

ΤΑΣΕΙΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΕΣ ΣΗΜΑΣΙΟΛΟΓΙΚΟΥ ΙΣΤΟΥ

Ερρίκος Καλτσόπουλος
Ηρακλής Θεοφανίδης
Αχιλλέας Γκέκας



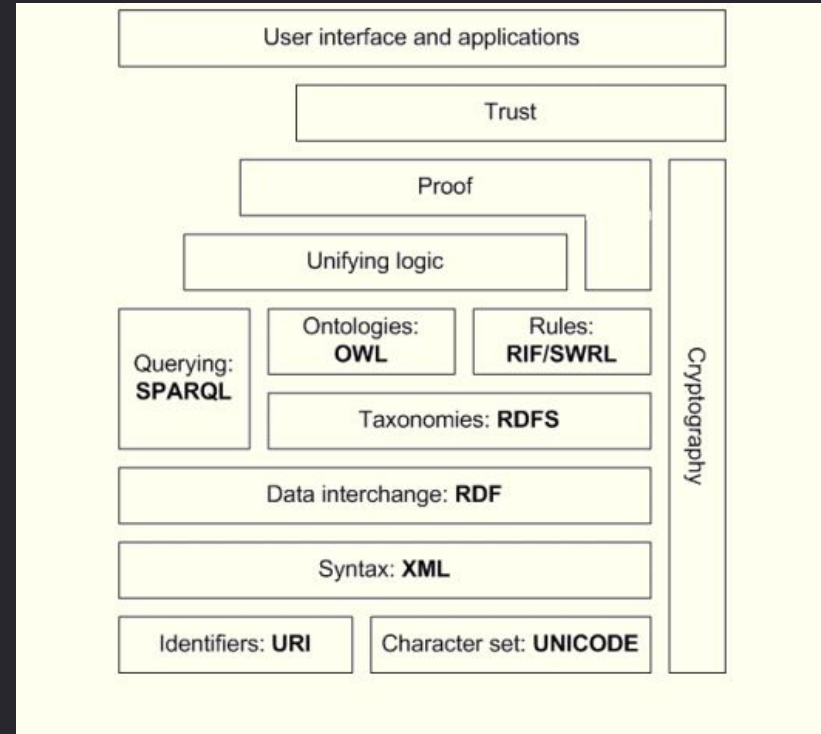
ΕΙΣΑΓΩΓΗ

Θα αναφερθούμε σε κάποιες από τις τεχνολογίες του Σημασιολογικού Ιστού και που αυτές χρησιμοποιούνται, καθώς και για κάποιες τάσεις και την σύνδεση τους με τον Σημασιολογικό Ιστό.



ΤΕΧΝΟΛΟΓΙΕΣ

- RDF & RDF SCHEMA
- OWL
- QUERY LANGUAGES
- RULES(SWRL,RIF)
- LOGIC (RacerPro,Pallet)



RDF & RDF SCHEMA

RDF: Είναι ένα πλαίσιο για την αντιπροσώπευση πληροφοριών σχετικά με τους πόρους σε μορφή γραφήματος. Όλα τα δεδομένα στον Σ.Ι χρησιμοποιούν το RDF ως την κύρια γλώσσα αναπαραγωγής δεδομένων.

RDF SCHEMA: Μπορεί να χρησιμοποιηθεί για να περιγράψει ταξινομίες κατηγοριών και ιδιοτήτων και να τις χρησιμοποιήσει για να δημιουργήσει οντολογίες.

OWL

- Είναι ένας συνδυασμός από Markup language η οποία είναι σχεδιασμένη για χρήση από εφαρμογές οι οποίες χρειάζονται να επεξεργαστούν δεδομένα.
- Είναι η επέκταση του RDFS, περιέχει το λεξιλόγιο του, αλλά προσφέρει μεγαλύτερη εκφραστικότητα στην οποία ο κόσμος είναι ένα σύνολο από κλάσεις, ιδιότητες και άτομα
- Χωρίζεται σε τρεις γλώσσες με διαφορετική εκφραστική ισχύ :
 - Owl Lite
 - Owl DL
 - Owl Full

QUERY LANGUAGES

- SPARQL
- SeRQL
- RDQL
- RQL
- N3
- N-Triple
- Turtle
- VERSA

RULES

SWRL:

- Είναι μια νέα τεχνολογία Σ.Ι. η οποία επεκτείνει την OWL-DL
- Αυτό που προσφέρει η SWRL είναι η δήλωση κανόνων εντός μιας οντολογίας.
- Μπορούν να γραφούν κανόνες τύπου Horn, για να εξάγουμε συμπεράσματα από υπάρχουσες βάσεις γνώσεις OWL.

RIF:

- Επικεντρώνεται στην ανταλλαγή και όχι στην οριοθέτηση μιας ενιαίας γλώσσας ενός κανόνα, σε αντίθεση με άλλα πρότυπα.
- Περιλαμβάνει 5 βασικές διαλέκτους, μια διάλεκτο Core που επεκτείνεται σε έναν βασικό λογικό διάλογο (BLD) και έναν διαλεκτικό κανόνα παραγωγής (PRD), την διάλεκτο FLD και την διάλεκτο DTB.

REASONERS

RacerPro:

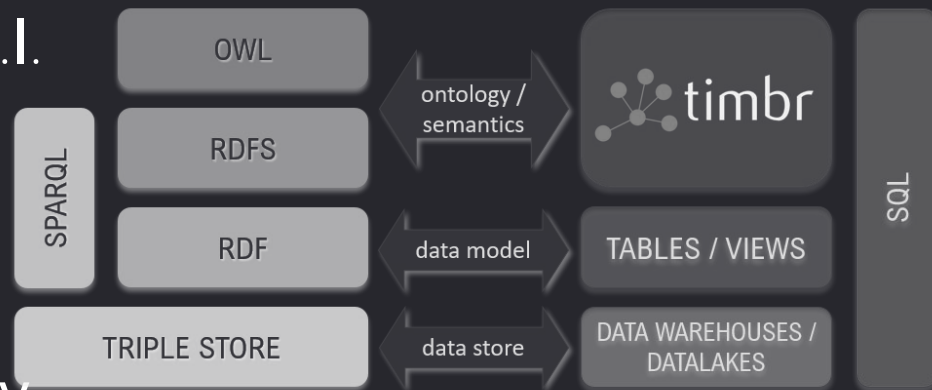
- Μπορεί να επεξεργαστεί οντολογίες γραμμένες σε OWL Lite και OWL DL
- Συμπεριλαμβάνει σε δοκιμαστικό στάδιο μία έκδοση της γλώσσας SWRL
- Προσφέρει ένα open-source API για τη δημιουργία επερωτήσεων μέσω ενός συστήματος OWL-QL
- Προσφέρει την δυνατότητα να δίνει απαντήσεις σε διαδικασίες ανάκτησης πληροφορίας

Pellet:

- Υποστηρίζει την εξαγωγή συμπερασμάτων με βάση την συνεπαγωγή.
- Υποστηρίζει αλγόριθμους που βελτιώνουν την ταχύτητα των ερωτημάτων.
- Επιτρέπει την εξαγωγή συμπερασμάτων από πολλαπλές οντολογίες

TIMBR

- Είναι μια πλατφόρμα που γεφυρώνει Sql και Σ.Ι..
- Εξάλειψη εμποδίων για την είσοδο και χρήση knowledge graph
- Εφαρμόζει τις αρχές του Σ.Ι. ώστε να παρέχει την λειτουργία του Knowledge graph
- Χρήση απλών ερωτημάτων Sql



ΤΑΣΕΙΣ

- Machine Learning
- Deep Learning
- Natural Language Processing
- Text Mining and Text Analytics
- Internet of things
- Big Data

MACHINE LEARNING

- Η μηχανική μάθηση οσον αφορά τον Σ.Ι έχει αναπτυχθεί για να υποστηρίξει τη διαχείριση οντολογιών, και να ενσωματώσει σημασιολογικές πληροφορίες στον web mining.
- Χρησιμοποιείται όλο και περισσότερο για την ανάλυση των κατανεμημένων πηγών δεδομένων που περιγράφονται σε format Σ.Ι.
- Η μηχανική μάθηση βοηθά τον σημασιολογικό ιστό μέσω της οντολογικής αξιολόγησης, της κατασκευής, της διαχείρισης και της βελτίωσης . Συγκρίνει πολλά μέτρα ομοιότητας και αλγόριθμους που χρησιμοποιούνται για τη χαρτογράφηση ή τη συγχώνευση δύο οντολογιών με αλγορίθμους μηχανικής μάθησης.

DEEP LEARNING

- Οι τεχνολογίες του Σημασιολογικού Ιστού χρησιμοποιούνται για τη βελτίωση των μεθόδων βαθιάς εκμάθησης, καθιστώντας τους αλγόριθμους ταξινόμησης πιο ενημερωμένους για τη σημασιολογία των δεδομένων από τα οποία μαθαίνουν.
- Μία χρήση του Deep learning είναι σε ένα πείραμα που έγινε σε μια οντολογία για τα καρδιαγγειακά νοσήματα (CVDO). Άλλη μία συνεργασία Σημασιολογικού ιστού και Deep learning ήταν για τη μείωση του θορύβου για το RDF.

TEXT MINING AND TEXT ANALYTICS

- Ευκολότερη εξαγωγή πληροφοριών από τεράστιους όγκους δεδομένων
 - Αποτελεσματική ευρετηρίαση και βελτιωμένη αναζήτηση
- Εξαγωγή χρήσιμης πληροφορίας από πηγές δεδομένων μέσω της αναγνώρισης και της διερεύνησης ενδιαφερόντων προτύπων.
- Ο συνδυασμός των SWT και της εξόρυξης ιστού :
 - Ενισχύει τις συνέπειες της εξόρυξης ιστού χρησιμοποιώντας τη σημασιολογική δομή στον ιστό
 - Απλούστερη και αποτελεσματικότερη διαδικασία.
 - Βοηθούν τους ανθρώπους να διερευνήσουν τα γεγονότα από τις λεπτομέρειες για να απαντήσουν συγκεκριμένες ερωτήσεις.

NATURAL LANGUAGE PROCESSING

- Το NLP και το SW είναι δύο διαφορετικές τεχνολογίες αλλά έχουν συμπληρωματικούς ρόλους στη διαχείριση δεδομένων.
- Ο συνδυασμός αυτών των δύο τεχνολογιών καθιστά δυνατή την ομοιόμορφη συγχώνευση δομημένων και αδόμητων δεδομένων, για παράδειγμα, ένας machine reader είναι ένα εργαλείο που μετατρέπει το κείμενο φυσικής γλώσσας σε σωστή δομημένη γνώση και το τελευταίο, σύμφωνα με την κοινή σημασιολογία, μπορεί να ερμηνευθεί από μηχανές. Αυτός τον συνδυασμός χρησιμοποιείται σε διάφορες εφαρμογές, όπου αντιπαραβάλλεται με ένα τεράστιο όγκο μη δομημένων ή ανοργάνωτων πληροφοριών. Εξάγονται δομημένα δεδομένα από ένα έγγραφο κειμένου χρησιμοποιώντας NLP και στη συνέχεια συνδέονται αυτά τα εξαγόμενα δεδομένα μέσω σημασιολογικών τεχνολογιών.

BIG DATA

- Οι SWT βοηθούν στις επιχειρήσεις, επιτρέποντας στη συνέχεια να κάνουν ακόμα καλύτερη κρίση σε πραγματικό χρόνο. Για τις οργανώσεις των επιχειρήσεων προς τους καταναλωτές (με επικεφαλής την Google), ο σημασιολογικός ιστός τους επιτρέπει να προσφέρουν στους καταναλωτές άμεσες απαντήσεις.
- Ένα Knowledge graph είναι ένα καλό παράδειγμα μεγάλων δεδομένων για τον σημασιολογικό ιστό.
- Ένα άλλο καλό παράδειγμα είναι το πρωτόκολλο Facebook Open Graph που δίνει τη δυνατότητα σε οποιαδήποτε ιστοσελίδα να γίνει ένα πλούσιο αντικείμενο σε ένα κοινωνικό γράφημα. Η Αναζήτηση Graph Facebook είναι σημασιολογική μηχανή αναζήτησης που εισήχθη το 2013 για να παρέχει απαντήσεις σε ερωτήματα φυσικής γλώσσας των χρηστών αντί για μια λίστα συνδέσμων.

INTERNET OF THINGS

- Ο Σημασιολογικός Ιστός μπορεί να βελτιώσει περαιτέρω την ικανότητά του να κατανοεί τα δεδομένα των πραγμάτων και να διευκολύνει τη λειτουργικότητά τους.
- Ένα από τα προβλήματα που έχει το IoT είναι ότι οι συσκευές μεταξύ τους δεν είναι συμβατές και έτσι δεν υπάρχει επικοινωνία μεταξύ τους, επειδή δεν υπάρχει μεταξύ τους κοινό λεξιλόγιο για την επεξεργασία των δεδομένων τους.
- Έτσι το IoT με την βοήθεια των σημασιολογικών τεχνολογιών όπως οντολογίες, Rdf μας δίνει μία λύση στο πρόβλημα που αναφέραμε.

ΕΥΧΑΡΙΣΤΟΥΜΕ

