



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

Τομέας Συστημάτων Μετάδοσης Πληροφορίας και Τεχνολογίας
Υλικών

**Δημιουργία Αλληλεπιδραστικού Γραφικού
Περιβάλλοντος για την οπτική αναπαράσταση των
συσχετίσεων μεταξύ οντολογιών**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΤΟΥ

Αθανάσιου Χ. Αποστόλου

Επιβλέπων: Θεοδώρα Βαρβαρίγου
Καθηγήτρια Ε.Μ.Π.

Αθήνα, Οκτώβρης 2020



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

Τομέας Συστημάτων Μετάδοσης Πληροφορίας και Τεχνολογίας
Υλικών

**Δημιουργία Αλληλεπιδραστικού Γραφικού
Περιβάλλοντος για την οπτική αναπαράσταση των
συσχετίσεων μεταξύ οντολογιών**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΤΟΥ

Αθανάσιου Χ. Αποστόλου

Επιβλέπων: Θεοδώρα Βαρβαρίγου
Καθηγήτρια Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 17η Αυγούστου 2020.

.....
Θεοδώρα Βαρβαρίγου
Καθηγήτρια Ε.Μ.Π.

.....
Εμμανουήλ Βαρβαρίγος
Καθηγητής Ε.Μ.Π.

.....
Συμεών Παπαβασιλείου
Καθηγητής Ε.Μ.Π.

Αθήνα, Οκτώβρης 2020

.....

Αθανάσιος Χ. Αποστόλου

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © Αθανάσιος Χ. Αποστόλου

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

δφασδ

$\gamma\varphi\delta\varsigma$

[illegible]

Λέξεις Κλειδιά

Abstract

δφσαδφ

φσδαφ

Ευχαριστίες

Η παρούσα διπλωματική εργασία εκπονήθηκε από τον Μάρτιο έως τον Οκτώβρη του 2020 για την ολοκλήρωση των σπουδών μου στην Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Εθνικού Μετσόβιου Πολυτεχνείου. Ο υπεύθυνος τομέας είναι ο Τομέας Επικοινωνιών, Ηλεκτρονικής & Συστημάτων Πληροφορικής υπό την επίβλεψη της καθηγήτριας Θεοδώρας Βρβαρίγου.

Εξαιρετική βοήθεια προήλθε από τον Ευθύμιο Χονδρογιάννη, ο οποίος μου είχε αναθέσει το θέμα της πτυχιακής και ήταν πάντα κατατοπιστικός σε οποιαδήποτε σύνθετα ζητήματα συνάντησα καθ' όλη την διάρκεια και τον ευχαριστώ θερμά.

Κατά την διάρκεια εκπόνησης της διπλωματικής εντρύφησα στις τεχνολογίες του Σημασιολογικού Ιστού και κατανόησα σε βάθος την έννοια της Οντολογίας. Απέκτησα διάφορες γνώσεις για επιστημονικά θέματα που αφορούν την επεξεργασία φαρμακευτικών και κλινικών δεδομένων. Δούλεψα πάνω σε πραγματικά δεδομένα έχοντας την ευκαιρία να συναντήσω ρεαλιστικά προβλήματα και να καταφέρω να τα επιλύσω. Επιπροσθέτως βελτίωσα και εξειδίκευσα τις γνώσεις μου που αφορούν την δημιουργία πραγματικών εφαρμογών τόσο για frontend γραφικών διεπαφών χρηστών, όσο και για backend/servers. Οι γνώσεις αυτές θα συνεισφέρουν θετικά στην αργότερα επαγγελματική μου πορεία.

Πίνακας Περιεχομένων

Περίληψη.....	5
Abstract.....	7
Ευχαριστίες.....	9
1. Εισαγωγή.....	14
2. State of the Art.....	16
2.1. Διαδικτυακές Εφαρμογές.....	16
2.1.1 Τεχνολογίες Ιστού και Javascript.....	16
2.1.2 Vuejs Framework.....	20
2.1.3 Http requests, Ajax και βιβλιοθήκη axios.....	24
2.1.4 Java, Http server και Vertx Framework.....	27
2.2. Σημασιολογικός Ιστός και Οντολογίες.....	31
2.2.1 Σημασιολογικός Ιστός.....	32
2.2.2 Οντολογίες.....	36
2.2.3 Δομή της OWL και βιβλιοθήκη Owlapi.....	38
2.3. Καθορισμός Συσχετίσεων μεταξύ Οντολογιών.....	40
3. Περιγραφή Συστήματος.....	43
4. Υλοποίηση Συστήματος και Παραδείγματα.....	44
5. Συμπεράσματα και μελλοντική εξέλιξη.....	45
6. Σύνοψη.....	46
7. Βιβλιογραφικές Αναφορές.....	47

Ευρετήριο Σχημάτων

Figure 1: δενδρική δομή HTML DOM [3].....	17
Figure 2: Event loop of firefox browser [5].....	18
Figure 3: nodejs event loop [6].....	19
Figure 4: Promise states [7].....	20
Figure 5: Reactivity στην VueJs [8].....	23
Figure 6: Vue Instance Life Cycle [9].....	24
Figure 7: Actor model [12].....	28
Figure 8: Vert.x event loop [14].....	30
Figure 9: Vert.x verticles [14].....	31
Figure 10: Vert.x event bus [14].....	31
Figure 11: Αρχιτεκτονική Σημασιολογικού Ιστού.....	33
Figure 12: OWL sublanguages.....	39
Figure 13: Πεδία ενός κανόνα συσχέτισης [1].....	42

Ευρετήριο Πινάκων

Table 1: Συστατικά των οντολογιών.....	37
Table 2: RDF Στοιχεία Σύνταξης βασισμένα στην XML.....	38

1. Εισαγωγή

Στις μέρες μας η αποθήκευση και η επεξεργασία της κλινικής και φαρμακευτικής γνώσης αποτελεί ζήτημα με μεγάλο ενδιαφέρον. Η γνώση αυτή αποκτάται από διαφορετικούς οργανισμούς (κλινικές, νοσοκομεία, φαρμακεία, κτλ) και αποθηκεύεται σε διαφορετικά σχήματα και δομές όπως απλά αρχεία ή διάφορες βάσεις δεδομένων. Ωστόσο πέρα από την αποθηκευμένη πληροφορία αυτή καθ' αυτή, μεγάλη σημασία έχει να καθορίσουμε τους όρους που χρησιμοποιούνται για να περιγράψουν αυτή την πληροφορία. Για αυτόν τον λόγο χρησιμοποιούμε Οντολογίες οι οποίες θα εξηγηθούν αναλυτικά στα επόμενα κεφάλαια.

Στην περίπτωση της διπλωματικής μας μελετάμε τέτοια δεδομένα που αφορούν φάρμακα και ασθένειες σε μορφή Οντολογιών που είναι αποθηκευμένα σε αρχεία γλώσσας OWL (Web Ontology Language). Αυτά τα αρχεία περιέχουν πολλές ίδιες έννοιες (π.χ. κάποιο συγκεκριμένο φάρμακο) οι οποίες όμως απεικονίζονται με διαφορετικά μοντέλα σε κάθε περίπτωση. Ο καθορισμός των συσχετίσεων μεταξύ τέτοιων διαφορετικών μοντέλων έχει ιδιαίτερη σημασία, καθώς μπορεί να μας οδηγήσει στην λύση διάφορων προβλημάτων. Τέτοια προβλήματα είναι η ενοποίηση δεδομένων και η απάντηση ερωτημάτων παρά την όποια διαφορετική φύση των μοντέλων μας.

Για τον καθορισμό των συσχετίσεων μεταξύ των οντολογιών υπάρχει αρκετή δουλειά στην βιβλιογραφία όσον αφορά τον εντοπισμό των συσχετίσεων μεταξύ των όρων των οντολογιών αλλά και την έκφραση των κανόνων σε μια μορφή που είναι κατανοητή από τον υπολογιστή. Στην παρούσα διπλωματική εργασία βασιζόμαστε στο εργαλείο *Ontology Alignment Tool* [1] το οποίο πραγματοποιεί αυτοματοποιημένη παραγωγή κανόνων συσχετισμών μεταξύ οντολογιών. Το εργαλείο αυτό θα αναλυθεί στα επόμενα κεφάλαια και θα εξηγηθεί ο τρόπος που προκύπτουν οι κανόνες καθώς και η μορφή τους.

Σκοπός της εργασίας μας είναι η οπτική αναπαράσταση αυτών των κανόνων καθώς και η αναπαράσταση των βασικών όρων των Οντολογιών με τρόπο κατανοητό. Για αυτόν τον σκοπό θα δημιουργήσουμε μια διαδικτυακή εφαρμογή η οποία θα επεξεργάζεται δύο Οντολογίες *owl* καθώς και τους κανόνες συσχέτισης τους οι οποίοι έχουν παραχθεί με το παραπάνω εργαλείο. Το αποτέλεσμα θα είναι σχηματικές αναπαραστάσεις των παραπάνω με δυνατότητα αλληλεπίδρασης του χρήστη για την πλήρη και εύκολη κατανόηση αυτών των συσχετίσεων. Αυτό θα μπορεί να οδηγήσει στην εύρεση λαθών του εργαλείου όπως για παράδειγμα η απουσία κανόνων που θα έπρεπε να υπάρχουν.

Στο κεφάλαιο 2 αναλύουμε τις βασικές θεωρητικές έννοιες που χρειάζονται για την κατανόηση της εργασίας μας καθώς και τις βασικές τεχνολογίες στις οποίες θα βασιστεί η εφαρμογή μας.

2. State of the Art

Εδώ θα αναλύσουμε τις βασικές έννοιες που χρησιμοποιούνται στην μελέτη μας καθώς και τις κύριες τεχνολογίες στις οποίες θα βασιστεί η εφαρμογή μας.

2.1. Διαδικτυακές Εφαρμογές

Θα εξηγήσουμε τις βασικές αρχές στις οποίες κατασκευάζονται οι διαδικτυακές εφαρμογές και θα αναλύσουμε σε λειτουργικό βαθμό τις τεχνολογίες που θα χρησιμοποιήσουμε.

2.1.1 Τεχνολογίες Ιστού και Javascript

Μια ιστοσελίδα διαδικτύου (*web page*) αποτελείται από αρχεία τα οποία έχουν μια συγκεκριμένη δομή ώστε να μπορούν οι περιηγητές ιστού (*web browsers*) να τα διαβάζουν για να τα αναπαραστήσουν στην οθόνη. Η λειτουργία αυτών των αρχείων περιγράφεται από την προδιαγραφή *html* (*html specification*) [2] . Αυτή η προδιαγραφή ορίζει μια γενικευμένη γλώσσα για την περιγραφή των εγγράφων και των εφαρμογών, καθώς και κάποιες προγραμματιστικές διεπαφές (*api*) για την αλληλεπίδραση με την αναπαράσταση στην μνήμη των πόρων που χρησιμοποιεί αυτή η γλώσσα.

Γενικά ορίζονται δύο βασικές συντάξεις με τις οποίες μπορούν να γραφούν τέτοια αρχεία. Η πρώτη είναι η *HTML* (*HyperText Markup Language*) και η δεύτερη είναι η *XML* (*eXtensible Markup Language*). Δεν θα μπούμε σε λεπτομέρειες περιγραφής τους αλλά και οι 2 γλώσσες που χρησιμοποιούνται για την σύνταξη τέτοιων αρχείων μπορούν να διαβαστούν από όλους τους σύγχρονους περιηγητές.

Η αναπαράσταση στην μνήμη των αντικειμένων που χρησιμοποιούνται στην ιστοσελίδα ή εφαρμογή ιστού μαζί με την προγραμματιστική διεπαφή που ορίζεται για να ελέγχουμε την κατάσταση αυτών των αντικειμένων είναι γνωστή ως *HTML DOM* (*Document Object Model*).

Καταλαβαίνουμε ότι το DOM είναι δύο έννοιες:

- Είναι αρχικά ένα μοντέλο αντικειμένων για ένα έγγραφο html. Για κάθε σελίδα που φορτώνει ο περιηγητής φτιάχνει ένα αντικείμενο DOM. Το αντικείμενο αυτό αποτελείται από όλα τα *html στοιχεία (html elements)* οργανωμένα σε δενδρική δομή. Για κάθε τέτοιο στοιχείο υπάρχει η πληροφορία για τις *ιδιότητες (properties)*, τις *μεθόδους (methods)* και τα *γεγονότα (events)* που σχετίζονται με αυτό το στοιχείο.
- Είναι επίσης μια προγραμματιστική διεπαφή μέσω της οποίας μπορούμε να επηρεάσουμε και να ελέγξουμε την κατάσταση αυτών των αντικειμένων. Η γλώσσα προγραμματισμού που χρησιμοποιείται από όλους τους περιηγητές για αυτή την λειτουργία είναι η *Javascript*.

Εναποθέτουμε ένα παράδειγμα της δενδρικής δομής των στοιχείων ενός html αντικειμένου:

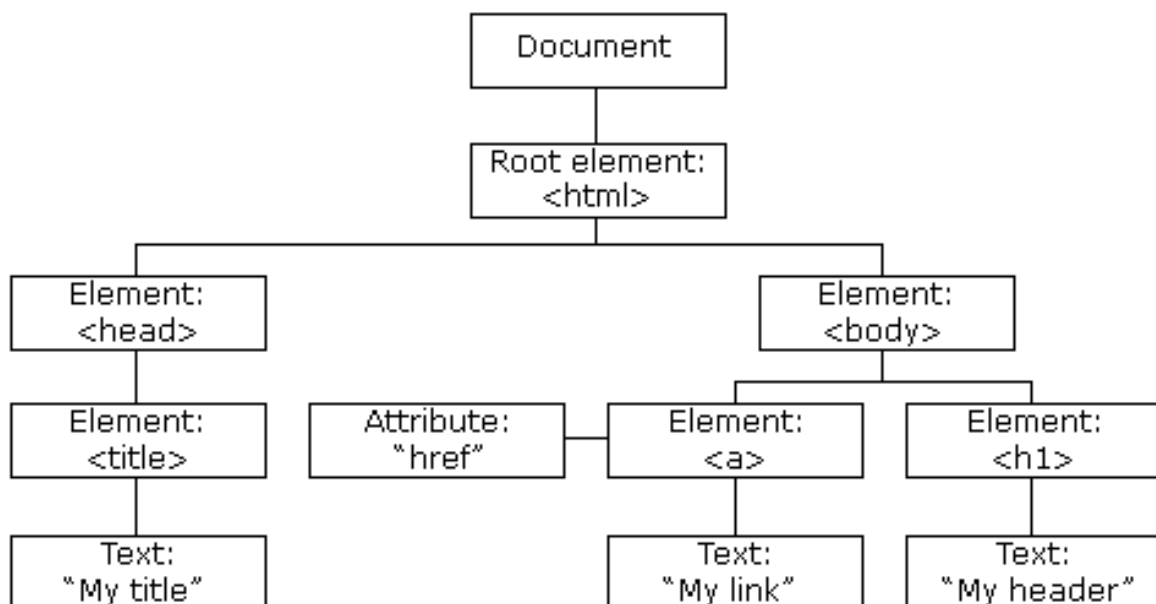


Figure 1: δενδρική δομή HTML DOM [3]

Αναφέραμε ότι η javascript είναι η κύρια γλώσσα επεξεργασίας του DOM. Ο κάθε περιηγητής έχει την δικιά του υλοποίηση javascript για να υποστηρίξει τις απαραίτητες λειτουργίες. Όλες οι υλοποιήσεις τηρούν τις προδιαγραφές για την γλώσσα όπως ορίζονται στο *ECMAScript specification*, τελευταία έκδοση του

οποίου είναι αυτή του 2020 [4]. Ο κώδικας javascript που συμπεριλαμβάνεται από κάποιο αρχείο html εκτελείται από τον περιηγητή κατά την φόρτωση της ιστοσελίδας. Υπάρχουν αρκετές διαφορές από υλοποίηση σε υλοποίηση αλλά για την κατανόηση των επόμενων μας ενδιαφέρουν να εξηγήσουμε κάποιες συμπεριφορές που ορίζονται στο πρότυπο και είναι ίδιες σε κάθε υλοποίηση.

Για διάφορους λόγους η javascript εκτελείται σε μόνο έναν νήμα επεξεργασίας. Αυτό σημαίνει ότι δεν υπάρχει δυνατότητα δημιουργίας πολλαπλών νημάτων από τον επεξεργαστή ώστε να μπορούν να εκτελεστούν πολλές λειτουργίες ταυτόχρονα. Για αυτό τον λόγο η javascript λειτουργεί με ένα μοντέλο ταυτοχρονισμού γνωστό ως *βρόγχος γεγονότων (event loop)*. Οι λεπτομέρειες διαφέρουν ανάλογα με την *μηχανή javascript (javascript engine)* του κάθε περιηγητή, αλλά η λογική είναι ότι υπάρχει μια ουρά μηνυμάτων ή γεγονότων καθένα με τα οποία συνδέονται με μια *συνάρτηση (function)*. Αυτά εκτελούνται κάθε φορά με την σειρά καθώς νέα γεγονότα προστίθενται στην ουρά.

Για καλύτερη κατανόηση δείχνουμε μια γενικευμένη εικόνα επεξήγησης του event loop από την javascript engine του περιηγητή mozilla firefox:

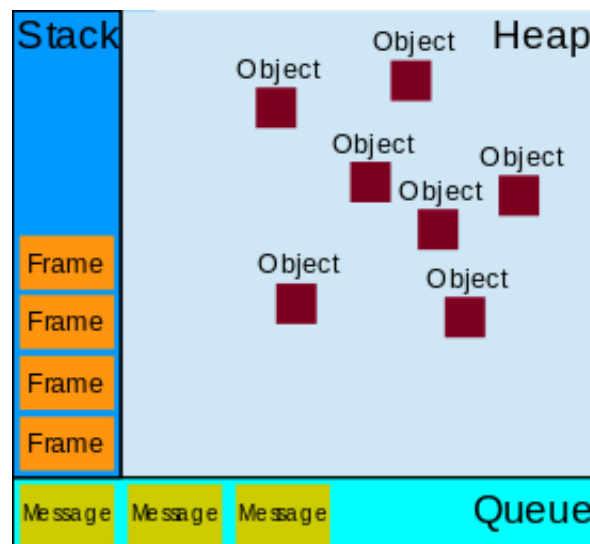


Figure 2: Event loop of firefox browser [5]

καθώς και μια εικόνα των φάσεων που περνάει το event loop του nodejs όπου είναι ένα περιβάλλον εκτέλεσης javascript (*javascript runtime*) βασισμένο στο javascript engine του περιηγητή chrome:

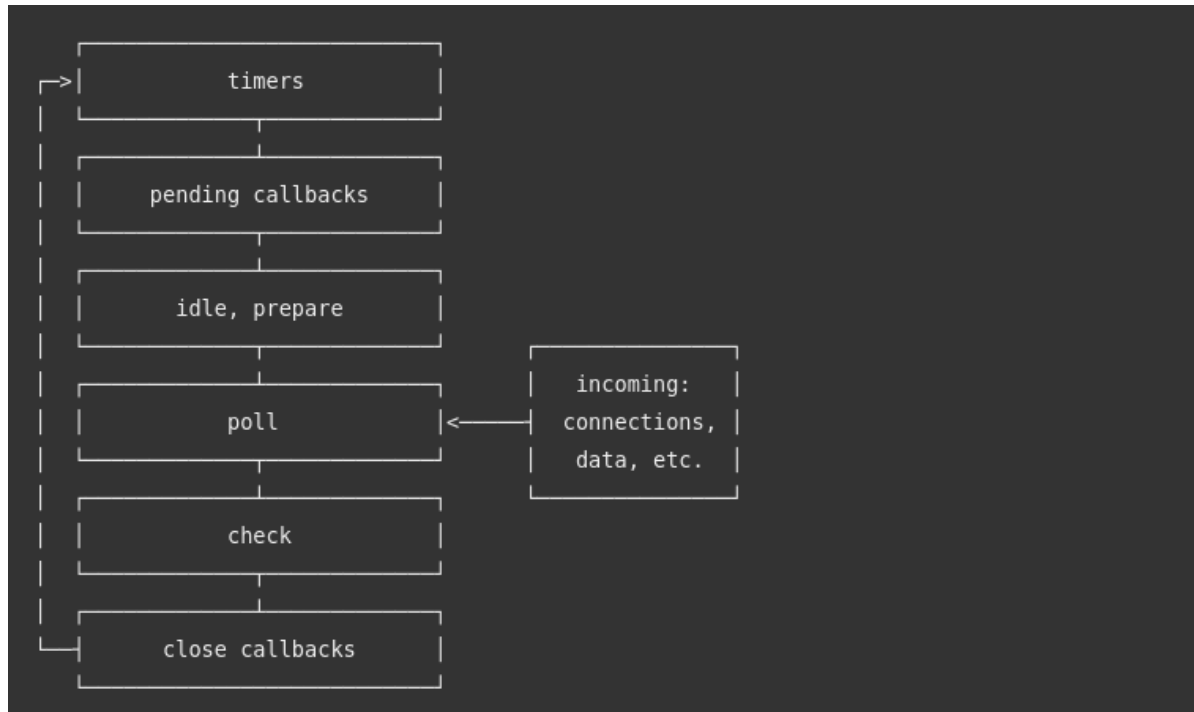


Figure 3: nodejs event loop [6]

Καταλαβαίνουμε ότι αν η συνάρτηση εκτέλεσης κάποιου γεγονότος είναι μεγάλης διάρκειας τότε όλη η ουρά μπλοκάρει και δεν εκτελείται κάποιο άλλο γεγονός μέχρι να τελειώσει το πρώτο. Για αυτό τον λόγο προσπαθούμε να φτιάχνουμε συναρτήσεις με τρόπο ασύγχρονο. Δηλαδή αρχικά τις καταμερίζουμε σε μικρότερα μέρη. Όταν καλούμε κάποια συνάρτηση, τότε αφού τελειώσει το μέρος της που εκτελούνταν, αυτή επιστρέφει τον έλεγχο ώστε να μπορέσει να εκτελεστεί κάποια άλλη συνάρτηση από την ουρά του event loop. Καταλαβαίνουμε ότι επειδή ο έλεγχος επιστρέφει μετά από την εκτέλεση ενός μικρού μέρους της αρχικής συνάρτησης που θέλαμε να υλοποιήσουμε και όχι αφού εκτελεστεί ολόκληρη, θέλουμε κάποιον τρόπο για να καταλαβαίνουμε πότε έχει ολοκληρωθεί όλη η εργασία που θέλαμε να κάνουμε. Στην σύγχρονη javascript η υλοποίηση των ασύγχρονων συναρτήσεων γίνονται μέσω του μηχανισμού των *υποσχέσεων*

(Promises). Ένα promise χρησιμοποιείται για να περιμένουμε κάποια τιμή που ακόμα δεν είναι διαθέσιμη. Μπορεί να έχει 3 καταστάσεις:

- *εκκρεμής (pending)* όταν δεν έχει ολοκληρωθεί ακόμα
- *ολοκληρωμένη (fulfilled)* όταν η εκτέλεση ολοκληρώθηκε και η τιμή είναι διαθέσιμη
- *απορριφθείσα (rejected)* όταν η λειτουργία υπολογισμού της τιμής απέτυχε

Έτσι μπορούμε να προγραμματίσουμε τι θα γίνει σε κάθε περίπτωση και να εκτελούμε ταυτόχρονα υπολογισμούς χωρίς να μπλοκάρουμε το event loop. Παραθέτουμε ένα διάγραμμα των καταστάσεων ενός promise σε προγραμματιστικό επίπεδο:

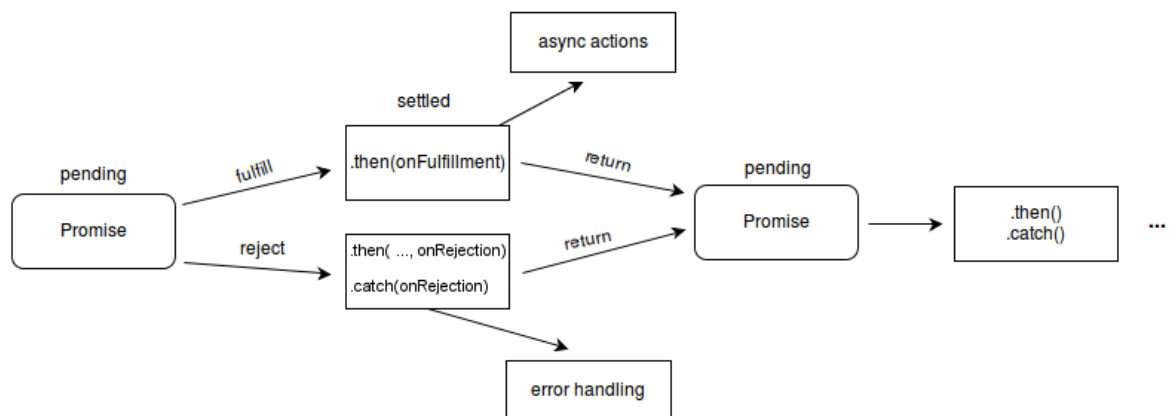


Figure 4: Promise states [7]

2.1.2 Vuejs Framework

Έχουμε εξηγήσει τις βασικές λειτουργίες των διαδικτυακών εφαρμογών και του DOM. Ωστόσο, δεδομένου ότι η βασικές λειτουργίες του σχεδιάστηκαν αρκετά παλιά, παρά τις βελτιώσεις του τα τελευταία χρόνια υπάρχουν ορισμένα μειονεκτήματα. Κύριο μειονέκτημα είναι η κακή απόδοσή του. Έχουμε αναφέρει ότι το DOM κρατάει μια δενδρική δομή των αντικειμένων html. Κάθε φορά που θέλουμε να αλλάξουμε κάποιο html element τότε πρέπει να βρεθεί αυτό το στοιχείο από το δέντρο, να επεξεργαστεί και μετά ο περιηγητής να το απεικονίσει στην οθόνη

(render). Όταν αλλάζουμε πολλά στοιχεία προγραμματιστικά με javascript υπάρχει μεγάλο υπολογιστικό κόστος. Το δεύτερο κύριο μειονέκτημα είναι ότι η προγραμματιστική διεπαφή που ορίζεται για την επεξεργασία του DOM δεν είναι πολύ εύχρηστη. Υπάρχουν βιβλιοθήκες όπως η γνωστή jQuery όπου προσφέρουν λίγο πιο εύχρηστες προγραμματιστικές διεπαφές, ωστόσο αν και έχουν βρει μεγάλη επιτυχία στην δημιουργία απλών ιστοσελίδων όταν πρόκειται για πιο σύνθετες εφαρμογές υστερούν σημαντικά. Για αυτόν τον σκοπό έχουν δημιουργηθεί διάφορα πιο σύνθετα frameworks που προσφέρουν αυξημένες δυνατότητες στους προγραμματιστές.

Ιστορικά τα παλιότερα χρόνια είχαν επικρατήσει τα *frameworks* από την μεριά του διακομιστή (*server side frameworks*). Η λογική είναι ότι οι ιστοσελίδες προσφέρονται από έναν υπολογιστή που έχει ρόλο server. Ωστόσο, πριν οι σελίδες σταλούν στο περιηγητές έχουν επεξεργαστεί με διάφορες *template engines* όπου έχουν δημιουργήσει από πριν το τελικό DOM. Έτσι ελαχιστοποιείται η χρήση javascript και η επεξεργασία του DOM από τους περιηγητές. Αν και αυτή η μέθοδος προτιμάται ακόμα σε πολλές περιπτώσεις, το κύριο μειονέκτημά της είναι η συνεχής εξάρτηση από τον server και οι συνεχείς κλήσεις για καινούριες σελίδες κάθε φορά που υπάρχει ανάγκη για επεξεργασία του DOM.

Ως εναλλακτική λύση έχουν δημιουργηθεί τα *frameworks* από την μεριά του πελάτη (*client side frameworks*). Αυτά προσφέρουν εξελιγμένες δυνατότητες επεξεργασίας του DOM με χρήση javascript χωρίς να χρειάζεται κλήση σε κάποιον server για κάθε σελίδα. Έτσι ο server μπορεί να χρησιμοποιείται αποκλειστικά για να προσφέρει δεδομένα όταν αυτό χρειάζεται χωρίς να έχει κάποιο ρόλο στην επεξεργασία των σελίδων και του DOM. Η διαδικτυακή εφαρμογή χρειάζεται τότε απλά να γίνει διαθέσιμη ως στατικές σελίδες όπου ο περιηγητής μπορεί να κατεβάσει. Η δυνατότητα αυτή της προσφοράς στατικών σελίδων υπάρχει σε πολλούς μικρούς και οικονομικούς από άποψης πόρων servers, αλλά επίσης προσφέρεται πάντα από όλους τους γνωστούς servers που συνήθως χρησιμοποιούνται για δυναμικές σελίδες σε συνδυασμό με κάποια γλώσσα για server side επεξεργασία (π.χ. Apache

ή nginx με γλώσσα php, Tomcat με γλώσσα java, κτλ.). Έτσι υπάρχει ευελιξία στον τρόπο με τον οποίο θα γίνει διαθέσιμη η εφαρμογή στους χρήστες.

Ένα τέτοιο client side javascript framework είναι το VueJs το οποίο θα αναλύσουμε εδώ. Πρόκειται καταρχάς για ένα, όπως αποκαλείται, *προοδευτικό framework* (*progressive framework*). Η έννοια αυτή σημαίνει ότι δεν είναι απαραίτητο να φτιαχτεί μια ολόκληρη ιστοσελίδα ή διαδικτυακή εφαρμογή βάση αυτού, αλλά μπορεί να χρησιμοποιηθεί μόνο για ένα συγκεκριμένο μέρος της εφαρμογής όπου χρειάζονται αυξημένες δυνατότητες επεξεργασίας όταν αυτό κρίνεται απαραίτητο. Η βασική ιδέα που διέπει την VueJs είναι αυτή του *εικονικού DOM* (*Virtual DOM*). Ο προγραμματιστής δεν ασχολείται με το να επεξεργαστεί απευθείας το DOM, αν και εξακολουθεί να υπάρχει αυτή η δυνατότητα. Αντιθέτως, ορίζονται μεταβλητές σε javascript οι οποίες δένονται με συγκεκριμένα elements του DOM (*data binding*). Λέμε ότι τα δεδομένα αυτά που ορίζονται με αυτόν τον τρόπο είναι *αντιδραστικά* (*reactive*). Αυτό σημαίνει ότι το framework παρακολουθεί αυτά τα δεδομένα για αλλαγές. Όταν αυτά αλλάξουν τότε αυτόματα ειδοποιείται η συνάρτηση απεικόνισης (*rendering*) και δημιουργείται το καινούριο πραγματικό DOM με τα νέα δεδομένα. Οι δυνατότητές του είναι αρκετά εξελιγμένες ώστε να γίνονται *rendered* μόνο οι περιοχές όπου άλλαξαν εξοικονομώντας έτσι επεξεργαστικούς πόρους. Επίσης, όταν αλλάζουν πολλές τέτοιες μεταβλητές ταυτόχρονα αλλά στην ίδια φάση λειτουργίας, θα γίνει μόνο μια φορά *render* μετά από όλες τις αλλαγές κάνοντας σαφές το πλεονέκτημα στην απόδοση σε σχέση με την απευθείας επεξεργασία του DOM. Παραθέτουμε μια εικόνα για την καλύτερη κατανόηση της παρακολούθησης αυτών των δεδομένων και του *rendering*:

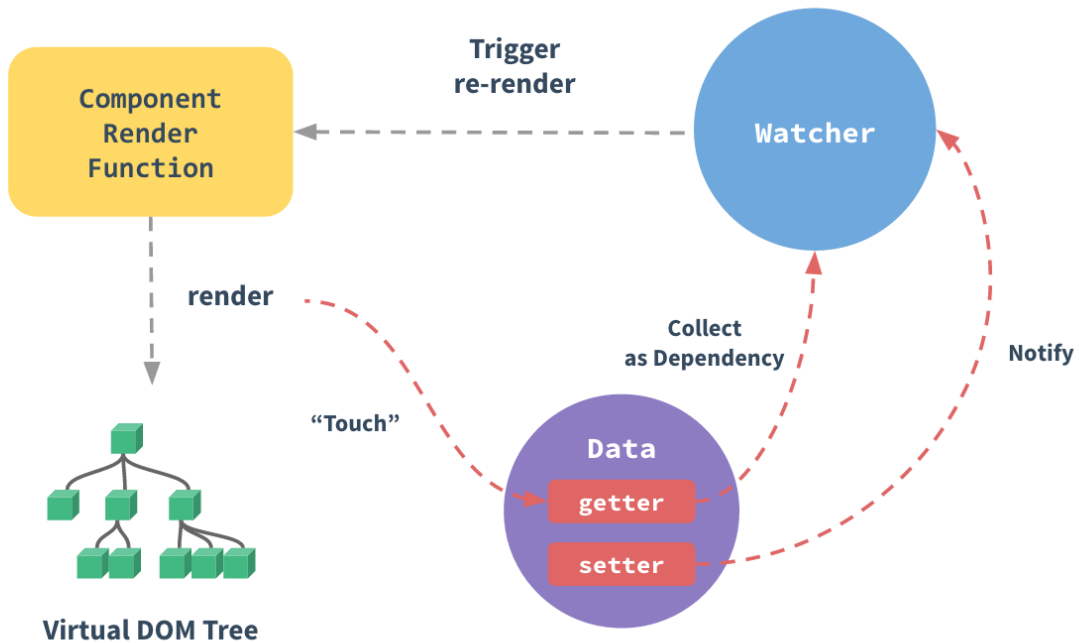


Figure 5: Reactivity στην VueJs [8]

Μια VueJs εφαρμογή οργανώνεται σε ξεχωριστά δομικά στοιχεία (*components*). Κάθε τέτοιο component έχει τον δικό του κώδικα html, την δικό του κώδικα javascript και την δικιά του css για την επεξεργασία της εμφάνισής του. Το αρχικό component είναι το κύριο και κάθε επόμενο προστίθεται σε κάποιο ήδη υπάρχων με σχέση πατέρα παιδιού. Ο πατέρας μπορεί να επικοινωνεί με το παιδί περνώντας του *ιδιότητες (properties)*, ενώ το παιδί επικοινωνεί με τον πατέρα στέλνοντάς του *γεγονότα (events)*. Κάθε component που δημιουργείται περνάει από διάφορα στάδια ενός κύκλου ζωής όπως φαίνεται στο παρακάτω σχήμα:

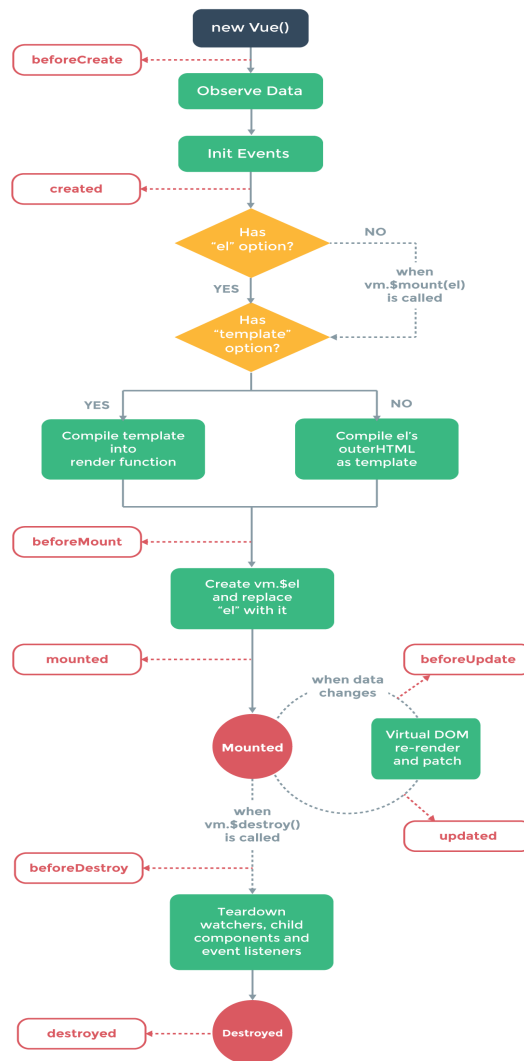


Figure 6: Vue Instance Life Cycle [9]

2.1.3 Http requests, Ajax και βιβλιοθήκη axios

Έχουμε εξηγήσει ότι με τα client side frameworks δεν χρειάζεται να υπάρχει κάποιος server που να επεξεργάζεται τις διαδικτυακές σελίδες με νέα δεδομένα. Ωστόσο, σε πολλές εφαρμογές, όπως θα δούμε και στην δική μας, χρησιμοποιείται ένας server για να επεξεργαστεί και να μας στείλει δεδομένα. Καταλαβαίνουμε ότι θέλουμε έναν τρόπο να μπορούμε να αλληλεπιδράσουμε με έναν server από την μεριά του client αφού έχει φορτωθεί μια σελίδα.

Τα αιτήματα που μας ενδιαφέρουν στην περίπτωση μας είναι τα *http requests*. Το *http* (*HyperText Transfer Protocol*) είναι ένα πρωτόκολλο δικτύου. Χρησιμοποιείται συνήθως από εφαρμογές για να στέλνουμε ή να λαμβάνουμε δεδομένα από έναν server. Δεν θα αναλύσουμε λεπτομέρειες καθώς την *προδιαγραφή* του (*specification*) μπορεί κανείς να την βρει στο διαδίκτυο [10]. Επιγραφικά αναφέρουμε ότι υποστηρίζει διάφορους μεθόδους ανάλογα με το τι θέλουμε να επιτύχουμε όπως οι μέθοδοι:

- GET
- HEAD
- POST
- PUT
- DELETE
- TRACE
- CONNECT
- OPTIONS

Η καλύτερη μέθοδος για να στείλουμε http αιτήματα όπως τα περιγράψαμε παραπάνω ονομάζεται *AJAX* (*Asynchronous JavaScript and XML*). Δεν πρόκειται για κάποια συγκεκριμένη υλοποίηση αλλά για μια τεχνική. Ο όρος XML περιλαμβάνεται για ιστορικούς λόγους αλλά δεν η ανάγκη η χρήση της XML. Με αυτή την τεχνική μπορούμε να στέλνουμε http αιτήματα σε κάποιον server ασύγχρονα χωρίς να μπλοκάρουμε το event loop (έχουμε εξηγήσει τη σημασία αυτού στις προηγούμενες ενότητες). Όταν το αίτημα ολοκληρωθεί καλείται κάποιος κώδικας που έχουμε ορίσει ανάλογα με το αν ήταν επιτυχές ή αν απέτυχε αυτό το αίτημα. Οι υλοποιήσεις javascript των περιηγητών διαθέτουν υποστήριξη για αυτή την λειτουργία μέσω του αντικειμένου *XMLHttpRequest* (*XMLHttpRequest object*).

Δεν θα αναλύσουμε τις λεπτομέρειες αυτού καθώς πρόκειται για μια συγκεκριμένη υλοποίηση της παραπάνω τεχνικής.

Εμάς μας ενδιαφέρει η βιβλιοθήκη *Axios*. Πρόκειται για μια βιβλιοθήκη javascript η οποία έχει δημιουργηθεί για την δημιουργία τέτοιων ασύγχρονων αιτημάτων.

Υποστηρίζει πολλές διαφορετικές μεθόδους http και είναι αρκετά ευέλικτο ως προς την παραμετροποίηση αυτών των αιτημάτων. Ο μηχανισμός με τον οποίο πετυχαίνει τα ασύγχρονα αιτήματα χωρίς να μπλοκάρει το event loop είναι αυτός των *Promises* τον οποίο έχουμε εξηγήσει σε προηγούμενη ενότητα. Αυτή η χρήση τους είναι πολύ σημαντική αφού επιτρέπει την εύκολη συνεργασία με άλλες βιβλιοθήκες όπου χρησιμοποιούν *Promises* ή και ακόμα να φτιάχνουμε δικές μας αφαιρετικές συναρτήσεις (abstractions) όπου θα χρησιμοποιούν αυτές την *axios* και θα επιστρέφουν το αντίστοιχο *Promise*, κρύβοντας τις λεπτομέρειες από το υπόλοιπο πρόγραμμά μας.

Τα δεδομένα που θα λαμβάνουμε από τον server θα είναι σε μορφή JSON αντί για XML. Το *JSON (JavaScript Object Notation)* είναι μια μορφοποίηση για ανταλλαγή δεδομένων. Είναι επίσης ένας τύπος αρχείου όπου αποθηκεύονται δεδομένα με την συγκεκριμένη μορφή. Έχει πολλά πλεονεκτήματα σε σχέση με την XML. Τα κυριότερα είναι

- Είναι πιο ευανάγνωστο και μπορεί να διαβαστεί ευκολότερα από τον άνθρωπο.
- Προσθέτει λιγότερους χαρακτήρες για την μορφοποίηση αφού βασίζεται κυρίως στην χρήση αγκύλων. Άρα τα αρχεία έχουν μικρότερο μέγεθος.
- Είναι πιο αποδοτικό στην ανάλυσή του και την επεξεργασία του από υπολογιστές με την χρήση διάφορων βιβλιοθηκών που υπάρχουν.

Το πρότυπο που ορίζεται είναι ανοιχτό και μπορεί κανείς να βρει περισσότερα για την χρήση του στο διαδίκτυο [11].

2.1.4 Java, Http server και Vertx Framework

Έχουμε αναφέρει ότι πολλές διαδικτυακές εφαρμογές χρησιμοποιούν κάποιον server είτε για να προσφέρει ιστοσελίδες είτε για να δέχεται και να στέλνει δεδομένα. Εμείς θα αναφερθούμε στους http servers όπου υλοποιούνται με την γλώσσα προγραμματισμού java και χρησιμοποιούνται από τους clients για να λαμβάνουν και να στέλνουν δεδομένα με τους τρόπους που εξηγήσαμε παραπάνω.

Η java είναι μια γλώσσα προγραμματισμού που έχει σχεδιαστεί κυρίως για να τρέχει σε μια εικονική μηχανή την λεγόμενη *JVM (Java Virtual Machine)*. Ο κώδικας πρώτα μεταγλωτίζεται (compiling) σε αρχεία που περιέχουν java bytecode η οποία είναι μια δυαδική μορφή αναπαράστασης του κώδικα αναγνωρίσιμη από το JVM για να μπορεί να εκτελεστεί. Λόγω της ιστορίας της υπάρχουν πληθώρα από βιβλιοθήκες και πολλές υλοποιήσεις διαφορετικών http servers. Υπάρχουν επίσης πολλές υλοποιήσεις με αρκετές διαφορές τέτοιων εικονικών μηχανών που αναγνωρίζουν java bytecode. Κύριο σημείο ενδιαφέροντος είναι ότι εδώ δεν έχουμε τον περιορισμό του μοναδικού thread ενός προγράμματος που υπάρχει στις υλοποιήσεις javascript των περιηγητών. Άρα ένα πρόγραμμα μπορεί να χρησιμοποιεί πολλά threads για να πετύχει τον σκοπό του.

Ένας http server πρέπει να μπορεί να δέχεται ένα http αίτημα με κάποια από τις http μεθόδους που έχουμε αναφέρει παραπάνω. Κάθε αίτημα θα δέχεται κάποια δεδομένα, θα εκτελεί κάποια συγκεκριμένη εργασία και θα επιστρέφει κάποιο αποτέλεσμα. Στόχος μας είναι να μπορούμε να εξυπηρετήσουμε πολλά τέτοια αιτήματα ταυτόχρονα. Κάθε αίτημα θα πρέπει να εκτελείται απομονωμένα χωρίς τα ενδιάμεσα αποτελέσματα κάποιας τρέχουσας εργασίας ενός αιτήματος να επηρεάζει τα αποτελέσματα κάποιας άλλης. Θα αναφέρουμε περιληπτικά 3 γνωστά μοντέλα με τα οποία πετυχαίνουμε αυτόν τον ταυτοχρονισμό (concurrency).

- Ο πιο ιστορικός τρόπος να πετύχουμε τον ταυτοχρονισμό είναι με *νήματα (threads)*. Ένα νήμα είναι απομονωμένο και έχει τον δικό χώρο μνήμης για μεταβλητές αν και μπορεί να διαμοιράζεται πόρους με διάφορους τρόπους. Με αυτό το μοντέλο επιλέγουμε να δημιουργούμε ένα καινούριο νήμα για κάθε καινούριο αίτημα που καλούμε να απαντήσουμε. Έτσι κάθε εργασία

που πρέπει να εκτελεστεί είναι απομονωμένη από τις υπόλοιπες.

Αναφέρουμε για παράδειγμα ότι ο Tomcat server χρησιμοποιεί αυτό το μοντέλο, αν και υπάρχουν αρκετοί τρόποι παραμετροποίησης του.

Μειονέκτημα αυτού του μοντέλου είναι ότι ένας υπολογιστής μπορεί να εκτελεί τόσα παράλληλα νήματα όσα μπορεί να υποστηρίξει ο επεξεργαστής του. Σε κρίσιμους servers με μεγάλο αριθμό αιτημάτων, τα νήματα που δημιουργούνται είναι αρκετές φορές πολλαπλάσια των νημάτων του επεξεργαστή, επιφέροντας μεγάλη μείωση στην απόδοση.

- Ένα άλλο πιο πρόσφατο μοντέλο είναι το *μοντέλο των ηθοποιών (actor model)*. Σε αυτό μια εργασία γίνεται σε κάποιον actor. Κάθε actor είναι απομονωμένος έχοντας την δικιά του *ιδιωτική κατάσταση (private state)* που δεν μπορεί να αλλάξει από κάποιον άλλον. Ο κάθε actor έχει ένα *γραμματοκιβώτιο (mailbox)* και επικοινωνούν ανταλλάσσοντας μηνύματα μεταξύ τους. Οι υλοποιήσεις διαφέρουν από σύστημα σε σύστημα αλλά οι actors πάντα χρειάζονται λιγότερους πόρους από τα threads και μπορούν να υπάρχουν σε μεγάλο βαθμό ταυτόχρονα. Μια τέτοια υλοποίηση βασισμένη στο actor model είναι αυτή της βιβλιοθήκης Akka. Δείχνουμε μια εικόνα για καλύτερη κατανόηση αυτού του μοντέλου.

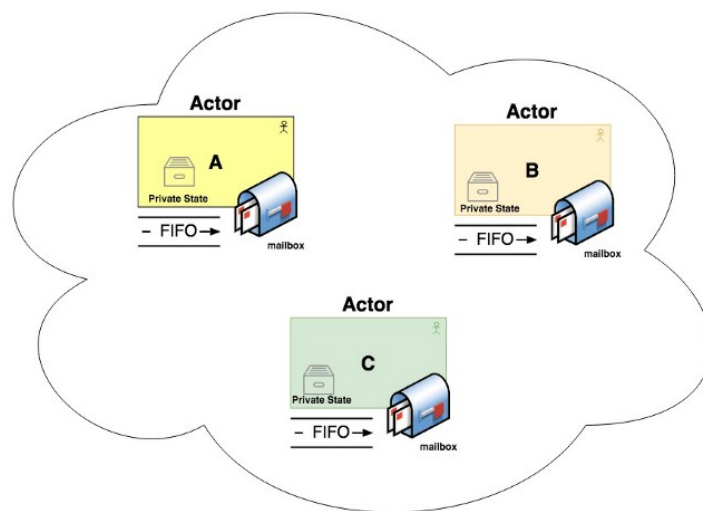


Figure 7: Actor model [12]

- Το τελευταίο μοντέλο που θα αναφέρουμε είναι αυτό του ασύγχρονου προγραμματισμού με κάποιο event loop. Το έχουμε ήδη εξηγήσει καθώς οι υλοποιήσεις javascript βασίζονται σε αυτό το μοντέλο. Πολλές φορές έχει την καλύτερη απόδοση από όλα τα μοντέλα. Ωστόσο, μειονέκτημά του είναι ότι δεν πρέπει ποτέ να μπλοκάρουμε το event loop με συνεχόμενες εργασίες που διαρκούν αρκετή ώρα (π.χ. επαναλήψεις). Έτσι, κάνει τον προγραμματισμό δυσκολότερο αφού πρέπει ο προγραμματιστής να μετατρέπει τις συναρτήσεις του σε ασύγχρονες καθώς και να χρησιμοποιεί όσο μπορεί βιβλιοθήκες με ασύγχρονη λογική.

Εμείς θα επικεντρωθούμε στο Vert.x framework [13]. Πρόκειται για ένα framework που είναι γραμμένο στην γλώσσα Polyglot, η οποία μπορεί να μεταφραστεί σε πολλές άλλες γλώσσες. Έτσι είναι εύκολα διαθέσιμο για όλες τις γλώσσες που μπορούν να τρέξουν στο JVM όπως Java, Scala, Kotlin, κτλ. Έχει εμπνευστεί από την NodeJs που είχαμε αναφέρει ως ένα περιβάλλον για εκτέλεση javascript. Βασίζεται στο μοντέλο του ασύγχρονου προγραμματισμού και μας προσφέρει την δυνατότητα να δημιουργούμε ασύγχρονες συναρτήσεις. Όπως όλες οι υλοποιήσεις αυτού του μοντέλου διαθέτει ένα event loop όπου εκτελεί τα γεγονότα που εισέρχονται σε μια ουρά:

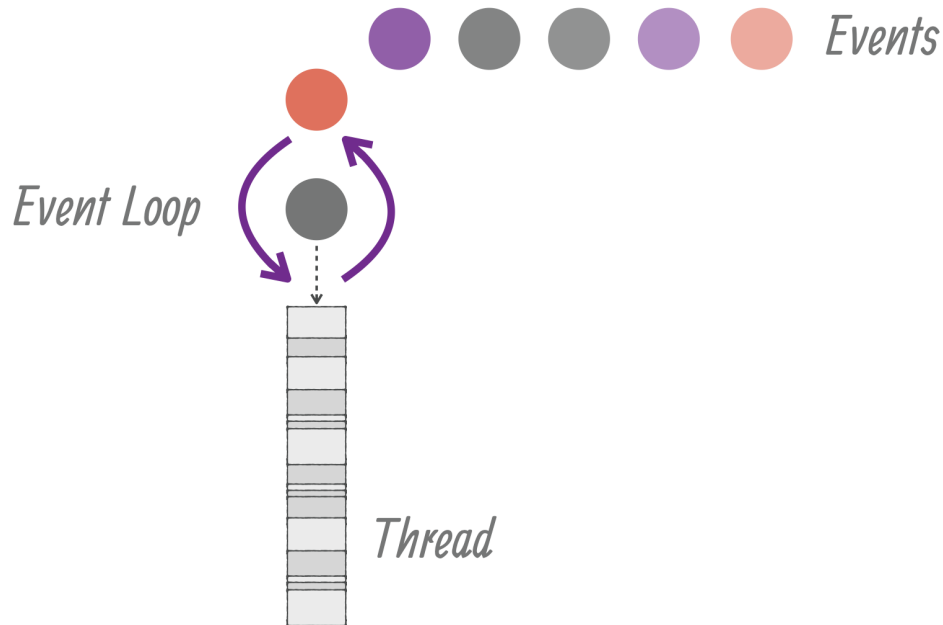


Figure 8: Vert.x event loop [14]

Ισχύουν όλα αυτά που έχουμε εξηγήσει σε προηγούμενες ενότητες για το event loop. Ο τρόπος για να περιμένουμε κάποιο ασύγχρονο αποτέλεσμα έχει την ίδια λογική, απλά οι ορολογίες είναι λίγο διαφορετικές καθώς χρησιμοποιούμε *Promises* και *Futures* για τα οποία δεν θα μπορούμε σε λεπτομέρειες αφού αφορούν την υλοποίηση της προγραμματιστικής διεπαφής.

Κύριο πλεονέκτημα αυτού του framework είναι ότι λόγω του ότι εκτελείται πάνω στο JVM είναι δυνατόν να χρησιμοποιήσει και πολλαπλά νήματα. Ο τρόπος που το πραγματοποιεί αυτό είναι με την έννοια του Verticle. Κάθε Verticle έχει το δικό του event loop και μπορεί να τρέχει σε ξεχωριστό νήμα αν υπάρχει διαθέσιμο στο σύστημα. Συνήθως υπάρχει ένα κύριο (main) Verticle το οποίο κάνει deploy τα υπόλοιπα.

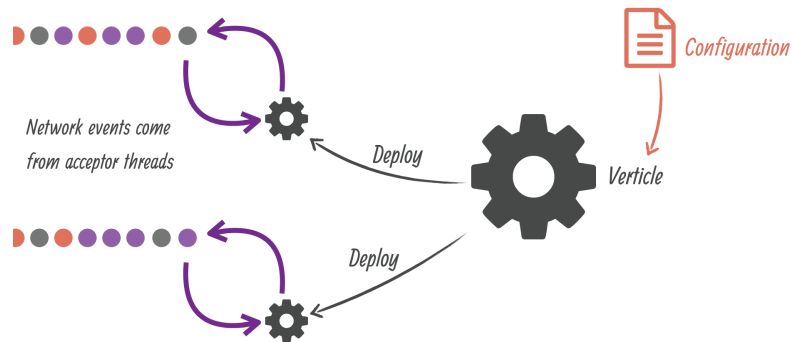


Figure 9: Vert.x verticles [14]

Τα verticles επικοινωνούν μεταξύ τους μέσω ανταλλαγής μηνυμάτων. Όλα τα μηνύματα περνούν από έναν διάδρομο γεγονότων (*event bus*) ο οποίος ελέγχεται από το κύριο Verticle. Ένα παράδειγμα φαίνεται στην παρακάτω εικόνα

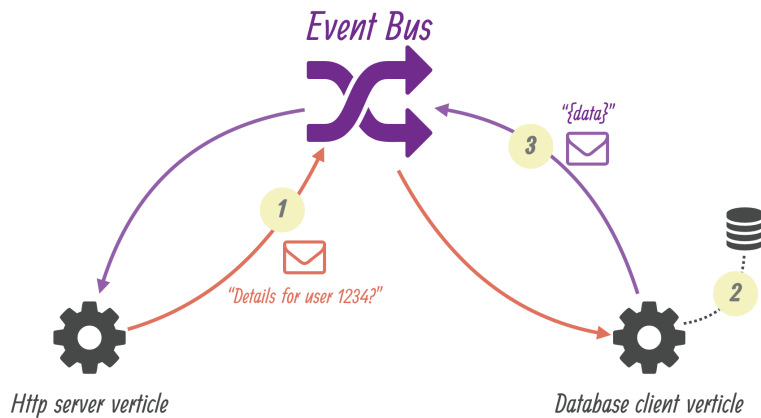


Figure 10: Vert.x event bus [14]

Το framework έχει ενσωματωμένη δυνατότητα δημιουργίας ενός http server όπου λειτουργεί με το ασύγχρονο μοντέλο βάση όλων των παραπάνω.

2.2. Σημασιολογικός Ιστός και Οντολογίες

Θα εξηγήσουμε τις βασικές έννοιες γύρω από τις Οντολογίες

2.2.1 Σημασιολογικός Ιστός

Ιστός των πραγμάτων (Web of things) ή *Ιστός των δεδομένων (Web of Data)*. Η αναζήτηση δεν βασίζεται σε αντιστοίχιση λέξεων αλλά σε σχετικά αντικείμενα και σχέσεις. Η επεξεργασία των δεδομένων δηλ. αναζήτηση και συμπερασμός (reasoning) γίνεται από μηχανές. Ο Σημασιολογικός Ιστός δεν είναι ένας ξεχωριστός Ιστός, αλλά μια επέκταση του τρέχοντος, στον οποίο οι πληροφορίες έχουν σαφώς καθορισμένη σημασία, επιτρέποντας στους υπολογιστές και τους ανθρώπους να συνεργάζονται καλύτερα. Τα πρώτα βήματα για την δόμηση του Σημασιολογικού Ιστού στη δομή του υπάρχοντος Ιστού βρίσκονται ήδη σε εξέλιξη. Στο εγγύς μέλλον, αυτές οι εξελίξεις θα οδηγήσουν σε σημαντική νέα λειτουργικότητα καθώς οι μηχανές γίνονται πολύ καλύτερα προετοιμασμένες να επεξεργαστούν και να «κατανοήσουν» τα δεδομένα που, προς το παρόν, απλώς εμφανίζουν [15] . Η κοινοπραξία WorldWide Web (W3C) είναι μία διεθνής κοινότητα που αναπτύσσει τα *ανοικτά πρότυπα (open standards)* προκειμένου για να εξασφαλίσει τη μακροπρόθεσμη ανάπτυξη του Παγκόσμιου Ιστού (Web).

Η αρχιτεκτονική του σημασιολογικού ιστού απεικονίζεται στο παρακάτω σχήμα. Το πρώτο επίπεδο, URI και Unicode, ακολουθεί τα σημαντικά χαρακτηριστικά του υπάρχοντος WWW. Το Unicode είναι ένα πρότυπο κωδικοποίησης διεθνών συνόλων χαρακτήρων και επιτρέπει τη χρήση όλων των γλωσσών (γραφής και ανάγνωσης) στον ιστό χρησιμοποιώντας μια τυποποιημένη φόρμα. Το *URI (Uniform Resource Identifier)* είναι μια συμβολοσειρά τυποποιημένης φόρμας που επιτρέπει τον μοναδικό προσδιορισμό πόρων (π.χ. έγγραφα). Ένα υποσύνολο του URI είναι το *URL (Uniform Resource Locator)*, το οποίο περιέχει το μηχανισμό πρόσβασης και μια θέση (δικτύου) ενός εγγράφου - όπως <http://www.example.org/>. Ένα άλλο υποσύνολο του URI είναι το URN που επιτρέπει τον εντοπισμό ενός πόρου χωρίς να υποδηλώνει τη θέση του και τα μέσα για την πρόσβαση στην τιμή ή το αντικείμενο - ένα παράδειγμα είναι το `urn:isbn:0-123-45678-9`. Η χρήση του URI είναι σημαντική για ένα καταμετρημένο σύστημα Διαδικτύου καθώς παρέχει κατανοητό προσδιορισμό όλων των πόρων. Μια διεθνής παραλλαγή του URI είναι το *IRI (Internationalized Resource Identifier)* που επιτρέπει τη χρήση χαρακτήρων

Unicode στο αναγνωριστικό (Identifier) και για την οποία ορίζεται μια αντιστοίχιση στο URI.

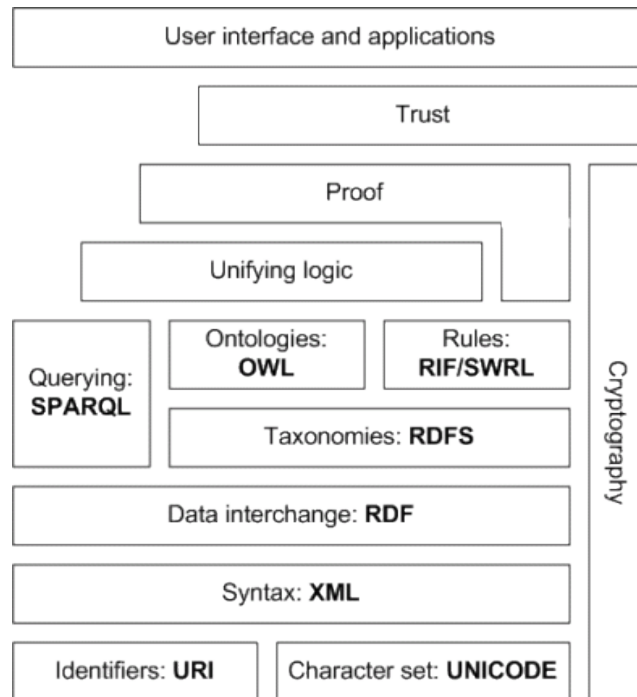


Figure 11: Αρχιτεκτονική Σημασιολογικού Ιστού

Το επίπεδο XML με τους ορισμούς *ονόματος χώρου XML (XML namespace)* και σχήματος XML (*XML schema*) διασφαλίζει ότι υπάρχει μια κοινή σύνταξη που χρησιμοποιείται στον σημασιολογικό ιστό. Το XML είναι μια γλώσσα σήμανσης (*markup*) γενικής χρήσης για έγγραφα που περιέχουν δομημένες πληροφορίες. Ένα έγγραφο XML περιέχει στοιχεία που μπορούν να εμφωλευθούν και μπορούν να έχουν χαρακτηριστικά και περιεχόμενο. Τα ονόματα χώρου XML επιτρέπουν τον καθορισμό διαφορετικών *λεξιλογίων σήμανσης (markup vocabularies)* σε ένα έγγραφο XML. Το *σχήμα XML (XML schema)* χρησιμεύει για την έκφραση σχήματος ενός συγκεκριμένου συνόλου εγγράφων XML.

Μια βασική μορφή αναπαράστασης δεδομένων για το σημασιολογικό Ιστό είναι το *RDF (Resource Description Framework)*. Το RDF είναι ένα πλαίσιο για την αναπαράσταση πληροφοριών σχετικά με τους πόρους σε μορφή γραφήματος. Προοριζόταν κυρίως για την αναπαράσταση μεταδεδομένων (*metadata*) σχετικά με τους πόρους του WWW, όπως ο τίτλος, ο συγγραφέας και η ημερομηνία

τροποποίησης μιας ιστοσελίδας, αλλά μπορεί να χρησιμοποιηθεί για την αποθήκευση οποιωνδήποτε άλλων δεδομένων. Βασίζεται σε τριάδες υποκείμενο-κατηγορήμα-αντικείμενο (subject-predicate-object) που σχηματίζουν γράφο δεδομένων. Όλα τα δεδομένα στον σημασιολογικό ιστό χρησιμοποιούν το RDF ως την κύρια γλώσσα αναπαράστασης. Η σύνταξη του RDF βασίζεται στην XML.

Από μόνο του το RDF χρησιμεύει ως περιγραφή ενός γραφήματος που σχηματίζεται από τριάδες. Ο καθένας μπορεί να ορίσει το λεξιλόγιο (vocabulary) των όρων που χρησιμοποιούνται για πιο λεπτομερή περιγραφή. Για να επιτραπεί η τυποποιημένη περιγραφή των *ταξινομιών (taxonomies)* και άλλων οντολογικών κατασκευών, δημιουργήθηκε ένα *Σχέδιο RDF (RDF Schema ή RDFS)* μαζί με την επίσημη σημασιολογία του στο RDF. Το RDFS μπορεί να χρησιμοποιηθεί για να περιγράψει τις ταξινομίες των κλάσεων (classes) και των ιδιοτήτων (properties) και να τις χρησιμοποιήσει για τη δημιουργία ελαφρών οντολογιών.

Μπορούμε να δημιουργήσουμε πιο λεπτομερείς οντολογίες με την *OWL (Web Ontology Language)*. Η OWL είναι μια γλώσσα περιγραφικής λογικής που συντακτικά ενσωματώνεται στο RDF όπως το RDFS, παρέχοντας επιπλέον τυποποιημένο λεξιλόγιο, και επίσημη σημασιολογία (formal semantics).

Τα RDFS και OWL έχουν καθορισμένη σημασιολογία και αυτή η σημασιολογία μπορεί να χρησιμοποιηθεί για διαδικασίες συλλογιστικής (reasoning) εντός οντολογιών και βάσεων γνώσεων που περιγράφονται χρησιμοποιώντας αυτές τις γλώσσες. Για την παροχή κανόνων πέρα από τις υφιστάμενες δομές αυτών των γλωσσών, έχουν αναδειχτεί δύο πρότυπα - RIF και SWRL.

Για την αναζήτηση δεδομένων RDF καθώς και RDFS και OWL οντολογιών σε βάσεις γνώσεων, υπάρχει η SPARQL (Simple Protocol and RDF Query Language). Η SPARQL είναι γλώσσα τύπου SQL, αλλά χρησιμοποιεί τριάδες (triples) και πόρους (resources) RDF τόσο για αντιστοίχιση τμήματος του ερωτήματος όσο και για την επιστροφή των αποτελεσμάτων του ερωτήματος. Η SPARQL δεν είναι μόνο γλώσσα ερωτημάτων (query language), αλλά είναι επίσης ένα πρωτόκολλο για την πρόσβαση σε δεδομένα RDF.

Όλη η σημασιολογία και οι κανόνες εκτελούνται στα επίπεδα κάτω από την Απόδειξη (Proof) και το αποτέλεσμα θα ελεγχθεί ως προς την αξιοπιστία του. Η επίσημη απόδειξη μαζί με αξιόπιστες εισόδους για την απόδειξη σημαίνει ότι τα αποτελέσματα μπορούν να είναι αξιόπιστα, οπότε εμφανίζονται στο επάνω επίπεδο του παραπάνω σχήματος. Για αξιόπιστες εισόδους, χρησιμοποιούνται μέσα κρυπτογράφησης, όπως *ψηφιακές υπογραφές (digital signatures)* για επαλήθευση της προέλευσης των πηγών. Πάνω από αυτά τα επίπεδα, μπορεί να δημιουργηθούν εφαρμογές με *διεπαφή χρήστη (user interface)*.

Ο Σημασιολογικός Ιστός βασίζεται σε ορισμένες τεχνολογίες.

- *Μεταδεδομένα (Metadata)* - Είναι δομημένες πληροφορίες που περιγράφουν, εξηγούν, εντοπίζουν ή διευκολύνουν με άλλο τρόπο την ανάκτηση, χρήση ή διαχείριση ενός πόρου πληροφοριών. Τα μεταδεδομένα ονομάζονται συχνά δεδομένα για τα δεδομένα ή πληροφορίες για τις πληροφορίες [16].

Τα μεταδεδομένα παρέχουν πληροφορίες που επιτρέπουν την κατανόηση των δεδομένων (π.χ. έγγραφα, εικόνες, σύνολα δεδομένων), έννοιες (π.χ. σχήματα ταξινόμησης) και οντότητες πραγματικού κόσμου (π.χ. άτομα, οργανισμοί, μέρη, πίνακες ζωγραφικής, προϊόντα).

Τύποι μεταδεδομένων:

- Περιγραφικά μεταδεδομένα, περιγράφουν ένα πόρο (resource) με σκοπό ανακάλυψης (discovery) και αναγνώρισης (identification).
- Δομικά μεταδεδομένα, π.χ. μοντέλα δεδομένων (datamodels) και δεδομένα αναφοράς (referencedata).
- Μεταδεδομένα διαχείρισης, παρέχουν πληροφορίες για τη διαχείριση ενός πόρου(resource).

Η διαχείριση των μεταδεδομένων πρέπει να διασφαλίζει:

- Διαθεσιμότητα (Availability): τα μεταδεδομένα πρέπει να είναι προσβάσιμα καταχωρημένα σε ευρετήρια ώστε να μπορούν να βρεθούν.

- Ποιότητα (Quality): τα μεταδεδομένα πρέπει να είναι σταθερής ποιότητας, ώστε οι χρήστες να γνωρίζουν ότι μπορούν να τα εμπιστευθούν.
- Ανθεκτικότητα (Persistence): τα μεταδεδομένα πρέπει να διατηρούνται με την πάροδο του χρόνου.
- Ανοιχτή άδεια χρήσης (OpenLicense): τα μεταδεδομένα πρέπει να είναι διαθέσιμα με άδεια Κοινού Κτήματος (PublicDomain) ώστε να επιτρέπουν την επαναχρησιμοποίησή τους.
- Ορολογίες (Terminologies) - παρέχουν κοινόχρηστα και κοινά λεξιλόγια (vocabularies) ενός πεδίου ορισμού (domain), ώστε οι μηχανές αναζήτησης (searchengines), οι πράκτορες (agents), οι συντάκτες (authors) και οι χρήστες (users) να μπορούν να επικοινωνούν. Όχι καλό, εκτός αν η κάθε μία σημαίνει το ίδιο πράγμα.
- Οντολογίες (Ontologies) - αποτελούν το δομικό στοιχείο του Σημασιολογικού Ιστού. Παρέχουν μια κοινή κατανόηση ενός πεδίου ορισμού (domain) που μπορεί να κοινοποιηθεί σε άτομα και εφαρμογές και θα διαδραματίσει σημαντικό ρόλο στην υποστήριξη της ανταλλαγής πληροφοριών και της ανακάλυψης. Εξετάζονται εκτενώς στο επόμενο κεφάλαιο.

2.2.2 Οντολογίες

Οι οντολογίες αναπτύχθηκαν στην Τεχνητή Νοημοσύνη (Artificial Intelligence) για να ικανοποιήσουν το διαμοιρασμό (sharing) και την επαναλαμβανόμενη χρήση (reuse) της γνώσης. Παρέχουν μία σημασιολογία επεξεργάσιμη από τη μηχανή (machine-processable semantics), πηγών πληροφοριών που μπορούν να είναι επικοινωνήσιμες μεταξύ διαφορετικών πρακτόρων (agents).

Μία οντολογία είναι μια κατηγορηματική (explicit), τυπική (formal), προδιαγραφή μιας διαμοιρασμένης (shared) εννοιολογικής αναπαράστασης (conceptualization), [17]. Ο όρος “κατηγορηματική” (explicit) σημαίνει ότι το είδος των εννοιών που χρησιμοποιούνται, και οι περιορισμοί που αφορούν την χρήση αυτών των εννοιών είναι προσδιορισμένοι με σαφήνεια. Ο όρος “τυπική” (formal) αναφέρεται στο ότι η

οντολογία πρέπει να είναι αναγνώσιμη από τη μηχανή. Ο όρος “διαμοιρασμένη” (shared) αναφέρεται στο ότι η οντολογία πρέπει να αποτυπώνει γνώση κοινής αποδοχής στα πλαίσια της κοινότητας. Τέλος, ο όρος “εννοιολογική αναπαράσταση” (conceptualization) αναφέρεται σε ένα αφηρημένο μοντέλο φαινομένων του κόσμου στο οποίο έχουν προσδιοριστεί οι έννοιες που σχετίζονται με τα φαινόμενα αυτά.

Στον παρακάτω πίνακα παρουσιάζονται τα βασικά συστατικά των οντολογιών:

Table 1: Συστατικά των οντολογιών

Κλάσεις (Classes)	Αναπαραστάσεις συλλογών πόρων (resources), που έχουν συγκεκριμένα, κοινά χαρακτηριστικά.
Άτομα (Individuals)	Στιγμιότυπα (instances) είναι τα μέλη μιας κλάσης.
Ιδιότητες (Attributes)	Οι ιδιότητες, χαρακτηριστικά ή παράμετροι των κλάσεων.
Σχέσεις (Relations)	Οι τρόποι με τους οποίους οι κλάσεις ή τα στιγμιότυπα σχετίζονται μεταξύ τους.
Συναρτήσεις (functions)	Μορφές σχέσεων, που βάσει κάποιων χαρακτηριστικών παράγουν κάποιο νέο χαρακτηριστικό.
Περιορισμοί (Restrictions)	Περιορισμοί στις σχέσεις και στις ιδιότητες των οντολογιών.
Κανόνες (Rules)	Δηλώσεις της μορφής Αν - Τότε .
Αξιώματα (Axioms)	Αναπαριστούν λογικές προτάσεις που θεωρούνται πάντα αληθείς.
Γεγονότα (Events)	Η αλλαγή ιδιοτήτων ή σχέσεων.

2.2.3 Δομή της OWL και βιβλιοθήκη Owlapi

Υπάρχουν αρκετές γλώσσες περιγραφής οντολογιών. Εμείς θα ασχοληθούμε με Οντολογίες που έχουν οριστεί με την γλώσσα *OWL (Web Ontology Language)*. Η OWL είναι μια οντολογική γλώσσα σχεδιασμένη για τον *Σημασιολογικό Ιστό (SemanticWeb)*. Παρέχει μια πλούσια συλλογή τελεστών (*operators*) για τη διαμόρφωση *περιγραφικών εννοιών (concept descriptions)*. Πρόκειται για ένα πρότυπο W3C (W3C standard), που προωθεί τη *διαλειτουργικότητα (interoperation)* και το *διαμοιρασμό (sharing)* μεταξύ εφαρμογών και έχει σχεδιαστεί ώστε να είναι συμβατή με τα υπάρχοντα *πρότυπα ιστού (web standards)*.

Όπως έχουμε αναφέρει η σύνταξή της είναι μια επέκταση του RDF και του RDFS τα οποία συντάσσονται με XML με τα παρακάτω βασικά στοιχεία.

Table 2: RDF Στοιχεία Σύνταξης βασισμένα στην XML

rdf:RDF	Ριζικό στοιχείο (rootelement) των εγγράφων RDF. Ορίζει το έγγραφο XML ως έγγραφο RDF. Περιέχει επίσης μια αναφορά στον χώρο ονομάτων RDF (RDF namespace).
rdf:Description	Στοιχείο που περιέχει την περιγραφή του πόρου (resource).
rdf:type	Στιγμιότυπο (instance) του ...
rdf:Bag	Ένα μη ταξινομημένο σύνολο πραγμάτων (container) πόρων (resources). Ενδεχομένως να περιέχει και διπλές τιμές.
rdf:Seq	Ένα ταξινομημένο σύνολο πόρων (resources). Ενδεχομένως να περιέχει και διπλές τιμές.
rdf:Alt	Ορίζει ένα σύνολο εναλλακτικών πόρων από τις οποίες ο χρήστης μπορεί να επιλέξει μόνο μία.
rdf:ID	Υποδεικνύει ένα νέο πόρο (resource).
rdf:about	Αναφορά σε ένα υπάρχοντα πόρο (resource).
rdf:resource	Επιτρέπει σε στοιχεία ιδιοτήτων (propertyelements) να οριστούν

	ως πόροι (resources).
--	-----------------------

Η επέκταση των συντάξεων των XML, RDF και RDFS που ορίζει η OWL γίνεται αντιληπτή από την κεφαλίδα των OWL αρχείων που έχουν την εξής μορφή

```
<rdf:RDF
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#">
```

Η owl ορίζει 3 υπογλώσσες τις OWL Lite, OWL DL, OWL Full κάθε μια με αυξημένες δυνατότητες σε εκφραστικότητα σε σχέση με την προηγούμενη.

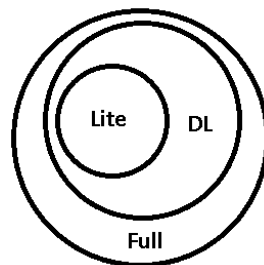


Figure 12: OWL sublanguages

Με το λεξιλόγιου του RDFS μπορούμε να περιγράψουμε κλάσεις (classes) και ιδιότητες (properties). Ωστόσο με την OWL μπορούμε επίσης να περιγράψουμε τις σχέσεις μεταξύ των κλάσεων και των ιδιοτήτων καθώς και τα χαρακτηριστικά τους. Οι κλάσεις και οι ιδιότητες μπορούν να έχουν και σχολιασμούς (annotations).

Η owlapi [18] είναι μια βιβλιοθήκη για java όπου χρησιμοποιείται στην επεξεργασία και την ανάγνωση OWL αρχείων. Προσφέρει διάφορες δυνατότητες και προγραμματιστικές διεπαφές σχετικές με οντολογίες. Εμείς θα επικεντρωθούμε μόνο στο διάβασμα για να μπορούμε να πάρουμε τις κλάσεις και τις ιδιότητες ενός OWL αρχείου μαζί με την επιπλέον πληροφορία που μας ενδιαφέρει. Συγκεκριμένα μιλάμε για τον OWL/XML αναλυτή (parser) που διαθέτει. Με αυτόν μπορούμε να διαβάζουμε OWL αρχεία σε οποιαδήποτε νόμιμη μορφή έχουν αυτά γραφεί.

Μπορούμε να αναλύουμε έτσι τα αξιώματα (axioms) που έχουν οριστεί και να χειριζόμαστε προγραμματιστικά την πληροφορία που μας ενδιαφέρει.

2.3. Καθορισμός Συσχετίσεων μεταξύ Οντολογιών

Έχουμε αναφέρει τα κύρια χαρακτηριστικά των Οντολογιών και συγκεκριμένα αυτών που ορίζονται με την OWL. Κύριος σκοπός της διπλωματικής αυτής είναι να μελετήσουμε τις συσχετίσεις που έχουν μεταξύ τους διαφορετικές τέτοιες οντολογίες. Για αυτόν τον λόγο βασιζόμαστε στο *Ontology Alignment Tool* [1]. Το εργαλείο αυτό δέχεται 2 οντολογίες owl και παράγει κανόνες αντιστοίχισης (*mapping rules*) τους οποίους εμείς θα χρησιμοποιήσουμε. Οι κανόνες αυτοί προκύπτουν με αυτοματοποιημένο τρόπο, αλλά οι χρήστες του εργαλείου έχουν την δυνατότητα να τους επεξεργαστούν και να προσθαφαιρέσουν κανόνες.

Σε κάθε κανόνα για να αντιστοιχιστούν τα στοιχεία (π.χ. κλάσεις, ιδιότητες, κτλ) της μιας οντολογίας με της άλλης, βασιζόμαστε στην έννοια των *μοτίβων οντολογιών* (*Ontologies Patterns*). Ένα τέτοιο μοτίβο περιγράφει επακριβώς τις παραμέτρους της οντολογίας που συμμετέχει στον κανόνα. Μπορεί να αναφέρεται σε ένα υπάρχων στοιχείο μιας οντολογίας (π.χ. μια κλάση), μπορεί να προσθέτει ένα νέο στοιχείο (π.χ. μια κλάση που θα έχει ρόλο domain σε μια ιδιότητα) ή γενικά να περιέχει έναν οποιοδήποτε συνδυασμών στοιχείων μιας οντολογίας. Κάποιο τέτοιο μοτίβο μπορεί να αποτελείται από άλλα τέτοια μοτίβα. Ορίζονται πολλά τέτοια μοτίβα ανάλογα με την φύση του κανόνα. Για παράδειγμα αναφέρουμε δύο τέτοια:

- Simple Relation Pattern

Το οποίο ορίζει μια σχέση (Relation) ενός συγκεκριμένου Object Property μιας οντολογίας.

- Relation Path Pattern

Με το οποίο ορίζεται μια Relation όπου σχηματίζεται από ένα ή περισσότερα Properties μιας οντολογίας.

Ο κάθε κανόνας που αντιστοιχίζει δύο οντολογίες έχει συγκεκριμένη μορφή. Κάποια πεδία είναι υποχρεωτικά και περιλαμβάνονται πάντα, ενώ άλλα είναι προαιρετικά και μπορεί να παραλείπονται. Υποχρεωτικά πεδία είναι

- Τα πεδία *entity 1* και *entity 2* τα οποία περιέχουν τα *Ontologies Patterns* που ορίσαμε πιο πάνω για την πρώτη και δεύτερη οντολογία αντίστοιχα. Επίσης πάντα έχει ένα πεδίο *Relation* που
- Το πεδίο *Relation* όπου καθορίζει την συσχέτιση των στοιχείων της πρώτης οντολογίας με την δεύτερη (π.χ. σχέση ισοδυναμίας).

Ορίζονται επίσης τα εξής προαιρετικά πεδία

- Το πεδίο *Transformation* όπου ορίζει κάποιον μετασχηματισμό που λαμβάνει μέρος στις τιμές των ιδιοτήτων της οντολογίας που συμμετέχουν στον κανόνα. Αυτό μπορεί να έχει την μορφή *directTransformation* ή *inverseTransformation* ανάλογα με το είδος του μετασχηματισμού.
- Το πεδίο *Direction* όπου δείχνει για ποια κατεύθυνση (από την 1η στην 2η ή το αντίστροφο) ισχύει η συσχέτιση του κανόνα.
- Το πεδίο *Origin* όπου δείχνει την προέλευση του κανόνα.
- Το πεδίο *Confidence Value* όπου δείχνει το πόσο σίγουρο ήταν το εργαλείο που για την ισχύ του κανόνα που παρήγαγε.
- Το πεδίο *Comments* όπου έχει την περιγραφή του συγκεκριμένου κανόνα.

Αυτά τα βλέπουμε και στον παρακάτω πίνακα που έχει οριστεί από τους δημιουργούς του εργαλείου:

Table 2. The (M)andatory and (O)ptional parameters of a Mapping Rule

Parameter	M/O	Description
Entity 1 and 2	M	Specifies the elements participating in the left and right side of a mapping rule.
Transformation	O	Specifies the transformation that should take place in the value(s) of properties defined
Relation	M	Specifies the relation of Entity 1 towards Entity 2 (e.g., equivalent terms)
Direction	O	Specifies the direction for which the correspondence is valid. (e.g., from ontology 1 to ontology 2)
Origin	O	Indicates the way the mapping rule has been produced (e.g., manually specified by the end user through OAT)
Confidence Value	O	Indicates how confident the OAT was for the suggested mapping rule, when accepted by the end user
Comments	O	Provides a human readable description of the mapping rule specified

Figure 13: Πεδία ενός κανόνα συσχέτισης [1]

Εμείς θα χρησιμοποιούμε τα αποτελέσματα αυτού του εργαλείου τα οποία θα είναι οι παραγόμενοι κανόνες συσχέτισης όπως τους περιγράψαμε παραπάνω σε ένα αρχείο JSON.

3. Περιγραφή Συστήματος

Asdfa

asdf

4. Υλοποίηση Συστήματος και Παραδείγματα

Φσαδφα

ασδφ

5. Συμπεράσματα και μελλοντική εξέλιξη

Δφασδ

δασφς

6. Σύνοψη

Δφασδ

7. Βιβλιογραφικές Αναφορές

- [1]: Efthymios Chondrogiannis, Vassiliki Andronikou, Efstathios Karanastasis, Theodora Varvarigou, *Ontology Alignment Tool*, , Ανακτήθηκε από <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.664.4075&rep=rep1&type=pdf>
- [2]: WHATWG, *HTML Living Standard*, 2020, Ανακτήθηκε από <https://html.spec.whatwg.org/print.pdf>
- [3]: , *The HTML DOM*, , Ανακτήθηκε από https://www.w3schools.com/whatis/whatis_html5dom.asp
- [4]: , *ECMAScript Specification*, 2020, Ανακτήθηκε από <https://www.ecma-international.org/ecma-262/>
- [5]: , *Concurrency model and the event loop*, , Ανακτήθηκε από <https://developer.mozilla.org/en-US/docs/Web/JavaScript/EventLoop>
- [6]: , *The Node.js Event Loop*, , Ανακτήθηκε από <https://nodejs.org/en/docs/guides/event-loop-timers-and-nexttick/>
- [7]: , *Promise*, , Ανακτήθηκε από https://developer.mozilla.org/en-US/docs/Web/JavaScript/Reference/Global_Objects/Promise
- [8]: , *Reactivity in VueJs*, , Ανακτήθηκε από <https://vuejs.org/v2/guide/reactivity.html>
- [9]: , *The Vue Instance*, , Ανακτήθηκε από <https://vuejs.org/v2/guide/instance.html>
- [10]: , *Hypertext Transfer Protocol -- HTTP/1.1*, , Ανακτήθηκε από <https://www.ietf.org/rfc/rfc2616.txt>
- [11]: , *Introducing JSON*, , Ανακτήθηκε από <https://www.json.org/json-en.html>
- [12]: , *Actor Model in Nutshell*, , Ανακτήθηκε από <https://medium.com/@KtheAgent/actor-model-in-nutshell-d13c0f81c8c7>
- [13]: , *Vertx*, , Ανακτήθηκε από <https://vertx.io/>
- [14]: Julien Ponge, Thomas Segismont, Julien Viet, *A gentle guide to asynchronous programming with Eclipse Vert.x for Java developers*, 2019, Ανακτήθηκε από <https://vertx.io/docs/guide-for-java-devs/guide-for-java-devs.pdf>
- [15]: Berners-Lee, T., Hendler, J. & Lassila, , 2001, Ανακτήθηκε από <https://web.archive.org/web/20081114135540/http://www.sciam.com/article.cfm?id=the-semantic-web&print=true>.
- [16]: Riley, J., *Understanding Metadata* Baltimore: National Information Standards Organization (NISO), 2017, Ανακτήθηκε από https://groups.niso.org/apps/group_public/download.php/17446/Understanding%20Metadata.pdf
- [17]: Gruber T.R., *A Translation Approach to Portable Ontology Specification*, Knowledge Acquisition, 1993
- [18]: , *Owlapi Wiki*, , Ανακτήθηκε από <https://github.com/owlcs/owlapi/wiki>