

Αξιολόγηση διαλειτουργικότητας μεταδεδομένων μεταξύ των  
ιδρυματικών αποθετηρίων και των καταλόγων (OPACs) των  
ελληνικών ακαδημαϊκών βιβλιοθηκών

**Μανόλης Πεπονάκης & Μιχάλης Σφακάκης**  
Τμήμα Αρχειονομίας Βιβλιοθηκονομίας / Ιόνιο Πανεπιστήμιο  
{peponakis, sfakakis} @ ionio.gr

17ο Πανελλήνιο Συνέδριο Ακαδημαϊκών Βιβλιοθηκών  
24-26 Σεπτεμβρίου 2008, Ιωάννινα

# Ορισμοί

- OPACs (On line Public Access Catalogues)
  - η πρόσβαση σε μια βάση δεδομένων η οποία αποτελείται από τις βιβλιογραφικές εγγραφές που περιγράφουν το υλικό της βιβλιοθήκης. Κατά κανόνα η πρόσβαση δίνεται μέσω web από γραφικό περιβάλλον (WebPac)
- Ιδρυματικά Αποθετήρια
  - a set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members (Clifford Lynch)
- Διαλειτουργικότητα
  - η δυνατότητα δύο ή περισσότερων συστημάτων (ή συστατικών στοιχείων συστημάτων) να ανταλλάσσουν πληροφορίες και να χρησιμοποιούν τις πληροφορίες που αντάλλαξαν



# Στόχος της έρευνας

- Η καταγραφή της υφιστάμενης κατάστασης στη δημιουργία καταλόγων για τα αποθετήρια
  - Η συγκρότηση των OPACs θεωρείται σε μεγάλο βαθμό δεδομένη και γνωστή
- Η ανίχνευση του βαθμού της διαλειτουργικότητας στις ελληνικές ακαδημαϊκές βιβλιοθήκες σε δύο επίπεδα:
  1. Διαλειτουργικότητα μεταξύ OPAC και καταλόγου αποθετηρίου στα πλαίσια του ίδιου ιδρύματος
  2. Διαλειτουργικότητα των καταλόγων (OPACs & αποθετηρίων) των πανεπιστημιακών βιβλιοθηκών σε εθνικό επίπεδο



# Μέθοδος έρευνας και επιλογή δείγματος

- Μέθοδος έρευνας
  - Ερωτηματολόγια
  - Απευθείας πρόσβαση στους καταλόγους
- Επιλογή δείγματος
  - 1<sup>ο</sup> κριτήριο: δραστηριοποίηση σε έργα που σχετίζονται με τις νέες τεχνολογίες
    - Επελέγησαν 10 πανεπιστημιακές βιβλιοθήκες
  - 2<sup>ο</sup> κριτήριο: ύπαρξη ιδρυματικού αποθετηρίου
    - Μόνο 6 από τις παραπάνω



# ΟΡΑΣ και κατάλογος αποθετηρίου στο ίδιο ίδρυμα

## Διατάξεις μεταδεδομένων και κανόνες καταλογογράφησης

- Οι δύο στις έξι βιβλιοθήκες (ποσοστό 33%) καταλογογραφούν στην ίδια διάταξη μεταδεδομένων (MARC 21) τόσο για τον κατάλογο της βιβλιοθήκης όσο και για τον κατάλογο του αποθετηρίου.
- Οι τέσσερις στις έξι (ποσοστό 67%) χρησιμοποιούν MARC (είτε MARC 21 είτε UNIMARC) για τον κατάλογο της βιβλιοθήκης και DC για τον κατάλογο του αποθετηρίου.
  - Μία από αυτές τις βιβλιοθήκες προτίθεται να διαθέσει τα δεδομένα σε MODS.
- Οι πέντε στις έξι χρησιμοποιούν τους AACR για τον κατάλογο του αποθετηρίου
  - Χρήζει διερεύνησης το πώς οι AACR εφαρμόζονται στο DC



## Δημιουργία σημείων πρόσβασης

- Καθιέρωση ονομάτων προσώπων
  - Οι ελληνικές ακαδημαϊκές βιβλιοθήκες τηρούν διαφορετικές πολιτικές καθιέρωσης στον κατάλογο της βιβλιοθήκης και σε αυτόν του αποθετηρίου
    - Το ένα τρίτο (ποσοστό 33%) των βιβλιοθηκών του δείγματος δημιουργεί επικεφαλίδες για τα ελληνικά ονόματα με διαφορετικό τρόπο στους δύο καταλόγους
    - Τα δύο τρίτα (ποσοστό 67%) των βιβλιοθηκών του δείγματος δημιουργούν επικεφαλίδες για τα ξενόγλωσσα ονόματα με διαφορετικό τρόπο στους δύο καταλόγους!



# ΟΡΑΣ και κατάλογος αποθετηρίου στο ίδιο ίδρυμα

## Δημιουργία σημείων πρόσβασης

### ▪ Απόδοση θεμάτων

- Δύο στις έξι βιβλιοθήκες χρησιμοποιούν ακριβώς την ίδια πολιτική
- Οι τέσσερις στις έξι ακολουθούν περισσότερο ή λιγότερο αποκλίνουσες πρακτικές
- Όλες οι βιβλιοθήκες του δείγματος χρησιμοποιούν ελεγχόμενο λεξιλόγιο στον κατάλογο της βιβλιοθήκης ενώ οι τρεις από τις έξι (ποσοστό 50%) χρησιμοποιούν μη ελεγχόμενο λεξιλόγιο στον κατάλογο του αποθετηρίου
  - Μόνο μία από αυτές χρησιμοποιεί αποκλειστικά μη ελεγχόμενο λεξιλόγιο για τον κατάλογο του αποθετηρίου



# ΟΡΑΣ και κατάλογος αποθετηρίου στο ίδιο ίδρυμα

## Δημιουργία σημείων πρόσβασης

- Ταξινόμηση του υλικού
  - Οι βιβλιοθήκες του δείγματος χρησιμοποιούν LCC, DDC και ένα τοπικό ταξινομικό σχήμα.
  - Το 67% των βιβλιοθηκών παράγει ταξινομικούς αριθμούς τόσο για τον κατάλογο της βιβλιοθήκης όσο και του αποθετηρίου με βάση το ίδιο ταξινομικό σχήμα
    - Η μία βιβλιοθήκη που χρησιμοποιεί LCC χρησιμοποιεί και ένα τοπικό. Η ίδια βιβλιοθήκη για το αποθετήριο χρησιμοποιεί μόνο το τοπικό.
  - Μία από τις βιβλιοθήκες δε χρησιμοποιεί καθόλου ταξινόμηση στον κατάλογο του αποθετηρίου, ενώ στον κατάλογο της βιβλιοθήκης χρησιμοποιεί το DDC





## Διαλειτουργικότητα συστημάτων

- Καμία από τις βιβλιοθήκες του δείγματος δεν διαθέτει κοινό λογισμικό για τους δύο καταλόγους (αποθετηρίου και βιβλιοθήκης)
- Οι 4 βιβλιοθήκες (67%) δίνουν τη δυνατότητα ταυτόχρονης αναζήτησης και στους δύο καταλόγους (βιβλιοθήκης και αποθετηρίου)
  - όμως, μόνο η μία από αυτές κάνει ενοποίηση των μεταδεδομένων από το πρόγραμμα της ταυτόχρονης αναζήτησης
    - Η βιβλιοθήκη που το επιτυγχάνει διαθέτει σε πολύ μεγάλο ποσοστό κοινές πολιτικές στη δημιουργία των μεταδεδομένων για τους δύο καταλόγους



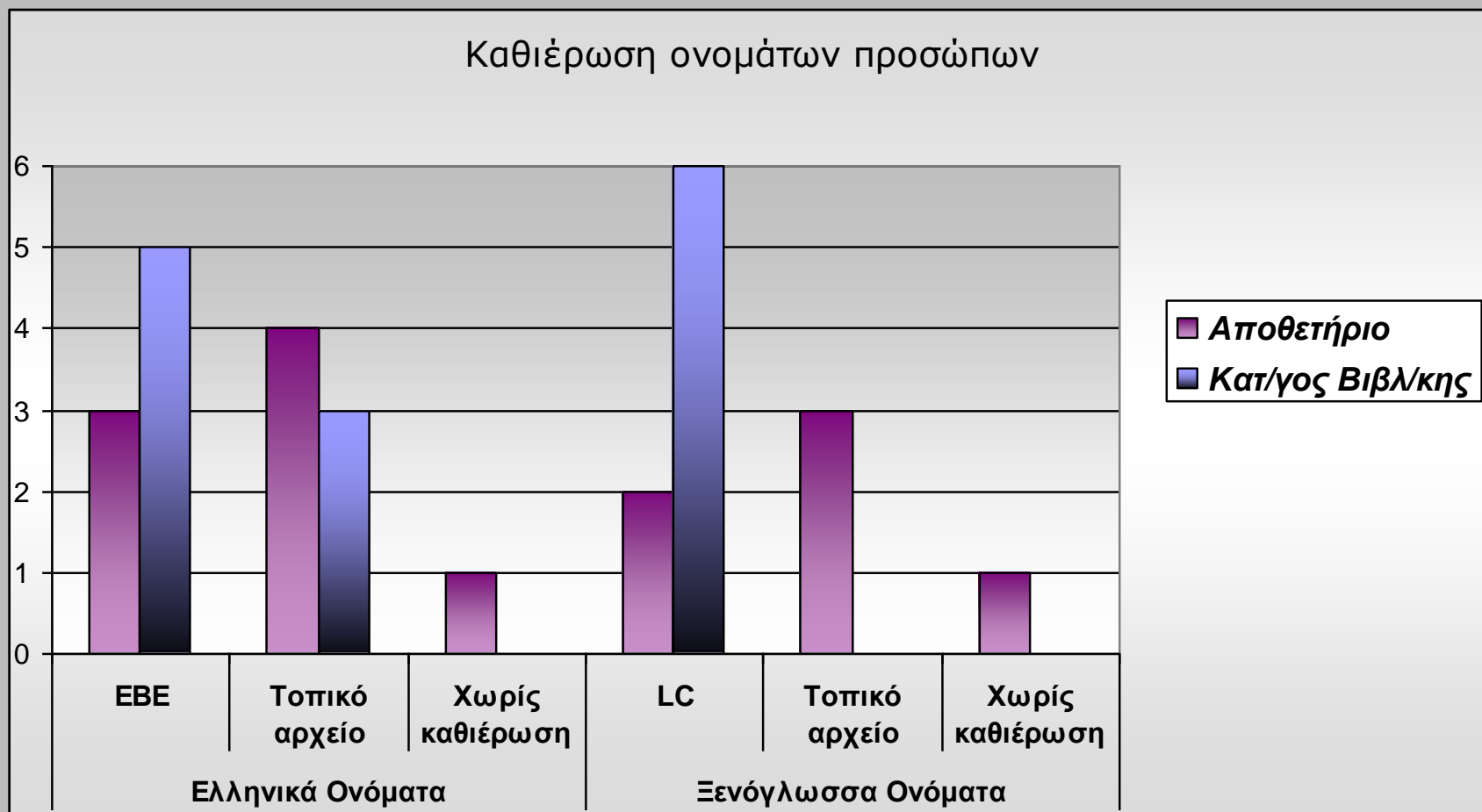
# Διαλειτουργικότητα μεταξύ OPACs και καταλόγων αποθετηρίων πανελλαδικά

## Πρωτόκολλα επικοινωνίας και συστήματα

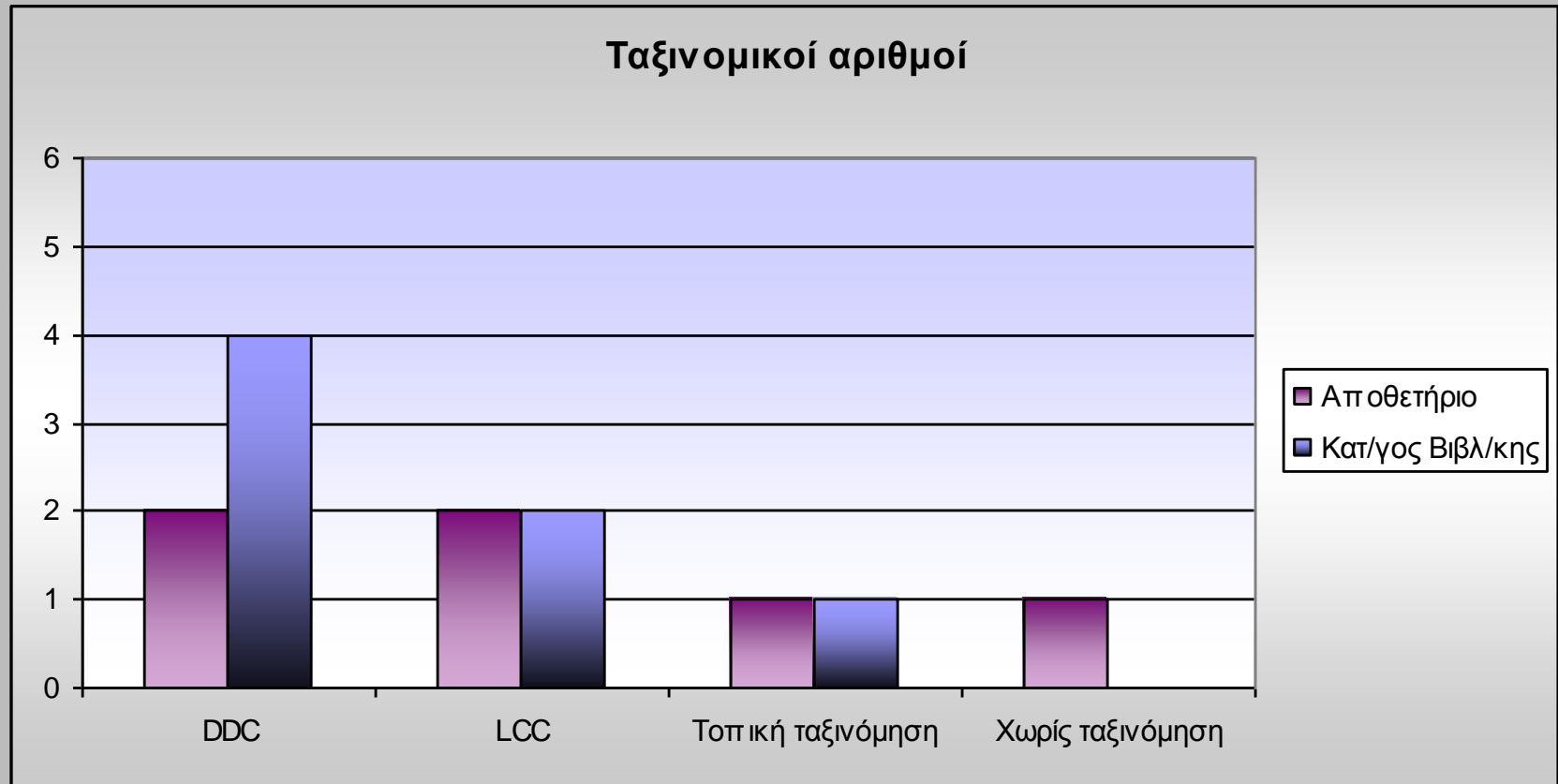
- Όλες οι βιβλιοθήκες του δείγματος παρέχουν πρόσβαση και στους δύο καταλόγους (αποθετηρίου και βιβλιοθήκης) μέσω ανοικτού προτύπου
  - Όλα τα αποθετήρια διαθέτουν τα μεταδεδομένα τους μέσω του πρωτοκόλλου συγκομιδής (harvesting) OAI-PMH
  - Για την αναζήτηση-ανάκτηση τα αποθετήρια διαθέτουν τα δεδομένα μέσω SRW/SRU (εκτός από ένα αποθετήριο)
  - Τα δεδομένα όλων των OPACs του δείγματος διατίθενται μέσω του πρωτοκόλλου Z39.50
- Ομοιογένεια στα συστήματα διαχείρισης του αποθετηρίου
  - 5 βιβλιοθήκες (ποσοστό 83%) χρησιμοποιούν DSpace και η μία CDS-WARE
    - Η μία βιβλιοθήκη δρομολογεί μετάβαση από το DSpace στο FEDORA
- Στα συστήματα βιβλιοθηκών παρατηρείται μεγαλύτερη ποικιλία
  - 3 Horizon, 2 Advance και 1 Aleph



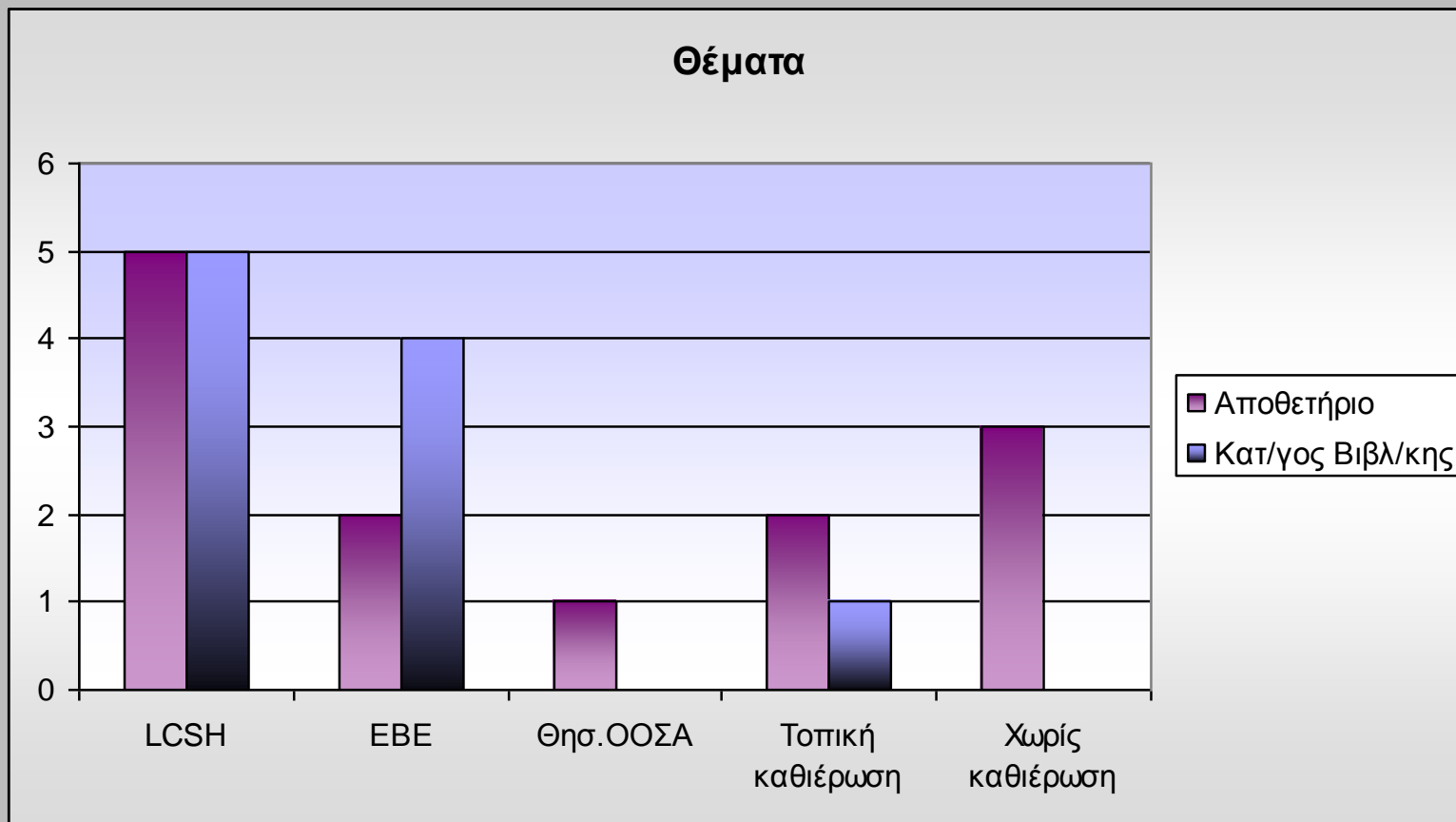
# Δημιουργία σημείων πρόσβασης: ονόματα



# Δημιουργία σημείων πρόσβασης: ταξινομικοί αριθμοί



# Δημιουργία σημείων πρόσβασης: θέματα



- Δεν υπάρχει ομοιομορφία στη γλώσσα των θεμάτων. Σε κάποιες βιβλιοθήκες χρησιμοποιείται και η ελληνική και η αγγλική στην ίδια εγγραφή, σε άλλες μόνο η αγγλική ή μόνο η ελληνική ενώ μία βιβλιοθήκη χρησιμοποιεί μόνο αγγλικά στον OPAC ενώ στον κατάλογο του αποθετηρίου η γλώσσα εξαρτάται από τον τύπο του υλικού

# Συμπεράσματα Ι

- Τα αποθετήρια των ελληνικών πανεπιστημίων εμφανίζουν ανομοιογένεια τόσο μεταξύ τους όσο και σε αντιστοιχία με τον κατάλογο της βιβλιοθήκης του οικείου ιδρύματος
  - Το υλικό της βιβλιοθήκης δεν αποτυπώνεται πλέον στον ένα και μοναδικό κατάλογο αλλά περιγράφεται από ένα σύνολο πηγών (OPACs, κατάλογοι αποθετηρίων, βάσεις άρθρων περιοδικών κ.α.). Γίνεται πιο επιτακτική η ανάγκη για δημιουργία συστημάτων ενιαίας πρόσβασης
    - Για να είναι όμως αποτελεσματική η πρόσβαση στο σύνολο του υλικού απαιτείται να αυξηθούν σημαντικά τα επίπεδα διαλειτουργικότητας των επιμέρους συστημάτων και των μεταδεδομένων που αυτά κομίζουν, καταρχάς στο πλαίσιο του ιδρύματος και στη συνέχεια στο σύνολο των ελληνικών ακαδημαϊκών βιβλιοθηκών
- Σημασιολογική διαλειτουργικότητα σε πολύ χαμηλά επίπεδα
  - Διαφορετικά θεματικά και ταξινομικά συστήματα, διαφορετικές γλώσσες κτλ.
  - Διαφαίνεται μια σαφής τάση απλούστευσης στα δεδομένα των αποθετηρίων με άμεσες συνέπειες στη σημασιολογική διαλειτουργικότητα
    - Στην εποχή των προσπαθειών για το σημασιολογικό ιστό (semantic web), των οντολογιών και εν γένει της εκφραστικής δομής, οι βιβλιοθήκες -φορείς με πραγματική παράδοση στα δομημένα δεδομένα- απλουστεύουν τόσο τα δεδομένα που παράγουν όσο και τις μεθόδους παραγωγής τους



# Συμπεράσματα II

- ▣ Διαλειτουργικότητα δομής φαινομενικά υψηλή
  - Σχετικά λίγα πρότυπα μεταδεδομένων (MARC & DC)
  - Αν, όμως, ληφθούν υπόψη τα τοπικά ορισμένα στοιχεία (π.χ. DC qualifiers) και συνδυαστούν με την απουσία κοινών πολιτικών μειώνεται ουσιαστικά ο βαθμός της
    - όλες οι βιβλιοθήκες του δείγματος που κάνουν καταλογογράφηση σε DC χρησιμοποιούν qualifiers και μάλιστα τοπικά ορισμένους
  - Το σύνολο των βιβλιοθηκών του δείγματος διαθέτει τα δεδομένα του καταλόγου του αποθετηρίου του με το OAI-PMH
    - Το συγκεκριμένο πρωτόκολλο παρότι είναι εγκατεστημένο σε όλες τις βιβλιοθήκες του δείγματος δεν χρησιμοποιείται από καμία αφού ούτε μία δεν κάνει harvesting μεταδεδομένων
      - Μήπως τα χαμηλά επίπεδα ομοιογένειας και εσωτερικής συνέπειας (consistency) των μεταδεδομένων οδηγούν και σε αδυναμία ανάπτυξης υπηρεσιών προστιθέμενης αξίας όπως για παράδειγμα η ανάπτυξη ενός θεματικού αποθετηρίου με χρήση διαδικασιών harvesting;



# Συμπεράσματα III

- Ο κατάλογος της βιβλιοθήκης έχει αντικατασταθεί από ένα σύνολο πηγών μέσα στο οποίο συμπεριλαμβάνεται ο τυπικός κατάλογος, αυτός του αποθετηρίου κ.α.
  - Για εφικτή ενιαία αναζήτηση απαιτούνται shareable μεταδεδομένα (δηλαδή μεταδεδομένα που μπορούν να κατανοηθούν και να χρησιμοποιηθούν έξω από το άμεσο περιβάλλον παραγωγής τους) ώστε να είναι δυνατό να αναπτυχθούν αναβαθμισμένες – προστιθέμενης αξίας υπηρεσίες
    - Το ΟΑΙ-ΡΜΗ ως πρωτόκολλο συγκομιδής (harvesting) -και όχι αναζήτησης- στοχεύει ακριβώς σε αυτή την έννοια του “shareability”
- Όλοι οι επιμέρους κατάλογοι (αποθετηρίου, βιβλιοθήκης κτλ.) θα πρέπει να μπορούν να ιδωθούν και ως συστατικά ενός και μοναδικού ενιαίου καταλόγου
  - Ο κατάλογος αυτός θα πρέπει να μπορεί να αντιμετωπιστεί από τον τελικό χρήστη ως μια αδιαίρετη ενότητα, όπου μία και μοναδική αναζήτηση θα του επιστρέψει το σύνολο των σχετικών πηγών που διαθέτει η βιβλιοθήκη ανεξάρτητα από τους πιθανούς τύπους αυτού του υλικού

