

Αξιολόγηση προτύπων μεταδεδομένων στο πλαίσιο ψηφιακών βιβλιοθηκών οπτικοακουστικού υλικού: Παρουσίαση της εργασίας “Evaluation of Metadata Standards in the Context of Digital Audio-Visual Libraries”

Όλγα Παρασκευά

Προλογικό σημείωμα

Οι Βιβλιοθήκες και μαζί ο βιβλιοθηκονομικός κόσμος αντιμετώπισαν την τελευταία δεκαετία το πέρασμα από τη συμβατικό τεκμήριο στην ψηφιοποίηση των συλλογών. Η τελευταία επέφερε μια σειρά από προβλήματα που ζητούσαν λύση. Το ψηφιακό υλικό αποτελείται από ηλεκτρονικά τεκμήρια με τη γνωστή μορφολογική ποικιλία. Πρόκειται για κείμενα (.DOC, .TXT, .RTF, .PDF), εικόνες (.BMP, .GIF, .JPEG, .TIF, .EPS), κινούμενες εικόνες (.ANI, .FLI, .FLC), ήχο (.WAV, .MID, .SND, .AUD), βίντεο (.AVI, .MOV, .MPG, .QT), ιστοσελίδες (.HTM, .HTML, .XML, .HTMLS, .DHTML) και προγράμματα (.COM, .EXE). Για την καταλογογράφηση και την περιγραφή αυτών των τεκμηρίων ακολουθούνται παραδοσιακές μέθοδοι (AACR2, ISBDs κ.λπ), αλλά και διατάξεις δεδομένων (MARC, XML) και πρότυπα μεταδεδομένων (Dublin Core, EAD κ.ά.).

Οι σύγχρονες εξελίξεις σχετικά με την δομή και την απρόσκοπτη λειτουργία της ψηφιακής βιβλιοθήκης έχουν να κάνουν με

1. Την ψηφιοποίηση του υπάρχοντος υλικού και τη δημιουργία νέου, πρωτογενώς ψηφιακού υλικού,
2. Τη δημιουργία προτύπων κωδικοποίησης και περιγραφής των δεδομένων,
3. Την ανάπτυξη λογισμικού για την εκμετάλλευση των ψηφιακών βιβλιοθηκών.

Τα μεταδεδομένα είναι δεδομένα που περιγράφουν τις ιδιότητες ενός τεκμηρίου και υποστηρίζουν διάφορες λειτουργίες: εντοπισμό, ανακάλυψη, τεκμηρίωση, αξιολόγηση, επιλογή κ.ά.. Αυτές οι δραστηριότητες μπορούν να πραγματοποιηθούν είτε από τους τελικούς χρήστες, είτε από τους αυτοματοποιημένους αντιπροσώπους τους. Ένας πιο διαδεδομένος ορισμός είναι ο εξής: «τα μεταδεδομένα είναι δεδομένα που συνδέονται με αντικείμενα και απελευθερώνουν τους πιθανούς τους χρήστες από την ανάγκη για πλήρη και εξειδικευμένη γνώση της ύπαρξης και των χαρακτηριστικών τους». Έχει δε παρατηρηθεί ότι «τα μεταδεδομένα που είναι απαραίτητα για την επιτυχή διαχείριση και χρήση των ψηφιακών αντικειμένων είναι εκτενέστερα και εν γένει διαφορετικά από τα μεταδεδομένα που χρησιμοποιούνται στην διαχείριση συλλογών συμβατικού υλικού».¹

Στις μέρες μας το πεδίο εφαρμογής πρακτικών «ψηφιακής βιβλιοθήκης» ολοένα και διευρύνεται με τον κατάλογο των Οργανισμών που μπορούν να επωφεληθούν να συμπεριλαμβάνει από Αρχεία και Βιβλιοθήκες έως Κρατικές Υπηρεσίες και Επιχειρήσεις. Οι ψηφιακές βιβλιοθήκες οπτικοακουστικού υλικού που δημιουργείται από τηλεοπτικά κανάλια και εταιρείες παραγωγής δεν γνωρίζει μεγάλη διάδοση, αφού οι περισσότεροι τέτοιοι οργανισμοί αποθηκεύουν και αρχειοθετούν το υλικό τους σε κασέτες, με κίνδυνο την απώλεια του περιεχομένου καθώς ο παράγοντας χρόνος επηρεάζει σε μεγάλο βαθμό τις κασέτες ως μέσο αποθήκευσης. Ωστόσο, το πέρασμα από την συμβατική στην ψηφιακή

αποθήκευση είναι εφικτό τόσο τεχνολογικά, όσο και οικονομικά και παρουσιάζει τα πλεονεκτήματα των ψηφιακών βιβλιοθηκών (μεγάλη ποσότητα πληροφοριών σε ελάχιστο χώρο, άμεση και απομωνομένη πρόσβαση, μείωση κόστους, λειτουργικότητα), ενώ ευνοεί και την άμεση κοινοποίησή του.

Εισαγωγή

Αιτία του αυξανόμενου ενδιαφέροντος για τα μεταδεδομένα είναι η ανάγκη δημιουργίας διαχειριστικών μηχανισμών οι οποίοι διευκολύνουν την ταυτοποίηση, τον εντοπισμό και την κατάλληλη χρήση ψηφιακών πληροφοριακών πόρων. Οι τομείς εφαρμογής που περιλαμβάνονται είναι πολλαπλοί, ωστόσο το εξέχουσα σημασία πρόβλημα αναφέρεται στην «πληροφοριακή έκρηξη» της εποχής. Από μια πιο αισιόδοξη οπτική γωνία, η σφαίρα της πληροφορίας χαρακτηρίζεται από «πληροφοριακό πλούτο».

Το πρόβλημα οξύνεται αν σκεφτεί κανείς το ετερογενές περιβάλλον διασποράς της πληροφορίας, μέσα στο οποίο λογικά συσχετισμένη πληροφορία αποθηκεύεται σε διαφορετικές τοποθεσίες σε διαφορετικά μορφότυπα και με διαφορετικό λογισμικό και μηχανισμούς πρόσβασης.

Η εργασία «Αξιολόγηση προτύπων μεταδεδομένων στο πλαίσιο ψηφιακών οπτικοακουστικών βιβλιοθηκών» που πραγματοποιήθηκε στο Εργαστήριο Πολυμέσων του Πανεπιστημίου του Ghent² προτείνει ένα αντικειμενικό πλαίσιο για την επιλογή του κατάλληλου σχήματος μεταδεδομένων για κάθε συγκεκριμένη εφαρμογή. Αφορά στο οπτικοακουστικό υλικό που προέρχεται από κανάλια ή/και εταιρείες παραγωγής, όπως προαναφέρθηκε. Συγκεκριμένα, ορίζει ένα σύνολο κριτηρίων που μπορούν να χρησιμοποιηθούν για την αξιολόγηση και την σύγκριση των προτύπων μεταδεδομένων με στόχο την επιλογή της βέλτιστης κατά περίπτωση λύσης. Εν συνεχεία, το πλαίσιο εφαρμόζεται σε τέσσερα γνωστά πρότυπα, το Dublin Core, το MPEG-7, το P/ Meta και το SMEF. Πρόκειται για έναν επίκαιρο, αν όχι «επείγοντα» προβληματισμό, καθώς ο όγκος του πολυμεσικού υλικού που παράγεται καθημερινά αυξάνεται εκθετικά. Η εργασία παρουσιάστηκε στα πλαίσια των εργασιών του 10ου Ευρωπαϊκού Συνεδρίου για την Έρευνα και την Τεχνολογία στις Ψηφιακές Βιβλιοθήκες³ και επιλέχθηκε να παρουσιαστεί στα πλαίσια αυτού του Μεταπτυχιακού Προγράμματος Σπουδών γιατί

- τα πολυμέσα είναι μια από τις πιο πολυσυζητημένες τεχνολογίες των αρχών της δεκαετίας του '90,
- τα πολυμέσα αποτελούν το σημείο συνάντησης πέντε μεγάλων βιομηχανιών: της πληροφορικής, των τηλεπικοινωνιών, των ηλεκτρονικών εκδόσεων, της βιομηχανίας audio και video καθώς και της βιομηχανίας της τηλεόρασης και του κινηματογράφου.

Ο βέλγικος Σύλλογος Κινηματογραφικών και Τηλεοπτικών Μηχανικών⁴ χρησιμοποιεί τον ακόλουθο ορισμό

¹ McDonough, J., Proffitt M., Smith M.: Structural, technical, and administrative metadata standards. A discussion document. Technical report, Digital Library Federation (2000) (<http://www.diglib.org/standards/stamdframe.htm>)

² De Sutter, R., Notebaert S., De Walle, R.: Evaluation of Metadata Standards in the Context of Digital Audio-Visual Libraries. ECDL 2006, LNCS 4172, σσ. 220-231, 2006.

³ 10th European Conference on Research and Advanced Technology for Digital Libraries (ECDL) <http://www.ecdl2006.org/index.jsp>

⁴ Society of Motion Pictures and Television Engineers (SMPTE) <http://www.smpete.org>. Πρόκειται για μια διεθνή επαγγελματική ένωση μηχανικών που εργάζονται στην κινηματογραφική βιομηχανία και παράγει πρότυπα διεθνούς αποδοχής, συνιστώμενες πρακτικές και οδηγίες για την τηλεόραση, τον κινηματογράφο, τον ψηφιακό κινηματογράφο, την ηχητική και ταρική απεικόνιση.

για τη σύνδεση περιεχομένου και μεταδεδομένων: Περιεχόμενο = νόημα + μεταδεδομένα, Όπου το νόημα αντιπροσωπεύει στην περίπτωση μας το οπτικοακουστικό υλικό. Ο ορισμός υποδηλώνει ότι χωρίς μεταδεδομένα δεν υφίσταται περιεχόμενο και έτσι το νόημα είναι άχρηστο. Ένας δεύτερος ορισμός διευρύνει τον πρώτο: Αγαθό = περιεχόμενο + δικαιώματα χρήσης του Αυτός ο ορισμός υποδηλώνει ότι το περιεχόμενο είναι αξιόλογο (δηλαδή ένα αγαθό) εάν ο ιδιοκτήτης του έχει το δικαίωμα να το χρησιμοποιεί.

Σε αυτό μπορεί να προστεθεί η πρόταση πως για να είναι χρήσιμα τα δεδομένα πρέπει να έχουν δομή, δηλαδή να ακολουθούν μια συγκεκριμένη σύνταξη, όπως αναφέρεται στα επόμενα.

Όλα τα παραπάνω καταδεικνύουν πόσο επιτακτική είναι η ανάγκη επιλογής του ιδανικού κατά περίπτωση προτύπου μεταδεδομένων προκειμένου μια ψηφιακή βιβλιοθήκη οπτικοακουστικού υλικού να μεγιστοποιήσει την λειτουργικότητα και την χρησιμότητά της. Αυτό δείχνει και η πληθώρα νέων προτύπων που εμφανίζονται με κύριο στόχο την διαχείριση του οπτικοακουστικού υλικού. Ωστόσο, το γεγονός πως αυτά τα πρότυπα μεταδεδομένων καλύπτουν ποικίλες εφαρμογές, συνδέει και εξαρτά άμεσα την επιλογή του βέλτιστου προτύπου από το υπό υλοποίηση έργο.

Βασικές έννοιες

Πληροφοριακός πόρος

Ένα πληροφοριακός πόρος είναι μια λογική μονάδα η οποία μπορεί να αποθηκευτεί, να υποβληθεί σε επεξεργασία, να διαβιβαστεί ή να αναπαραχθεί σαν μία μονάδα και μπορεί να αντιμετωπιστεί σαν μία μονάδα από τους παραγωγούς της πληροφορίας, τους φορείς παροχής υπηρεσιών και τους χρήστες. Μπορεί να συσχετίζεται με ένα μοναδικό σύνολο μεταδεδομένων.

Μεταδεδομένα⁵

Ο βασικός και διεθνώς αναγνωρισμένος ορισμός για τα μεταδεδομένα είναι:

«δεδομένα για τα δεδομένα» ή «πληροφορία για την πληροφορία».

Πολυμέσα⁶

Πολυμέσα στον χώρο της τεχνολογίας της πληροφορίας (information field) σημαίνει "πολλαπλοί μεσολαβητές" μεταξύ της πηγής και του παραλήπτη της πληροφορίας ή "πολλαπλά μέσα" μέσω των οποίων η πληροφορία αποθηκεύεται, μεταδίδεται, παρουσιάζεται ή γίνεται αντιληπτή.

Όταν η πληροφορία είναι ψηφιακή μιλάμε για ψηφιακά πολυμέσα τα οποία ορίζουμε ως εξής:

Ψηφιακά πολυμέσα είναι ο τομέας που ασχολείται με την ελεγχόμενη από υπολογιστή ολοκλήρωση κειμένου, γραφικών, ακίνητης και κινούμενης εικόνας, animation, ήχου, και οποιουδήποτε άλλου μέσου ψηφιακής αναπαράστασης, αποθήκευσης, μετάδοσης και επεξεργασίας της πληροφορίας.

Σχετικές εργασίες

Κατα τους De Sutter, Notebaert και De Walle, η εμπειρία των Βιβλιοθηκών και της Βιβλιοθηκονομίας

από την δημιουργία ψηφιακών βιβλιοθηκών μπορεί να χρησιμοποιηθεί γόνιμα από κάθε άλλη προσπάθεια ψηφιοποίησης, συμπεριλαμβανομένου του οπτικοακουστικού υλικού. Εξάλλου, το βασικό κίνητρο για την ανάπτυξη ψηφιακών βιβλιοθηκών είναι η πεποίθηση πως θα παρέχουν καλύτερη κατανομή της πληροφορίας, ιδιαίτερα σε μια εποχή στην οποία τα υπολογιστικά συστήματα και τα δίκτυα έχουν μεταβάλλει εν πολλοίς τον τρόπο επικοινωνίας των ανθρώπων. Από την οπτική γωνία των βιβλιοθηκονόμων ο στόχος των ψηφιακών βιβλιοθηκών περιγράφεται από τον ορισμό τους ως «ηλεκτρονικών βιβλιοθηκών στις οποίες μεγάλος αριθμός γεωγραφικά κατανεμημένων χρηστών μπορεί να έχει πρόσβαση στα περιεχόμενα μεγάλων και ποικίλων αποθετηρίων ηλεκτρονικών αντικειμένων – δικτυακού κειμένου, εικόνων, χαρτών, ήχων, βίντεο, εμπορικών καταλόγων, επιστημονικών, εμπορικών και κυβερνητικών συνόλων δεδομένων – καθώς και υπερκειμένου και πολυμεσικών αντικειμένων»⁷. Ο Arms έχει παρατηρήσει ότι «[με τις ψηφιακές βιβλιοθήκες] όλοι έχουν πρόσβαση [σε] πληροφορία που ήταν πριν διαθέσιμη μόνο σε επαγγελματίες. Με έναν προσωπικό ηλεκτρονικό υπολογιστή ο χρήστης μπορεί να συμβουλευτεί υλικό που βρίσκεται αποθηκευμένο σε υπολογιστές ανά τον κόσμο»⁸.

Από τα παραπάνω συνάγεται ότι η βιβλιοθηκονομική κοινότητα προσβλέπει κατ' αρχήν στην διαθεσιμότητα και την ανταλλαγή των ψηφιακών αντικειμένων, γεγονός που οδήγησε στην ανάπτυξη του Metadata Encoding & Transmission Standard (METS)⁹ από την Βιβλιοθήκη του Κογκρέσου. Το METS παρέχει ένα σχήμα για την κωδικοποίηση των μεταδεδομένων που χρησιμοποιούνται για την διαχείριση και ανταλλαγή των ψηφιακών αντικειμένων μιας βιβλιοθήκης. Ωστόσο, υπό το πρίσμα αυτής της παρουσίας, το METS παρουσιάζει αδυναμία στα εξής σημεία:

- δεν κανονικοποιεί την δομική, διοικητική και τεχνική αποτύπωση των μεταδεδομένων,
- αναφέρεται σε διαθέσιμες τεχνικές που αφορούν στην προ των ψηφιακών βιβλιοθηκών περίοδο όσον αφορά στα περιγραφικά μεταδεδομένα, όπως οι εγγραφές Machine-Readable Cataloging (MARC)¹⁰ και το Encoded Archival Description (EAD)¹¹. Αυτές οι τεχνικές κρίνονται ανεπαρκείς στην περίπτωση εξολοκλήρου ψηφιακών εφαρμογών.

Ωστόσο η έρευνα προς την κατεύθυνση καταλληλότερων προτύπων για τις ψηφιακές βιβλιοθήκες βρίσκεται σε συνεχή εξέλιξη και φαίνεται να αποδίδει με την εμφάνιση προτύπων, όπως το Dublin Core Element Set¹², το General International Standard Archival Description (ISAD (G))¹³ και το Multimedia Content Description Framework (γνωστό σαν MPEG-7).

Η πλειονότητα των προσπαθειών και των ερευνών στις ψηφιακές βιβλιοθήκες διεξάγονται με στόχο την ανταλλαγή ψηφιακών αντικειμένων μεταξύ αποθετηρίων και την διασφάλιση της διαλειτουργικότητας. Το Open Archive Initiative Protocol for Metadata Harvesting (OAI-PMH)¹⁴ αποτελεί μια προσπάθεια για την τεχνική υλοποίηση της διαλειτουργικότητας σε κατανεμημένα

⁵ Ο ορισμός προέρχεται από: CEN Workshop Agreement: CWA 13699, September 1999

⁶ Ο ορισμός προέρχεται από: http://www.it.uom.gr/project/ntua_java/chap1.htm

⁷ Μπρ. Από: Sreenivasulu, V.: *The Role of Digital Librarian in the Management of Digital Information Systems*. Aslib Proceeding 18 (2000) 12-20

⁸ Μπρ. Από: Arms, W.: *[The online edition of] Digital Libraries*, MIT Press (c2000)

⁹ Σε ελληνική απόδοση: Πρότυπο Κωδικοποίησης και Μεταφοράς Μεταδεδομένων. Περισσότερες πληροφορίες στο: <http://www.loc.gov/standards/mets>

¹⁰ Σε ελληνική απόδοση: Μηχαναγνώσιμη Καταλογογράφηση. Περισσότερες πληροφορίες στο: <http://www.loc.gov/marc/>

¹¹ Σε ελληνική απόδοση: Πρότυπο Κωδικοποίησης Αρχειακής Περιγραφής. Περισσότερες πληροφορίες στο: <http://www.loc.gov/ead/>

¹² Περισσότερες πληροφορίες στο: <http://dublincore.org/>

¹³ International Council on Archives. *ISAD(G): General International Standard Archival Description, Second Edition*. (1999)

¹⁴ Περισσότερες πληροφορίες στο: <http://www.openarchives.org/OAI/openarchivesprotocol.html>

αρχεία και βασίζεται στην αξιοποίηση του Simple Dublin Core Metadata Element Set.

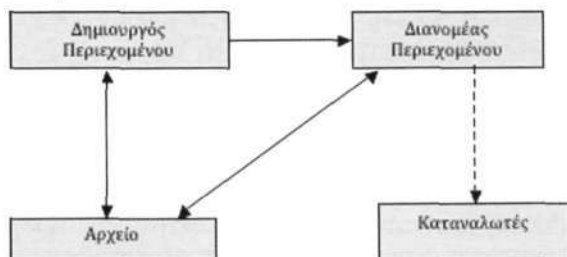
Οι χρησιμοποιούμενες τεχνικές για την δημιουργία και την διατήρηση ψηφιακών βιβλιοθηκών βιβλίων και εικόνων «στερούνται επαρκούς πρόβλεψης για την κωδικοποίηση περιγραφικών μεταδεδομένων, υποστηρίζοντας τεχνικά μεταδεδομένα για ορισμένες μόνο πηγές που περιλαμβάνουν κείμενο και εικόνα, ενώ δεν παρέχουν υποστήριξη για πολυμεσικό υλικό χρονικά εξαρτώμενο και προσφέρουν στοιχειώδεις λειτουργίες εσωτερικών και εξωτερικών συνδέσεων».¹⁵ Αυτό συνεπάγεται ότι η χρήση τέτοιων τεχνολογιών για την δημιουργία ψηφιακών βιβλιοθηκών με οπτικοακουστικό υλικό είναι ανεπαρκής, γεγονός που οδηγεί στην εμφάνιση νέων προτύπων μεταδεδομένων με κύριο στόχο την διαχείριση τέτοιων πόρων. Παρακάτω παρουσιάζεται ένα αντικειμενικό πλαίσιο σύγκρισης αυτών των προτύπων, όπως αυτό διαρθρώνεται στην εργασία «Αξιολόγηση προτύπων μεταδεδομένων στο πλαίσιο ψηφιακών οπτικοακουστικών βιβλιοθηκών» με τις απαραίτητες προσθήκες που επιβάλλει η κατανόηση και η απλούστευση των εννοιών.

Κριτήρια Επιλογής

Στη συνέχεια περιγράφονται τα κριτήρια που μπορούν να ληφθούν υπόψη προκειμένου να επιλεγεί το καταλληλότερο πρότυπο μεταδεδομένων για κάθε εφαρμογή. Τα κριτήρια δομήθηκαν μετά από συνεξέταση όλων των απόψεων, όπως η οργάνωση του περιεχομένου, αλλά ανεξάρτητα από επιμέρους περιορισμούς που μπορεί να επιβάλλει μία συγκεκριμένη υλοποίηση ή υλικό.

Κριτήριο 1: Εσωτερικά μοντέλα μεταδεδομένων, μοντέλα μεταδεδομένων ανταλλαγής

Σε αυτό το σημείο είναι σημαντικό να προσδιοριστούν τα εμπλεκόμενα μέρη που ανταλλάσσουν οπτικοακουστικό υλικό κατά την διάρκεια του κύκλου ζωής τους. Το European Broadcasting Union (EBU) προσδιορίζει τους καταναλωτές και τρεις εμπορικές οντότητες, δηλαδή τον δημιουργό του περιεχομένου, τον διανομέα και το αρχείο, ενώ διερεύνησε τις διαφορετικές σχέσεις που αναπτύσσονται ανάμεσά τους και παρουσίασε το αποτέλεσμα στο παρακάτω μοντέλο (βλ. Σχήμα 1). Αυτό το μοντέλο ροής δεδομένων είναι ανεξάρτητο από συγκεκριμένα πρότυπα μεταδεδομένων και εφαρμόζεται σε όλες τις εταιρείες παραγωγής και διακίνησης πολυμεσικού υλικού¹⁶, ενώ θα μπορούσε να εφαρμοστεί στον ευρύτερο χώρο των ψηφιακών βιβλιοθηκών και της ηλεκτρονικής δημοσίευσης με την προσθήκη της έννοιας του «χρήστη» δίπλα σε αυτή του «καταναλωτή».



Σχήμα 1. Διάγραμμα Ροής Δεδομένων EBU

Αφενός, αναπτύσσονται μοντέλα μεταδεδομένων που

αφορούν στην διαχείριση των μεταδεδομένων στο εσωτερικό ενός συστήματος, τα οποία αποκαλούνται εσωτερικά μοντέλα μεταδεδομένων ενώ αναπαρίστανται συνήθως σαν Διαγράμματα Οντότητας-Συσχέτισης¹⁷ και περιγράφουν την αρχιτεκτονική μιας βάσης δεδομένων που αποθηκεύει τα μεταδεδομένα για το οπτικοακουστικό υλικό.

Αφετέρου, υφίστανται μοντέλα μεταδεδομένων για την περιγραφή του τρόπου μεταφοράς της πληροφορίας από την πηγή στον προορισμό, τα λεγόμενα μοντέλα μεταδεδομένων ανταλλαγής, τα οποία είναι υπεύθυνα για την διακίνηση των μεταδεδομένων μεταξύ διαφορετικών συστημάτων. Στην υπό εξέταση περίπτωση, η ανταλλαγή παίρνει διευρυμένη έννοια, δηλαδή μπορεί να πραγματοποιηθεί μεταξύ οποιουδήποτε συνδυασμού δημιουργού, διανομέα, αρχείου και καταναλωτών.

Γενικά, ο παραπάνω διαχωρισμός σε εσωτερικά μοντέλα και σε μοντέλα ανταλλαγής αφορά περισσότερο τις δύο όψεις του ίδιου νομίσματος, δηλαδή τις δύο βασικές ανάγκες που πρέπει να καλύπτει ένα πρότυπο μεταδεδομένων.

Κριτήριο 2: Οργανωτική δομή της περιγραφής στα πρότυπα

Η οργανωτική δομή του νοήματος, δηλαδή του περιεχομένου αποτελεί το δεύτερο κριτήριο επιλογής προτύπου, σύμφωνα με την ίδια εργασία. Στην υπό εξέταση περίπτωση υπεύθυνη για τον βαθμό λεπτομέρειας των μεταδεδομένων που θα δημιουργηθούν είναι η εταιρεία παραγωγής/ κανάλι. Εν γένει, οι δύο απόψεις πιο κυριαρχούν συνοψίζονται στα εξής:

- Το περιεχόμενο θεωρείται αδιάσπαστη μονάδα την οποία καλούνται να περιγράψουν τα μεταδεδομένα με ενιαίο τρόπο, χωρίς περαιτέρω αναφορές στα ανεξάρτητα μέρη που την αποτελούν. Λόγου χάριν, εάν το περιεχόμενο συνδεθεί με ένα πρόγραμμα¹⁸, τα μεταδεδομένα θα περιγράψουν το συγκεκριμένο πρόγραμμα. Αυτό το μοντέλο αναφέρεται ως «επίπεδο μοντέλο μεταδεδομένων» (flat metadata model).

- Το περιεχόμενο χωρίζεται σε μικρά «πακέτα» καθένα από τα οποία περιγράφεται ξεχωριστά με πολλές λεπτομέρειες, πράγμα που επιτρέπει στους χρήστες να πραγματοποιούν πιο εξειδικευμένες αναζητήσεις. Για παράδειγμα, ένα πρόγραμμα μπορεί να διαιρεθεί σε πολλά «εκδοτικά αντικείμενα» που αντιπροσωπεύουν, λόγου χάριν, τις σκηνές.

Κάθε αντικείμενο μπορεί να περιγραφεί με επιπρόσθετα περιγραφικά μεταδεδομένα, ώστε να μπορεί να αναζητηθεί ανεξάρτητα από τα υπόλοιπα. Επιπροσθέτως, κάθε «εκδοτικό αντικείμενο» μπορεί να διαιρεθεί σε διαφορετικά «αντικείμενα μέσου», όπως ηχητικά μέρη, εικόνα, υπότιτλοι, κ.ο.κ. Η ίδια λογική μπορεί να εφαρμοστεί και με αντίθετη φορά, δηλαδή ένα σύνολο προγραμμάτων που συνδέονται μεταξύ τους μπορεί να ομαδοποιηθεί σε ένα «σύνολο προγραμμάτων», το οποίο θα περιγραφεί με μεταδεδομένα που αφορούν όλα τα προγράμματα που περιέχει και είναι κοινά, όπως ο τίτλος. Έτσι, δεν είναι απαραίτητο να επαναλαμβάνονται οι ίδιες πληροφορίες για κάθε πρόγραμμα, αφού το τελευταίο τις «κληρονομεί» από την ομάδα στην οποία

¹⁵ McDonough, J., Proffitt M., Smith M.: Structural, technical, and administrative metadata standards. A discussion document. Technical report, Digital Library Federation (2000) (<http://www.diglib.org/standards/stamdfame.htm>)

¹⁶ Hopper, R.: Metadata Exchange Scheme, v1.0. Technical Report No. 209, European Broadcasting Union (2002) Διαθέσιμο στο: http://www.ebu.ch/trev_290-hopper.pdf.

¹⁷ Τα γνωστά από τα Συστήματα Διαχείρισης Βάσεων Δεδομένων Entity Relationship Diagrams (ERDs) που παρέχουν υψηλού επιπέδου περιγραφή για τα εννοιολογικά μοντέλα δεδομένων. Γενικά, στο πρώτο στάδιο του σχεδιασμού ενός πληροφοριακού συστήματος χρησιμοποιούνται αυτά τα διαγράμματα για να περιγράψουν τις πληροφοριακές ανάγκες και τον τύπο των πληροφοριών που θα αποθηκευτούν στη βάση δεδομένων.

¹⁸ Ο όρος «πρόγραμμα» (αγγλικά "program") χρησιμοποιείται εδώ καταχρηστικά και περιλαμβάνει κάθε μορφής υλικό που θα μπορούσε να αποτελέσει αντικείμενο μιας ψηφιακής βιβλιοθήκης πολυμέσων.

ανήκει. Η βασική ιδέα που διέπει τα παραπάνω είναι η προσθήκη πληροφορίας, δηλαδή μεταδεδομένων, στα αντικείμενα στην κατάλληλη θέση και αναφέρεται ως «ιεραρχικό μοντέλο μεταδεδομένων» (hierarchical metadata model). Μία ανάλογη αρχιτεκτονική τεσσάρων επιπέδων, όπως αυτή περιγράφηκε παραπάνω, φαίνεται στο Σχήμα 2.



Σχήμα 2. Οργάνωση περιεχομένου

Παρά τα τεράστια πλεονεκτήματα που παρουσιάζουν ως προς την ταχύτητα και την αποτελεσματικότητα της αναζήτησης και της ανάκτησης, τα ιεραρχικά μοντέλα μεταδεδομένων δεν προτιμούνται πάντα από τους παραγωγούς/ κανάλια, με το σκεπτικό ότι έτσι περιορίζεται το κόστος το οποίο είναι αναλογικό με το πλήθος των μεταδεδομένων που χρειάζεται να συλλεχθούν. Δηλαδή, καθώς το μεταδεδομένα για ένα οπτικοακουστικό αντικείμενο πληθαίνουν, το περιθώριο κέρδους από τα πρόσθετα μεταδεδομένα μειώνεται, ενώ το κόστος παραγωγής τους αυξάνεται δυσανάλογα.. Στο μέλλον θα είναι αδύνατο να δημιουργηθούν πρόσθετα μεταδεδομένα χωρίς δαπάνες και έτσι, οι παραγωγοί/ κανάλια θα πρέπει να υπολογίζουν τη διαφορά ανάμεσα στο κόστος και την περιεκτικότητα σε μεταδεδομένα.

Κριτήριο 3: Υποστηριζόμενοι τύποι μεταδεδομένων

Όπως αναφέρθηκε σε αυτή την εισαγωγή τα μεταδεδομένα περιγράφουν το νόημα. Οι ίδιες οι απαιτήσεις των χρηστών καθορίζουν τους απαραίτητους τύπους μεταδεδομένων κατά περίπτωση. Δύο παρατηρήσεις που πρέπει να λαμβάνονται διαρκώς υπόψη είναι οι εξής:

- Το νόημα είναι άχρηστο χωρίς μεταδεδομένα και
- Το περιεχόμενο δεν έχει αξία όταν δεν συνοδεύεται από πληροφορίες για τα δικαιώματα που το αφορούν (δηλ. ποιός και κάτω από ποιες συνθήκες έχει πρόσβαση κ.λπ.)

Στα ακόλουθα αναφέρονται διάφοροι τύποι μεταδεδομένων σχετικοί με την διατήρηση του οπτικοακουστικού υλικού.

- **Προσδιοριστικά μεταδεδομένα**, που αναφέρονται στον μονοσήμαντο προσδιορισμό πληροφορίας και νοήματος. Κάτι τέτοιο μπορεί να υλοποιηθεί είτε μέσω επεξεργασιμών από τον άνθρωπο πεδίων, όπως ένας τίτλος ή ένα ευρετήριο, είτε μέσω μηχαναγώνισμων προσδιοριστικό, όπως το Unique Material Identifier

(UMID)¹⁹ ή το Uniform Resource Identifier (URI)²⁰. Εκτός από τα προσδιοριστικά μεταδεδομένα που συνδέονται με το νόημα, πολλές φορές είναι απαραίτητες και πρόσθετες πληροφορίες για τον εντοπισμό σχετικών αντικειμένων που είναι ενδεχομένως αποθηκευμένα σε άλλα συστήματα.

- **Μεταδεδομένα περιγραφής και ταξινόμησης**. Τα περιγραφικά μεταδεδομένα πρέπει να περιγράφουν αυτό που εκφράζει το νόημα, πράγμα που μπορεί να γίνει με την δημιουργία μιας λίστας λέξεων-κλειδιών οι οποίες θέτουν το νόημα σε ένα συγκεκριμένο πλαίσιο. Σε μερικές περιπτώσεις, οι λέξεις-κλειδιά επιλέγονται από ένα οργανωμένο λεξικό όρων, όπως είναι ένας θησαυρός. Εκτός από τον θησαυρό που εξυπηρετεί τους αμιγώς περιγραφικούς σκοπούς, μπορούν να χρησιμοποιηθούν και άλλα ταξινομικά σχήματα. Πράγματι, το περιεχόμενο μπορεί να κατηγοριοποιηθεί σε διάφορες προκαθορισμένες τάξεις σε συμφωνία με το είδος, το κοινό κ.ο.κ. Πρόκειται, δηλαδή, για τις δύο κατεξοχόν βιβλιοθηκονομικές δραστηριότητες της θεματικής περιγραφής και της ταξινόμησης. Ένα πολύ γνωστό ταξινομικό σύστημα είναι το Escort 2.4 του EBU το οποίο ομαδοποιεί το νόημα με τρόπο εννοιολογικό, διοικητικό, παραγωγικό, σχεδιαστικό, μεταφοράς, αναπαραγωγής και οικονομικό.

Ένας άλλος γνωστός τύπος περιγραφικών μεταδεδομένων περιλαμβάνει την περιγραφή του νοήματος με ένα μικρό κείμενο και χρησιμοποιείται σε κάθε σύστημα αρχειοθέτησης. Δυστυχώς, αυτά τα πεδία είναι επιρρεπή σε λάθη, όπως ορθογραφικά, και θα πρέπει να χρησιμοποιούνται προσεκτικά.

Τα ταξινομικά συστήματα των ραδιοηλεκτρονικών προγραμμάτων είναι δύο ειδών, πολλών (poly-dimensional) και πολλαπλών (multi-dimensional) διαστάσεων. Στα πρώτα, κάθε πρόγραμμα ταξινομείται μία φορά, από διαφορετικές οπτικές γωνίες. Για παράδειγμα, ένα εκπαιδευτικό αθλητικό πρόγραμμα για παιδιά ταξινομείται είτε ως εκπαιδευτικό (πρόθεση), είτε ως αθλητικό (περιεχόμενο), είτε ως παιδικό (κοινό). Στα ταξινομικά συστήματα πολλαπλών διαστάσεων, κάθε πρόγραμμα ταξινομείται τόσες φορές, όσες είναι οι διαστάσεις του συστήματος. Το πρόγραμμα του προηγούμενου παραδείγματος ταξινομείται ως εκπαιδευτικό στην σχετική παράμετρο, ως αθλητικό στην παράμετρο του περιεχομένου και ως παιδικό στην παράμετρο του κοινού.

- **Τεχνικά μεταδεδομένα**, τα οποία περιγράφουν τα τεχνολογικά χαρακτηριστικά του σχετιζόμενου νοήματος. Τα ελάχιστα απαιτούμενα τεχνικά μεταδεδομένα πρέπει να διευκρινίζουν τους αποκωδικοποιητές ήχου και εικόνας που μπορούν να χρησιμοποιηθούν για την αποκωδικοποίηση του οπτικοακουστικού υλικού. Με αυτή την πληροφορία, ο χρήστης έχει την δυνατότητα να αναπαράγει το νόημα. Ως εκ τούτου, τα τεχνικά μεταδεδομένα επιτρέπουν στο νόημα να χρησιμοποιηθεί, πράγμα που αποτελεί μια από τις βασικές απαιτήσεις για την δημιουργία νοήματος.

- **Μεταδεδομένα ασφάλειας και δικαιωμάτων**. Τα μεταδεδομένα ασφάλειας περιλαμβάνουν ποικίλες πληροφορίες που μπορούν να αφορούν διάφορες όψεις, από την ασφαλή μεταφορά (π.χ. μέθοδος κρυπτογράφησης)

¹⁹ Πρόκειται για ένα πρότυπο που εξασφαλίζει μια μέθοδο την παραγωγή μιας μοναδικής ετικέτας που συνδέεται με αρχεία

²⁰ Μία συμπληγής σειρά χαρακτήρων που χρησιμοποιούνται για τον προσδιορισμό ή την ονομασία πόρων με κύριο στόχο να επιτρέπει την αλληλεπίδραση των διάφορων παρουσιάσεων των πόρων μέσα σε ένα δίκτυο - κυρίως το Διαδίκτυο - με την χρήση ειδικών πρωτοκόλλων.

ως τα δικαιώματα πρόσβασης. Τα τελευταία είναι που μετατρέπουν το περιεχόμενο σε αγαθό. Τα μεταδεδομένα δικαιωμάτων πρόσβασης μπορούν να υποδιαιρεθούν σε πληροφορίες για τον κάτοχο των δικαιωμάτων και σε πληροφορίες για συμβόλαια σχετικά με την δημοσίευση του περιεχομένου και με τα άτομα που εμπλέκονται στην δημιουργία του νοήματος. Κάτοχος των δικαιωμάτων είναι ο οργανισμός στον οποίο ανήκουν τα πνευματικά δικαιώματα του οπτικοακουστικού υλικού.

• **Μεταδεδομένα έκδοσης.** Ο τελευταίος τύπος μεταδεδομένων περιγράφει την δημοσίευση ή τις δημοσιεύσεις του νοήματος. Κάθε δημοσίευση χαρακτηρίζεται από καθιερωμένη ημερομηνία δημοσίευσης, ώρα και διάρκεια, το κανάλι στο οποίο προβλήθηκε κ.ο.κ. Έτσι, οι υπεύθυνοι διαμορφώνουν μια ιδέα για την συχνότητα και την θεαματικότητα του νοήματος. Επιπλέον, αυτές οι πληροφορίες είναι απαραίτητες για την διαχείριση των δικαιωμάτων και των πληρωμών.

Κριτήριο 4: Σύνταξη και σημασιολογία

Όπως είναι γνωστό, μερικά πρότυπα ορίζουν μόνο την σύνταξη, ενώ άλλα μόνο την σημασιολογία²¹ ή και τα δύο. Η σύνταξη καθορίζει την παρουσίαση των μεταδεδομένων, η οποία μπορεί να πραγματοποιηθεί σε μορφή κειμένου ή σε δυαδική μορφή, πράγμα που αποτελεί και ένα από τα σημαντικότερα διλήμματα που εγείρονται. Η κειμενική παρουσίαση έχει το πλεονέκτημα ότι τα μεταδεδομένα είναι αναγνώσιμα από τον άνθρωπο, ωστόσο παρουσιάζουν μεγάλη φλυαρία. Αντίθετα, η δυαδική παρουσίαση παρέχει εξαιρετική πυκνότητα, αλλά μπορεί να αναγνωστεί μόνο από μηχανές. Στις περιπτώσεις που το προς επεξεργασία κείμενο είναι απλό, συνήθως, χρησιμοποιείται η γλώσσα XML (Extensible Markup Language)²² και το ίδιο το πρότυπο παρέχει ένα XML Σχήμα²³ που καθορίζει με ακρίβεια τη σύνταξη των μεταδεδομένων. Η χρήση του σχήματος δίνει την δυνατότητα ελέγχου της εγκυρότητας των μεταδεδομένων, πράγμα που ενισχύει την διαλειτουργικότητα.

Η σημασιολογία του προτύπου μεταδεδομένων καθορίζει το νόημα των στοιχείων. Απουσία σημασιολογικής περιγραφής, οι χρήστες υποθέτουν το νόημα, πιθανώς καταλήγοντας σε διαφορετικές ερμηνείες μεταξύ τους. Μόνο εάν η περιγραφή των μεταδεδομένων είναι κλειστή, δηλαδή κάθε στοιχείο περιγράφεται σημασιολογικά, όλοι οι χρήστες αντιλαμβάνονται το ίδιο νόημα για τα δεδομένα βελτιώνοντας την διαλειτουργικότητα.

Αξιολόγηση Προτύπων Μεταδεδομένων

Στο επόμενο μέρος, οι De Sutter, Notebaert και De Walle εφαρμόζουν τα κριτήρια αξιολόγησης που προηγήθηκαν σε τέσσερα γνωστά πρότυπα μεταδεδομένων, το Dublin Core, το MPEG-7, το P/Metan και το SMEF. Ο Πίνακας 1 παρέχει μία επισκόπηση των κριτηρίων για τα συγκεκριμένα πρότυπα.

Dublin Core

Το "Dublin Core"²⁴ είναι στενογραφία για το "Dublin Core Metadata Element Set". Πρόκειται για έναν κατάλογο στοιχείων μεταδεδομένων που συμφω-

νήθηκαν στο OCLC/NCSA Metadata Workshop τον Μάρτιο του 1995.

Το εργαστήριο οργανώθηκε από το OCLC και το Εθνικό Κέντρο για Εφαρμογές Υπερυπολογιστών (NCSA) για να προχωρήσει η ανάπτυξη ενός αρχείου μεταδεδομένων περιγραφής ηλεκτρονικών δικτυακών πληροφοριών. Το κίνητρο για την ανάπτυξη ενός τέτοιου προτύπου ήταν και είναι η συναίνεση για μια ελάχιστη περιγραφή των πόρων που μπορούν να χρησιμοποιηθούν προς όφελος όλων όσοι σχετίζονται με την δημιουργία, την αναζήτηση και την ανάκτηση ηλεκτρονικών πόρων. Υπήρξαν πολλές συμμετοχές από διάφορους επαγγελματίες (εκδότες, ειδικοί υπολογιστών, βιβλιοθηκονόμους).

Το Dublin Core χρησιμοποιείται συμπληρωματικά στις υπάρχουσες μεθόδους αναζήτησης και ευρετηρίωσης μεταδεδομένων του διαδικτύου. Το Dublin Core Element Set είναι ένα σύνολο 15 ορισμών και αντιπροσωπεύει μια ομάδα στοιχείων που χρησιμεύουν σε οποιοδήποτε οργανισμό οργανώνει και ταξινομεί πληροφορίες. Σε μια δεύτερη φάση, το μοντέλο επεκτάθηκε με τρία ακόμη στοιχεία και μία σειρά αλλαγές, που κατέληξαν στο Qualified Dublin Core Metadata Element Set (DMCES).

Το Dublin Core μπορεί να χρησιμοποιηθεί από καθέναν που θέλει να περιγράψει τους πόρους ενός πληροφοριακού συστήματος. Οι ιστοσελίδες είναι από τους πιο κοινούς τύπους πόρων στους οποίους μπορεί να εφαρμοστεί το Dublin Core, ενώ μια πληθώρα κοινοτήτων το ακολουθεί για την περιγραφή: εκπαιδευτικές οργανώσεις, βιβλιοθήκες, κυβερνητικά ιδρύματα, συντάκτες ιστοσελίδων κ.λπ.

Ο στόχος της έκδοσης του DMCES είναι η ανταλλαγή της περιγραφής των πόρων με σκοπό την πραγματοποίηση εφαρμογών μεταξύ διαφορετικών χώρων (κριτήριο 1). Τόσο η "simple", όσο και η "qualified" έκδοση του Dublin Core χαρακτηρίζονται από απλότητα, ωστόσο παρουσιάζουν ανεπάρκεια σε δύο σημεία. Αφενός, δεν υπάρχει πρόβλεψη για την περιγραφή ιεραρχικά δομημένου οπτικοακουστικού περιεχομένου, πρόβλημα που μπορεί να παρακαμφθεί με την δημιουργία αναφορών σε άλλα τμήματα της δομής του υπό περιγραφή αντικείμενου. Ως εκ τούτου, το DMCES είναι ένα επίπεδο πρότυπο μεταδεδομένων (κριτήριο 2). Αφετέρου, ο αριθμός των διαθέσιμων μεταδεδομένων είναι ιδιαίτερα περιορισμένος και εμποδίζει την λεπτομερή περιγραφή των οπτικοακουστικών πόρων στις ψηφιακές βιβλιοθήκες. Ειδικότερα, εξαιτίας του γενικού χαρακτήρα του DMCES, τα μεταδεδομένα περιγραφής των τεχνικών πληροφοριών, καθώς και αυτών που αφορούν στην ασφάλεια, στα δικαιώματα και στην έκδοση είναι περιορισμένα (κριτήριο 3).

Όσον αφορά στο κριτήριο 4, στο Dublin Core η σημασιολογία περιγράφεται συνοπτικά και ο χρήστης χαρακτηρίζεται από ελευθερία ερμηνείας αυτής της περιγραφής. Το DCMI παρέχει διαφορετικούς τρόπους συντακτικής περιγραφής των μεταδεδομένων, όπως βασικές αρχές ενσωμάτωσης του Dublin Core στην XML²⁵ και βασικές αρχές χρήσης του Dublin Core σε συνδυασμό με το RDF²⁶ (Resource Description Framework)²⁷.

²¹ Κατά τον Joram Borenstein, «η σημασιολογία καθορίζει το νόημα μιας έννοιας με τρόπο σαφή και οικονομικά ορθό».

²² Η XML είναι ένα απλό σύνολο κανόνων που επιτρέπει τη δημιουργία ενός συνόλου από ετικέτες που περιγράφουν τη δομή ενός αρχείου. Έτσι, παρέχει έναν απλό, ευέλικτο και παραμετροποιήσιμο τρόπο δόμησης δεδομένων.

²³ Το XML Σχήμα είναι ένα λεξιλόγιο σε XML για την έκφραση των κανόνων που διέπουν τα XML δεδομένα.

²⁴ Περισσότερες πληροφορίες στο <http://dublincore.org>

²⁵ Αυτές οι βασικές αρχές βρίσκονται στο <http://dublincore.org/documents/dc-xml-guidelines>

²⁶ «Το RDF είναι ένα γενικό πλαίσιο για την περιγραφή διαδικτυακών πόρων. Μια τέτοια περιγραφή μπορεί να περιέχει τους συγγραφείς ενός πόρου, την ημερομηνία δημιουργίας του κ.λπ.», κατά το γλωσσάριο του www.perfectxml.com. Επιπροσθέτως, προωθεί τη διαλειτουργικότητα και παρέχει την δυνατότητα σε μεμονωμένες κοινότητες να ορίσουν και να εκφράσουν το θέμα τους στο βασικό μοντέλο δεδομένων.

²⁷ Περισσότερες πληροφορίες για την χρήση του Dublin Core σε συνδυασμό με το RDF βρίσκονται στο <http://dublincore.org/documents/dca-rdf-xml>

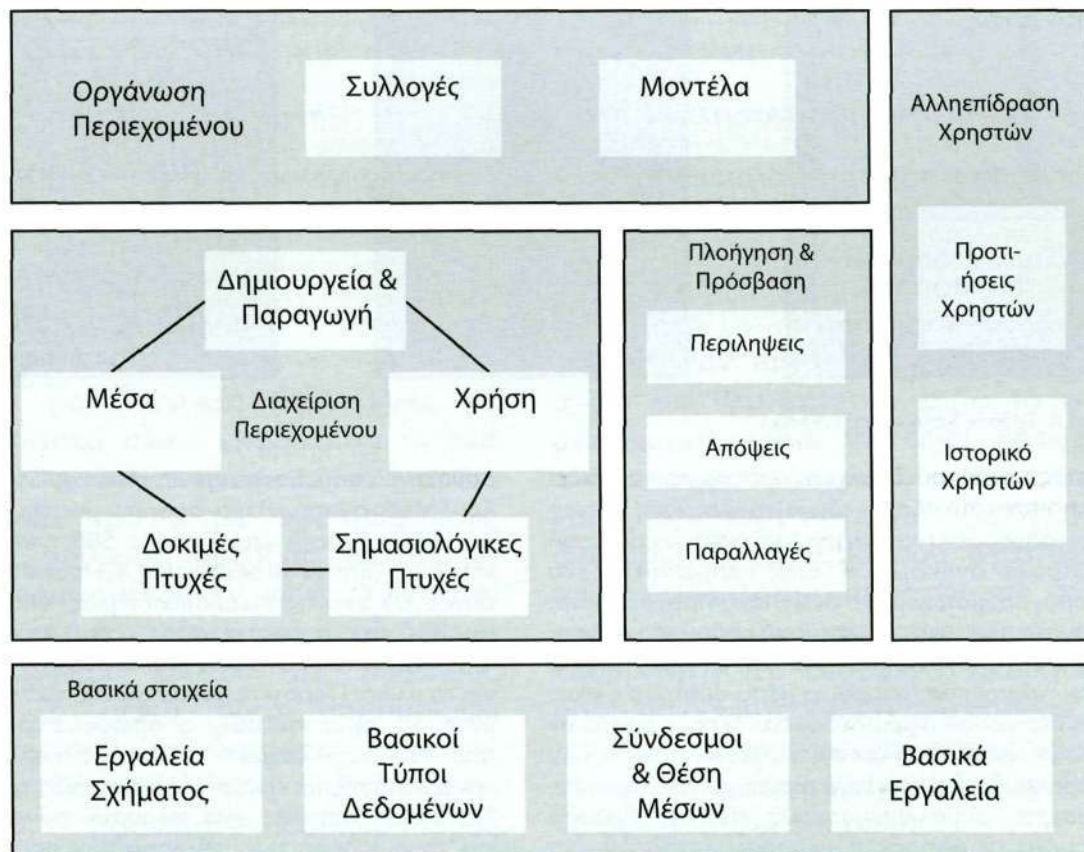
MPEG-7: Multimedia Content Description Interface

Ο Διεθνής Οργανισμός Προτυποποίησης και η Διεθνής Ηλεκτροτεχνική Επιτροπή (ISO/IEC) δημιούργησαν το Διεθνές Πρότυπο 15938, με την ονομασία Multimedia Content Description Interface²⁸, γνωστό ως MPEG-7, το οποίο παρέχει ένα πλούσιο σύνολο εργαλείων για την λεπτομερή περιγραφή πολυμεσικού περιεχομένου.

Το MPEG-7 αναπτύχθηκε μεταξύ άλλων για την ανταλλαγή μεταδεδομένων περιγραφής οπτικοακουστικού υλικού. Σχεδιάστηκε για να υποστηρίζει ένα ευρύ φάσμα εφαρμογών, χωρίς να εστιάζει σε κάποια συγκεκριμένη. Έτσι, πρόκειται για ένα μοντέλο ανταλλαγής (κριτήριο 1). Το MPEG-7 ορίζει την σύνταξη και την σημασιολογία όλων των στοιχείων μεταδεδομένων, τη μεν πρώτη με την χρήση ενός XML Σχήματος, ενώ την δεύτερη μέσα από ένα κανονιστικό κείμενο (κριτήριο 4). Τα στοιχεία δομούνται σαν «περιγραφείς» και «σχήματα περιγραφής». Για την αναπαράσταση ενός συγκεκριμένου χαρακτηριστικού του οπτικοακουστικού υλικού ορίζεται ένας «περιγραφέας». Στο ίδιο μοτίβο, «σχήμα περιγραφής» καλείται μια διαταγμένη δομή «περιγραφών» και άλλων «σχημάτων περιγραφής». Έτσι,

δημιουργείται ένα ιεραρχικό μοντέλο (κριτήριο 2). Παραδείγματος χάριν, ένα οπτικοακουστικό αντικείμενο μπορεί να περιγραφεί με βάση την χρονική του εξέλιξη ή την τμηματοποίηση της πηγής. Η τελευταία περιλαμβάνει περιγραφές για το ακουστικό και το οπτικό τμήμα, το οποίο χωρίζεται σε σκηνές, βασικά κάδρα και αντικείμενα. Οι υποστηριζόμενοι τύποι μεταδεδομένων (κριτήριο 3) είναι κυρίως τα περιγραφικά και τα τεχνικά, και σε μικρότερο βαθμό, τα προσδιοριστικά. Το πρότυπο δεν αναφέρεται σχεδόν καθόλου στα μεταδεδομένα έκδοσης, ασφάλειας και δικαιωμάτων, ωστόσο αυτές οι ανησυχίες εκφράζονται και στο πρότυπο MPEG-21.

Το τρίτο και το τέταρτο μέρος του προτύπου αφορούν στα τεχνικά μεταδεδομένα, τόσο για το οπτικό, όσο και για το ακουστικό υλικό. Στο πέμπτο μέρος του MPEG-7, που συχνά καλείται Multimedia Description Schemes (MDS), ορίζονται οι περιγραφείς και τα σχήματα περιγραφής για την περιγραφή και την ταξινόμηση του οπτικοακουστικού υλικού. Μία επισκόπηση των διάφορων λειτουργικών τομέων του MDS οπτικοποιείται στο Σχήμα 3²⁹.



Σχήμα 3. Επισκόπηση Multimedia Description Schemes

P / Meta

Το πρότυπο P/Meta αναπτύσσεται από την EBU σαν ένα λεξιλόγιο μεταδεδομένων για την ανταλλαγή προγραμμάτων στην βιομηχανία του θεάματος. Ως εκ τούτου, δεν αποτελεί μια εσωτερική αναπαράσταση, αλλά ένα σχήμα ανταλλαγής πληροφορίας που σχετίζεται με οπτικοακουστικά προγράμματα σε ενδοεπιχειρησιακό επίπεδο (κριτήριο1).

Το βασικό μοντέλο του P/Meta είναι ιεραρχικό και αποτελείται από πέντε επίπεδα (κριτήριο 2) – την εμπορική

ταυτότητα, την ομάδα στην οποία ανήκει το πρόγραμμα, το ίδιο το πρόγραμμα, το εκδοτικό αντικείμενο και το αντικείμενο μέσου. Κάθε ομάδα προγραμμάτων που συγκεντρώνεται από ένα κανάλι αποτελείται από ανεξάρτητα προγράμματα, τα οποία εν συνεχεία περιλαμβάνουν ανεξάρτητα εκδοτικά αντικείμενα, τα οποία θα μπορούσαν να βοηθούν για λόγους κατανόησης ως «επεισόδια». Τέλος, κάθε εκδοτικό αντικείμενο μπορεί να υποδιαιρεθεί σε αντικείμενα μέσου, όπως ηχητικά μέρη, εικόνες κ.ο.κ. Η ιεραρχία

²⁸ Σε ελεύθερη απόδοση: Διαπαθή Περιγραφής Πολυμεσικού Περιεχομένου

²⁹ Ελεύθερη απόδοση του Σχήματος από: Salembier, P., Smith, J. R.: MPEG-7 Multimedia Description Schemes. IEEE Transactions on Circuits, Systems and Video Technology 11 (221) 748-759

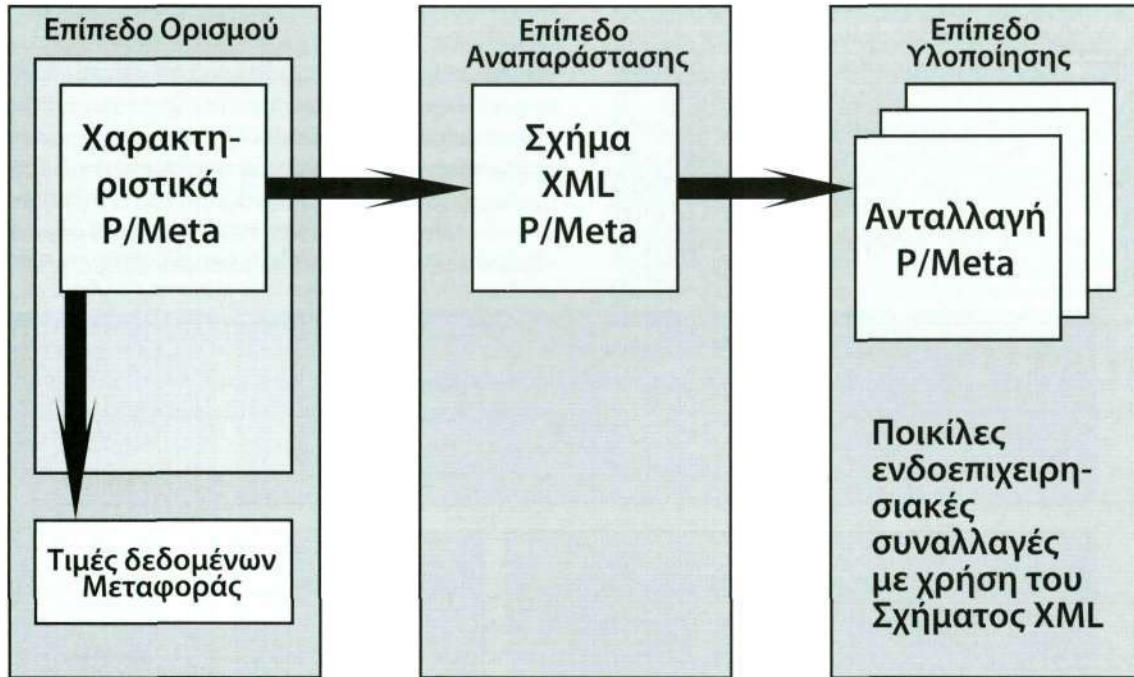
που περιγράφηκε παραπάνω είναι ανάλογη με αυτή που εικονίζεται στο Σχήμα 2.

Για να επιτευχθεί αυτή η ιεραρχική δομή, το πρότυπο ορίζει έναν αριθμό συνόλων (sets) και χαρακτηριστικών (attributes). Ως ένας επίπεδος κατάλογος χαρακτηριστικών, το P/Meta βοηθάει στον καθορισμό

- ενός κοινού συνόλου ορισμών για τα αντικείμενα μεταδεδομένων που απαιτούνται στην ενδοεπιχειρησιακή ανταλλαγή³⁰,
- ενός κοινού μορφοτύπου για τους τύπους πεδίων των μεταδεδομένων που θα εκφράζουν αυτά τα

μεταδεδομένα.

Επιπλέον, το P/Meta δίνει την δυνατότητα έκφρασης των χαρακτηριστικών μέσα από σύνολα, με τα οποία μπορεί να μεταβιβαστεί σημασιολογικά πολλή περισσότερη πληροφορία. Αυτό επιτρέπει την χρήση ενός σχετικά περιορισμένου αριθμού χαρακτηριστικών σε πολυάριθμους συνδυασμούς, πράγμα που οδηγεί σε έναν εύκαμπτο και ακριβή μηχανισμό συγκέντρωσης δεδομένων για τα πολυμεσικά αντικείμενα. Η φυσική υλοποίηση του P/Meta γίνεται πιθανότατα μέσω της παραγωγής ενός Σχήματος XML (Σχήμα 4).



Σχήμα 4. Τρόπος λειτουργίας P/Meta

Ένα σύνολο P/Meta ενσωματώνει χαρακτηριστικά, καθώς και άλλα σύνολα με τέτοιο τρόπο ώστε όλα τα σχετικά μεταδεδομένα να συλλέγονται για την περιγραφή του αντίστοιχου αντικείμενου. Έτσι, περιγράφεται ένα σύνολο προγραμμάτων με τα αντίστοιχα προγράμματα. Κάθε σύνολο προγραμμάτων και κάθε πρόγραμμα χαρακτηρίζεται από το προσδιοριστικό του (νούμερα και τίτλοι), την ταξινόμηση (σύμφωνα με το σύστημα Escort 2.4) και τα περιγραφικά μεταδεδομένα. Η περιγραφή των ανεξάρτητων προγραμμάτων συμπληρώνεται, εκτός των τριών παραπάνω στοιχειωδών τύπων μεταδεδομένων, από τέσσερις συμπληρωματικούς τύπους, δηλαδή μεταδεδομένα μεταφοράς ή έκδοσης, μεταδεδομένα που αφορούν στην εκδοτική ή την πολυμεσική διάσταση των αντικειμένων, τεχνικά μεταδεδομένα (προδιαγραφές ήχου και εικόνας, σχήματα συμπίεσης, κ.ο.κ.) και μεταδεδομένα δικαιωμάτων (συμβάσεις, λίστα δικαιωμάτων, κάτοχοι πνευματικών δικαιωμάτων). Αυτοί είναι και οι υποστηριζόμενοι τύποι μεταδεδομένων (κριτήριο 3).

Το P/Meta καθορίζει όλα τα σύνολα και τα χαρακτηριστικά και έτσι το πρότυπο περιέχει σαφείς ορισμούς. Η σύνταξη, όπως φαίνεται και στο Σχήμα 4, καθορίζεται από ένα XML Σχήμα (κριτήριο 4).

Standard Media Exchange Framework (SMEF)

Το SMEF αναπτύχθηκε από την Media Data Group του BBC, γνωστή ως Siemens SBS, για λογαριασμό του BBC (British Broadcasting Corporation). Η στενή συνεργασία και παρακολούθηση μιας ποικιλίας σχεδίων του BBC είχε ως αποτέλεσμα μια βαθιά κατανόηση των απαιτήσεων των καναλιών-παραγωγών σε πληροφορίες για το υλικό. Παρότι το μοντέλο αναπτύχθηκε για χρήση μέσα στα όρια του BBC, οι ορισμοί είναι ανεξάρτητοι από τον συγκεκριμένο οργανισμό και μπορούν να χρησιμοποιηθούν και από άλλα κανάλια-παραγωγούς. Το SMEF παρέχει ένα πλούσιο σύνολο ορισμών για τους τύπους των πληροφοριών που εμπλέκονται στην παραγωγή, την ανάπτυξη, την χρήση και την διαχείριση του πολυμεσικού υλικού. Στόχος του είναι να διασφαλίσει ότι διαφορετικά συστήματα αποθηκεύουν αυτή την πληροφορία κατά τρόπο ομοιογενή. Γι'αυτό το SMEF ορίζει ένα διάγραμμα οντοτήτων-συσχετίσεων που χρησιμεύει ως πλαίσιο για την αποθήκευση των μεταδεδομένων στο σύστημα. Επομένως, πρόκειται για ένα εσωτερικό μοντέλο μεταδεδομένων (κριτήριο 1).

Το SMEF καταγράφει όλη την πληροφορία που παράγεται κατά την διάρκεια του κύκλου παραγωγής, από την σύλληψη ενός προγράμματος ως την έκδοσή

³⁰ Αναφορά στον όρο B2B που περιγράφει την σχέση μεταξύ εμπορικών εταιρών χωρίς να εξυπηρετούν τον καταναλωτή (Business to business).

του Ένα εκδοτικό αντικείμενο, δηλαδή ένα πρόγραμμα μπορεί να χωριστεί σε διαφορετικά αντικείμενα μέσου, άρα πρόκειται για ένα ιεραρχικό μοντέλο μεταδεδομένων (κριτήριο 2). Κάθε αντικείμενο μέσου μπορεί να περιγραφεί εκτενώς με τεχνικά και περιγραφικά μεταδεδομένα. Οι οντότητες Εκδοτικό Αντικείμενο και Αντικείμενο Μέσου μπρούν να διασυνδεθούν με δύο άλλες οντότητες, την οντότητα Περιορισμού Χρήσης

και την οντότητα Πνευματικών Δικαιωμάτων. Από τα παραπάνω διαφαίνεται ότι το SMEF αναδεικνύει ιδιαίτερα τα μεταδεδομένα δικαιωμάτων (κριτήριο 3). Τέλος, το SMEF ορίζει την σημασιολογία όλων των οντοτήτων, των χαρακτηριστικών και των σχέσεων. Ο ορισμός των συντακτικών κανόνων εκφράζει τον τρόπο αναπαράστασης των μεταδεδομένων στο εσωτερικό του συστήματος (κριτήριο 4).

	Dublin Core	MPEG-7	P/Meta	SMEF
κριτήριο 1	ανταλλαγής	ανταλλαγής	ανταλλαγής	εσωτερικό
κριτήριο 2	επίπεδο	ιεραρχικό	ιεραρχικό	ιεραρχικό
κριτήριο 3	προσδιοριστικά, περιγραφής, τεχνικά	προσδιοριστικά, περιγραφής, τεχνικά	όλα	όλα (με έμφαση στα δικαιώματα)
κριτήριο 4	XML, RDF ανοιχτή σημασιολογία	XML κλειστή σημασιολογία	XML κλειστή σημασιολογία	ERD ανοιχτή σημασιολογία

Πίνακας 1. Επισκόπηση Αξιολόγησης Προτύπων Μεταδεδομένων

Συμπεράσματα

Η ανάγκη για ψηφιακές βιβλιοθήκες με οπτικοακουστικό υλικό εκφράζεται έντονα και γίνεται συχνά αντικείμενο συζήτησης τα τελευταία χρόνια με αφετηρία τους εντατικούς ρυθμούς αύξησης του πολυμεσικού περιεχομένου εν γένει. Υφίσταται μια αυξανόμενη ανάγκη για χρήση του οπτικοακουστικού υλικού με ποικίλους τρόπους. Το υλικό αυτό παρουσιάζει ιδιομορφίες που το διαφοροποιούν σε θέματα καταλογογράφησης, δηλαδή μεταδεδομένων. Βασικός προβληματισμός

σε αυτό τον τομέα είναι ο τρόπος με τον οποίο θα περιγραφούν αυτά τα ψηφιακά αντικείμενα, ώστε να είναι προσβάσιμα. Μια απάντηση σε αυτό το ερώτημα δίνει και η παραπάνω εργασία, αξιολογώντας τα διαφορετικά πρότυπα που υφίστανται και την καταλληλότητά τους για τις διάφορες εφαρμογές. Φυσικά, η έρευνα παράγει καθημερινά νέα δεδομένα, ακόμη και νέα πρότυπα όπως το VandA που προορίζεται για την αξιοποίηση ιστορικού οπτικοακουστικού υλικού, τα οποία μένει να αξιολογηθούν και να δοκιμαστούν εμπράκτως.

Ευχαριστίες

Ευχαριστώ τον κ. Σαράντο Καπιδάκη, στο μάθημα του οποίου εκπονήθηκε η παρούσα εργασία, για τη στήριξη.

Βιβλιογραφία

- BBC Technology: SMEF data model version 1.10. (Technical report) Διαθέσιμο στο <http://www.bbc.co.uk/guidelines/smef/>
- Bekaert, J., Van De Ville, D., Strauven, I., De Kooning, E., Van de Walle, R.: Metadata-based Access to Multimedia Architectural and Historical Archive Collections: a Review. *Aslib Proceeding 54* (2002) 362–371
- Digital Library Federation: METS: Metadata encoding and transmission standard. Technical report (2005) Διαθέσιμο στο <http://www.loc.gov/standards/mets>
- Digital Library Federation: The making of America II. Technical report (2005) Διαθέσιμο στο <http://sunsite.berkeley.edu/MOA2>
- Dublin Core Metadata Initiative: Dublin core metadata element set, version 1.1: Reference description. Technical report (2004) Διαθέσιμο στο <http://www.dublincore.org/documents/dces/>
- European Broadcasting Union: Escort: EBU System of Classification of RTV Programmes. Technical report (1995) Διαθέσιμο στο <http://www.ebu.ch/en/technical/metadata/specifications>
- European Broadcasting Union: P/Meta Metadata Exchange Scheme v1.1. Technical Report Tech. 3295 (2005) Διαθέσιμο στο http://www.ebu.ch/en/technical/metadata/specifications/notes_on_tech3295.php
- Hopper, R.: Metadata exchange scheme, v1.0. Technical Report Technical Report No. 290, European Broadcasting Union (2002) Διαθέσιμο στο <http://www.ebu.ch/trev/290-hopper.pdf>
- Hopper, R.: Metadata exchange standards. Technical Report Technical Report No. 284, European Broadcasting Union (2000) Διαθέσιμο στο http://www.ebu.ch/en/technical/trev/trev_284-hopper.pdf
- International Council on Archives: ISAD(G): General International Standard Archival Description, Second edition. (1999)
- Lagoze, C., Van de Sompel, H.: The making of the Open Archives Initiative Protocol for Metadata Harvesting. *Library Hi Tech 21* (2003) 118–128
- Library of Congress: Understanding Marc Authority Records. *Cataloging Distribution Service* (2003)
- Manjunath, B.S., Salembier, P., Sikora, T.: Introduction to MPEG-7: Multimedia Content Description Language. John Wiley & Sons (2002)
- Martinez, J.M., Koenen, R., Pereira, F.: MPEG-7: The Generic Multimedia Content Description Standard, Part 1. *IEEE MultiMedia 9* (2002) 78–87
- Martinez, J.M.: MPEG-7: Overview of MPEG-7 Description Tools, Part 2. *IEEE MultiMedia 9* (2002) 83–93
- McDonough, J., Proffitt, M., Smith, M.: Structural, technical, and administrative metadata standards. A discussion document. Technical report, Digital Library Federation (2000) Διαθέσιμο στο <http://www.diglib.org/standards/stamdfame.htm>
- Salembier, P., Smith, J.R.: MPEG-7 Multimedia Description Schemes. *IEEE Transactions on Circuits, Systems and Video Technology 11* (2001) 748–759
- Sreenivasulu, V.: The Role of a Digital Librarian in the Management of Digital Information Systems (dis). *Aslib Proceeding 18* (2000) 12–20
- The Society of American Archivists: Encoded Archival Description: Tag Library. Society of American Archivists (2002)
- Van de Sompel, H., Lagoze, C.: The Santa Fe Convention of the Open Archives Initiative. *D-Lib Magazine 6* (2000)
- Yu, S., Chen, H., Chang, H.: Building an Open Archive Union Catalog for Digital Archives. *The Electronic Library 23* (2005) 410–418