

Ηλεκτρονικός πολύγλωσσος θησαυρός : το μέλλον στην αυτοματοποιημένη ευρετηρίαση

A. Βασιλείου

Η διαδικασία καθώς και οι μέθοδοι βελτιστοποίησης των υπηρεσιών πληροφόρησης έχουν απασχολήσει επανειλημμένα τους ειδικούς με αποτέλεσμα να παρατηρείται μία αδιάκοπη αναζήτηση νέων, μοντέρνων και ευέλικτων εργαλείων για την ικανοποίηση των προηγμένων αναγκών που τίθενται τόσο στην τεκμηρίωση όσο και στην ευρετηρίαση των πληροφοριών. Οι θεματικές επικεφαλίδες αποτέλεσαν για μεγάλο χρονικό διάστημα, τον κύριο άξονα της ευρετηρίασης και προσέφεραν αναλυτική πληροφόρηση χωρίς όμως να επιτρέπουν πάντα άμεση πρόσβαση στον κατάλογο αφού θα πρέπει ο χρήστης να έχει βρει την κατάλληλη επικεφαλίδα με την πρώτη προσπάθεια. Αυτό προϋποθέτει ότι θα έχει διατυπώσει την ερώτηση του :

- α) Κατονομάζοντας το θέμα με τους ίδιους όρους που έχει χρησιμοποιήσει ο ταξινόμος,
- β) Κατονομάζοντας το θέμα στο επιθυμητό επίπεδο εξειδίκευσης (specific level),
- γ) Αναφέροντας αυτούς τους όρους με την ίδια σειρά αναφοράς που έχει χρησιμοποιήσει ο βιβλιοθηκάριος.

Οι περιορισμοί αυτοί έθεσαν τον προβληματισμό για την ποιότητα των υπηρεσιών των ευρετηρίων θεματικών επικεφαλίδων όσο και για την ενδεχόμενη δυνατότητα μετεξέλιξής τους. Η διάδοση και η χρήση των συστημάτων πληροφόρησης έθεσαν τον θεμέλιο λίθο για την δημιουργία Αυτοματοποιημένου θησαυρού, ενός ισχυρού εργαλείου για ευρετηρίαση και πολυγλωσσική αναζήτηση.

Δεδομένου ότι οι τάσεις, της τεχνολογικά αναπτυσσόμενης κοινωνίας των πληροφοριών, επιβάλλουν τη χρήση τέτοιων εργαλείων στην αυτοματοποιημένη μορφή τους, η υιοθέτηση ενός ηλεκτρονικού πολύγλωσσου θησαυρού αποτελεί βασική προϋπόθεση για την ολοκλήρωση της αυτοματοποιημένης τεκμηρίωσης, για την ουσιαστική συμβολή της ευρετηρίασης στη διάδοση της "γνώσης" καθώς και για την "χωρίς σύνορα", επικοινωνία των βιβλιογραφικών συστημάτων.

Αν και η δόμηση του θησαυρού είναι μία διανοητική εργασία, ο Η/Υ μπορεί να βοηθήσει πάρα πολύ, αφενός γιατί η ενημέρωση γίνεται πλέον άμεση, αφετέρου βοηθάει τον έλεγχο για την επισήμανση λαθών, τις διορθώσεις, την αντιστροφή των σχέσεων ανάμεσα στους όρους (π.χ. έχουμε δώσει εμείς σε κάθε όρο τους ευρύτερους, μπορεί ο Η/Υ να μας βρεί για κάθε ευρύτερο όρο όλους τους ειδικότερους, χωρίς να παραλήψει κανένα), την εκτύπωση τμήματος ή του συνόλου του θησαυρού κ.τ.λ.

Η συγκρότηση πολύγλωσσου θησαυρού δεν είναι τίποτε άλλο από την δημιουργία θεματικού καταλόγου όρων δομημένων στην δένδρική μορφή, μέσω σχέσεων που

* ΟΡΙΣΜΟΙ:

1. Περιγραφές : είναι ένας όρος που έχει επιλεγεί μέσα από ένα σύνολο ισοδύναμων όρων για να εκφράσει μία έννοια μονοσήμαντα. Αποτελείται από μία ή περισσότερες λέξεις.
2. Μη περιγραφές : είναι ένας από τους ισοδύναμους όρους που δεν χρησιμοποιείται στην τεκμηρίωση αλλά παραπέμπει στην επιλογή του κατάλληλου περιγραφέα.

αναπαρίστανται με σύμβολα, σύμφωνα με το διεθνές πρότυπο ISO 5964-1985 Symbolization of thesaural relationships. Όταν η διαδικασία αυτή πραγματοποιηθεί με τα σύγχρονα εργαλεία της πληροφορικής αναφερόμαστε στον Αυτοματοποιημένο Πολύγλωσσο Θησαυρό που προσφέρει την διεπαφή αναζήτησης μέσω των εννοιολογικών δένδρων.

Για τον διαχωρισμό των όρων του Θησαυρού, χρησιμοποιούνται σύμβολα με λατινικούς χαρακτήρες, τα οποία υποδηλώνουν το είδος της σχέσης του καθενός, με τους όρους που τον περιβάλλουν. Σε όσα ακολουθούν παρουσιάζονται τα σύμβολα εκείνα που συναντιούνται συχνότερα σε θησαυρούς, που έχουν εκδοθεί, καθώς επίσης και το πρότυπο ISO 5964, το οποίο χρησιμοποιεί γραφικά σύμβολα για την αναπαράσταση των σχέσεων :

SYMBOL	ISO 5964 SYMBOL	SIGNIFICANCE	ΜΕΤΑΦΡΑΣΗ
USE		use	χρησιμοποιείται
UF	=	use for	χρησιμοποιείται για
BT	<	broader term	ευρύτερος όρος
BT(A)	—<	associated broader term	συνδεδεμένος ευρύτερος όρος
NT	>	narrower term	στενότερος όρος
NT(A)		associated narrower term	συνδεδεμένος στενότερος όρος
RT	—	related term	σχετιζόμενος όρος
RT(A)	-----	associated related term	συνδεδεμένος σχετιζόμενος όρος
ET(x)	X	equivalent term in language X	ισοδύναμος όρος στη γλώσσα X

Οι σχέσεις των όρων, οι οποίες διαμορφώνουν την δομή του θησαυρού διακρίνονται σε τρεις κατηγορίες:

1. Σχέσεις ισοδυναμίας (Equivalence relationships). Τέτοιες είναι οι σχέσεις USE και UseFor.
Οι σχέσεις ισοδυναμίας εφαρμόζονται μεταξύ των περιγραφών και των αντίστοιχων τους μη περιγραφών.
2. Σχέσεις ιεραρχίας (Hierarchical relationships).
Οπως : BT (Broader Term), NT (Narrower Term), BT(A) (Associated broader term), NT(A) (Associated related term). Οι σχέσεις ιεραρχίας αναπτύσσονται μεταξύ περιγραφών.
3. Συνδεδεμένες σχέσεις (Associated relationships) RT (Related term) RT(A) (Associated related term) ET(x) (Equivalent term) *
* Σημείωση : Για πολύγλωσσους θησαυρούς μόνο.

* ΟΡΙΣΜΟΙ:

1. Περιγραφές : είναι ένας όρος που έχει επιλεγεί μέσα από ένα σύνολο ισοδύναμων όρων για να εκφράσει μία έννοια μονοσήμαντα. Αποτελείται από μία ή περισσότερες λέξεις.
2. Μη περιγραφές : είναι ένας από τους ισοδύναμους όρους που δεν χρησιμοποιείται στην τεκμηρίωση αλλά παραπέμπει στην επιλογή του κατάλληλου περιγραφέα.

Οι σχέσεις ιεραρχίας αναπτύσσονται μεταξύ περιγραφών. Η δομή του θησαυρού κατά συνέπεια είναι τρισδιάστατη. Η καθιέρωση των όρων εφαρμόζεται στις σχέσεις ισοδυναμίας από όπου και διοχετεύεται στις άλλες δύο κατηγορίες.

Όταν ο εννοιολογικός χάρτης ενός θησαυρού γίνεται με αλφαβητική κατάταξη των περιγραφών κάθε ενότητα περιλαμβάνει : Το βασικό περιγραφέα, τα συνώνυμα του, συνδεδεμένα με σχέσεις ισοδυναμίας, τους γενικότερους και ειδικότερους περιγραφείς συνδεδεμένους με σχέσεις ιεραρχίας, και τέλος τους συγγενικούς με τον όρο περιγραφείς, που παρουσιάζονται με συνδετικές σχέσεις.

Δίνουμε ένα παράδειγμα τέτοιας κατάταξης :

ΜΗΧΑΝΙΚΟΙ

UF Τεχνικοί επιστήμονες

BT Επιστήμονες

NT Ηλεκτρολόγοι Μηχανικοί

 Πολιτικοί Μηχανικοί

 Χημικοί Μηχανικοί

RT Φυσικοί Επιστήμονες

ET Engineers

Μέσα από την ενότητα του Αυτοματοποιημένου Θησαυρού προσφέρεται ένα ισχυρό εργαλείο για τις αναζητήσεις, γιατί δίνει στον χρήστη την δυνατότητα ξεκινώντας από γενικότερους όρους να καταλήξει σε κάποιο συγκεκριμένο που τον αντιπροσωπεύει με ακρίβεια και να ανακτήσει όλες τις σκορπισμένες συγγενικές έννοιες, ενώ ταυτόχρονα παρέχει έλεγχο των συνωνύμων όρων και λεκτικών μορφών, ώστε κάθε φορά να επιλέγεται ΕΝΑΣ όρος ευρετηρίασης, με άμεσο όφελος τον περιορισμό του όγκου του ευρετηρίου.

Ένα ολοκληρωμένο αυτοματοποιημένο σύστημα θησαυρού πρέπει να επιτρέπει την έρευνα του θησαυρού σε δύο επίπεδα :

- α) Σάρωση Θησαυρού.
Ο χρήστης πρέπει να έχει την δυνατότητα να εντοπίζει όχι μόνο έναν όρο και τις σχετικούς με αυτόν όρους, αλλά και τον τύπο των συσχετίσεων. Της πρέπει να μπορεί να επιλέγει μεταξύ διαφορετικών όρων, αλλάζοντας συνεχώς θέση στην ιεραρχική δομή του θησαυρού.
- β) Εντοπισμός βιβλιογραφικών εγγραφών.
Η σύνδεση όρων του θησαυρού και βιβλιογραφικών εγγραφών με τις οποίες συνδέεται πρέπει να είναι άμεση. Στο σημείο αυτό έγκειται η σπουδαιότητα του πολύγλωσσου θησαυρού η οποία περιγράφεται ως εξής:

Ο πολύγλωσσος θησαυρός συνδέει τις βιβλιογραφικές εγγραφές σε οποιαδήποτε γλώσσα και αν έχουν καταλογηθεί. Έτσι επιτρέπει στον χρήστη να ανακτήσει με μία αναζήτηση, τις συσχετιζόμενες με τον αναζητήσιμο όρο εγγραφές, σε περισσότερες από μία γλώσσες.

Π. χ.

A. ΣΥΝΕΔΡΙΟ

E.T. conference

ή

B. ΙΣΤΟΡΙΑ

E.T. History

Η αναζήτηση με οποιοδήποτε από τους τρεις όρους, στην μία ή στην άλλη περίπτωση, πρέπει να ανακτήσει όλες τις εγγραφές που έχουν καταλογογραφηθεί στα Ελληνικά είτε στα Αγγλικά είτε στα Γαλλικά.

Αυτό φυσικά διευκολύνει πολύ απομακρυσμένους χρήστες που κάνουν αναζήτηση σε βιβλιοθήκες που το υλικό τους είναι καταλογογραφημένο σε διαφορετική γλώσσα από την δική τους. Ερευνητές μάλιστα από άλλες χώρες, που αναζητούν υλικό στις ελληνικές βιβλιοθήκες, είναι πάρα πολύ δύσκολο να γνωρίζουν τον όρο αναζήτησης στην δική μας γλώσσα.

Κατά την διαδικασία αναζήτησης το σύστημα πρέπει να επιτρέπει χρήση χαρακτήρων μπαλαντέρ για αποκοπή χαρακτήρων. Επιπλέον, οι αναζητήσεις πρέπει να εκτελούνται με μεγάλη ταχύτητα ανεξάρτητα από το πλήθος των όρων του θησαυρού.

Παραγωγή καταλόγων

Το σύστημα πρέπει να επιτρέπει τη δημιουργία καταλόγων με τους όρους του θησαυρού και τις συσχετίσεις αυτών. Ένας αρχικός κι ένας τελικός όρος μπορούν να προσδιοριστούν με σκοπό να περιοριστεί το μέγεθος του καταλόγου. Κάθε κατάλογος μπορεί να εμφανίζεται στην οθόνη, να εκτυπώνεται και να καταγράφεται σε αρχείο ASCII ανάλογα με τις απαιτήσεις του χρήστη. Το σύστημα πρέπει να επιτρέπει τη δημιουργία καταλόγων διαφορετικών μορφών όπως:

- α) Κατάλογος όρων θησαυρού.
Αλφαβητικός κατάλογος με τους όρους του θησαυρού. Για κάθε όρο θα πρέπει να εμφανίζεται η κατάσταση του και ο ταξινομικός αριθμός.
- β) Ταξινομικός κατάλογος.
Περιέχει ότι και ο αλφαβητικός με τη διαφορά ότι ο κατάλογος ταξινομείται με βάση τον ταξινομικό αριθμό.
- γ) Κατάλογος περιγραφών (προτιμητέων όρων).
Στον κατάλογο αυτόν περιλαμβάνονται μόνο οι όροι που είναι περιγραφείς (προτιμητέοι όροι).
- δ) Αλφαβητικός κατάλογος όρων και συσχετίσεων σε κανονική μορφή.
Κάθε όρος εμφανίζεται μαζί με τους συσχετιζόμενους όρους, καθώς και τον τύπο της συσχέτισης.
- ε) Κατάλογος όρων και συσχετίσεων σε δενδρική μορφή
(με βάση το πρότυπο ISO 5964) .

Προσφορά ενός αυτοματοποιημένου θησαυρού στην αναβάθμιση των υπηρεσιών που προσφέρει ένα σύγχρονο κέντρο πληροφόρησης

Η χρήση του θησαυρού διευκολύνει άμεσα τον χρήστη αφού του επιτρέπει να χαράξει ο ίδιος την στρατηγική του και να επιλέξει τους όρους που έχουν

χρησιμοποιηθεί στον θησαυρό. Ο αυτοματισμός του θησαυρού είναι φανερό ότι επηρεάζει την υλική πλευρά της διαδικασίας, προσφέροντας ταχύτερες και βελτιωμένες υπηρεσίες. Δεν επηρεάζει τις αρχές της υστεροσυνδιασμένης ευρετηρίασης που παραμένουν οι ίδιες.

Οποιοσδήποτε θησαυρός όρων έχει εκδοθεί σε έντυπη μορφή στην παγκόσμια αγορά, μπορεί εύκολα να λειτουργήσει μέσα στο λογισμικό ενός αυτοματοποιημένου θησαυρού. Για να μπορέσουμε να διαπιστώσουμε και να εκτιμήσουμε τη συμβολή του θησαυρού στην αναβάθμιση των υπηρεσιών ενός κέντρου πληροφόρησης, πρέπει να το δούμε ως μία ενότητα ενός ολοκληρωμένου συστήματος αυτοματοποίησης βιβλιοθηκών, στο οποίο οι λειτουργίες που προσφέρει χρησιμοποιούνται από τα υπόλοιπα υποσυστήματα όπως καταλογογράφηση, OPAC, Δανεισμοί, SDI, κτλ.

Στην Ελλάδα δεν έχουμε παρόμοια βοηθήματα. Πολλές βιβλιοθήκες διατηρούν αρχεία των επικεφαλίδων τους, αυτά όμως δεν έχουν συγκροτηθεί στη βάση ενιαίων κανόνων. Επομένως η συνένωση των επιμέρους αρχείων θα χρειαστεί μια περαιτέρω επεξεργασία, ώστε η χρήση τους να είναι σε κάθε περίπτωση αποτελεσματική.

Τον τελευταίο καιρό αρχίζουν προσπάθειες από διάφορες υπηρεσίες για τη συγκρότηση θησαυρών. Η δημοσίευση ελληνικών προτύπων στον τομέα αυτό της ευρετηρίασης, θα βοηθούσε στην εφαρμογή ενιαίων μεθόδων, απαραίτητη προϋπόθεση για την συνεργασία και την ανταλλαγή πληροφοριών.

Ας εξετάσουμε όμως αναλυτικά, πως οι πληροφορίες που παρέχει ο θησαυρός χρησιμοποιούνται από τα υπόλοιπα υποσυστήματα.

Ξεκινώντας από την ενότητα της καταλογογράφησης μπορούμε εύκολα να διαπιστώσουμε ότι η ύπαρξη αυτοματοποιημένου θησαυρού εξασφαλίζει την συνέπεια εφόσον επιτρέπει μόνο τη χρήση της καθιερωμένης μορφής κάθε όρου. Επιπλέον η συνεργασία των δύο υποσυστημάτων επιτρέπει στον καταλογογράφο να μεταφέρει αυτόματα όρους από το θησαυρό και να τους συνδέει στην τρέχουσα βιβλιογραφική εγγραφή εξοικονομώντας σημαντικό χρόνο.

Η δυνατότητα που παρέχεται στους χρήστες του OPAC να σαρώνουν με μεγάλη ευκολία και ταχύτητα την δενδρική δομή του θησαυρού για να εντοπίζουν τους όρους και στην συνέχεια τις εγγραφές για τις οποίες ενδιαφέρονται: α) διευκολύνει την πρόσβαση στην ζητούμενη πληροφορία και β) μειώνει σημαντικά τον χρόνο ανεύρεσης, στοιχεία που αποτελούν το σκοπό της εργασίας της τεκμηρίωσης.

Η συνεργασία του θησαυρού με την ενότητα της επιλεκτικής διάδοσης των πληροφοριών (SDI)³ παίζει σημαντικό ρόλο στην πληρότητα και την εγκυρότητα της πληροφόρησης των χρηστών. Οι χρήστες του SDI έχουν τη δυνατότητα να σαρώσουν το θησαυρό και να μεταφέρουν αυτόματα τους όρους που τους ενδιαφέρουν στο προφίλ τους. Με τον τρόπο αυτό η ενημέρωσή τους μέσω των ανακοινωθέντων (bulletins) είναι πληρέστερη και εγκυρότερη.

Η εννοιολογική ομαδοποίηση (conceptual clustering) των όρων του θησαυρού βάση του ταξινομικού τους αριθμού επιτρέπει την παραγωγή βιβλιογραφικών δελτίων και

άλλων αναφορών βάση ευρύτερων θεματικών περιοχών. Για τα βιβλιογραφικά δελτία υπάρχει δυνατότητα αυτόματης παραγωγής αλφαβητικών θεματικών ευρετηρίων, εφόσον επιθυμείται από την υπηρεσία.

Στην εννοιολογική ομαδοποίηση των όρων του θησαυρού σε ευρύτερες θεματικές περιοχές μπορεί να βασιστεί και η έκδοση στατιστικών στοιχείων σχετικά με την κίνηση του υλικού που αποτελεί βασική πληροφορία για το ίδιο το κέντρο τεκμηρίωσης.

Τέλος η δυνατότητα χρησιμοποίησης πολύγλωσσου θησαυρού ως εργαλείο για την ανάκτηση πληροφοριών μέσω ενός δημοσίου δικτύου (public network), αποκτά ιδιαίτερη σημασία αφού είναι δυνατή η on-line παρουσίαση των όρων και των μεταξύ τους συσχετίσεων, σε όλες τις γλώσσες που υποστηρίζονται.

(3) S.D.I. (Selective Dissemination of Information)

Σύγχρονη ενότητα που προσφέρουν τα ολοκληρωμένα προγράμματα αυτοματοποίησης βιβλιοθηκών. Η επιλεκτική διάδοση πληροφοριών, όπως αναφέρεται στα Ελληνικά προσφέρει εξατομικευμένη πληροφόρηση για το νέο υλικό της υπηρεσίας πληροφόρησης με βάση το ιδιαίτερο profile πληροφοριακών αναγκών για κάθε χρήστη.