

Μεταδεδομένα  
για  
Ψηφιακές Βιβλιοθήκες

Γ. Δ. Μπώιος

# Μεταδεδομένα: Ο όρος

- Μεταδεδομένα: «**Δεδομένα σχετικά με Δεδομένα**»
  - Αναλυτικότερα: «Το σύνολο όσων θα μπορούσε να πει κανείς για ένα πληροφοριακό αντικείμενο σε κάθε συγκεκριμένη στιγμή της πληροφοριακής του ζωής»
  - Καλύτερα: «Πληροφορίες που αφορούν τη φυσική και διανοητική **μορφή**, το **περιεχόμενο**, το **περιβάλλον** παραγωγής και χρήσης (ποιος, τι, γιατί κλπ.), καθώς και τη **δομή** ενός πληροφοριακού αντικειμένου»

# Μεταδεδομένα: Προέλευση του όρου

- Ένας μάλλον σύγχρονος και «της μόδας» όρος για να υποδηλωθεί κάτι, γενικά, **γνωστό από πολύ καιρό**: η «προστιθέμενη αξία» που, με μορφή πληροφοριών, δημιουργούν τα αρχεία, οι βιβλιοθήκες, τα μουσεία κλπ. για την εξασφάλιση της πρόσβασης στις συλλογές τους
- Ο όρος χρησιμοποιείται για πρώτη φορά στο περιβάλλον των ψηφιακών αντικειμένων και, κυρίως, στο χώρο της διαχείρισης και του σχεδιασμού πληροφοριακών συστημάτων

# Μεταδεδομένα και συμβατικές βιβλιοθήκες

- Η συμβατική Βιβλιοθηκονομία έχει παρελθόν δεκαετιών στην παραγωγή «Μεταδεδομένων»
- Εγγραφές Καταλόγου, σε συμβατική ή ηλεκτρονική μορφή, διαμορφωμένες με βάση πρότυπα και εργαλεία όπως οι AACR ή τα ISBDs, τα LCSH, MeSH, τα ποικίλα MARC κλπ. αποτελούν τέτοια «Μεταδεδομένα»
- «Μεταδεδομένα» είναι, επίσης, τα ευρετήρια, οι περιλήψεις και άλλα ανάλογα προϊόντα
- Ανάλογα εργαλεία χρησιμοποιούνται πρόσφατα και στην αρχειονομική κοινότητα (MARC for AMC, ISAD(G), EAD κλπ.)

# Μεταδεδομένα: Διευκρινήσεις

- Δεν είναι μόνο ψηφιακά
- Αφορούν πολύ περισσότερα πράγματα από την απλή περιγραφή ενός τεκμηρίου
- Έχουν πολλαπλές πηγές προέλευσης
- Αρχίζουν με τη δημιουργία και συνεχίζουν να αυξάνονται στη διάρκεια ζωής του τεκμηρίου
- Τα Μεταδεδομένα ενός τεκμηρίου μπορεί να είναι Δεδομένα ενός άλλου

# Μεταδεδομένα: Οι χρήσεις

- Στο ψηφιακό περιβάλλον τα *Μεταδεδομένα* μπορούν να υποστηρίξουν πολλές δυνητικά χρήσεις: **Ενδεικτικά:**
  - Εντοπισμό πληροφοριακών πόρων (ψηφιακών αντικειμένων, τεκμηρίων)
  - Εντοπισμό και ειμετάλλευση των δομικών σχέσεων μεταξύ και στο εσωτερικό πληροφοριακών αντικειμένων
  - Αξιολόγηση περιεχομένου
  - Ηλεκτρονικό επιχειρείν
  - Έλεγχο και πιστοποίηση αυθεντικότητας και βαθμού πληρότητας περιεχομένου
  - Διαχείριση δεδομένων, τεκμηρίων και συλλογών
  - Διαχείριση δικαιωμάτων πνευματικής ιδιοκτησίας
  - Διατήρηση ψηφιακού υλικού

# Μεταδεδομένα: Οι κατηγορίες

## ■ Διοικητικά

- Χρησιμοποιούνται στη διαχείριση πληροφοριακών πόρων (π.χ. πληροφορίες προσκλήσεων, Έλεγχος και τεκμηρίωση απαιτήσεων νόμιμης πρόσβασης, πληροφορίες θέσης αντικειμένου κλπ.)

## ■ Περιγραφικά

- Χρησιμοποιούνται για την περιγραφή & ταύτιση πληροφοριακών πόρων (καταλογογραφικές εγγραφές, ευρετήρια, υπομνηματισμοί & σχόλια χρηστών κλπ.)

## ■ Διατήρησης

- Χρησιμοποιούνται για τη διαχείριση των θεμάτων διατήρησης των πληροφοριακών πόρων (Τεκμηρίωση συνθηκών παραγωγής ή αναπαραγωγής των ψηφιακών αντικειμένων κλπ.)

## ■ Τεχνικά

- Χρησιμοποιούνται για την τεκμηρίωση του τρόπου λειτουργίας ενός συστήματος (τεκμηρίωση εξοπλισμού και λογισμικού, πληροφορίες ψηφιοποίησης, πληροφορίες που αφορούν τον έλεγχο πρόσβασης κλπ.)

## ■ Χρήσης

- Αφορούν τη χρήση των ψηφιακών αντικειμένων

# Μεταδεδομένα: Προέλευση

## ■ Εσωτερικά Μεταδεδομένα

- Δημιουργούνται από τον παραγωγό του ψηφιακού αντικειμένου, την ώρα της παραγωγής (πρωτογενούς δημιουργίας ή ψηφιοποίησης) και είναι ενσωματωμένα σ' αυτό (ονόματα και τύποι αρχείων, μορφή συμπίεσης κλπ.)

## ■ Εξωτερικά Μεταδεδομένα

- Δημιουργούνται αργότερα, συνήθως από κάποιον άλλο, διαφορετικό από το δημιουργό, ή από άλλη αρμόδια υπηρεσία (π.χ. καταλογογραφικές εγγραφές)



# Μεταδεδομένα: Μέθοδοι παραγωγής

## ■ Αυτόματη παραγωγή

- Αυτόματη δημιουργία από το υπολογιστικό σύστημα (π.χ. ευρετήρια λέξεων-κλειδιών, καταγραφές χρήσης του ψηφιακού αντικειμένου κλπ.)

## ■ Χειροκίνητη παραγωγή

- Δημιουργούνται με ανθρώπινη προσπάθεια (π.χ. καταλογογραφικές εγγραφές, περιγραφές Dublin Core κλπ.)

# Η φύση των Μεταδεδομένων

## ■ Απλά

- Παράγονται από ανθρώπους που δεν είναι ειδικοί ούτε σε θέματα πληροφόρησης ούτε στο θεματικό χώρο του ψηφιακού αντικειμένου και, συνήθως, από το δημιουργό του αντικειμένου (π.χ. Μεταδεδομένα μια προσωπικής σελίδας web)

## ■ Εξειδικευμένα

- Παράγονται από ειδικούς της πληροφόρησης ή ειδικούς στο θεματικό αντικείμενο (π.χ. Θεματικές επικεφαλίδες, Εγγραφές MARC κλπ.)

# Η κατάσταση των Μεταδεδομένων

## ■ Στατικά

- Από τη στιγμή της δημιουργίας τους δεν μεταβάλλονται (π.χ. Τίτλος του τεκμηρίου)

## ■ Δυναμικά

- Μπορούν να μεταβάλλονται με το χειρισμό ή τη χρήση του τεκμηρίου (π.χ. Η ανάλυση μιας εικόνας)

## ■ Μακροπρόθεσμα

- Εξασφαλίζουν τη μακροπρόθεσμη χρηστικότητα του τεκμηρίου (π.χ. Πληροφορίες μορφής, επεξεργασίας, δικαιωμάτων ιδιοκτησίας)

## ■ Βραχυπρόθεσμα

- Συνήθως καταγραφές χρήσεως του ψηφιακού αντικειμένου

# Μεταδεδομένα: Η δομή

## ■ Δομημένα

- Δημιουργούνται σύμφωνα με μια δεδομένη πρότυπη ή όχι πρότυπη δομή (π.χ. MARC, TEI, EAD κλπ.)

## ■ Αδόμητα

- Δεν υπακούουν σε συγκεκριμένη, πρότυπη ή όχι, τυπική και προδιαγεγραμμένη δομή (π.χ. σημειώσεις και σχόλια)

# Μεταδεδομένα: Η σημασιολογία

## ■ Ελεγχόμενα

- Προέρχονται ή συμφωνούν με τις προβλέψεις τυποποιημένων και ελεγχόμενων λεξιλογίων ή μορφών

## ■ Όχι ελεγχόμενα

- Δεν προέρχονται ή συμφωνούν με τις προβλέψεις τυποποιημένων και ελεγχόμενων λεξιλογίων ή μορφών

# Μεταδεδομένα: Το επίπεδο αναφοράς

## ■ Συλλογής

- Μεταδεδομένα που αφορούν συλλογές ψηφιακών αντικειμένων

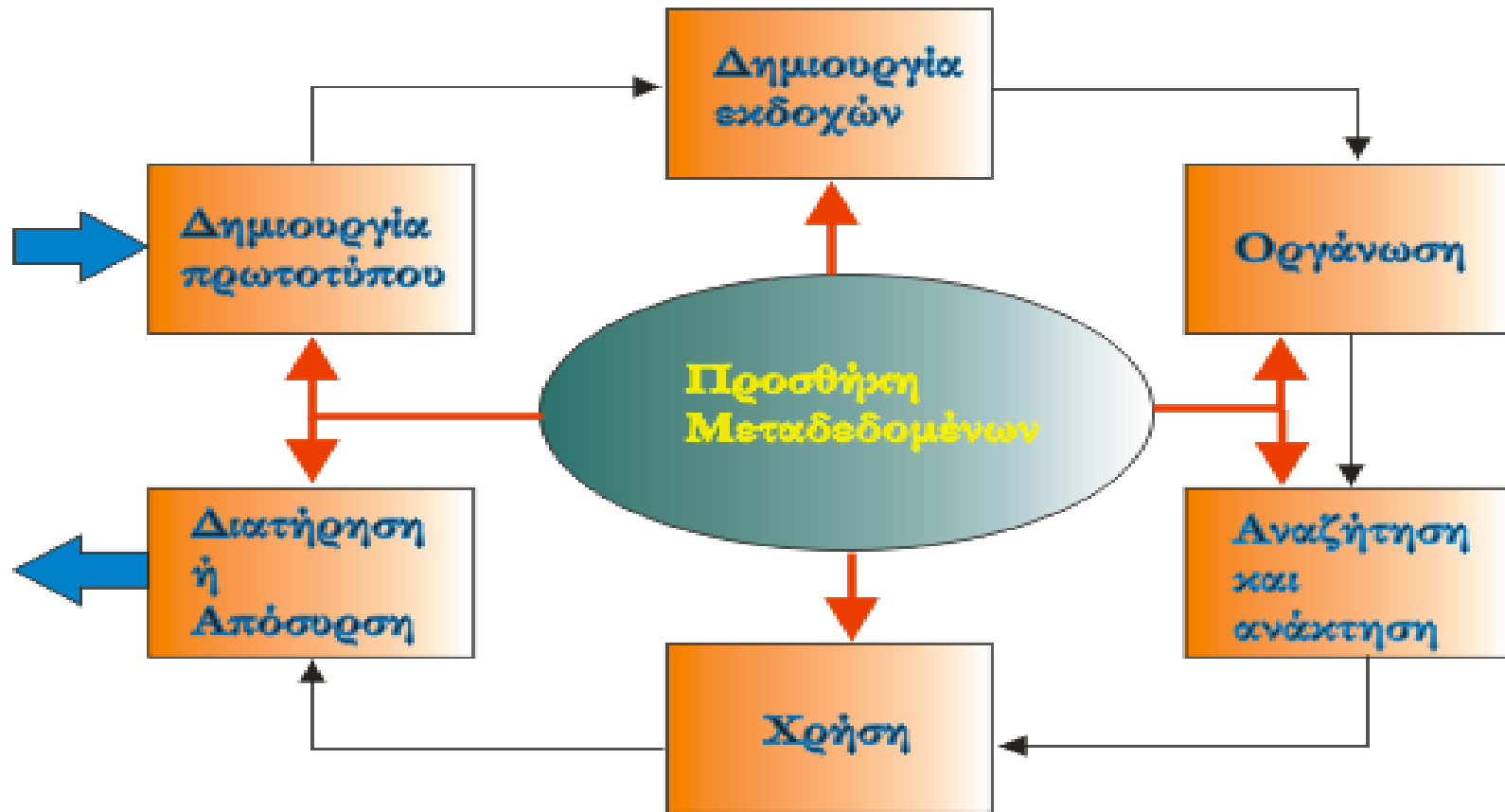
## ■ Μονάδας

- Μεταδεδομένα που αφορούν συγκεκριμένα ψηφιακά αντικείμενα, μέρος ή όχι αντίστοιχων συλλογών

# Μεταδεδομένα: Η σημασία

- Αυξημένες δυνατότητες πρόσβασης
  - Αξιόπιστοι μηχανισμοί αναζήτησης, αναζήτηση σε μεγάλο εύρος συλλογών, δημιουργία εικονικών συλλογών κλπ.
- Διεύρυνση της χρήσης των πληροφοριακών πόρων
  - Πολλαπλοί τρόποι αναζήτησης, συνδυασμού, παρουσίασης και χειρισμού πληροφοριακών πόρων
- Διευκόλυνση των πολλαπλών ειδοχών του ψηφιακού αντικειμένου
- Διευκόλυνση διαχείρισης νομικών θεμάτων
- Διατήρηση ψηφιακών αντικειμένων
- Προαγωγή της αποτελεσματικότητας των συστημάτων πληροφόρησης

# Μεταδεδομένα: Η παραγωγή





# Μεταδεδομένα: Τα σχήματα

## ■ Σχήμα Μεταδεδομένων

- Τυποποιημένη προδιαγραφή διάταξης, μορφής και/ή περιεχομένου για την παραγωγή Μεταδεδομένων
- Στο τυπικό περιβάλλον της συμβατικής βιβλιοθήκης τέτοια «σχήματα» είναι οι κανόνες καταλογογράφησης, τα ISBDs, οι διατάξεις MARC κλπ.

## ■ Πληθώρα και ποικιλία προέλευσης σχημάτων

- Δημιουργούνται από επιμέρους κοινότητες ενδιαφερομένων, που επιδιώκουν να περιγράψουν τα ψηφιακά αντικείμενα του χώρου τους

# Μεταδεδομένα: Η τυπολογία των σχημάτων

Απλό

Πολύπλοκο



| 1  | 2   | 3  | 4                                       |
|--|---|--|---|
| Ευρετήρια πλήρους κειμένου                           | Απλές, δομημένες, γενικού χαρακτήρα διατάξεις           | Πολυπλοκότερη δομή, προσαρμοσμένη σε ειδική θεματική περιοχή | Τμήμα ευρύτερου σημασιολογικού πλαισίου |
| Ειδικές, περιορισμένης χρήσης, διατάξεις, π.χ. Yahoo | Dublin Core<br>ROADS<br>IAFA/Whois++<br>PCC Core<br>CDP | FGDC<br>MARC 21<br>GILS                                      | TEI<br>EAD<br>CIMI                      |

Λιτό

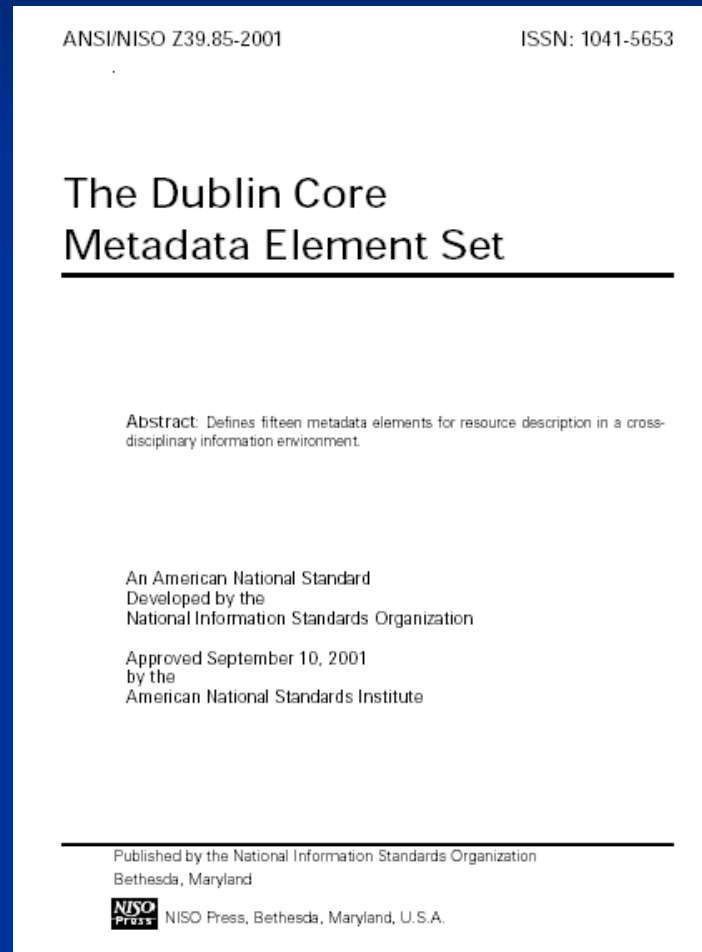
Πλούσιο



# Dublin Core Metadata Element Set

- Η διαδικασία ανάπτυξης (Dublin Core Metadata Initiative, DCMI) αρχίζει το 1995
- Στόχος:
  - Ο προσδιορισμός του ελάχιστου απαραίτητου συνόλου στοιχείων Μεταδεδομένων για τη διευκόλυνση της περιγραφής, αναζήτησης και ανεύρεσης ψηφιακών τεκμηρίων σε περιβάλλον Διαδικτύου
- Η προσπάθεια, διεπιστημονική και διεθνής, οδήγησε στην ανάπτυξη ενός συνόλου 15 στοιχείων περιγραφής
- Δεν στοχεύει στην υποκατάσταση άλλων σχημάτων, αλλά στη συνύπαρξη μαζί τους
- Είχε γρήγορη αποδοχή και ευρεία χρήση
  - Μέχρι το 2000 είχε 20 μεταφράσεις και υιοθέτηση από ποικίλα όργανα, χώρες κλπ.
- Το 2001 εντάχθηκε στα πρότυπα του NISO, ως ANSI/NISO Z39.85-2001

# Dublin Core: Το πρότυπο



# Dublin Core: Τα στοιχεία δεδομένων

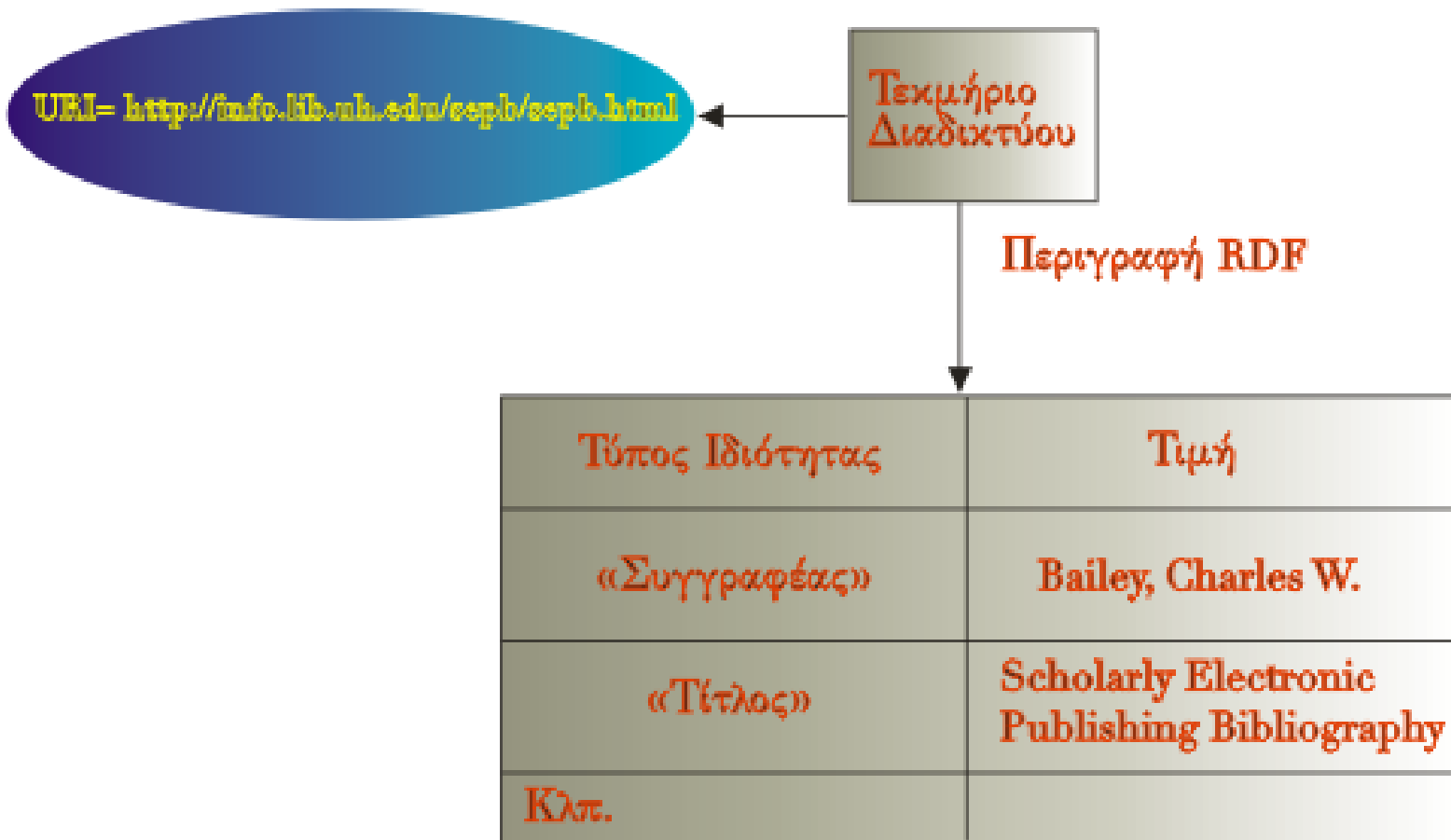
- Θέμα (Subject)
- Περιγραφή (Description)
- Δημιουργός (Creator)
- Τίτλος (Title)
- Εκδότης (Publisher)
- Υπεύθυνος συμβολής (Contributor)
- Χρονολογία (Date)
- Τύπος (Type)
- Ταύτιση (Identifier)
- Πηγή (Source)
- Γλώσσα (Language)
- Σχέση (Relation)
- Κάλυψη (Coverage)
- Δικαιώματα (Rights)
- Μορφή (Format)

# Resource Data Framework

## RDF

- Τα πολλαπλά και πολλαπλής προέλευσης σχήματα παραγωγής δημιουργούν την ανάγκη δημιουργίας ενός πλαισίου **υποδοχής, ανταλλαγής και ειμετάλλευσης** Μεταδεδομένων
- **Στόχος ανάπτυξης του RDF:**
  - Ενιαίο, διαλειτουργικό πλαίσιο για τη χρήση και ανταλλαγή Μεταδεδομένων, προερχομένων από διαφορετικά περιβάλλοντα
- **Η βασική ιδέα:**
  - Οποιοσδήποτε πληροφοριακός πόρος του Διαδικτύου μπορεί να περιγραφεί από ένα σύνολο «ιδιοτήτων». Κάθε «ιδιότητα» ανήκει σε ένα «τύπο» και έχει, για το κάθε τεκμήριο, μια «τιμή»
  - Μπορούν να χρησιμοποιηθούν «ιδιότητες» από ποικίλα σχήματα, με την προϋπόθεση ότι αυτό δηλώνεται
- **XML και RDF**
  - Συντακτικό όχημα μεταφοράς των Μεταδεδομένων αποτελεί για το RDF η γλώσσα XML

# RDF: Η λογική της περιγραφής



# Μεταδεδομένα: Τα εργαλεία παραγωγής

- Επεξεργαστές Μεταδεδομένων (Metadata editors)
- Διαθέσιμοι, ως εφαρμογές, στο Διαδίκτυο
  - Αυτόματη περιγραφή ψηφιακού αντικειμένου, αν δοθεί η διεύθυνσή του στο Διαδίκτυο, π.χ. DC dot
  - Συμπλήρωση ειδικής φόρμας και επιστροφή του δομημένου συνόλου Μεταδεδομένων, π.χ. Reggie Metadata Editor, Nordic Metadata Template
- Αυτόνομες εφαρμογές
  - NoteTab Light
- Εμπορικές εφαρμογές
  - Taggen, Taggendc



# DC dot



## Dublin Core metadata editor

---

Type the URL of the page you want to describe...

# DC dot



## Dublin Core metadata editor

---

Type the URL of the page you want to describe...

# DC dot



## Dublin Core metadata editor

---

Results for URL: <http://www.ionio.gr> [[summary](#)]

```
<link rel="schema.DC" href="http://purl.org/dc">
<meta name="DC.Title" content="Ionian University">
<meta name="DC.Publisher" content="European Regional Internet Registry/RIPE NCC">
<meta name="DC.Type" scheme="DCMIType" content="Text">
<meta name="DC.Format" content="text/html">
<meta name="DC.Format" content="2452 bytes">
<meta name="DC.Identifier" content="http://www.ionio.gr">
```

---

If necessary, edit the values in the boxes below, and

Convert metadata to

# DC assist



UKOLN

Application: **DC-Qualified**

## Title

Name

title

Definition

A name given to the resource.

Comment

Typically, a Title will be a name by which the resource is formally known.

Namespace

DCMES version 1.1

Refinements

[alternative](#)

Encoding schemes

[title](#)  
[creator](#)  
[subject](#)  
[description](#)  
[publisher](#)  
[contributor](#)  
[date](#)  
[type](#)  
[format](#)  
[identifier](#)  
[source](#)  
[language](#)  
[relation](#)  
[coverage](#)  
[rights](#)

[HTML example\(s\)](#)

# Reggie

Please choose a schema and language:

**Schema:**

Choose...



**Language:**

English



Or, enter the URL for another schema:

If describing a web page, enter the URL:

**Go!**

# Reggie

Please choose a schema and language:

**Schema:**

Dublin Core



**Language:**

English



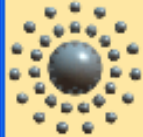

Or, enter the URL for another schema:

If describing a web page, enter the URL:

**Go!**

# Reggie

Reggie Metadata Editor

 **Dublin Core**  **About Reggie** **Quit**

**About Dublin Core** Send feedback to: [rdu-info@dstc.edu.au](mailto:rdu-info@dstc.edu.au) Reggie v1.62 © 1998 DSTC Pty Ltd

**Hide / Restore Fields** **Clear All Values** **Select a Syntax...** **Preview** **Export...**

? **Title:**  
+  X  
SubElement: **None** Language: **English**

? **Author or Creator:**  
+  X  
SubElement: **None** Language: **English**

? **Subject and Keywords:**  
+  X  
Scheme: **None** Language: **English**

? **Description:**  
+  X  
Language: **English**

# Nordic Metadata Template

**1 TITLE** of the resource to be described



Alternative title (Titles other than main title)



**2 CREATOR** (Name of the person or organization primarily responsible for creating the intellectual content)

Creator name



Creator's (Email) address

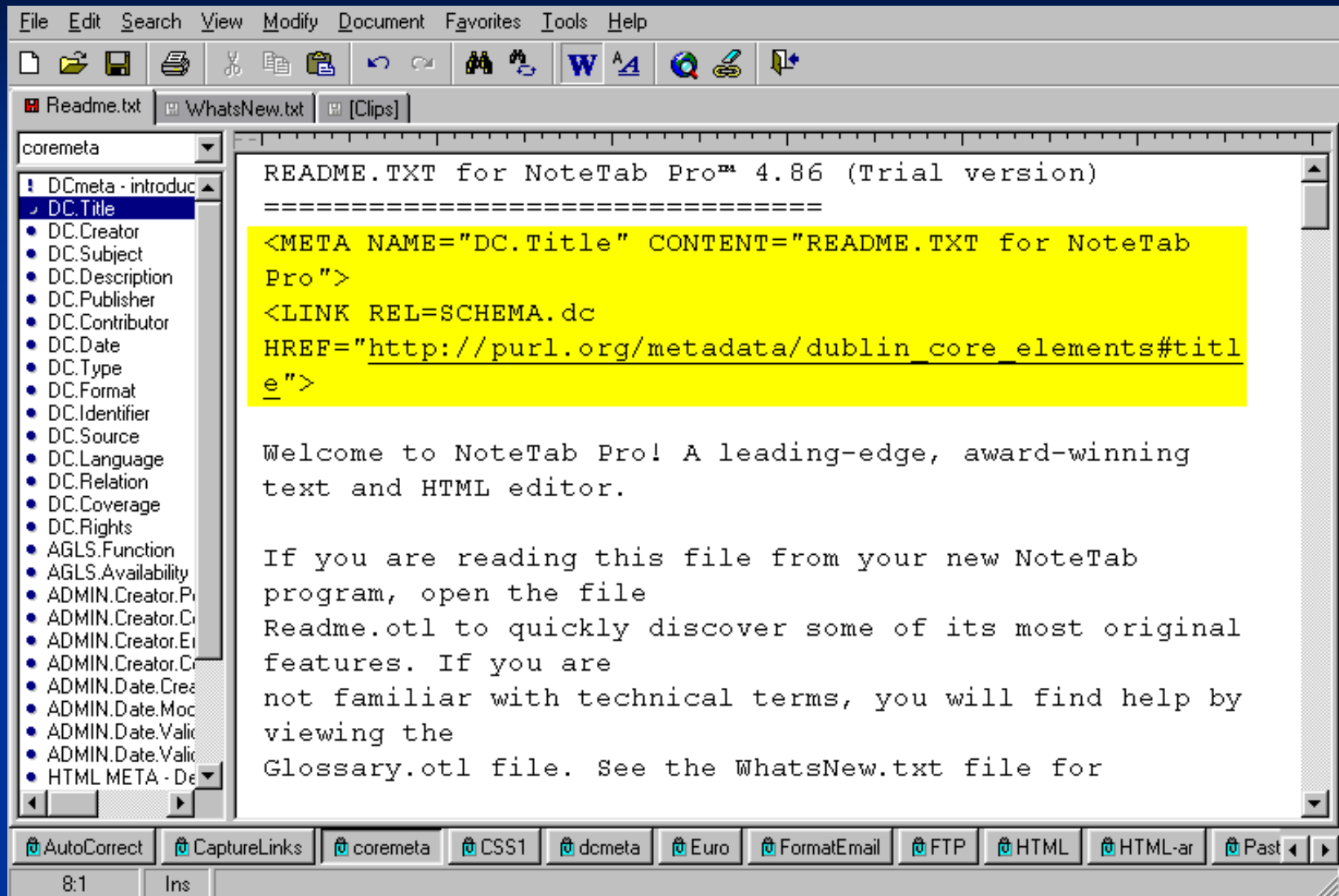


**3 SUBJECT: Keywords** (Your own keywords describing the topic of the resource, *one per box*)

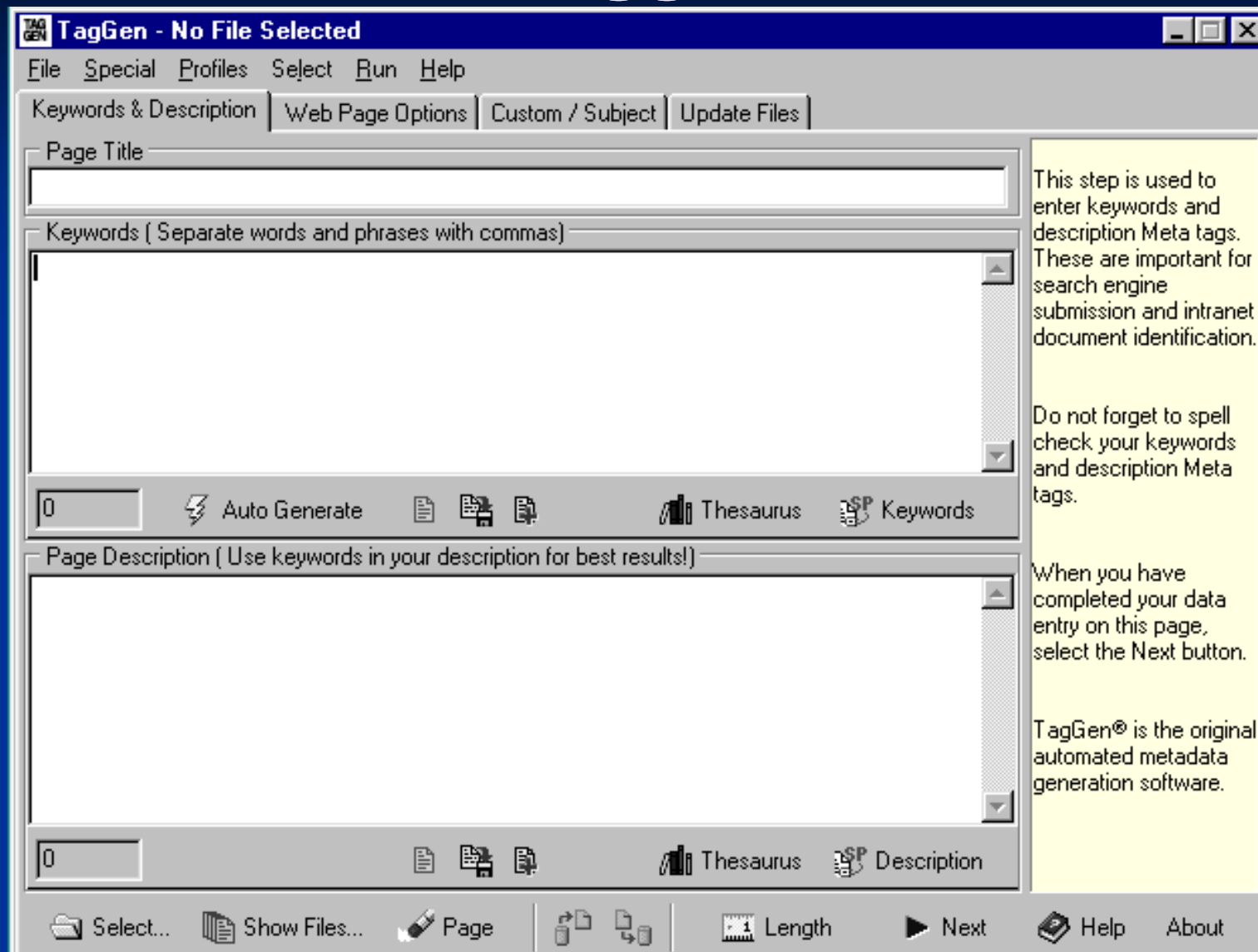




# NoteTabPro



# Taggen



# DC σε HTML

```
<link rel="schema.DC" href="http://purl.org/dc">
<meta name="DC.Title" content="UKOLN: DC-dot Dublin Core metadata editor">
<meta name="DC.Creator" content="Andy Powell">
<meta name="DC.Subject" content="Dublin Core; DC; generator; editor; Warwick Framework; SOIF; TEI; USMARC;
XML; GILS; ROADS; RDF; IMS">
<meta name="DC.Description" content="A CGI based Dublin Core metadata generator">
<meta name="DC.Publisher" content="UKOLN, University of Bath">
<meta name="DC.Date" scheme="W3CDTF" content="2001-12-11">
<meta name="DC.Type" content="Text">
<meta name="DC.Format" content="text/html">
<meta name="DC.Identifier" content="http://www.ukoln.ac.uk/metadata/dcdot/">
<meta name="DC.Rights" content="http://www.ukoln.ac.uk/metadata/dcdot/COPYING">
```

# DC σε RDF

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
         xmlns:dc="http://purl.org/dc/elements/1.1/">
  <rdf:Description>
    <dc:creator>Karl Mustermann</dc:creator>
    <dc:title>Algebra</dc:title>
    <dc:subject>mathematics</dc:subject>
    <dc:date>2000-01-23</dc:date>
    <dc:language>EN</dc:language>
    <dc:description>An introduction to algebra</dc:description>
  </rdf:Description>
</rdf:RDF>
```

# TEI Header

```
<TEIHEADER><FILEDESC>  
<TITLESTMT><TITLE>Liberty Lyrics (1895):  
a machine-readable transcription</TITLE>  
<AUTHOR>Bevington, Louisa Sarah (Guggenberger)  
(1845-?) </AUTHOR>  
<RESPSTMT><RESP>Transcribed and encoded by </RESP>  
<NAME>Felix Jung</NAME></RESPSTMT>  
<RESPSTMT><RESP>Edited by </RESP>  
<NAME>Perry Willett</NAME></RESPSTMT></TITLESTMT>  
<EXTENT>TEI formatted filesize uncompressed&colon;  
1426 bytes</EXTENT>  
<PUBLICATIONSTMT>  
<PUBLISHER>Library Electronic Text Resource Service  
(LETRS), Indiana University</PUBLISHER>  
<DATE>September 22, 1995</DATE>  
<AVAILABILITY><P>&copy; 1995, The Trustees of  
Indiana University. Indiana University makes a  
claim of copyright only to original contributions
```