



ΆΜΠΕΛ ΜΠΑΣΑ

CS121105@uniwa.gr

ΜΕΤΑΔΕΔΟΜΈΝΑ ΣΤΟ ΣΗΜΑΣΙΟΛΟΓΙΚΌ ΙΣΤΌ

ΕΡΓΑΣΊΑ ΣΤΟ ΜΆΘΗΜΑ: **ΑΝΆΚΤΗΣΗ ΠΛΗΡΟΦΟΡΊΑΣ ΚΑΙ ΕΞΌΡΥΞΗΣ ΔΕΔΟΜΈΝΩΝ**
ΚΑΘΗΓΉΤΡΙΑ: **ΕΛΈΝΗ ΓΑΛΊΩΤΟΥ**

ΕΙΣΑΓΩΓΙΚΑ

- Η τεράστια ανάπτυξη του παγκόσμιου ιστού και κατ' επέκταση του σημασιολογικού ιστού έχει δημιουργήσει την ανάγκη για μεθοδευμένη αναζήτηση ψηφιακών δεδομένων. Πολλές μηχανές αναζήτησης προσφέρουν την δυνατότητα εύρεσης με σχετικά κριτήρια αποτελεσμάτων. Οι μηχανές αναζήτησης αναλαμβάνουν το δύσκολο έργο της εύρεσης των κειμένων ή της πληροφορίας χρησιμοποιώντας αλγόριθμους αναζήτησης αλλά και δεδομένα που κατατοπίζουν αυτή την αναζήτηση.

ΜΕΤΑΔΕΔΟΜΈΝΑ

- Τα μεταδεδομένα είναι δεδομένα κατανοητά από τους υπολογιστές που περιγράφουν, εξηγούν, εντοπίζουν ή διευκολύνουν την πρόσβαση στους πληροφοριακούς πόρους και εμπεριέχονται συνήθως σε ένα XML αρχείο ως διαδεδομένος τύπος εγγράφου στον σημασιολογικό ιστό. Είναι αλλιώς «Ηλεκτρονικές πηγές πληροφορίας» που εξηγούν τα δεδομένα και τη σημασία τους.
- Παράλληλα χρησιμοποιούνται ως πληροφορίες κατατόπισης και γεωγραφικής σημασίας εμπεριέχοντας πληροφορίες καταλόγων, εφαρμογές χαρτογράφησης, μοντέλα δεδομένων αλλά και σχετικές ιστοσελίδες.
- Ο ευρέως γνωστός όρος στα Αγγλικά είναι «Metadata» που χρησιμοποιείται πλέον στον σημασιολογικό ιστό. Η λέξη προέρχεται από την ελληνική λέξη «μετά» και την λατινική «data» (πληθυντικός του datum=δεδομένο).

ΤΙ ΕΊΝΑΙ ΤΑ ΜΕΤΑΔΕΔΟΜΈΝΑ ΚΑΙ ΣΕ ΤΙ ΧΡΗΣΙΜΕΎΟΥΝ ;

- Τα μεταδεδομένα είναι η δομημένη γνώση που έχουμε για την πληροφορία που συλλέγουμε.
- Ένας σημαντικός λόγος για τη δημιουργία των μεταδεδομένων είναι ο τεράστιος όγκος δεδομένων που έχει δημιουργηθεί και το πόσο δύσκολη είναι η εύρεση κάποιου αντικειμένου το οποίο δεν γνωρίζουμε πού βρίσκεται. Γι' αυτόν τον λόγο δημιουργήθηκαν, για να διευκολύνουν την εύρεση, εντοπισμό και ταυτοποίηση των πόρων ή οντοτήτων (εξυπηρετητές ή βιβλιοθήκες) που περιέχουν τις πληροφορίες τις οποίες αναζητούμε. Τα μεταδεδομένα δεν είναι κάτι καινούριο, όμως υπάρχουν από τότε που ο άνθρωπος άρχισε να οργανώνεται.

ΚΙΝΔΥΝΟΙ ΚΑΙ ΟΡΘΗ ΧΡΗΣΗ ΜΕΤΑΔΕΔΟΜΕΝΩΝ

- Τα μεταδεδομένα, αν και πολύ χρήσιμα σε όλο το φάσμα της ψηφιακής πληροφορίας, μπορεί λόγω της ευαίσθητης φύσης τους να αποτελέσουν σοβαρή απειλή της ιδιωτικότητας ή της ανωνυμίας.
- Τίθεται σοβαρό ζήτημα για την μη ορθή χρήση τους και την εκμετάλλευση από φορείς ή πρόσωπα που εκμεταλλεύονται αυτού του είδους τις πληροφορίες.
- Το ζήτημα έχει μεγάλη βαρύτητα μετά από τις πρόσφατες ανακαλύψεις πως υπηρεσίες παρακολουθούν ανά τον κόσμο και συγκρατούν στοιχεία μέσα από τα μεταδεδομένα των αναζητήσεων ή της επισκεψιμότητας σε ιστοτόπους.

ΛΕΙΤΟΥΡΓΙΕΣ ΜΕΤΑΔΕΔΟΜΕΝΩΝ

Τα μεταδεδομένα έχουν τρεις βασικές λειτουργίες :

- Αναπαριστούν πληροφοριακές οντότητες στα συστήματα πρόσβασης στην πληροφόρηση.
- Εξυπηρετούν την δόμηση συστήματος.
- Εμφανίζουν συγκεκριμένα αποτελέσματα έτσι ώστε να είναι κατανοητά και ευανάγνωστα.

ΤΥΠΟΙ ΜΕΤΑΔΕΔΟΜΕΝΩΝ

Υπάρχουν 4 βασικοί τύποι μεταδεδομένων:

- Περιγραφικά (Descriptive Metadata)
- Διαρθρωτικά (Structural Metadata)
- Διοικητικά (Administrative Metadata)
- Γεωγραφικά (Geospatial Metadata)

ΠΕΡΙΓΡΑΦΙΚΑ ΜΕΤΑΔΕΔΟΜΈΝΑ (1)

- Είναι η πλήρης περιγραφή ενός αντικειμένου πχ : μια εικόνα, ένα κείμενο, μια εφαρμογή, αλλά και εξειδικευμένη ανάλυση. Εδώ ανήκει όλη εκείνη η πληροφορία που περιγράφει την πηγή με σκοπό την ανάσυρση πληροφορίας ή την ταυτοποίησή της.
- Οποιαδήποτε πληροφορία χρησιμοποιείται για την περιγραφή ψηφιακών αντικειμένων (με την ευρύτερη έννοια) αποτελεί μέρος των περιγραφικών μεταδεδομένων. Μια τυποποιημένη μορφή πληροφορίας εξυπηρετεί την εύρεση και ανάσυρση, βοηθάει την εμφάνιση και απλοποιεί την συντήρησή της.
- Στα περιγραφικά μεταδεδομένα υπάρχουν όλα εκείνα τα στοιχεία που περιγράφουν οποιαδήποτε αρχεία ή συλλογή αρχείων με σκοπό την καταγραφή ή την αποθήκευσή τους. Κάνει την οργάνωση και εύρεση των στοιχείων εύκολη και προσιτή. Μπορεί να γίνει εύκολα σύγκριση μεταξύ όμοιων ή παρόμοιων αντικειμένων. Οι λεπτομέρειες ενός αντικειμένου ή συλλογής αντικειμένων καθώς και πόρων χωρίζεται σε επίπεδα.

ΠΕΡΙΓΡΑΦΙΚΑ ΜΕΤΑΔΕΔΟΜΕΝΑ (2)

- `<dmdSec ID="dmd001"> <mdRef LOCTYPE="URN"
MIMETYPE="application/xml" MDTYPE="EAD" LABEL="Berol Collection
Finding Aid">urn:x-nyu:fales1735</mdRef> </dmdSec`
- Το `<mdRef>` στοιχείο αυτού του `<dmdSec>` εμπεριέχει τέσσερις ιδιότητες. Το `LOCTYPE` καθορίζει τον τύπο του εντοπιστή που εμπεριέχεται στο σώμα του στοιχείου αυτού. Θεωρούνται έγκυρες τιμές για το πεδίο 'LOCTYPE' η εξής 'URN,' 'URL,' 'PURL,' 'HANDLE,' 'DOI,' 'OTHER.'. Η ιδιότητα `MIMETYPE` προσδιορίζει το είδος MIME(Multipurpose Internet Mail Extension) το οποίο επιτρέπει την απόστολή αρχείων που δεν ανήκουν στην κατηγορία του κειμένου όπως εικόνες κλπ. Το `MDTYPE` με την σειρά του δείχνει σε ποια μορφή μεταδεδομένων γίνεται αναφορά.

ΠΕΡΙΓΡΑΦΙΚΑ ΜΕΤΑΔΕΔΟΜΕΝΑ (3)

- Έγκυρες τιμές για το MDTYPE θεωρούνται MARC, EAD, VRA (VRA Core), Dc (Dublin Core), NISOIMG (NISO Technical Metadata for Digital Still Images), LC-AV (Library of Congress Audiovisual Metadata), TEIHDR (TEI Header), DDI (Data Documentation Initiative), FGDC (Federal Geographic Data Committee Metadata Standard [FGDC-STD-001-1998]), OTHER. LABEL παρέχει έναν μηχανισμό για την περιγραφή των μεταδεδομένων στην επισκοπή ενός METS αρχείου, για παράδειγμα, σ' έναν 'Πίνακα Περιεχομένων' απεικονίζει ένα METS αρχείο.

ΠΕΡΙΓΡΑΦΙΚΑ ΜΕΤΑΔΕΔΟΜΈΝΑ (4)

Εσωτερικά Περιγραφικά Μεταδεδομένα(mdWrap)

- Ένα mdWrap στοιχείο παρέχει ένα περιτύλιγμα γύρω από μεταδεδομένα ενσωματωμένα μέσα σε ένα έγγραφο METS. Τέτοιου τύπου μεταδεδομένα μπορεί να είναι σε μια από τις δυο μορφές:
 - Κωδικοποιημένα σε XML μορφή με το XML το ίδιο να προσδιορίζει ότι ανήκει σε μια περιοχή ονομάτων εκτός της περιοχής του ονόματος του αρχείου METS.
 - Η οποιαδήποτε αυθαίρετη δυαδική μορφή ή μορφή κειμένου, υπό την προϋπόθεση ότι τα μεταδεδομένα είναι κωδικοποιημένα σε μορφή Base64 κωδικοποιημένα και ‘τυλιγμένα’ σε ένα <binData> mdWrap στοιχείο. Το παράδειγμα που ακολουθεί επιδεικνύει τη χρήση του στοιχείου mdWrap :

ΠΕΡΙΓΡΑΦΙΚΑ ΜΕΤΑΔΕΔΟΜΈΝΑ (5)

```
<dmdSec ID="dmd002">
```

```
<mdWrap MIMETYPE="text/xml" MDTYPE="DC" LABEL="Dublin Core Metadata">
```

```
<dc:title>Alice'sAdventuresinWonderland</dc:title> <dc:creator>Lewis  
Carroll</dc:creator>
```

```
<dc:date>between1872and1890</dc:date>
```

```
<dc:publisher>McCloughlinBrothers</dc:publisher>
```

```
<dc:type>text</dc:type> </mdWrap> </dmdSec>
```

```
<dmdSec ID="dmd003">
```

```
<mdWrap MIMETYPE="application/marc" MDTYPE="MARC" LABEL="OPAC Record">
```

```
<binData>MDI0ODdjam0gIDlyMDA1ODkgYSA0NU0wMDAxMDA...(etc.)
```

```
</binData>
```

```
</mdWrap>
```

```
</dmdSec>
```


ΔΙΑΡΘΡΩΤΙΚΑ ΜΕΤΑΔΕΔΟΜΈΝΑ

- Γνωστά και ως δομικά μεταδεδομένα, είναι εκείνη η ομαδοποιημένη πληροφορία που χρησιμοποιείται για να δομήσει ένα XML αρχείο που περιγράφει τα διανοητικά στοιχεία ενός ψηφιακού αντικειμένου.
- Περιλαμβάνουν χρήσιμες πληροφορίες σχετικά με την διάταξη μιας σελίδας. Αυτού του είδους τα μεταδεδομένα είναι χρήσιμα σε προηγμένες αναζητήσεις και εξελιγμένες ανακτήσεις πληροφορίας.
- Παρέχουν με ακρίβεια όλη εκείνη την πληροφορία που χρειάζεται για να χρησιμοποιηθούν όλα τα σχετικά αρχεία ή οντότητες προκειμένου να έρθουν σε επαφή με τον χρήστη.

ΔΙΟΙΚΗΤΙΚΑ ΜΕΤΑΔΕΔΟΜΈΝΑ

- Περιλαμβάνουν πληροφορίες για το αν ένας χώρος αποθήκευσης δεδομένων μπορεί να χρησιμοποιεί ή να διαχειρίζεται έναν πόρο που ανήκει σε έναν οργανισμό ή εταιρεία αλλά και πιθανόν χρεώσεις που εφαρμόζονται στην κάθε περίπτωση.
- Μπορεί να περιλαμβάνουν πληροφορίες σχετικά με δικαιώματα αναπαραγωγής ή τα κριτήρια επιλογής για κάποιο μέσο εγγραφής ή της εφαρμογή κάποιας πολιτικής αρχειοθέτησης ψηφιακού περιεχομένου.
- Είναι το κομμάτι στο οποίο γίνεται κατ' εξοχήν η διαχείριση των δικαιωμάτων και κατά συνέπεια η διαφύλαξη μεταδεδομένων. Επομένως τα διοικητικά μεταδεδομένα συχνά ορίζονται τοπικά και σπανίως εφαρμόζεται κοινό πρότυπο.

ΔΙΟΙΚΗΤΙΚΑ ΜΕΤΑΔΕΔΟΜΈΝΑ (2)

Υπάρχουν τέσσερις μορφές που προβλέπονται σε ένα έγγραφο METS:

- Τεχνικές μεταδεδομένων (πληροφορίες σχετικά με τη δημιουργία αρχείων, τη μορφή και τη χρήση χαρακτηριστικών).
- Μεταδεδομένα Πνευματικής Ιδιοκτησίας και πληροφορίες άδειας χρήσης των αρχείων.
- Μεταδεδομένα που περιγράφουν την αρχική πηγή από το οποίο μετατράπηκε σε ψηφιακή μορφή κάποιο αναλογικό υλικό.
- Η ψηφιακή προέλευση μεταδεδομένων όπως σχέση πηγής προορισμού μεταξύ αρχείων συμπεριλαμβανομένου και του κύριου υλικού ή τυχόν μετατροπές που έχουν υποστεί καθώς και την τρέχουσα ενσάρκωσή του ως ψηφιακό αντικείμενο βιβλιοθήκης).

ΔΙΟΙΚΗΤΙΚΑ ΜΕΤΑΔΕΔΟΜΈΝΑ (3)

- Το στοιχείο `< amdSec>` απασχολεί το ίδιο περιεχόμενο όπως `< dmdSec>` :

Μπορεί να περιέχει το `<mdref>` με σκοπό να δείξει σε εξωτερικά διοικητικά μεταδεδομένα, ή ένα `<mdWrap>` στοιχείο όταν ενσωματώνονται διοικητικά μεταδεδομένα στο ίδιο το METS αρχείο ή και τα δύο. Πολλαπλά `<amdSec>` στοιχεία μπορεί να βρίσκονται σε ένα τέτοιο αρχείο, επομένως το κάθε ένα από αυτά έχει ένα αναγνωριστικό ID χαρακτηριστικό με σκοπό την διασύνδεση των στοιχείων εντός του αρχείου, όπως διαχωρισμοί που υπάρχουν εντός του χάρτη δομής ή αρχεία(`<file>`) στοιχεία . Μπορεί επίσης να περιέχουν `<amdSec>` στοιχεία τα οποία περιέχουν τεχνικά μεταδεδομένα που αφορούν την προετοιμασία ενός αρχείου.

ΔΙΟΙΚΗΤΙΚΑ ΜΕΤΑΔΕΔΟΜΈΝΑ

(4)

- Τα στοιχεία:
<techMD>, <rightsMD>, <sourceMD> και <digiprovMD>,

Χρησιμοποιούν το ίδιο μοντέλο περιεχομένου, όπως <dmsec> τα οποία μπορεί να περιέχουν ένα στοιχείο <mdWrap> και χρησιμοποιούνται στην ενσωμάτωση των διοικητικών μεταδεδομένων μέσα σε ένα έγγραφο. Οι πολλαπλές εμφανίσεις των στοιχείων μέσα σε ένα έγγραφο πρέπει να φέρουν ένα χαρακτηριστικό ID έτσι ώστε άλλα στοιχεία εντός εγγράφου (όπως διαιρέσεις εντός του δομικού χάρτη ή <file> στοιχείων να μπορούν να συνδέονται με το <amdSec> υποστοιχείο που το χαρακτηρίζουν. Θα μπορούσαμε να δώσουμε ένα παράδειγμα όπου εμπεριέχεται ένα <techMD> στοιχείο το οποίο αφορά την προετοιμασία ενός αρχείου.

ΔΙΟΙΚΗΤΙΚΑ ΜΕΤΑΔΕΔΟΜΈΝΑ

(5)

```
<techMD ID="AMD001">
```

```
<mdWrap MIMETYPE="text/xml" MDTYPE="NISOIMG" LABEL="NISO Img. Data">
```

```
<xmlData>
```

```
<niso:MIMETYPE>image/tiff</niso:MIMETYPE>
```

```
<niso:Compression>LZW</niso:Compression>
```

```
<niso:PhotometricInterpretation>8</niso:PhotometricInterpretation>
```

```
<niso:Orientation>1</niso:Orientation>
```

```
<niso:ScanningAgency>NYU Press</niso:ScanningAgency>
```

```
</xmlData>
```

```
</mdWrap>
```

```
</techMD>
```

Ένα στοιχείο<file> εντός του <fileGrp> μπορεί να λειτουργεί ως αναγνωριστικό αυτών των διοικητικών μεταδεδομένων που αφορούν το αρχείο που προσδιορίζει χρησιμοποιώντας ένα χαρακτηριστικό ADMID που επισημαίνει κάτι σε αυτό το <techMD> στοιχείο:

```
<file ID="FILE001" ADMID="AMD001">
```

```
<FLocat LOCTYPE="URL">http://dlib.nyu.edu/press/testimg.tif</FLocat>
```

```
</file>
```

ΓΕΩΓΡΑΦΙΚΑ ΜΕΤΑΔΕΔΟΜΈΝΑ

- Η συσχέτιση γεωγραφικής πληροφορίας με αρχεία και πόρους σημασιολογικού ιστού ανέδειξε ένα νέο τύπο μεταδεδομένων. Πρόκειται για μια καταγραφή της γεωγραφικής τοποθεσίας και την ανάθεση συγκεκριμένων λειτουργιών ή αρχείων που ταυτίζονται με τη περιοχή.
- Η εύρεση γεωγραφικής πληροφορίας είναι η μέθοδος κατά την οποία εξάγεται η γεωγραφική πληροφορία που αφορά τον διαδικτυακό πόρο. Είναι χρήσιμο και απαραίτητο τα γεωγραφικά μεταδεδομένα να λάβουν πληροφορίες σε σχέση με την μέθοδο που έχει εφαρμοστεί για να παραχθούν τα συγκεκριμένα δεδομένα.
- Συγκρίνουμε δύο παραδείγματα μεταξύ τους: αυτό του δικτύου GSM (κινητής τηλεφωνίας) παράγει ακρίβεια μερικών εκατοντάδων μέτρων ενώ μέσω GPS παράγεται ακρίβεια μερικών μέτρων

ΤΜΗΜΑ ΑΡΧΕΙΟΥ

- Το τμήμα αρχείου (<fileSec>) εμπεριέχει ένα ή περισσότερα <fileGrp> στοιχεία τα οποία χρησιμεύουν στην ομαδοποίηση σχετικών αρχείων. Ένα <fileGrp> παραθέτει όλα τα αρχεία που περιλαμβάνουν μια ενιαία ηλεκτρονική έκδοση του ψηφιακού αντικειμένου βιβλιοθήκης. Για παράδειγμα, μπορεί να υπάρχουν ξεχωριστά στοιχεία <fileGrp> για τις μικρογραφίες, ή για τις εικόνες αρχείου, ή εκδόσεις pdf , ή κωδικοποιημένες εκδόσεις κειμένου ΤΕΙ κλπ.

ΔΙΑΡΘΡΩΤΙΚΟΣ ΧΑΡΤΗΣ

- Το τμήμα δομικού χάρτη του εγγράφου ορίζει μια ιεραρχική δομή που μπορεί να παρουσιάζεται στους χρήστες της ψηφιακής βιβλιοθήκης αντικειμένου που θα τους επιτρέψει να περιηγηθούν μέσα σε αυτό.
- Το στοιχείο `<structMap>` κωδικοποιεί αυτή την ιεράρχηση ως ένθετη σειρά στοιχείων `<div>`. Κάθε `<div>` φέρει χαρακτηριστικά για τα στοιχεία και περί ποιας διαίρεσης πρόκειται, και μπορεί επίσης να περιέχει πολλαπλούς δείκτες (`<mptr>`) και το αρχείο δείκτη (`<fptr>`) στοιχείο για τον προσδιορισμό του περιεχομένου αντίστοιχο με εκείνο `<div>`.
- Οι δείκτες προσδιορίζουν χωριστά έγγραφα καθώς περιέχουν τις σχετικές πληροφορίες αρχείων για το `<div>` που τα περιέχουν.
- Αυτό μπορεί να είναι χρήσιμο όταν κωδικοποιούνται μεγάλες συλλογές υλικού όπως για παράδειγμα όταν πρέπει να αποθηκευτεί η πορεία ενός περιοδικού εντύπου.

ΔΙΑΡΘΡΩΤΙΚΟΙ ΣΥΝΔΕΣΜΟΙ

- Το τμήμα των διαρθρωτικών συνδέσμων είναι η απλούστερη ενότητα από τις υπόλοιπες ενότητες, και περιέχει μόνο ένα στοιχείο `<smLink>` (αν το στοιχείο αυτό μπορεί να επαναληφθεί).
- Ο τομέας των διαρθρωτικών της δομής του υπερκειμένου, προσδιορίζει αν επιτρέπει να καταγραφεί η ύπαρξη των υπερσυνδέσμων μεταξύ των στοιχείων εντός του διαρθρωτικού χάρτη, συνήθως `<div>` στοιχείων.
- Αυτή είναι μια χρήσιμη δυνατότητα εάν θέλουμε να χρησιμοποιήσουμε METS κωδικοποίηση για την αρχειοθέτηση ιστοσελίδων και επιθυμούμε να διατηρήσουμε ένα αρχείο της δομής του υπερκειμένου, ή ιστοσελίδες ξεχωριστά από τα HTML αρχεία κάθε αυτού.

ΤΜΗΜΑ ΣΥΜΠΕΡΙΦΟΡΑΣ

- Ένα τμήμα της συμπεριφοράς μπορεί να χρησιμοποιηθεί για να συνδέσει την εκτελέσιμη συμπεριφορά με το περιεχόμενο στο αντικείμενο METS.
- Ένα τμήμα της συμπεριφοράς περιέχει ένα ή περισσότερα στοιχεία <behavior>, καθένα από τα οποία έχει ένα στοιχείο ορισμού διασύνδεσης που αντιπροσωπεύεται από ένα συγκεκριμένο τμήμα της συμπεριφοράς.
- Ένα <behavior> έχει επίσης ένα <mechanism> στοιχείο, το οποίο χρησιμοποιείται για να υποδείξει μια λειτουργική μονάδα εκτελέσιμου κώδικα που υλοποιεί και διαχειρίζεται την συμπεριφορά που ορίζεται αφηρημένα από τον ορισμό διεπαφής.

ΟΝΤΟΛΟΓΙΕΣ

- Η περαιτέρω ομαδοποίηση και οργάνωση των μεταδεδομένων οδηγεί στη δημιουργία μοντέλου πληροφοριών οντολογιών.
- Μια οντολογία ορίζει συγκεκριμένους όρους που καλούνται να αντιπροσωπεύουν ή αντίστοιχα να περιγράψουν μια περιοχή γνώσης.
- Χρησιμοποιούνται από χρήστες, εφαρμογές ή βάσεις δεδομένων με απώτερο σκοπό την ικανοποίηση των αναγκών που προκύπτουν σε κάποιο τομέα όπως η ιατρική, υλοποίηση κατασκευών, οικονομική διαχείριση και πολλές άλλες πιθανές χρήσεις.
- Το αναδυόμενο μοντέλο δημιούργησε την ανάγκη ύπαρξης μιας γλώσσας που αναλύει σε βάθος και δίνει περαιτέρω δυνατότητες. Η OWL (Ontology Web Language) κάνει δυνατή την «αντικειμενοστραφή μοντελοποίηση» αλλά μπορεί να χρησιμοποιηθεί ακόμη και για μια απλή περίπτωση ορισμού των όρων.

ΤΟ ΠΡΟΤΥΠΟ METS

- Σε βιβλιοθήκες ψηφιακών αντικειμένων προκύπτει η ανάγκη για την διατήρηση των μεταδεδομένων. Για να μπορεί να υπάρξει επιτυχής διαχείριση και χρήση των ψηφιακών δεδομένων πρέπει να υπάρχει μια ορθή συλλογή των εγγράφων και άλλων φυσικών μέσων.
- Το πρότυπο αυτό είναι μια προσπάθεια για τυποποιημένη μορφή κωδικοποίησης μέσα από το XML έγγραφο. Αυτό το πρότυπο αφορά και τους τρεις τύπους μεταδεδομένων, δηλ. περιγραφικά, δομικά και διοικητικά.

ΤΟ ΠΡΟΤΥΠΟ METS (2)

Ένα METS αρχείο αποτελείται από ενότητες οι οποίες είναι επτά στο σύνολο.

- Επικεφαλίδα METS
- Περιγραφικά μεταδεδομένα
- Διοικητικά μεταδεδομένα
- Τμήμα αρχείου
- Διαρθρωτικός χάρτης
- Διαρθρωτικοί σύνδεσμοι
- Συμπεριφορά

ΤΟ ΠΡΟΤΥΠΟ ΤΕΙ

- Η Text Encoding Initiative (TEI) είναι μια διεθνής πρωτοβουλία για την υλοποίηση και κατασκευή κατευθυντήριων γραμμών για την επισήμανση της ηλεκτρονικής πηγής. Αναφερόμαστε σε κείμενα όπως μυθιστορήματα, ποίηση, θεατρικά έργα και ό,τι χρειάζεται για την υποστήριξη της έρευνας στις ανθρωπιστικές επιστήμες.
- Το πρότυπο αυτό λέει πώς να κωδικοποιήσει το κείμενο ενός έργου με σκοπό την ακριβή εύρεση αυτού. Συνήθως είναι ενσωματωμένο στην πηγή (αυθεντικό έγγραφο).
- Η επικεφαλίδα TEI είναι ένα σύνολο από κανόνες που ορίζονται μέσα από την DTD (Document Type Definition), μέσα από την SGML (Standard Generalized Markup Language). Η SGML και αυτό ένα πρότυπο κωδικοποίησης βασισμένο στην ιδέα ότι τα έγγραφα έχουν δομή και σημασιολογικά στοιχεία. Θεωρείται πως η κωδικοποίηση TEI είναι ψηφιακές εκδοχές πραγματικού κειμένου.

ΤΟ ΠΡΟΤΥΠΟ MODS

- Το σχήμα αντικειμένων περιγραφής μεταδεδομένων θεωρείται παράγωγο του προτύπου MARC (21). Έχει σκοπό είτε να φέρει συγκεκριμένα αποτελέσματα είτε να επιτρέψει την δημιουργία περιγραφής αρχικών αρχείων.
- Περιλαμβάνει ένα υποσύνολο πεδίων MARC, χρησιμοποιεί ετικέτες βασισμένες σε γλωσσικά μοντέλα παρά αριθμητικές. Παρόλο που το πρότυπο MODS μπορεί να σταθεί από μόνο του μερικές φορές έρχονται άλλα δεδομένα και συμπληρώνουν το πεδίο.
- Λόγω της ευελιξίας και της χρήσης XML το οποίο είναι ευρέως διαδεδομένο γίνεται πλούσια περιγραφή των ηλεκτρονικών πηγών αλλά και άμεση εστίαση στο ζητούμενο, επομένως παρέχει κάποια πλεονεκτήματα σε σχέση με άλλα μεταδεδομένα.

ΒΙΒΛΙΟΓΡΑΦΙΑ - ΙΣΤΟΓΡΑΦΙΑ

1. *Introduction to Metadata*, Tony Gill- Anne J. Gilliland – Maureen Whalen – Mary S. Woodley, ed. Murtha Baca.
2. *Γεωγραφικά Μεταδεδομένα. Θέματα χρήσης και αξιοποίησης στο διαδίκτυο*, Βασιλική Σ. Μπεζάτη, Εθνικό Μετσόβιο Πολυτεχνείο
3. <http://www.wdl.org/>
4. <http://www.loc.gov/standards/mets/mets-home.html>
5. <http://www.photometadata.org/>