



ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΚΑΙ ΗΛΕΚΤΡΟΝΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ
« Σχεδιασμός και Ανάπτυξη Γνωσιακού Γράφου
Αντικειμένων Πολιτιστικής Κληρονομιάς »

Του φοιτητή
ΚΑΡΑΘΑΝΟΥ ΓΕΡΑΣΙΜΟΥ
Αρ. Μητρώου: 2020059

Επιβλέπων
ΧΑΡΑΛΑΜΠΟΣ ΜΠΡΑΤΣΑΣ
Βαθμίδα: Επίκουρος καθηγητής

Ημερομηνία: 28/05/2026

Τίτλος Δ.Ε.: Σχεδιασμός και Ανάπτυξη Γνωστικού Γράφου Αντικειμένων Πολιτιστικής Κληρονομιάς

Κωδικός Δ.Ε.: 26140

Ονοματεπώνυμο φοιτητή: Γεράσιμος Αιμίλιος Καραθάνος

Ονοματεπώνυμο εισηγητή: Χαράλαμπος Μπράτσας

Ημερομηνία ανάληψης Δ.Ε.: 11/03/2026

Ημερομηνία περάτωσης Δ.Ε. : 28/05/2026

Βεβαιώνω ότι είμαι ο συγγραφέας αυτής της εργασίας και ότι κάθε βοήθεια την οποία είχα για την προετοιμασία της είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης, έχω καταγράψει τις όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών, εικόνων και κειμένου, είτε αυτές αναφέρονται ακριβώς είτε παραφρασμένες. Επιπλέον, βεβαιώνω ότι αυτή η εργασία προετοιμάστηκε από εμένα προσωπικά, ειδικά ως διπλωματική εργασία, στο Τμήμα Μηχανικών Πληροφορικής και Ηλεκτρονικών Συστημάτων του ΔΙ.ΠΑ.Ε.

Η παρούσα εργασία αποτελεί πνευματική ιδιοκτησία του φοιτητή Γεράσιμου-Αιμίλιου Καραθάνου που την εκπόνησε. Στο πλαίσιο της πολιτικής ανοικτής πρόσβασης, ο συγγραφέας/δημιουργός εκχωρεί στο Διεθνές Πανεπιστήμιο της Ελλάδος άδεια χρήσης του δικαιώματος αναπαραγωγής, δανεισμού, παρουσίασης στο κοινό και ψηφιακής διάχυσης της εργασίας διεθνώς, σε ηλεκτρονική μορφή και σε οποιοδήποτε μέσο, για διδακτικούς και ερευνητικούς σκοπούς, άνευ ανταλλάγματος. Η ανοικτή πρόσβαση στο πλήρες κείμενο της εργασίας, δεν σημαίνει καθ' οιονδήποτε τρόπο παραχώρηση δικαιωμάτων διανοητικής ιδιοκτησίας του συγγραφέα/δημιουργού, ούτε επιτρέπει την αναπαραγωγή, αναδημοσίευση, αντιγραφή, πώληση, εμπορική χρήση, διανομή, έκδοση, μεταφόρτωση (downloading), ανάρτηση (uploading), μετάφραση, τροποποίηση με οποιονδήποτε τρόπο, τμηματικά ή περιληπτικά της εργασίας, χωρίς τη ρητή προηγούμενη έγγραφη συναίνεση του συγγραφέα/δημιουργού.

Η έγκριση της διπλωματικής εργασίας από το Τμήμα Μηχανικών Πληροφορικής και Ηλεκτρονικών Συστημάτων του Διεθνούς Πανεπιστημίου της Ελλάδος, δεν υποδηλώνει απαραίτητως και αποδοχή των απόψεων του συγγραφέα, εκ μέρους του Τμήματος.

«Αφιέρωση»

Πρόλογος

Η επιλογή του θέματος της παρούσας διπλωματικής εργασίας συνδέεται με το ιδιαίτερο ενδιαφέρον μου για τη συνάντηση της πολιτιστικής κληρονομιάς με τις τεχνολογίες του Σημασιολογικού Ιστού και της οντολογικής μοντελοποίησης. Η δυνατότητα οργάνωσης και αναπαράστασης σύνθετων πολιτιστικών και αρχαιολογικών δεδομένων μέσα από Γνωσιακούς Γράφους αποτέλεσε για μένα ένα πεδίο με έντονο θεωρητικό αλλά και πρακτικό ενδιαφέρον, καθώς συνδυάζει έννοιες όπως η γνώση, η τεκμηρίωση, η διασύνδεση και η αξιοποίηση της πληροφορίας.

Μέσα από την εκπόνηση της εργασίας αυτής είχα την ευκαιρία να εμβαθύνω σε σημαντικές οντολογίες και πρότυπα του χώρου, όπως το CIDOC CRM, να εξοικειωθώ περισσότερο με τεχνολογίες όπως το RDF και το SPARQL, και να κατανοήσω καλύτερα τις δυσκολίες αλλά και τις δυνατότητες που παρουσιάζει η σημασιολογική οργάνωση δεδομένων πολιτιστικής κληρονομιάς. Παράλληλα, η εργασία αυτή μου έδωσε τη δυνατότητα να συνδυάσω θεωρητική μελέτη με πρακτική εφαρμογή, σχεδιάζοντας και υλοποιώντας ένα συγκεκριμένο παράδειγμα Γνωσιακού Γράφου.

Συνολικά, η ενασχόληση με το συγκεκριμένο θέμα αποτέλεσε μια ιδιαίτερα χρήσιμη και δημιουργική εμπειρία, καθώς μου επέτρεψε να αναπτύξω τόσο τις τεχνικές όσο και τις ερευνητικές μου δεξιότητες σε ένα πεδίο με σύγχρονο επιστημονικό ενδιαφέρον και σημαντικές προοπτικές εφαρμογής.

Περίληψη

Η παρούσα διπλωματική εργασία εξετάζει τη σημασιολογική οργάνωση δεδομένων πολιτιστικής κληρονομιάς μέσω του σχεδιασμού και της υλοποίησης ενός Γνωσιακού Γράφου για αρχαιολογική πληροφορία. Η εργασία εστιάζει στον μετασχηματισμό ετερογενών, μη-RDF συνόλων δεδομένων σε μια σημασιολογικά δομημένη και διασυνδεδεμένη αναπαράσταση, βασισμένη σε τεχνολογίες του Σημασιολογικού Ιστού και σε αρχές οντολογικής μηχανικής.

Η προτεινόμενη προσέγγιση στηρίζεται κυρίως στο CIDOC CRM και συμπληρώνεται από επιλεγμένα στοιχεία των CRMarchaeo, GeoSPARQL, SKOS, PROV-O, DCAT και Heritage Digital Twin Ontology (HDTO). Αντί να αναπτυχθεί ένα πλήρως νέο σχήμα από την αρχή, η εργασία υιοθετεί μια στρατηγική επαναχρησιμοποίησης υφιστάμενων οντολογιών, σε συνδυασμό με περιορισμένες εφαρμοστικές εξειδικεύσεις όπου αυτό είναι αναγκαίο. Με τον τρόπο αυτό, το τελικό μοντέλο παραμένει τόσο σημασιολογικά εκφραστικό όσο και διαλειτουργικό.

Για την αξιολόγηση του μοντέλου χρησιμοποιήθηκαν δύο συμπληρωματικά use cases: το Archaeological Artefact Database of Finland (AADA), το οποίο υποστηρίζει κυρίως object-centric αναπαράσταση αρχαιολογικών ευρημάτων, και δεδομένα από το OpenContext, τα οποία αναδεικνύουν περισσότερο τη γεγονотоκεντρική διάσταση της αρχαιολογικής γνώσης, όπως ανασκαφικές δραστηριότητες, στρωματογραφικά συμφραζόμενα, actors, τεκμηριωτικές πράξεις και ψηφιακές αναπαραστάσεις. Με βάση το μοντέλο αυτό, τα επιλεγμένα δεδομένα μετασχηματίστηκαν σε RDF και οργανώθηκαν σε έναν ενιαίο Γνωσιακό Γράφο.

Η χρησιμότητα του παραγόμενου γράφου αξιολογήθηκε μέσω ερωτήσεων επάρκειας και SPARQL ερωτημάτων, τα οποία επέτρεψαν την ανάκτηση και εξερεύνηση πληροφορίας από πολλαπλά σημασιολογικά επίπεδα. Τα αποτελέσματα έδειξαν ότι η προτεινόμενη προσέγγιση υποστηρίζει συνεπή σημασιολογική αναπαράσταση και ουσιαστική εξερεύνηση αρχαιολογικών δεδομένων πέρα από την απλή περιγραφική αποθήκευσή τους.

«Ontology-Based Design and Development of a Knowledge Graph for Cultural Heritage Data»

«Gerasimos Karathanos»

Abstract

This thesis investigates the semantic organization of cultural heritage data through the design and implementation of a knowledge graph for archaeological information. The work focuses on the transformation of heterogeneous, non-RDF datasets into a semantically structured and interconnected representation based on Semantic Web technologies and ontology engineering principles.

The proposed approach is grounded primarily in CIDOC CRM and is further supported by selected elements from CRMarchaeo, GeoSPARQL, SKOS, PROV-O, DCAT and the Heritage Digital Twin Ontology (HDTO). Rather than developing an entirely new schema from scratch, the thesis adopts a reuse-oriented modeling strategy, combining established ontologies with limited application-specific extensions where required. This allows the resulting model to remain both semantically expressive and interoperable.

Two complementary use cases were employed for the evaluation of the proposed model: the Archaeological Artefact Database of Finland (AADA), which supports an object-centric representation of archaeological artefacts, and data from OpenContext, which better highlight the event-based dimension of archaeological knowledge, including excavation activities, stratigraphic contexts, actors, documentation events and digital representations. Based on this modeling framework, the selected data were transformed into RDF and organized as a knowledge graph.

The usefulness of the resulting graph was assessed through competency questions and SPARQL queries designed to retrieve and explore information across multiple semantic layers, including artefacts, materials, periods, places, excavation events, documentation and digital twins. The results show that the proposed approach supports consistent semantic representation and enables meaningful exploration of archaeological information beyond simple descriptive storage.

Overall, the thesis demonstrates that knowledge graph technologies can provide a coherent and practical framework for structuring, linking and querying cultural heritage data, while also offering a foundation for future enrichment, validation and digital heritage applications.

Ευχαριστίες

Θα ήθελα να εκφράσω τις θερμές μου ευχαριστίες προς τον επιβλέποντα καθηγητή μου και την ερευνητική του ομάδα για την πολύτιμη καθοδήγηση, τη συνεχή υποστήριξη και τις ουσιαστικές παρατηρήσεις τους, οι οποίες συνέβαλαν καθοριστικά στη βελτίωση της ποιότητας και στην ολοκλήρωση της παρούσας εργασίας.

Τέλος, θα ήθελα να ευχαριστήσω βαθιά την οικογένειά μου για τη διαρκή στήριξη, την υπομονή και την αγάπη που μου προσέφερε όλα αυτά τα φοιτητικά χρόνια. Η παρουσία τους δίπλα μου αποτέλεσε σημαντική δύναμη και στήριγμα σε κάθε δυσκολία και προσπάθεια.

Περιεχόμενα

Πρόλογος	v
Περίληψη	vi
Abstract	vii
Ευχαριστίες	viii
Περιεχόμενα	ix
Κατάλογος Σχημάτων	xi
Κατάλογος Πινάκων	xi
Συντομογραφίες	xii
Κεφάλαιο 1ο: Εισαγωγή	1
1.1 Αντικείμενο της εργασίας	1
1.2 Στόχοι της εργασίας	2
1.3 Σημασία του θέματος	4
1.4 Δομή της εργασίας	5
Κεφάλαιο 2ο: Θεωρητικό Υπόβαθρο	7
2.1 Πολιτιστική κληρονομιά και ψηφιακή αναπαράσταση	7
2.2 Γνωστικοί Γράφοι και βασικές έννοιες	9
2.3 RDF, οντολογίες, OWL και σημασιολογικά δεδομένα	10
2.4 Εισαγωγή στα ερωτήματα SPARQL	11
2.5 Μετατροπή και έλεγχος δεδομένων	12
Κεφάλαιο 3ο: Μελέτη των Σχετικών Οντολογιών	15
3.1 Το CIDOC CRM	16
3.2 Το CRMarchaeo	17
3.3 Το Heritage Digital Twin Ontology (HDTO)	18
3.4 Σύγκριση και επιλογή προσέγγισης	20
Κεφάλαιο 4ο: Σχεδιασμός της Εργασίας	22
4.1 Ανάλυση του προβλήματος	22
4.2 Παρουσίαση των επιλεγμένων συνόλων δεδομένων	24
4.2.1 Παρουσίαση του συνόλου δεδομένων AADA	24
4.2.2 Παρουσίαση του OpenContext ως δεύτερου use case	26
4.3 Βασικά σενάρια που θα μελετηθούν	26

4.4	Σχεδιασμός του μοντέλου δεδομένων	29
4.4.1	Γενική σχεδιαστική λογική	29
4.4.2	Βασικές οντότητες του μοντέλου	30
4.4.3	Βασικές σχέσεις του προτεινόμενου μοντέλου	32
4.4.4	Εννοιολογικό διάγραμμα.....	34
4.4.5	Use cases και ενδεικτική εφαρμογή σε δεδομένα	34
4.4.6	Ενδεικτικά RDF παραδείγματα.....	35
Κεφάλαιο 5ο:	Υλοποίηση του Γνωσιακού Γράφου.....	39
5.1	Επιλογή και δημιουργία συνόλου δεδομένων.....	39
5.2	Μετατροπή των δεδομένων σε RDF	41
5.3	Οργάνωση και αποθήκευση του Γνωσιακού Γράφου.....	44
5.4	Έλεγχος ορθότητας των δεδομένων.....	46
5.5	Εργαλεία που χρησιμοποιήθηκαν.....	48
Κεφάλαιο 6ο:	Εξερεύνηση Δεδομένων & Παρουσίαση Αποτελεσμάτων.....	50
6.1	Δημιουργία ερωτημάτων SPARQL	51
6.2	Παραδείγματα ερωτημάτων	52
6.3	Παρουσίαση και σχολιασμός αποτελεσμάτων.....	55
6.4	Πρωτότυπο και παρουσίαση του Γνωσιακού Γράφου.....	59
6.5	Ενδεικτικός έλεγχος εφαρμοσιμότητας σε τρίτο περιβάλλον: SearchCulture.gr.....	62
Κεφάλαιο 7ο:	Συμπεράσματα & Μελλοντικές Επεκτάσεις.....	65
7.1	Βασικά συμπεράσματα.....	65
7.2	Συνεισφορά της εργασίας.....	66
7.3	Πιθανές επεκτάσεις.....	68
Βιβλιογραφία	71	
Παραρτήματα	72	
Παράρτημα Α:	Ενδεικτικά RDF παραδείγματα.....	72
Παράρτημα Β:	Ερωτήματα SPARQL.....	84
Παράρτημα Γ:	Πλήρης αναπαράσταση της οντολογίας σε Turtle	92

Κατάλογος Σχημάτων

Σχήμα 1: Γενική ροή από δεδομένα σε Γνωσιακό Γράφο	7
Σχήμα 2: Ρόλος των οντολογιών στην εργασία	16
Σχήμα 3: Από το πρόβλημα στην υλοποίηση	23
Σχήμα 4: Εννοιολογικό μοντέλο του Γνωσιακού Γράφου	34
Σχήμα 5: Επίπεδα της προτεινόμενης οντολογίας	38
Σχήμα 6: Διάγραμμα ροής διαδικασίας υλοποίησης	42
Σχήμα 7: Από τα CQs στα SPARQL ερωτήματα και στα ανακτώμενα αποτελέσματα	52
Σχήμα 8: Παράδειγμα πολυεπίπεδης σημασιολογικής εξερεύνησης του artefact "Bone Ref# 1160" ...	56
Σχήμα 9: Ενδεικτική πρωτοτυπική παρουσίαση του Γνωσιακού Γράφου	60

Κατάλογος Πινάκων

Πίνακας 1: Βασικές τεχνολογίες του θεωρητικού υποβάθρου	14
Πίνακας 2: Σύγκριση CIDOC CRM, CRMarchaeo και HDTO	21
Πίνακας 3: Βασικά Χαρακτηριστικά του AADA	24
Πίνακας 4: Γιατί επιλέχθηκε το AADA	25
Πίνακας 5: Προτεινόμενα 14 Competency Questions	28
Πίνακας 6: Βασικοί εννοιολογικοί ρόλοι του προτεινόμενου μοντέλου και οντολογική αντιστοίχιση	31
Πίνακας 7: Βασικές σχέσεις του προτεινόμενου μοντέλου	33
Πίνακας 8: Τοπικές shortcut ιδιότητες και αιτιολόγηση χρήσης	34
Πίνακας 9: Ενδεικτική αντιστοίχιση use cases με την οντολογία	35
Πίνακας 10: Σύνολα δεδομένων και ρόλος τους στην υλοποίηση	41
Πίνακας 11: Ενδεικτικές αντιστοιχίσεις πεδίων των δεδομένων σε RDF και οντολογικούς όρους	43
Πίνακας 12: Εσωτερική οργάνωση του τελικού αρχείου TTL	45
Πίνακας 13: Έλεγχοι ορθότητας και αποτελέσματα	48
Πίνακας 14: Εργαλεία που χρησιμοποιήθηκαν στην υλοποίηση	49
Πίνακας 15: Αντιστοίχιση competency questions με τα βασικά στοιχεία του μοντέλου	53
Πίνακας 16: Κατηγοριοποίηση των ερωτημάτων SPARQL της εργασίας	54
Πίνακας 17: Ενδεικτικά αποτελέσματα ερωτημάτων SPARQL	56
Πίνακας 18: Σχολιασμός αποτελεσμάτων ανά κατηγορία ερωτήματος	59
Πίνακας 19: Βασικά συστατικά του πρωτότυπου και ο ρόλος τους	60
Πίνακας 20: Ενδεικτικός έλεγχος κάλυψης πεδίων SearchCulture.gr από το προτεινόμενο μοντέλο ..	62

Συντομογραφίες

Συντομογραφία	Επεξήγηση
AADA	Archaeological Artefact Database of Finland
AI	Artificial Intelligence
APA	American Psychological Association
CIDOC CRM	Conceptual Reference Model of the International Committee for Documentation
CQ	Competency Question
CQs	Competency Questions
CRMarchaeo	Archaeological extension of CIDOC CRM
DCAT	Data Catalog Vocabulary
DCTERMS	Dublin Core Terms
Δ.Ε.	Διπλωματική Εργασία
ΔΙ.ΠΑ.Ε.	Διεθνές Πανεπιστήμιο της Ελλάδος
GeoSPARQL	Geospatial extension of SPARQL / geospatial RDF vocabulary
HC2	Heritage Digital Twin
HC3	Tangible Heritage Entity
HDTO	Heritage Digital Twin Ontology
OWL	Web Ontology Language
OWL 2	Web Ontology Language 2
PROV-O	Provenance Ontology
RDF	Resource Description Framework
RML	RDF Mapping Language
SHACL	Shapes Constraint Language
SKOS	Simple Knowledge Organization System
SPARQL	SPARQL Protocol and RDF Query Language
TTL	Turtle
URI	Uniform Resource Identifier
URIs	Uniform Resource Identifiers
W3C	World Wide Web Consortium
WKT	Well-Known Text

Κεφάλαιο 1ο: Εισαγωγή

Η πολιτιστική κληρονομιά αποτελεί ένα πεδίο στο οποίο η πληροφορία είναι εκ φύσεως πολυδιάστατη, ετερογενής και έντονα εξαρτημένη από το εννοιολογικό πλαίσιο μέσα στο οποίο ερμηνεύεται. Τα πολιτιστικά και αρχαιολογικά δεδομένα δεν περιορίζονται σε απλές περιγραφές αντικειμένων, αλλά περιλαμβάνουν σχέσεις με τόπους, χρονικές περιόδους, κατηγοριοποιήσεις, πρόσωπα, ανασκαφικές δραστηριότητες, τεκμηριωτικές διαδικασίες και ψηφιακές αναπαραστάσεις. Η πολυπλοκότητα αυτή καθιστά συχνά ανεπαρκείς τις συμβατικές μορφές αποθήκευσης και οργάνωσης της πληροφορίας, ιδίως όταν ζητούμενο δεν είναι μόνο η καταγραφή των δεδομένων, αλλά και η διασύνδεση, η ερμηνεία και η επαναχρησιμοποίησή τους.

Στο πλαίσιο αυτό, οι τεχνολογίες του Σημασιολογικού Ιστού και ειδικότερα οι Γνωσιακοί Γράφοι προσφέρουν ένα κατάλληλο περιβάλλον για την οργάνωση της πολιτιστικής πληροφορίας με τρόπο περισσότερο σημασιολογικά συνεπή και υπολογιστικά αξιοποιήσιμο. Μέσω της χρήσης RDF, οντολογιών και σημασιολογικών ερωτημάτων, καθίσταται δυνατή η αναπαράσταση όχι μόνο μεμονωμένων δεδομένων, αλλά και των σχέσεων που τα συνδέουν, κάτι ιδιαίτερα σημαντικό για τον χώρο της πολιτιστικής κληρονομιάς. Παράλληλα, η αξιοποίηση καθιερωμένων οντολογιών, όπως το CIDOC CRM και οι σχετικές επεκτάσεις του, επιτρέπει τη διαμόρφωση μοντέλων που δεν είναι απομονωμένα ή ad hoc, αλλά εντάσσονται σε ευρύτερο πλαίσιο διαλειτουργικότητας.

Η παρούσα εργασία κινείται ακριβώς σε αυτό το σημείο συνάντησης ανάμεσα στην πολιτιστική κληρονομιά, στην οντολογική μοντελοποίηση και στις τεχνολογίες του Σημασιολογικού Ιστού. Αντικείμενό της είναι ο σχεδιασμός και η υλοποίηση ενός μικρής κλίμακας Γνωσιακού Γράφου για δεδομένα πολιτιστικής κληρονομιάς, με έμφαση σε αρχαιολογικά δεδομένα και στη σημασιολογική τους οργάνωση. Η εργασία δεν περιορίζεται στην τεχνική μετατροπή ενός συνόλου δεδομένων σε RDF, αλλά επιδιώκει να αναπτύξει ένα συνεκτικό οντολογικό μοντέλο, να το εφαρμόσει σε πραγματικά δεδομένα και να διερευνήσει στην πράξη τη χρησιμότητά του μέσω ερωτημάτων SPARQL.

Για την επίτευξη αυτού του στόχου επιλέχθηκαν δύο συμπληρωματικά use cases. Το πρώτο αφορά το Archaeological Artefact Database of Finland (AADA), το οποίο προσφέρεται κυρίως για object-centric αναπαράσταση αρχαιολογικών ευρημάτων. Το δεύτερο βασίζεται σε δεδομένα από το OpenContext και αξιοποιείται για να αναδείξει πιο καθαρά τη γεγονοκεντρική και τεκμηριωτική διάσταση του μοντέλου. Η συνδυαστική χρήση των δύο αυτών περιπτώσεων επιτρέπει στην εργασία να κινηθεί πέρα από μια μονοδιάστατη λογική μοντελοποίησης και να εξετάσει πιο ουσιαστικά τη δυνατότητα σύνδεσης αντικειμένων, γεγονότων, τόπων, actors, documentation events και ψηφιακών αναπαραστάσεων μέσα σε ενιαίο σημασιολογικό πλαίσιο.

Στις ενότητες που ακολουθούν παρουσιάζονται το αντικείμενο και το πλαίσιο της εργασίας, οι στόχοι και τα ερευνητικά της ερωτήματα, η βασική μεθοδολογική προσέγγιση που ακολουθήθηκε, καθώς και η συνολική δομή της μελέτης.

1.1 Αντικείμενο της εργασίας

Το αντικείμενο της παρούσας εργασίας είναι η σημασιολογική οργάνωση δεδομένων πολιτιστικής κληρονομιάς μέσω της ανάπτυξης ενός Γνωσιακού Γράφου βασισμένου σε οντολογίες του

Σημασιολογικού Ιστού. Ειδικότερα, η εργασία εστιάζει σε δεδομένα αρχαιολογικού ενδιαφέροντος και εξετάζει με ποιον τρόπο αυτά μπορούν να μετασχηματιστούν από συμβατικές, μη-RDF μορφές αποθήκευσης σε ένα διασυνδεδεμένο και σημασιολογικά συνεπές μοντέλο αναπαράστασης γνώσης.

Το πλαίσιο της εργασίας διαμορφώνεται από δύο βασικές παρατηρήσεις. Η πρώτη είναι ότι τα πολιτιστικά δεδομένα χαρακτηρίζονται από υψηλή εννοιολογική πολυπλοκότητα. Ένα αρχαιολογικό εύρημα, για παράδειγμα, δεν αποτελεί απλώς μια εγγραφή με ορισμένα πεδία, αλλά συνδέεται με τύπους, υλικά, περιόδους, τόπους, ενδεχομένως με γεγονότα εύρεσης ή παραγωγής, με πρόσωπα που συμμετείχαν στην τεκμηρίωσή του και με ψηφιακές αναπαραστάσεις που το αποδίδουν. Η δεύτερη παρατήρηση είναι ότι τα δεδομένα αυτά συχνά διατίθενται σε μορφές που είναι χρήσιμες για αποθήκευση ή δημοσίευση, αλλά όχι πάντοτε επαρκείς για σημασιολογική διασύνδεση και σύνθετη εξερεύνηση.

Η εργασία τοποθετείται επομένως στο πεδίο της εφαρμοσμένης οντολογικής μηχανικής για την πολιτιστική κληρονομιά. Δεν επιχειρεί να καλύψει ολόκληρο το φάσμα των πολιτιστικών δεδομένων ούτε να προτείνει ένα καθολικό μοντέλο για κάθε πιθανή περίπτωση. Αντίθετα, υιοθετεί ένα ελεγχόμενο και ρεαλιστικό εύρος, μέσα στο οποίο μπορεί να αναπτυχθεί και να αξιολογηθεί μια μικρής κλίμακας αλλά συνεκτική οντολογία, επαρκής ώστε να υποστηρίξει αντιπροσωπευτικές περιπτώσεις πολιτιστικής και αρχαιολογικής πληροφορίας.

Στο πλαίσιο αυτό, η εργασία βασίζεται κυρίως στο CIDOC CRM ως θεμελιώδες εννοιολογικό υπόβαθρο. Το CIDOC CRM επιλέγεται επειδή αποτελεί το πλέον αναγνωρισμένο εννοιολογικό μοντέλο για την αναπαράσταση πληροφορίας πολιτιστικής κληρονομιάς και επιτρέπει την απόδοση όχι μόνο αντικειμένων αλλά και γεγονότων, χωροχρονικών σχέσεων, τεκμηρίωσης και συσχετίσεων ανάμεσα σε ετερογενείς οντότητες. Παράλληλα, όπου αυτό κρίνεται αναγκαίο, το μοντέλο συμπληρώνεται από στοιχεία του CRMarchaeo για αρχαιολογικό context, του GeoSPARQL για γεωμετρίες και χωρική αναπαράσταση [17], του SKOS για ελεγχόμενα λεξιλόγια και ιεραρχίες εννοιών [14], καθώς και από επιλεγμένα στοιχεία provenance και ψηφιακής αναπαράστασης [15][18].

Η επιλογή των δεδομένων εντάσσεται επίσης στο ίδιο μεθοδολογικό πλαίσιο. Το AADA χρησιμοποιείται ως βασικό παράδειγμα επειδή παρέχει πραγματικά tabular δεδομένα για αρχαιολογικά ευρήματα και προσφέρεται για object-centric μοντελοποίηση. Το OpenContext, αντίθετα, λειτουργεί ως συμπληρωματική περίπτωση εφαρμογής που επιτρέπει να αναδειχθούν περισσότερο οι σχέσεις ανάμεσα σε artefacts, excavation events, stratigraphic contexts, actors και documentation processes. Με τον τρόπο αυτό, η εργασία δεν αντιμετωπίζει τα σύνολα δεδομένων ως απλές πηγές εισόδου, αλλά ως μέσα, μέσα από τα οποία ελέγχεται η επάρκεια και η λειτουργικότητα του οντολογικού μοντέλου.

Συνεπώς, το πλαίσιο της εργασίας είναι διττό: από τη μία πλευρά βρίσκεται το θεωρητικό και τεχνολογικό υπόβαθρο του Σημασιολογικού Ιστού και της οντολογικής μοντελοποίησης, και από την άλλη η πρακτική ανάγκη οργάνωσης πραγματικών δεδομένων πολιτιστικής κληρονομιάς σε μορφή που να υποστηρίζει διασύνδεση, επαναχρησιμοποίηση και σημασιολογική εξερεύνηση. Η παρούσα μελέτη επιχειρεί να γεφυρώσει τα δύο αυτά επίπεδα μέσα από ένα συγκεκριμένο, εφαρμόσιμο και αξιολογήσιμο παράδειγμα.

1.2 Στόχοι της εργασίας

Βασικός σκοπός της παρούσας εργασίας είναι ο σχεδιασμός και η υλοποίηση ενός μικρής κλίμακας Γνωσιακού Γράφου για δεδομένα πολιτιστικής κληρονομιάς, με έμφαση στην οντολογική αναπαράσταση αρχαιολογικής πληροφορίας και στη σημασιολογική αξιοποίησή της μέσω

τεχνολογιών του Σημασιολογικού Ιστού. Η εργασία επιδιώκει να δείξει πώς δεδομένα που προέρχονται από συμβατικές, μη-RDF μορφές μπορούν να μετασχηματιστούν σε ένα διασυνδεδεμένο σημασιολογικό σχήμα, το οποίο δεν αποθηκεύει απλώς πληροφορία, αλλά αποδίδει και τις σχέσεις που της προσδίδουν νόημα.

Ο σκοπός αυτός εξειδικεύεται σε μια σειρά από επιμέρους στόχους. Πρώτος στόχος είναι η μελέτη και επιλογή κατάλληλου οντολογικού πλαισίου για την αναπαράσταση πολιτιστικών και αρχαιολογικών δεδομένων. Στο πλαίσιο αυτό, η εργασία εξετάζει τη δυνατότητα αξιοποίησης του CIDOC CRM ως βασικού εννοιολογικού πυρήνα και τη συμπλήρωσή του με πρόσθετα πρότυπα, όπως το CRMarchaeo, το GeoSPARQL [17], το SKOS [14], καθώς και επιλεγμένα στοιχεία provenance και ψηφιακής αναπαράστασης [15][18]. Ο στόχος δεν είναι μόνο να επιλεγούν θεωρητικά κατάλληλες οντολογίες, αλλά να αξιολογηθεί στην πράξη κατά πόσο αυτές επαρκούν για την κάλυψη των αναγκών των επιλεγμένων δεδομένων.

Δεύτερος στόχος είναι η ανάπτυξη μιας συνεκτικής οντολογίας εφαρμογής, η οποία να μην αποτελεί απλή μηχανική αντανάκλαση ενός μόνο συνόλου δεδομένων, αλλά ένα πιο γενικό μοντέλο ικανό να εφαρμοστεί σε περισσότερες από μία περιπτώσεις πολιτιστικής πληροφορίας. Η εργασία επιδιώκει να διαμορφώσει ένα μοντέλο που να καλύπτει όχι μόνο τα ίδια τα αρχαιολογικά ευρήματα και τα περιγραφικά τους χαρακτηριστικά, αλλά και τις σχέσεις τους με περιόδους, τόπους, γεγονότα, actors, τεκμηριωτικές πράξεις και ψηφιακές αναπαραστάσεις. Επομένως, ένας βασικός στόχος είναι η συγκρότηση μιας πολυεπίπεδης σημασιολογικής δομής και όχι μόνο ενός απλού σχήματος καταγραφής αντικειμένων.

Τρίτος στόχος είναι η εφαρμογή της οντολογίας σε πραγματικά δεδομένα. Για τον λόγο αυτό, η εργασία δεν μένει σε θεωρητικό επίπεδο, αλλά αξιοποιεί δύο συγκεκριμένα και συμπληρωματικά use cases: το AADA και το OpenContext. Μέσα από αυτά επιδιώκεται να φανεί κατά πόσο το μοντέλο μπορεί πράγματι να απορροφήσει τόσο object-centric όσο και event-based πολιτιστική πληροφορία. Ο στόχος αυτός είναι ιδιαίτερα σημαντικός, καθώς επιτρέπει να αξιολογηθεί η λειτουργικότητα της οντολογίας σε συνθήκες που προσεγγίζουν πραγματικά προβλήματα σημασιολογικής οργάνωσης δεδομένων.

Τέταρτος στόχος είναι ο μετασχηματισμός των επιλεγμένων δεδομένων σε RDF και η συγκρότηση ενός λειτουργικού Γνωστικού Γράφου. Η εργασία επιδιώκει να δείξει όχι μόνο ποιο μοντέλο είναι κατάλληλο, αλλά και πώς αυτό μπορεί να υλοποιηθεί πρακτικά. Αυτό σημαίνει ότι στόχος της δεν είναι απλώς η περιγραφή ενός θεωρητικού mapping, αλλά η παραγωγή RDF τριάδων, η οργάνωση των δεδομένων σε συνεκτικό γράφο και η απόδοση των κατάλληλων σχέσεων ανάμεσα στις οντότητες που προκύπτουν από τα datasets.

Πέμπτος στόχος είναι η αξιολόγηση της χρησιμότητας του παραγόμενου Γνωστικού Γράφου μέσω συγκεκριμένων πληροφοριακών αναγκών. Για τον σκοπό αυτό, η εργασία αξιοποιεί ερωτήσεις επάρκειας (competency questions) ως μηχανισμό σύνδεσης των απαιτήσεων με το οντολογικό μοντέλο και στη συνέχεια μετατρέπει τα ερωτήματα αυτά σε SPARQL queries. Με τον τρόπο αυτό επιδιώκεται να αξιολογηθεί στην πράξη αν ο γράφος μπορεί να υποστηρίξει ουσιαστική εξερεύνηση της πληροφορίας και όχι μόνο στατική αποθήκευσή της. Η δυνατότητα ανάκτησης συνδυαστικής πληροφορίας από διαφορετικά επίπεδα του μοντέλου αποτελεί, επομένως, κεντρικό στόχο της εργασίας.

Έκτος στόχος είναι η ανάδειξη της αξίας της επαναχρησιμοποίησης και της ελεγχόμενης επέκτασης υπαρχουσών οντολογιών. Η εργασία επιδιώκει να δείξει ότι ένα λειτουργικό μοντέλο για την

πολιτιστική κληρονομιά δεν χρειάζεται να βασίζεται σε μαζική παραγωγή νέων κλάσεων και ιδιοτήτων, αλλά σε προσεκτική επιλογή ανάμεσα σε επαναχρησιμοποίηση, εξειδίκευση και περιορισμένη εισαγωγή τοπικών όρων όπου αυτό είναι πραγματικά αναγκαίο. Στο σημείο αυτό ο στόχος της εργασίας είναι και μεθοδολογικός, καθώς συνδέεται με μια πιο ώριμη προσέγγιση μηχανικής οντολογιών.

Συνολικά, οι παραπάνω στόχοι συγκροτούν μια εργασία που κινείται ταυτόχρονα σε θεωρητικό, σχεδιαστικό και εφαρμοστικό επίπεδο. Από τη μία πλευρά, επιχειρείται η κατανόηση και η αξιοποίηση κατάλληλων σημασιολογικών προτύπων. Από την άλλη, αναπτύσσεται ένα συγκεκριμένο οντολογικό μοντέλο και υλοποιείται ένας πραγματικός Γνωσιακός Γράφος πάνω σε επιλεγμένα δεδομένα. Τέλος, το αποτέλεσμα αξιολογείται μέσω ερωτημάτων και πρωτοτυπικής παρουσίασης, ώστε να καταδειχθεί η πρακτική σημασία της προσέγγισης.

1.3 Σημασία του θέματος

Το κεντρικό ερευνητικό ερώτημα της παρούσας εργασίας μπορεί να διατυπωθεί ως εξής: με ποιον τρόπο μπορούν ετερογενή δεδομένα πολιτιστικής κληρονομιάς, και ειδικότερα αρχαιολογικά δεδομένα που προέρχονται από διαφορετικές πηγές και μη-RDF μορφές, να οργανωθούν οντολογικά και να μετασχηματιστούν σε έναν Γνωσιακό Γράφο που να υποστηρίζει συνεπή σημασιολογική αναπαράσταση και ουσιαστική εξερεύνηση της πληροφορίας;

Το ερώτημα αυτό δεν είναι αμιγώς τεχνικό. Δεν αφορά μόνο το πώς ένα σύνολο δεδομένων μπορεί να μετατραπεί σε RDF, αλλά το πώς μπορεί να επιλεγεί και να εφαρμοστεί ένα κατάλληλο εννοιολογικό σχήμα, ικανό να αποδώσει όχι μόνο τα ίδια τα δεδομένα, αλλά και τις σχέσεις που συγκροτούν το νόημά τους. Με αυτή την έννοια, το ερευνητικό ερώτημα κινείται ταυτόχρονα σε τρία επίπεδα: στο επίπεδο της εννοιολογικής μοντελοποίησης, στο επίπεδο της τεχνικής υλοποίησης και στο επίπεδο της πρακτικής αξιοποίησης του παραγόμενου αποτελέσματος.

Ειδικότερα, το ερώτημα της εργασίας επιμερίζεται σε ορισμένα επιμέρους ζητήματα. Ένα πρώτο ζήτημα είναι κατά πόσο υφιστάμενες οντολογίες του χώρου, και κυρίως το CIDOC CRM μαζί με συναφείς επεκτάσεις και συμπληρωματικά πρότυπα, επαρκούν για την κάλυψη των αναγκών των επιλεγμένων δεδομένων. Ένα δεύτερο ζήτημα είναι σε ποιο βαθμό απαιτούνται τοπικές εφαρμοστικές εξειδικεύσεις, χωρίς όμως να διαρρηγνύεται η διαλειτουργικότητα με τα καθιερωμένα σημασιολογικά μοντέλα. Ένα τρίτο ζήτημα είναι αν ο παραγόμενος γράφος μπορεί πράγματι να λειτουργήσει ως χρήσιμο εργαλείο διερεύνησης, απαντώντας σε συγκεκριμένα competency questions μέσω SPARQL.

Η βασική προσέγγιση που ακολουθεί η εργασία για την απάντηση αυτού του ερωτήματος είναι σταδιακή και συνθετική. Αρχικά, μελετάται το θεωρητικό και τεχνολογικό υπόβαθρο των Γνωσιακών Γράφων, του RDF και των βασικών οντολογιών που σχετίζονται με την πολιτιστική κληρονομιά. Στη συνέχεια, επιλέγονται δύο συγκεκριμένα σύνολα δεδομένων, τα οποία λειτουργούν ως αντιπροσωπευτικές περιπτώσεις εφαρμογής και επιτρέπουν να εξεταστούν διαφορετικές όψεις της πολιτιστικής πληροφορίας. Ακολουθεί ο σχεδιασμός του οντολογικού μοντέλου, η αντιστοίχιση των δεδομένων στις κατάλληλες κλάσεις και ιδιότητες, η παραγωγή των RDF δεδομένων και η συγκρότηση του Γνωσιακού Γράφου.

Η προσέγγιση αυτή δεν είναι ούτε πλήρως data-driven ούτε αποκλειστικά ontology-driven. Αντίθετα, επιχειρεί να ισορροπήσει ανάμεσα στα χαρακτηριστικά των πραγματικών δεδομένων και στις εννοιολογικές απαιτήσεις ενός σημασιολογικά συνεπούς μοντέλου. Από τη μία πλευρά, η εργασία λαμβάνει σοβαρά υπόψη τη δομή, τις ιδιαιτερότητες και τους περιορισμούς των επιλεγμένων συνόλων δεδομένων. Από την άλλη, αποφεύγει να ταυτίσει το οντολογικό σχήμα με την εκάστοτε μορφή των

πηγών, επιδιώκοντας τη διαμόρφωση ενός πιο γενικού πυρήνα που να μπορεί να υποστηρίξει περισσότερες από μία περιπτώσεις εφαρμογής.

Κομβικό στοιχείο της βασικής προσέγγισης είναι επίσης η χρήση competency questions ως πρακτικού εργαλείου οντολογικού σχεδιασμού. Τα ερωτήματα αυτά δεν λειτουργούν απλώς ως τελικές ερωτήσεις αναζήτησης, αλλά ως μηχανισμός μέσω του οποίου αποσαφηνίζονται οι απαιτήσεις του μοντέλου: ποιες οντότητες πρέπει να υπάρχουν, ποιες σχέσεις οφείλουν να αποδίδονται και ποιες μορφές πληροφορίας πρέπει να μπορούν να ανακτηθούν. Με τον τρόπο αυτό, η οντολογία δεν σχεδιάζεται αφηρημένα, αλλά προσανατολίζεται εξ αρχής σε συγκεκριμένες ανάγκες σημασιολογικής αναπαράστασης και εξερεύνησης.

Επιπλέον, η εργασία υιοθετεί μια προσέγγιση προσεκτικής επαναχρησιμοποίησης υπάρχοντων σημασιολογικών προτύπων. Αντί να αναπτύσσει εκτεταμένα νέες κλάσεις και ιδιότητες, επιλέγει κατά προτεραιότητα την άμεση χρήση καθιερωμένων οντολογικών στοιχείων, ενώ εισάγει τοπικούς όρους μόνο όπου αυτό είναι απολύτως αναγκαίο για εφαρμοστικούς λόγους. Η προσέγγιση αυτή στηρίζεται στην παραδοχή ότι ένα καλό οντολογικό μοντέλο για πολιτιστικά δεδομένα πρέπει να είναι όχι μόνο εκφραστικό, αλλά και σαφές, περιορισμένο και διαλειτουργικό.

Τέλος, η βασική προσέγγιση της εργασίας ολοκληρώνεται με την αξιολόγηση του αποτελέσματος μέσω της εκτέλεσης SPARQL ερωτημάτων και της πρωτοτυπικής παρουσίας του γράφου. Έτσι, η μελέτη δεν μένει στο επίπεδο του σχεδιασμού και της υλοποίησης, αλλά επιχειρεί να δείξει στην πράξη ότι ο παραγόμενος Γνωσιακός Γράφος μπορεί να υποστηρίξει ουσιαστική εξερεύνηση της πληροφορίας. Με αυτή τη λογική, το ερευνητικό ερώτημα της εργασίας απαντάται όχι μόνο θεωρητικά, αλλά και μέσω ενός συγκεκριμένου, υλοποιημένου και ελεγχόμενου παραδείγματος.

1.4 Δομή της εργασίας

Η μεθοδολογία που ακολουθήθηκε στην παρούσα εργασία είναι σταδιακή και συνδυάζει θεωρητική διερεύνηση, οντολογικό σχεδιασμό, πρακτική υλοποίηση και αξιολόγηση του αποτελέσματος. Η επιλογή αυτής της μεθοδολογικής πορείας συνδέεται με τη φύση του ίδιου του προβλήματος, καθώς η ανάπτυξη ενός Γνωσιακού Γράφου για δεδομένα πολιτιστικής κληρονομιάς δεν μπορεί να περιοριστεί ούτε σε μια αμιγώς θεωρητική ανάλυση ούτε σε έναν καθαρά τεχνικό μετασχηματισμό δεδομένων. Αντίθετα, απαιτεί τη διαδοχική σύνδεση εννοιολογικών επιλογών, μοντελοποιητικών αποφάσεων και πρακτικών βημάτων υλοποίησης.

Σε πρώτο στάδιο, πραγματοποιήθηκε η θεωρητική μελέτη των βασικών εννοιών και τεχνολογιών που σχετίζονται με το αντικείμενο της εργασίας. Το στάδιο αυτό περιλάμβανε την εξέταση των Γνωσιακών Γράφων, του RDF, των οντολογιών και των μηχανισμών σημασιολογικής αναπαράστασης της γνώσης, καθώς και τη μελέτη καθιερωμένων οντολογιών του πεδίου της πολιτιστικής κληρονομιάς. Ιδιαίτερη έμφαση δόθηκε στο CIDOC CRM, καθώς και σε συμπληρωματικά πρότυπα, όπως το CRMarchaeo, το GeoSPARQL [17], το SKOS [14], το PROV-O [15], το DCAT [16] και επιλεγμένα στοιχεία της HDTO [18]. Το στάδιο αυτό ήταν αναγκαίο, ώστε να διαμορφωθεί το κατάλληλο εννοιολογικό υπόβαθρο πάνω στο οποίο θα στηριζόταν ο σχεδιασμός της οντολογίας.

Σε δεύτερο στάδιο, επιλέχθηκαν και μελετήθηκαν τα σύνολα δεδομένων που θα λειτουργούσαν ως περιπτώσεις εφαρμογής της εργασίας. Η επιλογή του AADA και του OpenContext δεν έγινε τυχαία, αλλά με γνώμονα τη συμπληρωματικότητά τους ως προς τις πληροφοριακές ανάγκες του μοντέλου. Το AADA προσέφερε ένα περισσότερο object-centric περιβάλλον αναπαράστασης, ενώ το OpenContext επέτρεψε να αναδειχθούν πιο καθαρά η γεγονотоκεντρική, τεκμηριωτική και ψηφιακή διάσταση της πολιτιστικής πληροφορίας. Στο σημείο αυτό πραγματοποιήθηκε και η ανάλυση της

δομής των δεδομένων, των βασικών πεδίων τους και των περιορισμών τους, ώστε να καταστεί δυνατή η αντιστοίχισή τους σε οντολογικούς όρους.

Σε τρίτο στάδιο, σχεδιάστηκε η προτεινόμενη οντολογία της εργασίας. Ο σχεδιασμός αυτός βασίστηκε στην αρχή της μέγιστης δυνατής επαναχρησιμοποίησης υφιστάμενων οντολογικών στοιχείων, με περιορισμένη και τεκμηριωμένη εισαγωγή τοπικών εφαρμοστικών όρων όπου αυτό κρίθηκε αναγκαίο. Η διαδικασία αυτή δεν ακολούθησε μια άκαμπτη, πλήρως τυποποιημένη μεθοδολογία μηχανικής οντολογιών, αλλά αξιοποίησε competency questions ως πρακτικό μηχανισμό αποσαφήνισης απαιτήσεων και καθορισμού των βασικών οντοτήτων, σχέσεων και επιπέδων του μοντέλου. Με τον τρόπο αυτό, η οντολογία διαμορφώθηκε ως ενδιάμεσος κρίκος ανάμεσα στις θεωρητικές αρχές των χρησιμοποιούμενων οντολογιών και στα πραγματικά χαρακτηριστικά των επιλεγμένων δεδομένων.

Σε τέταρτο στάδιο, πραγματοποιήθηκε η υλοποίηση του Γνωσιακού Γράφου μέσω του μετασχηματισμού των επιλεγμένων δεδομένων σε RDF. Το στάδιο αυτό περιλάμβανε τη δημιουργία αντιστοιχίσεων ανάμεσα στα πεδία των συνόλων δεδομένων και στις κλάσεις και ιδιότητες του μοντέλου, την παραγωγή RDF τριάδων, την οργάνωση του παραγόμενου γράφου και την ενσωμάτωση ενδεικτικών παραδειγμάτων που αναδεικνύουν τόσο την object-centric όσο και τη γεγονοκεντρική λογική της οντολογίας. Στο ίδιο πλαίσιο, λήφθηκε μέριμνα ώστε η υλοποίηση να παραμένει συνεπής με τη θεωρητική και σχεδιαστική λογική που είχε προηγουμένως διαμορφωθεί.

Σε πέμπτο στάδιο, η λειτουργικότητα του παραγόμενου μοντέλου αξιολογήθηκε μέσω της εκτέλεσης ερωτημάτων SPARQL. Τα ερωτήματα αυτά σχεδιάστηκαν με βάση τα competency questions και στόχευσαν στην ανάκτηση πληροφορίας από διαφορετικά επίπεδα του γράφου, όπως identifiers, types, materials, periods, places, excavation events, stratigraphic contexts, documentation events και digital representations. Με τον τρόπο αυτό, η αξιολόγηση δεν περιορίστηκε στην εσωτερική συνοχή της οντολογίας, αλλά επεκτάθηκε και στη χρηστικότητα της ως εργαλείου σημασιολογικής εξερεύνησης.

Η συνολική δομή της εργασίας ακολουθεί την παραπάνω μεθοδολογική πορεία. Στο δεύτερο κεφάλαιο παρουσιάζεται το θεωρητικό υπόβαθρο της εργασίας, με έμφαση στους Γνωσιακούς Γράφους, στο RDF, στις οντολογίες και στις τεχνολογίες του Σημασιολογικού Ιστού. Στο τρίτο κεφάλαιο εξετάζονται οι βασικές οντολογίες και τα πρότυπα που σχετίζονται με την πολιτιστική κληρονομιά και αποτελούν τη θεωρητική βάση της μοντελοποίησης. Στο τέταρτο κεφάλαιο παρουσιάζεται ο σχεδιασμός της εργασίας, η ανάλυση του προβλήματος, τα επιλεγμένα σύνολα δεδομένων, τα competency questions και ο σχεδιασμός του οντολογικού μοντέλου. Στο πέμπτο κεφάλαιο περιγράφεται η υλοποίηση του Γνωσιακού Γράφου, δηλαδή η επιλογή και προετοιμασία των δεδομένων, ο RDF μετασχηματισμός, η οργάνωση του γράφου, τα εργαλεία που χρησιμοποιήθηκαν και οι διαδικασίες ελέγχου. Στο έκτο κεφάλαιο παρουσιάζεται η εξερεύνηση του γράφου μέσω SPARQL, τα αποτελέσματα των ερωτημάτων και η πρωτοτυπική παρουσίαση του παραγόμενου μοντέλου. Τέλος, στο έβδομο κεφάλαιο συνοψίζονται τα βασικά συμπεράσματα της εργασίας, αναδεικνύεται η συνεισφορά της και συζητούνται πιθανές μελλοντικές επεκτάσεις.

Με βάση τα παραπάνω, η εργασία οργανώνεται έτσι ώστε να ακολουθεί μια συνεχή λογική μετάβασης από τη θεωρία στον σχεδιασμό, από τον σχεδιασμό στην υλοποίηση και από την υλοποίηση στην αξιολόγηση. Η διάρθρωση αυτή αποσκοπεί στο να καταστήσει σαφές ότι ο παραγόμενος Γνωσιακός Γράφος δεν αποτελεί ένα αποσπασματικό τεχνικό αποτέλεσμα, αλλά την κορύφωση μιας διαδοχικής και τεκμηριωμένης ερευνητικής διαδικασίας.

Κεφάλαιο 2ο: Θεωρητικό Υπόβαθρο

Η ανάπτυξη ενός Γνωσιακού Γράφου (Knowledge Graph) για δεδομένα πολιτιστικής κληρονομιάς προϋποθέτει την κατανόηση ενός συνόλου βασικών εννοιών και τεχνολογιών που σχετίζονται με τον Σημασιολογικό Ιστό, την οντολογική μοντελοποίηση και τη διαχείριση διασυνδεδεμένων δεδομένων. Η πολιτιστική κληρονομιά αποτελεί ένα πεδίο ιδιαίτερα πλούσιο σε πληροφορία, στην οποία συνυπάρχουν αντικείμενα, τόποι, πρόσωπα, χρονικές περίοδοι, επιστημονικές ερμηνείες και τεκμηριωτικές διαδικασίες. Η αναπαράσταση αυτής της πολυπλοκότητας με τρόπο δομημένο, επαναχρησιμοποιήσιμο και σημασιολογικά συνεπή απαιτεί την αξιοποίηση κατάλληλων εννοιολογικών και τεχνολογικών εργαλείων. [1][2][3]

Στο πλαίσιο αυτό, το θεωρητικό υπόβαθρο της παρούσας εργασίας στηρίζεται σε έννοιες όπως οι Γνωσιακοί Γράφοι, το Resource Description Framework (RDF), η Web Ontology Language (OWL 2), οι οντολογίες, η γλώσσα SPARQL (SPARQL Protocol and RDF Query Language), καθώς και οι διαδικασίες μετασχηματισμού και ελέγχου των δεδομένων. Οι έννοιες αυτές δεν αποτελούν απλώς τεχνικά μέσα υλοποίησης, αλλά συνθέτουν το ευρύτερο επιστημονικό πλαίσιο μέσα στο οποίο οργανώνεται η γνώση σε μορφή κατάλληλη για σημασιολογική επεξεργασία και διερεύνηση. Ειδικά όταν το σημείο εκκίνησης είναι ένα σύνολο δεδομένων που δεν βρίσκεται εξ αρχής σε RDF μορφή, όπως συμβαίνει στην παρούσα εργασία, η κατανόηση των παραπάνω τεχνολογιών είναι απαραίτητη τόσο για τον σχεδιασμό όσο και για την υλοποίηση του τελικού Γνωσιακού Γράφου. [2][4][5][6][7][8]

Ιδιαίτερη σημασία έχει επίσης η σύνδεση αυτών των τεχνολογιών με το ίδιο το πεδίο της πολιτιστικής κληρονομιάς. Τα πολιτιστικά δεδομένα δεν είναι απλές εγγραφές πληροφορίας, αλλά φέρουν ιστορικό, επιστημονικό και ερμηνευτικό χαρακτήρα. Για τον λόγο αυτό, η σημασιολογική τους αναπαράσταση δεν αποσκοπεί μόνο στην αποθήκευση ή ανάκτηση πληροφορίας, αλλά και στην καλύτερη οργάνωση της γνώσης, στη διασύνδεση ετερογενών πηγών και στη δυνατότητα παραγωγής πιο σύνθετων και ουσιαστικών ερωτημάτων. Το θεωρητικό υπόβαθρο του κεφαλαίου αυτού λειτουργεί επομένως ως απαραίτητη γέφυρα ανάμεσα στη γενική λογική του Σημασιολογικού Ιστού και στις ειδικότερες απαιτήσεις της παρούσας εφαρμογής. [1][3]

Στόχος του κεφαλαίου είναι να παρουσιαστούν συνοπτικά, αλλά με σαφήνεια, οι βασικές έννοιες που θα αξιοποιηθούν στα επόμενα στάδια της εργασίας. Αρχικά θα εξεταστεί η σχέση της πολιτιστικής κληρονομιάς με την ψηφιακή αναπαράσταση, στη συνέχεια θα παρουσιαστεί η έννοια των Γνωσιακών Γράφων και το θεωρητικό πλαίσιο των RDF, OWL και οντολογικών μοντέλων, ενώ έπειτα θα αναλυθεί ο ρόλος της SPARQL, καθώς και των διαδικασιών μετασχηματισμού και ελέγχου των δεδομένων. Μέσα από αυτή τη θεωρητική βάση θα γίνει πιο κατανοητό το πώς θα οργανωθεί η πρακτική υλοποίηση του Γνωσιακού Γράφου που αποτελεί το κύριο αντικείμενο της εργασίας.



Σχήμα 1: Γενική ροή από δεδομένα σε Γνωσιακό Γράφο

2.1 Πολιτιστική κληρονομιά και ψηφιακή αναπαράσταση

Η πολιτιστική κληρονομιά αποτελεί ένα σύνθετο και πολυδιάστατο πεδίο, το οποίο περιλαμβάνει υλικά και άυλα τεκμήρια της ανθρώπινης δραστηριότητας, όπως αρχαιολογικά ευρήματα, μνημεία,

έργα τέχνης, ιστορικά αντικείμενα, τόπους πολιτιστικού ενδιαφέροντος, καθώς και τις σχετικές γνώσεις, ερμηνείες και πρακτικές που τα συνοδεύουν. Η αξία των στοιχείων αυτών δεν περιορίζεται μόνο στη φυσική τους υπόσταση, αλλά επεκτείνεται και στο ιστορικό, κοινωνικό, επιστημονικό και πολιτισμικό τους νόημα. Για τον λόγο αυτό, η διαχείριση και η τεκμηρίωσή τους απαιτούν μεθόδους που να μπορούν να αποδώσουν όχι μόνο τις βασικές περιγραφικές πληροφορίες, αλλά και το πλέγμα σχέσεων που τα συνδέει με πρόσωπα, τόπους, χρονικές περιόδους, γεγονότα και ερμηνευτικά πλαίσια. [1][3]

Η ψηφιακή αναπαράσταση της πολιτιστικής κληρονομιάς προέκυψε ως ανάγκη οργάνωσης, τεκμηρίωσης, διατήρησης και διάχυσης αυτής της πληροφορίας σε ψηφιακά περιβάλλοντα. Μέσα από τη χρήση πληροφοριακών συστημάτων, βάσεων δεδομένων, ψηφιακών αποθετηρίων και διαδικτυακών εφαρμογών, κατέστη δυνατή η αποθήκευση και η διαχείριση μεγάλου όγκου πολιτιστικών δεδομένων, καθώς και η ευρύτερη πρόσβαση σε αυτά από ερευνητές, οργανισμούς και ενδιαφερόμενους πολίτες. Παρ' όλα αυτά, η απλή ψηφιοποίηση ή καταγραφή δεδομένων δεν αρκεί από μόνη της για την ουσιαστική αξιοποίησή τους, ιδίως όταν τα δεδομένα προέρχονται από διαφορετικές πηγές, ακολουθούν διαφορετικά σχήματα περιγραφής ή έχουν παραχθεί με διαφορετικές μεθοδολογικές και επιστημονικές λογικές. [1][3]

Ένα από τα βασικά προβλήματα της ψηφιακής διαχείρισης πολιτιστικής πληροφορίας είναι η ετερογένεια. Τα ίδια ή παρόμοια πολιτιστικά τεκμήρια μπορεί να καταγράφονται με διαφορετικό τρόπο από διαφορετικούς φορείς, ανάλογα με τις ανάγκες και τις προτεραιότητές τους. Ένα μουσείο, για παράδειγμα, μπορεί να δίνει έμφαση στη συλλογή, στη συντήρηση και στην έκθεση ενός αντικειμένου, ενώ ένας αρχαιολόγος μπορεί να επικεντρώνεται στο ανασκαφικό συμφραζόμενο, στη χρονολόγηση και στην επιστημονική ερμηνεία του. Αντίστοιχα, ένα ερευνητικό αποθετήριο μπορεί να οργανώνει την πληροφορία γύρω από δημοσιεύσεις, μελέτες ή τεκμηριωτικά αρχεία. Η ποικιλία αυτή καθιστά δύσκολη τη διασύνδεση των δεδομένων και τη δημιουργία ενός ενιαίου και κατανοητού πλαισίου αναφοράς. [1][3]

Η ανάγκη για πιο δομημένες μορφές ψηφιακής αναπαράστασης οδηγεί στη χρήση μοντέλων που δεν περιορίζονται σε απλή αποθήκευση εγγραφών, αλλά μπορούν να αποδώσουν με σαφήνεια τις σχέσεις ανάμεσα στις οντότητες του πεδίου. Σε αυτό το σημείο αποκτά ιδιαίτερη σημασία η έννοια της σημασιολογικής αναπαράστασης, δηλαδή της οργάνωσης της πληροφορίας με τρόπο που να αποτυπώνει όχι μόνο τα δεδομένα καθαυτά, αλλά και το νόημα των σχέσεων μεταξύ τους. Η προσέγγιση αυτή είναι ιδιαίτερα χρήσιμη στην πολιτιστική κληρονομιά, όπου η κατανόηση ενός τεκμηρίου εξαρτάται σε μεγάλο βαθμό από το πλαίσιο μέσα στο οποίο εντάσσεται. [1][3]

Η ψηφιακή αναπαράσταση της πολιτιστικής κληρονομιάς δεν πρέπει επομένως να αντιμετωπίζεται απλώς ως τεχνική διαδικασία μεταφοράς δεδομένων σε ηλεκτρονική μορφή. Αντιθέτως, αποτελεί μια διαδικασία οργάνωσης και εννοιολογικής αποτύπωσης της γνώσης, η οποία επηρεάζει τον τρόπο με τον οποίο τα πολιτιστικά δεδομένα μπορούν να διασυνδεθούν, να ερμηνευθούν και να αξιοποιηθούν. Η επιλογή κατάλληλων μοντέλων αναπαράστασης και κατάλληλων τεχνολογιών είναι κρίσιμη, επειδή καθορίζει το αν τα δεδομένα θα μπορούν να επαναχρησιμοποιηθούν, να διαλειτουργήσουν με άλλες πηγές και να υποστηρίξουν ουσιαστική αναζήτηση και ανάλυση. [1][2]

Στο πλαίσιο της παρούσας εργασίας, η ψηφιακή αναπαράσταση αφορά τη μετατροπή και οργάνωση δεδομένων πολιτιστικής κληρονομιάς σε μορφή κατάλληλη για τη δημιουργία ενός Γνωσιακού Γράφου. Το ζητούμενο δεν είναι μόνο η μεταφορά ενός συνόλου δεδομένων σε νέα τεχνική μορφή, αλλά η αναδόμησή του σύμφωνα με σημασιολογικές αρχές που να επιτρέπουν καλύτερη ερμηνεία, αναζήτηση και αξιοποίηση της πληροφορίας. Η προσέγγιση αυτή συνδέεται άμεσα με τις έννοιες που

θα παρουσιαστούν στις επόμενες ενότητες, αρχίζοντας από την έννοια των Γνωσιακών Γράφων, τα οποία αποτελούν ένα βασικό μοντέλο για την αναπαράσταση και διαχείριση διασυνδεδεμένης γνώσης. [2]

2.2 Γνωσιακοί Γράφοι και βασικές έννοιες

Οι Γνωσιακοί Γράφοι αποτελούν μια προσέγγιση αναπαράστασης γνώσης κατά την οποία η πληροφορία οργανώνεται με τη μορφή διασυνδεδεμένων οντοτήτων και σχέσεων. Σε αντίθεση με πιο παραδοσιακές μορφές οργάνωσης δεδομένων, όπου η πληροφορία αποθηκεύεται συχνά σε απομονωμένους πίνακες ή ανεξάρτητα αρχεία, ένας Γνωσιακός Γράφος επιτρέπει την περιγραφή του κόσμου ως ενός δικτύου από έννοιες, αντικείμενα, πρόσωπα, τόπους, γεγονότα και τις μεταξύ τους συνδέσεις. Με αυτόν τον τρόπο, η πληροφορία δεν αντιμετωπίζεται ως απλή συλλογή δεδομένων, αλλά ως δομημένη γνώση που μπορεί να διερευνηθεί, να ερμηνευθεί και να επαναχρησιμοποιηθεί. [2]

Η βασική ιδέα ενός Γνωσιακού Γράφου είναι ότι κάθε οντότητα αποκτά νόημα όχι μόνο από τα δικά της χαρακτηριστικά, αλλά και από τις σχέσεις της με άλλες οντότητες. Για παράδειγμα, ένα αρχαιολογικό εύρημα δεν περιγράφεται μόνο από το υλικό ή την κατηγορία του, αλλά και από τον τόπο εύρεσης, τη χρονολογική περίοδο στην οποία ανήκει, την επιστημονική τεκμηρίωση που το συνοδεύει και τις πιθανές συνδέσεις του με άλλα αντικείμενα ή γεγονότα. Η λογική αυτή είναι ιδιαίτερα κατάλληλη για πεδία όπως η πολιτιστική κληρονομιά, όπου η γνώση είναι εξ ορισμού πολυσύνθετη και βασίζεται σε πλούσια συμφοραζόμενα. [2][3]

Σε γενικές γραμμές, ένας Γνωσιακός Γράφος αποτελείται από κόμβους και σχέσεις. Οι κόμβοι αντιστοιχούν σε οντότητες ή έννοιες του πεδίου, ενώ οι σχέσεις δηλώνουν τον τρόπο με τον οποίο αυτές συνδέονται. Η δομή αυτή επιτρέπει την αναπαράσταση γνώσης με πιο φυσικό και εκφραστικό τρόπο σε σχέση με τα αυστηρά μοντέλα δεδομένων πινάκων. Ταυτόχρονα, καθιστά δυνατή την πλοήγηση μέσα στο δίκτυο της πληροφορίας και την εξαγωγή πιο σύνθετων συμπερασμάτων ή απαντήσεων μέσα από σημασιολογικά ερωτήματα. [2]

Ένα από τα σημαντικότερα πλεονεκτήματα των Γνωσιακών Γράφων είναι η δυνατότητα ενοποίησης ετερογενών δεδομένων. Στην πράξη, πολλά σύνολα δεδομένων έχουν δημιουργηθεί με διαφορετικά σχήματα, διαφορετικές ονομασίες πεδίων και διαφορετικούς τρόπους τεκμηρίωσης. Μέσω ενός Γνωσιακού Γράφου, τα δεδομένα αυτά μπορούν να χαρτογραφηθούν σε ένα κοινό σημασιολογικό πλαίσιο, επιτρέποντας τη μεταξύ τους σύνδεση και τη διαλειτουργικότητά τους. Η δυνατότητα αυτή είναι ιδιαίτερα κρίσιμη στον τομέα της πολιτιστικής κληρονομιάς, όπου τα δεδομένα συχνά προέρχονται από μουσεία, αρχαιολογικά αρχεία, ερευνητικά έργα και ψηφιακές συλλογές με πολύ διαφορετικές λογικές οργάνωσης. [2][3]

Παράλληλα, οι Γνωσιακοί Γράφοι επιτρέπουν πιο ουσιαστική αναζήτηση και εξερεύνηση της πληροφορίας. Αντί ο χρήστης να περιορίζεται σε απλές αναζητήσεις συγκεκριμένων πεδίων, μπορεί να διατυπώνει ερωτήματα που βασίζονται στις σχέσεις μεταξύ των οντοτήτων. Για παράδειγμα, είναι δυνατόν να αναζητηθούν όλα τα ευρήματα μιας συγκεκριμένης περιόδου που σχετίζονται με ορισμένη κατηγορία αντικειμένων ή που προέρχονται από συγκεκριμένο τόπο. Αυτή η δυνατότητα συνιστά σημαντικό πλεονέκτημα σε ερευνητικά περιβάλλοντα, όπου η πληροφορία σπάνια είναι γραμμική ή απομονωμένη. [2]

Σημαντικό είναι επίσης ότι οι Γνωσιακοί Γράφοι δεν αποτελούν μόνο μια τεχνική δομή αποθήκευσης, αλλά και μια εννοιολογική προσέγγιση οργάνωσης της γνώσης. Για να κατασκευαστεί ένας Γνωσιακός Γράφος, δεν αρκεί απλώς να μεταφερθούν δεδομένα από έναν πίνακα σε ένα νέο σύστημα. Απαιτείται να προηγηθεί η επιλογή ή ο σχεδιασμός ενός κατάλληλου μοντέλου που να ορίζει ποιες

είναι οι βασικές οντότητες του πεδίου, πώς σχετίζονται μεταξύ τους και ποια είναι η σημασία των σχέσεων αυτών. Σε αυτό ακριβώς το σημείο συνδέονται οι Γνωσιακοί Γράφοι με τις οντολογίες, οι οποίες παρέχουν το εννοιολογικό υπόβαθρο για τη σωστή και συνεπή αναπαράσταση της γνώσης. [2]

Στην παρούσα εργασία, ο στόχος είναι η δημιουργία ενός Γνωσιακού Γράφου για δεδομένα πολιτιστικής κληρονομιάς, με αφετηρία ένα μη-RDF σύνολο δεδομένων και με βάση κατάλληλες οντολογικές αρχές. Η έννοια του Γνωσιακού Γράφου αποτελεί επομένως κεντρικό άξονα ολόκληρης της μελέτης, καθώς συνδέει τη θεωρητική μοντελοποίηση με την πρακτική υλοποίηση. Στις επόμενες ενότητες θα παρουσιαστούν οι τεχνολογίες και τα πρότυπα που επιτρέπουν την υλοποίηση αυτής της προσέγγισης, αρχίζοντας από τα RDF, τις οντολογίες και γενικότερα τα σημασιολογικά δεδομένα.[2][13]

2.3 RDF, οντολογίες, OWL και σημασιολογικά δεδομένα

Η υλοποίηση ενός Γνωσιακού Γράφου βασίζεται σε τεχνολογίες και πρότυπα που επιτρέπουν τη δομημένη και σημασιολογικά συνεπή αναπαράσταση της πληροφορίας. Στο πλαίσιο του Σημασιολογικού Ιστού, ένα από τα βασικότερα πρότυπα είναι το RDF, το οποίο χρησιμοποιείται για την περιγραφή οντοτήτων και των μεταξύ τους σχέσεων με τρόπο τυπικό και μηχανικά επεξεργάσιμο. Το RDF προσφέρει το θεμελιώδες μοντέλο αναπαράστασης δεδομένων πάνω στο οποίο μπορούν να στηριχθούν πιο σύνθετες σημασιολογικές δομές. [4]

Η βασική μονάδα έκφρασης στο RDF είναι η τριάδα, η οποία αποτελείται από υποκείμενο, ιδιότητα και αντικείμενο. Μέσα από τη λογική αυτή, η πληροφορία περιγράφεται ως ένα σύνολο δηλώσεων που συνδέουν οντότητες μεταξύ τους. Για παράδειγμα, ένα αρχαιολογικό εύρημα μπορεί να συνδέεται με έναν τόπο, μια χρονολογική περίοδο, ένα υλικό ή μια κατηγορία αντικειμένου. Μέσα από πολλές τέτοιες τριάδες συγκροτείται ένα δίκτυο σχέσεων, το οποίο μπορεί να αποτελέσει τη βάση ενός Γνωσιακού Γράφου. Το RDF επομένως επιτρέπει την αναπαράσταση της πληροφορίας με ευέλικτο και διασυνδεδεμένο τρόπο, χωρίς να περιορίζεται σε αυστηρά μοντέλα συμβατικών πινάκων. [4]

Παρά τη σημασία του, το RDF από μόνο του δεν αρκεί για την πλήρη εννοιολογική περιγραφή ενός πεδίου γνώσης. Το RDF δίνει τον βασικό μηχανισμό για να δηλώνονται σχέσεις, αλλά δεν επαρκεί για να εκφράσει πιο σύνθετους ορισμούς, ιεραρχίες κλάσεων, περιορισμούς και λογικές συσχετίσεις που είναι αναγκαίες στην οντολογική μοντελοποίηση. Για τον λόγο αυτό, στον Σημασιολογικό Ιστό χρησιμοποιείται η OWL (Web Ontology Language), και ειδικότερα η OWL 2, η οποία αποτελεί το βασικό πρότυπο γλώσσας για την τυπική αναπαράσταση οντολογιών. [4][5]

Η OWL 2 επιτρέπει τον ορισμό κλάσεων, ιδιοτήτων, ιεραρχιών, σχέσεων ισοδυναμίας, καθώς και πιο σύνθετων περιορισμών που δεν μπορούν να αποδοθούν επαρκώς μόνο με απλό RDF. Με άλλα λόγια, ενώ το RDF παρέχει τη βασική γραμματική για την αναπαράσταση δεδομένων, η OWL 2 παρέχει τον μηχανισμό για την έκφραση της εννοιολογικής δομής και της λογικής οργάνωσης της γνώσης. Στο πλαίσιο αυτό, οι οντολογίες δεν είναι απλώς λεξιλόγια όρων, αλλά τυπικά μοντέλα γνώσης που μπορούν να περιγράψουν με σαφήνεια τις βασικές έννοιες ενός πεδίου και τις σχέσεις τους. [5]

Στην πολιτιστική κληρονομιά, η χρήση οντολογιών σε OWL είναι ιδιαίτερα σημαντική, καθώς το πεδίο αυτό χαρακτηρίζεται από πολυπλοκότητα, ετερογένεια και ισχυρή εξάρτηση από το συμφραζόμενο. Ένα πολιτιστικό ή αρχαιολογικό τεκμήριο δεν περιγράφεται μόνο από φυσικά χαρακτηριστικά, αλλά και από τη σχέση του με γεγονότα, πρόσωπα, τόπους, χρονικές περιόδους και διαδικασίες τεκμηρίωσης. Για να αποδοθεί με σαφήνεια αυτό το πλέγμα σχέσεων, απαιτείται ένα μοντέλο γνώσης που να μπορεί να εκφράζει όχι μόνο τα δεδομένα, αλλά και τις έννοιες που τα οργανώνουν. Η OWL 2, σε συνδυασμό με το RDF, προσφέρει ακριβώς αυτό το πλαίσιο. [1][5][9]

Ο όρος σημασιολογικά δεδομένα αναφέρεται σε δεδομένα που είναι οργανωμένα με τρόπο ώστε να έχουν σαφές, τυπικά ορισμένο και επαναχρησιμοποιήσιμο νόημα. Σε τέτοια δεδομένα, το νόημα των οντοτήτων και των σχέσεων τους δεν είναι απλώς υπονοούμενο ή εξαρτώμενο από ένα συγκεκριμένο λογισμικό, αλλά αποτυπώνεται ρητά μέσα από κατάλληλα μοντέλα και πρότυπα. Αυτό επιτρέπει την καλύτερη διαλειτουργικότητα μεταξύ συστημάτων, την ενιαία ερμηνεία της πληροφορίας και τη δυνατότητα εκτέλεσης πιο σύνθετων σημασιολογικών ερωτημάτων. [4][5]

Στην παρούσα εργασία, η διάκριση ανάμεσα στο RDF και στην OWL είναι ουσιώδης. Το RDF θα αποτελέσει τη βασική μορφή αναπαράστασης του Γνωσιακού Γράφου, ενώ η OWL 2 θα αποτελέσει τη γλώσσα μέσω της οποίας θα οριστεί η οντολογική δομή του μοντέλου. Με άλλα λόγια, ένα από τα τελικά παραδοτέα της εργασίας δεν θα είναι μόνο ένα σύνολο RDF τριάδων, αλλά και μια οντολογία σε OWL, η οποία θα αποτυπώνει τις βασικές έννοιες και σχέσεις που απαιτούνται για την οργάνωση των δεδομένων του επιλεγμένου πεδίου. [4][5]

Επομένως, το RDF, η OWL 2 και οι οντολογίες συνιστούν από κοινού το βασικό θεωρητικό και τεχνολογικό υπόβαθρο για την ανάπτυξη του προτεινόμενου Γνωσιακού Γράφου. Πάνω σε αυτή τη βάση θα στηριχθεί τόσο ο μετασχηματισμός του επιλεγμένου συνόλου δεδομένων όσο και η τελική σημασιολογική αναπαράσταση και διερεύνησή του. Στην επόμενη ενότητα θα παρουσιαστεί η γλώσσα SPARQL, η οποία επιτρέπει την αναζήτηση και εξερεύνηση της πληροφορίας μέσα σε ένα RDF Γνωσιακό Γράφο. [4][5]

2.4 Εισαγωγή στα ερωτήματα SPARQL

Η οργάνωση της πληροφορίας σε μορφή RDF και η μοντελοποίησή της μέσω οντολογιών αποκτούν ιδιαίτερη αξία όταν συνοδεύονται από τη δυνατότητα αναζήτησης, ανάκτησης και διερεύνησης της γνώσης που έχει παραχθεί. Για τον σκοπό αυτό, στο πλαίσιο του Σημασιολογικού Ιστού χρησιμοποιείται η γλώσσα SPARQL, η οποία αποτελεί το βασικό πρότυπο για την υποβολή ερωτημάτων πάνω σε RDF δεδομένα. Η SPARQL επιτρέπει την αναζήτηση πληροφορίας όχι μόνο με βάση μεμονωμένες τιμές, αλλά και με βάση τη δομή και τις σημασιολογικές σχέσεις που έχουν οριστεί μέσα στον Γνωσιακό Γράφο. [6]

Σε αντίθεση με τις παραδοσιακές γλώσσες ερωτημάτων που εφαρμόζονται κυρίως σε σχεσιακές βάσεις δεδομένων, η SPARQL έχει σχεδιαστεί για να λειτουργεί πάνω σε δεδομένα που οργανώνονται ως γράφος. Αυτό σημαίνει ότι τα ερωτήματα δεν αφορούν απλώς γραμμές και στήλες, αλλά πρότυπα σχέσεων ανάμεσα σε οντότητες. Με άλλα λόγια, ο χρήστης δεν αναζητά μόνο τιμές συγκεκριμένων πεδίων, αλλά μπορεί να εκφράσει ερωτήματα που ακολουθούν συνδέσεις μέσα στον γράφο και να εντοπίσει πώς σχετίζονται διαφορετικές οντότητες μεταξύ τους. Αυτή η δυνατότητα είναι εξαιρετικά σημαντική στην πολιτιστική κληρονομιά, όπου το νόημα της πληροφορίας βασίζεται συχνά στις σχέσεις μεταξύ αντικειμένων, τόπων, προσώπων, γεγονότων και περιόδων. [6]

Η βασική λογική της SPARQL στηρίζεται στην αντιστοίχιση προτύπων τριάδων πάνω στον RDF γράφο. Όπως το RDF αναπαριστά την πληροφορία ως τριάδες, έτσι και η SPARQL διατυπώνει ερωτήματα που αναζητούν συνδυασμούς τέτοιων τριάδων, χρησιμοποιώντας μεταβλητές αντί για συγκεκριμένες τιμές όπου χρειάζεται. Με αυτόν τον τρόπο, η γλώσσα επιτρέπει την εξαγωγή πληροφορίας με ευέλικτο τρόπο, ανάλογα με τα ερευνητικά ή τεκμηριωτικά ερωτήματα που έχουν τεθεί. Για παράδειγμα, σε ένα γράφο πολιτιστικής κληρονομιάς μπορεί να ζητηθεί η ανάκτηση όλων των ευρημάτων που συνδέονται με μια συγκεκριμένη χρονική περίοδο, ενός συνόλου αντικειμένων που ανήκουν στην ίδια κατηγορία ή όλων των οντοτήτων που σχετίζονται με έναν συγκεκριμένο τόπο. [6]

Η SPARQL δεν περιορίζεται μόνο στην απλή ανάκτηση αποτελεσμάτων. Παρέχει δυνατότητες φιλτραρίσματος, ταξινόμησης, ομαδοποίησης και συνδυασμού συνθηκών, επιτρέποντας τη διατύπωση πιο σύνθετων ερωτημάτων. Έτσι, μπορεί να χρησιμοποιηθεί τόσο για απλές αναζητήσεις όσο και για πιο απαιτητικές μορφές διερεύνησης της πληροφορίας. Επιπλέον, η χρήση της δεν είναι σημαντική μόνο για τον τελικό χρήστη ενός Γνωσιακού Γράφου, αλλά και για την ίδια τη διαδικασία αξιολόγησης του μοντέλου, καθώς μέσω των ερωτημάτων μπορεί να φανεί αν η οντολογική μοντελοποίηση και η οργάνωση των δεδομένων υποστηρίζουν ουσιαστικά την ανάκτηση γνώσης. [6]

Στην παρούσα εργασία, η SPARQL θα χρησιμοποιηθεί ως βασικό μέσο πειραματικής εξερεύνησης του γράφου γνώσης που θα παραχθεί από το επιλεγμένο σύνολο δεδομένων. Μέσα από ενδεικτικά ερωτήματα θα επιχειρηθεί να φανεί κατά πόσο το μοντέλο που θα υλοποιηθεί επιτρέπει την αποτελεσματική αναζήτηση και συνδυαστική αξιοποίηση των δεδομένων πολιτιστικής κληρονομιάς. Με τον τρόπο αυτό, η SPARQL δεν αντιμετωπίζεται απλώς ως ένα τεχνικό εργαλείο ερωτημάτων, αλλά ως μέσο επιβεβαίωσης της χρησιμότητας του ίδιου του Γνωσιακού Γράφου. [6][13]

Η σημασία της SPARQL στην εργασία αυτή είναι επομένως διπλή. Από τη μία πλευρά, αποτελεί το βασικό εργαλείο ανάκτησης και εξερεύνησης της σημασιολογικά οργανωμένης πληροφορίας. Από την άλλη, λειτουργεί ως μέσο ελέγχου του κατά πόσο η μοντελοποίηση που θα υιοθετηθεί είναι επαρκής και συνεπής σε σχέση με τα ερωτήματα που επιδιώκεται να απαντηθούν. Γι' αυτό και η παρουσία της στο θεωρητικό υπόβαθρο είναι ουσιώδης, καθώς συνδέει άμεσα την έννοια του RDF Γνωσιακού Γράφου με την πρακτική δυνατότητα αξιοποίησής του. [6]

Συνολικά, η SPARQL αποτελεί αναπόσπαστο μέρος του οικοσυστήματος του Σημασιολογικού Ιστού και βασικό εργαλείο για κάθε εφαρμογή που στηρίζεται σε RDF και οντολογίες. Στο πλαίσιο της παρούσας εργασίας, η χρήση της θα αναδείξει τον τρόπο με τον οποίο η σημασιολογική αναπαράσταση των δεδομένων μπορεί να υποστηρίξει ουσιαστικά ερωτήματα πάνω στην πολιτιστική πληροφορία. Στην επόμενη ενότητα θα παρουσιαστούν συνοπτικά οι διαδικασίες μετατροπής και ελέγχου των δεδομένων, οι οποίες είναι αναγκαίες για τη μετάβαση από το αρχικό σύνολο δεδομένων σε ένα συνεκτικό και αξιοποιήσιμο Γνωσιακό Γράφο. [6]

2.5 Μετατροπή και έλεγχος δεδομένων

Η δημιουργία ενός Γνωσιακού Γράφου δεν εξαρτάται μόνο από την επιλογή κατάλληλων οντολογιών και γλωσσών αναπαράστασης, αλλά και από τη διαδικασία μέσω της οποίας τα αρχικά δεδομένα μετατρέπονται σε σημασιολογικά οργανωμένη μορφή. Στην πράξη, τα περισσότερα δεδομένα δεν είναι εξαρχής διαθέσιμα ως RDF, αλλά βρίσκονται σε μορφές όπως πίνακες, αρχεία CSV, υπολογιστικά φύλλα ή άλλες συμβατικές δομές αποθήκευσης. Για τον λόγο αυτό, ένα από τα βασικά βήματα στην ανάπτυξη ενός Γνωσιακού Γράφου είναι ο μετασχηματισμός των δεδομένων αυτών σε RDF, σύμφωνα με ένα συγκεκριμένο εννοιολογικό μοντέλο. [4][7]

Η διαδικασία αυτή δεν αποτελεί απλή τεχνική μεταφορά δεδομένων από μία μορφή σε μία άλλη. Αντιθέτως, προϋποθέτει την ερμηνεία των πεδίων του αρχικού συνόλου δεδομένων, την αντιστοίχισή τους με κλάσεις και ιδιότητες της οντολογίας και τον καθορισμό του τρόπου με τον οποίο οι εγγραφές του συνόλου δεδομένων θα μετατραπούν σε οντότητες και σχέσεις μέσα στον Γνωσιακό Γράφο. Με άλλα λόγια, η μετατροπή των δεδομένων συνδέεται άμεσα με την οντολογική μοντελοποίηση, καθώς κάθε σχεδιαστική επιλογή επηρεάζει το πώς θα αναπαρασταθεί τελικά η πληροφορία και το ποιες ερωτήσεις θα μπορεί να υποστηρίξει ο παραγόμενος γράφος. [7]

Για την υποστήριξη αυτής της διαδικασίας έχουν αναπτυχθεί γλώσσες και τεχνικές χαρτογράφησης δεδομένων προς RDF. Μία από τις πιο γνωστές προσεγγίσεις είναι η RML (RDF Mapping Language),

η οποία επιτρέπει τον ορισμό κανόνων μετασχηματισμού από ετερογενείς πηγές δεδομένων προς RDF. Μέσω ενός συνόλου δηλώσεων χαρτογράφησης, η RML καθορίζει πώς οι δομές του αρχικού συνόλου δεδομένων αντιστοιχίζονται στις οντολογικές έννοιες που χρησιμοποιούνται στο μοντέλο. Με τον τρόπο αυτό, η διαδικασία μετασχηματισμού γίνεται πιο σαφής, πιο επαναλήψιμη και περισσότερο συμβατή με τις αρχές του Σημασιολογικού Ιστού. [7]

Στην παρούσα εργασία, η αναφορά στη RML έχει ιδιαίτερη σημασία, καθώς το επιλεγμένο σύνολο δεδομένων δεν βρίσκεται εξ αρχής σε RDF μορφή. Επομένως, η μετάβαση από την αρχική του δομή σε έναν RDF Γνωσιακό Γράφο προϋποθέτει έναν μηχανισμό αντιστοίχισης των πεδίων του συνόλου δεδομένων με τις κλάσεις και τις ιδιότητες που θα επιλεγούν από το μοντέλο. Ανεξάρτητα από το βάθος στο οποίο θα αξιοποιηθεί η RML στην πρακτική υλοποίηση, η θεωρητική και μεθοδολογική της σημασία είναι μεγάλη, επειδή εκφράζει με σαφήνεια τη λογική της μετατροπής δομημένων δεδομένων σε σημασιολογικά δεδομένα. [7][13]

Ωστόσο, η παραγωγή RDF δεδομένων δεν αρκεί από μόνη της για να θεωρηθεί επιτυχής η υλοποίηση ενός Γνωσιακού Γράφου. Εξίσου σημαντικός είναι ο έλεγχος της ορθότητας και της συνέπειας των παραγόμενων δεδομένων. Κατά τον μετασχηματισμό είναι δυνατόν να προκύψουν ελλείψεις, ασυνέπειες ή λανθασμένες αντιστοιχίσεις, οι οποίες μπορούν να επηρεάσουν τόσο την ποιότητα του μοντέλου όσο και τη δυνατότητα αναζήτησης της πληροφορίας. Για τον λόγο αυτό, είναι απαραίτητο να υπάρχει ένας μηχανισμός επικύρωσης, μέσω του οποίου να ελέγχεται αν τα δεδομένα ικανοποιούν ορισμένες βασικές δομικές και σημασιολογικές απαιτήσεις. [8]

Στο πλαίσιο του Σημασιολογικού Ιστού, μία από τις βασικές προσεγγίσεις για την επικύρωση RDF δεδομένων είναι η χρήση του SHACL (Shapes Constraint Language). Το SHACL επιτρέπει τον ορισμό κανόνων και περιορισμών που πρέπει να ικανοποιούν τα δεδομένα, όπως η ύπαρξη υποχρεωτικών ιδιοτήτων, η αναμενόμενη μορφή ορισμένων τιμών ή η σωστή σύνδεση μεταξύ συγκεκριμένων τύπων οντοτήτων. Με τον τρόπο αυτό, παρέχεται ένας πιο συστηματικός τρόπος ελέγχου του κατά πόσο ο παραγόμενος Γνωσιακός Γράφος συμμορφώνεται με τις βασικές απαιτήσεις του μοντέλου. [8]

Σημειώνεται ότι στην παρούσα εργασία το RML και το SHACL παρουσιάζονται ως σχετικές τεχνολογίες του οικοσυστήματος RDF και του Σημασιολογικού Ιστού, και όχι ως πλήρως υλοποιημένα εργαλεία της τελικής εφαρμογής. Η υλοποίηση του γράφου βασίστηκε σε ελεγχόμενη χειροκίνητη RDF/Turtle μοντελοποίηση, σε έλεγχο σύνταξης και σε SPARQL-based διερεύνηση των δεδομένων. Η συστηματική χρήση RML mappings ή SHACL shapes παραμένει πιθανή μελλοντική επέκταση της εργασίας.

Για την παρούσα εργασία, ο έλεγχος των δεδομένων έχει κυρίως λειτουργικό και ποιοτικό χαρακτήρα. Στόχος δεν είναι η ανάπτυξη ενός πλήρους συστήματος επικύρωσης μεγάλης κλίμακας, αλλά η διασφάλιση ότι το παραγόμενο αποτέλεσμα είναι δομικά σωστό και επαρκώς συνεπές ώστε να υποστηρίξει τη μετέπειτα εξερεύνησή του με SPARQL. Επομένως, η μετατροπή και ο έλεγχος των δεδομένων αποτελούν δύο αλληλένδετες φάσεις: η πρώτη αφορά τη μετάβαση από το αρχικό σύνολο δεδομένων σε RDF Γνωσιακό Γράφο, ενώ η δεύτερη την επιβεβαίωση ότι ο Γράφος αυτός ανταποκρίνεται ικανοποιητικά στις βασικές απαιτήσεις της μοντελοποίησης. [8]

Συνολικά, η διαδικασία μετασχηματισμού και ελέγχου των δεδομένων αποτελεί κρίσιμο στάδιο για την ανάπτυξη ενός αξιόπιστου Γνωσιακού Γράφου. Μέσα από αυτήν, η πληροφορία μεταβαίνει από μια αρχική, συμβατική μορφή σε μια πιο δομημένη και σημασιολογικά πλούσια αναπαράσταση, ενώ παράλληλα διασφαλίζεται ότι η νέα αυτή μορφή παραμένει συνεπής και αξιοποιήσιμη. Η ενότητα

αυτή ολοκληρώνει το βασικό θεωρητικό υπόβαθρο της εργασίας, το οποίο θα αποτελέσει τη βάση για τα επόμενα κεφάλαια, όπου θα παρουσιαστούν οι απαιτήσεις, τα σενάρια χρήσης και ο σχεδιασμός του προτεινόμενου μοντέλου. [7][8]

Το θεωρητικό υπόβαθρο που παρουσιάστηκε στο παρόν κεφάλαιο ανέδειξε τις βασικές έννοιες και τεχνολογίες που υποστηρίζουν την ανάπτυξη ενός Γνωσιακού Γράφου, όπως το RDF, η OWL 2, οι οντολογίες, η SPARQL, καθώς και οι διαδικασίες μετασχηματισμού και ελέγχου των δεδομένων. Η κατανόηση αυτών των στοιχείων είναι απαραίτητη, καθώς συγκροτούν το γενικό εννοιολογικό και τεχνολογικό πλαίσιο μέσα στο οποίο θα υλοποιηθεί η παρούσα εργασία. Με βάση αυτή τη θεωρητική αφετηρία, στο επόμενο κεφάλαιο παρουσιάζονται οι βασικές οντολογίες που σχετίζονται με το πεδίο της πολιτιστικής κληρονομιάς, ώστε να αποσαφηνιστεί ποιο εννοιολογικό μοντέλο είναι καταλληλότερο για τη μοντελοποίηση του προτεινόμενου Γνωσιακού Γράφου. [4][5][6][7][8]

Τεχνολογία / Έννοια	Ρόλος στην εργασία	Στάδιο αξιοποίησης
Γνωσιακός Γράφος	Τελική μορφή οργάνωσης της γνώσης	Σχεδιασμός και υλοποίηση
RDF	Αναπαράσταση δεδομένων σε τριάδες	Υλοποίηση γράφου
OWL 2	Ορισμός της οντολογικής δομής	Σχεδιασμός οντολογίας
SPARQL	Αναζήτηση και εξερεύνηση του γράφου	Αξιολόγηση / αποτελέσματα
RML	Θεωρητικά σχετική τεχνολογία για μετασχηματισμό μη-RDF δεδομένων σε RDF	Θεωρητικό υπόβαθρο / πιθανή μελλοντική επέκταση
SHACL	Θεωρητικά σχετική τεχνολογία για επικύρωση RDF γράφων	Θεωρητικό υπόβαθρο / πιθανή μελλοντική επέκταση

Πίνακας 1: Βασικές τεχνολογίες του θεωρητικού υποβάθρου

Κεφάλαιο 3ο: Μελέτη των Σχετικών Οντολογιών

Η αναπαράσταση γνώσης στον χώρο της πολιτιστικής κληρονομιάς αποτελεί μια ιδιαίτερα απαιτητική διαδικασία, καθώς τα δεδομένα του πεδίου αυτού χαρακτηρίζονται από ετερογένεια, πολυπλοκότητα και έντονη εξάρτηση από το ιστορικό, επιστημονικό και τεκμηριωτικό τους πλαίσιο. Αντικείμενα, χώροι, πρόσωπα, χρονικές περίοδοι, γεγονότα, ανασκαφικές δραστηριότητες και ερμηνευτικές προσεγγίσεις συνδέονται μεταξύ τους με πολλαπλούς τρόπους, γεγονός που καθιστά αναγκαία τη χρήση κατάλληλων εννοιολογικών μοντέλων για την οργανωμένη και σημασιολογικά συνεπή αναπαράστασή τους. [1][9]

Στο πλαίσιο αυτό, οι οντολογίες διαδραματίζουν καθοριστικό ρόλο, καθώς παρέχουν μια τυπική και κοινά κατανοητή δομή για την περιγραφή οντοτήτων, σχέσεων και γεγονότων. Μέσω αυτών, καθίσταται δυνατή όχι μόνο η οργάνωση της πληροφορίας, αλλά και η διαλειτουργικότητα μεταξύ διαφορετικών πηγών δεδομένων, η επαναχρησιμοποίηση της γνώσης και η δυνατότητα εκτέλεσης πιο σύνθετων ερωτημάτων πάνω σε έναν Γνωσιακό Γράφο. Ειδικά στον τομέα της πολιτιστικής κληρονομιάς, όπου η πληροφορία προέρχεται συχνά από μουσειακή τεκμηρίωση, αρχαιολογική έρευνα, επιστημονικές μελέτες και ψηφιακά αποθετήρια, η ύπαρξη κατάλληλων οντολογικών προτύπων είναι ουσιώδης. [1][9][10]

Για τις ανάγκες της παρούσας εργασίας επιλέχθηκαν τρεις βασικές οντολογικές προσεγγίσεις: το CIDOC CRM, το CRMarchaeo και το Heritage Digital Twin Ontology (HDTO). Η επιλογή τους δεν είναι τυχαία, αλλά βασίζεται στο γεγονός ότι καλύπτουν διαφορετικές, αλλά συμπληρωματικές, πτυχές της αναπαράστασης γνώσης στον χώρο της πολιτιστικής κληρονομιάς. Το CIDOC CRM αποτελεί το πιο καθιερωμένο εννοιολογικό πρότυπο για την περιγραφή πολιτιστικής πληροφορίας, το CRMarchaeo εξειδικεύει τη μοντελοποίηση αρχαιολογικών διαδικασιών και ανασκαφικών γεγονότων, ενώ το HDTO εκφράζει μια πιο σύγχρονη προσέγγιση που σχετίζεται με την ψηφιακή αναπαράσταση και τον εμπλουτισμό πολιτιστικών οντοτήτων σε σύγχρονα ψηφιακά περιβάλλοντα. [10][11][12]

Στόχος του κεφαλαίου αυτού είναι να παρουσιαστούν οι βασικές αρχές, η φιλοσοφία και η χρησιμότητα των παραπάνω οντολογιών, με έμφαση στα στοιχεία εκείνα που είναι πιο συναφή με την παρούσα εργασία. Δεν επιδιώκεται μια εξαντλητική ανάλυση ολόκληρου του περιεχομένου τους, αλλά μια στοχευμένη μελέτη των εννοιών και μηχανισμών που μπορούν να αξιοποιηθούν στον σχεδιασμό και την υλοποίηση ενός Γνωσιακού Γράφου για δεδομένα πολιτιστικής κληρονομιάς. Παράλληλα, μέσα από τη συγκριτική επισκόπηση των οντολογιών αυτών, επιχειρείται να αναδειχθεί ο τρόπος με τον οποίο μπορούν να συνδυαστούν ή να λειτουργήσουν συμπληρωματικά, ανάλογα με τις απαιτήσεις του υπό μελέτη σεναρίου. [1][10][11][12]

Στις επόμενες ενότητες παρουσιάζονται αρχικά το CIDOC CRM ως το βασικό εννοιολογικό υπόβαθρο της εργασίας, στη συνέχεια το CRMarchaeo ως εξειδίκευση για το αρχαιολογικό πεδίο, και τέλος το HDTO ως μια πιο πρόσφατη προσέγγιση σχετική με τον ψηφιακό μετασχηματισμό και την αναπαράσταση πολιτιστικών δεδομένων. Η ανάλυση αυτή θα αποτελέσει τη βάση για τις σχεδιαστικές επιλογές που θα ακολουθήσουν στα επόμενα κεφάλαια.



Σχήμα 2: Ρόλος των οντολογιών στην εργασία

3.1 Το CIDOC CRM

Το CIDOC CRM (Conceptual Reference Model) αποτελεί ένα από τα σημαντικότερα και πιο διαδεδομένα εννοιολογικά πρότυπα για την τεκμηρίωση και διασύνδεση δεδομένων πολιτιστικής κληρονομιάς. Αναπτύχθηκε με στόχο να παρέχει ένα κοινό σημασιολογικό πλαίσιο για την περιγραφή πολιτιστικών οντοτήτων, επιτρέποντας τη διαλειτουργικότητα μεταξύ ετερογενών συστημάτων, βάσεων δεδομένων και οργανισμών που δραστηριοποιούνται στον χώρο της μουσειακής, αρχαιακής και αρχαιολογικής τεκμηρίωσης. [9][10]

Η βασική φιλοσοφία του CIDOC CRM δεν είναι απλώς η καταγραφή αντικειμένων ως στατικών οντοτήτων, αλλά η αναπαράσταση της γνώσης γύρω από αυτά μέσα από ένα πλέγμα σχέσεων, γεγονότων και συμφραζομένων. Με άλλα λόγια, ένα πολιτιστικό αντικείμενο δεν αντιμετωπίζεται απομονωμένα, αλλά σε συνάρτηση με την παραγωγή του, τη χρήση του, την ανακάλυψή του, τη μελέτη του, τη μετακίνησή του, καθώς και με τα πρόσωπα, τους τόπους και τις χρονικές περιόδους που σχετίζονται με αυτό. Η προσέγγιση αυτή είναι ιδιαίτερα σημαντική για την πολιτιστική κληρονομιά, όπου η ερμηνεία και η τεκμηρίωση βασίζονται συχνά σε σύνθετες αλληλεπιδράσεις μεταξύ φυσικών αντικειμένων, ιστορικών γεγονότων και επιστημονικών παρατηρήσεων. [9][10]

Ένα από τα βασικά χαρακτηριστικά του CIDOC CRM είναι ότι βασίζεται σε μια γεγοντοκεντρική λογική μοντελοποίησης. Αντί να περιορίζεται μόνο στην περιγραφή χαρακτηριστικών ενός αντικειμένου, δίνει έμφαση στα γεγονότα και στις διαδικασίες μέσω των οποίων η γνώση για το αντικείμενο αυτό παράγεται και οργανώνεται. Για παράδειγμα, η δημιουργία ενός αντικειμένου, η εύρεσή του σε έναν αρχαιολογικό χώρο, η ταξινόμησή του από έναν ειδικό ή η ένταξή του σε μια συλλογή μπορούν να αποδοθούν ως ξεχωριστά γεγονότα που συνδέουν μεταξύ τους πρόσωπα, τόπους, χρόνους και υλικά τεκμήρια. Με αυτόν τον τρόπο, το μοντέλο επιτρέπει πιο εκφραστική και πλούσια σημασιολογική περιγραφή της πληροφορίας. [9][10]

Σε δομικό επίπεδο, το CIDOC CRM οργανώνεται γύρω από κλάσεις και ιδιότητες, οι οποίες επιτρέπουν την αναπαράσταση εννοιών του πεδίου και των σχέσεών τους. Οι κλάσεις χρησιμοποιούνται για να περιγράψουν βασικούς τύπους οντοτήτων, όπως αντικείμενα, πρόσωπα,

τόπους, χρονικές οντότητες, γεγονότα και εννοιολογικά αντικείμενα, ενώ οι ιδιότητες χρησιμοποιούνται για να δηλώσουν τις μεταξύ τους συνδέσεις. Η λογική αυτή είναι ιδιαίτερα συμβατή με τη μοντελοποίηση σε RDF, καθώς μπορεί να μεταφραστεί με φυσικό τρόπο σε τριάδες υποκειμένου–ιδιότητας–αντικειμένου (Subject, Predicate, Object), κάτι που την καθιστά κατάλληλη για την ανάπτυξη Γνωσιακών Γράφων. [9][10]

Ένα ακόμη σημαντικό πλεονέκτημα του CIDOC CRM είναι ότι λειτουργεί ως εννοιολογική “γέφυρα” ανάμεσα σε διαφορετικές πηγές και μορφές πολιτιστικής πληροφορίας. Σε πραγματικά περιβάλλοντα πολιτιστικής κληρονομιάς, τα δεδομένα συχνά προέρχονται από διαφορετικά ιδρύματα και έχουν παραχθεί με διαφορετικές λογικές τεκμηρίωσης. Ένα μουσείο μπορεί να καταγράφει ένα αντικείμενο με βάση τη συλλογή του, ένας αρχαιολόγος με βάση το ανασκαφικό του πλαίσιο και ένας ερευνητής με βάση τη μελέτη ή την ερμηνεία του. Το CIDOC CRM δεν επιδιώκει να αντικαταστήσει αυτά τα επιμέρους σχήματα, αλλά να προσφέρει ένα κοινό σημασιολογικό επίπεδο πάνω στο οποίο μπορούν να αντιστοιχιστούν και να συνδεθούν. [9][10]

Για την παρούσα εργασία, το CIDOC CRM αποτελεί τη βασική αφετηρία της μοντελοποίησης, καθώς παρέχει τις κεντρικές έννοιες που απαιτούνται για την περιγραφή αρχαιολογικών και πολιτιστικών δεδομένων σε ένα συνεκτικό σχήμα γνώσης. Μέσα από αυτό μπορούν να αναπαρασταθούν στοιχεία όπως τα αρχαιολογικά ευρήματα, οι τόποι προέλευσης ή εύρεσής τους, οι χρονικές περίοδοι στις οποίες εντάσσονται, οι διαδικασίες τεκμηρίωσης και οι σχέσεις τους με πρόσωπα ή οργανισμούς. Επιπλέον, η ευρεία αποδοχή του μοντέλου στη διεθνή κοινότητα της πολιτιστικής κληρονομιάς το καθιστά ιδιαίτερα κατάλληλο για μια εργασία που επιδιώκει να κινηθεί με βάση καθιερωμένα πρότυπα και όχι ad hoc λύσεις. [9][10]

Παρόλο που το CIDOC CRM είναι ισχυρό και ευέλικτο, δεν έχει σχεδιαστεί για να καλύπτει με την ίδια λεπτομέρεια όλες τις εξειδικευμένες ανάγκες κάθε υποπεδίου. Για παράδειγμα, στην περίπτωση της αρχαιολογίας, όπου υπάρχουν έντονα διαδικαστικά στοιχεία όπως οι ανασκαφικές δραστηριότητες, τα στρωματογραφικά συμφραζόμενα και οι παρατηρήσεις πεδίου, απαιτούνται συχνά πιο εξειδικευμένες επεκτάσεις. Για τον λόγο αυτό, στο επόμενο υποκεφάλαιο εξετάζεται το CRMarchaeo, το οποίο επεκτείνει τη λογική του CIDOC CRM με στόχο να υποστηρίξει με μεγαλύτερη ακρίβεια την αρχαιολογική τεκμηρίωση. [10][11]

3.2 Το CRMarchaeo

Το CRMarchaeo αποτελεί μια εξειδικευμένη επέκταση του CIDOC CRM, η οποία αναπτύχθηκε με σκοπό να καλύψει πιο συγκεκριμένες ανάγκες του αρχαιολογικού πεδίου. Ενώ το CIDOC CRM παρέχει ένα γενικό και ιδιαίτερα ισχυρό εννοιολογικό πλαίσιο για την αναπαράσταση δεδομένων πολιτιστικής κληρονομιάς, στην περίπτωση της αρχαιολογίας προκύπτουν συχνά πρόσθετες απαιτήσεις που σχετίζονται με τη φύση της ανασκαφικής διαδικασίας, τη στρωματογραφική παρατήρηση, την τεκμηρίωση πεδίου και τη συσχέτιση των ευρημάτων με το ακριβές συμφραζόμενό τους. Το CRMarchaeo επιχειρεί να καλύψει ακριβώς αυτό το κενό. [11]

Η αρχαιολογική τεκμηρίωση δεν περιορίζεται μόνο στην περιγραφή ενός ευρήματος ως φυσικού αντικειμένου, αλλά δίνει ιδιαίτερη σημασία στο πλαίσιο μέσα στο οποίο αυτό εντοπίστηκε. Η πληροφορία για το πού βρέθηκε ένα αντικείμενο, σε ποιο στρώμα, σε ποια ανασκαφική τομή, υπό ποιες συνθήκες, στο πλαίσιο ποιας ανασκαφικής πράξης και με ποια τεκμηριωτική διαδικασία, έχει συχνά εξίσου μεγάλη ή και μεγαλύτερη σημασία από το ίδιο το εύρημα. Για τον λόγο αυτό, η αρχαιολογία απαιτεί ένα μοντέλο που να δίνει έμφαση όχι μόνο στα αντικείμενα, αλλά και στις

διαδικασίες παρατήρησης, ανασκαφής, καταγραφής και ερμηνείας. Το CRMarchaeo έρχεται να προσφέρει ακριβώς αυτή τη δυνατότητα. [11]

Κεντρική επιδίωξη του CRMarchaeo είναι η μοντελοποίηση της ανασκαφής ως επιστημονικής και τεκμηριωτικής διαδικασίας. Στο πλαίσιο αυτό, το ενδιαφέρον μετατοπίζεται από το “τι είναι το αντικείμενο” και στο “πώς εντοπίστηκε, παρατηρήθηκε και τεκμηριώθηκε”. Με τον τρόπο αυτό, δίνεται η δυνατότητα να περιγραφούν πιο αναλυτικά οι σχέσεις ανάμεσα σε αρχαιολογικά ευρήματα, στρωματογραφικές ενότητες, τόπους, γεγονότα ανασκαφής και επιστημονικές πράξεις ερμηνείας. Πρόκειται επομένως για μια περισσότερο διαδικαστική και συμφραζομενική προσέγγιση της γνώσης, η οποία ανταποκρίνεται καλύτερα στις πραγματικές ανάγκες της αρχαιολογικής έρευνας. [11]

Σε σύγκριση με το CIDOC CRM, το CRMarchaeo μπορεί να θεωρηθεί ως ένα μοντέλο που εισάγει μεγαλύτερο βαθμό εξειδίκευσης στην περιγραφή αρχαιολογικών δεδομένων. Δεν αντικαθιστά το CIDOC CRM, αλλά βασίζεται σε αυτό και το επεκτείνει. Με άλλα λόγια, διατηρεί τη γενική φιλοσοφία της γεγοντοκεντρικής μοντελοποίησης, αλλά την εφαρμόζει με πιο συγκεκριμένο τρόπο σε έννοιες όπως η ανασκαφική δραστηριότητα, τα αρχαιολογικά συμφραζόμενα, τα στάδια τεκμηρίωσης και οι διαδικασίες δημιουργίας αρχαιολογικής γνώσης. Αυτό το καθιστά ιδιαίτερα χρήσιμο σε περιπτώσεις όπου ο ερευνητής θέλει να αποδώσει όχι μόνο τελικά αποτελέσματα, αλλά και τη λογική παραγωγής τους. [11]

Για την παρούσα εργασία, η σημασία του CRMarchaeo έγκειται κυρίως στο ότι προσφέρει ένα πιο κατάλληλο θεωρητικό πλαίσιο για τη σύνδεση των αρχαιολογικών δεδομένων με τις διαδικασίες μέσα από τις οποίες αυτά αποκτούν νόημα. Παρόλο που το σύνολο δεδομένων AADA επικεντρώνεται κυρίως σε αρχαιολογικά ευρήματα και όχι σε πλήρεις ανασκαφικές ροές εργασίας, η μελέτη του CRMarchaeo παραμένει σημαντική, επειδή βοηθά να γίνει κατανοητό πώς θα μπορούσε να επεκταθεί το μοντέλο σε πιο σύνθετα αρχαιολογικά σενάρια. Επομένως, στο πλαίσιο της παρούσας διπλωματικής, το CRMarchaeo δεν θα αξιοποιηθεί απαραίτητα στο ίδιο βάθος με το CIDOC CRM, αλλά θα λειτουργήσει ως συμπληρωματικό θεωρητικό εργαλείο που ενισχύει τη σύνδεση της εργασίας με την αρχαιολογική διάσταση του πεδίου. [11][13]

Η αξία του CRMarchaeo είναι ιδιαίτερα εμφανής όταν το ζητούμενο δεν είναι μόνο η αποθήκευση πολιτιστικών δεδομένων, αλλά και η ακριβέστερη τεκμηρίωση της αρχαιολογικής διαδικασίας. Σε ένα πιο σύνθετο σενάριο, όπου θα υπήρχαν αναλυτικά δεδομένα για σημεία εύρεσης, στρωματογραφικές σχέσεις, φάσεις ανασκαφής, παρατηρήσεις πεδίου και τεκμηριωτικά συμβάντα, η χρήση του θα ήταν ακόμη πιο ουσιαστική. Επομένως, η παρουσίασή του στο πλαίσιο της παρούσας εργασίας συμβάλλει όχι μόνο στη θεωρητική πληρότητα του κεφαλαίου, αλλά και στην ανάδειξη μελλοντικών κατευθύνσεων επέκτασης του μοντέλου. [11]

Συνολικά, το CRMarchaeo μπορεί να θεωρηθεί ως ένα πολύτιμο εργαλείο για την πιο λεπτομερή και αρχαιολογικά ακριβή μοντελοποίηση γνώσης. Παρότι η παρούσα εργασία δεν θα εξαντλήσει όλες τις δυνατότητές του, η ενσωμάτωσή του στη μελέτη των σχετικών οντολογιών είναι χρήσιμη, επειδή αναδεικνύει τη μετάβαση από μια γενική περιγραφή πολιτιστικής κληρονομιάς σε μια πιο εξειδικευμένη αναπαράσταση αρχαιολογικής πληροφορίας. Στην επόμενη ενότητα θα εξεταστεί το Heritage Digital Twin Ontology (HDTO), το οποίο εκφράζει μια διαφορετική και πιο σύγχρονη οπτική ως προς την ψηφιακή αναπαράσταση πολιτιστικών οντοτήτων. [11]

3.3 Το Heritage Digital Twin Ontology (HDTO)

Το Heritage Digital Twin Ontology (HDTO) αποτελεί μια πιο σύγχρονη οντολογική προσέγγιση για την αναπαράσταση δεδομένων πολιτιστικής κληρονομιάς σε ψηφιακά περιβάλλοντα. Σε αντίθεση με

πιο καθιερωμένα πρότυπα, όπως το CIDOC CRM, τα οποία έχουν ως βασικό στόχο τη σημασιολογική περιγραφή και διαλειτουργικότητα της πολιτιστικής πληροφορίας, το HDTO συνδέεται περισσότερο με τη λογική των ψηφιακών διδύμων (Digital Twins) και της δυναμικής, πολυεπίπεδης αναπαράστασης πολιτιστικών οντοτήτων σε ψηφιακή μορφή. Η προσέγγιση αυτή αντανακλά τη γενικότερη τάση για εμπλουτισμένες ψηφιακές αναπαραστάσεις μνημείων, αντικειμένων, χώρων και δεδομένων τεκμηρίωσης. [12]

Η έννοια του Digital Twin αναφέρεται, σε γενικές γραμμές, σε ένα ψηφιακό ισοδύναμο ή ψηφιακό αντίστοιχο μιας πραγματικής οντότητας, το οποίο μπορεί να συγκεντρώνει, να οργανώνει και να ενοποιεί διαφορετικά είδη πληροφορίας που σχετίζονται με αυτήν. Στο πεδίο της πολιτιστικής κληρονομιάς, αυτή η λογική αποκτά ιδιαίτερο ενδιαφέρον, καθώς τα πολιτιστικά τεκμήρια δεν περιορίζονται μόνο στη φυσική τους διάσταση, αλλά συνδέονται με ιστορικά, τεκμηριωτικά, γεωχωρικά, ερμηνευτικά και πολυμεσικά δεδομένα. Το HDTO επιχειρεί να υποστηρίξει ακριβώς αυτού του είδους την πολυδιάστατη αναπαράσταση. [12]

Σε σχέση με το CIDOC CRM και το CRMarchaeo, το HDTO εκφράζει μια διαφορετική έμφαση. Ενώ τα δύο προηγούμενα μοντέλα είναι περισσότερο προσανατολισμένα στη σημασιολογική οργάνωση της γνώσης και στη μοντελοποίηση γεγονότων, τεκμηριωτικών διαδικασιών και αρχαιολογικών συμφραζομένων, το HDTO συνδέεται πιο άμεσα με τις ανάγκες της ψηφιακής αναπαράστασης, της ενοποίησης πολλαπλών πηγών πληροφορίας και της δημιουργίας πιο σύνθετων ψηφιακών περιβαλλόντων. Αυτό σημαίνει ότι μπορεί να λειτουργήσει υποστηρικτικά σε περιπτώσεις όπου ένα πολιτιστικό αντικείμενο ή ένας χώρος δεν περιγράφεται μόνο ως τεκμήριο, αλλά και ως κόμβος σύνδεσης πολλών τύπων δεδομένων. [12]

Για την παρούσα εργασία, η μελέτη του HDTO έχει κυρίως συμπληρωματικό και προσανατολιστικό ρόλο. Το βασικό βάρος της μοντελοποίησης θα δοθεί στο CIDOC CRM, ενώ το CRMarchaeo συμβάλλει στη σύνδεση με την αρχαιολογική διάσταση. Το HDTO δεν επιλέγεται ως κύριος άξονας υλοποίησης, αλλά ως μια χρήσιμη επιπλέον οπτική που δείχνει προς πιο σύγχρονες κατευθύνσεις διαχείρισης και ψηφιακής αναπαράστασης πολιτιστικών δεδομένων. Με αυτόν τον τρόπο, το κεφάλαιο δεν περιορίζεται μόνο σε καθιερωμένα πρότυπα τεκμηρίωσης, αλλά λαμβάνει υπόψη και την εξέλιξη του πεδίου προς πιο εμπλουτισμένα και διασυνδεδεμένα ψηφιακά οικοσυστήματα. [12]

Η σημασία του HDTO γίνεται πιο εμφανής όταν η πολιτιστική κληρονομιά προσεγγίζεται όχι μόνο ως αντικείμενο επιστημονικής περιγραφής, αλλά και ως πεδίο ψηφιακής εμπειρίας, εικονικής αναπαράστασης και σύνδεσης ετερογενών πηγών τεκμηρίωσης. Σε τέτοια περιβάλλοντα, η πληροφορία μπορεί να αφορά όχι μόνο το ίδιο το αντικείμενο, αλλά και τρισδιάστατα μοντέλα, γεωχωρικά δεδομένα, αρχεία αισθητήρων, ιστορικές πηγές, αποκαταστάσεις, ψηφιακές απεικονίσεις ή άλλες μορφές τεκμηρίωσης. Παρόλο που η παρούσα εργασία δεν φτάνει σε αυτό το επίπεδο πολυπλοκότητας, η παρουσία του HDTO στη μελέτη των σχετικών οντολογιών είναι χρήσιμη, επειδή αναδεικνύει μια σημαντική μελλοντική προοπτική για τον τομέα. [12]

Στην παρούσα εργασία, η συμβολή του HDTO έγκειται κυρίως στο ότι ενισχύει τη θεωρητική πληρότητα της μελέτης και αναδεικνύει ότι η αναπαράσταση γνώσης στην πολιτιστική κληρονομιά δεν περιορίζεται μόνο στις παραδοσιακές μορφές τεκμηρίωσης. Παράλληλα, βοηθά να τοποθετηθεί η εργασία σε ένα ευρύτερο πλαίσιο, όπου οι Γνωσιακοί Γράφοι μπορούν να λειτουργήσουν όχι μόνο ως εργαλεία οργάνωσης δεδομένων, αλλά και ως βάση για πιο σύνθετες ψηφιακές εφαρμογές. Επομένως, ακόμη και αν το HDTO δεν αξιοποιηθεί εκτεταμένα στο πρακτικό μέρος της υλοποίησης, η αναφορά του στο θεωρητικό και οντολογικό υπόβαθρο της εργασίας είναι δικαιολογημένη και χρήσιμη. [12]

Συνοψίζοντας, το HDTO αντιπροσωπεύει μια νεότερη και περισσότερο προσανατολισμένη στο ψηφιακό περιβάλλον οντολογική κατεύθυνση, η οποία συμπληρώνει τα καθιερωμένα πρότυπα του πεδίου. Για την παρούσα εργασία δεν αποτελεί το βασικό εργαλείο μοντελοποίησης, αλλά λειτουργεί ως ένα πρόσθετο θεωρητικό σημείο αναφοράς που επιτρέπει να κατανοηθεί καλύτερα η σύγχρονη εξέλιξη της ψηφιακής πολιτιστικής κληρονομιάς. Στην επόμενη ενότητα θα επιχειρηθεί μια συνοπτική σύγκριση των τριών οντολογιών, ώστε να αποσαφηνιστεί ο ρόλος που μπορεί να διαδραματίσει η καθεμία στο πλαίσιο της παρούσας μελέτης. [12]

3.4 Σύγκριση και επιλογή προσέγγισης

Οι τρεις οντολογικές προσεγγίσεις που παρουσιάστηκαν στις προηγούμενες ενότητες, δηλαδή το CIDOC CRM, το CRMarchaeo και το HDTO, δεν λειτουργούν ανταγωνιστικά, αλλά καλύπτουν διαφορετικές ανάγκες και επίπεδα αναπαράστασης της γνώσης στην πολιτιστική κληρονομιά. Η συγκριτική εξέτασή τους είναι απαραίτητη, ώστε να αποσαφηνιστεί ο ρόλος που μπορεί να διαδραματίσει καθεμία στην παρούσα εργασία και να τεκμηριωθεί η τελική επιλογή της μοντελοποιητικής προσέγγισης. [10][11][12]

Το CIDOC CRM αποτελεί το πιο γενικό και καθιερωμένο εννοιολογικό πλαίσιο από τα τρία. Το βασικό του πλεονέκτημα είναι ότι προσφέρει μια ευρέως αποδεκτή σημασιολογική βάση για την περιγραφή πολιτιστικών οντοτήτων, γεγονότων, προσώπων, τόπων και χρονικών σχέσεων. Χάρη στον αφαιρετικό και διαλειτουργικό του χαρακτήρα, είναι ιδιαίτερα κατάλληλο για την οργάνωση ετερογενών δεδομένων και για τη μετατροπή τους σε ένα συνεκτικό Γνωσιακό Γράφο. Για μια εργασία που στοχεύει στη δημιουργία RDF Γνωσιακού Γράφου από ένα μη-RDF σύνολο δεδομένων, το CIDOC CRM προσφέρει την πιο σταθερή και κατάλληλη εννοιολογική αφετηρία. [9][10]

Το CRMarchaeo, από την άλλη πλευρά, είναι σαφώς πιο εξειδικευμένο και επικεντρώνεται περισσότερο στην αρχαιολογική τεκμηρίωση και στη μοντελοποίηση της ανασκαφικής διαδικασίας. Το βασικό του πλεονέκτημα είναι ότι αποδίδει με μεγαλύτερη ακρίβεια στοιχεία όπως το αρχαιολογικό συμφραζόμενο, η ανασκαφή ως διαδικασία, οι στρωματογραφικές σχέσεις και η παραγωγή αρχαιολογικής γνώσης μέσα από παρατήρηση και τεκμηρίωση. Ωστόσο, αυτή η μεγαλύτερη εξειδίκευση σημαίνει και αυξημένη πολυπλοκότητα. Για τον λόγο αυτό, η αξιοποίησή του εξαρτάται σε μεγάλο βαθμό από το αν το διαθέσιμο σύνολο δεδομένων περιλαμβάνει τέτοιου τύπου πληροφορία σε ικανοποιητικό βάθος. [11]

Στην παρούσα εργασία, το επιλεγμένο σύνολο δεδομένων AADA επικεντρώνεται κυρίως σε αρχαιολογικά ευρήματα και στις βασικές περιγραφικές τους ιδιότητες. Αυτό σημαίνει ότι, παρότι υπάρχει σαφής αρχαιολογική διάσταση, το σύνολο δεδομένων δεν φαίνεται να έχει ως κύριο άξονα τη λεπτομερή τεκμηρίωση ανασκαφικών γεγονότων ή στρωματογραφικών διαδικασιών. Επομένως, το CRMarchaeo είναι θεωρητικά πολύ σημαντικό και χρήσιμο ως επέκταση του πεδίου, αλλά στην πράξη δεν αναμένεται να αξιοποιηθεί στο ίδιο βάθος με το CIDOC CRM. Ο ρόλος του στο πλαίσιο της παρούσας μελέτης είναι περισσότερο υποστηρικτικός: βοηθά να τοποθετηθεί η εργασία πιο σωστά μέσα στο αρχαιολογικό πεδίο και αναδεικνύει πιθανές μελλοντικές επεκτάσεις του μοντέλου. [11][13]

Το HDTO διαφοροποιείται ακόμα περισσότερο από τα δύο προηγούμενα, καθώς σχετίζεται με μια πιο σύγχρονη λογική ψηφιακής αναπαράστασης πολιτιστικών οντοτήτων, ιδιαίτερα σε περιβάλλοντα όπου η πληροφορία είναι πολυεπίπεδη, διασυνδεδεμένη και συχνά συνδεδεμένη με ψηφιακές απεικονίσεις ή πιο σύνθετες εφαρμογές. Η αξία του έγκειται στο ότι ανοίγει τη συζήτηση προς τη μελλοντική κατεύθυνση των ψηφιακών διδύμων στον χώρο της πολιτιστικής κληρονομιάς. Παρ' όλα αυτά, για τους στόχους της παρούσας εργασίας, το HDTO δεν αποτελεί τον πιο άμεσο ή λειτουργικό

άξονα υλοποίησης. Αντιθέτως, λειτουργεί περισσότερο ως συμπληρωματικό θεωρητικό σημείο αναφοράς που διευρύνει το πλαίσιο της μελέτης και δείχνει μια σύγχρονη τάση του πεδίου. [12]

Με βάση τα παραπάνω, η πιο κατάλληλη προσέγγιση για την παρούσα εργασία είναι να χρησιμοποιηθεί το CIDOC CRM ως βασικός κορμός της μοντελοποίησης, ενώ το CRMarchaeo να αξιοποιηθεί επιλεκτικά ως αρχαιολογική εξειδίκευση, κυρίως σε σημεία όπου αυτό είναι συμβατό με το διαθέσιμο σύνολο δεδομένων και τις ανάγκες των σεναρίων της εργασίας. Το HDTO δεν θα αποτελέσει βασικό στοιχείο της πρακτικής υλοποίησης, αλλά θα διατηρηθεί στη θεωρητική μελέτη ως ένδειξη της σύγχρονης εξέλιξης των οντολογικών προσεγγίσεων στην πολιτιστική κληρονομιά. [10][11][12][13]

Η επιλογή αυτή θεωρείται κατάλληλη για δύο λόγους. Πρώτον, επιτρέπει στην εργασία να παραμείνει ρεαλιστική και διαχειρίσιμη, χωρίς να επιδιώκει υπερβολικά εκτεταμένη ή τεχνικά δυσανάλογη μοντελοποίηση σε σχέση με το εύρος μιας προπτυχιακής εργασίας. Δεύτερον, εξασφαλίζει ότι η μοντελοποίηση θα βασιστεί σε ένα καθιερωμένο και ισχυρό πρότυπο, το οποίο μπορεί να υποστηρίξει αποτελεσματικά τόσο τον μετασχηματισμό των δεδομένων σε RDF όσο και την επακόλουθη εξερεύνησή τους με SPARQL. [6][10][11]

Συνεπώς, η συνολική στρατηγική της παρούσας εργασίας είναι σαφής: το CIDOC CRM χρησιμοποιείται ως η κύρια εννοιολογική βάση για τον σχεδιασμό του Γνωσιακού Γράφου, το CRMarchaeo ενισχύει τη θεωρητική και αρχαιολογική διάσταση της μελέτης, ενώ το HDTO συμπληρώνει το κεφάλαιο με μια σύγχρονη οπτική σχετική με την ψηφιακή αναπαράσταση πολιτιστικής κληρονομιάς. Αυτή η ισορροπημένη προσέγγιση ανταποκρίνεται τόσο στις απαιτήσεις του θέματος όσο και στους πρακτικούς περιορισμούς της εργασίας. [10][11][12]

Η ανάλυση των σχετικών οντολογιών έδειξε ότι το CIDOC CRM αποτελεί το καταλληλότερο βασικό εννοιολογικό υπόβαθρο για τη μοντελοποίηση της παρούσας εργασίας, ενώ το CRMarchaeo προσφέρει χρήσιμη αρχαιολογική εξειδίκευση σε σημεία όπου αυτό είναι αναγκαίο. Παράλληλα, το HDTO συμπληρώνει τη θεωρητική θεώρηση του πεδίου, αναδεικνύοντας πιο σύγχρονες κατευθύνσεις ψηφιακής αναπαράστασης της πολιτιστικής κληρονομιάς. Με βάση τα παραπάνω, στο επόμενο κεφάλαιο θα προσδιοριστούν οι βασικές απαιτήσεις, τα σενάρια χρήσης και οι κεντρικές οντότητες που θα καθοδηγήσουν τον σχεδιασμό του μοντέλου δεδομένων και την υλοποίηση του Γνωσιακού Γράφου. [10][11][12]

Οντολογία	Κύριος σκοπός	Δυνατό σημείο	Ρόλος στην εργασία
CIDOC CRM	Γενική σημασιολογική αναπαράσταση πολιτιστικής κληρονομιάς	Διαλειτουργικότητα, γεγοντοκεντρική μοντελοποίηση, ώριμο πρότυπο	Βασικός εννοιολογικός κορμός
CRMarchaeo	Εξειδίκευση για αρχαιολογική τεκμηρίωση και ανασκαφικές διαδικασίες	Ανασκαφικά συμφραζόμενα, διαδικασίες, στρωματογραφική λογική	Συμπληρωματική αρχαιολογική εξειδίκευση
HDTO	Ψηφιακή αναπαράσταση πολιτιστικών οντοτήτων με λογική digital twin	Σύγχρονη οπτική για εμπλουτισμένα ψηφιακά περιβάλλοντα	Θεωρητικό / συμπληρωματικό πλαίσιο

Πίνακας 2: Σύγκριση CIDOC CRM, CRMarchaeo και HDTO

Κεφάλαιο 4ο: Σχεδιασμός της Εργασίας

Το παρόν κεφάλαιο επικεντρώνεται στον τρόπο με τον οποίο η θεωρητική και εννοιολογική βάση των προηγούμενων κεφαλαίων μετασχηματίζεται σε ένα συγκεκριμένο και εφαρμόσιμο σχέδιο υλοποίησης. Στο πλαίσιο αυτό παρουσιάζονται οι βασικές σχεδιαστικές επιλογές της εργασίας, τα επιλεγμένα σύνολα δεδομένων, τα σενάρια που θα καθοδηγήσουν τη μοντελοποίηση, καθώς και η γενική λογική με βάση την οποία θα οργανωθεί η προτεινόμενη οντολογία.

Στόχος του κεφαλαίου δεν είναι ακόμη η πλήρης τεχνική υλοποίηση του Γνωσιακού Γράφου, αλλά η αποσαφήνιση του πλαισίου μέσα στο οποίο αυτή θα πραγματοποιηθεί, ώστε να διασφαλιστεί ότι οι επόμενες φάσεις της εργασίας στηρίζονται σε σαφώς ορισμένες απαιτήσεις, επιλεγμένα δεδομένα και συνεκτικές σχεδιαστικές αρχές. Πιο συγκεκριμένα, το κεφάλαιο αυτό λειτουργεί ως γέφυρα ανάμεσα στο θεωρητικό υπόβαθρο και στο πρακτικό μέρος της εργασίας, μεταφέροντας τη συζήτηση από το επίπεδο των εννοιών και των οντολογιών στο επίπεδο του συγκεκριμένου οντολογικού σχεδιασμού και της εφαρμογής του πάνω σε πραγματικά δεδομένα.

Η παρούσα εργασία δεν περιορίζεται στη μετατροπή ενός μόνο dataset σε RDF, αλλά επιδιώκει την ανάπτυξη μιας οντολογικής πρότασης για δεδομένα πολιτιστικής κληρονομιάς, η οποία θα μπορεί να εφαρμοστεί και να αξιολογηθεί πάνω σε δύο συμπληρωματικά use cases. Το πρώτο use case είναι το Archaeological Artefact Database of Finland (AADA), το οποίο προσφέρεται κυρίως για object-centric αναπαράσταση αρχαιολογικών ευρημάτων. Το δεύτερο use case προέρχεται από το OpenContext και αξιοποιείται για να αναδειχθεί πιο καθαρά η event-based διάσταση του μοντέλου, δηλαδή η σύνδεση των ευρημάτων με ανασκαφικές δραστηριότητες, actors, χρονικά διαστήματα, τεκμηριωτικές πράξεις και ψηφιακές αναπαραστάσεις. Με τον τρόπο αυτό, το κεφάλαιο δεν αντιμετωπίζει τα σύνολα δεδομένων ως αυτοσκοπό, αλλά ως μέσα ελέγχου και αξιολόγησης μιας γενικότερης οντολογικής πρότασης.

4.1 Ανάλυση του προβλήματος

Η ανάπτυξη ενός Γνωσιακού Γράφου για δεδομένα πολιτιστικής κληρονομιάς δεν αποτελεί απλώς ένα τεχνικό εγχείρημα μετατροπής δεδομένων σε RDF, αλλά ένα πρόβλημα εννοιολογικής οργάνωσης και σημασιολογικής αναπαράστασης της γνώσης. Το βασικό ζητούμενο της παρούσας εργασίας είναι να εξεταστεί πώς ετερογενή σύνολα αρχαιολογικών και πολιτιστικών δεδομένων, τα οποία δεν είναι αρχικά δομημένα με βάση τις αρχές του Σημασιολογικού Ιστού, μπορούν να μετασχηματιστούν σε μια μορφή περισσότερο οργανωμένη, διασυνδεδεμένη και αξιοποιήσιμη για ερωτήματα και περαιτέρω επεξεργασία.

Το πρόβλημα αυτό συνδέεται άμεσα με τη φύση των πολιτιστικών και αρχαιολογικών δεδομένων. Τέτοιου είδους πληροφορία δεν περιορίζεται σε απλές περιγραφικές εγγραφές, αλλά ενσωματώνει σχέσεις ανάμεσα σε αντικείμενα, τόπους, χρονικές περιόδους, κατηγοριοποιήσεις, ανασκαφικές δραστηριότητες, πρόσωπα, τεκμηριωτικές πράξεις και ψηφιακές αναπαραστάσεις. Επομένως, ένα συμβατικό μοντέλο αποθήκευσης, όπως ένας πίνακας ή ένα υπολογιστικό φύλλο, μπορεί να επαρκεί για την καταγραφή της πληροφορίας σε βασικό επίπεδο, αλλά δεν αποδίδει πάντοτε με επάρκεια τις σημασιολογικές συνδέσεις που είναι κρίσιμες για την κατανόηση και την ερμηνεία των δεδομένων.

Στο πλαίσιο αυτό, το βασικό ερευνητικό πρόβλημα της εργασίας μπορεί να διατυπωθεί ως εξής: με ποιον τρόπο μπορούν δεδομένα πολιτιστικής κληρονομιάς που προέρχονται από διαφορετικές πηγές και έχουν διαφορετικό βαθμό δομικής και εννοιολογικής πολυπλοκότητας να οργανωθούν και να μοντελοποιηθούν οντολογικά, ώστε να παραχθεί ένας Γνωσιακός Γράφος που να υποστηρίζει συνεπή

αναπαράσταση της γνώσης και ουσιαστική εξερεύνηση της πληροφορίας; Η ερώτηση αυτή περιλαμβάνει τόσο τεχνικές όσο και εννοιολογικές διαστάσεις. Από τη μία πλευρά, απαιτείται η επιλογή τεχνολογιών και μηχανισμών μετασχηματισμού των δεδομένων. Από την άλλη, απαιτείται η επιλογή κατάλληλου οντολογικού πλαισίου, το οποίο να είναι συμβατό με το πεδίο της πολιτιστικής κληρονομιάς και να μπορεί να καλύψει περισσότερες από μία μορφές δεδομένων.

Για τις ανάγκες της παρούσας εργασίας, το πρόβλημα εξειδικεύεται μέσω δύο διαφορετικών αλλά συμπληρωματικών περιπτώσεων εφαρμογής. Η πρώτη περίπτωση αφορά το AADA, το οποίο επικεντρώνεται κυρίως σε αρχαιολογικά ευρήματα και στις βασικές περιγραφικές τους ιδιότητες. Η δεύτερη περίπτωση αφορά επιλεγμένα δεδομένα από το OpenContext, τα οποία προσφέρουν πλουσιότερο archaeological context και επιτρέπουν να εξεταστεί σε μεγαλύτερο βάθος η αναπαράσταση γεγονότων, όπως ανασκαφές, τεκμηριωτικές πράξεις και σχέσεις με actors. Με τον τρόπο αυτό, η εργασία δεν καλείται να αντιμετωπίσει ολόκληρη την πολυπλοκότητα του πεδίου της πολιτιστικής κληρονομιάς, αλλά να επικεντρωθεί σε ένα ελεγχόμενο και μεθοδολογικά σαφές υποσύνολο προβλημάτων.

Αυτό σημαίνει ότι το πρόβλημα της εργασίας δεν είναι η πλήρης μοντελοποίηση ολόκληρου του αρχαιολογικού πεδίου, αλλά η ανάπτυξη μιας οντολογικής πρότασης μικρής κλίμακας και ελεγχόμενου εύρους, μέσα από την οποία θα φανεί πώς μπορεί να εφαρμοστεί στην πράξη μια σημασιολογική και οντολογική προσέγγιση πάνω σε πραγματικά δεδομένα. Η επιλογή αυτή είναι σκόπιμη, καθώς επιτρέπει στην εργασία να παραμείνει εντός ενός διαχειρίσιμου εύρους, χωρίς να χάνει τη σύνδεσή της με πραγματικά προβλήματα οργάνωσης και αξιοποίησης πολιτιστικής πληροφορίας.

Παράλληλα, η ανάλυση του προβλήματος οδηγεί και σε έναν δεύτερο σημαντικό άξονα της εργασίας, ο οποίος αφορά τη χρηστικότητα του παραγόμενου Γνωσιακού Γράφου. Δεν αρκεί μόνο να παραχθεί ένα οντολογικά δομημένο αποτέλεσμα, αλλά πρέπει επίσης να εξεταστεί κατά πόσο αυτό το αποτέλεσμα μπορεί να υποστηρίξει ουσιαστικά ερωτήματα πάνω στα δεδομένα. Για τον λόγο αυτό, η εργασία δεν θα περιοριστεί μόνο στην παραγωγή της οντολογίας και των RDF τριάδων, αλλά θα επεκταθεί και στη διερεύνηση του γράφου μέσω SPARQL, ώστε να αξιολογηθεί στην πράξη η χρησιμότητα της μοντελοποίησης που θα υιοθετηθεί.

Συνεπώς, η ανάλυση του προβλήματος δείχνει ότι η παρούσα εργασία κινείται σε τρεις αλληλένδετες διαστάσεις:

Στην εννοιολογική οργάνωση των δεδομένων μέσω κατάλληλων οντολογιών,

στον τεχνικό μετασχηματισμό ενός μη-RDF συνόλου δεδομένων σε RDF Γνωσιακό Γράφο,

στην αξιοποίηση και εξερεύνηση του αποτελέσματος μέσω κατάλληλων ερωτημάτων.

Οι διαστάσεις αυτές θα εξειδικευτούν περαιτέρω στις επόμενες ενότητες, αρχίζοντας από την παρουσίαση των επιλεγμένων συνόλων δεδομένων, τα οποία αποτελούν τη βάση της πρακτικής εφαρμογής της οντολογίας.



Σχήμα 3: Από το πρόβλημα στην υλοποίηση

4.2 Παρουσίαση των επιλεγμένων συνόλων δεδομένων

Για τις ανάγκες της παρούσας εργασίας επιλέγονται δύο συμπληρωματικά σύνολα δεδομένων, τα οποία επιτελούν διαφορετικό αλλά αλληλένδετο ρόλο στη μελέτη. Το πρώτο είναι το Archaeological Artefact Database of Finland (AADA) και λειτουργεί ως βασική πηγή για την object-centric αναπαράσταση αρχαιολογικών ευρημάτων. Το δεύτερο προέρχεται από το OpenContext και αξιοποιείται ως πιο event-based παράδειγμα, ώστε να εξεταστεί σε μεγαλύτερο βάθος η γεγονотоκεντρική λογική του προτεινόμενου μοντέλου.

Η συνδυαστική επιλογή των δύο αυτών πηγών είναι σημαντική, επειδή επιτρέπει στην εργασία να μην περιοριστεί ούτε σε ένα υπερβολικά απλό artefact-only σενάριο ούτε, αντίστροφα, σε μια δυσανάλογα σύνθετη ανασκαφική μοντελοποίηση. Αντίθετα, τα δύο datasets λειτουργούν συμπληρωματικά: το AADA παρέχει σταθερό και διαχειρίσιμο υλικό για την περιγραφή αντικειμένων, ενώ το OpenContext ενισχύει τη δυνατότητα αναπαράστασης γεγονότων, contexts, actors και τεκμηρίωσης.

4.2.1 Παρουσίαση του συνόλου δεδομένων AADA

Το Archaeological Artefact Database of Finland (AADA) επιλέγεται ως η βασική πηγή δεδομένων της εργασίας. Πρόκειται για ανοικτό σύνολο δεδομένων που αφορά αρχαιολογικά ευρήματα από τη Φινλανδία και έχει σχεδιαστεί με στόχο την οργανωμένη συγκέντρωση και διάθεση πληροφοριών σχετικών με προϊστορικά artefacts. Το συγκεκριμένο dataset είναι ιδιαίτερα κατάλληλο για την παρούσα μελέτη, καθώς συνδυάζει επιστημονικό ενδιαφέρον, σαφή δομή και πρακτική αξιοποιησιμότητα για σκοπούς οντολογικής μοντελοποίησης και σημασιολογικού μετασχηματισμού.

Χαρακτηριστικό	Περιγραφή
Όνομα	Archaeological Artefact Database of Finland (AADA)
Αντικείμενο	Αρχαιολογικά ευρήματα / artefacts από τη Φινλανδία
Μορφή δεδομένων	Μη-RDF, κατάλληλο για μετασχηματισμό σε RDF
Κεντρική μονάδα πληροφορίας	Μεμονωμένο αρχαιολογικό εύρημα
Ενδεικτικά πεδία	Τύπος αντικειμένου, περίοδος, τόπος/πλαίσιο, περιγραφικά χαρακτηριστικά, μετρήσεις
Καταλληλότητα	Υποστηρίζει object-centric μοντελοποίηση με CIDOC CRM
Περιορισμός	Δεν καλύπτει πλήρως ανασκαφικές ροές εργασίας ή στρωματογραφικές σχέσεις

Πίνακας 3: Βασικά Χαρακτηριστικά του AADA

Το AADA περιλαμβάνει εγγραφές που αντιστοιχούν σε μεμονωμένα αρχαιολογικά ευρήματα, οργανωμένες σε δομημένη μορφή. Κάθε εγγραφή περιγράφει ένα artefact μέσα από ένα σύνολο χαρακτηριστικών, όπως ο τύπος του αντικειμένου, η χρονολογική ή πολιτισμική του περίοδος, η τοποθεσία ή το αρχαιολογικό πλαίσιο προέλευσης, καθώς και άλλες περιγραφικές ή μετρητικές πληροφορίες, όπου αυτές είναι διαθέσιμες. Η δομή αυτή το καθιστά ιδιαίτερα πρόσφορο για τη μετάβαση από μια συμβατική μορφή αποθήκευσης σε μια περισσότερο σημασιολογικά οργανωμένη αναπαράσταση.

Ένας βασικός λόγος για τον οποίο επιλέχθηκε το AADA είναι ότι πρόκειται για μη-RDF σύνολο δεδομένων, γεγονός που το καθιστά κατάλληλο για να υποστηρίζει τον βασικό στόχο της εργασίας, δηλαδή τον μετασχηματισμό συμβατικών πολιτιστικών δεδομένων σε RDF Γνωσιακό Γράφο. Με άλλα λόγια, το σύνολο δεδομένων προσφέρει μια ρεαλιστική αφετηρία για την εφαρμογή των

μεθόδων και τεχνολογιών που έχουν παρουσιαστεί στα προηγούμενα κεφάλαια, όπως η οντολογική μοντελοποίηση, η χρήση της OWL και η παραγωγή RDF τριάδων.

Παράλληλα, το AADA είναι ιδιαίτερα χρήσιμο επειδή η πληροφορία που περιέχει είναι αρκετά πλούσια ώστε να δικαιολογεί οντολογική μοντελοποίηση, χωρίς όμως να είναι τόσο σύνθετη ώστε να καθιστά το εύρος της εργασίας μη διαχειρίσιμο. Η ύπαρξη πεδίων που σχετίζονται με ταυτοποίηση αντικειμένων, κατηγοριοποίηση, χρονική ένταξη, χωρική πληροφορία και βασικές μετρήσεις επιτρέπει τη χαρτογράφηση των δεδομένων σε έννοιες που ταιριάζουν καλά με το CIDOC CRM.

Ένας ακόμη σημαντικός λόγος επιλογής του είναι ότι επιτρέπει μια object-centric προσέγγιση της μοντελοποίησης. Καθώς κάθε εγγραφή αντιστοιχεί ουσιαστικά σε ένα αρχαιολογικό εύρημα, είναι εφικτό να οργανωθεί η μοντελοποίηση γύρω από το ίδιο το αντικείμενο και τις βασικές σχέσεις του με άλλες οντότητες, όπως τόποι, χρονικές περίοδοι και ταξινομήσεις. Αυτή η λογική είναι ιδιαίτερα χρήσιμη, επειδή επιτρέπει να δομηθεί ένας καθαρός πυρήνας της οντολογικής πρότασης πάνω σε πραγματικά δεδομένα.

Κριτήριο επιλογής	Πώς το καλύπτει το AADA
Ανοικτό και διαθέσιμο dataset	Είναι δημόσια διαθέσιμο για ερευνητική χρήση
Μη-RDF μορφή	Επιτρέπει να φανεί η διαδικασία μετασχηματισμού σε RDF
Δομημένη πληροφορία	Περιέχει οργανωμένες εγγραφές για artefacts
Σχέση με πολιτιστική κληρονομιά	Αφορά αρχαιολογικά ευρήματα
Καταλληλότητα για CIDOC CRM	Τα αντικείμενα, οι τύποι, οι περίοδοι και οι τόποι χαρτογραφούνται φυσικά σε CIDOC CRM έννοιες
Διαχειρίσιμο scope	Είναι αρκετά πλούσιο χωρίς να είναι υπερβολικά σύνθετο

Πίνακας 4: Γιατί επιλέχθηκε το AADA

Ωστόσο, το AADA παρουσιάζει και ορισμένους περιορισμούς, οι οποίοι είναι σημαντικό να αναγνωριστούν εξ αρχής. Παρότι είναι ιδιαίτερα χρήσιμο για την περιγραφή αρχαιολογικών ευρημάτων και βασικών χαρακτηριστικών τους, δεν εστιάζει στον ίδιο βαθμό στη λεπτομερή αναπαράσταση ανασκαφικών γεγονότων, στρωματογραφικών σχέσεων ή πλήρων τεκμηριωτικών ροών εργασίας. Αυτό σημαίνει ότι το σύνολο δεδομένων ευνοεί περισσότερο την object-oriented πλευρά της οντολογίας, ενώ από μόνο του δεν επαρκεί για να αναδείξει όλο το event-based δυναμικό του μοντέλου.

Παρότι το AADA διαθέτει έναν κοινό πυρήνα πεδίων, η πραγματική δομή του δεν είναι απόλυτα ομοιόμορφη σε όλα τα υποσύνολα. Αντίθετα, περιλαμβάνει διαφοροποιήσεις ανά χρονολογική περίοδο και κατηγορία artefact, καθώς ορισμένα αρχεία διαθέτουν εξειδικευμένα πεδία για subperiod dating, context classification ή count-based representation, ιδίως στην περίπτωση της κεραμικής. Η παρατήρηση αυτή είναι σημαντική, επειδή δείχνει ότι το μοντέλο σχήμα δεν πρέπει να σχεδιαστεί ως μηχανική αντανάκλαση ενός μόνο dataset, αλλά ως γενικότερο μοντέλο που μπορεί να απορροφήσει τέτοιες διαφοροποιήσεις.

Συνεπώς, το AADA επιλέγεται ως βασικό σύνολο δεδομένων της εργασίας επειδή καλύπτει ταυτόχρονα τρεις σημαντικές απαιτήσεις:

1. παρέχει πραγματικά δεδομένα πολιτιστικής κληρονομιάς,
2. βρίσκεται σε μορφή κατάλληλη για σημασιολογικό μετασχηματισμό,

3. υποστηρίζει ένα ρεαλιστικό και διαχειρίσιμο σενάριο object-centric οντολογικής μοντελοποίησης.

4.2.2 Παρουσίαση του OpenContext ως δεύτερου use case

Εκτός από το AADA, στην παρούσα εργασία αξιοποιείται και ένα δεύτερο use case από το OpenContext, με στόχο να ενισχυθεί η event-based διάσταση της οντολογίας. Το OpenContext είναι μια ψηφιακή υποδομή δημοσίευσης και διάθεσης αρχαιολογικών δεδομένων, η οποία περιλαμβάνει item-level εγγραφές, αρχαιολογικά συμφραζόμενα, project-level πληροφορία, καθώς και διασυνδέσεις με άλλους πόρους του Linked Data οικοσυστήματος. Σε αντίθεση με το AADA, το οποίο οργανώνεται κυρίως γύρω από μεμονωμένα artefacts, το OpenContext προσφέρεται περισσότερο για τη σύνδεση ευρημάτων με ανασκαφικές δραστηριότητες, actors, τεκμηρίωση και πλουσιότερα contexts.

Ο ρόλος του OpenContext στην εργασία δεν είναι να λειτουργήσει ως δεύτερη κύρια βάση δεδομένων ίδιου βάρους με το AADA, αλλά ως συμπληρωματικό use case, το οποίο επιτρέπει να φανεί πιο καθαρά η γεγονотоκεντρική λογική του μοντέλου. Πιο συγκεκριμένα, το OpenContext αξιοποιείται για την αναπαράσταση περιπτώσεων όπου ένα εύρημα δεν αντιμετωπίζεται μόνο ως αντικείμενο με ιδιότητες, αλλά και ως πόρος που συνδέεται με ανασκαφική πράξη, τόπο, χρονικό διάστημα, actor, τεκμηριωτική δημιουργία και ενδεχομένως με ψηφιακή αναπαράσταση.

Η επιλογή αυτή είναι σημαντική, επειδή επιτρέπει στην εργασία να μη μείνει σε ένα καθαρά object-based σενάριο. Μέσω του δεύτερου αυτού use case, το μοντέλο μπορεί να ελεγχθεί σε πιο σύνθετες σχέσεις, όπως η σύνδεση ενός αντικειμένου με ανασκαφική δραστηριότητα, με stratigraphic unit, με actor και με documentation creation event. Με τον τρόπο αυτό, το OpenContext λειτουργεί ως αναγκαίο συμπλήρωμα του AADA και όχι ως ανταγωνιστική πηγή δεδομένων.

Συνολικά, η συνδυαστική χρήση των δύο datasets επιτρέπει στην εργασία να καλύψει ένα ευρύτερο φάσμα απαιτήσεων: από την καθαρή περιγραφή και κατηγοριοποίηση αρχαιολογικών ευρημάτων μέχρι τη σύνδεσή τους με γεγονότα, τεκμηρίωση και ψηφιακή αναπαράσταση. Η επιλογή αυτή ενισχύει τη μεθοδολογική πληρότητα της εργασίας και καθιστά το οντολογικό μοντέλο περισσότερο πειστικό και γενικεύσιμο.

4.3 Βασικά σενάρια που θα μελετηθούν

Μετά την επιλογή και παρουσίαση των επιλεγμένων συνόλων δεδομένων, είναι απαραίτητο να προσδιοριστούν τα βασικά σενάρια που θα καθοδηγήσουν τη μοντελοποίηση. Τα σενάρια αυτά λειτουργούν ως ενδιάμεσο βήμα ανάμεσα στα αρχικά δεδομένα και στο τελικό οντολογικό μοντέλο, καθώς βοηθούν να αποσαφηνιστεί ποιες πληροφορίες πρέπει να αναπαρασταθούν, ποιες σχέσεις είναι σημαντικές και τι είδους ερωτήματα θα πρέπει να μπορεί να υποστηρίξει ο παραγόμενος Γνωσιακός Γράφος.

Στην παρούσα εργασία, τα σενάρια δεν αποσκοπούν στην πλήρη κάλυψη όλων των πιθανών χρήσεων ενός συστήματος πολιτιστικής κληρονομιάς. Αντίθετα, επιλέγονται με βάση το περιεχόμενο και τη δομή των δύο use cases, ώστε να παραμείνουν ρεαλιστικά και άμεσα συνδεδεμένα με τα διαθέσιμα δεδομένα. Η επιλογή αυτή είναι σημαντική, επειδή αποφεύγει την υπερβολική διεύρυνση του εύρους της εργασίας και επιτρέπει την ανάπτυξη ενός μικρού αλλά συνεκτικού proof of concept.

Το πρώτο βασικό σενάριο αφορά την αναπαράσταση αρχαιολογικών ευρημάτων. Σε αυτό το σενάριο, το ενδιαφέρον επικεντρώνεται στο ίδιο το artefact ως κεντρική οντότητα του μοντέλου. Κάθε εύρημα περιγράφεται μέσα από βασικά χαρακτηριστικά, όπως ο τύπος του αντικειμένου, το υλικό, πιθανές μετρήσεις, η χρονολογική του ένταξη και η σχέση του με συγκεκριμένο τόπο ή context. Το σενάριο

αυτό είναι το πιο άμεσα συνδεδεμένο με το AADA και αποτελεί τον βασικό κορμό της object-oriented πλευράς της μοντελοποίησης.

Το δεύτερο σενάριο αφορά την κατηγοριοποίηση και χρονική ένταξη των ευρημάτων. Τα αρχαιολογικά αντικείμενα δεν αποκτούν νόημα μόνο ως μεμονωμένες εγγραφές, αλλά και μέσα από τη σύνδεσή τους με ευρύτερες κατηγορίες, περιόδους και ταξινομητικά σχήματα. Για παράδειγμα, ένα εύρημα μπορεί να ανήκει σε συγκεκριμένο τύπο αντικειμένου ή να σχετίζεται με μια προϊστορική περίοδο. Η αναπαράσταση αυτών των σχέσεων είναι σημαντική, επειδή επιτρέπει τη διατύπωση ερωτημάτων που δεν αναζητούν απλώς μεμονωμένα αντικείμενα, αλλά ομάδες αντικειμένων με κοινά χαρακτηριστικά.

Το τρίτο σενάριο αφορά τη σύνδεση των ευρημάτων με τόπους και συμφραζόμενα προέλευσης. Στον βαθμό που τα επιλεγμένα δεδομένα περιλαμβάνουν πληροφορίες για τόπους, περιοχές, sites ή archaeological contexts, οι πληροφορίες αυτές μπορούν να αξιοποιηθούν για τη δημιουργία σχέσεων ανάμεσα στα artefacts και στο χωρικό ή τεκμηριωτικό τους πλαίσιο. Το σενάριο αυτό είναι ιδιαίτερα σημαντικό, επειδή στην αρχαιολογία η αξία ενός ευρήματος συνδέεται συχνά με το πού εντοπίστηκε ή με το πλαίσιο μέσα στο οποίο τεκμηριώθηκε.

Το τέταρτο σενάριο αφορά τη γεγονотоκεντρική αναπαράσταση της πληροφορίας. Εδώ το ενδιαφέρον δεν περιορίζεται στο ίδιο το αντικείμενο, αλλά επεκτείνεται και στα γεγονότα που σχετίζονται με αυτό, όπως η παραγωγή, η ανασκαφή και η δημιουργία τεκμηρίωσης. Το σενάριο αυτό συνδέεται περισσότερο με το δεύτερο use case από το OpenContext και είναι καθοριστικό για να φανεί στην πράξη η αξιοποίηση του CIDOC CRM και του CRMarchaeo πέρα από ένα απλό object-based σχήμα.

Το πέμπτο σενάριο αφορά την ψηφιακή αναπαράσταση των ευρημάτων. Στο πλαίσιο αυτό, διερευνάται η δυνατότητα να συνδέεται το φυσικό αντικείμενο με μία ψηφιακή του όψη, όπως φωτογραφία, report, field note ή άλλη μορφή ψηφιακού τεκμηρίου. Το σενάριο αυτό δεν αποτελεί τον κύριο άξονα της εργασίας, αλλά είναι σημαντικό επειδή επιτρέπει να ενταχθεί στο μοντέλο και ένα περιορισμένο επίπεδο digital twin / digital representation.

Το έκτο σενάριο αφορά την εξερεύνηση του παραγόμενου Γνωσιακού Γράφου μέσω SPARQL ερωτημάτων. Τα προηγούμενα σενάρια οδηγούν στη διαμόρφωση συγκεκριμένων ερωτημάτων που θα χρησιμοποιηθούν για την αξιολόγηση της χρησιμότητας της οντολογίας. Με τον τρόπο αυτό, τα σενάρια χρήσης συνδέονται άμεσα με τη μετέπειτα πειραματική διερεύνηση των δεδομένων.

Τα παραπάνω σενάρια μπορούν να θεωρηθούν και ως ένα πρώτο σύνολο competency questions (CQs), δηλαδή ερωτημάτων που καθοδηγούν τον σχεδιασμό του μοντέλου και δείχνουν τι πρέπει να μπορεί να απαντήσει ο Γνωσιακός Γράφος. Στο πλαίσιο της παρούσας εργασίας, τα CQs δεν χρησιμοποιούνται ως πλήρης και αυστηρή μεθοδολογία ανάπτυξης οντολογίας, αλλά ως πρακτικό εργαλείο αποσαφήνισης των απαιτήσεων. Με τον τρόπο αυτό, το μοντέλο δεν σχεδιάζεται αφηρημένα, αλλά με βάση συγκεκριμένες πληροφοριακές ανάγκες.

A/A	Σενάριο	Competency question	Πληροφορία που πρέπει να μοντελοποιηθεί
CQ1	Βασική αναπαράσταση artefact	Ποια ευρήματα περιλαμβάνονται στον γράφο και ποια είναι τα βασικά identifiers και labels τους;	gerasim:Artefact, E18, E22 όπου εφαρμόζεται, E42, labels, source records
CQ2	Τυπολογία και υλικό	Σε ποιον τύπο ανήκει κάθε εύρημα και από ποιο υλικό αποτελείται;	E55 Type, E57 Material, P2, P45
CQ3	Χρονική ένταξη	Με ποια περίοδο ή υποπερίοδο συνδέεται ένα	E4 Period, E52 Time-Span,

		εύρημα ή ένα γεγονός;	skos:broader, PeriodO alignment
CQ4	Χωρική αναπαράσταση	Με ποιον τόπο και με ποια γεωμετρία συνδέεται ένα εύρημα ή μια ανασκαφή;	E53 Place, GeoSPARQL geometry, coordinates
CQ5	Excavation event	Ποια ευρήματα βρέθηκαν μέσω συγκεκριμένης ανασκαφικής δραστηριότητας;	A9 Archaeological Excavation, E7 Activity, P12i_was_present_at ή excavation shortcut
CQ6	Stratigraphic context	Ποια ευρήματα συνδέονται με συγκεκριμένη στρωματογραφική μονάδα ή context;	A8 Stratigraphic Unit, context relations
CQ7	Production layer	Πώς αναπαρίσταται η παραγωγή / κατασκευή ενός ευρήματος;	E12 Production, E39 Actor, E52 Time-Span, material/type relations
CQ8	Documentation layer	Πώς τεκμηριώνεται ένα εύρημα και μέσω ποιας πράξης δημιουργείται η τεκμηρίωση;	E65 Creation, documentation events, actors, provenance
CQ9	Same archaeologist across excavations	Ποια ευρήματα ανακαλύφθηκαν από τον ίδιο αρχαιολόγο σε διαφορετικές ανασκαφές;	excavation events, E39 Actor, place/time-span differentiation
CQ10	Same period across different places	Ποια αντικείμενα χρονολογούνται στην ίδια περίοδο αλλά προέρχονται από διαφορετικές τοποθεσίες;	E4 Period, E53 Place, artefact-event-place chains
CQ11	Digital twin layer	Ποια ευρήματα έχουν ψηφιακή αναπαράσταση / HC2 digital twin;	HC2 Heritage Digital Twin, HC3 Tangible Heritage Entity, documentation links
CQ12	Digital twin provenance	Ποια ευρήματα έχουν ψηφιακή αναπαράσταση και από ποιες ανασκαφές προέρχονται;	digital twin + excavation event + place + actor
CQ13	Temporal distribution by place	Ποια είναι η χρονολογική κατανομή ευρημάτων ανά τόπο εύρεσης;	E52, E4, E53, grouping / aggregation logic
CQ14	Multi-source graph exploitation	Πώς συνδυάζονται AADA object-centric δεδομένα και OpenContext event-based δεδομένα στο ίδιο οντολογικό σχήμα;	κοινός οντολογικός πυρήνας, δύο use cases, SPARQL over shared model

Πίνακας 5: Προτεινόμενα 14 Competency Questions

Ο παραπάνω πίνακας παρουσιάζει ένα ενδεικτικό σύνολο competency questions που θα καθοδηγήσει τον σχεδιασμό του μοντέλου. Τα ερωτήματα δεν αποσκοπούν στην εξαντλητική κάλυψη όλων των πιθανών τρόπων αναζήτησης στα προσφερόμενα σύνολα δεδομένων, αλλά στη διαμόρφωση ενός ρεαλιστικού πλαισίου απαιτήσεων για την παρούσα εργασία. Μέσω αυτών καθορίζονται οι βασικές οντότητες, σχέσεις και χαρακτηριστικά που πρέπει να αποδοθούν στον Γνωσιακό Γράφο, καθώς και η βάση για τη μετέπειτα διατύπωση των ερωτημάτων SPARQL.

Συνολικά, τα σενάρια που επιλέγονται διαμορφώνουν ένα ρεαλιστικό και συνεκτικό πλαίσιο για την παρούσα εργασία. Επικεντρώνονται στην αναπαράσταση, κατηγοριοποίηση και εξερεύνηση αρχαιολογικών ευρημάτων, αλλά επεκτείνονται και στη γεγονотоκεντρική, τεκμηριωτική και ψηφιακή διάσταση της πολιτιστικής πληροφορίας. Με βάση αυτά τα σενάρια, το επόμενο βήμα είναι

ο σχεδιασμός του μοντέλου δεδομένων, δηλαδή η επιλογή των βασικών οντοτήτων, σχέσεων και οντολογικών αντιστοιχίσεων που θα χρησιμοποιηθούν για την υλοποίηση του Γνωσιακού Γράφου.

4.4 Σχεδιασμός του μοντέλου δεδομένων

Το παρόν υποκεφάλαιο παρουσιάζει τον σχεδιασμό της προτεινόμενης οντολογίας για δεδομένα πολιτιστικής κληρονομιάς, η οποία αποτελεί το βασικό εννοιολογικό και τεχνικό υπόβαθρο της υλοποίησης. Η οντολογία στηρίζεται κυρίως στο CIDOC CRM και συμπληρώνεται από στοιχεία του CRMarchaeo, του GeoSPARQL, του SKOS, του PROV-O και του HDTO. Η λογική της δεν είναι η εκ νέου δημιουργία ενός πλήρους παράλληλου λεξιλογίου, αλλά η κατά προτεραιότητα αξιοποίηση υφιστάμενων κλάσεων και ιδιοτήτων από τις παραπάνω οντολογίες, με περιορισμένη εισαγωγή τοπικών όρων μόνο όπου αυτό κρίνεται απολύτως αναγκαίο για εφαρμοστική απλοποίηση. Ο σχεδιασμός της δεν προκύπτει ως απλή αντανάκλαση ενός μόνο συνόλου δεδομένων, αλλά ως γενικότερο μοντέλο που εφαρμόζεται και αξιολογείται πάνω σε δύο συμπληρωματικά use cases: το AADA, το οποίο υποστηρίζει κυρίως object-centric αναπαράσταση αρχαιολογικών ευρημάτων, και το OpenContext, το οποίο επιτρέπει την ενίσχυση της γεγονотоκεντρικής διάστασης του μοντέλου.

4.4.1 Γενική σχεδιαστική λογική

Η γενική σχεδιαστική λογική της οντολογίας βασίζεται στην παραδοχή ότι η πληροφορία πολιτιστικής κληρονομιάς δεν πρέπει να αναπαρίσταται μόνο ως σύνολο αντικειμένων και ιδιοτήτων, αλλά ως δίκτυο από αντικείμενα, γεγονότα, τόπους, χρονικά διαστήματα, actors, τεκμηριωτικές πράξεις και ψηφιακές αναπαραστάσεις. Για τον λόγο αυτό, η οντολογία της παρούσας εργασίας δεν περιορίζεται σε object-centric περιγραφή artefacts, αλλά ενσωματώνει ρητά μια event-based διάσταση, σύμφωνη με το πνεύμα του CIDOC CRM και με αρχαιολογικές επεκτάσεις τύπου CRMarchaeo. Το μοντέλο καλύπτει το ίδιο το εύρημα, την παραγωγή του, την ανασκαφική δραστηριότητα μέσω της οποίας συνδέεται με συγκεκριμένο τόπο και χρόνο, τη δημιουργία τεκμηρίωσης και τη σχέση του με ψηφιακές αναπαραστάσεις.

Η βασική σχεδιαστική αρχή που ακολουθήθηκε είναι η άμεση επαναχρησιμοποίηση υπάρχουσών κλάσεων και ιδιοτήτων από το CIDOC CRM, το CRMarchaeo, το GeoSPARQL [17], το SKOS [14], το PROV-O [15], το DCAT [16] και το HDTO [18], όπου αυτές καλύπτουν επαρκώς τη σημασιολογική απαίτηση. Τοπικές κλάσεις ή ιδιότητες ορίζονται μόνο όταν απαιτείται εφαρμοστική εξειδίκευση ή συντομευτική αποτύπωση ενός συχνά επαναλαμβανόμενου pattern. Ειδικά στο επίπεδο της ψηφιακής αναπαράστασης, η σχέση artefact–digital twin αποδίδεται πλέον με άμεση χρήση της ιδιότητας `hdt:HP1_has_digital_twin`, ενώ το `HC2_Heritage_Digital_Twin` χρησιμοποιείται ως η βασική κλάση του HDTO για το digital twin layer [18].

Το σύνολο δεδομένων AADA χρησιμοποιείται ως βασικό use case, επειδή παρέχει πραγματικά και πλούσια tabular δεδομένα για αρχαιολογικά ευρήματα, τύπους, περιόδους, τόπους, συντεταγμένες και μετρήσεις. Ωστόσο, δεν αντιμετωπίζεται ως το ίδιο το μοντέλο ούτε ως η μοναδική πηγή για την αξιολόγηση της οντολογίας. Η δομή του το καθιστά ιδιαίτερα χρήσιμο για την object-centric πλευρά του σχήματος, αλλά όχι επαρκές από μόνο του για να αναδείξει πλήρως τη γεγονотоκεντρική λογική του CIDOC CRM. Για τον λόγο αυτό, η οντολογία αξιολογείται και σε ένα δεύτερο use case με βάση το OpenContext, το οποίο επιτρέπει πλουσιότερη αναπαράσταση excavation events, stratigraphic contexts, actors και documentation processes.

Ο βασικός κορμός της οντολογίας στηρίζεται σε κλάσεις και ιδιότητες του CIDOC CRM, όπως E18 Physical Thing, E22 Human-Made Object, E12 Production, E7 Activity, E65 Creation, E39 Actor,

E53 Place, E52 Time-Span, E55 Type, E57 Material, E54 Dimension, E42 Identifier και E31 Document. Η χρήση του E18 Physical Thing επιτρέπει την κάλυψη τόσο ανθρωπογενών αρχαιολογικών αντικειμένων όσο και άλλων φυσικών αρχαιολογικών καταλοίπων, όπως οστά ή λοιπά ευρήματα που δεν είναι απαραίτητα ανθρώπινης κατασκευής. Όταν ένα συγκεκριμένο instance αφορά πράγματι ανθρωπογενές αντικείμενο, μπορεί επιπλέον να δηλώνεται και ως `crm:E22_Human-Made_Object`.

Η οντολογία ενσωματώνει επίσης ένα περιορισμένο επίπεδο ψηφιακής αναπαράστασης με βάση τους όρους `HC2_Heritage_Digital_Twin` και `HC3_Tangible_Heritage_Entity` της HDTO, οι οποίοι αξιοποιούνται για τη σύνδεση της φυσικής πολιτιστικής οντότητας με την ψηφιακή της αναπαράσταση. Στο πλαίσιο της παρούσας εργασίας, το επίπεδο αυτό χρησιμοποιείται με ελεγχόμενο και στοχευμένο τρόπο, ώστε να υποστηρίζονται περιπτώσεις όπως φωτογραφίες, τεκμηριωτικά αρχεία, field notes ή άλλες ψηφιακές αναπαραστάσεις που σχετίζονται με τα αρχαιολογικά ευρήματα. Με τον τρόπο αυτό, το μοντέλο επεκτείνεται πέρα από την περιγραφή του φυσικού artefact και καλύπτει και τη σύνδεσή του με το ψηφιακό του αποτύπωμα.

Τέλος, η οντολογία περιλαμβάνει περιορισμένο αριθμό τοπικών όρων στο namespace της εργασίας. Μετά από έλεγχο έναντι των ήδη διαθέσιμων όρων των χρησιμοποιούμενων οντολογιών, η τελική έκδοσή του μοντέλου δεν διατηρεί ένα εκτεταμένο σύνολο τοπικών κλάσεων και ιδιοτήτων, αλλά βασίζεται κυρίως στην άμεση χρήση των υφιστάμενων κλάσεων και ιδιοτήτων του CIDOC CRM, του CRMarchaeo, του GeoSPARQL, του SKOS, του PROV-O και του HDTO. Οι τοπικοί όροι που παραμένουν χρησιμοποιούνται είτε ως ελεγχόμενες application-level κλάσεις είτε ως shortcut ιδιότητες, όπου αυτό κρίνεται αναγκαίο για λόγους αναγνωσιμότητας, εφαρμοστικής απλοποίησης και διαχειρισιμότητας του μοντέλου, χωρίς να υποκαθιστούν τον βασικό εννοιολογικό κορμό των υφιστάμενων οντολογιών.

4.4.2 Βασικές οντότητες του μοντέλου

Η τοπική κλάση `gerasim:Artefact` χρησιμοποιείται ως application-level κλάση για τα φυσικά αρχαιολογικά ευρήματα που εμφανίζονται στα παραδείγματα της εργασίας. Επειδή τα δεδομένα δεν περιορίζονται αποκλειστικά σε ανθρωπογενή αντικείμενα, αλλά περιλαμβάνουν και φυσικά αρχαιολογικά κατάλοιπα, όπως οστά, η κλάση δεν δηλώνεται γενικά ως υποκλάση του `crm:E22_Human-Made_Object`. Αντίθετα, ευθυγραμμίζεται με το γενικότερο `crm:E18_Physical_Thing`. Όταν ένα συγκεκριμένο instance είναι πράγματι ανθρωπογενές αντικείμενο, μπορεί επιπλέον να δηλωθεί και ως `crm:E22_Human-Made_Object`.

Εννοιολογικός ρόλος στο μοντέλο	Περιγραφή	Κύρια κλάση / κλάσεις που χρησιμοποιούνται	Συμπληρωματική αντιστοίχιση	Ρόλος στο μοντέλο
Artefact	Φυσικό αρχαιολογικό εύρημα, συμπεριλαμβανομένων ανθρωπογενών αντικειμένων και άλλων φυσικών αρχαιολογικών καταλοίπων	<code>gerasim:Artefact</code> , ως subclass του <code>crm:E18_Physical_Thing</code>	<code>crm:E22_Human-Made_Object</code> μόνο όπου το instance είναι πράγματι ανθρωπογενές· <code>hdto:HC3_Tangible_Heritage_Entity</code> όπου υπάρχει digital twin layer	Κεντρικός πόρος του γράφου
Production	Γεγονός κατασκευής /	<code>crm:E12_Production</code>	—	Production layer

event	παραγωγή			
Excavation event	Ανασκαφική δραστηριότητα	crma:A9_Archaeological_Excavation	crm:E7_Activity	Excavation layer
Documentation creation	Δημιουργία τεκμηρίωσης	crm:E65_Creation	—	Documentation layer
Stratigraphic context	Στρωματογραφική μονάδα / context	crma:A8_Stratigraphic_Unit	—	Archaeological context layer
Place	Τόπος, site ή αρχαιολογική θέση	crm:E53_Place	geo:Feature	Χωρική αναπαράσταση
Time-span	Χρονικό διάστημα γεγονότος ή τεκμηρίωσης	crm:E52_Time-Span	PeriodO-compatible alignment	Χρονικός προσδιορισμός
Actor	Πρόσωπο ή οργανισμός	crm:E39_Actor	—	Archaeologist, researcher, maker, documenter
Artefact type	Τύπος ή κατηγορία ευρήματος	crm:E55_Type	skos:Concept	Ταξινόμηση
Material	Υλικό αντικειμένου	crm:E57_Material	skos:Concept	Φυσική σύσταση
Dimension	Μετρήσεις και ποσοτικά χαρακτηριστικά	crm:E54_Dimension	—	Quantitative layer
Identifier	Αναγνωριστικό	crm:E42_Identifier	—	Ταυτοποίηση
Dataset record / source record	Source record / provenance record	crm:E31_Document	prov:Entity [15]	Τεκμηριωτική προέλευση
Digital twin / digital representation	Ψηφιακή αναπαράσταση	hdto:HC2_Heritage_Digital_Twin	—	Digital layer
Period	Χρονική ή πολιτισμική περίοδος	crm:E4_Period	skos:Concept, PeriodO alignment	Period semantics
Find context	Context κατηγορίας εύρεσης	crm:E55_Type	skos:Concept	Lightweight contextual categorisation

Πίνακας 6: Βασικοί εννοιολογικοί ρόλοι του προτεινόμενου μοντέλου και οντολογική αντιστοίχιση

Ο πίνακας αυτός δείχνει ότι το προτεινόμενο μοντέλο καλύπτει τόσο το φυσικό αντικείμενο όσο και τα γεγονότα, τους τόπους, τους χρόνους, τους actors, την τεκμηρίωση και το ψηφιακό του αποτύπωμα, κυρίως μέσω άμεσης αξιοποίησης υφιστάμενων κλάσεων από τις οντολογίες που επαναχρησιμοποιούνται. Έτσι αποφεύγεται η εντύπωση ενός απλού object-based RDF σχήματος, αλλά και η άσκοπη επανάληψη ήδη διαθέσιμων οντολογικών όρων, ενώ ενισχύεται η δυνατότητα του μοντέλου να απαντά σύνθετα ερωτήματα γράφου.

Για τη χωρική αναπαράσταση δεν ορίζεται ξεχωριστή custom κλάση gerasim:Place, καθώς το dataset δεν παρέχει ειδικότερο τύπο τόπου, όπως archaeological place ή excavation place, που να απαιτεί πρόσθετη εφαρμοστική εξειδίκευση. Αντίθετα, χρησιμοποιείται απευθείας το crm:E53_Place, ενώ όπου υπάρχει διαθέσιμη γεωγραφική πληροφορία ο πόρος δηλώνεται συμπληρωματικά και ως geo:Feature. Με αυτόν τον τρόπο η χωρική αναπαράσταση παραμένει συμβατή με το CIDOC CRM

και μπορεί παράλληλα να αξιοποιεί GeoSPARQL ιδιότητες, όπως geo:hasGeometry και geo:asWKT [17].

Η μοντελοποίηση της περιόδου γίνεται μέσω ενός application-level period node, το οποίο χρησιμοποιείται τόσο ως πολιτισμική-ιστορική περίοδος όσο και ως έννοια ελεγχόμενου λεξιλογίου για ελαφριά ευθυγράμμιση των δεδομένων. Για τον λόγο αυτό, η κλάση gerasim:Period δηλώνεται ως υποκλάση τόσο του crm:E4_Period όσο και του skos:Concept [14]. Η επιλογή αυτή δεν αποτελεί την αυστηρότερη δυνατή μοντελοποίηση, αλλά εξυπηρετεί τις ανάγκες της παρούσας εργασίας, καθώς επιτρέπει τη σύνδεση τοπικών period values με εξωτερικές period authorities, όπως το PeriodO [19]. Επιπλέον, όπου είναι διαθέσιμη ή μπορεί να δηλωθεί σχετική χρονική έκταση, το period node συνδέεται με crm:E52_Time-Span μέσω της ιδιότητας crm:P4_has_time-span. Με αυτόν τον τρόπο, η περίοδος μπορεί να συνοδεύεται όχι μόνο από ετικέτα ελεγχόμενου λεξιλογίου, αλλά και από χρονικά όρια, όπως begin και end values.

4.4.3 Βασικές σχέσεις του προτεινόμενου μοντέλου

Σχέση	Από	Προς	Προτεινόμενη ιδιότητα	Οντολογία	Σκοπός
έχει αναγνωριστικό	artefact	identifier	crm:P1_is_identified_by	CIDOC CRM	Ταυτοποίηση artefact
έχει τύπο	artefact	artefact type	crm:P2_has_type	CIDOC CRM	Κατηγοριοποίηση
αποτελείται από υλικό	artefact	material	crm:P45_consists_of	CIDOC CRM	Material modelling
έχει διάσταση	artefact	dimension	crm:P43_has_dimension	CIDOC CRM	Measurements
έχει production event	artefact	production event	crm:P108i_was_produced_by	CIDOC CRM	Production relation
ήταν παρόν σε excavation	artefact	excavation event	crm:P12i_was_present_at	CIDOC CRM	Excavation relation
δημιουργείται από documentation creation	digital twin	documentation creation	crm:P94i_was_created_by	CIDOC CRM	Provenance της ψηφιακής αναπαράστασης
δημιουργεί ψηφιακή αναπαράσταση	documentation creation	digital twin	crm:P94_has_created	CIDOC CRM	Σύνδεση documentation event με digital object
carried out by	event	actor	crm:P14_carried_out_by	CIDOC CRM	Actor participation
took place at	excavation event	place	crm:P7_took_place_at	CIDOC CRM	Place of excavation
has time-span	event	time-span	crm:P4_has_time-span	CIDOC CRM	Time modelling
linked to period	crm:E1_CRM_Entity	gerasim:Period	gerasim:associatedWithPeriod	Local shortcut, CIDOC-	Σύνδεση CIDOC CRM οντότητας με application-

				aligned	level period node
has stratigraphic context	artefact	stratigraphic context	gerasim:hasStratigraphicContext	local shortcut / CRMarchaeo-compatible	Context relation
has find context	artefact	find context	gerasim:hasFindContext	local shortcut, subproperty of crm:P2_has_type	Categorical context
έχει τόπο εύρεσης / συνδέεται με τόπο	artefact	place	crm:P53_has_former_or_current_location	CIDOC CRM	Βασική χωρική σύνδεση αντικειμενοκεντρικών δεδομένων
has digital twin	artefact	digital twin	hdto:HP1_has_digital_twin	HDTO	Physical-digital bridge
refers to / represents	digital twin	artefact	crm:P67_refers_to, crm:P138_represents	CIDOC CRM	Digital representation semantics
documented in source record	artefact	dataset record	crm:P70_is_documented_in	CIDOC CRM	Provenance
source dataset	dataset record	dataset URI	prov:wasDerivedFrom, dcterms:source	PROV-O / DCTERMS	Source traceability
has geometry	place	geometry	geo:hasGeometry	GeoSPARQL	Spatial layer
as WKT	geometry	literal	geo:asWKT	GeoSPARQL	Coordinates

Πίνακας 7: Βασικές σχέσεις του προτεινόμενου μοντέλου

Ο πίνακας αυτός δείχνει πιο καθαρά ότι, όπου είναι δυνατό, το μοντέλο προτιμά άμεση χρήση του CIDOC CRM και των λοιπών οντολογιών που αξιοποιούνται ενώ οι local properties κρατιούνται μόνο εκεί όπου είναι πραγματικά απαραίτητες ως σύντομα application-level bridges. Οι τοπικές σχέσεις χρησιμοποιούνται μόνο όπου είναι απαραίτητο, ως ελεγχόμενες εφαρμοστικές απλοποιήσεις συγκεκριμένων CIDOC CRM ή CRMarchaeo patterns, χωρίς να υποκαθιστούν τον βασικό εννοιολογικό κορμό της οντολογίας. Η ιδιότητα gerasim:locatedInMunicipality χρησιμοποιείται αποκλειστικά μεταξύ τόπων, δηλαδή για τη σύνδεση ενός Place με μια διοικητική οντότητα τύπου Municipality, και όχι απευθείας μεταξύ artefacts και διοικητικών μονάδων.

Παρότι ο βασικός κορμός της οντολογίας στηρίζεται σε καθιερωμένες ιδιότητες του CIDOC CRM και των συμπληρωματικών οντολογιών που αξιοποιούνται στην εργασία, μετά από έλεγχο των διαθέσιμων όρων διατηρήθηκε μόνο μικρός αριθμός τοπικών shortcut ιδιοτήτων στο namespace gerasim:. Οι ιδιότητες αυτές δεν αποσκοπούν στην αντικατάσταση των υφιστάμενων προτύπων, αλλά στην εφαρμοστική απλοποίηση συγκεκριμένων εννοιολογικών μοτίβων, ώστε το μοντέλο να παραμένει αναγνώσιμο και διαχειρίσιμο στα δύο επιλεγμένα use cases.

Local shortcut property	Ρόλος στο μοντέλο	Γιατί χρησιμοποιείται	Ποιο pattern απλοποιεί	Use case
gerasim:associate	Συνδέει CIDOC	Παρέχει άμεση και αναγνώσιμη	Σύνδεση	AADA και

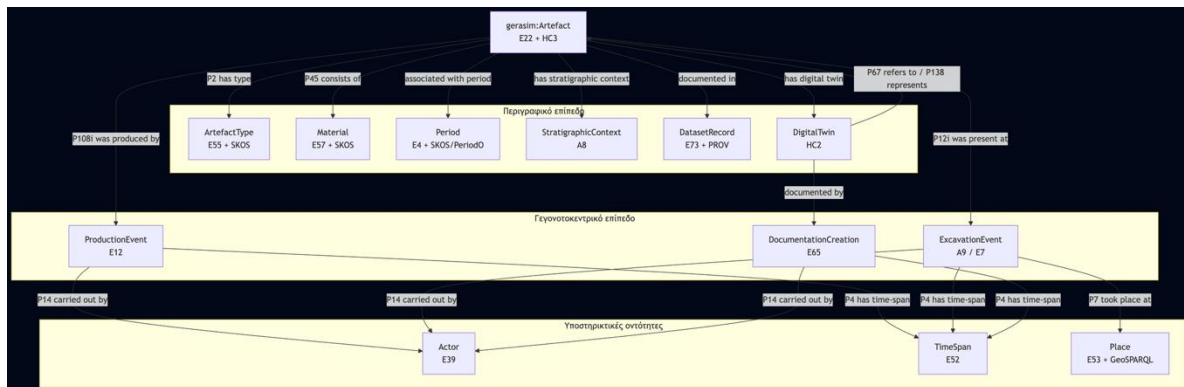
dWithPeriod	CRM οντότητα με application-level period node	σύνδεση με περίοδο, χωρίς να απαιτείται κάθε φορά πλήρης ανάπτυξη χρονολογικών ενδιάμεσων κόμβων	crm:E1_CRM_Entity με gerasim:Period, το οποίο λειτουργεί ως crm:E4_Period και skos:Concept	OpenContext
gerasim:hasFindContext	Συνδέει artefact με κατηγορία context εύρεσης	Το AADA δίνει συχνά context ως κατηγοριακή τιμή και όχι ως πλήρως αναπτυγμένο archaeological event/context pattern	Εξειδίκευση κατηγοριακής απόδοσης context, ως subproperty του crm:P2_has_type	Κυρίως AADA
gerasim:hasStratigraphicContext	Συνδέει artefact με stratigraphic unit	Επιτρέπει πιο άμεση εφαρμογή της σχέσης εύρημα-στρωματογραφικό context σε επίπεδο implementation	Πιο σύνθετο artefact-excavation-A8 pattern του CRMarchaeo	Κυρίως OpenContext

Πίνακας 8: Τοπικές shortcut ιδιότητες και αιτιολόγηση χρήσης

Ο πίνακας αυτός συνοψίζει τις τοπικές shortcut ιδιότητες που διατηρήθηκαν στην τελική εκδοχή του μοντέλου μετά τον έλεγχο έναντι των υφιστάμενων οντολογικών όρων. Δεν εισάγονται επιπλέον τοπικές κλάσεις ή ιδιότητες πέρα από όσες κρίθηκαν απολύτως απαραίτητες για τις συγκεκριμένες εφαρμοστικές ανάγκες της εργασίας.

Η ιδιότητα gerasim:associatedWithPeriod λειτουργεί ως application-level shortcut για τη σύνδεση μιας CIDOC CRM οντότητας, όπως artefact, event ή documentation-related resource, με έναν period node. Το domain της ιδιότητας περιορίζεται στο crm:E1_CRM_Entity, ώστε να παραμένει εντός της CIDOC CRM λογικής του μοντέλου και να μη δηλώνεται γενικά για οποιοδήποτε owl:Thing.

4.4.4 Εννοιολογικό διάγραμμα



Σχήμα 4: Εννοιολογικό μοντέλο του Γνωσιακού Γράφου

Το διάγραμμα παρουσιάζει την οντολογία ως μοντέλο που συνδυάζει το επίπεδο των αντικειμένων, το γεγονотоκεντρικό επίπεδο, τη χωροχρονική αναπαράσταση, την τεκμηρίωση και το επίπεδο της ψηφιακής αναπαράστασης.

4.4.5 Use cases και ενδεικτική εφαρμογή σε δεδομένα

Η οντολογία αξιολογείται πάνω σε δύο βασικά συμπληρωματικά use cases, καθώς και σε ένα επιπλέον enriched demonstrator που αναδεικνύει συνδυαστικά τις πιο σύνθετες δυνατότητες του μοντέλου. Το πρώτο βασικό use case είναι το AADA, το οποίο παρέχει artefact-oriented tabular δεδομένα και είναι κατάλληλο για την object-centric διάσταση του μοντέλου. Το δεύτερο βασικό use

case είναι ένα παράδειγμα βασισμένο σε OpenContext, το οποίο υποστηρίζει καλύτερα excavation context, actors, temporal framing και πλουσιότερη αρχαιολογική σημασιολογία. Το OpenContext project “Pınarbaşı 1994: Animal Bones” και η επιμέρους εγγραφή “Bone Ref# 1160” λειτουργούν εδώ ως βάση για την event-aware όψη της οντολογίας. Επιπλέον, το enriched demonstrator αξιοποιείται για να αναδείξει εντονότερα τη σημασιολογία της παραγωγής, της τεκμηρίωσης, της ψηφιακής αναπαράστασης και της αλληλοσύνδεσης γεγονότων.

Use case	Τι προσφέρει	Ποια μέρη της οντολογίας ενεργοποιεί
AADA	Artefact rows, identifiers, types, materials, periods, places, measurements	E22, E42, E55, E57, E4, E53, E54, provenance
OpenContext	Excavation context, actor, place, time-span, documentation logic	E7/A9, A8, E39, E52, E65, digital layer
PeriodO enrichment	Canonical period URI	E4, E52, period alignment
Optional place enrichment	Historical / external place linking	E53, GeoSPARQL, external authority linking
Digital twin layer	Photos, notes, digital surrogates	HC2_Heritage_Digital_Twin, HC3_Tangible_Heritage_Entity, E65, P67, P138

Πίνακας 9: Ενδεικτική αντιστοίχιση use cases με την οντολογία

Η λογική αυτή δείχνει ότι η οντολογία δεν αποσκοπεί στη μαζική μετατροπή ενός μόνο συνόλου δεδομένων, αλλά στην κατασκευή ενός συνεπούς μοντέλου που εφαρμόζεται σε επιλεγμένα παραδείγματα και καλύπτει τόσο object-centric όσο και event-centric πολιτιστικά δεδομένα.

4.4.6 Ενδεικτικά RDF παραδείγματα

Στο παρόν υποκεφάλαιο παρουσιάζονται ενδεικτικά RDF παραδείγματα, με σκοπό να αποσαφηνιστεί ο τρόπος με τον οποίο εφαρμόζεται στην πράξη η προτεινόμενη οντολογία. Τα παραδείγματα οργανώνονται σε τρεις συμπληρωματικές κατηγορίες. Η πρώτη περιλαμβάνει αντικειμενοκεντρικά παραδείγματα τύπου AADA, όπου το κύριο βάρος δίνεται στην περιγραφή του ίδιου του ευρήματος, του τύπου του, του υλικού όπου αυτό είναι διαθέσιμο, της χρονικής του ένταξης, του τόπου συσχέτισης ή εύρεσης και των βασικών μετρήσεών του. Η δεύτερη περιλαμβάνει περισσότερο γεγονотоκεντρικά παραδείγματα, στα οποία το εύρημα συνδέεται με ανασκαφική δραστηριότητα, στρωματογραφικό context, actor, τεκμηριωτική πράξη και ψηφιακή αναπαράσταση. Η τρίτη περιλαμβάνει ένα enriched demonstrator, το οποίο αναδεικνύει εντονότερα τη σημασιολογία της παραγωγής, της τεκμηρίωσης, της ψηφιακής αναπαράστασης και της αλληλοσύνδεσης γεγονότων.

Τα παραδείγματα αυτά δεν αποσκοπούν στην εξάντληση όλων των δυνατοτήτων της οντολογίας, αλλά στην παραστατική και συγκεντρωτική παρουσίαση του τρόπου με τον οποίο μπορούν να συνδυαστούν οι βασικές κλάσεις και σχέσεις του μοντέλου. Με τον τρόπο αυτό καθίσταται σαφές ότι η οντολογία μπορεί να υποστηρίξει τόσο object-oriented όσο και πιο σύνθετα event-based use cases.

Παραδείγματα A: Αντικειμενοκεντρικά παραδείγματα τύπου AADA

Τα παραδείγματα αυτά δείχνουν πώς μπορούν να αναπαρασταθούν ευρήματα με βάση πληροφορίες που είναι χαρακτηριστικές ενός συνόλου δεδομένων όπως το AADA. Τα ευρήματα συνδέονται με τύπο, περίοδο, τόπο, γεωμετρία, context εύρεσης ή κατηγοριακό context, διαστάσεις όπου αυτές είναι διαθέσιμες, καθώς και με source record.

[Example A.1](#) στα παραρτήματα,

[Example A.2](#) στα παραρτήματα,

[Example A.3](#) στα παραρτήματα.

Τα παραπάνω παραδείγματα αποτυπώνουν τη βασική λογική της οντολογίας ως προς την αναπαράσταση αρχαιολογικών ευρημάτων σε αντικειμενοκεντρικό επίπεδο. Το αντικείμενο μοντελοποιείται κυρίως ως E22 Human-Made Object, αποκτά αναγνωριστικό μέσω του E42 Identifier, συνδέεται με τύπο (E55 Type), με υλικό όπου αυτό είναι διαθέσιμο (E57 Material) και με πολιτισμική ή χρονολογική περίοδο. Παράλληλα, συνδέεται με τόπο (E53 Place), ο οποίος λειτουργεί στο παράδειγμα ως βασικός τόπος εύρεσης ή συσχέτισης του αντικειμένου, και με γεωμετρία μέσω GeoSPARQL, ενώ το context εύρεσης αποδίδεται ως ελεγχόμενη έννοια. Οι διαστάσεις αναπαρίστανται μέσω της κλάσης E54 Dimension, ώστε να διατηρείται η συμβατότητα με το CIDOC CRM και στις ποσοτικές περιγραφές. Τέλος, το εύρημα συνδέεται με source record, ώστε να τεκμηριώνεται και η προέλευση της πληροφορίας.

Τα KJ coordinates διατηρούνται ως ενιαίο string literal, π.χ. "3423130 6688380", ώστε να αποδίδεται η αρχική μορφή των συντεταγμένων του συνόλου δεδομένων χωρίς να δημιουργούνται δύο ασαφείς ανεξάρτητες τιμές. Η ιδιότητα gerasim:hasKJCoordinate λειτουργεί ως συμπληρωματική διατήρηση της αρχικής τιμής του συνόλου δεδομένων, ενώ η κανονική γεωμετρική αναπαράσταση αποδίδεται, όπου είναι διαθέσιμη, μέσω GeoSPARQL και της ιδιότητας geo:asWKT [17].

Τα παραδείγματα αυτά είναι κατάλληλα για να δείξουν πώς η οντολογία εφαρμόζεται σε δεδομένα που έχουν κυρίως περιγραφικό και αντικειμενοκεντρικό χαρακτήρα. Παρότι δεν ενεργοποιούν πλήρως τη γεγονотоκεντρική λογική του μοντέλου, παρουσιάζουν με καθαρό τρόπο τον ελάχιστο πυρήνα που απαιτείται για την αναπαράσταση αρχαιολογικών ευρημάτων σε RDF μορφή.

Παραδείγματα B: Γεγονотоκεντρικά παραδείγματα με ανασκαφή, τεκμηρίωση και ψηφιακή αναπαράσταση

Τα παραδείγματα αυτά δείχνουν πιο σύνθετες περιπτώσεις, στις οποίες το εύρημα δεν περιγράφεται μόνο ως αντικείμενο, αλλά εντάσσεται μέσα σε ένα δίκτυο γεγονότων, actors, τόπων, χρονικών διαστημάτων και ψηφιακών τεκμηρίων. Με τον τρόπο αυτό αναδεικνύεται πιο καθαρά η γεγονотоκεντρική λογική της οντολογίας.

[Example B.1](#) στα παραρτήματα.

[Example B.2](#) στα παραρτήματα.

[Example B.3](#) στα παραρτήματα.

Enriched demonstrator super-example

Στα παραδείγματα αυτά, το εύρημα αναπαρίσταται μέσω της εφαρμοστικής κλάσης gerasim:Artefact, η οποία καλύπτει φυσικά αρχαιολογικά ευρήματα χωρίς να τα περιορίζει απαραίτητα σε ανθρωπογενή αντικείμενα τύπου crm:E22_Human-Made_Object. Όπου απαιτείται από το παράδειγμα, το εύρημα συνδέεται και με το επίπεδο HC3_Tangible_Heritage_Entity, ώστε να αποδοθεί ρητά η φυσική όψη του πολιτιστικού αντικειμένου μέσα στο επίπεδο της ψηφιακής του αναπαράστασης. Το εύρημα συνδέεται με ανασκαφική δραστηριότητα μέσω του P12i_was_present_at, ενώ η ίδια η ανασκαφή αποδίδεται ως A9 Archaeological Excavation και E7 Activity. Έτσι αναπαρίσταται όχι μόνο η ύπαρξη

του αντικειμένου, αλλά και το γεγονός μέσω του οποίου εντάσσεται σε ένα αρχαιολογικό και τεκμηριωτικό πλαίσιο.

Η σύνδεση με A8 Stratigraphic Unit επιτρέπει να δηλωθεί ένα πιο ειδικό context εύρεσης, κάτι ιδιαίτερα χρήσιμο για αρχαιολογικά δεδομένα. Παράλληλα, ο τόπος της ανασκαφής αποδίδεται ως E53 Place και εμπλουτίζεται με γεωμετρία μέσω GeoSPARQL. Η παρουσία του actor (E39 Actor) καθιστά δυνατή τη μοντελοποίηση του ρόλου ερευνητών ή αρχαιολόγων στα επιμέρους γεγονότα του γράφου.

Ιδιαίτερη σημασία έχει και το επίπεδο της ψηφιακής αναπαράστασης. Το HC2_Heritage_Digital_Twin λειτουργεί εδώ ως ο βασικός πόρος του digital twin layer, μέσω του οποίου αποδίδεται η ψηφιακή όψη του ευρήματος. Η σύνδεση με το φυσικό αντικείμενο αποδίδεται με την ιδιότητα `hdto:HP1_has_digital_twin`, ενώ οι σχέσεις `crm:P67_refers_to` και `crm:P138_represents` χρησιμοποιούνται για να καταστήσουν σαφή τη σημασιολογική σύνδεση ανάμεσα στο digital twin και στο artefact. Στην παρούσα εκδοχή του μοντέλου, το digital twin δεν ταυτίζεται με την κλάση `crm:E36_Visual_Item`, καθώς αυτή μπορεί να αξιοποιηθεί καταλληλότερα για επιμέρους οπτικά τεκμήρια, όπως φωτογραφίες ή άλλες συγκεκριμένες visual representations.

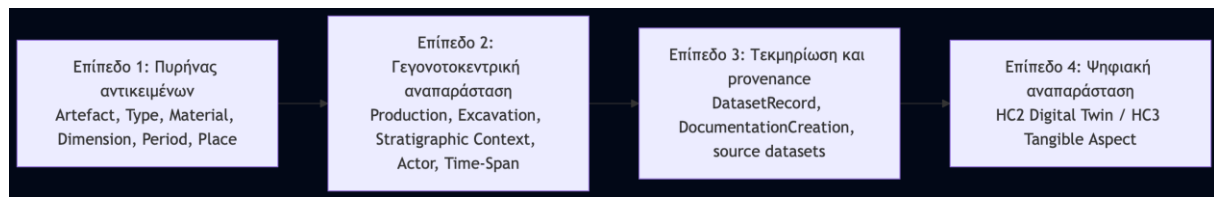
Τα παραδείγματα αυτά είναι ιδιαίτερα χρήσιμα, επειδή δείχνουν πώς η οντολογία μπορεί να υποστηρίξει σύνθετα ερωτήματα, όπως ποια ευρήματα συνδέονται με συγκεκριμένες ανασκαφές, ποια συνδέονται με τον ίδιο actor, ποια έχουν ψηφιακή αναπαράσταση, ποια τεκμηριώνονται μέσω συγκεκριμένης πράξης δημιουργίας και ποια τοποθετούνται στην ίδια χρονική περίοδο. Με τον τρόπο αυτό αναδεικνύεται η γεγονотоκεντρική και πολυεπίπεδη λογική του μοντέλου.

Επίπεδα της προτεινόμενης οντολογίας

Τα προηγούμενα παραδείγματα δείχνουν ότι η προτεινόμενη οντολογία δεν λειτουργεί σε ένα μόνο επίπεδο. Αντίθετα, οργανώνεται σε διακριτά αλλά αλληλένδετα επίπεδα αναπαράστασης, τα οποία επιτρέπουν τη σταδιακή μετάβαση από την απλή περιγραφή αντικειμένων σε μια πιο σύνθετη αναπαράσταση γεγονότων, τεκμηρίωσης και ψηφιακών αναπαραστάσεων.

Στο πρώτο επίπεδο βρίσκεται ο πυρήνας των πολιτιστικών αντικειμένων, δηλαδή τα ίδια τα artefacts, οι τύποι τους, τα υλικά τους, οι διαστάσεις τους, οι χρονικές τους εντάξεις και οι βασικές χωρικές τους σχέσεις. Στο δεύτερο επίπεδο βρίσκεται το γεγονотоκεντρικό επίπεδο, στο οποίο εντάσσονται η παραγωγή, η ανασκαφή, τα στρωματογραφικά contexts, οι actors και τα χρονικά διαστήματα. Στο τρίτο επίπεδο βρίσκεται το τεκμηριωτικό και provenance επίπεδο, το οποίο καλύπτει source records, documentation creation events και πηγές δεδομένων. Τέλος, στο τέταρτο επίπεδο βρίσκεται το ψηφιακό επίπεδο, όπου το φυσικό αντικείμενο συνδέεται με ένα digital twin ή άλλη ψηφιακή αναπαράσταση.

Η πολυεπίπεδη αυτή οργάνωση είναι σημαντική, επειδή επιτρέπει στην οντολογία να εφαρμοστεί τόσο σε datasets με περισσότερο περιγραφικό χαρακτήρα όσο και σε datasets με πλουσιότερο event-based αρχαιολογικό context. Έτσι, το AADA μπορεί να αξιοποιηθεί κυρίως στο επίπεδο των αντικειμένων και των βασικών γνωρισμάτων τους, ενώ το OpenContext και το enriched demonstrator ενεργοποιούν περισσότερο το γεγονотоκεντρικό, τεκμηριωτικό και ψηφιακό επίπεδο του μοντέλου.



Σχήμα 5: Επίπεδα της προτεινόμενης οντολογίας

Το διάγραμμα αυτό συνοψίζει τη συνολική αρχιτεκτονική του μοντέλου. Δείχνει ότι η οντολογία δεν είναι μια απλή επίπεδη λίστα κλάσεων, αλλά μια δομή στην οποία τα αντικείμενα, τα γεγονότα, η τεκμηρίωση και η ψηφιακή αναπαράσταση συνδέονται μεταξύ τους σε διαφορετικά επίπεδα. Με τον τρόπο αυτό καθίσταται σαφές ότι η προτεινόμενη οντολογία μπορεί να λειτουργήσει ως συνεπές σχήμα αναπαράστασης πολιτιστικής κληρονομιάς για περισσότερα από ένα είδη δεδομένων.

Κεφάλαιο 5ο: Υλοποίηση του Γνωσιακού Γράφου

Το παρόν κεφάλαιο παρουσιάζει το πρακτικό στάδιο της εργασίας, δηλαδή την υλοποίηση του προτεινόμενου Γνωσιακού Γράφου με βάση τη σχεδιαστική λογική και την οντολογία που αναπτύχθηκαν στο προηγούμενο κεφάλαιο. Ενώ στο Κεφάλαιο 4 δόθηκε έμφαση στον εννοιολογικό σχεδιασμό του μοντέλου, στις βασικές οντολογικές επιλογές και στα competency questions που καθοδηγούν την ανάπτυξή του, στο παρόν κεφάλαιο το ενδιαφέρον μετατοπίζεται στην εφαρμογή αυτών των επιλογών πάνω σε συγκεκριμένα δεδομένα και στην παραγωγή του τελικού RDF αποτελέσματος.

Η υλοποίηση της εργασίας δεν αντιμετωπίζεται ως μία απλή τεχνική διαδικασία μετατροπής αρχείων σε RDF, αλλά ως το στάδιο στο οποίο ελέγχεται στην πράξη η καταλληλότητα της οντολογίας και η δυνατότητά της να οργανώνει ετερογενή δεδομένα πολιτιστικής κληρονομιάς σε ένα ενιαίο, συνεκτικό και σημασιολογικά αξιοποιήσιμο σχήμα. Για τον λόγο αυτό, το κεφάλαιο επικεντρώνεται όχι μόνο στην ίδια τη διαδικασία μετασχηματισμού, αλλά και στον τρόπο επιλογής των δεδομένων, στην οργάνωση του τελικού γράφου, στον έλεγχο της ορθότητας της αναπαράστασης και στα εργαλεία που υποστήριζαν την υλοποίηση.

Ιδιαίτερη σημασία έχει το γεγονός ότι η παρούσα εργασία δεν επιδιώκει τη μαζική μετατροπή ολόκληρων datasets, αλλά την κατασκευή ενός ελεγχόμενου και αντιπροσωπευτικού demonstrator. Για τον σκοπό αυτό, αξιοποιούνται επιλεγμένα παραδείγματα από τα δύο use cases που παρουσιάστηκαν στο προηγούμενο κεφάλαιο, δηλαδή το AADA και το OpenContext. Η επιλογή αυτή είναι συμβατή με τον στόχο της εργασίας, καθώς επιτρέπει να φανεί με καθαρό τρόπο πώς το προτεινόμενο οντολογικό μοντέλο εφαρμόζεται τόσο σε object-centric όσο και σε event-based πολιτιστικά δεδομένα, χωρίς να απαιτείται εξαντλητική κάλυψη ολόκληρων πληροφοριακών συλλογών.

Με βάση τα παραπάνω, το παρόν κεφάλαιο παρουσιάζει διαδοχικά την επιλογή και τη διαμόρφωση του υλικού που χρησιμοποιήθηκε, τη μετατροπή των δεδομένων σε RDF, την οργάνωση και αποθήκευση του Γνωσιακού Γράφου, τον έλεγχο της ορθότητας και συνέπειας της αναπαράστασης, καθώς και τα εργαλεία που αξιοποιήθηκαν κατά τη διαδικασία υλοποίησης.

5.1 Επιλογή και δημιουργία συνόλου δεδομένων

Το πρώτο βήμα της υλοποίησης αφορούσε την επιλογή και διαμόρφωση του συνόλου δεδομένων που θα χρησιμοποιούνταν για την πρακτική εφαρμογή της οντολογίας. Στο πλαίσιο της παρούσας εργασίας, η επιλογή αυτή δεν έγινε με στόχο την πλήρη κάλυψη όλων των διαθέσιμων εγγραφών από τα επιλεγμένα datasets, αλλά με σκοπό τη δημιουργία ενός μικρού, ελεγχόμενου και σημασιολογικά αντιπροσωπευτικού συνόλου παραδειγμάτων. Με άλλα λόγια, η εργασία δεν στηρίχθηκε σε μαζική εξαγωγή και μετατροπή δεδομένων, αλλά σε επιλεγμένες εγγραφές που κρίθηκαν κατάλληλες για να ενεργοποιήσουν στην πράξη τα βασικά επίπεδα και τις σχέσεις της προτεινόμενης οντολογίας.

Για την υλοποίηση χρησιμοποιήθηκαν δύο διαφορετικές αλλά συμπληρωματικές πηγές δεδομένων. Η πρώτη ήταν το Archaeological Artefact Database of Finland (AADA), το οποίο λειτούργησε ως βασική πηγή για την αντικειμενοκεντρική πλευρά του μοντέλου. Το AADA παρείχε εγγραφές που περιγράφουν μεμονωμένα αρχαιολογικά ευρήματα, μαζί με βασικές πληροφορίες όπως identifiers, τύπους, υλικά, χρονικές εντάξεις, τόπους και μετρήσεις. Η φύση αυτών των δεδομένων το καθιστά ιδιαίτερα κατάλληλο για την αναπαράσταση του βασικού πυρήνα των artefacts και των χαρακτηριστικών τους, σύμφωνα με τη λογική που έχει ήδη αναπτυχθεί στο Κεφάλαιο 4.

Η δεύτερη πηγή ήταν το OpenContext, το οποίο αξιοποιήθηκε ως συμπληρωματικό και περισσότερο γεγονотоκεντρικό use case. Σε αντίθεση με το AADA, το OpenContext παρέχει richer archaeological context, επιτρέποντας τη σύνδεση ενός ευρήματος με ανασκαφικές δραστηριότητες, actors, stratigraphic units, χρονικά διαστήματα και τεκμηριωτικές πράξεις. Η ένταξη του OpenContext στο σύνολο υλοποίησης ήταν κρίσιμη, διότι έδωσε τη δυνατότητα να ελεγχθεί στην πράξη η event-based διάσταση της οντολογίας, δηλαδή ακριβώς εκείνα τα στοιχεία που δεν μπορούν να αναδειχθούν επαρκώς μόνο μέσα από ένα object-centric σύνολο δεδομένων.

Με βάση αυτές τις δύο πηγές, το τελικό υλικό της υλοποίησης οργανώθηκε ως ένα σύνολο επιλεγμένων παραδειγμάτων RDF, τα οποία κατασκευάστηκαν έτσι ώστε να είναι όσο το δυνατόν πιο πιστά στα πραγματικά δεδομένα των πηγών τους. Στην περίπτωση του AADA, τα παραδείγματα σχεδιάστηκαν με λογική faithful object-centric transformation, δηλαδή ως αναπαραστάσεις που ακολουθούν στενά τις διαθέσιμες τιμές των αντίστοιχων εγγραφών και αποδίδουν σε RDF τις βασικές πληροφορίες των artefacts χωρίς αυθαίρετο σημασιολογικό εμπλουτισμό. Στην περίπτωση του OpenContext, τα παραδείγματα οργανώθηκαν έτσι ώστε να αναδεικνύουν, πέρα από το ίδιο το εύρημα, και το σχετικό excavation context, τη stratigraphic μονάδα, τον actor, την documentation creation πράξη και, όπου ήταν εφικτό, το επίπεδο της ψηφιακής αναπαράστασης.

Η διαδικασία αυτή οδήγησε τελικά στη δημιουργία ενός μικρού αλλά πλούσιου demonstrator γράφου, στον οποίο συνυπάρχουν παραδείγματα διαφορετικής σημασιολογικής πυκνότητας. Ορισμένα παραδείγματα ενεργοποιούν κυρίως τον βασικό πυρήνα της οντολογίας, δηλαδή artefact, type, material, period, place και dimension, ενώ άλλα ενεργοποιούν επιπλέον τα επίπεδα excavation, documentation, provenance και digital representation. Η επιλογή αυτή είναι ιδιαίτερα σημαντική, καθώς επιτρέπει να φανεί ότι η προτεινόμενη οντολογία δεν σχεδιάστηκε για ένα μόνο είδος δεδομένων, αλλά μπορεί να λειτουργήσει ως κοινός εννοιολογικός πυρήνας για διαφορετικές όψεις της πολιτιστικής πληροφορίας.

Επομένως, το σύνολο δεδομένων της υλοποίησης δεν πρέπει να νοηθεί ως μία κλασική, πλήρης και ομοιογενής βάση δεδομένων, αλλά ως ένα επιλεγμένο σώμα παραδειγμάτων που κατασκευάστηκε με στόχο να ελέγξει τη λειτουργικότητα της οντολογίας. Η επιλογή αυτή είναι πλήρως συμβατή με τον χαρακτήρα της εργασίας, καθώς επιτρέπει να δοκιμαστούν στην πράξη οι κύριες σχεδιαστικές αποφάσεις του μοντέλου και να δημιουργηθεί μία κατάλληλη βάση για τα SPARQL ερωτήματα και την αξιολόγηση που θα ακολουθήσει στα επόμενα κεφάλαια.

Πηγή δεδομένων	Τύπος δεδομένων	Ρόλος στην εργασία	Τι αξιοποιήθηκε τελικά
Archaeological Artefact Database of Finland (AADA)	Object-centric tabular dataset	Βασική πηγή για την περιγραφή αρχαιολογικών ευρημάτων	Identifiers, types, materials, periods, places, measurements, find contexts
OpenContext	Event-oriented archaeological dataset	Συμπληρωματικό use case για τη γεγονотоκεντρική διάσταση του μοντέλου	Excavation events, stratigraphic contexts, actors, documentation processes, richer contextual links
PeriodO	Authority / controlled temporal resource	Εμπλουτισμός της χρονικής αναπαράστασης	Canonical URIs για περιόδους και σύνδεση period semantics
GeoSPARQL spatial layer	Γεωχωρικό λεξιλόγιο / πρότυπο	Υποστήριξη χωρικής αναπαράστασης	Places, geometries, WKT literals
HDTO	Οντολογικό επίπεδο ψηφιακής αναπαράστασης	Υποστήριξη digital twin layer	HC2 Heritage Digital Twin και HC3 Tangible Heritage Entity

CIDOC CRM / CRMarchaeo / SKOS / PROV-O / DCAT	Οντολογίες και vocabularies	Εννοιολογικό και σημασιολογικό υπόβαθρο της υλοποίησης	Βασικές κλάσεις, ιδιότητες, provenance και ελεγχόμενα λεξιλόγια
---	--------------------------------	--	--

Πίνακας 10: Σύνολα δεδομένων και ρόλος τους στην υλοποίηση

5.2 Μετατροπή των δεδομένων σε RDF

Αφού προσδιορίστηκε το σύνολο των επιλεγμένων παραδειγμάτων που θα χρησιμοποιούνταν για την υλοποίηση, το επόμενο στάδιο της εργασίας αφορούσε τη μετατροπή τους σε RDF, σύμφωνα με την οντολογία που σχεδιάστηκε στο προηγούμενο κεφάλαιο. Η διαδικασία αυτή αποτέλεσε το βασικό σημείο μετάβασης από την περιγραφική πληροφορία των αρχικών δεδομένων σε μία ρητά δομημένη σημασιολογική αναπαράσταση, στην οποία οι οντότητες, οι σχέσεις και τα επιμέρους επίπεδα του μοντέλου αποδίδονται με τυπικό και μηχανικά επεξεργάσιμο τρόπο.

Στο πλαίσιο της παρούσας εργασίας, η μετατροπή σε RDF δεν υλοποιήθηκε ως πλήρως αυτοματοποιημένη διαδικασία μαζικού μετασχηματισμού ολόκληρων συνόλων δεδομένων, αλλά ως ελεγχόμενη διαδικασία εννοιολογικής αντιστοίχισης και χειροκίνητης ή ημι-χειροκίνητης παραγωγής επιλεγμένων παραδειγμάτων σε Turtle συντακτικό. Η επιλογή αυτή ήταν σκόπιμη. Ο κύριος στόχος της εργασίας δεν ήταν να αναπτυχθεί ένα γενικής χρήσης ETL (Extract – Transform – Load) pipeline για χιλιάδες εγγραφές, αλλά να αποδειχθεί ότι το προτεινόμενο οντολογικό μοντέλο μπορεί να εφαρμοστεί με συνέπεια και επάρκεια σε πραγματικά πολιτιστικά δεδομένα, αποδίδοντας με σαφήνεια τόσο την object-centric όσο και την event-based όψη τους.

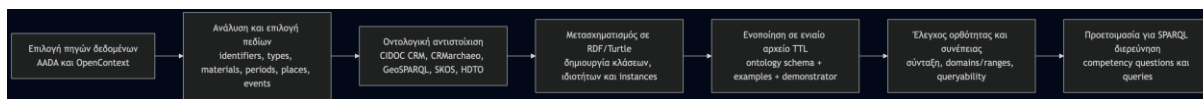
Η μετατροπή ακολούθησε ως βασική αρχή τη διατήρηση της εννοιολογικής συνέπειας με την οντολογία. Αυτό σημαίνει ότι κάθε πληροφοριακό στοιχείο των αρχικών δεδομένων εξετάστηκε σε σχέση με το αν αντιστοιχεί άμεσα σε υπάρχουσα κλάση ή ιδιότητα των χρησιμοποιούμενων οντολογιών ή αν απαιτεί κάποια ελεγχόμενη εφαρμοστική απλοποίηση μέσω των τοπικών όρων του namespace gerasim:. Με άλλα λόγια, η λογική της μετατροπής δεν ήταν να δημιουργηθούν αυθαίρετες RDF τριάδες βάσει των πεδίων των αρχικών αρχείων, αλλά να ενταχθεί κάθε πληροφορία σε ένα σημασιολογικά αιτιολογημένο μοτίβο αναπαράστασης.

Για τα παραδείγματα τύπου AADA, η διαδικασία μετατροπής ήταν κυρίως αντικειμενοκεντρική. Κάθε επιλεγμένο εύρημα μετασχηματίστηκε σε έναν RDF πόρο που δηλώνεται ως `crm:E22_Human-Made_Object` και συνδέεται με `identifier`, `type`, `material`, `period`, `place`, `dimension` και `source record`. Η πληροφορία αυτή αποδόθηκε με χρήση ήδη υπάρχουσών ιδιοτήτων του CIDOC CRM, όπως `crm:P1_is_identified_by`, `crm:P2_has_type`, `crm:P45_consists_of`, `crm:P43_has_dimension`, `crm:P70_is_documented_in` και `crm:P53_has_former_or_current_location`. Όπου τα δεδομένα του AADA παρείχαν κατηγοριακές τιμές που δεν αντιστοιχούσαν σε πλήρως ανεπτυγμένα γεγονотоκεντρικά μοτίβα, χρησιμοποιήθηκαν ελεγχόμενα `local shortcuts`, όπως `gerasim:associatedWithPeriod` και `gerasim:hasFindContext`, ώστε η αναπαράσταση να παραμείνει αναγνώσιμη χωρίς να χάνει τη συνάφειά της με το υποκείμενο εννοιολογικό μοντέλο.

Αντίθετα, στα παραδείγματα τύπου OpenContext, όπου εμφανίζονται οστά και άλλα φυσικά αρχαιολογικά κατάλοιπα, τα ευρήματα δεν δηλώνονται γενικά ως `crm:E22_Human-Made_Object`. Σε αυτές τις περιπτώσεις χρησιμοποιείται η εφαρμοστική κλάση `gerasim:Artefact`, η οποία ευθυγραμμίζεται με το γενικότερο `crm:E18_Physical_Thing`, ώστε να καλύπτει τόσο ανθρωπογενή αντικείμενα όσο και μη ανθρωπογενή φυσικά αρχαιολογικά ευρήματα.

Για τα πεδία συντεταγμένων τύπου KJ του AADA, οι δύο αριθμητικές τιμές δεν αποδόθηκαν ως δύο ανεξάρτητα string literals, καθώς αυτό θα καθιστούσε ασαφές ποια τιμή αντιστοιχεί σε ποια συντεταγμένη. Αντίθετα, διατηρήθηκαν ως ενιαίο string literal, σύμφωνα με τη μορφή "x y", ενώ η κανονική γεωμετρική αναπαράσταση αποδίδεται μέσω GeoSPARQL όπου αυτό είναι διαθέσιμο.

Για τα παραδείγματα τύπου OpenContext, η μετατροπή ακολούθησε πιο σύνθετη, γεγονотоκεντρική λογική. Το εύρημα δεν αποδόθηκε μόνο ως αντικείμενο, αλλά συνδέθηκε και με excavation event, stratigraphic context, actor, time-span, documentation creation event και digital representation. Στο επίπεδο αυτό αξιοποιήθηκαν κυρίως οι κλάσεις crma:A9_Archaeological_Excavation, crma:A8_Stratigraphic_Unit, crm:E65_Creation, crm:E39_Actor, crm:E52_Time-Span και hdto:HC2_Heritage_Digital_Twin, μαζί με τις αντίστοιχες ιδιότητες του CIDOC CRM και του CRMarchaeo. Με αυτόν τον τρόπο, η μετατροπή έδειξε ότι η ίδια οντολογία μπορεί να χρησιμοποιηθεί όχι μόνο για την περιγραφή isolated artefacts, αλλά και για την αναπαράσταση richer archaeological semantics, στις οποίες το εύρημα εντάσσεται μέσα σε ένα πλέγμα γεγονότων και τεκμηριωτικών σχέσεων.



Σχήμα 6: Διάγραμμα ροής διαδικασίας υλοποίησης

Το τελικό αποτέλεσμα της μετατροπής καταγράφηκε σε αρχείο Turtle, το οποίο λειτουργεί ταυτόχρονα ως υλοποίηση της οντολογίας και ως demonstrator γράφος δεδομένων. Στο αρχείο αυτό περιλαμβάνονται τόσο οι δηλώσεις του ίδιου του οντολογικού σχήματος όσο και τα ενδεικτικά instances που χρησιμοποιούνται για την πρακτική αξιολόγηση του μοντέλου. Η επιλογή της Turtle ως μορφής σειριοποίησης κρίθηκε κατάλληλη, επειδή είναι αρκετά εκφραστική ώστε να αποδίδει με σαφήνεια το σύνολο των RDF τριάδων, ενώ ταυτόχρονα παραμένει αναγνώσιμη από τον άνθρωπο και εύχρηστη για έλεγχο, επαλήθευση και φόρτωση σε σημασιολογικά εργαλεία.

Κατά τη διαδικασία μετατροπής δόθηκε ιδιαίτερη προσοχή και στην ονοματοδοσία των πόρων. Για τον σκοπό αυτό χρησιμοποιήθηκαν δύο διακριτά namespaces: το namespace της οντολογίας, δηλαδή το gerasim:, και το namespace των instances, δηλαδή το gres:. Η διάκριση αυτή ήταν σημαντική, επειδή επιτρέπει να ξεχωρίζει καθαρά το επίπεδο του σχήματος από το επίπεδο των δεδομένων. Με άλλα λόγια, οι κλάσεις και οι ιδιότητες της προτεινόμενης οντολογίας ανήκουν στο πρώτο namespace, ενώ τα συγκεκριμένα artefacts, events, places, records και λοιποί επιμέρους πόροι των παραδειγμάτων ανήκουν στο δεύτερο. Η πρακτική αυτή καθιστά τον γράφο περισσότερο συνεπή και διαφανή, τόσο σε επίπεδο τεκμηρίωσης όσο και σε επίπεδο μελλοντικής επαναχρησιμοποίησης.

Ιδιαίτερη σημασία είχε επίσης ο έλεγχος του βαθμού επέκτασης των χρησιμοποιούμενων οντολογιών. Η τελική εκδοχή της υλοποίησης περιόρισε τους τοπικούς όρους μόνο στις περιπτώσεις όπου αυτοί ήταν πράγματι απαραίτητοι ως εφαρμοστικές συντομεύσεις. Αυτό σημαίνει ότι η μετατροπή σε RDF βασίστηκε κατά το δυνατόν στην άμεση χρήση υπαρχουσών κλάσεων και ιδιοτήτων του CIDOC CRM, του CRMarchaeo, του GeoSPARQL, του SKOS, του PROV-O και του HDT, ενώ οι custom όροι διατηρήθηκαν μόνο εκεί όπου παρείχαν σαφές πρακτικό όφελος για τη μοντελοποίηση των δύο use cases. Με αυτόν τον τρόπο, το τελικό RDF αποτέλεσμα παραμένει συμβατό με τα καθιερωμένα πρότυπα, χωρίς όμως να χάνει τη λειτουργικότητα που απαιτεί η συγκεκριμένη εφαρμογή.

Αρχικό πεδίο / πληροφορία	Οντολογικός όρος ή ιδιότητα	Οντολογία	Ρόλος στη σημασιολογική αναπαράσταση
---------------------------	-----------------------------	-----------	--------------------------------------

Αναγνωριστικό ευρήματος	crm:P1_is_identified_by	CIDOC CRM	Ταυτοποίηση του artefact μέσω E42 Identifier
Τύπος αντικειμένου	crm:P2_has_type	CIDOC CRM	Κατηγοριοποίηση του artefact με E55 Type
Υλικό	crm:P45_consists_of	CIDOC CRM	Δήλωση της υλικής σύστασης με E57 Material
Διάσταση / μέτρηση	crm:P43_has_dimension	CIDOC CRM	Ποσοτική περιγραφή με E54 Dimension
Χρονική περίοδος	gerasim:associatedWithPeriod	Custom shortcut aligned with CIDOC CRM / PeriodO	Άμεση σύνδεση με E4 Period
Τόπος εύρεσης ή συσχέτισης	crm:P53_has_former_or_current_location	CIDOC CRM	Βασική object-centric χωρική σύνδεση
Γεωμετρία τόπου	geo:hasGeometry / geo:asWKT	GeoSPARQL	Χωρική περιγραφή τόπου με γεωμετρία
Context εύρεσης	gerasim:hasFindContext	Custom shortcut + SKOS	Ελαφριά κατηγοριοποίηση find context
Ανασκαφική δραστηριότητα	crm:P12i_was_present_at	CIDOC CRM	Σύνδεση artefact με excavation event
Στρωματογραφική μονάδα	gerasim:hasStratigraphicContext	Custom shortcut / CRMarchaeo-compatible	Σύνδεση artefact με stratigraphic unit
Χρονικό διάστημα γεγονότος	crm:P4_has_time-span	CIDOC CRM	Χρονικός προσδιορισμός event
Φορέας / πρόσωπο που συμμετέχει	crm:P14_carried_out_by	CIDOC CRM	Σύνδεση event με actor
Ψηφιακή αναπαράσταση	hdto:HC2_Heritage_Digital_Twin	HDTO	Σύνδεση φυσικού αντικειμένου με HC2 digital twin
Ψηφιακό τεκμήριο που αναφέρεται στο αντικείμενο	crm:P67_refers_to / crm:P138_represents	CIDOC CRM	Σημασιολογική σύνδεση digital representation με artefact
Τεκμηριωτική πράξη δημιουργίας	crm:P94_has_created / crm:P94i_was_created_by	CIDOC CRM	Σύνδεση documentation creation με digital object
Source record	crm:P70_is_documented_in	CIDOC CRM	Σύνδεση artefact με record τεκμηρίωσης
Πηγή προέλευσης record	dcterms:source / prov:wasDerivedFrom	DCTERMS / PROV-O	Δήλωση provenance και προέλευσης των δεδομένων

Πίνακας 11: Ενδεικτικές αντιστοιχίσεις πεδίων των δεδομένων σε RDF και οντολογικούς όρους

Συνολικά, η μετατροπή των δεδομένων σε RDF αποτέλεσε το στάδιο στο οποίο ο θεωρητικός σχεδιασμός της οντολογίας δοκιμάστηκε στην πράξη. Μέσα από την παραγωγή των επιλεγμένων Turtle παραδειγμάτων κατέστη δυνατό να φανεί ότι η προτεινόμενη οντολογία μπορεί να οργανώσει ετερογενείς πληροφορίες πολιτιστικής κληρονομιάς σε ενιαίο γράφο, να διατηρήσει συνέπεια ως προς τα χρησιμοποιούμενα σημασιολογικά πρότυπα και να δημιουργήσει κατάλληλη βάση για τη μετέπειτα αποθήκευση, επαλήθευση και διερεύνηση των δεδομένων μέσω SPARQL.

5.3 Οργάνωση και αποθήκευση του Γνωσιακού Γράφου

Μετά τη μετατροπή των επιλεγμένων δεδομένων σε RDF, το επόμενο βήμα της υλοποίησης αφορούσε την οργάνωση και την αποθήκευση του παραγόμενου Γνωσιακού Γράφου. Το στάδιο αυτό είναι ουσιώδες, διότι δεν αρκεί μόνο να παραχθούν σωστές RDF τριάδες, απαιτείται επίσης να οργανωθούν με τρόπο συνεπή, ώστε ο γράφος να είναι αναγνώσιμος, επεκτάσιμος, ελέγξιμος και κατάλληλος για μεταγενέστερη διερεύνηση μέσω SPARQL ερωτημάτων.

Στο πλαίσιο της παρούσας εργασίας, ο Γνωσιακός Γράφος οργανώθηκε σε ένα ενιαίο αρχείο Turtle, το οποίο περιλαμβάνει τόσο το οντολογικό σχήμα όσο και τα ενδεικτικά δεδομένα που χρησιμοποιούνται για την αξιολόγηση του μοντέλου. Η επιλογή αυτή έγινε σκόπιμα, επειδή το μέγεθος της υλοποίησης παραμένει ελεγχόμενο και ο στόχος της εργασίας είναι πρωτίστως αποδεικτικός και πειραματικός. Με άλλα λόγια, δεν επελέγη μια αρχιτεκτονική διάσπασης σε πολλαπλά αρχεία ή σε ανεξάρτητα named graphs, καθώς κάτι τέτοιο θα ήταν περισσότερο χρήσιμο σε μια μεγαλύτερη παραγωγική υποδομή παρά στη συγκεκριμένη εργασία.

Η οργάνωση του αρχείου ακολουθεί σαφή διαστρωμάτωση. Στην αρχή του αρχείου δηλώνονται τα prefixes και το βασικό metadata πλαίσιο της οντολογίας. Στη συνέχεια αποδίδονται οι κλάσεις και οι ιδιότητες που αποτελούν τον πυρήνα του μοντέλου, καθώς και οι σχέσεις τους με τις εξωτερικές οντολογίες που επαναχρησιμοποιούνται. Ακολουθούν τα παραδείγματα instances, τα οποία λειτουργούν ως demonstrators της εφαρμογής της οντολογίας πάνω στα δύο επιλεγμένα use cases. Η διάταξη αυτή καθιστά το αρχείο περισσότερο κατανοητό, επειδή επιτρέπει στον αναγνώστη να μεταβαίνει από το επίπεδο του σχήματος στο επίπεδο των δεδομένων με σαφή και φυσικό τρόπο.

Για την καλύτερη οργάνωση του γράφου χρησιμοποιήθηκε, επίσης, σαφής διαχωρισμός namespaces. Όπως αναφέρθηκε και στο προηγούμενο υποκεφάλαιο, το namespace gerasim: χρησιμοποιείται για την οντολογία, δηλαδή για τις τοπικές κλάσεις και ιδιότητες που διατηρήθηκαν στην τελική έκδοχή του μοντέλου, ενώ το namespace gres: χρησιμοποιείται για τα επιμέρους instances του γράφου, όπως artefacts, excavation events, places, time-spans, records και digital representations. Η διάκριση αυτή συμβάλλει σημαντικά στην καθαρότητα της αποθήκευσης, επειδή επιτρέπει να ξεχωρίζουν άμεσα οι όροι του σχήματος από τους πόρους των δεδομένων.

Ιδιαίτερη προσοχή δόθηκε και στη λογική ταυτοποίησης των πόρων. Κάθε βασική οντότητα του γράφου απέκτησε δικό της URI, ώστε να είναι δυνατή η σαφής αναφορά και διασύνδεσή της με άλλες οντότητες. Για παράδειγμα, τα artefacts, οι identifiers, τα places, τα documentation records και τα events δεν αποδίδονται ως απλές literal τιμές, αλλά ως διακριτοί πόροι, όπου αυτό είναι εννοιολογικά σκόπιμο. Η επιλογή αυτή ενισχύει τη graph-based φύση του μοντέλου, καθώς επιτρέπει οι σχέσεις να αποδίδονται ρητά ανάμεσα σε URI-identified οντότητες και όχι να περιορίζονται σε επίπεδο επιπέδων πεδίων ή συμβολοσειρών.

Όσον αφορά την αποθήκευση, το αρχείο Turtle ελέγχθηκε ώστε να είναι συντακτικά συμβατό με εργαλεία σημασιολογικού ιστού, όπως το Protégé και RDF parsers που υποστηρίζουν Turtle/RDF serialization. Ο έλεγχος αυτός ήταν αναγκαίος, επειδή σε τέτοιου τύπου υλοποιήσεις ακόμη και μικρές συντακτικές ασυνέπειες, όπως λανθασμένη χρήση URI, ειδικών χαρακτήρων ή λανθασμένων τερματισμών τριάδων, μπορούν να καταστήσουν ολόκληρο το αρχείο μη αναγνώσιμο. Για τον λόγο αυτό, η τελική μορφή του γράφου οργανώθηκε με τρόπο που να είναι όχι μόνο εννοιολογικά σωστός, αλλά και τεχνικά έγκυρος ως προς τη σύνταξη RDF/Turtle.

Η αποθήκευση του Γνωσιακού Γράφου σε Turtle παρέχει και ένα ακόμη σημαντικό πλεονέκτημα: επιτρέπει την εύκολη εισαγωγή του σε triple stores ή άλλα περιβάλλοντα επεξεργασίας RDF. Παρότι

στην παρούσα εργασία ο κύριος στόχος δεν ήταν η ανάπτυξη πλήρους production αποθετήριου, η μορφή αποθήκευσης που επιλέχθηκε αφήνει ανοιχτή τη δυνατότητα μελλοντικής επέκτασης. Το αρχείο μπορεί να φορτωθεί σε ένα RDF store, να αποτελέσει βάση για SPARQL queries, να επεκταθεί με νέα instances ή να διαχωριστεί σε επιμέρους μονάδες αν η υλοποίηση αποκτήσει μεγαλύτερη κλίμακα.

Σε εννοιολογικό επίπεδο, η οργάνωση του γράφου αποτυπώνει και την πολυεπίπεδη λογική της οντολογίας. Τα object-centric δεδομένα του AADA καταλαμβάνουν κυρίως το επίπεδο του πυρήνα των αντικειμένων, με σχέσεις προς type, material, period, place και source record. Αντίθετα, τα OpenContext-based παραδείγματα ενεργοποιούν περισσότερο το γεγονотоκεντρικό επίπεδο, καθώς περιλαμβάνουν excavation events, stratigraphic units, actors, time-spans, documentation creation και digital representation. Παρότι όλα αυτά συνυπάρχουν στο ίδιο αρχείο, η οργάνωσή τους παραμένει συνεπής, επειδή βασίζονται στον ίδιο οντολογικό πυρήνα και στις ίδιες σημασιολογικές αρχές.

Σημαντικό στοιχείο της αποθήκευσης ήταν και η διατήρηση των εξωτερικών συνδέσεων όπου αυτές ήταν διαθέσιμες ή κρίνονταν χρήσιμες. Για παράδειγμα, σε ορισμένα σημεία ο γράφος μπορεί να παραπέμπει σε canonical resources, όπως URI περιόδων από το PeriodO ή σε πηγές δεδομένων από τα αντίστοιχα σύνολα δεδομένων. Οι σύνδεσμοι αυτοί δεν λειτουργούν απλώς ως τεχνικά μεταδεδομένα, αλλά ενισχύουν τη διαλειτουργικότητα του μοντέλου, καθώς το τοποθετούν μέσα σε ένα ευρύτερο linked data περιβάλλον. Έτσι, ο γράφος δεν περιορίζεται σε ένα κλειστό σύστημα τοπικών identifiers, αλλά αποκτά και εξωστρεφή σημασιολογική συνδεσιμότητα.

Συνολικά, η οργάνωση και η αποθήκευση του Γνωσιακού Γράφου σχεδιάστηκαν έτσι ώστε να εξυπηρετούν ταυτόχρονα τέσσερις στόχους: εννοιολογική σαφήνεια, συντακτική ορθότητα, πρακτική αναγνωσιμότητα και δυνατότητα μελλοντικής επαναχρησιμοποίησης. Το τελικό Turtle αρχείο δεν λειτουργεί μόνο ως τεχνικό αποτέλεσμα της μετατροπής, αλλά και ως συγκροτημένη υλοποίηση της προτεινόμενης οντολογίας, στην οποία το επίπεδο του σχήματος και το επίπεδο των δεδομένων συνδυάζονται σε ενιαίο και συνεκτικό σημασιολογικό γράφο.

Αμέσως επόμενο και φυσικό βήμα μετά την οργάνωση και αποθήκευση του γράφου είναι ο έλεγχος της ορθότητας και της συνέπειάς του. Για τον λόγο αυτό, στο επόμενο υποκεφάλαιο εξετάζεται ο τρόπος με τον οποίο ελέγχθηκε η εγκυρότητα των δεδομένων και η συμβατότητά τους με το προτεινόμενο οντολογικό μοντέλο.

Τμήμα του αρχείου TTL	Περιεχόμενο	Σκοπός
Ontology header	Prefixes, ontology declaration, imports	Δήλωση ταυτότητας και εξαρτήσεων της οντολογίας
Schema layer	Κλάσεις, ιδιότητες, domains, ranges, alignments	Ορισμός του application profile
Audit-aligned custom layer	Μόνο οι αναγκαίοι custom όροι και shortcuts	Εφαρμοστική εξειδίκευση χωρίς περιττό πλεονασμό
Example A instances	Object-centric παραδείγματα τύπου AADA	Ανάδειξη της βασικής περιγραφικής χρήσης του μοντέλου
Example B instances	Event-based παραδείγματα τύπου OpenContext	Ανάδειξη excavation, documentation και digital layers
Enriched demonstrator	Super-example με πλουσιότερη διασύνδεση	Επίδειξη της εκφραστικής ισχύος της οντολογίας

Πίνακας 12: Εσωτερική οργάνωση του τελικού αρχείου TTL

5.4 Έλεγχος ορθότητας των δεδομένων

Μετά την οργάνωση και αποθήκευση του Γνωσιακού Γράφου, κρίθηκε αναγκαίο να πραγματοποιηθεί έλεγχος ορθότητας των παραγόμενων δεδομένων. Το στάδιο αυτό είναι ιδιαίτερα σημαντικό σε κάθε σημασιολογική υλοποίηση, επειδή η παραγωγή RDF τριάδων από μόνη της δεν εγγυάται ότι το τελικό αποτέλεσμα είναι και εννοιολογικά συνεπές, ούτε ότι ανταποκρίνεται σωστά στο οντολογικό σχήμα που έχει σχεδιαστεί. Επομένως, στην παρούσα εργασία ο έλεγχος δεν αντιμετωπίστηκε ως μια τυπική τεχνική διαδικασία, αλλά ως αναπόσπαστο μέρος της ίδιας της υλοποίησης.

Ο έλεγχος της ορθότητας πραγματοποιήθηκε σε τρία βασικά επίπεδα: στο συντακτικό επίπεδο του αρχείου RDF/Turtle, στο σημασιολογικό επίπεδο της αντιστοίχισης των δεδομένων με την οντολογία, και στο λειτουργικό επίπεδο της δυνατότητας του γράφου να απαντά τα competency questions μέσω ερωτημάτων SPARQL. Η πολυεπίπεδη αυτή προσέγγιση ήταν αναγκαία, επειδή ένα αρχείο μπορεί να είναι συντακτικά έγκυρο αλλά εννοιολογικά προβληματικό, ή αντιστρόφως να αποδίδει σωστές έννοιες αλλά να περιλαμβάνει τεχνικά σφάλματα που εμποδίζουν την επεξεργασία του.

Σε πρώτο επίπεδο, ελέγχθηκε η συντακτική εγκυρότητα του αρχείου Turtle. Ο έλεγχος αυτός περιλάμβανε την επιβεβαίωση ότι όλες οι τριάδες ακολουθούν σωστά τη σύνταξη της Turtle, ότι τα prefixes δηλώνονται ορθά, ότι οι URI αναφορές είναι έγκυρες, και ότι δεν υπάρχουν λάθη όπως ελλιπή statements, λανθασμένοι διαχωριστές ή μη αποδεκτή χρήση χαρακτήρων. Το στάδιο αυτό αποδείχθηκε κρίσιμο, καθώς ακόμη και μικρές αστοχίες μπορούν να καταστήσουν αδύνατη την εισαγωγή του γράφου σε ένα εργαλείο όπως το Protégé ή σε έναν RDF parser. Η τελική έκδοση του αρχείου ελέγχθηκε επαναληπτικά έως ότου μπορεί να διαβαστεί χωρίς σφάλματα από τα αντίστοιχα εργαλεία.

Σε δεύτερο επίπεδο, εξετάστηκε η εννοιολογική ορθότητα της μοντελοποίησης. Εδώ το ζητούμενο δεν ήταν απλώς αν τα δεδομένα «γράφονται σωστά» σε RDF, αλλά αν οι κλάσεις και οι ιδιότητες χρησιμοποιούνται με τρόπο συμβατό προς τη σημασία τους. Για παράδειγμα, ελέγχθηκε αν τα artefacts αποδίδονται ως instances κατάλληλων κλάσεων του CIDOC CRM, αν οι τύποι, τα υλικά, οι τόποι και οι χρονικές εντάξεις συνδέονται μέσω των κατάλληλων σχέσεων, και αν τα γεγονότα παραγωγής, ανασκαφής και τεκμηρίωσης αποδίδονται ως διακριτές σημασιολογικές οντότητες και όχι ως απλές περιγραφικές τιμές. Ο έλεγχος αυτός ήταν ιδιαίτερα σημαντικός κατά τις τελικές αναθεωρήσεις της οντολογίας, καθώς στόχος ήταν να περιοριστούν οι περιττές τοπικές κλάσεις και ιδιότητες και να αξιοποιηθούν, όπου ήταν δυνατό, υπάρχοντες όροι από το CIDOC CRM, το CRMarchaeo, το GeoSPARQL, το SKOS και το HDTO.

Στο ίδιο πλαίσιο εξετάστηκε και η συνέπεια ανάμεσα στο οντολογικό σχήμα και στα παραδείγματα δεδομένων. Με άλλα λόγια, ελέγχθηκε αν τα demonstrator instances του γράφου πράγματι ακολουθούν τις σχεδιαστικές επιλογές που περιγράφονται στο Κεφάλαιο 4. Αυτό περιλάμβανε, μεταξύ άλλων, τον έλεγχο ότι οι τοπικές shortcut ιδιότητες χρησιμοποιούνται μόνο εκεί όπου είχαν αιτιολογηθεί ως απαραίτητες, ότι οι εξωτερικές οντολογίες χρησιμοποιούνται άμεσα όταν δεν απαιτείται τοπική εξειδίκευση, και ότι οι αναπαραστάσεις των AADA και OpenContext παραδειγμάτων αντανakλούν σωστά τις δύο συμπληρωματικές όψεις του μοντέλου, δηλαδή την object-centric και τη γεγονοτοκεντρική.

Ένα τρίτο επίπεδο ελέγχου αφορούσε τη λογική πληρότητα των δεδομένων σε σχέση με τα competency questions. Στην πράξη, αυτό σημαίνει ότι ο γράφος ελέγχθηκε ως προς το κατά πόσο επιτρέπει την ανάκτηση της πληροφορίας που είχε τεθεί εξ αρχής ως απαίτηση του σχεδιασμού. Για παράδειγμα, επιβεβαιώθηκε ότι μπορούν να εντοπιστούν τα artefacts μαζί με τα identifiers και τα

labels τους, ότι είναι δυνατή η ανάκτηση τύπου και υλικού, ότι μπορούν να ανακτηθούν οι σχέσεις με περίοδο, τόπο και γεωμετρία, ότι εμφανίζονται ρητά οι excavation events και τα stratigraphic contexts, και ότι είναι δυνατή η σύνδεση των αντικειμένων με documentation events και digital representations. Με τον τρόπο αυτό, η αξιολόγηση της ορθότητας δεν έμεινε στο αφηρημένο επίπεδο της μοντελοποίησης, αλλά συνδέθηκε με τη λειτουργική χρησιμότητα του γράφου.

Ιδιαίτερη σημασία είχε επίσης ο έλεγχος των παραδειγμάτων που αφορούν την ψηφιακή αναπαράσταση. Επειδή το digital twin layer αποτελεί πιο περιορισμένο αλλά ταυτόχρονα πιο ευαίσθητο τμήμα του μοντέλου, εξετάστηκε εάν η χρήση των όρων του HDTO εντάσσεται με συνεπή τρόπο στον γράφο, αν οι σχέσεις μεταξύ φυσικού αντικειμένου, digital twin και documentation creation αποδίδονται σωστά, και αν αποφεύγεται η σύγχυση ανάμεσα στο ίδιο το ψηφιακό τεκμήριο και στο γεγονός της δημιουργίας του. Η διάκριση αυτή είναι εννοιολογικά σημαντική και επηρεάζει άμεσα τόσο την ποιότητα της μοντελοποίησης όσο και τη δυνατότητα υποστήριξης σύνθετων ερωτημάτων.

Πέρα από την εσωτερική λογική του γράφου, ελέγχθηκε και η ποιότητα των εξωτερικών συνδέσεων. Όπου χρησιμοποιούνται εξωτερικά URIs ή canonical references, όπως στην περίπτωση περιοδικών εντάξεων ή πηγών δεδομένων, εξετάστηκε αν οι αναφορές αυτές είναι συμβατές με το νόημα που επιδιώκεται να αποδοθεί. Ο στόχος εδώ δεν ήταν να οικοδομηθεί πλήρης linked open data ενοποίηση, αλλά να διασφαλιστεί ότι οι εξωτερικές παραπομπές που ενσωματώνονται στο μοντέλο είναι τουλάχιστον σημασιολογικά εύλογες και τεκμηριωμένες.

Έλεγχος	Τι ελέγχθηκε	Μέθοδος / εργαλείο	Αποτέλεσμα
Συντακτική εγκυρότητα Turtle	Ορθότητα της σύνταξης του αρχείου TTL	Protégé / RDF parser	Το αρχείο φορτώνεται χωρίς συντακτικά σφάλματα
Ορθότητα namespaces	Συνέπεια prefixes και URIs	Χειροκίνητος έλεγχος και δοκιμαστική φόρτωση	Τα namespaces χρησιμοποιούνται συνεκτικά
Συνέπεια οντολογικών όρων	Χρήση υπαρχουσών ιδιοτήτων και κλάσεων όπου ήταν διαθέσιμες	Εννοιολογικός έλεγχος βάσει CIDOC CRM, CRMarchaeo, HDTO κ.ά.	Περιορίστηκαν οι πλεονάζοντες custom όροι
Έλεγχος local shortcuts	Αναγκαιότητα και αιτιολόγηση των custom shortcut properties	Conceptual audit	Διατηρήθηκαν μόνο τα απολύτως αναγκαία shortcuts
Συνέπεια examples με την οντολογία	Αν τα RDF examples ακολουθούν το τελικό application profile	Χειροκίνητη αντιπαράβολή με το TTL schema	Τα examples προσαρμόστηκαν στο τελικό μοντέλο
Έλεγχος χωρικής αναπαράστασης	Ορθότητα χρήσης Place, Geometry και WKT	Έλεγχος GeoSPARQL patterns	Η χωρική πληροφορία αποδίδεται συνεκτικά
Έλεγχος temporal semantics	Σύνδεση periods και time-spans	Έλεγχος period modelling	Η χρονική αναπαράσταση υποστηρίζει period και event layers
Έλεγχος documentation layer	Σχέση record, documentation creation και provenance	Έλεγχος CIDOC CRM / PROV patterns	Το documentation layer αποδίδεται διακριτά
Έλεγχος digital twin layer	Σύνδεση φυσικού αντικειμένου και ψηφιακής αναπαράστασης	Έλεγχος HDTO/CIDOC patterns	Το digital layer ενσωματώνεται λειτουργικά στο μοντέλο
SPARQL readiness	Δυνατότητα υποστήριξης competency questions μέσω	Λογικός έλεγχος πάνω στα examples και στον	Το μοντέλο είναι κατάλληλο για SPARQL διερεύνηση

	queries	demonstrator graph	
--	---------	--------------------	--

Πίνακας 13: Έλεγχοι ορθότητας και αποτελέσματα

Από την όλη διαδικασία ελέγχου προέκυψε ότι η ορθότητα των δεδομένων δεν είναι αποτέλεσμα ενός μόνο βήματος, αλλά συνδυασμός επαναληπτικών διορθώσεων και συνεχούς σύγκρισης ανάμεσα στο θεωρητικό μοντέλο, στο πραγματικό σύνολο δεδομένων και στο τελικό RDF αποτέλεσμα. Στην πράξη, η διαδικασία αυτή λειτούργησε και ως μηχανισμός βελτίωσης της ίδιας της οντολογίας. Όπου εντοπίζονταν πλεονασμοί, ασάφειες ή αδύναμες σημασιολογικές επιλογές, το μοντέλο αναθεωρούνταν ώστε να γίνει πιο καθαρό και πιο συνεπές με τις αρχές επαναχρησιμοποίησης των υφιστάμενων οντολογιών.

Συνολικά, ο έλεγχος ορθότητας έδειξε ότι ο παραγόμενος Γνωσιακός Γράφος είναι όχι μόνο τεχνικά έγκυρος ως RDF/Turtle αρχείο, αλλά και επαρκώς συνεπής ως προς τη σημασιολογική του δομή. Επιπλέον, επιβεβαίωσε ότι ο γράφος μπορεί να λειτουργήσει ως σταθερή βάση για το επόμενο στάδιο της εργασίας, δηλαδή για την εξερεύνηση των δεδομένων μέσω ερωτημάτων και την παρουσίαση των αποτελεσμάτων. Με άλλα λόγια, το στάδιο αυτό έκλεισε τον κύκλο της υλοποίησης και προετοίμασε το έδαφος για τη μετάβαση από την κατασκευή του γράφου στη λειτουργική αξιοποίησή του.

Το επόμενο υποκεφάλαιο παρουσιάζει συνοπτικά τα βασικά εργαλεία που χρησιμοποιήθηκαν κατά την υλοποίηση της οντολογίας, τη δημιουργία του Γνωσιακού Γράφου και τον έλεγχο των δεδομένων.

5.5 Εργαλεία που χρησιμοποιήθηκαν

Για την υλοποίηση της παρούσας εργασίας αξιοποιήθηκε ένα σύνολο εργαλείων που υποστήριξαν διαφορετικά στάδια της διαδικασίας, από τον σχεδιασμό της οντολογίας και τη συγγραφή του αρχείου Turtle έως τον έλεγχο ορθότητας των δεδομένων και την προετοιμασία των ερωτημάτων. Η επιλογή των εργαλείων δεν έγινε με σκοπό την ανάπτυξη μιας σύνθετης αυτοματοποιημένης υποδομής, αλλά με γνώμονα την καταλληλότητά τους για ένα ελεγχόμενο proof of concept, στο οποίο προτεραιότητα είχε η σαφήνεια της μοντελοποίησης και η δυνατότητα επαλήθευσης του αποτελέσματος.

Το βασικό εργαλείο για τη συγγραφή, επιθεώρηση και αναθεώρηση της οντολογίας ήταν η επεξεργασία του αρχείου σε μορφή Turtle. Η χρήση της Turtle κρίθηκε ιδιαίτερα κατάλληλη, επειδή επιτρέπει ευανάγνωστη και σχετικά συμπαγή αναπαράσταση RDF τριάδων, ενώ παράλληλα διευκολύνει τον άμεσο έλεγχο των prefixes, των URI πόρων και της συνολικής δομής του γράφου. Σε σχέση με πιο εκτενείς σειριοποιήσεις RDF, η Turtle προσφέρει καλύτερη αναγνωσιμότητα για σκοπούς ανάπτυξης και τεκμηρίωσης, κάτι που ήταν ιδιαίτερα χρήσιμο στην παρούσα εργασία, όπου το αρχείο της οντολογίας και τα demonstrator examples αναθεωρήθηκαν επανειλημμένα.

Για την εννοιολογική επιθεώρηση της οντολογίας χρησιμοποιήθηκε το Protégé. Το εργαλείο αυτό αξιοποιήθηκε κυρίως για την επιβεβαίωση ότι η οντολογία μπορεί να φορτωθεί σωστά, ότι οι κλάσεις και οι ιδιότητες δημιουργούνται και οργανώνονται με συνεπή τρόπο, και ότι οι σχέσεις subclass και subproperty εμφανίζονται όπως αναμένεται. Η χρήση του Protégé δεν περιορίστηκε μόνο σε επίπεδο απλής φόρτωσης του αρχείου, αλλά λειτούργησε και ως μέσο εντοπισμού πιθανών προβλημάτων στη μοντελοποίηση, όπως ασυνέπειες σε declarations, λανθασμένες αναφορές σε namespaces ή περιττοί τοπικοί όροι που δεν εξυπηρετούσαν ουσιαστικό σκοπό. Με τον τρόπο αυτό, το Protégé συνέβαλε όχι μόνο στην αρχική δημιουργία και στον τεχνικό έλεγχο, αλλά και στη σταδιακή βελτίωση του ίδιου του οντολογικού σχήματος.

Παράλληλα, αξιοποιήθηκαν RDF parsing και validation μηχανισμοί για τον συντακτικό έλεγχο του Turtle αρχείου. Η ανάγκη για τέτοιο έλεγχο προέκυψε πρακτικά κατά τη διάρκεια της εργασίας, καθώς ακόμη και μικρά σφάλματα σύνταξης μπορούσαν να εμποδίσουν τη σωστή φόρτωση του αρχείου. Η χρήση parser επέτρεψε να επιβεβαιωθεί ότι η τελική έκδοση του γράφου είναι συντακτικά έγκυρη και μπορεί να επεξεργαστεί από συμβατά εργαλεία του RDF οικοσυστήματος. Το στάδιο αυτό ήταν κρίσιμο, επειδή διασφάλισε ότι η σημασιολογική εργασία που είχε προηγηθεί δεν θα υπονομευόταν από τεχνικά προβλήματα αναπαράστασης.

Για τη διατύπωση και δοκιμή των ερωτημάτων SPARQL χρησιμοποιήθηκαν περιβάλλοντα κατάλληλα για πειραματική ανάκτηση πληροφορίας πάνω στον γράφο. Στο πλαίσιο της παρούσας εργασίας, τα ερωτήματα δεν χρησιμοποιήθηκαν μόνο ως τελικό μέσο παρουσίασης αποτελεσμάτων, αλλά και ως εργαλείο επαλήθευσης της μοντελοποίησης. Με άλλα λόγια, η δυνατότητα επιτυχούς εκτέλεσης ενός ερωτήματος που ανταποκρίνεται σε συγκεκριμένο competency question λειτούργησε και ως έμμεσος έλεγχος της ορθότητας των RDF δεδομένων. Έτσι, τα εργαλεία SPARQL εντάχθηκαν ουσιαστικά στη διαδικασία ανάπτυξης και όχι μόνο στο τελικό στάδιο επίδειξης.

Σε επίπεδο πηγών και τεκμηρίωσης, σημαντικό ρόλο έπαιξαν και οι ίδιες οι επίσημες τεκμηριώσεις των οντολογιών και των συνόλων δεδομένων που χρησιμοποιήθηκαν. Η τεκμηρίωση του CIDOC CRM, του CRMarchaeo, του GeoSPARQL, του SKOS, του PROV-O, καθώς και οι διαθέσιμες πληροφορίες για το HDTO, λειτούργησαν ως βασικό σημείο αναφοράς για τη λήψη αποφάσεων σχετικά με την επαναχρησιμοποίηση ή μη υπαρχόντων όρων. Αντίστοιχα, τα διαθέσιμα δεδομένα και οι δομές των AADA και OpenContext χρησιμοποιήθηκαν όχι μόνο ως πηγές περιεχομένου, αλλά και ως οδηγοί για τον εντοπισμό των πραγματικών μοντελοποιητικών αναγκών της εργασίας. Με αυτή την έννοια, οι τεκμηριώσεις και οι ίδιες οι πηγές δεδομένων λειτούργησαν ως «εργαλεία γνώσης», τα οποία υποστήριζαν ουσιαστικά τις επιλογές σχεδιασμού.

Επιπλέον, χρησιμοποιήθηκαν γενικά εργαλεία επεξεργασίας κειμένου και δομημένης πληροφορίας για την οργάνωση των παραδειγμάτων, τη σύνταξη των πινάκων αντιστοίχισης και τη σταδιακή αναθεώρηση του υλικού της εργασίας. Παρότι τα εργαλεία αυτά δεν ανήκουν ειδικά στο οικοσύστημα του Σημασιολογικού Ιστού, ήταν απαραίτητα για την τεκμηρίωση των αποφάσεων, τη διατήρηση συνέπειας ανάμεσα στο κεφάλαιο σχεδιασμού και στο τελικό TTL αρχείο, καθώς και για την προετοιμασία του κειμένου της εργασίας σε επίσημη μορφή.

Εργαλείο	Ρόλος	Πού χρησιμοποιήθηκε	Παρατηρήσεις
Protégé	Επεξεργασία και έλεγχος της οντολογίας	Σχεδιασμός, φόρτωση και έλεγχος του TTL	Χρησιμοποιήθηκε για έλεγχο σύνταξης και δομής
Visual Studio Code	Επεξεργασία RDF/Turtle	Συγγραφή και αναθεώρηση ontology file και examples	Χρήσιμος για ακριβή επεξεργασία triples
Protégé parser	Συντακτικός έλεγχος RDF	Έλεγχος εγκυρότητας του αρχείου	Επιβεβαίωση σωστής σύνταξης Turtle
Protégé SPARQL Query tab	Διατύπωση και έλεγχος ερωτημάτων	Προετοιμασία query layer για τα competency questions	Υποστηρίζει τη διερεύνηση του γράφου
Word	Συγγραφή και μορφοποίηση της εργασίας	Σύνθεση των κεφαλαίων και πινάκων	Χρησιμοποιήθηκε για την τελική ακαδημαϊκή παρουσίαση

Πίνακας 14: Εργαλεία που χρησιμοποιήθηκαν στην υλοποίηση

Συνοψίζοντας, τα εργαλεία που χρησιμοποιήθηκαν στην παρούσα εργασία μπορούν να οργανωθούν σε τέσσερις βασικές κατηγορίες: εργαλεία συγγραφής και επεξεργασίας RDF/Turtle, εργαλεία εννοιολογικής επιθεώρησης της οντολογίας, εργαλεία ελέγχου συντακτικής ορθότητας και εργαλεία διατύπωσης και δοκιμής ερωτημάτων SPARQL. Ο συνδυασμός τους επέτρεψε να υποστηριχθεί με συνεπή τρόπο ολόκληρη η διαδικασία υλοποίησης, από την αρχική μοντελοποίηση μέχρι την τελική επιβεβαίωση ότι ο Γνωσιακός Γράφος είναι τεχνικά έγκυρος, σημασιολογικά συνεπής και λειτουργικά αξιοποιήσιμος.

Με την ολοκλήρωση και αυτού του υποκεφαλαίου κλείνει το Κεφάλαιο 5, το οποίο παρουσίασε το πρακτικό σκέλος της υλοποίησης του Γνωσιακού Γράφου. Στο επόμενο κεφάλαιο η έμφαση μετατοπίζεται από την κατασκευή του γράφου στην αξιοποίησή του, μέσω της διατύπωσης SPARQL ερωτημάτων και της παρουσίασης των αποτελεσμάτων που προκύπτουν από την εξερεύνηση των δεδομένων.

Κεφάλαιο 6ο: Εξερεύνηση Δεδομένων & Παρουσίαση Αποτελεσμάτων

Μετά τον σχεδιασμό της οντολογίας και την υλοποίηση του Γνωσιακού Γράφου, το επόμενο βήμα της εργασίας είναι η διερεύνηση του παραγόμενου γράφου και η αξιολόγηση της πρακτικής χρησιμότητάς του. Το παρόν κεφάλαιο επικεντρώνεται ακριβώς σε αυτή τη φάση, δηλαδή στον τρόπο με τον οποίο το οντολογικό μοντέλο και τα RDF δεδομένα μπορούν να αξιοποιηθούν για αναζήτηση, εξερεύνηση και εξαγωγή πληροφορίας.

Η σημασία του κεφαλαίου αυτού είναι ιδιαίτερα μεγάλη, διότι εδώ ελέγχεται στην πράξη κατά πόσο οι σχεδιαστικές επιλογές των προηγούμενων κεφαλαίων οδηγούν πράγματι σε ένα λειτουργικό και χρήσιμο σημασιολογικό αποτέλεσμα. Με άλλα λόγια, δεν αρκεί η οντολογία να είναι εννοιολογικά συνεπής ή τεχνικά ορθή θα πρέπει επίσης να μπορεί να υποστηρίξει ερωτήματα που έχουν ουσιαστικό πληροφοριακό ενδιαφέρον και να επιτρέπει την ανάκτηση γνώσης από τα δεδομένα με τρόπο σαφή και αξιοποιήσιμο.

Για τον σκοπό αυτό, η εξερεύνηση του γράφου πραγματοποιείται μέσω της γλώσσας SPARQL, η οποία αποτελεί τον βασικό μηχανισμό υποβολής ερωτημάτων σε RDF δεδομένα. Τα ερωτήματα που θα παρουσιαστούν στο κεφάλαιο αυτό δεν επιλέγονται αυθαίρετα, αλλά προκύπτουν άμεσα από τα competency questions που διαμορφώθηκαν στο Κεφάλαιο 4. Με τον τρόπο αυτό, εξασφαλίζεται ότι η πειραματική διερεύνηση παραμένει συνδεδεμένη με τις πραγματικές απαιτήσεις της μοντελοποίησης και ότι η αξιολόγηση του γράφου βασίζεται σε συγκεκριμένους και αιτιολογημένους στόχους.

Παράλληλα, το κεφάλαιο δεν περιορίζεται μόνο στην παράθεση τεχνικών ερωτημάτων. Αντίθετα, επιδιώκεται να παρουσιαστεί και να σχολιαστεί το είδος της πληροφορίας που ανακτάται, ώστε να αναδειχθεί η αξία του προτεινόμενου μοντέλου τόσο σε object-centric όσο και σε event-based περιπτώσεις. Ιδιαίτερη έμφαση δίνεται στη δυνατότητα του γράφου να συνδυάζει πληροφορία για ευρήματα, τύπους, υλικά, περιόδους, τόπους, ανασκαφικές δραστηριότητες, τεκμηριωτικές πράξεις και ψηφιακές αναπαραστάσεις μέσα σε ένα ενιαίο και συνεκτικό σημασιολογικό πλαίσιο.

Στις ενότητες που ακολουθούν παρουσιάζεται αρχικά η λογική με βάση την οποία διαμορφώθηκαν τα SPARQL ερωτήματα, στη συνέχεια δίνονται ενδεικτικά παραδείγματα ερωτημάτων, και τέλος σχολιάζονται τα αποτελέσματα που προέκυψαν από την εφαρμογή τους πάνω στον υλοποιημένο Γνωσιακό Γράφο.

6.1 Δημιουργία ερωτημάτων SPARQL

Η δημιουργία των ερωτημάτων SPARQL στην παρούσα εργασία βασίστηκε άμεσα στα competency questions που είχαν διατυπωθεί στο Κεφάλαιο 4. Η επιλογή αυτή δεν είναι τυχαία, αλλά ακολουθεί τη γενική λογική ότι ο Γνωσιακός Γράφος πρέπει να αξιολογείται με βάση το κατά πόσο μπορεί να απαντήσει συγκεκριμένες πληροφοριακές ανάγκες που είχαν τεθεί ήδη από το στάδιο του σχεδιασμού. Επομένως, τα SPARQL ερωτήματα λειτουργούν εδώ ως ο σύνδεσμος ανάμεσα στο θεωρητικό και εννοιολογικό επίπεδο της μοντελοποίησης και στην πρακτική αξιοποίηση του τελικού γράφου.

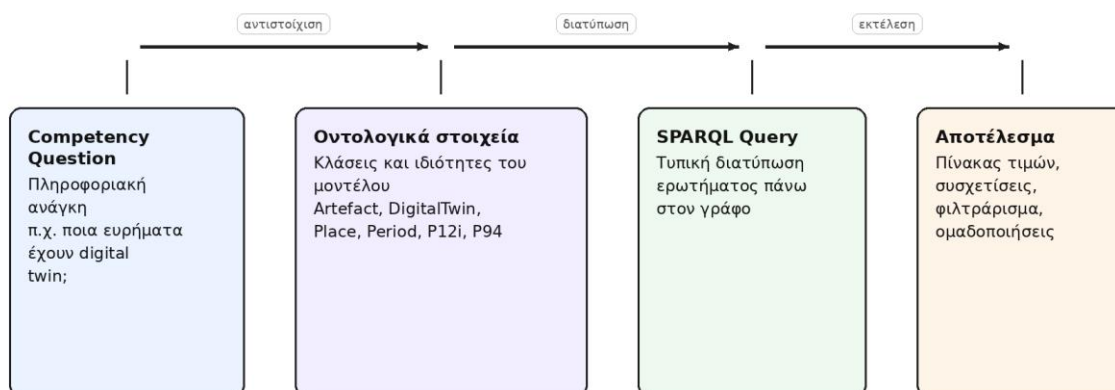
Η διαδικασία διαμόρφωσης των ερωτημάτων δεν ξεκίνησε από τη σύνταξη της SPARQL καθαυτής, αλλά από το σημασιολογικό περιεχόμενο που έπρεπε να ανακτηθεί. Για κάθε competency question εξετάστηκε πρώτα ποιες οντότητες, ποιες ιδιότητες και ποια μονοπάτια σχέσεων του γράφου εμπλέκονται στην αντίστοιχη πληροφοριακή ανάγκη. Στη συνέχεια, η ανάγκη αυτή μεταφράστηκε σε ερώτημα SPARQL με τρόπο που να είναι συμβατός με τη δομή της οντολογίας και με τα πρότυπα αναπαράστασης που υιοθετήθηκαν στο RDF αρχείο.

Η λογική αυτή ήταν ιδιαίτερα σημαντική, επειδή ο γράφος που κατασκευάστηκε στην παρούσα εργασία δεν περιλαμβάνει μόνο απλές σχέσεις τύπου «αντικείμενο, ιδιότητα, τιμή», αλλά και πιο σύνθετες διαδρομές ανάμεσα σε artefacts, periods, places, excavation events, stratigraphic contexts, actors, source records και digital twins. Επομένως, η διατύπωση των ερωτημάτων έπρεπε να λαμβάνει υπόψη όχι μόνο τις άμεσες τριάδες, αλλά και τα εννοιολογικά patterns που οργανώνουν το μοντέλο.

Στο πλαίσιο αυτό, τα ερωτήματα που διαμορφώθηκαν μπορούν να χωριστούν σε τέσσερις βασικές κατηγορίες. Η πρώτη κατηγορία αφορά ερωτήματα βασικής αναγνώρισης και περιγραφής των αντικειμένων, όπως η ανάκτηση identifiers, labels, types, materials και periods. Η δεύτερη κατηγορία αφορά ερωτήματα χωρικής και χρονικής διερεύνησης, στα οποία εξετάζεται η σύνδεση των ευρημάτων με τόπους, γεωμετρίες και χρονικές περιόδους. Η τρίτη κατηγορία αφορά πιο σύνθετα γεγονотоκεντρικά ερωτήματα, όπου διερευνώνται ανασκαφικές δραστηριότητες, stratigraphic contexts, actors και documentation creation events. Τέλος, η τέταρτη κατηγορία αφορά ερωτήματα που αξιοποιούν το digital twin layer και αναδεικνύουν τη δυνατότητα σύνδεσης του φυσικού αντικειμένου με την ψηφιακή του αναπαράσταση.

Ένα ακόμη σημαντικό σημείο είναι ότι τα ερωτήματα διαμορφώθηκαν με γνώμονα όχι μόνο την τυπική ορθότητά τους, αλλά και την αναγνωσιμότητα και ερμηνευσιμότητά τους. Για τον λόγο αυτό, όπου ήταν δυνατό, προτιμήθηκαν καθαρές και σχετικά ευανάγνωστες δομές query, ώστε να καθίσταται εμφανές ποιο μέρος της οντολογίας ενεργοποιείται κάθε φορά. Με τον τρόπο αυτό, η παρουσίαση των ερωτημάτων δεν λειτουργεί μόνο ως τεχνικό παράρτημα της υλοποίησης, αλλά και ως μέσο ερμηνείας του ίδιου του μοντέλου.

Συνεπώς, η SPARQL δεν χρησιμοποιείται στην παρούσα εργασία απλώς ως γλώσσα αναζήτησης, αλλά ως εργαλείο αξιολόγησης της οντολογικής επάρκειας του προτεινόμενου Γνωσιακού Γράφου. Αν ένα competency question μπορεί να μεταφραστεί με σαφήνεια σε ερώτημα και να επιστρέψει συνεκτικά αποτελέσματα, τότε αυτό αποτελεί ισχυρή ένδειξη ότι το μοντέλο υποστηρίζει επιτυχώς την αντίστοιχη πληροφοριακή απαίτηση. Για να αποτυπωθεί πιο καθαρά η λογική διατύπωσης των ερωτημάτων, η διαδικασία μετάβασης από τα competency questions στα SPARQL ερωτήματα και στα αποτελέσματα συνοψίζεται στο ακόλουθο διάγραμμα.



Σχήμα 7: Από τα CQs στα SPARQL ερωτήματα και στα ανακτώμενα αποτελέσματα

Στην επόμενη ενότητα παρουσιάζονται ενδεικτικά ερωτήματα SPARQL, τα οποία αντιστοιχούν στα βασικά competency questions της εργασίας και δείχνουν με πρακτικό τρόπο πώς μπορεί να διερευνηθεί ο παραγόμενος γράφος.

6.2 Παραδείγματα ερωτημάτων

Στην παρούσα ενότητα παρουσιάζονται ενδεικτικά ερωτήματα SPARQL, τα οποία αντιστοιχούν στα βασικά competency questions που είχαν διατυπωθεί στο Κεφάλαιο 4. Στόχος της ενότητας δεν είναι να παρουσιαστεί μια εξαντλητική συλλογή όλων των πιθανών ερωτημάτων πάνω στον Γνωσιακό Γράφο, αλλά να δοθεί ένα αντιπροσωπευτικό σύνολο ερωτημάτων που να δείχνει στην πράξη πώς το προτεινόμενο οντολογικό μοντέλο μπορεί να υποστηρίξει την αναζήτηση και την εξερεύνηση της πληροφορίας.

Ο σχεδιασμός των ερωτημάτων SPARQL στηρίχθηκε άμεσα στα competency questions που είχαν ήδη διαμορφωθεί σε προηγούμενο κεφάλαιο. Για τον λόγο αυτό, πριν από την αναλυτική παρουσίαση των ερωτημάτων, είναι χρήσιμο να αποτυπωθεί συνοπτικά η αντιστοίχιση ανάμεσα σε κάθε competency question, στον στόχο του ερωτήματος και στα βασικά στοιχεία του οντολογικού μοντέλου που ενεργοποιούνται.

CQ #	Θεματική	Σκοπός ερωτήματος	Κύρια στοιχεία του μοντέλου
CQ1	Βασική αναπαράσταση artefact	Ανάκτηση ευρημάτων, identifiers και labels	gerasim:Artefact, crm:P1_is_identified_by, rdfs:label
CQ2	Τυπολογία και υλικό	Ανάκτηση τύπου και υλικού	crm:P2_has_type, crm:P45_consists_of
CQ3	Χρονική ένταξη	Σύνδεση artefact ή event με period	gerasim:associatedWithPeriod, skos:prefLabel
CQ4	Χωρική αναπαράσταση	Σύνδεση με place και geometry	crm:P53_has_former_or_current_location, geo:hasGeometry, geo:asWKT
CQ5	Excavation event	Εύρεση artefacts ανά excavation	crm:P12i_was_present_at

		event	
CQ6	Stratigraphic context	Εύρεση artefacts ανά stratigraphic context	gerasim:hasStratigraphicContext
CQ7	Production layer	Ανάδειξη production event	crm:P108i_was_produced_by
CQ8	Documentation layer	Σύνδεση ευρήματος με documentation και provenance	crm:P70_is_documented_in, crm:P94_has_created
CQ9	Same archaeologist across excavations	Συσχέτιση διαφορετικών excavations μέσω κοινού actor	crm:P14_carried_out_by
CQ10	Same period across places	Σύγκριση ίδιας περιόδου σε διαφορετικούς τόπους	gerasim:associatedWithPeriod, crm:P53_has_former_or_current_location
CQ11	Digital twin layer	Εύρεση ευρημάτων με ψηφιακή αναπαράσταση	hdto:HP1_has_digital_twin, hdto:HC2_Heritage_Digital_Twi
CQ12	Digital twin provenance	Σύνδεση digital twin με excavation και place	digital twin, documentation, excavation chain
CQ13	Temporal distribution by place	Κατανομή ευρημάτων ανά τόπο και περίοδο	place + period aggregation
CQ14	Multi-source graph exploitation	Συνδυαστική ανάκτηση object-centric και event-based πληροφορίας	κοινός οντολογικός πυρήνας

Πίνακας 15: Αντιστοίχιση competency questions με τα βασικά στοιχεία του μοντέλου

Ο παραπάνω πίνακας δείχνει ότι τα ερωτήματα δεν προέκυψαν αποσπασματικά, αλλά ακολουθούν τη λογική του μοντέλου και των πληροφοριακών αναγκών της εργασίας. Με βάση την αντιστοίχιση αυτή, τα ερωτήματα μπορούν να οργανωθούν και σε ευρύτερες λειτουργικές κατηγορίες, ανάλογα με τον τύπο της σημασιολογικής πληροφορίας που ανακτούν.

Τα ερωτήματα που ακολουθούν καλύπτουν τόσο την object-centric πλευρά του γράφου, όπως η αναζήτηση artefacts, τύπων, υλικών, περιόδων και τόπων, όσο και τη γεγονотоκεντρική του διάσταση, δηλαδή τη σύνδεση των αντικειμένων με ανασκαφικές δραστηριότητες, stratigraphic contexts, actors, τεκμηριωτικά γεγονότα και ψηφιακές αναπαραστάσεις. Με τον τρόπο αυτό, η SPARQL αξιοποιείται όχι μόνο ως τεχνικό εργαλείο ανάκτησης δεδομένων, αλλά και ως μέσο ανάδειξης της εκφραστικής δύναμης του ίδιου του μοντέλου.

Πέρα από τη σύνδεση κάθε ερωτήματος με το αντίστοιχο competency question, είναι επίσης χρήσιμο να γίνει μια συνοπτική κατηγοριοποίηση των ερωτημάτων SPARQL με βάση τη θεματική τους εστίαση. Με τον τρόπο αυτό καθίσταται σαφέστερο ότι η εξερεύνηση του γράφου καλύπτει όχι μόνο τη βασική περιγραφή αντικειμένων, αλλά και χωρικές, χρονικές, γεγονотоκεντρικές, τεκμηριωτικές και ψηφιακές διαστάσεις της πληροφορίας.

Κατηγορία ερωτημάτων	Ενδεικτικά CQ	Εστίαση
Βασικά περιγραφικά ερωτήματα	CQ1, CQ2, CQ3	artefacts, identifiers, types, materials, periods
Χωρικά και χρονικά ερωτήματα	CQ4, CQ10, CQ13	places, geometries, periods, distribution
Γεγονотоκεντρικά ερωτήματα	CQ5, CQ6, CQ7, CQ9	excavation events, stratigraphic contexts, production, actors
Τεκμηρίωση και provenance	CQ8, CQ12	source records, documentation creation, provenance chains
Ψηφιακή αναπαράσταση	CQ11, CQ12	digital twins, representation links, digital layer

Συνδυαστικά / πολυεπίπεδα	CQ14	object-centric και event-based exploitation
---------------------------	------	---

Πίνακας 16: Κατηγοριοποίηση των ερωτημάτων SPARQL της εργασίας

Η κατηγοριοποίηση αυτή επιβεβαιώνει τον πολυεπίπεδο χαρακτήρα της αναζήτησης πάνω στον Γνωσιακό Γράφο. Στη συνέχεια παρουσιάζονται ενδεικτικά ερωτήματα SPARQL, τα οποία δείχνουν στην πράξη πώς οι παραπάνω κατηγορίες μεταφράζονται σε συγκεκριμένους μηχανισμούς ανάκτησης πληροφορίας.

Για λόγους αναγνωσιμότητας, τα ερωτήματα παρουσιάζονται με κοινά prefixes και με τρόπο ώστε να αντιστοιχούν όσο γίνεται πιο άμεσα στα competency questions της εργασίας.

Κοινά prefixes

PREFIX gerasim: <<https://example.org/gerasim/ontology/>>
 PREFIX gres: <<https://example.org/gerasim/resource/>>
 PREFIX crm: <<http://www.cidoc-crm.org/cidoc-crm/>>
 PREFIX crma: <<http://www.cidoc-crm.org/extensions/crmarchaeo/>>
 PREFIX hdto: <<https://isl.ics.forth.gr/ontology/echoes/>>
 PREFIX skos: <<http://www.w3.org/2004/02/skos/core#>>
 PREFIX geo: <<http://www.opengis.net/ont/geosparql#>>
 PREFIX prov: <<http://www.w3.org/ns/prov#>>
 PREFIX dcterms: <<http://purl.org/dc/terms/>>
 PREFIX rdfs: <<http://www.w3.org/2000/01/rdf-schema#>>

[Ερώτημα 1: Βασική Αναπαράσταση artefact \(CQ1\)](#)

Το πρώτο ερώτημα ανακτά τα βασικά artefacts του γράφου μαζί με το αναγνωριστικό και το label τους.

[Ερώτημα 2: Τυπολογία και υλικό \(CQ2\)](#)

Το ερώτημα αυτό εξετάζει σε ποιον τύπο ανήκει κάθε εύρημα και από ποιο υλικό αποτελείται.

[Ερώτημα 3: Χρονική ένταξη \(CQ3\)](#)

Εδώ ανακτάται η σύνδεση ενός ευρήματος με περίοδο ή υποπερίοδο.

[Ερώτημα 4: Χωρική αναπαράσταση \(CQ4\)](#)

Το ερώτημα αυτό ανακτά τη σύνδεση ευρημάτων ή ανασκαφών με τόπους και γεωμετρίες.

[Ερώτημα 5: Excavation event \(CQ5\)](#)

Το ερώτημα αυτό εξετάζει ποια ευρήματα συνδέονται με συγκεκριμένη ανασκαφική δραστηριότητα.

[Ερώτημα 6: Stratigraphic context \(CQ6\)](#)

Το ερώτημα αυτό αναζητά τα ευρήματα που ανήκουν σε συγκεκριμένο στρωματογραφικό context.

[Ερώτημα 7: Production layer \(CQ7\)](#)

Το ερώτημα αυτό δείχνει πώς αναπαρίσταται η παραγωγή ή κατασκευή ενός ευρήματος.

[Ερώτημα 8: Documentation layer \(CQ8\)](#)

Το ερώτημα αυτό ανακτά τον τρόπο τεκμηρίωσης ενός ευρήματος και τη σχετική πράξη δημιουργίας.

[Ερώτημα 9: Same archaeologist across excavations \(CQ9\)](#)

Εδώ εξετάζεται ποια ευρήματα συνδέονται με τον ίδιο actor σε διαφορετικές ανασκαφές.

[Ερώτημα 10: Same period across different places \(CQ10\)](#)

Το ερώτημα αυτό αναζητά αντικείμενα που χρονολογούνται στην ίδια περίοδο αλλά προέρχονται από διαφορετικούς τόπους.

[Ερώτημα 11: Digital twin layer \(CQ11\)](#)

Εδώ εξετάζεται ποια ευρήματα διαθέτουν ψηφιακή αναπαράσταση.

[Ερώτημα 12: Digital twin provenance \(CQ12\)](#)

Το ερώτημα αυτό συνδυάζει digital twin, ανασκαφή και τόπο.

[Ερώτημα 13 :Temporal distribution by place \(CQ13\)](#)

Το ερώτημα αυτό είναι πιο αναλυτικό και δείχνει τη χρονολογική κατανομή ευρημάτων ανά τόπο.

[Ερώτημα 14: Multi-source graph exploitation \(CQ14\)](#)

Το τελευταίο ερώτημα δείχνει πώς συνδυάζονται object-centric και event-based δεδομένα στο ίδιο σχήμα.

Το ερώτημα αυτό είναι ιδιαίτερα ενδεικτικό, διότι μέσα από ένα ενιαίο query καθίσταται δυνατή η ταυτόχρονη πρόσβαση τόσο στον object-centric πυρήνα της αναπαράστασης όσο και στα event-based και digital layers του μοντέλου.

Συνολικά, τα παραπάνω ερωτήματα δείχνουν ότι ο Γνωσιακός Γράφος που υλοποιήθηκε στην παρούσα εργασία μπορεί να υποστηρίξει ένα ευρύ φάσμα πληροφοριακών αναγκών. Από την απλή ανάκτηση identifiers, τύπων και υλικών μέχρι τη συνδυαστική διερεύνηση περιόδων, τόπων, ανασκαφικών δραστηριοτήτων, τεκμηριωτικών γεγονότων και ψηφιακών αναπαραστάσεων, η SPARQL καθιστά εφικτή την αξιοποίηση του σημασιολογικού πλούτου του προτεινόμενου μοντέλου.

Στην επόμενη ενότητα παρουσιάζονται και σχολιάζονται τα αποτελέσματα που επιστρέφουν τα παραπάνω ερωτήματα, ώστε να αξιολογηθεί πιο συστηματικά η χρησιμότητα και η επάρκεια του προτεινόμενου Γνωσιακού Γράφου.

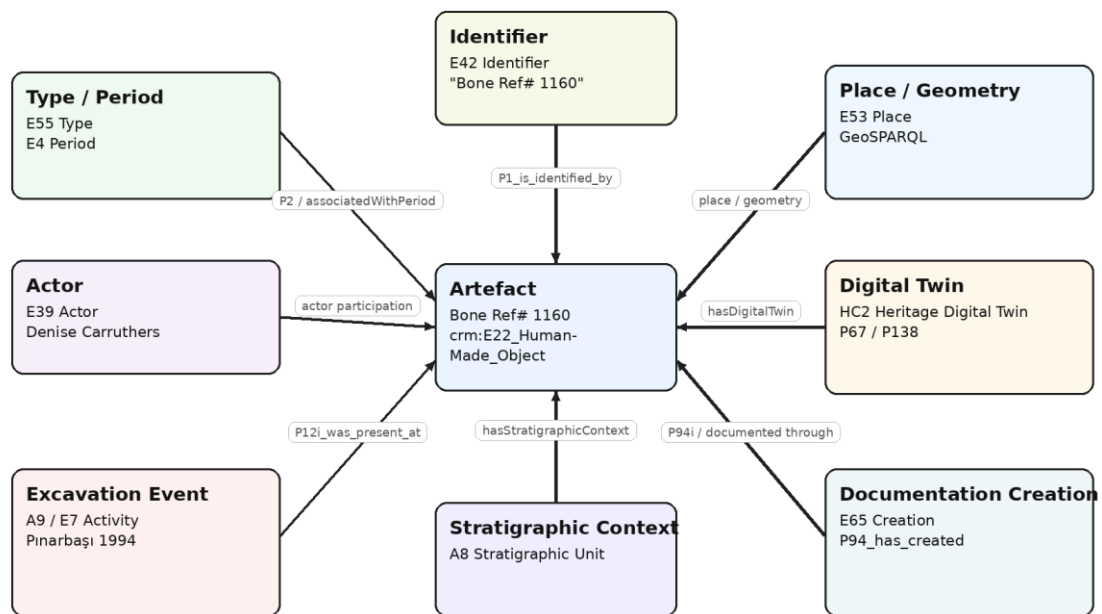
6.3 Παρουσίαση και σχολιασμός αποτελεσμάτων

Η εκτέλεση των ερωτημάτων SPARQL πάνω στον παραγόμενο Γνωσιακό Γράφο επέτρεψε την αξιολόγηση της λειτουργικότητας του μοντέλου στην πράξη. Τα αποτελέσματα που προέκυψαν δείχνουν ότι η οντολογία δεν παραμένει σε επίπεδο αφηρημένου εννοιολογικού σχεδιασμού, αλλά μπορεί να υποστηρίξει ουσιαστική ανάκτηση και συνδυαστική εξερεύνηση της πληροφορίας. Παράλληλα, τα αποτελέσματα καθιστούν εμφανές ότι ο γράφος επιτρέπει τη συνύπαρξη object-centric και event-based πληροφορίας μέσα σε ένα ενιαίο σημασιολογικό πλαίσιο. Τα queries επέστρεψαν αποτελέσματα που επιβεβαιώνουν τη λειτουργικότητα του μοντέλου. Πριν από τον αναλυτικό σχολιασμό των αποτελεσμάτων, είναι χρήσιμο να παρουσιαστεί ένας μικρός ενδεικτικός πίνακας, ώστε να γίνει πιο απτή η μορφή της πληροφορίας που μπορεί να ανακτηθεί από τον γράφο.

Artefact	Identifier	Type	Material	Period	Place	Digital Twin
Bone Ref# 1160	Bone Ref# 1160	Bone object	—	Pottery Neolithic	Pinarbasi Site B	Digital twin of Bone Ref# 1160
SA-stone- 0001	SA-stone-0001	Stone tool	Stone	Stone Age	Example AADA site	—

Πίνακας 17: Ενδεικτικά αποτελέσματα ερωτημάτων SPARQL

Τα ενδεικτικά αυτά αποτελέσματα δείχνουν ότι ο γράφος μπορεί να επιστρέφει συνδυαστικά πληροφορία από περισσότερα του ενός επίπεδα του μοντέλου. Με βάση την εικόνα αυτή, μπορεί πλέον να ακολουθήσει πιο συστηματικός σχολιασμός των αποτελεσμάτων ανά κατηγορία ερωτήματος. Η πολυεπίπεδη λογική του μοντέλου γίνεται ιδιαίτερα σαφής όταν εξεταστεί ένα μεμονωμένο artefact και οι σχέσεις του με τα υπόλοιπα στοιχεία του γράφου, όπως φαίνεται στο ακόλουθο σχήμα.



Σχήμα 8: Παράδειγμα πολυεπίπεδης σημασιολογικής εξερεύνησης του artefact "Bone Ref# 1160"

Στο επίπεδο της βασικής αναπαράστασης των αντικειμένων, τα πρώτα ερωτήματα έδειξαν ότι είναι δυνατή η άμεση ανάκτηση των artefacts μαζί με τα βασικά identifiers και labels τους. Το αποτέλεσμα αυτό είναι σημαντικό, επειδή επιβεβαιώνει ότι ο πυρήνας του μοντέλου λειτουργεί σωστά και ότι κάθε εύρημα αναπαρίσταται ως διακριτός πόρος με σαφή μηχανισμό ταυτοποίησης. Η πληροφορία αυτή αποτελεί τη βάση για κάθε επόμενο, πιο σύνθετο ερώτημα, καθώς χωρίς καθαρή αναγνώριση των βασικών οντοτήτων δεν θα ήταν δυνατή η συνεκτική εξερεύνηση του γράφου.

Εξίσου λειτουργικά αποδείχθηκαν και τα ερωτήματα που αφορούν την τυπολογία και τα υλικά των ευρημάτων. Μέσα από τις σχέσεις με τύπους και υλικά κατέστη δυνατό να ανακτηθούν ομάδες αντικειμένων που μοιράζονται κοινά χαρακτηριστικά, κάτι που αναδεικνύει την αξία της μοντελοποίησης με χρήση εννοιών όπως το crm:E55_Type και το crm:E57_Material. Τα

αποτελέσματα αυτής της κατηγορίας δείχνουν ότι ο γράφος δεν αποθηκεύει απλώς περιγραφικές τιμές, αλλά επιτρέπει ουσιαστική κατηγοριοποίηση και ομαδοποίηση της πληροφορίας.

Στο επίπεδο της χρονικής ένταξης, τα ερωτήματα που βασίστηκαν στη σύνδεση με περιόδους έδειξαν ότι το μοντέλο μπορεί να αποδώσει με συνεκτικό τρόπο τη χρονική ή πολιτισμική τοποθέτηση ενός ευρήματος. Η χρήση της σύνδεσης με περιόδους, σε συνδυασμό με SKOS labels και όπου χρειάζεται με broader period relations [14], διευκολύνει όχι μόνο την απλή ανάκτηση μιας χρονολογικής ένταξης, αλλά και τη μετάβαση σε ευρύτερες σημασιολογικές ομαδοποιήσεις. Αυτό είναι ιδιαίτερα σημαντικό για το αρχαιολογικό πεδίο, όπου τα ερωτήματα συχνά διατυπώνονται όχι μόνο για μεμονωμένα αντικείμενα, αλλά για αντικείμενα που σχετίζονται με συγκεκριμένες περιόδους ή υποπεριόδους.

Αντίστοιχα, τα ερωτήματα χωρικής αναπαράστασης έδειξαν ότι ο γράφος υποστηρίζει αποτελεσματικά τη σύνδεση αντικειμένων και γεγονότων με τόπους και γεωμετρίες. Η παρουσία τόπων ως διακριτών πόρων και η σύνδεσή τους με GeoSPARQL γεωμετρίες [17] καθιστά δυνατή την ανάκτηση όχι μόνο λεκτικών ονομάτων τοποθεσιών, αλλά και χωρικών αναπαραστάσεων σε μορφή WKT. Το αποτέλεσμα αυτό ενισχύει σημαντικά τη χρηστικότητα του μοντέλου, διότι δείχνει ότι ο γράφος μπορεί να αποτελέσει βάση όχι μόνο για σημασιολογική αναζήτηση, αλλά και για περαιτέρω χωρική ανάλυση ή οπτικοποίηση.

Ιδιαίτερο ενδιαφέρον παρουσιάζουν τα αποτελέσματα των γεγονотоκεντρικών ερωτημάτων. Μέσω της σύνδεσης των ευρημάτων με excavation events, stratigraphic contexts και actors, κατέστη δυνατό να αναδειχθεί ένα επίπεδο πληροφορίας που υπερβαίνει κατά πολύ μια απλή object-based περιγραφή. Τα ερωτήματα αυτά δείχνουν ότι το μοντέλο επιτρέπει να αντιμετωπιστεί το εύρημα όχι μόνο ως μεμονωμένο αντικείμενο, αλλά και ως οντότητα που εντάσσεται σε ένα ευρύτερο πλέγμα γεγονότων και σχέσεων. Το σημείο αυτό είναι ιδιαίτερα σημαντικό, επειδή επιβεβαιώνει στην πράξη τη γεγονотоκεντρική λογική που είχε τεθεί εξ αρχής ως βασικός στόχος της παρούσας εργασίας.

Πιο συγκεκριμένα, τα αποτελέσματα των ερωτημάτων για excavation events έδειξαν ότι είναι εφικτή η ανάκτηση των ευρημάτων που συνδέονται με συγκεκριμένες ανασκαφικές δραστηριότητες. Αυτό σημαίνει ότι ο γράφος μπορεί να απαντήσει όχι μόνο στο «τι είναι ένα εύρημα», αλλά και στο «μέσω ποιας αρχαιολογικής δραστηριότητας συνδέεται με το τεκμηριωτικό του πλαίσιο». Με παρόμοιο τρόπο, τα ερωτήματα για stratigraphic context κατέδειξαν ότι το μοντέλο υποστηρίζει και πιο εξειδικευμένες αρχαιολογικές σχέσεις, επιτρέποντας την ανάκτηση ευρημάτων που σχετίζονται με συγκεκριμένη στρωματογραφική μονάδα ή context.

Αντίστοιχα, τα ερωτήματα που περιλάμβαναν actors ανέδειξαν τη δυνατότητα του γράφου να συνδέει πολλαπλά γεγονότα μέσω κοινών προσώπων ή οργανισμών. Η δυνατότητα αυτή έχει ιδιαίτερη σημασία, διότι μετατρέπει τον γράφο από απλό αποθετήριο περιγραφών σε δομή μέσα στην οποία μπορούν να εντοπιστούν διασυνδέσεις ανάμεσα σε διαφορετικές πράξεις, ανασκαφές ή τεκμηριωτικές διαδικασίες. Σε μεγαλύτερης κλίμακας δεδομένα, η λογική αυτή θα μπορούσε να χρησιμοποιηθεί για πολύ πιο σύνθετες αναλύσεις δικτύου γύρω από πρόσωπα, ομάδες εργασίας ή διαχρονικές πρακτικές τεκμηρίωσης.

Σημαντικό τμήμα των αποτελεσμάτων αφορά επίσης το documentation layer. Τα σχετικά ερωτήματα έδειξαν ότι ο γράφος μπορεί να συνδέει ένα εύρημα όχι μόνο με source records, αλλά και με την ίδια την πράξη δημιουργίας τεκμηρίωσης. Η διάκριση αυτή έχει ουσιαστική σημασία, διότι επιτρέπει να διαχωρίζεται η ύπαρξη ενός τεκμηρίου από το γεγονός της παραγωγής του. Με άλλα λόγια, ο γράφος μπορεί να αναπαραστήσει όχι μόνο ότι ένα αντικείμενο τεκμηριώνεται, αλλά και ποιος το τεκμηρίωσε, πότε, και στο πλαίσιο ποιας τεκμηριωτικής πράξης. Το στοιχείο αυτό προσδίδει

μεγαλύτερο βάθος στη μοντελοποίηση provenance και ενισχύει τη σημασιολογική ακρίβεια του μοντέλου.

Ακόμη πιο χαρακτηριστικά είναι τα αποτελέσματα του digital twin layer. Τα σχετικά ερωτήματα κατέδειξαν ότι είναι εφικτή η ανάκτηση των ευρημάτων που διαθέτουν ψηφιακή αναπαράσταση, καθώς και η διασύνδεση της ψηφιακής αυτής όψης με το φυσικό αντικείμενο, το τεκμηριωτικό γεγονός και, έμμεσα, με το αρχαιολογικό του πλαίσιο. Η δυνατότητα αυτή είναι ιδιαίτερα σημαντική για την εργασία, διότι δείχνει ότι το μοντέλο δεν περιορίζεται σε παραδοσιακή τεκμηρίωση ευρημάτων, αλλά μπορεί να ενσωματώσει και ένα πρόσθετο ψηφιακό επίπεδο αναπαράστασης. Παρότι το επίπεδο αυτό χρησιμοποιείται εδώ με περιορισμένο και ελεγχόμενο τρόπο, αποτελεί ένδειξη ότι η οντολογία μπορεί να επεκταθεί προς περισσότερες πλούσιες μορφές πολιτιστικής ψηφιακής τεκμηρίωσης.

Εξίσου σημαντικά είναι και τα αποτελέσματα των συνδυαστικών ερωτημάτων, στα οποία αξιοποιούνται ταυτόχρονα πολλαπλά επίπεδα του μοντέλου. Ερωτήματα που συνδυάζουν περίοδο, τόπο, ανασκαφή, actor και digital twin έδειξαν ότι ο γράφος μπορεί να απαντήσει σε σύνθετες πληροφοριακές ανάγκες που δεν θα ήταν εύκολο να υποστηριχθούν σε μια συμβατική tabular αναπαράσταση. Η δυνατότητα αυτή αναδεικνύει την κύρια προστιθέμενη αξία του Γνωσιακού Γράφου: όχι την απλή αποθήκευση δεδομένων, αλλά την παραγωγή ενός διασυνδεδεμένου πλαισίου μέσα στο οποίο η πληροφορία μπορεί να επανασυνδυάζεται δυναμικά μέσω ερωτημάτων.

Ωστόσο, τα αποτελέσματα αυτά θα πρέπει να ερμηνευθούν και μέσα στο πραγματικό εύρος της παρούσας εργασίας. Ο γράφος που υλοποιήθηκε αποτελεί παράδειγμα μικρής κλίμακας και όχι πλήρη επιχειρησιακή υλοποίηση μεγάλου αρχαιολογικού αποθετηρίου. Επομένως, τα αποτελέσματα έχουν κυρίως επιδεικτικό και μεθοδολογικό χαρακτήρα: δείχνουν δηλαδή ότι το προτεινόμενο μοντέλο είναι σε θέση να υποστηρίξει τις ζητούμενες πληροφοριακές ανάγκες, όχι ότι έχει εξαντλήσει όλες τις πιθανές διαστάσεις του πεδίου. Παρ' όλα αυτά, ακόμη και σε αυτή τη μικρή κλίμακα, καθίσταται εμφανές ότι οι βασικές competency questions μπορούν να μεταφραστούν σε λειτουργικά ερωτήματα και να επιστρέψουν συνεκτικά αποτελέσματα.

Ο αναλυτικός σχολιασμός των αποτελεσμάτων δείχνει ότι κάθε κατηγορία ερωτημάτων φωτίζει διαφορετική όψη του μοντέλου. Για λόγους σύνοψης, ο παρακάτω πίνακας συγκεντρώνει το τι ανακτά κάθε βασική κατηγορία ερωτήματος και τι φανερώνει αυτό για τη λειτουργία και την εκφραστική ικανότητα του προτεινόμενου Γνωσιακού Γράφου.

Ερώτημα / κατηγορία	Τι ανακτά	Τι δείχνει για το μοντέλο
CQ1: Basic artefact retrieval	Artefacts, identifiers, labels	Ο γράφος υποστηρίζει σαφή ταυτοποίηση βασικών οντοτήτων
CQ2: Type and material	Τύπο και υλικό ευρημάτων	Το μοντέλο επιτρέπει κατηγοριοποίηση και ομαδοποίηση
CQ4: Spatial representation	Places και geometries	Υποστηρίζεται χωρική διάσταση μέσω GeoSPARQL
CQ5/CQ6: Event and context	Excavation events και stratigraphic contexts	Η αναπαράσταση υπερβαίνει την απλή object-centric λογική
CQ8: Documentation layer	Source records και documentation events	Το μοντέλο αποδίδει provenance και τεκμηριωτικές πράξεις
CQ11/CQ12: Digital twin layer	Ψηφιακές αναπαραστάσεις και τις σχέσεις τους	Ενσωματώνεται διακριτό digital layer

CQ14: Multi-source exploitation	Συνδυαστική πληροφορία από AADA και OpenContext	Επιβεβαιώνεται η συνύπαρξη object-centric και event-based πληροφορίας
---------------------------------	---	---

Πίνακας 18: Σχολιασμός αποτελεσμάτων ανά κατηγορία ερωτήματος

Ο παραπάνω πίνακας επιβεβαιώνει ότι το μοντέλο υποστηρίζει αποτελεσματικά τόσο τη βασική περιγραφική ανάκτηση όσο και πιο σύνθετες μορφές σημασιολογικής διερεύνησης. Η παρατήρηση αυτή οδηγεί φυσικά στην επόμενη ενότητα, όπου εξετάζεται και η πρωτοτυπική παρουσίαση του ίδιου του Γνωσιακού Γράφου.

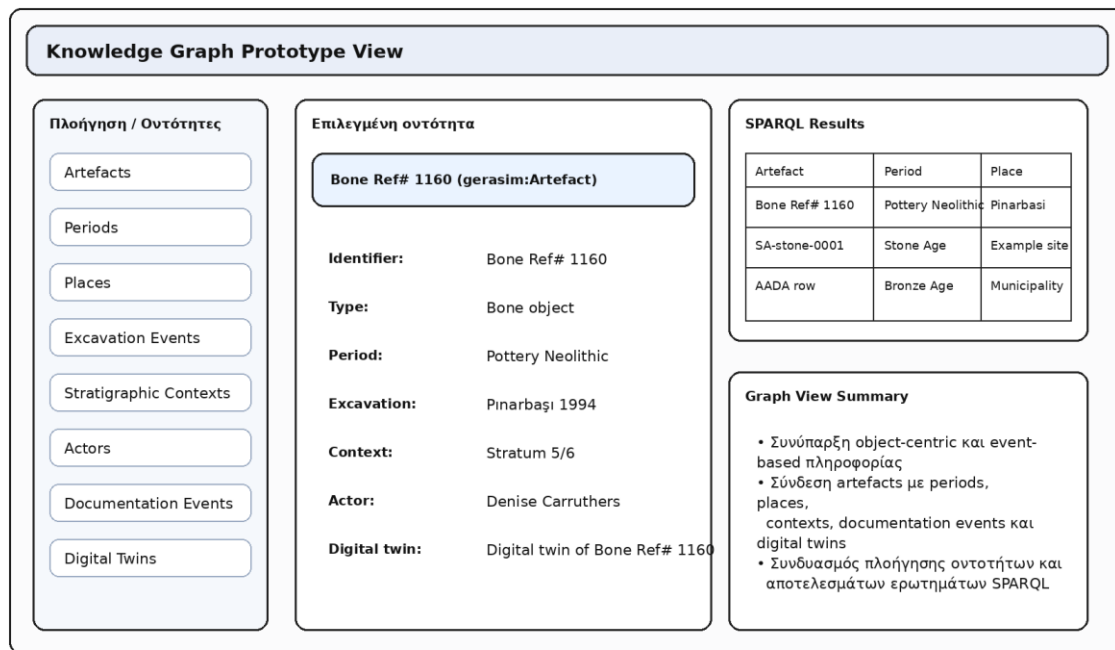
Συνολικά, η εκτέλεση και ο σχολιασμός των SPARQL ερωτημάτων δείχνουν ότι η προτεινόμενη οντολογία υποστηρίζει αποτελεσματικά τόσο την περιγραφή όσο και την εξερεύνηση των δεδομένων πολιτιστικής κληρονομιάς. Τα αποτελέσματα επιβεβαιώνουν ότι το μοντέλο είναι σε θέση να καλύψει τον object-centric πυρήνα των δεδομένων, να ενσωματώσει event-based αρχαιολογική πληροφορία, να αναπαραστήσει τεκμηριωτικές πράξεις και provenance, και να συνδέσει τα φυσικά ευρήματα με ψηφιακές αναπαραστάσεις. Με τον τρόπο αυτό, το παρόν κεφάλαιο τεκμηριώνει ότι ο παραγόμενος Γνωσιακός Γράφος δεν αποτελεί μόνο τεχνικό αποτέλεσμα μετασχηματισμού, αλλά και λειτουργικό μέσο σημασιολογικής διερεύνησης της πληροφορίας.

Στην επόμενη ενότητα παρουσιάζεται συνοπτικά και η πρωτοτυπική όψη του γράφου, δηλαδή ο τρόπος με τον οποίο η παραγόμενη πληροφορία μπορεί να εξεταστεί και να παρουσιαστεί σε ένα περιβάλλον πρακτικής χρήσης.

6.4 Πρωτότυπο και παρουσίαση του Γνωσιακού Γράφου

Πέρα από την οντολογική μοντελοποίηση, τον μετασχηματισμό των δεδομένων σε RDF και την εκτέλεση ερωτημάτων SPARQL, σημαντικό μέρος της παρούσας εργασίας αποτελεί και η πρακτική παρουσίαση του παραγόμενου Γνωσιακού Γράφου. Η παρουσίαση αυτή δεν αφορά την ανάπτυξη ενός πλήρως λειτουργικού πληροφοριακού συστήματος μεγάλης κλίμακας, αλλά τη δημιουργία ενός πρωτοτύπου, μέσω του οποίου καθίσταται δυνατή η επιθεώρηση, η πλοήγηση και η βασική εξερεύνηση των δεδομένων που οργανώθηκαν σημασιολογικά. Το παρόν υποκεφάλαιο δεν παρουσιάζει πλήρως ανεπτυγμένη τεχνική εφαρμογή ή ανεξάρτητο user interface. Αντίθετα, λειτουργεί ως conceptual design και ενδεικτική παρουσίαση του τρόπου με τον οποίο ο παραγόμενος Γνωσιακός Γράφος θα μπορούσε να αξιοποιηθεί σε ένα μελλοντικό περιβάλλον εξερεύνησης πολιτιστικών δεδομένων.

Η έννοια του πρωτοτύπου στην παρούσα εργασία συνδέεται κυρίως με την ανάγκη να αποδειχθεί ότι ο γράφος δεν αποτελεί μόνο ένα αφηρημένο ή θεωρητικό μοντέλο, αλλά ένα πρακτικά αξιοποιήσιμο αποτέλεσμα. Με άλλα λόγια, το ζητούμενο δεν είναι μόνο να υπάρχουν RDF τριάδες και ένα συνεκτικό σχήμα οντολογικής αναπαράστασης, αλλά και να μπορεί ο χρήστης ή ο ερευνητής να δει στην πράξη πώς οργανώνονται οι πληροφορίες, πώς συνδέονται οι οντότητες μεταξύ τους και πώς μπορεί να ανακτηθεί γνώση από το σύνολο των δεδομένων. Για να αποδοθεί πιο απτά η λογική του πρωτοτύπου, η ακόλουθη εικόνα παρουσιάζει ενδεικτικά έναν τρόπο με τον οποίο μπορεί να οργανωθεί η πλοήγηση, η προβολή οντοτήτων και η παρουσίαση αποτελεσμάτων πάνω στον Γνωσιακό Γράφο.



Σχήμα 9: Ενδεικτική πρωτοτυπική παρουσίαση του Γνωστικού Γράφου

Η παρουσίαση του Γνωστικού Γράφου μέσω πρωτοτύπου δεν αφορά ένα πλήρες πληροφοριακό σύστημα, αλλά ένα περιβάλλον πλοήγησης και επιθεώρησης της παραγόμενης σημασιολογικής δομής. Για να αποδοθεί πιο καθαρά η λειτουργική λογική του πρωτοτύπου, ο παρακάτω πίνακας συνοψίζει τα βασικά συστατικά του και τον ρόλο που επιτελεί το καθένα.

Συστατικό του πρωτοτύπου	Ρόλος
Πλοήγηση σε οντότητες	Επιλογή artefacts, periods, places, events και digital twins
Προβολή επιλεγμένης οντότητας	Παρουσίαση βασικών ιδιοτήτων και σχέσεων
Εκτέλεση SPARQL queries	Ανάκτηση και φιλτράρισμα σημασιολογικής πληροφορίας
Προβολή αποτελεσμάτων	Παρουσίαση των επιστρεφόμενων τιμών σε δομημένη μορφή
Οπτική επιθεώρηση σχέσεων	Κατανόηση της διασύνδεσης object-centric και event-based πληροφορίας

Πίνακας 19: Βασικά συστατικά του πρωτοτύπου και ο ρόλος τους

Ο πίνακας καθιστά σαφές ότι το πρωτότυπο λειτουργεί ως μέσο πλοήγησης, ανάκτησης και ερμηνείας της πληροφορίας του γράφου. Με βάση αυτή τη λογική, μπορεί να παρουσιαστεί στη συνέχεια το πώς η πρωτοτυπική αυτή όψη συμβάλλει στην κατανόηση και πρακτική αξιοποίηση του μοντέλου.

Στο πλαίσιο αυτό, το πρωτότυπο λειτουργεί ως ένα περιβάλλον επιβεβαίωσης της χρηστικότητας του παραγόμενου Γνωστικού Γράφου. Μέσα από αυτό μπορεί να παρατηρηθεί η παρουσία των βασικών artefacts, η σύνδεσή τους με τύπους, υλικά, περιόδους και τόπους, καθώς και η επέκταση της αναπαράστασης σε πιο σύνθετα στοιχεία, όπως ανασκαφικές δραστηριότητες, stratigraphic contexts, actors, source records και digital twins. Κατ' αυτόν τον τρόπο, το πρωτότυπο καθιστά ορατή τη μετάβαση από ένα συμβατικό σύνολο δεδομένων σε μια διασυνδεδεμένη σημασιολογική δομή.

Η παρουσίαση του γράφου μπορεί να γίνει σε δύο συμπληρωματικά επίπεδα. Το πρώτο επίπεδο αφορά την κειμενική και ερωτηματοκεντρική εξερεύνηση, δηλαδή την ανάκτηση πληροφορίας μέσω

SPARQL queries, όπως παρουσιάστηκε στις προηγούμενες ενότητες. Το δεύτερο επίπεδο αφορά την περισσότερο οπτική ή δομική επιθεώρηση των δεδομένων, όπου ο χρήστης μπορεί να δει τους κόμβους και τις σχέσεις του γράφου, να εντοπίσει πώς ένα artefact συνδέεται με μια περίοδο, έναν τόπο, ένα excavation event ή μια ψηφιακή αναπαράσταση, και να αποκτήσει έτσι μια συνολικότερη αίσθηση της δομής του μοντέλου.

Από τη σκοπιά της παρούσας εργασίας, το πρωτότυπο είναι ιδιαίτερα σημαντικό επειδή αναδεικνύει την πολυεπίπεδη λογική της οντολογίας. Ο χρήστης δεν βλέπει απλώς μεμονωμένες εγγραφές, αλλά ένα πλέγμα σχέσεων που οργανώνεται σε διακριτά επίπεδα: στο επίπεδο των αντικειμένων, στο γεγονотоκεντρικό επίπεδο, στο τεκμηριωτικό επίπεδο και στο επίπεδο της ψηφιακής αναπαράστασης. Η οπτική αυτή διάσταση βοηθά ιδιαίτερα στην κατανόηση του μοντέλου, διότι καθιστά σαφέστερο πώς η σημασιολογική αναπαράσταση υπερβαίνει τη λογική ενός απλού πίνακα δεδομένων.

Επιπλέον, η πρωτοτυπική παρουσίαση του Γνωσιακού Γράφου επιτρέπει να φανεί και η συμβολή της συνδυαστικής χρήσης των δύο use cases της εργασίας. Από τη μία πλευρά, το AADA προσφέρει το object-centric υπόβαθρο του μοντέλου, δηλαδή την περιγραφή artefacts με βασικά χαρακτηριστικά, περιόδους, τόπους και μετρήσεις. Από την άλλη πλευρά, το OpenContext ενισχύει την event-based διάσταση, εισάγοντας σχέσεις με ανασκαφικές δραστηριότητες, stratigraphic contexts, actors και documentation events. Το πρωτότυπο, επομένως, δεν αποτυπώνει μόνο μεμονωμένα δεδομένα, αλλά και την ίδια τη μεθοδολογική σύνθεση που επιχειρεί η εργασία.

Ένα ακόμη σημαντικό στοιχείο είναι ότι το πρωτότυπο λειτουργεί και ως μορφή ενδιάμεσης αξιολόγησης του ίδιου του μοντέλου. Κατά την επιθεώρηση του γράφου καθίσταται ευκολότερο να εντοπιστούν πλεονάζουσες σχέσεις, ελλείψεις, αδυναμίες στη διασύνδεση των οντοτήτων ή σημεία στα οποία απαιτείται περαιτέρω εξειδίκευση. Με αυτή την έννοια, η πρωτοτυπική παρουσίαση δεν είναι απλώς ένα εποπτικό συμπλήρωμα της εργασίας, αλλά και εργαλείο ελέγχου της συνοχής της μοντελοποίησης.

Παράλληλα, το πρωτότυπο δείχνει ότι ακόμη και ένας Γνωσιακός Γράφος μικρής κλίμακας μπορεί να αποτελέσει λειτουργική βάση για μελλοντικές επεκτάσεις. Εφόσον το μοντέλο υποστηρίζει ήδη συνδέσεις ανάμεσα σε artefacts, places, periods, excavation events, documentation events και digital twins, καθίσταται εφικτή η μελλοντική ενσωμάτωση περισσότερων datasets, πιο σύνθετων spatial ή temporal relations, καθώς και πλουσιότερων μορφών πολιτιστικής τεκμηρίωσης. Συνεπώς, το πρωτότυπο δεν αποτυπώνει μόνο το παρόν στάδιο της εργασίας, αλλά και τη δυνατότητα μελλοντικής εξέλιξής της.

Συνολικά, η παρουσίαση του Γνωσιακού Γράφου μέσω πρωτοτύπου ενισχύει την πρακτική διάσταση της παρούσας μελέτης και καθιστά από το αποτέλεσμα της οντολογικής μοντελοποίησης και της υλοποίησης. Μέσα από αυτή τη διαδικασία γίνεται εμφανές ότι ο παραγόμενος γράφος δεν αποτελεί απλώς ένα τεχνικό παραδοτέο, αλλά ένα οργανωμένο και εξερευνήσιμο σημασιολογικό περιβάλλον, ικανό να υποστηρίξει αναζήτηση, κατανόηση και συνδυαστική αξιοποίηση δεδομένων πολιτιστικής κληρονομιάς.

Με την ολοκλήρωση του παρόντος κεφαλαίου, έχει πλέον παρουσιαστεί τόσο η λογική δημιουργίας των ερωτημάτων όσο και η πρακτική αξιοποίηση και παρουσίαση των αποτελεσμάτων του Γνωσιακού Γράφου. Στο επόμενο κεφάλαιο συνοψίζονται τα βασικά συμπεράσματα της εργασίας και συζητούνται οι δυνατότητες μελλοντικής επέκτασης της προτεινόμενης προσέγγισης.

6.5 Ενδεικτικός έλεγχος εφαρμοσιμότητας σε τρίτο περιβάλλον δεδομένων: SearchCulture.gr

Πέρα από την εφαρμογή του προτεινόμενου μοντέλου στα δύο βασικά use cases της εργασίας, δηλαδή το AADA και το OpenContext, πραγματοποιήθηκε και ένας μικρής κλίμακας ενδεικτικός έλεγχος εφαρμοσιμότητας σε τρίτο περιβάλλον πολιτιστικών δεδομένων. Για τον σκοπό αυτό εξετάστηκε το SearchCulture.gr, ο Εθνικός Συσσωρευτής Ψηφιακού Πολιτιστικού Περιεχομένου, ο οποίος παρέχει κεντρική πρόσβαση σε ψηφιακές συλλογές πολιτισμού, ιστορίας και τέχνης από φορείς όπως μουσεία, αρχαία, βιβλιοθήκες, εφορείες αρχαιοτήτων και άλλους πολιτιστικούς οργανισμούς [20].

Ο έλεγχος αυτός δεν είχε ως στόχο την πλήρη μετατροπή τρίτου συνόλου δεδομένων σε RDF ούτε την επέκταση της υλοποίησης. Αντίθετα, λειτούργησε ως external applicability check, δηλαδή ως ενδεικτική αποτίμηση του βαθμού στον οποίο το application profile που αναπτύχθηκε μπορεί να χρησιμοποιηθεί και για πολιτιστικά μεταδεδομένα εκτός των δύο βασικών παραδειγμάτων της εργασίας.

Το SearchCulture.gr είναι ιδιαίτερα κατάλληλο για έναν τέτοιο έλεγχο, καθώς διαθέτει πεδία και δυνατότητες αναζήτησης που σχετίζονται άμεσα με την περιγραφή πολιτιστικών τεκμηρίων, όπως τύπος τεκμηρίου, θέμα, ιστορική περίοδος, χρονολόγηση, τόπος, φορέας/συλλογή, άδεια χρήσης ψηφιακού αρχείου και τύπος αρχείου [20], [21]. Επιπλέον, παρέχει μηχανισμούς διαλειτουργικότητας μέσω Web API και Linked Data, ενώ τα μεταδεδομένα των τεκμηρίων διατίθενται σύμφωνα με το Europeana Data Model (EDM), με δυνατότητα ανάκτησης RDF/XML και JSON-LD αναπαράστάσεων [21].

Ο παρακάτω πίνακας παρουσιάζει ενδεικτικά πώς βασικά πεδία ή πληροφορίες του SearchCulture.gr μπορούν να αντιστοιχιστούν στο προτεινόμενο μοντέλο της εργασίας.

Πίνακας 20: Ενδεικτικός έλεγχος κάλυψης πεδίων SearchCulture.gr από το προτεινόμενο μοντέλο

Πεδίο / πληροφορία από SearchCulture.gr	Καλύπτεται από το μοντέλο ;	Πώς μοντελοποιείται	Παρατήρηση
Τίτλος τεκμηρίου	Ναι	<code>rdfs:label</code>	Καλύπτεται άμεσα ως βασική ετικέτα του artefact ή resource
Αναγνωριστικό / URI τεκμηρίου	Ναι	URI πόρου, <code>crm:P1_is_identified_by</code> , <code>crm:E42_Identifier</code>	Το μοντέλο υποστηρίζει τόσο URI-based identification όσο και explicit identifiers
Τύπος τεκμηρίου	Ναι	<code>crm:P2_has_type</code> , <code>crm:E55_Type</code> , <code>skos:Concept</code>	Συμβατό με controlled vocabulary προσέγγιση
Θέμα / θεματική κατηγορία	Μερικώς	<code>crm:P2_has_type</code> , <code>skos:Concept</code>	Καλύπτεται ως lightweight classification, αλλά θα μπορούσε να επεκταθεί με ειδικότερο subject modelling
Ιστορική περίοδος	Ναι	<code>gerasim:associatedWithPeriod</code> , <code>gerasim:Period</code>	Καλύπτεται από το υβριδικό period modelling της εργασίας

Πεδίο / πληροφορία από SearchCult ure.gr	Καλύπτεται από το μοντέλο ;	Πώς μοντελοποιείται	Παρατήρηση
		crm:E4_Period, skos:Concept	
Χρονολόγηση	Ναι / μερικώς	crm:E52_Time-Span, crm:P4_has_time-span, crm:P82a_begin_of_the_begin , crm:P82b_end_of_the_end	Καλύπτεται όταν υπάρχουν διαθέσιμα begin/end ή χρονολογικό εύρος
Τόπος	Ναι	crm:E53_Place, geo:Feature, geo:hasGeometry, geo:asWKT	Καλύπτεται άμεσα μέσω CIDOC CRM και GeoSPARQL
Φορέας / συλλογή	Μερικώς	crm:E39_Actor, crm:E31_Document, prov:wasDerivedFrom, dcterms:source	Καλύπτεται ως provenance/provider information, αλλά όχι ως πλήρες aggregation metadata model
Εικόνα / ψηφιακή αναπαράσταση	Ναι / μερικώς	hdto:HC2_Heritage_Digital_Twin, hdto:HP1_has_digital_twin, crm:E36_Visual_Item	Το μοντέλο υποστηρίζει digital representation layer, αλλά μπορεί να χρειαστεί επέκταση για πολλά αρχεία ή IIIF-like structures
Δικαιώματα / άδεια χρήσης	Μερικώς	dcterms:rights, πιθανή επέκταση με rights vocabularies	Δεν αποτελεί κεντρικό μέρος του μοντέλου και θα χρειαζόταν περαιτέρω εξειδίκευση
Τύπος αρχείου / μορφότυπος	Μερικώς	dcterms:format ή επέκταση digital representation metadata	Καλύπτεται μόνο ως γενικό metadata πεδίο, όχι ως πλήρες media/file model
Σημασιολογικά εμπλουτισμένα πεδία	Μερικώς	skos:Concept, external URIs, owl:sameAs / alignment links όπου χρειάζεται	Το μοντέλο είναι συμβατό με external authorities, αλλά δεν υλοποιεί πλήρη enrichment pipeline

Από τον ενδεικτικό έλεγχο προκύπτει ότι το προτεινόμενο μοντέλο μπορεί να καλύψει άμεσα αρκετά από τα βασικά πεδία που εμφανίζονται σε ένα τρίτο περιβάλλον πολιτιστικών μεταδεδομένων, όπως το SearchCulture.gr. Συγκεκριμένα, πληροφορίες όπως τίτλος, identifier, τύπος τεκμηρίου, ιστορική περίοδος, χρονολόγηση, τόπος και βασική ψηφιακή αναπαράσταση μπορούν να αποδοθούν με όρους που ήδη υπάρχουν στο μοντέλο.

Παράλληλα, ο έλεγχος ανέδειξε και ορισμένα όρια του προτεινόμενου application profile. Πεδία όπως δικαιώματα χρήσης, τύπος αρχείου, αναλυτικό aggregation metadata και λεπτομερές provider/collection modelling μπορούν να καλυφθούν μόνο μερικώς ή θα απαιτούσαν πρόσθετη εξειδίκευση, πιθανώς με περαιτέρω χρήση Dublin Core Terms, DCAT, EDM ή εξειδικευμένων rights vocabularies.

Επομένως, το μοντέλο δεν είναι αποκλειστικά προσαρμοσμένο στα δύο αρχικά datasets της εργασίας, αλλά παρουσιάζει έναν βαθμό επαναχρησιμοποιησιμότητας και σε τρίτο πολιτιστικό περιβάλλον. Ταυτόχρονα, η δοκιμή αυτή δείχνει καθαρά τα όριά του: το μοντέλο είναι ισχυρό για την αναπαράσταση artefacts, τύπων, περιόδων, τόπων, provenance και digital representation σε βασικό επίπεδο, αλλά θα χρειαζόταν επέκταση αν ο στόχος ήταν η πλήρης κάλυψη aggregators όπως το SearchCulture.gr ή η Europeana.

Κεφάλαιο 7ο: Συμπεράσματα & Μελλοντικές Επεκτάσεις

Το παρόν κεφάλαιο συνοψίζει τα βασικά συμπεράσματα που προέκυψαν από τη θεωρητική διερεύνηση, τον οντολογικό σχεδιασμό, την υλοποίηση του Γνωσιακού Γράφου και την πειραματική εξερεύνηση των δεδομένων μέσω SPARQL ερωτημάτων. Στόχος του δεν είναι να επαναλάβει απλώς όσα παρουσιάστηκαν στα προηγούμενα κεφάλαια, αλλά να αποτιμήσει συνολικά την πορεία και τα αποτελέσματα της εργασίας, να αναδείξει τη συμβολή της και να εντοπίσει ρεαλιστικές δυνατότητες μελλοντικής επέκτασης.

Η εργασία κινήθηκε σε ένα πεδίο όπου η πολιτιστική κληρονομιά, η οντολογική μοντελοποίηση και οι τεχνολογίες του Σημασιολογικού Ιστού συναντώνται με στόχο την οργάνωση, τη διασύνδεση και την ουσιαστική αξιοποίηση ετερογενών δεδομένων. Μέσα από τη μελέτη δύο συμπληρωματικών περιπτώσεων εφαρμογής, του AADA και του OpenContext, επιχειρήθηκε όχι μόνο η παραγωγή RDF δεδομένων, αλλά και η συγκρότηση ενός μικρής κλίμακας αλλά συνεκτικού σημασιολογικού μοντέλου, το οποίο μπορεί να λειτουργήσει ως βάση για περαιτέρω ανάπτυξη.

Στις ενότητες που ακολουθούν παρουσιάζονται αρχικά τα βασικά συμπεράσματα της εργασίας, στη συνέχεια η ειδικότερη συνεισφορά της και τέλος ορισμένες πιθανές μελλοντικές επεκτάσεις, οι οποίες μπορούν να ενισχύσουν τόσο το ίδιο το οντολογικό μοντέλο όσο και την πρακτική αξιοποίησή του σε ευρύτερα πολιτιστικά και αρχαιολογικά περιβάλλοντα.

7.1 Βασικά συμπεράσματα

Η παρούσα εργασία έδειξε ότι η ανάπτυξη ενός Γνωσιακού Γράφου για δεδομένα πολιτιστικής κληρονομιάς αποτελεί μια εφικτή και μεθοδολογικά γόνιμη προσέγγιση για την οργάνωση και τη σημασιολογική αναπαράσταση ετερογενών αρχαιολογικών δεδομένων. Η μετάβαση από συμβατικά, μη-RDF σύνολα δεδομένων σε ένα διασυνδεδεμένο σημασιολογικό σχήμα δεν περιορίζεται σε μια απλή τεχνική μετατροπή μορφότυπου, αλλά προϋποθέτει εννοιολογικές επιλογές, οντολογική τεκμηρίωση και σαφή σχεδιαστική λογική.

Ένα πρώτο βασικό συμπέρασμα είναι ότι το CIDOC CRM, σε συνδυασμό με συμπληρωματικές οντολογίες όπως το CRMarchaeo, το GeoSPARQL [17], το SKOS [14] και επιλεγμένα στοιχεία από την HDTO [18], μπορεί να προσφέρει ένα ιδιαίτερα ισχυρό υπόβαθρο για τη μοντελοποίηση πολιτιστικών και αρχαιολογικών δεδομένων. Η εργασία έδειξε ότι δεν είναι απαραίτητο να αναπτυχθεί ένα εξ ολοκλήρου νέο οντολογικό σχήμα από την αρχή. Αντίθετα, η ουσιαστική αξιοποίηση και, όπου χρειάζεται, η ελεγχόμενη εξειδίκευση υπαρχόντων προτύπων οδηγεί σε πιο συνεκτικό, πιο τεκμηριωμένο και περισσότερο διαλειτουργικό αποτέλεσμα.

Ένα δεύτερο σημαντικό συμπέρασμα αφορά τη διάκριση ανάμεσα σε περιπτώσεις όπου αρκεί η άμεση επαναχρησιμοποίηση υπαρχουσών κλάσεων και ιδιοτήτων και σε περιπτώσεις όπου είναι σκόπιμη η εισαγωγή περιορισμένων τοπικών application-level shortcuts. Η αναθεώρηση της οντολογίας με βάση την κριτική αξιολόγηση των τοπικών όρων έδειξε ότι η μοντελοποίηση γίνεται πιο καθαρή όταν αποφεύγονται πλεονάζοντες επαναορισμοί και όταν οι custom όροι διατηρούνται μόνο εκεί όπου εξυπηρετούν πραγματικά μια ειδικότερη εφαρμοστική ανάγκη. Το συμπέρασμα αυτό είναι ιδιαίτερα σημαντικό, διότι αναδεικνύει ότι η ποιότητα μιας οντολογίας δεν εξαρτάται από το πόσους νέους όρους εισάγει, αλλά από το πόσο τεκμηριωμένα και φειδωλά τους χρησιμοποιεί.

Ένα τρίτο συμπέρασμα είναι ότι η συνδυαστική αξιοποίηση των δύο use cases ενίσχυσε ουσιαστικά τη μεθοδολογική πληρότητα της εργασίας. Το AADA λειτούργησε ως σταθερή βάση για την object-

centric αναπαράσταση αρχαιολογικών ευρημάτων, καλύπτοντας identifiers, τύπους, υλικά, περιόδους, τόπους και μετρήσεις. Το OpenContext, αντίθετα, ανέδειξε τη γεγονотоκεντρική διάσταση της πολιτιστικής πληροφορίας, επιτρέποντας τη σύνδεση των artefacts με ανασκαφικές δραστηριότητες, stratigraphic contexts, actors, documentation events και ψηφιακές αναπαραστάσεις. Η συνύπαρξη των δύο αυτών προσεγγίσεων μέσα σε κοινό οντολογικό πυρήνα έδειξε ότι το μοντέλο μπορεί να καλύψει περισσότερες από μία μορφές πολιτιστικής πληροφορίας χωρίς να χάνει τη συνοχή του.

Ένα ακόμη βασικό συμπέρασμα είναι ότι η χρήση competency questions αποδείχθηκε ιδιαίτερα χρήσιμη ως πρακτικό εργαλείο σχεδιασμού. Παρότι δεν εφαρμόστηκε μια πλήρως τυποποιημένη μεθοδολογία μηχανικής οντολογιών σε όλη της την αυστηρότητα, τα competency questions λειτούργησαν ως ουσιαστικός μηχανισμός σύνδεσης των απαιτήσεων της εργασίας με τις τελικές σχεδιαστικές επιλογές. Μέσα από αυτά, κατέστη δυνατό να καθοριστούν οι βασικές οντότητες, οι κρίσιμες σχέσεις και οι τύποι ερωτημάτων που όφειλε να υποστηρίζει ο παραγόμενος γράφος. Η μετατροπή των competency questions σε SPARQL queries επιβεβαίωσε στην πράξη ότι ο οντολογικός σχεδιασμός δεν παρέμεινε αφηρημένος, αλλά συνδέθηκε με συγκεκριμένες πληροφοριακές ανάγκες.

Ιδιαίτερα σημαντικό είναι επίσης το συμπέρασμα ότι ο παραγόμενος Γνωσιακός Γράφος δεν έχει μόνο περιγραφική, αλλά και διερευνητική αξία. Η εκτέλεση των SPARQL ερωτημάτων έδειξε ότι το μοντέλο μπορεί να υποστηρίζει όχι μόνο βασική ανάκτηση πληροφορίας, αλλά και πιο σύνθετες ερωτήσεις που αφορούν σχέσεις ανάμεσα σε περιόδους, τόπους, γεγονότα, πρόσωπα, τεκμηριωτικές διαδικασίες και ψηφιακές αναπαραστάσεις. Με άλλα λόγια, ο γράφος δεν λειτουργεί απλώς ως μέσο αποθήκευσης μετασχηματισμένων δεδομένων, αλλά ως οργανωμένο σημασιολογικό περιβάλλον εξερεύνησης της γνώσης.

Ένα ακόμη συμπέρασμα αφορά την αξία της πολυεπίπεδης αρχιτεκτονικής που υιοθετήθηκε στην οντολογία. Ο διαχωρισμός ανάμεσα στο επίπεδο των αντικειμένων, στο γεγονотоκεντρικό επίπεδο, στο επίπεδο τεκμηρίωσης/provenance και στο ψηφιακό επίπεδο αποδείχθηκε χρήσιμος τόσο εννοιολογικά όσο και πρακτικά. Η δομή αυτή επέτρεψε να φανεί ότι ένα artefact μπορεί να αποτελέσει κόμβο συνάντησης διαφορετικών τύπων πληροφορίας, από βασικά περιγραφικά χαρακτηριστικά μέχρι excavation events, digital twins και documentation creation events. Έτσι, η εργασία ανέδειξε ότι η σημασιολογική μοντελοποίηση στην πολιτιστική κληρονομιά αποκτά ιδιαίτερη αξία όταν αποτυπώνει πολλαπλά επίπεδα σχέσεων και όχι μόνο μεμονωμένες ιδιότητες αντικειμένων.

Τέλος, η εργασία κατέδειξε ότι ακόμη και μια μικρής κλίμακας υλοποίηση μπορεί να λειτουργήσει ως ουσιαστικό proof of concept για ευρύτερες σημασιολογικές εφαρμογές στον χώρο της πολιτιστικής κληρονομιάς. Παρά το περιορισμένο εύρος της, η προτεινόμενη προσέγγιση απέδειξε ότι η οντολογική μοντελοποίηση, η μετατροπή σε RDF, η τεκμηριωμένη χρήση καθιερωμένων οντολογιών και η εξερεύνηση μέσω SPARQL μπορούν να συνδυαστούν σε ένα ενιαίο και λειτουργικό αποτέλεσμα. Υπό αυτή την έννοια, η εργασία δεν περιορίζεται σε μια ακαδημαϊκή άσκηση μοντελοποίησης, αλλά προσφέρει ένα μικρό αλλά ουσιαστικό παράδειγμα του πώς μπορεί να οργανωθεί και να αξιοποιηθεί σημασιολογικά η πολιτιστική πληροφορία.

7.2 Συνεισφορά της εργασίας

Η συνεισφορά της παρούσας εργασίας μπορεί να εξεταστεί σε τρία αλληλένδετα επίπεδα: στο επίπεδο του οντολογικού σχεδιασμού, στο επίπεδο της πρακτικής υλοποίησης και στο επίπεδο της αξιοποίησης του παραγόμενου αποτελέσματος.

Σε επίπεδο οντολογικού σχεδιασμού, η βασική συνεισφορά της εργασίας έγκειται στη διαμόρφωση ενός συνεκτικού, μικρής κλίμακας αλλά λειτουργικού μοντέλου για δεδομένα πολιτιστικής

κληρονομιάς, το οποίο βασίζεται κυρίως στο CIDOC CRM και συμπληρώνεται από επιλεγμένα στοιχεία του CRMarchaeo, του GeoSPARQL [17], του SKOS [14], του PROV-O [15], του DCAT [16] και της HDTTO [18]. Η εργασία δεν περιορίστηκε σε μια απλή θεωρητική αναφορά των σχετικών οντολογιών, αλλά προχώρησε σε συγκεκριμένες επιλογές επαναχρησιμοποίησης, εξειδίκευσης και περιορισμένης τοπικής επέκτασης, με στόχο τη συγκρότηση ενός μοντέλου κατάλληλου για τις ανάγκες των επιλεγμένων συνόλων δεδομένων. Υπό αυτή την έννοια, η εργασία συνεισφέρει ένα πρακτικά τεκμηριωμένο παράδειγμα του πώς μπορούν να συνδυαστούν υφιστάμενα σημασιολογικά πρότυπα χωρίς να παράγεται ένα πλεονάζον ή ασύνδετο οντολογικό σχήμα.

Ιδιαίτερα σημαντική είναι και η συνεισφορά της εργασίας ως προς τη μεθοδολογική στάση απέναντι στην επαναχρησιμοποίηση υπαρχουσών οντολογιών. Μέσα από τη διαδικασία αναθεώρησης του μοντέλου, κατέστη σαφές ότι ο στόχος δεν είναι η δημιουργία όσο το δυνατόν περισσότερων νέων κλάσεων και ιδιοτήτων, αλλά η τεκμηριωμένη επιλογή ανάμεσα σε άμεση επαναχρησιμοποίηση, ελεγχόμενη εξειδίκευση και περιορισμένη δημιουργία τοπικών shortcuts. Η προσέγγιση αυτή αποτελεί ουσιαστική συνεισφορά της εργασίας, διότι δείχνει στην πράξη μια πιο ώριμη λογική μηχανικής οντολογιών, στην οποία το ζητούμενο είναι η σημασιολογική καθαρότητα, η διαλειτουργικότητα και η αναγνωσιμότητα του μοντέλου και όχι η άκριτη παραγωγή νέων όρων.

Σε επίπεδο πρακτικής υλοποίησης, η εργασία συνεισφέρει με τη δημιουργία ενός συγκεκριμένου RDF Γνωσιακού Γράφου που προέκυψε από τον μετασχηματισμό μη-RDF πολιτιστικών δεδομένων. Η συνεισφορά εδώ δεν αφορά μόνο το τελικό αποτέλεσμα, αλλά και τη συνολική διαδικασία που ακολουθήθηκε: επιλογή συνόλων δεδομένων, ανάλυση της δομής τους, αντιστοίχιση των πεδίων τους σε κλάσεις και ιδιότητες της οντολογίας, οργάνωση των παραγόμενων τριπλετών και ενσωμάτωση τους σε ενιαίο σημασιολογικό σχήμα. Η εργασία δείχνει έτσι με πρακτικό τρόπο πώς ένα πολιτιστικό σύνολο δεδομένων μπορεί να μετασχηματιστεί σε RDF όχι μηχανικά, αλλά με βάση μια σαφή οντολογική λογική.

Μια ακόμη σημαντική συνεισφορά στο επίπεδο της υλοποίησης είναι η συγκρότηση παραδειγμάτων και μοτίβων που αναδεικνύουν δύο διαφορετικές αλλά συμπληρωματικές λογικές αναπαράστασης: την object-centric και τη γεγονοκεντρική / event-based. Η συνύπαρξη των δύο αυτών λογικών μέσα στο ίδιο μοντέλο δείχνει ότι η εργασία δεν περιορίστηκε στην περιγραφή μεμονωμένων ευρημάτων, αλλά επιχείρησε να οργανώσει και τις σχέσεις τους με γεγονότα, stratigraphic contexts, actors, documentation events και digital representations. Με αυτόν τον τρόπο, η υλοποίηση υπερβαίνει το επίπεδο ενός απλού RDF export και λειτουργεί ως μικρό αλλά ουσιαστικό demonstrator σημασιολογικής ενοποίησης πολιτιστικής πληροφορίας.

Σε επίπεδο αξιοποίησης του αποτελέσματος, η εργασία συνεισφέρει μέσω της συστηματικής σύνδεσης του οντολογικού σχεδιασμού με competency questions και με SPARQL queries. Η συνεισφορά αυτή είναι σημαντική, διότι δείχνει ότι το μοντέλο δεν σχεδιάστηκε μόνο για να “υπάρχει” ως θεωρητικό σχήμα, αλλά για να απαντά συγκεκριμένα ερωτήματα. Η μετατροπή των competency questions σε ερωτήματα SPARQL και η παρουσίαση των αποτελεσμάτων τους επιβεβαιώνουν ότι ο γράφος μπορεί να λειτουργήσει ως περιβάλλον αναζήτησης, διερεύνησης και σύνθεσης πληροφορίας. Επομένως, η εργασία συνεισφέρει όχι μόνο με την παραγωγή ενός οντολογικού μοντέλου, αλλά και με την απόδειξη της λειτουργικής του αξίας.

Παράλληλα, η εργασία προσφέρει ένα μικρής κλίμακας αλλά σαφές πρωτότυπο λογικής παρουσίασης και εξερεύνησης του παραγόμενου Γνωσιακού Γράφου. Ακόμη και αν δεν αναπτύχθηκε ένα πλήρες τελικό πληροφοριακό σύστημα, η περιγραφή της πλοήγησης στις οντότητες, της προβολής των σχέσεων και της αξιοποίησης των αποτελεσμάτων των SPARQL queries δείχνει πώς ένα τέτοιο

μοντέλο μπορεί να μετατραπεί σε περιβάλλον χρήσιμο για επιθεώρηση και ερμηνεία της πολιτιστικής πληροφορίας. Η συνεισφορά αυτή είναι σημαντική, επειδή γεφυρώνει το χάσμα ανάμεσα στη μοντελοποίηση και στη χρήση.

Τέλος, η συνολική συνεισφορά της εργασίας μπορεί να ιδωθεί και ως proof of concept για τη σημασιολογική οργάνωση δεδομένων πολιτιστικής κληρονομιάς σε μικρό αλλά ρεαλιστικό εύρος. Η εργασία δεν ισχυρίζεται ότι επιλύει συνολικά το πρόβλημα της μοντελοποίησης του αρχαιολογικού πεδίου, ούτε ότι εξαντλεί όλες τις δυνατότητες των χρησιμοποιούμενων οντολογιών. Ωστόσο, προσφέρει ένα τεκμηριωμένο, λειτουργικό και αναπαραγώγιμο παράδειγμα που δείχνει πώς μπορούν να συνδυαστούν θεωρία, οντολογικός σχεδιασμός, RDF μετασχηματισμός και σημασιολογική εξερεύνηση σε μια ενιαία ερευνητική εργασία. Αυτή ακριβώς η σύνδεση θεωρητικού και πρακτικού επιπέδου συνιστά και την ουσιαστικότερη συνεισφορά της.

7.3 Πιθανές επεκτάσεις

Παρότι η παρούσα εργασία οδήγησε σε ένα συνεκτικό και λειτουργικό αποτέλεσμα, το προτεινόμενο μοντέλο παραμένει σκόπιμα περιορισμένο ως προς το εύρος και την κλίμακά του. Η επιλογή αυτή ήταν αναγκαία, ώστε η εργασία να παραμείνει μεθοδολογικά ελεγχόμενη και πρακτικά διαχειρίσιμη. Ωστόσο, ακριβώς επειδή συγκροτήθηκε ένας σαφής οντολογικός πυρήνας και ένας λειτουργικός RDF Γνωσιακός Γράφος, προκύπτουν αρκετές δυνατότητες μελλοντικής επέκτασης, τόσο σε επίπεδο μοντέλου όσο και σε επίπεδο δεδομένων, ερωτημάτων και εφαρμογών.

Μια πρώτη και προφανής κατεύθυνση επέκτασης αφορά τον εμπλουτισμό των συνόλων δεδομένων που χρησιμοποιήθηκαν. Στην παρούσα εργασία επιλέχθηκαν δύο συγκεκριμένα use cases, τα οποία λειτούργησαν συμπληρωματικά και επέτρεψαν να αναδειχθούν τόσο η object-centric όσο και η event-based διάσταση της πολιτιστικής πληροφορίας. Σε μελλοντική συνέχεια, θα μπορούσαν να ενσωματωθούν περισσότερα σύνολα δεδομένων από διαφορετικά αποθετήρια πολιτιστικής κληρονομιάς, ώστε να ελεγχθεί σε μεγαλύτερη κλίμακα η επάρκεια και η προσαρμοστικότητα του μοντέλου. Μια τέτοια διεύρυνση θα επέτρεπε να φανεί σε ποιο βαθμό ο οντολογικός πυρήνας παραμένει σταθερός όταν καλείται να υποστηρίξει περισσότερες κατηγορίες τεκμηρίων, περισσότερα αρχαιολογικά συμφραζόμενα και μεγαλύτερη ετερογένεια στη δομή των πηγών.

Μια δεύτερη σημαντική επέκταση αφορά τη λεπτομερέστερη γεγονотоκεντρική αναπαράσταση της αρχαιολογικής πληροφορίας. Στην παρούσα εργασία αξιοποιήθηκαν στοιχεία του CRMarchaeo σε ελεγχόμενο βαθμό, κυρίως για την απόδοση excavation events και stratigraphic contexts. Στο μέλλον, το μοντέλο θα μπορούσε να εμπλουτιστεί με πιο αναλυτική αναπαράσταση στρωματογραφικών σχέσεων, ακολουθιών ανασκαφικών δραστηριοτήτων, διαφορετικών τύπων τεκμηριωτικών πράξεων και πιο σύνθετων σχέσεων ανάμεσα σε archaeological features, units, observations και findings. Μια τέτοια επέκταση θα προσέφερε μεγαλύτερη αρχαιολογική ακρίβεια και θα επέτρεπε πιο εξειδικευμένα ερωτήματα πάνω στα δεδομένα.

Μια τρίτη κατεύθυνση αφορά την περαιτέρω ανάπτυξη του επιπέδου provenance και documentation. Στην παρούσα μορφή του, το μοντέλο αποδίδει source records,

documentation creation events και συνδέσεις με τα αρχικά σύνολα δεδομένων. Ωστόσο, θα μπορούσε στο μέλλον να επεκταθεί ώστε να περιλαμβάνει αναλυτικότερη τεκμηρίωση του κύκλου ζωής των δεδομένων, πληροφορία για το ποιος πραγματοποίησε τον μετασχηματισμό, πότε έγινε, ποιες πηγές χρησιμοποιήθηκαν, ποια mappings εφαρμόστηκαν και ποιες επεμβάσεις ή διορθώσεις πραγματοποιήθηκαν. Μια τέτοια προσέγγιση θα ενίσχυε σημαντικά τη διαφάνεια, την επαναχρησιμοποίηση και την αξιοπιστία του παραγόμενου Γνωσιακού Γράφου, ιδίως σε περιβάλλοντα όπου η τεκμηριωμένη προέλευση της πληροφορίας είναι κρίσιμη.

Ιδιαίτερο ενδιαφέρον παρουσιάζει και η περαιτέρω αξιοποίηση του digital twin / digital representation layer. Στην παρούσα εργασία το επίπεδο αυτό ενσωματώθηκε με ελεγχόμενο τρόπο, ώστε να αποδοθούν περιπτώσεις ψηφιακής αναπαράστασης πολιτιστικών αντικειμένων. Στο μέλλον, το επίπεδο αυτό θα μπορούσε να αναπτυχθεί περισσότερο, ώστε να περιλαμβάνει όχι μόνο απλές συνδέσεις με ψηφιακά τεκμήρια, αλλά και πιο σύνθετες μορφές ψηφιακής αναπαράστασης, όπως τρισδιάστατα μοντέλα, φωτογραμμετρικές αποτυπώσεις, διαφορετικές εκδοχές ανακατασκευής, επίπεδα ανάλυσης ή ακόμη και συσχετίσεις ανάμεσα σε φυσικά και ψηφιακά χαρακτηριστικά του ίδιου αντικειμένου. Με τον τρόπο αυτό, η εργασία θα μπορούσε να συνδεθεί πιο άμεσα με σύγχρονες προσεγγίσεις digital heritage και heritage digital twins.

Μια ακόμη σημαντική μελλοντική επέκταση αφορά τη βελτίωση και εμβάθυνση των SPARQL queries. Στην παρούσα εργασία τα ερωτήματα σχεδιάστηκαν ώστε να ανταποκρίνονται στα competency questions και να δείχνουν τη λειτουργικότητα του μοντέλου. Σε επόμενο στάδιο, θα μπορούσε να αναπτυχθεί ένα πιο εκτεταμένο σύνολο ερωτημάτων, με μεγαλύτερη χρήση φίλτρων, ομαδοποιήσεων, aggregation functions, property paths και συνδυαστικών μηχανισμών ανάκτησης. Επιπλέον, θα μπορούσε να διερευνηθεί και η αυτόματη παραγωγή ή υποβοήθηση διατύπωσης SPARQL queries με βάση φυσική γλώσσα ή προκαθορισμένα query templates, κάτι που θα έκανε το μοντέλο πιο προσιτό και σε μη εξειδικευμένους χρήστες.

Σε πρακτικό επίπεδο, μια ουσιαστική κατεύθυνση επέκτασης είναι η ανάπτυξη ενός πιο ολοκληρωμένου πρωτοτύπου ή ακόμη και μιας διαδραστικής εφαρμογής εξερεύνησης του Γνωσιακού Γράφου. Στην παρούσα εργασία παρουσιάστηκε μόνο μια ενδεικτική λογική πλοήγησης και προβολής της πληροφορίας. Στο μέλλον, θα μπορούσε να υλοποιηθεί ένα πραγματικό interface που να επιτρέπει πλοήγηση σε artefacts, periods, places, excavation events και digital twins, εκτέλεση SPARQL queries μέσω φιλικού περιβάλλοντος, οπτικοποίηση των σχέσεων του γράφου και σύνδεση με διαδραστικούς χάρτες ή χρονογραμμές. Μια τέτοια εφαρμογή θα ενίσχυε σημαντικά την πρακτική αξιοποίηση της εργασίας και θα μετέφερε το αποτέλεσμα από το επίπεδο της ερευνητικής απόδειξης εφικτότητας σε πιο εφαρμοστικό περιβάλλον χρήσης.

Μια ακόμη προοπτική αφορά την ενσωμάτωση συστηματικού μηχανισμού επικύρωσης με SHACL shapes. Αν και στην παρούσα εργασία ο έλεγχος βασίστηκε κυρίως στη συντακτική εγκυρότητα του TTL, στην οντολογική επιθεώρηση και στην εκτέλεση SPARQL

ερωτημάτων, μια μελλοντική εκδοχή θα μπορούσε να ορίσει ρητούς SHACL κανόνες για την υποχρεωτική παρουσία βασικών ιδιοτήτων, τη συμβατότητα των κλάσεων και τη συνέπεια των σχέσεων.

Τέλος, μια ευρύτερη κατεύθυνση μελλοντικής εργασίας θα μπορούσε να αφορά τη σύνδεση του παραγόμενου Γνωσιακού Γράφου με εξωτερικά authority files, controlled vocabularies και linked data resources. Στην παρούσα εργασία έγιναν ήδη ορισμένα βήματα προς αυτή την κατεύθυνση, ιδίως με τη χρήση περιοδικών εννοιών και σημασιολογικών αντιστοιχίσεων. Ωστόσο, σε μελλοντική συνέχεια θα μπορούσε να ενισχυθεί η διασύνδεση με εξωτερικές πηγές, όπως γεωγραφικά authority datasets, εξειδικευμένα αρχαιολογικά λεξιλόγια, χρονολογικά vocabularies ή άλλες υποδομές πολιτιστικής κληρονομιάς. Η διασύνδεση αυτή θα αύξανε τη διαλειτουργικότητα, θα ενίσχυε τη σημασιολογική πληρότητα του γράφου και θα επέτρεπε τη συμμετοχή του σε ευρύτερα οικοσυστήματα linked cultural data.

Συνολικά, οι παραπάνω επεκτάσεις δείχνουν ότι η εργασία, παρότι ολοκληρώνεται ως ένα μικρής κλίμακας και ελεγχόμενου εύρους εγχείρημα, αφήνει ανοιχτό ένα ευρύ πεδίο περαιτέρω ανάπτυξης. Ο οντολογικός πυρήνας που σχεδιάστηκε, ο RDF μετασχηματισμός που υλοποιήθηκε και η διερεύνηση μέσω SPARQL που παρουσιάστηκε συνιστούν μια σταθερή βάση πάνω στην οποία μπορούν να στηριχθούν πιο εκτεταμένες μελλοντικές εφαρμογές. Με αυτή την έννοια, η παρούσα εργασία δεν κλείνει απλώς με ένα σύνολο συμπερασμάτων, αλλά ανοίγει και μια σειρά από επόμενα βήματα για τη σημασιολογική οργάνωση και αξιοποίηση δεδομένων πολιτιστικής κληρονομιάς.

Βιβλιογραφία

- [1] Hyvönen, E. (2012). Publishing and using cultural heritage linked data on the Semantic Web. *Synthesis Lectures on the Semantic Web: Theory and Technology*, 2(1), 1–159. <https://doi.org/10.2200/S00452ED1V01Y201210WBE003>
- [2] Hogan, A., Blomqvist, E., Cochez, M., d’Amato, C., de Melo, G., Gutierrez, C., Kirrane, S., Gayo, J. E. L., Navigli, R., Neumaier, S., Ngonga Ngomo, A.-C., Polleres, A., Rashid, S. M., Rula, A., Schmelzeisen, L., Sequeda, J., Staab, S., & Zimmermann, A. (2021). Knowledge graphs. *ACM Computing Surveys*, 54(4), Article 71. <https://doi.org/10.1145/3447772>
- [3] Schmidt, S. C., Thiery, F., & Trognitz, M. (2022). Practices of linked open data in archaeology and their realisation in Wikidata. *Digital*, 2(3), 333–364. <https://doi.org/10.3390/digital2030019>
- [4] W3C. (2014). RDF 1.1 Primer. <https://www.w3.org/TR/rdf11-primer/>
- [5] W3C. (2012). OWL 2 Web Ontology Language Primer (Second Edition). <https://www.w3.org/TR/owl2-primer/>
- [6] W3C. (2013). SPARQL 1.1 Overview. <https://www.w3.org/TR/sparql11-overview/>
- [7] RML.io. (2024, June 20). RDF Mapping Language (RML). <https://rml.io/specs/rml/>
- [8] W3C. (2017). Shapes Constraint Language (SHACL). <https://www.w3.org/TR/shacl/>
- [9] Doerr, M. (2003). The CIDOC conceptual reference model: An ontological approach to semantic interoperability of metadata. *AI Magazine*, 24(3), 75–92. <https://doi.org/10.1609/aimag.v24i3.1720>
- [10] CIDOC CRM SIG. (2021). Definition of the CIDOC Conceptual Reference Model v7.1.1. <https://cidoc-crm.org/Resources/definition-of-the-cidoc-conceptual-reference-model-v7.1.1>
- [11] CIDOC CRM SIG. (2024). Definition of the CRMarchaeo v2.1: An extension of CRMBase to support the archaeological excavation process. https://cidoc-crm.org/sites/default/files/CRMarchaeo_v2.1%28site%29.pdf
- [12] Niccolucci, F., Markhoff, B., Theodoridou, M., Felicetti, A., & Hermon, S. (2023). The Heritage Digital Twin: A bicycle made for two. The integration of digital methodologies into cultural heritage research. *Open Research Europe*, 3, 64. <https://doi.org/10.12688/openreseurope.15496.1>
- [13] Pesonen, P., Nordfors, U., Roose, M., Saipio, J., Tiilikkala, J., Sanwal, M. U., Immonen, V., Vesakoski, O., & Onkamo, P. (2024). Archaeological Artefact Database of Finland (AADA). *Scientific Data*, 11, 815. <https://doi.org/10.1038/s41597-024-03602-8>
- [14] W3C. (2009). SKOS Simple Knowledge Organization System Reference. <https://www.w3.org/TR/skos-reference/>
- [15] W3C. (2013). PROV-O: The PROV Ontology. <https://www.w3.org/TR/prov-o/>
- [16] W3C. (2024). Data Catalog Vocabulary (DCAT) - Version 3. <https://www.w3.org/TR/vocab-dcat-3/>
- [17] Open Geospatial Consortium. (2024). GeoSPARQL - A Geographic Query Language for RDF Data. <https://www.ogc.org/standard/geosparql/>
- [18] ECHOES / FORTH-ICS. (n.d.). Heritage Digital Twin Ontology (HDTO) v1.1. https://isl.ics.forth.gr/ontology/echoes/html/HDT_v1.1.html
- [19] PeriodO. (n.d.). PeriodO: A gazetteer of scholarly definitions of historical, art-historical, and archaeological periods. <https://perio.do/>
- [20] SearchCulture.gr. (n.d.). Εθνικός Συσσωρευτής Ψηφιακού Πολιτιστικού Περιεχομένου. <https://www.searchculture.gr/aggregator/>
- [21] SearchCulture.gr. (n.d.). Διαλειτουργικότητα: Διασυνδεδεμένα Δεδομένα & SearchCulture API. <https://www.searchculture.gr/aggregator/portal/interoperability>

Παραρτήματα

Παράρτημα Α: Ενδεικτικά RDF παραδείγματα

A.1 Example A1

#####

Example A1 — object-centric AADA instance

#####

gres:artefact_KM_29954_10516

a crm:E22_Human-Made_Object, gerasim:Artefact ;

rdfs:label "KM 29954:10516"@en ;

crm:P1_is_identified_by gres:id_KM_29954_10516 ;

crm:P2_has_type gres:type_other_bone_artefact ;

crm:P53_has_former_or_current_location gres:place_92010043 ;

gerasim:associatedWithPeriod gres:period_n2 ;

crm:P43_has_dimension

gres:dim_frag_length_KM_29954_10516 ,

gres:dim_frag_width_KM_29954_10516 ,

gres:dim_frag_thickness_KM_29954_10516 ;

crm:P70_is_documented_in gres:record_KM_29954_10516 .

gres:id_KM_29954_10516 a crm:E42_Identifier ; rdfs:label "KM 29954_10516" .

gres:type_other_bone_artefact

a crm:E55_Type, skos:Concept ;

skos:prefLabel "Other bone artefact"@en ;

skos:altLabel "Luuesineen katkelma"@fi .

gres:period_n2

a crm:E4_Period, skos:Concept, gerasim:Period ;

skos:prefLabel "n2"@en .

gres:municipality_vantaa

a crm:E53_Place, gerasim:Municipality ;

rdfs:label "Vantaa"@en .

gres:place_92010043

a crm:E53_Place, geo:Feature ;
rdfs:label "Stenkulla"@en ;
crm:P1_is_identified_by gres:siteid_92010043 ;
geo:hasGeometry gres:geom_92010043 ;
gerasim:hasKKJCoordinate "3423130 6688380" ;
gerasim:locatedInMunicipality gres:municipality_vantaa .

gres:siteid_92010043 a crm:E42_Identifier ; rdfs:label "92010043" .

gres:geom_92010043

a geo:Geometry ;
geo:asWKT "POINT(25.06875558 60.29590896)"^^geo:wktLiteral .

gres:dim_frag_length_KM_29954_10516

a crm:E54_Dimension ;
crm:P2_has_type gres:fragment_length ;
crm:P90_has_value "10"^^xsd:decimal ;
crm:P91_has_unit gres:mm .

gres:dim_frag_width_KM_29954_10516

a crm:E54_Dimension ;
crm:P2_has_type gres:fragment_width ;
crm:P90_has_value "7"^^xsd:decimal ;
crm:P91_has_unit gres:mm .

gres:dim_frag_thickness_KM_29954_10516

a crm:E54_Dimension ;
crm:P2_has_type gres:fragment_thickness ;
crm:P90_has_value "5"^^xsd:decimal ;
crm:P91_has_unit gres:mm .

gres:record_KM_29954_10516

a crm:E31_Document, prov:Entity, gerasim:DatasetRecord ;
dcterms:source <<https://zenodo.org/records/11279777>> ;
rdfs:seeAlso
DATABase/blob/main/AADA%20Stone%20Age/SA%20bone.csv> .

<<https://github.com/AADA-DATA/AADA->

A.2 Example A2

#####

Example A2 — object-centric AADA instance

#####

gres:artefact_KM_30460_10967

a crm:E22_Human-Made_Object, gerasim:Artefact ;

rdfs:label "KM 30460:10967"@en ;

crm:P1_is_identified_by gres:id_KM_30460_10967 ;

crm:P2_has_type gres:type_other_bone_artefact ;

crm:P53_has_former_or_current_location gres:place_707010045 ;

gerasim:associatedWithPeriod gres:period_n2 ;

crm:P43_has_dimension gres:dim_length_KM_30460_10967 , gres:dim_width_KM_30460_10967 ;

crm:P70_is_documented_in gres:record_KM_30460_10967 .

gres:id_KM_30460_10967 a crm:E42_Identifier ; rdfs:label "KM 30460_10967" .

gres:municipality_raakkyla

a crm:E53_Place, gerasim:Municipality ;

rdfs:label "Rääkkylä"@fi .

gres:place_707010045

a crm:E53_Place, geo:Feature ;

rdfs:label "Vihi 1"@en ;

crm:P1_is_identified_by gres:siteid_707010045 ;

geo:hasGeometry gres:geom_707010045 ;

gerasim:hasKKJCoordinate "3636200 6907480" ;

gerasim:locatedInMunicipality gres:municipality_raakkyla .

gres:siteid_707010045 a crm:E42_Identifier ; rdfs:label "707010045" .

gres:geom_707010045

a geo:Geometry ;

geo:asWKT "POINT(29.97003938 62.31080103)"^^geo:wktLiteral .

gres:dim_length_KM_30460_10967

a crm:E54_Dimension ;

```

crm:P2_has_type gres:length ;
crm:P90_has_value "11"^^xsd:decimal ;
crm:P91_has_unit gres:mm .

gres:dim_width_KM_30460_10967
  a crm:E54_Dimension ;
  crm:P2_has_type gres:width ;
  crm:P90_has_value "4"^^xsd:decimal ;
  crm:P91_has_unit gres:mm .

gres:record_KM_30460_10967
  a crm:E31_Document, prov:Entity, gerasim:DatasetRecord ;
  dcterms:source <https://zenodo.org/records/11279777> ;
  rdfs:seeAlso
    <https://github.com/AADA-DATA/AADA-
    DATABASE/blob/main/AADA%20Stone%20Age/SA%20bone.csv> .

```

A.3 Example A3

```

#####
# Example A3 — object-centric AADA instance
#####

```

```

gres:artefact_KM_33461_305
  a crm:E22_Human-Made_Object, gerasim:Artefact ;
  rdfs:label "KM 33461:305"@en ;
  crm:P1_is_identified_by gres:id_KM_33461_305 ;
  crm:P2_has_type gres:type_other_bone_artefact ;
  crm:P53_has_former_or_current_location gres:place_433010038 ;
  gerasim:associatedWithPeriod gres:period_m3_n3 ;
  gerasim:locatedInMunicipality gres:municipality_loppi ;
  crm:P43_has_dimension
    gres:dim_frag_length_KM_33461_305 ,
    gres:dim_frag_width_KM_33461_305 ,
    gres:dim_frag_thickness_KM_33461_305 ;
  crm:P70_is_documented_in gres:record_KM_33461_305 .

gres:id_KM_33461_305 a crm:E42_Identifier ; rdfs:label "KM 33461_305" .

gres:period_m3_n3

```

a crm:E4_Period, skos:Concept, gerasim:Period ;
skos:prefLabel "m3 / n3"@en .

gres:municipality_loppi

a crm:E53_Place, gerasim:Municipality ;
rdfs:label "Loppi"@fi .

gres:place_433010038

a crm:E53_Place, geo:Feature ;
rdfs:label "Karhumäki"@fi ;
crm:P1_is_identified_by gres:siteid_433010038 ;
geo:hasGeometry gres:geom_433010038 ;
gerasim:hasKKJCoordinate "3365210 6732900" ;
gerasim:locatedInMunicipality gres:municipality_loppi .

gres:siteid_433010038 a crm:E42_Identifier ; rdfs:label "433010038" .

gres:geom_433010038

a geo:Geometry ;
geo:asWKT "POINT(24.4025334 60.70182809)"^^geo:wktLiteral .

gres:dim_frag_length_KM_33461_305

a crm:E54_Dimension ;
crm:P2_has_type gres:fragment_length ;
crm:P90_has_value "8"^^xsd:decimal ;
crm:P91_has_unit gres:mm .

gres:dim_frag_width_KM_33461_305

a crm:E54_Dimension ;
crm:P2_has_type gres:fragment_width ;
crm:P90_has_value "6"^^xsd:decimal ;
crm:P91_has_unit gres:mm .

gres:dim_frag_thickness_KM_33461_305

a crm:E54_Dimension ;
crm:P2_has_type gres:fragment_thickness ;
crm:P90_has_value "3"^^xsd:decimal ;
crm:P91_has_unit gres:mm .

```

gres:record_KM_33461_305
  a crm:E31_Document, prov:Entity, gerasim:DatasetRecord ;
  dcterms:source <https://zenodo.org/records/11279777> ;
  rdfs:seeAlso
    <https://github.com/AADA-DATA/AADA-
    DATABASE/blob/main/AADA%20Stone%20Age/SA%20bone.csv> .

```

A.4 Example B1

```

#####
# Example B1 — event-centric OpenContext instance
#####

```

```

<http://opencontext.org/subjects/TESTSPA0000005777>
  a gerasim:Artefact ;
  rdfs:label "Bone Ref# 1160"@en ;
  crm:P1_is_identified_by gres:id_TESTSPA0000005777 ;
  crm:P2_has_type gres:type_metatarsal , gres:taxon_ovis , gres:sex_male ;
  gerasim:associatedWithPeriod gres:period_-6700_-6000 ;
  crm:P12i_was_present_at gres:exc_TESTSPA0000005777 ;
  gerasim:hasStratigraphicContext gres:strat_BBH ;
  crm:P43_has_dimension gres:dim_weight_TESTSPA0000005777 ;
  crm:P70_is_documented_in gres:record_TESTSPA0000005777 .

gres:id_TESTSPA0000005777 a crm:E42_Identifier ; rdfs:label "TESTSPA0000005777" .

```

```

gres:type_metatarsal
  a crm:E55_Type, skos:Concept ;
  skos:prefLabel "Metatarsal"@en ;
  skos:exactMatch <http://purl.obolibrary.org/obo/UBERON_0003651> ,
    <http://purl.obolibrary.org/obo/UBERON_0003652> ,
    <http://purl.obolibrary.org/obo/UBERON_0003653> ,
    <http://purl.obolibrary.org/obo/UBERON_0013588> .

```

```

gres:taxon_ovis
  a crm:E55_Type, skos:Concept ;
  skos:prefLabel "Ovis"@en ;
  skos:exactMatch <http://eol.org/pages/39510> .

```

gres:sex_male

a crm:E55_Type, skos:Concept ;
skos:prefLabel "male organism"@en ;
skos:exactMatch <http://purl.obolibrary.org/obo/UBERON_0003101> .

gres:exc_TESTSPA0000005777

a crm:E7_Activity, crma:A9_Archaeological_Excavation, gerasim:ExcavationEvent ;
crm:P7_took_place_at gres:place_TESTSPA0000003444 ;
crm:P4_has_time-span gres:timespan_excavation_1994 ;
crm:P14_carried_out_by gres:actor_denise_carruthers .

gres:place_TESTSPA0000003444

a crm:E53_Place, geo:Feature ;
rdfs:label "Turkey / Pınarbaşı / Site B / Stratum 5 / Turkey / Pınarbaşı / Site B / Stratum 6 / Context BBH"@en ;
crm:P1_is_identified_by gres:id_place_TESTSPA0000003444 ;
geo:hasGeometry gres:geom_TESTSPA0000003444 .

gres:id_place_TESTSPA0000003444 a crm:E42_Identifier ; rdfs:label "TESTSPA0000003444" .

gres:geom_TESTSPA0000003444

a geo:Geometry ;
geo:asWKT "POINT(33.018551 37.49432)"^^geo:wktLiteral .

gres:strat_BBH

a crma:A8_Stratigraphic_Unit, gerasim:StratigraphicContext ;
rdfs:label "Context BBH"@en .

gres:dim_weight_TESTSPA0000005777

a crm:E54_Dimension ;
crm:P2_has_type gres:weight_type ;
crm:P90_has_value "12.8"^^xsd:decimal .

gres:record_TESTSPA0000005777

a crm:E31_Document, prov:Entity, gerasim:DatasetRecord ;
dcterms:source <<http://opencontext.org/projects/TESTPRJ0000000004>> ;
dcterms:creator gres:actor_denise_carruthers ;
rdfs:seeAlso <<http://opencontext.org/subjects/TESTSPA0000005777>> .

A.5 Example B2

#####

Example B2 — event-centric OpenContext instance

#####

<<http://opencontext.org/subjects/TESTSPA0000003836>>

a gerasim:Artefact ;
rdfs:label "Bone Ref# 2440"@en ;
crm:P1_is_identified_by gres:id_TESTSPA0000003836 ;
crm:P2_has_type gres:type_middle_phalanx , gres:taxon_canis_lupus ;
gerasim:associatedWithPeriod gres:period_-6700_-6000 ;
crm:P12i_was_present_at gres:exc_TESTSPA0000003836 ;
gerasim:hasStratigraphicContext gres:strat_BAZ ;
crm:P70_is_documented_in gres:record_TESTSPA0000003836 .

gres:id_TESTSPA0000003836 a crm:E42_Identifier ; rdfs:label "TESTSPA0000003836" .

gres:type_middle_phalanx

a crm:E55_Type, skos:Concept ;
skos:prefLabel "middle phalanx"@en ;
skos:exactMatch <http://purl.obolibrary.org/obo/UBERON_0004301> .

gres:taxon_canis_lupus

a crm:E55_Type, skos:Concept ;
skos:prefLabel "Canis lupus Linnaeus, 1758"@en ;
skos:exactMatch <<http://col.org/pages/328607>> .

gres:exc_TESTSPA0000003836

a crm:E7_Activity, crma:A9_Archaeological_Excavation, gerasim:ExcavationEvent ;
crm:P7_took_place_at gres:place_TESTSPA0000003554 ;
crm:P4_has_time-span gres:timespan_excavation_1994 ;
crm:P14_carried_out_by gres:actor_denise_carruthers .

gres:place_TESTSPA0000003554

a crm:E53_Place, geo:Feature ;
rdfs:label "Turkey / Pınarbaşı / Site B / Stratum 5 / Context BAZ"@en ;
crm:P1_is_identified_by gres:id_place_TESTSPA0000003554 ;
geo:hasGeometry gres:geom_TESTSPA0000003554 .

gres:id_place_TESTSPA0000003554 a crm:E42_Identifier ; rdfs:label "TESTSPA0000003554" .

gres:geom_TESTSPA0000003554

a geo:Geometry ;

geo:asWKT "POINT(33.018551 37.49432)"^^geo:wktLiteral .

gres:strat_BAZ

a crma:A8_Stratigraphic_Unit, gerasim:StratigraphicContext ;

rdfs:label "Context BAZ"@en .

gres:record_TESTSPA0000003836

a crm:E31_Document, prov:Entity, gerasim:DatasetRecord ;

dcterms:source <http://opencontext.org/projects/TESTPRJ0000000004> ;

dcterms:creator gres:actor_denise_carruthers ;

rdfs:seeAlso <http://opencontext.org/subjects/TESTSPA0000003836> .

A.6 Example B3

#####

Example B3 — event-centric OpenContext instance

#####

<http://opencontext.org/subjects/TESTSPA0000005212>

a gerasim:Artefact ;

rdfs:label "Bone Ref# 2553"@en ;

crm:P1_is_identified_by gres:id_TESTSPA0000005212 ;

crm:P2_has_type gres:type_bone , gres:type_not_identifiable ;

gerasim:associatedWithPeriod gres:period_-6700_-6000 ;

crm:P12i_was_present_at gres:exc_TESTSPA0000005212 ;

gerasim:hasStratigraphicContext gres:strat_BBI ;

crm:P43_has_dimension gres:dim_weight_TESTSPA0000005212 ;

crm:P70_is_documented_in gres:record_TESTSPA0000005212 .

gres:id_TESTSPA0000005212 a crm:E42_Identifier ; rdfs:label "TESTSPA0000005212" .

gres:type_bone

a crm:E55_Type, skos:Concept ;

skos:prefLabel "Bone"@en .

gres:type_not_identifiable

a crm:E55_Type, skos:Concept ;
skos:prefLabel "Not Identifiable"@en .

gres:exc_TESTSPA0000005212

a crm:E7_Activity, crma:A9_Archaeological_Excavation, gerasim:ExcavationEvent ;
crm:P7_took_place_at gres:place_TESTSPA0000003456 ;
crm:P4_has_time-span gres:timespan_excavation_1994 ;
crm:P14_carried_out_by gres:actor_denise_carruthers .

gres:place_TESTSPA0000003456

a crm:E53_Place, geo:Feature ;
rdfs:label "Turkey / Pınarbaşı / Site B / Stratum 6 / Context BBI"@en ;
crm:P1_is_identified_by gres:id_place_TESTSPA0000003456 ;
geo:hasGeometry gres:geom_TESTSPA0000003456 .

gres:id_place_TESTSPA0000003456 a crm:E42_Identifier ; rdfs:label "TESTSPA0000003456" .

gres:geom_TESTSPA0000003456

a geo:Geometry ;
geo:asWKT "POINT(33.018551 37.49432)"^^geo:wktLiteral .

gres:strat_BBI

a crma:A8_Stratigraphic_Unit, gerasim:StratigraphicContext ;
rdfs:label "Context BBI"@en .

gres:dim_weight_TESTSPA0000005212

a crm:E54_Dimension ;
crm:P2_has_type gres:weight_type ;
crm:P90_has_value "26.3"^^xsd:decimal .

gres:record_TESTSPA0000005212

a crm:E31_Document, prov:Entity, gerasim:DatasetRecord ;
dcterms:source <<http://opencontext.org/projects/TESTPRJ0000000004>> ;
dcterms:creator gres:actor_denise_carruthers ;
rdfs:seeAlso <<http://opencontext.org/subjects/TESTSPA0000005212>> .

A.7 Enriched demonstrator super-example

#####

Enriched demonstrator / event, documentation, production,

and digital twin layers combined in one illustrative example

#####

gres:artefact_demo_bone_1160_enriched

a gerasim:Artefact, hdto:HC3_Tangible_Heritage_Entity;
rdfs:label "Enriched demonstrator for Bone Ref# 1160"@en ;
crm:P1_is_identified_by gres:id_demo_bone_1160_enriched ;
crm:P2_has_type gres:type_metatarsal , gres:taxon_ovis , gres:sex_male ;
crm:P45_consists_of gres:material_bone ;
crm:P108i_was_produced_by gres:production_demo_bone_1160 ;
crm:P12i_was_present_at gres:exc_demo_bone_1160 ;
gerasim:associatedWithPeriod gres:period_-6700_-6000 ;
gerasim:hasStratigraphicContext gres:strat_BBH ;
hdto:HP1_has_digital_twin gres:digital_twin_demo_bone_1160 ;
crm:P43_has_dimension gres:dim_weight_demo_bone_1160 ;
crm:P70_is_documented_in gres:record_demo_bone_1160 .

gres:id_demo_bone_1160_enriched

a crm:E42_Identifier ;
rdfs:label "DEMO-BONE-1160-ENRICHED"@en .

gres:material_bone

a crm:E57_Material, skos:Concept ;
skos:prefLabel "bone"@en .

gres:production_demo_bone_1160

a crm:E12_Production ;
rdfs:label "Hypothetical production event for enriched demonstrator artefact"@en ;
crm:P4_has_time-span gres:timespan_production_demo_bone_1160 ;
crm:P7_took_place_at gres:place_production_region_demo ;
crm:P14_carried_out_by gres:actor_unknown_maker_demo .

gres:timespan_production_demo_bone_1160

a crm:E52_Time-Span ;
rdfs:label "Undetermined prehistoric production interval"@en ;

rdfs:comment "No precise begin/end is asserted because the demonstrator does not provide a reliable production date for the archaeological remain."@en .

gres:place_production_region_demo

a crm:E53_Place ;

rdfs:label "Undetermined production region"@en .

gres:actor_unknown_maker_demo

a crm:E39_Actor ;

rdfs:label "Unknown maker"@en .

gres:exc_demo_bone_1160

a crm:E7_Activity, crma:A9_Archaeological_Excavation, gerasim:ExcavationEvent ;

rdfs:label "Excavation event associated with Bone Ref# 1160 demonstrator"@en ;

crm:P7_took_place_at gres:place_TESTSPA0000003444 ;

crm:P4_has_time-span gres:timespan_excavation_1994 ;

crm:P14_carried_out_by gres:actor_denise_carruthers .

gres:dim_weight_demo_bone_1160

a crm:E54_Dimension ;

crm:P2_has_type gres:weight_type ;

crm:P90_has_value "12.8"^^xsd:decimal .

gres:record_demo_bone_1160

a crm:E31_Document, prov:Entity, gerasim:DatasetRecord ;

rdfs:label "Source record for enriched demonstrator based on Bone Ref# 1160"@en ;

dcterms:source <<http://opencontext.org/projects/TESTPRJ00000000004>> ;

rdfs:seeAlso <<http://opencontext.org/subjects/TESTSPA0000005777>> .

gres:doc_creation_demo_bone_1160

a crm:E65_Creation ;

rdfs:label "Documentation creation event for enriched demonstrator"@en ;

crm:P14_carried_out_by gres:actor_denise_carruthers ;

crm:P4_has_time-span gres:timespan_doc_demo_bone_1160 ;

crm:P94_has_created gres:photo_demo_bone_1160 , gres:fieldnote_demo_bone_1160 ,
gres:digital_twin_demo_bone_1160 ;

prov:wasDerivedFrom <<http://opencontext.org/projects/TESTPRJ00000000004>> .

gres:timespan_doc_demo_bone_1160

a crm:E52_Time-Span ;

```

rdfs:label "Documentation event timespan for enriched demonstrator"@en ;

rdfs:comment "No precise begin/end is asserted because the demonstrator only models the existence of a documentation
event, not its exact date."@en .

gres:photo_demo_bone_1160
a crm:E36_Visual_Item ;
rdfs:label "Illustrative photograph for enriched demonstrator"@en ;
crm:P138_represents gres:artefact_demo_bone_1160_enriched ;
crm:P94i_was_created_by gres:doc_creation_demo_bone_1160 .

gres:fieldnote_demo_bone_1160
a crm:E31_Document ;
rdfs:label "Illustrative field note for enriched demonstrator"@en ;
crm:P67_refers_to gres:artefact_demo_bone_1160_enriched ;
crm:P94i_was_created_by gres:doc_creation_demo_bone_1160 .

gres:digital_twin_demo_bone_1160
a hdto:HC2_Heritage_Digital_Twin ;
rdfs:label "Illustrative digital twin for enriched demonstrator"@en ;
crm:P67_refers_to gres:artefact_demo_bone_1160_enriched ;
crm:P138_represents gres:artefact_demo_bone_1160_enriched ;
crm:P94i_was_created_by gres:doc_creation_demo_bone_1160 .

```

Παράρτημα Β: Ερωτήματα SPARQL

Τα παρακάτω SPARQL ερωτήματα έχουν σχεδιαστεί ώστε να εκτελούνται πάνω στο τελικό TTL αρχείο της οντολογίας και των ενδεικτικών instances. Μπορούν να εκτελεστούν σε SPARQL περιβάλλον που υποστηρίζει φόρτωση RDF/Turtle αρχείων, όπως ενδεικτικά το Protégé SPARQL Query tab ή εναλλακτικά κάποιο triple store / query interface, π.χ. Apache Jena Fuseki ή GraphDB. Στην παρούσα εργασία τα ερωτήματα χρησιμοποιούνται ως μέσο διερεύνησης και ελέγχου του Γνωσιακού Γράφου σε επίπεδο proof of concept.

Για την εκτέλεση των ερωτημάτων χρησιμοποιείται το ακόλουθο prefix block. Σε περίπτωση που κάθε ερώτημα εκτελείται ανεξάρτητα σε SPARQL playground ή triple store query editor, το prefix block πρέπει να προηγείται του αντίστοιχου SELECT query.

```

PREFIX gerasim: <https://example.org/gerasim/ontology/>
PREFIX gres: <https://example.org/gerasim/resource/>
PREFIX crm: <http://www.cidoc-crm.org/cidoc-crm/>
PREFIX crma: <http://www.cidoc-crm.org/extensions/crmarchaeo/>

```

PREFIX hdto: <http://isl.ics.forth.gr/ontology/echoes/>
 PREFIX geo: <http://www.opengis.net/ont/geosparql#>
 PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
 PREFIX prov: <http://www.w3.org/ns/prov#>
 PREFIX dcterms: <http://purl.org/dc/terms/>

B.1 CQ1

```
SELECT ?artefact ?identifierLabel ?artefactLabel
WHERE {
  ?artefact a gerasim:Artefact ;
    crm:P1_is_identified_by ?identifier ;
    rdfs:label ?artefactLabel .
  ?identifier rdfs:label ?identifierLabel .
}
ORDER BY ?artefact
```

B.2 CQ2

```
SELECT ?artefact ?artefactLabel ?typeLabel ?materialLabel
WHERE {
  ?artefact a gerasim:Artefact ;
    rdfs:label ?artefactLabel ;
    crm:P2_has_type ?type .
  ?type skos:prefLabel ?typeLabel .
  OPTIONAL {
    ?artefact crm:P45_consists_of ?material .
    ?material skos:prefLabel ?materialLabel .
  }
}
ORDER BY ?artefactLabel
```

B.3 CQ3

```
SELECT ?artefact ?artefactLabel ?period ?periodLabel ?begin ?end
```

```

WHERE {
    ?artefact a gerasim:Artefact ;
        rdfs:label ?artefactLabel ;
        gerasim:associatedWithPeriod ?period .
    ?period skos:prefLabel ?periodLabel .
    OPTIONAL {
        ?period crm:P4_has_time-span ?timespan .
        OPTIONAL { ?timespan crm:P82a_begin_of_the_begin ?begin . }
        OPTIONAL { ?timespan crm:P82b_end_of_the_end ?end . }
    }
}
ORDER BY ?periodLabel ?artefactLabel

```

Αν υπάρχουν και ιεραρχικές σχέσεις περιόδων:

```

SELECT ?artefactLabel ?periodLabel ?broaderPeriodLabel
WHERE {
    ?artefact a gerasim:Artefact ;
        rdfs:label ?artefactLabel ;
        gerasim:associatedWithPeriod ?period .
    ?period skos:prefLabel ?periodLabel .
    OPTIONAL {
        ?period skos:broader ?broaderPeriod .
        ?broaderPeriod skos:prefLabel ?broaderPeriodLabel .
    }
}
ORDER BY ?artefactLabel

```

B.4 CQ4

```

SELECT ?artefactLabel ?placeLabel ?wkt
WHERE {
    ?artefact a gerasim:Artefact ;
        rdfs:label ?artefactLabel ;
        crm:P53_has_former_or_current_location ?place .
    ?place rdfs:label ?placeLabel ;
        geo:hasGeometry ?geometry .
}

```

```

?geometry geo:asWKT ?wkt .
}
ORDER BY ?artefactLabel

```

Για excavation events:

```

SELECT ?excavation ?placeLabel ?wkt
WHERE {
  ?excavation a gerasim:ExcavationEvent ;
    crm:P7_took_place_at ?place .
  ?place rdfs:label ?placeLabel ;
    geo:hasGeometry ?geometry .
  ?geometry geo:asWKT ?wkt .
}
ORDER BY ?excavation

```

B.5 CQ5

```

SELECT ?excavation ?artefact ?artefactLabel
WHERE {
  ?artefact a gerasim:Artefact ;
    rdfs:label ?artefactLabel ;
    crm:P12i_was_present_at ?excavation .
  ?excavation a gerasim:ExcavationEvent .
}
ORDER BY ?excavation ?artefactLabel

```

Για συγκεκριμένη ανασκαφή:

```

SELECT ?artefact ?artefactLabel
WHERE {
  ?artefact a gerasim:Artefact ;
    rdfs:label ?artefactLabel ;
    crm:P12i_was_present_at gres:exc_TESTSPA0000005777 .
}
ORDER BY ?artefactLabel

```

B.6 CQ6

```
SELECT ?artefact ?artefactLabel ?contextLabel
WHERE {
    ?artefact a gerasim:Artefact ;
        rdfs:label ?artefactLabel ;
        gerasim:hasStratigraphicContext ?context .
    ?context rdfs:label ?contextLabel .
}
ORDER BY ?contextLabel ?artefactLabel
```

B.7 CQ7

```
SELECT ?artefactLabel ?productionEvent ?actorLabel ?timespanLabel
WHERE {
    ?artefact a gerasim:Artefact ;
        rdfs:label ?artefactLabel ;
        crm:P108i_was_produced_by ?productionEvent .
    OPTIONAL {
        ?productionEvent crm:P14_carried_out_by ?actor .
        ?actor rdfs:label ?actorLabel .
    }
    OPTIONAL {
        ?productionEvent crm:P4_has_time-span ?timespan .
        ?timespan rdfs:label ?timespanLabel .
    }
}
ORDER BY ?artefactLabel
```

B.8 CQ8

```
SELECT DISTINCT ?artefactLabel ?record ?docEvent ?actorLabel ?timespanLabel
WHERE {
    ?artefact a gerasim:Artefact ;
        rdfs:label ?artefactLabel ;
        crm:P70_is_documented_in ?record .
```

```

OPTIONAL {
    ?digitalTwin a hdto:HC2_Heritage_Digital_Twin ;
        crm:P67_refers_to ?artefact ;
        crm:P94i_was_created_by ?docEvent .
OPTIONAL {
    ?docEvent crm:P14_carried_out_by ?actor .
    ?actor rdfs:label ?actorLabel .
}
OPTIONAL {
    ?docEvent crm:P4_has_time-span ?timespan .
    ?timespan rdfs:label ?timespanLabel .
}
}
}
ORDER BY ?artefactLabel

```

B.9 CQ9

```

SELECT ?actorLabel ?excavation ?artefactLabel
WHERE {
    ?excavation a gerasim:ExcavationEvent ;
        crm:P14_carried_out_by ?actor .
    ?actor rdfs:label ?actorLabel .
    ?artefact a gerasim:Artefact ;
        rdfs:label ?artefactLabel ;
        crm:P12i_was_present_at ?excavation .
}
ORDER BY ?actorLabel ?excavation ?artefactLabel

```

Αν θέλουμε μόνο actors που εμφανίζονται σε παραπάνω από μία ανασκαφές:

```

SELECT ?actorLabel (COUNT(DISTINCT ?excavation) AS ?excavationCount)
WHERE {
    ?excavation a gerasim:ExcavationEvent ;
        crm:P14_carried_out_by ?actor .
    ?actor rdfs:label ?actorLabel .
}
GROUP BY ?actorLabel

```

```
HAVING (COUNT(DISTINCT ?excavation) > 1)
ORDER BY DESC(?excavationCount)
```

B.10 CQ10

```
SELECT ?periodLabel ?artefactLabel ?placeLabel
WHERE {
    ?artefact a gerasim:Artefact ;
        rdfs:label ?artefactLabel ;
        gerasim:associatedWithPeriod ?period ;
        crm:P53_has_former_or_current_location ?place .
    ?period skos:prefLabel ?periodLabel .
    ?place rdfs:label ?placeLabel .
}
ORDER BY ?periodLabel ?placeLabel ?artefactLabel
```

B.11 CQ11

```
SELECT ?artefactLabel ?digitalTwin ?digitalTwinLabel
WHERE {
    ?artefact a gerasim:Artefact ;
        rdfs:label ?artefactLabel ;
        hdto:HP1_has_digital_twin ?digitalTwin .
    ?digitalTwin a hdto:HC2_Heritage_Digital_Twin ;
        rdfs:label ?digitalTwinLabel .
}
ORDER BY ?artefactLabel
```

B.12 CQ12

```
SELECT ?artefactLabel ?digitalTwinLabel ?excavation ?placeLabel
WHERE {
    ?artefact a gerasim:Artefact ;
        rdfs:label ?artefactLabel ;
        hdto:HP1_has_digital_twin ?digitalTwin ;
        crm:P12i_was_present_at ?excavation .
    ?digitalTwin rdfs:label ?digitalTwinLabel .
```

```

?excavation crm:P7_took_place_at ?place .
?place rdfs:label ?placeLabel .
}
ORDER BY ?artefactLabel

```

B.13 CQ13

```

SELECT ?placeLabel ?periodLabel (COUNT(?artefact) AS ?artefactCount)
WHERE {
  ?artefact a gerasim:Artefact ;
    gerasim:associatedWithPeriod ?period ;
    crm:P53_has_former_or_current_location ?place .
  ?period skos:prefLabel ?periodLabel .
  ?place rdfs:label ?placeLabel .
}
GROUP BY ?placeLabel ?periodLabel
ORDER BY ?placeLabel ?periodLabel

```

B.14 CQ14

```

SELECT ?artefactLabel ?periodLabel ?placeLabel ?excavation ?digitalTwinLabel
WHERE {
  ?artefact a gerasim:Artefact ;
    rdfs:label ?artefactLabel .
  OPTIONAL {
    ?artefact gerasim:associatedWithPeriod ?period .
    ?period skos:prefLabel ?periodLabel .
  }
  OPTIONAL {
    ?artefact crm:P53_has_former_or_current_location ?place .
    ?place rdfs:label ?placeLabel .
  }
  OPTIONAL {
    ?artefact crm:P12i_was_present_at ?excavation .
  }
  OPTIONAL {
    ?artefact hdto:HP1_has_digital_twin ?digitalTwin .
  }
}

```

```

    ?digitalTwin rdfs:label ?digitalTwinLabel .
  }
}

ORDER BY ?artefactLabel

```

Παράρτημα Γ: Πλήρης αναπαράσταση της οντολογίας σε Turtle

Το παρόν παράρτημα περιλαμβάνει την τελική έκδοση του οντολογικού σχήματος της εργασίας σε μορφή Turtle (.ttl), μαζί με τους κοινόχρηστους helper πόρους που αξιοποιούνται από τα RDF παραδείγματα. Τα αναλυτικά ενδεικτικά παραδείγματα δεδομένων παρατίθενται ξεχωριστά στο Παράρτημα Α.

```

@prefix gerasim: <https://example.org/gerasim/ontology/> .
@prefix gres: <https://example.org/gerasim/resource/> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix crm: <http://www.cidoc-crm.org/cidoc-crm/> .
@prefix crma: <http://www.cidoc-crm.org/extensions/crmarchaeo/> .
@prefix geo: <http://www.opengis.net/ont/geosparql#> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix prov: <http://www.w3.org/ns/prov#> .
@prefix dcterms: <http://purl.org/dc/terms/> .
@prefix dcat: <http://www.w3.org/ns/dcat#> .
@prefix hdto: <http://isl.ics.forth.gr/ontology/echoes/> .

#####
# Ontology header
#####

<https://example.org/gerasim/ontology>
  a owl:Ontology ;
  rdfs:label "Gerasim ontology for cultural heritage data"@en ;
  dcterms:title "Gerasim ontology for cultural heritage data"@en ;
  dcterms:description "A revised ontology for cultural heritage data, aligned primarily with CIDOC CRM and complemented by CRMarchaeo, GeoSPARQL, SKOS, PROV-O and a limited HDTO-based digital representation layer."@en ;
  dcterms:creator "Gerasimos Karathanos"@en ;
  owl:imports <http://www.cidoc-crm.org/cidoc-crm/> ,

```

```

    <http://www.cidoc-crm.org/extensions/crmarchaeo/> ,
    <http://www.opengis.net/ont/geosparql#> ,
    <http://www.w3.org/2004/02/skos/core#> ,
    <http://www.w3.org/ns/prov#> ,
    <http://purl.org/dc/terms/> ,
    <http://www.w3.org/ns/dcat#> ,
    <http://isl.ics.forth.gr/ontology/echoes/> ;

owl:versionInfo "v14.2"@en .

#####

# Local classes retained after audit

#####

gerasim:Artefact
    a owl:Class ;
    rdfs:label "Artefact"@en ;
    rdfs:comment "Local application-level class for archaeological finds represented in the Gerasim model, including human-made objects and other physical archaeological remains where appropriate."@en ;
    rdfs:subClassOf crm:E18_Physical_Thing .

gerasim:ExcavationEvent
    a owl:Class ;
    rdfs:label "Excavation event"@en ;
    rdfs:comment "Local application-level class for archaeological excavation events."@en ;
    rdfs:subClassOf crma:A9_Archaeological_Excavation .

gerasim:StratigraphicContext
    a owl:Class ;
    rdfs:label "Stratigraphic context"@en ;
    rdfs:comment "Local application-level class for stratigraphic context / unit."@en ;
    rdfs:subClassOf crma:A8_Stratigraphic_Unit .

gerasim:DatasetRecord
    a owl:Class ;
    rdfs:label "Dataset record"@en ;
    rdfs:comment "Local class for source records derived from datasets used in the thesis."@en ;
    rdfs:subClassOf crm:E31_Document , prov:Entity .

```

gerasim:FindContext

```
a owl:Class ;  
rdfs:label "Find context"@en ;  
rdfs:comment "Local lightweight categorical context of discovery, mainly for AADA-style values."@en ;  
rdfs:subClassOf crm:E55_Type , skos:Concept .
```

gerasim:Period

```
a owl:Class ;  
rdfs:label "Period"@en ;  
rdfs:comment "Application-level period node used both as a cultural-historical period and as a controlled vocabulary concept for lightweight dataset alignment. This modelling choice is used to support alignment with external period authorities such as PeriodO."@en ;  
rdfs:subClassOf crm:E4_Period , skos:Concept .
```

gerasim:Municipality

```
a owl:Class ;  
rdfs:label "Municipality"@en ;  
rdfs:comment "Administrative place used where municipality-level information is available in the dataset."@en ;  
rdfs:subClassOf crm:E53_Place .
```

#####

Local object properties retained after audit

#####

gerasim:associatedWithPeriod

```
a owl:ObjectProperty ;  
rdfs:label "associated with period"@en ;  
rdfs:comment "Application-level shortcut linking a CIDOC CRM entity, such as an artefact, event or documentation-related resource, directly to an application-level period node."@en ;  
rdfs:domain crm:E1_CRM_Entity ;  
rdfs:range gerasim:Period .
```

gerasim:hasFindContext

```
a owl:ObjectProperty ;  
rdfs:subPropertyOf crm:P2_has_type ;  
rdfs:label "has find context"@en ;  
rdfs:comment "Application-level shortcut linking an artefact to a lightweight find-context category."@en ;  
rdfs:domain gerasim:Artefact ;
```

```

rdfs:range gerasim:FindContext .

gerasim:hasStratigraphicContext
  a owl:ObjectProperty ;
  rdfs:label "has stratigraphic context"@en ;
  rdfs:comment "Application-level shortcut linking an artefact directly to a stratigraphic context."@en ;
  rdfs:domain gerasim:Artefact ;
  rdfs:range gerasim:StratigraphicContext .

gerasim:locatedInMunicipality
  a owl:ObjectProperty ;
  rdfs:label "located in municipality"@en ;
  rdfs:comment "Links a place to a municipality where municipality information is present in the source data."@en ;
  rdfs:domain crm:E53_Place ;
  rdfs:range gerasim:Municipality .

#####
# Local datatype properties retained after audit
#####

gerasim:hasKKJCoordinate
  a owl:DatatypeProperty ;
  rdfs:label "has KKJ coordinate"@en ;
  rdfs:comment "Stores Finnish KKJ / EPSG:2393 coordinate values as a single literal in the form \"x y\", preserving the
original dataset coordinate pair when a full GeoSPARQL geometry is not sufficient or when the original KKJ values need to
be retained."@en ;
  rdfs:domain crm:E53_Place ;
  rdfs:range xsd:string .

#####
# External classes referenced directly in the model
#####

hdto:HC2_Heritage_Digital_Twin a owl:Class .
hdto:HC3_Tangible_Heritage_Entity a owl:Class .
hdto:HP1_has_digital_twin a owl:ObjectProperty .

```

#####

Shared helper nodes

#####

gres:mm a crm:E58_Measurement_Unit ; rdfs:label "millimetre"@en .

gres:length a crm:E55_Type ; rdfs:label "Length (mm)"@en .

gres:width a crm:E55_Type ; rdfs:label "Width (mm)"@en .

gres:fragment_length a crm:E55_Type ; rdfs:label "Fragment length (mm)"@en .

gres:fragment_width a crm:E55_Type ; rdfs:label "Fragment width (mm)"@en .

gres:fragment_thickness a crm:E55_Type ; rdfs:label "Fragment thickness (mm)"@en .

gres:weight_type a crm:E55_Type ; rdfs:label "Weight"@en .

#####

Shared OpenContext helper nodes

#####

gres:actor_denise_carruthers

a crm:E39_Actor ;

rdfs:label "Denise Carruthers"@en .

gres:period_-6700_-6000

a crm:E4_Period, skos:Concept, gerasim:Period ;

skos:prefLabel "-6700 to -6000 BCE/CE"@en ;

crm:P4_has_time-span gres:timespan_-6700_-6000 .

gres:timespan_-6700_-6000

a crm:E52_Time-Span ;

rdfs:label "Timespan from -6700 to -6000 BCE/CE"@en ;

crm:P82a_begin_of_the_begin "-6700"^^xsd:integer ;

crm:P82b_end_of_the_end "-6000"^^xsd:integer .

gres:timespan_excavation_1994

a crm:E52_Time-Span ;

rdfs:label "Excavation in 1994"@en ;

crm:P82a_begin_of_the_begin "1994-01-01"^^xsd:date ;

crm:P82b_end_of_the_end "1994-12-31"^^xsd:date .