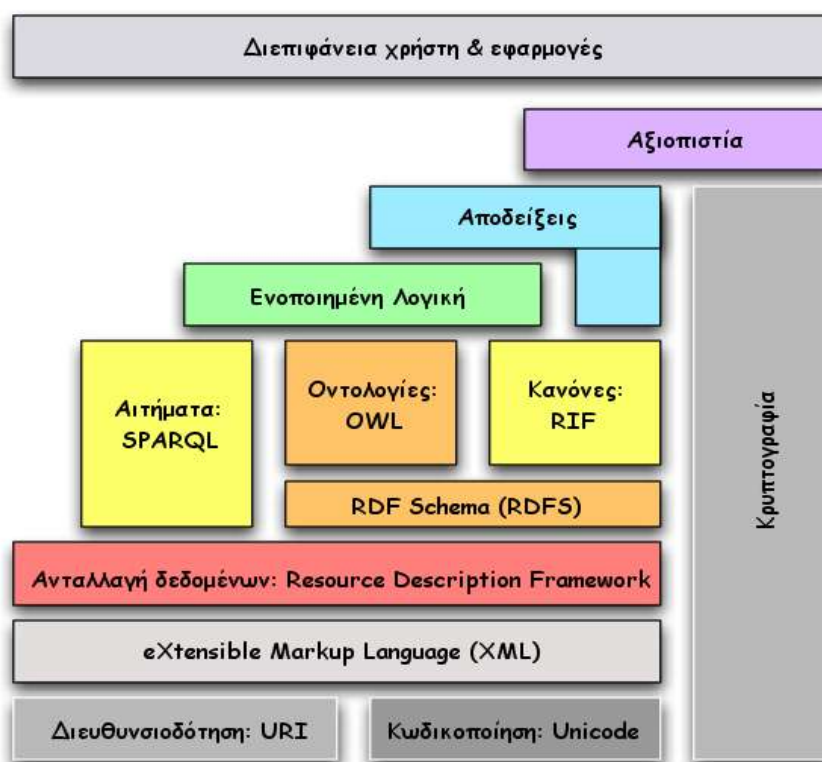




ΕΠΙΣΚΟΠΗΣΗ ΤΟΥ ΣΗΜΑΣΙΟΛΟΓΙΚΟΥ ΙΣΤΟΥ ΚΑΙ ΔΗΜΙΟΥΡΓΙΑ ΟΝΤΟΛΟΓΙΑΣ ΙΣΤΟΥ ΕΥΡΩΠΑΙΟΥ ΠΟΛΙΤΗ



ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ
ΑΘΑΝΑΣΙΟΣ ΚΑΣΑΜΠΑΛΗΣ

ΕΠΙΒΛΕΠΩΝ: ΔΡ. ΠΑΝΑΓΙΩΤΗΣ ΑΔΑΜΙΔΗΣ

ΘΕΣΣΑΛΟΝΙΚΗ 2007

Copyright (c) 2007 Alexander Technological Educational Institute of Thessaloniki.
Copyright (c) 2007 Athanasios Kasampalis.

Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.2 or any later version published by the Free Software Foundation; with the Invariant Sections being LIST THEIR TITLES, with the Front-Cover Texts being LIST, and with the Back-Cover Texts being LIST. A copy of the license is included in the section entitled "GNU Free Documentation License".

Συμβολισμοί και μορφές γραφής

Για να είναι πιο διαισθητικά τα κείμενα στον αναγνώστη, συνοδεύονται συχνά από ορισμένους συμβολισμούς ή γράφονται με διαφορετική μορφή από αυτήν που χρησιμοποιείται στο κυρίως κείμενο. Οι ερμηνείες των συμβολισμών και μορφών γραφής που χρησιμοποιούνται είναι οι εξής:

- ✓ πλεονέκτημα,
- × μειονέκτημα,
- χαρακτηριστικό/ιδιότητα,
- κανόνας/απαίτηση,
- υπόλοιπα (όταν δεν ταιριάζει κανένας από τους προηγούμενους συμβολισμούς),

<http://www.w3.org/People/EM/contact#me> έγκυρος Διαδικτυακός πόρος
(γραμματοσειρά Courier και υπογραμμισμένο κείμενο πόρου),

<http://www.w3.org/People/EM/contact#fake> υποθετικός Διαδικτυακός πόρος
(γραμματοσειρά Courier και μη υπογραμμισμένο κείμενο πόρου),

αυτό είναι ένα παράδειγμα τμήμα κώδικα ή παράδειγμα (γραμματοσειρά Courier και μη υπογραμμισμένο κείμενο).

Πίνακας Περιεχομένων

Abstract	6
Πρόλογος	7
1. Εισαγωγή	8
2. Εισαγωγή στον σημασιολογικό ιστό	11
2.1 Το πρόβλημα του Διαδικτύου σήμερα.....	11
2.2 Ορισμός.....	12
2.3 Οι πράκτορες του σημασιολογικού ιστού.....	13
2.3.1 Χαρακτηριστικά πρακτόρων.....	13
2.3.2 Σύγκριση ειδών πρακτόρων.....	13
2.3.3 Παραδείγματα χρήσης σημασιολογικών πρακτόρων.....	15
2.4 Αναπαράσταση γνώσης.....	17
2.5 Η αρχιτεκτονική του σημασιολογικού ιστού.....	18
2.5.1 Κωδικοποίηση χαρακτήρων.....	19
2.5.2 Διευθυνσιοδότηση εγγράφων.....	19
2.5.3 eXtensible Markup Language (XML).....	21
2.5.4 Ανταλλαγή δεδομένων.....	21
2.5.5 Σύγκριση eXtensible Markup Language με Resource Description Framework.....	23
2.5.6 Resource Description Framework Schema (RDFS).....	25
2.5.7 Οντολογίες.....	26
2.5.8 Κανόνες.....	27
2.5.9 Αιτήματα.....	27
2.5.10 Ενοποιημένη λογική.....	28
2.5.11 Αποδείξεις.....	29
2.5.12 Αξιοπιστία σημασιολογικού ιστού.....	30
2.5.13 Κρυπτογραφία.....	32
2.5.14 Γραφική παρουσίαση του σημασιολογικού ιστού.....	33
2.6 Τα κέρδη που θα επιφέρει η σημασιολογία.....	34
3. Οντολογίες	36
3.1 Ορισμός.....	36
3.2 Ανάγκες δημιουργίας.....	37
3.3 Οι οντολογίες ως μηχανισμός προδιαγραφών.....	37
3.4 Επιπλέον προβλήματα που λύνουν οι οντολογίες.....	39
3.5 Τα χαρακτηριστικά μίας οντολογίας.....	40
3.5.1 Άτομα (individuals).....	40
3.5.2 Κλάσεις (έννοιες).....	40
3.5.3 Ιδιότητες οντολογίας.....	43

3.5.4 Σχέσεις	43
3.5.5 Κατηγορίες οντολογιών.....	45
3.6 Τεχνολογίες δημιουργίας οντολογιών.....	46
3.7 Παραδείγματα οντολογιών.....	50
3.8 Εφαρμογές που χρησιμοποιούν οντολογίες.....	51
3.8.1 Ιστοσελίδα mindswap.org.....	51
3.8.2 Materialized Ontology View Extractor (MOVE) και Semantic Completeness Optimization Scheme (SCOS).....	53
3.8.3 Το σύστημα παροχής βοήθειας σε καθηγητές πληροφορικής Kasai.....	55
3.8.4 Μηχανή αναζήτησης semanticwebsearch.com.....	58
3.8.5 Βάση δεδομένων Oracle 10g.....	60
3.8.6 Πολιτική διαχείρισης ΚΑoS για υπηρεσίες του σημασιολογικού ιστού.....	62
4. Ανάλυση τεχνολογιών σημασιολογικού ιστού.....	68
4.1 eXtensible Markup Language (XML).....	69
4.1.1 Χαρακτηριστικά της eXtensible Markup Language.....	70
4.1.2 Δυνατά σημεία και αδυναμίες της eXtensible Markup Language.....	72
4.1.3 Παράδειγμα εγγράφου eXtensible Markup Language.....	74
4.2 eXtensible Markup Language Schema (XML Schema).....	74
4.2.1 Πότε χρησιμοποιείται η W3C XML Schema.....	75
4.2.2 Παράδειγμα εγγράφου W3C XML Schema.....	75
4.3 Resource Description Framework (RDF).....	76
4.3.1 Που χρειάζεται ο Resource Description Framework.....	76
4.3.2 Το μοντέλο δεδομένων του Resource Description Framework.....	78
4.3.3 Αναπαράσταση τριάδων Resource Description Framework σε γράφημα.....	80
4.3.4 Το συντακτικό του Resource Description Framework.....	81
4.4 Resource Description Framework Schema (RDFS).....	84
4.4.1 Περιγράφοντας κλάσεις στην Resource Description Framework Schema.....	87
4.4.2 Περιγράφοντας ιδιότητες στην Resource Description Framework Schema.....	91
4.5 Web Ontology Language (OWL).....	96
4.5.1 Οι τρεις υπογλώσσες της Web Ontology Language.....	97
4.5.2 Χαρακτηριστικά της Web Ontology Language Lite.....	99
5. Κατασκευή οντολογίας ιστού Ευρωπαίου πολίτη (EuroCitizen).....	111
5.1 Επιλογή τεχνολογίας.....	112
5.2 Επιλογή εργαλείου σύνταξης οντολογίας.....	114
5.3 Η οντολογία ιστού EuroCitizen.....	115
5.3.1 Σκοπός κατασκευής της οντολογίας EuroCitizen.....	115
5.3.2 Ομοιότητες και διαφορές μεταξύ των οντολογιών EuroCitizen και NTN.....	117
5.3.3 Χρήσεις της οντολογίας EuroCitizen.....	120

5.3.4 Ελλείψεις και προβλήματα σχετικά με την οντολογία EuroCitizen.....	121
Επίλογος / Συμπεράσματα.....	123
Βιβλιογραφία.....	125
Παράρτημα Α - Παραδείγματα εγγράφων.....	131
Παράρτημα Β - Τμήματα κώδικα εγγράφων.....	138
Παράρτημα Γ - Εργαλεία ελέγχου εγκυρότητας και μετατροπείς εγγράφων.....	145
Παράρτημα Δ - Άδεια Ελεύθερης Τεκμηρίωσης GNU.....	146

Abstract

This assignment is related with a huge and very important try which has officially started since 2004, and targets to extend the classic world wide web to a semantic web. By the time this try will have been completed, the quality of the Internet will be much better than the current. Both introductory and advanced topics are covered in the assignment. Initially, introductory topics are covered, such as the architecture of the semantic web and the role and characteristics of the web ontologies. Then, more advanced topics are covered, such as the most important technologies of the semantic web, and the methodology of developing a web ontology. Finally, the EuroCitizen ontology is analyzed, which is a web ontology that was developed using the OWL technology, to provide information about the common interests of the European citizens to the semantic agents. The EuroCitizen ontology is the first try to create a web ontology which is related to the European society, and therefore targets to cover an important missing part. The further support and extension of the ontology, will help to create a very useful knowledge base, appropriate to be used in a great number of applications.

The abstract is probably the only part of a writing, where the writer can write in singular. I am taking that chance, to say a great thank to my professor Dr. Panagiotis Adamidis, who gave me valuable help on successfully completing this assignment. I hope that you like it and that you will learn lots of new things by reading it.

Πρόλογος

Αυτή η πτυχιακή εργασία ασχολείται με μία τεράστια και πολύ σημαντική προσπάθεια που έχει ξεκινήσει επίσημα από το 2004, και έχει ως στόχο να επεκτείνει τον κλασικό παγκόσμιο ιστό πληροφοριών σε σημασιολογικό ιστό. Με την ολοκλήρωση αυτής της προσπάθειας, η ποιότητα του Διαδικτύου θα γίνει σαφώς καλύτερη από την τωρινή. Στην εργασία γίνεται αναφορά τόσο σε εισαγωγικά όσο και σε προχωρημένα ζητήματα. Αρχικά καλύπτονται εισαγωγικά ζητήματα όπως η αρχιτεκτονική του σημασιολογικού ιστού καθώς και ο ρόλος και τα χαρακτηριστικά των οντολογιών ιστού. Έπειτα καλύπτονται πιο προχωρημένα ζητήματα, όπως οι σημαντικότερες τεχνολογίες του σημασιολογικού ιστού και η μεθοδολογία κατασκευής μίας οντολογίας ιστού. Τέλος, προτείνεται η οντολογία EuroCitizen, μία οντολογία ιστού που κατασκευάστηκε με τη βοήθεια της τεχνολογίας OWL, ώστε να παρέχει στους σημασιολογικούς πράκτορες πληροφορίες για τις κοινές πεποιθήσεις των πολιτών της Ευρωπαϊκής ένωσης. Η οντολογία EuroCitizen αποτελεί την πρώτη προσπάθεια δημιουργίας μίας οντολογίας ιστού που σχετίζεται με τον Ευρωπαίο πολίτη, και έχει στόχο να καλύψει ένα σημαντικό κενό. Με την περαιτέρω υποστήριξη και επέκταση της, μπορεί να αποτελέσει μία ιδιαίτερα χρήσιμη βάση γνώσης κατάλληλη ώστε χρησιμοποιηθεί σε μία πληθώρα εφαρμογών.

Ο πρόλογος είναι ίσως το μοναδικό τμήμα ενός συγγράμματος, στο οποίο δίνεται η δυνατότητα στον συγγραφέα να μιλήσει σε πρώτο ενικό πρόσωπο. Εκμεταλλευόμενος λοιπόν αυτή την ευκαιρία, αισθάνομαι την ανάγκη να ευχαριστήσω θερμά τον επιβλέποντα καθηγητή μου Δρ. Παναγιώτη Αδαμίδη, η βοήθεια του οποίου υπήρξε πολύτιμη για την ολοκλήρωση αυτής της εργασίας. Ελπίζω να την απολαύσετε και να αποκομίσετε χρήσιμες πληροφορίες κατά την ανάγνωση της.

1. Εισαγωγή

Τα τελευταία χρόνια παρουσιάζεται εκρηκτική αύξηση του αριθμού των υπολογιστικών δικτύων δεδομένων (computer data networks) που δημιουργούνται παγκοσμίως. Τα περισσότερα από αυτά μάλιστα, είναι συνεχώς συνδεδεμένα στο Διαδίκτυο¹, ώστε οι χρήστες τους να μπορούν να επωφελούνται από τις υπηρεσίες του σε μόνιμη βάση.

Καθώς αυξάνονται οι χρήστες που χρησιμοποιούν το Διαδίκτυο, αυξάνεται και η πληροφορία που αυτό προσφέρει. Σήμερα, ο παγκόσμιος ιστός πληροφοριών (world wide web – www) περιέχει περίπου δέκα δισεκατομμύρια ιστοσελίδες! Εξαιτίας όμως αυτής της τεράστιας ποσότητας πληροφορίας που είναι διαθέσιμη σήμερα στο Διαδίκτυο, οι υπηρεσίες του αρχίζουν και γίνονται ολοένα και λιγότερο αξιόπιστες και εύχρηστες [STAA06].

Το καλύτερο παράδειγμα που μπορεί να δοθεί για να γίνει ιδιαίτερα αισθητή η ύπαρξη αυτού του φαινομένου, είναι η χρήση μίας μηχανής αναζήτησης. Ενώ στο παρελθόν ένας χρήστης μπορούσε πολύ εύκολα με τη χρήση μιας μηχανής να εισάγει τις λέξεις κλειδιά (keywords) που τον ενδιέφεραν, και η μηχανή επέστρεφε ικανοποιητικά αποτελέσματα, πλέον η κατάσταση έχει αλλάξει ριζικά.

Αρχικά άλλαξε ο τρόπος λειτουργίας των εφαρμογών που καταχωρούσαν πληροφορία στα ευρετήρια των βάσεων γνώσης (knowledge bases) των μηχανών αναζήτησης (γνωστές και ως αράχνες - spiders, ρομπότ - robots, κτλ), με την προσπάθεια μετάβασης από την σάρωση περιεχομένου στην σάρωση μεταδεδομένων (δεδομένα που χρησιμοποιούνται για να περιγράψουν άλλα δεδομένα). Τα μεταδεδομένα που υποστηρίζει όμως ο παγκόσμιος ιστός πληροφοριών στην τρέχουσα μορφή του, είναι πάρα πολύ φτωχά, και δεν βελτιώνουν τις υπηρεσίες του Διαδικτύου κατά τον απαιτούμενο βαθμό.

Όλα τα παραπάνω υποχρεώνουν έναν χρήστη να επιλέξει κάποια συγκεκριμένη μηχανή

¹ Ακολουθώντας τη σύμβαση της βιβλιογραφίας, η λέξη “Διαδίκτυο” (με κεφαλαίο δέλτα) θα χρησιμοποιείται όταν γίνεται αναφορά στο παγκόσμιο δίκτυο Internet, ενώ η λέξη “διαδίκτυο” (με πεζό δέλτα) όταν γίνεται αναφορά σε ιδιωτικά δίκτυα που χρησιμοποιούν την τεχνολογία TCP/IP.

αναζήτησης, και αφού εκπαιδευτεί ώστε να μάθει καλά τον τρόπο λειτουργίας της (διαφέρει από μηχανή σε μηχανή), να προσπαθεί να εντοπίσει πληροφορίες που σχετίζονται με το θέμα που τον ενδιαφέρει. Και κατά ένα μεγάλο ποσοστό τα αποτελέσματα συνεχίζουν να μην είναι ικανοποιητικά. Τα ίδια ισχύουν και για τις υπόλοιπες υπηρεσίες του Διαδικτύου.

Είναι ανάγκη συνεπώς, να βρεθεί κάποια λύση σε όλα τα παραπάνω προβλήματα. Ο καλύτερος τρόπος για να επιτευχθεί κάτι τέτοιο, σύμφωνα με τους ερευνητές, είναι να μπορούν πλέον οι Η/Υ να “κατανοήσουν” το είδος της πληροφορίας που αναζητούν, παύοντας να την αντιμετωπίζουν ως απλά ηλεκτρικά σήματα². Αν και αυτό ακούγεται αρκετά ουτοπικό, είναι πραγματικότητα και ονομάζεται σημασιολογικός ιστός (semantic web). Σκοπός του σημασιολογικού ιστού είναι να επεκτείνει τον παγκόσμιο ιστό πληροφοριών, περιγράφοντας την πληροφορία χρησιμοποιώντας σημασιολογικό περιεχόμενο (χρησιμοποιώντας κυρίως μεταδεδομένα), ώστε να υπάρχει ένας πρότυπος τρόπος επικοινωνίας και συνεργασίας μεταξύ των μηχανών. Αυτή η υποδομή είναι κατάλληλη, ώστε τις σύνθετες υπηρεσίες που αποτελούν πονοκέφαλο για τους χρήστες, να αναλαμβάνουν να τις φέρουν εις πέρας αυτόματα αυτόνομες εφαρμογές που ονομάζονται πράκτορες (agents).

Αυτή η εργασία, προσπαθεί να αναλύσει τα βασικότερα θέματα που σχετίζονται με το σημασιολογικό ιστό, ξεκινώντας από τις θεμελιώδεις έννοιες του (σημασιολογία, πράκτορες, οντολογίες, κτλ), και καταλήγοντας στο πώς μπορεί κάποιος να συμμετέχει ενεργά στη δημιουργία του (εφαρμογές, τεχνολογίες, παράδειγμα κατασκευής οντολογίας).

Αναλυτικότερα, το δεύτερο κεφάλαιο αποτελεί την πρώτη γνωριμία του αναγνώστη με τον σημασιολογικό ιστό, την αρχιτεκτονική, και τις υπηρεσίες του. Τονίζονται επίσης τα πλεονεκτήματα που παρέχει η χρήση της σημασιολογίας.

Το τρίτο κεφάλαιο αναλύει μία από τις σημαντικότερες τεχνολογικές αλλαγές που κάνουν το σημασιολογικό ιστό αισθητά καλύτερο από τον παγκόσμιο ιστό πληροφοριών. Η αλλαγή αυτή είναι η προσθήκη οντολογιών ιστού. Οι οντολογίες ιστού, που στην πιο απλή περίπτωση αποτελούν προδιαγραφές εννοιών, είναι αναμφισβήτητα ο καλύτερος τρόπος περιγραφής

² Αυτό φυσικά δε σημαίνει σε καμία περίπτωση ότι οι Η/Υ θα πάντων να αντιμετωπίζουν την πληροφορία ως ηλεκτρικά σήματα σε φυσικό επίπεδο (physical layer), η αλλαγή θα γίνει στο επίπεδο εφαρμογής (application layer) της στοιβάς πρωτοκόλλων TCP/IP.

όλων των ειδών δεδομένων (πχ είτε δημόσια είτε ιδιωτικά), έτσι ώστε να είναι εύκολη η αναζήτηση, η εξαγωγή (extraction), η αναπαράσταση (representation), η διερμηνεία (interpretation), και η συντήρηση (maintenance) τους. Στο τρίτο κεφάλαιο περιγράφονται τα προβλήματα που λύνουν οι οντολογίες, τα σημαντικότερα χαρακτηριστικά μίας οντολογίας ιστού, καθώς και εφαρμογές που τις αξιοποιούν στην πράξη.

Στο τέταρτο κεφάλαιο γίνεται ανάλυση των σημαντικότερων τεχνολογιών του σημασιολογικού ιστού. Η καθεμία τεχνολογία βάζει το δικό της λιθαράκι έτσι ώστε να χτιστεί η τελική υποδομή που θα επεκτείνει τον παγκόσμιο ιστό πληροφοριών σε σημασιολογικό ιστό. Συνεπώς, στο τρίτο κεφάλαιο περιγράφονται οι ρόλοι και τα χαρακτηριστικά της κάθε τεχνολογίας.

Στο πέμπτο και τελευταίο κεφάλαιο, αρχικά περιγράφεται η μεθοδολογία κατασκευής μίας οντολογίας ιστού. Έπειτα αναλύεται η οντολογία EuroCitizen, μία οντολογία ιστού για την περιγραφή των χαρακτηριστικών ενός Ευρωπαίου πολίτη και των κοινών πεπιοθήσεων του με τους υπόλοιπους πολίτες της Ευρώπης. Η κατασκευή της οντολογίας EuroCitizen αποτέλεσε τον τελικό στόχο αυτής της εργασίας, και η ιδέα είναι να υλοποιηθεί μία οντολογία ιστού που θα μπορεί να αξιοποιηθεί σε ερευνητικό και ακαδημαϊκό επίπεδο, σε δημόσιες και ιδιωτικές υπηρεσίες, ακόμη και από τελικούς χρήστες, για την υποστήριξη λήψης αποφάσεων και εξαγωγής συμπερασμάτων, για την “εκπαίδευση” και διαβούλευση ενός Ευρωπαίου πολίτη, κτλ. Η υλοποίηση της οντολογίας EuroCitizen θα καλύψει ένα σημαντικό κενό, αφού δεν υπάρχει (ή τουλάχιστον δεν έχει εντοπιστεί) κάποια οντολογία που να ασχολείται με αυτά τα θέματα. Περιγράφονται επίσης και επιπλέον λεπτομέρειες σχετικά με την κατασκευή της οντολογίας EuroCitizen, όπως η τεχνολογία και τα εργαλεία που χρησιμοποιήθηκαν για την ανάπτυξη της, καθώς και οι τρόποι με τους οποίους μπορεί να χρησιμοποιηθεί η οντολογία. Τέλος γίνεται μία αναφορά στα προβλήματα και στις ελλείψεις της οντολογίας στην τρέχουσα μορφή της.

2. Εισαγωγή στον σημασιολογικό ιστό

Ο σημασιολογικός ιστός είναι μία επέκταση του τωρινού παγκοσμίου ιστού πληροφοριών, με τη βοήθεια του οποίου άνθρωποι και Η/Υ θα μπορούν να συνεργάζονται με καλύτερο τρόπο. Η καλύτερη συνεργασία ανθρώπων και Η/Υ θα είναι εφικτή όταν οι Η/Υ θα μπορούν πλέον να “κατανοήσουν” τι είδους πληροφορία αναζητούν και επεξεργάζονται.

Αυτό το κεφάλαιο αρχικά παρουσιάζει στον αναγνώστη τις ανάγκες δημιουργίας του σημασιολογικού ιστού. Έπειτα γίνεται αναφορά στους πράκτορες του σημασιολογικού ιστού και τα χαρακτηριστικά τους. Η τελευταία ενότητα του κεφαλαίου αναλύει την αρχιτεκτονική του σημασιολογικού ιστού, η οποία έχει τη μορφή στοίβας που χωρίζεται σε επίπεδα. Στόχος αυτού του κεφαλαίου είναι να παρέχει στον αναγνώστη μία γενική εικόνα των χαρακτηριστικών και της αρχιτεκτονικής του σημασιολογικού ιστού.

2.1 Το πρόβλημα του Διαδικτύου σήμερα

Σύμφωνα με τους Βλαχάβα Ι, Κεφαλά Π, κ.α. [ΒΛΑΧ06 σελ. 696-700], κάθε είδος γλώσσας, είτε είναι φυσική (για παράδειγμα η ελληνική) είτε είναι τεχνητή (για παράδειγμα η γλώσσα προγραμματισμού C), πρέπει οπωσδήποτε να διαθέτει:

- 1) **Συντακτικό (syntax):** Σύμβολα και κανόνες συνδυασμού των συμβόλων ώστε να δημιουργούνται λέξεις που θεωρούνται έγκυρες.
- 2) **Σημασιολογία (semantics):** Καθορισμός των εννοιών που αποδίδονται στα σύμβολα και στους συνδυασμούς των συμβόλων που επιτρέπει το συντακτικό, ώστε να δημιουργούνται προτάσεις που να έχουν νόημα για τη γλώσσα.

Για παράδειγμα η πρόταση “η μπριζόλα έφαγε το σκύλο” είναι συντακτικά έγκυρη στην ελληνική γλώσσα, αλλά όταν τη διαβάσει ένας άνθρωπος αμέσως καταλαβαίνει ότι η πρόταση είναι σημασιολογικά μη αποδεκτή.

Το μεγαλύτερο ποσοστό του περιεχομένου που υπάρχει σήμερα στο Διαδίκτυο (Internet) είναι σχεδιασμένο ώστε να είναι κατανοητό μόνο από τους ανθρώπους, και όχι από τις εφαρμογές των Η/Υ. Βέβαια οι Η/Υ έχουν τη δυνατότητα να σαρώνουν ιστοσελίδες (websites) και να αναγνωρίζουν ορισμένα πεδία (tags), τις τιμές τους (values), κτλ. Για παράδειγμα μπορούν να καταλάβουν ότι σε μία ιστοσελίδα υπάρχει ένας σύνδεσμος (link) για μία άλλη ιστοσελίδα. Σε καμία περίπτωση όμως δεν μπορούν να “κατανοήσουν” τη σημασιολογία, λόγου χάριν ότι η ιστοσελίδα <http://www.w3.org/People/Berners-Lee/> είναι η προσωπική ιστοσελίδα του Tim Berners Lee, και ότι ο σύνδεσμος “Short Bio” εντός της οδηγεί σε μία σύντομη έκδοχή του βιογραφικού σημειώματος του [LEE01].

Υπηρεσίες όπως η αναζήτηση πληροφοριών, έχουν γίνει πλέον φοβερά κουραστικές για τους χρήστες. Ακόμη και οι καλύτερες μηχανές αναζήτησης αδυνατούν να ανταποκριθούν στις απαιτήσεις τους, λόγω του μεγάλου όγκου των αρχείων καθώς και των πολλαπλών ερμηνειών που μπορεί να έχει μία λέξη. Με την αναζήτηση αλφαριθμητικών (strings) που χρησιμοποιούν οι μηχανές, είναι αδύνατο να κατανοήσουν λόγου χάριν αν η λέξη “μοντέλο” αναφέρεται σε κάποιο μοντέλο μίας μηχανής αυτοκινήτου, σε κάποιο συλλεκτικό μίνι συναρμολογούμενο μοντέλο πολεμικών αεροπλάνων, ή σε κάποιον άνθρωπο που εργάζεται επαγγελματικά με την ιδιότητα του μοντέλου.

Βέβαια, όπως τονίστηκε ήδη, η αναζήτηση δεν είναι το μοναδικό πρόβλημα. Άλλα σοβαρά προβλήματα που σχετίζονται με την πληροφορία είναι η εξαγωγή (extraction), η αναπαράσταση (representation), η διερμηνεία (interpretation), και η συντήρηση (maintenance) της [FENS03].

2.2 Ορισμός

Ο σημασιολογικός ιστός ήρθε για να δώσει λύση σε όλα τα παραπάνω προβλήματα. Αποτελεί επέκταση του υπάρχοντος παγκοσμίου ιστού πληροφοριών (world wide web). Ουσιαστικά πρόκειται για τη δημιουργία ενός περιβάλλοντος το οποίο θα επιτρέπει σε αυτοματοποιημένα εργαλεία λογισμικού που ονομάζονται πράκτορες (software agents), να

διαμοιράζονται και να επεξεργάζονται πληροφορίες και δεδομένα για λογαριασμό των ανθρώπων-χρηστών τους. Με την ανάπτυξη αυτού του περιβάλλοντος, άνθρωποι και Η/Υ θα αποκτήσουν μεγαλύτερη ικανότητα συνεργασίας ([ΒΛΑΧ06 σελ. 640-641], [LEE01]).

2.3 Οι πράκτορες του σημασιολογικού ιστού

Σύμφωνα με τους Βλαχάβα Ι, Κεφαλά Π, κ.α. [ΒΛΑΧ06 σελ. 587], ένας πράκτορας είναι μία οντότητα που αντιλαμβάνεται το περιβάλλον μέσα στο οποίο βρίσκεται με τη βοήθεια αισθητήρων (sensors), είναι μέρος του περιβάλλοντος αυτού, κάνει συλλογισμούς για το περιβάλλον και δρα πάνω σε αυτό με τη βοήθεια μηχανισμών δράσης (effectors), για την επίτευξη κάποιων στόχων.

2.3.1 Χαρακτηριστικά πρακτόρων

Σύμφωνα με τους Βλαχάβα Ι, Κεφαλά Π, κ.α. [ΒΛΑΧ06 σελ. 587], αν και υπάρχουν πολλά είδη πρακτόρων, υπάρχει ένα κοινό χαρακτηριστικό που πρέπει να υπάρχει σε όλα. Το χαρακτηριστικό αυτό είναι η αυτονομία, δηλαδή η αυτενέργεια ενός πράκτορα για την επίτευξη των στόχων του. Για να υπάρχει αυτονομία σε ένα πράκτορα, πρέπει να υπάρχει σε αυτόν ενσωματωμένη “νοημοσύνη”, τουλάχιστον μέχρι κάποιο βαθμό (έστω και μικρό). Για το λόγο αυτό αναφέρεται ότι όλοι οι πράκτορες ανήκουν στην ευρύτερη κατηγορία των ευφυών πρακτόρων.

2.3.2 Σύγκριση ειδών πρακτόρων

Σύμφωνα με τους Βλαχάβα Ι, Κεφαλά Π, κ.α. [ΒΛΑΧ06 σελ. 643-645], αν γίνει μία σύγκριση των σημερινών πρακτόρων που υπάρχουν στο Διαδίκτυο, με τους πράκτορες του σημασιολογικού ιστού, θα εντοπιστούν σαφή πλεονεκτήματα στην δεύτερη περίπτωση. Ένας πράκτορας στο τωρινό διαδίκτυο:

× Πρέπει να γνωρίζει λεπτομέρειες μορφοποίησης κάθε ιστοσελίδας που αντλεί

πληροφορίες, για να ανακτήσει το τμήμα που τον ενδιαφέρει.

- x Πρέπει να χρησιμοποιεί διαφορετική μέθοδο ανάκτησης, για κάθε νέα ιστοσελίδα που ανακτά πληροφορίες. Επιπλέον, η παραμικρή αλλαγή στη δομή-μορφή της ιστοσελίδας συνεπάγεται επαναπρογραμματισμό της μεθόδου εύρεσης πληροφοριών.
- x Έχει μεγάλο κόστος επεκτασιμότητας και συντήρησης.

Αντίθετα, με τους πράκτορες του σημασιολογικού ιστού:

- ✓ Υπάρχει πρόσβαση σε πληροφορίες ανεξάρτητες από τη μορφή της παρουσίασής τους.
- ✓ Οι πληροφορίες είναι συνδεδεμένες με κοινόχρηστο πλαίσιο αναφοράς (οντολογία).
- ✓ Το νόημα των πληροφοριών είναι προσβάσιμο και κατανοητό από όλους τους πράκτορες, ανεξάρτητα από το σκοπό για τον οποίο οι πράκτορες αυτοί αναπτύχθηκαν.

Η λειτουργία των πρακτόρων αναβαθμίζεται συνεχώς. Μέχρι σήμερα οι πράκτορες χρησιμοποιούνταν κυρίως:

- για την εξαγωγή πληροφοριών από ιστοσελίδες με τη χρήση ευριστικών κανόνων (wrappers, crawlers, κτλ),
- ως διαμεσολαβητές (mediators) μεταξύ ετερογενών πηγών πληροφόρησης με αυστηρά καθορισμένη σύνταξη και σημασιολογία ανταλλαγής πληροφοριών.

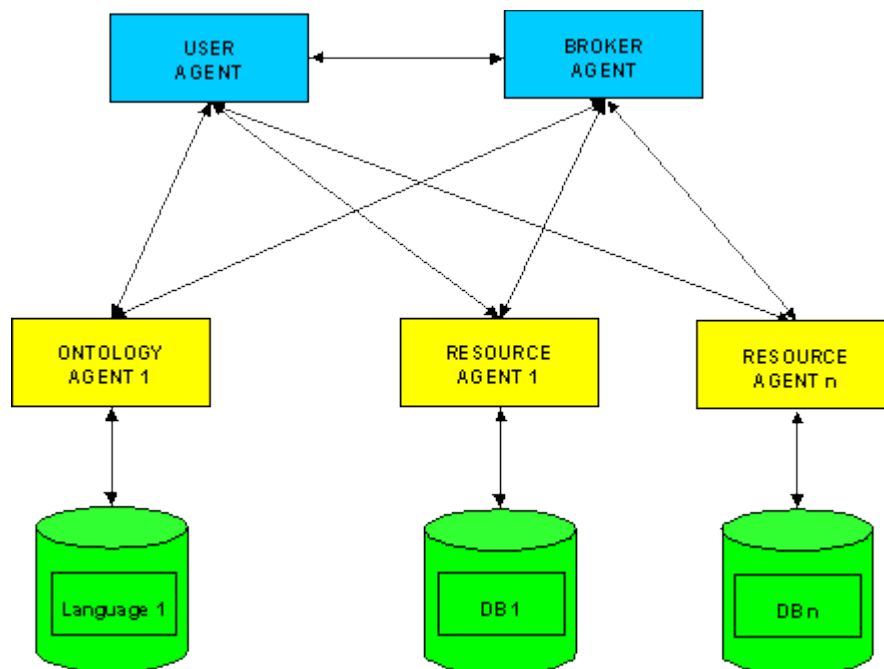
Στον σημασιολογικό ιστό οι πράκτορες αναμένεται να:

- ✓ δέχονται αιτήματα εξυπηρέτησης από τους χρήστες τους,
- ✓ βρίσκουν τις πιθανές απαντήσεις στις απαιτήσεις τους,
- ✓ παρουσιάζουν τις εναλλακτικές επιλογές δράσης, ώστε οι χρήστες να παίρνουν την τελική απόφαση.

Ένα παράδειγμα επικοινωνίας μεταξύ πολλαπλών ειδών πρακτόρων (πράκτορες χρηστών, πράκτορες πρόσβασης σε πόρους, κτλ) του σημασιολογικού ιστού φαίνεται στην εικόνα 2-1.

Κάθε χρήστης του σημασιολογικού ιστού, θα μπορεί να χρησιμοποιεί με ευκολία έναν πράκτορα χρήστη (user agent), ώστε να του υποβάλλει τα κατάλληλα αιτήματα, τα οποία διαφέρουν ανάλογα με τις ανάγκες του προβλήματος που αντιμετωπίζει (πχ αναζήτηση μεταχειρισμένου αυτοκινήτου για αγορά). Ο πράκτορας χρήστη, επικοινωνεί

χρησιμοποιώντας ένα κοινό πρωτόκολλο επικοινωνίας (πχ KQML), με πράκτορες οι οποίοι έχουν πρόσβαση σε οντολογίες (ontology agents), βάσεις δεδομένων (resource agents), κτλ. που είναι συνδεδεμένες με κοινόχρηστο πλαίσιο αναφοράς, και παρουσιάζει τα αποτελέσματα στον χρήστη. Το σημαντικό είναι ότι ο χρήστης δεν χρειάζεται να ασχολείται καθόλου με τις διαδικασίες που ακολουθεί ένας πράκτορας, από τη στιγμή που θα του υποβάλλει ένα αίτημα μέχρι τη στιγμή που θα λάβει κάποιο αποτέλεσμα. Όλες αυτές οι σύνθετες διαδικασίες είναι διαφανείς και δεν γίνονται αντιληπτές από τον χρήστη [LEE01].



Εικόνα 2-1 Επικοινωνία μεταξύ πολλαπλών ειδών πρακτόρων.

2.3.3 Παραδείγματα χρήσης σημασιολογικών πρακτόρων

Ένας πράκτορας του σημασιολογικού ιστού είναι σχεδιασμένος έτσι ώστε όχι μόνο να γνωρίζει για παράδειγμα ότι οι λέξεις κλειδιά της ιστοσελίδας του λειτουργικού συστήματος GNU είναι οι "λειτουργικό σύστημα GNU, Linux, Emacs, πυρήνας, Ίδρυμα Ελεύθερου Λογισμικού, Hurd" (κάτι το οποίο ισχύει με τη σημερινή κατάσταση), αλλά "ξέρει" επιπλέον ότι τα γραφεία του ιδρύματος ελεύθερου λογισμικού (Free Software Foundation) που είναι ο

σπόνσορας του GNU στεγάζονται στη Βοστώνη των Ηνωμένων Πολιτειών Αμερικής, ότι είναι ανοιχτά τις ημέρες Τετάρτη ως Παρασκευή και ώρες 10:00-14:00 και 18:00-20:00, ότι αρχικός ιδρυτής του Ιδρύματος Ελεύθερου Λογισμικού είναι ο Richard Stallman, κτλ.

Και όλα αυτά επιτυγχάνονται δίχως να χρειάζεται να κατασκευαστούν εφαρμογές που πρέπει να ενσωματώνουν νοημοσύνη παρόμοια με αυτή που παρουσιάστηκε σε ταινίες επιστημονικής φαντασίας όπως ο πόλεμος των άστρων. Με χρήση βοηθητικών εφαρμογών ανάπτυξης, η κατασκευή μίας ιστοσελίδας, εφαρμογής, βάσης δεδομένων, κτλ που ενσωματώνει σημασιολογία δεν είναι δύσκολη υπόθεση [LEE01].

Ένα καλό παράδειγμα χρήσεως ενός πράκτορα του σημασιολογικού ιστού, είναι το παρακάτω:

Έστω ότι ο x θέλει να πάει διακοπές με τη γυναίκα του y και αποφασίζει να κανονίσει όλες τις λεπτομέρειες που αφορούν τις διακοπές τους. Αποφασίζουν να επισκεφτούν το Ηράκλειο Κρήτης. Θέλουν να ξεκινήσουν αεροπορικώς από τη Θεσσαλονίκη στις 5 Αυγούστου. Οι πτήσεις που τους βολεύουν είναι μετά τις 14:00, και το αεροπορικό εισιτήριο θέλουν να κοστίζει το μέγιστο 200 ευρώ για κάθε άτομο. Το δωμάτιο που θα μείνουν θέλουν να είναι δίκλινο, με τηλεόραση και ψυγείο, και να απέχει το πολύ 1 χιλιόμετρο από τη θάλασσα.

Το μόνο που θα χρειαστεί να κάνει ο x για να αναζητήσει τις πληροφορίες που ψάχνει, είναι να δώσει τις σωστές παραμέτρους στον πράκτορα του. Έπειτα ο πράκτορας αρχικά επικοινωνεί με πράκτορες ταξιδιωτικών γραφείων της Θεσσαλονίκης, ώστε να εντοπίσει όλα όσα προσφέρουν πτήσεις στις 5 Αυγούστου για το Ηράκλειο, μετά τις 14:00 και με τιμή το πολύ 200 ευρώ. Αφού δημιουργήσει τη λίστα των γραφείων, η οποία ενδεχομένως να είναι κενή, την επιστρέφει στον x . Έπειτα ο x διαλέγει όποια επιλογή τον βολεύει περισσότερο, ή αν δεν έχει βρεθεί αποτέλεσμα που να ταιριάζει δίνει άλλες παραμέτρους στον πράκτορα. Κατά τον ίδιο τρόπο μετά την ολοκλήρωση της πρώτης φάσης, ο πράκτορας του x θα επικοινωνήσει με πράκτορες που γνωρίζουν πληροφορίες για τα ενοικιαζόμενα δωμάτια του Ηρακλείου ολοκληρώνοντας κάποια στιγμή τη διαδικασία.

Το σημαντικό και αυτό που πρέπει να γίνει κατανοητό από αυτό το παράδειγμα είναι ότι με τη

βοήθεια των πρακτόρων οι χρήστες θα πάψουν να ταλαιπωρούνται, αφού οι πράκτορες θα αναλαμβάνουν να επιλύσουν αυτόματα το πρόβλημα ανταλλάσσοντας πληροφορίες μεταξύ τους. Έτσι λοιπόν ένα παράδειγμα που για τα σημερινά δεδομένα του Διαδικτύου μοιάζει με σενάριο επιστημονικής φαντασίας, με τον σημασιολογικό ιστό μπορεί να γίνει πραγματικότητα.

2.4 Αναπαράσταση γνώσης

Σύμφωνα με τους Tim Berners-Lee, James Hendler, κ.α. [LEE01], η τεχνολογία που επιτρέπει στους Η/Υ να έχουν πρόσβαση σε δομημένες συλλογές πληροφοριών και σύνολα συμπερασματικών κανόνων έτσι ώστε να μπορούν να κάνουν αυτόματους συλλογισμούς, ονομάζεται αναπαράσταση γνώσης.

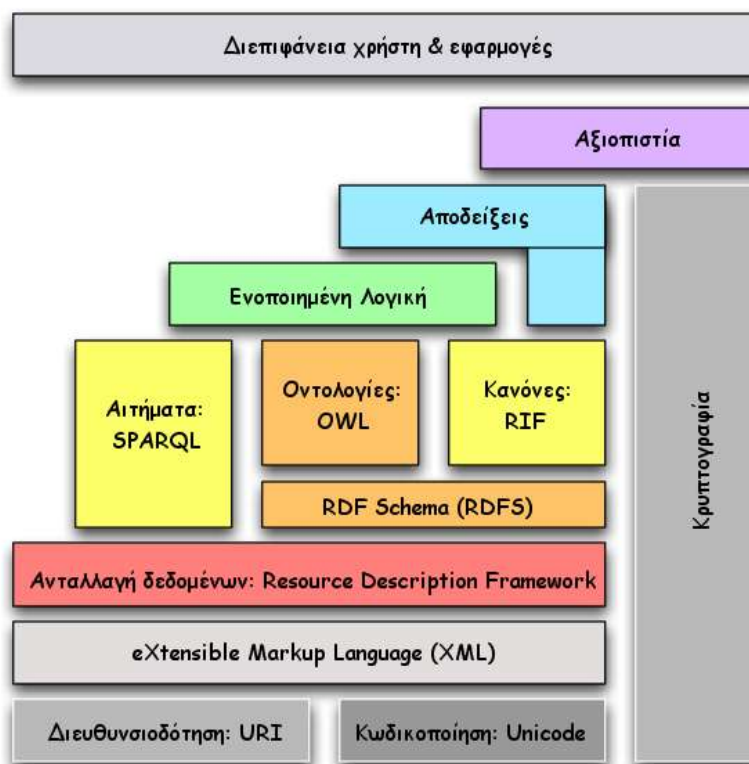
Η πρόκληση του σημασιολογικού ιστού είναι να παρέχει μία γλώσσα μέσω της οποίας θα είναι δυνατόν να εκφραστούν δεδομένα και συλλογιστικοί κανόνες που δρουν πάνω στα δεδομένα, και ταυτόχρονα να επιτρέπεται να εξάγονται στον ιστό του Διαδικτύου κανόνες από οποιοδήποτε ήδη υπάρχον σύστημα αναπαράστασης γνώσης. Για να επιτευχθεί κάτι τέτοιο θα πρέπει η αναπαράσταση της γνώσης στον σημασιολογικό ιστό να είναι ενιαία και ευρέως αποδεκτή.

Δύο σημαντικές τεχνολογίες που χρησιμοποιούνται για την αναπαράσταση της γνώσης και συμμετέχουν στην ανάπτυξη του σημασιολογικού ιστού είναι η επεκτάσιμη γλώσσα σήμανσης (eXtensible Markup Language - XML) και ο σκελετός περιγραφής πόρων (Resource Description Framework - RDF), οι οποίες θα περιγραφούν παρακάτω [ΚΟΥΡ02]. Η τρίτη τεχνολογία που ήρθε ώστε να συμπληρώσει τα κενά του RDF είναι οι οντολογίες (ontologies). Σύντομη αναφορά στις οντολογίες θα γίνει παρακάτω και αναλυτική περιγραφή τους θα γίνει σε επόμενη ενότητα.

2.5 Η αρχιτεκτονική του σημασιολογικού ιστού

Μία μορφή της αρχιτεκτονικής του σημασιολογικού ιστού φαίνεται στην εικόνα 2-2 [LEE06]. Η αρχιτεκτονική του σημασιολογικού ιστού έχει τη μορφή στοίβας που χωρίζεται σε οκτώ επίπεδα (το επίπεδο της διεπιφάνειας χρήστη και των εφαρμογών ουσιαστικά δεν ανήκει στη στοίβα). Ορισμένα επίπεδα χωρίζονται σε τμήματα, και άλλα παρέχουν υπηρεσίες στα υπόλοιπα (αιτήματα, κρυπτογραφία, κτλ). Τα χαμηλότερα επίπεδα σχετίζονται με λεπτομέρειες (διευθυνσιοδότηση, ονοματοδοσία, κωδικοποίηση, κτλ) που αφορούν τα έγγραφα που δημιουργούνται με τις διάφορες τεχνολογίες (XML, XML Schema, RDF, κτλ), ενώ τα ανώτερα επίπεδα ασχολούνται με υπηρεσίες που θεωρούνται αναγκαίες για την ορθή λειτουργία του σημασιολογικού ιστού (κρυπτογραφία, αποδείξεις, αξιοπιστία, κτλ).

Παρακάτω θα γίνει αναφορά σε καθένα από τα τμήματα των επιπέδων ξεχωριστά μόνο σε θεωρητικό επίπεδο. Θέματα όπως το συντακτικό των τεχνολογιών που συμμετέχουν στο σημασιολογικό ιστό θα περιγραφούν σε επόμενο κεφάλαιο.



Εικόνα 2-2 Η αρχιτεκτονική του σημασιολογικού ιστού (Ιούλιος 2006).

2.5.1 Κωδικοποίηση χαρακτήρων

Το πρώτο επίπεδο ξεκινά με τμήμα της κωδικοποίησης. Κωδικοποίηση είναι η διαδικασία μετατροπής των πληροφοριών (χαρακτήρες, λέξεις, φράσεις, κτλ) σε δεδομένα (δυαδικές τιμές). Αποκωδικοποίηση είναι η αντίστροφη διαδικασία.

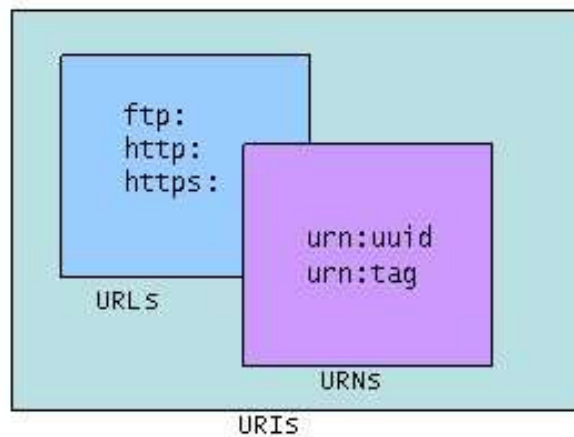
Η κωδικοποίηση των εγγράφων γίνεται με χρήση του διεθνώς καθιερωμένου προτύπου κωδικοποίησης Unicode. Το πρότυπο Unicode μπορεί και υποστηρίζει όλες τις γλώσσες, ακόμη και τις πιο πλούσιες και σύνθετες, όπως είναι για παράδειγμα τα παραδοσιακά Κινέζικα ([BLAX06 σελ. 650], [WICH06]).

2.5.2 Διευθυνσιοδότηση εγγράφων

Το αριστερό τμήμα του πρώτου επιπέδου αφορά τη διευθυνσιοδότηση. Διευθυνσιοδότηση είναι η χρήση αλφαριθμητικών που χαρακτηρίζουν μοναδικά οποιοδήποτε Διαδικτυακό πόρο.

Η διευθυνσιοδότηση των εγγράφων γίνεται με τον ενιαίο προσδιοριστή πόρων (Uniform Resource Identifier - URI), καθώς και τον διεθνή προσδιοριστή πόρων (Internationalized Resource Identifier - IRI), που διαφέρει από τον URI στο ότι παρέχει υποστήριξη χρήσης χαρακτήρων που δεν περιορίζονται στον κώδικα χαρακτήρων ASCII. Ο πιο γνωστός τύπος URI είναι ο ενιαίος εντοπιστής πόρων (Uniform Resource Locator - URL), μέσω του οποίου δηλώνονται οι σύνδεσμοι (links) στη γλώσσα σήμανσης υπερκειμένου (Hypertext Markup Language - HTML). Με τη χρήση των URI/IRI είναι πολύ εύκολο για τον καθένα να προσθέτει νέες έννοιες (concepts) στη σημασιολογία που κατασκευάζει. Άλλα πλεονεκτήματα των URI/IRI είναι ότι είναι ανεξάρτητοι περιεχομένου και πρωτοκόλλου επικοινωνίας, καθώς και το ότι η τεχνολογία τους είναι επεκτάσιμη ([BLAX06 σελ. 650], [LEE01], [WIIN06]). Παραδείγματα των διάφορων μορφών του URI φαίνονται στην εικόνα 2-3 [STEP04].

Όπως μπορεί να παρατηρηθεί και στην εικόνα, ένας URI μπορεί να έχει πολλές μορφές, είτε ενιαίου εντοπιστή πόρων (Uniform Resource Locator - URL) (πχ ftp:, http:, mailto:, news:, https:, κτλ), είτε ενιαίου ονόματος πόρων (Uniform Resource Name - URN) (urn:isbn:0-395-



Εικόνα 2-3 Παραδείγματα διάφορων μορφών του URI.

36341-1, urn:about:hello world, κτλ). Το URN ουσιαστικά είναι ένας URI που σχετίζεται με κάποιο ονοματοχώρο [WIUR06]. Αναφορά στους ονοματοχώρους θα γίνει σε μεταγενέστερη ενότητα.

Με χρήση των προσδιοριστών πόρων (URI/IRI) γίνεται επίσης και η ονοματοδοσία. Ονοματοδοσία είναι η διαδικασία δήλωσης ονομάτων στα στοιχεία (elements) που ανήκουν σε ένα έγγραφο (πχ XML, RDF, κτλ) [LEE01].

Η χρήση της ονοματοδοσίας στο σημασιολογικό ιστό είναι απαραίτητη ώστε να μην υπάρχει σύγχυση από τις εφαρμογές λογισμικού όταν δηλώνονται στοιχεία με το ίδιο όνομα. Για παράδειγμα έστω ότι μία εφαρμογή λογισμικού αναλαμβάνει να προσπελάσει ένα έγγραφο το οποίο περιέχει δύο φορές το στοιχείο αλληλογραφία. Στην μία περίπτωση το στοιχείο χρησιμοποιείται για να αναφερθεί στην παραδοσιακή διεύθυνση αλληλογραφίας ενός ατόμου, και στην άλλη στην ηλεκτρονική (email). Ο μόνος τρόπος για να μπορέσει να ξεχωρίσει τα δύο στοιχεία η εφαρμογή είναι η χρήση διαφορετικού προσδιοριστή. Μία άλλη λύση θα μπορούσε να είναι η δήλωση του δεύτερου στοιχείου με διαφορετικό όνομα, αλλά κάτι τέτοιο και δεν είναι καθόλου πρακτικό σε πολλές περιπτώσεις, και δεν είναι σίγουρο ότι θα το υιοθετήσουν όλοι οι δημιουργοί εγγράφων του κόσμου!

2.5.3 eXtensible Markup Language (XML)

Το δεύτερο επίπεδο αφορά τη χρήση της τεχνολογίας XML. Με τη βοήθεια της XML και των βοηθητικών τεχνολογιών που τη συνοδεύουν (DTD, XML Schema, XLL) επιτυγχάνεται η αναπαράσταση του περιεχομένου των εγγράφων. Η XML κατασκευάστηκε για να ξεπεραστεί ένα βασικό μειονέκτημα της HTML: η παροχή ενός σταθερού συνόλου προκαθορισμένων στοιχείων, τα οποία συμβολίζουν και οριοθετούν τα περιεχόμενα μιας ιστοσελίδας (πχ οι πίνακες, οι σύνδεσμοι, οι εικόνες, κτλ). Λόγω αυτού του σταθερού συνόλου στοιχείων που υποστηρίζει η HTML, κρίνεται ως ακατάλληλη για χρήση σε πολλά είδη δομημένων εγγράφων.

Η XML ανήκει στην κατηγορία των μετά-γλωσσών, δηλαδή είναι γλώσσα που χρησιμοποιείται για την περιγραφή μιας νέας γλώσσας. Διαθέτει ευέλικτο συντακτικό για τη δημιουργία οποιωνδήποτε στοιχείων και την περιγραφή οποιουδήποτε είδους εγγράφου. Η XML είναι κατασκευασμένη έτσι ώστε να επιτρέπει στον καθένα να δημιουργήσει μία δική του δομή για την περιγραφή ενός εγγράφου, αλλά δεν παρέχει καμία ερμηνεία για τη σημασία αυτής της δομής ([BLAX06 σελ. 651-652], [LEE01]).

2.5.4 Ανταλλαγή δεδομένων

Στο τρίτο επίπεδο της στοίβας υπάρχει αποκλειστικά η τεχνολογία Resource Description Framework (RDF), η οποία είναι ένας σκελετός περιγραφής πόρων. Με τον σκελετό RDF εκφράζεται η σημασία (meaning) των εννοιών που περιγράφονται σε ένα έγγραφο. Ο σκελετός είναι κοινός και χρησιμοποιείται από όλους τους δημιουργούς εγγράφων. Αυτό σημαίνει ότι με χρήση του RDF η ανταλλαγή δεδομένων είναι εύκολη, αφού όλα τα έγγραφα RDF έχουν κοινή μορφή. Ο RDF [BLAX06 σελ. 652-653]:

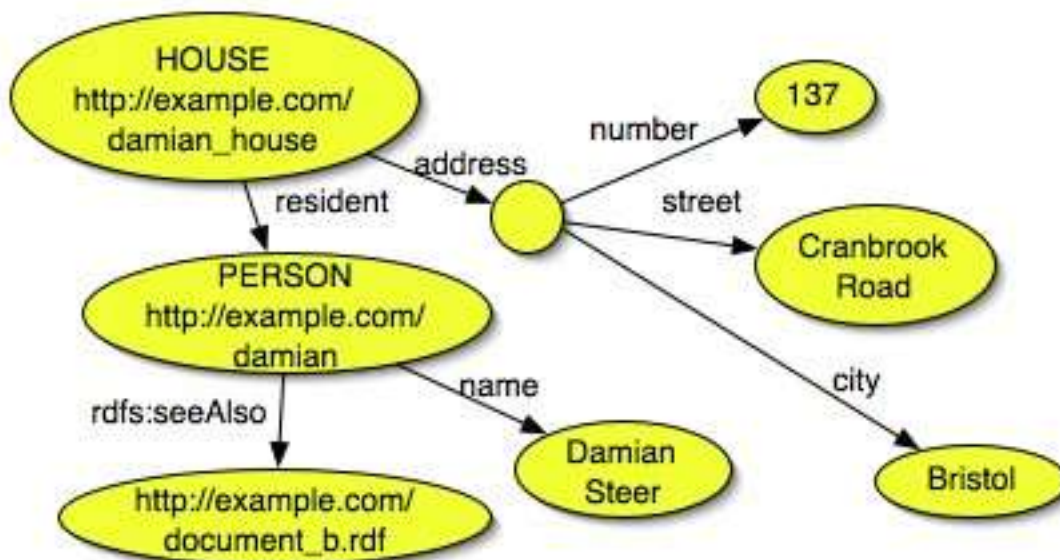
- αποτελείται από δηλώσεις ή ισχυρισμούς (statements) οι οποίες αφορούν μεταδεδομένα,
- δεν εξαρτάται καθόλου από το δεντρικό μοντέλο της XML,
- χρησιμοποιεί τη σύνταξη της XML για την αναπαράσταση δηλώσεων.

Ο RDF εκφράζει τη σημασία χωρίζοντας την σε τρία τμήματα: στο υποκείμενο (subject), στο κατηγορημα (predicate) και στο αντικείμενο (object). Αυτά τα τμήματα μπορούν να γραφούν

με τη μορφή πεδίων της XML. Η διευθυνσιοδότηση τους μπορεί να γίνει με χρήση των προσδιοριστών πόρων. Στον RDF, ένα έγγραφο περιέχει δηλώσεις αντικειμένων (όπως έπιπλα πχ “τραπέζι”, μουσικά κομμάτια πχ “Georgia on my mind”, ή οτιδήποτε άλλο) που έχουν ιδιότητες (όπως “έχει χρώμα”, “γράφηκε από τον”) με τιμές (όπως “καφέ”, “Ray Charles”). Αυτή η δομή καταλήγει να είναι ο φυσικός τρόπος περιγραφής του μεγαλύτερου μέρους των δεδομένων που επεξεργάζονται οι μηχανές. Το πλεονέκτημα της χρήσης των προσδιοριστών για την κωδικοποίηση των τριών τμημάτων του RDF, είναι η εξασφάλιση ότι οι έννοιες που περιγράφονται δεν είναι ανεξάρτητες λέξεις που υπάρχουν εντός ενός κειμένου, αλλά στενά συνδεδεμένες έννοιες ώστε να παράγουν έναν ενιαίο προσδιορισμό που μπορεί κάποιος να αναζητήσει στο Διαδίκτυο [LEE01].

Ένα πολύ χρήσιμο χαρακτηριστικό του RDF είναι ότι λόγω της δομής του μπορεί να απεικονιστεί εύκολα με τη μορφή γραφήματος. Ένα τέτοιο παράδειγμα φαίνεται στην εικόνα 2-4 [STEP04].

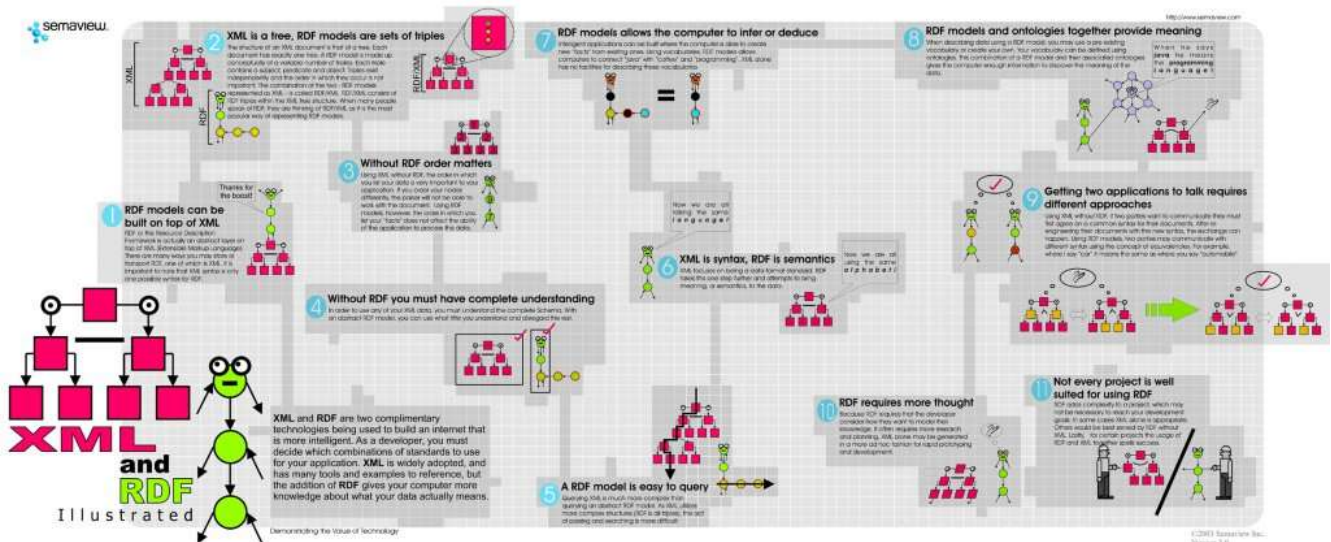
Η εικόνα περιγράφει ένα σπίτι (HOUSE). Η διεύθυνση (address) του σπιτιού βρίσκεται στην πόλη (city) Bristol, στην οδό (street) Cranbrook Road και στον αριθμό (number) 137. Άλλα χαρακτηριστικά του σπιτιού είναι ο κάτοικος (PERSON) του, το ονοματεπώνυμο του οποίου είναι Damian Steer. Τέλος, υπάρχει και μία παραπομπή (seeAlso) σε ένα άλλο έγγραφο RDF.



Εικόνα 2- 4 Το γράφημα της δομής ενός εγγράφου RDF.

2.5.5 Σύγκριση eXtensible Markup Language με Resource Description Framework

Οι λειτουργίες που αναλαμβάνουν να διεκπεραιώσουν οι τεχνολογίες XML και RDF στα πλαίσια του σημασιολογικού ιστού απεικονίζονται στην εικόνα 2-5 [LEE01].



Εικόνα 2-5 XML και RDF στα πλαίσια του σημασιολογικού ιστού.

Η εικόνα 2-5 συνοπτικά αναφέρει τα εξής:

1. Τα μοντέλα RDF χτίζονται "πάνω" στην XML. Αυτό σημαίνει ότι ο RDF υπάρχει σε αφαιρετικό επίπεδο, και συνεπώς η XML είναι ένας μόνο από τους τρόπους σύνταξης του RDF.

2. Η XML έχει δενδρική δομή, ενώ ο RDF είναι ένα σύνολο από τριάδες (triples). Κάθε τριάδα περιέχει το υποκείμενο (subject), το κατηγορημα (predicate) και το αντικείμενο (object). Οι τριάδες είναι ανεξάρτητες ως προς το περιεχόμενο και τη σειρά τοποθέτησης. Ο συνδυασμός των δύο - μοντέλων RDF που αναπαριστώνονται μέσω της XML - ονομάζεται RDF/XML. Η RDF/XML περιλαμβάνει τριάδες RDF εντός μίας δεντρικής δομής XML.

3. Δίχως τη χρήση του RDF, η σειρά τοποθέτησης των δηλώσεων έχει σημασία. Με χρήση της XML δίχως τον RDF, ο τρόπος τοποθέτησης των στοιχείων είναι σημαντικός, διότι σε περίπτωση που αλλάξει η εφαρμογή προσπέλασης δεν μπορεί να προσπελάσει το έγγραφο. Χρησιμοποιώντας μοντέλα RDF, η εφαρμογή καθίσταται ικανή να επεξεργαστεί τα δεδομένα ανεξαρτήτως της σειράς τοποθέτησης.

4. Χωρίς τον RDF, πρέπει να υπάρχει πλήρης κατανόηση ενός σχήματος XML, για να είναι δυνατή η χρήση των δεδομένων XML. Με χρήση ενός αφηρημένου μοντέλου RDF, μπορεί να χρησιμοποιηθεί μόνο το τμήμα των δεδομένων που θεωρείται κατανοητό και αναγκαίο, τα υπόλοιπα μπορούν να παραληφθούν.

5. Είναι εύκολη η υποβολή αιτημάτων σε ένα μοντέλο RDF. Η υποβολή αιτημάτων σε έγγραφα XML είναι μία πιο σύνθετη διαδικασία συγκριτικά με την υποβολή σε ένα αφηρημένο μοντέλο RDF. Επειδή η XML χειρίζεται σύνθετες δομές (ενώ ο RDF μόνο τριάδες), η προσπέλαση και η αναζήτηση δεδομένων είναι πιο δύσκολη.

6. Η XML είναι συντακτικό, ενώ ο RDF σημασιολογία. Ενώ η XML εστιάζει στη δημιουργία μίας πρότυπη μορφής δεδομένων, ο RDF επεκτείνει αυτή τη μορφή εισάγοντας τη λεγόμενη σημασιολογία, δηλαδή νόημα στα δεδομένα.

7. Τα μοντέλα RDF επιτρέπουν στον Η/Υ να συμπεράνει (infer) ή να συνάγει (induce). Με τη χρήση του RDF μπορούν να κατασκευαστούν ευφυείς εφαρμογές όπου ο Η/Υ μπορεί να δημιουργήσει νέα γεγονότα (facts) από τα ήδη υπάρχοντα. Πχ ένα μοντέλο RDF επιτρέπει τη συσχέτιση μεταξύ της γλώσσας προγραμματισμού "Java", του "καφέ" και του προγραμματισμού. Η XML δεν παρέχει υποστήριξη για περιγραφή τέτοιου είδους συσχετίσεων.

8. Τα μοντέλα RDF σε συνδυασμό με τις οντολογίες παρέχουν νόημα. Όταν περιγράφονται δεδομένα χρησιμοποιώντας ένα μοντέλο RDF, μπορεί να χρησιμοποιηθεί ένα ήδη υπάρχον λεξιλόγιο ή να κατασκευαστεί ένα νέο. Το νέο λεξιλόγιο μπορεί να οριστεί με χρήση των οντολογιών. Ο συνδυασμός ενός μοντέλου RDF και των συσχετιζόμενων οντολογιών του επιτρέπει στον Η/Υ να ανακαλύψει το νόημα των δεδομένων.

9. Δίχως τη χρήση του RDF, η επικοινωνία μεταξύ δύο εφαρμογών προαπαιτεί τη συμφωνία για χρήση ενός κοινού λεξιλογίου για τα έγγραφα τους. Χρησιμοποιώντας τον RDF, η επικοινωνία μπορεί να επιτευχθεί χρησιμοποιώντας τις λεγόμενες έννοιες (concepts) ή ισοδυναμίες (equivalencies). Πχ μπορεί να οριστεί ότι η λέξη "car" της μίας εφαρμογής ισοδυναμεί με τη λέξη "automobile" της άλλης, κτλ.

10. Ο RDF απαιτεί περισσότερη σκέψη. Αυτό ισχύει επειδή απαιτεί από το δημιουργό σχεδιασμό για τον τρόπο μοντελοποίησης της γνώσης, και συνεπάγεται ότι συνήθως απαιτείται περισσότερη έρευνα και σχεδιασμός. Υπάρχουν περιπτώσεις που η XML είναι καταλληλότερη για τη γρήγορη δημιουργία πρωτοτύπων και την ανάπτυξη εφαρμογών.

11. Ο RDF δεν είναι ο καταλληλότερος για όλα τα έργα. Ο RDF προσθέτει πολυπλοκότητα σε ένα έργο, κάτι το οποίο ίσως να μην είναι αναγκαίο για την επίτευξη των εκάστοτε στόχων μιας εφαρμογής. Υπάρχουν και περιπτώσεις που η XML από μόνη της αρκεί.

2.5.6 Resource Description Framework Schema (RDFS)

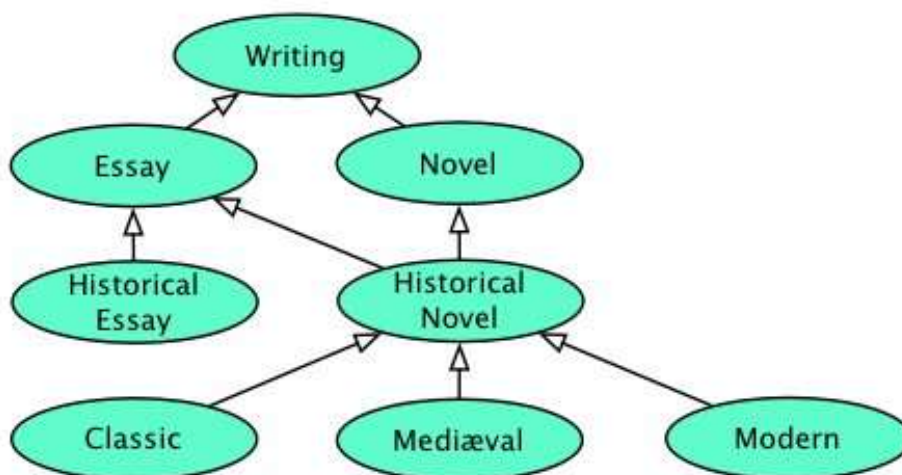
Το τέταρτο επίπεδο περιέχει την γλώσσα περιγραφής λεξιλογίων Resource Description Framework Schema (RDFS). Η τεχνολογία RDF είναι πολύ γενική και δεν παρέχει εργαλεία για δόμηση και οργάνωση της πληροφορίας του σημασιολογικού ιστού. Τη λύση σε αυτά τα προβλήματα δίνει η RDFS, η οποία προσφέρει τη δυνατότητα δηλώσεων ιδιοτήτων και τον ορισμό του τρόπου χρήσης αυτών. Με αυτό τον τρόπο μπορούν να διεξαχθούν στοιχειώδεις έλεγχοι σε ένα έγγραφο RDF. Η RDFS καθορίζει αν είναι επιτρεπτή κάποια δήλωση στο σκελετό RDF [BLAX06 σελ. 653]. Για παράδειγμα με την RDFS μπορεί να καθοριστεί ότι το πεδίο "Χρονολογία Γέννησης" στην ταυτότητα ενός Έλληνα πολίτη θα δέχεται μόνο αριθμούς και παύλες με τη μορφή HH-MM-EEEE (H=ημερομηνία, M=μήνας, E=έτος).

2.5.7 Οντολογίες

Βέβαια η ιστορία του σημασιολογικού ιστού δεν τελειώνει εδώ. Όποιος έχει εμπειρία με τον προγραμματισμό και με τη σχεδίαση βάσεων δεδομένων, θα γνωρίζει ήδη ότι είναι αδύνατο όλοι οι προγραμματιστές και οι σχεδιαστές βάσεων του κόσμου να χρησιμοποιούν τις ίδιες ονομασίες για τη δήλωση των μεταβλητών και των πεδίων των εφαρμογών τους. Αυτό έχει ως αποτέλεσμα να υπάρχουν πεδία με διαφορετικά ονόματα αλλά την ίδια ακριβώς σημασία.

Τέτοιου είδους πεδία πρέπει να ανακαλύπτονται από τους πράκτορες του σημασιολογικού ιστού όταν αναζητούν πληροφορίες. Η λύση σε αυτό το πρόβλημα δίνεται από τις οντολογίες, οι οποίες υπάρχουν στο πέμπτο επίπεδο της στοίβας. Ένας σύντομος ορισμός των οντολογιών του σημασιολογικού ιστού είναι: κοινά και διαμοιραζόμενα λεξικά ορολογίας με σκοπό την αμοιβαία αλληλοκατανόηση των όρων που περιλαμβάνονται σε έγγραφα που ανταλλάσσονται μεταξύ των πρακτόρων. Για τη δημιουργία οντολογιών έχει σχεδιαστεί η γλώσσα δημιουργίας οντολογιών ιστού (Web Ontology Language - OWL) από την κοινοπραξία του παγκόσμιου ιστού πληροφοριών (World Wide Web Consortium - W3C) ([BLAX06 σελ. 642-643], [LEE01]).

Για παράδειγμα ένα πρόβλημα που θα μπορούσε να λυθεί με τη σχεδίαση μίας οντολογίας είναι ο καθορισμός του είδους ενός εγγράφου (άρθρο, μυθιστόρημα, κτλ). Μία πιθανή μορφή της οντολογίας απεικονίζεται στην εικόνα 2-6 [BAIL05].



Εικόνα 2-6 Οντολογία για τον καθορισμό του είδους ενός εγγράφου.

2.5.8 Κανόνες

Στο πέμπτο επίπεδο υπάρχει επίσης και το τμήμα που αφορά τους κανόνες. Σύμφωνα με τη μαθηματική λογική αλλά και γενικότερα την επιστήμη της λογικής, δηλωτικός κανόνας (rule of inference) ενός συστήματος είναι ένα σχήμα δημιουργίας έγκυρων δηλώσεων. Τέτοιου είδους σχήματα επιτρέπουν τη δημιουργία συντακτικών σχέσεων μεταξύ ενός συνόλου τύπων (formulas) που ονομάζονται προϋποθέσεις (premises) και ενός ισχυρισμού (assertion) που ονομάζεται συμπέρασμα (conclusion). Αυτές οι συντακτικές σχέσεις χρησιμοποιούνται στην διαδικασία της δήλωσης (process of inference), και με τη βοήθεια τους μπορούν να δημιουργούνται νέοι ισχυρισμοί από αυτούς που είναι ήδη γνωστοί. Η προσθήκη λογικής στο Διαδίκτυο σχετίζεται με τη δημιουργία τέτοιων δηλωτικών κανόνων. Η διαφορά είναι ότι στο σημασιολογικό ιστό υπάρχει μία πρόκληση: η παροχή μίας γλώσσας που θα εκφράζει δεδομένα και κανόνες συλλογισμού που θα εφαρμόζονται στα δεδομένα. Ταυτόχρονα όμως, αυτή η γλώσσα θα πρέπει να επιτρέπει την εξαγωγή κανόνων για χρήση στο σημασιολογικό ιστό, από οποιοδήποτε σύστημα αναπαράστασης γνώσης ([WIRU06], [LEE01]).

Το επίπεδο αυτό εξελίσσεται συνεχώς. Έχει ιδρυθεί μία ομάδα που ασχολείται με την ανάπτυξη της μορφής ανταλλαγής κανόνων (Rule Interchange Format - RIF). Αρχικός σκοπός αυτής της ομάδας είναι η αναδιατύπωση ή η δημιουργία προτύπων που σχετίζονται με υπάρχουσες γλώσσες και συστήματα δημιουργίας κανόνων, ώστε να είναι συμβατά με το σημερινό σημασιολογικό ιστό αλλά και πιθανές μελλοντικές επεκτάσεις του. Περισσότερες πληροφορίες για την ομάδα και τη λειτουργία της υπάρχουν στην ιστοσελίδα <http://www.w3.org/2005/rules/>. Η πιο πρόσφατη δουλειά της ομάδας είναι η δημιουργία των περιπτώσεων χρήσης και των απαιτήσεων που αφορούν τη RIF (<http://www.w3.org/TR/rif-ucr/>) [GINS06]. Τελικός στόχος της ομάδας είναι η δημιουργία μίας γλώσσας κανόνων η οποία θα επιτρέπει τη διερμηνεία και τη μεταφορά κανόνων μεταξύ πολλαπλών γλωσσών και συστημάτων [COVE05].

2.5.9 Αιτήματα

Σύμφωνα με τον Eric Prud'hommeaux [PRUD06], [PRUD07], το τμήμα των αιτημάτων (queries) παρέχει υπηρεσίες στα επίπεδα τέσσερα και πέντε. Τα αιτήματα χρειάζονται ώστε

να είναι δυνατή η εξαγωγή και ο συνδυασμός των πληροφοριών που περιέχουν τα διάφορα έγγραφα που κατανέμονται σε όλο το εύρος του σημασιολογικού ιστού (RDF, RDFS, OWL), παρόμοια με τον τρόπο που υποβάλλονται αιτήματα στις σχεσιακές βάσεις δεδομένων. Μάλιστα έχει κατασκευαστεί μία γλώσσα υποβολής αιτημάτων η οποία έχει πολλές ομοιότητες με την SQL, και ονομάζεται SPARQL. Η SPARQL έχει κατασκευαστεί από την ομάδα εργασίας πρόσβασης δεδομένων RDF (RDF Data Access Group), που ανήκει στις ομάδες εργασίας του W3C που σχετίζονται με τον σημασιολογικό ιστό.

Η SPARQL είναι μία γλώσσα υποβολής αιτημάτων για την εξαγωγή πληροφορίας από γραφήματα RDF. Διευκολύνει την:

- ✓ εξαγωγή πληροφορίας στη μορφή που απεικονίζεται στα γραφήματα,
- ✓ εξαγωγή υπογραφημάτων (subgraphs) RDF,
- ✓ δημιουργία νέων γραφημάτων από τις πληροφορίες που εξάγονται έπειτα από υποβολή αιτημάτων στα υπάρχοντα γραφήματα.

Ως γλώσσα πρόσβασης δεδομένων, η SPARQL είναι κατάλληλη είτε για τοπική (local use) είτε για απομακρυσμένη χρήση (remote use). Η μορφή ενός αιτήματος SPARQL (καθώς και τα αποτελέσματα που παράγει) παρουσιάζονται στο παράρτημα Α (παράδειγμα Α-10). Όταν το αίτημα του παραδείγματος Α-10 υποβάλλεται σε μία οντολογία, εμφανίζει ως αποτέλεσμα όλα τα στοιχεία της που συνοδεύονται από την τιμή `owl:Class` στην ιδιότητα `rdf:type`, δηλαδή όλες τις κλάσεις OWL της οντολογίας. Αναφορά στην OWL και στις κλάσεις της θα γίνουν σε επόμενο κεφάλαιο.

2.5.10 Ενοποιημένη λογική

Σύμφωνα με τους Βλαχάβα Ι, Κεφαλά Π, κ.α. [ΒΛΑΧ06 σελ. 654], το έκτο επίπεδο της στοίβας του σημασιολογικού ιστού αποτελείται από ένα μόνο τμήμα, αυτό της ενοποιημένης λογικής. Αυτό το επίπεδο χρειάζεται ώστε να γίνονται κατανοητές διάφορες ενέργειες όπως:

- κανόνες μετατροπής εγγράφων από έναν τύπο σε κάποιο άλλο,
- έλεγχος ενός εγγράφου ως προς τη συνέπεια του (πχ έλεγχος αν ένα έγγραφο ικανοποιεί όλους τους περιορισμούς που θέτει μία ομάδα οντολογιών),
- αποτέλεσμα μίας ερώτησης με σκοπό την ανάκληση όρων ή την αναγνώριση άγνωστων

όρων.

Το επίπεδο της λογικής παρέχει τη δυνατότητα:

- ✓ αναπαράστασης της κατηγορηματικής λογικής (πχ τελεστές AND, OR, NOT),
- ✓ συσχέτισης των δεδομένων (πχ αν ισχύει το X, τότε ισχύουν όλα τα Y).

Τέλος το λογικό επίπεδο επιτρέπει την επεξεργασία των δεδομένων των κατώτερων επιπέδων με τη χρήση σημασιολογικών σχέσεων υπό μορφή κανόνων.

Ως παράδειγμα εφαρμογής αυτού του επιπέδου μπορεί να αναφερθεί ο αυτόματος μετασχηματισμός αιτημάτων (queries) δύο διαφορετικών Διαδικτυακών βάσεων δεδομένων. Επιπλέον μπορούν να χρησιμοποιηθούν και κανόνες που απεικονίζουν τη δομή της μίας βάσης στην άλλη, σε κατανοητή μορφή.

2.5.11 Αποδείξεις

Σύμφωνα με τους Βλαχάβα Ι, Κεφαλά Π, κ.α. [ΒΛΑΧ06 σελ. 654], το έβδομο επίπεδο περιλαμβάνει το τμήμα των αποδείξεων. Το επίπεδο των αποδείξεων χρησιμοποιεί το επίπεδο της λογικής, για την εξαγωγή συμπερασμάτων.

Αυτό το επίπεδο διαθέτει:

- μηχανισμούς και διαδικασίες εξαγωγής συμπερασμάτων που χρησιμοποιούν τη γνώση του λογικού επιπέδου,
- αναπαράσταση διαδικασιών απόδειξης σε XML ή RDF.

Η χρήση του λογικού επιπέδου από το επίπεδο των αποδείξεων παρέχει σημαντικά πλεονεκτήματα:

- ✓ εύκολη δημιουργία εξηγήσεων σχετικά με τα συμπεράσματα,
- ✓ εμφάνιση των διαδοχικών βημάτων του μηχανισμού εξαγωγής συμπερασμάτων.

Ο χρήστης έχει τη δυνατότητα να ζητήσει παρουσίαση των αποδείξεων με:

- ✓ παράθεση των λογικών προτάσεων και των δεδομένων που χρησιμοποιήθηκαν,

- ✓ χρήση εξελιγμένων τεχνικών παραγωγής εξηγήσεων σε φυσική γλώσσα που έχουν αναπτυχθεί στο πλαίσιο της έρευνας των έμπειρων συστημάτων,
- ✓ οπτικό τρόπο (πχ δέντρα αποδείξεων).

Η παροχή αποδείξεων αυξάνει την εμπιστοσύνη των χρηστών στις παρεχόμενες υπηρεσίες.

2.5.12 Αξιοπιστία σημασιολογικού ιστού

Σύμφωνα με τους Βλαχάβα Ι, Κεφαλά Π, κ.α. [BLAX06 σελ. 654-655], Theo Dimitrakos, Brian Matthews, κ.α. [DIMI01], και Jennifer Golbeck, Bijan Parsia, κ.α. [GOLB03], το όγδοο και υψηλότερο επίπεδο είναι αυτό της αξιοπιστίας. Η αξιοπιστία είναι ένα βασικό χαρακτηριστικό που πρέπει να παρέχει ο σημασιολογικός ιστός. Υλοποιείται κατά ένα μεγάλο βαθμό χρησιμοποιώντας το τμήμα της κρυπτογραφίας (κρυπτογράφηση, ψηφιακές υπογραφές, κτλ), που παρέχει τις υπηρεσίες του στο επίπεδο της αξιοπιστίας. Στον σημασιολογικό ιστό, όλοι οι χρήστες θα πρέπει να υπογράφουν ψηφιακά τα έγγραφα τους, για να υπάρχει μία μορφή εξακρίβωσης ότι ο χρήστης που ζητάει ή παρέχει μία υπηρεσία είναι όντως αυτός που παριστάνει ότι είναι.

Η αναζήτηση πληροφορίας θα πρέπει να περιλαμβάνει την απόδειξη ότι ο χρήστης έχει το δικαίωμα να λάβει απάντηση. Λόγου χάριν για να επιτραπεί σε κάποιον η πρόσβαση σε ένα δικτυακό τόπο πρέπει να παραδώσει έγγραφο (ηλεκτρονικό) το οποίο να αποδεικνύει το λόγο για τον οποίο του επιτράπηκε η πρόσβαση.

Οι ηλεκτρονικές συναλλαγές απαιτούν εχεμύθεια. Τα έγγραφα που διακινούνται θα πρέπει να περιλαμβάνουν επιπλέον πληροφορίες που αφορούν την αξιοπιστία τους. Επίσης, αν τεθεί θέμα, η απάντηση - απόδειξη θα είναι εύκολο να κατασκευαστεί υπολογιστικά.

Αν και η γλώσσα RDF Schema παρέχει υποστήριξη για την κωδικοποίηση πολιτικών ασφαλείας, χρειάζεται να κατασκευαστούν επιπλέον συστατικά (components) για την πλήρη αξιοποίηση του σημασιολογικού ιστού. Τα συστατικά αυτά θα βοηθήσουν:

- ➔ στον καθορισμό πιο σύνθετων μοντέλων δεδομένων και οντολογιών χρησιμοποιώντας μία ισχυρότερη γλώσσα καθορισμού περιορισμών από την RDF Schema,

- στην έκφραση κανόνων και ιδιοτήτων μοντέλων RDF με τη βοήθεια της λογικής (logic), και στη δημιουργία εργαλείων για την κληρονομικότητα αυτών των ιδιοτήτων,
- στη δημιουργία μίας γλώσσας μέσω της οποίας θα υποβάλλονται αιτήματα σε ένα σύνολο δηλώσεων RDF και θα αξιολογούνται με τη βοήθεια της λογικής.

Υπάρχουν ήδη εφαρμογές που υλοποιούν το επίπεδο της αξιοπιστίας (πχ TrustBot, TrustMail, κτλ).

Η αναφορά σε ένα παράδειγμα παροχής αξιοπιστίας θα βοηθήσει στην καλύτερη κατανόηση του ρόλου αυτού του επιπέδου.

Έστω ότι ο πράκτορας ενός ηλεκτρονικού καταστήματος αποστέλλει ένα ηλεκτρονικό τιμολόγιο στον πράκτορα ενός χρήστη με χρέωση ενός αγαθού. Το ηλεκτρονικό έγγραφο πρέπει να συνοδεύεται από:

- αρχείο καταγραφής συναλλαγών του ηλεκτρονικού καταστήματος (επιβεβαίωση ότι έγινε η συναλλαγή),
- ψηφιακή υπογραφή του αγοραστή (επιβεβαίωση ότι ο αγοραστής "μπήκε" στο ηλεκτρονικό κατάστημα),
- ηλεκτρονικό τιμοκατάλογο καταστήματος (επιβεβαίωση της τιμής του αγαθού),
- ηλεκτρονική απόδειξη παραλαβής του αγαθού μέσω εταιρείας χρηματοπιστωτικών (courier) (επιβεβαίωση της παράδοσης του αγαθού).

Ο πράκτορας του χρήστη:

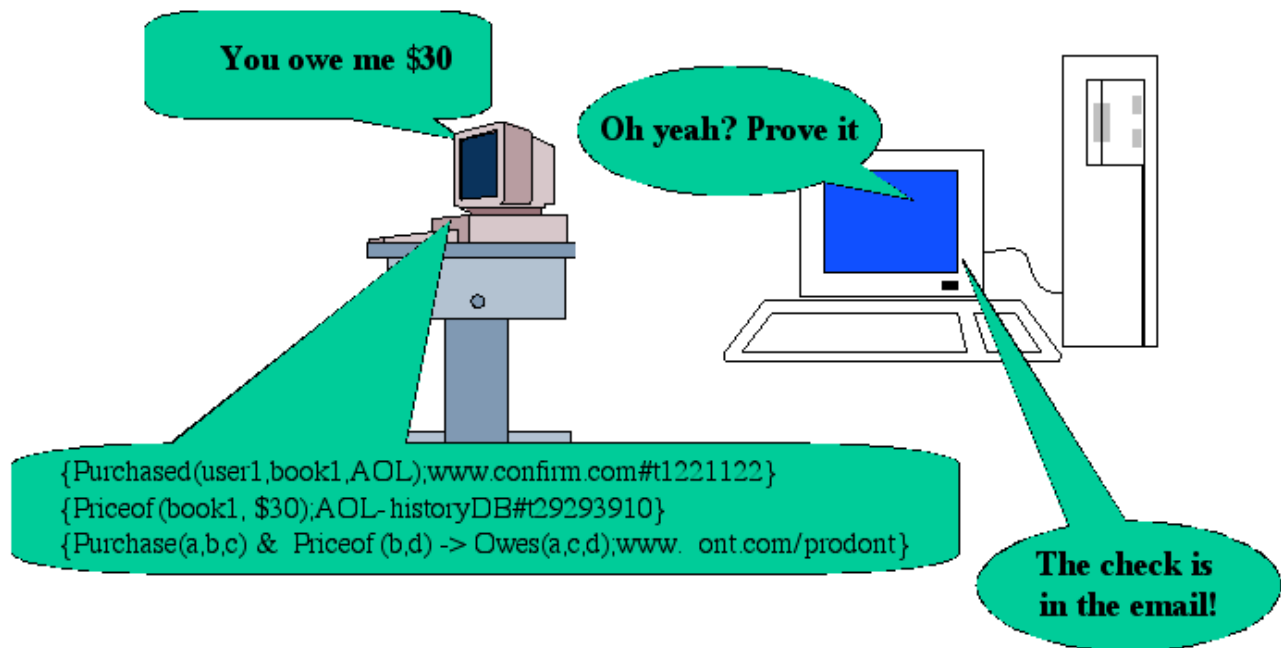
- επιβεβαιώνει την ορθότητα της χρέωσης,
- εξηγεί στο χρήστη γιατί προχώρησε σε χρέωση της πιστωτικής του κάρτας.

Ένα παρόμοιο παράδειγμα με το παραπάνω απεικονίζεται γραφικά στην εικόνα 2-7. Οι πληροφορίες που απεικονίζονται στην εικόνα 2-7 είναι οι εξής:

- ο χρήστης user1 έχει παραγγείλει το βιβλίο book1 από το ηλεκτρονικό κατάστημα της America On Line (AOL) {Purchased(user1, book1, AOL) },
- η τιμή του βιβλίου book1 είναι \$30 {PriceOf(book1, \$30) },
- όταν κάποιος χρήστης a παραγγέλλει ένα προϊόν b από το ηλεκτρονικό, κατάστημα c, και

η τιμή του προϊόντος b είναι d , τότε ο χρήστης a χρωστάει στο κατάστημα c , d χρήματα $\{Purchase(a, b, c) \ \& \ Priceof(b,d) \rightarrow Owes(a, c, d)\}$,

- οι πράκτορες επικοινωνούν μεταξύ τους για την επιβεβαίωση της ορθότητας της χρέωσης και την επεξήγηση των λόγων για τους οποίους έγινε χρέωση της πιστωτικής κάρτας του χρήστη.



Εικόνα 2-7 Παράδειγμα παροχής αξιοπιστίας.

2.5.13 Κρυπτογραφία

Σύμφωνα με τους Tim Berners-Lee, James Hendler, κ.α [LEE01], ένας επίσης ενδιαφέρον τομέας που πρέπει να καλύψει ο σημασιολογικός ιστός είναι αυτός της κρυπτογραφίας. Αυτό το επίπεδο παρέχει τις υπηρεσίες του σε όλα τα υπόλοιπα.

Με τις ψηφιακές υπογραφές, που ουσιαστικά είναι κρυπτογραφημένες ομάδες δεδομένων, Η/Υ και πράκτορες μπορούν να συνεργάζονται ώστε να εξασφαλίζουν ότι οι πληροφορίες που παρέχονται προέρχονται από αξιόπιστη πηγή.

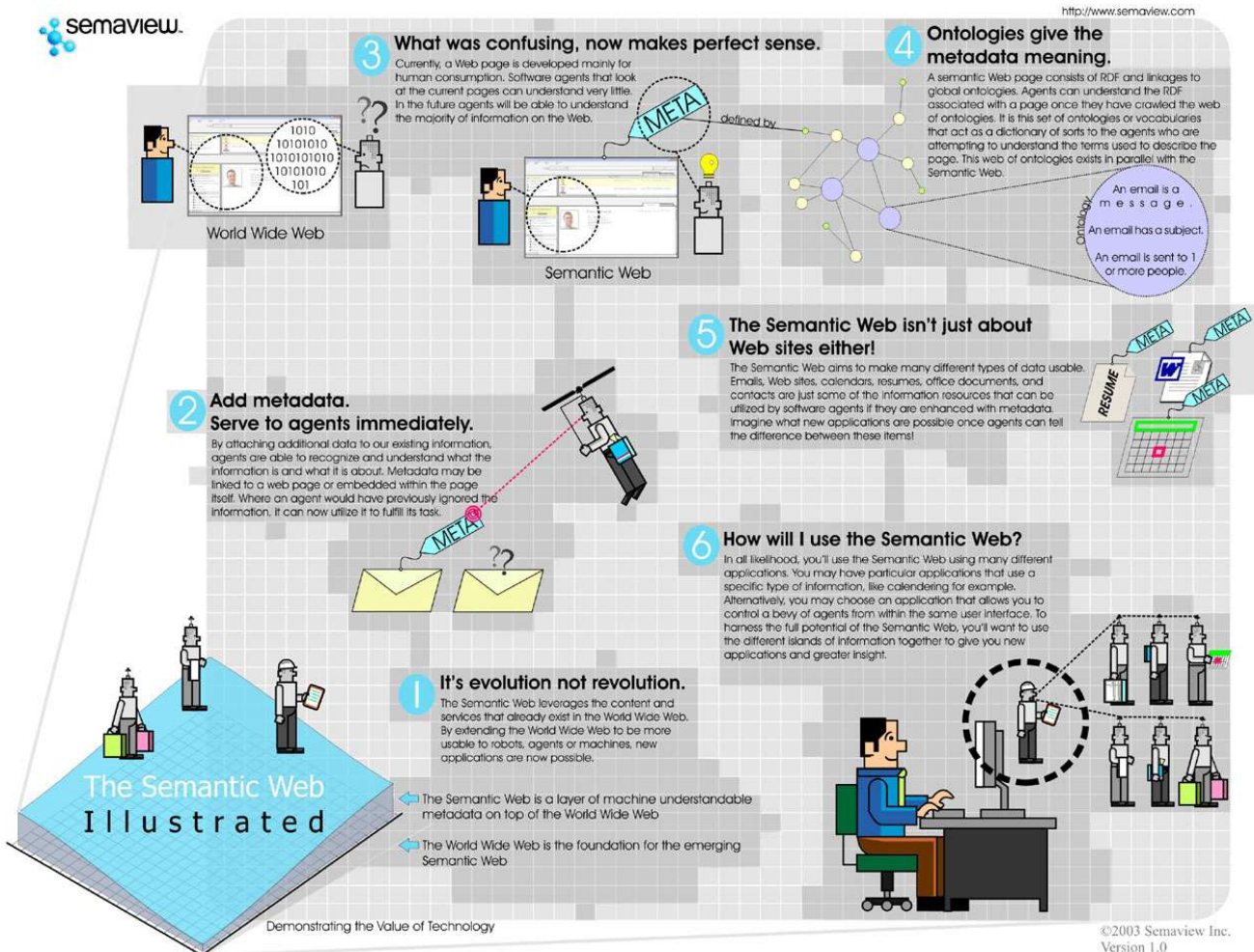
Πριν ένας πράκτορας λάβει μία σημαντική απόφαση όπως η μεταφορά χρημάτων σε ένα λογαριασμό, θα πρέπει να έχει πραγματοποιηθεί αυστηρός έλεγχος της αξιοπιστίας της πηγής που ζητάει να γίνει κάτι τέτοιο.

Η κρυπτογραφία χρειάζεται επιπλέον και για άλλες λειτουργίες, όπως η ασφαλής επικοινωνία, το ηλεκτρονικό εμπόριο, κτλ.

2.5.14 Γραφική παρουσίαση του σημασιολογικού ιστού

Μία γραφική παρουσίαση της δομής και της λειτουργίας του σημασιολογικού ιστού υπάρχει στην εικόνα 2-8 [LEE01]. Η εικόνα συνοπτικά αναφέρει τα εξής:

1. Ο σημασιολογικός ιστός είναι εξέλιξη του παγκοσμίου ιστού πληροφοριών και όχι επανάσταση.
2. Η προσθήκη μεταδεδομένων δίνει τη δυνατότητα στους πράκτορες να κατανοήσουν καλύτερα την πληροφορία και να αναγνωρίσουν με τι σχετίζεται.
3. Ενώ στο παρελθόν, το περιεχόμενο ήταν κατανοητό μόνο από τους ανθρώπους, πλέον είναι κατανοητό και από τις μηχανές με χρήση των πρακτόρων.
4. Οι οντολογίες (στις οποίες θα γίνει αναλυτική αναφορά στο επόμενο κεφάλαιο), ως λεξιλόγια με τα οποία δεσμεύονται οι πράκτορες, δίνουν το κατάλληλο νόημα στα μεταδεδομένα, ώστε να αποφεύγονται τυχόν συγχύσεις στην επικοινωνία και ανταλλαγή πληροφοριών μεταξύ των πρακτόρων.
5. Ο σημασιολογικός ιστός δεν σχετίζεται μόνο με ιστοσελίδες, αλλά με μία πληθώρα εφαρμογών όπως ηλεκτρονικά ημερολόγια, ηλεκτρονικά έγγραφα, ηλεκτρονικές ατζέντες, ηλεκτρονικές αλληλογραφίες, κτλ.
6. Η χρήση του σημασιολογικού ιστού θα γίνεται κυρίως με χρήση των πρακτόρων, οι οποίοι θα αναλαμβάνουν τη διεκπεραίωση σύνθετων εργασιών, αλλά και μέσω εφαρμογών συγκεκριμένου τύπου, πχ ηλεκτρονικών ημερολογίων.



Εικόνα 2-8 Ο σημασιολογικός ιστός.

2.6 Τα κέρδη που θα επιφέρει η σημασιολογία

Με τη βοήθεια του σημασιολογικού ιστού το Διαδίκτυο θα πάψει να είναι χαώδες κατά ένα μεγάλο βαθμό. Σε συνεργασία με τους πράκτορες, οι οποίοι θα είναι εύκολο να χρησιμοποιηθούν από οποιονδήποτε χρήστη δίχως να απαιτείται προηγούμενη εμπειρία,

διευκολύνονται και κυρίως βελτιώνονται διαδικασίες που σχετίζονται με την πληροφορία, όπως [ΒΛΑΧ06 σελ. 638-639, 647-649]:

- ✓ η αναζήτηση,
- ✓ η αυτοματοποίηση,
- ✓ η ομαδοποίηση,
- ✓ η επαναχρησιμοποίηση της από τελείως διαφορετικές διαδικτυακές εφαρμογές.

Πρέπει να γίνει κατανοητό ότι ο σημασιολογικός ιστός δεν βασίζεται στην κωδικοποίηση ιστοσελίδων με σημασιολογία (αν και παρέχει υποστήριξη για κάτι τέτοιο). Βασίζεται κυρίως στην εξαγωγή δεδομένων τα οποία είναι αναγνώσιμα από τις μηχανές - Η/Υ (machine readable) σε πολύ μεγαλύτερη κλίμακα από τη σημερινή. Αυτό σημαίνει ότι για την ανάπτυξη του σημασιολογικού ιστού θα πρέπει ολοένα και περισσότερα εργαλεία δημιουργίας εγγράφων (authoring tools), τεχνολογίες βάσεων δεδομένων, εφαρμογές μεγάλης κλίμακας (enterprise level applications), κτλ να χρησιμοποιούν τα πρότυπα του σημασιολογικού ιστού που έχει ορίσει το W3C για τη δημιουργία περιγραφών (RDF) και οντολογιών (OWL) [LEE05].

3. Οντολογίες

Οι οντολογίες σε θεωρητικό υπόβαθρο αποτελούν προδιαγραφές εννοιών (conceptualizations). Πρακτικά, οι οντολογίες του σημασιολογικού ιστού καθορίζουν λεξιλόγια με τα οποία δεσμεύονται οι πράκτορες ώστε η μεταξύ τους συνεργασία να είναι ευκολότερη και αποδοτικότερη.

Το κεφάλαιο αυτό αρχικά παρουσιάζει στον αναγνώστη έναν ορισμό της λέξης “οντολογία” στα πλαίσια του σημασιολογικού ιστού. Έπειτα αναφέρονται οι ανάγκες δημιουργίας οντολογιών και αναλύονται τα χαρακτηριστικά τους. Τέλος, παρουσιάζονται κάποιες ενδεικτικές εμπορικές και μη εφαρμογές που ορίζουν/χρησιμοποιούν οντολογίες. Στόχος του κεφαλαίου είναι να κατανοήσει ο αναγνώστης τους λόγους δημιουργίας των οντολογιών, και να μπορεί να αναφέρει τα βασικά χαρακτηριστικά και τις κατηγορίες τους.

3.1 Ορισμός

Η λέξη οντολογία χρησιμοποιείται σε πολλά επιστημονικά πεδία με διαφορετική σημασία και για το λόγο αυτό έχει δημιουργήσει μεγάλη σύγχυση σε πολλές συζητήσεις που σχετίζονται με τον τομέα της τεχνητής νοημοσύνης, της φιλοσοφίας, της γνώσης, κτλ [GRUB93].

Οι οντολογίες στον σημασιολογικό ιστό, χρειάζονται ώστε να μπορούν να επικοινωνούν και να ανταλλάσσουν (διαμοιράζονται) πληροφορίες μεταξύ τους οι πράκτορες. Άρα ενδιαφέρει ο καθορισμός της λέξης οντολογίας από τη σκοπιά του τομέα που ασχολείται με το διαμοιρασμό γνώσης.

Στον διαμοιρασμό γνώσης (knowledge sharing), οντολογία είναι η προδιαγραφή μίας έννοιας. Συγκεκριμένα, οντολογία είναι μία περιγραφή των εννοιών και των σχέσεων που μπορούν να γνωστοποιηθούν σε ένα πράκτορα ή σε μία κοινότητα πρακτόρων.

Στον σημασιολογικό ιστό, ο πιο κοινός τύπος οντολογίας είναι μία ταξινόμηση (taxonomy) και ένα σύνολο συμπερασματικών κανόνων. Η ταξινόμηση καθορίζει κλάσεις αντικειμένων και σχέσεις μεταξύ τους. Για παράδειγμα, ένας παπαγάλος είδους *Forpus Coelestis* μπορεί να καθοριστεί να είναι τύπου *παπαγάλος*, ένας *παπαγάλος* να καθοριστεί να είναι τύπου *πουλί*, η ικανότητα των *πουλιών* να μιλάνε να καθοριστεί να ισχύει μόνο για ορισμένους τύπους *παπαγάλων*, κτλ. Κλάσεις, υποκλάσεις και σχέσεις ανάμεσα τους είναι ένα ιδιαίτερα ισχυρό εργαλείο που χαρακτηρίζεται πολύτιμο για το σημασιολογικό ιστό [LEE01].

3.2 Ανάγκες δημιουργίας

Αυτό που έχει τελικά μεγάλη σημασία δεν είναι ο ορισμός της οντολογίας, αλλά οι λόγοι για τους οποίους κατασκευάστηκε. Τα προβλήματα που οδήγησαν στη δημιουργία των οντολογιών είναι [LEE01]:

- × η δυσκολία διαμοιρασμού και επαναχρησιμοποίησης της γνώσης,
- × το μεγαλύτερο ποσοστό των πληροφοριών των οργανισμών είναι κωδικοποιημένο σε μορφή κειμένου,
- × το μη δομημένο ή ήμι-δομημένο κείμενο το οποίο είναι δύσκολο να ανακτηθεί, να ερμηνευτεί και να αναλυθεί,
- × οι ανολοκλήρωτες περιγραφές,
- × οι άσκοπα επαναλαμβανόμενες τιμές,
- × τα ασύμβατα έγγραφα,
- × οι διαφορετικές ορολογίες / γλώσσες / συντομογραφίες.

Οι οντολογίες σχεδιάζονται έτσι ώστε να είναι δυνατός ο διαμοιρασμός γνώσης μεταξύ των πρακτόρων.

3.3 Οι οντολογίες ως μηχανισμός προδιαγραφών

Σύμφωνα με τον Tom Gruber [GRUB93], οι οντολογίες γράφονται ως ένα σύνολο ορισμών

που ανήκουν σε ένα λεξιλόγιο (vocabulary). Αν και αυτός δεν είναι ο μοναδικός τρόπος δημιουργίας οντολογιών, είναι πολύ βολικός για τη δημιουργία οντολογικών δεσμεύσεων (ontological commitments).

Πρακτικά, μία οντολογική δέσμευση είναι μία συμφωνία για τη χρήση ενός λεξιλογίου (πχ για την υποβολή αιτημάτων και τη δημιουργία δηλώσεων) με τρόπο συνεπή αλλά όχι πλήρη, δίχως σε καμία περίπτωση αυτή η χρήση να παραβαίνει τη θεωρία που έχει καθορίσει η οντολογία. Τελικός σκοπός είναι η δημιουργία πρακτόρων που δεσμεύονται με οντολογίες, έτσι ώστε να είναι εύκολη η μεταξύ τους επικοινωνία και ο διαμοιρασμός γνώσης.

Όταν η γνώση ενός χώρου/περιοχής (domain) αντιπροσωπεύεται με ένα δηλωτικό τρόπο, το σύνολο των αντικειμένων τα οποία μπορούν να αναπαρασταθούν ονομάζεται υπερσύνολο αναφοράς (universe of discourse). Αυτό το σύνολο αντικειμένων, καθώς και οι σχέσεις μεταξύ τους, αντιπροσωπεύονται από ένα λεξιλόγιο με τη βοήθεια του οποίου αναπαριστά τη γνώση μία εφαρμογή (πράκτορας).

Η οντολογία ενός προγράμματος μπορεί να περιγραφεί καθορίζοντας ένα σύνολο αντιπροσωπευτικών όρων. Σε μια τέτοιου είδους οντολογία, οι όροι συνδέουν τα ονόματα των οντοτήτων στο υπερσύνολο αναφοράς (πχ κλάσεις, σχέσεις, συναρτήσεις, ή άλλα αντικείμενα) με το αναγνώσιμο από τον άνθρωπο κείμενο (human readable) που περιγράφει τι σημαίνουν αυτά τα ονόματα, και επίσημα αξιώματα (formal axioms) που περιορίζουν την ερμηνεία και την καλοσχηματισμένη (well-formed) χρήση αυτών των όρων.

Οι οντολογίες χρησιμοποιούνται για την περιγραφή οντολογικών δεσμεύσεων για ένα σύνολο πρακτόρων, έτσι ώστε οι πράκτορες να μπορούν να επικοινωνήσουν για να ανταλλάξουν πληροφορίες σχετικά με ένα υπερσύνολο αναφοράς, δίχως να χρειάζεται να λειτουργούν με μία κοινά διαμοιρασμένη θεωρία (global shared theory).

Πρακτικά, μία συνήθης οντολογία καθορίζει ένα λεξιλόγιο μέσω του οποίου υποβάλλονται αιτήματα και ανταλλάσσονται δηλώσεις μεταξύ των πρακτόρων. Οι οντολογικές δεσμεύσεις είναι συμφωνίες για κοινή χρήση αυτού του λεξιλογίου με ξεκάθαρο και συνεπή τρόπο. Οι πράκτορες που μοιράζονται ένα λεξιλόγιο δε χρειάζεται να μοιράζονται και μία κοινή βάση

γνώσης. Ο καθένας γνωρίζει πληροφορίες που δεν γνωρίζει ο άλλος. Ένας πράκτορας που δεσμεύεται με μία οντολογία δεν απαιτείται να είναι ικανός να απαντήσει σε όλα τα πιθανά αιτήματα που μπορούν να υποβληθούν στο κοινό λεξιλόγιο.

Οι οντολογίες συχνά συγχέονται με ταξινομημένες ιεραρχίες κλάσεων. Αν και μπορούν να αναπαρασταθούν με αυτό τον τρόπο, δεν είναι απαραίτητο να περιορίζονται σε κάτι τέτοιο. Οι οντολογίες δεν περιορίζονται σε προσδιορισμούς της παραδοσιακής λογικής (traditional logic sense), οι οποίοι το μόνο που κάνουν είναι εισάγουν ορολογία και να μη προσθέτουν καθόλου γνώση σχετικά με τον κόσμο. Για να καθοριστεί μία έννοια αυτό που χρειάζεται είναι ο ορισμός αξιωμάτων που περιορίζουν τις πιθανές ερμηνείες των όρων που έχουν δηλωθεί.

3.4 Επιπλέον προβλήματα που λύνουν οι οντολογίες

Σύμφωνα με τους Tim Berners-Lee, James Hendler, κ.α. [LEE01], όταν οι εφαρμογές χρησιμοποιούν οντολογίες, προβλήματα όπως αυτά της ορολογίας αρχίζουν και εξαλείφονται.

Το νόημα των όρων που χρησιμοποιούνται σε μία ιστοσελίδα μπορεί να καθοριστεί χρησιμοποιώντας δείκτες σε μία οντολογία. Βέβαια τα προβλήματα θα συνεχίσουν να υπάρχουν εάν η οντολογία που αναφέρονται οι δείκτες ερμηνεύει τους όρους που χρησιμοποιούνται στην ιστοσελίδα με διαφορετικό τρόπο από αυτόν που επιθυμεί ο δημιουργός της. Αυτή η σύγχυση μπορεί να αποφευχθεί εάν οι οντολογίες (ή κάποιες άλλες υπηρεσίες του ιστού) παρέχουν σχέσεις ισοδυναμίας (πχ αν υπάρχουν δύο πεδία σε μία φόρμα, ένα με όνομα `postal code` και ένα άλλο με όνομα `zip code`, να γίνεται κατανοητό ότι και οι δύο τιμές αφορούν το ίδιο αντικείμενο, δηλαδή τον ταχυδρομικό κώδικα μιας περιοχής).

Οι οντολογίες μπορούν να ενισχύσουν τη λειτουργία του σημασιολογικού ιστού σε πολλούς τομείς, και ένας από αυτούς είναι ο τομέας της αναζήτησης. Η εφαρμογή (μηχανή) αναζήτησης μπορεί να εκμεταλλευτεί τις οντολογίες ώστε να αναζητά περιεχόμενο μόνο εντός

των σελίδων που αναφέρονται σε ακριβείς έννοιες, αποκλείοντας όσες αναφέρονται σε διφορούμενες λέξεις κλειδιά.

3.5 Τα χαρακτηριστικά μίας οντολογίας

Όλες οι οντολογίες, ανεξαρτήτως της τεχνολογίας η οποία τις περιγράφει, περιέχουν ορισμένες ομοιότητες, δηλαδή κοινά χαρακτηριστικά. Παρακάτω θα γίνει αναφορά σε αυτά τα χαρακτηριστικά, για να ξεκαθαριστεί ακόμη περισσότερο η μορφή μίας οντολογίας.

3.5.1 Άτομα (individuals)

Σύμφωνα με την Wikipedia [WION07], και τον Leo Obrst [LEO06], τα άτομα ή στιγμιότυπα (instances) είναι τα βασικά συστατικά “χαμηλού επιπέδου” μίας οντολογίας. Τα άτομα μίας οντολογίας μπορεί να είναι διακριτά αντικείμενα όπως άνθρωποι, ζώα, οχήματα, πλανήτες, καθώς επίσης και αφηρημένα στοιχεία όπως λέξεις ή αριθμοί.

Για την ακρίβεια, μία οντολογία δεν είναι απαραίτητο να περιέχει άτομα, αλλά ένας από τους βασικούς σκοπούς της είναι να παρέχει ένα τρόπο ταξινόμησης των ατόμων, ακόμη κι αν αυτά δεν αποτελούν τμήμα της.

3.5.2 Κλάσεις (έννοιες)

Οι κλάσεις³ ή έννοιες είναι αφηρημένες ομάδες, σύνολα ή συλλογές αντικειμένων. Μπορούν να εμπεριέχουν άτομα, άλλες κλάσεις ή και συνδυασμό των δύο. Μερικά παραδείγματα κλάσεων είναι:

- αριθμός, αναπαριστά όλους τους αριθμούς,
- αυτοκίνητο, αναπαριστά όλα τα αυτοκίνητα,
- αετός, αναπαριστά όλους τους αετούς,

³ Οι κλάσεις των οντολογιών έχουν σημαντικές ομοιότητες αλλά και διαφορές με τις κλάσεις του αντικειμενοστραφούς προγραμματισμού. Περισσότερες λεπτομέρειες υπάρχουν στο [KNUB06].

- ψάρι, αναπαριστά όλα τα ψάρια,
- κλάση, αναπαριστά όλες τις κλάσεις.

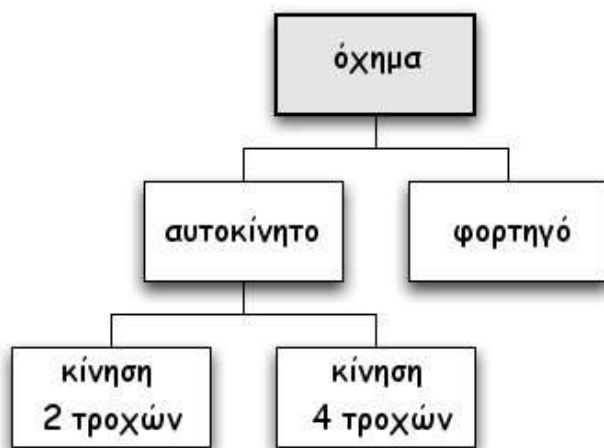
Μία κλάση μπορεί να εμπεριέχει άλλες κλάσεις, μπορεί να ανήκει στον εαυτό της ή μπορεί να είναι μία ολότητα - δηλαδή να περιέχει τα πάντα, κτλ. Τέτοιου είδους διαφορές εντοπίζονται στα διάφορα ήδη οντολογιών. Πολλές φορές χρειάζεται να τεθούν περιορισμοί στις παραπάνω περιπτώσεις, για να αποφευχθεί η δημιουργία ορισμένων ευρέως γνωστών παραδόξων, όπως για παράδειγμα το να δηλωθεί σε μία ιεραρχία που περιγράφει ανθρώπους, ότι ένας συγκεκριμένος άνθρωπος είναι ταυτόχρονα και άντρας και γυναίκα.

Οι κλάσεις μίας οντολογίας μπορεί να είναι υποθετικές (intensional) ή συμπερασματικές (extensional). Μία υποθετική κλάση είναι μία περιγραφή με ιδιότητες, από την οποία μπορούν να περιγραφούν πολλά στιγμιότυπα. Για παράδειγμα η κλάση άνθρωπος μπορεί να θεωρηθεί ως μία περιγραφή (με ιδιότητες όπως το ονοματεπώνυμο, η ημερομηνία γέννησης, το χρώμα ματιών, κτλ), από την οποία μπορούν να προκύψουν πολλά στιγμιότυπα, για παράδειγμα ο “Στέλιος Παπαδόπουλος”, η “Μαρία Καραγιάννη”, κτλ. Για αυτό το λόγο το σχήμα δημιουργίας μίας σχεσιακής βάσης δεδομένων (εντολές SQL) ονομάζεται υποθετική βάση δεδομένων (intensional database), ενώ οι εγγραφές των δεδομένων ονομάζονται συμπερασματική βάση δεδομένων (extensional database): Από το ίδιο σχήμα, μπορούν να προκύψουν διαφορετικά αποτελέσματα (εγγραφές). Η κλάση άνθρωπος μπορεί όμως να αντιμετωπιστεί και ως συμπερασματική, όπου σε αντιστοιχία με τη μαθηματική θεωρία των συνόλων, αποτελεί το στοιχείο ενός συνόλου, πχ του $\text{άνθρωπος} = \{\text{Γιάννης}, \text{Μαρία}, \text{Πέτρος}\}$.

Η διάκριση μεταξύ των δύο ειδών κλάσεων είναι ιδιαίτερα λεπτή. Η υπόθεση (intension) σχετίζεται με κάποια ιδιότητα, ενώ το συμπέρασμα (extension) σχετίζεται με τα αντικείμενα που έχουν αυτή την ιδιότητα εντός ενός προκαθορισμένου πλαισίου (context). Ένα καλό παράδειγμα αποτελεί η ιδιότητα “κόκκινο”. Πολλά μπορούν να ειπωθούν για τη συγκεκριμένη ιδιότητα, για παράδειγμα ότι πρόκειται για ένα φάσμα συχνοτήτων συγκεκριμένου μήκους κύματος, ότι είναι ορατή στο ανθρώπινο μάτι, κτλ. Αυτό είναι ένα παράδειγμα υπόθεσης. Αντίθετα, μπορεί να γίνει αναφορά σε συγκεκριμένα αντικείμενα τα οποία έχουν τη συγκεκριμένη ιδιότητα (κόκκινο χρώμα), πχ μία κόκκινη μπάλα, ένα κόκκινο τριαντάφυλλο, κτλ. Αυτό είναι ένα παράδειγμα συμπεράσματος.

Αν και οι συμπερασματικές κλάσεις παρουσιάζουν καλύτερη συμπεριφορά και είναι περισσότερο κατανοητές από μαθηματικής απόψεως, δεν επιτρέπουν τη δημιουργία λεπτών διακρίσεων που συνήθως κρίνονται απαραίτητες για την κατασκευή οντολογιών. Για παράδειγμα, μία οντολογία ίσως χρειαστεί να υλοποιήσει σαφή διαχωρισμό μεταξύ μίας κλάσης όλων των πλασμάτων τα οποία έχουν νεφρά, και μίας κλάσης όλων των πλασμάτων με καρδιά, ακόμη κι αν τυχαίνει οι δύο αυτές κλάσεις να έχουν ακριβώς τα ίδια μέλη.

Ένα επίσης σημαντικό χαρακτηριστικό, είναι ότι μία κλάση μπορεί να εντάσσει (subsume) ή να εντάσσεται σε άλλες κλάσεις. Για παράδειγμα, στην εικόνα 3-1, η κλάση όχημα εντάσσει την κλάση αυτοκίνητο, εφόσον οποιοδήποτε αυτοκίνητο εντάσσεται και στην ευρύτερη κατηγορία των οχημάτων. Η σχέση της ένταξης χρησιμοποιείται ώστε να δημιουργηθεί μία ιεραρχία κλάσεων, με την γενικότερη να βρίσκεται στην κορυφή (πχ αντικείμενο), και τις ειδικότερες στα κάτω σημεία (πχ *bmw 316 compact*).



Εικόνα 3-1 Μερικό διάγραμμα ιεραρχίας κλάσεων οντολογίας.

Μία διανομή (partition) είναι ένα σύνολο συσχετιζόμενων κλάσεων και κανόνων, που επιτρέπουν στα αντικείμενα να τοποθετούνται στην κατάλληλη κλάση. Για παράδειγμα, στην εικόνα 3-1, υπάρχει μία διανομή για την κλάση αυτοκίνητο που την χωρίζει σε δύο υποκλάσεις (ένα επίπεδο κατώτερες στην ιεραρχία): η μία αναπαριστά όλα τα αυτοκίνητα που υποστηρίζουν κίνηση δύο τροχών, και η άλλη όλα όσα υποστηρίζουν κίνηση τεσσάρων τροχών.

Εάν η διανομή εγγυάται ότι ένα αυτοκίνητο δεν μπορεί να υπάρχει και στις δύο υποκλάσεις, τότε ονομάζεται διαχωρίσιμη (disjoint). Εάν οι κανόνες της διανομής εξασφαλίζουν ότι κάθε διακριτό αντικείμενο της υπερκλάσης (της ανώτερης κατά ένα επίπεδο στην ιεραρχία) είναι στιγμιότυπο τουλάχιστον μίας εκ των υποκλάσεων, η διανομή ονομάζεται εξαντλητική (exhaustive).

3.5.3 Ιδιότητες οντολογίας

Τα αντικείμενα μίας οντολογίας μπορούν να περιγραφούν με ανάθεση ιδιοτήτων. Κάθε ιδιότητα έχει τουλάχιστον όνομα και τιμή, και χρησιμοποιείται ώστε να καταχωρήσει πληροφορία που αφορά το αντικείμενο στο οποίο αναφέρεται. Για παράδειγμα, το αντικείμενο `bmw 316 compact` έχει ιδιότητες όπως [WION07]:

- όνομα: `bmw 316 compact`,
- αριθμός θυρών: `2`,
- μηχανή: `{ 4.0 λίτρα, 4.5 λίτρα }`,
- κίνηση: `2 τροχών`.

Η τιμή μίας ιδιότητας μπορεί να είναι ένας σύνθετος τύπος δεδομένων. Στο παράδειγμα, η τιμή της ιδιότητας μηχανή είναι μία λίστα τιμών, ενώ η τιμή κίνηση είναι ένα στιγμιότυπο της κλάσης κίνηση 2 τροχών.

3.5.4 Σχέσεις

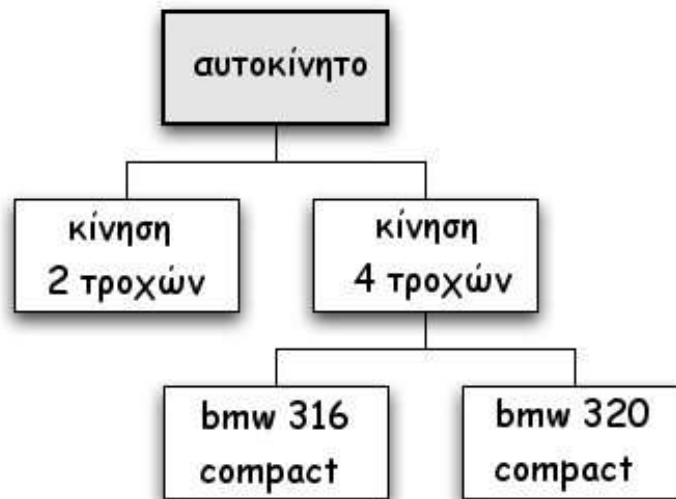
Σύμφωνα με τη Wikipedia [WION07], οι ιδιότητες είναι χρήσιμες για την περιγραφή των σχέσεων (relationships ή relations) μεταξύ των αντικειμένων μίας οντολογίας. Μία σχέση είναι μία ιδιότητα η τιμή της οποίας είναι ένα άλλο αντικείμενο της οντολογίας. Στην εικόνα 3-2, μία πιθανή ιδιότητα του αντικειμένου `bmw 316 compact` είναι η:

- διάδοχος: `bmw 320 compact`.

Αυτή η πληροφορία γνωστοποιεί ότι το `bmw 320 compact` είναι το μοντέλο που διαδέχθηκε το μοντέλο `bmw 316 compact`. Οι οντολογίες αντλούν μεγάλη δύναμη από την ικανότητα τους

για περιγραφή τέτοιου είδους σχέσεων. Το σύνολο όλων των σχέσεων μίας οντολογίας περιγράφουν τη σημασιολογία της.

Η σημαντικότερη σχέση μίας οντολογίας είναι η ένταξη (γνωστή και ως *είναι-ένα*, *είναι-υποτύπος*, *είναι-υποκλάση*). Αυτή η σχέση καθορίζει ποια αντικείμενα είναι μέλη σε μία ιεραρχία κλάσεων.



Εικόνα 3-2 Ιεραρχία κλάσεων οντολογίας αυτοκινήτου.

Παραδείγματος χάριν, στην εικόνα 3-2 φαίνεται ότι το bmw 316 compact είναι-ένα όχημα με κίνηση τεσσάρων τροχών, και με τη σειρά του ένα όχημα με κίνηση τεσσάρων τροχών είναι-ένα αυτοκίνητο, το οποίο με τη σειρά του είναι-ένα όχημα, κτλ.

Η χρήση της σχέσης είναι-ένα δημιουργεί μία ιεραρχική ταξινόμηση (hierarchical taxonomy)· μία δενδρική δομή που απεικονίζει με σαφήνεια πως σχετίζονται μεταξύ τους τα αντικείμενα. Σε μία τέτοιου είδους ιεραρχία, κάθε αντικείμενο είναι το “παιδί” μίας “γονικής” κλάσης (εκτός φυσικά από τη ρίζα ή τις ρίζες του δέντρου). Μία άλλη σχέση είναι η σχέση μέρους (meronymy [WIME06]) που γράφεται ως *είναι-μέρος*. Η σχέση μέρους αναπαριστά πώς αντικείμενα συνδυάζονται μεταξύ τους ώστε να διαμορφώσουν σύνθετα (composite) αντικείμενα. Λόγου χάριν, αν επεκταθεί το παράδειγμα της εικόνας 3-2 ώστε η οντολογία να περιέχει αντικείμενα όπως το αντικλεπτικό σύστημα bmw, είναι γνωστό ότι “το αντικλεπτικό σύστημα bmw *είναι-μέρος* του bmw 316 compact”.

Σε περίπτωση που εισαχθούν σχέσεις του τύπου *είναι-μέρος* σε μία οντολογία, παρατηρείται ότι η απλή και κομψή δένδρική δομή μετατρέπεται σε σύνθετη και επομένως είναι δυσκολότερο να ερμηνευτεί. Η εξήγηση του λόγου για τον οποίο εμφανίζεται αυτό το πρόβλημα δεν είναι δύσκολη. Μία οντότητα που *είναι-μέρος* μίας οντότητας, μπορεί να *είναι-μέρος* και κάποιας άλλης γενικότερης οντότητας, κτλ. Η δομή που προκύπτει σε αυτή την περίπτωση ονομάζεται κατευθυνόμενο μη κυκλικό γράφημα (directed acyclic graph - dag) [WIDA06]. Σχέσεις όπως οι *είναι-μέρος* πρέπει να χρησιμοποιούνται με ιδιαίτερη προσοχή, για να αποφεύγονται προβλήματα όπως το dag.

Πέρα από τις πρότυπες σχέσεις του τύπου *είναι-ένα* και *είναι-μέρος*, νέες σχέσεις μπορούν να προκύψουν στα πλαίσια επίλυσης ενός προβλήματος. Για παράδειγμα, στον τομέα των αυτοκινήτων θα μπορούσε να καθοριστεί η σχέση προέρχεται-από, η οποία περιγράφει τη χώρα προέλευσης ενός αυτοκινήτου. Έτσι, για το παράδειγμα της εικόνας 3-2 μπορεί να αναφερθεί ότι το bmw 316 compact προέρχεται-από το Βερολίνο, το οποίο *είναι-μία* από τις πόλεις της Γερμανίας. Το λογισμικό το οποίο θα χρησιμοποιούσε αυτή την οντολογία θα ήταν ικανό να απαντήσει σε ερωτήσεις του τύπου "ποια αυτοκίνητα *προέρχονται-από* τη Γερμανία;".

3.5.5 Κατηγορίες οντολογιών

Σύμφωνα με τη Wikipedia [WION07], υπάρχουν δύο κατηγορίες οντολογιών, οι οντολογίες περιοχής (domain ή domain-specific) και οι θεμελιώδεις (upper ή foundation) οντολογίες. Η κατηγοριοποίηση των οντολογιών γίνεται έχοντας ως βάση τη μοντελοποίηση που χρησιμοποιούν.

Μία οντολογία περιοχής μοντελοποιεί μία συγκεκριμένη περιοχή, δηλαδή ένα συγκεκριμένο τμήμα του κόσμου. Αναπαριστά τη σημασία των όρων που είναι συγκεκριμένη και αφορά μία περιοχή. Λόγου χάριν, η λέξη κάρτα έχει πολλές διαφορετικές σημασίες. Μία οντολογία που αφορά την περιοχή του παιχνιδιού roker, θα μοντελοποιήσει τη σημασία της "κάρτας παιχνιδιού", ενώ μία οντολογία που αφορά το υλικό των Η/Υ, θα μοντελοποιήσει τη σημασία της "διάτρητης κάρτας" ή της "κάρτας οθόνης" (πιο γνωστή και ως κάρτα γραφικών).

Μία θεμελιώδης οντολογία είναι ένα μοντέλο αντικειμένων τα οποία εφαρμόζονται σε μία μεγάλη γκάμα οντολογιών περιοχής. Περιέχει ένα βασικό γλωσσάριο με τους όρους τους οποίους μπορούν να περιγραφούν αντικείμενα που ανήκουν σε ένα σύνολο περιοχών.

Εφόσον οι οντολογίες περιοχής αναπαριστούν έννοιες με πολύ συγκεκριμένους τρόπους, πολλές φορές χαρακτηρίζονται ως ασύμβατες. Καθώς τα συστήματα που βασίζονται σε οντολογίες περιοχής αυξάνουν, χρειάζεται συχνά να γίνεται συγχώνευση πολλών οντολογιών περιοχής, για τη δημιουργία μίας πιο γενικευμένης αναπαράστασης. Αυτό γεννά μία νέα πρόκληση για τους μηχανικούς οντολογιών. Υπάρχει επίσης το ενδεχόμενο να προκύψουν διαφορετικές οντολογίες εντός της ίδιας περιοχής. Αυτό οφείλεται είτε στο διαφορετικό πολιτιστικό, μορφωτικό και ιδεολογικό υπόβαθρο, είτε επειδή επιλέχθηκε διαφορετική γλώσσα αναπαράστασης.

Προς το παρόν, η συγχώνευση οντολογιών είναι μία μεγάλη χειρωνακτική διαδικασία, η οποία είναι χρονοβόρα, εξαντλητική και ακριβή. Η χρήση μίας θεμελιώδους οντολογίας για την παροχή κοινών ορισμών των βασικών όρων μπορεί να διευκολύνει αρκετά αυτή τη διαδικασία. Έχουν γίνει μελέτες σχετικά με την εύρεση τεχνικών για τη συγχώνευση οντολογιών, αλλά αυτός ο ερευνητικός τομέας βρίσκεται ακόμη σε θεωρητικό επίπεδο.

3.6 Τεχνολογίες δημιουργίας οντολογιών

Υπάρχουν πολλές τεχνολογίες που έχουν κατασκευαστεί και προταθεί στο παρελθόν για τη δημιουργία οντολογιών στο σημασιολογικό ιστό⁴, αλλά σύμφωνα με τα μέλη του W3C η καταλληλότερη είναι η OWL, επειδή κατασκευάστηκε ώστε να εξαλείψει τις αδυναμίες των παλαιότερων προτεινόμενων τεχνολογιών (πχ XOL, DAML+OIL⁵, RDF, RDFS) [CORC00].

Οι αδυναμίες που παρουσιάζει καθεμία από τις παλαιότερες προτεινόμενες τεχνολογίες ποικίλουν. Παρακάτω θα περιγραφούν ορισμένες από αυτές.

4 Υπάρχουν και παραδοσιακές τεχνολογίες δημιουργίας οντολογιών που δεν χρησιμοποιούνται στο σημασιολογικό ιστό, όπως πχ οι KIF, OCLM, LOOM, KM, κτλ ([WION06], [CORC00]).

5 DAML+OIL web ontology language [CONN01]

Ο σκελετός RDF είναι ακατάλληλος για τη δημιουργία οντολογιών. Αδυνατεί να επιβάλλει αυστηρούς περιορισμούς στις τιμές των στοιχείων και δεν παρέχει εργαλεία για δόμηση και οργάνωση της πληροφορίας. Επίσης, η έλλειψη ικανότητας περιγραφής της σημασιολογίας των εννοιών και των σχέσεων που υπάρχουν μεταξύ τους και παρέχονται από έναν μηχανισμό κληρονομικότητας (inheritance mechanism), τον καθιστά πολύ αδύναμο ακόμη και για την περιγραφή ιδιαίτερα απλουστευμένων συστημάτων που βασίζονται σε γνώση (knowledge based) ([BLAX06 σελ. 653], [STAA00]).

Σύμφωνα με τους Steffen Staab, Michael Erdmann, κ.α. [STAA00], οι ρόλοι που αναλαμβάνει ο RDF είναι ευδιάκριτοι. Παρέχει υποστήριξη για την περιγραφή τμημάτων γνώσης με τη μορφή τριάδων, με τη βοήθεια των οποίων είναι εύκολη η δημιουργία σχέσεων μεταξύ αντικειμένων. Η ανταλλαγή και επεξεργασία μοντέλων RDF είναι εύκολη έπειτα από τη σειριοποίηση (serialization) τους. Ο RDF μπορεί να εκμεταλλευτεί το συντακτικό της XML για τη δήλωση διακριτών ονοματοχώρων και τη δημιουργία συνδέσεων μεταξύ τους, και ως εκ τούτου, να προσφέρει μία γενικευμένη γνώση που θα μπορούν να διαμοιράζονται οι μηχανές.

“Πάνω από” τον RDF έχει οριστεί και προταθεί για χρήση από το W3C η γλώσσα RDF Schema (RDFS), η οποία προσφέρει ένα λεξιλόγιο για την μοντελοποίηση κλάσεων, ιεραρχιών ιδιοτήτων, καθώς και άλλων βασικών αρχέγονων σχημάτων (primitive schemas). Ο ρόλος της RDFS στα πλαίσια της μηχανικής γνώσης, είναι ο εξής: καθορίζει μία απλή οντολογία μέσω της οποίας μπορούν να κατασκευαστούν συγκεκριμένα έγγραφα RDF.

Αν και οι τεχνολογίες RDF και RDF Schema - RDF(S)⁶ παρέχουν εκτεταμένη υποστήριξη για τη μοντελοποίηση οντολογικών εννοιών, δεν ισχύει το ίδιο σχετικά με το θέμα των οντολογικών αξιωμάτων (ontological axioms) - τα οποία αποτελούν συστατικό κλειδί στον ορισμό οντολογιών και ένα από τα σημαντικά οφέλη των εφαρμογών που χρησιμοποιούν οντολογίες.

Σε όλες τις επίσημες γλώσσες (formal languages) υπάρχουν πλεονεκτήματα και μειονεκτήματα, αλλά για την επιλογή κάποιας ως καταλληλότερης για την επίτευξη ορισμένων στόχων, πρέπει να δοθεί έμφαση στην ικανότητα της για συμβιβασμό μεταξύ της ευκολίας

⁶ Για λόγους συντομίας, οι τεχνολογίες RDF και RDF Schema θα αναφέρονται ως RDF(S).

έκφρασης (tractability) και της δυνατότητας έκφρασης (expressiveness). Οι τεχνολογίες RDF (S) παρέχουν μεγάλη δυνατότητα έκφρασης, επειδή έχουν κατασκευαστεί ώστε να χρησιμοποιούνται για την περιγραφή των απέραντων πόρων που περιέχει ο ιστός.

Σε αντίθεση με τα υπόλοιπα συστήματα αναπαράστασης γνώσης, οι τεχνολογίες RDF(S) δεν κατασκευάστηκαν ώστε να είναι η καθοριστική απάντηση σε όλα τα προβλήματα αναπαράστασης γνώσης, αλλά ως βασικές επεκτάσιμες γλώσσες (extensible core languages). Οι μηχανισμοί των ονοματοχώρων και της ειδίκευσης επιτρέπουν στους χρήστες να καθορίζουν τα δικά τους πρότυπα με τις τεχνολογίες RDF(S) - επεκτείνοντας τους βασικούς ορισμούς και σημασιολογίες.

Η τεχνολογία DAML+OIL είναι μία σημασιολογική γλώσσα σήμανσης για την περιγραφή πόρων του ιστού (web). Χρησιμοποιεί τα υπάρχοντα πρότυπα του W3C RDF(S), και τα επεκτείνει με πλουσιότερες μεθόδους μοντελοποίησης. Κατασκευάστηκε για να εξαλείψει τα μειονεκτήματα των τεχνολογιών RDF(S), σχετικά με τη δημιουργία οντολογιών. Η DAML+OIL παρέχει μεθόδους μοντελοποίησης παρόμοιες με αυτές των γλωσσών που βασίζονται στα πλαίσια (frame-based). Η μοντελοποίηση με χρήση των πλαισίων είναι πιο διαισθητική για πολλούς χρήστες [HORR00]. Η DAML+OIL (2001) επεκτείνει το πρότυπο της DAML+OIL (2000) χρησιμοποιώντας τιμές από τους τύπους δεδομένων της XML Schema. Η τεχνολογία DAML+OIL κατασκευάστηκε από την αρχική τεχνολογία δημιουργίας οντολογιών DAML (DARPA Agent Markup Language) και DAML-ONT (Οκτώβριος 2000) σε μία προσπάθεια συνδυασμού της DAML με πολλά από τα συστατικά της OIL (Ontology Inference Layer) [DECK00]. Η γλώσσα που προκύπτει από αυτό το συνδυασμό αποτελείται από ξεκάθαρη και καλά καθορισμένη σημασιολογία.

Σύμφωνα με τον Jim Hendler [HEND4], όλες οι παραπάνω τεχνολογίες παρουσιάζουν ένα κοινό πρόβλημα: αν και αποτελούν τεχνολογίες που χρησιμοποιούνται για τη δημιουργία οντολογιών, δεν είναι τεχνολογίες που κατασκευάστηκαν για τη δημιουργία οντολογιών ιστού. Αν και αυτές οι τεχνολογίες έχουν χρησιμοποιηθεί για τη δημιουργία οντολογιών που αφορούν συγκεκριμένες κοινότητες χρηστών (ειδικότερα επιστημονικές, αλλά και εταιρικές για χρήση κυρίως στον τομέα του ηλεκτρονικού εμπορίου), δεν έχουν καθοριστεί ώστε να είναι συμβατές με την αρχιτεκτονική του παγκοσμίου ιστού πληροφοριών γενικότερα, και του

σημασιολογικού ιστού ειδικότερα. Η τεχνολογία που κατασκευάστηκε για να καλύψει αυτές τις ανάγκες είναι η γλώσσα δημιουργίας οντολογιών ιστού (Web Ontology Language - OWL).

Η OWL λύνει όλα τα παραπάνω προβλήματα παρέχοντας μία γλώσσα που χρησιμοποιεί το μηχανισμό σύνδεσης (linking) της τεχνολογίας RDF, και επιπλέον παρέχει τα παρακάτω πλεονεκτήματα για τη δημιουργία οντολογιών ιστού:

- ✓ μπορεί να εκτελείται σε μια πληθώρα συστημάτων,
- ✓ είναι επεκτάσιμη ώστε να μπορεί να καλύπτει τις ανάγκες του ιστού,
- ✓ είναι συμβατή με τα πρότυπα του ιστού που αφορούν την προσβασιμότητα (accessibility) και τη διεθνοποίηση (internationalization),
- ✓ είναι ανοιχτή τεχνολογία (ανεξάρτητη από ιδιόκτητα πρότυπα).

Οι οντολογίες ιστού που μπορούν να δημιουργηθούν με την OWL είναι πολύτιμες για πολλές περιοχές εφαρμογών, μερικές εκ των οποίων είναι οι παρακάτω:

- ✓ **πύλες ιστού (web portals)**
 - παρέχοντας κανόνες κατηγοριοποίησης για πιο αποδοτικές αναζητήσεις,
- ✓ **συλλογές πολυμέσων**
 - αναζητήσεις βάση περιεχομένου για μη αναγνώσιμα αρχεία (non-text),
- ✓ **διαχείριση ιστοσελίδων**
 - αυτόματη ταξινόμηση των δεδομένων και εγγράφων των οργανισμών,
- ✓ **σχεδίαση τεκμηρίωσης**
 - ρητή (explicit) διαχείριση των περιορισμών,
- ✓ **ευφυείς πράκτορες**
 - για την έκφραση των προτιμήσεων και/ή ενδιαφερόντων των χρηστών,
 - χαρτογράφηση περιεχομένων (content mapping) μεταξύ ιστοσελίδων,
- ✓ **υπηρεσίες ιστού και υπόλοιπες υπολογιστικές υπηρεσίες**
 - ανακάλυψη και σύνθεση υπηρεσιών του ιστού,
 - διαχείριση δικαιωμάτων και έλεγχος πρόσβασης.

3.7 Παραδείγματα οντολογιών

Έχουν κατασκευαστεί ήδη πολλές οντολογίες, είτε περιοχής για την επίλυση πολύ συγκεκριμένων προβλημάτων, είτε θεμελιώδεις για ευρεία χρήση σε πολλές περιοχές. Μερικές θεμελιώδεις οντολογίες είναι οι [WION07]:

- **Dublin Core** [WIDU06], η οποία αποτελείται από ένα σύνολο συμβάσεων για την περιγραφή αντικειμένων (ψηφιακού υλικού όπως αρχείων ήχου, βίντεο, κειμένου, εικόνες, ιστοσελίδες, κτλ) ώστε να είναι ευκολότερη η εύρεση τους κατά τη διαδικασία αναζήτησης.
- **GFO** [WIGE06], η οποία ενσωματώνει διαδικασίες και αντικείμενα παρέχοντας την κατάλληλη υποδομή για τη δημιουργία οντολογιών περιοχής (πχ φαρμακευτικές, βιολογικές, οικονομικές, κοινωνικές).
- **OpenCyc/ResearchCyc** [WICY06], οι οποίες σχετίζονται με το Cyc, ένα έργο τεχνητής νοημοσύνης σκοπός του οποίου είναι η δημιουργία μίας περιεκτικής οντολογίας και βάσης δεδομένων, που θα περιέχει κοινή γνώση της καθημερινότητας (every day common sense knowledge), έτσι ώστε οι εφαρμογές τεχνητής νοημοσύνης να μπορούν να εκτελούν ανθρώπινους συλλογισμούς. Ένας από τους στόχους των οντολογιών είναι να παρέχουν ένα απόλυτα ελεύθερο [PANT06] και απεριόριστο λεξιλόγιο σημασιολογίας για χρήση στον σημασιολογικό ιστό.
- **SUMO** [WISU06], η οποία προσφέρεται για χρήση σε μία μεγάλη ποικιλία υπολογιστικών συστημάτων επεξεργασίας πληροφοριών.
- **WordNet** [40], η οποία είναι ένα σημασιολογικό λεξικό (semantic lexicon) για την αγγλική γλώσσα. Η WordNet ομαδοποιεί ένα σύνολο αγγλικών λέξεων σε σύνολα συνώνυμα που ονομάζει synsets και παρέχει σύντομους, γενικούς προσδιορισμούς των διάφορων σημασιολογικών σχέσεων μεταξύ αυτών των συνόλων. Ο σκοπός της είναι διπλός - να παρέχει έναν συνδυασμό λεξικού (dictionary) και θησαυρού (thesaurus) το οποίο είναι διαισθητικά χρησιμότερο από την αυτόνομη και ανεξάρτητη χρήση τους, και να βοηθήσει στην αυτόματη ανάλυση κειμένου (text analysis) και στην κατασκευή εφαρμογών τεχνητής νοημοσύνης.
- **DOLCE** [MASO03], η οποία ανήκει στο έργο WonderWeb Foundation Ontologies Library (WFOL), σκοπός του οποίου είναι να παρέχει ένα σύνολο από απομονωμένες θεμελιώδεις οντολογίες που θα ανήκουν σε πολλές διαφορετικές κατηγορίες, σε αντίθεση με την παροχή μίας μονολιθικής ενότητας (module). Η οντολογία DOLCE σχετίζεται με τους

τομείς της γλωσσικής και γνωστικής μηχανικής (linguistic and cognitive engineering).

- **Gene Ontology** [WIGO06], η οποία χωρίζεται σε δύο τμήματα. Το πρώτο είναι η ίδια η οντολογία, δηλαδή ένα ελεγχόμενο λεξιλόγιο όρων που χωρίζεται σε τρεις οντολογίες και καλύπτει τους βασικούς όρους της μοριακής βιολογίας (molecular biology). Το δεύτερο τμήμα αφορά τον χαρακτηρισμό των προϊόντων γονιδίων (gene products) χρησιμοποιώντας όρους αυτής της οντολογίας.
- **Generalized Upper Model** (GUM) [BATE95], η οποία σχετίζεται με τη δημιουργία μίας ανεξαρτήτου περιοχής (domain independent) “οντολογίας με γλωσσολογικό κίνητρο” (linguistically motivated), που προορίζεται για την οργάνωση πληροφορίας ώστε να μπορεί να εκφραστεί σε φυσική γλώσσα. Η οντολογία είναι σχεδιασμένη ώστε να είναι πολυγλωσσική (multilingual) όσο αυτό είναι δυνατόν.

Οι οντολογίες που αναφέρθηκαν παραπάνω είναι ενδεικτικές. Επιπλέον αναλυτικές και καλοσχεδιασμένες λίστες οντολογιών δίνεται από τα SchemaWeb [SCHE06], Daml Ontology Librarian [DAML04], και Protégé Community Wiki [PROT06].

3.8 Εφαρμογές που χρησιμοποιούν οντολογίες

Οι οντολογίες χρησιμοποιούνται ήδη σε μία μεγάλη γκάμα εφαρμογών, και για ποικίλους λόγους. Μερικά παραδείγματα χρήσης οντολογιών σε πραγματικές εφαρμογές αναφέρονται παρακάτω.

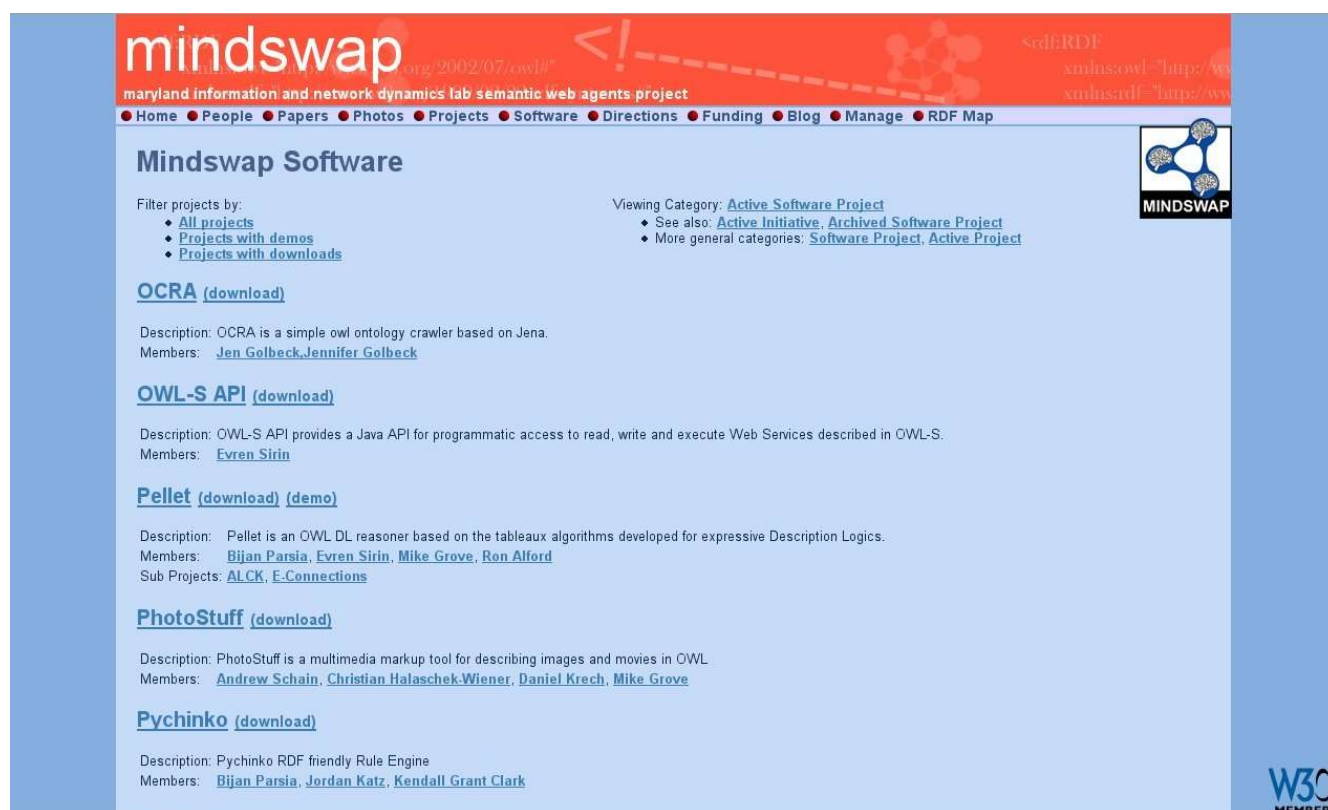
3.8.1 Ιστοσελίδα mindswap.org

Η ιστοσελίδα <http://www.mindswap.org> κατασκευάστηκε από την ερευνητική ομάδα του σημασιολογικού ιστού MINDSWAP, που δημιουργήθηκε στα πλαίσια του Maryland Information And Dynamics Laboratory (MIND LAB). Σύμφωνα με τους κατασκευαστές της, είναι η πρώτη (χρονολογικά) ιστοσελίδα που κατασκευάστηκε ώστε να υποστηρίζει σημασιολογία. Όπως αναφέρουν και οι κατασκευαστές της, η ιστοσελίδα συνδυάζει πολλές τεχνολογίες του κλασικού παγκοσμίου ιστού πληροφοριών (XHTML, PHP, κτλ) με

τεχνολογίες του σημασιολογικού ιστού (RDF, RDFS, DAML+OIL, OWL).

Η ιστοσελίδα υποστηρίζει την επίσημη γλώσσα δημιουργίας οντολογιών ιστού (OWL). Υπάρχουν σύνδεσμοι που όταν επιλεχθούν εμφανίζεται περιεχόμενο με μορφή αναγνώσιμη από τον άνθρωπο (πχ ο σύνδεσμος "Software", εικόνα 3-3), καθώς και σύνδεσμοι που όταν επιλεχθούν εμφανίζεται περιεχόμενο με μορφή αναγνώσιμη από τις μηχανές (πχ ο σύνδεσμος "Photos", εικόνα 3-4).

Εντός της ιστοσελίδας υπάρχουν χρήσιμα επιστημονικά άρθρα (σύνδεσμος "Papers"), πληροφορίες για τα μέλη της ομάδας (σύνδεσμος "People"), χάρτης RDF ("RDF map"), κτλ.



Εικόνα 3-3 Ο σύνδεσμος Software της ιστοσελίδας mindswap.org.

Ο σύνδεσμος "Photos" παρουσιάζει μεγάλο ενδιαφέρον, επειδή χρησιμοποιεί οντολογίες περιοχής, όπως η NATO2.owl για την παρουσίαση πληροφοριών που αφορούν πολεμικά οχήματα του NATO, καθώς και θεμελιώδεις οντολογίες, όπως η friend of a friend (foaf), για

την περιγραφή "γνωστών" φυσιογνωμιών.

Τέλος, ο σύνδεσμος "Software" περιέχει πολύ χρήσιμες εφαρμογές στα πλαίσια του σημασιολογικού ιστού.

3.8.2 Materialized Ontology View Extractor (MOVE) και Semantic Completeness Optimization Scheme (SCOS)

Σύμφωνα με τους Mehul Bhatt, Carlo Wouters, κ.α. [BHAT04], το σύστημα MOVE και το σχήμα SCOS κατασκευάστηκαν από ένα πλήθος επιστημόνων των πανεπιστημίων La Trobe και Monash της Αυστραλίας.

Οι δημιουργοί των MOVE και SCOS εντόπισαν ότι οι οντολογίες του σημασιολογικού ιστού αυξάνουν πάρα πολύ σε μέγεθος αλλά και σε πολυπλοκότητα, με αποτέλεσμα οι εφαρμογές να δεσμεύονται με όλα τα τμήματα τους άσκοπα, αφού ουσιαστικά χρησιμοποιούν μόνο ένα μικρό κομμάτι τους. Αυτό φυσικά σημαίνει αυτόματα και μείωση της αποδοτικότητας των εφαρμογών.

Με το σύστημα MOVE έγινε προσπάθεια να λυθεί το παραπάνω πρόβλημα προσεγγίζοντας τη δημιουργία/βελτιστοποίηση υπό-οντολογιών (sub-ontologies) από μία μεγάλη και πολύπλοκη οντολογία. Η εφαρμογή MOVE υποστηρίζει πολλαπλές τεχνολογίες δημιουργίας οντολογιών (OWL, DAML+OIL, Ontolingua) σε αφαιρετικό επίπεδο. Το σχήμα SCOS κατασκευάστηκε για να είναι δυνατός ο έλεγχος εγκυρότητας των υπό-οντολογιών που προκύπτουν από τις οντολογίες. Επίσης, η αρχιτεκτονική του SCOS είναι κατάλληλη ώστε να μπορεί να λειτουργεί και σε περιβάλλον συστοιχίας (cluster).

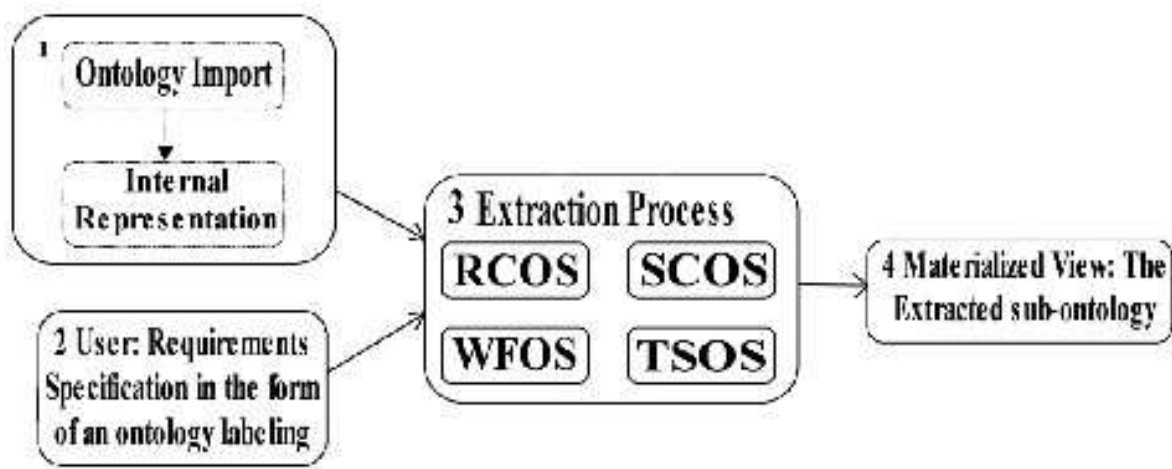


Εικόνα 3-4 Ο σύνδεσμος Photos της ιστοσελίδας mindswap.org.

Η αρχιτεκτονική του συστήματος MOVE με τα διάφορα σχήματα που υποστηρίζει φαίνεται στην εικόνα 3-5. Όπως φαίνεται στην εικόνα, η αρχιτεκτονική του MOVE χωρίζεται σε επίπεδα. Συνοπτικά οι λειτουργίες του κάθε επιπέδου είναι οι παρακάτω:

- **Επίπεδο 'εισαγωγής οντολογίας' (ontology import layer - 1):** Είναι υπεύθυνο για τη διαχείριση διαφόρων προτύπων αναπαράστασης οντολογιών, με τα οποία πρέπει να είναι συμβατή η διαδικασία εξαγωγής. Αυτό επιτυγχάνεται με μετατροπή της εξωτερικής αναπαράστασης της οντολογίας σε εσωτερική μορφή που γνωρίζει το σύστημα.
- **Επίπεδο 'εισαγωγής ετικετών' (labeling - 2):** Αυτό το επίπεδο είναι πολύ σημαντικό στην αλληλεπίδραση των χρηστών και του αλγορίθμου εξαγωγής πληροφοριών της βασικής οντολογίας στην υπό-οντολογία που θα προκύψει. Επιτρέπει στον χρήστη να παρέχει πληροφορία, μέσω της οποίας καθορίζει τι επιθυμεί/δεν επιθυμεί να υπάρχει στην υπό-οντολογία. Κάθε οντολογικό στοιχείο μπορεί να έχει τρεις ετικέτες, selected αν επιλέγεται να είναι παρών στην υπό-οντολογία, deselected αν επιλέγεται να είναι απών, και void όταν είναι επιθυμητό να λάβει απόφαση ο αλγόριθμος.

- **Επίπεδο 'διαδικασίας εξαγωγής' (extraction process - 3):** Το επίπεδο αυτό περιλαμβάνει διάφορα σχήματα (schemes) ελέγχου της συνέπειας (consistency), εγκυρότητας (validity) και βελτιστοποίησης (optimization) της υπό-οντολογίας. Τέτοιου είδους σχήματα είναι το Semantic Completeness Optimization Scheme (SCOS1-SCOS3), WellFormedness Optimization Scheme (WFOS1-WFOS5) και το Total Simplicity Optimization Scheme (TSOS1-TSOS3).
- **Επίπεδο 'υλοποιημένης άποψης οντολογίας' (materialized ontology view - 4):** Η υλοποιημένη άποψη μια οντολογίας χρειάζεται έτσι ώστε να εξασφαλισθεί ότι η υπό-οντολογία που έχει προκύψει είναι πλήρως ανεξάρτητη από την βασική οντολογία, και θα πρέπει να είναι απόλυτα σωστή ακόμη κι αν η βασική οντολογία αφαιρεθεί.



Εικόνα 3-5 Η αρχιτεκτονική του συστήματος MOVE.

3.8.3 Το σύστημα παροχής βοήθειας σε καθηγητές πληροφορικής Kasai

Σύμφωνα με τους Toshinobu Kasai και Haruhisa Yamaguchi [KASA05], σε πολλές χώρες όπως η Ιαπωνία, οι καθηγητές πληροφορικής δεν είναι πολύ εξειδικευμένοι, και είναι ιδιαίτερα δύσκολο για αυτούς να αποκτήσουν την απαραίτητη γνώση και τις ικανότητες που απαιτούνται για τη διδασκαλία, μιας και δεν έχει ξεκαθαριστεί ακόμη τι χρειάζεται να γνωρίζει ένας καθηγητής πληροφορικής για να θεωρείται ικανός. Αν και το ενδιαφέρον για την

επιστήμη της πληροφορικής αυξάνει συνεχώς και στην Ιαπωνία, οι αξιόλογοι καθηγητές πληροφορικής είναι ελάχιστοι.

Αν και πολλοί οργανισμοί μέσω ιστοσελίδων παρέχουν εκπαιδευτικό υλικό και πολλών ειδών βοηθήματα στους καθηγητές πληροφορικής (σχέδιο διδασκαλίας, ψηφιακό περιεχόμενο, διδακτικές μεθόδους, κτλ), τέτοιου είδους βοηθήματα είναι δύσκολα στην εύρεση, και παρουσιάζουν το πρόβλημα των διφορούμενων εννοιών. Πρέπει να βρεθεί κάποιος τρόπος ώστε οι πληροφορίες να είναι καλά οργανωμένες, εύκολες στην εύρεση, και οι σημασίες των εννοιών να είναι ξεκάθαρες. Το σύστημα Kasai⁷ κατασκευάστηκε έπειτα από συνεργασία επιστημονικών συνεργατών των ιαπωνικών πανεπιστημίων Okayama, Sacred heart, και Osaka, με σκοπό την επίλυση των παραπάνω προβλημάτων.

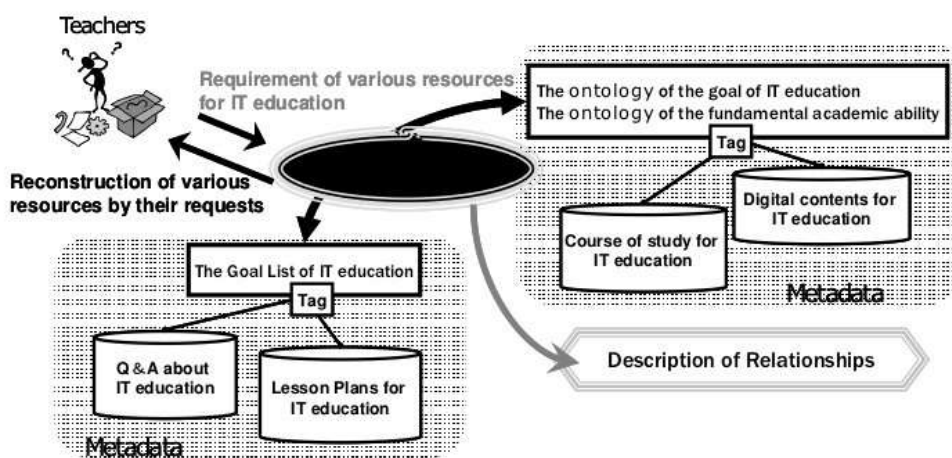
Οι δημιουργοί του συστήματος κατασκεύασαν μία αφηρημένη (abstract) οντολογία που περιέχει ξεκάθαρες έννοιες που σχετίζονται με την διδασκαλία στην επιστήμη της πληροφορικής. Εκτός από τις ξεκάθαρες έννοιες που υπάρχουν στην επιστήμη της πληροφορικής, υπάρχουν και οι αμφιλεγόμενες έννοιες, που προκύπτουν έπειτα από έρευνα. Αυτό σημαίνει ότι πρέπει να βρεθεί ένας τρόπος έτσι ώστε να υπάρχει πρόσβαση σε όλες τις έννοιες, απλά να γνωστοποιείται στους καθηγητές ότι οι ερευνητικές έννοιες είναι χαμηλότερης ποιότητας.

Για να λυθεί αυτό το πρόβλημα, οι δημιουργοί της οντολογίας κατασκεύασαν την κατάλληλη υποδομή (framework), ώστε να είναι δυνατή και η χρήση αυτής της οντολογίας, και η επαναχρησιμοποίηση μίας οντολογίας που καθορίζει μία λίστα στόχων (Goal list) σχετικά με την επιμόρφωση των καθηγητών πληροφορικής (IT education), η οποία έχει προκύψει έπειτα από έρευνα.

⁷ Αν και το σύστημα δε συνοδεύεται από κάποια επίσημη ονομασία, αναφέρεται ως Kasai που είναι το επώνυμο του Toshinobu Kasai, ενός από τους δημιουργούς του.

Τέλος, κατασκευάστηκε μία εφαρμογή σημασιολογικού ιστού που χρησιμοποιεί τη συγκεκριμένη υποδομή, έτσι ώστε να είναι εύκολο για έναν καθηγητή πληροφορικής δίχως προηγούμενη εμπειρία, να αναζητήσει πληροφορίες για θέματα διδασκαλίας με ιδιαίτερη ευκολία και με συνέπεια. Για τη δημιουργία των οντολογιών χρησιμοποιήθηκαν οι τεχνολογίες RDF και RDF Schema .

Μία προσέγγιση της εξωτερικής μορφής της αρχιτεκτονικής του συστήματος Kasai υπάρχει στην εικόνα 3-6. Οι καθηγητές (Teachers) θέλουν να αντλήσουν πληροφορίες σχετικά με διάφορα θέματα της επιστήμης της πληροφορικής. Οι πόροι (resources) που μπορούν να αντλήσουν χωρίζονται σε δύο τμήματα μεταδεδομένων (Metadata). Το ένα τμήμα (αριστερό) συνοδεύεται από τον τίτλο “λίστα στόχων για την επιμόρφωση της επιστήμης της πληροφορικής” (Goal List of IT education), και αποτελείται από ερωτήσεις και απαντήσεις σχετικά με την διδασκαλία της πληροφορικής (Q & A about IT education), και πλάνα διδασκαλίας (Lesson Plans for IT education). Το άλλο τμήμα (δεξί) συνοδεύεται από τον τίτλο “οντολογία των στόχων της επιμόρφωσης της επιστήμης της πληροφορικής” (The ontology of the goal of IT education) και “οντολογία της ακαδημαϊκής ικανότητας” (The ontology of the fundamental academic ability), και περιέχει τις γνώσεις που χρειάζεται κάποιος για να γίνει καθηγητής πληροφορικής (Course of study for IT education), και τα ψηφιακά περιεχόμενα που προσφέρονται στους καθηγητές (Digital contents for IT education).

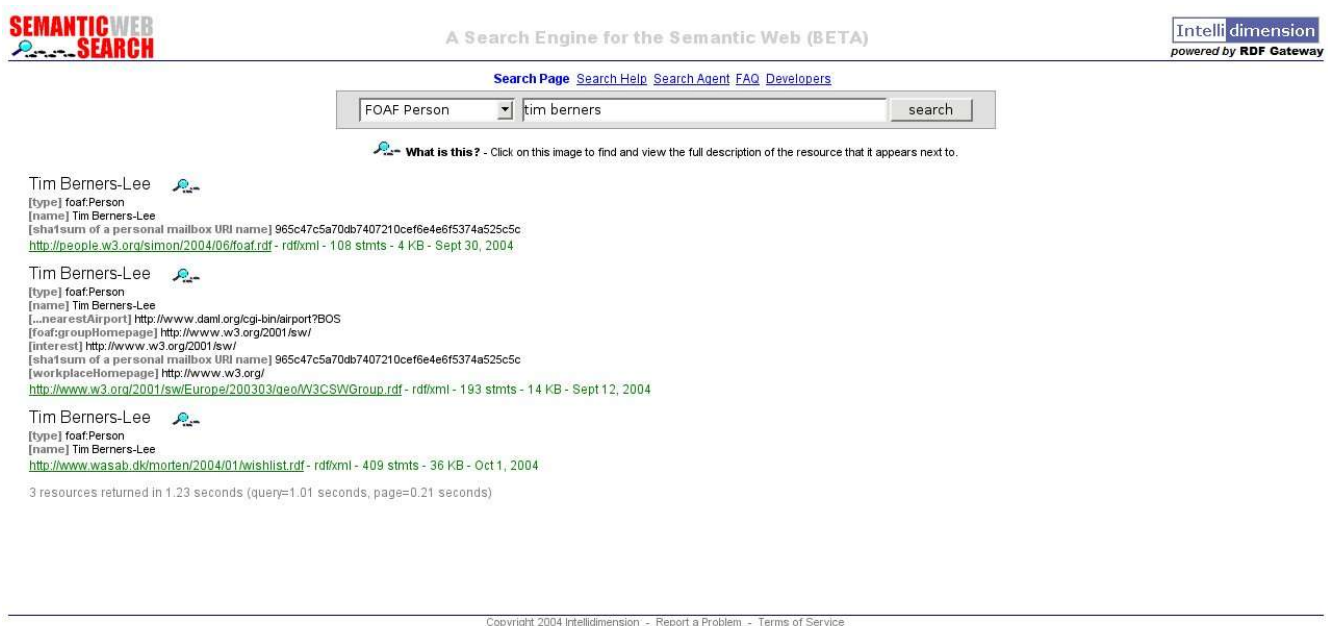


Εικόνα 3-6 Η εξωτερική μορφή της αρχιτεκτονικής του συστήματος Kasai.

3.8.4 Μηχανή αναζήτησης semanticwebsearch.com

Μία δοκιμαστική έκδοση (beta) μίας μηχανής αναζήτησης για το σημασιολογικό ιστό είναι η <http://www.semanticwebsearch.com/>. Η μηχανή που έχει κατασκευαστεί από την εταιρία Intellidimension, δίνει τη δυνατότητα αναζήτησης πόρων με συγκεκριμένες σημασιολογικές ιδιότητες (πχ κανάλι RSS, κλάση RDF, κτλ) ή οποιαδήποτε εκ των ιδιοτήτων που υποστηρίζει. Αξίζει να αναφερθεί ότι η μηχανή συνοδεύεται από πολύ καλή τεκμηρίωση συμπεριλαμβανομένων και παραδειγμάτων για τους διαφορετικούς τρόπους αναζήτησης που υποστηρίζει. Τέλος, υπάρχει και ειδικός πράκτορας ο οποίος αναλαμβάνει να επιστρέψει μία λίστα αποτελεσμάτων από μία σειρά σύνθετων κριτηρίων αναζήτησης που δέχεται ως είσοδο από τον χρήστη [INTEL04].

Στην εικόνα 3-7 φαίνονται τα αποτελέσματα που εξήγαγε η μηχανή έπειτα από την αναζήτηση εντός της οντολογίας FOAF (Friend Of A Friend) με κριτήριο το “tim berners”.



Εικόνα 3-7 Αποτελέσματα αναζήτησης στην οντολογία FOAF.

Στην εικόνα 3-8 φαίνονται τα κριτήρια αναζήτησης που ανατέθηκαν στον σημασιολογικό πράκτορα της μηχανής. Συνοπτικά, του ανατέθηκε να αναζητήσει όλες τις RDF κλάσεις με

ιδιότητες που περιέχουν τον όρο “owl”. Στην εικόνα 3-9 φαίνονται τα αποτελέσματα αυτής της αναζήτησης.

The screenshot shows the Semantic Web Search interface. At the top left is the "SEMANTIC WEB SEARCH" logo. In the center is the text "A Search Engine for the Semantic Web (BETA)". At the top right is the "Intelli dimension" logo with the text "powered by RDF Gateway" below it. Below the logo is a navigation bar with links: "Search Page", "Search Help", "Search Agent", "FAQ", and "Developers". The main section is titled "Search Agent" and contains the instruction: "Use the hyper-text below to create a sentence that accurately describes the resource(s) you are trying to find." To the right of this instruction are links for "Advanced Edit" and "Clear Search". Below the instruction is a large text box containing a query: "Find [any resource](#) that is a type of [RDF Class](#) with ☒ [any property](#) that [contains](#) the term ['owl'](#) and [any property](#) that [contains](#) [any value](#)". Below the text box is a "search" button. At the bottom of the page is a footer with the text: "Copyright 2004 Intellidimension - [Report a Problem](#) - [Terms of Service](#)".

Εικόνα 3-8 Κριτήρια αναζήτησης του πράκτορα της μηχανής semanticwebsearch.com.

RDF Class

 **What is this?** - Click on this image to find and view the full description of the resource that it appears next to.

[Night_bird](#)

[rdfs:description] any bird associated with night: owl, nightingale, nighthawk; etc
[type] rdfs:Class
[label] Night_bird
[comment] any bird associated with night: owl, nightingale, nighthawk; etc
[subClassOf] wn:Bird
<http://xmains.com/wordnet/1.6/Bird - rdf/xml> - 231 stmts - 20 KB - Oct 15, 2004

[Bubo virginianus](#)

[rdfs:description] brown North American horned owl
[type] rdfs:Class
[label] Bubo_virginianus
[comment] brown North American horned owl
[subClassOf] wn:Horned_owl
http://xmains.com/wordnet/1.6/Horned_owl - rdf/xml - 62 stmts - 5 KB - Oct 18, 2004

[Laughing_owl\[1\]](#)

[rdfs:description] almost extinct owl of New Zealand
[type] rdfs:Class
[label] Laughing_owl[1]
[comment] almost extinct owl of New Zealand
[subClassOf] wn:Owl
http://xmains.com/wordnet/1.6/Laughing_owl - rdf/xml - 52 stmts - 4 KB - Oct 18, 2004

[Hoot](#)

[rdfs:description] a loud raucous cry (as of an owl)
[type] rdfs:Class
[label] Hoot
[comment] a loud raucous cry (as of an owl)
[comment] a cry or noise made to express displeasure or contempt
[subClassOf] wn:Outcry
<http://xmains.com/wordnet/1.6/Outcry - rdf/xml> - 230 stmts - 18 KB - Oct 18, 2004

[Owl](#)

[type] rdfs:Class
[label] Owl
[comment] nocturnal bird of prey with hawk-like beak and claws and large head with front-facing eyes
[rdfs:description] nocturnal bird of prey with hawk-like beak and claws and large head with front-facing eyes
[subClassOf] wn:Raptor
<http://xmains.com/wordnet/1.6/Raptor - rdf/xml> - 98 stmts - 8 KB - Oct 18, 2004

[Old World scops owl](#)

[rdfs:description] European scops owl

Εικόνα 3-9 Αποτελέσματα αναζήτησης του πράκτορα της μηχανής semanticwebsearch.com.

3.8.5 Βάση δεδομένων Oracle 10g

Σύμφωνα με τους Chuck Murray [MURR05], Xavier Lopez, και Susie Stephens [LOPE06], το σύστημα διαχείρισης βάσεων δεδομένων Oracle με την τεχνολογία Oracle Spatial 10g, εισάγει την πρώτη ανοιχτή, επεκτάσιμη, ασφαλή και αξιόπιστη πλατφόρμα διαχείρισης δεδομένων με πλήρη υποστήριξη των τεχνολογιών RDF(S). Οι κατασκευαστές εφαρμογών μπορούν να προσθέτουν νόημα στα δεδομένα και μεταδεδομένα που χρησιμοποιούν, ορίζοντας ένα σύνολο όρων καθώς και σχέσεις μεταξύ αυτών των όρων. Τα σύνολα όρων που προκύπτουν (“οντολογίες”) επιτρέπουν την υποβολή αιτημάτων, την ανάλυση, και άλλου είδους ενέργειες, που βασίζονται πλέον σε σημασιολογικό περιεχόμενο, και όχι σε απλές τιμές δεδομένων.

Οι οντολογίες χρησιμοποιούνται ολόένα και περισσότερο για την περιγραφή γνώσης

συγκεκριμένων περιοχών (domain-specific). Τα σύνολα δεδομένων που προκύπτουν από τις οντολογίες και συχνά περιέχουν εκατοντάδες εκατομμύρια όρων και σχέσεων, μπορούν να αποθηκευτούν με τη μορφή τριάδων, χρησιμοποιώντας το μοντέλο RDF. Η Oracle 10g επιτρέπει την δήλωση δισεκατομμυρίων τριάδων, ώστε να μπορεί να ικανοποιήσει τις ανάγκες ακόμη και των πιο απαιτητικών εφαρμογών.

Η χρήση σημασιολογικών μοντέλων δεδομένων με τη βοήθεια της Oracle 10g, εισάγει σημαντικά πλεονεκτήματα έναντι της προσέγγισης που βασίζεται σε αρχεία (file-based) και άλλου είδους προσεγγίσεων (πχ της αντικειμενοστραφούς) που χρησιμοποιούν τα υπόλοιπα συστήματα διαχείρισης. Ορισμένα από αυτά είναι:

- ✓ **Χαμηλό ιδιοκτησιακό κόστος (Low cost of ownership):** Οι εφαρμογές του σημασιολογικού ιστού μπορούν να συνδυαστούν με άλλες εφαρμογές και να επεκταθούν σε εταιρικό επίπεδο, με τα δεδομένα να αποθηκεύονται κεντρικά (centrally), μειώνοντας τα ιδιοκτησιακά κόστη. Πέρα από τα πλεονεκτήματα της κεντρικής αποθήκευσης και υποβολής αιτημάτων, οι αρχιτεκτονικές που βασίζονται σε υπηρεσίες (Service Oriented Architectures) ελαχιστοποιούν τις ανάγκες εγκατάστασης και συντήρησης εφαρμογών στην πλευρά του πελάτη (client-side), καθώς και τις ανάγκες αποθήκευσης και διαχείρισης δεδομένων εκτός της εταιρικής βάσης δεδομένων, που είναι εγκατεστημένη στην πλευρά του διακομιστή (server-side).
- ✓ **Χαμηλό ρίσκο (Low risk):** Τα μοντέλα RDF και OWL μπορούν να τοποθετούνται εξολοκλήρου στην εταιρική βάση δεδομένων, σε συνδυασμό με υπάρχοντα δεδομένα, έγγραφα κειμένου, έγγραφα XML, κτλ. Αυτό έχει ως αποτέλεσμα τη δημιουργία ολοκληρωμένων, επεκτάσιμων, ασφαλών και αποδοτικών εφαρμογών. Οι πελάτες μπορούν να επιλέξουν την πλατφόρμα του διακομιστή που επιθυμούν (Unix, GNU/Linux, Windows NT) ώστε να διαχειρίζονται τις εφαρμογές τους.
- ✓ **Ανοιχτή αρχιτεκτονική (Open architecture):** Οι καλύτεροι κατασκευαστές σημασιολογικών εφαρμογών, συμπεριλαμβανομένων των εταιριών Cerebra και Siderean, παρέχουν ήδη υποστήριξη στο μοντέλο Oracle 10g RDF. Επιπρόσθετα, υπάρχουν ήδη συνδεόμενες υπομονάδες (plug-in) υποστήριξης της Oracle 10g από πολλούς κατασκευαστές ελεύθερου λογισμικού, όπως οι Protégé, Sesame, και Cytospace.

Η βάση δεδομένων Oracle 10g RDF διασφαλίζει ότι οι κατασκευαστές εφαρμογών θα

ωφεληθούν από την κλιμάκωση της ώστε να κατασκευάζουν επεκτάσιμες και ασφαλείς σημασιολογικές εφαρμογές. Μερικοί τομείς εφαρμογής της 10g είναι:

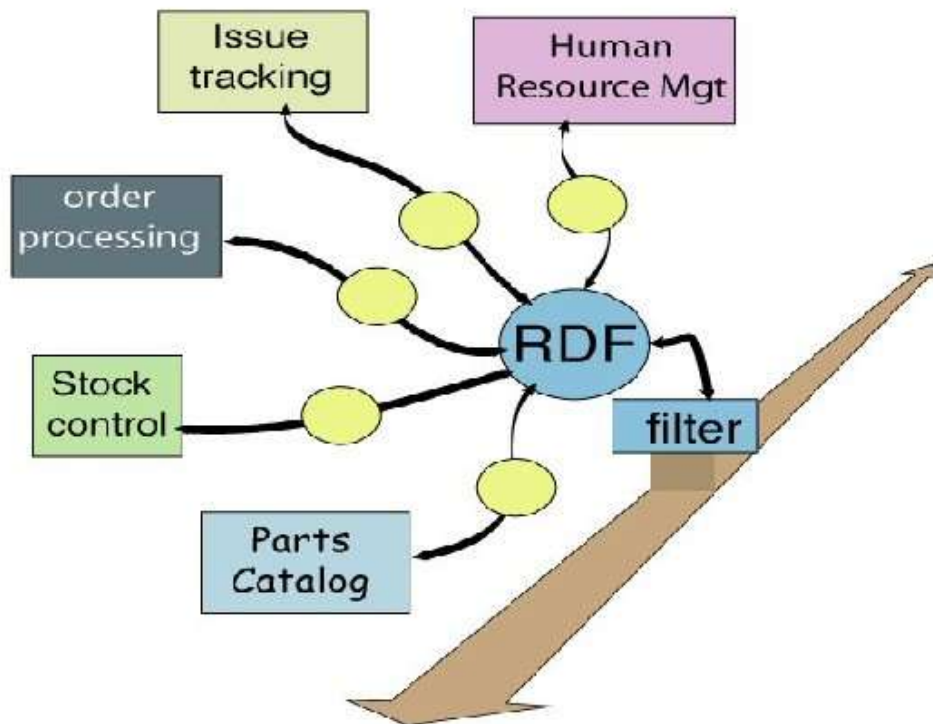
- ✓ **Κοινωνικές επιστήμες:** βιολογικές αναλύσεις, ανακάλυψη και ενισχυμένη αναζήτηση.
- ✓ **Νοημοσύνη:** ολοκλήρωση (integration) περιεχομένων και δεδομένων, δημιουργία συλλογισμών και δηλώσεων.
- ✓ **Ολοκληρωμένες επιχειρηματικές εφαρμογές:** ολοκλήρωση δεδομένων και συστημάτων, ολοκλήρωση σημασιολογικών επιχειρησιακών πρωτοβουλιών (semantic enterprise integration) και σημασιολογικές υπηρεσίες ιστού.
- ✓ **CRM/ERP:** ολοκληρωμένες αλυσίδες ανεφοδιασμού (supply chain integration), βελτιστοποίηση πηγών (sourcing optimization) και αυτοματοποίηση εξυπηρέτησης πελατών.

Μία ολοκληρωμένη επιχειρηματική εφαρμογή, φαίνεται στην εικόνα 3-10. Η εφαρμογή χρησιμοποιεί μία επιχειρησιακή οντολογία περιοχής, η οποία είναι αποθηκευμένη σε μορφή RDF. Η οντολογία παρέχει σχέσεις μεταξύ των επιμέρους επιχειρησιακών εφαρμογών (παραγγελίες, αλυσίδες ανεφοδιασμού, ανθρωπίνοι πόροι, κτλ). Οι χρήστες της εφαρμογής μπορούν να υποβάλλουν αιτήματα στο σύστημα μέσω του καθολικού (global) μοντέλου RDF, ενώ σε κάθε άλλη περίπτωση θα ήταν υποχρεωτικό να υλοποιηθούν πολλαπλά σχήματα, ένα ξεχωριστό για τη σχέση με καθεμία επιχειρησιακή εφαρμογή.

3.8.6 Πολιτική διαχείρισης ΚΑoS για υπηρεσίες του σημασιολογικού ιστού

Σύμφωνα με τους Andrzej Uszok, Jeffrey M. Bradshaw, κ.α. [USZO04], το ΚΑoS είναι μία προσπάθεια προσέγγισης της δημιουργίας πολιτικών (policies) για τις υπηρεσίες του σημασιολογικού ιστού (semantic web services), με χρήση του σημασιολογικού ιστού, συγκεκριμένα χρησιμοποιώντας την τεχνολογία OWL. Παραδείγματα μερικών υπηρεσιών του σημασιολογικού ιστού αποτελούν τα σημασιολογικά τείχη προστασίας (semantic firewalls) και η συνεργατική αναζήτηση και διάσωση (Coalition Search and Rescue) ([USZO04], [KIFE02]). Οι πολιτικές έχουν την έννοια του δυναμικού καθορισμού περιορισμών και ρύθμισης της συμπεριφοράς ενός συστήματος, δίχως να απαιτούνται αλλαγές στον πηγαίο κώδικα του.

Πολιτικές υιοθετούνται ολοένα και περισσότερο τόσο σε ακαδημαϊκό όσο και σε βιομηχανικό επίπεδο, για τη δημιουργία δυναμικά προσαρμοζόμενων (dynamic adjustable) συστημάτων.



Εικόνα 3-10 Ολοκληρωμένη επιχειρηματική εφαρμογή με τη βοήθεια της Oracle 10g.

Οι προσεγγίσεις που βασίζονται σε πολιτικές (policy-based) έχουν πολλά πλεονεκτήματα, όπως:

- ✓ επαναχρησιμοποίηση (reusability),
- ✓ αποδοτικότητα (efficiency),
- ✓ επεκτασιμότητα (extensibility),
- ✓ επαλήθευση (verifiability),
- ✓ υποστήριξη απλών και περίπλοκων συστατικών (components),
- ✓ προστασία από τη δημιουργία κακοσχεδιασμένων (poorly designed), με σφάλματα (buggy), ή κακόβουλων (malicious) συστατικών,
- ✓ αιτιολόγηση συμπεριφοράς.

Πολιτικές εφαρμόζονται πολύ συχνά και για την αυτοματοποίηση εργασιών που αφορούν τη διαχείριση υπολογιστικών δικτύων, όπως η ασφάλεια, η επανάκαμψη (recovery), η

παραμετροποίηση (configuration), και η ποιότητα υπηρεσίας (quality of service).

Σήμερα, το πεδίο εφαρμογής των πολιτικών διαχείρισης διευρύνεται πέρα από τις προαναφερόμενες κλασικές εφαρμογές. Δημιουργούνται νέες προκλήσεις που πρέπει να αντιμετωπιστούν από τις πολιτικές, όπως:

- ➔ Η διαχείριση των ψηφιακών δικαιωμάτων (digital-rights management), το φιλτράρισμα (filtering) και ο μετασχηματισμός (transformation) πληροφοριών, και η πρόσβαση βάση των ικανοτήτων (capability-based access).
- ➔ Τα ενεργά δίκτυα (active networks), η ευέλικτη υπολογιστική (agile computing), τα διεισδυτικά (pervasive) και φορητά (mobile) συστήματα.
- ➔ Η οργανωτική μοντελοποίηση (organizational modeling) και η τυποποίηση οργανωτικών συμφωνιών (formalizing cross-organizational agreements).
- ➔ Τα μοντέλα πεποιθήσεων (trust models), η προέλευση της πληροφορίας (information pedigrees), και η έμπιστη διαχείριση (trust management).
- ➔ Η αποδοτική αλληλεπίδραση ανθρώπου-Η/Υ (human-machine interaction): προσαρμοζόμενη αυτονομία (adjustable autonomy), ασφάλεια (safety), διευκόλυνση ομαδικής εργασίας (teamwork facilitation), διαχείριση διακοπών και ειδοποιήσεων (interruption and notification management), κτλ.
- ➔ Η υποστήριξη της προσπάθειας των ανθρώπων να ανακτήσουν, κατανοήσουν, και αναλύσουν όλες τις πολιτικές που σχετίζονται με κάποια κατάσταση.

Οι υπηρεσίες και τα εργαλεία του ΚΑoS επιτρέπουν τον καθορισμό, τη διαχείριση, την ανάλυση διενέξεων (conflict resolution) και την επιβολή πολιτικών, σε συγκεκριμένα πλαίσια που καθιερώνονται από σύνθετες οργανωτικές δομές και παριστάνονται με τη μορφή περιοχών (domains). Μερικές περιοχές που έχει εφαρμοστεί επιτυχώς το ΚΑoS είναι οι εξής:

- πολιτική διαχείρισης σε περιβάλλοντα αρχιτεκτονικής πλέγματος (grid-computing),
- εξακρίβωση της συμμόρφωσης με κάποια πολιτική (policy compliance) για τη σύνθεση χρονοδιαγραμμάτων (workflow composition) που αφορούν τις υπηρεσίες του σημασιολογικού ιστού,
- επιβολή πολιτικών κατά τη θέσπιση των χρονοδιαγραμμάτων (workflow enactment).

Το ΚΑoS χρησιμοποιεί έννοιες οντολογιών (κωδικοποιημένες σε OWL) για την κατασκευή

πολιτικών. Κατά τη διαδικασία εκκίνησης του (bootstrap), το KAOs αρχικά φορτώνει μία βασική οντολογία (KAOs Policy Ontology), στην οποία ορίζονται έννοιες που χρησιμοποιούνται για την περιγραφή ενός γενικευμένου περιβάλλοντος και πολιτικών. Έπειτα, το KAOs φορτώνει επιπλέον οντολογίες “πάνω από” (on top) τη βασική οντολογία, επεκτείνοντας έτσι τις έννοιες της για το συγκεκριμένο επιθυμητό ελεγχόμενο περιβάλλον (controlled environment) και την επιθυμητή περιοχή εφαρμογής (application domain).

Η πολιτική του KAOs που είναι κωδικοποιημένη με τη βοήθεια της OWL, είναι ένα στιγμιότυπο μίας εκ των τεσσάρων βασικών κλάσεων που περιγράφουν πολιτικές:

- PositiveAuthorization,
- NegativeAuthorization,
- PositiveObligation,
- NegativeObligation.

Οι τιμές των ιδιοτήτων καθορίζουν πληροφορίες διαχείρισης για μία συγκεκριμένη πολιτική (πχ την προτεραιότητα της). Ο τύπος του στιγμιότυπου της πολιτικής καθορίζει το είδος του περιορισμού που θα εφαρμόσει το KAOs σε κάποια δεδομένη ενέργεια, ενώ η κλάση ενέργειας (action class) της πολιτικής χρησιμοποιείται για τον καθορισμό της καταλληλότητας μίας πολιτικής για μία δεδομένη περίπτωση.

Η πολιτική υπηρεσίας του KAOs (KAOs Policy Service) κάνει διάκριση μεταξύ θετικών και αρνητικών εξουσιοδοτήσεων (authorizations – περιορισμοί που επιτρέπουν ή απαγορεύουν κάποια ενέργεια) και μεταξύ θετικών και αρνητικών υποχρεώσεων (obligations – περιορισμοί που απαιτούν κάποια ενέργεια όταν προκύπτει μία κατάσταση ή ενεργοποιείται ένα έναυσμα - trigger που βασίζεται σε γεγονότα). Άλλου είδους πολιτικές (πχ αντιπροσωπευτικές ή βασισμένες σε ρόλους) χτίζονται από τα αρχέγονα της βασικής περιοχής (basic domain primitives) σε συνδυασμό με τους τέσσερις βασικούς τύπους πολιτικής.

Η αναπαράσταση των πολιτικών σε OWL διευκολύνει τους συλλογισμούς (reasoning) σχετικά με το ελεγχόμενο περιβάλλον, τη σχέση μεταξύ των πολιτικών και την κοινοποίηση (disclosure), την ανίχνευση διενέξεων (conflict detection) στις πολιτικές, και την εναρμόνιση (harmonization). Διευκολύνει επίσης τους συλλογισμούς σχετικά με τη δομή της περιοχής

(domain structure) και τις έννοιες που εκμεταλλεύονται αλγορίθμους ταξινόμησης και λογικές περιγραφές. Το KAOs μπορεί να αναγνωρίσει, και αν είναι επιθυμητό να διορθώσει/εναρμονίσει (harmonize) συγκρουόμενες πολιτικές με τη βοήθεια αλγορίθμων που έχουν υλοποιηθεί με το εργαλείο Java Theorem Prover (JTP) του πανεπιστημίου Stanford.

Τα σημαντικότερα χαρακτηριστικά του KAOs είναι:

→ **Ομοιογενής αναπαράσταση πολιτικών (Homogeneous policy representation).**

Επειδή όλες οι πολιτικές του KAOs είναι υλοποιημένες σε OWL, οποιοδήποτε εργαλείο τρίτων κατασκευαστών ή περιβάλλον που υποστηρίζει OWL μπορεί να εκτελέσει εξειδικευμένες αναλύσεις της πλήρους βάσης γνώσης ανεξαρτήτως του KAOs.

→ **Ωριμότητα (Maturity).** Οι υπηρεσίες του KAOs έχουν χρησιμοποιούνται εδώ και αρκετά χρόνια σε ένα μεγάλο εύρος εφαρμογών και λειτουργικών πλατφορμών.

→ **Περιεκτικότητα (Comprehensiveness).** Σε αντίθεση με πολλές προσεγγίσεις οι οποίες συμβιβάζονται μόνο με απλές φόρμες ελέγχου πρόσβασης ή εξουσιοδότησης, το KAOs υποστηρίζει ταυτόχρονα πολιτικές εξουσιοδότησης και υποχρεώσεων. Επιπλέον, έχει υλοποιηθεί μία πλήρης υποδομή για την διαχείριση πολιτικών που περιλαμβάνει ένα μεγάλο εύρος ικανοτήτων, από περίπλοκες διεπιφάνειες χρήστη για τον καθορισμό προδιαγραφών, ως την ανάλυση μίας γενικευμένης πολιτικής μηχανισμού κοινοποίησης.

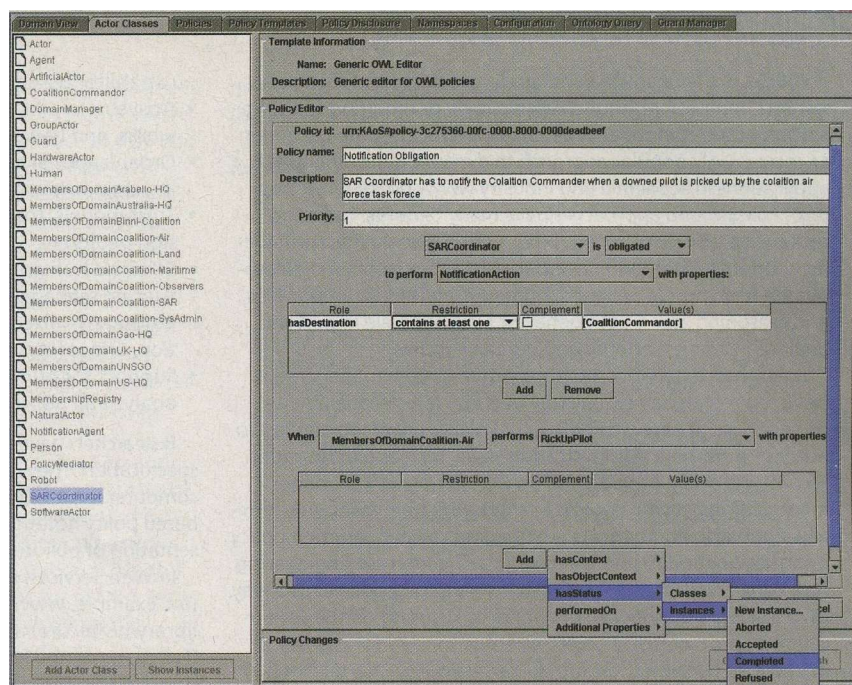
→ **Συνδεσιμότητα (Pluggability).** Μία οντολογία συγκεκριμένης πλατφόρμας και συγκεκριμένης εφαρμογής μπορεί εύκολα να φορτωθεί “πάνω από” τις βασικές έννοιες (core concepts). Επιπλέον, τα στοιχεία επιβολής πολιτικών έχουν προσαρμοστεί ώστε να συνεργάζονται με ένα εύρος υπολογιστικών περιβάλλοντων, τόσο παραδοσιακών καταμεμημένων πλατφορμών (πχ Web Services, grid computing, Corba), καθώς και διάφορων πλατφορμών λογισμικού και ρομποτικών πρακτόρων (πχ Nomads, Brahms, SFS, CoABS, Grid, Cougar).

→ **Επεκτασιμότητα και αποδοτικότητα (Scalability and performance).** Οι μέθοδοι κοινοποίησης της πολιτικής είναι βελτιστοποιημένες ώστε να αποκρίνονται σε ένα αίτημα που υποβάλλει κάποιος “επιβολέας” (enforcer) σε μέσο χρόνο λιγότερο από 1 ms. Επιπλέον, πολλαπλά αιτήματα που υποβάλλουν οι “υποβολείς” μπορούν να εκτελούνται ταυτόχρονα, επιτρέποντας στο KAOs να εκμεταλλευτεί τις δυνατότητες των συστημάτων πολυεπεξεργασίας.

Στην εικόνα 3-11 φαίνεται η γραφική διεπιφάνεια του συντάκτη και διαχειριστή πολιτικών του ΚΑoS. Οι πολιτικές δημιουργούνται με τη βοήθεια της τεχνολογίας OWL. Το εργαλείο ονομάζεται ΚΑoS Policy Administration Tool και παρέχει πολλές επιλογές στο χρήστη, όπως καθορισμό ονόματος πολιτικής, περιγραφής, προτεραιότητας, κτλ.

Περισσότερα σχετικά με αυτή την ενότητα

Περισσότερες εφαρμογές που χρησιμοποιούν/ορίζουν οντολογίες στα πλαίσια του σημασιολογικού ιστού αναφέρουν οι Susie Stephens [STEP06] και Peter Clark [CLAR05].



Εικόνα 3-11: Το εργαλείο ΚΑoS Policy Administration Tool.

4. Ανάλυση τεχνολογιών σημασιολογικού ιστού

Μέχρι τώρα έχει γίνει απλή αναφορά στις τεχνολογίες του σημασιολογικού ιστού, στα πλαίσια της ανάλυσης των επιπέδων της στοίβας της αρχιτεκτονικής του. Οι τεχνολογίες που αποτελούν τα βασικά συστατικά του σημασιολογικού ιστού είναι οι: XML, XML Schema, RDF, RDF Schema (RDFS), και OWL.

Σε αυτό το κεφάλαιο γίνεται ανάλυση των χαρακτηριστικών των παραπάνω τεχνολογιών, και του συντακτικού το οποίο χρησιμοποιείται για τη δημιουργία εγγράφων με χρήση αυτών. Δίνεται μεγαλύτερη έμφαση στις τεχνολογίες RDF, RDFS, και OWL, επειδή αυτές είναι οι τεχνολογίες που προσθέτουν σημασιολογία στους διαδικτυακούς πόρους (έγγραφα, βάσεις δεδομένων, εφαρμογές, κτλ), με αποτέλεσμα ο σημασιολογικός ιστός να είναι ποιοτικά καλύτερος από τον παραδοσιακό. Ένας επιπλέον λόγος για τον οποίο δεν δίνεται έμφαση στις τεχνολογίες XML και XML Schema, είναι ότι θεωρούνται περισσότερο τεχνολογίες γενικών εφαρμογών, δηλαδή δεν κατασκευάστηκαν αποκλειστικά για χρήση στον σημασιολογικό ιστό, σε αντίθεση με τις υπόλοιπες τρεις.

Σε καμία περίπτωση δεν πρέπει να θεωρηθεί ότι οι αναφορές που θα γίνουν στις τεχνολογίες του σημασιολογικού αυτού στα πλαίσια αυτού του κεφαλαίου, είναι ικανές να τις καλύψουν πλήρως, είτε σε θεωρητικό (πλεονεκτήματα/μειονεκτήματα, εφαρμογές, κτλ) είτε σε πρακτικό επίπεδο (πλήρες συντακτικό που υποστηρίζουν). Ο αναγνώστης που επιθυμεί να μάθει περισσότερα για αυτές τις τεχνολογίες, μπορεί να αναφερθεί στην προτεινόμενη βιβλιογραφία.

Στόχος αυτού του κεφαλαίου είναι να μπορεί ο αναγνώστης να αναφέρει τις περιπτώσεις που είναι κατάλληλη για χρήση η καθεμία τεχνολογία, να μπορεί να κατανοήσει το περιεχόμενο των εγγράφων που αναλύονται, και τέλος να είναι ικανός να δημιουργήσει δικά του έγγραφα με αντίστοιχο περιεχόμενο.

4.1 eXtensible Markup Language (XML)

Σύμφωνα με την Wikipedia [WIXM07], η XML είναι μία γενικευμένη γλώσσα σήμανσης που κατασκευάστηκε από το W3C, και προτείνεται για τη δημιουργία γλωσσών σήμανσης ειδικού σκοπού, που είναι ικανές να περιγράφουν πολλά διαφορετικά είδη δεδομένων. Με λίγα λόγια, η XML είναι ένας τρόπος περιγραφής δεδομένων, αλλά ένα αρχείο XML μπορεί επίσης να ενσωματώνει δεδομένα, παρόμοια με μία βάση δεδομένων. Αποτελεί ένα απλοποιημένο υποσύνολο της πρότυπης γενικευμένης γλώσσας σήμανσης (Standard Generalized Markup Language - SGML), η οποία είναι η μητέρα όλων των γλωσσών σήμανσης.

Ο κυριότερος σκοπός της XML είναι να διευκολύνει το διαμοιρασμό δεδομένων μεταξύ διαφορετικών συστημάτων, και πιο συγκεκριμένα συστημάτων που είναι συνδεδεμένα στο Διαδίκτυο. Οι γλώσσες που βασίζονται στην XML, για παράδειγμα η γλώσσα σήμανσης γεωγραφίας (Geography Markup Language - GML), η RDF/XML, το RSS, το Atom, η MathML, η XHTML, η SVG, το Klip, η MusicXML, κτλ. καθορίζονται με επίσημο τρόπο, τέτοιο ώστε να επιτρέπεται σε εφαρμογές να τροποποιούν και να ελέγχουν την εγκυρότητα των εγγράφων τους δίχως να γνωρίζουν τίποτα για τη μορφή τους (δηλαδή για τη δομή των εγγράφων που δημιουργούνται με αυτές τις γλώσσες).

4.1.1 Χαρακτηριστικά της eXtensible Markup Language

Σύμφωνα με την Wikipedia [WIXM07], η XML παρέχει έναν τρόπο αρχικά περιγραφής της πληροφορίας σε μορφή κειμένου και έπειτα εφαρμογής μίας δένδρικής δομής στην πληροφορία. Στο βασικό επίπεδο της XML, όλη η πληροφορία παρουσιάζεται ως κείμενο που αναμιγνύεται με σήμανση, ώστε να φαίνεται ο διαχωρισμός της πληροφορίας σε:

- μία ιεραρχία χαρακτήρων δεδομένων (character data),
- στοιχεία που μοιάζουν με τύπους περιεχομένων (container-like elements),
- ιδιότητες (attributes) αυτών των στοιχείων.

Η θεμελιώδης μονάδα της XML είναι ο χαρακτήρας, όπως καθορίζεται από το διεθνές σύνολο χαρακτήρων. Οι χαρακτήρες συνδυάζονται με συγκεκριμένους σειραϊκούς (serial)

συνδυασμούς ώστε να συνθέσουν ένα έγγραφο XML. Το έγγραφο αποτελείται από μία ή περισσότερες οντότητες (entities), καθεμία από τις οποίες είναι ένα τμήμα των χαρακτήρων του εγγράφου κωδικοποιημένο ως ακολουθία ψηφίων (bit sequence) και αποθηκευμένο σε ένα αρχείο κειμένου.

Τα αρχεία XML μπορούν να εξυπηρετούνται από πολλούς τύπους μέσων (media types). Το RFC3023 καθορίζει τους τύπους “application/xml” και “text/xml”, οι οποίοι δίνουν την πληροφορία ότι τα δεδομένα υπάρχουν εντός του αρχείου XML, αλλά δεν δίνουν καμία πληροφορία για τη σημασιολογία του. Η χρήση του “text/xml” έχει κατακριθεί ως η πιθανότερη πηγή δημιουργίας προβλημάτων κωδικοποίησης. Το πρότυπο RFC3023 καθορίζει επίσης ότι οι γλώσσες που βασίζονται στην XML θα πρέπει να εξυπηρετούν τύπους μέσων που να ξεκινούν με το πρόθεμα “application/” και να τερματίζουν με το επίθεμα “+xml”. Για παράδειγμα για την υπηρεσία Atom ο τύπος είναι “application/atom+xml”.

Η παρουσία αμέτρητων εφαρμογών επεξεργασίας κειμένου πλέον διευκολύνει τη γρήγορη δημιουργία και συντήρηση αρχείων XML, αν εκτιμηθεί ότι πριν την εμφάνιση της XML υπήρχαν ελάχιστες γλώσσες περιγραφής οι οποίες ήταν γενικής χρήσης, φιλικές με τα πρωτόκολλα του Διαδικτύου, και πολύ εύκολες στην εκμάθηση και στην σύνταξη. Εξάλλου, οι περισσότερες τεχνικές ανταλλαγής δεδομένων ήταν ιδιόκτητες, ειδικού σκοπού και δυαδικής φύσεως (βασισμένες σε ακολουθίες ψηφίων και όχι σε χαρακτήρες), έτσι δεν ήταν εύκολο να τις μοιραστούν οι διάφορες εφαρμογές λογισμικού ή οι διάφορες υπολογιστικές πλατφόρμες.

Επιτρέποντας τον καθορισμό ονομάτων, την δήλωση της επιτρεπόμενης ιεραρχίας και τον καθορισμό σημασιών για τα στοιχεία να πραγματοποιείται με τρόπο ανοιχτό και καθορισμένο από ένα εξατομικευμένο (customizable) σχήμα, η XML παρέχει τη συντακτική βάση για τη δημιουργία ατομικών γλωσσών σήμανσης βασισμένων σε αυτήν. Η γενική σύνταξη (syntax) τέτοιου είδους γλωσσών είναι αυστηρή - τα έγγραφα πρέπει να υπακούν στους γενικούς κανόνες της XML, διασφαλίζοντας ότι όλο το λογισμικό που την χρησιμοποιεί θα μπορεί το λιγότερο να τα διαβάσει (προσπελάσει) και να καταλάβει τον τρόπο που είναι δομημένη η πληροφορία. Το σχήμα συμπληρώνει τους συντακτικούς κανόνες ορίζοντας ένα σύνολο περιορισμών. Τα σχήματα περιορίζουν τα ονόματα των στοιχείων και των ιδιοτήτων, καθώς και τις επιτρεπόμενες ιεραρχίες που μπορούν να σχηματιστούν. Για παράδειγμα μπορεί να

οριστεί ότι ένα στοιχείο με όνομα “γενέθλια” θα έχει εντός του μόνο δύο στοιχεία. Ένα στοιχείο με όνομα “μήνας” και ένα στοιχείο με όνομα “ημέρα”, και οι τιμές που θα μπορούν να λάβουν τα δύο στοιχεία θα αποτελούνται μόνο από δεδομένα χαρακτήρων. Οι περιορισμοί ενός σχήματος μπορούν να περιέχουν και αναθέσεις τύπων δεδομένων, οι οποίοι επηρεάζουν τον τρόπο επεξεργασίας της πληροφορίας.

Από τα παραπάνω είναι εμφανές ότι η XML αντιτίθεται με την HTML, η οποία είναι μη ευέλικτη και χρησιμοποιεί ένα λεξιλόγιο στοιχείων και ιδιοτήτων που δεν μπορεί να επαναχρησιμοποιηθεί για πολλούς σκοπούς. Με την XML, είναι πολύ ευκολότερη η συγγραφή λογισμικού που μπορεί να προσπελάσει τις πληροφορίες ενός εγγράφου, εφόσον η δομή των δεδομένων εκφράζεται με έναν επίσημο και σχετικά απλό τρόπο.

Η XML δεν εισάγει περιορισμούς στον τρόπο με τον οποίο χρησιμοποιείται. Αν και είναι βασισμένη σε κείμενο (text-based), το λογισμικό που την χρησιμοποιεί συνήθως την ενσωματώνει σε πλουσιότερες μορφές (richer formats) δεδομένων με χρήση της αφαιρετικότητας (abstraction), κυρίως με τη χρήση σχημάτων που στρέφονται σε τύπους δεδομένων (datatype oriented) και αντικειμενοστραφών (object oriented) παραδειγμάτων (στα οποία το έγγραφο έχει το ρόλο του αντικειμένου).

4.1.2 Δυνατά σημεία και αδυναμίες της eXtensible Markup Language

Σύμφωνα με την Wikipedia [WIXM07], μερικά χαρακτηριστικά της XML που την καθιστούν κατάλληλη για τη μεταφορά δεδομένων είναι:

- ✓ είναι αναγνώσιμη από τον άνθρωπο και από τις μηχανές,
- ✓ υποστηρίζει το πρότυπο χαρακτήρων Unicode, επιτρέποντας τη χρήση σχεδόν οποιασδήποτε γλώσσας,
- ✓ είναι ικανή να αναπαραστήσει τις γενικότερες δομές δεδομένων της επιστήμης των Η/Υ: εγγραφές, λίστες και αρχεία,
- ✓ η μορφή της αυτό-τεκμηρίωσης που περιγράφει τη δομή και τα ονόματα των πεδίων, καθώς και καθορισμένες τιμές για αυτά,

- ✓ το ακριβές (strict) συντακτικό και οι απαιτήσεις προσπέλασης που επιτρέπουν στους αναγκαίους αλγορίθμους προσπέλασης να είναι απλοί, ευέλικτοι και ακριβείς.

Η XML χρησιμοποιείται επίσης και για αποθήκευση και επεξεργασία εγγράφων (document storage and processing), είτε άμεσα (online) είτε έμμεσα (offline), και προσφέρει αρκετά οφέλη:

- ✓ η λογικά επαληθεύσιμη μορφή της είναι βασισμένη σε διεθνή πρότυπα,
- ✓ η ιεραρχική δομή της είναι βολική για τους περισσότερους (αν όχι για όλους) τους τύπους εγγράφων,
- ✓ εκδίδεται σαν τα κοινά αρχεία κειμένου, ανεξάρτητη από άδειες ή περιορισμούς,
- ✓ είναι ανεξάρτητη από κάποια υπολογιστική πλατφόρμα (platform-independent) και συνεπώς δεν επηρεάζεται από τεχνολογικές αλλαγές,
- ✓ η πρόγονος της SGML χρησιμοποιείται από το 1986, κάτι το οποίο σημαίνει ότι υπάρχει μεγάλη εμπειρία και βοηθητικό λογισμικό για τις δύο αυτές τεχνολογίες.

Για ορισμένες εφαρμογές, η XML παρουσιάζει κάποιες αδυναμίες:

- × Η σύνταξη της είναι αρκετά φλύαρη και μερικώς περιττή. Αυτό μπορεί να βλάψει την αναγνωσιμότητα και την ευελιξία των εφαρμογών, και να δημιουργήσει υψηλότερες δαπάνες αποθηκευτικού χώρου. Μπορεί επίσης να κάνει δυσκολότερη τη χρήση της XML σε περιπτώσεις που το εύρος (bandwidth) είναι περιορισμένο, αν και η συμπίεση μπορεί να μειώσει μερικώς αυτό το πρόβλημα σε ορισμένες περιπτώσεις. Μία τέτοια περίπτωση είναι τα κινητά τηλέφωνα που χρησιμοποιούν πολυμεσικές εφαρμογές καθώς και οι προσωπικοί ψηφιακοί βοηθοί (Personal Digital Assistants - PDA's), οι οποίοι χρειάζεται να χρησιμοποιούν την XML για την περιγραφή των εικόνων και των βίντεο που προβάλλουν.
- × Οι εφαρμογές προσπέλασης πρέπει να σχεδιάζονται έτσι ώστε να υποστηρίζουν την αναδρομική προσπέλαση των εμφωλευμένων δομών δεδομένων και την εφαρμογή επιπλέον ελέγχων λανθασμένης σύνταξης ή σειράς τοποθέτησης των στοιχείων (αυτό απαιτείται επειδή η σήμανση είναι φλύαρη και μερικώς περιττή, όπως αναφέρθηκε παραπάνω). Κάτι τέτοιο προκαλεί σημαντικό φόρτο (overhead) σε αρκετές χρήσεις της XML, ειδικά όταν οι πόροι είναι περιορισμένοι - για παράδειγμα στα ενσωματωμένα συστήματα. Επιπλέον, θέματα που αφορούν την ασφάλεια προκύπτουν όταν τα δεδομένα εισόδου της XML προέρχονται από αναξιόπιστες πηγές, και είναι πιθανή η εξάντληση των

πόρων (resource exhaustion) και η δημιουργία εξαιρέσεων λόγω υπερχειλίσης της στοίβας (stack overflow).

- x Πολλοί πιστεύουν ότι το συντακτικό της XML περιέχει κάποια περιττά και δυσνόητα χαρακτηριστικά που οφείλονται στη συμβατότητα της με την SGML. Μία προσπάθεια δημιουργίας ενός υποσυνόλου που θα ονομάζονταν “Ελάχιστη XML” (Minimal XML) δεν στέφθηκε με επιτυχία, διότι τελικά δεν είναι καθόλου ξεκάθαρο ποια χαρακτηριστικά της XML θεωρούνται δυσνόητα ή περιττά.
- x Η βασική διαδικασία προσπέλασης δεν υποστηρίζει έναν μεγάλο αριθμό τύπων δεδομένων, με αποτέλεσμα να απαιτείται επιπλέον δουλειά σε αρκετές περιπτώσεις για την επεξεργασία των επιθυμητών δεδομένων. Λόγου χάριν, δεν υπάρχει τρόπος ώστε η XML να είναι ικανή να καταλάβει ότι το “3.14159” είναι αριθμός κινητής υποδιαστολής και όχι ένα αλφαριθμητικό επτά χαρακτήρων. Μερικές γλώσσες δημιουργίας σχημάτων προσθέτουν τέτοιου είδους υποστήριξη.
- x Χρησιμοποιεί το ιεραρχικό μοντέλο αναπαράστασης, το οποίο είναι περιορισμένο συγκριτικά με το σχεσιακό μοντέλο, αφού υποστηρίζει μόνο μία σταθερή απεικόνιση (fix view) της πληροφορίας.
- x Η μοντελοποίηση επικαλυπτόμενων (μη ιεραρχικών) δομών δεδομένων απαιτεί επιπλέον προσπάθεια.
- x Η συσχέτιση της XML με σχεσιακά ή αντικειμενοστραφή παραδείγματα είναι συχνά δύσκολη.
- x Μερικοί συμφωνούν ότι η XML μπορεί να χρησιμοποιηθεί μόνο σε αρχεία μικρού μεγέθους, αλλά αυτό αληθεύει μόνο αν ληφθούν υπόψιν κριτήρια όπως η αρχιτεκτονική, τα δεδομένα, η υλοποίηση και άλλα θέματα.
- x Η σειρά των πλήκτρων για τη δημιουργία εκφράσεων XML με τη χρήση ενός κοινού πληκτρολογίου είναι άβολη [CUNN06].

4.1.3 Παράδειγμα εγγράφου eXtensible Markup Language

Το παράδειγμα A-2 του παραρτήματος Α παρουσιάζει ένα έγγραφο XML. Το έγγραφο αποτελεί περιγραφή μίας υπενθύμισης (ένας άνθρωπος υπενθυμίζει σε κάποιον άλλο μία ερχόμενη συνάντηση).

4.2 eXtensible Markup Language Schema (XML Schema)

Σύμφωνα με την Wikipedia [WIXM06], ένα σχήμα (schema) XML είναι μία περιγραφή του τύπου ενός εγγράφου XML, η οποία εκφράζεται με ένα σύνολο περιορισμών ως προς τη δομή και το περιεχόμενο των εγγράφων αυτού του τύπου [79]. Πρακτικά πρόκειται για ένα σύνολο κανόνων που πρέπει να υπακούει ένα έγγραφο XML, ώστε να θεωρείται “έγκυρο” (valid), πάντα στα πλαίσια του συγκεκριμένου σχήματος που χρησιμοποιεί.

Υπάρχουν πολλές γλώσσες δημιουργίας σχημάτων XML. Ο καθορισμός τύπου εγγράφου (Document Type Definition – DTD) ο οποίος υπάρχει στην επίσημη προδιαγραφή της XML (XML specification), είναι μία γλώσσα δημιουργίας σχημάτων περιορισμένων δυνατοτήτων. Άλλες τεχνολογίες είναι η TREX [80], η Relax [WIRE06], η Schematron [WISC06], κτλ. Παρόλα αυτά, η πιο σημαντική και καταλληλότερη γλώσσα δημιουργίας σχημάτων είναι η W3C XML Schema. Ένα στιγμιότυπο (instance) της γλώσσας W3C XML Schema είναι ένα XML Schema Definition (XSD), τα αρχεία του οποίου συνηθίζεται να συνοδεύονται από την κατάληξη “.xsd”. Για το λόγο αυτό, η W3C XML Schema καταχρηστικά αποκαλείται πολύ συχνά λανθασμένα XSD, αν και το όνομα WXS⁸ (αρχικά των W3C XML Schema) θεωρείται πιο ακριβές.

4.2.1 Πότε χρησιμοποιείται η W3C XML Schema

Σε αντίθεση με τις υπόλοιπες γλώσσες δημιουργίας σχημάτων, η WXS σχεδιάστηκε με την πρόθεση τα έγκυρα αρχεία της να περιέχουν συλλογές πληροφοριών που σχετίζονται με συγκεκριμένους τύπους δεδομένων⁹, κάτι το οποίο είναι πολύ χρήσιμο για την ανάπτυξη λογισμικού επεξεργασίας εγγράφων XML, αλλά ταυτόχρονα δέχτηκε και πολλές κριτικές [WIXM06].

Σύμφωνα με την Wikipedia [WIXS06], η WXS χρησιμοποιείται για τον έλεγχο εγκυρότητας της δομής και των τύπων δεδομένων ενός εγγράφου XML. Η δομή συγκεκριμένα σχετίζεται με τον καθορισμό:

⁸ Στο εξής, θα χρησιμοποιείται το όνομα WXS ως συντομογραφία του W3C XML Schema.

⁹ Ένας πολύ χρήσιμος πίνακας με τους τύπους δεδομένων της WXS υπάρχει στο [WIXS06].

- των στοιχείων δεδομένων (data elements) του εγγράφου,
- της σειράς (order) με την οποία δηλώνονται τα στοιχεία δεδομένων,
- της εμφώλευσης (nesting) των στοιχείων δεδομένων,
- των προαιρετικών και υποχρεωτικών στοιχείων δεδομένων.

Η WXS είναι επίσης πολύ χρήσιμη για την γρήγορη εξαγωγή δεδομένων από βάσεις (databases) και την μεταφορά τους σε αρχεία XML.

4.2.2 Παράδειγμα εγγράφου W3C XML Schema

Το έγγραφο του παραδείγματος A-3 που υπάρχει στο παράρτημα Α, είναι ένα έγγραφο WXS που περιγράφει τη δομή ενός στοιχείου που χρησιμοποιείται για τη δήλωση χωρών. Στο παράδειγμα A-4, φαίνεται πώς ένα έγγραφο XML χρησιμοποιεί με σωστό τρόπο το σχήμα του παραδείγματος A-3. Αναλυτικότερα, στο έγγραφο A-3 περιγράφεται το στοιχείο `country`, το οποίο είναι τύπου `Country`. Ο τύπος `Country` είναι ένας σύνθετος τύπος δεδομένων, που αποτελείται από δύο ιδιότητες. Την ιδιότητα `name`, που είναι τύπου `string`, σύμφωνα με τους εγγενείς τύπους δεδομένων (builtin datatypes) της XML Schema, και την ιδιότητα `population` που είναι τύπου `decimal`, επίσης σύμφωνα με τους εγγενείς τύπους δεδομένων της XML Schema. Το έγγραφο A-4, χρησιμοποιεί το σχήμα του εγγράφου A-3, και δηλώνει το στοιχείο `country`, με τιμή “France” στην ιδιότητα `name`, και τιμή “59.7” στην ιδιότητα `population`.

4.3 Resource Description Framework (RDF)

Ο παγκόσμιος ιστός πληροφοριών προσφέρει απεριόριστη πρόσβαση σε πληροφορίες που βρίσκονται κατανεμημένες οπουδήποτε. Τα μεταδεδομένα, δηλαδή τα δομημένα δεδομένα που χρησιμοποιούνται για την περιγραφή άλλων δεδομένων, βελτιώνουν αισθητά την εύρεση και την πρόσβαση σε αυτές τις πληροφορίες. Παρόλα αυτά, η αποδοτική χρήση των μεταδεδομένων μεταξύ των εφαρμογών, απαιτεί περισσότερες συμβάσεις (conventions) σχετικά με τις σημασιολογίες (semantics), το συντακτικό (syntax), και τη δομή (structure). Οι

σημασιολογίες, δηλαδή οι έννοιες των μεταδεδομένων που καλύπτουν τις ιδιαίτερες ανάγκες ενός προβλήματος, καθορίζονται από κοινότητες περιγραφής πόρων (resource description communities) – δηλαδή ομάδες ανθρώπων (είτε ερευνητών είτε απλών χρηστών) που ασχολούνται με την περιγραφή πόρων. Το συντακτικό, δηλαδή η συστηματική τοποθέτηση των στοιχείων δεδομένων που επεξεργάζονται οι μηχανές (machine-processing), διευκολύνει την ανταλλαγή και χρήση των μεταδεδομένων μεταξύ πολλών εφαρμογών. Η δομή μπορεί να θεωρηθεί ως ο καθορισμός επίσημων περιορισμών στο συντακτικό, για τη συνεπή αναπαράσταση των σημασιολογιών [USZO04].

4.3.1 Που χρειάζεται ο Resource Description Framework

Σύμφωνα με τους Andrzej Uszok, Jeffrey M. Bradshaw, κ.α. [USZO04], και Eric Miller, [MILL98], ο σκελετός περιγραφής πόρων (Resource Description Framework - RDF) που έχει κατασκευαστεί από το W3C, παρέχει την κατάλληλη υποδομή που επιτρέπει την κωδικοποίηση, την ανταλλαγή, και την επαναχρησιμοποίηση των δομημένων μεταδεδομένων. Μερικά παραδείγματα μεταδεδομένων που μπορούν να αναπαρασταθούν με τον RDF είναι ο τίτλος, ο συγγραφέας, και η ημερομηνία τροποποίησης μιας ιστοσελίδας, πληροφορίες που αφορούν τα πνευματικά δικαιώματα και την άδεια χρήσης ενός Διαδικτυακού εγγράφου, ή ακόμη η διαθεσιμότητα ενός διαμοιραζόμενου Διαδικτυακού πόρου, κτλ. Αυτή η υποδομή επιτρέπει τη διαλειτουργικότητα (interoperability) των μεταδεδομένων μέσω της σχεδίασης μηχανισμών που υποστηρίζουν κοινές έννοιες στις σημασιολογίες, στο συντακτικό, και στη δομή. Ο RDF δεν ορίζει σημασιολογίες για κάθε κοινότητα περιγραφής πόρων, αντίθετα παρέχει την ικανότητα στις κοινότητες να μπορούν να καθορίσουν τις έννοιες των μεταδεδομένων τους όπως αυτές επιθυμούν.

Ο RDF δε σχετίζεται μόνο με την κωδικοποίηση μεταδεδομένων. Γενικά χρησιμοποιείται για την αναπαράσταση πληροφοριών που σχετίζονται με "Διαδικτυακούς πόρους" (Web resources). Αυτό σημαίνει ότι ο RDF μπορεί να χρησιμοποιηθεί και για την αναπαράσταση πληροφοριών που σχετίζονται με αντικείμενα τα οποία μπορούν να αναγνωριστούν στον σημασιολογικό ιστό, ακόμη και αν δεν μπορούν να ανακτηθούν άμεσα. Παραδείγματα αποτελούν οι πληροφορίες σχετικά με προϊόντα που προσφέρονται από ηλεκτρονικά καταστήματα μέσω του Διαδικτύου (on-line shopping facilities) - πχ πληροφορίες σχετικά με

προδιαγραφές, τιμές, διαθεσιμότητα, κτλ - ή πχ η περιγραφή των προτιμήσεων ενός χρήστη του σημασιολογικού ιστού για τη λήψη πληροφοριών (ηλεκτρονική αλληλογραφία, παραδοσιακή αλληλογραφία, τηλεφωνία, κτλ), κτλ.

Ο RDF προορίζεται για χρήση σε καταστάσεις όπου οι πληροφορίες πρέπει να υποστούν επεξεργασία από εφαρμογές, και όχι απλά να απεικονιστούν στους χρήστες. Ο RDF παρέχει ένα κοινό πλαίσιο (common framework) για την έκφραση πληροφοριών έτσι ώστε να μπορούν να ανταλλάσσονται από τις εφαρμογές, δίχως να αλλοιώνεται η σημασία τους. Επειδή το πλαίσιο του είναι κοινό, οι σχεδιαστές εφαρμογών μπορούν να ενισχύσουν τη διαθεσιμότητα των εφαρμογών προσπέλασης και επεξεργασίας εγγράφων RDF. Η ικανότητα ανταλλαγής πληροφοριών μεταξύ διαφορετικών εφαρμογών, σημαίνει ότι η πληροφορία μπορεί να γίνει διαθέσιμη σε εφαρμογές τελείως διαφορετικές από αυτές για τις οποίες δημιουργήθηκε. Αυτό εισάγει το πλεονέκτημα της επαναχρησιμοποίησης της πληροφορίας.

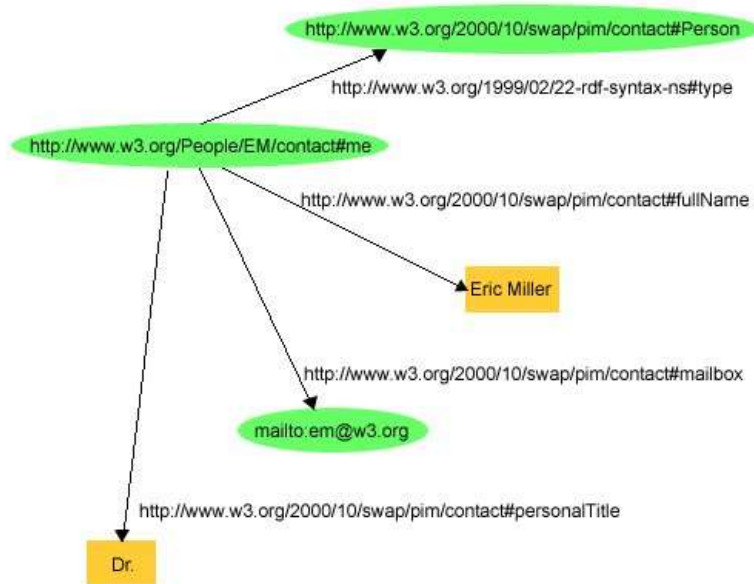
Ο RDF υποστηρίζει τη δήλωση συμβάσεων οι οποίες διευκολύνουν την διαλειτουργικότητα μεταξύ ξεχωριστών συνόλων στοιχείων μεταδεδομένων. Αυτές οι συμβάσεις περιέχουν πρότυπους μηχανισμούς για την αναπαράσταση σημασιολογιών οι οποίες στηρίζονται σε ένα απλό αλλά ταυτόχρονα δυναμικό μοντέλο δεδομένων, που θα περιγραφεί στην αμέσως επόμενη ενότητα. Επιπρόσθετα, με τη βοήθεια του RDF μπορούν να δημιουργούνται λεξιλόγια (vocabularies) που είναι αναγνώσιμα και από τους ανθρώπους (human-readable), και από τις μηχανές (machine-readable). Ένα λεξιλόγιο είναι ένα σύνολο ιδιοτήτων, ή στοιχείων μεταδεδομένων, που καθορίζονται από τις κοινότητες περιγραφής πόρων [MANO04].

4.3.2 Το μοντέλο δεδομένων του Resource Description Framework

Σύμφωνα με τους Joshua Tauberer [TAUB06] και Eric Miller [MILL98], και Frank Manola, Eric Miller, κ.α. [MANO04], το Διαδίκτυο αποτελεί ένα αποκεντρωμένο (decentralized) σύστημα, στο οποίο οι πληροφορίες είναι κατανεμημένες (distributed) διάσπαρτα και ανομοιόμορφα. Για να λειτουργήσει καλά ένα τέτοιο σύστημα, επιβάλλεται η χρήση κοινά αποδεκτών

προτύπων (standards), έτσι ώστε να διευκολύνεται η κατασκευή εφαρμογών που θα λειτουργούν στα πλαίσια αυτού του συστήματος. Ένα πρότυπο που σχετίζεται με την κατανομημένη γνώση του σημασιολογικού ιστού, πρέπει να λαμβάνει υπόψιν του τα εξής:

- Τα αρχεία του σημασιολογικού ιστού πρέπει να εκφράζουν την πληροφορία με τρόπο διαισθητικό και έτσι ώστε να τροποποιείται εύκολα. Δεν είναι υποχρεωτικό, η γνώση να είναι τακτοποιημένη ούτε σε πίνακες όπως ισχύει στις σχεσιακές βάσεις δεδομένων, ούτε σε ιεραρχίες όπως ισχύει στην XML. Ένας απλούστερος και ευκολότερα κατανοητός τρόπος απεικόνισης είναι με τη μορφή γραφήματος. Ο RDF παρέχει ένα μοντέλο για την περιγραφή πόρων. Οι πόροι περιγράφονται στα πλαίσια ιδιοτήτων (properties) οι οποίες λαμβάνουν τιμές (property values). Ο RDF καθορίζει ένα πόρο ως ένα οποιοδήποτε αντικείμενο που προσδιορίζεται μοναδικά (uniquely identifiable) με τη βοήθεια των προσδιοριστών πόρων (γνωστοί ως URI/IRI)¹⁰. Αυτό του επιτρέπει να αναπαριστά απλές δηλώσεις που αφορούν τους πόρους ως ένα γράφημα με κόμβους και ακμές. Για παράδειγμα, η πρόταση “υπάρχει ένας άνθρωπος, ο οποίος αναγνωρίζεται από το <http://www.w3.org/People/EM/contact#me>, του οποίου το όνομα είναι Eric Miller, η ηλεκτρονική αλληλογραφία του είναι το em@w3.org, και η ιδιότητα του είναι Δρ.”, μπορεί να αναπαρασταθεί με το γράφημα RDF της εικόνας 4-1.



Εικόνα 4-1: Το γράφημα RDF για την περιγραφή των στοιχείων του Eric Miller.

¹⁰ Εκτός από την ονομασία URI/IRI, οι προσδιοριστές πόρων είναι γνωστοί και ως URIs όταν χρησιμοποιούν αποκλειστικά URI για τον προσδιορισμό πόρων.

➤ Τα αρχεία του σημασιολογικού ιστού πρέπει να μπορούν να σχετίζονται μεταξύ τους. Έστω ότι υπάρχει ένα αρχείο που περιέχει τιμές προϊόντων και έχει δημιουργηθεί από κάποιον πωλητή (vendor), και ένα αρχείο με σχολιασμούς των ίδιων προϊόντων (product reviews) που έχει δημιουργηθεί ανεξάρτητα από έναν καταναλωτή. Πρέπει να υπάρχει κάποιος τρόπος ένδειξης ότι οι δύο άνθρωποι αναφέρονται στα ίδια προϊόντα. Η ονομασία των προϊόντων δεν αρκεί, επειδή είτε μπορεί να διαφέρει αλλά τα προϊόντα να είναι ουσιαστικά τα ίδια, είτε μπορεί να είναι όμοια αλλά τα προϊόντα να είναι τελείως διαφορετικά. Επομένως με κάποιο τρόπο πρέπει να πιστοποιήσουμε ότι δεν υπάρχει ασάφεια στις πληροφορίες που υπάρχουν στον σημασιολογικό ιστό, ώστε να μπορούν οι μηχανές να τις επεξεργάζονται με βεβαιότητα. Όλα αυτά τα προβλήματα, λύνονται με τους προσδιοριστές πόρων. Στο γράφημα της εικόνας 4-1, οι προσδιοριστές πόρων χρησιμοποιούνται για τον προσδιορισμό:

- ατόμων (individuals), πχ ο Eric Miller, που προσδιορίζεται από το <http://www.w3.org/People/EM/contact#me>,
- ειδών αντικειμένων (kind of things), πχ άτομο (Person), που προσδιορίζεται από το <http://www.w3.org/2000/10/swap/pim/contact#Person>,
- ιδιοτήτων αντικειμένων (properties of things), πχ ταχυδρομικό κουτί (mailbox), που προσδιορίζεται από το <http://www.w3.org/2000/10/swap/pim/contact#mailbox>,
- τιμές των ιδιοτήτων (values of properties), πχ <mailto:em@w3.org> ως η τιμή της ιδιότητας mailbox.

Ο RDF χρησιμοποιεί επίσης και αλφαριθμητικά ως ιδιότητες, πχ "Eric Miller", καθώς και τιμές άλλων τύπων δεδομένων όπως οι ακέραιοι αριθμοί, οι ημερομηνίες, κτλ, ως τιμές των ιδιοτήτων.

➤ Η χρήση λεξιλογίων για τη δημιουργία δηλώσεων σχετικά με τα αντικείμενα είναι πολύ σημαντική, αλλά τα λεξιλόγια πρέπει να έχουν τη δυνατότητα να αναμιγνύονται. Πχ ένα λεξιλόγιο σχετικά με τηλεοπτικές σειρές που δημιουργήθηκε από μία ομάδα τηλεθεατών, και ένα λεξιλόγιο σχετικά με κινηματογραφικές ταινίες που δημιουργήθηκε από μία κριτική επιτροπή ανεξάρτητα, θα πρέπει να μπορούν να συνδυαστούν σε ένα κοινό αρχείο, ώστε

να είναι δυνατή η αναφορά στα ίδια αντικείμενα (πχ σε έναν ηθοποιό που έχει συμμετάσχει και σε τηλεοπτικές σειρές και σε κινηματογραφικές ταινίες).

Όλες οι παραπάνω προϋποθέσεις πληρούνται από το πρότυπο του RDF, καθιστώντας τον ως μία από τις βασικότερες τεχνολογίες του σημασιολογικού ιστού.

4.3.3 Αναπαράσταση τριάδων Resource Description Framework σε γράφημα

Έχει γίνει ήδη αναφορά στις τριάδες με τις οποίες ο RDF εκφράζει τη σημασία, στο πρώτο κεφάλαιο. Κάθε τριάδα αποτελείται από το υποκείμενο (subject), το κατηγορημα (predicate), και το αντικείμενο (object). Σε αυτό το σημείο αξίζει να γίνει μία αναφορά στον τρόπο με τον οποίο απεικονίζονται οι τριάδες RDF σε ένα γράφημα. Στο γράφημα της εικόνας 4-1, η τριάδα που περιγράφει την ιδιότητα του Eric Miller είναι η εξής:

ο <http://www.w3.org/People/EM/contact#me> έχει την ιδιότητα του Δρ.

Σύμφωνα με τους Frank Manola, Eric Miller, κ.α. [MANO04], αυτή η τριάδα RDF μπορεί να αναπαρασταθεί με:

- υποκείμενο το <http://www.w3.org/People/EM/contact#me>,
- κατηγορημα το <http://www.w3.org/2000/10/swap/pim/contact#personalTitle>
- αντικείμενο το Δρ.

Όπως παρατηρείται, δεν είναι υποχρεωτικό κάθε στοιχείο μίας τριάδας να συνοδεύεται από έναν προσδιοριστή πόρων. Αυτό είναι απόλυτα λογικό, αφού όπως αναφέρθηκε ήδη, μια ιδιότητα μπορεί να έχει ως τιμή κάποιο αλφαριθμητικό. Παρόλα αυτά, στην απεικόνιση του γραφήματος, όλες οι δηλώσεις παριστάνονται υποχρεωτικά με:

- ένα κόμβο (node) για το υποκείμενο,
- ένα κόμβο για το αντικείμενο,
- μία ακμή (edge) για το κατηγορημα, κατευθυνόμενη από τον κόμβο του υποκειμένου προς τον κόμβο του αντικειμένου.

Τα παραπάνω μπορούν να παρατηρηθούν και στο γράφημα της εικόνας 4-1.

4.3.4 Το συντακτικό του Resource Description Framework

Στο πρώτο κεφάλαιο, έγινε αναφορά σε μία ευελιξία που παρέχει ο RDF. Ο RDF υπάρχει σε αφαιρετικό επίπεδο “πάνω από” άλλες τεχνολογίες, με αποτέλεσμα να μπορεί να γράφεται με πολλούς τρόπους σύνταξης.

Σύμφωνα με τους Frank Manola, Eric Miller, κ.α [MANO04], και Joshua Tauberer [TAUB06], ένας όχι πολύ γνωστός, αλλά ιδιαίτερα χρήσιμος τρόπος σύνταξης του RDF, λόγω της μεγάλης εκφραστικότητας που προσφέρει, αλλά και της ομοιότητας με τις τριάδες (υποκείμενο, κατηγορημα, αντικείμενο) των γραφημάτων αφού μετατραπούν σε μορφή όπως η παρακάτω, ονομάζεται “σημειογραφία 3” (Notation 3):

ο <http://www.w3.org/People/EM/contact#me> έχει την ιδιότητα του Δρ.

Το παράδειγμα A-5 του παραρτήματος Α, αποτελεί την κωδικοποίηση του γραφήματος της εικόνας 4-1, με σύνταξη σημειογραφίας 3. Αν και αυτός ο τρόπος σύνταξης μοιάζει περισσότερο με όσα απεικονίζονται στο γράφημα, γενικά δεν είναι διαισθητικός (intuitive) προς τους ανθρώπους. Για το λόγο αυτό δεν είναι ιδιαίτερα διάσημος, οπότε δε θα δοθεί μεγάλη έμφαση στο συντακτικό που χρησιμοποιεί.

Ο πιο γνωστός από τους τρόπους σύνταξης έχει την ονομασία RDF/XML, επειδή χρησιμοποιεί τη δένδρική σύνταξη της XML για την περιγραφή των τριάδων RDF. Αυτός ο τρόπος είναι περισσότερο οικείος προς τους ανθρώπους, μιας και η XML είναι ευρέως γνωστή στο Διαδίκτυο, αφού η γενικότητα που προσφέρει την καθιστά χρήσιμη για πάρα πολλούς σκοπούς. Αξίζει να τονιστεί ότι η μετατροπή ενός εγγράφου από σημειογραφία 3 σε RDF/XML και το αντίστροφο επιτυγχάνεται εύκολα, και υπάρχουν πολλά εργαλεία που κάνουν αυτή τη δουλειά αυτόματα. Το παράδειγμα A-6 παρουσιάζει την κωδικοποίηση του γραφήματος της εικόνας 4-1, με σύνταξη RDF/XML. Παρακάτω ακολουθεί επεξήγηση του συντακτικού RDF/XML στα πλαίσια του παραδείγματος A-6. Απαιτείται καλή εξοικείωση με τις τεχνολογίες XML, XML Schema, καθώς και με τους ονοματοχώρους της XML (XML namespaces) για την κατανόηση του παραδείγματος.

Η πρώτη γραμμή θεωρείται γνωστή. Είναι η οδηγία (instruction) που καθορίζει την έκδοση της XML (XML declaration) στις εφαρμογές προσπέλασης του εγγράφου:

```
<?xml version="1.0"?>
```

Στις επόμενες δύο γραμμές γίνεται χρήση των ονοματοχώρων της XML. Ο πρώτος ονοματοχώρος κάνει χρήση του λεγόμενου “γνωστού προθέματος” (“well-known” QName prefix) για να καθορίσει τον ονοματοχώρο του RDF που χρησιμοποιείται για όλο το έγγραφο, ώστε να χρειάζεται να επαναλαμβάνει μόνο το πρόθεμα, με αποτέλεσμα ο κώδικας να είναι πιο κομψός και ευανάγνωστος. Ο ονοματοχώρος του RDF είναι υποχρεωτικό να δηλώνεται στην αρχή όλων των εγγράφων RDF, εντός του υποχρεωτικού στοιχείου `rdf:RDF`. Ο δεύτερος ονοματοχώρος (με πρόθεμα `contact`) χρησιμοποιείται για να προσδιορίσει μοναδικά τα στοιχεία που δηλώνονται στο έγγραφο RDF:

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:contact="http://www.w3.org/2000/10/swap/pim/contact#">
```

Η επόμενη γραμμή χρησιμοποιεί μία δεσμευμένη ιδιότητα του RDF, την ιδιότητα `rdf:about`. Ένας προφανής τρόπος για το χαρακτηρισμό οποιασδήποτε δήλωσης (statement) RDF, είναι να θεωρηθεί ότι είναι μία περιγραφή (*description*), η οποία σχετίζεται (*is about*) με το θέμα της δήλωσης. Με αυτό τον τρόπο γίνεται η αναπαράσταση των δηλώσεων στην RDF/XML. Το πεδίο `rdf:Description` είναι επίσης δεσμευμένο και σηματοδοτεί την έναρξη της περιγραφής ενός πόρου. Στη συνέχεια, σε συνδυασμό με την ιδιότητα `rdf:about`, προσδιορίζει το υποκείμενο του πόρου με τη βοήθεια ενός προσδιοριστή πόρων. Συνήθως το δεσμευμένο πεδίο `rdf:Description` συνυπάρχει με την δεσμευμένη ιδιότητα `rdf:about`, αν και αυτό δεν είναι απαραίτητο (λόγου χάριν στο παράδειγμα που περιγράφεται δε συμβαίνει). Συνοψίζοντας, η παρακάτω γραμμή προσδιορίζει μοναδικά το υποκείμενο `Person`:

```
<contact:Person
  rdf:about="http://www.w3.org/People/EM/contact#me">
```

Η επόμενη γραμμή μπορεί να θεωρηθεί ως μία τυπική γραμμή XML, αφού δεν περιέχει

κανένα ιδιαίτερο στοιχείο του RDF.

```
<contact:fullName>Eric Miller</contact:fullName>
```

Η παρακάτω γραμμή είναι ενδιαφέρουσα, επειδή χρησιμοποιεί τη δεσμευμένη ιδιότητα `rdf:resource`. Αυτή η ιδιότητα δηλώνεται εντός ενός κενού στοιχείου XML (empty element), δηλαδή του στοιχείου που συνδυάζει τα πεδία έναρξης και τερματισμού σε ένα πεδίο (που τερματίζεται με `>`). Τα κενά στοιχεία δεν έχουν περιεχόμενο (content), δηλαδή δε μεσολαβεί τίποτα εντός των πεδίων έναρξης και τερματισμού τους. Στη συγκεκριμένη περίπτωση, απαιτείται το ηλεκτρονικό ταχυδρομικό κουτί (mailbox) του Eric Miller να συνοδεύεται από έναν προσδιοριστή πόρων, συγκεκριμένα το `mailto:`, ώστε να είναι δυνατή η άμεση χρήση της υπηρεσίας (ηλεκτρονικής αλληλογραφίας) με την προσπέλαση του εγγράφου. Η ιδιότητα `rdf:resource` διαφέρει από την ιδιότητα `rdf:about`, επειδή η τελευταία χρησιμοποιείται μόνο για να καθορίσει έναν προσδιοριστή που οδηγεί σε πόρο που σχετίζεται με ότι περιγράφεται στο έγγραφο. Στη συγκεκριμένη περίπτωση περιγράφεται ο Eric Miller, άρα ο προσδιοριστής πρέπει να οδηγεί σε έναν σύνδεσμο που δίνει πρόσβαση σε προσωπικές πληροφορίες που αφορούν τον Eric Miller. Αντίθετα, η ιδιότητα `rdf:resource` χρησιμοποιείται για να καθορίσει γενικότερα ζεύγη υποκειμένων - προσδιοριστών.

Υπάρχει επίσης κάτι σημαντικό που είναι χρήσιμο να αναφερθεί. Η σύνταξη RDF/XML παρέχει τη δυνατότητα αντί για τη χρήση της ιδιότητας `rdf:resource`, να γίνει χρήση ενός τυπικού στοιχείου XML, το οποίο θα έχει ως περιεχόμενο τον προσδιοριστή `mailto:em@w3.org` παρόμοια με το περιεχόμενο του στοιχείου `<contact:fullName>`, πχ:

```
<contact:mailbox>mailto:em@w3.org</contact:mailbox>
```

Σε αυτή την περίπτωση όμως, το έγγραφο δε θα διερμηνεύεται όπως θα έπρεπε, επειδή ο προσδιοριστής θα διερμηνεύεται ως απλό αλφαριθμητικό, και δε θα είναι δυνατή η χρήση της υπηρεσίας αποστολής ηλεκτρονικής αλληλογραφίας άμεσα.

Οι τρεις τελευταίες γραμμές είναι επίσης τυπικά στοιχεία της XML. Η πρώτη περιγράφει την ιδιότητα (`personalTitle`) του Eric Miller, και οι δύο τελευταίες τερματίζουν με έγκυρο τρόπο τα στοιχεία `<contact:Person>` και `<rdf:RDF>`, ώστε το έγγραφο να θεωρείται καλοσχηματισμένο (well-formed) [WIXM07]:

```
<contact:personalTitle>Dr.</contact:personalTitle>
</contact:Person>
</rdf:RDF>
```

Η σύνταξη RDF/XML βέβαια δεν περιορίζεται σε όσα αναφέρθηκαν στα πλαίσια του παραδείγματος A-6. Υποστηρίζει επίσης το σύνολο των εγγενών τύπων δεδομένων της XML Schema. Μία πολύ καλή αναφορά σε αυτούς τους τύπους δεδομένων δίνεται από τους Paul V. Biron και Ashok Malhotra [BIRO01], και παραδείγματα χρήσεως τους από τους Frank Manola, Eric Miller, κ.α. [MANO04].

4.4 Resource Description Framework Schema (RDFS)

Σύμφωνα με τους Frank Manola, Eric Miller, κ.α. [MANO04], ο RDF παρέχει έναν τρόπο για την έκφραση απλουστευμένων δηλώσεων (ή ισχυρισμών)¹¹ σχετικά με Διαδικτυακούς πόρους, με χρήση ιδιοτήτων και τιμών. Παρόλα αυτά, οι κοινότητες χρηστών του RDF πρέπει να έχουν τη δυνατότητα να καθορίζουν λεξιλόγια (όρων) που τείνουν να χρησιμοποιούν στις δηλώσεις των εγγράφων τους. Συγκεκριμένα, πρέπει να μπορούν να περιγράφουν επιλεγμένα είδη ή κλάσεις πόρων, και να χρησιμοποιούν επιλεγμένες ιδιότητες στην περιγραφή αυτών των πόρων.

Για παράδειγμα, οι άνθρωποι που ενδιαφέρονται για την περιγραφή βιβλιογραφικών πόρων, θέλουν να περιγράφουν κλάσεις όπως οι `bib:Book` και `bib:MagazineArticle`, και να χρησιμοποιούν ιδιότητες όπως οι `bib:author`, `bib:title`, και `bib:subject` για την περιγραφή τους. Άλλου είδους εφαρμογές ενδέχεται να χρειαστεί να περιγράψουν κλάσεις όπως οι `world:Person` και `organ:company`, και ιδιότητες όπως οι `world:age`,

¹¹ Στο εξής, οι λέξεις δήλωση και ισχυρισμός θα χρησιμοποιούνται δίχως κανένα διαχωρισμό, για να δηλώσουν τον τρόπο έκφρασης που χρησιμοποιεί ο RDF για την περιγραφή πληροφοριών.

`world:jobTitle`, `organ:stockSymbol`, και `organ:numberOfEmployees`. Ο RDF δεν παρέχει κανένα τρόπο για την περιγραφή τέτοιου είδους κλάσεων εφαρμογών (application-specific) και ιδιοτήτων τους. Αντίθετα, τέτοιου είδους κλάσεις και ιδιότητες περιγράφονται ως ένα λεξιλόγιο RDF, επεκτείνοντας τον RDF με τη γλώσσα περιγραφής λεξιλογίων RDF (RDF Vocabulary Description Language), ευρύτερα γνωστή ως RDF Schema (RDFS).

Η RDFS δεν παρέχει ένα λεξιλόγιο με έτοιμες κλάσεις εφαρμογών όπως οι `bib:Book`, `world:Person`, `furn:Sofa`, και ιδιότητες όπως οι `bib:author`, `world:age`, `furn:type`. Αντίθετα, παρέχει υποστήριξη για την περιγραφή τέτοιων κλάσεων και ιδιοτήτων, και για την ένδειξη των κλάσεων και ιδιοτήτων που συνδυάζονται (πχ για να οριστεί ότι η ιδιότητα `world:jobTitle` χρησιμοποιείται κατά την περιγραφή του `world:Person`). Με άλλα λόγια, η RDFS παρέχει ένα σύστημα τύπων (type system) στον RDF. Το σύστημα τύπων της RDFS είναι κατά κάποιο τρόπο παρόμοιο με τα συστήματα τύπων των αντικειμενοστραφών γλωσσών προγραμματισμού (Java, Smalltalk, κτλ). Λόγου χάριν, η RDFS επιτρέπει να καθορίζονται οι πόροι ως στιγμιότυπα μίας ή περισσότερων κλάσεων. Επιπρόσθετα, επιτρέπει οι κλάσεις να οργανώνονται με ιεραρχικό τρόπο. Για παράδειγμα μία κλάση `an:Dog` μπορεί να καθοριστεί ως υποκλάση της `an:Mammal`, η οποία πιθανόν να είναι υποκλάση της `an:Animal`, με αποτέλεσμα όποιος πόρος είναι δηλωμένος εντός της κλάσης `an:Dog` να είναι έμμεσα (implicitly) και εντός της κλάσης `an:Animal` (εξαιτίας της ιεραρχίας). Παρόλα αυτά, οι ιδιότητες και κλάσεις της RDFS παρουσιάζουν εμφανείς διαφορές σε ορισμένα ζητήματα, αν συγκριθούν με τις κλασικές αντικειμενοστραφείς γλώσσες προγραμματισμού¹².

Τα χαρακτηριστικά της RDFS παρέχονται με τη μορφή ενός λεξιλογίου RDF, δηλαδή ως ένα εξειδικευμένο σύνολο προκαθορισμένων πόρων RDF με τις δικές τους εξειδικευμένες σημασίες. Οι πόροι εντός του λεξιλογίου της RDFS χρησιμοποιούν προσδιοριστές πόρων με το πρόθεμα <http://www.w3.org/2000/01/rdf-schema#> (το οποίο κατά σύμβαση σχετίζεται με το αποδεκτό πρόθεμα `rdfs:`). Οι περιγραφές λεξιλογίων (σχήματα – schemas) που γράφονται στη γλώσσα RDFS αποτελούν έγκυρα γραφήματα RDF. Τουτέστιν, το λογισμικό RDF που δεν έχει γραφεί ώστε να επεξεργάζεται έγγραφα με λεξιλόγιο RDF

¹² Οι κλάσεις και ιδιότητες της RDFS δε δημιουργούν μία αλυσίδα (straightjacket) εντός της οποίας πρέπει να τοποθετηθεί η πληροφορία. Αντίθετα, παρέχουν επιπρόσθετη πληροφορία για τους πόρους RDF που περιγράφουν. Αυτή η πληροφορία μπορεί να χρησιμοποιηθεί με πολλούς τρόπους. Μία καλή αναφορά στις διαφορές μεταξύ των κλάσεων RDFS και των αντικειμενοστραφών κλάσεων υπάρχει στο [MANO04].

Schema, μπορεί να διερμηνεύσει ένα schema ως ένα έγκυρο γράφημα RDF που περιέχει διάφορους πόρους και ιδιότητες, αλλά δεν μπορεί να “κατανοήσει” τις επιπλέον σημασίες των όρων της RDF Schema. Για την κατανόηση των επιπλέον σημασιών, το λογισμικό RDF πρέπει να γραφεί ώστε να επεξεργάζεται μία εκτεταμένη γλώσσα που δεν περιέχει μόνο το λεξιλόγιο `rdf:`, αλλά και το λεξιλόγιο `rdfs:`, και τα δύο με τους ενσωματωμένους (built-in) μηχανισμούς τους. Ένα τέτοιο παράδειγμα θα δοθεί στην επόμενη ενότητα.

4.4.1 Περιγράφοντας κλάσεις στην Resource Description Framework Schema

Σύμφωνα με Frank Manola, Eric Miller, κ.α. [MANO04], ένα βασικό βήμα σε κάθε τρόπο περιγραφής, είναι η αναγνώριση των διαφόρων ειδών αντικειμένων που πρόκειται να περιγραφούν. Η RDFS αναφέρεται σε αυτά τα “είδη αντικειμένων” (kinds of things) ως κλάσεις. Μία κλάση στην RDFS αντιστοιχεί στην γενική έννοια του *τύπου* ή της *κατηγορίας*, παρόμοια κατά κάποιο τρόπο με την έννοια της κλάσης σε αντικειμενοστραφείς γλώσσες προγραμματισμού, όπως η Java. Οι κλάσεις RDF μπορούν να χρησιμοποιηθούν για να αναπαραστήσουν σχεδόν οποιαδήποτε κατηγορία αντικειμένου, όπως ιστοσελίδες, ανθρώπους, τύπους εγγράφων, βάσεις δεδομένων, ή αφηρημένες έννοιες. Οι κλάσεις περιγράφονται χρησιμοποιώντας τους πόρους της RDFS `rdfs:Class` και `rdfs:Resource`, και τις ιδιότητες `rdf:type`¹³ και `rdfs:subClassOf`.

Για παράδειγμα, έστω ότι ο οργανισμός `example.org` θέλει να χρησιμοποιήσει τον RDF για να παρέχει πληροφορίες για διάφορα είδη μηχανοκίνητων οχημάτων (motor vehicles). Στην RDFS, ο οργανισμός χρειάζεται αρχικά μία κλάση που θα αναπαριστά την κατηγορία των αντικειμένων που είναι μηχανοκίνητα οχήματα. Οι πόροι που ανήκουν σε μία κλάση ονομάζονται *στιγμιότυπα* (*instances*) της. Στην περίπτωση που παρουσιάζεται, ο `example.org` φροντίζει ώστε τα στιγμιότυπα αυτής της κλάσης να είναι πόροι που αφορούν μηχανοκίνητα οχήματα.

¹³ Όταν ένας πόρος RDF περιγράφεται από την ιδιότητα `rdf:type`, η τιμή αυτής της ιδιότητας θεωρείται πόρος που αναπαριστά μία κατηγορία ή *κλάση* αντικειμένων. Το υποκείμενο αυτής της ιδιότητας θεωρείται ένα *στιγμιότυπο* (*instance*) της κατηγορίας ή κλάσης που αυτή αναπαριστά [MANO04].

Στην RDFS, κλάση είναι οποιοσδήποτε πόρος έχει την ιδιότητα `rdf:type`, τιμή της οποίας είναι ο πόρος `rdfs:Class`. Άρα η κλάση για τα μηχανοκίνητα οχήματα μπορεί να περιγραφεί αν της ανατεθεί ένας προσδιοριστής πόρων, πχ ο `veh:MotorVehicle` (χρησιμοποιώντας το πρόθεμα `veh:` για τους πόρους του λεξιλογίου του `example.org`, με το `veh:` να δείχνει στο `http://www.example.org/schemas/vehicles`) και ο πόρος αυτός περιγραφεί με την ιδιότητα `rdf:type`, τιμή της οποίας θα είναι ο πόρος `rdfs:Class`. Δηλαδή, ο οργανισμός `example.org`, πρέπει να δημιουργήσει μία δήλωση RDF παρόμοια με την παρακάτω:

```
veh:MotorVehicle    rdf:type    rdfs:Class .
```

Ο παραπάνω, είναι ένας ακόμη απλός και κατανοητός τρόπος αναπαράστασης των τριάδων RDF, και ονομάζεται “σημειογραφία στηλών” (tabular notation). Όπως έχει αναφερθεί ήδη, η ιδιότητα `rdf:type` χρησιμοποιείται ώστε να “δείξει” ότι κάποιος πόρος είναι στιγμιότυπο μίας κλάσης. Άρα, έχοντας περιγράψει τον όρο `veh:MotorVehicle` ως κλάση, για την περιγραφή του πόρου `vehthings:companyCar` ως μηχανοκίνητο όχημα, απαιτείται μία δήλωση παρόμοια με την παρακάτω¹⁴:

```
vehthings:companyCar    rdf:type    rdfs:MotorVehicle .
```

Η παραπάνω δήλωση φανερώνει ότι ο οργανισμός `example.org` αποφάσισε να ορίσει διαφορετικά λεξιλόγια για την περιγραφή των κλάσεων αντικειμένων (`veh:`) και των στιγμιότυπων αντικειμένων (`vehthings:`) αντίστοιχα.

Έπειτα από την περιγραφή της κλάσης `vehthings:MotorVehicle`, ο οργανισμός `example.org` θέλει να περιγράψει επιπλέον κλάσεις, οι οποίες αναπαριστούν είδη μηχανοκίνητων οχημάτων, πχ επιβατικά οχήματα (passenger vehicles), φορτηγά (vans), μικρά φορτηγά (minivans), κοκ. Αυτές οι κλάσεις μπορούν να καθοριστούν ανάλογα, δηλαδή με ανάθεση ενός προσδιοριστή πόρων για κάθε νέα κλάση, και την περιγραφή των πόρων ως κλάσεις με τη βοήθεια του RDF, πχ:

¹⁴ Στα παραδείγματα που παρουσιάζονται, έχει χρησιμοποιηθεί η συνήθης σύμβαση όλα τα ονόματα των κλάσεων να ξεκινούν με κεφαλαίο χαρακτήρα, ενώ τα ονόματα των στιγμιότυπων, των ιδιοτήτων, και των τύπων δεδομένων με πεζά, με κάθε νέα λέξη που υπάρχει στο όνομα ενός στιγμιότυπου να ξεκινά με κεφαλαίο χαρακτήρα. Αν και η RDFS δεν παρουσιάζει ασυμφωνία μεταξύ πεζών και κεφαλαίων χαρακτήρων, προτείνεται η υιοθέτηση κάποιου από τους συνήθεις τρόπους ονοματοδοσίας.

```
veh:Van      rdf:type    rdfs:Class .  
veh:Truck    rdf:type    rdfs:Class .
```

Βέβαια αυτή η αναπαράσταση δεν είναι αρκετή, επειδή ναι μεν καθορίζει ότι ο πόροι `veh:Van` και `veh:Truck` είναι κλάσεις, αλλά δεν καθορίζει πουθενά ότι αποτελούν εξειδικευμένα (specialized) είδη (*kinds*) μηχανοκίνητων οχημάτων.

Αυτό το είδος της σχέσης ειδίκευσης μεταξύ δύο κλάσεων περιγράφεται με την ιδιότητα `rdfs:subClassOf`. Για παράδειγμα, για να δηλωθεί ότι το `veh:Van` είναι ένα εξειδικευμένο είδος `veh:MotorVehicle`, η δήλωση RDF είναι η εξής:

```
veh:Van      rdfs:subClassOf    veh:MotorVehicle .
```

Με της χρήση της ιδιότητας `rdfs:subClassOf` στην παραπάνω δήλωση, κάθε στιγμιότυπο της κλάσης `veh:Van`, είναι επίσης στιγμιότυπο της κλάσης `veh:MotorVehicle`. Άρα αν υποθεθεί ότι ο πόρος `vehthings:companyVan` είναι ένα στιγμιότυπο της κλάσης `veh:Van`, και έχει γραφεί λογισμικό RDF που κατανοεί το λεξιλόγιο της RDFS, τότε το λογισμικό αυτό βασιζόμενο στη σχέση `rdfs:subClassOf`, μπορεί να συμπεράνει επίσης ότι το `vehthings:companyVan` είναι και στιγμιότυπο της κλάσης `veh:MotorVehicle`.

Τα παραπάνω αποδεικνύουν στην πράξη την επεκτασιμότητα που παρέχει η γλώσσα RDFS στον σκελετό RDF. Σε περίπτωση που δεν έχει γραφεί λογισμικό RDF ώστε να είναι κατανοητό το λεξιλόγιο της RDFS, το αρχείο θα είναι έγκυρο και η σχέση `rdfs:subClassOf` θα αποτελεί ένα ακόμη κατηγορήμα στις τριάδες RDF, δίχως να είναι κατανοητή η ειδική σημασία του. Αυτό θα έχει ως αποτέλεσμα να μην μπορούν να εξαχθούν συμπεράσματα σχετικά με τις ιεραρχίες κλάσεων που δηλώνονται.

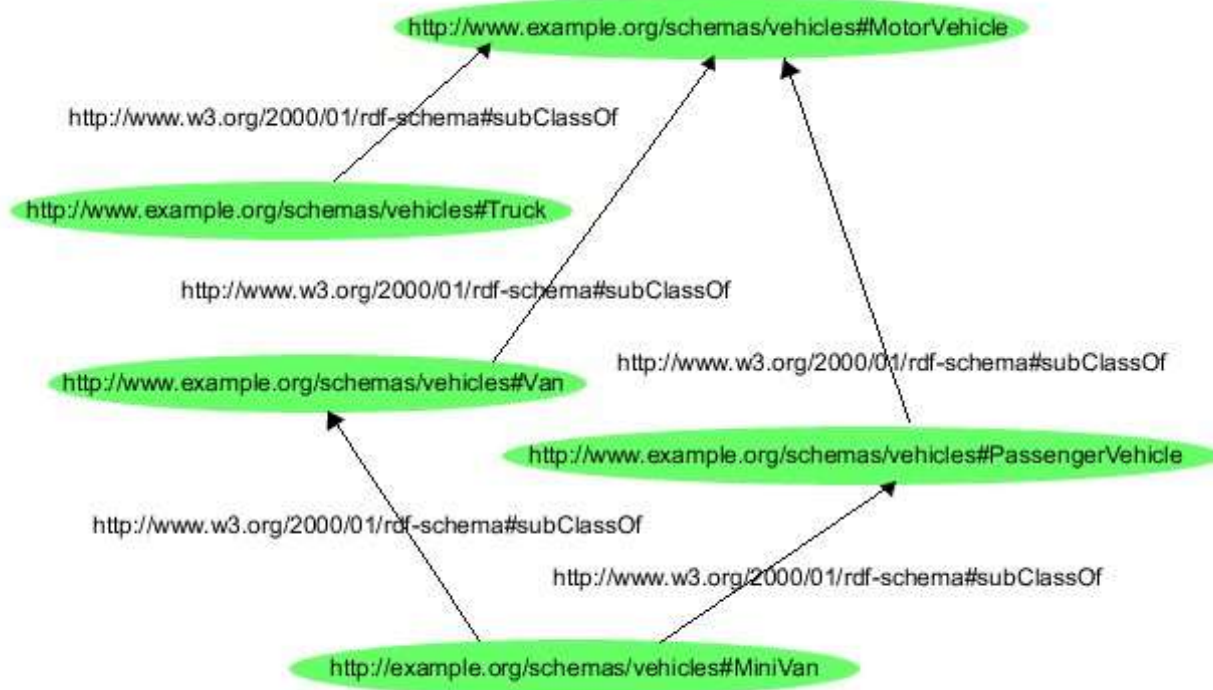
Σημαντικό είναι να ξεκαθαριστεί ότι η ιδιότητα `rdfs:subClassOf` είναι *μεταβατική* (*transitive*). Αυτό συνεπάγεται ότι μία κλάση μπορεί να είναι υποκλάση πολλών (υπέρ) κλάσεων. Στον αντικειμενοστραφή προγραμματισμό, το χαρακτηριστικό αυτό είναι γνωστό με την ονομασία “πολλαπλή κληρονομικότητα” (multiple inheritance)¹⁵. Η RDFS ορίζει όλες τις κλάσεις ως υποκλάσεις της κλάσης `rdfs:Resource` (εφόσον τα στιγμιότυπα όλων των

¹⁵ Οι απόψεις για τα πλεονεκτήματα και τα μειονεκτήματα που εισάγει η πολλαπλή κληρονομικότητα δίστανται. Περισσότερες πληροφορίες σχετικά με αυτό το θέμα υπάρχουν στο [WIMU06].

κλάσεων είναι πόροι). Ένα παράδειγμα πολλαπλής κληρονομικότητας είναι το παρακάτω, στο οποίο η κλάση `veh:Minivan` δηλώνεται ως υποκλάση της `veh:Van` και της `veh:PassengerVehicle`:

```
veh:Minivan    rdfs:subClassOf    veh:Van .  
veh:Minivan    rdfs:subClassOf    veh:PassengerVehicle .
```

Στην εικόνα 4-2 φαίνεται η πλήρης ιεραρχία όλων όσων συζητήθηκαν σε αυτή την ενότητα. Στην εικόνα φαίνεται γραφικά η πολλαπλή κληρονομικότητα και τα χαρακτηριστικά για τα οποία μπορούν να εξαχθούν συμπεράσματα που αφορούν τη δομή της ιεραρχίας των κλάσεων. Για απλοποίηση του γραφήματος, από την εικόνα 4-2 έχει παραλειφθεί η ιδιότητα `rdf:type` που σχετίζεται με την ιδιότητα `rdf:Class` για καθεμία κλάση.



Εικόνα 4-2: Ένα γράφημα ιεραρχίας κλάσεων οχημάτων.

Το γράφημα της εικόνας 4-2 μπορεί να περιγραφεί εύκολα με σημειογραφία στηλών. Η περιγραφή αυτή παρουσιάζεται στο παράδειγμα A-7 του παραρτήματος Α. Το παράδειγμα A-7 μπορεί με λίγη προσπάθεια να μετασχηματιστεί σε σύνταξη RDF/XML (παράδειγμα A-8) που θα υποστηρίζει τη γλώσσα RDFS.

Το παράδειγμα A-8 παρουσιάζει μία πολύ χρήσιμη ιδιότητα του RDF, στην οποία δεν έχει γίνει αναφορά, την ιδιότητα `rdf:ID`. Αυτή η ιδιότητα γενικά χρησιμοποιείται για τον καθορισμό σχετικών προσδιοριστών πόρων (relative URIs), η έννοια των οποίων είναι παρόμοια με αυτή της “σχετικής διαδρομής” (relative path) στους καταλόγους ενός λειτουργικού συστήματος. Πχ στο παράδειγμα A-8 υπάρχει ο σχετικός προσδιοριστής `#MotorVehicle`. Με την προϋπόθεση ότι ο προσδιοριστής αυτού του εγγράφου (base URI)¹⁶ είναι ο `http://example.org/schemas/vehicles`, ο πλήρης προσδιοριστής της κλάσης είναι ο `http://example.org/schemas/vehicles#MotorVehicle`. Οι σχετικοί προσδιοριστές χρησιμοποιούνται ως συντομεύσεις (abbreviations) για την αποφυγή επαναλαμβανόμενων μακροσκελών δηλώσεων.

4.4.2 Περιγράφοντας ιδιότητες στην Resource Description Framework Schema

Σύμφωνα με τους Frank Manola, Eric Miller, κ.α. [MANO04], και την Wikipedia [WIRD06], οι κοινότητες χρηστών, εκτός από την περιγραφή συγκεκριμένων κλάσεων, συχνά χρειάζεται να περιγράψουν και συγκεκριμένες *ιδιότητες* (*properties*) που χαρακτηρίζουν αυτές τις κλάσεις (πχ η ιδιότητα `age` για την περιγραφή της ηλικίας ενός ανθρώπου). Στην RDFS, οι ιδιότητες περιγράφονται με χρήση της κλάσης του RDF `rdf:Property`, και των ιδιοτήτων της RDFS `rdfs:domain`, `rdfs:range`, και `rdfs:subPropertyOf`.

Όλες οι ιδιότητες του RDF περιγράφονται ως στιγμιότυπα της κλάσης `rdf:Property`. Πχ η ιδιότητα `externs:weightInKg` μπορεί να περιγραφεί ορίζοντας της έναν προσδιοριστή, και περιγράφοντας την με χρήση της ιδιότητας `rdf:type`, η τιμή της οποίας είναι ο πόρος

¹⁶ Ο βασικός προσδιοριστής (base URI) ενός εγγράφου συνήθως προσδιορίζει το ίδιο το έγγραφο (document URI). Περιγράφεται με τη δήλωση `xml:base`, και γενικά θεωρείται πολύ πρακτικό να χρησιμοποιείται [MANO04].

`rdf:Property`, ως εξής:

```
exterms:weightInKg    rdf:type    rdf:Property .
```

Η RDFS παρέχει επίσης λεξιλόγιο για την περιγραφή του τρόπου με τον οποίο οι κλάσεις και οι ιδιότητες προορίζονται ώστε να χρησιμοποιούνται μαζί. Η πιο σημαντική πληροφορία αυτού του είδους παρέχεται μέσω των ιδιοτήτων `rdfs:range` και `rdfs:domain`.

Η ιδιότητα `rdfs:range` δηλώνει ότι οι τιμές μίας συγκεκριμένης ιδιότητας πρέπει να είναι στιγμιότυπα μίας συγκεκριμένης κλάσης. Για παράδειγμα, αν ο οργανισμός `example.org` θέλει να δηλώσει ότι η ιδιότητα `ex:author` πρέπει να λαμβάνει τιμές που είναι στιγμιότυπα της κλάσης `ex:Person`, πρέπει να γράψει τις παρακάτω δηλώσεις:

```
ex:Person    rdf:type    rdfs:Class .
ex:author    rdf:type    rdf:Property .
ex:author    rdfs:range  ex:Person .
```

Αυτές οι τρεις δηλώσεις δείχνουν ότι:

- η `ex:Person` είναι κλάση,
- η `ex:author` είναι ιδιότητα,
- οι δηλώσεις RDF (τριάδες) που χρησιμοποιούν την ιδιότητα `ex:author` πρέπει να λαμβάνουν τιμές που να είναι στιγμιότυπα της κλάσης `ex:Person` ως αντικείμενα (objects).

Μία ιδιότητα, πχ η `ex:hasMother`, μπορεί να έχει καμία, μία, ή και περισσότερες ιδιότητες `range`. Εάν η `ex:hasMother` δεν έχει καμία `range`, τότε δεν υπάρχει κανένας περιορισμός στις τιμές που μπορεί να λάβει. Αν έχει μία, πχ την `ex:Person`, τότε αυτό σημαίνει ότι οι τιμές της θα πρέπει να είναι στιγμιότυπα της `ex:Person`. Αν έχει παραπάνω από μία `range`, αυτό σημαίνει ότι οι τιμές της πρέπει να είναι πόροι που αποτελούν στιγμιότυπα όλων των κλάσεων που έχουν οριστεί με την ιδιότητα `range`. Ένα παράδειγμα για την τελευταία περίπτωση είναι να οριστεί ότι οι τιμές της ιδιότητας `ex:hasMother` είναι στιγμιότυπα και της κλάσης `ex:Female` και της κλάσης `ex:Person`. Είναι προφανές ότι αυτές οι δύο αναθέσεις απαιτούν δύο ξεχωριστές δηλώσεις:

```
ex:hasMother    rdfs:range    ex:Female .
ex:hasMother    rdfs:range    ex:Person .
```

Έστω ότι υπάρχει η παρακάτω δήλωση:

```
exstaff:frank    ex:hasMother    exstaff:frances .
```

Για να είναι έγκυρες οι δύο δηλώσεις `rdfs:range` του προηγούμενου παραδείγματος, πρέπει η `exstaff:frances` να είναι στιγμιότυπο και της κλάσης `ex:Female` και της κλάσης `ex:Person`.

Η ιδιότητα `rdfs:range` μπορεί να χρησιμοποιηθεί και για τον καθορισμό της τιμής που μπορεί να λάβει μία ιδιότητα, σύμφωνα με τους τύπους δεδομένων της WXS. Πχ έστω ότι ο οργανισμός `example.org` θέλει να καθορίσει ότι η ιδιότητα `ex:age` θα δέχεται τιμές που είναι συμβατές με τον τύπο δεδομένων της WXS `xsd:integer`. Οι δηλώσεις που χρειάζονται είναι οι εξής:

```
ex:age    rdf:type        rdf:Property .
ex:age    rdfs:range      xsd:integer .
```

Φυσικά οι τύποι δεδομένων της WXS έχουν με τη σειρά τους συγκεκριμένους προσδιοριστές πόρων, πχ ο `xsd:integer` προσδιορίζεται από το <http://www.w3.org/2001/XMLSchema#integer>.

Συνοψίζοντας, η ιδιότητα `rdfs:range` καθορίζει την κλάση ή τον τύπο δεδομένων του αντικειμένου (object) της τριάδας της ιδιότητας που τη χρησιμοποιεί ως κατηγορημα (predicate).

Η ιδιότητα `rdfs:domain` χρησιμοποιείται για να καθοριστεί ότι μία συγκεκριμένη ιδιότητα ισχύει για μία ορισμένη κλάση. Πχ έστω ότι ο `example.org` θέλει να καθορίσει ότι η ιδιότητα `ex:author` ισχύει για στιγμιότυπα της κλάσης `ex:Book`. Οι δηλώσεις που θα πρέπει να γραφούν είναι οι:

```
ex:Book      rdf:type      rdfs:Class .
ex:author    rdf:type      rdf:Property .
ex:author    rdfs:domain    ex:Book .
```

Αυτές οι τρεις δηλώσεις δείχνουν ότι:

- η `ex:Book` είναι κλάση,
- η `ex:author` είναι ιδιότητα,
- οι δηλώσεις RDF (τριάδες) που χρησιμοποιούν την ιδιότητα `ex:author` πρέπει να λαμβάνουν τιμές που να είναι στιγμιότυπα της κλάσης `ex:Book` ως υποκείμενα (subjects).

Παρόμοια με όσα ισχύουν για την `range`, μία δεδομένη ιδιότητα, πχ η `externs:weight`, μπορεί να έχει καμία, μία, ή και περισσότερες ιδιότητες `domain`. Εάν η `externs:weight` δεν έχει καμία `domain`, τότε δεν υπάρχει κανένας περιορισμός στις τιμές που μπορεί να λάβει. Εάν έχει μία, πχ την `ex:Book`, αυτό σημαίνει ότι η ιδιότητα `externs:weight` ισχύει για στιγμιότυπα της κλάσης `ex:Book`. Αν έχει παραπάνω από μία `domain`, αυτό σημαίνει ότι οι τιμές της πρέπει να είναι πόροι που αποτελούν στιγμιότυπα όλων των κλάσεων που έχουν οριστεί με την ιδιότητα `domain`. Ένα παράδειγμα για την τελευταία περίπτωση είναι να οριστεί ότι οι τιμές της ιδιότητας `externs:weight` είναι στιγμιότυπα και της κλάσης `ex:Book` και της κλάσης `ex:MotorVehicle`. Είναι προφανές ότι αυτές οι δύο αναθέσεις απαιτούν δύο ξεχωριστές δηλώσεις:

```
externs:weight  rdfs:domain  ex:Book .
externs:weight  rdfs:domain  ex:MotorVehicle .
```

Έστω ότι υπάρχει η παρακάτω δήλωση:

```
exthings:companyCar  externs:weight  "2500"^^xsd:integer .
```

Για να είναι έγκυρες οι δύο δηλώσεις `rdfs:domain` του προηγούμενου παραδείγματος, πρέπει η `exthings:companyCar` να είναι στιγμιότυπο και της κλάσης `ex:Book` και της κλάσης `ex:MotorVehicle`.

Συνοψίζοντας, η ιδιότητα `rdfs:domain` καθορίζει την κλάση του υποκειμένου (subject) της τριάδας της ιδιότητας που τη χρησιμοποιεί ως κατηγορημα (predicate).

Η χρήση των ιδιοτήτων `range` και `domain` μπορεί να παρουσιαστεί επεκτείνοντας το σχήμα των οχημάτων (παράδειγμα A-8), με την προσθήκη των ιδιοτήτων `ex:registeredTo` και `ex:rearSeatLegRoom`¹⁷, της κλάσης `ex:Person`, και περιγράφοντας αποκλειστικά τον `xsd:integer` ως τύπο δεδομένων¹⁸. Η ιδιότητα `ex:registeredTo` ισχύει για κάθε στιγμιότυπο της κλάσης `ex:MotorVehicle` (υποκείμενο), με τιμή στιγμιότυπο της `ex:Person` (αντικείμενο). Η ιδιότητα `ex:rearSeatLegRoom` ισχύει μόνο για στιγμιότυπα της κλάσης `ex:PassengerVehicle`, με τιμή `xsd:integer`, που δίνει τα εκατοστά του διαθέσιμου χώρου στο πίσω κάθισμα. Στο παράδειγμα B-1 υπάρχουν μόνο οι δηλώσεις των ιδιοτήτων που πρέπει να προστεθούν στο σχήμα του παραδείγματος A-8 για να επεκταθεί (δηλαδή το παράδειγμα B-1 δεν αποτελεί πλήρες σχήμα).

Η RDFS παρέχει επίσης έναν τρόπο για την εξειδίκευση των ιδιοτήτων και των κλάσεων. Αυτή η σχέση ειδίκευσης μεταξύ δύο ιδιοτήτων περιγράφεται με την προκαθορισμένη ιδιότητα `rdfs:subPropertyOf`. Πχ αν οι `ex:primaryDriver` και `ex:driver` είναι ιδιότητες, ο οργανισμός `example.org` είναι καλό να τις περιγράψει λαμβάνοντας υπόψιν ότι η ιδιότητα `ex:primaryDriver` είναι εξειδίκευση της ιδιότητας `ex:driver`. Οι δηλώσεις που απαιτούνται είναι οι:

<code>ex:driver</code>	<code>rdf:type</code>	<code>rdf:Property</code> .
<code>ex:primaryDriver</code>	<code>rdf:type</code>	<code>rdf:Property</code> .
<code>ex:primaryDriver</code>	<code>rdfs:subPropertyOf</code>	<code>ex:driver</code> .

Η σημασία της σχέσης `rdfs:subPropertyOf` στις παραπάνω δηλώσεις εξηγείται ως εξής: αν ο `exstaff:fred` είναι ένας `ex:primaryDriver` του `ex:companyVan`, τότε η RDFS καθορίζει τον `exstaff:fred` και ως `ex:driver` του `ex:companyVan`. Οι τελευταίες ιδιότητες που προστίθενται στο παράδειγμα A-8 για να σχηματιστεί το πλήρες σχήμα RDFS του παραδείγματος A-9, υπάρχουν στο παράδειγμα B-2.

Μία ιδιότητα μπορεί να είναι υπό-ιδιότητα (subproperty) καμίας, μίας, ή περισσότερων

¹⁷ Η συντόμευση προέρχεται από τις λέξεις `rear seat legroom`, που είναι ο διαθέσιμος χώρος για κάποιον που κάθεται πίσω από κάποιο κάθισμα [MANO04].

¹⁸ Αν και η αποκλειστική περιγραφή ενός τύπου δεδομένων της WXS σε ένα έγγραφο RDF δεν είναι υποχρεωτική, γενικά είναι καλή πρακτική και προτείνεται από το W3C [MANO04].

ιδιοτήτων. Όλες οι ιδιότητες `rdfs:range` και `rdfs:domain` ισχύουν για μία ιδιότητα ισχύουν αυτόματα και για όλες τις υπό-ιδιότητες της. Επομένως, στο παράδειγμα B-2, η RDFS ορίζει ότι η ιδιότητα `ex:primaryDriver` έχει επίσης ως `rdfs:domain` την κλάση `ex:MotorVehicle`, εξαιτίας της υπό-ιδιότητας που τη συσχετίζει με την ιδιότητα `ex:driver`.

Υπάρχουν και άλλες ενσωματωμένες βοηθητικές ιδιότητες που παρέχει η RDFS, όπως η `rdfs:comment` για την παροχή της περιγραφής ενός πόρου με τρόπο εύκολα αναγνώσιμο από τους ανθρώπους (human-readable), η `rdfs:seeAlso` για την αναφορά σε κάποιον πόρο που παρέχει επιπλέον πληροφορίες για τον πόρο του υποκειμένου (subject), κτλ. Περισσότερες πληροφορίες για αυτές τις ιδιότητες δίνονται από τους Dan Brickley και R.V. Guha [BRIC04].

4.5 Web Ontology Language (OWL)

Η RDFS παρέχει ορισμένες δυνατότητες για την περιγραφή λεξιλογίων RDF, οι οποίες όμως δεν είναι αρκετές. Επιπλέον δυνατότητες μπορούν να υπάρξουν, και αποδεικνύονται ιδιαίτερα χρήσιμες. Μερικές από αυτές (που δεν παρέχονται από την RDFS) είναι [MANO04]:

- Περιορισμός εύρους τιμών (cardinality constraints) στις ιδιότητες, πχ ένας άνθρωπος (Person) έχει ακριβώς ένα (και μόνο ένα) βιολογικό πατέρα.
- Ορισμός μίας δεδομένης ιδιότητας (όπως η `ex:hasAncestor`) ως μεταβατικής (transitive), πχ εάν `A ex:hasAncestor B` και `B ex:hasAncestor C`, τότε `A ex:hasAncestor C`.
- Ορισμός μίας δεδομένης ιδιότητας ως μοναδικό προσδιοριστή (unique identifier) ή κλειδί (key) για στιγμιότυπα μίας δεδομένης κλάσης.
- Ορισμός ότι δύο διαφορετικές κλάσεις (δηλαδή με διαφορετικούς προσδιοριστές) αντιπροσωπεύουν ουσιαστικά την ίδια κλάση.
- Ορισμός ότι δύο διαφορετικά στιγμιότυπα (δηλαδή με διαφορετικούς προσδιοριστές) αντιπροσωπεύουν ουσιαστικά το ίδιο άτομο (individual).

- Περιορισμός του εύρους τιμών μίας ιδιότητας, εξαρτημένος από την κλάση του πόρου στον οποίο αυτή εφαρμόζεται. Για παράδειγμα, να υπάρχει η δυνατότητα να δηλωθεί ότι η ιδιότητα `ex:hasPlayers` ισούται με έντεκα όταν περιγράφεται μία ποδοσφαιρική ομάδα, ενώ ισούται με πέντε όταν περιγράφεται μία ομάδα καλαθοσφαίρισης.
- Ικανότητα περιγραφής νέων κλάσεων χρησιμοποιώντας συνδυασμούς (πχ ενώσεις και τομές) άλλων κλάσεων, καθώς και ικανότητα ορισμού δύο κλάσεων ως διακριτές (disjoint). Πχ να δίνεται η δυνατότητα να απαγορευτεί οποιοσδήποτε πόρος να είναι στιγμιότυπο δύο κλάσεων που έχουν χαρακτηριστεί διακριτές.

Όλες οι παραπάνω ικανότητες (καθώς και άλλες), παρέχονται από τις διάφορες γλώσσες δημιουργίας οντολογιών. Η πιο πρόσφατη από αυτές (και καταλληλότερη για τη δημιουργία οντολογιών ιστού σύμφωνα με το W3C), είναι η γλώσσα δημιουργίας οντολογιών ιστού (Web Ontology Language – OWL). Η OWL έχει σχεδιαστεί για χρήση από εφαρμογές που χρειάζεται να επεξεργάζονται το πληροφοριακό περιεχόμενο των εγγράφων, σε αντίθεση με τις περιπτώσεις που απαιτείται απλά παρουσίαση του περιεχομένου στους ανθρώπους. Η OWL μπορεί να χρησιμοποιηθεί για την περιγραφή της σημασίας των όρων που τοποθετούνται σε λεξιλόγια, και των σχέσεων μεταξύ αυτών των όρων. Αυτή η αναπαράσταση των όρων και των σχέσεων που τους διέπουν ονομάζεται οντολογία. Η OWL έχει περισσότερες δυνατότητες για την έκφραση των σημασιών και των σημασιολογιών από τις τεχνολογίες XML, RDF και RDFS, και επομένως μπορεί να περιγράψει ακόμη καλύτερα το Διαδικτυακό περιεχόμενο που διερμηνεύουν οι μηχανές (machine interpretable). Η OWL είναι μία αναθεώρηση της DAML+OIL, που ενσωματώνει παθήματα που έγιναν μαθήματα κατά τη σχεδίαση και την εφαρμογή της τελευταίας [MCGU04].

4.5.1 Οι τρεις υπογλώσσες της Web Ontology Language

Σύμφωνα με τους Deborah L. McGuinness και Frank van Harmelen [MCGU04], η OWL παρέχει τρεις υπογλώσσες με αυξανόμενες δυνατότητες έκφρασης (increasingly expressive sublanguages)¹⁹, οι οποίες έχουν σχεδιαστεί για χρήση από συγκεκριμένες κοινότητες χρηστών:

- Η **OWL Lite** απευθύνεται σε χρήστες που χρειάζονται κυρίως μία ταξινομημένη ιεραρχία

¹⁹ Αυτό σημαίνει ότι η OWL Lite παρέχει μικρότερες δυνατότητες έκφρασης από την OWL DL, και αντίστοιχα η OWL DL παρέχει μικρότερες δυνατότητες έκφρασης από την OWL Full.

(classification hierarchy) και απλούς περιορισμούς. Λόγου χάριν, αν και η OWL Lite υποστηρίζει τους περιορισμούς εύρους τιμών (cardinality constraints), οι επιτρεπτές τιμές του εύρους είναι μόνο το μηδέν ή το ένα. Η δημιουργία ενός εργαλείου που θα υποστηρίζει την OWL Lite είναι πολύ ευκολότερη από ένα αντίστοιχο με υποστήριξη των δύο άλλων υπογλωσσών, που προσφέρουν μεγαλύτερες δυνατότητες έκφρασης. Η OWL Lite παρέχει ένα γρήγορο τρόπο δημιουργίας θησαυρών (thesauri) και άλλων ταξινομιών (taxonomies). Έχει επίσης μικρότερη πολυπλοκότητα από την OWL DL.

- Η **OWL DL** απευθύνεται σε χρήστες που επιθυμούν τη μέγιστη εκφραστικότητα, διατηρώντας ταυτόχρονα την υπολογιστική πληρότητα (computational completeness) – όλα τα συμπεράσματα είναι σίγουρα υπολογίσιμα (computable), και αποφασιστικότητα (decidability) – όλοι οι υπολογισμοί ολοκληρώνονται σε πεπερασμένο χρονικό διάστημα. Αν και η OWL DL περιλαμβάνει όλα τα γλωσσικά κατασκευάσματα (language constructs) της OWL, αυτά μπορούν να χρησιμοποιηθούν μόνο υπό συγκεκριμένους περιορισμούς (πχ ενώ μία κλάση μπορεί να είναι υποκλάση άλλων κλάσεων, μία κλάση δεν μπορεί να είναι στιγμιότυπο μίας άλλης κλάσης). Το όνομα της OWL DL προκύπτει από την ομοιότητα της με την περιγραφική λογική (description logics)²⁰, ένα ερευνητικό πεδίο που έχει μελετήσει τη λογική στην οποία στηρίζεται η OWL.
- Η **OWL Full** απευθύνεται σε χρήστες που επιθυμούν τη μέγιστη εκφραστικότητα και ελευθερία σύνταξης του RDF, δίχως να εξασφαλίζονται υπολογιστικές εγγυήσεις (computational guarantees). Πχ μία κλάση στην OWL Full μπορεί να αντιμετωπιστεί ταυτόχρονα ως μία συλλογή ατόμων και ως ένα αυτόνομο άτομο (με δικά του δικαιώματα). Η OWL Full επιτρέπει σε μία οντολογία να εμπλουτίσει (augment) τη σημασία του προκαθορισμένου (RDF ή OWL) λεξιλογίου. Είναι γενικά ασυνήθιστο ένα οποιοδήποτε σύστημα δημιουργίας συλλογισμών να είναι ικανό να υποστηρίξει πλήρως τη δημιουργία συλλογισμών για κάθε χαρακτηριστικό της OWL Full.

Καθεμία από αυτές τις υπογλώσσες αποτελεί επέκταση του απλούστερου προκατόχου της, και σε όσα μπορεί να εκφράσει νόμιμα (legally expressed), αλλά και σε όσα μπορεί να συμπεράνει έγκυρα (validly concluded). Οι παρακάτω σχέσεις ισχύουν, ενώ οι αντίστροφες τους δεν ισχύουν:

²⁰ Η περιγραφική λογική αποτελεί το σημαντικότερο επίσημο τρόπο αναπαράστασης γνώσης στην τεχνητή νοημοσύνη, ο οποίος αποτελεί τη λογική βάση για την ανάπτυξη των σημασιολογικών δικτύων (Semantic Networks), των συστημάτων που βασίζονται σε πλαίσια (Frame-based systems), κτλ [LUTZ07].

- Κάθε νόμιμη οντολογία της OWL Lite είναι μία νόμιμη οντολογία της OWL DL.
- Κάθε νόμιμη οντολογία της OWL DL είναι μία νόμιμη οντολογία της OWL Full.
- Κάθε έγκυρο συμπέρασμα της OWL Lite είναι ένα έγκυρο συμπέρασμα της OWL DL.
- Κάθε έγκυρο συμπέρασμα της OWL DL είναι ένα έγκυρο συμπέρασμα της OWL Full.

Οι κατασκευαστές οντολογιών που υιοθετούν την OWL πρέπει να εξετάζουν ποια υπογλώσσα ικανοποιεί τις ανάγκες τους. Η επιλογή μεταξύ των OWL Lite και OWL DL εξαρτάται από το βαθμό που απαιτούν οι χρήστες τα περισσότερα εκφραστικά κατασκευάσματα που παρέχει η OWL DL (πχ εύρος τιμών μεγαλύτερο του ένα). Η επιλογή μεταξύ των OWL DL και OWL Full εξαρτάται από το βαθμό απαίτησης των εκφραστικών κατασκευασμάτων που παρέχει η RDFS (πχ ο ορισμός κλάσεων από κλάσεις, η προσθήκη ιδιοτήτων στις κλάσεις, κτλ). Όταν χρησιμοποιείται η OWL Full σε σύγκριση με την OWL DL, η υποστήριξη συλλογισμών είναι λιγότερο προβλέψιμη (predictable) εφόσον δεν υπάρχουν ακόμη ολοκληρωμένες υλοποιήσεις της OWL Full [MCGU04]. Κάθε δήλωση της OWL Lite και της OWL DL μπορεί να αποφασιστεί (δηλαδή να δημιουργηθεί συλλογισμός για τη δήλωση) σε πεπερασμένο χρόνο. Αντίθετα, η OWL Full είναι πιθανόν να περιέχει ατέρμονες βρόχους (endless loops) [WIOU06].

Η OWL Full μπορεί να θεωρηθεί επέκταση του RDF, ενώ οι OWL Lite και OWL DL μπορούν να θεωρηθούν ως επεκτάσεις μίας περιορισμένης έκδοσης του RDF. Κάθε έγγραφο OWL (Lite, DL, Full) είναι ένα έγγραφο RDF, και κάθε έγγραφο RDF είναι ένα έγγραφο OWL Full, αλλά μόνο ορισμένα έγγραφα RDF είναι έγκυρα έγγραφα OWL Lite ή OWL DL. Για αυτούς τους λόγους, πρέπει να υπάρξει προσοχή σε περίπτωση που κάποιος χρήστης θέλει να αναβαθμίσει ένα έγγραφο RDF σε έγγραφο OWL. Όταν η εκφραστικότητα της OWL DL ή της OWL Lite κρίνεται κατάλληλη, πρέπει να ληφθούν υπόψιν κάποιες προφυλάξεις για να διασφαλιστεί ότι το αρχικό έγγραφο RDF συμμορφώνεται με τους επιπλέον περιορισμούς που επιβάλλουν οι OWL DL και OWL Lite. Πέραν των άλλων, κάθε προσδιοριστής πόρων που χρησιμοποιείται ως όνομα κλάσης πρέπει να δηλωθεί ρητά ως τύπου `owl:Class`. Παρομοίως για τις ιδιότητες, κάθε άτομο πρέπει να ανήκει σε τουλάχιστον μία κλάση (έστω και μόνο στην `owl:Thing`). Επίσης, οι προσδιοριστές των κλάσεων, των ιδιοτήτων, και των ατόμων, πρέπει να είναι αμοιβαία διαχωρίσιμοι (mutually disjoint).

Στις ενότητες που ακολουθούν, θα γίνει περιγραφή των κυριότερων χαρακτηριστικών των υπογλωσσών της OWL.

4.5.2 Χαρακτηριστικά της Web Ontology Language Lite

Η OWL Lite χρησιμοποιεί μόνο ορισμένα από τα χαρακτηριστικά της OWL, και εισάγει περισσότερους περιορισμούς στη χρήση αυτών, σε σύγκριση με τις OWL DL και OWL Full. Στη συνέχεια, ακολουθεί περιγραφή των χαρακτηριστικών της OWL Lite [MCGU04].

Χαρακτηριστικά της Resource Description Framework Schema που υπάρχουν στην Web Ontology Language Lite

Η OWL, παρόμοια με την RDFS, αποτελεί ένα ακόμη λεξιλόγιο που επεκτείνει τον RDF. Αυτό έχει ως αποτέλεσμα να είναι εύκολο να υιοθετήσει πολλά από τα χαρακτηριστικά της RDFS. Συγκεκριμένα, τα χαρακτηριστικά που υιοθετεί είναι τα εξής ([MCGU04], [SMIT04]):

- ➔ **owl:Class**: Μία κλάση είναι ένα σύνολο ατόμων που ανήκουν στην ίδια ομάδα και μοιράζονται ορισμένες ιδιότητες (πχ η Deborah και ο Frank είναι και οι δύο μέλη της κλάσης `Person`). Υπάρχουν οι κλάσεις που ορίζουν οι χρήστες (user-defined), και οι ριζικές κλάσεις περιοχής (domain root classes), οι οποίες διαφέρουν από τις πρώτες στο ότι δεν εμπεριέχουν στοιχεία (περιγράφεται απλά το όνομα τους). Οι κλάσεις μπορούν να οργανωθούν σε μία ιεραρχία εξειδίκευσης με χρήση της ιδιότητας `rdfs:subClassOf`. Υπάρχει μία έμφυτη γενική κλάση που ονομάζεται `owl:Thing`, και αποτελεί την κλάση όλων των ατόμων (όλα τα άτομα είναι στιγμιότυπα της) και την υπερκλάση όλων των κλάσεων OWL. Δηλαδή όλες οι κλάσεις που ορίζουν οι χρήστες είναι υποκλάσεις της `owl:Thing`. Επίσης, υπάρχει μία ειδική κλάση που ονομάζεται `owl:Nothing`, η οποία δεν έχει στιγμιότυπα (κενή κλάση) και αποτελεί υποκλάση όλων των κλάσεων OWL. Το παράδειγμα B-3 παρουσιάζει τη δήλωση των τριών ριζικών κλάσεων περιοχής `Winery`, `Region`, και `ConsumableThing`.
- ➔ **rdfs:subClassOf**: Η ιδιότητα λειτουργεί ακριβώς όπως λειτουργεί και στην RDFS.
- ➔ **rdf:Property**: Στην OWL Lite υπάρχουν δύο ειδών ιδιότητες. Οι ιδιότητες που συσχετίζουν κλάσεις (πχ `hasChild`, `hasRelative`, κτλ) ονομάζονται ιδιότητες

αντικειμένων (object properties), και αποτελούν στιγμιότυπα της κλάσης `owl:ObjectProperty`. Οι ιδιότητες που συσχετίζουν στιγμιότυπα κλάσεων και τύπους δεδομένων της XML Schema ονομάζονται ιδιότητες τύπων δεδομένων (datatype properties), και αποτελούν στιγμιότυπα της `owl:DatatypeProperty`. Με τη σειρά τους οι κλάσεις `owl:ObjectProperty` και `owl:DatatypeProperty` είναι υποκλάσεις της κλάσης `rdf:Property`. Στα παραδείγματα B-4 και B-5 γίνεται παρουσίαση των δύο ειδών ιδιοτήτων που παρέχει η OWL Lite. Στο παράδειγμα B-4, η ιδιότητα `madeFromGrape` είναι μία ιδιότητα αντικειμένων, στην οποία ορίζεται για περιοχή (domain) η κλάση `Wine`, και για εύρος (range) η κλάση `WineGrape`. Αντίθετα, στο παράδειγμα B-5, η ιδιότητα `yearValue` είναι μία ιδιότητα τύπων δεδομένων, στην οποία ορίζεται για περιοχή (domain) η κλάση `VintageYear`, και για εύρος (range) ο εγγενής τύπος δεδομένων της XML Schema `positiveInteger`.

- **`rdfs:subPropertyOf`**: Η ιδιότητα λειτουργεί ακριβώς όπως λειτουργεί και στην RDFS.
- **`rdfs:domain`**: Η ιδιότητα λειτουργεί ακριβώς όπως λειτουργεί και στην RDFS.
- **`owl:Individual`**: Τα άτομα είναι στιγμιότυπα κλάσεων, δηλαδή περιγράφονται κυρίως με τη σχέση *έχει-ένα* και όχι με τις σχέσεις *είναι-ένα*, *μοιάζει-με-ένα*, κτλ.

Ισότητα (Equality) και ανισότητα (Inequality) στην Web Ontology Language Lite

Ακολουθούν τα χαρακτηριστικά της OWL Lite που σχετίζονται με την ισότητα και την ανισότητα ([MCGU04], [SMIT04]):

- **`owl:equivalentClass`**: Δύο ισοδύναμες (equivalent) κλάσεις έχουν υποχρεωτικά τα ίδια στιγμιότυπα. Η ισοδυναμία χρησιμοποιείται για τη δημιουργία συνώνυμων κλάσεων. Πχ μπορούν να οριστούν ως ισοδύναμες οι κλάσεις `Car` και `AutoMobile`. Αυτό δίνει τη δυνατότητα σε ένα μηχανισμό συλλογιστικής (reasoner)²¹ να συμπεράνει ότι κάθε άτομο που είναι στιγμιότυπο της κλάσης `Car` είναι επίσης στιγμιότυπο της κλάσης `AutoMobile` και αντίστροφα. Το παράδειγμα B-6 παρουσιάζει τη χρήση του χαρακτηριστικού `owl:equivalentClass`. Στο παράδειγμα δηλώνεται ότι η κλάση `Wine` του

²¹ Από εδώ και στο εξής, ο μηχανισμός συλλογιστικής θα αναφέρεται απλά ως μηχανισμός για λόγους συντομίας.

ονοματοχώρου που χρησιμοποιεί η τρέχουσα οντολογία, είναι ισοδύναμη με την κλάση `Wine` κάποιας άλλης οντολογίας, που ορίζεται μοναδικά από τον ονοματοχώρο `&vin;` (συντόμευση του πλήρους ονόματος του ονοματοχώρου).

- **owl:equivalentProperty**: Δύο ισοδύναμες (equivalent) ιδιότητες σχετίζουν ένα άτομο με το ίδιο σύνολο άλλων ατόμων. Με αυτό τον τρόπο ορίζονται συνώνυμες ιδιότητες. Πχ μπορούν να οριστούν ως ισοδύναμες οι ιδιότητες `studiesC++` και `isProgrammer`. Από αυτή τη σχέση ο μηχανισμός μπορεί να συμπεράνει ότι αν το `X` σχετίζεται με το `Y` από την ιδιότητα `studiesC++`, τότε το `X` σχετίζεται με το `Y` και από την ιδιότητα `isProgrammer`, και αντίστροφα.
- **owl:sameAs**: Όταν υπάρχουν διαφορετικά ονόματα που αναφέρονται ουσιαστικά στο ίδιο άτομο, πρέπει να χρησιμοποιείται η ιδιότητα `owl:sameAs`. Πχ το άτομο `trash` μπορεί να οριστεί ως όμοιο με το άτομο `wasteBasket`. Το παράδειγμα B-7 παρουσιάζει τη χρήση της ιδιότητας `owl:sameAs`. Στο παράδειγμα δηλώνεται ότι η κλάση `Wine` με αναγνωριστικό `MikesFavoriteWine`, είναι ίδια (πανομοιότυπη) με την κλάση `StGenevieveTexasWhite`.
- **owl:differentFrom**: Με αυτή την ιδιότητα ένα άτομο μπορεί να δηλωθεί ως διαφορετικό από άλλα άτομα. Το αποτέλεσμα που προκύπτει είναι το αντίθετο της ιδιότητας `owl:sameAs`. Πχ το άτομο `Frank` μπορεί να δηλωθεί ως διαφορετικό από τα άτομα `Deborah` και `Jim`. Αν σε αυτή την περίπτωση, τα άτομα `Frank` και `Deborah` δηλωθούν ως τιμές μιας ιδιότητας, υπάρχει αντίφαση (contradiction), επειδή έχει δηλωθεί ρητά ότι τα δύο άτομα διαφέρουν. Σε τεχνολογίες όπως η OWL που δεν εξασφαλίζουν ότι τα άτομα θα έχουν μία (και μόνο μία) τιμή, τέτοιου είδους δηλώσεις είναι πολύ σημαντικές, διότι δίχως αυτές ο μηχανισμός αδυνατεί να συμπεράνει ότι δύο άτομα είναι διακριτά (distinct). Το παράδειγμα B-8 παρουσιάζει τη χρήση της ιδιότητας `owl:differentFrom`. Σε αυτό το παράδειγμα εξασφαλίζεται ότι δεν υπάρχει περίπτωση να υπάρχει πχ ένα άτομο τύπου `WineSugar` που να έχει την τιμή `Sweet`, και ταυτόχρονα να έχει και την τιμή `Dry`. Η δήλωση είναι απόλυτα σωστή, επειδή είναι γνωστό ότι ένας οίνος δεν γίνεται να είναι ταυτόχρονα και ξηρός (dry), και γλυκός (sweet).
- **owl:AllDifferent**: Αυτή η ιδιότητα χρησιμοποιείται όταν χρειάζεται να δηλωθούν ως αμοιβαία διακριτά (mutually distinct) ένας αριθμός από άτομα. Πχ έστω ότι οι `Frank`, `Jim`, και `Deborah` δηλώνονται ως αμοιβαία διακριτά άτομα. Σε αυτή την περίπτωση σε

αντίθεση με την ιδιότητα `owl:differentFrom`, κανένα από τα άτομα δεν μπορεί να χρησιμοποιηθεί μαζί με κάποιο άλλο ως τιμή ιδιότητας. Το παράδειγμα B-9 παρουσιάζει τη χρήση της ιδιότητας `owl:AllDifferent`. Με τη χρήση της ιδιότητας `AllDiferrent` στο παράδειγμα B-9, εξασφαλίζεται ότι το χρώμα ενός οίνου (`WineColor`) δεν γίνεται να είναι πχ ταυτόχρονα και κόκκινο (`red`), και λευκό (`white`).

Χαρακτηριστικά ιδιοτήτων (Property Characteristics) της Web Ontology Language Lite

Στην OWL Lite, υπάρχουν ειδικοί προσδιοριστές που παρέχουν πληροφορίες σχετικά με τις ιδιότητες και τις τιμές που αυτές λαμβάνουν. Η διάκριση μεταξύ της ιδιότητας `ObjectProperty` και της `DataProperty` έχει αναφερθεί ήδη. Τα υπόλοιπα σημαντικά χαρακτηριστικά των ιδιοτήτων της OWL είναι τα παρακάτω ([MCGU04], [SMIT04]):

- ➔ **`owl:inverseOf`:** Μία ιδιότητα μπορεί να δηλωθεί ως η αντίστροφη (*inverse*) μίας άλλης. Αν η ιδιότητα `P1` δηλωθεί ως αντίστροφη της ιδιότητας `P2`, τότε αν το `x` σχετίζεται με το `y` μέσω της ιδιότητας `P2`, το `y` σχετίζεται με το `x` μέσω της ιδιότητας `P1`. Πχ αν η ιδιότητα `hasChild` είναι αντίστροφη της `hasParent` και ισχύει η σχέση "`Deborah hasParent Louise`", τότε ο μηχανισμός μπορεί να συμπεράνει ότι ισχύει επίσης η σχέση "`Louise hasChild Deborah`". Το παράδειγμα B-10 παρουσιάζει τη χρήση αυτού του χαρακτηριστικού. Στο παράδειγμα B-10 δηλώνονται ως αντίστροφες οι ιδιότητες `hasMaker` και `producesWine`. Έτσι, αν για παράδειγμα ισχύει η σχέση "`Pinot Noir hasMaker David Bruce`", τότε ο μηχανισμός θα συμπεράνει αυτόματα ότι ισχύει η σχέση "`David Bruce producesWine Pinot Noir`".
- ➔ **`owl:TransitiveProperty`:** Οι ιδιότητες μπορούν να οριστούν ως μεταβατικές (*transitive*). Έστω ότι μία ιδιότητα `P` είναι μεταβατική. Ισχύει ότι αν τα ζευγάρια (x, y) και (y, z) είναι στιγμιότυπα της `P`, τότε και το ζευγάρι (x, z) είναι επίσης στιγμιότυπο της. Πχ αν η ιδιότητα `ancestor` (πρόγονος) έχει δηλωθεί μεταβατική, και ισχύει η σχέση "`Sarah is an ancestor of Louise`" και η "`Louise is an ancestor of Deborah`", τότε ο μηχανισμός μπορεί να συμπεράνει ότι ισχύει επίσης η σχέση "`Sarah`

is an ancestor of Deborah". Το παράδειγμα B-11 παρουσιάζει τη χρήση του χαρακτηριστικού `owl:TransitiveProperty`. Στο παράδειγμα B-11 δηλώνεται ως μεταβατική η ιδιότητα `locatedIn`. Έτσι, εφόσον ισχύει η σχέση "`SantaCruzMountainsRegion locatedIn CaliforniaRegion`", και η σχέση "`CaliforniaRegion locatedIn USRegion`", ο μηχανισμός θα δημιουργήσει αυτόματα τη σχέση "`SantaCruzMountainsRegion locatedIn USRegion`".

→ **owl:SymmetricProperty**: Οι ιδιότητες μπορούν να δηλωθούν ως συμμετρικές (`symmetric`). Έστω ότι μία ιδιότητα `P` είναι συμμετρική. Ισχύει ότι αν το ζευγάρι (x, y) είναι στιγμιότυπο της συμμετρικής ιδιότητας `P`, τότε και το ζευγάρι (y, x) είναι επίσης στιγμιότυπο της. Πχ έστω ότι η ιδιότητα `friend` είναι συμμετρική. Αν ισχύει η σχέση "`Jim is a friend of Nicol`", ο μηχανισμός μπορεί να συμπεράνει ότι ισχύει επίσης η σχέση "`Nicol is a friend of Jim`". Το παράδειγμα B-12 παρουσιάζει τη χρήση του χαρακτηριστικού `owl:SymmetricProperty`. Στο παράδειγμα δηλώνεται ως συμμετρική η ιδιότητα `adjacentRegion` (γειτονική περιοχή). Δηλώνεται επίσης ότι ισχύει η σχέση "`MendocinoRegion adjacentRegion SonomaRegion`", συνεπώς δίνεται η δυνατότητα στον μηχανισμό να δημιουργήσει αυτόματα τη σχέση "`SonomaRegion adjacentRegion MendocinoRegion`".

→ **owl:FunctionalProperty**: Οι ιδιότητες μπορούν να δηλωθούν ως συναρτησιακές (`functional`). Αν μία ιδιότητα δηλωθεί ως συναρτησιακή, τότε δεν μπορεί να έχει παραπάνω από μία τιμή για κάθε άτομο (βέβαια δεν είναι λάθος να μην έχει καμία τιμή για κάποιο άτομο). Αυτό το χαρακτηριστικό της OWL είναι γνωστό και ως "unique property". Η `owl:FunctionalProperty` είναι στενογραφία (`shorthand`) για τη δήλωση ελάχιστης τιμής εύρους (`minimum cardinality`) μίας ιδιότητας μηδέν, και μέγιστης τιμής εύρους (`maximum cardinality`) ένα. Πχ έστω ότι η ιδιότητα `hasPrimaryEmployer` είναι συναρτησιακή. Από αυτή τη δήλωση ο μηχανισμός μπορεί να συμπεράνει ότι **κανένα** άτομο δεν μπορεί να έχει παραπάνω από έναν βασικό εργοδότη (`primary employer`). Αυτό φυσικά δεν υπονοεί ότι **όλα** τα άτομα πρέπει να έχουν τουλάχιστον ένα βασικό εργοδότη. Το παράδειγμα B-13 παρουσιάζει τη χρήση του χαρακτηριστικού `owl:FunctionalProperty`. Στο παράδειγμα δηλώνεται ως συναρτησιακή η ιδιότητα `hasVintageYear` (έτος παραγωγής οίνου). Δηλώνοντας αυτή την ιδιότητα, το έτος παραγωγής ενός οίνου περιορίζεται σε καμία (δεν έχει παραχθεί καθόλου) ή μία μόνο τιμή

(συγκεκριμένο έτος παραγωγής).

→ **owl:InverseFunctionalProperty**: Οι ιδιότητες μπορούν να δηλωθούν ως αντίστροφα συναρτησιακές (inverse functional). Αν μία ιδιότητα δηλωθεί ως αντίστροφα συναρτησιακή, τότε η αντίστροφη αυτής της ιδιότητας είναι συναρτησιακή. Επομένως η αντίστροφη αυτής της ιδιότητας έχει το πολύ μία τιμή για κάθε άτομο. Αυτό το χαρακτηριστικό της OWL είναι γνωστό και ως “unambiguous property”. Πχ έστω ότι η ιδιότητα `hasUSSocialSecurityNumber`²² είναι αντίστροφα συναρτησιακή. Τότε η αντίστροφη αυτής της ιδιότητας (έστω ότι είναι η `isTheSocialSecurityNumberFor`) έχει το πολύ μία τιμή για οποιοδήποτε άτομο. Επομένως η μοναδική τιμή που μπορεί να έχει η ιδιότητα `isTheSocialSecurityNumberFor` είναι ο SSN ενός ατόμου. Από αυτή τη δήλωση ο μηχανισμός μπορεί να συμπεράνει ότι δύο διαφορετικά στιγμιότυπα ατόμων (πχ της κλάσης `Person`) δεν μπορούν να έχουν τον ίδιο SSN. Από την ίδια δήλωση, ο μηχανισμός μπορεί να συμπεράνει επίσης ότι αν δύο στιγμιότυπα της κλάσης `Person` έχουν τον ίδιο SSN, τότε τα δύο στιγμιότυπα αναφέρονται στο ίδιο ακριβώς άτομο. Το παράδειγμα B-14 παρουσιάζει τη χρήση του χαρακτηριστικού `owl:InverseFunctionalProperty`. Στο παράδειγμα δηλώνεται η ιδιότητα `hasMaker` ως αντίστροφα συναρτησιακή της ιδιότητας `producesWine`. Έτσι λοιπόν αν ισχύει η σχέση “`Pinot Noir hasMaker David Bruce`”, τότε ο μηχανισμός θα δημιουργήσει αυτόματα τη σχέση “`David Bruce producesWine Pinot Noir`”.

Περιορισμοί ιδιοτήτων (Property Restrictions) της Web Ontology Language Lite

Η OWL Lite επιτρέπει την υποβολή περιορισμών σχετικά με τον τρόπο που θα χρησιμοποιούνται οι ιδιότητες από τα στιγμιότυπα μίας κλάσης. Αυτοί οι περιορισμοί (καθώς και οι περιορισμοί εύρους που περιγράφονται στην επόμενη ενότητα) χρησιμοποιούν ως υπερκλάση την `owl:Restriction`. Το στοιχείο `owl:onProperty` περιγράφει την ιδιότητα που περιορίζεται. Οι δύο παρακάτω περιορισμοί περιορίζουν τις τιμές που μπορούν να χρησιμοποιηθούν (δηλαδή ορίζουν ποιες είναι οι αποδεκτές τιμές), ενώ οι περιορισμοί της

²² Ο Social Security Number (SSN) είναι ένας μοναδικός αριθμός εννέα ψηφίων που παρέχεται σε κάθε πολίτη των Ηνωμένων Πολιτειών Αμερικής για την εξυπηρέτηση των αναγκών του [WISO07].

επόμενης ενότητας περιορίζουν το εύρος των τιμών (δηλαδή ορίζουν πόσες τιμές μπορούν να χρησιμοποιηθούν). Οι περιορισμοί ιδιοτήτων της OWL Lite είναι ([MCGU04], [SMIT04]):

→ **owl:allValuesFrom:** Αυτός ο περιορισμός δηλώνεται σε μία ιδιότητα πάντα σε συνδυασμό με μία συγκεκριμένη κλάση. Σημαίνει ότι η συγκεκριμένη ιδιότητα της συγκεκριμένης κλάσης σχετίζεται με έναν τοπικό περιορισμό εύρους (local range restriction). Πχ έστω ότι η κλάση *Person* έχει μία ιδιότητα *hasDaughter* που περιορίζεται με την ιδιότητα *owl:allValuesFrom* στην κλάση *Woman*. Αυτό σημαίνει ότι όλες οι τιμές της ιδιότητας *hasDaughter* είναι υποχρεωτικά στιγμιότυπα της κλάσης *Woman*. Συνεπώς αν ισχύει η σχέση “*Louise hasDaughter Deborah*”, τότε ο μηχανισμός μπορεί να συμπεράνει ότι η *Deborah* είναι στιγμιότυπο της κλάσης *Woman*. Το παράδειγμα B-15 παρουσιάζει τη χρήση του περιορισμού ιδιοτήτων *owl:allValuesFrom*. Σε αυτό το παράδειγμα περιορίζεται η ιδιότητα *hasMaker* με τον περιορισμό *allValuesFrom* στην κλάση *Winery*. Αυτή η δήλωση επιβάλλει σε όλες τις σχέσεις του τύπου “*X hasMaker Y*” (όπου *X* στιγμιότυπο της κλάσης *Wine*), το *Y* να είναι πάντα στιγμιότυπο της κλάσης *Winery*.

→ **owl:someValuesFrom:** Παρόμοια με τον περιορισμό *owl:allValuesFrom*, ο περιορισμός *owl:someValuesFrom* δηλώνεται σε μία ιδιότητα πάντα σε συνδυασμό με μία συγκεκριμένη κλάση. Αυτός ο περιορισμός επιβάλλει στη συγκεκριμένη ιδιότητα της συγκεκριμένης κλάσης που ορίζεται να έχει τουλάχιστον μία τιμή που θα είναι συγκεκριμένου τύπου. Πχ έστω ότι η κλάση *SemanticWebPaper* έχει μία ιδιότητα *hasKeyword*, που περιορίζεται με την ιδιότητα *owl:someValuesFrom*, έτσι ώστε **κάποια** τιμή της να είναι στιγμιότυπο της κλάσης *SemanticWebTopic*. Αυτός ο περιορισμός επιτρέπει στο έγγραφο (paper) να έχει πολλές λέξεις κλειδιά (keywords) που είναι στιγμιότυπα οποιασδήποτε κλάσης, εκ των οποίων μόνο η μία λέξη κλειδί πρέπει να είναι υποχρεωτικά στιγμιότυπο της κλάσης *SemanticWebTopic*, σε αντίθεση με τον περιορισμό *owl:allValuesFrom*. Το παράδειγμα B-16 παρουσιάζει τη χρήση του περιορισμού ιδιοτήτων *owl:someValuesFrom*. Το παράδειγμα B-16 είναι παρόμοιο με το B-15, με μόνη διαφορά τη χρήση του περιορισμού *someValuesFrom* αντί του *allValuesFrom*.

Περιορισμός εύρους τιμών (Restricted Cardinality) της Web Ontology Language Lite

Η OWL Lite υποστηρίζει μία (περιορισμένη) μορφή περιορισμών εύρους τιμών. Οι περιορισμοί εύρους τιμών στην OWL αναφέρονται ως τοπικοί περιορισμοί (local restrictions), επειδή δηλώνονται σε ιδιότητες πάντα σε συνδυασμό με μία συγκεκριμένη κλάση. Δηλαδή οι περιορισμοί περιορίζουν το εύρος μίας ιδιότητας στα στιγμιότυπα μίας κλάσης. Οι περιορισμοί εύρους τιμών της OWL Lite είναι περιορισμένοι επειδή επιτρέπουν τη δημιουργία δηλώσεων με τιμές εύρους μόνο μηδέν ή ένα (δεν επιτρέπουν τη χρήση αυθαίρετων τιμών όπως η OWL DL και η OWL Full). Οι περιορισμοί εύρους τιμών που υποστηρίζει η OWL Lite είναι οι εξής ([MCGU04], [SMIT04]):

→ **owl:minCardinality:** Όπως τονίστηκε ήδη, το εύρος τιμών (cardinality) δηλώνεται σε μία ιδιότητα πάντα σε συνδυασμό με μία συγκεκριμένη κλάση. Αν δηλωθεί σε μία ιδιότητα ελάχιστο εύρος (**owl:minCardinality**) ίσο με ένα, τότε οποιοδήποτε στιγμιότυπο της κλάσης που ανήκει η ιδιότητα πρέπει να σχετίζεται υποχρεωτικά με τουλάχιστον ένα άτομο. Αυτός είναι ένας τρόπος με τον οποίο μπορεί να δηλωθεί ότι μία ιδιότητα απαιτείται να έχει τιμή για όλα τα στιγμιότυπα της κλάσης που ανήκει. Πχ η κλάση *Person* δεν πρέπει να έχει περιορισμούς στο ελάχιστο εύρος της ιδιότητας *hasOffspring* επειδή δεν έχουν υποχρεωτικά όλοι οι άνθρωποι απογόνους. Αντίθετα, η κλάση *Parent* πρέπει να έχει ελάχιστο εύρος ίσο με ένα στην ιδιότητα *hasOffspring*, επειδή για να είναι κάποιος άνθρωπος γονέας πρέπει να έχει τουλάχιστον έναν απόγονο. Αν ο μηχανισμός γνωρίζει τη σχέση “*Louise is a Person*”, δε μπορεί να συμπεράνει τίποτα σχετικά με την ιδιότητα *hasOffspring*. Αν όμως ανιχνεύσει ότι ισχύει η σχέση “*Louise is a Parent*”, τότε μπορεί να συμπεράνει ότι η *Louise* σχετίζεται με τουλάχιστον ένα άτομο μέσω της ιδιότητας *hasOffspring*. Βέβαια ο μηχανισμός δεν μπορεί να συμπεράνει τίποτα σχετικά με το μέγιστο αριθμό των απογόνων των στιγμιότυπων της κλάσης *Parent*, χρησιμοποιώντας μόνο αυτή την πληροφορία. Όπως έχει ειπωθεί ήδη, στην OWL Lite επιτρέπονται μόνο ελάχιστες τιμές εύρους ίσες με μηδέν ή ένα. Η τιμή εύρους μηδέν σε μία ιδιότητα σημαίνει ότι η ιδιότητα είναι προαιρετική (στην κλάση που αναφέρεται).

→ **owl:maxCardinality:** Αν δηλωθεί σε μία ιδιότητα μέγιστο εύρος (**owl:maxCardinality**) ίσο με ένα, που είναι η μέγιστη τιμή εύρους που επιτρέπει η

OWL Lite (πάντα σε συνδυασμό με μία συγκεκριμένη κλάση), τότε οποιοδήποτε στιγμιότυπο της κλάσης που ανήκει η ιδιότητα μπορεί να σχετίζεται το πολύ με ένα άτομο της. Πχ όταν η ιδιότητα `hasRegisteredVotingCountry` σχετίζεται με την κλάση `EuropeanCitizens` είναι σωστό να έχει μέγιστη τιμή εύρους ίση με ένα (επειδή ένας Ευρωπαίος πολίτης έχει δικαίωμα να ψηφίζει μόνο σε μία χώρα). Από αυτή τη σχέση ο μηχανισμός μπορεί να συμπεράνει ότι τα στιγμιότυπα της κλάσης `EuropeanCitizens` δεν μπορούν να σχετίζονται με δύο (ή παραπάνω) διαφορετικά άτομα μέσω της ιδιότητας `hasRegisteredVotingCountry`. Βέβαια ο μηχανισμός δεν μπορεί να συμπεράνει τίποτα σχετικά με τον ελάχιστο αριθμό των χωρών που μπορεί να ψηφίσει ένας Ευρωπαίος πολίτης, χρησιμοποιώντας μόνο αυτή την πληροφορία.

➔ **owl:cardinality**: Αυτή η ιδιότητα χρησιμοποιείται ώστε να υπάρχει συνέπεια σε περίπτωση που μία ιδιότητα κάποιας κλάσης έχει ορίσει και στους δύο περιορισμούς εύρους τιμών (`owl:minCardinality` και `owl:maxCardinality`) την τιμή μηδέν, ή την τιμή ένα. Πχ η κλάση `Person` είναι σωστό να δέχεται μία (ακριβώς) τιμή στην ιδιότητα `hasBirthMother`, επειδή κάθε άνθρωπος έχει μόνο μία βιολογική μητέρα. Από αυτή τη σχέση ο μηχανισμός μπορεί να συμπεράνει ότι απαγορεύεται δύο διαφορετικά στιγμιότυπα της κλάσης `Mother` να είναι τιμές της ιδιότητας `hasBirthMother` του ίδιου ατόμου. Το παράδειγμα B-17 παρουσιάζει τη χρήση του περιορισμού εύρους τιμών `owl:cardinality`. Στο παράδειγμα B-17, ο περιορισμός `cardinality` εφαρμόζεται στην κλάση `Vintage`, και συγκεκριμένα στην ιδιότητα `hasVintageYear`. Το έτος παραγωγής ενός συγκεκριμένου τύπου οίνου είναι μόνο ένα, οπότε το `cardinality` ορίζεται ίσο με την τιμή ένα.

Τομή κλάσεων (Class Intersection) στην Web Ontology Language Lite

Η OWL Lite παρέχει έναν τελεστή τομής (intersection constructor), αλλά περιορίζει τη χρήση του. Αναλυτικότερα [MCGU04]:

➔ **owl:intersectionOf**: Η OWL Lite επιτρέπει την τομή (intersection) κλάσεων και περιορισμών. Πχ η κλάση `EmployedPerson` (υπάλληλος με πρόσληψη) μπορεί να περιγραφεί ως μία σχέση “`Person intersectionOf EmployedThings`” (όπου η κλάση `EmployedThings` ορίζει αντικείμενα/things με περιορισμό `owl:minCardinality`

ίσο με ένα στην ιδιότητα `hasEmployer`). Αυτό δίνει τη δυνατότητα στο μηχανισμό να συμπεράνει ότι κάθε υπάλληλος με πρόσληψη (`EmployedPerson`) έχει τουλάχιστον έναν εργοδότη (`employer`). Τούτο το συμπέρασμα εξάγεται επειδή λόγω του περιορισμού `owl:minCardinality` είναι υποχρεωτικό να υπάρχει τουλάχιστον ένα στιγμιότυπο της κλάσης `EmployedThings`, με το οποίο θα σχετίζεται η ιδιότητα `hasEmployer` (δηλαδή θα υπάρχει μία σχέση "`EmployedThings hasEmployer Person`", πχ αν ένα θεωρηθεί ότι το άτομο `nick` αποτελεί στιγμιότυπο της κλάσης `EmployedThings` και το άτομο `mary` στιγμιότυπο της κλάσης `Person`, θα υπάρχει μία σχέση "`nick hasEmployer mary`".

Επιπλέον περιγραφικά χαρακτηριστικά των Web Ontology Language DL και Web Ontology Language Full

Οι γλώσσες OWL DL και OWL Full χρησιμοποιούν το ίδιο λεξιλόγιο, αν και η OWL DL υπόκειται σε κάποιους περιορισμούς. Πχ η OWL DL απαιτεί διαχωρισμό τύπων (μία κλάση δεν μπορεί να είναι επιπλέον και άτομο ή ιδιότητα, και μία ιδιότητα δεν μπορεί να είναι επιπλέον και άτομο ή κλάση). Επιπρόσθετα, η OWL DL απαιτεί οι ιδιότητες να είναι είτε ιδιότητες τύπων δεδομένων (`datatype properties`) είτε ιδιότητες αντικειμένων (`object properties`), κάτι το οποίο δεν ισχύει στην OWL Full. Οι διαφορές μεταξύ των λεξιλογίων των δύο γλωσσών²³ δημιουργούν ένα κοινό λεξιλόγιο που επεκτείνει τα χαρακτηριστικά της OWL Lite προσθέτοντας τα παρακάτω στοιχεία ([MCGU04], [SMIT04]):

→ **owl:oneOf**: Οι κλάσεις μπορούν να περιγραφούν ως μία απαρίθμηση (`enumeration`) των ατόμων που τις συνθέτουν (απαριθμημένες κλάσεις). Πχ η κλάση `DaysOfTheWeek` μπορεί να περιγράψει απαριθμώντας τα άτομα `sunday`, `monday`, `tuesday`, `wednesday`, `thursday`, `friday`, `saturday`. Από αυτό το χαρακτηριστικό ο μηχανισμός μπορεί να συμπεράνει ποιο είναι το μέγιστο εύρος τιμών (επτά) που έχει οποιαδήποτε ιδιότητα χρησιμοποιεί τον περιορισμό `owl:allValuesFrom` με τιμή την κλάση `DaysOfTheWeek`. Το παράδειγμα B-18 παρουσιάζει τη χρήση του απαριθμητή κλάσεων `owl:oneOf`. Στο παράδειγμα ορίζεται ότι η κλάση `WineColor` πρέπει να λαμβάνει μία εκ των τιμών `White`, `Rose`, και `Red`. Συνεπώς αν δηλωθεί κάπου η ιδιότητα `hasColor` η οποία θα χρησιμοποιεί

²³ Αναλυτική περιγραφή σχετικά με τις διαφορές που παρουσιάζουν τα λεξιλόγια της OWL DL και της OWL Full, υπάρχει στο [PATE04].

τον περιορισμό `owl:allValuesFrom` με τιμή την κλάση `WineColor`, ο μηχανισμός μπορεί να συμπεράνει ότι το μέγιστο εύρος τιμών της ιδιότητας `hasColor` ισούται με την τιμή τρία.

- **`owl:hasValue`**: Μία ιδιότητα συχνά απαιτείται να έχει ως τιμή ένα συγκεκριμένο άτομο (τιμές ιδιοτήτων). Πχ τα στιγμιότυπα της κλάσης `DutchCitizens` μπορούν να χαρακτηριστούν ως άνθρωποι που έχουν εθνικότητα το στιγμιότυπο `theNetherlands` (όπου το `theNetherlands` είναι με τη σειρά του στιγμιότυπο της κλάσης `Nationalities`). Το παράδειγμα B-19 παρουσιάζει τη χρήση του χαρακτηρισμού `owl:hasValue`. Στο παράδειγμα δηλώνεται ότι ένας συγκεκριμένος τύπος οίνου (`Burgundy`) λαμβάνει πάντα στην ιδιότητα του `hasSugar` την κλάση `Dry`. Δηλαδή δηλώνεται ότι ο `Burgundy` είναι ξηρός (και μόνο ξηρός) οίνος.
- **`owl:disjointWith`**: Οι κλάσεις μπορούν να δηλωθούν ως διαχωρίσιμες (`disjoint`) από τις υπόλοιπες. Πχ έστω ότι οι κλάσεις `Man` και `Woman` έχουν δηλωθεί ως διαχωρίσιμες. Από αυτή τη δήλωση, ο μηχανισμός μπορεί να συμπεράνει ότι υπάρχει ασυνέπεια όταν ένα άτομο δηλωθεί ως στιγμιότυπο και των δύο κλάσεων. Επίσης, ο μηχανισμός μπορεί να συμπεράνει ότι αν το `A` είναι στιγμιότυπο της κλάσης `Man`, τότε το `A` δεν είναι στιγμιότυπο της κλάσης `Woman`. Το παράδειγμα B-20 παρουσιάζει τη χρήση της δήλωσης `owl:disjointWith`. Στο παράδειγμα B-20, γίνεται διαχωρισμός μεταξύ των διαφορετικών κατηγοριών τροφίμων, πχ ζυμαρικά (`Pasta`), κρεατικά (`Meat`), φρούτα (`Fruit`), κτλ. Έτσι ο μηχανισμός προστατεύει την οντολογία από ασυνέπειες όπως το να δηλωθεί ότι ένα τρόφιμο είναι ταυτόχρονα ζυμαρικό και φρούτο.
- **`owl:unionOf`, `owl:complementOf`, `owl:intersectionOf`**: Οι γλώσσες OWL DL και OWL Full επιτρέπουν τον αυθαίρετο συνδυασμό κλάσεων και περιορισμών (`Boolean combinations`), με εφαρμογή των τελεστών συνόλων (ή δυαδικών τελεστών): της ένωσης (`union`), της τομής (`intersection`), και του συμπληρώματος (`complement`). Πχ με χρήση του τελεστή ένωσης (`owl:unionOf`) μπορεί να δηλωθεί ότι οι κάτοικοι της Θεσσαλονίκης είναι είτε Έλληνες (`Greek`) είτε αλλοδαποί (`Foreigners`). Με χρήση του τελεστή συμπληρώματος (`owl:complementOf`) μπορεί να δηλωθεί πχ ότι τα παιδιά (`Children`) **δεν** θεωρούνται πολίτες (`SeniorCitizens`). Αυτό μπορεί να γίνει αν πχ η κλάση `Children` είναι υποκλάση του συμπληρώματος της κλάσης `SeniorCitizens`. Το

παράδειγμα B-21 παρουσιάζει τη χρήση των τελεστών συνόλων `owl:unionOf` και `owl:complementOf`. Η χρήση του τελεστή `unionOf` δηλώνει ότι ένα φρούτο (`Fruit`) μπορεί να έχει είτε γλυκιά γεύση (`SweetFruit`) είτε πικρή γεύση (`NonSweetFruit`). Ο τελεστής `complementOf` διαχωρίζει τα αγαθά που χρειάζονται συχνά ανανέωση (`ConsumableThing`), από όλα τα υπόλοιπα, τα οποία προκύπτουν αποτελούν το συμπλήρωμα της κλάσης `ConsumableThing`.

- ➔ **`owl:minCardinality`, `owl:maxCardinality`, `owl:cardinality`:** Σε αντίθεση με την OWL Lite που το εύρος τιμών περιορίζεται σε μηδέν ή ένα, οι γλώσσες OWL DL και OWL Full επιτρέπουν τη δημιουργία δηλώσεων με αυθαίρετες (καθώς και μη ακέραιες) τιμές εύρους. Πχ η ιδιότητα `teamMembers` θα έχει ελάχιστο εύρος τιμών ίσο με τρία (επειδή μία ομάδα αποτελείται από τουλάχιστον τρία μέλη, αφού τα δύο άτομα θεωρούνται ζεύγος).
- ➔ **Σύνθετες κλάσεις (complex classes):** Σε αντίθεση με την OWL Lite που περιορίζει το συντακτικό με χρήση απλών κλάσεων (πχ με δηλώσεις `owl:subClassOf` και `owl:equivalentClass`), η OWL Full επιτρέπει την περιγραφή σύνθετων κλάσεων, που ενσωματώνουν απαριθμημένες κλάσεις, περιορισμούς ιδιοτήτων, και δυαδικούς συνδυασμούς (Boolean combinations).

5. Κατασκευή οντολογίας ιστού Ευρωπαίου πολίτη (EuroCitizen)

Τελικός σκοπός αυτής της εργασίας είναι η δημιουργία μίας οντολογίας ιστού ή οποία θα παρέχει τη βασική υποδομή για την περιγραφή ενός Ευρωπαίου πολίτη (εξ ου και το όνομα της οντολογίας, EuroCitizen). Μεγαλύτερη έμφαση δίδεται στα στοιχεία που κάνουν έναν Ευρωπαίο πολίτη να διαφέρει από όλες τις υπόλοιπες κατηγορίες πολιτών (πχ πολίτες ΗΠΑ).

Στο κεφάλαιο αυτό περιγράφονται οι τεχνολογίες και τα εργαλεία που χρησιμοποιήθηκαν για την κατασκευή της οντολογίας EuroCitizen. Επίσης, γίνεται αναφορά σε ορισμένα θέματα που πρέπει να προσέξει κάποιος αρχάριος, προτού ξεκινήσει να κατασκευάσει μία οντολογία ιστού. Τέλος, παρουσιάζονται οι διάφοροι τρόποι και τομείς στους οποίους μπορεί να αξιοποιηθεί η οντολογία EuroCitizen, καθώς και οι ελλείψεις και τα προβλήματα που παρουσιάστηκαν κατά την ανάπτυξη της.

Μετά την ολοκλήρωση αυτού του κεφαλαίου, ο αναγνώστης θα πρέπει να νιώθει έτοιμος ώστε να επιλέξει τα κατάλληλα εργαλεία και τεχνολογίες, και με βάση όσα αναφέρονται σε αυτό αλλά και στα προηγούμενα κεφάλαια, να είναι ικανός να κατασκευάζει μία απλή οντολογία ιστού.

5.1 Επιλογή τεχνολογίας

Η πρώτη απόφαση που πρέπει να λάβει κάποιος αφού αποκτήσει τις προδιαγραφές (και ότι επιπλέον πληροφορίες κριθούν απαραίτητες) για τη δημιουργία της οντολογίας του, είναι η τεχνολογία που θα χρησιμοποιήσει. Για την οντολογία EuroCitizen, αποφασίστηκε να χρησιμοποιηθεί η τεχνολογία OWL, επειδή είναι:

- ✓ η καταλληλότερη τεχνολογία για τη δημιουργία οντολογιών **ιστού**,

- ✓ η συνιστώμενη τεχνολογία για τη δημιουργία οντολογιών ιστού από το W3C,
- ✓ ανοιχτή τεχνολογία,
- ✓ συμβατή με τα πρότυπα του ιστού που αφορούν την προσβασιμότητα και τη διεθνοποίηση.

Επιλέγοντας την OWL ως τεχνολογία υλοποίησης, προκύπτει ένα επιπλέον ερώτημα: Ποια έκδοση της OWL είναι αυτή που θεωρείται καταλληλότερη για τη δημιουργία της οντολογίας; Αυτή η απάντηση πρέπει να δοθεί σε πρώιμο στάδιο, επειδή παίζει καθοριστικό ρόλο στο αν είναι δυνατό να επαναχρησιμοποιηθεί κάποια οντολογία (είτε θεμελιώδης είτε περιοχής) από όλες αυτές που υπάρχουν στο Διαδίκτυο και διατίθενται για ελεύθερη χρήση.

Έχει γίνει ήδη αναφορά στις τρεις εκδόσεις της OWL (Lite, DL, Full). Ο εκάστοτε κατασκευαστής, θα πρέπει να διαλέξει μία από τις εκδόσεις της ανάλογα με τις ανάγκες του. Παρόλα αυτά υπάρχει ένα σημαντικό θέμα στο οποίο δεν έχει γίνει αναφορά μέχρι τώρα. Ένας μηχανισμός συλλογιστικής (reasoner) είναι ένα ιδιαίτερα χρήσιμο εργαλείο, που είναι ικανό να υπολογίζει αυτόματα την ταξινομημένη ιεραρχία (classification hierarchy) μίας οντολογίας, και να εντοπίζει πιθανές ασυνέπειες (inconsistencies) που υπάρχουν σε αυτήν. Για να είναι δυνατή η χρήση ενός μηχανισμού συλλογιστικής, η έκδοση της OWL που θα πρέπει να χρησιμοποιηθεί είναι είτε η OWL DL, είτε η OWL Lite [KNUB04]. Για όλους τους παραπάνω λόγους, αποφασίστηκε ότι η καταλληλότερη έκδοση της OWL για τη δημιουργία της οντολογίας EuroCitizen είναι η OWL DL.

Ένας άτυπος κανόνας που ορίζεται στην συγγραφή λογισμικού (καθώς και σε πολλούς άλλους τομείς), είναι αυτός που λέει ότι “αν υπάρχει ήδη κάτι έτοιμο που μπορεί να χρησιμοποιηθεί για τη διευκόλυνση του κατασκευαστή, το καλύτερο που μπορεί να κάνει είναι να το χρησιμοποιήσει”. Πριν την έναρξη της κατασκευής της οντολογίας EuroCitizen, έγινε μεγάλη προσπάθεια να χρησιμοποιηθούν αρκετές θεμελιώδεις οντολογίες (πχ WordNet, SUMO, OpenCyc, κτλ) [NOY00]. Το πρόβλημα με τις περισσότερες από αυτές είναι ότι είναι γραμμένες στην έκδοση OWL Full (ως παράδειγμα μπορούν να αναφερθούν οι οντολογίες GUM και SUMO), καθιστώντας συνεπώς αδύνατη τη χρήση ενός μηχανισμού συλλογιστικής. Άλλες, ήταν αδύνατο να φορτωθούν στο εργαλείο σύνταξης της οντολογίας, στο οποίο γίνεται αναφορά στην αμέσως επόμενη ενότητα, λόγω του τεράστιου μεγέθους τους (ως παράδειγμα

μπορούν να αναφερθούν οι οντολογίες OpenCyc και WordNet). Τελικά, βρέθηκε μία αξιόλογη οντολογία περιοχής που κρίθηκε κατάλληλη για χρήση. Πρόκειται για την οντολογία New Testament Names (NTN), η οποία ανήκει στο έργο δημιουργίας μίας Βίβλου (Χριστιανικής) με σημασιολογικό περιεχόμενο (Semantic Bible). Η οντολογία NTN βασίζεται στην οντολογία SUMO, και η οντολογία EuroCitizen με τη σειρά της, βασίζεται σε μία τροποποιημένη έκδοση της οντολογίας NTN ([BOIS04], [BOIS05]).

5.2 Επιλογή εργαλείου σύνταξης οντολογίας

Η δημιουργία μίας οντολογίας δίχως τη βοήθεια κάποιου εργαλείου, είναι τόσο σύνθετη όσο είναι και η κατασκευή οποιουδήποτε λογισμικού δίχως τη χρήση ενός ολοκληρωμένου περιβάλλοντος ανάπτυξης (Integrated Development Environment – IDE). Η επιλογή ενός καλού εργαλείου είναι το δεύτερο βήμα που πρέπει να κάνει κάποιος, αν θέλει να μειώσει την πιθανότητα να καταλήξει σε κάποιο σφάλμα που θα τον ταλαιπωρεί για μεγάλο χρονικό διάστημα.

Για τη δημιουργία της οντολογίας EuroCitizen, χρησιμοποιήθηκε το εργαλείο Protégé. Οι λόγοι για τους οποίους επιλέχθηκε το συγκεκριμένο εργαλείο είναι πολλοί. Καταρχάς έχει χρησιμοποιηθεί ήδη σε ένα μεγάλο αριθμό έργων [PROT07], αφήνοντας πολύ καλές εντυπώσεις στους χρήστες του. Ένας ακόμη λόγος είναι ότι το εργαλείο μπορεί να εκτελεστεί σε όλες τις πλατφόρμες που περιέχουν ένα περιβάλλον εκτέλεσης εφαρμογών Java (Java Runtime Environment – JRE). Αυτό ήταν απαραίτητο, επειδή η οντολογία αναπτύχθηκε σε μηχανήμα που έχει εγκατεστημένο το λειτουργικό σύστημα Debian GNU/Linux. Τέλος, δεν πρέπει να παραλειφθεί ότι το εργαλείο Protégé ανήκει στην κατηγορία του ελεύθερου λογισμικού [ΠΑΝΤ06], καθώς και το πλήθος των υπό-μονάδων (plugins) που παρέχει.

Παρακάτω παρατίθεται ένας συγκριτικός πίνακας των χαρακτηριστικών τριών δημοφιλών εργαλείων που παρέχονται σήμερα σε έναν κατασκευαστή οντολογιών με χρήση της τεχνολογίας OWL. Όπως μπορεί να παρατηρηθεί και στον πίνακα, ένα μεγάλο πλεονέκτημα που έχει η εφαρμογή Protégé έναντι των άλλων δύο, είναι η ισχυρή κοινότητα χρηστών που

διαθέτει. Αυτό το χαρακτηριστικό συνεπάγεται ότι η εφαρμογή είναι ευρείας αποδοχής, κάτι το οποίο παρέχει κάποια σιγουριά στον αρχάριο κατασκευαστή σχετικά με την αξιοπιστία της εφαρμογής. Επίσης, αυτό σημαίνει ότι τα ενδεχόμενα προβλήματα που αντιμετωπίζει ένας κατασκευαστής οντολογιών μπορούν να επιλυθούν γρήγορα με τη βοήθεια της εμπειρίας των υπόλοιπων κατασκευαστών/χρηστών.

Πίνακας 5-1: Σύγκριση τριών δημοφιλών εργαλείων σύνταξης οντολογιών ιστού.

Έκδοση	Άδεια Χρήσης	Κόστος	Κοινότητα Χρηστών	Υπομονάδες (plugins)	Πλατφόρμες	Τεκμηρίωση
Protégé 3.2	MPL	δωρεάν	μεγάλη	87	όλες οι Java-based	διαθέσιμη
Swoop 2.2.2	MIT	δωρεάν	μικρή	1	όλες οι Java-based	μη διαθέσιμη
SemanticWorks TM 2007	EULA	€ 99.00	μικρή	0	Microsoft Windows	διαθέσιμη
Υποστήριξη πολλαπλών reasoner		Απεικόνιση οντολογίας		Περιβάλλον χρήσης		Φόρτωση πολλαπλών οντολογιών
Ναι		μέτρια		καλό		Όχι
Όχι		καλή		μέτριο		Ναι
Όχι		καλή		πολύ καλό		Όχι

5.3 Η οντολογία ιστού EuroCitizen

Η οντολογία ιστού EuroCitizen αποτελεί μία απλουστευμένη μορφή μία οντολογίας που περιγράφει έναν Ευρωπαίο πολίτη. Όπως έχει αναφερθεί ήδη, έχει γραφεί σε έκδοση OWL DL, ώστε να είναι δυνατή η χρήση ενός μηχανισμού συλλογιστικής. Όλα τα χαρακτηριστικά της οντολογίας, είναι τεκμηριωμένα χρησιμοποιώντας σχόλια από διάφορες πηγές, η πλειοψηφία των οποίων προέρχεται από την οντολογία SUMO και το λεξικό WordNet [106]. Η τεκμηρίωση παρέχεται και σε μορφή HTML από το εργαλείο OWLDoc, σε μορφή παρόμοια με αυτή που εξάγει το εργαλείο javadoc για τη γλώσσα προγραμματισμού Java. Έγιναν προσπάθειες ώστε η οντολογία EuroCitizen να είναι μικρή σε μέγεθος (περίπου 107 KB),

γενική, και εύκολη στην επαναχρησιμοποίηση. Οι πληροφορίες που παρέχει η οντολογία, πάρθηκαν από επίσημες πηγές, όσο αυτό ήταν επιτρεπτό ([ΟΜΑΔ07], [WIKI07]).

5.3.1 Σκοπός κατασκευής της οντολογίας EuroCitizen

Ο απώτερος σκοπός κατασκευής της οντολογίας EuroCitizen, είναι η δημιουργία μίας Οντολογίας Ιστού για την περιγραφή των κοινών πεποιθήσεων που σχετίζονται και με το σύνολο των Ευρωπαίων πολιτών, αλλά και με συγκεκριμένους Ευρωπαίους πολίτες. Επομένως, στην τελική μορφή της, η οντολογία EuroCitizen θα είναι κατάλληλη για πολλές χρήσεις, όπως:

- ✓ **Η υποστήριξη λήψης αποφάσεων και εξαγωγής συμπερασμάτων.** Για παράδειγμα, έστω ότι ο Ευρωπαίος πολίτης x θέλει να κάνει ως δώρο-έκπληξη στον Ευρωπαίο πολίτη y , εισιτήρια για έναν αθλητικό αγώνα. Δεν γνωρίζει όμως, ποιο είναι το αγαπημένο άθλημα του y . Οπότε υποβάλλει το αίτημα στον πράκτορα του, ο οποίος “μαθαίνει” έπειτα από αναζήτηση στην οντολογία EuroCitizen, κάποια αθλήματα που θεωρούνται γενικά τα αγαπημένα των Ευρωπαίων. Ο πράκτορας προσπαθεί να “μάθει” ακόμη περισσότερα, αναζητώντας συγκεκριμένες πληροφορίες για τον χρήστη y . Σε περίπτωση που ανακαλύψει κάποια συγκεκριμένα αγαπημένα άθληματα του y , εμφανίζει τις προτιμήσεις του y στον x , καθώς και όλες τις πιθανές προτάσεις για κλείσιμο εισιτηρίων, πάντα σύμφωνα με τις υπόλοιπες παραμέτρους που έχει εισάγει ο x (ημερομηνία και ώρα, τιμή, κτλ). Ακόμη και σε περίπτωση που ο πράκτορας δεν καταφέρει να εντοπίσει συγκεκριμένες πληροφορίες για τον y , ακολουθώντας τις γενικές προτιμήσεις των Ευρωπαίων πολιτών, αυξάνει την πιθανότητα λήψης μίας σωστής απόφασης, που θα αφήσει ικανοποιημένο και τον x και τον y .
- ✓ **Η “εκπαίδευση” και η διαβούλευση ενός Ευρωπαίου πολίτη.** Ένας Ευρωπαίος πολίτης θα μπορεί να “εκπαιδευτεί” από την οντολογία EuroCitizen, μαθαίνοντας πολύτιμες πληροφορίες για την Ευρωπαϊκή νομοθεσία και τον Ευρωπαϊκό τρόπο ζωής, την Ευρωπαϊκή κουλτούρα, κτλ. Για παράδειγμα έστω ότι ο Ευρωπαίος πολίτης x θέλει να ταξιδέψει σε κάποια Ευρωπαϊκή χώρα, διαφορετική από αυτήν που κατοικεί. Ζητάει από τον πράκτορα του να του γνωστοποιήσει τα απαραίτητα έγγραφα που θα χρειαστεί (διαβατήριο, ταυτότητα, κτλ), ποια έγγραφα απαιτούνται για την έκδοση τους σε

περίπτωση που δεν τα διαθέτει (πιστοποιητικό γεννήσεως για ταυτότητα, αριθμό δημοτολογίου για διαβατήριο, κτλ), και το συνολικό κόστος των εγγράφων που θα προσκομιθούν (πχ 100 ευρώ για τα παράβολα του διαβατηρίου).

Στην τρέχουσα μορφή της, η οντολογία παρέχει μία πρότυπη και διαμοιραζόμενη αναπαράσταση των ευρωπαϊκών πεποιθήσεων. Συνεπώς, αποτελεί μία επαναχρησιμοποιήσιμη βάση γνώσης για χρήση από μία μεγάλη ποικιλία ειδών εφαρμογών.

Βέβαια, η οντολογία καθώς έχει απέχει αρκετά από την τελική μορφή της, αλλά το σημαντικό βήμα έχει ήδη γίνει, αφού υπάρχει η υποδομή πάνω στην οποία μπορούν να επενδύσουν μελλοντικοί σχεδιαστές.

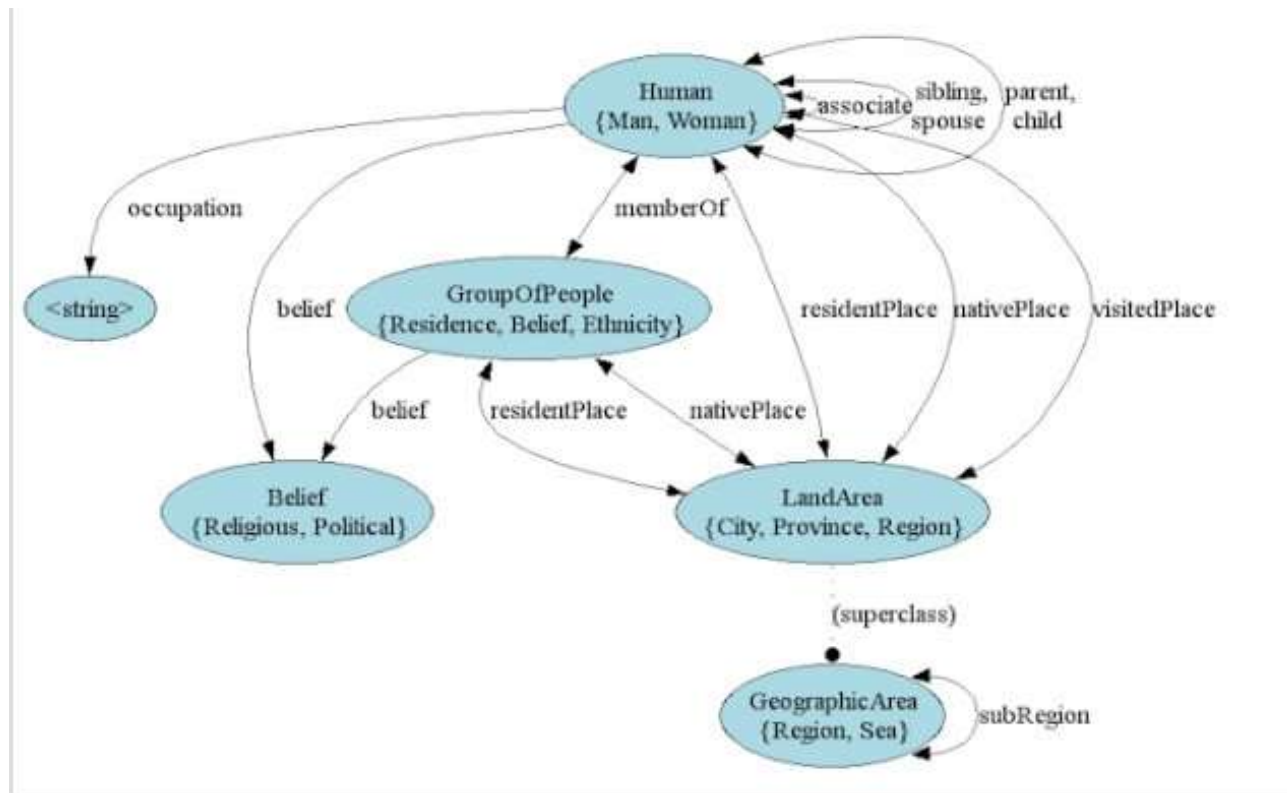
5.3.2 Ομοιότητες και διαφορές μεταξύ των οντολογιών EuroCitizen και NTN

Όπως έχει αναφερθεί ήδη, η οντολογία EuroCitizen αποτελεί επέκταση της οντολογίας NTN. Αυτό σημαίνει ότι η οντολογία EuroCitizen παρουσιάζει πολλές ομοιότητες με την οντολογία NTN. Παρόλα αυτά, υπάρχουν ορισμένες αλλαγές που κρίθηκε σκόπιμο να γίνουν, επειδή η οντολογία NTN σχετίζεται μόνο με ένα συγκεκριμένο κομμάτι (θρησκευτικό) που κλίνεται να καλύψει η οντολογία EuroCitizen. Συνεπώς, τα τμήματα της οντολογίας NTN που θεωρήθηκαν ακατάλληλα για χρήση στην οντολογία EuroCitizen έχουν αφαιρεθεί, και νέα τμήματα τα οποία θεωρήθηκαν αναγκαία έχουν προστεθεί.

Μία απεικόνιση των κλάσεων των δύο οντολογιών ίσως να δώσει στον αναγνώστη μία πληρέστερη εικόνα τους.

Στην εικόνα 5-1 φαίνεται ένα τμήμα των κλάσεων και των ιδιοτήτων της οντολογίας NTN, σε σχήμα που δίνεται από τον κατασκευαστή της. Το πλήρες τμήμα της προτιμήθηκε να μην απεικονιστεί, λόγω του υπερβολικά μεγάλου μεγέθους και της πολυπλοκότητας της εικόνας που προκύπτει. Το σημαντικό εδώ είναι να γίνει κατανοητό ότι η οντολογία NTN υποστηρίζει τη δόμηση γενικευμένων πληροφοριών που αφορούν οποιονδήποτε άνθρωπο ή ομάδα

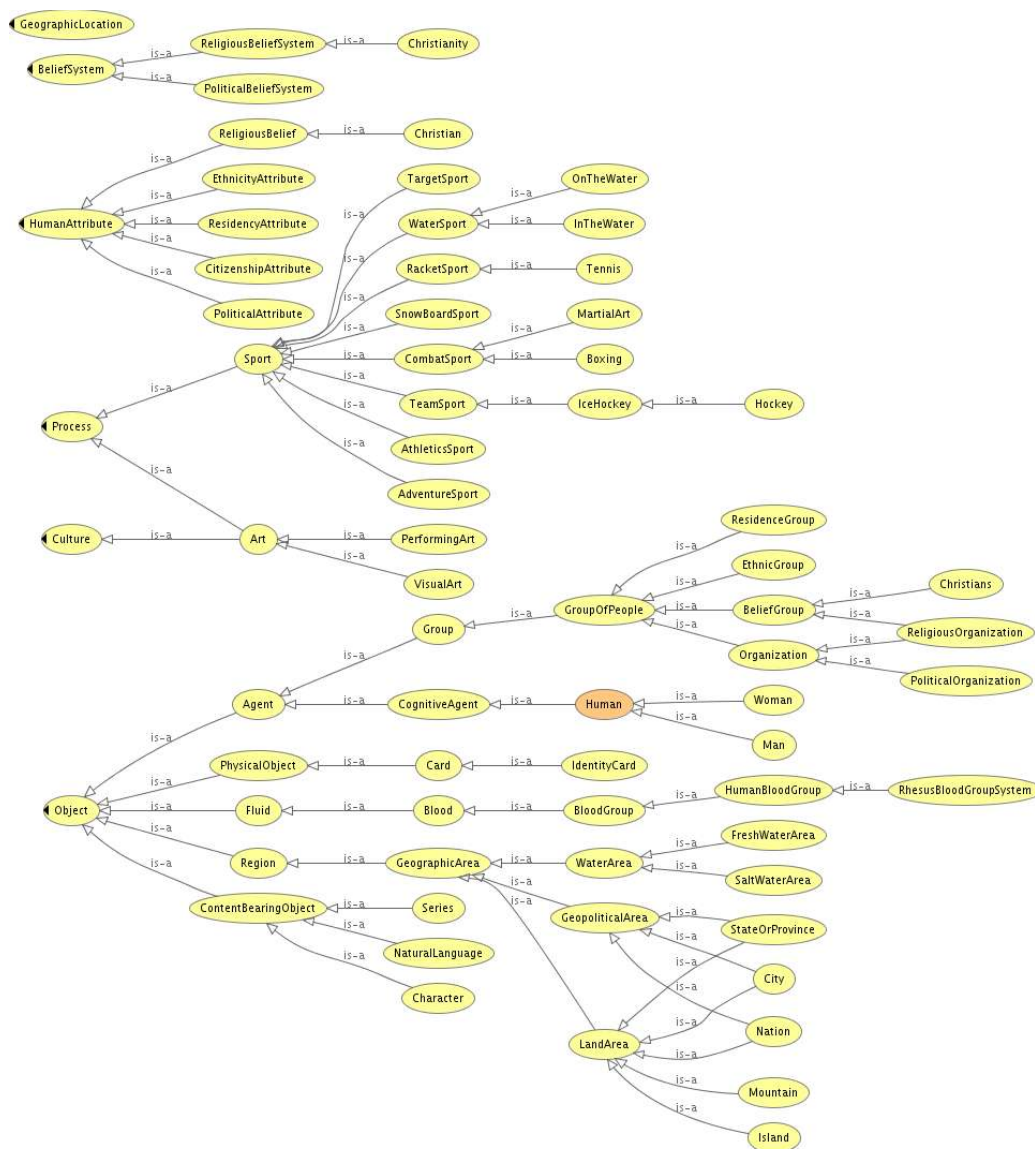
ανθρώπων. Για παράδειγμα, υποστηρίζει τη δημιουργία σχέσεων όπως η “ο x είναι άνθρωπος (Human), συγκεκριμένα άντρας (Man), και έχει επισκεφτεί (visitedPlace) την εδαφική περιοχή (landArea) Θεσσαλονίκη, η οποία είναι πόλη (City)”. Αυτό σημαίνει ότι απαιτούνται κάποιες ενέργειες για τον μετασχηματισμό της οντολογίας, ώστε να αφορά έναν Ευρωπαίο πολίτη. Για παράδειγμα, θα πρέπει η χώρα διαμονής ενός ανθρώπου να ανήκει αυστηρά και μόνο στην Ευρωπαϊκή ένωση, η ταυτότητα του να ακολουθεί αυστηρά τις Ευρωπαϊκές προδιαγραφές, κτλ. Επίσης, απαιτούνται περαιτέρω στοιχεία που θα παρέχουν πληροφορίες για έναν Ευρωπαίο πολίτη. Για παράδειγμα, τα θρησκευματα που είναι συνήθη στην Ευρώπη και πιθανόν να πιστεύει, τα δημοφιλέστερα αθλήματα που αρέσκονται να παρακολουθούν αλλά και να ασχολούνται οι Ευρωπαίοι, η κοινή κουλτούρα τους, κτλ.



Εικόνα 5-1: Τμήμα της οντολογίας NTN.

Στην εικόνα 5-2 φαίνεται η ιεραρχία των κλάσεων της οντολογίας EuroCitizen. Τα υπόλοιπα

τμήματα της οντολογίας (ιδιότητες, άτομα, κτλ) δεν απεικονίζονται, λόγω του υπερβολικά μεγάλου μεγέθους και της πολυπλοκότητας της εικόνας που προκύπτει.



Εικόνα 5-2: Οι κλάσεις της οντολογίας EuroCitizen.

Εύκολα μπορεί να παρατηρηθεί ότι στην οντολογία EuroCitizen έγινε προσπάθεια ώστε να παρέχονται πληροφορίες που είναι τόσο συγκεκριμένες, δηλαδή αφορούν έναν Ευρωπαίο πολίτη, όσο και γενικές, δηλαδή σχετίζονται με όλη την Ευρωπαϊκή κοινότητα και όχι μόνο με

κάποιο κομμάτι της. Την αυστηρότητα που απαιτεί η οντολογία EuroCitizen μπορεί και την παρέχει η OWL, μέσω των σχέσεων που υποστηρίζει. Επίσης, ο εμπλουτισμός της οντολογίας με τις νέες κλάσεις που απαιτούνται (Sport, Art, Card, κτλ) είναι επίσης εύκολος, αφού η OWL παρέχει τις κατάλληλες σχέσεις για τη δημιουργία απλών ως και ιδιαίτερα πολύπλοκων ιεραρχιών.

5.3.3 Χρήσεις της οντολογίας EuroCitizen

Η οντολογία EuroCitizen απευθύνεται σε μία μεγάλη γκάμα εφαρμογών και υπηρεσιών. Στην τελική μορφή της, θα είναι χρήσιμη σε:

- ✓ **δημόσιες και ιδιωτικές υπηρεσίες**, για εύρεση χρήσιμων πληροφοριών και εγγράφων σχετικά με την Ευρωπαϊκή πολιτική, νομοθεσία, κτλ,
- ✓ **ακαδημαϊκά ιδρύματα και ερευνητές**, για επέκταση της οντολογίας, ανάπτυξη λογισμικού που θα την χρησιμοποιεί, περαιτέρω έρευνα για την χρησιμότητα της και τις υπηρεσίες που μπορεί να προσφέρει, κτλ,
- ✓ **τελικούς χρήστες**, για την “εκπαίδευση” τους σε πολλά Ευρωπαϊκά ζητήματα (κοινωνικά, θρησκευτικά, πολιτικά, κτλ).

Στην τρέχουσα μορφή της, η οντολογία ιστού EuroCitizen είναι κατάλληλη για ορισμένες περιπτώσεις χρήσεως, ανάλογα με το αν συνοδεύεται με δεδομένα (άτομα) που αφορούν συγκεκριμένους Ευρωπαίους πολίτες ή όχι. Ορισμένες από αυτές είναι:

→ δίχως δεδομένα

→ αναζήτηση

- ✓ αγαπημένα αθλήματα,
- ✓ κοινή Ευρωπαϊκή κουλτούρα,
- ✓ καθομιλούμενες γλώσσες,
- ✓ περισσότερο διαδεδομένες θρησκείες,
- ✓ πιθανές εθνικότητες,

→ με δεδομένα

→ αναζήτηση

- ✓ καταγωγή του Ευρωπαίου πολίτη x,

- ✓ ποιες πόλεις έχει επισκεφτεί ο X,
- ✓ ομάδα αίματος του,
- απεικόνιση
 - ✓ σε ποια μέρη έχει συναναστραφεί με άλλους πολίτες ο X,
- κοινωνικά δίκτυα
 - ✓ ποιους Ευρωπαίους πολίτες γνωρίζει ο Ευρωπαίος πολίτης X.

Βέβαια πρέπει να γίνει ξεκάθαρο ότι η οντολογία EuroCitizen (και γενικότερα μία οντολογία ιστού) παρέχει απλά την υποδομή για χρήση όλων των παραπάνω υπηρεσιών. Για να υλοποιηθούν στην πράξη χρειάζεται να κατασκευαστεί επιπλέον λογισμικό, όπως σημασιολογικοί πράκτορες, λογισμικό απεικόνισης οντολογιών, κτλ.

5.3.4 Ελλείψεις και προβλήματα σχετικά με την οντολογία EuroCitizen

Όπως έχει αναφερθεί ήδη, η οντολογία ιστού EuroCitizen είναι ακόμη σε πρώιμο στάδιο, κάτι το οποίο σημαίνει ότι δεν υποστηρίζει πολλές από τις επιθυμητές υπηρεσίες. Συγκεκριμένα, η οντολογία δεν παρέχει πληροφορίες σχετικά με την Ευρωπαϊκή νομοθεσία και πολιτική, και αυτό συναινεί ώστε να μην μπορεί να χρησιμοποιηθεί για οτιδήποτε σχετίζεται με τέτοιου είδους ζητήματα. Τέτοιου είδους πληροφορίες σκόπιμα δεν ενσωματώθηκαν στην οντολογία, επειδή η γνώση που πρέπει να διαθέτει κάποιος για να προχωρήσει σε μία τέτοιου είδους ενέργεια ξεφεύγει από τα όρια μίας πτυχιακής εργασίας. Για να γίνει κάτι τέτοιο με βεβαιότητα ότι δεν θα υπάρχουν σφάλματα, χρειάζεται να προσληφθεί ένας άνθρωπος με βαθιές γνώσεις σε αυτά τα θέματα.

Επίσης, επειδή η οντολογία δεν έχει εφοδιαστεί με δεδομένα (άτομα), κάτι το οποίο έγινε επίσης σκόπιμα για να παραμείνει απλή και μικρή σε μέγεθος, δεν μπορεί να χρησιμοποιηθεί για υπηρεσίες που αφορούν συγκεκριμένους Ευρωπαίους πολίτες. Το να εφοδιαστεί με δεδομένα η οντολογία δεν είναι κάτι δύσκολο, αλλά είναι σίγουρα ιδιαίτερα χρονοβόρο.

Κατά τη διάρκεια της σχεδίασης της οντολογίας EuroCitizen, ένα από τα σημαντικότερα

προβλήματα που προέκυψαν είναι η δυσκολία εύρεσης δεδομένων που να αφορούν όλη την Ευρωπαϊκή κοινότητα. Έπειτα από χρονοβόρες αναζητήσεις με τη βοήθεια μηχανών και εγκεκριμένων Διαδικτυακών πυλών, όπως η Europa, βρέθηκαν κάποιες (λιγοστές) έρευνες και ορισμένα δεδομένα που κρίθηκαν χρήσιμα και ενσωματώθηκαν στην οντολογία.

Τέλος, όπως αναφέρθηκε και σε προηγούμενη παράγραφο, για την πλήρη εκμετάλλευση της οντολογίας, θα πρέπει να κατασκευαστεί επιπλέον λογισμικό υποστήριξης και να επεκταθεί η δομή της οντολογίας, με την προσθήκη επιπλέον κλάσεων, σχέσεων, ατόμων, κτλ.

Επίλογος / Συμπεράσματα

Σύμφωνα με τους Steffen Staab, Nigel Shadbolt, κ.α. [STAA06], η διάθεση, η ανάγκη, και η υποχρέωση των οργανισμών και των χρηστών του Διαδικτύου να διαθέσουν δημόσια τα δεδομένα τους, γίνεται ολοένα και μεγαλύτερη. Αυτό το συμπέρασμα εξάγεται εύκολα αν παρατηρηθεί ο εκρηκτικός ρυθμός αύξησης των ιστοσελίδων που εμφανίζονται καθημερινά στο Διαδίκτυο. Αυτή η αλλαγή είναι καλή επειδή η πληροφορία σήμερα είναι οικονομικά ακριβή και δυσεύρετη. Εύκολα αποδεικνύεται όμως (με μία απλή αναζήτηση) ότι ο παγκόσμιος ιστός πληροφοριών και οι τεχνολογίες που υποστηρίζουν τις υπηρεσίες του έχουν φτάσει σε επίπεδο κορεσμού.

Η απάντηση αναπτύσσεται ήδη και ονομάζεται σημασιολογικός ιστός. Ο σημασιολογικός ιστός επιδιώκει να χρησιμοποιήσει με ουσιαστικό τρόπο τις οντολογίες και τα δεδομένα που υπάρχουν ήδη στο Διαδίκτυο. Αρχικός στόχος του είναι ο εμπλουτισμός αυτών των οντολογιών και δεδομένων, ώστε να είναι ευκολότερο να τα ανακαλύψουν και να τα χρησιμοποιήσουν οι χρήστες, με εύκολο τρόπο (δίνοντας τις κατάλληλες παραμέτρους στους διαθέσιμους πράκτορες).

Επόμενος στόχος του είναι να δημιουργηθούν οντολογίες (θεμελιώδεις και περιοχής) που θα παρέχουν την κατάλληλη υποδομή για να αντικατασταθούν οι υπηρεσίες του παγκοσμίου ιστού με υπηρεσίες που θα χρησιμοποιούν σημασιολογικό περιεχόμενο (πχ σημασιολογικές μηχανές αναζήτησης, σημασιολογικές βάσεις δεδομένων, κτλ), και φυσικά οι κατάλληλοι πράκτορες που θα φροντίζουν να αυτοματοποιούν όλες τις σύνθετες διαδικασίες διευκολύνοντας τους χρήστες κατά το μέγιστο δυνατό βαθμό.

Στην εργασία καλύπτονται τα σημαντικότερα θέματα που αφορούν τον σημασιολογικό ιστό, και δίνεται μεγαλύτερη έμφαση στις κυριότερες τεχνολογίες του. Η τεχνολογία XML είναι μία γενικευμένη γλώσσα σήμανση, και χρησιμοποιείται για την δημιουργία γλωσσών σήμανσης ειδικού σκοπού. Επειδή η XML δεν εισάγει κανένα περιορισμό σχετικά με την δομή ενός εγγράφου, πολλές φορές είναι απαραίτητη της τεχνολογίας XML Schema, η οποία αποτελεί

ένα σύνολο περιορισμών ως προς τη δομή και το περιεχόμενο των εγγράφων XML. Ο RDF είναι η τεχνολογία που παρέχει την κατάλληλη υποδομή για την κωδικοποίηση, ανταλλαγή, και επαναχρησιμοποίηση των μεταδεδομένων, δηλαδή των δομημένων δεδομένων που χρησιμοποιούνται για την περιγραφή άλλων δεδομένων. Υπάρχουν όμως περιπτώσεις, στις οποίες οι απλοποιημένες δηλώσεις που επιτρέπει να ορίζονται ο RDF, δεν είναι αρκετές, και εμφανίζεται η ανάγκη δημιουργίας απλουστευμένων λεξιλογίων όρων, τα οποία απαιτείται να ενσωματώνουν κλάσεις εφαρμογών με ιδιότητες και αυστηρή ιεραρχία. Σε αυτές τις περιπτώσεις, απαιτείται η χρήση της τεχνολογίας RDF Schema. Τέλος, σε περίπτωση που είναι αναγκαία η δημιουργία ενός σύνθετου λεξιλογίου όρων, το οποίο υποστηρίζει τη δημιουργία πολύπλοκων συσχετίσεων, την επιβολή αυστηρών περιορισμών στις ιδιότητες, τη δημιουργία σύνθετων κλάσεων, κτλ., η τεχνολογία OWL είναι η καταλληλότερη.

Οι κοινότητες χρηστών και ερευνητών έχουν αναπτύξει πολλές θεμελιώδεις οντολογίες και οντολογίες περιοχής με τη βοήθεια της OWL τα τελευταία χρόνια. Παρόλα αυτά, καμία από αυτές τις οντολογίες δεν σχετίζεται με τον Ευρωπαίο πολίτη. Μία τέτοιου είδους οντολογία θεωρείται πολύ χρήσιμη, επειδή μπορεί να βοηθήσει στην εκμετάλλευση όλων των πλεονεκτημάτων που παρέχει ο σημασιολογικός ιστός σε σύγκριση με τον κλασικό ιστό, σε πολλές περιοχές εφαρμογών (ακαδημαϊκό επίπεδο, ερευνητικό επίπεδο, δημόσιες και ιδιωτικές υπηρεσίες, τελικούς χρήστες, κτλ), με διάφορους τρόπους (υποστήριξη λήψης αποφάσεων και εξαγωγής συμπερασμάτων δίχως σφάλματα, “εκπαίδευση” και διαβούλευση ενός Ευρωπαίου πολίτη, αναζήτηση, απεικόνιση, δημιουργία κοινωνικών δικτύων). Αυτοί είναι οι τελικοί στόχοι της οντολογίας EuroCitizen, η οποία όμως προς το παρόν βρίσκεται σε πρώιμο στάδιο, και παρέχει μόνο ορισμένες από τις παραπάνω υπηρεσίες, όπως η αναζήτηση, καθώς και η υποστήριξη λήψης αποφάσεων και εξαγωγής συμπερασμάτων.

Με τις γνώσεις που έχει αποκτήσει ο αναγνώστης διαβάζοντας αυτήν την εργασία, ενθαρρύνεται ώστε να γίνει ενεργό μέλος της κοινότητας επέκτασης του παγκοσμίου ιστού σε σημασιολογικό ιστό. Έτσι θα βοηθήσει ώστε να μετατραπεί το χαώδες Διαδίκτυο σε μία πολύ καλά δομημένη δημόσια βάση γνώσης, με αξιόπιστες και εύχρηστες υπηρεσίες.

Βιβλιογραφία

- [ΒΛΑΧ06] Ι. Βλαχάβας, Π. Κεφαλάς, Ν. Βασιλειάδης, Φ. Κόκκορας, Η. Σακελλαρίου. *Τεχνητή Νοημοσύνη - Β' Έκδοση*, ISBN: 960-7013-35-2, Εκδόσεις Γαρταγάνη, 2005.
- [ΚΟΥΡ02] Φώτης Κουρέτας. "XML σε 10 σημεία". Ανακτήθηκε από το Διαδίκτυο: <http://www.w3c.gr/office/docs/XML10points.el.htm>.
- [ΟΜΑΔ07] Ομάδα δικτυακού συντονισμού πύλης EUROPA. Ανακτήθηκε από το Διαδίκτυο: http://europa.eu/index_el.htm.
- [ΠΑΝΤ06] Βασίλης Παντής. "Ο Ορισμός του Ελεύθερου Λογισμικού". Ανακτήθηκε από το Διαδίκτυο: <http://www.gnu.org/philosophy/free-sw.el.html>.
- [BAIL05] James Bailey, François Bry, Tim Furche, and Sebastian Schaffert. "Web and Semantic Web Query Languages: A Survey". Ανακτήθηκε από το Διαδίκτυο: http://reasoningweb.org/2005/teaching-material/bry-francois_bailey-james_web-and-semantic-web-query-languages.html.
- [BATE95] John A. Bateman, Renate Henschel, Fabio Rinaldi. "The Generalized Upper Model 2.0". Ανακτήθηκε από το Διαδίκτυο: <http://www.fb10.uni-bremen.de/anglistik/langpro/webpace/jb/gum/index.htm>.
- [BHAT04] Mehul Bhatt, Carlo Wouters, Andrew Flahive, Wenny Rahayu, and David Taniar. "Semantic completeness in sub-ontology extraction using distributed methods." *ICCSA 2004*.
- [BIRO01] Paul V. Biron, Ashok Malhotra. "XML Schema Part 2: Datatypes". Ανακτήθηκε από το Διαδίκτυο: <http://www.w3.org/TR/2001/REC-xmlschema-2-20010502/#built-in-datatypes>.
- [BOIS04] Sean Boisen. "New Testament Names: a Semantic Knowledge Base". Ανακτήθηκε από το Διαδίκτυο: <http://www.semanticbible.com/ntn/ntn-overview.html>.
- [BOIS05] Sean Boisen. "View and Download NT Names". Ανακτήθηκε από το Διαδίκτυο: <http://www.semanticbible.com/ntn/ntn-view.html>.
- [BRIC04] Dan Brickley, R.V. Guha. "RDF Vocabulary Description Language 1.0: RDF Schema". Ανακτήθηκε από το Διαδίκτυο: <http://www.w3.org/TR/rdf-schema/>.
- [CONN01] Dan Connolly, Frank van Harmelen, Ian Horrocks, Deborah L. McGuinness, Peter F. Patel-Schneider, Lynn Andrea Stein. "DAML+OIL (March 2001) Reference Description". Ανακτήθηκε από το Διαδίκτυο: <http://www.w3.org/TR/daml+oil-reference>.

- [CORC00]** Óscar Corcho, Asunción Gómez-Pérez. “A Roadmap to Ontology Specification Languages”. *Proceedings of the 12th European Workshop on Knowledge Acquisition, Modeling and Management*.
- [COVE05]** Robin Cover. “Business Rules and Web Architecture: W3C Creates Rule Interchange Format WG”. Ανακτήθηκε από το Διαδίκτυο: <http://xml.coverpages.org/ni2005-11-09-a.html>.
- [CLAR05]** Peter Clark. “Some Ongoing KBS/Ontology Projects and Groups”. Ανακτήθηκε από το Διαδίκτυο: <http://www.cs.utexas.edu/users/mfkb/related.html#a>.
- [CUNN06]** Cunningham & Cunningham, Inc. “XML Sucks”. Ανακτήθηκε από το Διαδίκτυο: <http://c2.com/cgi/wiki?XmlSucks>.
- [DAML04]** DAML ontology librarian. “DAML Ontology Library”. Ανακτήθηκε από το Διαδίκτυο: <http://www.daml.org/ontologies/>.
- [DECK00]** Stefan Decker, Dieter Fensel, Frank van Harmelen, Ian Horrocks, Sergey Melnik, Michel Klein and Jeen Broekstra. “Knowledge Representation on the Web”. *Proceedings of the International Workshop on Description Logics (DL2000)*.
- [DIMI01]** Theo Dimitrakos, Brian Matthews, Juan Bicarregui. “Towards Supporting Security and Trust Management Policies on the Web”. *ERCIM workshop 'The Role of Trust in e-Business' in conjunction with IFIP I3E conference, 2001*.
- [FENS03]** Dieter Phensel, Christoph Bussler, “Semantic Web enabled Web Services”. *Proceedings of the International Semantic Web Conference 2002, Sardinia, Italy*.
- [GINS06]** Allen Ginsberg, David Hirtle, Frank McCabe, Paula-Lavinia. “RIF Use Cases and Requirements”. Ανακτήθηκε από το Διαδίκτυο: <http://www.w3.org/TR/rif-ucr/>.
- [GOLB03]** Jennifer Golbeck, Bijan Parsia, James Hendler. “Trust Networks on the Semantic Web”. *Proceedings of Cooperative Intelligent Agents 2003, Helsinki, Finland, August 27-29*.
- [GRUB93]** Tom Gruber. “What is an ontology?”. Ανακτήθηκε από το Διαδίκτυο: <http://www-ksl.stanford.edu/kst/what-is-an-ontology.html>.
- [HEND04]** Jim Hendler. “Frequently Asked Questions on W3C's Web Ontology Language (OWL)”. Ανακτήθηκε από το Διαδίκτυο: <http://www.w3.org/2003/08/owlfaq.html>.
- [HERM07]** Ivar Herman. “Semantic Web”. Ανακτήθηκε από το Διαδίκτυο: <http://www.w3.org/2001/sw/>.
- [HORR00]** Ian Horrocks, Dieter Fensel, Jeen Broekstra, Stefan Decker, Michael Erdmann, Carole Goble, Frank van Harlemen, Michel Klein, Steffen Staab, Rudi Studer, Enrico Motta. “The Ontology Inference Layer OIL”. Ανακτήθηκε από το Διαδίκτυο: <http://www.ontoknowledge.org/oil/TR/oil.long.html>.
- [INTEL04]** Intellidimension. “A Search Engine for the Semantic Web (BETA)”. Ανακτήθηκε από το Διαδίκτυο: <http://www.semanticwebsearch.com/>.

- [KASA05]** Toshinobu Kasai, Haruhisa Yamaguchi. "A Semantic Web System for Helping Teachers Plan Lessons Using Ontology Alignment". *SW-EL*, 18 July 2005.
- [KIFE02]** Michael Kifer, David Martin. "Bringing Services to the Semantic Web and Semantics to Web Services". Ανακτήθηκε από το Διαδίκτυο: <http://www.swsi.org/resources/swsc-lc-intro.pdf>.
- [KNUB04]** Holger Knublauch, Alan Rector, Robert Stevens, Chris Wroe. "A Practical Guide To Building OWL Ontologies Using The Protégé-OWL Plugin and CO-ODE Tools Edition 1.0". Ανακτήθηκε από το Διαδίκτυο: <http://www.co-ode.org/resources/tutorials/ProtegeOWLTutorial.pdf>.
- [KNUB06]** Holger Knublauch, Daniel Oberle, Phil Tetlow, Evan Wallace. "A Semantic Web Primer for Object-Oriented Software Developers". Ανακτήθηκε από το Διαδίκτυο: <http://www.w3.org/TR/sw-oosd-primer/>.
- [LEE01]** Tim Berners-Lee, James Hendler and Ora Lassila. "The Semantic Web". *Scientific American*, May 2001.
- [LEE06]** Tim Berners-Lee. "Artificial Intelligence and the Semantic Web". *AAAI*, 18 July 2006.
- [LEO06]** Leo Obrst. "Future directions for ontologies and terminologies". Ανακτήθηκε από το Διαδίκτυο: <http://colab.cim3.net/forum/ontac-forum/2006-03/msg00006.html>.
- [LOPE06]** Xavier Lopez, Susie Stephens. "Semantic Data Integration for the Enterprise". Ανακτήθηκε από το Διαδίκτυο: http://www.oracle.com/technology/tech/semantic_technologies/pdf/semantic_grid_wp_0603.pdf.
- [LUTZ07]** Carsten Lutz. "DESCRIPTION LOGICS". Ανακτήθηκε από το Διαδίκτυο: <http://dl.kr.org/>.
- [MANO04]** Frank Manola, Eric Miller, Brian McBride. "RDF Primer". Ανακτήθηκε από το Διαδίκτυο: <http://www.w3.org/TR/rdf-primer/>.
- [MASO03]** Claudio Masolo, Stefano Borgo, Aldo Gangemi, Nicola Guarino, Alessandro Oltramari, Luc Schneider. "The WonderWeb Library of Foundational Ontologies Preliminary Report". Ανακτήθηκε από το Διαδίκτυο: <http://www.loa-cnr.it/Papers/DOLCE2.1-FOL.pdf>.
- [MCGU04]** Deborah L. McGuinness, Frank van Harmelen. "OWL Web Ontology Language Overview". Ανακτήθηκε από το Διαδίκτυο: <http://www.w3.org/TR/owl-features/>.
- [MILL98]** Eric Miller. "An Introduction to the Resource Description Framework". *D-Lib Magazine*, May 1998.
- [NOY00]** Natalya F. Noy and Deborah L. McGuinness. "Ontology Development 101: A Guide to Creating Your First Ontology". Ανακτήθηκε από το Διαδίκτυο: http://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html.
- [MURR05]** Chuck Murray. "Oracle® Spatial Resource Description Framework (RDF) 10g Release 2 (10.2)". Ανακτήθηκε από το Διαδίκτυο: http://download.oracle.com/otndocs/tech/semantic_web/pdf/rdfm.pdf.

- [PATE04]** Peter F. Patel-Schneider, Patrick Hayes, Ian Horrocks. "OWL Web Ontology Language Semantics and Abstract Syntax". Ανακτήθηκε από το Διαδίκτυο: <http://www.w3.org/TR/owl-semantics/>.
- [PROT06]** Protégé Community Wiki. "ProjectsThatUseProtege". Ανακτήθηκε από το Διαδίκτυο: <http://protege.cim3.net/cgi-bin/wiki.pl?ProjectsThatUseProtege>.
- [PROT07]** Protégé Community Wiki. "ProtegeOntologiesLibrary". Ανακτήθηκε από το Διαδίκτυο:
- [PRUD06]** Eric Prud'hommeaux, Andy Seaborne. "SPARQL Query Language for RDF". Ανακτήθηκε από το Διαδίκτυο: <http://www.w3.org/TR/rdf-sparql-query/>.
- [PRUD07]** Eric Prud'hommeaux. "RDF Data Access Working Group". Ανακτήθηκε από το Διαδίκτυο: <http://www.w3.org/2001/sw/DataAccess/>.
- [SCHE06]** SchemaWeb. "A directory of RDF schemas expressed in the RDFS, OWL and DAML+OIL schema languages". Ανακτήθηκε από το Διαδίκτυο: <http://www.schemaweb.info/default.aspx>.
- [SMIT04]** Michael K. Smith, Chris Welty, Deborah L. McGuinness. "OWL Web Ontology Language Guide". Ανακτήθηκε από το Διαδίκτυο: <http://www.w3.org/TR/owl-guide/>.
- [STAA00]** Steffen Staab, Michael Erdmann, Alexander Maedche, Stefan Decker. "An Extensible Approach for Modeling Ontologies in RDF(S)". *ECDL 2000 Workshop on the Semantic Web, 21 September 2000, Lisbon Portugal*.
- [STAA06]** Steffen Staab, Nigel Shadbolt, Wendy Hall, Tim Berners-Lee. "The Semantic Web Revisited". IEEE Intelligent Systems, January/February 2006, σελ. 96-101.
- [STEP04]** David Stephenson. "XML fundamentals checklist". Ανακτήθηκε από το Διαδίκτυο: <http://devresource.hp.com/drc/resources/xmlfundCklist/index.jsp>.
- [STEP06]** Susie Stephens. "Semantic Web Commercial Products". Ανακτήθηκε από το Διαδίκτυο: <http://esw.w3.org/topic/CommercialProducts>.
- [TAUB06]** Joshua Tauberer. "What is RDF". Ανακτήθηκε από το Διαδίκτυο: <http://www.xml.com/pub/a/2001/01/24/rdf.html>.
- [UNDE05]** Andrew Undergrove. "The Semantic Web: An interview with Tim Berners-Lee". Ανακτήθηκε από το Διαδίκτυο: <http://www.consortiuminfo.org/bulletins/semanticweb.php>.
- [USZO04]** Andrzej Uszok, Jeffrey M. Bradshaw, Matthew Johnson, Renia Jeffers, Austin Tate, Jeff Dalton, and Stuart Aitken. "KAoS Policy Management for Semantic Web Services". IEEE Intelligent Systems, July/August 2004, σελ. 32- 41.
- [WICH06]** Wikipedia. "Character encoding". Ανακτήθηκε από το Διαδίκτυο: http://en.wikipedia.org/wiki/Character_encoding.
- [WICY06]** Wikipedia. "Cyc". Ανακτήθηκε από το Διαδίκτυο: <http://en.wikipedia.org/wiki/OpenCyc>.
- [WIDA06]** Wikipedia. "Directed acyclic graph". Ανακτήθηκε από το Διαδίκτυο: http://en.wikipedia.org/wiki/Directed_acyclic_graph.

- [WIDU06]** Wikipedia. "Dublin Core". Ανακτήθηκε από το Διαδίκτυο:
http://en.wikipedia.org/wiki/Dublin_Core.
- [WIGE06]** Wikipedia. "General Formal Ontology". Ανακτήθηκε από το Διαδίκτυο:
http://en.wikipedia.org/wiki/General_Formal_Ontology.
- [WIGO06]** Wikipedia. "Gene Ontology". Ανακτήθηκε από το Διαδίκτυο:
http://en.wikipedia.org/wiki/Gene_ontology.
- [WIIN06]** Wikipedia. "Internationalized Resource Identifier". Ανακτήθηκε από το Διαδίκτυο: http://en.wikipedia.org/wiki/Internationalized_Resource_Identifier.
- [WIKI07]** Wikipedia. "Welcome to Wikipedia". Ανακτήθηκε από το Διαδίκτυο:
<http://en.wikipedia.org/>.
- [WIME06]** Wikipedia. "Meronymy". Ανακτήθηκε από το Διαδίκτυο:
<http://en.wikipedia.org/wiki/Meronymy>.
- [WIMU06]** Wikipedia. "Multiple inheritance". Ανακτήθηκε από το Διαδίκτυο:
http://en.wikipedia.org/wiki/Multiple_inheritance.
- [WION06]** Wikipedia. "Ontology Language (computer science)". Ανακτήθηκε από το Διαδίκτυο: http://en.wikipedia.org/wiki/Ontology_Languages_%28computer_science%29.
- [WION07]** Wikipedia. "Ontology (computer science)". Ανακτήθηκε από το Διαδίκτυο:
http://en.wikipedia.org/wiki/Ontology_%28computer_science%29.
- [WIRD06]** Wikipedia. "RDF Schema". Ανακτήθηκε από το Διαδίκτυο:
http://en.wikipedia.org/wiki/RDF_Schema.
- [WIRE06]** Wikipedia. "Regular Language description for XML". Ανακτήθηκε από το Διαδίκτυο: <http://en.wikipedia.org/wiki/RELAX>.
- [WIRU06]** Wikipedia. "Rule of inference". Ανακτήθηκε από το Διαδίκτυο:
http://en.wikipedia.org/wiki/Rule_of_inference.
- [WISC06]** Wikipedia. "Schematron". Ανακτήθηκε από το Διαδίκτυο:
<http://en.wikipedia.org/wiki/Schematron>.
- [WISE06]** Wikipedia. "Semantic Web". Ανακτήθηκε από το Διαδίκτυο:
http://en.wikipedia.org/wiki/Semantic_web.
- [WISO07]** Wikipedia. "Social Security number". Ανακτήθηκε από το Διαδίκτυο:
http://en.wikipedia.org/wiki/Social_security_number.
- [WISU06]** Wikipedia. "Suggested Upper Merged Ontology". Ανακτήθηκε από το Διαδίκτυο:
http://en.wikipedia.org/wiki/Suggested_Upper_Merged_Ontology.
- [WITR05]** Wikipedia. "TREX". Ανακτήθηκε από το Διαδίκτυο:
<http://en.wikipedia.org/wiki/TREX>.
- [WIUN06]** Wikipedia. "Uniform Resource Identifier". Ανακτήθηκε από το Διαδίκτυο:
http://en.wikipedia.org/wiki/Uniform_Resource_Identifier.
- [WIWO06]** Wikipedia. "WordNet". Ανακτήθηκε από το Διαδίκτυο:
<http://en.wikipedia.org/wiki/WordNet>.

- [WLOW06]** Wikipedia. “Web Ontology Language”. Ανακτήθηκε από το Διαδίκτυο:
http://en.wikipedia.org/wiki/Web_Ontology_Language.
- [WIXM06]** Wikipedia. “XML Schema (W3C)”. Ανακτήθηκε από το Διαδίκτυο:
http://en.wikipedia.org/wiki/XML_Schema.
- [WIXM07]** Wikipedia. “XML”. Ανακτήθηκε από το Διαδίκτυο:
<http://en.wikipedia.org/wiki/XML>.
- [WIXS06]** Wiki Books. “XML Schema”. Ανακτήθηκε από το Διαδίκτυο:
http://en.wikibooks.org/wiki/XML_Schema.
- [WIXS07]** Wikipedia. “XML schema”. Ανακτήθηκε από το Διαδίκτυο:
http://en.wikipedia.org/wiki/XML_schema.

Παράρτημα Α - Παραδείγματα εγγράφων

Παράδειγμα Α-1: Η απλούστερη μορφή ενός εγγράφου HTML.

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN"
"http://www.w3.org/TR/html4/strict.dtd">
<html lang="utf-8">

  <head>

    <title>Τίτλος εγγράφου</title>
    <meta http-equiv="Content-Type" content="text/html; charset=utf-8"
>
  </head>

  <body>

    <h1>Αυτό είναι ένα έγγραφο HTML.</h1>

  </body>

</html>
```

Παράδειγμα Α-2: Περιγραφή μίας υπενθύμισης με έγγραφο XML.

```
<?xml version="1.0" encoding="ISO-8859-1" standalone="yes"?>
<!DOCTYPE note [
  <!ELEMENT note (to, from, heading, body)>
  <!ELEMENT to (#PCDATA)>
  <!ELEMENT from (#PCDATA)>
```

```

<!ELEMENT heading (#PCDATA)>
<!ELEMENT body (#PCDATA)>
]>
<note>
  <to>Tove</to>
  <from>Jani</from>
  <heading>Reminder</heading>
  <body>Don't forget me this weekend!</body>
</note>

```

Παράδειγμα A-3: Η δομή ενός στοιχείου σε WXS για τη δήλωση χωρών.

```

<xs:schema
  xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="country" type="Country"/>
  <xs:complexType name="Country">
    <xs:sequence>
      <xs:element name="name" type="xs:string"/>
      <xs:element name="population" type="xs:decimal"/>
    </xs:sequence>
  </xs:complexType>
</xs:schema>

```

Παράδειγμα A-4: Χρήση του σχήματος WXS του παραδείγματος A-3.

```

<country
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:noNamespaceSchemaLocation="A-3.xsd">
  <name>France</name>
  <population>59.7</population>
</country>

```

Παράδειγμα A-5: Το γράφημα της εικόνας 4-1 με σύνταξη σημειογραφίας 3.

```

@prefix contact: <http://www.w3.org/2000/10/swap/pim/contact#> .

```

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .

<http://www.w3.org/People/EM/contact#me> rdf:type contact:Person ;
contact:fullName "Eric Miller" ;
contact:mailbox <mailto:em@w3.org> ;
contact:personalTitle "Dr." .
```

Παράδειγμα A-6: Το γράφημα της εικόνας 4-1 σε κωδικοποίηση RDF/XML.

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
        xmlns:contact="http://www.w3.org/2000/10/swap/pim/contact#">

  <contact:Person rdf:about="http://www.w3.org/People/EM/contact#me">
    <contact:fullName>Eric Miller</contact:fullName>
    <contact:mailbox rdf:resource="mailto:em@w3.org"/>
    <contact:personalTitle>Dr.</contact:personalTitle>
  </contact:Person>

</rdf:RDF>
```

Παράδειγμα A-7: Το γράφημα της εικόνας 4-2 με σύνταξη σημειογραφίας στηλών.

veh:MotorVehicle	rdf:type	rdfs:Class .
veh:PassengerVehicle	rdf:type	rdfs:Class .
veh:Van	rdf:type	rdfs:Class .
veh:Truck	rdf:type	rdfs:Class .
veh:MiniVan	rdf:type	rdfs:Class .
veh:PassengerVehicle	rdfs:subClassOf	veh:MotorVehicle .
veh:Van	rdfs:subClassOf	veh:MotorVehicle .
veh:Truck	rdfs:subClassOf	veh:MotorVehicle .
veh:MiniVan	rdfs:subClassOf	veh:Van .

```
veh:MiniVan      rdfs:subClassOf    veh:PassengerVehicle .
```

Παράδειγμα A-8: Το γράφημα της εικόνας 4-2 σε RDFS (κωδικοποίηση RDF/XML).

```
<?xml version="1.0"?>
<!DOCTYPE rdf:RDF [<!ENTITY xsd "http://www.w3.org/2001/XMLSchema#">]
>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xml:base="http://example.org/schemas/vehicles">

  <rdfs:Class rdf:ID="MotorVehicle"/>

  <rdfs:Class rdf:ID="PassengerVehicle">
    <rdfs:subClassOf rdf:resource="#MotorVehicle"/>
  </rdfs:Class>

  <rdfs:Class rdf:ID="Truck">
    <rdfs:subClassOf rdf:resource="#MotorVehicle"/>
  </rdfs:Class>

  <rdfs:Class rdf:ID="Van">
    <rdfs:subClassOf rdf:resource="#MotorVehicle"/>
  </rdfs:Class>

  <rdfs:Class rdf:ID="MiniVan">
    <rdfs:subClassOf rdf:resource="#Van"/>
    <rdfs:subClassOf rdf:resource="#PassengerVehicle"/>
  </rdfs:Class>

</rdf:RDF>
```

Παράδειγμα A-9: Το πλήρες σχήμα RDFS της εικόνας 4-2 (κωδικοποίηση RDF/XML).

```
<?xml version="1.0"?>
<!DOCTYPE rdf:RDF [<!ENTITY xsd "http://www.w3.org/2001/XMLSchema#">]
>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xml:base="http://example.org/schemas/vehicles">

  <rdfs:Class rdf:ID="MotorVehicle"/>

  <rdfs:Class rdf:ID="PassengerVehicle">
    <rdfs:subClassOf rdf:resource="#MotorVehicle"/>
  </rdfs:Class>

  <rdfs:Class rdf:ID="Truck">
    <rdfs:subClassOf rdf:resource="#MotorVehicle"/>
  </rdfs:Class>

  <rdfs:Class rdf:ID="Van">
    <rdfs:subClassOf rdf:resource="#MotorVehicle"/>
  </rdfs:Class>

  <rdfs:Class rdf:ID="MiniVan">
    <rdfs:subClassOf rdf:resource="#Van"/>
    <rdfs:subClassOf rdf:resource="#PassengerVehicle"/>
  </rdfs:Class>

  <rdfs:Class rdf:ID="Person"/>

  <rdfs:Datatype rdf:about="&xsd;integer"/>
```

```

<rdf:Property rdf:ID="registeredTo">
  <rdfs:domain rdf:resource="#MotorVehicle"/>
  <rdfs:range rdf:resource="#Person"/>
</rdf:Property>

<rdf:Property rdf:ID="rearSeatLegRoom">
  <rdfs:domain rdf:resource="#PassengerVehicle"/>
  <rdfs:range rdf:resource="&xsd;integer"/>
</rdf:Property>

<rdf:Property rdf:ID="driver">
  <rdfs:domain rdf:resource="#MotorVehicle"/>
</rdf:Property>

<rdf:Property rdf:ID="primaryDriver">
  <rdfs:subPropertyOf rdf:resource="#driver"/>
</rdf:Property>

</rdf:RDF>

```

Παράδειγμα A-10: Η μορφή ενός αιτήματος SPARQL.

```

SELECT ?X WHERE
?X rdf:type owl:Class .

```

Η παραπάνω είναι μία απλούστατη μορφή ενός αιτήματος SPARQL. Όταν υποβάλλεται σε μία οντολογία, εμφανίζει ως αποτέλεσμα όλα τα στοιχεία της που συνοδεύονται από την τιμή `owl:Class` στην ιδιότητα `rdf:type`, δηλαδή όλες τις κλάσεις OWL της οντολογίας.

Ένα τμήμα των αποτελεσμάτων έπεται από υποβολή του παραπάνω αιτήματος στην οντολογία EuroCitizen, φαίνεται στην εικόνα A-1. Η συγγραφή και υποβολή του αιτήματος έγινε μέσα από το εργαλείο Protégé.

● EthnicGroup
■ Human or GroupOfPeople
● MartialArt
■ Human or GroupOfPeople
● SaltWaterArea
● Island
● IceHockey
● InTheWater
■ Human or GroupOfPeople
● NaturalLanguage
■ Object or BeliefSystem or HumanAttribute
● FreshWaterArea
● Nation
● BeliefGroup
● WaterArea
● Hockey
■ Human
● BloodGroup

Εικόνα A-1: Τμήμα των αποτελεσμάτων SPARQL της οντολογίας EuroCitizen.

Παράρτημα Β – Τμήματα κώδικα εγγράφων

Παράδειγμα Β-1: Επέκταση του σχήματος οχημάτων με την προσθήκη ιδιοτήτων RDFS.

```
<rdf:Property rdf:ID="registeredTo">
  <rdfs:domain rdf:resource="#MotorVehicle"/>
  <rdfs:range rdf:resource="#Person"/>
</rdf:Property>

<rdf:Property rdf:ID="rearSeatLegRoom">
  <rdfs:domain rdf:resource="#PassengerVehicle"/>
  <rdfs:range rdf:resource="&xsd;integer"/>
</rdf:Property>

<rdfs:Class rdf:ID="Person"/>

<rdfs:Datatype rdf:about="&xsd;integer"/>
```

Παράδειγμα Β-2: Οι τελευταίες ιδιότητες RDFS για το πλήρες σχήμα οχημάτων.

```
<rdf:Property rdf:ID="driver">
  <rdfs:domain rdf:resource="#MotorVehicle"/>
</rdf:Property>

<rdf:Property rdf:ID="primaryDriver">
  <rdfs:subPropertyOf rdf:resource="#driver"/>
</rdf:Property>
```

Παράδειγμα Β-3: Δήλωση τριών ριζικών κλάσεων περιοχής.

```
<owl:Class rdf:ID="Winery"/>
<owl:Class rdf:ID="Region"/>
<owl:Class rdf:ID="ConsumableThing"/>
```

Παράδειγμα B-4: Δήλωση ιδιότητας αντικειμένου.

```
<owl:ObjectProperty rdf:ID="madeFromGrape">
  <rdfs:domain rdf:resource="#Wine"/>
  <rdfs:range rdf:resource="#WineGrape"/>
</owl:ObjectProperty>
```

Παράδειγμα B-5: Δήλωση ιδιότητας τύπου δεδομένων.

```
<owl:Class rdf:ID="VintageYear" />

<owl:DatatypeProperty rdf:ID="yearValue">
  <rdfs:domain rdf:resource="#VintageYear" />
  <rdfs:range rdf:resource="&xsd;positiveInteger"/>
</owl:DatatypeProperty>
```

Παράδειγμα B-6: Χρήση της ιδιότητας owl:equivalentClass.

```
<owl:Class rdf:ID="Wine">
  <owl:equivalentClass rdf:resource="&vin;Wine"/>
</owl:Class>
```

Παράδειγμα B-7: Χρήση της ιδιότητας owl:sameAs.

```
<Wine rdf:ID="MikesFavoriteWine">
  <owl:sameAs rdf:resource="#StGenevieveTexasWhite" />
</Wine>
```

Παράδειγμα B-8: Χρήση της ιδιότητας owl:differentFrom.

```
<WineSugar rdf:ID="Dry" />

<WineSugar rdf:ID="Sweet">
  <owl:differentFrom rdf:resource="#Dry"/>
</WineSugar>

<WineSugar rdf:ID="OffDry">
```

```
<owl:differentFrom rdf:resource="#Dry"/>
<owl:differentFrom rdf:resource="#Sweet"/>
</WineSugar>
```

Παράδειγμα B-9: Χρήση της ιδιότητας `owl:AllDifferent`.

```
<owl:AllDifferent>
  <owl:distinctMembers rdf:parseType="Collection">
    <vin:WineColor rdf:about="#Red" />
    <vin:WineColor rdf:about="#White" />
    <vin:WineColor rdf:about="#Rose" />
  </owl:distinctMembers>
</owl:AllDifferent>
```

Παράδειγμα B-10: Χρήση της ιδιότητας `owl:inverseOf`.

```
<owl:ObjectProperty rdf:ID="hasMaker">
  <rdf:type rdf:resource="&owl;FunctionalProperty" />
</owl:ObjectProperty>

<owl:ObjectProperty rdf:ID="producesWine">
  <owl:inverseOf rdf:resource="#hasMaker" />
</owl:ObjectProperty>
```

Παράδειγμα B-11: Χρήση της ιδιότητας `owl:TransitiveProperty`.

```
<owl:ObjectProperty rdf:ID="hasMaker">
  <rdf:type rdf:resource="&owl;FunctionalProperty" />
</owl:ObjectProperty>

<owl:ObjectProperty rdf:ID="producesWine">
  <owl:inverseOf rdf:resource="#hasMaker" />
</owl:ObjectProperty>
```

Παράδειγμα B-12: Χρήση της ιδιότητας `owl:SymmetricProperty`.

```

<owl:ObjectProperty rdf:ID="adjacentRegion">
  <rdf:type rdf:resource="&owl;SymmetricProperty" />
  <rdfs:domain rdf:resource="#Region" />
  <rdfs:range rdf:resource="#Region" />
</owl:ObjectProperty>

<Region rdf:ID="MendocinoRegion">
  <locatedIn rdf:resource="#CaliforniaRegion" />
  <adjacentRegion rdf:resource="#SonomaRegion" />
</Region>

```

Παράδειγμα B-13: Χρήση της ιδιότητας `owl:FunctionalProperty`.

```

<owl:Class rdf:ID="VintageYear" />

<owl:ObjectProperty rdf:ID="hasVintageYear">
  <rdf:type rdf:resource="&owl;FunctionalProperty" />
  <rdfs:domain rdf:resource="#Vintage" />
  <rdfs:range rdf:resource="#VintageYear" />
</owl:ObjectProperty>

```

Παράδειγμα B-14: Χρήση της ιδιότητας `owl:InverseFunctionalProperty`.

```

<owl:ObjectProperty rdf:ID="hasMaker" />

<owl:ObjectProperty rdf:ID="producesWine">
  <rdf:type rdf:resource="&owl;InverseFunctionalProperty" />
  <owl:inverseOf rdf:resource="#hasMaker" />
</owl:ObjectProperty>

```

Παράδειγμα B-15: Χρήση της ιδιότητας `owl:allValuesFrom`.

```

<owl:Class rdf:ID="Wine">
  <rdfs:subClassOf rdf:resource="&food;PotableLiquid" />
  <rdfs:subClassOf>

```

```

<owl:Restriction>
  <owl:onProperty rdf:resource="#hasMaker" />
  <owl:allValuesFrom rdf:resource="#Winery" />
</owl:Restriction>
</rdfs:subClassOf>
</owl:Class>

```

Παράδειγμα B-16: Χρήση της ιδιότητας `owl:someValuesFrom`.

```

<owl:Class rdf:ID="Wine">
  <rdfs:subClassOf rdf:resource="&food;PotableLiquid" />
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#hasMaker" />
      <owl:someValuesFrom rdf:resource="#Winery" />
    </owl:Restriction>
  </rdfs:subClassOf>
</owl:Class>

```

Παράδειγμα B-17: Χρήση του περιορισμού `owl:cardinality`.

```

<owl:Class rdf:ID="Vintage">
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#hasVintageYear"/>
      <owl:cardinality
rdf:datatype="&xsd;nonNegativeInteger">1</owl:cardinality>
    </owl:Restriction>
  </rdfs:subClassOf>
</owl:Class>

```

Παράδειγμα B-18: Χρήση του απαριθμητή `owl:oneOf`.

```

<owl:Class rdf:ID="WineColor">
  <rdfs:subClassOf rdf:resource="#WineDescriptor"/>

```

```

<owl:oneOf rdf:parseType="Collection">
  <owl:Thing rdf:about="#White"/>
  <owl:Thing rdf:about="#Rose"/>
  <owl:Thing rdf:about="#Red"/>
</owl:oneOf>
</owl:Class>

```

Παράδειγμα B-19: Χρήση του χαρακτηρισμού `owl:hasValue`.

```

<owl:Class rdf:ID="Burgundy">
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#hasSugar" />
      <owl:hasValue rdf:resource="#Dry" />
    </owl:Restriction>
  </rdfs:subClassOf>
</owl:Class>

```

Παράδειγμα B-20: Χρήση της δήλωσης `owl:disjointWith`.

```

<owl:Class rdf:ID="Pasta">
  <rdfs:subClassOf rdf:resource="#EdibleThing"/>
  <owl:disjointWith rdf:resource="#Meat"/>
  <owl:disjointWith rdf:resource="#Fowl"/>
  <owl:disjointWith rdf:resource="#Seafood"/>
  <owl:disjointWith rdf:resource="#Dessert"/>
  <owl:disjointWith rdf:resource="#Fruit"/>
</owl:Class>

```

Παράδειγμα B-21: Χρήση των τελεστών συνόλων `owl:unionOf` και `owl:complementOf`.

```

<owl:Class rdf:ID="Fruit">
  <owl:unionOf rdf:parseType="Collection">
    <owl:Class rdf:about="#SweetFruit" />
    <owl:Class rdf:about="#NonSweetFruit" />
  </owl:unionOf>
</owl:Class>

```

```
    </owl:unionOf>
</owl:Class>

<owl:Class rdf:ID="ConsumableThing" />

    <owl:Class rdf:ID="NonConsumableThing">
        <owl:complementOf rdf:resource="#ConsumableThing" />
    </owl:Class>
```

Παράρτημα Γ – Εργαλεία ελέγχου εγκυρότητας και μετατροπείς εγγράφων

Έλεγχος εγκυρότητας εγγράφων HTML, XHTML, XML:

- ♦ <http://validator.w3.org/>

Έλεγχος εγκυρότητας εγγράφων XML Schema:

- ♦ <http://tools.decisionsoft.com/schemaValidate/>,
- ♦ <http://www.w3.org/2001/03/webdata/xsv/>

Έλεγχος εγκυρότητας εγγράφων RDF, OWL:

- ♦ <http://phoebus.cs.man.ac.uk:9999/OWL/Validator>,
- ♦ <http://www.w3.org/RDF/Validator/>

Μετατροπέας εγγράφων RDF σε σημειογραφία 3 και αντίστροφα:

- ♦ <http://rdfabout.com/demo/validator/validate.xpd>

Παράρτημα Δ – Άδεια Ελεύθερης Τεκμηρίωσης GNU

GNU Free Documentation License
Version 1.2, November 2002

Copyright (C) 2000,2001,2002 Free Software Foundation, Inc.
51 Franklin St, Fifth Floor, Boston, MA 02110-1301 USA
Everyone is permitted to copy and distribute verbatim copies
of this license document, but changing it is not allowed.

0. PREAMBLE

The purpose of this License is to make a manual, textbook, or other functional and useful document "free" in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondly, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of "copyleft", which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

1. APPLICABILITY AND DEFINITIONS

This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The "Document", below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as "you". You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A "Modified Version" of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A "Secondary Section" is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document's overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a

textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The "Invariant Sections" are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

The "Cover Texts" are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.

A "Transparent" copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart

or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not "Transparent" is called "Opaque".

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The "Title Page" means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material

this License requires to appear in the title page. For works in formats which do not have any title page as such, "Title Page" means the text near the most prominent appearance of the work's title, preceding the beginning of the body of the text.

A section "Entitled XYZ" means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following

text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as "Acknowledgements", "Dedications", "Endorsements", or "History".) To "Preserve the Title"

of such a section when you modify the Document means that it remains a

section "Entitled XYZ" according to this definition.

The Document may include Warranty Disclaimers next to the notice which

states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties: any other implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

2. VERBATIM COPYING

You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other

conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept

compensation in exchange for copies. If you distribute a large enough

number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

3. COPYING IN QUANTITY

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy

a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an

Opaque copy (directly or through your agents or retailers) of that

edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

4. MODIFICATIONS

You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

A. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version

if the original publisher of that version gives permission.

B. List on the Title Page, as authors, one or more persons or entities

responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the

Document (all of its principal authors, if it has fewer than five),

unless they release you from this requirement.

- C. State on the Title page the name of the publisher of the Modified Version, as the publisher.
- D. Preserve all the copyright notices of the Document.
- E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.
- F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.
- G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.
- H. Include an unaltered copy of this License.
- I. Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.
- J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.
- K. For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.

- L. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.
- M. Delete any section Entitled "Endorsements". Such a section may not be included in the Modified Version.
- N. Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section.
- O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties--for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or

by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

5. COMBINING DOCUMENTS

You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but

different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number.

Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled "History" in the various original documents, forming one section Entitled "History"; likewise combine any sections Entitled "Acknowledgements",

and any sections Entitled "Dedications". You must delete all sections Entitled "Endorsements".

6. COLLECTIONS OF DOCUMENTS

You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

7. AGGREGATION WITH INDEPENDENT WORKS

A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, is called an "aggregate" if the copyright resulting from the compilation is not used to limit the legal rights of the compilation's users beyond what the individual works permit.

When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document's Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form.

Otherwise they must appear on printed covers that bracket the whole aggregate.

8. TRANSLATION

Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled "Acknowledgements",

"Dedications", or "History", the requirement (section 4) to Preserve its Title (section 1) will typically require changing the actual title.

9. TERMINATION

You may not copy, modify, sublicense, or distribute the Document except

as expressly provided for under this License. Any other attempt to copy, modify, sublicense or distribute the Document is void, and will automatically terminate your rights under this License. However, parties who have received copies, or rights, from you under this License will not have their licenses terminated so long as such parties remain in full compliance.

10. FUTURE REVISIONS OF THIS LICENSE

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See <http://www.gnu.org/copyleft/>.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License "or any later version" applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version

number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation.

ADDENDUM: How to use this License for your documents

To use this License in a document you have written, include a copy of the License in the document and put the following copyright and license notices just after the title page:

Copyright (c) YEAR YOUR NAME.

Permission is granted to copy, distribute and/or modify this document

under the terms of the GNU Free Documentation License, Version 1.2

or any later version published by the Free Software Foundation;

with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts.

A copy of the license is included in the section entitled "GNU Free Documentation License".

If you have Invariant Sections, Front-Cover Texts and Back-Cover Texts,

replace the "with...Texts." line with this:

with the Invariant Sections being LIST THEIR TITLES, with the

Front-Cover Texts being LIST, and with the Back-Cover Texts being LIST.

If you have Invariant Sections without Cover Texts, or some other combination of the three, merge those two alternatives to suit the

situation.

If your document contains nontrivial examples of program code, we recommend releasing these examples in parallel under your choice of free software license, such as the GNU General Public License, to permit their use in free software.

Ευρετήριο όρων

A

αναζήτηση 12, 24, 30, 35, 58, 62, 123
αναπαράσταση γνώσης 17
ανοιχτή τεχνολογία 49
ανταλλαγή δεδομένων 21
αντίστροφη ιδιότητα 103, 104
αντίφαση 102
αξιοπιστία 18, 30
απόδειξη 29, 30
αρχιτεκτονική 11, 18, 48, 53, 54, 55, 61, 74
άτομο 40, 100, 101, 102, 104, 107, 109, 111
αυτοματοποίηση 35, 62, 63

Γ

γεγονός 24, 65
γλώσσα δημιουργίας οντολογιών ιστού 26, 49, 52, 96
γλώσσα περιγραφής λεξιλογίων 25, 85
γλώσσα σήμανσης υπερκειμένου 19
γλώσσα υποβολής αιτημάτων 28
γλωσσικό κατασκεύασμα 97

Δ

δεντρικό μοντέλο 21
δήλωση 21, 22, 38, 78, 81, 85, 92, 93, 94, 95, 102, 111
δηλωτικός κανόνας 27
δηλωτικός τρόπος 38
Διαδίκτυο 8, 12, 13, 22, 27, 34, 69, 78, 82, 113, 123, 124
διαμοιρασμός γνώσης 37, 38
διανομή 42, 43
διευθυνσιοδότηση 18, 19, 22

Ε

εκφραστικότητα 81, 97, 98, 99
έλεγχος εγκυρότητας 53
ενιαίος εντοπιστής πόρων 19
έννοια 9, 19, 22, 25, 36, 40, 46, 56, 64, 65, 66, 76, 77, 87
ενοποιημένη λογική 28
ένταξη 42
εξαγωγή δεδομένων 35, 75
επαναχρησιμοποίηση 35, 37, 56, 63, 76, 77, 115
επεκτάσιμη 17, 19, 49, 60
επεκτάσιμη γλώσσα σήμανσης 17

εύρος τιμών 106, 109, 110

Θ

θεμελιώδης οντολογία 46

Ι

ιεραρχία κλάσεων 42, 44
ιεραρχία χαρακτήρων δεδομένων 70
ιεραρχική ταξινόμηση 44
ισοδυναμία 25
ισχυρισμός 21

Κ

αντικείμενο 21, 23, 39, 42, 43, 44, 78, 80, 81, 94
καλοσχηματισμένο έγγραφο 84
κατευθυνόμενο μη κυκλικό γράφημα 45
κατηγορήμα 21, 23, 80, 81, 89, 93, 94
κατηγορήμα 81
κατηγορηματική λογική 29
Κλάση 37, 38, 40, 41, 42, 58, 85, 86, 87, 88, 89, 91, 96, 98, 100, 101, 109, 110, 111, 119
κοινοπραξία του παγκόσμιου ιστού πληροφοριών 26
κρυπτογραφία 18, 32, 33
κωδικοποίηση 18, 19, 22, 30, 35, 76, 77, 81, 82

Λ

λέξη κλειδί 8, 15, 40, 106
λεξιλόγιο 24, 38, 39, 47, 50, 51, 71, 78, 80, 85, 86, 89, 91, 100, 109

Μ

μεταβατική ιδιότητα 89, 103
μεταδεδομένα 8, 9, 21, 33, 60, 76
μηχανή αναζήτησης 8, 58
μηχανισμός συλλογιστικής 113
μορφή ανταλλαγής κανόνων 27

Ν

νοημοσύνη 13, 16, 62

Ο

ομάδα εργασίας πρόσβασης δεδομένων RDF 28
ομαδοποίηση 35
ονοματοδοσία 18, 20

οντολογία 9, 17, 24, 26, 33, 36, 37, 38, 39, 40, 43, 45, 46, 47, 49, 50, 51, 53, 60, 61, 65, 67, 113, 123
οντολογία περιοχής 45, 62, 113
οντολογική δέσμευση 38
οντότητα 13, 45
ορθότητα 31

Π

παγκόσμιος ιστός πληροφοριών 8, 76, 123
περιγραφική λογική 98
πολιτική 62, 64, 65
πολλαπλή κληρονομικότητα 89, 90
πολυπλοκότητα 25, 53, 97
προδιαγραφή 36, 74
πρότυπη γενικευμένη γλώσσα σήμανσης 69
πρωτόκολλο επικοινωνίας 15

Σ

σημασιολογία 11
σημασιολογικό περιεχόμενο 9, 60, 113, 123
σημασιολογικός ιστός 9, 11, 12, 30, 32, 33, 34, 35, 68, 123
σημειογραφία 3 81, 82
σημειογραφία στηλών 88, 90
σκελετός περιγραφής πόρων 17, 21, 76
στιγμιότυπο 40, 41, 86, 87, 89, 91, 92, 93, 94, 96, 100, 101, 103, 104, 105, 106, 107, 108, 109
στοίβα 18
συγχώνευση οντολογιών 46
συλλογιστικός κανόνας 17
συμπερασματική κλάση 41
σύνδεσμος 12, 52, 53, 54
σύνθετος τύπος δεδομένων 43
συντακτικό 11, 18, 21, 24, 69, 72, 73, 76, 77, 81, 82, 111
σύνταξη 14, 21, 70, 71, 72, 81, 82, 84, 90
σύστημα δημιουργίας κανόνων 27
σύστημα διαχείρισης βάσεων δεδομένων 60
σχέση 42, 43, 44, 45, 62, 65, 89, 95, 101, 103, 105, 107, 108
σχήμα 27, 53, 71, 74, 75, 94, 95

Τ

τεχνολογία δημιουργίας οντολογιών 48
τομή 108
τριάδα 23, 24, 80, 81, 89, 92, 93
τύπος δεδομένων 48, 71, 75, 84, 93, 100

Υ

υπερσύνολο αναφοράς 38
υποθετική κλάση 41
υποκείμενο 21, 23, 80, 81, 83, 94
υποκλάση 37, 42, 43, 89, 100
υπολογιστική πληρότητα 97

Ψ

ψηφιακή υπογραφή 31

A

agent 14, 48

C

cardinality 96, 97, 104, 106, 107, 108, 110
Cyc 50

D

DAML+OIL 46, 48, 52, 53, 97
datatype 71, 100, 109
directed acyclic graph 45
disjoint 43, 96, 99, 110
DOLCE 50
domain 38, 45, 51, 61, 65, 66, 91, 93, 94, 95, 100, 101
Dublin Core 50

E

edge 81
EuroCitizen 111, 112, 113, 114, 115, 118, 119
exhaustive 43
expressiveness 48
eXtensible Markup Language 17, 21, 23, 69, 70, 72, 74
extensional 41

F

functional property 104

G

Gene Ontology 51
Generalized Upper Model 51
GFO 50
GUM 51

H

HTML 19, 21, 71
Hypertext Markup Language 19

I

inheritance mechanism 47
instance 40, 87
intensional 41
Internationalized Resource Identifier 19
Internet 12
inverse functional property 104
inverse property 103, 104
IRI 19, 20, 78

K

KAoS 62, 64, 65, 66, 67
Kasai 55, 56, 57
keyword 8, 106
knowledge sharing 36

L

link 12
logic 31, 39

M

Materialized Ontology View Extractor 53
meronymy 44
metadata 57
MOVE 53, 54, 55
multiple inheritance 89

N

node 81
notation 3 81

O

object 21, 23, 71, 80, 93, 100, 109
ontology 17, 50, 53
OWL 26, 28, 35, 46, 49, 52, 53, 59, 61, 62, 64, 65, 66, 67, 68, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 115
OWL DL 97, 98, 99, 106, 109, 110, 113, 115
OWL Full 98, 99, 106, 109, 110, 111, 113
OWL Lite 97, 98, 99, 100, 101, 102, 105, 106, 107, 108, 109, 110, 111, 113

P

partition 42
predicate 21, 23, 80, 93, 94
Property 78, 91, 92, 93, 95, 100, 102, 104, 105

Q

query 27

R

rdf 17, 18, 20, 21, 22, 23, 24, 25, 28, 29, 30, 31, 35, 46, 47, 48, 49, 52, 57, 58, 60, 61, 62, 68, 69, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 95, 96, 97, 98, 99, 100
RDFS 25, 28, 46, 47, 52, 68, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 100, 101
reasoner 101, 113
relation 43
Resource Description Framework 17, 21, 23, 25, 76, 78, 80, 81, 84, 87, 91, 99
Resource Description Framework Schema 25, 84, 87, 91, 99
reusability 63
rif 27
Rule Interchange Format 27
rule of inference 27

S

schema 47, 86, 87, 91
SCOS 53

Semantic Completeness Optimization Scheme	53, 55
semantic web	9, 62
serialization	47
SGML	69, 72, 73
SPARQL	28
Standard Generalized Markup Language	69
statement	83
string	75
structure	66, 76
subject	21, 23, 80, 85, 94, 95
subproperty	95
SUMO	50, 113, 115
symmetric property	103
syntax	11, 71, 76, 82

T

tabular notation	88
taxonomy	37, 44, 97
tractability	48
transitive property	89, 96, 103

U

Uniform Resource Identifier	19
Uniform Resource Locator	19
Uniform Resource Name	19
universe of discourse	38
URI	19, 20, 78, 91
URL	19
URN	19, 20

V

Vocabulary	38, 77, 85
------------	------------

W

W3C	26, 28, 35, 46, 47, 48, 69, 74, 75, 76, 96
Web Ontology Language	26, 49, 96, 97, 99, 101, 102, 105, 106, 108, 109
WordNet	50, 113, 115
World Wide Web	8, 12, 26
World Wide Web Consortium	26

X

XML 17, 18, 20, 21, 22, 23, 24, 25, 29, 47, 48, 61, 68, 69, 70, 71, 72, 73, 74, 75, 78, 82, 83, 84, 90, 97, 100

