



ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ Η/Υ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ ΚΑΙ ΔΙΚΤΥΩΝ

ΗΥ628—Βάσεις δεδομένων και Παγκόσμιος Ιστός

Εργασία 2:

Σημασιολογικός Ιστός – Semantic Web

Ganatsiou Aikaterini

Physicist (A.U.TH)
M. Sc. in Physics of Materials (A.U.TH)
M. Sc. in Computer Science Telecommunications and Networks (U.TH)

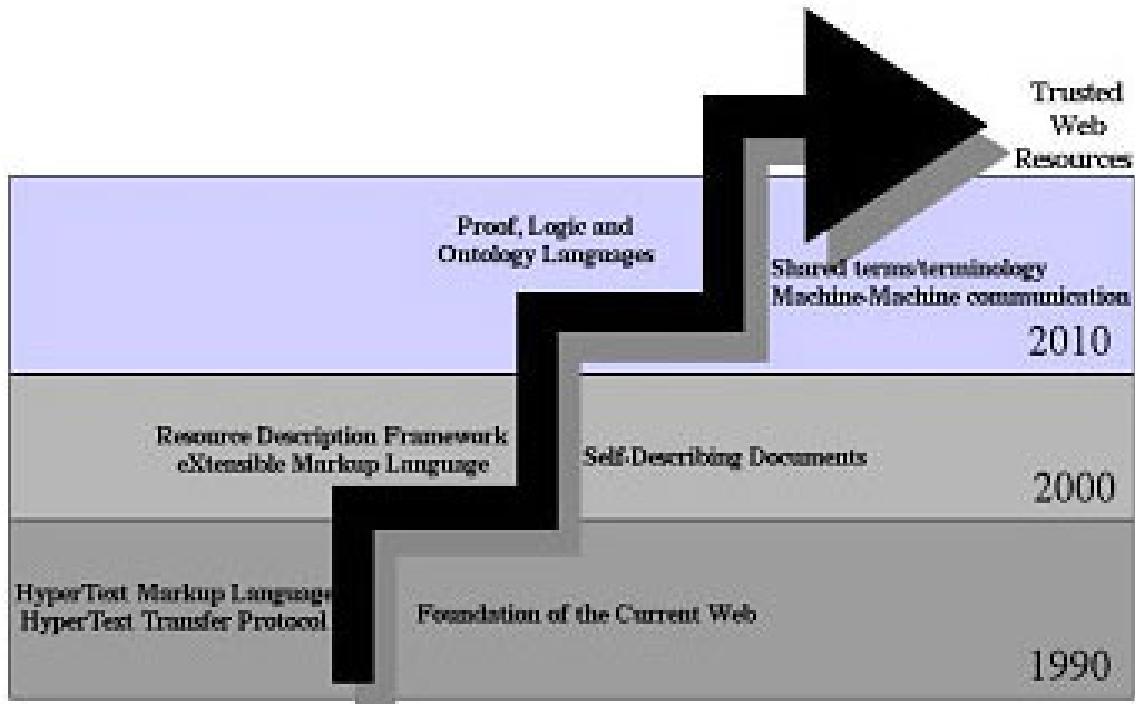
Orfanakis Stilianos

Physicist (A.U.TH)
M. Sc. in Electronics - Radioelectrology (A.U.TH)
Certified in Bank Management (U.O.A)
M. Sc. in Computer Science Telecommunications and Networks (U.TH)

Βόλος, 21/05/04

Εισαγωγή

Τα τελευταία δύο χρόνια γίνεται πολύς λόγος και για το Semantic Web, ενός διαδικτύου δηλαδή στο οποίο η πληροφορία είναι καλύτερα ορισμένη επιτρέποντας με αυτόν τον τρόπο την ολοκλήρωση, αυτοματοποίηση και επαναχρησιμοποίηση των δεδομένων.



Ο Ιστός (Web) του μέλλοντος προβλέπεται να αποτελεί μια παγκόσμια βάση δεδομένων και γνώσης με πληροφορίες οι οποίες θα είναι "κατανοητές" από μηχανές (machine-understandable information). Οι κύριες τεχνολογίες για την υλοποίηση του Σημασιολογικού Ιστού είναι η σημασιολογικό εμπλουτισμός και η χρήση των οντολογιών.

Η λέξη "Σημασιολογία" έχει ρίζα τις Ελληνικές λέξεις "σημάδι", "σημαίνω" και "σημαντικός" και σήμερα αναφέρεται στο νόημα συχνά σε επίπεδο γλώσσας. Μπορούμε να πούμε ότι ο Σημασιολογικός Ιστός αποτελεί το μεγαλύτερο σε παγκόσμιο επίπεδο έργο έξυπνης ενσωμάτωσης συστημάτων ώστε να συνεργάζονται δια-λειτουργικά.

Συντάκτες:

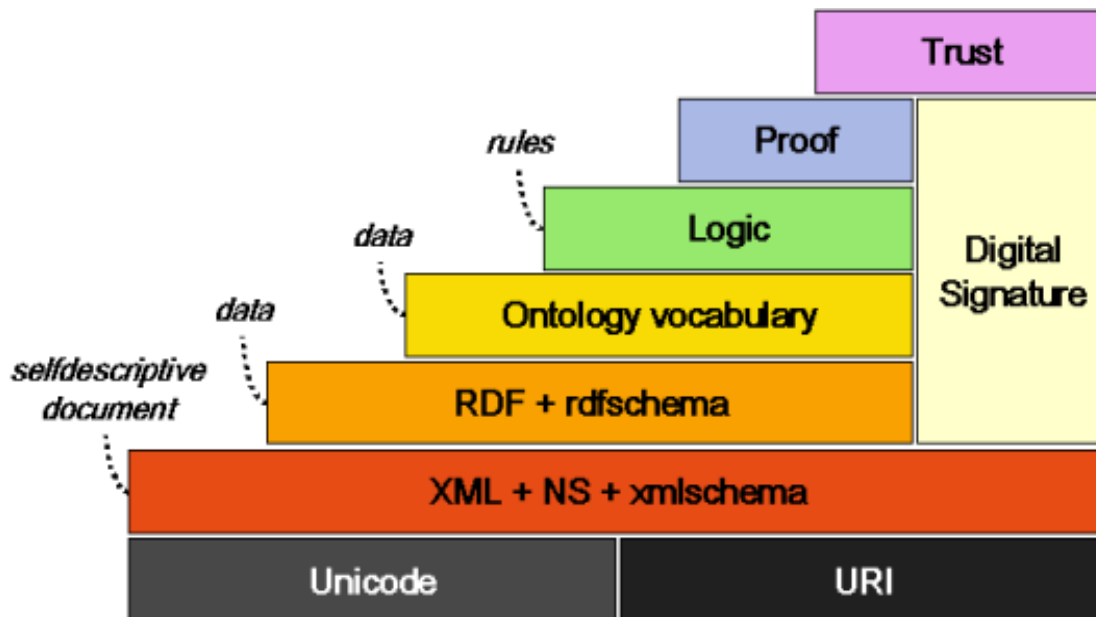
Γκανάτσου Αικατερίνη

Ορφανάκης Στυλιανός

Ο Tim Berners-Lee, που επινόησε τον Παγκόσμιο Ιστό το 1989, είχε το όραμα, που τώρα συμμερίζονται πολλοί - ενός ιστού δεδομένων που μπορούν να επεξεργαστούν από μηχανές. "Ο Σημασιολογικός Ιστός είναι μια επέκταση του σημερινού ιστού όπου η πληροφορία έχει καλά καθορισμένο νόημα, καθιστώντας τη συνεργασία μεταξύ ανθρώπων και υπολογιστών πιο αποτελεσματική", Tim Berners-Lee, James Hendler, Ora Lassila, The Semantic Web, Scientific American, Μάιος 2001.

Το κέντρο βάρους του περιεχομένου του Ιστού μετατοπίζεται συνεχώς από τον άνθρωπο προς προς τα δεδομένα. Για να φτάσει ο Ιστός το μέγιστο των δυνατοτήτων του, πρέπει να εξελιχθεί σε ένα Σημασιολογικό Ιστό, ο οποίος παρέχει μια διεθνώς προσβάσιμη πλατφόρμα που επιτρέπει σε αυτοματοποιημένα εργαλεία αλλά και σε ανθρώπους να μοιράζονται και να επεξεργάζονται δεδομένα.

Ο Σημασιολογικός Ιστός αποτελεί πρωτοβουλία της Κοινοπραξίας του Παγκοσμίου Ιστού (W3C) και η σχετική Δραστηριότητα (W3C Semantic Web Activity) έχει δημιουργηθεί για να εξυπηρετήσει έναν ηγετικό ρόλο, τόσο στο σχεδιασμό προδιαγραφών, όσο και στην ανοικτή ανάπτυξη της τεχνολογίας μέσω της συνεργασίας.



Συντάκτες:

Γκανάτσου Αικατερίνη

Ορφανάκης Στυλιανός

1. Ποιες οι τάσεις στην διαχείριση της ψηφιακής πληροφορίας

Σύμφωνα με πρόσφατες μελέτες, η ανθρωπότητα έχει παράγει, στο διάστημα 1999-2002, τόσες νέες πληροφορίες όσες παρήγαγε όλα τα προηγούμενα χρόνια της ιστορίας της. Σε αυτό το διάστημα των τριών τελευταίων χρόνων παρήχθησαν 12 exabytes πληροφορίας υπό την μορφή έντυπου, ηχητικού και οπτικού υλικού. Η αυξανόμενη παραγωγή και η συνεχής βελτίωση των μεθόδων ψηφιοποίησης συμβάλλουν στην παραγωγή ενός ωκεανού ψηφιακών δεδομένων που πρόκειται να δημιουργήσει μία σωρεία προβλημάτων.

Ένα από τα σημαντικότερα προβλήματα που αναδύονται είναι η αποτελεσματική διαχείριση όλου αυτού του όγκου των ψηφιακών δεδομένων, αφού η ικανότητα της παραγωγής, αποθήκευσης και μετάδοσης έχει ξεπεράσει κατά πολύ τις δυνατότητες αναζήτησης, πρόσβασης και παρουσίασης. Η ερευνητική περιοχή της Έξυπνης Πρόσβασης στην Πληροφορία επικεντρώνεται στην επίλυση προβλημάτων διαχείρισης, ενοποιημένης και προσαρμοζόμενης πρόσβασης σε πολύ μεγάλες συλλογές ψηφιακών δεδομένων. Το φάσμα που καλύπτει η κατεύθυνση αφορά τις παρακάτω θεματικές υπο-περιοχές:

- ☒ Διαχείριση Ψηφιακής Πληροφορίας με βάση το περιεχόμενο. Η διαχείριση πληροφορίας με βάση το περιεχόμενο στηρίζεται στη ανάλυση του περιεχομένου της πληροφορίας για την επίτευξη όσο το δυνατόν καλύτερων αποτελεσμάτων τόσο στον τομέα της ταχύτητας όσο και στον τομέα της ποιότητας. Στην συγκεκριμένη θεματική περιοχή η έρευνα επικεντρώνεται σε τεχνικές οι οποίες θα εκμεταλλεύονται το περιεχόμενο όλων των τύπων πληροφορίας (κείμενο, εικόνα και κινούμενη εικόνα) για την επίτευξη αποτελεσματικής, ενοποιημένης πρόσβασης σε ψηφιακές συλλογές.

☒ Γνωστικές μέθοδοι για την διαχείριση Ψηφιακής Πληροφορίας. Η συγκεκριμένη θεματική ενότητα επικεντρώνεται σε μεθόδους διαχείρισης πληροφορίας μέσω της εκμετάλλευση των μετα-δεδομένων και την ανάπτυξη ειδικών δομών, των οντολογιών, που περιγράφουν αποδοτικά αντικείμενα, ιδέες, διαδικασίες καθώς και των σχέσεων που τις διέπουν. Στόχος είναι η έρευνα της αποτελεσματικότητας οντολογικών προσεγγίσεων για την ανάπτυξη συστημάτων διαχείρισης ψηφιακής πληροφορίας μέσω της κωδικοποίησης και εκμετάλλευσης της σημασιολογίας της πληροφορίας.

☒ Ολοκλήρωση μεθόδων διαχείρισης Ψηφιακής Πληροφορίας σε πληροφοριακά συστήματα. Η ολοκλήρωση μεθόδων διαχείρισης ψηφιακής πληροφορίας σε πληροφοριακά συστήματα και μεγάλες βάσεις δεδομένων παίρνοντας υπόψη τις απαιτήσεις για on-line πρόσβαση, αυθεντικότητα, ασφάλεια, διάχυση των σύγχρονων εφαρμογών. Επίσης έρευνα πάνω στην επίδραση που έχουν τα τρέχοντα και μελλοντικά επιχειρηματικά μοντέλα πάνω στην σχεδίαση, ανάπτυξη, εφαρμογή και συντήρηση συστημάτων διαχείρισης ψηφιακής πληροφορίας. Έρευνα για την ολοκλήρωση των μεθόδων διαχείρισης με νέα τεχνολογικά μοντέλα και εφαρμογές όπως το Grid, Augmented Reality και Semantic Web.

2. Το πρότυπο RDF – RDF/S

2.1 Εισαγωγή - Χαρακτηριστικά του

Στην ενότητα αυτή θα επιχειρήσουμε να παρουσιάσουμε μια αναλυτική περιγραφή του *RDF* (*Resource Description Framework*) - Πλαισίου Περιγραφής Πόρων. Το *RDF* έχει προταθεί από το *World Wide Web Consortium (W3C)* όπως θα δούμε και το *RDF Schema Specification (RDF/S)*, η οποία είναι μια γλώσσα περιγραφής σχημάτων περιγραφής πληροφοριακών πόρων, που ορίζουν τη σημασιολογία των περιεχομένων των μεταδεδομένων για διάφορες εφαρμογές, που εξ ορισμού περιέχουν από κοινού κλάσεις και ιδιότητες, και σχετίζεται με το *RDF* όπως θα περιγράψουμε με πάσα λεπτομέρεια στην συνέχεια.

Η χρήση του *RDF* σαν επιμέρους βέλτιστο μηχανισμό περιγραφής των πληροφοριακών πόρων, έγκειται επιμέρους στα εξής χαρακτηριστικά γνωρίσματα της:

1. *ανεξαρτησίας* του πεδίου εφαρμογής σε κάθε περίπτωση και η παροχή της δυνατότητας σε οποιονδήποτε να ορίζει κλάσεις ή ιδιότητες και να της επαναχρησιμοποιεί με βάση την επιβεβλημένη σημασιολογία.
2. *ευκολίας ανταλλαγής και μεταφοράς* των *RDF* περιγραφών τόσο από ανθρώπους αλλά κι από μηχανές διαχείρισης.
3. στην *δυνατότητα* της αναπαράστασης της γνώσης μέσω της *περιγραφής ιδιοτήτων/κλάσεων*, σε ένα πεδίο εφαρμογής μέσω μηχανισμών που υποστηρίζουν κοινές συμβάσεις για την σημασιολογία, τη σύνταξη και την δομή των μεταδεδομένων.
4. Μάλιστα μπορεί να αξιοποιηθεί και για μια σημασιολογικά πλούσια περιγραφή πληροφοριακών πόρων με την δημιουργία μεταδεδομένων μεταξύ υπαρχόντων και επαναχρησιμοποιούμενων ή κατ' επέκταση σχημάτων.

Εν γένει ο πυρήνας του *RDF* είναι ένα μοντέλο που αποτελείται από κατευθυνόμενους γράφους με ετικέτες (*labels*), που μπορούν να χρησιμοποιηθούν για την περιγραφή των σχέσεων μεταξύ τόσο των πόρων, όσο και των γνωρισμάτων των πόρων. Τα λεξιλόγια αυτά επιμέρους αποτελούνται από ιεραρχίες κλάσεων που στηρίζονται, καθώς κι από τύπους ιδιοτήτων που τις συσχετίζουν, αλλά και μπορούν να επεκταθούν μέσω των σχέσεων εξειδίκευσης (*isA*) τόσο για κλάσεις όσο και ιδιότητες.

2.2 RDF μοντέλο δεδομένων

Το μοντέλο δεδομένων του *RDF* είναι ένα πολύ γενικό και απλό μοντέλο για την περιγραφή των πληροφοριακών πόρων. Στους πόρους αποδίδονται ιδιότητες, οι οποίες αντιστοιχούν είτε σε *σχέσεις* μεταξύ πόρων είτε σε *γνωρίσματα* του πόρου που αποδίδονται. Συγκεκριμένα, το μοντέλο δεδομένων του *RDF* αποτελείται από τρεις βασικούς τύπους δεδομένων: *Πόρους* (*Resources*), *Ιδιότητες* (*Properties*) και *Δηλώσεις* (*Statements*) τους οποίους και θα παρουσιάσουμε εν συνεχεία :

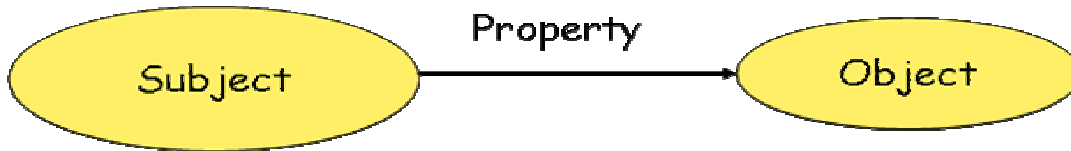
- **Πόρος**, ονομάζεται οτιδήποτε που μπορεί να περιγραφεί με *RDF* εκφράσεις. Αναφορικά, μπορεί να είναι οτιδήποτε από τα παρακάτω:
 - ο μια σελίδα του διαδικτύου,
 - ο έγγραφα ή σελίδες με την μορφή *HTML*, *XML*
 - ο ένα τμήμα τέτοιων εγγράφων ή αποσπασματικά κάποιο στοιχείο μιας σελίδας μας

ο με τον χαρακτηρισμό πόροι μπορούμε κι ονομάζουμε και τα αντικείμενα τα οποία δεν είναι προσπελάσιμα μέσω του δια-δικτύου, όπως είναι για παράδειγμα ένα άτομο.

Εν γένει πόρος μπορεί να θεωρηθεί οτιδήποτε στο οποίο αποδίδεται ένα μοναδικό *URI (Uniform Resource Identifier)- Ομοιόμορφο Αναγνωριστικό Πόρων* στο οποίο μπορεί να προστεθεί και μια άγκυρα (anchor ids) - τα καλούμενα *qualified URIs*, τα οποία είναι *URIs*, που μπορεί να περιέχουν ένα αναγνωριστικό τμήματος (fragment identifier), και το οποίο διαχωρίζεται από το υπόλοιπο URI με το χαρακτήρα '#'.
 Εάν θεωρήσουμε για παράδειγμα το URI [“http://dmoz.org”](http://dmoz.org) αποδίδεται στην σελίδα με την αντίστοιχη διεύθυνση της τιμής του *URI*, ενώ το URI [“http://www.ics.forth.gr/RDF/Dublin_Core.rdf#Title”](http://www.ics.forth.gr/RDF/Dublin_Core.rdf#Title) αντιστοιχεί την ιδιότητα *Title* που ορίζεται στο αρχείο [“http://www.../Dublin_Core.rdf”](http://www.../Dublin_Core.rdf)

- **Ιδιότητα**, είναι ένα γνώρισμα που χρησιμοποιείται για να περιγράψει είτε ένα πόρο είτε μια σχέση μεταξύ πόρων. Μια ιδιότητα έχει καθορισμένη σημασιολογία και πιθανόν έχει κάποιο πεδίο ορισμού, πεδίο τιμών καθώς και συσχετίσεις με άλλες ιδιότητες. Κι οι ιδιότητες με την σειρά τους είναι επιμέρους πόροι με απόλυτα διακριτά *URIs* και άρα μπορούν να περιγράφουν με αντίστοιχες των κλάσεων *RDF* δηλώσεις. Αποτελούν αυτόνομες οντότητες και δεν χαρακτηρίζονται από την κλάση στην οποία ορίζονται και προσδιορίζουν, όπως συμβαίνει στα διάφορα υπάρχοντα επικρατούντα οντοκεντρικά μοντέλα. Στις ιδιότητες έχει πολύ μεγάλη σημασία η κατεύθυνση της ιδιότητας από την άποψη ότι δεικνυοδοτεί το αντικείμενο της *δήλωσης* που ανήκει. Ένα σύνολο ιδιοτήτων που αποδίδονται σε ένα πόρο ονομάζεται *περιγραφή* του πόρου αυτού.
- **Δήλωση**, με την σειρά της ονομάζεται ο συνδυασμός της τριάδας *πόρος - ιδιότητα - τιμή ιδιότητας*. Συγκεκριμένα, τα τρία αυτά μέρη κάθε μίας δήλωσης ονομάζονται θέμα (*subject*), κατηγορημα (*predicate*) και αντικείμενο

(object) αντίστοιχα με την σειρά αναφοράς τους (βλ. Σχήμα 1 παράδειγμα δήλωσης ενός γράφου).



Σχήμα 1: RDF Γράφος

Το αντικείμενο μίας δήλωσης μπορεί να είναι είτε πόρος - ο οποίος με την σειρά του μπορεί να έχει κι άλλες ιδιότητες - , είτε κι ένα *Literal*, δηλαδή μια συγκεκριμένη κλάση που χρησιμοποιείται για να περιγράψει τιμές όπως π.χ. αριθμοί και ημερομηνίες μέσω λεκτικών σταθερών (*literals*), όπως είναι απλές συμβολοσειρές οι τιμές άλλων βασικών τύπων δεδομένων *XML*. Οτιδήποτε που μπορεί να αναπαρασταθεί με ένα *literal* μπορεί και να αναπαρασταθεί με ένα *URI*, εάν κι η χρήση του *literal* είναι πιο βολική και δημιουργική συνάμα. Ένα *literal* μπορεί να είναι το αντικείμενο μιας δήλωσης *RDF*, αλλά ποτέ το υποκείμενο της.

Το μοντέλο του *RDF* μπορεί να συντελέσει τόσο στην δημιουργία, περιγραφή αλλά και διαμόρφωση ενός μοντέλου μεταδεδομένων με σημασιολογία κατανοητή από τις μηχανές. Μάλιστα για την επιθυμητή αναπαράσταση του *RDF* υπάρχουν τρεις διαφορετικοί τρόποι προσέγγισης και περιγραφής του:

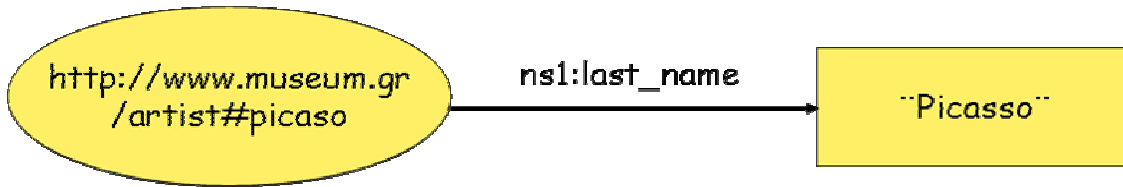
1. Με βάση μια οντοκεντρική προσέγγιση προχωρούμε στη χρήση κατευθυνόμενων γράφων με ετικέτες, που αποτελούν την απεικόνιση του μοντέλου δεδομένων με τρόπο ανεξάρτητο από το χρησιμοποιούμενο συντακτικό.
2. Με βάση μια σχεσιακή περιγραφή, μπορούμε να περιγράψουμε το *RDF* μέσω ενός μοντέλου δηλώσεων-των προαναφερθέντων σχηματιζόμενων *τριάδων* που επίσης εκφράζει την σημασιολογία των περιγραφών.
3. Τέλος, με βάση μια σειριοποίηση των μεταδεδομένων σε αρχεία, και μάλιστα στην συντακτική αναπαράσταση του *RDF*, που με την σειρά του

βασίζεται στο συντακτικό που διέπει την XML. Η αναπαράσταση με την μορφή αυτή είναι βέλτιστη υπό την άποψη ότι είναι κατανοητή κι από τον άνθρωπο, αλλά κι άμεσα επεξεργάσιμη από διάφορα υπολογιστικά προγράμματα να τα διαχειριστούν περαιτέρω.

Εκ του σημείου αυτού και ύστερα θα προχωρήσουμε, μόνο με την χρήση της αναπαράστασης των RDF περιγραφών σε μορφή κατευθυνόμενων γράφων με ετικέτες σε μια ανάλυση του RDF. Κι αυτό γιατί θεωρούμε πως με αυτήν την μορφή η περιγραφή του RDF απλοποιείται στο μέγιστο βαθμό και παρουσιάζει με πάσα ευκρίνεια τα χαρακτηριστικά του μοντέλου δεδομένων της RDF, που θέλουμε άλλωστε να αναδείξουμε μέσω ενός βέλτιστου σχήματος κωδικοποίησης.

2.3 RDF Γράφοι

Οι RDF γράφοι είναι κατευθυνόμενοι γράφοι με ετικέτες, όπου όπως προείπαμε οι μεν κόμβοι αντιστοιχούν στους πόρους ή σε *Literal* και οι δε ακμές αντιστοιχούν στις ιδιότητες. Διαφοροποίηση υπάρχει και ανάμεσα στην γραφική απεικόνιση των κόμβων που αντιστοιχούν σε πόρους να απεικονίζονται με ελλείψεις, ενώ οι κόμβοι που αντιστοιχούν σε *literal* να απεικονίζονται με παραλληλόγραμμα. Δεδομένου ότι το πεδίο ορισμού μιας ιδιότητας είναι πάντα ένας πόρος, ενώ η τιμή της είναι είτε ένας πόρος είτε μια τιμή βασικού τύπου. Άρα μια ακμή στον γράφο ξεκινά πάντα από ένα ελλειψοειδές κόμβο και καταλήγει σε ένα ελλειψοειδή ή παραλληλόγραμμο κόμβο, όπως βλέπουμε και στο σχήμα 2 ένα απλό παράδειγμα γράφου.



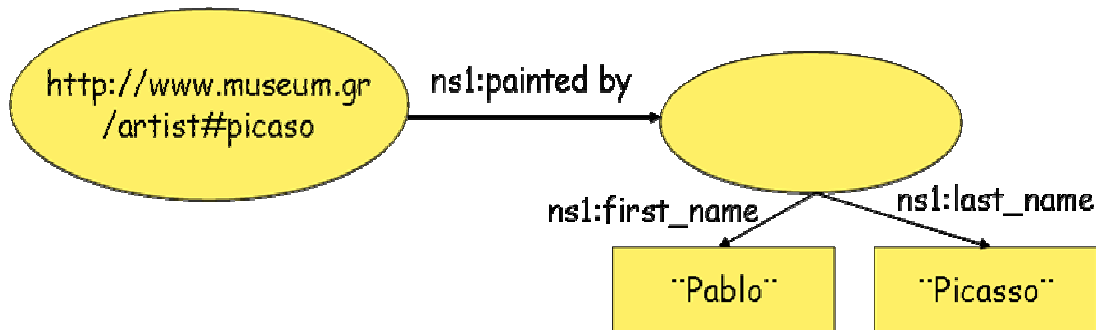
Σχήμα 2: Παράδειγμα RDF γράφου (με τιμή ιδιότητα Literal)

Ο γράφος του σχήματος 2, μεταφράζοντας τον σε φυσική γλώσσα, μας αναπαριστά την φράση: Το επώνυμο (`ns1:last_name`) του <http://www.museum.gr/artist#picaso> είναι *"Picasso"*. Στο σημείο αυτό θα πρέπει να προσθέσουμε πως το πρόθεμα `ns1` που χρησιμοποιήσαμε μπροστά από την ονομασία της ιδιότητας `last_name` είναι μια έκφραση συντομογραφίας του πλήρους URI που θα μπορούσε να χρησιμοποιηθεί, δηλαδή http://www.culture.gr/schema.rdf#last_name. Η δυνατότητα αυτή συντομογραφίας προέρχεται από την αντίστοιχη δυνατότητα ονοματοδοσίας της XML, η οποία καθορίζει τον ορισμό προθεμάτων που υποδεικνύουν το έγγραφο στο οποίο έχει οριστεί το αντίστοιχο στοιχείο.

Τα αναγνωριστικά τόσο των κόμβων όσο και των βελών είναι τα *URIs*. Δεδομένου ότι ένας πόρος μπορεί να εμφανίζεται σε πολλές περιγραφές, η μοναδικότητα των *URIs* κάνει εφικτό τον συνδυασμό μεταδεδομένων που αναφέρονται στον ίδιο πόρο και βρίσκονται σε διαφορετικά αρχεία ή σε διαφορετικές περιγραφές.

Ας προχωρήσουμε σε ένα πιο δύσκολο παράδειγμα γράφου, που είναι αυτό του παρακάτω σχήματος 3. Σε αυτό το σχήμα περιγράφουμε το άτομο με όνομα "Pablo" και επίθετο "Picasso" πρέπει να εισάγουμε ένα καινούριο πόρο (όχι *literal*) ώστε να μπορούμε να του αποδώσουμε ιδιότητες. Το αναγνωριστικό του πόρου αυτού είναι άγνωστο και γι' αυτό χρησιμοποιείται ένα άδειο οβάλ σχήμα για την αναπαράστασή του. Σε επίπεδο εφαρμογής θα πρέπει να αποδοθεί κάποιο

μοναδικό αναγνωριστικό σε όλους τους ανώνυμους πόρους ώστε να μπορούν να διαφοροποιηθούν απολύτως μεταξύ τους.



Σχήμα 3: Σύνθετο παράδειγμα RDF γράφου (με τιμή ιδιότητα έναν πόρο)

Οι τιμές μιας ιδιότητας συχνά χρειάζεται να είναι μια συλλογή από πόρους ή από (αλφαριθμητικά). Οι *RDF* συλλογές (*RDF Containers*) στο *RDF* μοντέλο δεδομένων είναι μια ειδική κατηγορία πόρων οι οποίοι περιέχουν συλλογές από πόρους ή *literals*. Υπάρχουν 3 είδη *RDF* συλλογών:

1. **Bag**, ονομασία με την οποία καλείται ένα επιμέρους σύνολο από πόρους ή *literals*. Χρησιμοποιείται για να δηλώσει ότι μία ιδιότητα έχει πολλαπλές τιμές και ότι η σειρά των τιμών αυτών δεν έχει σημασία. Τα πολυσύνολα μπορούν να έχουν περιέχουν πολλές φορές τα ίδια μέλη.
2. **Sequence**, με την οποία καλείται μια ορισμένη λίστα από πόρους ή *literals*. Χρησιμοποιείται για να δηλώσει ότι μία ιδιότητα έχει πολλαπλές τιμές και ότι η σειρά των τιμών αυτών είναι σημαντική. Οι ακολουθίες μπορούν να έχουν περιέχουν πολλές φορές τα ίδια μέλη.
3. **Alternative**, που είναι μια συλλογή από πόρους ή *literals*. Τα μέλη του συνόλου αντιπροσωπεύουν εναλλακτικές τιμές για την τιμή μιας ιδιότητας. Εφαρμογές οι οποίες διαχειρίζονται παραλλαγές 'γνωρίζουν' ότι μπορούν να επιλέξουν οποιαδήποτε από τα μέλη της λίστας σαν τιμή της ιδιότητας. Το πρώτο μέλος μίας παραλλαγής αποτελεί την προκαθορισμένη τιμή, αλλά βέβαια μπορεί να επιλεγεί οποιαδήποτε άλλο μέλος της παραλλαγής σαν τιμή της ιδιότητας.

2.4 Γλώσσα Περιγραφής RDF Σχημάτων (RDF/S)

Η *RDF/S* (*RDF Schema Specification Language*) αποτελεί μια γλώσσα ορισμού σχημάτων, σχήματα στα οποία περιλαμβάνονται επιμέρους κλάσεις και ιδιότητες, κατορθώνει να ικανοποιεί και να συμπληρώνει την εξαρχής αδυναμία του προτύπου του *RDF* να παρέχει επιμέρους μηχανισμούς δήλωσης κι εν συνεχεία περιγραφής των ιδιοτήτων και των κλάσεων που χρησιμοποιούνται στις *RDF* περιγραφές. Η επέκταση του προτύπου αυτού είναι αναγκαία δεδομένου ότι το απλό μοντέλο δεδομένων του *RDF* είναι ένα από μοντέλο περιγραφής των σχέσεων μεταξύ των πόρων.

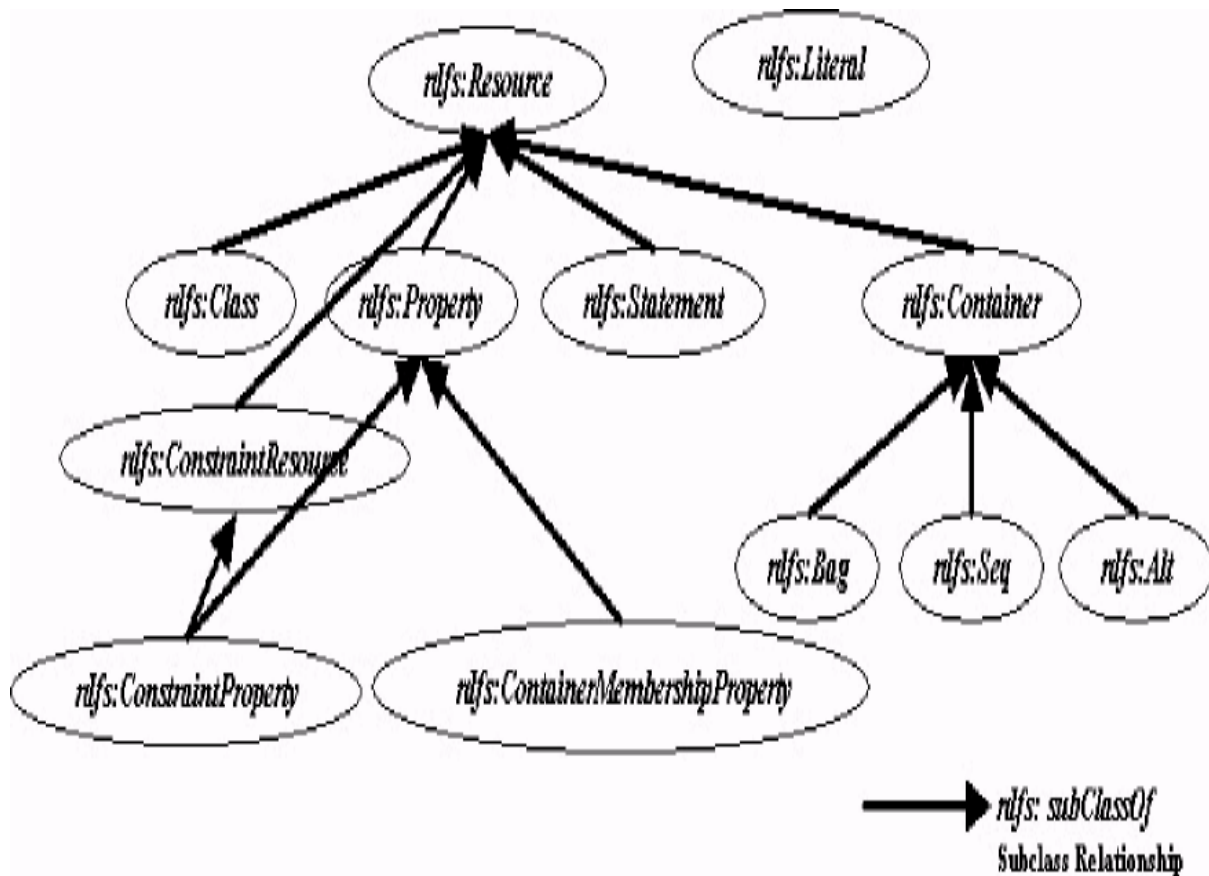
Η *RDF/S* με την σειρά της παρέχει ένα σύνολο από βασικούς τύπους - κλάσεις και ιδιότητες -, που χρησιμοποιούνται για την δημιουργία των *RDF* σχημάτων. Τέτοια χαρακτηριστικά παραδείγματα είναι ότι:

1. Η *RDF/S* ορίζει ιδιότητες που εκφράζουν περιορισμούς για το πεδίο ορισμού και εξίσου για το πεδίο τιμών των ιδιοτήτων.
2. Επίσης η *RDF/S* παρέχει πόρους (αντιστοίχως κι ιδιότητες), οι οποίες με την σειρά τους δηλώνουν σχέσεις μεταξύ κλάσεων ή μεταξύ ιδιοτήτων.

Τα *RDF* σχήματα ορίζονται χρησιμοποιώντας τους τύπους που παρέχει η *RDF/S* και βασίζονται στο *RDF* μοντέλο δεδομένων. Άρα τα σχήματα είναι κι αυτά επιμέρους *RDF* μεταδεδομένα. Το γεγονός αυτό παρέχει ευελιξία στις εφαρμογές δεδομένου ότι δεν χρειάζεται να γνωρίζουν εκ των προτέρων την σημασιολογία των σχημάτων στα οποία στηρίζονται τα μεταδεδομένα που αναλύουν αλλά μπορούν να την εξάγουν. Τα *RDF* σχήματα παρέχουν δυνατότητα ερμηνείας των μεταδεδομένων και έμμεσα επηρεάζουν την δομή των μεταδεδομένων λόγω των περιορισμών που υπάρχουν στο πεδίο τιμών (*rdfs:domain*) και ορισμού (*rdfs:range*) των ιδιοτήτων.

2.5 Βασικές RDF/S Κλάσεις - Ιδιότητες

Στη συνέχεια θα προχωρήσουμε στην ανάλυση των βασικών κλάσεων που σε κάθε περίπτωση είτε ορίζονται είτε περιγράφονται στο RDF σχήμα. Οι κλάσεις συσχετίζονται με σχέσεις υποσυνόλου /υπερσυνόλου κι άρα δημιουργούν μια μορφής ιεραρχία απεικονίζεται στο σχήμα 4 με κορυφή της την κλάση *rdfs:Resource*



Σχήμα 4: Ιεραρχία RDF/S κλάσεων

☑ **rdfs:Resource:** Μέλη της κλάσης αυτής είναι οτιδήποτε μπορεί να περιγραφεί με RDF δηλώσεις. Μπορούμε σε αυτήν να αντιστοιχήσουμε οτιδήποτε προαναφέραμε ανήκει στο σύνολο *Πόροι* (που προαναφέραμε στο 2.2 *RDF Μοντέλο Δεδομένων*).

☑ ***rdfs:Class***: Η σημασιολογία της κλάσης αυτής έγκειται άμεσα με τις γενικές έννοιες *Τύπος*, *Κατηγορία* ή *Κλάση* στις διάφορες οντοκεντρικές γλώσσες. Αναλυτικά, θα πρέπει ο κάθε πόρος που αντιπροσωπεύει μια επιμέρους RDF κλάση να δηλώνεται σαν μέλος της δηλαδή συνοδεύεται απαρεγκλίτως από την ιδιότητα *rdf:type* με τιμή *rdfs:Class*.

☑ ***rdf:Property***: Αναπαριστάνει το υποσύνολο των πόρων που είναι ιδιότητες. ☑ ***rdfs:Literal***: Αντιπροσωπεύει το σύνολο *Literal*, στο οποίο ανήκουν ατομικές τιμές όπως αλφαριθμητικά και ακέραιοι.

☑ ***rdfs:ConstraintResource***: Η κλάση αυτή όπως φαίνεται και στο σχήμα είναι υποκλάση της ***rdfs:Resource***. Μέλη της είναι το υποσύνολο των κλάσεων και ιδιοτήτων αυτών που χρησιμοποιούνται για να εκφράζουν περιορισμούς και με αυτόν τον τρόπο να ελέγχεται η ορθότητα μεταξύ των εξεταζόμενων RDF σχημάτων. Βέβαια ελλοχεύει ο κίνδυνος. Περιορισμοί που δεν έχουν οριστεί στην *RDF/S*, δεν δύνανται να επεξεργαστούν από RDF επεξεργαστή δεν είναι σε θέση να αποφανθεί για την συνέπεια τους ή ασυμβατότητα τους.

☑ ***rdfs:ConstraintProperty***: Αποτελεί υποκλάση της ***rdfs:ConstraintResource*** και της ***rdf:Property*** και μέλη της είναι όλες οι ιδιότητες που δηλώνουν κάποιο περιορισμό. Οι ιδιότητες ***rdfs:range*** και ***rdfs:domain*** που θα περιγραφούν παρακάτω και περιορίζουν το πεδίο τιμών και το πεδίο ορισμού μιας ιδιότητας ανήκουν στη κλάση αυτή.

☑ ***rdfs:Container***: Η κλάση αυτή αντιπροσωπεύει το σύνολο των συλλογών πόρων (*containers*). Υποκλάσεις της είναι οι κλάσεις ***rdf:Bag***, ***rdf:Seq*** και ***rdf:Alt***.

☑ ***rdf:Bag***, ***rdf:Seq*** και ***rdf:Alt***: Αναπαριστούν τα σύνολα *Bag*, *Sequences* και *Alternative* κι είναι υποκλάσεις της *rdfs:Containers*.

☑ ***rdfs:ContainerMembershipProperty***: Περιέχει τις ιδιότητες *rdf:_1*, *rdf:_2...* οι οποίες δηλώνουν τα μέλη μιας συλλογής. Οι ιδιότητες είναι άπειρες. Είναι υποκλάση της *rdf:Property*.

☑ ***rdf:Statement***: Αναπαριστάνει το σύνολο των υποστασιοποιημένων δηλώσεων. Η υποστασιοποίηση αποτελεί μία από τις κύριες ιδιότητες του *RDF/S* σχήματος, καθώς χρησιμοποιείται στην περιγραφή των ίδιων των

μεταδεδομένων. Κι αυτό για να περιγράψουμε μια δήλωση θα πρέπει να δημιουργήσουμε την υποστασιοποιημένη δήλωση της. Η υποστασιοποιημένη δήλωση είναι ένας νέος πόρος με τέσσερις ιδιότητες ***rdf:subject*** (αντίστοιχο θέμα υποστασιοποίησης της δήλωσης), ***rdf:predicate*** (ομοίως το θέμα της δήλωσης), ***rdf:object*** (ομοίως το θέμα της δήλωσης) και ***rdf:type*** (για κάθε πόρο που αντιστοιχεί σε υποστασιοποιημένη δήλωση η κλάση ***rdf:Statement***).

2.6 Βασικές RDF/S Ιδιότητες

Παρακάτω αναλύονται οι πλέον βασικές ιδιότητες που ορίζονται κι εν συνεχεία χρησιμοποιούνται στο *RDF* σχήμα.

☑ ***rdf:type***: Δηλώνει ότι ένας πόρος είναι μέλος μιας κλάσης. Στο *RDF* δεν υπάρχει η έννοια της αυστηρής κληρονομικότητας των ιδιοτήτων και οι κλάσεις δεν επιβάλλουν καμιά δομή στα μέλη τους κι ένας πόρος μπορεί να είναι μέλος πολλαπλών κλάσεων. Οι κλάσεις που ορίζονται στην *RDF/S* άλλα και οι κλάσεις που ορίζονται στα *RDF* σχήματα έχουν την ιδιότητα ***rdf:type*** με τιμή ***rdfs:Class***. Αντίστοιχα, η ιδιότητα ***rdf:type*** αποδίδεται τόσο στις ιδιότητες που ορίζονται στην *RDF/S* όσο και στις απλές ιδιότητες με τιμή ***rdf:Property***. Το πεδίο ορισμού της ιδιότητας ***rdf:type*** είναι η κλάση ***rdfs:Resource***.

☑ ***rdfs:subClassOf***: Δηλώνει την σχέση υποσυνόλου-υπερσυνόλου μεταξύ επιμέρους κλάσεων κι είναι μεταβατική ιδιότητα.. Θα πρέπει ο πόρος στον οποίο εφαρμόζεται η ιδιότητα ***rdfs:subClassOf*** όσο και η τιμή της ιδιότητας πρέπει να είναι κλάσεις. Μια κλάση μπορεί να είναι υποκλάση πολλών κλάσεων. Δεν μπορεί όμως να δηλωθεί σαν υποκλάση του

εαυτού της ή κάποιας από τις υποκλάσεις της, κι αυτό γιατί δεν επιτρέπεται η δημιουργία κύκλων στην ιεραρχία των υποκλάσεων.

☑ ***rdfs:range/rdfs:domain:*** Από κοινού μέλη της κλάσης ***rdfs:ConstraintProperty*** και χρησιμοποιούνται για να ορίσει τις κλάσεις οι οποίες αποτελούν το πεδίο τιμών ή στις οποίες μπορεί να εφαρμοστεί μια ιδιότητα. Μια ιδιότητα μπορεί να έχει μηδέν, μία ή περισσότερες *rdfs:domain* ιδιότητες.

☑ ***rds:predicate, rdf:subject, rdf:object:*** Αντιπροσωπεύουν τις ιδιότητες κατηγορημα (predicate), θέμα (subject) και αντικείμενο (object) και χρησιμοποιούνται για να σχηματίσουν την υποστασιοποιημένη δήλωση.

Συμπερασματικά

Ήδη κυκλοφορούν κάποια συστήματα υπό τη μορφή πρωτοτύπων, που λειτουργούν κατανεμημένοι χώροι αποθήκευσης (repository) για τον Παγκόσμιο Ιστό (WWW). Τα προϊόντα αυτά, υποστηρίζουν την ομοσπονδιακή (federated) ανάπτυξη και αποθήκευση των μοντέλων αναφοράς και των οντολογιών (ontologies), καθώς και την αντιστοίχιση και προβολή ανάμεσα σε διαφορετικά συστήματα μοντέλα/σχήματα, ενώ υποστηρίζουν τα πρότυπα του W3C (όπως XML και RDF).

Ο σημασιολογικός Ιστός (Semantic Web) αποτελεί μια καινοτομία "εν τη γενέσει" της, η οποία υπόσχεται την ενσωμάτωση όψεων (views) βάσεων δεδομένων, "συναλλαγών" μεταξύ βάσεων δεδομένων (database transactions), λογικών αναπαραστάσεων, συνδέσμων Ιστού (Web links) και αντικειμενοστρεφών αναπαραστάσεων, σε μια σημασιολογική βάση.

Τα παραπάνω χαρακτηριστικά οδηγούν σε ένα νέο τρόπο στην επεξεργασία ερωτήσεων (query processing). Η τεχνολογία που παρουσιάζεται θα είναι υποχρεωτική, κατά τους δημιουργούς της, στις περισσότερες κατανεμημένες εφαρμογές-πελάτη, ενώ δεν υπάρχει περιορισμός ως προς τα πεδία (αγορές) εφαρμογής της.

Το όραμα της δημιουργίας του Semantic Web στηρίζεται στην επέκταση των υπάρχοντων πλαισίων περιγραφής μετα- δεδομένων και ειδικότερα στην ύπαρξη σημασιολογικού περιεχομένου που είναι δυνατόν να υπόκειται σε αυτόματη επεξεργασία από τον υπολογιστή χωρίς την επέμβαση του ανθρώπινου παράγοντα.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- ❶ T. Berners-Lee, R. Fielding, L. Masinter. “Uniform Resource Identifiers (URI): Generic Syntax” . RFC 2396. August 1998.
Available at <http://www.ietf.org/rfc/rfc2396.txt>
- ❷ Tim Berners-Lee, James Hendler, Ora Lassila. “The Semantic Web”, Scientific American, May 2001.
Available at <http://www.sciam.com/2001/0501issue/0501berniers-lee.html>
- ❸ D. Brickley, R.V. Guha. “RDF Vocabulary Description Language 1.0: RDF Schema”. W3C Working Draft. April 30,2002.
Available at <http://www.w3.org/TR/2000/rdf-schema>
- ❹ D. Brickley, R.V. Guha. “Resource Description Framework Schema (RDF/S) Specification 1.0”. W3C Candidate Recommendation. March 27, 2000.
Available at <http://www.w3c.org/TR/2000/CR-rdf-schema-20000327/>
- ❺ D. Fensel, C. Bussler, Y. Ding, V. Kartseva, M. Klein, M. Korotkiy, B. Omelayenko, R. Siebes. “Semantic Web Application Areas”. In Procceding of 7th International Workshop on Application of Natural Language to Information Systems, Stockholm, Sweden, June 27-28, 2002