

Τεχνολογία ανάπτυξης νέων ιστοσελίδων στο Semantic Web.

ΒΕΖΕΡΗΣ ΔΗΜΗΤΡΙΟΣ

Τμήμα Ηλεκτρολόγων Μηχανικών & Μηχανικών Η/Υ

Δημοκρίτειο Πανεπιστήμιο Θράκης

Λευκίππου 6, 67100 Ξάνθη

ΕΛΛΑΔΑ

leader@cosmos4u.com, www.vezeris.com

Abstract: - Ζούμε στην εποχή του Web 3.0, του Semantic Web – Σημασιολογικού ιστού, το οποίο θα έχει τον πλέον ενεργό ρόλο στην δεκαετία που διανύουμε. Σε αυτό το επίπεδο ανάπτυξης του Web έχουν εμφανιστεί και διαβουλεύονται νέες τεχνολογίες ανάπτυξης ιστοσελίδων, οι οποίες τεχνικές προάγουν τη σημασιολογία και ταξινόμηση της γνώσης και των δεδομένων. Έτσι όλη η επιστημονική κοινότητα του είδους, περιστρέφεται γύρω από ορισμούς, οι οποίοι έχουν να κάνουν με τον ορισμό των δεδομένων και των ορισμό της σχέσης που έχουν αυτά μεταξύ τους. Στο παρόν συγκεντρώνουμε το επικρατέστερο μέχρι σήμερα σχήμα τεχνολογιών ανάπτυξης του Web 3.0, λαμβάνοντας υπόψη συνέδρια και υιοθετήσεις του W3C καθώς και τη σχετική βιβλιογραφία.

Key-Words: - Τεχνολογίες του σημασιολογικού ιστού, σημασιολογικός ιστός, web 3.0, semantic web, XML,RDF,OWL, Semantic technologies.

1 Εισαγωγή

Το Semantic Web είναι η εξέλιξη του World Wide Web (WWW). Το όλο αυτό σχέδιο βασίζεται σε μια ιδέα του Tim Berners-Lee, ιδρυτή του World Wide Web. Σκοπός του Semantic Web είναι η αξιολόγηση και συσχέτιση των πληροφοριών.

Οι πληροφορίες του διαδικτύου θα κατανοούνται από μηχανές και θα επεξεργάζονται αυτόματα. Πληροφορίες που αφορούν πρόσωπα, περιοχές και αντικείμενα θα μπορούν να συσχετιστούν με τη βοήθεια του Semantic Web. Κατά τη διαδικασία της επεξεργασίας των πληροφοριών σε ένα Semantic Web θα μπορούσαν να ανακαλυφθούν νέοι συσχετισμοί οι οποίοι μέχρι τώρα δεν ήταν αναγνωρίσιμοι.

Ενώ το World Wide Web παρέχει τη δυνατότητα της δικτύωσης όλων των δεδομένων του κόσμου, το Semantic Web ανοίγει το δρόμο για τη διασύνδεση των δεδομένων σε επίπεδο της σημασίας τους.

Το World Wide Web μπορεί να κατανοηθεί μόνο από τον ανθρώπινο νου. Σε ένα Semantic Web οι πληροφορίες είναι ταξινομημένες, έτσι ώστε να είναι κατανοητές και να μπορούν να επεξεργάζονται από υπολογιστές.

Επιπλέον οι υπολογιστές μπορούν να λαμβάνουν από τα δεδομένα κάποιες πληροφορίες από τις οποίες θα είναι δυνατόν να δημιουργήσουν νέες πληροφορίες. Κάποια στοιχεία της προέλευσης του Semantic Web τα συναντάμε στο χώρο έρευνας της τεχνητής νοημοσύνης

2 Περιορισμοί και στόχοι

Στον τομέα της πληροφορικής βρισκόμαστε πολλές φορές σε θέση να παρουσιάσουμε κάτι το οποίο έχουμε αναγνωρίσει ή έχουμε σκεφτεί. Ο ανθρώπινος νους μπορεί να χρησιμοποιήσει αποθηκευμένες βασικές γνώσεις ανακαλώντας τις από τη μνήμη του, να χρησιμοποιήσει λεξικά, οδηγίες, βιβλία εκμάθησης κλπ. και να συνδυάσει όλα τα παραπάνω μεταξύ τους.

Κάτι ανάλογο δεν είναι δυνατόν να πραγματοποιηθεί σήμερα στον τομέα της πληροφορικής με τα μέσα που διατίθενται σήμερα. Το World Wide Web δεν είναι σε θέση να αντλήσει πληροφορίες από διάφορες βάσεις δεδομένων, να τις συνδυάσει και να μας δώσει απαντήσεις σε ερωτήματα που θα θέσουμε.

Σήμερα λαμβάνουμε απαντήσεις σε ερωτήματα τύπου:

- Ποιες ιστοσελίδες περιέχουν τον όρο «X»;
- Ποιες είναι οι καλύτερες ιστοσελίδες που περιέχουν τον όρο «X»;

Θα επιθυμούσαμε να μπορούμε να πάρουμε απάντηση μέσω διαδικτύου στο εξής ερώτημα:

- Τι θερμοκρασία έχει σήμερα η Ξάνθη;
- Από ποιον προμηθευτή σε ακτίνα 3 χλμ. από το πανεπιστήμιο της Ξάνθης μπορώ να αγοράσω φορητό υπολογιστή με λιγότερο από 600 €;

2.1 Η αναζήτηση πληροφορίας σήμερα

Η περαιτέρω επεξεργασία των πληροφοριών και η διασύνδεση των δεδομένων υποστηρίζεται από λογισμικό σε ελάχιστες περιπτώσεις.

Το κυριότερο εργαλείο στο διαδίκτυο είναι οι **μηχανές αναζήτησης**.

Με τις μηχανές αναζήτησης μπορούμε να εντοπίσουμε ιστοσελίδες βασισμένες σε λέξεις κλειδιά, αλλά δεν υπάρχει η δυνατότητα να διασυνδέσουμε τις πληροφορίες των ιστοσελίδων αυτών.

Τις πληροφορίες τις οποίες αναζητούμε πρέπει να τις αντλήσουμε ανοίγοντας την κάθε ιστοσελίδα ξεχωριστά.

Προβλήματα των μηχανών αναζήτησης:

- Υψηλή ανάκληση, χαμηλή ακρίβεια. Ακόμη και έχουν ανακτηθεί οι κύριες σχετικές σελίδες.
- Συχνά συμβαίνει να μην παίρνουμε καμία σχετική απάντηση για το αίτημά μας.
- Τα αποτελέσματα είναι ιδιαίτερα ευαίσθητα στο λεξιλόγιο
- Τα αποτελέσματα είναι πάντοτε σελίδες και ποτέ απαντήσεις σε ερωτήματα. Για να λάβουμε τις απαιτούμενες πληροφορίες απαιτείται αρκετός χρόνος και προσπάθεια για την περιήγηση στις ιστοσελίδες και την ανάκτηση των έγγραφων για τις σχετικές πληροφορίες
- Τα αποτελέσματα δεν είναι προσβάσιμα από άλλα εργαλεία λογισμικού
- Η σημασία του περιεχομένου μιας ιστοσελίδας δεν είναι προσβάσιμη και κατανοητή από μια μηχανή αναζήτησης.

Παρά τις συνεχείς βελτιώσεις των μηχανών αναζήτησης, δεν είναι αισθητή η βελτίωση των παραπάνω προβλημάτων.

2.2 Το Web 3.0

Στο Semantic Web, η πληροφορία θα είναι ταξινομημένη και τυποποιημένα καταχωρημένη μέσα στην ίδια την ιστοσελίδα. Συνεπώς και η εμφάνιση της πληροφορίας θα είναι τυποποιημένη. Κατά συνέπεια πλέον των ανθρώπων, το λογισμικό θα μπορεί να διαχειριστεί και επεξεργαστεί τα δεδομένα και τη γνώση του Web, θα μπορεί να το ταξινομήσει λόγω του ότι το Web θα έχει μετασηματιστεί σε μία τεράστια βάση δεδομένων.

Με την ολοκλήρωση του παραπάνω μετασηματισμού του ιστού, θα αναπτυχθούν εργαλεία λογισμικού, τα οποία αυτόματα θα εισάγουν ερωτήματα σε όλο το γίγνεσθαι του ιστού. Θα μπορεί πλέον ο χρήστης να υποβάλει στο

εργαλείο ερώτημα όπου η απάντηση δεν θα είναι πιθανοί ιστοχώροι όπου μπορεί να υπάρχει η πληροφορία που ζητάμε, αλλά η απάντηση θα είναι η ίδια η πληροφορία μέσω στοιχείων που έχουν αντληθεί από το σημασιολογικό web.

Στα e-shop για παράδειγμα, σήμερα επισκεπτόμαστε και φιλτράρουμε οι ίδιοι μετά από έρευνα τα αποτελέσματα, βάση προδιαγραφών και περιορισμών που έχουμε αποφασίσει. Στο Web 3.0 αυτή η εργασία θα γίνεται από λογισμικό (το οποίο το ονομάζουμε πράκτορα), όπου θα συλλέγει τις πληροφορίες από διάφορα e-shop και θα μας δίνει την πληροφορία –ως απάντηση σύμφωνα με τα κριτήρια που θα έχουμε θέσει.

Στα wikis σήμερα μπορούμε να προσθαφαιρέσουμε σελίδες, γίνεται brainstorming, εξάγονται ιδέες κλπ. Στο Web 3.0, η μηχανή θα μπορεί ακλουθώντας τις συνδέσεις που υπάρχουν στις σελίδες, θα ανακτά τη γνώση και θα μπορεί να αποφασίζει για την απάντηση, βασισμένη πάνω στη κοινή λογική.

2.3 Ο στόχος του Web 3.0

Σήμερα ο παγκόσμιος ιστός μετατρέπεται σε κοινωνία της γνώσης. Οι υπολογιστές ήταν για υπολογισμούς ενώ σήμερα είναι συστήματα επεξεργασίας πληροφοριών, βάσεις δεδομένων, εφαρμογές επεξεργασίας γραπτού λόγου, παιχνιδιομηχανές, κοινωνικές συνάξεις, ηλεκτρονικό εμπόριο και μηχανές αναζήτησης.

2.3.1 Σύγχρονοι περιορισμοί.

Στις αναζητήσεις πληροφορίας στο σύγχρονο ιστό, η απάντηση από της μηχανές, πολλές φορές φέρνουν και σελίδες που είναι ανεπαρκείς ή άσχετες με την αναζήτησή μας. Επιπλέον όλα αυτά τα αποτελέσματα δεν είναι προσβάσιμα από λογισμικό, για περαιτέρω επεξεργασία.

Οι πληροφορίες παρουσιάζουν ασθενές δομημένη μορφή ενώ η διαχείριση της γνώσης γίνεται ως:

A) Αναζήτηση με λέξη κλειδί,

B) Χρειάζεται ανθρώπινος χρόνος για εξαγωγή πληροφοριών,

Γ) Η πληροφορία είναι κάτι που ανακαλύπτεται,

Δ) Οι προβολές περιεχομένου γίνεται βάση της λέξεις κλειδί και πλέον αυτού είναι ανεξέλεγκτες.

2.3.1 Σύγχρονοι στόχοι.

Σε αντιπαραβολή με τους παραπάνω περιορισμούς, στο σύγχρονο σημασιολογικό ιστό:

A) Η γνώση θα οργανωθεί σύμφωνα με την έννοια (εννοιολογικά - σημασιολογικά),

B) Θα συντηρείτε μέσω αυτόματων εργαλείων,

Γ) Η λέξη κλειδί θα αντικατασταθεί από ερωτήματα,

Δ) Κύριο λόγω στην αναζήτηση θα έχουν τα ερωτήματα τύπου βάσεων δεδομένων – queries,

E) Θα έχουμε ορισμούς δικαιωμάτων πρόσβασης και θα εξοικειωθούμε με αξιόπιστες και μη πηγές.

Θα έχουμε πλέον να αντιλαμβανόμαστε και να πληροφορούμαστε αν μία πηγή είναι ή όχι αξιόπιστη.

3 Τεχνολογίες Semantic Web.

Η λύση στα θέματα που προαναφέρθηκαν είναι η χρήση του Semantic Web.

3.1 Τεχνολογίες Semantic Web

Για να ολοκληρωθεί η μετάβαση στο Semantic Web, δεν απαιτούνται νέα επιστημονικά επιτεύγματα, απλά πρέπει να οριστούν πρότυπα, βάση των οποίων θα γίνει η ανάπτυξη του σημασιολογικού web και θα οριστούν οι κανόνες.

Αυτά που έχει πλέον υιοθετήσει το W3C και που διαβουλεύεται ολόκληρη η επιστημονική κοινότητα, περιληπτικά είναι τα παρακάτω στοιχεία, που συνθέτουν τις τεχνολογίες του σημασιολογικού ιστού:

A) Η XML σε αντίθεση με την HTML, εισάγει δομή στα έγγραφα DTDs και XSLT,

B) Για να δηλώνουμε πληροφορίες σχετικά με αντικείμενα, ορίζουμε την RDF και την RDF schema,

Γ) Στη διαβούλευση για την χρήση γλώσσας οντολογίας επικρατούν οι OWL και οι RDF schema,

Δ) Προάγουμε θέματα λογικής (Logic) και ορίζουμε κανόνες (rules).

3.1.1 XML

Για να μπορέσουμε να κατασκευάσουμε σελίδες, φιλικές προς το web 3.0, θα πρέπει να βρούμε έναν πιο νοικοκυρεμένο τρόπο σύνταξης και παρουσίασης των δεδομένων. Η HTML δεν ενδείκνυται διότι ορίζει παραγράφους χωρίς να δίνει έννοιες στα δεδομένα που έχει. Η XML όμως με τη χρήση των meta-λέξεων μπορεί να ορίζει τη σημασία των δεδομένων που περιέχει. Έτσι λοιπόν δεν χρειάζεται να ορίσουμε νέα πρότυπα απλά να υιοθετήσουμε πάγια την τεχνολογία της XML στην ανάπτυξη ιστοσελίδων.

Έχουμε λοιπόν μία HTML με μέτα-δεδομένα (δεδομένα σχετικά με δεδομένα). Έχουμε λοιπόν δεδομένα επεξεργαζόμενα από υπολογιστές.

Η HTML, έχει υιοθετηθεί, ως η γλώσσα πρότυπο για την ανάπτυξη ιστοσελίδων. Αυτό είχε ως συνέπεια να οριστεί ως πρότυπο και να υποστηρίζεται από όλα τα εργαλεία ανάπτυξης ιστοσελίδων. Ο σημασιολογικός ιστός θα είναι το νέο βήμα όπου η XML θα είναι ορισμένη ως πρότυπο ιστοσελίδας ή τύπου XML (πχ HTML 5.0) το οποίο πρότυπο θα ακολουθεί τον κανόνα σημασιολογίας των δεδομένων.

3.1.2 Οντολογίες

Οι οντολογίες δίνουν τη δυνατότητα να επιβληθεί μία σαφής «εννοιολογική» σύλληψη οποιουδήποτε υλικού, το οποίο ενδεχομένως να είναι ψηφιοποιημένο ή/και πολυμεσικό.

Ο ορισμός από τον R.Studer είναι «AN ontology is an explicit and formal specification of a conceptualization» και σε μετάφραση του παρόντος συγγραφέα, «Οντολογία είναι η ρητή και επίσημη ορισμένη προδιαγραφή μίας εννοιολογίας».

Οντολογία λοιπόν είναι η τυποποιημένη και ακριβής προδιαγραφή την έννοιας. Οι οντολογίες έχουν όρους (ορισμούς) και σχέσεις μεταξύ των. Στη γλώσσα της οντολογίας έχουμε ιδιότητες στη σχέσεις, περιορισμούς, λογικές σχέσεις κλπ. Ο ρόλος της οντολογίας είναι η διαλειτουργικότητα και η οργάνωση και πλοήγηση στο site. Είναι εργαλείο καλύτερης ακρίβειας έρευνας & εύρεσης.

Σήμερα οι πιο βασικές γλώσσες οντολογίας είναι: RDF (μοντέλο δεδομένων για αντικείμενα (resources), και σχέσεις μεταξύ τους (relations)), OWL (Πιο πλούσια γλώσσα περιγραφής λέξεων από την RDF).

Στις βασικές σχέσεις που χαρακτηρίζουν αυτού του τύπου τις προσεγγίσεις συγκαταλέγεται η σχέση επωνυμίας. Μία επικρατούσα ερμηνεία είναι η του συνόλου – υποσυνόλου, π.χ. το ανθρώπινο είδος είναι υποσύνολο του ζωικού βασιλείου. Άλλες σχέσεις που επίσης χρησιμοποιούνται είναι αυτή της μερωνυμίας (σχέση όλου – μέρους) ή της αντωνυμίας (π.χ. σε δεδομένο σημασιολογικό πεδίο η ηλιοφάνεια είναι αντώνυμο της βροχής).

Η οργάνωση του υλικού βάσει οντολογιών

1. διακρίνεται από ευελιξία και επεκτασιμότητα
2. δεν εξαρτάται από συγκεκριμένη φυσική γλώσσα
3. μπορεί να υλοποιηθεί σε κάποια πρότυπη γλώσσα κωδικοποίησης (π.χ. XML, RDF κλπ)
4. επιτρέπει την εξαγωγή συμπερασμάτων (inferencing) με σχετική ασφάλεια, έτσι ώστε να αναδεικνύονται οι όποιες σχέσεις μεταξύ διαφορετικών αντικειμένων του πολυμεσικού υλικού
5. προσφέρει δυνατότητες πρόσβασης στο υλικό, οι οποίες είναι εξατομικευμένες, λεπτομερείς, ασφαλείς, μη πεπερασμένες και μη εξαρτημένες από γεωγραφικούς περιορισμούς
6. εξασφαλίζει τη βιωσιμότητα του υλικού στο πέρασμα του χρόνου, στοιχείο εξόχως αναγκαίο στην περίπτωση που το ψηφιοποιημένο υλικό είναι 'μοναδικό', όπως

στην περίπτωση του πολιτισμικού υλικού (έργα τέχνης, χειρόγραφα κλπ).

Σήμερα οι σημαντικότερες γλώσσες οντολογίας είναι οι:

RDF-Schema: Η γλώσσα RDF (Resource Description Framework - 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. Πόρος (resource) είναι οτιδήποτε θέλουμε να δηλώσουμε ή να περιγράψουμε. Παραδείγματος χάρη, πόρος μπορεί να είναι μία ιστοσελίδα, ένας δικτυακός τόπος, ένα αντικείμενο, μία έννοια κτλ. Κάθε πόρος προσδιορίζεται με το Καθολικό Αναγνωριστικό Πόρου (Universal Resource Identifier-URI). Η γλώσσα RDF είναι ένα απλό μοντέλο δεδομένων, στο οποίο όλες οι προτάσεις αποτελούν μια τριπλέτα της μορφής: P(O, V) όπου Property (ιδιότητα), Object (αντικείμενο), Value(τιμή). Η συντριπτική πλειοψηφία της γνώσης που θέλουμε να αναπαραστήσουμε στους υπολογιστές 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 συντακτικό XML, δηλαδή τον τρόπο με τον οποίο οι προτάσεις RDF εκφράζονται ως XML κείμενα.

Web Ontology Language (OWL): Η οντολογική γλώσσα OWL είναι μία σημασιολογικής σήμανσης γλώσσα για τη δημοσίευση και ανταλλαγή οντολογιών στο διαδίκτυο. Χρησιμοποιεί τη σύνταξη των RDF/XML (McGuinness & Van Harmelen, 2004). Αποτελεί το πιο πρόσφατο επίτευγμα στα πρότυπα των οντολογικών γλωσσών, και είναι πιστοποιημένη από την Κοινοπραξία του Παγκόσμιου Ιστού ότι προάγει το όραμα του Σημασιολογικού Ιστού. Η OWL χρησιμοποιείται από εφαρμογές, που χρειάζονται να επεξεργαστούν το περιεχόμενο των πληροφοριών αντί να επεξεργαστούν απλά την πληροφορία που απλά παρουσιάζεται στους ανθρώπους. Η OWL πηγαίνει πέρα από τις γλώσσες XML και OWL με την ικανότητα της να αναπαριστά μηχανικά εξηγήσιμο περιεχόμενο στο διαδίκτυο, επειδή έχει πιο πολλά μέσα για την έκφραση νοημάτων και εννοιών από ότι αυτές (McGuinness & Van Harmelen, 2004)

DAML+OIL: Η DAML-OIL (γλώσσα σήμανσης πρακτόρων DARPA – γλώσσα διεπαφών

οντολογίας) είναι μια βασισμένη σε πλήρη-pledgedWeb γλώσσα οντολογίας που αναπτύχθηκε πάνω από RDFS.7, που χαρακτηρίζει μια βασισμένη σε XML σύνταξη και μια αρχιτεκτονική σε στρώσεις. Η DAML-OIL παρέχει τους πρωτόγονους μοντελοποίησης που χρησιμοποιούνται συνήθως σε βασισμένες σε πλαίσιο προσεγγίσεις στην εφαρμοσμένη μηχανική οντολογίας (ontology engineering), και την επίσημη υποστήριξη σημασιολογίας και συλλογισμού που βρίσκεται σε προσεγγίσεις λογικής περιγραφής. Ενσωματώνει επίσης τους τύπους δεδομένων σχημάτων XML για τη σημασιολογική διαλειτουργικότητα σε XML.

3.1.3 Λογική

Η λογική ορίζεται στενά ως η ικανότητα ή η διαδικασία εξαγωγής λογικών συμπερασμάτων.

Δεν περιλαμβάνει μόνο την δυνατότητά μας για λογικό συλλογισμό, αλλά επίσης την ικανότητα που έχουμε να διεξάγουμε έρευνα, να λύνουμε προβλήματα, να εκτιμούμε, να κριτικάρουμε, να αποφασίζουμε για το πώς θα πράξουμε και να φτάνουμε σε μια κατανόηση του εαυτού μας, των άλλων ανθρώπων και του κόσμου.

Η λογική στον τομέα της πληροφορικής παρέχει την απαραίτητη δομή, ώστε να διαμορφώνονται κανόνες, βάσει των οποίων το σύστημα του υπολογιστή μπορεί να βγάζει συμπεράσματα και:

A) Είναι η επιστήμη που μελετά τις αρχές της συλλογιστικής,

B) Είναι η επίσημη γλώσσα έκφρασης της γνώσης,
Γ) Σύμφωνα με τα δεδομένα μπορεί να αποφασίζει τι ενέργειες και τι απάντηση θα δώσει,

Δ) Μπορεί να επιχειρηματολογήσει για μία απόφαση ή εξαγωγή συμπεράσματος,

E) Υποστηρίζεται από τις γλώσσες DAML+OIL και OWL.

3.1.4 Πράκτορες

Οι πράκτορες (agents) είναι κομμάτια κώδικα (λογισμικό). Ο πράκτορας που θα επιλέγει ο χρήστης θα επιλέγει την καλύτερη ανάμεσα στις λύσεις που θα προκύπτουν και θα την προτείνει στον χρήστη.

Αυτό θα γίνεται εξάγοντας δεδομένα μέσα από τα μέτα-δεδομένα.

Η λογική που προαναφέραμε θα χρησιμοποιείτε για την επεξεργασία και την ανάκτηση πληροφοριών για την σύνταξη των συμπερασμάτων.

3.1.5 Προσέγγιση βάση επιπέδων

Η ανάπτυξη του σημασιολογικού ιστού έχει βηματική πρόοδο. Κάθε βήμα αναπτύσσεται στην κορυφή ενός άλλου (του προηγούμενου) επιπέδου.

Αυτό γίνεται διότι α) είναι μία ήδη επιτυχημένη συνταγή, β) είναι πολύ εύκολο και στιβαρό να αναπτύσσετε κάθε επίπεδο χωριστά. Συνήθως πολλές ομάδες μπορούν να αναπτύξουν την επιστήμη και σε διαφορετικές πορείες για κάθε επίπεδο. Αυτό είναι και το ζητούμενο. Να προκύψει το βέλτιστο για το πρωτόκολλο, για το στάνταρ, για το πρότυπο.

Όσο λοιπόν ερευνητές συμφωνούν σε κάποια σημεία και διαφωνούν σε κάποια άλλα, το ζητούμενο είναι να συμφωνηθεί η συμφωνία.

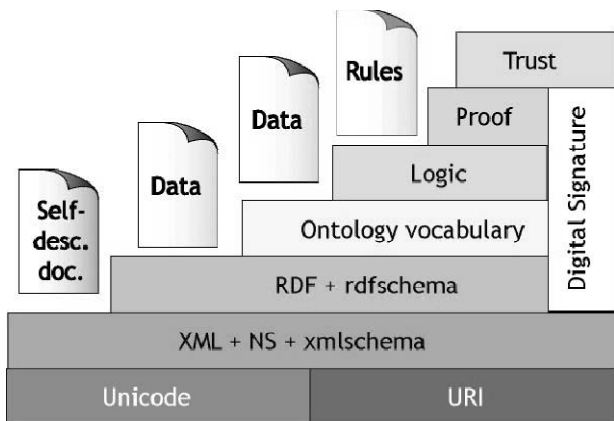
Από τη στιγμή που ορίζεται ένα πρότυπο, τότε αυτό εξελίσσεται από ολόκληρη την κοινωνία.

Χτίζοντας ένα επίπεδο στο σημασιολογικό ιστό, ακολουθούμε δύο αρχές:

A) Η προς τα πίσω συμβατότητα. Οι πράκτορες πρέπει να μπορούν να χειρίζονται δεδομένα όλων των επιπέδων. Αν οι πράκτορες μπορούν να χειριστούν σημασιολογίες της OWL πρέπει να μπορούν να χειρίζονται και σε RDF & RDF schema.

B) Η προς τα εμπρός ανάλογη συμπεριφορά. Λόγω ανοιχτής αρχιτεκτονικής πρέπει οι πράκτορες να μπορούν να χειριστούν υψηλότερα επίπεδα, πχ οι πράκτορες που χειρίζονται RDF θα πρέπει να μπορούν να χειριστούν μερικών και OWL.

Τη στιγμή που αυτά αποτελούν θεωρητικό υπόβαθρο για την ανάπτυξη του σημασιολογικού web, πρέπει να γίνουν πράξη στην υλοποίηση μέσω κάποιων παραδοχών. Στο παρακάτω σχήμα έχουμε τα βασικά επίπεδα του σημασιολογικού ιστού.



Στον πυθμένα βρίσκουμε την XML, μέσω της οποίας γράφουμε δομημένα δεδομένα με λεξιλόγιο ορισμένο από τον χρήστη. Η XML είναι ειδική στο να στέλνουμε έγγραφα μέσω του web.

Η RDF είναι το βασικό μοντέλο δεδομένων, ως το μοντέλο σχέσεων, για να γράφουμε απλές δηλώσεις σχετικά με τα αντικείμενα του ιστού. Το RDF δε βασίζεται στην XML αλλά έχει σύνταξη που μοιάζει με της XML. Κατά τα άλλα είναι ένα επίπεδο πάνω από την XML.

Το σχήμα RDF, RDF schema, παρέχει μοντελοποίηση καταγωγής ώστε να μπορεί να ορίζεται η ιεράρχηση σχέσεων και αντικειμένων του ιστού.

Βασίζεται πάνω στην RDF και μπορείς να γράψεις οντολογία.

Λόγω όμως ότι θέλουμε πολύπλοκη σχέση και εκτενής γλώσσα για να περιγράψουμε πλήρως την οντολογία του ιστού, έχουμε το επίπεδο της λογικής πάνω από όλα τα οντολογικά επίπεδα.

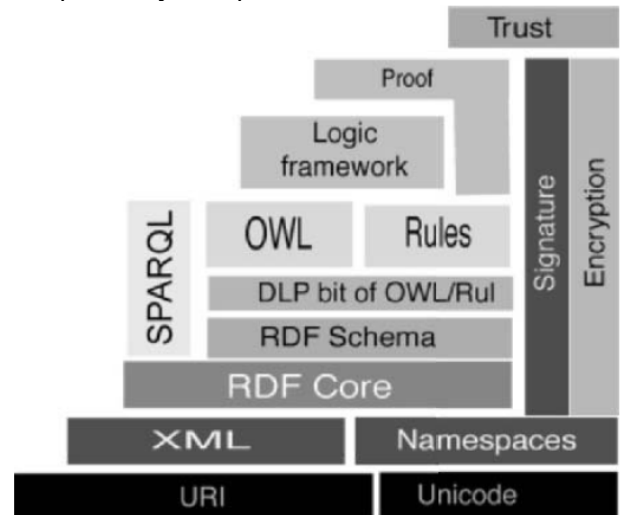
Το λογικό επίπεδο χρησιμοποιείτε για να εκτείνει την οντολογία και να επιτρέψει την εφαρμογή ειδικών δηλώσεων για την πληροφορία της γνώσης.

Το επίπεδο της απόδειξης (proof), είναι το επίπεδο όπου θα δίδει στις νέες γλώσσες του web τις αποδείξεις από τα χαμηλότερα επίπεδα και θα τις επικυρώνει.

Τελικά το επίπεδο εμπιστοσύνης (trust layer) θα διαχειρίζεται τις ψηφιακές υπογραφές και άλλου είδους γνώση, βασιζόμενο σε συστάσεις από πράκτορες που εμπιστεύεται ή εμπιστευτικές πηγές υπηρεσιών – πρεσβείες.

Εδώ πρέπει να καταλάβουμε ότι ο παγκόσμιος ιστός, πρέπει να αναπτύξει την εμπιστοσύνη με ένα τρόπο που να μοιάζει με τον ανθρώπινο. Πρέπει να υπάρχουν κανόνες και αρχές που να ορίζουν την εμπιστοσύνη. Μόνο αν οι χρήστες επιστεφούν την λειτουργία του σημασιολογικού ιστού αυτό θα έχει επιτυχία.

Βέβαια η στοίβα των επιπέδων που είναι υπό διαβούλευση αυτή τη στιγμή έχει και εναλλακτικά σενάρια όπως το παρακάτω:



Οι σημαντικές διαφορές των σχημάτων είναι:

- Το επίπεδο της οντολογίας αντικαθίστανται από δύο εναλλακτικά, που είναι η γλώσσα οντολογίας OWL και η γλώσσα βασισμένοι σε

κανόνες (OWL Language & rule – based language),

- DLP είναι η τομή της OWL και της λογικής και χρησιμεύει ως κοινή αρχή.

Ο σημασιολογικός ιστός χτίζεται στην παρούσα φάση και μπορεί να έχουμε τροποποιήσεις και βελτιώσεις στο μέλλον.

4 Προκλήσεις για το μέλλον

Είναι γεγονός πλέον το λογισμικό ανοιχτής αρχιτεκτονικής. Τα μελλοντικά συστήματα πληροφοριών πρέπει να αλληλεπιδρούν ομαλά με μια μεγάλη ποικιλία ανεξάρτητων πηγών δεδομένων και εφαρμογών που εκτελούνται σε ετερογενείς πλατφόρμες.

Τα meta-δεδομένα θα διαδραματίσουν κρίσιμο ρόλο, στην περιγραφή του περιεχομένου των εν λόγω πηγών δεδομένων.

Τέτοιες αλληλεπιδράσεις μπορούν να περιλαμβάνουν πλοήγηση, υποβολή ερωτήσεων και ανάκτηση, παράλληλα με ασφάλεια και εχεμύθεια.

Η κοινότητα είναι προσανατολισμένη στις πύλες πληροφοριών (portals or information gateways). Σήμερα γίνονται διαθέσιμες βάσεις δεδομένων στο web, σε κάποια μορφή, ανοιχτές σε χρήστες και εφαρμογές. Εδώ η σημασιολογία των δεδομένων πρέπει να υπάρχει μαζί με τα δεδομένα, διότι με διαφορετικό τρόπο παρουσιάζονται στον άνθρωπο – χρήστη, διαφορετικά γίνονται προσπελάσιμα από τις εφαρμογές.

Στο ηλεκτρονικό εμπόριο ο σημασιολογικός ιστός, χειρίζεται και βρίσκει τη σωστή πληροφορία και κρατάει την ιδιωτικότητα, αξιολογεί και ερευνά τη φήμη και την αξιοπιστία των καταστημάτων, είναι ευέλικτο σε συγκερασμό προϋποθέσεων αναζήτησης.

Οι προσωπικοί πράκτορες, (μελλοντικό σενάριο), αποφασίζουν και επιλέγουν βάση των ορισμάτων που τους δίνουμε. Εδώ και είναι η κατακλείδα του σημασιολογικού ιστού δείχνοντας το πραγματικό του πρόσωπο στον τελικό χρήστη. Εδώ θα αποδώσουμε και εμείς το Web 3.0 με ένα παράδειγμα αναζήτησης μέσω πράκτορα μίας πληροφορίας:

Ποιος οφθαλμίατρος είναι διαθέσιμος, κοντά στο σπίτι μου, για την Πέμπτη το απόγευμα ?

To Web 3.0 θα σου απαντήσει.!!!

Αναφορές:

- [1] Jagadeesh Nandigam, Semantic Web Services, CCSC: Midwestern Conference, 2009

- [2] G. Antoniou and F. van Harmelen. A Semantic Web Primer. (Second Edition) The MIT Press, Cambridge, MA, 2008. ISBN: 978-0-262-01242-3.

- [3] Liyang Yu. Introduction to the Semantic Web and Semantic Web Services, Chapman & Hall/CRC, Taylor & Francis Group, 6000 Broken Sound Parkway NW, Suite 300, 2007, ISBN-13:978-1-58488-933-5.

- [4] The Semantic Web, Michael C Daconta, Wiley Publishing 2003, ISBN:0-471-43257-1.

- [5] International Semantic Web Conferences, <http://iswc.semanticweb.org/> last visited 14/1/2009.

- [6] Extended Semantic Web Conference, <http://www.eswc2010.org/>, last visited 14/1/2009.