

# Θεωρητική προσέγγιση του Σημασιολογικού Ιστού στο χώρο της πολιτισμικής πληροφορίας: μία πρότυπη εφαρμογή στη βιβλιοθηκονομία

## A theoretical approach of Semantic Web technologies in the cultural information domain: a case application in library science

Σοφία Ζαπουνίδου  
Πανεπιστήμιο Θεσσαλίας - Κεντρική Βιβλιοθήκη

Sofia Zapounidou  
University of Thessaly - Central Library

### Περίληψη

*Η ιδέα της δημιουργίας ενός νέου Παγκόσμιου Ιστού, όπου τα υπολογιστικά συστήματα κατανοούν την πληροφορία και -βασισμένα στο περιεχόμενό της- προσφέρουν προηγμένες υπηρεσίες, έχει προσελκύσει το ενδιαφέρον πολλών επιστημονικών κοινοτήτων, συμπεριλαμβανομένου και αυτής των επιστημόνων της πληροφορίας. Οι ειδικοί στο χώρο της πληροφορίας αναμένεται, ως διαχειριστές της, να έχουν πρωτεύοντα ρόλο στην εξέλιξη του Παρόντος Παγκόσμιου Ιστού (World Wide Web) στο Σημασιολογικό Ιστό του μέλλοντος (Semantic Web). Η εφαρμογή των νέων τεχνολογιών Σημασιολογικού Ιστού στο χώρο της διαχείρισης της πολιτισμικής πληροφορίας αποτελεί πρόκληση για τους ειδικούς τόσο των επιστημών της πληροφορίας, όσο και της πολιτισμικής πληροφορικής. Η επιτυχία του εγχειρήματος εξαρτάται αποκλειστικά από το βαθμό κατανόησης του θεωρητικού πλαισίου πάνω στο οποίο βασίζονται αυτές οι τεχνολογίες. Η παρούσα εργασία επιχειρεί να συμβάλει προς αυτήν την κατεύθυνση αναλύοντας, στο πρώτο της μέρος, τη θεωρία και συνεχίζοντας, στο δεύτερο, με σχετικό παράδειγμα.*

*Κύριο άξονα στη θεωρητική προσέγγιση του θέματος αποτελεί η διαστρωματωμένη αρχιτεκτονική του Σημασιολογικού Ιστού. Οι νέες τεχνολογίες που βασίζονται στη σημασιολογία των δεδομένων και των πληροφοριών θεωρούνται δύσκολες στην εφαρμογή τους, με αποτέλεσμα να προτιμώνται γνωστές και -κατά συνέπεια- ασφαλέστερες λύσεις πληροφορικής οργάνωσης. Η εισαγωγή της καινοτομίας, ωστόσο, σε βιβλιοθήκες, και σε κέντρα διαχείρισης της πολιτισμικής πληροφορίας γενικότερα, αποτελεί μία συνηθισμένη πρακτική, λόγος για τον οποίο εκδηλώνεται από την πλευρά των επιστημόνων που απασχολούνται σε αυτούς τους χώρους ολοένα αυξανόμενο ενδιαφέρον. Το ενδιαφέρον τους εστιάζεται στην κατανόηση του τρόπου σημασιολογικής αναπαράστασης της γνώσης. Ο απώτερος στόχος είναι η υιοθέτηση και η αξιοποίηση αυτής της μεθοδολογικής προσέγγισης στη βελτίωση των ήδη παρεχόμενων υπηρεσιών και στην ανάπτυξη νέων, πιο προηγμένων, οι οποίες θα προσφέρουν πρωτοποριακές -για τα σύγχρονα διαθέσιμα τεχνολογικά εργαλεία- λύσεις.*

*Στο δεύτερο μέρος της εργασίας παρατίθεται ενδεικτικό παράδειγμα χρήσης τεχνολογιών Σημασιολογικού Ιστού από το χώρο της πολιτισμικής πληροφορίας. Ο Παγκόσμιος Ιστός του μέλλοντος προσφέρει νέες προοπτικές στο χώρο της διαχείρισης της πληροφορίας (Information Management) μπορεί να δώσει ώθηση σε νέες εφαρμογές διαχείρισης γνώσης (Knowledge Management) και να εισάγει “νοήμονες” υπηρεσίες που θα πραγματοποιούν εργασίες αντί των*

ανθρώπων με πολύ ικανοποιητικά αποτελέσματα. Παρουσιάζεται μία υποθετική εφαρμογή από το χώρο της βιβλιοθηκονομίας, η μεθοδολογία της οποίας μπορεί να αξιοποιηθεί και σε άλλους τομείς της πολιτισμικής πληροφορίας. Η διαπίστωση ότι έχουν ήδη αρχίσει να δημιουργούνται νησίδες πολιτισμικής πληροφορίας οργανωμένης με σημασιολογικό τρόπο θα αποτελέσει έναυσμα για τη δημιουργία νέων νησίδων και για τη διασύνδεσή τους με τις υπάρχουσες, προβάλλοντας ταυτόχρονα στους επιστήμονες της πληροφορίας ρόλους, μέχρι στιγμής, άγνωστους.

Λέξεις-Κλειδιά: Σημασιολογικός Ιστός, Πλαίσιο Περιγραφής Πόρων, μεταδεδομένα, πολιτισμική πληροφορική, οντολογίες

## Abstract

*The idea of developing a new World Wide Web, where computers understand the meaning of the information and offer -based upon it- advanced services has drawn many scientific communities' attention, including the information community. Its professionals are expected to be the pioneers in the evolution process from the existing World Wide Web to the Semantic Web of the future. Implementing new Semantic Web technologies in the cultural information disciplines challenges both the information science and cultural informatics experts. Addressing this challenge successfully depends on the understanding of the theoretical framework upon which the Semantic Web is based. This paper seeks to contribute to this goal by firstly analyzing the basic Semantic Web concepts and by secondly introducing a relevant application example.*

*Primary axis of the theoretical approach is the Semantic Web's layered architecture. New technologies based upon data's and information's semantics are considered difficult in implementing. As a result other known, and therefore more secure, solutions of computer technology are preferred. Yet, introducing innovation in libraries and cultural information centres is a common practice; a fact not ignored by the experts working in these institutions. Their growing interest focuses on understanding the method of organizing knowledge semantically within the scope of adopting and exploiting new ways in improving rendered services and in developing new advanced ones. These new advanced services are going to offer pioneer -for the state of art technological tools- solutions.*

*In the second part of the paper a case semantic application in the cultural information domain is mentioned. The future World Wide Web offers new perspective in the information management field; it launches new knowledge management applications and introduces "intelligent" services, able to perform tasks for people with very satisfying results. An indicative example of a hypothetical library application is presented, the methodology of which may be used in other cultural information applications too. Ascertaining that semantic islands of cultural information are being developed will contribute in developing new ones and connecting them with those that already exist. Within this context information scientists will be expected to play new roles -unknown to them so far.*

Keywords: Semantic Web, Resource Description Framework (RDF), metadata, cultural informatics, ontologies

## 1. Εισαγωγή

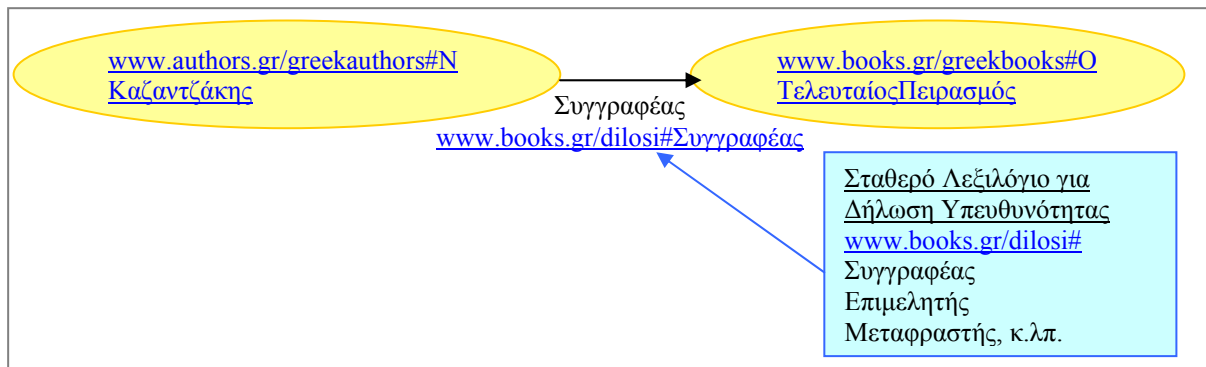
Ο Σημασιολογικός Ιστός αποτελεί ένα σχετικά καινούριο όραμα για την κοινωνία της πληροφορίας του μέλλοντος, όπως το εμπνεύστηκε ο πατέρας του σύγχρονου παγκόσμιου Ιστού, Berners-Lee (1998). Ο ίδιος υποστήριξε πως η κατανόηση του περιεχομένου που είναι διαθέσιμο στον Παγκόσμιο Ιστό από τους ηλεκτρονικούς υπολογιστές μπορεί να προσφέρει μία επανάσταση από νέες δυνατότητες και πληροφορικές εφαρμογές (Berners-Lee 2001). Στη διαμόρφωση της κοινωνίας της πληροφορίας του μέλλοντος συμμετέχουν ενεργά οι σύγχρονοι διαχειριστές της, οι ειδικοί από το χώρο των επιστημών της πληροφορίας.

Η παρουσίαση που ακολουθεί θα θεωρείται πως έχει πετύχει την αποστολή της, αν καταφέρει να διασαφηνίσει στους Έλληνες επιστήμονες της πληροφορίας τον τρόπο λειτουργίας του Σημασιολογικού Ιστού, τον τρόπο αναπαράστασης δεδομένων και νοημάτων στα πλαίσιά του, καθώς και τις δυνατότητες αξιοποίησης αυτών των εργαλείων στην ανάπτυξη “νοήμονων” υπηρεσιών. Η παρουσίαση διακρίνεται σε δύο μέρη: αρχικά αναπτύσσεται το θεωρητικό υπόβαθρο του Σημασιολογικού Ιστού και προβάλλονται οι δυνατότητές του. Ακολουθεί μία μελέτη περίπτωσης στη βιβλιοθηκονομία με στόχο να λειτουργήσει ως απτό παράδειγμα των όσων αναφέρονται στο θεωρητικό πρώτο μέρος.

## 2. Αρχιτεκτονική Σημασιολογικού Ιστού

### 2.1 Επίπεδο Δεδομένων

Ο Σημασιολογικός Ιστός αποτελεί επέκταση του παρόντος Ιστού. Η διαφορά έγκειται στο γεγονός ότι αποδίδεται πολύ καλά το νόημα της διαθέσιμης πληροφορίας σε μορφή κατανοητή τόσο από ανθρώπους, όσο και από ηλεκτρονικούς υπολογιστές, διευκολύνοντας τη μεταξύ τους συνεργασία. Η σημασιολογική περιγραφή των δεδομένων και του περιεχομένου πραγματοποιείται σε σταθερά λεξιλόγια, τα οποία είναι διαθέσιμα στον Ιστό και λειτουργούν ως σημεία αναφοράς. Για την περιγραφή των δεδομένων και τη σύνδεσή τους με σταθερά λεξιλόγια απαιτείται η χρήση ενός απλού μοντέλου μεταδεδομένων και μίας σύνταξης για μεταδεδομένα. Το επικρατέστερο μοντέλο μεταδεδομένων είναι το Πλαίσιο Περιγραφής Πόρων (Resource Description Framework), το οποίο χρησιμοποιώντας τριαδικούς συνδυασμούς πόρων Ιστού καταφέρνει να προβάλλει μία σαφή μέθοδο έκφρασης σημασιολογίας σε μηχαναγνώσιμη κωδικοποίηση (Miller 1998).



Σχήμα 1: Μία RDF δήλωση

Οι Πόροι (Resources) είναι ένας μόνο από τους τρεις τύπους αντικειμένων που αποτελούν το βασικό RDF μοντέλο. Οι άλλοι δύο είναι οι Ιδιότητες (Properties) και οι Δηλώσεις (Statements). Συγκεκριμένα (Lassila 1999):

- Πόροι (Resources) Μπορούν να χαρακτηριστούν όλα τα αντικείμενα που περιγράφονται από RDF δηλώσεις, π.χ. μία ιστοσελίδα, ένα μέρος μίας ιστοσελίδας, ένας ολόκληρος ιστότοπος (Web site). Ακόμα, Πόρος μπορεί να είναι ένα αντικείμενο που δεν είναι άμεσα προσβάσιμο μέσω του Παγκόσμιου Ιστού, όπως ένα τυπωμένο βιβλίο. Οι Πόροι του παραδείγματος είναι ο “N. Καζαντζάκης” και “Ο Τελευταίος Πειρασμός”.
- Ιδιότητες (Properties) Εκφράζουν συγκεκριμένες απόψεις, ένα χαρακτηριστικό, μία σχέση που χρησιμοποιείται για να περιγράψει τον Πόρο. Ιδιότητα στο παράδειγμα είναι ο “Συγγραφέας”.
- Δηλώσεις (Statements) Ένας συγκεκριμένος Πόρος μαζί με μία δηλωμένη Ιδιότητα συν την αξία της ιδιότητας για τον Πόρο αποτελούν μία RDF δήλωση. Στο παράδειγμα: ο N. Καζαντζάκης είναι συγγραφέας του βιβλίου “Ο Τελευταίος Πειρασμός”.

## 2.2 Επίπεδο Σχήματος

Η σημασιολογία των Πόρων εκφράζεται με μία αναφορά (URI) σε ένα σταθερό λεξιλόγιο. Τα σταθερά αυτά λεξιλόγια που ονομάζονται σχήματα–οντολογίες καθορίζουν τους όρους που χρησιμοποιούνται στις RDF δηλώσεις και δίνουν συγκεκριμένο νόημα σε αυτές. Το καθένα από αυτά αναφέρεται σε συγκεκριμένο επιστημονικό, επιχειρηματικό πεδίο και περιλαμβάνει τόσο τους όρους που κρίθηκαν ως αντιπροσωπευτικοί για την περιγραφή τους, όσο και τις σχέσεις που υπάρχουν μεταξύ τους (Swartz 2002).

Το RDF περιλαμβάνει μία προκαταρκτική εργασία σχετικά με το ορισμό σχημάτων μεταδεδομένων (metadata schemas). Η εργασία αυτή βασίζεται στο RDF μοντέλο και ονομάζεται RDF Schema (Ianella 1999). Το RDF Schema αν και αποτελεί ένα χρήσιμο εργαλείο για τον καθορισμό λεξιλογίου, δομής και περιορισμών και την έκφραση μεταδεδομένων σχετικά με Πόρους Ιστού, δεν μπορεί να αποδώσει περίπλοκες σχέσεις μεταξύ των εννοιών. Επειδή ακριβώς παρέχει περιορισμένες δυνατότητες εκφρασιμότητας, οι οποίες δεν αρκούν για μία ολοκληρωμένη οντολογική μοντελοποίηση και λογική, απαιτείται ένα επιπλέον επίπεδο πάνω από αυτό του ορισμού λεξιλογίου μέσω σχημάτων. Αυτήν

ακριβώς τη διαστρωματωμένη αρχιτεκτονική που δημιουργείται αποκαλεί ο Tim Berners-Lee “Σημασιολογικό Ιστό (*Semantic Web*)” (Broekstra et. al. 2000: 1).

<b>Λογικό Επίπεδο (Logical layer)</b> Επίσημη σημασιολογία και υποστήριξη λειτουργίας εξαγωγής συμπερασμάτων OIL, DAML-OIL, OWL
<b>Επίπεδο Σχήματος (Schema layer)</b> Ορισμός του λεξιλογίου RDF Schema
<b>Επίπεδο Δεδομένων (Data layer)</b> Απλό μοντέλο δεδομένων και σύνταξη για μεταδεδομένα RDF

Σχήμα 2: Η διαστρωματωμένη αρχιτεκτονική του Σημασιολογικού Ιστού<sup>1</sup>

### 2.3 Λογικό Επίπεδο

Στο τρίτο και ανώτερο επίπεδο εισάγονται οι οντολογίες. Ο όρος προέρχεται από το χώρο της φιλοσοφίας και εκφράζει ένα τμήμα της το οποίο μελετά τις βασικές αρχές και τις νομοτέλειες της ύπαρξης (Εγκυκλοπαίδεια Υδρία-Cambridge-Ήλιος 1992: Λήμμα Οντολογία). Στην πληροφορική, οντολογία ορίζεται ως ο *σαφής προσδιορισμός ενός εννοιολογικού συλλογισμού* (Gruber 1993: 1)<sup>2</sup>. Περιλαμβάνουν και αυτές, όπως τα σχήματα, ένα ελεγχόμενο λεξιλόγιο όρων και τη διευκρίνιση του νοήματός τους. Το ελεγχόμενο λεξιλόγιο εκφράζεται σε μία γλώσσα αναπαράστασης οντολογιών, της οποίας η γραμματική περιλαμβάνει επίσημους περιορισμούς σχετικά με το πώς οι όροι του λεξιλογίου μπορούν να χρησιμοποιηθούν μαζί (Metamodel 2003). Πρωταρχικός στόχος μίας οντολογικής αναπαράστασης είναι η μείωση της ασάφειας στη χρήση κάποιων όρων, έτσι ώστε οι όροι να χρησιμοποιούνται με την ερμηνεία που τους δόθηκε μέσα στην οντολογία και όχι με κάποια άλλη που πιθανώς αυτοί να επιδέχονται (Ζαπουνίδου 2004: 12).

Ουσιαστική διαφορά μεταξύ των όρων *σχήμα-οντολογία* δεν υπάρχει. Έχει επικρατήσει στην Κοινότητα του Σημασιολογικού Ιστού ο πρώτος όρος να αναφέρεται σε απλούς εννοιολογικούς συλλογισμούς, οι οποίοι εκφράζονται από λιγότερο εκφραστικές γλώσσες αναπαράστασης της γνώσης, όπως είναι το RDF Schema. Ο δεύτερος όρος αφορά σε πιο περίπλοκα μοντέλα, τα οποία περιγράφονται από εκφραστικότερες γλώσσες, όπως είναι οι OIL, DAML+OIL, Ontolingua, OWL, κ.λπ. (ESW Wiki FAQs 2003).

<sup>1</sup> Το συγκεκριμένο σχήμα με ελάχιστες τροποποιήσεις βασίζεται σε σχήμα που χρησιμοποιήθηκε στην παρουσίαση της γλώσσας αναπαράστασης οντολογιών OIL (Broekstra et.al. 2000: 1)

<sup>2</sup> Η μετάφραση του ορισμού του Gruber ανακτήθηκε από το άρθρο των Τομαή, Κάβουρα (2002: 1).

### 2.3.1 Σχέση θησαυρού – οντολογιών

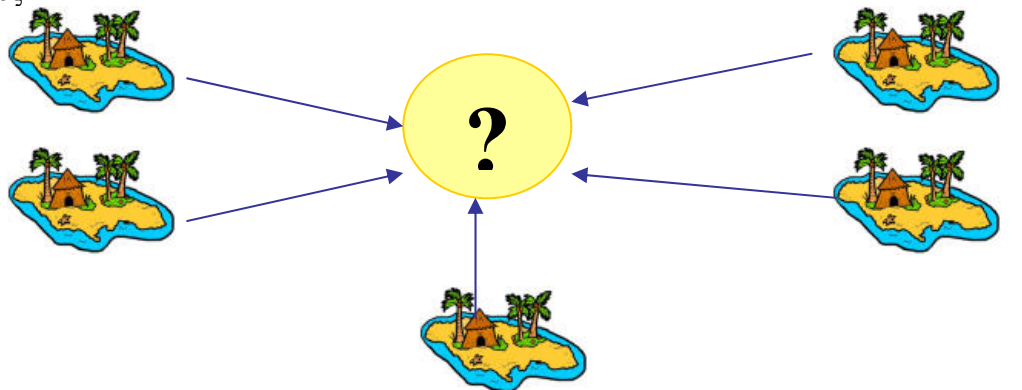
Η σχέση των οντολογιών με τους θησαυρούς είναι εμφανής, καθώς και τα δύο εργαλεία αποτελούν “προσεγγίσεις δόμησης, ταξινόμησης, μοντελοποίησης και αναπαράστασης εννοιών και σχέσεων ενός γνωστικού πεδίου” (Metamodel 2003). Και στις δύο περιπτώσεις η κοινή αντίληψη που δημιουργείται για ένα γνωστικό πεδίο εκφράζεται μέσα από ένα ελεγχόμενο λεξιλόγιο όρων, το οποίο έχουν συμφωνήσει να χρησιμοποιούν τα μέλη των σχετικών επιστημονικών κοινοτήτων. Και στα δύο εργαλεία αναπαράστασης γνώσης οι όροι του λεξιλογίου ομαδοποιούνται και το νόημα τους εκφράζεται με κάποιον τρόπο. Δε θα ήταν υπερβολή να υποστηριχθεί ότι οι οντολογίες αποτελούν τη φυσική εξέλιξη των θησαυρών. Η διαφορά εντοπίζεται στο γεγονός ότι οι θησαυροί προορίζονται για χρήση από ανθρώπους, ενώ μία οντολογία μπορεί να χρησιμοποιηθεί εξίσου από ανθρώπους και υπολογιστικά συστήματα. Το χαρακτηριστικό τους αυτό αποτελεί τον ακρογωνιαίο λίθο του Σημασιολογικού Ιστού και για ό,τι αναμένεται αυτός να παρέχει στην κοινωνία της πληροφορίας και γενικότερα στην καθημερινότητα του μέλλοντος.

### 3. Δυνατότητες Σημασιολογικού Ιστού

Ο τρόπος με τον οποίο θα αναπτυχθεί ο Σημασιολογικός Ιστός προβλέπεται (Semaview 2002: 6) να είναι ανάλογος με αυτόν του σύγχρονου Παγκόσμιου Ιστού. Το πρώτο βήμα στη διαμόρφωση του Παγκόσμιου Ιστού του μέλλοντος είναι η δημιουργία νησίδων πληροφορίας οργανωμένης σημασιολογικά. Οι νησίδες αυτές σιγά-σιγά θα διασυνδεθούν μεταξύ τους προσφέροντας περισσότερες δυνατότητες για την ανάπτυξη προηγμένων εφαρμογών. Το επόμενο βήμα είναι η αξιοποίηση του πλέγματος πληροφοριών, που δημιουργείται, σύμφωνα με τις προοπτικές που παρέχει η σημασιολογική τους οργάνωση.

Το καθένα από τα στάδια εξέλιξης του Σημασιολογικού Ιστού χαρακτηρίζεται από συγκεκριμένες τεχνολογικές δυνατότητες. Η διάθεση μέσω του Παγκόσμιου Ιστού δεδομένων που έχουν οριστεί και διασυνδεθεί μεταξύ τους αποτελεί τη βασική υποδομή για ό,τι υπόσχεται ο μελλοντικός Παγκόσμιος Ιστός: αποτελεσματικότερη αναζήτηση δεδομένων και πληροφοριών, αυτοματοποίηση και ολοκλήρωση υπηρεσιών, αξιοποίηση της υπάρχουσας γνώσης σε ποικίλες εφαρμογές (Hendler 2002).

1. Οι νησίδες δημιουργούνται
2. Οι νησίδες διασυνδέονται
3. Αξιοποίηση πληροφοριακού πλέγματος



Σχήμα 3: Νησίδες Πληροφορίας<sup>3</sup>

### 3.1 Αναζήτηση δεδομένων

Η προφανέστερη δυνατότητα του Σημασιολογικού Ιστού αφορά στην ανάκτηση πληροφοριών τόσο στα πλαίσια του Παγκόσμιου Ιστού, όσο και σε εκείνα μίας Βάσης δεδομένων – Ψηφιακής Βιβλιοθήκης. Η σημασιολογική οργάνωση της πληροφορίας έχει ως επακόλουθο τη βελτίωση των μηχανισμών αναζήτησης με πιο ποιοτικά αποτελέσματα κατά τις διαδικασίες ανάκτησης δεδομένων και πληροφοριών. Επιπλέον η χρήση μεταδεδομένων στην περιγραφή Πόρων Ιστού καθιστά δυνατή την αυτόματη ταξινόμηση και αποθήκευσή τους με βάση κανόνες εξαγωγής συμπερασμάτων που περιλαμβάνονται σε σχετικές οντολογίες. Μία τέτοια προοπτική προσφέρει επιπλέον δυνατότητες στην οργάνωση του ψηφιακού περιεχομένου με σημασιολογικά κριτήρια, στην αναζήτηση πληροφοριών και στην εξέλιξη των Ψηφιακών Βιβλιοθηκών και των εφαρμογών Διαχείρισης Γνώσης (Knowledge Management).

### 3.2 Αυτοματοποίηση και ολοκλήρωση υπηρεσιών

Η ύπαρξη οντολογιών που καλύπτουν ποικίλους τομείς της ανθρώπινης ζωής αποτελεί μία σημαντική υποδομή για την ανάπτυξη προηγμένων υπηρεσιών· η ελεύθερη πρόσβαση μέσω Ιστού σε οντολογίες στην ουσία αποτελεί ελεύθερη πρόσβαση σε γνώση που, επιπλέον, είναι οργανωμένη σύμφωνα με λογικά και σημασιολογικά κριτήρια. Ειδικά προγράμματα, βασιζόμενα σε αυτή τη γνώση, μπορούν να αναζητούν, να εκτελούν ή και να συνθέτουν ακόμη νέες υπηρεσίες, προς όφελος του χρήστη. Τα προγράμματα αυτά ονομάζονται - ανάλογα με την τεχνολογία στην οποία στηρίζονται- Πράκτορες (Agents) ή Υπηρεσίες Ιστού (Web Services). Και τα δύο είδη προγραμμάτων, αν και αποτελούν διαφορετικές προσεγγίσεις του Σημασιολογικού Ιστού, στηρίζονται στη γνώση που προσφέρουν οι

<sup>3</sup> Το σχήμα βασίζεται σε εικόνα που χρησιμοποιήθηκε στην παρουσίαση της εταιρείας Semaview (2002: 7).

οντολογίες. Οδηγούνται μέσα στον Ιστό ή περιγράφονται από αυτές, προκειμένου να εκτελέσουν κάποια ενέργεια για λογαριασμό του χρήστη.

Μία επιτυχής συνεργασία - αλληλεπίδραση μεταξύ συστημάτων από διαφορετικούς επιστημονικούς ή επιχειρηματικούς κλάδους εξαρτάται, εκτός από τα προγράμματα που αναλαμβάνουν το διαδικαστικό κομμάτι εκτέλεσης ενεργειών, κυρίως από τις οντολογίες που τους περιγράφουν. Το πρόβλημα σε αυτήν την περίπτωση είναι ότι ο κάθε κλάδος χρησιμοποιεί δική του ορολογία, με αποτέλεσμα στο μέλλον να χρειαστεί να αντιμετωπισθούν θέματα ορολογίας, πολυγλωσσίας, συνωνυμίας ή αμφισημίας όρων και εννοιών. Η σημασιολογική χαρτογράφηση εννοιών και όρων μεταξύ διαφορετικών οντολογιών θα αποτελέσει σημείο αιχμής στην εξέλιξη του Σημασιολογικού Ιστού. Μία τέτοια εργασία επιλύοντας τα προβλήματα ορολογίας, ταυτόχρονα θα συμβάλει στην προώθηση της συνεργασίας μεταξύ των ανθρώπων και των τομέων που εκπροσωπούν, στη βελτίωση των διαδικασιών ανάκτησης δεδομένων και πληροφοριών, καθώς και στην αποτελεσματικότητα Πρακτόρων και Υπηρεσιών Ιστού. Στο δύσκολο αυτό εγχείρημα, το τέλος του οποίου θα σηματοδοτήσει μία νέα εποχή για την κοινωνία της πληροφορίας, η αξία της εμπειρίας των βιβλιοθηκονόμων στην κατασκευή και διαχείριση θησαυρών είναι αδιαμφισβήτητη.

### *3.3 Αξιοποίηση της υπάρχουσας γνώσης*

Η πληθώρα μεταδεδομένων και γνώσης διαθέσιμη στον Παγκόσμιο Ιστό, θα αυξήσει τις δυνατότητες εξαγωγής συμπερασμάτων του Σημασιολογικού Ιστού, δίνοντας ώθηση σε “νοήμονες” εφαρμογές τεχνητής νοημοσύνης (Semaview 2002: 7). Αν και είναι δεδομένο πως ο συνδυασμός της υπάρχουσας γνώσης με άλλα στοιχεία και πληροφορίες μπορεί να αξιοποιηθεί σε ποικίλες και πρωτοποριακές εφαρμογές, μία πρόβλεψη του τί επιφυλάσσει ο Σημασιολογικός Ιστός κατά το στάδιο της πλήρους ανάπτυξής του πιθανότατα να χαρακτηριστεί σε αυτή τη χρονική στιγμή ως παρακινδυνευμένη. Το πώς θα αξιοποιηθούν η γνώση και οι μηχανισμοί εξαγωγής συμπερασμάτων αποτελεί θέμα ανθρώπινης έμπνευσης και τεχνολογικής ανάπτυξης. Δεν υπάρχει ακόμα η εμπειρία και η τεχνογνωσία του δεύτερου σταδίου ανάπτυξης του Σημασιολογικού Ιστού, το οποίο αναμένεται τόσο να ικανοποιήσει υπάρχουσες ανάγκες και να δημιουργήσει νέες (προκαλώντας την έμπνευση και τη δημιουργικότητα των ανθρώπων), όσο και να σηματοδοτήσει την εξέλιξη νέων, ακόμα πιο προηγμένων τεχνολογικών υποδομών.

## **4. “Νοήμων” Υπηρεσία Διαδανεισμού Επιστημονικών Δημοσιευμάτων**

Η πρότυπη εφαρμογή στη βιβλιοθηκονομία που παρουσιάζεται βασίζεται σε τεχνολογίες που ανήκουν στο δεύτερο στάδιο ανάπτυξης του Σημασιολογικού Ιστού, όπου αναμένεται να πραγματοποιηθεί η αυτοματοποίηση και η ολοκλήρωση υπηρεσιών. Αφορά στην υπηρεσία διαδανεισμού επιστημονικών δημοσιευμάτων.

Σήμερα ο διαδανεισμός επιστημονικών δημοσιευμάτων πραγματοποιείται μέσω του Συστήματος Διαδανεισμού Βιβλιοθηκών (Εθνικό Κέντρο Τεκμηρίωσης 2004) του Εθνικού Δικτύου Επιστημονικών και Τεχνολογικών Βιβλιοθηκών, το οποίο υποστηρίζεται τεχνολογικά από το Εθνικό Κέντρο Τεκμηρίωσης. Ο χρήστης αιτείται να διαδανειστεί ένα άρθρο, παρέχοντας τα πλήρη στοιχεία του. Μετά την αίτηση του χρήστη, ο βιβλιοθηκονόμος

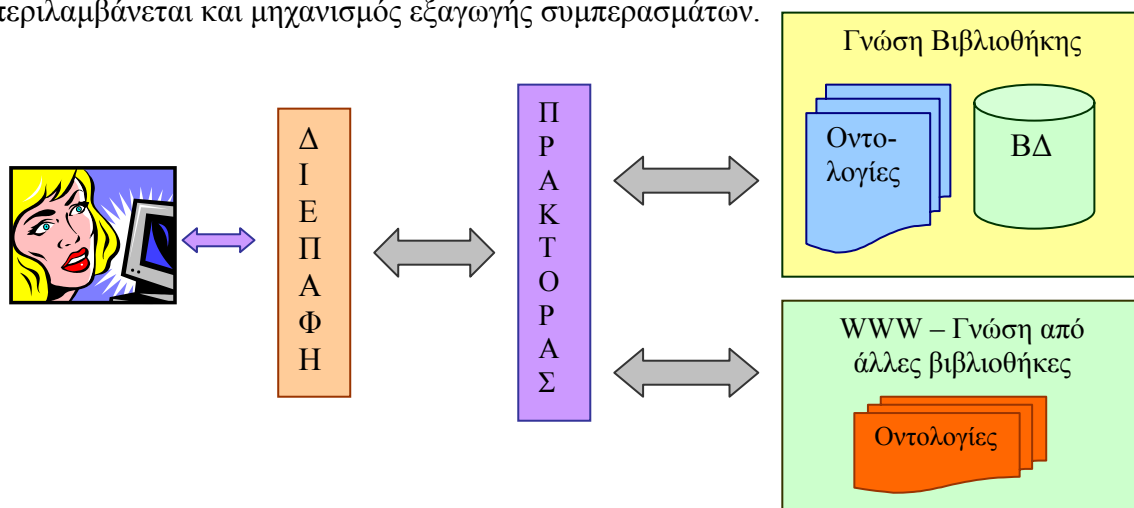


ελέγχει πρώτα αν υπάρχει το συγκεκριμένο τεκμήριο στη συλλογή της βιβλιοθήκης. Αν δεν υπάρχει, προχωρά σε αναζήτηση του τεκμηρίου στο Συλλογικό Κατάλογο Περιοδικών του Εθνικού Κέντρου Τεκμηρίωσης (ΕΚΤ). Η αναζήτηση επιστρέφει έναν κατάλογο βιβλιοθηκών, οι οποίες διαθέτουν το τεύχος περιοδικού που περιλαμβάνει το ζητούμενο άρθρο. Ο βιβλιοθηκονόμος, έπειτα, επιλέγει σε ποιες βιβλιοθήκες επιθυμεί να απευθυνθεί το ΕΚΤ για λογαριασμό του. Η επιλογή αυτή γίνεται με βάση αντικειμενικά και υποκειμενικά κριτήρια, όπως για παράδειγμα ο χρόνος ανταπόκρισης αυτής της βιβλιοθήκης σε παλαιότερα αιτήματα. Τα κριτήρια αυτά αποτελούν γνώση της βιβλιοθήκης, η οποία μπορεί να χαθεί και να μην αξιοποιηθεί σε μία επόμενη διαδικασία διαδανεισμού, αν ο συγκεκριμένος υπάλληλος δεν είναι διαθέσιμος.

Η υπηρεσία διαδανεισμού που προτείνεται προσπαθεί να καταγράψει και να αξιοποιήσει τέτοιου είδους πληροφορίες προς όφελος του χρήστη. Η πετυχημένη λειτουργία της προϋποθέτει την ύπαρξη της σχετικής γνώσης στον Παγκόσμιο Ιστό, την ύπαρξη, δηλαδή, οντολογιών που να παρέχουν την κατάλληλη πληροφόρηση. Συγκεκριμένα:

1. Οντολογία με βασικές έννοιες διαδανεισμού: π.χ. χρήστης, αίτηση, άρθρο, τεύχος, περιοδικό, διαδικασίες διαδανεισμού, κόστος, παραγγελία, κ.λπ.
2. Οντολογία με πολιτική διαδανεισμού βιβλιοθήκης: π.χ. δε φωτοτυπούνται άρθρα τευχών πριν το 1970. Η παρούσα οντολογία θα συγκριθεί από το μηχανισμό εξαγωγής συμπερασμάτων με τα στοιχεία του ζητούμενου άρθρου και τα κριτήρια που ο χρήστης έχει θέσει στην αίτησή του.
3. Οντολογία με κριτήρια επιλογής της βιβλιοθήκης, από όπου θα πραγματοποιηθεί ο διαδανεισμός του τεκμηρίου: π.χ. χρόνος ανταπόκρισης σε παλαιότερα αιτήματα, έξοδα αποστολής, κόστος διαδανεισμού για χρήστη. Στην ουσία αυτά τα κριτήρια αποτελούν κανόνες εξαγωγής συμπερασμάτων, με βάση τους οποίους το πρόγραμμα θα κάνει προτάσεις στο χρήστη για τις καλύτερες επιλογές διαδανεισμού.
4. Οντολογία με προτιμήσεις χρηστών, π.χ. παράγοντες κόστος διαδανεισμού, έξοδα αποστολής, χρόνος αποστολής, χιλιομετρική απόσταση βιβλιοθήκης που διαθέτει το τεκμήριο, τρόπος πληρωμής, κ.λπ. Η ύπαρξη αυτής της οντολογίας κρίνεται ως απαραίτητη, καθώς θα διευκολύνει στη διαδικασία εξαγωγής συμπερασμάτων σχετικά με τις προτιμήσεις των χρηστών.

Την τεχνολογική υποδομή της υπηρεσίας αποτελούν η διεπαφή χρήστη-προγράμματος (interface), οι οντολογίες που έχει αναπτύξει η βιβλιοθήκη, μία βάση, όπου αποθηκεύονται παλαιότερες ενέργειες διαδανεισμού για εξαγωγή συμπερασμάτων, και τέλος το πρόγραμμα – πράκτορας που λειτουργεί ως ο εκτελεστής ενεργειών. Στο πρόγραμμα – πράκτορας περιλαμβάνεται και μηχανισμός εξαγωγής συμπερασμάτων.



#### Σχήμα 4: Η τεχνολογική υποδομή της υπηρεσίας

Ο χρήστης, μέσω του συστήματος διεπαφής, πραγματοποιεί την αίτησή του με όλες τις σχετικές παραμέτρους (χρόνος, κόστος, τρόπος πληρωμής, κ.λπ.). Το πρόγραμμα “κατανοεί” την αίτηση με τη βοήθεια της πρώτης οντολογίας. Κατανοεί ότι υπάρχει ένας άνθρωπος, ότι αυτός είναι χρήστης της βιβλιοθήκης και ότι επιθυμεί να προμηθευθεί ένα συγκεκριμένο άρθρο που βρίσκεται σε συγκεκριμένο τεύχος ενός περιοδικού. Κατανοεί με τη βοήθεια της τέταρτης οντολογίας πως ο χρήστης έχει κάποιες προτιμήσεις αναφορικά με τη διαδικασία διαδανεισμού, π.χ. επιθυμεί το κόστος να μην υπερβαίνει ένα συγκεκριμένο ποσό και να έχει στα χέρια του το άρθρο μέσα σε ορισμένα χρονικά πλαίσια. Ο χρόνος κατανόησης των παραπάνω δεδομένων είναι μηδαμινός, καθώς πρόκειται για γνώση που ήδη κατέχει το σύστημα.

Το πρόγραμμα αναζητά στον OPAC της βιβλιοθήκης αν υπάρχει το συγκεκριμένο τεύχος περιοδικού. Αν ναι, ενημερώνει το χρήστη και αυτός επιλέγει είτε να το αναζητήσει μόνος του, είτε να του εκτυπωθεί/φωτοτυπηθεί (ανάλογα με το υπόστρωμα του περιοδικού) και να του παραδοθεί από το προσωπικό της βιβλιοθήκης. Αν ο πράκτορας δεν εντοπίσει το τεύχος στη συλλογή της βιβλιοθήκης, ξεκινά την αναζήτηση σε βιβλιοθήκες οι οποίες δηλώνουν μέσα από τους ιστοχώρους τους ότι διαθέτουν υπηρεσία διαδανεισμού. Η διαδικασία αναζήτησης αποκλείει κάποιες βιβλιοθήκες με γνώμονα το αν ταιριάζει η πολιτική διαδανεισμού τους με τα κριτήρια που έχει θέσει ο χρήστης. Σε αυτό το σημείο αξιοποιείται η οντολογία που εκφράζει την πολιτική διαδανεισμού της κάθε βιβλιοθήκης (2<sup>η</sup> οντολογία). Έπειτα το πρόγραμμα αποκλείει όσες δε διαθέτουν το τεύχος/τον τόμο, όπου περιλαμβάνεται η αιτούμενη επιστημονική δημοσίευση. Η αναζήτηση επιστρέφει έναν κατάλογο βιβλιοθηκών οι οποίες μπορούν να εξυπηρετήσουν το χρήστη. Το πρόγραμμα ταυτόχρονα προτείνει κάποια επιλογή ως τη καλύτερη· λαμβάνει υπόψη τις προτιμήσεις του χρήστη, τα κριτήρια επιλογής βιβλιοθήκης για συνεργασία (3<sup>η</sup> οντολογία) και τη γνώση που εξάγεται από τις αποθηκευμένες ενέργειες διαδανεισμού.

Αν το πρόγραμμα δεν έχει στη διάθεσή του πολλά στοιχεία από μέρος του χρήστη, προχωρά σε εξαγωγή συμπερασμάτων από προτιμήσεις άλλων χρηστών σε παλαιότερες ενέργειες διαδανεισμού, π.χ. το 70% των χρηστών που έκαναν χρήση της υπηρεσίας έδινε βαρύτητα στον παράγοντα κόστος. Ο χρήστης επιλέγει ανάμεσα στις βιβλιοθήκες που του προτείνει το σύστημα και δίνει την τελική έγκριση για την πραγματοποίηση του διαδανεισμού. Θα υπάρχουν σίγουρα και περιπτώσεις, όπου η αναζήτηση δε θα μπορεί να επιστρέψει κατάλογο βιβλιοθηκών, που να ικανοποιούν πλήρως τις προτιμήσεις του χρήστη. Σε αυτήν την περίπτωση, το πρόγραμμα αναλαμβάνει να προτείνει εναλλακτικές λύσεις βασίζόμενο πάντα στις προτιμήσεις του χρήστη. Μπορεί για παράδειγμα, αν ο χρήστης όρισε ως σημαντικότερο παράγοντα για αυτόν το κόστος, να του προταθεί η επίσκεψη σε μια κοντινή βιβλιοθήκη, η οποία δε διαδανείζει το τεύχος που αυτός επιθυμεί, αλλά επιτρέπει τη μελέτη του μέσα στους χώρους της.

Η ύπαρξη των σχετικών οντολογιών αποτελεί το πρώτο βήμα για την ανάπτυξη προηγμένων υπηρεσιών, όπως αυτή που περιγράφηκε. Κατά το σχεδιασμό τέτοιων υπηρεσιών απαιτείται διεξοδική μελέτη κυρίως όσον αφορά στον καθορισμό των παραμέτρων, σύμφωνα με τις οποίες αυτές θα λειτουργούν και θα εξάγουν συμπεράσματα.. Στόχος του παρόντος παραδείγματος ήταν η επίδειξη του τρόπου αξιοποίησης των οντολογιών και όχι η λεπτομερής περιγραφή ενός πολύπλοκου συστήματος, όπως είναι ένα σύστημα διαδανεισμού επιστημονικών δημοσιευμάτων.

## 5. Συμπεράσματα

Το ότι ένας ηλεκτρονικός υπολογιστής ποτέ δε θα μπορεί πλήρως να κατανοήσει –με την έννοια της τεχνητής νοημοσύνης- τι σημαίνει ένα έγγραφο είναι δεδομένο. Ο Σημασιολογικός Ιστός αποτελεί μία λύση με την οποία οι Η/Υ συμφωνώντας στο νόημα εννοιών και όρων μπορούν να ανταλλάσσουν μηνύματα και πληροφορίες. Η περιγραφή των πληροφοριών πραγματοποιείται με το Πλαίσιο Περιγραφής Πόρων-RDF και η απόδοση του νοήματος με τη βοήθεια των οντολογιών. Τα βασικά εργαλεία ανάπτυξης του Σημασιολογικού Ιστού ήδη υπάρχουν, γεγονός που επιτρέπει στους επιστήμονες της πληροφορίας να αξιοποιήσουν την εμπειρία τους στη διαχείριση θησαυρών και να εστιάσουν στην ανάπτυξη οντολογιών.

Η αναπαράσταση της πληροφορίας σε μορφή κατανοητή από τα υπολογιστικά συστήματα ανοίγει νέους ορίζοντες στην εξέλιξη εφαρμογών και υπηρεσιών. Το οικοδόμημα του Σημασιολογικού Ιστού, που τώρα θεμελιώνεται, προβάλλει νέους ρόλους στους σύγχρονους διαχειριστές της πληροφορίας: οι υπάρχουσες πληροφορίες και γνώσεις θα μεταφερθούν σε μηχαναγνώσιμη μορφή και θα αποτελέσουν τη μνήμη, τη γνώση που θα αξιοποιηθεί από πλήθος επιστημονικών, επιχειρηματικών και καθημερινών εφαρμογών του μέλλοντος.

Μέσα σε αυτά τα πλαίσια οι Έλληνες επιστήμονες της πληροφορίας καλούνται να συμμετάσχουν ενεργά στη μετατροπή της εγχώριας γνώσης σε μορφή κατανοητή από τους ηλεκτρονικούς υπολογιστές και στην ανάπτυξη των τεχνολογικών εκείνων υποδομών που θα επιτρέψουν σε αυτή τη γνώση να αξιοποιηθεί και πέρα από τα ελληνικά σύνορα.

## 6. Ευχαριστίες

Θερμές ευχαριστίες οφείλονται στον καθηγητή μου στο Πανεπιστήμιο Μακεδονίας, κύριο Κωνσταντίνο Ταραμπάνη, που με εισήγαγε στον κόσμο του Σημασιολογικού Ιστού. Ακόμη, θα ήθελα να ευχαριστήσω τους συναδέλφους από το Πανεπιστήμιο Θεσσαλίας και συγκεκριμένα την Ιωάννα Ζορμπά, της οποίας οι απορίες αποτέλεσαν το έναυσμα για το παρόν άρθρο, και τους Ιφιγένεια Βαρσαμά, Δημήτρη και τη Μίνα Πατρινού για την υποστήριξή τους.

## 7. Βιβλιογραφία

Εθνικό Κέντρο Τεκμηρίωσης (2004) Σύστημα Διαδανεισμού Βιβλιοθηκών του ΕΚΤ [Διαθέσιμο στο: <http://www.ekt.gr/tolib-serv/edetb/interlib-loan.htm>] (23/08/2004)

Ζαπουνίδου, Σοφία (2004) *RDF μοντελοποίηση και εφαρμογές: περιγραφή και ανάκτηση αρχειακών συλλογών με τη χρήση του RDF*. Διατριβή, Πανεπιστήμιο Μακεδονίας Μεταπτυχιακό Πρόγραμμα στα Πληροφοριακά Συστήματα..

Τομαή, Ελένη, Κάβουρας Μαρίνος (2002) “Ανάπτυξη Γεωγραφικών Οντολογιών” [Διαθέσιμο στο: [http://ontogeo.ntua.gr/publications/Greek%20publications/Tomai\\_Kavouras\\_hellasgi2002.pdf](http://ontogeo.ntua.gr/publications/Greek%20publications/Tomai_Kavouras_hellasgi2002.pdf)] (24/12/2003)

- Berners-Lee, Tim (1998) *Web Architecture from 50,000 feet* [Διαθέσιμο στο <http://www.w3.org/DesignIssues/Architecture.html>] (07/07/2004)
- Berners-Lee, Tim, James Hendler και Ora Lassila (2001) “The Semantic Web: A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities” *Scientific American* May [Διαθέσιμο στο <http://www.sciam.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21>] (07/07/2004)
- Broekstra, Jeen et.al. (2000), “Adding formal semantics to the Web: building on top of RDF Schema”. [Διαθέσιμο στην ηλεκτρονική διεύθυνση: <http://www.ontoknowledge.org/oil/extending-rdfs.pdf>] (16/10/2003)
- ESW Wiki FAQs (2003) Απάντηση στο ερώτημα “*What is the difference between an ontology (OWL) and a schema (RDFS)?*”. Το ερώτημα αυτό στάλθηκε στις 29/09/2003 από το χρήστη nmg στη λίστα rdfweb-dev, από όπου και αντιγράφηκε. [Διαθέσιμο στην ηλεκτρονική διεύθυνση: <http://esw.w3.org/topic/SchemaVsOntology>] (10/12/2003)
- Fisseha, Frehiwot, Anita Liang, Johannes Keizer (2003) “Reengineering AGROVOC to Ontologies: Step towards better semantic structure”, Στο *NKOS Workshop*, Houston, Texas: Rice University, 31 May. [Διαθέσιμο στην ηλεκτρονική διεύθυνση: <http://nkos.slis.kent.edu/2003workshop/fissehaontology.ppt>] (16/10/2003)
- Gruber, Thomas R. (1993) “Toward Principles for the Design of Ontologies Used for Knowledge Sharing”. Technical Report KSL 93-04, Knowledge Systems Laboratory, Stanford University. [Διαθέσιμο στο: <http://www.cise.ufl.edu/~jhammer/classes/6930/XML-FA02/papers/gruber93ontology.pdf>] (27/07/2004).
- Hendler, James, Tim Berners-Lee και Eric Miller (2002) “Integrating Applications on the Semantic Web”, *Journal of the Institute of Electrical Engineers of Japan* Vol 122 (10), pp. 676-680. [Διαθέσιμο στο [www.w3.org/2002/07/swint](http://www.w3.org/2002/07/swint)] (20/10/2003)
- Ianella, Renato (1999) “An Idiot's Guide to the Resource Description Framework”. [Διαθέσιμο στην ηλεκτρονική διεύθυνση: <http://archive.dstc.edu.au/RDU/reports/RDF-Idiot/>] (23/07/2004)
- Lassila, Ora, Ralph R. Swick, Resource Description Framework (RDF) Model and Syntax Specification: W3C Recommendation 22 February 1999, [Διαθέσιμο σε <http://www.w3.org/TR/1999/REC-rdf-syntax-19990222>] (20/01/2004).
- Lu, Shiyong, Ming Dong και Farshad Fotouhi (2002) “The Semantic Web: opportunities and challenges for next-generation Web applications” *Information Research* 7(4) [Διαθέσιμο σε <http://InformationR.net/ir/7-4/paper134.html>] (18/05/2004)]
- Metamodel.com (2003) “What are the differences between a vocabulary, a taxonomy, a thesaurus, an ontology, and a meta-model?” [Διαθέσιμο στο: <http://www.metamodel.com/article.php?story=20030115211223271>] (10/10/2003).
- Miller, Eric (1998) “An introduction to the Resource Description Framework”, *D-Lib Magazine* May. [Διαθέσιμο στην ηλεκτρονική διεύθυνση [www.dlib.org/dlib/may98/miller/05miller.html](http://www.dlib.org/dlib/may98/miller/05miller.html)] (07/09/2003)

Semaview Inc (2002), *Concept to Reality: What the emerging Semantic Web means to your business* [Διαθέσιμο στο <http://www.semaview.com/d/whitepaper.pdf>] (07/07/2004).

Swartz, Aaron (2002) “The Semantic Web in Breadth”. [Διαθέσιμο στην ηλεκτρονική διεύθυνση: <http://logicerror.com/semanticWeb-long>] (07/09/2003)