

Σχεδιασμός ενός Περιβάλλοντος Διαχείρισης Κοινωνικών Δεδομένων

Γ. Κάλλας Α. Λιναρδής

Αθήνα, 2004

Περιεχόμενα

1.	ΕΙΣΑΓΩΓΗ.....	3
2.	ΤΟ ΔΙΦΑΣΙΚΟ ΜΕΘΟΔΟΛΟΓΙΚΟ ΥΠΟΔΕΙΓΜΑ	4
2.1.	Η παραγωγικές διαδικασίες.....	5
2.2.	Η φάση της καταγραφής	7
2.3.	Η φάση της ανάλυσης	7
2.4.	Η λειτουργία των υποδομών	8
2.4.1.	Τι είναι ένα ΠΔΕΚΔ	8
4.	ΤΟ ΠΕΡΙΒΑΛΛΟΝ ΔΙΑΧΕΙΡΙΣΗΣ ΚΑΙ ΕΠΕΞΕΡΓΑΣΙΑΣ ΚΟΙΝΩΝΙΚΩΝ ΔΕΔΟΜΕΝΩΝ (ΠΔΕΚΔ).....	9
3.1.	Γιατί ένα Περιβάλλον.....	9
3.1.1.	Πλεονεκτήματα στο σχεδιασμό	9
3.1.2.	Πλεονεκτήματα στην αναζήτηση	10
3.1.3.	Πλεονεκτήματα στην επεξεργασία	10
3.1.4.	Πλεονεκτήματα στην τεκμηρίωση.....	10
3.1.5.	Πλεονεκτήματα στην συντήρηση	10
3.2.	Τα χαρακτηριστικά και οι λειτουργίες ενός ΠΔΕΚΔ	10
3.2.1.	Υπηρεσίες σε επίπεδο Η/Υ και λειτουργικού συστήματος.....	11
3.2.2.	Υπηρεσίες διαχείρισης δεδομένων	11
3.2.3.	Υπηρεσίες διαχείρισης πληροφοριών	11
3.2.4.	Υπηρεσίες ελέγχου των παραγωγικών διαδικασιών.....	11
3.2.5.	Υπηρεσίες μεθοδολογικής υποστήριξης.....	12
3.2.6.	Υπηρεσίες υποστήριξης στρατηγικών αξόνων	12
3.3.	Η δομή ενός ΠΔΕΚΔ	12
3.3.1.	Ο πυρήνας διαχείρισης των δεδομένων.....	12
3.3.2.	Το στρώμα των βασικών εργαλείων.....	13
3.3.3.	Το στρώμα των εργαλείων για την υποστήριξη των χρηστών.....	13
3.4.	Η ανάπτυξη του ΠΔΕΚΔ με τη βοήθεια ενός ΣΣΔΒΔ	14
3.4.1.	Οι ιδιαιτερότητες του ΠΔΕΚΔ	14
3.4.2.	Η υλοποίηση του Π.Δ.Ε.Κ.Δ. σε ένα Σ.Σ.Δ.Β.Δ.....	14
5.	ΣΥΜΠΕΡΑΣΜΑΤΑ	15
6.	ΒΙΒΛΙΟΓΡΑΦΙΚΕΣ ΑΝΑΦΟΡΕΣ.....	16

1. ΕΙΣΑΓΩΓΗ

Στις κοινωνικές επιστήμες εμφανίζεται ένα σημαντικό πρόβλημα δεδομένων. Η ανάλυση και η ερμηνεία είναι οι διαδικασίες που ορίζουν τα δεδομένα, που τα συγκροτούν ως μέρος ενός ευρύτερου θεωρητικού εγχειρήματος. Η ανάλυση και η ερμηνεία προϋποθέτουν ταυτόχρονα τα δεδομένα. Ανάλυση και καταγραφή αποτελούν επομένως τις δύο όψεις του ίδιου νομίσματος. Καμιά δεν μπορεί να υπάρξει χωρίς την άλλη. Η αναγνώριση αυτής της πραγματικότητας οδήγησε σε μια προφανή ερευνητική επιλογή. Ανάλυση και καταγραφή να αποτελέσουν τα στάδια μιας ενιαίας ερευνητικής διαδικασίας που συγκροτείται σαν ένα ενιαίο έργο μιας συγκεκριμένης ερευνητικής ομάδας. Το προϊόν της καταγραφής είναι ένα δομημένο σώμα δεδομένων που ονομάζουμε σύνολο δεδομένων. Η δομή ενός συνόλου δεδομένων καθορίζεται με βάση τις ανάγκες της ανάλυσης και τους περιορισμούς της καταγραφής και περιγράφεται από ένα συγκεκριμένο μοντέλο δεδομένων που προσδιορίζει τις απαιτούμενες μεταβλητές καθώς και τις μεταξύ τους σχέσεις. Η ενότητα ανάλυσης και καταγραφής διασφαλίζεται μέσα από το κοινό σύνολο δεδομένων στο οποίο αναφέρονται με τον καλύτερο τρόπο. Αυτή η μεθοδολογική προσέγγιση υπήρξε αποτελεσματική και κυριάρχησε στην εμπειρική έρευνα διασφαλίζοντας μάλιστα την πρωταρχική συσσώρευση δεδομένων με τη μορφή ανεξάρτητων συνόλων δεδομένων. Η φύση ωστόσο των κοινωνικών φαινομένων έδειξε γρήγορα τα όρια αυτής της διαδικασίας.

Τα κοινωνικά φαινόμενα είναι πολύπλοκα. Η πολυπλοκότητα αυτή μπορεί σχηματικά να περιγραφεί από τρία χαρακτηριστικά: α) είναι δυναμικά και εξελίσσονται στο χρόνο, β) είναι στατιστικού χαρακτήρα, αναφέρονται δηλαδή σε ένα πληθυσμό και γ) αλληλεπιδρούν και εμφανίζουν ασυνέχειες με αποτέλεσμα να μην μπορούν να αναπαρασταθούν παρά ως τοπικότητες στο χωροχρόνο πεδίο.

Λόγω αυτών των χαρακτηριστικών ο επιχειρησιακός ορίζοντας ενός μόνο έργου αποδεικνύεται ανεπαρκής για τη παραγωγή των δεδομένων που απαιτούνται για την ερμηνεία ή ακόμη και για την περιγραφή ευρύτερων κοινωνικών φαινομένων. Η εμπειρική έρευνα περιορίζεται έτσι σε φαινόμενα πολύ πιο εντοπισμένα τόσο χωροχρονικά όσο ως προς τον πληθυσμό αναφοράς.

Ακόμη όμως και αν η δειγματοληψία κατορθώνει να ξεπερνά τα προβλήματα της αντιπροσωπευτικότητας του παρατηρούμενου πληθυσμού, τα προβλήματα της συγκυρίας παραμένουν ενεργά. Η επανάληψη της καταγραφής σε διαφορετικές χωροχρονικές συντεταγμένες εμφανίζεται ως μονόδρομος για την αναγνώριση τάσεων σε ευρύτερα κοινωνικά φαινόμενα. Αυτή η προσέγγιση ωστόσο οδηγεί σε μια διαφορετική μεθοδολογική αντιμετώπιση: πολλαπλές καταγραφές οργανωμένες με το ίδιο σχήμα δεδομένων επιτρέπουν την απαιτούμενη συσσώρευση. Παρά τη χρήση ενός σταθερού σχήματος για όλες τις καταγραφές, η συσσώρευση των δεδομένων που επιτυγχάνεται δημιουργεί ένα σύνολο δεδομένων ανοικτό σε πολλαπλές αναγνώσεις, που επιτυγχάνονται μέσα από τη δευτερογενή ανάλυση. Η γενικευμένη χρήση της δευτερογενούς ανάλυσης οδηγεί ωστόσο σε μια εντελώς νέα πραγματικότητα: καταγραφή και ανάλυση καθίστανται σχετικά αυτόνομες διαδικασίες που μπορούν να επαναλαμβάνονται ανεξάρτητα. Ο τρόπος που επικοινωνούν είναι πλέον η αναφορά σε ένα κοινό σύνολο δεδομένων. Αυτή η αυτόνομη συνεπάγεται τελικά ένα συγκεκριμένο καταμερισμό εργασίας ανάμεσα στους παραγωγούς και αναλυτές δεδομένων. Αυτή η αυξανόμενη αυτόνομη ανάλυση και καταγραφής έχει ωστόσο και μία δεύτερη επίπτωση. Το παραγόμενο κατά την καταγραφή σύνολο δεδομένων δεν καλύπτει απολύτως τις ανάγκες της ανάλυσης. Ο αναλυτής οφείλει πρώτα να το μετασχηματίσει, ή να το συμπληρώσει με δεδομένα από άλλα σύνολα δεδομένων, εφόσον είναι βέβαια δυνατό. Ξεκινά έτσι μία διαδικασία ομογενοποίησης δεδομένων εκ των υστέρων. Ο αναλυτής παράγει στη συγκεκριμένη περίπτωση τα δεδομένα του δευτερογενώς.

Για όλους τους παραπάνω λόγους φαίνεται ότι το πρόβλημα της συσσώρευσης των δεδομένων καθίσταται κεντρικό για τις κοινωνικές επιστήμες, ιδιαίτερα για την πρακτική αξιοποίησή τους. Η συσσώρευση των δεδομένων δεν είναι φυσικά δυνατόν να επιτευχθεί από μία μόνη ερευνητική ομάδα. Επιπλέον η συσσώρευση δεν είναι απλά μια διαδικασία προσθήκης δεδομένων σε ένα θησαυρό. Τα δεδομένα που προστίθενται σε ένα σύνολο δεδομένων πρέπει να είναι συμβατά με τα προηγούμενα. Αυτή η συμβατότητα είναι δυνατή μέσα από μία διαδικασία που ονομάζουμε ομογενοποίηση. Η ομογενοποίηση μπορεί να προκύψει είτε ως αποτέλεσμα καταγραφών που οργανώνονται με το ίδιο μοντέλο δεδομένων. Η συσσώρευση λοιπόν των δεδομένων προϋποθέτει την παραγωγή τους σύμφωνα με ένα ενιαίο μοντέλο. Αυτή υπήρξε εξάλλου η αιτία που συνέδεσε την ανάλυση με την καταγραφή σε μια ενιαία ερευνητική διαδικασία. Η ενιαία διαδικασία ανάλυσης και καταγραφής επιβάλει ωστόσο όρια τόσο στην καταγραφή

όσο και στην ανάλυση και περιορίζει τις εμπλεκόμενες ερευνητικές ομάδες. Αντίθετα η διάκριση της ανάλυσης από την καταγραφή επιτρέπει μεγαλύτερο καταμερισμό εργασίας διαχωρίζοντας σαφέστερα τους ρόλους όσων εμπλέκονται στην παραγωγική διαδικασία.

Όπως και κάθε καταμερισμός εργασίας μια τέτοια οργάνωση της παραγωγικής διαδικασίας απαιτεί υποστήριξη από ανάλογη τεχνολογία και υποδομές ώστε να διασφαλίζεται η ασύγχρονη λειτουργία ανάλυσης και καταγραφής από μία μηχανική μνήμη. Στόχος αυτής της εργασίας είναι να προσδιορίσει τα χαρακτηριστικά και τις λειτουργίες των απαιτούμενων υποδομών.

2. ΟΙ ΕΡΕΥΝΗΤΙΚΕΣ ΥΠΟΔΟΜΕΣ ΤΩΝ ΚΟΙΝΩΝΙΚΩΝ ΕΠΙΣΤΗΜΩΝ.

Έχουν οι κοινωνικές επιστήμες ερευνητικές υποδομές;

Οι κοινωνικές επιστήμες, όπως κάθε επιστήμη, είναι υποχρεωμένες να ελέγχουν τους ισχυρισμούς τους μέσα από τις πειραματικές συναλλαγές που οι ίδιες προσδιορίζουν για τον εαυτό τους (βλέπε Κάλλας σελ. 30).

Τα δεδομένα που απαιτούν οι κοινωνικές επιστήμες καθορίζονται από το αντικείμενο τους.

Το αντικείμενο τους είναι τα κοινωνικά φαινόμενα

Τα κοινωνικά φαινόμενα έχουν τα εξής χαρακτηριστικά.

- Χαρακτηρίζονται από έντονη χωροχρονική τοπικότητα που απαιτεί συστηματική καταγραφή
- Είναι στατιστικού χαρακτήρα. Διαμορφώνονται δηλαδή από την δραστηριότητα πληθυσμών και όχι ατόμων. Οι παρατηρήσεις ωστόσο δεν είναι συνήθως δυνατόν να συγκεντρώνονται παρά σε επίπεδο ατόμων. Απαιτείται λοιπόν η συσσώρευση μεγάλου πλήθους παρατηρήσεων
- Ορίζονται σε πολλαπλά επίπεδα γενίκευσης
- Αλληλεπιδρούν μεταξύ τους σε μεγάλο βαθμό

Απαιτούν επομένως μία μεγάλη συσσώρευση δεδομένων, μία δόμηση αυτών των δεδομένων έτσι ώστε να απεικονίζονται και οι σχέσεις ανάμεσα στα μετρούμενα μεγέθη. Αυτοί οι περιορισμοί απαιτούν συγκεκριμένη προσέγγιση για τη διαχείριση και την επεξεργασία των δεδομένων που απαιτούν και ειδικά μαθηματικά (σχεσιακή άλγεβρα). Ακριβώς λόγω του όγκου των δεδομένων η διαχείριση αυτή είναι αδύνατη χωρίς ειδικές τεχνολογικές υποδομές.

Οι τεχνολογικές υποδομές των κοινωνικών επιστημών απαιτούν δυο βασικούς τύπους εργαλείων :

- α. Μεγάλες και πολύπλοκες βάσεις δεδομένων
- β Ένα περιβάλλον διαχείρισης των δεδομένων

2.1. Οι μεγάλες και πολύπλοκες βάσεις δεδομένων.

Με τον όρο μεγάλες και πολύπλοκες βάσεις δεδομένων εννοούμε μεγάλους όγκους δεδομένων που προσδιορίζονται από ένα μεγάλο πλήθος μεταβλητών δομημένων κατά πολύπλοκο τρόπο.

2.2. Το περιβάλλον διαχείρισης των δεδομένων.

Κάθε βάση δεδομένων απαιτεί εξ ορισμού ένα σύνολο από εργαλεία διαχείρισης. Τα εργαλεία αυτά συνήθως οργανώνονται σε ένα σύστημα που ονομάζεται Σύστημα Διαχείρισης Βάσεων Δεδομένων. Η εξέλιξη της τεχνολογίας έχει οδηγήσει σε αρκετά τέτοια συστήματα. Η πιο δημοφιλής κατηγορία τέτοιων συστημάτων υπήρξαν τα σχεσιακά. Ένα ΣΔΒΔ δεν είναι ωστόσο ουδέτερο απέναντι στην οργάνωση μιας βάσης δεδομένων εφόσον διευκολύνει ή παρεμποδίζει συγκεκριμένους τρόπους οργάνωσης καθοδηγώντας έτσι ουσιαστικά το αποτέλεσμα.

3. ΤΟ ΔΙΦΑΣΙΚΟ ΜΕΘΟΔΟΛΟΓΙΚΟ ΥΠΟΔΕΙΓΜΑ

Για την αντιμετώπιση του προβλήματος της συσσώρευσης και την ανάπτυξη των απαιτούμενων ερευνητικών υποδομών προτάθηκε μια νέα μεθοδολογική προσέγγιση της εμπειρικής έρευνας που ονομάστηκε διφασικό μεθοδολογικό υπόδειγμα (βλέπε Κάλλας[2001]). Σύμφωνα με την υπόθεση του διφασικού μεθοδολογικού υποδείγματος η διαδικασία παραγωγής δεδομένων χωρίζεται σε δύο διακριτές μεταξύ τους φάσεις: την φάση της καταγραφής και τη φάση της ανάλυσης. Οι δύο αυτές φάσεις δεν ορίζουν τόσο δύο διαφορετικά βήματα σε χρονική αλληλουχία όπου το ένα απλώς προηγείται του άλλου σε μία γραμμική παραγωγική διαδικασία που ξεκινά με τη συγκρότηση ενός συστήματος μεταβλητών και τελειώνει με την παρουσίαση των επεξεργασμένων αποτελεσμάτων στο πλαίσιο μιας και μόνης έρευνας. Αντίθετα υποστηρίζεται ότι η παραγωγή δεδομένων είναι τελικά μια συλλογική διαδικασία που συγκροτείται από ένα σύνολο αυτόνομων φορέων και επιτυγχάνεται μέσα από ένα σύστημα αυτόνομες και ασύγχρονες λειτουργίες οι οποίες αλληλεπιδρούν μεταξύ τους με την βοήθεια σταθερών υποδομών επιστημονικής επικοινωνίας. Αυτές οι υποδομές συγκροτούν τα σημεία συσσώρευσης των δεδομένων. Συσσωρεύουν λοιπόν τα σύνολα δεδομένων πολλών ανεξάρτητων μεταξύ τους καταγραφών. Όπως ωστόσο κάθε συλλογική παραγωγή έτσι και αυτή δεν είναι τελικά τόσο αυτονομημένη όσο φαίνεται. Οι παραγωγοί είναι υποχρεωμένοι να προσαρμόζονται στις ανάγκες των αναλυτών και οι αναλυτές πρέπει να συμβιβάζονται με αυτό που οι παραγωγοί μπορούν να προσφέρουν. Την ποιότητα της συλλογικής παραγωγής διασφαλίζει επομένως το σύστημα των διαθέσιμων υποδομών και η πρόσβαση των παραγωγών σε αυτές. Σημειώνουμε ότι τα ίδια τα παραγόμενα σύνολα δεδομένων αυτονομούνται τελικά τόσο από την ανάλυση και αποτελούν μέρος της υποδομής στην οποία συσσωρεύονται. Η διαχείριση αυτών των δεδομένων, η δημιουργία συστημάτων ταξινόμησης και τεκμηρίωσης τους συγκροτεί σιγά-σιγά μια αυτόνομη λειτουργία και ένα νέο ρόλο μέσα στο νέο καταμερισμό εργασίας αυτόν του διαθέτη.

3.1. Η παραγωγικές διαδικασίες

Σύμφωνα λοιπόν με την υπόθεση του διφασικού μεθοδολογικού υποδείγματος η παραγωγή των δεδομένων οργανώνεται συλλογικά με τη βοήθεια ενός καταμερισμού εργασίας που προβλέπει διαφορετικούς ρόλους και παραγωγικές διαδικασίες οργανωμένες σε δύο μεγάλες φάσεις, της καταγραφής και της ανάλυσης που αλληλεπιδρούν με τη βοήθεια συγκεκριμένων υποδομών.

Μπορούμε σχηματικά να περιγράψουμε αυτές τις παραγωγικές διαδικασίες:

- α. Η πρωτογενής καταγραφή
- β. Η δευτερογενής καταγραφή
- γ. Η ομογενοποίηση
- δ. Η αρχική επεξεργασία
- ε. Η κυρίως ανάλυση.

Θα ονομάσουμε πρωτογενή καταγραφή τη διαδικασία σύμφωνα με την οποία παράγονται κοινωνικά δεδομένα από την απευθείας ερώτηση ή παρατήρηση των φορέων της δράσης.

Θα ονομάσουμε δευτερογενή καταγραφή τη διαδικασία κατά την οποία παράγονται κοινωνικά δεδομένα από την παρατήρηση των ιχνών που αφήνουν οι φορείς της δράσης

Θα ονομάσουμε ομογενοποίηση τη διαδικασία μετασχηματισμού δεδομένων με στόχο τη δημιουργία συνόλων δεδομένων κατάλληλων για συγκεκριμένου τύπου αναλύσεις σύμφωνα με ένα συνεκτικό και σχετικά σταθεροποιημένο εννοιολογικό σύστημα (Ομογενοποίηση, χρονοσειρές, μακροδοδεδομένα).

Θα ονομάσουμε αρχική επεξεργασία τη διαδικασία μετασχηματισμού των δεδομένων με στόχο τη δημιουργία των κατάλληλων μεταβλητών ενός συγκεκριμένου αναλυτικού κώδικα.

Θα ονομάσουμε κυρίως ανάλυση τη διαδικασία μετασχηματισμού των δεδομένων με στόχο την τεκμηρίωση συγκεκριμένων υποθέσεων.

Η πρωτογενής και η δευτερογενής καταγραφή αποτελούν διαδικασίες που εντάσσονται σαφώς στην φάση της καταγραφής ενώ η αρχική επεξεργασία και η κυρίως ανάλυση σαφώς στη φάση της ανάλυσης.

Όλες οι διαδικασίες που περιγράψαμε είτε ως διαδικασίες που ανήκουν στη φάση της καταγραφής είτε ως διαδικασίες που ανήκουν στη φάση της ανάλυσης μπορούν να αποτελέσουν τις διαδικασίες είτε ενός μεμονωμένου έργου είτε περισσότερων διαφορετικών έργων. Παρόλα αυτά η διάκριση ισχύει διότι δεν

αναφέρεται απλά σε μία χρονική αλληλουχία

Ένα σύνολο δεδομένων είναι ένα δομημένο σώμα δεδομένων που μπορεί να εξελίσσεται τόσο ως προς τη δομή όσο και ως προς το περιεχόμενο αλλά που κάθε φορά έχει μία μόνο συγκεκριμένη δομή. Ένα σύνολο δεδομένων εξελίσσεται διότι:

α. αλλάζουν οι τιμές ταξινομητικών μεταβλητών καθώς εξελίσσεται η κοινωνία γεγονός που συνεπάγεται την μεταβολή της δομής του συνόλου.

β. με τη συνεχή συσσώρευση δεδομένων που προέρχονται από διαφορετικές καταγραφές δημιουργείται η ανάγκη προσθήκης νέων στοιχείων δεδομένων.

γ. αλλάζει το ίδιο το εννοιολογικό μοντέλο ως αποτέλεσμα των εξελίξεων στη θεωρία

Η ομογενοποίηση είναι λοιπόν μία πρόσθετη ταξινομητική διαδικασία που προκύπτει ως αποτέλεσμα της ασύγχρονης παραγωγής των επί μέρους στοιχείων που συνθέτουν τελικά ένα σύνολο δεδομένων. Με αυτή την έννοια η ομογενοποίηση είναι μια διαδικασία που πρέπει να ενταχθεί στη φάση της καταγραφής παρόλο που είναι μια διαδικασία καθοδηγημένη εξ ολοκλήρου από τη διαδικασία της ανάλυσης.

Η παραγωγή θεματικών βάσεων δεδομένων είναι λοιπόν μια διαδικασία ομογενοποίησης και ως τέτοια παρουσιάζει ορισμένα σταθερά χαρακτηριστικά:

α. Είναι μία διαδικασία ταξινομητική.

β. Είναι μία διαδικασία που προϋποθέτει τη γνώση των αναγκών της ανάλυσης

γ. Είναι μία διαδικασία δευτερογενής, που προϋποθέτει δηλαδή την ύπαρξη προηγούμενων καταγραφών

Σύμφωνα με όσα είπαμε παραπάνω, η διαδικασία παραγωγής δεδομένων δεν πρέπει να θεωρείται υπόθεση μιας μεμονωμένης ερευνητικής ομάδας αλλά μια συλλογική εργασία που αφορά όλους εκείνους που με τον ένα ή τον άλλο τρόπο παράγουν, διαχειρίζονται ή χρησιμοποιούν κοινωνικά δεδομένα και που με τον ένα ή τον άλλο τρόπο αλληλεξαρτώνται. Έτσι ακόμη και όταν μια απομονωμένη ερευνητική ομάδα παράγει από την αρχή τα δικά της δεδομένα για να στηρίξει τη δική της ανάλυση υπόκειται σε περιορισμούς ως προς τον ορισμό αυτών των δεδομένων που επιβάλλει μια ευρύτερη κοινότητα. Οι περιορισμοί αυτοί αφορούν τον ορισμό των δεδομένων, τα εργαλεία επεξεργασίας τους, τον τρόπο παρουσίασης και διάχυσης των δεδομένων, τις πιθανές χρήσεις τους κ.λ.π.

Η προσαρμογή της ομάδας σε αυτούς τους περιορισμούς γίνεται με την αναφορά στο έργο άλλων επιστημόνων αλλά για ορισμένους από αυτούς με την χρήση κατάλληλων εργαλείων, που ταυτόχρονα επιταχύνουν την παραγωγή.

Η εξέλιξη λοιπόν της τεχνολογίας επιτρέπει σήμερα την ανάπτυξη σύνθετων υποδομών ικανών να υποστηρίζουν με ολοκληρωμένο τρόπο πολύ περισσότερες από αυτές τις λειτουργίες και να επιτρέψουν καλύτερη και ευκολότερη προσαρμογή στους περιορισμούς της επιστήμης. Τέτοιες υποδομές βρίσκονται σήμερα σε εξέλιξη και στόχος αυτής της εργασίας είναι να παρουσιάσει ορισμένα από τα βασικά χαρακτηριστικά τους.

Εννοιολογική εξάρτηση από ορισμούς και από μοντέλα, αναφορές σε ορισμούς και σε δημοσιεύσεις, και σε κείμενα λόγω της σχετικής σημασίας.

3.2. Η φάση της καταγραφής

Σύμφωνα με το διφασικό μεθοδολογικό υπόδειγμα στόχος της καταγραφής είναι η παραγωγή δεδομένων κατάλληλων να χρησιμοποιηθούν από όσο το δυνατόν περισσότερες αναλύσεις και ταυτόχρονα κατάλληλων να ομογενοποιηθούν κατά το δυνατόν ώστε να καταστεί δυνατή η συσσώρευση δεδομένων. Η καταγραφή είναι λοιπόν μια λειτουργία που δίνει προτεραιότητα στην ταξινόμηση απέναντι στον υπολογισμό.

Η ασύγχρονη παραγωγή του καταγραφικού σχήματος έγινε από την ανάγκη να εκφράσει ως μεταμοντέλο τη δομή του διφασικού της ταυτόχρονης δηλαδή ύπαρξης πολλών σχημάτων δεδομένων: ενός του τύπου

παρατηρησιακό και πολλά του τύπου αναλυτικό.

Η δομή έχει νόημα διότι επιτρέπει πράξης ανάμεσα στα σύνολα δεδομένων (σχεσιακές πράξεις, όχι πράξης διοίκησης αλλά πράξης λογικής οργάνωσης όπως πχ τα virtual data sets, semantic dependency, ανάμεσα στα data sets που ελέγχεται αντίστοιχα όπως και το referential integrity (είναι λοιπόν ανάγκη να αναπτυχθούν εργαλεία διαχείρισης τους

3.3. Η φάση της ανάλυσης

Η φάση της ανάλυσης περιλαμβάνει το σύνολο των αναλύσεων που μπορούν να γίνουν επάνω σε ένα συγκεκριμένο σύνολο δεδομένων. Κάθε διαφορετική ανάλυση βασίζεται σε μια διαφορετική εικόνα του συγκεκριμένου συνόλου δεδομένων που προκύπτει από την ιδιαίτερη οπτική της και περιγράφεται με τις υποθέσεις εργασίας της. Αυτή η διαφορετική εικόνα αποτυπώνεται σε ένα συγκεκριμένο σχήμα δεδομένων που μετασχηματίζει μέρος ή το σύνολο των στοιχείων δεδομένων του αρχικού συνόλου σε μια νέα δομή δεδομένων που αντιστοιχεί στις μεταβλητές της ανάλυσης. Η μετάβαση από το σχήμα της καταγραφής στο σχήμα της ανάλυσης γίνεται με τη βοήθεια μετασχηματισμών που βασίζονται κυρίως σε σχεσιακές πράξεις και δευτερευόντως σε πράξεις υπολογισμού.

Αντίθετα η κυρίως επεξεργασία, που υπολογίζει από το σύνολο δεδομένων της ανάλυσης συγκεκριμένα αποτελέσματα βασίζεται σε υπολογισμούς

Συχνά το σύνολο των δεδομένων που αντιστοιχεί στο σχήμα της ανάλυσης περιγράφει καλύτερα το ίδιο το παρατηρούμενο φαινόμενο και μπορεί να αντικαταστήσει πλήρως ή εν μέρει το σχήμα καταγραφής και να αποτελέσει τη βάση για τις μετέπειτα αναλύσεις. Σε αυτή τη περίπτωση έχουμε μία εξέλιξη του σχήματος καταγραφής.

Για να είναι δυνατή η κοινή διαχείριση τόσο των διαφορετικών αναλυτικών σχημάτων όσο και του σχήματος καταγραφής είναι ανάγκη να περιγράφονται τα σχήματα αυτά με ένα ενιαίο μεταμοντέλο.

Διάκριση ανάμεσα στη στατιστική επεξεργασία και την οργάνωση των δεδομένων

3.4. Η λειτουργία των υποδομών

Σύμφωνα με όσα είπαμε υποστηρίζοντας το διφασικό μεθοδολογικό υπόδειγμα, η καταγραφή καθίσταται συλλογικό ζήτημα από τη στιγμή που οφείλει να ανοιχθεί σε περισσότερες από μία αναλύσεις. Εφόσον λοιπόν τα δεδομένα της καταγραφής καθίστανται αντικείμενο διαχείρισης και επεξεργασίας πολλών ανεξάρτητων μεταξύ τους ερευνητών η οργάνωση και η διαχείριση τους απαιτεί ένα σταθερό σύστημα διαχείρισης βασισμένο σε σταθερές προδιαγραφές. Αυτές οι σταθερές προδιαγραφές καθορίζονται από τη διαθέσιμη σήμερα τεχνολογία για τη διαχείριση δεδομένων που υλοποιείται με την ανάπτυξη συστημάτων διαχείρισης βάσεων δεδομένων (ΣΔΒΔ). Τα ΣΔΒΔ είναι κεντρικά ή κατακεντρωμένα συστήματα λογισμικού που παρέχουν εργαλεία για τον ορισμό βάσεων δεδομένων, για την επιλογή δομών δεδομένων κατάλληλων για την αποθήκευση και την αναζήτηση των δεδομένων. Η τεχνολογία αυτή εξελίσσεται συνεχώς τα τελευταία 40 περίπου χρόνια.

3.4.1. Τι είναι ένα ΠΔΕΚΔ

Το Περιβάλλον Διαχείρισης και Επεξεργασίας Κοινωνικών Δεδομένων είναι ένα σύνολο από εργαλεία λογισμικού που επιτρέπουν την οργάνωση, την διαχείριση και την επεξεργασία κοινωνικών δεδομένων. Υλοποίηση τέτοιου περιβάλλοντος αποτελεί η Τράπεζα Κοινωνικών Δεδομένων του ΕΚΚΕ

Ένα ΠΔΕΚΔ πρέπει λοιπόν να αντιμετωπίζεται ως μια εφαρμογή αυτής της τεχνολογίας.

Μια τέτοια επιλογή διασφαλίζει τα ακόλουθα:

1. no redundancy
2. Referential integrity
3. Entity integrity

4. ΤΟ ΠΕΡΙΒΑΛΛΟΝ ΔΙΑΧΕΙΡΙΣΗΣ ΚΑΙ ΕΠΕΞΕΡΓΑΣΙΑΣ ΚΟΙΝΩΝΙΚΩΝ ΔΕΔΟΜΕΝΩΝ (ΠΔΕΚΔ)

4.1. Γιατί ένα Περιβάλλον

Στο βαθμό που η παραγωγή των δεδομένων καθίσταται μία πολύπλοκη συλλογική διαδικασία που στηρίζεται στη συσσώρευση, η ανάλυση και η επεξεργασία των δεδομένων καθίσταται επίσης ένα πολύπλοκο εγχείρημα. Καθώς ο καταμερισμός της εργασίας επιβάλει το χωρισμό της καταγραφής από την ανάλυση, εφόσον ο αναλυτής δεν παράγει τα δεδομένα που χρησιμοποιεί αλλά πρέπει ταυτόχρονα να τα διαμορφώνει ώστε να καθίστανται κατάλληλα για τις ανάγκες της ανάλυσης του και συγχρόνως ο παραγωγός αντιμετωπίζει τις δυσκολίες της καταγραφής με βάση τις προτεραιότητες της ανάλυσης στην οποία ωστόσο δεν συμμετέχει είναι απαραίτητη η συγκρότηση μιας γλώσσας ενός κώδικα παρατήρησης. Ο κώδικας αυτός υποστηρίζει όλα τα εργαλεία τόσο της ανάλυσης όσο και της αναγραφής και τα ολοκληρώνει από άποψη προδιαγραφών σε ένα ενιαίο περιβάλλον διαχείρισης και επεξεργασίας.

Το περιβάλλον παρέχει ορισμένα βασικά πλεονεκτήματα:

4.1.1. Πλεονεκτήματα στο σχεδιασμό

Ο σχεδιασμός πρέπει να διασφαλίζει εννοιολογική αυστηρότητα και ακεραιότητα των δεδομένων

Η εννοιολογική αυστηρότητα

Η εννοιολογική αυστηρότητα δεν είναι μόνο πρόβλημα ορισμού της έννοιας αλλά κυρίως ένταξης της έννοιας σε ένα συγκεκριμένο εννοιολογικό πλαίσιο.

Λόγω του καταμερισμού εργασίας και της πολυπλοκότητας των κοινωνικών φαινομένων οι έννοιες πρέπει να ανήκουν σε ένα ευρύτερα αποδεκτό ταξινομικό σχήμα μέρος μόνο του οποίου μπορεί να κατασκευαστεί από μία ερευνητική ομάδα.

Η ακεραιότητα

Η ακεραιότητα των δεδομένων περιλαμβάνει την ακεραιότητα των οντοτήτων και την ακεραιότητα των αναφορών.

Η χρήση σχεσιακών μοντέλων διασφαλίζει την αυστηρότητα στις σχεσιακές πράξεις που απαιτούνται.

4.1.2. Πλεονεκτήματα στην αναζήτηση

Η αναζήτηση είναι μία πράξη που αποκτά ιδιαίτερη σημασία για ένα περιβάλλον πολλαπλών χρηστών. Από τη στιγμή λοιπόν που η ανάγκη για συσσώρευση δεδομένων και για δευτερογενή, συγκριτική ή διαχρονική ανάλυση αποκτά προτεραιότητα γίνεται κρίσιμη για τη διαχείριση των δεδομένων εφόσον είναι αυτή που επιτρέπει την δυνατότητα αξιοποίησης των δεδομένων από ερευνητές που δεν ανήκουν στην ερευνητική ομάδα που τα παρήγε.

Η αναζήτηση απαιτεί τη διασφάλιση ενιαίου ταξινομικού συστήματος ανεξάρτητου από το συγκεκριμένο κάθε φορά σύνολο δεδομένων ώστε να καθίσταται δυνατή από οποιοδήποτε ερευνητή.

4.1.3. Πλεονεκτήματα στην επεξεργασία

Η επεξεργασίες που είναι απαραίτητες για την ανάλυση ενός συνόλου δεδομένων είναι δύο ειδών:

α. Επεξεργασίες σχεσιακού τύπου

β. Επεξεργασίες υπολογιστικού τύπου

Η ανάπτυξη των συνόλων δεδομένων σύμφωνα με το μονοφασικό μεθοδολογικό υπόδειγμα περιορίζει την επεξεργασία των δεδομένων στον στατιστικό υπολογισμό. Εφόσον λοιπόν το σχήμα της ανάλυσης ταυτίζεται με το σχήμα της μεταγραφής και ο όγκος των δεδομένων παρέμενε σχετικά περιορισμένος, οι ανάγκες για σχεσιακές πράξεις παρέμεναν περιορισμένες. Ιδιαίτερα εφόσον δεν υπήρχε πρόβλεψη αξιοποίηση των δυνατοτήτων που προσφέρουν τα σχεσιακά μοντέλα δεδομένων στο σωστό σχεδιασμό.

4.1.4. Πλεονεκτήματα στην τεκμηρίωση

Η τεκμηρίωση απαιτεί πρόσθετες πληροφορίες σχετικά με τον τρόπο παραγωγής των δεδομένων (μεταδεδομένα). Τα μεταδεδομένα αυτά πρέπει να αποθηκεύονται και να ανασύρονται μαζί με τα δεδομένα.

Μεταδεδομένα που αφορούν στη σχεδίαση (ορισμοί, παραδοχές, ταξινόμηση)

Μεταδεδομένα που αφορούν στη συγκυρία

Μεταδεδομένα που αφορούν στην εξέλιξη

4.1.5. Πλεονεκτήματα στην συντήρηση

Ασύγχρονη συμπλήρωση και διόρθωση των δεδομένων που επιτρέπει ένα διευρυμένο καταμερισμό εργασίας ανάμεσα σε πολλούς παραγωγούς με σχετικά μειωμένο συντονισμό ο οποίος επιτυγχάνεται μέσω των υποδομών και της γνώσης που ενσωματώνεται σε αυτές.

Διαχείριση διαφορετικών εκδόσεων.

4.2. Τα χαρακτηριστικά και οι λειτουργίες ενός ΠΔΕΚΔ

Ένα τέτοιο περιβάλλον έχει ως κύριο στόχο να επιτρέψει σε πολλές διαφορετικές ερευνητικές ομάδες να επεξεργάζονται και να παράγουν δεδομένα. Το περιβάλλον αυτό θα πρέπει να εξυπηρετεί τις ακόλουθες ανάγκες:

α) Να είναι ολοκληρωμένο. Να καλύπτει δηλαδή ένα σημαντικό φάσμα λειτουργιών, κατά τρόπο που να επιτρέπει την κοινή χρήση των δεδομένων,

β) Να υποστηρίζει την ανάλυση

γ) Να υποστηρίζει την καταγραφή

δ) Να παρέχει πληροφόρηση

ε) Να διασφαλίζει τη διαχείριση των δεδομένων σύμφωνα με συγκεκριμένα πρότυπα.

Το ΠΔΕΚΔ ικανοποιεί τις ανάγκες αυτές μέσα από τις ακόλουθες υπηρεσίες.

4.2.1. Υπηρεσίες σε επίπεδο Η/Υ και λειτουργικού συστήματος

Η λειτουργία του διαφασικού μεθοδολογικού υποδείγματος βασίζεται στη λειτουργία μιας μηχανικής μνήμης που διασφαλίζει την ασύγχρονη επικοινωνία ανάμεσα στις διαφορετικές επιστημονικές ομάδες που

εμπλέκονται στη διαδικασία παραγωγής δεδομένων. Η λειτουργία αυτής της μηχανικής μνήμης απαιτεί υπηρεσίες σε επίπεδο υλικού και λειτουργικού συστήματος που το περιβάλλον πρέπει να διασφαλίζει.

4.2.2. Υπηρεσίες διαχείρισης δεδομένων

Το περιβάλλον πρέπει να διαθέτει στον πυρήνα του ένα σύστημα διαχείρισης δεδομένων που να επιτρέπει την κοινή χρήση των δεδομένων από πολλές διαφορετικές επιστημονικές ομάδες.

4.2.3. Υπηρεσίες διαχείρισης πληροφοριών

Εκτός ωστόσο από πρόσβαση στα δεδομένα το περιβάλλον πρέπει να παρέχει την απαιτούμενη πληροφόρηση ώστε να καθίσταται η διαχείριση των δεδομένων αποτελεσματική. Η πληροφόρηση σε ένα τέτοιο σύστημα ξεπερνά κατά πολύ τις ανάγκες απλής τεκμηρίωσης των δεδομένων αφού απαιτεί και μια σειρά πραγματολογικές πληροφορίες που επιτρέπουν όχι απλά τη χρήση των δεδομένων αλλά και την αξιολόγηση τους και την ένταξη σε μια ευρύτερη συγκυρία.

4.2.4. Υπηρεσίες ελέγχου των παραγωγικών διαδικασιών

Το περιβάλλον πρέπει να υποστηρίζει την ταυτόχρονη εξέλιξη διαφορετικών παραγωγικών διαδικασιών. Απαραίτητη προϋπόθεση για αυτό είναι η διαχείριση των συνόλων δεδομένων.

Η διφασική φύση της παραγωγικής διαδικασίας επιτρέπει την εξέλιξη των δεδομένων τόσο ως προς το περιεχόμενο όσο και ως προς τη δομή τους. Βρισκόμαστε λοιπόν αντιμέτωποι με σχήματα δεδομένων που μεταβάλλονται είτε για τις ανάγκες μιας συγκεκριμένης ανάλυσης είτε διότι η αναπαράσταση των κοινωνικών φαινομένων απαιτεί τροποποιημένα σχήματα καταγραφής. Το περιβάλλον πρέπει λοιπόν να υποστηρίζει υπηρεσίες ελέγχου των εκδόσεων των αυτόνομων συνόλων δεδομένων (version control).

4.2.5. Υπηρεσίες μεθοδολογικής υποστήριξης

Το περιβάλλον πρέπει να διασφαλίζει την υποστήριξη συγκεκριμένων μεθόδων οργάνωσης των κοινωνικών δεδομένων παρέχοντας τα απαιτούμενα εργαλεία για την υλοποίησή τους και επιβάλλοντας έτσι συγκεκριμένους περιορισμούς και τυποποιήσεις. Το περιβάλλον επιβάλλει λοιπόν μεθοδολογίες που ξεπερνούν τον μεμονωμένο χρήστη και διασφαλίζουν τη συλλογικότητα της παραγωγικής διαδικασίας.

4.2.6. Υπηρεσίες υποστήριξης συγκεκριμένων στρατηγικών

Τα δεδομένα είναι οργανωμένα σε αυτόνομα σύνολα δεδομένων. Η δομή τους δεν πρέπει να καθορίζεται τόσο από αυθαίρετες θεωρητικές υποθέσεις που επιβάλλει η ομάδα που τα παράγει. Αντίθετα πρέπει να καθορίζεται κυρίως από τους στρατηγικούς στόχους που καλούνται να εξυπηρετήσουν τα συγκεκριμένα σύνολα δεδομένων και οι οποίοι μπορεί να ξεπερνούν τους στόχους τους ορίζοντες της ομάδας παραγωγής. Έτσι εφόσον τα δεδομένα αποτελούν προϊόν συλλογικής εργασίας μέσα στο χρόνο αναπτύσσονται και εξελίσσονται σύμφωνα με ένα στρατηγικό άξονα που υπερκαθορίζει τους επί μέρους στόχους των μεμονωμένων επιστημονικών ομάδων. Ένας τέτοιος άξονας υλοποιείται στο σύστημα με τη μορφή συγκεκριμένων σχημάτων δεδομένων, μεταδεδομένων και πρόσθετων πληροφοριών που έχουν ως στόχο

την ομογενοποίηση περισσότερων συνόλων δεδομένων σε μια ενιαία συλλογή. Η συλλογή αυτή μπορεί να προσδιορίζεται από μία απλή θεματική ταξινόμηση συγκεκριμένων συνόλων δεδομένων, μέχρι συγκεκριμένες πράξεις και διαδικασίες ομογενοποίησης ή κοινά σχήματα δεδομένων. (Η παραγωγή θεματικών βάσεων δεδομένων όπου ενσωματώνονται δεδομένα από διάφορα αυτόνομα σύνολα δεδομένων μετά από κατάλληλους μετασχηματισμούς είναι ένα τέτοιο παράδειγμα.)

4.3. Η δομή ενός ΠΔΕΚΑ

Για την ανάπτυξη ενός τέτοιου περιβάλλοντος προτείνεται ένα μοντέλο πολλών στρωμάτων που αποτελείται από:

- Έναν πυρήνα που διασφαλίζει τη διαχείριση των δεδομένων
- Ένα στρώμα βασικών εργαλείων
- Ένα στρώμα εργαλείων για την υποστήριξη των χρηστών

4.3.1. Ο πυρήνας διαχείρισης των δεδομένων.

Ο πυρήνας του συστήματος πρέπει να εξασφαλίζει τις ακόλουθες υπηρεσίες:

1. Υπηρεσίες σε επίπεδο H/W και λειτουργικού συστήματος
2. Υπηρεσίες διαχείρισης δεδομένων
3. Υπηρεσίες διαχείρισης πληροφοριών

Το διφασικό μεθοδολογικό υπόδειγμα προβλέπει λοιπόν δύο εννοιολογικά σχήματα για την οργάνωση των κοινωνικών δεδομένων. Αυτό της ανάλυσης και αυτό της καταγραφής. Τα εννοιολογικά σχήματα υλοποιούνται με τη βοήθεια εργαλείων που επιτρέπουν την ακεραιότητα τους.

Το μεταμοντέλο για την περιγραφή όλων των δεδομένων μπορεί να περιγραφεί από ένα σχεσιακό μοντέλο δεδομένων με τρεις βασικές οντότητες: την οντότητα σύνολο δεδομένων, την οντότητα πίνακας, και την οντότητα πεδίο.

4.3.2. Το στρώμα των βασικών εργαλείων

πρέπει να διασφαλίζει τις ελάχιστες λειτουργίες ώστε όλοι οι χρήστες να ανταποκρίνονται στις προδιαγραφές του βασικού μοντέλου δεδομένων με βάση το οποίο συγκροτείται ο πυρήνας.

Εργαλεία οργάνωσης των δεδομένων

Εργαλεία για τη συντήρηση των διαθέσιμων δεδομένων

Εργαλεία για την αναζήτηση και την διάγνωση των διαθέσιμων δεδομένων

Εργαλεία για το σχεδιασμό νέων συνόλων δεδομένων

4.3.3. Το στρώμα των εργαλείων για την υποστήριξη των χρηστών

1. Υπηρεσίες ελέγχου των παραγωγικών διαδικασιών
2. Υπηρεσίες μεθοδολογικής υποστήριξης
3. Υπηρεσίες υποστήριξης συγκεκριμένων στρατηγικών

Εργαλεία στατιστικής ανάλυσης

Γεωγραφικά συστήματα δεδομένων

Εργαλεία σύνδεσης των εργαλείων υποστήριξης των χρηστών με τα βασικά εργαλεία

4.4. Η ανάπτυξη του ΠΔΕΚΔ με τη βοήθεια ενός ΣΣΔΒΔ

4.4.1. Οι ιδιαιτερότητες του ΠΔΕΚΔ

Βάση Δεδομένων για εμπορικές εφαρμογές	Βάση δεδομένων για υποστήριξη ενός Αρχείου κοινωνικών δεδομένων
Η πληροφορία που περιέχεται είναι στατική και Περιγράφεται α priori. Επομένως και το σχήμα της βάσης είναι στατικό	Συνεχής εξέλιξη των πληροφοριών, Δεδομένα που αναφέρονται στο Περιβάλλον (εργαλεία, μέθοδοι) και τα προϊόντα που παράγονται
Μεταβολές στο σχήμα είναι σπάνιες και ελέγχονται από τους διαχειριστές της βάσης	Οι αλλαγές του σχήματος αναμένονται και είναι σχετικά συχνές. Αρκετοί χρήστες χρειάζεται να κάνουν αλλαγές στο σχήμα
Τα δεδομένα είναι ατομικά και σταθερού μήκους (π.χ προτάσεις και αριθμοί)	Τα Δεδομένα είναι ατομικά αλλά ταυτόχρονα και δομημένα. Επίσης τα Πεδία μπορεί να είναι μεγάλα σύνθετα και μεταβλητού μήκους
Μικρός αριθμός τύπων οντοτήτων και μεγάλος αριθμός από εγγραφές ίδιου τύπου	Μεγάλος αριθμός οντοτήτων με λίγες σχετικά εγγραφές για κάθε τύπο. Συχνά υπάρχουν σύνθετες σχέσεις ανάμεσα στις οντότητες και δημιουργούνται νέες
Αρκετά δεδομένα εισάγονται αρχικά και αυξάνονται σταδιακά	Λίγα δεδομένα εισάγονται αρχικά και αυξάνονται γρήγορα
Πεδία με μία τιμή που ενημερώνεται επί τόπου	Οι διαφορετικές εκδόσεις των πεδίων είναι ζωτικής σημασίας
Οι διαντιδράσεις είναι μικρές, ατομικές, και μπορούν να χρησιμοποιηθούν ως βάση για την ταυτόχρονη πρόσβαση	Οι διαντιδράσεις είναι μακράς διάρκειας, και μπορούν να αφήσουν τη βάση για μεγάλο διάστημα σε μεταβατική κατάσταση

4.4.2. Η υλοποίηση του Π.Δ.Ε.Κ.Δ. σε ένα Σ.Σ.Δ.Β.Δ.

Το σχεσιακό μοντέλο οργάνωσης των δεδομένων

Το αντικειμενοστραφές μοντέλο οργάνωσης των δεδομένων

Ένα επιχειρησιακό μοντέλο για τη διαχείριση των κοινωνικών δεδομένων

Ένα αυτόνομο σύνολο δεδομένων είναι ένα σύνθετο αντικείμενο που περιλαμβάνει συγκεκριμένες οντότητες και σχέσεις. Οι οντότητες και οι σχέσεις ενός αυτόνομου συνόλου δεδομένων μπορούν να αναπαρασταθούν από ένα σχήμα σχέσεων οντοτήτων που περιγράφει το συγκεκριμένο σύνολο δεδομένων. Κάθε σύνολο δεδομένων μπορεί λοιπόν να περιγραφεί με ένα μοντέλο που συσχετίζει την οντότητα σύνολο δεδομένων με την οντότητα πίνακας.

Υπάρχουν δύο βασικά αντικείμενα

- α. Το σύνολο δεδομένων
- β. Ο πίνακας

Το προτεινόμενο περιβάλλον πρέπει να μπορεί να διαχειρίζεται ικανοποιητικά και τα δύο αντικείμενα. Το σύνολο δεδομένων μπορεί να αντιμετωπισθεί ως ένα σύνθετο αντικείμενο αποτελούμενο από πίνακες. Αυτοί οι πίνακες είναι δύο ειδών: Απλοί πίνακες δεδομένων που ποικίλουν σε κάθε διαφορετικό σύνολο δεδομένων και πίνακες πληροφοριών που παραμένουν σταθεροί και ανεξάρτητοι από το συγκεκριμένο σύνολο δεδομένων.

5. ΣΥΜΠΕΡΑΣΜΑΤΑ

Η αντιμετώπιση των ιστορικών φαινομένων απαιτεί συσσώρευση δεδομένων. Η επεξεργασία των δεδομένων αυτών απαιτεί ένα σύνθετο περιβάλλον επεξεργασίας που διασφαλίζει τόσο την σχεσιακή όσο και την υπολογιστική επεξεργασία και διαθέτει τα απαιτούμενα εργαλεία για αυτές τις πράξεις. Ταυτόχρονα ωστόσο ακριβώς λόγω της τοπικότητας που παρουσιάζει το χωροχρονικό πεδίο των ιστορικών φαινομένων, τοπικότητες που περιγράφονται με την έννοια της συγκυρίας η επεξεργασία απαιτεί ταυτόχρονα πρόσθετη πληροφορία που προσδιορίζει κατά το δυνατόν αυτή την τοπικότητα. Έτσι τα δεδομένα απαιτούν πρόσθετη πληροφορία που τα τοποθετεί μέσα στο χωροχρονικό πεδίο και η οποία αποτυπώνεται σε πολύπλοκα συστήματα ταξινόμησης. Η διαμόρφωση τέτοιων ταξινομικών συστημάτων προϋποθέτει την συσσώρευση.

6. ΒΙΒΛΙΟΓΡΑΦΙΚΕΣ ΑΝΑΦΟΡΕΣ

1. Κάλλας Γιάννης
Η πληροφορική τεχνολογία στην κοινωνική έρευνα. Το πρόβλημα των δεδομένων
Εκδόσεις Νεφέλη-ΕΚΚΕ, 2001
2. Bertino Elisa, Martino Lorenzo
Object - Oriented Database Systems
Concepts and Architectures
Addison Wesley, 1993
3. Brown W. Allan
Object Oriented Databases Applications in Software Engineering
Mc Graw Hill, 1991
4. Chan and Shoshani
SUBJECT: A Directory Driven System for Large Statistical Databases, Proc. Of the
First Working Conf. SSDBM, 1981, pp 61-64
5. Chen P.P. S.
“The entity Relationship Model -Toward a Unified View of Data”
ACM TODS 1, No 1 (March 1976)
6. Darius P, Bourcheau P. de Greef, De Feber E., Froeschl K.
Modelling Metadata
Statistical Journal of the United Nations ECE 10 (1993) pp171-179
7. Date C. J.
An Introduction to Database Systems
Vol I, II, Addison Wesley, 1986
8. De Vaus D.A.
Surveys in Social Research
UCL Press, 1996
9. Franzosi Roberto
“From Words to Numbers : A Set Theory Framework”
in Sociological Methodology Vol 24 (1994)
10. Glos, S.P
“Statistical Relation Model”
in “Statistical and Scientific Databases”
Vol (1991), pp267 - 305
11. Hyman, Herbert.
“Secondary Analysis of Sample Surveys: Principles, Procedures and Potentialities”
Wiley, N.York, 1972
12. Malvestuto, F.M.,
Universal -Schem Approach to Statistical Databases, ACM Database Systems, Vol 18,
No 4, 1993, pp 678-708.

13. Navathe S.B.
Evolution of Data Modeling for Databases
Communications of the ACM, Vol. 35, No 9, p112 - 23, 1992
14. Rafanelli M. and Shoshani A.
STORM: A Statistical Object Representation Model
Proc. of the fifth Int. Conf. On SSDBM Vol (1990) pp 14 -29
15. Rafanelli M. and Ricci F.L.
MEFISTO : A Function Model for Statistical Entities
IEEE Knowledge and Data Engineering Vol 5 (1993) pp 670 - 681
16. Soshani A. and Wong H.K.T
Statistical and Scientific Database Issues
IEEE Software Enginnering Vol 11 (1985) pp 1090 - 1097
17. Su S.Y.W
A Semantic Assotiation Model for Corporate and Scientific-Statistical Databases
Information Sciences Vol 29, (1983) pp151 -199
18. Sundgren, B.
Statistical Metainformation and Metainformation Systems, R&D Reports Statistics ,
Sweden, 1991.
19. Van Meter, Karl M.(1994), "Sociological Methodology"
Bulletin de Methodologie Sociologique March 1994 N0 42 p 72 - 94
20. Wall W.D., Williams H.L Longitudinal Studies and the Social Sciences
London : Heinemann, 1970
21. Wongsaroje M and Vangdijck, E.
An Object Oriented Statistical Database: A Proposal for the National StatisticalOffice
Proc. Of the sixth Int. Working Conf. SSDBM Vol. (1992) pp178-194