

Η Χρήση των Θησαυρών στις Οντολογίες

Σταματίνα Τσάφου*, Ματίνα Πολίτη**

**stsafou@teiath.gr*, ** *mpfatsa@gmail.com* ή *matina.politi@covidien.com*

Περίληψη

Στις σύγχρονες μεθόδους θεματικής αναπαράστασης, οργάνωσης και ανάκτησης πληροφοριών από ηλεκτρονικές πηγές πληροφόρησης εντάσσονται οι θησαυροί και οι οντολογίες. Οι θησαυροί αποτελούν τη νοηματική οργάνωση οριζοντίων και κάθετων ιεραρχικών συσχετισμών των εννοιών. Οι οντολογίες, ως βασικό μέσο αναπαράστασης γνώσης στο σημασιολογικό ιστό, επιτρέπουν το διαμοιρασμό της γνώσης και των πληροφοριών και παρέχουν τους μηχανισμούς για την αποτελεσματική αξιοποίηση αυτών, τόσο από ανθρώπους όσο και από πράκτορες λογισμικού.

Στον τομέα των Ψηφιακών Βιβλιοθηκών όπου η πληροφορία πρέπει να παρουσιάζεται μονοσήμαντα, μέσω ενός καλά ορισμένου τρόπου αναπαράστασής της, ο ρόλος των θησαυρών και των οντολογιών είναι εξαιρετικά σημαντικός.

Στην εισήγηση αυτή εξετάζεται η σχέση των θησαυρών με τις οντολογίες. Παρουσιάζονται οι ομοιότητες στη διαδικασία δόμησης θησαυρού και οντολογίας και η δυνατότητα χρήσης των θησαυρών στην οργάνωση των οντολογιών. Ακόμα παρουσιάζεται η δυνατότητα χρήσης του θησαυρού στις οντολογίες, μέσω συγκεκριμένων κανόνων και η χρήση των οντολογιών και των θησαυρών στην οργάνωση και ανάκτηση των πληροφοριών.

Οι σημασιολογικές τεχνολογίες επιτρέπουν την περιγραφή των αντικειμένων και των θησαυρών γνώσεων. Ο κυριότερος στόχος τους είναι να υπάρχει η δυνατότητα πρόσβασης συνακόλουθα και κατανοητά σε παρόμοιες τάξεις-κατηγορίες ψηφιακών αντικειμένων και υπηρεσιών, κατανεμημένα δια μέσου ετερογενών θησαυρών γνώσεων.

Λέξεις κλειδιά: θησαυρός, οντολογία, σημασιολογικός ιστός, ψηφιακές βιβλιοθήκες, ανάκτηση πληροφοριών

Abstract

Thesauri and ontologies, as structures representing digital information sources, play an important role among the modern methods of knowledge representation and management and information retrieval. Thesauri constitute the horizontal and vertical meaning management of hierarchic correlations of significances. Ontologies, as the basic tool for knowledge representation in Semantic Web, allow knowledge distribution and provide mechanisms for effective exploitation of information both from humans and software agents.

Concerning the domain of Digital Libraries, where information should be presented unambiguously via a well-defined representational way, the role of thesauri and ontologies is exceptionally important.

In this paper the relations between thesauri and ontologies are examined by presenting the resemblances and differences in the procedure of structuring a thesaurus and ontology. Moreover, the use of a thesaurus into ontology is proposed through certain rules that map the elements and relations contained in the thesaurus to concepts and semantic relations of the ontology.

Semantic technologies allow the description of objects and thesauri of knowledge. Our main objective is to have the ability to access consecutively and comprehensibly categories of digital objects and services, distributed by means of heterogeneous thesauri of knowledge.

Keywords: thesaurus, ontology, semantic web, digital libraries, information retrieval

1. Πεδίο Ανάπτυξης

Η εξέλιξη της τεχνολογίας επιτάχυνε σημαντικά τους ρυθμούς ανάπτυξης όλων των επιστημών, με επακόλουθο την αύξηση των επιστημονικών κειμένων και των πληροφοριών. Η αύξηση αυτή δυσχεραίνει τον εντοπισμό μιας συγκεκριμένης πληροφορίας που αναζητά ο χρήστης.

Η ταυτόχρονη εξέλιξη της τεχνολογίας της πληροφόρησης επέφερε μεταβολές στις μεθόδους οργάνωσης και αποθήκευσης και επιτάχυνε τους ρυθμούς εντοπισμού και διακίνησης των πληροφοριών. Βασικό μέλημα των επιστημόνων της πληροφόρησης ήταν και είναι η βελτίωση των υπάρχοντων ή η εξεύρεση νέων, πλέον αποτελεσματικών μεθόδων θεματικής οργάνωσης των πληροφοριών. Ο χρήστης ερευνητής, στην άσκηση της επιστημονικής εργασίας, έχει ανάγκη να έχει στη

διάθεσή του την κατάλληλη πληροφορία όταν την χρειάζεται. Για το λόγο αυτό, η θεματική επεξεργασία των πληροφοριών, που εμπεριέχονται στα τεκμήρια αποτελεί κυρίαρχο πρόβλημα σε κάθε σύστημα πληροφόρησης.

Πρώτη προσπάθεια θεματικής οργάνωσης των πληροφοριών ήταν η δημιουργία θεματικών ευρετηρίων. Η αποτελεσματικότητα των θεματικών ευρετηρίων, ως μέσου για την ταύτιση και την ανάκτηση τεκμηρίων, εξαρτιόταν από μια καλά δομημένη γλώσσα ευρετηρίασης. «Η ανάγκη απόδοσης των όρων ευρετηρίασης με τρόπο μοναδικό αντιμετωπίστηκε με τη χρήση γλώσσας ανάκτησης πληροφοριών, η οποία διαχρονικά είχε τη μορφή ταξινομικού συστήματος, δομημένου καταλόγου θεματικών επικεφαλίδων και πρόσφατα, δομημένου θησαυρού¹». Η γλώσσα ευρετηρίασης χρησιμοποιείται, τόσο για την ένταξη των πληροφοριών κάτω από ενιαίους όρους, όσο και για τη διεξαγωγή της έρευνας για την ανεύρεση των πηγών πληροφόρησης. Κύριο χαρακτηριστικό των γλωσσών ευρετηρίασης είναι η χρήση ενός δομημένου ελεγχόμενου λεξιλογίου.

Η ανάπτυξη των υστεροσυνδυασμένων μεθόδων ευρετηρίασης με τη δημιουργία και χρήση θησαυρών συμβάλλει αποτελεσματικά στη θεματική οργάνωση και αναζήτηση των πληροφοριών. Με τον όρο θησαυρός χαρακτηρίζεται το λεξιλόγιο όρων μιας ελεγχόμενης γλώσσας ευρετηρίασης, δομημένο με τέτοιο τρόπο ώστε να υποδηλώνονται με σαφήνεια οι προϋπάρχουσες μεταξύ των εννοιών σχέσεις. Αποτελεί ένα εργαλείο για την ένταξη των γνώσεων και των πληροφοριών σε εννοιολογικές ομάδες, διευκολύνει τη θεματική οργάνωση των συλλογών, των καταλόγων των βιβλιοθηκών, των ευρετηρίων και των ηλεκτρονικών βάσεων πληροφοριών.

Η σάρωση του θησαυρού κατά την αναζήτηση επιτρέπει στον ερευνητή να καθορίζει το επίπεδο ανάκτησης πληροφοριών από γενικές σε ειδικές πληροφορίες και το αντίστροφο, ρυθμίζοντας το επίπεδο ανάκτησης που εκείνος επιθυμεί.

Για την εξυπηρέτηση των χρηστών τους πολλοί διαθέτες και διεθνείς οργανισμοί προχώρησαν στη δημιουργία θησαυρών σε συγκεκριμένα και σε σχετιζόμενα με αυτά θεματικά πεδία.

Η διάσπαση των θησαυρών σε επιστημονικά πεδία δεν διευκόλυνε τους επιστήμονες στην εξαντλητική αναζήτηση πληροφοριών. Αυτό οφείλεται στη μη ταύτιση των όρων ή της εννοιολογικής τους ανάλυσης ή της ένταξής τους σε ιεραρχίες. Η δυσκολία αυτή προβλημάτισε τους επιστήμονες, κυρίως της πληροφορικής, οι οποίοι προσπάθησαν να αναπτύξουν αυτόματες μεθόδους θεματικού προσδιορισμού των πληροφοριών με την ανάπτυξη οντολογιών².

Οι περισσότεροι από τους επιστήμονες στον τομέα της πληροφορικής αποδέχονται τον ορισμό της οντολογίας που έδωσε ο Gruber: “An ontology is a specification of a conceptualization”, δηλαδή οντολογία είναι ο προσδιορισμός μιας αντίληψης με την έννοια της λειτουργίας, με την οποία ο άνθρωπος αποκτά γνώση της πραγματικότητας με παρέμβαση του λογικού³. Είναι ένας τρόπος αναπαράστασης της γνώσης.

¹ Σκανδάλη Αλκμήνη, 2003, σ.31

² Zavitsanos E.,... 2007

³ Τριανταφυλλίδης, 1998

Σύμφωνα με τους Biolchini και Patel⁴ η οντολογία αναπαριστά τύπους γνώσης, οι οποίοι χρησιμοποιούνται από επιστήμονες στη λογική ανάλυση (reasoning) και την εξαγωγή συμπερασμάτων (decision processes): την πραγματική (factual), τη σημασιολογική (semantic) και τη σχηματική (schematic). Η προσέγγιση αυτή συμβάλλει στην ανάπτυξη του αυτοματισμού της οντολογίας, τόσο στην απόκτηση γνώσης, όσο και στις φάσεις οργάνωσής της.

Οι οντολογίες βοηθάνε στην εξάπλωση του σημασιολογικού ιστού. Θα μπορούσε να λεχθεί ότι είναι το μέσο για τη μετάβαση από τον παγκόσμιο ιστό στη σημασιολογική του έκδοση (Semantic Web).

Ο Σημασιολογικός Ιστός αποτελεί την εξέλιξη του σημερινού παγκόσμιου ιστού, από τον οποίο διαφέρει στο ότι σ' αυτόν παρέχεται η σημασιολογία των πληροφοριών που διαθέτει. Με τον τρόπο αυτό η πληροφορία μαζί με τη μεταπληροφορία που τη συνοδεύει είναι κατανοητή και κατάλληλη για να τη διαχειριστούν άνθρωποι και μηχανές.

Ο θησαυρός συμβάλλει στην αποτελεσματική οργάνωση των πληροφοριών ενός πληροφοριακού συστήματος και στην ανάκτηση των πληροφοριών αυτών από το χρήστη.

Οι οντολογίες στην πληροφόρηση, χρησιμοποιούνται για την ανάκτηση πληροφοριών από διαφορετικά πληροφοριακά συστήματα. Και οι δύο μέθοδοι συμβάλλουν στην εξαντλητική ανάκτηση των κατάλληλων για το χρήστη πληροφοριών περιορίζοντας το θόρυβο της ανάκτησης.

2. Η Χρήση των Θησαυρών σε Οντολογίες

2.1. Οντολογίες και επιστήμη της πληροφόρησης (βιβλιοθήκες)

Στον κλάδο της Βιβλιοθηκονομίας και της Επιστήμης της Πληροφόρησης, τα τελευταία χρόνια πολύς λόγος γίνεται για τις οντολογίες και για το κατά πόσο και με ποιο τρόπο μπορούν να συμβάλλουν στη λειτουργία και ανάπτυξη των Ψηφιακών, κυρίως, Βιβλιοθηκών.

Οι Ψηφιακές Βιβλιοθήκες προσφέρουν πρόσβαση σε πλήθος ψηφιακών πηγών πληροφόρησης στο διαδίκτυο. Παραδοσιακές βιβλιοθήκες με τη μέθοδο της σάρωσης, μετατρέπουν τα έντυπα σε ηλεκτρονικά τεκμήρια έτσι ώστε οι πηγές πληροφόρησης, μέσω του παγκόσμιου ιστού, να είναι διαθέσιμες σε ένα ευρύτερο κοινό. Η ψηφιοποίηση από μόνη της δεν είναι αρκετή για την αξιοποίηση των πληροφοριών των ηλεκτρονικών τεκμηρίων, αλλά και γενικότερα, η ανάπτυξη των Ψηφιακών Βιβλιοθηκών περιλαμβάνει την επεξεργασία, την ανάλυση και τη διάθεση όλων των τύπων ψηφιακών πληροφοριών.

Ο στόχος της οντολογίας, στην εφαρμογή της στις ψηφιακές βιβλιοθήκες, είναι η ηλεκτρονική ανάλυση των όρων και γενικά των εννοιών, που χρησιμοποιούνται για τη θεματική περιγραφή των καταχωρημένων πληροφοριών και κειμένων,

⁴ Biolchini Jorge, Patel Vilma L. 2004

προκειμένου να εξακριβωθεί η ταύτισή τους σε διαφορετικά πληροφοριακά συστήματα, ώστε να ανακτηθούν οι ζητούμενες πληροφορίες.

Οι οντολογίες αξιοποιούν τις ψηφιακές βιβλιοθήκες. «Ακόμα και ο τομέας των ψηφιακών βιβλιοθηκών έχει αρχίσει να περιγράφεται μέσω οντολογιών»⁵, οπότε αυτό που απομένει είναι οι υπηρεσίες των ψηφιακών βιβλιοθηκών αλλά και η γνώση που περιέχεται σ' αυτές να περάσουν σε "σημασιολογικό" επίπεδο.

Ο Fischer⁶ το 1998, εν αναμονή της δημοσίευσης της Οντολογίας (Cyc Upper Ontology) και με αφορμή τη δημοσίευση του προτύπου 2788 του ISO, συζητά τη χρήση θησαυρού ως βάση σε ένα νέο είδος οντολογίας για εφαρμογές στη τεχνητή Νοημοσύνη.

Ο Sebastian Ryszard Kruc⁷ το Σεπτέμβριο του 2006 αναφέρθηκε στο ρόλο των Οντολογιών στις Σημασιολογικές Ψηφιακές Βιβλιοθήκες. Υποστήριξε ότι χρειάζονται οντολογίες, οι οποίες να υποστηρίζουν τις βασικές οπτικές γωνίες τρόπων επεξεργασίας της γνώσης, όπως οι βιβλιογραφικές περιγραφές κ.ά.

Ο Boris Lauser⁸ προσπάθησε να αποδείξει σε τι διαφέρουν οι οντολογίες από τους θησαυρούς, στηριζόμενος σε ένα απλό παράδειγμα από τον τομέα των Safety Food χρησιμοποιώντας το θησαυρό AGROVOC. Στο παράδειγμα παρουσιάστηκε η δυνατότητα της χρήσης των γλωσσών προγραμματισμού RDFS, ώστε ένας θησαυρός να εξελιχθεί σε ένα σύνθετο μοντέλο οντολογίας με τη χρήση της γλώσσας OWL.

Σε εθνικό επίπεδο, μερικές μελέτες αναφέρονται στις Οντολογίες και στις Ψηφιακές Βιβλιοθήκες και όχι στις Οντολογίες και στους Θησαυρούς. Περισσότερο περιστρέφονται γύρω από τη Σημασιολογική Ολοκλήρωση Δεδομένων⁹, τη Διαχείριση οντολογιών με μελέτη και εμβάθυνση στα βασικά τους προβλήματα, την παρουσίαση υφισταμένων οντολογιών¹⁰ και την Οντολογία στις Ψηφιακές Βιβλιοθήκες¹¹.

Στην εισήγηση αυτή επιχειρείται η μετατροπή ενός Βιβλιοθηκονομικού Θησαυρού σε Οντολογία με τη βοήθεια ενός συντάκτη οντολογιών (ontology editor), ώστε να αποφευχθεί η εκ του μηδενός σύνταξη της οντολογίας στη γλώσσα περιγραφής οντολογιών OWL.

2.2 Μετατροπή ενός θησαυρού σε οντολογία – Προτεινόμενη μέθοδος

⁵ Αντωνάκης Δημήτρης, κ.ά. 2006

⁶ Fischer Dietrich H...,1998

⁷ <http://www.slideshare.net/tag/corrib>

⁸ Lauser Boris.

⁹ Μπουντούρη, Λίνα, Γεργατσούλης, Μανόλης. 2006

¹⁰ Γαϊτάνου, Πανωραία, Γεργατσούλης Μανόλης, 2006

¹¹ Αντωνάκης Δημήτρης, κ.ά. 2006

Η οργάνωση των θησαυρών αποτελεί προϊόν διανοητικής εργασίας ειδικών επιστημόνων, οι οποίοι αποδίδουν τους όρους σύμφωνα με την ισχύουσα ορολογία του επιστημονικού πεδίου στον οποίο ανήκει ο όρος. Αντίθετα η οντολογία αποτελεί προϊόν αυτόματης επεξεργασίας των Η/Υ.

Οι Η/Υ δεν διαθέτουν «νόηση» για την ανάλυση των εννοιών. Για τη μηχανική επεξεργασία οι όροι κωδικοποιούνται κατά τέτοιο τρόπο, ώστε μια μηχανή να μπορεί να τους επεξεργαστεί και να καταλήξει σε συμπεράσματα, που ένας άνθρωπος μπορεί να οδηγηθεί μέσω του λογικού συλλογισμού.

Σε ένα βιβλιοθηκονομικό θησαυρό η εννοιολογική ανάλυση ενός όρου περιλαμβάνει:

- ❖ ΔΣ (SN) – Διευκρινιστική Σημείωση
- ❖ ΧΡ (USE) – Χρησιμοποίησε
- ❖ ΧΑ (UF) – Χρησιμοποίησε Αντί
- ❖ ΟΚ (TT) – Όρος Κορυφή
- ❖ ΠΟ (BT) – Πλατύτερος Όρος
- ❖ ΕΟ (NT) – Ειδικότερος Όρος
- ❖ ΣΟ (RT) – Σχετικός Όρος
- ❖ ΠΡΟΣΔΙΟΡΙΣΤΗΣ ο οποίος δίνει την άποψη από την οποία χρησιμοποιείται ο όρος.

Σε μια οντολογία χρησιμοποιούμε αντίστοιχα:

- ❖ Top Concept (όρος / έννοια κορυφή)
- ❖ Slot / Properties (ιδιότητες των εννοιών)
- ❖ Super class (υπερκατηγορία) ή super concept (υπερέννοια)
- ❖ Subclass (υποκατηγορία) ή sub concept (υποέννοια)
- ❖ Individual ή instance (τύπος, άτομο ή στιγμιότυπο)
- ❖ rdfs [Resource Description Frameworks (Πηγή Περιγραφής Πλαισίων)]:
comment (σχόλιο)

Μια οντολογία αναπαριστά με σαφή και αποδοτικό τρόπο ένα θεματικό τομέα. Συγκεκριμένα, αναπαριστά τη συσσωρευμένη γνώση που υπάρχει σ' αυτόν τον τομέα. Το βασικό της πλεονέκτημα είναι ότι, μπορεί να χρησιμοποιηθεί για την εξαγωγή συμπερασμάτων σχετικά με τις ιδιότητες του εν λόγω τομέα. Μάλιστα, η εξαγωγή συμπερασμάτων μπορεί να γίνει αποδοτικότερη, αν για την περιγραφή της οντολογίας έχει χρησιμοποιηθεί μια κατάλληλη γλώσσα, όπως η OWL DL, η οποία βασίζεται σε λογικές περιγραφής (Description Logics) και επιτρέπει τον ορισμό της λογικής, που διέπει τις έννοιες και τις σημασιολογικές σχέσεις μεταξύ των εννοιών μιας οντολογίας.

Βασισμένη σε αυτές τις αρχές, μια οντολογία μπορεί να περιέχει έννοιες (αφηρημένες ή μη), σχέσεις υπαγωγής μεταξύ των εννοιών αυτών, στιγμιότυπα των εννοιών και άλλες σημασιολογικές σχέσεις μεταξύ των εννοιών. Το βασικό πλεονέκτημα στο σημείο αυτό είναι ότι, παρέχει την ελευθερία για ορισμό σημασιολογικών σχέσεων, ανάλογα με τις απαιτήσεις του χρήστη και του θεματικού τομέα που πρόκειται να αναπαρασταθεί. Επιπρόσθετα, προσφέρει τη δυνατότητα εισαγωγής ελέγχων (cardinality restrictions), οι οποίοι μέσω μηχανισμού εξαγωγής συμπερασμών, επιτρέπουν την αυτόματη παροχή σύνθετων πληροφοριών μεταξύ στιγμιότυπων δύο εννοιών, που συμμετέχουν σε μια σημασιολογική σχέση. Επίσης επιτρέπουν, τον ορισμό μιας έννοιας ως ενότητα (union) άλλων εννοιών, τον ορισμό μιας έννοιας ως ανεξάρτητη (disjoint) από κάποια άλλη, κλπ. Όλα τα τελευταία ομαδοποιούνται σε ένα σύνολο κανόνων, που ονομάζονται αξιώματα (axioms) στην οντολογία και αποτελούν τον πυρήνα για ένα μηχανισμό εξαγωγής συμπερασμάτων.

Οι παραπάνω ιδιότητες της οντολογίας, βασίζονται κυρίως στο γεγονός ότι στην ουσία οι έννοιες είναι σύνολα από στιγμιότυπα (τύποι, ή άτομα) που τις ορίζουν. Θα μπορούσαμε να πούμε πως για την οντολογία δεν έχει τόσο σημασία το όνομα κάποιας έννοιάς της, αλλά οι ιδιότητές της και τα στιγμιότυπά της. Επομένως, είναι φανερό πως μέσω των εννοιών, των σημασιολογικών σχέσεων, των στιγμιότυπων και των ιδιοτήτων μιας οντολογίας, μπορούν να απαντηθούν ερωτήσεις και να εξαχθούν συμπεράσματα, βασισμένα στις ιδιότητες των εννοιών και όχι σε απλή αναζήτηση στην οντολογία με λέξεις-κλειδιά.

Για να μετατρέψουμε τον βιβλιοθηκονομικό θησαυρό σε οντολογία, θα χρησιμοποιήσουμε ένα συντάκτη οντολογίας (ontology editor). Οι συντάκτες οντολογίας (ontology editors) υποστηρίζουν την επεξεργασία των οντολογιών, διευκολύνουν την ανάπτυξη και την διαχείριση τους, τον καθορισμό και την τροποποίηση των εννοιών και μερικοί από αυτούς επιτρέπουν τη διερεύνηση και το «περιδιάβασμά» (browsing) τους.

Στην εργασία αυτή χρησιμοποιήθηκε ο συντάκτης οντολογίας Protégé ο οποίος διατίθεται ελεύθερα από το διαδίκτυο. Το Protégé είναι ένα ολοκληρωμένο εργαλείο λογισμικού, το οποίο χρησιμοποιείται από επιστήμονες της πληροφορικής σε συνεργασία με ειδικούς επιστήμονες (domain experts), για να αναπτύξουν συστήματα πληροφόρησης, βασισμένα στην οργάνωση της ειδικευμένης γνώσης (knowledge – based systems). Οι εφαρμογές που αναπτύσσονται με το Protégé χρησιμοποιούνται για την επίλυση προβλημάτων, που δημιουργούν οι διαφορετικές αποδόσεις των εννοιών.

Στο Protégé τους χαρακτηρισμούς των μεταξύ των όρων σχέσεων ενός βιβλιοθηκονομικού θησαυρού, μπορούμε να τους αντικαταστήσουμε αντίστοιχα με χαρακτηρισμούς μιας οντολογίας ως εξής:

Θησαυρός	Οντολογία
Όρος	Έννοια (concept)
ΔΣ	rdfs: comment (rdfs:σχόλιο)
OK	Top Concept (όρος / έννοια κορυφή)
XP / XA	Slot/Properties (ιδιότητες των εννοιών)
ΠΟ	Super class (subsumption relations)
ΕΟ	Subclass (subsumption relations)
ΣΟ	Individuals (στιγμιότυπα)

Αναλυτικότερα:

Με τον όρο έννοια (concept) νοείται ένα ή σύνολο αντικειμένων αλλά και μια αφηρημένη έννοια π.χ. η διασκέδαση. Αυτό ισχύει και στο θησαυρό και στην οντολογία και αποδίδεται με έναν όρο ή ένα λήμμα.

Η Top Concept (όρος ή έννοια κορυφή), αποτελεί τον όρο που εκφράζει την γενικότερη κατηγορία, στην οποία ανήκει η έννοια που αναλύεται στην οντολογία και αντιστοιχεί στον OK ενός θησαυρού.

Στο rdfs:comment καταχωρείται σχόλιο με το οποίο, στην αντίστοιχη θέση επεξηγούνται, ο όρος, οι κατηγορίες ή οι υποκατηγορίες, οι ιδιότητες των εννοιών και των στιγμιότυπων. Στο θησαυρό η εργασία αυτή γίνεται με τη διευκρινιστική σημείωση (ΔΣ) και τον προσδιοριστή.

Τα slot είναι οι ιδιότητες των εννοιών. Στη θέση αυτή γίνεται ο προσδιορισμός σημασιολογικών σχέσεων μεταξύ δύο εννοιών μιας οντολογίας, πέρα από την καθορισμένη σχέση υπαγωγής. Στο θησαυρό η λειτουργία αυτή αντιμετωπίζεται με το χαρακτηρισμό του όρου ως προτιμώμενος, με την παραπομπή (XA) στον μη προτιμώμενο και αντίστροφα από τον μη προτιμώμενο στον προτιμώμενο με την παραπομπή XP.

Μια ταξινόμια (taxonomy) ορίζεται από σχέσεις τύπου υπαγωγής ή ταξινομικές (inclusion or subsumption relations), οι οποίες ορίζουν υπέρ (Super class: subsumption relations) και υπό (Subclass: subsumption relations) ιεραρχικές έννοιες που έχουν άλλες σημασιολογικές σχέσεις. Στο θησαυρό οι σχέσεις αυτές δίνονται για τους ευρύτερους όρους ως ΠΟ (Πλατύτεροι όροι) και για το διαχωρισμό της έννοιας του όρου σε επιμέρους με τους ΕΟ (ειδικοί όροι).

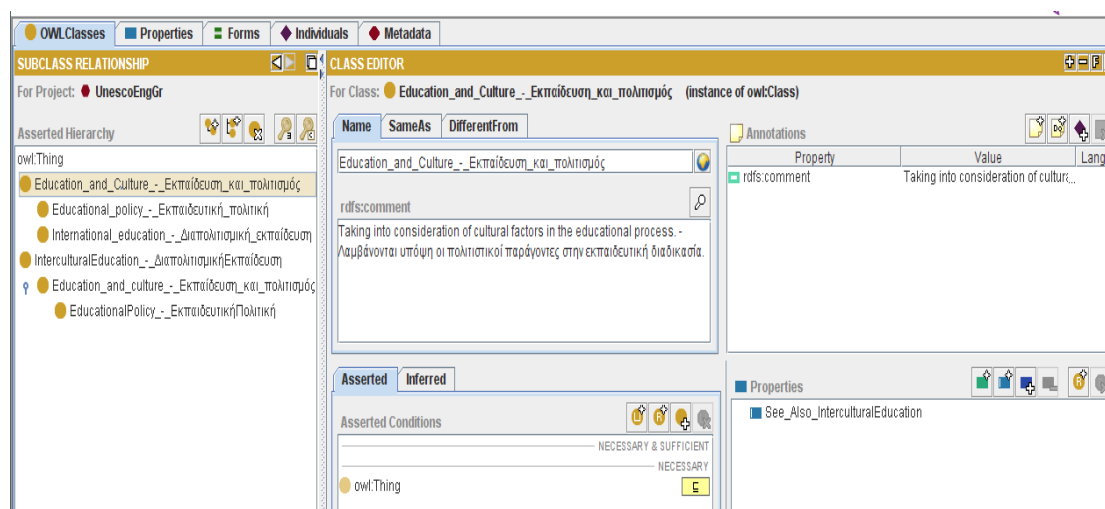
Τα στιγμιότυπα ή οι πραγματώσεις (individuals) εκφράζουν μια έννοια, που περιέχεται σε μια οντολογία και είναι τελικά ένα σύνολο από στιγμιότυπα που μοιράζονται κοινά χαρακτηριστικά. Θα μπορούσε να λεχθεί ότι αποτελούν τους σχετικούς όρους (ΣΟ) ενός περιγραφέα σε ένα θησαυρό.

2.3 Παραδείγματα μετασχηματισμού λημμάτων θησαυρού σε οντολογία

Για την εφαρμογή των παραπάνω στην πράξη, επιλέχτηκε ο όρος «Εκπαίδευση και πολιτισμός» (Education and culture) από την αλφαβητική λίστα του Θησαυρού της Unesco με διευκρινιστική σημείωση, με πλατύτερους, ειδικούς και σχετικούς όρους,

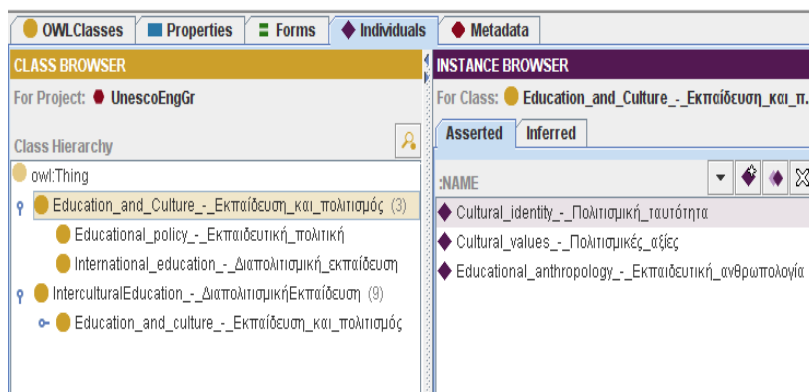
όπου η πρώτη εμφανίζεται με τη μορφή “rdfs:comment”, οι ΠΟ και ΕΟ ως Super class και Sub class αντίστοιχα και οι ΣΟ ως “Individuals”. Να σημειωθεί ότι οι όροι μπορούν να καταχωρούνται σε δύο γλώσσες, όπως στο παράδειγμα που εμφανίζεται σε Ελληνικά και Αγγλικά.

Στη μεταφορά όρων από το συγκεκριμένο θησαυρό παρουσιάστηκε πρόβλημα από τη χρησιμοποίηση αρκετές φορές του ίδιου όρου σε πεδίο και υποπεδίο (class and subclass). Όμως, στην οντολογία δεν είναι δυνατόν δύο πεδία που συνδέονται μεταξύ τους με σχέση ιεραρχίας και εμπεριέχουν λογικές έννοιες να αποδίδονται με τον ίδιο όρο. Αλλά και σ’ ένα θησαυρό οι όροι πρέπει να χαρακτηρίζονται από μονοσημαντότητα. Το παρακάτω παράδειγμα επιλέχτηκε επειδή δεν παρουσιάζει το πρόβλημα αυτό, έτσι ώστε να έχουμε μια πλήρη μεταφορά όρου θησαυρού σε οντολογία. Το Protégé δέχεται τα link του θησαυρού και μας παραπέμπει στους ΠΟ ή ΕΟ του όρου.



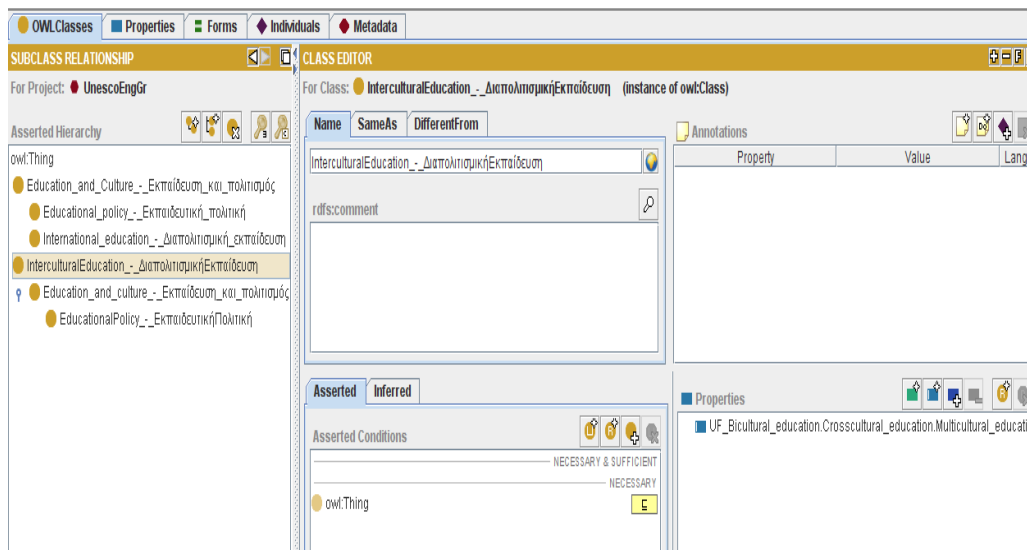
Εικόνα 1

Στην εικόνα 1 ξεκινώντας από αριστερά παρατηρούμε τους όρους “Education and Culture” και “InterculturalEducation”, με τους Πλατύτερους (ΠΟ) (Super Class) και Ειδικούς (ΕΟ) (Sub class) όρους στους οποίους αναλύονται. Στα δεξιά παρουσιάζεται η Διευκρινιστική Σημείωση ΔΣ (SN) ή rdfs:comment για το λήμμα “Education and Culture”. Κάτω και δεξιά παρουσιάζεται η ιδιότητα (Property) “See Also”, η οποία δεν αντιστοιχίζεται με την παραπομπή «Βλέπε επίσης» μιας θεματικής επικεφαλίδας, αλλά μας δίνει τη δυνατότητα να ορίζουμε τους αντίστοιχους υπερσυνδέσμους (links) του συγκεκριμένου όρου του θησαυρού. Στην παραπάνω οντολογία θα μπορούσε επίσης να οριστεί η “Έννοια Κορυφή” (Top Concept), δηλαδή ο όρος κορυφής (OK) ή η ιεραρχία που ανήκει ο όρος.



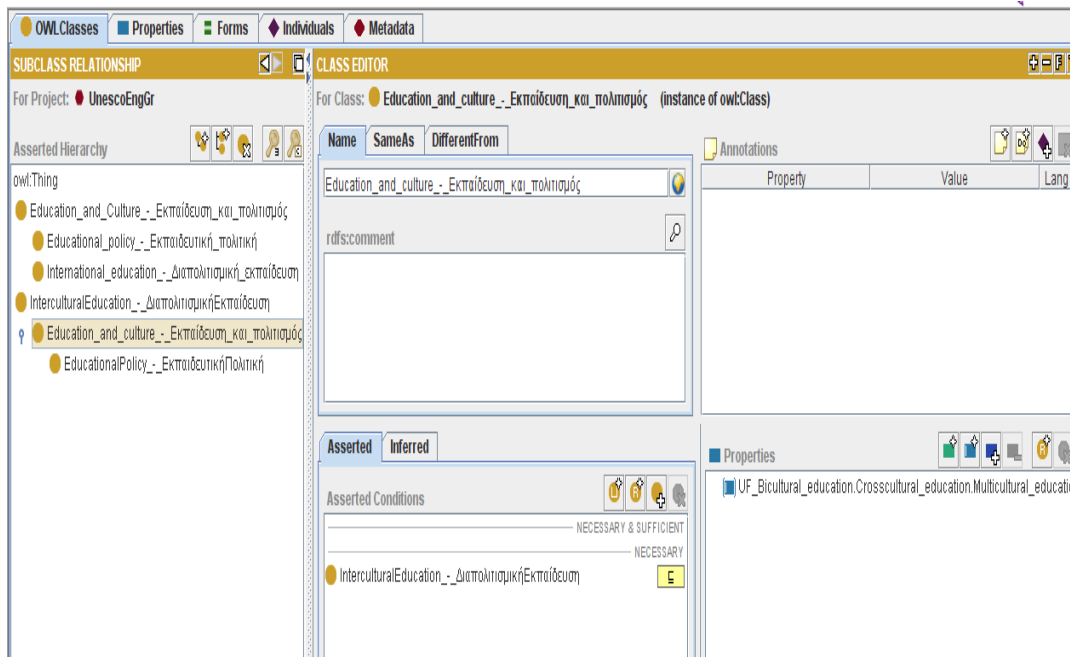
Εικόνα 2

Στην εικόνα 2 παρουσιάζονται οι Σχετικοί Όροι [ΣΟ (RT) ή Individuals], οι οποίοι αναφέρονται στο λήμμα “Education and Culture”. Στην παρουσίαση της ιεραρχικής δομής εμφανίζεται ο αριθμός των Individuals που ανήκουν στον όρο. Για παράδειγμα στη συγκεκριμένη εικόνα δίπλα στο λήμμα “Education and Culture” και “InterculturalEducation” μέσα σε παρένθεση εμφανίζεται ο αριθμός 3 και 9 αντίστοιχα, γιατί τόσα είναι τα Individuals που σχετίζονται με τα συγκεκριμένα λήμματα.



Εικόνα 3

Στην εικόνα 3 αναλύεται η εντολή παραπομπής από προτιμώμενο όρο σε μη προτιμώμενο με την εντολή XA (UF) που στην οντολογία αποτελεί Property για το λήμμα “InterculturalEducation”.



Εικόνα 4

Στην εικόνα 4 παρουσιάζονται οι πληροφορίες σχετικά με τον ΠΟ (BT) ή Super class του λήμματος “Education and Culture”, αλλά επιπλέον κάτω δεξιά παρουσιάζεται και ο μη προτιμώμενος όρος με την παραπομπή XA (UF), που στην οντολογία είναι Property του ΠΟ ή Supper Concept “InterculturalEducation”, όπου το σύμβολό του εμφανίζεται μέσα σε παρένθεση.

3. Σχόλια

Στο παραπάνω παράδειγμα ο όρος του θησαυρού μετατράπηκε σε οντολογία, χωρίς να χαθούν πληροφορίες και όλη η λειτουργία του θησαυρού μεταφέρθηκε στη λειτουργικότητα της οντολογίας.

Ο θησαυρός βοηθάει την οντολογία στο σαφή ορισμό των εννοιών της, καθορίζει τη μονοσημαντότητα των όρων, συμβάλλει στην εννοιολογική ανάλυση των εννοιών που εμπλουτίζουν μια οντολογία και καθορίζει τις σχέσεις τους στην ανάλυση της ιεραρχίας της οντολογίας.

Η οντολογία ενισχύει τη χρήση των θησαυρών οι οποίοι παρέχουν μόνο σχέσεις εννοιών. Από ένα θησαυρό δεν μπορεί να εξαχθεί γνώση, ενώ με τη χρήση γλωσσών οντολογιών όπως η OWL DL, μέσω μηχανισμών εξαγωγής συμπερασμάτων, επιτρέπεται η διαμόρφωση άποψης και γνώσης¹².

Η δημιουργία πολυθεματικών θησαυρών θα διευκόλυne την εξ αρχής θεματική οργάνωση και ανάκτηση των πληροφοριών σε συγκεκριμένα θεματικά πεδία και τη

¹² Lauser Boris.

θεματική συσχέτιση των πληροφοριών για την ανάκτηση των σχετιζομένων με οποιοδήποτε τρόπο πληροφοριών και θα διευκόλυε την ανάπτυξη της οντολογίας.

Βιβλιογραφία

Έντυπη:

1. ΑΝΤΩΝΑΚΗΣ Δημήτρης, ΜΗΤΡΕΛΗΣ Άγγελος, ΠΑΠΟΥΤΣΗΣ Κωνσταντίνος –Θεόδωρος, ΣΙΩΧΟΣ Βασίλειος. Οντολογία Ψηφιακής Βιβλιοθήκης. 15^ο Πανελλήνιο Συνέδριο Ακαδημαϊκών Βιβλιοθηκών. Βιβλιοθήκη και Υπηρεσία Πληροφόρησης, Πανεπιστήμιο Πατρών:1-3 Νοεμβρίου 2006
2. ΓΑΪΤΑΝΟΥ Πανωραία, ΓΕΡΓΑΤΣΟΥΛΗΣ Μανόλης. Διαχείριση Οντολογιών: μελέτη και εμβάθυνση στα βασικά προβλήματα που την αφορούν και παρουσίαση υπαρχόντων βιβλιοθηκών οντολογιών. 15^ο Πανελλήνιο Συνέδριο Ακαδημαϊκών Βιβλιοθηκών. Βιβλιοθήκη και Υπηρεσία Πληροφόρησης. Πανεπιστήμιο Πατρών:1-3 Νοεμβρίου 2006
3. ΕΛΟΤ1321. *Τεκμηρίωση-κατευθυντήριες οδηγίες για τη συγκρότηση και ανάπτυξη μονόγλωσσων θησαυρών*, Αθήνα 1993.
4. ΚΥΡΙΑΚΗ – ΜΑΝΕΣΗ, Δάφνη. *Κέντρα τεκμηρίωσης και πληροφόρησης*. Αθήνα: Ίων, 1999, 100 σ.
5. ΜΠΟΥΝΤΟΥΡΗ Λίνα, ΓΕΡΓΑΤΣΟΥΛΗΣ Μανόλης. Σημασιολογική Ολοκλήρωση Δεδομένων με τη χρήση Οντολογιών. 15^ο Πανελλήνιο Συνέδριο Ακαδημαϊκών Βιβλιοθηκών. Βιβλιοθήκη και Υπηρεσία Πληροφόρησης, Πανεπιστήμιο Πατρών:1-3 Νοεμβρίου 2006
6. ΜΠΩΚΟΣ Γ. *Τεχνολογία και Πληροφόρηση*. Αθήνα: Παπασωτηρίου, 2002
7. ΣΚΑΝΔΑΛΗ, Αλκμήνη. *Σχέδιο Σημειώσεων στην Επιστήμη της Πληροφόρησης*. επιμ. Τσάφου Σταματίνα, Αθήνα: 2002, σ. 31
8. ΤΡΙΑΝΤΑΦΥΛΛΙΔΗΣ. *Λεξικό της Κοινής Νεοελληνικής*. Ινστιτούτο Νεοελληνικών Σπουδών Αριστοτέλειο Πανεπιστήμιο, Θεσσαλονίκη: 1998
9. BIOLCHINI Jorge, PATEL Vilma L.. *From Thesauri towards Ontologies?*. Ios: Amsterdam, 2004
10. DAGOBERT, S. Thesauri and ontologies in digital libraries. *Joint Conference on digital libraries*. Portland, 2002
11. FISCHER, Dietrich H.. *From Thesauri towards Ontologies?* GMD – ISI: 1998
12. LAUSER, Boris. *From Thesauri to Ontologies: A short case study in the food safety area in low ontologies is more powerful than thesauri*. Food and Agriculture Organization of the United Nations, Rome
13. NABIL R. Adam, BHARAT K. Bhorgava, MILTON Halem, YESHA Yelena (Eds). *Digital Libraries – Research and Technology Advances*. Germany: Springer, 1996, Chapter10, p. 116, 122 – 123
14. ZAVITSANOS E., PALIOURAS G., VOUIROS G. A., PETRIDIS S.. *Discovering Subsumption Hierarchies of Ontology Concepts from Text Corpora*. *IEEE/WIC/ACM International Conference on Web Intelligence*, 2007

15. Ηλεκτρονική:

16. <http://www.aifb.uni-karlsruhe.de/Publikationen> (YORK, Sure, RUDI Studer. *Semantic Web Technologies for Digital Libraries: Library Management* 26 (415): 190 – 195. April 2005, Special Issue: Semantic Web, ISSN: 0143-5124
17. <http://www.aifb.uni-karlsruhe.de/Publikationen> (DE BRUIJN, Jos. *Using Ontologies – Enabling Knowledge Sharing and Reuse on the Semantic Web: Report on a Study of Literature*, October 2003.
18. <http://ontology.buffalo.edu>
19. <http://protege.stanford.edu>
20. <http://www.wikipedia.org>
21. <http://www.slideshare.net/tag/corrib>