

Προβλήματα διαλειτουργικότητας κατά την ταυτόχρονη πρόσβαση σε πηγές μέσω του πρωτοκόλλου Z39.50 & Το περιβάλλον πρόσβασης "Η ΑΡΓΩ"

Μανόλης Πεπονάκης

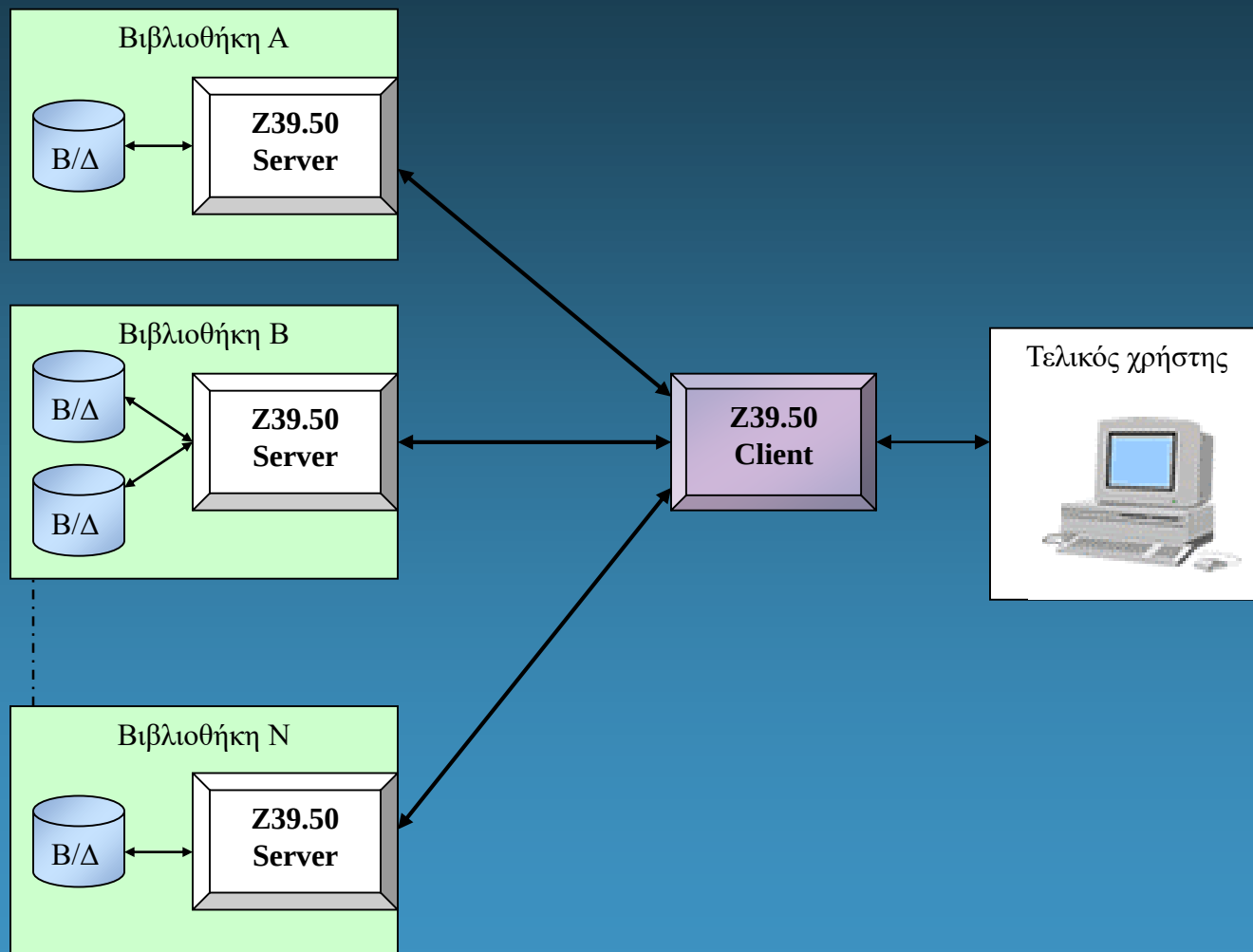
Ντίνα Τρούτπεγλη

Μιχάλης Σφακάκης



Εθνικό Κέντρο Τεκμηρίωσης

Περιβάλλον λειτουργίας του πρωτοκόλλου Z39.50



Χαρακτηριστικά ερώτησης (Bib-1)

- Use = Ευρετήριο
- Truncation = Αποκοπή
- Structure = Δομή
- Position = Θέση
- Completeness = Πληρότητα
- Relation = Σχέση

Παράδειγμα ερώτησης

- Όρος αναζήτησης => Parker, John
 - Ευρετήριο => Συγγραφέας
 - Δομή => Φράση
 - Αποκοπή => Όχι
 - Πληρότητα => Ατελής υποπεδίο
 - Θέση => Οπουδήποτε στο πεδίο
 - Σχέση => Ίσο

Τα πιο διαδεδομένα πεδία

- Σύμφωνα με την Index Data πέντε από τα πιο διαδεδομένα πεδία αναζήτησης είναι τα εξής :
 - Τίτλος (αναζήτηση το 97% των servers)
 - Θέμα (αναζήτηση το 94% των servers)
 - ISBN (αναζήτηση το 90% των servers)
 - Συγγραφέας (αναζήτηση το 88% των servers)
 - Οποιοδήποτε (αναζήτηση το 81% των servers)

Ταυτόχρονη πρόσβαση

- Επιλογή δέκα ακαδημαϊκών βιβλιοθηκών
 - Κριτήρια επιλογής
 - Χωροταξική κατανομή (όλη η Ελλάδα)
 - Πολλαπλότητα συστημάτων αυτοματοποίησης (πέντε διαφορετικά συστήματα)
 - Διαφορετικές διατάξεις βιβλιογραφικών εγγραφών (MARC 21 & UNIMARC)

Διαδικασία Αναζήτησης

- Ποια είναι τα κοινά ευρετήρια που μπορούν να αναζητηθούν στις Ελληνικές Ακαδημαϊκές βιβλιοθήκες?
 - Τίτλος (100%)
 - Συγγραφέας (100%)
 - Θέμα (100%)
 - ISBN (90%)
 - Οποιοδήποτε (60%)

Αναζήτηση με βάση το Συγγραφέα

- Στόχοι
 - Να βρούμε τα έργα του συγγραφέα X
 - Να βρούμε μόνο τα έργα του συγγραφέα X
 - Να βρούμε όλα τα καταλογογραφημένα έργα του συγγραφέα X

1η προσέγγιση

Ερώτηση με μοναδικό χαρακτηριστικό αυτό του ευρετηρίου

- π.χ. Συγγραφέας => Smith, John
 - Οι 2 στους 10 servers εκτελούν την ερώτηση ως φράση με δεξιά αποκοπή
 - Οι 8 στους 10 servers εκτελούν την ερώτηση ως λέξεις, χωρίς αποκοπή
 - ❖ Αποτέλεσμα 245 εγγραφές εκ των οποίων ένας μεγάλος αριθμός δεν έχει για συγγραφέα τον Smith, John (λόγω του ότι οι servers θεώρησαν ότι πρόκειται για λέξεις)
- ✓ Συμπέρασμα : υπερβολικός θόρυβος

2η προσέγγιση

Ερώτηση με προσθήκη και χαρακτηριστικού δομής

- π.χ. Συγγραφέας => Smith, John & Δομή => Φράση
 - ❖ Αποτέλεσμα 48 εγγραφές
 - Δεν εντοπίστηκαν, σε 5 βιβλιοθήκες, εγγραφές που είχαν για συγγραφέα το: Smith, John, 1945-
 - Δεν εντοπίστηκαν, σε 7 βιβλιοθήκες, εγγραφές που είχαν για συγγραφέα το: Parker - Smith, John
- ✓ Συμπέρασμα: υπερβολική σιωπή

3η προσέγγιση

Χρήση των καθιερωμένων αποδόσεων για τον εντοπισμό του ακριβούς ονόματος του Συγγραφέα

- Αναζήτηση του βιβλίου με τίτλο *"Two studies in ancient Macedonian topography"*
 - Το βιβλίο εντοπίστηκε σε 4 βιβλιοθήκες
 - Το όνομα του συγγραφέα εμφανιζόταν:
 - Χατζόπουλος, Μ.Β. (Μιλτιάδης Β.)
 - Χατζόπουλος, Μιλτιάδης Β [2 βιβλιοθήκες]
 - Χατζόπουλος , Μιλτιάδης Β. , 1944-
- ✓ Συμπέρασμα: απουσία κοινών Αρχείων Καθιερωμένων Αποδόσεων

Αναζήτηση με βάση το θέμα

Λεκτικές αποδόσεις

- Απουσία Αρχείων Καθιερωμένων Αποδόσεων
(ύπαρξη μη ελεγχόμενων θεματικών όρων
π.χ. πεδία 610 του UNIMARC)
- Απουσία κοινών Αρχείων Καθιερωμένων Αποδόσεων
- Πολλαπλότητα γλωσσών (ελληνικές και αγγλικές αποδόσεις)
- ✓ Συμπέρασμα: αδύνατη η ταυτόχρονη αναζήτηση με βάση τη λεκτική απόδοση των θεμάτων

Αναζήτηση με βάση το θέμα

Ταξινομικοί αριθμοί

- Χρήση τριών διαφορετικών ταξινομικών συστημάτων (LC, DDC, UDC)
- Οι 3 στους 10 servers δεν υποστηρίζουν αναζήτηση με ταξινομικό αριθμό
 - (ο ένας από αυτούς δεν στέλνει καν διαγνωστικό, όμως δεν επιστρέφει αποτελέσματα σε καμία αναζήτηση!)
- ✓ Συμπέρασμα: ανέφικτη η ταυτόχρονη αναζήτηση με ταξινομικό αριθμό

Γενικά Συμπεράσματα για τη Διαδικασία της Αναζήτησης

- Κάθε Z39.50 server διαθέτει τα δικά του προεπιλεγμένα χαρακτηριστικά
- Δεν υποστηρίζονται οι ίδιοι συνδυασμοί χαρακτηριστικών από τους επιμέρους servers
- Απουσία κοινών Αρχείων Καθιερωμένων Αποδόσεων
 - Αδύνατη η ταυτόχρονη προσέγγιση των καταλόγων θεματικά
 - Θορυβώδης ή υπερβολικά σιωπηρή η προσέγγιση με βάση τον συγγραφέα
- ✓ Αποτέλεσμα: αδυναμία ουσιαστικής αναζήτησης

Διαδικασία Ανάκτησης

- Στόχος να ενταχθούν τα επιμέρους αποτελέσματα σε ένα ενιαίο ομοιογενές σύνολο

Προβλήματα

- Διαφορετικά MARC
- Διαφορετικά σύνολα χαρακτήρων
- Ασυμβατότητα στα Holdings
- Μη έγκυρες εγγραφές (είτε κατά ISO 2709 είτε κατά MARC)
- ✓ Συμπέρασμα: εξαιρετικά δυσχερής η ουσιαστική ομογενοποίηση των αποτελεσμάτων

Γενικά συμπεράσματα

- Η πρόσβαση, μέσω Z39.50, μπορεί να επιτευχθεί σε ένα πρώτο στάδιο
- Κατά την ταυτόχρονη πρόσβαση σε πολλές πηγές τα προβλήματα πολλαπλασιάζονται
- Όσο οι αναζητήσεις αποκτούν ακρίβεια τόσο μειώνεται και η συμβατότητα των servers

«Η ΑΡΓΩ»

- Το νέο web περιβάλλον πρόσβασης σε πηγές πληροφόρησης
 - Πύλη ταυτόχρονης πρόσβασης σε πηγές που διατίθενται μέσω Z39.50
- URL: <http://argo.ekt.gr>

«Η ΑΡΓΩ»: Στόχοι

- Ταυτόχρονη πρόσβαση σε πηγές που διατίθενται μέσω Z39.50
 - Εργαλείο έρευνας
 - Ανάγκες πληροφόρησης
 - Εργαλείο ανάπτυξης καταλόγων
 - Ανάπτυξη των τοπικών καταλόγων με χρήση «έτοιμων» βιβλιογραφικών εγγραφών (copy cataloguing)

«Η ΑΡΓΩ»

ως εργαλείο έρευνας

- Δυνατότητα αποστολής ερωτήσεων χωρίς χαρακτηριστικά
 - Πλεονεκτήματα: ευκολία στη χρήση, απάντηση από μεγάλο αριθμό servers
 - Μειονεκτήματα: αμφιβόλου ποιότητας αποτελέσματα
- Δυνατότητα αποστολής σύνθετων ερωτήσεων
 - Πλεονεκτήματα: ακρίβεια
 - Μειονεκτήματα: σχετική ασυμβατότητα υλοποιήσεων (δεν υποστηρίζονται τα χαρακτηριστικά από τους servers) & απαιτείται υψηλή τεχνογνωσία από το χρήστη ο οποίος υποβάλει το ερώτημα

«Η ΑΡΓΩ»

ως εργαλείο ανάπτυξης του καταλόγου

- Παρέχει στο χρήστη τη δυνατότητα να αποθηκεύσει σε διάταξη UNIMARC τα αποτελέσματα της αναζήτησης
 - Αν η πηγή στέλνει τα δεδομένα σε MARC 21 αυτά μετατρέπονται αυτόματα σε UNIMARC
- Ενοποιεί το σύνολο χαρακτήρων των επιμέρους αποτελεσμάτων
 - Ο χρήστης έχει τη δυνατότητα να επιλέξει, κατά την αποθήκευση, το σύνολο χαρακτήρων που «διαβάζει» το σύστημα του
- ✓ Αποτέλεσμα: ο χρήστης μπορεί να αποθηκεύσει στο δίσκο του ένα ομογενοποιημένο αρχείο (τόσο σε σύνολο χαρακτήρων όσο και σε διάταξη εγγραφών) το οποίο μπορεί να εισαγάγει χωρίς επιπλέον μετατροπές

Συνοψίζοντας

- Το πρωτόκολλο Z39.50 παρέχει τις δυνατότητες για την ταυτόχρονη πρόσβαση αλλά:
 - Υπάρχει ασυμβατότητα
 - Σε επίπεδο υλοποίησης του πρωτοκόλλου από τα συστήματα (π.χ. χαρακτηριστικά)
 - Σε βιβλιοθηκονομικό επίπεδο (π.χ. σημεία πρόσβασης)
- Για τη βελτιστοποίηση της επικοινωνίας απαιτείται συνεργασία:
 - Σε επίπεδο παραμετροποίησης συστημάτων
 - Θα βοηθούσε η δημιουργία (ή η υιοθέτηση) ενός profile (π.χ. Bath Profile)
 - Σε σημασιολογικό επίπεδο (semantics)