

6ο ΠΑΝΕΛΛΗΝΙΟ ΣΥΝΕΔΡΙΟ
ΒΙΒΛΙΟΘΗΚΑΡΙΩΝ

**νέες
τεχνολογίες**

**και
ελληνικές
βιβλιοθήκες**

ΕΝΩΣΗ ΕΛΛΗΝΩΝ ΒΙΒΛΙΟΘΗΚΑΡΙΩΝ

ΑΘΗΝΑ 22-23 ΝΟΕΜΒΡΙΟΥ 1990

ΙΝΣΤΙΤΟΥΤΟ GOETHE
Ομήρου 14-16

Νικήτας Καρανικόλας – Ιωάννα Πολυδώρου
Οικονομικό Πανεπιστήμιο Αθηνών (ΑΣΟΕΕ)
Πατησίων 76, 10434 Αθήνα, τηλ. 8225268, 8819354.

Το μηχανογραφημένο σύστημα της Βιβλιοθήκης
του Οικονομικού Πανεπιστημίου Αθηνών (ΑΣΟΕΕ).

ΠΕΡΙΛΗΨΗ

Τα συστήματα αποθήκευσης και ανάκτησης δεδομένων μπορούν να σχεδιαστούν χρησιμοποιώντας προχωρημένες τεχνικές που προκύπτουν από το χώρο της έρευνας. Ένα τέτοιο σύστημα Ανάκτησης δεδομένων είναι αυτό που παρουσιάζουμε. Το σύστημα χρησιμοποιεί μια διαδικασία αυτόματης δεικτοδότησης των βιβλίων και αποθηκεύει τους όρους ανακτησής τους σε ανεστραμμένες λίστες. Οι ερωτήσεις για ανάκτηση υποβάλλονται στο σύστημα με τη μορφή λογικών (boolean) φορμών.

1. ΕΙΣΑΓΩΓΗ

Στην εισήγηση αυτή περιγράφεται το ιστορικό της δημιουργίας ενός μηχανοργανωμένου συστήματος, πώς εγκαταστάθηκε και λειτουργεί στη Βιβλιοθήκη του Οικονομικού Πανεπιστημίου Αθηνών, προκειμένου να διευκολυνθούν οι βιβλιοθηκονομικές εργασίες και υπηρεσίες. Οι σκοπιμότητες που επέβαλαν τη δημιουργία του, καθώς και οι τεχνικές προδιαγραφές με τις δυνατότητες ανάπτυξης και εγκατάστασης και σε άλλες Βιβλιοθήκες.

Σκοπός αυτής της εισήγησης είναι να ενημερώσουμε τους συναδέλφους που εργάζονται στις Βιβλιοθήκες για το πώς καταλήξαμε να δημιουργήσουμε κι' εμείς ένα δικό μας σύστημα κοντά στα τόσα άλλα, που ήδη υπάρχουν και στα όσα διόλου απίθανο να ακολουθήσουν. Τις αιτίες, τα προβλήματα και τις δυσκολίες που αντιμετωπίσαμε καθώς και τα αποτελέσματα των προσπαθειών μας. Δεν αποσκοπούμε να ωραιοποιήσουμε τα γεγονότα και τις καταστάσεις. Επιχειρούμε να καταδείξουμε τα στάδια μιας μακράς και επίπονης πορείας προκειμένου να ξεφύγουμε από την αιραξία που μας καταδικάζει η χωρίς μέτρο θεωρητική προσκόλληση για την αντιμετώπιση πρακτικών προβλημάτων και η ανασμύνη για την έξωθεν και άνωθεν έλευση του τέλειου στην πληροφόρηση και επικοινωνία.

Οι πρώτες πρωτόβουλες προσπάθειες, ήδη από το 1983, κατέληξαν σε κάποια εφαρμογή για τον κατάλογο των περιοδικών και μόνο με τη φιλότιμη προσφορά του τότε προγραμματιστή στο νεόκτιστο υπολογιστικό κέντρο του Πανεπιστημίου. Τον καιρό εκείνο ούτε συζήτηση για μηχανογράφηση της Βιβλιοθήκης μιας και άλλα τμήματα είχαν το προβάδισμα ενώ εμείς κρατούσαμε ακόμα τη θέση της τελευταίας προτεραιότητας. Ακολούθησε η γνωστή σε όλους μας συνεργασία στο πρόγραμμα του *Εθνικού Κέντρου Τεκμηρίωσης* (ΕΚΤ) για το συλλογικό κατάλογο των περιοδικών και η ανήσυχη αναμονή και αναζήτηση για το πώς και πότε θα κατέληγαν οι Βιβλιοθήκες Έρευνας να συνεννοηθούν και να εξεύρουν κοινές λύσεις και εφαρμογές χάριν της οικονομικότερης και ευρύτερης αποτελεσματικότητας και επικοινωνίας. Οι χρήστες που διαρκώς αυξάνονται ενώ το προσωπικό αποδεκατίζεται σχεδόν, το *Cardex* των περιοδικών που δεν ελέγχεται πια με τα ισχύοντα περί γραφικής ύλης και προμηθειών, το πλήθος των δανεισμών που καταδυναστεύει τα συρτάρια των υπαλλήλων δεν σφηνούν πια άλλα περιθώρια μοιρολατρικής αναμονής. Η άποψη "ο καθένας το δικό του" που πάντα αποφεύγαμε σαν το διάβολο, φάνηκε στη διαδρομή πως θα ήταν προτιμότερη από το τίποτα και ώσπου κάποια μεγάλη εταιρεία να αναλάβει να μας δικτυώσει, ποιος ξέρει με ποιά πενταετή αναπτυξιακό πρόγραμμα και πότε. Γι' αυτά όταν η Διοίκηση το 1989 έκρινε πως θά πρεπε να αρχίσει να αντιμετωπίζει κατ'είδαν τη μηχανοργάνωση της Βιβλιοθήκης του Πανεπιστημίου, βρεθήκαμε "σαν έτοιμοι από καιρό". Τη μέριμνα για τη δημιουργία του συστήματος

που παρουσιάζεται ανέλαβε το Κεντρο Οικονομικής Έρευνας του Πανεπιστημίου ύστερα απο σχετική πρόταση του κυρίου Ιωάννη Κάβουρα Καθηγητή Πληροφορικής ο οποίος ορίστηκε Επιστημονικός Υπεύθυνος για την παρακολούθηση του έργου. Την εργασία εκπόνησε ομάδα απο τους:

Καρανικόλα Νικήτα (Ανάλυση, Σχεδίαση, Υλοποίηση, Έλεγχος)

Πλέσσα Παναγιώτα (Συμμετοχή στα Παραπάνω και στην Τεκμηρίωση)

Ζωγραφίδη Γιώργο (Τεκμηρίωση)

Πολυδώρου Ιωάννα (Φορέας-Εκπρόσωπος Βιβλιοθήκης)

2. ΣΥΣΤΗΜΑΤΑ ΑΝΑΚΤΗΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ

Αυτό που συχνά αναφέρουμε σαν *Information Retrieval*[1] (Ανάκτηση Πληροφορίας) χωρίζεται σε δυό μεγάλες κατηγορίες. *Text Retrieval*[7] (Ανάκτηση Κειμένων) ή *Document Retrieval*[6] (Ανάκτηση Τεκμηρίωσης) είναι η κατηγορία εκείνη στην οποία γίνεται ανάκτηση εκείνων των κειμένων τα οποία ομοιάζουν σε περιεχόμενο με ένα κείμενο-πρότυπο το οποίο εισάγει ο χρήστης των συστημάτων αυτών. Στην κατηγορία αυτή δεν απαιτείται το πλήρη ταίριασμα κάποιων όρων του κειμένου-πρότυπου με τα ανακτώμενα κείμενα ούτε η χρήση λογικών φορμών μεταξύ όρων ανάκτησης. Ένα τέτοιο σύστημα πλεονεκτεί έναντι των συστημάτων *Data Retrieval* λόγω της εισαγωγής του κειμένου-πρότυπου σε φυσική γλώσσα και της αποφυγής λογικών φορμών. Τα συστήματα της κατηγορίας *Data Retrieval* αν και περιοριστικά για τον τρόπο εισόδου των ερωτήσεων για ανάκτηση (*Queries*) είναι πολύ ταχύτερα στην αναζήτηση δεδομένων.

Η απόφαση για το αν θα χρησιμοποιηθούν μέθοδοι *Text Retrieval* ή *Data Retrieval* δεν επηρεάζονται μόνο απο τα προαναφερθέντα πλεονεκτήματα και μειονεκτήματα αλλά και απο τη μορφή των δεδομένων για εισαγωγή και ανάκτηση. Έτσι σε ένα σύστημα που αφορά την καταλογογράφηση των βιβλίων μιάς Επιστημονικής Βιβλιοθήκης είναι προτιμότερη η χρήση μεθόδων *Data Retrieval* ενώ αντίθετα σ'ένα σύστημα ανάκτησης *abstracts* απο *papers* ή και ολόκληρων *papers* είναι προτιμότερη η χρήση μεθόδων *Text Retrieval*. Ειδικά στις περιπτώσεις Δανειστικών Βιβλιοθηκών στις οποίες απαιτείται η αποθήκευση και ανάκτηση και άλλων στοιχείων σχετικών με τους εκκρεμείς δανεισμούς τα οποία απο τη μία περιορίζουν το διαθέσιμο χώρο αποθήκευσης και απο την άλλη επιβάλλουν την επιτάχυνση της ανάκτησης των βιβλίων είναι επιβεβλημένη η χρήση μεθόδων *Data Retrieval*.

3. Η ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΤΟΥ ΣΥΣΤΗΜΑΤΟΣ

Σ' αυτή την ενότητα παρουσιάζεται μια σύντομη σκιαγράφιση των χαρακτηριστικών γνωρισμάτων του συστήματος και της οργάνωσης των αρχείων του. Οι στόχοι του συστήματος, των δύο προγραμμάτων τα οποία το αποτελούν, είναι η αποθήκευση, η ανάκτηση και η πρόσβαση μιας συλλογής απο βιβλία. Το ένα από τα δύο προγράμματα διαχειρίζεται τα βιβλία και το άλλο τα περιοδικά της βιβλιοθήκης.

Ειδικά το πρόγραμμα διαχείρισης βιβλίων είναι δανειστικό, δηλαδή μας επιτρέπει να παρακολουθούμε δανεισμούς βιβλίων σε αναγνώστες και να γνωρίζουμε ανα πάσα στιγμή αν ένα βιβλίο είναι δανεισμένο ή όχι. Ακόμη μπορούμε να γνωρίζουμε σε ποιόν είναι δανεισμένο ένα βιβλίο καθώς επίσης ποιά βιβλία οφείλει κάθε αναγνώστης.

Το πρόγραμμα διαχείρισης περιοδικών μας επιτρέπει να έχουμε *Cardex* για κάθε τίτλο περιοδικού. Στο *Cardex* κάθε περιοδικού φυλάγονται στοιχεία που αφορούν όλα τα παραληφθέντα τεύχη του και στοιχεία σχετικά με τις βιβλιοδετήσεις του για ένα μεγάλο χρονικό διάστημα. Για παράδειγμα για ένα μηνιαίο περιοδικό μπορούμε να έχουμε *Cardex* απο το 1950 μέχρι το 2020.

Και τα δύο προγράμματα επιτρέπουν την ανάκτηση βιβλίων (ή περιοδικών) με βάση οποιαδήποτε λέξη του τίτλου τους, οποιαδήποτε λέξη κλειδί (θεματικός κατάλογος), οποιοδήποτε πρόσωπο (Συγγραφέας, Εκδότης κ.λ.π.), με τον ταξινομικό του αριθμό και τέλος με συνδυασμό των παραπάνω. Μερικές ακόμα απο τις δυνατότητες των προγραμμάτων είναι η έκδοση ετικετών (καρτελάκια), τήρηση στατιστικών στοιχείων, εκτύπωση καταστάσεων τελευταίων παραλαβών στην περίπτωση των περιοδικών, εκτύπωση καταστάσεων καθυστερημένων βιβλίων καθώς και άλλες βοηθητικές εργασίες.

Τα θέματα που μας απασχόλησαν στην σχεδίαση και υλοποίηση του συστήματος και αφορούν τη δεικτοδότηση (*Indexing*) των δεδομένων, την ανάκτηση τους και την οργάνωση των αρχείων περιγράφονται στις επόμενες δύο υποενότητες.

3.1. Δεικτοδότηση και Δομές Αρχείων

Για τη δεικτοδότηση των εγγραφών (*records*) χρησιμοποιούνται τρία ανεστραμμένα αρχεία (*Inverted Files*) ή αρχεία ανεστραμμένων λιστών καθώς και άλλα δύο κοινά *Index* ένα για τον ταξινομικό και ένα για το κωδικός (*Identifier*) του βιβλίου. Το πρώτο *Inverted File* χρησιμοποιείται για την δεικτοδότηση των λέξεων του τίτλου, το δεύτερο για άλλες λέξεις κλειδιά (θέματα) και το τρίτο για τα επώνυμα των προσώπων που σχετίζονται με το βιβλίο είτε γιατί

είναι εκδότες, συγγραφείς, μεταφραστές κ.λ.π..

Τα ανεστραμμένα αρχεία υλοποιούνται στο σύστημα μας από εγγραφές σταθερού μήκους. Λέμε σταθερού μήκους γιατί κάθε εγγραφή ενός ανεστραμμένου αρχείου αποτελείται από τη λέξη που αντιπροσωπεύει, ένα πίνακα σταθερού μήκους του οποίου η κάθε θέση περιέχει τον κωδικό μιας εγγραφής του αρχείου των βιβλίων στην οποία (στο πεδίο της που αντιπροσωπεύεται από το εν λόγω ανεστραμμένο αρχείο) περιέχεται η λέξη και ένα μετρητή κατειλημμένων θέσεων. Στην περίπτωση που μια ανεστραμμένη λίστα είναι μεγαλύτερη από το μήκος του πίνακα των εγγραφών του ανεστραμμένου αρχείου δεσμεύεται και δεύτερη τρίτη κ.ο.κ. εγγραφή του ανεστραμμένου αρχείου για την αναπαράσταση της ανεστραμμένης λίστας. Την απόφαση να υλοποιήσουμε μ' αυτό τον τρόπο τα ανεστραμμένα αρχεία πήραμε γιατί το DataBase ToolBox[3] που χρησιμοποιήσαμε επέτρεπε το πρωτεύον κλειδί ενός Δεικτοδοτημένου Αρχείου να έχει επαναλήψεις (Duplicates). Για κάθε νέα εγγραφή που εισάγεται αφού βέβαια γίνει το αναγκαίο Data Validation εισάγεται ο κωδικός της εγγραφής και ο ταξινομικός στα αντίστοιχα Index και ενημερώνονται τα τρία Inverted Files με βάση τις λέξεις του τίτλου, τις λέξεις κλειδιά και τα επώνυμα προσώπων.

Λέγοντας ενημέρωση ενός Inverted File με μια λέξη εννοούμε την ανάκτηση της εγγραφής ή των εγγραφών που συνθέτουν την ανεστραμμένη λίστα γι' αυτή τη λέξη και την τοποθέτηση του κωδικού του βιβλίου στο οποίο ανήκει η λέξη στην πρώτη ελεύθερη θέση της πρώτης μη συμπληρωμένης από τις ανακτηθείσες εγγραφές. Στην περίπτωση που δε βρεθεί καμία εγγραφή γι' αυτή τη λέξη μιλάμε για τη δημιουργία μιας ανεστραμμένης λίστας που ισοδυναμεί με την τοποθέτηση μίας νέας εγγραφής στο Inverted File με τη νέα λέξη, μετρητή κατειλημμένων θέσεων ίσο με 1 και τον κωδικό του βιβλίου (στο οποίο ανήκει η λέξη) στην πρώτη θέση του πίνακα της εγγραφής. Το ίδιο γίνεται και στην περίπτωση που βρεθεί μία ή περισσότερες εγγραφές, γι' αυτή τη λέξη, στο Inverted File αλλά είναι όλες συμπληρωμένες.

Πρέπει να διευκρινιστεί ότι κάθε ανεστραμμένο αρχείο ενημερώνεται μόνο μία φορά ακόμα και όταν μια λέξη υπάρχει περισσότερες από μία φορές στο πεδίο της εγγραφής του βιβλίου που αντιστοιχεί στο ανεστραμμένο αρχείο. Ακόμα μια ομάδα λέξεων, οι λεγόμενες *Τετριμμένες Λέξεις* (άρθρα συνδεδετικά κ.λ.π.) δεν χρησιμοποιούνται για την ενημέρωση των Inverted Files.

3.2. Ανάκτηση

Η ανάκτηση βιβλίων (ή περιοδικών) μπορεί να γίνει με τον ταξινομικό τους αριθμό ή με σύνθετη αναζήτηση. Κατά την ανάκτηση με ταξινομικό αριθμό ο χρήστης πληκτρολογεί τον ταξινομικό αριθμό

αναζήτησης και το πρόγραμμα ανακτά όλα τα βιβλία των οποίων ο ταξινομικός (ή τα πρώτα ψηφία του) ταυτίζονται με τον ταξινομικό αριθμό αναζήτησης.

Όταν ο χρήστης επιλέξει σύνθετη αναζήτηση στην οθόνη του εμφανίζεται ένα παράθυρο το οποίο έχει τρεις κενές θέσεις για να συμπληρώσει λέξεις τίτλου, δύο κενές θέσεις για να συμπληρώσει λέξεις κλειδιά (θέματα) και δύο θέσεις για να συμπληρώσει επώνυμα προσώπων. Από τις επτά συνολικά κενές θέσεις ο χρήστης είναι υποχρεωμένος να συμπληρώσει τουλάχιστο μία χωρίς να υπάρχει περιορισμός ποιά ή ποιές θα συμπληρώσει. Οι θέσεις που συμπληρώνει συνδέονται με λογικό ΚΑΙ (Boolean AND). Παραδείγματος χάρη αν συμπληρώσει μόνο μία από τις τρεις θέσεις λέξεων τίτλου με τιμή "ΙΣΤΟΡ" και καμμιά από τις άλλες θέσεις τότε το πρόγραμμα θα ανακτήσει όλα τα βιβλία που στον τίτλο τους έχουν τη λέξη ΙΣΤΟΡΙΑ, ΙΣΤΟΡΙΚΟ, ΙΣΤΟΡΙΚΗ κ.ο.κ.. Αν πάλι συμπληρώσει μια λέξη τίτλου όπως στο προηγούμενο παράδειγμα και μια από τις δύο θέσεις προσώπων που σχετίζονται με το βιβλίο με τιμή "ΠΑΠ" τότε το πρόγραμμα θα ανακτήσει όλα τα βιβλία που έχουν στον τίτλο τους λέξεις που παράγονται από το "ΙΣΤΟΡ" και ταυτόχρονα κάποιο από τα συμμετέχοντα πρόσωπα έχει επώνυμο που αρχίζει από "ΠΑΠ" (ΠΑΠΑΝΙΚΟΛΑΟΥ, ΠΑΠΑΝΟΥΤΣΟΣ κ.λ.π.).

Όλα τα βιβλία τα οποία ανακτώνται με κάποια από τις παραπάνω μεθόδους είναι δυνατό να εκτυπωθούν ή να εμφανιστούν ένα-ένα στην οθόνη κάτω από τον έλεγχο του χρήστη.

4. ΣΥΜΠΕΡΑΣΜΑΤΑ

Με την οικονομική υποστήριξη του Κ.Ο.Ε. του Ο.Π.Α. (πρώην Α.Σ.Ο.Ε.Σ.) δημιουργήθηκε το σύστημα Μηχανοργάνωσης Βιβλιοθήκης που παρουσιάσαμε. Το σύστημα αυτό είναι εγκατεστημένο και σε άλλες δύο Βιβλιοθήκες εκτός της Βιβλιοθήκης του Ο.Π.Α.. Η πρώτη από αυτές είναι στη βιβλιοθήκη του Ο.Α.Ε.Α. στους Τράχωνες. Η δεύτερη στο Λεόντειο Λύκειο Πατησίων είναι η πιο ανεπτυγμένη από πλευράς όγκου στοιχείων, όσον αφορά το πρόγραμμα διαχείρισης βιβλίων, αριθμοί περίπου 7000 βιβλία και διαρκώς εισάγονται μέχρις ότου ολοκληρωθεί όλη η αναδρομική δουλειά. Όσον αφορά το πρόγραμμα διαχείρισης περιοδικών η βιβλιοθήκη του Ο.Π.Α. έχει το μεγαλύτερο όγκο στοιχείων με πλέον των 550 τίτλων περιοδικών με πλήρως ενημερωμένο Cardex.

Μια από τις αντικειμενικές δυσκολίες σ'όλες τις περιπτώσεις είναι η αναδρομική δουλειά, δηλαδή η εισαγωγή των παλαιών βιβλίων στο σύστημα, κάτι που δυσχεραίνει ακόμη περισσότερο στις περιπτώσεις που το προσωπικό είναι μειωμένο. Πρέπει να σημειωθεί ότι το σύστημα που περιγράψαμε δεν διαχειρίζεται τις παραγγελίες βιβλίων δεδομένου ότι η Διοίκηση του Πανεπιστημίου σχεδιάζει να αντιμετωπίσει όλα τα συστήματα προμηθειών με ενιαίο σύστημα.

BIBΛΙΟΓΡΑΦΙΚΕΣ ΑΝΑΦΟΡΕΣ

- [1] Agosti, M., Grestani, F., and Gradenigo, G. Towards Data Modelling in information retrieval. *Journal of Information Science* 15 (1989) 307-319.
- [2] Antonacci, F. et. al. A system for text analysis and lexical knowledge acquisition. *Data & Knowledge Engineering* 4 (1989) 1-20.
- [3] Borland, Turbo Pascal DataBase Toolbox 1987.
- [4] Chartron, G. Lexicon management tools for large textual databases: the Lexinet system. *Journal of Information Science* 15 (1989) 339-344.
- [5] Karanikolas, N.N., and Cavouras J.C. An Integrated Natural Language Processing System. Submitted for publication.
- [6] Lucarella, D. A document retrieval system based on nearest neighbour searching. *Journal of Information Science* 14 (1988) 25-33.
- [7] Salton, G. Another look at automatic text retrieval systems. *Commun. ACM* 29, 7 (July 1986). 648-656.