



Πληροφοριακά Συστήματα Ανάκτησης Ιατρικής Πληροφορίας

Μ. Μποτσιβάλη, Β. Σπυρόπουλος

Τμήμα Τεχνολογίας Ιατρικών Οργάνων
Τ.Ε.Ι. Αθήνας

Email: botsivali.maria@gmail.com **URL:** www.bmtl.bme.teiath.gr



Εισαγωγή – Η Πληροφορία

- Τι είναι η *Πληροφορία*;
 - ◆ *Συνδυασμός δεδομένων και γνώσεων.*

- Τι είναι τα *Δεδομένα* ;
 - ◆ *Παρατηρήσεις ή μετρήσεις που χαρακτηρίζουν την τιμή μίας μεταβλητής ενός συγκεκριμένου αντικειμένου (για παράδειγμα, ενός ασθενούς) μια συγκεκριμένη χρονική στιγμή.*

- Τι είναι η *Γνώση* ;
 - ◆ *Προϊόν της ερμηνείας των δεδομένων.*



Η αξία της πληροφορίας

- Η πληροφορία θεωρείται *περιουσία* και *πάγιο* για το άτομο ή τον οργανισμό που την κατέχει και βέβαια αναλώνεται πολύτιμος χρόνος και χρήμα για την συγκέντρωσή της.
- Σε μία εταιρία η κατοχή πληροφοριών δίνει μόνον ανταγωνιστικό πλεονέκτημα, ενώ στην ιατρική η γνώση σώζει ζωές!
- Οι ελλειπείς πληροφορίες οδηγούν σε λάθος αποφάσεις:
 - ◆ *Στην ιατρική οι λάθος αποφάσεις σημαίνουν σοβαρές συνέπειες στη ζωή ενός ασθενή.*
- Στην Υγεία η πληρότητα και η διαθεσιμότητα της πληροφορίας είναι ζωτικής σημασίας. Ημιτελής πληροφορία μπορεί να οδηγήσει σε κακή διάγνωση και θεραπεία, σε σπατάλη χρημάτων και πόρων και σε καταστάσεις οι οποίες απειλούν τη ζωή του ασθενούς.



Η ζούγκλα της πληροφορίας

- Το μέγεθος της ιατρικής και βιοϊατρικής πληροφορίας που παράγεται καθημερινά στο σύγχρονο κόσμο είναι τεράστιο και δύσκολο ακόμα και να εκφραστεί με τρόπο απλό.
- Οι πηγές της ιατρικής - βιοϊατρικής πληροφορίας είναι πολλές: οι φορείς υγείας, η εκπαίδευση, οι ερευνητικοί φορείς, οι θεσμικοί φορείς, κλπ.
- Η ιατρική και βιοϊατρική πληροφορία που υπάρχει σήμερα διαθέσιμη σε ψηφιακή μορφή, έχει ήδη χαρακτηριστεί «ζούγκλα πληροφορίας» (information jungle).



Πηγές Ιατρικής Πληροφορίας (Γνώσης)

- Η ιατρική γνώση είναι ανακτήσιμη από:
 - ◆ *την ιατρική βιβλιογραφία (τεκμηριωμένη γνώση) και*
 - ◆ *τους διάφορους εμπειρογνώμονες σε μια συγκεκριμένη περιοχή (κλινική εμπειρία)*

- Τα κύρια χαρακτηριστικά και των δύο είναι:
 - ◆ *Η αφηγηματική μορφή.*
 - ◆ *Η ιδιαίτερη δυσκολία αξιοποίησής τους.*

- Κανένα από τα δύο δεν εκφράζεται σε μορφή προσιτή ηλεκτρονικά και ερμηνεύσιμη από υπολογιστή



Ποιος είναι ο στόχος

- Ο στόχος σήμερα είναι:
 - ◆ *Η μετάβαση σε μία νέα κατάσταση, όπου η πληροφορία θα είναι δομημένη και η οποία θα επιτρέπει τη χρήση προηγμένων τεχνολογιών διαχείρισης γνώσης (knowledge management).*
 - ◆ *Η συγχώνευση και συγκέντρωση της ιατρικής πληροφορίας σε μεγάλες και διασυνδεδεμένες βάσεις δεδομένων*
- Έτσι, θα υπάρχει εύκολη, άμεση και επιτυχής αναζήτηση ιατρικών πληροφοριών, τόσο από εξουσιοδοτημένα άτομα (ιατρούς, υγειονομικούς φορείς κτλ) όσο και από απλούς πολίτες.
- Από τα μέσα της δεκαετίας του '80, αυξανόμενος αριθμός από βάσεις γνώσης έχει αναπτυχθεί στις επιστήμες υγείας. Αυτές οι βάσεις επιχειρούν να συνδυάσουν:
 - ◆ *στοιχεία βιβλιογραφίας (από άρθρα σε περιοδικά και εγχειρίδια) και*
 - ◆ *πραγματικά στοιχεία (π.χ., που καθοδηγούνται από τους εμπειρογνώμονες και που προκύπτουν από καλά τεκμηριωμένες περιπτώσεις ασθενών).*



Μερικοί Ορισμοί

- **Λεξικό:**
 - ◆ Βιβλίο αναφοράς που περιέχει αλφαβητικό κατάλογο λέξεων, με αντίστοιχη πληροφορία για κάθε λέξη, που περιέχει ερμηνεία (μετάφραση), προφορά και ετυμολογία.
- **Ορολογία:**
 - ◆ Ένα σύστημα που περιέχει τους όρους που χρησιμοποιούνται σε μια τέχνη ή σε μια επιστήμη, π.χ. ορολογία της Ορυκτολογίας.
- **Γνωστικός θησαυρός:**
 - ◆ Ένα βιβλίο επιλεγμένων λέξεων ή εννοιών παρόμοιο με ένα εξειδικευμένο λεξικό ενός συγκεκριμένου επιστημονικού χώρου, όπως η ιατρική, η γεωλογία, η μουσικολογία κτλ.
- **Ταξινόμηση:**
 - ◆ Η συστηματοποιημένη ομαδοποίηση των οργανισμών σε κατηγορίες στη βάση των εξελικτικών ή δομικών σχέσεων που έχουν μεταξύ τους. Στη Βιολογία οι οργανισμοί έχουν ταξινομηθεί σε βασιλεία, υποβασιλεία, κατηγορίες, σειρές, οικογένειες, είδη, υποείδη, ποικιλίες και μορφές.
- **Κατάλογος:**
 - ◆ Ένα σύνολο από τίτλους, που συνήθως περιέχει περιγραφική πληροφορία ή σχήματα.
- **Κωδικοποίηση:**
 - ◆ Η αναγωγή ενός συνόλου πληροφοριών σε κώδικα (συνοπτική τυποποίηση). Οι κώδικες είναι συνήθως αριθμητικοί ή γραμματο-αριθμητικοί.



Μερικοί Ορισμοί (συνέχεια)

- *Βάση δεδομένων (data base):*
 - ◆ είναι μια συλλογή απλών παρατηρήσεων οι οποίες ταξινομούνται με ένα δομημένο τρόπο.
 - ◆ Ένα αυτοματοποιημένο (ηλεκτρονικό) σύστημα αρχειοθέτησης θεωρείται πρωτογενώς ως μια βάση δεδομένων, δηλαδή ως το μέρος όπου αποθηκεύονται όλα τα στοιχεία των ασθενών.

- *Βάση γνώσης (Knowledge base):*
 - ◆ είναι μια συλλογή εμπειριών ευρετικών διαδικασιών και συνδυαστικών μοντέλων που μπορούν να χρησιμοποιηθούν για επίλυση προβλημάτων.



Συστήματα κωδικοποίησης

- Μεγάλο μέρος των εφαρμογών στο χώρο της υγείας σχετίζονται με επιδημιολογικές μελέτες ή τις τάσεις ορισμένων ασθενειών στον πληθυσμό.
- Κωδικοί για την ταξινόμηση των ασθενειών (*Ενιαίος κώδικας διαγνώσεων ICD-10*).
- Κωδικοί για τις κλινικές και εργαστηριακές πράξεις, όπως και για άλλες οντότητες του υγειονομικού χώρου.



Ιστορική αναδρομή της Διεθνούς Ταξινόμησης Ασθενειών (ICD)

1853	W. Farr (1807-1883) Τυποποίηση της ορολογίας αιτιών θανάτου
1893	J. Bertillon (1851-1922) Νέα ταξινόμηση αιτιών θανάτου
1946	WHO (Διεθνής Οργανισμός Υγείας) Διεθνής στατιστική ταξινόμηση ασθενειών, τραυμάτων και αιτιών θανάτου
1975	Ένατη διορθωμένη έκδοση (ICD-9)
1979	Κλινική τροποποίηση (ICD-9-CM, HCIMO)
1992	Διεθνής στατιστική ταξινόμηση των ασθενειών και συσχετιζόμενων προβλημάτων υγείας (ICD-10)



Κατηγορίες του συστήματος ICD-10

I	Λοιμώδη και παρασιτικά νοσήματα
II	Νεοπλασίες
III	Ασθένειες του αίματος και των αιμοποιητικών οργάνων, καθώς και διαταραχές που αναφέρονται σε ανοσοποιητικούς μηχανισμούς
IV	Ενδοκρινικά, διατροφικά και μεταβολικά νοσήματα
V	Διαταραχές νόησης και συμπεριφοράς
VI	Ασθένειες του νευρικού συστήματος
VII	Οφθαλμολογικές ασθένειες
VIII	Ασθένειες των ώτων και της μαστοειδούς διαδικασίας
IX	Ασθένειες του κυκλοφορικού συστήματος
X	Ασθένειες του αναπνευστικού συστήματος
XI	Ασθένειες του πεπτικού συστήματος
XII	Ασθένειες του δέρματος και του υποδόριου ιστού
XIII	Ασθένειες του μυοσκελετικού συστήματος και του συνδετικού ιστού
XIV	Ασθένειες του ουροποιητικού συστήματος
XV	Εγκυμοσύνη, τοκετός και επιλόχειες επιπλοκές
XVI	Ανωμαλίες προερχόμενες από την προγενετική περίοδο
XVII	Συγγενείς παραμορφώσεις και χρωμοσωμικές ανωμαλίες
XVIII	Συμπτώματα, σημεία και παθολογικά κλινικά ή εργαστηριακά ευρήματα μη ταξινομημένα σε ειδικές κατηγορίες
XIX	Τραύματα, δηλητηριάσεις και συνέπειες άλλων εξωτερικών παραγόντων
XX	Εξωτερικές αιτίες νοσηρότητας και θνητότητας
XXI	Παράγοντες που επηρεάζουν την υγειονομική κατάσταση και τις υγειονομικές υπηρεσίες



Παράδειγμα ταξινόμησης ιατρικών πράξεων στο σύστημα ICD-10

35-39	Επεμβάσεις στο καρδιαγγειακό σύστημα
35	Επεμβάσεις στις καρδιακές βαλβίδες και καρδιακό διάφραγμα Περιέχει: στερνοτομή και θωρακοτομή ως χειρουργικές προσεγγίσεις, καθώς επίσης και το καρδιοπνευμονικό bypass (39.61).
	<i>35.0 Βαλβιδοτομή κλειστής καρδιάς μη συμβατή με: εξωδερματική βαλβιδοπλαστική (μπαλόνη) (35.96)</i>
	<i>35.00 Βαλβιδοτομή κλειστής καρδιάς σε μη προσδιορισμένη βαλβίδα</i>
	<i>35.01 Βαλβιδοτομή κλειστής καρδιάς στην αορτική βαλβίδα</i>
	<i>35.02 Βαλβιδοτομή κλειστής καρδιάς στη μιτροειδή βαλβίδα</i>
35.1	Βαλβιδοπλαστική ανοικτής καρδιάς χωρίς αντικατάσταση
35.2	Αντικατάσταση καρδιακής βαλβίδας

- Το αμιγές ICD-10 είναι ταξινόμηση που ταιριάζει περισσότερο σε στατιστικούς και όχι τόσο σε ιατρούς.



Το σύστημα SNOMED (Systemized Nomenclature of Human and Veterinary MEDicine International)

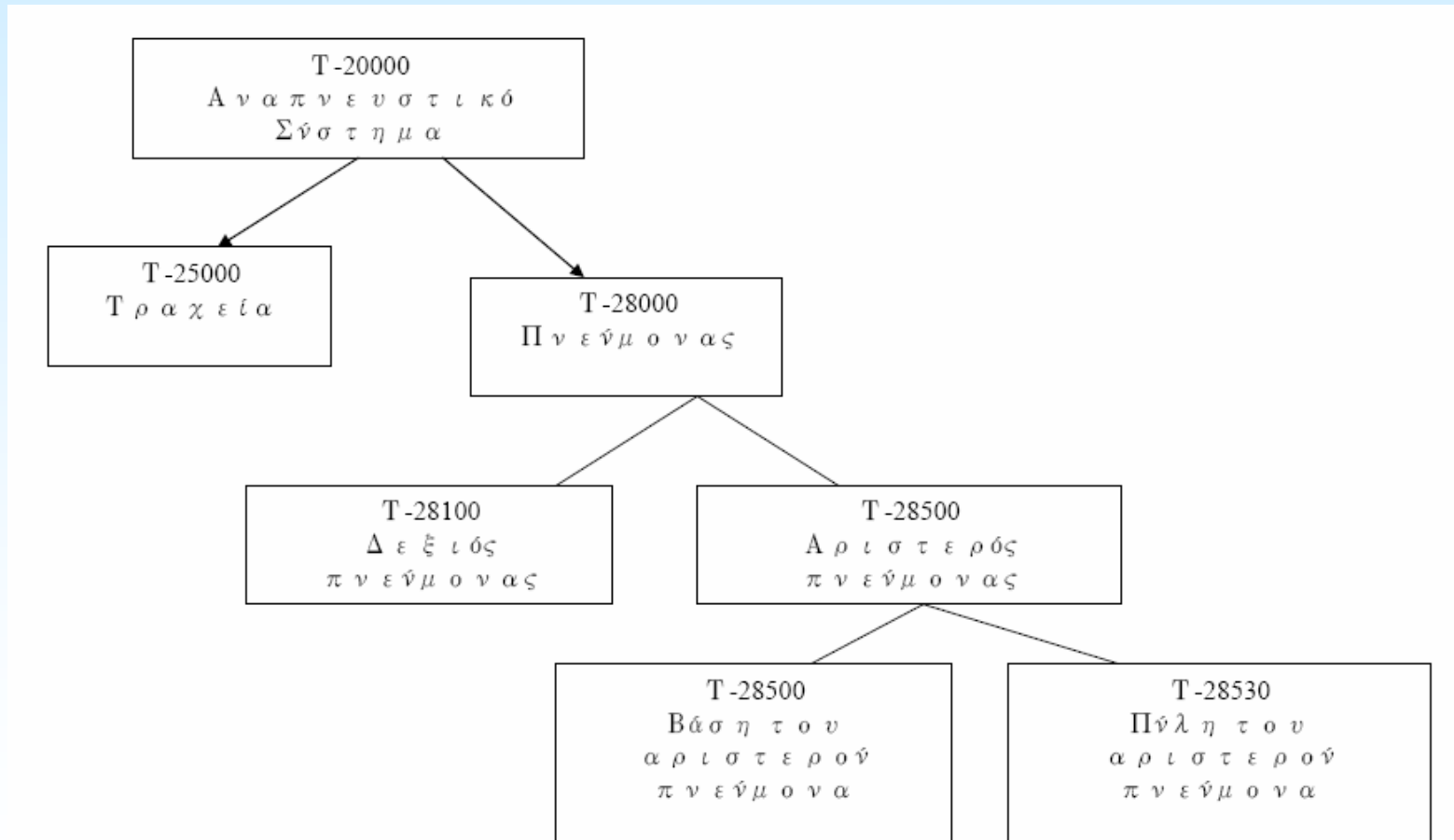
- *T για τοπογραφία:*
 - ◆ Περιγράφει τα μέρη του ανθρώπινου σώματος, όργανα και περιοχές (12.385 όροι).
- *M για μορφολογία:*
 - ◆ Αναφέρεται σε εκ γενετής ή προκληθείσες ανατομικές και κυτταρικές αλλοιώσεις. (4.991 όροι).
- *L για ζώντες οργανισμούς:*
 - ◆ Ταξινόμηση ζώων και φυτών που είναι βασικά παθογόνοι φορείς ασθενειών (25.265 όροι).
- *C για χημικά, φαρμακευτικά και βιολογικά προϊόντα:*
 - ◆ Παρουσιάζονται οι διάφορες κατηγορίες φαρμάκων και θεραπειών, καθώς επίσης και το σύνολο των χημικών ή φυτικών ουσιών που έχουν ιατρικές εφαρμογές (14.075 όροι).
- *A για φυσικούς παράγοντες, δυνάμεις και δραστηριότητες:*
 - ◆ Περιέχει έναν κατάλογο από δραστηριότητες, εργαλεία και συσκευές που σχετίζονται με ασθένειες και τραύματα (1353 όροι).



Το σύστημα SNOMED (συνέχεια)

- *J για την επαγγελματική απασχόληση:*
 - ◆ *Κωδικοποίηση των επαγγελματών σύμφωνα με την ταξινόμηση που έχει καθιερωθεί από το Διεθνές Γραφείο Εργασίας (ILO) (1.886 όροι).*
- *S για το κοινωνικό περιβάλλον:*
 - ◆ *Λίστα από κοινωνικές συνθήκες που παρουσιάζουν σημαντικό ιατρικό ενδιαφέρον (433 όροι).*
- *D για τους ασθενείς:*
 - ◆ *Περιγράφει τις ασθένειες και τους συνδυασμούς σημείων και συμπτωμάτων. Επιπλέον, δίνει τις αντιστοιχίες με τους διαγνωστικούς όρους που περιέχονται στο ICD-9-CM (28.622 όροι).*
- *P για το σύνολο των υγειονομικών διαδικασιών:*
 - ◆ *Περιγράφει τις διοικητικές, διαγνωστικές και θεραπευτικές δραστηριότητες που χρησιμοποιούνται για την πρόληψη και θεραπεία των ασθενειών (περίπου 25.000 όροι).*
- *G για την εκτέλεση διασυνδέσεων και τροποποιήσεων:*
 - ◆ *Καθορίζει έναν κατάλογο όρων που χρησιμοποιούνται για το χαρακτηρισμό ή τη διασύνδεση όρων που ανήκουν σε διαφορετικές κατηγορίες.*

Ιεραρχική δομή του διεθνούς συστήματος ταξινόμησης SNOMED





Ένας κανόνας

- Με μια απλή διατύπωση της μορφής
 - ◆ $T + M + F + E = D$,
- και χρησιμοποιώντας τους κατάλληλους όρους από κάθε κατηγορία, μπορεί να τυποποιηθεί η παρακάτω έκφραση:
 - ◆ Σε μια συγκεκριμένη τοπογραφική περιοχή (T) υπάρχει μια ειδική μορφολογική αλλοίωση (M), συσχετιζόμενη με μια ορισμένη δυσλειτουργία (F), που οφείλεται σε ένα συγκεκριμένο αιτιολογικό παράγοντα (E). Το σύνολο όλων αυτών των στοιχείων προσδιορίζει μια ασθένεια ή σύνδρομο D



Χρησιμότητα

- Η υλοποίηση τέτοιων διαδικασιών επιτρέπει τη δημιουργία προτύπων μέσω των οποίων μπορεί να ελεγχθεί η ποιότητα των ιατρικών πράξεων του κάθε ιατρού ξεχωριστά ή συνολικά μιας ιατρικής μονάδας (κλινική ή εργαστήριο).
- Έτσι, με τέτοια πρότυπα και χρησιμοποιώντας δείγματα επαρκούς μεγέθους μπορεί να ελεγχθεί στατιστικά αν:
 - ◆ *η διάρκεια νοσηλείας για ένα συγκεκριμένο νόσημα σε μια συγκεκριμένη κλινική ήταν πάνω ή κάτω από το κατά μέσο όρο (norma) προβλεπόμενο,*
 - ◆ *ένας ιατρός παρήγγειλε, κατά μέσο όρο, περισσότερες από τις προβλεπόμενες εργαστηριακές εξετάσεις για τη διάγνωση ενός νοσήματος.*



Μειονέκτημα συστήματος SNOMED

- Η κωδικοποίηση SNOMED έχει πάνω από 65000 όρους και θεωρείται δύσχρηστη για άμεση χρήση από άνθρωπο.
- Έχει αξία μόνο σε αυτοματοποιημένα συστήματα κωδικοποίησης ιατρικών κειμένων.



Κωδικοποίηση LOINC (Laboratory Observation Identifier Names. and Codes)

- Η κωδικοποίηση LOINC είναι μία βάση δεδομένων με πάνω από 6.000 τύπους κωδικοποιημένων εργαστηριακών ευρημάτων, που αναπτύχθηκε στις ΗΠΑ.
- Έχει ευρεία χρήση αλλά δεν θεωρείται ιδιαίτερος πρακτική καθώς έχει πάρα πολλούς όρους.



Παραδείγματα Κωδικοποίησης LOINC

Class	LOINC #	Short Name	Related Names
ALLERGY	15570-5	Honey Bee IgE RAST Ql	i1; immune globulin e; immunoglobulin e; honey bee; honeybee; honey bee; arbitrary concentration; point in time; random; serum; sr; ql; ordinal; ql; qualitative; qual; screen; antibodies; autoantibody; antibody; autoantibodies; antby; aby; anti; allergy testing
CHAL	40069-7	Calcium 1.5H p chal SerPl-mCnc	cal; ca; 1.5h p chal; mass concentration; level; point in time; random; serpl; serplas; serp; serum; sr; plasma; pl; plsm; quantitative; qnt; quant; quan; pst; 90 minutes; 1 1/2 hours; 90m; 1.5hr; 1 1/2 hr; chemistry.challenge testing; chemistry.challenge testing



Το πρότυπο ICPC (International Classification in Primary Care)

- Το πρότυπο ICPC (Διεθνής Ταξινόμηση Προβλημάτων Υγείας Πρωτοβάθμιας Περίθαλψης) αναφέρεται σε ιατρικές και νοσηλευτικές πράξεις της πρωτοβάθμιας φροντίδας υγείας (ΠΦΥ), κωδικοποιεί επιτυχώς το 90% των επισκέψεων στην ΠΦΥ, ενώ κωδικοποιεί και κοινωνικού χαρακτήρα προβλήματα.
- Σήμερα βρισκόμαστε στη δεύτερη έκδοση του προτύπου (ICPC-2), η οποία έχει αποδοθεί και στα Ελληνικά.
- Τα δύο βασικότερα πλεονεκτήματα του προτύπου είναι ότι είναι πολύγλωσσο και ότι είναι αναγνωρισμένο από την επιστημονική βιβλιογραφία και τη βιβλιογραφία της δημόσιας υγείας.



Παράδειγμα κωδικοποίησης κατά ICPC-2

Κωδικός ICPC	Ορισμός κωδικού
A03	Πυρετός
R05	Βήχας
R21	Συμπτώματα / ενοχλήσεις λαιμού
R07	Πταρμός / ρινική συμφόρηση



Το πρότυπο ATC (Anatomic Therapeutic Classification System)

- Η κωδικοποίηση κατά ATC της παγκόσμιας οργάνωσης υγείας είναι η επικρατέστερη διεθνής κωδικοποίηση που αφορά στα φάρμακα και τις δραστικές ουσίες.
- Είναι πολύ πρακτική για το ιατρό γιατί:
 - ◆ *Επιτρέπει την επιλογή συγγενών φαρμάκων.*
 - ◆ *Προσδιορίζει και τις προτεινόμενες DDD (Defined Daily Doses) για κάθε δραστική ουσία.*



Παράδειγμα ταξινόμησης κατά ATC

				(DDD)	(U)
A04	A04	ANTIEMETICS AND ANTINAUSEANTS			
A04A	A04A	ANTIEMETICS AND ANTINAUSEANTS			
A04AA	A04AA	<i>Serotonin (5HT₃) antagonists</i>			
	A04AA01	01	Ondansetron	16	mg
	A04AA01			16	mg
	A04AA01			16	mg
	A04AA02	02	Granisetron	2	mg
	A04AA02			3	mg
	A04AA03	03	Tropisetron	5	mg
	A04AA03			5	mg
	A04AA04	04	Dolasetron	0.2	g
	A04AA04			0.1	g
	A04AA05	05	Palonosetron	0.25	mg
A04AD	A04AD	<i>Other antiemetics</i>			
	A04AD01	01	Scopolamine		
	A04AD02	02	Cerium oxalate		
	A04AD04	04	Chlorobutanol		
	A04AD05	05	Metopimazine	15	mg
	A04AD05			15	mg
	A04AD10	10	Dronabinol		
	A04AD11	11	Nabilone		
	A04AD12	12	Aprepitant	95	mg
	A04AD51	51	Scopolamine, combinations		
	A04AD54	54	Chlorobutanol, combinations		



Χαρακτηριστικά Συστημάτων Κωδικοποίησης

- *Πληρότητα*
 - ◆ Μια πλήρης περιγραφή των αντικειμένων του ιατρικού χώρου είναι πολύ δύσκολο να επιτευχθεί.
- *Απουσία αμφιλογίας*
 - ◆ Οι όροι του συστήματος ταξινόμησης πρέπει να αναφέρονται μόνο σε μια έννοια. Εάν ένας όρος είναι αμφιλεγόμενος, τότε τουλάχιστον δύο διαφορετικοί τύποι δεδομένων καταγράφονται κάτω από τον ίδιο όρο, γεγονός που επηρεάζει την ειδικότητα των διερευνήσεων.
- *Απουσία πλεονασμών*
 - ◆ Κάθε έννοια πρέπει να εκφράζεται μονοσήμαντα. Εάν δύο όροι αναφέρονται στην ίδια έννοια, τότε μειώνεται η ευαισθησία των διερευνήσεων που επιτελούνται στα πλαίσια της βάσης δεδομένων.



Χαρακτηριστικά Συστημάτων Κωδικοποίησης (συνέχεια)

- *Συνωνυμίες*
 - ◆ Είναι πολύ σημαντικό να έχει το σύστημα ταξινόμησης τη δυνατότητα διαχείρισης των συνωνυμιών. Η ύπαρξη συνωνυμιών πρέπει να διαχωρίζεται από τον πλεονασμό, ο οποίος πρέπει να αποφεύγεται. Τα συνώνυμα επιτρέπονται ως ενδιάμεσοι όροι που αναφέρονται σε ένα μοναδικό όρο (εσωτερική κωδικοποίηση), ο οποίος χρησιμοποιείται για την κωδικοποίηση και εύρεση της χρήσιμης πληροφορίας.
- *Σαφείς συσχετίσεις*
 - ◆ Όταν οι τύποι των συσχετίσεων μεταξύ των όρων δεν είναι σαφείς, τότε η ποιότητα των διερευνήσεων υποβαθμίζεται. Οι πιο συχνές συσχετίσεις έχουν τις εξής μορφές: είναι ένα, είναι μέρος του, σχετίζεται με, ισοδύναμο του, είναι εντός. Εάν θέλουμε να διερευνήσουμε μια βάση δεδομένων για να βρούμε τις πνευμονικές ασθένειες ενός ατόμου, θα συνδέσουμε όλους τους σχετικούς όρους με την έννοια «πνευμονική ασθένεια» χρησιμοποιώντας τη σχέση που έχει τη μορφή είναι ένα.



Ηλεκτρονικές βιβλιοθήκες

- Υπολογίζεται σήμερα ότι στο τομέα της Ιατρικής δημοσιεύονται:
 - ◆ *600.000 άρθρα το χρόνο*
 - ◆ *80.000 επιστημονικά περιοδικά*
 - ◆ *200 άρθρα το χρόνο ανά έμπειρο αναγνώστη*
- Ένας γιατρός που διαβάζει 2 επιστημονικά άρθρα την ημέρα από την τρέχουσα ιατρική βιβλιογραφία, χρειάζεται 800 χρόνια για να διαβάσει την ετήσια βιβλιογραφία
- Τα δεδομένα αυτά οδηγούν στην ανάγκη ύπαρξης βιβλιοθηκών οι οποίες θα προσφέρουν:
 - ◆ *Κωδικοποίησης της πληροφορίας.*
 - ◆ *«Έξυπνες» αναζητήσεις για την ανάγνωση μόνο των σημαντικών για αυτών άρθρων.*



Ο ρόλος του διαδικτύου

- Το διαδίκτυο και ο παγκόσμιος ιστός αποτελούν την πιο διαδεδομένη και ισχυρή εφαρμογή του 21ου αιώνα και επηρεάζουν ουσιαστικά κάθε πλευρά της κοινωνικής ζωής και της επιστήμης.
- Ειδικότερα στην ιατρική, οι νέες αυτές μαζικές τεχνολογικές δυνατότητες επιδρούν θετικά και δραστικά στην καθημερινή της πρακτική και επιβάλλουν την σταδιακή απορρόφηση και την αξιοποίηση τους
- Ο κυβερνοχώρος (WWW) και σε αυτή την περίπτωση είναι το μέσο που υπερισχύει με βάση τα μεγέθη. Το αναπόφευκτο ερώτημα είναι αν και με ποιο τρόπο είναι δυνατό να αξιοποιηθεί πλέον αυτή η υπερπληθώρα πληροφορίας και να μετατραπεί σε γνώση.



Οι περιορισμοί του διαδικτύου

- Η γρήγορη, ακριβής και έγκυρη πληροφόρηση μέσω του διαδικτύου, δεν είναι ξεκάθαρη, τουλάχιστον για τους ανειδίκευτους χρήστες, οι οποίοι θα επιθυμούσαν να χρησιμοποιούν ερωτήματα σε φυσική γλώσσα.
- Τα keywords στις αναζητήσεις είναι προφανές ότι δεν αντικατοπτρίζουν τη γνώση οπότε συχνά καταλήγουν σε έναν αυξημένο όγκο πληροφορίας ο οποίος σε μεγάλο βαθμό δεν ικανοποιεί τα αντιληπτικά ή σημασιολογικά κριτήρια του ενδιαφερόμενου.
- Οι υπάρχουσες απλές μηχανές αναζήτησης καλύπτουν μόνο τη δυνατότητα κάποιων λογικών εκφράσεων. Αυτό είναι ιδιαίτερα προβληματικό σε γλώσσες όπως η Ελληνική, η οποία έχει πλούσια μορφολογία.



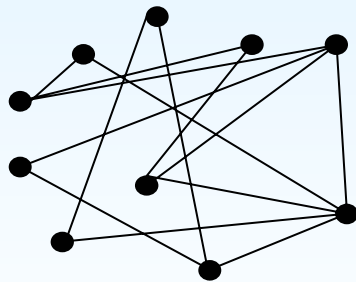
Το όνειρο του σημασιολογικού ιστού (semantic web)

- Εάν οι πόροι του διαδικτύου γίνουν περισσότερο *σημασιολογικοί* για τους υπολογιστές και τους ανθρώπους, υπάρχει η ελπίδα ότι θα δημιουργηθεί ένα *ευφρές διαδίκτυο* που θα παρέχει τη σωστή πληροφόρηση στα διάφορα ιατρικά προβλήματα και άμεσες απαντήσεις στα συγκεκριμένα ερωτήματα των απλών πολιτών αλλά και των εξουσιοδοτημένων παροχέων υγείας.

Το όνειρο του σημασιολογικού ιστού (συνέχεια)

Προβλήματα του WWW

Οι αναζητήσεις στο WWW εξετάζουν μόνο το 20% των σχετικών ιστοσελίδων και επιστρέφουν μεγάλη ποσότητα άσχετης πληροφορίας



Διπλασιάζεται
σε μέγεθος
κάθε 6 μήνες

Η πληροφορίες στο www δεν μπορούν εύκολα να επεξεργαστούν από software agents

Σημασιολογικός Ιστός



Η πληροφορία στο διαδίκτυο είναι καθορισμένη και συνδεδεμένη έτσι ώστε να μπορεί να χρησιμοποιηθεί από τους υπολογιστές για την αυτοματοποίηση, την ενσωμάτωση και την επαναχρησιμοποίηση δεδομένων σε πολλές εφαρμογές



Online βάσεις ιατρικών δεδομένων και βιβλιοθήκες

- Σχεδόν παράλληλα με την ανάπτυξη του παγκόσμιου ιστού δημιουργήθηκαν ηλεκτρονικές βιβλιοθήκες σχετικές με τον τομέα της ιατρικής επιστήμης.
- Οι βιβλιοθήκες ιατρικού περιεχομένου, αρχικά είχαν ως στόχο να ενημερώνουν άμεσα τον απλό πολίτη σχετικά με ερωτήματα και απορίες που πιθανόν θα προέκυπταν σε ιατρικά θέματα. Επιπλέον, διασυνδεδεμένες βάσεις και ιατρικά forums, συντελούσαν στη διακίνηση της ιατρικής πληροφορίας από ειδικούς (ιατρούς, υγειονομικούς φορείς) και την προαγωγή της έρευνας μέσω της ανταλλαγής απόψεων και ερευνητικών αποτελεσμάτων.
- Πέρα από την απλή εγκυκλοπαιδική πληροφόρηση γίνεται φανερό ότι παρόμοια συστήματα θα μπορούσαν να διενεργούν ένα είδος αυτόματης προκαταρκτικής εξέτασης στον ασθενή, με μεγάλη συνεισφορά στην έγκαιρη διάγνωση και πρόληψη συμβάντων υγείας.
 - ◆ *Θέτοντας ο ασθενής τα συμπτώματά του θα μπορούσε να λάβει μια πρώτη διάγνωση.*
 - ◆ *Επιπλέον, ο εξειδικευμένος ιατρός θα μπορούσε να έχει στα χέρια του ένα πολύ ισχυρό στατιστικό εργαλείο μέσω της συγκέντρωσης παρόμοιων περιστατικών και αντίστοιχων διαγνώσεων.*



Ομογενοποίηση της Ιατρικής πληροφορίας

- Ο μεγάλος όγκος της ιατρικής πληροφορίας δημιούργησε το πρόβλημα της ασυμβατότητας των ιατρικών δεδομένων.
- Η ομογενοποίηση της ιατρικής πληροφορίας αποτέλεσε και αποτελεί μία μεγάλη πρόκληση για όλους τους παροχής ιατρικής πληροφόρησης.
- Τα κύρια προβλήματα που παρουσιάζονται αφορούν:
 - ◆ *Εξάρτηση από το περιεχόμενο της πηγής πληροφόρησης*
 - ☞ Συμφραζόμενα
 - ☞ Ιστορική εξέλιξη όρων και δεδομένων
 - ☞ Χωροχρονική κατάσταση χρήσης όρων
 - ☞ Οπτική γωνία και πεποιθήσεις του πληροφοριοδότη
 - ☞ Χρήση του ίδιου όρου από διαφορετικά υποπεδία της ιατρικής
 - ◆ *Συσχετίσεις όρων*
 - ☞ Εντοπισμός όρων σε παραπάνω από μία ιεραρχικές δομές
 - ☞ Διαφορετικοί όροι για την ίδια έννοια
 - ☞ Διαφορετικές περιγραφές για την ίδια έννοια
 - ◆ *Επίπεδα λεπτομέρειας στη χρήση όρων και περιγραφών*
- Μέχρι σήμερα έχουν γίνει μεγάλες προσπάθειες για την ομογενοποίηση των ιατρικών όρων από ετερογενείς πηγές αποθήκευσης. Η σημαντικότερη είναι αυτή που διεξάχθηκε από την NLM και οδήγησε στην δημιουργία της οντολογίας MeSH και του εξυπηρετητή πηγής γνώσης (Knowledge Source Server) UMLS (Unified Medical Language System).



Τμήμα Τεχνολογίας Ιατρικών Οργάνων
Εργαστήριο Βιοϊατρικής Τεχνολογίας



United States
National Library of Medicine
National Institutes of Health

- Η National Library of Medicine είναι η μεγαλύτερη ιατρική βιβλιοθήκη παγκοσμίως και είναι ένα από τα Εθνικά Ινστιτούτα Υγείας (National Institutes of Health - NIH) των Ηνωμένων Πολιτειών.



Η NLM παρέχει

- Πληροφορίες υγείας
MEDLINE: βιβλιογραφική βάση δεδομένων
PubMed: μηχανή αναζήτησης
- Υπηρεσίες βιβλιοθήκης
Κατάλογοι, Βάσεις Δεδομένων, MeSH (οντολογία ιατρικών όρων),
Ιστορικές πληροφορίες, Δημοσιεύσεις
- Ερευνητικά προγράμματα
Computational Molecular Biology, UMLS (knowledge source server),
Medical Informatics, Visible Human Project



MEDLINE®

- Η MEDLINE είναι η αρχαιότερη βιβλιογραφική βάση δεδομένων της NLM (National Library of Medicine).
- Καλύπτει τα πεδία της φαρμακευτικής, παιδιατρικής, οδοντιατρικής, κτηνιατρικής, του συστήματος περίθαλψης υγείας και τις προ-κλινικές έρευνες όπως η μοριακή βιολογία.
- Περιέχει βιβλιογραφικές αναφορές σε δημοσιευμένες ιατρικές πληροφορίες (με τις παραπομπές και τις περιλήψεις τους) από περίπου 4.600 βιοιατρικά περιοδικά της τρέχουσας επικαιρότητας τα οποία έχουν εκδοθεί στις Ηνωμένες Πολιτείες και σε άλλες 70 ακόμη χώρες.
- Η βάση δεδομένων της περιέχει περίπου συνολικά 12 εκατομμύρια δημοσιεύματα τα οποία χρονολογούνται από τα μέσα της δεκαετίας του 1960.



MEDLINE®

- Χρησιμοποιεί την ιατρική οντολογία **MeSH** μέσω της οποίας γίνεται η σημασιολογική ερμηνεία των εγγράφων.

Descriptors

PMID: Pubmed Identifier	PT: Publication Type
UID: Unique Identifier	DA: Date of Acceptance
TI: Title	DP: Date of Publication
AU: Author	AB: ABstract
LA: LAnguage	SO: Source of Publication
MH: MeSH term related (σημασιολογικά μεταδεδομένα)	



Τμήμα Τεχνολογίας Ιατρικών Οργάνων
Εργαστήριο Βιοϊατρικής Τεχνολογίας



MEDLINE[®]

- Η πρόσβαση στη MEDLINE γίνεται μέσω της μηχανής αναζήτησης PubMed.

PubMed
www.pubmed.gov

A service of the [U.S. National Library of Medicine](#)
and the [National Institutes of Health](#)



Medical Subject Headings



- Το MeSH (Medical Subject Headings) είναι ο θησαυρός ορολογιών που έχει αναπτύξει η NLM.
- Οι ορολογίες του MeSH είναι αποθηκευμένες σε αλφαβητική σειρά αλλά και σε προκαθορισμένη ιεραρχική δομή (δέντρο).
- Στο πιο γενικό επίπεδο της ιεραρχικής δομής υπάρχουν έννοιες όπως «Ανατομία» ή «Ψυχικές Διαταραχές», ενώ σε ειδικότερα επίπεδα υπάρχουν έννοιες όπως «Αστράγαλος», «Βιταμίνη C».



Medical Subject Headings

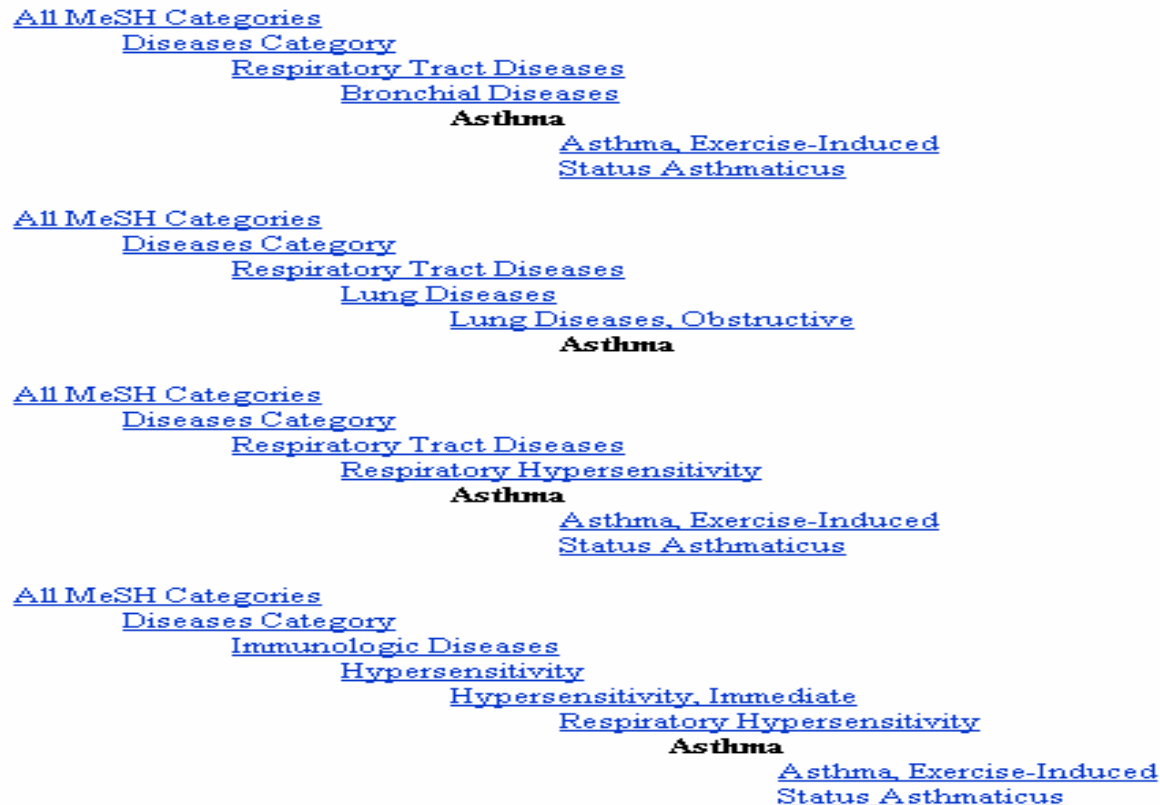


- Υπάρχουν 21.973 όροι στο MeSH και κάθε όρος μπορεί να εμφανιστεί σε περισσότερα από ένα κλαδιά.
- Ο θησαυρός MeSH χρησιμοποιείται από την NLM για indexing άρθρων.
- Τα άρθρα αυτά προέρχονται από τα 4.600 κορυφαία βιοιατρικά περιοδικά και χρησιμοποιούνται στη βάση MEDLINE.
- Το MeSH χρησιμοποιείται σε πολλές βάσεις της NLM που περιέχουν καταλόγους βιβλίων και αρχείων

Medical Subject Headings



- Ο ίδιος όρος εμφανίζεται σε πολλαπλά κλαδιά του ιεραρχικού πολυδένδρου.





Medical Subject Headings



- Κάθε όρος περιγράφεται από τις παρακάτω ιδιότητες:

Name: Name of the term
Definition: Medical Definition
Related Terms: Other related terms
Subheadings: Allowed Subheadings to modify and complement the meaning of this term
Position on the polytree: Fathers and sons on all locations of the term



A service of the U.S. National Library of Medicine
and the National Institutes of Health

- Το σύστημα αναζήτησης **PubMed** αναπτύχθηκε από το National Center for Biotechnology Information (NCBI) σε συνεργασία με εκδότες βιοϊατρικής βιβλιογραφίας.
- Είναι ένα εργαλείο αναζήτησης βιβλιογραφικών παραπομπών το οποίο χρησιμοποιείται και ως σύνδεσμος για τίτλους περιοδικών των εκδοτών που συμμετέχουν.
- Οι εκδότες προμηθεύουν στην NLM βιβλιογραφικές εγγραφές, η οποία μετά τις προσθέτει στο σύστημα αναζήτησης PubMed.



Τμήμα Τεχνολογίας Ιατρικών Οργάνων
Εργαστήριο Βιοϊατρικής Τεχνολογίας



A service of the [U.S. National Library of Medicine](#)
and the [National Institutes of Health](#)

NCBI PubMed A service of the U.S. National Library of Medicine and the National Institutes of Health My NCBI [Sign In] [Register]

All Databases PubMed Nucleotide Protein Genome Structure OMIM PMC Journals Books

Search PubMed for medical records Go Clear [Advanced Search \(beta\)](#) [Save Search](#)

Limits Preview/Index History Clipboard Details

Display Summary Show 20 Sort By Send to

All: 99449 Review: 4257

Items 1 - 20 of 99449 Page 1 of 4973 Next

1: [Sinawat S, Sanguansak T, Angkawinijwong T, Ratanapakorn T, Intapan PM, Sinawat S, Yospaiboon Y.](#) Related Articles, Links
Ocular angiostrongyliasis: clinical study of three cases.
Eye. 2008 Jun 6. [Epub ahead of print]
PMID: 18535614 [PubMed - as supplied by publisher]

2: [Davis JL, Morris A, Kallet RH, Powell K, Chi AS, Bensley M, Luce JM, Huang L.](#) Related Articles, Links
Low Tidal Volume Ventilation Is Associated with Reduced Mortality in HIV-infected Patients with Acute Lung Injury.
Thorax. 2008 Jun 5. [Epub ahead of print]
PMID: 18535118 [PubMed - as supplied by publisher]

3: [Ulrichsen J, Lützen L, Poulsen LM.](#) Related Articles, Links
[Respiratory depression in delirium tremens patients treated with phenobarbital.]
Ugeskr Laeger. 2008 Jun 2;170(23):2018-2022. Danish.
PMID: 18534165 [PubMed - as supplied by publisher]

4: [Toth-Pal E, Wårdh I, Strender LE, Nilsson G.](#) Related Articles, Links
A guideline-based computerised decision support system (CDSS) to influence general practitioners management of chronic heart failure.
Inform Prim Care. 2008;16(1):29-39.
PMID: 18534075 [PubMed - in process]

5: [Soudry E, Marmor Y, Hazan A, Marx S, Sadov R, Feinmesser R.](#) Related Articles, Links
Supracricoid partial laryngectomy: an alternative to total laryngectomy for locally advanced laryngeal cancers.



- Το PubMed επιτρέπει διαφορετικούς τύπους αναζήτησης για ένα Keyword χρησιμοποιώντας περιορισμούς που ονομάζονται Search modifiers.

MAJR: Mesh term as major topic

MH:NOEXP: Mesh term without searching the descendants

MH: Like previous expanding the search through descendants

TI: The term in the Title

TW: The same, plus looking the term in the Abstract

PT: This term as Publication Type

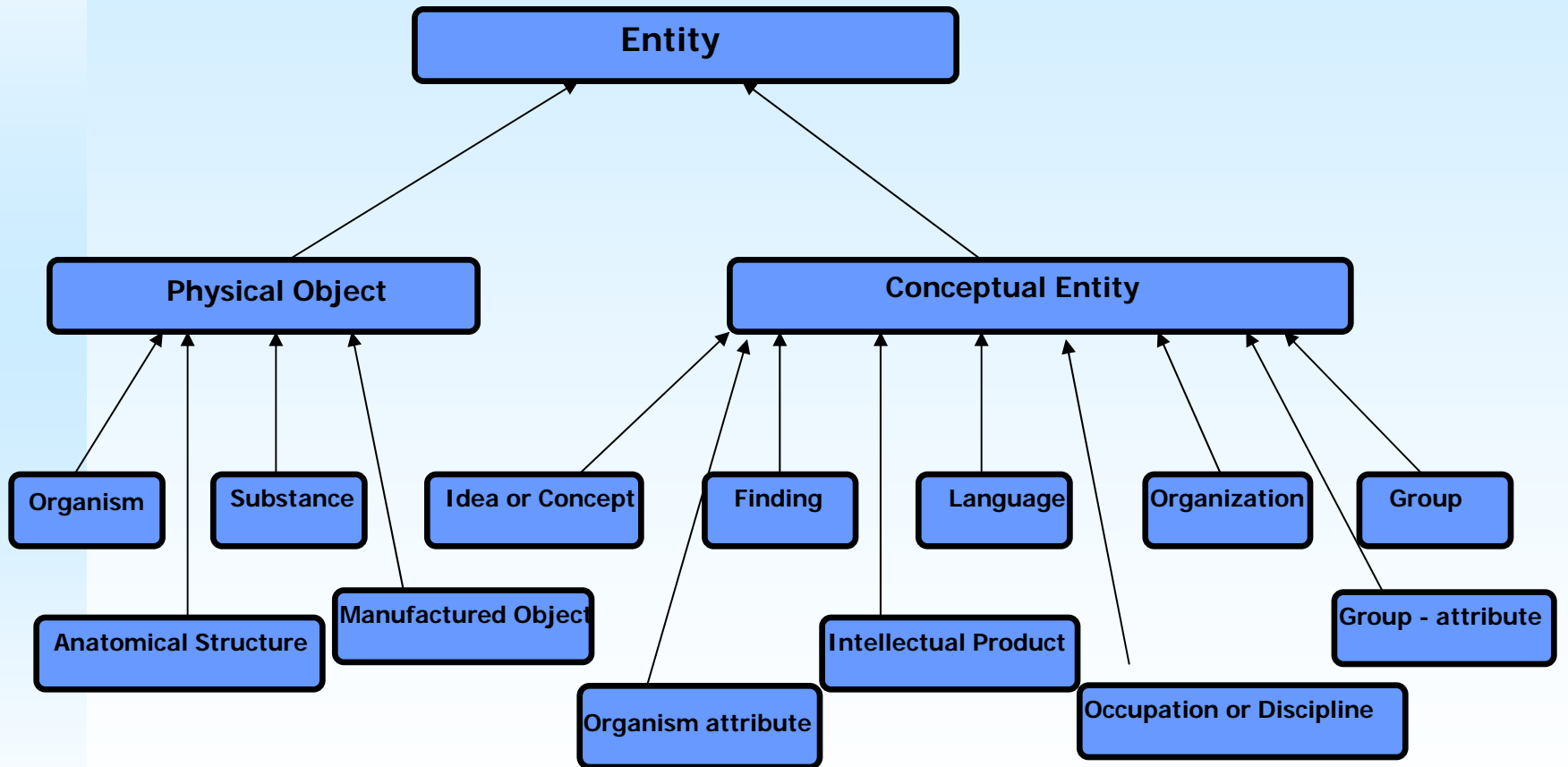


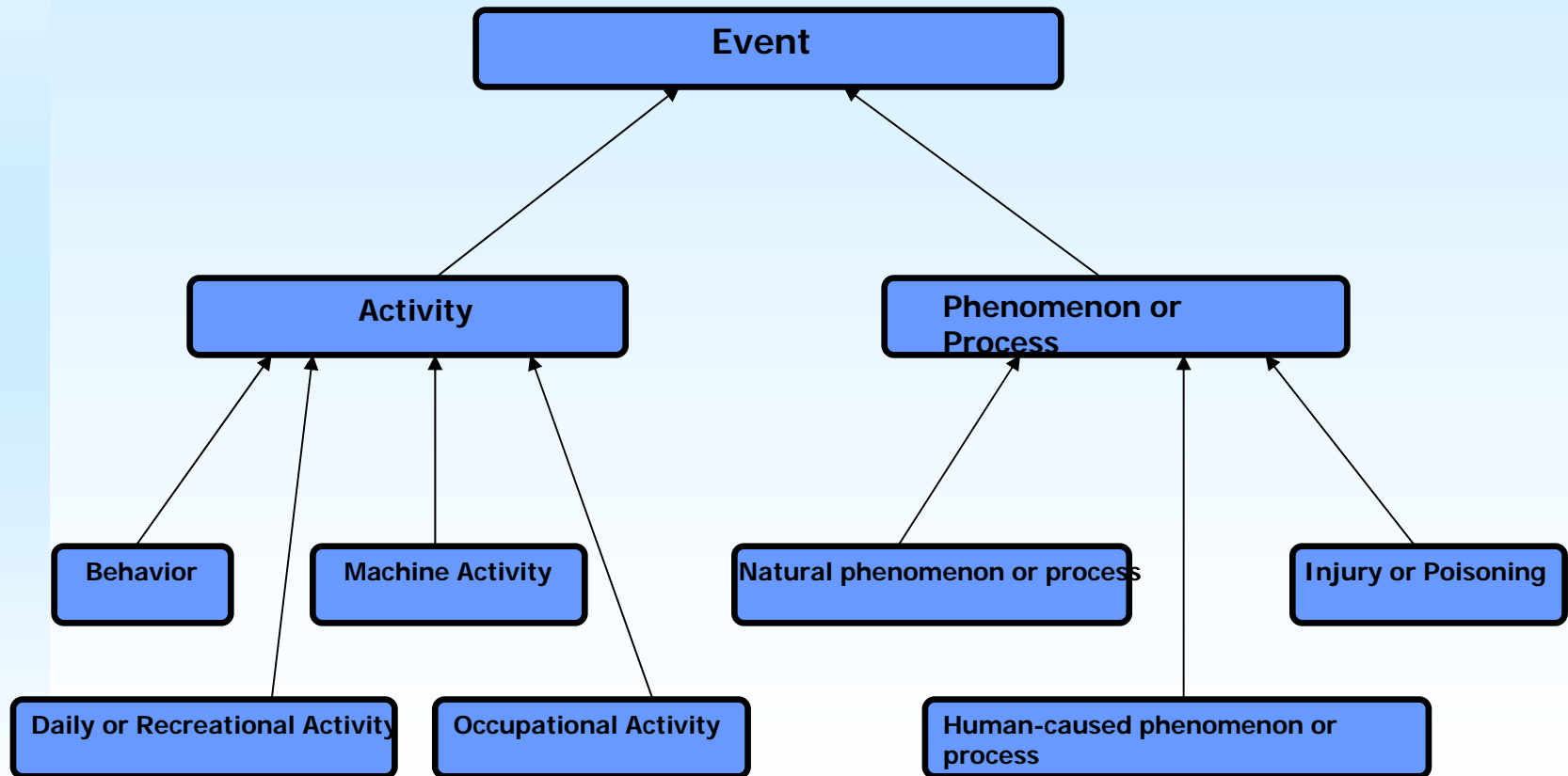
Το **Unified Medical Language System (UMLS)**:

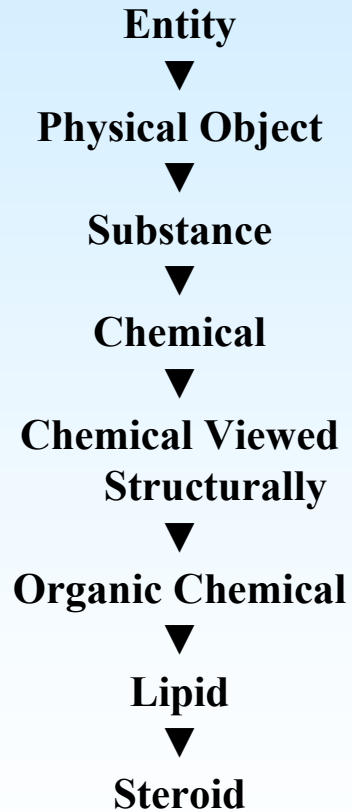
- Έχει ως στόχο τη δημιουργία ενός ενοποιημένου μεταθησαυρού που θα συνδυάζει επιμέρους θησαυρούς.
- Διευκολύνει την ανάκτηση πλήρους πληροφορίας η οποία προέρχεται από κατανεμημένες πηγές όπως βιοιατρική βιβλιογραφία, ιατρικούς φακέλους υγείας, βάσεις γνώσης και καταλόγους για ιατρικούς οργανισμούς.



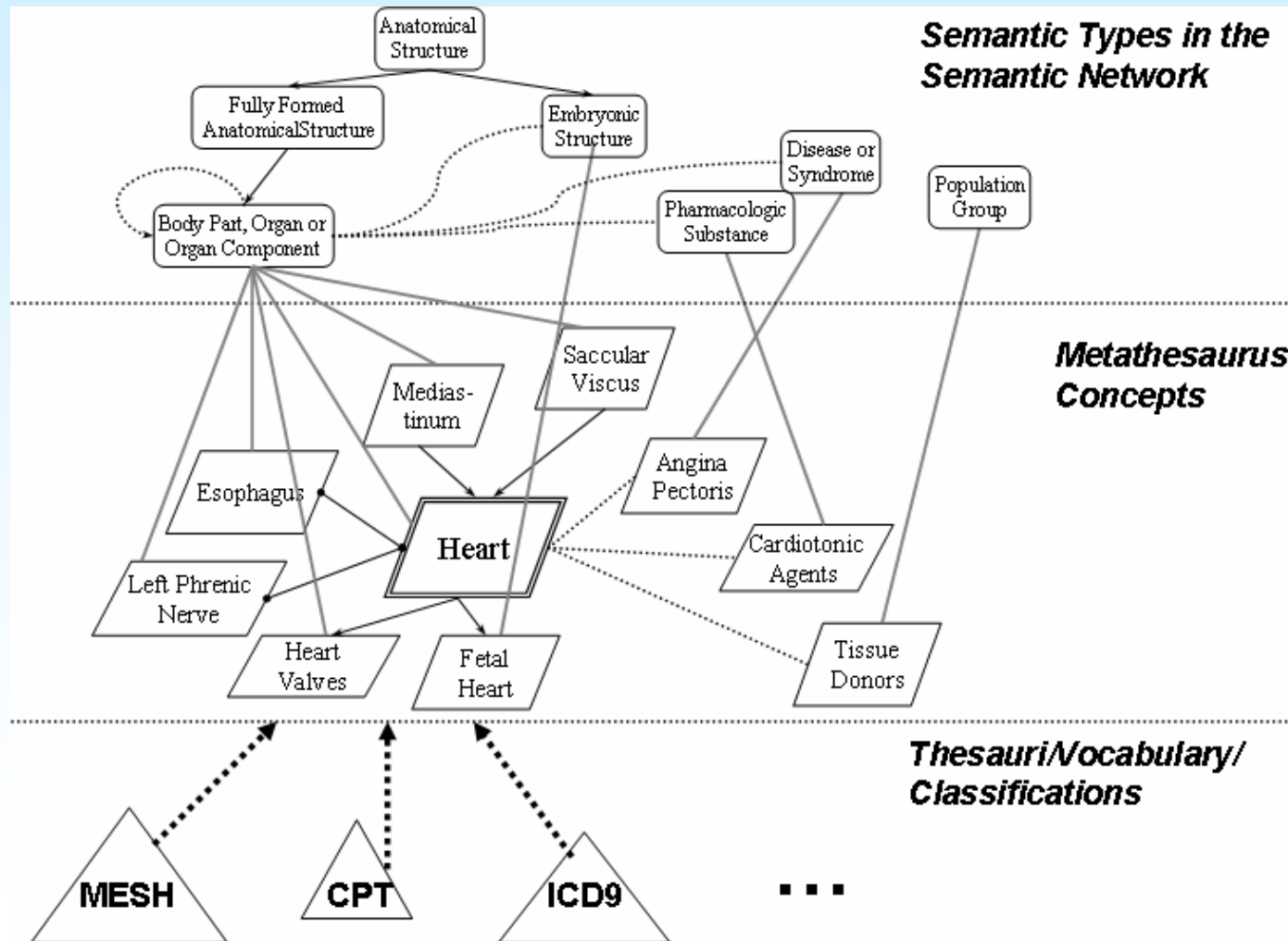
- Τρεις πηγές γνώσης:
 - ◆ *The Metathesaurus* : Περιέχει σημασιολογική πληροφορία για βιοιατρικές έννοιες, ονόματα και συσχετίσεις μεταξύ τους.
 - ☞ Περίπου 60.000 όπους από ανατομία, φάρμακα κ.λ.π.
 - ☞ Π.χ. Στομάχι, συκώτι κ.λ.π.
 - ☞ Οι όροι προέρχονται από MeSH, ICD κλπ.
 - ◆ *The Semantic Network* : Αποτελεί ένα δίκτυο γενικών κατηγοριών και σημασιολογικών τύπων για όλους τους όρους του *Metathesaurus*.
 - ☞ Είναι ένα πολυδέντρο που ξεκινά από γενικές κατηγορίες και αναλύεται,
 - ☞ Κάθε όρο του μεταθησαυρού ανήκει σε μία κατηγορία.
 - ◆ *The Specialist Lexicon* : Περιέχει συντακτική πληροφορία για βιοιατρικούς όρους.





UI	T110
STY	Steroid
Definition	One of a group of polycyclic, 17-carbon-atom, fused-ring compounds occurring both in natural and synthetic forms. Included here are naturally occurring and synthetic steroids, bufanolides, cardanolides, homosteroids, norsteroids, and secosteroids.
Examples	Bufanolides, Ketosteroids, Norandrostanes, Prednisone





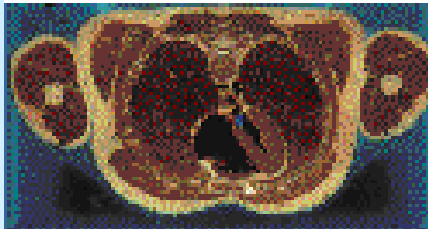
- Με την ανάπτυξη σύγχρονων μεθόδων τρισδιάστατης οπτικοποίησης είναι δυνατή η 3D αναπαράσταση των ανατομικών δομών του ανθρωπίνου σώματος.
- Οποιοδήποτε υποσύνολο από τα 3D αντικείμενα μπορεί να αναπαρασταθεί είτε μόνο του είτε ως μέλος μιας μεγαλύτερης ομάδας.
- Δημιουργείται η ανάγκη για ένα εργαλείο που να διευκολύνει την πρόσβαση στα 3D αντικείμενα.



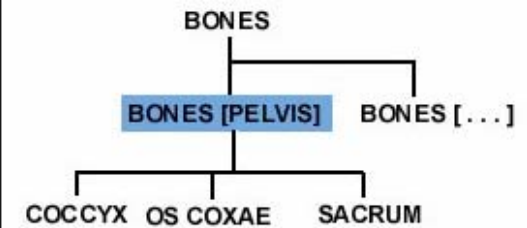
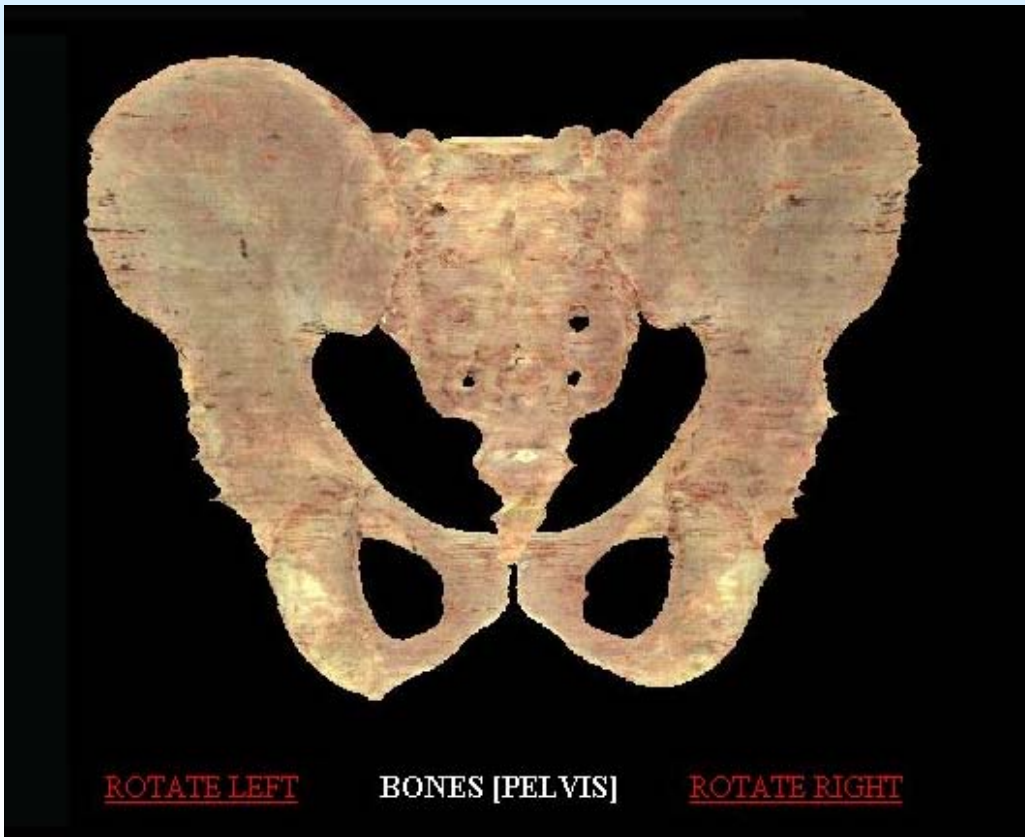
- Δημιουργήθηκε ένα **σύστημα πλοήγησης** σε 3D εικόνες ανατομίας.
- Αυτές συσχετίζονται με έννοιες ανατομίας τις οποίες συνδέει μια σημασιολογική οντολογία.

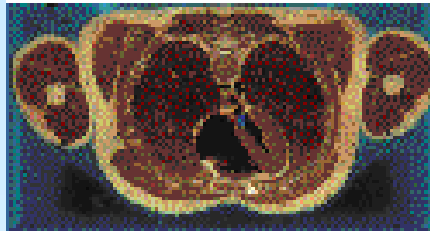


Τμήμα Τεχνολογίας Ιατρικών Οργάνων
Εργαστήριο Βιοϊατρικής Τεχνολογίας



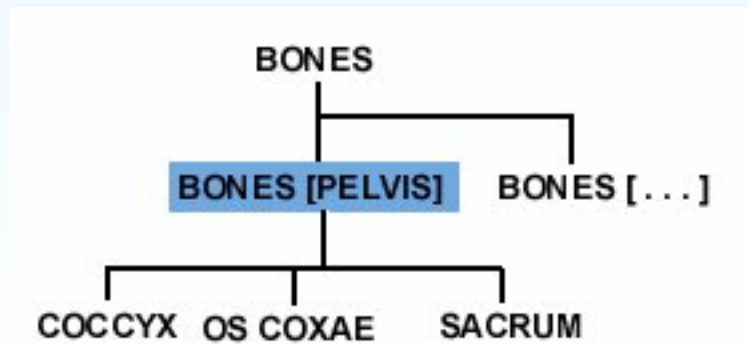
THE VISIBLE HUMAN PROJECT®



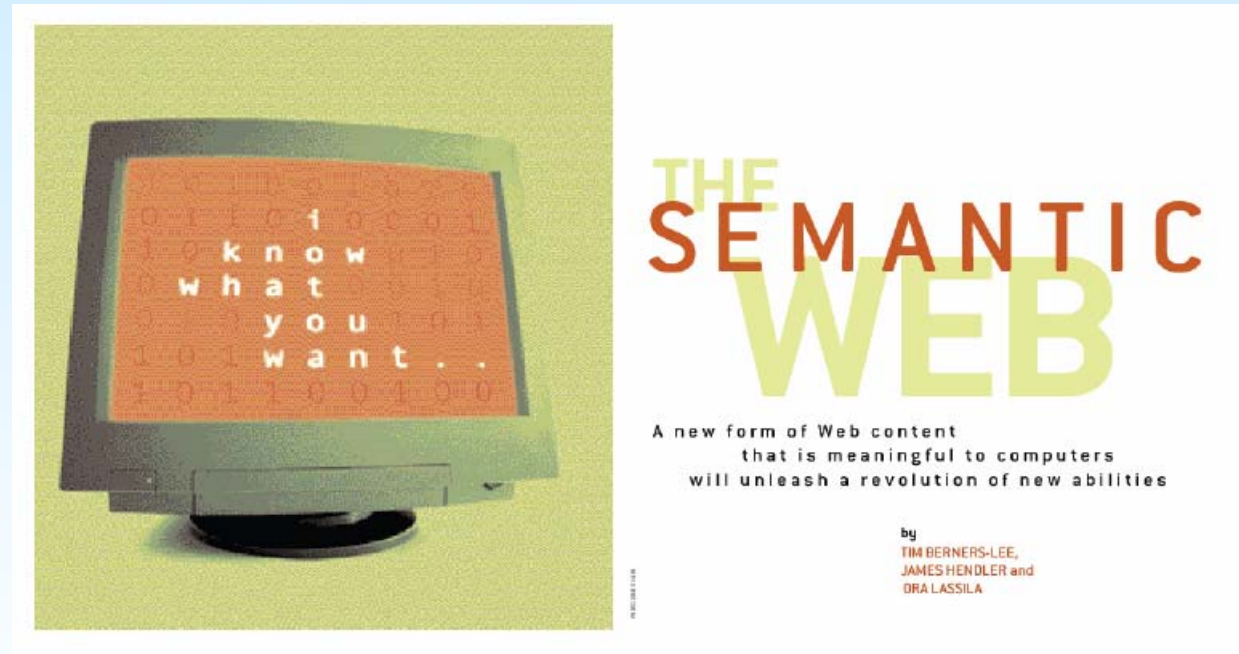


THE VISIBLE HUMAN PROJECT®

- Ανάλογα με το μέρος του σώματος που απεικονίζει η 3D εικόνα, παρουσιάζεται το αντίστοιχο τμήμα της οντολογίας.
 - ◆ Η λεκάνη και όλες οι υπόλοιπες κατηγορίες οστών είναι *IsA* της γενικότερης κλάσης *BONES*.
 - ◆ Τα οστά *COCCYX*, *OS COXAE*, *SACRUM* είναι *IsA* της κλάσης *BONES [PELVIS]*



Ο σημασιολογικός ιστός





Η Ιστορία του Σημασιολογικού Ιστού

- Ο Ιστός(Web) «εφευρέθηκε», μεταξύ άλλων, από τον Tim Berners-Lee, Φυσικό στο Κέντρο Πυρηνικών Μελετών και Ερευνών (CERN).
- Το πρωταρχικό όραμα του Tim Berners-Lee για τον Ιστό υπήρξε πολύ πιο φιλόδοξο από την πραγματικότητα του υπαρκτού σήμερα Συντακτικού Ιστού (Syntactic Web).
- Ο Tim Berners-Lee εργάζεται έκτοτε προς την πραγματοποίηση του οράματος, το οποίο έγινε ευρέως γνωστό με την ονομασία Σημασιολογικός Ιστός (Semantic Web).



Τι είναι ο σημασιολογικός ιστός

- Ο Σημασιολογικός Ιστός δεν είναι ένας ξεχωριστός ιστός αλλά η επέκταση του συντακτικού ιστού, στον οποίο η πληροφορία είναι καλά καθορισμένη ώστε να διευκολύνεται η συνεργασία ανθρώπων και υπολογιστών
- Ο Σημασιολογικός Ιστός
 - ◆ Θα είναι μία νέα μορφή του *Web content* του οποίου η σημασία είναι κατανοητή στους υπολογιστές
 - ◆ Θα οδηγήσει σε μία επανάσταση νέων τεχνολογιών και δυνατοτήτων για τους υπολογιστές
 - ◆ Θα είναι ένα σύνολο από πηγές και συνδέσεις προσβάσιμες και κατανοητές από υπολογιστές οι οποίοι θα μπορούν, εκμεταλλευόμενοι την κατανοητή σε αυτούς γνώση, να εξάγουν συμπεράσματα



Τι είναι ο Ιστός σήμερα; – Συντακτικός Ιστός

- Μια ψηφιακή βιβλιοθήκη
 - ◆ *Μια βιβλιοθήκη «εγγράφων» που λέγονται ιστοσελίδες συνδεδεμένα μεταξύ τους με hypermedia of links (hyperlinks -υπερσυνδέσεις)*
- Μια πλατφόρμα εφαρμογών
 - ◆ *Μια κοινή πύλη για εφαρμογές οι οποίες*
 - ☞ είναι προσβάσιμες μέσω ιστοσελίδων, και
 - ☞ εκθέτουν τα αποτελέσματά τους σε ιστοσελίδες
- Μια πλατφόρμα πολυμέσων
 - ◆ *Διαφημίσεις*
- Είναι ένα μέρος όπου οι άνθρωποι κάνουν τη «δύσκολη / σκληρή» δουλειά (διασύνδεση και ερμηνεία των πληροφοριών των πηγών) και οι υπολογιστές είναι υπεύθυνοι μόνο για την παρουσίαση.
- *Γιατί να μην κάνουν οι υπολογιστές τη «δύσκολη» δουλειά;*



Τι είναι αδύνατο στον συντακτικό ιστό;

- Πολύπλοκα ερωτήματα που απαιτούν background knowledge
 - ◆ πχ. *«εύρεση ζώων που χρησιμοποιούν sonar εκτός από τις νυχτερίδες και τα δελφίνια»*
- Εύρεση πληροφορίας σε data repositories
 - ◆ *Ταξιδιωτικές πληροφορίες*
 - ◆ *Τιμές από προϊόντα και υπηρεσίες*
- Ανάθεση πολύπλοκων εργασιών σε web agents
 - ◆ *Κράτηση πακέτου διακοπών σε μέρος με συγκεκριμένες απαιτήσεις πχ. Ψηλή θερμοκρασία, θάλασσα, αγγλική γλώσσα*



Ποιο είναι τελικά το πρόβλημα στον συντακτικό ιστό;

- Σε μία τυπική web page
 - ◆ *Η markup πληροφορία αποτελείται από πληροφορίες παρουσίασης (π.χ. μέγεθος γραμματοσειράς) και διασυνδέσεις σε σχετικό περιεχόμενο*
 - ◆ *Το σημασιολογικό περιεχόμενο είναι προσβάσιμο στους χρήστες, αλλά όχι στους υπολογιστές*
 - ◆ *Απαιτείται τουλάχιστον κατανόηση της φυσικής γλώσσας*
- Σε μια τυπική Μηχανή Αναζήτησης
 - ◆ *Η χρήση λέξεων κλειδιών περιορίζει σημαντικά τις δυνατότητες εύρεσης σελίδων.*
 - ◆ *Όταν δεν υπάρχει μια λέξη κλειδί σε μία σελίδα τότε είναι δύσκολο για μια μηχανή να την ανακτήσει.*
- Έλλειψη κοινής ορολογίας, οργάνωσης και σημασιολογίας στα δεδομένα που διακινούνται στον Ιστό.
 - ◆ *Δεν μπορούμε να υποχρεώσουμε τους χρήστες του Ιστού να οργανώνουν τα δεδομένα τους όλοι με τον ίδιο τρόπο.*



Η Λύση είναι ο Σημασιολογικός Ιστός

- Ο Σημασιολογικός Ιστός στοχεύει στον εμπλουτισμό του Ιστού με σημασιολογική πληροφορία:
 - ◆ *Τα δεδομένα να είναι κατανοητά σε επίπεδο μηχανής ώστε να διευκολύνεται η επεξεργασία και αναζήτησή τους*
 - ◆ *Να υπάρχει κοινή ορολογία, οργάνωση και σημασιολογία στα δεδομένα που διακινούνται στον Ιστό, στο βαθμό που αυτό είναι εφικτό*
 - ◆ *Να υπάρχει η ευελιξία σύνθεσης της σημασιολογικής πληροφορίας και δημιουργία νέας*



Σημασιολογικός Εμπλουτισμός Ιστού

- Προσθήκη σημασιολογικού σχολιασμού / markup στις ιστοσελίδες
- Η σημασιολογία δίνεται χρησιμοποιώντας οντολογίες
 - ◆ *Οι οντολογίες παρέχουν το λεξιλόγιο των όρων που χρησιμοποιούνται στο σχολιασμό*
 - ◆ *Νέοι όροι μπορούν να σχηματιστούν από το συνδυασμό υπαρχόντων*
 - ◆ *Η Σημασιολογία (το νόημα) αυτών των όρων είναι τυπικά ορισμένη*
- Μία βασική απαίτηση είναι η ύπαρξη μίας standard web ontology language



Το μυστικό είναι τα μεταδεδομένα

- **Μεταδεδομένα: Δεδομένα για τα δεδομένα**
 - ◆ *Κάνουν σαφή την πληροφορία που είναι αόριστη*
 - ◆ *Επιτρέπουν την αναζήτηση, επεξεργασία και ενοποίηση της πληροφορίας*
 - ◆ *Μεταδεδομένα περιγραφών: περιγράφουν το περιεχόμενο και τη σημασία των πηγών (προϋποθέτουν την ύπαρξη κοινής γλώσσας ή γλώσσας αναφοράς)*
- **Οντολογίες**
 - ◆ *Διαμοιραζόμενα και κοινά λεξιλόγια*
 - ◆ *Χρήση σε μηχανές εύρεσης και πράκτορες*
 - ◆ *Διαμοιραζόμενη και κοινή κατανόηση ενός πεδίου ενδιαφέροντος*
 - ◆ *Απαραίτητες για ανταλλαγή και αναζήτηση*
- **Συμπερασματολογία**
 - ◆ *Εξαγωγή καινούργιας γνώσης και μεταδεδομένων από την υπάρχουσα γνώση*



Πού βρισκόμαστε σήμερα?

- Καθορίζονται οι γλώσσες που θα πλαισιώσουν το σημασιολογικό ιστό
- Καθορίζεται και αναπτύσσεται η αρχιτεκτονική και τα εργαλεία που θα χρησιμοποιηθούν στο σημασιολογικό ιστό
 - ◆ *Επόμενο βήμα η ανάπτυξή τους*
- Χρήση των εργαλείων και των γλωσσών για να σχολιάσουν το υλικό στις διάφορες εφαρμογές.



Οντολογία

- Κλάδος της φιλοσοφίας που ασχολείται με τη φύση και την οργάνωση των όντων
 - ◆ *Η επιστήμη της ύπαρξης (Αριστοτέλης, Μεταφυσικά, IV, 1)*
- Στην επιστήμη και στην τεχνητή νοημοσύνη
 - ◆ *Οι οντολογίες είναι εργαλεία αναπαράστασης (representation) και συλλογιστικής (reasoning).*
 - ◆ *Μια ρητή προδιαγραφή της αντίληψης για τον κόσμο*
“An explicit specification of a conceptualisation” (Gruber)
- Περιγράφει ένα πεδίο γνώσης (knowledge domain) ώστε
 - ◆ *Να υπάρχει κοινός τρόπος επικοινωνίας (ορολογία / λεξιλόγιο) μεταξύ όλων αυτών που χρησιμοποιούν το συγκεκριμένο πεδίο γνώσης.*
 - ◆ *Να είναι αυστηρά θεμελιωμένη (formal model).*
 - ◆ *Να είναι κατανοητή και επεξεργάσιμη (machine understandable, manipulable) σε επίπεδο μηχανής.*



Οντολογία

- Κάθε Ontology έχει
 - ◆ *Classes*
 - ☞ Subclasses
 - Individual Instances
- Δομή:
 - ◆ *Κάθε Subclass is a “kind of” Class.*
 - ◆ *Κάθε Individual Instance είναι ένα “τελικό προϊόν”*



Οντολογία

- Κάθε Class και Subclass χαρακτηρίζεται από:
 - ◆ *Slots (=roles/properties)*. Περιγράφουν τα *features* και τα *attributes* κάθε *class* και κάθε *subclass*.
 - ◆ Κάθε *Subclass* κληρονομεί τα *Slots* του *Class* στο οποίο ανήκει
 - ◆ Σε κάθε *Subclass* προστίθενται τα *Slots* τα οποία ανήκουν αποκλειστικά σε αυτό.
 - ◆ Κάθε *Individual Instance* κληρονομεί αυτόματα τα *Slots* του *Subclass* και του *Class* στα οποία ανήκει.
- Κάθε *Slot* περιγράφεται από τα αντίστοιχα:
 - ◆ *Facets (=role restrictions)*. Περιγράφουν τους ορισμούς / περιορισμούς του *Slot*.
 - ◆ Παράδειγμα: *value type, allowed values, the number of the values*



Παράδειγμα Οντολογίας

<i>Ontology : Wine</i>	
Class	Wine
Subclass 1	White wine
Subclass 2	Rose wine
Subclass 3	Red wine
Subclass 1	Beaujolais
Subclass 2	Red Burgundy
Subclass 3	Red Bordeaux



Παράδειγμα Οντολογίας (συνέχεια)

- Το Class Wine έχει τα εξής Slots (features και attributes):

• Color	
• Flavor	
• Grape	Από τι είδους σταφύλια προέρχεται
• Maker	Ποιος το παρασκευάζει
• Region	Σε ποια περιοχή
• Name	Ονομασία κρασιού
• Sugar	Περιεκτικότητα σε ζάχαρη

- Τα χαρακτηριστικά αυτά αφορούν σε όλα τα κρασιά, οπότε είναι Slots του Class Wine.



Παράδειγμα Οντολογίας (συνέχεια)

- Τα facets (role restrictions) των slots είναι:

<i>Name</i>	<i>Type</i>	<i>Cardinality</i>	<i>Other facets</i>
Color	Symbol	Single	Allowed-values={White, Rose, Red}
Flavor	Symbol	Single	Allowed-values={Delicate, Moderate, Strong}
Grape	Instance	Multiple	Classes={Wine-grape}
Maker	Instance	Single	Classes={Winery}
Region	Instance	Single	Classes={Wine-Region}
Name	String	Single	
Sugar	Symbol	Single	Allowed-values={Dry, Sweet, Off Dry}

- **** Όταν ένα slot είναι τύπου *Instance*, παίρνει τις τιμές του από *Individual Instances* άλλων *Classes*.**



Παράδειγμα Οντολογίας (συνέχεια)

- Το Subclass3 (Red Wine) κληρονομεί όλα τα slots (και τα facets) του class στο οποίο ανήκει, αλλά έχει και ένα δικό του slot.
 - ◆ *Tannin level*
- *Τα facets αυτού του slot είναι:*

<i>Name</i>	<i>Type</i>	<i>Cardinality</i>	<i>Other facets</i>
Tannin level	Symbol	Single	Allowed-values={Low, Moderate, High}



Παράδειγμα Οντολογίας (συνέχεια)

- Έστω ότι θέλουμε να ορίσουμε ένα Individual Instance (ένα συγκεκριμένο κρασί), το οποίο είναι κόκκινο. Επομένως, πρέπει να συμπληρώσουμε τα εξής slots:

<i>Name</i>	•Chateau-Morgon-Beaujolais
<i>Color</i>	•Red
<i>Flavor</i>	•Delicate
<i>Tannin level</i>	•Low
<i>Grape</i>	•Gamay (instance of the Wine-grape class)
<i>Maker</i>	•Chateau-Morgon (instance of the Winery class)
<i>Region</i>	•Beaujolais (instance of the Wine-Region class)
<i>Sugar</i>	•Dry



Τέλος ...

Σας ευχαριστώ για την υπομονή σας.

