

**Μανόλης Πεπονάκης, Παναγιώτης Στάικος,
Μιχάλης Σφακάκης, Σαράντος Καπιδάκης**
*Εργαστήριο Ψηφιακών Βιβλιοθηκών και Ηλεκτρονικής Δημοσίευσης
Τμήμα Αρχιαιονομίας – Βιβλιοθηκονομίας, Ιόνιο Πανεπιστήμιο
{peponakis, pstaikos, sfakakis, sarantos}@ionio.gr*

FRBRization: **αναζητώντας το “κλειδί” για τον εντοπισμό των Έργων**

19ο Πανελλήνιο Συνέδριο Ακαδημαϊκών Βιβλιοθηκών

Αθήνα, 3-5 Νοεμβρίου 2010

Πάντειο Πανεπιστήμιο

FRBR

■ Τι είναι οι FRBR

- Functional Requirements for Bibliographic Records (FRBR)
 - ◆ Ένα εννοιολογικό μοντέλο για τη δημιουργία βιβλιογραφικών εγγραφών
 - “an entity-relationship model as a generalized view of the bibliographic universe, intended to be **independent of any cataloging code or implementation**” (Tillet, 2004)
 - ◆ Αναπτύχθηκαν από την IFLA
 - IFLA Study Group on Functional Requirements for Bibliographic Records
 - Το final report εκδόθηκε το 1998

■ Τι δεν είναι οι FRBR

- Σχήμα κωδικοποίησης μεταδεδομένων
- Κανόνες καταλογογράφησης
 - ◆ Οι RDA (Resource Description and Access), διάδοχοι των AACR, έχουν ενσωματώσει τη φιλοσοφία των FRBR

FRBR αξιώματα

- Υπάρχουν διακριτές οντότητες (το μοντέλο ορίζει δέκα)
- Κάθε οντότητα έχει μια σειρά από ιδιότητες (attributes)
- Οι οντότητες συσχετίζονται μεταξύ τους



ΟΙ ΟΝΤΟΤΗΤΕΣ: σε τρεις ομάδες

Ομάδα 1

- Work = Έργο
- Expression = Έκφραση
- Manifestation = «;»
- Item = Αντίτυπο

Η πρώτη ομάδα αφορά στις διαφορετικές οπτικές για το πνευματικό ή καλλιτεχνικό δημιούργημα

Ομάδα 2

- Person = Πρόσωπο
- Corporate Body = Συλλογικό Όργανο

Η δεύτερη ομάδα αφορά στα πρόσωπα που σχετίζονται με το πνευματικό ή καλλιτεχνικό δημιούργημα

Ομάδα 3

- Concept = Έννοια
- Object = Αντικείμενο
- Event = Γεγονός
- Place = Τόπος

Η τρίτη ομάδα αφορά στα όσα πραγματεύεται το πνευματικό ή καλλιτεχνικό δημιούργημα

Σημείωση: όλες οι οντότητες μπορούν να ανήκουν στην τρίτη ομάδα

Βασικές σχέσεις οντοτήτων 1^{ης} ομάδας



FRBR: βασικές καινοτομίες

- Αφαιρετική οπτική για την καταλογογράφηση, θεμελιωμένη σε στιβαρό θεωρητικό υπόβαθρο
 - Από την απαρίθμηση κανόνων στον ορισμό κάθε πληροφορίας και του σκοπού που αυτή εξυπηρετεί
- Αποστασιοποίηση από το φυσικό αντικείμενο
 - Δυνατότητα περιγραφής αφαιρετικών εννοιών (π.χ. Έργο)
- Εισαγωγή νέων εννοιών-οντοτήτων *(σε σχέση με MARC & AACR)*
 - Work, Expression, Event, Object, Concept
- Διάκριση σημασιολογίας – σύνταξης

FRBR: προβλήματα

- Ευρύ πλαίσιο με δυσκολίες στους ορισμούς
 - π.χ. πότε δημιουργείται νέα Έκφραση;
- Απουσία σχημάτων κωδικοποίησης
 - Σε κάποιες περιπτώσεις υλοποιούνται μέσω τοπικών επεκτάσεων του MODS ή με RDF
- ❖ Δεν τίθεται θέμα επανακαταλογογράφησης των τεκμηρίων με βάση τους FRBR
 - Είναι αναγκαίο να αναπτυχθούν αλγόριθμοι μετατροπής για τα εκατομμύρια των MARC εγγραφών

FRBRization

- Ορισμός FRBRization
 - Διαδικασία ανίχνευσης FRBR οντοτήτων, ιδιοτήτων και σχέσεων, από εγγραφές οι οποίες έχουν καταλογογραφηθεί πρωτογενώς σε διάφορα σχήματα κωδικοποίησης

- Στάδια FRBRization
 - Εντοπισμός των εγγραφών που αποτελούν ένα Έργο
 - Εντοπισμός των εγγραφών που αποτελούν Εκφράσεις εντός του συγκεκριμένου Έργου
 - Εντοπισμός των εγγραφών που αποτελούν τα πολλαπλά Manifestations
 - ◆ Μόνο στην περίπτωση πολλών πηγών ή λανθασμένης καταλογογράφησης
 - Αντιστοίχιση των πληροφοριών (πεδίων-υποπεδίων) των βιβλιογραφικών εγγραφών στις αντίστοιχες οντότητες και ιδιότητες των FRBR
 - Δημιουργία της νέας δομής

Παράδειγμα

3 εγγραφές: 1 Έργο, 2 Εκφράσεις, 3 Manifestations

Εγγραφή 1 - Βιβλίο	
Τίτλος / Συγγραφέας	Ομήρου Ιλιάδα / μετάφραση Ν. Καζαντζάκη
Δημοσίευση	Αθήνα : Εστία, 2000
Ομοιόμορφος τίτλος	Ιλιάς
Συγγραφέας	Όμηρος, 8ος αι. π.Χ.
Μεταφραστής	Καζαντζάκης, Νίκος ; Κακριδής, Ιωάννης

Εγγραφή 2 - Βιβλίο	
Τίτλος / Συγγραφέας	Ιλιάδα / Όμηρος ; έμμετρη μετάφραση Σ. Χρήστου
Δημοσίευση	Αθήναι : Γερ. Σ. Χρήστου, 1955
Ομοιόμορφος τίτλος	Ιλιάς
Συγγραφέας	Όμηρος, 8ος αι. π.Χ.
Μεταφραστής	Καζαντζάκης, Νίκος ; Κακριδής, Ιωάννης

Εγγραφή 3 - Βιβλίο	
Τίτλος / Συγγραφέας	The Iliad of Homer / translated by Lord
Δημοσίευση	London : Dent, 1944
Ομοιόμορφος τίτλος	Ιλιάς
Συγγραφέας	Όμηρος, 8ος αι. π.Χ.
Μεταφραστής	Derby, Lord

Οι τρεις εγγραφές αποτελούν

• **Ένα Έργο:**

Η Ιλιάδα του Ομήρου

• **Δύο Εκφράσεις:**

Η Ελληνική μετάφραση Καζαντζάκη και Κακριδή και η Αγγλική του Derby

• **Τρία Manifestations:**

Δύο που εκδίδονται στην Αθήνα από Εστία και Γερ. Σ. Χρήστου το 2000 και το 1955 αντίστοιχα και ένα που εκδίδεται στο Λονδίνο από το Dent το 1944

Χρήση κλειδιών για την ομαδοποίηση των εγγραφών



Τα 3 τμήματα του κλειδιού για τον εντοπισμό Έργων

1. Τίτλος

- Πολλά διαφορετικά πεδία (19 πεδία στο UNIMARC, χωρίς τα πεδία διασύνδεσης)
 - ◆ Σειρά επιλογής πεδίου τίτλου βάσει αλγορίθμου OCLC
 - Ομοιόμορφος Τίτλος (Κύρια Αναγραφή), (UNIMARC 500, MARC21 130)
 - Ομοιόμορφος Τίτλος (Μη κύρια αναγραφή), (UNIMARC 500 , MARC21 240)
 - Μετάφραση Τίτλου από τον Καταλογογράφο, (UNIMARC 541, MARC21 242)
 - Κύριος Τίτλος, (UNIMARC 200, MARC21 245)
 - Παραλλαγές του Τίτλου, (UNIMARC 517, MARC21 246)
 - Προηγούμενος Τίτλος (UNIMARC 520, MARC21 247)

Τα 3 τμήματα του κλειδιού για τον εντοπισμό Έργων

2. Δημιουργός

- Όπου ακολουθείται η έννοια της κύριας αναγραφής επιλογή του συγκεκριμένου πεδίου
 - ◆ Στο MARC 21 υποχρεωτική
 - ◆ Στο UNIMARC προαιρετική

3. Μορφή

- Απόλυτα κωδικοποιημένη στην ετικέτα εγγραφής
- Γλώσσας και κανόνων καταλογογράφησης εξαρτώμενη στο υποπεδίο 200\$b UNIMARC

Εφαρμογή στην ελληνική πραγματικότητα: ΣΚΕΑΒ & Δείγμα

- Οι συλλογικοί κατάλογοι αποτελούν πολύ καλό πεδίο δοκιμών για FRBRization
 - περιέχουν πολλές *Εκφράσεις* και *Manifestation* ενός Έργου
- Στην Ελλάδα δύο μείζονες συλλογικοί κατάλογοι, οι οποίοι αναπτύσσονται με τελείως διαφορετικό τρόπο
 - Συλλογικός Δημοσίων Βιβλιοθηκών
 - ◆ η εισαγωγή των δεδομένων ελέγχεται από άνθρωπο
 - ◆ υπάρχει πολύ αυξημένη συνέπεια στην καθιέρωση των ονομάτων
 - ◆ υπάρχει σχεδόν παντελής απουσία του πεδίου του ομοιόμορφου τίτλου
 - Συλλογικός Κατάλογος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών (ΣΚΕΑΒ)
 - ◆ παράγεται αυτόματα χωρίς ελέγχους από ανθρώπους σε κάθε εγγραφή
 - ◆ μεγαλύτερο (τόσο από άποψη εγγραφών όσο και από άποψη βιβλιοθηκών) δείγμα της ελληνικής κατάστασης

ΣΚΕΑΒ & συχνότητα ομοιόμορφων τίτλων

Τίτλος / Συγγραφέας	Εγγραφές που αντλήθηκαν από ΣΚΕΑΒ	Εγγραφές με ομοιόμορφο τίτλο	Ποσοστό εγγραφών με ομοιόμορφο τίτλο
Αποκάλυψη / Ιωάννης ο Απόστολος	17	3	17,6 %
Ερωτόκριτος / Κορνάρος	98	7	7,1 %
Ευμενίδες / Αισχύλος	94	30	31,9 %
Ζορμπάς / Καζαντζάκης	33	11	33,3 %
Η Πάππιστα Ιωάννα / Ροΐδης	70	5	7,1 %
Θεσμοφοριάζουσες / Αριστοφάνης	43	21	48,9 %
Ιλιάδα / Όμηρος	40	15	37,5 %
Καπετάν Μιχάλης / Καζαντζάκης	43	17	40 %
Ο ζητιάνος / Καρκαβίτσας	45	1	2,2 %
Οθέλλος / Shakespeare	66	29	43,9 %
Τελευταίος πειρασμός / Καζαντζάκης	28	7	25 %

Η πάππια Ιωάννα

Το όνομα του Ροΐδη εμφανίζεται με 4 παραλλαγές στις 70 εγγραφές του δείγματος ως εξής:

α/α	Όνομα μετά από κανονικοποίηση	Εγγραφές που αντιστοιχούν
1	ΡΟΙΔΗΣ ΕΜΜΑΝΟΥΗΛ 18361904	23
2	ΡΟΙΔΗΣ ΕΜΜΑΝΟΥΗΛ	3
3	ΡΟΙΔΗΣ ΕΜΜΑΝΟΥΗΛ Δ 18361904	17
4	ΡΟ ΙΔΗΣ ΕΜΜΑΝΟΥΗΛ Δ 18361904	27

- Ομοιόμορφος τίτλος υπάρχει σε 5 από τις 70 εγγραφές
 - Εμφανίζεται με 2 παραλλαγές (με ή χωρίς άρθρο)
- Προκύπτουν 15 διαφορετικά υπό-κλειδιά τίτλων
- Ολοκληρωμένο κλειδί (συγγραφέας-τίτλος)
 - Συνολικά 22 κλειδιά στις 70 εγγραφές

Όμηρος

Το όνομα του Ομήρου εμφανίζεται με 6 παραλλαγές στις 40 εγγραφές του δείγματος ως εξής:

α/α	Όνομα Συγγραφέα μετά από κανονικοποίηση	Εγγραφές που αντιστοιχούν
1	HOMER	1
2	ΒΟΜΗΡΟΣ 8ΟΣ ΑΙ ΠΧ	2
3	ΟΜΗΡΟΣ	1
4	ΟΜΗΡΟΣ 8ΟΣ ΑΙ ΠΧ	20
5	ΟΜΗΡΟΣ 8ΟΣ ΑΙΩΝΑΣ ΠΧ	12
6	ΟΜΗΡΟΣ ΠΕΡ 8ΟΣ ΑΙ ΠΧ	4

Ιλιάδα του Ομήρου

- Για το τμήμα του κλειδιού τίτλος
 - Ομοιόμορφος Τίτλος υπάρχει σε 15 από τις 40 εγγραφές
 - ◆ Εμφανίζεται σε 4 παραλλαγές
 - Ιλιάδα
 - Ιλιάς
 - Iliad
 - Ilias
 - Προκύπτουν 23 διαφορετικά υπό-κλειδιά τίτλων
- Ολοκληρωμένο κλειδί (συγγραφέας-τίτλος)
 - Συνολικά 28 κλειδιά στις 40 εγγραφές!
 - ◆ Με απλό ταίριασμα κλειδιών για το ένα και μοναδικό Έργο Ιλιάδα του Ομήρου θα παραχθούν 28 διαφορετικά Έργα από 40 εγγραφές

Αριστοφάνης

Το όνομα του Αριστοφάνη εμφανίζεται με 9 παραλλαγές στις 43 εγγραφές του δείγματος ως εξής:

α/α	Όνομα Συγγραφέα μετά από κανονικοποίηση	Εγγραφές που αντιστοιχούν
1	ARISTOPHANES	6
2	ARISTOPHANES CA 446CA 385 BC	5
3	ΑΡΙΣΤΟΦΑΝΗΣ 438338 ΠΧ	3
4	ΑΡΙΣΤΟΦΑΝΗΣ 444388	4
5	ΑΡΙΣΤΟΦΑΝΗΣ 444388 ΠΧ	12
6	ΑΡΙΣΤΟΦΑΝΗΣ 446385 ΠΧ	9
7	ΑΡΙΣΤΟΦΑΝΗΣ 450380 ΠΧ	1
8	ΑΡΙΣΤΟΦΑΝΗΣ 450385 ΠΧ	1
9	ΑΡΙΣΤΟΦΑΝΗΣ Π 446385 ΠΧ	2

Θεσμοφοριάζουσai

- Για το τμήμα του κλειδιού τίτλος
 - Σε σύνολο 43 εγγραφών οι 21 έχουν ομοιόμορφο τίτλο
 - ◆ Οι παραλλαγές του ομοιόμορφου τίτλου είναι 3
 - Thesmorphoriazusae
 - Thesmorphoriazousai
 - Θεσμοφοριάζουσai
 - Προκύπτουν 6 διαφορετικά υπό-κλειδιά τίτλων
- Ολοκληρωμένο κλειδί (συγγραφέας-τίτλος)
 - Συνολικά 17 κλειδιά στις 43 εγγραφές
 - ◆ Με απλό ταίριασμα κλειδιών για το ένα και μοναδικό Έργο Θεσμοφοριάζουσai του Αριστοφάνη θα παραχθούν 17 διαφορετικά Έργα από 43 εγγραφές

Συγκεντρωτικά κλειδιά Τίτλων-Συγγραφέων

	Πάππιστα Ιωάννα	Ιλιάδα	Θεσμοφοριάζουσες
Διαφορετικά κλειδιά / Εγγραφές δείγματος	22 / 70 Ανά 100 εγγραφές 31,4 κλειδιά	28 / 40 Ανά 100 εγγραφές 70 κλειδιά	17 / 43 Ανά 100 εγγραφές 40 κλειδιά

Σημείωση: ο αριθμός των κλειδιών είναι αντιστρόφως ανάλογος του αριθμού των εγγραφών

- Αυτό συμβαίνει διότι περιορίζεται η “πολυφωνία” των ομοιόμορφων τίτλων και ονομάτων

Συμπεράσματα

- Επιβεβαιώνεται, για άλλη μια φορά, τόσο η ανυπαρξία χρήσης αρχείου καθιερωμένων αποδόσεων, όσο και η αναγκαιότητα χρήσης του
- Η χρήση των ομοιόμορφων τίτλων είναι αποσπασματική και οι τίτλοι κάθε άλλο παρά ομοιόμορφοι είναι
- ❖ Εάν οι καθιερωμένες αποδόσεις ήταν συνεπείς ως προς το όνομά τους
 - Παρά τη σχετικά μικρή παρουσία ομοιόμορφων τίτλων *[στις 153 εγγραφές του δείγματος των τριών Έργων υπάρχουν ομοιόμορφοι τίτλοι σε 41 εγγραφές (ποσοστό 26,7)]* τα ποσοστά επιτυχίας σχεδόν θα διπλασιάζονταν
 - ◆ Τα κλειδιά θα μειώνονταν ως εξής:
 - από 22 σε 12 για την Πάππισα Ιωάννα
 - από 28 σε 20 για την Ιλιάδα
 - από 17 σε 5 για τις Θεσμοφοριάζουσες

...σε συνεχή εξέλιξη...

- μέτα-Composer (<http://dlib.ionio.gr/metacomposer/>)
 - Σύστημα μετα-Αναζήτησης βιβλιογραφικών πηγών που συνθέτει και παρουσιάζει τα αποτελέσματα αναζήτησης στο χρήστη σε Έργο-κεντρικές συστάδες σύμφωνα με το εννοιολογικό μοντέλο *Functional Requirements for Bibliographic Records* (FRBR)
- Key Generator (<http://dlib.ionio.gr/frbrwkey/>)
 - Εφαρμογή λογισμικού για την παραγωγή κλειδιών εντοπισμού Έργων
 - ◆ Δυνατότητα επιλογής διαφόρων διατάξεων και συνόλων χαρακτήρων

Μελλοντική έρευνα

- Πόσο θα βελτιώνε τα αποτελέσματα η χρήση τεχνικών ελέγχου βαθμού ομοιότητας (similarity);
 - Ένα από τα βασικά προβλήματα είναι οι διαφορετικές γλώσσες
- Σε μια προσπάθεια βελτίωσης του αλγορίθμου του OCLC και προσαρμογής του στην ελληνική πραγματικότητα επιχειρήθηκε η προσθήκη στη σειρά επιλογής του αλγορίθμου του πεδίου διασύνδεσης “Μετάφραση τού” (454 UNIMARC)
 - Τα κλειδιά μειώνονται ως εξής:
 - ◆ από 15 σε 12 για την Πάππια Ιωάννα
 - ◆ από 23 σε 22 στην Ιλιάδα
 - ◆ μένουν σταθερά στο 6 αυτά από τις Θεσμοφοριάζουσες
 - Παρότι με μια πρώτη ματιά οι αριθμοί δείχνουν μια σαφή (αλλά όχι θεαματική) βελτίωση των αποτελεσμάτων μένει να διερευνηθεί η ορθότητά της
- Πόσο θα βελτιώνε τα αποτελέσματα η εκ των υστέρων χρήση ενός αρχείου καθιερωμένων αποδόσεων, κατά την παραγωγή των κλειδιών;
 - Τα ποσοστά βελτίωσης θα ήταν άρρηκτα συνδεδεμένα με την ποιότητα του αρχείου καθιερωμένων αποδόσεων