

Διαχείριση Λεξιλογίων και Θησαυρών

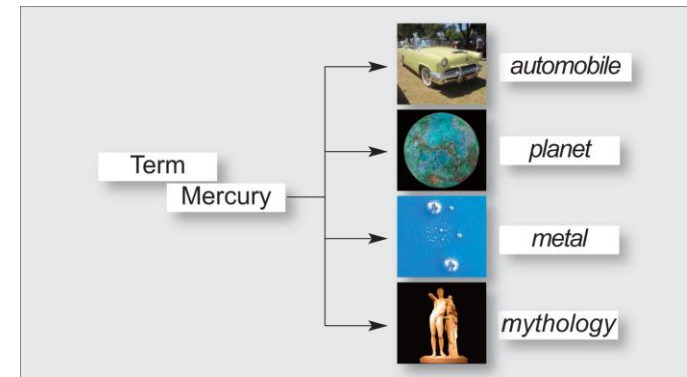
Λήδα Χαραμή, Χρυσούλα Μπεκιάρη
14 Σεπτεμβρίου 2021

Εισαγωγή

- **Έννοια** (concept): η μονάδα γνώσης που δημιουργείται από έναν μοναδικό συνδυασμό χαρακτηριστικών - ιδιοτήτων.
- **Όρος** (term): η λεκτική μορφή που χρησιμοποιούμε για να δηλώσουμε αυτή την έννοια
- Οι έννοιες που φιλοξενεί ένας θησαυρός όρων προσδιορίζουν σύνολα αντικειμένων του πραγματικού κόσμου και οι όροι χρησιμοποιούνται για να αναφερθούμε στις έννοιες.

Ορισμός - Χαρακτηριστικά Θησαυρών

- Ανήκουν στα Συστήματα Οργάνωσης Γνώσης - Knowledge Organization Systems (KOS)
- Είναι **λεξιλόγια** που **οργανώνουν** τις **έννοιες και όρους** ενός επιστημονικού πεδίου με βάση **συγκεκριμένους κανόνες** με σκοπό την ελαχιστοποίηση της σημασιολογικής ασάφειας που προκύπτει στη φυσικής γλώσσα.
 - Η ίδια λέξη μπορεί να δηλώνει περισσότερες από μία έννοιες



- Μια έννοια μπορεί να αναπαρασταθεί με δύο ή περισσότερες λέξεις που έχουν την ίδια ή παρόμοια σημασία

Words	Concept
Artificial consciousness Biocomputers Electronic brains Mechanical brains Synthetic consciousness	Conscious automata

Χαρακτηριστικά Θησαυρών

- Παρέχουν όρους με αυστηρά **οργανωμένο τρόπο** που αποδεκτούς από τους ειδικούς τομείς, καθένας από τους οποίους αντιπροσωπεύει μόνο μία έννοια.
- Ο όρος που επιλέγεται για να αντιπροσωπεύσει μία έννοια ονομάζεται **προτιμώμενος όρος** ή περιγραφέας.
- **Περιορίζουν το νόημα** μιας έννοιας στον τομέα ενός συγκεκριμένου επιστημονικού πεδίου
- Παρέχουν **έλεγχο συνωνύμων**
- Δημιουργούν **ιεραρχίες** όρων
- **Διασυνδέουν έννοιες** που σχετίζονται νοηματικά

Χρήση (1)

- Μπορούν να χρησιμοποιηθούν κατά τη διαδικασία της **τεκμηρίωσης**
 - Σαν τιμές των πεδίων σε ένα σύστημα βάσης δεδομένων όπου
 - Κάθε αντικείμενο/ τεκμήριο μπορεί να ευρετηριαστεί από τους προτιμώμενους όρους

- Αποτελούν ένα **εργαλείο αναζήτησης** και **ανάκτησης** της πληροφορίας και χρησιμοποιούνται για να **βελτιώσουν τις επερωτήσεις** σε ένα σύστημα βάσεως δεδομένων
 - Οδηγούν του χρήστες να επιλέγουν τον ίδιο όρο για να εκφράσουν το ίδιο αντικείμενο/ τεκμήριο
 - Παρέχουν τη δυνατότητα επέκτασης της αναζήτησης σε προτιμώμενους όρους (που μπορεί να είναι ευρύτεροι, στενότεροι ή συναφείς του αρχικού όρου αναζήτησης)
 - Επιτυγχάνουν ανάκτηση της πληροφορίας τη μέγιστη σαφήνεια και πληρότητα

Χρήση (2)

- Εξασφαλίζει την ομοιομορφία και συνέπεια στην ταξινόμηση και ανάκτηση του περιεχομένου των συλλογών
- Διευκολύνει την πρόσβαση, συμβατότητα και σύγκριση ετερογενών συστημάτων ταξινόμησης και διαχείρισης ψηφιακών συλλογών
- Τυποποιεί τις εκφράσεις ταξινόμησης και καταλογογράφησης ώστε να δημιουργηθεί μία συστηματική γλώσσα επικοινωνίας.

Οι όροι του Θησαυρού

- Συγκεκριμένες οντότητες - Αποδίδονται σε πληθυντικό αριθμό π.χ. αγγεία, κτίρια, βιβλία, όπλα
- Αφηρημένες έννοιες- Αποδίδονται σε ενικό αριθμό π.χ. δίκαιο, χημεία, γλωσσολογία, μετανάστευση, φιλοσοφία
- Οι όροι ευρετηρίασης συνηθέστερα αποδίδονται με
 - Ουσιαστικά – είναι οι συνηθέστεροι τύποι όρων στο θησαυρό π.χ. Διδασκαλία [όχι Διδάσκω]
 - Φράση με ουσιαστικό → σύνθετοι όροι π.χ. Θαλάσσια Πτηνά
- Θα πρέπει να αποφεύγονται
 - Επίθετα και επιρρήματα
 - Ρήματα, απαρέμφατα και μετοχές
- Εξαιρέσεις: Αν έχει αποκτήσει μία ειδική έννοια στο πλαίσιο ενός τομέα π.χ. πολύ υψηλή συχνότητα

Οι όροι του Θησαυρού (3)

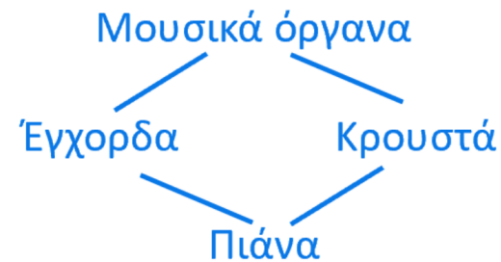
- Τα Ομοιόγραφα και τα πολύσημα - δηλαδή λέξεις με την ίδια ορθογραφία αλλά με διαφορετική σημασία.
 - Θα πρέπει να πρέπει να συνοδεύονται από μία προσδιοριστική λέξη ή φράση που θα τα διαφοροποιεί, συνήθως σε παρένθεση ή με πλάγια γράμματα
 - π.χ. Εγχειρίδια (όπλα), Εγχειρίδια (βιβλία)
- Κάθε όρος που εισάγεται στο θησαυρό θα πρέπει να συνοδεύεται από ένα **διευκρινιστικό σημείωμα**, ένα κείμενο που αποτελεί τη λεκτική περιγραφή και χρήση της έννοιας.
- Καταγραφή των βιβλιογραφικών αναφορών που έχουν χρησιμοποιηθεί για τη συγκέντρωση και τον προσδιορισμό των όρων και των εννοιών προς ταξινόμηση.

Σχέσεις σε ένα θησαυρό (1)

Σχέση Ιεραρχίας (IsA)

- Είναι η σχέση ανάμεσα σε δύο έννοιες, όταν το πεδίο εφαρμογής της μιας έννοιας περιέχεται εξ ολοκλήρου στο πεδίο της άλλης
 - OK: Όρος κορυφής,
 - ΠΟ: Πλατύτερος όρος,
 - ΕΟ: Ειδικότερος όρος
- π.χ. OK: Κινητά Αντικείμενα, ΠΟ: Αγγεία, ΕΟ: Αμφορείς
- Πολυιεραρχικές σχέσεις έχουμε στις περιπτώσεις που μία έννοια ανήκει σε πάνω από μία κατηγορίες ταυτόχρονα

π.χ.



- Ετικέτες Δεσμού: είναι φράσεις που παρεμβάλλονται στην ιεραρχία για να δείξουν ποιο χαρακτηριστικό χρησιμοποιήθηκε για να διαιρέσει την τάξη στα είδη της. Αποδίδονται μέσα σε < >

π.χ <ως προς τη χρήση>, <ως προς τη μορφή>

Σχέσεις σε ένα θησαυρό (2)

Σχέση Ισοδυναμίας

- είναι η σχέση μεταξύ δύο ή περισσότερων **όρων** που θεωρούνται ότι αναφέρονται στην ίδια έννοια.
 - Σε ένα θησαυρό επιλέγεται ένας όρος σαν αντιπροσωπευτικός της έννοιας ενώ όλοι οι άλλοι, δηλώνονται ως συνώνυμοι ή αλλιώς μη προτιμώμενοι.
 - XP: Χρησιμοποίησε,
 - XA: Χρησιμοποίησε αντί
 - π.χ. Καμάρες
 - XP: Αψίδες

Σχέση Συσχέτισης

- είναι σχέση ανάμεσα σε όρους που ανήκουν σε διαφορετικούς ιεραρχικούς κλάδους, αλλά που σχετίζονται νοηματικά έτσι ώστε ο δεσμός να πρέπει να γίνει φανερός στον θησαυρό
 - ΣΟ: Σχετικός όρος
 - π.χ. Διδασκαλία ΣΟ: Βιβλία

Δημιουργία Θησαυρών – Βήμα (1)

Κατά τη συγκρότηση ενός θησαυρού τα **γενικά χαρακτηριστικά του καθορίζονται από το πεδίο εφαρμογής του θησαυρού**

Επομένως:

- Θα πρέπει να καθοριστεί το **επιστημονικό πεδίο**
- Να συγκεντρωθούν οι έννοιες και να **οριστούν** ώστε να αποδίδουν το νόημα με το οποίο γίνονται αντιληπτές στο συγκεκριμένο πεδίο
- Να οριστούν οι έννοιες με βάση τις ουσιώδεις τους ιδιότητες.
- Να βρεθούν οι όροι που είναι νοηματικά ισοδύναμοι, δηλαδή να αναγνωριστούν οι νοηματικά συνώνυμοι ή οίωνοί συνώνυμοι όροι
- Από αυτούς να επιλεγεί ο προτιμώμενος όρος που θα αντιπροσωπεύει μια συγκεκριμένη έννοια και να οριστούν οι υπόλοιποι ως μη προτιμώμενοι

Δημιουργία Θησαυρών – Βήμα (2)

- Να οργανωθούν βάσει *θεμελιωδών κατηγοριών (Facets – Θέματα)* που εκφράζουν ουσιώδεις «όψεις» όλων όσων αντιλαμβανόμαστε.
- Οι κατηγορίες είναι γενικές έννοιες, αποσπασμένες από το συγκεκριμένο περιεχόμενο κάθε επιστήμης (πχ. υλικό αντικείμενο), οι οποίες μας βοηθούν σε μια νέα ταξινόμηση των εννοιών των ειδικών επιστημών.
- Κάθε κατηγορία περιλαμβάνει μία ενιαία (ομογενή) κλάση πραγμάτων με κοινά χαρακτηριστικά π.χ.
 - Υλικά αντικείμενα
 - Γεωπολιτικές μονάδες
- Οι κατηγορίες εμφανίζονται μέσω της ανάλυσης των συνθετότερων εννοιών των ειδικών επιστημών σε απλούστερες και πιο στοιχειώδεις έννοιες.

Δημιουργία Θησαυρών – Βήμα (3)

- Να δημιουργηθούν **ιεραρχίες** μεταξύ των εννοιών - να ορίσουμε τις πλατύτερες και ειδικότερες έννοιες
- Να καθοριστούν οι **συσχετισμοί** μίας έννοια με άλλες έννοιες, όπου αυτό είναι απαραίτητο.

Πρότυπα Θησαυρών

1. ISO 25964 'Thesauri and interoperability with other vocabularies' 2013.
2. ANSI/NISO Z39.19 Guidelines for the construction, format, and management of monolingual controlled vocabularies. 1993 (revised 2005, 2010).
3. ΕΛΟΤ 1321 Τεκμηρίωση - Κατευθυντήριες οδηγίες για τη συγκρότηση και ανάπτυξη μονόγλωσσων θησαυρών, 1993
4. BS 8723-1:2005, BS 8723-2:2005, BS 8723-3:2007, BS 8723-4:2007: Structured Vocabularies for Information Retrieval.
5. ISO 2788 'Documentation - Guidelines for the Establishment and Development of Monolingual Thesauri' was published in 1074. It was revised in 1986 (superseded by ISO 25964).
6. ISO 5964 'Documentation - Guidelines for the establishment and development of multilingual thesauri' was released for multilingual thesauri In 1985 (superseded by ISO 25964).
7. DIN 1463 Guidelines for the establishment and development of monolingual thesauri. 1972 (followed by later editions)

Ευχαριστώ για την προσοχή σας

Λήδα Χαραμή
lida@ics.forth.gr