



ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ ΔΥΤΙΚΗΣ ΕΛΛΑΔΑΣ
ΣΧΟΛΗ ΕΠΑΓΓΕΛΜΑΤΩΝ ΥΓΕΙΑΣ ΚΑΙ ΠΡΟΝΟΙΑΣ
ΤΜΗΜΑ ΦΥΣΙΚΟΘΕΡΑΠΕΙΑΣ

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

**ΑΠΟΤΕΛΕΣΜΑΤΙΚΕΣ ΜΕΘΟΔΟΛΟΓΙΕΣ
ΔΙΕΡΕΥΝΗΣΗΣ ΙΑΤΡΙΚΩΝ ΒΑΣΕΩΝ ΔΕΔΟΜΕΝΩΝ**

ΚΑΝΕΛΛΑΚΗ ΑΡΤΕΜΙΣ

ΣΠΑΝΟΥ-ΧΑΖΕΛΜΠΕΡΓΚΕΡ ΕΥΑ-ΜΑΡΙΑ

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ: ΚΟΥΤΣΟΓΙΑΝΝΗΣ ΚΩΝΣΤΑΝΤΙΝΟΣ

ΑΙΓΙΟ -2019

ΠΡΟΛΟΓΟΣ

Με την ολοκλήρωση της πτυχιακής μας εργασίας θα θέλαμε να ευχαριστήσουμε όλους όσους μας βοήθησαν σε αυτή μας την προσπάθεια, τους καθηγητές μας, τους συμφοιτητές μας που συμμετείχαν στην έρευνα μας και κυρίως τους γονείς μας που μας στήριξαν με κάθε τρόπο καθ' όλη την διάρκεια των σπουδών μας.

Ξεχωριστές ευχαριστίες στην Ομάδα Τεχνητής Νοημοσύνης (Καθ. Ιωάννη Χατζηλυγεροδη, Δρ Φωτεινή Γριβοκωστοπούλου και Δρ Ισίδωρο Περρίκο) του MMLAB (εργαστήριο Γραφικών και Πολυμέσων) του τμήματος Μηχανικών ΗΥ και Πληροφορικής της Πολυτεχνικής Σχολής του Πανεπιστημίου Πατρών που έδωσε την δυνατότητα της χρήσης του συστήματος Biosite το οποίο σχεδιάστηκε και υλοποιήθηκε στα πλαίσια της μεταπτυχιακής εργασίας της μεταπτυχιακής φοιτήτριας Μηνά Ιωάννας το 2017, από την οποία και χρησιμοποιήθηκε το τεχνικό κομμάτι περιγραφής του συστήματος στα κεφάλαια 3 και 4.

ΠΕΡΙΛΗΨΗ

Η πτυχιακή μας εργασία αναφέρεται στην αξιολόγηση ενός ήδη ολοκληρωμένου συστήματος που περιλαμβάνει την συλλογή κειμένων βιοϊατρικού περιεχομένου. Έτσι το σύστημα αυτό περιλαμβάνει την ευκαιρία συλλογής άρθρων από διαδικτυακές πηγές, την ανάλυσή τους με εργαλεία επεξεργασίας φυσικής γλώσσας και τον προσδιορισμό των εννοιών που παρουσιάζονται, με σκοπό την υποστήριξη ερωτημάτων από χρήστες.

Για τους παραπάνω λόγους πραγματοποιήθηκε αξιολόγηση στο σύστημα BioSite που αποτελείται από έξι υποσυστήματα ώστε η διαδοχική εκτέλεση τους να συγκεντρώνει κείμενα από την PubMed Central βιβλιοθήκη, τα επεξεργάζεται και τα αποθηκεύει σε τοπική βάση.

Το BioSite αναλύει τα άρθρα με τη βοήθεια εργαλείων επεξεργασίας φυσικής γλώσσας, με προσανατολισμό την ανάλυση τους ως προς το βιοϊατρικό περιεχόμενο. Αφού γίνει η κατάλληλη ανάλυση, εισάγει την πληροφορία για τις έννοιες που προσδιορίστηκαν με κατάλληλη μορφή στην οντολογία του συστήματος, που αποτελεί συνένωση και επέκταση οντολογιών του βιοϊατρικού τομέα.

Στην συνέχεια εφαρμόζει αλγόριθμους κατηγοριοποίησης των άρθρων, για την ταξινόμηση τους σε ομάδες ανάλογα με το περιεχόμενό τους. Όλη η πληροφορία που προκύπτει, χρησιμοποιείται για την υποστήριξη αναζητήσεων άρθρων και απάντησης σε ερωτήματα χρηστών μέσω της ιστοσελίδας του συστήματος.

Οι χρήστες χρησιμοποιώντας την ιστοσελίδα του συστήματος μπορούν να αναζητήσουν άρθρα, να βρουν τα άρθρα του συγγραφέα που τους ενδιαφέρει ή άρθρα συγκεκριμένης ημερομηνίας έκδοσης. Μπορούν να θέσουν ερωτήματα σε φυσική γλώσσα ή σε βιοϊατρικούς όρους όπως σε οποιαδήποτε βάση δεδομένων όπως αυτής της Google και να πάρουν ως απαντήσεις τα σχετικά άρθρα που συνάδουν με το ερώτημα που έθεσαν. Μπορούν επίσης να αναζητήσουν όμοια με το άρθρο που τους ενδιαφέρει άρθρα, όσον αφορά το περιεχόμενό τους από βιοϊατρική σκοπιά.

Οι 400 σπουδαστές που συμμετείχαν στην παρούσα έρευνα για την αξιολόγηση της ευχρηστίας και λειτουργίας του πληροφοριακού συστήματος, εξοικειώθηκαν πολύ σύντομα στη χρήση του αφού είχαν και τις κατάλληλες γνώσεις διερεύνησης ιατρικών βάσεων δεδομένων που διδάσκονται στα πλαίσια των μαθημάτων τους που αφορούν την ιατρική πληροφορική και την μεθοδολογία έρευνας. Το τελικό σύστημα με την εφαρμογή όλων των υποσυστημάτων του όπως χρησιμοποιήθηκε από δύο διαδοχικά έτη σπουδών των φοιτητών του τμήματος Φυσιοθεραπείας και τέσσερα συνολικά εξάμηνα φοίτησης, αποδείχθηκε ότι καλύπτει όλες τις απαιτήσεις που τέθηκαν, όπως φαίνεται και από την αξιολόγηση της λειτουργίας του.

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΡΟΛΟΓΟΣ	2
ΠΕΡΙΛΗΨΗ	3
ΠΕΡΙΕΧΟΜΕΝΑ	5
ΕΙΣΑΓΩΓΗ	7
ΓΕΝΙΚΟ ΜΕΡΟΣ	12
ΚΕΦΑΛΑΙΟ 1	12
1.1 ΕΦΑΡΜΟΓΕΣ ΤΩΝ ΤΕΧΝΟΛΟΓΙΩΝ ΤΗΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΟΥ ΔΙΑΔΙΚΤΥΟΥ ΣΤΟ ΧΩΡΟ ΤΗΣ ΥΓΕΙΑΣ	12
1.2 ΤΙ ΕΙΝΑΙ ΤΟ INTERNET	13
1.3 ΠΟΙΕΣ ΥΠΗΡΕΣΙΕΣ ΠΡΟΣΦΕΡΕΙ ΤΟ INTERNET	13
1.4 Η ΣΥΝΔΕΣΗ ΜΕ ΤΟ INTERNET	17
1.5 Η ΙΑΤΡΙΚΗ ΠΛΗΡΟΦΟΡΙΑ ΣΤΟ INTERNET	17
1.6 ΑΞΙΟΠΟΙΗΣΗ ΤΟΥ INTERNET ΑΠΟ ΤΗΝ ΙΑΤΡΙΚΗ ΚΟΙΝΟΤΗΤΑ	20
ΚΕΦΑΛΑΙΟ 2	22
2.1 ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ	22
2.2 ΙΑΤΡΙΚΕΣ ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ	26
2.2.1 ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ ΗΛΕΚΤΡΟΝΙΚΗΣ ΑΝΑΖΗΤΗΣΗΣ	26
2.3 ΠΡΟΒΟΛΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ ΑΝΑΖΗΤΗΣΗΣ	31
ΚΕΦΑΛΑΙΟ 3	36
3.1 ΣΥΣΤΗΜΑ BioSite	36
3.2 Υποσύστημα CreateBase	40
3.3 Υποσύστημα Crawler	53
3.4 Υποσύστημα GateProcess	57
3.5 Υποσύστημα OntologyProcess	60
3.6 Clustering Process	63
BioSite Interface	67
ΚΕΦΑΛΑΙΟ 4	71

Παρουσίαση του συστήματος BioSite	71
ΕΙΔΙΚΟ ΜΕΡΟΣ	87
ΚΕΦΑΛΑΙΟ 5	87
5.1 Σχεδιασμός της έρευνας	87
A. ΑΠΟΤΕΛΕΣΜΑΤΑ	89
5.2.1 ΠΕΡΙΓΡΑΦΙΚΗ ΚΑΤΑΝΟΜΗ ΤΟΥ ΔΕΙΓΜΑΤΟΣ	89
5.2.2 ΔΗΜΟΓΡΑΦΙΚΑ ΣΤΟΙΧΕΙΑ	89
ΚΕΦΑΛΑΙΟ 6	101
6.1 ΓΕΝΙΚΑ ΣΥΜΠΕΡΑΣΜΑΤΑ	101
6.2 ΣΥΜΠΕΡΑΣΜΑΤΑ ΜΕ ΒΑΣΕΙ ΤΑ ΣΤΑΤΙΣΤΙΚΑ	101
ΒΙΒΛΙΟΓΡΑΦΙΑ	103
ΠΑΡΑΡΤΗΜΑ Ι	106

ΕΙΣΑΓΩΓΗ

Για την περιγραφή του όρου «επιστημονική έρευνα» έχουν δοθεί αρκετοί ορισμοί με την πάροδο των χρόνων, σημαντικότεροι είναι οι εξής (Λαγουμιντζής Γ. και συν 2015):

- Έρευνα είναι μία συστηματική και καλώς σχεδιασμένη διαδικασία για την επίλυση προβλημάτων με βάση την εμπειρική πραγματικότητα.
- Έρευνα είναι η συστηματική χρήση επιστημονικής γνώσης, βάσει συγκεκριμένου σκοπού, για την βελτίωση επιτευγμάτων του ανθρώπου, έστω και αν μερικές από αυτές τις δραστηριότητες δεν τυγχάνουν καθολικής αποδοχής.

Σκοπός της έρευνας είναι να προσδιορίσει πώς είναι κάποια δεδομένα στην πραγματικότητα σε σύγκριση με το πώς θα έπρεπε να είναι ιδανικά. Για την επίτευξη του σκοπού αυτού η έρευνα θα πρέπει να προϋποθέτει προσεκτικό και συστηματικό τρόπο επίλυσης προβλημάτων, περιλαμβάνοντας τα προκείμενα χαρακτηριστικά (Λαγουμιντζής Γ. και συν 2015):

1. Συστηματική → Η επίλυση των προβλημάτων επιτυγχάνεται μέσω του εντοπισμού και της ονομασίας ορισμένων μεταβλητών και ακολουθείται από το σχεδιασμό της έρευνας που εξετάζει τις σχέσεις μεταξύ αυτών. Μετέπειτα συλλέγονται τα δεδομένα που επιτρέπουν την αξιολόγηση του προβλήματος και των ερευνητικών υποθέσεων.

2. Λογική → Η εξέταση των διαδικασιών που χρησιμοποιούνται επιτρέπει στους ερευνητές την αξιολόγηση των συμπερασμάτων.
3. Εμπειρική → Ο ερευνητής συλλέγει ορισμένα δεδομένα στα οποία βασίζει τις αποφάσεις του.
4. Αναγωγική → Ο ερευνητής συλλέγει δεδομένα από διάφορα (ξεχωριστά) γεγονότα, ώστε να τα χρησιμοποιήσει για την απόδειξη γενικών σχέσεων.
5. Επαναληπτική → Η διαδικασία καταγράφεται με τρόπο που επιτρέπει σε άλλους ερευνητές να εξετάσουν τα ευρήματα επαναλαμβάνοντας την έρευνα ή να σχεδιάσουν μία μελλοντική έρευνα βασισμένη στα αποτελέσματα αυτής.

Στην προκείμενη εργασία θα ασχοληθούμε με το **βιβλιογραφικό σκέλος** της εκάστοτε έρευνας. Η διερεύνηση της βιβλιογραφίας αποτελεί ένα από τα σημαντικότερα στάδια της διαδικασίας μίας έρευνας, ιδιαίτερα εκπαιδευτικής, και είναι ουσιώδες μέρος της ερευνητικής αναφοράς. Η ανασκόπηση ερευνών και άρθρων σχετικών με το θέμα που έχει επιλέξει ο φοιτητής πρέπει να οδηγεί στην ερευνητική ερώτηση με βάση την καλύτερη διαθέσιμη πληροφόρηση. Έτσι, θα καταφέρει να διεκπεραιώσει μία ποιοτική έρευνα, η οποία θα χαρακτηρίζεται από εγκυρότητα και αξιοπιστία (Λαγουμιντζής Γ. και συν 2015).

Όλα τα είδη πειραματικών σχεδιασμών μίας έρευνας έχουν πλεονεκτήματα και μειονεκτήματα που δημιουργούν απειλές για την εγκυρότητά της. Βασική αρχή είναι η διάκριση μεταξύ εσωτερικής κι εξωτερικής εγκυρότητας. Η εσωτερική εγκυρότητα θέτει το ερώτημα αν κάνουν στην πραγματικότητα διαφορά στο συγκεκριμένο σχεδιασμό οι πειραματικοί χειρισμοί, ενώ η εξωτερική εγκυρότητα αφορά το ερώτημα σε ποιους πληθυσμούς, περιβάλλον ή μεταβλητές μπορεί να γενικευτεί το αποτέλεσμα που θα προκύψει. Ωστόσο, τα δύο αυτά είδη εγκυρότητας συνήθως έρχονται σε σύγκρουση, καθώς η εξασφάλιση της εσωτερικής περιλαμβάνει τον έλεγχο όλων των μεταβλητών έτσι ώστε ο ερευνητής να μπορέσει να αποκλείσει όλες τις ανταγωνιστικές υποθέσεις που ενέχονται στην ερμηνεία των αποτελεσμάτων που παρατηρούνται. Με

αυτόν τον τρόπο θέτεται σε κίνδυνο η γενίκευση των ευρημάτων που περιλαμβάνει η εξωτερική εγκυρότητα. Το φαινόμενο αυτό παρατηρείται διότι το οικολογικά έγκυρο περιβάλλον δεν υπάρχει ο απόλυτος έλεγχος των ερεθισμάτων όπως στο χώρο ενός εργαστηρίου (Λαγουμιντζής Γ. και συν 2015).

ΑΠΕΙΛΕΣ ΤΗΣ ΕΣΩΤΕΡΙΚΗΣ ΕΓΚΥΡΟΤΗΤΑΣ

- I. Ιστορικό, γεγονότα που συμβαίνουν κατά τη διάρκεια του πειράματος χωρίς να αποτελούν μέρος των χειρισμών
- II. Ωρίμανση, διαδικασίες που επηρεάζουν αποκλειστικά τα υποκείμενα που σχετίζονται άμεσα με την επίδραση του χρόνου (γήρανση, κόπωση, πείνα)
- III. Αρχική μέτρηση, επιδράσεις του test στις επόμενες μετρήσεις του ίδιου test
- IV. Επίδραση οργάνων μέτρησης, αλλαγές στην βαθμονόμηση ενός οργάνου, συμπεριλαμβανομένης και της έλλειψης συμφωνίας ενός και μεταξύ των παρατηρητών
- V. Στατιστική παλινδρόμηση, το γεγονός ότι οι ομάδες επιλέγονται με βάση ακραίων τιμών, που δεν είναι τόσο ακραίες σε μια επακόλουθη δοκιμασία
- VI. Μεροληψία στην επιλογή υποκειμένων, καθορισμός ομάδων σύγκρισης με άλλη διαδικασία αντί της τυχαίας
- VII. Πειραματική θνησιμότητα, απώλεια υποκειμένων, απ' τις εξεταζόμενες ομάδες που δεν οφείλεται σε τυχαίους λόγους
- VIII. Αλληλεπίδραση επιλογής-ωρίμανσης, παρατηρείται συχνά σε μη ισοδύναμο σχεδιασμό των ομάδων κατά τον οποίο η πάροδος του χρόνου μπορεί να επηρεάσει τη μία ομάδα και όχι την άλλη

→ Επιπρόσθετη απειλή εντοπισμένη από την Rosenthal, 1966: η προσδοκία αναφέρεται στους ερευνητές που αναμένουν ότι ορισμένα υποκείμενα θα έχουν καλύτερη επίδοση.

ΑΠΕΙΛΕΣ ΤΗΣ ΕΞΩΤΕΡΙΚΗΣ ΕΓΚΥΡΟΤΗΤΑΣ

- i. Επίδραση ή αλληλεπίδραση της αρχικής μέτρησης
- ii. Αλληλεπίδραση μεταξύ μεροληψίας στην επιλογή των υποκειμένων και του πειραματικού χειρισμού, όταν επιλέγεται μία ομάδα βάσει ενός συγκεκριμένου χαρακτηριστικού
- iii. Επιδράσεις απ' τις πειραματικές διαδικασίες, δηλαδή πειραματικοί χειρισμοί που είναι αποτελεσματικοί σε τεχνητό περιβάλλον ή συνθήκες μπορεί να μην είναι αποτελεσματικοί σε μη-πειραματικές συνθήκες
- iv. Εφαρμογή πολλαπλών πειραματικών χειρισμών, όταν τα υποκείμενα υπόκεινται σε περισσότερους από έναν οι επιδράσεις των προηγούμενων επηρεάζουν τους επόμενους

Όλες οι έρευνες πρέπει να περιλαμβάνουν τις έννοιες της εσωτερικής κι εξωτερικής εγκυρότητας (Λαγουμιντζής Γ. και συν 2015):

- **πόση εμπιστοσύνη** έχετε στην ποιότητα της περιγραφής και την ερμηνεία των συμβάντων στο συγκεκριμένο περιβάλλον έρευνας; (εσωτερική εγκυρότητα)

- **ποια είναι η εκτίμησή σας** για την έκταση στην οποία μπορούν τα αποτελέσματα αυτής της μελέτης να εφαρμοστούν σε ένα διαφορετικό περιβάλλον με το οποίο είστε εξοικειωμένοι; (εξωτερική εγκυρότητα)

Τελικά καταλήγουμε σε πέντε είδη εγκυρότητας στις ποιοτικές μεθόδους πραγματοποίησης έρευνας:

1. *Περιγραφική*: η τεκμηριωμένη ακρίβεια του απολογισμού, η οποία δεν είναι κατασκευασμένη, επιλεκτική ή διαστρεβλωμένη. Υπό αυτή την έννοια η αξιοπιστία εντάσσεται στο πλαίσιο της εγκυρότητας.
2. *Ερμηνευτική*: η ικανότητα της έρευνας να συλλάβει το νόημα, τις ερμηνείες, τους όρους, τις προθέσεις των καταστάσεων και γεγονότων, δηλαδή τα δεδομένα, για τους ίδιους τους συμμετέχοντες.

3. *Θεωρητική*: οι θεωρητικές κατασκευές που φέρνει ο ερευνητής στην ερευνητική διαδικασία (η θεωρία μπορεί να ιδωθεί ως εξήγηση).
4. *Γενικευσιμότητα*: η άποψη ότι η θεωρία που διατυπώνεται ενδεχομένως να είναι χρήσιμη στην κατανόηση κι άλλων χρήσιμων καταστάσεων.
5. *Αξιολογική*: η εφαρμογή ενός αξιολογικού πλαισίου, το οποίο έχει μία κριτική διάθεση απέναντι σε αυτό που ερευνάται και είναι λιγότερο περιγραφικό, επεξηγηματικό και ερμηνευτικό (Λαγουμιντζής Γ. και συν 2015).

Σκοπός της εργασίας αυτής είναι η διευκόλυνση των ερευνητών-φοιτητών στην αναζήτηση βιβλιογραφίας για την διεκπεραίωση μίας έρευνας. Με την πάροδο των χρόνων έχουν δημιουργηθεί αμέτρητες βάσεις δεδομένων στις οποίες μπορεί να απευθυνθεί ο καθένας. Τί είναι όμως μία βάση δεδομένων; Για το λόγο αυτό στα επόμενα κεφάλαια δημιουργούμε όλο το γνωστικό πλαίσιο το οποίο απαιτείται για μια σύγχρονη βιβλιογραφική έρευνα μέσω Η/Υ και με τη βοήθεια του σημαντικότερου εργαλείου των τελευταίων ετών στη πληροφορική, αυτό του **internet**, με έμφαση στο χώρο της υγείας και γενικότερα των βιοϊατρικών επιστημών.

ΓΕΝΙΚΟ ΜΕΡΟΣ

ΚΕΦΑΛΑΙΟ 1

1.1 ΕΦΑΡΜΟΓΕΣ ΤΩΝ ΤΕΧΝΟΛΟΓΙΩΝ ΤΗΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΟΥ ΔΙΑΔΙΚΤΥΟΥ ΣΤΟ ΧΩΡΟ ΤΗΣ ΥΓΕΙΑΣ

Στις εφαρμογές της Πληροφορικής συγκαταλέγονται το Internet, η Βιοπληροφορική και η Τηλεματική – Τηλεϊατρική.

Το Παγκόσμιο διαδίκτυο Ηλεκτρονικών Υπολογιστών, γνωστό ως Internet ανοίγει νέους ορίζοντες στην πρόσκληση επιστημονικής γνώσης, αλλά και μεταφέρει την ευθύνη της επιλογής στον ίδιο το χρήστη. Παράλληλα, το Internet προσφέρει ένα νέο εργαλείο με μεγάλες προοπτικές αξιοποίησης στον τομέα της Συνεχιζόμενης Ιατρικής Εκπαίδευσης.

Με το ψηφιακό αυτό δίκτυο μεταφέρονται σε χρόνο μηδέν σε όλο τον κόσμο πληροφορίες που δεν έχουν μόνο το χαρακτήρα κειμένου και σταθερής εικόνας, αλλά επεκτείνονται σε μορφές που δεν μπορούν να αναπαραχθούν σε έντυπα, όπως η κινητή εικόνα-video ή η φωνή και γενικά, ο ήχος. Επιπλέον, η σημερινή τεχνολογία κάνει προσιτές μέσω του Internet νέες υπηρεσίες όπως η videoconference, το video-text κ.α.

Το Internet, το Παγκόσμιο διαδίκτυο Ηλεκτρονικών Υπολογιστών, που απλώνεται σε όλες τις χώρες της Υφηλίου δεν είναι πια άγνωστο όπως μερικά χρόνια πριν. Σήμερα οι τεράστιες ποσότητες της ψηφιακής πληροφορίας που διακινούνται στο Internet, αφορούν όλες τις ανθρώπινες δραστηριότητες και όλα τα επαγγέλματα. Η Ιατρική πληροφορία, η πληροφορία που αφορά κάθε τομέα της Ιατρικής, που αφορά τον κάθε ιατρό κάθε ειδικότητας, τον κάθε ιατρό κάθε ειδικότητας, τον κάθε λειτουργό υγείας, τα διαφορετικά Συστήματα Υγείας, διακινείται μέσα από το Internet σε ελάχιστο χρόνο και σε οποιοδήποτε σημείο της υφηλίου (Μάργαρης και συν 2015).

Η ύπαρξη και η ραγδαία ανάπτυξη του Internet με τις προσφερόμενες υπηρεσίες, δημιουργούν μια νέα κατάσταση στον τρόπο και τις μορφές επικοινωνίας μεταξύ των μελών της Παγκόσμιας Ιατρικής κοινότητας, που ανατρέπει τα σημερινά δεδομένα. Το Internet εμφανίζεται σαν «**εργαλείο στα χέρια κάθε ιατρού**» και επιδρά στη διαμόρφωση νέων συνθηκών απόκτησης και επεξεργασίας εξειδικευμένης επιστημονικής γνώσης, προσιτής στον κάθε ενδιαφερόμενο. Οι συνθήκες επιτρέπουν την ταχεία ποιοτική αναβάθμιση της Συνεχιζόμενης Ιατρικής Εκπαίδευσης, Ενημέρωσης και Επικοινωνίας, με ό,τι αυτό συνεπάγεται. Για πρώτη φορά αναπτύσσονται προβληματισμοί για προοπτικές αξιοποίησης του Internet από τον Ιατρικό κόσμο, ενώ σκεπτικισμός και επιφυλάξεις έχουν διατυπωθεί για την αποτελεσματικότητά του.

1.2 ΤΙ ΕΙΝΑΙ ΤΟ INTERNET

Το Internet απλά αποτελεί το μέσο που δίνει τη δυνατότητα να συνδεθούν μεταξύ τους τα δίκτυα των Ηλεκτρονικών Υπολογιστών σε όλο τον κόσμο. Το κάθε δίκτυο Η/Υ, όπου και να βρίσκεται, στο Νοσοκομείο, το Πανεπιστήμιο, στους χώρους εργασίας μπορεί να συνδεθεί μέσω του Internet, με άλλα δίκτυα σε οποιοδήποτε μέρος της υφηλίου. Με τον τρόπο αυτό ο κάθε χρήστης του τοπικού δικτύου έχει τη δυνατότητα πρόσβασης σε άλλα δίκτυα (Μάργαρης και συν 2015).

Σήμερα υπάρχουν και λειτουργούν δίκτυα Η/Υ σε ιατρικές Σχολές, σε νοσοκομεία, στη χώρα μας και το εξωτερικό, που συνδέουν μεταξύ τους Η/Υ οι οποίοι βρίσκονται σε εργαστήρια, σε τμήματα, σε κλινικές διάσπαρτα σε διαφορετικούς χώρους, στο ίδιο κτίριο ή την ίδια περιοχή. Αυτά τα τοπικά δίκτυα Η/Υ σε άλλα σημεία της υφηλίου, μέσω του Internet και των υπηρεσιών που αυτό προσφέρει.

1.3 ΠΟΙΕΣ ΥΠΗΡΕΣΙΕΣ ΠΡΟΣΦΕΡΕΙ ΤΟ INTERNET

Πολλές και σημαντικές υπηρεσίες παρέχει το Internet. Με τη χρήση αυτών των υπηρεσιών που συνεχώς βελτιώνονται, ενώ νέες προστίθενται, λόγω της τεχνολογικής ανάπτυξης του software και hardware, η ψηφιακή επικοινωνία γίνεται ταχύτερη και φιλικότερη από όσο ήταν πριν. Οι πιο γνωστές από αυτές είναι:

1. E-MAIL ή Ηλεκτρονικό Ταχυδρομείο:

Ο κάθε χρήστης του Internet (λ.χ. από την Πάτρα) μπορεί να ανταλλάσσει μηνύματα, αρχεία κειμένου, προγράμματα κ.α. με έναν ή περισσότερους χρήστες ταυτόχρονα, σε οποιοδήποτε μέρος του κόσμου και αν βρίσκονται, χρησιμοποιώντας τη δική του μοναδική διεύθυνση. Η ηλεκτρονική διεύθυνση σχηματίζεται από το όνομα του χρήστη που είναι μοναδικό (λ.χ. niva) και από το όνομα του συστήματος (που παρέχει τη σύνδεση) που και αυτό είναι μοναδικό στο Internet (λ.χ. niva) και από το όνομα του συστήματος (που παρέχει τη σύνδεση) που και αυτό είναι μοναδικό στο Internet (λ.χ. diavlos.gr).

Απο το συνδυασμό αυτών των ονομάτων προκύπτει και η ηλεκτρονική διεύθυνση του κάθε χρήστη του Internet με τη μορφή niva@diavlos.gr. Με τον τρόπο αυτό η διεύθυνση είναι μοναδική και κάθε μήνυμα φτάνει στο συγκεκριμένο χρήστη σε ελάχιστα δευτερόλεπτα. Για τη χρήση του Ηλεκτρονικού Ταχυδρομείου υπάρχουν πολλά προγράμματα που ονομάζονται mailers και διατίθενται από διάφορους κόμβους δωρεάν στο Internet.

2.Υπηρεσία Telnet:

Η υπηρεσία αυτή παρέχει τη δυνατότητα σύνδεσης και χρήσης από τον υπολογιστή μας, μέσω του Internet, απομακρυσμένων υπολογιστών.

3.Υπηρεσία FTP:

Η υπηρεσία αυτή εξασφαλίζει τη μεταφορά αρχείων από έναν υπολογιστή του Internet σε κάποιον άλλο. σε κάποιον άλλο. Πολλοί υπολογιστές διαθέτουν για τους

χρήστες του Internet πλήθος αρχείων που μπορούν να τα μεταφέρουν στον υπολογιστή τους.

Η μεταφορά αυτή γίνεται με προγράμματα που μπορεί ο καθένας να προμηθευτεί δωρεάν από το Internet (FTP).

4.Υπηρεσία Usenet:

Η υπηρεσία αυτή επιτρέπει την ηλεκτρονική διάσκεψη πολλών χρηστών στο Internet. Ο κάθε χρήστης μπορεί να διατυπώσει κάποια ερώτηση, να κάνει κάποια ανακοίνωση ή να κοινοποιήσει μια άποψή του, την οποία έχουν δυνατότητα να διαβάσουν όλοι οι χρήστες του Internet και κατόπιν να πάρουν θέση ή να απαντήσουν. Υπάρχουν πάνω από 300.000 ομάδες (groups) που καλύπτουν οποιοδήποτε θέμα μπορεί να φανταστεί κανείς. Βέβαια υπάρχουν και πάρα πολλά groups με ιατρικά θέματα (Μάργαρης και συν 2015).

6. Υπηρεσία Talk και ICR:

Η πρώτη υπηρεσία επιτρέπει την άμεση και σε πραγματικό χρόνο επικοινωνία δύο χρηστών του Internet, όπου και να βρίσκονται αυτοί, εφόσον επιτευχθεί η σύνδεση μεταξύ τους. Με τον τρόπο αυτό,τι πληκτρολογεί ο ένας εμφανίζεται στην οθόνη του Η/Υ του άλλου.

Με τη δεύτερη υπηρεσία, το ICR έχουμε άμεση και σε πραγματικό χρόνο επικοινωνία πολλών χρηστών του Internet. Όλοι οι χρήστες συνδέονται σε ICR Servers και αφού επιλέξουν το κανάλι που φιλοξενεί το θέμα συζήτησης που τους ενδιαφέρει, ό,τι πληκτρολογούν μεταφέρεται στις οθόνες όλων των άλλων χρηστών που συμμετέχουν στη συζήτηση.

6.Υπηρεσία Gopher:

Πρόκειται για μια υπηρεσία όπου η πληροφορία παρουσιάζεται με τη χρησιμοποίηση ιεραρχικών επιλογών (μενού) που οδηγούν σε συγκεκριμένες περιοχές πληροφοριών.

7. World Wide Web:

Είναι η υπηρεσία που έφερε «επανάσταση» στο Internet. Λόγω της εμφάνισής της έγινε προσιτό το Internet σε εκατομμύρια ανθρώπους σε ελάχιστο χρονικό διάστημα. Η φιλικότητα της χρήσης της υπηρεσίας αυτής, η χρήση των εικόνων, των video, των ήχων, των κειμένων που όλα μαζί μπορούν να παρουσιαστούν στις οθόνες των Η/Υ του κάθε χρήστη στο Internet, δημιούργησε νέες δυνατότητες. Η χρήση του Hypertext (του υπερκειμένου) δημιούργησε τις προϋποθέσεις για αλληλεπίδραση από μακριά, εξασφαλίζοντας έτσι σε ελάχιστα χρονικά διαστήματα την ανθρώπινη συμμετοχή σε διαδικασίες απόκτησης και επεξεργασίας πληροφοριών που μέχρι χθες ήταν δυνατό να πραγματοποιηθούν μόνο με άμεση πρόσβαση.

Η δημιουργία σελίδων αλληλεπίδρασης με multimedia (ήχος, κίνηση, κείμενο κ.α.) που περιέχουν Hypertext (δηλαδή λέξεις κλειδιά που συνδέουν τη σελίδα με άλλες παρόμοιες στο Internet) επέτρεψε την αλματώδη αύξηση των χρηστών και της διακινούμενης πληροφορίας.

Το World Wide Web επεκτείνεται και εμπλουτίζεται συνεχώς με νέους κόμβους (Web Servers) ποικίλης πληροφορίας, με νέα προγράμματα που εξασφαλίζουν την αξιοποίηση κάθε πληροφορίας σε μικρότερο χρόνο και τη φιλικότερη παρουσίαση και διαχείρισή της.

Για να χρησιμοποιήσουμε αυτή την υπηρεσία πρέπει να χρησιμοποιήσουμε προγράμματα τα οποία διατίθενται δωρεάν στο Internet όπως το Netscape, Mosaic κ.α.

Όσον αφορά την Ιατρική, υπάρχουν χιλιάδες κόμβοι σε όλο τον κόσμο, με τη μορφή των Web Servers, που παρέχουν πολλές εξειδικευμένες ιατρικές πληροφορίες και μια δυσκολία που δημιουργείται από τη συνεχιζόμενη αύξηση της ποσότητας των πληροφοριών είναι ο εντοπισμός εκείνων που μας ενδιαφέρουν. Η ανεύρεση τέτοιων πληροφοριών στο World Wide Web γίνεται εύκολη με την ύπαρξη των μηχανών αναζήτησης, που καταγράφουν τα περιεχόμενα των σελίδων του Web. Θέτουμε το ερώτημα και οι μηχανές αυτές μας επιστρέφουν ένα πλήρη κατάλογο με σελίδες που περιέχουν την πληροφορία που ψάχνουμε.

8.Videoconferencing:

Είναι υπηρεσία που υπάρχει στο Internet και είναι ελάχιστα διαδεδομένη. Επιτρέπει τη δυνατότητα επικοινωνίας μεταξύ δύο ή περισσότερων ατόμων που βρίσκονται σε απόσταση μεταξύ τους με τέτοιο τρόπο ώστε να υπάρχει οπτική και ακουστική επαφή. Για την υλοποίηση μιας τέτοιας σύνδεσης μέσα από ένα απλό τηλεφωνικό δίκτυο (με όλα τα μειονεκτήματα που αυτό συνεπάγεται) απαιτούνται ειδικά προγράμματα και κατάλληλος εξοπλισμός (κάμερα, κάρτα ήχου, μικρόφωνο).

1.4 Η ΣΥΝΔΕΣΗ ΜΕ ΤΟ INTERNET

Για μια τυπική σύνδεση με το Internet απαιτείται η ύπαρξη ενός Ηλεκτρονικού Υπολογιστή, ενός modem, μιας απλής τηλεφωνικής γραμμής του ΟΤΕ, των κατάλληλων προγραμμάτων επικοινωνίας και ενός Internet Service Provider (που θα μας επιτρέπει την πρόσβαση στο Internet).

Τα προγράμματα επικοινωνίας που επιτρέπουν τη σύνδεση με το Internet είναι εύκολο να τα προμηθευτούμε χωρίς οικονομική επιβάρυνση.

Βασικός κρίκος για τη σύνδεσή μας με το Internet είναι ο Internet Service Provider. Είναι εταιρίες που μπορούν να μας δώσουν τη δυνατότητα σύνδεσης με το Διαδίκτυο, διαθέτοντάς μας ένα λογαριασμό (account) στους υπολογιστές της, με κάποιο προσιτό οικονομικό κόστος. Σήμερα υπάρχουν τέτοιες εταιρίες που διαθέτουν κόμβους σε πολλές πόλεις της χώρας μας.

Με τον τρόπο αυτό από τον Υπολογιστή μας, με το κατάλληλο πρόγραμμα επικοινωνίας και με το modem, χρησιμοποιώντας την απλή τηλεφωνική γραμμή (με κόστος μιας αστικής μονάδας) συνδεόμαστε με τον Internet Provider και μέσω αυτού με οποιοδήποτε δίκτυο Η/Υ στο Internet.

1.5 Η ΙΑΤΡΙΚΗ ΠΛΗΡΟΦΟΡΙΑ ΣΤΟ INTERNET

Στο Internet υπάρχουν και διακινούνται τεράστιες ποσότητες ψηφιακής πληροφορίας που αφορούν την Ιατρική και τους λειτουργούς της. Πληροφορίες λιγότερο ή περισσότερο εξειδικευμένες που έχουν σχέση με όλες τις Ιατρικές ειδικότητες και είναι διεσπαρμένες σε κόμβους (Web Servers) σε όλο τον κόσμο. Οι πληροφορίες αυτές έχουν μερικά σημαντικά χαρακτηριστικά. Είναι πληροφορίες που παράγονται από κέντρα, ομάδες ή ακόμα και μεμονωμένους επιστήμονες και είναι έγκυρες και χρήσιμες. Συνήθως είναι πληροφορίες που δημοσιοποιούνται άμεσα και επώνυμα, γεγονός που εξασφαλίζει την εγκυρότητά τους. Οι πληροφορίες αυτές απευθύνονται σε ιατρούς διαφόρων ειδικοτήτων και είναι δυνατό να είναι εξειδικευμένες. Διατίθενται ελεύθερα και χωρίς ιδιαίτερο κόστος στον κάθε ενδιαφερόμενο ιατρό. Είναι πολύμορφες πληροφορίες, με τη μορφή επιστημονικών κειμένων, φωτογραφιών, ήχου και video που έχουν ψηφιοποιηθεί και μπορούν εύκολα να μεταφερθούν από την πηγή σε κάθε μέρος του πλανήτη. Οι χρήσιμες αυτές πληροφορίες διατίθενται στο Internet μέσω των υπηρεσιών του και κυρίως του World Wide Web. Βέβαια υπάρχουν και πληροφορίες που απευθύνονται σε απλούς χρήστες του Internet και είναι προσιτές και κατανοητές μια και τις περισσότερες φορές διατίθενται από επιστήμονες και αφορούν συγκεκριμένα θέματα υγείας, πρώτες βοήθειες κ.α.

Αυτά τα χαρακτηριστικά των πληροφοριών, με την ολοένα αυξανόμενη ανάπτυξη του Internet και το προσιτό κόστος του εξοπλισμού και της εύκολης σύνδεσης του Η/Υ με τον provider, έχουν μεταβάλλει τη δυνατότητα της διακίνησης της Ιατρικής πληροφορίας, σε σύγκριση με τους κλασσικούς τρόπους μεταφοράς και διάδοσης. Ολοένα και περισσότεροι ιατροί χρησιμοποιούν το Internet. Όλοι πια έχουν τη δυνατότητα πρόσβασης στην εξειδικευμένη Ιατρική πληροφορία, χωρίς διακρίσεις και απαγορεύσεις που οφείλονται στο χρόνο και στο χώρο. Ο γιατρός του Αγροτικού Ιατρείου σε μια ακριτική περιοχή έχει την ίδια δυνατότητα πρόσβασης στην Ιατρική πληροφορία με τον Νοσοκομειακό ή Πανεπιστημιακό Ιατρό, γεγονός που μέχρι σήμερα τουλάχιστον ήταν εξαιρετικά δύσκολο. Ένα Κέντρο Παραγωγής Ιατρικής πληροφορίας, λ.χ. ένα εργαστήριο ή μια κλινική ενός νοσοκομείου ή ένας μεμονωμένος ιατρός ακόμα, μπορεί να παρουσιάσει το επιστημονικό του έργο στην Παγκόσμια Ιατρική κοινότητα και να δεχτεί τις

παρατηρήσεις και τις κριτικές της. Μέσα από το Internet γνωστοποιείται ισότιμα και κρίνεται η δουλειά όλων των επιστημονικών κέντρων, ενώ εμφανίζονται ολοένα και περισσότερα κέντρα παροχής Ιατρικών πληροφοριών, στα οποία εύκολα και γρήγορα μπορεί να φτάσει ο κάθε ιατρός. Έτσι, μέσα από μια συνεχή ψηφιακή διακίνηση ιατρικής πληροφορίας, αναβαθμίζεται η δια βίου εκπαίδευση των ιατρών, η εξειδικευμένη πληροφορία συμβάλλει στον εμπλουτισμό της παγκόσμιας γνώσης και βέβαια αξιοποιείται από όλους. Μπορούμε να ισχυριστούμε ότι δίπλα στις κλασσικές μεθόδους της συνεχιζόμενης εκπαίδευσης (έντυπη ενημέρωση, χρήση σύγχρονων οπτικοακουστικών μέσων, συναντήσεων σε τοπικό επίπεδο, ημερίδων, συμποσίων, συνεδρίων κ.α.) των ιατρών, το Internet αποτελεί ένα σημαντικό εργαλείο που αν αξιοποιηθεί κατάλληλα (με τις υπάρχουσες υλικοτεχνικές δυνατότητες) μπορεί να αποδώσει πλούσιους καρπούς. Πρόσφατα έγινε στο Internet το Πρώτο Ελληνικό Ιατρικό Συνέδριο με γενικό τίτλο "Παθήσεις και Κακώσεις του Χεριού", που υπήρξε και το πρώτο συνέδριο που εξολοκλήρου διεξήχθη στον κυβερνοχώρο. Η ορθοπεδική οικογένεια πήρε την πρωτοβουλία να χρησιμοποιήσει αυτή τη μορφή επικοινωνίας και επιστημονικής ενημέρωσης, για πρώτη φορά στην Ελλάδα και μάλιστα αυτή η πρωτοβουλία ξεκίνησε από τη Βόρεια Ελλάδα, από τη Θεσσαλονίκη. Για την ιστορία και μόνο αναφέρουμε την Ορθοπεδική Κλινική του Β' Νοσοκομείου ΙΚΑ Θεσσαλονίκης που είχε την ευθύνη και το συντονισμό του τριμήνου αυτού του συνεδρίου. Στο συνέδριο αυτό οι σύνεδροι (και μπορούμε να πούμε ότι δεν ήταν λίγοι, πάνω από 400 άτομα) παρακολούθησαν τις διαλέξεις των εισηγητών και διάβασαν τις εργασίες από το σπίτι τους ή από τους χώρους εργασίας τους. Η εμπειρία που αποκτήθηκε υπήρξε θετική και σίγουρα ενθαρρύνει και άλλους ιατρούς να τη μιμηθούν και να τη βελτιώσουν.

Η δυνατότητα αξιοποίησης των multimedia στο Internet και η αλληλεπίδραση με τις σελίδες του World Wide Web (WWW) από το χρήστη, δημιουργεί νέες δυνατότητες στην εκπαίδευση από μακριά. Είναι μια μορφή τηλεματικής μέσα από το Internet, που μπορεί να συνδέσει για εκπαιδευτικούς λόγους (εκπαίδευση ειδικευομένων) τα επιστημονικά κέντρα (Νοσοκομείο, Πανεπιστήμιο) με ιατρούς σε απομακρυσμένα σημεία της υφελίου. Μέσα από τις προηγούμενες αναφορές διακρίνεται καθαρά η παγκοσμιότητα και η

αμεσότητα, η φιλικότητα, η ευκολία πρόσβασης και χρήσης αυτού του είδους της ψηφιακής επικοινωνίας, που μαζί με το χαμηλό κόστος της, την κάνει ολοένα και πιο δημοφιλή στον Ιατρικό κόσμο της χώρας μας.

1.6 ΑΞΙΟΠΟΙΗΣΗ ΤΟΥ INTERNET ΑΠΟ ΤΗΝ ΙΑΤΡΙΚΗ ΚΟΙΝΟΤΗΤΑ

Τον κάθε ειδικό ιατρό, τον ενδιαφέρει κατά κύριο λόγο η εξειδικευμένη πληροφορία. Η πληροφορία που παράγεται από τα εξειδικευμένα κέντρα και προορίζεται για αυτόν και τους υπόλοιπους συναδέλφους του. Συνεπώς η υπόθεση αυτή αφορά το σύνολο των ιατρών κάθε ειδικότητας, που εκφράζεται στην Ιατρική κοινότητα, μέσα από τις Επιστημονικές Εταιρίες λ.χ. η Ε.Ε.Χ.Ο.Τ. για τους ορθοπεδικούς ή η Ορθοπεδική Εταιρία Μακεδονίας Θράκης για τους Ορθοπεδικούς της Βορείου Ελλάδος. Οι επιστημονικοί φορείς είναι οι καταλληλότεροι για να αναλάβουν τη δημιουργία εξειδικευμένων Web Servers και να χρησιμοποιήσουν τις δυνατότητες που προσφέρονται από το Internet. Μέχρι σήμερα όμως κάτι τέτοιο δεν έχει επιτευχθεί. Στην Αμερική, οι ιατρικές επιστημονικές εταιρίες διαφορετικών ειδικοτήτων έχουν ήδη τους δικούς τους εξειδικευμένους κόμβους στο Internet και παρέχουν πλήθος εξειδικευμένων πληροφοριών λ.χ. η Αμερικανική Ορθοπεδική Ακαδημία (AAOS) διαθέτει έναν από τους πιο έγκυρους και τεκμηριωμένους κόμβους στο Internet με πληθώρα ορθοπεδικής πληροφορίας. Ούτε ο χρόνος, ούτε και ο τόπος αποτελούν εμπόδιο στη μεταφορά της αμερικανικής εμπειρίας στον ιατρικό κόσμο της χώρας μας. Και η τεχνογνωσία υπάρχει και το κατάλληλο υλικό για να στηθούν εξειδικευμένοι ιατρικοί κόμβοι διαφόρων ειδικοτήτων.

Ήδη υπάρχουν και λειτουργούν πάνω από χρόνο ορισμένοι Ελληνικοί ιατρικοί κόμβοι. Ένας από αυτούς είναι η ΟΡΘΟΠΕΔΙΚΗ ή οι Ορθοπεδικές Σελίδες, από τη Θεσσαλονίκη. Στον κόμβο αυτό περιέχεται και διακινείται μόνον ορθοπεδική πληροφορία. Κάθε πληροφορία που ενδιαφέρει τον ορθοπεδικό γιατρό έχει τη θέση της στον κόμβο αυτό. Ο κόμβος βασίζεται στην αρχή της συμμετοχής των ορθοπεδικών γιατρών, των

ορθοπεδικών κλινικών και επιστημονικών ορθοπεδικών εταιριών στον εμπλουτισμό του, με εξειδικευμένη πληροφορία στην οποία διαθέτει σε κάθε χρήστη γιατρό του Internet. Η πρόσβαση στον κόμβο είναι ελεύθερη. Ο κόμβος διαθέτει όλες τις υπηρεσίες του Internet και είναι WEB Server, με δυνατότητες αλληλεπίδρασης, χρήσης των Hypertext και πολυμέσων μέσα από τις σελίδες του.

ΚΕΦΑΛΑΙΟ 2

2.1 ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ

Πληροφορική είναι η επιστήμη που ασχολείται με τις επιστημονικές και τεχνολογικές έννοιες. Που έχει ως αντικείμενο την έρευνα, συλλογή, αποθήκευση, επεξεργασία, παραγωγή και μετάδοση των πληροφοριών, χρησιμοποιώντας ως κύριο εργαλείο τον ηλεκτρονικό υπολογιστή. Η έννοια της πληροφορίας είναι το αποτέλεσμα που μας δίνει ο ηλεκτρονικός υπολογιστής μετά από την επεξεργασία των δεδομένων. Η έγκαιρη και αξιόπιστη πληροφορία είναι το πλέον απαραίτητο στοιχείο για τη λήψη σωστών αποφάσεων. Οι πληροφορίες μπορεί να είναι οποιασδήποτε μορφής στατικές δηλαδή δεν αλλάζουν στην πάροδο του χρόνου ή δυναμικές όπου αυτές αλλάζουν στην πάροδο του χρόνου. Έτσι, η πληροφορία ορίζεται ως στοιχείο γνώσης που μετουσιώνει και δίνει αξία στα πράγματα, τα οποία αποκτούν σημασία με αυτόν τον τρόπο. Πρέπει πάντως να γίνει ο διαχωρισμός μεταξύ πληροφορίας και γνώσης: η πληροφορία είναι το μέγεθος που, καθώς αυξάνει και επεξεργάζεται, προσδίδει σημασία στα πράγματα και τα γεγονότα και παράγει γνώση. Αυτό μπορεί να γίνει καταληπτό στο σημερινό τρόπο ζωής που, ενώ κατακλυζόμαστε από πληροφορίες, η επεξεργασία του και τελικά η παραγωγή γνώσης δεν είναι αυτονόητη (Μάργαρης και συν 2015).

Η έννοια της πληροφορίας ορίζεται συχνά σε σχέση με την έννοια των δεδομένων (data). Οι δύο έννοιες πολλές φορές συγχέονται ακριβώς γιατί αποκτώντας τα δεδομένα σημασία μεταπίπτουν σε πληροφορία. Τα δεδομένα δεν είναι πληροφορία, όταν δεν έχουν λάβει από τη νόηση συγκεκριμένη σημασία, ενώ η πληροφορία είναι δεδομένα με σημασία, δεδομένα που έχουν «ουσιαστικό» περιεχόμενο. Η μετάβαση από τα δεδομένα στην πληροφορία μπορεί να αφορά την επίλυση κάποιου συγκεκριμένου ζητήματος, την απάντηση κάποιου συγκεκριμένου ερωτήματος και γενικότερα παραπέμπει σε καινούριο στοιχείο γνώσης για κάτι. Σε απλοποιημένη ή σχηματική μορφή, δεδομένα + σημασία = πληροφορία.

Σε αντίθεση με τα δεδομένα η πληροφορία έχει ορισμένο νόημα και είναι οργανωμένη για συγκεκριμένο σκοπό. Με την προσθήκη αξιών στα δεδομένα, με διαμόρφωση,

οργάνωση, μαθηματική ή στατιστική ανάλυση, διόρθωση λαθών ή συμπίεση παράγεται η πληροφορία. Τη διαδικασία αυτή την ονομάζουμε επεξεργασία δεδομένων. Στην Πληροφορική συχνά δεν γίνεται διάκριση ανάμεσα στα δεδομένα και την πληροφορία. Έτσι, τα δεδομένα μπορούν να οριστούν ως τρόποι αναπαράστασης εννοιών και γεγονότων που δύνανται να υποστούν διαχείριση και επεξεργασία από τον άνθρωπο, ή από ένα αυτοματοποιημένο υπολογιστικό σύστημα. Αυτοί οι τρόποι αναπαράστασης, συσχετίζονται άμεσα με τη φύση της πληροφορίας που μεταφέρεται από τα δεδομένα. Η διαδικασία επιλογής του κατάλληλου είδους αναπαράστασης ώστε να μπορεί να αποθηκευτεί η επιθυμητή πληροφορία χωρίς να υπάρξουν απώλειες και χωρίς να γίνει σπατάλη πόρων ονομάζεται (Data Modeling).

Καθώς προσπαθούμε να διακρίνουμε τις διαφορές μεταξύ πληροφορίας και δεδομένων, πρέπει να κάνουμε αναφορά και στην έννοια των αδόμητων ή ακατέργαστων δεδομένων(raw data). Συνήθως πρόκειται για δεδομένα που συλλέγονται από κάποια πηγή χωρίς επεξεργασία και αναφέρονται και ως πρωτεύοντα δεδομένα (primary data). Τα αδόμητα δεδομένα μπορεί να προέρχονται, π.χ., από χειρονακτική εισαγωγή στοιχείων σε ένα πρόγραμμα. Έτσι, τα αδόμητα δεδομένα εν γένει περιέχουν σφάλματα και διαφορετική ή καθόλου μορφοποίηση (format): μια εισαγωγή δεδομένων ημερομηνίας : "31st January 1999", "31/01/1999", "31/1/99", "31 Jan". Άπαξ και αντληθούν τα αδόμητα δεδομένα, πρέπει να υποστούν επεξεργασία για να αποκτήσουν ενιαία μορφή, απαλλαγμένη από σφάλματα, ώστε να παρέχουν πραγματική πληροφορία.

Στον προγραμματισμό τα δεδομένα χαρακτηρίζονται από τον τύπο τους (data type). Ο τύπος δεδομένων είναι μια ταξινόμηση που καθορίζει :

- τις πιθανές τιμές που μπορεί να έχουν τα δεδομένα μας (π.χ. ακέραιος αριθμός ή πραγματικός, συμβολοσειρά)
- τις πράξεις που μπορούν να γίνουν με τις τιμές αυτού του τύπου
- το νόημα των δεδομένων,
- τον τρόπο αποθήκευσης των δεδομένων.

Έτσι μπορούμε να πούμε ότι ο τύπος των δεδομένων είναι μια τυπική περιγραφή των τιμών που μπορεί να αποθηκευτούν σε μια σταθερά ή μεταβλητή καθώς και των πράξεων που μπορεί να εκτελεστούν με τις τιμές αυτές. Οι τύποι δεδομένων χρησιμοποιούνται εντός συστημάτων τύπων, τα οποία προσφέρουν διαφορετικούς τρόπους για τον ορισμό, την ανάπτυξη και τη χρήση τους και διαφοροποιούνται ως προς τους βαθμούς ασφάλειας των τύπων. Έτσι, είναι άλλο το σύστημα δεδομένων της γλώσσας C και άλλο της PHP. Ο κάθε τύπος δεδομένων αναπαριστά και έναν περιορισμό για το σύστημα τύπων δεδομένων, περιγράφοντας μια αναπαράσταση που μεταφράζεται σε μια δομή τιμών ή αντικειμένων τα οποία αποθηκεύονται στη μνήμη του υπολογιστή. Το σύστημα τύπων δεδομένων χρησιμοποιεί την πληροφορία των τύπων για να ελέγξει την ορθότητα των προγραμμάτων που προσεγγίζουν ή χειρίζονται τα δεδομένα (Μάργαρης και συν 2015, Μηνά I. 2017).

Η έννοια της δομής δεδομένων (data structure) αναφέρεται στους διαφορετικούς δυνατούς τρόπους οργάνωσης και αποθήκευσης δεδομένων, ώστε αυτά να μπορούν να χρησιμοποιηθούν αποδοτικά στον προγραμματισμό. Σε συγκεκριμένες εφαρμογές χρησιμοποιούνται συγκεκριμένες δομές δεδομένων, δηλαδή αυτές που είναι οι πιο αποδοτικές αλγοριθμικά για το κάθε είδος εφαρμογής. Παρέχουν έναν τρόπο αποδοτικής διαχείρισης τεράστιου όγκου δεδομένων, όπως μεγάλες βάσεις δεδομένων και υπηρεσίες ευρετηρίου στο διαδίκτυο. Οι δομές δεδομένων διακρίνονται επίσης σε στατικές και σε δυναμικές.

Βάση δεδομένων ονομάζεται ένα σύνολο πληροφοριών που έχουν οργανωθεί, έτσι ώστε η χρήση τους να είναι γρήγορη και αποτελεσματική και χρησιμοποιείται σε χώρους σπιτιού, εργασίας, σχολείων και στο διαδίκτυο.

Τα δεδομένα που αποθηκεύονται σε μία βάση δεδομένων οργανώνονται σύμφωνα με δύο βασικά μοντέλα:

1. Ιεραρχικό, το οποίο έχει άκαμπτη μορφή αφού το περιεχόμενό του διαμορφώνεται και δομείται με διάταξη δέντρου σαν τα διαγράμματα ροής. Το μοντέλο αυτό καθορίζει τα στοιχεία δεδομένων και τις αναμεταξύ τους σχέσεις, και παρόλο που ποικίλει σε πολυπλοκότητα χαρακτηρίζεται από την

καθορισμένη με σαφήνεια δομή του. Οι περισσότερες ιεραρχικές βάσεις δεδομένων περιέχουν όλες τις πληροφορίες σε ένα αρχείο ενιαίο, μέσα στο οποίο τα στοιχεία μπορούν να ομαδοποιηθούν από τα κοινά χαρακτηριστικά. Εννοιολογικά μια τέτοια βάση δεδομένων είναι εύκολο να εξεταστεί, περιορίζεται όμως σοβαρά από την δυνατότητα να αλλάξει οργανωτική δομή ανάλογα με τις ανάγκες του χρήστη.

2. Σχηματικό, στο οποίο τα στοιχεία των διαφορετικών πινάκων συνδέονται μεταξύ τους με σχέσεις που βασίζονται στο ταίριασμα πεδίων. Το μοντέλο αυτό επιτρέπει την εύκολη αναδιάρθρωση της βάσης δεδομένων με χρήση επαναπροσδιορισμού των σχέσεων μεταξύ των πινάκων χωρίς να επηρεάζει τα περιεχόμενα των πινάκων για την οποία και είναι δημοφιλές.

Το πλεονέκτημα των βάσεων δεδομένων είναι η ευκολία εύρεσης συγκεκριμένων πληροφοριών. Αν η βάση δεδομένων είναι μηχανογραφημένη οι πληροφορίες ανακαλούνται ταχύτερα με περισσότερη ακρίβεια και ο χειρισμός τους είναι πιο προχωρημένος.

Το μειονέκτημα μίας μηχανογραφημένης βάσης δεδομένων είναι ο χρόνος που απαιτείται για τη δημιουργία και ανάπτυξή της.

Ενδεικτικά παραδείγματα βάσεων δεδομένων είναι το πελατολόγιο μίας επιχείρησης, το προσωπικό, κατάλογος επαφών, κατάλογος εμπορευμάτων και οι εγκυκλοπαίδειες. Κάθε βάση δεδομένων διαθέτει ένα ευρετήριο, μέσω του οποίου ταξινομούνται τιμές σε μία ή περισσότερες στήλες ενός πίνακα της βάσης. Το ευρετήριο βασίζεται σε πεδία που επιτρέπουν την διάκριση μεταξύ εγγραφών και την εύρεση συγκεκριμένων αρχείων.

Για να δημιουργηθεί μία βάση δεδομένων χρειάζεται ένα λογισμικό εφαρμογής και σχεδιάζεται ανάλογα με τις απαιτήσεις του κατασκευαστή. Υπάρχουν 4 βήματα στη δημιουργία και χρήση μίας βάσης δεδομένων:

- Ο σχεδιασμός, δηλαδή το θεωρητικό διάγραμμα των πινάκων και των σχέσεων που θα τους συνδέουν
- Ο ορισμός των πινάκων στους οποίους αποθηκεύονται τα δεδομένα

- Η προσθήκη δεδομένων συνήθως μέσω μιας φόρμας
- Η εξαγωγή πληροφοριών από τη βάση δεδομένων χρησιμοποιώντας ένα φίλτροερώτημα για την παραγωγή μιας αποτελεσματικής παρουσίασής τους μέσω μιας έκθεσης.

Εν κατακλείδι ,οι βάσεις δεδομένων είναι ένα σύνολο δεδομένων που χαρακτηρίζονται από κάποια λογική οργάνωση και ομαδοποίηση έτσι ώστε να είναι εύκολη και αποτελεσματική η διαχείριση τους .Οι βάσεις δεδομένων κατέχουν κεντρική θέση στην επιστήμη της πληροφορικής, διότι επιτυγχάνουν διαχείριση της πληροφορίας. Σύστημα διαχείρισης βάσεων δεδομένων ο οποίος χρησιμοποιείται για τις ειδικές εφαρμογές λογισμικού που χρησιμοποιούνται για την υλοποίηση, τη συντήρηση, την επεξεργασία και γενικά τη διαχείριση των βάσεων δεδομένων(Μάργαρης και συν 2015).

2.2 ΙΑΤΡΙΚΕΣ ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ

Για την διεκπεραίωση μίας επιστημονικής έρευνας ο φοιτητής θα απευθυνθεί για πληροφορίες στους βασικούς τύπους πηγών, οι οποίοι είναι:

- βιβλία του επιλεγμένου κλάδου
- επιστημονικά CD-ROM
- Παγκόσμιος Ιστός (World Wide Web)
- τοπικό δίκτυο web site

2.2.1 ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ ΗΛΕΚΤΡΟΝΙΚΗΣ ΑΝΑΖΗΤΗΣΗΣ

Πλέον στην σύγχρονη κοινωνία ο τομέας της βιοϊατρικής εξελίσσεται ταχύτατα, Με αποτέλεσμα να αυξάνεται και η ζήτηση περισσότερων πληροφοριών για τις κατάλληλες βάσεις βιοϊατρικών δεδομένων, που θα επιτρέπουν την εύκολη και γρήγορη

πρόσβαση για αξιοποίηση των δεδομένων αυτών. Καθώς είναι καινούργια και η ανάπτυξη εργαλείων και συστημάτων που θα αυτοματοποιούν τις

διαδικασίες αναζήτησης και ανάκτησης πληροφορίας σε αυτά τα δεδομένα (Μάργαρης και συν 2015).

- Education Resources Information Center (ERIC) : κλάδοι διοίκησης, εκπαιδευτικών προγραμμάτων, διδασκαλίας, μάθησης
- MEDLINE: βιοϊατρική, μέριμνα υγείας, γεροντολογία κ.α.
- PsyINFO : ψυχολογία και σχετικές επιστήμες συμπεριφοράς που προέρχεται από το Psychological Abstracts με CD-ROM το PsyLit
- Sociological Abstracts: κοινωνική, αναπτυξιακή και κλινική ψυχολογία με CDROM το Sociolife
- Dissertation Abstracts Online: διατριβές
- UnCover: υπηρεσία βασιζόμενη σε αμοιβή που προσφέρει την ευκαιρία απόκτησης των περισσότερων άρθρων με FAX.

Η Εθνική Βιβλιοθήκη Ιατρικής των Ηνωμένων Πολιτειών (NLM) που εδρεύει στο Εθνικό Ινστιτούτο Υγείας (Bethesda, Maryland) έχει αναπτύξει το σύστημα αναζήτησης πληροφοριών Medical Literature Analysis and Retrieval System, το οποίο παρέχει στους χρήστες πρόσβαση στην βιοϊατρική πληροφόρηση. Η ηλεκτρονική της διεύθυνση είναι (www.nlm.nih.gov) και έχει ως πρώτη κι ευρέως γνωστή βιβλιογραφική βάση δεδομένων την MEDLINE. Αυτό περιλαμβάνει άρθρα που χρονολογούνται από το 1996 έως σήμερα. Επίσης άρθρα από το 1963-1965 που καταχωρήθηκαν μετά το 1966. Περιοδικά από το 1953 έως το 1965 ανασύνται από το OLDMEDLINE. Περιέχει πληροφορίες άρθρων , ονόματα συγγραφέων, τίτλος άρθρου, όνομα περιοδικού, έτος έκδοσης , τόμος, τεύχος και σελίδες, Κωδικός στο Pubmed ,κατάσταση του άρθρου. Άλλες πληροφορίες (γλώσσα, review, no abstract available).

Έτσι το Pubmed έχει αναφορές σε βιοϊατρικού περιεχομένου κείμενα από την Medline επιστημονικά άρθρα και ηλεκτρονικής μορφής βιβλία. Οι αναφορές μπορεί να περιλαμβάνουν από συνδέσμους μέχρι και πλήρους μεγέθους άρθρα, από την Medline και άλλες διαδικτυακές πηγές. Οι παραπομπές και οι περιλήψεις που παρέχει, περιλαμβάνουν τα πεδία της βιοϊατρικής και υγείας, καλύπτοντας επιστήμες ζωής, συμπεριφοράς, χημείας και βιομηχανικής(Μάργαρης και συν 2015).

Η PubMed Central παρέχει στους χρήστες μία διεπαφή μέσω της οποίας αυτοί μπορούν να υποβάλλουν όρους για σχετική αναζήτηση κειμένων.Επίσης παρέχει μια σειρά από ειδικού τύπου αναζητήσεων μέσω των οποίων οι χρήστες μπορούν να αναζητήσουν υλικό βάση του συγγραφέα και άλλων πεδίων (Μηνά Ι. 2017).

Ακόμη άλλη λειτουργία που παρέχεται από την PubMed Central, είναι η επιστροφή ενός συνόλου σχετικών κειμένων με το άρθρο που βλέπει ο χρήστης, τα οποία υπολογίζονται μέσω αλγορίθμου συσταδοποίησης. Τα σύνολα σχετικών κειμένων είναι προϋπολογισμένα και έτσι επιταχύνεται η διαδικασία επιστροφής και εμφάνισής τους (Μηνά Ι. 2017).

Η PubMed Central παρέχει επίσης στους χρήστες την δυνατότητα να Βασικά χαρακτηριστικά:

- Εύχρηστο μηχανισμό αναζήτησης
- Επεξήγηση του τρόπου επεξεργασίας των αναζητήσεων του χρήστη
- Ελεγχόμενο λεξιλόγιο MESH (Medical Subject Headings)
- Εύχρηστο οδηγό χρήσης με ξεναγήσεις (tour) σε συγκεκριμένες δυνατότητες της υπηρεσίας
- Εξατομικευμένες υπηρεσίες (απαιτείτε εγγραφή) Παρέχει την δυνατότητα:
- Ανάκτησης λίστας αποτελεσμάτων μέσω εκτύπωσης ηλεκτρονικού ταχυδρομείου και αποθήκευσης
- Εξαγωγής αποτελεσμάτων σε ειδικό λογισμικό διαχείρισης βιβλιογραφικών αναφορών (Endnote, Reference Manager, ProCite)
- Δημιουργία ιστορικού αναζητήσεων

- Ενεργοποίησης υπηρεσιών ειδοποίησης με την τεχνολογία RSS
. Η πρόσβαση σε αυτήν γίνεται μέσω του Pubmed και του Internet Grateful Med από την ηλεκτρονική διεύθυνση (www.nlm.nih.gov/databases/freemedl.html).

Το σύστημα αναζήτησης PubMed αναπτύχθηκε απ' το National Center for Biotechnology Information (NCBI) της NLM που εδρεύει στο National Institutes of Health

(NIH). Αναπτύχθηκε σε συνεργασία με εκδότες βιοϊατρικής βιβλιογραφίας σαν εργαλείο αναζήτησης βιβλιογραφικών παραπομπών. Περιέχει βιβλιογραφικές εγγραφές

Κάθε φορά που πραγματοποιείται αναζήτηση βιβλιογραφίας στο PubMed, αρχικά πρέπει να καθορίζεται με σαφήνεια το ερευνητικό ερώτημα που πρόκειται να διερευνηθεί. Έπειτα, καθορίζονται οι λέξεις-κλειδιά (keywords) ή, αλλιώς, οι όροι αναζήτησης (search terms) που χρησιμοποιούνται στην αναζήτηση βιβλιογραφίας στο PubMed.. Η επιτυχής αναζήτηση και ανεύρεση των κατάλληλων αναφορών στο PubMed καθορίζεται σε μεγάλο βαθμό από τις λέξεις-κλειδιά που χρησιμοποιούνται από τους ερευνητές. Απαιτείται ιδιαίτερη προσοχή στον καθορισμό των λέξεων-κλειδιών, καθώς το PubMed δεν αντιλαμβάνεται τις σκέψεις των ερευνητών, αλλά απλώς προσαρμόζει την αναζήτηση σύμφωνα με τις λέξεις-κλειδιά που χρησιμοποιούνται από τους ερευνητές. Ουσιαστικά, οι λέξεις-κλειδιά είναι αυτές που καθορίζουν την εγκυρότητα της ανεύρεσης των κατάλληλων αναφορών στο PubMed.



Εικόνα1. Αρχική σελίδα της υπηρεσίας

Στην αρχική σελίδα της υπηρεσίας προβάλλεται ο βασικός μηχανισμός αναζήτησης, από όπου είναι δυνατή η επιλογή της βάσης όπου θα εκτελεστεί η αναζήτηση (search) και η εισαγωγή όρων αναζήτησης (for). Οι όροι αναζήτησης μπορεί να αφορούν συγκεκριμένο θέμα (topic), συγγραφέα (author) ή περιοδικό (journal). Στην αρχική σελίδα της υπηρεσίας ακόμη προβάλλονται πέντε καρτέλες χειρισμού των αναζητή που εκτελεί ο χρήστης.

Συγκεκριμένα:

- Περιορισμοί (limits). Καθορισμός περιορισμών στην αναζήτηση σχετικά με τον συγγραφέα, τίτλο περιοδικού, πλήρες κείμενο, γλώσσα, τύπος άρθρου και τα λοιπά...
- Περισκόπηση / Ευρετήριο (preview/intex). Προεπισκόπηση του αριθμού των αποτελεσμάτων για εκτέλεση συγκεκριμένης αναζήτησης. Το κομμάτι του ευρετηρίου προβάλλει στενότερους όρους σε σχέση με αυτό που εισήχθη ως όρος αναζήτησης, π.χ.

ADHDADHD/autism. Δίπλα από κάθε σετ όρων παρουσιάζεται ο αριθμός των αποτελεσμάτων που αντιστοιχούν σε αυτόν.

- Ιστορικό αναζητήσεων(history). Προβολή όλων των αναζητήσεων που εκτελέστηκαν κατά την τρέχουσα είσοδο(session) του χρήστη στην υπηρεσία. Οι πραγματοποιηθείσες αναζητήσεις είναι δυνατόν να: συνδυαστούν μεταξύ τους με BOOLEAN λογική (AND-OR-NOT), διαγραφούν (delete), εκτελεστούν ξανά (go), προβληθούν λεπτομέρειες εκτέλεσής τους από τον μηχανισμό της υπηρεσίας (details), αποθηκεύουν στον προσωπικό λογαριασμό του χρήστη (save in my NCBI).
- Πρόχειρο (clipboard). Προβολή επιλεγμένων εγγράφων κατά την τρέχουσα είσοδο (session) του χρήστη στην υπηρεσία.
- Λεπτομέρειες (details). Προβολή λεπτομερειών σχετικά με τον τρόπο με τον οποίο το σύστημα αναζήτησης κατανοεί και επεξεργάζεται τους όρους αναζήτησης του χρήστη. Παρουσιάζεται η μετάφραση των όρων αναζήτησης στις θεματικές επικεφαλίδες ιατρικού περιεχομένου που χρησιμοποιείται στην medline (Μηνά Ι. 2017).

2.3 ΠΡΟΒΟΛΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ ΑΝΑΖΗΤΗΣΗΣ

Η λίστα των αποτελεσμάτων μιας επιλογής αναζήτησης ενός τίτλου άρθρου οδηγεί στην σελίδα της συγκεκριμένης εγγραφής. Οι δυνατότητες χειρισμού της λίστας αποτελεσμάτων είναι: προβολή (display), καθορίζεται η μορφή με την οποία θα προβάλλονται οι εγγραφές, αριθμός αποτελεσμάτων (show) ανά σελίδα, ταξινόμηση αποτελεσμάτων (sort by) και αποθήκευση επιλεγμένων εγγράφων (sent to). Τα εικονίδια που μπορεί να εμφανίζονται μπροστά από κάθε αποτέλεσμα έχουν μια ιδιαίτερη σημασία: -no abstract δεν διατίθεται περίληψη του άρθρου

-abstract διατίθεται περίληψη του άρθρου

-free full-text διατίθεται ελεύθερο το πλήρες κείμενο του άρθρου

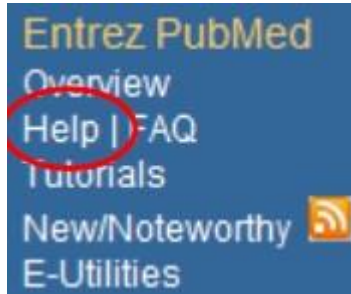
-free in PNC διατίθεται το πλήρες κείμενο του άρθρου από την ψηφιακή βιβλιοθήκη PubMed

Δίπλα από κάθε αποτέλεσμα παρέχονται υπερσύνδεσμοι για σχετικά άρθρα (**Related Articles**) και προς εξωτερικές πηγές πληροφόρησης (**Links**)

Εικόνα 1. Σελίδα προβολής εγγραφής/βιβλιογραφικής παραπομπής.

Η συγκεκριμένη εγγραφή διαθέτει περίληψη του άρθρου και σχετικούς συνδέσμους .

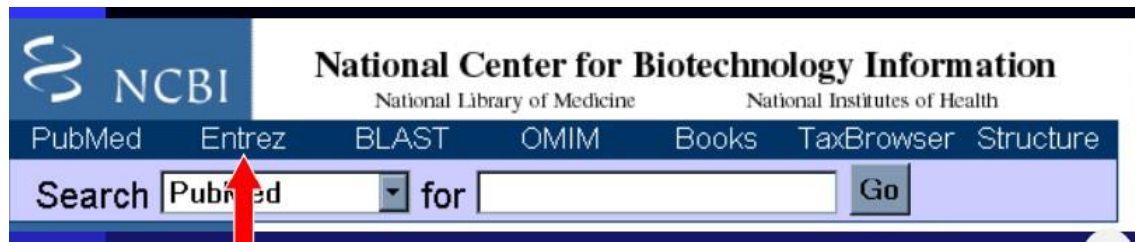
Ο Οδηγός ή βοήθεια είναι η επιλογή help του βασικού μενού οδηγεί στην αρχική σελίδα της βοήθειας. Ταυτόχρονα παρέχεται δυνατότητα αναζήτησης (search) στα περιεχόμενα της βοήθειας, καθώς και μετάβασης (navigation) οποιαδήποτε άλλο σημείο της μέσω της λίστας θεμάτων το αριστερό μέρος της σελίδας.



Εικόνα 3. Πρόσθετα εργαλεία- Υπηρεσίες

Το θεματικό πεδίο που καλύπτει το medline, καθώς και το μεγάλο εύρος χρηστών τόσο σε μορφωτικό επίπεδο όσο και σε εξειδικευμένα ερευνητικά ενδιαφέροντα .

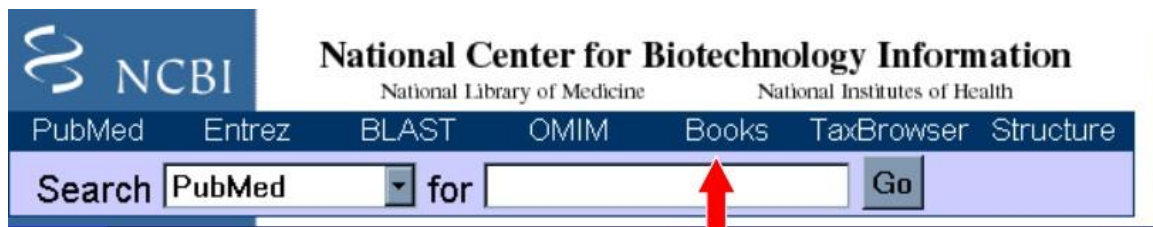
- Αναζητήσεις με συγκεκριμένα κριτήρια (JOURNALS DATABASE, MESE DATABASE)
- Αναζητήσεις συγκεκριμένων παραπομπών (Single Citation Matcher, Batch Citation Matcher).
- Εξειδικευμένες αναζητήσεις (Clinical Queries, Special Clinical Queries)
- Σύνδεση με σχετικό περιεχόμενο διαθέσιμο από άλλους φορείς (LinkOut), όπως βιβλιοθήκες, εκδότες και τα λοιπά.
- Εξατομικευμένες υπηρεσίες (My NCBI).



Εικόνα 4. Εργαλείο Entrez

Το Entrez είναι ένα εργαλείο βάσεων δεδομένων και περιέχει:

- Επιστημονικά κείμενα
- Βάσεις αλληλουχίων DNA και πρωτεϊνών
- Δεδομένα τριτοτάγους και τεταρτοτάγους δομής πρωτεϊνών
- Στοιχεία πληθυσμιακών μελετών



Εικόνα 5. Εργαλείο Books

Το BOOKS είναι online κατάλογος βιβλίων ο οποίος επιτρέπει την αναζήτηση σε αυτά με την χρήση λέξεων- κλειδιών και την εμφάνιση πλήρους κειμένου..

ΚΕΦΑΛΑΙΟ 3

3.1 ΣΥΣΤΗΜΑ BioSite

Το BioSite, ένα σύστημα επεξεργασίας κι ανάλυσης της βιοϊατρικής βιβλιογραφίας, δημιουργήθηκε με σκοπό την υποστήριξη ταχύτερων κι αποτελεσματικών αναζητήσεων στις ηλεκτρονικές βιβλιοθήκες. Ο γενικός τρόπος λειτουργίας του συστήματος είναι η συλλογή κειμένων από την PubMed Central, το περιεχόμενο των οποίων θα είναι δημοσίως διαθέσιμο, και παρέχει στον χρήστη τη δυνατότητα να λαμβάνει άρθρα συσχετιζόμενα άμεσα με τους όρους που χρησιμοποίησε στην αναζήτηση. Στην ουσία, η ιδιότητα που διαφοροποιεί το BioSite από την PubMed Central είναι η απλοποίηση της διαδικασίας αναζήτησης άρθρων με ταχύτερα επιθυμητά προς τον χρήστη αποτελέσματα (Μηνά Ι. 2017).

Στις περισσότερες ηλεκτρονικές βιβλιοθήκες ο χρήστης έχει πρόσβαση μόνο στις βασικές πληροφορίες και την περίληψη του άρθρου. Αντιθέτως, στο BioSite τα επιστρεφόμενα αποτελέσματα αποτελούν ένα υποσύνολο άρθρων, που περιλαμβάνουν κείμενα άμεσα προσβάσιμα στον χρήστη έτσι ώστε να μπορεί να διαβάσει το πλήρες άρθρο στην ιστοσελίδα του συστήματος. Με τον τρόπο αυτό επιτυγχάνεται η μείωση του χρόνου αναζήτησης κι αποφεύγεται η ανίχνευση του επιθυμητού αποτελέσματος σε μία πληθώρα κειμένων (Μηνά Ι. 2017).

Ένας επιπλέον στόχος του συστήματος BioSite είναι η παρουσίαση αποτελεσμάτων με τρόπο που διευκολύνει τον εντοπισμό της επιθυμητής πληροφορίας που αναζητά ο χρήστης αλλά και την περαιτέρω αναζήτηση άρθρων βάση συγκεκριμένων πεδίων. Εφόσον παρουσιαστούν τα επιστρεφόμενα ως αποτελέσματα άρθρα μπορούν εύκολα να εντοπιστούν πληροφορίες για αυτά, όπως ημερομηνία εκδόσεως, τίτλος, συγγραφείς κ.α. . Μία ακόμη λειτουργία του συστήματος BioSite είναι ότι τα πεδία, στα οποία έχει νόημα να παρέχεται, θα αποτελούν υπερσυνδέσμους που θα πραγματοποιούν μία νέα αναζήτηση βάση της τιμής του συγκεκριμένου πεδίου. Δεν θα απαιτείται επανεισαγωγή των τιμών των πεδίων αυτών ή κάποια άλλη διαδικασία, αλλά αυτομάτως θα παρέχονται στον χρήστη τα αποτελέσματα των νέων αυτών αναζητήσεων.

Για την επίτευξη όλων αυτών των στόχων που προαναφέρθηκαν ορίζονται κάποιες λειτουργίες που πρέπει να υλοποιηθούν, ώστε να ικανοποιηθούν οι σκοποί του συστήματος.

Η πρώτη λειτουργία αφορά την αναζήτηση κειμένων στην PubMed Central, τα οποία θα είναι δωρεάν διαθέσιμα και θα προσφέρεται το πλήρες κείμενο τους. Τα κείμενα θα ανακτώνται και θα αποθηκεύονται σε τοπική βάση δεδομένων.

Η δεύτερη λειτουργία αφορά την επεξεργασία τους με εργαλεία επεξεργασίας φυσικής γλώσσας και σχολιασμού τους με κατάλληλο τρόπο, ο οποίος θα βοηθήσει στην περαιτέρω επεξεργασία τους. Ο σχολιασμός θα αφορά βασικές βιοϊατρικές έννοιες, καθώς και πληροφορίες σχετικά με την σύνταξη των προτάσεων και πληροφορίες σχετικά με τις λέξεις μέσα σε αυτές. Η πληροφορία που θα προκύψει, θα χρησιμοποιηθεί για την εξαγωγή των βασικών βιοϊατρικών όρων ως εννοιολογικές οντότητες και των σχέσεων μεταξύ αυτών. Η τρίτη βασική λειτουργία αφορά την δημιουργία της οντολογίας που θα χρησιμεύσει ως βάση αποθήκευσης της επιθυμητής πληροφορίας. Η οντολογία αυτή, θα αποτελεί μία επέκταση με όλα τα νέα απαραίτητα στοιχεία που χρειάζεται το συγκεκριμένο σύστημα, της οντολογίας που θα προκύψει από την συνένωση κάποιων υπάρχοντων οντολογιών στο τομέα της βιοϊατρικής.

Η τέταρτη λειτουργία που απαιτείται στο σύστημα είναι η δημιουργία αλγορίθμου που θα κάνει επεξεργασία των σχολιασμένων κειμένων και θα εισάγει την πληροφορία που θα προκύπτει μετά την επεξεργασία, στην οντολογία – βάση του συστήματος (Μηνά Ι. 2017). Η πέμπτη λειτουργία του συστήματος αφορά την δημιουργία ιστοσελίδας που θα αποτελεί το μέσο με το οποίο θα αλληλεπιδρά ένας χρήστης. Οι λειτουργίες που θα παρέχονται περιλαμβάνουν την υποστήριξη υποβολής διάφορων μορφών ερωτημάτων στην βάση, ανάλογα με τις ανάγκες του και την επιθυμητή μορφή αποτελεσμάτων και στη συνέχεια εμφάνιση των αποτελεσμάτων αυτών.

Η έκτη λειτουργία που απαιτείται είναι η δημιουργία αλγορίθμου που θα επεξεργάζεται το ερώτημα του χρήστη ώστε να το μετατρέπει σε μορφή ερωτήματος που μπορεί να εφαρμοστεί στην οντολογία προκειμένου

να παραχθούν τα αποτελέσματα. Η έβδομη βασική λειτουργία είναι η δημιουργία αλγορίθμου που θα υπολογίζει σε ποια από τις προκαθορισμένες κλάσεις της οντολογίας ανήκει κάθε κείμενο, ώστε να δημιουργηθεί η κατηγοριοποίηση των κειμένων.

Τέλος, η όγδοη βασική λειτουργία είναι η δημιουργία αλγορίθμου που θα χρησιμοποιεί την συγκεκριμένη δομή της οντολογίας και θα εκτελεί συσταδοποίηση των κειμένων σε αυτόματα παραγόμενες συστάδες. Ο αλγόριθμος συσταδοποίησης θα εστιάζει στο περιεχόμενο των κειμένων από βιοϊατρικής άποψης.

Αυτές οι λειτουργίες και στοιχεία μπορούν να ομαδοποιηθούν και να υλοποιηθούν σε ξεχωριστά μέρη – υποσυστήματα του συστήματος, ώστε να είναι αυτόνομα εργαλεία, αναλόγως με την λειτουργικότητά τους. Για παράδειγμα η αναζήτηση και η ανάκτηση των κειμένων μπορεί να γίνει από ένα εργαλείο – προσκομιστή εγγράφων το οποίο θα λειτουργεί ανεξάρτητα από τα άλλα υποσυστήματα και θα έχει την δική του λειτουργικότητα, έτσι ώστε να μπορεί να χρησιμοποιηθεί από το ολοκληρωμένο σύστημα που σχεδιάζουμε, αλλά να μπορεί να χρησιμοποιηθεί και από άλλες εφαρμογές, εφόσον καλύπτει τις απαιτήσεις τους.

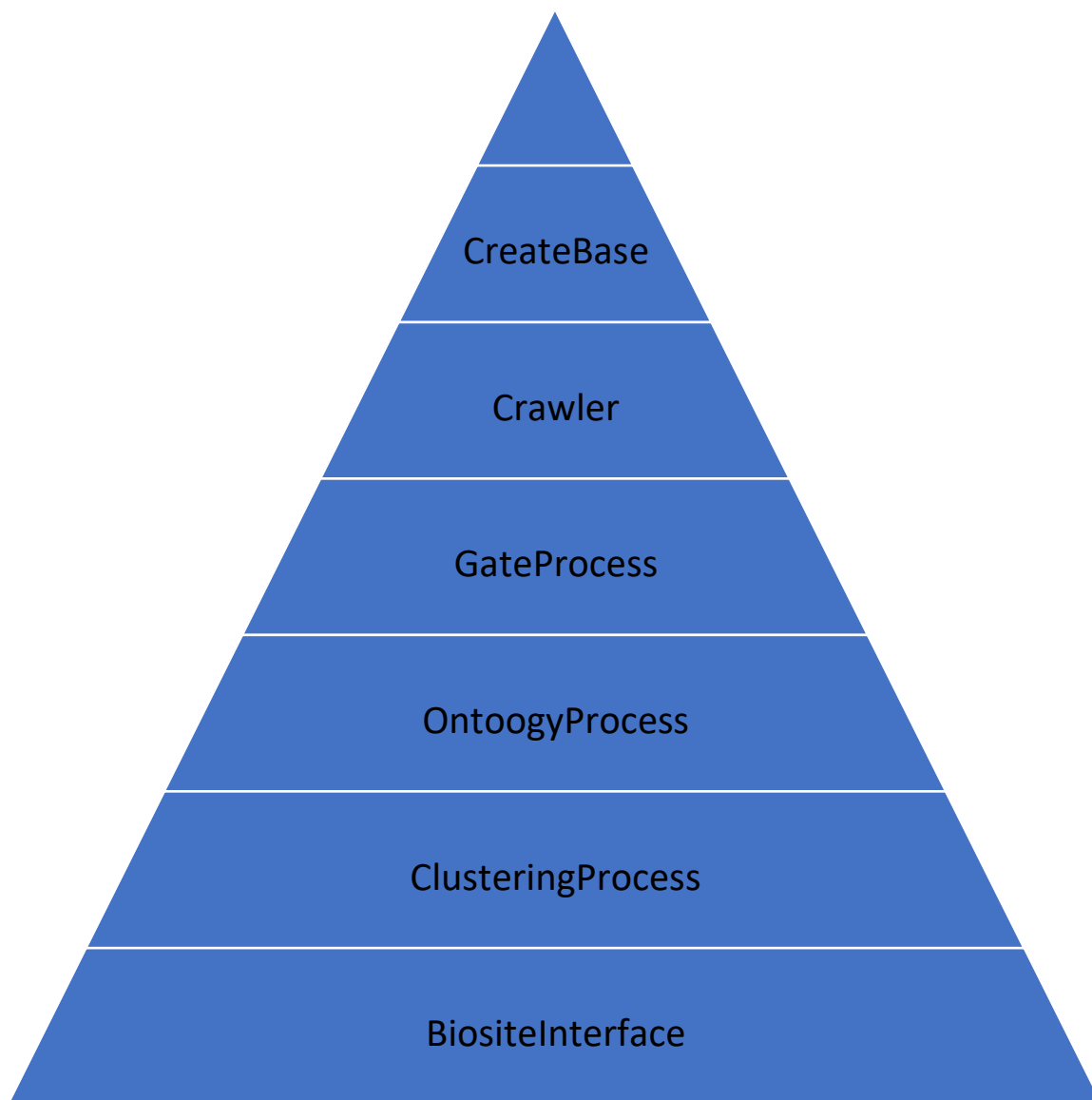
Για αυτόν τον λόγο θα δημιουργηθούν έξι βασικά υποσυστήματα τα οποία θα έχουν το καθένα την δική του λειτουργικότητα και θα υλοποιούν κάποιες από τις παραπάνω λειτουργίες, έτσι ώστε το σύνολο των λειτουργιών να καλύπτεται πλήρως. Τα ονόματα των υποσυστημάτων αυτών παρουσιάζονται στον Πίνακα 1 που ακολουθεί και στην συνέχεια παρουσιάζονται αναλυτικά τα υποσυστήματα.

CreateBase
Crawler
GateProcess
OntologyProcess

ClusteringProcess
BioSiteInterface

Πίνακας 1: Υποσυστήματα του συστήματος BioSite

Το τελικό σύστημα BioSite αποτελείται από την διαδοχική εκτέλεση αυτών των υποσυστημάτων. Η σειρά με την οποία θα εκτελεστεί κάθε υποσύστημα έχει μεγάλη σημασία, αφού κάθε υποσύστημα εκτός του πρώτου, χρειάζεται, για να λειτουργήσει σωστά, ως είσοδό του την πληροφορία που παράγει το προηγούμενο υποσύστημα. Η σειρά με την οποία πρέπει να εκτελεστούν τα υποσυστήματα και από που λαμβάνει πληροφορία ως είσοδο το κάθε ένα, παρουσιάζεται στο Διάγραμμα 1.



Διάγραμμα 1: Διαδοχική εκτέλεση των υποσυστημάτων του BioSite (Μηνά Ι. 2017).

3.2 Υποσύστημα CreateBase

Το CreateBase είναι το πρώτο υποσύστημα και είναι υπεύθυνο για την δημιουργία της οντολογίας – βάσης του συστήματος. Η οντολογία που δημιουργείται είναι το αποτέλεσμα ενοποίησης δύο άλλων οντολογιών και

σε αυτό το αποτέλεσμα προστίθενται κάποιες επιπλέον κλάσεις και ιδιότητες που είναι απαραίτητες για την λειτουργία της οντολογίας και ως βάση του τελικού συστήματος.

Μία οντολογία είναι ένα είδος ελεγχόμενου λεξιλογίου σαφώς καθορισμένων όρων και σχέσεων μεταξύ αυτών των όρων, ώστε να μπορούν να ερμηνευτούν τόσο από ανθρώπους όσο και από υπολογιστές. Στον χώρο της βιοϊατρικής, είναι απαραίτητη η ύπαρξη οντολογιών που θα επικεντρώνονται σε όρους και έννοιες βιοϊατρικής σημασίας. Μία βιοϊατρική οντολογία αναπτύσσεται για να εκφράσει βιοϊατρικά δεδομένα με τρόπο που να μπορούν να διαβαστούν και να επεξεργαστούν από τους υπολογιστές, αλλά και να μπορούν να διαμοιραστούν εύκολα μεταξύ των πηγών και των εφαρμογών που τα χρειάζονται. Είναι ένας κατάλληλος τρόπος αποθήκευσης και εύκολης διαχείρισης των συνεχώς αυξανόμενων δεδομένων, σε αυτή τη νέα εποχή της πληροφορίας (Μηνά Ι. 2017). Έτσι έχουν δημιουργηθεί πολλές βιοϊατρικές οντολογίες οι οποίες χρησιμοποιούνται για την αποθήκευση και την επεξεργασία πληροφορίας και εννοιών στους τομείς της βιολογίας, της βιοπληροφορικής και άλλων τέτοιων πεδίων. Τέτοιες οντολογίες, που δημιουργήθηκαν για την υποστήριξη εργασιών και συστημάτων στον τομέα της βιοϊατρικής, μπορούν να βρεθούν εύκολα στον ιστότοπο National Center for Biomedical Ontology – NCBO , που αποτελεί ένα μέσο πρόσβασης στις περισσότερες βιοϊατρικές οντολογίες καθώς και στις πληροφορίες σχετικά με τα ερευνητικά έργα που τις χρησιμοποιούν.

Παραδείγματα τέτοιων βιοϊατρικών οντολογιών είναι και οι δύο οντολογίες, που ονομάζονται GENIAterm40.owl και GENIAevent.owl. Οι οντολογίες αυτές σχεδιάστηκαν και δημιουργήθηκαν για να υποστηρίξουν

εργασίες σχολιασμού βιοϊατρικού κειμένου και ειδικότερα το GENIA Project .

Η οντολογία GENIAterm40.owl, ορίζει κλάσεις που αποσκοπούν στην αποθήκευση βιοϊατρικών όρων, δηλαδή κλάσεις αποθήκευσης όρων που αποτελούν υποκείμενα ή αντικείμενα σε προτάσεις.

Η GENIAevent.owl ορίζει και ταξινομεί σχέσεις μεταξύ βιοϊατρικών όρων. Με άλλα λόγια, η GENIAevent.owl περιέχει κλάσεις για την αποθήκευση σχέσεων δηλαδή κατηγορημάτων που εκφράζονται συνήθως από τα ρήματα μιας πρότασης.

Τα δεδομένα που αποθηκεύονται σε αυτές τις οντολογίες προέρχονται από προτάσεις κειμένων βιοϊατρικού χαρακτήρα και χρησιμοποιούνται για την υποστήριξη ερευνητικών έργων, σχολιασμών κειμένων κ.ά., σε αυτόν τον τομέα.

Για την υποστήριξη του συστήματος BioSite δημιουργήθηκε μία νέα βιοϊατρική οντολογία, η οποία έχει ως σκοπό της να αποτελέσει το μέσο αποθήκευσης των δεδομένων που θα προκύψουν από την επεξεργασία των βιοϊατρικών κειμένων, αλλά επίσης να αποτελέσει και την βάση δεδομένων που απαιτείται για την υποστήριξη και λειτουργία της ιστοσελίδας του συστήματος BioSite. Η νέα αυτή οντολογία είναι σχεδιασμένη και υλοποιημένη έτσι, ώστε να μπορεί να συγκρατήσει πληροφορία και δεδομένα βιοϊατρικού περιεχομένου που προκύπτουν μετά την επεξεργασία των άρθρων, πληροφορία σχετικά με την ομοιότητα και τη σχετικότητα των άρθρων, πληροφορία για την κατηγοριοποίηση και συσταδοποίησή τους, αλλά και πληροφορία που αφορά τους χρήστες, για την υποστήριξη και παροχή υπηρεσιών σε αυτούς (Μηνά Ι. 2017).

Πιο συγκεκριμένα, στην οντολογία αυτή αποθηκεύονται στις αντίστοιχες κλάσεις, τα ονόματα των άρθρων και των δημοσιεύσεων που χρησιμοποιούνται από το

σύστημα. Αποθηκεύονται επίσης οι όροι των κειμένων, αλλά και όροι, οι οποίοι μετά την επεξεργασία των κειμένων αυτών, αναγνωρίστηκαν ως βιοϊατρικοί όροι. Με τη βοήθεια των κατάλληλων ιδιοτήτων αποθηκεύεται πληροφορία σχετικά με τις σχέσεις μεταξύ των άρθρων και των όρων, των όρων μεταξύ τους, αλλά και πληροφορία σχετικά με την ομοιότητα των άρθρων. Ακόμη, αποθηκεύεται πληροφορία σχετικά με τις προτάσεις των κειμένων και το ποιους όρους αυτές περιλαμβάνουν και επίσης πληροφορία σχετικά με τα άρθρα όπως ο τίτλος τους, οι συγγραφείς κ.ά.. Τέλος, αποθηκεύεται πληροφορία σχετικά με τους χρήστες της ιστοσελίδας του συστήματος.

Για τον λόγο αυτό, δημιουργείται η νέα οντολογία την συνένωση των δύο οντολογιών GENIAterm40.owl (Διάγραμμα 2) και GENIAevent.owl (Διάγραμμα 3) σε μία οντολογία που περιέχει όλες τις κλάσεις που περιέχουν και αυτές, διατηρώντας την ιεραρχία σε περίπτωση κοινών κλάσεων. Η νέα οντολογία στην συνέχεια προσαυξάνεται με νέες κλάσεις και ιδιότητες για την υποστήριξη των απαιτούμενων λειτουργιών.

Στο υποσύστημα CreateBase γίνεται η δημιουργία της οντολογίας Αυτής. Οι δύο αρχικές οντολογίες συγχωνεύονται σε μία, με την νέα οντολογία να περιέχει όλες τις κλάσεις των άλλων δύο οντολογιών με την ίδια ιεραρχία. Στην περίπτωση κοινών κλάσεων, αυτές αντιπροσωπεύονται από μία, που διατηρεί τις σχέσεις τους με τις κλάσεις και των δύο αρχικών οντολογιών.

Κάποιες από τις κλάσεις αυτές δεν χρησιμοποιούνται, με την έννοια δεν λαμβάνουν κανένα στιγμιότυπο, ούτε μετά από την χρήση κάποιου εργαλείου εξαγωγής συμπερασμάτων. Συμπεριλήφθηκαν όμως για λόγους πληρότητας, όπως επίσης και για την περίπτωση μελλοντικών αλλαγών ή βελτιώσεων του συστήματος BioSite.

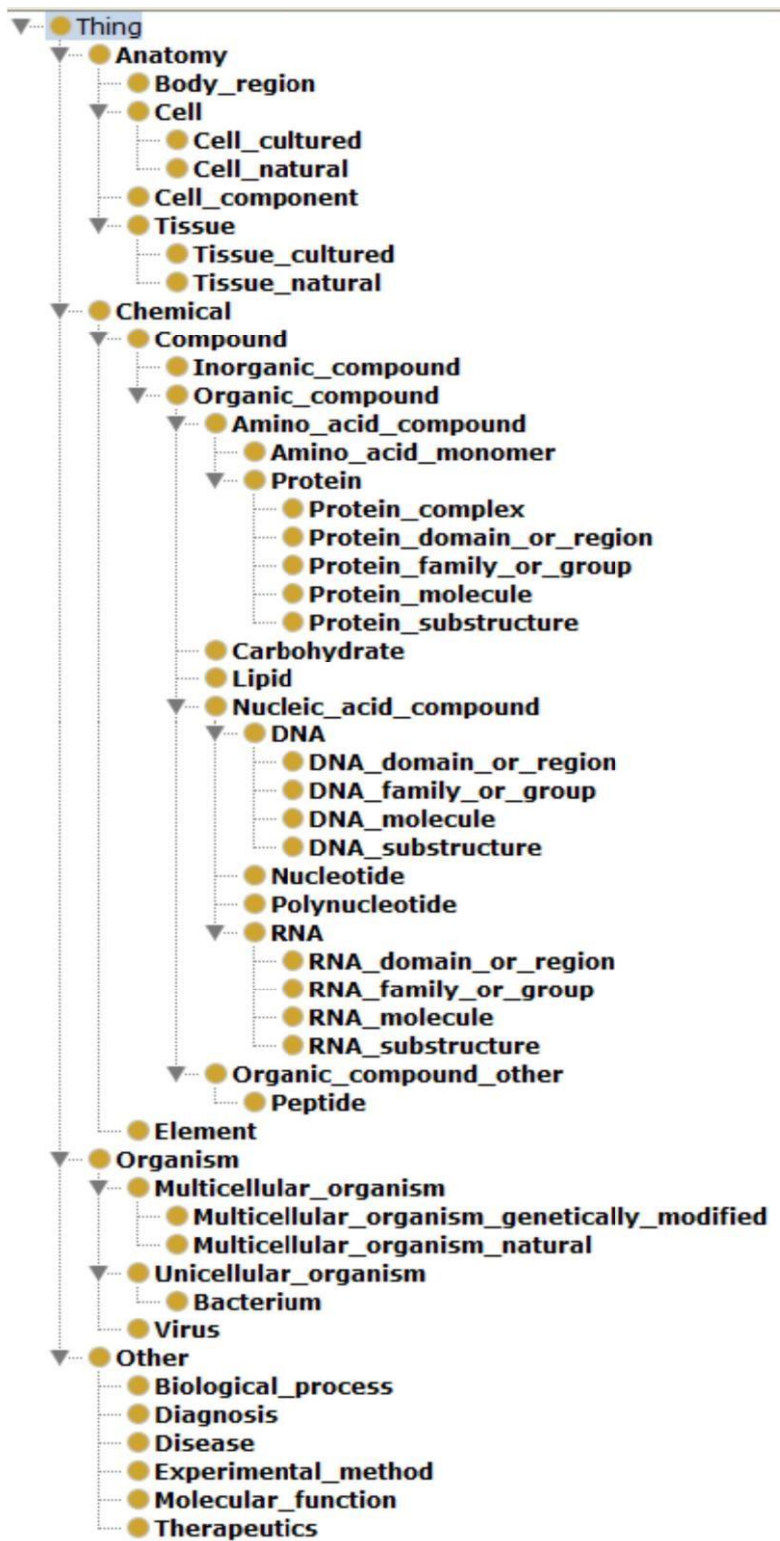
Στις προκύπτουσες κλάσεις, αποθηκεύονται ως στιγμιότυπα ολοι όροι των άρθρων, που μετά από την επεξεργασία τους από τα κατάλληλα εργαλεία,

αποφασίζεται ότι έχουν σημασία από βιοϊατρικής άποψης και χρησιμεύουν στην εξαγωγή γνώσης και πληροφορίας από το σύστημα.

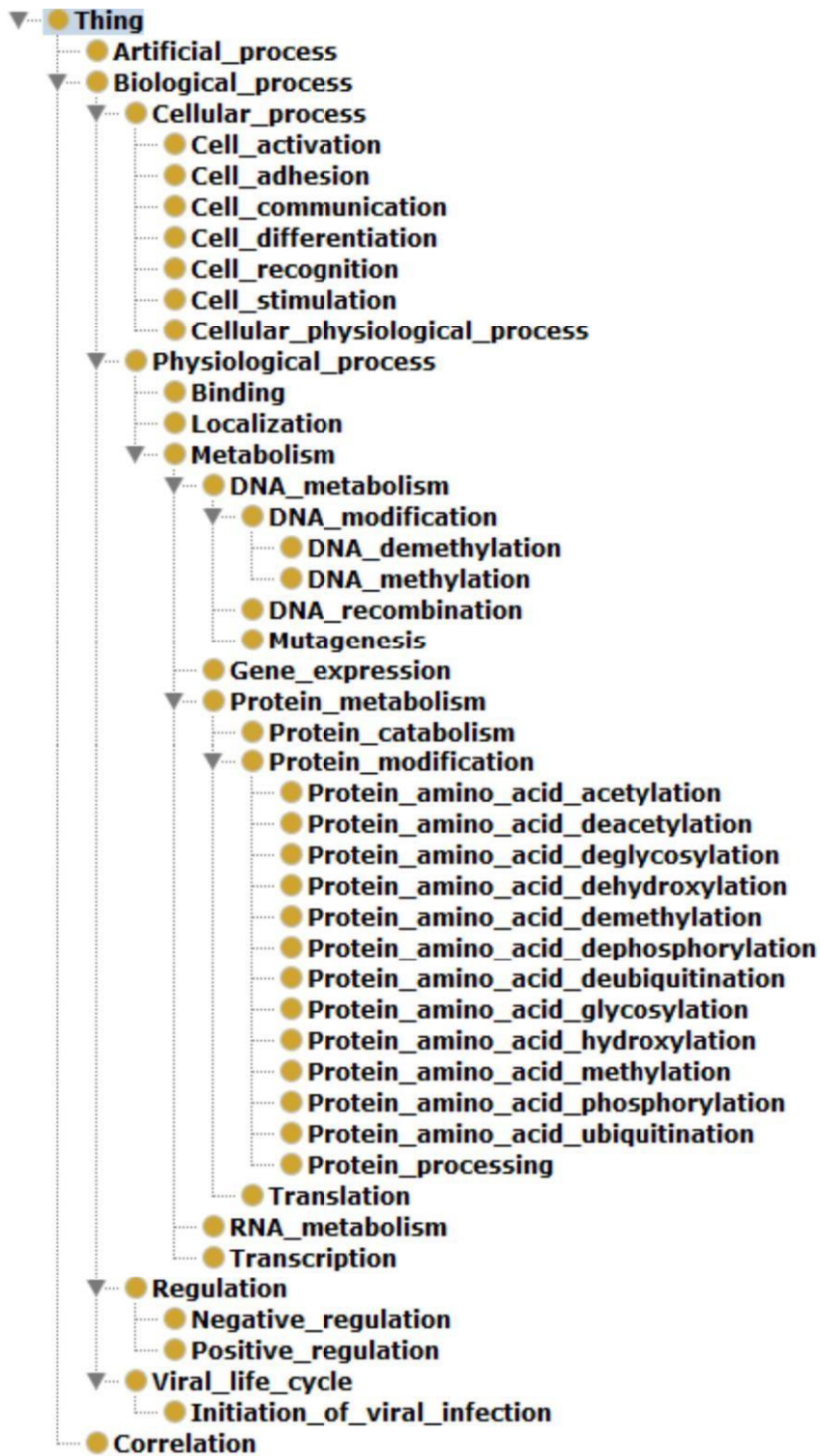
Ένας βιοϊατρικός όρος, όπως ονομάζεται ένας τέτοιος όρος, μπορεί να αποτελείται από μία ή περισσότερες λέξεις.

Υπάρχει επίσης η προσθήκη κάποιων νέων κλάσεων, οι οποίες είναι απαραίτητες για την σωστή λειτουργία του συστήματος και την χρήση της οντολογίας ως βάση με την οποία θα αλληλεπιδρά και η ιστοσελίδα του. Παράδειγμα αποτελεί η κλάση που χρησιμεύει για την αποθήκευση πληροφορίας σχετικά με τις συστάδες που χρησιμοποιούνται από τον αλγόριθμο συσταδοποίησης, αλλά και η κλάση στην οποία αποθηκεύεται πληροφορία σχετικά με τους χρήστες της ιστοσελίδας του συστήματος. Όλες οι κλάσεις που προστίθενται στην οντολογία που προκύπτει μετά την συνένωση των GENIAterm40.owl και GENIAevent.owl και οδηγούν στην τελική μορφή της νέας οντολογίας, περιγράφονται στον Πίνακα 2, ενώ στον Πίνακα 3 περιγράφονται οι απαραίτητες ιδιότητες που χρησιμεύουν για την αποθήκευση της απαιτούμενης πληροφορίας. Η νέα οντολογία ονομάζεται BioSiteOntology.owl (Διάγραμμα 4). Εδώ αποθηκεύεται όλη η πληροφορία που εξάγεται από την επεξεργασία των άρθρων, κάτι που γίνεται με την βοήθεια των επόμενων υποσυστημάτων. Αποθηκεύεται επίσης και η πληροφορία που προκύπτει από την εφαρμογή του αλγορίθμου συσταδοποίησης και από την χρήση της ιστοσελίδας του συστήματος από τους χρήστες (Μηνά Ι. 2017).

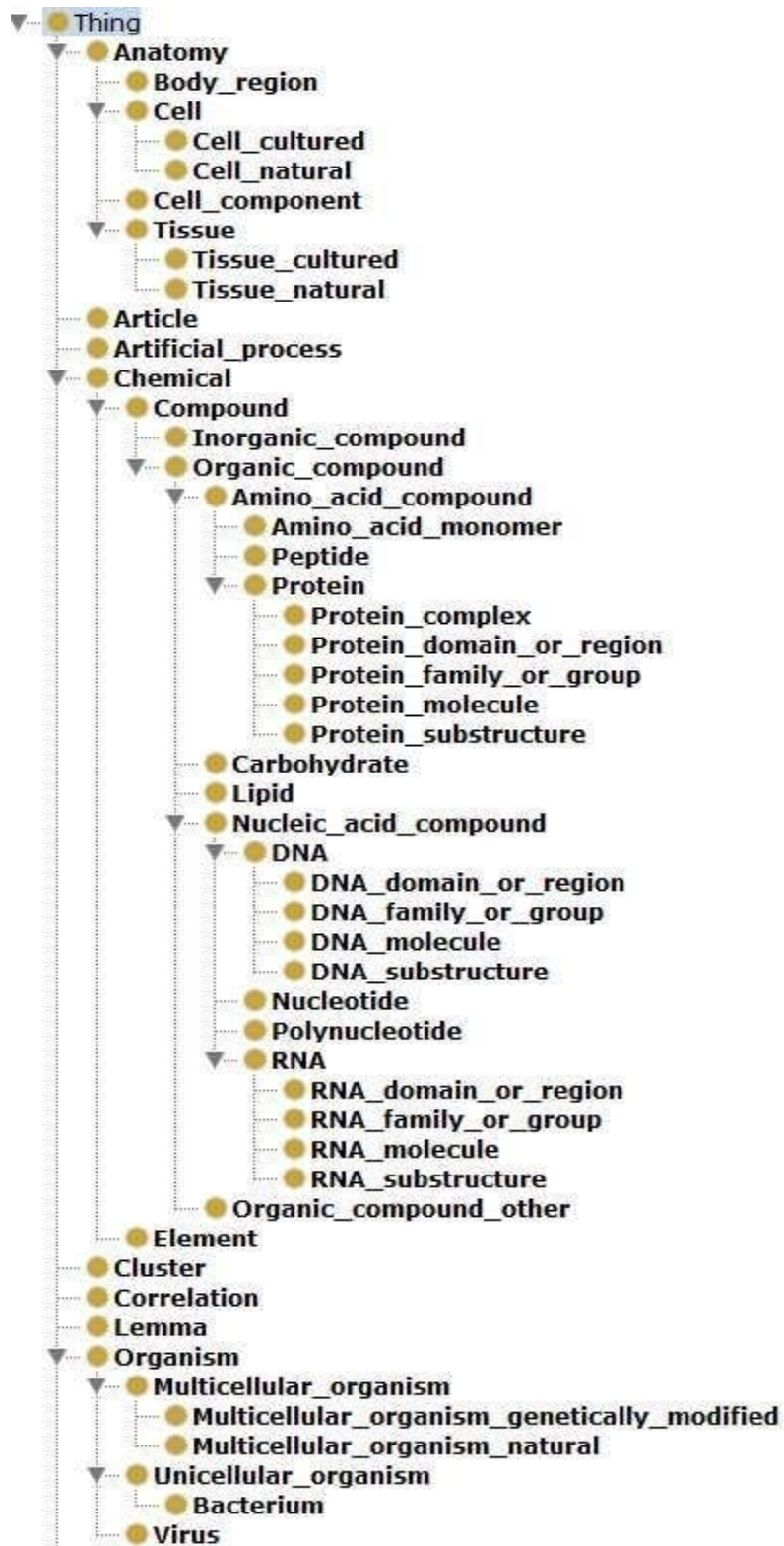
Ακολουθούν οι εικόνες που δείχνουν τις ιεραρχίες όλων των οντολογιών και οι πίνακες που αναφέρουν τις νέες κλάσεις και ιδιότητες και περιέχουν μια σύντομη περιγραφή τους.

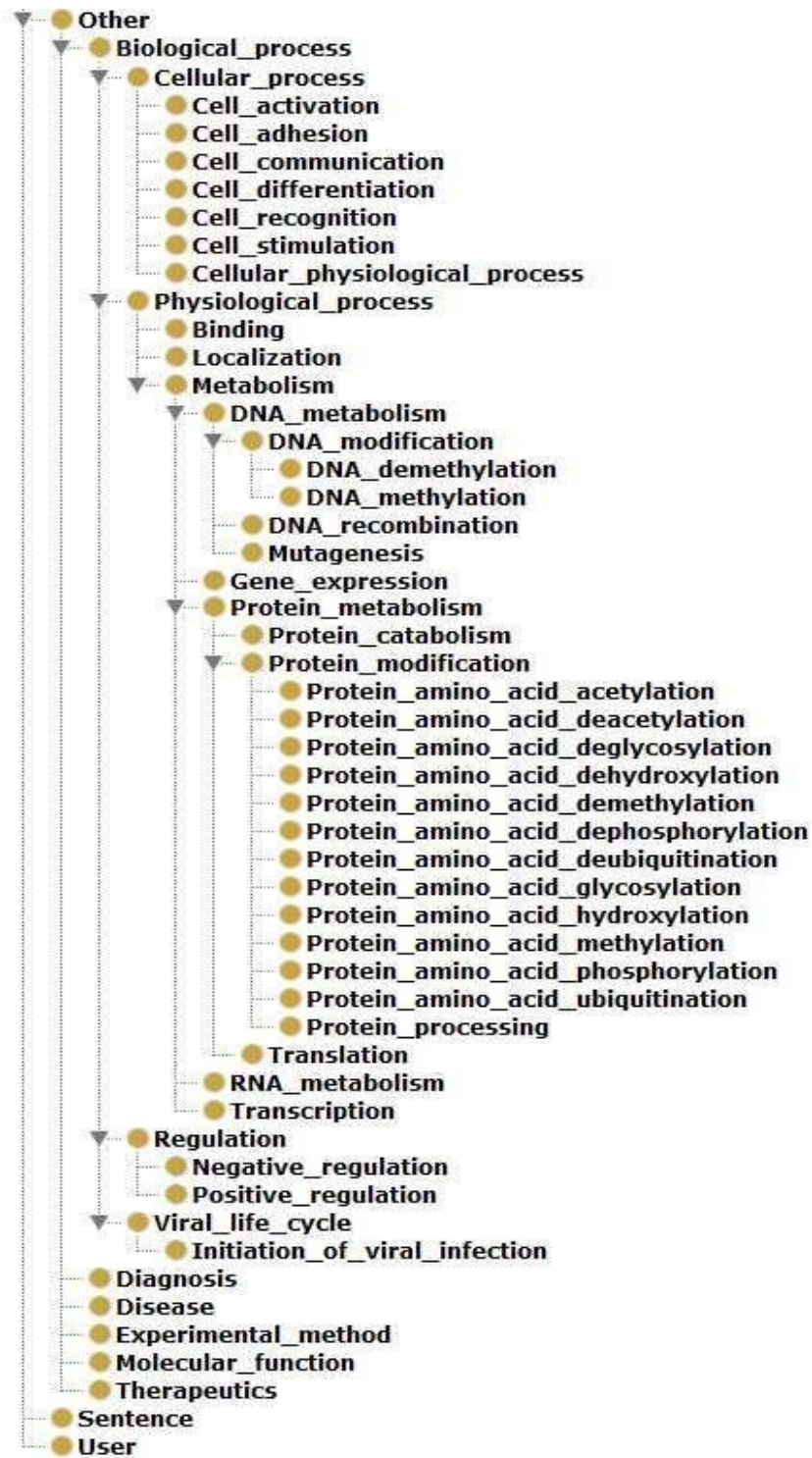


Διάγραμμα 2: Ιεραρχία των κλάσεων της GENIAterm40.owl οντολογίας



Διάγραμμα 3: Ιεραρχία των κλάσεων της GENIAevent.owl οντολογίας





Διάγραμμα 4: Ιεραρχία των κλάσεων της BioSiteOntology.owl οντολογίας

Κλάση		Περιγραφή
Article		Χρησιμοποιείται για την αποθήκευση των στιγμιότυπων που αντιπροσωπεύουν τα ids των άρθρων που επεξεργαζόμαστε, για την εξαγωγή πληροφορίας
Sentence		Χρησιμοποιείται για την αποθήκευση των στιγμιότυπων που αντιπροσωπεύουν τα ids των προτάσεων που υπάρχουν στα άρθρα, και τα οποία χρειαζόμαστε για την εξαγωγή πιο ειδικών σχέσεων ανάμεσα στους όρους των άρθρων
Lemma		Χρησιμοποιείται για την αποθήκευση των στιγμιότυπων που αντιπροσωπεύουν τους όρους των κειμένων

Cluster		Χρησιμοποιείται για την αποθήκευση των στιγμιotypων που αντιπροσωπεύουν τις ομάδες – clusters που θα προκύψουν από την εφαρμογή του
---------	--	---

		αλγόριθμου clustering και θα συνδέουν τα σχετικά μεταξύ τους άρθρα
--	--	--

User		Χρησιμοποιείται για την αποθήκευση των στιγμιotypων που αντιπροσωπεύουν τους χρήστες της οντολογίας, όταν αυτοί κάνουν εγγραφή στο σύστημα
------	--	--

Πίνακας 2: Κλάσεις που προστέθηκαν στην BioSiteOntology.owl (Μηνά Ι. 2017).

Τύπου Object	Περιγραφή
inCluster	Σύνδεση άρθρων και των συστάδων στις οποίες θα ανήκουν ως ομάδες σχετικών άρθρων. Δείχνει σε ποιες συστάδες ανήκει ένα άρθρο
clusterHas	Σύνδεση συστάδων και άρθρων. Δείχνει ποια άρθρα περιέχει κάποια συστάδα
articleHasWord	Σύνδεση άρθρων και όρων. Δείχνει ποιους όρους περιέχει κάποιο άρθρο
wordIsFromArticle	Σύνδεση όρων και άρθρων στα οποία ανήκουν. Δείχνει σε ποια άρθρα εμφανίζεται κάποιος όρος
sentenceHasWord	Δείχνει ποιους όρους περιέχει κάποια πρόταση ενός άρθρου
inSentence	Σύνδεση όρων και προτάσεων των άρθρων στις οποίες αυτοί ανήκουν. Δείχνει σε ποιες προτάσεις εμφανίζεται κάποιος όρος

hasReadArticle	Αποθήκευση πληροφορίας σχετικά με τα άρθρα που έχει διαβάσει ένας χρήστης, με σκοπό να χρησιμοποιηθεί στις λειτουργίες της ιστοσελίδας για τους εγγεγραμμένους χρήστες

Πίνακας 3: Ιδιότητες τύπου *Object* που προστέθηκαν στην *BioSiteOntology.owl*

Τύπου Datatype	Περιγραφή
hasAuthor	Αποθήκευση των συγγραφέων ενός άρθρου
hasPubYear	Αποθήκευση της ημερομηνίας έκδοσης ενός άρθρου
hasTitle	Αποθήκευση του τίτλου ενός άρθρου
hasTag	Αποθήκευση της τιμής που επέστρεψε ο GENIAtagger σχετικά με τον σχολιασμό ενός όρου
bioTerm	Αποθήκευση του ονόματος ενός βιοϊατρικού όρου με σκοπό την γρηγορότερη και ευκολότερη χρήση αυτού του όρου όπου χρειάζεται

password	Αποθήκευση του password ενός χρήστη που επιθυμεί να κάνει εγγραφή στην ιστοσελίδα του συστήματος
email	Αποθήκευση του email ενός χρήστη που επιθυμεί να κάνει εγγραφή στην ιστοσελίδα του συστήματος

Πίνακας 4: Ιδιότητες τύπου Datatype που προστέθηκαν στην BioSiteOntology.owl

3.3 Υποσύστημα Crawler

Το Crawler είναι το δεύτερο στη σειρά υποσύστημα. Η βασική λειτουργία του αφορά τόσο την συλλογή άρθρων – κειμένων από την online βάση PubMed Central, όσο και μία αρχική επεξεργασία τους. Όσον αφορά τη συλλογή των κειμένων, υπάρχει ένας έλεγχος σχετικά με ποια από όλα τα άρθρα θα επιλεγούν για αποθήκευση, όχι βάση του περιεχομένου τους και του τομέα της βιοϊατρικής στον οποίο ανήκουν, αλλά βάση του τύπου του άρθρου και του λόγου δημιουργίας του. Όσον αφορά την επεξεργασία, αφορά μία πρώτη επιλογή των μερών των άρθρων που θα είναι χρήσιμα για το σύστημα (Μηνά Ι. 2017).

Αναλυτικότερα οι λειτουργίες του υποσυστήματος είναι αρχικά ο έλεγχος εάν υπάρχει κατάλληλο μέρος αποθήκευσης των άρθρων που πρόκειται να προσκομμιστούν. Εάν δεν υπάρχει, δημιουργείται ο χώρος, ο οποίος θα χρησιμεύσει για την αποθήκευση των άρθρων που θα επιλεγθούν, για την

αποθήκευση της οντολογίας και των άλλων εργαλείων που χρειάζεται το σύστημα BioSite και τέλος για την αποθήκευση των απαραίτητων ενδιάμεσων αποτελεσμάτων (μετά την επεξεργασία των κειμένων για παράδειγμα).

Αφού δημιουργηθεί ο χώρος αυτός, το υποσύστημα αρχίζει την ανάκτηση των άρθρων και δημοσιεύσεων από την PubMed Central. Αφού οριστούν οι επιθυμητές παράμετροι, όπως για παράδειγμα το πόσα άρθρα θέλουμε να επιστρέφονται σε κάθε εκτέλεση, το υποσύστημα ανακτά ώστε να αποθηκεύσει στην τοπική βάση το κείμενο και τις πληροφορίες κάθε άρθρου. Το κάθε άρθρο που ανακτάται, ανήκει στην κατηγορία άρθρων των οποίων το πλήρες κείμενό τους είναι διαθέσιμο δημόσια δωρεάν. Στην συνέχεια ελέγχεται για κάποια αρχικά χαρακτηριστικά και συνθήκες οι οποίες πρέπει να ισχύουν.

Ο έλεγχος αυτός, αφορά αρχικά το είδος του άρθρου και το εάν ανήκει σε κάποιο από τα επιτρεπτά είδη. Αυτό γίνεται με την βοήθεια των ετικετών που παρέχονται από την PubMed Central Tagging . Αυτές οι ετικέτες ορίζουν το είδος των άρθρων βάση του περιεχομένου και του σκοπού για τον οποίο δημιουργήθηκε. Τα επιτρεπτά είδη, δηλαδή τα είδη των οποίων η επεξεργασία έχει λογικό νόημα και σημασία για τους σκοπούς του συστήματος BioSite, περιγράφονται στον Πίνακα 5 (Μηνά Ι. 2017).

Ετικέτα	Περιγραφή
Abstract	Περιλήψεις άρθρων
Research-article	Άρθρα ερευνητικού σκοπού και περιεχομένου
Oration	Παρουσιάσεις ομιλιών από ερευνητικά συνέδρια

Brief-report	Σύντομες αναφορές ερευνητικών έργων
Case-report	Αναφορές ερευνητικών έργων
Addendum	Προσθήκες νέων ερευνητικών αποτελεσμάτων σε παλαιότερα ερευνητικά έργα
News	Παρουσιάσεις νέων ευρημάτων

Πίνακας 5: Επιτρεπτά είδη άρθρων - δημοσιεύσεων

Τα μη επιτρεπτά είδη αφορούν για παράδειγμα κριτικές βιβλίων, δημοσιεύσεων και έργων, διορθώσεις άλλων άρθρων, πρακτικά συζητήσεων ερευνητών κ.ά.

Εάν το άρθρο ικανοποιεί τις προδιαγραφές, τότε αποθηκεύεται στην τοπική βάση ένα αντίγραφο του πλήρους κειμένου του, από το οποίο θα εξαχθούν οι επιθυμητές πληροφορίες.

Οι λειτουργίες του υποσυστήματος, εκτός από τον πρώτο έλεγχο και ανάκτηση των κειμένων, περιλαμβάνουν και κάποιες ακόμη. Αρχικά, σε κάθε εκτέλεση του υποσυστήματος υπάρχει έλεγχος σχετικά με το ποια άρθρα έχουν ήδη ανακτηθεί, ώστε να αποφύγουμε την επεξεργασία ήδη ανεκτημένων και επεξεργασμένων άρθρων (Μηνά Ι. 2017). Στην συνέχεια, αφού ανακτηθούν τα άρθρα και το πλήρες περιεχόμενό τους αποθηκευτεί στη τοπική βάση, υπάρχει μία πρώτη

ανάλυση τους και αποθήκευση κάποιων πληροφοριών στην οντολογία BioSiteOntology.owl του συστήματος. Οι πληροφορίες αυτές αφορούν τον τίτλο, τους συγγραφείς και την ημερομηνία έκδοσης και εισάγονται ως τιμές στις αντίστοιχες ιδιότητες κάθε άρθρου στην οντολογία, για το οποίο αποθηκεύεται ταυτόχρονα και ένα μοναδικό αναγνωριστικό, για γρήγορη πρόσβαση σε αυτό.

Για να επιτύχουμε γρήγορη και εύκολη επεξεργασία των άρθρων στα επόμενα υποσυστήματα, το παρόν υποσύστημα εκτελεί μία ακόμη λειτουργία. Από κάθε άρθρο, επιλέγεται μόνο το καθαρό κείμενο – σώμα του και αποθηκεύεται σε ξεχωριστό αρχείο. Έτσι δεν περιλαμβάνεται πληροφορία όπως βιβλιογραφία, ή εξωτερικοί σύνδεσμοι ή άλλα δεδομένα (τα οποία όμως υπάρχουν στη τοπική βάση σε άλλα αρχεία). Με αυτόν τον τρόπο, κάποια από τα επόμενα υποσυστήματα μειώνουν τον χρόνο επεξεργασίας και ανάλυσης των κειμένων, αφού δεν χρειάζεται να ασχοληθούν και να διατρέξουν μη χρήσιμη για τους σκοπούς τους, πληροφορία.

Έτσι, το υποσύστημα σε κάθε εκτέλεση δημιουργεί έναν νέο αριθμό από έγγραφα, που αφορούν μόνο επιτρεπτά είδη άρθρων και περιέχουν μόνο απλό κείμενο προς επεξεργασία. Τα κείμενα αυτά αποθηκεύονται ώστε να είναι προσβάσιμα από τα υπόλοιπα υποσυστήματα που θα τα χρησιμοποιήσουν για επεξεργασία. Οι απαραίτητες πληροφορίες όπως συγγραφείς κ.ά. αποθηκεύονται στην οντολογία για μετέπειτα χρήση από την ιστοσελίδα του συστήματος.

Το παρόν υποσύστημα υλοποιήθηκε με τέτοιο τρόπο ώστε η λειτουργία του να μπορεί να αλλάξει εύκολα ανάλογα με τις εκάστοτε ανάγκες αυτού που το χρησιμοποιεί. Έτσι, ενώ δημιουργήθηκε για να καλύπτει τις

ανάγκες του συστήματος Biosite, είναι πλήρως παραμετροποιημένο έτσι ώστε να μπορεί να προσαρμοστεί ανάλογα η λειτουργία του.

Οι παράμετροι αυτοί περιλαμβάνουν τον αριθμό των επιθυμητών άρθρων, τον έλεγχο ή μη του τύπου αυτών των άρθρων, την επιλογή άλλων ή και όλων των υπαρχόντων τύπων που θα αποδέχεται, αλλά και την μορφή αποθήκευσης των άρθρων.

Υπάρχει επίσης η επιλογή των πεδίων που θα ανακτηθούν από κάθε άρθρο, όπως για παράδειγμα μόνο η περίληψη ή η βιβλιογραφία, αλλά και η επιλογή ολόκληρου του άρθρου με όλες τις πληροφορίες και τα μεταδεδομένα που περιέχει. Αυτό σημαίνει ότι το υποσύστημα μπορεί να χρησιμοποιηθεί και για την απόκτηση όλων των δημόσια δωρεάν άρθρων που παρέχει η PubMed Central και την αποθήκευσή τους, χωρίς αυτά να έχουν υποστεί καμία αλλαγή στο περιεχόμενό τους (δεδομένα και μεταδεδομένα) και χωρίς έλεγχο του τύπου τους. Άρα σε περίπτωση που κάποιος το επιθυμεί, μπορεί να χρησιμοποιήσει το υποσύστημα ως αυτόνομο εργαλείο, το οποίο θα επιστρέφει τους επιθυμητούς τύπους άρθρων και θα αποθηκεύει σε τοπική βάση τον επιθυμητό αριθμό από άρθρα, σε επιθυμητή μορφή (Μηνά Ι. 2017).

3.4 Υποσύστημα GateProcess

Το GateProcess είναι το τρίτο υποσύστημα του συστήματος BioSite. Ο σκοπός αυτού του υποσυστήματος είναι να επεξεργαστεί τα κείμενα των άρθρων που ανακτήθηκαν από το προηγούμενο υποσύστημα και να παρέχει κατάλληλο σχολιασμό του περιεχομένου τους, των προτάσεων και των όρων τους. Ο σχολιασμός αφορά παροχή πληροφοριών σχετικά με την συντακτική και λεξικολογική ανάλυση των κειμένων, αλλά και πληροφορίες όσον αφορά την εμφάνιση όρων και εννοιών που έχουν σχέση

με την βιοϊατρική και αποτελούν έτσι το ουσιαστικό αντικείμενο της ανάλυσης και επεξεργασίας που γίνεται.

Έτσι με την εκτέλεση αυτού του εργαλείου, προστίθεται στο κείμενο κάθε άρθρου, πληροφορία που περιγράφει τις βασικές έννοιες του κειμένου, το ποιοι όροι είναι σημαντικοί από βιοϊατρική άποψη, ποιοι όροι αφορούν οντότητες της βιοϊατρικής και σε ποιο συγκεκριμένο πεδίο αναφέρονται ή ανήκουν.

Για να υλοποιηθεί ο σχολιασμός των κειμένων και η εύρεση των σημαντικών όρων από βιοϊατρικής άποψης, ώστε η πληροφορία αυτή να μπορεί να χρησιμοποιηθεί για τους σκοπούς του συστήματος BioSite, απαιτείται η χρήση κάποιων εργαλείων. Η λειτουργία του παρόντος υποσυστήματος αφορά την εκτέλεση και κατάλληλη χρήση αυτών των εργαλείων, καθώς και τον χειρισμό των αποτελεσμάτων τους με μετατροπή τους σε κατάλληλη μορφή, ώστε να μπορούν να χρησιμοποιηθούν από τα επόμενα υποσυστήματα.

Έτσι, στο παρόν υποσύστημα υλοποιείται η κατάλληλη διασύνδεση με το GATE API για να επιτραπεί η απαραίτητη χρήση του GENIA tagger, μέσα από το Genia application. Το Genia application είναι ένα εργαλείο που δημιουργήθηκε μέσω του εργαλείου GATE, ειδικά για χρήση από το BioSite σύστημα. Η λειτουργία του αφορά την επεξεργασία κειμένων με χρήση του GENIA tagger και άλλων απαραίτητων εργαλείων. Το Genia application μπορεί να ενσωματωθεί εύκολα σε εφαρμογές ανάλυσης βιοϊατρικών κειμένων και να χρησιμοποιηθεί σε συνδυασμό με το GATE Embedded, το οποίο αποτελεί ένα API για το GATE που μπορεί να χρησιμοποιηθεί από εφαρμογές JAVA. Η δομή του Genia application φαίνεται στον Πίνακα 6.

Το Genia Application είναι υπεύθυνο για την διασύνδεση των επιμέρους εργαλείων από τα οποία αποτελείται, την διαδοχική εκτέλεσή τους, την σωστή ρύθμιση των παραμέτρων εισόδου όπως το είδος κειμένου που θα δέχεται για επεξεργασία και το είδος των σχολιασμών. Επίσης περιέχει μέσα του όλα τα απαραίτητα αρχεία και εργαλεία που απαιτούνται για την εκτέλεσή του, έτσι ώστε να είναι πλήρως μεταφέρσιμο και διαχειρίσιμο όχι μόνο από το σύστημα BioSite, αλλά και οποιαδήποτε άλλη εφαρμογή χρειάζεται την χρήση του για επεξεργασία και σχολιασμό κειμένων σε βιοϊατρικό επίπεδο (Μηνά Ι. 2017).

Επιμέρους εργαλεία	Περιγραφή
Document Reset PR	Προετοιμασία ενός εγγράφου για την επερχόμενη επεξεργασία του.
Gate Unicode Tokeniser	Διαχωρισμός του κειμένου των εγγράφων στις λέξεις από τις οποίες αποτελείται.
ANNIE Sentence Splitter	Διαχωρισμός του κειμένου των εγγράφων στις προτάσεις από τις οποίες αποτελείται.
GENIATagger-NoTokenization	Το βασικό στοιχείο του Genia application για τον σχολιασμό, με βιοϊατρικού περιεχομένου και σημασίας ετικέτες, των όρων του κειμένου των εγγράφων.

Με τη χρήση των παραπάνω εργαλείων, γίνεται η επεξεργασία του κειμένου των άρθρων που επέστρεψε ο προσκομιστής εγγράφων του προηγούμενου υποσυστήματος, με σκοπό τον κατάλληλο σχολιασμό τους.

Έτσι, το GateProcess υποσύστημα εκτελεί αρχικά την διασύνδεση με τα απαραίτητα εργαλεία που αναλύθηκαν παραπάνω. Στην συνέχεια αναζητά τα άρθρα που υπάρχουν στην βάση δεδομένων αλλά δεν έχουν ήδη υποστεί την απαραίτητη επεξεργασία. Αφού τα ανακαλύψει, τα επεξεργάζεται μέσω των κατάλληλων εργαλείων και δημιουργεί νέα αρχεία, τα οποία περιέχουν τα κείμενα των άρθρων εμπλουτισμένα με την νέα πληροφορία που παρέχει ο σχολιασμός τους.

Άρα μετά την εκτέλεση του GateProcess υποσυστήματος, έχουν προστεθεί στην βάση δεδομένων οι σχολιασμένες εκδόσεις των άρθρων, οι οποίες θα χρησιμοποιηθούν ως είσοδος για το επόμενο υποσύστημα.

3.5 Υποσύστημα OntologyProcess

Το OntologyProcess είναι το τέταρτο υποσύστημα. Είναι υπεύθυνο για την επεξεργασία των σχολιασμένων κειμένων των άρθρων, την εξαγωγή πληροφορίας μέσω αυτής της επεξεργασίας και την εισαγωγή των αποτελεσμάτων στην οντολογία του συστήματος. Η πληροφορία που θα εισαχθεί μετά από κατάλληλη τροποποίηση στην οντολογία, θα χρησιμοποιηθεί από την ιστοσελίδα του συστήματος για την παρουσίαση των αποτελεσμάτων και την υποστήριξη ερωτημάτων από τους χρήστες.

Ο σκοπός αυτού του υποσυστήματος λοιπόν, είναι να βρεθούν και να αναγνωριστούν οι έννοιες και όροι που έχουν σημασία σε βιοϊατρικό επίπεδο μέσω

της ανάλυσης και της χρήσης των σχολιασμών που δημιουργήσε το προηγούμενο υποσύστημα, για τα κείμενα κάθε άρθρου. Στη συνέχεια, αφού βρεθούν αυτές οι έννοιες, πρέπει να κατασκευαστεί και η απαραίτητη πληροφορία που τις περιγράφει και να μετατραπεί σε κατάλληλη μορφή ώστε να είναι δυνατή η εισαγωγή της στην οντολογία του συστήματος.

Για κάθε κείμενο άρθρου, η πληροφορία που θα εισαχθεί στην οντολογία μετά την ανάλυση των σχολιασμών, είναι πληροφορία σχετικά με τους βασικούς όρους που περιέχει κάθε κείμενο, τους όρους που περιέχει και αποτελούν βιοϊατρικές έννοιες και τις προτάσεις από τις οποίες αποτελείται το κείμενο. Αφορά επίσης το ποιοι όροι, βιοϊατρικοί και μη, ανήκουν σε κάθε κείμενο και σε κάθε πρόταση, ώστε να είναι εύκολη η εξαγωγή συμπερασμάτων σχετικά με τις σχέσεις μεταξύ όρων και άρα και μεταξύ κειμένων, στα επόμενα υποσυστήματα.

Επιπλέον, για κάθε βιοϊατρικό όρο, δηλαδή όρο που αντιστοιχεί σε βιοϊατρική έννοια, καθορίζεται και η κλάση στην οποία θα εισαχθεί βάση της έννοιας και του πεδίου που αντιπροσωπεύει. Έτσι, είναι δυνατός και ο χαρακτηρισμός ενός κειμένου, αναλόγως των κλάσεων στις οποίες ανήκουν οι βιοϊατρικοί όροι που περιέχει. Επίσης, η πληροφορία σχετικά με το σε ποια κλάση ανήκει κάθε βιοϊατρικός όρος, θα χρησιμεύσει και στην διαδικασία της συσταδοποίησης, η οποία θα εκμεταλλεύεται την πρόσθετη πληροφορία, προς όφελος της απόδοσης του αλγορίθμου. Η λειτουργία αυτού του υποσυστήματος, περιλαμβάνει την ανάκτηση των σχολιασμένων εκδόσεων των κειμένων των άρθρων από την τοπική βάση και τον έλεγχο ώστε να μην γίνει ξανά επεξεργασία των ήδη επεξεργασμένων. Στην συνέχεια με τη βοήθεια των σχολιασμών, αναγνωρίζονται οι προτάσεις κάθε κειμένου.

Κάθε πρόταση αναλύεται ώστε να βρεθούν οι βιοϊατρικοί όροι που αυτή περιέχει, αλλά και οι βασικοί όροι που περιέχονται σε αυτήν και παρόλο που δεν αντιστοιχούν σε βιοϊατρικές έννοιες, είναι χρήσιμοι για την αναζήτηση άρθρων.

Για κάθε βιοϊατρικό όρο, ελέγχεται με ποιους άλλους όρους συνυπάρχει (στην ίδια πρόταση).

Αφού η ανάλυση ολοκληρωθεί, εισάγεται στην οντολογία η αντίστοιχη πληροφορία που προκύπτει. Έτσι εισάγονται στις κατάλληλες κλάσεις οι βιοϊατρικοί όροι και μέσω των ιδιοτήτων μεταξύ των κλάσεων, οι πληροφορίες σχετικά με τις σχέσεις συνύπαρξής τους, σε επίπεδο πρότασης ή άρθρου. Επίσης εισάγεται η πληροφορία σχετικά με το ποιες προτάσεις περιέχει ένα κείμενο άρθρου και ποιοι βασικοί όροι εμφανίζονται σε αυτές και στο ίδιο το κείμενο.

Έτσι μετά την εκτέλεση του υποσυστήματος, για κάθε άρθρο η οντολογία περιέχει ένα αντικείμενο της κλάσης Article που αντιπροσωπεύει το άρθρο, αντικείμενα της κλάσης Lemma που αντιπροσωπεύουν τους βασικούς όρους του άρθρου, αντικείμενα της κλάσης Sentence που αντιπροσωπεύουν μοναδικά κάθε πρόταση του άρθρου, αντικείμενα στις κατάλληλες κάθε φορά κλάσεις, που αντιπροσωπεύουν τους βιοϊατρικούς όρους του άρθρου. Επίσης περιέχεται, με την βοήθεια των ιδιοτήτων articleHasWord και wordIsFromArticle, πληροφορία σχετικά με το άρθρο στο οποίο ανήκουν όλοι οι βιοϊατρικοί και βασικοί όροι και τέλος, με την βοήθεια των ιδιοτήτων sentenceHasWord και inSentence, πληροφορία σχετικά με το ποιους όρους συνυπάρχουν, σε πρώτο επίπεδο στην ίδια πρόταση ενός άρθρου και συνεπώς σε δεύτερο επίπεδο στο ίδιο άρθρο.

Όλη η πληροφορία που περιέχει η οντολογία μετά την εκτέλεση του OntologyProcess υποσυστήματος, χρησιμοποιείται από αλγορίθμους στα επόμενα υποσυστήματα, ώστε να γίνει η συσταδοποίηση των άρθρων αλλά και η εύρεση συγκεκριμένης πληροφορίας, ώστε να είναι δυνατή η απάντηση διαφόρων τύπων ερωτημάτων από τους χρήστες του συστήματος (Μηνά Ι. 2017).

Αξίζει να αναφερθεί, ότι κατά την διάρκεια υλοποίησης του συγκεκριμένου υποσυστήματος, υπήρξε η δοκιμή διαφόρων εργαλείων εξαγωγής σχέσεων μεταξύ όρων. Κάποια από αυτά τα εργαλεία είναι τα εξής:

- Open Language Learning for Information Extracting (Ollie)[59]
- Open Information Extraction (Open IE) [60]
- StanfordNLP[61]
- Multipax[62]

Τα εργαλεία αυτά είναι εργαλεία εξαγωγής σχέσεων. Δηλαδή, δέχονται προτάσεις και εξάγουν πληροφορία σχετικά με τις σχέσεις που βρέθηκαν μεταξύ των όρων στις προτάσεις αυτές. Η εφαρμογή τους όμως από το σύστημα οδήγησε σε σημαντικές καθυστερήσεις εκτέλεσης και τα επιστρεφόμενα αποτελέσματα δεν είχαν την αναμενόμενη χρησιμότητα στην παρούσα εργασία, ώστε να δικαιολογούν την χρήση τους. Για παράδειγμα το εργαλείο Multipax , όταν εφαρμοζόταν στις προτάσεις με πολλούς βιοϊατρικούς όρους, αποτύγγανε να αναγνωρίσει τα ουσιαστικά που τους αποτελούσαν και τα ρήματα τα οποία τους συνέδεαν. Τα προβλήματα αυτά δεν αντικατοπτρίζουν την αξία αυτών των εργαλείων, αλλά την αδυναμία τους να εφαρμοστούν με αποδεκτά αποτελέσματα τόσο από άποψη χρόνου όσο και από άποψη ανοχής λαθών, σε ένα τόσο συγκεκριμένο πεδίο όσο αυτό της Βιοϊατρικής.

3.6 ClusteringProcess

Η Εξόρυξη Δεδομένων (Data Mining) αφορά στην ανακάλυψη γνώσης από βάσεις δεδομένων και στις τεχνικές για την εξαγωγή και ανάλυση της, από σύνολα δεδομένων.

Υπάρχουν πολλές μέθοδοι που χρησιμοποιούνται, όπως η Κατηγοριοποίηση (Classification), Συσταδοποίηση (Clustering), Κανόνες Συσχέτισης (Association Rules) , Παλινδρόμηση (Regression) και Περιληπτική Παρουσίαση Πληροφορίας.

Μία από αυτές τις μεθόδους είναι η συσταδοποίηση. Αποτελεί μία μέθοδο με την οποία ένα σύνολο δεδομένων αντιστοιχίζεται σε ένα σύνολο συστάδων

(clusters). Οι συστάδες, μπορεί να έχουν προκαθορισμένο αριθμό ή να δημιουργούνται δυναμικά, με τον αριθμό τους να υπολογίζεται με αυτόματο τρόπο. Τα δεδομένα στη συνέχεια, ανατίθενται στις συστάδες βάση της ομοιότητας που έχουν μεταξύ τους.

Ο πιο διαδεδομένος αλγόριθμος συσταδοποίησης, είναι ο k-means [42]. Ο k-means αλγόριθμος λύνει το πρόβλημα της συσταδοποίησης, ομαδοποιώντας τα σχετικά κείμενα βάση της ομοιότητάς τους που υπολογίζεται από τις λέξεις που περιέχονται σε αυτά. Η διαδικασία που ακολουθεί είναι ένας απλός τρόπος ταξινόμησης των κειμένων σε έναν αριθμό συστάδων. Καθορίζονται k κέντρα, ένα για κάθε συστάδα, όπου το κάθε κέντρο είναι κάποιο κείμενο από αυτά για τα οποία γίνεται η ομαδοποίηση. Στην συνέχεια τα υπόλοιπα κείμενα τοποθετούνται στην συστάδα, με της οποίας το κέντρο είναι πλησιέστερα, βάση μίας μετρικής (Μηνά Ι. 2017).

Στο υποσύστημα ClusteringProcess, υλοποιείται η μέθοδος της συσταδοποίησης με την δημιουργία και εφαρμογή ενός αλγορίθμου, που βασίζεται στον kmeans αλγόριθμο. Ο αλγόριθμος αυτός που δημιουργήθηκε είναι ειδικά σχεδιασμένος, ώστε να ακολουθεί την λογική του k-means αλγορίθμου, αλλά και να εκμεταλλεύεται την πληροφορία που υπάρχει στην οντολογία του συστήματος, ώστε να επιτυγχάνει καλύτερα αποτελέσματα.

Σκοπός του αλγορίθμου είναι να εκμεταλλευτεί τις κλάσεις στις οποίες αποθηκεύονται οι βιοϊατρικοί όροι, έτσι ώστε η συσταδοποίηση που εκτελείται να συμπεριλαμβάνει και να βασίζεται στην πρόσθετη πληροφορία, με όφελος την καλύτερη ομαδοποίηση των άρθρων. Αυτό συμβαίνει καθώς η μετρική ομοιότητας δεν βασίζεται σε όλες τις λέξεις των κειμένων, όπως γίνεται μέχρι τώρα, αλλά βασίζεται ειδικότερα μόνο στους βιοϊατρικούς όρους των κειμένων αυτών. Έτσι, τα άρθρα που υπολογίζεται ότι ανήκουν στην ίδια συστάδα και άρα αναμένεται να έχουν σχετικό περιεχόμενο, στην

πραγματικότητα έχουν σχετικό περιεχόμενο με πιο αυστηρά κριτήρια, καθώς θα δίνεται έμφαση στις βιοϊατρικές έννοιες τις οποίες περιέχουν και περιγράφουν.

Η λειτουργία του αλγορίθμου αυτού είναι αρχικά ο υπολογισμός των ειδών των συστάδων που θα χρησιμοποιηθούν, αλλά και ο αριθμός τους. Έπειτα, για κάθε άρθρο, υπολογίζεται σε ποιες συστάδες ανήκει, ακολουθώντας την λογική του kmeans αλγορίθμου, εστιάζοντας όμως στους βιοϊατρικούς όρους αντί όλων των όρων των κειμένων. Αυτό σημαίνει ότι η ομοιότητα των άρθρων υπολογίζεται από την ευκλείδεια απόσταση των κειμένων τους, για τον υπολογισμό της οποίας χρησιμοποιούνται μόνο οι βιοϊατρικοί όροι που έχουν αναγνωρισθεί από τα προηγούμενα υποσυστήματα. Η αντίστοιχη πληροφορία σχετικά με τις συστάδες και τα κείμενα από τα οποία αποτελούνται, αποθηκεύεται στη συνέχεια στην οντολογία.

Χάρη στην εκτέλεση των προηγούμενων υποσυστημάτων, η οντολογία που δημιουργήθηκε έχει τέτοια δομή και περιλαμβάνει τέτοια πληροφορία, ώστε η διαδικασία συσταδοποίησης να γίνει με τέτοιο τρόπο ώστε να αποδίδει περισσότερη πληροφορία για κάθε άρθρο. Αυτό σημαίνει ότι χάρη στην δομή της οντολογίας, είναι δυνατόν να δημιουργηθούν περισσότερα του ενός, είδη συστάδων, ανάλογα με τις τέσσερις βασικές κλάσεις στις οποίες ανήκει κάθε άρθρο. Παρόλο που μετά την εφαρμογή εργαλείων εξαγωγής συμπερασμάτων, τα άρθρα που έχουν υποστεί επεξεργασία καταλήγουν να ανήκουν σε περισσότερες από μία κλάσεις, υπάρχουν τέσσερις βασικές κλάσεις που χρησιμοποιούνται κυρίως από το σύστημα. Αυτές είναι οι Cell, Protein, DNA και RNA κλάσεις της οντολογίας BioSiteOntology.owl

Οι βιοϊατρικοί όροι που εισάγονται στην οντολογία, είναι βέβαιο ότι θα ανήκουν σε μία τουλάχιστον από αυτές τις κλάσεις, λόγω του τρόπου λειτουργίας του εργαλείου σχολιασμών Genia tagger. Χρησιμοποιώντας αυτήν την πληροφορία, δημιουργήθηκαν τέσσερα διαφορετικά είδη συστάδων, ένα για κάθε μία από τις

βασικές κλάσεις. Έτσι κάθε άρθρο ανατίθεται σε τέσσερις διαφορετικές συστάδες, εάν περιέχει βιοϊατρικούς όρους των κλάσεων στις οποίες αντιστοιχούν οι συστάδες.

Για κάθε είδος, υπάρχει ένα σύνολο συστάδων, του οποίου ο αριθμός υπολογίζεται βάση του πλήθους των άρθρων που ανήκουν στην κατηγορία που αντιπροσωπεύει το συγκεκριμένο είδος. Αυτό σημαίνει αν εισαχθούν νέα άρθρα στην βάση, ο αριθμός αυτός πρέπει να επαναυπολογιστεί.

Κάθε άρθρο, μπορεί συνεπώς να ανήκει ταυτόχρονα σε τέσσερις διαφορετικές συστάδες. Για κάθε βασική κλάση, υπολογίζεται ένα πλήθος σχετικών άρθρων που τοποθετούνται στην ίδια συστάδα. Άρα, ανάλογα με την περίπτωση που μας ενδιαφέρει, το πλήθος των σχετικών άρθρων με ένα συγκεκριμένο άρθρο είναι διαφορετικό ανάλογα με την κλάση για την οποία ενδιαφερόμαστε. Αυτό προκύπτει από το γεγονός, ότι ένα άρθρο μπορεί να περιλαμβάνει βιοϊατρικούς όρους που ανήκουν σε περισσότερες της μίας βασικής κλάσης. Άρα για παράδειγμα, αν περιέχει βιοϊατρικούς όρους, τέτοιους ώστε να αποφασιστεί ότι ανήκει στην κλάση Protein και DNA, τότε μέσω του αλγορίθμου συσταδοποίησης, θα ανατεθεί σε δύο clusters. Το ένα θα περιλαμβάνει τα σχετικά άρθρα που ανήκουν και εκείνα στην κλάση Protein και το άλλο θα περιλαμβάνει ένα πλήθος άρθρων που ανήκουν στην κλάση DNA και είναι σχετικά με το άρθρο σε αυτόν τον τομέα. Τα δύο αυτά σύνολα, μπορεί να περιέχουν εντελώς διαφορετικά άρθρα, αλλά μπορεί να περιέχουν και κοινά στοιχεία.

Μετά την εκτέλεση αυτού του υποσυστήματος, η οντολογία έχει εμπλουτιστεί με νέα πληροφορία που ομαδοποιεί τα άρθρα βάση του περιεχομένου τους ως προς τους βιοϊατρικούς όρους. Άρα η ομοιότητα των κειμένων δεν αφορά το περιεχόμενό τους γενικά, αλλά εστιάζει στο περιεχόμενο σε βιοϊατρικό επίπεδο. Τα διανύσματα που αντιπροσωπεύουν τα άρθρα και πάνω στα οποία υπολογίζεται η απόσταση των άρθρων, είναι κατασκευασμένα ώστε να περιέχουν πληροφορία

μόνο για τους βιοϊατρικούς όρους που υπάρχουν στην οντολογία και όχι για όλους τους βασικούς όρους των κειμένων.

BioSite Interface

Το BioSite Interface είναι το τελευταίο υποσύστημα του συστήματος BioSite. Σε αυτό δημιουργείται η ιστοσελίδα του συστήματος που αποτελεί την διεπαφή χρήστη, έτσι ώστε να μπορούν να υποβληθούν ερωτήματα στην βάση – οντολογία του συστήματος, να υπολογιστούν τα αποτελέσματα και να παρουσιαστούν στον χρήστη.

Σκοπός της ιστοσελίδας είναι να προσφέρει στον χρήστη ένα εύκολο τρόπο υποβολής ερωτημάτων, που θα τον διευκολύνει στην εύρεση της επιθυμητής πληροφορίας στο πεδίο της βιοϊατρικής. Για τον λόγο αυτό, στην ιστοσελίδα πρέπει να παρέχεται ένας εύχρηστος τρόπος υποβολής ερωτημάτων και αναζήτησης άρθρων, που θα διευκολύνει τον χρήστη και θα τον απαλλάσσει από την απαίτηση παροχής ειδικής μορφής ερωτημάτων αλλά θα του παρέχει όλη την διαθέσιμη πληροφορία με εύκολο και γρήγορο τρόπο. Τα αποτελέσματα θα πρέπει να παρουσιάζονται με τρόπο τέτοιο, ώστε ο εντοπισμός της αναζητούμενης πληροφορίας να γίνεται εύκολα και γρήγορα, χωρίς την απαίτηση περαιτέρω προσπάθειας και έρευνας από την μεριά του χρήστη.

Για τον λόγο αυτό η ιστοσελίδα του συστήματος υλοποιεί πολλές λειτουργίες, τόσο για την εύκολη χρήση του συστήματος από τους χρήστες όσο και για τον υπολογισμό των απαντήσεων στα ερωτήματα που αυτοί θα υποβάλλουν, την παρουσίαση των κειμένων αλλά και της πληροφορίας που υπολογίστηκε από τις εκτελέσεις των προηγούμενων υποσυστημάτων (Μηνά Ι. 2017).

Οι λειτουργίες που παρέχονται στον χρήστη είναι αρχικά η εγγραφή στην ιστοσελίδα, η είσοδος εφόσον είναι ήδη εγγεγραμμένος, αλλά και η έξοδος από αυτήν.

Στην συνέχεια, παρέχεται η δυνατότητα αναζήτησης άρθρων βάση βιοϊατρικών όρων. Ένας χρήστης μπορεί να δώσει όρους αναζήτησης (query) στο αντίστοιχο πεδίο αναζήτησης. Τα αποτελέσματα που θα επιστραφούν θα είναι άρθρα που περιέχουν αυτούς τους όρους ως βιοϊατρικούς και όχι άρθρα που απλά τυχαίνει να περιέχουν τους όρους χωρίς βιοϊατρική σημασία.

Παρέχεται επίσης η δυνατότητα γενικής αναζήτησης βάση οποιουδήποτε όρου. Ένας χρήστης μπορεί να εκτελέσει αναζήτηση στα άρθρα δίνοντας όρους αναζήτησης, χωρίς αυτοί να αποτελούν βιοϊατρικούς όρους. Σε αυτήν την περίπτωση, επιστρέφονται άρθρα που τους περιέχουν, είτε αποτελούν όρους βιοϊατρικής σημασίας είτε όχι.

Μία ακόμα λειτουργία της ιστοσελίδας είναι η δυνατότητα αναζήτησης άρθρων βάση τίτλου, συγγραφέων και ημερομηνίας άρθρου. Ένας χρήστης μπορεί να πραγματοποιήσει αναζήτηση άρθρων παρέχοντας όρους αναζήτησης σε συγκεκριμένο πεδίο του άρθρου. Έτσι αν ενδιαφέρεται για παράδειγμα για τα άρθρα που έχει γράψει ένας συγκεκριμένος συγγραφέας, μπορεί να πραγματοποιήσει αναζήτηση δίνοντας ως όρο αναζήτησης το όνομα του συγγραφέα.

Σε κάθε αναζήτηση που πραγματοποιεί ο χρήστης, επιστρέφονται ως αποτελέσματα τα άρθρα που είναι σχετικά με το ερώτημα που έθεσε. Αυτά εμφανίζονται σε έναν πίνακα μαζί με τις βασικές πληροφορίες του κάθε ενός. Σε αυτό το σημείο παρέχεται στον χρήστη η δυνατότητα να εκτελέσει νέα αναζήτηση βάση της τιμής που έχει κάποιο συγκεκριμένο πεδίο, επιλέγοντας το. Αυτό θα έχει ως αποτέλεσμα να επιστραφεί ένα νέο σύνολο άρθρων που έχουν ως κοινό γνώρισμα, την ίδια τιμή, με αυτήν που είχε το αρχικό άρθρο, στο πεδίο που επέλεξε ο χρήστης.

Σε περίπτωση που ο χρήστης υποβάλλει ερώτημα με όρους, γενικούς ή βιοϊατρικούς, αυτοί υποβάλλονται σε επεξεργασία ώστε να αφαιρεθούν οι κοινές λέξεις που μπορεί να περιέχονται όπως σύνδεσμοι, άρθρα κ.ά. Επίσης υλοποιείται επεξεργασία τους, όπου υπολογίζεται η βάση της λέξης από την οποία προέρχεται

κάθε όρος, ώστε να μην έχει σημασία ο αριθμός των όρων (ενικός, πληθυντικός) ή η συγκεκριμένη μορφή στην οποία δόθηκαν, αλλά η έννοια της λέξης.

Μία επιπλέον λειτουργία που παρέχει η ιστοσελίδα είναι η εμφάνιση των λεπτομερειών των άρθρων. Εάν ο χρήστης επιλέξει ένα άρθρο από το σύνολο των επιστρεφόμενων ως αποτέλεσμα της αναζήτησής του, τότε εμφανίζονται οι αναλυτικές πληροφορίες του άρθρου και το πλήρες κείμενο του.

Για κάθε άρθρο παρέχεται στον χρήστη η δυνατότητα εμφάνισης του στην ηλεκτρονική βιβλιοθήκη PubMed Central, με την χρήση του αντίστοιχου υπερσυνδέσμου.

Μία από τις σημαντικές λειτουργίες της ιστοσελίδας είναι ότι εάν ο χρήστης οδηγήθηκε στις λεπτομέρειες ενός άρθρου μετά από αναζήτηση γενικών ή βιοϊατρικών όρων, τότε εκτός από την εμφάνιση του πλήρους κειμένου του άρθρου, εμφανίζονται τα μέρη του κειμένου στα οποία βρέθηκαν οι όροι αναζήτησης, με αυτούς να παρουσιάζονται με ειδική επισήμανση για εύκολο εντοπισμό τους. Τα μέρη του κειμένου που εμφανίζονται, είναι ταξινομημένα έτσι ώστε να εμφανίζονται πρώτα αυτά που περιέχουν, όσο το δυνατόν περισσότερους από τους όρους του ερωτήματος (Μηνά Ι. 2017).

Επιπλέον, για κάθε άρθρο παρέχεται στον χρήστη η δυνατότητα εμφάνισης ενός συνόλου σχετικών άρθρων, τα οποία έχουν υπολογιστεί μέσω συσταδοποίησης. Η λειτουργία παροχής σχετικών κειμένων, είναι γρήγορη και αποτελεσματική, χωρίς να υπάρχει καθυστέρηση στην εμφάνιση των αποτελεσμάτων, εφόσον οι συστάδες έχουν προϋπολογιστεί. Έτσι εμφανίζονται άμεσα στον χρήστη, χωρίς να απαιτείται χρόνος υπολογισμού των σχετικών κειμένων εκείνη τη στιγμή.

Μία επιπλέον λειτουργία του συστήματος είναι η ταξινόμηση των άρθρων σχετικά με την σχετικότητά τους με το ερώτημα του χρήστη και η σελιδοποίηση των αποτελεσμάτων ανά δέκα, με δυνατότητα περιήγησης στα υπόλοιπα με εύκολο

τρόπο. Για κάθε λίστα αποτελεσμάτων που επιστρέφεται μετά από κάποια αναζήτηση, τα αποτελέσματα ταξινομούνται, ώστε πρώτα να εμφανίζονται τα άρθρα των οποίων τα κείμενα πληρούν τις προϋποθέσεις πιο ικανοποιητικά. Για παράδειγμα εάν η λίστα αφορά αποτελέσματα που προέκυψαν μετά την αναζήτηση βάση όρων, πρώτα στη λίστα θα τοποθετηθούν τα κείμενα που περιέχουν τους περισσότερους από τους όρους αναζήτησης και με την μεγαλύτερη συχνότητα για κάθε έναν από αυτούς.

ΚΕΦΑΛΑΙΟ 4

Παρουσίαση του συστήματος BioSite

Σε αυτό το κεφάλαιο θα παρουσιαστεί η χρήση του ολοκληρωμένου συστήματος BioSite και θα δοθούν εικόνες και αποτελέσματα της λειτουργίας του. Αφού εκτελεστούν οι λειτουργίες των πέντε πρώτων υποσυστημάτων, η οντολογία του συστήματος έχει ενημερωθεί και εμπλουτιστεί με την νέα πληροφορία που αφορά τα άρθρα, τις πληροφορίες σχετικά με αυτά, τις σχέσεις των διάφορων στοιχείων αλλά και την πληροφορία που αφορά τον υπολογισμό των σχετικών όρων μετά την συσταδοποίηση (Μηνά Ι. 2017).

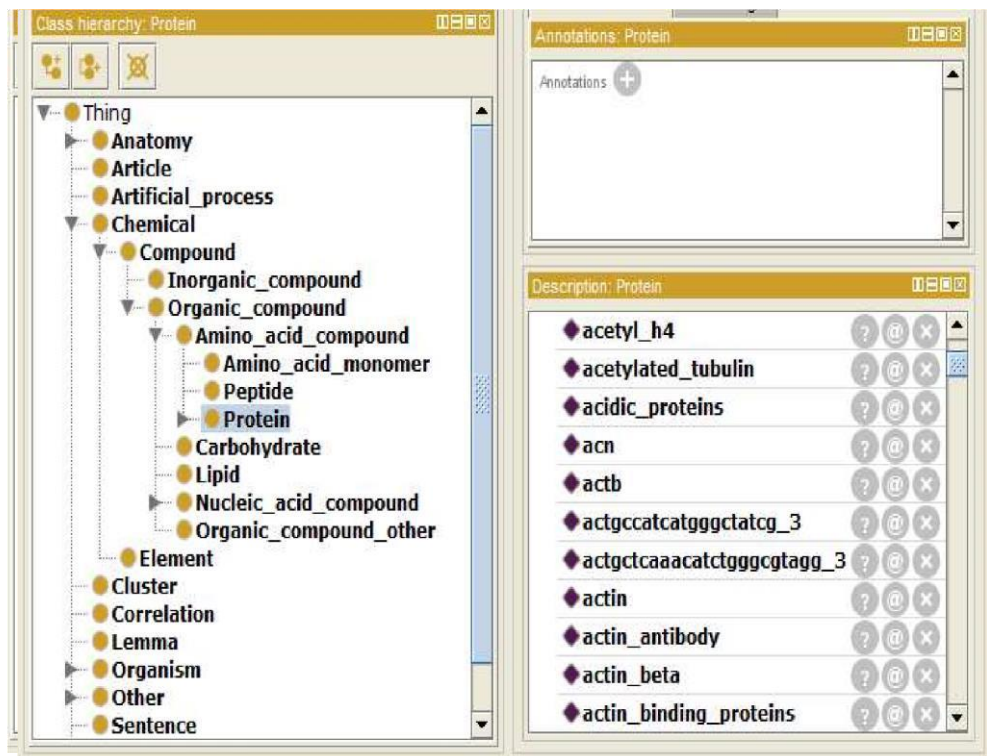
Αρχικά λοιπόν δίνεται μία σειρά εικόνων όπου παρουσιάζονται τα περιεχόμενα της οντολογίας του συστήματος και στην συνέχεια παρουσιάζεται η ιστοσελίδα του συστήματος και γίνεται μία επίδειξη χρήσης των παρεχόμενων λειτουργιών της.

Στις Εικόνες 6 και 7, παρουσιάζονται μέρη από τα περιεχόμενα μερικών αντιπροσωπευτικών κλάσεων της οντολογίας, δηλαδή τα αντικείμενα αυτών των κλάσεων.

Στην Εικόνα 8 παρουσιάζονται τα αντικείμενα μίας αντιπροσωπευτικής κλάσης, καθώς και οι ιδιότητες που τα συνδέουν με αντικείμενα άλλων κλάσεων. Στην Εικόνα 9 παρουσιάζεται ένα μέρος των αντικειμένων της κλάσης Cluster, η οποία μετά την εκτέλεση του αλγορίθμου συσταδοποίησης, περιέχει τις συστάδες που δημιουργήθηκαν για κάθε μία από τις βασικές κλάσεις.

Στην Εικόνα 10 παρουσιάζονται τα περιεχόμενα μίας συστάδας και ένα μέρος των άρθρων που περιέχει ως αντικείμενα.

Στην Εικόνα 11 παρουσιάζονται τα ενημερωμένα περιεχόμενα ενός άρθρου, που μετά την εκτέλεση του αλγορίθμου συσταδοποίησης περιέχουν και πληροφορία σχετικά με τις συστάδες στις οποίες ανήκει το άρθρο.



Εικόνα 6: Περιεχόμενα της Κλάσης Protein (Μηνά Ι. 2017).

The screenshot displays a software interface with a class hierarchy on the left and a list of data property assertions on the right. The class hierarchy includes categories like Thing, Anatomy, Article, Artificial_process, Chemical, Cluster, Correlation, Lemma, Organism, Other, Sentence, User, and User. The data property assertions list includes:

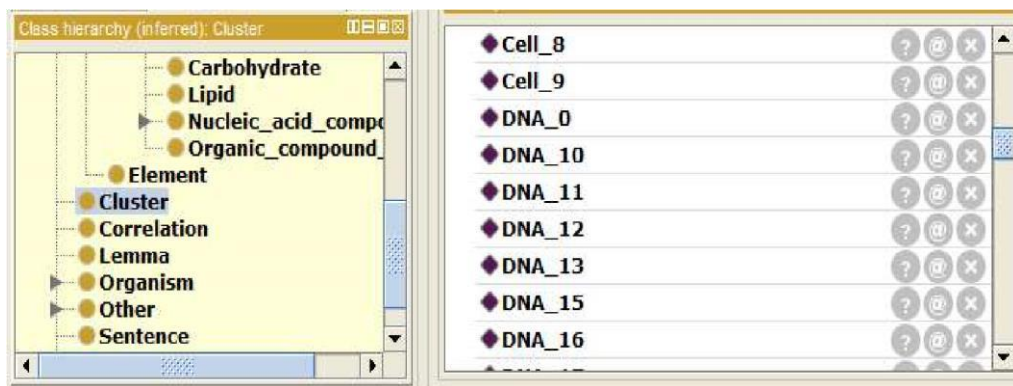
- articleHasWord ceus
- articleHasWord mri
- articleHasWord platelet_count_40_109
- articleHasWord ct
- articleHasWord fetoprotein
- articleHasWord hyperechoic_region
- articleHasWord spss
- hasAuthor "Ping Liang"^^string
- hasPubYear "2017"^^string
- hasTitle "A clinical study of thermal monitoring techniques of ultrasound-guided microwave ablation for hepatocellular carcinoma in high-risk locations"^^string
- hasAuthor "Zhi-gang Cheng"^^string
- hasAuthor "Xiao-ling Yu"^^string
- hasAuthor "Fang-yi Liu"^^string
- hasAuthor "Zhi-yu Han"^^string
- hasAuthor "Jie Yu"^^string

Εικόνα 7: Ιδιότητες ενός άρθρου

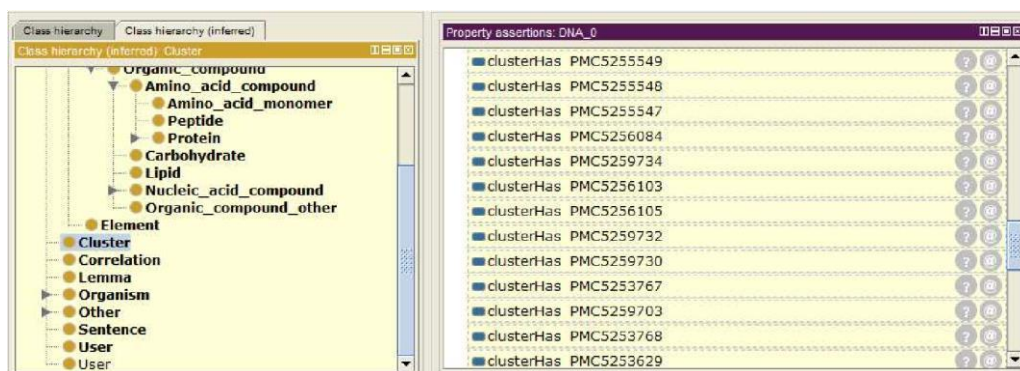
The screenshot displays a software interface with a class hierarchy on the left and a list of data property assertions on the right. The class hierarchy includes categories like Thing, Anatomy, Article, Artificial_process, Chemical, Compound, Inorganic_compound, Organic_compound, Amino_acid_compound, Amino_acid_monomer, Peptide, Protein, Carbohydrate, Lipid, Nucleic_acid_compound, Organic_compound_other, Element, Cluster, Correlation, Lemma, Organism, Other, Sentence, User, and User. The data property assertions list includes:

- inSentence PMCS253729_10292
- inSentence PMCS259722_11784
- wordIsFromArticle PMCS259722
- inSentence PMCS259722_11869
- inSentence PMCS259718_12434
- wordIsFromArticle PMCS253729
- inSentence PMCS259722_11774
- inSentence PMCS259722_11658
- wordIsFromArticle PMCS253721
- inSentence PMCS259714_13503
- inSentence PMCS259722_11678
- inSentence PMCS259722_11776
- inSentence PMCS259722_11676
- inSentence PMCS259722_11663
- inSentence PMCS259722_11778
- inSentence PMCS259722_11792
- wordIsFromArticle PMCS259718
- inSentence PMCS253721_7890
- inSentence PMCS259722_11659
- bioTerm "actin"^^string

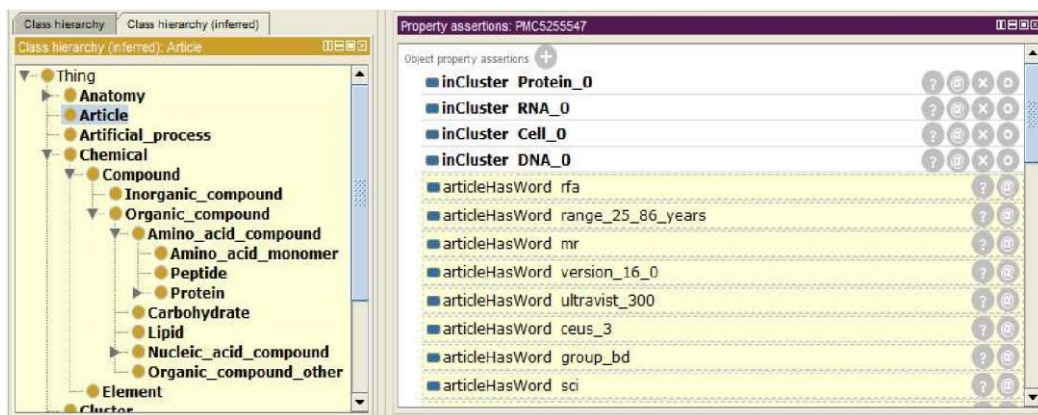
Εικόνα 8: Ιδιότητες ενός βιοϊατρικού όρου



Εικόνα 9: Περιεχόμενα της κλάσης Cluster



Εικόνα 10: Ιδιότητες μίας συστάδας



Εικόνα 11: Ιδιότητες ενός άρθρου μετά την εκτέλεση της συσταδοποίησης

Ακολουθεί μία σειρά εικόνων που δείχνουν τις βασικές λειτουργίες της ιστοσελίδας BioSite Interface και της λειτουργίας του συστήματος σε κάθε βήμα.

Στην Εικόνα 12 παρουσιάζεται η αρχική σελίδα του συστήματος. Από εδώ ο χρήστης μπορεί να εκτελέσει εγγραφή ή είσοδο στο σύστημα. Παρέχονται επίσης σύνδεσμοι προς την φόρμα αναζήτησης βάση συγκεκριμένου πεδίου, προς την σελίδα αναζήτησης βάση βιοϊατρικών όρων και το πεδίο αναζήτησης βάση γενικών όρων.



Εικόνα 12: Αρχική σελίδα του συστήματος BioSite

Εάν ο χρήστης επιλέξει να κάνει εγγραφή κατευθύνεται στην σελίδα όπου υπάρχει η αντίστοιχη φόρμα συμπλήρωσης των στοιχείων του, όπως φαίνεται στην Εικόνα 13.



Εικόνα 13: Σελίδα εγγραφής χρήστη

Εάν ο χρήστης επιλέξει τον σύνδεσμο για την αναζήτηση βάση πεδίων, εμφανίζονται τα πεδία που μπορεί να επιλέξει, όπως φαίνεται και στην Εικόνα 14.



Εικόνα 14: Εμφάνιση πεδίων αναζήτησης

Αφού επιλέξει το πεδίο που επιθυμεί όπως για παράδειγμα συγγραφέα, όπως μπορούμε να δούμε και από την Εικόνα 15, κατευθύνεται σε μία σελίδα όπου μπορεί να εισάγει το όνομα του συγγραφέα που επιθυμεί να αναζητήσει άρθρα γραμμένα από αυτόν.



Εικόνα 15: Σελίδα εισαγωγής τιμής σε πεδίο αναζήτησης

Στη συνέχεια εμφανίζονται τα αποτελέσματα, που περιέχουν άρθρα στα οποία βρέθηκε ως συγγραφέας, αυτός που εισήγαγε ο χρήστης. Όπως φαίνεται και από την Εικόνα 16 των αποτελεσμάτων, εμφανίζεται ο τίτλος των άρθρων, τα ονόματα των συγγραφέων που αποτελούν υπερσυνδέσμους και εάν επιλεγτούν, θα εκτελεστεί μία νέα αναζήτηση στο πεδίο αυτό με τιμή την νέα επιλογή του χρήστη. Εμφανίζεται ακόμη η ημερομηνία έκδοσης του άρθρου και ένας υπερσύνδεσμος προς την PubMed Central όπου μπορεί να βρει την εκεί έκδοση του άρθρου. Τέλος για κάθε άρθρο παρέχεται υπερσύνδεσμος προς την σελίδα λεπτομερειών του άρθρου, η οποία περιγράφεται στην συνέχεια.

Title	Publication Year	Article's Details Page	Authors	Pubmed link
A clinical study of thermal monitoring techniques of ultrasound-guided microwave ablation for hepatocellular carcinoma in high-risk locations	2017	Details	Ping Lian Zhi-xiang Cheng Xiao-ling Yu Yang-yi Liu Zhi-yu Han Jie Yu	Go


Εικόνα 16: Αναζήτηση βάση ονόματος συγγραφέα (Μηνά Ι. 2017).

Στην εικόνα 17, φαίνεται ο τρόπος εισαγωγής ερωτημάτων με γενικούς όρους. Στο συγκεκριμένο παράδειγμα εισάγεται το ερώτημα “cancer in women”



Εικόνα 17: Αναζήτηση βάση γενικών όρων : "cancer in women"

Στην Εικόνα 18, παρουσιάζεται η σελίδα με τα δέκα πρώτα αποτελέσματα που εμφανίζονται.



Title	Publication Year	Article's Details Page	Authors	PubMed link:
Prostate Cancer in a Patient with a Family History of BRCA Mutation: a Case Report and Literature Review	2017	Details	Seo Ho Kyung Kim Sung Han Chung Junsoo Song Won Hoon Joung Jae Young Park Weon Seo Lee Kang Hvun	Go
TLR2 2258 G>A single nucleotide polymorphism and the risk of congenital infection with human cytomegalovirus	2017	Details	Studzinska Miroslawa Wujcicka Wioletta Wilczyński Jan Paradowska Edyta Nowakowska Dorota	Go
National Trends in Surgery for Rotator Cuff Disease in Korea	2017	Details	Lee Kwang-Hyun Lee Bong-Gun Jo Young-Hoon Kim Sung-Jae Kim Jaedong	Go
Parathyroid Hormone-Like Hormone is a Poor Prognosis Marker of Head and Neck Cancer and Promotes Cell Growth via RUNX2 Regulation	2017	Details	Hsiao Jenn-Ren Shieh Yi-Shang Chen Chi-Long Shuoh Shue-Gwo Peng Hsuan-Yu Lin Yuan-Feng Su Chia-Yi Chang Wei-Min Chang Yu-Chan Hsiao Michael Chang Jang-Yang	Go
The role of total cell-free DNA in predicting outcomes among trauma patients in the intensive care unit: a systematic review	2017	Details	Gögenur Ismail Burcharth Jakob Gögenur Mikail	Go
GSAR: Bioconductor package for Gene Set analysis in R	2017	Details	Emmert-Streib Frank Zybailov Boris Glazko Galina Rahmatallah Yasir	Go
Systems biology analysis of drivers underlying hallmarks of cancer cell metabolism	2017	Details	Zielinski Daniel C. Thomas Alex Palsson Bernhard O. Corbett Austin J. Jamshidi Neema Bordbar Arash	Go
MicroRNA-199a-5p promotes tumour growth by dual-targeting PIAS3 and p27 in human osteosarcoma	2017	Details	Wang Chen Zhou Guangxin Sun Defang Guo Yu Chen Jianming Jiang Haoyang Li Wentao Ba Ximing Huang Zhen Zhang	Go
MicroRNA-199a-5p promotes tumour growth by dual-targeting PIAS3 and p27 in human osteosarcoma	2017	Details	Wang Chen Zhou Guangxin Sun Defang Guo Yu Chen Jianming Jiang Haoyang Li Wentao Ba Ximing Huang Zhen Zhang Junfeng Wu Sunja	Go
Primary neuroendocrine breast carcinomas are associated with poor local control despite favourable biological profile: a retrospective clinical study	2017	Details	Jukkola-Vuorinen Aija Mattson Johanna Karhitala Peeter Heikkilä Paivi Rominen Nelli Takala Sari Haapasari Kirsi-Maria	Go
Immunotherapy with Dendritic Cells Modified with Tumor-Associated Antigen Gene Demonstrates Enhanced Antitumor Effect Against Lung Cancer1	2017	Details	Jiang Tao Zhou Wei Chen Xiao Fan Guoxin Zhang Jun Zhou Caicun Zhao Peilin Ren Shengxiang	Go

Εικόνα 18: Αποτελέσματα ερωτήματος γενικών όρων

Η Εικόνα 19, αποτελεί την σελίδα λεπτομερειών του πρώτου άρθρου από τα επιστρεφόμενα αποτελέσματα. Εκτός από τις βασικές πληροφορίες των άρθρων όπως τίτλος, συγγραφείς, ημερομηνία έκδοσης και τον υπερσύνδεσμο προς την PubMed Central, εμφανίζονται δύο νέα στοιχεία.

Το πρώτο από αυτά είναι ένας σύνδεσμος προς την σελίδα όπου παρουσιάζονται τα σχετικά με το παρόν, άρθρα που έχουν υπολογιστεί βάση συσταδοποίησης.

Το δεύτερο είναι ότι επειδή η σελίδα λεπτομερειών προέκυψε μετά από ερώτημα του χρήστη βάση όρων, εμφανίζονται τα μέρη των κειμένων που τους περιέχουν με ειδική επισήμανση αυτών των όρων, για εύκολο εντοπισμό τους. Όπως φαίνεται, υπάρχει επισήμανση και των δύο βασικών όρων του ερωτήματος και όχι τον όρων χωρίς ιδιαίτερη σημασία όπως συνδέσμους κτλ.

Εάν ο χρήστης επιθυμεί να έχει πρόσβαση στο πλήρες κείμενο του άρθρου, μπορεί να επιλέξει τον διαθέσιμο υπερσύνδεσμο “Details” (Μηνά Ι. 2017).

Title	Publication Year	Article's Details Page	Similar Articles Page	Authors	PubMed link:
Prostate Cancer in a Patient with a Family History of BRCA Mutation: a Case Report and Literature Review	2017	Details	Similar Articles	Seo Ho Kyung Kim Sung Han Chung Jinsoo Song Won Hoon Joung Jae Young Park Weon Seo Lee Kang Hyun	Go

Article

Show

... cancer-prone families (), and suggest that the family members who do not carry the deleterious BRCA family-specific mutation might inherit other underlying genetic instabilities that increase their risk of cancer. A recent study by Dite et al. () indicated that first-degree relatives of women diagnosed with breast cancer at a ...

... agonist. After 2 months of therapy, his bone pain subsided, and his PSA level decreased to 740 ng/mL. His family pedigree revealed a strong history of breast cancer with his 2 nieces being diagnosed at the age of 36 and 39 years (). Furthermore, he had 2 nuts and 2 ...

Εικόνα 19: Το πρώτο άρθρο των αποτελεσμάτων του ερωτήματος γενικών όρων

Στην Εικόνα 20 παρουσιάζεται το δεύτερο αποτέλεσμα της αναζήτησης με τους όρους “cancer in women”. Επιλέγοντας εμφάνιση του άρθρου, εμφανίζεται το περιεχόμενο της Εικόνας 21. Εκεί εμφανίζονται τα μέρη του κειμένου με επισήμανση του όρου “cancer”. Παρόλο που δεν υπάρχει η λέξη “woman” ή “women”, στο κείμενο, αυτό κατατάχθηκε ως δεύτερο στην σειρά των αποτελεσμάτων, καθώς το περιεχόμενο του αφορά τον καρκίνο του στήθους, όπως μπορεί να διαπιστωθεί από το κείμενο γύρω από την λέξη “cancer”.

Title	Publication Year	Article's Details Page	Similar Articles Page	Authors	PubMed link
Primary neuroendocrine breast carcinomas are associated with poor local control despite favourable biological profile: a retrospective clinical study	2017	Details	Similar Articles	Jukkola-Muurinen Ari Mattson Johanna Karhinen Petter Heikkilä Päivi Rouvinen Nejli Takala Sari Hannu Saar Kirsi Maria	Go

Εικόνα 20: Το δεύτερο άρθρο των αποτελεσμάτων του ερωτήματος γενικών όρων

... sites arose during the follow-up. Clinical and pathological data of the 43 patients with NEBC are described in Table . Five patients had a history of earlier breast cancer. From all the NEBC patients, 19 (44.2%) were operated with mastectomy and axillary evacuation, 11 (25.6%) with partial breast resection and sentinel ...
... Breast carcinomas with neuroendocrine features (NEBC) are usually estimated to represent <1% of all breast cancers [.]. Their actual incidence is in fact difficult to assess since neuroendocrine markers (NE) are not routinely used in breast cancer diagnostics and NEBC are challenging to recognize clinically or with basic immunohistochemistry. ...
... than patients with ductal carcinomas. Plasma chromogranin A correlated only to higher age at diagnosis ($p = 0.0028$). Relapse-free survival ($p = 0.0013$), disease-free survival ($p = 0.024$) and overall survival ($p = 0.0028$) favoured ductal carcinomas compared to NEBC, while no difference was observed in distant disease-free survival or in breast cancer-specific survival. Conclusions There is remarkable ...
... analysis. Disease-free (DFS), relapse-free (RFS), distant disease-free (DDFS), breast cancer-specific (BCSS) and overall (OS) survival were calculated from the time of diagnosis to disease recurrence at any site (DFS), in the ipsilateral breast, scar, or axilla (RFS), at distant sites (DDFS) or to the time of confirmed breast cancer-related death ...
... The World Health Organization (WHO) definition for breast cancer with neuroendocrine features has been revised lately [.]. According to the definition of 2012, there is no specific threshold for the positivity of neuroendocrine marker expression, namely synaptophysin or chromogranin but either expression is required for the diagnosis. According to ...

Εικόνα 21: Μέρος κειμένου άρθρου με επισήμανση όρων

Στην Εικόνα 22 παρουσιάζεται η αναζήτηση βάση βιοϊατρικών όρων. Η λέξη που αποτελεί το ερώτημα είναι ο όρος “cancer”. Στα αποτελέσματα που εμφανίζονται στην Εικόνα 23, εμφανίζονται άρθρα τα οποία περιέχουν τον όρο αυτό με βιοϊατρική έννοια. Δηλαδή το εργαλείο σχολιασμού επέδωσε σε αυτόν τον

όρο σχολιασμό βιοϊατρικής έννοιας και αναγνώρισε από τα συμφραζόμενα ότι αποτελεί αναφορά στην ασθένεια και όχι σε κάποια άλλη έννοια με την ίδια ονομασία.



Εικόνα 22: Αναζήτηση βάση βιοϊατρικού όρου : "cancer"

Title	Publication Year	Article's Details Page	Authors	PubMed link:
Prostate Cancer in a Patient with a Family History of BRCA Mutation: a Case Report and Literature Review	2017	Details	Seo Ho Kyung Kim Song Han Chung Jinsoo Song Wei Hoon Joong Jae Young Park Weon Seo Lee Kang Hyun	Go
An evaluation of the sonoporation potential of low-boiling point phase-change ultrasound contrast agents in vitro	2017	Details	Arena Christopher B. Fox Sumantika M. Novell Anthony Dayton Paul A. Yun Yooheung	Go
Calcium-independent disruption of microtubule dynamics by nanosecond pulsed electric fields in U87 human glioblastoma cells	2017	Details	Carl Lynn Leyeque Philippe Arnaud-Cornos Delia Burke Ryan C. Bardeh Sylvia M. O'Connor Rodrigo P.	Go
SESN2 correlates with advantageous prognosis in hepatocellular carcinoma	2017	Details	Xu Xinyu Yan Weizang Liao Weiva Liu Yulong Dao Hongzhang Xu Li Yu Jing Chen Shaosen	Go
Single-cell epigenomic variability reveals functional cancer heterogeneity	2017	Details	Shun Yant Wu Benjamin Goodrich William J. Buenostro Jason D. Latzenberger Ulrike M. Chung Howard Y. Sheffield Nathan C. Karlsson Arvan	Go

Εικόνα 23: Αποτελέσματα αναζήτησης βάση βιοϊατρικού όρου

Στην Εικόνα 24, παρουσιάζεται ένα από τα άρθρα που επιστράφηκαν ως απάντηση στο ερώτημα με βιοϊατρικό όρο “cancer”. Το άρθρο αφορά τον καρκίνο του προστάτη. Επιλέγοντας την επιλογή εμφάνιση, παρουσιάζονται τα μέρη του

κειμένου με επισήμανση του βιοϊατρικού όρου αναζήτησης, όπως φαίνεται στην Εικόνα 25.

Title	Publication Year	Article's Details Page	Similar Articles Page	Authors	PubMed link:
Prostate Cancer in a Patient with a Family History of BRCA Mutation: a Case Report and Literature Review	2017	Details	Similar Articles	Seo Ho Kyung Kim Sunghwan Chong Jinsou Song Won Hoon Joungr Jae Young Park Weon Seo Lee Kang Hyun	Go

Article

Εικόνα 24: Σελίδα λεπτομερειών άρθρου

Article

... agonist. After 2 months of therapy, his bone pain subsided, and his PSA level decreased to 740 ng/mL. His family pedigree revealed a strong history of breast cancer with his 2 nieces being diagnosed at the age of 36 and 39 years (). Furthermore, he had 2 aunts and 2 ...

... uncles from both sides who had been diagnosed with stomach cancer, ovarian cancer, or colorectal cancer. His genomic sequencing analysis of the BRCA genes revealed the same BRCA2 NM_000059.3:c.3744_3747delTGAG (p.Ser1248Argfs*10) mutation that his breast cancer-affected nieces carried (). Further PC surveillance, including digital rectal examination and PSA screening, was planned ...

... One of the most significant risk factors for prostate cancer (PC) is a family history of the disease. To date, germ-line mutations in the breast cancer predisposition gene (BRCA) 2 are the genetic events that have been shown to confer the highest risk of PC at 8.6-fold in men ≤ ...

... bone metastases. The patient had a strong family history of breast cancer with his 2 nieces affected. Furthermore, his aunts and uncles from both sides were diagnosed with stomach, ovarian, and colorectal cancers. His genomic sequencing analysis of the BRCA genes revealed the same BRCA2 deleterious mutation that his breast ...

... stratification, particularly to identify those with potentially lethal forms of the disease. We herein present a case of BRCA2-mutated PC with a family history of breast cancer. In addition, we reviewed the literature on germ-line mutations of the BRCA genes and their predictive values of patient outcomes. A 56-year-old patient ...

... for the patient's male relatives. Informed written consent was submitted by the patient and the review was exempted by the Institutional Review Board of the National Cancer Center.(August 25, 2015) The role of BRCA germ-line mutations in male cancer has been extensively studied although they are present in only 1%–2% ...

Εικόνα 25: Μέρος κειμένου άρθρου με επισήμανση όρων

Εάν ο χρήστης επιθυμεί να έχει πρόσβαση στο πλήρες κείμενο του άρθρου και όχι μόνο στα συγκεκριμένα σημεία με τον βιοϊατρικό όρο, μπορεί να επιλέξει τον υπερσύνδεσμο λεπτομερειών και την επιλογή εμφάνισης του. Στην Εικόνα 26 παρουσιάζεται ένα μέρος του πλήρους κειμένου του συγκεκριμένου άρθρου.

Title	Publication Year	Article's Details Page	Similar Articles Page	Authors	PubMed link
Prostate Cancer in a Patient with a Family History of BRCA Mutation: a Case Report and Literature Review	2017	Details	Similar Articles	Seo Ho Kyung Kim Song Han Chong Jinsoo Song Won Hoon Joung Jae Young Park Woon Seo Lee Kang Hyun	Go

Article

Show ▾

One of the most significant risk factors for prostate cancer (PC) is a family history of the disease, with germ-line mutations in the breast cancer predisposition gene (BRCA) 2 conferring the highest risk. We here report a 56-year-old man presented with painful gait disturbance and diagnosed PC with multiple disseminated bone metastases. The patient had a strong family history of breast cancer with his 2 nieces affected. Furthermore, his aunts and uncles from both sides were diagnosed with stomach, ovarian, and colorectal cancers. His genome sequencing analysis of the BRCA genes revealed the same BRCA2 deleterious mutation that his breast cancer-affected nieces carried. Previous studies have suggested that BRCA2-mutated PC is associated with a more aggressive phenotype and poor prognosis. Our experience in the present case also indicated the urgent needs for novel treatment modality and PC screening in this high-risk group of patients. One of the most significant risk factors for prostate cancer (PC) is a family history of the disease. To date, germ-line mutations in the breast cancer predisposition gene (BRCA) 2 are the genetic events that have been shown to confer the highest risk of PC at 8.6-fold in men \leq 65 years of age (1). Although it remains unclear how BRCA2 affect prostate tumorigenesis, deleterious mutations in BRCA2 genes have

Εικόνα 26: Εμφάνιση του πλήρους κειμένου ενός άρθρου

Στην Εικόνα 27 παρουσιάζονται τα σχετικά άρθρα με αυτό το οποίο προέκυψε μετά το ερώτημα βάση του βιοϊατρικού όρου “cancer”. Τα άρθρα αυτά είναι ταξινομημένα βάση του βαθμού ομοιότητας τους με αυτό που υπολογίστηκε από τον αλγόριθμο συσταδοποίησης και αν επιλέξουμε κάποιο από αυτά, εμφανίζεται η σελίδα λεπτομερειών τους. Στην Εικόνα 28 παρουσιάζεται ένα από αυτά, που

αποδεικνύει την σωστή λειτουργία του αλγορίθμου καθώς αναφέρεται και αυτό στην ασθένεια του καρκίνου (Μηνά Ι. 2017).



Εικόνα 27: Εμφάνιση σχετικών άρθρων βάση συσταδοποίησης

Title	Publication Year	Article's Details Page	Similar Articles Page	Authors	PubMed link:
Systems biology analysis of drivers underlying hallmarks of cancer cell metabolism	2017	Details	Similar Articles	Zielinski Daniel C. Palaion Bernhard O. Thomas Alex Corbett Austin J. Janshadi Neena Bordbar Arash	Go

Article
hide >

Εικόνα 28: Σελίδα λεπτομερειών σχετικού άρθρου

ΕΙΔΙΚΟ ΜΕΡΟΣ

ΚΕΦΑΛΑΙΟ 5

5.1 Σχεδιασμός της έρευνας

Η μέθοδός μας στηρίχθηκε στο περιγραφικό μοντέλο έρευνας με βάση το οποίο περιγράφονται μεταβλητές και συγκρίνονται ομάδες ατόμων για κάποια μεταβλητή (Λαγουμιζτής και συν 2016).

Για την συλλογή των στοιχείων της έρευνάς μας χρησιμοποιήθηκε ερωτηματολόγιο κατάλληλα σχεδιασμένο το οποίο και απευθυνόταν σε σπουδαστές, του τμήματος Φυσικοθεραπείας του ΤΕΙ Δυτικής Ελλάδος. Η επιλογή του δείγματος έγινε ανεξάρτητα από καταγωγή, οικογενειακή και κοινωνικοοικονομική κατάσταση.

Ως όργανο μέτρησης χρησιμοποιήθηκε γραπτό ερωτηματολόγιο, αποτελούμενο από 9 ερωτήσεις όλες κλειστού τύπου. Όλες ήταν εναλλακτικών απαντήσεων.

Τα στοιχεία συλλέχθηκαν από τον Νοέμβριο του 2017 έως τον Ιανουάριο του 2019 στο χώρο του ΤΕΙ στο Αίγιο.

Για να επιτευχθεί υψηλή εγκυρότητα περιεχομένου το ερωτηματολόγιο συντάχθηκε από την ερευνητική ομάδα με βάση ελληνικές και διεθνείς μελέτες. Τα στοιχεία συλλέχθηκαν με προσωπική συνέντευξη, αφού επισημάνθηκε σε κάθε ερωτώμενο, ότι μπορούσαν να μην απαντήσουν στις ερωτήσεις μας αλλά και ότι ανά πάσα στιγμή μπορούσαν να διακόψουν τη συμπλήρωση του ερωτηματολογίου.

Η συμπλήρωση του ερωτηματολογίου διαρκούσε περίπου 20 λεπτά της ώρας. Κριτήρια εισαγωγής στην έρευνά μας ήταν:

- Η ιδιότητα του ερωτώμενου σε σχέση με το ΤΕΙ
- Ο χώρος εργασίας του ερωτώμενου και κριτήρια αποκλεισμού ήταν:
- μη πλήρως συμπληρωμένα ερωτηματολόγια

- όχι άμεση σχέση ερωτώμενου με το ΤΕΙ

Τελικά χρησιμοποιήθηκαν όλα τα ερωτηματολόγια από αυτά που διανεμήθηκαν (σύνολο 400).

Ακολουθήθηκε πιστά ο κώδικας της Νυρεμβέργης και η διακήρυξη του Ελσίνκι για την προστασία των ανθρώπων από κάθε μορφής έρευνας με βάση τα δικαιώματα που έχει κανείς (να μην υποστεί κάποια βλάβη φυσική, συγκινησιακή κλπ, πλήρους διαφάνειας, ανωνυμίας και εχεμύθειας και αυτοδιάθεσης).

Για το λόγο αυτό πριν αρχίσει η συμπλήρωση του ερωτηματολογίου (κλειστού τύπου με δυνατότητες πολλαπλών απαντήσεων), εξηγήσαμε το σκοπό της έρευνάς μας, επιδιώκαμε τη μη παρεμπόδιση της φυσιολογικής ζωής και της παρεχόμενης εργασίας, σημειώναμε ότι το ερωτηματολόγιο ήταν ανώνυμο και το δείγμα (δηλαδή τα συμμετέχοντα πρόσωπα) τυχαίο, και τον φορέα της έρευνας - σχολή της φοίτησής μας. Αναλυτικά το ερωτηματολόγιο παρατίθεται στο Παράρτημα της παρούσας έρευνας.

Κάθε πιθανή απάντηση σε μία ερώτηση κωδικοποιήθηκε με ένα ακέραιο αριθμό ανάλογα με τον αριθμό των δυνατών απαντήσεων. Έπειτα τα δεδομένα εισήχθησαν στον ηλεκτρονικό υπολογιστή σε μεταβλητές που η κάθε μία αντιπροσώπευε μία ερώτηση. Το πρόγραμμα που χρησιμοποιήθηκε για την εισαγωγή των κωδικοποιημένων δεδομένων και τη στατιστική επεξεργασία τους ήταν το SPSS 20.00 για Windows 10. Τα αποτελέσματα που προέκυψαν συντάχθηκαν σε πίνακες στους οποίους αναφέρεται το όνομα της μεταβλητής καθώς και η αντίστοιχη ερώτηση στην οποία αναφέρεται. Επίσης αναφέρονται οι εξεταζόμενες ομάδες καθώς και τα σύνολα των απαντήσεων.

Με βάση τα παραπάνω έχουν εξαχθεί και τα συμπεράσματα από την έρευνά μας τα οποία και αναλύονται στην ΣΥΖΗΤΗΣΗ

Α. ΑΠΟΤΕΛΕΣΜΑΤΑ

5.2.1 ΠΕΡΙΓΡΑΦΙΚΗ ΚΑΤΑΝΟΜΗ ΤΟΥ ΔΕΙΓΜΑΤΟΣ

Τα αποτελέσματα παρουσιάζονται αναλυτικά με μορφή πινάκων, ενώ ακολουθεί αντίστοιχο σχήμα με ανάλογη γραφική παράσταση των αποτελεσμάτων για σαφέστερη παρουσίαση τους.

5.2.2 ΔΗΜΟΓΡΑΦΙΚΑ ΣΤΟΙΧΕΙΑ

ΠΙΝΑΚΑΣ 1: Κατανομή των απαντήσεων των ερωτηθέντων σε σχέση με το φύλο τους.

ΑΠΑΝΤΗΣΗ	ΑΡΙΘΜΟΣ	ΠΟΣΟΣΤΟ
ΑΝΔΡΑΣ	160	40
ΓΥΝΑΙΚΑ	240	60
ΣΥΝΟΛΟ	400	100 %

Οι περισσότεροι ερωτηθέντες στην παρούσα έρευνα ήταν γυναίκες (60 %).

ΠΙΝΑΚΑΣ 3: Κατανομή των απαντήσεων 400 ερωτηθέντων σε σχέση με την ιδιότητά τους.

ΑΠΑΝΤΗΣΗ	ΑΡΙΘΜΟΣ	ΠΟΣΟΣΤΟ
1 ^ο έτος	280	70
4 ^ο έτος	80	20
Άλλα έτη	40	10
ΣΥΝΟΛΟ	400	100 %

Οι περισσότεροι ερωτηθέντες στην παρούσα έρευνα ήταν σπουδαστές του 1^{ου} έτους.

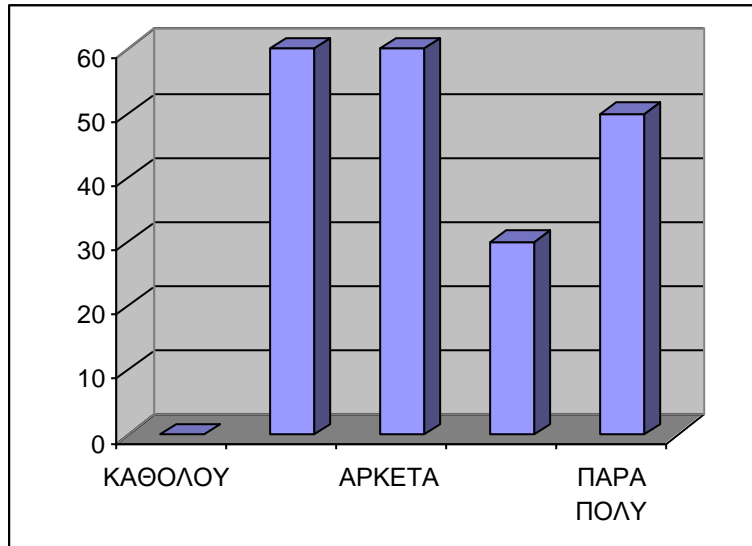
Η ηλικίες των ερωτώμενων ήταν από 18 έως 27 ετών με μέση τιμή τα 20 έτη.

5.2.3 ΕΡΩΤΗΜΑΤΟΛΟΓΙΟ

ΠΙΝΑΚΑΣ 4: Κατανομή των απαντήσεων 400 ερωτηθέντων σε σχέση με τις γνώσεις τους στους Η/Υ και το Internet.

ΑΠΑΝΤΗΣΗ	ΑΡΙΘΜΟΣ	ΠΟΣΟΣΤΟ
ΚΑΘΟΛΟΥ	0	0
ΛΙΓΟ	120	30
ΑΡΚΕΤΑ	190	30
ΠΟΛΥ	61	15
ΠΑΡΑ ΠΟΛΥ	100	25
ΣΥΝΟΛΟ	400	100 %

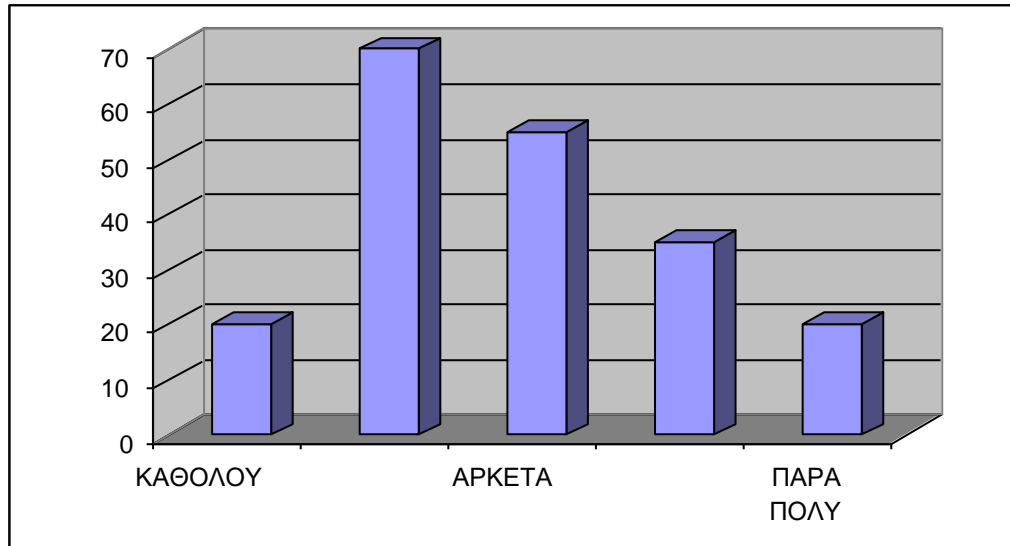
Οι περισσότεροι ερωτηθέντες (60 %) στην παρούσα έρευνα χαρακτηρίζουν τις γνώσεις τους ως λίγες ή αρκετές.



ΠΙΝΑΚΑΣ 5: Κατανομή των απαντήσεων 400 ερωτηθέντων σε σχέση με την πρόσβαση στο internet στο χώρο του ΤΕΙ.

ΑΠΑΝΤΗΣΗ	ΑΡΙΘΜΟΣ	ΠΟΣΟΣΤΟ
ΚΑΘΟΛΟΥ	41	10
ΛΙΓΟ	139	35
ΑΡΚΕΤΑ	111	27,5
ΠΟΛΥ	69	17,5
ΠΑΡΑ ΠΟΛΥ	40	10
ΣΥΝΟΛΟ	400	100 %

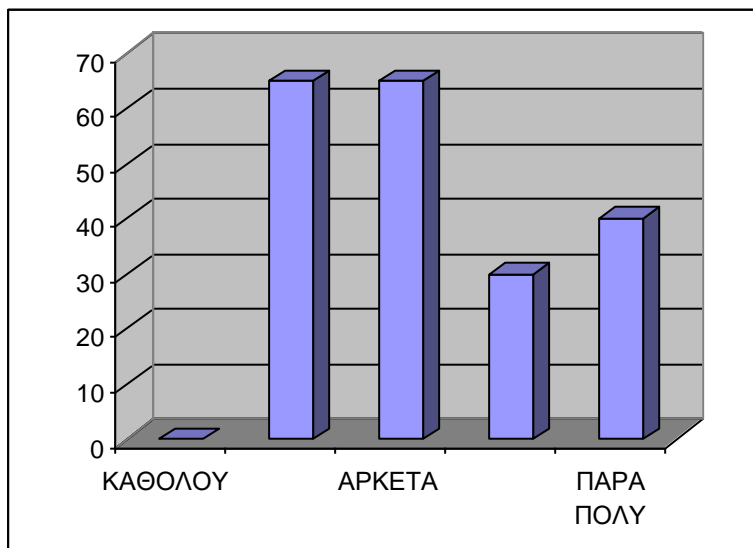
Οι περισσότεροι ερωτηθέντες στην παρούσα έρευνα χαρακτηρίζουν ως ικανοποιητική την πρόσβαση του Internet στο χώρο του ΤΕΙ



ΠΙΝΑΚΑΣ 6: Κατανομή των απαντήσεων 400 ερωτηθέντων σε σχέση με το αν γνωρίζουν για τις υπηρεσίες παροχής πληροφοριών και εκπαίδευσης σε θέματα υγείας από το internet.

ΑΠΑΝΤΗΣΗ	ΑΡΙΘΜΟΣ	ΠΟΣΟΣΤΟ
ΚΑΘΟΛΟΥ	0	0
ΛΙΓΟ	132	32,5
ΑΡΚΕΤΑ	130	32,5
ΠΟΛΥ	57	15
ΠΑΡΑ ΠΟΛΥ	81	20
ΣΥΝΟΛΟ	400	100 %

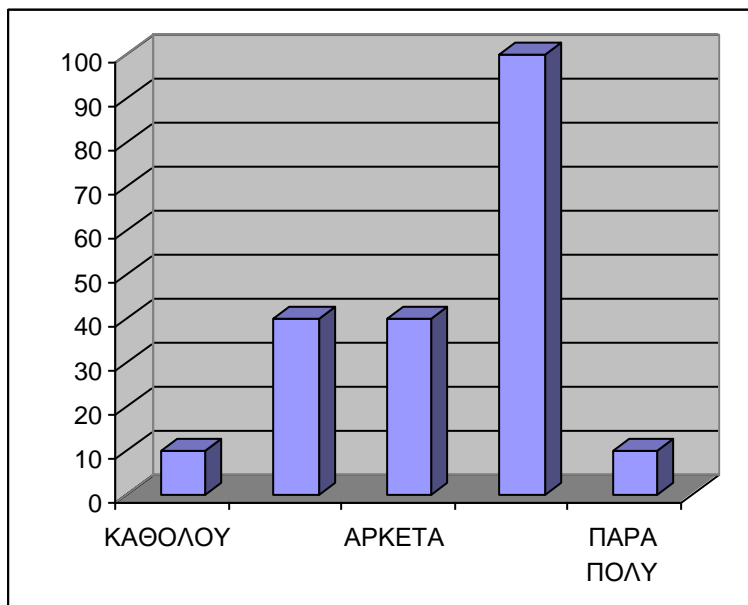
Οι περισσότεροι ερωτηθέντες στην παρούσα έρευνα γνωρίζουν για τις υπηρεσίες παροχής πληροφοριών και εκπαίδευσης σε θέματα υγείας από το internet.



ΠΙΝΑΚΑΣ 7: Κατανομή των απαντήσεων 400 ερωτηθέντων σε σχέση με το αν βρήκαν χρήσιμο και θα χρησιμοποιούσαν στο μέλλον το εργαλείο Biosite.

ΑΠΑΝΤΗΣΗ	ΑΡΙΘΜΟΣ	ΠΟΣΟΣΤΟ
ΚΑΘΟΛΟΥ	21	5
ΛΙΓΟ	80	20
ΑΡΚΕΤΑ	80	20
ΠΟΛΥ	202	50
ΠΑΡΑ ΠΟΛΥ	23	5
ΣΥΝΟΛΟ	400	100 %

Οι περισσότεροι ερωτηθέντες στην παρούσα έρευνα θα χρησιμοποιούσαν ΠΟΛΥ το σύστημα Biosite.



ΠΙΝΑΚΑΣ 8: Κατανομή των απαντήσεων 400 ερωτηθέντων σε σχέση με το αν τους προσφέρει περισσότερες πληροφορίες για τις οποίες θα ενδιαφέρονταν σε σχέση με άλλες σελίδες αναζήτησης ιατρικών πληροφοριών.

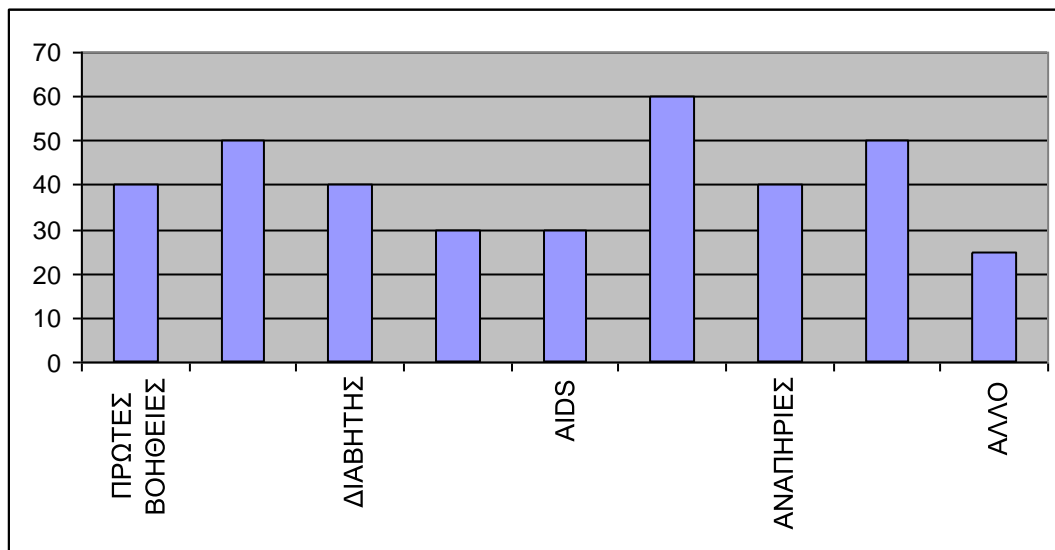
ΑΠΑΝΤΗΣΗ	ΑΡΙΘΜΟΣ	ΠΟΣΟΣΤΟ
Ίδιες	181	45
Περισσότερες	139	35
Λιγότερες	42	10
Δεν απαντώ	38	10
ΣΥΝΟΛΟ	400	100 %

Οι περισσότεροι ερωτηθέντες στην παρούσα έρευνα τους προσφέρει περισσότερες πληροφορίες για τις οποίες θα ενδιαφέρονταν σε σχέση με άλλες σελίδες αναζήτησης ιατρικών πληροφοριών

ΠΙΝΑΚΑΣ 9: Κατανομή των απαντήσεων 400 ερωτηθέντων σε σχέση με την προτίμησή τους σε υπηρεσίες παροχής εκπαίδευσης σε θέματα υγείας από το internet τα οποία θα πρότειναν να συμπεριληφθούν στο σύστημα Biosite με τις κατάλληλες οντολογίες που θα προστεθούν.

ΑΠΑΝΤΗΣΗ	ΑΡΙΘΜΟΣ	ΠΟΣΟΣΤΟ
ΠΡΩΤΕΣ ΒΟΗΘΕΙΕΣ	40	20
ΝΑΡΚΩΤΙΚΑ	7	25
ΑΝΑΤΟΜΙΑ	40	20
ΨΥΧΙΚΕΣ ΝΟΣΟΙ	60	15
AIDS	30	15
ΚΑΡΚΙΝΟΣ	60	30
ΑΝΑΠΗΡΙΕΣ	47	20
ΑΣΦΑΛΕΙΑ ΚΑΙ ΥΓΙΕΙΝΗ ΣΤΟ ΧΩΡΟ ΕΡΓΑΣΙΑΣ	95	25
ΑΛΛΟ	25	12,5

Οι περισσότεροι ερωτηθέντες στην παρούσα έρευνα επιθυμούν υπηρεσίες παροχής εκπαίδευσης σε θέματα υγείας από το internet που να έχουν σχέση με μεγάλη ποικιλία θεμάτων, τα οποία και πρότειναν να συμπεριληφθούν στο σύστημα Biosite με τις κατάλληλες οντολογίες που θα προστεθούν.



5.3 ΣΤΑΤΙΣΤΙΚΕΣ ΣΥΓΚΡΙΣΕΙΣ

Για να διαπιστωθεί αν ορισμένες κατηγορίες ερωτηθέντων έδωσαν διαφοροποιημένες απαντήσεις σε σχέση με κάποιο χαρακτηριστικό, χρησιμοποιήθηκαν ενδεικτικά και για λίγες περιπτώσεις λόγω του μικρού αριθμού του δείγματος μας, crosstabs με τα οποία συνδυάζονται οι απαντήσεις των 2 ερωτήσεων που μας ενδιαφέρουν. Κάθε κελί δίνει τον αριθμό και το επόμενο το ποσοστό επί του συνόλου των ερωτηθέντων.

Στο τέλος των crosstabs αναγράφονται τα αποτελέσματα που προέκυψαν από τον στατιστικό έλεγχο. Πιο συγκεκριμένα τα στατιστικά αποτελέσματα αποτελούνται από:

1. Μέγεθος του δείγματος
2. Πιθανότητα στατιστικής σημαντικότητας (**p**)

Θεωρούμε σαν στατιστικώς σημαντική μία διαφορά ως προς κάποιο χαρακτηριστικό, αν και μόνο αν το αποτέλεσμα που δίνεται από το στατιστικό έλεγχο οδηγεί σε μία πιθανότητα $p < 0.05$. Η στατιστική μέθοδος που χρησιμοποιήθηκε για τον έλεγχο των παρατηρούμενων διαφορών μεταξύ των εξεταζομένων ομάδων, ήταν το chi-square.

Πίνακας 1: Αποτελέσματα συσχέτισης των απαντήσεων των ερωτηθέντων με το έτος σπουδών τους.

A/A	Ερώτηση	1 ^ο έτος	4 ^ο έτος	Άλλα έτη	P
4	ΕΧΕΤΕ ΓΝΩΣΕΙΣ ΧΕΙΡΙΣΜΟΥ Η/Υ ΚΑΙ ΧΡΗΣΗΣ ΤΟΥ INTERNET;				> 0,05
	ΚΑΘΟΛΟΥ	140	40	20	
	ΛΙΓΟ	55	5	0	

	ΑΡΚΕΤΑ	45	5	10	
	ΠΟΛΥ	15	10	5	
	ΠΑΡΑ ΠΟΛΥ	25	20	0	
5	ΕΧΕΤΕ ΠΡΟΣΒΑΣΗ ΣΤΟ INTERNET ΣΤΟ ΧΩΡΟ ΤΟΥ Α.Τ.Ε.Ι.				< 0,05
	ΚΑΘΟΛΟΥ	10	0	0	
	ΛΙΓΟ	55	20	5	
	ΑΡΚΕΤΑ	15	35	5	
	ΠΟΛΥ	5	30	0	
	ΠΑΡΑ ΠΟΛΥ	5	5	10	
6	ΓΝΩΡΙΖΕΤΕ ΥΠΗΡΕΣΙΕΣ ΠΑΡΟΧΗΣ ΠΛΗΡΟΦΟΡΙΩΝ ΥΓΕΙΑΣ ΑΠΟ ΤΟ INTERNET;				> 0,05
	ΚΑΘΟΛΟΥ	0	0	0	
	ΛΙΓΟ	37	23	5	
	ΑΡΚΕΤΑ	10	0	10	
	ΠΟΛΥ	25	5	0	
	ΠΑΡΑ ΠΟΛΥ	30	10	0	
7	Αν βρήκαν χρήσιμο και θα χρησιμοποιούσαν στο μέλλον το εργαλείο Biosite				> 0,05
	ΚΑΘΟΛΟΥ	7	2	1	
	ΛΙΓΟ	15	15	5	
	ΑΡΚΕΤΑ	20	10	0	

	ΠΟΛΥ	85	3	12	
	ΠΑΡΑ ΠΟΛΥ	8	0	2	
8	Προσφέρει περισσότερες πληροφορίες για τις οποίες θα ενδιαφέρονταν σε σχέση με άλλες σελίδες αναζήτησης ιατρικών πληροφοριών				< 0,05
	Ίδιες	70	15	5	
	Περισσότερες	25	25	15	
	Λιγότερες	20	0	0	
	Δεν απαντώ	0	0	20	
9	Τα θέματα υγείας από το internet τα οποία θα πρότειναν να συμπεριληφθούν στο σύστημα Biosite με τις κατάλληλες οντολογίες που θα προστεθούν				> 0,05
	ΠΡΩΤΕΣ ΒΟΗΘΕΙΕΣ	20	10	10	
	ΝΑΡΚΩΤΙΚΑ	40	5	5	
	ΑΝΑΤΟΜΙΑ	30	0	10	
	ΨΥΧΙΚΕΣ ΝΟΣΟΙ	10	15	5	
	AIDS	13	12	5	
	ΚΑΡΚΙΝΟΣ	35	25	5	
	ΑΝΑΠΗΡΙΕΣ	20	10	10	

ΑΣΦΑΛΕΙΑ ΚΑΙ ΥΓΙΕΙΝΗ ΣΤΟ ΧΩΡΟ ΕΡΓΑΣΙΑΣ	25	15	10
ΑΛΛΟ	15	5	5

Με βάση τις παραπάνω συσχετίσεις διαπιστώνουμε ότι δεν υπάρχει στατιστικά σημαντική διαφορά μεταξύ των ερωτώμενων σπουδαστών, στα θέματα που αφορούν τη χρήση του Biosite και τις υπηρεσίες Υγείας μέσα από αυτό εκτός από

1. Την πρόσβαση τους στο internet με ιδιαίτερα παράπονα κυρίως από τους σπουδαστές των μεγάλων εξαμήνων.
2. Την τελική αξιολόγηση του συστήματος το οποίο και έχει σχέση περισσότερο με το αν είχαν ήδη εξοικειωθεί σε άλλο όπως το ίδιο το Medline η το ScholarGoogle.

ΚΕΦΑΛΑΙΟ 6

6.1 ΓΕΝΙΚΑ ΣΥΜΠΕΡΑΣΜΑΤΑ

Το σύστημα BioSite είναι ένα ολοκληρωμένο σύστημα συλλογής και ανάλυσης κειμένων βιοϊατρικού περιεχομένου και παροχής υποστήριξης αναζητήσεων και ερωτημάτων χρηστών. Παρόλο που οι στόχοι που τέθηκαν κατά το ξεκίνημα της παρούσας εργασίας επιτεύχθηκαν από την υλοποίηση και λειτουργία του συστήματος, υπάρχουν μερικά σημεία στα οποία θα μπορούσαν να υπάρξουν κάποιες μελλοντικές βελτιώσεις ή προσθήκες.

6.2 ΣΥΜΠΕΡΑΣΜΑΤΑ ΜΕ ΒΑΣΕΙ ΤΑ ΣΤΑΤΙΣΤΙΚΑ

Οι περισσότεροι ερωτηθέντες στην παρούσα έρευνα ήταν γυναίκες (60 %), και τα κεντρικά συμπεράσματα της μελέτης μας ήταν:

Οι περισσότεροι ερωτηθέντες (60 %) στην παρούσα έρευνα

- χαρακτηρίζουν τις γνώσεις τους γύρω από τους Η/Υ και το Internet ως λίγες ή απλά αρκετές.
- ως ικανοποιητική την πρόσβαση του Internet στο χώρο του ΤΕΙ (55%)
- γνωρίζουν για τις υπηρεσίες παροχής πληροφοριών και υπηρεσιών υγείας από το internet (67,5 %).
- θα χρησιμοποιούσαν πολύ υπηρεσίες παροχής πληροφοριών και υπηρεσιών υγείας από Biosite (75 %).
- θα ενδιαφέρονταν και για πληροφορίες μέσω του Biosite.
- επιθυμούν υπηρεσίες παροχής πληροφοριών και υπηρεσιών υγείας από το internet που να έχουν σχέση με μεγάλη ποικιλία θεμάτων.

Με βάση τις παραπάνω συσχετίσεις διαπιστώνουμε ότι δεν υπάρχει στατιστικά σημαντική διαφορά μεταξύ των ερωτώμενων σπουδαστών, στο ΤΕΙ στα θέματα που αφορούν τη χρήση του

Διαδικτύου και τις υπηρεσίες Υγείας μέσα από αυτό εκτός από

- την πρόσβαση τους στο internet με ιδιαίτερα παράπονα από τους σπουδαστές.
- τις υπηρεσίες τηλεϋγείας που θα επιθυμούσαν.

Πρέπει εδώ να σημειωθεί ότι όλοι σχεδόν οι ερωτώμενοι επέμειναν στο γεγονός ότι τις υπηρεσίες τηλεϋγείας πρέπει να τις υποστηρίζουν ειδικοί επιστήμονες ανά υπηρεσία οι οποίοι με εμφάνιση των προσωπικών τους στοιχείων να αναλαμβάνουν και την ευθύνη της υποστήριξης των χρηστών των υπηρεσιών.

Με βάση τα παραπάνω διαπιστώνουμε την μεγάλη αναγκαιότητα δημιουργία και υποστήριξης υπηρεσιών τηλεϋγείας από τις υπηρεσίες του ιδρύματος του ΤΕΙ.

ΒΙΒΛΙΟΓΡΑΦΙΑ

1. Μηνά Ι. (2017), Σύστημα Επεξεργασίας και Ανάλυσης Βιο-ιατρικής Βιβλιογραφίας με Στόχο την Υποστήριξη Ερωτημάτων, Μεταπτυχιακή Διπλωματική Εργασία, Υπευθ. Καθηγ. Ι Χατζηλυγερούδης, Τμήμα Μηχ. ΗΥ και Πληροφορικής, Πάτρα 2017.
2. Μάργαρης Α., Εισαγωγή στις Βάσεις Δεδομένων Κεχρής Ε. (2015), Σχεσιακές Βάσεις Δεδομένων
3. Λαγουμιντζής Γ., Βλαχόπουλος Γ., Κουτσογιάννης Κ. (2015), Μεθοδολογία Έρευνας στις Επιστήμες Υγείας. Εκδ ΚΑΛΛΙΠΟΣ (e-book).
4. Matthew S. Simpson & Dina Demner-Fushman (2012). Biomedical Text Mining: A survey of recent Progress, in Mining Text Data, 465-517.
5. Arvind C. Rajasekaran (2014). Categorization of biomedical articles with spectral clustering , Final Project Report.
6. Jon Rune Paulsen & Heri Ramampiaro (2009). Combining Latent Semantic Indexing and Clustering to Retrieve and Cluster Biomedical Information: A2-step Approach, presented at NIK-2009 conference.
7. Yuka Tateisi & Jun-ichi Tsujii (2004). Part-of-Speech Annotation of Biology Research Abstracts, Proceedings of 4th International Conference on Language Resources and Evaluation (LREC).
8. Jin-Dong Kim, Tomoko Ohta & Jun-ichi Tsujii (2008). Corpus annotation for mining biomedical events from literature, BMC Bioinformatics.
9. Yuka Tateisi, Akane Yakushiji, Tomoko Ohta & Jun-ichi Tsujii (2005). Syntax annotation for the GENIA corpus.
10. Jin-Dong Kim, Tomoko Ohta, Yuka Tateisi & Jun-ichi Tsujii (2006).GENIA Corpus Manual Encoding schemes for the corpus and annotation , Report from Tsujii Laboratory, University of Tokyo.
11. Jin-Dong Kim & Jun-ichi Tsujii (2006). GENIA Corpus Curation Framework , Report from Tsujii Laboratory, University of Tokyo.
12. Jin-Dong Kim, Tomoko Ohta, Yuka Tateisi & Jun-ichi Tsujii (2006). GENIA Ontology, Report from Tsujii Laboratory, University of Tokyo
13. Tomoko Ohta, Jin-Dong Kim & Jun-ichi Tsujii (2007). Guidelines for event annotation , Department of Information Science, Graduate School of Science, University of Tokyo.
14. Schweiger D, Trajanoski Z & Pabinger S.(2014). SPARQLGraph: aweb-based platform for graphically querying biological Semantic Web databases, BMC Bioinformatics.

15. Fabio Rinaldi, Gerold Schneider, Kaarel Kaljurand, Michael Hess & Martin Romacker (2006). An environment for relation mining over richly annotated corpora: the case of GENIA, BMC Bioinformatics.
16. Medline : <https://www.nlm.nih.gov/bsd/pmresources.html>
17. National Library of Medicine : <https://www.nlm.nih.gov/>
18. PubMed Central : <https://www.ncbi.nlm.nih.gov/pmc/>
19. Springerlink : <http://link.springer.com/>
20. LinkOut : <https://www.ncbi.nlm.nih.gov/projects/linkout/>
21. GATE : <https://gate.ac.uk/>
22. GENIA Tagger : <http://www.nactem.ac.uk/GENIA/tagger/>
23. Basic Local Alignment Search Tool – BLAST: <https://blast.ncbi.nlm.nih.gov/Blast.cgi>
24. SmartBlast:<https://blast.ncbi.nlm.nih.gov/smartblast/>
LINK_LOC=BlastHomeLink
25. Μπεσμπέας Σταύρος, (1994). *Τεχνολογική Εξέλιξη, Πρόληψη και Έγκαιρη Διάγνωση του Καρκίνου, Πλεονεκτήματα-Μειονεκτήματα Μικροϋπολογιστών, Βοηθήματα Νοσηλευτικής Πρόληψης και Έγκαιρης Διάγνωσης του Καρκίνου, Επίτομος, Έκδοση Πρώτη, Εκδόσεις «Αντικαρκινική Εταιρεία», Αθήνα 1994, σ.7*
26. Μπότσαρης Χαράλαμπος, *Υγεία και Πληροφορική, Πληροφορική Νέες Τεχνολογίες και Υγεία, Τεύχος 3, Τόμος 1, Θεσσαλονίκη 1991, σ.7-8.*
27. Βενιεράκης Γεώργιος, *Εξέλιξη της Πληροφορικής, Ιστορία, Τύποι και Επιλογές Υπολογιστών, Θεωρία και Πράξη, Βοηθήματα Νοσηλευτικής Πρόληψης και Έγκαιρης Διάγνωσης του Καρκίνου, Επίτομος, Έκδοση Πρώτη, Εκδόσεις «Αντικαρκινική Εταιρεία», Αθήνα 1994, σ.21-23.*
28. GoIdschIager Les and Lister Andrew (1996), *Εισαγωγή στη Σύγχρονη Επιστήμη των Υπολογιστών, Μετάφραση Χαλάτσης Κώστας, Επίτομος, Έκδοση Τρίτη, Εκδόσεις Δίαυλος, Αθήνα 1996, σ.25.*
29. Elmasri R.- Navathe S.B. (1996), *Θεμελιώδεις Αρχές Συστημάτων Βάσεων Δεδομένων, Μετάφραση Χατζόπουλος Μιχάλης, Τόμος 1, Έκδοση Δεύτερη, Εκδόσεις Δίαυλος, Αθήνα 1996, σ.26.*

30. Tanenbaum S. Andrew, *Δίκτυα Υπολογιστών*, Μετάφραση Στυλιανάκης Βασίλειος, Επίτομος, Έκδοση Τρίτη, Εκδόσεις Παπασωτηρίου, Αθήνα 2000, σ.2
31. 17.Παναγοπούλου Μαρία, Διπλωματική Εργασία «*Αλγόριθμοι Και Μοριακή Βιοπληροφορική*», Επιβλέπων Τσακαλίδης Αθ.,Τμήμα Μηχ. Η/Υ.και Πληροφορικής ,Πάτρα, Οκτώβριος 1994, σ.1-12
32. Σβύνου Κωνσταντίνα, Μεταπτυχιακή Διπλωματική Εργασία «*Τεχνολογίες Δικτύων Με Εφαρμογές Στην Τηλεϊατρική*», Υπεύθυνος Καθηγητής Λυκοθανάσης Ι Τμήμα Πληροφορικής Πάτρα 2000, σ.1-30.
33. Πραστάκος Π. Γρηγ., *Αλληλεπίδραση Ανθρώπου --Υπολογιστή και Επιπτώσεις στο Ανθρώπινο Δυναμικό*, Βοηθήματα Νοσηλευτικής Πρόληψης και Έγκαιρης Διάγνωσης του Καρκίνου, Επίτομος, Έκδοση Πρώτη, Εκδόσεις Αντικαρκινική Εταιρία, Αθήνα 1994, σ.79-82

ΠΑΡΑΡΤΗΜΑ Ι

ΕΡΩΤΗΜΑΤΟΛΟΓΙΟ

1. ΦΥΛΟ

ΑΝΔΡΑΣ ΓΥΝΑΙΚΑ

2. ΗΛΙΚΙΑ _____

3. ΕΤΟΣ ΣΠΟΥΔΩΝ

1^ο ΕΤΟΣ 4^ο ΕΤΟΣ ΑΛΛΟ ΕΤΟΣ

4. ΕΧΕΤΕ ΓΝΩΣΕΙΣ ΧΕΙΡΙΣΜΟΥ Η/Υ ΚΑΙ ΧΡΗΣΗΣ ΤΟΥ INTERNET;

ΚΑΘΟΛΟΥ ΛΙΓΟ ΑΡΚΕΤΑ ΠΟΛΥ ΠΑΡΑ ΠΟΛΥ

5. ΕΧΕΤΕ ΠΡΟΣΒΑΣΗ ΣΤΟ INTERNET ΣΤΟ ΧΩΡΟ ΤΟΥ Α.Τ.Ε.Ι.

ΚΑΘΟΛΟΥ ΛΙΓΟ ΑΡΚΕΤΑ ΠΟΛΥ ΠΑΡΑ ΠΟΛΥ

6. ΓΝΩΡΙΖΕΤΕ ΥΠΗΡΕΣΙΕΣ ΠΑΡΟΧΗΣ ΠΛΗΡΟΦΟΡΙΩΝ ΚΑΙ ΥΠΗΡΕΣΙΩΝ ΥΓΕΙΑΣ ΑΠΟ ΤΟ INTERNET;

ΚΑΘΟΛΟΥ ΛΙΓΟ ΑΡΚΕΤΑ ΠΟΛΥ ΠΑΡΑ ΠΟΛΥ

7. ΒΡΗΚΑΤΕ ΧΡΗΣΙΜΟ ΚΑΙ ΘΑ ΧΡΗΣΙΜΟΠΟΙΟΥΣΑΝ ΣΤΟ ΜΕΛΛΟΝ ΤΟ ΕΡΓΑΛΕΙΟ BIOSITE;

ΚΑΘΟΛΟΥ ΛΙΓΟ ΑΡΚΕΤΑ ΠΟΛΥ ΠΑΡΑ ΠΟΛΥ

8. ΠΡΟΣΦΕΡΕΙ ΠΕΡΙΣΣΟΤΕΡΕΣ ΠΛΗΡΟΦΟΡΙΕΣ ΓΙΑ ΤΙΣ ΟΠΟΙΕΣ ΘΑ ΕΝΔΙΑΦΕΡΟΝΤΑΝ ΣΕ ΣΧΕΣΗ ΜΕ ΑΛΛΕΣ ΣΕΛΙΔΕΣ ΑΝΑΖΗΤΗΣΗΣ ΙΑΤΡΙΚΩΝ ΠΛΗΡΟΦΟΡΙΩΝ.

ΙΔΙΕΣ ΠΕΡΙΣΣΟΤΕΡΕΣ ΛΙΓΟΤΕΡΕΣ ΔΕΝ ΑΠΑΝΤΩ

9. ΣΗΜΕΙΩΣΤΕ ΤΙΣ ΠΡΟΤΙΜΗΣΕΙΣ ΣΑΣ ΣΤΙΑ ΠΑΡΑΚΑΤΩ ΠΙΘΑΝΕΣ ΥΠΗΡΕΣΙΕΣ ΥΓΕΙΑΣ ΜΕΣΩ INTERNET ΠΟΥ ΘΑ ΘΕΛΑΤΕ ΝΑ ΣΥΜΠΕΡΙΛΗΦΘΟΥΝ ΣΤΟ Biosite.

ΠΡΩΤΕΣ ΒΟΗΘΕΙΕΣ □ ΝΑΡΚΩΤΙΚΑ □ ΑΝΑΤΟΜΙΑ □ ΨΥΧΙΚΕΣ ΝΟΣΟΙ □
AIDS □ ΚΑΡΚΙΝΟΣ □ ΑΝΑΠΗΡΙΕΣ □ ΑΣΦΑΛΕΙΑ ΚΑΙ ΥΓΙΕΙΝΗ ΣΤΟ ΧΩΡΟ
ΕΡΓΑΣΙΑΣ □ ΑΛΛΟ