



ΜΟΝΑΔΕΣ ΑΡΙΣΤΕΙΑΣ  
ΑΝΟΙΧΤΟΥ ΛΟΓΙΣΜΙΚΟΥ



# Μεταδεδομένα ψηφιακού περιεχομένου

Ελεύθερο λογισμικό και λογισμικό  
ανοιχτού κώδικα για τη δημιουργία  
ψηφιακών βιβλιοθηκών - αποθετηρίων

Αλέξανδρος Ταγκούλης



Ευρωπαϊκή Ένωση  
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ  
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Αριστεία ΕΛ/ΛΑΚ ΤΕΙ Αθήνας



2

# Μεταδεδομένα

Δεδομένα για τα δεδομένα

# Μεταδεδομένα - Ορισμός

**Τα μεταδεδομένα ορίζονται ως τα δεδομένα για τα δεδομένα**

Είναι δομημένη πληροφορία η οποία χρησιμοποιείται για την περιγραφή, την επεξήγηση, τον εντοπισμό, τη χρήση και τη διαχείριση μίας πληροφοριακής πηγής, όπως ένα βιβλίο ή ένα μουσειακό αντικείμενο.

# Μεταδεδομένα - Είναι παντού

4

**Συναντώνται καθημερινά, π.χ., σε ένα σύστημα αρχείων:**

- Όνομα αρχείου
- Τύπος αρχείου
- Μέγεθος αρχείου
- Ημερομηνία αλλαγής αρχείου
- Επιπλέον σε αρχεία εικόνων
  - Φυσικές διαστάσεις εικόνας
  - Ανάλυση, μορφή αποθήκευσης, τρόπος συμπίεσης

# Μεταδεδομένα - Παραδείγματα

## Για την περιγραφή βιβλίων:

- Τίτλος
- Συγγραφέας
- Εκδότης
- Ημερομηνία έκδοσης
- ISBN
- Το πρότυπο MARC περιγράφει τέτοια μεταδεδομένα

# Μεταδεδομένα - Παραδείγματα

6

## Για την περιγραφή μουσικών κομματιών:

- Τίτλος κομματιού
- Καλλιτέχνης κομματιού
- Τίτλος δίσκου
- Είδος
- Αριθμός κομματιού
- Αριθμός δίσκου
- Έτος

# Μεταδεδομένα - Παραδείγματα

7

## Για την περιγραφή αρχείων βίντεο

- Ημερομηνία παραγωγής
- Διάρκεια
- Ανάλυση
- Κωδικοποίηση ήχου
- Κωδικοποίηση βίντεο
- Μέγεθος αρχείου
- Σύνομη περίληψη του τι είναι στο κλιπ

# Μεταδεδομένα - Σχήμα (schema)

- Τα μεταδεδομένα αποθηκεύονται σε κατάλληλο αρχείο με συγκεκριμένη δομή και σύνταξη = σχήμα (schema)
- Metadata syntax: Αναφέρεται στους κανόνες για τη δόμηση των πεδίων και των στοιχείων περιγραφής
- Ένα σχήμα μεταδεδομένων μπορεί να αποτυπωθεί σε διάφορες γλώσσες σήμανσης ή προγραμματισμού, όπως HTML, XML και RDF.



# Μεταδεδομένα - Σύνταξη & Πρότυπα

- Πολλά σχήματα χρησιμοποιούν την XML για τον καθορισμό της σύνταξης
- Τα σχήματα μεταδεδομένων αναπτύσσονται και συντηρούνται από οργανισμούς προτύπων (π.χ. ISO) ή οργανισμούς που έχουν αναλάβει τη συγκεκριμένη ευθύνη π.χ. Dublin Core Metadata Initiative ή **IEEE** Learning Technology Standards Committee

# Μεταδεδομένα - Πρότυπα

10

Πρότυπο	Εφαρμογή	Περιγραφή
<u>MARC</u>	Βιβλιοθήκες	<u>MARC - M</u> Achine Readable Cataloging - Πρότυπο για την αναπαράσταση βιβλιογραφικής πληροφορίας σε μορφή αναγνώσιμη από μηχανή
<u>MEI</u>	Μουσική	<u>Music Encoding Initiative</u> Κοινότητα με σκοπό τη δημιουργία κοινής ψηφιακής αναπαράστασης της πληροφορίας σχετικής με τη μουσική
<u>MPEG-7</u>	Πολυμέσα	Το <u>MPEG-7</u> is είναι ένα πρότυπο ISO/IEC το οποίο καθορίζει ένα σύνολο πεδίων για την περιγραφή διαφόρων τύπων πολυμεσικής πληροφορίας
<u>CSDGM</u>	Γεωγραφικά δεδομένα	Content Standard for Digital Geospatial Metadata συντηρείται από την Federal Geographic Data Committee (FGDC).
<u>IEEE LOM</u>	Εκπαίδευση	<u>Learning Objects Metadata</u> - Καθορίζει τη σύνταξη και τη σημασιολογία για τα μεταδεδομένα μαθησιακών αντικειμένων

# Μεταδεδομένα - Εκπαιδευτικά αντικείμενα

11

## Για την περιγραφή εκπαιδευτικών αντικειμένων

- Δημιουργός
- Ημερομηνία δημιουργίας
- Επιστήμη / Θεματική Ενότητα
- Βαθμίδα εκπαίδευσης
- Βαθμός αλληλεπίδρασης
- Βαθμός δυσκολίας
- Τύπος αντικειμένου (προσομοίωση, παιχνίδι, άσκηση, ανάγνωσμα, παρουσίαση, βίντεο)

# Μεταδεδομένα - Πνευματικά Δικαιώματα

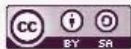
12

## CREATIVE COMMONS SIMPLIFIED

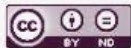
### When I see this license



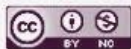
Attribution  
CC BY



Attribution-ShareAlike  
CC BY-SA



Attribution-NonCommercial  
CC BY-NC



Attribution-NonCommercial-ShareAlike  
CC BY-NC-SA



Attribution-NonCommercial-NoDerivs  
CC BY-NC-ND

### I can

share, remix, make money

share, remix, make money

share, make money

share, remix

share, remix

share

### If I ...

mention author

mention author, keep the same license

mention author, no remix

mention author, don't make money

mention author, keep the same license  
don't make money

mention author, no remix  
keep the same license  
don't make money



-this is intended as general information only, not a legal advice-

cc by-nc-nd [www.milosjanata.com](http://www.milosjanata.com)  
based on [creativecommons.org](http://creativecommons.org)

# Ελεύθερο & Ανοιχτού κώδικα λογισμικό

για την ανάπτυξη ψηφιακών βιβλιοθηκών και αποθετηρίων

# ΕΛ/ΛΑΚ - Ορισμοί

- Ο όρος «**Ελεύθερο Λογισμικό / Λογισμικό Ανοικτού Κώδικα**» (**ΕΛ/ΛΑΚ**) ομαδοποιεί το Ελεύθερο Λογισμικό (ΕΛ) και το Λογισμικό Ανοικτού Κώδικα (ΛΑΚ)
- Το ΛΑΚ είναι λογισμικό για το οποίο ο πηγαίος κώδικας είναι ελεύθερα και δημόσια διαθέσιμος, παρόλο που οι συμφωνίες άδειας ποικίλουν ως προς το τι επιτρέπεται να κάνει κανείς με αυτόν τον κώδικα
- Το Ελεύθερο λογισμικό παρέχει στους χρήστες την ελευθερία να εκτελούν, αντιγράφουν, διανέμουν, μελετούν, τροποποιούν και βελτιώνουν το Ελεύθερο λογισμικό

# Ελεύθερο λογισμικό - Ελευθερίες

15

## Αναφέρεται σε 4 βασικές ελευθερίες:

- Την ελευθερία να εκτελείται το πρόγραμμα για οποιονδήποτε σκοπό (ελευθερία 0).
- Την ελευθερία να μελετάται ο τρόπος λειτουργίας του προγράμματος και να προσαρμόζεται στις εκάστοτε ανάγκες (ελευθερία 1). Η πρόσβαση στον πηγαίο κώδικα είναι προϋπόθεση για να ισχύει κάτι τέτοιο.
- Την ελευθερία για αναδιανομή αντιγράφων του προγράμματος ώστε να διαμοιράζεται με άλλους (ελευθερία 2).
- Την ελευθερία βελτίωσης του προγράμματος και της δημοσίευσης των βελτιώσεων στο ευρύ κοινό, ώστε να επωφεληθεί ολόκληρη η κοινότητα (ελευθερία 3). Η πρόσβαση στον πηγαίο κώδικα είναι προϋπόθεση για να ισχύει κάτι τέτοιο.

# ΕΛ/ΛΑΚ - Αδειοδότηση

16

## OSI-approved licenses



## GPL-Compatible Free Software Licenses

- ❑ Apache License 2.0
- ❑ BSD 3-Clause "New" or "Revised" license
- ❑ BSD 2-Clause "Simplified" or "FreeBSD" license
- ❑ GNU General Public License (GPL)
- ❑ GNU Library or "Lesser" General Public License (LGPL)
- ❑ MIT license
- ❑ Mozilla Public License 2.0
- ❑ Common Development and Distribution License
- ❑ Eclipse Public License

- ❑ Apache License 2.0
- ❑ FreeBSD license
- ❑ Modified BSD license
- ❑ GNU General Public License (GPL)
- ❑ GNU Library or "Lesser" General Public License (LGPL)
- ❑ MIT licenses
- ❑ Mozilla Public License (MPL) v 2.0

<https://www.gnu.org/licenses/license-list.html#GPLCompatibleLicenses>

<http://opensource.org/licenses>



# D SPACE - Εισαγωγικά

17

- Το χρησιμοποιούν περισσότεροι από 1000 οργανισμούς
- Η πιο κοινή χρήση είναι για ερευνητικές βιβλιοθήκες και αποθετήρια
- Έχει περισσότερους από 100 contributors ανά τον κόσμο (αναπτύσσουν κώδικα, διορθώνουν σφάλματα κ.τ.λ.)
- Το λογισμικό D SPACE διαχειρίζεται μια μικρότερη ομάδα εθελοντών προγραμματιστών και μηχανικών (committers) οι οποίοι συνεργάζονται για να αναπτύξουν νέες εκδόσεις να ενσωματώσουν νέα χαρακτηριστικά και να διορθώσουν σφάλματα που αναφέρουν οι contributors



**D SPACE**

# D SPACE - Χαρακτηριστικά

- **Προσαρμογή διεπαφής**
- **Προσαρμογή μεταδεδομένων** – Η προεπιλεγμένη μορφή μεταδεδομένων είναι το Dublin Core ενώ υποστηρίζει και MARC
- **Προσαρμογή αναζήτησης και πλοήγησης**
- **Συμβατότητα με πρότυπα** - Συμβατό με διάφορα πρωτόκολλα. Υποστηρίζει: OAI-PMH, OpenSearch, OpenURL, RSS, ATOM.
- **Προσαρμογή βάσης** - Επιλογή ανάμεσα σε PostgreSQL ή Oracle για τη βάση όπου γίνεται και η διαχείριση των μεταδεδομένων

# D SPACE - Δυνατότητες

19

- Αναγνωρίζει και διαχειρίζεται ένα μεγάλο αριθμό τύπων αρχείων. Τα πιο κοινά είναι PDF, Word, JPEG, MPEG, TIFF files
- Παρέχει μητρώο (registry) όπου οποιοσδήποτε μη αναγνωρίσιμος τύπος αρχείου μπορεί να καταχωρηθεί εκεί ώστε να αναγνωρίζεται στο μέλλον

# E-prints - Εισαγωγικά

- Λογισμικό για την ανάπτυξη αποθετηρίων συμβατά με ΟΑΙ
- Εύκολο και γρήγορο στην εγκατάσταση για την ανάπτυξη αποθετηρίων ανοιχτής πρόσβασης σχετικά με επιστήμες, λογοτεχνία, διδακτορικές διατριβές, πολυμέσα
- Διακριτούς ρόλους-> διαχειριστές αποθετηρίου, συνεισφέροντες, ερευνητές, τεχνικούς διαχειριστές

# E-prints - Δυνατότητες

21

- Μπορούν να δημιουργηθούν διάφοροι τύποι πλοήγησης στα αντικείμενα
- Αναζήτηση με πολύπλοκα κριτήρια
- Η παρουσίαση ενός αντικειμένου μπορεί να περιέχει
  - Εικόνα, βίντεο, pdf thumbnails
  - Περισσότερες από 50 γραμμές μεταδεδομένων για το Google Scholar
  - Κουμπί αίτησης (email στο δημιουργό) για κάποιο αντικείμενο που έχει κλειστή πρόσβαση

# E-prints - Πεδία περιγραφής αντικειμένων

22

Demo Repository



Home About Browse by Year Browse by Subject  
Logged in as Dr Leslie Carr | [Manage deposits](#) | [Profile](#) | [Saved searches](#) | [Review](#) | [Admin](#) | [Logout](#)

## View Item: Observations on the Use of EPrints

You are both a [depositor](#) and an [editor](#) of this item. This is a depositor screen.  
This item is in review. It will not appear in the repository until it has been checked by an editor.

Details Summary Actions Export History

Item ID: 111  
Revision: 12  
Item Status: Under Review  
Depositing User: [Dr Leslie Carr](#)  
Directory: disk0/00/00/01/11  
Last Modified: 18 Apr 2007 09:42  
Last Status Change: 18 Apr 2007 09:42  
Item Type: Article  
Metadata Visibility: Always Show

### Links to files:



### Creators:

Creators	Email
Carr, Leslie A.	lac@ecs.soton.ac.uk
Hammad, Stevan	hammad@ecs.soton.ac.uk
Oppenheim, Charles	

Title: Observations on the Use of EPrints  
Status: Published  
Subjects: [K Law > K Law \(General\)](#)  
Divisions: Faculty of Engineering, Science and Mathematics > School of Chemistry  
Full Text Status: Restricted  
Abstract: Demo abstract  
Journal or Publication Title: Early Years: an international journal of research and development  
Publisher: Taylor & Francis  
Refereed: Yes  
ISSN: 0957-5146

Format: MS Word. Language: English. Visible to: Anyone.

Document: MS Word  
• [almaquestionnaire.doc](#)

### Unspecified fields

Date Deposited, Later Version of, Commentary on, Replaced by, Contact Email Address, Latitude, Longitude, Corporate Creators, Monograph Type, Presentation Type, Uncontrolled Keywords, Additional Information, Comments and Suggestions, Date, Date Type, Series Name, Volume, Number, Place of Publication, Page Range, Number of Pages, Event Title, Event Location, Event Dates, Event Type, Identification Number, Patent Applicant, Institution, Department, Thesis Type, ISBN, Title of Book, Editors, Official URL, Related URLs, References, Funders, Projects, Media of Output, Exhibitors, Number of Pieces, Composition Type, Producers, Conductors, Lyricists, Accompaniment, Type of Data, Pedagogic Type, Completion Time, Task Purpose, Skill Areas, Copyright Holders, Learning Level

# E-prints - Αρχεία

23

## Διάφοροι τύποι αντικειμένων

- Έγγραφα ( Άρθρα, αναφορές, πατέντες, διδακτορικές διατριβές )
- Πολυμέσα (Αρχεία βίντεο, αρχεία ήχου, εικόνες )
- Επιστημονικά δεδομένα & πειράματα
- Τέχνη
- Εκπαιδευτικούς πόρους
- Εκθέσεις

Demo Repository

eprints

Home About Browse by Year Browse by Subject

Logged in as Dr Leslie Carr | [Manage deposits](#) | [Profile](#) | [Saved searches](#) | [Review](#) | [Admin](#) | [Logout](#)

Search

Edit item: Article #112

You are both a [depositor](#) and an [editor](#) of this item. This is a depositor screen.

Type → Upload → Details → Subjects → Deposit

Save for Later Next >

### ★ Item Type

- Article**  
An article in a journal, magazine, newspaper. Not necessarily peer-reviewed. May be an electronic-only medium, such as an online journal or news website.
- Book Section**  
A chapter or section in a book.
- Monograph**  
A monograph. This may be a technical report, project report, documentation, manual, working paper or discussion paper.
- Conference or Workshop Item**  
A paper, poster, speech, lecture or presentation given at a conference, workshop or other event. If the conference item has been published in a journal or book then please use "Book Section" or "Article" instead.
- Book**  
A book or a conference volume.
- Thesis**  
A thesis or dissertation.
- Patent**  
A published patent. Do not include as yet unpublished patent applications.
- Artefact**  
An artist's artefact or work product.
- Show/Exhibition**  
An artist's exhibition or site specific performance-based deposit.
- Composition**  
A musical composition.
- Performance**  
Performance of a musical event.
- Image**  
A digital photograph or visual image.
- Video**  
A digital video.
- Audio**  
A sound recording.
- Dataset**  
A bounded collection of quantitative data (e.g. spreadsheet or XML data file).
- Experiment**  
Experimental data with intermediate analyses and summary results.
- Teaching Resource**  
Lecture notes, exercises, exam papers or course syllabuses.
- Other**  
Something within the scope of the repository, but not covered by the other categories.

# E-prints - Εξειδικευμένα μεταδεδομένα

24



Journal autocompletion information is derived from the [RoMEO](#) database which is compiled by [SHERPA](#) and has been modified for use here.

## Publication Details

- Υποστήριξη για επιστημονικά περιοδικά
- Όνομα Περιοδικού + ISSN + Εκδοτικός Οίκος

★ Refereed:	<input type="radio"/> Yes, this version has been refereed. <input type="radio"/> No, this version has not been refereed.	?
★ Status:	<input type="radio"/> Published <input type="radio"/> In Press <input type="radio"/> Submitted <input type="radio"/> Unpublished	?
★ Journal or Publication Title:	<input type="text" value="Intelligent"/>	?
ISSN:	"Chemometrics and Intelligent Laboratory Systems" published by "Elsevier" (a Green publisher)	?
Publisher:	"International Journal of Intelligent Systems Technologies and Applications (IJISTA)" published by "Inderscience" (a Green publisher) ISSN: 1740-8865	?
Official URL:	"IEEE Transactions on Intelligent Transportation Systems" published by "Institute of Electrical and Electronics Engineers" (a Green publisher)	?
Volume:		?
Number:	"Intelligent Data Analysis" published by "IOS Press" (a Green publisher)	?



# Fedora Commons - Εισαγωγικά

25

- Το **Fedora** (ή **Flexible Extensible Digital Object Repository Architecture**) είναι ένα σύστημα διαχείρισης ψηφίδων για τη δημιουργία αποθετηρίων, ψηφιακών αρχειοθηκών και ψηφιακών βιβλιοθηκών.
- Περιλαμβάνει εισαγωγή και εξαγωγή ψηφιακών αντικειμένων σε διάφορες μορφές XML αρχείων επιτρέποντας έτσι τη διαλειτουργικότητα με πολλές άλλες εφαρμογές



# Fedora Commons - Δυνατότητες

26

- Το **Digital Object Model** του FEDORA επιτρέπει καλή διαχείριση των μεταδεδομένων και του ψηφιακού περιεχομένου ανεξαρτήτως του τύπου των αντικειμένου
- Έχει ευέλικτο σύστημα για εύκολη συσχέτιση με εξωτερικά αντικείμενα και άλλα αποθετήρια
- Παρέχει χαρακτηριστικά που υποστηρίζουν την ψηφιακή διατήρηση των αντικειμένων
  - Ένα από αυτά είναι ότι υποστηρίζει όλους τους τύπους ψηφιακού περιεχομένου

# Fedora Commons - Εξωτερικά αντικείμενα

27

**Control Groups:** τα αντικείμενα διακρίνονται/ανήκουν σε 4 groups

Δύο groups για περιπτώσεις όπου το περιεχόμενο βρίσκεται σε απομακρυσμένη τοποθεσία

- ***Externally Referenced Content*** - Το αρχείο XML του αντικειμένου μπορεί να ανακτήσει το αντικείμενο και όλη τη σχετική πληροφορία από την απομακρυσμένη τοποθεσία. Φαίνεται ως ο άμεσος εξυπηρετητής
- ***Redirect Referenced Content*** - Εκτελεί ανακατεύθυνση σε απομακρυσμένη τοποθεσία

# Fedora Commons - Οργανισμοί & Χρήσεις

28

## **Οργανισμοί που έχουν υλοποιήσει το Fedora Commons**

- ❑ broadcasting and media
- ❑ consortia
- ❑ corporations
- ❑ government agencies
- ❑ IT-related institutions
- ❑ medical centers and libraries
- ❑ museums and cultural organizations
- ❑ national and public libraries and archives
- ❑ professional societies
- ❑ publishing
- ❑ research groups and projects
- ❑ semantic and virtual library projects
- ❑ university libraries and archives

# Αναφορές - Πηγές

29

## Μεταδεδομένα

- Metadata, wikipedia <http://en.wikipedia.org/wiki/Metadata>
- Metadata Training Resources, Dublin Core Metadata Initiative <http://dublincore.org/resources/training/>
- Understanding Metadata, National Information Standards Organization <http://www.niso.org/publications/press/UnderstandingMetadata.pdf>
- XSD wikipedia <http://en.wikipedia.org/wiki/XSD>

# Αναφορές - Πηγές

30

## **ΕΛ/ΛΑΚ και αδειοδότηση**

- Open Source Initiative <http://opensource.org/>
- Free Software – GNU <http://www.gnu.org/>
- Εταιρία ΕΛ/ΛΑΚ <http://www.ellak.gr/>

## **Ανοικτές άδειες περιεχομένου**

Creative Commons <http://www.creativecommons.gr/>

## **Λογισμικά**

- Eprints <http://www.eprints.org/>
- Fedora Commons <http://fedorarepository.org/>
- Dspace <http://www.dspace.org/>

# Υποστηρικτικό υλικό

31



## Metadata: Organizing and Discovering Information

Metadata is an unsung hero of the modern world, the plumbing that makes the information age possible. This course describes how Metadata is used as an information tool for the Web, for databases, and for the software and computing applications around us.



### About the Course

If you use nearly any digital technology, you make use of metadata. Use an ATM today? You interacted with metadata about your account. Searched for songs in iTunes or Spotify? You used metadata about those songs. We use and even create metadata constantly, but we rarely realize it. Metadata -- or data about data -- describes real and digital objects, so that those objects may be organized now and found later.

Metadata is a tool that enables the information age functions performed by humans as well as those performed by computers. Metadata is important to many fields, particularly Computer Science; but this course is not purely a Computer Science course. This course approaches Metadata from the perspective of Information Science, which is a broad interdisciplinary field that studies how people create and manage information.

### Sessions

Jul 14th 2014

Starts in 2 months

Earn a Verified Certificate

### Course at a Glance

- 8 weeks
- 4-6 hours of work / week
- English
- English subtitles

[evergreen-ils.org/understanding-open-source-webinar](http://evergreen-ils.org/understanding-open-source-webinar)

[coursera.org/course/metadata](https://coursera.org/course/metadata)

# Σας ευχαριστώ πολύ

## Ερωτήσεις;