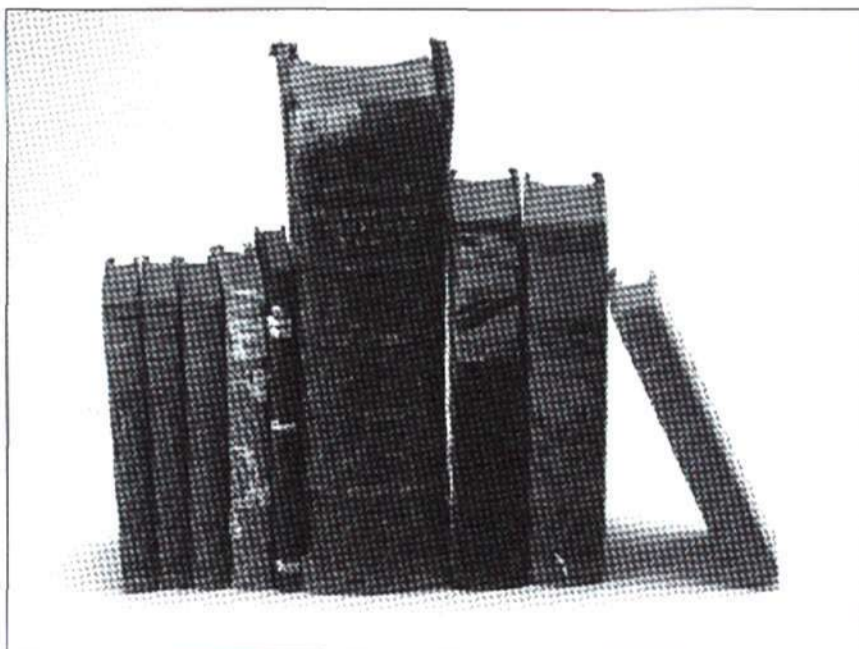


# Η Συμβολή του Σημσιολογικού Ισοτύπου στις Ακαδημαϊκές Βιβλιοθήκες

Δημήτριος Κανελλόπουλος - Ανδριανή Ζαφειροπούλου

## 1. Ακαδημαϊκές Βιβλιοθήκες

Οι ακαδημαϊκές βιβλιοθήκες αποτελούν κύρια πηγή οργανωμένης γνώσης, ενώ παράλληλα υποστηρίζουν το εκπαιδευτικό και ερευνητικό έργο που διεξάγεται στα Ακαδημαϊκά Ιδρύματα. Οι βασικές υπηρεσίες των ακαδημαϊκών βιβλιοθηκών είναι: ο δανεισμός, ο διαδανεισμός, η πρόσκτηση υλικού, η καταλογογράφηση, η διακίνηση, η αποθήκευση και προστασία του υλικού, η αναζήτηση, η εύρεση, καθώς και η δυνατότητα πρόσβασης και διάθεσης του πληροφοριακού υλικού. Σήμερα, το πληροφοριακό υλικό των ακαδημαϊκών βιβλιοθηκών δεν υπάρχει μόνο σε έντυπη, αλλά και σε ψηφιακή μορφή. Η ραγδαία εξέλιξη της Πληροφορικής και των Επικοινωνιών έχει επιφέρει ριζικές αλλαγές στη μορφή και στο υλικό των ακαδημαϊκών βιβλιοθηκών. Οι ψηφιακές ακαδημαϊκές βιβλιοθήκες παρέχουν μέσα από το Διαδίκτυο τη δυνατότητα πρόσβασης και ανάκτησης μεγάλου όγκου πληροφοριών. Η τεχνολογική αυτή εξέλιξη επιφέρει μεγάλες προκλήσεις και προβλήματα που σχετίζονται με την περιγραφή, ταξινόμηση, οργάνωση, αξιοπιστία και εμπιστευτικότητα των διατιθέμενων πληροφοριών. Σαφώς, η ετερογένεια των δεδομένων (data heterogeneity) είναι το κυριότερο πρόβλημα, διότι τα διάφορα πληροφοριακά συστήματα των βιβλιοθηκών χρησιμοποιούν διαφορετικά μεταδεδομένα για την αναπαράσταση των βιβλιογραφικών τους δεδομένων (SDLIP, 1999).



## 1.1 Μεταδεδομένα βιβλιοθηκών

Τα μεταδεδομένα (metadata) είναι απαραίτητα για τη σωστή περιγραφή, οργάνωση, αποθήκευση, αναζήτηση, εύρεση και ανάκτηση των διατιθέμενων πληροφοριών, διότι οι πληροφορίες στο Διαδίκτυο είναι ανομοιογενείς. Σε πολλές ακαδημαϊκές βιβλιοθήκες, χρήσιμες ακαδημαϊκές πηγές του Διαδικτύου καταλογογραφούνται και εισάγονται στους καταλόγους τους. Συγκεκριμένα, ορισμένες βιβλιοθήκες παρέχουν τις βιβλιογραφικές τους εγγραφές σε αρχεία μεταδεδομένων, έτσι ώστε οι εγγραφές των φυσικών τους συλλογών υλικού να είναι προσπελάσιμες από αυτές των ψηφιακών βιβλιοθηκών. Ο καθορισμός των μεταδεδομένων εξαρτάται από: 1) τον τύπο του πληροφοριακού υλικού που φιλοξενεί μια βιβλιοθήκη, 2) τις απαιτήσεις αναζήτησης, και 3) τις ανάγκες πληροφόρησης των χρηστών. Σήμερα, έχουν αναπτυχθεί διάφορα πρότυπα μεταδεδομένων με διαφορετικά επίπεδα πολυπλοκότητας ως προς την περιγραφή των δεδομένων (π.χ. Dublin Core Metadata Initiative, MARC, PRISM, ONIX for books). Ωστόσο, δεν έχει ακόμη καθιερωθεί ένα διεθνές πρότυπο μεταδεδομένων για τις ακαδημαϊκές βιβλιοθήκες. Κάθε αρχιτεκτονική μεταδεδομένων βασίζεται στην περιγραφή των πεδίων του καταλόγου (π.χ. συγγραφέας, τίτλος, ημερομηνία, εκδότης), καθώς και στη χρήση λεξιλογίων (vocabularies) που ελέγχονται και που επιτρέπουν την πρόσβαση στους καταλόγους των βιβλιοθηκών.

Οι ακαδημαϊκές βιβλιοθήκες χρησιμοποιούν πολύπλοκες δομές περιγραφής των βιβλιογραφικών τους δεδομένων. Ωστόσο, οι αρχιτεκτονικές μεταδεδομένων προσφέρουν μόνο ένα πολύ μικρό ποσοστό αυτής της πολυπλοκότητας, γεγονός που οδηγεί σε ικανοποιητικά επίπεδα διαλειτουργικότητας (interoperability) και χρηστικότητας (usability) των μεταδεδομένων στο Διαδίκτυο (Paercke et al., 1998). Επιπλέον, για την καταλογογράφηση των δεδομένων και την περιγραφή των επιστημονικών εγγραφών, απαιτείται η πρόσληψη βιβλιοθηκονόμων που θα έχουν εξειδικευμένες γνώσεις σε συγκεκριμένες θε-

ματικές ενότητες. Σήμερα, οι πηγές των ακαδημαϊκών βιβλιοθηκών καλύπτουν ένα ευρύ φάσμα θεματικών περιοχών και οι μη ειδικευμένοι βιβλιοθηκονόμοι δεν μπορούν να αντεπεξέλθουν στις απαιτήσεις που επιβάλλει μια λεπτομερειακή και πλήρη περιγραφή των βιβλιογραφικών δεδομένων.

Στην εργασία αυτή παρουσιάζουμε τις τεχνολογίες του Σημασιολογικού Ιστού που προάγουν τη διαλειτουργικότητα, τη δυνατότητα επαναχρησιμοποίησης και διαμοιρασμού του ψηφιακού υλικού των ακαδημαϊκών βιβλιοθηκών. Τεκμηριώνεται η άποψη ότι ο Σημασιολογικός Ιστός αποτελεί μία πλατφόρμα για σημασιολογικές μηχανές αναζήτησης, διαμεσολαβητές πληροφορίας και έξυπνους πράκτορες λογισμικού. Οι δυνατότητες αυτές του Σημασιολογικού Ιστού βελτιώνουν σημαντικά τον τρόπο λειτουργίας των ακαδημαϊκών βιβλιοθηκών.

## 2. Τεχνολογίες Σημασιολογικού Ιστού στις Ακαδημαϊκές Βιβλιοθήκες

### 2.1 Σημασιολογικός Ιστός

Ο Σημασιολογικός Ιστός (Semantic Web) υιοθετήθηκε από το World Wide Web Consortium (W3C), έναν Οργανισμό που στοχεύει στην προώθηση, ανάπτυξη και εξέλιξη του Web και των πρωτοκόλλων που το υποστηρίζουν. Ο Σημασιολογικός Ιστός (<http://www.w3c.org/2001/sw/>) είναι μία επέκταση και βελτίωση του σημερινού Web στην κατεύθυνση, κυρίως της δόμησης της πληροφορίας, έτσι ώστε αυτή να είναι προσιτή από εφαρμογές υπολογιστών, με τελικό στόχο την αυτοματοποίηση πολλών λειτουργιών στο Διαδίκτυο. Η σημερινή αναπαράσταση της πληροφορίας που προορίζεται για χρήση από ανθρώπους θα αντικατασταθεί από μία αναπαράσταση κατανοητή από υπολογιστές. Ο Σημασιολογικός Ιστός επιτρέπει στο περιεχόμενο του Διαδικτύου να

αυτοπροσδιορίζεται (Berners-Lee et al., 2001). Ο αυτοπροσδιορισμός αυτός επιτρέπει στους χρήστες (λ.χ. στους βιβλιοθηκονόμους) να υποβάλλουν ερωτήσεις και να επάγουν γνώση από τις πληροφορίες του Διαδικτύου γρήγορα και αυτόματα. Ο Σημασιολογικός Ιστός αποτελεί μία πλατφόρμα για έξυπνες μηχανές αναζήτησης, διαμεσολαβητές πληροφορίας και έξυπνους πράκτορες λογισμικού. Επίσης, προάγει τη διαλειτουργικότητα, τη δυνατότητα επαναχρησιμοποίησης και διαμοιρασμού του πληροφοριακού υλικού των ακαδημαϊκών βιβλιοθηκών (Sure & Studer, 2005). Επιπρόσθετα, οι σημασιολογίες (semantics) μπορούν να χρησιμοποιηθούν για την ανακάλυψη, σύνθεση και τον έλεγχο των υπηρεσιών Ιστού (Web Services) (Ouzzani & Bouguettaya, 2004). Με την υποστήριξη του Σημασιολογικού Ιστού, τα απομονωμένα σημασιολογικά τμήματα του πληροφοριακού υλικού των ακαδημαϊκών βιβλιοθηκών θα μπορούν να συνδεθούν μεταξύ τους, και ο χρήστης θα μπορεί να βρει εύκολα τις βιβλιογραφικές πηγές.

Η ανάπτυξη του Σημασιολογικού Ιστού προχωρά σταδιακά, και σε κάθε στάδιο κτίζεται ένα επίπεδο στην κορυφή ενός άλλου (Σχήμα 1). Στην αρχιτεκτονική του Σημασιολογικού Ιστού προσφέρεται ένα καθολικό σύστημα διευθυνσιοδότησης πόρων. Πόρος (resource) είναι οτιδήποτε θέλουμε να δηλώσουμε ή να περιγράψουμε. Για παράδειγμα, πόρος μπορεί να είναι μία ιστοσελίδα, μία δικτυακή τοποθεσία, ένα αντικείμενο, μία έννοια κτλ. Κάθε πόρος προσδιορίζεται με το Καθολικό Αναγνωριστικό Πόρου (Universal Resource Identifier-URI).

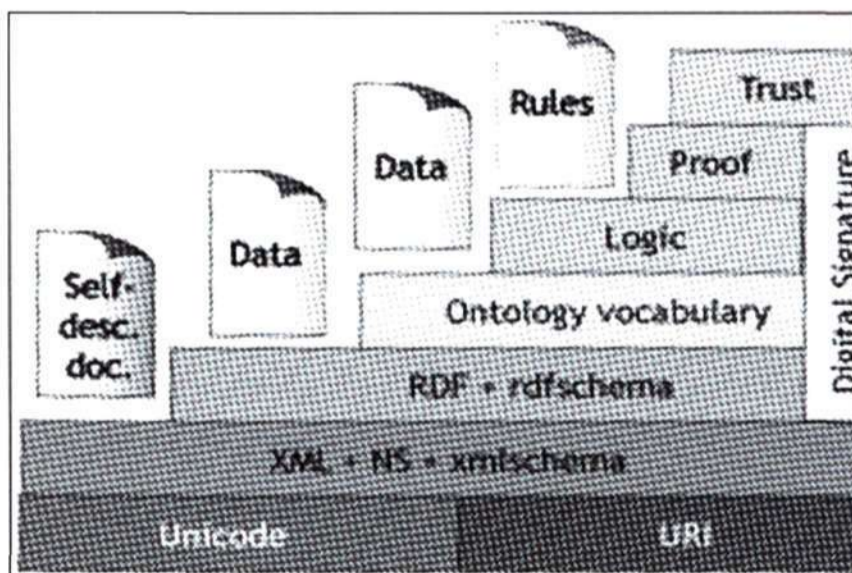
- Στο χαμηλότερο επίπεδο συναντάμε την XML, μια γλώσσα που επιτρέπει τη συγγραφή δομημένων εγγράφων Web με ένα καλώς ορισμένο λεξιλόγιο. Η XML είναι κατάλληλη για την αποστολή εγγράφων μέσω του Web, και ως εκ τούτου υποστηρίζει τη συντακτική διαλειτουργικότητα (syntactic interoperability) των δεδομένων.

- Το πρότυπο RDF είναι ένα βασικό μοντέλο δεδομένων για τη συγγραφή απλών δηλώσεων που αναφέρονται σε αντικείμενα του Web (Web objects).

- Το RDF Schema βασίζεται στη γλώσσα RDF και παρέχει τρόπους μοντελοποίησης των αντικειμένων του Web προκειμένου να οργανωθούν ιεραρχικά.

- Για την αναπαράσταση πιο πολύπλοκων σχέσεων μεταξύ των αντικειμένων του Web, υπάρχουν κατάλληλότερες γλώσσες οντολογιών, όπως η OWL.

- Το επίπεδο λογικής (logic) βελτιώνει περαιτέρω



Σχήμα 1. Η αρχή του σημασιολογικού ιστού

τη γλώσσα οντολογίας και επιτρέπει τη συγγραφή δηλωτικής γνώσης που είναι ειδική με την εφαρμογή.

- Το επίπεδο απόδειξης (proof) υλοποιεί τη διεργασία επαγωγής, καθώς και την αναπαράσταση και επικύρωση των αποδείξεων στις γλώσσες Web. Μία αξιολογη υποδομή οντολογιών που υποστηρίζει κανόνες επαγωγής γνώσης είναι η OIL (Fensel et al., 1999).

- Τέλος, το επίπεδο εμπιστοσύνης (trust) βασίζεται στη χρήση των ψηφιακών υπογραφών (digital signatures) και ασχολείται με τη γνησιότητα της πληροφορίας.

### *XML και XML Schema*

Η XML (eXtensible Markup Language - επεκτάσιμη γλώσσα σήμανσης) είναι η επικρατέστερη γλώσσα για την περιγραφή και ανταλλαγή κειμένων στο Διαδίκτυο. Η XML παρέχει τη δυνατότητα δημιουργίας κειμένων με απεριόριστα πολύπλοκη δομή και συντακτικό (Liebetrau, 2004). Έτσι, μπορούν να δομηθούν οι πληροφορίες που περιέχονται στα κείμενα για να επεξεργάζονται πιο εύκολα από τους υπολογιστές. Το πρότυπο XML συμπληρώνεται από το πρότυπο XML Schema (<http://www.w3.org/XML/Schema/>), μια γλώσσα με την οποία γράφουμε "λεξικά" και "γραμματικές" για XML κείμενα. Το XML Schema ορίζει τα επιτρεπόμενα στοιχεία, τις ιδιότητές τους, καθώς και τον τρόπο με τον οποίο συνδυάζονται μεταξύ τους μέσα στο XML κείμενο. Με απλά λόγια, το XML Schema αποτελεί το "συντακτικό" του XML κειμένου.

### *RDF και RDF Schema*

Η γλώσσα RDF (Resource Description Framework) ή Περιβάλλον Περιγραφής Πόρων είναι το πρότυπο που υιοθετήθηκε από το W3C για την περιγραφή πληροφοριακών πόρων και γενικότερα για την αναπαράσταση της γνώσης στο περιβάλλον του Διαδικτύου. Μέσω του RDF είναι δυνατή η μετατροπή της πληροφορίας σε σημασιολογική (Miller et al., 1999). Η γλώσσα RDF είναι ένα απλό μοντέλο δεδομένων, στο οποίο όλες οι προτάσεις αποτελούν μια συνάρτηση τριών μεταβλητών της μορφής: P(O, V) όπου Property (ιδιότητα), Object (αντικείμενο), Value(τιμή). Η συντομική πλειοψηφία της γνώσης των βιβλιοθηκών που θέλουμε να αναπαραστήσουμε μπορεί να αναπαρασταθεί με αυτή την μορφή. Στο περιβάλλον RDF, το αντικείμενο και η ιδιότητα δηλώνονται με ένα URI. Η τιμή μπορεί να δηλώνεται με ένα URI ή μπορεί να είναι ένα αλφαριθμητικό ή μια λέξη. Τέλος, το πρότυπο RDF ορίζει το συντα-

κτικό XML, δηλαδή τον τρόπο με τον οποίο οι προτάσεις RDF εκφράζονται ως XML κείμενα. Το RDF Schema (<http://www.w3.org/TR/2000/CR-rdf-schema-20000327/>) είναι η οντοκεντρική επέκταση του RDF. Είναι μια γλώσσα με την οποία το μοντέλο δεδομένων του RDF εμπλουτίζεται με χαρακτηριστικά αντικειμενοστραφούς αναπαράστασης, όπου ο πόρος αντιστοιχεί σε αντικείμενο. Συγκεκριμένα, το RDF Schema ορίζει ένα λεξικό για να εκφράζονται οι κατηγορίες (κλάσεις) των πόρων, οι πόροι, οι ιδιότητές τους και οι μεταξύ τους σχέσεις.

### *2.3 Οντολογίες και OWL*

Σύμφωνα με τον Hendler (2001: 30) μία οντολογία (ontology) είναι «μία ομάδα όρων γνώσης που περιλαμβάνει το λεξιλόγιο, τις σημασιολογικές διασυνδέσεις, και μερικούς απλούς κανόνες επαγωγής και λογικής για κάποιο συγκεκριμένο θέμα». Μια οντολογία εμπεριέχει τις κλάσεις των οντοτήτων, τις σχέσεις μεταξύ των οντοτήτων, και τα αξιώματα που εφαρμόζονται σε αυτές (Mizoguchi, 2004). Για την ολοκλήρωση και κατανόηση του πληροφοριακού υλικού των βιβλιοθηκών από τους υπολογιστές, είναι απαραίτητη η ανάπτυξη οντολογιών βιβλιοθήκης (library ontologies). Στη βιβλιοθηκονομία, η επιτυχία του Σημασιολογικού Ιστού θα εξαρτηθεί κυρίως από τη δημιουργία κατάλληλων οντολογιών βιβλιοθήκης που θα προσφέρουν την υποδομή για την αντιμετώπιση των ετερογενών αναπαραστάσεων του πληροφοριακού υλικού. Οι οντολογίες βιβλιοθήκης περιέχουν γνώση σχετική με τις βιβλιοθήκες και δημιουργούνται με επεξεργαστές οντολογιών, όπως ο Protigi 2000 (Protigi, 2000), ο οποίος βασίζει τη λειτουργία του στη γλώσσα Java με OWL Plugin. Αυτό έχει ως αποτέλεσμα κάθε οντολογία βιβλιοθήκης που δημιουργείται με το Protigi 2000 να αποτελεί μία εφαρμογή διαδικτύου (applet). Στις οντολογίες βιβλιοθήκης, η δυσκολία εντοπίζεται στο ότι οι κοινότητες των βιβλιοθηκονόμων θα πρέπει να συμφωνήσουν στην οντολογική περιγραφή των βιβλιογραφικών τους δεδομένων. Η ετερογένεια των δεδομένων θα επιλυθεί με την παροχή σημασιολογικής διαμεσολάβηση μεταξύ των διαφορετικών οντολογιών και ετερογενών συστημάτων. Στο μέλλον, οι σημασιολογικές υπηρεσίες βιβλιοθηκονομίας που θα συναντάμε στο Διαδίκτυο θα συνδέονται συχνά με διαφορετικές οντολογίες βιβλιοθήκης. Στην περίπτωση αυτή, θα πρέπει να επιλυθεί το πρόβλημα της

αντιστοίχισης μεταξύ των διαφορετικών οντολογιών (ontology mapping). Πρόσφατα, στο πρόβλημα αυτό προτάθηκαν διάφορες λύσεις, όπως αυτή που υιοθετείται στο σύστημα GLUE (Doan et al, 2002).

Η OWL (Web Ontology Language - Γλώσσα Οντολογιών Ιστού) είναι μία γλώσσα που χρησιμοποιείται για την περιγραφή των οντολογιών (McGuinness and Harmelen, 2003). Η κοινοπραξία W3C έχει θεσπίσει πρόσφατα την OWL (<http://www.w3.org/2004/OWL/>) ως την προκαθορισμένη γλώσσα για την αναπαράσταση των οντολογιών. Η OWL παρέχει μεγαλύτερη δυνατότητα διερμηνείας του περιεχομένου του Web, από ότι η XML, η RDF και το RDF-Schema.

#### 2.4 Οι πράκτορες λογισμικού

Οι πράκτορες λογισμικού (software agents) μπορούν να εκμεταλλευτούν την κωδικοποιημένη σημασιολογία στο Web (Hendler, 2001). Είναι προγράμματα που βασίζουν τη λειτουργία τους στο Διαδίκτυο και δημιουργούνται για να ενεργούν αυτόνομα. Συνήθως, περιδιαβαίνουν το Διαδίκτυο και επεξεργάζονται τις πληροφορίες που βρίσκουν στις ιστοσελίδες που επισκέπτονται. Οι πράκτορες λογισμικού συχνά χρησιμοποιούνται για λειτουργίες όπως: εύρεση, ταξινόμηση και επιλογή δεδομένων. Στις ακαδημαϊκές βιβλιοθήκες, μερικά παραδείγματα λειτουργιών τους είναι: η αρχειοθέτηση πληροφοριακού υλικού, η ειδοποίηση για την εμφάνιση νέου υλικού σε άλλες ψηφιακές βιβλιοθήκες κ.α. Αυτόματοι πράκτορες λογισμικού θα μπορούσαν να χρησιμοποιούν έξυπνα το πληροφοριακό υλικό, αν προστεθούν μεταδεδομένα στο υλικό, και επιπλέον αναπαρασταθούν κατάλληλα τα μεταδεδομένα και το υλικό.

Ο Σημασιολογικός Ιστός θα χρησιμοποιεί μία ποικιλία πρακτόρων λογισμικού που θα είναι προσανατολισμένοι: (1) στον αναγνώστη-χρήστη, (2) στον βιβλιοθηκονόμο και (3) στο υλικό, προκειμένου να βελτιωθούν διάφορες διαδικασίες. Για παράδειγμα, ένας πράκτορας λογισμικού που είναι προσανατολισμένος σε εργασίες του βιβλιοθηκονόμου, θα αναλαμβάνει πολλές απλές, αλλά χρονοβόρες εργασίες του: τήρηση αρχείων, έλεγχος εγγράφων κλπ. Οι πράκτορες λογισμικού εργασιών βιβλιοθηκονόμου θα επικοινωνούν με τους πράκτορες λογισμικού αναγνώστη και θα βοηθούν στον χρονοπρογραμματισμό των εργασιών, καθώς θα προγραμματίζουν τον προσωπικό χρόνο μεταξύ των βιβλιοθηκονόμων και των αναγνωστών για να θα μεγιστοποιήσουν την επί-

δραση αυτής της επικοινωνίας. Από την άλλη μεριά, οι πράκτορες λογισμικού αναγνώστη θα βοηθούν τους αναγνώστες να εργαστούν συνεργατικά, βρίσκοντας τις πηγές εμπειρίας τους. Επιπλέον, θα βοηθούν τους αναγνώστες στην τεκμηρίωση και στην αρχειοθέτηση των προϊόντων τους.

### 3. Έργα εφαρμογής του Σημασιολογικού Ιστού στις Ακαδημαϊκές Βιβλιοθήκες

- Το έργο SWAD (SWAD) παρέχει το Απλό Σύστημα Οργάνωσης Γνώσης (Simple Knowledge Organisation System), ένα σύνολο δηλαδή εργαλείων για την περιγραφή ελεγχόμενων λεξιλογίων και ταξινομήσεων (SKOS). Το σύστημα παρέχει ορισμένα παραδείγματα θησαυρών που χρησιμοποιούν αυτά τα σχήματα, έτσι ώστε οι φυλλομετρητές (browsers) να επιλέγουν όρους από έναν συγκεκριμένο θησαυρό. Το έργο αυτό έχει αναλάβει η Ομάδα Εργασίας του W3C: Semantic Web Best Practices and Deployment Working Group.

- Το έργο SIMILE (Semantic Interoperability of Metadata and Information in unLike Environments: <http://simile.mit.edu/>) είναι ένα πρόγραμμα, στο οποίο συμμετέχουν το W3C, οι βιβλιοθήκες του MIT, και το Εργαστήριο MIT CSAIL (Computer Science and Artificial Intelligence Laboratory). Το έργο στοχεύει στη διαλειτουργικότητα των ψηφιακών σχημάτων των βιβλιοθηκών, λεξιλογίων, οντολογιών, μεταδεδομένων και υπηρεσιών (Butler et al., 2004). Βασικός στόχος του έργου είναι οι συχνή διανομή των συλλογών των βιβλιοθηκών στα μεμονωμένα, κοινοτικά και θεσμικά παραρτήματα των βιβλιοθηκών. Στόχος του SIMILE είναι η παροχή υπηρεσιών τελικού χρήστη, οι οποίες θα βασίζονται σε συγκεκριμένα σχήματα, λεξιλόγια, οντολογίες και μεταδεδομένα που βρίσκονται στα παραρτήματα των βιβλιοθηκών. Το SIMILE εφαρμόζει μια ψηφιακή αρχιτεκτονική διάδοσης ψηφιακών περιουσιακών στοιχείων, η οποία βασίζεται στα πρότυπα Ιστού (Web standards). Η αρχιτεκτονική διάδοσης παρέχει έναν μηχανισμό που θα προσθέτει χρήσιμες προβολές (views) σε συγκεκριμένα ψηφιακά αντικείμενα βιβλιοθήκης που έχουν δημιουργηθεί με χειρονακτικό τρόπο.

- Το CiteSeer ([citeseer.ist.psu.edu/](http://citeseer.ist.psu.edu/)) δεν έχει εξελιχθεί ακόμα σε σημασιολογική υπηρεσία, και συνεπώς τα τρέχοντα μεταδεδομένα του δεν χρησιμοποιούν μηχανισμούς του Σημασιολογικού

Ιστού. Το CiteSeer είναι μια μεγάλη πηγή μεταδεδομένων, τα οποία θα αποτελέσουν το βασικό υλικό για το Σημαιολογικό Ιστό. Η πολυπλοκότητα του CiteSeer θα συμβάλει στη μετάβασή του σε μια αποτελεσματική σηματολογική υπηρεσία.

- Το ScholOnto (Buckingham et al., 2000). Στα πλαίσια του έργου αυτού, ένας κεντρικός υπολογιστής ψηφιακών βιβλιοθηκών (που βασίζεται σε μια οντολογία βιβλιοθήκης) υποστηρίζει την ερμηνεία ακαδημαϊκών εννοιών, ομιλιών, και επιστημονικών εγγράφων. Δίνει τη δυνατότητα στους ερευνητές να περιγράφουν και να συζητούν μέσω ενός σηματολογικού δικτύου για τη συμβολή ενός εγγράφου, μιας επιστημονικής εργασίας και τη σχέση του εγγράφου με τη Λογοτεχνία. Το σύστημα ScholOnto υποστηρίζει βασικά ερωτήματα, όπως: α) Ποια είναι η πηγή προέλευσης των ιδεών; β) Ποια είναι η επίδραση των ιδεών; γ) Ποιες προοπτικές, ασυνέπειες αλλά και συγκλίσεις μπορούν να διαπιστωθούν σε αυτές;

Η εφαρμογή του ScholOnto επιτρέπει:

1. Να προσδιορίζονται οι σχέσεις μεταξύ των εννοιών και άλλων ισχυρισμών που διατυπώνονται σε διάφορα έγγραφα.

2. Να εκφέρουν οι ερευνητές την άποψή τους για τη σημασία αυτών των ιδεών και εννοιών.

- Το έργο Bibster (Haase et al., 2004) στοχεύει στη δημιουργία ενός κατανεμημένου περιβάλλοντος γνώσης, που θα επιτρέπει στους ερευνητές να μοιράζονται εύκολα τα βιβλιογραφικά δεδομένα με τη χρήση τεχνολογιών P2P (Peer-to-Peer), να εξοικονομούν χρόνο στην προσπάθεια αναζήτησης αυτών των δεδομένων και να διαμορφώνουν μια σαφή άποψη γι' αυτά. Η αναζήτηση των βιβλιογραφικών δεδομένων γίνεται τόσο με τη χρήση λέξεων-κλειδιών (keywords), όσο και με πιο προηγμένες τεχνικές σηματολογικής αναζήτησης. Οι χρήστες μπορούν να ενσωματώσουν τα αποτελέσματα ενός ερωτήματος σε μια τοπική βάση γνώσεων για μελλοντική χρήση. Το Bibster εκμεταλλεύεται τις οντολογίες στην εισαγωγή δεδομένων, στη διατύπωση και επεξεργασία των ερωτημάτων, καθώς και στην παρουσίαση των απαντήσεων. Οι βιβλιογραφικές καταχωρήσεις στο Bibster είναι δομημένες σύμφωνα με δύο διαφορετικές οντολογίες. Η πρώτη οντολογία (SWRC) περιγράφει τις διάφορες γενικές πτυχές των βιβλιογραφικών μεταδεδομένων (ανά ερευνητική περιοχή), ενώ η δεύτερη οντολογία (ACM Topic Hierarchy) περιγράφει συγκεκριμένες κατηγορίες

της Λογοτεχνίας από την μεριά της Πληροφορικής. Αυτή η οντολογική δομή αξιοποιείται για να διαμορφώνουν οι χρήστες κατάλληλα τα ερωτήματα τους. Το Bibster είναι κτισμένο σε πλατφόρμα JXTA (Gong, 2001) και διανέμεται στη δικτυακή τοποθεσία: <http://bibster.semanticweb.org>. Βρίσκεται ακόμη στο αρχικό του στάδιο, ενώ στα επόμενα στάδια προβλέπεται η βελτιστοποίησή του και η επέκτασή του για εξατομικευμένες σηματολογικές δομές.

- UK Berkeley Digital Library Project (<http://elib.cs.berkeley.edu/>). Η βιβλιοθήκη του Berkeley υλοποιεί αυτό το έργο, στο οποίο μελετάται η προσθήκη σηματολογικών εννοιών σε αρχεία εικόνων, προκειμένου να αξιοποιηθεί ο Σηματολογικός Ιστός στις ψηφιακές βιβλιοθήκες. Στα πλαίσια του έργου αυτού, έχει γίνει αξιολογή έρευνα για την παρουσίαση των εικόνων, ενώ παρέχεται πρόσβαση σε μεγάλες συλλογές εικόνων μέσα από την πλοήγηση και την αναζήτηση. Μεγάλες συλλογές από εικόνες χρησιμοποιούνται ως δεδομένα για τη μελέτη διαφόρων προβλημάτων απεικόνισης (όπως λ.χ. αυτό της αναγνώρισης αντικειμένων).

#### 4. Τα πλεονεκτήματα εφαρμογής του Σηματολογικού Ιστού στις Ακαδημαϊκές Βιβλιοθήκες

Ο Σηματολογικός Ιστός αναπτύσσει και εξελίσσει τις ακαδημαϊκές βιβλιοθήκες επηρεάζοντας όχι μόνο τον τρόπο με τον οποίο λειτουργεί μια βιβλιοθήκη, αλλά και τον τρόπο με τον οποίο οι χρήστες τους αναζητούν και εντοπίζουν τις επιθυμητές πληροφορίες (Campbell & Fast, 2004). Μετατρέπει τις ακαδημαϊκές βιβλιοθήκες από αρχειοφύλακες της γνώσης σε συστήματα σήμανσης του υλικού των βιβλιοθηκών (Sure & Studer, 2005). Τα σχήματα, οι οντολογίες, και οι πράκτορες λογισμικού παρέχουν τη δυνατότητα για έξυπνη επεξεργασία και διανομή των βιβλιογραφικών πληροφοριών. Στο Σηματολογικό Ιστό, όλες οι βιβλιογραφικές πληροφορίες θα είναι αποτελεσματικά ανακτήσιμες και θα συγκεντρώνονται για την επαγωγή λογικών συμπερασμάτων. Θα παρέχονται εξατομικευμένες υπηρεσίες βιβλιογραφικής πληροφόρησης, ενώ παράλληλα θα επιτυγχάνεται η διαλειτουργικότητα, καθώς και η μέγιστη δυνατή ανταλλαγή και επαναχρησιμοποίηση του υλικού των βιβλιοθηκών (Ding, 2004). Παρόμοιες τάξεις ψηφιακών αντικειμένων και υπηρεσίες βιβλιοθηκονομίας θα διανέμονται ανάμεσα σε ετερογενείς ακαδημαϊκές βιβλιοθήκες. Επιπλέον, οι σηματολογικές

μηχανές αναζήτησης θα επιστρέφουν ακριβή αποτελέσματα αναζήτησης, καθώς θα κατανοούν τα ερωτήματα σημασιολογικά και όχι απλά ως λέξεις-κλειδιά. Οι χρήστες θα εκφράζουν νέες έννοιες (concepts) και μία κοινή γλώσσα λογικής θα επιτρέπει σε αυτές τις έννοιες να συνδέονται σταδιακά σε ένα καθολικό Δίκτυο (Campbell & Fast, 2004).

#### 4.1 Σημασιολογική μοντελοποίηση του υλικού

Σημασιολογική σήμανση είναι η διεργασία εισαγωγής ετικετών (tags) στα έγγραφα Web για να αποδώσουμε σημασιολογίες στο κείμενο τους. Η αποτελεσματική λειτουργία του Σημασιολογικού Ιστού απαιτεί υπολογιστές, οι οποίοι θα δημιουργούν ετικέτες που θα περιγράφουν το περιεχόμενο των βιβλιοθηκών. Οι ετικέτες αυτές θα είναι οργανωμένες στην κατάλληλη μορφή που απαιτείται από τις επίσημες οντολογίες βιβλιοθήκης. Οι χρήστες και οι πράκτορες λογισμικού θα χρησιμοποιούν τις ετικέτες αυτές για να ανακτούν και να επεξεργάζονται τις πληροφορίες των βιβλιοθηκών που βρίσκονται στο Διαδίκτυο. Η σημασιολογική μοντελοποίηση της πληροφορίας των βιβλιοθηκών καθιστά τα πληροφοριακά συστήματα βιβλιοθηκονομίας πιο έξυπνα και ικανά να παρέχουν εξατομικευμένες υπηρεσίες (Brooks, 2002). Ένα έξυπνο πληροφοριακό σύστημα βιβλιοθηκονομίας περιέχει οντολογίες και αποθήκες για το υλικό των βιβλιοθηκών. Η επιτυχία του Σημασιολογικού Ιστού θα εξαρτηθεί από τη διαθεσιμότητα κατάλληλων οντολογιών βιβλιοθήκης, καθώς και την αύξηση των ιστοσελίδων που θα επισημειώνονται με μεταδεδομένα συμμορφούμενα με αυτές τις οντολογίες.

#### 4.2 Υποβολή ερωτήσεων σε σημασιολογική βάση για την εύρεση πληροφοριακού υλικού

Η σημασιολογική αναζήτηση βελτιώνει τις τρέχουσες μηχανές αναζήτησης με σημασιολογίες. Παρέχεται από εργαλεία του Σημασιολογικού Ιστού, όπως το σύστημα Ontobroker που παρέχει μία υπηρεσία αναζήτησης/απάντησης βασιζόμενη σε οντολογία (Fensel et al., 1999). Σημασιολογικοί φυλλομετρητές (λ.χ. ο Magpie) θα χρησιμοποιούν τις οντολογίες βιβλιοθήκης για να προσδιορίζουν σημαντικές έννοιες μέσα σε ένα έγγραφο και θα παρέχουν πρόσβαση στο σχετικό υλικό (Dzbor et al., 2003)

#### 4.3 Σημασιολογική διαλειτουργικότητα

Στα πληροφοριακά συστήματα βιβλιοθηκονομίας, η τεχνική διαλειτουργικότητα θα πρέπει να συμπληρώ-

νεται με τη σημασιολογική διαλειτουργικότητα. Για αυτό απαιτείται η δημιουργία κατάλληλης πληροφορίας σχετική με τα στοιχεία των βιβλιογραφικών δεδομένων, έτσι ώστε η συγκεκριμένη πληροφορία να επεξεργάζεται από πράκτορες λογισμικού και από άλλες εφαρμογές του Σημασιολογικού Ιστού. Επίσης, απαιτείται η κατάλληλη διασύνδεση των διαφορετικών στοιχείων των βιβλιογραφικών δεδομένων για να γίνει η σημασιολογική επέκταση, και το ραφινάρισμα των λεξικών.

#### 4.4 Σημασιολογική διαχείριση γνώσης

Η εξόρυξη από σημασιολογικά δεδομένα (semantic data mining) επιτρέπει: 1) τον ακριβή προσδιορισμό μιας ομάδας χρηστών/αναγνωστών με συγκεκριμένα χαρακτηριστικά χρήσης πληροφοριακού υλικού, 2) παροχή εξατομικευμένων υπηρεσιών βιβλιοθηκονομίας, και 3) μέτρηση της χρήσης του πληροφοριακού υλικού. Όλα αυτά μπορούν να συμβάλλουν στην υλοποίηση αποτελεσματικών στρατηγικών μάρκετινγκ στο χώρο των ακαδημαϊκών βιβλιοθηκών. Η εξόρυξη από σημασιολογικά δεδομένα μπορεί να εφαρμοστεί για να καταγράψει και να αναλύσει τις προτιμήσεις των χρηστών/αναγνωστών ως προς τα ειδικά στοιχεία ενός δομοστοιχείου βιβλιογραφικής πληροφορίας (library information module). Τα μελλοντικά πληροφοριακά συστήματα βιβλιοθηκονομίας θα καταγράφουν τις αναγνωστικές προτιμήσεις των χρηστών τους και θα παράγουν προφίλ χρήστη (user profile). Το προφίλ χρήστη εκφράζει τα χαρακτηριστικά ενός χρήστη και αποτελείται από το στατικό μέρος (π.χ. δημογραφική πληροφορία: όνομα, φύλο, ηλικία, χώρα προέλευσης...) και το δυναμικό μέρος (αναγνωστικά ενδιαφέροντα, φίλτρα, ίχνη). Τα φίλτρα περιγράφουν τον μηχανισμό που εκφράζει τα αναγνωστικά ενδιαφέροντα του χρήστη (λ.χ. Κλασική Λογοτεχνία), ενώ τα ίχνη περιγράφουν τις αλληλεπιδράσεις των χρηστών με το πληροφοριακό σύστημα, καθώς και τον μηχανισμό καταγραφής αυτών των ενεργειών.

### 5. Συμπεράσματα

Ο Σημασιολογικός Ιστός θα επιφέρει αλλαγές στον τρόπο που οι βιβλιοθήκες καταλογογραφούν το υλικό τους (Gradmann, 2005). Η αλλαγή θα επέλθει στον τρόπο με τον οποίο θα αναπτύσσονται οι συλλογές των βιβλιοθηκών. Με τη χρήση σημασιολογικών εργαλείων, οι καταλογογράφοι θα κάνουν έξυπνες επιλογές, θα εντοπίζουν, και θα εξάγουν τις κατάλληλες

πληροφορίες από άλλες πηγές στις βιβλιογραφικές τους εγγραφές. Ο στόχος του καταλογογράφου θα μετατοπιστεί από την μεταγραφή των βιβλιογραφικών δεδομένων στην ανεύρεση βιβλιογραφικών δεδομένων και τη διερεύνηση εκείνου του στοιχείου που χαρακτηρίζεται από αυθεντικότητα και χρησιμότητα. Έτσι, ο καταλογογράφος θα ενεργεί ως διαμεσολαβητής της πληροφορίας (info broker) χρησιμοποιώντας την πείρα και τις γνώσεις του. Συμπερασματικά, οι ψηφιακές βιβλιοθήκες θα πρέπει να αξιοποιήσουν τα χαρακτηριστικά του Σημασιολογικού Ιστού: τη διαλειτουργικότητα, τις οντολογίες και τον σημασιολογικό συμβολισμό της πληροφορίας. Ο Σημασιολογικός Ιστός μπορεί να προσφέρει περισσότερη ευελιξία και νέες προοπτικές στις ψηφιακές βιβλιοθήκες μέσω της χρήσης νέων τεχνολογιών όπως είναι τα εργαλεία συμβολισμού και οι πράκτορες λογισμικού.

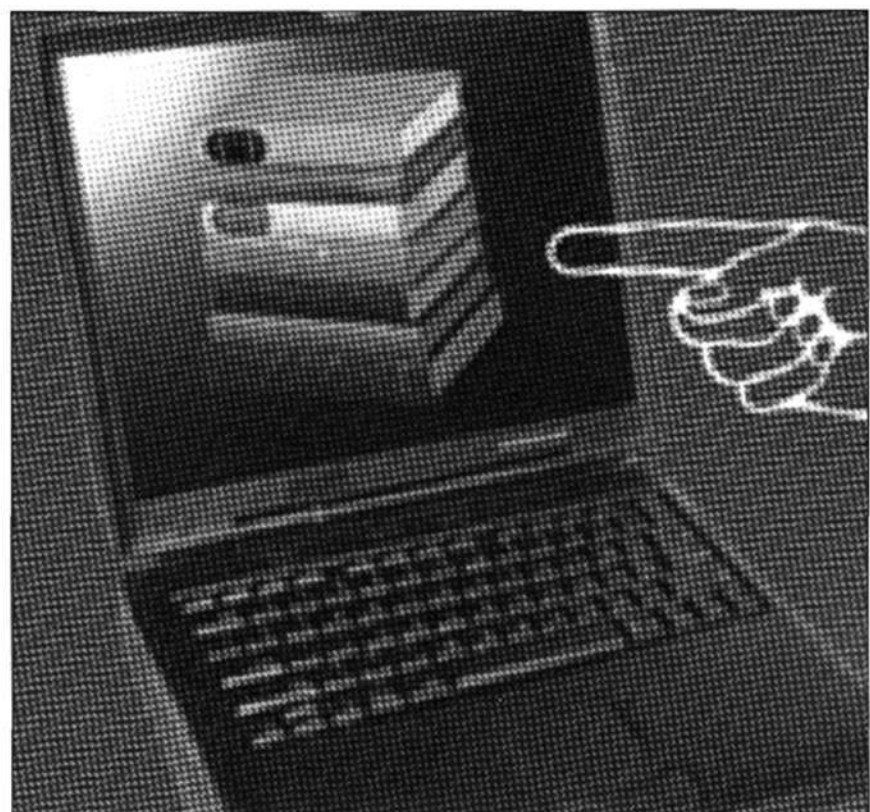
#### Αναφορές

Berners-Lee, T., Hendler, J. and Lassila, O. (2001) The Semantic Web, *Scientific American*, 279(5), 34-43.

Brooks, T. (2002) The Semantic Web, universalist ambition and some lessons from librarianship, *Information Research*, 7(4).

Buckingham Shum, S., Motta, E., Domingue, J. (2000) ScholOnto: An Ontology-Based Digital Library Server for Research Documents and Discourse, *Int. Journal Digit. Libr.*, 3, 237-248.

Butler, M., Gilbert, J., Scaborne, A. and Smathers, K. (2004) Data conversion, extraction and record linkage using XML and RDF tools in Project



SIMILE, Tech Report:

<http://www.hpl.hp.com/techreports/2004/HPL-2004-147.html>

Campbell, D. and Fast, V. (2004) Academic Libraries and the Semantic Web: What the Future May Hold for Research-Supporting Library Catalogues, *The Journal of Academic Librarianship*, 30(5).

Ding, H. (2004) Challenges in constructing Semantic Interoperable Digital Library System. <http://www.idi.ntnu.no/yunl/seminar/presentations/2003-04-25/research-talk.pdf>.

Doan, A. H., Madhavan, J., Domingos, P. and Halevy, A. (2002) Learning to Map Ontologies on the Semantic Web. In: *Proceedings of the Eleventh International World Wide Web Conference 2002*, 7-11 May, Honolulu, USA.

Dublin Core Metadata Initiative homepage: <http://dublincore.org>

Dzbor, M., Domingue, J.B., and Motta, E. (2003) Magpie - Towards a Semantic Web Browser, In: *Proceedings of the 2nd International Conference (ISWC 2003)*, 690-705. Florida, USA, October. <http://kmi.open.ac.uk/people/domingue/papers/magpie-iswc-03.pdf>

Fensel, D., Angele, J. Erdmann, M., Schnurr, H., Staab, S. Studer, R. and Witt, A. (1999) On2broker: Semantic-based access to information sources at the WWW, In: *Proceedings of WebNet*, 366-371.

[http://www.aifb.uni-karlsruhe.de/WBS/Publ/1999/webnet\\_dflectal\\_1999.pdf](http://www.aifb.uni-karlsruhe.de/WBS/Publ/1999/webnet_dflectal_1999.pdf)

Gong, L. (2001) Project JXTA: A technology

overview, Tech. Report, SUN Microsystems, <http://www.jxta.org/project/www/docs/TechOverview.pdf>, April 2001.

Gradmann, S. (2005) rdfs:frbr - Towards an implementation model for library catalogs using semantic web technology. In: Cataloguing and Classification Quarterly, FRBR Theme Issue, Spring 2005.

Haase, P. et al. (2004) Bibster □ A Semantics-Based Bibliographic Peer-to-Peer System, Web Semantics: Science, Services, and Agents on the World Wide Web, 2(1), 99-103.

Hendler, J. (2001) Agents and the Semantic Web, IEEE Intelligent Systems, 16(2), 30-37.

Liebetrau, P. (2004) XML for Digital Libraries, IFLA- Africa Section Workshop.

MARC 21 XML Schema official webpage: <http://www.loc.gov/standards/marcxml/>

McGuinness, D. and Van Harmelen, F. (2003) OWL Overview: <http://www.w3.org/TR/owl-features/>

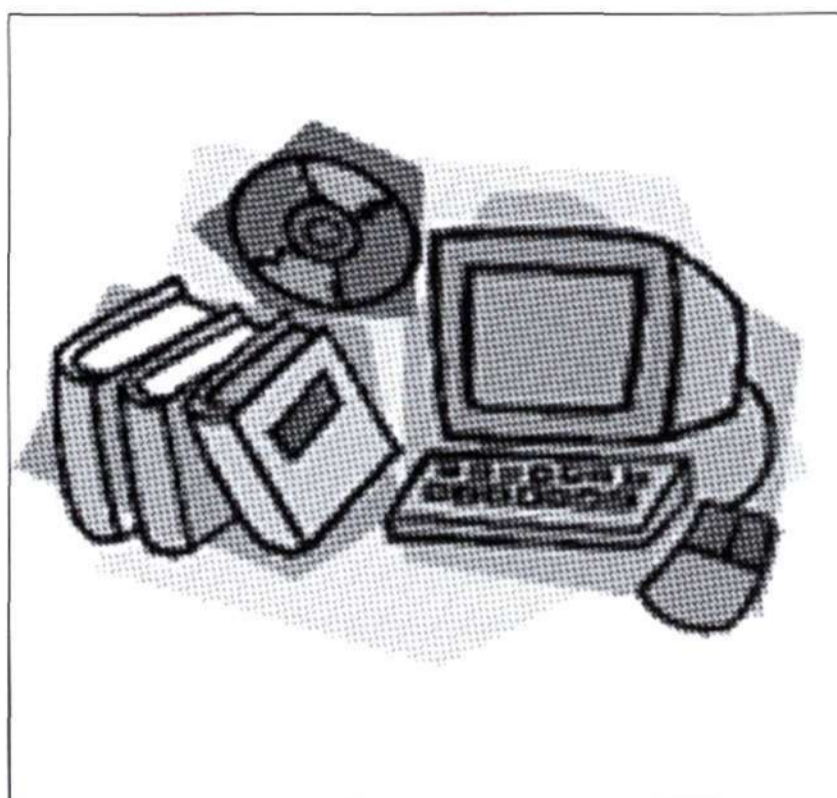
Miller, E., Miller, P. and Brickley D. (1999) Guidance on expressing the Dublin Core within the Resource Description Framework (RDF). Dublin Core Metadata Initiative Draft Proposal. Available at:

<http://www.ukoln.ac.uk/metadata/resources/dc/data-model/WD-dc-rdf/>

Mizoguchi, R. (2004) Ontology Engineering Environments, In S. Staab, R. Studer (Eds.) Handbook on Ontologies, Berlin: Springer, 275-298.

ONIX for Books. homepage: <http://www.editeur.org/onix.html>

Ouzzani, M. and Bouguettaya, A. (2004) Efficient



Access to Web Services, IEEE Internet Computing, March-April 2004, 8(2), 34-44.

Paepcke, A., Chang, C., Garcia-Molina, H. and Winograd, T. (1998) Interoperability for digital libraries worldwide, Communications of the ACM, 41(4): April 1998.

Protege-2000 tool (2000) User's Guide.

<http://protege.stanford.edu/publications/UserGuide.pdf>

Publishing Requirements for Industry Standard Metadata (PRISM) homepage:

<http://www.prismstandard.org/about/>

Simple Digital Library Interoperability Protocol (SDLIP) (1999) [http://www-diglib.stanford.edu/\\_testbed/doc2/SDLIP/](http://www-diglib.stanford.edu/_testbed/doc2/SDLIP/)

Simple Knowledge Organisation System (SKOS) homepage: <http://www.w3.org/2004/02/skos/>

Sure, Y. and Studer, R. (2005) Semantic Web technologies for Digital Libraries, Library Management, 26(4/5), 190-195(6).

SWAD-Europe Thesaurus Activity homepage: <http://www.w3.org/2001/sw/Europe/reports/thes/>

### **Δημήτριος Κανελλόπουλος**

Δρ. Πληροφορικής, Επιστημονικός Συνεργάτης της Α.Σ.ΠΑΙ.Τ.Ε., [dkanellop@teipat.gr](mailto:dkanellop@teipat.gr)

### **Ανδριανή Ζαφειροπούλου**

Αρχιονόμος - Βιβλιοθηκονόμος στο Ελληνικό Ανοικτό Πανεπιστήμιο (Ε.Α.Π.)

