

Πολυγλωσσικοί Θησαυροί Όρων : Ζητήματα Τεκμηρίωσης και Αναζήτησης

Χρύσα Καππή
Δήμητρα Κονδύλη
Χριστίνα Φρέντζου

ΕΚΚΕ Κείμενα Εργασίας 2001/16

*Χρύσα Καππή, κοινωνική ψυχολόγος, ερευνήτρια, Ινστιτούτο Κοινωνικής Πολιτικής
Δήμητρα Κονδύλη, κοινωνιολόγος, Γραφείο Διαμεσολάβησης
Χριστίνα Φρέντζου, πληροφορικός, Διεύθυνση Υποστήριξης Ερευνών*

[...] *οι γλώσσες, οι λεγόμενες περιορισμένης χρήσης, δεν πρέπει να καταντούν φράκτες όπου πνίγεται ο παλμός της ανθρώπινης καρδιάς*

Γιώργος Σεφέρης

1 ΕΙΣΑΓΩΓΗ

Η επίδραση των τεχνολογικών εξελίξεων είναι σημαντική ως προς την παραγωγή και τη μετάδοσή της γνώσης. Είναι λάθος να χωρίζουμε την παραγωγή της γνώσης (έρευνα, θεωρητική παραγωγή) από τη μετάδοσή της (πληροφορία). Ο πολλαπλασιασμός τόσο των τεχνολογικών μέσων όσο και της ανάγκης για πληροφορία, ως οντότητα γνώσης και δύναμης, επηρεάζει την κυκλοφορία των γνώσεων τόσο από χρονική όσο και από χωρική άποψη. Αυτό το γεγονός αποτελεί ένα μετασχηματισμό για τον οποίο εγείρονται πολλά ζητήματα, ένα από τα οποία είναι η αποτελεσματικότητα της γνώσης. Ο Lyotard, το 1979, έγραφε: « η γνώση δεν μπορεί να είναι αποτελεσματική παρά μόνο αν μετατραπεί σε δομημένες ομάδες πληροφόρησης [...] όλα όσα δεν έχουν τη δυνατότητα να μεταφραστούν έτσι μέσα στη συγκροτημένη γνώση, θα εγκαταλειφθούν και ο προσανατολισμός των νέων ερευνών θα υπαχθεί στην προϋπόθεση του να μπορούν να μεταφράζονται τα ενδεχόμενα αποτελέσματα στη γλώσσα της μηχανής. Οι 'παραγωγοί' της γνώσης καθώς και οι χρήστες της οφείλουν και θα πρέπει να έχουν τα μέσα να μεταφράσουν σε αυτές τις γλώσσες ό,τι οι μεν ζητούν να επινοήσουν και οι δε να μάθουν» [Lyotard, (1988), σελ.32]. Αυτή η προσέγγιση περιγράφει, ίσως, μια μελλοντική κατάσταση, καταδεικνύει όμως και μια τάση προς την οριοθέτηση προδιαγραφών για τη μετάδοση των προϊόντων της γνώσης, η οποία δεν είναι άγνωστη στις μέρες μας. Αντιθέτως, αποτελεί προτεραιότητα η διαδικασία οριοθέτησης των προδιαγραφών μετάδοσης του περιεχομένου, μέσα σε ένα φάσμα που προσδιορίζεται από την απλή ενημέρωση έως το εξειδικευμένο προϊόν της ερευνητικής ή/και θεωρητικής παραγωγής. Καθίσταται δε αναγκαία η συνέργεια πολλών επιστημονικών κλάδων που συμβάλλουν στην ανάπτυξη μεθόδων και εργαλείων προς την κατεύθυνση αυτή.

Τα συνεχώς αυξανόμενα αιτήματα πρόσβασης σε πληροφορίες και δεδομένα από μεγάλο αριθμό χρηστών, διεθνώς, απορρέει αφ' ενός από την άρση πολιτικών και κοινωνικών φραγμών, λόγω της ενοποίησης κοινωνικών σχηματισμών (π.χ. ΕΕ) και αφ' ετέρου από την παγκοσμιοποίηση της βιομηχανίας και του εμπορίου. Τα άμεσα αποτελέσματα αυτών των εξελίξεων είναι, ανάμεσα σε πολλά άλλα, η ανάγκη σχεδιασμού πολιτικών και λήψης αποφάσεων σε ευρεία κλίμακα, οι προσπάθειες εφαρμογής πολιτικών ρυθμίσεων σε διαφορετικές πληθυσμιακές ομάδες, η διεύρυνση των εκπαιδευτικών ευκαιριών. Είναι φανερό πως τέτοιας κλίμακας μετασχηματισμοί συναντούν εμπόδια στο επίπεδο του σχεδιασμού, πόσο μάλλον στο επίπεδο της εφαρμογής. Ως εκ τούτου, και αναφορικά με τη μετάδοση της γνώσης, η προϋπόθεση που θέτει ο Lyotard –να μεταφράζονται τα αποτελέσματα της γνώσης στη γλώσσα της μηχανής- αφορά ένα συνεχώς αυξανόμενο πλήθος 'παραγωγών' και 'χρηστών'.

Η διαδικασία της μετάβασης από τη γνώση στην πληροφόρηση, δηλαδή τη δημοσιοποίησή της, κάτω από το πρίσμα των νέων εξελίξεων αφορά, αναπόφευκτα, και την τροποποίηση ή εξέλιξη των μέσων επικοινωνίας. Όσο και αν αυτό φαντάζει αυτονόητο, στην πράξη έχει ξεπεραστεί ένα μικρό μόνο μέρος των δυσκολιών που αφορούν τη μετατροπή της γνώσης σε «δομημένες ομάδες πληροφόρησης». Οι

ιδιότητες της δομής (δομημένη πληροφόρηση) και της ταξινόμησης (ομάδες πληροφόρησης) αφορούν κυρίως τα αποτελέσματα της ερευνητικής παραγωγής, αν και, όπως θα δούμε στη συνέχεια, αποτελούν πρότυπο για θεωρητικά υποδείγματα.

Στην εργασία αυτή θα επιχειρήσουμε μια παρουσίαση κάποιων ζητημάτων που αφορούν τη μετάδοση της γνώσης που προκύπτει από την ερευνητική διαδικασία των κοινωνικών επιστημών κάτω από το πρίσμα της εξέλιξης της τεχνολογίας. Ειδικότερα, θα αναφερθούμε σε ζητήματα της τεκμηρίωσης και της αναζήτησης των αποτελεσμάτων της ερευνητικής παραγωγής. Αφετηρία για αυτή την εργασία αποτελεί ο προβληματισμός μας στο ζήτημα των *γλωσσικών κωδίκων επικοινωνίας*, οι οποίοι αποτελούν το πεδίο σύνδεσης *μεταξύ* ανθρώπων με κοινή γλώσσα στους παραδοσιακούς τρόπους επικοινωνίας, ενώ συγχρόνως αναδεικνύεται η ανάγκη εξέλιξής τους για να αποτελέσουν το πεδίο σύνδεσης ενός αυξανόμενου αριθμού ανθρώπων με διαφορετικές γλώσσες, *μέσω* των μηχανών (υπολογιστές).

2 ΓΝΩΣΗ ΚΑΙ ΠΛΗΡΟΦΟΡΙΑ

Ο ανθρώπινος νους οδηγείται στη γνωστική εμπειρία μέσα από μια αδιάκοπη διαδικασία σύνθεσης και ανάλυσης ερεθισμάτων –προερχόμενα είτε από το εξωτερικό είτε από το εσωτερικό περιβάλλον. Το αποτέλεσμα της διαδικασίας αυτής, όπως και αν ορίζεται ή ερμηνεύεται -φιλοσοφικά ή επιστημονικά- εφ’ όσον είναι μεταβιβάσιμο προς το εξωτερικό περιβάλλον με ένα τουλάχιστον σύστημα επικοινωνίας του ανθρώπου, μπορεί να αναγνωρισθεί ως γνώση. Ο μηχανισμός απόκτησης της γνώσης αποκτά επιστημονική και φιλοσοφική σημασία στο βαθμό που αποτελεί το αντικείμενο της διερεύνησης ή της πραγματείας. Όποια όμως και αν είναι η οπτική της προσέγγισης, η συλλογιστική υπαγορεύεται από τη λειτουργικότητα που ενέχει η απάντηση, με άλλα λόγια, ο τελικός στόχος είναι το αποτέλεσμα της γνώσης. Στην περίπτωση της επιστήμης το αποτέλεσμα της γνώσης είναι η διατύπωση των γεγονότων που αποκαλύπτονται, αποδεικνύονται ή διαψεύδονται μέσω των επιστημονικών μεθόδων.

Τα κρίσιμα ζητήματα στην περίπτωση της επιστημονικής γνώσης είναι όχι μόνον η μεθόδευση για την αποκάλυψη των γεγονότων, αλλά και η διαχείριση των αποδείξεων της έρευνας (Lyotard, 1988). Η ικανότητα διαχείρισης των αποδείξεων βρίσκεται σε άμεση συνάρτηση με την ύπαρξη τεχνικών κριτηρίων, τα οποία εν τέλει επηρεάζουν το κριτήριο της αλήθειας. Έτσι, πραγματικότητα καθίσταται ό,τι στηρίζεται στην επιστημονική επιχειρηματολογία. Ό,τι αποτελεί γνώση μέσα από αυτή τη διαδικασία αποτελεί ίσως και τη μόνη μορφή αλήθειας για την επιστήμη. Αυτό προϋποθέτει έναν μηχανισμό διατύπωσης και μετάδοσης της επιστημονικής γνώσης. Η ίδια η επιστημονική γνώση δεν είναι ποτέ δυνατόν να μεταδοθεί και να κατανοηθεί με τον ίδιο τρόπο που είναι αντιληπτή στον φορέα της γνώσης (ειδήμονα), αλλά διατυπώνεται, δηλαδή σχηματοποιείται, σε ένα *δομημένο σύστημα πληροφορίας*. Σύμφωνα με αυτή την οπτική, η γνώση βρίσκεται σε ένα πιο αφηρημένο επίπεδο από την πληροφορία· η πληροφορία είναι η *μετα-γνώση*, η έκφραση, η αναγνώριση της γνώσης. Είναι το λειτουργικό αποτέλεσμα της γνώσης, αλλά δεν αποτελεί οντότητα παρά μόνο όταν βρεθεί μέσα σε ένα σύστημα *πομπού* και *αποδέκτη*. Για να λειτουργήσει το σύστημα αυτό, είναι απαραίτητη η διαμεσολάβηση ενός *συστήματος διαχείρισης συμβόλων*, ενός κοινού κώδικα επικοινωνίας.

Πώς όμως γίνεται αντιληπτό ένα *δομημένο σύστημα πληροφορίας*, όταν το γεγονός της διαμεσολάβησης ενός κώδικα επικοινωνίας τροποποιεί, εξ ορισμού, το αρχικό ερέθισμα; Η απάντηση στο ερώτημα αυτό απασχολεί μεγάλη μερίδα φιλοσόφων, οι οποίοι στη συζήτηση περί της γλώσσας, ως τον πιο εξελιγμένο κώδικα επικοινωνίας,

υποστηρίζουν πως, τελικά, εργαλείο για τη μετάδοση της δομής της πληροφορίας αποτελεί η γλώσσα, η δε δομή κατανοείται –ή δεν κατανοείται- πέρα από το λεκτικό επίπεδο γίνεται γνώση, εμπειρία, αντίληψη, αναζήτηση (Ayer,1956). Αντίθετα, στα μοντέλα ερμηνείας που προσέφεραν οι πρώτες επιστημονικές θεωρίες της ανθρώπινης νόησης, αγνοούσαν συστηματικά τη διάσταση της δομής στο επίπεδο του ερεθίσματος από το εξωτερικό περιβάλλον. Σε αυτό συνέβαλλε δραματικά τόσο ο μπεχαβιορισμός όσο και οι μεταλλάξεις του, που οδήγησαν την επιστημονική σκέψη προς ένα παράδειγμα (paradigm) περί λειτουργίας της ανθρώπινης νόησης κατ'αναλογία των τεχνολογικών κατασκευών της εποχής εκείνης. Στο περιθώριο όμως του “συρμού” αναδύθηκαν θεωρητικοί όπως ο Chomsky, ο Piaget και άλλοι, που έστρεψαν την προσοχή της επιστήμης σε –ανερμήνευτες, ακόμη- λειτουργίες της νόησης, με βάση την αντίληψη μιας ενυπάρχουσας δομής μέσα στο ερέθισμα. Σ' αυτή την περίπτωση η δομή δεν αποτελεί οντότητα αλλά ερμηνευτικό σχήμα, χωρίς το οποίο καταρρέει το επιστημονικό μοντέλο.

Η Gibson εξέφρασε με πιο σαφή τρόπο αυτό το ερμηνευτικό σχήμα, αναπτύσσοντας τη θεωρία της συναλλαγής της ανθρώπινης νόησης με ένα δομημένο κόσμο ερεθισμάτων προς τον οποίο στρέφεται με επιλεκτική προσοχή, ανάλογα με τους στόχους, για να εκμαιεύσει την πληροφορία που βρίσκεται διαθέσιμη [Miller (1992), κεφ.6]. Σύμφωνα με τη συλλογιστική της Gibson, η πληροφορία αποκτά οντότητα διαμέσου της δομής που κατανοείται εμπειρικά και, αυτό που τελικά μεταβιβάζεται από το εξωτερικό περιβάλλον, δεν είναι μονάδες πληροφορίας αλλά ένα σύστημα νοημάτων. Στην πληροφορική τεχνολογία αυτό το “μοντέλο” επικοινωνίας παραμένει ανεφάρμοστο, ενώ η επικοινωνία ανθρώπου-μηχανής στο επίπεδο δομημένων γλωσσικών κωδικών έχει επιτευχθεί.

2.1 Κώδικες επικοινωνίας

Οι τρόποι μετάδοσης της γνώσης έχουν πάντοτε μια χωροχρονική διάσταση και είναι εφικτοί με τη χρήση εργαλείων. Η γλώσσα αποτελεί εργαλείο με την ευρύτερη έννοια και συμβάλλει στην εξέλιξη της νόησης, επειδή σχηματοποιεί τη σκέψη και ενισχύει την επικοινωνία [Vygotsky, στο: Miller (1992), κεφ.7]. Γλώσσα αποτελεί κάθε είδους κώδικας που είναι αποδεκτός και κατανοητός από δύο τουλάχιστον οντότητες. Ο λόγος -προφορικός και γραπτός- αποτελεί τον πιο εξελιγμένο κώδικα επικοινωνίας μεταξύ των ανθρώπων.

Ο μηχανισμός ανάπτυξης των γλωσσικών κωδικών αποτελεί αντικείμενο έρευνας στα επιστημονικά πεδία της ψυχολογίας, κοινωνιολογίας, γλωσσολογίας, ανθρωπολογίας, ιατρικής, αλλά και της πληροφορικής. Η είσοδος των υπολογιστών ως εργαλείων επικοινωνίας, εισήγαγε μια νέα διάσταση στην επικοινωνία, ξεπερνώντας τον εργαλειακό τους ρόλο. Η δυνατότητα αυτής της ‘νέου τύπου’ επικοινωνίας επιφέρει σταδιακά σημαντικές αλλαγές στον τρόπο σκέψης και –κατ'επέκτασιν- στον τρόπο επικοινωνίας μεταξύ των μελών του σύγχρονου πολιτισμού. Έτσι, μετεξελίσσονται οι παραδοσιακοί γλωσσικοί κώδικες. Μία ένδειξη αυτής της εξέλιξης είναι η ‘γλωσσική υστέρηση’ που παρατηρείται, για παράδειγμα, στην ελληνική γλώσσα ως προς την ορολογία της πληροφορικής επιστήμης. Λόγω της υστέρησης αυτής, αναγκαστικά, εισάγονται αναγκαστικά νέοι όροι, οι οποίοι είτε χρησιμοποιούνται αυτούσιοι είτε αφομοιώνονται ηχητικά στην ελληνική. Μια παρόμοια εξέλιξη, η οποία σήμερα θεωρείται αυτονόητη, αναδεικνύεται με τη συνεχή ανάπτυξη της επιστημονικής γνώσης. Πρόκειται για τη δημιουργία “αξιοματικής γλώσσας”¹, η οποία έχει τη

¹ Ο όρος αναφέρεται από τον Lyotard.

μορφή ενός κλειστού κώδικα επικοινωνίας στα όρια της ακριβούς επιστημονικής και τεχνικής σκέψης που μοιράζονται οι ‘μυημένοι’ του κάθε κλάδου. Για να είναι, βέβαια, μεταδίδσιμη η αξιωματική γλώσσα είναι απαραίτητη η εκ των προτέρων ύπαρξη της “μετα-γλώσσας”² -της γλώσσας της λογικής- που προσδιορίζει τις συνθήκες της αξιωματικής.

Η εξέλιξη της πληροφορικής τεχνολογίας και κυρίως του διαδικτύου, ανέδειξε ως μέγιστο περιορισμό στην παγκόσμια επικοινωνία την ύπαρξη ποικιλίας γλωσσικών συστημάτων, με αποτέλεσμα την έλλειψη κοινώς αποδεκτών κωδίκων για τον προσδιορισμό επιστημονικών και τεχνικών όρων. Περαιτέρω, δημιουργήθηκε η ανάγκη εφαρμογής κοινών προδιαγραφών και τυποποιήσεων για έναν μεγάλο αριθμό ενεργειών στο πλαίσιο της χωρίς σύνορα επικοινωνίας. Η διεύρυνση των δυνατοτήτων επικοινωνίας με αυτό το μέσο αναδεικνύει όλο και πιο έντονα το εξής παράδοξο: τη σύγκλιση προς μία κοινή γλώσσα, η οποία με τα σημερινά δεδομένα συμπίπτει με τη γλώσσα των κοινωνιών εκείνων που πρωτοπορούν στη νέα τεχνολογία και, συγχρόνως, την αυξανόμενη ανάγκη για ερμηνεία και απόδοση εννοιών, οι οποίες υποκρύπτουν πολιτισμικές ιδιαιτερότητες και, ενδεχομένως, διαφορές που προκύπτουν από τα διαφορετικά πεδία θεώρησης των επιστημών.

Είναι επομένως αναγκαία η ανάπτυξη κοινώς αποδεκτών *ειδικών κωδίκων* –της αξιωματικής γλώσσας- και *τυπολογιών*, που θα έχουν τη μέγιστη δυνατή σαφήνεια ως προς το αντικείμενο το οποίο προσδιορίζουν και την ευρύτερη δυνατή συμφωνία της σχετικής με το αντικείμενο κοινότητας. Οι κώδικες αυτοί δεν μπορούν να είναι αυθύπαρκτοι αλλά αποκτούν νόημα με την επικουρία του *γενικού κώδικα* –της φυσικής γλώσσας- η οποία συγκρίνει, ερμηνεύει, οριοθετεί και ορίζει το περιεχόμενο. Ο ειδικός κώδικας δεν είναι εργαλείο που λειτουργεί ανεξάρτητα από τη νόηση, δεν είναι το ανάλογο των εργαλείων της φυσικής επιστήμης (π.χ. μέτρα και σταθμά), αλλά αποτελεί την οριοθετημένη έκφραση της επιστημονικής σκέψης. Όπως χαρακτηριστικά υποστηρίζει ο Vygotsky, οι επιστημονικές έννοιες είναι λογικά προσδιορισμένες έννοιες, που εμπεριέχουν κοινωνικό και επιστημονικό περιεχόμενοˆ συναντώνται μέσα από τους επίσημους μηχανισμούς μάθησης και εσωτερικεύονται με σύγχρονη ανάμειξη των αυθόρμητων εννοιών [Vygotsky, στο: Miller (1992), κεφ.7].

2.2 Εισαγωγή στο πρόβλημα του λεξιλογίου

Όσο η γλώσσα εξαρτάται από το νοητικό μας σύστημα άλλο τόσο το νοητικό σύστημα λειτουργεί ανεξάρτητα από τη γλώσσα. Οι περιορισμοί στην επικοινωνία προκύπτουν, επειδή η γλώσσα αποτελεί το μόνο μέσο διατύπωσης των εννοιών. Οι λέξεις αποτελούν τον ‘καθρέπτη’ των νοημάτων. Σύμφωνα με την οπτική ορισμένων φιλοσοφικών θεωρήσεων, οι ειδικοί (κλειστοί) κώδικες συνθέτουν έναν “ιδιωτικό λόγο” σε αντίθεση με τον “δημόσιο λόγο”, τη φυσική, καθημερινή γλώσσα. Είναι απαραίτητοι μόνο στον βαθμό που βοηθούν στην αποσαφήνιση εσωτερικευμένων εννοιών και πρέπει να λειτουργούν ως διαμεσολαβητικό μέσο για την έκφραση και μετάδοσή τους. (Ayer,1956). Γι’ αυτό προσλαμβάνουν το χαρακτηριστικό της καθολικότητας.

Η καθολικότητα στην επικοινωνία μέσω ειδικών κωδίκων –ή της αξιωματικής γλώσσας- αποτέλεσε πρόβλημα και στις φυσικές επιστήμες. Όμως, τα σύμβολα, οι μαθηματικές εκφράσεις και οι *λειτουργικοί ορισμοί* έγιναν κοινή γνώση των

² Ο όρος αναφέρεται από τον Lyotard

μυημένων στις επιστήμες αυτές. Στις κοινωνικές επιστήμες, το ζήτημα προέκυψε πρώτον, με τη μετεξέλιξη των μεθόδων των κοινωνικών επιστημών σε μεθόδους ακριβείας, όπου η ποσοτικοποίηση κοινωνικών φαινομένων είχε νόημα και, δεύτερον, με την ανάπτυξη της πληροφορικής τεχνολογίας, η οποία οδήγησε, λίγο ως πολύ, στην ενοποίηση των μέσων της επιστήμης. Η καθολικότητα όμως σπάνια χαρακτηρίζει τα κοινωνικά φαινόμενα· επιπροσθέτως, δεν μπορεί να υποστηρίξει κανείς ότι οι μέθοδοι επιστημονικής ανάλυσης και η απορρέουσα γνώση των κοινωνικών φαινομένων απολαμβάνουν την καθολική αποδοχή της παγκόσμιας επιστημονικής κοινότητας. Έτσι, όταν, για παράδειγμα, αναφερόμαστε στο ρατσισμό ξέρουμε ότι δεν εννοούμε το ίδιο πράγμα όταν αναφερόμαστε σε διαφορετικές χώρες και σε διαφορετικές χρονικές περιόδους. Παρ'όλα αυτά ο ρατσισμός είναι ένα φαινόμενο που ενδιαφέρει όλους, κοινωνικά και επιστημονικά και, όσο περισσότερες θεωρητικές οπτικές και πραγματικά δεδομένα είναι διαθέσιμα, τόσο θα βελτιώνεται η αντίληψή μας για το φαινόμενο αυτό. Σ' αυτή την περίπτωση, οι πολιτισμικές διαστάσεις που απορρέουν τόσο από τη χωροχρονική εμπειρία, όσο και από τις γλωσσικές ιδιαιτερότητες, πρέπει να αναδειχθούν με τη βοήθεια μιας καθολικά αποδεκτής ειδικής γλώσσας και της τυπολογίας της. Αυτή η προοπτική συμβάλλει στη διάχυση της γνώσης, όσο και αν φαίνεται αντιφατικό το γεγονός ότι ένας κλειστός κώδικας –μια λειτουργική τυποποίηση– η οποία κατά κανόνα εξυπηρετεί τους λίγους, χρησιμοποιείται ως εργαλείο επικοινωνίας.

3 ΛΕΙΤΟΥΡΓΙΚΗ ΤΥΠΟΠΟΙΗΣΗ: Η ΣΥΡΡΙΚΝΩΣΗ ΤΟΥ ΝΟΗΜΑΤΟΣ Ή Ο ΘΡΙΑΜΒΟΣ ΤΗΣ ΓΛΩΣΣΑΣ;

Όπως έγινε ήδη φανερό από την ανάπτυξη που προηγήθηκε, η δομή αποτελεί απαραίτητο συστατικό στοιχείο της επικοινωνίας, εφ'όσον και όπου η γλώσσα χρησιμοποιείται ως μέσον. Η καταγραφή και συστηματοποίηση των γλωσσικών στοιχείων αποτελεί έναν πρωταρχικό τύπο “γραφής”, με την έννοια που προσδίδει στον όρο ο Derrida (Ayer, 1956) δηλαδή, αρχέγονη, προ-γραφή –ένα κωδικοποιημένο σύστημα πολιτισμικών συμβόλων, γραμμάτων, γραμματικών κανόνων και ταξινομήσεων, μέσω των οποίων ο λόγος και η “εμπειρική γραφή” (το γραπτό κείμενο) μορφοποιούνται. Αυτή η μορφή γραφής περιγράφει με τον πιο γλαφυρό, ίσως, τρόπο την επικοινωνιακή λογική της γλώσσας. Προϊόντα αυτής της λογικής είναι τα λεξικά, τα κάθε μορφής ευρετήρια, οι θησαυροί, που δημιουργούνται για λειτουργικούς σκοπούς, για να εξυπηρετούν την εμπειρική γραφή. Όμως, αυτό που συμβαίνει στην πράξη με την ταξινόμηση της γλώσσας, είναι αυτό που ο Foucault ονομάζει “γνωσιολογική τυραννία” δηλαδή, η συρρίκνωση των νοημάτων προς χάριν της τυποποίησης και, τελικά, η απομάκρυνση της “γραφής” από τον *Λόγο*, σήμερα μάλιστα, περισσότερο από ποτέ άλλοτε, που οι ανθρώπινες κοινωνίες αντιμετωπίζουν την πρόκληση να καλλιεργήσουν τις συνθήκες ισορροπίας μεταξύ του *Λόγου* και της *κοινής λογικής*.

3.1 Λεξικό, ευρετήριο, θησαυρός: Διακριτές οντότητες

Στην καθημερινή χρήση της γλώσσας είναι πολλές φορές δύσκολο να ξεχωρίσει κανείς τις εννοιολογικές μικροδιαφορές μεταξύ λέξεων, οι οποίες τυχαίνει να χρησιμοποιούνται με λανθασμένους εννοιολογικούς προσδιορισμούς. Τα εργαλεία καταγραφής των γλωσσικών στοιχείων δημιουργούνται με σκοπό να προσδιορίζουν με σαφήνεια τη σωστή χρήση της γλώσσας και, όπως κάθε εργαλείο, εξειδικεύονται ανάλογα με τη λειτουργία που επιτελούν. Έτσι, υπάρχει σαφής λειτουργική διάκριση μεταξύ των λεξικών, των ευρετηρίων και των θησαυρών.

3.1.1 Λεξικό

Το λεξικό είναι η συστηματοποιημένη καταγραφή των λέξεων, του τρόπου γραφής τους, των συνηθέστερων ερμηνειών και χρήσεών τους, των κυριότερων γραμματικών και συντακτικών αναφορών που σχετίζονται με αυτές και αφορά μία γλώσσα, ένα ανοικτό σύνολο λέξεων. Αφορά τον *δημόσιο λόγο* -την καθημερινή γλώσσα. Το λεξικό δεν εμπεριέχει ταξινομική δομή. Οι αναφορές σε εξειδικευμένους όρους είναι περιορισμένες· συνήθως τέτοιοι όροι περιλαμβάνονται όταν η χρήση τους έχει μια ευρύτερη συχνότητα (π.χ. συνήθεις ιατρικοί όροι). Το λεξικό, επειδή αφορά τη συνήθη χρήση της γλώσσας, έχει δυναμικές ιδιότητες, είναι δηλαδή ανανεώσιμο, απαιτεί επιστημονική έρευνα για την εξέλιξή του, αλλά είναι ανεπαρκές για χρήση στην επιστημονική κοινότητα.

3.1.2 Ευρετήριο

Το ευρετήριο είναι ένας κατάλογος λέξεων και όρων, που αφορά μεν ένα συγκεκριμένο και πεπερασμένο κείμενο, εξασφαλίζει δε την 'πλοήγηση' μέσα σε αυτό. Τα ευρετήρια αποτελούν αναπόσπαστο μέρος των βιβλίων, βοηθούν στην αναζήτηση λέξεων και όρων που αναφέρονται στο περιεχόμενο και, σε πολλές περιπτώσεις, περιέχουν αναφορές για τις σχέσεις μεταξύ των όρων. Οποιοδήποτε κείμενο μπορεί να αποκτήσει ευρετήριο, αλλά η αξία του κάθε ευρετηρίου αφορά μόνο το συγκεκριμένο κείμενο. Το ευρετήριο αναφέρεται στην ταξινομική διαδικασία αναζήτησης κειμένων, βιβλίων, δεδομένων και ως εκ τούτου ένα ευρετήριο είναι πεπερασμένο και στατικό.

3.1.3 Θησαυρός

Ο θησαυρός αποτελεί ένα εξειδικευμένο αναλυτικό εργαλείο, που είναι δομημένο σε τρία επίπεδα: το επίπεδο των λέξεων-όρων-φράσεων, το επίπεδο των επιτρεπόμενων χρήσεων με τους περιορισμούς τους και το επίπεδο της δομής και των σχέσεων των λέξεων-όρων μεταξύ τους. Οι βασικές συνθήκες που πρέπει να ικανοποιεί ένας θησαυρός είναι:

- α) η συνεχής ανανέωση,
- β) η προσαρμογή σε μεγάλο αριθμό χρηστών και στους ρυθμούς αύξησής τους,
- γ) η ταξινόμηση της πληροφορίας σε ευρετηριακό σχήμα ώστε να διευκολύνει την αναζήτηση,
- δ) η μετατροπή του ερωτήματος σε αποσαφηνισμένο εννοιολογικό σχήμα,
- ε) η καλύτερη δυνατή αντιπροσώπευση/ αναπαράσταση του περιεχομένου,
- στ) η οργάνωση και η παρουσίαση της δομής και των σχέσεων των εννοιών,
- ζ) η σύνδεση της απλής γλώσσας με την αξιωματική,
- η) η μετατροπή της περιορισμένης, τοπικής ή ιδιωτικής ορολογίας σε δημόσια ορολογία.

Εξάλλου, όσον αφορά τη χρήση του θησαυρού μέσω της πληροφορικής, θα πρέπει να εξασφαλίζονται:

- θ) ο αυτοματισμός των παραπάνω ενεργειών και
- ι) η δυνατότητα μετάφρασης σε πολλές γλώσσες χωρίς απώλεια ή παραποίηση του εννοιολογικού περιεχομένου.

Ο θησαυρός είναι ένα δυναμικό εργαλείο και μια από τις σημαντικές του χρήσεις είναι ο σαφής εννοιολογικός προσδιορισμός των λέξεων –γι'αυτό και στους θησαυρούς αναφερόμαστε σε *όρους* και όχι σε *λέξεις*. Για τον εννοιολογικό προσδιορισμό των όρων χρησιμοποιείται η γνώση ανά επιστημονικό τομέα, στα όρια

ενός ελεγχόμενου λεξιλογίου (controlled vocabulary). Ως εκ τούτου οι θησαυροί εξυπηρετούν αποκλειστικά τον επιστημονικό χώρο για τον οποίο δημιουργήθηκαν. Η American Society of Indexers (ASI)³ αναφέρει ότι όλοι οι υπάρχοντες δημοσιευμένοι θησαυροί έχουν αναπτυχθεί για να εξυπηρετήσουν τις ανάγκες των ειδικών βάσεων για τις οποίες αναπτύχθηκαν. Συστήνουν δε να μην υιοθετείται αυτούσιος ένας θησαυρός αν η χρήση του δεν έχει τον ίδιο στόχο και δεν αναφέρεται στο ίδιο επίπεδο αναζήτησης. Είναι σαφές λοιπόν η ιδιότητα της εξειδίκευσης.

Στην εποχή μας, με την ανάπτυξη της υπολογιστικής τεχνολογίας και της επικοινωνίας μέσω του διαδικτύου, υπάρχουν ήδη διαθέσιμοι θησαυροί για τους περισσότερους επιστημονικούς τομείς. Ένας κατάλογος των υπαρχόντων θησαυρών στο διαδίκτυο διατίθεται από τον οδηγό πηγών που δημοσιεύει το Τεχνολογικό Πανεπιστήμιο Queensland. Όμως, παρά την εξειδίκευση, υπάρχουν ποιοτικές διαφορές μεταξύ των θησαυρών, οι οποίες απορρέουν αφ' ενός από τον διαφορετικό βαθμό συσσώρευσης εμπειρίας αφ' ετέρου από τις γλωσσικές ιδιαιτερότητες των διαφόρων επιστημονικών κλάδων. Επανερχόμαστε έτσι στο ζήτημα της σχέσης μεταξύ γλώσσας και εννοιών. Όπως χαρακτηριστικά αναφέρεται από την ASI ο βαθμός εξυπηρέτησης των επιστημονικών πεδίων ποικίλλει ανάλογα με το επίπεδο κωδικοποίησης των επιστημονικών κλάδων. Οι θησαυροί των θετικών επιστημών και της τεχνολογίας, με εξαίρεση αυτόν της Τέχνης και Αρχιτεκτονικής, έχουν σαφώς καλύτερη κάλυψη από τους θησαυρούς των ανθρωπιστικών επιστημών, ενώ οι θησαυροί για τις κοινωνικές επιστήμες καλύπτουν σχετικά ικανοποιητικά τους κλάδους, με περιθώρια βελτίωσης (ASI, 2001).

3.2 Η σύγκλιση τεχνητής & φυσικής γλώσσας: Θησαυροί

Τα τελευταία χρόνια, με την ανάπτυξη της *Υπολογιστικής Γλωσσολογίας*, έχουν δημιουργηθεί στατιστικά μοντέλα επεξεργασίας της φυσικής γλώσσας (στατιστική υπολογιστική γλωσσολογία), τα οποία αφορούν τη μελέτη των γλωσσικών φαινομένων μέσα από τη στατιστική ανάλυση μεγάλων συλλογών κειμένου, με στόχο την κατανόηση –από υπολογιστές– της δομής της γλώσσας με μεθόδους συσχέτισης και επαγωγικές διαδικασίες (Νικολαΐδης, 2001). Αφετηρία της υπολογιστικής γλωσσολογίας είναι οι εσωτερικευμένοι κανόνες της ανθρώπινης γλώσσας και στόχος η αποκωδικοποίηση αυτών των κανόνων ώστε να γίνουν ‘κτήμα’ των υπολογιστών. Είναι ένας μακρινός και αμφιβόλου προοπτικής στόχος, αλλά ήδη τα ενδιάμεσα στάδια έρευνας έχουν προσφέρει εντυπωσιακές εφαρμογές ‘έξυπνου’ λογισμικού για μια πιο ουσιαστική επικοινωνία ανθρώπου-μηχανής.

Τέτοιου τύπου μελέτες, ακόμη και αν δεν έχουν ικανοποιητικά αποτελέσματα ως προς τους τελικούς τους στόχους, δημιουργούν νέες κατευθύνσεις στη διεπιστημονική συνεργασία για τη βελτιστοποίηση των τεχνολογικών εφαρμογών στο ζήτημα της επικοινωνίας μέσω υπολογιστών. Στο θέμα των ηλεκτρονικών θησαυρών βρισκόμαστε ακόμη στην αρχή της πορείας. Οι στατιστικοί θησαυροί στη σημερινή φάση της εξέλιξής τους στηρίζονται στη συχνότητα εμφάνισης των σχετικών όρων, αλλά δεν έχουν λύσει το πρόβλημα της κοινής αποδοχής του νοήματος που φέρει ο κάθε όρος. Σύμφωνα με την E.D. Liddy η εξέλιξη από την ανθρώπινη στη στατιστικά ελεγχόμενη ταξινόμηση όρων παρουσιάζει τα εξής στάδια: από τα εγχειρίδια ελεγχόμενων λεξιλογίων, στην κοινή πρωτοβουλία ανθρώπου-μηχανής για την εκμάθηση της γλώσσας από την τελευταία και, τέλος, στην αυτοματοποιημένη

³ American Society of Indexers (ASI) [Αμερικανική Εταιρεία Συντακτών Ευρετηρίων, (η μετάφραση είναι των συγγραφέων)]

στατιστική επεξεργασία για την ταξινόμηση μέσω της φυσικής γλώσσας. Το κυριότερο πρόβλημα στην πορεία αυτή εντοπίζεται στην εννοιολογική αποσαφήνιση, η οποία δεν μπορεί να επιτευχθεί αυτόματα (Liddy, 1999). Η εννοιολογική αποσαφήνιση είναι δύσκολο να επιτευχθεί ακόμη και στον φυσικό λόγο της ανθρώπινης επικοινωνίας και ειδικότερα στη διαδικασία της *ερωτηματοθεσίας*, η οποία αποτελεί μια “εκλεπτυσμένη χρήση του γλωσσικού παιχνιδιού”: Ποιός είναι ο κατάλληλος διάυλος για εκείνο που θέλουμε να μάθουμε; Ποιό είναι το πιο αποτελεσματικό εργαλείο; Πώς να διατυπώσουμε τον λόγο για να αποφύγουμε τις παρεξηγήσεις; (Lyotard, 1988)

4 ΑΝΑΖΗΤΗΣΗ: Η ΜΗΧΑΝΙΚΗ ΤΗΣ ΕΡΩΤΗΜΑΤΟΘΕΣΙΑΣ

Η διευρυνόμενη πολυπλοκότητα του είδους και της μορφής της πληροφορίας και η δυνατότητα αποθήκευσης μεγάλου αριθμού πληροφοριών κάνει αναγκαία τη διαμόρφωση μιας νέας προσέγγισης στην αναζήτηση, που θα είναι *απαλλαγμένη από φράγματα*: τεχνικά, γλωσσικά, πολιτισμικά, επιστημονικής πειθαρχίας. Επιπλέον, η συσσώρευση των πληροφοριών και η οργάνωσή τους απαιτεί την ύπαρξη ενός διευρυνμένου επιπέδου οργάνωσης “της πληροφορίας για την πληροφορία”. Ένα νέο μοντέλο αναζήτησης καλείται να καλύψει αυτές ακριβώς τις ανάγκες. Οι αρχές σχεδιασμού του είναι απαραίτητο να διευκολύνουν τη μετατροπή της πληροφορίας σε ερώτημα μέσω ενός λεξιλογίου το οποίο θα ικανοποιεί τρία επίπεδα:

α) το επίπεδο της *απλής γλώσσας*, για μια *βιβλιοθηκονομικού τύπου* ταξινόμηση (σε αυτόν τον τύπο βασίζονται τα λεξικά αναζήτησης του διαδικτύου)

β) το επίπεδο της *αξιοματικής γλώσσας*, για την ταξινόμηση *εξειδικευμένων όρων* (οι θησαυροί που έχουν αναπτυχθεί για λογαριασμό των διαφόρων επιστημονικών κλάδων καλύπτουν, ως ένα βαθμό, αυτόν τον τύπο αναζήτησης)

γ) το επίπεδο του *παρατηρησιακού κώδικα* για την ταξινόμηση όσων πληροφοριών δεν είναι δυνατόν να εντοπισθούν και να διευκρινισθούν εννοιολογικά, χωρίς την περιγραφή των θεωρητικών και πρακτικών σκοπών για τους οποίους δημιουργήθηκαν. Οι μεν θεωρητικοί σκοποί αποτελούν την «οπτική μέσα από την οποία “βλέπει” την πραγματικότητα η κάθε επιστήμη, ορίζει δηλαδή τα φαινόμενα που ανήκουν στο αντικείμενό της και τα διαχωρίζει από τα υπόλοιπα» (Κάλλας, 2000, σελ.258) οι δε πρακτικοί σκοποί αποτελούν τις μετρήσιμες οντότητες που προκύπτουν από τις «πειραματικές συναλλαγές»⁴ των κοινωνικών επιστημών.

Με βάση αυτό το μοντέλο και για τη διασφάλιση της πληρέστερης πληροφόρησης του κοινού και των ειδικών, η ηλεκτρονική αρχειοθέτηση πρέπει να περιλαμβάνει ορισμένες βασικές ενέργειες:

α) περιορισμό του λεξιλογίου σε εννοιολογικά ‘καθαρούς’ όρους, ως προς τους οποίους να συμφωνεί η σχετική επιστημονική πειθαρχία, αλλά και η ευρύτερη επιστημονική κοινότητα

β) εφαρμογή των όρων στα δεδομένα και στα μεταδεδομένα ώστε να αυξάνονται οι πιθανότητες επιτυχίας της αναζήτησης

γ) μία συμφωνημένη δομή περιγραφής των μεταδεδομένων

δ) πολυγλωσσική αντιστοιχία της ορολογίας

ε) αποθήκευση των μεταδεδομένων σε ένα περιβάλλον αποθήκευσης και διαχείρισης

στ) ανάπτυξη εργαλείων πρόσβασης και αναζήτησης.

⁴ Ο όρος αναφέρεται στο: Κάλλας, Γιάννης. *Η Πληροφορική Τεχνολογία στην Κοινωνική Έρευνα*, Αθήνα: ΕΚΚΕ / Νεφέλη, 2000.

4.1 Η βιβλιογραφική αναζήτηση

Η βιβλιογραφική αναζήτηση από την οπτική του κοινωνικού επιστήμονα ως “χρήστη” και όχι του βιβλιοθηκονόμου ως “κατασκευαστή” εντάσσεται στο ζήτημα της ταξινόμησης. Είναι το πέρασμα από τη γενική ταξινόμηση στην ειδική γλώσσα. Η ειδική ή αξιωματική γλώσσα αναφέρεται σε συγκεκριμένο ερευνητικό πεδίο ο δε χρήστης την μαθαίνει μέσα από τη μύηση στο επιστημονικό πεδίο. Οι βιβλιοθήκες χρησιμοποιούν ένα ανοιχτό σύστημα ταξινόμησης που είναι αυτό της θεματολογίας .

Ένα ιδιαίτερα διαδεδομένο σύστημα βιβλιοθηκονομικής ταξιθέτησης, το οποίο εφαρμόζεται στη βιβλιοθήκη του ΕΚΚΕ και σε ένα σημαντικό αριθμό άλλων ελληνικών βιβλιοθηκών είναι το Dewey Decimal Classification . Το σύστημα βοηθά τον χρήστη να αναζητήσει χωροταξικά το θεματικό πεδίο που τον ενδιαφέρει. Με την εδραίωση της χρήσης των Η/Υ το κέντρο βάρους στην αναζήτηση μετατοπίζεται για τον χρήστη στο γνωστικό πεδίο ενδιαφέροντός του.

Ενώ παλαιότερα η ταξιθέτηση αποτελούσε το βασικό στοιχείο αναζήτησης, μετατοπίζεται σήμερα σε συνθετότερους τρόπους αναζήτησης μέσα από συνδυασμούς. Η κυρίως ταξινόμηση βασίζεται στο σύστημα της Library of Congress, που είναι η θεματική ευρετηρίαση (subject headings). Η βιβλιογραφική αναζήτηση στον ελληνικό χώρο είναι ως επί το πλείστον δίγλωσση και, ως εκ τούτου, η συνέργεια κοινωνικών επιστημόνων και βιβλιοθηκονόμων στο ζήτημα της αντιστοιχίας των ελληνικών και των αγγλικών όρων είναι ιδιαίτερα χρήσιμη, δεδομένου ότι η αγγλική γλώσσα επικρατεί στην επιστημονική συναλλαγή. Η αναζήτηση πραγματοποιείται τόσο μέσα από τον τίτλο του εντύπου που αναζητά ο χρήστης όσο και μέσα από το όνομα του συγγραφέα και το θέμα. Η αναζήτηση του θέματος είναι συνθετότερη, γιατί υποστηρίζεται από τις παραπομπές (θεματικές κατηγορίες), που ξετυλίγονται όπως το λεξικό και προσφέρουν ευρύτερο πλέγμα δυνατότητας αναζητήσεων για τον χρήστη. Από το συνολικό μετατοπίζεται η αναζήτηση στο επιμέρους . Σε κάθε περίπτωση ακολουθούνται εσωτερικοί κανόνες που υποβοηθούν την ταξινόμηση σε πολλαπλά πεδία για τη διευκόλυνση του αναγνώστη. Τα δε όρια μεταξύ της βιβλιογραφικής αναζήτησης σε ηλεκτρονικές ή έντυπου υλικού βιβλιοθήκες είναι πλέον δυσδιάκριτα.

4.2 Τα μεταδεδομένα και τα προβλήματα στην ανάπτυξη κοινού κώδικα περιγραφής τους

Τα μεταδεδομένα είναι δεδομένα που παρέχουν την τεκμηρίωση που απαιτείται για να περιγραφούν τα διάφορα σύνολα δεδομένων, που παράγονται μέσα από μία έρευνα. Τα ίδια δεν αποτελούν δεδομένα με τη στενή έννοια του όρου, εφόσον δεν είναι αποτελέσματα καταγραφών ή αναλύσεων. Αναφέρονται στον τρόπο οργάνωσης της καταγραφής και της επεξεργασίας μιας έρευνας και δεν προκύπτουν από την παρατήρηση αλλά από τη διαδικασία της ταξινόμησης. Τα μεταδεδομένα είναι λοιπόν η τεκμηρίωση των δεδομένων μιας έρευνας, που πρέπει να παράγονται από κάθε ερευνητική ομάδα, αν και πολύ συχνά αυτά δεν είναι πλήρη ή λείπουν εντελώς.

Άρα είναι δύσκολο να μιλήσουμε και για την οργάνωσή τους. Πράγματι παρουσιάζεται δυσκολία στην οργάνωση των μεταδεδομένων, διότι στις κοινωνικές επιστήμες δεν υπάρχει ένας σταθερός επιστημονικός κώδικας οργάνωσης, που να είναι γενικά αποδεκτός, όπως πιθανόν γίνεται σε κάποιες φυσικές επιστήμες. Αυτό έχει ως αποτέλεσμα τα δεδομένα να παρουσιάζουν ασυμβατότητες, όπως επίσης και διαφοροποιήσεις από έρευνα σε έρευνα.

Γίνεται επομένως φανερή η ανάγκη δημιουργίας ενός σώματος πρόσθετων πληροφοριών και κανόνων σχεδιασμού, που θα βοηθά τον ερευνητή να ξεπεράσει αυτές τις αδυναμίες. Το σώμα αυτό θα βασίζεται σε ένα μοντέλο μεταδεδομένων, που θα επιτρέπει την περιγραφή όλων των σχημάτων δεδομένων με ενιαίο τρόπο, και το οποίο θα περιέχει όλες τις πρόσθετες πληροφορίες για τον τρόπο που παρήχθησαν τα δεδομένα. Ορίζεται έτσι μία σταθερή δομή για όλες τις πληροφορίες που είναι απαραίτητες για να ορισθεί ένα μοντέλο κοινωνικών δεδομένων. Είναι λοιπόν φανερή η ανάγκη να δημιουργηθεί ένα σύστημα, το οποίο να εξασφαλίζει την πληρότητα και τη συστηματοποίηση των μεταδεδομένων, καθώς επίσης και την αποτελεσματική διαχείρισή τους, αυξάνοντας έτσι την αξία και τη χρησιμότητα αλλά κυρίως τη συμβατότητα των ίδιων των δεδομένων.

Το όραμα μιας παγκόσμιας εννοιολογικής οντολογίας που θα χρησιμοποιείται για τον ορισμό των διαφόρων εννοιών κάποιου λεξικού σε πολλές γλώσσες, έχει την απαρχή του στην πλατωνική φιλοσοφία (Miller & Matthews, 2001). Στην απουσία της παγκόσμιας οντολογίας, κάποια Αρχεία⁵ έχουν δημιουργήσει κάποια εθνικά και, σε μερικές περιπτώσεις, διεθνή μεταδεδομένα, όπως επίσης και κάποια πρότυπα δημιουργίας λεξικών, τα οποία υποστηρίζουν. Για τις κοινωνικές επιστήμες έχει αναπτυχθεί μια διεθνής πρωτοβουλία, με στόχο τη δημιουργία ενός προτύπου μεταδεδομένων των κοινωνικών επιστημών που ονομάζεται Data Documentation Initiative⁶ (Miller & Matthews, 2001). Αυτό το πρότυπο βασίστηκε σε ένα παλαιότερο πρότυπο δημιουργίας μεταδεδομένων, που δεν αφορά αποκλειστικά τις κοινωνικές επιστήμες, και καλύπτει βασικά κριτήρια αναζήτησης και ονομάστηκε Dublin Core από την αντίστοιχη πρωτοβουλία Dublin Core Initiative. Είναι φανερό ότι τα μεταδεδομένα είναι πιο αποτελεσματικά όταν έχουν δημιουργηθεί με συμβατό τρόπο, βασιζόμενα δηλαδή σε κάποιο πρότυπο. Έτσι, τα ψηφιακά δεδομένα μπορούν να καταστούν συμβατά και διαχειρίσιμα από κοινά εργαλεία λογισμικού. Το Dublin Core είναι ένας κανόνας που μπορεί να δημιουργήσει την καρδιά κάποιου παγκόσμιου μοντέλου μεταδεδομένων. Ο ρόλος του είναι να ικανοποιεί ένα πρώτο επίπεδο αρχικών κριτηρίων αναζήτησης, που θα εξειδικευθούν στη συνέχεια από τον κάθε επιστημονικό κλάδο. Η προσέγγιση στο βασικό μοντέλο μεταδεδομένων είναι η απαραίτητη δημιουργία 15 στοιχείων: title, author, subject, description, publisher, other contributor, date, resource type, format resource identifier, source language, relation, coverage, rights management⁷. Τα στοιχεία αυτά μπορεί να είναι επαναλαμβανόμενα καθώς και επεκτάσιμα. Το Dublin Core έχει γίνει διεθνώς αποδεκτό και χρησιμοποιείται στα αρχικά κριτήρια αναζήτησης των βάσεων δεδομένων (Miller & Matthews, 2001).

4.3 Το μοντέλο αναζήτησης στο Internet

Το διαδίκτυο έχει αποδειχθεί ότι είναι σήμερα μια παγκόσμια αστείρευτη πηγή

⁵ Τράπεζες δεδομένων.

⁶ Data Documentation Initiative (DDI) [Πρωτοβουλία Τεκμηρίωσης Δεδομένων, (η μετάφραση είναι των συγγραφέων)]

⁷ Με την παράθεση των στοιχείων στην αγγλική γλώσσα, χωρίς την ελληνική απόδοση, ο αναγνώστης “παίρνει” μια πρώτη γεύση από τις δυσκολίες και προκλήσεις που αντιμετωπίζουν τα έργα πολυγλωσσικής ταξινόμησης. Το συγκεκριμένο πρότυπο τυποποίησης δεν δημιουργήθηκε κατ’ αρχήν για τις κοινωνικές επιστήμες και επομένως, υπάρχει ένας βαθμός ελευθερίας στην προσαρμογή του για τις ανάγκες της τεκμηρίωσης ηλεκτρονικών δεδομένων. Κατά τη διαδικασία της εφαρμογής του προτύπου θα αντιμετωπισθεί η μετάφραση κάθε ενός από τα δεκαπέντε (15) στοιχεία σύμφωνα με το ακριβές εννοιολογικό περιεχόμενο που προσδίδεται στις υπάρχουσες βάσεις δεδομένων.

πληροφοριών. Παρέχει τη δυνατότητα άντλησης πληροφοριών από πολλές μεμονωμένες βάσεις δεδομένων. Έχει παρατηρηθεί ότι, ενώ ο τρόπος αναζήτησης μέσω του διαδικτύου είναι πολύ εύκολος και γρήγορος, τα αποτελέσματα των αναζητήσεων, πολλές φορές, είτε δεν είναι αυτά που θέλουμε, είτε δεν προσεγγίζουν ουσιαστικά το θέμα μας, είτε δεν μπορούν να χρησιμοποιηθούν. Η αναζήτηση σε ετερογενώς καταναμημένες βάσεις δεδομένων είναι σχεδόν αδύνατη. Και αυτό παρατηρείται διότι υπάρχει μια ανομοιογένεια ανάμεσα στα δεδομένα. Για να λυθεί αυτό το πρόβλημα θα έπρεπε: πρώτον, να δημιουργείται πάντα η περιγραφή των δεδομένων, που δεν είναι τίποτα άλλο από τα μεταδεδομένα και, δεύτερον, να υπάρχει ένα ευρέως αποδεκτό λεξικό, το οποίο θα χρησιμοποιείται για την ταξινόμηση των μεταδεδομένων. Με τη χρήση του λεξικού αυτού στην αναζήτηση είναι σίγουρο ότι τα αποτελέσματα θα είναι πολύ πιο ουσιαστικά και ίσως και πιο πολλά. Η λύση αυτή προϋποθέτει: α) ένα μοντέλο για την περιγραφή των μεταδεδομένων, β) ένα λεξικό δεδομένων και γ) την ταξινόμηση των περιγραφών των δεδομένων με βάση το λεξικό. Όταν η παραπάνω λύση δόθηκε στη δεκαετία του 1970, πολλά Αρχεία που θέλησαν να την υιοθετήσουν δημιούργησαν τα δικά τους μοντέλα μεταδεδομένων και τα δικά τους λεξικά, χωρίς φυσικά να ακολουθήσουν μια καθολική παγκόσμια λύση στο πρόβλημα.

4.4 Η σχέση τεκμηρίωσης & αναζήτησης

Όπως αναφέρθηκε ήδη, το διαδίκτυο έχει αποδειχθεί ότι είναι μια παγκόσμια, αστείρευτη πηγή αναζήτησης πληροφοριών. Για να γίνουν αναζητήσιμα τα μεταδεδομένα, θα πρέπει να ακολουθηθούν κάποιοι κανόνες όσον αφορά τη μορφή παρουσίασής τους στο διαδίκτυο. Οι κανόνες τίθενται από το Resource Description Framework (RDF)⁸. Σύμφωνα με αυτόν τον κανονισμό τα μεταδεδομένα δομούνται και γίνονται διαχειρίσιμα ανάλογα με το πρότυπο διαχείρισης των συλλογών δεδομένων.

Για να εγκολπωθεί το διαδίκτυο τις διαφορετικές τυπικές κοινότητες περιγραφής πηγών όπως: μουσεία, βιβλιοθήκες, ερευνητικά κέντρα, πανεπιστήμια κ.ά. αλλά και άτυπες, που διαθέτουν συλλογές, επιβάλλεται να ακολουθηθούν κάποιοι κανόνες συμβατότητας ανάμεσα στα μεταδεδομένα. Η συμβατότητα θα πρέπει να διέπει: α) τη σημασιολογία, δηλαδή την έννοια των στοιχείων, που προϋποθέτει την ύπαρξη ενός κοινού κώδικα, ενός κοινού τύπου επικοινωνίας ανάμεσα σε εκείνον που τον “κατασκευάζει” και σε εκείνον που τον αναζητά, β) τη δόμηση, δηλαδή τα μεταδεδομένα θα πρέπει να διαβάζονται από τους ανθρώπους, αλλά και από τα μηχανήματα και γ) τη σύνταξη, δηλαδή θα πρέπει να γίνει το πάντρεμα μεταξύ των εννοιών και της δόμησης. Το τελευταίο επιτυγχάνεται με τις γλώσσες προγραμματισμού HTML, SGML και XML. Οι γλώσσες αυτές είναι αποδεκτές και προτείνονται από το World Wide Web Consortium (W3C)⁹.

Για την πρόσβαση στα μεταδεδομένα έχουν αναπτυχθεί διάφορα εργαλεία. Ένα από αυτά, το οποίο χρησιμοποιείται για την πρόσβαση μεταδεδομένων σε Αρχεία Κοινωνικών Δεδομένων μέσω του διαδικτύου είναι το NESSTAR. Μέσω του

⁸ Resource Description Framework (RDF) [Πλαίσιο Περιγραφής Πόρων, (η μετάφραση είναι των συγγραφέων)]

⁹ World Wide Web Consortium (W3C) [Κοινοπραξία του Ευρύτερου Παγκόσμιου Ιστού, (η μετάφραση είναι των συγγραφέων)]

NESSTAR επιτυγχάνεται πρόσβαση στα Αρχεία των Βρετανών (UK-DA), των Δανών (DDA), και των Νορβηγών (NSD). Τα παραπάνω Αρχεία χρησιμοποιούν το DDI ως πρότυπο για την οργάνωση των μεταδεδομένων τους και την XML για την εμφάνιση των μεταδεδομένων τους. Είναι ένα εργαλείο που διατίθεται μόνο στην αγγλική γλώσσα και συνδέει τα Αρχεία Κοινωνικών Δεδομένων των Βρετανών, των Δανών και των Νορβηγών.

4.5 Πολυγλωσσική αναζήτηση

Η πολυγλωσσική αναζήτηση εμπεριέχει το ζήτημα της ασυμβατότητας μεταξύ των γλωσσών, γεγονός που συνιστά πρόβλημα τεκμηρίωσης μάλλον παρά αναζήτησης. Το ζήτημα των γλωσσικών φραγμών αφορά αφ' ενός τα προβλήματα αναζήτησης και τον περιορισμό του αριθμού αλλά και του είδους των τελικών χρηστών των βάσεων κοινωνικών δεδομένων, αφ' ετέρου την υπάρχουσα ποικιλία και, εν τέλει, την ανακολουθία στη χρήση λειτουργικών όρων, που περιγράφουν έννοιες μέσα στον ίδιο επιστημονικό κλάδο. Από μια ευρύτερη οπτική αφορά επιπλέον και τα όρια στην επικοινωνία ανάμεσα σε διαφορετικές επιστημονικές πειθαρχίες (Επιτροπή Gulbenkian, 1998).

Οι προσπάθειες να αμβλυνθούν οι γλωσσικοί περιορισμοί ενεργοποιούν και αναπτύσσουν παράλληλα μια σειρά από δράσεις, οι οποίες αφορούν: α) τη δημιουργία μηχανισμών τυποποίησης της τεκμηρίωσης – με τον τρόπο που ήδη περιγράψαμε στο κεφάλαιο περί μεταδεδομένων· β) τον σχεδιασμό διαμεσολαβητικού περιβάλλοντος (interfaces) πολυγλωσσικής αντιστοιχίας όρων και γ) ελεγχόμενους πολυγλωσσικούς θησαυρούς κοινωνικών όρων. Πρόκειται για φιλόδοξες προσπάθειες, με έμφαση στην ανάπτυξη τεχνικών μέσων υποστήριξης του περιεχομένου και η πρωτοτυπία τους έγκειται στο ότι το περιεχόμενο αποτελεί *διαχειρίσιμη οντότητα* από όποια οπτική και αν το προσεγγίσει κανείς. Υπ' αυτήν την έννοια ο διαμεσολαβητής θα πρέπει να είναι επιστήμονας-μεταφραστής και όχι απλά μεταφραστής δεδομένου ότι ο πολυγλωσσικός μηχανισμός κτίζεται πάνω στην ασύγχρονη επικοινωνία. Οι πληροφορίες μέσω του H/Y ψηφιοποιούνται για να γίνουν διαχειρίσιμες χωρίς τη φυσική παρουσία του παραγωγού και να μετασηματισθεί η επικοινωνία από «σύγχρονη» σε «ασύγχρονη» (Κάλλας, 2000).

4.5.1 Το πρόγραμμα Limber

Ένα πρόγραμμα για την άμβλυνση των γλωσσικών περιορισμών κατά την αναζήτηση δεδομένων από το δίκτυο των Αρχείων των ευρωπαϊκών χωρών – το Language Independent Metadata Browsing of European Resources (LIMBER)¹⁰ υλοποιείται από κοινοπραξία ερευνητικών και πανεπιστημιακών ιδρυμάτων¹¹. Στόχος είναι η δημιουργία ενός πολυγλωσσικού θησαυρού, του European Language Social Science Thesaurus (ELSST)¹² για τις περισσότερες ευρωπαϊκές γλώσσες. Ο θησαυρός αυτός βασίζεται στον θησαυρό Humanities and Social Science Electronic Thesaurus

¹⁰ Language Independent Metadata Browsing of European Resources (LIMBER) [Γλωσσικά Ανεξάρτητη Αναζήτηση Ευρωπαϊκών Πηγών Δεδομένων, (η μετάφραση είναι των συγγραφέων)]

¹¹ Το έργο χρηματοδοτείται από την ΕΕ στο πλαίσιο του προγράμματος IST (Information Societies Technology). Ανάδοχος φορέας είναι το Πανεπιστήμιο του Essex και συμμετέχουν φορείς από πολλές ευρωπαϊκές χώρες.

¹² European Language Social Science Thesaurus (ELSST) [Θησαυρός Κοινωνικών Επιστημών της Ευρωπαϊκής Γλώσσας, (η μετάφραση είναι των συγγραφέων)]. Η ονομασία αναφέρεται στην *Ευρωπαϊκή Γλώσσα* με την έννοια της ενιαίας κατανόησης όρων και της επικοινωνίας χωρίς γλωσσικούς περιορισμούς που αποτελεί στόχο του συγκεκριμένου προγράμματος

(HASSET)¹³, ο οποίος προήλθε από τον θησαυρό της UNESCO. Σε πρώτο στάδιο ο πολυγλωσσικός θησαυρός ELSST θα περιλαμβάνει όρους στα αγγλικά, γερμανικά, ισπανικά και γαλλικά. Ο μελλοντικός προγραμματισμός προβλέπει και την ενσωμάτωση θησαυρών άλλων ευρωπαϊκών γλωσσών – περιλαμβανομένης και της ελληνικής.

Οι ανάγκες που οδήγησαν στη δημιουργία αυτού του προγράμματος απορρέουν από την ύπαρξη καταρχήν γλωσσικών ορίων στην ενοποιούμενη Ευρώπη, που δυσχεραίνουν το άνοιγμα συγκριτικών ερευνών στον ευρωπαϊκό, πλέον, κοινωνικό χώρο, καθώς και τη δυνατότητα της συνδρομής των ευρημάτων άλλων επιστημονικών κλάδων (γεωγραφία, περιβάλλον, επιστήμες της υγείας) για τη διαμόρφωση των συμπερασμάτων της κοινωνικής έρευνας (Miller & Matthews, 2001). Η μέθοδος με την οποία προσεγγίζεται αυτή η προσπάθεια ακολουθεί τις αρχές του “νέου μοντέλου αναζήτησης” και το επιθυμητό αποτέλεσμα-στόχος είναι η διάθεση των δεδομένων, διεθνώς, σε μεγάλο αριθμό χρηστών, στη φυσική τους γλώσσα (Limber, Δεκ. 21, 2000). Αναλυτικά, το πρόγραμμα στοχεύει στα εξής:

α) Τον σχεδιασμό και την ανάπτυξη ενός κοινού μοντέλου αναπαράστασης μεταδεδομένων, που θα επιτρέψει την ενσωμάτωσή τους στα Αρχεία δεδομένων όλων των χωρών.

β) Την ανάπτυξη ενός πολύγλωσσου θησαυρού κοινωνικών όρων, ο οποίος θα περιέχεται στα σύνολα δεδομένων των κοινωνικών επιστημών και θα χρησιμοποιείται κατά την τεκμηρίωση και την αναζήτησή τους.

γ) Την ανάπτυξη ενός εργαλείου αναζήτησης και ανεύρεσης δεδομένων, που θα λειτουργεί σε πολλές γλώσσες, θα χρησιμοποιεί τον πολύγλωσσο θησαυρό με αυτόματη μετάφραση φράσεων και λέξεων-κλειδιά, που περιέχονται στα μεταδεδομένα.

δ) Την ανάπτυξη εργαλείων υποστήριξης για την κατασκευή και συντήρηση βάσεων δεδομένων με χρήση αυτόματου ευρετηρίου (indexing).

ε) Την παρουσίαση ενός προτύπου XML διακομιστή (server) μεταδεδομένων, με εφαρμογή του μοντέλου αναπαράστασης μεταδεδομένων.

(UK Data Archive/ Limber, 2000)

4.5.2 Η λειτουργία του συστήματος μετάφρασης στο επίπεδο του τελικού χρήστη

Στην τελική του μορφή το σύστημα θα αξιοποιεί το διαδίκτυο για την επικοινωνία μέσω του δικτύου CESSDA¹⁴ από οποιοδήποτε σημείο εργασίας. Με τη βοήθεια των πολυγλωσσικών διαμεσολαβητικών μέσων ο χρήστης διαμορφώνει την επιφάνεια εργασίας ανάλογα με τη φυσική του γλώσσα, και δημιουργεί ερωτήματα με την επιλογή όρων από το περιορισμένο λεξιλόγιο του θησαυρού στη γλώσσα που επιθυμεί. Κατόπιν, το σύστημα μεταφράζει αυτόματα τα ερωτήματα, μέσω του πολύγλωσσου θησαυρού, σε όλες τις διαθέσιμες γλώσσες. Οι περιγραφές των μεταδεδομένων που περιέχουν τους όρους επιστρέφουν στον χρήστη στη γλώσσα που έχει επιλέξει για να κάνει την αναζήτηση των δεδομένων.

¹³ Humanities and Social Science Electronic Thesaurus (HASSET) [Ηλεκτρονικός Θησαυρός Ανθρωπιστικών και Κοινωνικών Επιστημών (η μετάφραση είναι των συγγραφέων)]

¹⁴ Council of European Social Science Data Archive (CESSDA) [Συμβούλιο Ευρωπαϊκών Αρχείων Δεδομένων των Κοινωνικών Επιστημών]

Για να λειτουργεί το σύστημα είναι απαραίτητο τα κείμενα των μεταδεδομένων να περιλαμβάνουν τις σχετικές φράσεις ή λέξεις-κλειδιά που βρίσκονται στον πολύγλωσσο θησαυρό. Αυτό δεν σημαίνει ότι θα αποδίδεται πλήρως το νόημα που εμπεριέχεται στα κείμενα και ότι δεν θα υπάρχουν δυσνόητα σημεία και αποκλίσεις από την αρχική εννοιολογική αναζήτηση. Η χρήση όμως του περιορισμένου θησαυρού όρων θα ανοίξει τη δυνατότητα πρόσβασης σε βάσεις δεδομένων που, λόγω του γλωσσικού περιορισμού, έχουν μόνο τοπική χρήση. Επίσης, σημαντική είναι η διάσταση της ιεραρχικής ταξινόμησης που θα περιέχεται στον πολύγλωσσο θησαυρό όρων, και ενισχύει την αποτελεσματικότητα της αναζήτησης, όταν διαμεσολαβείται από τη μετάφραση. Δηλαδή, οι όροι αναζήτησης θα περιέχουν έναν αριθμό από σχετικούς, παράλληλους και ειδικούς όρους, καθώς και συνώνυμα, που θα μπορούν να εμπλουτίσουν το ερώτημα της αναζήτησης και να αυξήσουν την επιτυχία.

4.6 Ηλεκτρονικοί θησαυροί κοινωνικών όρων στην ελληνική γλώσσα

Όπως προκύπτει από τη διεθνή εμπειρία στις βάσεις δεδομένων, τα ταξινομικά συστήματα τεκμηρίωσης και αναζήτησης βελτιστοποιούνται με τη συνεχή εφαρμογή και την ανανέωσή τους σύμφωνα με τις ανάγκες του περιβάλλοντος διαχείρισης για το οποίο αναπτύχθηκαν, και του περιεχομένου του. Στην Ελλάδα δεν υπήρξαν οι συγκυρίες εκείνες που θα επέτρεπαν την ανάπτυξη μιας ελληνικής ορολογίας των κοινωνικών επιστημών, αφ'ενός λόγω της *γλωσσικής ανακολουθίας*, που είναι αποτέλεσμα της άνισης ανάπτυξης της πρωτογενούς επιστημονικής παραγωγής σε σχέση με άλλες δυτικές χώρες αφ'ετέρου λόγω έλλειψης τεκμηρίωσης της ερευνητικής διαδικασίας. Είναι δε γεγονός ότι μέχρι πρότινος δεν συνέτρεχε ίσως, λόγος για ένα τέτοιο έργο, εφ'όσον ο πλούτος της ελληνικής γλώσσας κάλυπτε τις ανάγκες επικοινωνίας σε εθνικό επίπεδο. Στην ελληνική γλώσσα διατίθενται λεξικά σε έντυπη μορφή [βλ. για παράδειγμα το *Ελληνο-Αγγλικό Λεξικό Κοινωνικών Όρων* Π. Σταθόπουλου (1989), ή το *Χρηστικό Λεξικό Κοινωνιολογίας*, Δ.Τσαούση (1984)]. Παράλληλα, εξ' όσων γνωρίζουμε διατίθενται ευρετήρια όρων σε επιμέρους τομείς των Κοινωνικών Επιστημών όπως η έκδοση *European Employment & Industrial Relations Glossary: Greece*, Yota Kravaritou et al. (1994). Εξάλλου, η πλοήγηση στην ιστοσελίδα του Εθνικού Κέντρου Τεκμηρίωσης (ΕΚΤ) σε ό,τι αφορά τα ελληνικά λεξικά μας οδηγεί σε λεξικά γενικής χρήσης. Ως εξειδικευμένα λεξικά αναφέρονται λεξικά και ευρετήρια τεχνικών όρων, πληροφορικής και διαδικτύου (internet).

Οι νέες μορφές επικοινωνίας, η αλλαγή των ρυθμών της, η αναπόφευκτη σύνδεση των κοινωνικών επιστημών με την επιστήμη της πληροφορικής, η αξιόλογη διατύπωση επιστημονικού λόγου των σύγχρονων ελλήνων επιστημόνων πέρα από τα εθνικά μας όρια, είναι μερικοί από τους λόγους για τους οποίους είναι απαραίτητη η συσσώρευση και η διαχείριση του επιστημονικού έργου στον τόπο μας. Απαραίτητη προϋπόθεση για αυτόν τον σκοπό είναι η καταρχήν συμφωνία των επιστημόνων και φορέων που υπηρετούν την κοινωνική έρευνα στην Ελλάδα στη διατύπωση αποδεκτού/ αποδεκτών ειδικών κωδικών κοινωνικών όρων. Σε ένα δεύτερο επίπεδο αυτός ο κώδικας πρέπει να “συνομιλεί” με τους αντίστοιχους κώδικες της ίδιας επιστήμης σε άλλες γλώσσες και σε ένα τρίτο επίπεδο πρέπει η εφαρμογή του στη γλώσσα μας να επιτρέπει την αυτόματη προσαρμογή του σε οποιαδήποτε άλλη γλώσσα, τόσο από εννοιολογική όσο και από τεχνική άποψη.

5 ΑΝΑΠΤΥΞΗ ΥΠΟΔΟΜΩΝ ΓΙΑ ΤΗΝ ΚΟΙΝΩΝΙΚΗ ΕΡΕΥΝΑ: Η ΤΡΑΠΕΖΑ ΚΟΙΝΩΝΙΚΩΝ ΔΕΔΟΜΕΝΩΝ ΤΟΥ ΕΚΚΕ

Το Εθνικό Κέντρο Κοινωνικών Ερευνών ξεκίνησε από το 1998 τη δημιουργία της Τράπεζας Κοινωνικών Δεδομένων (ΤΚΔ), με σκοπό να λειτουργήσει ως το πλέον αξιόπιστο Αρχείο Κοινωνικών Δεδομένων για τη χώρα. Η συμμετοχή της ΤΚΔ. στο δίκτυο CESSDA σηματοδοτεί τον επίσημο ρόλο της Τράπεζας στην Ελλάδα και προωθεί τη συσσώρευση μεγάλου όγκου δεδομένων, ενώ εγγυάται την απαιτούμενη αξιοπιστία του εγχειρήματος. Η εξέλιξη και ολοκλήρωση του έργου της ΤΚΔ. θα καλύψει πολλά κενά στην τεκμηρίωση της εμπειρικής κοινωνικής έρευνας στη χώρα, καθώς υπάρχουν βέβαια και άλλες αξιόλογες προσπάθειες, είναι ωστόσο αποσπασματικές και μεσοπρόθεσμες.

Η ΤΚΔ. του ΕΚΚΕ, πέρα από την κάλυψη της -τόσο επιτακτικής στη χώρα μας - ανάγκης συσσώρευσης των κοινωνικών δεδομένων φιλοδοξεί να αποτελέσει ένα ολοκληρωμένο πλαίσιο διαχείρισης του συνόλου της πληροφορίας που σχετίζεται με τις κοινωνικές επιστήμες. Έτσι, είναι σχεδιασμένη σε ένα Σχεσιακό Μοντέλο Διαχείρισης Βάσεων Δεδομένων, που αποτελεί μια πιο αποτελεσματική φιλοσοφία σχεδιασμού των βάσεων (για τις λεπτομέρειες της δομής και λειτουργίας της ΤΚΔ., βλ.: Κάλλας, 2001). Το σχεσιακό μοντέλο διαχείρισης συμβάλλει στην ταχύτερη και αποτελεσματικότερη αναζήτηση των δεδομένων και, ως εκ τούτου, στο κτίσιμο της επικοινωνίας μεταξύ της μελών και φορέων της επιστημονικής κοινότητας.

5.1 Η ανάπτυξη Θησαυρού Κοινωνικών Όρων στην Ελληνική

Ένας από τους επιμέρους στόχους της ΤΚΔ. είναι και η συμβολή της στην ανάπτυξη μεθόδων και εργαλείων για την τυποποίηση των δεδομένων και μεταδεδομένων της ελληνικής εμπειρικής κοινωνικής έρευνας. Στο πλαίσιο αυτό έχει προγραμματιστεί και η δημιουργία *Θησαυρού Κοινωνικών Όρων στην Ελληνική*, ο οποίος θα εξυπηρετεί τις ανάγκες τεκμηρίωσης και αναζήτησης των έργων και δεδομένων της ΤΚΔ. Για τον σκοπό αυτό έχει σχηματισθεί μια διεπιστημονική ομάδα εργασίας, με τη συμμετοχή μελών των Ινστιτούτων και της Διεύθυνσης Υποστήριξης Ερευνών του ΕΚΚΕ¹⁵. Το έργο της δημιουργίας ενός ελληνικού ταξινομικού συστήματος όρων των κοινωνικών επιστημών θα είναι βασισμένο σε διεθνή ηλεκτρονικά ταξινομικά συστήματα για λόγους συγκρισιμότητας και επιστημονικής πειθαρχίας. Στόχος της ομάδας είναι να παραδώσει ένα ολοκληρωμένο ταξινομικό σύστημα όρων, συμβατό αφ'ενός με την ελληνική ορολογία των κοινωνικών επιστημών αφ'ετέρου με τα αντίστοιχα ταξινομικά συστήματα που υπάρχουν στις κυριώτερες ευρωπαϊκές γλώσσες.

Στο εννοιολογικό επίπεδο θα επιχειρηθεί η προσαρμογή στην ελληνική γλώσσα των αγγλόφωνων όρων των κοινωνικών επιστημών, που προέρχονται από τον θησαυρό HASSET, ο οποίος έχει αναπτυχθεί με βάση τον θησαυρό της UNESCO. Οι όροι και οι ταξινομήσεις του HASSET αποτελούν την κυριώτερη πηγή τεκμηρίωσης και αναζήτησης για τις βάσεις δεδομένων του πανεπιστημίου του Essex στη Μεγάλη Βρετανία, που συγκεντρώνει τον μεγαλύτερο όγκο κοινωνικών δεδομένων στην Ευρώπη. Η αντιστοίχιση του ελληνικού θησαυρού με τον αγγλόφωνο θεωρείται προτιμότερη για τρεις κυρίως λόγους:

¹⁵ Στην ομάδα εργασίας συμμετέχουν οι : Ε. Αλλισον (ΙΑΑΚ) , Α. Δανέλη (ΔΥΕ), Χ. Καππή (ΙΝΚΠΟ), Δ. Κονδύλη (Γ.Δ.), Γ. Κοτταρίδη (ΙΝΚΠΟ), Θ. Σταθοπούλου (ΠΠΟΚΟΙΝΩ), Μ. Τοπάλη (Γ.Δ.), Χ. Φρέντζου (ΔΥΕ), Α. Χατζηγιάννη (ΙΑΑΚ).

α) ο αγγλόφωνος θησαυρός καλύπτει τις ανάγκες μεγάλου όγκου κοινωνικών δεδομένων και ως εκ τούτου αποτελεί μια μεγάλη και διαθέσιμη πηγή κοινωνικών όρων,

β) η όσο το δυνατόν καλύτερη αντιστοιχία των ελληνικών όρων με τους όρους ενός διεθνώς αναγνωρισμένου θησαυρού θα ενισχύσει το κριτήριο της συγκρισιμότητας των βάσεων που θα σχεδιασθούν και,

γ) δεδομένου ότι η ελληνική επιστημονική κοινότητα αντιμετωπίζει δυσκολίες στην ανεύρεση και εδραίωση χρήσης κοινά αποδεκτών όρων, το έργο δημιουργίας ενός τέτοιου θησαυρού θα ενεργοποιήσει τον προβληματισμό για τις εννοιολογικές ιδιαιτερότητες των κοινωνικών όρων που προκύπτουν από τα ιδιαίτερα ελληνικά πολιτισμικά στοιχεία.

5.2 Δημιουργία ενός ελληνικού ταξινομικού συστήματος όρων των κοινωνικών επιστημών

Το έργο της δημιουργίας ενός ελληνικού ταξινομικού συστήματος όρων των κοινωνικών επιστημών στο πλαίσιο του ΕΚΚΕ αποτελεί την απαρχή μιας μακροπρόθεσμης προσπάθειας στον ευρύτερο ορίζοντα της τεκμηρίωσης της ελληνικής κοινωνικής έρευνας. Η πραγμάτωση του έργου θα είναι ανέφικτη χωρίς τη συνεργασία και επικοινωνία με την ευρύτερη ελληνική επιστημονική κοινότητα. Το έργο εντάσσεται στον πάγιο ερευνητικό και αναπτυξιακό άξονα του ΕΚΚΕ με τίτλο «Ανάπτυξη Υποδομών για την Κοινωνική Έρευνα» (Επιχειρησιακό Σχέδιο Δράσης ΕΚΚΕ 2001-2006) και αποτελείται από ενδιάμεσα στάδια εργασίας για την εξασφάλιση της σταδιακής βελτίωσης και ανατροφοδότησής του από την εμπειρία που θα συσσωρεύεται σε κάθε στάδιο. Συγκεκριμένα προβλέπονται τα εξής:

1 Μετάφραση στην ελληνική γλώσσα όρων των κοινωνικών επιστημών από θησαυρούς κοινωνικών όρων που χρησιμοποιούνται από τα Αρχεία άλλων χωρών για:

- την εννοιολογική αποσαφήνιση των όρων
- την απόδοσή τους στην ελληνική γλώσσα

2. Αντιστοίχιση των παραχθέντων ελληνικών όρων με το αγγλόφωνο ταξινομικό σύστημα ELSST του προγράμματος LIMBER για:

- εννοιολογικές συγκρίσεις
- αντίστοιχη σχεσιακή δόμηση των όρων

3. Εμπλουτισμός του ελληνικού ταξινομικού συστήματος με όρους που δεν περιέχονται στο ταξινομικό σύστημα του προγράμματος LIMBER αλλά χρησιμοποιούνται στην ελληνική γλώσσα από τις κοινωνικές επιστήμες για:

- την ολοκλήρωση του *ειδικού κώδικα* για την ελληνική γλώσσα
- τη διαμόρφωση σχεσιακής δόμησης, η οποία θα ανταποκρίνεται στις εννοιολογικές κατηγοριοποιήσεις της ελληνικής γλωσσικής και πολιτισμικής πραγματικότητας.

4. Πιλοτική εφαρμογή και αξιολόγηση του ελληνικού ταξινομικού συστήματος σε επιλεγμένα σύνολα δεδομένων, τα οποία αφορούν τομείς αιχμής της ερευνητικής παραγωγής του ΕΚΚΕ για:

- την αξιολόγηση του *θησαυρού* ως εργαλείου τεκμηρίωσης και αναζήτησης ελληνικών δεδομένων και μεταδεδομένων της κοινωνικής έρευνας στην

Ελλάδα

5. Εφαρμογή του ελληνικού ταξινομικού συστήματος στα έργα, τα δεδομένα και τα μεταδεδομένα της Τράπεζας Κοινωνικών Δεδομένων του ΕΚΚΕ για:

- την ανάπτυξη εργαλείου τεκμηρίωσης και διαχείρισης της ΤΚΔ.

6. Διάχυση του παραχθέντος ταξινομικού συστήματος κοινωνικών όρων στην ευρύτερη επιστημονική κοινότητα για:

- την προώθηση διαλόγου που θα επιτρέψει τη βελτίωση και περαιτέρω ανάπτυξη ενός *Θησαυρού Κοινωνικών Όρων στην Ελληνική* με την ευρύτερη δυνατή αποδοχή.

6 ΠΑΡΑΤΗΡΗΣΕΙΣ

Τα Αρχεία λειτουργούν ως διεθνή ερευνητικά εργαστήρια αφ' ενός για τη συλλογή και καταγραφή της πληροφορίας, αφ' ετέρου για την ανάπτυξη νέων μεθόδων επιστημονικής συναλλαγής. Στο σημερινό στάδιο της ανάπτυξης των Αρχείων και των εργαλείων τεκμηρίωσης και διαχείρισής του περιεχομένου τους υπάρχουν αρκετές προσεγγίσεις για την αναπαράσταση των τεκμηρίων και τη διατύπωση των ερωτημάτων. Κάθε μια από αυτές πρέπει να δοκιμαστεί και να αξιολογηθεί ως προς τη δυνατότητα επιτυχίας των σκοπών της εφαρμογής για την οποία δημιουργήθηκε. Οι εφαρμογές που προέκυψαν εισήγαγαν ένα νέο είδος μεθόδων και παραλλαγών στις παραδοσιακές ταξινομικές προσεγγίσεις. Σε μια μελλοντική κατάσταση θα είναι διαθέσιμος ένας τεράστιος όγκος δεδομένων χωρίς εννοιολογικούς φραγμούς, και θα αρκεί το ενδιαφέρον και η δημιουργικότητα του ερευνητή, από οποιοδήποτε σημείο του πλανήτη, για να σχηματοποιήσει ένα νέο ερώτημα, να ανακαλύψει νέες σχέσεις.

Οι προοπτικές και οι συνέπειες της διάθεσης των δεδομένων μέσω της νέας τεχνολογίας είναι δύσκολο, ακόμη, να εκτιμηθούν σε όλη τους την έκταση. Η έντονη προσπάθεια ανάπτυξης νέων μεθόδων και εργαλείων επικοινωνίας αναδεικνύει προβλήματα με διαστάσεις τεχνικές, εννοιολογικές, πρακτικές. Ωστόσο, είναι βέβαιο ότι η σύγκλιση προς έναν συμβατό και κοινώς αποδεκτό τρόπο διάθεσης της επιστημονικής παραγωγής δεν ακυρώνει τους παραδοσιακούς τρόπους επικοινωνίας. Η συνύπαρξη του εγγράφου με το ψηφιοποιημένο τεκμήριο είναι εφικτή. Το ψηφιοποιημένο τεκμήριο είναι αποτελεσματικό μόνον ως το αποτέλεσμα της δημιουργικής παραγωγής της ανθρώπινης νόησης, όπως συμβαίνει με οποιαδήποτε 'καταγραφή' της ανθρώπινης σκέψης.

ΠΗΓΕΣ

- 1 Εθνικό Κέντρο Τεκμηρίωσης,
<http://www.yourdictionary.com/languages/hellenic.html>
- 2 Έκθεση της Επιτροπής Gulbenkian για την Αναδιάρθρωση των Κοινωνικών Επιστημών, *Για το Άνοιγμα των Κοινωνικών Επιστημών*. ΕΚΚΕ. Κείμενα Εργασίας, τ.3 (1998)
- 3 Κάλλας, Γιάννης. *Η Πληροφορική Τεχνολογία στην Κοινωνική Έρευνα*, Αθήνα: ΕΚΚΕ / Νεφέλη (2000)
- 4 Κάλλας, Γιάννης. Σχεδιασμός ενός Περιβάλλοντος Διαχείρισης Κοινωνικών Δεδομένων. ΕΚΚΕ. Κείμενα Εργασίας, τ.15 (2001)
- 5 Λέκκας, Παντελής. *Σύγχρονα Ρεύματα Κοινωνικής Θεωρίας*, Αθήνα: Πάντειο Πανεπιστήμιο (1995)
- 6 Νικολαΐδης, Δημήτριος. *Υπολογιστική Γλωσσολογία: Όταν οι Υπολογιστές Καταλαβαίνουν τη Δική μας Γλώσσα*, Περισκόπιο της Επιστήμης, τ.246, (Ιανουάριος 2001), σελ.78-83.
- 7 Σταθόπουλος, Π. Α. *Αγγλοελληνικό & Ελληνοαγγλικό Λεξικό Όρων Κοινωνικών Επιστημών*, Αθήνα: Καρδαμίτσας (1989)
- 8 Τράπεζα Κοινωνικών Δεδομένων, <http://www.ekke.gr/databank/index.html>
- 9 Τσαούσης, Δ. *Χρηστικό Λεξικό Κοινωνιολογίας*, Αθήνα: Gutenberg (1984)
- 10 American Society of Indexers, *Thesaurus Construction*, University of Western Ontario: (January 2001)
<http://instruct.uwo.ca/gplis/677/thesaur/main00.htm>
- 11 Ayer, Alfred. *The Problem of Knowledge*, Middlesex: Penguin (1956)
- 12 DDC20 Dewey Decimal Classification, edition 20, New York: Forest Press (1989)
- 13 HASSET. Humanities And Social Sciences Electronic Thesaurus at University of Essex . Based upon *Unesco Thesaurus*, comp. J. Aitchison, Unesco, Paris (1977) <http://biron.essex.ac.uk/searching/zhasset.html>
- 14 Kravaritou Y. *European Employment & Industrial Relations Glossary: Greece* – European Foundation for the Improvement of Living and Working Conditions. Office for Official Publications of the European Communities: Sweet and Maxwell (1994)
- 15 Liddy, Elizabeth D. (Dr.) **[Whither Come the Words?](#)** Centre for Natural Language Processing School of Information Studies Syracuse University. Presentation at the Subject Analysis and Retrieval Working Group

Conference *Controlled Vocabulary and the Internet* (September 29, 1999)

- 16 Limber Annual Report 2000.
http://venus.cis.rl.ac.uk/limber/External/report_2000.com
- 17 Limber: Language Independent Metadata Browsing of European Resources,
<http://www.data-archive.ac.uk/home/limber.asp>
- 18 Lyotard, Jean-Francois (μτφ. Κ. Παπαγιώργης). *Η Μεταμοντέρνα Κατάσταση*, Αθήνα: Γνώση (1988).
- 19 Matthews, B.M., K. Miller, A. Ramfos, J. Ryssevic, M.D. Wilson, “Internationalising data access through LIMBER”, in D.L. Day & I.M. Dunckley (eds) *Designing for Global Markets 3: Proceedings of IWIPS 2001*, p.p. 129-142, Open University: Milton Keynes.
- 20 Miller, Ken, B. Matthews. “Having the Right Connections: the LIMBER Project”, *Journal of Digital Information*, Vol. 1 issue 8 Themes: Information Discovery (2001-02-05)
- 21 Miller, Patricia H. *Theories of Developmental Psychology*, New York: W.H. Freeman & Co (1992)
- 22 Queensland University of Technology. *Controlled Vocabularies*, Controlled Vocabularies Resource Guide (September 2001):
http://www2.fit.qut.edu.au/InfoSys/middle/cont_voc.html
- 23 University of Essex, U.K., <http://www.data-archive.ac.uk/>
- 24 Usher, Robin & Richard Edwards. *Post-modernism and Education*, London: Routledge (1994)

ΕΠΕΞΗΓΗΣΗ ΑΚΡΩΝΥΜΙΩΝ

ΓΔ -Γραφείο Διαμεσολάβησης.....	16
ΔΥΕ -Διεύθυνση Υποστήριξης Ερευνών.....	16
ΕΕ -Ευρωπαϊκή Ένωση.....	1, 12
ΕΚΤ -Εθνικό Κέντρο Τεκμηρίωσης.....	15
ΙΑΑΚ -Ινστιτούτο Αστικής και Αγροτικής Κοινωνιολογίας.....	16
ΙΝΚΠΟ - Ινστιτούτο Κοινωνικής Πολιτικής.....	16
ΙΠΟΚΟΙΝΩ - Ινστιτούτο Πολιτικής Κοινωνιολογίας.....	16
ΤΚΔ - Τράπεζα Κοινωνικών Δεδομένων.....	15, 16, 17
ASI -American Society of Indexers.....	7, 12
CESSDA -Council of European Social Science Data Archives.....	14, 15
DDA -Danish Data Archive.....	12
DDI -Data Documentation Initiative.....	9, 11, 12
ELSST - European Language Social Science Thesaurus.....	13, 17
HASSET -Humanities And Social Science Electronic Thesaurus.....	13, 16, 19
HTML -Hypertext Markup Language.....	12
LIMBER -Language Independent Metadata Browsing of European Resources.....	13, 17, 19
NESSTAR -Network European Social Science Tools and Resources.....	12
NSD -Norwegian Social Science Data Services.....	12
RDF -Resource Description Framework.....	12
SGML -Standard Generalised Markup Language.....	12
UK-DA -United Kingdom Data Archives.....	12
UNESCO -United Nations Educational Scientific and Cultural Organization.....	13, 16
W3C - World Wide Web Consortium.....	12
XML - eXtensible Markup Language.....	12, 14