



Πανεπιστήμιο Πειραιώς – Τμήμα Πληροφορικής
Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακός Πολιτισμός, Έξυπνες Πόλεις, IoT και Προηγμένες Ψηφιακές
Τεχνολογίες»

Μεταπτυχιακή Διατριβή

Τίτλος Διατριβής	Η Καταγραφή αρχαίων πηγών με τη χρήση Οντολο- γιών: Η περίπτωση της Δυτικής Λοκρίδας Recording ancient sources using Ontologies: The case of Western Locris
Όνοματεπώνυμο Φοιτητή	Θεοφάνης Ξάνθης
Πατρώνυμο	Θεόδωρος
Αριθμός Μητρώου	ΨΠΟΛ 2230
Επιβλέπων	Γεράσιμος Ραζής, Διδάσκων

Ημερομηνία Παράδοσης **Μήνας Έτος**

Τριμελής Εξεταστική Επιτροπή

(υπογραφή)

(υπογραφή)

(υπογραφή)

Γεράσιμος Ραζής
Διδάσκων

Ιωάννης Αναγνωστόπουλος
Καθηγητής

Δημήτριος Βέργαδος
Καθηγητής

ΠΕΡΙΕΧΟΜΕΝΑ	
ΠΕΡΙΕΧΟΜΕΝΑ	2
ΠΑΡΑΡΤΗΜΑ ΕΙΚΟΝΩΝ	3
ΠΕΡΙΛΗΨΗ- ABSTRACT	5
ΕΙΣΑΓΩΓΗ	6
Ενότητα 1- ΘΕΩΡΗΤΙΚΟ ΠΛΑΙΣΙΟ	9
1.1. Σημασιολογικός Ιστός και Οντολογίες	10
α. Σημασιολογικός Ιστός	10
β. Οντολογίες	12
1.2. Η Καταγραφή των Αρχαίων Πηγών	15
Ενότητα 2- ΠΡΑΚΤΙΚΟ ΜΕΡΟΣ	20
2.1. Δημιουργία Οντολογίας	21
α. Η Θεωρητική Διαδικασία Δημιουργίας μιας οντολογίας	21
β. Τεκμηρίωση του προτύπου καταγραφής πηγών	27
2.2. Συγκεκριμένη μελέτη περίπτωσης	36
α. Επιλογή και εξήγηση του θέματος	36
β. Τεκμηρίωση της μελέτης περίπτωσης	38
γ. SPARQL ερωτήματα	44
ΕΠΙΛΟΓΟΣ	53
ΒΙΒΛΙΟΓΡΑΦΙΑ	55

ΠΑΡΑΡΤΗΜΑ ΕΙΚΟΝΩΝ	
Εικόνες 1 & 2: Σχεδιαγράμματα οπτικής αναπαράστασης της οντολογίας σε φυσική γλώσσα	28
Εικόνα 3: το μοναδικό χαρακτηριστικό (IRI) της Οντολογίας	28
Εικόνα 4: Ιεραρχία των κλάσεων- εικόνα στιγμιότυπου οθόνης από το Protégé	29
Εικόνα 5: Classes IRI- Labels- Οντολογίες ενσωμάτωσης των κλάσεων	30
Εικόνα 6: οι Object Properties της οντολογίας - εικόνα στιγμιότυπου οθόνης από το Protégé	31
Εικόνα 7: Object Properties- Labels- Οντολογία ενσωμάτωσης	32
Εικόνα 8: εύρος πεδίου τιμής των object properties- εικόνα στιγμιότυπου οθόνης από το Excel	33
Εικόνα 9 : οι Object Properties της οντολογίας - εικόνα στιγμιότυπου οθόνης από το Protégé	33
Εικόνα 10: Property IRI, Prefixes, Labels, εύρος πεδίου τιμής και τύπος των data properties- εικόνα στιγμιότυπου οθόνης από το Excel	34
Εικόνα 11: οπτική αναπαράσταση της οντολογίας- στιγμιότυπο οθόνης από το WebVOWL	37
Εικόνα 12: η κλάση "dbpedia:Author" και τα στιγμιότυπα της- εικόνα στιγμιότυπου οθόνης από το Protégé	37
Εικόνα 13: στιγμιότυπο οθόνης από το WebVOWL- εστίαση στην κλάση "dbpedia:Author"	38
Εικόνα 14: περιγραφή της οντότητας "sao:Aeschines", κλάσης "dbpedia:Author"- εικόνα στιγμιότυπου οθόνης από Protégé	38
Εικόνα 15: στιγμιότυπο οθόνης από το WebVOWL- εστίαση στην κλάση "bibo:Book"	39
Εικόνα 16: η κλάση "bibo:Book" και τα στιγμιότυπά της- εικόνα στιγμιότυπου οθόνης από το Protégé	39
Εικόνα 17: περιγραφή της οντότητας "sao:Ellinika", κλάσης "bibo:Book"- εικόνα στιγμιότυπου οθόνης στο Protégé	40
Εικόνα 18: στιγμιότυπο οθόνης από το WebVOWL- εστίαση στην κλάση "sao:Passage"	40
Εικόνα 19: περιγραφή οντότητας "3.100-102.5" κλάσης "sao:Passage"- εικόνα στιγμιότυπου οθόνης στο Protégé	41
Εικόνα 20: η καρτέλα "Usage" του στιγμιότυπου "sao:Athens"- εικόνα στιγμιότυπου οθόνης από το Protégé	42
Εικόνα 21: η καρτέλα "Usage" του στιγμιότυπου "Historical"- εικόνα στιγμιότυπου οθόνης από το Protégé	42
Εικόνα 22: Στιγμιότυπο οθόνης από την Query επέκταση του Protégé, SPARQL ερώτημα 1 "Ποια βιβλία έχει γράψει ο Αισχίνης"	45
Εικόνα 23: Στιγμιότυπο οθόνης από την Query επέκταση του Protégé, SPARQL ερώτημα 2 "Ποιοι συγγραφείς κατάγονται από την Αθήνα;"	46
Εικόνα 24: Στιγμιότυπο οθόνης από την Query επέκταση του Protégé, SPARQL ερώτημα 3 "Σε ποιο Είδος ανήκει το κάθε Βιβλίο;"	46
Εικόνα 25: Στιγμιότυπο οθόνης από την Query επέκταση του Protégé, SPARQL ερώτημα 4 "Πόσα είναι τα βιβλία από το κάθε είδος;"	47
Εικόνα 26: Στιγμιότυπο οθόνης από την Query επέκταση του Protégé, SPARQL ερώτημα 5 "Πόσα αποσπάσματα έχουμε από το κάθε βιβλίο;"	48
Εικόνα 27: Στιγμιότυπο οθόνης από την Query επέκταση του Protégé, SPARQL ερώτημα 6 "Ποια είναι τα βιβλία που ανήκουν στο Είδος Ιστορικά;"	49
Εικόνα 28: Στιγμιότυπο οθόνης από την Query επέκταση του Protégé, SPARQL ερώτημα 7 "ποια χωρία έχουν ως θέμα τους τον Πελοποννησιακό	49

Εικόνα 29: Στιγμιότυπο οθόνης από την Query επέκταση του Protégé, SPARQL ερώτημα 8 “Σε ποιο βιβλίο και σε ποιόν συγγραφέα ανήκει το κάθε	50
Εικόνα 30: Στιγμιότυπο οθόνης από την Query επέκταση του Protégé, SPARQL ερώτημα 9 “Ποια είναι τα π.Χ. αποσπάσματα που αφορούν τον Πε-	51

ΠΕΡΙΛΗΨΗ

Η παρούσα Μεταπτυχιακή Διατριβή διερευνά τη χρήση των Οντολογιών ως εργαλείων υποστήριξης και ανάπτυξης της ιστορικής έρευνας, εστιάζοντας στην οργάνωση και τη διαχείριση των απαραίτητων μεταδεδομένων. Για το σκοπό αυτό, αναπτύχθηκε μια οντολογία, μέσω του λογισμικού ανοιχτού κώδικα Protégé, η οποία λειτουργεί ως οντολογικό πρότυπο για την καταγραφή των αρχαίων πηγών, προτείνοντας ένα ευέλικτο και αποδοτικό σύστημα οργάνωσης. Απευθύνεται στους ερευνητές και επιτρέπει αποτελεσματική οργάνωση, γρήγορη πρόσβαση και στοχευμένες αναζητήσεις. Σε αυτή, βασίστηκε η οντολογία που εφαρμόστηκε για τη μελέτη περίπτωσης της παρουσίασης των πηγών της δυτικής Λοκρίδας. Η οντολογική εφαρμογή αποδεικνύει τη λειτουργικότητα του προτύπου, καλύπτοντας παράλληλα και μερικώς ένα σημαντικό ερευνητικό κενό. Το πρότυπο και η μελέτη περίπτωσης συνοδεύονται από το θεωρητικό τους πλαίσιο, την ανάλυση της μεθοδολογίας και της δομής τους και την τεκμηρίωση των επιλογών των εννοιών. Τέλος, ενδεικτικά SPARQL ερωτήματα χρησιμοποιούνται για να υπογραμμίσουν τις πρακτικές δυνατότητες της εφαρμογής και να αναδείξουν τα όρια της μελλοντικής της εξέλιξης. Η παρούσα εργασία προσφέρει μια βάση για περαιτέρω μελλοντική ανάπτυξη τέτοιων εργαλείων που θα εκμεταλλεύονται τις δυνατότητες διασύνδεσης, οργάνωσης και ανάλυσης δεδομένων για την προώθηση της ιστορικής έρευνας.

ABSTRACT

This Master Thesis explores the use of Ontologies as tools to support and develop historical research, focusing on the organization and management of the necessary metadata. For this purpose, an ontology was developed, through the open-source software Protégé, which acts as an ontological standard for the recording of ancient sources, proposing a flexible and efficient system of organization. It is aimed at researchers and allows efficient organization, quick access and targeted searches. On this, the ontology applied for the case study of the presentation of the sources of western Locris was based. The ontological implementation proves the functionality of the standard, while also partially filling an important research gap. The model and the case study are accompanied by their theoretical framework, the analysis of their methodology and structure, and the documentation of the choices of concepts. Finally, indicative SPARQL queries are used to highlight the practical capabilities of the application and highlight the limits of its future development. This paper offers a basis for further future development of such tools that will take advantage of the possibilities of data interconnection, organization and analysis to advance historical research.

ΕΙΣΑΓΩΓΗ

Η χρήση του διαδικτύου έχει μεγιστοποιήσει την ταχύτητα και τον όγκο των πληροφοριών που δημιουργούνται, διαδίδονται, ανακτώνται και επεξεργάζονται, επηρεάζοντας όλες τις πτυχές της ανθρώπινης δραστηριότητας. Δημιουργεί συνεχώς νέες ευκαιρίες επαφής, επεξεργασίας και αξιοποίησης του υλικού, αλλά αυτή η πρόοδος επιφέρει και προκλήσεις. Ο Παγκόσμιος Ιστός λειτουργεί ως μια τεράστια βάση δεδομένων στην οποία ο χρήστης αναζητά οτιδήποτε με τη βοήθεια μηχανών αναζήτησης και «λέξεων-κλειδίων». Η διαρκής ανάπτυξη του Παγκόσμιου Ιστού έχει αναδείξει τα δομικά του προβλήματα αναφορικά με τη διαχείριση του όγκου των πληροφοριών και την αξιολόγησή τους, την ακρίβεια των αποτελεσμάτων και τελικά την αξιοποίηση του υλικού. Ο Παγκόσμιος Ιστός είναι δομημένος με βάση τις φυσικές γλώσσες των χρηστών του, προς διευκόλυνσή τους, αλλά αυτό δημιουργεί πρόβλημα στην «επικοινωνία» με τους υπολογιστές. Αυτοί, δεν έχουν τη δυνατότητα να ερμηνεύουν και να αξιολογούν το περιεχόμενο της πληροφορίας, αλλά μόνο να «διαβάζουν» τη συντακτική περιγραφή της. Το γεγονός έχει επιπτώσεις στην ακρίβεια, την ποιότητα και την αξιοπιστία των ανακλήσεων και απαιτείται επιπλέον εργασία από το χρήστη, ο οποίος «επιβαρύνεται» σε χρόνο και ενέργεια προκειμένου να βρει, να οργανώσει και να αξιολογήσει μόνος του τα αποτελέσματα που επιθυμεί.

Καθώς αυξάνεται η ταχύτητα διάδοσης, ο όγκος των πληροφοριών και οι απαιτήσεις των χρηστών, μεγεθύνονται επίσης και τα προβλήματα του Παγκόσμιου Ιστού οπότε παρουσιάζεται όλο και εντονότερη η ανάγκη βελτιστοποίησης των «υπηρεσιών» του. Αυτό θα επιτευχθεί μέσω μιας εξελικτικής διαδικασίας δομικού μετασχηματισμού του, κατά την οποία ο «Παγκόσμιος» θα μετατραπεί στον «Σημασιολογικό» Ιστό. Σύμφωνα με τον οραματιστή του, Tim Berners-Lee, ο Σημασιολογικός Ιστός αποτελεί προέκταση του Παγκόσμιου και πρόκειται για ένα σύνολο μεθόδων καταγραφής και διασύνδεσης πληροφοριών το οποίο θα επιτρέπει στους υπολογιστές να «κατανοούν» τη σημασία των πληροφοριών τις οποίες αναζητούν οι χρήστες. Επομένως θα έχουν πλέον τη δυνατότητα και να την «αξιολογήσουν», «αφαιρώντας» τη δουλειά από το χρήστη. Το αποτέλεσμα αυτής της μετατροπής θα είναι η αυτόματη αξιολόγηση και διαχείριση του τεράστιου και ποικίλου όγκου πληροφοριών και οι αναζητήσεις θα οδηγούν σε ακριβέστερα και πιο αξιόπιστα αποτελέσματα.

Το σημαντικότερο ρόλο σ' αυτή την εξελικτική διαδικασία του Ιστού έχει ο τρόπος αναπαράστασης της πληροφορίας. Ανεξάρτητα με την παρουσίασή τους σε φυσική γλώσσα για την επικοινωνία των χρηστών, το νόημα των εννοιών και η δομή του πλαισίου τους θα πρέπει να περιγράφεται σε μορφή αναγνώσιμη από τους υπολογιστές. Γι' αυτή τη «σημασιολογική» περιγραφή απαιτείται η δημιουργία/ χρήση μιας κοινής «γλώσσας» κατανοητής από χρήστες και υπολογιστές και η υιοθέτηση ενός πλαισίου το οποίο θα «επιτρέπει» την αποτελεσματικότερη συνεργασία τους. Για το σκοπό αυτό δημιουργούνται και αναπτύσσονται οι 'Οντολογίες', σημασιολογικές δομές που βασίζονται σε τυποποιημένα και συστηματοποιημένα γλωσσικά μοντέλα αναπαράστασης πληροφορίας. Αποτελούν τη «λύση» βασικών προβλημάτων του Παγκόσμιου Ιστού: επιτρέπουν την αναπαράσταση της γνώσης σε γλώσσα κατανοητή από χρήστες και υπολογιστές (knowledge representation) και παρέχουν τη συλλογιστική μορφή που επιτρέπει την καλύτερη οργάνωση που οδηγεί στην εξαγωγή γρήγορων και ακριβών συμπερασμάτων (reasoning).

Η πρόοδος της ερευνητικής διαδικασίας και η εξαγωγή επιστημονικών συμπερασμάτων βασίζεται στις εξελίξεις της τεχνολογίας και την «εκμετάλλευση» των νέων εργαλείων. Οι οντολογίες βοηθούν τις επιστήμες στην οργάνωση του μεγάλου και αυξανόμενου όγκου των πληροφοριών της και στη διασύνδεση δεδομένων διαφορετικού τύπου και μορφής. Με τη συνεργασία διαφορετικών επιστημονικών κλάδων αυξάνεται επίσης η ακρίβεια και η εγκυρότητα των αποτελεσμάτων.

Ως αρχαίες -φιλολογικές- πηγές εννοούμε τα γραπτά τεκμήρια που προέρχονται από την αρχαιότητα και μας παρέχουν πληροφορίες αναφορικά με οποιαδήποτε ανθρώπινη δραστηριότητα. Η ύπαρξη αυτών των πηγών και η γνώση την οποία προσφέρουν είναι καθοριστικές για την έρευνα και τη μελέτη της ιστορίας. Η καταγραφή περιλαμβάνει τη συλλογή, καταλογογράφηση, αρχειοθέτηση, ανάλυση και ερμηνεία των αρχαίων κειμένων, με σκοπό την εξαγωγή καθοριστικών συμπερασμάτων για την ανασύσταση της ιστορικής πραγματικότητας. Η ορθή, συστηματική και αποτελεσματική καταγραφή διασφαλίζει τη διατήρησή τους και προωθεί την πρόσβαση σε αυτές από το ευρύ κοινό (συμβάλλοντας στη συλλογική ιστορική γνώση). Επιπλέον, κρίνεται επιβεβλημένη για την πληρότητα της ιστορικής γνώσης και την εξαγωγή τεκμηριωμένων συμπερασμάτων.

Η ανάγκη για τη συλλογή τεκμηρίων με σκοπό την κατανόηση του παρελθόντος είναι

διαχρονική. Υπήρχε από την αρχαιότητα και συνεχίζεται ως τις μέρες μας και αυτό που έχει αλλάξει ριζικά είναι ο τρόπος που συλλέγονται, οργανώνονται και επεξεργάζονται. Από την προφορική παράδοση και ατεκμηρίωτες ιστορίες και σημειώσεις, περάσαμε στην συγκέντρωσή της γνώσης σε βιβλιοθήκες και στη διάσωσή της μέσω της καταλογογράφησης της αντιγραφής και του σχολιασμού. Σήμερα η διαδικασία αποθήκευσης και καταγραφής έχει εξελιχθεί στο ψηφιακό πλαίσιο και πραγματοποιείται αρχικά μέσω της ψηφιοποίησης του αρχαιικού υλικού. Τα ψηφιοποιημένα αρχεία έπειτα συλλέγονται και καταχωρούνται σε βάσεις δεδομένων, επικουρούμενα από την καταγραφή των βασικών πληροφοριών τους, σύμφωνα με τα πρότυπα αρχειοθέτησης. Εξακολουθούν όμως να υφίστανται τα προβλήματα, ενώ πρέπει να αντιμετωπιστούν οι προκλήσεις της σύγχρονης εποχής. Τα κυριότερα προβλήματα είναι η δυσκολία διαχείρισης του μεγάλου και ποικίλου όγκου πληροφοριών και η έλλειψη διασύνδεσης. Στην αντιμετώπισή τους προσβλέπει ο Σημασιολογικός Ιστός με την ανάπτυξη των οντολογιών, προάγοντας τη διασύνδεση και των συσχετισμό των πληροφοριών, την αποδοτικότερη οργάνωσή, ενώ επιτρέπει και την ταχύτερη και ακριβέστερη αναζήτηση και ανάκτησή, βελτιώνοντας τις συνθήκες διεξαγωγής της έρευνας.

Η διπλωματική εργασία συνεισφέρει στην ανάγκη ευέλικτης και αποδοτικής αρχειοθέτησης των γραμματειακών πηγών. Διαθέτει αρχικά ένα πρακτικό κομμάτι το οποίο περιλαμβάνει την κατασκευή μιας οντολογίας η οποία λειτουργεί ως πρότυπο καταγραφής των αρχαίων ελληνικών πηγών με τη βοήθεια του εργαλείου ανοιχτού κώδικα Protégé (<https://Protégé.stanford.edu>). Στόχος της κατασκευής ήταν ο απλός, ευέλικτος και αποδοτικός τρόπος αποτύπωσης των βασικών πληροφοριών, με στόχο την ορθότερη και ταχύτερη αναζήτησή τους για την διευκόλυνση της ερευνητικής διαδικασίας. Η παρούσα εργασία λειτουργεί υποστηρικτικά ως προς την εφαρμογή, επεξηγώντας το πλαίσιο και τους κύριους όρους, περιγράφοντας τη διαδικασία κατασκευής και τη μεθοδολογία και παρουσιάζοντας αντιπροσωπευτικά για την κατανόηση των αναγνωστών παραδείγματα.

Στο γραπτό της κομμάτι η δομή της εργασίας ακολουθεί τους στόχους και τον σκοπό της. Η 'Εισαγωγή' (σ.σ. το παρόν) περιλαμβάνει αρχικά μια συνοπτική ανασκόπηση της σημερινής κατάστασης στο πεδίο της καταγραφής των πηγών και των προβλημάτων που παρουσιάζει ο Παγκόσμιος Ιστός. Στη συνέχεια εξηγούνται συνοπτικά οι βασικοί θεωρητικοί όροι που αποτέλεσαν το υπόβαθρο τόσο για τη δημιουργία της οντολογίας, όσο και για το θεωρητικό κομμάτι της εργασίας και θα σχολιαστούν στα κεφάλαιά της (Σημασιολογικός Ιστός, οντολογίες, κλάσεις, καταγραφή, πηγές κ.α.). Ολοκληρώνεται με τη διάρθρωση της εργασίας και των περιεχομένου της.

Το κύριο μέρος της Διπλωματικής χωρίζεται σε δύο ενότητες, με την πρώτη να αποτελεί το θεωρητικό υπόβαθρο και τη δεύτερη την πρακτική τεκμηρίωση της οντολογίας που κατασκευάστηκε. Η πρώτη ενότητα, χωρίζεται σε δύο υποενότητες και αφορά το πλαίσιο στο οποίο εντάσσεται η πρακτική κατασκευή. Η πρώτη υποενότητα με τη σειρά της χωρίζεται σε δύο μικρότερα μέρη. Στο πρώτο εξετάζεται ο 'Σημασιολογικός Ιστός': δίδεται ο ορισμός, σχολιάζεται η αναγκαιότητα επέκτασης του Ιστού, η επίδρασή του στη διάδοση της γνώσης και εξηγείται ο τρόπος λειτουργίας του. Στο δεύτερο μέρος εμβαθύνουμε σε ένα κύριο άξονα του 'Σημασιολογικού Ιστού', τις 'Οντολογίες': γίνεται προσπάθεια ορισμού τους, σχολιάζονται η αναγκαιότητα χρήσης τους, οι δυνατότητες αξιοποίησης και η μελλοντική προσφορά από την ανάπτυξή τους. Περιγράφεται ο τρόπος λειτουργίας και αναφέρονται τα βασικά μέρη που τις αποτελούν.

Η δεύτερη υποενότητα του πρώτου μέρους επικεντρώνεται στο ιδιαίτερο θέμα της 'Καταγραφής των αρχαίων πηγών': πρώτα δίνεται ο ορισμός της και πραγματοποιείται μια συνοπτική ανασκόπηση της παραδοσιακής διαδικασίας. Στη συνέχεια αναλύονται οι προκλήσεις που αφενός αντιμετωπίζει η διαδικασία και καλούνται να «λύσουν» οι οντολογίες και ο Σημασιολογικός Ιστός, αφετέρου αυτές που (θα) προκύψουν από την ανάπτυξή τους. Σχολιάζεται η αναγκαιότητα βελτιστοποίησής της και η επίδραση (θετική ή αρνητική) της χρήσης των οντολογιών.

Η δεύτερη ενότητα της εργασίας περιλαμβάνει την «τεκμηρίωση» την ανάλυση των οντολογιών που κατασκευάστηκαν. Χωρίζεται επίσης σε δύο υποενότητες. Η πρώτη υποενότητα πραγματεύεται τη διαδικασία δημιουργίας μιας οντολογίας. Στο πρώτο κομμάτι της σχολιάζονται οι κύριοι όροι που αφορούν μια οντολογία και αναλύονται τα βήματα που ακολουθούνται για την ανάπτυξή της. Η ανάλυση των βημάτων περιλαμβάνει την περιγραφή του, κρίνεται η αναγκαιότητά του και σχολιάζεται η λειτουργία και η σημασία του. Το δεύτερο κομμάτι της πραγματεύεται την κατασκευή ενός οντολογικού προτύπου για την καταγραφή των αρχαίων πηγών με σκοπό τη χρήση του ως γρήγορο και βασικό εργαλείο των ερευνητών. Χρησιμοποιώντας κατάλληλα screenshots από το Protégé (το πρόγραμμα ανοιχτού κώδικα το οποίο χρησιμοποιήθηκε για την κατασκευή της οντολογίας) θα παρουσιαστούν ενδεικτικά παραδείγματα από τις κλάσεις και τις

ιδιότητες με τις οποίες σχετίζονται και επιλέχθηκαν ως κατάλληλες για την εξυπηρέτηση του σκοπού της οντολογίας. Οι επιλογές αυτές αιτιολογούνται, αξιολογούνται και κρίνεται η σημασία τους.

Το πρότυπο που αναπτύχθηκε πρέπει να μπορεί να εφαρμοστεί σε μια υπαρκτή υπόθεση εργασίας προκειμένου να αποδεικνύεται η λειτουργικότητα και η χρησιμότητα του μοντέλου. Μια τέτοια μελέτη περίπτωσης παρουσιάζεται στη δεύτερη υποενότητα, η οποία επίσης χωρίζεται σε τρία μέρη: στο πρώτο, συνοπτικό, κομμάτι σχολιάζεται η διαδικασία επιλογή του θέματος, πως ξεκίνησε και διαμορφώθηκε, αποτιμάται η ιδέα και γίνεται αναφορά στην ως τώρα κατάσταση. Στο δεύτερο, σχολιάζεται η εφαρμογή του προτύπου που αναλύθηκε στην προηγούμενη υποενότητα. Με τη βοήθεια και πάλι των *printscreens* από το Protégé, παρουσιάζονται παραδείγματα 'στιγμιότυπων' (*instances*) από κλάσεις, ιδιότητες ή τις σχέσεις τους, προκειμένου να αναδειχθεί ο τρόπος κατασκευής αλλά και πρακτικής εφαρμογής της λειτουργίας του προτύπου. Στο τρίτο και τελευταίο κομμάτι παρουσιάζονται ορισμένα ενδεικτικά SPARQL ερωτήματα. Επιλέχθηκαν για να αναδειχθεί η δυνατότητα επεξεργασίας του υλικού, το εύρος των απαντήσεων που μπορούν να ανακτηθούν, οι περιορισμοί του παραδείγματος αλλά και η συντακτική αρτιότητα της οντολογίας.

Το τελευταίο κεφάλαιο της εργασίας είναι ο 'Επίλογος'. Περιλαμβάνει μια συνοπτική ανακεφαλαίωση της κατάστασης και των κύριων όρων που χρησιμοποιήθηκαν για το θεωρητικό μέρος της εργασίας. Ακολουθεί η αποτίμηση της οντολογίας που κατασκευάστηκε (προτύπου και ιδιαίτερης μελέτης περίπτωσης), αναφορικά με τη διαδικασία δημιουργίας, αλλά και τη χρήση τους. Τέλος, προτείνονται τρόποι μελλοντικής εξέλιξης των οντολογιών που παρουσιάστηκαν.

Στο σημείο αυτό και πριν την παρουσίαση του κύριου μέρους της εργασίας αισθάνομαι την ανάγκη να εκφράσω τις ευχαριστίες μου σε όσους συνέβαλαν στην ολοκλήρωσή της. Πρώτα από όλους θα ήθελα να ευχαριστήσω τους γονείς μου για την υποστήριξη, ενθάρρυνση και υπομονή τους όλα τα χρόνια των σπουδών μου. Επιπλέον εκφράζω τις ευχαριστίες μου προς τους καθηγητές μου στο Τμήμα Ιστορίας και Αρχαιολογίας του Πανεπιστημίου Ιωαννίνων, οι οποίοι τόσο στο προπτυχιακό όσο και στο μεταπτυχιακό επίπεδο, μου μετέδωσαν τη σωστή μεθοδολογία της έρευνας και φρόντισαν να καλλιεργήσουν τις ερευνητικές μου ανησυχίες. Επίσης το Π.Μ.Σ. «Ψηφιακός Πολιτισμός, Έξυπνες Πόλεις, IoT Και Προηγμένες Ψηφιακές Τεχνολογίες» για την ευκαιρία που μου προσέφερε και τους καθηγητές του που μου έδειξαν τον τεράστιο και άγνωστο κόσμο της τεχνολογίας και τις δυνατότητές του. Ιδιαίτερη μνεία αξίζει στον επόπτη της διπλωματικής μου εργασίας, Δρ. Ραζή, το μάθημα του οποίου πυροδότησε την ιδέα της παρούσας εργασίας. Από την αρχή έδειξε ενδιαφέρον για την ιδέα μου και συμφώνησε να υποστηρίξει την προσπάθειά μου. Καθ' όλη τη διάρκεια της συνεργασίας μας, η καθοδήγησή του υπήρξε καθοριστική. Με την υπομονή του, τις χρήσιμες συμβουλές του και την εμπιστοσύνη που μου έδειξε, μου παρείχε τη στήριξη που χρειαζόμουν για να ολοκληρώσω αυτή την εργασία.

Ενότητα 1- ΘΕΩΡΗΤΙΚΟ ΠΛΑΙΣΙΟ

1.1. Σημαιολογικός Ιστός και Οντολογίες

α. Σημαιολογικός Ιστός

Ο Σημαιολογικός Ιστός αποτέλεσε το όραμα του Tim Berners-Lee, ιδρυτή του Παγκόσμιου Ιστού ήδη από τη δημιουργία του ως το επόμενο βήμα στην εξέλιξή του. Είναι μια βελτιωμένη έκδοση του Παγκόσμιου Ιστού, ο σκοπός δεν είναι η αντικατάσταση αλλά η ενίσχυση της πληροφορίας που έχει. Για να το επιτύχει αυτό καλείται να αναπτύξει ένα πλαίσιο ανάγνωσης, επεξεργασίας, διαχείρισης και ανάλυσης από τους υπολογιστές των διαθέσιμων στο διαδίκτυο εγγράφων των χρηστών με τρόπο που θα επιτρέπει την άμεση και ακριβή αναζήτηση (Iακωβάκης 2019, σ. 7; Mehra & Kumar 2011, σ.σ. 217, 222; Σαμοθρακίτου 2015, σ. 19; Τσαρτσάλη 2015, σ.σ. 12-3; Ψυχογιός 2012, σ.σ. 46-7). Πρόκειται για ένα Ιστό πληροφοριών και δεδομένων με αυστηρά καθορισμένα νοήματα, τα οποία θα διέπονται από κανόνες και η σημαιολογική ανάγνωση και ερμηνεία τους θα προκύπτει από την ανάπτυξη των οντολογιών και τη δυνατότητα των υπολογιστών να κατανοούν, ερμηνεύουν, διαχειρίζονται και επεξεργάζονται τα μεταδεδομένα τους, με τον τρόπο που το κάνουν οι χρήστες.

Η ανάγκη επέκτασης του Ιστού προέκυψε επιτακτική λόγω της αδυναμίας του Παγκόσμιου Ιστού να καλύψει την ανάγκη των χρηστών του να διαχειριστούν αποτελεσματικότερα τον τεράστιο όγκο πληροφοριών που συσσωρεύει. Κατά κύριο λόγο το πρόβλημα προκύπτει καθώς ο Παγκόσμιος Ιστός δε διαθέτει λογική στη δομή του. Οι πληροφορίες που δημοσιεύονται σε αυτόν βασίζονται στην συντακτική περιγραφή του περιεχομένου τους, τα οποία αποδίδονται σε κάποια από τις φυσικές γλώσσες που γνωρίζουν οι χρήστες αλλά δεν είναι κατανοητές από τους υπολογιστές, επομένως δεν μπορούν να τις διαβάσουν για να τις επεξεργαστούν. Το αποτέλεσμα είναι πως οι χρήστες αναγκάζονται να επιτελούν επιπλέον εργασίες για την αξιολόγηση, ανάλυση, ερμηνεία και επιλογή των δεδομένων που ανακτούν από τον Ιστό. Ο Σημαιολογικός Ιστός χρησιμοποιεί γλωσσικά και μορφολογικά πρότυπα με τα οποία διαχειρίζεται την έκφραση των (μετα)δεδομένων και στοιχείων (όπως για παράδειγμα ορισμό όρων, σχέσεων, περιορισμούς και αξιώματα) τα οποία η HTML κ η XML δε μπορούν να περιγράψουν¹. Η OWL (Web Ontology Language) είναι μια γλώσσα προγραμματισμού ειδικά σχεδιασμένη για τον Σημαιολογικό Ιστό, την ανάπτυξη των Οντολογιών και την αναπαράσταση των πολύπλοκων σχέσεων των δεδομένων (Μαυρογιάννη 2009, σ. 103; Τριανταφύλλου 2022, σ. 24; Ψυχογιός 2012, σ.σ. 44-6). Εκφράζονται επίσης οι λογικοί κανόνες που δίνουν τη δυνατότητα στους υπολογιστές να εξάγουν μόνοι τους συμπεράσματα από τα δεδομένα που περιγράφονται. Το βασικό μοντέλο αναπαράστασης των δεδομένων (πόρων/resources) και της σημαιολογίας τους είναι η RDF (Resource Description Framework), μια απλή γλώσσα προγραμματισμού η οποία βασίζεται στη δημιουργία τριπλετών. Κάθε τριπλέτα εκφράζει μια δήλωση και αποτελείται από τρία μέρη, σε ένα σχήμα που θα μπορούσε να παραλληλιστεί με μια πολύ απλή πρόταση σε φυσική γλώσσα: το υποκείμενο (subject- το δεδομένο που περιγράφεται), το κατηγορημα (predicate- μια ιδιότητα του πόρου που περιγράφεται ή μια σχέση που έχει αυτός ο πόρος με έναν άλλο) και το αντικείμενο (object- ένας πόρος του μέσου του οποίο ενώνεται με έναν άλλο ή μια μοναδική τιμή που χαρακτηρίζει μια ιδιότητα ή σχέση)². Η RDFS είναι η σημαιολογική επέκτασή της και ουσιαστικά παρέχει ένα λεξικό μοντελοποίησης των δεδομένων, επιτρέποντας τον ορισμό των κλάσεων και των 'ιδιοτήτων' (Bikakis 2013, σ.σ. 7-10; Τσαρτσάλη 2015, σ.σ. 15-9; Χατζησταμούλου 2020, σ.σ. 31-3; Χρυσούλας 2009, σ.σ. 63-4). Κάθε πόρος μπορεί – μεταξύ άλλων- να είναι ένα URI (Uniform Resource Identifiers- Καθολικό Αναγνωριστικό Πόρων), το οποίο είναι μια παγκόσμια σύμβαση ονομασίας με μοναδικά χαρακτηριστικά, σαν μια ταυτότητα. Η συσχέτιση ενός URI με έναν πόρο σημαίνει ότι ο καθένας

¹ Η HTML και η XML είναι κατάλληλες για την αναπαράσταση δομημένων εγγράφων, ενώ η RDF(S) και η OWL χρησιμοποιούνται για την περιγραφή και την αναπαράσταση των σχέσεων του περιεχομένου, της σημαιολογικής ερμηνείας του και επιτρέπουν την ανάκτηση του, μέσω «λογικής».

² Για παράδειγμα: έστω το Υποκείμενο είναι ο πόρος για το βιβλίο «Ιστορία», το Κατηγορημα είναι η ιδιότητα «Συγγραφέας» και το Αντικείμενο είναι ο πόρος «Θουκυδίδης». Σε αυτή την περίπτωση η τριπλέτα διαβάζεται σε φυσική γλώσσα ως «Η 'Ιστορία' έχει συγγραφέα το Θουκυδίδη».

θα μπορεί να συνδεθεί με αυτόν, να τον αναζητήσει και να τον χρησιμοποιήσει. Στόχος των RDF και RDF Shema είναι η δημιουργία ενός πλαισίου στο οποίο περιγράφονται οι πόροι και οι σχέσεις τους με τρόπο κατανοητό και επεξεργάσιμο από τις μηχανές. Οι πληροφορίες στον Ιστό θα είναι πλέον αναγνώσιμες και κατανοητές τόσο από τους χρήστες, όσο και τους υπολογιστές (Αγλαΐνη 2016, σ. 13; Ιακωβάκης 2019, σ. 14; Mehra & Kumar 2011, σ. 217; Σαμοθρακίτου 2015, σ. 19; Τριανταφύλλου 2022, σ.σ. 15-6; Χρυσούλας 2009, σ.σ. 61-2 & 89-90). Αυτό το σύστημα οργάνωσης προωθεί τη διασύνδεση των δεδομένων και τη διαλειτουργικότητα των εφαρμογών λόγω της σχεδόν αυτόματης ανταλλαγής δεδομένων, ακόμα και εντός του Παγκόσμιου Ιστού.

Κύρια επιδίωξη του Σημασιολογικού Ιστού είναι η καλύτερη συνεργασία μεταξύ υπολογιστών και χρηστών, γι' αυτό και το ζητούμενο είναι η παρουσίαση των πληροφοριών με τρόπο που προσεγγίζει την ανθρώπινη νόηση και είναι κατανοητός από τους υπολογιστές, στοιχείο που θα βελτιώσει τη διαχείριση και την επεξεργασία τους. Για το σκοπό αυτό ο Σημασιολογικός Ιστός βασίζει τη λειτουργία του σε τρεις αλληλένδετους άξονες: τα Δομημένα Δεδομένα, τις Οντολογίες και την Σημασιολογική Ανάλυση. Με τον όρο Δομημένα Δεδομένα εννοούμε πως τα δεδομένα που καταγράφονται στον Ιστό θα πρέπει να είναι καταγεγραμμένα με τέτοιο τρόπο ώστε να μπορούν να κατανοούνται και να επεξεργάζονται από τους υπολογιστές. Οι Οντολογίες με μια αρκετά απλοϊκή ανάλυση είναι σύνολα τριπλετών, οι οποίες ορίζουν αντικείμενα, έννοιες, ιδιότητες και τις σχέσεις τους περιγράφοντας τα στο πλαίσιο ενός συγκεκριμένου γνωστικού πεδίου. Μέσω των οντολογιών μπορούν οι υπολογιστές να επεξεργάζονται τα δεδομένα και να βελτιώνουν την ποιότητα των ανακρίσεων. Μια οντολογία μπορεί να κατασκευαστεί ή να επεξεργαστεί από οποιονδήποτε. Η σημασιολογία που διατυπώνει πρέπει να ορίζεται από εγκεκριμένους και διατυπωμένους κανόνες, αλλά επίσης πάντα πρέπει να λαμβάνεται υπόψη ότι μια οντολογία, σα 'ζωντανός' οργανισμός, μπορεί να αναπτυχθεί περαιτέρω και να εξελιχθεί, να οδηγήσει σε πιο συγκεκριμένο παρακλάδι, να συγχωνευθεί ή να τροποποιηθεί (Μαυρογιάννη 2007, σ.σ. 103). Η αλλαγή στα νοήματα των οντολογιών ονομάζεται Σημασιολογική Ερμηνεία και είναι η σύνδεση μεταξύ του δομημένου λεξιλογίου με τα πρότυπα της ανθρώπινης λογικής και νόησης. Η αυτοματοποίηση της ερμηνευτικής διαδικασίας είναι ένας από τους στόχους του Σημασιολογικού Ιστού και μια 'φιλοδοξία' της ανάπτυξης των Οντολογιών είναι να μετατοπιστεί το σημασιολογικό ερμηνευτικό φορτίο από το χρήστη -που είναι σήμερα λόγω των μειονεκτημάτων του Παγκόσμιου Ιστού- στον υπολογιστή, ο οποίος θα προσεγγίσει την ανθρώπινη ερμηνευτική ικανότητα. Η Σημασιολογική Ανάλυση είναι η διαδικασία στην οποία οι υπολογιστές αναλύουν και κατανοούν τα δεδομένα και τις σχέσεις που καταγράφονται σε μια οντολογία, τα ερμηνεύουν και τα αξιολογούν. Επιτυγχάνεται με τη γλωσσική περιγραφή, τη χρήση δομημένων λεξιλογίων και τις οντολογίες και τους επιτρέπει να ανακτούν τις πληροφορίες που ζητούν οι χρήστες με μεγαλύτερη ακρίβεια. Με αυτόν τον τρόπο η επεξεργασία των υπολογιστών θα προσεγγίσει την ανθρώπινη νόηση, σύμφωνα με λογικές αρχές και τα ερμηνευτικά της πρότυπα (Μαυρογιάννη 2007, σ.σ. 101-3).

Οι προκλήσεις που αντιμετωπίζει ο Σημασιολογικός Ιστός είναι σημαντικές και πηγάζουν από την ανταπόκριση των βασικών συστατικών του στις νέες απαιτήσεις, στην εξέλιξη και διαχείρισή τους, καθώς και τη γενίκευση της χρήσης τους. Κύριο πρόβλημα είναι ο όγκος των πληροφοριών που πρέπει να διασυνδεθούν και πως η οργάνωση θα καταφέρει να περιορίσει και την ποσότητα, αυξάνοντας ταυτόχρονα την ακρίβεια και την αξιοπιστία των αποτελεσμάτων. Η διαχείριση του πλήθους των δεδομένων έχει να κάνει και με την ασυνέπεια, με τις αντιφάσεις, επαναλήψεις και τα λογικά σφάλματα που θα προκύψουν από τις μεγάλες και πολύπλοκες οντολογίες. Αφού ο Σημασιολογικός Ιστός αποτελείται από Οντολογίες και τις τεχνολογίες που τις συνοδεύουν, η πρόοδος του βασίζεται από την ποιότητα και την δική τους εξέλιξη. Επομένως εξαρτάται από παραμέτρους όπως το κόστος ανάπτυξης και συντήρησής, η σωστή, λογική και δομημένη διατύπωση, η διαχείριση από τους χρήστες, η εξέλιξη των συμφωνημένων γλωσσικών προτύπων που τον υποστηρίζουν. Η μεγαλύτερη πρόοδος και η καθολική χρήση των οντολογιών θα οδηγήσουν σε πλήρη ανάπτυξη των δυνατοτήτων του Ιστού, στην καλύτερη οργάνωση του υλικού και την αποτελεσματικότερη ανάκτηση των πληροφοριών από τους χρήστες, εκπληρώνοντας την κύρια επιδίωξη όταν ο Σημασιολογικός Ιστός ήταν μόνο μια θεωρητική μελλοντική εξέλιξη (Ιακωβάκης 2019, σ.σ. 9-10).

β. Οντολογίες

Η Οντολογία για τη φιλοσοφία είναι η μελέτη των όντων και των ιδιοτήτων τους³. Στον κόσμο της πληροφορικής αποτελούν ένα εργαλείο αναπαράστασης της γνώσης και κύριο άξονα της λειτουργίας του Σημαιολογικού Ιστού. Οι Οντολογίες δεν αποκλίνουν από τη φιλοσοφική ερμηνεία τους καθότι αφορούν επίσημα, καθορισμένα και τυποποιημένα σύνολα από έννοιες στα οποία επίσης περιγράφονται οι ιδιότητες και οι μεταξύ τους σχέσεις. Στοχεύουν στην κάλυψη της ανάγκης διαμοιρασμού και επαναχρησιμοποίησης της γνώσης, μετατοπίζοντας το ενδιαφέρον στη σχεδίαση ενός λειτουργικού πλαισίου αναπαράστασης και οργάνωσης των δεδομένων και όχι στο ίδιο το περιεχόμενο.

Υπάρχουν πολλοί παραπλήσιοι ορισμοί στη διεθνή βιβλιογραφία από επιστήμονες που προσπάθησαν από διάφορες οπτικές να καλύψουν την πολυπλοκότητά της και να περιγράψουν μια οντολογία⁴. Συλλογικά, εξάγουμε το συμπέρασμα πως μια Οντολογία αποτελεί το πλαίσιο μιας αυστηρής, τυπικής, σαφούς και ακριβούς καθολικής περιγραφής μιας αναπαράστασης εννοιών, ιδιοτήτων, σχέσεων και περιορισμών ενός συγκεκριμένου πεδίου γνώσης και ενδιαφέροντος. Σαφής (ή ρητή) σημαίνει ότι οι έννοιες και οι ορισμοί είναι προσδιορισμένοι με ακρίβεια και συνέπεια, χωρίς λογικές υπερβάσεις ή παρανοήσεις. Αυστηρή, καθώς πρέπει να διέπεται από κανόνες που την διατηρούν μηχανικά αναγνώσιμη. Τυπική, ώστε να γίνεται κατανοητή και να επεξεργάζεται από τους υπολογιστές και να είναι πλήρως εκμεταλλεύσιμη χωρίς σφάλματα. Τέλος, μια οντολογία οφείλει να είναι δομημένη με τρόπο ούτως ώστε το περιεχόμενό της να γίνεται συλλογικά αποδεκτό και να καλύπτει τα ανθρώπινα ερμηνευτικά πρότυπα. Αυτή η ανθρώπινη μορφή οργάνωσης, ονομάζεται εννοιολογική σύλληψη και νοείται ως η νοητική αναπαράσταση του τρόπου οργάνωσης και κατανόησης των σχέσεων των όρων ενός γνωστικού τομέα.

Με βάση το “βαθμό τυπικότητας” τους οι οντολογίες χωρίζονται σε κατηγορίες: τις Άτυπες (γραμμένες σε φυσική γλώσσα και κατανοητές από τον άνθρωπο), Ημιάτυπες (χρησιμοποιούν φυσικές γλώσσες με περιορισμένα δομημένο τρόπο), Ημιτυπικές (εκφράζονται με δομημένες γλώσσες προγραμματισμού και διαθέτουν σημειολογία και σύνταξη) και τις (Αυστηρά) Τυπικές (βασίζονται σε λογικές γλώσσες προγραμματισμού και έχουν σημαιολογία και αυστηρή σύνταξη). Οι περισσότερες οντολογίες ανήκουν στις ημιτυπικές διότι αυτή η μορφή συνδυάζει στοιχειά από τις άτυπες και τις αυστηρά τυπικές ή αποτελούν συνδυασμό ημιτυπικών και αυστηρά τυπικών (Ιακωβάκης 2019, σ. 3; Κατσουλάκου et al. 2015, σ.σ. 27-8; Τριανταφύλλου 2022, σ. 10; Χρυσούλας 2009, σ.σ. 93-5). Υπάρχει η πρόταση και για μια άλλου είδους κατηγοριοποίηση, με βάση τη “γενικότητα”, δηλαδή τον τομέα στον οποίο μπορεί να χρησιμοποιηθεί το πρότυπο της οντολογίας. Χωρίζονται στις “Οντολογίες υψηλού επιπέδου” στις οποίες περιγράφονται γενικές έννοιες, στις “Οντολογίες πεδίο ορισμού” στις οποίες καταγράφεται ένα συγκεκριμένο γνωστικό πεδίο, στις “Οντολογίες μεθοδολογίας ή εργασιών” στις οποίες περιγράφουμε τα δεδομένα μιας συγκεκριμένης εργασίας και τέλος στις “Οντολογίες εφαρμογής” για την περιγραφή εννοιών ενός πεδίου ενδιαφέροντος ή μιας εφαρμογής (Ιακωβάκης 2019, σ. 3; Κατσουλάκου et al. 2015, σ. 29; Πετράκης 2011, σ.σ. 31-8; Σαραντοπούλου 2014, σ.σ. 59-61; Τριανταφύλλου 2022, σ. 11). Ανεξάρτητα από τον τρόπο με τον οποίο γίνονται, οι κατηγοριοποιήσεις των οντολογιών είναι καθαρά τεχνικό ζήτημα. Οι κατηγορίες δεν είναι στεγανές και κατά συνθήκη μια οντολογία ανάλογα την οπτική ενδέχεται να έχει μεικτά χαρακτηριστικά και να εντάσσεται σε πολλαπλές κατηγορίες. Σε κάθε περίπτωση η κατηγοριοποίηση δεν επηρεάζει τους χρήστες, ενώ συνήθως δεν έχει ιδιαίτερη σημασία και για τους δημιουργούς της οντολογίας. Αυτό που είναι σημαντικό, είναι η συνέπεια στις αρχές στις οποίες βασίζεται μια οντολογία, όπως προκύπτουν από τον ορισμό της.

³ Αριστοτέλης, Μετά τα Φυσικά, 3.1003a.21 “ἔστιν ἐπιστήμη τις ἢ θεωρεῖ τὸ ὄν ἢ ὄν καὶ τὰ τούτῳ ὑπάρχοντα καθ’ αὐτό.... οὐδεμία γὰρ τῶν ἄλλων ἐπισκοπεῖ καθόλου περὶ τοῦ ὄντος ἢ ὄν, ἀλλὰ μέρος αὐτοῦ τι ἀποτεμύμεται....” (Υπάρχει μια επιστήμη που εξετάζει το *ον* ως *ον* και τις ιδιότητες που αυτό έχει...καμία άλλη δεν εξετάζει με κάποιο τρόπο το *ον* καθ’ αυτό αλλά ασχολούνται με κάποιο κομμάτι του...)

⁴ Συγκεντρωτικά οι ορισμοί στα Falbo & de Menezes & da Rocha 1998, σ. 350; Ιακωβάκης 2019, σ. 1; Κουσιώρας 2018, σ.σ. 13-5; Πετράκης 2011, σ.σ. 29-30; Σαραντοπούλου 2014, σ.σ. 51-8; Τριανταφύλλου 2022, σ.σ. 8-12; Χρυσούλας 2009, σ.σ. 67-9. Στα παραπάνω αναφέρονται οι ορισμοί διάφορων άλλων επιστημόνων όπως προτάθηκαν ανά τα χρόνια και καταγράφονται οι βασικές αρχές στις οποίες βασίζεται η κατασκευή μιας οντολογίας.

Η οντολογία πρέπει να βασίζει την οργάνωση και το περιεχόμενό της κατασκευή της στις απαντήσεις κάποιων ερωτημάτων που θέτει στο ξεκίνημα ο κατασκευαστής της: να έχει σαφή προσανατολισμό και στόχους, να είναι φτιαγμένη με ευέλικτο τρόπο ώστε να επαναχρησιμοποιείται (για το σκοπό αυτό πρέπει να μπορεί να επεξεργάζεται, προσαρμόζεται, αναθεωρείται και επεκτείνεται εύκολα), να είναι λειτουργική και δομημένη με τρόπο που να διαβάζεται από τους υπολογιστές αναπαριστώντας κατανοητά την ανθρώπινη εννοιολογική σύλληψη, να είναι αξιοποιήσιμη από ψηφιακά εργαλεία και εφαρμογές και ωφέλιμη σε όσο το δυνατόν ευρύτερο κοινό και τέλος, να μην είναι ένα απλά θεωρητικό πρότυπο αλλά να προσφέρει πρακτική και ουσιαστική λύση σε κάποιο συγκεκριμένο πρόβλημα (το οποίο τέθηκε εξ αρχής) είτε αυτό πρόκειται για μια ανάλυση και εύρεση λύσεων, είτε οργάνωση και εύκολη πρόσβαση σε αρχεία⁵.

Μια οντολογία είναι μια συνολική αποτύπωση του κατάλληλου συνδυασμού των στοιχείων που την αποτελούν. Κατασκευάζεται με σύμμεκτο πρότυπο ιεραρχίας προκειμένου να εξυπηρετείται ο σκοπός της αναπαράστασης των πολύπλοκων σχέσεων που την αποτελούν (Μαυρογιάννη 2007, σ.σ. 104-8 & 116-8). Στη βάση της αποτελείται από μια ιεραρχική ταξινόμηση των κλάσεων. Οι κλάσεις είναι το κύριο συστατικό μιας οντολογίας (Μαυρογιάννη 2007, σ.σ. 103, 109-122, 115-7; Uschold & Gruninger 1996, σ.σ. 93-155) και είναι κομβικές καθώς πρέπει να αποτυπώνουν το σκοπό και τους στόχους της οντολογίας, γι' αυτό πρέπει να είναι προσεκτικά επιλεγμένες και ακριβείς. Κάθε κλάση είναι μια οντότητα που αντιπροσωπεύει μια έννοια ή ένα αντικείμενο και μπορεί ανήκει σε μια ευρύτερη ομάδα (την υπερκλάση) ή και να διαθέτει άλλες κλάσεις ως νοηματικά της υποσύνολα (τις υποκλάσεις). Όλες οι κλάσεις μπορούν να ανήκουν σε πολλές υπερκλάσεις ή (και) να διαθέτουν πολλές υποκλάσεις. Τα συγκεκριμένα παραδείγματα της οντότητας της κλάσης, λέγονται στιγμιότυπα. Το σύστημά των οντολογιών όμως είναι πολύπλοκότερο από μια πυραμοειδή τύπου οργάνωση πληροφοριών εξαιτίας των ιδιοτήτων των κλάσεων (Μαυρογιάννη 2007, 11-2). Οι ιδιότητες είναι έννοιες, συνήθως ρηματικοί τύποι σε μια φυσική γλώσσα και χωρίζονται σε δύο κατηγορίες: αυτές που σχετίζονται με αντικείμενα ("object properties") και συνδέουν δύο στιγμιότυπα κλάσεων και αυτές που αφορούν δεδομένα ("datatype properties") και συνδέουν ένα στιγμιότυπο με μια συγκεκριμένη τιμή που το χαρακτηρίζει⁶. Εκτός από την οργάνωση των κλάσεων, των ιδιοτήτων και των στιγμιότυπών τους, για μια οντολογία απαιτείται και ένα λειτουργικό πλαίσιο από περιορισμούς (κανόνες για τη σωστή σύνδεση των κλάσεων και των ιδιοτήτων) και αξιώματα (λογικές που καθορίζουν τη σημασιολογία και τις επιθυμητές τιμές των αποτελεσμάτων) που ρυθμίζουν την ανάλυση των στοιχείων και το φάσμα των διαθέσιμων αποτελεσμάτων κατά την ανάκτηση των δεδομένων (Falbo et al. 1998, σ.σ. 352-3). Οι περιορισμοί είναι κανόνες που εφαρμόζονται στις κλάσεις και τις ιδιότητες και διασφαλίζουν τη σωστή έκφρασή τους, οπότε τη συνέπεια και τη λειτουργικότητα της οντολογίας συνολικά. Οι περιορισμοί μπορεί να είναι πολλών ειδών και να υπολογιστούν ως ένα είδος αξιώματος: ιδιαίτερα σημαντικοί είναι οι περιορισμοί Domain-Range⁷, η μοναδικότητα, η ελάχιστη και μέγιστη δυνατή τιμή και το πλήθος των ιδιοτήτων ενός στιγμιότυπου. Τα αξιώματα υπόκεινται σε κατηγορίες ανάλογα το σκοπό τους: τα αξιώματα παραγωγής για την εξαγωγή νέας γνώσης, τα αξιώματα ενοποίησης τα οποία προσδιορίζουν τους περιορισμούς στον ορισμό μιας έννοιας ή στην ανάκτηση του επιθυμητού στιγμιότυπου, τα επιστημολογικά αξιώματα τα οποία καθοδηγούν τη διαμόρφωση των εννοιών και τη δομική μορφή της οντολογίας και τα οντολογικά αξιώματα για τους περιορισμούς των σχέσεων και των χαρακτηριστικών των κλάσεων και των ιδιοτήτων.

Ο όγκος των πληροφοριών είναι τεράστιος και εξακολουθεί να αναπτύσσεται, επομένως δεν είναι δυνατή η διαχείρισή του χωρίς τις κατάλληλες συνδέσεις. Γι' αυτό και οι οντολογίες αποτελούν αναγκαιότητα. Στο πλαίσιο του Σημασιολογικού Ιστού κορυφαίος στόχος για την ανάπτυξη και τη χρήση των οντολογιών είναι η διάδοση της πληροφορίας και του περιεχομένου της μεταξύ των ανθρώπων (ακόμα και με διαφορετικά επιστημονικά ενδιαφέροντα και βάσεις), των υπολογιστών ή και στις δύο πλευρές, με γρήγορο, κατανοητό και αξιόπιστο τρόπο. Στο ίδιο πνεύμα

⁵ Βλ. στο παρόν, κεφάλαιο 2.1.α. σχετικά με τη διαδικασία και τις απαιτήσεις για τη δημιουργία μιας οντολογίας.

⁶ Παράδειγμα από την οντολογία της εργασίας: η object property όπως "cd:isAuthorOf" ("είναι συγγραφέας του...") ενώνει στιγμιότυπα της κλάσης "dbpedia:Author" με το κατάλληλο στιγμιότυπο της κλάσης "bibo:Book". Μια "datatype property", παράδειγμα η "dcterms:title" ("έχει τίτλο") χαρακτηρίζει την κλάση "bibo:Book" με μια μοναδική τιμή (κάθε βιβλίο έχει έναν μοναδικό, δικό του τίτλο).

⁷ "Πεδίο Ορισμού- Σύνολο τιμών". Ο περιορισμός "Domain" προσδιορίζει την κλάση όπου εφαρμόζεται μια ιδιότητα και ο περιορισμός "Range" την κλάση στην οποία πρέπει να ανήκουν οι τιμές της ιδιότητας. Παράδειγμα, η ιδιότητα "creator" έχει "Domain" την κλάση "Βιβλίο" και "Range" την κλάση "Συγγραφέας". Επομένως όλα τα στιγμιότυπα της κλάσης "Βιβλίο" υποχρεωτικά πρέπει να "έχουν συγγραφέα" ένα στιγμιότυπο της κλάσης "Συγγραφέας".

τοποθετείται και η επαναχρησιμοποίηση των οντολογιών. Η χρήση της επιστημονικής γνώσης και η δυνατότητα πλέον μέσω της τεχνολογίας αυτή να χρησιμοποιείται, ενώνεται, εμπλουτίζεται και τροποποιείται εξελίσσει την έρευνα προς άλλα γνωστικά πεδία βελτιώνοντας τα τελικά πορίσματα. Οι ρητές προδιαγραφές των εννοιών οποιοδήποτε γνωστικού πεδίου και η οργάνωσή τους διευκολύνουν την κατανόηση των όρων και τη σημασιολογία, στοιχείο που εξυπηρετεί κυρίως τους νέους ή άπειρους ερευνητές. Προσφέρεται επίσης ένα πλαίσιο μέσα στο οποίο τα (μετα)δεδομένα διατηρούνται ανεξάρτητα από το ίδιο το αντικείμενο. Οι οργανωτικές δυνατότητες των οντολογιών επιτρέπουν στους μελετητές να χρησιμοποιήσουν (ή και να δημιουργήσουν) λειτουργικές, ακριβείς και πολυεπίπεδες βάσεις δεδομένων. Η δυνατότητα της συμπερίληψης διάφορων τύπων δεδομένων και αρχείων ενισχύει επίσης τη διεπιστημονική έρευνα. Τέλος, μια σωστά δομημένη, καλά οργανωμένη και “πλούσια” οντολογία επιτρέπει την εύκολη, ποιοτική και αποτελεσματική ανάκτηση δεδομένων, επομένως εξασφαλίζονται περισσότερο ορθά, συλλογικά και αξιόπιστα αποτελέσματα. Για την ιστορική επιστήμη, ειδικότερα, η δομημένη και συστηματική προσέγγιση οργάνωσης και αναπαράστασης των δεδομένων που προσφέρουν οι οντολογίες έχει πολλαπλά οφέλη. Η διάδοση των δεδομένων μεταξύ των επιστημόνων (ακόμα και διαφορετικών επιστημονικών πεδίων) και των εργαλείων τους προάγει την έρευνα εφόσον βοηθά την επικοινωνία και τη συνεργασία τους, ώστε να προκύπτουν καθολικά, αξιόπιστα και πολύπλευρα αποτελέσματα που θα αξιολογηθούν συλλογικά. Τέλος, προσφέρουν ένα χρήσιμο εργαλείο για τη συγκέντρωση, οργάνωση, παρουσίαση και συσχέτιση των δεδομένων (μεταξύ άλλων και των πρωτογενών γραμματειακών πηγών) εξυπηρετώντας τη δημιουργία προσωπικού αρχείου και καθιστούν την προσωπική έρευνα ευκολότερη, ταχύτερη και αποτελεσματικότερη (Ιακωβάκης 2019, σ. 5; Σαραντοπούλου 2014, σ.σ. 55-7) ⁸.

⁸ <http://www.Protégé.stanford.edu> (Noy & McGuiness, Ontology Development 101: A Guide to Creating Your First Ontology, pp. 1-2.)

1.2. Η Καταγραφή των Αρχαίων Πηγών

Η ιστορική επιστήμη⁹ είναι η συστηματική μελέτη όψεων του παρελθόντος με στόχο την κατανόηση και τη γνώση κάθε ανθρώπινης δραστηριότητας. Για το σκοπό αυτό ο ιστορικός χρησιμοποιεί “πηγές”, «...το σύνολο των ιχνών που έχουν αφήσει οι φορείς τις δράσης στο παρελθόν και πάνω στα οποία ο ερευνητής στηρίζει τη δουλειά του» (Offenstadt 2005, σσ.162-3). Οι πηγές αποκαλύπτουν πληροφορίες, στοιχεία και γεγονότα ενός πολύ συγκεκριμένου χώρου μιας μοναδικής στιγμής της ανθρώπινης ιστορικής πορείας¹⁰. Ως πηγές μπορεί να θεωρηθεί κάθε είδους - γραπτό ή όχι- τεκμήριο (κείμενο, έγγραφο, κατάλογος, επιγραφή, χάρτες, εφημερίδα, νομίσματα, αρχαιολογικά κατάλοιπα κ.α.) που μπορεί να προσφέρει πληροφόρηση για οποιαδήποτε πτυχή της ανθρώπινης ζωής και δραστηριότητας. Κάθε ιστορικός, όποιο κι αν είναι το θέμα με το οποίο καταπιάνεται, πρέπει και οφείλει να ξεκινά την έρευνά του από τις πηγές και να βασίζεται τα συμπεράσματά του στην ανάλυσή τους. Αν οι πηγές για οποιοδήποτε λόγο δεν είναι οργανωμένες και προσβάσιμες δεν μπορούν να αξιοποιηθούν από τον ιστορικό, επομένως “χάνονται” (Βόγλη 2015, σ.σ. 100-1). Η μελέτη των πηγών περιλαμβάνει κάποια βήματα, τα οποία, όπως και τα στάδια ανάπτυξης μιας οντολογίας, δεν είναι ξεχωριστά και αυστηρά ιεραρχημένα, αλλά πραγματοποιούνται ως ένα βαθμό παράλληλα και εξελίσσονται καθώς προχωράει και η έρευνα. Καθένα περιλαμβάνει και επιμέρους εργασίες. Ούτε αυτές είναι δυνατόν να προσδιοριστούν με χρονική ακρίβεια, καθώς η εξοικείωση του ιστορικού με τις πηγές και το θέμα του επηρεάζει το χρόνο που κάθε βήμα θα τον απασχολήσει. Ενδεχομένως να προστίθενται και επιπλέον διαδικασίες προκειμένου ο ιστορικός να καλύψει κάποιο γνωστικό κενό ή άλλο πρόβλημα. Γενικά η ακολουθία διαμορφώνεται ως εξής: 1. συλλογή των πηγών (περιλαμβάνεται η φυσική και η ψηφιακή αναζήτηση), 2. οργάνωση (καταλογογράφηση) η οποία απαιτεί ταξινόμηση και διαλογή του υλικού, 3. αρχειοθέτηση (καταχώριση με συστηματικό τρόπο) και αποθήκευση, 4. ανάλυση, μέσω προσεκτικής μελέτης και αξιολόγησης του υλικού και 5. ερμηνεία, κατά την οποία τα συμπεράσματα από την ανάλυση τοποθετούνται σε συγκεκριμένο πλαίσιο και σχολιάζονται. Περιγράφοντας πιο αναλυτικά:

Αφού ο ερευνητής αποφασίσει και προσδιορίσει το είδος των πηγών με τις οποίες θα καταπιαστεί, το πρώτο ουσιαστικά βήμα που ξεκινά είναι η συλλογή τους. Για τον μελετητή της αρχαίας ιστορίας οι πηγές μπορεί να είναι κείμενα επιγραφών, χειρόγραφα, πάπυροι, νομίσματα και αρχαιολογικά ευρήματα. Κάθε μια τους, δεν αποτελεί μια μοναδική πηγή πληροφόρησης, για παράδειγμα ένα νόμισμα δίνει πληροφορίες μέσω της παράστασης που φέρει, αλλά και τυχόν επιγραφές του, καθώς και το είδος του μετάλλου στο οποίο κόπηκε. Στην παρούσα εργασί, θα ασχοληθούμε με τις γραμματειακές πηγές. Όμως, όλες “απαιτούν” από τον μελετητή ανάλογη εξοικείωση με το είδος, χρόνο για τη συλλογή τους και επιμέλεια στη συλλογή και την ανάλυση.

Επόμενο βήμα από τη στιγμή που θα συλλεχθούν οι πηγές- αν και πρακτικά είναι μια ταυτόχρονη εργασία- είναι η καταλογογράφηση¹¹ και αρχειοθέτησή τους. Πρακτικά είναι ταυτόχρονες ενέργειες με τη συλλογή, καθώς δεν ωφελεί ο ιστορικός να ανακαλύπτει μια πηγή και να μη προχωρά σε μια έστω τυπική διαλογή. Αν προχωρήσει στην αναζήτηση, στο τέλος θα χρειαστεί πολύ χρόνο και κόπο καθώς το υλικό θα είναι τεράστιο. Η καταλογογράφηση είναι η ταξινόμηση και η καταγραφή, ενώ η αρχειοθέτηση είναι η οργάνωση και η αποθήκευσή τους. Στη διάρκεια αυτών των -επίσης παράλληλων- διαδικασιών, οι πηγές χωρίζονται σε κατηγορίες που θα ορίσει ο ιστορικός. Η ταξινόμησή τους μπορεί να γίνεται με βάση τη χρονολόγηση, το είδος της πηγής, την γεωγραφική προέλευση, τον συγγραφέα ή το ιδιαίτερο θέμα. Συνήθως η κατηγοριοποίηση γίνεται προσμετρώντας συνδυαστικά περισσότερα του ενός στοιχεία για τη διευκόλυνση

⁹ Βλ. συγκεντρωτικά Βόγλη 2015, σ.σ. 25-40, όπου στα κεφάλαια 1.1.2 κ 1.1.3, παρατίθεται η γνώμη πολλών ιστορικών και θεωρητικών του 19^{ου}-20^{ου} αιώνα για την αναγνώριση της ιστορίας ως επιστήμης και την υποστήριξη του ισχυρισμού. Οι οπτικές και οι ορισμοί καθενός διαφέρουν, όμως η επιστημονικότητα της Ιστορίας δεν αμφισβητήθηκε ποτέ.

¹⁰ Αυτή η αποσπασματικότητα μπορεί να εμποδίζει την εξαγωγή ποιοτικότερων και εκτενών συμπερασμάτων γι' αυτό και οι πηγές πρέπει να αξιολογούνται και να μελετώνται με ιδιαίτερη προσοχή. Κάθε πηγή όμως έχει μεγάλη σημασία για τον τόπο ή την κατάσταση στην οποία αναφέρεται. Αυτό γίνεται φανερό και για την ιδιαίτερη μελέτη περίπτωσης της εργασίας και αποτέλεσε και ένα από τα κίνητρα για την πραγματοποίησή της.

¹¹ Σύμφωνα με τον ορισμό στον Sales 2005, 18 "Cataloging is the process of creating bibliographic records of works according to accepted rules or standards" (Η καταλογογράφηση είναι η διαδικασία δημιουργίας βιβλιογραφικών εγγραφών σύμφωνα με αποδεκτούς κανόνες ή πρότυπα). Για την καταλογογράφηση και την πρακτική της, ο.π., 18-9.

της διαχείρισης του υλικού και τη διαμόρφωση επιμέρους ομάδων- αυτό θα ωφελήσει αργότερα, καθώς για κάποιο συγκεκριμένο θέμα χρειάζεται μια συγκεκριμένη υποομάδα πηγών και όχι το σύνολο του υλικού-. Όσο αναλυτικότερη μπορεί να γίνει αυτή η ομαδοποίηση είναι προτιμότερη για τη συνέχεια. Για τη μεγαλύτερη συνέπεια και ακρίβειά της, βασική προϋπόθεση είναι μια πρώτη, έστω και επιφανειακή, επαφή, με το περιεχόμενο της πηγής. Η μερική γλωσσική κατανόησή του, απαιτεί τη συνεργασία των ιστορικών με άλλους επιστήμονες παραπλήσιων κλάδων (Βόγλη 2015, σσ. 45-7). Των “άμεσων συνεργατών” τους, αρχαιολόγων και φιλόλογων και ορισμένων εξειδικευμένων επιστημόνων (λ.χ. παλαιογράφοι, νομισματολόγοι, επιγραφολόγοι...). Έπειτα ακολουθεί η αρχειοθέτηση των πηγών, η οποία περιλαμβάνει την οργανωμένη και τυποποιημένη καταχώρηση και συστηματική αποθήκευσή τους σε αναλογικά ή ψηφιακά αρχεία (λίστες εγγράφων, καταλόγους βιβλιοθηκών, βάσεις δεδομένων κ.α.). Η αρχειοθέτηση δεν μπορεί να είναι ανεξάρτητη από την ταξινόμηση και την ομαδοποίηση που προηγήθηκε. Εξασφαλίζει τη διατήρηση της πληροφορίας¹², την δυνατότητα αυξημένης και εύκολης πρόσβασης και αναζήτησης, ενώ επίσης διευκολύνει την γρήγορη και απλή ανάκτηση της όταν θα χρειαστεί.

Αφού έχει κατηγοριοποιήσει και οργανώσει τις πηγές του, η επόμενη εργασία για τον ερευνητή είναι η ανάλυση και η ερμηνεία τους. Πρώτα πραγματοποιεί και μια επιπλέον διαλογή, καθώς δεν είναι όλο το υλικό που επέλεξε ποιοτικό και αξιοποιήσιμο τελικά για το θέμα του. Κατόπιν ο ιστορικός μελετά προσεκτικά τις πηγές του και εξετάζει κριτικά τις πηγές του λαμβάνοντας υπόψη το πλαίσιο στο οποίο εντάσσονται – γι’ αυτό απαιτείται η πολύ καλή γνώση της γλώσσας, του θέματος και του περιεχομένου-, τα αξιολογεί και γενικά συγκεντρώνει τα απαραίτητα στοιχεία για την εξαγωγή των συμπερασμάτων. Πλέον ο ιστορικός έχει φθάσει στο τελευταίο στάδιο, την ερμηνεία. Είναι η διαδικασία κατά την οποία τα παραπάνω στοιχεία, που προέκυψαν από την ανάλυση των πηγών, εντάσσονται σε ένα συγκεκριμένο αλλά και ευρύτερο πολιτιστικό και ιστορικό πλαίσιο, αποκτούν νόημα και σχολιάζονται για να εξαχθούν τα τελικά συμπεράσματα.

Οι πηγές συμβάλλουν τα μέγιστα στην διαμόρφωση μιας πιο ολοκληρωμένης εικόνας για τη ζωή στην αρχαιότητα. Για την μελέτη τους το σημαντικότερο ρόλο έχει η καταλογογράφηση και η αρχειοθέτησή τους. Οι δύο αυτές διαδικασίες περιλαμβάνουν τις εργασίες που περιγράψαμε και μπορούν συνολικά να συνοψιστούν στον όρο “καταγραφή”. Αυτή αφορά τη συγκέντρωση και καταλογογράφηση των τεκμηρίων, την καταχώρησή τους σε αρχεία μαζί με τις βασικές πληροφορίες τους να τα “συνοδεύουν” και την αποθήκευσή τους με μορφή λίστας σε κάποια μορφή αποθετήριο. Σκοπός της είναι η διατήρηση των συλλογών του υλικού και των βασικών (μετα)δεδομένων του στο χρόνο και η δυνατότητα επαφής τους με ένα ευρύτερο κοινό. Στις μέρες μας η καταγραφή περιλαμβάνει πλέον και την ψηφιοποίηση, τη μετατροπή δηλαδή των φυσικών αρχείων στο ψηφιακό τους ανάλογο. Η ψηφιοποίηση θεωρείται ένα σύγχρονο μέσο διατήρησης, αλλά σύμφωνα και με τον ορισμό της δεν συμπίπτει ακριβώς με την καταγραφή.

Η ψηφιοποίηση αποτελεί σημαντικό εργαλείο της τεχνολογικής εξέλιξης και της προσφοράς της στην πρόοδο των επιστημών. Για την ιστορική επιστήμη είναι ένα ιδιαίτερα σημαντικό εργαλείο, στο οποίο βασίζεται πλέον η διαδικασία της καταγραφής των πηγών, η διατήρηση και κατ’ επέκταση η πρόσβαση στις πληροφορίες. Ψηφιοποίηση ονομάζουμε τη μετατροπή κάθε είδους φυσικής μορφής τεκμηρίου στο ψηφιακό ανάλογο του μέσω της χρήσης εργαλείων της τεχνολογίας (φωτογραφικές μηχανές, σαρωτές κ.α.) και έπειτα την ηλεκτρονική αποθήκευσή του σε ένα ηλεκτρονικό αποθετήριο (Hitchcock & Shoemaker 2015, σ.σ. 83). Ως πρακτική δεν είναι σύνθετη, ούτε ξένη για την καθημερινότητα, όμως τα οφέλη για τις επιστήμες δεν είναι πάντα αντιληπτά.

Η ψηφιοποίηση των τεκμηρίων συμβάλει με πολλούς τρόπους στην πρόοδο της έρευνας. Ένα από τα ζητήματα της ψηφιακής εποχής είναι η διατήρηση των φυσικών αντικειμένων. Με την ψηφιοποίηση όχι μόνο διατηρείται η απόδειξη της ύπαρξής τους ακόμα και αν χαθούν στο φυσικό κόσμο, αλλά επιπλέον προφυλάσσονται από την φθορά που θα προξενούσε η συνεχής έκθεση και επαφή του με τους χρήστες. Κάθε ενδιαφερόμενος μπορεί πλέον να αποκτήσει ένα αντίγραφο, να το αποθηκεύσει και να το έχει στην κατοχή του, να το διαμοιράζει και να το επεξεργάζεται. Η έρευνα επιταχύνεται καθώς τα αρχεία μπορούν να ανακαλούνται άμεσα και από οπουδήποτε. Όποιο στοιχείο “ανεβαίνει” στο Διαδίκτυο, γνωστοποιείται σε όλο τον κόσμο αυτόματα και γίνεται αντικείμενο συζήτησης, αξιολόγησης και προβληματισμών. Προωθείται η επικοινωνία και η

¹² Αυτό αποτελεί και την ουσία της Σημαιολογικής εξέλιξης. Στόχος είναι να (δια)σώζεται όχι η ίδια η πηγή στην αναλογική της μορφή, αλλά περισσότερο η περιγραφή της. Ειδικά για την αρχαία ιστορία, είναι σημαντικό ο ερευνητής να γνωρίζει την ύπαρξη κάποιου τεκμηρίου, έστω και αν για οποιοδήποτε λόγο δεν έχει πρόσβαση σε αυτό.

συνεργασία των μελετητών (και από διαφορετικές ειδικότητες), αλλά και η αλληλεπίδραση με το κοινό. Η ευρεία χρήση αντανάκλαται τόσο στο πλήθος των ανθρώπων στους οποίους “φτάνει” ένα αρχείο, όσο και στον τρόπο, ο οποίος δεν δεσμεύεται από την τοποθεσία ή την κατάσταση του φυσικού αντικειμένου. Η ανάκτηση δεδομένων από μια βάση ψηφιοποιημένων αρχείων γίνεται επίσης πιο γρήγορα και εύκολα σε σχέση με την παραδοσιακή μέθοδο έρευνας (αναζήτηση αντιγράφων, διάβασμα, σημειώσεις κ.α.) και επιταχύνεται περισσότερο αν διατίθενται τα κατάλληλα προγράμματα για τη σωστή αρχειοθέτηση και επεξεργασία τους (Kelly 2013, σ.σ. 13-4; Kuny 1997, σ.σ. 1-6; Chouchene 2019, σ.σ. 74-9; Hitchcock & Shoemaker 2015, 83).

Κάθε νέα τεχνολογία που φαίνεται να “λύνει” κάποιο ζήτημα δεν πρέπει να αντιμετωπίζεται ως πανάκεια. Από το “ξεκίνημα” υπήρξαν φωνές που τόνιζαν τους προβληματισμούς γι’ αυτά που επιφέρει η ψηφιοποίηση και τονίζονταν οι ενέργειες που πρέπει να γίνουν προκειμένου να περιοριστούν οι αρνητικές προεκτάσεις και πως πρέπει να διαχειριζόμαστε γενικά την τεχνολογία. Πολλά προβλήματα διαπιστώνονται μέχρι και τις μέρες μας που η ψηφιοποίηση των εγγράφων δεν είναι πολυτέλεια αλλά ένας σύγχρονος τρόπος προσέγγισης της πληροφορίας.

Η δημιουργία ενός ψηφιακού αντιγράφου για ένα αντικείμενο μπορεί να το προφυλάσσει από τις φθορές της καθημερινής χρήσης, αλλά επίσης εγκυμονεί τον κίνδυνο ακόμα και για την ανάγκη ύπαρξης του πρωτότυπου. Ενδεχόμενη απώλειά του βέβαια υποβαθμίζει την αξιοπιστία και του ψηφιακού αρχείου, το οποίο μπορεί εύκολα να παραποιηθεί ή να αλλοιωθεί. Είναι μια σημαντική πρόκληση για μια εποχή στην οποία προσπαθούμε να εξισορροπήσουμε τη διατήρηση των φυσικών αντικειμένων με την εξασφάλιση των ψηφιακών αντιγράφων τους. Όπως όλοι οι επιστήμονες, έτσι και οι ιστορικοί στηρίζονται στην τεχνολογία για την πρόοδο των επιστημών τους. Την ίδια στιγμή, όμως, παρατηρείται το φαινόμενο να αγνοούν – ακούσια ή ηθελημένα- το μη ψηφιοποιημένο υλικό, γεγονός που τους οδηγεί σε ελλειπτικά ή εσφαλμένα συμπεράσματα.

Ένα από τα προβλήματα της ψηφιοποίησης το οποίο έγινε πολύ γρήγορα αντιληπτό ήταν η –ακούσια ή συνειδητή- απώλεια του υλικού. Ακούσια, γιατί υπάρχει υλικό χαμένο ή κατεστραμμένο το οποίο δεν ψηφιοποιήθηκε ή αρχειοθετήθηκε σωστά (ή και καθόλου). Καθώς αυτή η τεχνολογία εξελίσσεται ταχύτατα ακόμα και ψηφιοποιημένο υλικό που δεν ακολουθεί τις τεχνολογικές εξελίξεις μέσα σε σύντομο διάστημα γίνεται ανενεργό καθώς αποθηκεύεται σε μορφή που δε μπορεί να διαβαστεί από το εξελιγμένο εργαλείο. Η συνειδητή απώλεια υλικού προκύπτει από τις πρακτικές ανάγκες της εποχής. Ο μεγάλος όγκος πληροφοριών δημιουργεί προβλήματα χωρητικότητας και διαχείρισης, ακόμα και αν μιλάμε για διαδικτυακό χώρο και σύγχρονες βάσεις δεδομένων. Ζήτημα αποτελεί και η οικονομική επιβάρυνση από τη διαδικασία της διατήρησης του αρχείου. Οι συγκεκριμένοι οικονομικοί τους πόροι, θα οδηγήσουν αναπόφευκτα τους οργανισμούς στην διαδικασία επιλογής «τι πρέπει να ψηφιοποιηθεί και με ποια προτεραιότητα» με αποτέλεσμα μέρος του υλικού να μη ψηφιοποιηθεί ποτέ και ένα άλλο να παροπλιστεί εσκεμμένα καθώς δεν επικαιροποιήθηκε το μέσο ψηφιοποίησης ή το σύστημα αρχειοθέτησής του. Τα κριτήρια επιλογής για την ψηφιοποίηση του υλικού μπορεί να είναι και οικονομικά αλλά εξαρτάται επίσης και από πολλούς άλλους παράγοντες: η ζήτηση του κοινού (σχετίζεται με τα οικονομικά κριτήρια και την εμπορικότητα), τα προσωπικά ενδιαφέροντα αυτού που αναλαμβάνει την ψηφιοποίηση, τεχνικά ζητήματα και δυσκολίες (η ανάγκη για εξειδικευμένη εργασία απαιτεί χρόνο, χρήμα και την εύρεση του κατάλληλου προσωπικού) η κατάσταση του υλικού (αν “κινδυνεύει” η διάσωσή του μπορεί να κριθεί ως προτεραιότητα)¹³. Οι ίδιοι οι ερευνητές και οι χρήστες καλούνται πολλές φορές να επιλέξουν τι θέλουν να αποκτήσουν και πώς θα το χρησιμοποιήσουν. Ο όγκος του υλικού απαιτεί μια πρώτη αξιολόγησή του, πριν την όποια επεξεργασία. Για τον ιστορικό (και ειδικά τον μελετητή της αρχαίας ιστορία) το πρόβλημα εντείνεται καθώς ασχολείται με πολλά είδη πηγών, πολλές από τις οποίες δεν είναι αξιοποιήσιμες (Βόγλη 2015, 82-6; Raju et al. 2020, σ.σ. 11-12).

Η διαδικασία που περιγράψαμε σχετικά με την εργασία του ιστορικού πάνω στις πηγές εξακολουθεί να υφίσταται παρά τις τεχνολογικές εξελίξεις, με τη διαφορά ότι έχει μεταβληθεί αρκετά ο τρόπος που εκτελούνται. Η ψηφιακή επανάσταση που βιώνουμε προσφέρει μια σειρά από πλεονεκτήματα (Mehra & Kumar 2011, σ. 222; Χρυσούλας 2009, σ.σ. 61-2) τα οποία έχουν μεταβάλλει τον τρόπο διεξαγωγής της ιστορικής έρευνας. Δημιουργήθηκαν νέες ευκαιρίες για τους

¹³ Τα κριτήρια επιλογής για την ψηφιοποίηση συνήθως έχουν συνδυαστικές προεκτάσεις: η ζήτηση του κοινού σχετίζεται με την οικονομία, την εμπορικότητα και τους στόχους του φορέα. Τα προσωπικά ενδιαφέροντα που αναλαμβάνει την ψηφιοποίηση έχουν πιθανόν και οικονομικό στόχο, αλλά το κριτήριο είναι ιδιαίτερο και καλύπτει επίσης τις προσωπικές ερευνητικές ανάγκες ή επιδιώξεις. Τα τεχνικά ζητήματα και η αντιμετώπιση των δυσκολιών χρειάζονται εξειδικευμένη εργασία άρα επιβαρύνει οικονομικά τον φορέα, απαιτούν περισσότερο χρόνο, καθώς και την εύρεση του κατάλληλου προσωπικού. Τέλος, η κατάσταση του υλικού είναι πιθανό να επηρεάσει την ιεράρχηση των αναγκών.

ερευνητές, αλλά ταυτόχρονα τέθηκαν και σημαντικές προκλήσεις που καλούμαστε να αντιμετωπίσουμε, συχνά και με τη βοήθεια της ίδια της τεχνολογίας. Η ψηφιοποίηση είναι ένα σημαντικό στάδιο για την καταγραφή των πηγών και το οποίο προστέθηκε λόγω των τεχνολογικών εξελίξεων. Όμως το ψηφιοποιημένο υλικό δεν μπορούμε να το εκμεταλλευτούμε και να το αξιοποιήσουμε πλήρως αν δεν έχει προηγηθεί η οντολογική του προσέγγιση, δηλαδή η οργάνωση, η κατάλληλη ευρετηρίαση, η σωστή διαχείριση και η δημιουργία των μεταδεδομένων.

Η αλλαγή στη μέθοδο της έρευνας και στους τρόπους που ασκείται, δημιουργεί και μια αυξανόμενη εξάρτηση από την τεχνολογία και τα εργαλεία της. Οι συλλογές αρχείων ψηφιοποιούνται και εμπλουτίζονται συνεχώς και δημοσιεύονται στο διαδίκτυο, όπου “φιλοξενείται” επίσης μια πληθώρα δευτερεύουσας βιβλιογραφίας. Το διαδίκτυο και η ψηφιοποίηση έχουν καταστήσει τις πρωτογενείς και δευτερογενείς πηγές γνώσης προσβάσιμες και διαθέσιμες για χρήση στο ευρύ κοινό, ανεξάρτητα από την κατάσταση του φυσικού αντιτύπου. Πρόκειται για ένα “ιστορικό” άνοιγμα. Αυτή η εύκολη πρόσβαση είναι πολύτιμη για την ιστορική έρευνα και ωφελεί τόσο τους απλούς χρήστες του διαδικτύου, όσο και τους ιστορικούς. Όλοι εξοικονομούν χρόνο καθώς μπορούν να έχουν πρόσβαση στο επιθυμητό υλικό χωρίς να χρειαστεί να το αναζητήσουν ως φυσικό αντικείμενο. Η αναζήτησή και η καταγραφή σημειώσεων γίνονται οργανωμένα, ταχύτερα και ευκολότερα. Η διάθεση και η πρόσβαση συνδέονται με τη δυνατότητα και την ευκαιρία να έχει οποιοσδήποτε επιθυμεί περισσότερα δεδομένα από ποτέ. Καθένας μπορεί να δημιουργήσει ένα προσωπικό αρχείο, εξατομικευμένο στις ανάγκες του. Αυτό διευκολύνει τους ακαδημαϊκούς, αλλά είναι χρήσιμο εργαλείο και για μια νέα ομάδα ερευνητών η οποία μπορεί να διακατέχεται μόνο από το προσωπικό της ενδιαφέρον, όμως διαθέτει πλέον πληθώρα υλικού το οποίο ως τώρα ήταν απρόσιτο. Το διαδίκτυο διευκολύνει την επικοινωνία και τη συνεργασία όλων αυτών των ανθρώπων, ανεξαρτήτως ειδικότητας, θέσης ή τοποθεσίας (Chouchene 2019, σ.σ. 74-7).

Οι ενδείξεις είναι θετικές, όμως καμία επανάσταση δεν είναι αναίμακτη. Η μετάβαση στην ψηφιακή εποχή επιφυλάσσει και αρνητικές προεκτάσεις, πολλές από τις οποίες είχαν εντοπιστεί ήδη πριν από την εκτίναξή της και οφείλονται στα ίδια εργαλεία που προσφέρουν διευκολύνσεις. Συχνά είναι κεκαλυμμένες και παράλληλες με κάποια θεωρητικά θετική εξέλιξη. Καλούμαστε να τις ξεπεράσουμε προκειμένου να διαχειριστούμε καλύτερα την αλλαγή και να μπορέσουμε να εκμεταλλευτούμε πλήρως τις νέες δυνατότητες και τα εργαλεία που προσφέρονται.

Ένα από τα μεγαλύτερα πλεονεκτήματα της ψηφιακής εποχής θεωρείται το “άνοιγμα” της πρόσβασης στην πληροφόρηση. Πράγματι κερδίζεται πολύτιμος χρόνος και το ευρύ κοινό έρχεται σε επαφή με υλικό περισσότερο από ποτέ. Αλλά αυτή η δυνατότητα δεσμεύεται από δύο παράγοντες που τελικά δημιουργούν σημαντικές ανισότητες: την γνώση και την εξοικείωση με τα νέα ψηφιακά εργαλεία που δεν είναι βέβαιη για μεγάλο μέρος του πληθυσμού (Fridlund 2020, σ.σ. 73-4; Gibbs & Owens 2012) και περιορισμούς στην πρόσβαση που σχετίζονται με δικαιώματα ή οικονομικές συναλλαγές. Οι περιορισμοί λόγω πνευματικής ιδιοκτησίας και η ιδιωτική εκμετάλλευση του υλικού (οι εταιρίες και ιδιωτικοί οργανισμοί δεν υποχρεούνται να δημοσιεύουν τα δεδομένα τους, τα οποία τα αντιμετωπίζουν ως περιουσιακά στοιχεία) κρατάνε μεγάλο μέρος των πιθανόν διαθέσιμων δεδομένων κρυφά στο παρόν, ενώ δεν υπάρχει εγγύηση ότι θα γίνουν δημόσια μελλοντικά.

Εκτός από το ψηφιοποιημένο υλικό για το οποίο δεν δίνεται δυνατότητα πρόσβασης, υπάρχει και υλικό που δεν έχει ψηφιοποιηθεί. Συχνά οι ερευνητές αγνοούν ακούσια ή ηθελημένα αυτές τις δύο κατηγορίες τεκμηρίων. Η ύπαρξη του τεράστιου ποσότητας υλικού εντείνει το πρόβλημα καθώς δυσκολεύει την οργάνωσή του από τους ιστορικούς, με αποτέλεσμα στις αναζητήσεις τους να συλλέγουν τις πρώτες ανακτήσεις (οι οποίες στον Παγκόσμιο Ιστό γίνονται με τα κριτήρια επιλογής της μηχανής αναζήτησης ως καλύτερες απαντήσεις, κάτι που δεν ανταποκρίνεται πάντα στην αλήθεια). Και στην παραδοσιακή μέθοδο ο ερευνητής θα έπρεπε να διαβάσει ένα πολύ μεγάλο όγκο πληροφοριών πριν αποφασίσει τι θα συμπεριλάβει στην έρευνά του. Με τα ψηφιακά εργαλεία αυτή η δημιουργία ερευνητικών περιορισμών γίνεται αυτόματα, αλλά η πολύ συγκεκριμένη αναζήτηση, περιορίζει σημαντικά και την γενική επισκόπηση του ιστορικού για το θέμα. Όλα αυτά επηρεάζουν την ποιότητα και την αξιοπιστία των συμπερασμάτων τους, στοιχεία που για μια έρευνα είναι σημαντικότερα και από τα ίδια τα αποτελέσματα. Η ψηφιοποίηση, η ψηφιακή οργάνωση και η ευρεία διάθεση των δεδομένων είναι “επιλεκτικές” διαδικασίες με την έννοια πως εξαρτώνται από τις επιλογές αυτού που είναι επιφορτισμένος να τις εκτελέσει. Είτε πρόκειται για άνθρωπο είτε για οργανισμό, οι αποφάσεις βασίζονται σε μια σειρά από υποκειμενικά κριτήρια τα οποία καθορίζουν τις προτεραιότητες: το προσωπικό ενδιαφέρον του και το ιδιαίτερο περιεχόμενο των πηγών, τη ζήτηση από τους χρήστες (το κοινό στόχευσης), η πιθανή

εμπορική και οικονομική εκμετάλλευση του υλικού, το κόστος και οι τεχνικές δυσκολίες.

Όσον αφορά το αρχαιολογικό υλικό, το πρόβλημα είναι πως εξαιτίας της ψηφιοποίησης και της ανάγκης για την ψηφιακή εκμετάλλευσή του, είναι αναγκαστικά αποστασιοποιημένο από το εννοιολογικό του πλαίσιο (context). Αυτό είναι αποτέλεσμα ακόμα και της παραδοσιακής αντιγραφής σε βιβλίο ή την έκθεση αντικειμένων σε ένα μουσείο. Στην ψηφιακή μορφή τους όμως, ο μελετητής δεν έχει πρόσβαση και στα συμφραζόμενα και παρά την αυξημένη δυνατότητα επαφής, αγνοεί πολύ σημαντικά στοιχεία για την κατανόηση των πηγών του και τη διαδικασία της ιστορικής έρευνας συνολικά. Οι ψηφιακές μορφές των αρχείων έχουν περισσότερα προβλήματα από την απώλεια του πλαισίου τους. Η ίδια η ύπαρξη ψηφιοποιημένου υλικού υποβαθμίζει τη σημασία του αυθεντικού και την ανάγκη διατήρησής του. Ιδιαίτερα αν το φυσικό αντικείμενο έχει χαθεί τίθεται και ζήτημα αξιοπιστίας του ψηφιακού, επειδή αυτά τα αρχεία μπορεί να έχουν σφάλματα, να είναι αλλοιωμένα, παραποιημένα ή η ποιότητα ανάλυσής τους είναι κακή (είτε σκόπιμα είτε ακούσια τα σφάλματα, σε κάθε περίπτωση επηρεάζουν αρνητικά τη διαδικασία της ψηφιοποίησης). Όλα αυτά καθιστούν απαραίτητη την επιπλέον επεξεργασία των ψηφιοποιημένων αρχείων από άλλους ειδικούς. Σε συνδυασμό με την εξάρτηση από τα ψηφιακά εργαλεία και τις νέες πρακτικές υπάρχει ο κίνδυνος να χαθεί υλικό. Οτιδήποτε δεν ψηφιοποιήθηκε ή αρχειοθετήθηκε “σωστά”, αν πλέον δεν υπάρχει σε φυσική μορφή τότε χάνεται εντελώς. Κι αυτό γιατί τα ψηφιακά εργαλεία που επιτρέπουν την αποθήκευση και επεξεργασία των αρχείων αλλάζουν συνεχώς. Μεγάλο μέρος του υλικού μπορεί να μην είναι προσβάσιμο, είτε επειδή η πληροφορία του είναι καταγεγραμμένη σε κάποιο παλαιότερο τύπο επεξεργαστή κειμένου, είτε είναι αποθηκευμένο σε μη αναβαθμισμένη μορφή βάση δεδομένων και δεν μπορεί πλέον να διαβαστεί. Η συνεχόμενη πρόοδος στον τομέα αναγκάζει τους φορείς να προβαίνουν σε ενέργειες προκειμένου να συμβαδίζουν μαζί της και να αναβαθμίζουν τις υποδομές τους ώστε το υλικό να παραμένει ενεργό και αξιοποιήσιμο. Σε αντίθεση με την παραδοσιακή διαδικασία διατήρησης, για τα ψηφιακά αντικείμενα απαιτείται συνεχή αναβάθμιση των εργαλείων, των αρχείων και των βάσεων. Αυτό απαιτεί χρόνο, σημαντική δέσμευση οικονομικών πόρων και μεγάλη τεχνική εμπειρία. Η απαξίωση του πρωτότυπου υλικού αποτελεί απειλή και για τη διατήρηση του παραδοσιακού τρόπου με τον οποίο διεξαγόταν ως τώρα η έρευνα και η συνήθης επαφή με τις πρωτογενείς πηγές (Kuny 1997, σ.σ. 1-10; Chouchene 2019, σ.σ.74-9; Hitchcock & Shoemaker 2015, σ.σ. 78).

Ενότητα 2- ΠΡΑΚΤΙΚΟ ΜΕΡΟΣ

2.1. Δημιουργία Οντολογίας

α. Η Θεωρητική Διαδικασία Δημιουργίας μιας οντολογίας

Μια οντολογία αποτελεί το δομικό πλαίσιο αναπαράστασης για σύνολα εννοιών, ιδιοτήτων και των μεταξύ τους σχέσεων και χρησιμοποιείται για να περιγράψει ένα πεδίο γνώσης με συγκεκριμένο τρόπο, σκοπό και στόχους. Κάθε οντολογία καλύπτει ένα συγκεκριμένο γνωστικό πεδίο, παρέχοντας ένα ακριβές αλλά και ευέλικτο πλαίσιο οργάνωσης των εννοιών και των σχέσεων που το αποτελούν. Η δημιουργία μιας αποτελεσματικής οντολογίας απαιτεί προσεκτικό σχεδιασμό και οργάνωση και επηρεάζεται από πολλές παραμέτρους. Είναι ένα πλαίσιο σύνταξης, που καθορίζει το λεξιλόγιο, τους όρους, την ταξινόμηση των σχέσεων και τους περιορισμούς τους (Μαυρογιάννη 2007, σ.σ. 111-2). Ανάλογα τον σκοπό, την τυπικότητα, την χρήση της και τον τρόπο δομής και οργάνωσης των κλάσεων που την αποτελούν μια οντολογία μπορεί να υπάγεται σε διάφορους τύπους. Οι διαχωρισμοί δεν απασχολούν τους τελικούς χρήστες, αλλά η μορφή της προαποφασίζεται από τον κατασκευαστή της οντολογίας και επηρεάζει τη μέθοδο εργασίας του και τις απαιτήσεις που θα έχει η εφαρμογή από αυτόν¹⁴.

Πυρήνας της οντολογίας και κύριο στοιχείο στην αναπαράσταση της γνώσης είναι οι κλάσεις, έννοιες οι οποίες περιγράφουν σύνολα που μοιράζονται κοινά χαρακτηριστικά. Κάθε κλάση αποτελείται από τα στιγμιότυπα της, οντότητες που αντιπροσωπεύουν συγκεκριμένα παραδείγματα της και φέρουν μια κοινή χαρακτηριστική ιδιότητα. Μια κλάση μπορεί να είναι μια συλλογή από στιγμιότυπα, είτε μια έννοια με ιδιότητες, είτε – συνηθέστερα- και τα δύο (δηλαδή και στιγμιότυπα και ιδιότητες με τις οποίες ενώνονται με τις άλλες κλάσεις). Κάθε κλάση μπορεί να “ανήκει” σε μια μεγαλύτερη νοηματικά ομάδα οντοτήτων που ονομάζεται υπερκλάση. Ταυτόχρονα μπορεί να έχει και η ίδια κάποιες άλλες κλάσεις, τις υποκλάσεις. Σε μια πρώτη φάση πρόκειται για μια ιεραρχικού τύπου ταξινόμηση, όπου οι κλάσεις τοποθετούνται σε επίπεδα ανάλογα με τη γενικότητά τους, από το γενικό στο πιο ειδικό. Η αναζήτηση και η καταγραφή των κλάσεων είναι ένα σημαντικό βήμα για τη διαμόρφωση μιας οντολογίας καθώς πρέπει να είναι τέτοιες που να αποτυπώνουν το σκοπό, τους στόχους και το πεδίο εφαρμογής της. Τα στιγμιότυπα κάθε κλάσης έχουν ιδιότητες, πράξεις ή έννοιες που τους προσδίδουν κάποιο χαρακτηριστικό. Μέσω αυτών ενώνονται με άλλες κλάσεις (λ.χ. η ιδιότητα “έχει Συγγραφέα” συνδέει την κλάση “Βιβλίο” με την κλάση “Συγγραφέας”- οπότε κάθε στιγμιότυπο της πρώτης πρέπει να διαθέτει σύνδεση με ένα στιγμιότυπο της δεύτερης. Οι ιδιότητες μπορούν επίσης να προσδίδουν σε κάθε στιγμιότυπο κάποιο μοναδικό χαρακτηριστικό, μέσα από το οποίο ξεχωρίζει από τα υπόλοιπα στιγμιότυπα.

Με τις κλάσεις και τις ιδιότητες συνδέονται τα αξιώματα, λογικές που καθορίζουν τη σημασιολογία (ερμηνεία) και τους περιορισμούς των όρων με σκοπό συγκεκριμένα επιθυμητά αποτελέσματα. Ανάλογα το σκοπό για τον οποίο θέτονται, τα αξιώματα χωρίζονται σε κατηγορίες: σε αυτά που παράγουν νέα γνώση και αυτά που επιβάλλουν περιορισμούς, είτε στα δεδομένα που θα ανακτηθούν από ερωτήματα στην οντολογία, είτε στη μορφή των οντοτήτων, είτε στα χαρακτηριστικά των κλάσεων και των ιδιοτήτων. Παρόμοια σημασία και χρήση με τα αξιώματα έχουν για την οντολογία οι περιορισμοί, οι οποίοι είναι κανόνες που εξασφαλίζουν την ομαλή λειτουργία της οντολογίας και τις σωστές συνδέσεις των κλάσεων και των ιδιοτήτων.

Μια οντολογία είναι το πλαίσιο στο οποίο οργανώνονται οι έννοιες και αποτυπώνονται οι μεταξύ τους σχέσεις εντός ενός συγκεκριμένου γνωσιακού συνόλου. Αν και η ιεραρχία “δεντρικού” τύπου (Μαυρογιάννη 2007, σ.σ. 104-6) είναι βασικός, απλός και συνηθισμένος τρόπος οργάνωσης στοιχείων, δεν επαρκεί για να περιγράψει τις πολύπλοκες σχέσεις των εννοιών. Ενώ είναι η πιο απλή και εύκολη αποτύπωση οργάνωσης σύμφωνα με την ανθρώπινη λογική, είναι χρήσιμη μόνο για τις απλές κατηγοριοποιήσεις εξειδίκευσης (από τη γενικότερη ομάδα στην ειδικότερη). Σε αυτή την περίπτωση ιεραρχίας, κάθε κλάση αποτελεί υποκλάση της προηγούμενης και υπερκλάση της επόμενης, δημιουργώντας μια δομή όπου η μια έννοια οδηγεί στην επόμενη ή αποτελεί μέρος της προηγούμενης, “κληρονομώντας” τα χαρακτηριστικά της, τα οποία θα

¹⁴ Για τα χαρακτηριστικά τους, βλ. στο παρόν, κεφάλαιο 1.1.β. Η συνηθέστερη μορφή είναι η Ημιτυπική γιατί συνδυάζει και στοιχεία των άλλων. Η επιλογή εξαρτάται από την οικειότητα του κατασκευαστή με την δημιουργία οντολογιών και τις γλώσσες προγραμματισμού. Αντίστοιχα η Άτυπη οντολογία δεν είναι χρήσιμη σε υπολογιστικά συστήματα, επομένως χρησιμοποιείται στην ουσία στην αρχή της διαδικασίας βοηθητικά (ωφελεί κυρίως λιγότερο έμπειρους χρήστες). Η εφαρμογή που κατασκευάστηκε για την παρούσα Διπλωματική είναι ένας συνδυασμός ημιτυπικής και αυστηρά τυπικής μορφής.

πρέπει να μοιράζεται με τα υπόλοιπα μέλη αυτής της κλάσης. Ένα τέτοιο μοντέλο δεν είναι δυνατό να περιγράψει όλες τις σχέσεις και τις ιδιαιτερότητες των εννοιών που βρίσκονται μέσα στα πολύπλοκα γνωστικά πλαίσια, μπορούν να αποτυπώνονται όμως σε μια οντολογία.

Το ζητούμενο είναι η δημιουργία ενός προτύπου που να μπορεί να αποτυπώνει τη σημασιολογική ερμηνεία όλων των εννοιών με τρόπο που να αντανακλά την πολυπλοκότητα της ανθρώπινης σκέψης. Σχόλιο για την σημασιολογική ερμηνεία πραγματοποιήθηκε αναφορικά με τους άξονες ανάπτυξης του Σημασιολογικού Ιστού. Τα ίδια ισχύουν και εδώ, εφόσον οι οντολογίες αποτελούν κομμάτι του. Είναι αναγκαίο ο υπολογιστής μέσω της γλώσσας αναπαράστασης, να μπορεί αρχικά να αναπαριστά και έπειτα να κατανοεί την πληροφορία με παρόμοιο τρόπο με τον άνθρωπο, αποτυπώνοντας τα δικά του εννοιολογικά πρότυπα (Μαυρογιάννη 2007, σ.σ. 100-3).

Οι οντολογίες προσφέρουν αυτή τη λύση. Παρότι βασίζονται στην αρχή τους σε μια ιεραρχική ταξινόμηση επιπέδων και οι κλάσεις-οντότητες ανήκουν σε μια μεγαλύτερη νοηματικά ομάδα και να περιέχουν μικρότερες, ταυτόχρονα μπορούν να ενσωματώνουν πολλές σχέσεις σε οποιοδήποτε στιγμιότυπο της κάθε κλάσης ανεξάρτητα από την κατάσταση των υπόλοιπων¹⁵, όπως συμβαίνει στην ιεραρχική οργάνωση (που βασίζεται στη λογική γονέας-παιδί). Άλλωστε η κατανόηση της σημασιολογικής ερμηνείας από τους υπολογιστές είναι ένας από τους στόχους της ανάπτυξης του Σημασιολογικού Ιστού. Δεν δομούνται δεντρικά, αλλά με περίπλοκους γράφους που βασίζονται σε τριπλέτες (διακλαδώσεις). Κάτι τέτοιο παρέχει τη δυνατότητα σχεδιασμού και αναπαράστασης ιδιαίτερα πολύπλοκων συστημάτων για κάθε αντικείμενο.

Η κατασκευή μιας οντολογίας είναι μια πολύπλοκη διαδικασία. Πρόκειται για ένα ζήτημα που απασχόλησε από την αρχή της διαφαινόμενης ανάγκης εξέλιξης προς τον Σημασιολογικό Ιστό και συνεχίζει να αποτελεί αντικείμενο προτάσεων, χωρίς να έχει προσδιοριστεί με ακρίβεια. Περιλαμβάνει μια σειρά από βήματα από τη σύλληψη της ιδέας μέχρι και την τελική κατασκευή της σε ένα λογισμικό ή πρόγραμμα ανοιχτού κώδικα. Συνυπολογίζοντας όλες τις απόψεις, μπορούμε να καθορίσουμε τα βήματα με την παρακάτω ακολουθία:

1. Προσδιορισμός του θέματος, του πεδίου εφαρμογής και της σκοπιμότητας της οντολογίας
2. Προσδιορισμός των βασικών όρων και των εννοιών
3. Ιεραρχική ταξινόμηση των όρων
4. Καθορισμός των περιορισμών
5. Κωδικοποίηση ή Αποτύπωση σε πρόγραμμα αναπαράστασης
6. Ενοποίηση υπαρχουσών οντολογιών
7. Αξιολόγηση και τεκμηρίωση¹⁶

Ο χρόνος που απαιτείται για την πραγματοποίηση κάθε βήματος, καθώς και ο τρόπος προσέγγισής του, δεν είναι δυνατόν να υπολογιστούν εκ των προτέρων. Κι αυτό γιατί η ανάλυση και οι επιμέρους διαδικασίες στην περάτωσή τους εξαρτώνται και επηρεάζονται κάθε φορά από αστάθμητους παράγοντες, όπως είναι η οικειότητα του δημιουργού της οντολογίας με τη διαδικασία (είτε με γραφή κώδικα είτε με εξοικείωση στη χρήση προγραμμάτων ανοιχτού κώδικα, όπως το Protégé που χρησιμοποιήθηκε στην παρούσα εργασία- η προσέγγιση και η προεργασία για τα βήματα πάντως είναι περίπου ίδια ανεξάρτητα του τρόπου κατασκευής της οντολογίας) και η εμπειρία στη χρήση τους. Τέλος, θα πρέπει να έχουμε υπόψη ότι μιλώντας για βήματα δεν εννοούμε μια ακολουθία πράξεων από την αρχή και μέχρι την ολοκλήρωση της διαδικασίας, αλλά περισσότερο για μια λίστα προϋποθέσεων που λαμβάνονται υπόψη, εφόσον πολλά από αυτά τα βήματα είναι αυτόματοι, ταυτόχρονοι υπολογισμοί του ανθρώπινου μυαλού και δεν είναι δυνατόν να υπολογίζονται ξεχωριστά.

1. Προσδιορισμός του θέματος, του πεδίου εφαρμογής και της σκοπιμότητας:

Μια οντολογία είναι ένα μοντέλο παρουσίασης των εννοιών και των σχέσεων ενός γνωστικού τομέα και έχει συγκεκριμένο κοινό στο οποίο (θα) απευθύνεται, φτιάχνεται για κάποιον ιδιαίτερο σκοπό και έχει σαφείς στόχους. Όλα αυτά τα ζητούμενα θα πρέπει να καθορίζονται από το δημιουργό της οντολογίας πριν ξεκινήσει να συγκεντρώνει με ακρίβεια τις έννοιες και τις σχέσεις που θέλει να τις διέπουν. Ο εντοπισμός αυτών των στοιχείων γίνεται όταν ο δημιουργός της

¹⁵ Για παράδειγμα μια κλάση ή ένα στιγμιότυπό της μπορεί να συνδέεται μέσω διάφορων ιδιοτήτων με πολλές άλλες κλάσεις ή να “μοιράζεται” μια ιδιότητα με μια άλλη κλάση ή στιγμιότυπό της χωρίς να έχουν κάποιο άλλο κοινό στοιχείο στην οργανωτική δομή της οντολογίας.

¹⁶ Για την περιγραφή των βημάτων για την κατασκευή μιας οντολογίας ή κάποια διαδικασία που περιλαμβάνεται σε αυτά, βλ., Falbo et al. 1998, σ. 352; Ιακωβάκης 2019, σ. 4; Μαυρογιάννη 2007, σ.σ. 104-6, 115-8; Σαραντοπούλου 2014, σ.σ. 83-4; Uschold & Gruninger 1996, σ.σ. 93-155; Uschold & King 1995, σ.σ. 3-13; Χριστίδου 2005, σ.σ. 71-5.

οντολογίας θέσει στον εαυτό του μια σειρά από ερωτήματα: ποιος είναι ο γενικός τομέας στον οποίο θα κινείται η οντολογία, που θα χρησιμοποιηθεί, από ποιους (ή και πόσους) και για ποιο σκοπό; Ο σκοπός ανταποκρίνεται στην αναγκαιότητα των προβλεπόμενων χρήσεων, η οποία αποκαλύπτεται από τον υπολογισμό των ερωτημάτων στα οποία η οντολογία θα μπορεί να “απαντά” ή από τις εργασίες που θα προσφέρει στο χρήστη. Η ποσότητα και η ιδιότητα των χρηστών μιας οντολογίας εκτός από το πεδίο και το σκοπό και το στόχο της υπολογίζεται και από την επαναχρησιμοποίησή της διαδικασίας χρήσης μιας ήδη υπάρχουσας οντολογίας (ή κάποιου μέρους της) σε μια νέα εφαρμογή, πλαίσιο ή άλλη, παρόμοια οντολογία. Λέγοντας το που θα χρησιμοποιηθεί, εννοούμε σε τι εργαλείο ή λογισμικό θα ενσωματωθεί και ως τι θα λειτουργεί: πλαίσιο οργάνωσης μια ιστοσελίδας εμπορικού περιεχομένου, βάση ανάλυσης δεδομένων, καταλογογράφηση βιβλίων κ.α.. Οι απαντήσεις στα παραπάνω ερωτήματα ήδη από την αρχή επιτρέπουν τον ακριβέστερο σχεδιασμό της οντολογίας και θα καθορίσουν τον ορισμό των εννοιών που ακολουθεί (Falbo et al. 1998, σ. 352).

2. Προσδιορισμός των βασικών όρων και των εννοιών (κλάσεις και ιδιότητες).

Όταν καθοριστούν στο μυαλό του δημιουργού οι απαντήσεις στα ερωτήματα που τέθηκαν παραπάνω, επόμενο βήμα είναι η καταγραφή των κλάσεων και των ιδιοτήτων τους. Οι κλάσεις αποτελούν το κυριότερο στοιχείο μιας οντολογίας καθώς αποτελούν τα σύνολα που περιγράφουν τις βασικές έννοιες του γνωστικού πεδίου της οντολογίας. Επομένως, πρέπει να είναι ακριβείς, ευέλικτες και στη σωστή ποσότητα ώστε να καλύπτουν τους στόχους που τέθηκαν και να ανταποκρίνονται στην ποιότητα της ανακτημένης γνώσης (Μαυρογιάννη 2007, σ.σ. 103, 109, 111-2, 115-8; Uschold & Gruninger 1996, σ.σ. 93-155). Το ίδιο οι σχέσεις αλληλεπίδρασης και οι περιορισμοί τους (ιδιότητες και αξιώματα), πρέπει να είναι από την αρχή αντικείμενο απασχόλησης, γιατί βοηθούν τις κλάσεις να αποτυπώσουν το επιθυμητό αποτέλεσμα και να καταγράφουν όλα τα στοιχεία που κρίνονται απαραίτητα για να αποτυπώσουν την οπτική του θέματος. Επομένως, αντίστροφα, μια κλάση μπορεί να γίνει αντιληπτή ως απάντηση σε μια τέτοια σύνδεση που κρίνεται απαραίτητη. Σε αυτή τη φάση της δημιουργίας αρκεί να καταγράψει τους όρους που αφορούν την οντολογία του, σε άμεση εξάρτηση με τα ερωτήματα του πρώτου βήματος και σε κάθε περίπτωση πρέπει να (μπορούν να) προστίθενται καθώς προχωράει η κατασκευή και παρουσιάζονται ανάγκες για την κάλυψη πεδίων που δεν είχε ο δημιουργός προσμετρήσει αρχικά.

3. Ιεραρχική ταξινόμηση των όρων

Στο προηγούμενο μέρος είπαμε πως σε πρώτο στάδιο αρκεί απλά να καταγραφούν οι όροι. Σε αυτό το βήμα δημιουργείται το οντολογικό πλαίσιο και οι όροι που επιλέχθηκαν προηγουμένως οργανώνονται καταρχήν ιεραρχικά. Δηλαδή οι όροι που υπολογίστηκε να αποτελέσουν τις κλάσεις τοποθετούνται σε μια ιεραρχία κλάσεων και υποκλάσεων με τη διαβάθμιση να γίνεται από το γενικότερο στο ειδικότερο. Η ιεραρχική ταξινόμηση διασφαλίζει την σωστή οργάνωση των όρων, αποσαφηνίζει τις μεταξύ τους σχέσεις και ορίζει τη θέση κάθε έννοιας εντός του οντολογικού πλαισίου. Η διαδικασία είναι σημαντική για την εύρυθμη λειτουργία της οντολογίας. Συνεπώς η επιλογή των όρων και των λέξεων που θα αποτελέσουν τις κλάσεις πρέπει να γίνεται με μεγάλη προσοχή για να ανταποκρίνονται όχι μόνο στα ερωτήματα που αναφέρθηκαν, αλλά και στην ιεραρχική ταξινόμηση γονεϊκού τύπου (τα στιγμιότυπα κάθε υποκλάσης μοιράζονται κοινά χαρακτηριστικά μεταξύ τους, αλλά και με την κλάση στην οποία υπάγεται η υποκλάση τους).

Η ιεραρχική ταξινόμηση των όρων σε μια οντολογία αφορά την οργάνωση των κλάσεων. Αυτό όμως το βήμα περιλαμβάνει επίσης και τον υπολογισμό της συσχέτισης των ιδιοτήτων με τις κλάσεις τους. Είναι μια διαδικασία που νοητικά έτσι κ αλλιώς γίνεται παράλληλα. Εφόσον δημιουργούμε ένα πλαίσιο, είναι λογικό να ορίζουμε ταυτόχρονα την ιεράρχηση των κλάσεων με τις συσχετίσεις που θα έχουν. Οι ιδιότητες δεν υπάγονται σε παρόμοια με τις κλάσεις ιεραρχική δομή, δηλαδή δε ταξινομούνται με σχέσεις διαβάθμισης ή ιεραρχίας είτε μεταξύ τους είτε ως προς τις κλάσεις. Δεν υπάρχουν κάποιες ανώτερες ή κατώτερες, απλώς ιδιότητες που δίνουν πολλές και διαφορετικές πληροφορίες για κάθε κλάση. Επίσης οι ιδιότητες δε “κληρονομούν” τα χαρακτηριστικά τους όπως συμβαίνει στις ιεραρχημένες κλάσεις. Αντίθετα, τοποθετούνται δεντρικά όπου μια κλάση μπορεί να χαρακτηρίζεται από πολλές ιδιότητες, οι οποίες δε σχετίζονται άμεσα μεταξύ τους, αλλά συνδέουν και κλάσεις (έννοιες) μεταξύ τους ορίζοντας τις σχέσεις που τις διέπουν.

4. Καθορισμός των περιορισμών

Ως επόμενο στάδιο, το οποίο συχνά γίνεται ταυτόχρονα με το προηγούμενο, μπορούμε να θεωρήσουμε τον “καθορισμό των περιορισμών”. Οι περιορισμοί είναι οι κανόνες και οι

συνθήκες που επιβάλλουν το πως οι κλάσεις, τα στιγμιότυπα τους και οι ιδιότητές μπορούν να συνδυαστούν και να αλληλοεπιδρούν. Διαμορφώνουν μαζί με τις κλάσεις και τις ιδιότητες τη μορφή των αποτελεσμάτων που θα ανακτώνται (reasoning- υποστήριξη λογικών συμπερασμάτων¹⁷) και εξασφαλίζουν τη συνέπεια, την ορθότητα και την εγκυρότητα της οντολογίας, αποτρέποντας ασάφειες και αντιφάσεις¹⁸. Αυτό το βήμα αποτελεί μια αποτύπωση των απαιτήσεων της οντολογίας, προκειμένου να υποστηρίζει αποτελεσματικά το συγκεκριμένο ζήτημα για το οποίο σχεδιάζεται. Διαμορφώνουν το πλαίσιο μέσα στο οποίο οργανώνεται η πληροφορία με τρόπο που εξισορροπεί την ανθρώπινη λογική και την κατανόηση των υπολογιστών. Οι περιορισμοί καθορίζουν το εύρος των τιμών ή τη μοναδικότητά τους, το είδος των τιμών των ιδιοτήτων (αριθμοί, σύμβολα κ.α.), το πλήθος των ιδιοτήτων, την σύνθεση των κλάσεων από μια ιδιότητα (Domain-Range). Μέσω των περιορισμών διαπιστώνεται αν τελικά πληρούνται οι απαιτήσεις κ προκύπτουν τα επιθυμητά αποτελέσματα που τέθηκαν κατά τη σύλληψη της ιδέας της οντολογίας και μπορούν να εξαχθούν χρήσιμα και αξιόπιστα συμπεράσματα.

5. Κωδικοποίηση ή Αποτύπωση σε πρόγραμμα αναπαράστασης

Όλα τα προηγούμενα βήματα εννοιολογικού σχεδιασμού που περιγράψαμε, παρά την κρισιμότητά τους και την προσοχή που χρειάζονται, είναι αρκετά “θεωρητικά”, εφόσον μπορούν να πραγματοποιηθούν με μια πρόχειρη καταγραφή (σε χαρτί ή σε κάποιον φυλλομετρητή) ή ακόμη και εντελώς αφηρημένα και σχηματικά στο μυαλό του σχεδιαστή της οντολογίας. Σε αυτό το στάδιο ξεκινά ουσιαστικά η κατασκευή της οντολογίας καθώς μετατρέπεται η θεωρητική σχεδίαση σε σημασιολογικό εργαλείο που θα ενσωματωθεί σε κάποιο λογισμικό ή πρόγραμμα. Η καταγραφή και των όρων και των εννοιών που αποφασίστηκε μπορεί να γίνει είτε με τη γραφή κώδικα σε εργαλεία χρησιμοποιώντας γλώσσα προγραμματισμού (λ.χ. RDF ή OWL), είτε με την οπτική αναπαράστασή τους με τη βοήθεια προγραμμάτων ανοιχτού κώδικα. Η δεύτερη περίπτωση δε προϋποθέτει βαθιά γνώση γλωσσών προγραμματισμού, σίγουρα όμως είναι αναγκαία η μερική επαφή με το αντικείμενο, προκειμένου η οντολογία να γίνει άρτια και λειτουργική. Από τη στιγμή που για την υλοποίηση της παρούσας εργασίας χρησιμοποιήθηκε το λογισμικό ανοιχτού κώδικα Protégé, στο εξής θα αναφερόμαστε σε αυτό τον τρόπο. Πάντως όποιος τρόπος εργασίας και να επιλεγεί, η διαδικασία κατασκευής μιας οντολογίας περιλαμβάνει τα βήματα που αναφέρθηκαν.

Οι όροι και οι έννοιες στα προηγούμενα βήματα καταγράφηκαν σε κάποια φυσική γλώσσα. Σε αυτό το στάδιο καλούμαστε να αποτυπώσουμε το θεωρητικό μας σχέδιο σε ένα εργαλείο, με τη βοήθεια της κωδικοποίησης. Τη “μεταφορά” και τη ρητή αναπαράσταση των δεδομένων που θέλουμε να περιλαμβάνονται στην οντολογία (κλάσεις, ιδιότητες, περιορισμοί) σε μορφή κατανοητή από τον υπολογιστή (σε κάποια “επίσημη” γλώσσα). Όλα τα στοιχεία τοποθετούνται και οργανώνονται με τρόπο που μπορούν να επεξεργάζονται από τους υπολογιστές και να χρησιμοποιούνται για την εξαγωγή συμπερασμάτων και αναλύσεων. Γι’ αυτό και πολλές φορές μπορεί να εντοπιστούν σφάλματα στη διατύπωση της αναπαράστασης, ασάφειες ή ελλείψεις που χρήζουν συμπλήρωσης. Σε εργαλεία ανοιχτού κώδικα, προστίθενται οι κλάσεις, ορίζεται η ιεραρχία (αν και ποιες από αυτές θα αποτελούν υποκλάσεις ποιων κλάσεων) και καταγράφονται τα στιγμιότυπα τους. Έπειτα καταγράφονται οι ιδιότητες και καθορίζονται οι περιορισμοί οι τιμές τους (εφόσον είναι ιδιότητες πληροφορίας και όχι ενοποιητικές). Έπειτα, με τις ιδιότητες, ενώνονται τα στιγμιότυπα των διάφορων κλάσεων.

6. Ενοποίηση υπαρχουσών οντολογιών

Ένας από τους στόχους της ανάπτυξης των οντολογιών στο πλαίσιο του Σημασιολογικού Ιστού είναι η επαναχρησιμοποίησή τους, η οποία επιτρέπει την αξιοποίηση της προηγούμενης γνώσης και έχει ως αποτέλεσμα τη διασύνδεση των διάφορων οντολογιών οι οποίες θεματικά εντάσσονται στα ίδια πεδία (ή υπάρχουν κάποιοι όροι που μπορούν να χρησιμοποιηθούν εξίσου).

¹⁷ Για παράδειγμα αν οριστεί πως ένα “Βιβλίο” πρέπει να έχει ένα “Συγγραφέα”, τότε όλα τα στιγμιότυπα της κλάσης “Βιβλίο” θα πρέπει να συνδέονται με ένα στιγμιότυπο της κλάσης “Συγγραφέας”. Το σύστημα μπορεί να αναγνωρίσει και να επισημάνει αν για κάποιο “Βιβλίο” δεν οριστεί “Συγγραφέας” ή δεν θα επιτρέψει ένα “Βιβλίο” να ενωθεί με πάνω από ένα “Συγγραφέα”.

¹⁸ Αυτές προκύπτουν είτε από “κατασκευαστικό” λάθος της οντολογίας, είτε από “λογικό” πρόβλημα στην οργάνωσή της. Οι υπολογιστές, προς το παρόν, δεν μπορούν να ελέγχουν την εγκυρότητα των δεδομένων που καταγράφουν οι χρήστες. Αν μια δήλωση είναι λάθος -για παράδειγμα ότι ο Χ έγραψε το βιβλίο Ψ, ενώ γνωρίζουμε ότι αυτό δεν ισχύει, αλλά είναι διατυπωμένη σωστά εντός της οντολογίας, το πρόγραμμα δε θα αναγνωρίσει το λάθος. Οι περιορισμοί ελέγχουν την εγκυρότητα της σχέσης των δεδομένων με τους προκαθορισμένους κανόνες και τη λογική ακολουθία. Αν δεν έχουμε ορίσει ότι μια ιδιότητα συνδέει δύο κλάσεις, αυτόματα δενμπας επιτρέπεται να τη χρησιμοποιήσουμε για να ενώσουμε κάποια στιγμιότυπά τους.

Γι' αυτό και η χρήση πολύ συγκεκριμένων όρων στην οντολογία περιορίζει την επαναχρησιμοποίηση και έπειτα την ανταλλαγή πληροφοριών με άλλα συστήματα. Καλύτερη πρακτική, συνήθως, είναι να δημιουργούνται πιο γενικές οντολογίες που να καλύπτουν ευρύτερα πεδία. Αυτές, θα έχουν και μεγαλύτερες δυνατότητες επαναχρησιμοποίησης για σε διάφορα γνωστικά πεδία και μπορούν εύκολα να επεκτείνονται ή να προσαρμόζονται οι όροι τους σε άλλες οντολογίες.

Η επαναχρησιμοποίηση επιτυγχάνεται μέσω της ενσωμάτωσης, τη διαδικασία κατά την οποία (επιανα)χρησιμοποιούνται ή επεκτείνονται στοιχεία άλλων, επίσημων οντολογιών στην οντολογία που κατασκευάζεται. Η νέα οντολογία ωφελείται πολλαπλά, αρχικά “κερδίζοντας” για τον κώδικά της αξιοπιστία. Αυτή προκύπτει με δύο μορφές: στο επίπεδο της μείωσης της πιθανότητας λάθους (καθώς χρησιμοποιούνται έτοιμες και καλά ορισμένες έννοιες και σχέσεις) και της απόδειξης ότι ο κατασκευαστής της οντολογίας μελέτησε καλά το θέμα του. Έτσι, συμβάλει στην μείωση του όγκου των πληροφοριών και την οργάνωσή τους, στο Δίκτυο¹⁹ (Falbo et al. 1998, σ. 353).

Στο συγκεκριμένο βήμα, η διαδικασία ενσωμάτωσης παρουσιάζει διαφορές ανάλογα με το μέσο κατασκευής της οντολογίας. Αν γίνεται γράφοντας κώδικα, η ενοποίηση προηγείται της κωδικοποίησης. Σε αυτό το σενάριο είναι αδύνατο να επανελέγχεται συνεχώς όλη η οντολογία και να γίνονται αλλαγές καθώς γράφεται. Ο κίνδυνος σφαλμάτων στην οντολογική λογική είναι μεγάλος. Αντίθετα όταν η οντολογία κατασκευάζεται σε κάποιο πρόγραμμα ανοιχτού κώδικα, η ενοποίηση μπορεί να γίνεται μετά την καταγραφή σε πρόγραμμα αναπαράστασης. Ειδικά αν ο κατασκευαστής δεν είναι έμπειρος με τη διαδικασία ή με τον προγραμματισμό, είναι ευκολότερο να καταγράψει τις έννοιες σε μια φυσική γλώσσα και μετά να αντικαταστήσει τις έννοιες που βρίσκει.

Σε αυτή την περίπτωση η διαδικασία περιλαμβάνει μια σειρά από ενέργειες, αφού πρώτα έχει καταγραφεί η δομή της οντολογίας σε φυσική γλώσσα: ο κατασκευαστής πρώτα πρέπει να αναζητήσει τους υπάρχοντες όρους που σχετίζονται με τις έννοιες που επιθυμεί να περιλαμβάνει η οντολογία του. Η αναζήτηση μπορεί να γίνεται σε επιμέρους βάσεις δεδομένων και αποθετήρια οντολογιών (DBpedia, Wikidata, Linked Open Vocabularies κ.α.). Σε αυτές υπάρχουν πολλοί όροι που θα φαίνονται αντίστοιχοι, αλλά με την ανάγνωση των στοιχείων τους προκύπτουν μικροδιαφορές. Πρέπει να χρησιμοποιούνται μόνο όροι που ταιριάζουν απόλυτα με το σκοπό της οντολογίας που φτιάχνεται. Για παράδειγμα, στην οντολογία που θα περιγράψουμε στη συνέχεια, χρησιμοποιήσαμε την κλάση “Language”, η αντιστοιχία της οποίας βρέθηκε στην αναγνωρισμένη οντολογία της BIBFRAME. Επομένως αντικαταστάθηκε το URI της κλάσης της οντολογίας με το URI της κλάσης από την BIBFRAME (“http://id.loc.gov/ontologies/bibframe/Language”) και πλέον το όνομα της κλάσης δεν εμφανίζεται ως “Language” αλλά ως “bibframe:language”. Με αυτόν τον τρόπο, η έννοια και η σημασία της άλλαξε από μια φυσική γλώσσα (που καταλαβαίνει ο χρήστης) σε μια μορφή που καταλαβαίνει ο υπολογιστής και μάλιστα μπορεί να χρησιμοποιηθεί και από άλλες οντολογίες. Η ενοποίηση αφορά όλους τους όρους της οντολογίας (κλάσεις και ιδιότητες) και το πιθανότερο είναι για τις περισσότερες να βρεθεί μια έννοια που να πληροί τις προϋποθέσεις για να συνδεθεί στη νέα οντολογία.

7. Αξιολόγηση και τεκμηρίωση

Με την ολοκλήρωση του προηγούμενου βήματος η κατασκευή έχει ουσιαστικά ολοκληρωθεί. Όμως της τελικής παρουσίασης και χρήσης της οντολογίας προηγείται η αξιολόγηση και η τεκμηρίωση της, δυο διαδικασίες σημαντικές για τη λειτουργία και την αξιοποίησή της. Η αξιολόγηση είναι ο τελικός έλεγχος ο οποίος απαιτείται για να διασφαλιστεί η λειτουργικότητα και η αξιοπιστία της οντολογίας, ενώ η τεκμηρίωση αποτελεί έναν οδηγό παρουσίασης και χρήσης της.

Η αξιολόγηση είναι απαραίτητη για να αποδειχθεί πως η οντολογία που κατασκευάστηκε δεν είναι ένα θεωρητικό μοντέλο, αλλά ένα πρακτικά εφαρμόσιμο πρότυπο που θα μπορεί να αξιοποιηθεί. Αρχικά πραγματοποιείται έλεγχος για να επιβεβαιωθεί ότι οι όροι που επιλέχθηκαν και η σχεδίαση της οντολογίας συνολικά ανταποκρίνονται στους στόχους και το σκοπό που ορίστηκαν στο ξεκίνημα και είναι ακριβείς. Έπειτα εξετάζεται πόσο ευέλικτη (να μπορεί να επαναχρησιμοποιηθεί εύκολα) και αποδοτικά κατασκευασμένη είναι. Στη συνέχεια ελέγχεται για λογικά και συντακτικά σφάλματα που θα επηρεάσουν τη λειτουργικότητά της, λ.χ. σε κάποια σύνδεση κλάσης-ιδιότητας, σε καθορισμένο περιορισμό κ.α. ή λάθος στις πληροφορίες που εισήγαγε ο κατασκευαστής.

¹⁹ Πρόκειται για την αποφυγή της διπλοκαταγραφής ή επανεφεύρεσης όρων. Οι έννοιες που αποτελούν τις κλάσεις και τις ιδιότητες αντί να δημιουργούνται και να προστίθενται ως νέες οντότητες, αντικαθίστανται στην νέα οντολογία με ίδιες ή παρόμοιες που υπάρχουν ήδη.

Την αξιολόγηση μιας οντολογίας διενεργεί καταρχήν ο κατασκευαστής της. Είναι αυτός που πρώτος μπορεί να εντοπίσει σφάλματα, ειδικά αναφορικά με την πληρότητά της και τα δεδομένα της. Εκτός από τον προσωπικό του έλεγχο, τα λογισμικά ανοιχτού κώδικα προσφέρουν τρόπους ελέγχου της οντολογίας: με την επέκταση DL Query τίθενται ερωτήματα μέσω των οποίων αποδεικνύεται η λειτουργικότητα και η αξιοπιστία της οντολογίας. Αν η εκτέλεση των ερωτημάτων δεν παράγει τα αναμενόμενα αποτελέσματα με τον επιθυμητό τρόπο, πολύ πιθανό να υπάρχουν λογικά λάθη ή κενά στο σχεδιασμό. Τα προγράμματα αυτά έχουν εγκατεστημένες διάφορες επεκτάσεις (reasoners) ή αλλιώς “εργαλεία αυτόματης επαλήθευσης”, τα οποία ελέγχουν την οντολογία για λογικές και συντακτικές ασυνέπειες. Για τον έλεγχο της οντολογίας μπορούν να κληθούν και ειδικοί του πεδίου που εντάσσεται η οντολογία (Domain Experts), προκειμένου να κρίνουν αν οι έννοιες και οι σχέσεις τους είναι ακριβείς και η οντολογία ανταποκρίνεται στο στόχο της. Σε μετέπειτα χρόνο και οι ίδιοι οι χρήστες θα διαπιστώσουν τυχόν προβλήματα κατά τη χρήση της.

Τελευταία ενέργεια για μια νέα οντολογία είναι η τεκμηρίωσή της. Ο όρος αναφέρεται στην αναλυτική καταγραφή των πληροφοριών που σχετίζονται με την οντολογία και παρέχουν τις απαραίτητες δηλώσεις για την εγκυρότητα της και την επικοινωνία με τις μηχανές (Χρυσούλας 2009, σ.σ. 63-4). Πραγματοποιείται από αυτόν που ασχολήθηκε με την ανάπτυξη της οντολογίας και συνήθως καταγράφεται σε κείμενο το οποίο μπορεί να αποθηκευτεί σε διάφορες μορφές: αρχείο κειμένου (π.χ. σε PDF), online στον ιστότοπο που θα φιλοξενήσει και την εφαρμογή, σε ένα εργαλείο που θα βασίζεται στην οντολογία ή αν η οντολογία κατασκευάστηκε σε πρόγραμμα ανοιχτού κώδικα να καταγραφούν ως σχόλια μέσα στο ίδιο το εργαλείο (για παράδειγμα στο Protégé μπορούν να ενσωματωθούν πληροφορίες στις κλάσεις και τις ιδιότητες ως σχόλια [Annotations])²⁰.

Η διαδικασία της τεκμηρίωσης περιλαμβάνει την λεπτομερή επεξήγηση των όρων, των εννοιών και των σχέσεων. Σχολιάζεται η επιλογή και τα χαρακτηριστικά τους, η χρήση και η σημασία τους. Περιγράφεται ακόμα η μεθοδολογία που ακολουθήθηκε και εξηγούνται οι λόγοι για τις επιλογές και τις αποφάσεις που πάρθηκαν. Είναι σημαντική διαδικασία, διότι μπορεί να χρησιμοποιηθεί ως οδηγία χρήσης για το πως θα χρησιμοποιείται η οντολογία, αλλά και πως μπορεί να επεξεργαστεί και να επεκταθεί. Η εξήγηση των όρων είναι απαραίτητη για την επαναχρησιμοποίησή της. Το ίδιο και η περιγραφή της μεθόδου και της χρήσης της, αφορούν κυρίως την κατανόηση για την μελλοντική επεξεργασία της. Στην τεκμηρίωση περιλαμβάνονται και παραδείγματα (πιθανόν και με τη μορφή SPARQL ερωτημάτων) που αποδεικνύουν τη λειτουργικότητα, το εύρος και την αποτελεσματικότητα της στην ανάκτηση δεδομένων.

²⁰ Για την οντολογία που δημιουργήθηκε στο πλαίσιο αυτής της Διπλωματικής εργασίας, τα επόμενα κεφάλαια αποτελούν την τεκμηρίωσή της.

β. Τεκμηρίωση του προτύπου καταγραφής πηγών

Στο παρόν κεφάλαιο θα παρουσιαστεί το οντολογικό πρότυπο το οποίο κατασκευάστηκε και προτείνεται για την καταγραφή των αρχαίων πηγών και θα αναλυθεί η δομή του και οι επιλογές που το αποτελούν. Αυτή η παρουσίαση είναι ιδιαίτερα σημαντική για την υποστήριξη του προτύπου ως ένα λειτουργικό, αξιόλογο και αξιόπιστο μέσο, εφόσον θα εξηγηθεί η μεθοδολογία της κατασκευής και η δομή του. Η τεκμηρίωση λειτουργεί ως ένας οδηγός και η εξήγηση των όρων είναι απαραίτητη για την κατανόηση, τη μελλοντική χρήση και την επαναχρησιμοποίηση της οντολογίας. Δίνεται επίσης βαρύτητα στην περιγραφή της δομής και της λειτουργικότητας της οντολογίας. Η οντολογία κατασκευάστηκε στο Protégé, ένα πρόγραμμα ανοιχτού κώδικα, κατάλληλου για τη δομημένη μοντελοποίηση των δεδομένων, χωρίς να απαιτείται από τους χρήστες του ιδιαίτερη εξοικείωση με τον προγραμματισμό²¹. Η παρουσίαση που θα ακολουθήσει θα γίνει με την υποστηρικτική προσθήκη εικόνων καταγραφής οθόνης από το πρόγραμμα (printscreens).

Η δημιουργία μιας λειτουργικής αλλά και πρακτικά χρήσιμης οντολογίας προϋποθέτει προσεκτική οργάνωση και σωστό σχεδιασμό. Στο προηγούμενο υποκεφάλαιο περιγράψαμε εκτενώς τα βήματα της διαδικασίας κατασκευής μιας οντολογίας. Τα πρώτα από τα βήματα (είναι τα 1-3, στα οποία περιλαμβάνεται ο προσδιορισμός του θέματος και των βασικών εννοιών και η ιεράρχησή τους) είναι αρκετά θεωρητικά, αλλά και τα σημαντικότερα για μια νέα οντολογία. Πρώτο ζητούμενο είναι η αναζήτηση απαντήσεων σε υποθετικές ερωτήσεις, όπως ο σκοπός, οι στόχοι, το κοινό και ο τρόπος χρήσης. Οι απαντήσεις θα αποτελέσουν στη συνέχεια το βασικό κριτήριο για την επιλογή των εννοιών. Στην περίπτωση μας, ο στόχος ήταν η δημιουργία ενός προτύπου οργάνωσης της καταγραφής των αρχαίων γραμματειακών πηγών σε μια βάση δεδομένων που θα απευθύνεται σε κάθε λογής ερευνητές, οι οποίοι θα αναζητούν την γρήγορη και εύκολη ανάκτηση των αποσπασμάτων που επιθυμούν μαζί με τα κύρια υποστηρικτικά τους δεδομένα. Το δεύτερο βήμα είναι ο προσδιορισμός των βασικών όρων και των εννοιών (κλάσεις και ιδιότητες) της οντολογίας. Οι κλάσεις και οι ιδιότητες διαμορφώνονται με γνώμονα την ακρίβεια, τη σαφήνεια και την κάλυψη των προϋποθέσεων που τέθηκαν εξ αρχής ώστε να αποτυπώνουν το επιθυμητό αποτέλεσμα. Σε ένα πρώιμο στάδιο αρκεί μια απλή καταγραφή τους. Ενώ εξελίσσεται η διαδικασία, ο δημιουργός της οντολογίας συνεχώς επεξεργάζεται το μοντέλο του, το επαναξιολογεί και προχωρά σε τροποποιήσεις, προσθέτοντας, αφαιρώντας ή αλλάζοντας έννοιες και όρους (Μαυρογιάννη 2007, σ.σ. 103-18; Uschold & Gruninger 1996, σ.σ. 93-155). Μετά από τον προσδιορισμό των όρων, επόμενο βήμα είναι η ταξινόμηση. Πρόκειται για ένα σημαντικό βήμα που καθορίζει την εξέλιξη της οντολογίας, καθώς εδώ δημιουργείται το οντολογικό πλαίσιο και οι όροι ταξινομούνται δομικά και ιεραρχικά. Η ιεραρχική ταξινόμηση των κλάσεων δεν είναι αρκετή ώστε να εξασφαλίζει τις πολύπλοκες σχέσεις που εξυπηρετεί μια οντολογία, συνεπώς χρησιμοποιείται ένα δεντρικό μοντέλο αναπαράστασής της. Είναι συνετή η δημιουργία ενός σχεδίου, καθώς η οπτική απεικόνιση της ταξινόμησης είναι χρήσιμη πρακτική -ειδικά για τους λιγότερο εξοικειωμένους δημιουργούς- που επιτρέπει την καλύτερη οργάνωση του οντολογικού πλαισίου, την αποφυγή λογικών λαθών ή παραλήψεων. Εφόσον αποδειχθεί επαρκής, στη συνέχεια να την αποτυπώνει μέσω της κωδικοποίησης. Συνεπώς δεν είναι μόνο μια βοηθητική σχεδιαστική αποτύπωση των σκέψεων, αλλά αποτελεί έναν οδηγό για τη σύνταξη του κώδικα.

Ακολουθούν δύο σχεδιαγράμματα (“Εικόνες 1 & 2”) οπτικής απεικόνισης του οντολογικού σχεδιασμού. Η “Εικόνα 1” αποτελεί την αποτύπωση της ανάγκης για αναπαράσταση της ιδέας με βάση την ανθρώπινη “σημασιολογική ερμηνεία” για το θέμα της οντολογίας. Είναι γραμμένο σε φυσική γλώσσα και εφαρμόζει τις θεωρητικές σκέψεις για τις έννοιες της οντολογίας που θα κατασκευαζόταν και τις σχέσεις τους. Πρόκειται για μια αναπαράσταση δεντρικού τύπου, όπου οι κλάσεις και οι ιδιότητες των δεδομένων εντάσσονται σε ορθογώνιο σχήμα και οι ιδιότητές τους αποτυπώνονται με τη μορφή βελών. Το σχεδιάγραμμα στην “Εικόνα 2” αναπαριστά το πρώτο

²¹ Παραδείγματα τεκμηρίωσης και επεξήγησης μιας οντολογίας και τεχνικά ζητήματα ανάπτυξης οντολογιών στο Protégé, βλ. στα Debelliss 2021, σ.σ. 1- 61; Ιακωβάκης 2019, σ.σ. 50-85; Malviya et al. 2011, σ.σ. 3-5; Τριανταφύλλου 2022, σ.σ. 31-42. Η οντολογία είναι διαθέσιμη στον παρακάτω σύνδεσμο: https://drive.google.com/file/d/1xRvYJDQZmYt3LIEuPBLmLgRhlLQ4ILS/view?usp=drive_link

σχεδιάγραμμα έπειτα από τα βήματα της κωδικοποίησης και την μερική αντικατάσταση της φυσικής γλώσσας με την οντολογική της αντιστοιχία.



Εικόνες 1 & 2: Σχεδιαγράμματα οπτικής αναπαράστασης της οντολογίας σε φυσική γλώσσα

Η θεωρητική σύλληψη της ιδέας της οντολογίας, ξεκινά να κωδικοποιείται με την επιλογή του ονόματός της. Η επιλογή του βασίζεται στην ανάγκη να αποτυπώνεται με ακρίβεια το περιεχόμενο και ο σκοπός της, να διασφαλιστεί η αναγνωσιμότητα της και να βοηθά την αναζήτησή της παρέχοντας πληροφορίες για το αντικείμενό της. Το όνομα πρέπει να είναι περιγραφικό και να περιλαμβάνει λέξεις-κλειδιά που να αναδεικνύουν το περιεχόμενό της, αλλά ταυτόχρονα κατανοητό, σύντομο και ευέλικτο (ώστε να αναγνωρίζεται και να χρησιμοποιείται).



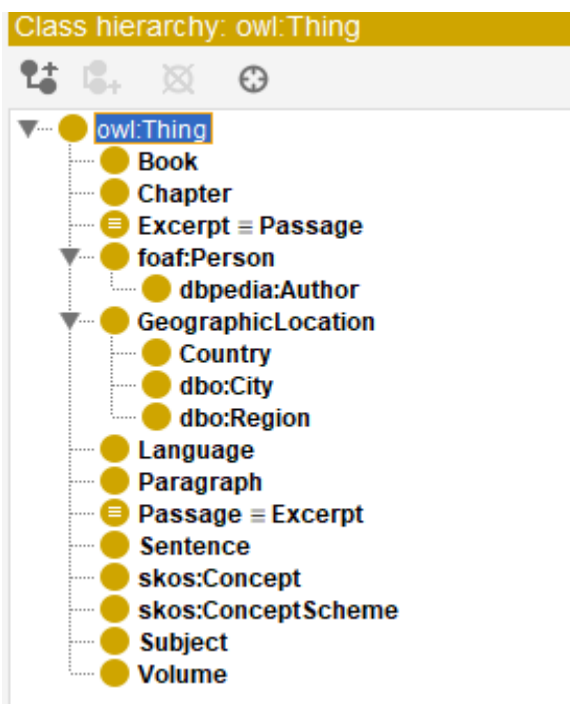
Εικόνα 3: το μοναδικό χαρακτηριστικό (IRI) της Οντολογίας

Για τον σκοπό μας επιλέχθηκε η ονομασία (IRI) “http://www.SourceArchiveOnto-ontology.com/”, η οποία δηλώνει πως η οντολογία επικεντρώνεται στην αρχειοθέτηση πηγών και διαθέτει αυστηρή σημασιολογία και δομή. Διευκρινίζοντας τμηματικά το “SourceArchiveOnto”, αιτιολογούμε ως εξής: Το “Source” αναφέρεται στο αντικείμενο της οντολογίας, το οποίο είναι οι πηγές. Το “Archive” συνδέεται με τις πηγές καθώς είναι η αρχειοθετημένη συλλογή πηγών, ο οποίος είναι ο κύριος στόχος της οντολογίας. Το “Onto(-logy)” σημειώνει ότι το πρότυπο βασίζεται σε μια δομημένη σημασιολογική αναπαράσταση. Για να συνομειώσει ο όρος ώστε να είναι διαχειρισμός κατά την ανάλυση της οντολογίας, επιλέξαμε τη χρήση του prefix “sao:”

Ο πυρήνας μιας οντολογίας είναι οι κλάσεις της (βλ. “Εικόνες 4 και 5”), οι έννοιες που περιγράφουν το γνωστικό πεδίο. Επομένως, οφείλουν να πληρούν τα κριτήρια της ακρίβειας (να ανταποκρίνονται στο θέμα και το σκοπό) και της ευελιξίας (να μπορούν να επαναχρησιμοποιηθούν). Για το πρότυπο που αναπτύξαμε δημιουργήθηκαν 17 κλάσεις. Η ιεράρχηση που εφαρμόζεται από το Protégé, με εξαίρεση τις υποκλάσεις, είναι αλφαβητική και δε σχετίζεται με τη νοηματική ή τη σημασιολογική ερμηνεία των κλάσεων. Συγκεκριμένα επιλέχθηκαν οι εξής (βλ. Εικόνα 5): “bibo:Book” (“Βιβίο”), “bibo:Chapter” (“Κεφάλαιο”), “bibo:Excerpt” (“Απόσπασμα”), “foaf:Person” (“Ατομο”), “dbpedia:Author” (“Συγγραφέας”), “vivo:GeographicLocation” (“Τόπος Καταγωγής”), “juso:Country” (“Χώρα”), “dbo:City” (“Πόλη”), “dbo:Region” (“Περιοχή”), “bibframe:Language” (“Γλώσσα”), “doco:Paragraph” (“Παράγραφος”), “sao:Passage” (“Απόσπασμα”), “doco:Sentence” (“Πρόταση”), “skos:Concept” (“Είδος”), “skos:ConceptScheme”, “getty:Subject” (“Θέμα”) και η “tribont:Volume” (“Τόμος”).

Ζητούμενο της οντολογίας είναι η πρόταση ενός λειτουργικού και ευέλικτου προτύπου αρχειοθέτησης των αρχαίων γραμματειακών πηγών, το οποίο θα απευθύνεται στους ερευνητές και θα τους δίνει τη δυνατότητα να αναζητήσουν εύκολα και γρήγορα βασικές πληροφορίες για τα αποσπάσματα που τους ενδιαφέρουν. Η επιλογή των όρων στηρίχθηκε στην προσωπική μου ερευνητική εμπειρία και κατανόηση των αναγκών του πεδίου, αλλά και στο γενικότερο ανθρώπινο

σημασιολογικό πρότυπο. Αυτό αντικατοπτρίζει συλλογικά αποδεκτές αρχές για την οργάνωση και τη διασύνδεση δεδομένων, ενσωματώνοντας τις αντιλήψεις για τις πιο συνηθισμένες και σημαντικές αναζητήσεις των χρηστών. Κεντρική θέση στο συλλογισμό μου είχαν το “Απόσπασμα”, ο “Συγγραφέας” και το “Θέμα”²². Οι υπόλοιπες κλάσεις προέκυψαν από τις απαντήσεις διευκρινιστικών ερωτημάτων με βάση αυτό το “πρότυπο”. Έτσι (βλ. “Εικόνα 1”) το “Απόσπασμα” προέρχεται από κάποιο “Βιβλίο” και πραγματεύεται ένα συγκεκριμένο “Θέμα”. Το “Βιβλίο” μπορεί να έχει “Τόμο”, “Κεφάλαιο”, “Παράγραφο” και “Πρόταση” ή κάποιο συνδυασμό αυτών. Σίγουρα θα ανήκει σε κάποιο “Είδος”, θα είναι γραμμένο σε μια “Γλώσσα” και θα έχει ένα “Συγγραφέα”. Ο “Συγγραφέας” είναι “Πρόσωπο” και έχει μια “Καταγωγή” η οποία μπορεί να είναι “Χώρα”, “Περιοχή” ή “Πόλη”.



Εικόνα 4: Ιεραρχία των κλάσεων- εικόνα στιγμιότυπου οθόνης από το Protégé

Η κλάση “foaf:Person” επιλέχθηκε ως υπερκλάση για την κλάση “dbpedia:Author” για την λογική πληρότητα του προσδιορισμού (<http://xmlns.com/foaf/0.1/Person>: *The Person class represents people. Something is a Person if it is a person...*). Αντίθετα, ως υπερκλάση με την έννοια μιας μεγαλύτερης ομάδας ενός συνόλου υποκλάσεων με κοινά χαρακτηριστικά έχει η κλάση “vino:GeographicLocation” (“Καταγωγή”) και συγκεκριμένα τις κλάσεις “juso:Country” (Χώρα), “dbo:Region” (Περιοχή) και “dbo:City” (Πόλη), αν και εμφανίζονται ως ισότιμες κλάσεις. Ειδικά για τα τεκμήρια που προέρχονται από την αρχαιότητα, ο τόπος καταγωγής δεν είναι πάντα γνωστός με ακρίβεια ώστε να μπορεί να ενσωματώνεται σε μια ιεραρχική ταξινόμηση· αυτή η πληροφορία συνήθως δεν είναι χρήσιμη ή απαραίτητη. Παρότι εύλογη νοηματικά υπόθεση, οι κλάσεις “Απόσπασμα” (“sao:Passage”) ή “Βιβλίο” (“bibo:Book”) δεν είναι υπερκλάσεις για το “Κεφάλαιο” (“bibo:Chapter”) και αυτό έπειτα του “Τόμου” (“tribont:Volume”) και αντίστοιχα για την “Παράγραφο” (“doco:Paragraph”) και την “Πρόταση” (“doco:Sentence”), όπως προκύπτει από δύο τεχνικούς λόγους. Προκειμένου να είναι ορθή η οντολογική δομή και σύμφωνα με την ιεραρχική ταξινόμηση, κάθε στιγμιότυπο μιας υποκλάσης πρέπει υποχρεωτικά να “ανήκει” σε κάποιο στιγμιότυπο της ευρύτερης κλάσης. Στο συγκεκριμένο πεδίο αυτό δε μπορεί να υποστηριχθεί. Σύμφωνα με τη φιλολογική προσέγγιση η απόδοση των χωρίων (αποσπάσματα, passage) γίνεται με βάση της αρίθμηση των εκδόσεων των πηγών που χρησιμοποιήθηκε. Αυτή γίνεται με διάφορα κριτήρια (ανάλογα το μέγεθος του έργου, την κατάσταση των χειρόγραφων κ.α.) και δεν υπάρχει μια συγκεκριμένη μορφή στην οποία απαρένγκλιτα κάθε απόσπασμα θα έχει όλες τις πιθανές

²² Όπως έχει ειπωθεί, αυτοί ήταν και οι πρώτοι όροι που ξεχωρίζουν σε μια διαδικασία αρχειοθέτησης.

υποδιαιρέσεις. Η σημαντικότερη έννοια μιας οντολογίας για την καταγραφή πηγών είναι το "sao:Passage" (απόσπασμα, χωρίο) καθώς με αυτή την ορολογία έρχεται σε επαφή ο ιστορικός στην έρευνά του στις δευτερογενείς πηγές. Έχουμε επισημάνει τη σημασία της ακρίβειας στις έννοιες που αντιπροσωπεύουν οι κλάσεις. Σε αυτή την περίπτωση δεν κατέστη δυνατό να βρεθεί σε μια επίσημη οντολογία ένας όρος που να ταιριάζει απόλυτα η περιγραφή του με τη σωστή σημασία που θέλουμε να αποδοθεί. Γι' αυτό το λόγο δημιουργήσαμε μια δική μας, χρησιμοποιώντας την ονομασία της οντολογίας μας. Προκειμένου όμως να είναι αξιόπιστη και να μπορεί να επαναχρησιμοποιηθεί με ασφάλεια, επιλέχθηκε και η έννοια "bibo:Excerpt" (από την BIBO: <http://purl.org/ontology/bibo/Excerpt>, όπου description: "A passage selected from a larger work"). Αφού το "sao:Passage" είναι η βιβλιογραφική απόδοση ενός αποσπάσματος, επιλέχθηκε η χρήση της δήλωση "Equivalent to", δηλαδή πως οι κλάσεις ":sao:Passage" και "bibo:Excerpt" είναι "ισοδύναμες" και όπου χρησιμοποιείται το "Passage", αυτό γίνεται με την σημασία του "Excerpt". Η επαναχρησιμοποίηση είναι από τους βασικές απαιτήσεις στη διαδικασία κατασκευής μιας οντολογίας και για να είναι δυνατή και ευκολότερη, απαιτείται η ενσωμάτωση²³. Για το πρότυπο που φτιάξαμε, όλες οι κλάσεις που επιλέχθηκαν υπήρξε εφικτό να βρεθούν οι επίσημοι όροι με την επιθυμητή σημασιολογία και να ενσωματωθούν στην οντολογία αντικαθιστώντας τα URI τους.

Στον παρακάτω πίνακα (Εικόνα 5) βλέπουμε πως 3 κλάσεις ενσωματώθηκαν από την DBPEDIA ("dbpedia:Author", "dbp:City", "dbo:Region"), 3 από την BIBO ("bibo:Book", "bibo:Chapter", "bibo:Excerpt"), 2 από την SKOS ("skos:Concept", "skos:ConceptScheme"), 2 από την DoCO ("doco:Paragraph", "doco:Sentence"), και από 1 από τις JUSO ONTOLOGY ("juso:Country"), FOAF ("foaf:Person"), BIBFRAME ("bubframe:Language"), GETTY VOCABULARY ("getty:Subject"), VIVO ("vivo:GeographicLocation") και TribOnt ("tribont:Volume"). Τέλος το IRI της κλάσης "sao:Passage" κατασκευάστηκε για τις ανάγκες της οντολογίας και έχει το Prefix της, SAO.

CLASSES

Class IRI	label EN	label EL	Ontology
http://dbpedia.org/resource/Author	Author	Συγγραφέας	DBPEDIA
http://purl.org/ontology/bibo/Book	Book	Βιβλίο	BIBO
http://www.SourceArchiveOnto-ontology.com/Passage	Passage	Χωρίο	SAO
http://id.loc.gov/ontologies/bibframe/Language	Language	Γλώσσα	BIBFRAME
https://w3id.org/tribont/core/Volume	Volume	Τόμος	TribOnt
http://purl.org/ontology/bibo/Chapter	Chapter	Κεφάλαιο	BIBO
http://purl.org/spar/doco/Paragraph	Paragraph	Παράγραφος	DoCO
http://purl.org/spar/doco/Sentence	Sentence	Περίοδος	DoCO
http://vocab.getty.edu/ontology/Subject	Subject	Θέμα	GETTY VOCABULARY
http://www.w3.org/2004/02/skos/core/Concept	Concept	Είδος	SKOS
http://www.w3.org/2004/02/skos/core/ConceptScheme	Concept Scheme	Κατηγορία	SKOS
http://vivoweb.org/ontology/core/GeographicLocation	Geographic Location	Καταγωγή	VIVO
http://dbpedia.org/ontology/Region	Region	Περιοχή	DBPEDIA
http://dbpedia.org/ontology/City	City	Πόλη	DBPEDIA
http://rdfs.co/juso/Country	Country	Χώρα	JUSO ONTOLOGY
http://xmlns.com/foaf/0.1/Person	Person	Άτομο	FOAF
http://purl.org/ontology/bibo/Excerpt	Excerpt	Χωρίο	BIBO

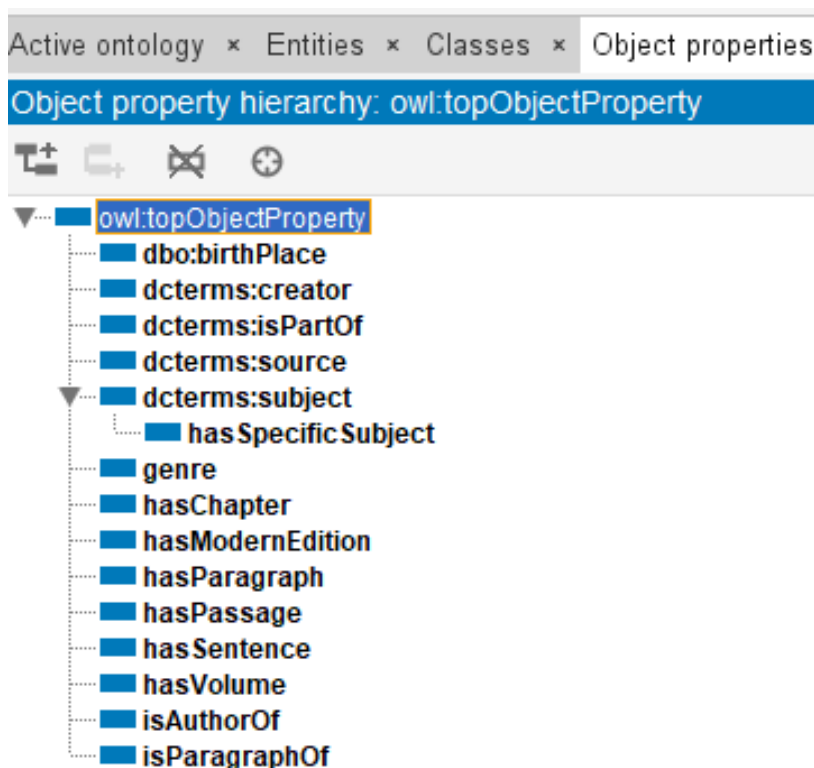
Εικόνα 5: Classes IRI- labels- Οντολογίες ενσωμάτωσης των κλάσεων

Μετά από τον καθορισμό των εννοιών και των κλάσεων, ακολουθεί ο ορισμός των ιδιοτήτων. Και εδώ, τα κριτήρια για την επιλογή τους βασίστηκαν στο στόχο της οντολογίας για γρήγορη και αποτελεσματική αρχειοθέτηση των αρχαίων πηγών. Ορίστηκαν με γνώμονα συγκεκριμένες δομικές και λειτουργικές ανάγκες της οντολογίας, αλλά και το σημασιολογικό πρότυπο για τις βασικές ανάγκες ενός ερευνητή. Αποτυπώνουν το είδος των πληροφοριών που οι ερευνητές αναζητούν πρώτες και επιθυμούν να βρύνουν εύκολα και γρήγορα, αλλά και τον τρόπο που

²³ Τι είναι και τι προσφέρουν σε μια οντολογία θεωρητικά, έχει σχολιαστεί παραπάνω. Περισσότερα βλέπε σε επόμενο κεφάλαιο, αναφορικά με τη μελλοντική ανάπτυξη της οντολογίας.

αντιλαμβάνονται και οργανώνουν τη γνώση²⁴. Οι ιδιότητες χωρίζονται σε δύο κατηγορίες, τις “object properties” οι οποίες αντιπροσωπεύουν τη σύνδεση δύο κλάσεων δημιουργώντας λογικές προτάσεις²⁵, και τις “datatype properties”, οι οποίες προσθέτουν πληροφορίες στα στιγμιότυπα, αποδίδοντάς τους κάποιο ιδιαίτερο χαρακτηριστικό ή μια συγκεκριμένη τιμή²⁶. Συνήθως αποτελούνται από μια φράση που περιλαμβάνει ρηματικό τύπο και ουσιαστικό (λ.χ. “hasPassage”) ή από ένα ουσιαστικό που μεταφράζεται με αυτόν τον τρόπο σε φυσική γλώσσα (λ.χ. “creator”, “έχει συγγραφέα”).

Όσον αφορά τις “object properties” (βλ. “Εικόνες 6-8”) κρίθηκε απαραίτητη η δημιουργία 15 συνδέσεων μεταξύ ορισμένων κλάσεων. Αυτές είναι: “cd:isAuthorOf” (“είναι συγγραφέας του”), “dbo:birthPlace” (“έχει καταγωγή”), “lod:genre” (“ανήκει σε είδος”), “sao:hasModernEdition” (“έχει σύγχρονη πηγή”), “dcterms:creator” (“έχει συγγραφέα”), “dcterms:subject” (“έχει κεντρικό θέμα”), “sao:hasSpecificSubject” (“έχει συγκεκριμένο θέμα”), “sao:hasPassage” (“έχει απόσπασμα”), “dcterms:source” (“βρίσκεται στο”), “arpreteur:hasVolume” (“έχει τόμο”), “dcterms:isPartOf” (“περιέχεται σε”), “gdprt:hasChapter” (“έχει κεφάλαιο”), “sao:hasParagraph” (“έχει παράγραφο”), “sao:isParagraphOf” (“περιέχεται σε παράγραφο”) και “sao:hasSentence” (“περιέχει γραμμές”).



Εικόνα 6: οι Object Properties της οντολογίας - εικόνα στιγμιότυπου οθόνης από το Protégé

Για λόγους αξιοπιστίας και τη δυνατότητα επαναχρησιμοποίησης, όπως συνέβη και με τις κλάσεις, έγινε προσπάθεια να βρεθούν και να ενσωματωθούν πόροι από επίσημες οντολογίες (Εικόνα 7). Και στις “ιδιότητες αντικειμένων” οι περισσότερες (4) ενσωματώθηκαν από την

²⁴ Ωστόσο καθώς η δομή μιας οντολογίας την οδηγεί να επεκτείνεται, δεν ήταν δυνατό να χρησιμοποιηθεί κάθε πιθανή ιδιότητα. Χρησιμοποιήθηκαν αυτές που θεωρήσα πως καλύπτουν τις βασικότερες ανάγκες και με γνώμονα ότι στο πλαίσιο πρότασης για μια Διπλωματική Εργασία, έπρεπε να διατηρηθεί η οντολογία λειτουργική, συγκεκριμένη, ευέλικτη και όχι ιδιαίτερα πολύπλοκη.

²⁵ Για παράδειγμα, θα δούμε και παρακάτω, η ιδιότητα “dc:isAuthorOf” (“είναι συγγραφέας του”) συνδέει την κλάση “dbpedia:Author” με την κλάση “bibo:Book” (ο Συγγραφέας είναι συγγραφέας του Βιβλίου), οπότε όλα τα στιγμιότυπα της κλάσης Συγγραφέας πρέπει υποχρεωτικά να συνδέονται με ένα τουλάχιστον Βιβλίο.

²⁶ Παράδειγμα η ιδιότητα “dcterms:title” συνδέει ένα ‘στιγμιότυπο’ της κλάσης του Βιβλίου με τον Τίτλο του (που είναι σε μορφή κειμένου), ενώ η ιδιότητα “fabio:hasPublicationYear” συνδέει το Βιβλίο με ένα μοναδικό αριθμό, που αποδίδει την χρονιά δημοσίευσής του.

DUBLINCORE (“dcterms:Creator”, “dcterms:Source”, “dcterms:isPartOf”, “dcterms:Subject”) και από 1 οι ARCO (“cd:isAuthorOf”), ARPEUTEUR (“arpenteur:hasVolume”), LOD (“lod:genre”) και DBPEDIA (“dbo:birthPlace”). Αντίθετα με τις κλάσεις δεν μπόρεσαν να βρεθούν έτοιμοι οντολογικοί πόροι για όλες τις ιδιότητες και έτσι δημιουργήθηκαν 7 (“sao:hasPassage”, “sao:hasChapter”, “sao:hasModernEdition”, “sao:hasParagraph”, “sao:hasSentence”, “sao:isParagraphOf”, “sao:hasSpecificSubject”).

OBJECT PROPERTIES			
OBJECT PROPERTIES IRI	Label EN	Label EL	Ontology
https://w3id.org/arco/ontology/context-description/isAuthorOf	Is Author Of	είναι συγγραφέας του	ARCO
http://dbpedia.org/ontology/birthPlace	Birth Place	έχει καταγωγή	DBPEDIA
http://purl.org/dc/terms/creator	creator	έχει συγγραφέα	DUBLINCORE
http://lod.nl.go.kr/ontology/genre	genre	ανήκει σε είδος	LOD
http://www.SourceArchiveOnto-ontology.com/hasModernEdition	has Modern Edition	έχει σύγχρονη πηγή	SAO
http://purl.org/dc/terms/subject	subject	έχει κεντρικό θέμα	DUBLINCORE
http://www.SourceArchiveOnto-ontology.com/hasSpecificSubject	has Specific Subject	έχει ιδιαίτερο θέμα	SAO
http://www.SourceArchiveOnto-ontology.com/hasPassage	has Passage	έχει χωρίο	SAO
http://purl.org/dc/terms/source	source	Βρίσκεται στο	DUBLINCORE
http://www.arpenteur.org/ontology/Arpenteur.owl/hasVolume	has Volume	έχει τόμο	ARPEUTEUR ONTOLOGY
http://purl.org/dc/terms/isPartOf	Is Part Of	περιέχεται	DUBLINCORE
https://w3id.org/GDPRtEXT#hasChapter	has Chapter	Περιέχει κεφάλαιο	GDPRtEXT
http://www.SourceArchiveOnto-ontology.com/hasParagraph	has Paragraph	περιέχει παράγραφο	SAO
http://www.SourceArchiveOnto-ontology.com/isParagraphOf	is Paragraph Of	Βρίσκεται στην παράγραφο	SAO
http://www.SourceArchiveOnto-ontology.com/hasSentence	has Sentence	περιέχει γραμμές	SAO

Εικόνα 7: Object Properties- Labels- Οντολογία ενσωμάτωσης

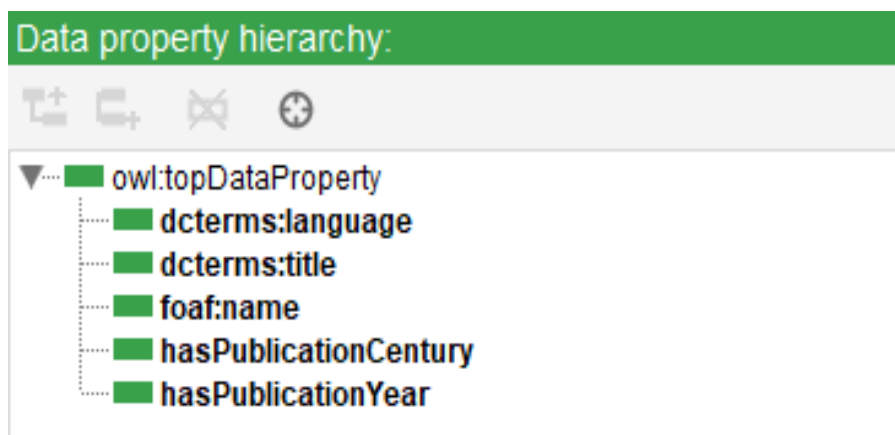
Η σύνδεση των κλάσεων αποτυπώνεται μέσω του περιορισμού για το ‘εύρος πεδίου τιμής. Για κάθε κλάσεων ορίζεται μια κλάση στο πεδίο “Domain” και μια άλλη στο πεδίο “Range”. Αυτό σημαίνει ότι από τη στιγμή που συνδέονται οι δύο αυτές κλάσεις μέσω αυτής της ιδιότητας, τότε τα στιγμιότυπα της μιας παίρνουν ως υποχρεωτική τιμή κάποιο στιγμιότυπο της άλλης. Ο περιορισμός αυτός οδηγεί στη δημιουργία προτάσεων οι οποίες αναδεικνύουν τη λειτουργικότητα της οντολογίας αλλά επίσης σχηματίζουν λογικές ακολουθίες. Για παράδειγμα, η ιδιότητα “cd:isAuthorOf” έχει ως “Domain” την κλάση “dbpedia:Author” και ως “Range” την κλάση “bibo:Book”. Επομένως ένα στιγμιότυπο της κλάσης “dbpedia:Author”, όπως π.χ. ο Θουκυδίδης, μπορεί να συνδεθεί μέσω αυτής με ένα στιγμιότυπο της κλάσης “bibo:Book”, όπως την οντότητα “sao:Istoria”, σχηματίζοντας τη λογική πρόταση “Ο Θουκυδίδης είναι συγγραφέας του βιβλίου Ιστορία”. Αυτός ο περιορισμός είναι σημαντικός για την ανάκτηση δεδομένων, όπως για παράδειγμα μέσω των SPARQL ερωτημάτων (βλ. κεφ. 2.2.γ.) γιατί ορίζει το πλαίσιο στο οποίο αναζητούνται οι πληροφορίες. Αν δεν καθοριστούν σωστά οι σχέσεις μεταξύ των στιγμιότυπων, δεν θα γίνεται ολοκληρωμένη η ανάκτηση των δεδομένων και τα στιγμιότυπα δεν θα εμφανιστούν ως αποτελέσματα στην αναζήτηση. Οι περισσότερες ιδιότητες έχουν ως “Domain” τους την κλάση του Βιβλίου. Διευκρινίζουν το είδος του (“lod:genre”), το (ιδιαίτερο) θέμα (“dcterms:subject” και “sao:hasSpecificSubject”) και τη σύγχρονη έκδοση του αρχαίου έργου (“sao:hasModernEdition”). Το Βιβλίο και το Απόσπασμα συνδέονται μεταξύ τους με την ιδιότητα “sao:hasPassage” (και την αντίστροφη της -inverse of- “dcterms:source”). Οι ιδιότητες “arpenteur:hasVolume”, “gtrpt:hasChapter”, “sao:hasParagraph”, “sao:hasSentence”, αλλά και οι “sao:isParagraphOf” και “dcterms:isPartOf” αφορούν τις υποδιαιρέσεις της οντότητας των Βιβλίων. Ο Συγγραφέας συνδέεται με το Βιβλίο του (μέσω της “cd:isAuthorOf”, αλλά και με της αντίστροφη της “dcterms:creator”) και τον Τόπο Καταγωγής του (“dbo:birthPlace”). Στο Protégé οι αντίθετες ιδιότητες εννοούνται αυτόματα, μέσω της λογικής ακολουθίας πως εφόσον ορίζεται μια σχέση, τότε η αντίστροφη της σίγουρα θα πρέπει να ισχύει. Παρότι εννοούνται, για την πληρότητα της

οντολογίας μπορούν να προστίθενται από τον κατασκευαστή της. Για να μη φορτωθεί το μοντέλο με πληροφορίες, εδώ δε δημιουργήσαμε χειροκίνητα όλες τις πιθανές αντίστροφες ιδιότητες.

Property IRI	Domain	Range
https://w3id.org/arco/ontology/context-description/isAuthorOf	dbpedia:Author	bibo:Book
http://dbpedia.org/ontology/birthPlace	dbpedia:Author	Vivo:GeographicLocation
http://purl.org/dc/terms/creator	bibo:Book	dbpedia:Author
http://lod.nl.go.kr/ontology/genre	bibo:Book	skos:Concept
http://www.SourceArchiveOnto-ontology.com/hasMdemEdition	bibo:Book	bibo:Book
http://purl.org/dc/terms/subject	sao:Passage	getty:Subject
http://www.SourceArchiveOnto-ontology.com/hasSpecificSubject	sao:Passage	getty:Subject
http://www.SourceArchiveOnto-ontology.com/hasPassage	bibo:Book	sao:Passage
http://purl.org/dc/terms/source	sao:Passage	bibo:Book
http://www.arpenteur.org/ontology/Arpenteur.owl/hasVolume	sao:Passage	tribont:Volume
http://purl.org/dc/terms/isPartOf	tribont:Volume bibo:Chapter doco:Sentence	sao:Passage doco:Paragraph
https://w3id.org/GDPRtEXT#hasChapter	sao:Passage	bibo:Chapter
http://www.SourceArchiveOnto-ontology.com/hasParagraph	sao:Passage	doco:Paragraph
http://www.SourceArchiveOnto-ontology.com/is-ParagraphOf	doco:Paragraph	sao:Passage
http://www.SourceArchiveOnto-ontology.com/hasSentence	sao:Passage	doco:Sentence

Εικόνα 8: εύρος πεδίου τιμής των object properties- εικόνα στιγμιότυπου οθόνης από το Excel

Εκτός από τις ιδιότητες που συνδέουν τις κλάσεις μεταξύ τους, υπάρχουν και οι “datatype properties” ή αλλιώς “ιδιότητες δεδομένων” (βλ. Εικόνες 9, 10, 11). Προσφέρουν μοναδικά και συγκεκριμένα χαρακτηριστικά σε κάθε στιγμιότυπο μιας κλάσης. Οι “datatype properties” χρησιμοποιούνται όταν η πληροφορία που θα αποδώσουν είναι μια συγκεκριμένη απάντηση, η οποία δε χρειάζεται περαιτέρω ανάλυση ή διευκρίνιση. Συμβάλουν στην απλότητα της οντολογικής δομής και στην ευελιξία της, καθώς σε κάθε νέα προσθήκη στιγμιότυπου μπορεί να προστεθεί η τιμή της ιδιότητας εύκολα και γρήγορα χωρίς τη δημιουργία πολύπλοκων συνδέσεων. Για τις ανάγκες της οντολογίας καταγραφής πηγών, δημιουργήθηκαν 5 “datatype properties”. Αυτές αφορούν τον τίτλο (“dcterms:title”), τη γλώσσα (“dcterms:language”), τη χρονολογία ή τον αιώνα δημοσίευσης του βιβλίου (“fabio:hasPublicationYear” και “sao:hasPublicationCentury”) και το όνομα του Συγγραφέα (“foaf:name”)



Εικόνα 9 : οι Object Properties της οντολογίας - εικόνα στιγμιότυπου οθόνης από το Protégé

Το “Domain” αυτών των ιδιοτήτων είναι κλάση, αλλά το “Range” δεν είναι μια άλλη όπως στις object properties, αλλά μια συγκεκριμένη ξεχωριστή τιμή (literal). Οι τιμές αυτές συχνά είναι

μορφής “String” (κείμενο) ή “Integer” (ακέραιος αριθμός), αλλά υπάρχει επίσης πλήθος άλλων τύπων literal τιμών για πιο εξειδικευμένες καταχωρήσεις, όπως οι δεκαδικοί αριθμοί (Float/Double), οι λογικές τιμές για μια αληθή ή ψευδή δήλωση (Boolean), η ημερομηνία (Date), το URI (μοναδικό αναγνωριστικό πόρου) και άλλοι. Στο πρότυπό μας έχουμε 3 datatype properties τύπου “string” (τίτλος, γλώσσα, όνομα) και 2 τύπου “integer” (χρονολογία δημοσίευσης και αιώνας δημοσίευσης). Η ιδιότητα “dcterms:language” (τύπου string) προσφέρει πληροφορία για τη γλώσσα που είναι γραμμένη η πηγή. Παρότι στο παρόν πρότυπο η γλώσσα είναι αποκλειστικά η Αρχαία Ελληνική, χρησιμοποιήθηκε η ιδιότητα αυτή και μάλιστα υπάρχει και κλάση “Language” σε περίπτωση μελλοντικής προσθήκης. Η ιδιότητα “dcterms:title” (επίσης string- κείμενο) αποδίδει τον τίτλο του Βιβλίου γραμμένο σε φυσική γλώσσα. Οι ιδιότητες “hasPublicationYear” (integer/ αριθμός) και “hasPublicationCentury” (πρόκειται για αριθμητικό προσδιορισμό με επιπρόσθετη χρήση κειμένου ή συμβόλων) προσδιορίζουν την ακριβή χρονολογία ή τον αιώνα δημοσίευσης ενός έργου. Σε έργα που προέρχονται από την αρχαιότητα η διάκριση αυτή είναι επιβεβλημένη καθώς δε γνωρίζουμε πάντα με ακρίβεια αυτή την πληροφορία. Η ιδιότητα “foaf:name” συνδέει τις οντότητες από τα στιγμιότυπα της κλάσης “dbpedia:Author” με το “πραγματικό” τους όνομα.

Και εδώ έγινε προσπάθεια ενσωμάτωσης των ιδιοτήτων: 2 ιδιότητες ενσωματώθηκαν από την οντολογία της DUBLICORE (“dcterms:language”, “dcterms:title”), 1 από τη FAbiO (“fabio:hasPublicationYear”), 1 από τη FOAF (“foaf:name”) και 1 δημιουργήθηκε καθώς δεν βρέθηκε ο αντίστοιχος πόρος (“sao:hasPublicationCentury”). Οι ιδιότητες επιλέχθηκαν γιατί προσδιορίζουν μερικές από τις βασικότερες αναζητήσεις των ερευνητών σε σχέση με τις πηγές που συλλέγουν.

Property IRI	Label EN	Label EL	Domain	Range	Type
http://purl.org/dc/terms/title	Title	Έχει τίτλο	bibo:Book	literal	String
http://purl.org/dc/terms/language	Language	Είναι σε γλώσσα	bibo:Book	literal	String
http://xmlns.com/foaf/0.1/name	Name	Έχει όνομα	dbpedia:Author	literal	String
http://www.SourceArchive-Onto-ontology.com/hasPublicationCentury	has Publication Century	Γράφτηκε αιώνα	bibo:Book	literal	Integer
http://purl.org/spar/fabio/hasPublicationYear	has Publication Year	Γράφτηκε χρόνο	bibo:Book	literal	Integer

Εικόνα 10: Property IRI, Prefixes, Labels, εύρος πεδίου τιμής και τύπος των data properties- εικόνα στιγμιότυπου οθόνης από το Excel

Όπως θα δούμε και εκτενέστερα στο επόμενο κεφάλαιο από την εφαρμογή του οντολογικού προτύπου, τα στιγμιότυπα μιας κλάσης συμβολίζουν οντότητες που φέρουν μια πληροφορία και οι οποίες χρειάζονται τις ιδιότητες για να αποκτήσουν νόημα και λειτουργία στο πλαίσιο της οντολογίας. Στο σημείο αυτό όμως είναι χρήσιμη η επεξήγηση ενός παραδείγματος προκειμένου να γίνει αντιληπτή η χρησιμότητα και ο ρόλος των “data properties” σε μια οντολογία. Έχουμε για παράδειγμα το στιγμιότυπο “sao:Istoria” από την κλάση “bibo:Book”. Αυτό το στιγμιότυπο δεν είναι αποτύπωση ενός φυσικού αντικείμενου, αλλά μια οντολογική οντότητα στην οποία μπορούμε να δώσουμε όλες τις πληροφορίες σχετικά με το αντικείμενο αυτό. Οι “datatype properties” συμβάλουν και αυτές ώστε να γίνει αυτή η οντότητα πιο συγκεκριμένη: με την ιδιότητα “dcterms:title” λαμβάνει τον τίτλο “Ιστορία”, με την “dcterms:language” διευκρινίζεται ότι το πρωτότυπο κείμενο είναι γραμμένο στα “Αρχαία Ελληνικά” και με την ιδιότητα “sao:hasPublicationCentury” προσδιορίζεται η έκδοσή του στο “β’ μισό του 5^{ου} π.Χ. αιώνα”. Με αυτόν τον τρόπο η οντότητα λαμβάνει μια συγκεκριμένη ταυτότητα και χαρακτηρίζεται από σταθερές τιμές. Αυτός ο ακριβής προσδιορισμός και η αναπαράστασή του, επιτρέπει στους ερευνητές να αναζητούν είτε συγκεκριμένες πληροφορίες για μια ιδιαίτερη οντότητα εύκολα και γρήγορα, είτε να αναζητούν με βάση τις μοναδικές τιμές (θα δούμε στα SPARQL ερωτήματα, -βλ. κεφ. 2.2.γ) για παράδειγμα “αναζήτηση των ιστορικών έργων του 5^{ου} π.Χ. αιώνα”).

2. Μελέτη Περίπτωσης: οι αρχαίες πηγές της δυτικής Λοκρίδας

α. Επιλογή και επεξήγηση του θέματος

Οι αρχαίες γραμματειακές πηγές είναι κομβικές για την ιστορική έρευνα, καθώς παρέχουν πολύτιμες πληροφορίες απαραίτητες για την κατανόηση του ανθρώπινου παρελθόντος. Η τεχνολογία έχει αλλάξει ριζικά τους τρόπους και τη μέθοδο που οι ιστορικοί οργανώνουν, διασυνδέουν και τελικά παρουσιάζουν αυτές τις πηγές, με τη χρήση νέων εργαλείων, όπως η ψηφιοποίηση και οι οντολογίες. Πλέον πολλοί οργανισμοί και ιδιώτες ψηφιοποιούν, επεξεργάζονται και δημοσιεύουν τα δεδομένα τους, ωστόσο, υπάρχουν ακόμα πολλά πεδία που παραμένουν εκτός της διαδικασίας, γεγονός που περιορίζει σημαντικά την πρόσβαση, γνώση και αξιοποίησή τους στην έρευνα.

Το 2017 συγκέντρωνα υλικό για τη συγγραφή της Διπλωματικής εργασίας μου για το Π.Μ.Σ. “*Αρχαίος Κόσμος: Αρχαία Ελληνική και Ρωμαϊκή Ιστορία*” του τμήματος Ιστορίας και Αρχαιολογίας στο Πανεπιστήμιο Ιωαννίνων, με θέμα “*Η Πολιτική ιστορία της δυτικής Λοκρίδας (6^{ος}-4^{ος} αιώνες π.Χ.)*”. Η έρευνα γρήγορα με οδήγησε στη διαπίστωση σημαντικών προβλημάτων: πρώτα από όλα οι πηγές είναι εξαιρετικά διάσπαρτες, καθιστώντας τη συγκέντρωση και τη συστηματική μελέτη τους δύσκολο εγχείρημα. Μάλιστα τελικά αποδείχθηκε πως ήταν λίγες οι πραγματικά αξιόλογες και αξιοποιήσιμες πληροφορίες οι που εξάγονται από αυτές. Η περιορισμένη δευτερεύουσα βιβλιογραφία ήταν το λογικό επακόλουθο αυτής της κατάστασης.

Η Δυτική Λοκρίδα ήταν μια μικρή περιοχή στη σημερινή κεντρική Στερεά Ελλάδα που εκτεινόταν παραλιακά από την περιοχή του Αντιρρίου ως τον Παρνασσό, με ενδοχώρα κυρίως προς την πεδιάδα της Άμφισσας. Συνόρευε στα δυτικά και τα βόρεια με τους Αιτωλούς και στα ανατολικά με τους Φωκείς. Για την διαφαινόμενη ήδη από την αρχαιότητα έλλειψη ενδιαφέροντος για τους Λοκρούς συνέβαλε ένας συνδυασμός από πολλές αιτίες. Οι κυριότερες είναι το “αφιλόξενο” γεωφυσικό τοπίο και η περιορισμένη οικονομική δυνατότητά της περιοχής τους, η “χαλαρή” πολιτειακή συνοχή και η φυλετικού τύπου οργάνωσή των Λοκρών και ο ιδιαίτερα περιορισμένος ρόλος τους στα Πανελλήνια στρατιωτικά γεγονότα (Ξάνθης 2021).

Εξαιτίας των παραπάνω, η ποσότητα των αποσπασμάτων των πηγών που αφορούν τους Λοκρούς είναι εξαιρετικά περιορισμένη. Και ποιοτικά η εικόνα που παρουσιάζουν δεν είναι καλύτερη: τα περισσότερα χωρία αποτελούν συνοπτικές αναφορές σε διάφορα γεωγραφικά στοιχεία της περιοχής ή στο μυθικό της υπόβαθρο (κυρίως μέσω της απλής χρήσης ονομάτων μυθικών προσώπων ή μικρών μυθικών ιστοριών χωρίς αξία για τη συνολική αφήγηση). Αντίθετα, σημαντικά θέματα όπως ζητήματα πολιτειακής οργάνωσης ή εξωτερικής πολιτικής και στρατιωτικής δράσης αποσιωπώνται σχεδόν ολοκληρωτικά, εκτός από την εξαιρετική περίπτωση στην οποία αυτά είχαν κάποιο αντίκτυπο σε γεγονότα ευρύτερου πανελληνίου ενδιαφέροντος. Πολλές φορές συναντάμε ανακρίβειες οι οποίες προέρχονται από την άγνοια, την έλλειψη πληροφόρησης και την αδιαφορία (τόσο του “συγγραφέα” όσο και του “κοινού” στο οποίο απευθυνόταν) για την παρουσίαση της πραγματικότητας. Συχνά οι αρχαίοι κατέφευγαν σε γενικεύσεις ή απλουστεύσεις προκειμένου να καλύψουν αυτή την άγνοια ή απλώς δεν αιτιολογούσαν τα λεγόμενά τους.

Έχοντας αυτή την εικόνα από τις πηγές, εύλογα διαπιστώνονται τα ίδια προβλήματα και στη σύγχρονη βιβλιογραφία. Η γεωγραφία της περιοχής συνδυαστικά με τα αρχαιολογικά της κατάλοιπα απασχολεί κεφάλαια βιβλίων γεωγραφικού προσανατολισμού του ευρύτερου ελληνικού χώρου. Ταυτόχρονα, υπάρχει μια αρκετά πλούσια αρθρογραφία για μεμονωμένα αρχαιολογικά κατάλοιπα, τα οποία συνήθως αποτελούν αναστηλώσεις μνημείων ή προέρχονται από τυχαία ευρήματα σύγχρονων έργων. Για τα υπόλοιπα ζητήματα (πολιτειακά και στρατιωτικά) βρίσκονται μόνο διάσπαρτες, περιστασιακές και νοηματικά φτωχές αναφορές σε συλλογικά έργα ευρύτερης θεματικής. Το μοναδικό έργο το οποίο ασχολείται αποκλειστικά με τη Δυτική Λοκρίδα είναι η διτομή διατριβή του Lucien Lerat, *Les Locriens de l' Ouest (I-II)*. Πρόκειται για ένα θεμελιώδες έργο το οποίο συντάχθηκε με την αξιοποίηση της πλειοψηφίας των διαθέσιμων στο συγγραφέα πηγών (γραμματειακών, επιγραφικών, αρχαιολογικών κ.α.). Η διατριβή δημοσιεύθηκε το 1952, γεγονός προβληματικό για την έρευνα σήμερα, καθώς οι πηγές που έχουν στο μεταξύ αποκαλυφθεί ή επανεξεταστεί έχουν μεταβάλει την εικόνα μας. Πολλές από τις υποθέσεις του πλέον έχουν αποδειχθεί εσφαλμένες και τον είχαν οδηγήσει σε λανθασμένα, ελλιπή ή αυθαίρετα συμπεράσματα, σε απλουστεύσεις και λογικές υπερβάσεις. Επομένως, πρέπει να αναθεωρηθούν.

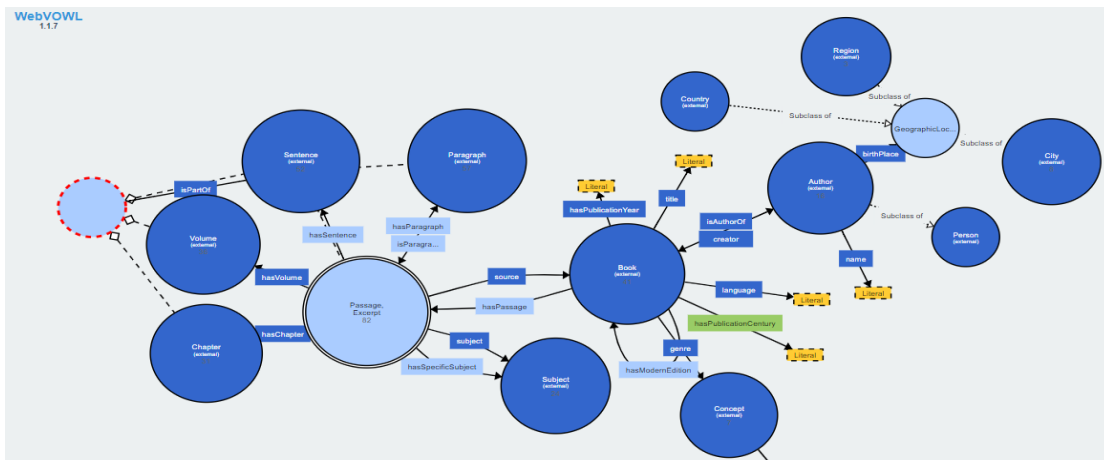
Η τεχνολογία έχει μεταβάλει τον τρόπο και τα μέσα συγκέντρωσης υλικού. Οι ψηφιακές

βιβλιοθήκες είναι βάσεις που διαθέτουν ψηφιοποιημένες συλλογές γραμματειακών πηγών, μαζί με ορισμένα μεταδεδομένα, υποστηρικτικά για την αναζήτηση από τους χρήστες τους. Ένα σπουδαίο εργαλείο, το οποίο συνδέεται και άμεσα με τις αρχές και τις επιδιώξεις του Σηματολογικού Ιστού όπως τους περιγράψαμε, είναι η εφαρμογή ToposText.org (<https://topostext.org/>). Η εφαρμογή αποτελεί μια διαδραστική βάση, με πολλούς τρόπους αναζήτησης του υλικού, με τον σημαντικότερο να επιτυγχάνεται μέσω τοποσημών από ένα χάρτη. Αφού επιλέξει ο χρήστης, στη συνέχεια ανακατευθύνεται σε μια νέα σελίδα η οποία περιλαμβάνει υπό μορφή λίστας τα αποσπάσματα της αρχαίας ελληνικής γραμματείας τα οποία σχετίζονται με αυτό. Τα αποσπάσματα διατίθενται στα Αγγλικά και η αρίθμηση γίνεται με βάση την αρίθμηση των χωρίων από την έντυπη έκδοση που χρησιμοποιήθηκε για την ψηφιοποίηση. Χαρακτηρίζονται από το συγγραφέα, τον τίτλο του έργου και από μια γενική χρονολόγηση, στοιχεία που επιτρέπουν στον ερευνητή να επιλέξει το υλικό που θα χρησιμοποιήσει πριν ασχοληθεί καθαυτό με το περιεχόμενό του. Η εφαρμογή πάντως αντιμετωπίζει ορισμένα προβλήματα: σε γενικότερο επίπεδο θα ήταν ωφέλιμο να αναθεωρηθεί η βάση της ώστε να περιλαμβάνει ακόμα περισσότερα αποσπάσματα, ενώ επίσης θα μπορούσε να εμπλουτιστεί και με άλλα είδη πηγών, ίσως χρησιμοποιώντας συνδέσεις με άλλες βάσεις. Αναφορικά με την περίπτωση των Λοκρών που εξετάζουμε, εκτός των βελτιώσεων που αναφέρθηκαν, κρίνεται ακόμη αναγκαία η επικαιροποίηση των τοπογραφικών δεδομένων.

Όλη αυτή η κατάσταση συνολικά δημιουργεί προβλήματα στην έρευνα και οι μελετητές καλούνται να τα ξεπεράσουν. Όπως επισημάνθηκε σε προηγούμενο κεφάλαιο αναφορικά με τον τρόπο που ο ιστορικός αντιμετωπίζει τις πηγές του (βλ. στο 1.1.2), η προσεκτική μελέτη των πηγών οδηγεί σε ορθότερα και πληρέστερα συμπεράσματα. Επομένως, η ταχύτερη και ευκολότερη πρόσβαση σε αυτές είναι καθοριστικής σημασίας και η σωστή οργάνωση και αποθήκευσή τους κρίνεται αναγκαία. Για να καλύψω αυτές τις προϋποθέσεις, παράλληλα με τη συγκέντρωση του υλικού, προχώρησα στη δημιουργία ενός αρχείου Excel στο οποίο κατέγραφα τα βασικά μεταδεδομένα των αποσπασμάτων που έβρισκα. Οι πρώτες στήλες του αρχείου που δημιουργήθηκε (“Απόσπασμα”, “Συγγραφέας”, “Έργο”, “Γενικό Θέμα”) αποτέλεσαν τον αρχικό πυρήνα των κλάσεων για την οντολογία που κατασκεύασα αργότερα. Όσο η διαδικασία συνεχιζόταν τόσο το αρχείο εμπλουτιζόταν με νέες γραμμές αποσπασμάτων και επιπλέον στήλες (μετα)δεδομένων. Για τις παρούσες ανάγκες αυτό το σύστημα φαινόταν αρκετά αποτελεσματικό. Όμως με έκανε να συνειδητοποιήσω τη σημασία της σωστής και αποτελεσματικής καταγραφής των πηγών και το ρόλο της για την διεξαγωγή ιστορικής έρευνας. Σε επόμενο στάδιο έθεσα το ερώτημα, το πως θα μπορούσε η δουλειά μου (το αρχείο που είχα φτιάξει) να χρησιμοποιηθεί από την τεχνολογία – και να ωφεληθεί από αυτήν- ώστε να καλυφθεί το επιστημονικό κενό που είχα κληθεί να αντιμετωπίσω. Η απάντηση ήρθε το 2023 στο πλαίσιο του μαθήματος “*Παγκόσμιος Ιστός και Διαχείριση Πολιτιστικών Συλλογών*” στο Π.Μ.Σ. “*Ψηφιακός Πολιτισμός, Έξυπνες Πόλεις, IoT και Προηγμένες Ψηφιακές Τεχνολογίες*” του Πανεπιστημίου Πειραιώς. Εκεί έγινε λόγος για τη συστηματική οργάνωση και διασύνδεση των δεδομένων, το σχηματισμό “τριπλετών” και την ανάπτυξη οντολογιών, μέσω των οποίων αναπαρίστανται οι πληροφορίες. Η οντολογία είναι το μέσο για την αξιοποίηση του αρχείου μου και αυτό το δεδομένο διαμόρφωσε τη θεματική της παρούσας εργασίας. Στο πρακτικό της κομμάτι βασίζεται στην πρόταση, τη δημιουργία και το σχεδιασμό ενός ευέλικτου προτύπου για την καταγραφή πηγών που θα καλύπτει την ανάγκη για ταχύτητα, ορθότερη διασύνδεση και αποδοτική αρχειοθέτησή τους προς όφελος της ερευνητικής διαδικασίας. Το θεωρητικό κομμάτι περιλαμβάνει σχολιασμό του πλαισίου στο οποίο χτίστηκε η οντολογία και η τεκμηρίωση των όρων της. Η έρευνα για την ιδιαίτερη μελέτη περίπτωσης, υπήρξε η αφορμή, αλλά η επιλογή αυτού του παραδείγματος δεν είναι μόνο μια προσωπική επιδίωξη, αλλά έχει πολλαπλή αξία για την εργασία. Λειτουργεί ως παράδειγμα απόδειξης της λειτουργικότητας του προτεινόμενου μοντέλου και της πρακτικής εφαρμογής του. Ανταποκρίνεται έτσι και σε μια βασική προϋπόθεση για την ανάπτυξη οντολογιών, την “*παροχή λύσης με πρακτικό τρόπο για ένα υπαρκτό ζήτημα*”. Η προϋπόθεση βρίσκει σημαντική εφαρμογή για τη συλλογή και την ευέλικτη αρχειοθέτηση των γραμματειακών πηγών για τη Λοκρίδα καθώς είναι ένα ιδιαίτερο πεδίο το οποίο δεν έχει απασχολήσει επαρκώς την έρευνα.

β. Τεκμηρίωση της μελέτης περίπτωσης

Σε προηγούμενο κεφάλαιο (2.1.β.) παρουσιάστηκε το οντολογικό πρότυπο για μια οντολογία καταγραφής αρχαίων πηγών. Περιγράψαμε τη δομή της και τη λειτουργία των κύριων στοιχείων της. Ωστόσο είναι απαραίτητη η εφαρμογή πραγματικών δεδομένων στην οντολογία για να διαπιστωθούν η λειτουργικότητα και η χρησιμότητα. Σε αυτό το κεφάλαιο θα εξετάσουμε τη λειτουργία της οντολογίας μέσω των παραδειγμάτων της μελέτης περίπτωσης που επιλέχθηκε και αφορά τις αρχαίες γραμματειακές πηγές της δυτικής Λοκρίδας²⁷. Παραθέτοντας συγκεκριμένα στιγμιότυπα, θα δείξουμε πως χρησιμοποιούνται και διασυνδέονται οι κλάσεις και οι ιδιότητες του προτύπου. Όπως και για την τεκμηρίωση του προτύπου, θα χρησιμοποιήσουμε επικουρικά της ανάλυσης τις εικόνες αναπαράστασης από το Protégé. Ταυτόχρονα, θα βοηθήσουμε την οπτικοποίηση της οντολογίας με τη χρήση του WebVOWL²⁸, εργαλείο το οποίο προσφέρει μια τυπική διαγραμματική αναπαράστασή. Το παρακάτω διάγραμμα (Εικόνα 12) αποτυπώνει τη συνολική δομή της οντολογίας όπως δημιουργήθηκε στο Protégé και απεικονίζεται στο WebVOWL.



Εικόνα 11: οπτική αναπαράσταση της οντολογίας- στιγμιότυπο οθόνης από το WebVOWL

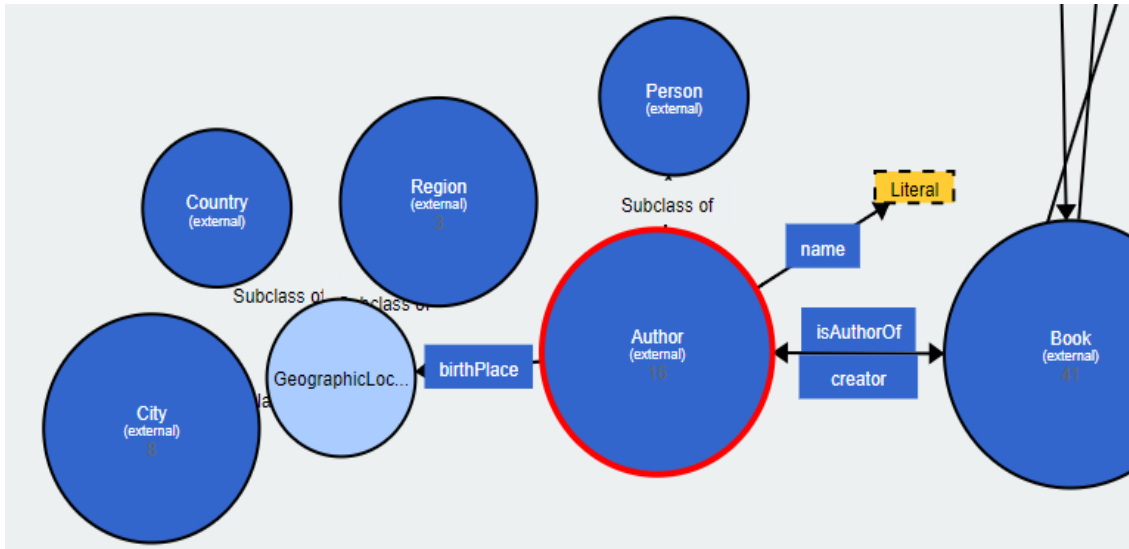
Κάθε κλάση έχει στιγμιότυπα τα ιδιαίτερα δεδομένα της οντολογίας. Στη περίπτωση που εξετάζουμε, δημιουργήσαμε, για παράδειγμα, την κλάση “dbpedia:Author” (“Συγγραφέας”) και προσθέσαμε σε αυτήν στιγμιότυπα. Συγκεκριμένα για την κλάση “dbpedia:Author” καταγράψαμε

Εικόνα 12: η κλάση “dbpedia:Author” και τα στιγμιότυπα της- εικόνα στιγμιότυπου οθόνης από το Protégé

²⁷ Η οντολογία είναι διαθέσιμη στον παρακάτω σύνδεσμο: <https://drive.google.com/file/d/1q36CI0ixJbbJvEiWByMomEcit-3GbfAk/view?usp=sharing>

²⁸ Το WebVOWL είναι εργαλείο απλής οπτικής απεικόνισης των οντολογιών. Πρόκειται για εξελικτικό στάδιο του σχεδιαγράμματος που παρουσιάσαμε στο κεφάλαιο 2.1β.

16 στιγμιότυπα, καθώς τόσο συγγραφείς προκύπτουν από τα αποσπάσματα. Κάθε στιγμιότυπο είναι μια οντότητα (individual) που χρησιμοποιείται για να συγκεντρώσει όλα τα στοιχεία που την αφορούν, μέσω των ιδιοτήτων και των σχέσεων. Εστιάζοντας στο σχεδιάγραμμα της Εικόνας 12 (βλ. Εικόνα 14), επικεντρωνόμαστε στην κλάση του Συγγραφέα. Βλέπουμε να τοποθετείται ως υποκλάση της “foaf:Person” και να ενώνεται με την κλάση του “Βιβλίου” με την ιδιότητα “cd:isAuthorOf” και με τον “Τόπο Καταγωγής” μέσω της ιδιότητας “dbo:birthPlace” και να έχει ένα name, ένα όνομα, μία string literal τιμή σε φυσική γλώσσα.



Εικόνα 13: στιγμιότυπο οθόνης από το WebVOWL- εστίαση στην κλάση “dbpedia:Author”

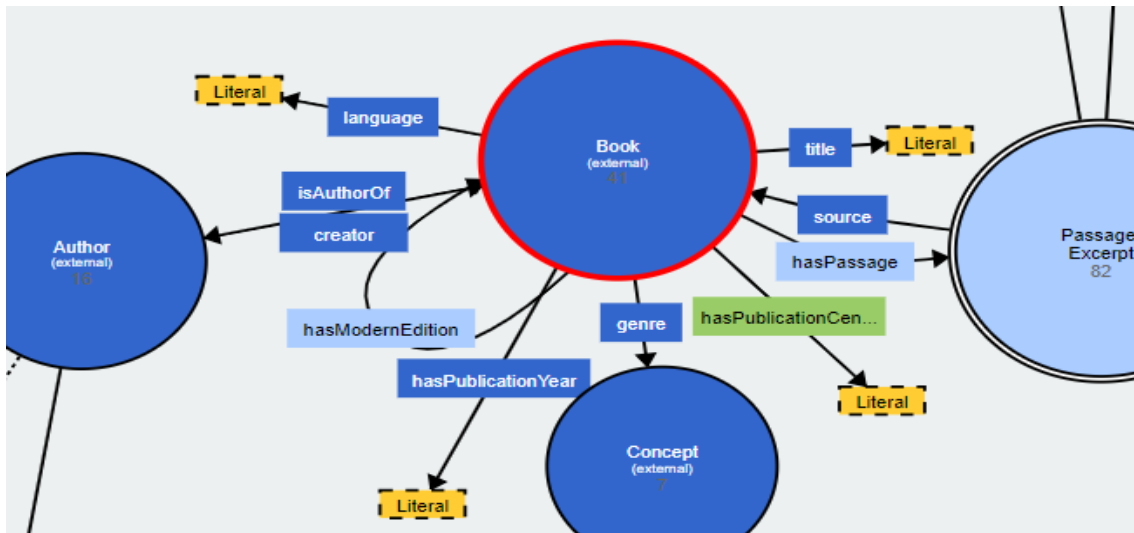
Την παραπάνω σχηματική αναπαράσταση (Εικόνα 13) μπορούμε να αντιληφθούμε καλύτερα με τη χρήση ενός παραδείγματος από το Protégé. Ενδεικτικά, παίρνουμε το στιγμιότυπο της κλάσης “dbpedia:Author”, “sao:Aeschines”:

Εικόνα 14: περιγραφή της οντότητας “sao:Aeschines”, κλάσης “dbpedia:Author”- εικόνα στιγμιότυπου οθόνης από Protégé

Βλέπουμε στην περιγραφή (Description) πως το στιγμιότυπο ανήκει σωστά στην κλάση “dbpedia:Author” (υποκλάση της “foaf:Person”). Στις “object properties”, ο “Συγγραφέας” συνδέθηκε με τα έργα του με την ιδιότητα “cd:isAuthorOf”. Επιλέγοντας το παράδειγμα του “Aeschines”, εμφανίζονται καταγεγραμμένα όλα τα βιβλία του (που έχουμε καταχωρήσει στην κλάση “bibo:Book”), εδώ οι οντότητες “sao:PeriTisParapresveias” και “sao:KataKtisiphontos”. Δηλώνεται ακόμη μέσω της ιδιότητας “dbo:birthPlace” η σύνδεσή του με έναν “Τόπο Καταγωγής”, εδώ με το στιγμιότυπο “sao:Athens” της κλάσης “dbo:City”. Στις “datatype properties”, χρησιμοποιεί

μόνο τον προσδιορισμό του ονόματός "foaf:name" γραμμένο σε φυσική γλώσσα -τα Ελληνικά-, "Αισχίνης".

Εστιάζοντας εκ νέου στο σχεδιάγραμμα του WebVOWL (Εικόνα 11), η επόμενη μεγάλη κλάση που διακρίνουμε είναι αυτή των βιβλίων, "bibo:Book" (Εικόνα 15). Συνδέεται με τις άλλες δύο μεγάλες κλάσεις της οντολογίας ("dbpedia:Author", "sao:Passage"). Αρχικά με την αντίστροφη ιδιότητα "dcterms:creator" δηλώνεται πως το κάθε Βιβλίο "έχει συγγραφέα". Με την ιδιότητα "lod:genre" δηλώνεται το γραμματειακό "Είδος" στο οποίο ανήκει. Η κλάση του "Βιβλίου" συνδέεται με την κύρια κλάση της οντολογίας, το "sao:Passage" και έτσι κάθε στιγμιότυπο της συνδέεται με όλα τα χωρία τα οποία βρίσκονται σε αυτό. Με την ιδιότητα "sao:hasModernEdition" συνδέεται το αρχαίο κείμενο με την σύγχρονη έκδοσή του. Η κλάση έχει και τις τέσσερις διαθέσιμες "data properties", δύο "integer" τιμές ("sao:hasPublicationCentury", "fabio:hasPublicationYear") και δύο "string" τιμές ("dcterms:language", "dcterms:title").



Εικόνα 15: στιγμιότυπο οθόνης από το WebVOWL- εστίαση στην κλάση "bibo:Book"

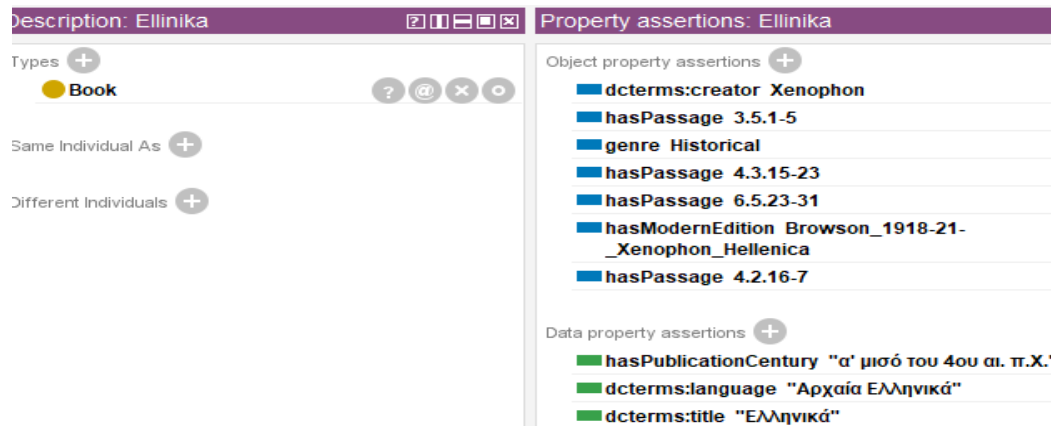
Παρατηρούμε πως υπάρχει η ιδιότητα "sao:hasModernEdition", της οποίας το βέλος ξεκινά και καταλήγει στην κλάση "bibo:Book", έχει δηλαδή ίδιο "Domain" και "Range". Αυτό συμβαίνει γιατί στην κλάση "bibo:Book" περιέχονται οντότητες τόσο από αρχαία κείμενα όσο και από σύγχρονα έργα, όπως επιβεβαιώνεται με την Εικόνα 16, επιλέγοντας να δούμε instances από την

The screenshot shows the Protégé interface. On the left, the 'Class hierarchy: Book' is expanded, showing 'Book' as a subclass of 'owl:Thing'. On the right, the 'Annotations' section shows 'rdfs:label [language: el] Βιβλίο' and 'rdfs:label [language: en] Book'. Below that, the 'Description: Book' section is visible. At the bottom, the 'Instances' section lists 14 instances, including 'Jones_1917-1932-The_Geography_of_Strabo_vols._II_V', 'Adams_1919-The_Speeches_Of_Aeschines', 'Agisilaos', 'Babbitt_1927-1969-Plutarch's_Moralia_vols._II-VII', 'Bekker_1970-1985-Diodori_Bibliotheca_Historica_Vols.I-VI', 'Billerbeck_2006-Stephani_Byzantij_Ethnica_Vol._I-V', 'Browson_1918-21_Xenophon_Hellenica', 'Chambers_1993-Hellenica_Oxyrhynchia', 'ElladosPeriigisis', 'Ellinika', and 'EllinikaTisOksirinchou'.

Εικόνα 16: η κλάση "bibo:Book" και τα στιγμιότυπά της- εικόνα στιγμιότυπου οθόνης από το Protégé

κλάση στο Protégé (λ.χ. "sao:ElladosPeriigisis" κ.α., αλλά και "sao:Jones_1927-1932-The_Geography_of_Strabo_vols_II_V" ...).

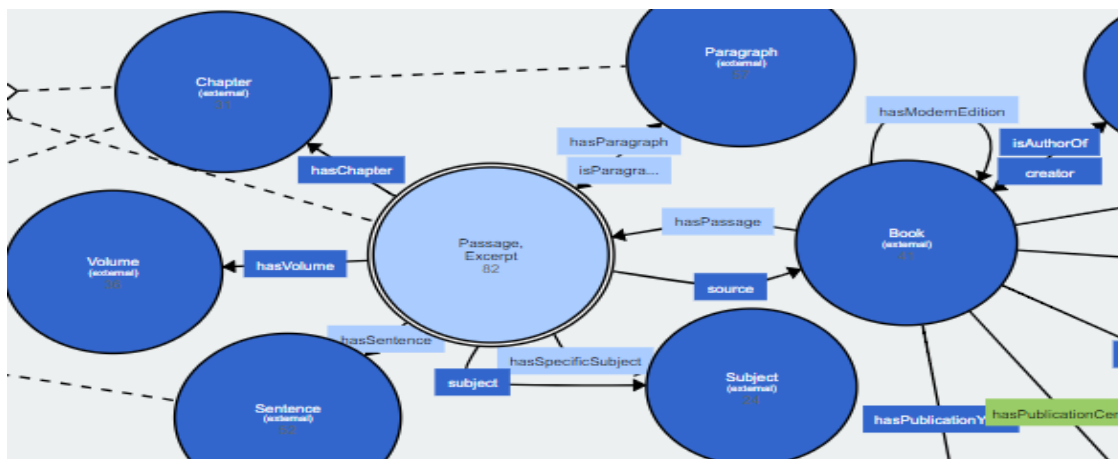
Επιλέγοντας ένα στιγμιότυπο από τα 41 διαθέσιμα για την κλάση "bibo:Book", για παράδειγμα το βιβλίο "sao:Ellinika", ανακατευθυνόμαστε στην καρτέλα της συγκεκριμένης οντότητας (Εικόνα 17):



Εικόνα 17: περιγραφή της οντότητας "sao:Ellinika", κλάσης "bibo:Book"- εικόνα στιγμιότυπου οθόνης στο Protégé

Στο "Description" διαπιστώνεται η σωστή λειτουργία της οντολογίας, καθώς το στιγμιότυπο περιλαμβάνεται στην κλάση του, την "bibo:Book". Επομένως, οι ιδιότητες που μπορεί να έχει είναι αυτές που έχουν ορισμένο ως "Domain" τους την κλάση του Βιβλίου (βλ. Εικόνα 8 και Εικόνα 10). Όσον αφορά τις "object properties", διαβάζουμε πως το βιβλίο έχει συγγραφέα ("dcterms:creator" "sao:Xenophon") και ανήκει στο "Είδος" του "Ιστορικού Βιβλίου" (lod:genre "sao:Historical"). Επίσης, είναι σημειωμένη και μια σύγχρονη έκδοσή του, η οποία χρησιμοποιήθηκε για να βρούμε τα αποσπάσματα ("sao:hasModernEdition" "sao:Browson_1918-21-Xenophon_Hellenica"). Τέλος, βλέπουμε καταγεγραμμένα όλα τα αποσπάσματα του βιβλίου που βρίσκονται στη βάση μας ("sao:hasPassage": "sao:3.5.1-5", "sao:4.3.15-23", "sao:6.5.23-31" και "sao:4.2.16-7"). Το συγκεκριμένο παράδειγμα διαθέτει και τις τρεις δυνατές "data properties" με τις οποίες μπορούν να χαρακτηρίζονται μοναδικά οι οντότητες από την κλάση του Βιβλίου. Αρχικά σημειώνεται ο πραγματικός τίτλος του έργου ("dcterms:title" "Ελληνικά"). Διευκρινίζεται η γλώσσα γραφής του πρωτότυπου ("dcterms:language" "sao:Αρχαία Ελληνικά"). Τέλος ορίζεται η χρονολογία του έργου ("sao:hasPublicationCentury" "α' μισό του 4ου αιώνα π.Χ."). Με τη χρήση του παραδείγματος αποδεικνύεται πρακτικά η χρησιμότητα και η αναγκαιότητα των ιδιοτήτων που σχεδιάστηκαν στο πρότυπο (κεφ. 2.1.β), προκειμένου να παρέχονται οργανωμένα, εύκολα και άμεσα τα χρησιμότερα δεδομένα για τους ερευνητές, προς διευκόλυνση της έρευνάς τους.

Η κύρια κλάση και η βασική έννοια γύρω από την οποία χτίστηκε, είναι το "sao:Passage" ("χωρίο"). Αντιπροσωπεύει την αριθμηση η οποία είναι απαραίτητη στον ερευνητή προκειμένου να εντοπίζει (ή και να παραπέμπει άλλους) ένα ακριβές απόσπασμα που περιέχει τις πληροφορίες



Εικόνα 18: στιγμιότυπο οθόνης από το WebVOWL- εστίαση στην κλάση "sao:Passage"

Η Καταγραφή αρχαίων πηγών με τη χρήση Οντολογιών:

Η περίπτωση της Δυτικής Λοκρίδας

που θέλει. Στο παράδειγμά μας, βλέπουμε στο σχεδιάγραμμα της WebVOWL (Εικόνα 18) πως υπάρχουν 82 στιγμιότυπα “Αποσπασμάτων” τα οποία αντιστοιχούν στα “Βιβλία” από τα οποία προέρχονται μέσω της ιδιότητας “dcterms: source”²⁹ (“βρίσκεται στο ...”). Κάθε απόσπασμα πραγματεύεται ένα θέμα, το οποίο μπορεί να είναι αρκετά γενικό (λ.χ. ένας πόλεμος) μέχρι πολύ ειδικό (λ.χ. μια συγκεκριμένη μάχη του). Αυτή τη σύνδεση μεταξύ των κλάσεων του “Χωρίου” (“sao:Passage”) και του “Θέματος” (“getty:Subject”) γίνεται μέσω της ιδιότητας “dcterms:subject” και της υπο-ιδιότητάς της “sao:hasSpecificSubject”. Σύμφωνα με το σχεδιάγραμμα (Εικόνα 19) το “sao:Passage” χωρίζεται στις κλάσεις “tribont:Volume” (Τόμος), “bibo:Chapter” (Κεφάλαιο), “doco:Paragraph” (Παράγραφος) και “doco:Sentence” (Πρόταση) οι οποίες συνδέονται τόσο με το “sao:Passage”, όσο και κατά συνθήκη και μεταξύ τους (με τις ιδιότητες “sao:isParagraphOf” και “dcterms:isPartOf”). Αυτές είναι ξεχωριστές κλάσεις και όχι υποκλάσεις, καθώς όπως έχει εξηγηθεί (βλ. στο), κάτι τέτοιο θα συνεπαγόταν υποχρεωτικά πως όλα τα αποσπάσματα θα έπρεπε να έχουν αυτόν τον διαχωρισμό, κάτι που πρακτικά, για διάφορους λόγους³⁰, δεν υφίσταται³¹.

Στο παρακάτω παράδειγμα έχουμε επιλέξει ένα απόσπασμα “3.100-102.5”, σημαντικό για τις πληροφορίες που παρέχει για τη Λοκρίδα, αλλά και χρήσιμο προκειμένου να εξηγήσουμε τις σχέσεις του “Χωρίου” με τις άλλες κλάσεις:

The screenshot shows the Protégé interface for the class "3.100-102.5". The left pane, titled "Description: 3.100-102.5", shows the class is of type "Passage". The right pane, titled "Property assertions: 3.100-102.5", lists several object property assertions:

- has Specific Subject EkstrateiaDimostheni
- has Sentence 5
- dcterms:source Istoría
- has Volume 3
- has Paragraph 100-2
- dcterms:subject PeloponnisiakosPolemos

Εικόνα 19: περιγραφή οντότητας “3.100-102.5” κλάσης “sao:Passage”- εικόνα στιγμιότυπου οθόνης Protégé

Στην αποτύπωση της περιγραφής του στιγμιότυπου από το Protégé (Εικόνα 19) επιβεβαιώνουμε στην καρτέλα “Description” ότι ανήκει στην κατηγορία “sao:Passage”. Στην καρτέλα “property assertions” βλέπουμε τις διαθέσιμες “object properties” που περιγράψαμε. Με την ιδιότητα “dcterms: source” (“βρίσκεται στο...”) μαθαίνουμε ότι ανήκει στην οντότητα βιβλίου “sao:Istoria”. Συνδέεται με 2 στιγμιότυπα της κλάσης “getty:Subject”, μέσω και των δύο πιθανών ιδιοτήτων της, “dcterms:subject” (οντότητα “sao: Peloponnisiakos Polemos”) και της υποκλάσης της “sao:hasSpecificSubject” (οντότητα “sao:EkstrateiaDimostheni”). Η σύνδεση με τα μέρη του έργου τα οποία καταγράφει, γίνεται μέσω των ιδιοτήτων που ως “Range” έχουν τις αντίστοιχες κλάσεις (“tribont:Volume”, “doco:Paragraph” και “doco:Sentence” στη συγκεκριμένη περίπτωση). Με την ιδιότητα “arpenenteur:hasVolume” ορίζεται ότι ο “Τόμος” για την αναζήτηση του χωρίου είναι ο “3ος” και με την “sao:hasParagraph” εστιάζουμε στις παραγράφους “100-2”. Η “sao:hasSentence” μας δείχνει μέχρι “που φτάνει” το απόσπασμα, εδώ στην πρόταση “5”.

Οι 3 κλάσεις (“dbpedia:Author”, “bibo:Book” και “sao:Passage”) είναι οι σημαντικότερες στη διαδικασία αρχειοθέτησης αρχαίων πηγών. Γι’ αυτό και χρησιμοποιήθηκαν παραδείγματα από τη μελέτη περίπτωσης για να αναδειχθεί ο ρόλος τους και να παρουσιαστεί η λειτουργία τους στο οντολογικό πλαίσιο. Για την περιγραφή τους χρειάστηκε η αναφορά των υπόλοιπων κλάσεων με τις οποίες συνδέονται μέσω συγκεκριμένων ιδιοτήτων. Επομένως, δε κρίθηκε σκόπιμο να αναλυθεί κάθε μια ξεχωριστά. Είναι όμως σημαντικό να σημειωθούν ορισμένες διευκρινήσεις, αναφορικά με τις ιδιαιτερότητες λειτουργίας και περιεχομένου της περίπτωσης για κάποιες από αυτές.

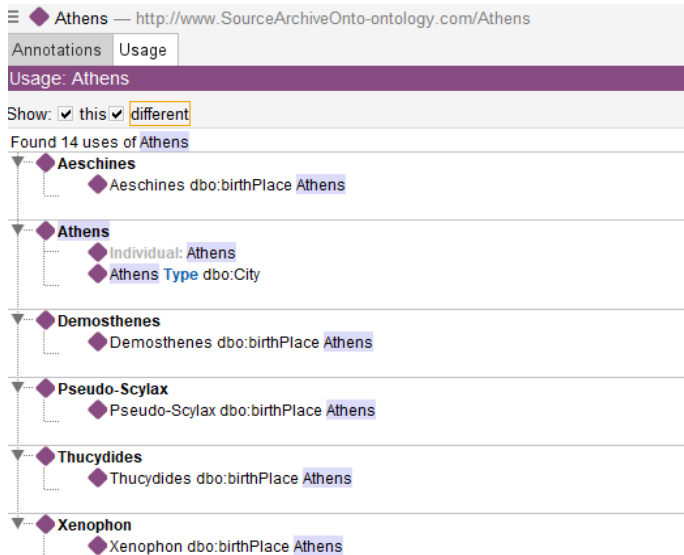
Για την κλάση “vino:GeographicLocation”, τον “τόπο καταγωγής”, να θυμίσουμε πως είναι

²⁹ Η αντίστροφη της είναι η “sao:hasPassage”, “έχει απόσπασμα”, η οποία έχει αντίστοιχα “Domain” “bibo:Book” και “Range” “sao:Passage”.

³⁰ Βλ. στο παρόν, σ.σ. 27-8.

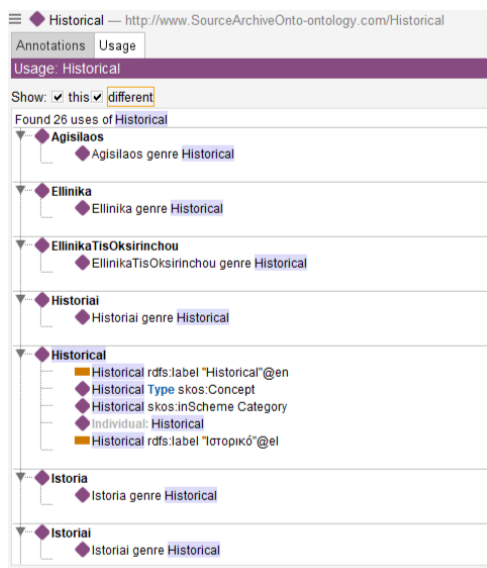
³¹ Στη μελέτη περίπτωσης καταγράψαμε 82 αποσπάσματα που υπάγονται σε 36 Τόμους, 31 Κεφάλαια, 57 Παραγράφους και 52 Πρότασεις. Όμως οι συνδυασμοί τους είναι πολλοί. Για παράδειγμα, από τους 36 τόμους οι 30 χωρίζονται σε Κεφάλαια, από τα 31 συνολικά Κεφάλαια, μόνο τα 15 χωρίζονται σε Παραγράφους. Τα 14 που έχουν Παράγραφο, έχουν επίσης και Κεφάλαιο και Τόμο, κ.ο.κ..

υπερκλάση που περιλαμβάνει τις κλάσεις “juso:Country”, “dbo:Region” και “dbo:City”. Για τη μελέτη περίπτωσης επιλέγεται αυτή η ταξινόμηση και όχι η ιεραρχία, καθώς αφενός γιατί δεν είναι απαραίτητη η εξειδίκευση για τον “Συγγραφέα” ενός κειμένου. Αφετέρου γιατί στην έρευνα των αρχαίων συγγραφέων δεν είναι πάντα γνωστή με ακρίβεια η πόλη ή αν είναι δεν έχει σημασία η περιοχή, ενώ δεν υπήρχαν κράτη με τη σύγχρονη έννοια. Γι’ αυτό στη μελέτη μας δεν υπάρχει στιγμιότυπο για την κλάση “juso:Country”. Αντίθετα έχουμε 3 “Περιοχές” (“dbo:Region”) και 8 “Πόλεις” (“dbo:City”). Με τα εργαλεία του Protégé μπορούμε να δούμε με την καρτέλα “usage” το που και πως έχει χρησιμοποιηθεί το κάθε στιγμιότυπο. Αυτό εξυπηρετεί τις γρήγορες αναζητήσεις πληροφοριών, όπως για παράδειγμα τους “Συγγραφείς” που έχουμε ορίσει την καταγωγή τους από την πόλη “sao:Athens”, μέσω της ιδιότητας “dbo:birthPlace”(Εικόνα 20)³².



Εικόνα 20: η καρτέλα “Usage” του στιγμιότυπου “sao:Athens”- εικόνα στιγμιότυπου οθόνης από το Protégé

Η κλάση “Concept” αναφέρεται στο λογοτεχνικό “Είδος” των Βιβλίων σύμφωνα με τη σύγχρονη κατάταξή τους. Στην περίπτωση μας αυτά ομαδοποιήθηκαν σε 7 κατηγορίες, τις: “sao:Geographical Lexicon” (Γεωγραφικό Λεξικό), “sao:Symposium” (Συμπόσιο),



Εικόνα 21: η καρτέλα “Usage” του στιγμιότυπου “Historical”- εικόνα στιγμιότυπου οθόνης από το Protégé

³² Αυτή είναι και η λογική των SPARQL ερωτημάτων και των ανακλήσεων τους που θα δούμε στο 2.2.γ.

“sao:Geographical” (Γεωγραφία), “sao:Biography” (Βιογραφία), “sao:Rhetoric” (Ρητορική), “sao:Historical” (Ιστορία) και “sao:Mythology” (Μυθολογία). Τα στιγμιότυπα της κλάσης χρησιμοποιούνται με την ίδια λογική αναφέραμε προηγουμένως (Εικόνα 22).

Πρέπει να διευκρινιστεί πως η κλάση “getty:Subject” έχει 24 στιγμιότυπα. Από αυτά, τα 11 χρησιμοποιούνται ως “θέματα” των αποσπασμάτων με την ιδιότητα “dcterms:subject” και τα 16 ως “ιδιαίτερα θέματα” μέσω της υποιδιότητάς της “sao:hasSpecificSubject”. Υπάρχουν και 4 στιγμιότυπα τα οποία χρησιμοποιούνται και από τις 2 ιδιότητες κατά περίπτωση. Ο διαχωρισμός γίνεται μετά από την ερμηνεία του περιεχομένου του αποσπάσματος. Το να θεωρηθεί ένα θέμα γενικό ή ειδικό, υπόκειται στην κρίση του κατασκευαστή και στον προσανατολισμό της οντολογίας συνολικά.

Τέλος, αναφορικά με τα “μέρη” των αποσπασμάτων [“tribont:Volume” (Τόμος), “bibo:Chapter” (Κεφάλαιο), “doco:Paragraph” (Παράγραφος) και “doco:Sentence” (Πρόταση)]. Η αριθμητική αποτύπωση κάθε αποσπάσματος εξυπηρετεί στην ανάγνωσή του. Για να την κατανοήσουμε, διαβάζεται “διασπώντας” τους όρους του. Στο παράδειγμα που είδαμε, η πληροφορία “Ιστορία 3.100-2.5” σημαίνει πως μπορούμε να βρούμε το απόσπασμα στη σειρά Ιστορία, συγκεκριμένα στον 3^ο τόμο της, στην παράγραφο 100-102. Το απόσπασμα που μας ενδιαφέρει ολοκληρώνεται στην σειρά 5 της παραγράφου 102. Η επιλογή ώστε να αποτελούν τα μέρη του “sao:Passage” ξεχωριστές κλάσεις προσβλέπει στην λογική και λειτουργική ομαδοποίηση των αποσπασμάτων αυτών. Με αυτό τον τρόπο προσφέρεται μεγαλύτερη ακρίβεια στην διαχείριση των αποσπασμάτων και στην ανάκτηση συγκεκριμένων τμημάτων ενός έργου (για παράδειγμα υπάρχουν 4 αποσπάσματα από τον Τόμο 3 του Βιβλίου “Ιστορία” του Θουκυδίδη). Επιπλέον, αυτή η τακτική εξυπηρετεί τη μελλοντική ανάπτυξη της οντολογίας, επιτρέποντας την προσθήκη των ίδιων των αρχαίων κειμένων.

γ. SPARQL Ερωτήματα

SPARQL (SPARQL Protocol and RDF Query Language) ονομάζουμε την επίσημη, τυποποιημένη γλώσσα ερωτημάτων και αναζήτησης, η οποία σχεδιάστηκε για την εκμετάλλευση (αναζήτηση, ανάκτηση και διαχείριση) των δεδομένων που είναι αποθηκευμένα σε μορφή RDF. Τα ερωτήματα αυτά βασίζονται σε μοτίβα τριπλετών (υποκείμενο- κατηγορήμα- αντικείμενο) και μεταβλητές. Μπορούν να εκτελεστούν σε οποιοδήποτε σύστημα του Σημασιολογικού Ιστού που υποστηρίζει την RDF και να ενσωματώνονται σε πλατφόρμες, εργαλεία και προγράμματα ανοιχτού κώδικα τα οποία προσφέρουν κατάλληλο γραφικό περιβάλλον για την ανάπτυξη και εκτέλεσή τους. Επιτρέπουν στους χρήστες να εκφράσουν μια μεγάλη ποικιλία από άποψη εύρους και συνθετότητας ερωτημάτων πάνω στα δομημένα δεδομένα μιας οντολογίας και να ανακτήσουν ακριβείς και συγκεκριμένες πληροφορίες μέσω συσχετισμού των στοιχείων και φιλτραρίσματος (Χρυσούλας 2009, σ. 91). Οι ερωτήσεις που είναι σε θέση να απαντήσει μια οντολογικά δομημένη Βάση προσδιορίζουν το εύρος της και έτσι ο δημιουργός της οντολογίας έχει τη δυνατότητα να την εξελίξει περαιτέρω. Αν οι οντολογίες δημιουργούνται για την οργάνωση και διασύνδεση της γνώσης, τα SPARQL ερωτήματα προσφέρουν ένα μέσο απόδειξης των δυνατοτήτων αξιοποίησής τους οντολογιών, αποδεικνύοντας την λειτουργικότητά και την πρακτική αξία τους.

Συμπεριέλαβα τα SPARQL ερωτήματα στην εργασία μου για να επιβεβαιώσω τη λειτουργικότητα³³της οντολογίας. Καθώς μπορούν να τεθούν ακριβή ερωτήματα και να ανακτηθούν αξιόπιστα αποτελέσματα, αποδεικνύεται η ορθή δόμηση της οντολογίας η οποία «επιτρέπει» στα δεδομένα της να είναι διαχειρίσιμα και να αναλύονται. Έπειτα για να αναδειχθεί η πρακτική ερευνητική της εφαρμογή. Διατυπώνοντας παραδείγματα ερωτήσεων παρουσιάζονται οι δυνατότητες της οντολογίας στην εξαγωγή ποικίλων συμπερασμάτων για την έρευνα και καταδεικνύονται οι περιορισμοί της. Μέσω της αναπαραγωγιμότητας³⁴ άλλοι ερευνητές μπορούν να επιβεβαιώσουν τα αποτελέσματα, να λάβουν απαντήσεις στα δικά τους ερωτήματα, να διαπιστώσουν τους περιορισμούς και να προχωρήσουν σε περαιτέρω ανάπτυξη της οντολογίας.

Υπάρχουν συνολικά τέσσερα είδη SPARQL ερωτημάτων που μπορούν να χρησιμοποιηθούν (Ιακωβάκης 2019, σ.σ. 16-7; Μερτής 2010, σ.σ. 29, 33-4; Τριανταφύλλου 2022, σ.σ. 28-34):

1. "SELECT" Βασίζεται στις τριπλέτες και "ζητά" από τον επεξεργαστή να ανακτήσει όλες τις τριπλέτες που ταιριάζουν σε το ζητούμενο και να παρουσιάσει τις τιμές τους.
2. "CONSTRUCT" Στην ίδια λογική με το "SELECT" όμως δεν επιστρέφει ιδιαίτερες τριπλέτες, αλλά γράφους.
3. "DESCRIBE" Επιστρέφει την περιγραφή όλων των πληροφοριών για κάποιο συγκεκριμένο πόρο
4. "ASK" Περιορίζει το φάσμα των επιτρεπόμενων ερωτημάτων διευκολύνοντας τον επεξεργαστή να αντλήσει τα δεδομένα από μικρότερο δυνατό εύρος

Στην περίπτωση μας θα ασχοληθούμε με το ερώτημα "SELECT". Το αποτέλεσμα που επιστρέφεται εξαρτάται από τις τριπλέτες που θα ορίσουμε σε αυτό. Ο επεξεργαστής θα βρεί τις τριπλέτες που ταιριάζουν στο ερώτημα και θα παρουσιάσει τις τιμές που "πληρούν" τα κριτήρια. Για τον ορισμό της πρότασης αποτελεσμάτων χρησιμοποιούμε τη λέξη-κλειδί "WHERE". Η απάντηση μπορεί να προσδιοριστεί ακόμα περισσότερο χρησιμοποιώντας μετατροπείς λύσεων ερωτήματος όπως η ταξινόμηση, η μοναδικότητα, το όριο, ο περιορισμός των διπλότυπων ("ORDER", "SELECT", "DISTINCT", "REDUCED", "LIMIT"). Ένα ερώτημα μπορεί να περιέχει λογικούς τελεστές σύζευξης ή διάζευξης, προαιρετικούς περιορισμούς (χρόνος, πλήθος, είδος αποτελεσμάτων κ.α.) που συγκεκριμενοποιούν ακόμη περισσότερο το αποτέλεσμα της ανάκτησης (Bikakis et al. 2013, σ.σ. 11-2; Κουτσιώρας 2018, σ.σ. 29-30; Μερτής 2010, σ.σ. 31-34; Τριανταφύλλου 2022, σ. 29; <https://www.w3.org/TR/rdf-sparql-query/>).

³³ Η λειτουργικότητα ενός εργαλείου αφορά την ικανότητά να εκτελεί τις λειτουργίες του με ακρίβεια και αποδοτικότητα. Η λειτουργικότητα των SPARQL ερωτημάτων σε σχέση με την οντολογία έγκειται στη δυνατότητά τους να ανακτούν συγκεκριμένα δεδομένα από την οντολογία, να απαντούν σε πολύπλοκα ερωτήματα και να παρέχουν χρήσιμες και ακριβείς πληροφορίες στους χρήστες.

³⁴ Η δυνατότητα που έχουν τα SPARQL ερωτήματα να επαναληφθούν από τον οποιονδήποτε, όποια στιγμή και να φέρουν τα ίδια αποτελέσματα κάθε φορά.

Ακολουθούν ενδεικτικά SPARQL ερωτήματα (Εικόνα 22-30, SPARQL ερωτήματα 1-9) με τα οποία ανακτώνται δεδομένα από την οντολογία της μελέτης περίπτωσης που προηγήθηκε. Με τα ερωτήματα αυτά διαπιστώνεται το εύρος των ερωτήσεων που μπορούν να τεθούν και των απαντήσεων που ανακτώνται, παρουσιάζεται η εξέλιξη των ερωτημάτων με την προσθήκη των κατάλληλων τελεστών και μεταβλητών και αναδεικνύονται οι περιορισμοί της.

Σε μια βάση δεδομένων η οποία θα περιέχει έναν κατάλογο των αρχαίων πηγών, η πιο απλή και συνήθης ερώτηση που μπορεί να γίνει από το χρήστη είναι η αναζήτηση των βιβλίων που έγραψε ένας συγκεκριμένος συγγραφέας (Εικόνα 22, SPARQL ερώτημα 1):

```
PREFIX sao: <http://www.SourceArchiveOnto-ontology.com/>
PREFIX cd: <https://w3id.org/arco/ontology/context-description/>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX dct: <http://purl.org/dc/terms/>
```

```
SELECT ?authorName ?book ?bookTitle
WHERE {
  sao:Aeschines cd:isAuthorOf ?book;
  foaf:name ?authorName.
  ?book dct:title ?bookTitle.
}
```

Στο ερώτημα ζητάμε να ανακτηθεί το “όνομα του Συγγραφέα” (?authorName), οι οντότητες από την κλάση του “Βιβλίου” (?bibo:Book) και ο “Τίτλος” του Βιβλίου (?bibo:BookTitle). Έπειτα εισάγουμε τους απαιτούμενους περιορισμούς. Εδώ για παράδειγμα ρωτάμε για τα έργα του “Αισχίνη” (sao:Aeschines), το στιγμιότυπο της οντότητας του οποίου συνδέεται με τα “Βιβλία” του μέσω της ιδιότητας “cd:isAuthorOf” και με το όνομά του μέσω της ιδιότητας “foaf:name”. Τα στιγμιότυπα από την κλάση του “Βιβλίου” που αναζητούμε συνδέονται με τον τίτλο τους μέσω της ιδιότητας “dct:title”. Εκτελώντας το ερώτημα ανακτώνται τρεις στήλες, το όνομα του συγγραφέα (“Αισχίνης”), τα έργα του ως οντότητες της κλάσης του “Βιβλίου” (“sao:KataKtisiphontos”, “sao:PeriTisParapresveias”) και ο τίτλος τους (“Κατὰ Κτησιφώντος”, “Περὶ τῆς παραπροσβείας”).

```
SELECT ?authorName ?book ?bookTitle
WHERE {
  sao:Aeschines cd:isAuthorOf ?book;
  foaf:name ?authorName.
  ?book dct:title ?bookTitle.
}
```

authorName	book	bookTitle
"Αισχίνης"	KataKtisiphontos	"Κατὰ Κτησιφώντος"
"Αισχίνης"	PeriTisParapresveias	"Περὶ τῆς Παραπροσβείας"

Εικόνα 22: Στιγμιότυπο οθόνης από την Query επέκταση του Protégé, SPARQL ερώτημα 1 “Ποια βιβλία έχει γράψει ο Αισχίνης”.

Ένα επόμενο απλό ερώτημα, γεωγραφικού προσδιορισμού αυτή τη φορά, είναι η αναζήτηση των συγγραφέων με Καταγωγή από μια συγκεκριμένη πόλη (Εικόνα 23, SPARQL ερώτημα 2):

```
PREFIX vivo: <http://vivoweb.org/ontology/>
PREFIX db: <http://dbpedia.org/ontology/>
PREFIX sao: <http://www.SourceArchiveOnto-ontology.com/>
```

```
SELECT ?author ?geographicLocation
WHERE {
  ?author db:birthPlace sao:Athens.
}
```

Στο ερώτημα ζητάμε την ανάκτηση των τιμών της κλάσης “Συγγραφέας” (?author) και της κλάσης “Γεωγραφική Περιοχή” (?geographicLocation). Ορίζουμε πως θέλουμε αποτελέσματα μέσω της σχέσης τους από την ιδιότητα “έχει καταγωγή” (“db:birthPlace”) και συγκεκριμένα εδώ μόνο τους συγγραφείς που έχουν καταγωγή από την Αθήνα (sao:Athens). Εκτελώντας το ερώτημα ανακτάται η στήλη “Author” και περιέχει τα στιγμιότυπα της κλάσης που έχουν στις ιδιότητές τους την καταγωγή από την Αθήνα.

```
SELECT ?author ?geographicLocation
WHERE {
  ?author db:birthPlace sao:Athens.
}
```

author
Aeschines
Xenophon
Demosthenes
Pseudo-Scylax
Thucydides

Εικόνα 23: Στιγμιότυπο οθόνης από την Query επέκταση του Protégé, SPARQL ερώτημα 2 “Ποιοι Συγγραφείς κατάγονται από την Αθήνα;”

Μια απλή ερώτηση, αλλά ιδιαίτερα βοηθητική σε μια βάση για την οργάνωση της μελέτης είναι η κατηγοριοποίησή των βιβλίων της σε “Είδος” (Εικόνα 24, SPARQL ερώτημα 3):

Στο ερώτημα ζητάμε την ανάκτηση “Βιβλίου” (?biblio:Book) και “Είδος” (?BookGenre), μέσω της τριπλέτας που ορίζει η ιδιότητα του βιβλίου, “έχει είδος” (“lod:genre”) και η οποία θα μας δώσουν τη στήλη bookGenre. Με την εκτέλεση του ερωτήματος εμφανίζονται δύο στήλες, η μια με στιγμιότυπα της κλάσης του “Βιβλίου” και δίπλα την αντιστοιχία του “Είδους” τους.

PREFIX sao: <http://www.SourceArchiveOnto-ontology.com/>

PREFIX dct: <http://purl.org/dc/terms/>

PREFIX lod: <http://lod.nl.go.kr/ontology/>

SELECT ?book ?bookGenre

WHERE {

?book lod:genre ?bookGenre.

}

```
SELECT ?book ?bookGenre
WHERE {
  ?book lod:genre ?bookGenre.
}
```

book	bookGenre
Historiai	Historical
Istoria	Historical
KataKtisisiphontos	Rhetoric
Agisilaos	Historical
YperKtisisiphontosPeriTouStefanou	Rhetoric
PeriodosTouNikomidi	Geographical
IthikaAltiaEllinika	Historical
IthikaTwnEptaSofwnSymposion	Symposium
ElladosPeriigisis	Geographical
VioiParalliloiPeriklis	Biography
IstoriKivVioithiki	Historical
PeriTisParapresveias	Rhetoric

Εικόνα 24: Στιγμιότυπο οθόνης από την Query επέκταση του Protégé, SPARQL ερώτημα 3 “Σε ποιο Είδος ανήκει το κάθε Βιβλίο;”

Σε ένα αντίστοιχο με το προηγούμενο αλλά πιο σύνθετης μορφής ερώτημα, εκτός από τα

“Είδη”, εάν θέλουμε να ομαδοποιήσουμε τις πηγές με βάση την ποσότητα του υλικού, μπορούμε να ρωτήσουμε πόσα βιβλία από το κάθε “Είδος” υπάρχουν (Εικόνα 25, SPARQL ερώτημα 4):

PREFIX sao: <http://www.SourceArchiveOnto-ontology.com/>

PREFIX dct: <http://purl.org/dc/terms/>

PREFIX lod: <http://lod.nl.go.kr/ontology/>

```
SELECT ?genre (COUNT(?book)AS ?bookCounter)
  WHERE {
    ?book lod:genre ?genre
  }
GROUP BY (?genre)
```

Το ερώτημα είναι σύνθετο καθώς περιέχει τελεστές που διαμορφώνουν την απάντηση. Ζητάμε να ανακτηθούν στιγμιότυπα της κλάσης “Είδος” (?genre). Μέσω του COUNT(?bibo:Book) ο επεξεργαστής μετράει τον αριθμό των “Βιβλίων” (?bibo:Book) που ανήκουν σε κάθε “Είδος” και το αποτέλεσμα θα παρουσιαστεί στην στήλη με το όνομα “bookCounter”, μέσω της συνάρτησης AS ?BookCounter. Έπειτα χρησιμοποιούμε τη ρήτρα WHERE όπως στο παραπάνω ερώτημα (βλ. Εικόνα 24) για τα είδη των βιβλίων (?bibo:Book lod:genre ?genre). Μετά την ολοκλήρωση του ερωτήματος ορίζουμε και τον μετατροπέα (GROUP BY (?genre)) προκειμένου να ζητήσουμε η ομαδοποίηση να γίνει σύμφωνα με κάθε μοναδική τιμή της μεταβλητής “Είδος”. Η εκτέλεση του ερωτήματος εμφανίζει ως αποτέλεσμα δύο στήλες, την “genre” με τα διαφορετικά στιγμιότυπα της κλάσης του “Είδους” και δίπλα τους, υπό την στήλη “bookCounter”, τον αριθμό των αντίστοιχων βιβλίων τους, όπως προκύπτει από την τριπλέτα με την ιδιότητα. Με αυτόν τον τρόπο βλέπουμε για παράδειγμα πως τα “Ιστορικά” βιβλία είναι “9”, της “Ρητορικής” είναι “4” και τα “Γεωγραφικά” είναι “3”. Δίπλα σε κάθε αριθμό υπάρχει URI το οποίο επιβεβαιώνει ότι είναι ακέραιος αριθμός (integer) σύμφωνα και με τους κανόνες της XML Schema.

```
SELECT ?genre (COUNT(?book) AS ?bookCounter)
  WHERE {
    ?book lod:genre ?genre
  }
GROUP BY (?genre)
```

genre	bookCounter
Historical	"9" ^{^^} <http://www.w3.org/2001/XMLSchema#integer>
Rhetoric	"4" ^{^^} <http://www.w3.org/2001/XMLSchema#integer>
Geographical	"3" ^{^^} <http://www.w3.org/2001/XMLSchema#integer>
Symposium	"1" ^{^^} <http://www.w3.org/2001/XMLSchema#integer>
Biography	"2" ^{^^} <http://www.w3.org/2001/XMLSchema#integer>
Geographical_Lexicon	"1" ^{^^} <http://www.w3.org/2001/XMLSchema#integer>
Geography	"1" ^{^^} <http://www.w3.org/2001/XMLSchema#integer>
Mythology	"1" ^{^^} <http://www.w3.org/2001/XMLSchema#integer>

Εικόνα 25: Στιγμιότυπο οθόνης από την Query επέκταση του Protégé, SPARQL ερώτημα 4 “Πόσα είναι τα βιβλία από το κάθε είδος;”

Εκτός από το Είδος του κάθε βιβλίου, σημαντική πληροφορία για την αξιολόγηση και αξιοποίηση του υλικού από έναν μελετητή είναι το πόσα αποσπάσματα αντιστοιχούν στο κάθε βιβλίο (πάντα για το θέμα της βάσης και σε όσα βιβλία αυτή περιέχει) (Εικόνα 26, SPARQL ερώτημα 5):

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

PREFIX dct: <http://purl.org/dc/terms/>

PREFIX sao: <http://www.SourceArchiveOnto-ontology.com/>

PREFIX cd: <https://w3id.org/arco/ontology/context-description/>

PREFIX purl: <http://purl.org/dc/terms/>

```
SELECT DISTINCT ?authorName
          ?bookTitle
(COUNT(?passage) AS ?passageCounter)
```



```

WHERE {
?authorName cd:isAuthorOf ?book.
?book sao:hasPassage ?passage .
?passage purl:subject ?subject ;
rdf:type sao:Passage.
?book dct:title ?bookTitle.
}
GROUP BY ?book ?bookTitle ?authorName

```

Στο ερώτημα αυτό επιλέγουμε τις λίστες τιμών που θέλουμε να σχηματιστούν (?authorName, ?bookTitle) και χρησιμοποιώντας τη ρήτρα DISTINCT για μοναδικά αποτελέσματα και τον τελεστή COUNT, ζητάμε και τα ?passage του κάθε συνδυασμού αυτών των μεταβλητών, τα αποτελέσματα να καταγραφούν στη λίστα “passageCounter” ((COUNT(?passage) AS ?passageCounter)). Με τη λέξη-κλειδί WHERE ορίζουμε με ακρίβεια τις σχέσεις των οντοτήτων που ζητήσαμε. Τα στιγμιότυπα του “Συγγραφέα” συνδέονται με τα βιβλία τους μέσω της ιδιότητας “είναι συγγραφέας του” (“cd:isAuthorOf”). Με τη μεταβλητή ?book sao:hasPassage ?passage δηλώνουμε πως το οποιοδήποτε βιβλίο (?book) συνδέεται με όλα του τα αποσπάσματα (?passage) μέσω της ιδιότητας “έχει χωρίο” (sao:hasPassage). Οι μεταβλητές ?passage purl:subject ?subject; και rdf:type sao:Passage ορίζει πως αναζητούμε τα αποσπάσματα τα οποία είναι “sao:Passage” (δηλαδή είναι στιγμιότυπα τα οποία ανήκουν στην κλάση “Passage”) και πρέπει υποχρεωτικά να “έχουν θέμα” (“purl:subject”) ένα στιγμιότυπο της κλάσης “getty:Subject”. Τέλος το ?bookTitle θα προκύψει από την αναζήτηση της ονομασίας των οντοτήτων των βιβλίων μέσω μιας literal μοναδικής τιμής από την ιδιότητά τους “έχει τίτλο” (“dct:title”) Με τη ρήτρα GROUP BY, ορίζουμε πως τα αποτελέσματα θα ομαδοποιηθούν βάσει των τιμών των μεταβλητών που ορίστηκαν (?book ?bookTitle ?authorName). Με την εκτέλεσή του ερωτήματος ανακτάται πίνακας με 3 στήλες με το όνομα του Συγγραφέα, το Βιβλίο του και τον αριθμό των αποσπασμάτων του βιβλίου αυτού. Έτσι, διαβάζουμε, για παράδειγμα, για την οντότητα Συγγραφέα “Thucydides”, από το βιβλίο του “Ιστορία”, στην οντολογία είναι καταγεγραμμένα “14” αποσπάσματα.

```

SELECT DISTINCT ?authorName
?bookTitle
(COUNT(?passage) AS ?passageCounter)
WHERE {
?authorName cd:isAuthorOf ?book.
?book sao:hasPassage ?passage.
?passage purl:subject ?subject ;
rdf:type sao:Passage.
?book dct:title ?bookTitle.
}
GROUP BY ?book ?bookTitle ?authorName

```

authorName	bookTitle	passageCounter
Anonymous	"Ὀμηρικὰ τῶς ἑυρρύχου"	"1" http://www.w3.org/2001/XMLSchema#integer
Herodotus	"Ἱστορία"	"6" http://www.w3.org/2001/XMLSchema#integer
Plutarch	"Ἐθικά, τῶν ἑσσοφ συμποσίων"	"1" http://www.w3.org/2001/XMLSchema#integer
Diodorus	"Ἱστορικὴ Βιβλιοθήκη"	"14" http://www.w3.org/2001/XMLSchema#integer
Pseudo-Apollodorus	"Βιβλιοθήκη"	"1" http://www.w3.org/2001/XMLSchema#integer
Pseudo-Scylax	"Περσικὸς Ἱστορικός"	"1" http://www.w3.org/2001/XMLSchema#integer
Aeschines	"Κατὰ Κτησιφῶντος"	"2" http://www.w3.org/2001/XMLSchema#integer
Strabo	"Γεωγραφικά"	"9" http://www.w3.org/2001/XMLSchema#integer
Thucydides	"Ἱστορία"	"14" http://www.w3.org/2001/XMLSchema#integer
Pausanias	"Ἐλλάδος Περιήγησις"	"8" http://www.w3.org/2001/XMLSchema#integer
Demosthenes	"Ἔπερ Κτησιφῶντος περὶ τοῦ Στεφάνου"	"1" http://www.w3.org/2001/XMLSchema#integer

Εικόνα 26: Στιγμιότυπο οθόνης από την Query επέκταση του Protégé, SPARQL ερώτημα 5 “Πόσα αποσπάσματα έχουμε από το κάθε βιβλίο;”

Μια ακόλουθη ερώτηση μπορεί να είναι η λίστα των βιβλίων που ανήκουν σε κάποιο “Είδος” (Εικόνα 27, SPARQL ερώτημα 6):

```

PREFIX sao: <http://www.SourceArchiveOnto-ontology.com/>
PREFIX cd: <https://w3id.org/arco/ontology/context-description/>
PREFIX dct: <http://purl.org/dc/terms/>
PREFIX lod: <http://lod.nl.go.kr/ontology/>

```

```

SELECT ?book
WHERE {
?book lod:genre sao:Historical.
}

```

Σε αυτό το ερώτημα ζητάμε ανάκτηση από την κλάση “Βιβλίο” (SELECT ?bibo:Book). Με τη λέξη-κλειδί WHERE ορίζουμε την επιθυμητή τριπλέτα: οι τιμές που θα ανακτηθούν θα πρέπει να είναι οντότητες από την κλάση “bibo:Book” και να “έχουν Είδος” (“Iod:genre”) και πιο συγκεκριμένα να ανήκουν στο είδος “Ιστορικά” (“sao:Historical”). Η εκτέλεση του ερωτήματος φέρνει μια λίστα book, όπου βρίσκονται οντότητες από την κλάση “bibo:Book” με αυτή την τιμή στην ιδιότητά τους.

```
SELECT ?book
WHERE {
?book Iod:genre sao:Historical.
}
```

	book
Historiai	
Istoria	
Agisilaos	
IthikaAitiaEllinika	
IstorikiViviothiki	
Ellinika	
Istoriai	
Stratigimata	
EllinikaTisOksirinchou	

Εικόνα 27: Στιγμιότυπο οθόνης από την Query επέκταση του Protégé, SPARQL ερώτημα 6 “Ποια είναι τα βιβλία που ανήκουν στο Είδος Ιστορικά;”

Εξειδικεύοντας την αναζήτησή μας ώστε να γίνει πιο συγκεκριμένη ως προς τον τομέα που ερευνούμε, μπορούμε να αναζητήσουμε όλα τα αποσπάσματα που αφορούν μια ιδιαίτερη θεματική ενότητα (Εικόνα 28, SPARQL ερώτημα 7):

```
PREFIX sao: <http://www.SourceArchiveOnto-ontology.com/>
PREFIX purl: <http://purl.org/dc/terms/>
```

```
SELECT ?passage
WHERE {
?passage purl:subject sao:PeloponnesiakosPolemos.
}
```

Σε αυτό το ερώτημα ζητάμε να ανακτηθούν σε μια λίστα τα χωρία τα οποία αναφέρονται στον Πελοποννησιακό πόλεμο. Με τη μεταβλητή SELECT δηλώνουμε πως θέλουμε αποτελέσματα από τα στιγμιότυπα της κλάσης “sao:Passage” (?passage). Έπειτα με το WHERE κάνουμε την αναζήτηση από όλα τα αποσπάσματα πιο συγκεκριμένα, διευκρινίζοντας ότι το απόσπασμα (?passage) θα πρέπει να συνδέεται με την ιδιότητα “έχει θέμα” (purl:subject) με στιγμιότυπο από την κλάση “getty:Subject”. Στο συγκεκριμένο παράδειγμα “χρειαζόμαστε” μόνο τις τιμές που μέσω της ιδιότητας του θέματος συνδέονται με τη θεματική ενότητα “Πελοποννησιακός πόλεμος” (sao:PeloponnesiakosPolemos). Η εκτέλεση του ερωτήματος εμφανίζει απλά και μόνο μια στήλη passage στην οποία καταγράφονται όλα τα χωρία με αυτό το θέμα, ανεξάρτητα από οποιοδήποτε άλλο περιορισμό που θα μπορούσε να υπάρξει, όπως για παράδειγμα το βιβλίο στο οποίο ανήκει, ο συγγραφέας του ή η χρονολογία.

```
SELECT ?passage
WHERE {
?passage purl:subject sao:PeloponnesiakosPolemos.
}
```

	passage
7.19.5	
10.38.10	
15.57.1-2	
11.85.1-2	
2.102.1	
12.80.4	
7.17.4	
3.98.3-5	
17	
12.42-8	
11.84.6-8	
2.69.1	

Εικόνα 28: Στιγμιότυπο οθόνης από την Query επέκταση του Protégé, SPARQL ερώτημα 7 “ποια χωρία έχουν ως θέμα τους τον Πελοποννησιακό πόλεμο;”

Η αναζήτηση αυτή είναι αρκετά απλή και γρήγορη, όχι όμως ιδιαίτερα χρήσιμη για ερευνητικούς σκοπούς. Παρακάτω (Εικόνες 29-30, SPARQL ερωτήματα 8-9) προσθέτουμε στο ερώτημα περισσότερες μεταβλητές και βλέπουμε να λαμβάνονται πιο περιεκτικά και αξιολογα αποτελέσματα. Ξεκινάμε με την ερώτηση σε ποιο βιβλίο και ποιον συγγραφέα ανήκουν τα αποσπάσματα που αφορούν τον Πελοποννησιακό Πόλεμο (Εικόνα 29, SPARQL ερώτημα 8):

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX sao: <http://www.SourceArchiveOnto-ontology.com/>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX db: <http://dbpedia.org/ontology/>
PREFIX dct: <http://purl.org/dc/terms/>
```

```
SELECT DISTINCT ?book ?passage ?name
WHERE {
  ?passage purl:subject sao:PeloponniasiakosPolemos ; rdf:type sao:Passage.
  ?book sao:hasPassage ?passage ; dct:creator ?author.
  ?author foaf:name ?name.
}
```

Σε αυτό το ερώτημα δεν αναζητούμε απλά όλα τα αποσπάσματα, αλλά θέλουμε να φανεί η σύνδεση κάθε αποσπάσματος με το βιβλίο και το συγγραφέα του. Με τη λέξη-κλειδί SELECT ορίζουμε πως θέλουμε να ανακτήσουμε αποτελέσματα σε τρεις στήλες (book, passage και name) και με την DISTINCT εξασφαλίζουμε ότι κάθε αποτέλεσμα που θα εμφανιστεί θα είναι μοναδικό. Με τη ρήτρα WHERE περιγράφουμε την προέλευση που θέλουμε να έχουν τα αποτελέσματά μας. Το “Απόσπασμα” (“sao:Passage”) συνδέεται μέσω της ιδιότητας “έχει θέμα” (“purl:subject”) με το θέμα “sao:PeloponniasiakosPolemos”, μια σχέση που είδαμε και στο προηγούμενο ερώτημα. Το απόσπασμα πλέον συνδέεται και με την κλάση του “bibo:Book” μέσω της ιδιότητας “έχει απόσπασμα” (“sao:hasPassage”). Το βιβλίο σχετίζεται με το Συγγραφέα του μέσω της αντίστροφης ιδιότητάς “dct:creator” (“έχει συγγραφέα”). Οι οντότητες-στιγμιότυπα της κλάσης του “db:Author” με την data property “foaf:name” έχουν ονόματα στη μητρική φυσική γλώσσα. Η εκτέλεση του ερωτήματος έχει ως αποτέλεσμα τρεις στήλες: στην μεσαία στήλη με το όνομα “passage” εμφανίζεται –όπως και προηγουμένως- η λίστα με τα αποσπάσματα. Πλέον όμως αυτή η πληροφορία εμπλουτίστηκε με την πρώτη στήλη με τίτλο book, από όπου προκύπτουν τα στιγμιότυπα της κλάσης “bibo:Book” από τα οποία προέρχονται αυτά τα αποσπάσματα. Κάθε βιβλίο συνδέεται με ένα συγγραφέα και κάθε στιγμιότυπο συγγραφέα έχει ένα όνομα μέσω της foaf:name. Επομένως, διαβάζουμε κατά σειρά, για παράδειγμα, “Istoria”- “7.19.5”- “Θουκυδίδης” και καταλαβαίνουμε πως το απόσπασμα 7.19.5 -που μιλά για τον Πελοποννησιακό Πόλεμο- βρίσκεται στο βιβλίο Istoria της βάσης το οποίο επίσης έχει για συγγραφέα το Θουκυδίδη.

```
SELECT DISTINCT ?book ?passage ?name
WHERE {
  ?passage purl:subject sao:PeloponniasiakosPolemos ; rdf:type sao:Passage.
  ?book sao:hasPassage ?passage ; dct:creator ?author.
  ?author foaf:name ?name .
}
```

book	passage	name
Istoria	7.19.5	"Θουκυδίδης"
ElladosPerigisis	10.38.10	"Παυσανίας"
IstoriKiVivliothiki	15.57.1-2	"Διόδωρος"
IstoriKiVivliothiki	11.85.1-2	"Διόδωρος"
Istoria	2.102.1	"Θουκυδίδης"
IstoriKiVivliothiki	12.80.4	"Διόδωρος"
Istoria	7.17.4	"Θουκυδίδης"
Istoria	3.98.3-5	"Θουκυδίδης"
VioiParallioiPeriklis	17	"Πλούταρχος"
IstoriKiVivliothiki	12.42-8	"Διόδωρος"
IstoriKiVivliothiki	11.84.6-8	"Διόδωρος"
Istoria	2.69.1	"Θουκυδίδης"

Εικόνα 29: Στιγμιότυπο οθόνης από την Query επέκταση του Protégé, SPARQL ερώτημα 8 “Σε ποιο βιβλίο και σε ποιόν συγγραφέα ανήκει το κάθε χωρίο που αφορά τον Πελοποννησιακό πόλεμο”.

Μπορούμε να εξειδικεύσουμε την ερώτηση για να λάβουμε ακόμα πιο συγκεκριμένα (περιορισμένα) αποτελέσματα, όταν θέλουμε να ασχοληθούμε με ένα ιδιαίτερα ορισμένο πεδίο, όπως είναι η αναζήτηση των πηγών για ένα θέμα σε πηγές συγκεκριμένου χρονικού εύρους (Εικόνα 30, SPARQL ερώτημα 9):

PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

PREFIX sao: <http://www.SourceArchiveOnto-ontology.com/>

PREFIX dct: <http://purl.org/dc/terms/>

PREFIX purl: <http://purl.org/dc/terms/>

PREFIX foaf: <http://xmlns.com/foaf/0.1/>

SELECT DISTINCT ?book ?pubCent ?passage ?name

WHERE {

?passage purl:subject sao:PeloponniakosPolemos ; rdfs:type sao:Passage.

?book sao:hasPassage ?passage ; dct:creator ?author ; sao:hasPublicationCentury ?pubCent.

?author foaf:name ?name.

}

FILTER regex(?pubCent, "π.X")

Το ερώτημα έχει στην αναζητήσή του τις περισσότερες μεταβλητές του ίδιες με το προηγούμενο. Ζητούμενα είναι και πάλι τα μοναδικά αποτελέσματα που θα υπάγονται σε τρεις στήλες (book, passage και name) και επεκτείνοντάς με τη μεταβλητή ?pubCent προσθέτουμε στην απάντηση μια στήλη χρονολογικού προσδιορισμού. Η λέξη-κλειδί WHERE θα έχει επίσης τις ίδιες μεταβλητές, με τη διαφορά ότι το απόσπασμα δε θα συνδέεται μόνο με το βιβλίο και το συγγραφέα του, αλλά και με τον "Αιώνα Έκδοσης" μέσω της ιδιότητάς του "sao:hasPublicationCentury". Για να περιορίσουμε τα αποτελέσματα ώστε να ανακτηθούν μόνο αυτά που περιέχουν το "π.X." στην τιμή της μεταβλητής, προσθέτουμε το φίλτρο regex(?pubCent, "π.X."). Η εκτέλεση του ερωτήματος ανακτά τέσσερις στήλες, όπου τα δεδομένα που εμφανίζονται πρέπει να καλύπτουν όλες τις προϋποθέσεις: η πρώτη (book) περιέχει τα στιγμιότυπα των οντοτήτων των βιβλίων. Η δεύτερη (pubCent) περιλαμβάνει τη χρονολόγηση του κάθε έργου. Η τρίτη (passage) είναι η λίστα των χωρίων που αναφέρονται στον Πελοποννησιακό πόλεμο, όπως και στα δύο προηγούμενα ερωτήματα και στην τέταρτη (name) έχουμε τα ονόματα των στιγμιότυπων της κλάσης "db:Author" που αντιστοιχούν στο κάθε έργο.

```
SELECT DISTINCT ?book ?pubCent ?passage ?name
WHERE {
  ?passage purl:subject sao:PeloponniakosPolemos ; rdfs:type sao:Passage.
  ?book sao:hasPassage ?passage ; dct:creator ?author ; sao:hasPublicationCentury ?pubCent.
  ?author foaf:name ?name.
  FILTER regex(?pubCent, "π.X.")
}
```

book	pubCent	passage	name
Istoria	"β' μισό του 5ου αι. π.Χ."	2.102.1	"Θουκυδίδης"
Istoria	"β' μισό του 5ου αι. π.Χ."	7.31-34.8	"Θουκυδίδης"
Istoria	"β' μισό του 5ου αι. π.Χ."	7.19.5	"Θουκυδίδης"
Istoria	"β' μισό του 5ου αι. π.Χ."	3.94-96.3	"Θουκυδίδης"
Istoria	"β' μισό του 5ου αι. π.Χ."	3.97.1-2	"Θουκυδίδης"
Istoria	"β' μισό του 5ου αι. π.Χ."	2.9.2-4	"Θουκυδίδης"
Istoria	"β' μισό του 5ου αι. π.Χ."	7.17.4	"Θουκυδίδης"
Istoria	"β' μισό του 5ου αι. π.Χ."	2.69.1	"Θουκυδίδης"
Istoria	"β' μισό του 5ου αι. π.Χ."	2.80-92.7	"Θουκυδίδης"
Istoria	"β' μισό του 5ου αι. π.Χ."	3.100-102.5	"Θουκυδίδης"
Istoria	"β' μισό του 5ου αι. π.Χ."	5.32.2	"Θουκυδίδης"
Istoria	"β' μισό του 5ου αι. π.Χ."	3.98.3-5	"Θουκυδίδης"

Εικόνα 30: Στιγμιότυπο οθόνης από την Query επέκταση του Protégé, SPARQL ερώτημα 9 "Ποια είναι τα π.Χ. αποσπάσματα που αφορούν τον Πελοποννησιακό πόλεμο".

Τα SPARQL ερωτήματα που παρουσιάστηκαν (Εικόνες 22-30, SPARQL 1-9) είναι ενδεικτικά. Χρησιμοποιήθηκαν ως απόδειξη της λειτουργικότητας και της δομής του οντολογικού προτύπου που αναπτύχθηκε και αναδεικνύουν τη χρησιμότητα και τα όρια του, αποκαλύπτοντας το εύρος των απαντήσεων που μπορεί να ανακτήσει. Αυτό φαίνεται στα πρώτα ερωτήματα, SPARQL 1-6), όπου θέσαμε απλές ερωτήσεις για τα έργα ή την καταγωγή ενός συγγραφέα (SPARQL 1-2), καθώς και διευκρινιστικά και ποσοτικά ερωτήματα σχετικά με τα βιβλία (SPARQL 3-6). Συγκεκριμένα διαχωρίζουμε τα είδη στα οποία ανήκουν και έπειτα ρωτήσαμε πόσα είναι τα βιβλία που ανήκουν σε κάθε είδος αλλά και πόσα αποσπάσματα υπάρχουν αυτό το βιβλίο. Στο τελευταίο κομμάτι (SPARQL 7-9) αναλύθηκε ένα ερώτημα το οποίο αποδείξαμε πως με μικρές μεταβολές μπορεί να γίνει σύνθετο και να φέρνει την ανάκτηση πολύ συγκεκριμένων απαντήσεων. Ξεκινώντας από ένα απλό ερώτημα σχετικά με “ποια χωρία έχουν θέμα τον Πελοποννησιακό Πόλεμο” (SPARQL 7), εμπλουτίσαμε το ερώτημα ώστε να μην ανακτηθούν μόνο τα αποσπάσματα, αλλά και η πληροφορία για το ποιος είναι ο συγγραφέας του και σε ποιο βιβλίο βρίσκεται (SPARQL 8). Το ερώτημα τέλος έγινε περισσότερο συγκεκριμένο, ζητώντας να ανακτηθούν μόνο τα αποτελέσματα που χρονολογούνται μια συγκεκριμένη χρονική περίοδο (π.Χ.) (SPARQ 9).

Εικόνα 22, SPARQL 1	Ποια βιβλία έχει γράψει ο Αισχίνης;
Εικόνα 23, SPARQL 2	Ποιοι συγγραφείς κατάγονται από την Αθήνα;
Εικόνα 24, SPARQL 3	Τι είδους είναι το κάθε Βιβλίο;
Εικόνα 25, SPARQL 4	Πόσα είναι τα βιβλία από το κάθε είδος;
Εικόνα 26, SPARQL 5	Πόσα αποσπάσματα έχουμε από το κάθε βιβλίο;
Εικόνα 27, SPARQL 6	Ποια είναι τα βιβλία που ανήκουν στο Είδος Ιστορικά;
Εικόνα 28, SPARQL 7	Ποια χωρία έχουν ως θέμα τους τον Πελοποννησιακό πόλεμο;
Εικόνα 29, SPARQL 8	Σε ποιο βιβλίο και σε ποιόν συγγραφέα ανήκει το κάθε χωρίο που αφορά τον Πελοποννησιακό πόλεμο;
Εικόνα 30, SPARQL 9	Ποια είναι τα π.Χ. αποσπάσματα που αφορούν τον Πελοποννησιακό πόλεμο;

ΕΠΙΛΟΓΟΣ

Τα εργαλεία της τεχνολογίας συνεχώς εξελίσσονται, θέτοντας την επιστημονική έρευνα σε νέα βάση, καθώς μεταβάλλουν τον τρόπο με τον οποίο διαχειριζόμαστε, αποθηκεύουμε και αναλύουμε τις πληροφορίες. Η αδυναμία του Παγκόσμιου Ιστού να οργανώνει και να αξιολογήσει τον όγκο των δεδομένων αποτελεσματικά και η δυσκολία επικοινωνίας με τους χρήστες, έχουν αναδείξει την ανάγκη αναβάθμισής του στον Σημασιολογικό Ιστό. Η εξέλιξη στοχεύει στη δημιουργία μιας μεθόδου και ενός πλαισίου αποτύπωσης και διασύνδεσης των δεδομένων με τρόπο κατανοητό από άνθρωπο και μηχανή. Κύριο ρόλο στη μετάβαση θα έχει ο τρόπος αναπαράστασης, ο οποίος επιτυγχάνεται με την ανάπτυξη των οντολογιών, οι οποίες με τη χρήση κατάλληλων δομών και τυποποιημένων γλωσσικών μοντέλων, θα καθιστούν τα δεδομένα κατανοητά και συνδεδεμένα. Οι οντολογίες συμβάλουν επιπλέον στην επιστημονική έρευνα, επιτρέποντας τη συστηματική και αποτελεσματική οργάνωση και τη διασύνδεσή με δεδομένα και άλλης μορφής.

Η πρόοδος της επιστημονικής διαδικασίας εξαρτάται διαχρονικά από τα διαθέσιμα εργαλεία της τεχνολογίας. Στο χώρο της ιστορικής έρευνας, η τεχνολογία έχει τροποποιήσει ουσιαστικά τον τρόπο με τον οποίο ασκείται. Η ερευνητική της διαδικασία απαιτεί τη συλλογή, οργάνωση, μελέτη και ανάλυση των πηγών για την εξαγωγή ποιοτικών και ακριβέστερων συμπερασμάτων. Ως πηγές εννοούμε κάθε τεκμήριο το οποίο προέρχεται από την αρχαιότητα και παρέχει πληροφορίες για οποιαδήποτε ανθρώπινη δραστηριότητα. Η παραδοσιακή διαδικασία μελέτης έχει ενταχθεί σε ένα ψηφιακό πλαίσιο και οι πηγές πλέον ψηφιοποιούνται, οργανώνονται σε ένα οντολογικό πλαίσιο και στη συνέχεια καταχωρούνται σε ηλεκτρονικές βάσεις δεδομένων.

Η εργασία ανταποκρίνεται στην ανάγκη για την απλή, ευέλικτη και αποδοτική αρχειοθέτηση των αρχαίων φιλολογικών πηγών. Το βασικότερο κομμάτι της αποτέλεσε η ανάπτυξη μιας οντολογίας μέσω του προγράμματος ανοιχτού κώδικα Protégé. Η οντολογία αποτελεί ένα πρότυπο αποτύπωσης των αναγκών αρχειοθέτησης των αρχαίων πηγών. Απευθύνεται στους ερευνητές, καλύπτοντας την επιθυμία για εύκολη και γρήγορη πρόσβαση στις πηγές και τα απαραίτητα μεταδεδομένα. Για τον έλεγχο της λειτουργικότητας, της χρησιμότητας και κατ' επέκταση την μερική κάλυψη ενός σημαντικού ερευνητικού κενού, κρίθηκε απαραίτητη η δημιουργία μιας οντολογίας με τη χρήση της μελέτης περίπτωσης των αρχαίων γραμματειακών πηγών της δυτικής Λοκρίδας. Η εργασία περιλαμβάνει επίσης και ένα γραπτό μέρος (σ.σ. το παρόν) το οποίο λειτουργεί επικουρικά και υποστηρικτικά στις οντολογίες. Σε αυτήν ορίσαμε το πλαίσιο, περιγράψαμε τη θεωρητική διαδικασία κατασκευής, καταγράψαμε τους κύριους όρους και παρουσιάσαμε αντιπροσωπευτικά παραδείγματα, απαραίτητα για την τεκμηρίωση των θεωρητικών δεδομένων.

Το κύριο μέρος της χωρίστηκε σε δύο μέρη, το θεωρητικό υπόβαθρο και την πρακτική ανάλυση. Στο θεωρητικό υπόβαθρο είναι το πλαίσιο στο οποίο εντάσσεται η κατασκευή της οντολογίας. Αρχικά δόθηκε ο ορισμός και εξηγήθηκε ο τρόπος λειτουργίας του Σημασιολογικού Ιστού. Σχολιάστηκε η αναγκαιότητα για εξέλιξη και η επίδραση στη διάδοση της γνώσης. Έπειτα εμβαθήναμε σε ένα βασικό άξονα του Σημασιολογικού Ιστού, τις Οντολογίες. Έγινε προσπάθεια ορισμού της πολυπλοκότητάς τους και σχολιάστηκαν η δομή, ο τρόπος λειτουργίας, η ανάγκη αξιοποίησης, οι δυνατότητες τους. Αξιολογήθηκαν τα μελλοντικά οφέλη και οι προκλήσεις από την ανάπτυξή τους. Κατόπιν επικεντρωθήκαμε στο ιδιαίτερο θέμα, την “Καταγραφή των αρχαίων πηγών”. Δόθηκε ο ορισμός και στη συνέχεια αναλύθηκαν οι προκλήσεις από την εξέλιξη της διαδικασίας λόγω των ψηφιακών εργαλείων. Εστιάσαμε στη λειτουργία της ψηφιοποίησης και στη συνολική επίδραση των τεχνολογικών εργαλείων στην ιστορική έρευνα και επιστήμη.

Η δεύτερη ενότητα αποτελεί την πρακτική ανάλυση των οντολογιών που αναπτύχθηκαν. Στο πρώτο κομμάτι περιγράψαμε και αναλύσαμε τα βήματα στα οποία βασίζεται η διαδικασία κατασκευής μιας οντολογίας. Έπειτα αναλύσαμε το οντολογικό πρότυπο που κατασκευάσαμε. Παρουσιάζοντας εικόνες καταγραφής οθόνης (screenshots) από το Protégé, αναλύσαμε τους βασικούς όρους της οντολογίας (κλάσεις και ιδιότητες) και αξιολογήσαμε τις επιλογές. Καθώς το πρότυπο πρέπει να μπορεί να εφαρμοστεί σε μια πραγματική υπόθεση εργασίας, επικεντρωθήκαμε στη μελέτη περίπτωσης που επιλέχθηκε, τις γραμματειακές πηγές της δυτικής Λοκρίδας. Πρώτα σχολιάστηκε συνοπτικά ο συλλογισμός για την επιλογή του θέματος, παρουσιάστηκε η υπάρχουσα ερευνητική κατάσταση και αξιολογήθηκε η σημασία της επιλογής. Ακολουθώντας το πρότυπο, δόθηκαν παραδείγματα από στιγμιότυπα των κλάσεων για να αναδειχθεί η δομή και η λειτουργικότητα της οντολογίας. Τέλος, παραθέσαμε κάποια ενδεικτικά, απλά ως σύνθετα SPARQL ερωτήματα για να διαπιστώσουμε τη συντακτική αρτιότητα, το εύρος και τους περιορισμούς των δυνατοτήτων του προτύπου και να αναδείξουμε τη χρησιμότητά του για την ιστορική έρευνα. Στο

τελευταίο κομμάτι της εργασίας αναφερθήκαμε στις μελλοντικές προεκτάσεις των οντολογιών.

Η φύση των οντολογιών τους υπαγορεύει να αναπτύσσονται συνεχώς και μερικοί από τους στόχους τους είναι να είναι ευέλικτες ώστε αυτό να γίνεται με ευκολία, καθώς και να μπορούν να εμπλουτίζονται, να τροποποιούνται και να επαναχρησιμοποιούνται. Οι οντολογίες που αναπτύχθηκαν ανταποκρίνονται σε αυτά τα ζητούμενα, ωστόσο η δομή τους είναι τέτοια που επιτρέπει μελλοντικές επεκτάσεις και βελτιώσεις. Τονίστηκε πως οι οντολογίες που κατασκευάστηκαν αποτελούν μέρος μιας διπλωματικής εργασίας, επομένως η δομή και το περιεχόμενό τους υπάγονται σε ορισμένους περιορισμούς, χρόνου και πολυπλοκότητας. Οι περιορισμοί συνεπάγονται πως η οντολογία είναι λειτουργική αλλά δεν καλύπτει όλες τις δυνατές λεπτομέρειες. Αυτό ισχύει τόσο για το πρότυπο, όσο και για τη μελέτη περίπτωσης. Ζητούμενο ήταν να παρουσιαστεί μια ιδέα και να εφαρμοστεί σε ένα παράδειγμα και όχι να αναπαρασταθεί ολιστικά μια οντολογία.

Μελλοντικά η οντολογία θα μπορούσε να επεκταθεί με την προσθήκη περισσότερων κλάσεων και ιδιοτήτων. Αυτό ήταν και ένα από τα κριτήρια ώστε να παραμείνει η οντολογία απλή και ευέλικτη. Στην υπάρχουσα δομή αρχικά μπορούν να προστεθούν οι αντίστροφες ιδιότητες, οι οποίες στο Protégé “εννοούνται” όμως για λόγους πληρότητας μπορούν να προστεθούν (εδώ αναπτύξαμε αυτές που θεωρήθηκαν χρήσιμες στην τεκμηρίωση). Μια άλλη κλάση θα μπορούσαν να είναι τα “Κείμενα”, ώστε να συνδεθούν τα χωρία με τα αποσπάσματά τους σε μορφή κειμένου. Οι “Σύγχρονες Εκδόσεις” θα μπορούσαν να αποτελέσουν ξεχωριστή κλάση με δικές τους “data properties” (εκδοτικός οίκος, χρονολογία έκδοσης, συγγραφέας, χειρόγραφα αναφοράς κ.α.). Ιδιότητες για ενσωμάτωση επιπλέον πληροφοριών θα μπορούσε να έχουν και οι “Συγγραφείς” των πηγών, λ.χ. να προστεθούν το επάγγελμα ή το εύρος ζωής τους. Με όλες τις προσθήκες και την ενσωμάτωση συνεχώς συμπληρωματικών δεδομένων η οντολογία γίνεται εργαλείο που επιτρέπει στους ερευνητές να ανακτούν εύκολα, γρήγορα και οργανωμένα μια ολοκληρωμένη εικόνα, φιλτράροντας τις απαντήσεις ανάλογα με τις κατά συνθήκη ερευνητικές ανάγκες τους.

Η οντολογία θα μπορούσε να επεκταθεί και με άλλον τρόπο πέρα από την προσθήκη κλάσεων και ιδιοτήτων στη δομή της. Μέσω της σύνδεσης με άλλες, υφιστάμενες οντολογίες ή βάσεις δεδομένων σχετικές με την αρχαία ιστορία, δεν επιβαρύνεται η ίδια η οντολογία μας, αλλά ενσωματώνεται σε ένα ευρύτερο σύστημα προαγωγής της γνώσης. Αυτή η λειτουργία, είναι ένας ακόμη λόγος που το πρότυπο έπρεπε να είναι αρκετά απλό και ευέλικτο, ώστε να μην είναι προβληματικό το να συνδεθεί με άλλα εργαλεία του Ιστού. Στο ίδιο πλαίσιο, η οντολογία μπορεί να ενταχθεί σε ένα σύστημα βάσεων που αφορούν άλλα είδη πηγών (επιγραφές, χειρόγραφα, πάπυροι, αρχαιολογικά κατάλοιπα). Αυτό θα άνοιγε νέους ορίζοντες στην ιστορική έρευνα, προσφέροντας στον ερευνητή την προοπτική της παρακολούθησης του συνόλου των διαθέσιμων πηγών, οπότε και την εξαγωγή πληρέστερων, περισσότερο αξιόπιστων και αξιόλογων συμπερασμάτων.

Το πρότυπο χρησιμοποιήθηκε για τη δημιουργία οντολογίας για τις αρχαίες πηγές της δυτικής Λοκρίδας. Το παράδειγμα θέτει τις βάσεις για την κάλυψη ενός ερευνητικού κενού και έχει μεγάλο ιστορικό ενδιαφέρον. Αναφορικά με ενδεχόμενες προεκτάσεις του, ισχύουν οι προτάσεις που ειπώθηκαν προηγουμένως: η παρούσα οντολογία μπορεί να εμπλουτιστεί με περισσότερες κλάσεις και ιδιότητες, με περισσότερες πηγές και με την προσθήκη κατόπιν μετατροπών άλλου είδους πηγών για την απόκτηση μιας συνολικότερη και πιο εμπειριστατωμένη εικόνας σχετικά με τη Λοκρίδα. Επίσης μπορεί να συνδεθεί με άλλες πηγές για άλλες γεωγραφικές περιοχές και να ενταχθεί σε μια μεγαλύτερη βάση αναφορικά με τις πηγές περιοχών του ελληνικού χώρου. Για την εργασία κρίθηκε απαραίτητη η δημιουργία του παραδείγματος ώστε να ελεγχθεί η πρακτική εφαρμογή και η λειτουργικότητα του προτύπου. Επομένως συνεπάγεται πως μπορεί να εφαρμοστεί και για οποιαδήποτε γεωγραφική περιοχή ή συγκεκριμένη ιστορική περίοδο.

Η εργασία συμβάλει στην διαμόρφωση της οργάνωσης και της διασύνδεσης των δεδομένων που σχετίζονται με τις αρχαίες πηγές, αξιοποιώντας τις δυνατότητες της τεχνολογίας και του Σηματολογικού Ιστού. Οι οντολογίες που αναπτύχθηκαν μπορούν να αποτελέσουν μια καλή βάση για αντίστοιχες προσπάθειες στο μέλλον και αποδεικνύουν τις νέες δυνατότητες στην ιστορική έρευνα αλλά και τις προεκτάσεις της “συνεργασίας” των επιστημών και των εργαλείων της τεχνολογίας.

ΒΙΒΛΙΟΓΡΑΦΙΑ

Αγλαΐνη, Γ. Α. (2016). Σημασιολογικός παγκόσμιος ιστός και εφαρμογές του.

Bikakis, N., Tsinaraki, C., Gioldasis, N., Stavrakantonakis, I., & Christodoulakis, S. (2013). The XML and semantic web worlds: technologies, interoperability and integration: a survey of the state of the art. *Semantic hyper/multimedia adaptation: Schemes and applications*, 319-360.

Chouchene, A. M. (2019). Historical research in the digital age: opportunities and challenges. *Int. J. Humanit. Cult. Stud.*, 6(2), 73-83.

Debellis, M. (2021). A practical guide to building owl ontologies using Protégé 5.5 and plugins. The University of Manchester.

Falbo, R. D. A., de Menezes, C. S., & da Rocha, A. R. C. (1998, October). A systematic approach for building ontologies. In *Ibero-American Conference on Artificial Intelligence* (pp. 349-360). Berlin, Heidelberg: Springer Berlin Heidelberg.

Fridlund, M., Oiva, M., & Paju, P. (2020). *Digital histories: emergent approaches within the new digital history*. Helsinki University Press.

Gibbs, F., & Owens, T. (2012). Building Better Digital Humanities Tools: Toward broader audiences and user-centered designs. *Digital Humanities Quarterly*, 6(2).

Hitchcock, T., & Shoemaker, R. (2015). Making history online. *Transactions of the Royal Historical Society*, 25, 75-93.

Ιακωβάκης, Α. (2019). Η εξέλιξη των οντολογιών στον κλάδο του ΤΕΙ Δυτικής Ελλάδας.

Κατσουλάκου, Ε. Α., Μαΐστρος, Α. Α., & Τσουτσούρη, Σ. Α. (2015). Εισαγωγή στον σημασιολογικό παγκόσμιο ιστό και ανάλυση οντολογιών OWL.

Kelly, B. L. (2013). Historical research in the digital age: Examining issues relating to incorporation of expanding online resources and their impact on contemporary archaeology.

Κουτσιώρας, Β. (2018). *Αναζήτηση Εικόνων βάσει Σημασιολογικής Ομοιότητας* (Doctoral dissertation, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης).

Kuny, T. (1997). The digital dark ages? Challenges in the preservation of electronic information. *International preservation news*, 17, 8-13.

Malviya, N., Mishra, N., & Sahu, S. (2011). Developing university ontology using protégé owl tool: Process and reasoning. *International Journal of Scientific & Engineering Research*, 2(9), 1-8.

Μαυρογιάννη, Σ. (2007). *Semantic web tutorial* (Master's thesis).

Mehra, M., & Kumar, N. (2011). Semantic Web Applications. *DESIDOC Journal of Library & Information Technology*, 31(4).

Μερτής, Α. (2010). *Μελέτη τεχνολογιών σημασιολογικού ιστού και ανάπτυξη συστήματος διαχείρισης πολιτισμικών δεδομένων* (Doctoral dissertation).

Offenstadt, N., & Dufaud, G. (2005). *Les mots de l'historien*. Presses Univ. du Mirail.

Πετράκης, Α. Δ. (2011). *Σημασιολογικά κατευθυνόμενη εξόρυξη γνώσης με χρήση οντολογιών* (Master's thesis, Πανεπιστήμιο Πειραιώς).

Ψυχογιός, Α. Γ. (2012). *Σημασιολογικά δίκτυα και εφαρμογές* (Master's thesis).

Sales, M. H. (2005). Cataloging and classification for library technicians. *Journal of Adventist Education*, 67(2), 18-21.

Σαμοθρακίτου, Ζ. (2015). *Ανοιχτά Διασυνδεδεμένα Δεδομένα και βιβλιοθήκες: η διάθεση και η διασύνδεση των βιβλιογραφικών δεδομένων της Δημοτικής Βιβλιοθήκης Ευόσμου* (Master's thesis, Πανεπιστήμιο Πειραιώς).

Σαραντοπούλου, Ο. (2014). *Δημιουργία οντολογίας στη σχολική ηγεσία* (Master's thesis, Χαροκόπειο Πανεπιστήμιο Αθηνών).

Τριανταφύλλου, Π. (2022). *Ανάπτυξη οντολογίας και συναφών δεδομένων (triples) για την περιγραφή της λειτουργίας του Ανοικτού Πανεπιστημίου* (Master's thesis, Ελληνικό Ανοικτό Πανεπιστήμιο).

Τσαρτσάλη, Α. (2015). *Ο σημασιολογικός ιστός και τα ανοιχτά διασυνδεδεμένα δεδομένα στην εκπαίδευση: ανάπτυξη θεματικής πύλης δυναμικής παρουσίασης ανοικτών και διασυνδεδεμένων εκπαιδευτικών δεδομένων* (Master's thesis, Πανεπιστήμιο Πειραιώς).

Uschold, M., & Gruninger, M. (1996). Ontologies: Principles, methods and applications. *The knowledge engineering review*, 11(2), 93-136.

Uschold, M., & King, M. (1995). *Towards a methodology for building ontologies* (pp. 1-13). Edinburgh: Artificial Intelligence Applications Institute, University of Edinburgh.

Βόγλη, Ε. (2016). Τι πρέπει να γνωρίζει ο ιστορικός για την επιστήμη και το επάγγελμά του.

Ξάνθης, Θ. (2021). *Η πολιτική ιστορία της Δυτικής Λοκρίδας (6ος-4ος αι. π. Χ.)* (Master's thesis, Πανεπιστήμιο Ιωαννίνων)

Χατζησταμούλου, Β. (2021). Η εξέλιξη του παγκόσμιου ιστού και το μέλλον: Περίοδοι ανάπτυξης και τα χαρακτηριστικά τους.

Χριστίδου, Φ. (2005). Η ιδέα και τα συστατικά του σημασιολογικού ιστού, με έμφαση στις εφαρμογές σε επιχειρηματικούς τομείς.

Χρυσούλας, Χ. (2009). *Ανάπτυξη, σχεδιασμός και υλοποίηση δικτυακών συστημάτων διαχείρισης πόρων και παροχής υπηρεσιών* (Doctoral dissertation, Σχολή Πολυτεχνική. Τμήμα Ηλεκτρολόγων Μηχανικών και Τεχνολογίας Υπολογιστών, Πανεπιστήμιο Πατρών).