

ΔΙΑΧΕΙΡΙΣΗ, ΕΠΙΜΕΛΕΙΑ ΚΑΙ ΕΚΔΟΣΗ ΨΗΦΙΑΚΩΝ ΠΟΡΩΝ

Βίκη Ντρίτσου

ΠΜΣ Ψηφιακές Μέθοδοι στις Ανθρωπιστικές Επιστήμες

Οικονομικό Πανεπιστήμιο Αθηνών

2024-2025

ΠΕΡΙΕΧΟΜΕΝΑ

Dublin Core



Προφίλ Εφαρμογών



XML



6. Dublin Core

Dublin Core

- Πρότυπο ISO 15836, ANSI/NISO Z39.85, και IETF RFC 5013
- Διατυπώθηκε αρχικά στο Δουβλίνο του Οχάιο (Dublin)
- Αποτελείται από έναν πυρήνα (core) βασικών στοιχείων μεταδεδομένων
 - Η σημασιολογία τους διατυπώνεται από διεπιστημονική ομάδα ειδικών στη βιβλιοθηκονομία, την πληροφορική, την κωδικοποίηση κειμένου, τα μουσεία και άλλα
- Πρωταρχικός σκοπός η βελτίωση της ανακάλυψης ψηφιακών πόρων στο διαδίκτυο

Στοιχεία Dublin Core

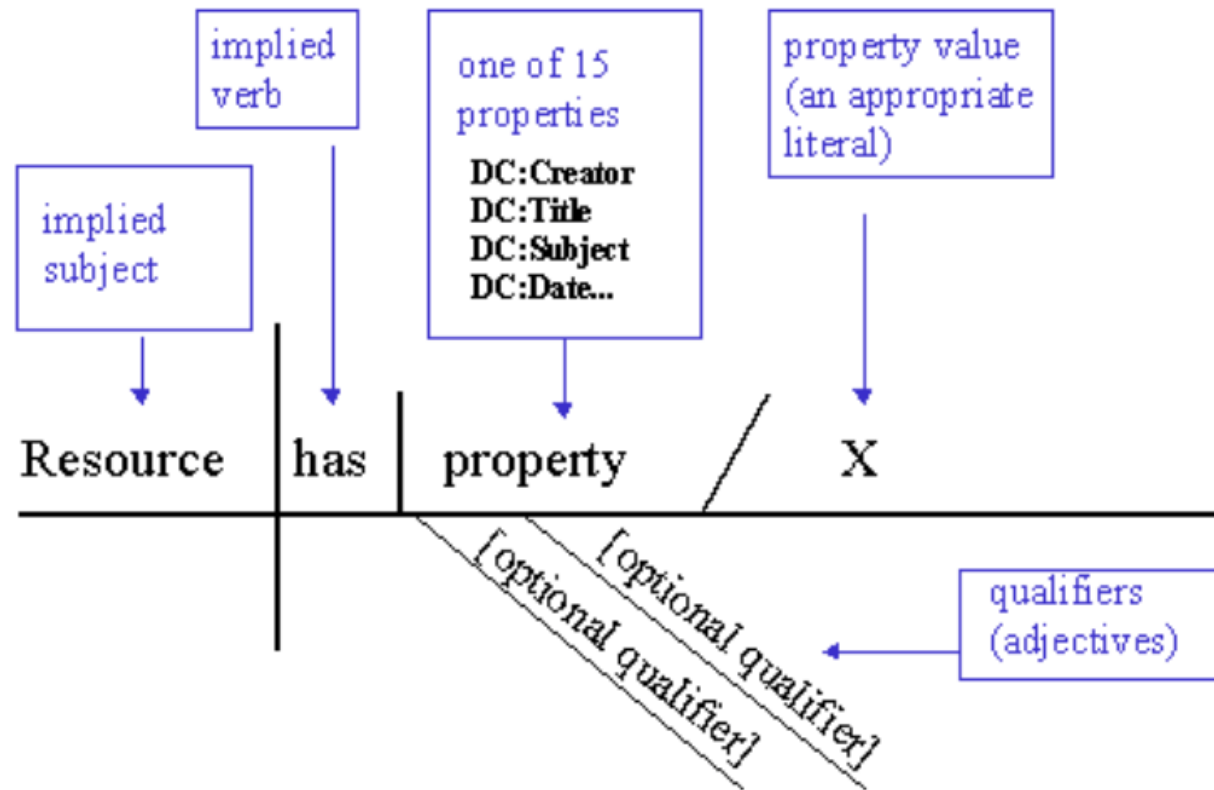
- 15 στοιχεία μεταδεδομένων
 - Προαιρετικά
 - Επαναλαμβανόμενα

1. Contributor
2. Coverage
3. Creator
4. Date
5. Description

6. Format
7. Identifier
8. Language
9. Publisher
10. Relation

11. Rights
12. Source
13. Subject
14. Title
15. Type

Διαγραμματικές δηλώσεις Dublin Core



Κωδικοποίηση Dublin Core

- Στην κεφαλή ενός HTML εγγράφου:

```
<head>
```

```
<meta name="DC.Format" content="video/mpeg; 32 minutes" />
```

```
<meta name="DC.Language" content="en" />
```

```
<meta name="DC.Creator" content="Vicky Dritsou" />
```

```
<meta name="DC.Title" content="My first video" />
```

```
</head>
```

- HTML: γλώσσα σήμανσης (markup), σκοπός της η [απεικόνιση](#) πόρων
- XML: γλώσσα σήμανσης (markup), σκοπός της η [περιγραφή](#) των πόρων

Σύνοψη Dublin Core

- Βασικός πυρήνας 15 στοιχείων μεταδεδομένων
 - Ορισμένα διεπιστημονικώς
 - Προαιρετικά και επαναλαμβανόμενα
 - Αρκούν;
- Πληροφορία διατυπωμένη με "απλά" στοιχεία
 - Ευκολία χρήσης (και από μη εξειδικευμένους χρήστες)
 - Παρέχει διαλειτουργικότητα (συντακτική και σημασιολογική)
 - Επεκτάσιμο

Δεν ενδείκνυται για την αποτύπωση στοιχείων μεγάλης πολυπλοκότητας!

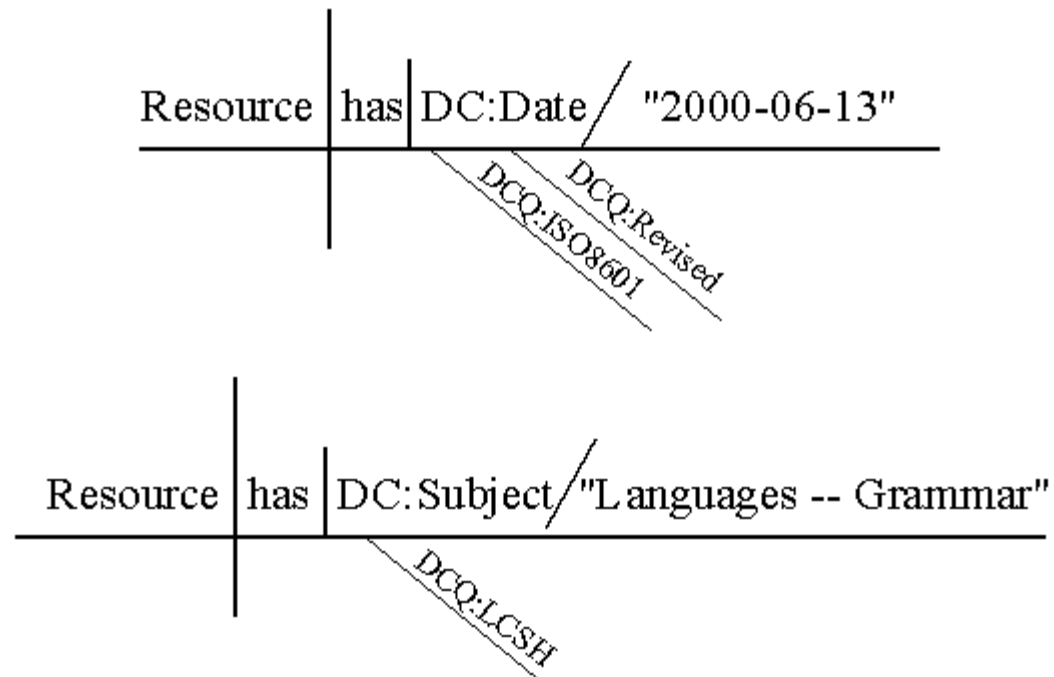
(όλη η πληροφορία ανάγεται, "απλοποιείται" στα 15 στοιχεία)

Επέκταση DC: Qualified Dublin Core

- Εξειδικευμένο Dublin Core: επέκταση/εξειδίκευση του αρχικού DC
- Επέκταση των 15 αρχικών στοιχείων
 - Audience, Provenance, Rights Holder, Instructional Method, Accrual Method, Accrual Periodicity, Accrual Policy
- Ακριβέστερο σημασιολογικά με χρήση των εξειδικευτών (qualifiers)
 - Εξειδίκευση στοιχείου (element refinement): στενότερη, πιο εξειδικευμένη σημασιολογία του στοιχείου
 - Σχήμα κωδικοποίησης (encoding scheme): προσδιορισμός σχήματος (ελεγχόμενα λεξιλόγια, τυπικοί συμβολισμοί, κανόνες ανάλυσης) για την ερμηνεία της τιμής ενός στοιχείου

Χρήση των εξειδικευτών του Dublin Core



Παραδείγματα δηλώσεων



Qualified Dublin Core

- Οι ερμηνείες των στοιχείων **εξειδικεύονται**
 - Γίνονται στενότερες
 - Δεν γίνονται ποτέ ευρύτερες!
- Ισχύει η αρχή "dump-down principle"
 - Η μηχανή μπορεί να αγνοήσει όποιον εξειδικευτή δεν μπορεί να ερμηνεύσει
 - Η πληροφορία του εξειδικευτή χρησιμοποιείται σε αυτές τις περιπτώσεις ως μη εξειδικευμένη

Συμπεράσματα

- Σκοπός χρήσης των μεταδεδομένων: διαλειτουργικότητα, διαμοιρασμός πόρων, σαφής σημασιολογία
- Όσο πιο εξειδικευμένα, τόσο δυσκολότερα επιτυγχάνεται η διαλειτουργικότητα
 - Γιατί;
- Χρήση προτύπων
- Διαλειτουργικότητα επιτυχής διασύνδεση/διαμοιρασμός πόρων
επιτυχής ανακάλυψη, νάκτηση και περιήγηση στους ψηφιακούς πόρους




Ξεφύλλισέ το

Κύρια χαρακτηριστικά

[Αναλυτικές λεπτομέρειες](#)

Συγγραφέας Murakami, Haruki, 1949-	Εκδότης Ψυχογιός	ISBN 9789604966530
Αριθμός Σελίδων 453	Εξώφυλλο Μαλακό εξώφυλλο	Διαστάσεις 21x14

Άσκηση

Χαρακτηριστικά

Συγγραφέας:	Murakami, Haruki, 1949-	Εκδότης:	Ψυχογιός
ISBN:	9789604966530	Εξώφυλλο:	Μαλακό εξώφυλλο
Διαστάσεις:	21x14	Αριθμός Σελίδων:	453
Ημερ/νία έκδοσης:	01/04/2012	Γλώσσα Γραφής:	ελληνικά
Βάρος βιβλίου:	699		

Βασικές πληροφορίες

Χρονολογία έκδοσης:	2012	Τόπος έκδοσης:	Αθήνα
Γλώσσα πρωτοτύπου:	ιαπωνικά	Τίτλος πρωτοτύπου:	Norwegian Wood
Συντελεστές:	Μετάφραση: Αργυρώ Μαντόγλου	ISBN 10:	960-496-653-7
Κατηγορία Βιβλίου:	Μεταφρασμένη Λογοτεχνία		

ΠΕΡΙΕΧΟΜΕΝΑ

Dublin Core



Προφίλ Εφαρμογών



XML



Προφίλ εφαρμογών

- Αποτελείται από
 - ένα σύνολο στοιχείων μεταδεδομένων (προερχόμενα από ένα ή περισσότερα σχήματα μεταδεδομένων–χώρων ονομάτων–λεξιλογίων),
 - ένα σύνολο πολιτικών, και
 - ένα σύνολο οδηγιώνπου έχουν οριστεί για μια συγκεκριμένη εφαρμογή.
- Έγγραφο (ή ένα σύνολο εγγράφων) που περιγράφει τα απαραίτητα μεταδεδομένα
 - για συγκεκριμένη εφαρμογή
 - + τρόπος χρήσης τους

Χαρακτηριστικά προφίλ εφαρμογών

- Τα στοιχεία μεταδεδομένων μπορούν να προέρχονται από εξωτερικά σχήματα ή/και από τοπικά ορισμένα σχήματα
- Ανάπτυξη Προφίλ Εφαρμογών (και από ομάδες εργασίας του Dublin Core)
 - προς επέκταση/εξειδίκευση των στοιχείων μεταδεδομένων
 - προσαρμοσμένων σε συγκεκριμένες εφαρμογές
- Παραδείγματα: [DC Library Application Profile](#), [UK e-Government Metadata Standard Application Profile](#)

Χαρακτηριστικά προφίλ εφαρμογών II

- Προσφέρουν μεγάλη εξειδίκευση στοιχείων μεταδεδομένων
- ΔΕΝ αποτελούν αυτόνομα ελεγχόμενα λεξιλόγια
 - Χρησιμοποιούν ένα ή περισσότερα υπάρχοντα λεξιλόγια
 - Μπορούν να περιέχουν και τοπικά ορισμένα λεξιλόγια

ΠΕΡΙΕΧΟΜΕΝΑ

Dublin Core



Προφίλ Εφαρμογών



XML



7. XML

Γλώσσα XML

- XML: eXtensible Markup Language
 - Γλώσσα σήμανσης (markup)
 - Διάλεκτος της Standard Generalized Markup Language (SGML)
 - Δημιουργήθηκε για τη διευκόλυνση της ανταλλαγής και επεξεργασίας δεδομένων στον παγκόσμιο ιστό
- Η XML δεν αντικαθιστά την HTML !
 - HTML: σκοπός η **προβολή** των δεδομένων, εστιάζοντας στο πώς τα δεδομένα εμφανίζονται
 - XML: σκοπός η **περιγραφή** των δεδομένων, εστιάζοντας στο τι πραγματικά είναι τα δεδομένα αυτά
 - Η XML συμπληρώνει την HTML ορίζοντας το περιεχόμενό τους

Σήμανση (markup)

- Είδη σήμανσης (markup)
 - Παρουσίασης -> εμφάνιση του τεκμηρίου
 - Δομική -> δομή του τεκμηρίου
 - Σημασιολογική -> περιεχόμενο του τεκμηρίου
- Στην XML μας ενδιαφέρει μόνο η **δομική** σήμανση

Δομή τεκμηρίων xml

- Κάθε xml τεκμήριο ξεκινά με τη δηλωτική εντολή της XML

```
<?xml version="1.0"?>
```

- Στη συνέχεια, με χρήση ετικετών (tags) δίνονται οι δηλώσεις των στοιχείων (elements)
- Ετικέτες (tags): περιγραφικές, ελεύθερου περιεχομένου
- Στοιχεία (elements):
 - βασική δομική μονάδα ενός XML τεκμηρίου
 - οριοθετείται από την ετικέτα αρχής και την ετικέτα τέλους
 - κατάλληλα εμφωλευμένα
 - μπορούν να διαθέτουν γνωρίσματα (attributes), για παροχή πρόσθετης πληροφορίας

Γνωρίσματα στοιχείων

- Η δήλωση των γνωρισμάτων πραγματοποιείται αμέσως μετά την ετικέτα αρχής του στοιχείου
 - Οι τιμές των γνωρισμάτων πρέπει να περικλείονται από εισαγωγικά
<ετικέτα γνώρισμα = “τιμή_γνωρίσματος”> </ετικέτα>
 - Κάθε στοιχείο μπορεί να διαθέτει ένα ή περισσότερα γνωρίσματα <ετικέτα γνώρ_1 = “τιμή_1” γνώρ_2 = “τιμή_2” > </ετικέτα>

```
<note id="501" date="22/11/2021">  
  <to>Dave</to>  
  <from>Jack</from>  
  <heading>Reminder</heading>  
  <body>Don't forget the XML exercises!</body>  
</note>
```


Δενδρική δομή XML τεκμηρίων

- Κάθε τεκμήριο XML περιέχει ένα βασικό στοιχείο (root element)
- Δημιουργία δενδρικής δομής μέσω εμφώλευσης (nesting) στοιχείων

```
<root>  
  <child>  
    <subchild> ... </subchild>  
  </child>  
</root>
```

Τύποι στοιχείων XML (elements)

- Απλά στοιχεία: περιέχουν μόνο κείμενο

```
<title>Harry Potter</title>
```

- Σύνθετα στοιχεία: περιέχουν άλλα στοιχεία στις δηλώσεις τους

```
<book>  
  <title>Harry Potter</title>  
  <author>J. K. Rowling</author>  
</book>
```

Τύποι στοιχείων XML (elements) II

- Σύνθετα στοιχεία με ανάμικτο περιεχόμενο: περιέχουν άλλα στοιχεία στις δηλώσεις τους, ταυτόχρονα με κείμενο

```
<book>
```

```
  The book with title <title>Harry Potter</title>
```

```
  written by <author>J. K. Rowling</author>
```

```
</book>
```

Κανόνες σύνταξης XML

- Σε κάθε ετικέτα αρχής του τεκμηρίου πρέπει να αντιστοιχεί μια ετικέτα τέλους
 - Εκτός από την δηλωτική εντολή της `<xml` στην αρχή του τεκμηρίου (δεν αποτελεί μέρος του ίδιου του XML τεκμηρίου)
- Διάκριση πεζών-κεφαλαίων στις ετικέτες
- Όλα τα στοιχεία πρέπει να είναι κατάλληλα εμφωλευμένα
 - Οι ετικέτες τέλους πρέπει να εμφανίζονται με την αντίστροφη σειρά από τη σειρά εμφάνισης των ετικετών αρχής

Κανόνες σύνταξης XML II

- Κάθε XML τεκμήριο πρέπει να έχει ένα στοιχείο-ρίζα (root element)
- Επιτρέπονται τα κενά στοιχεία
- Διατηρούνται τα κενά διαστήματα
- Εισαγωγή σχολίων:

`<!--Comment-->`

Ειδικοί χαρακτήρες - αναφορές οντοτήτων

<	<	less than
>	>	greater than
&	&	ampersand
'	'	apostrophe
"	"	quotation mark

This XML file does not appear to have any style information associated with it. The document tree is shown below.

```
▼ <breakfast_menu>
  ▼ <food>
    <name>Belgian Waffles</name>
    <price>$5.95</price>
    <description>Two of our famous Belgian Waffles with plenty of real maple syrup</description>
    <calories>650</calories>
  </food>
  ▼ <food>
    <name>Strawberry Belgian Waffles</name>
    <price>$7.95</price>
    <description>Light Belgian waffles covered with strawberries and whipped cream</description>
    <calories>900</calories>
  </food>
  ▼ <food>
    <name>Berry-Berry Belgian Waffles</name>
    <price>$8.95</price>
    <description>Light Belgian waffles covered with an assortment of fresh berries and whipped cream</description>
    <calories>900</calories>
  </food>
  ▼ <food>
    <name>French Toast</name>
    <price>$4.50</price>
    <description>Thick slices made from our homemade sourdough bread</description>
    <calories>600</calories>
  </food>
  ▼ <food>
    <name>Homestyle Breakfast</name>
    <price>$6.95</price>
    <description>Two eggs, bacon or sausage, toast, and our ever-popular hash browns</description>
    <calories>950</calories>
  </food>
</breakfast_menu>
```

```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <breakfast_menu>
3   <food>
4     <name>Belgian Waffles</name>
5     <price>$5.95</price>
6     <description>Two of our famous Belgian Waffles with plenty of real maple syrup</description>
7     <calories>650</calories>
8   </food>
9   <food>
10    <name>Strawberry Belgian Waffles</name>
11    <price>$7.95</price>
12    <description>Light Belgian waffles covered with strawberries and whipped cream</description>
13    <calories>900</calories>
14  </food>
15  <food>
16    <name>Berry-Berry Belgian Waffles</name>
17    <price>$8.95</price>
18    <description>Light Belgian waffles covered with an assortment of fresh berries and whipped cream</description>
19    <calories>900</calories>
20  </food>
21  <food>
22    <name>French Toast</name>
23    <price>$4.50</price>
24    <description>Thick slices made from our homemade sourdough bread</description>
25    <calories>600</calories>
26  </food>
27  <food>
28    <name>Homestyle Breakfast</name>
29    <price>$6.95</price>
30    <description>Two eggs, bacon or sausage, toast, and our ever-popular hash browns</description>
31    <calories>950</calories>
32  </food>
33 </breakfast_menu>
```




Κύρια χαρακτηριστικά

[Αναλυτικές λεπτομέρειες](#)

Συγγραφέας Murakami, Haruki, 1949-	Εκδότης Ψυχογιός	ISBN 9789604966530
Αριθμός Σελίδων 453	Εξώφυλλο Μαλακό εξώφυλλο	Διαστάσεις 21x14

Άσκηση

Χαρακτηριστικά			
Συγγραφέας:	Murakami, Haruki, 1949-	Εκδότης:	Ψυχογιός
ISBN:	9789604966530	Εξώφυλλο:	Μαλακό εξώφυλλο
Διαστάσεις:	21x14	Αριθμός Σελίδων:	453
Ημερ/νία έκδοσης:	01/04/2012	Γλώσσα Γραφής:	ελληνικά
Βάρος βιβλίου:	699		
Βασικές πληροφορίες			
Χρονολογία έκδοσης:	2012	Τόπος έκδοσης:	Αθήνα
Γλώσσα πρωτοτύπου:	ιαπωνικά	Τίτλος πρωτοτύπου:	Norwegian Wood
Συντελεστές:	Μετάφραση: Αργυρώ Μαντόγλου	ISBN 10:	960-496-653-7
Κατηγορία Βιβλίου:	Μεταφρασμένη Λογοτεχνία		

Δήλωση του DC namespace: `<metadata xmlns:dc="http://purl.org/dc/elements/1.1/">`