

Κεφάλαιο 4

Θεματική επεξεργασία: θεωρία και πράξη

Η ανάγκη προσδιορισμού του νοηματικού περιεχομένου ενός τεκμηρίου και η απόδοσή του με συνοπτικό τρόπο αποτελεί τη βασική αρχή ανάπτυξης όλων των συστημάτων θεματικής πρόσβασης. Ουσιαστικά όλα τα συστήματα αποσκοπούν στην απόδοση εννοιών με συγκεκριμένο τρόπο κάθε φορά και την εξασφάλιση της πρόσβασης στο περιεχόμενο ενός τεκμηρίου με βάση το θέμα του. Πολύ απλά, είναι σα να δίνουμε ένα σαφή νοηματικό χαρακτηρισμό «μια πινακίδα» θεματική που μας λέει τι θέμα πραγματεύεται ένα βιβλίο, (ή ένα άρθρο, μια ιστοσελίδα, ένα μουσειακό αντικείμενο) ή οτιδήποτε και αν είναι το τεκμήριο που έχουμε μπροστά μας.

Θεματική ευρετηρίαση και ταξινόμηση

Ακολουθώντας τώρα το φιλοσοφικό ερώτημα αν «η κότα γέννησε το αυγό ή το αυγό την κότα» εξετάζουμε αν δίνουμε «πινακίδες» που εκφράζουν θέματα για να περιγράψουμε το νοηματικό περιεχόμενο ενός βιβλίου ή αν τοποθετούμε το βιβλίο κάτω από ήδη έτοιμες «πινακίδες» που εκφράζουν θέματα που ανήκουν σε συγκεκριμένες ευρύτερες κατηγορίες.

Αν θεωρήσουμε ότι η κότα γέννησε το αυγό και κάθε αυγό αντιστοιχεί με νέες «πινακίδες» που δημιουργούμε για να περιγράψουμε το θέμα ενός τεκμηρίου τότε ουσιαστικά δημιουργούμε θεματική πρόσβαση.



Δημιουργία Θεματικής Πρόσβασης ή πιο γνωστή ως θεματική επεξεργασία

Αν θεωρήσουμε ότι τα αυγά ήδη προϋπήρχαν – όπως οι ήδη έτοιμες «πινακίδες» μας με τα θέματα- τότε πρέπει να εντάξουμε κάθε κότα που θα προκύψει κάτω από μια τέτοια κατηγορία.



Η διαδικασία αυτή είναι γνωστή ως ταξινόμηση

Τι είναι λοιπόν θεματική επεξεργασία και τι είναι ταξινόμηση:

Οι διαδικασίες είναι πολύ κοντινές και στην καθημερινότητα της βιβλιοθηκονομικής εργασίας σπάνια κάνουμε το διαχωρισμό. Εδώ αναπτύσσονται στο πλαίσιο της κατανόησης της λειτουργίας των θεμάτων και της χρήσης τους.

Θεματική επεξεργασία είναι το να εντοπίσω όλα τα θέματα που περιέχει ένα τεκμήριο και να προσδιορίσω το κυρίως θέμα του. Παίρνοντας ως οδηγό το κυρίως θέμα θα το εντάξω σε μια θεματική κατηγορία χρησιμοποιώντας ένα ταξινομικό σύστημα. Αυτή η τελευταία διαδικασία είναι η ταξινόμηση. Ποιο αναλυτικά, η θεματική επεξεργασία περιλαμβάνει:

Τον προσδιορισμό των εννοιών ενός τεκμηρίου

Τον προσδιορισμό των σημασιολογικών σχέσεων μεταξύ των θεμάτων ενός τεκμηρίου

Τον προσδιορισμό της κεντρικής έννοιας του τεκμηρίου και σε ποιες κατηγορίες εντάσσεται.

Η ταξινόμηση περιλαμβάνει:

τον προσδιορισμό των θεματικών κατηγοριών

τον προσδιορισμό των σχέσεων μεταξύ θεματικών κατηγοριών και επομένως την δημιουργία ή περαιτέρω ανάπτυξη ταξινομικών συστημάτων

την ένταξη τεκμηρίων σε μια κατηγορία ενός ταξινομικού συστήματος και την απόδοση των κωδικών ώστε να τοποθετηθούν όλα μαζί.

Ταξινόμηση και θεματική ευρετηρίαση

Συνοψίζοντας λοιπόν βλέπουμε ότι

Θεματική επεξεργασία	Ταξινόμηση
Η θεματική επεξεργασία ξεκινάει με τον προσδιορισμό της εννοιολογικής κατηγορίας στην οποία εντάσσεται το τεκμήριο	Η Ταξινόμηση στηρίζεται στο θέμα του τεκμηρίου και το εντάσσει σε μια εννοιολογική κατηγορία
Η θεματική επεξεργασία προχωρά στο προσδιορισμό δευτερευουσών θεματολογικών κατηγοριών και ειδικότερων εννοιών του περιεχομένου	Η ταξινόμηση στη συνέχεια προσδιορίζει τις υποκατηγορίες στις οποίες εντάσσεται το τεκμήριο και ολοκληρώνεται με τη διαδικασία της απόδοσης των ταξινομικών αριθμών (κωδικών)

Γλώσσα, επικοινωνία και έκφραση

Και οι δυο διαδικασίες της θεματικής επεξεργασίας και της ταξινόμησης -που είναι αναπόσπαστα μέρη της θεματικής οργάνωσης των τεκμηρίων- συναντώνται στο σημείο των **λεξιλογίων**. Δηλαδή, και οι δύο διαδικασίες έχουν ανάγκη να προσδιορίζουν έννοιες που η μεν θεματική ευρετηρίαση τις εκφράζει με λέξεις, η δε ταξινόμηση με αριθμούς ή συνδυασμούς γραμμάτων και αριθμών.

Η γλώσσα αποτελεί μια «συνθήκη» μεταξύ των μελών μιας ομάδας ανθρώπων η οποία λέει ότι «συγκεκριμένες λέξεις εκφράζουν συγκεκριμένες έννοιες». Η γλώσσα όντας το κεντρικό μέσο επικοινωνίας του ανθρώπου και μία από τις πλέον πολυποίκιλες εκφάνσεις της ανθρώπινης δραστηριότητας μοιραία είναι όχι μόνο διαρκώς εξελισσόμενη και αναπτυσσόμενη αλλά στην κυριολεξία χαρτογραφεί την ανθρώπινη πορεία πάνω στη γη. Είναι δε τέτοια η ποικιλία του κόσμου που μας περιβάλλει τόσο των πραγμάτων όσο και των εννοιών που σχεδόν κάθε λέξη μπορεί να έχει δύο και τρεις ερμηνείες ή να σημαίνει ταυτόχρονα διαφορετικά πράγματα κάτω από διαφορετικές συνθήκες.

Για παράδειγμα η λέξη εργασία μπορεί να σημαίνει: δουλειά, επαγγελματική απασχόληση αλλά και έργο.

Έχει επίσης παρατηρηθεί ότι οι άνθρωποι παρότι εννοούν το ίδιο πράγμα ή γράφουν για το ίδιο πράγμα, δεν χρησιμοποιούν τις ίδιες λέξεις. Έτσι για παράδειγμα ένας γιατρός μπορεί να γράφει για «αρτηριακή πίεση», ένας άλλος για «υπέρταση και υπόταση» και ένας τρίτος απλά για «πίεση». Ακόμα περισσότερο οι άνθρωποι δεν είναι απαραίτητο να ψάξουν τις έννοιες όλοι με τις ίδιες ακριβώς λέξεις εφόσον υπάρχει η ποικιλία. Είναι εμφανές λοιπόν ότι ο έλεγχος του λεξιλογίου και κυρίως ο περιορισμός έκφρασης εννοιών με συγκεκριμένες λέξεις και όχι άλλες συνώνυμες τους φιλοδοξεί να δώσει λύση στο πρόβλημα αυτό. Αυτό δημιούργησε τα **ελεγχόμενα λεξιλόγια**. Δηλαδή, πρόκειται για λεξιλόγια που οι λέξεις που περιλαμβάνουν έχουν προσδιοριστεί ώστε να εκφράζουν συγκεκριμένες έννοιες, έχουν επιλεγεί από πριν και έχουν καθιερωθεί για να χρησιμοποιούνται με συγκεκριμένο τρόπο από τους δημιουργούς του λεξιλογίου. Είναι αντίθετα προς τη χρήση της ελεύθερης φυσικής γλώσσας που δεν γνωρίζει περιορισμούς στη χρήση των λέξεων.

Στη βιβλιοθηκονομία τα ελεγχόμενα λεξιλόγια περιλαμβάνουν λέξεις ή συνδυασμούς λέξεων που συγκροτούν αυτοτελείς έννοιες και που αποδιδόμενες σε τεκμήρια τα προσδιορίζουν και λειτουργούν ως σημεία θεματικής πρόσβασης σε αυτά.

Πρέπει επίσης να επισημάνουμε ότι τα ελεγχόμενα λεξιλόγια δίνουν λύσεις σε μια σειρά από προβλήματα που αφορούν τη χρήση των συνωνύμων, των αμφίσημων και πολύσημων (πολυσήμαντων) λέξεων καθώς και την καθιέρωση ή μη λέξεων μέσω παραπομπών. Στόχος των ελεγχόμενων λεξιλογίων είναι η τυποποίηση στη

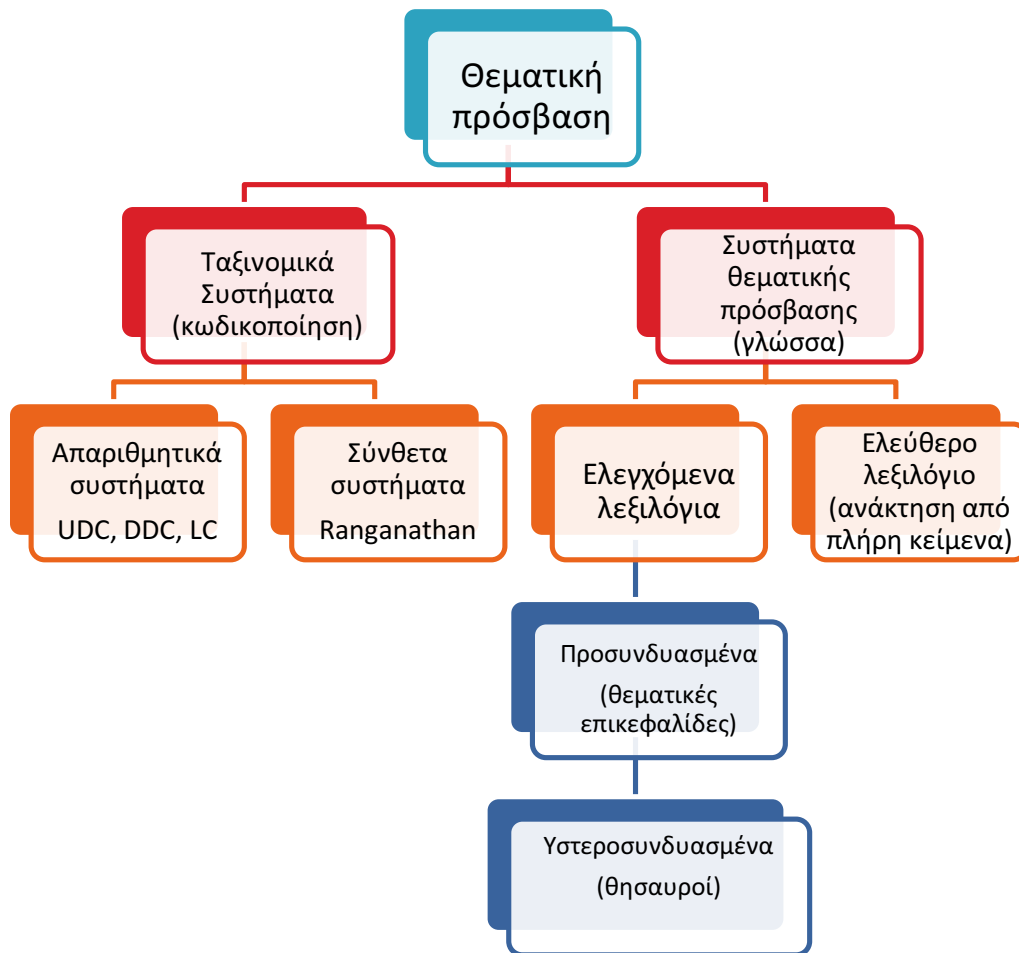
χρήση των λέξεων για την απόδοση εννοιών. Αν λάβουμε υπόψη μας ότι η εξέλιξη της γλώσσας είναι δυναμική και διαρκής, η τυποποίηση από μόνη της είναι αρκετή για να δημιουργήσει ένα τεχνητό πλαίσιο, που διαρκώς επιζητεί βελτίωση, ενημέρωση και συντήρηση.

Η ανάπτυξη διαφορετικών τρόπων συγκρότησης και δόμησης των ελεγχόμενων λεξιλογίων συνιστά και διαφορετικά συστήματα θεματικής πρόσβασης που θα μας απασχολήσουν παρακάτω σε επόμενα κεφάλαια. Η ανάπτυξη διαφορετικών τρόπων οργάνωσης των κατηγοριών και κωδικοποίησης των κατηγοριών αυτών και των υποδιαίρέσεών τους δημιουργεί και διαφορετικά ταξινομικά συστήματα που θα σας απασχολήσουν σε επόμενο μάθημα.

Το μόνο σύστημα που διαφέρει για τον απλούστατο λόγο ότι δεν στηρίζεται σε ελεγχόμενο λεξιλόγιο είναι η χρήση του ελεύθερου λεξιλογίου. Στον αντίποδα λοιπόν, του ελεγχόμενου λεξιλογίου βρίσκεται το ελεύθερο λεξιλόγιο που δηλώνει αυτό ακριβώς που λέει ο όρος, δηλαδή ότι δεν υπόκειται σε κανένα περιορισμό και δεν υπάρχει έλεγχος στο τι εντάσσεται και τι όχι. Είναι δηλαδή η χρήση της γλώσσας όπως την ξέρουμε στη γραπτή ή και την προφορική μορφή της ανάλογα με τα τεκμήρια που επεξεργαζόμαστε.

Με τη χρήση του ελεύθερου λεξιλογίου ουσιαστικά αναφερόμαστε στην πρόσβαση που επιτυγχάνεται με αναζήτηση μέσα στα πλήρη κείμενα των τεκμηρίων και η οποία επιτρέπει την ανάκτηση δεδομένων αδιακρίτως χρήσης λέξεων απευθείας από το κείμενο, όπως το έχει γράψει ο συγγραφέας. Τα συστήματα αυτά δεν είναι δομημένα ή παρουσιάζουν μερική μόνο δόμηση μέσα από τεχνικές “tagging” που θα αναλυθούν στο τελευταίο κεφάλαιο. Τα συστήματα αυτά μπορεί να συνοδεύονται από ταξινομικά συστήματα που εντάσσουν τις λέξεις κλειδιά σε ευρείες θεματικές κατηγορίες, αλλά μπορεί και να μη συνοδεύονται από διαδικασίες ταξινόμησης, απλά να υπάρχουν για χρήση μέσα από ισχυρές μηχανές αναζήτησης, με κυριότερο παράδειγμα τη Google.

Το σχήμα που ακολουθεί δείχνει τη θεματική πρόσβαση και τη σχέση της με τα ελεγχόμενα λεξιλόγια, τα ταξινομικά συστήματα και την ελεύθερη ανάκτηση από πλήρη κείμενα.



Σχήμα 2. Η χαρτογράφηση της θεματικής πρόσβασης

Επεξηγήσεις στο σχήμα 2:

Απαριθμητικά συστήματα (enumerative systems) είναι αυτά που παρουσιάζουν πίνακες και λίστες με όλες τις συγκεκριμένες θεματικές κατηγορίες που περιλαμβάνουν. Απαριθμούν δηλαδή τις θεματικές κατηγορίες και υποκατηγορίες των εννοιών.

Σύνθετα συστήματα: επίσης γνωστά και ως αναλυτικοσυνθετικά συστήματα και ως συστήματα «απόψεων» (faceted systems, analytic synthetic systems)²⁵. Είναι αυτά που στηρίζονται σε δύο διαδικασίες, την ανάλυση και τη σύνθεση. Η ανάλυση συνίσταται στην αποδόμηση του νοηματικού περιεχομένου του τεκμηρίου σε όλα του τα επιμέρους θέματα. Ακολουθεί η σύνθεση η οποία στοχεύει στην επανένωση των συστατικών

²⁵ http://www.iva.dk/bh/lifeboat_ko/CONCEPTS/facet_and_facet_analysis.htm Ανακτήθηκε 4/10/2013.

αυτών με την εξέτασή τους μέσα από τις ακόλουθες οπτικές γωνίες (απόψεις): α) Προσωπικότητα (αναφέρεται στο ιδιαίτερο χαρακτηριστικό ενός θέματος), β) Ύλη (το φυσικό υλικό στο οποίο αναφέρεται το θέμα, γ) Ενέργεια (κάθε κίνηση που γίνεται σε σχέση πάντα με το θέμα) δ) τόπος (η γεωγραφική περιοχή που αναφέρεται το θέμα) ε) Χρονική περίοδος που σχετίζεται με το θέμα.

Πρέπει να τονίσουμε ότι παρά το γεγονός ότι τα αναλυτικοσυνθετικά συστήματα θα μπορούσαν να αποδώσουν πολύ περισσότερο αξιοποιώντας τις τεχνολογικές εξελίξεις, τα απεριθμητικά είναι τα συστήματα εκείνα που χρησιμοποιούν οι βιβλιοθήκες σήμερα και κυρίως το DDC (Dewey Decimal Classification System).

Προσυνδυασμένα συστήματα είναι αυτά που οι θεματικοί όροι παρατίθενται έτοιμοι (είναι δηλαδή συνδυασμένοι από πριν, προσυνδυάζονται) και εκφράζουν έννοιες υποδιαίρεσεις εννοιών, γεωγραφικά ονόματα, ιστορικές περιόδους και μορφές παρατάσσονται με τη σειρά που προαναφέρθηκε και σχηματίζουν μια θεματική αλυσίδα. Οι θεματικές επικεφαλίδες είναι το κυριότερο προσυνδυασμένο σύστημα. Ο βιβλιοθηκονόμος λαμβάνει έτοιμα τα στοιχεία των θεματικών επικεφαλίδων μέσα από ένα αυστηρά ελεγχόμενο λεξιλόγιο.

Έτσι για παράδειγμα, αν έχω ένα παιδικό βιβλίο για τη νεότερη ιστορία της Ελλάδας η θεματική μου επικεφαλίδα θα ήταν:

Ιστορία, Νεότερη – Ελλάδα – 19^{ος} αι – Βιβλία για παιδιά

Υστεροσυνδυασμένα συστήματα είναι αυτά στα οποία ο βιβλιοθηκονόμος καλείται να συνδυάσει επί τόπου τους όρους που λαμβάνει από ένα ελεγχόμενο λεξιλόγιο ανάλογα με τις ανάγκες του τεκμηρίου. Στα υστεροσυνδυασμένα συστήματα ο βιβλιοθηκονόμος μπορεί να συνδυάσει δύο όρους που βρίσκονται ανεξάρτητοι στο ελεγχόμενο λεξιλόγιο και να δημιουργήσει νέους συνδυασμούς και νέες υποδιαίρεσεις των εννοιών. Τα υστεροσυνδυασμένα συστήματα δεν παρουσιάζουν τα θέματα σε αλυσίδες που περιγράφουν έννοιες, υποδιαίρεσεις των εννοιών, γεωγραφικά ονόματα και χρονικές περιόδους όπως τα προσυνδυασμένα. Αντίθετα παρουσιάζουν συνδυασμούς και συσχετισμούς των εννοιών που δημιουργούν εξειδικεύσεις. Οι Θησαυροί είναι το κυριότερο υστεροσυνδυασμένο σύστημα.

Πρέπει να σημειωθεί ότι με την πάροδο των ετών και τη διαρκή αναβάθμιση των θεματικών επικεφαλίδων, οι διαφορές ανάμεσα στα προσυνδυασμένα και τα

υστεροσυνδυασμένα συστήματα έχουν ελαχιστοποιηθεί. Η διάκριση δε, τείνει να μην αναφέρεται πλέον στη βιβλιογραφία. Για το λόγο αυτό θα εξεταστούν παρακάτω ως αυτόνομα συστήματα οι θεματικές επικεφαλίδες και οι θησαυροί χωρίς να γίνεται ιδιαίτερη μνεία στο αν είναι προσυνδυασμένα ή υστεροσυνδυασμένα.

Τα συστήματα θεματικής πρόσβασης αναλύονται παρακάτω σε ειδικό κεφάλαιο.

Ελεύθερο λεξιλόγιο (απευθείας ανάκτηση από πλήρη κείμενα). Πρόκειται για ελεύθερο λεξιλόγιο, που μπορεί να υπάγεται σε κάποιο σύστημα, συνήθως ταξινομικό, κάτω από το οποίο εντάσσονται οι λέξεις κλειδιά που προκύπτουν, ή και για τελείως ελεύθερο λεξιλόγιο που δεν εντάσσεται σε κανενός είδους σύστημα. Το ελεύθερο λεξιλόγιο έχει ένα συγκριτικό πλεονέκτημα, ότι η γλώσσα του είναι φυσική, δεν υπάγεται σε περιορισμούς και δεν απαιτεί επεξηγήσεις. Χρησιμοποιείται από όλους και δε χρειάζεται κάποιον να μεσολαβήσει και να ερμηνεύσει τυποποιήσεις και παραπομπές για τον απλούστατο λόγο ότι δεν υπάρχουν.

Λεξιλόγια και αναζήτηση πληροφοριών

Οι χρήστες των βιβλιοθηκών κάνουν αναζητήσεις με βάση το θέμα και στηρίζονται σε αυτές κυρίως για τον εντοπισμό τεκμηρίων που δεν ξέρουν ότι υπάρχουν, για την ερευνητική διαδικασία και για την ικανοποίηση απλών πληροφοριακών αναγκών που σχετίζονται με το «θέμα» που τους ενδιαφέρει. Η διαδικασία των βιβλιοθηκών να μετατρέπουν τις λέξεις που χρησιμοποίησαν οι συγγραφείς σε ένα τυποποιημένο λεξιλόγιο περιόρισε τις δυνατότητες αναζήτησης αλλά σίγουρα ομαδοποίησε έννοιες και θεματικές περιοχές. Το λεξιλόγιο αυτό συχνά είναι δύσκαμπτο και θέτει περιορισμούς στην έκφραση αλλά και στην ακριβή αποτύπωση νοημάτων.

Αυτό είχε ως αποτέλεσμα την αναζήτηση εννοιών που οι χρήστες νόμιζαν ότι θα βρουν αλλά δεν έβρισκαν, ή την αναζήτηση με συναφείς όρους που δεν έδιναν πάντα την ακριβή πληροφορία που ήθελαν. Παρατηρήθηκε το φαινόμενο όπου οι χρήστες αντί να επιλέγουν την αναζήτηση με βάση το θέμα επέλεξαν την αναζήτηση με βάση τις λέξεις του τίτλου ελπίζοντας να βρουν εκεί το νόημα που αναζητούσαν. Αυτό βέβαια, ως ένα βαθμό ήταν αποδοτικό, αλλά αχρήστευε ουσιαστικά τη θεματική επεξεργασία. Παράλληλα, αυτού του είδους οι αναζητήσεις απέδιδαν συνήθως υπερβολικά πολλά τεκμήρια που αποθάρρυναν τους χρήστες από κάθε είδους έρευνα.

Η επιτομή της αναζήτησης με βάση το θέμα είναι το διαδίκτυο, και όσα περιγράφηκαν παραπάνω είναι πολύ γνωστά στους χρήστες του Google και των άλλων μηχανών αναζήτησης που χρησιμοποιούν ελεύθερο λεξιλόγιο και αξιοποιούν τις νέες τεχνολογίες για να βελτιώσουν τα αποτελέσματα των αδόμητων δεδομένων.

Ελεύθερο λεξιλόγιο και Σημασιολογικός Ιστός (Semantic Web)

Το ελεύθερο λεξιλόγιο είναι η φυσική γλώσσα που επιτρέπει την τοποθέτηση ενός τεκμηρίου σε όποιο σημείο ορίζουν οι λέξεις που το απαρτίζουν και την ανάκτησή του από οποιοδήποτε σημείο παρέχουν πρόσβαση οι λέξεις που εμπεριέχονται σε αυτό. Ελεύθερο, γρήγορο και ελκυστικό όπως ακούγεται το ελεύθερο λεξιλόγιο έχει πολλά πλεονεκτήματα αλλά εξακολουθεί να έχει κάποια μειονεκτήματα.

Η απόλυτη ελευθερία του λεξιλογίου συχνά αποδίδει πληθώρα αποτελεσμάτων, άλλες φορές ελάχιστα, συχνά δεν υπάρχουν μηχανισμοί που να συνδυάζουν ταυτόσημες έννοιες, οπότε χάνεται πληροφορία που έχει εκφραστεί με διαφορετικό τρόπο ή λέξεις που γράφονται ή ακούγονται όμοιες σημαίνουν διαφορετικά πράγματα με βάση το ευρύτερο πλαίσιο που εμπεριέχεται στο τεκμήριο. Η ανάπτυξη των συστημάτων του σημασιολογικού ιστού (semantic web) προσπαθεί να καλύψει τις αδυναμίες αυτές και μέχρι στιγμής έχει αποδώσει εξαιρετικά αποτελέσματα. Η χρήση της φυσικής γλώσσας και του ελεύθερου λεξιλογίου δεν προϋποθέτει την ύπαρξη ενός ελεγχόμενου λεξιλογίου, ούτε τη δόμησή του με ιεραρχίες και συσχετισμούς. Δεν απαιτεί την καθιέρωση όρων και τη χρήση παραπομπών. Αξιοποιεί απλά το γραπτό λόγο και στηρίζεται στην αξιολόγηση κάποιων στοιχείων του. Τα κυριότερα από αυτά τα στοιχεία είναι:

- Ο προσδιορισμός προθέσεων και συνθετικών των λέξεων. Ξεχωρίζει δηλαδή το θέμα (θεματική ρίζα) των λέξεων από τα συνθετικά
- Ξεχωρίζει τα σημεία στίξης από τις λέξεις
- Αναλύει τη συντακτική δομή των προτάσεων για να δείξει το συσχετισμό των λέξεων (υποκείμενο, ρήμα, αντικείμενο)
- Ξεχωρίζει τη «γραμματική» αξία των λέξεων και προσδιορίζει την πληροφοριακή τους δυνατότητα. Προαναφέρθηκε ότι οι λέξεις που φέρουν πληροφορία είναι τα ουσιαστικά και τα επίθετα. Τα ρήματα δείχνουν ενέργεια

ή κατάσταση και έχουν δευτερεύουσα πληροφοριακή χρησιμότητα. Οι λοιπές γραμματικές κατηγορίες δεν έχουν πληροφοριακή αξία.

- Εξετάζει τη δομή των προτάσεων και του συσχετισμού τους.
- Κάνει σημασιολογική ανάλυση ώστε να καταλήγει κανείς στο νόημα ενός τεκμηρίου και όχι απλά στις λέξεις που εμφανίζονται σε αυτό.
- Χρησιμοποιεί ως βασική αρχή τους αυτοματοποιημένους τρόπους εντοπισμού των λέξεων που φέρουν πληροφορία και τη συστηματική αντιπαραβολή τους με ελεγχόμενα λεξιλόγια.

Ο D. Grant Cambell (2007)²⁶ αναπτύσσει τη θεωρία ότι το βασικότερο σημείο σύνδεσης μεταξύ του Σημασιολογικού Ιστού και των βιβλιοθηκονομικών εργασιών είναι ακριβώς οι τρόποι διαχείρισης της θεματικής πρόσβασης και αναφέρει ότι «ο Σημασιολογικός ιστός προσφέρει μια πυραμίδα προτύπων πληροφόρησης που επιτρέπουν στην μηχανή (υπολογιστή) να κατανοεί και να διαβάζει την πληροφορία». Αυτά τα πρότυπα, που αφορούν τη θεματική επεξεργασία είναι τα μετεξελιγμένα βιβλιοθηκονομικά πρότυπα που θα μας απασχολήσουν στις επόμενες εβδομάδες και είναι κυρίως οι θησαυροί και οι οντολογίες με τη μορφή που έχουν πλέον αποκτήσει και τη χρήση που έχουν στο σημασιολογικό ιστό. Επισημαίνουμε εδώ, ότι όπως όλα τα βιβλιοθηκονομικά πρότυπα εξελίσσονται προσαρμοζόμενα στις τεχνολογικές εξελίξεις, το ίδιο ισχύει και για τα πρότυπα που αφορούν τα θέματα. Για παράδειγμα οι Θησαυροί απέχουν πολύ από τις θεματικές επικεφαλίδες και οι Αγγλοαμερικάνικοι κανόνες καταλογογράφησης είναι πολύ μακριά από τα μεταδεδομένα που αποτυπώνουμε σε RDA.

Ανάλογα, πρότυπα κωδικοποίησης γίνονται πιο ευέλικτα και πιο εύκολα στη χρήση καθώς η τεχνολογία μπορεί να υποστηρίξει σύνθετες ενέργειες που στο παρελθόν γινόταν με τη βοήθεια των προτύπων. Αναφέρουμε χαρακτηριστικά ότι οι βιβλιοθήκες εξετάζουν και εφαρμόζουν τρόπους μετατροπής των MARC και Unimarc εγγραφών τους σε εγγραφές Dublin Core. Αυτό τους δίνει ευελιξία, διαλειτουργικότητα και ευκολία στη διαχείριση των θεμάτων. Ανάλογη είναι και η εφαρμογή προτύπων που αφορούν τα θέματα.

²⁶ D.G. Cambell, Part I: Semantic Web Foundations, Standards and Tools, “The Birth of the New Web: A Foucauldian Reading of the Semantic Web” *Cataloging & Classification Quarterly* Vol. 43, No. 3/4, 2007, pp. 9-20; Available online at <http://ccq.haworthpress.com> doi:10.1300/J104v43n03_02 9 (Ανακτήθηκε 26 Σεπτεμβρίου 2012).

Τα στοιχεία αυτά θα γίνουν κατανοητά παρακάτω όπου εξετάζονται τα πρότυπα που αφορούν τα θέματα και τα συστήματα που υπάρχουν σήμερα για τη διαχείριση των θεμάτων. Οι ανάγκες της θεματικής οργάνωσης της γνώσης είναι πιο εμφανείς στην ανάγκη υιοθέτησης ενός ταξινομικού συστήματος που θα μπορεί να εντάσσει το ελεύθερο λεξιλόγιο που χρησιμοποιούν οι δημιουργοί σε μεγάλες θεματικές κατηγορίες και θα μπορεί με βάση τη συνάφεια και το συσχετισμό των εννοιών, να οργανώνει τους πληροφοριακούς πόρους του διαδικτύου σε μεγάλες θεματικές ενότητες. **Οι οντολογίες είναι αυτή τη στιγμή, το πλησιέστερο τέτοιο σχήμα που αντλεί από το ελεύθερο λεξιλόγιο για να σχηματίσει ένα «ελεγχόμενο λεξιλόγιο» που δεν είναι τίποτα άλλο από μια λίστα λέξεων με μια αναγωγή της κάθε λέξης σε μια ή περισσότερες θεματικές κατηγορίες που προσομοιάζουν ένα ταξινομικό σύστημα.** Αυτό επιτρέπει την ανάπτυξη συσχετισμών μεταξύ των πληροφοριακών πόρων και την επίτευξη της συνάφειας που μέχρι στιγμής μπορούσε να δοθεί μόνο μέσα από τη χρήση των ελεγχόμενων λεξιλογίων.

Ελεγχόμενα λεξιλόγια

Με τον όρο ελεγχόμενο λεξιλόγιο εννοούμε «μια καθιερωμένη λίστα τυποποιημένης ορολογίας που χρησιμοποιείται στη θεματική ευρετηρίαση και την ανάκτηση της πληροφορίας. Παράδειγμα ελεγχόμενου λεξιλογίου αποτελούν οι θεματικές επικεφαλίδες ή οι όροι ενός θησαυρού που χρησιμοποιούνται για την θεματική περιγραφή του υλικού των βιβλιοθηκών. Το ελεγχόμενο λεξιλόγιο εγγυάται ότι ένα θέμα θα περιγράφεται πάντοτε με τον ίδιο προτιμώμενο όρο ή θεματική επικεφαλίδα κάθε φορά που θα ευρετηριάζεται. Αυτό στοχεύει στο να βρίσκει κανείς με ευκολία όλες τις πληροφορίες για το συγκεκριμένο θέμα κατά την αναζήτηση» .

Η δημιουργία καταλόγων ελεγχόμενου λεξιλογίου αποτέλεσε βασική προσπάθεια των βιβλιοθηκονόμων με στόχο την τυποποίηση της ορολογίας και τη διευκόλυνση τόσο της δουλειάς τους στην πλευρά της οργάνωσης όσο και των χρηστών στην πλευρά της αναζήτησης.

Για την απόδοση των θεμάτων που έχουν συμπεριληφθεί σε ελεγχόμενα λεξιλόγια χρησιμοποιούνται ειδικά πρότυπα πάνω στα οποία μπορούμε να δημιουργήσουμε σημεία για θεματικά δεδομένα. Τα πρότυπα αυτά αφορούν τη διαχείριση των θεμάτων

στις βιβλιογραφικές βάσεις και είναι διαφορετικά από τα πρότυπα των συστημάτων που αναπτύσσουμε για να αναπτύξουμε και να φτιάξουμε τους συσχετισμούς μεταξύ των θεμάτων. Ο παρακάτω πίνακας παρουσιάζει τα διαφορετικά αυτά πρότυπα:

Πρότυπα περιγραφής (και επομένως και περιγραφής θεμάτων)	Πρότυπα ανάπτυξης ελεγχόμενων λεξιλογίων (συστήματα)	Ταξινομικά συστήματα
Unimarc: Πεδία από το <600> έως το <661> (Πεδία που αφορούν το θέμα) (ειδικότερα 608, 610 και 615) Και <670> έως <686> πεδία που αφορούν την ταξινόμηση	LC Θεματικές επικεφαλίδες Όροι θησαυρών Λέξεις κλειδιά	UDC DDC LC Classification Scheme Άλλα ταξινομικά συστήματα
TEI Header (TEI): πεδία <Keywords><classcode>	Λέξεις κλειδιά (ελεύθερο λεξιλόγιο), όροι θησαυρών Θεματικές επικεφαλίδες	UDC DDC LC Classification Scheme Άλλα ταξινομικά συστήματα
Dublin Core: <Subject> (πεδίο Θέμα)	LC Θεματικές επικεφαλίδες Όρους Θησαυρών Όρους οντολογιών Όρους θεματικών χαρτογραφήσεων Ταξινομικούς αριθμούς Ελεύθερο λεξιλόγιο (λέξεις κλειδιά)	Οτιδήποτε

Πίνακας 3. Πρότυπα που χρησιμοποιούνται για τη θεματική οργάνωση

Βλέπουμε δηλαδή, ότι η αναγραφή των θεμάτων στις εγγραφές που κάνουμε σε ένα κατάλογο ακολουθούν το βασικό πρότυπο αναγραφής του καταλόγου, είτε αυτό είναι Unimarc, Marc, TEI ή DC. Στην περίπτωση δηλαδή που η βιβλιοθήκη χρησιμοποιεί Unimarc θα ακολουθήσουμε τις οδηγίες των πεδίων από το 6XX και μετά για να καταγράψουμε σε κάθε πεδίο το αντίστοιχο θεματικό στοιχείο (στήλη 1 του πίνακα 3). Το περιεχόμενο αυτών των στοιχείων θα προέλθει από τη δεύτερη στήλη του πίνακα 3 και θα έρχεται από κάποιο σύστημα θεματικής πρόσβασης (οργάνωσης) που χρησιμοποιεί η βιβλιοθήκη και μπορεί να είναι Θεματικές επικεφαλίδες ή κάποιος

θησαυρός. Με βάση το θέμα που θα έχουμε βρει θα κοιτάξουμε το ταξινομικό σύστημα που χρησιμοποιεί η βιβλιοθήκη το οποίο μπορεί να είναι ένα από τα αναφερόμενα στη στήλη 3 του πίνακα 3.

Παράδειγμα εγγραφής Unimarc από τη Βιβλιοθήκη του ΤΕΙ Αθήνας

LDR: 00639nam 2200205 4500

001: GR-AtTEI43655

008: 080123s gr gr 00010 gre d

020: \$a 9602181834

049: \$b 01

082: 0 \$a 025.17

100: 1 \$4 aut \$a Μπάγιας, Ανδρέας Φ.

245: 10 \$a Εγχειρίδιο αρχειονομίας : \$b η επεξεργασία ενός ιστορικού αρχείου / \$c Ανδρέας Μπάγιας.

260: \$a Αθήνα : \$b Κριτική, \$c 1999.

300: \$a 251 σ. ; \$c 20 εκ.

504: \$a Περιέχει βιβλιογραφία.

650: 10 \$a Αρχεία

650: 10 \$a Archives

650: 10 \$a Αρχειονομία

Θέματα

676: 025.17 ΜΠΑ

Ταξινομικός
αριθμός

Στην περίπτωση του προτύπου TEI, το οποίο-αναφέρεται εδώ γιατί αποτελεί μια περιεκτική μορφή του LC, διακρίνουμε τέσσερες βασικές ενότητες: την περιγραφή του αρχείου (file description) την περιγραφή της κωδικοποίησης (encoding description), την περιγραφή του περιγράμματος (profile description) και την περιγραφή της αναθεώρησης (revision description). Στην ενότητα της περιγραφής του περιγράμματος του τεκμηρίου περιλαμβάνονται τα δύο κομβικά στοιχεία που αφορούν το θέμα του τεκμηρίου, δηλαδή οι λέξεις κλειδιά και ο κωδικός της ταξινόμησης που απαρτίζουν ουσιαστικά το θεματικό περίγραμμα (profile) του τεκμηρίου. Επισημαίνουμε και πάλι ότι το τι θα γραφτεί μέσα σε αυτά τα δύο πεδία μπορεί να προέρχεται για μεν τα θέματα από ένα οποιοδήποτε σύστημα, συνήθως θεματικές επικεφαλίδες ή θησαυρούς αλλά και από ελεύθερο λεξιλόγιο δηλαδή «λέξεις κλειδιά» ενώ ο κωδικός ταξινόμησης (δηλαδή ο ταξινομικός αριθμός) μπορεί να προέρχεται από ένα οποιοδήποτε

ταξινόμικό σύστημα από αυτά που αναφέρονται στη στήλη 3 του πίνακα. Τα συνηθέστερα ταξινόμικά συστήματα που θα συναντήσουμε είναι το LCCS (Library of Congress Classification Scheme) και το DDC (Dewey Decimal Classification System).

Παράδειγμα εγγραφής σε TEI Header από τη Βιβλιοθήκη του Πανεπιστημίου του Michigan²⁷

```

<teiHeader>
  <fileDesc>
    <titleStmt>
      <title type="245">The Wade-Davis manifesto. A
last appeal to the democracy. By Caspar Butz.</title>
      <author> Butz, Caspar, 1825-1885.</author>
    </titleStmt>
    <extent>4 600 dpi G4 TIFF images</extent>
    <publicationStmt>
      <publisher>University of Michigan
Library</publisher>
      <pubPlace>Ann Arbor, Michigan</pubPlace>
      <date>2005</date>
      <idno type="dlps">ADG9061.0001.001</idno>
      <idno type="lccallno">E458.4 .B99</idno>
      <availability>
        <p>These pages may be freely searched
and displayed. Permission must be received for
subsequent distribution in print or electronically. Please go to
http://www.umd.umich.edu/ for more information.</p>
      </availability>
    </publicationStmt>
    <sourceDesc>
      <biblFull>
        <titleStmt>
          <title type="245">The Wade-Davis
manifesto. A last appeal to the democracy. By Caspar Butz.</title>
          <author> Butz, Caspar, 1825-
1885.</author>
        </titleStmt>
        <extent>[4] p. 27 cm. </extent>
        <publicationStmt>
          <pubPlace>[Chicago?</pubPlace>
          <date>1864]</date>
        </publicationStmt>
        <notesStmt>
          <note>Caption title.</note>
          <note>At head of title: Printed by
order of the Fremont central committee.</note>
          <note>"From the German American
monthly for September."</note>
          <note>"Messr. Wade and Davis ...
propose to hold a new convention ...
in order to trot out a new presidential candidate against Mr.
Lincoln, in case ... [he] does not ... withdraw voluntarily ...
We only know of one man, who ... can ... re-establish this Union, and
his name is John C. Fremont.--p. [2], [4]</note>

```

²⁷ <http://old.diglib.org/standards/teiheaders.html>

```

        </notesStmt>
        </biblFull>
    </sourceDesc>
    </fileDesc>
    <encodingDesc>
        <projectDesc>
            <p>Header created with script
mrcb2teiutf.xsl on 2005-05-25.</p>
        </projectDesc>
        <editorialDecl n="1">
            <p>This electronic text file was
created by Optical Character Recognition (OCR).
No corrections have been made to the OCR-ed text and no editing has
been done to the content of the original document.
Encoding has been done using the recommendations for Level 1 of the
TEI in Libraries Guidelines. Digital page images are
linked to the text file.</p>
        </editorialDecl>
    </encodingDesc>
    <profileDesc>
        <langUsage id="eng">
            <language>eng</language>
        </langUsage>
        <textClass>
            <keywords>
                <term>Campaign literature -
- United States</term>
                <term>Fremont, John
Charles, -- 1813-1890.</term>
                <term>Lincoln, Abraham, --
1809-1865.</term>
            </keywords>
        </textClass>
    </profileDesc>
</teiHeader>

```



Διαχείριση θεμάτων σε Dublin Core

Η περιγραφή των θεμάτων των πληροφοριακών πόρων με τη χρήση του Dublin Core γίνεται μέσα από το πεδίο <Θέμα> (subject), το οποίο είναι κομβικής σημασίας. Το θέμα εκφράζει το γνωστικό πεδίο στο οποίο ανήκει το τεκμήριο και μπορεί να είναι μια γενικότερη περιγραφή που να αφορά μια επιστήμη ή μια σειρά θεμάτων που να αφορούν εξειδικεύσεις ενός γνωστικού πεδίου. Συγκεκριμένα το στοιχείο «Θέμα» στο Dublin Core έχει τα ακόλουθα χαρακτηριστικά:

«Όνομα πεδίου: Θέμα²⁸

Όνομασία πεδίου (label): Θέμα

Ορισμός: το θέμα ενός πληροφοριακού πόρου²⁹. Τυπικά, το θέμα παρουσιάζεται μέσα από λέξεις κλειδιά, φράσεις κλειδιά ή ταξινομικούς αριθμούς.

Σχόλιο: η χρήση ελεγχόμενου λεξιλογίου ή η χρήση ενός τυπικού ταξινομικού σχήματος είναι η προτιμώμενη»

Οι οδηγίες για την ανάπτυξη των πεδίων στο Dublin Core δίνουν τις ακόλουθες διευκρινίσεις σε ότι αφορά το στοιχείο «Θέμα»³⁰:

«Οδηγίες για τη δημιουργία του περιεχομένου του πεδίου «Θέμα»: Επιλέξτε λέξεις κλειδιά που εκφράζουν θέματα από τον τίτλο ή το πεδίο «Περιγραφή» ή μέσα από το κείμενο του πληροφοριακού πόρου. Αν το θέμα είναι ένα πρόσωπο ή ένα συλλογικό όργανο, χρησιμοποιείστε την ίδια μορφή του ονόματος (του φυσικού προσώπου ή του συλλογικού οργάνου) που θα χρησιμοποιούσατε αν τα πρόσωπα αυτά ήταν οι δημιουργοί ή οι συντελεστές του τεκμηρίου³¹.

²⁸ <http://dublincore.org/documents/2012/06/14/dcmi-terms/?v=elements#subject>

²⁹ Με τον όρο πληροφοριακό πόρο μεταφράζουμε τον αγγλικό όρο resource. Θα μπορούσε να χρησιμοποιηθεί και ο όρος «πηγή», ωστόσο ο όρος πόρος είναι ακριβέστερος και πιο κοντά στην έννοια της λέξης resource τουλάχιστον με τον τρόπο που έχει χρησιμοποιηθεί εδώ για να εκφράσει το πληροφοριακό αγαθό (πόρο) και όχι το σημείο προέλευσης (πηγή). Για διευκόλυνση του κειμένου, θα μπορούσαμε να πούμε ότι ο «πληροφοριακός πόρος» του Dublin Core που έχει σχεδιαστεί για την κάλυψη ηλεκτρονικών πληροφοριακών πόρων θα μπορούσε να αντιστοιχηθεί με τον όρο «τεκμήριο» που χρησιμοποιούμε για τα άλλα πρότυπα ή με τον όρο «βιβλίο» που θα βλέπατε σε μια παραδοσιακή βιβλιοθήκη.

³⁰ <http://dublincore.org/documents/usageguide/elements.shtml>

³¹ Ουσιαστικά εδώ σας δίνει την οδηγία να καταγράψετε τα ονόματα αυτά ακολουθώντας κάποιους τυποποιημένους κανόνες. Οι κανόνες αυτοί θα μπορούσαν να είναι οι AACR2 ή ένα οποιοδήποτε καθιερωμένο αρχείο έχετε στη διάθεσή σας και το οποίο έχει προκύψει από μια βιβλιογραφική διαδικασία.

Γενικά, επιλέξτε τις πιο σημαντικές και μοναδικές λέξεις για να καταγράψετε ως λέξεις κλειδιά, αποφεύγοντας πολύ γενικές λέξεις για να περιγράψετε συγκεκριμένα πράγματα. Το πεδίο «Θέμα» μπορεί να περιλαμβάνει ταξινομικούς αριθμούς αν αυτοί είναι διαθέσιμοι (για παράδειγμα, τον αριθμό LC ή DDC που έχει το τεκμήριο), μπορεί να περιλαμβάνει θεματικές επικεφαλίδες, όρους θησαυρού και λέξεις κλειδιά (ελεύθερο λεξιλόγιο). Αν τα θέματα που καταγράφετε προέρχονται από πολλαπλά λεξιλόγια ξεχωρίστε τα επαναλαμβάνοντας το πεδίο «θέμα». Αν τα θέματα που καταγράφετε προέρχονται από πολλαπλά λεξιλόγια ή από ελεύθερο λεξιλόγιο ξεχωρίστε τα είτε βάζοντας ένα ερωτηματικό ανάμεσα τους (;) ή απλά επαναλαμβάνοντας το πεδίο «θέμα».

Παραδείγματα:

Θέμα: Αεροπλάνα, Ενοικιάσεις; Αεροπλάνα, μισθώσεις

Θέμα: Καταδρομικά αεροπλάνα

Θέμα: Πιλότοι

Παράδειγμα εγγραφής σε Dublin Core από το Ιδρυματικό Αποθετήριο του Εθνικού Ιδρύματος Ερευνών «Ήλιος», ΕΚΤ

Title:	Εικόνες εξουσίας στα βασιλικά και σατραπικά νομίσματα του περσικού κράτους
Author/Creator:	Ζουρνατζή, Αντιγόνη
Contributor:	Sponsor: Χορηγός: Τράπεζα Alpha Bank
Subject Category:	Auxiliary sciences of History::Numismatics Auxiliary sciences of History::Inscriptions. Epigraphy History Geography, Anthropology, Recreation::Anthropology
Date:	20-Apr-2010
Publisher:	Εθνικό Κέντρο Τεκμηρίωσης (ΕΚΤ)
Part Of Event:	Συμβολισμός και Ιδεολογία: Η ανθρώπινη μορφή στο αρχαίο νόμισμα, 2η ημέρα
URL:	http://hdl.handle.net/10442/8329
Publisher web location (link):	http://www.eie.gr/epistimiskoinonia/2009-2010/e-cycle.pdf
Duration:	29:10
Source:	Επιστήμης κοινωνία: ειδικές μορφωτικές εκδηλώσεις, Ιανουάριος-Ιούνιος 2010, Ε' Κύκλος Ομιλιών, Συμβολισμός και Ιδεολογία: Η ανθρώπινη μορφή στο αρχαίο νόμισμα (13.4.2010-27.4.2010), 2η μέρα Αθήνα, Εθνικό Ίδρυμα Ερευνών, 2010-04-20

Type: Video
Presentation

Keywords: απεικόνιση; αρχαίο νόμισμα; Πέρσες; Coins; Ancient history; Signs and symbols

Notes: Υπεύθυνη Μορφωτικών Εκδηλώσεων ΕΙΕ: κ. Ελένη Γραμματικοπούλου, Τηλ . 210.7273.501, fax: 210.72.46.618, E-mail: gramma@eie.gr.
Στον εκθεσιακό χώρο του ΕΙΕ εκτίθενται φωτογραφίες νομισμάτων από τη συλλογή της Τράπεζας Alpha Bank

File	Size	Format
Εικόνες εξουσίας 20-4- 2010 pdf.pdf	80.59 MB	Adobe PDF

Σχολιάζοντας θα πρέπει να αναφέρουμε την ευκολία που μας δίνει το Dublin Core με την αποδοχή πολλαπλών συστημάτων θεματικής πρόσβασης και τη σαφή προώθηση των λέξεων κλειδιών που δεν εντάσσονται σε ελεγχόμενα λεξιλόγια, αλλά προκύπτουν από τα κείμενα των συγγραφέων και τη χρήση της επιστημονικής ορολογίας που αυτοί έχουν επιλέξει. Η ελευθερία αυτή προκύπτει καθώς το Dublin Core έχει σχεδιαστεί για να εξυπηρετήσει ηλεκτρονικούς πληροφοριακούς πόρους, οι οποίοι συνήθως διαθέτουν τα πλήρη κείμενα. Στο σημείο αυτό θα πρέπει να δούμε την εξέλιξη των πληροφοριακών τεκμηρίων, την προοπτική τους και την προσέγγιση προς αυτούς τους τύπους τεκμηρίων, αποφεύγοντας την προσκόλληση σε παρελθόντα συστήματα. Ας πούμε απλά, ότι νέες μορφές τεκμηρίων απαιτούν νέες προσεγγίσεις διαχείρισης.

Προβληματισμοί στη δημιουργία των ελεγχόμενων λεξιλογίων

Η διαδικασία της δημιουργίας των καταλόγων ελεγχόμενου λεξιλογίου δημιούργησε μια σειρά από προβληματισμούς, πολλοί από τους οποίους είναι κοινοί και στη διαδικασία επιλογής θεμάτων που αναφέρθηκαν παραπάνω για τη θεματική επεξεργασία των τεκμηρίων. Συγκεκριμένα:

1. Ο Βαθμός εξειδίκευσης του λεξιλογίου

Ο βαθμός εξειδίκευσης που δίνουμε για ένα θέμα και η καθιέρωσή του στο ελεγχόμενο λεξιλόγιο είναι από τους βασικούς προβληματισμούς. Για παράδειγμα, οι θεματικές επικεφαλίδες της ΕΒΕ έχουν το γενικό όρο <Γάτες> ενώ οι Θεματικές επικεφαλίδες του Κογκρέσου έχουν και τον ειδικότερο όρο <Γάτες του Σιάμ>. Είναι απόφαση

δηλαδή του παραγωγού του καταλόγου σε πόση εξειδίκευση θέλει να φτάσει το λεξιλόγιο που αντιστοιχεί, επίσης στο πόση εξειδίκευση θα παρουσιάζει ο κατάλογος και η δυνατότητα του χρήστη να φτάσει άμεσα στην εξειδικευμένη πληροφορία. Συνήθως η εξειδίκευση είναι μεγαλύτερη στο ελεγχόμενο λεξιλόγιο που έχει δημιουργηθεί για να εξυπηρετήσει επιστημονικό υλικό και μεγάλες συλλογές. Για παράδειγμα, το ελεγχόμενο λεξιλόγιο των επικεφαλίδων του Κογκρέσου LCSH παρέχει εξαιρετική εξειδίκευση και καλύπτει κάθε θέμα. Ενημερώνεται δε διαρκώς ώστε να συμπεριλαμβάνει κάθε νέα ορολογία που προκύπτει από την εξέλιξη των διαφόρων επιστημονικών κλάδων και την εξέλιξη της ίδιας της γλώσσας. Ταυτόχρονα καταργεί όρους που δε χρησιμοποιούνται πια και δημιουργεί παραπομπές σε νέους δόκιμους όρους που τους αντικαθιστούν. Αντίθετα, ελεγχόμενο λεξιλόγιο που έχει δημιουργηθεί για να εξυπηρετήσει γενικές συλλογές ή συλλογές σχολικών ή και λαϊκών βιβλιοθηκών τείνει να είναι γενικότερο αποφεύγοντας τις μεγάλες εξειδικεύσεις. Έτσι, ανάλογα με το προηγούμενο παράδειγμα στο «Θησαυρό» του ΕΚΤ, που έχει δημιουργηθεί για τις λαϊκές βιβλιοθήκες της χώρας μας, θα βρει κανείς απλά τον όρο <γάτες>.

2. Διαχείριση των συνωνύμων

Συμβαίνει συχνά να έχουμε λέξεις που εκφράζουν το ίδιο πράγμα

πχ ενδύματα και ρούχα

ή έννοιες που η σημασία τους είναι τόσο κοντινή που χρησιμοποιούνται χωρίς διάκριση,

πχ οι λέξεις τροφή και φαγητό.

Κάθε συγγραφέας χρησιμοποιεί τη δική του ορολογία και το δικό του τρόπο γραφής και διαμορφώνει τον τίτλο της δουλειάς του ανάλογα. Ωστόσο νοηματικά μπορεί κάποιες εργασίες να ταιριάζουν. Στην ανάκτηση, αυτός που αναζητά θα πρέπει να ψάξει με όλους τους δυνατούς τρόπους έκφρασης μιας έννοιας εκτός αν το λεξιλόγιο είναι ελεγχόμενο και συνώνυμες έννοιες εκφράζονται με έναν καθιερωμένο όρο. Η διαχείριση λοιπόν των συνωνύμων αποτελεί βασικό μέλημα της παραγωγής ενός ελεγχόμενου λεξιλογίου. Ο συνήθης τρόπος με τον οποίο διαχειριζόμαστε τους συνώνυμους όρους είναι με παραπομπή από το μη καθιερωμένο όρο προς τον καθιερωμένο, δηλαδή:

Θησαυρός	Θεματικές Επικεφαλίδες
Γέννα Χρ. Τοκετός	Αρχέτυπα – Καταλογογράφηση Χρ. Καταλογογράφηση Αρχετύπων
Τοκετός Συνώνυμο: Γέννα	Καταλογογράφηση Αρχετύπων Χρ. αντί: Αρχέτυπα - Καταλογογράφηση

Η συνήθης πρακτική για να οριστεί ο καθιερωμένος όρος και ποιοι είναι εκείνοι που θα παραπέμπουν σε αυτόν, είναι να προσδιοριστεί το ποιος είναι αυτός που χρησιμοποιείται περισσότερο. Επίσης, λαμβάνεται υπόψη η καθιερωμένη ορολογία ενός επιστημονικού κλάδου και η συχνότητα εμφάνισης των όρων στα γραπτά κείμενα.

3. Μία λέξη ή πολλές;

Η επιλογή έκφρασης μιας έννοιας με μία ή περισσότερες λέξεις συχνά υπαγορεύεται από τις ίδιες τις έννοιες. Για παράδειγμα, η έννοια <Ενέργεια> είναι γενική και πρέπει να εκφραστεί για να δημιουργήσει μια ομάδα ειδικότερων εννοιών που είτε θα αναφέρονται σε αυτή είτε θα αποτελούν υποδιαίρεσεις της ανάλογα με το σύστημα. Αντίστοιχα, η έννοια «Εξοικονόμηση ενέργειας» απαιτεί δύο λέξεις και δηλώνει ένα υποσύνολο της έννοιας <Ενέργεια>.

Πρέπει να επισημάνουμε ότι η τάση στα ελεγχόμενα λεξιλόγια είναι να παρέχουν πρόσβαση απευθείας στο πιο εξειδικευμένο όρο παρά στο γενικότερο, γεγονός που οδηγεί στη δημιουργία όρων που απαρτίζονται από δύο ή και τρεις λέξεις. Για παράδειγμα, μπορεί να έχουμε τον όρο «Εξοικονόμηση πηγών ενέργειας». Στο σημείο αυτό θα πρέπει να επισημανθεί ότι αποτελεί απόφαση αυτού που αναπτύσσει το λεξιλόγιο αν θα συμπεριλάβει όρους με περισσότερες λέξεις που δίνουν συγκριτικά μεγαλύτερη εξειδίκευση σε έναν ήδη εξειδικευμένο όρο ή αν θα περιορίσει τους όρους στη σύνθεση των δύο. Για παράδειγμα:

<Ενέργεια>

<Εξοικονόμηση ενέργειας>

<Εξοικονόμηση πηγών ενέργειας>

Σημασία έχει ότι η απόφαση πρέπει να ισχύσει για το σύνολο του λεξιλογίου και όχι κατά ομάδες όρων ή ανεξέλεγκτα, οπότε και θα δημιουργήσει άνιση ανάπτυξη του λεξιλογίου. Άρθρα, προθέσεις, αντωνυμίες, σύνδεσμοι και γενικά άκλιτα μέρη του λόγου δε χρησιμοποιούνται στη σύνθεση των επικεφαλίδων ή των όρων ενός θησαυρού.

Αύξηση του αριθμού των λέξεων μπορεί να προκύπτει από την ανάγκη επεξήγησης μιας έννοιας. Για παράδειγμα στον όρο <Λόγος> προσθέτουμε τη διευκρίνιση (Φιλοσοφία) προκειμένου να υποδηλώσουμε ότι ο <Λόγος> εκλαμβάνεται με την έννοια της φιλοσοφίας και όχι της Ρητορικής ή της γραμματικής, ή των Μαθηματικών. Επομένως ο όρος μας διαμορφώνεται:

<Λόγος (Φιλοσοφία)>

Επισημαίνεται και πάλι ότι η πρακτική αυτή πρέπει να διακρίνεται από μια ενιαία πολιτική.

5. Ορθογραφία

Η χρήση του συνηθέστερου ορθογραφικού τύπου είναι η προτιμότερη.

πχ **Ταξίδι** αντί *Ταξείδι*.

Σήμερα οι μηχανές αναζήτησης έχουν τη δυνατότητα να παρακάμπτουν τις ορθογραφικές αποκλείσεις «προτείνοντας» τον ορθό τρόπο γραφής. «Μήπως εννοείτε: ταξίδι;» ωστόσο η τυποποίηση είναι απαραίτητη και κυρίως η αδιάκριτη χρήση πότε του ενός τρόπου γραφής και πότε του άλλου δεν είναι δόκιμη πρακτική.

6. Διαχείριση συντομογραφιών

Κατά παράδοση στα ελεγχόμενα λεξιλόγια, οι συντομογραφίες και τα ακρόνυμα παίρνουν την πλήρη τους μορφή πχ. *EIE* → Εθνικό Ίδρυμα Ερευνών.

Ωστόσο, κάποια ακρόνυμα, όταν αναπτυχθούν μετατρέπονται σε άγνωστες οντότητες ή έννοιες. Για παράδειγμα στις Θεματικές Επικεφαλίδες της ΕΒΕ θα δούμε *AIDS* (βλ) → Σύνδρομο Επίκτητης Ανοσοποιητικής Ανεπάρκειας

Που ουσιαστικά μας οδηγεί με παραπομπή από τη συντομογραφία στη χρήση του όρου στην πλήρη του μορφή. Η παραπομπή «θεραπεύει» την αχρηστία του πλήρους όρου

και ταυτόχρονα για τις Επικεφαλίδες της ΕΒΕ διατηρεί την ομοιομορφία της χρήσης των αναπτυγμένων μορφών των συντομογραφιών που έχει καθιερώσει.

Αντίθετα, σε άλλα εργαλεία ελεγχόμενων λεξιλογίων, μπορεί η μεταχείριση των συντομογραφιών να είναι διαφορετική και να διατηρούν τη σύντομη γραφή τους. Για παράδειγμα στο Θησαυρό του ΕΚΤ θα δούμε να καθιερώνει τη συντομογραφία αντί για τον πλήρη όρο.

<AIDS>

Συνώνυμο: Σύνδρομο Επίκτητης Ανοσοποιητικής Ανεπάρκειας

Γενικά, όροι οι οποίοι έχουν καθιερωθεί αποκλειστικά ως ακρώνυμα είναι προτιμότερο να διατηρούνται ως έχουν μια και αυτός είναι ο συνηθέστερος τρόπος να τους αναζητήσει κάποιος. Σε κάθε άλλη περίπτωση πρέπει να ακολουθείται η ανάπτυξη τους.

7. Εκκλαϊκευμένο ή Επιστημονικό/ Τεχνικό λεξιλόγιο

Το ελεγχόμενο λεξιλόγιο πρέπει να κρατά μια ισορροπία σε σχέση με τη χρήση για την οποία απευθύνεται. Για παράδειγμα, ο Θησαυρός του ΕΚΤ που έχει δημιουργηθεί για τις ανάγκες των λαϊκών βιβλιοθηκών και των χρηστών τους έχει τον όρο

<Καρκίνος>

Αντίθετα ο Θησαυρός Ιατρικών όρων του ΕΚΤ που απευθύνεται στις ειδικές ιατρικές βιβλιοθήκες δεν έχει καθόλου τον όρο καρκίνος αλλά αντί αυτού έχει τον όρο

<Νεοπλάσματα>

Πρέπει να σημειωθεί ότι με τη χρήση των νέων τεχνολογιών πληροφόρησης όροι εκκλαϊκευμένοι και επιστημονικοί μπορεί να παρουσιάζονται στο ίδιο λήμμα με τη μορφή συνωνύμων και να ανακτώνται ανεξάρτητα από τη μορφή με την οποία έχουν αναζητηθεί. Δηλ. Να αναζητώ με τον όρο «Καρκίνος» και να ανακτώ και όλα τα λήμματα που έχουν καταχωρηθεί με τον όρο «Νεοπλάσματα». Αυτό αποτελεί επίτευγμα της χρήσης των μηχανών αναζήτησης και της ανάπτυξης σχετικών οντολογιών.

Συμπεράσματα

Η ανάπτυξη των ελεγχόμενων λεξιλογίων στηρίζει όλα τα συστήματα θεματικής πρόσβασης και ταυτόχρονα αποτελεί τη βάση πάνω στην οποία έχουν στηριχθεί τα

αυτοματοποιημένα συστήματα αναζήτησης και ανάκτησης πληροφοριών που ακολουθούν τη λογική των ελεύθερων λεξιλογίων και την απόδοση «Λέξεων κλειδιών».

Η χρήση των ελεγχόμενων λεξιλογίων έχει πλεονεκτήματα και μειονεκτήματα:

Τα πλεονεκτήματα είναι:

- Αποδίδουν πληροφoρία με ακρίβεια και ταχύτητα
- Συνδέουν έννοιες και αποδίδουν εξειδικεύσεις
- Δομούν τα τεκμήρια σε ενότητες και υποβοηθούν την έρευνα
- Αποδίδουν ποιοτικές υπηρεσίες προς τους χρήστες
- Μπορούν να δημιουργούν συνδέσμους με ταξινομικά συστήματα
- Μπορούν να δημιουργούν συνδέσμους μεταξύ διαφορετικών γλωσσών

Τα μειονεκτήματα είναι:

- Τα ελεγχόμενα λεξιλόγια περιορίζουν τη φυσική γλώσσα
- Έχουν κόστος παραγωγής και συντήρησης
- Συχνά χάνουν σε εξειδίκευση
- Η έλλειψη διαρκούς ενημέρωσης τους και επέκτασής τους συχνά τα αχρηστεύουν ή μειώνουν την ικανότητά τους για απόδοση εξειδικεύσεων ή/και ομαδοποίηση τεκμηρίων

Βιβλιογραφία

Cambell, D.G. (2007). Part I: Semantic Web Foundations, Standards and Tools, “The Birth of the New Web: A Foucauldian Reading of the Semantic Web” *Cataloging & Classification Quarterly* Vol. 43, No. 3/4, 2007, pp. 9-20. doi:10.1300/J104v43n03_02 9 (Ανακτήθηκε 26 Σεπτεμβρίου 2012).

Dublin Core Metadata Initiative. <http://dublincore.org/documents/2012/06/14/dcmi-terms/?v=elements#subject> (Ανακτήθηκε 1 Οκτωβρίου 2012).

Facet analysis.

http://www.iva.dk/bh/lifeboat_ko/CONCEPTS/facet_and_facet_analysis.htm
(Ανακτήθηκε 4 Οκτωβρίου 2012).

Hjorland, Birger (2002). “Domain analysis in information science: Eleven approaches - traditional as well as innovative”. *Journal of Documentation*; 2002; 58, 4; ProQuest pg. 422 . DOI: 10.1108/00220410210431136.

University of Michigan. Library. TEI Headers.
<http://old.diglib.org/standards/teiheaders.html> (Ανακτήθηκε 26 Οκτωβρίου 2012).