



Μοντέλα διεργασιών και ροών εργασίας

Παραδοτέο Π2 / Ενότητα Εργασίας 2

Εμβληματική Δράση: «Το αναδυόμενο τοπίο ψηφιακής εργασίας στις Ανθρωπιστικές Επιστήμες στο πλαίσιο των ευρωπαϊκών υποδομών DARIAH και CLARIN» (Αριθμός Έργου:7982)

Τελική έκδοση

Έγγραφο:	7982_Π2-Μοντέλα διεργασιών και ροών εργασίας
Τίτλος παραδοτέου:	Π2: Μοντέλα διεργασιών και ροών εργασίας / D2: Process and workflow models
Ενότητα εργασίας:	ΕΕ2: Διεργασίες ψηφιακής εργασίας στις Ανθρωπιστικές και Κοινωνικές Επιστήμες / WP2: Digital Work Processes in the Humanities and Social Sciences
Υπεύθυνος φορέας:	Ε.Κ. ΑΘΗΝΑ
Συμμετέχων φορέας:	Ακαδημία Αθηνών (ΑΑ)
Συντάκτες:	Βίκη Ντρίτσου (Ε.Κ. Αθηνά / ΙΠΣΥ)
Με τη συμβολή των:	Βουκούτης Λέων (Ε.Κ. Αθηνά/ΙΕΛ), Γκίρτζου Κατερίνα (Ε.Κ. Αθηνά/ΙΕΛ), Γκούμας Δημήτρης (Ε.Κ. Αθηνά/ΙΕΛ), Γουλή Ελένη (ΑΑ), Ηλβανίδου Μαρία (Ε.Κ. Αθηνά/ΙΠΣΥ), Καλαφατά Πατρίτσια (ΑΑ), Κωνσταντόπουλος Πάνος (Ε.Κ. Αθηνά/ΙΠΣΥ), Μπουγά Σοφία (ΑΑ), Σπηλιωτοπούλου Μαρία (ΑΑ), Σουγιουλτζόγλου Ηράκλειτος (ΑΑ), Τζεδόπουλος Γιώργος (ΑΑ)
Έκδοση εγγράφου:	Τελική
Ημερομηνία έκδοσης:	Απρίλιος 2024

ΠΕΡΙΕΧΟΜΕΝΑ

Περιεχόμενα	2
Περίληψη.....	3
1. Εισαγωγή	4
2. Οντολογικό πλαίσιο.....	5
2.1 Οντολογία αναφοράς: Scholarly Ontology.....	6
2.2 Εξειδίκευση της Οντολογίας Αναφοράς.....	8
2.2.1 Τύπος Δραστηριότητας	8
2.2.2 Στάδια έρευνας.....	14
3. Ερευνητικές προκλήσεις και διεργασίες	16
4. Γράφος γνώσης.....	18
5. Μοντέλο ροής εργασίας	20
Βιβλιογραφία	22

ΠΕΡΙΛΗΨΗ

Στην παρούσα αναφορά περιγράφονται τα αποτελέσματα των εργασιών που εκτελέστηκαν στο πλαίσιο των Εργασιών 2.1 και 2.2. της Ενότητας Εργασίας 2 (ΕΕ2) του έργου. Πιο συγκεκριμένα, εστιάζουμε στην οντολογική πλαισίωση των ψηφιακών πρακτικών εργασίας στις Ανθρωπιστικές και Κοινωνικές επιστήμες και εισάγουμε μια σημασιολογική αναπαράσταση αυτών. Για την πλαισίωση τους χρησιμοποιήσαμε μια οντολογία αναφοράς, σύμφωνα με την οποία αναπαραστήσαμε την πληροφορία που συλλέχθηκε στην ΕΕ1, την οποία και κωδικοποιήσαμε στη συνέχεια σε γράφο γνώσης. Ταυτόχρονα μελετήσαμε και αναπαραστήσαμε σημασιολογικά την εκτέλεση των πρακτικών εργασίας ως ροή εργασιών. Αποτέλεσμα της μελέτης που πραγματοποιήσαμε στην εκτέλεσης των εργασιών ως ροές αποτέλεσε η δημιουργία ενός μοντέλου ροής εργασίας, το οποίο αποτυπώσαμε σε μορφή διαγράμματος BPMN, και αποτελεί μια πρότυπη ροή εργασίας βασισμένη επίσης στην οντολογίας αναφοράς που υιοθετήσαμε.

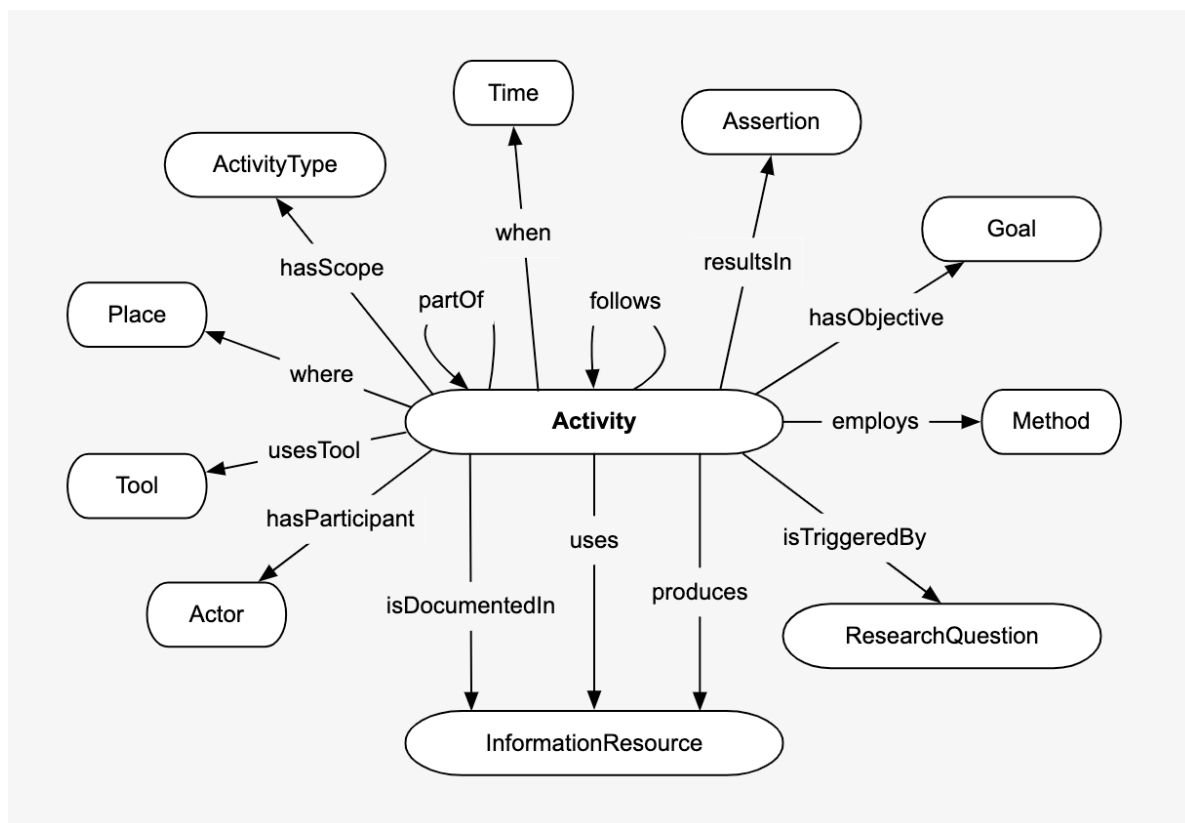
1. ΕΙΣΑΓΩΓΗ

Κύριο στόχο της έρευνάς μας στην ΕΕ2 αποτέλεσε η ανάπτυξη σημασιολογικής αναπαράστασης για την πληροφορία που αφορά στις ψηφιακές πρακτικές στις ανθρωπιστικές και κοινωνικές επιστήμες. Για την αναπαράσταση αυτή βασικό μας μέλημα ήταν τόσο η αποτύπωση της δομής των ψηφιακών πρακτικών, αναδεικνύοντας τα ιδιαίτερα χαρακτηριστικά τους και τις μεταξύ τους σχέσεις, όσο και η εκτέλεσή τους ως ροή εργασίας. Για το λόγο αυτό στραφήκαμε αρχικά στην αποτύπωση της δομής τους, χρησιμοποιώντας μια οντολογία αναφοράς για τη σημασιολογική πλαισίωση του εγχειρήματος. Στη συνέχεια, οι πρακτικές που αναδείχθηκαν μελετήθηκαν ως προς τη διαδικαστική εκτέλεσή τους, με σκοπό να αναπαρασταθούν σε ένα μοντέλο ροής εργασίας, ικανό να τυποποιήσει και να συστηματοποιήσει την εκτέλεση τέτοιων ροών. Τα συνολικά αποτελέσματα της ΕΕ2, τόσο αυτά που περιέχονται στην παρούσα αναφορά όσο και εκείνα του παραδοτέου «Π3: Οδηγός προτύπων και μεταδεδομένων», έχουν γίνει δεκτά για δημοσίευση ως αναρτημένη παρουσίαση (poster) στην Ετήσια Συνάντηση του DARIAH 2024 που θα πραγματοποιηθεί στη Λισαβόνα (18-21 Ιουνίου 2024).

2. ΟΝΤΟΛΟΓΙΚΟ ΠΛΑΙΣΙΟ

Η εργασία T2.1 έχει ως κύριο σκοπό την ανάπτυξη μιας σημασιολογικής αναπαράστασης για τις ψηφιακές πρακτικές εργασίας στον τομέα των Ανθρωπιστικών και Κοινωνικών Επιστημών, όπως αυτές περιγράφονται στην ΕΕ1. Με σκοπό την ανάδειξη αυτών των πρακτικών εργασίας, επιλέξαμε να κωδικοποιήσουμε την πλούσια πληροφορία που προέκυψε από το ερωτηματολόγιο και τις ομάδες εστίασης της ΕΕ1 σε έναν γράφο γνώσης. Ο γράφος γνώσης εξυπηρετεί στη σημασιολογική αναπαράσταση της συλλεγόμενης πληροφορίας και στην υποστήριξη και εκτέλεση πολύπλοκων ερωτήσεων που ανακτούν και αναδεικνύουν σημαντικά χαρακτηριστικά του τομέα. Αυτό επιτυγχάνεται μέσω μιας τυπικής σημασιολογικής αναπαράστασης με μορφή γράφου, όπου οι οντότητες συμβολίζονται ως κόμβοι και οι συσχετίσεις μεταξύ αυτών ως ακμές. Επιπλέον, πέραν της ρητής γνώσης που αναπαρίσταται, ένας γράφος γνώσης επιτρέπει την εξαγωγή λογικών συμπερασμάτων και, επομένως, την παραγωγή νέας γνώσης.

Η δημιουργία ενός γράφου γνώσης βασίζεται σε ένα σαφώς ορισμένο σημασιολογικό σχήμα, μια σαφή οντολογία, η οποία ορίζει τις έννοιες και τις σχέσεις μεταξύ αυτών. Το σχήμα αναφοράς ή η οντολογία αναφοράς μας δίνει έτσι το πλαίσιο για την κατανόηση και την ερμηνεία των δεδομένων του πεδίου. Για την οντολογική πλαισίωση και τη δημιουργία του γράφου γνώσης του παρόντος έργου, στηριχθήκαμε στην οντολογία αναφοράς Scholarly Ontology (SO) [1], η οποία περιγράφεται περιληπτικά παρακάτω.



Εικ. 1. Scholarly Ontology - Σημασιολογική αναπαράσταση της Δραστηριότητας (Activity)

2.1 ΟΝΤΟΛΟΓΙΑ ΑΝΑΦΟΡΑΣ: SCHOLARLY ONTOLOGY

Με σκοπό την οντολογική πλαισίωση του ψηφιακών πρακτικών εργασίας που αναδεικνύονται από την έρευνα της ΕΕ1, υιοθετήσαμε την οντολογία Scholarly Ontology ως οντολογία αναφοράς. Πρόκειται για μια οντολογία που σκοπό έχει την παράσταση της επιστημονικής εργασίας, βασισμένη στη μοντελοποίηση επιχειρησιακών διαδικασιών και η οποία αναπτύχθηκε μελετώντας εμπειρικές έρευνες και προγενέστερα μοντέλα. Η αρθρωτή δομή της επιτρέπει την παράσταση των βασικών εννοιών του πεδίου μέσα από τέσσερις όψεις της επιστημονικής εργασίας: Δραστηριότητα, Διαδικασία, Πόρος και Δράστης.

Κεντρική στην οντολογία αυτή είναι η οντότητα Δραστηριότητα (Activity), η οποία συνδέεται σημασιολογικά με Δράστες και Εργαλεία, Τόπους και Χρόνους στους οποίους λαμβάνει χώρα η δραστηριότητα αυτή, Σκοπούς που εξυπηρετεί, Μεθόδους που χρησιμοποιεί, Δηλώσεις και Ερευνητικά Ερωτήματα που παράγει/απαντά, καθώς επίσης και με τους Πληροφοριακούς Πόρους που χρησιμοποιεί, παράγει ή τεκμηριώνεται σε αυτούς. Η χρήση της συσχέτισης ακολουθεί (follows) επιτρέπει τη δήλωση ακολουθίας δραστηριοτήτων, ενώ σύμπλοκες δραστηριότητες σχετίζονται με τα συνιστώντα μέρη τους με τη μερωνυμική σχέση μέρος-του (partOf). Η σχηματική απεικόνιση της σημασιολογικής αναπαράστασης της οντότητας Δραστηριότητα δίνεται στην Εικ. 1.

- ☐ 1. Acquiring
 - 1.1. Collecting
 - 1.2. Gathering
 - 1.3. Learning
- ☐ 2. Communicating
 - ☒ 2.1. Collaborating
 - ☒ 2.2. Disseminating
- ☐ 3. Conceiving
 - 3.1. Concept Formation
 - 3.2. Hypothesis Formulation
 - 3.3. Problem Recognition
 - 3.4. Research Question / Topic Identification
- ☐ 4. Processing
 - ☒ 4.1. Analyzing
 - ☒ 4.2. Modifying
 - ☒ 4.3. Organizing
 - ☒ 4.4. Preserving
 - ☒ 4.5. Producing
- ☐ 5. Seeking
 - 5.1. Associative Searching
 - 5.2. Browsing
 - 5.3. Data Mining
 - 5.4. Direct Accessing
 - 5.5. Filtering
 - 5.6. Locating
 - 5.7. Retrieving
 - 5.8. Tracking

Εικ. 2. Activity Types - Ανώτερα επίπεδα ιεραρχικής δομής τύπων δραστηριοτήτων

Κάθε Δραστηριότητα (Activity) κατατάσσεται σε (hasScope) ένα Τύπο Δραστηριότητας (Activity Type). Οι τύποι αυτοί δραστηριότητας αποτελούν ιεραρχικά οργανωμένους όρους, δίνοντας τη φύση των δραστηριοτήτων, ενώ ταυτόχρονα λειτουργούν ως ευρετήριο δραστηριοτήτων [2]. Τα ανώτερα επίπεδα της ιεραρχικής αυτής οργάνωσης φαίνονται στην Εικ. 2.

2.2 ΕΞΕΙΔΙΚΕΥΣΗ ΤΗΣ ΟΝΤΟΛΟΓΙΑΣ ΑΝΑΦΟΡΑΣ

Κατόπιν συστηματικής μελέτης των στοιχείων του Ερωτηματολογίου και των Ομάδων Εστίασης της ΕΕ1 και έχοντας ως σημείο αναφοράς την οντολογία SO, αξιοποιήσαμε το μοντέλο της τελευταίας για τη σημασιολογική αναπαράσταση της πληροφορίας που συλλέχθηκε. Έχοντας ως αφετηρία το μοντέλο της SO, αποφασίσαμε να αναπτύξουμε μια εξειδίκευση αυτής, ικανής να αναπαραστήσει την υπό εξέταση πληροφορία με πληρότητα και σαφήνεια. Οι εξειδικεύσεις που προέκυψαν αφορούν σε ένα τμήμα της SO, το οποίο παρουσιάζεται αναλυτικά παρακάτω.

Ειδικότερα, ιδιαίτερη προσοχή δόθηκε στην ορθή αποτύπωση των δραστηριοτήτων που οι ερωτώμενοι δήλωσαν ότι αξιοποιούν/χρησιμοποιούν ψηφιακές μεθόδους για την εκτέλεσή τους. Μάλιστα, μέσω του ερωτηματολογίου κλήθηκαν να επιλέξουν για ποιες από τις παρακάτω 5 κατηγορίες δραστηριοτήτων χρησιμοποιούν ψηφιακές μεθόδους:

1. Ανακάλυψη, συλλογή και δημιουργία δεδομένων,
2. Δημοσίευση, διαμοιρασμός ή κοινοποίηση δεδομένων,
3. Επεξεργασία, ανάλυση ή οπτικοποίηση δεδομένων,
4. Επισημείωση, εμπλουτισμός ή επιμέλεια δεδομένων,
5. Οργάνωση, δόμηση ή διαχείριση δεδομένων.
- 6.

Παρατηρούμε ότι οι παραπάνω επιλογές αποτελούν σύνολα από διαφορετικές δραστηριότητες και αποτελούν εξειδικεύσεις της οντότητας *Τύπος Δραστηριότητας*. Ως εκ τούτου μπορούν να αναπαρασταθούν αξιοποιώντας και εξειδικεύοντας όπου κρίνεται αναγκαίο την ιεραρχική δομή του *Τύπου Δραστηριότητας* που παρουσιάστηκε στην Εικ. 2.

2.2.1 Τύπος Δραστηριότητας

Όπως ήδη αναφέρθηκε παραπάνω, οι πέντε επιλογές που δόθηκαν στους συμμετέχοντες του ερωτηματολογίου, σχετικά με τις δραστηριότητες για τις οποίες χρησιμοποιούν ψηφιακές μεθόδους, αποτελούν σύνολα από διαφορετικές δραστηριότητες που εκφράστηκαν ομαδοποιημένα κατά τη σύνταξη του ερωτηματολογίου προς υποβοήθηση του χρήστη. Για την ορθή αναπαράσταση αυτών μέσω εξειδίκευσης των υπάρχοντων τύπων αξιοποιούμε τη συσχέτιση *isA*, η οποία αναπαριστά τη σημασιολογική σχέση από μία στενότερη έννοια σε μία ευρύτερη, διατηρώντας την κληρονομικότητα των χαρακτηριστικών και των σχέσεων της ευρύτερης έννοιας προς τη στενότερη.

Αναλύοντας περαιτέρω τα πέντε σύνολα δραστηριοτήτων, παρατηρούμε ότι ορισμένα εξ αυτών χρησιμοποιούν διάζευξη (ή) στη διατύπωσή τους, ενώ ένα εξ αυτών χρησιμοποιεί σύζευξη (και). Προς αποφυγή παρερμηνείας, να σημειώσουμε εδώ ότι οι διατυπώσεις αυτές δεν ερμηνεύονται αυστηρά ως αποκλειστικές διαζεύξεις ή συζεύξεις μαθηματικής λογικής. Αντιθέτως, θεωρούνται γλωσσολογικές εκφράσεις που χρησιμοποιούνται για να διατυπώσουν βασικές κατηγορίες δραστηριοτήτων και για το λόγο αυτό ερμηνεύονται στο σύνολό τους στο εννοιολογικό μοντέλο με χρήση λογικής διάζευξης. Η μοντελοποίηση βασισμένη στην SO για καθένα από τα πέντε αυτά σύνολα δραστηριοτήτων δίνεται παρακάτω. Για διευκόλυνση του αναγνώστη, έννοιες των οποίων τα ονόματα ξεκινούν με το πρόθεμα “SO” προέρχονται από το μοντέλο αναφοράς, ενώ όλες

οι υπόλοιπες αποτελούν προσθήκες/εξειδικεύσεις της παρούσας εργασίας. Η ανάλυση των πέντε δοθαισών επιλογών δραστηριοτήτων με χρήση του Τύπου Δραστηριότητας και οι παραγόμενες εξειδικεύσεις αυτού δίνονται αμέσως παρακάτω.

1. Ανακάλυψη, συλλογή και δημιουργία δεδομένων.

Αποτελεί σύνολο τύπων δραστηριοτήτων που αναλύεται στους παρακάτω επιμέρους τύπους.

1.1 Ανακάλυψη δεδομένων:

Η ανακάλυψη των δεδομένων αποτελεί τύπο δραστηριότητας που δεν αποτυπώνεται πλήρως στη SO και για τον λόγο αυτό απαιτείται εξειδίκευση προϋπάρχουσας οντότητας. Για την ακρίβεια, η οντότητα *SO:Seeking* δηλώνει τη δραστηριότητα της προσπάθειας εύρεσης, χωρίς όμως να διατυπώνει την επιτυχή έκβαση της εύρεσης αυτής που υποδηλώνεται μέσω της Ανακάλυψης. Επομένως, η Ανακάλυψη (*Discovering*) μοντελοποιείται ως εξειδίκευση της *SO:Seeking*, όπως φαίνεται σχηματικά στην Εικ. 3.

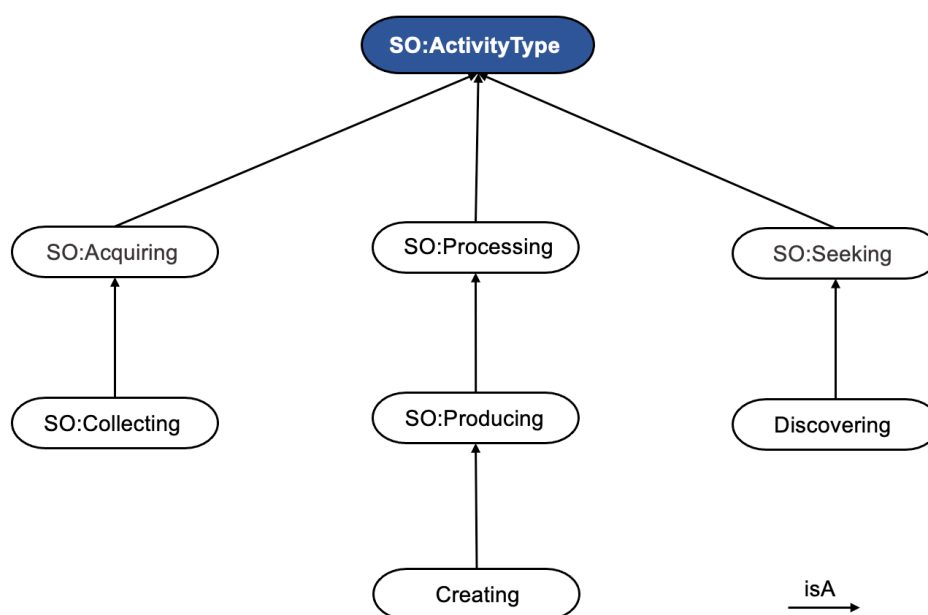
1.2 Συλλογή δεδομένων:

Η συλλογή δεδομένων αποτελεί τύπο δραστηριότητας που διατυπώνεται σαφώς στην SO μέσω της οντότητας *SO:Collecting*, η οποία αναπαριστά τη δραστηριότητα συστηματικής και μεθοδικής απόκτησης και συγκέντρωσης αντικειμένων. Επομένως, δεν απαιτείται δημιουργία νέας οντότητας.

1.3 Δημιουργία δεδομένων:

Η δημιουργία των δεδομένων αποτελεί τύπο δραστηριότητας που δεν αποτυπώνεται πλήρως στη SO και για τον λόγο αυτό απαιτείται εξειδίκευση προϋπάρχουσας οντότητας. Για την ακρίβεια, η οντότητα *SO:Producing* δηλώνει την παραγωγή ή κατασκευή προϊόντων/αντικειμένων, στην οποία εμπεριέχεται και η έννοια της δημιουργίας. Επομένως, η Δημιουργία (*Creating*) μοντελοποιείται ως εξειδίκευση της *SO:Producing*, όπως φαίνεται σχηματικά στην Εικ. 3.

Το σύνολο των δραστηριοτήτων που περιέχονται στην τρέχουσα περίπτωση αποτυπώνονται γραφικά από το μέρος της ιεραρχικής δομής που εμφανίζεται στην Εικ. 3.



Εικ. 3. Μοντελοποίηση δραστηριοτήτων περίπτωσης 1

2. Δημοσίευση, διαμοιρασμός ή κοινοποίηση δεδομένων.

Αποτελεί σύνολο τύπων δραστηριοτήτων που αναλύεται στους παρακάτω επιμέρους τύπους.

2.1 Δημοσίευση δεδομένων:

Η δημοσίευση δεδομένων αποτελεί τύπο δραστηριότητας που διατυπώνεται σαφώς στην SO μέσω της οντότητας *SO:Publishing*, η οποία αναπαριστά την επίσημη διάθεση οποιουδήποτε είδους αντικειμένου στο ευρύτερο κοινό. Επομένως, δεν απαιτείται δημιουργία νέας οντότητας.

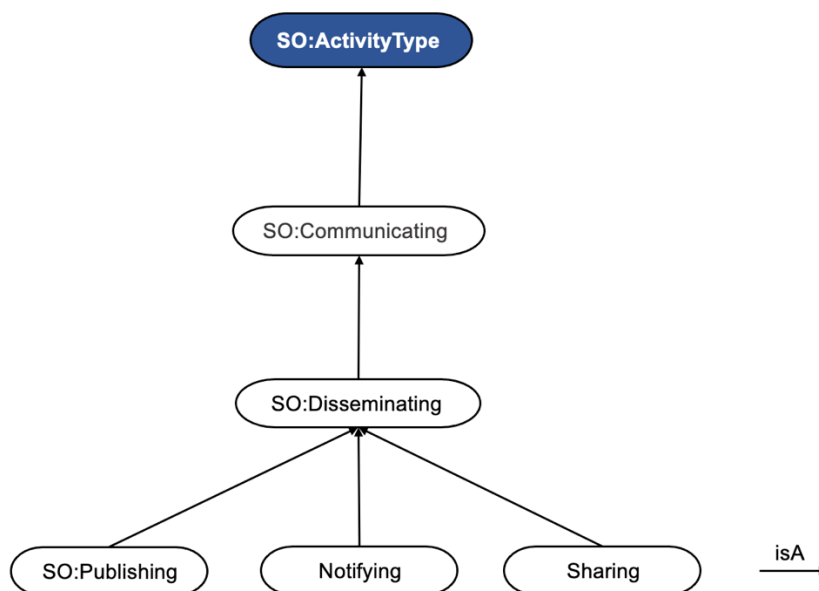
2.2 Διαμοιρασμός δεδομένων:

Ο διαμοιρασμός των δεδομένων αποτελεί τύπο δραστηριότητας που δεν αποτυπώνεται πλήρως στη SO και για τον λόγο αυτό απαιτείται εξειδίκευση προϋπάρχουσας οντότητας. Για την ακρίβεια, η οντότητα *SO:Disseminating* δηλώνει τον τύπο δραστηριότητας της διάθεσης και διάχυσης αντικειμένων, στην οποία εμπεριέχεται και η έννοια του διαμοιρασμού. Επομένως, ο Διαμοιρασμός (*Sharing*) μοντελοποιείται ως εξειδίκευση της *SO:Disseminating*, όπως φαίνεται σχηματικά στην Εικ. 4.

2.3 Κοινοποίηση δεδομένων:

Ομοίως με παραπάνω, η κοινοποίηση των δεδομένων αποτελεί τύπο δραστηριότητας που δεν αποτυπώνεται πλήρως στη SO και για τον λόγο αυτό απαιτείται εξειδίκευση προϋπάρχουσας οντότητας. Πράγματι, η Κοινοποίηση (*Notifying*) μοντελοποιείται ως εξειδίκευση της *SO:Disseminating*, όπως φαίνεται σχηματικά στην Εικ. 4.

Το σύνολο των δραστηριοτήτων που περιέχονται στην τρέχουσα περίπτωση αποτυπώνονται γραφικά από το μέρος της ιεραρχικής δομής που εμφανίζεται στην Εικ. 4.



Εικ. 4. Μοντελοποίηση δραστηριοτήτων περίπτωσης 2

3. Επεξεργασία, ανάλυση ή οπτικοποίηση δεδομένων.

Αποτελεί σύνολο τύπων δραστηριοτήτων που αναλύεται στους παρακάτω επιμέρους τύπους.

3.1 Επεξεργασία δεδομένων:

Η επεξεργασία δεδομένων αποτελεί τύπο δραστηριότητας που διατυπώνεται σαφώς στην SO μέσω της οντότητας *SO:Processing*, η οποία αναπαριστά την εκτέλεση μιας σειράς ενεργειών πάνω σε αντικείμενα με σκοπό την επίτευξη κάποιου αποτελέσματος. Επομένως, δεν απαιτείται δημιουργία νέας οντότητας.

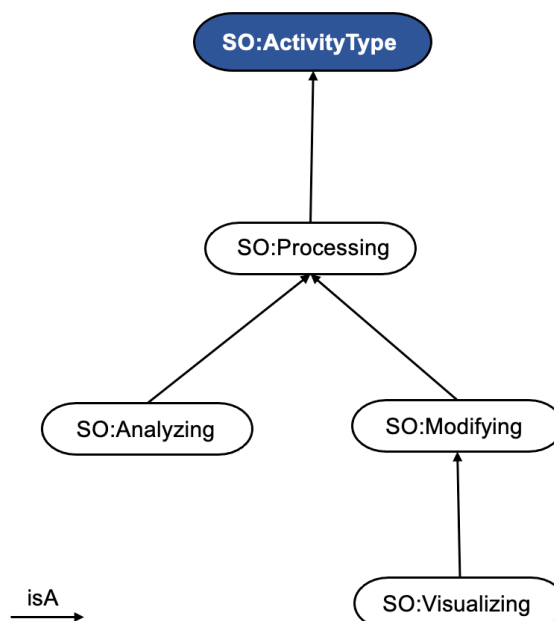
3.2 Ανάλυση δεδομένων:

Η ανάλυση δεδομένων αποτελεί τύπο δραστηριότητας που διατυπώνεται σαφώς στην SO μέσω της οντότητας *SO:Analyzing*, η οποία αναπαριστά τον τύπο δραστηριότητας κατά τον οποίο πραγματοποιείται εξαγωγή οποιουδήποτε είδους πληροφορίας από δεδομένα. Επομένως, δεν απαιτείται δημιουργία νέας οντότητας.

3.3 Οπτικοποίηση δεδομένων:

Η οπτικοποίηση δεδομένων αποτελεί τύπο δραστηριότητας που διατυπώνεται σαφώς στην SO μέσω της οντότητας *SO:Visualizing*, η οποία αναπαριστά τον τύπο δραστηριότητας σύνοψης και παρουσίασης σε γραφική μορφή. Επομένως, δεν απαιτείται δημιουργία νέας οντότητας.

Το σύνολο των δραστηριοτήτων που περιέχονται στην τρέχουσα περίπτωση αποτυπώνονται γραφικά από το μέρος της ιεραρχικής δομής που εμφανίζεται στην Εικ. 5.



Εικ. 5. Μοντελοποίηση δραστηριοτήτων περίπτωσης 3

4. Επισημείωση, εμπλουτισμός ή επιμέλεια δεδομένων.

Αποτελεί σύνολο τύπων δραστηριοτήτων που αναλύεται στους παρακάτω επιμέρους τύπους.

4.1 Επισημείωση δεδομένων:

Η επισημείωση δεδομένων αποτελεί τύπο δραστηριότητας που διατυπώνεται σαφώς στην SO μέσω της οντότητας *SO:Annotating*, η οποία αναπαριστά τον τύπο δραστηριότητας προσθήκης πληροφορίας σχετικής με ορισμένα στοιχεία ενός αντικειμένου. Επομένως, δεν απαιτείται δημιουργία νέας οντότητας.

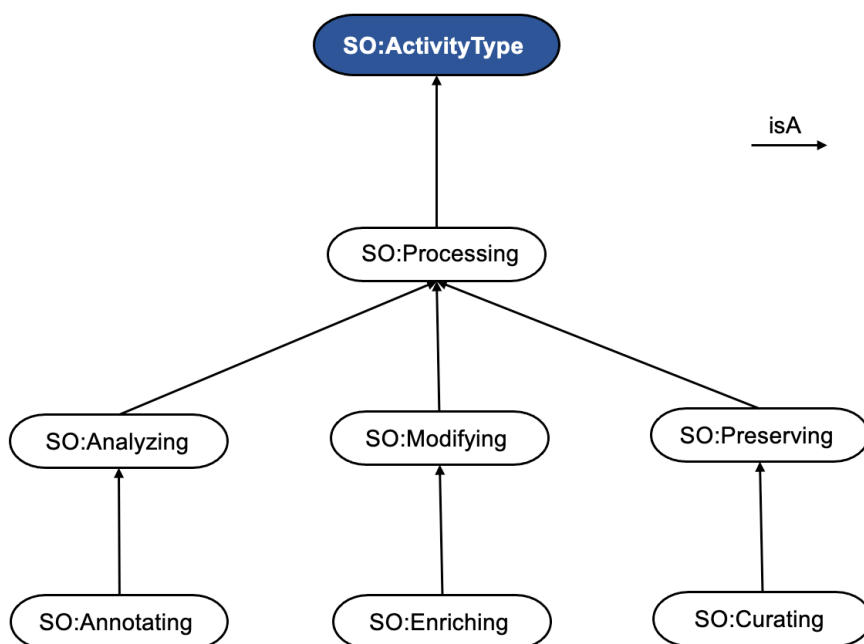
4.2 Εμπλουτισμός δεδομένων:

Ο εμπλουτισμός δεδομένων αποτελεί τύπο δραστηριότητας που διατυπώνεται σαφώς στην SO μέσω της οντότητας *SO:Enriching*, η οποία αναπαριστά τον τύπο δραστηριότητας κατά την οποία προστίθεται πληροφορία σε ένα αντικείμενο, διατυπώνοντας σαφώς την προέλευση, τη φύση, τη δομή ή τα στοιχεία του. Επομένως, δεν απαιτείται δημιουργία νέας οντότητας.

4.3 Επιμέλεια δεδομένων:

Η επιμέλεια δεδομένων αποτελεί τύπο δραστηριότητας που διατυπώνεται σαφώς στην SO μέσω της οντότητας *SO:Curating*, η οποία αναπαριστά τον τύπο δραστηριότητας της επιλογής, της οργάνωσης και της φροντίδας συγκεκριμένων αντικειμένων, συνήθως με τη χρήση επαγγελματικών ή ειδικών γνώσεων. Επομένως, δεν απαιτείται δημιουργία νέας οντότητας.

Το σύνολο των δραστηριοτήτων που περιέχονται στην τρέχουσα περίπτωση αποτυπώνονται γραφικά από το μέρος της ιεραρχικής δομής που εμφανίζεται στην Εικ. 6.



Εικ. 6. Μοντελοποίηση δραστηριοτήτων περίπτωσης 4

5. Οργάνωση, δόμηση ή διαχείριση δεδομένων.

Αποτελεί σύνολο τύπων δραστηριοτήτων που αναλύεται στους παρακάτω επιμέρους τύπους.

5.1 Οργάνωση δεδομένων:

Η οργάνωση δεδομένων αποτελεί τύπο δραστηριότητας που διατυπώνεται σαφώς στην SO μέσω της οντότητας `SO:Organizing` και η οποία αναπαριστά τον τύπο δραστηριότητας της διάταξης/τακτοποίησης των αντικειμένων, με τρόπο που διευκολύνει άλλες δραστηριότητες. Επομένως, δεν απαιτείται δημιουργία νέας οντότητας.

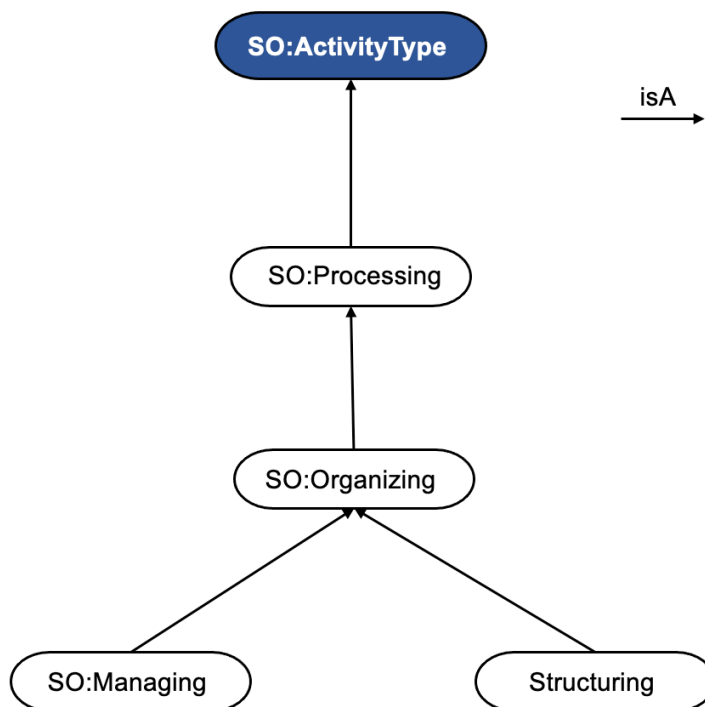
5.2 Δόμηση δεδομένων:

Η δόμηση των δεδομένων αποτελεί τύπο δραστηριότητας που δεν αποτυπώνεται πλήρως στη SO και για τον λόγο αυτό απαιτείται εξειδίκευση προϋπάρχουσας οντότητας. Για την ακρίβεια, η οντότητα `SO:Organizing` αναπαριστά τον τύπο δραστηριότητας της διάταξης/τακτοποίησης των αντικειμένων όπως είδαμε και παραπάνω, όμως η έννοια της Δόμησης παραπέμπει σε πιο στενή έννοια διάταξης με βάση συγκεκριμένο μοτίβο ή δομή δεδομένων. Για τον λόγο αυτό δημιουργείται η εξειδίκευση `Structuring` ως εξειδίκευση της `SO:Organizing`, όπως παριστάνεται σχηματικά στην Εικ. 7.

5.3 Διαχείριση δεδομένων:

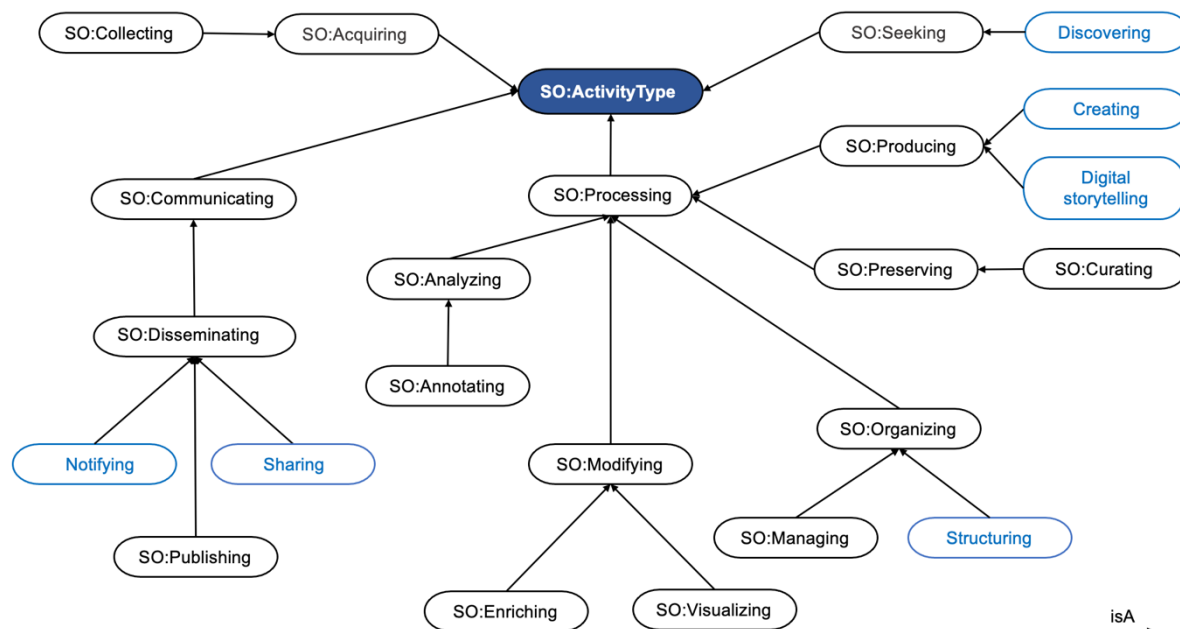
Η διαχείριση δεδομένων αποτελεί τύπο δραστηριότητας που διατυπώνεται σαφώς στην SO μέσω της οντότητας `SO:Managing` και η οποία αναπαριστά τον τύπο δραστηριότητας της οργάνωσης, του συντονισμού, της παρακολούθησης και της προσαρμογής εργασιών και πόρων. Επομένως, δεν απαιτείται δημιουργία νέας οντότητας.

Το σύνολο των δραστηριοτήτων που περιέχονται στην τρέχουσα περίπτωση αποτυπώνονται γραφικά από το μέρος της ιεραρχικής δομής που εμφανίζεται στην Εικ. 7.



Εικ. 7. Μοντελοποίηση δραστηριοτήτων περίπτωσης 5

Το σύνολο των απαιτούμενων τύπων δραστηριοτήτων για την αποτύπωση της υφιστάμενης πληροφορίας (τόσο εκείνες που προέρχονται από την SO όσο και οι εξειδικεύσεις αυτών που παρήχθησαν από την παραπάνω ανάλυση) δίνεται σχηματικά στην Εικ. 8.



Εικ. 8. Οντολογικό πλαίσιο αναπαράστασης κατηγοριών δραστηριοτήτων με αξιοποίηση των Τύπων Δραστηριοτήτων της SO και εξειδικεύσεις αυτών

2.2.2 Στάδια έρευνας

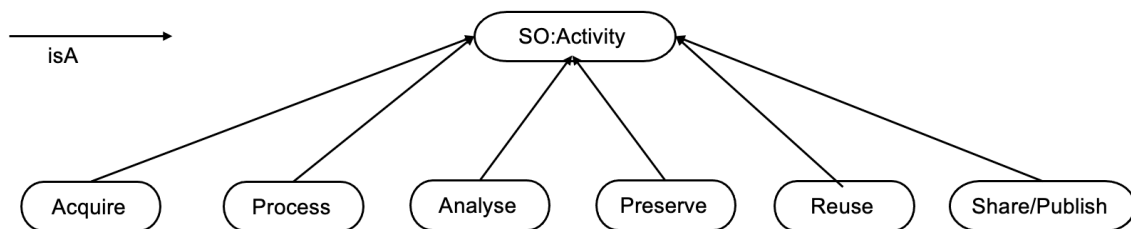
Για την ανάλυση των αποτελεσμάτων του ερωτηματολογίου της ΕΕ1, πραγματοποιήθηκε επίσης μελέτη των σταδίων του κύκλου έρευνας/εργασίας που αξιοποιήθηκαν. Παρά το γεγονός ότι η ανάδειξη των σταδίων πραγματοποιήθηκε μετά τη συμπλήρωση του ερωτηματολογίου και συγκεκριμένα κατά τη φάση της ανάλυσης των αποτελεσμάτων, εντούτοις περιέχει σημαντική πληροφορία επιτρέποντας την αναγωγή των καταγεγραμμένων δραστηριοτήτων σε ομάδες αυτών με κοινά σημασιολογικά χαρακτηριστικά. Τα έξι στάδια που αναδείχθηκαν είναι τα εξής:

1. Αποκτώ (Acquire)
2. Επεξεργάζομαι (Process)
3. Αναλύω (Analyse)
4. Διατηρώ (Preserve)
5. Επαναχρησιμοποιώ (Reuse)
6. Διαμοιράζομαι/Δημοσιεύω (Share/Publish)

Καθένα από τα παραπάνω στάδια αποτελεί μια γενίκευση ενός συνόλου από εξειδικευμένες δραστηριότητες/ενέργειες, όπως αυτές αποτυπώθηκαν στις αποκρίσεις των συμμετεχόντων. Ως προς την SO που υιοθετούμε ως οντολογία αναφοράς στην παρούσα πλαισίωση, τα 6 αυτά στάδια

αποτελούν σημασιολογικά εξειδικεύσεις της οντότητας Δραστηριότητα (Activity) (βλέπε Εικ. 1).

Πράγματι, τα ερευνητικά αυτά στάδια ομαδοποιούν/γενικεύουν συγκεκριμένες Δραστηριότητες (Activities) που εκτελούνται από κάποιον Δράστη (Actor) αξιοποιώντας Μεθόδους (Methods) και Εργαλεία (Tool). Ταυτόχρονα αποτελούνται από επιμέρους δραστηριότητες (partOf), οι οποίες μπορούν να εκτελούνται σε ακολουθία (follows). Οδηγούμαστε έτσι στο συμπέρασμα ότι τα αναφερόμενα στάδια έρευνας αποτελούν εξειδικεύσεις της οντότητας Activity της SO. Η την αναπαράσταση της εξειδίκευσης αυτής δίνεται σχηματικά στην Εικ. 9.



Εικ. 9. Εξειδίκευση οντότητας Δραστηριότητα βάσει των σταδίων έρευνας

3. ΕΡΕΥΝΗΤΙΚΕΣ ΠΡΟΚΛΗΣΕΙΣ ΚΑΙ ΔΙΕΡΓΑΣΙΕΣ

Κατά τη διάρκεια του πρώτου έτους εκτέλεσης του έργου διεξήχθη ποιοτική έρευνα με τη μέθοδο των Ομάδων Εστίασης. Συγκεκριμένα, πραγματοποιήθηκαν έξι συναντήσεις ομάδων εστίασης, οι αναφορές των οποίων περιέχονται στο Παραδοτέο “D1: Trends in digital methods in the Humanities and Social Sciences”. Με σκοπό να καταγράψουμε και να αναλύσουμε τις κοινές προκλήσεις που αντιμετωπίζουν οι ερευνητές στο πεδίο των Ψηφιακών Ανθρωπιστικών Επιστημών, πραγματοποιήσαμε συγκριτική μελέτη των αποτελεσμάτων των 6 ομάδων εστίασης. Η παρακάτω λίστα περιέχει συμπεράσματα αλλά και προκλήσεις αντιμετωπίζουν οι ερευνητές κατά την εργασία τους, όπως καταγράφηκαν από τις υπό εξέταση ερευνητικές ομάδες, καθώς και προτεινόμενες ενέργειες προς αντιμετώπιση αυτών. Οι προτεινόμενες ενέργειες βασίζονται στην αξιοποίηση ψηφιακών εργαλείων προς βελτίωση της μεθοδολογίας έρευνας και των ερευνητικών αποτελεσμάτων.

1. Συλλογή και Διαχείριση Δεδομένων
Αναμφίβολα η συλλογή των δεδομένων διευκολύνεται με τη χρήση ψηφιακών εργαλείων. Η επιλογή των κατάλληλων όμως εργαλείων είναι πολύ σημαντική, ώστε να εξυπηρετούν τις ερευνητικές ανάγκες. Θέματα που θα πρέπει να εξετάζονται κατά την επιλογή των εργαλείων είναι το κατά πόσο εξυπηρετούν την αποτελεσματική οργάνωση, την αποθήκευση και την ανάκτηση των δεδομένων.
2. Ανάλυση και Επεξεργασία Δεδομένων
Λόγω του αυξημένου πλήθους ψηφιακών εργαλείων, πριν την επιλογή του εργαλείου θα πρέπει να εξετάζονται οι δυνατότητες αυτού και οι τύποι δεδομένων που υποστηρίζει. Το ψηφιακό εργαλείο θα πρέπει να συνεισφέρει στην επίτευξη των ερευνητικών στόχων. Σημαντικές διεργασίες για την επίτευξη αυτή αποτελούν η εκτέλεση ποιοτικής και ποσοτικής ανάλυσης των δεδομένων, ο σχολιασμός και η οπτικοποίηση.
3. Οργάνωση και Διαχείριση Δεδομένων
Η ετερογένεια των δεδομένων αποτελεί σημαντική πρόκληση για την αποτελεσματική οργάνωση των δεδομένων, επομένως απαιτείται μελέτη των δεδομένων και ανάπτυξη στρατηγικών για την οργάνωσή τους. Για την αντιμετώπιση της ετερογένειας και την επίτευξη διαλειτουργικότητας συνιστάται η υιοθέτηση προτύπων και καλών πρακτικών διαχείρισης δεδομένων.
4. Αξιοποίηση Χωρικής Αναπαράστασης
Σε περίπτωση ύπαρξης γεωχωρικών δεδομένων, αυτά θα πρέπει να αξιοποιούνται κατάλληλα με χρήση εργαλείων αναπαράστασης χωρικής πληροφορίας. Περαιτέρω ενίσχυση στην αποτύπωση και παρουσίαση των γεωχωρικών δεδομένων προσδίδει η υιοθέτηση τεχνικών ψηφιακής αφήγησης.
5. Συνεργασία και Διεπιστημονικότητα
Ιδιαίτερης σημασίας είναι η συμβολή των διεπιστημονικών συνεργασιών για την επίλυση των ερευνητικών ζητημάτων. Πρόκληση στις συνεργασίες αυτές αποτελεί η επίτευξη αποτελεσματικής επικοινωνίας και αμοιβαίας κατανόησης πάνω στα ερευνητικά ερωτήματα.

6. Υιοθέτηση κοινών προτύπων
Μέσω της υποστήριξης και της συμβολής στην ανάπτυξη κοινών προτύπων στις ψηφιακές υποδομές και της συνεργασίας με άλλους ερευνητές της περιοχής έρευνας, μπορούν να θεσπιστούν κατευθυντήριες γραμμές για την αναπαράσταση και τον διαμοιρασμό δεδομένων.
7. Επίδραση ψηφιακών μεθόδων στις Ανθρωπιστικές και Κοινωνικές Επιστήμες
Καθώς οι ψηφιακές μέθοδοι συνεχώς εξελίσσονται, η ενεργός συμμετοχή σε ερευνητικά έργα που περιέχουν την ψηφιακή συνιστώσα εξασφαλίζει τη δυνατότητα εξέλιξης των ερευνητών. Ταυτόχρονα, η ψηφιακότητα επιδρά και στην ίδια την αποτύπωση των ερευνητικών ερωτημάτων, επομένως οι ερευνητές καλούνται να αναγνωρίσουν και να αξιοποιήσουν κατάλληλα τις ψηφιακές μεθόδους.
8. Εκπαίδευση στις ψηφιακές μεθόδους
Σε άμεση συνάρτηση με την σημαντική επίδραση των ψηφιακών μεθόδων που αναφέρθηκε παραπάνω, προτεραιότητα θα πρέπει να δίνεται στην εκπαίδευση των ερευνητικών ομάδων στις ψηφιακές μεθόδους. Το επίπεδο χρήσης και αξιοποίησης ψηφιακών δεξιοτήτων αυξάνεται σημαντικά μέσω δράσεων εκπαίδευσης στα ψηφιακά εργαλεία, οι οποίες θα πρέπει να υποστηρίζονται.
9. Προκλήσεις ψηφιακών εργαλείων
Η χρήση των ψηφιακών εργαλείων στην ερευνητική εργασία, παρόλο καθόλα χρήσιμη, θα πρέπει ταυτόχρονα να αντιμετωπίζεται με κριτική σκέψη και προσοχή. Τόσο οι λειτουργικές απαιτήσεις των εργαλείων (π.χ. πλήθος αναγκαίων πόρων) όσο και το κόστος τους πρέπει να λαμβάνονται υπόψη κατά την επιλογή, ώστε η χρήση τους να αποτελεί βιώσιμη πρακτική για την ερευνητική εργασία. Συνιστάται η χρήση εργαλείων ανοιχτού κώδικα στο μέγιστο δυνατό βαθμό.
10. Επιπτώσεις της πανδημίας
Οι μεγάλες αλλαγές που έφερε η πανδημία ώθησε τους ερευνητές στην υιοθέτηση και αξιοποίηση συνεργατικών ψηφιακών πλατφορμών, τόσο για επικοινωνία όσο και για συνεργατική εργασία. Οι νέες αυτές ψηφιακές δεξιότητες και γνώσεις θα πρέπει να ενσωματωθούν στην ερευνητική εργασία και να αποτυπωθούν στις ερευνητικές ροές εργασίας.

4. ΓΡΑΦΟΣ ΓΝΩΣΗΣ

Τόσο η οντολογία αναφοράς όσο και τα δεδομένα που μοντελοποιήθηκαν βάσει αυτής αποφασίστηκε να κωδικοποιηθούν στην συνέχεια σε έναν γράφο γνώσης, ο οποίος επιτρέπει την εκτέλεση πολύπλοκων επερωτήσεων και την ανάκτηση πλούσιας πληροφορίας. Ο γράφος γνώσης αναπτύχθηκε με χρήση της ανοιχτού κώδικα βάσης γράφου Neo4j [3] (έκδοση Neo4j Server 4.0.9, έκδοση Neo4j Browser 4.2.0) και είναι προσβάσιμος εδώ (η πρόσβαση επιτρέπεται κατόπιν αυθεντικοποίησης):

<http://services.apollonis-infrastructure.gr:7474/browser/>

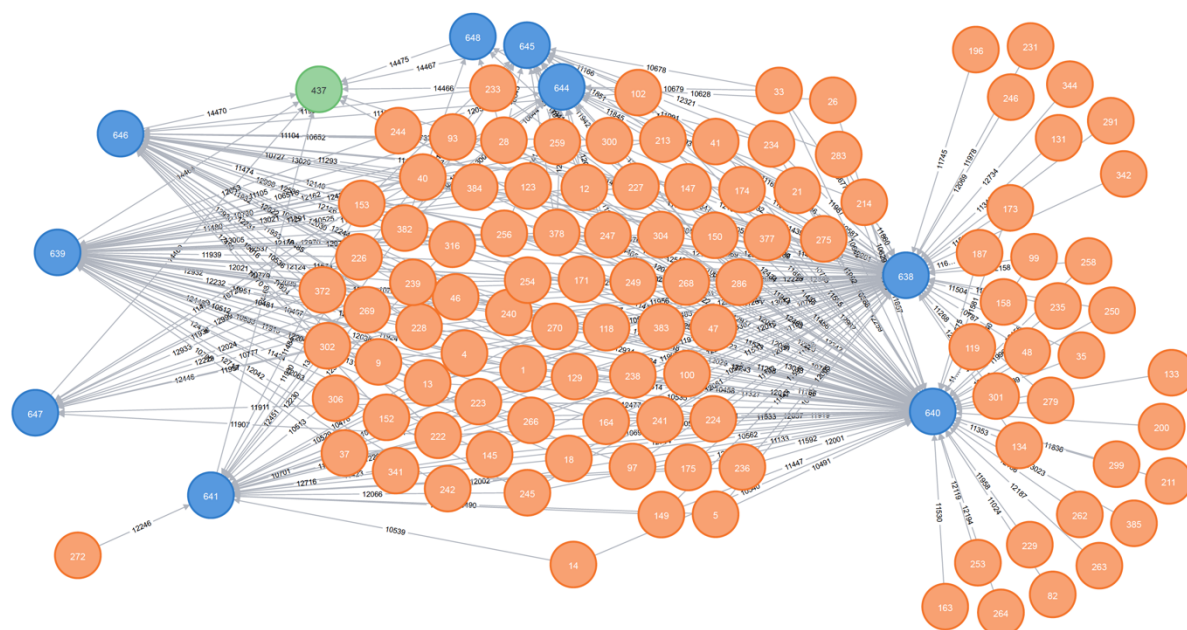
Το συνολικό μέγεθος του γράφου που αναπτύχθηκε περιέχει 1199 κόμβους, οι οποίοι είναι διασυνδεδεμένοι με 30808 ακμές.

Όπως αναφέρθηκε παραπάνω, ο γράφος γνώσης επιτρέπει την εκτέλεση επερωτήσεων μεγάλης πολυπλοκότητας, δίνοντας έτσι στους χρήστες τη δυνατότητα να ανακαλύψουν συσχετίσεις μεταξύ οντοτήτων που τους ενδιαφέρουν και να εξάγουν συμπεράσματα. Για την ανάκτηση αποτελεσμάτων από τον γράφο Neo4j είναι απαραίτητη η διατύπωση επερωτήσεων στη γλώσσα επερωτήσεων γράφου (graph query language) Cypher [4]. Λόγω της πολυπλοκότητας της γλώσσας και προς υποβοήθηση της ομάδας εργασίας του έργου, διατυπώθηκε ένα σύνολο επερωτήσεων με χρήση παραμέτρων, τις οποίες στη συνέχεια αξιοποίησε η ομάδα εργασίας της ΕΕ1 προς εξαγωγή συμπερασμάτων και ανάλυση τάσεων, όπως περιγράφεται στο Παραδοτέο Π1:Ανάλυση Τάσεων [5]. Παράδειγμα μιας τέτοιας επερώτησης αποτελεί η παρακάτω, στην οποία αναζητούμε πρόσωπα που έχουν δηλώσει ότι έχουν πολλαπλό επιστημονικό πεδίο, συμπεριλαμβάνοντας τα πεδία “ιστορία” και “προφορική ιστορία”:

```
MATCH (p:Person)-[:has_scientificField]->(r:ScientificField)
WITH p, COLLECT(r.label) as fields
WHERE ALL (x IN ["προφορική ιστορία","ιστορία"] WHERE x IN fields)
RETURN p
```

Τα αποτελέσματα των επερωτήσεων στον γράφο Neo4j μπορούν να επιστραφούν

- σε μορφή γράφου, όπως φαίνεται στην Εικ. 10, με δυνατότητα μεταφόρτωσης ως εικόνα σε μορφότυπο png και svg,
- σε μορφή πίνακα, με δυνατότητα μεταφόρτωσης σε μορφότυπο csv και json,
- σε μορφή κειμένου, με δυνατότητα μεταφόρτωσης σε μορφότυπο csv και json, και
- σε μορφή κώδικα, με δυνατότητα μεταφόρτωσης σε μορφότυπο csv και json.



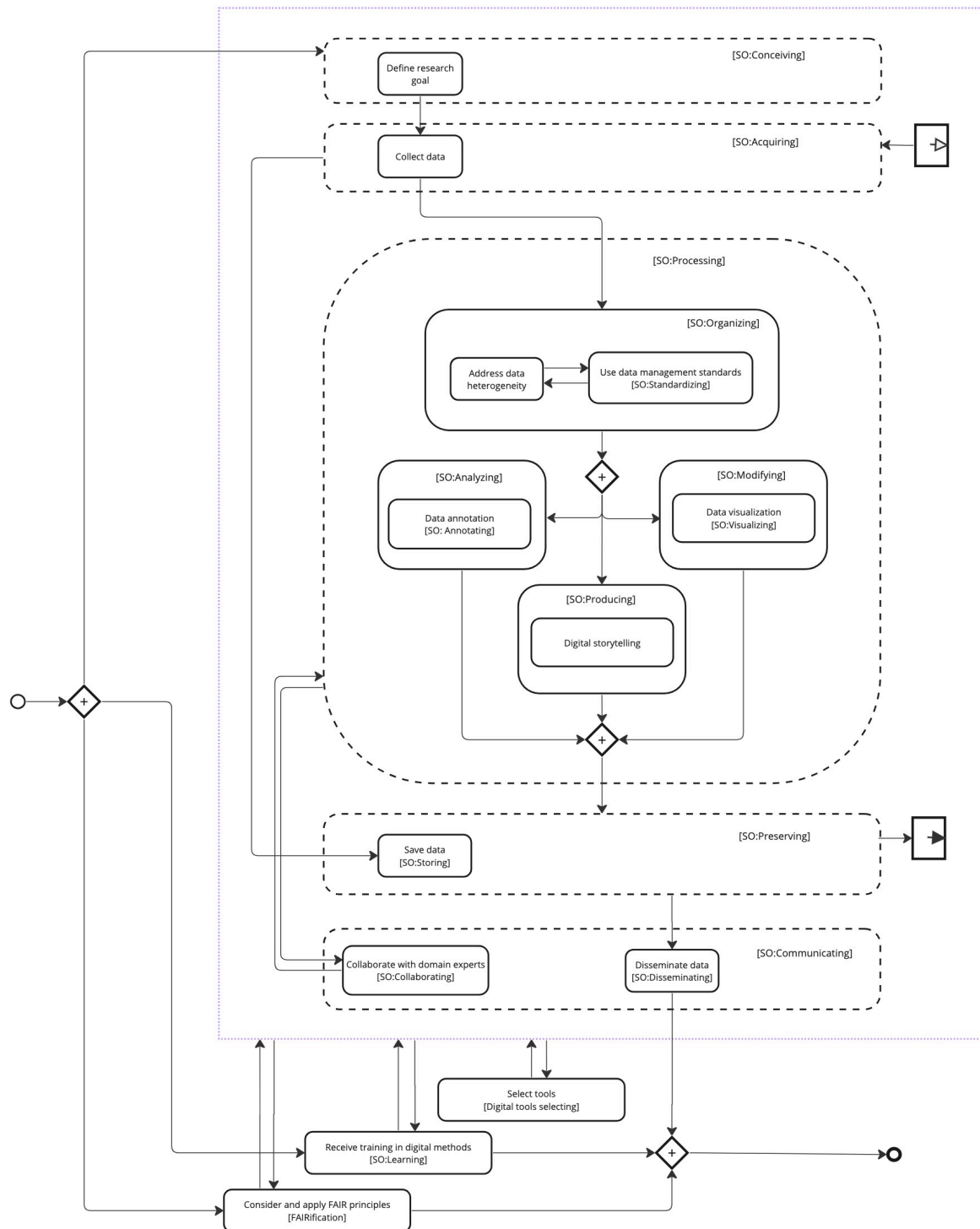
Εικ. 10. Αποτέλεσμα εκτέλεσης επερώτησης στον γράφο Neo4j σε μορφή γράφου

5. ΜΟΝΤΕΛΟ ΡΟΗΣ ΕΡΓΑΣΙΑΣ

Πέρα από τη δομική αναπαράσταση των πρακτικών εργασίας που αναδείχθηκαν στην ΕΕ1 και πλαισιώθηκαν οντολογικά στην προηγούμενη ενότητα, ιδιαίτερη σημασία δόθηκε επίσης στη διαδικαστική αναπαράσταση των πρακτικών αυτών μέσω ρών εργασίας. Για το λόγο αυτό στη συνέχεια της εργασίας μας εστιάσαμε σε μια βήμα προς βήμα αναπαράσταση της εκτέλεσης των ενεργειών (ή ακολουθιών ενεργειών), με σκοπό την ανάπτυξη ρών εργασίας αξιοποιήσιμων και επαναχρησιμοποιήσιμων από τους ερευνητές της περιοχής. Έτσι, εξυπηρετούν ως οδηγοί εφαρμογής για την εκτέλεση συχνών ενεργειών, όπως αυτές καταγράφηκαν στην έρευνά μας. Οι ροές εργασίας απαιτείται να είναι εξειδικεύσιμες και ικανές να αναπαραστήσουν τις πρακτικές που αναδείχθηκαν κατά την ανάλυση τάσεων, προκειμένου να παρέχουν συγκεκριμένες οδηγίες και να διευκολύνουν στην εφαρμογή τους.

Για την ανάπτυξη των ρών εργασίας επικεντρωθήκαμε στα δεδομένα της μελέτης από τις έξι ομάδες εστίασης της ΕΕ1. Αυτό συνέβη, διότι οι επιστημονικές συζητήσεις στις ομάδες εστίασης περιλάμβαναν τόσο προκλήσεις που αντιμετωπίζουν οι ερευνητές όσο και περιγραφές από τρόπους εργασίας και αντιμετώπισης των προκλήσεων. Οι περιγραφές αυτές εξυπηρετούν άριστα στην αναπαράσταση διαδικασιών μέσω ρών εργασίας. Μετά τη συγκριτική ανάλυση των αποτελεσμάτων των δεδομένων αυτών, όπως παρουσιάστηκε στην παραπάνω ενότητα, καταγράψαμε το σύνολο των ενεργειών που εκτελούνται εντοπίσαμε κοινές ερευνητικές προκλήσεις καθώς και προκλήσεις που εστιάζουν σε νέες (πρόσφατες) πρακτικές. Για την αναπαράστασή τους σε ροές εργασίας χρησιμοποιήσαμε την ίδια οντολογία αναφοράς SO.

Στη συνέχεια, οι επιμέρους ροές ανάχθηκαν σε μία γενικότερη, πιο αφαιρετική αναπαράσταση/ροή εργασίας. Το αποτέλεσμα διατυπώθηκε ως διάγραμμα τύπου Business Process Modeling Notation (BPMN) και απεικονίζεται στην Εικ. 11, ενώ αποτελεί εξέλιξη προηγούμενης ερευνητικής ροής εργασίας [6]. Αυτή η ενιαία αφαιρετική αναπαράσταση είναι ικανή να εκφράσει το σύνολο των πρακτικών που αναδείχθηκαν, ενώ ταυτόχρονα με τη χρήση εξειδικεύσεων προσαρμόζεται ώστε να αντικατοπτρίζει τις συγκεκριμένες ανάγκες και τις ειδικές πτυχές της κάθε πρακτικής. Αποτελεί μια γεννήτρια δομή ικανή να παράξει επιμέρους ροές εργασίας και βασίζεται στις ανώτερες τάξεις της ιεραρχικής δομής Τύπου Δραστηριοτήτων της SO.



Εικ. 11. Ροή ερευνητικής εργασίας

ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] Pertsas, V., Constantopoulos, P. (2017). *Scholarly Ontology: modelling scholarly practices*. Int J Digit Libr 18, 173–190 (2017). <https://doi.org/10.1007/s00799-016-0169-3> .
- [2] Pertsas, V., Constantopoulos, P. (2016). *Scholarly Ontology Activity Types v1.3*. Available at <http://nedimah.dcu.gr/resources/> .
- [3] Webber, J. (2020). *Introducing Neo4j Graph Database 4.0 [GA Release]*. Available at <https://neo4j.com/blog/neo4j-graph-database-4-0-ga-release/> .
- [4] openCypher. (2017). *Cypher Query Language Reference*, Version 9, Nov. 2017. Available at <https://github.com/opencypher/openCypher/blob/master/docs/openCypher9.pdf> .
- [5] Ilvanidou, M., Pouli, K., Souyiultzoglou, I., Tzedopoulos, Y. (2023). *Trends in digital methods in the Humanities and Social Sciences in Greece*. Zenodo, DOI <https://zenodo.org/records/10797453>
- [6] Constantopoulos P., Dritsou V., Ilvanidou M., Chroni A. (2022). *Aggregation and Curation of Historical Archive Information*. In: Karagiannis D., Lee M., Hinkelmann K., Utz W. (eds) *Domain-Specific Conceptual Modeling*. Springer, Cham. https://doi.org/10.1007/978-3-030-93547-4_23