



ΧΑΡΟΚΟΠΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ

ΣΧΟΛΗ ΨΗΦΙΑΚΗΣ ΤΕΧΝΟΛΟΓΙΑΣ

ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΜΑΤΙΚΗΣ

**Σημασιολογικός Ιστός: Αξιολόγηση Εργαλείων Συγχώνευσης
Οντολογιών Με Πραγματικά Δεδομένα**

Πτυχιακή Εργασία

Χρήστος Νάτσος

Αθήνα, 2018

1



ΧΑΡΟΚΟΠΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ

ΣΧΟΛΗ ΨΗΦΙΑΚΗΣ ΤΕΧΝΟΛΟΓΙΑΣ

ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΜΑΤΙΚΗΣ

Τριμελής εξεταστική επιτροπή

Δημοσθένης Αναγνωστόπουλος

Μάρα Νικολαΐδου

Αγγελική Καραγιαννάκη

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

Περίληψη στα Ελληνικά.....	5
Περίληψη στα Αγγλικά	6
Κατάλογος Εικόνων.....	7
Συντομογραφίες	8
Κεφάλαιο 1: Εισαγωγή.....	9
1.1 Δήλωση προβλήματος.....	10
1.2 Σκοπός της πτυχιακής	10
1.3 Δομή της πτυχιακής	11
Κεφάλαιο 2: Βιβλιογραφική Επισκόπηση	12
2.1 4 ^η βιομηχανική επανάσταση (Industry 4.0)	12
2.1.1 Τι ονομάζουμε 4 ^η βιομηχανική επανάσταση (Industry 4.0);	12
2.1.2 Οι τεχνολογίες της 4 ^{ης} βιομηχανικής επανάστασης	16
2.2 Σημασιολογικός Ιστός.....	19
2.2.1 Τι είναι ο σημασιολογικός ιστός	19
2.2.2 Οι τεχνολογίες του σημασιολογικού ιστού.....	20
2.2.3 Οντολογίες.....	23
2.3 Συγχώνευση Οντολογιών	25
2.3.1 Ορισμοί.....	25
2.3.2 Εργαλεία/Προσεγγίσεις	26
Κεφάλαιο 3: Μεθοδολογία.....	29
3.1 Δεδομένα που θα χρησιμοποιηθούν	31
3.1.1 Οντολογία δημοσιεύσεων IFX.....	31
3.1.2 Οντολογία διατριβών IFX	33
Κεφάλαιο 4: Εφαρμογή και αποτελέσματα	37
4.1 προσέγγιση LogMap	37
4.1.1 LogMap – Διαδικασία εφαρμογής	38

4.1.2 LogMap – Αποτελέσματα	39
4.2 Προσέγγιση Protégé.....	42
4.2.1 Protégé – Διαδικασία εφαρμογής.....	42
4.2.2 Protégé – Αποτελέσματα	43
Συμπέρασμα	47
Βιβλιογραφία	48

Περίληψη στα Ελληνικά

Οι Τεχνολογίες Σημασιολογικού Ιστού (Semantic Web Technologies) αντιπροσωπεύουν την επόμενη σημαντική εξέλιξη στη σύνδεση των πληροφοριών. Ο Σημασιολογικός Ιστός είναι μια επέκταση του World Wide Web μέσω προτύπων από την Κοινοπραξία του World Wide Web (World Wide Web Consortium). Η τέταρτη βιομηχανική επανάσταση, η επονομαζόμενη Industry 4.0, θα εκμεταλλευτεί αυτή την πρωτοποριακή τεχνολογία, η οποία θα επιτρέψει τη χρήση μεγάλων δεδομένων, συντακτικών αναφορών, συνδετικότητας, αυτοματισμού, ευελιξίας στις επιχειρήσεις, διαδίκτυο των πραγμάτων (Internet of Things) και πολλών άλλων. Οι τοπικές δομές μπορούν να μοντελοποιηθούν με τη χρήση οντολογιών που θεωρούνται ως ένας από τους πυλώνες του Σημασιολογικού Ιστού. Η συγχώνευση οντολογιών σημαίνει δημιουργία μιας νέας οντολογίας, συνδέοντας τις υπάρχουσες. Η συγχώνευση των οντολογιών με ελεγχόμενο και κλιμακωτό τρόπο είναι κρίσιμη για την επιτυχία του Σημασιολογικού Ιστού. Έχουν γίνει πολλές έρευνες και προσπάθειες κατα καιρούς, αλλά εξακολουθεί να υπάρχει ένα κενό στο πώς μπορούν να συγχωνευθούν οι οντολογίες με επιτυχία. Σε αυτή τη πτυχιακή θα προσπαθήσουμε να εφαρμόσουμε μερικές προσεγγίσεις σχετικά με την συγχώνευση οντολογιών, προκειμένου να βρεθεί μια αποτελεσματική λύση.

Λέξεις κλειδιά: Σημασιολογικός Ιστός, Τεχνολογίες Σημασιολογικού Ιστού, Οντολογίες, Συγχώνευση Οντολογιών

Περίληψη στα Αγγλικά

Semantic Web Technologies represent the next major evolution in connecting information. The Semantic Web is an extension of the World Wide Web through standards by the World Wide Web Consortium. The fourth industrial revolution, the so-called Industry 4.0 will take advantage of this innovative technology enabling Big Data, Prescriptive Analytics, Connectivity, Automation, Business Agility, Internet of Things and many more. Local structures can be modeled by ontologies, which are considered one of the pillars of the Semantic Web. Merging of ontologies means the creation of a new ontology by linking up the existing ones. Merging ontologies in a controlled and scalable way is crucial for the success of the Semantic Web. There has been a lot of research and effort, but there is still a gap in how to merge ontologies. In this thesis, we will try several approaches on ontology merging in order to find which one is more suitable and effective.

Keywords: Semantic Web, Semantic Web Technologies, Ontologies, Ontology Merging

Κατάλογος Εικόνων

- Εικόνα 1: Οντολογία δημοσιεύσεων IFX - Κλάσεις
- Εικόνα 2: Οντολογία δημοσιεύσεων IFX - Ιδιότητες Αντικειμένων
- Εικόνα 3: Οντολογία δημοσιεύσεων IFX – Ιδιότητες Δεδομένων
- Εικόνα 4: Οντολογία διατριβών IFX - Κλάσεις
- Εικόνα 5: Οντολογία διατριβών IFX - Ιδιότητες Αντικειμένων
- Εικόνα 6: Οντολογία διατριβών IFX - Ιδιότητες Δεδομένων
- Εικόνα 7: LogMap – Διαδικτυακή εφαρμογή
- Εικόνα 8: LogMap – Βήματα συγχώνευσης
- Εικόνα 9: LogMap – Ολοκληρωμένη οντολογία - Κλάσεις
- Εικόνα 10: LogMap – Ολοκληρωμένη οντολογία - Ιδιότητες Αντικειμένων
- Εικόνα 11: LogMap – Ολοκληρωμένη οντολογία - Ιδιότητες Δεδομένων
- Εικόνα 12: Protégé – Λειτουργία Συγχώνευσης
- Εικόνα 13: Protégé's συγχωνευμένη οντολογία – 1^η φάση - Κλάσεις
- Εικόνα 14: Protégé's συγχωνευμένη οντολογία – 1^η φάση - Ιδιότητες Αντικειμένων
- Εικόνα 15: Protégé's συγχωνευμένη οντολογία – 1^η φάση - Ιδιότητες Δεδομένων
- Εικόνα 16: Protégé's συγχωνευμένη οντολογία – 2^η φάση - Κλάσεις
- Εικόνα 17: Protégé's συγχωνευμένη οντολογία – 2^η φάση - Ιδιότητες Αντικειμένων
- Εικόνα 18: Protégé's συγχωνευμένη οντολογία – 2^η φάση - Ιδιότητες Δεδομένων

Συντομογραφίες

AR	Augmented Reality
AS	Automation Systems
ATOM	Automatic Target-driven Ontology Merging
CPS	Cyber Physical Systems
DP	Digital Production
HUA	Harokopio University of Athens
IFX	Infineon Technologies AG
IoS	Internet of Services
IoT	Internet of Things
IT	Information Technologies
PLM	Product Lifecycle Management
RDF	Resource Description Framework
SCN	Supply Chain Networks
SPARQL	SPARQL Protocol and RDF Query Language
URI	Universal Resource Identifier
URL	Uniform Resource Locator
URN	Uniform Resource Name
VR	Virtual Reality
WoT	Web of Things
WP	Work Package
XML	eXtensible Markup Language

Κεφάλαιο 1: Εισαγωγή

Η εφεύρεση του ατμού εισήγαγε την πρώτη εποχή της βιομηχανίας με μηχανήματα μέσα στα εργοστάσια. Η δεύτερη εποχή εισήγαγε τη μαζική παραγωγή με τη χρήση ηλεκτρισμού. Η τρίτη εποχή κατέληξε σε συστήματα πληροφορικής, αυτοματισμούς, ρομπότ και μηχανήματα που αντικαθιστούν τους ανθρώπινους εργάτες. Τώρα μπαίνουμε στην τέταρτη βιομηχανική επανάσταση που ονομάζεται "Industry 4.0", όπου οι υπολογιστές επικοινωνούν με συστήματα αυτοματισμού. Τα ρομποτικά συνδέονται εξ αποστάσεως με συστήματα υπολογιστών τα οποία είναι εξοπλισμένα με αλγόριθμους μηχανικής μάθησης που μπορούν να μάθουν και να ελέγχουν τη ρομποτική με πολύ λίγες εισροές από ανθρώπινους χειριστές. Οι Τεχνολογίες Σημαιολογικού Ιστού αποτελούν μέρος αυτής της επανάστασης. Η σημασιολογία είναι ένα μεγάλο βήμα προς τη Βιομηχανία 4.0, επιτρέποντας τα Μεγάλα Δεδομένα, Προπαρασκευαστικά Analytics, Συνδεσιμότητα, Αυτοματισμό, Επιχειρηματική Ευελιξία, Διαδίκτυο των πραγμάτων και πολλά άλλα. Σύμφωνα με τον Tim Berners-Lee et al (2001), ο Σημαιολογικός Ιστός είναι μια επέκταση του τρέχοντος ιστού στον οποίο δίνεται σαφώς καθορισμένη σημασία στις πληροφορίες, επιτρέποντας καλύτερα στους υπολογιστές και τους ανθρώπους να δουλεύουν σε συνεργασία. Το 2017, η Infineon Technologies AG (IFX) ξεκίνησε με άλλες εταιρείες το έργο Productive 4.0 με χορηγία της Ευρωπαϊκής Ένωσης. Το Productive 4.0 είναι ένα έργο καινοτομίας που αποσκοπεί να ανοίξει τις πόρτες στις δυνατότητες της Ψηφιακής Βιομηχανίας που εστιάζει στην ψηφιακή παραγωγή, στα δίκτυα εφοδιαστικής αλυσίδας (SCN) και στη διαχείριση του κύκλου ζωής του προϊόντος (PLM). Όλα αυτά αλληλεπιδρούν και αλληλεπιδρούν. Το έργο αυτό υπολογίζεται ότι θα διαρκέσει μέχρι το τέλος Απριλίου 2020.

1.1 Δήλωση προβλήματος

Το έργο Productive 4.0 αποτελείται από 10 πακέτα εργασίας (WPs), το καθένα με μεμονωμένους τεχνικούς στόχους. Το τμήμα εφοδιαστικής αλυσίδας της IFX στην Γερμανία είναι υπεύθυνο για το WP 7 σε συνεργασία με το Χαροκόπειο Πανεπιστήμιο (HUA) της Αθήνας στην Ελλάδα. Αυτό το WP αναφέρεται ως προς το πλαίσιο εκμετάλλευσης του έργου. Χωρίς σωστή στρατηγική εκμετάλλευσης και καλή έρευνα, κάθε έργο μπορεί να παραμείνει απλώς ένα ερευνητικό έργο, χωρίς να φτάσει ποτέ στο μέγιστο των δυνατοτήτων του και να μην δημιουργήσει ποτέ αντίκτυπο στην κοινωνία μας. Η εκμετάλλευση βοηθά στην επιτάχυνση της επιστημονικής προόδου. Υπάρχουν διαφορετικές προσεγγίσεις για την αξιοποίηση / αξιοποίηση των αποτελεσμάτων της έρευνας. Η IFX και το HUA χρησιμοποιούν την περίπτωση του "Open Science Model", όπου η εκμετάλλευση πραγματοποιείται κυρίως μέσω της εκπαίδευσης και της δημοσίευσης των αποτελεσμάτων της έρευνας. Προκειμένου να επεξεργαστούν οι τεχνολογίες σημασιολογικού ιστού, οι εσωτερικές και εξωτερικές δημοσιεύσεις του έργου πρέπει να έχουν μια οντολογία. Μια οντολογία ορίζει τις κοινές λέξεις και έννοιες (έννοιες) που χρησιμοποιούνται για να περιγράψουν και να αντιπροσωπεύσουν ένα χώρο γνώσης, και τυποποιεί τις έννοιες. Οι οντολογίες περιλαμβάνουν επίσης χρησιμοποιήσιμους ορισμούς βασικών εννοιών στον τομέα και τις σχέσεις μεταξύ τους.

1.2 Σκοπός της πτυχιακής

Στόχος της παρούσας εργασίας είναι να εισαγάγει τον αναγνώστη στο θέμα των τεχνολογιών σημασιολογικού ιστού, την εφαρμογή του και τον βασικό του ρόλο στην τέταρτη βιομηχανική επανάσταση. Στη συνέχεια να εξηγήσει το νόημα και τον ρόλο της οντολογίας, μίας από τις βασικότερες τεχνολογίες του Σημασιολογικού Ιστού. Τέλος, μετά την ανάγνωση σχετικών εργασιών που έχουν ήδη πραγματοποιηθεί σε μεμονωμένες προσεγγίσεις, να διεξαχθούν δοκιμές σε διαφορετικές προσεγγίσεις και εργαλεία για να βρεθεί μία ιδανική λύση σχετικά με τη σύζευξη οντολογιών μέσα από διάφορες δοκιμές και πειράματα.

1.3 Δομή της πτυχιακής

Η παρούσα πτυχιακή είναι δομημένη ως εξής. Στο Κεφάλαιο 2 παρουσιάζεται η βιβλιογραφική ανασκόπηση, ξεκινώντας από τη θεμελιώδη έννοια της τεχνολογίας σημασιολογικού ιστού, συνεχίζει με την περίληψη της 4^{ης} βιομηχανικής γενιάς και κλείνει με μεθόδους συγχώνευσης / ευθυγράμμισης οντολογιών. Στο Κεφάλαιο 3 εξηγείται η μεθοδολογία αυτής της εργασίας. Στο κεφάλαιο 4 γίνεται εφαρμογή των εργαλείων συγχώνευσης και παρουσιάζονται τα αποτελέσματα. Τέλος η πτυχιακή ολοκληρώνεται με το συμπέρασμα του ερευνητή.

Κεφάλαιο 2: Βιβλιογραφική Επισκόπηση

Σε αυτή την ενότητα θα βρείτε μια ανασκόπηση της βιβλιογραφίας που αναλύθηκε. Το πρώτο θέμα που εξετάζεται είναι ο σημασιολογικός ιστός που εξηγεί τι είναι. Στη συνέχεια αναφέρεται η 4^η βιομηχανική γενιά και θα εξηγήτε γιατί οι εμπειρογνώμονες την ονομάζουν τέταρτη βιομηχανική επανάσταση και πώς σχετίζεται με τεχνολογίες σημασιολογικού ιστού. Τέλος, θα διαβάσουμε διαφορετικές προσεγγίσεις για τη συγχώνευση οντολογιών προκειμένου να βρούμε ποια μπορεί να μας βοηθήσει να επιτύχουμε την αποτελεσματικότερη συγχώνευση.

2.1 4^η βιομηχανική επανάσταση (Industry 4.0)

2.1.1 Τι ονομάζουμε 4^η βιομηχανική επανάσταση (Industry 4.0);

Ο Bernard Marr (2016) εξήγησε την εξέλιξη της βιομηχανίας στο άρθρο του "Τι πρέπει να γνωρίζουμε όλοι για το Industry 4.0". Πρώτα ήρθε ο ατμός και τα πρώτα μηχανήματα. Στη συνέχεια η ηλεκτρική ενέργεια, η γραμμή συναρμολόγησης και η γέννηση της μαζικής παραγωγής. Η τρίτη εποχή της βιομηχανίας προήλθε από την εμφάνιση των υπολογιστών και τις αρχές της αυτοματοποίησης, όταν τα ρομπότ και τα μηχανήματα άρχισαν να αντικαθιστούν τους ανθρώπινους εργαζόμενους σε αυτές τις γραμμές συναρμολόγησης. Η βιομηχανία 4.0 είναι η τέταρτη βιομηχανική επανάσταση στην οποία οι υπολογιστές και ο αυτοματισμός θα συναντηθούν με έναν εντελώς νέο τρόπο, ενώ η ρομποτική θα συνδέεται εξ αποστάσεως με υπολογιστικά συστήματα εξοπλισμένα με αλγόριθμους μηχανικής μάθησης που μπορούν να μάθουν και να ελέγχουν τη ρομποτική με πολύ λίγες εισροές από τους ανθρώπους. Η 4^η βιομηχανική γενιά θα εισάγει το "έξυπνο εργοστάσιο", όπου τα φυσικά συστήματα θα γίνονται στο Διαδίκτυο των Πράξεων, επικοινωνώντας και

συνεργάζοντας τόσο μεταξύ τους όσο και με τους ανθρώπους σε πραγματικό χρόνο μέσω του ασύρματου ιστού.

Σύμφωνα με το SAP (2015), μια ευρωπαϊκή πολυεθνική εταιρία λογισμικού με έδρα τη Γερμανία που κάνει επιχειρηματικό λογισμικό για τη διαχείριση επιχειρηματικών δραστηριοτήτων και σχέσεων με τους πελάτες, η 4^η βιομηχανική γενιά (Industry 4.0) είναι ένας συλλογικός όρος για τις τεχνολογίες και τις έννοιες της οργάνωσης της αλυσίδας αξίας. Με βάση τις τεχνολογικές ιδέες των κυβερνο-φυσικών συστημάτων, το Διαδίκτυο των πραγμάτων και το Διαδίκτυο των Υπηρεσιών, διευκολύνει το όραμα του έξυπνου εργοστασίου. Μέσα στα δομοστοιχειωτά δομημένα έξυπνα εργοστάσια της 4^{ης} βιομηχανικής γενιάς, τα κυβερνο-φυσικά συστήματα παρακολουθούν τις φυσικές διαδικασίες, δημιουργούν ένα εικονικό αντίγραφο του φυσικού κόσμου και κάνουν αποκεντρωμένες αποφάσεις. Μέσω του Διαδικτύου των πραγμάτων, τα κυβερνο-φυσικά συστήματα επικοινωνούν και συνεργάζονται μεταξύ τους και με τον άνθρωπο σε πραγματικό χρόνο. Μέσω του Διαδικτύου των Υπηρεσιών, οι εσωτερικές και δια-οργανωτικές υπηρεσίες προσφέρονται και αξιοποιούνται από συμμετέχοντες της αλυσίδας αξίας.

Πιστεύεται από την BDI (2016), την οργάνωση ομπρέλα των γερμανικών βιομηχανιών και των υπηρεσιών που σχετίζονται με τη βιομηχανία, ότι η 4^η βιομηχανική γενιά (Industry 4.0) θα είναι η τέταρτη βιομηχανική επανάσταση. Μετά τη μηχανοποίηση (Industry 1.0), τη μαζική παραγωγή (Industry 2.0) και την αυτοματοποίηση (Industry 3.0), τώρα το "διαδίκτυο των πραγμάτων και των υπηρεσιών" γίνεται αναπόσπαστο μέρος της κατασκευής. Οι τεχνολογίες της 4^{ης} βιομηχανικής γενιάς (Industry 4.0) έχουν τη δυνατότητα να δημιουργήσουν εξαιρετικές ευκαιρίες ανάπτυξης και ανταγωνιστικά πλεονεκτήματα για τη Γερμανία ως επιχειρηματική θέση. Οι ειδικοί προβλέπουν ότι οι επιχειρήσεις θα είναι σε θέση να αυξήσουν την παραγωγικότητά τους κατά 30% χρησιμοποιώντας την 4^η βιομηχανική γενιά (Industry 4.0).

Η Gartner (2015), μια αμερικανική ερευνητική και συμβουλευτική εταιρεία που παρέχει πληροφόρηση σχετικά με τις τεχνολογίες πληροφορικής για την τεχνολογία πληροφορικής, διαπίστωσε ότι η 4^η βιομηχανική γενιά (Industry 4.0) είναι ένα όραμα για την προηγμένη κατασκευή που υποστηρίζεται από την κυβέρνηση της Γερμανίας. Η βασική ιδέα του Industry 4.0 είναι να συνδέσει ενσωματωμένα συστήματα και έξυπνες εγκαταστάσεις παραγωγής για να δημιουργήσει μια ψηφιακή σύγκλιση μεταξύ της βιομηχανίας, των επιχειρήσεων και των εσωτερικών λειτουργιών και διαδικασιών. Η 4^η βιομηχανική γενιά (Industry 4.0) αναφέρεται σε μια τέταρτη βιομηχανική επανάσταση (μετά την ισχύ του νερού / ατμού, μαζική παραγωγή και αυτοματοποίηση μέσω της πληροφορικής και της ρομποτικής) και εισάγει την έννοια των «κυβερνο-φυσικών συστημάτων» για να διαφοροποιήσει αυτή τη νέα εξελικτική φάση από τον ηλεκτρονικό αυτοματισμό που έχει προηγηθεί.

Το Ευρωπαϊκό Κοινοβούλιο (2015) περιέγραψε την 4^η βιομηχανική γενιά (Industry 4.0) ως έναν όρο που εφαρμόζεται σε μια ομάδα ταχέων μετασχηματισμών στον σχεδιασμό, την κατασκευή, τη λειτουργία και την εξυπηρέτηση των βιομηχανικών συστημάτων και προϊόντων. Ο χαρακτηρισμός "4.0" σημαίνει ότι αυτή είναι η τέταρτη βιομηχανική επανάσταση στον κόσμο, ο διάδοχος σε τρεις προηγούμενες βιομηχανικές επαναστάσεις που προκάλεσαν κβαντικά άλματα στην παραγωγικότητα και άλλαξαν τη ζωή των ανθρώπων σε όλο τον κόσμο. Σύμφωνα με τα λόγια της καγκελάριας της Γερμανίας Άνγκελα Μέρκελ, η 4^η βιομηχανική γενιά (Industry 4.0) είναι ο ολοκληρωμένος μετασχηματισμός ολόκληρης της βιομηχανικής παραγωγής μέσω της συγχώνευσης της ψηφιακής τεχνολογίας και του διαδικτύου με τη συμβατική βιομηχανία. Ο όρος "Industry 4.0" προέρχεται από τη Γερμανία, αλλά η ιδέα επικαλύπτει σε μεγάλο βαθμό τις εξελίξεις που μπορεί να επισημαίνονται με διαφορετικό τρόπο σε άλλες ευρωπαϊκές χώρες: τα έξυπνα εργοστάσια, το βιομηχανικό διαδίκτυο των πραγμάτων, η έξυπνη βιομηχανία ή η προηγμένη κατασκευή.

Σύμφωνα με τον John Donovan (2013), το όραμα της 4^{ης} βιομηχανικής γενιάς (Industry 4.0) είναι για τα "συστήματα φυσικής παραγωγής στον κυβερνοχώρο", στα οποία τα «έξυπνα προϊόντα» με αισθητήρα λένε μηχανές για το πώς πρέπει να υποστούν επεξεργασία. Οι διαδικασίες θα κυβερνούσαν τώρα σε ένα αποκεντρωμένο, αρθρωτό σύστημα. Οι έξυπνες ενσωματωμένες συσκευές αρχίζουν να εργάζονται μαζί ασύρματα είτε άμεσα είτε μέσω του cloud Internet - το Διαδίκτυο των πραγμάτων - για να αλλάξουν πάλι την παραγωγή. Τα άκαμπτα, συγκεντρωτικά συστήματα εργοστασιακών ελέγχων δίνουν τη θέση τους στην αποκεντρωμένη νοημοσύνη καθώς η επικοινωνία μηχανή με μηχανή χτυπά στο πάτωμα του συνεργείου. Αυτό είναι το όραμα της 4^{ης} βιομηχανικής γενιάς (Industry 4.0) της Βιομηχανικής Επανάστασης.

Οι Aparna Saisree Thuluvu et al (2017) δήλωσαν ότι η 4^η βιομηχανική γενιά (Industry 4.0) στοχεύει σε μαζική εξατομικευμένη παραγωγή με χαμηλό κόστος και μικρότερο κύκλο παραγωγής, με ψηφιοποίηση και αυτοματοποίηση της παραγωγικής διαδικασίας. Εξηγούν επίσης ότι τα Συστήματα Αυτοματισμού (AS) που χρησιμοποιούνται στη διαδικασία κατασκευής πρέπει να είναι ευέλικτα για την επίτευξη αυτού του στόχου. Μια προσέγγιση βασισμένη στη σημασιολογία για μηχανική και διαμόρφωση χαμηλής προσπάθειας και ανακατασκευή ενός WoT-enabled AS με ελάχιστη ανθρώπινη παρέμβαση καθιστά ένα AS ευέλικτο για μαζική εξατομικευμένη παραγωγή.

Η i-SCOOP είναι μια εταιρεία με έδρα το Βέλγιο που παρέχει εκδόσεις, εκπαιδευτικούς πόρους, εκπαίδευση, πρακτικές συμβουλές και υπηρεσίες σχετικά με το ολοκληρωμένο μάρκετινγκ, την ψηφιακή επιχείρηση, τις μεταμορφώσεις και τις οργανωτικές διαδικασίες. Το i-SCOOP κυκλοφόρησε στον ιστότοπό του ένα άρθρο με τίτλο "Industry 4.0: η τέταρτη βιομηχανική επανάσταση - οδηγός για την Industrie 4.0" όπου αναφέρει ότι ο βασικός στόχος της 4^{ης} βιομηχανικής γενιάς (Industry 4.0) είναι να καταστήσει τις βιομηχανίες παραγωγής και τις συναφείς βιομηχανίες, όπως η εφοδιαστική, ταχύτερη και αποτελεσματικότερη και πιο

πελατο-κεντρική, ενώ συγχρόνως ξεπερνά την αυτοματοποίηση και τη βελτιστοποίηση και εντοπίζει νέες επιχειρηματικές ευκαιρίες και μοντέλα. Σύμφωνα με αυτά, τα μεγαλύτερα οφέλη της 4^{ης} βιομηχανικής γενιάς (Industry 4.0) είναι - προφανώς - παρόμοια με τα οφέλη του ψηφιακού μετασχηματισμού της βιομηχανίας, τη χρήση του IoT στην κατασκευή, τη βελτιστοποίηση των λειτουργικών και επιχειρησιακών διαδικασιών, των οικοσυστημάτων αξίας, των ψηφιακών μετασχηματισμών, Βιομηχανικό Διαδίκτυο και πολλά άλλα.

2.1.2 Οι τεχνολογίες της 4^{ης} βιομηχανικής επανάστασης

Σύμφωνα με την Gradient (2016), ένα ισπανικό ιδιωτικό μη κερδοσκοπικό ίδρυμα που στοχεύει στη δημιουργία γνώσεων στον τομέα των ICT, η πορεία της καινοτομίας στην 4^η βιομηχανική γενιά (Industry 4.0) υπογραμμίζεται από την ψηφιακή σύγκλιση βιομηχανικών και επιχειρηματικών συνιστωσών. Περιγράφουν έξι από τις τεχνολογίες πάνω στις οποίες θα εξελιχθεί το μελλοντικό βιομηχανικό μοντέλο.

- 1. IIoT and cyberphysical systems** – "Η έννοια του IIoT (Industrial Internet of Things) αναφέρεται στη χρήση της τεχνολογίας IoT στις διαδικασίες παραγωγής. Τα cyberphysical συστήματα είναι συσκευές που ενσωματώνουν δυνατότητες επεξεργασίας, καθώς και δυνατότητες αποθήκευσης και επικοινωνίας, προκειμένου να ελέγχουν μία ή περισσότερες φυσικές διεργασίες. Τα κυβερνητικά συστήματα αλληλοσυνδέονται μεταξύ τους και στο παγκόσμιο δίκτυο μέσω του παραδείγματος του IoT "(Gradient, 2016)
- 2. Additive manufacturing, 3D printing** – "Επιτρέπει, μεταξύ άλλων, τον υπερσυντονισμό του προϊόντος -που είναι εγγενές στο πρότυπο Industry 4.0 και στην έννοια της servitilization-, φθηνότερες διαδικασίες κατασκευής προϊόντων ανεξάρτητα από τον αριθμό των παραγόμενων αντικειμένων και καθιστά πολύ ευκολότερη την παραγωγή μικρών σειρά πρωτοτύπων." (Gradient, 2016)

3. Big Data, Data Mining and Data Analytics – "Η ποσότητα των πληροφοριών που αποθηκεύονται επί του παρόντος σε σχέση με διαφορετικές διαδικασίες και συστήματα (τόσο βιομηχανικές όσο και υλικοτεχνικές), τις υπηρεσίες (πωλήσεις, συνδέσεις μεταξύ χρηστών, κατανάλωση ενέργειας κλπ.) ή η κυκλοφορία δεδομένων (logs σε δρομολογητές και εξοπλισμό κλπ. , έτσι δεν είναι δυνατή η χειρωνακτική διαχείριση. Μια περιεκτική ανάλυση αυτών των δεδομένων μπορεί να παρέχει πολύτιμες πληροφορίες σχετικά με αυτές τις διαδικασίες. και μπορεί επίσης να αποτρέψει προβλήματα σε συγκεκριμένες βιομηχανικές διεργασίες μέσω της ανίχνευσης μη φυσιολογικών μέτρων (χωρίς να χρειάζεται προηγουμένως να καθοριστεί σε ποιο βαθμό τα αποτελέσματα μέτρησης είναι ή όχι μη φυσιολογικά) · ή ακόμα και να καθορίσουν τα γεγονότα που σχετίζονται με πιο σύνθετες διαδικασίες, διευκολύνοντας έτσι τη διαχείριση μέσω πρόβλεψης, γνωρίζοντας ότι ένα γεγονός μπορεί να προκαλέσει ένα άλλο με μια ορισμένη πιθανότητα. Οι προσομοιώσεις μπορούν επίσης να προβλέψουν ποιοι πόροι θα χρειαστούν και να βελτιστοποιήσουν τη χρήση αυτών των πόρων αυτομάτως ή να προβλέψουν προληπτικά μελλοντικά γεγονότα και ανάγκες." (Gradient, 2016)

4. Artificial Intelligence – "Σε αυτό το τοπίο μιας βιομηχανίας δεδομένων, υπάρχει μια επείγουσα ανάγκη να εξάγεται η αξία από μεγάλες ποσότητες πληροφοριών, επομένως τα εργαλεία και οι τεχνολογίες που είναι σε θέση να επεξεργαστούν σε πραγματικό χρόνο απαιτούν αυτές τις μεγάλες ποσότητες πληροφοριών, είναι σε θέση να μάθουν ανεξάρτητα από τις πληροφορίες που λαμβάνουν, ανεξάρτητα από τις πηγές και τις αντιδράσεις των χρηστών και των χειριστών (τεχνολογίες μηχανικής μάθησης ή βαθιάς εκμάθησης)." (Gradient, 2016)

5. Collaborative robotics (Cobot) – "Αυτός ο όρος ορίζει μια νέα γενιά ρομπότ που συνεργάζεται στενά με τους ανθρώπους, χωρίς τους συνήθως

απαιτούμενους περιορισμούς ασφαλείας που εφαρμόζονται στις τυπικές εφαρμογές βιομηχανικής ρομποτικής. Χαρακτηρίζεται, μεταξύ άλλων, για την ευελιξία, την προσβασιμότητα και τη σχετική ευκολία προγραμματισμού.” (Gradient, 2016)

6. Virtual Reality and Augmented Reality – “Χάρη στην αυξημένη προσβασιμότητα και την ευκολία χρήσης αυτών των τεχνολογιών τα τελευταία χρόνια, τα VR και AR έχουν καταστεί χρήσιμα εργαλεία για τη βελτιστοποίηση των σχεδίων, την αυτοματοποίηση των διαδικασιών, τον έλεγχο της κατασκευής και της κατασκευής, την κατάρτιση των εργαζομένων και για δραστηριότητες συντήρησης και παρακολούθησης περιβάλλοντα.” (Gradient, 2016)

Σύμφωνα με το ισπανικό εθνικό ινστιτούτο για την ασφάλεια στον κυβερνοχώρο (INCIBE) (2015), η ασφάλεια στον κυβερνοχώρο της 4^{ης} βιομηχανικής γενιάς (Industry 4.0) έχει άμεση επίδραση στα συστήματα φυσικών συστημάτων Cyber, το Διαδίκτυο των πραγμάτων και το Διαδίκτυο των Υπηρεσιών.

- **“CPSs (Cyber Physical Systems):** “Είναι συστήματα που παρακολουθούν τα φυσικά συστήματα, δημιουργούν ένα εικονικό αντίγραφο και κάνουν αποκεντρωμένες αποφάσεις. Το σύστημα αισθητήρων και τα στοιχεία ελέγχου είναι ικανά να συνδέουν τα μηχανήματα και τις συσκευές με εγκαταστάσεις, δοχεία, δίκτυα και ανθρώπους.” (INCIBE, 2015)
- **“Internet of Things (IoT):** “Οι CPS ενεργούν στο Διαδίκτυο των πραγμάτων. Αυτό το παράδειγμα διασφαλίζει ότι οι συσκευές και τα στοιχεία είναι συνδεδεμένα σε πρωτόκολλα που βασίζονται στο Internet (TCP / IP).” (INCIBE, 2015)

- **“Internet of Services (IoS):** “Αυτό το παράδειγμα προσφέρει πολλές εσωτερικές εργοστασιακές υπηρεσίες οριζόντια σε όλη την αλυσίδα αξίας.” (INCIBE, 2015)

Αυτά είναι τα χαρακτηριστικά γνωρίσματα της 4^{ης} βιομηχανικής γενιάς (Industry 4.0), από την άποψη της τεχνολογίας και των διαδικασιών. Ο καθένας φέρνει το δικό του φάσμα απειλών, τρωτών σημείων και πιθανών προκλήσεων / διασφαλίσεων όσον αφορά την ασφάλεια στον κυβερνοχώρο.

2.2 Σημασιολογικός Ιστός

2.2.1 Τι είναι ο σημασιολογικός ιστός

Ο Tim Berners-Lee, εφευρέτης του World Wide Web, δημιούργησε τον όρο "Σημασιολογικός Ιστός" εξηγώντας ότι δεν είναι ένας ξεχωριστός ιστός αλλά μια επέκταση του σημερινού, όπου οι πληροφορίες δίδονται με σαφώς καθορισμένο νόημα, επιτρέποντας καλύτερα στους υπολογιστές και τους ανθρώπους να δουλεύουν σε συνεργασία. Ο Tim Berners-Lee et al (2001) δήλωσε επίσης ότι ο σημασιολογικός ιστός θα φέρει δομή στο ουσιαστικό περιεχόμενο των ιστοσελίδων, δημιουργώντας ένα περιβάλλον όπου οι πράκτορες λογισμικού που μετακινούνται από σελίδα σε σελίδα μπορούν εύκολα να εκτελούν εξελιγμένες εργασίες για τους χρήστες.

Το Cambridge Semantics, μια εταιρία λογισμικού διαχείρισης μεγάλων δεδομένων και εξερευνητικής αναλυτικής λογισμικού με σημασιολογικές εφαρμογές ιστού, ανέφερε στην ιστοσελίδα τους ότι ο Σημασιολογικός Ιστός που ονομάζεται επίσης Web 3.0, Linked Data Web, Web of Data αποτελεί την επόμενη σημαντική εξέλιξη των πληροφοριών σύνδεσης. Επιτρέπει τη σύνδεση δεδομένων από πηγή σε οποιαδήποτε άλλη πηγή και να γίνεται κατανοητή από τους υπολογιστές ώστε να μπορούν να εκτελούν ολοένα και πιο εξελιγμένα καθήκοντα για λογαριασμό μας.

Οι Grigoris Antoniou et al (2012) δήλωσαν ότι το γενικό όραμα ενός "σημασιολογικού ιστού" μπορεί να συνοψιστεί σε μια ενιαία φράση: "να γίνει ο διαδικτυακός κόσμος πιο προσβάσιμος στους υπολογιστές". Ανέφερε επίσης ότι ο τρέχων ιστός είναι ένας ιστός κειμένου και εικόνων. Όλες οι έξυπνες εργασίες (επιλογή, συνδυασμός, συγκέντρωση δεδομένων κ.λπ.) πρέπει να γίνουν από ανθρώπους. Αν θέλουμε να κάνουμε τον ιστό πιο προσβάσιμο στους υπολογιστές, πρέπει να σταματήσουμε να περιορίσουμε τον ρόλο τους στον τρέχοντα ιστό. Αυτή τη στιγμή, απλώς προσθέτουν λέξεις-κλειδιά και πληροφορίες πλοίων από διακομιστές σε πελάτες, και αυτό είναι όλο.

Η Techopedia, ένας ιστότοπος για την εκπαίδευση ηλεκτρονικών υπολογιστών, περιγράφει στον ιστότοπό της σημασιολογικό ιστό ως ένα δίκτυο στοιχείων που συνδέονται με τέτοιο τρόπο ώστε να μπορούν εύκολα να υποβάλλονται σε επεξεργασία από μηχανές αντί για ανθρώπινους χειριστές. Περιλαμβάνουν επίσης ότι ο βασικός στόχος του Σημασιολογικού Ιστού είναι να ενεργοποιήσει την εξέλιξη του υπάρχοντος ιστού, ώστε να επιτρέπει στους χρήστες να αναζητούν, να ανακαλύπτουν, να μοιράζονται και να ενώνουν πληροφορίες με λιγότερες προσπάθειες.

2.2.2 Οι τεχνολογίες του σημασιολογικού ιστού

Υπάρχει ένα σύνολο τεχνολογιών, εργαλείων και προτύπων που υποστηρίζουν ολόκληρο το όραμα του σημασιολογικού ιστού και σύμφωνα με τον Tim Berners-Lee et al (2001), οι σημαντικότερες τεχνολογίες είναι η εκτατή γλώσσα σήμανσης (XML) και το πλαίσιο περιγραφής πόρων (RDF).

Οι Tim Berners-Lee et al (2001) εξήγησαν ότι η XML επιτρέπει σε όλους να δημιουργήσουν τις δικές τους ετικέτες. Τα σενάρια ή τα προγράμματα μπορούν να κάνουν χρήση αυτών των ετικετών με εξελεγμένους τρόπους, αλλά ο συγγραφέας

script πρέπει να γνωρίζει τι χρησιμοποιεί ο κάθε συγγραφέας σελίδας για κάθε ετικέτα. Εν ολίγοις, η XML επιτρέπει στους χρήστες να προσθέσουν αυθαίρετη δομή στα έγγραφά τους, αλλά δεν λέει τίποτα για το τι σημαίνουν οι δομές.

Ο Allen Taylor (2015) αναφέρεται στο RDF, που ονομάζεται επίσης παγκόσμιο πλαίσιο για τη σημασιολογία και είναι ένα πρότυπο για τη μοντελοποίηση δεδομένων που χρησιμοποιεί τρία απλά στοιχεία: ένα θέμα, ένα πρόβατο και ένα αντικείμενο. Για το λόγο αυτό, οι άνθρωποι συχνά καλούν απλά δεδομένα RDF τριπλά. Όταν συνδέονται τα τριπλάσια, το αντικείμενο ενός τριπλού είναι το αντικείμενο άλλου τριπλού, σχηματίζουν μια γραφική παράσταση δεδομένων με κόμβους και άκρα που δεν έχουν ιεραρχία και είναι μηχανικά αναγνώσιμα, εύκολο να μοιράζονται και να συνδυάζονται εύκολα. Είναι εύκολο να μοιραστούν και να συνδυαστούν επειδή κάθε τρίκλινο είναι ατομικό, πράγμα που σημαίνει ότι δεν μπορεί να αναλυθεί περαιτέρω χωρίς να καταστραφεί το νόημα του γεγονότος που αντιπροσωπεύει.

Η Ontotext, εταιρεία λογισμικού που αναπτύσσει προϊόντα λογισμικού και λύσεις βασισμένες σε γλώσσες και πρότυπα σημασιολογικού ιστού, ανέφερε άλλα δύο σημαντικά στοιχεία του σημασιολογικού ιστού. Σύμφωνα με αυτά, το ένα είναι το URI που είναι μια ακολουθία χαρακτήρων που περιγράφουν πόρους όπως εντοπιστές (URL) ή ονόματα (URN). Τα URI χρησιμοποιούνται για να δώσουν μοναδικά ονόματα σε οτιδήποτε - από το ψηφιακό περιεχόμενο που διατίθεται στο Web σε αντικείμενα πραγματικού κόσμου και αφηρημένες έννοιες. Είναι το δομικό στοιχείο του RDF. Η δεύτερη είναι η SPARQL, η οποία είναι μόνιμη για το πρωτόκολλο SPARQL και τη γλώσσα ερωτημάτων RDF. Χρησιμοποιείται για την αναζήτηση, ανάκτηση και χειρισμό δεδομένων που είναι αποθηκευμένα σε μορφή RDF. Ως εκ τούτου, μας επιτρέπει να ψάξουμε στο Web των Δεδομένων (ή οποιαδήποτε βάση δεδομένων) και να ανακαλύψουμε σχέσεις. Με άλλα λόγια, εάν εξετάσουμε τον Σημασιολογικό Ιστό ως μια παγκόσμια συλλογή βάσεων δεδομένων, το SPARQL μας δίνει τη δυνατότητα να επικεντρωθούμε σε αυτό που θα

θέλαμε να μάθουμε αντί για τον τρόπο με τον οποίο οργανώνεται μια βάση δεδομένων.

Ο Rene Millman (2017) δήλωσε ότι ένα από τα βασικά οφέλη του σημασιολογικού ιστού είναι ότι έχει μεγάλα ποσά δεδομένων, γνώσεων και πληροφοριών που έγιναν κατανοητά και προσβάσιμα σε μηχανές, ειδικά τεχνητά νοήμονα bots, εικονικούς βοηθούς και πράκτορες.

Σύμφωνα με τον Tim Berners-Lee et al (2001), η πραγματική δύναμη του Σημασιολογικού Ιστού θα πραγματοποιηθεί όταν οι άνθρωποι δημιουργούν πολλά προγράμματα που συλλέγουν το περιεχόμενο Ιστού από διαφορετικές πηγές, επεξεργάζονται τις πληροφορίες και ανταλλάσσουν τα αποτελέσματα με άλλα προγράμματα. Η αποτελεσματικότητα τέτοιων πρακτόρων λογισμικού θα αυξηθεί εκθετικά, καθώς θα είναι διαθέσιμα περισσότερο ευανάγνωστο από το μηχάνημα περιεχόμενο στο Web και αυτοματοποιημένες υπηρεσίες (συμπεριλαμβανομένων και άλλων παραγόντων). Ο Σημασιολογικός Ιστός προωθεί αυτή τη συνέργεια: ακόμη και οι πράκτορες που δεν σχεδιάστηκαν ρητά για να συνεργαστούν μπορούν να μεταφέρουν δεδομένα μεταξύ τους όταν τα δεδομένα έρχονται με σημασιολογία. Ένα σημαντικό γεγονός για τη λειτουργία των πρακτόρων θα είναι η ανταλλαγή «αποδείξεων» γραμμένων στην ενοποιητική γλώσσα του Σημασιολογικού Ιστού (η γλώσσα που εκφράζει λογικές συμπεράσματα με κανόνες και πληροφορίες όπως αυτές που ορίζονται από τις οντολογίες).

Ένα άλλο ζωτικό χαρακτηριστικό που παρουσιάζουν οι Tim Berners-Lee et al (2001) θα είναι οι ψηφιακές υπογραφές, οι οποίες είναι κρυπτογραφημένες μονάδες δεδομένων που μπορούν να χρησιμοποιήσουν υπολογιστές και πράκτορες για να επαληθεύσουν ότι οι συνημμένες πληροφορίες έχουν παρασχεθεί από μια συγκεκριμένη αξιόπιστη πηγή. Ο Σημασιολογικός Ιστός, αντίθετα, είναι πιο ευέλικτος. Οι καταναλωτές και οι παραγωγοί πράκτορες μπορούν να επιτύχουν μια κοινή κατανόηση με την ανταλλαγή οντολογιών, οι οποίες παρέχουν το λεξιλόγιο

που απαιτείται για συζήτηση. Ισχυρίστηκε επίσης ότι ορισμένοι πράκτορες θα εκμεταλλευτούν τεχνολογίες τεχνητής νοημοσύνης επιπλέον του Σημασιολογικού Ιστού. Το Σημασιολογικό Δίκτυο θα παράσχει τα θεμέλια και το πλαίσιο για να καταστήσει τέτοιες τεχνολογίες πιο εφικτές.

2.2.3 Οντολογίες

Ο Tim Berners-Lee et al (2001) προσδιόρισε τις οντολογίες ως βασικό συστατικό του Σημασιολογικού Ιστού που περιγράφονται ως συλλογές πληροφοριών. Το πιο χαρακτηριστικό είδος οντολογίας για το Web έχει μια ταξινόμηση και ένα σύνολο κανόνων συμπερασμάτων. Η ταξινομία ορίζει κατηγορίες αντικειμένων και σχέσεις μεταξύ τους. Οι κλάσεις, οι υποκατηγορίες και οι σχέσεις μεταξύ των οντοτήτων είναι ένα πολύ ισχυρό εργαλείο για χρήση στο Web. Μπορούμε να εκφράσουμε ένα μεγάλο αριθμό σχέσεων μεταξύ οντοτήτων, αναθέτοντας ιδιότητες στις τάξεις και επιτρέποντας σε υπο-κλάσεις να κληρονομήσουν τέτοιες ιδιότητες. Οι κανόνες συμπερασμάτων στις οντολογίες παρέχουν περαιτέρω ισχύ. Ο υπολογιστής δεν κατανοεί πραγματικά καμία από αυτές τις πληροφορίες, αλλά μπορεί πλέον να χειριστεί τους όρους πολύ πιο αποτελεσματικά με τρόπους που είναι χρήσιμοι και χρήσιμοι για τον ανθρώπινο χρήστη. Με τις σελίδες οντολογιών στον Ιστό, αρχίζουν να αναδύονται λύσεις σε ορολογικά (και άλλα) προβλήματα. Η έννοια των όρων ή των κωδικών XML που χρησιμοποιούνται σε μια ιστοσελίδα μπορεί να οριστεί από τους δείκτες από τη σελίδα σε μια οντολογία. Οι οντολογίες μπορούν να βελτιώσουν τη λειτουργία του ιστού με πολλούς τρόπους. Μπορούν να χρησιμοποιηθούν με απλό τρόπο για να βελτιώσουν την ακρίβεια των αναζητήσεων στο Web - το πρόγραμμα αναζήτησης μπορεί να αναζητήσει μόνο εκείνες τις σελίδες που αναφέρονται σε μια ακριβή αντί για όλες αυτές που χρησιμοποιούν διαφορετικές λέξεις-κλειδιά. Οι πιο προηγμένες εφαρμογές θα χρησιμοποιούν οντολογίες για τη συσχέτιση των πληροφοριών μιας σελίδας με τις σχετικές δομές γνώσης και τους κανόνες συμπερασμάτων.

Οι Zeeshan Ahmed και Detlef Gerhard (2010) δήλωσαν ότι ο Σημασιολογικός Ιστός είναι μια έξυπνη ενσάρκωση και εξέλιξη στον Παγκόσμιο Ιστό για τη συλλογή, χειρισμό και σχολιασμό των πληροφοριών, παρέχοντας κατηγοριοποίηση, ομοιόμορφη πρόσβαση στους πόρους και δόμηση των πληροφοριών σε μηχανικά επεξεργάσιμη μορφή. Η Οντολογία είναι η συλλογή αλληλοσυνδεόμενων σημασιολογικά διαμορφωμένων εννοιών βασισμένων σε ήδη καθορισμένα πεπερασμένα σύνολα όρων και εννοιών που χρησιμοποιούνται στην ενοποίηση της πληροφορίας και τη διαχείριση γνώσης. Για την απόκτηση των επιθυμητών αποτελεσμάτων, η οντολογία κατηγοριοποιείται σε τρεις υπο-κατηγορίες οντολογιών: Οντολογία Φυσικής Γλώσσας (NLO), Οντολογία Περιοχής (DO) και Ontology Instance (OI). Το NLO δημιουργεί σχέσεις μεταξύ των δημιουργημένων λεξικών μαρκών που λαμβάνονται από τις δηλώσεις φυσικής γλώσσας, το DO περιέχει τη γνώση ενός συγκεκριμένου τομέα και το OI δημιουργεί αυτόματες αντικειμενοστρεφείς ιστοσελίδες. Οι οντολογίες κατασκευάζονται με τη χρήση ορισμένων οντολογιών που υποστηρίζουν γλώσσες όπως RDF, OWL κλπ. Και συνδέονται μεταξύ τους με αποκεντρωμένο τρόπο ώστε να εκφράζουν σαφώς τα σημασιολογικά περιεχόμενα και να οργανώνουν σημασιολογικά όρια για να εξάγουν συγκεκριμένες πληροφορίες.

Ο Nidhi Malik (2015) εξήγησε ότι οι οντολογίες χρησιμοποιούνται ευρέως για την παροχή σημασιολογικά άφθονων λεξιλογίων και μετα-δεδομένων για την έκφραση και την ανακάλυψη πληροφοριών. Ο λόγος για τη δημοτικότητα των οντολογιών είναι ότι παρέχουν κοινή και κοινή κατανόηση ενός τομέα που μπορεί να μεταδοθεί σε ανθρώπους και υπολογιστές.

Οι Patrick Lambrix και He Tan (2005) κατέληξαν στο συμπέρασμα ότι οι οντολογίες αποτελούν σημαντική τεχνολογία για τον Σημασιολογικό Ιστό. Σε διαφορετικές περιοχές έχουν ήδη αναπτυχθεί οντολογίες και πολλές από αυτές τις οντολογίες περιέχουν επικαλυπτόμενες πληροφορίες.

Ο V. Richard Benjamins (2016) δήλωσε ότι χωρίς τις οντολογίες ή τις εννοιολογίσεις που αποτελούν τη βάση της γνώσης, δεν μπορεί να υπάρχει λεξιλόγιο για την εκπροσώπηση της γνώσης. Οι οντολογίες επιτρέπουν την ανταλλαγή γνώσεων που μας επιτρέπουν να χτίσουμε συγκεκριμένες βάσεις γνώσεων που περιγράφουν συγκεκριμένες καταστάσεις.

2.3 Συγχώνευση Οντολογιών

Σύμφωνα με την Maria Ganzha et al (2014), υπάρχει ένα μεγάλο μέρος επιστημονικής βιβλιογραφίας που αφιερώνεται στην οντολογία που ταιριάζει, ευθυγραμμίζει, χαρτογραφεί τη μετάφραση και τη συγχώνευση. Το επίπεδο δραστηριότητας παραμένει υψηλό και το ενδιαφέρον αυξάνεται. Τα αποτελέσματα από τις δοκιμές είναι κάπως αποθαρρυντικά. Η πλειοψηφία των υφιστάμενων αποτελεσμάτων έρευνας είναι καθαρά θεωρητική, δηλαδή τελειώνει με δημοσίευση, αλλά δεν διατίθενται εργαλεία. Ένας μεγάλος αριθμός εργαλείων εξαφανίζεται / σταματάει να διατηρείται εντός 1-2 ετών από τη δημοσίευση του τελευταίου εγγράφου.

2.3.1 Ορισμοί

"Η οντολογική ευθυγράμμιση αναφέρεται στην εύρεση αντιστοιχιών μεταξύ δύο ή περισσότερων οντολογιών. Το αποτέλεσμα αυτής της διαδικασίας είναι μια ευθυγράμμιση - ένα σύνολο αντιστοιχιών μεταξύ οντοτήτων (ατομική ευθυγράμμιση) ή ομάδων οντοτήτων και υπο-δομών (σύνθετη ευθυγράμμιση) από διαφορετικές οντολογίες. Μια αντιστοιχία μπορεί να είναι είτε ένα κατηγορημα για την ομοιότητα, που ονομάζεται αντιστοίχιση, είτε ένα λογικό αξίωμα - μια χαρτογράφηση. Τυπικά, τα χρησιμοποιούμενα αξιώματα χαρτογράφησης είναι η

ισοδυναμία και η συνύπαρξη. Στην πράξη, τα εργαλεία ευθυγράμμισης οντολογίας συχνά δηλώνουν ένα βαθμό εμπιστοσύνης για κάθε αλληλογραφία στη χαρτογράφηση. Ένα αξίωμα ισοδυναμίας με ένα βαθμό εμπιστοσύνης είναι πολύ στενό στο νόημα σε ένα πρόβατο για την ομοιότητα (μια αντιστοίχιση). Οι όροι "ring χάρτη" και "ταιριάζουν" συχνά δεν διακρίνεται στην ορολογία που χρησιμοποιείται από τα εργαλεία ευθυγράμμισης. Ένα σύνολο αντιστοιχιών μπορεί να ονομαστεί "ευθυγράμμιση", "αντιστοίχιση" ή "χαρτογράφηση" πρακτικά εναλλακτικά. "(Maria Ganzha et al, 2014)

“Η συγχώνευση οντολογιών είναι μια διαδικασία συνδυασμού δύο ή περισσότερων οντολογιών σε μία. Συνεπώς, η προκύπτουσα οντολογία αποθηκεύει γνώσεις από όλες τις συγχωνευμένες. Η συγχώνευση συχνά χρησιμοποιεί ένα σύνολο ευθυγραμμίσεων για να δημιουργήσει βαθιές διασυνδέσεις μεταξύ οντολογιών και, τελικά, να τις συγχωνεύσει σε μία.” (Maria Ganzha et al, 2014)

“Η συγχώνευση οντολογιών χρησιμοποιεί ευθυγράμμιση για να κάνει μια οντολογία από δύο ή περισσότερες οντολογίες από συναφείς τομείς. Συνήθως στην οντολογία που αποκτάται, δεν είναι δυνατόν να προσδιοριστεί η πηγή οντολογίας του συγκεκριμένου στοιχείου. Το κύριο μέρος της συγχώνευσης οντολογιών είναι η οντολογική αντιστοίχιση, η οποία σχετίζεται με την αντιστοιχία ή τον καθορισμό της σχέσης μεταξύ στοιχείων διαφορετικών οντολογιών. Ως αποτέλεσμα της αντιστοίχισης, υπάρχει ευθυγράμμιση, που περιέχει το σύνολο αντιστοιχιών μεταξύ στοιχείων οντολογιών.” (Vita Graudina et al, 2012)

2.3.2 Εργαλεία/Προσεγγίσεις

Μια ενδιαφέρουσα προσέγγιση σχετικά με τη συγχώνευση οντολογιών έγινε από τον Hovy (1998). Περιγράφονται πολλά θεωρητικά στοιχεία για την αναγνώριση των αντιστοιχών εννοιών σε διαφορετικές οντολογίες, π.χ. συγκρίνοντας τα ονόματα και

τους ορισμούς της φυσικής γλώσσας για δύο έννοιες και ελέγχοντας την εγγύτητα δύο έννοιων στην ιεραρχία της έννοιας.

Το σύστημα OntoMorph από τον Chalupsky (2000) προσφέρει δύο είδη μηχανισμών για τη μετάφραση και τη συγχώνευση των οντολογιών. Η συντακτική επανεγγραφή υποστηρίζει τη μετάφραση μεταξύ δύο διαφορετικών γλωσσών εκπροσώπησης της γνώσης, η σημασιολογική επανεγγραφή προσφέρει μέσα για μετασχηματισμούς που βασίζονται σε συναγωγή. Επιτρέπει ρητά να παραβιάζεται η διατήρηση της σημασιολογίας στην ανταλλαγή απόψεων για έναν πιο ευέλικτο μηχανισμό μετασχηματισμού.

Οι McGuinness et al (2000) περιγράφουν το σύστημα Chimaera. Το σύστημα αυτό παρέχει υποστήριξη για τη συγχώνευση οντολογικών όρων από διαφορετικές πηγές, για τον έλεγχο της κάλυψης και την ορθότητα των οντολογιών και για τη διατήρηση των οντολογιών με την πάροδο του χρόνου. Το Chimaera προσέφερε επίσης μια ευρεία συλλογή λειτουργιών, αλλά οι υποκείμενες υποθέσεις σχετικά με τις δομικές ιδιότητες των οντολογιών δεν είναι σαφείς.

Σύμφωνα με τους Noy και Musen (2000), ο Prompt είναι ένας αλγόριθμος συγχώνευσης και ευθυγράμμισης οντολογιών που ενσωματώνεται στο Protégé (2000), ένας ελεύθερος επεξεργαστής οντολογιών ανοιχτού κώδικα και ένα σύστημα διαχείρισης γνώσης. Αρχίζει με την αναγνώριση των ονομάτων τάξεων που ταιριάζουν. με βάση αυτό το αρχικό βήμα πραγματοποιείται επαναληπτική προσέγγιση για την πραγματοποίηση αυτόματων ενημερώσεων, εύρεση συγκρουόμενων αποτελεσμάτων και προτάσεις για την κατάργηση αυτών των συγκρούσεων.

Ο Ernesto Jimenez-Ruiz και ο Bernardo Cuenca Grau (2011) δήλωσαν ότι το LogMap είναι το μόνο σύστημα που ταιριάζει με σημασιολογικά πλούσιες οντολογίες που περιέχουν δεκάδες (και εκατοντάδες) χιλιάδες τάξεις. Το LogMap δημιουργήθηκε

στην ομάδα "Γνώση Αντιπροσώπευσης και Λογοτεχνίας" του Τμήματος Επιστήμης Υπολογιστών του Πανεπιστημίου της Οξφόρδης από τον Ernesto Jiménez-Ruiz, τον Bernardo Cuenca Grau και τον Ian Horrocks. Το LogMap εφαρμόζει επίσης αλγόριθμους για ανίχνευση και επισκευή μη ικανοποιητικής ανάλυσης "εν πτήσει". Τα πειράματά τους με τις οντολογίες NCI, FMA και SNOMED CT επιβεβαίωσαν ότι το σύστημά τους θα μπορούσε να ταιριάζει αποτελεσματικά ακόμη και με τις μεγαλύτερες υπάρχουσες βιο-ιατρικές οντολογίες. Επιπλέον, το LogMap είναι σε θέση να παράγει σε πολλές περιπτώσεις ένα "καθαρό" σύνολο χαρτογραφικών εξόδων, υπό την έννοια ότι η οντολογία που προκύπτει από την ενσωμάτωση των αντιστοιχιών εξόδου του LogMap με τις οντολογίες εισόδου είναι συνεπής και δεν περιέχει μη ικανοποιητικές κατηγορίες.

Το Πανεπιστήμιο της Λειψίας ανέπτυξε το COMA 3.0, ένα εργαλείο σχεδίασης και οντολογίας. Επεκτείνει τα προηγούμενα πρωτότυπα COMA και COMA ++ με βελτιωμένη διαχείριση ροής εργασίας και πρόσθετα χαρακτηριστικά γνωρίσματα όπως οντολογία συγχώνευσης. Επιπλέον, προσφέρει μια ολοκληρωμένη υποδομή για την επίλυση μεγάλων προβλημάτων πραγματικού κόσμου. Το γραφικό περιβάλλον προσφέρει μια ποικιλία αλληλεπιδράσεων, επιτρέποντας στον χρήστη να επηρεάζει τη διαδικασία αντιστοίχισης με πολλούς τρόπους. Η λειτουργία COMA 3.0 χρησιμοποιείται στο νέο πρωτότυπο QuickMig που εστιάζει στη δημιουργία εκτελέσιμων χαρτογραφήματων για τη μετανάστευση δεδομένων. Το νέο COMA 3.0 ενσωματώνει επίσης το πρωτότυπο ATOM για αυτόματη συγχώνευση σχημάτων και οντολογιών.

Κεφάλαιο 3: Μεθοδολογία

Το ερευνητικό υλικό αυτής της εργασίας σχετικά με τη συγχώνευση οντολογιών είναι σχετικές μελέτες, άρθρα και βιβλιογραφίες. Η Μηχανή Αναζήτησης Google και ο Μελετητής Google χρησιμοποιήθηκαν για την αναζήτηση άρθρων και βιβλιογραφιών. Οι λέξεις-κλειδιά που χρησιμοποιήθηκαν για την έρευνα στον τομέα του Σημασιολογικού Ιστού ήταν "σημασιολογία", "σημασιολογική διαδικτυακή" και "τεχνολογίες σημασιολογικού ιστού". Για τη Βιομηχανία 4.0 χρησιμοποιήσαμε "βιομηχανία 4.0", "βιομηχανία 4.0 τεχνολογίες", και για τη συγχώνευση οντολογιών χρησιμοποιήσαμε "ενοποίηση οντολογιών", "συγχωνεύοντας οντολογίες για σημασιολογικό ιστό", "πώς να συγχωνεύσουμε δύο οντολογίες" και "εργαλεία ενοποίησης οντολογιών".

Όπως αναφέρθηκε στη βιβλιογραφία, υπάρχει ένας μεγάλος αριθμός ενοποιημένων εργαλείων οντολογίας που εξαφανίζονται / σταματούν να διατηρούνται εντός 1-2 ετών από τη δημοσίευσή τους. Μετά την ανάγνωση της σχετικής βιβλιογραφίας σχετικά με τη συγχώνευση οντολογιών, την επιλογή εργαλείων / πλαισίων που απάντησαν επιτυχώς στην ερευνητική ερώτηση. "Το εργαλείο αυτό / το πλαίσιο είναι ικανό να συγχωνεύσει δύο οντολογίες;" πραγματοποιήθηκε και συνεπώς διεξήχθησαν δοκιμές συγχώνευσης οντολογιών. Αυτά τα εργαλεία και τα πλαίσια που εντοπίστηκαν ήταν Coma 3.0, LogMap και Protégé's build in merging function.

Για να εκτελέσει τη δοκιμή συγχώνευσης οντολογιών, ο ερευνητής έχει ελέγξει τις οντολογίες που θα μπορούσε να χρησιμοποιήσει για συγχώνευση συγκρίσεων στη βάση δεδομένων οντολογιών που παρέχεται από την Infineon Technologies. Δεν μπορούσε να βρει δύο οντολογίες που ήταν κατάλληλες για τη δοκιμή. Ως εκ τούτου,

ο ερευνητής αποφάσισε να χρησιμοποιήσει μία από τις οντολογίες που κυκλοφόρησε στο Infineon με τίτλο "Δημοσιεύσεις IFX" από την Lisa Kredig, η οποία εργαζόταν στην Infineon Technologies AG περίπου την ίδια περίοδο με τον ερευνητή. Τα πλεονεκτήματα αυτής της οντολογίας είναι ότι βασίζεται σε πληροφορίες Infineon που έχουν ήδη δημοσιευτεί εκτός της εταιρείας και είναι δυνατό να αναπτυχθεί μια οντολογία με παρόμοιο θέμα. Αυτό το ερευνητικό έργο ανέπτυξε μια οντολογία που ονομάζεται "έργα διδακτικών έργων IFX" κατά τη διάρκεια του χρόνου του ως ασκούμενος στην Infineon Technologies AG. Οι λεπτομέρειες αυτών των δύο οντολογιών θα περιγραφούν στην επόμενη ενότητα.

Η πρώτη απόπειρα συγχώνευσης οντολογιών ήταν στο COMA 3.0, ένα εργαλείο αντιστοίχισης σχήματος και οντολογίας που αναπτύχθηκε στο Πανεπιστήμιο της Λειψίας. Το COMA 3.0 φορτώνει τις δύο οντολογίες μέσω του GUI, εκτελεί μια αντιστοίχιση οντολογιών και στη συνέχεια τα αποτελέσματα εμφανίζονται στην οθόνη του χρήστη. Η κοινοτική έκδοση δεν υποστηρίζει την οντολογική ικανότητα συγχώνευσης, αυτό είναι διαθέσιμο μόνο στην επιχειρηματική έκδοση του COMA 3.0. Ο ερευνητής προσπάθησε να επικοινωνήσει με την εταιρεία στις 24 Αυγούστου 2018 για να αγοράσει το COMA 3.0 Business Edition στέλνοντας ένα μήνυμα ηλεκτρονικού ταχυδρομείου στη διεύθυνση "WebData Solutions" (info@webdata-solutions.com) αλλά δεν έλαβε καμία απάντηση μέχρι το 18ο Σεπτέμβριος 2018. Ο ερευνητής τελικά έτρεξε τις δοκιμές του στο LogMap μέσω της εφαρμογής διεπαφής ιστού και με τη συγχώνευση του Protégé.

Η προσδοκία των αποτελεσμάτων του ερευνητή από αυτά τα εργαλεία συγχώνευσης οντολογιών είναι η δημιουργία μιας νέας οντολογίας όπου όλες οι ανάλογες ιδιότητες Κατηγοριών, αντικειμένων και δεδομένων μεταξύ των δημοσιεύσεων IFX και της διατριβής IFX θα αντιστοιχούν μεταξύ τους. Ο χρόνος, η προσπάθεια και η ποιότητα των αποτελεσμάτων θα αξιολογηθούν για να επιλέξετε το καλύτερο.

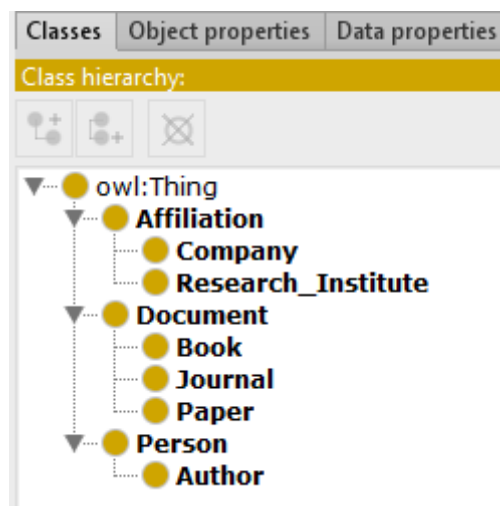
3.1 Δεδομένα που θα χρησιμοποιηθούν

3.1.1 Οντολογία δημοσιεύσεων IFX

Αυτή η οντολογία αναπτύχθηκε για να περιγράψει τα εξωτερικά έντυπα δημοσίευσης της Infineon Technologies. Οι κλάσεις, οι ιδιότητες αντικειμένου και δεδομένων αυτής της οντολογίας που παρουσιάζονται στα ακόλουθα σχήματα 1, 2 και 3 όπως φαίνονται στο Protégé.

Κλάσεις:

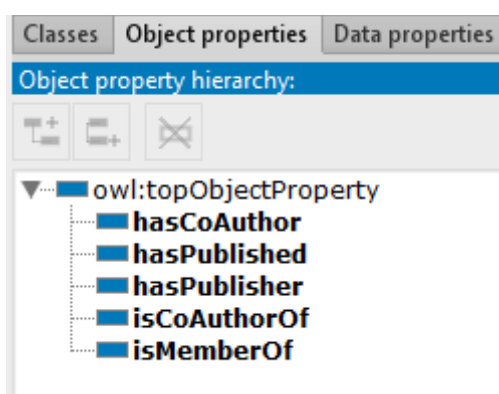
- “Affiliation”, η οποία περιέχει τις κλάσεις “Company” και “Research Institute” ως υπο-κλάσεις.
- “Document”, η οποία περιέχει τις κλάσεις “Book”, “Journal” και “Paper” ως υπο-κλάσεις.
- “Person”, η οποία περιέχει την κλάση “Author” ως υπο-κλάση.



Εικόνα 1: Οντολογία δημοσιεύσεων IFX – Κλάσεις

Ιδιότητες Αντικειμένου:

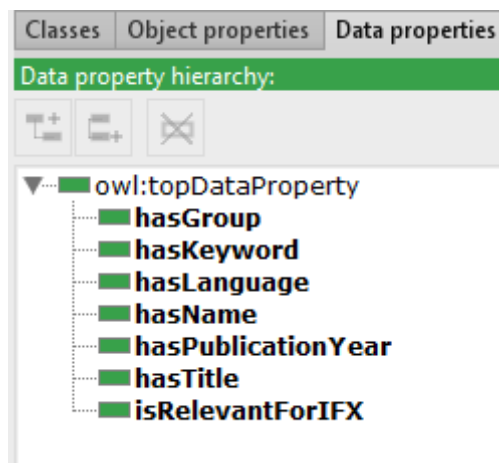
- Το "hasCoAuthor" σχετίζει το έγγραφο δημοσίευσης με τον συν-συγγραφέα του. Η κλάση εγγράφου είναι το πεδίο ορισμού, η κλάση του συγγραφέα η εμβέλεια και είναι αντίστροφη της ιδιότητας αντικειμένου isCoAuthorOf.
- Το "hasPublished" σχετίζει τον συγγραφέα με το έγγραφο δημοσίευσης του. Η κλάση του συγγραφέα είναι το πεδίο ορισμού, η κλάση του εγγράφου η εμβέλεια και είναι αντίστροφη της ιδιότητας αντικειμένου hasPublisher.
- Το "hasPublisher" σχετίζει το έγγραφο δημοσίευσης με τον συγγραφέα του. Η κλάση εγγράφου είναι το πεδίο ορισμού, η κλάση του συγγραφέα η εμβέλεια και είναι αντίστροφη της ιδιότητας αντικειμένου hasPublished.
- "isCoAuthorOf" σχετίζει τον συν-συγγραφέα με το έγγραφο δημοσίευσης του. Η κλάση του συγγραφέα είναι το πεδίο ορισμού, η κλάση του εγγράφου η εμβέλεια και είναι αντίστροφη της ιδιότητας αντικειμένου hasCoAuthor.
- "isMemberOf" σχετίζει τον συγγραφέα με τον δεσμό που έχει. Η κλάση του συγγραφέα είναι το πεδίο ορισμού και η κλάση του δεσμού η εμβέλεια του.



Εικόνα 2: Οντολογία δημοσιεύσεων IFX – Ιδιότητες Αντικειμένων

Ιδιότητες Δεδομένων:

- “hasGroup”, “hasKeyword”, “hasLanguage”, “hasTitle” ορίζει ότι η κλάση εγγράφου έχει τύπο δεδομένων συμβολοσειράς.
- “hasName” ορίζει ότι η κλάση ατόμου έχει τύπο δεδομένων συμβολοσειράς.
- “hasPublicationYear” ορίζει ότι η κλάση εγγράφου έχει μη αρνητικό τύπο δεδομένων.
- “isRelevantForIFX” ορίζει ότι η κλάση εγγράφου έχει έναν τύπο δεδομένων τύπου boolean.



Εικόνα 3: Οντολογία δημοσιεύσεων IFX – Ιδιότητες Δεδομένων

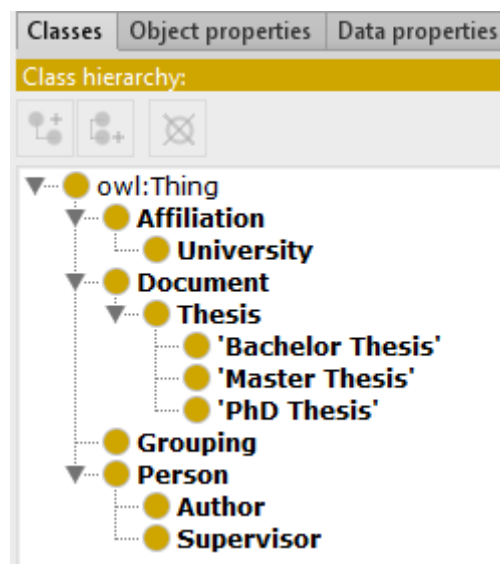
3.1.2 Οντολογία διατριβών IFX

Αυτή η οντολογία αναπτύχθηκε για να περιγράψει τα έργα εσωτερικής διατριβής που έγραψαν οι φοιτητές διδακτορικού, μεταπτυχιακού και προπτυχιακού τίτλου στο

τμήμα Μηχανικών Εφοδιαστικής Αλυσίδας της Infineon Technologies AG. Οι κλάσεις, οι ιδιότητες αντικειμένου και δεδομένων αυτής της οντολογίας παρουσιάζονται στις ακόλουθες εικόνες 4, 5 και 6 όπως φαίνονται στο Protégé.

Κλάσεις:

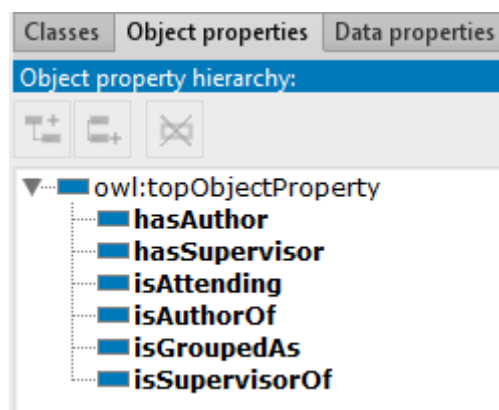
- “Affiliation”, η οποία περιέχει την κλάση “University” ως υπο-κλάση.
- “Document”, η οποία περιέχει την κλάση “Thesis” ως υπο-κλάση, και αυτή με τη σειρά της τις “PhD Thesis”, “Master Thesis” και “Bachelor Thesis” υπο-κλάσεις.
- “Grouping”.
- “Person”, η οποία περιέχει την κλάση “Author” και “Supervisor” ως υπο-κλάσεις.



Εικόνα 4: Οντολογία διατριβών IFX - Κλάσεις

Ιδιότητες Αντικειμένου:

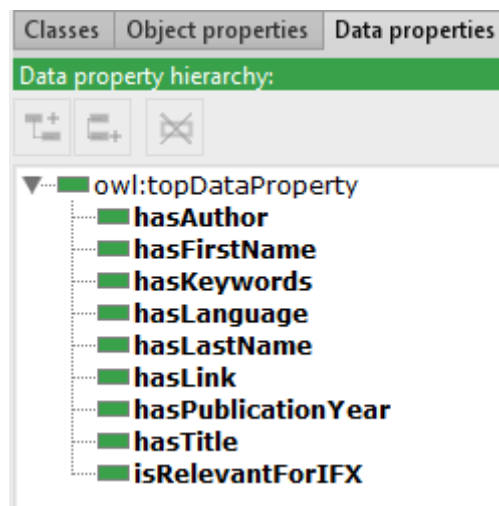
- Το "hasAuthor" σχετίζει το έγγραφο διατριβής με τον συν-συγγραφέα του. Η κλάση εγγράφου είναι το πεδίο ορισμού, η κλάση του συγγραφέα η εμβέλεια και είναι αντίστροφη της ιδιότητας αντικειμένου IsAuthorOf.
- "hasSupervisor" σχετίζει το έγγραφο διατριβής με τον επόπτη του. Η κλάση εγγράφου είναι το πεδίο ορισμού, η κλάση του επόπτη η εμβέλεια και είναι αντίστροφη της ιδιότητας αντικειμένου IsSupervisorOf.
- "isAttending" σχετίζει τον συγγραφέα της διατριβής με τον δεσμό του. Η κλάση του συγγραφέα είναι το πεδίο ορισμού και η κλάση του πανεπιστημίου η εμβέλεια του.
- "IsAuthorOf" σχετίζει τον συγγραφέα με το έγγραφο διατριβής του. Η κλάση του συγγραφέα είναι το πεδίο ορισμού, η κλάση του εγγράφου η εμβέλεια και είναι αντίστροφη της ιδιότητας αντικειμένου hasAuthor.
- "isGroupedAs" σχετίζει το έγγραφο διατριβής με την κατηγορία του. Η κλάση του εγγράφου είναι το πεδίο ορισμού και η κλάση κατηγορίας η εμβέλεια του.
- "IsSupervisorOf" σχετίζει τον επόπτη με το έγγραφο διατριβής. Η κλάση επόπτη είναι το πεδίο ορισμού, η κλάση εγγράφου η εμβέλεια και είναι αντίστροφη της ιδιότητας αντικειμένου hasSupervisor.



Εικόνα 5: Οντολογία διατριβών IFX - Ιδιότητες Αντικειμένων

Ιδιότητες Δεδομένων:

- “hasAuthor”, “hasKeywords”, “hasLanguage”, “hasLink”, “hasTitle” ορίζει ότι η κλάση εγγράφου (Document) έχει τύπο δεδομένων συμβολοσειράς.
- “hasFirstName” and “hasLastName” ορίζει ότι η κλάση ατόμου (Person) έχει τύπο δεδομένων συμβολοσειράς.
- “hasPublicationYear” ορίζει ότι η κλάση εγγράφου (Document) έχει μη αρνητικό τύπο δεδομένων.
- “isRelevantForIFX” ορίζει ότι η κλάση εγγράφου έχει έναν τύπο δεδομένων τύπου boolean.



Εικόνα 6: Οντολογία διατριβών IFX - Ιδιότητες Δεδομένων

Κεφάλαιο 4: Εφαρμογή και αποτελέσματα

4.1 προσέγγιση LogMap

Εργαλείο ανοιχτού κώδικα που αναπτύχθηκε στο Πανεπιστήμιο της Οξφόρδης. Δέχεται όλες τις μορφές τύπου OWL (Web Ontology Language), πραγματοποιεί αντιστοίχιση μεταξύ κλάσεων, ιδιοτήτων και αντικειμένων, και στη συνέχεια πραγματοποιεί συγχώνευση. Είναι γραμμένο σε Java, και μπορεί να χρησιμοποιηθεί από τη γραμμή εντολών (command-line) καθώς και από την διαδικτυακή του εφαρμογή (web interface application) (<http://krrwebtools.cs.ox.ac.uk/logmap/>). Είναι διαθέσιμο από την υπηρεσία "SourceForge", η οποία προσφέρει στους προγραμματιστές έργα δωρεάν λογισμικού και ανοιχτού κώδικα. Ο ερευνητής χρησιμοποίησε την διαδικτυακή εφαρμογή όπως φαίνεται στην Εικόνα 7 για να εκτελέσει τη δοκιμή του. Λεπτομέρειες σχετικά με τη διαδικασία υποβολής αιτήσεων εξηγούνται στο σημείο 4.1.1.

The screenshot shows the LogMap web interface. At the top, there is a header with the LogMap logo and project information. Below the header, there is a red notice about a fixed issue with email service. The main form is divided into two sections: 'Requester Information' and 'Input Ontologies'. The 'Requester Information' section contains fields for Name, Institution, and E-mail, all marked as mandatory. The 'Input Ontologies' section contains two identical blocks for 'Ontology 1' and 'Ontology 2'. Each block has a radio button for 'Web ontology URI' and a radio button for 'Local ontology file'. The 'Web ontology URI' option has a dropdown menu for 'URI Ontology 1' and 'URI Ontology 2'. The 'Local ontology file' option has a 'Choose File' button and an 'Upload Ontology' button. The 'Upload Ontology' button is disabled.

4.1.1 LogMap – Διαδικασία εφαρμογής

Η διαδικασία ήταν η εξής. Πρώτα, ο ερευνητής ανέβασε τις οντολογίες (Log Files) στο LogMap μέσω της διαδικτυακής εφαρμογής και έπειτα αφού έδωσε τα προσωπικά του στοιχεία, μετά από μερικά δευτερόλεπτα έλαβε ένα μήνυμα ηλεκτρονικού ταχυδρομείου που περιείχε έναν κατάλογο αναλυτικής διαδικασίας όπως φαίνεται στην Εικόνα 8 με όλα τα βήματα που έγιναν για τη συγχώνευση των δύο οντολογιών με το LogMap. Η οντολογία διατριβών IFX έχει 11 κλάσεις (classes), 15 ιδιότητες (properties) και 8 αντικείμενα (Individuals). Η οντολογία δημοσιεύσεων IFX έχει 9 κλάσεις (classes), 12 ιδιότητες (properties) και 10 αντικείμενα (Individuals). Περαιτέρω, πραγματοποιείται λεξικολογική ευρετηρίαση. Έπειτα, ακολουθήθηκε μια εξαγωγή και διάγνωση αξιόπιστων χαρτογραφημάτων, στη συνέχεια μια διάγνωση υποψηφίων χαρτογραφημάτων, μια εξαγωγή των αντιστοιχιών ιδιοτήτων και μια εξαγωγή των χαρτογραφικών στιγμιότυπων. Η εργασία αντιστοίχισης ολοκληρώθηκε με επιτυχία και τα αρχεία εξόδου αποθηκεύτηκαν σε 2 δευτερόλεπτα σύμφωνα με το σενάριο καταγραφής (Εικόνα 8). Τέλος, οι χαρτογραφήσεις εξόδου ενσωματώθηκαν πλήρως χρησιμοποιώντας επικαλύψεις / ενότητες και δημιουργήθηκε η νέα οντολογία η οποία είναι προσβάσιμη μέσω της τελευταίας γραμμής του σεναρίου καταγραφής.

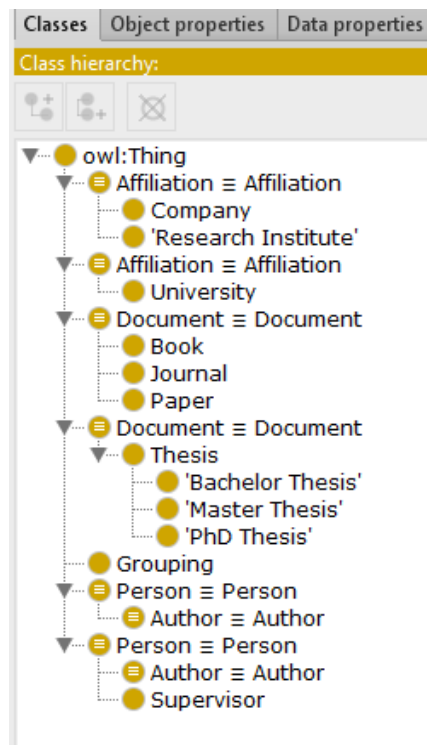
Progress of the matching task on 27/08/2018 by Christos Natsos using LogMap with reasoning: [Refresh](#)

- 09:05:57: LogMap is processing the request.
 - **Ontology 1:** [JFXThesis.owl](#)
 - **Ontology 2:** [PublicationOntology.owl](#)
- 09:05:58: Loading and overlapping...done
- 09:05:58: Overlapping ontology 1: 11 classes, 15 properties, 8 individuals.
- 09:05:58: Overlapping ontology 2: 9 classes, 12 properties, 10 individuals.
- 09:05:58: Lexical indexing...done
- 09:05:58: Extraction and diagnosis of reliable mappings...done
- 09:05:58: Diagnosis of candidate mappings...done
- 09:05:58: Extraction of property mappings...done
- 09:05:58: Extraction of instance mappings...done
- 09:05:58: Matching task finished. Storing output files...
- 09:05:58: Output mappings (class mappings = 4, property mappings = 2, instance mappings = 8): [\[OWL format\]](#), [\[TXT format\]](#), [\[OAEI Alignment format\]](#).
- 09:05:58: Full integrated ontology: [\[Original ontologies + mappings\]](#), (note that this ontology only imports the respective OWL files)
- 09:05:58: Overlapping ontology modules: [\[Module 1\]](#) [\[Module 2\]](#)
- 09:05:58: Integrated ontology using overlappings/modules: [\[Overlapping ontologies + mappings\]](#) (note that this ontology only imports the respective OWL files)

Εικόνα 8: LogMap – Βήματα συγχώνευσης

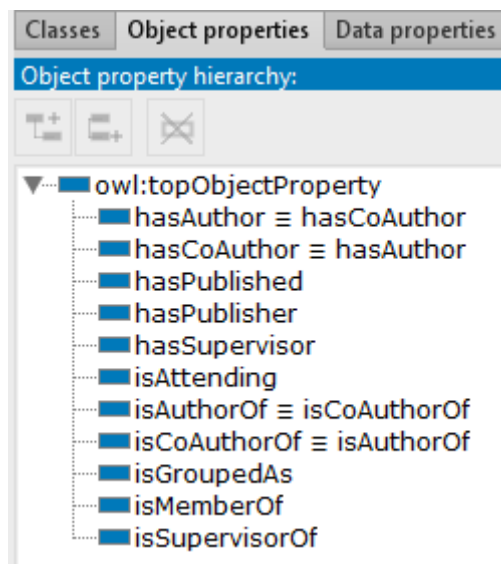
4.1.2 LogMap – Αποτελέσματα

Η όλη διαδικασία συγχώνευσης οντολογιών σύμφωνα με το σενάριο καταγραφής (Εικόνα 8) διήρκεσε 2 δευτερόλεπτα. Μετά την ολοκλήρωση της διαδικασίας, ο ερευνητής άνοιξε την ενσωματωμένη οντολογία μέσω του URI που παρείχε το LogMap (τελευταία γραμμή του σεναρίου καταγραφής). Όπως φαίνεται στις παρακάτω Εικόνες 9, 10 και 11, οι παρόμοιες κλάσεις και ιδιότητες από αμφότερες τις οντολογίες αντισταθμίστηκαν επιτυχώς και συγχωνεύθηκαν. Εδώ, πρέπει να πούμε ότι οι οντολογίες που χρησιμοποιήθηκαν ήταν μικρές και μη-περίπλοκες. Δεν γνωρίζουμε τι θα συμβεί όταν πρέπει να αντιμετωπίσουμε μεγάλες και σύνθετες οντολογίες. Η Maria Ganzha et al (2014) ανέφεραν ότι ο αλγόριθμος του LogMap βρίσκει ομοιότητες μεταξύ των εννοιών, χρησιμοποιώντας λεξιλόγια των οντολογιών εισόδου. Επομένως, το αποτέλεσμα μπορεί να μην είναι ικανοποιητικό αν οι οντολογίες είναι σοβαρά λεξικά διαφορετικές ή δεν παρέχουν αρκετές λεξικές πληροφορίες.



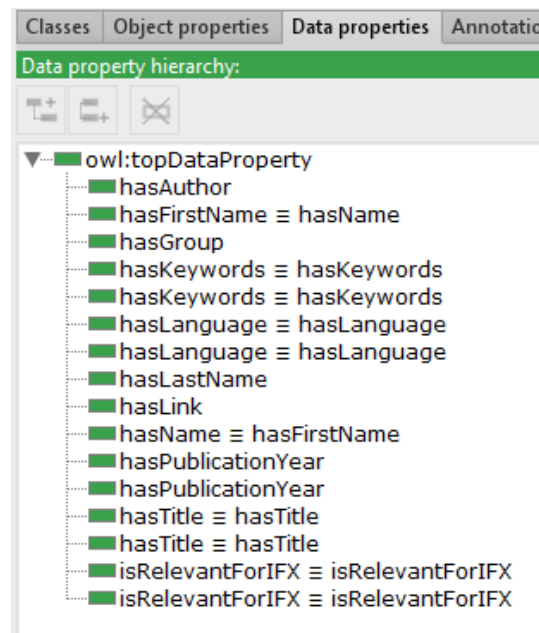
Εικόνα 9: LogMap – Ολοκληρωμένη οντολογία - Κλάσεις

LogMap αντιστόιχισε και έθεσε ισοδύναμες τις κλάσεις Δεσμού, Εγγράφου και Ατόμου μεταξύ των δύο οντολογιών.



Εικόνα 10: LogMap – Ολοκληρωμένη οντολογία – Ιδιότητες Αντικειμένων

Αντιστοίχισε και έθεσε ισοδύναμη την ιδιότητα αντικειμένου “hasAuthor” από την οντολογία διατριβών με την “hasCoAuthor” από την οντολογία δημοσιεύσεων, καθώς και την “isAuthorOf” από την οντολογία διατριβών με την “isCoAuthorOf” από την οντολογία δημοσιεύσεων.

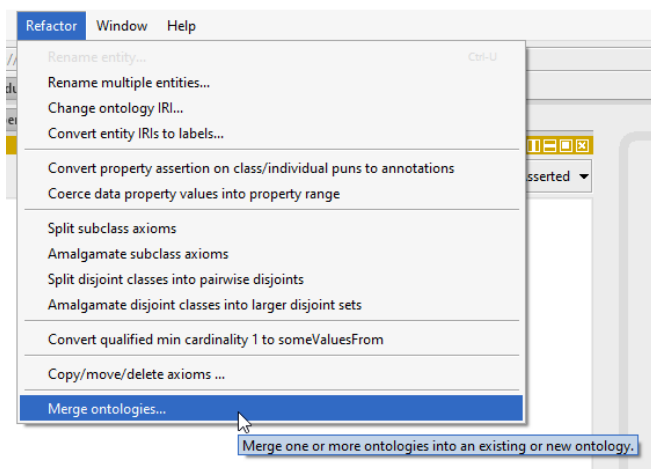


Εικόνα 11: LogMap – Ολοκληρωμένη οντολογία – Ιδιότητες Δεδομένων

Αντιστοίχισε και έθεσε ισοδύναμη τις ιδιότητες αντικειμένου “hasKeywords”, “hasLanguage”, “hasTitle” και “isRelevantForIFX” μεταξύ των δύο οντολογιών. Αντιστοίχισε επίσης μόνο την “hasFirstName” από την οντολογία διατριβών με την “hasName” της οντολογίας δημοσιεύσεων, παραλείποντας την “hasLastName”.

4.2 Προσέγγιση Protégé

Το Protégé είναι το πιο γνωστό περιβάλλον επεξεργασίας οντολογιών. Όπως φαίνεται στην Εικόνα 12, η τελευταία έκδοση του Protégé εισήγαγε τη λειτουργία συγχώνευσης στην γραμμή εργαλείων του. Λεπτομέρειες σχετικά με τη διαδικασία υποβολής αιτήσεων εξηγούνται στο σημείο 4.2.1.



Εικόνα 12: Protégé – Λειτουργία Συγχώνευσης

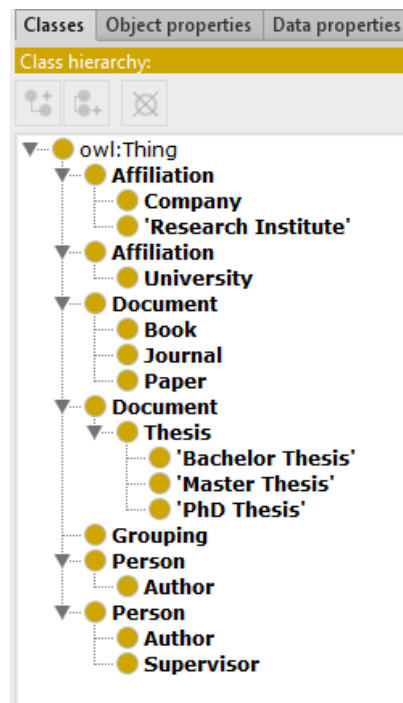
4.2.1 Protégé – Διαδικασία εφαρμογής

Η διαδικασία συγχώνευσης της Protégé είχε δύο φάσεις. Στη 1^η φάση ο ερευνητής άνοιξε αρχικά την οντολογία Διατριβών IFX και στη συνέχεια εισήγαγε την οντολογία Δημοσιεύσεων IFX από την καρτέλα "Εισαγωγές Οντολογίας" → "Άμεσες εισαγωγές" επιλέγοντας "Εισαγωγή οντολογίας σε συγκεκριμένο αρχείο". Ξεκίνησε τη διαδικασία συγχώνευσης επιλέγοντας "Συγχώνευση οντολογιών" από την καρτέλα "Refactor". Επίλεξε την οντολογία που ήθελε να ενοποιήσει στην άλλη οντολογία (την εισαγόμενη), μετά επέλεξε την επιλογή "Συγχώνευση σε υπάρχουσα οντολογία" και τελικά, επέλεξε την οντολογία την οποία θέλησε να ενοποιήσει (πρώτη φορτωμένη οντολογία). Στη φάση 2, ο ερευνητής έπρεπε να ορίσει

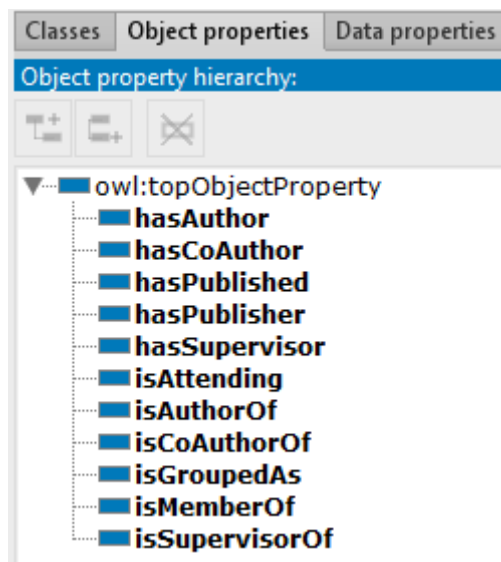
χειροκίνητα όλες τις κλάσεις και ιδιότητες (αντικειμένων και δεδομένων) που είναι παρόμοιες και στις δύο οντολογίες με την εντολή "Equivalent to".

4.2.2 Protégé – Αποτελέσματα

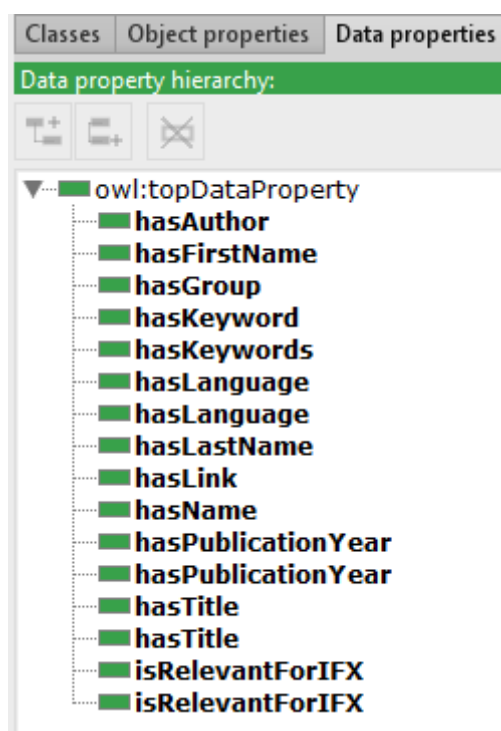
Όπως φαίνεται στις Εικόνες 13, 14 και 15, η συγχώνευση της Protégé από την 1^η φάση είναι μία απευθείας εισαγωγή της δεύτερης οντολογίας στην πρώτη, δημιουργώντας με αυτόν τον τρόπο μία ενιαία.



Εικόνα 13: Protégé's συγχωνευμένη οντολογία – 1^η φάση - Κλάσεις

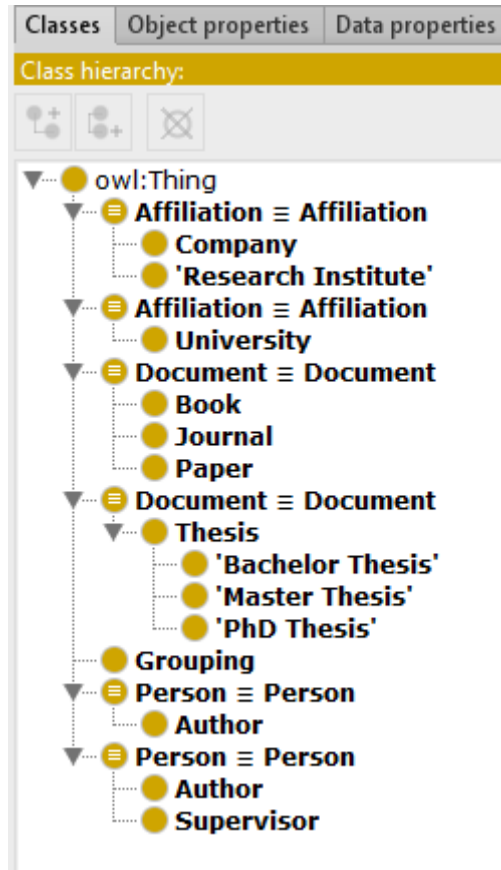


Εικόνα 14: Protégé's συγχωνευμένη οντολογία – 1^η φάση – Ιδιότητες Αντικειμένων

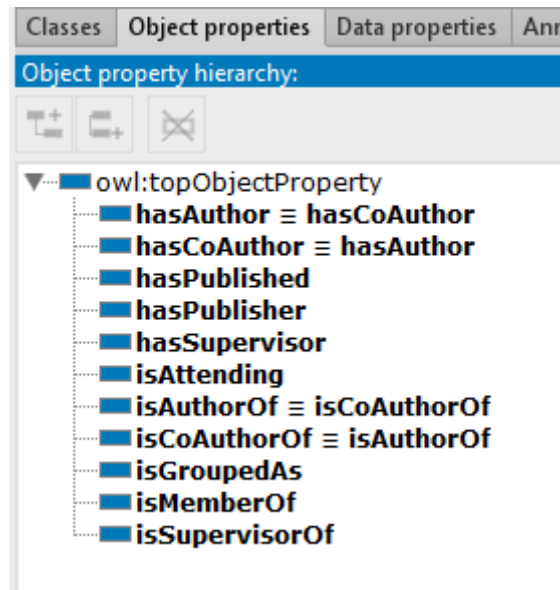


Εικόνα 15: Protégé's συγχωνευμένη οντολογία – 1^η φάση – Ιδιότητες Δεδομένων

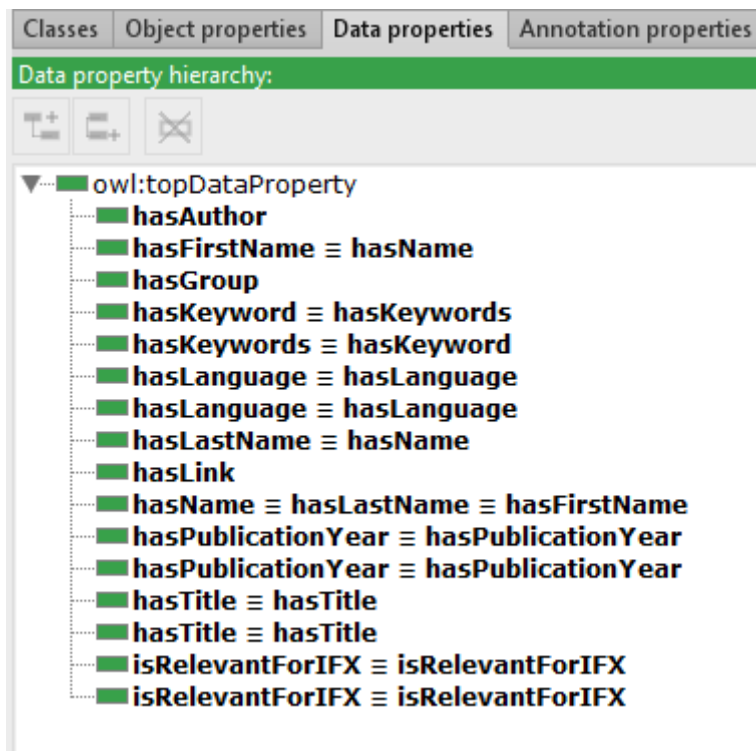
Το επόμενο βήμα του ερευνητή ήταν η 2^η φάση, όπου όρισε χειροκίνητα όλες τις όμοιες κλάσεις και ιδιότητες μεταξύ των δύο οντολογιών ισοδύναμες μεταξύ τους, η διαδικασία καταναλώθηκε περισσότερο χρόνο, αλλά το αποτέλεσμα όπως φαίνεται στις Εικόνες 16, 17 και 18 ήταν ακριβώς όπως το ήθελε ο ερευνητής.



Εικόνα 16: Protégé's συγχωνευμένη οντολογία – 2^η φάση - Κλάσεις



Εικόνα 17: Protégé's συγχωνευμένη οντολογία – 2^η φάση – Ιδιότητες Αντικειμένων



Εικόνα 18: Protégé's συγχωνευμένη οντολογία – 2^η φάση – Ιδιότητες Δεδομένων

Συμπέρασμα

Υπάρχει πολύ βιβλιογραφία για τη συγχώνευση των οντολογιών. Το ενδιαφέρον για αυτό το θέμα εξακολουθεί να αυξάνεται. Δυστυχώς, τα διαθέσιμα εργαλεία για το σκοπό αυτό είναι ελάχιστα. Τα περισσότερα από αυτά είτε δεν έχουν καλή τεκμηρίωση είτε έχουν σταματήσει να διατηρούνται και εξαφανίζονται πολύ γρήγορα μετά την εισαγωγή τους.

Ο ερευνητής χρησιμοποίησε μικρές και απλοποιημένες οντολογίες για να εκτελέσει τις δοκιμές του. Το αποτέλεσμα που έλαβε από το LogMap, το οποίο κατηγοριοποιείται ως αυτόματο εργαλείο συγχώνευσης οντολογιών, ήταν όπως περίμενε. Το LogMap αντιστοίχησε όλες τις παρόμοιες κλάσεις και ιδιότητες μεταξύ των παρεχόμενων οντολογιών και τις καθόρισε ισοδύναμες μεταξύ τους. Η διαδικασία αξιολογήθηκε γρήγορη και αποτελεσματική, αφήνοντας όμως στον ερευνητή μία αμφιβολία αφού δεν αντιστοίχισε την ιδιότητα `hasLastName` με την `hasName` όπως έκανε με την `hasFirstName`. Η προσέγγιση μέσω του Protégé κατανάλωσε περισσότερο χρόνο, αλλά τελικά ο ερευνητής έλαβε και πάλι το επιθυμητό αποτέλεσμα, αφού έπρεπε να ορίσει χειροκίνητα όλες τις ομοιότητες μεταξύ των δύο οντολογιών.

Ο ερευνητής θα συνεχίσει να αναζητά νέα εργαλεία και τεχνικές σχετικά με τη συγχώνευση οντολογιών. Οι επόμενες δοκιμές συγχώνευσης θα εκτελεστούν με μεγαλύτερες και πιο σύνθετες οντολογίες προκειμένου να λάβει αποτελέσματα που θα τον ικανοποιήσουν και θα τον πείσουν ότι οι υπολογιστές μπορούν να ταιριάξουν και να συγχωνεύσουν οντολογίες όπως ο άνθρωπος.

Βιβλιογραφία

- Allen Taylor (2015): Semantics for Dummies
- Aparna Saisree Thuluva, Darko Anicic, Sebastian Rudolph (2017): Semantic Web of Things for Industry 4.0
- BDI (2016): What is Industry 4.0?. Διαθέσιμο στο: <https://english.bdi.eu/article/news/what-is-industry-40/> [Πρόσβαση στις 06 Αυγούστου 2018]
- Bernard Marr (2016): What Everyone Must Know About Industry 4.0. Διαθέσιμο στο: <https://www.forbes.com/sites/bernardmarr/2016/06/20/what-everyone-must-know-about-industry-4-0/#27eff7c9795f> [Πρόσβαση στις 30 Ιουλίου 2018]
- Brian Matthews (2005): Semantic Web Technologies
- Cambridge Semantics: Getting Started with Semantic Technologies. Διαθέσιμο στο: <https://www.cambridgesemantics.com/blog/semantic-university/intro-semantic-web/> [Πρόσβαση στις 16 Ιουλίου 2018]
- D. L. McGuinness, R. Fikes, J. Rice, and S. Wilder (2000): An environment for merging and testing large ontologies
- E. Hovy (1998): Combining and standardizing large-scale, practical ontologies for machine translation and other uses
- Ernesto Jimenez-Ruiz and Bernardo Cuenca Grau (2011): LogMap: Logic-based and Scalable Ontology Matching
- European Parliament (2015): Industry 4.0 Digitalisation for productivity and growth
- Fausto Giunchiglia, Aliaksandr Autayeu and Juan Pane (2012): S-Match: an open source framework for matching lightweight ontologies

- Gartner (2015): What Is Industrie 4.0 and What Should CIOs Do About It?. Διαθέσιμο στο: <https://www.gartner.com/newsroom/id/3054921> [Πρόσβαση στις 07 Αυγούστου 2018]
- Gradiant (2016): Six technologies for Industry 4.0. Διαθέσιμο στο: <https://www.gradiant.org/noticia/10173/> [Πρόσβαση στις 09 Αυγούστου 2018]
- Grigoris Antoniou, Paul Groth, Frank van Harmelen and Rinke Hoekstra (2012): A Semantic Web Primer Third Edition
- H. Chalupsky: OntoMorph (2000): A translation system for symbolic knowledge
- Hayden Richards (2014): 5 Reasons Why Industry 4.0 is a game changer
- INCIBE (2015): Cybersecurity in Industry 4.0. Διαθέσιμο στο: <https://www.certs.es/en/blog/cybersecurity-in-industry-4-0> [Πρόσβαση στις 09 Αυγούστου 2018]
- i-Scoop: Industry 4.0: the fourth industrial revolution – guide to Industrie 4.0. Διαθέσιμο στο: <https://www.i-scoop.eu/industry-4-0/> [Πρόσβαση στις 08 Αυγούστου 2018]
- John Donovan (2013): The 4th Industrial Revolution is upon us. Διαθέσιμο στο: <https://www.ecnmag.com/article/2013/10/4th-industrial-revolution-upon-us> [Πρόσβαση στις 08 Αυγούστου 2018]
- Lorena Otero-Cerdeira, Francisco J. Rodríguez-Martínez, Alma Gómez-Rodríguez (2015): Ontology matching: A literature review
- Maria Ganzha, Marcin Paprzycki, Wieslaw Pawlowski, Pawel Szmaja, Katarzyna Wasielewska and Giancarlo Fortino (2014): Tools for ontology matching - practical considerations
- N. Fridman Noy, M. A. Musen (2000): PROMPT: algorithm and tool for automated ontology merging and alignment
- Nidhi Malik, Aditi Sharan, Deena Hijam (2015): Ontology Development for Agriculture Domain, International Conference on Computing for Sustainable Global Development

- Ontotext: What Is the Semantic Web?. Διαθέσιμο στο:
<https://ontotext.com/knowledgehub/fundamentals/what-is-the-semantic-web/> [Πρόσβαση στις 20 Ιουλίου 2018]
- Pascal Hitzler and Markus Krötzsch and Marc Ehrig and York Sure (2005):
What Is Ontology Merging?
- Rene Millman (2017): What is the semantic web?. Διαθέσιμο στο:
<http://www.itpro.co.uk/strategy/29606/what-is-the-semantic-web>
[Πρόσβαση στις 20 Ιουλίου 2018]
- SAP (2015): Industry 4.0 – fourth industrial revolution. Διαθέσιμο στο:
<https://blogs.sap.com/2015/06/30/industry-40-fourth-industrial-revolution/>
[Πρόσβαση στις 02 Αυγούστου 2018]
- Techopedia: Semantic Web. Διαθέσιμο στο:
<https://www.techopedia.com/definition/27961/semantic-web> [Πρόσβαση
στις 18 Ιουλίου 2018]
- Tim Berners-Lee, James Hendler and Ora Lassila (2001): The Semantic Web
- University of Leipzig: COMA 3.0. Διαθέσιμο στο: <https://dbs.uni-leipzig.de/Research/coma.html> [Πρόσβαση στις 15 Αυγούστου 2018]
- V. Richard Benjamins (2016): What Are Ontologies, and Why Do We Need Them?
- Vita Graudina, Janis Grundspenkis, Sigita Milasevica (2012): Ontology Merging in the Context of Concept Maps
- Zeeshan Ahmed and Detlef Gerhard (2010): Role of Ontology in Semantic Web Development