

Η ΣΥΜΒΟΛΗ ΤΩΝ ΨΗΦΙΑΚΩΝ ΣΩΜΑΤΩΝ (E-CORPUS) ΤΕΚΜΗΡΙΩΝ ΣΤΗΝ ΚΟΙΝΩΝΙΚΗ ΕΡΕΥΝΑ

Δημήτρης Παρασκευόπουλος
Νίκος Ναγόπουλος

Περίληψη

Οι συνεχώς εξελισσόμενες τεχνολογίες πληροφορίας και επικοινωνίας παρέχουν νέες προοπτικές στον σχεδιασμό, την ανάπτυξη και τη λειτουργία των ερευνητικών υποδομών. Εντός των τελευταίων, ξεπερνώντας περιορισμούς του παρελθόντος, αναπτύσσονται ψηφιακά σώματα (e-corpus) τεκμηρίων με δυνατότητες που αγγίζουν την «πλανητική εναρχίωση» πληροφοριών, συμβάλλοντας ταυτόχρονα στη μελέτη πεδίων που εμπίπτουν στο χώρο των κοινωνικών, ανθρωπιστικών, νομικών και πολιτικών επιστημών. Ο χώρος αναφοράς που καλύπτουν συλλογές, υποσυλλογές και ομάδες τεκμηρίων, μέσα από μια διαδικασία αναζήτησης-επισκόπησης, μετατρέπεται σε ένα καταλογογραφημένο και ταξινομημένο πεδίο βάσει του χρονοτοπικού πλαισίου και της τυπολογίας περιεχομένου. Η χρήση ψηφιακών σωμάτων (e-corpus) τεκμηρίων για ερευνητικούς σκοπούς υπόκειται κατά περίπτωση σε τυπικούς και δεοντολογικούς περιορισμούς. Στο παρόν άρθρο υποστηρίζεται ότι η πολλά υποσχόμενη νέα συνθήκη λειτουργεί αναδομητικά στον τρόπο πρόσληψης της «κοινωνικής έρευνας» και ειδικότερα στις στρατηγικές δευτερογενούς ανάλυσης δεδομένων. Σκοπός της παρούσας συμβολής είναι αφενός η προσέγγιση των λειτουργιών που δύνανται να επιτελέσουν τα ψηφιακά σώματα (e-corpus) τεκμηρίων στις νέες πολύπλευρες και πολύπλοκες διαστάσεις της ερευνητικής διαδρομής και αφετέρου η σκιαγράφηση ενός μοντέλου δόμησης e-corpus τεκμηρίων.

Λέξεις-κλειδιά: ψηφιακά σώματα (e-corpus) τεκμηρίων, κοινωνική έρευνα, ανάλυση τεκμηρίων, μετα-ανάλυση.

Abstract

The constantly evolving information and communication technologies provide new perspectives on the design, development and function of research infrastructures. Within the latter, overcoming the limitations of the past, e-corpora of documents are being developed with capabilities that reach a sort of “global information archiving”, while contributing to the study of disciplines that fall within the fields of social sciences, humanities, law and political sciences. The space covered by collections, sub-collections and groups of documents is transformed through a search-and-review process into a field catalogued and classified in terms of space-time frame and content typology. The use of e-corpora of documents for research purposes is subject in some cases to formal and deontological restrictions. In this article, it is argued that the all-promising new setup has a restructuring effect on the perception of “social research” and particularly in secondary data analysis strategies. The purpose of this article is on the one hand to consider the functions that can be performed by e-corpora of documents in the new multifaceted and complex dimensions of the research process, and on the other hand to outline a model for structuring e-corpora of documents.

Keywords: *e-corpora of documents, social research, document analysis, meta-analysis.*

Εισαγωγή

Πριν φτάσουμε στις προκλήσεις της ύστερης νεωτερικότητας –και τη συζήτηση για την παραγωγή της επιστημονικής γνώσης στο κατώφλι του 21ου αιώνα–, πρόδρομες μορφές συνεργασίας στην επιστήμη εντός ή και μεταξύ εθνικών κρατών καταγράφονται στο πέρας της σύγχρονης παγκόσμιας ιστορίας¹. Ο όρος *ερευνητικές*

1. Ο Hallonsten (2020) υποστηρίζει ότι κατά τη διάρκεια του Β΄ Παγκοσμίου Πολέμου, οι ερευνητικές υποδομές αποτέλεσαν σημαντικές πηγές κινητοποίησης της επιστήμης και της τεχνολογίας για στρατιωτικούς και στρατηγικούς σκοπούς από την πλευρά των εθνικών κρατών. Μεταπολεμικά και υπό το βάρος του σχεδίου ανοικοδόμησης της Ευρώπης, το CERN (Conseil Européen pour la Recherche Nucléaire, Γενεύη-Ελβετία, 1954), το JRC (Joint Research Centre, Ispra, Ιταλία, 1960), το E.S.O (European Southern Observatory, 1964) αποτέλεσαν δείγματα μιας ευρωπαϊκής συνεργασίας στην «επιστήμη» (Hallonsten, 2020). Σε επίπεδο Ευρωπαϊκών Κοινοτήτων, η συμπερίληψη της μέριμνας μιας πολιτικής για την επιστήμη και την τεχνολογία τοποθετείται χρονικά στα τέλη της δεκαετίας του 1970 και εκφράστηκε μέσα από την υλοποίηση πλήθους προγραμμάτων εφαρμοσμένης έρευνας και ανάπτυξης (R&D) κατά τη δεκαετία του 1980, φτάνοντας στο πρώτο και δεύτερο πρόγραμμα-πλαίσιο για την Έρευνα και την Τεχνολογική Ανάπτυξη (FP1, 1984-1987 και FP2, 1987-1991) (Hallonsten, 2020). Μετά την εφαρμογή του σχεδίου ανοικοδόμησης της Ευρώπης και το σχέδιο Marshall, τέθηκε σε πλήρη ισχύ ο Οργανισμός Οικονομικής Συνεργασίας και Ανάπτυξης (ΟΟΣΑ), εντός του

υποδομές μπορεί να περιγράψει ένα ευρύ φάσμα δομών, όπως α) περιπτώσεις τεχνικών εγκαταστάσεων που στηρίζουν την επιστημονική παρατήρηση, β) μηχανισμούς τεχνικής στήριξης της επιστημονικής ανάλυσης, γ) εγκαταστάσεις στήριξης της επιστημονικής επικοινωνίας, της συσσώρευσης και της διάχυσης των επιστημονικών τεκμηρίων, καθώς και δ) δομές υψηλής συνθετότητας που απαρτίζονται από επιμέρους υποδομές και η λειτουργία τους χαρακτηρίζεται από έναν σύνθετο καταμερισμό εργασίας (Κάλλας, 2015: 281). Οι ερευνητικές υποδομές σε επίπεδο συγκρότησης, λειτουργίας και διαχείρισης συνιστούν ένα πολύπλοκο εγχείρημα μέσα σε ένα κοινωνικοεπιστημονικό τοπίο που χαρακτηρίζεται αφενός από τη συνεχή επίδραση των εξελισσόμενων Τεχνολογιών Πληροφορικής και Επικοινωνίας (ΤΠΕ) και αφετέρου από τη διεθνοποίηση-παγκοσμιοποίηση των κοινωνικών διεργασιών και τις προεκτάσεις της στην ερευνητική διαδικασία (Μήτρου, 2011: 75-78). Στην οικοδόμηση των ερευνητικών υποδομών κατά το πέρασμα στον 21ο αιώνα, η χρήση των Τεχνολογιών Πληροφορικής και Επικοινωνίας (ΤΠΕ) συνέδραμε στην αντιμετώπιση δυσκολιών του παρελθόντος όσον αφορά την προσβασιμότητα, καθώς και στην οργάνωση της *διάσπαρτης και απομονωμένης* πληροφορίας (Κάλλας, 2015: 284). Βασική αρχή που διέπει τα συστήματα οργάνωσης των ερευνητικών υποδομών είναι η *κοινοχρησία των δεδομένων*, που προϋποθέτει-διασφαλίζει τη «ροή της ερευνητικής πληροφορίας» –με δυνατότητες κατάθεσης και τεκμηρίωσης ερευνητικών έργων και δεδομένων και πρόσβασης στα μεταδεδομένα– καθώς και την τήρηση του κανονιστικού πλαισίου –«πληροφοριακές υποχρεώσεις»– από πλευράς ερευνητικών φορέων (Μήτρου, 2011: 84-85).

Στρέφοντας τη συζήτηση στο πεδίο των διεθνοποιημένων κοινωνικών επιστημών (Mosbah-Natanson & Gingras, 2014) και των μεθόδων κοινωνικής έρευνας που εφαρμόζονται, οι Κάλλας και Κονδύλη (2008: 102-108) αναφέρονται στην ανάδυση του *διαφασικού υποδείγματος* εντός των ορίων ενός νέου επικοινωνιακού ορθολογισμού, ο οποίος περιλαμβάνει διαφορετικούς ερευνητικούς χειρισμούς στην οργάνωση της παρατήρησης και στηρίζεται στις λειτουργίες των Τεχνολο-

οποίου συστάθηκε το *Megascience Forum* (1992), που το 1999 μετονομάστηκε σε Παγκόσμιο Επιστημονικό Φόρουμ (GSF), με βασικό άξονα την ανάλυση και μελέτη της θέσης των ερευνητικών υποδομών στα πεδία οικονομικής συνεργασίας και ανάπτυξης. Καθ' όλη τη δεκαετία του 1990 δόθηκε έμφαση στο ζήτημα της καινοτομίας, διαμορφώνοντας παράλληλα το γόνιμο έδαφος ώστε στις αρχές της δεκαετίας του 2000 να συζητείται η ευρωπαϊκή στρατηγική για τις ερευνητικές υποδομές (Ευρωπαϊκό Στρατηγικό Φόρουμ για τις Ερευνητικές Υποδομές – ESFRI) (Hallonsten, 2020). Βασικό καθήκον που ανατέθηκε στο φόρουμ είναι η ανάπτυξη οδικού χάρτη –ανασκόπηση και χαρτογράφηση– των ερευνητικών υποδομών σε όλους τους τομείς της έρευνας στην Ευρώπη. Σε πρόσφατη επικαιροποίηση, ο οδικός χάρτης καταγράφει συνολικά εξήντα τρεις (63) ερευνητικές υποδομές που λειτουργούν στην Ευρώπη, υπαγόμενες στις παρακάτω θεματικές κατηγορίες: α) ψηφιακότητα, β) ενέργεια, γ) περιβάλλον, δ) υγεία και διατροφή, ε) φυσικές επιστήμες και μηχανική και ζ) κοινωνική και πολιτισμική καινοτομία (ESFRI, 2021, σ. 17-18).

γιών Πληροφορικής και Επικοινωνίας (ΤΠΕ). Το διφασικό υπόδειγμα παρουσιάζει μεγάλη εξάρτηση από τις ερευνητικές υποδομές και κινείται αντίθετα προς το μονοφασικό μεθοδολογικό υπόδειγμα, στη στρατηγική του οποίου κυριαρχεί η προτεραιοποίηση της πρωτογενούς έναντι της δευτερογενούς έρευνας και η οργάνωση της έρευνας σε αυτόνομα ερευνητικά έργα (Κάλλας, 2015: 290, 293-294). Οι ερευνητικές υποδομές συντίθεται πάνω στην τεκμηρίωση ερευνητικών έργων και πόρων. Ως ερευνητικό έργο² εντός των ερευνητικών υποδομών για τις κοινωνικές επιστήμες νοείται μια συντονισμένη και στοχοθετημένη προσπάθεια η οποία εντάσσεται στις θεωρητικές παραδόσεις των κοινωνικών επιστημών και επιδιώκει την απάντηση σε ερευνητικά ερωτήματα εφαρμόζοντας διαδικασίες κοινωνικής έρευνας. Σε κάθε στάδιο της όλης διαδικασίας (στρατηγική εννοιολόγησης, διατύπωση ερευνητικών ερωτημάτων, σχεδιασμός δειγματοληπτικού πλαισίου, συλλογή και ανάλυση δεδομένων, συγγραφή ερευνητικής έκθεσης κ.ά) αναδύονται πόροι που αφορούν είτε τα ίδια τα συλλεχθέντα/παραγόμενα δεδομένα είτε τους εκάστοτε ερευνητικούς χειρισμούς.

Εντός της Κοινωνίας της Πληροφορίας³, ο σχεδιασμός και η λειτουργία ερευνητικών υποδομών «ειδικού τύπου» (με τη μορφή συστημάτων τεκμηρίωσης) συμβάλλει στην παραγωγή νέας γνώσης μέσω της συσσώρευσης και διαχείρισης των δεδομένων (Κάλλας, 2011: 69· Κάλλας, 2015: 296). Οι Κάλλας και Κονδύλη (2008: 102-108) επισημαίνουν ότι το δεσμό στη συγκρότηση ολοκληρωμένων σειρών αποτελεί ένα σώμα μεταδεδομένων του οποίου τη διαχείριση αναλαμβάνουν οι ερευνητικές υποδομές. Βασισμένες στο διφασικό υπόδειγμα, οι ερευνητικές υποδομές συνιστούν μηχανισμούς υποστήριξης και ολοκλήρωσης επιμέρους ερευνητικών έργων μέσα σε ευρύτερα ερευνητικά προγράμματα, καθώς παρέχουν δυνατότητες όχι μόνο ανταλλαγής ερευνητικών δεδομένων αλλά και ένταξης των δεδομένων σε σειρές με κοινή πραγματολογική βάση. Η ανάλυση των κοινωνικών δεδομένων «αυτονομείται» από το πλαίσιο καταγραφής-παραγωγής/συλλογής. Τα δεδομένα που έχουν συλλεχθεί/παραχθεί δεν υπόκεινται μονάχα σε διαδικασίες αρχιοθέτησης και τεκμηρίωσης, αλλά και διασύνδεσης –μέσω της πρόσθετης γνώσης που παράγουν– με νέες αναλύσεις (Κάλλας, 2015: 293).

Στις κοινωνικές επιστήμες, η συγκρότηση και λειτουργία των ερευνητικών υποδομών υπόκειται στη διάκριση πρωτογενούς και δευτερογενούς ανάλυσης δεδομένων, ποσοτικών (Καλογεράκη, 2011· Κυριαζή, 2004· Schnell κ.ά, 2014· Singh, 2017) και ποιοτικών (Flick, 2017· Τσιώλης, 2011· Τσιώλης, 2014). Ένα ερευνητικό πρόγραμμα στο πλαίσιο του οποίου έχει υλοποιηθεί πρωτογενής ποσοτική, ποιοτική ή μεικτής μεθοδολογίας έρευνα (Creswell, 2014), εντός των ερευνητικών υποδομών μπορεί να αποτελέσει σώμα δευτερογενούς ανάλυσης στο πλαίσιο ενός

2. Βλ. και Bell (2005), Denscombe (2003).

3. Βλ. Κάλλας (2019).

νέου ερευνητικού έργου. Η παράθεση πληροφοριών (μεταδεδομένων) –τόσο σε επίπεδο συνολικού ερευνητικού έργου, όσο και σε επίπεδο επιμέρους πόρων που το πλαισιώνουν– από τον/την καταθέτη/τρια του ερευνητικού προγράμματος σε ένα αποθετήριο αποτελεί σημαντική διεργασία, καθώς συνιστά το σημείο απόθεσης, οργάνωσης και σύνδεσης του υλικού με τις προοπτικές δευτερογενούς χρήσης του (Λιναρδής, 2011: 21). Πρόκειται για τη διαφοροποίηση ανάμεσα στη λογική δόμησης μιας βάσης δεδομένων (database), η οποία διαθέτει λειτουργίες αποθήκευσης και αναζήτησης πληροφοριών, και μιας βάσης γνώσεων (knowledge base), η οποία διαθέτει λειτουργίες αναπαράστασης της υπάρχουσας γνώσης με σκοπό την υποβοήθηση, μεταξύ άλλων, σχεδίων επιστημονικής έρευνας.

Κατ' αυτό τον τρόπο –και σε μια αναδιάταξη της στρατηγικής που αφορά το εύρος των στόχων και την πολιτική διαχείριση του ερευνητικού σχεδιασμού μέσα από τη συγκέντρωση, αξιολόγηση και συμπληρωματικότητα μιας πληθώρας πληροφοριακών πηγών, στοιχείων και διαθέσιμων δεδομένων– δημιουργείται η βάση και η προοπτική μιας διαχρονικής αξιοποίησης των παραπάνω μέσα από ενισχυμένες μελέτες και μεθοδολογίες μετα-ανάλυσης. Επιπλέον, διευρύνονται οι εξειδικευμένες συλλογές δεδομένων, κατανοούνται με καλύτερο και ασφαλέστερο τρόπο τα περιεχόμενα των στοιχείων που εισάγονται και συναπαρτίζουν το συνολικό σώμα της βάσης δεδομένων και ενισχύονται οι δυνατότητες σύγκρισης και ανάδειξης της ερμηνευτικής σημαντικότητας των στοιχείων για ερευνητικούς αλλά και κοινωνικούς σκοπούς ή και διεύρυνσης των προοπτικών αξιοποίησης αναβαθμισμένου επιπέδου γνώσης για τη διάχυση εμπεριστατωμένου υλικού από τα μέσα πληροφόρησης και επικοινωνίας. Ιδιαίτερα χρήσιμη είναι η ανάλυση δευτερογενών δεδομένων σε εκείνες τις περιπτώσεις κατά τις οποίες οι επιχειρούμενες μετα-αναλύσεις οδηγούν σε απρόβλεπτα συμπεράσματα και απροσδόκητες ανακαλύψεις, που επιτρέπουν νέες εννοιολογικές συνθέσεις ευρημάτων και προοπτικές αξιοποίησης νέας γνώσης με την ανασύνθεση των ερευνητικών ερωτημάτων.

Βάσει αυτού, τα δεδομένα που εντάσσονται σε διαδικασία περαιτέρω ανάλυσης και χρησιμοποιούνται ως μεταδεδομένα μπορούν να αξιοποιηθούν σε ένα νέο ψηφιακό περιβάλλον συσχετισμών και διασυνδέσεων με άλλα δεδομένα, ενώ, λόγω της δυναμικά υψηλότερης σημαντικότητας που αποκτούν, μπορούν να διασφαλίζουν μια ευρύτερη και διαχρονική προβολή σε εθνικό και διεθνές περιβάλλον. Έτσι τα αποθηκευμένα δεδομένα αφενός αποτελούν ψηφιακές αναπαραστάσεις που υπόκεινται σε διαδικασία κοινής κωδικοποίησης με σταθερούς προσδιοριστές διαχείρισης περιεχομένου και εντάσσονται σε διακριτές συλλογές, αφετέρου συνιστούν ένα ευρύ περιβάλλον πλοήγησης, διάθεσης μεταδεδομένων και διαμοιρασμού πληροφορίας, που μπορεί να αναδείξει καινοτόμες προοπτικές σε επίπεδο μετα-ερευνητικών προκλήσεων και νέων σημασιολογικών ερμηνειών των ψηφιακών σωμάτων για περαιτέρω αξιοποίηση της γνώσης.

Σε συνέχεια των παραπάνω εισαγωγικών παρατηρήσεων, το παρόν κείμε-

νο: α) εξειδικεύεται στις προδιαγραφές των ψηφιακών σωμάτων (e-corpus) τεκμηρίων (μιας ειδικής κατηγορίας δεδομένων) εντός των ερευνητικών υποδομών, β) προχωρά στην περιγραφή των βασικών χαρακτηριστικών της σχέσης μεταξύ ερευνητικών προγραμμάτων και πόρων στο στάδιο εφαρμογής τεχνικών ανάλυσης τεκμηρίων, καταλήγοντας γ) στην περιγραφή των δυνατοτήτων μετασύνθεσης ερευνητικών έργων και κατ' επέκταση επικαιροποίησης επιστημολογικο-οντολογικών κατασκευών στα όρια της ανάδρασης με την κοινωνική πραγματικότητα της ύστερης νεωτερικότητας.

Η δημιουργία ενός ψηφιακού σώματος (e-corpus) τεκμηρίων εντός των ερευνητικών υποδομών⁴

Το πλαίσιο της τεκμηρίωσης ερευνητικών έργων προχωρά στην καταγραφή συγκεκριμένων πρωτογενών δεδομένων και την κειμενική επεξεργασία, μεταγραφή και απόδοση δεδομένων σε ψηφιακή μορφή, που μπορούν να αξιοποιηθούν ως μεταδεδομένα και αφορούν αρχεία που προκύπτουν από τη χρήση πολυμέσων – γραπτά ερωτηματολόγια, προφορικές συνεντεύξεις, αφηγήσεις και γενικότερα έρευνες πεδίου ή έρευνες δράσης (action researches). Ο πολυτροπικός χαρακτήρας των ψηφιακών βάσεων δεδομένων παρέχει τη δυνατότητα συνύπαρξης διαφορετικών ειδών και συλλογών πληροφορίας. Ας επιχειρήσουμε μια παραδειγματική αναφορά ως προς τον εντοπισμό και την καταγραφή γλωσσικών πραγματώσεων, υπογραμμίζοντας τον επιτελεστικό χαρακτήρα της γλώσσας αναφορικά με τα πρωτογενή δεδομένα. Στην περίπτωση αυτή, οι γλωσσικές πληροφορίες που προκύπτουν ως πρωτογενή δεδομένα υπόκεινται σε γλωσσολογική ανάλυση και περαιτέρω επισημειώνονται κατά τη διαδικασία κωδικοποίησης ανάλογα με τα γλωσσικά επίπεδα στα οποία αναφέρονται. Σημειώνεται ότι στο πεδίο αυτό υπάρχει μέχρι και σήμερα δυσκολία και έλλειψη όσον αφορά μια κάποια κοινή αναγνώριση και αποδοχή ενός προαποφασισμένου υποστρώματος πληροφοριών που να επισημειώνονται ως πρωταρχικές και να συστήνουν μια διαδικασία ανεξάρτητη από τις προθέσεις και τους στόχους των ερευνητών. Αντίθετα, οι γλωσσικές πληροφορίες που συνήθως επισημειώνονται σε μια βάση δεδομένων συσχετίζονται πάντα με τους ερευνητικούς στόχους που θέτει το ερευνητικό πρόγραμμα στο πλαίσιο του οποίου πραγματοποιείται η συλλογή, ανάλυση και επεξεργασία του γλωσσικού υλικού. Ωστόσο, από την έως σήμερα εμπειρία προκύπτει ότι η επισημείωση των μεταδεδομένων, όπως επίσης η οργάνωση και κωδικοποίηση γραπτών ή προφορικών πηγών σε πολυτροπικές βάσεις δεδομένων αιχμής και ψηφιακά μέσα αποθήκευσης, πραγματοποιείται κυρίως με βάση τη δομημένη μέθοδο (Παπαζαχαρίου & Καρασίμος, 2016), η οποία

4. Η πλαισίωση του ζητήματος αντλεί από τη διαχείριση ερευνητικών δεδομένων που εφαρμόζεται εντός της Εθνικής Ερευνητικής Υποδομής για τις Κοινωνικές Επιστήμες (So.Da.Net).

ακολουθεί προδιαγεγραμμένες κατηγορίες και δομές, καθώς και προαποφασισμένα σύνολα τιμών, περιορίζοντας με αυτό τον τρόπο τόσο την αδόκιμη εισαγωγή τιμών στα αποθετήρια όσο και πιθανές ασυμβατότητες στον προσδιορισμό των δεδομένων όταν αναδεικνύονται ως μεταδεδομένα και προσφέρονται για δευτερογενείς αναλύσεις.

Προσανατολισμένο στις αρχές δευτερογενούς ανάλυσης, ένα ερευνητικό έργο εντός των ερευνητικών υποδομών δύναται να απαντήσει σε ερευνητικά ερωτήματα θέτοντας είδη τεκμηρίων ως πηγή των δεδομένων. Η διάκριση μεταξύ ερευνητικών δεδομένων και τεκμηρίων υφίσταται σε περιπτώσεις αυστηρής οριοθέτησης του υλικού που θα τεθεί προς δευτερογενή ανάλυση. Ποσοτικά και ποιοτικά ερευνητικά δεδομένα συνιστούν οι πόροι προγενέστερων ερευνών, ενώ υπό τον όρο «τεκμήρια» θεωρούνται τα κοινωνικά και προσωπικά παράγωγα που έχουν παραχθεί εκτός ερευνητικής διαδικασίας (Τσιώλης, 2011: 130). Πρόκειται για ένα ετερογενές τοπίο πηγών, που συμπεριλαμβάνει προσωπικά τεκμήρια (ημερολόγια, επιστολές κ.ά.), τεκμήρια-παράγωγα θεσμικών διαδικασιών (νομοθετικά-διοικητικά κείμενα κ.ά.), οπτικά/ηχητικά τεκμήρια (φωτογραφίες, φιλικές κατασκευές κ.ά.), τεκμήρια μέσω επικοινωνίας-ενημέρωσης (εφημερίδες, περιοδικά), ενσωματώνοντας έτσι σε ένα παγκοσμιοποιημένο περιβάλλον τις μεθοδολογικές προκλήσεις των «τεράστιων δεδομένων». Ο A. Bryman (2017: 591-613) χαρτογραφεί αυτό το φάσμα πηγών στις παρακάτω κατηγορίες: α) προσωπικά τεκμήρια σε γραπτή ή οπτική μορφή, β) επίσημα τεκμήρια κρατικής προέλευσης⁵, γ) επίσημα τεκμήρια ιδιωτικής προέλευσης, δ) υλικό από τα μέσα μαζικής επικοινωνίας⁶ και ε) διαδικτυακά τεκμήρια.

Τα ψηφιακά σώματα (e-corpus) τεκμηρίων έχουν ασκήσει σημαντική επίδραση σε ποικίλα πεδία της επιστήμης και ειδικότερα στο χώρο των επιστημών του λόγου⁷. Πλήθος ερευνητικών εγχειρημάτων, αντλώντας από ψηφιακές συλλογές

5. Βλ. και Scott (1990: 63).

6. Βλ. και Babbie (2011: 517).

7. Στο χώρο της κοινωνιογλωσσολογίας, οι συνεχώς εξελισσόμενες τεχνολογίες πληροφορικής έχουν ασκήσει «επαναστατική επίδραση» (Krishnamurthy, 2018). Οι κοινωνιογλωσσικές προσεγγίσεις, μέσω της τροφοδότησης από τα e-corpus, δύναται να συμβάλλουν στη μελέτη της χρήσης της γλώσσας και φαινομένων γλωσσικής παραλλαγής (Hansen, 2019), μέσα από τη σύνδεση της περιφερειακής διαλεκτολογίας, της έμφυλης διάστασης και των σεξουαλικότητων, της ηλικίας, της φυλής, του έθνους, του χώρου εργασίας, των μέσων κοινωνικής δικτύωσης κ.ά. (Friginal, 2018). Στο πεδίο μελέτης της γλώσσας καταγράφονται συλλογές τεκμηρίων ανά τον παγκόσμιο χάρτη, οι οποίες εμφανίζουν πεδία διαφοροποίησης τόσο σε επίπεδο οργάνωσης, όσο και σε επίπεδο περιεχομένου. Σε επίπεδο γλωσσικής τεχνολογίας (Fernandez et al., 2019), ο τρόπος εισαγωγής γλωσσικών πόρων σε ένα e-corpus κινείται βάσει αρχών κωδικοποίησης και παραγωγής μεταδεδομένων που μπορεί να αξιοποιηθούν από το χώρο των κοινωνικών και ανθρωπιστικών επιστημών (βλ. *Guidelines for Electronic Text Encoding and Interchange*, 2021). Η δόμηση ενός e-corpus, που μπορεί να καλύψει παιδαγωγικές/διδασκτικές και ερευνητικές ανάγκες, σε άλλες περιπτώσεις κινείται με αποκλειστικό άξονα την καταγραφή

(e-corpus) τεκμηρίων, μελέτησαν τη δομή της γλώσσας, τη σχέση της με την κοινωνική πραγματικότητα, καθώς και αναδυόμενα πεδία γλωσσικής διαφοροποίησης και ποικιλομορφίας εντός των σύγχρονων κοινωνιών.

Η δόμηση ενός e-corpus –και ταυτόχρονα η καταγραφή τεκμηρίων που εμπύπτουν σε ένα θεματικό πεδίο– δύναται να καλύψει ανάγκες πολλαπλών κατηγοριών (παιδαγωγικές, διδακτικές, ερευνητικές κ.ά.). Συνεπώς τα τεκμήρια, ενώ δεν αποτελούν παράγωγα μιας ερευνητικής διαδικασίας, μπορούν να αποτελέσουν πηγή ερευνητικών δεδομένων (στο πλαίσιο ενός ερευνητικού έργου ή στην περίπτωση υλοποίησης ή και μετασύνθεσης πολλαπλών ερευνητικών έργων στα όρια ενός ερευνητικού προγράμματος). Στις κατηγορίες ερευνητικών έργων και ερευνητικών προγραμμάτων⁸, τα σώματα τεκμηρίων πλαισιώνουν ως πόροι την όλη ερευνητική διαδικασία. Ωστόσο, εντός των ερευνητικών υποδομών, η δόμηση ψηφιακών σωμάτων (e-corpus) τεκμηρίων για ερευνητικούς σκοπούς πιθανώς και να αυτονομείται από επιμέρους ερευνητικά έργα και προγράμματα. Στη συγκεκριμένη περίπτωση, πραγματοποιείται μια διαδικασία προδιαμόρφωσης ενός χώρου δεδομένων, που αποτελεί μηχανισμό υποβοήθησης της ερευνητικής διαδικασίας με δυνατότητες συμβολής στη μελλοντική υλοποίηση συγχρονικών, διαχρονικών, συγκριτικών και περιπτώσιολογικών σχεδίων έρευνας και σε εφαρμογές δευτερογενούς ανάλυσης.

Η δημιουργία ενός ψηφιακού σώματος (e-corpus) τεκμηρίων κινείται βάσει *μονάδων παρατήρησης* που είτε έχουν τεθεί σε προηγούμενο χρόνο (στο πλαίσιο εφαρμογής σχεδίων έρευνας που έχουν ολοκληρωθεί) είτε τίθενται ως βάση συγκρότησης ενός νέου σώματος τεκμηρίων (παρέχοντας πρόσβαση σε πηγές δεδομένων που πρόκειται να υλοποιηθούν μελλοντικά). Στη βάση αυτή, η συγκρότηση ενός σώματος τεκμηρίων αφενός δεν λειτουργεί αποκλειστικά και μόνο με στόχο τη συσσώρευση και αρχειοθέτηση τεκμηρίων, αλλά επίσης με άξονα τη συμβολή στην κοινωνική έρευνα, προϋποθέτοντας ένα μοντέλο τεκμηρίωσης δεδομένων. Προς αυτή την κατεύθυνση έχουν σχεδιαστεί διεθνή πρότυπα περιγραφής δεδομένων, όπως το Data Documentation Initiative (DDI)⁹ ή και το Dataverse Documentation¹⁰.

διαλέκτων (Johannessen, 2011). Επιπρόσθετα, στο πεδίο εφαρμογής προστίθενται εμπορικοί τομείς συνδεδεμένοι με τον ακαδημαϊκό χώρο στην επεξεργασία της γλώσσας, τη λεξικογραφία και τη μετάφραση (Anderson, 2008). Καλές πρακτικές στο πεδίο δόμησης e-corpus δεδομένων ψηφιακής γλώσσας (digital language data) αναδεικνύουν τη συμβολή της «εναρχείωσης» στην εμπειρική κοινωνική έρευνα. Σε ευρωπαϊκό επίπεδο, στην ερευνητική υποδομή CLARIN (Common Language Resources and Technology Infrastructure) συμμετέχουν φορείς από χώρες όπως η Ελλάδα, η Αυστρία, η Βουλγαρία, η Τσεχία, η Δανία, η Εσθονία, η Γερμανία, η Ολλανδία, η Πολωνία κ.ά. (βλ. Common Language Resources and Technology Infrastructure, CLARIN, <https://www.clarin.eu/>). Πρόκειται για ένα δίκτυο «λεξιλογικών και εννοιολογικών σχέσεων» εντός και μεταξύ των χωρών που συμμετέχουν (Krauer & Hinrichs, 2014).

8. Βλ. Lakatos (1978) και Kuhn (1962).

9. Βλ. DDI-Codebook 2.5 (2021).

10. Βλ. Dataverse Documentation v. 5.9 (2021).

Τα συγκεκριμένα πρότυπα λειτουργούν με σκοπό την τεκμηρίωση και διαχείριση των ερευνητικών δεδομένων, καλύπτοντας τα επιμέρους ερευνητικά στάδια (ερευνητική προβληματική, συλλογή, επεξεργασία, αρχειοθέτηση δεδομένων κ.ά.) και παρέχοντας δυνατότητες αξιολόγησης της ποιότητας του ψηφιακού υλικού.

Τα ψηφιακά τεκμήρια εντάσσονται σε αποθετήρια και σε διάφορες συλλογές, ακολουθώντας οργανωμένο σύστημα αρχειοθέτησης σύμφωνα με τον τύπο που ανήκουν και τη συνάφεια περιεχομένου που αναδεικνύουν. Ήδη η επιλογή ένταξης του υλικού σε συγκεκριμένες ψηφιακές συλλογές συνιστά ένα εγχείρημα σημασιολογικής απόδοσης και αναπαραγωγής με στόχο τον περαιτέρω εμπλουτισμό μιας κατηγορίας δεδομένων που δυνητικά προσφέρονται για περαιτέρω συσχετισμούς, δευτερογενή επεξεργασία και αξιοποίηση σε συμπληρωματική προοπτική¹¹. Η ένταξη των δεδομένων σε χώρους όπου τεκμηριώνεται το ψηφιακό περιεχόμενο, όπως τα αποθετήρια, παρέχει, πέρα από τη δυνατότητα ασφαλούς διαφύλαξης και διαχείρισης αδειών χρήσης και δικαιωμάτων πνευματικής ιδιοκτησίας, πολυδιάστατες δυναμικές και διαρκώς αναβαθμιζόμενες δυνατότητες πλοήγησης και αξιοποίησης εξελιγμένων υπηρεσιών περιεχομένου. Η όλη διαδικασία αναπτύσσεται στη βάση διεθνών προδιαγραφών και προτύπων, που διασφαλίζουν τόσο τη διατήρηση της ιδιαιτερότητας του περιεχομένου των δεδομένων όσο και τη διαλειτουργικότητα, η οποία παρέχει πολυσχιδείς δυνατότητες συγκρίσεων και συσχετισμών, που διευρύνουν το σημασιολογικό φάσμα του περιεχομένου τους και συνιστούν νέους τρόπους διαχείρισης και διάθεσης στους ενδιαφερόμενους. Το όλο εγχείρημα λαμβάνει υπόψη το γεγονός ότι το υλικό προέρχεται από φυσικούς ομιλητές και επενδύεται γλωσσικά μέσα από επικοινωνιακές διεργασίες που το καθιστούν νοηματικά επαρκές σε ένα κοινωνικοπολιτισμικό πλαίσιο. Πέρα από τα διακριτά αξιακά και πολιτισμικά στοιχεία που αποτελούν σημασιολογικές δομές, στις οποίες εμπεριέχονται ερμηνείες λόγου, η διεύρυνση σε έναν ορίζοντα κοινής γνώσης γλωσσικών κωδίκων, που παρέχεται μέσα από τον εμπλουτισμό των προτύπων, διανοίγει νέες προοπτικές στη δευτερογενή ανάλυση, καθώς και στη συγκριτική δυνατότητα ανάμεσα σε διαφορετικές συλλογές δεδομένων.

Οι πληροφορίες που τίθενται αποτελούν τη βάση από την οποία ένας ερευνητής/μία ερευνήτρια, χρήστης του αποθετηρίου δεδομένων, δύναται σε επόμενο στάδιο να χειριστεί ένα ολόκληρο σύνολο πόρων ή επιμέρους πόρους συνόλων τεκμηρίων. Στην οργάνωση μιας συλλογής τεκμηρίων η οποία δεν δημιουργήθηκε στα όρια ενός ερευνητικού έργου, αλλά με σκοπό την κάλυψη ενός θεματικού πεδίου και την συμβολή στην υλοποίηση μελλοντικού ερευνητικού έργου, προσμετρώνται οι ακόλουθες πληροφορίες:

11. Όσον αφορά τις προδιαγραφές και τις λειτουργικές ιδιότητες του αποθετηρίου δεδομένων της Εθνικής Ερευνητικής Υποδομής για τις Κοινωνικές Επιστήμες (So.Da.Net), βλ. σε αυτό τον τόμο τη διεξοδική πραγμάτευση των Λιναρδή, Αλεξάνδρη και Κληρονόμου` επιρροσθετα βλ. *Dataverse, R & Geoserver Services*.

Ο *τίτλος* συνοψίζει τα βασικά χαρακτηριστικά των πόρων που εμπεριέχονται στο σώμα της συλλογής τεκμηρίων. Στην κατηγορία *Δημιουργός της συλλογής τεκμηρίων* παρατίθενται τα στοιχεία (ονοματεπώνυμο, στοιχεία επικοινωνίας, ίδρυμα/ερευνητικός φορέας) των βασικών μελών του εγχειρήματος συγκρότησης της συλλογής. Η *περίληψη* συνοψίζει περιεκτικά τη συλλογή τεκμηρίων, τον τρόπο συλλογής και τις δυνατότητες συμβολής στην κοινωνική έρευνα. Οι *λέξεις-κλειδιά* που παρατίθενται συμβάλλουν στην εννοιολογική ταξινόμηση και στη συμπερίληψη της συλλογής σε ενδεχόμενη μελλοντική αναζήτηση σχετικού όρου από επόμενο χρήστη-ερευνητή/τρια. Η θεματική ταξινόμηση εντάσσει το όλο εγχείρημα σε *θεματικές κατηγορίες* ερευνητικού ενδιαφέροντος. Πρόσθετη πληροφορία είναι η επιλογή της γλώσσας του υλικού. *Ημερομηνία και τόπος παραγωγής της συλλογής* καθώς και *Πληροφορίες επιχορήγησης* αποτελούν πρόσθετες πληροφορίες. Στην κατηγορία *Καταθέτης/τρια* παρατίθενται τα στοιχεία των καταθετών, καθώς και η ημερομηνία κατάθεσης. Άλλα στοιχεία αποτελούν η *χρονική περίοδος* που καλύπτουν τα τεκμήρια, καθώς και η *χρονική περίοδος συλλογής*, κατά την οποία συλλέχθηκαν. Η παράθεση της πληροφορίας σχετικά με το *είδος* του τεκμηρίου αποτελεί σημαντική παράμετρο σε σχέση με τις τεχνικές δευτερογενούς ανάλυσης που μπορούν να εφαρμοστούν κατά περίπτωση. Επίσης, η καταγραφή των *πηγών των τεκμηρίων* αποτελεί βασικό μεταδεδομένο σχετικά με τις νοηματοδοτήσεις της κοινωνικής πραγματικότητας που εμπερικλείουν. Τέλος, η *γεωγραφική κάλυψη* αφορά την περιγραφή των γεωγραφικών ορίων εντός των οποίων αναδύθηκαν τα τεκμήρια.

Το στοιχείο –ο πόρος– κάθε ολοκληρωμένης συλλογής τεκμηριώνεται και ταξινομείται μέσα από την παράθεση, μεταξύ άλλων, των παρακάτω πληροφοριών: α) όνομα αρχείου, β) ταξινομητική κατηγορία, γ) περιγραφή περιεχομένου, δ) γλώσσα στοιχείου, ε) ταυτοποιητικός σύνδεσμος/ διαδικτυακός σύνδεσμος (εφόσον υπάρχει), ζ) πηγή του αρχείου, η) δυνατότητες πρόσβασης στο αρχείο σε επίπεδο χρηστών της ερευνητικής υποδομής, θ) λοιπές σημειώσεις, ι) τύπος αρχείου (αρχείο κειμένου κ.ά.). Οι πόροι κάθε συλλογής παρατίθενται με τη μορφή καταλόγου ή δέντρου βάσει κριτηρίων ταξινόμησης.

Ο Α. Bryman (2017: 607) αναφέρει ότι «κάθε τεκμήριο θα πρέπει να ιδωθεί στη σύνδεσή του με άλλα τεκμήρια, διότι τα διάφορα τεκμήρια αναφέρονται ή/και αποκρίνονται σταθερά σε άλλα τεκμήρια», προσθέτοντας ότι «κάποια άλλα τεκμήρια αποτελούν μέρος του πλαισίου ή υποβάθρου σύνταξης ενός τεκμηρίου». Κατ'επέκταση, το περιβάλλον μιας ερευνητικής υποδομής ξεπερνά τα όρια ενός συστήματος αρχειοθέτησης. Η αρχή της διακειμενικότητας προϋποθέτει δυνατότητες προσθήκης/επικαιροποίησης της υπάρχουσας βάσης και περιγραφής της πιθανής αλληλοσύνδεσης μεταξύ των τεκμηρίων. Στη βάση αυτή, η λειτουργία του συστήματος προβάλλει με γνώμονα τον σημασιολογικό εμπλουτισμό ενός αποθετηρίου το οποίο διαρθρώνεται πατώντας πάνω σε ένα κοινό λεξιλόγιο όρων.

Ο χώρος/το πεδίο αναφοράς (βλ. προηγούμενως θεματική ταξινόμηση) των

συλλογών τεκμηρίων συνιστά στοιχείο ομοιογένειας, διαπερνώντας το σώμα της συλλογής, υποσυλλογής ή ομάδας τεκμηρίων μέσα από μια διαδικασία αναζήτησης/επισκόπησης και συλλογής/αρχειοθέτησης. Πρόκειται για μια διαδικασία μετατροπής των πηγών σε ένα καταλογογραφημένο και ταξινομημένο χώρο βάσει του χρονοτοπικού πλαισίου και της τυπολογίας περιεχομένου – βασικών στοιχείων ετερογένειας. Στη βάση αυτή, υφίσταται διάκριση μεταξύ γενικών σωμάτων, εξειδικευμένων σωμάτων, καθώς και τοπικά και ιστορικά προσδιορισμένων σωμάτων τεκμηρίων (Anderson, 2008).

Η χρήση ενός e-corpus τεκμηρίων υπόκειται στους κανόνες του προτύπου διαχείρισης/χρήσης, δημοσίευσης και αρχειοθέτησης των δεδομένων. Το είδος της ερευνητικής πληροφορίας που αρχειοθετείται –συμπεριλαμβανομένων των νομικών και δεοντολογικών δεσμεύσεων προστασίας– αποτελεί σημαντική παράμετρο για τις δυνατότητες χρήσης και επεξεργασίας που παρέχονται σε επίπεδο χρηστών της ερευνητικής υποδομής.

Κάθε e-corpus τεκμηρίων πριν από την ανάρτηση αξιολογείται βάσει των όρων χρήσης της πλατφόρμας από τον/την Υπεύθυνο/η του Αποθετηρίου του αποθετηρίου/υποδομής. Επιπροσθέτως, ανάλογα με το είδος και τη φύση των πόρων λαμβάνονται υπόψιν ζητήματα προστασίας προσωπικών δεδομένων, πνευματικής ιδιοκτησίας και όρων χρήσης που εφαρμόζονται στην πηγή των δεδομένων καθώς και των όρων πρόσβασης και χρήσης που εφαρμόζονται στην ερευνητική υποδομή.

Ερευνητικά έργα και πόροι: μέθοδοι ανάλυσης των e-corpus τεκμηρίων, διακειμενικότητα και παραγωγή νέας ερευνητικής γνώσης

Όπως εύστοχα επισημαίνει ο Κάλλας (2015: 295), «με την ανάδυση του διφασικού υποδείγματος και με την ανάπτυξη ερευνητικών υποδομών δημιουργήθηκε η δυνατότητα μια ευρύτερης παραγωγής και επεξεργασίας τεκμηρίων, που ξεπερνά τα όρια ενός μεμονωμένου ερευνητικού έργου». Συνεπώς τα τεκμήρια μπορεί να αποτελέσουν πόρους πολύ περισσότερων από ένα ερευνητικό έργο, ενώ νέα ερευνητικά έργα μέσω προσθήκης πόρων μπορεί να επικαιροποιήσουν το περιεχόμενο υπαρχουσών συλλογών τεκμηρίων. Στην ανάλυση της «πραγματικότητας» των τεκμηρίων, σημαντική παράμετρος είναι η παραδοχή ότι τα τεκμήρια «συντάσσονται για συγκεκριμένους σκοπούς και όχι απλώς για να αναπαραστήσουν την πραγματικότητα». Οι ερευνητές/ερευνήτριες πιθανά να πρέπει να προσεγγίσουν άλλες πηγές δεδομένων με σκοπό την κατανόηση του πλαισίου παραγωγής τους (Bryman, 2017: 607).

Στην κοινωνική έρευνα, η επεξεργασία των τεκμηρίων πραγματοποιείται κυρίως μέσω της ανάλυσης περιεχομένου¹² (Mayring, 2000· Bryman, 2017) ή της

12. Βασικός στόχος στην εφαρμογή της ανάλυσης περιεχομένου είναι η καταγραφή των

ανάλυσης λόγου¹³ (Gee, 2011) ή της κριτικής ανάλυσης λόγου¹⁴ (Fairclough, 2010). Ωστόσο το είδος των τεκμηρίων μπορεί και να προσφέρει εφαρμογή περισσότερο εξειδικευμένων τεχνικών ανάλυσης, όπως η αφηγηματική ανάλυση, η οποία δύναται να εφαρμοστεί πέραν των δεδομένων που έχουν παραχθεί στα όρια της κοινωνικής έρευνας –μέσα από την τεχνική της βιογραφικής αφήγησης (Τσιώλης, 2010)– και σε προσωπικά τεκμήρια που καλούνται να περιγράψουν τη ζωή ενός ατόμου. Τα σώματα τεκμηρίων μπορούν να αποτελέσουν χώρο άντλησης δεδομένων συγχρονικών σχεδίων, διαχρονικών σχεδίων, συγκριτικών σχεδίων ή σχεδίων περιπτώσιολογικών μελετών (Bryman, 2017: 74-96) με τη χρήση τεχνικών δευτερογενούς ανάλυσης. Επιπλέον, υπό το πρίσμα αυτών των δυνατοτήτων παρέχεται η δυνατότητα εφαρμογής μεικτών τεχνικών δευτερογενούς ανάλυσης. Ένα παράδειγμα μεικτής στρατηγικής δευτερογενούς ανάλυσης συναντάμε στη μελέτη περιεχομένου των μηνυμάτων των Μέσων Μαζικής Επικοινωνίας. Ένα θέμα «επικαιρότητας» και οι μιντιακές αναπαραστάσεις επ’ αυτού ως πεδίο ερευνητικού προβληματισμού μπορούν να αποτελέσουν τη βάση συγκρότησης ενός e-corpus τεκμηρίων σε εθνικό επίπεδο, με υψηλό αριθμό εγγραφών και αντίστοιχη δυσκολία εφαρμογής τεχνικών ανάλυσης. Στη βάση αυτή, μια αρχική ποσοτική περιγραφή μέσω ανάλυσης περιεχομένου θα μπορούσε να καταγράψει τις κεντρικές τάσεις του περιεχομένου των αναπαραστάσεων, ενώ σε δεύτερο χρόνο μια σε βάθος σημειωτική ανάλυση θα συνέβαλε στην περιγραφή του κώδικα, δηλαδή ενός συστήματος σημείων που διαπερνά τις αναπαραστάσεις γύρω από το φαινόμενο¹⁵.

τάσεων στο περιεχόμενο της επικοινωνίας, με τη χρησιμοποίηση τεχνικών κατηγοριοποίησης, σύγκρισης και συσχέτισης των στοιχείων που έχουν συλλεχθεί. Ως μονάδα καταγραφής «προσδιορίζεται το τμήμα του κειμένου που θα αποτελέσει τη βάση για την κατηγοριοποίηση: η λέξη, η πρόταση, η παράγραφος, το θέμα [...] ή το συνολικό κείμενο [...]» (Κυριαζή, 1998: 289). Για την πραγματοποίηση σύγκρισης, συσχέτισης και κατηγοριοποίησης (Robson, 2007: 647), τίθενται με όσον το δυνατόν περισσότερη ακρίβεια άξονες (κατηγορίες) ανάλυσης (Berelson, 1971) ή «υποθετικές κατασκευές» (Silvermann, 2000: 825). Μέσα από αυτή τη διαδικασία ανάλυσης (Finch, 1990) απαιτείται συνεπώς η «ανάπτυξη μιας διαχειρίσιμης κατηγοριοποίησης ή ενός κωδικοποιημένου σχήματος» (Patton, 2002: 463) με βάση την οποία μελετάται γραμμή-γραμμή ο γραπτός λόγος (Straus & Corbin, 1996: 56).

13. Μέσα από την εφαρμογή της ανάλυσης λόγου διερευνάται ο τρόπος χρήσης λέξεων/λεκτικών συνόλων στην κατασκευή και μετάδοση ενός νοήματος (Gee, 2011).

14. Στην κριτική διάσταση της τεχνικής (κριτική ανάλυση Λόγου), τα ερευνητικά δεδομένα τοποθετούνται μέσα σε κοινωνιο-πολιτισμικά συμφραζόμενα (Fairclough, 2010), εμπειρικλείοντας την περιγραφή της διαλεκτικής σχέσης ανάμεσα στην καταγεγραμμένη επικοινωνία και τις συνθήκες, τους θεσμούς και τις κοινωνικές δομές που την περιβάλλουν.

15. Η ανάλυση περιεχομένου είναι μια τεχνική έρευνας με σκοπό την αντικειμενική, συστηματική και ποσοτική περιγραφή του έκδηλου περιεχομένου της επικοινωνίας (Σεραφετινίδου, 2005: 271). Στη συγκεκριμένη περίπτωση, η εννοιολόγηση είναι μια διαδικασία ορισμού αναλυτικών κατηγοριών. Το «σπάσιμο του μηνύματος» σε συστατικές κατηγορίες ή στοιχεία παρέχει τη δυνατότητα μετρήσεων συμβόλων (λέξεων, φράσεων, εικόνων) και συσχετίσεων (Σεραφετινίδου, 2005: 271-272). Η παραπάνω τεχνική δύναται να συμβάλλει στην καταγραφή

Ως επίκαιρο παράδειγμα ψηφιακών corpus που παρέχουν τη δυνατότητα περαιτέρω επεξεργασίας και ανάλυσης τεκμηρίων αναφέρουμε το Coronavirus Corpus (Davies, 2021), που έχει σχεδιαστεί για την καταγραφή των κοινωνικών, πολιτιστικών και οικονομικών επιπτώσεων του κορωνοϊού (COVID-19) από το 2020, και το οποίο δεν περιορίζεται στις επιλογές αναζήτησης και στην πυκνότητα αυτών των επιλογών, αλλά διευρύνει τις δυνατότητες της αποθηκευμένης πληροφόρησης ως προς το πώς εκφράζονται πραγματικά οι άνθρωποι σε διάφορα μέσα επικοινωνίας, σε ηλεκτρονικά περιοδικά και διαδικτυακές εφημερίδες σε είκοσι (20) διαφορετικές χώρες. Είναι σημαντικό να διαπιστώνει κανείς στην πορεία του χρόνου τις συχνότητες εμφάνισης, αλλά ενδεχομένως και τη διακύμανση του περιεχομένου εννοιών όπως «κοινωνική απόσταση», «απομόνωση», «ατομική ευθύνη» κ.λπ. και να επιχειρεί συνθέσεις αναλύοντας τα σημασιολογικά μοτίβα των λέξεων και των εννοιών, ανάλογα π.χ. με το μέγεθος της κρίσης, την αύξηση των κρουσμάτων ή της ανάληψης συγκεκριμένων μέτρων πολιτικής. Με αυτό τον τρόπο η αναζήτηση, χρήση και ανανέωση των ψηφιακών τεκμηρίων στα αποθετήρια βρίσκεται στο επίκεντρο ενός συνεχούς ερευνητικού και συνάμα κοινωνικού ενδιαφέροντος, καθιστώντας την όλη διαδικασία μια ενεργή ερευνητική κουλτούρα ανάλυσης και ανάδειξης νέων στοιχείων στη δημόσια σφαίρα. Ιδιαίτερα χρήσιμη είναι και η σύγκριση των στοιχείων ανάμεσα σε διαφορετικά, μικρά ή μεγαλύτερα διαστήματα, αλλά και ο τρόπος που εννοιολογούνται σε διαφορετικές χώρες.

Αντί επιλόγου:

Δυνατότητες μετασύνθεσης επιστημολογικών παραδειγμάτων

Το πλαίσιο κατανόησης της διαδικασίας μετατροπής των δεδομένων σε πληροφορίες, μετέπειτα σε γνώση και τέλος σε σοφία («πυραμίδα της γνώσης») αρθρώνεται σε διάφορα στάδια. Τα δεδομένα καλύπτουν τη βάση της πυραμίδας. Σε ένα αρχικό στάδιο, η διαδικασία δόμησης των δεδομένων τα καθιστά σώμα πληροφοριών, οι

των τάσεων που κυριαρχούν στο ετερόκλητο πλήθος των μέσων και των μηνυμάτων μαζικής επικοινωνίας, καθώς και των τάσεων που αναδύονται περιπτωσιολογικά. Μέσα από αυτή τη διαδικασία, ένας ερευνητής μπορεί να εστιάσει σε μεταγενέστερο χρόνο στην ανάλυση συμβόλων, εφαρμόζοντας την επιστήμη των σημείων ή τεχνική της σημειωτικής. Η ανάλυση συμβόλων μέσω σημειωτικής δύναται να εφαρμοστεί όχι μόνο σε τεκμήρια, αλλά και σε όλα τα υπόλοιπα είδη δεδομένων. Το σημείο απαρτίζεται από το σημαίνον και το σημαϊνόμενο. Το σημαίνον είναι το πράγμα που δείχνει προς μια υποκείμενη σημασία. Το σημαϊνόμενο είναι η σημασία προς την οποία δείχνει το σημαίνον (Bryman 2017: 611-612). Η καταδηλωτική σημασία είναι η φανερή ή πιο προφανής σημασία ενός σημαϊνοντος και ως τέτοια αποτελεί ένδειξη της λειτουργίας του. Η συνδηλωτική σημασία είναι μια σημασία πρόσθετη στην καταδηλωτική, που συνδέεται με συγκεκριμένο κοινωνικό πλαίσιο. Μέσα από αυτή τη διαδικασία, περιγράφεται ένας κώδικας (σύστημα σημείων) και παρέχεται η δυνατότητα πολλαπλής ερμηνείας (πολυσημία) (Bryman 2017: 611-612).

οποίες σε ένα δεύτερο στάδιο –μέσα από τη χρήση υπολογιστικών συστημάτων και τη διαδικασία κωδικοποίησης– συμβάλλουν στην περαιτέρω απόκτηση γνώσης.

Η παραγωγή πολλαπλών ερευνητικών έργων που αντλεί από κοινούς πόρους δύναται να αναδείξει το φάσμα των επιστημολογικών παραδειγμάτων που διέπει το χώρο των κοινωνικών επιστημών. Δεν πρόκειται για μια εύκολη διαδικασία, και εξάλλου στην όλη συνθήκη προστίθενται ενστάσεις και σκεπτικισμοί σχετικά με την ποιότητα των συνεχώς αναδυόμενων μελετών. Στην όλη συζήτηση, ζητήματα αξιοπιστίας, δυνατοτήτων επαναληψιμότητας και εγκυρότητας στα όρια αξιολόγησης των ερευνητικών έργων επανέρχονται συνεχώς (Bryman, 2017: 70). Ο Appadurai (2001) επισημαίνει τις πολλαπλές προκλήσεις που αναδύονται για την κοινωνική έρευνα στο πεδίο των σύγχρονων παγκόσμιων κοινωνικών διεργασιών. Η διαδικασία σταθερής ερευνητικής παραγωγής στη βάση κυρίαρχων παραδειγμάτων και ποσοτικών κριτηρίων αναδεικνύει την αδυναμία ύπαρξης αναγκαίου χρόνου αναστοχασμού και εφαρμογής διαδικασιών ποιοτικού ελέγχου και επικύρωσης των αποτελεσμάτων (Cunha, 2013).

Οι ερευνητικές υποδομές, μέσω της τεκμηρίωσης ερευνητικών έργων, μπορούν, πέρα από βασικές λειτουργίες όπως η πρόσβαση σε ερευνητικούς πόρους κ.ά., να παράσχουν χρήσιμα εργαλεία μετασύνθεσης ή μετα-ανάλυσης¹⁶ επιστημολογικών και οντολογικών σχημάτων που υιοθετούνται σε επιμέρους ερευνητικές δράσεις¹⁷.

Σε μια προσπάθεια αντίστοιχης μετασύνθεσης, μπορεί να προσεγγιστούν οι πολλαπλές τροπικότητες στον καθορισμό του ερευνητικού αντικειμένου, και ειδικότερα όσον αφορά τον προσδιορισμό α) της φύσης των φαινομένων ή των οντοτήτων ή της κοινωνικής «πραγματικότητας» που τέθηκαν προς διερεύνηση, β) του τι συνιστά γνώση ή ένδειξη των οντοτήτων ή της κοινωνικής πραγματικότητας, γ) του θέματος ή του πεδίου που αφορά η έρευνα, δ) του νοητικού γρίφου και των ερευνητικών ερωτημάτων και ε) του σκοπού της έρευνας (Mason, 2003: 32-34).

Ειδικότερα στις εφαρμογές εμπειρικής κοινωνικής έρευνας, επιμέρους σημεία αποτελούν ο τρόπος πλαισίωσης της ποιοτικής, της ποσοτικής και της μεικτής στρατηγικής με αφετηρία τη θέση της υπάρχουσας γνώσης εντός του ερευνητικού σχεδιασμού και τις διαφοροποιήσεις μεταξύ θετικιστικού μοντέλου, κονστрукτιβιστικών αρχών και ρεαλιστικού υποδείγματος (προσδιορισμοί της πορείας στην περίπτωση της παραγωγικής, της επαγωγικής και των μεικτών σχεδίων) (Bryman,

16. Στην περίπτωση θεματικής σύνθεσης ο ερευνητής μετά τον καθορισμό του «ερωτήματος επισκόπησης» προχωρά στην αναζήτηση μελετών οι οποίες πληρούν «τα κριτήρια που θέτει το ερευνητικό ερώτημα», απορρίπτοντας τις περιπτώσεις μελετών όπου τα συγκεκριμένα κριτήρια δεν πληρούνται (Bryman, 2017: 645-646). Η μετέπειτα ανάλυση προβλέπει την κωδικοποίηση του κειμένου των μελετών και εν συνεχεία τη διαμόρφωση περιγραφικών και –μέσα από μια διαδικασία συμπύκνωσης αυτών– αναλυτικών θεματικών μοτίβων (Bryman: 647).

17. Ως προς το ρόλο των ερευνητικών υποδομών νέου τύπου, βλ. Κάλλας (2015: 296).

2017). Επιπλέον, και στη βάση της παραπάνω κατηγορικής διάκρισης, σύνθετο πεδίο μελέτης αποτελεί η μετα-ανάλυση του περιεχομένου των διαδικασιών και των προϊόντων εννοιολόγησης. *Ποια τα όρια θεωρητικής αφαίρεσης και ποιες οι δυνατότητες στη διατύπωση λεκτικών συνόλων με δυναμική οριοθέτησης και εμπειρικής σύλληψης της πολυπλοκότητας ενός, προς εξέταση, κοινωνικού φαινομένου; Ποιοι οι χειρισμοί του ερευνητή/της ερευνήτριας στην κατάρτιση του δειγματοληπτικού πλαισίου και την συλλογή/παραγωγή των δεδομένων; Τόσο η επιλογή δειγματοληπτικού πλαισίου όσο και η επιλογή τεχνικής για τη συλλογή ερευνητικών δεδομένων (μέθοδος της δομημένης συνέντευξης, ερωτηματολόγια ατομικής συμπλήρωσης ή διαδικασία της συμμετοχικής παρατήρησης, διεξαγωγή ποιοτικών συνεντεύξεων, υλοποίηση ομάδων εστίασης) βρίσκονται σε τροχιά αλληλεξάρτησης με διλήμματα που υποδηλώνουν έναν επιστημολογικο-οντολογικό σχήμα: α. να βασιστώ σε μια δοκιμασμένη θεωρία με σκοπό την απόρριψη ή την επαλήθευση μιας οπτικής, ή να προσπαθήσω διερευνητικά να αντλήσω θεωρία και γνώση μέσα από τα ερευνητικά δεδομένα; β. να αναζητήσω δεδομένα γενικεύσιμα σε έναν πληθυσμό (μακροεπίπεδο) ή να προβώ στην κατανόηση του πλαισίου, των κρυφών νοημάτων (μικροεπίπεδο); (Mason, 2003: 240-241).*

Στη βάση των παραπάνω, εντός των ερευνητικών υποδομών για τις κοινωνικές επιστήμες, η σύνθεση ερευνητικών έργων αφενός συμβάλλει στην καταγραφή της υπάρχουσας γνώσης ενός πεδίου, λειτουργώντας ως «βάση για την περαιτέρω προώθηση της μελλοντικής έρευνας» (Bryman, 2017: 644), αφετέρου δύναται να προσεγγίσει το φάσμα επιστημολογικών και οντολογικών σχημάτων που υιοθετούνται στην εκάστοτε πορεία μελέτης ενός φαινομένου. Συνεπώς, οι ερευνητικές υποδομές αναδύονται ως κόμβοι διαχείρισης της ερευνητικής πληροφορίας στα όρια ενός διεπιστημονικού διαλόγου, παρέχοντας τη δυνατότητα καταγραφής του διαχρονικά ετερόκλητου πλήθους των ιδεών και αφηγήσεων στο χώρο των επιστημών, σε συνδυασμό με τις πολιτισμικές επιρροές που διαπερνούν την καταγραφόμενη τοπικο-παγκοσμιοποιημένη κοινωνική πραγματικότητα.

Βιβλιογραφικές αναφορές

- Anderson, W. (2008). Corpus Linguistics in the UK: Resources for Sociolinguistic Research. *Language and Linguistics Compass*, 2, 352-371.
- Appadurai, A. (Coord.). (2001). *Globalization*. Durham: Duke University Press.
- Babbie, E. (2011). *Εισαγωγή στην κοινωνική έρευνα*. Αθήνα: Κριτική.
- Bell, J. (2005). *Doing Your Research Project: A Guide For the First-Time Researchers In Education, Health And Social Science* (4th ed.). Maidenhead: Open University Press.
- Berelson, B. (1971). *Content Analysis in Communication Research*. New York: Hafner.
- Bryman, A. (2017). *Μέθοδοι κοινωνικής έρευνας*. Αθήνα: Gutenberg.
- Creswell, J., W. (2014). *Research design: qualitative, quantitative, and mixed methods approaches*. Thousand Oaks: Sage.

- Cunha, I. F. (2013). The Globalization of Research in Social Sciences: The case of Communication Studies in the Ibero-American and Lusophone Space, *MARTIZES*, Ano 7 N° 1. January/June, 149 -165.
- Dataverse Project (2021 Dec. 13). *Dataverse Documentation v. 5.9*. Retrieved from <https://guides.dataverse.org/en/latest/>
- DDI-Codebook 2.5. (n.d.). Retrieved January 11, 2022, from DDI website, <https://ddialliance.org/Specification/DDI-Codebook/2.5/>
- Davies, M. (2021). The Coronavirus Corpus Design, construction, and use. *International Journal of Corpus Linguistics*, 26(4), 583-598.
- Denscombe, M. (2003). *The Good Research Guide: for small-scale social research projects*. Maidenhead: Open University Press.
- European Strategy Forum on Research Infrastructures (ESFRI). (2021). *Roadmap 2021 Strategy report on research infrastructures*. Ανακτήθηκε από <https://roadmap2021.esfri.eu/media/1295/esfri-roadmap-2021.pdf>
- Fairclough, N. (2010). *Critical Discourse Analysis: The Critical Study of Language*. New York: Routledge.
- Fernandez, P. D., Samy, D., Arenas-Garcia, J., & Gonzalez, J. de dios L. (2019). Computers that Read and Understand (almost): Language Technologies and Scientific Information Management. Στο B. Maegaard, and R. Pozzo (επιμ.), A. Melloni and M. Woollard (συνεπιμ.). *Stay tuned to the future. Impact of research infrastructures for social sciences and humanities* (101-108). Firenze: Istituto per il Lessico Intellettuale Europeo, Storia delle Idee, Leo S. Olschki Editore.
- Finch, H. (1990). Analysing qualitative material. In: M., Slater (Ed.), *Research Methods in Library and Information Studies* (128-147). London: Library Association.
- Flick, U. (2017). *Εισαγωγή στην Ποιοτική Έρευνα*. Αθήνα: Προπομπός.
- Friginal, E. (Ed). (2018). *Studies in Corpus-Based Sociolinguistics*. Routledge.
- Gee J.P. (2011). *An Introduction to Discourse Analysis, Theory and Method* (3η έκδ.). New York and London: Routledge.
- Hallonsten, O. (2020). Research Infrastructures in Europe: The Hype and the Field. *European Review*, 28 (4), 617-635.
- Hansen, B. (2018). *Corpus Linguistics and Sociolinguistics. A Study of Variation and Change in the Modal Systems of World Englishes*. Brill.
- Hymes, D. H. (Ed.) (1964). *Language in culture and society: A reader in linguistics and anthropology*. New York: Harper & Row.
- Johannessen, B. J. (2011). The Nordic Dialect Corpus – a joint research infrastructure. *Language Variation Infrastructure. Papers on selected projects*, 3 (2).
- Κάλλας, Ι. & Κονδύλη, Δ. (2008). Το διφασικό υπόδειγμα και οι ερευνητικές υποδομές ως προϋπόθεση για τη σύζευξη της ποσοτικής και της ποιοτικής μεθοδολογίας. Στο Γ. Κάλλας, Δ. Κονδύλη & Γ. Καραγιάννης (επιμ.), *Μεθοδολογικά ζητήματα και ερευνητικές υποδομές των κοινωνικών επιστημών* (96-109). Αθήνα - Μυτιλήνη: Πανεπιστήμιο Αθηνών - Πανεπιστήμιο Αιγαίου.
- Κάλλας, Ι. (2011). Η οργάνωση της παρατήρησης στις κοινωνικές επιστήμες και οι συνέπειές της στη συγκρότηση των ερευνητικών υποδομών. Στο Γ. Τσιώλης, Ν. Σερντεδάκης, και Γ. Κάλλας (επιμ.), *Ερευνητικές υποδομές και δεδομένα στην εμπειρική κοινωνική έρευνα* (27-74). Αθήνα: νήσος.
- Κάλλας, Ι. (2015). *Θεωρία, Μεθοδολογία και Ερευνητικές Υποδομές στις Κοινωνικές Επιστήμες*. Αθήνα: Κριτική.

- Κάλλας, Ι. (2019). *Αναλύοντας τις Κοινωνίες της Πληροφορίας*. Αθήνα: Προπομπός.
- Krauwier, S., & Hinrichs, E. (2014). The CLARIN Research Infrastructure: Resources and Tools for e-Humanities Scholars. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC-2014)*, (pp. 1525 – 1531).
- Καλογεράκη, Σ. (2011). Δευτερογενής ανάλυση ποσοτικών δεδομένων: Μια μεθοδολογική επιλογή με πολλαπλά οφέλη και περιορισμούς. Στο Γ. Τσιώλης, Ν. Σερντεδάκης, και Γ. Κάλλας (επιμ.), *Ερευνητικές υποδομές και δεδομένα στην εμπειρική κοινωνική έρευνα* (161-188). Αθήνα: νήσος.
- Krishnamurthy, R. (2018). Modes of analysis: An autoergography of corpus linguistics from lexis to discourse. Στο Α. Čermáková, & Μ. Mahlberg (επιμ.), *The Corpus Linguistics Discourse. In honour of Wolfgang Teubert* (pp. 35-75). John Benjamins.
- Kuhn, S. T. (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- Κυριαζή, Ν. (1998). *Η Κοινωνιολογική έρευνα. Κριτική επισκόπηση των μεθόδων και των τεχνικών*. Αθήνα: Ελληνικές Επιστημονικές Εκδόσεις.
- Κυριαζή, Ν. (2004). Η Κοινωνιολογική Έρευνα και η Κατασκευή της Κοινωνικής Πραγματικότητας. Το παράδειγμα της ποσοτικής προσέγγισης. Στο Γ. Παπαγεωργίου (επιμ.), *Μέθοδοι στην Κοινωνιολογική Έρευνα*. Αθήνα: Τυπωθήτω - Γ. Δαρδανός.
- Lakatos, I. (1978). *The methodology of scientific research programmes: Philosophical papers, vol I*. Cambridge: Cambridge University Press.
- Λιναρδής, Α. (2011). Η σημασία της ψηφιακής τεκμηρίωσης. Τα πρότυπα τεκμηρίωσης ποσοτικών δεδομένων. Στο Γ. Τσιώλης, Ν. Σερντεδάκης, και Γ. Κάλλας (επιμ.), *Ερευνητικές υποδομές και δεδομένα στην εμπειρική κοινωνική έρευνα* (217-247). Αθήνα: νήσος.
- Mason, J. (2003). *Η διεξαγωγή της ποιοτικής έρευνας*. Αθήνα: Ελληνικά Γράμματα.
- Mayring, P. (2000). Qualitative content analysis. *Qualitative Social Research* [Online], 1(2), 20. <https://www.qualitative-research.net/index.php/fqs/article/view/1089/2385>
- Μήτρου, Λ. (2011). Θεσμικά ζητήματα διαχείρισης και λειτουργίας ερευνητικών υποδομών. Στο Γ. Τσιώλης, Ν. Σερντεδάκης, και Γ. Κάλλας (επιμ.), *Ερευνητικές υποδομές και δεδομένα στην εμπειρική κοινωνική έρευνα* (75-127). Αθήνα: νήσος.
- Mosbah-Natanson, S., & Gingras, Y. (2014). The globalization of social sciences? Evidence from a quantitative analysis of 30 years of production, collaboration and citations in the social sciences (1980–2009). *Current Sociology*, 62 (5), 626-646.
- Παπαζαχαρίου, Δ., & Καρασίμος, Α. (2016). Οργάνωση και κωδικοποίηση προφορικών πηγών σε πολυτροπική βάση δεδομένων αιχμής: η περίπτωση του AMIGRE corpus. Στο Α. Ralli και Σ. Μρομπολας (επιμ.), *Πρόγραμμα ΘΑΛΗΣ: «Πόντος, Καπαδοκία, Αίβαλί: στα χνάρια της Μικρασιατικής Ελληνικής»* (55-68). Εργαστήριο Νεοελληνικών Διαλέκτων, Πανεπιστήμιο Πάτρας.
- Patton, M. (2002). *Qualitative research and evaluation methods*. London: Sage.
- Robson, C. (2007). *Η Έρευνα του Πραγματικού Κόσμου. Ένα Μέσον για Κοινωνικούς Επιστήμονες και Επαγγελματίες Ερευνητές*. Αθήνα: Gutenberg.
- Scott, J. (1990). *A Matter of Record: Documentary Sources in Social Research*. Cambridge, UK: Polity Press.
- Σεραφετινίδου, Μ. (2005). *Κοινωνιολογία των Μέσων Μαζικής Επικοινωνίας: Ο ρόλος των Μέσων στην Αναπαραγωγή του Σύγχρονου Καπιταλισμού*. Αθήνα: Gutenberg.
- Schnell, R., Hill, P. & Esser, E. (2014). *Μέθοδοι εμπειρικής έρευνας*. Αθήνα: Προπομπός.
- Silverman, D. (2000). *Doing Qualitative Research: A Practical Handbook*. London: Sage.
- Singh, K. (2007). *Quantitative Social Research Methods*. Los Angeles: Sage.

- Straus, A. & Corbin, J. (1996). *Basics of qualitative research*. London: Sage.
- Dataverse, R. & Geoserver Services. (n.d.). Ανακτήθηκε στις 10 Ιανουαρίου 2022, από τον ιστότοπο Social Data Network, <https://www.sodanet.gr/other-services/dataverse>
- TEI (2021 Aug 31). *Guidelines for Electronic Text Encoding and Interchange*. Ανακτήθηκε από <https://tei-c.org/release/doc/tei-p5-doc/en/html/AB.html>
- Τσιώλης, Γ. (2010). Η Επικαιρότητα της Βιογραφικής Προσέγγισης στην Ποιοτική Κοινωνική έρευνα. Στο Μ. Πουρκός & Μ. Δαφέρμος (επιμ.), *Ποιοτική Έρευνα στις Κοινωνικές Επιστήμες: Επιστημολογικά, Μεθοδολογικά και Ηθικά Ζητήματα* (347-370). Αθήνα: Τόπος.
- Τσιώλης, Γ. (2011). Δευτερογενής ανάλυση ποιοτικών δεδομένων: μια ερευνητική στρατηγική συμβατή με την ποιοτική προσέγγιση; Στο: Γ. Τσιώλης, Ν. Σερντεδάκης και Γ. Κάλλας (επιμ.), *Ερευνητικές υποδομές και δεδομένα στην εμπειρική κοινωνική έρευνα* (129-159). Αθήνα: νήσος.
- Τσιώλης, Γ. (2014). *Μέθοδοι και Τεχνικές Ανάλυσης στην Ποιοτική Κοινωνική Έρευνα*. Αθήνα: Κριτική.