

Εργαστήριο Ψηφιακών Βιβλιοθηκών και Ηλεκτρονικής Δημοσίευσης

Τμήμα Αρχιονομίας – Βιβλιοθηκονομίας

Ιόνιο Πανεπιστήμιο

Προσεγγίσεις
Διαλειτουργικότητας στα
Μεταδεδομένα

Σαράντος Καπιδάκης

sarantos@ionio.gr

Επισκόπηση Παρουσίασης

- Ο ρόλος των μεταδεδομένων
- MARC
- Dublin Core
- Dublin Core Qualified
- Europeana Semantic Elements
- Europeana Data Model

Ο Ρόλος των Μεταδεδομένων

Εμφανίσεις Μεταδεδομένων

- Τα μεταδεδομένα είναι **δομημένη** πληροφορία που περιγράφει, εξηγεί, εντοπίζει ή διευκολύνει την ανάκτηση, τη χρήση ή την διαχείριση ενός πληροφοριακού πόρου.
- Συχνά καλούνται («**δομημένα**») «δεδομένα για άλλα δεδομένα» ή «πληροφορία για άλλη πληροφορία»

Πού Βρίσκονται τα Μεταδεδομένα

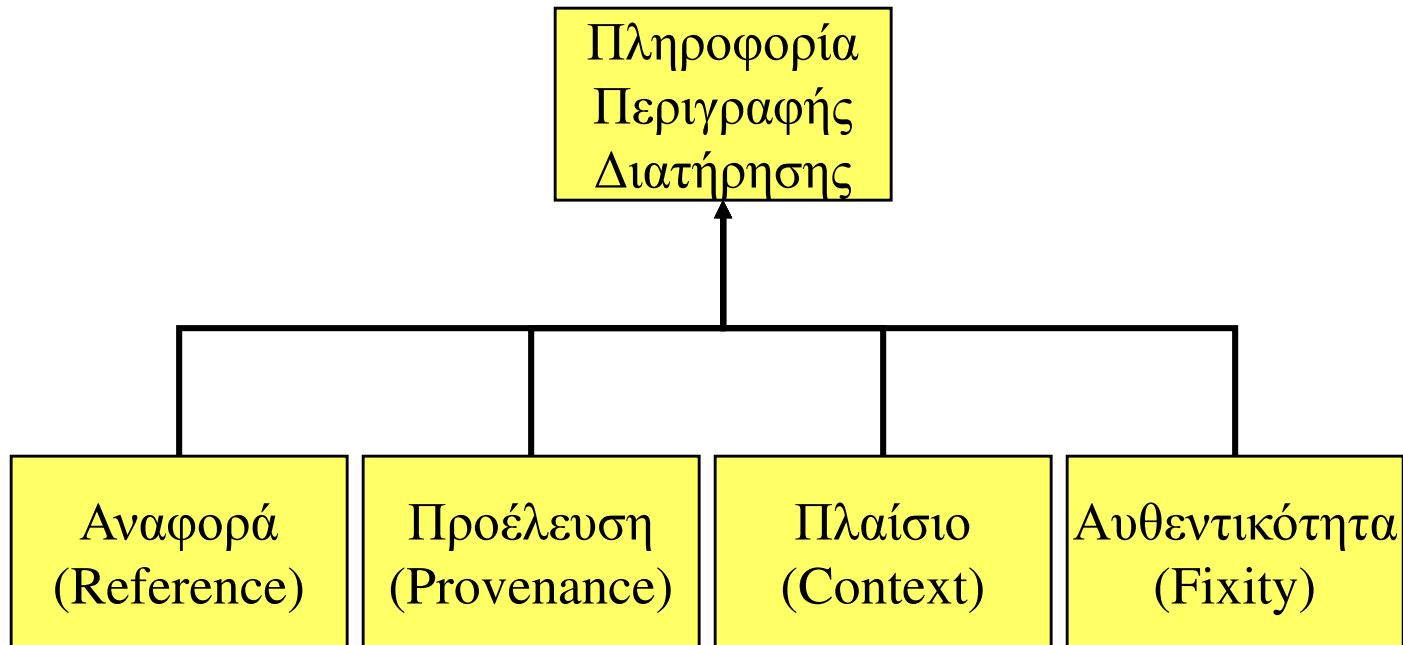
- Στην ετικέτα ενός CD (ορατά)
- Στη σελίδα τίτλου ενός τεκμηρίου (ορατά)
- Στην κορυφή μιας ιστοσελίδας (ορατά)
- Στην ηλεκτρονική μορφή μέσα στον πόρο (ορατά)
- Ενσωματωμένα στο ηλεκτρονικό δημοσίευμα (μέρος της κωδικοποίησης, μη ορατά)
- Σε διαφορετική εγγραφή
- ...

Κατηγορίες Μεταδεδομένων

- Περιγραφικά (descriptive)
 - τίτλος, συγγραφέας, θεματικοί όροι, περιγραφή
- Δομικά (structural)
 - πώς δομούνται σύνθετα αντικείμενα
- Διαχειριστικά (administrative)
 - κάτοχος, ημερομηνία πρόσβασης, κατηγορία αρχείου, διαχείριση δικαιωμάτων
- Το Dublin Core: περιγραφικά μεταδεδομένα

Το Μοντέλο OAIS

Ταξινόμια πακέτων πληροφοριών: Πληροφορία περιγραφής διατήρησης (Preservation Description Information) – Open Archival Information System



Πληροφορία Περιγραφής Διατήρησης

- **Αναφορά:** προσδιορίζει (εσωτερικά ή εξωτερικά) μοναδικά το αντικείμενο
 - ISBN, URN, ψηφιακά προσδιοριστικά: DOI, URL, ...
- **Προέλευση:** τεκμηριώνει την ιστορία του αντικειμένου
 - προέλευση, αλλαγές κηδεμονίας, ενέργειες και αποτελέσματα διατήρησης, στοιχεία αρχειοθέτησης
- **Πλαίσιο:** σχέση του (ψηφιακού) αντικειμένου με το περιβάλλον του
 - Γιατί δημιουργήθηκε, σχέση με άλλα αντικείμενα
- **Αυθεντικότητα:** πιστοποίηση / επικύρωση
 - checksum, ψηφιακές υπογραφές

Εμφανίσεις των Μεταδεδομένων

- Τρεις συνηθισμένες εμφανίσεις των μεταδεδομένων:
 - Συστήματα διαχείρισης βάσεων δεδομένων
 - Σχήματα των σχεσιακών βάσεων δεδομένων
 - Παραδοσιακοί κατάλογοι βιβλιοθήκης
 - Παγκόσμιος Ιστός

Παράδειγμα Διαλειτουργικότητας με Μεταδεδομένα

- Σύμφωνα με την ελάχιστη κοινή δομή, π.χ.
 - Συλλογή με **Βιβλία**
 - TITLE, AUTHOR, PUBLISHER, ABSTRACT
 - Συλλογή με **Διατριβές**
 - TITLE, AUTHOR, SUMMARY
 - Συλλογή με **Τραγούδια**
 - TITLE, CREATOR, VERSE
- Κοινά ανά 2; Ανά 3;
 - ABSTRACT = SUMMARY;
 - AUTHOR = CREATOR;

Ανταλλαγή (Μετά)? Δεδομένων

Τα δεδομένα (και τα μεταδεδομένα) τα

- Χρησιμοποιούμε (εσωτερικά στην εφαρμογή)
- Αποθηκεύουμε (σε αρχεία)
- Μετατρέπουμε, σε άλλες μορφές [εύκολα;]
- **Μεταφέρουμε – Δίνουμε – Ανταλλάσσουμε**
 - Με αποθήκευση σε αρχεία (και φύλαξη για μετά)
 - Μέσω πρωτοκόλλων (μεταξύ εφαρμογών)
- Η διαλειτουργικότητα μεταδεδομένων απαιτεί
 - Χρήση κανόνων περιεχομένου και προτύπων
 - Σαφήνεια για τους πόρους που περιγράφονται

Ανταλλαγή και Πρότυπα

Κάθε ανταλλαγή χρειάζεται πρότυπα:

- Για εικόνα, ήχο, βίντεο, κλπ
- Για γραφήματα, πίνακες
- Για μορφοποιημένα κείμενα (HTML, PDF, ...)
- Για κωδικοποίηση γραμμάτων / γραμματοσειρές
- Για βάσεις δεδομένων
- Για ονοματολογία – προσδιορισμό πόρων
- Για **μεταδεδομένα**
- Για πρωτόκολλα επικοινωνίας (Z39.50, Dienst,...)
- ...

Πρότυπα και Συστάσεις

Για να είναι τα μεταδεδομένα χρήσιμα και επωφελή είναι ουσιαστικό

- Η Σύνταξη, Σημασιολογία και Δομή να **συμμορφώνονται** με τα πρότυπα
- Να απεικονίζουν την **ουσία** των πηγών (την πληροφορία που έχουν να δώσουν)

Πρώιμη έκφραση των πρότυπων, για χρήση και κριτική, είναι οι **συστάσεις** – RFC (Request For Comment), που χρησιμοποιούνται παρόμοια

Αναγκαιότητα Προτύπων Μεταδεδομένων

- Χρειάζονται για να έχουμε κοινή αντίληψη των δεδομένων που περιγράφονται με αυτά
- Μας προσφέρουν μεγαλύτερη **δομή**
- \Rightarrow Μας περιορίζουν στην **ευελιξία**
- \Rightarrow Επεξεργάζονται ευκολότερα **μηχανικά**
- Είναι αναγκαιότερα σε **ψηφιακά αντικείμενα**
 - Αφού στα συμβατικά έχουμε αισθητήρια αντίληψη για αναγνώριση, θέση, θέματος, ...

Πόσα Πρότυπα Μεταδεδομένων

- Χρειάζονται **κοινά** κατανεμημένα μοντέλα μεταδεδομένων
- Δεν υπάρχει **ένα** μοναδικό διεθνές πρότυπο για μεταδεδομένα, γιατί:
 - Χρειαζόμαστε διαφορετικά **επίπεδα πολυπλοκότητας**, από πλούσιες μέχρι απλές περιγραφές
 - **Υπάρχουν** κάμποσα σχήματα μεταδεδομένων, για διαφορετικά επίπεδα και απαιτήσεις
 - **Επεκτείνουμε** τα υπάρχοντα πρότυπα
- Πότε πρέπει να αναζητούμε **άλλο** πρότυπο – ή σύσταση;

Διαλειτουργικότητα με Μεταδεδομένα

Κύριες προσεγγίσεις διαλειτουργικότητας

- Αναζήτηση σε διαφορετικά συστήματα
 - Δεν έχουν τα ίδια μεταδεδομένα, αλλά αντιστοιχούν τις δυνατότητες αναζήτησης σε ένα κοινό σύνολο
 - Π.χ. Z39.50
- Συγκομιδή μεταδεδομένων
 - Μεταφράζουν τα δικά τους μεταδεδομένα σε ένα κοινό βασικό σύνολο στοιχείων, που διαθέτουν για συγκομιδή
 - Π.χ. OAI

Ο Μύθος της Κατηγοριοποίησης

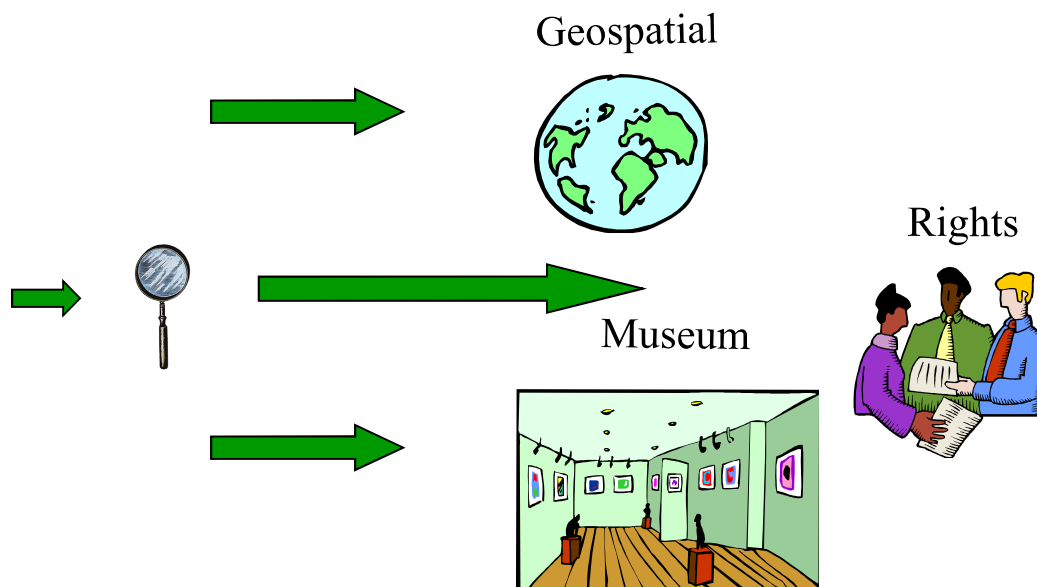
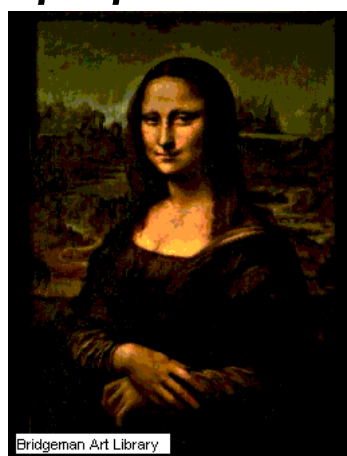
...there is no classification of the universe that is not fictional and conjectural.

...δεν υπάρχει κατηγοριοποίηση του σύμπαντος που να μην είναι φανταστική και αυθαίρετη

– Jorge Luis Borges

Φακοί και Όψεις

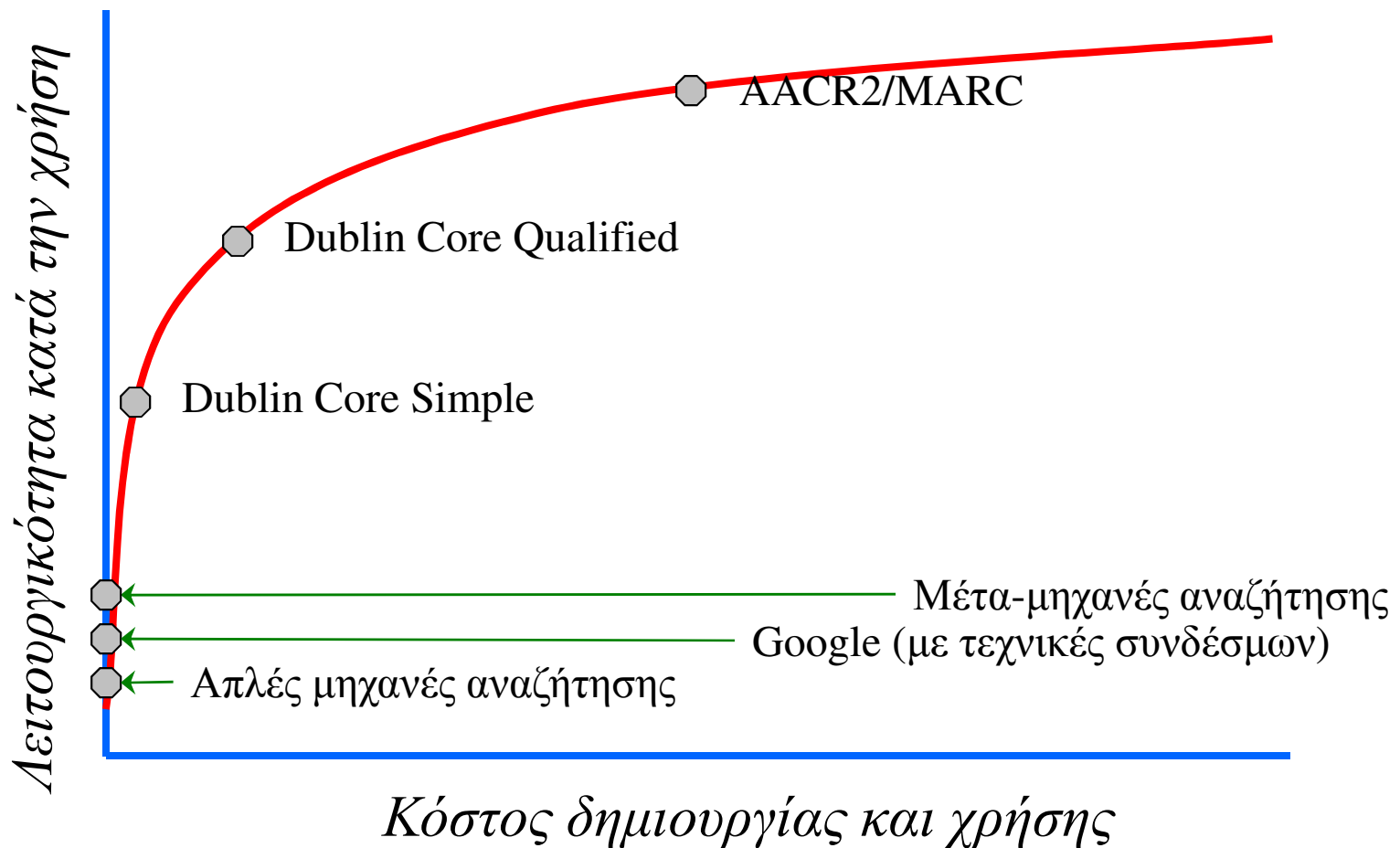
- Όλες οι κατηγοριοποιήσεις παρέχουν και **πρέπει** να παρέχουν μια μεροληπτική όψη της πραγματικότητας
- Κάθε όψη τονίζει συγκεκριμένα χαρακτηριστικά και κρύβει άλλα



Προκλήσεις και Τάσεις στα Μεταδεδομένα

- «Απλού και ευρύ» και «πλούσιου αλλά συγκεκριμένου» (απλότητα \leftrightarrow λειτουργικότητα)
- «Γενικού» και «προσαρμοσμένου» (διαλειτουργικότητα \leftrightarrow επεκτασιμότητα)
- «Αναγνώσιμου από ανθρώπους» και «επεξεργάσιμου από μηχανές» (σε δημιουργία και χρήση)
- «Κατανομή» και «Συγκέντρωση» σε ένα μέρος
- Αυτόματη παραγωγή ή με ανθρώπινη παρέμβαση

Κόστος και Λειτουργικότητα Αναζήτησης και Μεταδεδομένων



MARC

Πλαίσια Μεταδεδομένων

- ... Μοντελοποίηση περίπλοκων αντικειμένων, και του τρόπου συμπεριφοράς τους
- Δεν έχουμε **ήδη** καταλήξει σε μεταδεδομένα;

MARC STANDARDS

*Library of Congress
Network Development and MARC Standards Office*

- Η οικογένεια των προτύπων MARC είναι το πιο επιτυχές πρότυπο περιγραφής πόρων στον κόσμο

Τι Δεν Είναι Καλό στο Μοντέλο του MARC

- Ακριβό
 - Περίπλοκο (ακόμα και για τον αρχικό του στόχο;)
 - Επαγγελματική μεσολάβηση (υποθέτει ότι μία μόνο επαγγελματική κοινότητα έχει την απαιτούμενη γνώση)
- Μονολιθικό
 - Ένα μέγεθος ταιριάζει για όλα
 - Απεικονίζει την κεντριοποιημένη του προέλευση
- Προκαταλαμβάνει προς φυσικά τεχνουργήματα
 - Περιγράφει προκαθορισμένους πόρους
 - Ελλιπής κάλυψη της εξέλιξης πόρων και άλλων σχέσεων των πόρων
- Άγγλο-κεντρικό

Χαρακτηριστικά του MARC

- Πολύ περίπλοκο, απαιτεί υψηλά εκπαιδευμένο προσωπικό και εξειδικευμένα συστήματα εισαγωγής
- Πολύ επικεντρωμένο στο έντυπο υλικό, δύσκολα επεκτάσιμο σε ψηφιακές συλλογές
- Παρά το μέγεθος και την περιπλοκότητά του, είναι πρότυπο ενός μόνο αντικειμένου, τα μεταδεδομένα συνδέονται με μόνο ένα αντικείμενο ανά εγγραφή

Η Μέθοδος των MARC

- 10/1000 κατηγορίες/πεδία
- Με υποπεδία
- Με σαφή μονοσήμαντη σημασία
- Με υποχρεωτικά και προαιρετικά πεδία,
- Με πεδία (αδόμητων) σημειώσεων
- Επεκτείνονται πεπερασμένα για νέα πληροφορία
 - Λειτουργούν σαν ένα σύνολο γραμματοκιβωτίων
- Επαρκές για αυτόνομες συλλογές
- Συστήνουν ανταλλαγή τους με ISO2709

Μειονεκτήματα του MARC

- Μη σχεσιακό, δόμηση με περιορισμούς
- Δύσκολοι χαρακτηρισμοί συχνά καταγράφονται σε λάθος θέση
- Επεκτείνεται, αλλά τοπικά, περιορισμένα και αδόμητα
- Χρειάζεται αλλαγές στα δεδομένα για να παρακολουθεί την εξέλιξη
 - Π.χ. CDROM -> μεγάλο ή μικρό!
- Κοστίζει \$50 / πλήρη εγγραφή

Dublin Core

Το Πρότυπο «Dublin Core»

- Πρωτοβουλία για να βελτιώσει την ανακάλυψη πόρων στο Διαδίκτυο
- Κοινός παρονομαστής για επικοινωνία και διαλειτουργικότητα
- Έναυσμα για σύγκλιση των προτύπων

ANSI/NISO Z39.85-2001

ISSN: 1041-5653

The Dublin Core Metadata Element Set

Abstract: Defines fifteen metadata elements for resource description in a cross-disciplinary information environment.

An American National Standard
Developed by the
National Information Standards Organization

Approved September 10, 2001
by the
American National Standards Institute

Published by the National Information Standards Organization
Bethesda, Maryland



NISO Press, Bethesda, Maryland, U.S.A.

Χαρακτηριστικά του «Dublin Core»

- «Dublin Core» σημαίνει «Dublin, Ohio» / OCLC
- Πρωτοβουλία για να βελτιώσει την ανακάλυψη πόρων στο Διαδίκτυο
- Κοινός παρονομαστής για επικοινωνία / διαλειτουργικότητα
- Έναυσμα για σύγκλιση των προτύπων
- Επεκτάσιμο, για να καλύψει τις επιπρόσθετες ανάγκες ανακάλυψης πόρων των διαφορετικών εφαρμογών / περιοχών

Στόχοι του Dublin Core

- Απλότητα δημιουργίας και διατήρησης
 - Μη ειδικοί να δημιουργούν περιγραφικές εγγραφές για αποτελεσματική ανάκτηση σε δικτυωμένο περιβάλλον
- Κοινά κατανοητή (διαθεματική) σημασιολογία
 - Σύγκλιση κοινών, περισσότερο γενικών στοιχείων
 - Αυξημένη ορατότητα και προσβασιμότητα
 - Κατάλληλο και για τον μη ειδικό της αναζητητής
 - Τον «ψηφιακό τουρίστα»

Το Πρότυπο Dublin Core

- Προτυποποίηση
 - ISO 15836-2003
 - US: NISO Z39.85-2001
 - Ευρώπη: αναγνώριση από το CEN/ISSS Workshop Agreement 13874-2000
- Συστήνεται ανταλλαγή με RDF/XML

Χρήση του Dublin Core

- Είναι βασικός πυρήνας στοιχείων
- *Δεν* είναι υποκατάστατο σε πλουσιότερα περιγραφικά πρότυπα
- Παρέχει 15 «παράθυρα» από πλουσιότερη περιγραφή πόρων
 - Φανερώνει πλούσιες περιγραφές σε απλή μορφή
 - Σημασιολογικά σταυροδρόμια, αντιστοιχίσεις σε υπάρχοντα δεδομένα

Η Αποστολή του DCMΙ

- Να διευκολύνει την εύρεση πόρων από το Διαδίκτυο, μέσω των ακόλουθων δραστηριοτήτων:
 - Ανάπτυξη προτύπων μεταδεδομένων για ανακάλυψη μεταξύ περιοχών (π.χ. Dublin Core)
 - Ορισμό πλαισίων για διαλειτουργικότητα των συνόλων μεταδεδομένων
 - Διευκόλυνση της ανάπτυξης ειδικευμένων συνόλου μεταδεδομένων για μια κοινωνία ή κλάδο
 - Που είναι συνεπή με τα παραπάνω

Τα Στοιχεία Μεταδεδομένων του Dublin Core

- Διεπιστημονική ομοφωνία σε απλά σύνολα στοιχείων για ανακάλυψη πόρων
 - 15 στοιχεία (πεδία μεταδεδομένων)
 - όλα προαιρετικά
 - όλα επαναλαμβανόμενα
- Δεν προορίζεται για περιγραφή περίπλοκων πόρων
 - Η αρχική ιδέα των «απλών αντικειμένων – σαν τεκμήρια»
 - Απλότητα στη σημασιολογία, ευκολία χρήσης
- Παρέχει βασική «σημασιολογική διαλειτουργικότητα»
 - Μεταξύ επιστημονικών περιοχών, μεταξύ γλωσσών
 - Δεν παρέχει λεπτομερείς κανόνες καταλογογράφησης
- Επιτρέπει επεκτασιμότητα – σε άλλες κατηγορίες πόρων
- Ένα σύνολο από 15 ευρύχωρα «καλάθια» ...

Dublin Core: Τα 15 Στοιχεία

- Τίτλος / **Title**
- Θέμα / **Subject**
- Περιγραφή / **Description**
- Πηγή / **Source**
- Γλώσσα / **Language**
- Σχέση / **Relation**
- Κάλυψη / **Coverage**
- Δημιουργός / **Creator**
- Εκδότης / **Publisher**
- Συντελεστής / **Contributor**
- Δικαιώματα / **Rights**
- Ημερομηνία / **Date**
- Τύπος / **Type**
- Μορφότυπο / **Format**
- Αναγνωριστικό / **Identifier**

Εισαγωγή στο Dublin Core

- Η διαλειτουργικότητα απαιτεί
 - Χρήση κανόνων περιεχομένου και προτύπων
 - Σαφήνεια για τους πόρους που περιγράφονται
 - Π.χ. εργασία, έκφραση, παρουσίαση, επιμέρους θέματα
- Οι πραγματικοί πόροι είναι πιο περίπλοκοι από τα (μόνιμα) αντικείμενα τύπου τεκμηρίων
 - Τα χαρακτηριστικά των πόρων αλλάζουν με το χρόνο
 - Οι ενέργειες των πρακτόρων προκαλούν αλλαγές

Dublin Core – Simple

- Το Dublin Core έχει 15 στοιχεία. Κάθε ένα από αυτά είναι προαιρετικό και επαναλαμβανόμενο
- Τα 15 στοιχεία χωρίζονται σε 3 κατηγορίες:
 - Περιεχόμενο:
 - Περιγράφουν το αντικείμενο
 - Πνευματική Ιδιοκτησία:
 - Περιγράφουν το copyright και τη δημιουργία
 - Στιγμιότυπο:
 - Περιγράφουν την εισαγωγή και διαχείριση

DC – Περιεχόμενο

- Τίτλος / **Title** – (ονομασία πηγής)
- Θέμα / **Subject**, π.χ. λέξεις-κλειδιά, ταξινομικοί κωδικοί
- Περιγραφή / **Description**
 - Π.χ. περίληψη, περιεχόμενα, περιγραφή εικόνας
- Πηγή (ή «Προέλευση») / **Source** – (παραγωγής)
- Γλώσσα / **Language** – (του περιεχομένου)
- Σχέση / **Relation** – (αναφορά σε σχετική πηγή)
 - Π.χ. έκδοση του ...
- Κάλυψη / **Coverage** – (γεωγραφική ή χρονική)

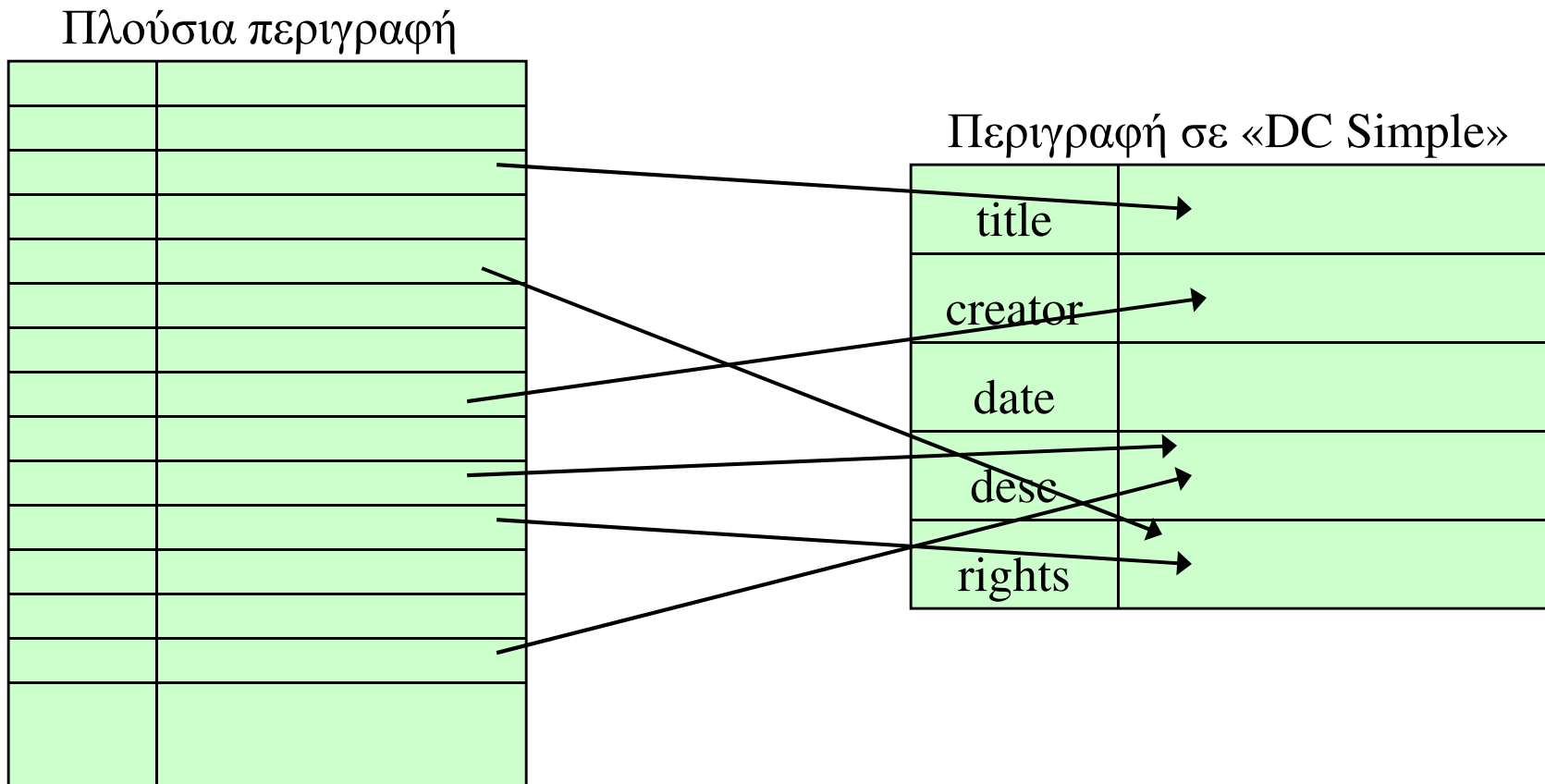
DC – Πνευματική Ιδιοκτησία

- Δημιουργός / **Creator** – (πρόσωπο, οργανισμός, υπηρεσία)
- Εκδότης / **Publisher** – (πρόσωπο, οργανισμός, υπηρεσία)
- Συντελεστής (ή «Συνεργάτης» ή «Υπεύθυνος συμβολής») / **Contributor** – (πρόσωπο, οργανισμός, υπηρεσία που συμβάλλει στο περιεχόμενο)
 - Π.χ. μεταφραστής, εικονογράφος, κριτής
- Δικαιώματα / **Rights** – (κείμενο σχετικά με την πνευματική ιδιοκτησία)

DC – Στιγμιότυπο

- Ημερομηνία / **Date**
 - Π.χ. δημιουργίας, έκδοσης, μετάφρασης, πρόσκτησης, καταλογογράφησης, ...
- Τύπος / **Type** – (κατηγορία, σχετικά με το περιεχόμενο)
 - Π.χ. ποίημα, λεξικό, software, home-page
- Μορφότυπο / **Format** – (φυσική ή ψηφιακή μορφή)
 - Π.χ. Macintosh-software, pdf, html, διαστάσεις, διάρκεια
- Αναγνωριστικό (ή Προσδιοριστής ή Κωδικός Ταύτισης) / **Identifier**
 - Μοναδικό προσδιοριστικό, π.χ. URL, ISBN, ...

Περιγραφή σε Dublin Core



Το Dublin Core είναι περισσότερο αντικείμενο
εξαγωγής – παρά δημιουργίας

Συναίνεση και Dublin Core

- Δημιουργήθηκε σαν συναίνεση από μία διεθνή, δια-θεματική ομάδα που αντιπροσωπεύει
 - Βιβλιοθήκες και Πληροφόρηση
 - Πληροφορική
 - Κωδικοποίηση κειμένου (TEI)
 - Μουσεία
 - Σχετικούς επιστημονικούς τομείς
- Τα στοιχεία του αναπαριστούν μια ευρεία δια-θεματική συναίνεση
- Διεθνείς συναντήσεις, ημερίδες, κλπ για ανάπτυξη συναίνεσης

Παρανοήσεις για Dublin Core

- *Το Dublin Core είναι άλλο ένα πρότυπο περιγραφής ή ανταλλαγής μεταδεδομένων, και πρέπει να επιλέξουμε αν θα το προτιμήσουμε;*
- Το Dublin Core είναι η **σημασιολογική τομή** των (κυριότερων) **προτύπων** (που χρησιμοποιούνται σήμερα)
- Μπορεί να χρησιμοποιηθεί ως βάση στοιχείων σε «**προφίλ εφαρμογών**», για ευκολότερη (προσχεδιασμένη) διαλειτουργικότητα

Παρανοήσεις για Dublin Core

- Το Dublin Core δεν είναι μορφότυπο **ανταλλαγής** μεταδεδομένων (αποθηκευμένα π.χ. σε MARC)
- Μπορούμε να απεικονίσουμε («μετατρέψουμε») πολλά πρότυπα σε Dublin Core (**export DC**)
- Δεν μας ενδιαφέρει να «μετατρέψουμε» από Dublin Core πίσω σε άλλα πρότυπα (**import DC**)
- **Δεν καταργούμε** τις αρχικές (λεπτομερείς) εγγραφές – δημιουργούμε **καινούργιες** (συνοπτικές) για άμεση χρήση
- Το Dublin Core είναι σχεδιασμένο για εύκολη / **προφανή** απεικόνιση (άλλων προτύπων) μεταδεδομένων σε αυτό
- Το Dublin Core έχει περισσότερο σχέση με τη **διαλειτουργικότητα** παρά με τα μεταδεδομένα

Dublin Core Qualified

Dublin Core – Qualified

Ως επέκταση του απλού (simple ή unqualified)

Dublin Core έχουμε το εξειδικευτικό (qualified)

Dublin Core, που προσφέρει:

Βελτίωση της σημασιολογικής ακρίβειας του Dublin Core ορίζοντας τους **εξειδικευτές** (qualifiers)

- Εξειδίκευση του στοιχείου (element refinement)
 - Για περισσότερη λεπτομέρεια στην περιγραφή
- Σύστημα κωδικοποίησης (encoding scheme)
 - Σαν χρήση απλών κανόνων καταλογογράφησης

Ποικιλίες Εξειδικευτών:

Εξειδίκευση Στοιχείων

- Προσφέρει μεγαλύτερη λεπτομέρεια σε αυτόν που τη χρειάζεται
- Κάθε εξειδικευτής, κάνει την σημασία ενός στοιχείου στενότερη ή πιο ειδική
 - «*Date Created*» και «*Date Modified*»
 - «*IsReplacedBy Relation*» και «*Replaces Relation*»
 - Δεν την τροποποιεί ή επεκτείνει, σε καμία περίπτωση
- Αν το λογισμικό δεν καταλαβαίνει κάποιο εξειδικευτή, μπορεί με ασφάλεια να τον αγνοήσει!

Εξειδίκευση σε Dublin Core

- Δομημένο αλλά μη περιορισμένο σύνολο!
 - Κάθε στοιχείο μπορεί να έχει και άτυπες εξειδικεύσεις
- Έσχατη λύση: νέοι προσδιορισμοί / επίπεδα
 - π.χ. Degree = τύπος πτυχίου
- Κανόνας «ακρωτηριασμού» (γενίκευσης), π.χ:
 - TITLE
 - TITLE.ALTERNATIVE
 - TITLE.TRANSLATED
 - TITLE.ALTERNATIVE.TRANSLATED
 - TITLE.FULL

Dublin Core: μη Εξειδίκευση Στοιχείων

- Creator
- Subject
- Publisher
- Contributor
- Type
- Identifier
- Source
- Language
- Rights

Dublin Core: Εξειδίκευση Στοιχείων

Στοιχείο	Εξειδίκευση
Title	Alternative
Description	Table Of Contents Abstract
Date	Created Valid Available Issued Modified
Format	Extent Medium
Coverage	Spatial Temporal

Dublin Core: Εξειδίκευση Στοιχείων

Στοιχείο	Εξειδίκευση
Relation	Is Version Of Has Version Is Replaced By Replaces Is Required By Requires Is Part Of Has Part Is Referenced By References Is Format Of Has Format

Ποικιλίες Εξειδικευτών: Σύστημα Κωδικοποίησης Τιμών

- Δηλώνει ότι μια τιμή είναι
 - Ένας όρος από ελεγχόμενο λεξικό (π.χ., *Library of Congress Subject Headings*)
 - Χαρακτήρες μορφοποιημένοι με συγκεκριμένο τρόπο (π.χ., «2003-05-02» σημαίνει «2 Μαΐου», όχι «5 Φεβρουαρίου»)
- Ακόμα και αν ένα σχήμα δεν είναι γνωστό από το λογισμικό, η τιμή πρέπει να είναι «κατάλληλη» και χρησιμοποιήσιμη για ανακάλυψη πόρων

Dublin Core: μη Κωδικοποίηση Στοιχείων

- Title
- Creator
- Description
- Publisher
- Contributor
- Rights

Dublin Core: Κωδικοποίηση Στοιχείων

Στοιχείο	Σχήμα κωδικοποίησης
Date	W3C-DTF DCMI
Type	DCMI Type Vocabulary
Format	IMT
Identifier	URI
Source	URI
Language	ISO 639-2 RFC 1766
Relation	URI

Dublin Core: Κωδικοποίηση Στοιχείων

Στοιχείο	Σχήμα κωδικοποίησης
Subject	LCSH MeSH DDC UDC LCC
Coverage	DCMI Point ISO 3166 DCMI Box DCMI Period W3C-DTF

Κωδικοποίησης Τιμών του «Type»

- Text (βιβλία, άρθρα, email, fax, ...)
- Image (εικόνες, κινούμενες εικόνες, διαγράμματα, ...)
- Sound (μουσικό CD, αρχείο μουσικής, ομιλία, ...)
- Interactive Resource (υπηρεσίες συζήτησης, ...)
- Service (τραπεζικές υπηρεσίες, υπηρεσία Z3950, ...)
- Software (λογισμικό υπολογιστή)
- Dataset (πίνακες, βάσεις δεδομένων, ...)
- Event (συνάντηση, έκθεση, συνέδριο, δίκη, γιορτή, ...)
- Collection (συλλογή τεκμηρίων)

Europeana Rhine

Europeana Semantic Elements

Europeana Semantic Elements

- ESE v3.2.1
- Στηρίζεται στο DC και DC terms
 - <http://dublincore.org/documents/dcmi-terms/>
- Προσθέτει μερικά επιπλέον στοιχεία Europeana
- Κάθε παροχέας απεικονίζει τα τοπικά μεταδεδομένα στο ESE
- Και με συγκομιδή συλλέγονται στην Europeana

Europeana Semantic Elements

v3.2.1

- Περιλαμβάνει στοιχεία από τα
- Dublin Core και Dublin Core terms
 - <http://purl.org/dc/terms/><στοιχείο>
- Europeana
 - <http://www.europeana.eu/schemas/ese/><στοιχείο>

Στοιχεία DC στο ESE v3.2.1

- Creator
- Subject
- Publisher
- Contributor
- Type
- Identifier
- Source
- Language
- Rights

Εξειδικευτές DC στο ESE v3.2.1

Element	Element Refinement(s)
Title	alternative
Description	tableOfContents
Date	created; issued
format	extent; medium
Relation	isVersionOf; hasVersion; isReplacedBy; replaces; isRequiredBy; requires; isPartOf; hasPart; isReferencedBy; references; isFormatOf; hasFormat; conformsTo
Coverage	spatial; temporal

Στοιχεία DC.terms και Europeana

- provenance
 - unstored
 - object
 - provider
 - type
 - isShownBy; isShownAt
 - country
 - hasObject
 - language
 - userTag
 - uri
 - year

Europeana Danube

Europeana Data Model

Europeana Data Model

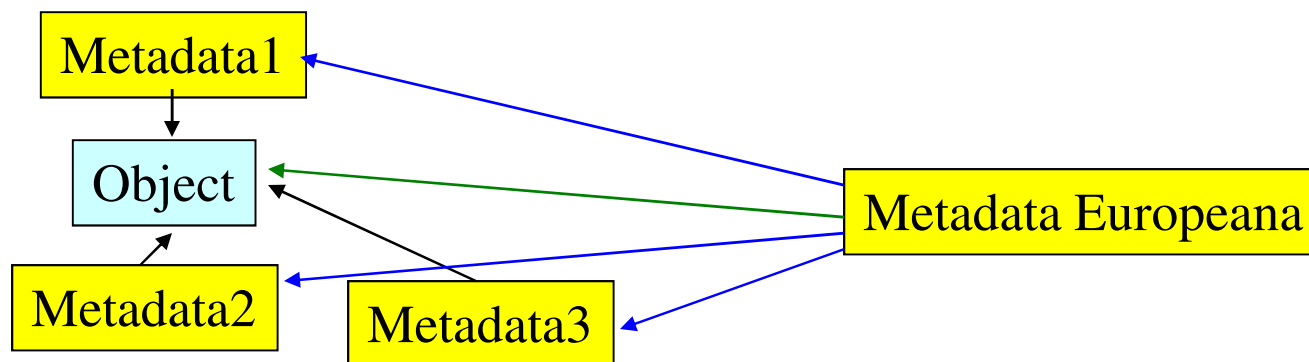
- Ετοιμάζεται η επόμενη έκδοση της Europeana, Danube
- Έχει επιπλέον υπηρεσίες και λειτουργικότητα
- Πιο λεπτομερή μεταδεδομένα
- ...

Σχήμα Μεταδεδομένων στο Europeana Danube

- Βασίζεται στο CIDOC
 - Δεν είναι ζευγάρια τιμών/σχέσεων (όπως το DC) αλλά αποτελείται από σχέσεις / ιδιότητες
 - Π.χ. ένα Αντικείμενο [έχει:]
 - Δημιουργία [με ιδιότητες:]
 - Δράστης
 - Τόπος
 - Ημερομηνία
 - Δημοσίευση [με ιδιότητες:]
 - Ημερομηνία ...

Αρχιτεκτονική Μεταδεδομένων

- Κάθε ψηφιακό αντικείμενο έχει συνήθως μια εγγραφή μεταδεδομένων.
- Στη Europeana θα υπάρχουν πολλές εγγραφές μεταδεδομένων για κάθε αντικείμενο
 - Για κάθε διαφορετική «όψη» για αυτό



Συμπεράσματα

- Χρήση προτύπων και καθιερωμένων κανόνων
- Όχι αυτοσχεδιασμοί στην περιγραφή των δεδομένων
- Έστω και να φαινομενικά επαρκούν για την εφαρμογή
- Δεν φτάνει να δουλεύει μια εφαρμογή
- Πρέπει να μπορεί να **δια-**λειτουργεί στο μέλλον

ΤΕΛΟΣ – ΕΡΩΤΗΣΕΙΣ;

Σαράντος Καπιδάκης

sarantos@ionio.gr

Εργαστήριο Ψηφιακών Βιβλιοθηκών και
Ηλεκτρονικής Δημοσίευσης

<http://dlib.ionio.gr>

Τμήμα Αρχιονομίας – Βιβλιοθηκονομίας
Ιόνιο Πανεπιστήμιο

<http://www.ionio.gr>