επιτευχθούν τα παραπάνω βεβαίως και απαιτούνται πρότυπα. Όμως τα πρότυπα από τη φύση είναι γενικά και δεν επαρκούν. Ιδιαίτερη προσοχή πρέπει να δοθεί στις πολιτικές υλοποίησης διότι αυτές τελικά θα καθορίσουν και την αποτελεσματικότητα της διαλειτουργικότητας (Moen, 2000).

Η σημασιολογική διαλειτουργικότητα επιτρέπει στα διαφορετικά συστήματα να επικοινωνήσουν μεταξύ τους χωρίς να χάνεται η πληροφορία η οποία αρχικώς δόθηκε από τους δημιουργούς των δεδομένων. Όμως για να επιτευχθεί αυτό είναι προϋπόθεση να μπορέσει να υπάρξει προηγουμένως συντακτική διαλειτουργικότητα. Χωρίς συντακτική διαλειτουργικότητα τα δεδομένα και οι πληροφορίες δεν μπορούν να διαχειριστούν προβλεπόμενα με τις σωστές αναγωγές στο δικό τους format, κωδικοποίηση κτλ. Από την άλλη χωρίς σημασιολογική διαλειτουργικότητα το νόημα της χρησιμοποιούμενης γλώσσας δεν μπορεί να γίνει αρκούντως κατανοητό ή να επικοινωνήσει σωστά (DELOS 2005, σ. 8-9). Ο Doerr (Doerr, 2005) ορίζει τρία επίπεδα για τη σημασιολογική διαλειτουργικότητα.

- **Δομές δεδομένων:** είτε για μεταδεδομένα, είτε για δεδομένα περιεχομένου (content data), είτε δεδομένα διαχείρισης συλλογής, είτε περιγραφές υπηρεσιών
- **Δεδομένα κατηγοριοποίησης:** π.χ. ταξινόμηση, γενικά θέματα
- **Πραγματικά δεδομένα:** δηλ. στοιχεία που αναφέρονται στις λεπτομέρειες, όπως οι άνθρωποι, στοιχεία, θέσεις

Είναι γεγονός πως ένα από τα βασικά ζητούμενα της εποχής μας είναι η σημασιολογική διαλειτουργικότητα. Η οποία μάλιστα συνδέεται με όλο τον κύκλο ζωής της πληροφορίας από τη δημιουργία της έως και τη διατήρηση της (DELOS 2005, σ.13). Νέες γλώσσες και νέοι τρόποι έρχονται να προστεθούν ευελπιστώντας να καλυφθεί το κενό που υπάρχει. Πολλές προσπάθειες γίνονται προς αυτή την κατεύθυνση. Η πλατφόρμα Protégé δημιουργήθηκε από το Ερευνητικό Κέντρο Βιοϊατρικής Πληροφορικής του Πανεπιστημίου του Stanford για την ανάπτυξη οντολογιών έγινε αλλά με την υποστήριξη της National Library of Medicine των

ΗΠΑ. Πρόσφατα η Βιβλιοθήκη του Κογκρέσου ανακοίνωσε τη μετατροπή των γνωστών και ευρέως χρησιμοποιούμενων θεματικών επικεφαλίδων σε SKOS³³.

Σημσιολογικός ιστός προ των πυλών. Οι βιβλιοθήκες σε διεθνές επίπεδο παίζουν έναν ρόλο άλλοτε περισσότερο και άλλοτε λιγότερο καθοριστικό. Την ίδια στιγμή εντός των συνόρων η κατάσταση δεν είναι ιδιαίτερα παρήγορη. Η διαλειτουργικότητα δεν ήταν ποτέ σε υψηλά επίπεδα (και αυτό το δηλώνουν οι υπογράφωντες ως βασικοί συντελεστές δημιουργίας του περιβάλλοντος πρόσβασης «ΑΡΓΩ³⁴», βλέπε και Πεπονάκης, 2004) και η κατάσταση δε φαίνεται να οδεύει προς το καλύτερο. Σε μια πρόσφατη έρευνα³⁵ με μια πρόχειρη ανάλυση των δεδομένων φαίνεται ότι η ανάπτυξη των ηλεκτρονικών αποθετηρίων από τις ελληνικές ακαδημαϊκές βιβλιοθήκες οδηγεί στη δημιουργία ενός νέου καταλόγου εντός της βιβλιοθήκης ο οποίος στην πλειοψηφία των περιπτώσεων δεν επικοινωνεί (διαλειτουργεί) με τον κατάλογο των βιβλίων. Με άλλα λόγια στις ελληνικές βιβλιοθήκες στην εποχή της διαλειτουργικότητας παρατηρείται: ένας κατάλογος για τα βιβλία, άλλος για το ηλεκτρονικό αποθετήριο, άλλος για τα περιοδικά, άλλος για τα άρθρα των περιοδικών από τον Elsevier ή τον ψ εκδότη κοκ... Όχι ο εντοπισμός του τι υπάρχει εντός των συνόρων είναι ανέφικτος αλλά και η αναζήτηση στη συλλογή μιας και μοναδικής βιβλιοθήκης καταντάει πραγματική περιπέτεια.

³³ Στο σχετικό διαδικτυακό τόπο (<http://lcsb.info/>) αναφέρει πως πρόκειται για 2,441,494 triples και δίνει τη δυνατότητα download όλου του συνόλου των δεδομένων.

³⁴ <http://argo.ekt.gr>

³⁵ Πρόκειται να δημοσιευθεί στο προσεχές συνέδριο ακαδημαϊκών βιβλιοθηκών που θα λάβει χώρα στα Ιωάννινα

Βιβλιογραφία

- Bahr, Ellen (2007). “Dreaming of a Better ILS”. *Computers in Libraries* No.9, pp. 10-14.
- Bainbridge, David, et. al. (2006). “Document Level Interoperability for Collection Creators”. *Proceedings of the 6th ACM/IEEE-CS joint conference on Digital libraries, Chapel Hill, NC, USA*, pp. 105-106. Διαθέσιμο στο <http://doi.acm.org/10.1145/1141753.1141773>.
- Chan, Lois Mai (2005). “Metadata Interoperability: A Study of Methodology”. Διατίθεται στο διαδίκτυο <http://white-clouds.com/iclc/cliej/cl19chan.htm> (τελευταία επίσκεψη 2/6/2008).
- Chan, Lois Mai and Zeng, Marcia Lei (2006). “Metadata Interoperability: A Study of Methodology part I”. *D-Lib magazine* Vol.12, No.6. Διατίθεται στο διαδίκτυο <http://www.dlib.org/dlib/june06/chan/06chan.html> (τελευταία επίσκεψη 2/6/2008).
- Chan, Lois Mai and Zeng, Marcia Lei (2006). “Metadata Interoperability: A Study of Methodology part II”. *D-Lib magazine* Vol.12, No.6. Διατίθεται στο διαδίκτυο <http://www.dlib.org/dlib/june06/zeng/06zeng.html> (τελευταία επίσκεψη 2/6/2008).
- Coyle, Karen (2005). “Libraries and Standards”. *The Journal of Academic Librarianship* Vol.31, No.4, pp. 373–376.
- DELOS: A Network of Excellence on Digital Libraries (2005). “D5.3.1: Semantic Interoperability in Digital Library Systems”. Διατίθεται στο διαδίκτυο <http://delos-wp5.ukoln.ac.uk/project-outcomes/SI-in-DLs/SI-in-DLs.pdf> (τελευταία επίσκεψη 2/6/2008).
- Doerr, Martin (2005). “Knowledge Extraction and Semantic Interoperability” Διατίθεται στο διαδίκτυο http://www.delos.info/index.php?Itemid=333&id=420&option=com_content&task=view (τελευταία επίσκεψη 2/6/2008).
- Ebenezer, Catherine (2003). “Trends in integrated library systems”. *VINE* Vol.32, No.4, pp. 19-45.
- Haslhofer, Bernhard (2006). “A Service Oriented Approach for Integrating Metadata from Heterogeneous Digital Libraries”. Διατίθεται στο διαδίκτυο http://www.cs.univie.ac.at/upload//550/papers/iiWAS2006_bhaslhofer_cr_20061015.pdf (τελευταία επίσκεψη 2/6/2008).
- Helly, John J. et. al. (1999). “A method for interoperable digital libraries and data repositories”. *Future Generation Computer Systems* No.16, pp. 21–28.
- Kelly, Brian et. al. (2004). “Interoperable Digital Library Programmes? We Must Have QA!”. *ECDL LNCS* Vol. 3232, pp. 80-85.
- Kelly, Brian et. al. (2005). “Gateway Standardization: A Quality Assurance Framework for Metadata”. *Library Trends* Vol.53, No.4, pp. 637–650.

- Moen, William E. (2000). "Interoperability for Information Access: Technical Standards and Policy Considerations". *The Journal of Academic Librarianship* Vol.26, No.2, pp. 129–132.
- Paepcke, Andreas et. al. (1998). "Interoperability for Digital Libraries Worldwide" *Communications of the ACM* Vol.41, No.4, pp. 33-43.
- Payette, Sandra et. al. (1999). "Interoperability for Digital Objects and Repositories: The Cornell/CNRI Experiments" *D-Lib Magazine* Vol.5, No.5. Διατίθεται στο διαδίκτυο <http://dlib.org/dlib/may99/payette/05payette.html> (τελευταία επίσκεψη 2/6/2008).
- Rogers, Philip (2005). "Metadata Interoperability: Literature Review". Διατίθεται στο διαδίκτυο http://www.ils.unc.edu/~gerogers/papers/rogers_litreview_210.doc (τελευταία επίσκεψη 2/6/2008).
- Shepherd, Michael (2003). "Interoperability for Digital Libraries". Στο *DRTC Workshop on Semantic Web*. Διατίθεται στο διαδίκτυο http://drtc.isibang.ac.in:8080/bitstream/1849/95/2/C_%20shepherd_interoperability.pdf (τελευταία επίσκεψη 2/6/2008).
- Stvilia, Besiki et. al. (2004). "Metadata Quality for Federated Collections". Στο *ICIQ04 - 9th International Conference on Information Quality*. Διατίθεται το pre-print στο http://www.ideals.uiuc.edu/bitstream/2142/721/2/iciq_144_final_v1.pdf (τελευταία επίσκεψη 2/6/2008).
- Suleman, Hussein and Fox, Edward A. (2001). "A Framework for Building Open Digital Libraries". *D-Lib Magazine* Vol.7, No.12. Διατίθεται στο διαδίκτυο <http://www.dlib.org/dlib/december01/suleman/12suleman.html> (τελευταία επίσκεψη 2/6/2008).
- Η-Επιχειρείν. "Web services". Διατίθεται στο διαδίκτυο http://www.go-online.gr/ebusiness/specials/article.html?article_id=214 (τελευταία επίσκεψη 2/6/2008).
- Wilson, Andrew (2005). "Strategies for managing digital content formats". *Library Review* Vol.54, No. 9, pp. 508-513.
- Πεπονάκης, Μανόλης; Τρούτπεγλη, Ντίνα και Σφακάκης, Μιχάλης (2004). "Προβλήματα διαλειτουργικότητας κατά την ταυτόχρονη πρόσβαση σε πηγές μέσω του πρωτοκόλλου Z39.50 και το περιβάλλον πρόσβασης *Η Αργώ*". *Πρακτικά 13^{ου} Πανελληνίου Συνεδρίου Ακαδημαϊκών Βιβλιοθηκών*, 10-13 Οκτωβρίου 2004. Διαθέσιμο στο http://www.ionio.gr/libconf/pdfs/PeponakisZ3950_Interoperability.pdf (τελευταία επίσκεψη 2/6/2008).