

Ιόνιο Πανεπιστήμιο

Τμήμα Αρχειονομίας - Βιβλιοθηκονομίας

Πρόγραμμα Μεταπτυχιακών Σπουδών στην Επιστήμη της Πληροφορίας

«Διοίκηση και Οργάνωση Βιβλιοθηκών με έμφαση στις Νέες Τεχνολογίες της Πληροφορίας»

Μάθημα:

Ηλεκτρονική Δημοσίευση

Θέμα εργασίας:

NewsML - News Markup Language

Ζούτσου Κυριακή

Διδάσκοντες καθηγητές:

Μ. Γεργατσούλης

Σ. Καπιδάκης

Ακαδημαϊκό Έτος 2005 - 2006

Περιεχόμενα

Περιεχόμενα	2
Εισαγωγή	3
Κεφάλαιο 1 ^ο	4
Πρότυπα διαχείρισης ειδήσεων	4
Κεφάλαιο 2 ^ο	7
Η ανάπτυξη της NEWSML	7
Κεφάλαιο 3 ^ο	9
Η δομή NewsML	9
Το επίπεδο του περιεχομένου – ContentItem	11
Το επίπεδο δομής - NewsComponent	21
Μεταδεδομένα περιεχομένου- NewsComponent	30
Διαχειριστικά μεταδεδομένα	31
Περιγραφικά μεταδεδομένα	36
Μεταδεδομένα δικαιωμάτων	40
NewsLines	41
Το επίπεδο διαχείρισης – NewsItem	44
Το επίπεδο ανταλλαγής - NewsML envelope	50
Βιβλιογραφία	53

Εισαγωγή

Ένα από τα χαρακτηριστικά της σύγχρονης κοινωνίας στην οποία ζούμε, είναι η πληθώρα πληροφοριών που μας κατακλύζουν. Ο χρήστης αντιμετωπίζει πλέον ένα πολύ σοβαρό πρόβλημα στην προσπάθεια του να εντοπίζει την πληροφορία που τον ενδιαφέρει και αυτό δεν είναι άλλο από τον εντοπισμό της σωστής και κατάλληλης πληροφορίας και την αντιμετώπιση του φαινομένου του «θορύβου».

Την ύπαρξη του συγκεκριμένου προβλήματος αντιμετώπισε τα τελευταία χρόνια η βιομηχανία των ειδήσεων με την ταχύτατη εξάπλωση του Διαδικτύου, η οποία νωρίς αντιλήφθηκε ότι το πρόβλημα δεν επικεντρωνόταν στη διάθεση της πληροφορίας, αλλά στη συλλογή των ειδήσεων που επιθυμούσαν. Για το λόγο αυτό στράφηκε στην αναζήτηση εργαλείων και προτύπων για την οργάνωση, δόμηση και διαχείριση των σημαντικών στο χώρο τους πληροφοριών, δηλαδή των ειδήσεων.

Σήμερα πλέον στο ειδησιογραφικό περιβάλλον έχουν ήδη αναπτυχθεί και χρησιμοποιούνται αρκετά πρότυπα τα οποία διαρκώς εξελίσσονται και βελτιώνονται., όπως για παράδειγμα τα NewsML, Nitf, Prism, XMLNews. Αυτά τα πρότυπα είναι σημαντικά και βασικά εργαλεία για την περιγραφή, τη δόμηση, την οργάνωση, την τεκμηρίωση, ανταλλαγή και διαχείριση ειδησιογραφικών δεδομένων μεγάλων φορέων, όπως είναι οι εφημερίδες, τα πρακτορεία ειδήσεων, οι εκδοτικοί οίκοι, κλπ.

Το ενδιαφέρον της παρούσας εργασίας επικεντρώνεται, όπως διαφαίνεται στη συνέχεια, σε ένα από τα πρότυπα που υπάρχουν στον κόσμο των ειδήσεων και συγκεκριμένα τη NewsML. Η επιλογή της προαναφερόμενης γλώσσας δόμησης και ανταλλαγής αντικειμένων ειδήσεων έγινε βάση της ευρείας διάδοσής της, αλλά κυρίως λόγω του γεγονότος ότι είναι βασισμένη στη διαδεδομένη γλώσσα σήμανσης XML. Πριν περάσουμε στο κύριο μέρος της εργασίας αυτής, κρίνεται χρήσιμο να δούμε επιγραμματικά ποια είναι τα πιο γνωστά πρότυπα που έχουν αναπτυχθεί στο χώρο των ειδήσεων, καθώς σκοπός της συντάκτριας είναι να δοθεί μια πλήρης εικόνα της NewsML.

Κεφάλαιο 1^ο

Πρότυπα διαχείρισης ειδήσεων

Πέρα από τη NewsML, όπως έχει ήδη αναφερθεί, έχουν αναπτυχθεί αρκετά εργαλεία για τη διαχείριση ειδήσεων στο ειδησιογραφικό περιβάλλον. Τα κυριότερα από αυτά αναπτύσσονται συνοπτικά στη συνέχεια:

- **NITF.** Το πρότυπο **Nitf** (News Industry Text Format) αναπτύχθηκε μέσα από τη συνεργασία των δύο σημαντικότερων οργανισμών προτυποποίησης για τη βιομηχανία ειδήσεων οι οποίοι είναι το Διεθνές Συμβούλιο Τύπου και Τηλεπικοινωνιών – **IPTC** (International Press Telecommunications Council) στις Η.Π.Α. και η Αμερικανική Ένωση Εφημερίδων – **NAA** (National Association of America). Πρόκειται για ένα δωρεάν πρότυπο, καλά τεκμηριωμένο, καλά υποστηριζόμενο και ευρέως χρησιμοποιούμενο από τους σημαντικότερους εκδότες εφημερίδων καθώς και πρακτορεία ειδήσεων σε ολόκληρο τον κόσμο. Το **Nitf** χρησιμοποιεί τη γλώσσα σήμανσης **XML** για να καθορίζει το περιεχόμενο και τη δομή των ειδησιογραφικών άρθρων. Τα μεταδεδομένα, που χρησιμοποιεί, εφαρμόζονται μέσα σε όλο το περιεχόμενο της είδησης και με αυτόν τον τρόπο οι πληροφορίες είναι πολύ περισσότερο αναζητήσιμες και χρήσιμες από ότι οι πληροφορίες που δίνονται σε σελίδες **HTML** (Hypertext Markup Language).
- **XMLNews.** Το πρότυπο XMLNews αποτελεί ίσως ένα από τα πιο ανεπτυγμένα μορφότυπα που χρησιμοποιούνταν από τη βιομηχανία ειδήσεων στον Παγκόσμιο Ιστό μέχρι πρόσφατα. Σχεδιάστηκε από τον David Megginson και είναι υποδιέστερο του Nitf. Το κομμάτι της περιγραφής είναι γνωστό ως «XMLNews-Story». Επιπροσθέτως, το XMLNews περιείχε το λεγόμενο «XMLNews-Meta», μία εφαρμογή σε RDF (Resource Description Framework) για την περιγραφή του περιεχομένου των ειδήσεων.
- **PRISM.** Το πρότυπο PRISM (Publishing Requirements for Industry Standard Metadata), το οποίο δημιουργήθηκε και λειτούργησε κάτω από

την αιγίδα της εταιρείας IDEAlliance (International Digital Enterprise Alliance), καθορίζει ένα λεξιλόγιο μεταδεδομένων βασισμένο στη γλώσσα XML και χρησιμοποιείται για τη διαχείριση, τη συγκομιδή, τη μετεξέλιξη και την πολυσκοπιμότητα περιοδικών, ειδήσεων, καταλόγων, βιβλίων και το δεσπόζων περιεχόμενό τους. Το πεδίο ενδιαφέροντος για τη δημιουργία του προτύπου PRISM ξεκινά από τις ανάγκες των εκδοτών να λαμβάνουν, να παρακολουθούν και να μεταφέρουν πολυτμηματικό περιεχόμενο.

- **ICE.** Η ανακοίνωση της δημιουργίας του πρωτοκόλλου **ICE** (Information and Content Exchange) τον Οκτώβριο του 1998 από την εταιρεία **IDEAlliance** έδωσε μία νέα ώθηση όσον αφορά το αντικείμενο «Ανταλλαγή Περιεχομένου και Πληροφορίας». Η δημιουργία και ολοκλήρωση του πρωτοκόλλου που βασιζόταν στην γλώσσα σήμανσης XML και προοριζόταν για χρήση στον Παγκόσμιο Ιστό αποτέλεσε ένα σημαντικό βήμα για την υποστήριξη της διανομής πληροφοριών μέσα στην εκδοτική και όχι μόνο κοινωνία.
- **IPTC 7901.** Το **IPTC** (International Press Telecommunications Council) **7901** αναπτύχθηκε από τον οργανισμό **IPTC**, χρησιμοποιούνταν για τη μετάδοση μηνυμάτων κειμένου σε εφημερίδες, πρακτορεία ειδήσεων και άλλους παραλήπτες ειδήσεων. Η πρώτη του έκδοση εμφανίστηκε στις αρχές της δεκαετίας του 1980 και ενημερωνόταν σε τακτά χρονικά διαστήματα μέχρι και την τελευταία του έκδοση (**IPTC 7901 5.0**), η οποία εγκρίθηκε το 1995. Από τότε, όμως, και μετά η ανάπτυξη του **IPTC 7901** σταμάτησε παρόλο που χρησιμοποιείται σε πολλές χώρες ακόμη και σήμερα.
- **RSS.** Το **RSS** (Really Simple Syndication) αποτελεί ένα μορφότυπο που αφορά τη διανομή ειδήσεων και τα περιγραφικά μεταδεδομένα της όλης διαδικασίας. Είναι και αυτό μία εφαρμογή βασισμένη στην XML, προσαρμόζεται με την W3C's RDF προδιαγραφή και είναι επεκτάσιμο μέσα από το XML – namespace και την εφαρμογή του RDF.
- **EVENTSML.** Το **EventsML** (Events Markup Language) ξεκίνησε να αναπτύσσεται από το **IPTC** και αποτελεί ένα πολύ συγκεκριμένο και ιδιαίτερο πρότυπο για την ανταλλαγή δομημένων δεδομένων, όπως είναι

τα «γεγονότα». Στην ουσία ανταλλάσσει πληροφορίες σχετικά με εξειδικευμένα γεγονότα, τα οποία είναι αξιόπιστα από ειδησεογραφικής πλευράς.

- **SPORTSML.** Το **SportsML** (Sports Markup Language) έχει ως σκοπό να αποτελέσει ένα παγκόσμιο βασισμένο στην XML γλώσσα πρότυπο για την ανταλλαγή αθλητικών δεδομένων. Έχει σχεδιαστεί με στόχο να είναι εύκολα κατανοητό και ευπροσάρμοστο, όσο το δυνατόν περισσότερο στα υπολογιστικά συστήματα. Επιτρέπει την ανταλλαγή πληροφοριών, που έχουν να κάνουν με αθλητικά σκορ, προγράμματα και στατιστικά στοιχεία μέσα στα πλαίσια μίας ευρείας ποικιλίας αθλητικών διαγωνισμών.
- **PROGRAMGUIDEML.** Το **ProgramguideML** (Programguide Markup Language) αποτελεί μία λύση για τη μεταφορά ραδιοτηλεοπτικών πληροφοριών και βασίζεται στο πρότυπο **NewsML**. Η δημιουργία του επιτρέπει την ανταλλαγή όλων των ραδιοτηλεοπτικών δεδομένων, που αφορούν τους εκδότες ειδήσεων και τους διάφορους ραδιοτηλεοπτικούς σταθμούς.

Κεφάλαιο 2ο

Η ανάπτυξη της NEWSML

Το NewsML (News Markup Language) έχει αναπτυχθεί και επικυρωθεί ως ένα ανοικτό πρότυπο από το Διεθνές Συμβούλιο Τύπου και Τηλεπικοινωνιών – IPTC (International Press Telecommunications Council) για τη δόμηση και διαχείριση ειδήσεων. Το Διεθνές Συμβούλιο Τύπου και Τηλεπικοινωνιών ιδρύθηκε το 1965 με σκοπό τη διαφύλαξη των τηλεπικοινωνιακών συμφερόντων και ενδιαφερόντων του παγκόσμιου τύπου και μία από τις πρωταρχικές δραστηριότητές του είναι η δημιουργία εργαλείων και προτύπων για την ανταλλαγή δεδομένων ειδήσεων. Έχοντας ένα καλό παρελθόν στην ανάπτυξη και διατήρηση μορφοτύπων ανταλλαγής ειδήσεων, το IPTC δημιούργησε την NewsML ως το πιο κατανοητό και πολύπλευρο τρόπο μετάδοσης όλων αυτών των πληροφοριών μεταξύ συστημάτων προκειμένου να γίνει η ανταλλαγή πιο αποτελεσματική και έγκυρη.

Ο σκοπός της NewsML είναι να καταστήσει εφικτή τη ανταλλαγή ειδήσεων είτε είναι κείμενο, φωτογραφία είτε είναι video, με ακρίβεια και ταχύτητα δένοντας το περιεχόμενο με έναν τρόπο ο οποίος επιτρέπει διαδικασίες αυτοματισμού. Η ανταλλαγή ειδήσεων είναι η διαδικασία μεταφοράς όχι μόνο των κυρίως ειδήσεων, αλλά επίσης και δεδομένων που περιγράφουν αυτό το περιεχόμενο σε μία πιο περιληπτική μορφή (μεταδεδομένα), πληροφορίες πώς να χειριζόμαστε τις ειδήσεις με έναν πιο αποτελεσματικό τρόπο (δεδομένα διαχείρισης ειδήσεων) και τέλος πληροφορίες για την μετάδοση και κατεύθυνση των ειδήσεων.

Αποτελεί ένα συμβατό, επεκτάσιμο και ευέλικτο πλαίσιο για τα ειδησεογραφικά τεκμήρια. Είναι βασισμένο στη διαδεδομένη γλώσσα σήμανσης XML και σε άλλα κατάλληλα πρότυπα και προδιαγραφές. Έχει ως σκοπό να εμφανίζει ξανά και να διαχειρίζεται κομμάτια ηλεκτρονικών ειδήσεων, συλλογές τέτοιων κομματιών, τις σχέσεις μεταξύ τους, καθώς και τα σχετικά μεταδεδομένα τους μέσα από τον ευρύτερο κύκλο ζωής τους.

Επιτρέπει την παροχή πολλαπλών εμφανίσεων της ίδιας πληροφορίας και χειρίζεται αυθαίρετες αναμίξεις τύπων, μορφοτύπων, γλωσσών και κωδικοποιήσεων.

Οι επίσημες προϋποθέσεις όσον αφορά τη χρήση του αντικατοπτρίζουν τις προκλήσεις, οι οποίες εγείρονται από ορισμένες νέες απαιτήσεις. Οι ειδήσεις μπορεί να μεταφέρονται ως μοναδικά κομμάτια ή ως πακέτα διαφόρων σχετιζόμενων κομματιών, και πρέπει να υπάρχουν και τα μεταδεδομένα για να επιτρέπουν αποτελεσματική παραγωγή, μεταφορά και χρήση, συμπεριλαμβανομένων των ταξινομήσεων και της αναζήτησης. Τα πακέτα ειδήσεων μπορούν να περιέχουν ποικίλους τύπους περιεχομένου - κείμενο, εικόνες, βίντεο, ήχο - καθένα από τα οποία είναι ισοδύναμα μεταξύ τους. Το ίδιο πακέτο ειδήσεων μπορεί επίσης να υπάρχει σε έναν αριθμό διαφορετικών μορφών, όπως μεταφράσεις ενός κειμένου σε διαφορετικές γλώσσες, ή παρουσίαση εικόνων σε εναλλακτικά μορφότυπα.

Η κύρια χρήση του προτύπου πιθανόν να γίνεται για την ανταλλαγή ειδήσεων, θα πρέπει όμως να μπορεί να χρησιμοποιηθεί και για τη δημιουργία, τη διαχείριση και τη δημοσίευση των ειδήσεων σε συστήματα δικτύου καθώς και για εφαρμογές αρχείων. Επιπροσθέτως, η NewsML πρέπει να υποστηρίζει τη διαχείριση και την ανάπτυξη των κομματιών των ειδήσεων συνεχώς. Τα ειδησιογραφικά γεγονότα συχνά αναπτύσσονται βαθμιαία και έτσι προκύπτει η ανάγκη να ενημερώνονται, να προστίθενται ή και να αντικαθίστανται νεώτερες εκδόσεις. Κομμάτια ειδήσεων σε διαφορετικά μέσα μπορεί να μην είναι διαθέσιμα την ίδια χρονική στιγμή και να χρειάζεται να μεταφερθούν μαζί.

Τέλος πολύ σημαντικό παράγοντας στο ειδησιογραφικό περιβάλλον, αποτελεί η αυθεντικοποίηση και η υπογραφή των μεταδεδομένων και του περιεχομένου της είδησης. Ας μην ξεχνάμε ότι η αξία του περιεχομένου της είδησης καθώς και τα συσχετιζόμενα μεταδεδομένα της εξαρτώνται από την αξιοπιστία της.

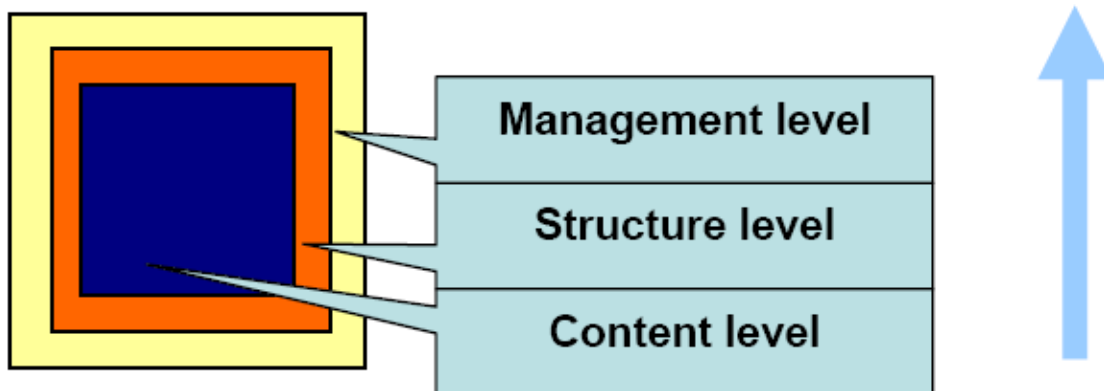
Κεφάλαιο 3^ο

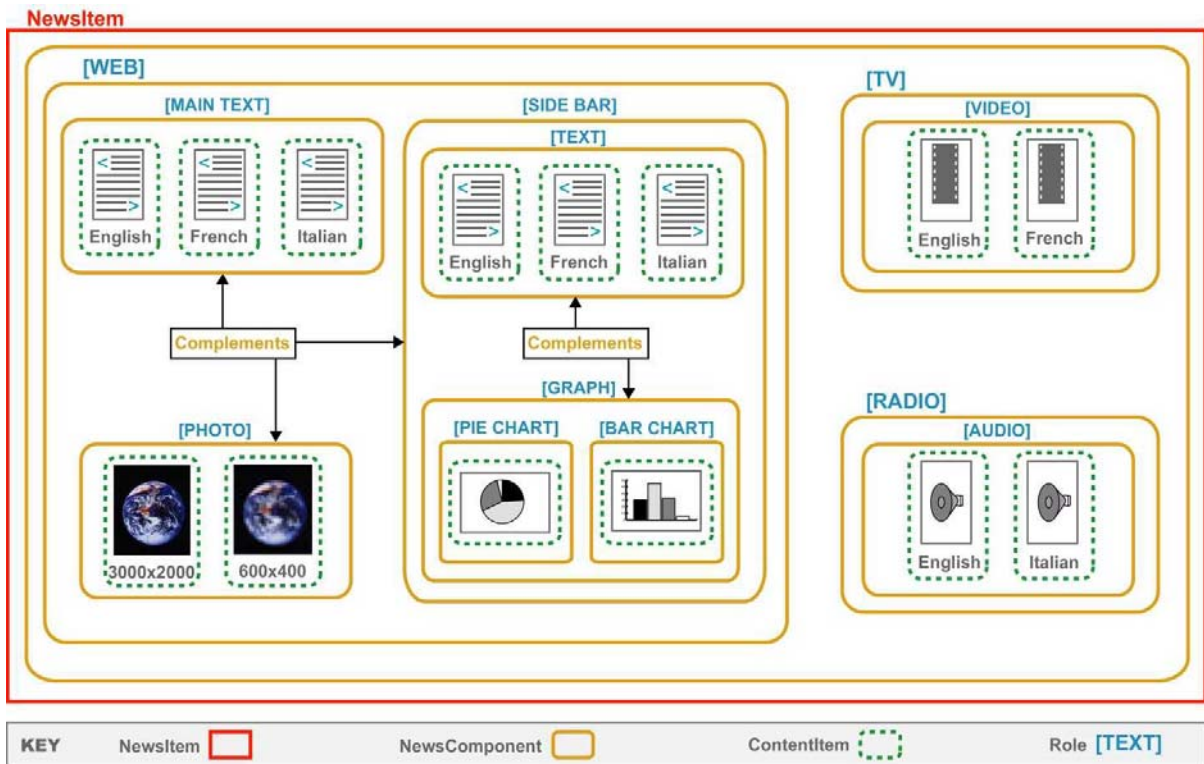
Η δομή NewsML

Η NewsML έχει δομή τεσσάρων επιπέδων. Κάθε επίπεδο αντιπροσωπεύεται από ένα news object. Ο προσδιορισμός του όρου βρίσκεται στα NewsML specifications, και είναι ένα από τα βασικά συστατικά των τεκμηρίων NewsML. Τα διαφορετικά είδη των news objects είναι το NewsEnvelope, NewsItem, NewsComponent και το ContentItem.

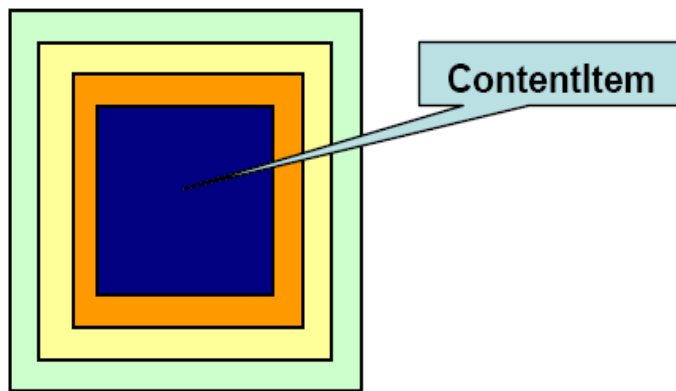
Αυτά τα επίπεδα παρέχουν συμπληρωματικές λειτουργίες και ονομάζονται από το εξωτερικό ως το εσωτερικό ως εξής:

1. Το **Exchange** level - αντιπροσωπεύεται από το στοιχείο **NewsEnvelope**
2. Το **Management** level - αντιπροσωπεύεται από το στοιχείο **NewsItem**
3. Το **Structure** level - αντιπροσωπεύεται από το στοιχείο **NewsComponent**
4. Το **Content** level - αντιπροσωπεύεται από το στοιχείο **ContentItem**





Το επίπεδο του περιεχομένου – ContentItem

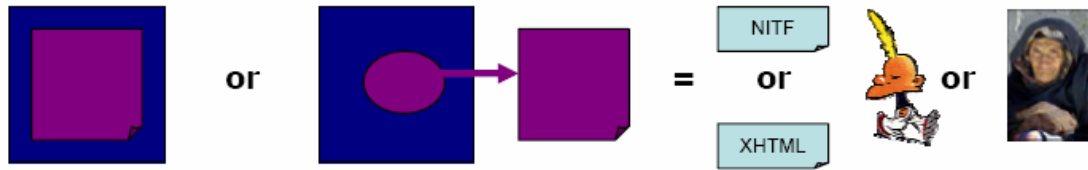


Ο πυρήνας της NewsML επικεντρώνεται αναμφίβολα στο ίδιο το περιεχόμενο. Το στοιχείο ContentItem παρέχει ένα ενιαίο πλαίσιο περιγραφής

του περιεχομένου, που δεν περιορίζεται από το μέσο στο οποίο είναι αποθηκευμένο το συγκεκριμένο περιεχόμενο.

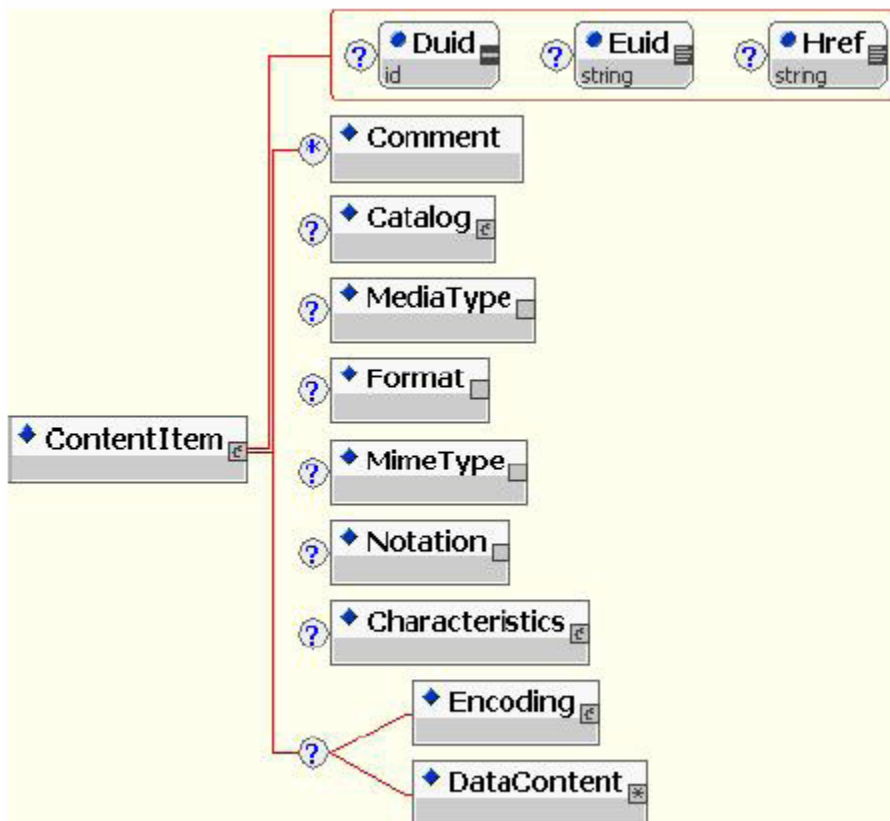
Το ContentItem αποτελεί στην ουσία την ενότητα που περικλείονται τα δεδομένα περιεχομένου σε ένα ειδησιογραφικό περιβάλλον και έχει τη δυνατότητα να αντιπροσωπεύει κάθε είδους πληροφοριακού αρχείου, από ένα κείμενο και μια φωτογραφία μέχρι γραφικά (EPSF), Flash animation, βίντεο ή οποιοδήποτε άλλο πολυμεσικό αντικείμενο. Αποτελεί ένα περίβλημα/ πακέτο που καθιστά κατανοητό, σε ένα επεξεργαστή της NewsML, όχι μόνο το ίδιο το περιεχόμενο αλλά και τα μεταδεδομένα του που περιγράφουν είτε τεχνικές ιδιότητες ή φυσικά χαρακτηριστικά του περιεχομένου. Το στοιχείο ContentItem δεν αποτελεί σε καμιά περίπτωση μεμονωμένο κομμάτι που φέρει πληροφορίες σε ένα ειδησιογραφικό περιβάλλον. Αντίθετα, συμπεριλαμβάνεται στο στοιχείο NewsComponent.

Το στοιχείο ContentItem καθιστά το περιεχόμενο του διαθέσιμο με δύο διακριτούς τρόπους. Στο μεν πρώτο τρόπο το περιεχόμενο περικλείεται μέσα στο ίδιο το στοιχείο, ενώ στη δεύτερη περίπτωση στο στοιχείο εμπεριέχεται μόνο μια παραπομπή μέσω της οποίας μπορεί να ανακτηθεί το περιεχόμενο που μπορεί να είναι αποθηκευμένο αλλού, παραδείγματος χάρη σε έναν απομακρυσμένο διακομιστή (server).



Στο σημείο αυτό πιθανόν να προκύπτει το ερώτημα γιατί περιλαμβάνονται μεταδεδομένα στο στοιχείο ContentItem. Για το λόγο αυτό, θα πρέπει στο σημείο αυτό να διευκρινιστεί ότι η παροχή μεταδεδομένων στο επίπεδο αυτό του περιεχομένου εξυπηρετεί έναν βασικότατο παράγοντα, την εξοικονόμηση χρόνου, καθώς μια εφαρμογή NewsML δεν χρειάζεται να ανακτήσει το ίδιο το τεκμήριο στην περίπτωση που φυλάσσεται σε έναν απομακρυσμένο διακομιστή, ώστε να προσδιοριστούν τα χαρακτηριστικά του.

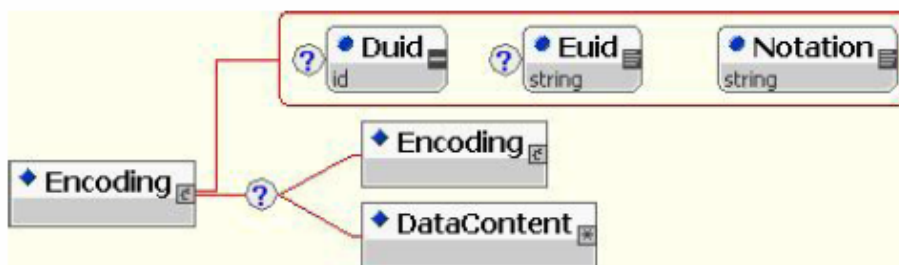
Οι διαφορετικοί τύποι των μεταδεδομένων που προβλέπει η NewsML στο επίπεδο περιεχομένου δηλώνονται με τα εξής στοιχεία: media type (τύπος μέσου), format (μορφότυπο), mime type (τύπος μίμησης), notation (σημειώσεις), καθώς και ένα σύνολο χαρακτηριστικών.



Ενσωμάτωση των δεδομένων

Τα πραγματικά δεδομένα, όπως έχουμε ήδη αναφέρει, είτε βρίσκονται αποθηκευμένα σε απομακρυσμένη θέση, για την οποία υπάρχει παραπομπή στο γνώρισμα Href, είτε βέβαια περιλαμβάνονται εξολοκλήρου στο στοιχείο ContentItem. Στην τελευταία περίπτωση τα δεδομένα φυλάσσονται στο υποστοιχείο DataContent που είναι τύπου ANY. Αυτό σημαίνει ότι στο συγκεκριμένο υποστοιχείο μπορούν να εισαχθούν οποιαδήποτε συμβατά με την XML δεδομένα, ακόμα και περιεχόμενο κωδικοποιημένο στο δυαδικό σύστημα. Όταν το περιεχόμενο που εισάγεται στο υποστοιχείο DataContent είναι ένα κείμενο, τότε το σύνολο των χαρακτήρων πρέπει να είναι οπωσδήποτε κωδικοποιημένο με τον ίδιο ακριβώς τρόπο με το σύνολο χαρακτήρων που έχει εισαχθεί στην εγγραφή NewsML. Σε αυτή την περίπτωση συνίσταται η χρήση του UTF-8.

Όπως έχουμε ήδη αναφέρει παραπάνω, αυτούσιο το περιεχόμενο μπορεί να συμπεριληφθεί στο στοιχείο DataContent. Σε αυτή την περίπτωση το στοιχείο Encoding περικλείει το DataContent. Το γνώρισμα Notation του στοιχείου Encoding υποδηλώνει τον αλγόριθμο κωδικοποίησης.



Όταν μία πηγή είναι κωδικοποιημένη ποικιλοτρόπως, όπως για παράδειγμα ένα αρχείο Zip που ακολούθως κωδικοποιείται με Base64, τότε το στοιχείο Encoding επαναλαμβάνεται όπως φαίνεται στο παράδειγμα που ακολουθεί:

```

<ContentItem>
  <MediaType FormalName="Graphic"/>
  <Format FormalName="Freehand"/>
  <Encoding Notation="base64">
    <Encoding Notation="zip">
      <DataContent>SDF35GZsDFGZER5R6RTGSE5 ... SRTYERT</DataContent>
    </Encoding>
  </Encoding>
</ContentItem>

```

Πολλές είναι οι επιλογές που επιτρέπουν την εισαγωγή δεδομένων κειμένου στο στοιχείο DataContent. Ενδεικτικά αναφέρονται:

- –Η χρήση χώρου ονομάτων (namespace). Ένας συγκεκριμένος χώρος ονομάτων μπορεί να δηλωθεί μέσα στο στοιχείο DataContent. Η τακτική αυτή κρίνεται χρήσιμη όταν το τεκμήριο χρειάζεται να είναι έγκυρο σύμφωνα με ένα XML schema ή ένα DTD. Το στοιχείο που εισάγεται στο στοιχείο DataContent θα πρέπει να αποτελεί το βασικό στοιχείο που περιλαμβάνεται στη δομή της XML, π.χ. <nitf>, <xhtml> κ.λπ.

```

<ContentItem>
  <MediaType FormalName="Text"/>
  <Format FormalName="NITF"/>
  <DataContent>
    <nitf xmlns="urn:newsml:iptc.org:20011012:NITF">
      <body>
        <body.content>
          <p>Today, <person>Clinton</person> visited...</p>
          <p><person>Al Gore</person> also attended the...</p>
        </body.content>
      </body>
    </nitf>
  </DataContent>
</ContentItem>

```

- Η απευθείας εισαγωγή χωρίς τη δήλωση χώρου ονομάτων. Η εισαγωγή δεδομένων σε XML χωρίς τη δήλωση χώρου ονομάτων είναι επίσης πιθανή. Σε αυτή την περίπτωση ο έλεγχος αξιοπιστίας της XML μπορεί να γίνει με ένα DTD. Στην περίπτωση αυτή, στην αρχή του τεκμηρίου γίνεται η δήλωση του DTD είτε με παραπομπή στο πρωτότυπο αρχείο στο διαδίκτυο είτε με αναφορά του μονοπατιού όταν το DTD είναι αποθηκευμένο τοπικά.

```

<!DOCTYPE NewsML SYSTEM "http://www.company.com/dtd/NewsML_1.2.dtd" [
<ENTITY % NITF SYSTEM "http://www.company.com/dtd/nitf-3-1.dtd">
%nitf,
]>
<NewsML>
  <!-- ----- -->
  <ContentItem>
    <MediaType FormalName="Text"/>
    <Format FormalName="NITF"/>
    <DataContent>
      <nitf>
        <body>
          <body.content>
            <p>Today, <person>Clinton</person> visited...</p>
            <p><person>Al Gore</person> also attended the...</p>
          </body.content>
        </body>
      </nitf>
    </DataContent>
  </ContentItem>
  <!-- ----- -->
</NewsML>

```

- Η χρήση του CDATA. Υπάρχει επίσης η δυνατότητα της διατήρησης του πρωτότυπου κειμένου με την εισαγωγή του στην ενότητα CDATA που είναι απολύτως αδιαφανής από τον επεξεργαστή NewsML. Το χαρακτηριστικό αυτό κρίνεται ιδιαίτερα χρήσιμο όταν το κείμενο εισάγεται με τη χρήση της HTML και όχι της XML, δεν συνίσταται όμως όταν πρόκειται να γίνει χρήση της XML.

```

<ContentItem>
  <MediaType FormalName="Text"/>
  <Format FormalName="HTML"/>
  <DataContent>
    <![CDATA[
      <html>...</html>
    ]>
  </DataContent>
</ContentItem>

```

Η επιλογή αυτή της αξιοποίησης της ενότητας CDATA είναι δυνατή μόνο στην περίπτωση που το ίδιο το περιεχόμενο δεν περιέχει δική του ενότητα CDATA ή έστω δεν περιέχει τη εξής σειρά χαρακτήρων «]]> ». Για λόγους συνέπειας και συνοχής της NewsML, δεν συνίσταται η χρήση της ενότητας CDATA για ενσωματωμένα XML τεκμήρια, καθώς όταν γίνεται η επεξεργασία της ενότητας CDATA που περιέχει δεδομένα σε XML πιθανόν να διαφέρουν από το πραγματικό κείμενο.

- Η εισαγωγή καθαρού κειμένου. Η συγκεκριμένη επιλογή αποτελεί την τελευταία λύση που προτείνεται από τη NewsML και είναι δυνατή με την προϋπόθεση ότι δεν υπάρχει στο πρωτότυπο κείμενο κανένας χαρακτήρας «mark-up».

```
<ContentItem>
  <MediaType FormalName="Text"/>
  <MimeType FormalName="text/plain"/>
  <DataContent>The quick brown fox jumped over the lazy dog. The lazy dog & his aunt showed
quite some surprise! "That's more than (>) I was bargaining for," she said.</DataContent>
</ContentItem>
```

Η παραπομπή στο περιεχόμενο

Το γνώρισμα Href δέχεται ως τιμή ένα URL (Uniform Resource Locator) που υποδεικνύει επακριβώς τη θέση του περιεχομένου από την οποία μπορεί να ανακτηθεί. Η συγκεκριμένη μέθοδος συνίσταται όταν μάλιστα τα δεδομένα μας είναι σε δυαδική μορφή.

```
<ContentItem Href="http://www.mycompany.com/images/BER90-082802a.jpg">
  <MediaType FormalName="Photo"/>
  <Format FormalName="JPEG Baseline"/>
  <Characteristics>
    <Property FormalName="Width" Value="2048"/>
    <Property FormalName="Height" Value="1543"/>
  </Characteristics>
</ContentItem>
```

Στο παραπάνω παράδειγμα, βλέπουμε ότι τα δεδομένα μας είναι διαθέσιμα μέσω ενός URL που στηρίζεται στο πρωτόκολλο μεταφοράς http, αν και οποιαδήποτε άλλη μορφή URL είναι εφικτή. Παραδείγματος χάρη, ο σύνδεσμος που οδηγεί στα πραγματικά δεδομένα μπορεί να είναι πολύ απλούστερος όπως <ContentItem Href="./BER90-082802a.jpg">. Η μορφή αυτή χρησιμοποιείται βέβαια μόνο στην περίπτωση που το τεκμήριο της NewsML και η πραγματική πηγή των δεδομένων βρίσκονται κάτω από το ίδιο directory.

Στο σημείο αυτό χρήσιμο θεωρείται να διευκρινιστεί το γεγονός ότι το URL δεν αποτελεί σε καμιά περίπτωση ένα μοναδικό αριθμό ταυτοποίησης. Για κάθε πηγή υπάρχει η δυνατότητα της επανάληψης, με αποτέλεσμα ταυτόσημοι πόροι να υποδεικνύονται από διαφορετικά URLs. Ακόμα κι αν ένας πόρος δεν επαναλαμβάνεται, υπάρχει η πιθανότητα να υποδεικνύεται από διαφορετικά

URLs με τη χρήση εναλλακτικής σύνταξης π.χ. με αριθμούς IP ή με χώρους ονομάτων (host names).

Οι ιδιότητες ενός ContentItem

Τα ContentItems μπορούν να κατηγοριοποιηθούν ποικιλοτρόπως. Ο τύπος του μέσου προσδιορίζεται στο υποστοιχείο MediaType κάνοντας χρήση του γνωρίσματος FormalName που συνδέεται με τα γνωρίσματα Vocabulary και Scheme. Οι τιμές που εισάγονται στο γνώρισμα FormalName είναι περιορισμένες και προέρχονται από ένα ελεγχόμενο λεξιλόγιο που προβλέπει η γλώσσα. Ενδεικτικά αναφέρονται τα εξής μορφότυπα: NITF, XHTML, JPEG Baseline, Waveform Audio, FLA, MPEG. Με αυτό τον τρόπο ο παραλήπτης του εγγράφου της NewsML γνωρίζει το είδος του τεκμηρίου και κατ' επέκταση τον τρόπο επεξεργασίας του.



Όμοιο ρόλο διαδραματίζει το στοιχείο MimeType. Για παράδειγμα το συγκεκριμένο στοιχείο παίρνει τιμές ως εξής: **"text/vnd.IPTC.NITF+xml"**, **"image/jpeg"**, **"audio/x-wav"**, **"video/mpeg"**.

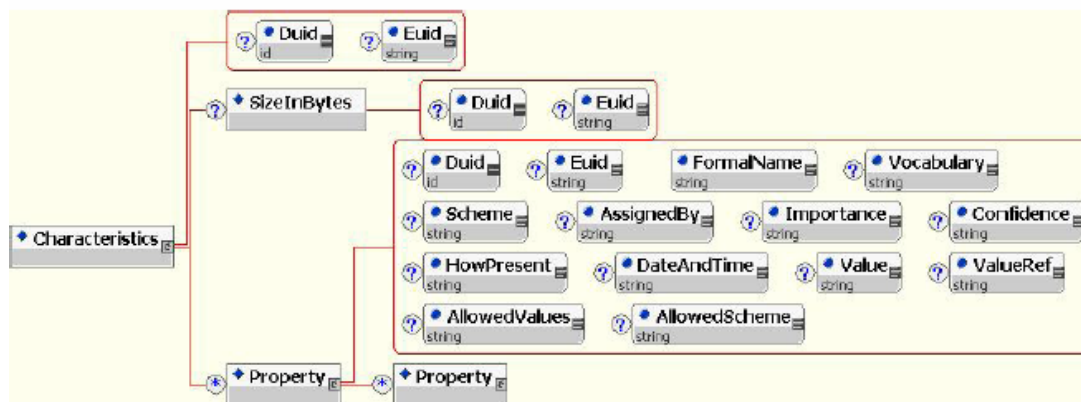
Η ίδια πληροφορία σχετικά με τη δομή του περιεχομένου μπορεί να δηλωθεί επίσης στο στοιχείο **Notation**. Το Διεθνές Συμβούλιο Τύπου και Τηλεπικοινωνιών (IPTC) προσδιορίζει τρία εναλλακτικά σχήματα για το λεξιλόγιο του συγκεκριμένου γνωρίσματος, που είναι τα εξής: IPTCNotation scheme, το Formal Public Identifier scheme και το NewsML URN scheme. Παραδείγματος χάρη, λαμβάνει την τιμή «NITF» από το σχήμα IptcNotation scheme, ή το μορφότυπο δηλώνεται ως εξής «-//IPTC-NAA//DTD NITF-XML 1.0//EN» όταν προέρχεται από το σχήμα Formal Public Identifier. Στην περίπτωση γίνεται χρήση του NewsML URN scheme στο γνώρισμα Notation μπαίνει η τιμή με την εξής μορφή «urn:newsml:iptc.org:20001006:NITF».

Η δήλωση του format είτε στο στοιχείο MediaType είτε στο στοιχείο MimeType συνίσταται από την NewsML ως ορθή τακτική. Στην περίπτωση

βέβαια που χρησιμοποιούνται και τα δύο αυτά στοιχεία, τότε θα πρέπει σαφώς να δηλώνουν το ίδιο ακριβώς μορφότυπο. Σε περίπτωση βέβαια που για οποιοδήποτε λόγο οι τιμές των δύο στοιχείων διαφέρουν, τότε λαμβάνεται υπόψη μόνο η τιμή που δίδεται στο MediaType. Η χρήση του γνωρίσματος Notation συνίσταται ως έσχατη λύση από την NewsML, καθώς η καταγωγή του γνωρίσματος είναι από τη SGML και πολλοί λίγοι είναι αυτοί που το χρησιμοποιούν.

Χαρακτηριστικά του περιεχομένου

Το ContentItem υποστηρίζει διαφορετικά προαιρετικά χαρακτηριστικά (**Characteristics**) επιτρέποντας με αυτόν τον τρόπο στους εκδότες να αποφασίσουν ποια από αυτά θα χρησιμοποιήσουν για να προσδιορίσουν με τον καλύτερο δυνατό τρόπο τα φυσικά χαρακτηριστικά του περιεχομένου.



Το υποστοιχείο **SizeInBytes** αποτελεί το μοναδικό στοιχείο από την κατηγορία της φυσικής περιγραφή που ορίζεται από την IPTC, καθώς η πληροφορία αυτή είναι κοινή για όλους τους τύπους πληροφοριακών μέσων.

Format version

Το υποστοιχείο DataContent του ContentItem μπορεί να εμφανίζεται σε ποικίλα μορφότυπα. Τα προαναφερόμενα μπορούν να τροποποιούνται και συχνά εμφανίζονται σε ανανεωμένες εκδοχές. Η εκάστη έκδοση δηλώνεται στο υποστοιχείο Property του Characteristics όπως φαίνεται στο παράδειγμα που ακολουθεί.

```

<ContentItem>
  <MediaType FormalName="Text"/>
  <Format FormalName="NITF"/>
  <Characteristics>
    <Property Scheme="CharacteristicsProperty" FormalName="FormatVersion" Value="3.1"/>
  </Characteristics>
  <DataContent>
    <nitf>.....</nitf>
  </DataContent>
</ContentItem>

```

Χαρακτηριστικά διαφορετικών μέσων

- **Κείμενο:** όταν η πληροφορία μας εκφράζεται με τη μορφή κειμένου τα χαρακτηριστικά που προσδιορίζονται από την IPTC είναι τα εξής: λέξεις (Words), αλφάβητο (Alphabet), γραμματοσειρά (Font).
- **Φωτογραφία:** στην περίπτωση των φωτογραφιών τα προβλεπόμενα χαρακτηριστικά διαφέρουν και είναι τα εξής: Ποιότητα (Quality level), πλάτος (Width), ύψος (Height), χρώματα (ColorSpace), ανάλυση (PixelDepth), περιστροφή (Rotation) και ICCProfile.

```

<ContentItem Href="http://www.company.com/photo.jpg">
  <MediaType FormalName="Photo" />
  <Format FormalName="JPEG Baseline" />
  <Characteristics>
    <SizeInBytes>401230</SizeInBytes>
    <Property FormalName="Width" Value="3072"/>
    <Property FormalName="Height" Value="2536" />
    <Property FormalName="ColorSpace" Value="RGB"/>
    <Property FormalName="ICCProfile" Value="Nikon D1"/>
  </Characteristics>
</ContentItem>

```

- **Γραφικό:** όταν το πληροφοριακό αντικείμενο είναι γραφικό τα χαρακτηριστικά που προβλέπονται σχετίζονται με τον τύπο της δυναμικής κίνησης (AnimationType), η ανάλυση (Resolution) και η αναλογία ύψους-πλάτους (HeightWidthRatio). Συχνά βέβαια τα γραφικά δανείζονται χαρακτηριστικά από τις φωτογραφίες όπως πλάτος (Width) και ύψος (Height), ενώ όταν πρόκειται για διαδραστικό animation (π.χ. Macromedia Flash) η διάρκεια (Duration) αποτελεί χρήσιμο χαρακτηριστικό περιγραφής.
- **Ήχος:** τα ηχητικά αντικείμενα προσδιορίζονται από τα εξής χαρακτηριστικά: κωδικοποίηση ήχου (AudioCoder), εκδοχή κωδικοποιητή ήχου (AudioCoderVersion), συνολική διάρκεια (TotalDuration), μέση

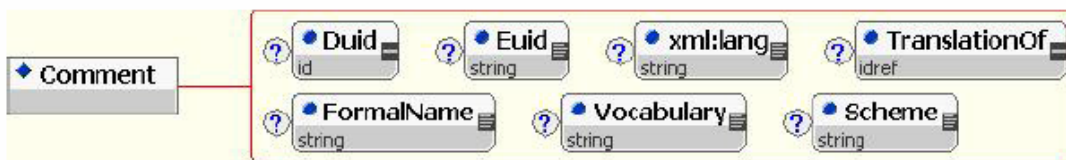
συχνότητα bits (AverageBitRate), μέγεθος (SampleSize), συχνότητα (SampleRate) και τα κανάλια ήχου (AudioChannels).

```
<ContentItem Href="http://www.company.com/audio-report.mp3">
  <MediaType FormalName="Audio" />
  <Format FormalName="MP3" />
  <Characteristics>
    <SizeInBytes>123456</SizeInBytes>
    <Property FormalName="FileExtension" Value=".mp3"/>
    <Property FormalName="AudioCoder" Value="MP3" />
    <Property FormalName="AudioCoderVersion" Value="1.0"/>
    <Property FormalName="TotalDuration" Value="30"/>
    <Property FormalName="AverageBiteRate" Value="128"/>
    <Property FormalName="SampleSize" Value="16"/>
    <Property FormalName="SampleRate" Value="44.100"/>
    <Property FormalName="AudioChannels" Value="2"/>
  </Characteristics>
</ContentItem>
```

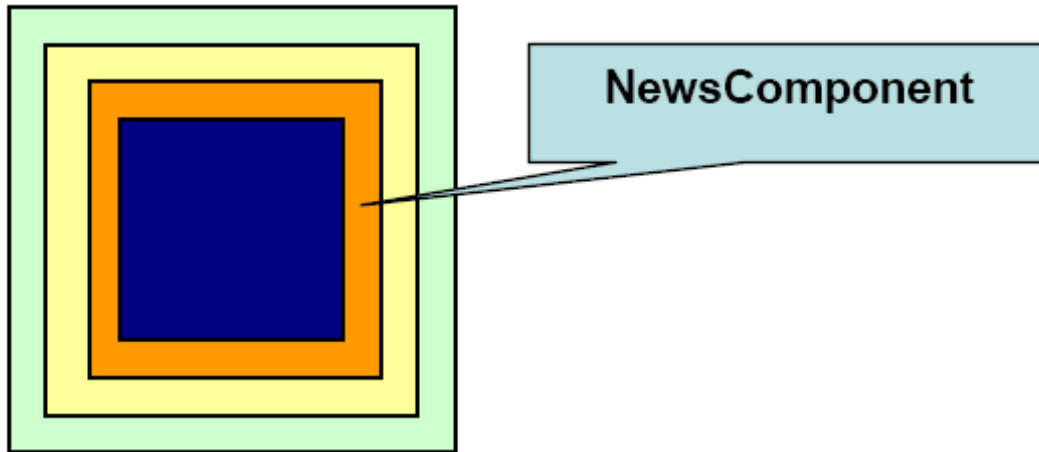
- **Βίντεο:** τέλος το βίντεο προσδιορίζεται από χαρακτηριστικά όπως ο κωδικοποιητής βίντεο (VideoCoder), η εκδοχή του τελευταίου (VideoCoderVersion), τα Vbr, το πλάτος (Width), το ύψος (Height), η συνολική διάρκεια (TotalDuration), το σύνολο των frames (FramesTotal), τα KeyFrames, η ανάλυση των pixels (PixelDepth), ο ρυθμός των frames (FrameRate), μέση συχνότητα bits (AverageBitRate), η δειγματοληψία (Sampling) και τέλος ο Redirector.

Το στοιχείο Comment

Όπως πολλά από τα στοιχεία της NewsML, έτσι και το ContentItem μπορεί να έχει το υποστοιχείο Comment μία ή περισσότερες φορές που περικλείει επιπρόσθετες πληροφορίες για εσωτερική χρήση στην εκδοτική διαδικασία.



Το επίπεδο δομής - NewsComponent

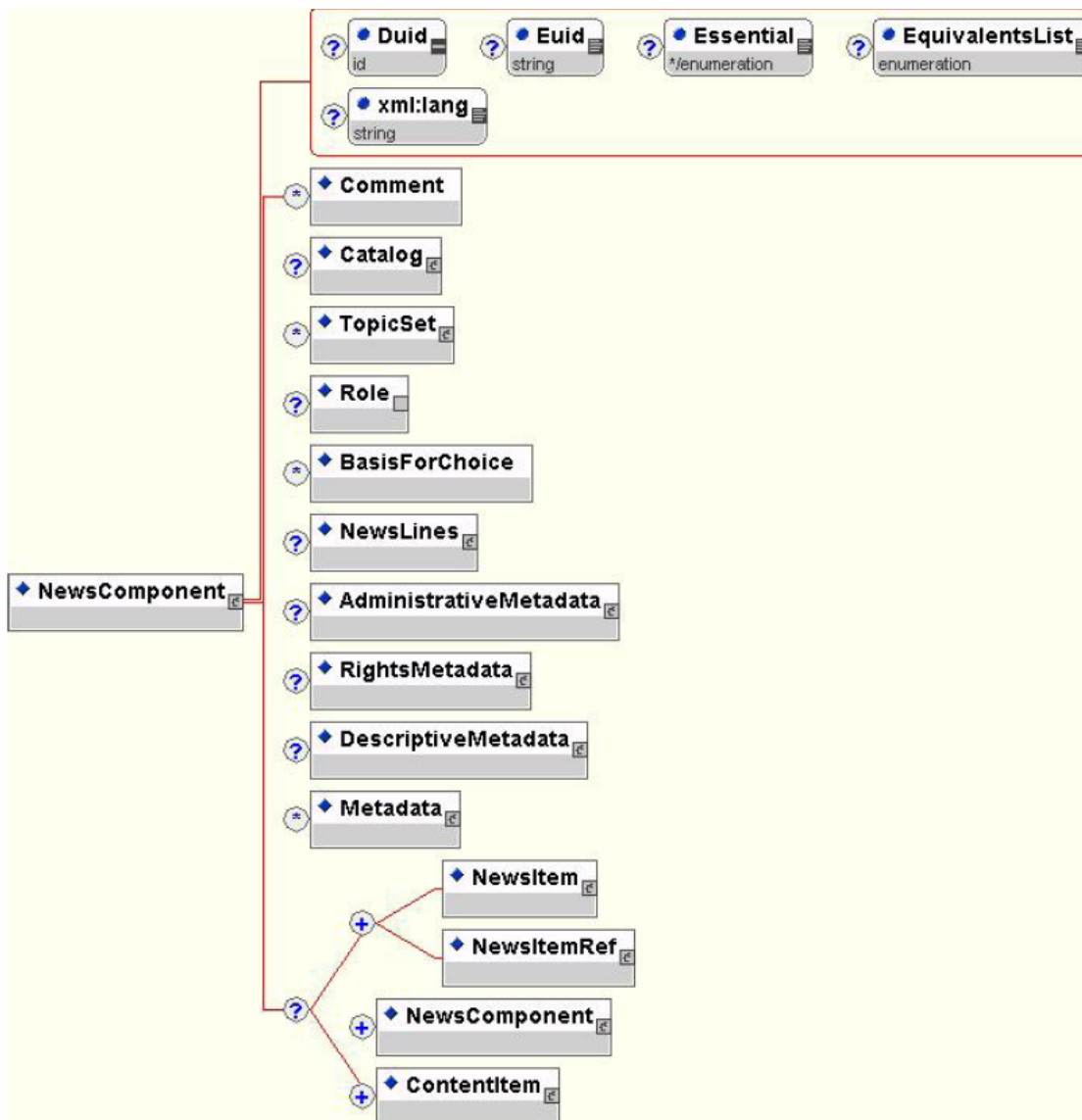


Πάνω από το επίπεδο περιεχομένου είναι το επίπεδο δομής, που αντιπροσωπεύεται από το NewsComponent. Το NewsComponent εξυπηρετεί πολλούς σκοπούς σε σχέση με το NewsItem:

1. Ενεργεί ως περιέκτης - **container** για τα news objects (π.χ. ContentItems, NewsComponents ή NewsItems) στη συνέχεια θα καλούμε αυτά τα αντικείμενα ως συστατικά - **constituents** του NewsComponent.
2. Επιτρέπει την επισύναψη των μεταδεδομένων τα οποία είναι σχετικά.
3. Επιτρέπει την επισύναψη των **NewsLines** τα οποία περιέχουν πληροφορίες αναγνώσιμες από τον άνθρωπο.

Η δομή του NewsComponent

Το NewsComponent προσθέτει δομή προκειμένου να συνθέσει ένα τεκμήριο. Ομαδοποιεί σχετικά news objects σε σχετικές συλλογές, τα οποία κατά κάποιο τρόπο ανήκουν στην ίδια ομάδα. Για παράδειγμα το NewsComponent μπορεί να αναπαραστήσει μία φωτογραφία ή πολλά κείμενα σε διαφορετικές γλώσσες. Τα NewsComponents δεν είναι κάτι το οποίο μπορεί να δημιουργηθεί. Κολλούν μαζί αντικείμενα τα οποία μοιράζονται κάτι (μοιράζονται μεταδεδομένα ή ομαδοποιούνται σε λογικές συλλογές).



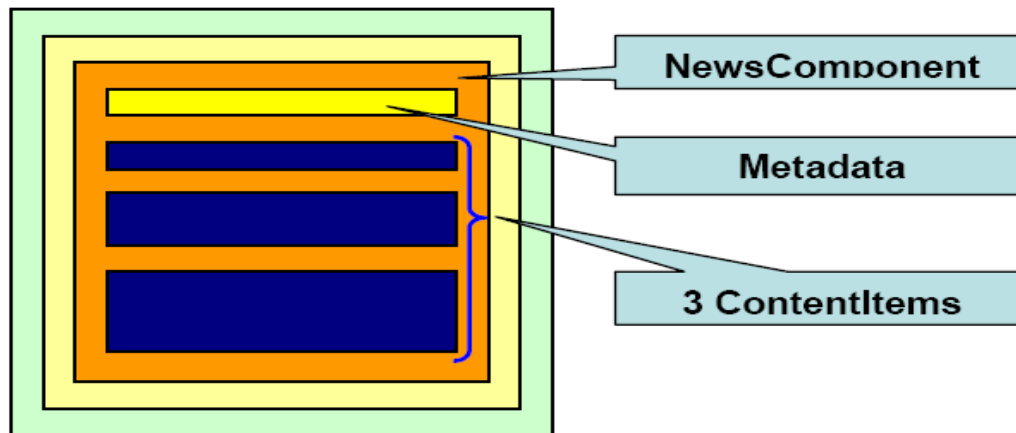
Τα **NewsComponents** μπορούν να αντιπροσωπεύουν συλλογές δύο ειδών: τα συστατικά μπορούν να είναι εναλλακτικά μεταξύ τους. (π.χ. an equivalents list) ή συμπληρωματικά μεταξύ τους can be complementary to each other (π.χ. a complementary list). Το **NewsComponents** μπορούν επίσης να δηλώσουν τι ρόλο παίζουν τα συστατικά στο **NewsItem** όπου εμφανίζονται.

Αντικείμενα του **NewsComponent**

Το **NewsComponent** μπορεί να περιέχει τα **NewsItems**, **NewsItemRefs**, **NewsComponents** ή **ContentItems**. **NewsItems** και **NewsItemRefs** μπορούν να

συνδυαστούν μεταξύ τους αλλά άλλοι συνδυασμοί πέρα από αυτόν δεν είναι επιτρεπτοί.

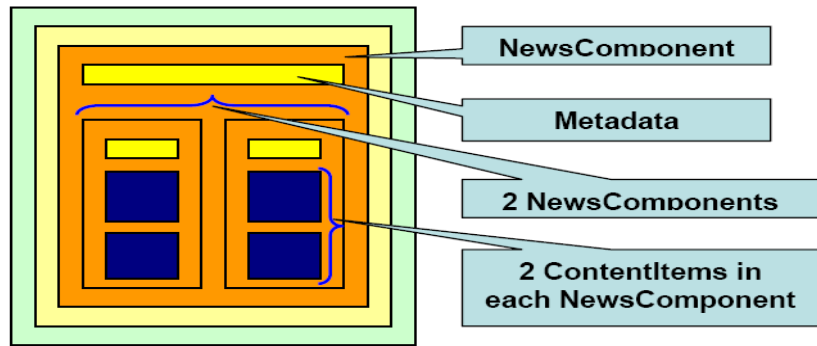
Παράδειγμα 1^ο: το NewsComponent έχει τρία ContentItems. Αυτά μπορεί να είναι τρεις εικόνες σε τρεις διαφορετικές μορφές (thumbnail, preview και full-resolution images).



```
<NewsComponent>
  <AdministrativeMetadata>
    <Creator>
      <Party FormalName="john.doe@mycompany.com"
Vocabulary="urn:newsml:mycompany.com:20031010:people"/>
      <Contribution FormalName="Photographer"/>
    </Creator>
  </AdministrativeMetadata>
  <DescriptiveMetadata>
    <SubjectCode><Subject FormalName="1506206"/></SubjectCode>
  </DescriptiveMetadata>
  <ContentItem Href="/photo1-thumbnail.jpg">
    <MediaType FormalName="Photo"/>
    <Format FormalName="JPEG Baseline"/>
    <Characteristics>
      <Property FormalName="Width" Value="128"/>
      <Property FormalName="Height" Value="98"/>
    </Characteristics>
  </ContentItem>
  <ContentItem Href="/photo1-preview.jpg">
    <MediaType FormalName="Photo"/>
    <Format FormalName="JPEG Baseline"/>
    <Characteristics>
      <Property FormalName="Width" Value="512"/>
      <Property FormalName="Height" Value="430"/>
    </Characteristics>
  </ContentItem>
  <ContentItem Href="/photo1-full.jpg">
    <MediaType FormalName="Photo"/>
    <Format FormalName="JPEG Baseline"/>
    <Characteristics>
      <Property FormalName="Width" Value="2048"/>
      <Property FormalName="Height" Value="1543"/>
    </Characteristics>
  </ContentItem>
</NewsComponent>
```

Παράδειγμα 2^ο:

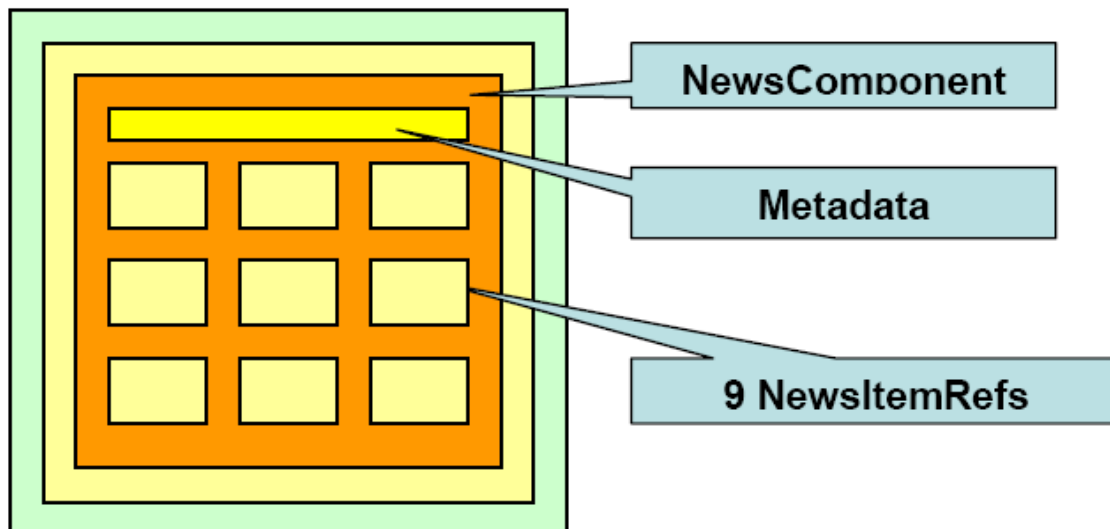
το NewsComponent έχει
δύο NewsComponents;
το κάθε ένα
περιλαμβάνει δύο
ContentItems. Αυτό
μπορεί να είναι μία



μικρή ή μεγάλη εκδοχή μίας ιστορίας στα ιταλικά καθώς το NewsComponent
έχει δύο ισοδύναμα ContentItems.

```
<NewsComponent>
  <AdministrativeMetadata>
    <Provider>
      <Party FormalName="MyCompany"/>
    </Provider>
  </AdministrativeMetadata>
  <NewsComponent EquivalentsList="yes">
    <AdministrativeMetadata ...>
      <Creator>
        <Party FormalName="Enzo Massimo"/>
      </Creator>
    </AdministrativeMetadata>
    <DescriptiveMetadata>
      <Language FormalName="it"/>
    </DescriptiveMetadata>
    <ContentItem>
      <MediaType FormalName="Text"/>
      <Characteristics>
        <SizeInBytes>266</SizeInBytes>
      </Characteristics>
      <DataContent> ... </DataContent>
    </ContentItem>
    <ContentItem>
      <MediaType FormalName="Text"/>
      <Characteristics>
        <SizeInBytes>1356</SizeInBytes>
      </Characteristics>
      <DataContent> ... </DataContent>
    </ContentItem>
  </NewsComponent>
  <NewsComponent EquivalentsList="yes">
    <Creator>
      <Party FormalName="John Doe" />
    </Creator>
    <ContentItem Href="/photo1-thumbnail.jpg">
      <MediaType FormalName="Photo"/>
      <Format FormalName="JPEG Baseline"/>
      <Characteristics>
        <Property FormalName="Width" Value="128"/>
        <Property FormalName="Height" Value="98"/>
      </Characteristics>
    </ContentItem>
    <ContentItem Href="/photo1-preview.jpg">
      <MediaType FormalName="Photo"/>
      <Format FormalName="JPEG Baseline"/>
      <Characteristics>
        <Property FormalName="Width" Value="512"/>
        <Property FormalName="Height" Value="430"/>
      </Characteristics>
    </ContentItem>
  </NewsComponent>
</NewsComponent>
```


Παράδειγμα 3^ο: το NewsComponent έχει 9 NewsItemRefs. Αυτό μπορεί να είναι μία λίστα από NewsItems.



```
<NewsComponent>
  <AdministrativeMetadata>
    <Provider>
      <Party FormalName="MyCompany"/>
    </Provider>
  </AdministrativeMetadata>
  <NewsItemRef NewsItem="urn:newsml:mycompany.com:20031010:25ab03:1"/>
  <NewsItemRef NewsItem="urn:newsml:mycompany.com:20031010:26cb02:1"/>
  <NewsItemRef NewsItem="urn:newsml:mycompany.com:20031010:03bg23:1"/>
  <NewsItemRef NewsItem="urn:newsml:mycompany.com:20031010:06kl43:1"/>
  <NewsItemRef NewsItem="urn:newsml:mycompany.com:20031010:58fn34:1"/>
  <NewsItemRef NewsItem="urn:newsml:mycompany.com:20031010:38ke56:1"/>
  <NewsItemRef NewsItem="urn:newsml:mycompany.com:20031010:27hn89:1"/>
  <NewsItemRef NewsItem="urn:newsml:mycompany.com:20031010:46ab34:1"/>
  <NewsItemRef NewsItem="urn:newsml:mycompany.com:20031010:25av03:1"/>
</NewsComponent>
```

Τοπικός προσδιορισμός του NewsComponent

Ο τοπικός προσδιοριστής (identifier) – το γνώρισμα **Duid** (attribute) – συνδέεται συνήθως με το NewsComponent έτσι ώστε να μπορεί να δείχνεται από άλλο σημείο του τεκμηρίου.

```

<NewsComponent>
  <NewsComponent>
    <ContentItem>
      <MediaType FormalName="Text"/>
      <Format FormalName="NITF"/>
      <DataContent>
        <nitf><body><body.content>
          <media media-type="image">
            <media-reference source="#photo0"/>
          </media>
          <p>...</p>
        </body.content></body></nitf>
      </DataContent>
    </ContentItem>
  </NewsComponent>
  <NewsComponent Duid="photo0">
    <ContentItem Href="/photo0.jpg"/>
  </NewsComponent>
</NewsComponent>

```

Καθώς το NewsComponents δεν έχουν έναν μοναδικό παγκόσμιο προσδιοριστή, δεν θεωρούνται πηγές του διαδικτύου, δεν είναι επαναχρησιμοποιήσιμοι στο NewsML περιβάλλον: δεν υπάρχει τρόπος να γνωρίζουμε αν δύο NewsComponents είναι ίδια, και δεν υπάρχει άμεσος τρόπος στο να αναφερθούμε σε ένα εξωτερικό NewsComponent. Οι παγκόσμιοι προσδιοριστές προσδιορίζονται στο επίπεδο NewsItem level.

Ιδιότητες του container

Κάθε NewsComponent σε ένα τεκμήριο ειδήσεων μπορεί να περιέχει ένα στοιχείο ρόλου. Ο ρόλος είναι το διακριτό χαρακτηριστικό του NewsComponent ή της σχέσης του με τα άλλα με τα οποία συσχετίζεται στην ίδια ομάδα του NewsComponent.

Το στοιχείο ρόλος είναι κενό. Η αξία του αντιπροσωπεύεται μέσω του FormalName και των σχετιζόμενων ιδιοτήτων λεξιλογίου.



Σε ένα ευρύτερο πλαίσιο, δεν μιλάμε για τον ρόλο του NewsComponent τον ίδιο, αλλά των αντικειμένων ειδήσεων που είναι ενσωματωμένα στο NewsComponent στο πλαίσιο του περιέκτη του NewsComponent. Αυτός ο περιέκτης είναι μία συμπληρωματική λίστα. Στην περίπτωση λίστας ομοίων αντικειμένων, όλα τα στοιχεία της λίστας μοιράζονται το ρόλο του στοιχείου που τα περιέχει.

Είναι συχνά χρήσιμο να διαχωρίζουμε τα συστατικά του NewsComponent από το ρόλο τους. Οι ρόλοι μπορούν να δείξουν την πιθανή χρήση ενός αντικειμένου σε μία εφαρμογή client. Η εφαρμογή μπορεί να επιλέξει όσα περισσότερα συστατικά από τη λίστα που ανταποκρίνονται σε αυτή, όπως προσδιορίζεται από το ρόλο τους.

Προκειμένου να γίνει αυτό, κάθε συστατικό του NewsComponent μπορεί να είναι από μόνο του ένα NewsComponent περιλαμβάνοντας ένα στοιχείο ρόλου. Μερικά NewsComponents δημιουργούνται μόνο προκειμένου να συνδεθεί ο ρόλος με το περιεχόμενο.

Για παράδειγμα, σε ένα τεκμήριο ειδήσεων το οποίο έχει δύο NewsComponents και αναπαριστά φωτογραφίες, ο ρόλος της πρώτης εικόνας μπορεί να είναι 'Thumbnail', και ο ρόλος της δεύτερης 'Preview'. Ένα NewsComponent που αντιπροσωπεύει ένα άρθρο μπορεί να έχει 'Main' Role και ένα άλλο να έχει 'Sidebar' Role.

```
<NewsComponent>
  <NewsComponent>
    <Role FormalName="Thumbnail"/>
    <ContentItem Href="http://www.mycompany.com/content/images/photo1-th.jpg" />
  </NewsComponent>
  <NewsComponent>
    <Role FormalName="Preview"/>
    <ContentItem Href="http://www.mycompany.com/content/images/photo1-pr.jpg" />
  </NewsComponent>
</NewsComponent>
```

Το IPTC διατηρεί ένα σύνολο από αξίες ρόλων (π.χ. 'Main', 'Supporting', 'Thumbnail', 'Preview', 'Caption', 'Abstract' etc.), σύμφωνα με τις περιγραφές τους σε ένα σύνολο θεμάτων στο δικό του ιστοχώρο. Οι αξίες για αυτά τα μεταδεδομένα είναι ουδέτερες προκειμένου να αντιπροσωπεύουν όλα τα μέσα που περιέχουν περιεχόμενο π.χ. audio ή video.

```

<NewsComponent EquivalentsList="yes">
  <NewsComponent>
    <Role FormalName="Main"/>
    <NewsItemRef NewsItem="urn:newsml:mycompany.com:20031010:03bg23:1"/>
  </NewsComponent>
  <NewsComponent>
    <Role FormalName="Abstract"/>
    <ContentItem>
      <MediaType FormalName="Text"/>
      <DataContent> ... </DataContent>
    </ContentItem>
  </NewsComponent>
</NewsComponent>

```

Επιλογή ανάμεσα στα αντικείμενα

Μία λίστα από όμοια αντικείμενα πρέπει να έχει το στοιχείο **BasisForChoice**, το οποίο υποδεικνύει την ιδιότητα ή τις ιδιότητες που κατά την άποψη του παροχέα επιτρέπει στον αποδέκτη στο να κάνει μία επιλογή από τη λίστα. Η επιλογή επιτυγχάνεται διαμέσου του στοιχείου **BasisForChoice** που θα έχει ως αποτέλεσμα ένα και μόνο ένα αντικείμενο. Εάν δεν συμβεί αυτό, η λ' θση που συστήνεται είναι να επιλέγεται το πρώτο αντικείμενο, το **BasisForChoice** χρησιμοποιεί μία σύνταξη **XPath** προκειμένου να επιλεχθεί η ιδιότητάτο (Language, Format ...). Το πρώτο στοιχείο που θα φέρει η αναζήτηση αντιπροσωπεύει την πληροφορία που μπορεί να χρησιμοποιηθεί.

Παράδειγμα: Σε αυτό το παράδειγμα η επιλογή που πρέπει να γίνει είναι η ιδιότητα της γλώσσας.

```

<NewsComponent EquivalentsList="yes">
  <BasisForChoice>./DescriptiveMetadata/Language/@FormalName</BasisForChoice>
  <NewsComponent>
    <DescriptiveMetadata>
      <Language FormalName="de" />
    </DescriptiveMetadata>
    <ContentItem Href="http://www.mycompany.com/content/de/content-de.xml" />
  </NewsComponent>
  <NewsComponent>
    <DescriptiveMetadata>
      <Language FormalName="es" />
    </DescriptiveMetadata>
    <ContentItem Href="http://www.mycompany.com/content/es/content-es.xml" />
  </NewsComponent>
</NewsComponent>

```

Πολλά στοιχεία BasisForChoice μπορούν να προστεθούν. Σε αυτή την περίπτωση εάν πολλά αντικείμενα θεωρούνται ίδια επιτρέπεται στον παροχέα να τοποθετήσει νούμερα σύμφωνα με την σημασία που έχουν γι αυτόν βάζοντας το 1 σε αυτό που έχει μεγαλύτερη σημασία.

Παράδειγμα: Σε αυτό το παράδειγμα το πρώτο φίλτρο BasisForChoice με ιεραρχία 1 μπορεί να εφαρμοστεί στη γλώσσα και το δεύτερο στο επιδιωκόμενο κοινό μέσω του.

```
<NewsComponent EquivalentList="yes">
  <BasisForChoice Rank="1">./DescriptiveMetadata/Language/@FormalName</BasisForChoice>
  <BasisForChoice Rank="2">./ DescriptiveMetadata /OfInterestTo/@FormalName</BasisForChoice>
  <NewsComponent>
    <DescriptiveMetadata>
      <Language FormalName="de" />
      <OfInterestTo FormalName="CEO" />
    </DescriptiveMetadata>
    <ContentItem Href="http://www.mycompany.com/content/de/summary-de.xml" />
  </NewsComponent>
  <NewsComponent>
    <DescriptiveMetadata>
      <Language FormalName="de" />
      <OfInterestTo FormalName="Analyst" />
    </DescriptiveMetadata>
    <ContentItem Href="http://www.mycompany.com/content/de/analysis-de.xml" />
  </NewsComponent>
  <NewsComponent>
    <DescriptiveMetadata>
      <Language FormalName="es" />
      <OfInterestTo FormalName="CEO" />
    </DescriptiveMetadata>
    <ContentItem Href="http://www.mycompany.com/content/es/summary-es.xml" />
  </NewsComponent>
  <NewsComponent>
    <DescriptiveMetadata>
      <Language FormalName="es" />
      <OfInterestTo FormalName="Analyst" />
    </DescriptiveMetadata>
    <ContentItem Href="http://www.mycompany.com/content/es/analysis-es.xml" />
  </NewsComponent>
</NewsComponent>
```

Μεταδεδομένα περιεχομένου- NewsComponent

Στο προηγούμενο κεφάλαιο περιγράψαμε το στοιχείο NewsComponent ως ευέλικτο περιέκτη για αντικείμενα ειδήσεων. Στο κεφάλαιο αυτό θα δούμε το NewsComponent ως στοιχείο που περιέχει μεταδεδομένα για περιεχόμενο. Το στοιχείο NewsComponent φέρει μεταδεδομένα που περιγράφουν τα συστατικά του (π.χ. ContentItems, NewsComponents ή NewsItems).

Κλάσεις μεταδεδομένων

Η NewsML διαιρεί τον κόσμο των μεταδεδομένων σε κλάσεις στο επίπεδο NewsComponent σε τέσσερις κλάσεις:

- Διαχειριστικά μεταδεδομένα – πληροφορίες για το πακετάρισμα των αντικειμένων, τη δημιουργία του περιεχομένου κτλ.
- Περιγραφικά μεταδεδομένα – πληροφορίες για το περιεχόμενο
- Μεταδεδομένα δικαιωμάτων – πληροφορίες για τη χρήση και τα πνευματικά δικαιώματα του περιεχομένου
- Διάφορα – άλλα μεταδεδομένα

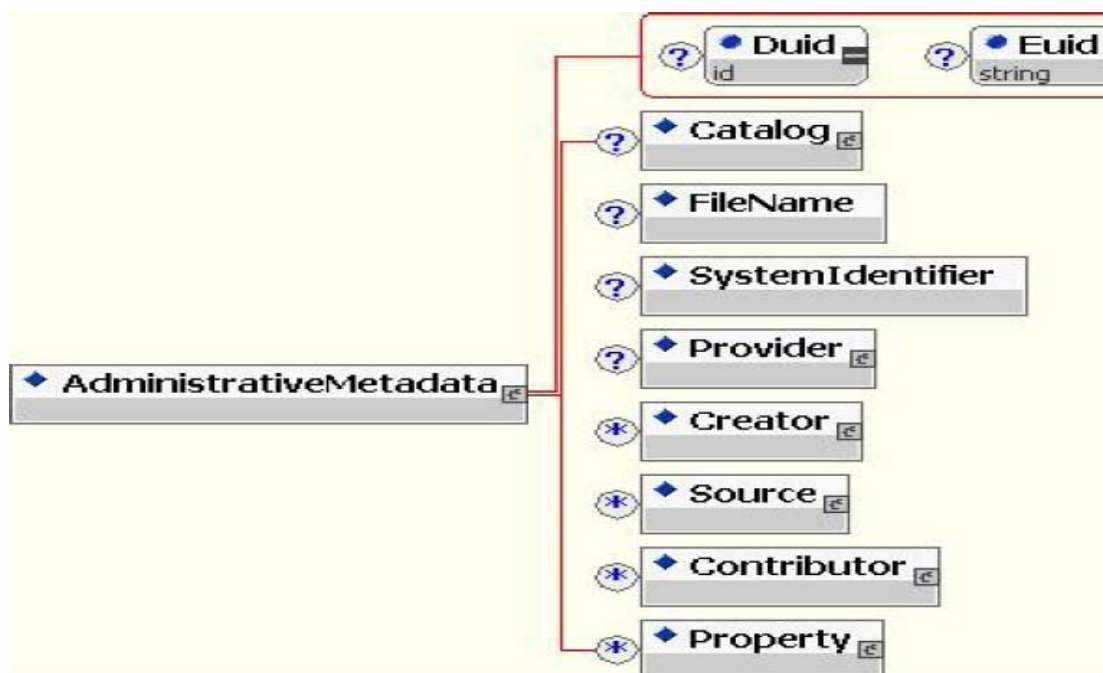
Η NewsML χρησιμοποιεί ελεγχόμενο λεξιλόγιο. Οι ελεγχόμενες τιμές μεταδεδομένων αναπαριστούνται στο FormalName που συνδέεται με προαιρετικό σχήμα και ιδιότητες λεξιλογίου.

«Κληρονομικότητα» των μεταδεδομένων

Το NewsComponent φέρει μεταδεδομένα που περιγράφουν τα συστατικά του. Τα μεταδεδομένα που επισυνάπτονται στο NewsComponent κληροδοτούνται στα συστατικά στοιχεία του. Προκειμένου να αποφευχθούν οι επαναλήψεις τα μεταδεδομένα τοποθετούνται κοντά στο στοιχείο ρίζας.

Διαχειριστικά μεταδεδομένα

Τα διοικητικά δεδομένα παρέχουν πληροφορίες που σχετίζονται με το πακετάρισμα των αντικειμένων ειδήσεων, π.χ. το `NewsItem`. Κάποια από αυτά τα στοιχεία παρέχουν πληροφορίες σχετικά με την δημιουργία του περιεχομένου. Τα δύο πρώτα στοιχεία περιγράφουν που μπορεί να βρεθεί η πληροφορία και τα άλλα στοιχεία περιγράφουν ποιος δημιούργησε την πληροφορία.



FileName και SystemIdentifier

Το προαιρετικό στοιχείο `FileName` προσδιορίζει το όνομα με το οποίο έχει αποθηκευτεί το αρχείο για ένα `NewsItem`.

Το προαιρετικό `SystemIdentifier` υποδεικνύει μία διεύθυνση (such as an http URL pointing at a file or a dynamic web page) όπου το `NewsItem` ή το `NewsML` μπορεί να βρεθεί.

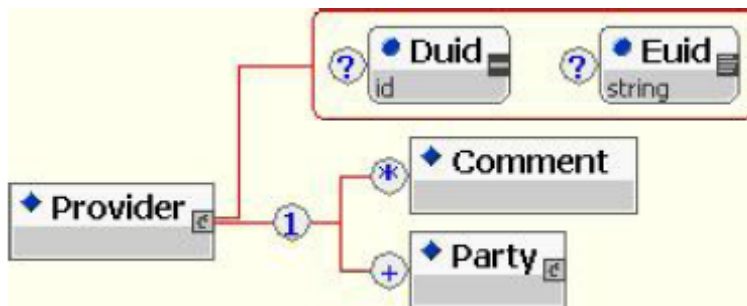
```

<NewsItem>
  <Identification> ... </Identification>
  <NewsManagement> ... </NewsManagement>
  <NewsComponent>
    <AdministrativeMetadata>
      <FileName>NewsMLStory.xml</FileName>
      <SystemIdentifier>http://www.mycompany.com/stories/NewsMLStory.xml</SystemIdentifier>
    </AdministrativeMetadata>
  </NewsComponent>
</NewsItem>

```

Παροχέας - Provider

Το προαιρετικό στοιχείο Provider προσδιορίζει το οργανισμό και την εταιρία όπου παρήγαγε το NewsItem και το έκανε διαθέσιμο.



Το υποχρεωτικό Party χρησιμοποιείται για να παρέχουμε περισσότερες πληροφορίες για το παροχέα, ενώ συγχρόνως θα πρέπει να σημειώσουμε ότι το στοιχείο Provider αντιστοιχεί στο στοιχείο 'publisher' του Dublin Core.

```

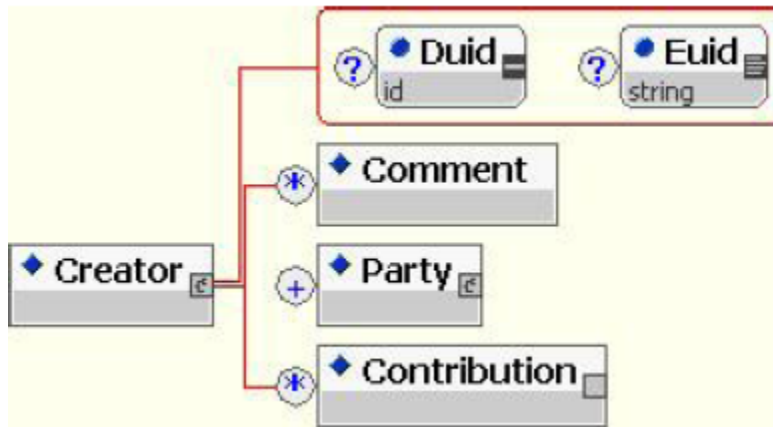
<NewsItem>
  <Identification> ... </Identification>
  <NewsManagement> ... </NewsManagement>
  <NewsComponent>
    <AdministrativeMetadata>
      <Provider>
        <Party FormalName="MyCompany" />
      </Provider>
    </AdministrativeMetadata>
  </NewsComponent>
</NewsItem>

```

Creator και Contributor

Κάποια στοιχεία αναφέρονται στο NewsItem όπως και στα παιδιά τους. Αυτή είναι η περίπτωση του στοιχείου Creator και Contributor, που έχουν όμοια δομή και βρίσκονται στο NewsComponent του NewsItem. Το προαιρετικό και επαναλαμβανόμενο στοιχείο Creator προσδιορίζει αυτόν που

δημιούργησε το NewsItem. Αν κάποια από τα παιδιά τους δεν φέρουν τα ίδια μεταδεδομένα είναι απαραίτητο να εισάγουμε ένα NewsComponent με AdministrativeMetadata.



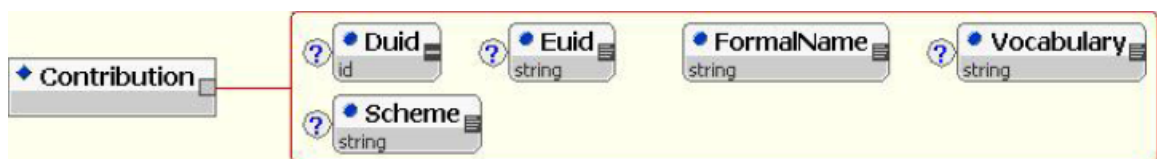
Το προαιρετικό και επαναλαμβανόμενο στοιχείο Contributor προσδιορίζει ένα άτομο ή ένα οργανισμό όπου έχει μετατρέψει ή έχει κάνει κάποιες αλλαγές ή έχει συμβάλει στο NewsItem ή στα συστατικά του NewsComponent μετά την δημιουργία.

Το Creator και το Contributor έχουν την ίδια σημασία με τα αντίστοιχα του Dublin Core.

Το προαιρετικό και επαναλαμβανόμενο στοιχείο Comment χρησιμοποιείται προκειμένου να εισάγει ο χρήστης ελεύθερο κείμενο σχετικά με τον δημιουργό ή τον συντελεστή.

Το υποχρεωτικό στοιχείο χρησιμοποιείται για πληροφορίες για τον δημιουργό και τον συντελεστή. Ο προσδιοριστής party είναι υποχρεωτικός χρησιμοποιώντας το FormalName.

Το προαιρετικό και επαναλαμβανόμενο στοιχείο **Contribution** δημιουργήθηκε από το NewsMLv1.2 προκειμένου να αναπαρασταθεί ο δημιουργός και ο συντελεστής.

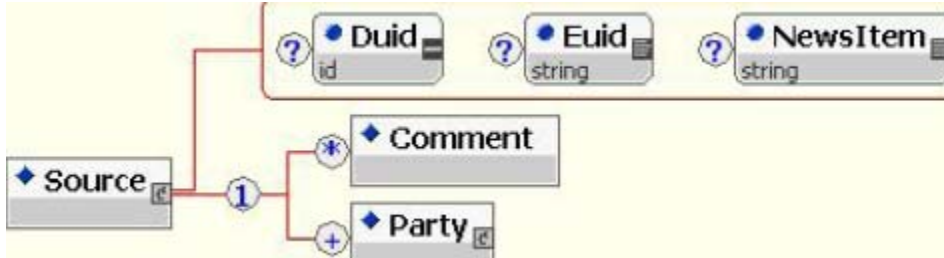


Το IPTC δεν έχει ακόμα προσδιορίσει ένα ελεγχόμενο λεξιλόγιο αλλά παραδείγματα τιμών είναι 'Reporter', 'Writer', 'Editor', 'Photographer', 'CaptionWriter', 'Translator'.

```
<NewsComponent>
  <AdministrativeMetadata>
    <Creator>
      <Party FormalName="john.doe@mycompany.com"
        Vocabulary="urn:newsml:mycompany.com:20031010:people"/>
      <Contribution FormalName="Writer"/>
    </Creator>
    <Contributor>
      <Comment xml:lang="fr">Traduction française</Comment>
      <Party FormalName="Julien Dupond" />
      <Contribution FormalName="Translator"/>
    </Contributor>
  </AdministrativeMetadata>
</NewsComponent>
```

Η πηγή

Το προαιρετικό και επαναλαμβανόμενο στοιχείο **Source** προσδιορίζεται ως «ένα άτομο ή ένας οργανισμός που παρείχε το υλικό για το object».

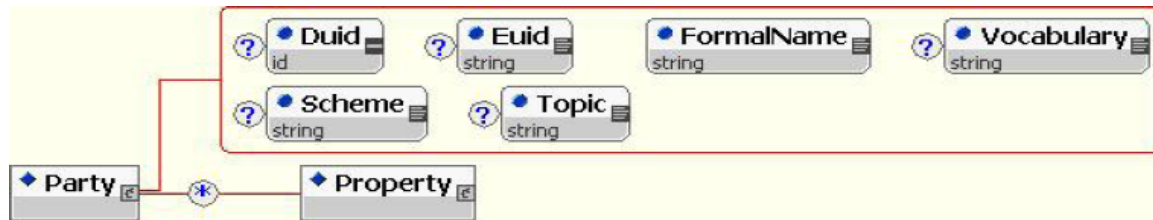


Το υποχρεωτικό **Party** χρησιμοποιείται για να παρέχουμε περισσότερες πληροφορίες για την πηγή. Και πάλι το **Comment** μπορεί να παρέχει πρόσθετες σχετικές πληροφορίες.

```
<NewsItem>
  <Identification> ... </Identification>
  <NewsManagement> ... </NewsManagement>
  <NewsComponent>
    <AdministrativeMetadata>
      <Source>
        <Party FormalName="ACompany" NewsItem="urn:newsml:acompany:20031010:NewsMLStory"/>
      </Source>
    </AdministrativeMetadata>
  </NewsComponent>
</NewsItem>
```

Το στοιχείο Party

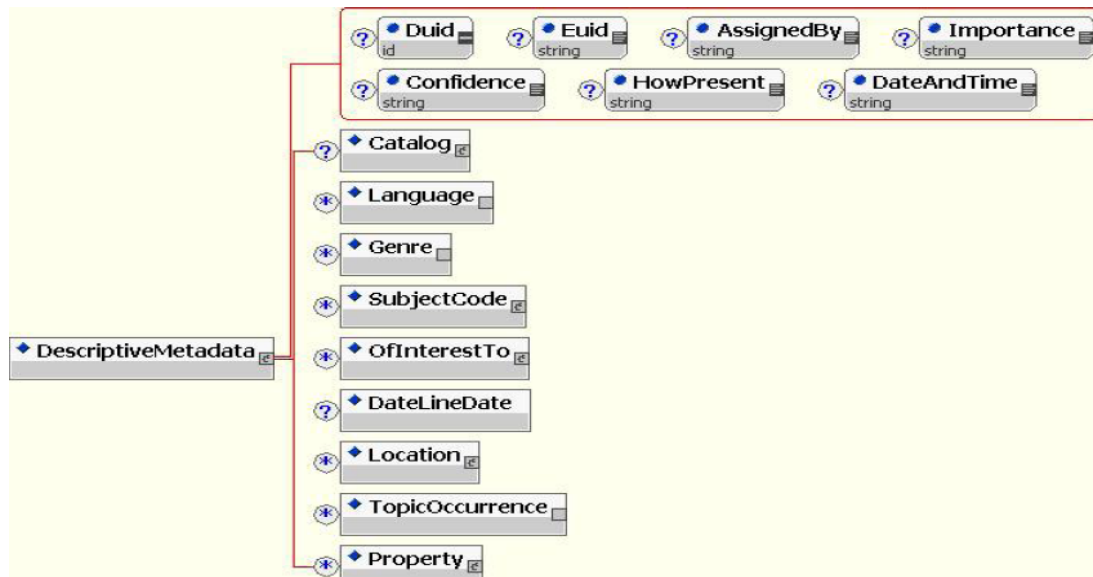
Το στοιχείο Party χρησιμοποιείται για πληροφορίες για την πηγή, τον παροχέα τον δημιουργό και για το συντελεστή.



```
<NewsComponent>
  <AdministrativeMetadata>
    <Creator>
      <Party FormalName="jdoe">
        <Property FormalName="vCard">
          <Property FormalName="FN" Value="John Doe"/>
          <Property FormalName="Email" Value="john.doe@mycompany.com"/>
          <Property FormalName="Tel" Value="+44 (20) 4455 6677"/>
        </Property>
      </Party>
      <Contribution FormalName="Writer"/>
    </Creator>
    <Contributor>
      <Comment xml:lang="fr">Traduction française</Comment>
      <Party FormalName="jdupond">
        <Property FormalName="vCard">
          <Property FormalName="FN" Value="Julien Dupond"/>
          <Property FormalName="Email" Value="jules.dupond@masociete.com"/>
          <Property FormalName="Tel" Value="+33 0144556677"/>
        </Property>
      </Party>
      <Contribution FormalName="Translator"/>
    </Contributor>
  </AdministrativeMetadata>
</NewsComponent>
```

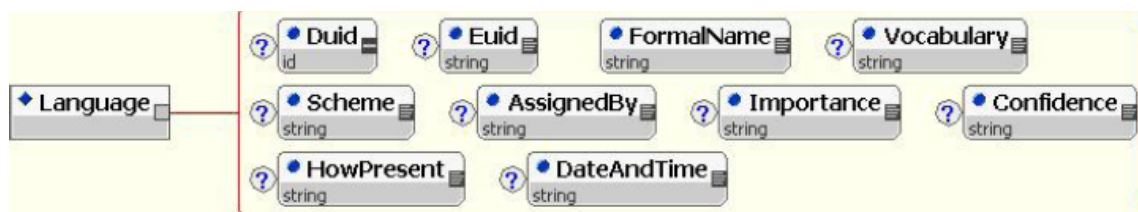
Περιγραφικά μεταδεδομένα

Τα περιγραφικά μεταδεδομένα παρέχουν πληροφορίες για το περιεχόμενο των συστατικών του NewsComponent.



Language

Το προαιρετικό και επαναλαμβανόμενο στοιχείο **Language** προσδιορίζει τη γλώσσα που χρησιμοποιείται στα συστατικά του NewsComponent και παρουσιάζει αντιστοιχία με το Dublin Core. Περισσότερες από μία γλώσσα μπορούν να οριστούν. Το πλήρες σύνολο των ετικετών γλώσσας - language tags – προσδιορίζεται από το **RFC 3066** (π.χ. 'en-US', 'fr-BE'. Το RFC 3066 έχει χρησιμοποιηθεί στη χρήση της `xml:lang` και για κάθε άλλο XML document.



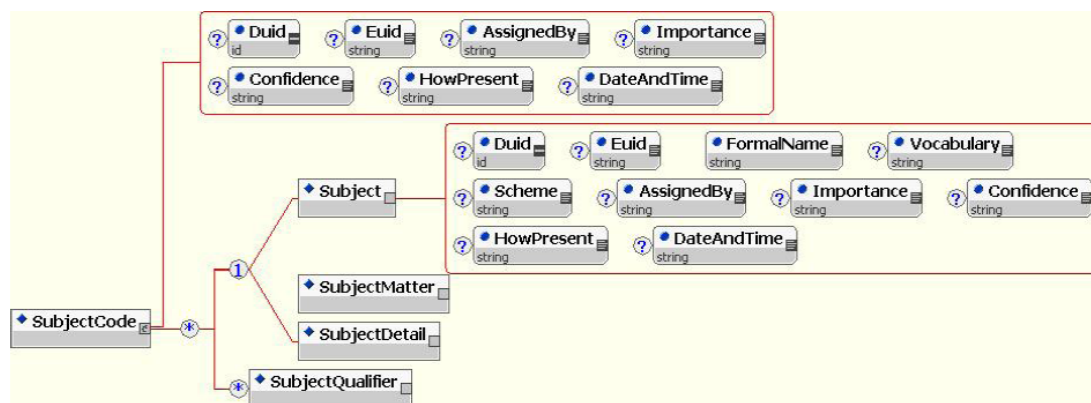
Το είδος

Το προαιρετικό και επαναλαμβανόμενο στοιχείο Genre υποδεικνύει το «στυλ της έκφρασης» που χρησιμοποιείται στο NewsComponent. Το στοιχείο Genre είναι κενό. Το κατάλληλο είδος- genre – αντιπροσωπεύεται από το FormalName. Το IPTC διατηρεί κάποιες τιμές Genre (π.χ. 'Current', 'Analysis', 'Feature', 'Obituary', 'Profile', 'Interview' etc.). Οι τιμές των μεταδεδομένων πρέπει να αντιπροσωπεύουν τα διαφορετικά μέσα. Το Genre έχει το ίδιο νόημα και στο Dublin Core που είναι η φύση ή το είδος του περιεχομένου της πηγής.

Το SubjectCode

Το προαιρετικό και επαναλαμβανόμενο στοιχείο SubjectCode προσδιορίζει το θέμα ή την κατηγορία των ειδήσεων που σχετίζονται με το NewsComponent. Το SubjectCode περιέχει τα Subject, SubjectMatter ή SubjectDetail και επίσης υποστηρίζει το στοιχείο SubjectQualifier. Το SubjectQualifier προσθέτει πληροφορίες στο θέμα και ειδικά στα θέματα που σχετίζονται με τα αθλήματα. Το SubjectCode έχει το ίδιο νόημα με το Dublin Core που είναι το θέμα του περιεχομένου της πηγής.

Τα Subject, SubjectMatter, SubjectDetail και SubjectQualifier έχουν όλα τη ίδια δομή.



```

<DescriptiveMetadata>
  <SubjectCode>
    <!-- sport, swimming, 50 m butterfly, final, men-->
    <Subject FormalName="1506206"/>
    <SubjectQualifier FormalName="15000001"/>
    <SubjectQualifier FormalName="15000024"/>
  </SubjectCode>
  <SubjectCode>
    <!-- applied science -->
    <Subject FormalName="13001000"/>
  </SubjectCode>
</DescriptiveMetadata>

```

DateLineDate και Location προέλευσης των ειδήσεων

Είναι σχετικά με τη δημιουργία των ειδήσεων π.χ την ημερομηνία που γράφτηκε η ιστορία, η τοποθεσία του αρχείου, η ημέρα και η τοποθεσία που τραβήχτηκε η φωτογραφία.

Το DateLineDate παρέχεται προκειμένου να προσδιοριστεί η ημερομηνία προέλευσης του ειδησεογραφικού αντικειμένου. Το περιεχόμενο του είναι σε ISO8601 Basic Date Format.

Αντίστοιχα το Location element παρουσιάζει την τοποθεσία προέλευσης του ειδησεογραφικού αντικειμένου.

```

<DescriptiveMetadata>
  <DateLineDate>20031010</DateLineDate>
  <Location HowPresent="Origin">
    <Property FormalName="Country" Value="US"/>
    <Property FormalName="CountryArea" Value="DC"/>
    <Property FormalName="City" Value="Washington"/>
    <Property FormalName="SubLocation" Value="The White House"/>
    <Property FormalName="WorldRegion" Value="North America"/>
  </Location>
</DescriptiveMetadata>

```

Τοποθεσία

Πέρα από την προέλευση, η δομή επιτρέπει την αναπαράσταση της τοποθεσίας που παρουσιάζεται μέσα στην είδηση. Η τιμή HowPresent πρέπει να δίνεται σύμφωνα με:

- RelatesTo: Τα συστατικά του NewsComponent να σχετίζονται με την τοποθεσία.
- Event: Τα συστατικά των NewsComponent αναφέρονται στα γεγονότα που λαμβάνουν χώρα στην τοποθεσία.

Θέμα ειδήσεων

Το στοιχείο TopicOccurrence υποδεικνύει το θέμα που έχουν τα στοιχεία του NewsComponent.

Το προαιρετικό HowPresent υποδεικνύει τη φύση της εμφάνισης.

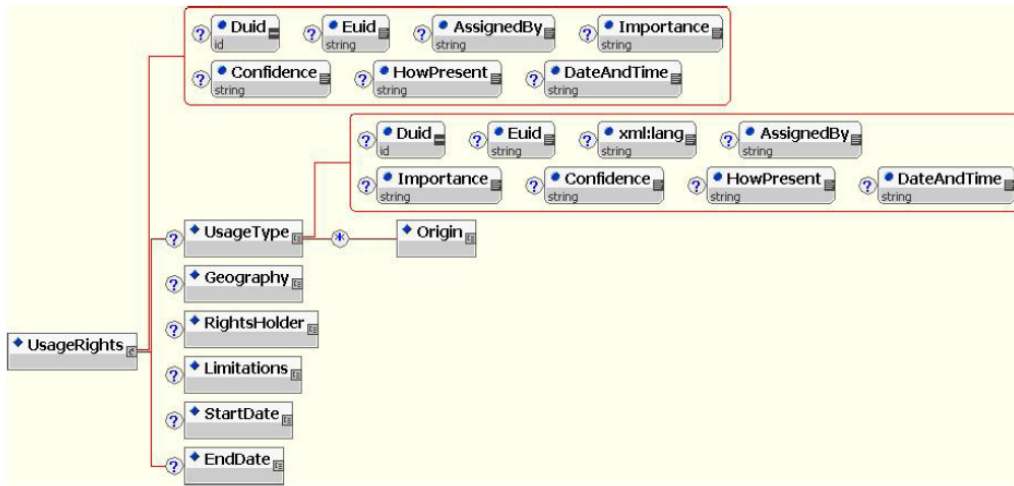
Παράδειγμα:

```
<NewsComponent>
  <TopicSet Duid="LocalTopicSet">
    <Topic Duid="topic1" Details="http://www.afghan-info.com/Politics/Massoud/Biography.htm">
      <!-- Details given in a web page -->
      <!-- extracted from urn:newsml:afp.com:20020701:topicset.AfghanPeople:3 -->
      <TopicType FormalName="Person" />
      <FormalName>Ahmed Shah Massoud</FormalName>
      <Description xml:lang="en-US">Cdr Massoud</Description>
      <Property FormalName="Nationality" Value="AFG"
ValueVocabulary="urn:newsml:iptc.org:20001006:topicset.iso-country:3" ValueScheme="ISO3166-alpha3"/>
    </Topic>
    <Topic Duid="topic2">
      <!-- extracted from urn:newsml:afp.com:20001006:topicset.IslamistOrganizations:5 -->
      <TopicType FormalName="Organization" />
      <FormalName>Al Qaeda</FormalName>
    </Topic>
  </TopicSet>
  <DescriptiveMetadata>
    <TopicOccurrence Topic="#topic1" HowPresent="Prominent" />
    <TopicOccurrence Topic="#topic2" HowPresent="RelatesTo" />
  </DescriptiveMetadata>
</NewsComponent>
```


Usage rights

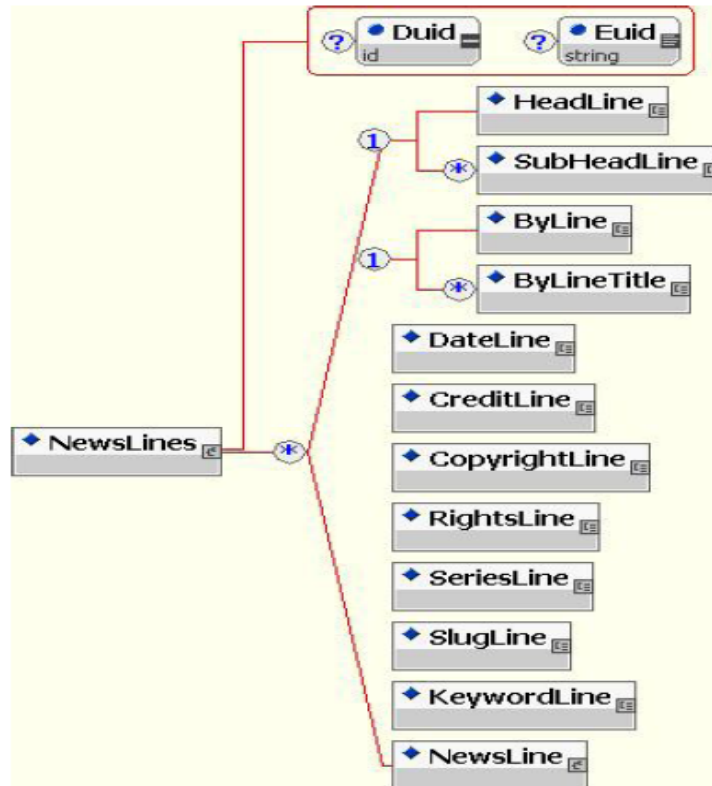
Το UsageRights περιέχει πληροφορίες όπως:

- **UsageType** δηλωση που αναφέρεται στη χρήση.
- **Geography** ο γεωγραφικός χώρος που ισχύουν τα δικαιώματα.
- **RightsHolder** – ποιος έχει τα δικαιώματα
- **Limitations** περιορισμοί δικαιωμάτων
- **StartDate** και **EndDate**, - η περίοδος που ισχύουν τα δικαιώματα.



NewsLines

Το NewsLines παραθέτει τομείς των ειδήσεων σε φυσική γλώσσα. Αντιπροσωπεύουν αντικείμενα ειδήσεων όπως η επικεφαλίδα η ημερομηνία συμβάντος. Η σειρά των NewsLines είναι ελεύθερη και ο κάθε παροχέας μπορεί να ακολουθήσει όποια σειρά θέλει.



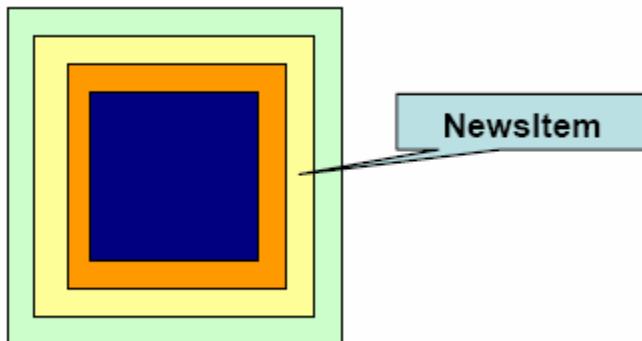
Τα NewsLines μπορούν να θεωρηθούν ως ένα ευρετήριο καθώς αρχικά είχαν αναπτυχθεί προκειμένου να υπάρξει μία αναπαράσταση των μεταδεδομένων που συνοδεύουν τα περιεχόμενα.

Ένα σύνολο NewsLines περιλαμβάνει:

- Το στοιχείο HeadLine περιλαμβάνει την επικεφαλίδα και το SubHeadLine την υποκεφαλίδα.
- Το ByLine παρέχει πληροφορίες για τον συγγραφέα και δημιουργό της πληροφορίας.
- Το DateLine παρέχει πληροφορίες σε φυσική γλώσσα για την ημερομηνία δημιουργίας του περιεχομένου.
- Το CopyrightLine δηλώνει τον κάτοχο των πνευματικών δικαιωμάτων.
- Το RightsLine δηλώνει ποιος έχει το δικαίωμα χρήσης. Διαφέρει από το παραπάνω στοιχείο που αναφέρει τον κάτοχο.
- Το στοιχείο SeriesLine παρέχει πληροφορίες που αφορά ειδήσεις που έχουν κάποια ακολουθία- σειρά.

- Το στοιχείο SlugLine παρέχει πληροφορίες που προσδιορίζονται από τον εκάστοτε παροχέα κάθε φορά.
- Το στοιχείο KeywordLine παρέχει πληροφορίες για τις λέξεις κλειδιά που σχετίζονται με το περιεχόμενο.

Το επίπεδο διαχείρισης – `NewsItem`



Ένα επίπεδο πιο πάνω από το επίπεδο δομής που είδαμε πιο πάνω βρίσκεται το στοιχείο `NewsItem` που αφορά στη διαχείριση των ειδήσεων στη NewsML.

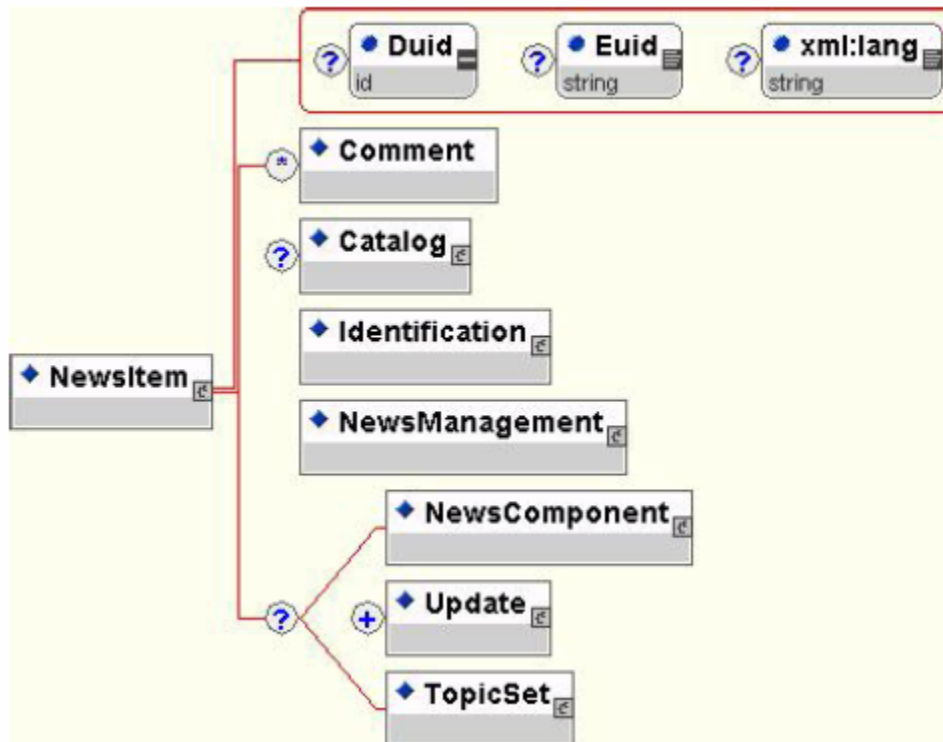
Το στοιχείο `NewsItem` είναι το αναγνωρίσιμο και δημοσιεύσιμο κομμάτι των ειδήσεων. Αποτελεί στη ουσία το αντικείμενο των ειδήσεων που ο προμηθευτής δημιουργεί, αποθηκεύει, διαχειρίζεται, επαναχρησιμοποιεί και συνδέει με άλλα ισότιμα στοιχεία `NewsItems`. Στο ειδησιογραφικό περιβάλλον το στοιχείο αυτό χρησιμοποιείται για την ανταλλαγή ειδήσεων ή ακόμα και ως σημείο εισόδου σε πολλά `NewsItems` που το ένα παραπέμπει ή περικλείει το άλλο.

Το στοιχείο αυτό μπορεί να αντιπροσωπεύει:

- Είτε μια ενιαία πηγή όπως για παράδειγμα, κείμενο, φωτογραφία, ήχος ή βίντεο.
- Είτε ένα πακέτο πολυμέσων που περιλαμβάνει διαφορετικά ειδησιογραφικά αντικείμενα. Σαν παράδειγμα μπορούμε να θέσουμε μια στήλη ειδήσεων στην οποία πέρα από το κείμενο υπάρχουν μερικές φωτογραφίες.
- Είτε τέλος μια συλλογή από συναφή ειδησιογραφικά αντικείμενα, π.χ. οι πέντε καλύτερες φωτογραφίες της ημέρας. Μια τέτοια συλλογή μπορεί να περιέχει επίσης συνδέσμους σε άλλα `NewsItems`

με προαιρετική την εμφάνιση κάποιων ετικετών ή μεταδεδομένων σχετικών με τα αντικείμενα που παραπέμπονται.

Η δομή του **NewsItem**

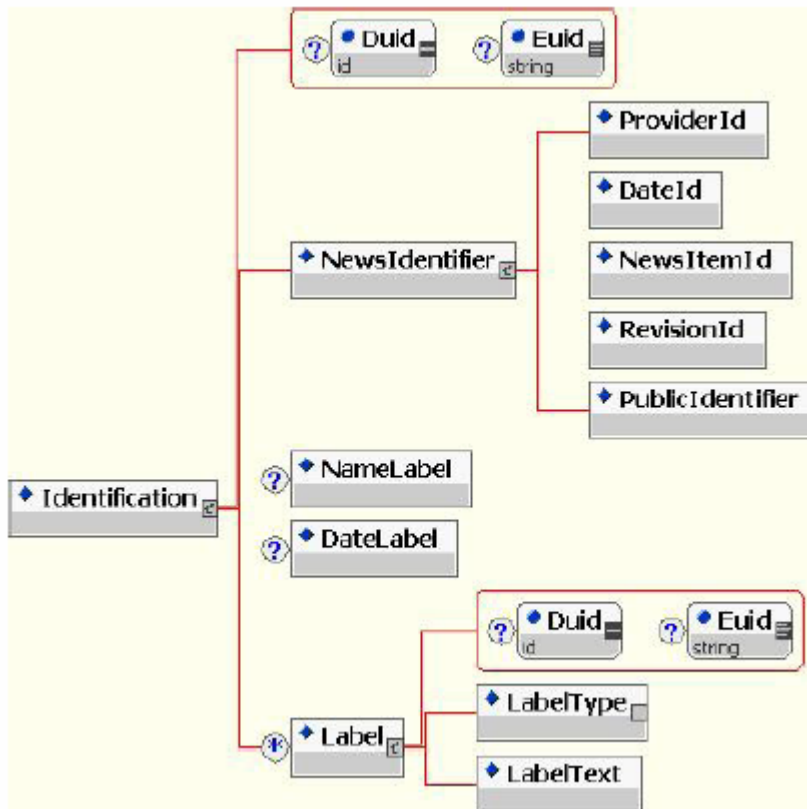


Τυπικό στοιχείο ταυτοποίησης του **NewsItem**

Είναι πολύ σημαντικό να μπορεί να αναγνωριστεί το **NewsItem**, καθώς διακινείται και μεταφέρεται από μέρος σε μέρος και από σύστημα σε σύστημα. Η NewsML προβλέπει και απαιτεί τα στοιχεία **NewsItems** να έχουν ένα μοναδικό προσδιοριστή ταυτότητας στο στοιχείο **NewsIdentifier**.

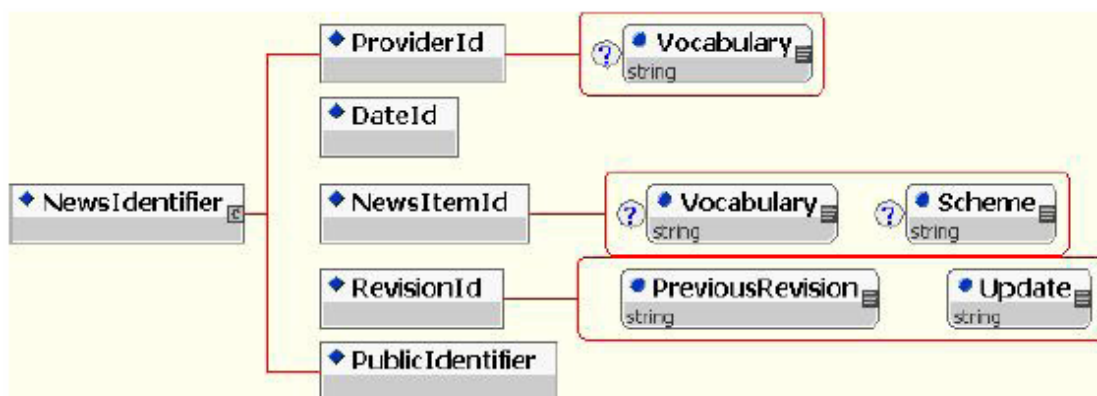
Αναγνώριση (**Identification**)

Η δομή του στοιχείου **Identification** περιλαμβάνει τα εξής υποστοιχεία όπως διαφαίνονται στο σχήμα που ακολουθεί.



NewsIdentifier

Το στοιχείο NewsIdentifier περιέχει τέσσερα βασικά υποστοιχεία, το ProviderId, το DateId, το NewsItemId και το RevisionId, καθώς και το υποστοιχείο PublicIdentifier που συνδέει αλυσιδωτά τα προαναφερόμενα τέσσερα υποστοιχεία σε μια ενιαία σειρά χαρακτήρων. Το NewsIdentifier, όπως έχουμε αναφέρει και πιο πάνω, παρέχει στο NewsItem ένα παγκόσμια μοναδικό προσδιοριστή ταυτότητας. Σε καμία περίπτωση δηλαδή δεν πρέπει δύο διαφορετικά NewsItems να περικλείουν τα ίδια ProviderId, DateId, NewsItemId και RevisionId. Στην περίπτωση βέβαια που ένα NewsItem ανανεωθεί και υπάρξει η παραμικρή έστω αλλαγή στο περιεχόμενο, τότε ένα νέο υποστοιχείο RevisionId θα πρέπει να εισαχθεί στην ανανεωμένη και νέα έκδοχή.



Στον παρακάτω πίνακα που ακολουθεί δίδεται μια σύντομη περιγραφή των τεσσάρων υποστοιχείων του NewsIdentifier:

Υποστοιχείο	Περιγραφή
ProviderId	Ο ταυτοποιητής του παροχού, που συνήθως εμφανίζεται με τη μορφή ενός χώρου ονομάτων διαδικτύου (internet domain name) π.χ. iptc.org
DateId	Προσδιορίζει την ημερομηνία, χωρίς να γίνεται αναφορά στην ώρα, με μορφή που προσδιορίζεται από το ISO 8601, δηλ. CCYYMMDD, π.χ. 20031231
NewsItemId	Πρόκειται για τον προσδιοριστή ταυτότητας του ειδησιογραφικού αντικειμένου, που είναι μοναδικό για το χώρο ονομάτων του συγκεκριμένου παροχέα και στη συγκεκριμένη ημερομηνία.
RevisionId	Πρόκειται για ένα θετικό ακέραιο αριθμό. Η προεπιλεγμένη τιμή είναι πάντα το «1» (ένα) και η τιμή αυτή αυξάνεται κάθε φορά που προκύπτει μια διόρθωση ή αναθεώρηση.

PublicIdentifier

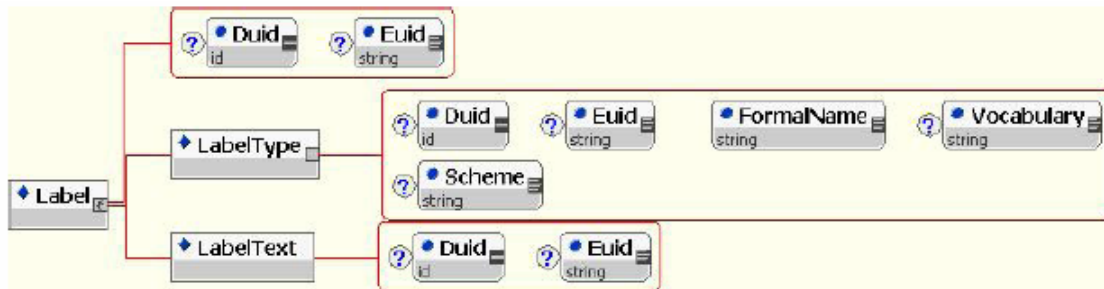
Στην ουσία το στοιχείο PublicIdentifier ενός NewsItem αποτελεί το URN της NewsML. Συγκεκριμένα πρόκειται για ένα URI (Uniform Resource Identifier) που προσδιορίζει ένα αντικείμενο, αλλά δεν το εντοπίζει ρητά. Η ύπαρξη ενός URN καθιστά δυνατό σε ένα NewsItem να αποτελεί αναμφίβολα σημείο αναφοράς από δείκτες άλλων στοιχείων σε XML ή άλλων πόρων, π.χ. από το στοιχείο NewsItemRef.

Το URN της NewsML σχηματίζεται από την αλληλουχία των τεσσάρων προαναφερόμενων υποστοιχείων, δηλαδή μια τυπική μορφή ενός URN είναι ως εξής:

urn:newsml:{ProviderId}:{DateId}:{NewsItemId}:{RevisionId Update}

NameLabel, DateLabel και Label

Τα προαιρετικά στοιχεία **NameLabel** και **DateLabel** και το γενικής χρήσης **Label** αποτελούν σειρές χαρακτήρων αναγνώσιμες από τον άνθρωπο με σκοπό τον προσδιορισμό ενός ειδησιογραφικού αντικειμένου. Εύκολα γίνεται αντιληπτό ότι τέτοιου είδους στοιχεία δεν είναι δυνατόν να αποτελούν αντικείμενο επεξεργασίας, ενώ συγχρόνως δεν είναι δυνατόν να εξάγονται από τα ίδια πεδία μοναδικά προσδιοριστικά ταυτότητας. Αντιθέτως αποτελούν στοιχεία που μπορούν να ανανεώνονται οποτεδήποτε αναθεωρείται το NewsItem. Πιο αναλυτικά το στοιχείο **NameLabel** αποτελεί μια σύντομη ονομασία που παραπέμπει στο NewsItem και δίδεται από τον ίδιο το δημιουργό, ενώ το **DateLabel**, εκφράζει την ημερομηνία δημοσίευσης των ειδήσεων. Τέλος το στοιχείο **Label** χρησιμοποιείται όταν συγκεκριμένοι προσδιοριστές ταυτότητας προσδίδονται στο NewsItem από τον παροχέα.



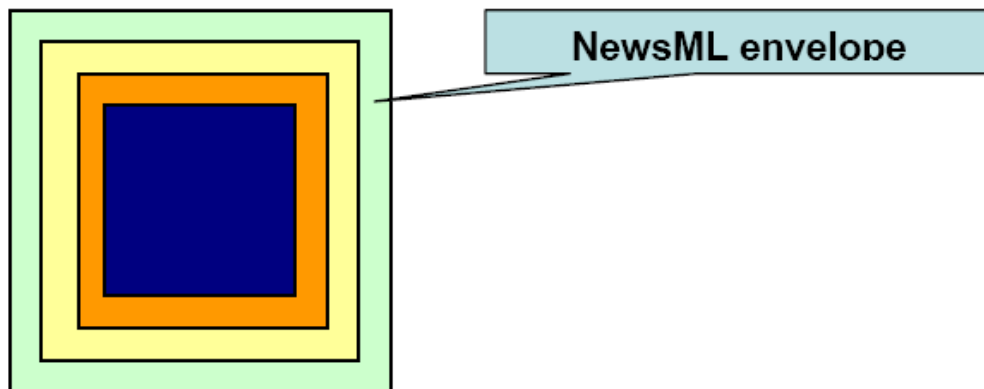
Η δομή του στοιχείου **Label** διαφαίνεται στο παραπάνω σχήμα. Το υποστοιχείο **LabelText** φιλοξενεί δεδομένα, ενώ το υποστοιχείο **LabelType** δέχεται ορισμένες μόνο τιμές από ένα ελεγχόμενο λεξιλόγιο. Παραδείγματος χάρη, σε μια μετάδοση ειδήσεων ένας συγκεκριμένος κωδικός μπορεί να δοθεί σε μια φωτογραφία ενός φωτογράφου, χωρίς όμως με αυτό τον τρόπο η

φωτογραφία να προσδιορίζεται μοναδικά στο χώρο ονομάτων του παροχέα. Η χρησιμότητα του συγκεκριμένου κωδικού έγκειται στη διευκόλυνση του φωτογράφου να ελέγξει ότι μεταδόθηκε η σωστή φωτογραφία και όχι κάποια άλλη.

```
<NewsItem>
  <Identification>
    <NewsIdentifier>
      <ProviderId>iptc.org</ProviderId>
      <DateId>20030602</DateId>
      <NewsItemId>25ab03</NewsItemId>
      <RevisionId PreviousRevision="0" Update="N">1</RevisionId>
      <PublicIdentifier>urn:newsml:afp.com:20020602:25ab03:1</PublicIdentifier>
    </NewsIdentifier>
    <NameLabel>IPTC 2003 AGM Aarhus</NameLabel>
    <DateLabel>6 june 2003</DateLabel>
    <Label>
      <LabelType FormalName="CodeName" />
      <LabelText>AAR01</LabelText>
    </Label>
  </Identification>
  ...
</NewsItem>
```

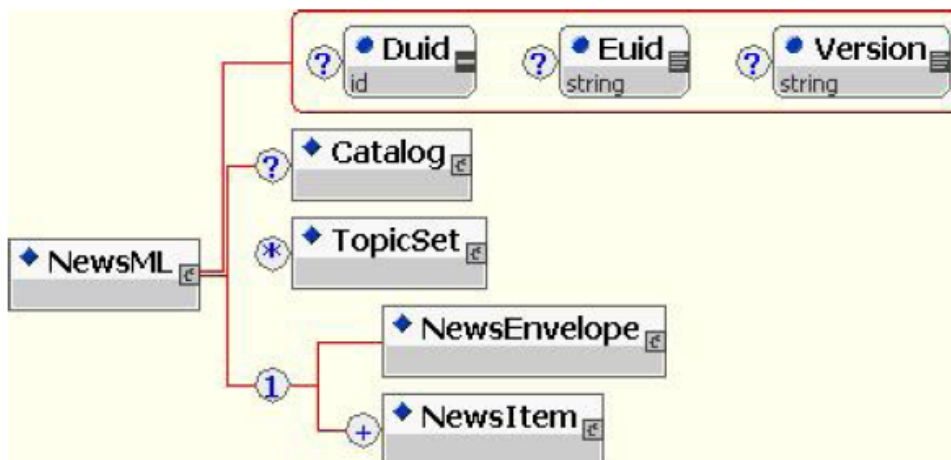
Το επίπεδο ανταλλαγής - NewsML envelope

Τα NewsItems ανταλλάσσονται μεταξύ συστημάτων σε φακέλους (NewsML envelopes). NewsML περιέχει μεταδεδομένα που είναι χρήσιμα για την ανταλλαγή των δεδομένων



NewsML root element

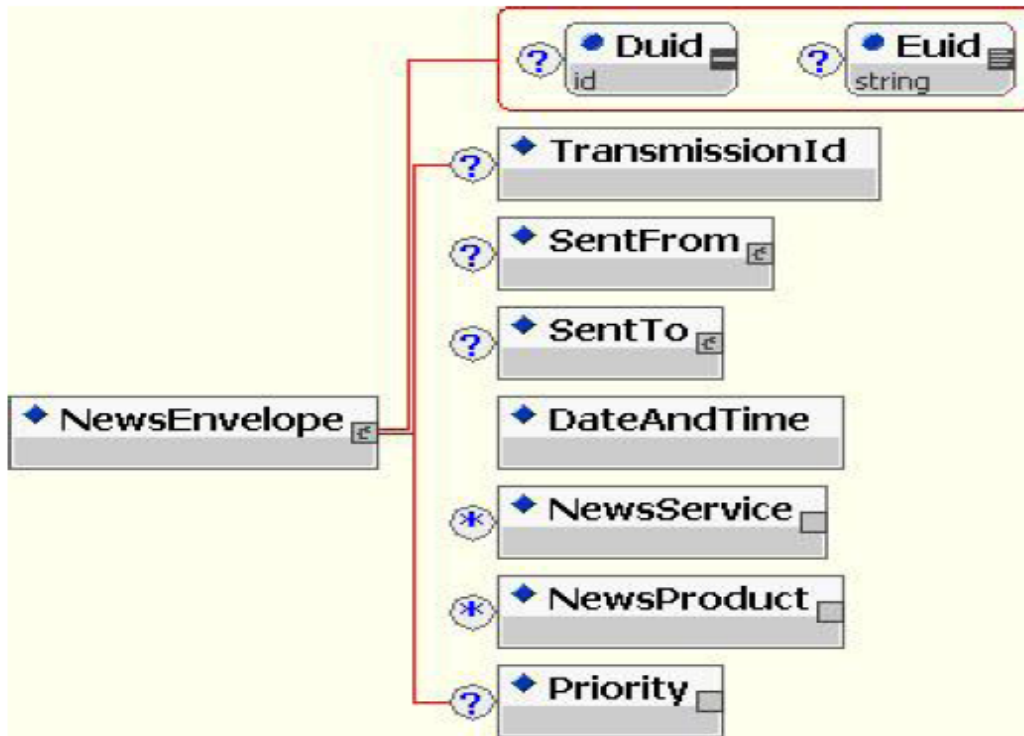
Το NewsML root element περιέχει ένα ή περισσότερα NewsItems. Το NewsEnvelope χρησιμοποιείται για να παρέχει πληροφορίες που σχετίζονται με την μετάδοση των NewsItems.



NewsEnvelope

Το στοιχείο NewsEnvelope περιέχει πληροφορίες για το πώς το NewsML τεκμήριο χρησιμοποιείται μέσα στην όλη διαδικασία και σε σχέση μεταξύ του παροχέα και του αποδέκτη.

Το NewsEnvelope έχει την παρακάτω δομή:



7.3.1 DateAndTime

Το μοναδικό υποχρεωτικό στοιχείο είναι το **DateAndTime**, που υποδηλώνει την ημερομηνία και την ώρα μετάδοσης του αντικειμένου σύμφωνα με μορφή του το ISO 8601 (π.χ. **20031231T000001+0100**).

SentFrom, SentTo

SentFrom χρησιμοποιείται προκειμένου να προσδιοριστεί ποιος στέλνει το NewsML instance. Ο αποστολέας προσδιορίζεται στο υποστοιχείο Party, ενώ το **SentTo** χρησιμοποιείται προκειμένου να προσδιοριστεί ποιος αποδέχεται την NewsML.

Βιβλιογραφία

1. <http://www.newsml.org>
2. <http://www.iptc.org/EventsML/>
3. <http://www.iptc.org/>
4. <http://www.sportsml.org/>
5. <http://www.xmlnews.org>
6. <http://www.prism.org>
7. International Press Telecommunications Council, "NewsML™ 1.2, Guidelines V 1.00", 2004.
8. Tony Allday, "NewsML: enabling a standards – led revolution in news publishing?" Reuters Media Group, Ebu Technical Review, 2001.