

**ΙΟΝΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ**  
**ΤΜΗΜΑ ΑΡΧΕΙΟΝΟΜΙΑΣ – ΒΙΒΛΙΟΘΗΚΟΝΟΜΙΑΣ**  
Μεταπτυχιακό Πρόγραμμα Σπουδών στην Επιστήμη της Πληροφορίας  
Διοίκηση και Οργάνωση Βιβλιοθηκών με έμφαση στις Νέες Τεχνολογίες της  
Πληροφορίας



ΕΡΓΑΣΙΑ ΣΤΑ ΠΛΑΙΣΙΑ ΤΟΥ ΜΑΘΗΜΑΤΟΣ

«ΤΕΧΝΟΛΟΓΙΕΣ ΤΗΣ ΠΛΗΡΟΦΟΡΙΑΣ»

ΘΕΜΑ:

«ΑΝΑΖΗΤΗΣΗ ΚΑΙ ΑΝΑΚΤΗΣΗ ΠΛΗΡΟΦΟΡΙΩΝ ΑΠΟ  
ΔΙΑΦΟΡΕΣ ΠΗΓΕΣ»

ΚΑΘΗΓΗΤΡΙΑ: Κ. ΤΟΡΑΚΗ

ΕΠΙΜΕΛΕΙΑ: ΕΛΕΝΗ ΚΑΡΑΝΙΚΟΛΑ  
ΜΑΡΙΑ ΜΑΖΝΩΚΗ

ΙΟΥΝΙΟΣ 2008

ΑΘΗΝΑ

## Πίνακας Περιεχομένων

1. Βιβλιοθήκες-Αναζήτηση-Ανάκτηση Πληροφοριών.....	2
2. Αναζήτηση και ανάκτηση πληροφοριών στο συμβατικό περιβάλλον.....	2
2.1. Κατάλογοι .....	2
2.1.1. Είδη καταλόγων .....	2
2.2. Βιβλιογραφία .....	3
2.3. Συστήματα ευρετηρίασης .....	4
2.4. Κατάλογοι θεματικών επικεφαλίδων και Θησαυροί .....	5
3. Αναζήτηση και ανάκτηση πληροφοριών στο ψηφιακό περιβάλλον.....	7
3.1. Διαδίκτυο .....	8
3.1.1. Άλλες υπηρεσίες του Διαδικτύου .....	9
3.2. Μηχανές αναζήτησης.....	11
3.3. Μετα-μηχανές αναζήτησης.....	12
3.4. Συλλογές μηχανών αναζήτησης.....	13
3.5. Οδηγοί- Κατάλογοι .....	14
3.6. Βάσεις δεδομένων.....	14
3.7. Αυτοματοποιημένος κατάλογος.....	15
3.8. Crawlers .....	16
3.9. Γλώσσες υποβολής ερωτημάτων .....	16
3.10. XML.....	18
3.11. Πρωτόκολλα .....	18
3.11.1. Z39.50 .....	18
3.11.2. OAI .....	19
3.11.3. SRU.....	21
3.12. Συστήματα αναζήτησης και ανάκτησης πολυμεσικού υλικού .....	23
3.12.1. Συστήματα ανάκτησης εικόνων.....	23
3.12.2. Συστήματα ανάκτησης ήχου .....	23
3.12.3. Συστήματα ανάκτησης βίντεο.....	24
3.13. Υπηρεσίες αναζήτησης .....	25
4. Η Διαδικασία της Αναζήτησης .....	25
5. Τεχνικές Ανάκτησης Πληροφοριών .....	27
6. Ρόλος Βιβλιοθηκονόμου .....	28
7. Συμπεράσματα .....	29
8. Βιβλιογραφία .....	31

## ***1. Βιβλιοθήκες-Αναζήτηση-Ανάκτηση Πληροφοριών***

Ο ρόλος, η θέση και η μορφή της βιβλιοθήκης ως θεσμού στην κοινωνία της πληροφορίας επανεξετάζεται. Οι νέες τεχνολογίες στην παραγωγή, διακίνηση, δημοσίευση, οργάνωση και διάθεση της πληροφορίας αμφισβητούν τα υπάρχοντα μοντέλα οργάνωσης και λειτουργίας της βιβλιοθήκης.

Η μεγάλη ποσότητα και οι ποικίλες μορφές πληροφοριών που παράγονται και συγκεντρώνονται από της βιβλιοθήκες, καθώς και η αβέβαιη ποιότητας τους δημιουργούν νέες προκλήσεις ως προς την αναζήτηση ανάκτηση και απόδοση τους. Η «άνεση» στην τεχνολογία της πληροφόρησης απαιτεί τεχνικές και πνευματικές δεξιότητες που θα πρέπει να έχει ο τελικός αποδέκτης της πληροφορίας για την σωστή και αποτελεσματική απόδοση της. Στο σύγχρονο περιβάλλον οι βιβλιοθήκες διευρύνουν τον ρόλο τους, ο οποίος προσανατολίζεται προς την κατεύθυνση του εντοπισμού, της συγκέντρωσης και οργάνωσης των πληροφοριών.

Η ανάγκη για αλλαγή δεν είναι τυχαία αλλά προέρχεται από την ζωτική σημασία που έχει η πληροφορία για τον άνθρωπο σε μορφωτικό, ερευνητικό, ψυχαγωγικό και οικονομικό επίπεδο.

## ***2. Αναζήτηση και ανάκτηση πληροφοριών στο συμβατικό περιβάλλον***

### ***2.1. Κατάλογοι***

Ο κατάλογος του υλικού της βιβλιοθήκης ποικίλλει από απλό κατάλογο μέχρι πλήρη βιβλιογραφία. Από τους πιο συνηθισμένους είναι ο κατάλογος με τα βιβλία ή με τα περιοδικά που παρέλαβε η βιβλιοθήκη σε ένα συγκεκριμένο χρονικό διάστημα. Τα στοιχεία που περιλαμβάνονται για τα βιβλία είναι συνήθως ο ταξινομικός αριθμός, ο συγγραφέας, ο τίτλος, ο εκδότης, ο τόπος και ο χρόνος έκδοσης. Για τα περιοδικά ο τίτλος, ο τόμος, το τεύχος, ο εκδότης και ο χρόνος έκδοσης. Για τα βιβλία μπορεί να ακολουθηθεί αλφαβητική σειρά κατά συγγραφέα ή κατά τίτλο και για τα περιοδικά κατά τίτλο. Εκτός από τους καταλόγους αυτούς που έχουν σαν στόχο να επισύρουν την προσοχή του κοινού στο καινούργιο υλικό, μπορούν να συνταχθούν κατάλογοι με υλικό σε ένα ορισμένο τομέα, σε μία ειδική συλλογή, με υλικό που είναι διασκορπισμένα από το ταξινομικό σύστημα κι από την ταξινότηση του στα ράφια, με αφορμή την επικαιρότητα ή να προταθούν βιβλία ή άλλου είδους υλικό σε διάφορους τομείς. Οι κατάλογοι αυτοί είναι χρήσιμοι και για το προσωπικό και για τους χρήστες. ( Μονιάρου- Παπακωνσταντίνου Βαλεντίνη, 1998)

#### ***2.1.1. Είδη καταλόγων***

Σε ένα συμβατικό σύστημα βιβλιοθήκης – υπηρεσίας πληροφόρησης χρησιμοποιήθηκαν κατά το πέρασμα των χρόνων τα παρακάτω είδη καταλόγων (Σταϊκούρα, 1996):

- Ως προς το σκοπό

- ✍ Δημόσιος : Χρησιμοποιείται για την εξυπηρέτηση του κοινού της βιβλιοθήκης και για αυτό το λόγο βρίσκεται πάντα σε εμφανές μέρος της βιβλιοθήκης.
  - ✍ Επίσημος : Χρησιμοποιείται από το προσωπικό της βιβλιοθήκης και βρίσκεται στο τμήμα καταλογογράφησης.
  - ✍ Συλλογικός : Καταγράφει τις συλλογές πολλών βιβλιοθηκών και κατά συνέπεια βοηθάει τον χρήστη να βρει ποια βιβλιοθήκη έχει τι. Οι συλλογικοί κατάλογοι είναι τοπικοί και εθνικοί.
- Ως προς το σχήμα ή τη μορφή
    - ✍ Τυπωμένος κατάλογος ή βιβλιοκατάλογος : Αποτελείται από έναν ή περισσότερους τόμους και είναι η πιο παλιά μορφή καταλόγου. Αν και εύκολος στη χρήση η παραγωγή του стоίχιζε αρκετά χρήματα ενώ δεν παρείχε τη δυνατότητα ενσωμάτωσης αναγραφών νέου υλικού. Ο βιβλιοκατάλογος άρχισε να πεθαίνει νωρίς, στις αρχές του εικοστού αιώνα όταν αντικαταστάθηκε από τον ευέλικτο δελτιοκατάλογο.
    - ✍ Δεματοκατάλογος : Σε αυτή τη μορφή καταλόγου οι αναγραφές καταχωρούνται σε φύλλα χαρτιού χωριστά και ταξιθετούνται σε ειδικά φτιαγμένα κλασέρ. Ουσιαστικά απετέλεσε έναν συνδυασμό της άνεσης του βιβλιοκαταλόγου και της ευκαμψίας για πρόσθεση και αφαίρεση των αναγραφών του δελτιοκαταλόγου. Δεν αποτελεί ευρέως γνωστό τύπο καταλόγου και χρησιμοποιήθηκε κυρίως σε πολλές βιβλιοθήκες της Μεγάλης Βρετανίας.
    - ✍ Δελτιοκατάλογος : Αποτελούσε την πλέον συνηθισμένη μορφή καταλόγου που μπορούσαμε να συναντούμε σε μια συμβατικού τύπου βιβλιοθήκη. Οι αναγραφές καταχωρούνται σε χωριστά δελτία σταθερού μεγέθους 12,5 εκ. x 7,5 εκ. και φυλάσσονται σε ειδικό έπιπλο (εικόνα 1). Στο κάτω μέρος του δελτίου υπάρχει μια διάτρηση μέσα από την οποία περνάει μια μεταλλική βέργα ασφαλείας. Το μεγαλύτερο πλεονέκτημά του είναι η ευελιξία του. Οι νέες αναγραφές προστίθενται εύκολα και γρήγορα στο κατάλογο ενώ με την ίδια ευκολία αφαιρούνται τα δελτία που αναφέρονται σε υλικό το οποίο έχει αποσυρθεί από τη βιβλιοθήκη. Αν και η μετακίνηση του δεν είναι εύκολη μπορεί να χρησιμοποιηθεί από πολλούς αναγνώστες αν δεν χρειάζονται το ίδιο συρτάρι.
    - ✍ Microform : Γενικά ο τύπος του καταλόγου σε μορφή μικροφόρμας παρουσιάζεται σε δύο μορφές : α) Οι αναγραφές μειώνονται στο ελάχιστο και τυπώνονται σε φιλμ. Για να τα συμβουλευτεί κανείς έπρεπε να χρησιμοποιήσει το ανάλογο μηχάνημα που μεγεθύνει το φιλμ και το προβάλλει σε οθόνη. β) Microfiche. Είναι ένας τύπος διαφανούς κάρτας που μπορεί σε σμίκρυνση να δώσει και διακόσιες αναγραφές σε κάθε κάρτα. Το microfiche έχει το πλεονέκτημα της απευθείας πρόσβασης σε μία αναγραφή, ενώ το microfilm απαιτεί μια διαδοχική αναζήτηση της αναγραφής πάνω στο φιλμ.

## 2.2. Βιβλιογραφία

Η βιβλιογραφία σε οποιαδήποτε μορφή κι αν είναι (έντυπη, on-line, CD-Rom) παρέχει τα στοιχεία εκείνα που χρειαζόμαστε για να εντοπίσουμε μια πηγή. Έχουν

γίνει προσπάθειες διατύπωσης ορισμών για τη βιβλιογραφία παρ' όλα αυτά κανείς δεν είναι σε θέση να συγκεράσει όλες τις απόψεις. Για τους περισσότερους είναι ένας κατάλογος βιβλίων που μεταξύ άλλων μας πληροφορεί για τον συγγραφέα ενός βιβλίου, τον εκδότη, τον χρόνο και τον τόπο έκδοσης κ.α. οι βιβλιογραφίες σήμερα δεν περιορίζονται κατά ανάγκη στα βιβλία, αλλά μπορεί να συμπεριλαμβάνουν και άλλους τύπους τεκμηρίων όπως films, φωτογραφίες κ.α.

Είναι σημαντικό μια βιβλιογραφία να καλύπτει μια πραγματική ανάγκη και να μην είναι επανάληψη μιας προηγούμενης. Θα πρέπει να είναι όσο το δυνατόν πιο πλήρης και τα έργα που περιλαμβάνει θα πρέπει να περιγράφονται με ένα συγκεκριμένο βιβλιογραφικό στυλ αναφέροντας τα βασικά στοιχεία μιας βιβλιογραφικής αναγραφής. Η δομή της βιβλιογραφίας πρέπει να είναι τέτοια ώστε να μπορεί να χρησιμοποιηθεί εύκολα, καθώς επίσης να συνοδεύεται από ευρετήρια ώστε η πρόσβαση στα λήμματα να είναι πολλαπλή. (Σταϊκούρα, 1996)

Παρακάτω αναφέρονται οι σημαντικότερες κατηγορίες βιβλιογραφιών.

- **Ανάλογα με το περιεχόμενο:**

- ✍ *Γενικές* : περιλαμβάνουν όλων των ειδών τα τεκμήρια και δεν περιορίζονται χρονικά, θεματικά ή ποσοτικά.

- ✍ *Ειδικές* : δημοσιεύουν τίτλους οι οποίοι αναφέρονται σε ένα θέμα, ή σε ένα πρόσωπο, ή σε ένα τίτλο.

- **Ανάλογα με τον τόπο και τη γλώσσα που είναι γραμμένες:**

- ✍ *Εθνικές* : οι οποίες περιλαμβάνουν τα έργα τα οποία εκδίδονται σε μια χώρα ή τα έργα τα οποία είναι γραμμένα στην επίσημη ή επίσημες γλώσσες μιας χώρας.

- ✍ *Διεθνείς* : περιλαμβάνουν τεκμήρια που χρησιμοποιούν διάφορες γλώσσες.

- **Ανάλογα με το χρόνο**

- ✍ *Τρέχουσες* : Η τρέχουσα βιβλιογραφία παρουσιάζει την πνευματική παραγωγή ενός τόπου αμέσως μετά την έκδοσή της και εκδίδεται κατά τακτά χρονικά διαστήματα.

- ✍ *Αναδρομικές* : Καλύπτουν μια καθορισμένη χρονική περίοδο.

### **2.3. Συστήματα ευρετηρίασης**

Η θεματική επεξεργασία των πληροφοριών που εμπεριέχονται στα τεκμήρια αποτελεί κυρίαρχο πρόβλημα σε κάθε σύστημα ανάκτησης πληροφοριών.

Η αποτελεσματικότητα ενός θεματικού ευρετηρίου ως μέσου για την ταύτιση και την ανάκτηση τεκμηρίων εξαρτάται από μια καλά δομημένη γλώσσα ευρετηρίασης. Αυτό ισχύει για κάθε σύστημα όπου η επιλογή των όρων απαιτεί την ανθρώπινη διανοητική κρίση και απόφαση.

Η ανάγκη απόδοσης των όρων ευρετηρίασης με τρόπο μοναδικό αντιμετωπίστηκε διαχρονικά με τη χρήση μιας γλώσσας, η οποία είχε τη μορφή ταξινομικού συστήματος, δομημένου καταλόγου θεματικών επικεφαλίδων και, πρόσφατα, δομημένου θησαυρού.

Οι τυποποιημένες αυτές γλώσσες ανάκτησης πληροφοριών διευκολύνουν τη σύνδεση εννοιών, (κυρίως λέξεων) για να δημιουργήσουν θεματικούς όρους, που περιγράφουν με ακρίβεια το θεματικό περιεχόμενο των τεκμηρίων.

Η γλώσσα που χρησιμοποιείται για την καταχώριση και στη συνέχεια, την ανεύρεση των πληροφοριών σε ένα πληροφοριακό σύστημα, ονομάζεται γλώσσα ή σύστημα ευρετηρίασης ή ανάκτησης πληροφοριών.

Οι γλώσσες ευρετηρίασης χρησιμοποιούνται τόσο για την ένταξη των τεκμηρίων των συλλογών μιας βιβλιοθήκης κάτω από ενιαίους όρους, όσο και για τη διεξαγωγή της έρευνας για την ανεύρεσή τους. Κύριο χαρακτηριστικό των γλωσσών ευρετηρίασης είναι το ελεγχόμενο λεξιλόγιο στη θεματική απόδοση του περιεχομένου των τεκμηρίων.

Ανάλογα με το πώς χειρίζονται τα σύνθετα θέματα, τα συστήματα ευρετηρίασης μπορούν να διακριθούν σε:

- *Προσυνδυσασμένα συστήματα:* οι όροι ευρετηρίασης συνδυάζονται κατά την ευρετηρίαση. Τα συστήματα αυτά διαχειρίζονται τα σύνθετα θέματα με ολοκληρωμένες θεματικές περιγραφές. Στα προσυνδυσασμένα ευρετήρια, οι σχέσεις μεταξύ των όρων δηλώνονται με σύμβολα, καθένα από τα οποία εκφράζει ειδική σχέση, με τη θέση των όρων στην κάθε αναγραφή, την τυπογραφία και με τη στίξη που τους συνδέει. Η πληροφορία ανακτάται, αν κατά την αναζήτηση η θεματική περιγραφή αποδοθεί ως έχει.

- *Υστεροσυνδυσασμένα συστήματα:* οι όροι δεν συνδυάζονται κατά την ευρετηρίαση, αλλά κατά την αναζήτηση. Οι όροι που αποδίδουν το περιεχόμενο του τεκμηρίου καταχωρούνται ανεξάρτητοι. Ο συνδυασμός των όρων γίνεται κατά τη διάρκεια της αναζήτησης των πληροφοριών με διάφορες τεχνικές. Στα υστεροσυνδυσασμένα συστήματα, το τεκμήριο ανακτάται, εάν όλοι ή ένας από τους όρους αυτούς χρησιμοποιηθούν για την αναζήτησή του.

Στα προσυνδυσασμένα συστήματα υπάγονται τα ταξινομικά συστήματα και οι κατάλογοι θεματικών επικεφαλίδων ενώ στα υστεροσυνδυσασμένα οι θησαυροί. (Τσάφου Σταματίνα- Χατζημαρή Στέλλα)

#### **2.4. Κατάλογοι θεματικών επικεφαλίδων και Θησαυροί**

Το στοιχείο όμως που προσδίδει στη δομή του λεξιλογίου τη μορφή δικτύου σχέσεων και παρέχει στον χρήστη τη δυνατότητα να ξεφύγει από την ομαδοποίηση των όρων και να κινηθεί σε εναλλακτικές διαδρομές είναι η «οριζόντια» διασύνδεση των εννοιών με μη ιεραρχικές σχέσεις (π.χ. ο όρος «πτηνά» έχει ως σχετικό τον όρο «ορνιθολογία». ακολουθώντας τη διαδρομή αυτή, ο χρήστης περνά από την ομάδα ή κατηγορία των «έμβιων όντων» σε εκείνη των «επιστημών»).

Αυτές οι σχέσεις είναι που διασυνδέουν τις έννοιες με τρόπο που παρακάμπτει τις λογικές κατηγοριοποιήσεις και αξιοποιεί λιγότερο προφανείς και, γι' αυτό, πολύτιμες «συνάφειες». Το όχημα για την αναζήτηση της πληροφορίας δεν είναι πλέον ένας ανελκυστήρας μεταξύ ορόφων ενός κτηρίου, αλλά ένας υπόγειος σιδηρόδρομος που μας μεταφέρει γρήγορα και αποτελεσματικά σε άλλο σημείο της ίδιας περιοχής ή και στην άλλη άκρη της πόλης.

Τέτοια δομή διαθέτουν οι *κατάλογοι θεματικών επικεφαλίδων* και οι *θησαυροί*, δύο τύποι γλωσσών ευρετηρίασης που χρησιμοποιούνται ευρέως σε συστηματικά τεκμηριωμένες συλλογές.

Η μεταξύ τους διάκριση δυσχεραίνεται από αυτήν ακριβώς την δομική συγγενειά τους, η οποία εντούτοις είναι σε μεγάλο βαθμό επιφανειακή.

Στην πραγματικότητα, μεγάλες είναι οι διαφορές τους τόσο στο θεωρητικό υπόβαθρο όσο και στην πρακτική εφαρμογή τους, και αυτό οφείλεται κατ' αρχήν στο γεγονός ότι εκπροσωπούν δύο

διαφορετικές προσεγγίσεις στο δίπτυχο ευρετηρίαση-αναζήτηση.

Πιο συγκεκριμένα οι θεματικές επικεφαλίδες:

- δομούνται με βάση τη θεματική περιγραφή συγκεκριμένων τεκμηρίων,
- συχνά αποτελούνται από συνδυασμούς όρων που δηλώνουν μια ευκαιριακή, ή αλλιώς συντακτική, σχέση μεταξύ των εννοιών, η οποία δεν έχει νόημα παρά στο πλαίσιο της ευρετηρίασης του εκάστοτε τεκμηρίου,
- χρησιμοποιούνται στην προσυνδυασμένη ευρετηρίαση.

Ο τυχαίος τρόπος με τον οποίο προστίθενται νέες επικεφαλίδες στον κατάλογο, βάσει των αναγκών της ευρετηρίασης, αφήνει μεγάλα περιθώρια για θεματικά «κενά» ή «επικαλύψεις».

Στους θησαυρούς, οι όροι:

- εκφράζουν αυτόνομες έννοιες που από κοινού καλύπτουν θεματικά συγκεκριμένο εύρος και βάθος,
- διασυνδέονται με σταθερές σημασιολογικές σχέσεις που απορρέουν από τις ίδιες τις έννοιες και δεν εξαρτώνται από κανένα συγκεκριμένο τεκμήριο,
- χρησιμοποιούνται στην υστεροσυνδυασμένη ευρετηρίαση, όπου σε κάθε τεκμήριο αποδίδεται ένας αριθμός ξεχωριστών θεματικών όρων, και μόνο στη φάση της αναζήτησης μπορεί να γίνει συνδυασμός των όρων αυτών.

Ο συγκροτημένος τρόπος με τον οποίο δομούνται εξαρχής οι θησαυροί διευκολύνει την πληρότητα και τη σαφήνεια εννοιών και σχέσεων στο θεματικό τους πλαίσιο.

Η φιλοσοφία των θησαυρών ενισχύει την «οικονομία» της γλώσσας ευρετηρίασης. Σχηματικά, θα λέγαμε ότι η δομική μονάδα στους θησαυρούς είναι η «λέξη», ενώ στους καταλόγους θεματικών

επικεφαλίδων είναι η «φράση». Οι επικεφαλίδες αυξάνονται συνεχώς και με τυχαίο τρόπο προκειμένου να περιγραφούν νέα τεκμήρια· αντίθετα, αλλαγές στο περιεχόμενο θησαυρών γίνονται μόνο στον βαθμό που εμφανίζονται νέες έννοιες ή επέρχονται τροποποιήσεις στις ήδη υπάρχουσες, πάντοτε στο πλαίσιο της θεματικής τους κάλυψης.

Η φαινομενική ομοιότητα των δύο αυτών γλωσσών ευρετηρίασης ως προς τη δομή του πλέγματος των σχέσεων οφείλεται περισσότερο στο γεγονός ότι ορισμένοι κατάλογοι θεματικών επικεφαλίδων (όπως της Βιβλιοθήκης του Κογκρέσου, από τον οποίο προέρχονται και τα παραδείγματα που ακολουθούν) υιοθέτησαν τα τελευταία χρόνια τη σημειογραφία που αναπτύχθηκε για την περιγραφή της δομής των θησαυρών. Παλαιότερα χρησιμοποιούσαν (και πολλοί χρησιμοποιούν ακόμη και σήμερα) τις παραπομπές του τύπου «βλέπε επίσης» για να αποτυπώσουν αδιακρίτως τόσο τις κάθετες όσο και τις οριζόντιες σχέσεις μεταξύ των όρων. Η αλλαγή αποφασίστηκε ακριβώς επειδή διαπιστώθηκε η αδυναμία των συγκεκριμένων παραπομπών να αποδώσουν αυτήν τη θεμελιώδη δομική διάκριση. Ωστόσο, σε μεγάλο βαθμό το εγχείρημα απέτυχε: το σύστημα παραπομπών μεταξύ των επικεφαλίδων, παρά τη χρήση της σημειογραφίας των θησαυρών, εξακολουθεί να μη διαχωρίζει σαφώς τις

κατηγορίες σχέσεων και να μην ακολουθεί τις κατευθυντήριες οδηγίες για τους θησαυρούς.

Ο λόγος της αποτυχίας είναι απλός: από τη στιγμή που κάθε επικεφαλίδα δύναται να απαρτίζεται από περισσότερες της μίας έννοιες, η εφαρμογή της δομής των θησαυρών είναι ουσιαστικά

ανέφικτη. Είναι φανερό ότι, με τους συνδυασμούς των εννοιών, ο καθορισμός μιας ιεραρχικής σχέσης δεν μπορεί παρά να είναι αυθαίρετος.

Αλλά ακόμη και στις σχέσεις μεταξύ απλών εννοιών, επικρατεί ανάλογη σύγχυση. Το αποτέλεσμα είναι ότι συνολικά η δομή του καταλόγου θεματικών επικεφαλίδων χαρακτηρίζεται από έλλειψη συνοχής και τελείως υποκειμενική οργάνωση.

Είναι σημαντικό πάντως να πούμε ότι οι θεματικές επικεφαλίδες πήραν τη μορφή τους υπηρετώντας τις ανάγκες θεματικής πρόσβασης στις κλασικές συλλογές των βιβλιοθηκών και καλύπτοντας αντίστοιχα με ανοικτό τρόπο όλο το φάσμα της γνώσης, ενώ οι θησαυροί εμφανίστηκαν από την αρχή ως εργαλείο ανάκτησης της πληροφορίας στο πλαίσιο προηγμένων συστημάτων πληροφόρησης και οι περισσότεροι από αυτούς αποτελούν το λεξιλόγιο πρόσβασης σε μεγάλες βάσεις δεδομένων επιστημονικού και τεχνικού περιεχομένου. Και είναι χαρακτηριστικό ότι η μετεξέλιξη των λεξιλογίων μέσα από μια διαρκή διερεύνηση των εφαρμογών και της απόδοσής τους είναι σήμερα περισσότερο από ποτέ στο προσκήνιο της επιστήμης της πληροφόρησης.

### ***3. Αναζήτηση και ανάκτηση πληροφοριών στο ψηφιακό περιβάλλον***

Η ανάπτυξη της πληροφορίας σε ηλεκτρονική μορφή, σε συνδυασμό με την ανάπτυξη του διαδικτύου οδήγησαν τις βιβλιοθήκες σε μια νέα ψηφιακή εποχή. Έγιναν πιο προσιτές στο χρήστη, ανέπτυξαν πολλές υπηρεσίες και εμφανίστηκαν τα πρώτα δείγματα απομακρυσμένης πρόσβασης.

Οι κατάλογοι των βιβλιοθηκών ανέβηκαν στο διαδίκτυο και έγιναν προσβάσιμοι σε όλους, χωρίς να απαιτείται η φυσική τους παρουσία στη βιβλιοθήκη για την αναζήτηση και την ανάκτηση τεκμηρίων. Οι παράγοντες αλλαγής ωθούν τις βιβλιοθήκες να αναθεωρήσουν τις προτεραιότητές τους και να θέσουν νέους στόχους στην εξυπηρέτηση των πληροφοριακών αναγκών των χρηστών. Ανταποκρινόμενες στα νέα τεχνολογικά δεδομένα και συμπεριφορές, επεκτείνουν την κλασική πληροφοριακή εξυπηρέτηση μέσα στο χώρο της βιβλιοθήκης σε ένα διαφορετικό, εικονικό και ψηφιακό επίπεδο. Απαραίτητο θεωρείται πια να ικανοποιηθούν οι απαιτήσεις των χρηστών με τη χρήση των νέων διαδικτυακών τεχνολογιών επικοινωνίας ώστε να προσαρμοστεί η βιβλιοθήκη στο νέο περιβάλλον αλλά και να το διαμορφώσει ως προς το συμφέρον της γνώσης και της έρευνας.

Οι βιβλιοθήκες βρίσκονται σε μια διαδικασία συνολικού επανακαθορισμού το ρόλου τους και λήψης αποφάσεων για το είδος και τη μορφή των υπηρεσιών που θα προσφέρουν στους χρήστες τους στο νέο περιβάλλον που ορίζει η ψηφιακή τεχνολογία και το Διαδίκτυο.

Οι βιβλιοθήκες όλων των τύπων σήμερα προσφέρουν πρόσβαση σε όλο και περισσότερες πηγές όπου οι χρήστες μπορούν να έχουν πρόσβαση απομακρυσμένα. Αυτό είναι το μοντέλο της βιβλιοθήκης που έχει επικρατήσει τα τελευταία χρόνια και αποκαλείται Υβριδικό.

Η ανάπτυξη της τεχνολογίας των πληροφοριών και επικοινωνιών επέφερε ριζικές αλλαγές στα συστήματα αναζήτησης και ανάκτησης πληροφοριών. Η ανάπτυξη του Διαδικτύου επηρεάζει όλους τους τομείς της βιβλιοθήκης και ιδιαίτερα τις υπηρεσίες που προσφέρονται στους χρήστες με τη χρήση νέων εργαλείων και μεθόδων παροχής πληροφόρησης. Η ραγδαία εξέλιξη του Παγκόσμιου Ιστού προσφέρει ένα πολυδύναμο μέσο για τις βιβλιοθήκες για σημαντικές αλλαγές στην ποιότητα και ποσότητα των παρεχόμενων υπηρεσιών προς την κοινότητα των



χρηστών τους. Η εκτεταμένη χρήση του Διαδικτύου διεύρυνε τόσο τις μεθόδους που χρησιμοποιούσε κάθε βιβλιοθήκη για την παροχή των υπηρεσιών της όσο και την εξοικείωση των χρηστών με στρατηγικές αναζήτησης και ανάκτησης πληροφοριών.

Το Διαδίκτυο έχει χαρακτηριστεί ως η ουσία της σύγχρονης ζωής αφού μέσω της άμεσης διάχυσης της πληροφορίας σε όλο το φάσμα της ανθρώπινης δραστηριότητας γίνεται το εργαλείο της μετάβασης σε μια νέα μορφή κοινωνίας, την διαδικτυακή κοινωνία ή κοινωνία της πληροφορίας (“network society”). Οι χρήστες και το προσωπικό των βιβλιοθηκών δρουν μέσα σε ένα πληροφοριακό περιβάλλον που καθορίζεται από τον Παγκόσμιο Ιστό σε όλες του τις εκφάνσεις.

Με την εμφάνιση πλήθους ψηφιακών υπηρεσιών ο αριθμός χρηστών που αναζητούν απομακρυσμένη πρόσβαση στις υπηρεσίες τις βιβλιοθήκης αυξήθηκε ραγδαία. σε αντιδιαστολή με τους φυσικούς χρήστες που συνεχώς μειώνονται.

Η πρόσβαση στο Διαδίκτυο είναι σχετικά εύκολη καθώς ο χρήστης δεν χρειάζεται να έρθει στο φυσικό χώρο της βιβλιοθήκης για να ζητήσει βοήθεια και καθοδήγηση αλλά χρησιμοποιεί τον ηλεκτρονικό κατάλογο και τις μηχανές αναζήτησης. Δρα απρόσωπα, αυτόνομα και ανεξάρτητα και εκμεταλλεύεται τον διαθέσιμο όγκο δεδομένων του Ιστού θεωρώντας ότι η διαδικασία της αναζήτησης έχει πια απλουστευτεί, ακόμα και αν υποψιάζεται πως δεν είναι τόσο αξιόπιστος και έγκυρος όσο το παραδοσιακό έντυπο μέσο. Αναπτύσσει παράλληλα μια αντίληψη για τις βιβλιοθήκες και τις υπηρεσίες πληροφόρησης, που τις περιγράφει όλο και λιγότερο ως φυσικούς χώρους αναζήτησης πληροφορίας και όλο και περισσότερο ως εικονικούς διαύλους πρόσβασης σε αυτήν.

### **3.1. Διαδίκτυο**

Η προσπάθεια κάλυψης αναγκών για πληροφόρηση μέσα από το διαδίκτυο μπορεί να είναι μία πολύπλοκη διαδικασία ή να αποτελέσει μία απλή αναζήτηση ανάλογα με το είδος της πληροφορίας και το λόγο για τον οποίο ο χρήστης χρειάζεται αυτή την πληροφορία. Υπάρχουν άπειρες πληροφορίες στο διαδίκτυο, αλλά η πρόσβαση σε αυτές που είναι συναφείς κάθε φορά είναι το ζητούμενο αποτελεί μία σύνθετη διαδικασία. Παρ όλο που η χρήση του διαδικτύου κάθε μέρα γίνεται πιο φιλική, ωστόσο δεν είναι τόσο εύκολο όσο το να ανοίξουμε την τηλεόραση ή το βίντεο. Μέρος της λύσης αποτελεί να γνωρίζουμε όσο πιο αναλυτικά είναι δυνατόν το είδος των πληροφοριών που θα εντοπίσουμε στο διαδίκτυο.

Οι βιβλιοθηκονόμοι για να μπορέσουν να συνεχίσουν να παρέχουν πληροφοριακές υπηρεσίες στους χρήστες θα πρέπει να αξιοποιήσουν στο έπακρο τις δυνατότητες που τους δίνονται και το διαδίκτυο να αποτελέσει σημαντικό μέρος των πληροφοριακών πηγών που διαθέτουν. Γενικότερα για να εντοπίσει ο χρήστης πληροφορίες στο διαδίκτυο μπορεί:

- Να οδηγηθεί κατ ευθείαν σε μία σελίδα, αν γνωρίζει τη διεύθυνση (URL)
- Να ξεφυλλίσει έναν θεματικό οδηγό (Subject Directory)
- Να εκτελέσει μία αναζήτηση χρησιμοποιώντας μία μηχανή αναζήτησης (Search Engine)
- Να εκτελέσει μία αναζήτηση σε μία μηχανή πολλαπλής αναζήτησης (Metasearch Engine)
- Να συμμετέχει σε μία από τις μορφές επικοινωνίας

Αυτή τη στιγμή υπάρχει στενή σχέση ανάμεσα στα διάφορα εργαλεία αναζήτησης ώστε οι περισσότερες πληροφορίες να μπορούν να προσεγγιστούν και με ένα θεματικό κατάλογο (Subject tree ή directory) και με μία μηχανή αναζήτησης. Ο

χρήστης του διαδικτύου μπορεί να χρησιμοποιήσει ένα θεματικό οδηγό, αν η αναζήτηση του αφορά γενικό θέμα, ενώ θα πρέπει να χρησιμοποιήσει μία μηχανή αναζήτησης ή μία μηχανή πολλαπλής αναζήτησης αν το θέμα είναι συγκεκριμένο. Στην πρώτη περίπτωση η αναζήτηση γίνεται ιεραρχικά, αν και σε ορισμένες περιπτώσεις υπάρχει η δυνατότητα χρήσης λέξεων- κλειδιών, ενώ στη δεύτερη χρησιμοποιούνται λέξεις- κλειδιά και σε πολλές περιπτώσεις και αναπτυγμένες τεχνικές ανάκτησης. Γενικότερα μπορεί να πει κανείς ότι ο χρήστης πρέπει να χρησιμοποιήσει ένα θεματικό κατάλογο όταν:

- Το θέμα της αναζήτησης είναι γενικό ή υπάρχει μία γενική ιδέα
- Θέλει να δει μία κατάσταση με ιστοσελίδες σχετικές με το θέμα οι οποίες προτείνονται και σχολιάζονται από ειδικούς
- Θέλει να ανακτήσει μία κατάσταση με ιστοσελίδες σχετικές με το θέμα, αντί για πολλές ξεχωριστές σελίδες
- Θέλει να αναζητήσει τους τίτλους ιστοσελίδων, σχολιασμό και αν υπάρχει η δυνατότητα να χρησιμοποιηθούν λέξεις- κλειδιά για να ανακτηθούν συναφή θέματα
- Θέλει να αποφύγει να ανακτήσει τεκμήρια που έχουν μικρή σχέση με το θέμα της αναζήτησης, κάτι που είναι συνηθισμένο στις μηχανές αναζήτησης

Ενώ ο χρήστης πρέπει να χρησιμοποιήσει μηχανές αναζήτησης όταν:

- Το θέμα της αναζήτησης είναι συγκεκριμένο
- Θέλει να εντοπίσει μία συγκεκριμένη ιστοσελίδα
- Θέλει να διεξάγει αναζήτηση στο πλήρες κείμενο πολλών σελίδων
- Θέλει να ανακτήσει ένα μεγάλο αριθμό τεκμηρίων
- Θέλει να αναζητήσει συγκεκριμένους τύπους τεκμηρίων
- Θέλει να αξιοποιήσει τις σύγχρονες τεχνικές ανάκτησης

Ωστόσο ο βιβλιοθηκονόμος πρέπει να έχει υπόψη του ότι κατά τη διαδικασία αναζήτησης εμφανίζονται σημαντικά προβλήματα:

- *Έλλειψη ελέγχου καθιέρωσης*: Στο διαδίκτυο τα τεκμήρια ευρετηριάζονται με όποιο τρόπο αποφασίσει ο κάθε δημιουργός, χωρίς να συμβουλευεται καθιερωμένες επικεφαλίδες
- *Έλλειψη ελέγχου ποιότητας*: Ο κάθε χρήστης μπορεί να αποκτήσει τη δική του σελίδα στο διαδίκτυο χωρίς κανέναν έλεγχο σχετικό με την ποιότητα
- *Εφήμερη φύση των πηγών του διαδικτύου*: Πολλές σελίδες δημιουργούνται καθημερινά, αλλά κανείς δεν γνωρίζει για πόσο διάστημα θα συνεχίσουν να συντηρούνται και να ενημερώνονται
- *Τεχνική αστάθεια και περιορισμοί*: Συχνά τεχνικά προβλήματα εμποδίζουν την αποτελεσματική αναζήτηση στο διαδίκτυο (Κορομπίλη- Ξαντινίδου Στέλλα, 2000)

### **3.1.1. Άλλες υπηρεσίες του Διαδικτύου**

Υπάρχουν και άλλες υπηρεσίες του διαδικτύου που προσπαθούν να διευκολύνουν το χρήστη στην αναζήτηση πληροφοριών. Τέτοιες είναι:

- Οι ηλεκτρονικές βιβλιοθήκες (Virtual libraries) και
- Τα δελτία ενημέρωσης ( Newsletters) που αποτελούν ευρετήρια των πηγών του διαδικτύου χρησιμοποιώντας κριτήρια αξιολόγησης για να περιληφθούν ή όχι στη βάση

Οι αξιολογημένες βάσεις αποτελούν ένα πολύ χρήσιμο βοήθημα για τους χρήστες που δεν έχουν ούτε το χρόνο, αλλά ούτε και την ειδικότητα να αξιολογήσουν και να οργανώσουν τις πηγές του διαδικτύου. Οι βιβλιοθήκες υπήρξαν

αρχικά οι φορείς εκείνοι που οργάνωσαν μία τέτοια υπηρεσία, έχοντας στο νου τους φοιτητές και το εκπαιδευτικό προσωπικό του Ιδρύματος. Οι υπηρεσίες αυτές, όμως, δεν διαθέτουν μία ενιαία πολιτική αξιολόγησης, δεν υπάρχει καν συμφωνία στο περιεχόμενο των όρων αξιολόγησης που χρησιμοποιούν. Τέτοια αξιολόγηση προσπάθεια θεωρείται η LIBRARIAN'S INDEX TO THE INTERNET (<http://www.lii.org/>).



Η μηχανή αυτή περιλαμβάνει πηγές που καλύπτουν κυρία τις ανάγκες για πληροφόρηση χρηστών λαϊκών βιβλιοθηκών. Οι χρήστες δεν κάνουν έρευνα στις σελίδες του διαδικτύου, αλλά στις αξιολογημένες πηγές της Library's Index to the Internet, οι οποίες αποτελούν εγγραφές της βάσης. Ο χρήστης μπορεί να ξεφυλλίσει την κατάσταση με τα θέματα που διατηρεί η βάση. Επίσης μπορεί να επιλέξει να εκτελέσει την αναζήτηση/ ανάκτηση με:

- **Λέξεις- κλειδιά:** Η αναζήτηση εκτελείται σε λέξεις του πεδίου του τίτλου, της περιγραφής και των θεμάτων
- **Θέματα:** Η αναζήτηση εκτελείται στο πεδίο των θεμάτων των εγγραφών
- **Τίτλο:** Η αναζήτηση εκτελείται στο πεδίο του τίτλου
- **Περιγραφή:** Η αναζήτηση εκτελείται σε λέξεις του πεδίου της περιγραφής
- **Σύνδεσμοι:** Η αναζήτηση εκτελείται σε URLs που περιλαμβάνονται στα πεδία του τίτλου και της περιγραφής

Οι τεχνικές ανάκτησης που χρησιμοποιεί η απλή αναζήτηση επιτρέπουν:

- Την εξ ορισμού χρήση του τελεστή AND χωρίς να πληκτρολογηθεί
- Χρήση του τελεστεί OR και NOT
- Χρήση του συμβόλου (\*) για αποκοπή
- Χρήση παρενθέσεων για ομαδοποίηση των όρων

Η αναπτυγμένη αναζήτηση κάνει χρήση των ίδιων τεχνικών όπως η απλή, αλλά επιτρέπει τη δυνατότητα περιορισμού της αναζήτησης σε πεδία:

- ~au: Αναζήτηση στο πεδίο του τίτλου
- ~de: Αναζήτηση στο πεδίο της περιγραφής
- ~in: Αναζήτηση στα αρχικά αυτού που ευρετηριάζει
- ~kw: Αναζήτηση στα πεδία του τίτλου, θεμάτων και της περιγραφής
- ~li: Αναζήτηση σε λέξεις των URLs
- ~pu: Αναζήτηση στο πεδίο του εκδότη
- ~su: Αναζήτηση σε λέξεις του πεδίου των θεμάτων
- ~ti: Αναζήτηση στο πεδίο του τίτλου
- ~tu: Αναζήτηση στο URL του τίτλου

Τα κριτήρια επιλογής που χρησιμοποιούνται για την εισαγωγή πηγών στη Librarian's Index to the Internet, όπως αναφέρονται στη σελίδα της βάσης είναι:

- **Περιεχόμενο:** Αν η πηγή είναι ελεύθερη, ή με συνδρομή. Είναι η πληροφορία ακριβής; Οι πληροφορίες είναι συγκεκριμένα γεγονότα ή γνώμες και απόψεις; Η σελίδα συμβάλει στο θέμα σημαντικά; Πόσο επίκαιρη και τρέχουσα είναι η πληροφόρηση; Πόσο συχνή είναι η ενημέρωση; Αν υπάρχουν σύνδεσμοι σε άλλες σελίδες, κατά πόσο λειτουργούν, ή παραπέμπουν σε σωστές σελίδες. Είναι το

περιεχόμενο καλογραμμένο;

- *Μνεία ευθύνης*: Ποια είναι η φήμη του συγγραφέα και τα προσόντα του σε σχέση με το θέμα που καλύπτει; Υπάρχει η δυνατότητα επικοινωνίας με το συγγραφέα για διευκρινίσεις ή για ενημέρωση;
- *Σκοπός*: Ο σκοπός της πηγής δηλώνεται καθαρά; Είναι η σελίδα κατάλληλη για το κοινό που προορίζεται; Πώς κρίνετε σε σχέση με άλλες σελίδες που αναφέρονται στο ίδιο θέμα;
- *Σχεδιασμός*: Είναι οργανωμένη η σελίδα ώστε να βρίσκει ο χρήστης την πληροφορία που θέλει εύκολα; Απαιτείται ειδικό λογισμικό ή υλικό; Μπορεί να φορτωθεί στον υπολογιστή του χρήστη σε σχετικά σύντομο χρόνο; (Κορομπίλη-Ξαντινίδου Στέλλα, 2000)

### 3.2. Μηχανές αναζήτησης

Μία μηχανή αναζήτησης είναι ένα εργαλείο που σου επιτρέπει να εξερευνήσεις τις βάσεις δεδομένων που περιέχουν το κείμενο δεκάδων εκατομμυρίων σελίδων Web. Όταν το λογισμικό της μηχανής αναζήτησης εντοπίζει σελίδες που ταιριάζουν στην ερώτηση της έρευνας ενός χρήστη, τις παρουσιάζει με σύντομες περιγραφές και συνδέσμους που σε οδηγούν αλλού. Οι μηχανές αναζήτησης συνεχώς ανανεώνονται και περιλαμβάνουν τεράστιο αριθμό πληροφοριών. Αν και διαφέρουν μεταξύ τους ως προς το πως λειτουργούν, ωστόσο όλες αποτελούνται από τρία μέρη:

- Μία αράχνη (spider)
- Μία βάση δεδομένων, η οποία περιέχει όλες τις πληροφορίες που συλλέγει η αράχνη και
- Ένα εργαλείο αναζήτησης, το οποίο οι χρήστες χρησιμοποιούν για να πραγματοποιήσουν αναζήτηση στις βάσεις δεδομένων

Η αράχνη είναι ένα αυτοματοποιημένο πρόγραμμα που αναζητάει στο διαδίκτυο νέες ιστοσελίδες και τοποθετεί τις διευθύνσεις τους μαζί με πληροφορίες για το περιεχόμενό τους σε μία βάση δεδομένων, στην οποία μπορεί να υπάρξει πρόσβαση μέσω μηχανής αναζήτησης. Η κάθε μηχανή αναζήτησης έχει διαφορετικό τρόπο ευρετηρίασης των πληροφοριών. Μερικές ευρετηριάζουν κάθε λέξη του τεκμηρίου, άλλες ευρετηριάζουν μόνο εκατό λέξεις- κλειδιά σε κάθε ένα τεκμήριο και άλλες μόνον τον τίτλο. Επίσης κάθε μηχανή αναζήτησης εμφανίζει τα αποτελέσματα με διαφορετικό τρόπο. Μερικές δείχνουν τη συνάφεια των τεκμηρίων, μερικές τις πρώτες προτάσεις του τεκμηρίου, μερικές τον τίτλο του τεκμηρίου όπως και το URL. Οι μηχανές αναζήτησης διαφέρουν και ως προς τη μέθοδο που χρησιμοποιούν για να μετρήσουν τη συνάφεια των κειμένων προς το ζητούμενο. Για αποτελεσματική αναζήτηση/ ανάκτηση στις μηχανές προτείνεται:

- Χρήση μοναδικών όρων
- *Διεκπεραίωση της έρευνας σε περισσότερα από ένα στάδια*: Ο χρήστης θα πρέπει να εξετάσει με προσοχή τα πρώτα αποτελέσματα και να εντοπίσει τους όρους που χρησιμοποιούνται, ώστε να επαναλάβει την έρευνα χρησιμοποιώντας αυτούς τους όρους
- *Επανάληψη της έρευνας στα αποτελέσματα*: Αν τα αποτελέσματα παρέχουν χρήσιμες πληροφορίες, αλλά είναι πολλές οι σελίδες που ανακτήθηκαν, τότε προτείνεται η επανάληψη της αναζήτησης, μόνο σε αυτά τα αποτελέσματα προσθέτοντας έναν καινούργιο όρο

- Αναζήτηση του μοναδικού όρου στον τίτλο της σελίδας Web
- Έλεγχος αν γίνεται διάκριση πεζών/ κεφαλαίων
- Χρήση κεφαλαίων του πρώτου γράμματος κυρίων ονομάτων: Αν ο χρήστης ενδιαφέρεται για το όνομα ενός ατόμου, για μία γεωγραφική τοποθεσία, ένα βιβλίο, τότε συνιστάται να γράφει ο όρος με κεφάλαιο το πρώτο γράμμα, ώστε να αποφευχθούν πολλές εσφαλμένες ανακτήσεις
- Έλεγχος της ορθογραφίας: Αν ο χρήστης έχει χρησιμοποιήσει τους σωστούς όρους και η μηχανή αναζήτησης καταλήγει σε μηδέν αποτελέσματα, τότε ο χρήστης θα πρέπει να ελέγξει την ορθογραφία πριν κάνει οτιδήποτε άλλο (Κορομπίλη-Ξαντινίδου Στέλλα, 2000)

### 3.3. Μετα-μηχανές αναζήτησης

Με τόσες μηχανές αναζήτησης είναι επόμενο ότι θα υπάρχουν και μηχανές που ευρετηριάζουν άλλες μηχανές αναζήτησης. Είναι οι μηχανές μετα- αναζήτησης, οι οποίες εκτελούν την έρευνα σε διάφορες μηχανές αναζήτησης και εμφανίζουν τα αποτελέσματα. Οι μηχανές μετά- αναζήτησης κάνουν έρευνα σε αρκετά ευρετήρια του Web, σε οδηγούς, σε ομάδες συζητήσεων, αλλά και σε συγκεκριμένες ιστοσελίδες. Υπάρχουν θετικά και αρνητικά στη χρήση τέτοιων μηχανών.

Η διαδικασία αναζήτησης σε κάθε μία μηχανή χωριστά απαιτεί χρόνο, ενώ οι μηχανές μετά αναζήτησης έλυσαν αυτό το πρόβλημα εκτελώντας μία αναζήτηση σε πολλές μηχανές συγχρόνως. Οι περισσότερες από αυτές τις μηχανές αφαιρούν τις διπλοεγγραφές και εμφανίζουν μία κατάσταση με τα συνολικά αποτελέσματα. Σε κάθε εγγραφή εμφανίζεται σύνδεσμος στην αντίστοιχη ιστοσελίδα. Επίσης αυτές οι μηχανές μπορούν να ανακτήσουν μερικές από τις μοναδικές έγγραφες από κάθε μία βάση. Σημαντικό είναι επίσης ότι ο χρήστης δεν χρειάζεται να γνωρίζει τη γλώσσα εντολών της κάθε μηχανής αναζήτησης. Μερικές από τις παράλληλες μηχανές έχουν τη δυνατότητα να ταξινομούν τα αποτελέσματα με ορισμένα στοιχεία, όπως ανάλογα με τη μηχανή αναζήτησης, ή τις λέξεις- κλειδιά, ή ακόμη και την ημερομηνία των εγγραφών. Επειδή μία μηχανή μετα- αναζήτησης μπορεί να εντοπίσει πολλές εγγραφές, ανακτά δέκα ή δεκαπέντε εγγραφές από κάθε βάση και τις συνδυάζει με τις εγγραφές από τις άλλες βάσεις. Τέλος, όπως έδειξαν αρκετές έρευνες, δεν υπάρχει μεγάλη επικάλυψη ανάμεσα στις διάφορες μηχανές αναζήτησης και από αυτή την άποψη η χρήση μίας μηχανής μετα- αναζήτησης μπορεί να ανακτήσει πολύ σημαντικά αποτελέσματα.

Ωστόσο οι μηχανές αυτές δεν παρέχουν τη δυνατότητα χρήσης αναπτυγμένων τεχνικών αναζήτησης, όπως χρήση των τελεστών Boole, αναζήτηση σε συγκεκριμένα πεδία, χρήση των τελεστών εγγύτητας. Επίσης μερικές μεγάλες μηχανές, ή οδηγοί δεν περιλαμβάνονται στις μηχανές μετά- αναζήτησης. Η χρήση μίας μηχανής αναζήτησης έχει σαν αποτέλεσμα να αυξηθεί η ακρίβεια στην ανάκριση. Τέλος έχοντας σαν δεδομένο ότι δεν περιλαμβάνονται όλες οι μηχανές αναζήτησης, με τη χρήση μηχανών μετα- αναζήτησης, δεν έχουμε συχνά πρόσβαση σε πολύ πρόσφατη, ή εξειδικευμένη πληροφορία.

Καμία από τις δύο μεθόδους, δεν είναι ιδανική. Αν το θέμα είναι συγκεκριμένο, η καλύτερη αντιμετώπιση είναι να ξεκινήσει ο χρήστης από μία μηχανή μετα- αναζήτησης η οποία θα φέρει σε λιγότερο χρόνο τα επιθυμητά αποτελέσματα. Αν πάλι ο χρήστης δεν γνωρίζει με βεβαιότητα ότι μία μεγάλη μηχανή αναζήτησης θα ανακτήσει αυτό που αναζητά, τότε είναι προτιμότερη η

χρήση της μιας μηχανής αναζήτησης. Μηχανή μετα- αναζήτησης είναι η MetaCrawler (<http://www.metacrawler.com>).



Τα θέματα που καλύπτει αυτή η μηχανή μετα- αναζήτησης είναι:

- Υπολογιστές
- Εκπαίδευση
- Ψυχαγωγία
- Οικονομικά
- Υγεία
- Νέα
- Αγορές
- Σπορ
- Ταξίδια

Η MetaCrawler δίνει τη δυνατότητα στο χρήστη να επιλέξει αν:

- Θα εκτελεστεί η έρευνα σε μερικούς από τους όρους της αναζήτησης, σε όλους τους όρους της αναζήτησης, ή σε φράσεις
- Τα αποτελέσματα θα περιοριστούν σε συγκεκριμένες περιοχές (domains), ή σε συγκεκριμένες χώρες
- Η αναζήτηση θα γίνει γρήγορα ή θα γίνει πιο εμπεριστατωμένη
- Τον αριθμό των αποτελεσμάτων που θα εμφανίζονται σε κάθε σελίδα
- Το μέγιστο αριθμό αποτελεσμάτων που θα ανακτηθούν από κάθε μηχανή
- Τα αποτελέσματα θα είναι ταξινομημένα ανάλογα με το ποσοστό συνάφειας, ή ανάλογα με την τοποθεσία, ή την πηγή

Η MetaCrawler έχει τις δικές της τεχνικές ανάκτησης:

- Χρήση εισαγωγικών για λέξεις που θα πρέπει να βρίσκονται μαζί
- Χρήση συμβόλων(+) και (-) για να περιληφθούν η να αποκλειστούν λέξεις (Κορομπίλη- Ξαντινίδου Στέλλα, 2000)

### 3.4. Συλλογές μηχανών αναζήτησης

Υπάρχουν μηχανές αναζήτησης που περιλαμβάνουν καταστάσεις με μηχανές αναζήτησης και πληροφορίες για το πώς λειτουργούν, πώς είναι ευρετήριασμένες. Μία από αυτές είναι η Collection of Special Search Engines (<http://www.leidenuniv.nl/locatieub/>).



Αποτελεί μία συλλογή με συνδέσμους σε μηχανές αναζήτησης και οδηγούς η οποία σε πολλές περιπτώσεις μπορεί να χρησιμοποιηθεί σαν εναλλακτική λύση για έρευνα αντί για τις μεγάλες μηχανές αναζήτησης. Επίσης περιλαμβάνονται σελίδες οι οποίες είναι χρήσιμες για την εκμάθηση της αναζήτησης στο www, μία συλλογή από

λεξικά με συνώνυμα ή θησαυρούς, καθώς και μίας σειράς υπηρεσιών που προσφέρονται δωρεάν για έρευνα. (Κορομπίλη- Ξαντινίδου Στέλλα, 2000)

### **3.5. Οδηγοί- Κατάλογοι**

Οι οδηγοί ή κατάλογοι προσπαθούν να οργανώσουν τα τεκμήρια χρησιμοποιώντας ένα σύστημα ταξινόμησης. Υπάρχουν γενικοί κατάλογοι και θεματικοί κατάλογοι, που καλύπτουν διάφορες θεματικές κατηγορίες. Ένας από τους πιο γνωστούς και ολοκληρωμένους γενικούς κατάλογους στο διαδίκτυο είναι ο Yahoo. Ο καλύτερος τρόπος για να εντοπίσει ο χρήστης θεματικούς οδηγούς είναι να χρησιμοποιήσει τη βάση Clearinghouse for Subject- Oriented Internet Resource Guides.

Αποτελεί έναν κατάλογο των σελίδων του Web τις οποίες οργανώνει σε θεματικές κατηγορίες, όπως ο χρυσός οδηγός. Είναι ο πιο ολοκληρωμένος και ενημερωμένος κατάλογος, ένα ιεραρχικό θεματικό ευρετήριο και μία από τις μεγαλύτερες βάσεις με συνδέσμους σε πηγές, οι οποίες προστίθενται καθημερινά. Είναι μία καλή αρχή όταν ο χρήστης δεν ξέρει που ακριβώς να ψάξει και μπορεί να εντοπίσει μερικά καλά αρχεία για αρχή, τα οποία μπορεί να χρησιμοποιήσει για να συνεχίσει την αναζήτηση του σε πιο συγκεκριμένα θέματα. Μπορεί να βρει πληροφορίες για μεγάλη γκάμα θεμάτων και είναι χρήσιμη πηγή για έρευνα σε βάθος αλλά και για παροχή σύντομων πληροφοριακών υπηρεσιών. (Κορομπίλη- Ξαντινίδου Στέλλα, 2000)

### **3.6. Βάσεις δεδομένων**

Τα δεδομένα μπορούν να οριστούν σαν γράμματα ή αριθμοί που αποθηκεύονται σε έναν υπολογιστή και αντιπροσωπεύουν οντότητες (entities). Οι βάσεις δεδομένων (databases) είναι μία συλλογή από τέτοια δεδομένα, η οποία υπάρχει με σκοπό να παρέχει πληροφορίες. Τα δεδομένα μαζί με το σύστημα που τα διαχειρίζεται ονομάζεται σύστημα βάσης δεδομένων (database system). Στους παραπάνω ορισμούς θα πρέπει να προσθέσουμε δύο βασικά κριτήρια:

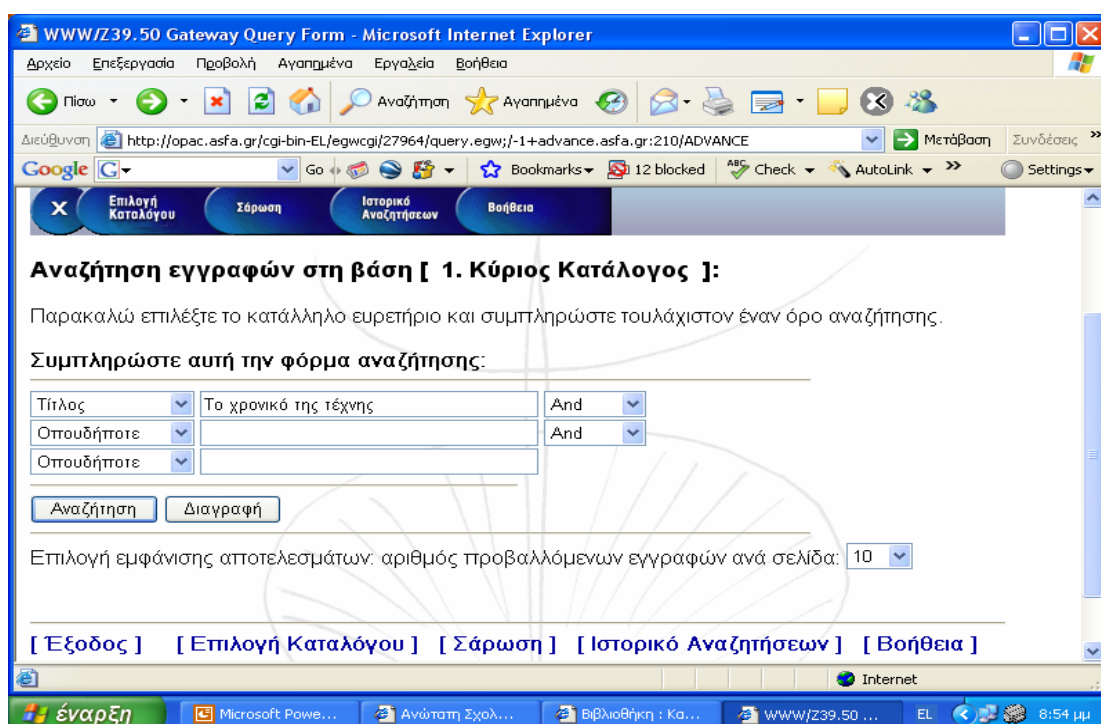
- Κάθε πληροφορία έχει μία μοναδική τιμή
- Η βάση δεδομένων πρέπει να αποτελείται από συγγενείς πληροφορίες

Όταν λένε ότι μία πληροφορία ή ένα δεδομένο έχει μία μοναδική τιμή, εννοούμε ότι δεν μπορεί να διασπαστεί σε μικρότερα κομμάτια χωρίς να χάσει σημαντικό νόημά. Είναι σημαντικό να διαχωρίσουμε τον όρο "βάση δεδομένων", που είναι η συλλογή δεδομένων, από τον όρο "σύστημα βάσης δεδομένων", στο οποίο όλα τα μέρη του συνεργάζονται για να γίνει συλλογή, διαχείριση και μετάδοση πληροφοριών (Κορομπίλη Στέλλα, 1995). Υπάρχουν πολλά είδη συστημάτων βάσεων δεδομένων:

- Το σύστημα διαχείρισης δεδομένων (DBMS)
- Το σύστημα ανάκτησης πληροφοριών (information retrieval system/ IRS), στο οποίο ανήκουν και οι εμπορικές βάσεις δεδομένων
- Το υπερκείμενο σύστημα (hypertext system)
- Το έμπειρο σύστημα βάσης δεδομένων (expert database system) και
- Το σύστημα σύνθετης βάσης δεδομένων (multimedia database system)

### 3.7. Αυτοματοποιημένος κατάλογος

Αυτοματοποιημένος κατάλογος (OPAC) : Υποστηρίζεται από ηλεκτρονικό υπολογιστικό σύστημα. Απετέλεσε αναπόφευκτο παράγωγο της εξέλιξης των υπολογιστών και της υποδομής δικτύωσης τους και βοήθησε καταλυτικά στη βελτίωση των παρεχόμενων υπηρεσιών της βιβλιοθήκης προς τους αναγνώστες της. Η πρόσβαση στο συγκεκριμένο κατάλογο γίνεται μέσω τερματικών τοποθετημένων στον χώρο της βιβλιοθήκης, ή και από απόσταση μέσω του διαδικτύου. Ο αυτοματοποιημένος κατάλογος αντικατέστησε σταδιακά τον δελτιοκατάλογο από τις αρχές της δεκαετίας του 80 ενώ σήμερα αποτελεί βασικό τμήμα των ολοκληρωμένων συστημάτων αυτοματοποίησης των λειτουργιών των βιβλιοθηκών.



Ο κατάλογος αυτός αποτελεί το σημείο πρόσβασης των χρηστών στο σύστημα, θέτοντας βιβλιογραφικά ερωτήματα για κάθε είδους υλικό που περιλαμβάνεται στη βάση. Στο παρελθόν με τη χρήση των δελτιοκαταλόγων είχε κανείς την πιθανότητα να βρει την πηγή που αναζητούσε αν γνώριζε: το επώνυμο του συγγραφέα, τουλάχιστον την πρώτη λέξη του τίτλου, τουλάχιστον την πρώτη λέξη του θέματος ή της σειράς ή του συλλογικού οργάνου. Οι όροι αναζήτησης στον αυτοματοποιημένο κατάλογο περιλαμβάνουν τις περισσότερες από τις λέξεις των τίτλων (εκτός από άρθρα, συνδέσμους, προθέσεις κ.α.) και των θεματικών επικεφαλίδων, με αποτέλεσμα να δίνεται η δυνατότητα στους χρήστες αλλά και στους βιβλιοθηκονόμους, να αναζητούν με πολλαπλούς τρόπους την επιθυμητή πληροφορία, χωρίς να χρειάζεται να απομνημονεύουν οπωσδήποτε τα αναγκαία βιβλιογραφικά στοιχεία.(Κατσιρίκου,2003)



### 3.8. Crawlers

Οι crawlers ή τα ρομπότ συνήθως έχουν ως αφετηρία ιστοχώρους με μεγάλη επισκεψιμότητα από χρήστες ή ιστοχώρους που ταξινομούν πολλές περιοχές πληροφοριών που ονομάζονται και *Πύλες*. Βρίσκοντας ένα μεγάλο ποσό από υπερσυνδέσεις τα προγράμματα αναζήτησης εξαπλώνονται στον Κυβερνοχώρο με μεγάλο ρυθμό ανάπτυξης.

Ένα παράδειγμα αποτελεί η ευρέως και πλέον σύγχρονη αυτόματη M.A. Google, η οποία ξεκίνησε ως ένα πανεπιστημιακό ερευνητικό πρόγραμμα. Το σύστημα χρησιμοποιεί τρία αυτόματα προγράμματα κάθε φορά τα οποία με την σειρά τους είναι συνδεδεμένα με περίπου 300 συνδέσεις Εντοπιστών Ομοιόμορφων Πόρων (URL, Uniform Resource Locator) σε ιστοσελίδες. Τα URL είναι μία ειδικά διαμορφωμένη ταινία κειμένου που προσδιορίζει την τοποθεσία στο Ίντερνετ. Οι ιστοσελίδες χρησιμοποιούν τα URLs για να συνδεθούν με άλλες σελίδες. Εκτός από αυτό τα URLs περιέχουν και πολλές άλλες δημοφιλείς δικτυακές εντολές όπως δείκτες προς FTP αρχεία, μηνύματα Usenet κτλ. Το URL καθορίζει τον τύπο του πόρου (όπως FTP, HTTP, ή Gopher) που θα προσπελαστεί, τη συγκεκριμένη θέση όπου είναι αποθηκευμένες οι πληροφορίες, και το ακριβές σημείο όπου βρίσκονται οι πληροφορίες μέσα στη θέση.

Περιέχει τρία τμήματα:

- δικτυακό πρωτόκολλο (network protocol), που χρησιμοποιείται για την πρόσβαση στην τοποθεσία του διαδικτύου.
- όνομα και διεύθυνση του φιλοξενούμενου (host name and address), που ακολουθεί τον προσδιορισμό του ανωτέρω πρωτοκόλλου. Οι φιλοξενούμενοι μπορούν να προσδιορίζονται από την πρότυπη ονοματοδοσία του διαδικτύου (Internet-standard naming) ή από τη διεύθυνση του διαδικτυακού πρωτοκόλλου (IP address).
- τοποθεσία φακέλου (file location), που καθορίζει την τοποθεσία του πόρου του δικτύου. Πόροι (resources) είναι φάκελοι που μπορεί να είναι φάκελοι κειμένου, έγγραφα, γραφικά ή προγράμματα. Τα ονόματα των πόρων (resource names) είναι σχετικά με ένα τοπικό διευθυνσιολόγιο.

Τεχνικά, ένα URL, περιέχει μία τεκμαρτή τοποθεσία φακέλου την οποία οι διακομιστές του διαδικτύου (Web servers), όπως ο Apache, μετατρέπουν αυτόματα σε ένα όνομα ειδικού φακέλου. Όλοι οι άλλοι ειδικοί φάκελοι υπάρχουν σε μία ιεραρχία ή δενδρικό διευθυνσιολόγιο όπως το επόμενο:

<http://compnetworking.about.com/library/glossary/blglossary/htm>.

Στην μέγιστη απόδοση λειτουργίας του συστήματος αυτού, χρησιμοποιούνται τέσσερα αυτόματα προγράμματα και “σαρώνονται” πάνω από 100 ιστοσελίδες το δευτερόλεπτο, δημιουργώντας έτσι γύρω στα 600 Kb δεδομένων το δευτερόλεπτο.

### 3.9. Γλώσσες υποβολής ερωτημάτων

Η αναζήτηση στις ηλεκτρονικές πηγές πραγματοποιείται με τη βοήθεια των γλωσσών υποβολής ερωτημάτων (Query languages). Μερικά παραδείγματα τέτοιων γλωσσών αποτελούν οι:

- SQL – Structured Query Language για αναζήτηση σε σχεσιακές βάσεις δεδομένων
- XQuery για την αναζήτηση σε XML έγγραφα

- SPARQL - Protocol and RDF Query Language για την αναζήτηση σε RDF γράφους

Κάθε γλώσσα υποβολής ερωτημάτων χαρακτηρίζεται από συγκεκριμένες δυνατότητες. Παρακάτω αναφέρονται συνοπτικά σε τρεις κατηγορίες οι δυνατότητες τους αποδεικνύοντας πως οι χρήστες έχουν στη διάθεσή τους ισχυρά εργαλεία για τη δόμηση απλών αλλά και σύνθετων ερωτημάτων με στόχο τη βελτιωμένη ανάκτηση πληροφοριών.

Έτσι, οι χρήστες μπορούν να υποβάλουν τα ερωτήματά τους με τις παρακάτω μορφές:

- Λέξεις κλειδιά (Keyword-based querying):
  - ✍ Μία μόνο λέξη
  - ✍ Context queries (Αναζήτηση με φράση)
  - ✍ Συνδυασμός λέξεων και χρήση τελεστών Boole
  - ✍ Φυσική γλώσσα (π.χ. Τι είναι η τεχνητή νοημοσύνη?)
- Ταίριασμα προτύπων (Pattern matching):
  - ✍ Λέξεις
  - ✍ Μια συμβολοσειρά, η οποία πρέπει να αποτελεί λέξη σε ένα κείμενο
  - ✍ Προθέματα
  - ✍ Επιθέματα
  - ✍ Υποσυμβολοσειρές (substrings)
  - ✍ Διαστήματα (ranges)
  - ✍ Ερωτήματα με λάθη (allowing errors)
  - ✍ Κανονικές εκφράσεις (regular expressions) και
  - ✍ Επεκτάσεις προτύπων (extended patterns)
- Ερωτήματα δομής (Structural queries)
  - ✍ Σταθερή μορφή (fixed structure)
  - ✍ Τα τεκμήρια είναι δομημένα σε πεδία π.χ. φόρμες και η αναζήτηση πραγματοποιείται με τη βοήθεια προτύπων (patterns) σε συγκεκριμένα πεδία
  - ✍ Δομή με τη μορφή υπερκειμένου (hypertext)
  - ✍ Ιεραρχική δομή (hierachical structure)

Οι δυνατότητες των γλωσσών αναζήτησης είναι αμέτρητες, γεγονός το οποίο έχει συγχρόνως θετικές και αρνητικές επιπτώσεις για τους χρήστες κατά την αναζήτηση και ανάκτηση πληροφοριών. Είναι βέβαιο, πως μέσα από την μεγάλη γκάμα τεχνικών αναζήτησης οι χρήστες έχουν τη δυνατότητα σύνθετης αναζήτησης με στόχο την καλύτερη ανάκτηση πληροφοριών. Ωστόσο, το πλήθος των δυνατοτήτων αυτών καθιστά για το μέσο χρήστη ένα ιδιαίτερο πρόβλημα διότι προκειμένου να ανακτήσει τις πληροφορίες που επιθυμεί πρέπει να έχει προηγουμένως μάθει τον τρόπο με τον οποίο κάθε μηχανή αναζήτησης ή βάση δεδομένων δέχεται την υποβολή των ερωτήσεων. Το πρόβλημα αυτό, ωστόσο, παύει να υφίσταται στην περίπτωση της φυσικής γλώσσας, διότι ο χρήστης δεν είναι αναγκασμένος να μάθει τεχνικές, τελεστές και τρόπους αναζήτησης αλλά υποβάλει το ερώτημά του σε φυσική γλώσσα, π.χ. τι είναι η τεχνητή νοημοσύνη;

Χαρακτηριστικό παράδειγμα μηχανής αναζήτησης, η οποία απαντά σε ερωτήσεις που υποβάλλονται σε φυσική γλώσσα αποτελεί η START η οποία είναι βασισμένη σε σύστημα επεξεργασίας φυσικής γλώσσας. Η START αναλύει τα εισερχόμενα ερωτήματα, αντιστοιχίζει τα ερωτήματα που έχουν πλέον αναλυθεί σε δένδροειδή μορφή στα περιεχόμενα της βάσης και παρουσιάζει τα κατάλληλα τμήματα πληροφορίας στον χρήστη. Ακολουθώντας την παραπάνω τεχνολογία, η START παρέχει άμεσα πληροφορίες σε ανεκπαιδευτους χρήστες εξοικονομώντας τους πολύτιμο χρόνο και χρήμα (START, χ.χ).

Ορισμένοι θα μπορούσαν να ισχυριστούν πως μια τέτοια μηχανή αναζήτησης θα μπορούσε να αντικαταστασεί τον ανθρώπινο παράγοντα, στη συγκεκριμένη περίπτωση τον βιβλιοθηκονόμο, ο οποίος σε ένα συμβατικό περιβάλλον θα απαντούσε το πληροφοριακό ερώτημα του χρήστη. Ωστόσο, τα αποτελέσματα που επιστρέφονται από ένα σύστημα και έναν άνθρωπο διαφέρουν ουσιαστικά. Παραδείγματος χάριν στο ερώτημα που θέσαμε στη μηχανή αναζήτησης START “Give me journal articles on artificial intelligence” το αποτέλεσμα ήταν “Unfortunately, I don't have that information” σε αντίθεση με έναν βιβλιοθηκονόμο, ο οποίος θα μας παρέπεμπε σε περιοδικά σχετικά με την τεχνητή νοημοσύνη.

### **3.10. XML**

Η γενικού χαρακτήρα γλώσσα σήμανσης XML (eXtensible Markup Language) δημιουργήθηκε για την διευκόλυνση της ανταλλαγής και της κοινής χρήσης δεδομένων από διαφορετικά πληροφοριακά συστήματα και ειδικότερα για να καλύψει τις ανάγκες του Διαδικτύου (Bray, 2006). Η προσθήκη σημασιολογικών περιορισμών οδήγησε στην δημιουργία γλωσσών εφαρμογής, όπως η MathML για την περιγραφή μαθηματικής σήμανσης, η GraphML για τα γραφήματα και η MusicXML για την σήμανση μουσικών εγγράφων.

Ο W3C πρότεινε την χρήση του προτύπου RDF (Resource Description Framework) για την περιγραφή του περιεχομένου ψηφιοποιημένων φωτογραφιών, ώστε να είναι αναζητήσιμες και ανακτήσιμες μέσω του HTTP πρωτοκόλλου (Lafon, 2002). Αυτό το σύστημα απαρτίζεται από τρία βασικά τμήματα:

- Την σύρωση της φωτογραφίας και την διάσωσή της ως αρχείου JPEG.
- Την δημιουργία RDF μεταδεδομένων και την αποθηκευσή τους μέσα στο ίδιο το αρχείο.
- Την χρήση εξυπηρετητών που υποστηρίζουν HTTP για την μεταφορά της φωτογραφίας ή των μεταδεδομένων της σε ένα πελάτη.

### **3.11. Πρωτόκολλα**

#### **3.11.1. Z39.50**

Το Z39.50 είναι ένα δικτυακό πρωτόκολλο το οποίο επιτρέπει αναζήτηση (συνήθως απομακρυσμένων) ετερογενών βάσεων δεδομένων και ανάκτηση δεδομένων μέσω ενός ενδιάμεσου περιβάλλοντος διεπαφής του χρήστη (user interface). Συνήθως χρησιμοποιείται για ανάκτηση βιβλιογραφικών εγγράφων, αν και υπάρχουν επίσης μη-βιβλιογραφικές εφαρμογές.

Ο χρήστης (ή ο Βιβλιοθηκονόμος) επικοινωνεί μέσα από το οικείο περιβάλλον του δικού του συστήματος με οποιοδήποτε σύστημα βιβλιοθήκης. Τα βασικά οφέλη του είναι:

- Αποστολή ερωτήσεων (ίδιων ή/και διαφορετικών) από Z-client σε πολλές βιβλιοθήκες ταυτόχρονα
- Ενιαίος τρόπος ερώτησης και παρουσίασης αποτελεσμάτων
- Δεν υπάρχει ανάγκη διαφορετικού client για κάθε σύστημα
- Αξιοποίηση πολυάριθμων και τεράστιων βάσεων δεδομένων
- Εξοικονόμηση χρόνου, κόπου εκμάθησης διαφορετικών λογισμικών

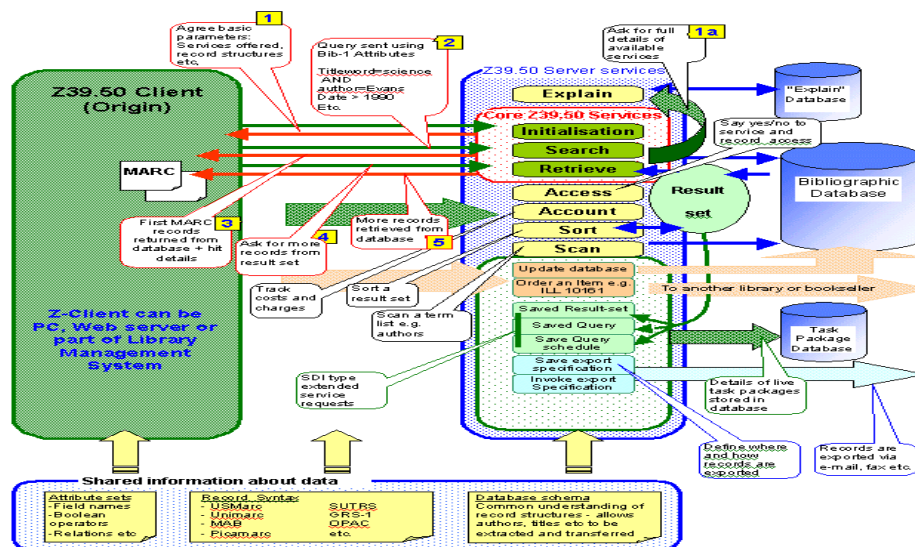
- Αξιοποίηση μορφότυπου ανταλλαγής δεδομένων MARC. Παρουσίαση αποτελεσμάτων και δυνατότητα περαιτέρω επεξεργασίας (καταλογογράφηση βιβλίων με έτοιμες εγγραφές από άλλες βιβλιοθήκες)
- Τυποποίηση βασικών λειτουργιών αναζήτησης και ανάκτησης
- Εκτεταμένες υπηρεσίες (Extended services) για παραγγελία εγγράφων (ordering documents), ενημέρωση βάσεων δεδομένων, αποθήκευση στρατηγικών αναζήτησης.
- Άνοιγμα λειτουργιών βιβλιοθήκης (ILL, δανεισμός υλικού. Προκρατήσεις, ανανεώσεις δανεισμών κλπ)

Τα χαρακτηριστικά αναζήτησης του:

- Δυνατότητες πολύπλοκων ερωτήσεων με χρήση τελεστών Boolean (AND, OR, NOT)
- Σύγκριση τελεστών ημερομηνιών π.χ. Μεγαλύτερο από, ίσο με κλπ.
- Αναζήτηση με τελεστές γειννίασης (proximity searching)
- Αποκοπή χαρακτήρων (Truncation)
- Τελεστές πληρότητας πεδίων π.χ. Μέρος πεδίου, πλήρες πεδίο, αρχή κλπ.

Οι εφαρμογές του στις Βιβλιοθήκες

- Εξασφάλιση βιβλιογραφικών εγγράφων (Bibliographic record sourcing)
- Συλλογικοί Κατάλογοι (Union Catalogues)
- Διαδανεισμός Βιβλιοθηκών (Inter-Library Loan)
- Επιλεκτική Διάθεση Πληροφοριών (Selective Dissemination of Information)
- Εμπορικές βάσεις δεδομένων
- Αναζήτηση και διύλιση του web (Web Searching and filtering)
- Ενημέρωση βάσεων δεδομένων
- Προσωπικά Βιβλιογραφικά εργαλεία (Personal bibliographic tools)



### 3.11.2. OAI

Τα αρχικά OAI του πρωτοκόλλου OAI-PMH προέρχονται από το Open Archives Initiative. Πρόκειται για έναν οργανισμό που δημιουργεί και προωθεί πρότυπα διαχείρισης που στοχεύουν στη διευκόλυνση της αποτελεσματικής παροχής

περιεχομένου. Το OAI αποτελεί εγχείρημα του Cornell University, αλλά είναι ευρέως αποδεκτό και υποστηριζόμενο από διάφορους οργανισμούς, όπως το Ίδρυμα Ψηφιακών Βιβλιοθηκών(Digital Library Foundation).

Το OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting) είναι ένα απλό πρωτόκολλο που επιτρέπει την συγκέντρωση και μεταφορά εγγραφών μεταδεδομένων από κάποιο σύστημα σε κάποιο άλλο. Οι παροχείς περιεχομένου χρησιμοποιούν το OAI-PMH για να κάνουν τις πληροφορίες, δηλαδή τα μεταδεδομένα τους διαθέσιμα σε τρίτους. Μπορούν να θέσουν περιορισμούς σχετικά με το ποιοι μπορούν να συλλέξουν τις εγγραφές μεταδεδομένων τους, με βάση είτε τις IP διευθύνσεις των παροχών υπηρεσιών, είτε πιο πολύπλοκους μηχανισμούς όπως τα πρωτόκολλα HTTP Basic Authentication και SSL.

Σύμφωνα με την ορολογία που χρησιμοποιεί το OAI-PMH, ένας παροχέας δεδομένων (data provider) κάνει τα δεδομένα διαθέσιμα για συγκομιδή και ένας παροχέας υπηρεσιών(service provider) συγκεντρώνει τα μεταδεδομένα αυτά και τα καθιστά διαθέσιμα για αναζήτηση. Σύμφωνα με το μοντέλο πελάτη – εξυπηρετητή, ο παροχέας δεδομένων είναι ένας εξυπηρετητής και ο παροχέας υπηρεσιών ένας πελάτης.

Το OAI-PMH χρησιμοποιεί τον OAI aggregator, έναν παροχέα υπηρεσιών και δεδομένων. Πρόκειται για μία υπηρεσία που συγκεντρώνει εγγραφές μεταδεδομένων από διάφορους παροχείς πληροφοριών και έπειτα καθιστά τις εγγραφές αυτές διαθέσιμες για συγκέντρωση από τρίτους που χρησιμοποιούν το πρωτόκολλο OAI-PMH. Το OAI-PMH βασίζεται σε πρότυπα του διαδικτύου, όπως το HTTP και η XML, κάτι που το καθιστά εύκολο στην εγκατάσταση σε οποιονδήποτε εξυπηρετητή διαδικτύου. Ο όρος «πηγή πληροφοριών» OAI (OAI repository) αναφέρεται στην ουσία σε μια βάση δεδομένων που χρησιμοποιεί το πρωτόκολλο OAI-PMH και αποτελεί, έναν παροχέα δεδομένων OAI. Αυτή η πηγή πληροφοριών μπορεί να αποθηκεύει είτε κείμενο και μεταδεδομένα είτε μόνο μεταδεδομένα (όπως για παράδειγμα ένα κατάλογο θεμάτων).

Το OAI-PMH ver. 1.0 στηρίζεται κατά κύριο λόγο στη συμφωνία της Santa Fe, στις συναντήσεις της Digital Library Federation (Ψηφιακή Ομοσπονδία Βιβλιοθηκών), καθώς και σε εργασίες που έλαβαν χώρα στο Πανεπιστήμιο του Cornell.

Τα κύρια χαρακτηριστικά του είναι:

- ότι προσδιορίζει ένα χαμηλό επίπεδο διαλειτουργικότητας, το οποίο δεν εξυπηρετούσε σε μεγάλο βαθμό τους χρήστες.
- ότι το μοντέλο συγκομιδής συγκεντρώνει και καθιστά τα μεταδεδομένα, διαθέσιμα για αναζήτηση μέσω δύο κατηγοριών που συμμετέχουν στη συγκομιδή των μεταδεδομένων με βάση το OAI και είναι:
  - ✍ οι διαχειριστές των συστημάτων που υποστηρίζουν το OAI, OAI-PMH ως μέσω έκθεσης των μεταδεδομένων, οι γνωστοί ως προμηθευτές στοιχείων και
  - ✍ οι φορείς παροχής υπηρεσιών, οι οποίοι χρησιμοποιούν τα μεταδεδομένα που συγκεντρώνονται μέσω του OAI-PMH ως βάση για τις υπηρεσίες.
- ότι βασίζεται στα πρωτόκολλα HTTP τα οποία εκφράζονται με τα ίδια αιτήματα μέσω μιας χαρακτηριστικής εφαρμογής που χρησιμοποιεί ένα κεντρικό υπολογιστή δικτύου που διαμορφώνεται για να αποστείλει τα αιτήματα OAI και OAI-PMH στο λογισμικό που ήδη χρησιμοποιούν.
- Και τέλος βασίζεται στο XML δηλαδή σε ένα κωδικοποιημένο αρχείο μεταδεδομένων, το οποίο απαντά σε ένα αίτημα OAI, OAI-PMH. Το OAI

χρησιμοποιεί XML Schemas για να ορίσει το format των εγγράφων. Μπορεί να ανταλλάξει μεταδεδομένα σε οποιαδήποτε format αρκεί να είναι κωδικοποιημένα σε XML και να ορίζεις ένα XML Schemas για αυτά τα format.

### 3.11.3. SRU

Το SRU αποτελεί πρωτόκολλο το οποίο βασίζεται στις διαδικτυακές υπηρεσίες για την επερώτηση βάσεων δεδομένων και την επιστροφή αποτελεσμάτων. Με απλά λόγια είναι ένα πρωτόκολλο αναζήτησης και ανάκτησης που χρησιμοποιεί της υπηρεσίες του Internet και του Ιστού για να μεταφέρει μηνύματα μεταξύ client και server.

Το SRU έχει έναν σημαντικό πρόγονο, το Z39.50 το οποίο είναι διαδεδομένο παγκοσμίως. Το μεγαλύτερο μέρος των λειτουργιών του SRU απορρέει από το παλαιότερο πρωτόκολλο, παρ' όλα αυτά μόνο τα σημαντικότερα μεταφέρθηκαν στο νέο, σε απλούστερη μορφή. Τον καιρό που δημιουργήθηκε το SRU αρκετά ιδρύματα - κυρίως το προσωπικό της Βασιλικής Βιβλιοθήκης στην Ολλανδία (Royal Library in the Netherlands) – μελετούσαν τη χρήση του URL σε υπηρεσίες αναζήτησης. Μία διεθνής ομάδα ειδικών από την κοινότητα του Z39.50 συνεργάστηκαν στο σχέδιο του νέου πρωτοκόλλου για το περιβάλλον Internet/Web/XML. Οι προδιαγραφές SRU εκδόθηκαν για πρώτη φορά το 2002 και η χρήση τους είναι αρκετά δημοφιλής σε νέες χρήσεις λόγω της εύκολης εφαρμογής τους.

Το SRU είναι πολύ εύκαμπτο, καθώς βασίζεται σε XML και η συχνότερη εφαρμογή του είναι η SRU μέσω URL, που χρησιμοποιεί την HTTP GET για την μεταφορά μηνυμάτων. Άλλες εκδόσεις όμως μπορούν να λειτουργήσουν μέσω του πρωτοκόλλου SOAP (SRU μέσω SOAP), που υποστηρίζει περισσότερα χαρακτηριστικά του ιστού και μέσω HTTP POST (SRU POST), που αποφεύγει τους περιορισμούς στο μήκος και το είδος των χαρακτήρων που παρουσιάζονται στο HTTP GET. Οι εγγραφές που παρουσιάζονται ως απάντηση σε μια αναζήτηση μπορεί να παρουσιαστούν σε οποιαδήποτε σωστά-ορισμένη XML τυποποίηση.

Το λειτουργικό μοντέλο του SRU, για αναζήτηση σε διαφορετικές ιστοσελίδες είναι κοινό με το μοντέλο που χρησιμοποιείται για το πρωτόκολλο Z39.50. Ο τελικός χρήστης δημιουργεί ένα αίτημα αναζήτησης από το σύστημα του σπιτιού του, το οποίο έχει ένα συγκεκριμένο συντακτικό τοπικής αναζήτησης. Για την αναζήτηση στο σύστημα του server, με το διαφορετικό συντακτικό αναζήτησης, τη διαφορετική σχεδίαση της βάσης δεδομένων και την διαφορετική ευρετηρίαση, το τοπικό αίτημα αναζήτησης του τελικού χρήστη μετατρέπεται σε τυποποιημένη μορφή. Η σελίδα του server δέχεται το πρότυπο μήνυμα αναζήτησης, το οποίο αποτελείται από το πρωτόκολλο και την ερώτηση, και το μεταφράζει στη σύνταξη που μπορεί να καταλάβει η βάση δεδομένων του συστήματός του. Αν η ερώτηση της αναζήτησης είναι πολύ λεπτομερής ή οι παράμετροι της αναζήτησης δεν υποστηρίζονται από τον server, τότε ο αυτός μπορεί να απορρίψει την αναζήτηση ως μη υποστηριζόμενη ή σε μερικές περιπτώσεις να αγνοήσει και να αλλάξει τα χαρακτηριστικά που δεν γνωρίζει και να συνεχίσει με την αλλαγμένη αναζήτηση.

Το πρωτόκολλο που συνοδεύει την ερώτηση, ορίζει για τον server: την μορφή των ανακτώμενων εγγράφων, τον αριθμό των αρχείων που επιστρέφουν κατά την αναζήτηση καθώς και πολλές άλλες πληροφορίες. Από την στιγμή που ένα από τα

βασικά θέματα στην αναζήτηση μέσα σε ετερογενή συστήματα είναι η διαφορά στην ευρετηρίαση των διαφορετικών σελίδων, οι κατασκευαστές του SRU σχεδίασαν – για την βελτίωση της ποιότητας της αναζήτησης μέσα σε αυτά – μία λειτουργία, που επιτρέπει στον χρήστη να ρωτήσει το σύστημα τι υποστηρίζει, περιμένοντας από το σύστημα να ανταποκριθεί στο αίτημα αυτό με καθορισμένο τρόπο, όπως άλλωστε ορίζει και το πρωτόκολλο. Το σύστημα του χρήστη, χρησιμοποιώντας τις πληροφορίες που αποκομίζει από το σύστημα ευρετηρίασης του server, προσδοκείται ότι θα διατυπώσει μία πιο επιτυχημένη αναζήτηση.

Το SRU ορίζει τρεις βασικές λειτουργίες: Την *επεξήγηση* (explain) , την *σάρωση* (scan) και την *αναζήτηση-ανάκτηση* (SearchRetrieve). Αναλυτικότερα:

- *Η λειτουργία Επεξήγησης*: είναι αίτηση από τους πελάτες για πληροφόρηση σχετικά με τις βάσεις του εξυπηρετητή (π.χ. θέση, περιεχόμενο και χαρακτηριστικά που πρωτοκόλλου που υποστηρίζει ο εξυπηρετητής). Πιο συγκεκριμένα, είναι η λειτουργία που περιγράφηκε πρωτύτερα και επιτρέπει στο server να ενημερώνει τον client τι μπορεί να κάνει. Ο server έχει την δική του περιγραφή π.χ. σε πρότυπο XML η οποία μπορεί εύκολα να ανακτηθεί από τον client μέσω του αιτήματος επεξήγησης και η περιγραφή αυτή μπορεί να βοηθήσει τον client να σχηματοποιήσει αναζητήσεις. Το αρχείο επεξήγησης έχει όπως έχει ήδη αναφερθεί, έναν αριθμό περιγραφικών πεδίων όπως π.χ. τύπου XML που μπορούν να χρησιμοποιηθούν για ανάκτηση αρχείων ή για τον αριθμό των αρχείων που μπορούν να δοθούν από το σύστημα. Οι server SRU δεν χρειάζεται να διατηρούν αρχείο επεξήγησης, αλλά παρακινούνται να το κάνουν.
- *Η λειτουργία Σάρωσης*: Η σάρωση SRU επιτρέπει στον client να απαριθμήσει τους όρους που βρίσκονται στο ευρετήριο της απομακρυσμένης βάσης δεδομένων, μία διαδικασία συναφής με τη φυλλομέτρηση του πίνακα περιεχομένου ενός βιβλίου. Έτσι ο χρήστης μπορεί να δει πόσα αποτελέσματα θα είχε κατά την αναζήτηση με κάθε όρο. Η σάρωση συχνά χρησιμοποιείται από τον client για να επιλέξει όρους για διαδοχική αναζήτηση ή να επαληθεύσει οπτικά μία αρνητική αναζήτηση.
- *Η λειτουργία Αναζήτησης/Ανάκτησης*: είναι η πιο βασική λειτουργία του SRU. Παρέχει τα μέσα για την επερώτηση μίας απομακρυσμένης βάσης δεδομένων και την ανάσυρση αποτελεσμάτων αναζήτησης. Οι επερωτήσεις στις βάσεις δεδομένων πρέπει να διατυπωθούν χρησιμοποιώντας την Κοινή Γλώσσα Επερώτησης (Common Query Language – CQL), η οποία επιτρέπει τόσο απλές όσο και σύνθετες αναζητήσεις με τελεστές Boole. Τα αποτελέσματα της λειτουργίας αναζήτησης και ανάκτησης μπορούν να επιστραφούν σε οποιοδήποτε αριθμό τύπων, όπως σε MARC, XML, MARCXML, MODS ή ακόμα και DC. Τέλος η κατάταξή τους είναι ανάλογη με την σχετικότητα

Το SRU χρησιμοποιείται καλύτερα έναντι των λογικά δομημένων καταλόγων, όπου ο χρήστης επιθυμεί τη δυνατότητα περισσότερο δομημένης αναζήτησης και περισσότερο ελέγχου των αποτελεσμάτων. Οι ερευνητές που είναι εξοικειωμένοι με έναν τοπικό κατάλογο βιβλιοθήκης π.χ. θα βρουν το SRU περισσότερο χρήσιμο για τις ομοιότητές του με την τοπική αναζήτηση. Μία εφαρμογή του πρωτοκόλλου SRU σε συνδυασμό με τα πρωτόκολλα Z39.50 και OAI παρατηρείται από την Ευρωπαϊκή Βιβλιοθήκη (European Library). Η Ευρωπαϊκή Βιβλιοθήκη αποτελεί μία υπηρεσία όλων των Ευρωπαϊκών Εθνικών Βιβλιοθηκών για την πρόσβαση στους καταλόγους και τις ψηφιακές συλλογές τους, μέσω μίας κεντρικής πύλης (portal).

### **3.12. Συστήματα αναζήτησης και ανάκτησης πολυμεσικού υλικού**

Μέχρι πρότινος η έρευνα στην ανάκτηση πληροφοριών είχε επικεντρωθεί αποκλειστικά στην αναζήτηση και ανάκτηση τεκμηρίων με τη μορφή κειμένου (text). Στο σύγχρονο περιβάλλον όμως, η πληροφορία μπορεί να έχει τη μορφή εικόνας, ήχου ή βίντεο. Πλέον, η έρευνα στην ανάκτηση πληροφοριών έχει στραφεί προς το πολυμεσικό υλικό και αφορά ουσιαστικά στην αναζήτηση και ανάκτηση εικόνων, ήχου (ομιλία, μουσική και ήχοι από το γενικότερο περιβάλλον) και βίντεο, το οποίο αποτελεί ένα συνδυασμό των δύο παραπάνω (εικόνα + ήχο). Στις επόμενες ενότητες θα αναπτύξουμε με περισσότερες λεπτομέρειες τους τρόπους αναζήτησης και ανάκτησης κάθε κατηγορίας πολυμεσικού υλικού, ξεκινώντας από τις εικόνες και καταλήγοντας στο βίντεο.

#### **3.12.1. Συστήματα ανάκτησης εικόνων**

Τα συστήματα ανάκτησης εικόνων επιτρέπουν την αναζήτηση και ανάκτηση δισδιάστατων, τριδιάστατων, σταθερών (still) και κινούμενων (moving) εικόνων από μεγάλες βάσεις δεδομένων. Ωστόσο, πρέπει να έχει προηγηθεί η καταλογογράφηση των εικόνων αυτών, η οποία ουσιαστικά αφορά στη χειροκίνητη ή αυτόματη προσθήκη μεταδεδομένων για κάθε αρχείο εικόνας και εκφράζεται συνήθως με την λεκτική περιγραφή του περιεχομένου της. Δηλαδή, αποδίδονται λέξεις – κλειδιά, λεζάντες, ή άλλο συνοδευτικό κείμενο που περιγράφουν το περιεχόμενο της κάθε εικόνας (Pavel, 2001). Σε αυτή την περίπτωση η αναζήτηση πραγματοποιείται επίσης με λέξεις κλειδιά και ανακτώνται οι εικόνες των οποίων η λεκτική περιγραφή ταιριάζει περισσότερο με το ερώτημα που έχει θέσει ο χρήστης.

Τα πιο σύγχρονα συστήματα αναζήτησης και ανάκτησης εικόνων βασίζονται πλέον στο ίδιο το περιεχόμενο της εικόνας (content based image retrieval). Συγκεκριμένα, γίνεται αυτόματη εξαγωγή των στοιχείων εκείνων που αφορούν το σχήμα, το χρώμα, και την υφή των εικόνων, ενώ η αναζήτηση σε αυτά τα συστήματα πραγματοποιείται συνήθως με τη βοήθεια των μεθόδων που είδαμε σε παραπάνω ενότητα και αφορούν κυρίως το query by example (Goodrum, 2000).

#### **3.12.2. Συστήματα ανάκτησης ήχου**

Τα συστήματα ανάκτησης ήχου αφορούν στην αναζήτηση και ανάκτηση ομιλίας, μουσικής και ήχων που προέρχονται από το γενικότερο περιβάλλον (π.χ. ήχοι που προκαλούνται από αντικείμενα ή ζώα). Προκειμένου να είναι αναζητήσιμα τα ηχητικά τεκμήρια έχουν υιοθετηθεί διάφορες προσεγγίσεις. Ο πιο κοινός τρόπος αναζήτησης των τεκμηρίων αυτών είναι με λέξεις-κλειδιά εφόσον έχει προηγηθεί κάποιο είδος λεκτικής περιγραφής. Στην περίπτωση που το τεκμήριο περιλαμβάνει και ομιλία, η τεχνολογία της αυτόματης αναγνώρισης φωνής (automatic speech recognition), μειώνει τα προβλήματα ανάκτησης, καθώς η ομιλία μετατρέπεται σε κείμενο και η αναζήτηση πραγματοποιείται πλέον στο μεταγεγραμμένο κείμενο (Foote, 1999). Επίσης, υπάρχει η δυνατότητα αναγνώρισης των διαφορετικών ομιλιτών βάσει των παραλλαγών στη φωνή διευκολύνοντας με αυτό τον τρόπο την



ανάκτηση συγκεκριμένης ομιλίας από ένα μακρύ σε διάρκεια ηχητικό τεκμήριο, π.χ. δημόσιος διάλογος πολιτικών.

Οι συλλογές μουσικής αποτελούν μια από τις πιο δημοφιλείς κατηγορίες πολυμεσικού υλικού, η διακίνηση των οποίων έχει λάβει τεράστιες διαστάσεις στο χώρο του διαδικτύου. Αυτό πρακτικά σημαίνει πως, η περιγραφή βασισμένη σε λέξεις-κλειδιά δεν επαρκεί, ιδιαίτερα για τους χρήστες, οι οποίοι μπορεί να μην γνωρίζουν τα μεταδεδομένα με τα οποία έχει περιγραφεί ένα μουσικό κομμάτι, αλλά να γνωρίζουν πολύ περισσότερες πληροφορίες για το περιεχόμενο, τους στίχους, την ενορχήστρωση, τις μελωδίες, τους ρυθμούς ή τις αρμονίες του. Για την ικανοποίηση των παραπάνω χρηστών χρειάζονται νέοι τρόποι περιγραφής και ισχυρά εργαλεία αναζήτησης και ανάκτησης μουσικών πληροφοριών.

Σήμερα, ο τομέας της ανάκτησης μουσικής πληροφορίας επικεντρώνεται κυρίως στην αποτελεσματική βασισμένη στο περιεχόμενο (content-based) αναζήτηση και ανάκτηση από τοπικές και on-line βάσεις δεδομένων. Οι πρώτες βασισμένες στο περιεχόμενο μουσικές βάσεις δεδομένων επικεντρώνονται στη μελωδία. Χαρακτηριστικό παράδειγμα αποτελεί το Meldex, το οποίο παρέχεται από το University of Waikato της Νέας Ζηλανδίας και επιτρέπει στους χρήστες του την αναζήτηση σε μια βάση δεδομένων, η οποία περιέχει περισσότερες από 9.000 Ευρωπαϊκές και Κινέζικες παραδοσιακές μελωδίες παίζοντας απλώς ο χρήστης μια μελωδία σε ένα εικονικό πληκτρολόγιο που εμφανίζεται στην ιστοσελίδα του. Αμέσως εντοπίζονται τα μελωδικά όμοια αποσπάσματα και επιστρέφονται στο χρήστη υπό τη μορφή καταλόγου.

Ένας άλλος τρόπος αναζήτησης και ανάκτησης μουσικών τεκμηρίων, είναι αυτός που επιτρέπει στους χρήστες την ταυτοποίηση τραγουδιών μέσω της αναπαραγωγής έστω μερικών δευτερολέπτων ήχου μέσω του κινητού τους τηλεφώνου. Πρόκειται για την υπηρεσία μουσικής αναγνώρισης, η οποία βοηθάει τους ενδιαφερόμενους χρήστες να ταυτοποιήσουν συγκεκριμένες μουσικές ηχογραφήσεις μέσω της αναγνώρισης ενός δείγματος ήχου διάρκειας δέκα δευτερολέπτων σταλμένο μέσω τηλεφώνου ακόμα και από χώρους με πολύ θόρυβο.

Αν και οι τεχνολογίες αναζήτησης και ανάκτησης μουσικών τεκμηρίων έχει εξελιχθεί ιδιαίτερος, απαιτείται πρόοδος στο διαχωρισμό των πηγών (εξαγωγή μεμονωμένων πηγών ήχου από μία ηχογράφηση που ταυτοχρόνως περιέχει πολλαπλές πηγές ήχου) και την αναγνώριση και ταυτοποίηση των πηγών αυτών (μεμονωμένη φωνή, μουσικό όργανο ή οποιοσδήποτε άλλος ήχος μιας ηχογράφησης).

Τέλος, η σύγχρονη τάση της έρευνας σχετικά με την αναζήτηση και ανάκτηση μουσικών τεκμηρίων έχει επικεντρωθεί στην ανάπτυξη συστημάτων που επιτρέπουν την απάντηση ερωτημάτων που διατυπώνονται με τη μορφή σιγοτραγουδισμένης μελωδίας (Foote, 1999).

### **3.12.3. Συστήματα ανάκτησης βίντεο**

Τα συστήματα αναζήτησης και ανάκτησης πολυμέσων απασχολούν ένα πολύ σημαντικό τομέα της έρευνας που στόχο έχει την ανάκτηση συγκεκριμένης πληροφορίας που εμπεριέχεται σε ένα αρχείο βίντεο. Παραδείγματος χάριν, ένας χρήστης ενδιαφέρεται να δει μια κινηματογραφική σκηνή από μια ταινία, δίχως όμως να γνωρίζει σε ποιο σημείο του αρχείου βρίσκεται ακριβώς αυτή η σκηνή. Παραδοσιακά, πρέπει να δει όλη την ταινία και να χρησιμοποιεί το πλήκτρο “forward” μέχρις ότου βρει τη συγκεκριμένη σκηνή που τον ενδιαφέρει. Σήμερα, με τη βοήθεια των τεχνολογιών XML, και των νέων συστημάτων αναζήτησης και

ανάκτησης βίντεο θα έχει τη δυνατότητα εντοπισμού της συγκεκριμένης πληροφορίας που τον ενδιαφέρει.

Τα συστήματα αναζήτησης και ανάκτησης βίντεο παρέχουν δύο βασικές δυνατότητες αναζήτησης βασισμένες η πρώτη σε λέξεις-κλειδιά και η δεύτερη στο περιεχόμενο (Ruzomberka, 2002). Όπως συμβαίνει και στην ανάκτηση εικόνων και ήχου, έτσι και στην περίπτωση του βίντεο η περιγραφή γίνεται με λέξεις-κλειδιά και η αναζήτηση πραγματοποιείται επίσης με λέξεις-κλειδιά στο κείμενο που συνοδεύει ένα αρχείο βίντεο.

Ωστόσο, οι χρήστες έχουν ανάγκη υποβολής σύνθετων ερωτημάτων τα οποία είναι δυνατό να απαντηθούν μόνο εφόσον έχει πραγματοποιηθεί αυτόματη ευρετηρίαση των αρχείων βίντεο. Συγκεκριμένα, τα συστήματα αναζήτησης και ανάκτησης βίντεο βασισμένα στο περιεχόμενο ευρετηριάζουν αυτόματα το περιεχόμενο ενός αρχείου βίντεο τεμαχίζοντάς το σε κλιπ και εξάγοντας χαρακτηριστικά όπως κείμενο, χρώμα, υφή, κίνηση από κάθε κλιπ για την υποστήριξη της αναζήτησης. Τα παραπάνω συστήματα παρέχουν πρόσβαση στο περιεχόμενο μέσω αναζήτησης πλήρους κειμένου, οπτικό query-by-example, ενώ σε ορισμένες περιπτώσεις παρέχεται η δυνατότητα υποβολής ερωτημάτων με τη μορφή σχεδίου. Τα συγκεκριμένα εργαλεία αναζήτησης του περιεχομένου αρχείων βίντεο συνήθως είναι ενσωματωμένα με τις δυνατότητες περιήγησης (video browsing) και αναπαραγωγής με στόχο την προεπισκόπηση του βίντεο και την ευκολότερη ανάκτηση της επιθυμητής πληροφορίας (McDonald, 2005).

### **3.13. Υπηρεσίες αναζήτησης**

Μία υπηρεσία αναζήτησης (search service), ή διαθέτης (host) της έρευνας, είναι ένας οργανισμός, συνήθως κερδοσκοπικός, που παρέχει τη δυνατότητα σε πελάτες ή σε ειδικούς της πληροφόρησης, της πρόσβασης σε βάσεις δεδομένων. Μία υπηρεσία αναζήτησης/ ανάκτησης διαθέτει ή μερικές φορές παράγει βάσεις δεδομένων που πιστεύουν ότι χρειάζονται οι χρήστες τους. Οι διαθέτες τις περισσότερες φορές συμβάλλονται με τους παραγωγούς, οι οποίοι είναι συνήθως μη κερδοσκοπικοί φορείς και είναι αυτοί που συλλέγουν τις πληροφορίες και τις μετατρέπουν σε μορφή αναγνώσιμη από υπολογιστή. Οι περισσότεροι διαθέτες είναι ιδιωτικές επιχειρήσεις, που χρεώνουν για τις υπηρεσίες που παρέχουν. Συνήθως το μεγαλύτερο ποσό αφορά το χρόνο έρευνας. Ορισμένοι παραγωγοί αποτελούν συγχρόνως και υπηρεσίες αναζήτησης/ ανάκτησης, όπως ο OCLC. (Κορομπίλη Στέλλα, 1995)

## **4. Η Διαδικασία της Αναζήτησης**

Η χρήση δύο, τριών ή περισσότερων εννοιών σε μία έρευνα προϋποθέτει στρατηγική αναζήτησης. "Στρατηγική αναζήτησης" αποτελεί ένα σχέδιο ή μία προσέγγιση για τη λύση ενός προβλήματος αναζήτησης. Με άλλα λόγια αποτελεί ένα σχέδιο για την επιτυχία ενός στόχου. Η στρατηγική αναζήτησης μπορεί να συντελέσει στην αποτελεσματικότητα της ανάκτησης πληροφοριών και πρέπει να προετοιμαστεί πριν την σύνδεση με την online βάση δεδομένων. Αν σκεφτούμε ότι η αναζήτηση έχει σημαντικό κόστος, κατανοούμε γιατί ο ερευνητής πρέπει να είναι προετοιμασμένος καλά πριν αρχίσει την έρευνα. Όσο καλύτερη είναι η στρατηγική αναζήτησης τόσο

καλύτερα αποτελέσματα θα ανακτηθούν. "Τακτική αναζήτησης" (search tactic) είναι μία κίνηση που γίνεται για την επέκταση της έρευνας. Ενώ "διατύπωση αναζήτησης" (search formulation) είναι οι προτάσεις που εκφράζουν την ερώτηση σε μορφή που να είναι κατανοητή από την υπηρεσία.

Σε κάθε αναζήτηση είτε σε έντυπο ευρετήριο, είτε σε βάση δεδομένων μπορούν να ανακτηθούν μία σειρά από επαφές που είναι συναφείς με το θέμα της έρευνας και άλλες που δεν ενδιαφέρουν καθόλου το χρήστη. Ακόμη και αυτές που είναι συναφείς, άλλες έχουν άμεση σχέση με το θέμα και άλλες έμμεση. Στην ανάκτηση πληροφοριών ο ερευνητής σπάνια μπορεί να ανακτήσει όλες τις πληροφορίες που χρειάζεται και μόνον αυτές. Τις περισσότερες φορές θα ανακτηθούν πάνω από μια εγγραφή που δεν είναι σχετική με την ερώτηση. Την ίδια στιγμή ο ερευνητής δεν είναι σίγουρος αν μετά το τέλος της έρευνας δεν έχουν μείνει σημαντικές επαφές στη βάση δεδομένων που δεν έχουν ανακτηθεί. Όταν ο ερευνητής ανακτά μία σειρά από εγγραφές είναι πολύ δύσκολο να γνωρίζει αν έχει ανακτήσει όλες τις χρήσιμες εγγραφές ή αν έχει ανακτήσει τουλάχιστον τις πιο χρήσιμες από τις εγγραφές.

Η ποιότητα και η ποσότητα των αποτελεσμάτων, καθώς και η ταχύτητα με την οποία ο ερευνητής παίρνει μία ικανοποιητική απάντηση εξαρτάται από τις συλλογικές αποφάσεις που θα πάρει κατά τη διάρκεια της έρευνας. Ο ερευνητής πρέπει να προβλέψει με ποιο τρόπο έχουν ευρετηριαστεί οι εγγραφές και να ανακτήσει ένα αριθμό τεκμηρίων που να είναι αρκετά μικρός ώστε να μπορεί να ρίξει μια ματιά και να βρει τα πιο χρήσιμα τεκμήρια. Αν αυτός που κάνει την ευρετηρίαση χρησιμοποιεί πέντε όρους για να ευρετηριάσει ένα τεκμήριο, τότε και ο ερευνητής πρέπει να προβλέψει τους όρους αυτούς για να κάνει αποτελεσματική ανάκτηση. Η πρόβλεψη των όρων που χρησιμοποιούνται για ευρετηρίαση στα τεκμήρια που είναι σχετικά με την ερώτηση ονομάζεται κριτήριο πρόβλεψης (prediction criterion).

Έχουμε ήδη αναφέρει ότι πρέπει να γίνει προσπάθεια να ανακτηθούν τα πιο συναφή ή χρήσιμα τεκμήρια. Αλλά τι εννοούμε με τον όρο "συνάφεια"; Η online ανάκτηση έχει παράδοση στη μέτρηση και στην αξιολόγηση των ανακτώμενων αποτελεσμάτων. Ωστόσο μέχρι τώρα δεν υπάρχει συστηματικός τρόπος που να περιγράφει τη σχέση ανάμεσα στην ποιότητα και στη μέτρηση αυτής της ποιότητας με ποσοτικά κριτήρια. Έτσι οι ερευνητές και οι χρήστες της online ανάκτησης εφαρμόζουν τυπικά κριτήρια όπως μέγιστο και ελάχιστο. Με άλλα λόγια δίνεται ένα σύμβολο σε ένα ανακτημένο αποτέλεσμα το οποίο σε μία ιεραρχία βρίσκεται σε υψηλότερο σημείο από ένα σύμβολο που δίνεται σε ένα δεύτερο αποτέλεσμα και η έννοια της αποτελεσματικότητας εκφράζεται με ποσοτικά κριτήρια. Γενικότερα, για να μετρήσουμε ένα αποτέλεσμα, πρέπει να μετρήσουμε τον αριθμό των συναφών τεκμηρίων που έχουν ανακτηθεί.

Ο όρος "συνάφεια" αποτελεί βασική έννοια στην επιστήμη της πληροφόρησης και παίζει σημαντικό ρόλο στη διαδικασία της αξιολόγησης. Ωστόσο δεν υπάρχει σύμφωνη γνώμη για το τι σημαίνει ακριβώς. Συνήθως οι άνθρωποι καταλαβαίνουν τι σημαίνει, αλλά δεν μπορούν να δώσουν ορισμό. Άλλοι χρησιμοποιούν τον όρο "χρήσιμα τεκμήρια", γιατί ο χρήστης μπορεί να μην ξέρει τι είναι συναφές με τις πληροφοριακές του ανάγκες, αλλά ξέρει τι είναι χρήσιμο για αυτόν. Αυτό που θεωρείται δεδομένο είναι ότι η έννοια της "συνάφειας" των τεκμηρίων βασίζεται στις απόψεις των χρηστών, στο πως οι χρήστες αντιλαμβάνονται την πληροφόρηση και τις πληροφοριακές ανάγκες τους. Κυρίως μας ενδιαφέρουν τα κριτήρια των χρηστών για τη μέτρηση της συνάφειας των τεκμηρίων. Οι κρίσεις των χρηστών μπορεί να επηρεαστούν από εσωτερικούς παράγοντες, όπως στόχους και αξίες και από εξωτερικούς παράγοντες.

Συνοψίζοντας, πριν ο χρήστης φτάσει να ξεφυλλίσει, ηλεκτρονικά ή συμβατικά, έναν κατάλογο ή ένα ευρετήριο πρέπει να έχει καθορίσει με σαφήνεια τι είναι αυτό που αναζητά. Τα τέσσερα στάδια που ακολουθούν περιγράφουν αυτή τη διαδικασία:

- Ανακεφαλαίωση του θέματος της αναζήτησης σε μία ή δύο προτάσεις.
- Προσδιορισμός των ιδεών που διέπουν το θέμα, υπογράμμιση των ιδεών στις προτάσεις. Έτσι, το θέμα διαιρείται σε δύο ή περισσότερες ενότητες.
- Δημιουργία λίστας λέξεων ή φράσεων που περιγράφουν κάθε ιδέα. Κατά το στάδιο αυτό μπορούν να λαμβάνονται υπόψη οι θεματικές επικεφαλίδες της Βιβλιοθήκης του Κογκρέσου ή κάποιο άλλο ελεγχόμενο λεξιλόγιο, αλλά και πιθανές ιεραρχικές σχέσεις μεταξύ των ιδεών.
- Σύνδεση λέξεων και ιδεών με λογικούς τελεστές (συνήθως Boolean). Αυτή η διαδικασία μπορεί να πραγματοποιηθεί από επαγγελματίες ή να διδαχθεί άμεσα στους χρήστες, ώστε να πραγματοποιούν οι ίδιοι αποτελεσματικές αναζητήσεις.

Πολλές είναι οι Βιβλιοθήκες που παρέχουν στους χρήστες φόρμες δημιουργίας τέτοιων στρατηγικών, όπως αυτή που φαίνεται παρακάτω προερχόμενη από την Βιβλιοθήκη του Πανεπιστημίου της Αριζόνα.

## 5. Τεχνικές Ανάκτησης Πληροφοριών

Οι περισσότερες από τις παρακάτω τεχνικές χρησιμοποιούνται από όλες σχεδόν τις βιβλιογραφικές τράπεζες πληροφοριών.

- **Λογικοί τελεστές:** Οι χρήστες συνήθως χρησιμοποιούν ερωτήσεις που αποτελούνται από σύνθετες έννοιες. Η πορεία αναζήτησης είναι συνδεδεμένη με τους τελεστές Boole or, and, not. Δύο σημαντικά προβλήματα υπάρχουν στην αναζήτηση με τελεστές Boole:

- ✍ Ορισμένες φορές είναι απαραίτητο να υπάρχει μεγαλύτερη ακρίβεια στη σύνταξη των όρων. Οι όροι μπορεί να βρίσκονται στο ίδιο κείμενο, χωρίς όμως να σχετίζονται νοηματικά ή να εμφανίζονται σε διαφορετική από την επιθυμητή σειρά.
- ✍ Είναι δύσκολο και κουραστικό να περιλαμβάνονται σε μία ερώτηση όλες οι λέξεις που παράγονται από την ίδια ρίζα

Για να λυθούν τα παραπάνω προβλήματα μπορούμε να χρησιμοποιήσουμε τελεστές εγγύτητας (proximity operators) και αποκοπή (truncation):

- *Τελεστές εγγύτητας*: Είναι οι λειτουργίες που δηλώνουν εγγύτητα των όρων μεταξύ τους
- *Αποκοπή*: Η αποκοπή δίνει τη δυνατότητα να περιλάβουμε στην έρευνα όλες τις λέξεις ή όσες χρειάζεται η έρευνα, που παράγονται από την ίδια ρίζα

Η αποκοπή, όπως και οι τελεστές εγγύτητας μπορούν να χρησιμοποιηθούν και σε συνδυασμό με τους τελεστές Boole. (Κορομπίλη Στέλλα, 1995)

## 6. Ρόλος Βιβλιοθηκονόμου

Ένα από τα σημαντικότερα καθήκοντα ενός βιβλιοθηκονόμου είναι η αποτελεσματική παροχή πληροφοριών σε κάθε χρήστη. Επίσης θα πρέπει να έχει τη δυνατότητα να προβλέπει μελλοντικές ανάγκες των χρηστών, που θα οδηγήσουν στη δημιουργία νέων υπηρεσιών. Είναι το πρόσωπο το οποίο μεταφράζει την ερώτηση σε όρους αποδεκτούς από τις πληροφοριακές πηγές, εντοπίζει τις ακριβείς πηγές και μαζί με τον χρήστη αποφασίζουν αν η απάντηση είναι επαρκής ή όχι.

Ο πληροφοριακός βιβλιοθηκονόμος θα πρέπει να αναπτύξει τις εξής δεξιότητες:

- ικανότητα οργάνωσης πληροφοριών για μελλοντική χρήση από τους χρήστες
- γνώση χρήσης βιβλιογραφικών εργαλείων
- δυνατότητα επικοινωνίας με τους χρήστες
- ικανότητα διεξαγωγής πληροφοριακού διαλόγου

Ο συνυπολογισμός των αναγκών των χρηστών και των ανθρώπινων κοινωνιογνωστικών χαρακτηριστικών γνωρισμάτων στη διαδικασία πληροφόρησης μπορεί να συμβάλει στην καλύτερη λειτουργία των συστημάτων, των υπηρεσιών και των προϊόντων πληροφόρησης.

Υπάρχουν αρκετές διαδικτυακές τεχνολογίες που επιτρέπουν στους βιβλιοθηκονόμους να αναπτύξουν νέες υπηρεσίες επικοινωνίας και πληροφόρησης για τους χρήστες. Αυτές συχνά χρησιμοποιούνται σε συνδυασμό για να αποτελέσουν ένα είδος υπηρεσίας και συνήθως αποτελούν διαφορετικά μέσα για την εκτέλεση της αίτησης του χρήστη.

Μπορούμε να διακρίνουμε τρία είδη υπηρεσιών ανάλογα με το ευρύτερο / συχνότερα χρησιμοποιούμενο τεχνολογικό – επικοινωνιακό μέσο / τεχνική που εφαρμόζεται. Η πιο απλή και συχνότερα χρησιμοποιούμενη μορφή της ψηφιακής υπηρεσίας πληροφόρησης είναι η δημιουργία μιας ιστοσελίδας όπου περιέχονται οι «συχνές ερωτήσεις» (FAQs). Η επόμενη κατηγορία υπηρεσιών επικοινωνίας είναι το ηλεκτρονικό ταχυδρομείο (e-mail) που περιλαμβάνει και τις έτοιμες φόρμες ερωτήσεων του Ιστού (Web forms). Η τρίτη κατηγορία ψηφιακών υπηρεσιών πληροφόρησης είναι αυτές όπου ο χρήστης συνομιλεί με το βιβλιοθηκονόμο μέσω

ενός λογισμικού άμεσης επικοινωνίας που μπορεί να είναι αυτό της ταυτόχρονης συνομιλίας (chat) ή άμεσης ανταλλαγής μηνυμάτων (instant messaging).

Ο προσανατολισμός για την ανάπτυξη στο Διαδίκτυο αποτελεσματικών υπηρεσιών πληροφόρησης προϋποθέτει:

- Κριτική εξέταση των θεμάτων διαχείρισης και διοίκησης των πληροφοριών και των πηγών τους, με τη στελέχωση των βιβλιοθηκών με ειδικευμένο προσωπικό στην αξιολόγηση και εγκατάσταση ηλεκτρονικών πηγών πληροφόρησης.
- Καλή γνώση των αναγκών και των προσδοκιών των χρηστών με τον κατάλληλο σχεδιασμό για εκπαιδευτικά προγράμματα, δημοσιότητα και συνεχή αξιολόγηση της αποτελεσματικότητας των υπηρεσιών που τους προσφέρονται.
- Προσαρμογή και/ή ανάπτυξη νέων πολιτικών και διαδικασιών των ίδιων των συλλογών των βιβλιοθηκών. Η ανάπτυξη ψηφιακών βιβλιοθηκών αποτελεί ιδανικό κίνητρο για την αποθήκευση, οργάνωση και χρήση της γνώσης στην παγκόσμια κοινωνία των πληροφοριών, γιατί η ψηφιακή βιβλιοθήκη δημιουργεί ένα "περιβάλλον στο οποίο συνενώνονται συλλογές, υπηρεσίες και άνθρωποι για να υποστηρίξουν τον πλήρη κύκλο της δημιουργίας, διάχυσης, χρήσης και διατήρησης δεδομένων, πληροφοριών και γνώσης".

## **7. Συμπεράσματα**

Σε έναν κόσμο συνεχώς μεταβαλλόμενο, η ανάγκη για γρήγορη και αποτελεσματική ανάκτηση πληροφοριών έχει γίνει επιτακτική. Οι τρόποι και τα εργαλεία που εξυπηρετούν αυτή την ανάγκη έχουν αλλάξει με το πέρασμα του χρόνου και φυσικά αναμένονται και άλλες αλλαγές. Η βιβλιοθήκη πρέπει να προσαρμόζεται στις ανάγκες της κοινωνίας και να έχει το βασικό ρόλο διάδοσης και φύλαξης της πληροφορίας ως κοινωνικό αγαθό. Οι επιστήμονες της πληροφόρησης, εμείς, πρέπει να βοηθήσουμε στη δημιουργία μιας τέτοιας κοινωνίας με υψηλό μορφωτικό επίπεδο.

Το ψηφιακό περιβάλλον έχει φέρει πολλές αλλαγές στο περιβάλλον αναζήτησης και ανάκτησης της πληροφορίας. Στις αλλαγές αυτές οφείλονται όλοι οι μηχανισμοί που αναλύθηκαν και παρουσιάστηκαν παραπάνω. Οι μέχρι τώρα εξελίξεις έχουν:

- βελτιώσει τις υπηρεσίες προς το χρήστη
- προσφέρει ευκολία αναζήτησης
- προσφέρει ευρύτητα πληροφοριακών πηγών
- εξαλείφει τους χωροταξικούς περιορισμούς (χώρος βιβλιοθήκης)

Βέβαια οι εξελίξεις δεν σταματούν εδώ αλλά διαρκώς ακολουθείται μια πρόοδος όσον αφορά τον τομέα της τεχνολογίας που εισάγεται στις βιβλιοθήκες.

Το σύγχρονο περιβάλλον χαρακτηρίζεται από τεράστιο πλήθος πληροφοριών και μεγάλη ποικιλία μορφοτύπων. Η πληροφορία αποτυπωμένη σε κείμενο, αν και επικρατεί στο διαδίκτυο, δεν αποτελεί τη μοναδική μορφή καταγραφής των πληροφοριών. Ο χρήστης σήμερα έχει τη δυνατότητα να αναζητήσει και ανακτήσει πληροφορίες που έχουν δημοσιευθεί με τη μορφή εικόνων, ήχου και βίντεο. Οι τεχνολογικές εξελίξεις επιτρέπουν την υποβολή τόσο απλών όσο και σύνθετων ερωτημάτων, με στόχο την ανάκτηση της επιθυμητής πληροφορίας. Η έρευνα, στα πλαίσια της τεχνητής νοημοσύνης, έχει αναπτύξει ιδιαίτερα συστήματα επεξεργασίας φυσικής γλώσσας, γεγονός που εξυπηρετεί τους χρήστες του διαδικτύου και τους απαλλάσσει από τον κόπο εκμάθησης ποικίλων τεχνικών και μεθόδων αναζήτησης και ανάκτησης πληροφοριών. Εξίσου ευοίωνα είναι τα πρώτα επιτεύγματα της

τεχνολογίας σχετικά με τη φωνητική αναζήτηση εξυπηρετώντας όχι μόνο τα άτομα με ειδικές ανάγκες αλλά και τους χρήστες φορητών και συχνά μικρών σε μέγεθος συσκευών για την πρόσβαση στο διαδίκτυο.

Αν και το πλήθος των ηλεκτρονικών πηγών συχνά είναι σε θέση να καλύψει τις πληροφοριακές ανάγκες ακόμη και των πιο απαιτητικών χρηστών, δεν πρέπει να αγνοούμε τον θησαυρό που κρύβουν οι βιβλιοθήκες και το έντυπο υλικό που διαθέτουν. Ωστόσο, βελτιωμένες τεχνικές αναζήτησης και ανάκτησης αυτού του υλικού πρέπει να εφαρμοσθούν ώστε ο χρήστης να αποκτή για κάθε πληροφοριακή του ανάγκη μια γενικότερη επισκόπηση των διαθέσιμων πηγών πληροφόρησης, ηλεκτρονικών και έντυπων.

Στον σύγχρονο κόσμο της πληροφορίας η φυλλομέτρηση και η αναζήτηση πληροφοριών σε αποθετήρια ανά τον κόσμο έχει γίνει ρουτίνα και πραγματοποιείται χωρίς την ανάγκη μετακίνησης. Έτσι, η παραδοσιακή έννοια της βιβλιοθήκης, με περιορισμούς στον τόπο και τον χρόνο πρόσβασης, αμφισβητείται και η αξία της ως χώρου αποθήκευσης και ως υπηρεσίας πληροφόρησης συχνά παρανοείται (Ercegovac, 2003). Σύμφωνα με αρκετά σενάρια, οι Ακαδημαϊκές Βιβλιοθήκες θα εγκαταλειφθούν σε φυσικό επίπεδο (κτιριακές εγκαταστάσεις, έντυπο υλικό), αφού οι χρήστες τους θα έχουν πρόσβαση στα πάντα μέσω του υπολογιστή τους (Lesk, 1995). Επιπλέον, τα επόμενα χρόνια, το πολυμεσικό υλικό αναμένεται να πολλαπλασιαστεί και θα αποτελέσει ενδεχομένως έναν ισχυρό ανταγωνιστή του κειμένου. Τα προβλήματα που θα διογκωθούν στον χώρο της αναζήτησης και ανάκτησης πληροφορίας κατά την προσεχή περίοδο, θα αφορούν κυρίως τα πνευματικά δικαιώματα, το φιλτράρισμα της πληροφορίας που παρέχει το Διαδίκτυο και τον έλεγχο του τελευταίου από ομάδες ατόμων ή κρατών

Στη σημερινή, μεταβαλλόμενη εποχή, η οποία χαρακτηρίζεται από τη συνεχόμενη αύξηση των παραγόμενων πληροφοριών και ως συνέπεια την αύξηση της ανάγκης για διαμοιρασμό και πρόσβαση σε όλες αυτές τις πληροφορίες, κρίνεται απαραίτητη η οργάνωση του περιεχομένου της πληροφορίας.

Οι συνεχείς εξελίξεις που παρατηρούνται στον τομέα της αναζήτησης και ανάκτησης της πληροφορίας, θα βοηθήσουν σημαντικά στον διαμοιρασμό αυτής. Απαιτείται όμως πρωταρχικά να γίνει προσπάθεια από όλους τους εμπλεκόμενους για εξοικείωση με τα νέα εργαλεία και για συνεχή ενημέρωση των τρεχόντων εξελίξεων.

Τέλος, πρέπει να σημειωθεί ότι πολλές από τις παραπάνω τεχνικές και τεχνολογίες που παρουσιάστηκαν βρίσκονται σε πιλοτικό ή πειραματικό στάδιο, γεγονός που εξηγεί ίσως και την αδράνεια των βιβλιοθηκών να τις εφαρμόσουν στα δικά τους συστήματα αναζήτησης και ανάκτησης πληροφοριών.

## 8. Βιβλιογραφία

Κωνσταντόπουλος, Π., Μαγκαναράκη, Α. Και Ντερ, Μ. (2005). 'Οδηγός Πολιτισμικής Τεκμηρίωσης Και Δια-Λειτουργικότητας'. Κέντρο Πολιτισμικής Πληροφορικής, Ινστιτούτο Πληροφορικής ΙΤΕ.

Ding, Y. et al. (2002). 'The semantic web: yet another hip?'. *Data & Knowledge Engineering*, 41, 205-227.

McCallum, S. (2006). 'A look at New Information Retrieval Protocols: SPU, OpenSearch/A9, CQL, and XQuery'. Seoul, Korea: World Library and Information Congress, 72<sup>nd</sup> General Conference and Council.

Merholz, P. (2004). 'Ethnolocation and vernacular vocabularies'. August 30, 2004. [Ιστοσελίδα] Διαθέσιμο στο: <<http://www.peterme.com/archives/000387.html>>. [Ημερομηνία τελευταίας πρόσβασης 12 Μαΐου 2008].

Neches, R. et al. (1991). 'Enabling technology for knowledge sharing'. *AI Magazine*, 12 (3), 36-56.

Rainie, L. (2007). '28% of online Americans have used the internet to tag content', *Pew Internet and American Life Project*. Διαθέσιμο στο: <[http://www.pewinternet.org/pdfs/PIP\\_Tagging.pdf](http://www.pewinternet.org/pdfs/PIP_Tagging.pdf)>. [Ημερομηνία τελευταίας πρόσβασης 12 Μαΐου 2008].

Studer, R., Benjamins, V. R. and Fensel, D. (1998). 'Knowledge engineering: principles and methods'. *Data & Knowledge Engineering*, 25, 161-197.

Tringology. (2007). 'Danbury, CT Kicks off Library Thing for Libraries!'. [Ιστοσελίδα]. Διαθέσιμο στο: <<http://www.librarything.com/thingology/2007/05/danbury-ct-kicks-off-librarything-for.php>>. [Ημερομηνία τελευταίας πρόσβασης 12 Μαΐου 2008].

Wexler, M. (2005). 'I continue to despise tagging...'. August 1, 2005. *The Net Takeaway*. Διαθέσιμο στο: <<http://www.nettakeaway.com/tp/article/175/i-continue-to-despise-tagging>>. [Ημερομηνία τελευταίας πρόσβασης 12 Μαΐου 2008].

Marchionini, Gary (2006). Toward Human-Computer Information Retrieval. *ASIST Bulletin*. Τεκμήριο διαθέσιμο στο Διαδίκτυο: <http://www.asis.org/Bulletin/Jun-06/marchionini.html>. [Ημερομηνία τελευταίας πρόσβασης: 12. Μαΐου. 2008]

Hafner, Katie (2004). Old Search Engine, the Library, Tries to Fit Into a Google World. *The New York Times*, June 21 2004.



Twentyman, Jessica (2006). Security special report: Who sees your data? Τεκμήριο διαθέσιμο στο Διαδίκτυο στο: <http://www.computerweekly.com/Articles/2006/04/25/215622/security-special-report-who-sees-your-data.htm>. [Ημερομηνία τελευταίας πρόσβασης 12.Μαΐου. 2007]

Long, Fuhui, Hongjiang Zhang and David Dagan Feng (2003). Fundamentals of content-based image retrieval. Τεκμήριο διαθέσιμο στο Διαδίκτυο στο: [http://research.microsoft.com/asia/dload\\_files/group/mcomputing/2003P/ch01\\_Long\\_v40-proof.pdf](http://research.microsoft.com/asia/dload_files/group/mcomputing/2003P/ch01_Long_v40-proof.pdf) [Ημερομηνία τελευταίας πρόσβασης: 12 Μαΐου 2008]

Rochkind, Jonathan (2007) (Meta)search like Google. Τεκμήριο διαθέσιμο στο Διαδίκτυο στο: <http://www.libraryjournal.com/article/CA6413442.html> [Ημερομηνία τελευταίας πρόσβασης: 12 Μαΐου 2008]

Vauhini, Vara *New Search Engines Help Users Find Blogs*. The Wall Street Journal. 7 Sept. 2005. Τεκμήριο διαθέσιμο στο Διαδίκτυο στο: [http://online.wsj.com/public/article/SB112605609071333521-YJ\\_zDNeKEGDupVAyFZ96e1xw8tY\\_20060907.html](http://online.wsj.com/public/article/SB112605609071333521-YJ_zDNeKEGDupVAyFZ96e1xw8tY_20060907.html). [Ημερομηνία τελευταίας πρόσβασης: 12 Μαΐου 2008.]

Kasowitz-Scheer, Abby-Pasqualoni, Michael (2002) *Information Literacy Instruction in Higher Education: Trends and Issues*. ERIC Clearinghouse on Information and Technology Syracuse NY. Τεκμήριο διαθέσιμο στο Διαδίκτυο στο: <http://www.ericdigests.org/2003-1/information.htm>. [Ημερομηνία τελευταίας πρόσβασης: 12 Μαΐου 2008]

Scherrer, Carol S. (2004) *Reference librarians' perceptions of the issues they face as academic health information professionals*. Journal of the Medical Library Association, 92(2).

Bray, Tim, Jean Paoli [et al.] (2006) Extensible Markup Language (XML) 1.0 – Origin and Goals. World Wide Web Consortium. Τεκμήριο διαθέσιμο στο Διαδίκτυο στο: <http://www.w3.org/TR/2006/REC-xml-20060816/#sec-origin-goals> [Ημερομηνία τελευταίας πρόσβασης: 12 Μαΐου 2008]

Lafon, Yves, Bert Bos (2002) Describing and retrieving photos using RDF and HTTP: W3C Note 19 April 2002. Τεκμήριο διαθέσιμο στο Διαδίκτυο στο: <http://www.w3.org/TR/photo-rdf/> [Ημερομηνία τελευταίας πρόσβασης: 12 Μαΐου 2008]

Brown, Cecilia M. (1999) Information seeking behavior of Scientists in the Electronic Information Age: Astronomers, Chemists, Mathematicians, and Physicists. Journal of the American Society for Information Science. 50(10):929-943.

Cool, C., Park, S., Belkin, N.J., Koeneman, J., & Ng, KB. (1996). Information seeking behavior in new searching environments. *Proceedings of the Second International Conference on Conceptions of Library and Informaion Science*. Copenhagen, Denmark.

Maurer, Donna (2006). Four Modes of Seeking Information and How to Design for Them. Τεκμήριο διαθέσιμο στο Διαδίκτυο στο: [http://www.boxesandarrows.com/view/four\\_modes\\_of\\_seeking\\_information\\_and\\_how\\_to\\_design\\_for\\_them](http://www.boxesandarrows.com/view/four_modes_of_seeking_information_and_how_to_design_for_them) [Ημερομηνία τελευταίας πρόσβασης: 12 Μαΐου 2008]

Ercegovac, Zorana and Erika Yamasaki (2003). Information Literacy: Search Strategies, Tools and Resources. Τεκμήριο διαθέσιμο στο Διαδίκτυο στο: <http://www.libraryinstruction.com/infosearch.html> [Ημερομηνία τελευταίας πρόσβασης: 12 Μαΐου 2008]

Lesk, Michael. (1995) The seven ages of information retrieval. International Federation of Library Associations and Institutions, Occasional Paper. Τεκμήριο διαθέσιμο στο Διαδίκτυο στο: <http://www.ifla.org/VI/5/op/udtop5/udt-op5.pdf> [Ημερομηνία τελευταίας πρόσβασης: 12 Μαΐου 2008]

“Query expansion: From Wikipedia, the free encyclopedia”. (2007). [Ιστοσελίδα]. <[http://en.wikipedia.org/wiki/Query\\_expansion](http://en.wikipedia.org/wiki/Query_expansion)>. [Ημερομηνία τελευταίας πρόσβασης: 12 Μαΐου 2008.]

“Query Expansion” in Free On-line Dictionary of Computing. (2000). [Ιστοσελίδα]. <<http://www.cacs.louisiana.edu/~mgr/404/burks/foldoc/>>. [Ημερομηνία τελευταίας πρόσβασης: 12 Μαΐου 2008.]

Baeza-Yates, R και Riberio-Neto, B. (1999). *Modern Information Retrieval*. New York: ACM press, Harlow: Addison-Wesley.

Foote, J. (1999). “An overview of audio information retrieval” *Multimedia Systems*, 7, pp. 2–10. <[http://web.archive.org/web/20060517012614/http://www.ee.columbia.edu/~mjr59/reviews/foote\\_overview.pdf](http://web.archive.org/web/20060517012614/http://www.ee.columbia.edu/~mjr59/reviews/foote_overview.pdf)>. [Ημερομηνία τελευταίας πρόσβασης: 12 Μαΐου 2008.]

Goodrum, A. A. (2000). ‘Image information retrieval: an overview of current research’ *Informing Science*, vol.3, no. 2, pp. 63-67.

Google. (2007). “Google Voice Local Search” [Ιστοσελίδα]. <<http://labs.google.com/goog411/>>. [Ημερομηνία τελευταίας πρόσβασης: 12 Μαΐου 2008.]

Lager, M. (1996). “Spinning a Web Search” [Ιστοσελίδα]. <<http://www.library.ucsb.edu/untangle/lager.html>>. [Ημερομηνία τελευταίας πρόσβασης: 12 Μαΐου 2008.]

Pavel, M. (2001). “Multimedia Information Retrieval” [Ιστοσελίδα]. <<http://www.cse.ogi.edu/class/cse580ir/handouts/14%20November/Multimedia%20Information%20Retrieval/index.htm>>. [Ημερομηνία τελευταίας πρόσβασης: 12 Μαΐου 2008.]

Ramkrishnan, R and Gehrke, J (2003). *Database Management Systems*. New York: McGraw-Hill.

START. (χ.χ.) “The START Natural Language Question Answering System” [Ιστοσελίδα]. <<http://start.csail.mit.edu/>>. [Ημερομηνία τελευταίας πρόσβασης: 12 Μαΐου 2008.]

Youssef, M. A. (2001). “Cross Language Information Retrieval” [Ιστοσελίδα]. <<http://www.otal.umd.edu/uupractice/clir/>>. [Ημερομηνία τελευταίας πρόσβασης 12 Μαΐου 2008.]

Παπαθεοδώρου, Χ. (2007). “Ανάκτηση Πληροφοριών” [Παρουσίαση Powerpoint]. <<http://www.ionio.gr/~papatheodor/lessons/IONIO-msc-RETRIEVAL.PPT#1>>. [Ημερομηνία τελευταίας πρόσβασης 12 Μαΐου 2008.]

McGettrick, S. (2002). ‘Query Expansion’ [Παρουσίαση Powerpoint] <[http://www.ist.psu.edu/faculty\\_pages/giles/IST497/presentations/McGettrick.ppt#1](http://www.ist.psu.edu/faculty_pages/giles/IST497/presentations/McGettrick.ppt#1)> [Ημερομηνία τελευταίας πρόσβασης 12 Μαΐου 2008.]

McCarthy, J. (2004). “What is artificial intelligence?” [Ιστοσελίδα]. <<http://www-formal.stanford.edu/jmc/whatisai/>>. [Ημερομηνία τελευταίας πρόσβασης 12 Μαΐου 2008.]

McDonald, K. (2005). “Discrete Language Models for Video Retrieval” [Ιστοσελίδα]. <<http://www.computing.dcu.ie/~kmcdon/kmd-thesis/thesis.pdf>>. [Ημερομηνία τελευταίας πρόσβασης 12 Μαΐου 2008.]

Ruzomberka, M. (2002). “Video Information Retrieval” [Παρουσίαση Powerpoint]. <[http://www.ist.psu.edu/faculty\\_pages/giles/IST497/presentations/Ruzomberka.ppt](http://www.ist.psu.edu/faculty_pages/giles/IST497/presentations/Ruzomberka.ppt)>. Ημερομηνία τελευταίας πρόσβασης 1 Ιουνίου 2007.

Σταθούλα Θ. (2003), «Από τον αυτοματισμό των βιβλιοθηκών στη διαχείριση του ψηφιακού περιεχομένου», Αθήνα: [χ.ο.]

Κορομπίλη- Ξαντινίδου Στέλλα. (2000), « Online ανάκτηση πληροφοριών», Αθήνα: [χ.ο.]

Τριανταφύλλου Ι. (2004), «Βιβλιοθηκονομικές εφαρμογές Η/Υ», Αθήνα: [χ.ο.]

Μονιάρου-Παπακωνσταντίνου, Β. (2000), «Ανάπτυξη και διαχείριση συλλογής», Αθήνα: [χ.ο.]

‘Δίκτυο Ελληνικών Ακαδημαϊκών Βιβλιοθηκών’ [Ιστοσελίδα], [χ.χ.]. <<http://www.heallink.gr/journals/about.jsp>>. Ημερομηνία τελευταίας πρόσβασης 15 Μαΐου 2007.

Τσιμπόγλου Φ. (1998), «Κοινοπραξίες Ελληνικών Βιβλιοθηκών. Το “αναγκαίο καλό” των ψηφιακών βιβλιοθηκών», 7ο Πανελλήνιο Συνέδριο Ακαδημαϊκών Βιβλιοθηκών. Βόλος, 4-6 Νοεμβρίου 1998.

Γ. Βουγιουκλής κ.α, (2007), «Ελληνικός Συλλογικός Κατάλογος – Εξελίξεις και Προοπτικές», Διημερίδα Ακαδημαϊκών Βιβλιοθηκών 7 – 8 Μαΐου, Αθήνα

Μπόκος, Γεώργιος. *Εισαγωγή στην επιστήμη της πληροφόρησης*. Αθήνα: Παπασωτηρίου, 2001.

Σκανδάλη, Αλκμήνη Α. *Θεματική ευρετηρίαση : προσυνδυασμένα συστήματα*. Αθήνα: Ι. Γ. Βασιλείου, c1990.

Σταϊκούρα, Βασιλική. *Περιγραφική καταταλογογράφηση Ι*. Αθήνα: ΤΕΙ Αθήνας, 1996.

Χατζημαρή, Στέλλα. *Υστεροσυνδιασμένη θεματική ευρετηρίαση*. Αθήνα: ΤΕΙ Αθήνας, 2004.