

2. Μορφές οργάνωσης των όρων στην επιστήμη της πληροφόρησης

Διονύσιος Γιαννίμπας

ΠΕΡΙΛΗΨΗ

Στην πορεία της μέχρι σήμερα, η επιστήμη της πληροφόρησης χρησιμοποίησε και χρησιμοποιεί σύνολα όρων για τη θεματική περιγραφή των τεκμηρίων και τη διευκόλυνση της ανάκτησης της πληροφορίας. Τα λεξιλόγια αυτά χαρακτηρίζονται από διαφορετικούς τύπους και βαθμούς οργάνωσης, ως απόρροια της ανάπτυξης και χρήσης τους σε διαφορετικά πλαίσια απαιτήσεων.

Προσδιορίζονται τα είδη των λεξιλογίων, αποσαφηνίζονται σχετικές έννοιες και στοιχειοθετούνται κρίσιμες λεπτές διακρίσεις. Κωδικοποιούνται οι βασικές παράμετροι που καθορίζουν τη φυσιογνωμία των λεξιλογίων, παρουσιάζονται τυπικές περιπτώσεις χρήσης τους και περιγράφεται αντίστοιχα η συμπεριφορά τους. Αναδεικνύονται τα ισχυρά και αδύνατα σημεία τους, καθώς και οι μεταξύ τους ομοιότητες και διαφορές.

Τέλος, ιδιαίτερη βαρύτητα δίνεται αφενός στον έλεγχο του λεξιλογίου ως μια ειδική περίπτωση τυποποίησης ορολογίας, και αφετέρου στη δομή του λεξιλογίου ως τον πλέον καθοριστικό παράγοντα για την απόδοση του στο πλαίσιο ενός συστήματος πληροφόρησης.

Types of term organization in LIS

Dionyssios Giannibas

SUMMARY

Library and Information Science has used over the years term groups for the subject description of documents and the facilitation of information retrieval. These vocabularies are characterized by various kinds and degrees of organization, as a result of their development and use in different contexts.

The paper determines the different types of vocabularies, explains relevant notions and outlines critical, subtle distinctions. It encodes the basic parameters defining the character of the vocabularies and presents typical cases of their use, describing their behaviour. It also highlights their pros and cons, as well as the similarities and differences between them.

Especially, it discusses at length both vocabulary control as a specific case of terminology standardization and vocabulary structure as the factor contributing most decisively to their efficiency in the context of an information system.

1 Η ΓΛΩΣΣΑ, ΚΛΕΙΔΙ ΓΙΑ ΤΗΝ ΑΝΑΚΤΗΣΗ ΤΗΣ ΠΛΗΡΟΦΟΡΙΑΣ

Ο θαυμαστός σύγχρονος κόσμος της πληροφόρησης, με τον ασύλληπτο όγκο δεδομένων, τις αστρονομικές ταχύτητες μετάδοσης, τα δαιδαλώδη δίκτυα υπολογιστών, θυμίζει έναν πανίσχυρο εγκέφαλο, που αντιμετωπίζει όμως τεράστιες δυσκολίες στη διαχείριση των πληροφοριών.

Ο ανθρώπινος εγκέφαλος χρησιμοποιεί τη γλώσσα, αυτό το μοναδικό εργαλείο επικοινωνίας, για να βρίσκει τον δρόμο του στον κόσμο της σκέψης, σε έναν κόσμο που έχει σμιλευτεί με το ίδιο αυτό εργαλείο. Ο άνθρωπος ορίζει, ονοματίζει και συσχετίζει έννοιες διαρκώς, δημιουργώντας έναν νοηματικό «χάρτη» που προσδίδει μορφή και συνοχή στον κόσμο γύρω του. Οι συνάψεις των νευρώνων χαράσσουν τις διαδρομές και εξυπηρετούν τους συνειρμούς ανάμεσα σε πλήθος λέξεων και φράσεων με σύνθετο σημασιολογικό περιεχόμενο.

Όταν ο εγκέφαλος διαχειρίζεται πληροφορίες, προϊόντα εξωτερικών ερεθισμάτων, αξιοποιεί αυτόν τον «χάρτη» τόσο για την καταγραφή και αποθήκευσή τους όσο και για την επεξεργασία και ανάκτησή τους. Στη θαυμαστή αυτή οργανωτική ικανότητα του ανθρώπινου νου στηρίζεται όλο το γνωστικό οικοδόμημα· αντίληψη, μνήμη, επικοινωνία, μάθηση, αλλά και η ίδια η συνειδηση, δεν μπορούν να υπάρξουν χωρίς αυτήν. Δεν είναι λοιπόν τόσο παράδοξο που οι πανίσχυροι επεξεργαστές και τα ταχύτατα δίκτυα αδυνατούν έστω να συλλάβουν ό,τι εμείς κατανοούμε σχεδόν «δαισθητικά».

Καθώς σήμερα, περισσότερο από ποτέ, ζητούμενο είναι η συλλογική διαχείριση των αναρίθμητων πληροφοριακών πόρων που παράγει καθημερινά ο ανθρώπινος πολιτισμός, απαιτούνται «τεχνητοί» μηχανισμοί θεματικής περιγραφής και ανάκτησης που να μιμούνται στοιχεία του αντίστοιχου «φυσικού» μηχανισμού της νόησης, ώστε η σκέψη να ακολουθεί αβίαστα γνωστές διαδρομές σε κάθε αναζήτηση πληροφορίας. Προϋπόθεση είναι η συγκρότηση του ανάλογου νοηματικού «χάρτη», με τη βοήθεια γλωσσικών και άλλων συμβάσεων, σε μορφή τέτοια που να μπορεί να αξιοποιηθεί αφενός από τον συλλογικό ανθρώπινο νου ως κοινό σύστημα αναφοράς και αφετέρου από το σύστημα πληροφόρησης ως μέσο διαχείρισης της «επικοινωνίας» ανθρώπου και αποθηκευμένης πληροφορίας. Τα λεξιλόγια που η επιστήμη της πληροφόρησης –ή η βιβλιοθηκονομία, αν προτιμάτε– χρησιμοποίησε και χρησιμοποιεί για τη θεματική περιγραφή των τεκμηρίων και την υποστήριξη της πρόσβασης σε αυτά, είναι αυτός ακριβώς ο ενδιάμεσος κρίκος στην ανάκτηση της πληροφορίας.

2 Ο ΕΛΕΓΧΟΣ ΤΟΥ ΛΕΞΙΛΟΓΙΟΥ

Το πρώτο βήμα σε ένα τέτοιο λεξιλόγιο είναι η εφαρμογή των συμβάσεων εκείνων που διασφαλίζουν την τυποποίησή του κατά τρόπον ώστε να δηλώνεται ρητά η αντιστοιχία κάθε έννοιας με συγκεκριμένους όρους και κάθε όρου με συγκεκριμένη έννοια. Η τυποποίηση, ή αλλιώς έλεγχος, του λεξιλογίου έρχεται να αντιμετωπίσει το «πρόβλημα» της μεγάλης ποικιλομορφίας του ανθρώπινου λόγου, η οποία δυσχεραίνει την ταύτιση του όρου αναζήτησης με τον όρο ευρετηρίασης.

Τονίζεται ότι αυτό που τυποποιείται εδώ είναι η ισοδυναμία των όρων που αντιπροσωπεύουν την ίδια έννοια και όχι ο διαχωρισμός ενός από τους όρους αυτούς με παράλληλη απόρριψη των υπολοίπων. Κάτι τέτοιο θα είχε κατά την ανάκτηση το ακριβώς αντίθετο αποτέλεσμα από το επιθυμητό, αφού θα αφαιρούσε τη δυνατότητα για πολλαπλούς όρους εισόδου στο λεξιλόγιο και αντίστοιχα σημεία πρόσβασης στην πληροφορία. Ο χαρακτηρισμός του «προτιμώμενου» όρου που χρησιμοποιείται σε ορισμένα ελεγχόμενα λεξιλόγια είναι περισσότερο μια υπόδειξη ευρετηρίασης και δεν έχει ουσιαστικές προεκτάσεις.

Χωρίς τον έλεγχο του λεξιλογίου, η αναζήτηση βασίζεται σε μη καθιερωμένες λέξεις-κλειδιά ή σε όρους ελεύθερου κειμένου. Ειδικά ο ταχύτατος εντοπισμός των όρων αναζήτησης σε μηχαναγνώσιμο κείμενο, εφικτός χάρη στην πρόοδο της πληροφορικής, δημιούργησε ελπίδες για επαρκή θεματική πρόσβαση δίχως την ανάγκη μιας συστηματικής και απαιτητικής εργασίας υποδομής όπως η ευρετηρίαση. Δεν άργησε, ωστόσο, να φανεί ότι η συγκεκριμένη τεχνική έχει περιορισμένες δυνατότητες ως προς την ποιότητα των αποτελεσμάτων. Είναι βέβαια γεγονός ότι τη χρήση της αναζήτησης ελεύθερου κειμένου εξακολουθούν να ευνοούν διάφοροι παράγοντες, με κυριότερο τον ρυθμό αύξησης του όγκου του περιεχομένου. Ακόμη και έτσι όμως, η ποιότητα της ανάκτησης βελτιώνεται σημαντικά όταν γίνεται χρήση ελεγχόμενου –πόσω μάλλον δομημένου, όπως θα δούμε στη συνέχεια– λεξιλογίου για την επιλογή του όρου αναζήτησης. Είναι χαρακτηριστικό ότι συχνά εφαρμόζεται από τα συστήματα η τεχνική των *δακτυλίων συνωνύμων* (synonym rings) για τον εντοπισμό των διαφόρων όρων που αντιπροσωπεύουν την ίδια έννοια, χωρίς μάλιστα να προαπαιτείται σχετική επιλογή από τον χρήστη.

Την απλούστερη μορφή ελεγχόμενου λεξιλογίου (ή αλλιώς γλώσσας ευρετηρίασης) συναντάμε στον *κατάλογο καθιερωμένων όρων*, ο οποίος, στην κλασική έντυπη μορφή του, περιλαμβάνει παραπομπές του τύπου «βλέπε» (π.χ. «πουλιά, βλέπε πτηνά»). Εντούτοις, τα οφέλη από τη

χρήση ενός τέτοιου λεξιλογίου σε ένα σύστημα πληροφόρησης εξαντλούνται εν πολλοίς στην άρση της πολυσημίας και της συνωνυμίας που χαρακτηρίζουν τη φυσική γλώσσα. Για την καθοδήγηση του χρήστη στην επιλογή του κατάλληλου όρου αναζήτησης, μέσα από τη παροχή ενός φάσματος εναλλακτικών θεωρήσεων του θέματος που τον ενδιαφέρει, προϋπόθεση είναι η διασύνδεση των εννοιών με επαρκώς και σαφώς καθορισμένες σχέσεις. Με άλλα λόγια, το λεξιλόγιο είναι απαραίτητο να διαθέτει δομή.

3 Η ΔΟΜΗ ΤΟΥ ΛΕΞΙΛΟΓΙΟΥ

Τα λεξιλόγια με την απλούστερη δομή είναι οι διαδεδομένες στο Διαδίκτυο «ταξινομίες», όρος πρόσφατος γύρω από τον οποίο ακόμη υπάρχει κάποια σύγχυση, καθώς απουσιάζει από την βασική βιβλιογραφία του τομέα. Οι ταξινομίες, που αναπτύχθηκαν παράλληλα με τους «εμπορικούς» ιστότοπους και τις διαδικτυακές πύλες, επιβλήθηκαν στην ειδική ορολογία της πληροφόρησης έχοντας κάνει πρώτα την εμφάνισή τους στον χώρο της πληροφορικής τεχνολογίας και της ανάπτυξης συστημάτων. Ένα πολύ γνωστό σε όλους παράδειγμα είναι το λεξιλόγιο που χρησιμοποιεί το περίφημο «Yahoo!».

Οι ταξινομίες έχουν ιεραρχική δομή: οι σχέσεις μεταξύ των εννοιών αναπτύσσονται κάθετα και χρησιμοποιούνται κυρίως για την ομαδοποίηση των όρων σε κατηγορίες (π.χ. ο όρος «πτηνά» έχει ως πλατύτερο τον όρο «ζώα» και ως ειδικότερο τον όρο «περιστέρια»). Η συγκεκριμένη δομή καθιστά εφικτή, εκτός των άλλων, τη διεύρυνση ή εξειδίκευση της αναζήτησης, αλλά και τη λεγόμενη επέκταση ερωτήματος που συνήθως αναφέρεται στην αξιοποίηση, κατά την αναζήτηση, και των όρων εκείνων που είναι ειδικότεροι του όρου που επέλεξε ο χρήστης (αν και η τελευταία αυτή δυνατότητα δεν θεωρείται συνήθως απαραίτητη στα «περιβάλλοντα» πληροφόρησης που χρησιμοποιούν ταξινομίες).

Η δυνατότητα πρόσβασης που παρέχεται στον χρήστη περιορίζεται κατά κανόνα στην πλοήγηση από το γενικότερο στο ειδικότερο, ιδίως όταν η κάθετη ανάπτυξη της ταξινομίας δεν έχει μεγάλο βάθος. Αυτός είναι και ο λόγος που έλεγχος του λεξιλογίου στις ταξινομίες δεν σημαίνει συνήθως τίποτε παραπάνω από τη χρήση συγκεκριμένου όρου για κάθε θέμα, καθώς απουσιάζουν οι σχέσεις ισοδυναμίας και οι αντίστοιχοι όροι εισόδου στο λεξιλόγιο.

Ανάλογη δομή έχουν και τα *ταξινομικά συστήματα*, αλλά, επειδή αυτά αναπτύχθηκαν για να εξυπηρετήσουν τις ανάγκες της βιβλιογραφικής ταξινόμησης, αφενός χρησιμοποιούν για τη θεματική περιγραφή σύμβολα (αριθμούς ή γράμματα ή και τα δύο), τα οποία βεβαίως συνδέονται

με συγκεκριμένα θέματα μέσω πινάκων και ευρετηρίων, και αφετέρου εφαρμόζουν ως επί το πλείστον διάφορες τεχνικές συνδυασμού των θεμάτων. Τα πλέον διαδεδομένα ταξινομικά συστήματα, που χρησιμοποιούνται διεθνώς ως πρότυπα εργαλεία και καλύπτουν θεματικά όλο το φάσμα του επιστητού, είναι το DDC (Dewey Decimal Classification), το UDC (Universal Decimal Classification) και το LCC (Library of Congress Classification). Τα συστήματα αυτά συντηρούνται και εμπλουτίζονται από μεγάλους πληροφοριακούς οργανισμούς, ενώ η «γλώσσα» των συμβόλων βοηθά την εφαρμογή τους σε πολύγλωσσο περιβάλλον. Επειδή ακριβώς η «γλώσσα» τους δεν βασίζεται στη φυσική γλώσσα του ανθρώπου, δεν θα αναπτύξουμε εδώ περισσότερο τα ιδιαίτερα χαρακτηριστικά τους.

3.1 ΘΗΣΑΥΡΟΙ ΚΑΙ ΘΕΜΑΤΙΚΕΣ ΕΠΙΚΕΦΑΛΙΔΕΣ

Το στοιχείο όμως που προσδίδει στη δομή του λεξιλογίου τη μορφή δικτύου σχέσεων και παρέχει στον χρήστη τη δυνατότητα να ξεφύγει από την ομαδοποίηση των όρων και να κινηθεί σε εναλλακτικές διαδρομές είναι η «οριζόντια» διασύνδεση των εννοιών με μη ιεραρχικές σχέσεις (π.χ. ο όρος «πτηνά» έχει ως σχετικό τον όρο «ορνιθολογία»· ακολουθώντας τη διαδρομή αυτή, ο χρήστης περνά από την ομάδα ή κατηγορία των «έμβιων όντων» σε εκείνη των «επιστημών»). Αυτές οι σχέσεις είναι που διασυνδέουν τις έννοιες με τρόπο που παρακάμπτει τις λογικές κατηγοριοποιήσεις και αξιοποιεί λιγότερο προφανείς και, γι' αυτό, πολύτιμες «συνάφειες». Το όχημα για την αναζήτηση της πληροφορίας δεν είναι πλέον ένας ανελκυστήρας μεταξύ ορόφων ενός κτηρίου, αλλά ένας υπόγειος σιδηρόδρομος που μας μεταφέρει γρήγορα και αποτελεσματικά σε άλλο σημείο της ίδιας περιοχής ή και στην άλλη άκρη της πόλης.

Τέτοια δομή διαθέτουν οι *κατάλογοι θεματικών επικεφαλίδων* και οι *θησαυροί*, δύο τύποι γλώσσας ευρετηρίασης που χρησιμοποιούνται ευρέως σε συστηματικά τεκμηριωμένες συλλογές. Η μεταξύ τους διάκριση δυσχεραίνεται από αυτήν ακριβώς την δομική συγγενεία τους, η οποία εντούτοις είναι σε μεγάλο βαθμό επιφανειακή.

Στην πραγματικότητα, μεγάλες είναι οι διαφορές τους τόσο στο θεωρητικό υπόβαθρο όσο και στην πρακτική εφαρμογή τους, και αυτό οφείλεται κατ' αρχήν στο γεγονός ότι εκπροσωπούν δύο διαφορετικές προσεγγίσεις στο δίπτυχο ευρετηρίαση-αναζήτηση.

Πιο συγκεκριμένα:

- Οι θεματικές επικεφαλίδες

- δομούνται με βάση τη θεματική περιγραφή συγκεκριμένων τεκμηρίων,
- συχνά αποτελούνται από συνδυασμούς όρων που δηλώνουν μια ευκαιριακή, ή αλλιώς συντακτική, σχέση μεταξύ των εννοιών, η οποία δεν έχει νόημα παρά στο πλαίσιο της ευρετηρίασης του εκάστοτε τεκμηρίου,
- χρησιμοποιούνται στην προσυνδυασμένη ευρετηρίαση.

Ο τυχαίος τρόπος με τον οποίο προστίθενται νέες επικεφαλίδες στον κατάλογο, βάσει των αναγκών της ευρετηρίασης, αφήνει μεγάλα περιθώρια για θεματικά «κενά» ή «επικαλύψεις».

➤ Στους θησαυρούς, οι όροι

- εκφράζουν αυτόνομες έννοιες που από κοινού καλύπτουν θεματικά συγκεκριμένο εύρος και βάθος,
- διασυνδέονται με σταθερές σημασιολογικές σχέσεις που απορρέουν από τις ίδιες τις έννοιες και δεν εξαρτώνται από κανένα συγκεκριμένο τεκμήριο,
- χρησιμοποιούνται στην υστεροσυνδυασμένη ευρετηρίαση, όπου σε κάθε τεκμήριο αποδίδεται ένας αριθμός ξεχωριστών θεματικών όρων, και μόνο στη φάση της αναζήτησης μπορεί να γίνει συνδυασμός των όρων αυτών.

Ο συγκροτημένος τρόπος με τον οποίο δομούνται εξ αρχής οι θησαυροί διευκολύνει την πληρότητα και τη σαφήνεια εννοιών και σχέσεων στο θεματικό τους πλαίσιο.

Η φιλοσοφία των θησαυρών ενισχύει την «οικονομία» της γλώσσας ευρετηρίασης. Σχηματικά, θα λέγαμε ότι η δομική μονάδα στους θησαυρούς είναι η «λέξη», ενώ στους καταλόγους θεματικών επικεφαλίδων είναι η «φράση». Οι επικεφαλίδες αυξάνονται συνεχώς και με τυχαίο τρόπο προκειμένου να περιγραφούν νέα τεκμήρια· αντίθετα, αλλαγές στο περιεχόμενο θησαυρών γίνονται μόνο στον βαθμό που εμφανίζονται νέες έννοιες ή επέρχονται τροποποιήσεις στις ήδη υπάρχουσες, πάντοτε στο πλαίσιο της θεματικής τους κάλυψης.

Η φαινομενική ομοιότητα των δύο αυτών γλωσσών ευρετηρίασης ως προς τη δομή του πλέγματος των σχέσεων οφείλεται περισσότερο στο γεγονός ότι ορισμένοι κατάλογοι θεματικών επικεφαλίδων (όπως της Βιβλιοθήκης του Κογκρέσου, από τον οποίο προέρχονται και τα παραδείγματα που ακολουθούν) υιοθέτησαν τα τελευταία χρόνια τη σημειογραφία που αναπτύχθηκε για την περιγραφή της δομής των θησαυρών. Παλαιότερα χρησιμοποιούσαν (και

πολλοί χρησιμοποιούν ακόμη και σήμερα) τις παραπομπές του τύπου «βλέπε επίσης» για να αποτυπώσουν αδιακρίτως τόσο τις κάθετες όσο και τις οριζόντιες σχέσεις μεταξύ των όρων. Η αλλαγή αποφασίστηκε ακριβώς επειδή διαπιστώθηκε η αδυναμία των συγκεκριμένων παραπομπών να αποδώσουν αυτήν τη θεμελιώδη δομική διάκριση. Ωστόσο, σε μεγάλο βαθμό το εγχείρημα απέτυχε: το σύστημα παραπομπών μεταξύ των επικεφαλίδων, παρά τη χρήση της σημειογραφίας των θησαυρών ('BT-NT', 'RT'), εξακολουθεί να μη διαχωρίζει σαφώς τις κατηγορίες σχέσεων και να μην ακολουθεί τις κατευθυντήριες οδηγίες για τους θησαυρούς [1].

Ο λόγος της αποτυχίας είναι απλός: από τη στιγμή που κάθε επικεφαλίδα δύναται να απαρτίζεται από περισσότερες της μίας έννοιες, η εφαρμογή της δομής των θησαυρών είναι ουσιαστικά ανέφικτη. Η επικεφαλίδα «Automobiles – Pollution control devices» εμφανίζεται ως ειδικότερος όρος ('NT') της επικεφαλίδας «Air – Pollution»... Είναι φανερό ότι, με τους συνδυασμούς των εννοιών, ο καθορισμός μιας ιεραρχικής σχέσης δεν μπορεί παρά να είναι αυθαίρετος.

Αλλά ακόμη και στις σχέσεις μεταξύ απλών εννοιών, επικρατεί ανάλογη σύγχυση: η επικεφαλίδα «Automobiles» εμφανίζεται ως 'NT' της επικεφαλίδας «Transportation, Automotive», η οποία με τη σειρά της δεν εμφανίζεται ως 'NT' της επικεφαλίδας «Transportation». Ή, πάλι, η επικεφαλίδα «Marriage» εμφανίζεται ως 'NT' της επικεφαλίδας «Love»! Ο θεμελιώδης κανόνας που θέλει τις έννοιες που συνδέονται με ιεραρχική σχέση να ανήκουν στην ίδια γενική κατηγορία, στον ίδιο «βασικό τύπο εννοιών» (π.χ. «πράγματα», «ενέργειες» ή «ιδιότητες», βλ. [1] § 8.3.2), αγνοείται συστηματικά. Το αποτέλεσμα είναι ότι συνολικά η δομή του καταλόγου θεματικών επικεφαλίδων χαρακτηρίζεται από έλλειψη συνοχής και τελείως υποκειμενική οργάνωση.

Είναι σημαντικό πάντως να πούμε ότι οι θεματικές επικεφαλίδες πήραν τη μορφή τους υπηρετώντας τις ανάγκες θεματικής πρόσβασης στις κλασικές συλλογές των βιβλιοθηκών και καλύπτοντας αντίστοιχα με ανοικτό τρόπο όλο το φάσμα της γνώσης, ενώ οι θησαυροί εμφανίστηκαν από την αρχή ως εργαλείο ανάκτησης της πληροφορίας στο πλαίσιο προηγμένων συστημάτων πληροφόρησης και οι περισσότεροι από αυτούς αποτελούν το λεξιλόγιο πρόσβασης σε μεγάλες βάσεις δεδομένων επιστημονικού και τεχνικού περιεχομένου. Και είναι χαρακτηριστικό ότι η μετεξέλιξη των λεξιλογίων μέσα από μια διαρκή διερεύνηση των εφαρμογών και της απόδοσής τους είναι σήμερα περισσότερο από ποτέ στο προσκήνιο της επιστήμης της πληροφόρησης.

3.2 ΟΝΤΟΛΟΓΙΕΣ, Η ΕΠΟΜΕΝΗ ΜΕΡΑ

Κεντρική θέση στη σκηνή αυτή έχουν πάρει τα τελευταία χρόνια οι οντολογίες. Πρόκειται για έναν ακόμη «ξενόφερτο» στην οικογένεια των λεξιλογίων που μας απασχολούν. Οι οντολογίες προέρχονται από τον χώρο της τεχνητής νοημοσύνης και της αναπαράστασης της γνώσης: βασικός σκοπός τους είναι να δώσουν στις μηχανές και τα συστήματα τη δυνατότητα να χειριστούν αποτελεσματικά την πληροφορία, και όχι να προσφέρουν στον χρήστη το λεξιλόγιο που θα τον καθοδηγήσει στην επιλογή του κατάλληλου όρου αναζήτησης. Αντλούν την ισχύ τους από το γεγονός ότι πρόκειται για ανοικτές γλώσσες, υπό την έννοια ότι η δομή τους δεν περιορίζεται σε συγκεκριμένες σχέσεις και κατηγορίες εννοιών αλλά περιλαμβάνει «ιδιότητες», που αφήνουν κάθε ευχέρεια για επέκταση και εμβάθυνση. Γνωστό παράδειγμα οντολογίας είναι το Wordnet.

Θα μπορούσε να πει κανείς ότι είναι ένα είδος εξελιγμένου θησαυρού, που έχει την αφετηρία του στη φιλόδοξη προσπάθεια περιγραφής ενός κομματιού του κόσμου μας με τρόπο τέτοιο ώστε η συγκεκριμένη γνώση να δύναται να «επικοινωνηθεί» από υπολογιστή σε υπολογιστή. Σημαντικό ρόλο παίζουν εδώ η σαφήνεια και η λεπτομέρεια με την οποία ορίζονται και αποτυπώνονται οι έννοιες και οι κανόνες που τις πλαισιώνουν. Στην κατεύθυνση αυτή, χαρακτηριστική είναι η συστηματική εξειδίκευση των πλέον «χαλαρών» σχέσεων των θησαυρών, των σχέσεων συσχέτισης.

Οι οντολογίες μπορεί να έχουν ένα ευρύ φάσμα εφαρμογών, αλλά στο πλαίσιο της επιστήμης της πληροφόρησης επιστρατεύονται και αυτές στη μάχη της ανάκτησης της πληροφορίας. Και αν τα χαρακτηριστικά των λεξιλογίων διαφέρουν, με συγκεκριμένες πρακτικές συνέπειες στη χρήση και απόδοση των συστημάτων, η κοινή τους αποστολή παραμένει το στοιχείο που τα φέρνει κοντά.

Αυτός είναι και ο λόγος που πρόσφατα έχει αναπτυχθεί και ένα πεδίο έρευνας που αφορά τους τρόπους με τους οποίους μπορεί να διασυνδεθεί ή να αντιστοιχιστεί το περιεχόμενό τους. Υπάρχει μια μακρά παράδοση στον τομέα, η οποία έχει αποφέρει πολλά μέχρι σήμερα, και ο πλούτος αυτός οφείλει να αξιοποιείται με τον καλύτερο δυνατό τρόπο.

Μόνον έτσι θα μπορούν τα συστήματα πληροφόρησης να «τροφοδοτούν» τον καθένα από εμάς με δεδομένα που δεν θα αποτελούν απλά «κομμάτια» πληροφορίας αλλά θα δίνουν ουσιαστικό νόημα στις όλο και πιο δημοφιλείς αναφορές σε «συστήματα διαχείρισης περιεχομένου», ή

γενικότερα στη λεγόμενη «διαχείριση της γνώσης». Είναι ένα στοίχημα που θα κερδηθεί όταν εμπεδωθεί η πεποίθηση ότι το μέλλον της σχετικής τεχνολογίας δεν εξαρτάται τόσο από την πρόοδο της πληροφορικής όσο από την επανάσταση που θα φέρει η προσαρμογή της ανθρώπινης γλώσσας σε εργαλείο επικοινωνίας για τις μηχανές.

Βιβλιογραφία

1. Ελληνικό Πρότυπο ΕΛΟΤ 1321, *Τεκμηρίωση – Κατευθυντήριες οδηγίες για τη συγκρότηση και ανάπτυξη μονόγλωσσων θησαυρών*
2. *Θησαυρός Ελληνικών Όρων – Μεθοδολογία Ανάπτυξης*, Εθνικό Κέντρο Τεκμηρίωσης, 2005
3. Γιαννίμπας Δ, *Το στοιχείο της δομής στις γλώσσες ευρετηρίασης*, Ενημερωτικό Δελτίο για τις Δημόσιες Βιβλιοθήκες, τεύχος 3, 2005

Διονύσιος Γιαννίμπας

Μεταφραστής

Συντονιστής σε θέματα ορολογίας – ΕΚΤ

Διεύθυνση αλληλογραφίας: Στίλπωνος 15 – 116 36 Αθήνα

τηλ.: 210-7273946, ηλ. ταχ.: giannibas@ekt.gr