



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

Προηγμένη Εφαρμογή Ιστού Διαχείρισης Δεδομένων Βιοεπιστημών

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

ΓΕΩΡΓΙΟΥ ΠΡΕΚΑ

Επιβλέπων : Τιμολέων Σελλής
Καθηγητής Ε.Μ.Π.

Αθήνα, Δεκέμβριος 2009

Η σελίδα αυτή είναι σκόπιμα λευκή.



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

Προηγμένη Εφαρμογή Ιστού Διαχείρισης Δεδομένων Βιοεπιστημών

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

ΓΕΩΡΓΙΟΥ ΠΡΕΚΑ

Επιβλέπων : Τιμολέων Σελλής
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 14^η Ιανουαρίου 2010.

(Υπογραφή)

.....
Τιμολέων Σελλής
Καθηγητής Ε.Μ.Π.

(Υπογραφή)

.....
Ιωάννης Βασιλείου
Καθηγητής Ε.Μ.Π.

(Υπογραφή)

.....
Νεκτάριος Κοζύρης
Αν. Καθηγητής Ε.Μ.Π.

Αθήνα, Δεκέμβριος 2009

(Υπογραφή)

.....

ΓΕΩΡΓΙΟΣ ΠΡΕΚΑΣ

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

© 2009 – All rights reserved

Περίληψη

Σκοπός της διπλωματικής εργασίας ήταν η ανάπτυξη μίας προηγμένης εφαρμογής ιστού για την αναζήτηση και παρουσίαση βιολογικών πληροφοριών. Η εφαρμογή αφορά την έρευνα γύρω από τα μόρια microRNA των οποίων η μελέτη της συνεισφοράς στη λειτουργία των οργανισμών παρουσιάζει ιδιαίτερη ανάπτυξη τα τελευταία χρόνια. Συγκεκριμένα, οι ερευνητές - βιολόγοι ενδιαφέρονται για τη γνώση των βιολογικών στόχων των συγκεκριμένων μορίων πάνω στα μόρια που κωδικοποιούν τις πρωτεΐνες, ούτως ώστε να εξάγουν χρήσιμα συμπεράσματα τα οποία μπορούν να δώσουν λύσεις σε διάφορες παθήσεις. Η εκπόνηση της διπλωματικής πραγματοποιήθηκε σε συνεργασία με το Ερευνητικό Κέντρο Βιοϊατρικών Επιστημών «Αλέξανδρος Φλέμινγκ», το οποίο έχει αναπτύξει κατάλληλο αλγόριθμο για την πρόβλεψη των συγκεκριμένων στόχων. Επιπρόσθετα, η εφαρμογή παρέχει υπηρεσίες εξατομίκευσης, οι οποίες στοχεύουν στη βελτίωση της εμπειρίας των χρηστών. Οι συγκεκριμένες υπηρεσίες περιλαμβάνουν την τήρηση του ιστορικού των αναζητήσεων καθώς και τη δυνατότητα αποθήκευσης και επεξεργασίας αγαπημένων αναζητήσεων στον προσωπικό χώρο του κάθε χρήστη. Η ανάπτυξη της εφαρμογής ακολούθησε σύγχρονες τεχνικές και μεθοδολογίες, όπως είναι η χρήση του αρχιτεκτονικού προτύπου Model - View - Controller καθώς και η χρήση unit testing, έτσι ώστε να διευκολύνει τη μετέπειτα συντήρηση και αναβάθμιση των δυνατοτήτων του.

Λέξεις Κλειδιά: εφαρμογή ιστού, βιολογία, μόρια microRNA, γονίδια, στόχοι microRNA, Yii PHP framework, unit testing, πρότυπο Model - View - Controller

Η σελίδα αυτή είναι σκόπιμα λευκή.

Abstract

The scope of this thesis was the development of an advanced web application for searching and presenting biological information. The application concerns the research surrounding microRNA molecules which are studied for their effects on the function of the living organisms with great interest in the past few years. More specifically, researchers - biologists are interested in knowing the biological targets of the specific molecules onto the molecules that code proteins, so that they can arrive at useful conclusions which can give solution for different diseases. The preparation of the thesis was realized in cooperation with the Biomedical Sciences Research Center "Alexander Fleming", which has developed a suitable algorithm for predicting such targets. Moreover, the application provides personalization services, which aim to improve the user experience. These services comprise of history keeping of searches and the ability to save and edit favorite searches into the user's personal space. The application's development followed modern techniques and methodologies, such as the use of the Model - View - Controller architectural pattern and the use of unit testing, so that it facilitates the subsequent maintenance and upgrade of the capabilities offered.

Keywords: web application, biology, microRNA molecules, genes, microRNA targets, Yii PHP framework, unit testing, Model - View - Controller pattern

Η σελίδα αυτή είναι σκόπιμα λευκή.

Πίνακας περιεχομένων

1	Εισαγωγή.....	13
1.1	Βιολογική έρευνα των μορίων microRNA	13
1.2	Αντικείμενο διπλωματικής.....	14
1.2.1	Συνεισφορά.....	15
1.3	Οργάνωση κειμένου.....	15
2	Θεωρητικό υπόβαθρο και σχετικές εργασίες	17
2.1	Το κεντρικό δόγμα της μοριακής βιολογίας	17
2.2	Η ανακάλυψη των microRNA	20
2.3	Πρόβλεψη στόχων των microRNA.....	21
2.4	Σχετικές εργασίες.....	21
2.4.1	DIANA-microT 3.0	22
2.4.2	miRanda.....	22
2.4.3	MicroCosm	23
2.4.4	EIMMo	23
2.4.5	Pictar.....	24
2.4.6	Targetscan 5.0.....	24
2.4.7	PITA	24
2.4.8	RNA22.....	25
3	Ανάλυση Απαιτήσεων Συστήματος.....	27
3.1	Αρχιτεκτονική.....	27
3.2	Μοντέλο Οντοτήτων Συσχετίσεων	28
3.3	Περιγραφή Λειτουργιών	31
3.3.1	Εισαγωγή στοιχείων	31
3.3.2	Επαλήθευση στοιχείων	32
3.3.3	Παροχή προτάσεων για τους όρους της αναζήτησης	32
3.3.4	Αναζήτηση βιολογικών στόχων.....	33
3.3.5	Παρουσίαση αποτελεσμάτων	33
3.3.6	Διαχείριση χρηστών.....	33

3.3.7	Ιστορικό αναζητήσεων	34
3.3.8	Τήρηση αγαπημένων αναζητήσεων (bookmarks)	34
3.4	Περιπτώσεις χρήσης.....	35
3.4.1	Περίπτωση χρήσης «Αναζήτηση»	37
3.4.2	Περίπτωση χρήσης «Είσοδος στο σύστημα»	38
3.4.3	Περίπτωση χρήσης «Προσθήκη στις αγαπημένες αναζητήσεις»	39
3.4.4	Περίπτωση χρήσης «Δημιουργία λογαριασμού».....	40
3.4.5	Περίπτωση χρήσης «Αποσύνδεση από το σύστημα»	41
3.4.6	Περίπτωση χρήσης «Προβολή αγαπημένης αναζήτησης»	42
3.4.7	Περίπτωση χρήσης «Διαγραφή αγαπημένης αναζήτησης».....	42
3.4.8	Περίπτωση χρήσης «Καθαρισμός ιστορικού αναζητήσεων»	43
3.4.9	Περίπτωση χρήσης «Μεταβολή κριτηρίων αναζήτησης»	44
3.4.10	Περίπτωση χρήσης «Αλλαγή περιγραφής αγαπημένης αναζήτησης»	45
3.4.11	Περίπτωση χρήσης «Μεταβολή βιολογικών στόχων αγαπημένης αναζήτησης».....	46
4	Σχεδίαση Συστήματος	49
4.1	Αρχιτεκτονική.....	50
4.1.1	Model.....	51
4.1.2	View.....	51
4.1.3	Controller.....	51
4.2	Yii PHP framework.....	52
4.2.1	Ταχύτητα.....	53
4.2.2	Ποιότητα κώδικα	54
4.2.3	Data Access Objects και σχεδιαστικό πρότυπο Active Record.....	54
4.2.4	Ενσωμάτωση του jQuery.....	54
4.2.5	Πιστοποίηση χρηστών	55
4.2.6	Εισαγωγή δεδομένων από HTML φόρμες.....	55
4.2.7	Δυνατότητες παραμετροποίησης της εμφάνισης της εφαρμογής.....	55
4.2.8	Υποστήριξη για web services	55
4.2.9	Υποστήριξη πολλαπλών γλωσσών	55
4.2.10	Δυνατότητες caching πολλαπλών επιπέδων.....	56
4.2.11	Διαχείριση σφαλμάτων και καταγραφή συμβάντων	56

4.2.12	Ασφάλεια.....	56
4.2.13	Συμβατότητα με πρότυπα.....	56
4.2.14	Εργαλεία αυτόματης δημιουργίας κώδικα	56
4.2.15	Τεκμηρίωση.....	57
4.2.16	Δυνατότητα επεκτάσεων	57
4.2.17	Ενεργή ανάπτυξη και υποστήριξη.....	57
4.3	Περιγραφή Κλάσεων.....	57
4.3.1	Κλάσεις Active Record.....	58
4.3.2	Κλάσεις Form Model.....	61
4.3.3	Άλλα μοντέλα	63
4.3.4	Ελεγκτές.....	64
4.3.5	Widgets.....	66
4.4	Βάση Δεδομένων	74
5	Υλοποίηση.....	77
5.1	Επιλογή μηχανής αποθήκευσης	77
5.2	Βελτιστοποίηση των ερωτημάτων στη βάση δεδομένων.....	79
5.3	Full text search.....	81
5.4	Ευρετήρια.....	81
5.5	Μηχανή προτάσεων για τις αναζητήσεις των χρηστών του συστήματος	83
5.6	Επαναχρησιμοποιήσιμο στοιχείο λίστας με λεπτομέρειες.....	86
5.7	Πλατφόρμες και προγραμματιστικά εργαλεία	88
5.7.1	Apache HTTP Server.....	88
5.7.2	MySQL	89
5.7.3	PHP.....	90
5.7.4	PHPUnit.....	90
6	Έλεγχος	93
6.1	Test fixture	93
6.2	Unit testing.....	95
6.2.1	Παράδειγμα unit testing #1.....	97
6.2.2	Παράδειγμα unit testing #2.....	97
6.2.3	Παράδειγμα unit testing #3.....	98
6.3	System testing	98

7	Επίλογος.....	105
7.1	Σύνοψη και συμπεράσματα.....	105
7.2	Μελλοντικές επεκτάσεις.....	108
8	Βιβλιογραφία.....	111

1

Εισαγωγή

1.1 Βιολογική έρευνα των μορίων microRNA

Είναι πλέον γεγονός πως η σύγχρονη ερευνητική δραστηριότητα σε όλους τους τομείς και κλάδους της επιστήμης ανατρέπει διαρκώς την εδραιωμένη γνώση προσθέτοντας νέες πληροφορίες που έρχονται να συμπληρώσουν τον τρόπο με τον οποίο οι επιστήμες βλέπουν και περιγράφουν τον κόσμο. Από αυτή τη διαδικασία δε θα μπορούσε να λείπει η βιολογία και ιδιαίτερα ο κλάδος της μοριακής βιολογίας. Από το 1953, οπότε και ανακαλύφθηκε η δομή του DNA από τους James Watson και Francis Crick, η μοριακή βιολογία γνωρίζει ιδιαίτερη ανάπτυξη προσθέτοντας πληροφορίες για τη λειτουργία των βιολογικών οργανισμών που μέχρι πρότινος ήταν εντελώς άγνωστες.

Μέσω αυτής της αέναης διαδικασίας έρευνας και αναζήτησης των μοριακών βιολόγων προέκυψε και η γνώση των μορίων microRNA. Τα συγκεκριμένα μόρια περιγράφηκαν το 1993 ενώ ο όρος χρησιμοποιήθηκε για πρώτη φορά το 2001 σε μία σειρά από τρία άρθρα στο επιστημονικό περιοδικό Science. Καθώς η εξερεύνηση των συγκεκριμένων μορίων επεκτάθηκε, γινόταν σαφής ο σημαντικός ρόλος που διαδραματίζουν για τη λειτουργία των οργανισμών. Η γνώση γύρω από το microRNA αποτελεί εφόδιο για τους ερευνητές – βιολόγους στην προσπάθειά τους να κατανοήσουν τους μηχανισμούς της ζωής αλλά και να προσφέρουν στο κοινωνικό σύνολο, προτείνοντας θεραπείες για ασθένειες.

Τα μόρια microRNA είναι βιολογικά μόρια, τα οποία ελέγχουν την παραγωγή πρωτεϊνών του οργανισμού. Κάθε μόριο microRNA έχει τη δυνατότητα να συνδεθεί με ένα μόριο mRNA, το οποίο είναι υπεύθυνο για την παραγωγή μίας πρωτεΐνης. Εφόσον επιτευχθεί η σύνδεση τότε το μόριο mRNA καταστρέφεται και αναστέλλεται η παραγωγή της συγκεκριμένης πρωτεΐνης. Κατά αυτόν τον τρόπο λειτουργούν ρυθμιστικά, στη σύνθεση πρωτεϊνών από το κύτταρο. Αυτή η δράση έχει ιδιαίτερη σημασία μιας και η σύνθεση ή όχι των πρωτεϊνών είναι ζωτικής σημασίας για τον οργανισμό. Πολλές ασθένειες, για παράδειγμα, οφείλονται στην παραγωγή συγκεκριμένων πρωτεϊνών από τον οργανισμό.

Για τη μελέτη της λειτουργίας των συγκεκριμένων μορίων είναι επιτακτική η ανάγκη γνώσης των δυνατών συνδέσεων με τα μόρια mRNA. Η εύρεση των λεγόμενων στόχων (targets) μπορεί να πραγματοποιηθεί είτε πειραματικά είτε μέσω προσομοίωσης. Η ερευνητική ομάδα του Ερευνητικού Κέντρου Βιοϊατρικών Επιστημών «Αλέξανδρος Φλέμινγκ», η οποία συνεργάστηκε για την πραγματοποίηση της παρούσας διπλωματικής εργασίας έχει αναπτύξει έναν αλγόριθμο πρόβλεψης στόχων, του οποίου τα αποτελέσματα φιλοδοξεί να παρουσιάσει το σύστημα που αναπτύχθηκε στα πλαίσια της παρούσας διπλωματικής εργασίας.

1.2 Αντικείμενο διπλωματικής

Η παρούσα διπλωματική εργασία μαζί με το λογισμικό που αναπτύχθηκε κατά τη διάρκεια εκπόνησής της έρχεται να συνεισφέρει στον τομέα της ερευνητικής δραστηριότητας των βιολόγων γύρω από τα μόρια microRNA. Το Ερευνητικό Κέντρο Βιοϊατρικών Επιστημών «Αλέξανδρος Φλέμινγκ» έχει αναπτύξει έναν αλγόριθμο που έχει τη δυνατότητα να προβλέπει τη λειτουργία των μορίων microRNA έτσι ώστε να έχει τη δυνατότητα ο βιολόγος να αντιλαμβάνεται τον τρόπο δράσης των συγκεκριμένων μορίων στους οργανισμούς. Η εφαρμογή του συγκεκριμένου αλγορίθμου παράγει μεγάλο όγκο από δεδομένα, τα οποία στην πρωταρχική μορφή τους είναι δύσκολα στην επεξεργασία και κατανόηση. Το αντικείμενο της διπλωματικής είναι η άντληση των συγκεκριμένων δεδομένων, η επεξεργασία τους και η παρουσίασή τους με τρόπο απλό και κατανοητό στον ερευνητή – βιολόγο, ο οποίος χρησιμοποιεί το σύστημα έτσι ώστε να βοηθηθεί στην έρευνά του.

Το σύστημα που αναπτύχθηκε λειτουργεί στον παγκόσμιο ιστό έτσι ώστε η πρόσβαση σε αυτό να είναι ανεμπόδιστη και, επιπλέον, να υπάρχει δυνατότητα χρήσης αυτού από βιολόγους παγκοσμίως. Οι τεχνικές δυσκολίες ανάπτυξης ενός συστήματος όπως το συγκεκριμένο είναι ο χειρισμός μεγάλου όγκου δεδομένων και οι απαιτήσεις ταχύτητας και ευχρηστίας που πρέπει να διαθέτει ένα σύστημα το οποίο φιλοδοξεί να απευθύνεται σε ένα ευρύ κοινό. Επιπλέον, βασικός στόχος στην ανάπτυξη του συστήματος ήταν η μελέτη και η προσφορά υπηρεσιών πρόσθετης αξίας προς τους χρήστες του συστήματος. Αυτές οι υπηρεσίες προσδιορίστηκαν

στην παροχή εξατομίκευσης, δηλαδή στην εξυπηρέτηση των προσωπικών απαιτήσεων του κάθε χρήστη ξεχωριστά.

1.2.1 Συνεισφορά

Η παρούσα διπλωματική εργασία και το σύστημα που τη συνοδεύει πραγματοποιήθηκε έτσι ώστε να επιτευχθούν μία σειρά από σημαντικούς στόχους. Αυτοί οι στόχοι είχαν αρχικά τεθεί από την κοινότητα των ερευνητών – βιολόγων και ορίζονταν από τις απαιτήσεις που έχουν λόγω της καθημερινής ερευνητικής δραστηριότητάς τους. Στο συγκεκριμένο πλαίσιο, η συγκεκριμένη διπλωματική εργασία συνεισέφερε με μοναδικό τρόπο όπως περιγράφεται στη συνέχεια:

- Δημιουργία ενός χρήσιμου εργαλείου για την ερευνητική κοινότητα των βιολόγων που φιλοδοξεί να βρεθεί ανάμεσα στα κορυφαία παρόμοια προϊόντα σε παγκόσμια κλίμακα.
- Επίτευξη των στόχων ταχύτητας και ευχρηστίας του συστήματος, έτσι ώστε να παρέχεται άριστη εμπειρία κατά τη χρήση του συστήματος από τους χρήστες του.
- Ανάπτυξη του συστήματος χρησιμοποιώντας το σύγχρονο αρχιτεκτονικό πρότυπο Model – View – Controller, το οποίο διαχωρίζει το επίπεδο του business logic από το επίπεδο της παρουσίασης παρέχοντας με αυτόν τον τρόπο ευελιξία κατά την ανάπτυξη και συντήρηση του συστήματος.
- Συγγραφή κώδικα, ο οποίος εκτός από τη στήριξη της δημιουργία του συστήματος της παρούσας διπλωματικής εργασίας μπορεί να λειτουργήσει και ως πλατφόρμα ανάπτυξης μελλοντικών παρόμοιων εφαρμογών.
- Χρήση σύγχρονων τεχνικών ανάπτυξης και ελέγχου του συστήματος με στόχο την ελάττωση των προβλημάτων του λογισμικού, τη δυνατότητα εύκολης παραμετροποίησής του και τη δυνατότητα για εύκολη αναβάθμιση του συστήματος σύμφωνα με τις ανάγκες για μελλοντικές επεκτάσεις.
- Υλοποίηση καινοτόμων υπηρεσιών εξατομίκευσης προς τον ερευνητή – βιολόγο, όπως είναι η παροχή ενός προσωπικού χώρου μέσω του οποίου έχει τη δυνατότητα να προβάλλει το ιστορικό των τελευταίων αναζητήσεών του στο σύστημα, καθώς επίσης και τη δυνατότητα να αποθηκεύσει τις συχνές αναζητήσεις του σε bookmarks, τα οποία μπορεί να επεξεργαστεί και να παραμετροποιήσει περαιτέρω.

1.3 Οργάνωση κειμένου

Η παρούσα διπλωματική εργασία έχει οργανωθεί σε κεφάλαια που ακολουθούν και περιγράφουν τη μεθοδολογία ανάπτυξης του συστήματος, η οποία εφαρμόστηκε κατά τη διάρκεια

ανάλυσης, σχεδίασης και υλοποίησης του συστήματος της εργασίας. Τα κεφάλαια που ακολουθούν περιγράφονται συνοπτικά στη συνέχεια.

Στο 2^ο κεφάλαιο πραγματοποιείται μία σύντομη επισκόπηση των διαθέσιμων συστημάτων λογισμικού που επιτελούν παρόμοια λειτουργία με το σύστημα που αναπτύχθηκε στα πλαίσια της διπλωματικής εργασίας. Στο 3^ο κεφάλαιο δίνεται μία περιεκτική εισαγωγή σε θέματα βιολογίας, των οποίων η γνώση είναι σημαντική έτσι ώστε ο αναγνώστης να είναι σε θέση να αντιληφθεί και να αξιολογήσει το σύστημα που αναπτύχθηκε.

Στο 4^ο κεφάλαιο ξεκινάει η ουσιαστική ανάπτυξη του συστήματος με την ανάλυση απαιτήσεων αυτού. Δίνεται ένα περίγραμμα της αρχιτεκτονικής που πρόκειται να χρησιμοποιηθεί και περιγράφονται αναλυτικά τα υποσυστήματα που αποτελούν την εφαρμογή. Στη συνέχεια δίνονται αναλυτικές περιπτώσεις χρήσης για όλα τα σενάρια που οφείλει να καλύψει το σύστημα. Τέλος, παρουσιάζεται το εννοιολογικό μοντέλο της σχεσιακής βάσης δεδομένων που συνοδεύει την εφαρμογή.

Στο 5^ο κεφάλαιο ακολουθεί η λεπτομερής σχεδίαση του συστήματος. Αναλύεται η αρχιτεκτονική Model – View – Controller στα συστατικά της, γίνεται αναφορά και αξιολόγηση στο πλαίσιο ανάπτυξης (framework) που πρόκειται να χρησιμοποιηθεί κατά την ανάπτυξη (Yii PHP framework) και περιγράφονται οι κλάσεις από τις οποίες αποτελείται το σύστημα. Επίσης, παρουσιάζεται και το σχεσιακό σχήμα της βάσης δεδομένων που αναπτύχθηκε.

Στο 6^ο κεφάλαιο αναφέρονται σημαντικά σημεία της υλοποίησης, τα οποία αξίζουν περαιτέρω αναφοράς. Πρόκειται για θέματα απόδοσης και αξιοπιστίας που προέκυψαν και αντιμετωπίστηκαν με επιτυχία όσον αφορά τη βάση δεδομένων του συστήματος. Επιπλέον, γίνεται αναφορά σε αξιόλογα συστατικά του πηγαίου κώδικα της εφαρμογής, τα οποία παρουσιάζουν ιδιαίτερο ενδιαφέρον τόσο για τη συγκεκριμένη εφαρμογή όσο και για την ανάπτυξη παρόμοιων εργαλείων μιας και έχουν σχεδιαστεί με γνώμονα τη δυνατότητα επαναχρησιμοποίησής τους. Τέλος, παρουσιάζονται τα εργαλεία και οι πλατφόρμες ανάπτυξης του συστήματος.

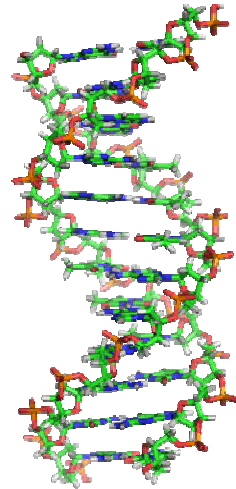
Στο 7^ο κεφάλαιο ακολουθεί η περιγραφή της διαδικασίας του αναλυτικού ελέγχου του συστήματος. Ο έλεγχος πραγματοποιήθηκε στα πλαίσια σύγχρονων τεχνικών, όπως είναι το unit testing και το system testing. Οι συγκεκριμένες τεχνικές αναλύονται ως προς τα χαρακτηριστικά τους και αξιολογούνται ως προς τις επιδόσεις και την αξιοπιστία τους.

Στο 8^ο κεφάλαιο καταλήγει η παρούσα διπλωματική εργασία συνοψίζοντας τους στόχους που επιτεύχθηκαν και παραθέτοντας τις δυνατότητες για μελλοντική επέκταση αυτής.

2

Θεωρητικό υπόβαθρο και σχετικές εργασίες

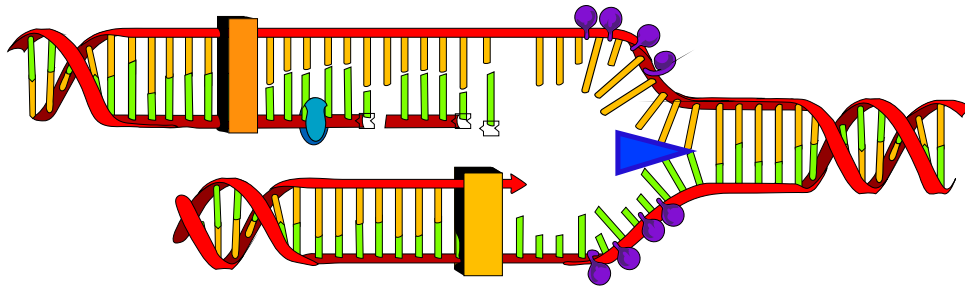
Η παρούσα διπλωματική εργασία εντάσσεται στον ευρύτερο χώρο συνεργασίας της πληροφορικής με τη βιολογία. Πρόκειται για ένα σύγχρονο σύστημα λογισμικού σχεδιασμένο ώστε να παρέχει υπηρεσίες υψηλού επιπέδου που παρουσιάζουν ενδιαφέρον και χρησιμότητα στην ερευνητική κοινότητα των βιολόγων. Επομένως, κρίνεται σκόπιμη η σύντομη αναφορά σε ορισμένα θέματα που άπτονται του χώρου της βιολογίας έτσι ώστε να είναι ευκολότερη η μετέπειτα κατανόηση του αντικειμένου της διπλωματικής εργασίας. Πρόκειται για μία συνοπτική αναφορά ζητημάτων ξεκινώντας από το λεγόμενο «κεντρικό δόγμα της μοριακής βιολογίας» και καταλήγοντας στην ανακάλυψη της ύπαρξης των microRNA.



2.1 Το κεντρικό δόγμα της μοριακής βιολογίας

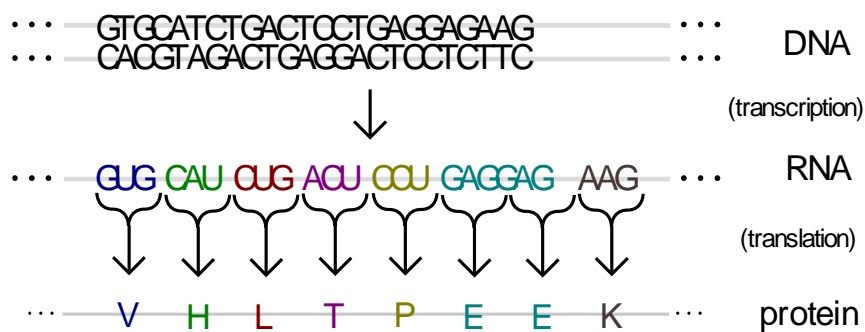
Το «κεντρικό δόγμα της μοριακής βιολογίας» είναι μία έκφραση που ειπώθηκε το 1958 από τον Francis Crick, ο οποίος το 1953 μαζί με τον James Watson είχε ανακαλύψει τη δομή της διπλής έλικας του μορίου DNA, το οποίο απεικονίζεται σχηματικά παραπάνω. Αυτό το δόγμα διαδραματίζει καθοριστικό ρόλο στην κατανόηση της ροής της γενετικής πληροφορίας στους ζωντανούς οργανισμούς. Η ροή της πληροφορίας μπορεί να πραγματοποιηθεί ανάμεσα σε τρεις μεγάλες οικογένειες βιολογικών μορίων: το DNA, το RNA και τις πρωτεΐνες. Επομέ-

νωσ, υπάρχουν 9 διαφορετικοί τρόποι άμεσης ροής της πληροφορίας από το ένα μόριο στο άλλο. Το δόγμα χωρίζει τους συγκεκριμένους τρόπους σε τρεις κατηγορίες, από τις οποίες η μία αντιστοιχεί στη φυσιολογική ροή της βιολογικής πληροφορίας: το DNA μπορεί να αντιγραφεί σε DNA, το DNA μπορεί να μεταγραφεί σε mRNA και οι πρωτεΐνες μπορούν να συντεθούν από το mRNA σε μια διαδικασία που ονομάζεται μετάφραση.



Εικόνα 1: Σχηματική αναπαράσταση της αντιγραφής του DNA σε DNA.

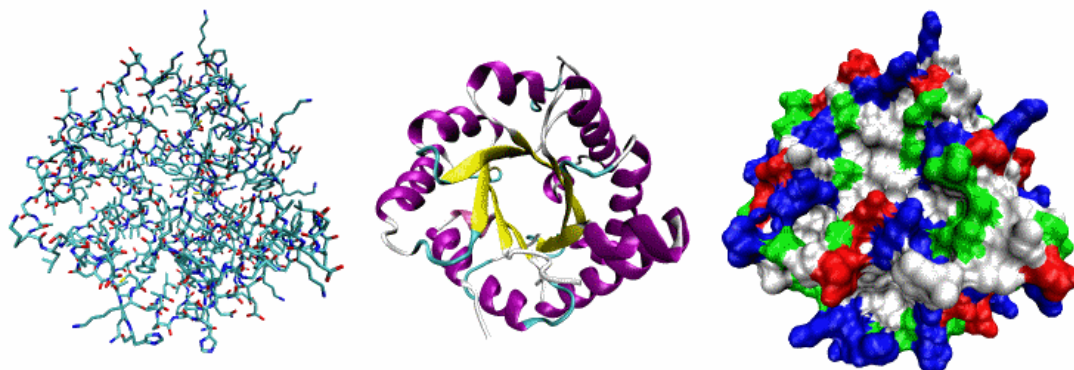
Το DNA είναι ένα βιολογικό μόριο, το οποίο βρίσκεται στον πυρήνα όλων των κυττάρων του οργανισμού και περιλαμβάνει τη γενετική πληροφορία από την οποία κατασκευάζονται όλες οι πρωτεΐνες του οργανισμού και κατ' επέκταση και ο ίδιος ο οργανισμός. Με άλλα λόγια το DNA κωδικοποιεί το γονιδίωμα του οργανισμού. Το μόριο του DNA αποτελείται από δύο έλικες που συστρέφονται και η κάθε έλικα αποτελείται από μία αλληλουχία βάσεων. Η αλληλουχία των βάσεων είναι αυτή που καθορίζει την κατασκευή των πρωτεϊνών.



Εικόνα 2: Σχηματικό διάγραμμα της ροής της γενετικής πληροφορίας από το μόριο DNA έως την πρωτεΐνη.

Το RNA είναι ένα βιολογικό μόριο, το οποίο διαδραματίζει σπουδαίο ρόλο στη λειτουργία των οργανισμών. Διαφέρει από το DNA μιας και παρουσιάζει ορισμένες δομικές διαφορές από αυτό, όπως το γεγονός ότι δεν έχει τη μορφή της διπλής έλικας αλλά συνήθως αποτελείται από μόνο μία αλληλουχία βάσεων. Ο ρόλος του στα κύτταρα των οργανισμών είναι να

μεταφέρει έξω από τον πυρήνα, όπου βρίσκεται το DNA του κυττάρου, τη βιολογική πληροφορία για την παραγωγή των πρωτεϊνών. Αυτή η μεταφορά πραγματοποιείται μέσω ενός ειδικού τύπου RNA, το οποίο καλείται αγγελιοφόρος RNA (mRNA).



Εικόνα 3: Σχηματική αναπαράσταση μίας πρωτεΐνης με τρεις διαφορετικούς τρόπους.

Οι πρωτεΐνες είναι βιολογικά μόρια, τα οποία αποτελούνται από αμινοξέα διατεταγμένα σε μία γραμμική αλυσίδα και διπλωμένα σε μία σφαιρική διάταξη. Η σειρά των αμινοξέων στην πρωτεΐνη προσδιορίζεται από την αλληλουχία βάσεων του γενετικού κώδικα και είναι αυτή που καθορίζει τη λειτουργικότητα του μορίου της πρωτεΐνης. Στους περισσότερους οργανισμούς, το γονιδίωμα ορίζει 20 αμινοξέα με τα οποία κατασκευάζονται όλες οι απαραίτητες πρωτεΐνες. Οι πρωτεΐνες μπορούν να λειτουργήσουν είτε ως ένζυμα και να λάβουν μέρος στις διαδικασίες του μεταβολισμού του οργανισμού είτε μπορούν να έχουν δομικές και μηχανικές λειτουργίες, όπως π.χ. στους μυς του οργανισμού. Κάθε συγκεκριμένη πρωτεΐνη παράγεται με την πληροφορία που περιέχεται σε ένα συγκεκριμένο τμήμα του γονιδιώματος. Τα τμήματα του γονιδιώματος που περιγράφουν την κατασκευή των πρωτεϊνών ονομάζονται γονίδια.

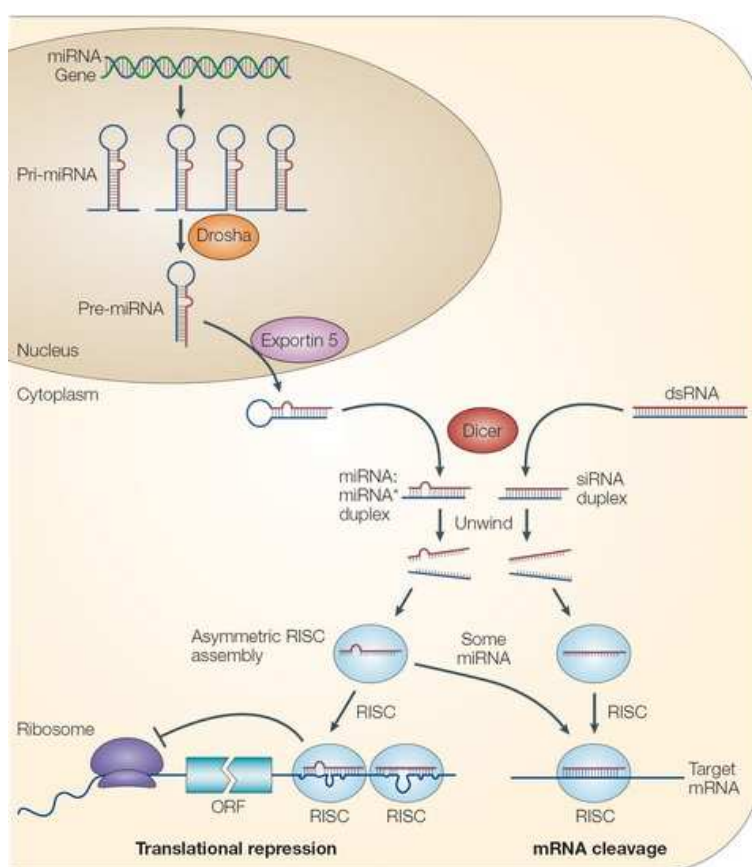
Για την παραγωγή πρωτεϊνών ακολουθείται η εξής διαδικασία:

- Διαβάζεται το κατάλληλο γονίδιο από το DNA που βρίσκεται στον πυρήνα του κυττάρου και μεταγράφεται στο μόριο mRNA.
- Το μόριο mRNA μετακινείται έξω από τον πυρήνα και πηγαίνει στα ριβοσώματα, τα οποία είναι τα «εργοστάσια» παραγωγής πρωτεϊνών του κυττάρου.

Η μετακίνηση της γενετικής πληροφορίας με το συγκεκριμένο τρόπο είναι απαραίτητη μιας και η κατασκευή των πρωτεϊνών δε μπορεί να λάβει χώρα στον πυρήνα των κυττάρων. Επομένως, η διαδικασία της μεταγραφής προστατεύει το μόριο του DNA, το οποίο μένει πάντα προφυλαγμένο στο εσωτερικό του πυρήνα του κυττάρου ενώ ταυτόχρονα δίνει λύση στην ανάγκη σύνθεσης των πρωτεϊνών.

2.2 Η ανακάλυψη των *microRNA*

Όταν ολοκληρώθηκε το πρόγραμμα χαρτογράφησης του ανθρώπινου γονιδιώματος έγινε γνωστό ότι μόνο το 1% με 2% του γονιδιώματος χρησιμοποιείται για την κωδικοποίηση πρωτεϊνών. Μία προσέγγιση στο ζήτημα του «άχρηστου γονιδιώματος», όπως ονομάστηκε, ήταν η πεποίθηση ότι υπάρχει για να προστατεύει το DNA του οργανισμού από μεταλλάξεις. Σταδιακά, όμως έγινε αντιληπτό πως και το υπόλοιπο γονιδίωμα μεταγράφεται σε RNA του οποίου, όμως, η λειτουργία δεν ήταν ακόμη γνωστή. Το συγκεκριμένο RNA ονομάστηκε ncRNA (non-coding RNA), ακριβώς διότι δεν προοριζόταν για την κωδικοποίηση πρωτεϊνών.



Εικόνα 4: Σχηματική αναπαράσταση της λειτουργίας των μορίων *microRNA*.

Μία περίπτωση ncRNA είναι το λεγόμενο *microRNA*. Πρόκειται για μικρά μόρια RNA, τα οποία παράγονται από συγκεκριμένα κομμάτια του γονιδιώματος. Η λειτουργία τους δεν είναι η παραγωγή πρωτεΐνης γι' αυτό και άργησαν να γίνουν αντιληπτά από τους βιολόγους. Ο τρόπος δράσης τους είναι να συνδέονται σε συγκεκριμένα σημεία συγκεκριμένων μορίων mRNA με αποτέλεσμα να καταστρέφουν αυτά τα μόρια και, επομένως, να εμποδίζουν την παραγωγή συγκεκριμένων πρωτεϊνών. Λειτουργούν ρυθμιστικά, επομένως, στη σύνθεση πρωτεϊνών από το κύτταρο. Αυτή η δράση έχει ιδιαίτερη σημασία μιας και η σύνθεση ή όχι

μίας πρωτεΐνης είναι ζωτικής σημασίας για το σύνολο του οργανισμού. Πολλές ασθένειες, για παράδειγμα, οφείλονται στην παραγωγή συγκεκριμένων πρωτεϊνών από τον οργανισμό.

2.3 Πρόβλεψη στόχων των *microRNA*

Τα σημεία στα οποία προσδένονται τα μόρια *microRNA* πάνω στους αγγελιοφόρους RNA (*mRNA*) ονομάζονται στόχοι (*targets*). Η γνώση των στόχων έχει ιδιαίτερη σημασία για τους ερευνητές – βιολόγους διότι με αυτόν τον τρόπο μπορούν να μάθουν για τη λειτουργία των μορίων *microRNA*. Για τον εντοπισμό των στόχων υπάρχουν δύο διαφορετικές τεχνικές. Η πρώτη είναι η πειραματική μέθοδος, η οποία δυστυχώς είναι δαπανηρή και χρονοβόρα επομένως δεν είναι εύκολη η πραγματοποίησή της. Η δεύτερη τεχνική αφορά τη χρήση εξειδικευμένων αλγορίθμων από το χώρο της πληροφορικής, οι οποίοι μπορούν μέσω μίας σειράς από υπολογιστικές διαδικασίες να προβλέψουν την ύπαρξη στόχων ενός *microRNA* σε ένα *mRNA*. Αυτή η διαδικασία καλείται *target prediction* και είναι ένα σημείο συνεργασίας μεταξύ βιολογίας και πληροφορικής που γνωρίζει ιδιαίτερη ανάπτυξη τα τελευταία χρόνια. Η ερευνητική ομάδα του Ερευνητικού Κέντρου Βιοϊατρικών Επιστημών «Αλέξανδρος Φλέμινγκ», η οποία συνεργάστηκε για την πραγματοποίηση της παρούσας διπλωματικής εργασίας έχει αναπτύξει έναν αλγόριθμο πρόβλεψης στόχων, του οποίου τα αποτελέσματα φιλοδοξεί να παρουσιάσει το σύστημα που αναπτύχθηκε στα πλαίσια της παρούσας διπλωματικής εργασίας.

2.4 Σχετικές εργασίες

Στο παρόν κεφάλαιο παρουσιάζονται συνοπτικά παρόμοιες δικτυακές εφαρμογές που προσφέρουν πρόσβαση σε βάσεις δεδομένων με πληροφορίες για βιολογικούς στόχους *microRNA*. Η συγκεκριμένη ερευνητική περιοχή έχει γνωρίσει ιδιαίτερη ανάπτυξη τα τελευταία χρόνια λόγω της διαφαινόμενης σπουδαιότητας των *microRNA* στην εξήγηση διάφορων βιολογικών διαδικασιών αλλά και στην αντιμετώπιση ασθενειών. Η έρευνα στο συγκεκριμένο πεδίο προωθείται με τη χρήση σύγχρονων αλγορίθμων, οι οποίοι έχουν την ικανότητα να προβλέπουν τους στόχους των *microRNA* πάνω στα μόρια των *mRNA*. Η εφαρμογή των αλγορίθμων, συνήθως, παράγει ένα μεγάλο όγκο πληροφορίας ο οποίος, πρωτογενώς, είναι δύσκολος στην χρήση και αξιοποίησή του. Επομένως, είναι απαραίτητες οι εφαρμογές εκείνες που έχουν τη δυνατότητα να παρουσιάζουν τις πληροφορίες στον ερευνητή – βιολόγο σε τέτοια μορφή ώστε να τον διευκολύνουν στην ερευνητική δραστηριότητά του.

2.4.1 DIANA-microT 3.0

DIANA LAB DNA Intelligent Analysis

HOME SOFTWARE DATABASES MEMBERS PUBLICATIONS HELP

DIANA - microT v3.0

NTUA CSLab

How to cite:

1/ M. Maragkakis; P. Alexiou; G. L. Papadopoulos; M. Reczko; T. Dalamagas; G. Giannopoulos; G. Goumas; E. Koukis; K. Kourtis; V. A. Simossis; P. Sethupathy; T. Vergoulis; N. Koziris; T. Sellis; P. Tsanakas; A. G. Hatzigeorgiou. **Accurate microRNA target prediction correlates with protein repression levels.** *BMC Bioinformatics* 2009, 10:295. [VIEW](#)

2/ M. Maragkakis; M. Reczko; V. A. Simossis; P. Alexiou; G. L. Papadopoulos; T. Dalamagas; G. Giannopoulos; G. Goumas; E. Koukis; K. Kourtis; T. Vergoulis; N. Koziris; T. Sellis; P. Tsanakas; A. G. Hatzigeorgiou. **DIANA-microT web server: elucidating microRNA functions through target prediction.** *Nucleic Acids Research* 2009, Jul 15; 37(Web Server issue):W273-6. [VIEW](#)

[Download all microT v3.0 results](#)

Search for targets of a specific miRNA

Insert a miRNA identifier (eg.hsa-let-7a)

[VIEW RESULTS](#)

[OR predict targets for user defined microRNA sequence](#)

E.S.R.C. Alexander Fleming - 34 Fleming Street, 15672, Athens, Greece

ALEXANDER FLEMING
Biomedical Sciences Research Center

Εικόνα 5: Ο δικτυακός τόπος της εφαρμογής DIANA-microT 3.0.

Ο δικτυακός τόπος <http://diana.cslab.ntua.gr/microT/> αποτελεί την προγενέστερη έκδοση της εφαρμογής που παρουσιάζεται στην παρούσα διπλωματική εργασία. Προσφέρει τη δυνατότητα αναζήτησης στόχων είτε με το όνομα του microRNA είτε με το όνομα του γονιδίου είτε με συνδυασμό των δύο τελευταίων. Τα αποτελέσματα παρουσιάζονται σε μορφή πίνακα όπου δίνεται η δυνατότητα στον ερευνητή να προβάλει, επιπλέον, τα σημεία πρόσδεσης για τον κάθε στόχο. Επιπλέον, δίνει τη δυνατότητα στον ερευνητή να προβλέψει τους στόχους για κάποιο microRNA, το οποίο μπορεί να εισάγει ο ίδιος στην περίπτωση που δεν υπάρχει ακόμη στην ενσωματωμένη βάση δεδομένων της εφαρμογής.

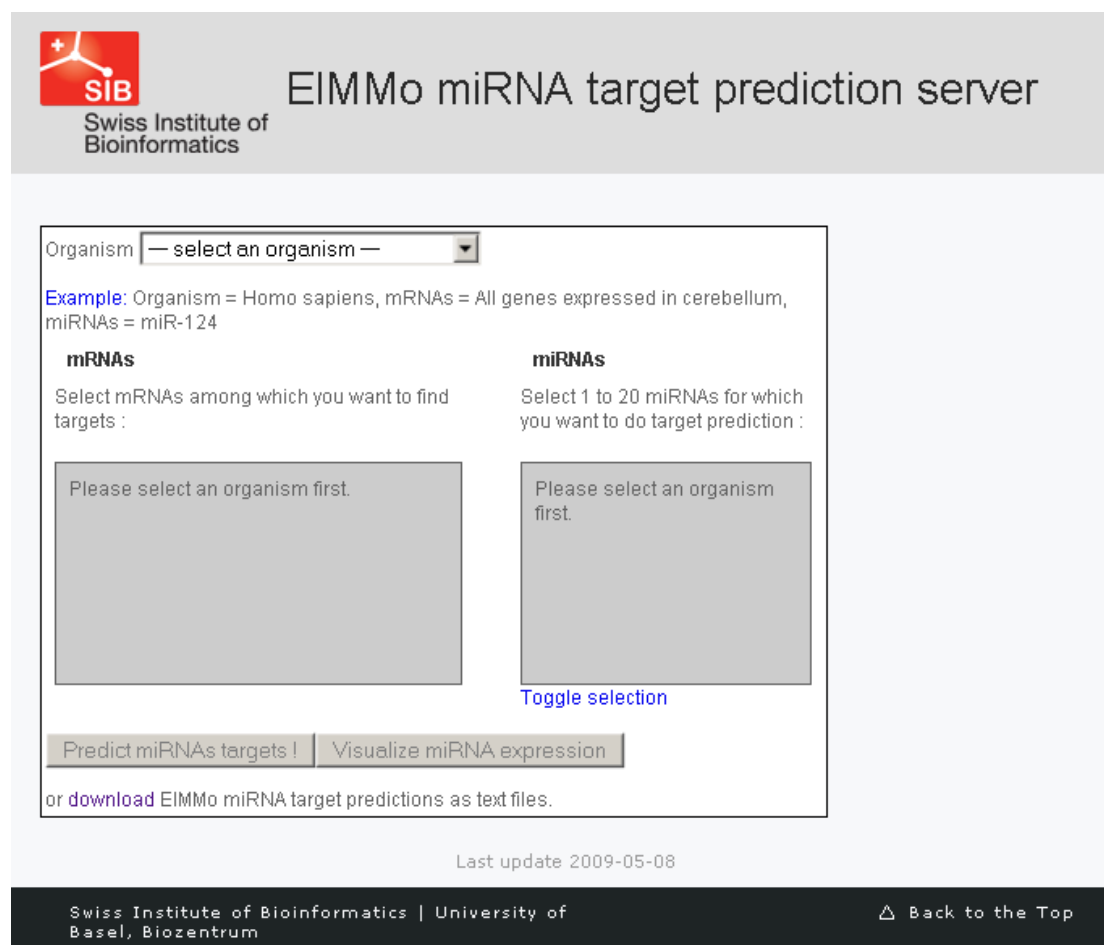
2.4.2 miRanda

Η συγκεκριμένη δικτυακή εφαρμογή βρίσκεται στη διεύθυνση: <http://www.microrna.org/> και παρέχει δυνατότητες αναζήτησης στόχων προσδιορίζοντας είτε το microRNA είτε το mRNA. Τα αποτελέσματα οργανώνονται με βάση το mRNA, όπου για κάθε μόριο στο οποίο βρέθηκε στόχος εμφανίζεται ένα σχήμα που αναπαριστά τα σημεία πρόσδεσης όλων των microRNA στο συγκεκριμένο μόριο.

2.4.3 *MicroCosm*

Η συγκεκριμένη δικτυακή εφαρμογή παρέχει δυνατότητες τόσο απλής αναζήτησης με το microRNA ή το γονίδιο όσο και σύνθετης αναζήτησης προσδιορίζοντας και τα δύο. Τα αποτελέσματα εμφανίζονται σε ένα πίνακα ταξινομημένα ανά γονίδιο και ο χρήστης έχει τη δυνατότητα να προβάλλει τα ξεχωριστά σημεία πρόσδεσης σε κάθε γονίδιο επιλέγοντας τον αντίστοιχο σύνδεσμο. Η εφαρμογή βρίσκεται στη διεύθυνση: <http://www.ebi.ac.uk/enright-srv/microcosm/cgi-bin/targets/v5/search.pl>.

2.4.4 *EIMMo*



The screenshot shows the EIMMo miRNA target prediction server interface. At the top left is the SIB logo (Swiss Institute of Bioinformatics). The main title is "EIMMo miRNA target prediction server". Below the title is a form with the following elements:

- An "Organism" dropdown menu with the text "— select an organism —".
- An "Example" text: "Example: Organism = Homo sapiens, mRNAs = All genes expressed in cerebellum, miRNAs = miR-124".
- Two columns for selection:
 - mRNAs**: "Select mRNAs among which you want to find targets :". Below it is a grey box with the text "Please select an organism first."
 - miRNAs**: "Select 1 to 20 miRNAs for which you want to do target prediction :". Below it is a grey box with the text "Please select an organism first."
- A "Toggle selection" link between the two columns.
- Two buttons: "Predict miRNAs targets !" and "Visualize miRNA expression".
- A link: "or download EIMMo miRNA target predictions as text files."

At the bottom of the form area, it says "Last update 2009-05-08". The footer contains the SIB logo and text: "Swiss Institute of Bioinformatics | University of Basel, Biozentrum" and a "Back to the Top" link.

Εικόνα 6: Ο δικτυακός τόπος της εφαρμογής EIMMo.

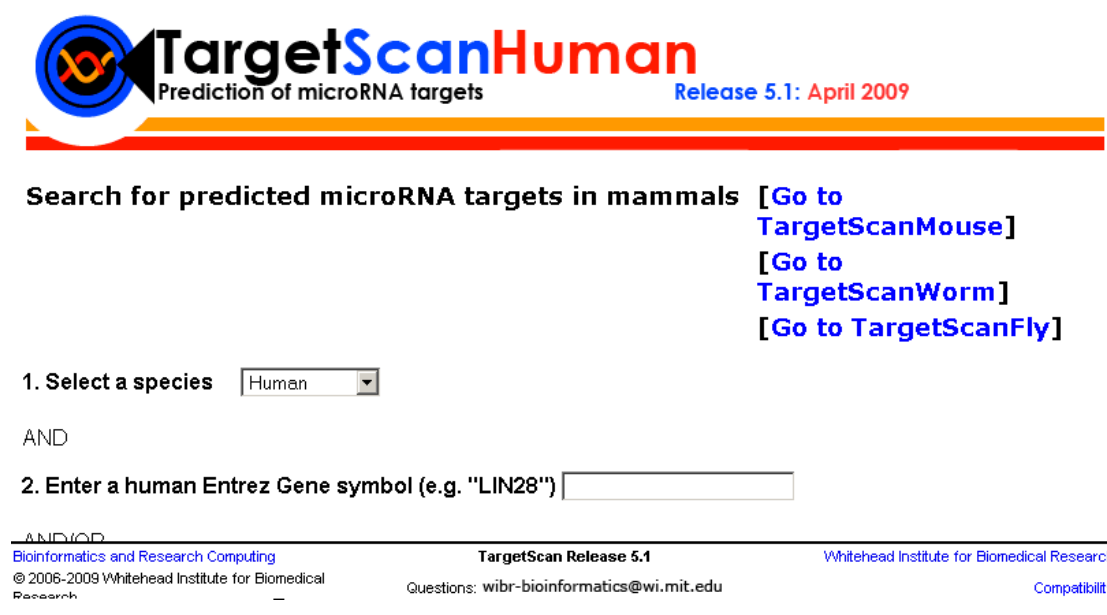
Η συγκεκριμένη δικτυακή εφαρμογή βρίσκεται στη διεύθυνση: <http://www.mirz.unibas.ch/> και παρέχει δυνατότητες αναζήτησης βιολογικών στόχων προσδιορίζοντας είτε το microRNA είτε το mRNA είτε το συνδυασμό τους. Ο προσδιορισμός των microRNA πραγματοποιείται επιλέγοντας από έναν κατάλογο επιλογών που δίνεται στο χρήστη, ενώ για την επιλογή των mRNA ο χρήστης έχει τη δυνατότητα να πληκτρολογήσει τα ονόματά τους. Τα αποτελέσματα περιλαμβάνουν τα mRNA στα οποία βρέθηκαν στόχοι. Ο χρήστης έχει τη δυνατότητα να

επιλέξει ένα από τα mRNA και να προβάλει σε ξεχωριστή σελίδα τα σημεία πρόσδεσης στο συγκεκριμένο μόριο.

2.4.5 Pictar

Η συγκεκριμένη δικτυακή εφαρμογή βρίσκεται στη διεύθυνση: <http://pictar.mdc-berlin.de/> και παρέχει δυνατότητες αναζήτησης στόχων με βάση είτε το microRNA (το οποίο μπορεί ο χρήστης να επιλέξει από ένα κατάλογο που προβάλλεται) είτε το γονίδιο. Η προβολή των αποτελεσμάτων πραγματοποιείται σε ένα πίνακα όπου ο χρήστης έχει τη δυνατότητα να προβάλει τα σημεία πρόσδεσης ενός δεδομένου βιολογικού στόχου επιλέγοντας τον αντίστοιχο σύνδεσμο.

2.4.6 Targetscan 5.0



Search for predicted microRNA targets in mammals [Go to TargetScanMouse]
[Go to TargetScanWorm]
[Go to TargetScanFly]

1. Select a species

AND

2. Enter a human Entrez Gene symbol (e.g. "LIN28")

AND/OR

Bioinformatics and Research Computing TargetScan Release 5.1 Whitehead Institute for Biomedical Research
© 2006-2009 Whitehead Institute for Biomedical Research Questions: wibr-bioinformatics@wi.mit.edu Compatibility

Εικόνα 7: Ο δικτυακός τόπος της εφαρμογής Targetscan 5.0.

Ο δικτυακός τόπος <http://www.targetscan.org/> παρέχει τη δυνατότητα αναζήτησης στόχων σε μία πληθώρα διαφορετικών ειδών. Οι δυνατότητες αναζήτησης περιλαμβάνουν αναζήτηση μέσω microRNA, αναζήτηση μέσω γονιδίου και συνδυαστική αναζήτηση. Επιπλέον, υπάρχει η δυνατότητα αναζήτησης μέσω διάφορων οικογενειών microRNA. Τα αποτελέσματα εμφανίζονται στη συνηθισμένη μορφή του πίνακα ανά γονίδιο και υπάρχει η δυνατότητα προβολής των σημείων πρόσδεσης σε ξεχωριστή σελίδα επιλέγοντας τον αντίστοιχο σύνδεσμο.

2.4.7 PITA

Η συγκεκριμένη προσέγγιση στο ερευνητικό πεδίο των microRNA δεν παρέχει κάποια εφαρμογή αντίστοιχη με τις προηγούμενες, η οποία να επιτρέπει την άμεση πρόσβαση στα δεδομένα. Αντιθέτως, δίνεται η δυνατότητα στον ερευνητή – βιολόγο να αποκτήσει το πρωτογε-

νές υλικό της μεθόδου αποθηκεύοντας στον προσωπικό υπολογιστή του το κατάλληλο αρχείο δεδομένων από την ακόλουθη διεύθυνση: <http://genie.weizmann.ac.il/pubs/mir07/>.

2.4.8 RNA22

Η συγκεκριμένη δικτυακή εφαρμογή ακολουθεί μία διαφορετική προσέγγιση στην παροχή της πληροφορίας για τους στόχους των microRNA. Οι περισσότερες εφαρμογές του κλάδου είναι εφοδιασμένες με βάση δεδομένων μέσω της οποίας προβάλλουν τα αποτελέσματά τους. Η συγκεκριμένη εφαρμογή πραγματοποιεί τον εντοπισμό των στόχων τη στιγμή της αναζήτησης, επομένως ο χρήστης έχει την ευχέρεια να αναζητήσει στόχους για microRNA, τα οποία ενδεχομένως δεν περιλαμβάνονται στις βάσεις δεδομένων των υπολοίπων εφαρμογών. Η δυσκολία στη χρήση της εφαρμογής είναι η ανάγκη εισαγωγής της πλήρους αλληλουχίας βάσεων για το microRNA και το mRNA, στα οποία θα πραγματοποιηθεί η αναζήτηση στόχων. Η εφαρμογή βρίσκεται στη διεύθυνση: http://cbcsrv.watson.ibm.com/rna22_targets.html.

3

Ανάλυση Απαιτήσεων Συστήματος

Στο παρόν κεφάλαιο παρουσιάζεται η γενικότερη αρχιτεκτονική του συστήματος και η ανάλυση απαιτήσεων αυτού βάσει της οποίας προχώρησε ο σχεδιασμός της εφαρμογής. Η αρχιτεκτονική της εφαρμογής συνοψίζει τα λογικά τμήματα τα οποία την αποτελούν και κατά κάποιο τρόπο μπορούν να λογιστούν ως ξεχωριστά υποσυστήματα της. Για την τεκμηρίωση και υποστήριξη της ανάλυσης απαιτήσεων έχουν χρησιμοποιηθεί όπου κρίνεται απαραίτητο διαγράμματα UML, όπως διαγράμματα περιπτώσεων χρήσης και διαγράμματα ευρωστίας. Με αυτόν τον τρόπο επιτυγχάνεται η μεγιστοποίηση της ακρίβειας της ανάλυσης απαιτήσεων ούτως ώστε ο σχεδιασμός του συστήματος που ακολουθεί να μπορεί να βασιστεί στο συγκεκριμένο υλικό.

3.1 Αρχιτεκτονική

Το σύστημα που αναπτύχθηκε στο πλαίσιο της παρούσας διπλωματικής εργασίας είναι μία προηγμένη δικτυακή εφαρμογή αναζήτησης βιολογικών στόχων microRNA. Παράλληλα, προσφέρει μία σειρά από ξεχωριστά χαρακτηριστικά που στόχο έχουν τη βελτίωση της εμπειρίας του χρήστη αυτής. Συγκεκριμένα, προσφέρονται δυνατότητες εξατομίκευσης μέσα από υπηρεσίες ειδικά σχεδιασμένες ώστε να προσαρμόζουν την εφαρμογή ανάλογα με τις προτιμήσεις του χρήστη.

Η ανάπτυξη μίας εφαρμογής που ακολουθεί τα σύγχρονα πρότυπα και τεχνολογίες συγγραφής λογισμικού απαιτεί το διαχωρισμό του έργου σε μικρότερες ενότητες ούτως ώστε να εί-

ναι ευκολότερη η ανάπτυξη, η τεκμηρίωση και ο έλεγχος του συστήματος. Ταυτόχρονα, με αυτόν τον τρόπο επιτυγχάνεται η μεγιστοποίηση της επαναχρησιμοποίησης του κώδικα σε παρόμοια έργα λογισμικού. Για αυτό το λόγο το πρώτο βήμα για την κατασκευή του συστήματος είναι ο εντοπισμός και η καταγραφή των απαραίτητων υποσυστημάτων. Κάθε υποσύστημα οφείλει να ικανοποιεί μία σειρά από κριτήρια, όπως είναι:

- Να είναι χρήσιμο, δηλαδή να μπορεί να φέρει σε πέρας μία χρήσιμη ενέργεια για το σύστημα.
- Να είναι αυτοτελές, δηλαδή να έχει τη δυνατότητα να διεκπεραιώσει το ρόλο του αυτόνομα χωρίς να εξαρτάται από την ύπαρξη άλλων υποσυστημάτων.
- Να μπορεί να αντικατασταθεί, δηλαδή η αφαίρεσή του και αντικατάστασή του από ένα άλλο υποσύστημα που επιτελεί το ίδιο καθήκον να μην απαιτεί τροποποιήσεις σε όλο το υπόλοιπο σύστημα.

Μία άλλη όψη των παραπάνω κριτηρίων, η οποία είναι ευρέως διαδεδομένη στο χώρο της σύγχρονης ανάπτυξης λογισμικού είναι τα κριτήρια υψηλής συνεκτικότητας και χαμηλής σύζευξης.

- Η συνεκτικότητα ενός υποσυστήματος έγκειται στην ύπαρξη συγκεκριμένων και αλληλένδετων ρόλων που καλείται να διαδραματίσει το συγκεκριμένο υποσύστημα. Είναι επιθυμητό κάθε υποσύστημα να χαρακτηρίζεται από υψηλή συνεκτικότητα, δηλαδή να έχει συγκεκριμένες αρμοδιότητες, οι οποίες μάλιστα να έχουν υψηλό βαθμό αλληλεξάρτησης. Με αυτόν τον τρόπο επιτυγχάνεται η μεγιστοποίηση της σταθερότητας, της αξιοπιστίας, της επαναχρησιμοποίησης, της κατανόησης και της δυνατότητας ελέγχου του υποσυστήματος.
- Η σύζευξη ενός υποσυστήματος χαρακτηρίζει το βαθμό εξάρτησής του από τα υπόλοιπα υποσυστήματα της εφαρμογής. Η επίτευξη χαμηλής σύζευξης δηλώνει την καλή σχεδίαση του υποσυστήματος και δίνει τη δυνατότητα εύκολης προσαρμογής του στις εκάστοτε ανάγκες της εφαρμογής. Επιπλέον, επιτρέπει την αυτόνομη επαναχρησιμοποίηση του υποσυστήματος σε άλλες εφαρμογές.

3.2 Μοντέλο Οντοτήτων Συσχετίσεων

Για την υποστήριξη των απαιτούμενων λειτουργιών αναζήτησης και παροχής εξατομικευμένων υπηρεσιών από το σύστημα σχεδιάστηκε κατάλληλη βάση δεδομένων. Για το σχεδιασμό της κατασκευάστηκε ως πρώτο βήμα το μοντέλο οντοτήτων συσχετίσεων. Το συγκεκριμένο μοντέλο αποτελεί μία αφηρημένη αναπαράσταση των δεδομένων που πρόκειται να αποθηκευτούν στη βάση δεδομένων και χρησιμοποιείται συνήθως κατά τη διάρκεια της ανάλυσης απαιτήσεων έτσι ώστε να εντοπιστούν και να καταγραφούν οι απαιτήσεις αναπαράστασης

των δεδομένων που διαχειρίζεται η αναπτυσσόμενη εφαρμογή λογισμικού. Στη συνέχεια περιγράφονται οι οντότητες και οι συσχετίσεις που περιλαμβάνονται στο μοντέλο ενώ παρουσιάζεται και το αντίστοιχο διάγραμμα.

Gene: Αναπαριστά ένα γονίδιο. Το κάθε γονίδιο χαρακτηρίζεται από την αλληλουχία βάσεων του, από την περιγραφή του και από την τοποθεσία του, η οποία είναι σύνθετη ιδιότητα και περιλαμβάνει το είδος, το χρωμόσωμα, την αλυσίδα του χρωμοσώματος, την αρχή και το τέλος της αλληλουχίας βάσεων. Ένα γονίδιο μπορεί να είναι είτε protein gene είτε microRNA gene. Επίσης, μπορεί να σχετίζεται με ένα σύνολο από CPG, ένα σύνολο από transcription factors και ένα σύνολο από hairpin microRNAs.

Protein gene: Αναπαριστά ένα γονίδιο που χρησιμοποιείται για την κωδικοποίηση μίας πρωτεΐνης. Χαρακτηρίζεται από ένα μοναδικό αναγνωριστικό που αντιστοιχεί στην εγγραφή της δικτυακής βάσης δεδομένων Ensembl, από ένα σύνολο ονομάτων και από σύνολο αναγνωριστικών για τη δικτυακή βάση δεδομένων Refseq. Ένα πρωτεϊνικό γονίδιο μπορεί να αντιστοιχίζεται σε ένα σύνολο από KEGG και σε ένα σύνολο από transcripts.

MicroRNA gene: Περιγράφει ένα γονίδιο που χρησιμοποιείται για την κωδικοποίηση ενός microRNA.

Transcript: Αναπαριστά ένα μετάγραφο ενός γονιδίου και χαρακτηρίζεται από ένα μοναδικό αναγνωριστικό που αντιστοιχεί στην εγγραφή της δικτυακής βάσης δεδομένων Ensembl.

Transcription factor: Αναπαριστά έναν παράγοντα μεταγραφής και χαρακτηρίζεται από το όνομά του και από το πρότυπό του.

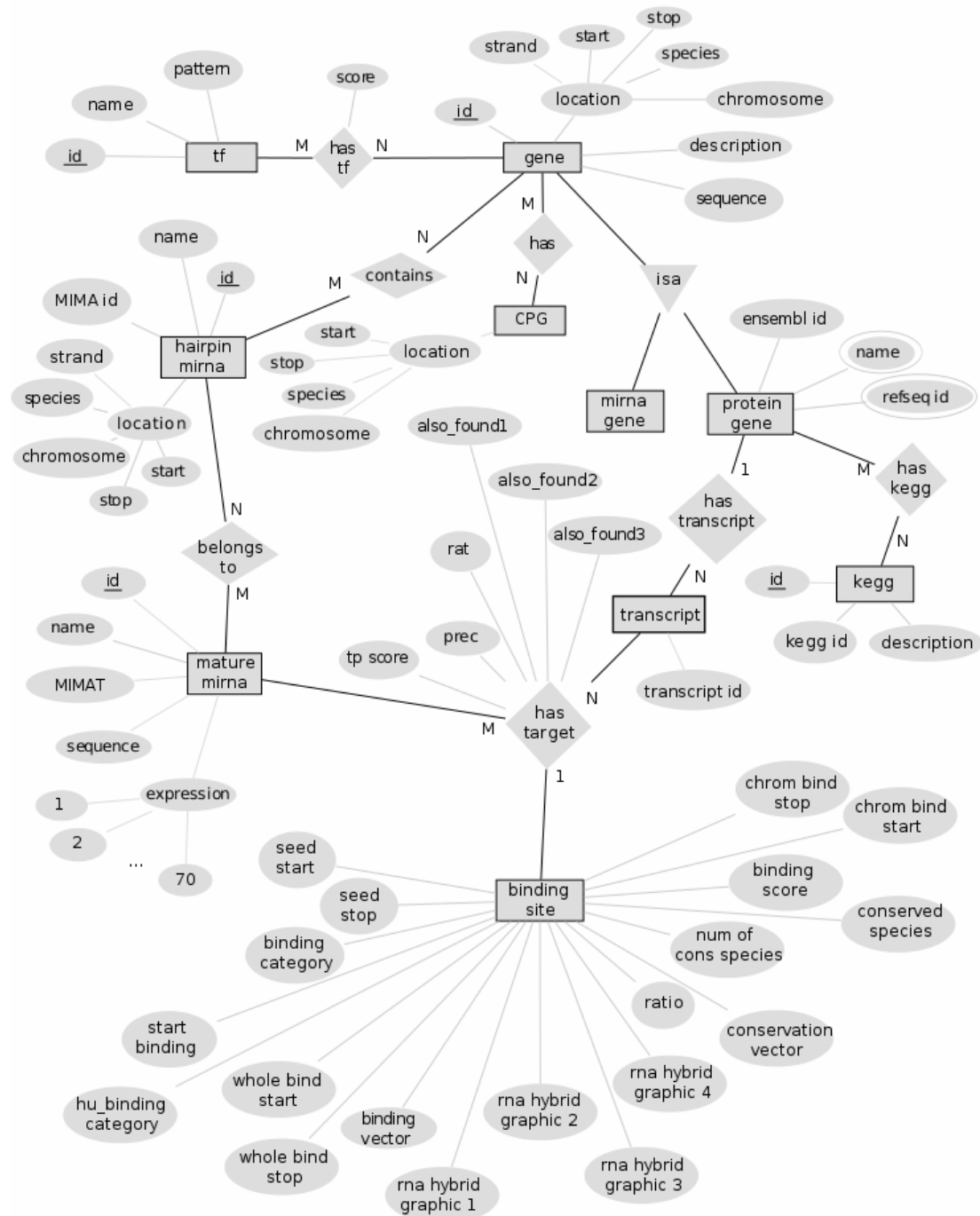
CPG: Περιγράφει μία περιοχή του DNA όπου υπάρχουν διαδοχικά μία βάση C (κυτοσίνη) και μία βάση G (γουανίνη) στην ίδια αλυσίδα βάσεων και χαρακτηρίζεται από την τοποθεσία του, η οποία είναι σύνθετη ιδιότητα και ορίζεται όπως και παραπάνω.

Hairpin microRNA: Αναπαριστά ένα microRNA το οποίο παρουσιάζει μία ιδιαίτερη δομή, δομή φουρκέτας. Προκύπτει από τη μεταγραφή ενός γονιδίου και χαρακτηρίζεται από το όνομά του, από ένα μοναδικό αναγνωριστικό που αντιστοιχεί στην εγγραφή της δικτυακής βάσης δεδομένων miRBase και από την τοποθεσία του, η οποία είναι σύνθετη ιδιότητα και ορίζεται όπως και παραπάνω.

Mature microRNA: Αναπαριστά ένα τελικό microRNA, το οποίο προκύπτει από ένα hairpin microRNA και χαρακτηρίζεται από το όνομά του, από ένα μοναδικό αναγνωριστικό για τη δικτυακή βάση δεδομένων miRBase, από την αλληλουχία των βάσεων του και από ένα σύνολο επιπέδων έκφρασης.

KEGG: Περιγράφει μία εγγραφή στη δικτυακή βάση δεδομένων KEGG Pathway. Χαρακτηρίζεται από ένα μοναδικό αναγνωριστικό και μία περιγραφή.

Binding site: Αναπαριστά μία θέση πρόσδεσης μεταξύ ενός mature microRNA και ενός transcript και χαρακτηρίζεται από ένα σύνολο ιδιοτήτων, όπως ο βαθμός πρόσδεσης, τα σημεία αρχής και τέλους της πρόσδεσης, κλπ. Σχετίζεται με ένα mature microRNA και ένα transcript μέσω της συσχέτισης has target, η οποία έχει ένα σύνολο από ιδιότητες που την χαρακτηρίζουν.

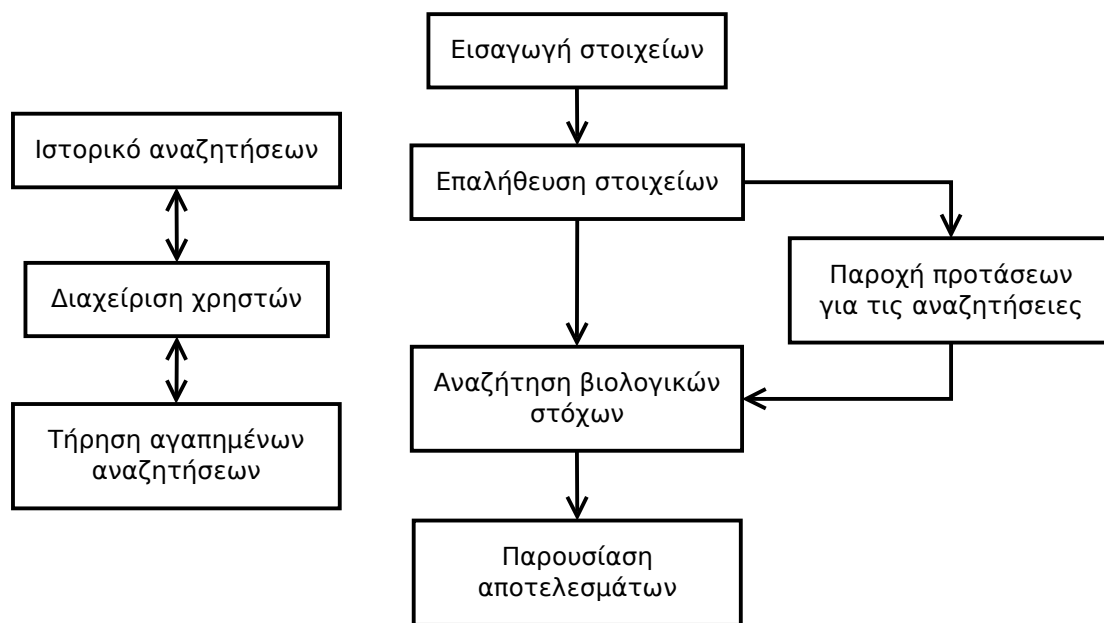


Εικόνα 8: Διάγραμμα οντοτήτων – συσχετίσεων για τη σχεσιακή βάση δεδομένων που αναπτύχθηκε για την υποστήριξη του συστήματος της παρούσας διπλωματικής εργασίας.

3.3 Περιγραφή Λειτουργιών

Η εφαρμογή που αναπτύχθηκε στο πλαίσιο της παρούσας διπλωματικής εργασίας αποτελεί ένα σύστημα αναζήτησης πληροφοριών βιολογικού ενδιαφέροντος. Παρέχει ένα σύνολο από εξειδικευμένες δυνατότητες αναζήτησης με πολλαπλά κριτήρια ούτως ώστε ο χρήστης να μπορεί να εντοπίσει με ευκολία την επιθυμητή πληροφορία. Επιπλέον, παρέχονται δυνατότητες εξατομίκευσης της εφαρμογής ώστε να αυξάνεται η χρησιμότητα αυτής από τον εκάστοτε χρήστη. Οι δυνατότητες αυτές υλοποιούνται στη βάση ενός προσωπικού λογαριασμού μέσω του οποίου ταυτοποιείται ο χρήστης της υπηρεσίας και αποκτά πρόσβαση στον κατάλογο των προσφάτων αναζητήσεων του και στη δυνατότητα να αποθηκεύει αναζητήσεις για μελλοντική χρήση.

Στη συνέχεια ακολουθεί ένα συνοπτικό διάγραμμα των επιμέρους υποσυστημάτων που απαρτίζουν την εφαρμογή καθώς και η λεπτομερής ανάλυσή τους. Οδηγός για το σωστό και τεκμηριωμένο διαχωρισμό της εφαρμογής είναι τα παραπάνω κριτήρια.



Εικόνα 9: Σχηματικό διάγραμμα των υποσυστημάτων που απαρτίζουν την εφαρμογή της παρούσας διπλωματικής εργασίας.

3.3.1 Εισαγωγή στοιχείων

Το συγκεκριμένο υποσύστημα αναλαμβάνει τη λήψη των παραμέτρων της αναζήτησης που επιθυμεί να πραγματοποιήσει ο χρήστης του συστήματος. Οι παράμετροι μπορεί να είναι μία ή περισσότερες από τις ακόλουθες:

- Το όνομα του microRNA, όπως hsa-let-7a. Όταν πραγματοποιηθεί αναζήτηση με βάση το συγκεκριμένο όνομα τότε επιστρέφονται βιολογικοί στόχοι που σχετίζονται με το συγκεκριμένο microRNA.
- Το όνομα του γονιδίου, όπως ENSG00000156273. Στη συγκεκριμένη περίπτωση η αναζήτηση επιστρέφει τους βιολογικούς στόχους που σχετίζονται με το συγκεκριμένο γονίδιο. Μπορεί να συνδυαστεί με το όνομα του microRNA ώστε να περιοριστούν ακόμη περισσότερο οι βιολογικοί στόχοι των αποτελεσμάτων.
- Η περιγραφή του γονιδίου, όπως “TRANSCRIPTION REGULATOR PROTEIN BACH1”. Στη συγκεκριμένη περίπτωση επιστρέφεται μία λίστα με όλα τα γονίδια που ταιριάζουν στην περιγραφή. Αν δοθεί το όνομα του γονιδίου, τότε αγνοείται η περιγραφή του γονιδίου διότι δεν έχει νόημα η συνδυαστική αναζήτηση σε αυτή την περίπτωση.
- Το KEGG pathway, όπως “Focal adhesion”. Αυτή η παράμετρος μπορεί να χρησιμοποιηθεί εφόσον έχει πραγματοποιηθεί μία αναζήτηση με βάση το όνομα του microRNA ή / και του γονιδίου και έχει ως λειτουργία των περιορισμό των εμφανιζόμενων αποτελεσμάτων έτσι ώστε να περιλαμβάνονται μόνο οι βιολογικοί στόχοι των οποίων το KEGG pathway είναι συναφές με την παράμετρο που δίνεται.
- Ο ελάχιστος βαθμός του βιολογικού στόχου, ο οποίος είναι ένας δεκαδικός αριθμός που κυμαίνεται από 0,20 έως 1,12. Αυτή η παράμετρος λειτουργεί εφόσον έχει ήδη πραγματοποιηθεί μία αναζήτηση και δρα περιορίζοντας τα αποτελέσματα σε εκείνα που επιτυγχάνουν βαθμό ίσο ή ανώτερο με το δοσμένο.

3.3.2 *Επαλήθευση στοιχείων*

Το συγκεκριμένο υποσύστημα αναλαμβάνει να επαληθεύσει την εγκυρότητα των παραμέτρων αναζήτησης. Σε περίπτωση που οι παράμετροι δεν είναι έγκυρες τότε ενημερώνεται ο χρήστης για την ύπαρξη εσφαλμένων παραμέτρων αναζήτησης και καλείται να προσπαθήσει εκ νέου.

3.3.3 *Παροχή προτάσεων για τους όρους της αναζήτησης*

Το συγκεκριμένο υποσύστημα ενεργοποιείται όταν οι όροι που δίνει ο χρήστης για την αναζήτηση περιέχουν ασάφειες, οι οποίες μπορούν να εμποδίσουν την ομαλή προβολή των αποτελεσμάτων προς το χρήστη. Συγκεκριμένα, η λειτουργία του υποσυστήματος συνοψίζεται στις εξής δυνατότητες:

- Αν ο χρήστης εισάγει ένα όνομα microRNA ή γονιδίου, το οποίο δεν μπορεί να εντοπιστεί στη βάση δεδομένων, τότε γίνεται προσεγγιστική αναζήτηση του ονόματος σε

ορισμένα πιθανά πεδία με τα οποία μπορεί να ταιριάζει και επιστρέφονται στο χρήστη οι εναλλακτικές προτάσεις.

- Αν ο χρήστης εισάγει την περιγραφή ενός γονιδίου, τότε πραγματοποιείται αναζήτηση στο αντίστοιχο πεδίο των γονιδίων και επιστρέφονται τα γονίδια εκείνα των οποίων η περιγραφή ταιριάζει με την αναζήτηση του χρήστη.

3.3.4 Αναζήτηση βιολογικών στόχων

Το συγκεκριμένο υποσύστημα αναλαμβάνει να εντοπίσει τους βιολογικούς στόχους που ανταποκρίνονται στις παραμέτρους που έχει ορίσει ο χρήστης. Παράλληλα, για κάθε βιολογικό στόχο που εντοπίζεται και περιλαμβάνεται στον κατάλογο των αποτελεσμάτων αναζητούνται οι περιοχές πρόσδεσης που αντιστοιχούν και προστίθενται στα αποτελέσματα. Η κατασκευή του καταλόγου των αποτελεσμάτων ακολουθεί μία δομημένη μορφή έχοντας στο πρώτο επίπεδο τους βιολογικούς στόχους και στο δεύτερο επίπεδο τις περιοχές πρόσδεσης. Με αυτό τον τρόπο διευκολύνεται η μετέπειτα παρουσίαση των αποτελεσμάτων στον browser του χρήστη.

3.3.5 Παρουσίαση αποτελεσμάτων

Το συγκεκριμένο υποσύστημα αναλαμβάνει την απεικόνιση των αποτελεσμάτων σε κατάλληλους πίνακες. Είναι σχεδιασμένο έτσι ώστε να προσαρμόζεται ανάλογα με το είδος των αναζητούμενων στοιχείων. Συγκεκριμένα, υπάρχουν τρεις διαφορετικές εκδοχές της σελίδας αποτελεσμάτων:

- Αποτελέσματα για αναζήτηση με βάση ένα microRNA
- Αποτελέσματα για αναζήτηση με βάση ένα γονίδιο
- Αποτελέσματα για συνδυαστική αναζήτηση με βάση ένα microRNA και ένα γονίδιο.

Σε κάθε ξεχωριστή περίπτωση παρουσιάζονται με διαφορετικό τρόπο τα αποτελέσματα ούτως ώστε να είναι ευκολότερη η πρόσβαση στην απαιτούμενη πληροφορία από το χρήστη. Οι πίνακες που χρησιμοποιούνται αναδεικνύουν τη δομή των αποτελεσμάτων παρουσιάζονται τους βιολογικούς στόχους στο πρώτο επίπεδο και εμφωλευμένα τις περιοχές πρόσδεσης. Η παρουσίαση πραγματοποιείται σε συνοπτική μορφή εξ αρχής και ο χρήστης έχει τη δυνατότητα να εισέλθει σε περισσότερες λεπτομέρειες επιλέγοντας από τα αντίστοιχα στοιχεία ελέγχου της σελίδας των αποτελεσμάτων.

3.3.6 Διαχείριση χρηστών

Το συγκεκριμένο υποσύστημα είναι υπεύθυνο για τη διαχείριση των λογαριασμών των χρηστών του συστήματος. Η ύπαρξή του δίνει τη δυνατότητα στην εφαρμογή να παρέχει υπηρε-

σίες εξατομίκευσης στο σύστημα, επιτρέποντας στον εκάστοτε χρήστη να χρησιμοποιήσει μία σειρά από διαθέσιμες λειτουργίες βάσει των προτιμήσεών του και των επιλογών του. Δίνονται οι εξής δυνατότητες στον χρήστη:

- Δημιουργία νέου λογαριασμού χρήστη στο σύστημα
- Είσοδο και ταυτοποίηση στο σύστημα

Για τον προσδιορισμό των χρηστών χρησιμοποιείται μοναδικό όνομα χρήστη, το οποίο επιλέγεται κατά τη διάρκεια της εγγραφής του χρήστη στην εφαρμογή ενώ για την πιστοποίηση των χρηστών χρησιμοποιείται προσωπικός κωδικός πρόσβασης.

3.3.7 Ιστορικό αναζητήσεων

Το συγκεκριμένο υποσύστημα διατηρεί μία συνεχή ιστορική καταγραφή όλων των αναζητήσεων που πραγματοποιούνται από τους χρήστες του συστήματος και η οποία αποθηκεύεται σε μόνιμο αποθηκευτικό χώρο. Ο εκάστοτε χρήστης έχει τη δυνατότητα να προβάλλει τις τελευταίες αναζητήσεις που έχει πραγματοποιήσει προς διευκόλυνσή του. Ο διαχειριστής της εφαρμογής έχει τη δυνατότητα να εξάγει τις πληροφορίες από το ιστορικό καταγραφής όλων των χρηστών, τις οποίες και έχει τη δυνατότητα να επεξεργαστεί περαιτέρω με σκοπό την άντληση χρήσιμων στατιστικών στοιχείων. Αξίζει να σημειωθεί πως η ιστορική καταγραφή δεν έχει χρονικούς ή χωρικούς περιορισμούς με αποτέλεσμα να μπορεί να καταγράψει τις αναζητήσεις που έχουν πραγματοποιηθεί στο σύστημα για απεριόριστο χρονικό διάστημα. Ο καθαρισμός του ιστορικού από το μόνιμο μέσο αποθήκευσης είναι ευθύνη του διαχειριστή του συστήματος.

3.3.8 Τήρηση αγαπημένων αναζητήσεων (bookmarks)

Το συγκεκριμένο υποσύστημα επιτρέπει σε κάθε εγγεγραμμένο χρήστη της εφαρμογής να αποθηκεύει για μελλοντική χρήση αναζητήσεις που πραγματοποιεί μέσω της εφαρμογής. Ο χρήστης έχει τη δυνατότητα αφού πραγματοποιήσει μία αναζήτηση να την επιλέξει για προσθήκη στον προσωπικό κατάλογο αγαπημένων αναζητήσεων που διατηρεί το σύστημα για κάθε χρήστη. Επιπλέον, έχει τη δυνατότητα να αντιστοιχίσει σε αυτή ένα περιγραφικό όνομα ώστε να καταστήσει αποδοτικότερη την αρχειοθέτηση των αγαπημένων αναζητήσεών του. Επίσης, έχει τη δυνατότητα να επιλέξει συγκεκριμένους βιολογικούς στόχους από την αναζήτηση σύμφωνα με τις προτιμήσεις του. Αυτή η δυνατότητα λειτουργεί ως ένα προσαρμοσμένο φίλτρο πάνω στην αναζήτηση με στόχο την απομόνωση συγκεκριμένων βιολογικών στόχων, οι οποίοι παρουσιάζουν το μεγαλύτερο ενδιαφέρον για το χρήστη και πιθανόν επιθυμεί να τους μελετήσει με λεπτομέρεια στο μέλλον. Για να ενεργοποιήσει τη συγκεκριμένη δυνατότητα, ο χρήστης οφείλει να εκτελέσει την αναζήτηση που επιθυμεί, στη συνέχεια να την προσθέσει στον κατάλογο με τις αγαπημένες αναζητήσεις του και, τέλος, να προχωρήσει

στην επιλογή των μεμονωμένων βιολογικών στόχων της αναζήτησης που επιθυμεί να κρατήσει σε αυτή. Στη συνέχεια, όταν προβάλλεται η αγαπημένη αναζήτηση τότε το σύστημα εμφανίζει μόνο τους επιλεγμένους βιολογικούς στόχους.

Η συγκεκριμένη δυνατότητα αποδεικνύει τη χρησιμότητά της όταν οι αναζητήσεις που πραγματοποιεί ο χρήστης επιστρέφουν ένα μεγάλο αριθμό από αποτελέσματα. Τότε είναι πιθανό ο χρήστης να αφιερώσει ένα σημαντικό χρονικό διάστημα στην επιλογή των βιολογικών στόχων που τον ενδιαφέρουν κατά κύριο λόγο. Η επιλογή των βιολογικών στόχων (π.χ. 30 ενδιαφέροντες στους 1000 των αποτελεσμάτων) είναι μία κοπιαστική εργασία, την οποία ο χρήστης επιθυμεί να μην εκτελεί σε κάθε αναζήτησή του. Προς αυτή την κατεύθυνση λειτουργεί η συγκεκριμένη δυνατότητα, επιτρέποντας στο χρήστη να αποθηκεύσει μόνιμα τους βιολογικούς στόχους που παρουσιάζουν το μέγιστο ενδιαφέρον. Επιπλέον, δίνεται η δυνατότητα μεταβολής των επιλεγμένων βιολογικών στόχων ανά πάσα στιγμή έτσι ώστε ο χρήστης να μπορεί να κρατάει ενήμερες τις αγαπημένες αναζητήσεις του σύμφωνα με τα τρέχοντα ενδιαφέροντά του.

3.4 Περιπτώσεις χρήσης

Μία απαραίτητη ενέργεια κατά τη διάρκεια της ανάλυσης απαιτήσεων του συστήματος είναι ο εντοπισμός και η λεπτομερής καταγραφή όλων των δυνατών ενεργειών που μπορούν να πραγματοποιήσουν οι χρήστες του συστήματος καθώς και η συμπεριφορά αυτού σε κάθε πιθανή ενέργεια. Για το σκοπό αυτό έχει αναπτυχθεί η μεθοδολογία των περιπτώσεων χρήσης.

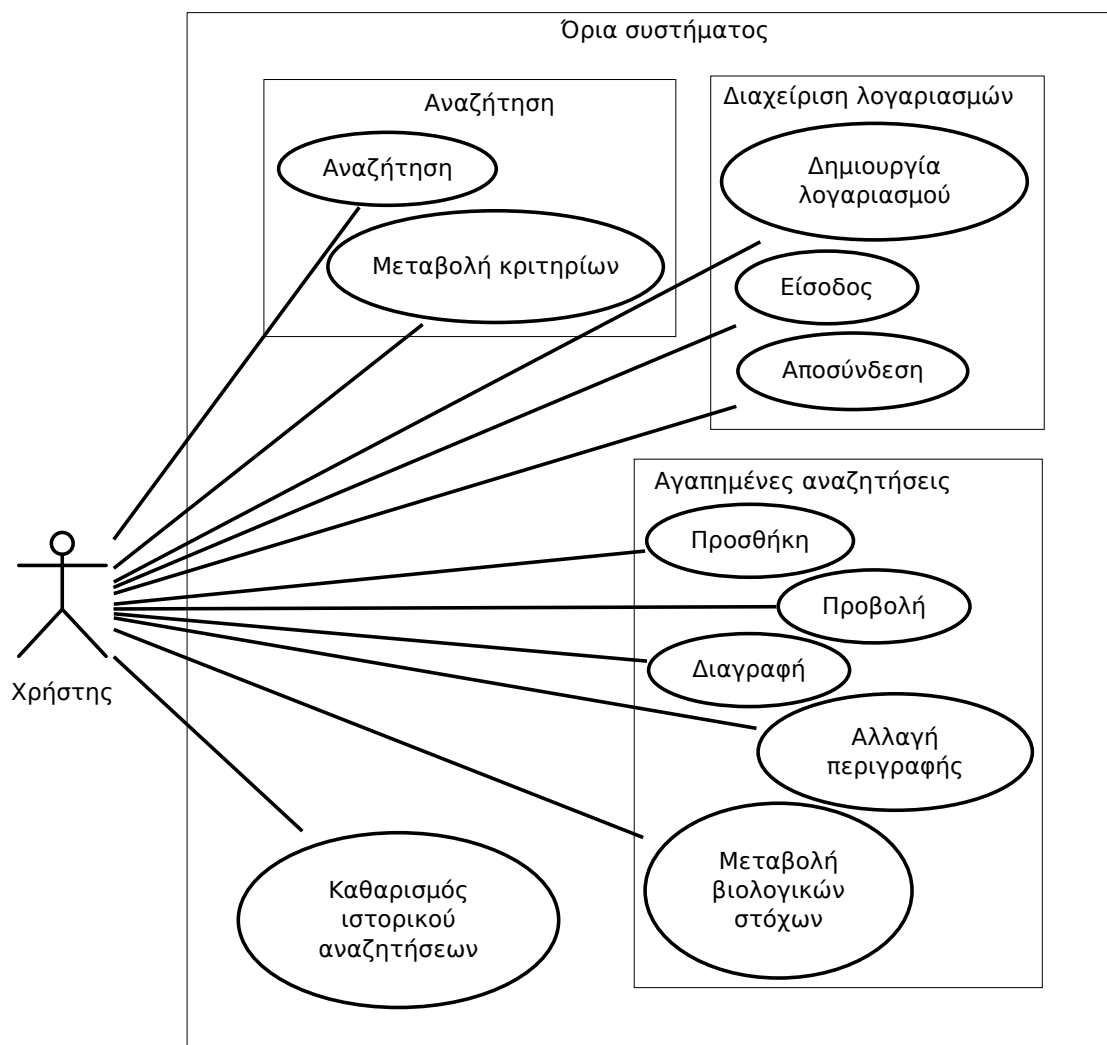
Πρώτο βήμα στην εφαρμογή της συγκεκριμένης μεθοδολογίας είναι η εύρεση των χειριστών (actors) του συστήματος. Σε πολύπλοκα συστήματα είναι πιθανό να υπάρχουν διάφοροι ρόλοι και χειριστές, όπως π.χ. απλός χρήστης και διαχειριστής. Στην εφαρμογή που αναπτύχθηκε στα πλαίσια της διπλωματικής εργασίας ωστόσο υπάρχει ένας χειριστής, ο οποίος χρησιμοποιεί την εφαρμογή μέσω του διαδικτύου.

Δεύτερο βήμα για την εφαρμογή της μεθοδολογίας των περιπτώσεων χρήσης είναι ο εντοπισμός των διακριτών δράσεων που δύναται να πραγματοποιήσει ο χειριστής αξιοποιώντας το σύστημα. Για κάθε δράση οφείλουν να καταγραφούν οι απαιτούμενες ενέργειες του χειριστή αλλά και η αναμενόμενη ανταπόκριση του συστήματος. Συνοπτικά, έχουμε τις εξής περιπτώσεις χρήσης:

- Αναζήτηση
- Είσοδος στο σύστημα
- Προσθήκη στις αγαπημένες αναζητήσεις
- Δημιουργία λογαριασμού
- Αποσύνδεση από το σύστημα

- Προβολή αγαπημένης αναζήτησης
- Διαγραφή αγαπημένης αναζήτησης
- Καθαρισμός ιστορικού αναζητήσεων
- Μεταβολή κριτηρίων αναζήτησης
- Αλλαγή περιγραφής αγαπημένης αναζήτησης
- Μεταβολή βιολογικών στόχων αγαπημένης αναζήτησης

Οι περιπτώσεις χρήσης του συστήματος απεικονίζονται στο αντίστοιχο διάγραμμα περιπτώσεων χρήσης. Αξίζει να σημειωθεί ότι στο συγκεκριμένο διάγραμμα περικλείεται σε παραλληλόγραμμο το σύστημα έτσι ώστε να είναι ευδιάκριτα τα όριά του. Επιπλέον, περικλείονται σε εσωτερικά παραλληλόγραμμα τα σημαντικότερα υποσυστήματα της εφαρμογής.



Εικόνα 10: Διάγραμμα περιπτώσεων χρήσης για το σύστημα που αναπτύχθηκε στο πλαίσιο της παρούσας διπλωματικής εργασίας.

3.4.1 Περίπτωση χρήσης «Αναζήτηση»

Τίτλος περίπτωσης χρήσης:

Αναζήτηση

Σύντομη περιγραφή:

Ο χρήστης εισάγει συγκεκριμένους όρους αναζήτησης και το σύστημα παρουσιάζει τους βιολογικούς στόχους που ικανοποιούν τα κριτήρια της αναζήτησης.

Ροή γεγονότων

Βασική ροή

1. Ο χρήστης εισάγει τα στοιχεία του microRNA ή / και του γονιδίου για τα οποία θέλει να αναζητήσει τους βιολογικούς στόχους.
2. Ο χρήστης υποβάλει τη φόρμα αναζήτησης επιλέγοντας το αντίστοιχο πλήκτρο.
3. Εμφανίζεται η σελίδα αποτελεσμάτων που περιλαμβάνει τους στόχους που βρέθηκαν σύμφωνα με την αναζήτηση του χρήστη.

Εναλλακτικές ροές

Εναλλακτική ροή 1

1. Τα στοιχεία που έχουν εισαχθεί είναι διαφορετικά, δηλαδή προσδιορίζουν περισσότερα από ένα microRNA ή / και γονίδια.
2. Εμφανίζεται φόρμα διασαφήνισης της αναζήτησης.
3. Ο χρήστης διευκρινίζει τους όρους της αναζήτησής του επιλέγοντας από τη λίστα των προτάσεων που του παρέχει το σύστημα.
4. Ο χρήστης υποβάλει εκ νέου την αναζήτηση επιλέγοντας το αντίστοιχο πλήκτρο.

Μη λειτουργικές απαιτήσεις

1. Ο χρόνος επεξεργασίας της αναζήτησης από το σύστημα οφείλει να είναι αποδεκτής διάρκειας ώστε να δίνεται η αίσθηση αμεσότητας στο χρήστη.
2. Όταν είναι απαραίτητη η εμφάνιση της φόρμας διασαφήνισης, οφείλει να γίνεται εμφανές ποιος όρος της αναζήτησης του χρήστη χρήζει περαιτέρω διασαφήνισης.
3. Οι προτάσεις διασαφήνισης που δίνονται στο χρήστη, όταν είναι απαραίτητο, οφείλουν να είναι ταξινομημένες με τέτοιο τρόπο ώστε η επιλογή του χρήστη να πραγματοποιηθεί ανάμεσα στις προτάσεις που βρίσκονται στην κορυφή της λίστας.
4. Στη σελίδα αποτελεσμάτων οφείλουν να εμφανίζονται με σαφήνεια και με λεπτομέρειες οι όροι της αναζήτησης του χρήστη.

Κατάσταση εισόδου:

Δεν υπάρχουν απαιτήσεις εισόδου για τη συγκεκριμένη περίπτωση χρήσης.

Κατάσταση εξόδου:

1. Αν ο χρήστης είναι συνδεδεμένος στο σύστημα με τον προσωπικό λογαριασμό του, τότε η αναζήτηση αποθηκεύεται στο προσωπικό ιστορικό αναζητήσεων του χρήστη.

3.4.2 Περίπτωση χρήσης «Είσοδος στο σύστημα»

Τίτλος περίπτωσης χρήσης:

Είσοδος στο σύστημα

Σύντομη περιγραφή:

Ο χρήστης εισάγει τα στοιχεία πρόσβασής του για να αποκτήσει πρόσβαση στις προσωπικές υπηρεσίες του συστήματος.

Ροή γεγονότων

Βασική ροή

1. Ο χρήστης επιλέγει τον σύνδεσμο σύνδεσης στο σύστημα που βρίσκεται στο δεξιό μέρος της σελίδας.
2. Εμφανίζεται η φόρμα εισαγωγής των στοιχείων του χρήστη, δηλαδή του ονόματος χρήστη και του κωδικού πρόσβασης.
3. Ο χρήστης εισάγει το όνομα χρήστη και τον κωδικό πρόσβασης στα αντίστοιχα πεδία.
4. Ο χρήστης υποβάλει τη φόρμα μέσω του αντίστοιχου πλήκτρου.
5. Γίνεται επαλήθευση των στοιχείων του χρήστη από το σύστημα.
6. Εμφανίζεται η κεντρική σελίδα της εφαρμογής και στο δεξιό μέρος φαίνονται οι προσωπικές ρυθμίσεις του χρήστη.

Εναλλακτικές ροές

Εναλλακτική ροή 1

1. Τα στοιχεία εισόδου του χρήστη δεν είναι έγκυρα.
2. Εμφανίζεται ξανά η φόρμα εισαγωγής των στοιχείων με κατάλληλο μήνυμα προς το χρήστη για την εκ νέου προσπάθειά του.

Μη λειτουργικές απαιτήσεις

1. Τα στοιχεία πρόσβασης των χρηστών του συστήματος αποθηκεύονται στη σχεσιακή βάση δεδομένων του συστήματος.

Κατάσταση εισόδου:

1. Ο χρήστης του συστήματος αναγνωρίζεται από το σύστημα ως χρήστης – επισκέ-

πτης, δηλαδή δεν έχει ήδη συνδεθεί στο σύστημα.

Κατάσταση εξόδου:

1. Ο χρήστης του συστήματος αναγνωρίζεται από το σύστημα ως συνδεδεμένος χρήστης με μοναδικό αναγνωριστικό.
2. Ο χρήστης έχει πρόσβαση στις προσωπικές υπηρεσίες του συστήματος, όπως είναι το ιστορικό αναζητήσεων και η τήρηση προσωπικών αγαπημένων αναζητήσεων.

3.4.3 Περίπτωση χρήσης «Προσθήκη στις αγαπημένες αναζητήσεις»

Τίτλος περίπτωσης χρήσης:

Προσθήκη στις αγαπημένες αναζητήσεις

Σύντομη περιγραφή:

Ο χρήστης επιλέγει να προσθέσει μία αναζήτηση στον κατάλογο με τις αγαπημένες αναζητήσεις του.

Ροή γεγονότων

Βασική ροή

1. Ο χρήστης επιλέγει το σύνδεσμο προσθήκης της τρέχουσας αναζήτησης στα αγαπημένα.
2. Η αναζήτηση αποθηκεύεται στις αγαπημένες αναζητήσεις του χρήστη.
3. Η αναζήτηση εμφανίζεται στο δεξί μέρος της σελίδας του συστήματος στο χώρο των προσωπικών υπηρεσιών του χρήστη.

Εναλλακτικές ροές

Δεν υπάρχουν εναλλακτικές ροές για τη συγκεκριμένη περίπτωση χρήσης.

Μη λειτουργικές απαιτήσεις

1. Εφόσον η τρέχουσα αναζήτηση έχει ήδη προστεθεί στις αγαπημένες αναζητήσεις του χρήστη, δεν εμφανίζεται σύνδεσμος προσθήκης αλλά στη θέση του εμφανίζεται κατάλληλο μήνυμα ενημέρωσης.

Κατάσταση εισόδου:

1. Ο χρήστης είναι συνδεδεμένος με τα προσωπικά στοιχεία πρόσβασής του στο σύστημα.
2. Ο χρήστης έχει εκτελέσει μία αναζήτηση στο σύστημα.
3. Ο χρήστης δεν έχει ήδη προσθέσει στα αγαπημένα του την τρέχουσα αναζήτηση.

Κατάσταση εξόδου:

1. Η επιλεγμένη αναζήτηση προστίθεται στον κατάλογο των αγαπημένων αναζητήσεων

του χρήστη και εμφανίζεται στον αντίστοιχο χώρο.

3.4.4 Περίπτωση χρήσης «Δημιουργία λογαριασμού»

Τίτλος περίπτωσης χρήσης:

Δημιουργία λογαριασμού

Σύντομη περιγραφή:

Ο χρήστης δημιουργεί ένα προσωπικό λογαριασμό πρόσβασης στο σύστημα.

Ροή γεγονότων

Βασική ροή

1. Ο χρήστης επιλέγει το σύνδεσμο σύνδεσης στο σύστημα που βρίσκεται στο δεξί μέρος της σελίδας.
2. Εμφανίζεται η φόρμα σύνδεσης στο σύστημα.
3. Ο χρήστης επιλέγει το σύνδεσμο δημιουργίας νέου λογαριασμού στο σύστημα.
4. Εμφανίζεται η φόρμα δημιουργίας νέου λογαριασμού στο σύστημα.
5. Ο χρήστης εισάγει το επιθυμητό όνομα χρήστη και εισάγει εις διπλούν τον επιθυμητό κωδικό πρόσβασης.
6. Ο χρήστης επιλέγει τη δημιουργία του νέου προσωπικού λογαριασμού μέσω του αντίστοιχου πλήκτρου.

Εναλλακτικές ροές

Εναλλακτική ροή 1

1. Το όνομα χρήστη που επέλεξε ο χρήστης είναι ήδη καταχωρημένο και χρησιμοποιείται.
2. Εμφανίζεται εκ νέου η φόρμα δημιουργίας νέου λογαριασμού με κατάλληλο μήνυμα προς τον χρήστη.

Εναλλακτική ροή 2

1. Ο χρήστης δεν έχει εισάγει τον ίδιο κωδικό στα δύο αντίστοιχα πεδία.
2. Εμφανίζεται εκ νέου η φόρμα δημιουργίας νέου λογαριασμού με κατάλληλο μήνυμα προς τον χρήστη.

Εναλλακτική ροή 3

1. Ο χρήστης έχει αφήσει κενό κάποιο από τα πεδία της φόρμας δημιουργίας νέου λογαριασμού.
2. Εμφανίζεται εκ νέου η φόρμα δημιουργίας νέου λογαριασμού με κατάλληλο μήνυμα

προς τον χρήστη.

Μη λειτουργικές απαιτήσεις

Δεν υπάρχουν μη λειτουργικές απαιτήσεις για τη συγκεκριμένη περίπτωση χρήσης.

Κατάσταση εισόδου:

1. Ο χρήστης του συστήματος αναγνωρίζεται από το σύστημα ως χρήστης – επισκέπτης, δηλαδή δεν έχει ήδη συνδεθεί στο σύστημα.

Κατάσταση εξόδου:

1. Δημιουργείται ένας νέος λογαριασμός χρήστη για το σύστημα και αποθηκεύονται τα στοιχεία πρόσβασής του στη σχεσιακή βάση δεδομένων.

3.4.5 Περίπτωση χρήσης «Αποσύνδεση από το σύστημα»

Τίτλος περίπτωσης χρήσης:

Αποσύνδεση από το σύστημα

Σύντομη περιγραφή:

Ο χρήστης αποσυνδέεται από το σύστημα και λειτουργεί πλέον ως χρήστης – επισκέπτης.

Ροή γεγονότων

Βασική ροή

1. Ο χρήστης επιλέγει το σύνδεσμο αποσύνδεσης από το σύστημα που βρίσκεται στο δεξί μέρος της σελίδας.
2. Το σύστημα αποσυνδέει το χρήστη.
3. Εμφανίζεται η κεντρική σελίδα της εφαρμογής και στο δεξί μέρος φαίνεται προτροπή για εκ νέου σύνδεση του χρήστη.

Εναλλακτικές ροές

Δεν υπάρχουν εναλλακτικές ροές για τη συγκεκριμένη περίπτωση χρήσης.

Μη λειτουργικές απαιτήσεις

Δεν υπάρχουν μη λειτουργικές απαιτήσεις για τη συγκεκριμένη περίπτωση χρήσης.

Κατάσταση εισόδου:

1. Ο χρήστης του συστήματος αναγνωρίζεται από το σύστημα ως συνδεδεμένος χρήστης με μοναδικό αναγνωριστικό.

Κατάσταση εξόδου:

1. Ο χρήστης του συστήματος αναγνωρίζεται από το σύστημα ως χρήστης – επισκέπτης, δηλαδή δεν έχει ήδη συνδεθεί στο σύστημα.

3.4.6 Περίπτωση χρήσης «Προβολή αγαπημένης αναζήτησης»

<p>Τίτλος περίπτωσης χρήσης: Προβολή αγαπημένης αναζήτησης</p> <p>Σύντομη περιγραφή: Ο χρήστης επιλέγει για προβολή μία από τις αποθηκευμένες στον προσωπικό του χώρο αγαπημένες αναζητήσεις.</p> <p>Ροή γεγονότων</p> <p>Βασική ροή</p> <ol style="list-style-type: none">1. Ο χρήστης επιλέγει το σύνδεσμο της προβολής για την αγαπημένη αναζήτηση που επιθυμεί.2. Εμφανίζεται η σελίδα αποτελεσμάτων που προβάλλει τα προσαρμοσμένα στοιχεία της αγαπημένης αναζήτησης. <p>Εναλλακτικές ροές Δεν υπάρχουν εναλλακτικές ροές για τη συγκεκριμένη περίπτωση χρήσης.</p> <p>Μη λειτουργικές απαιτήσεις Δεν υπάρχουν μη λειτουργικές απαιτήσεις για τη συγκεκριμένη περίπτωση χρήσης.</p> <p>Κατάσταση εισόδου:</p> <ol style="list-style-type: none">1. Ο χρήστης του συστήματος αναγνωρίζεται από το σύστημα ως συνδεδεμένος χρήστης με μοναδικό αναγνωριστικό.2. Ο χρήστης έχει προσθέσει τουλάχιστον μία αναζήτηση στον προσωπικό κατάλογο των αγαπημένων αναζητήσεών του. <p>Κατάσταση εξόδου: Δεν υπάρχει.</p>
--

3.4.7 Περίπτωση χρήσης «Διαγραφή αγαπημένης αναζήτησης»

<p>Τίτλος περίπτωσης χρήσης: Διαγραφή αγαπημένης αναζήτησης</p> <p>Σύντομη περιγραφή: Ο χρήστης επιλέγει να διαγράψει μία από τις αποθηκευμένες στον προσωπικό του χώρο αγαπημένες αναζητήσεις.</p> <p>Ροή γεγονότων</p> <p>Βασική ροή</p>
--

1. Ο χρήστης επιλέγει το σύνδεσμο της διαγραφής για την αγαπημένη αναζήτηση που επιθυμεί.
2. Η αγαπημένη αναζήτηση διαγράφεται από τον προσωπικό χώρο του χρήστη.
3. Εμφανίζεται η σελίδα στην οποία βρισκόταν ο χρήστης προτού εκτελέσει τη συγκεκριμένη περίπτωση χρήσης.

Εναλλακτικές ροές

Δεν υπάρχουν εναλλακτικές ροές για τη συγκεκριμένη περίπτωση χρήσης.

Μη λειτουργικές απαιτήσεις

Δεν υπάρχουν μη λειτουργικές απαιτήσεις για τη συγκεκριμένη περίπτωση χρήσης.

Κατάσταση εισόδου:

1. Ο χρήστης του συστήματος αναγνωρίζεται από το σύστημα ως συνδεδεμένος χρήστης με μοναδικό αναγνωριστικό.
2. Ο χρήστης έχει προσθέσει τουλάχιστον μία αναζήτηση στον προσωπικό κατάλογο των αγαπημένων αναζητήσεων του.

Κατάσταση εξόδου:

1. Η επιλεγμένη αγαπημένη αναζήτηση διαγράφεται από τον κατάλογο των αγαπημένων αναζητήσεων του χρήστη.

3.4.8 Περίπτωση χρήσης «Καθαρισμός ιστορικού αναζητήσεων»**Τίτλος περίπτωσης χρήσης:**

Καθαρισμός ιστορικού αναζητήσεων

Σύντομη περιγραφή:

Ο χρήστης επιλέγει να απαλείψει το προσωπικό ιστορικό αναζητήσεων που εμφανίζεται στον προσωπικό χώρο του.

Ροή γεγονότων**Βασική ροή**

1. Ο χρήστης επιλέγει το σύνδεσμο του καθαρισμού ιστορικού αναζητήσεων.
2. Το ιστορικό των αναζητήσεων του συγκεκριμένου χρήστη απαλείφεται.
3. Εμφανίζεται η σελίδα στην οποία βρισκόταν ο χρήστης προτού εκτελέσει τη συγκεκριμένη περίπτωση χρήσης.

Εναλλακτικές ροές

Δεν υπάρχουν εναλλακτικές ροές για τη συγκεκριμένη περίπτωση χρήσης.

Μη λειτουργικές απαιτήσεις

Δεν υπάρχουν μη λειτουργικές απαιτήσεις για τη συγκεκριμένη περίπτωση χρήσης.

Κατάσταση εισόδου:

1. Ο χρήστης του συστήματος αναγνωρίζεται από το σύστημα ως συνδεδεμένος χρήστης με μοναδικό αναγνωριστικό.

Κατάσταση εξόδου:

1. Το ιστορικό αναζητήσεων του χρήστη είναι κενό.

3.4.9 Περίπτωση χρήσης «Μεταβολή κριτηρίων αναζήτησης»

Τίτλος περίπτωσης χρήσης:

Μεταβολή κριτηρίων αναζήτησης

Σύντομη περιγραφή:

Ο χρήστης επιθυμεί τη μεταβολή των κριτηρίων της αναζήτησης που προβάλλεται στη σελίδα των αποτελεσμάτων. Τα κριτήρια αναζήτησης είναι το microRNA, το γονίδιο, ο βαθμός του βιολογικού στόχου και το KEGG pathway.

Ροή γεγονότων

Βασική ροή

1. Ο χρήστης εισάγει ή τροποποιεί τα νέα στοιχεία της αναζήτησης. Για το σκοπό αυτό υπάρχουν δύο ειδικά διαμορφωμένοι χώροι άνω και κάτω από τον πίνακα των αποτελεσμάτων.
2. Ο χρήστης επιλέγει το πλήκτρο της ανανέωσης της αναζήτησης.
3. Το σύστημα επεξεργάζεται τα νέα στοιχεία και υπολογίζει τα ανανεωμένα αποτελέσματα.
4. Το σύστημα εμφανίζει την ανανεωμένη σελίδα των αποτελεσμάτων.

Εναλλακτικές ροές

Δεν υπάρχουν εναλλακτικές ροές για τη συγκεκριμένη περίπτωση χρήσης.

Μη λειτουργικές απαιτήσεις

Δεν υπάρχουν μη λειτουργικές απαιτήσεις για τη συγκεκριμένη περίπτωση χρήσης.

Κατάσταση εισόδου:

1. Ο χρήστης έχει πραγματοποιήσει μία αναζήτηση στο σύστημα και βρίσκεται στη σελίδα των αποτελεσμάτων.

Κατάσταση εξόδου:

Δεν υπάρχει

3.4.10 Περίπτωση χρήσης «Αλλαγή περιγραφής αγαπημένης αναζήτησης»

Τίτλος περίπτωσης χρήσης:

Αλλαγή περιγραφής αγαπημένης αναζήτησης

Σύντομη περιγραφή:

Ο χρήστης επιθυμεί την αλλαγή της περιγραφής μίας αποθηκευμένης αγαπημένης αναζήτησης.

Ροή γεγονότων

Βασική ροή

1. Ο χρήστης επιλέγει το σύνδεσμο τροποποίησης για την αγαπημένη αναζήτηση που επιθυμεί να επεξεργαστεί.
2. Το σύστημα εμφανίζει τη σελίδα επεξεργασίας της αγαπημένης αναζήτησης.
3. Ο χρήστης εισάγει την επιθυμητή περιγραφή στον αντίστοιχο χώρο.
4. Ο χρήστης επιλέγει το πλήκτρο ανανέωσης των στοιχείων της αγαπημένης αναζήτησης.
5. Το σύστημα αποθηκεύει τη νέα περιγραφή της αγαπημένης αναζήτησης.

Εναλλακτικές ροές

Δεν υπάρχουν εναλλακτικές ροές για τη συγκεκριμένη περίπτωση χρήσης.

Μη λειτουργικές απαιτήσεις

Δεν υπάρχουν μη λειτουργικές απαιτήσεις για τη συγκεκριμένη περίπτωση χρήσης.

Κατάσταση εισόδου:

1. Ο χρήστης του συστήματος αναγνωρίζεται από το σύστημα ως συνδεδεμένος χρήστης με μοναδικό αναγνωριστικό.
2. Ο χρήστης έχει προσθέσει τουλάχιστον μία αναζήτηση στον προσωπικό κατάλογο των αγαπημένων αναζητήσεών του.

Κατάσταση εξόδου:

1. Μεταβάλλεται η περιγραφή της αγαπημένης αναζήτησης σύμφωνα με τις ενέργειες του χρήστη.

3.4.11 Περίπτωση χρήσης «Μεταβολή βιολογικών στόχων αγαπημένης αναζήτησης»

Τίτλος περίπτωσης χρήσης:

Μεταβολή βιολογικών στόχων αγαπημένης αναζήτησης

Σύντομη περιγραφή:

Ο χρήστης επιθυμεί να προσθέσει ή να αφαιρέσει βιολογικούς στόχους σε μία αποθηκευμένη αναζήτηση.

Ροή γεγονότων

Βασική ροή

1. Ο χρήστης επιλέγει το σύνδεσμο τροποποίησης για την αγαπημένη αναζήτηση που επιθυμεί να επεξεργαστεί.
2. Το σύστημα εμφανίζει τη σελίδα επεξεργασίας της αγαπημένης αναζήτησης.
3. Ο χρήστης χρησιμοποιεί τα checkboxes που εμφανίζονται δίπλα στους βιολογικούς στόχους για να προσθέσει ή να αφαιρέσει τον εκάστοτε βιολογικό στόχο στην αγαπημένη αναζήτηση.
4. Ο χρήστης επιλέγει το πλήκτρο ανανέωσης των στοιχείων της αγαπημένης αναζήτησης.
5. Το σύστημα τροποποιεί τους αγαπημένους στόχους για την αναζήτηση, λαμβάνοντας υπόψη τη σελιδοποίηση των αποτελεσμάτων.

Εναλλακτικές ροές

Εναλλακτική ροή 1

1. Ο χρήστης αλλάζει σελίδα στα εμφανιζόμενα αποτελέσματα χρησιμοποιώντας το αντίστοιχο χειριστήριο πλοήγησης.
2. Στη συνέχεια η περίπτωση χρήσης συνεχίζει στο βήμα 3 της βασικής ροής.

Μη λειτουργικές απαιτήσεις

Δεν υπάρχουν μη λειτουργικές απαιτήσεις για τη συγκεκριμένη περίπτωση χρήσης.

Κατάσταση εισόδου:

1. Ο χρήστης του συστήματος αναγνωρίζεται από το σύστημα ως συνδεδεμένος χρήστης με μοναδικό αναγνωριστικό.
2. Ο χρήστης έχει προσθέσει τουλάχιστον μία αναζήτηση στον προσωπικό κατάλογο των αγαπημένων αναζητήσεών του.

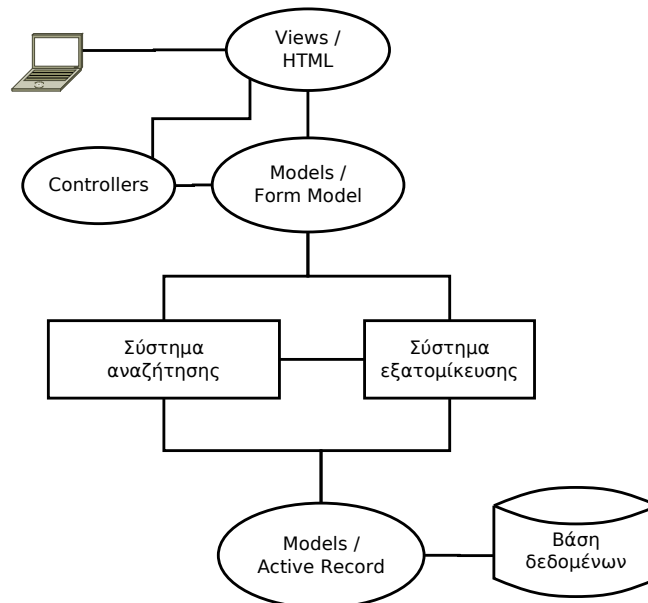
Κατάσταση εξόδου:

1. Μεταβάλλονται οι αγαπημένοι βιολογικοί στόχοι που περιλαμβάνονται στην επιλεγμένη αναζήτηση σύμφωνα με τις ενέργειες του χρήστη.

4

Σχεδίαση Συστήματος

Στο παρόν κεφάλαιο παρουσιάζεται και αναλύεται η σχεδίαση του συστήματος. Η σχεδίαση πραγματοποιήθηκε στη βάση της ανάλυσης απαιτήσεων αλλά και της γενικότερης αρχιτεκτονικής του συστήματος όπως ορίστηκαν στο προηγούμενο κεφάλαιο. Παρακάτω, παρατίθεται ένα σχηματικό διάγραμμα του συστήματος και των τμημάτων που το αποτελούν.



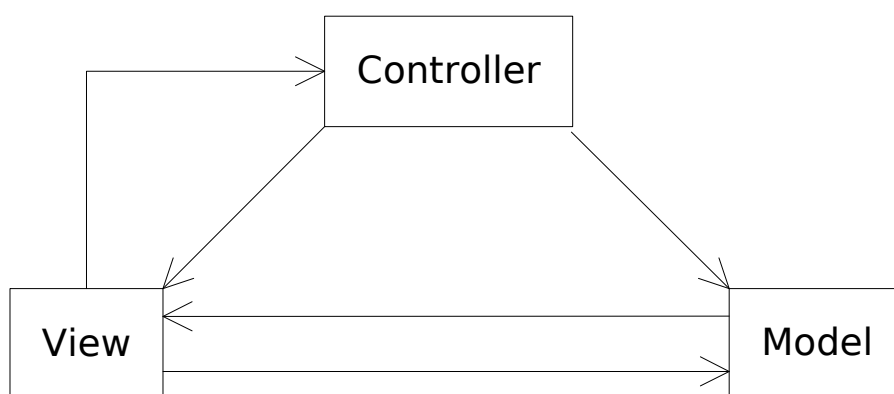
Εικόνα 11: Το block διάγραμμα του συστήματος που αναπτύχθηκε στο πλαίσιο της παρούσας διπλωματικής εργασίας. Διακρίνονται τα υποσυστήματα που το απαρτίζουν.

4.1 Αρχιτεκτονική

Η αρχιτεκτονική του συστήματος ακολουθεί το αρχιτεκτονικό πρότυπο Model – View – Controller. Το συγκεκριμένο αρχιτεκτονικό πρότυπο επιτρέπει το διαχωρισμό της λογικής από την παρουσίαση στο υπό ανάπτυξη σύστημα με αποτέλεσμα να προσφέρει μεγαλύτερη ευελιξία στην ανάπτυξη και στη συντήρηση του συστήματος και ευκολότερο έλεγχο των επιμέρους υποσυστημάτων. Μία εφαρμογή που αναπτύσσεται ακολουθώντας τη συγκεκριμένη τεχνοτροπία απαρτίζεται από ξεχωριστές δομικές μονάδες που υλοποιούν τους τρεις διακριτούς ρόλους που ορίζονται από τη συγκεκριμένη αρχιτεκτονική. Τα μοντέλα είναι κομμάτια κώδικα που αναπαριστούν τα δεδομένα της εφαρμογής. Οι όψεις (views) είναι κομμάτια κώδικα που αναλαμβάνουν να παρουσιάσουν το αποτέλεσμα του συστήματος στο χρήστη. Οι ελεγκτές είναι τμήματα κώδικα που δέχονται και διεκπεραιώνουν τις εντολές του χρήστη.

Στη σύγχρονη ανάπτυξη δικτυακών εφαρμογών το αρχιτεκτονικό πρότυπο Model – View – Controller βρίσκει ιδιαίτερη εφαρμογή και αποδεικνύεται σταδιακά ως ιδανική λύση για την ανάπτυξη πολύπλοκων εφαρμογών μιας και το μοντέλο ανάπτυξης δικτυακών συστημάτων επιβάλλει εγγενώς το διαχωρισμό της παρουσίασης, η οποία γίνεται μέσω HTML, και της λογικής, η οποία πραγματοποιείται στο διακομιστή της εφαρμογής. Σε αυτό το πλαίσιο ταίριαζει απόλυτα η λογική της λειτουργίας του συγκεκριμένου αρχιτεκτονικού προτύπου.

Στο σχήμα που ακολουθεί παρουσιάζονται διαγραμματικά τα επιμέρους στοιχεία της αρχιτεκτονικής και αναπαρίστανται οι συσχετίσεις μεταξύ αυτών. Επιγραμματικά, αξίζει να αναφερθεί ότι ο ελεγκτής έχει πρόσβαση τόσο στις όψεις, τις οποίες δίνει προς απεικόνιση, όσο και στα μοντέλα, από τα οποία αντλεί πληροφορίες ενώ και οι όψεις, με τη σειρά τους, έχουν τη δυνατότητα άντλησης πληροφοριών από τα μοντέλα. Στη συνέχεια ακολουθεί η ενδελεχής ανάλυση του αρχιτεκτονικού προτύπου Model – View – Controller στα επιμέρους τμήματα που το αποτελούν.



Εικόνα 12: Σχηματικό διάγραμμα του αρχιτεκτονικού μοντέλου Model – View – Controller και των σχέσεων μεταξύ των τμημάτων του.

4.1.1 Model

Αποτελεί την αναπαράσταση των δεδομένων τα οποία χρησιμοποιεί και επεξεργάζεται η εφαρμογή. Τα μοντέλα προσθέτουν λογική στην πληροφορία των εκάστοτε δεδομένων, π.χ. υπολογίζοντας αν σήμερα είναι η μέρα γενεθλίων ενός χρήστη του συστήματος. Όταν ένα μοντέλο μεταβάλλει την κατάστασή του ενημερώνει τα views που σχετίζονται με αυτό ώστε να ανανεωθούν με τη νέα κατάσταση. Τα περισσότερα συστήματα χρησιμοποιούν ένα μηχανισμό μόνιμης αποθήκευσης, π.χ. μία σχεσιακή βάση δεδομένων, για την αποθήκευση των δεδομένων. Τα μοντέλα δεν είναι απαραίτητο να γνωρίζουν για το μέσο αποθήκευσης της πληροφορίας. Παρόλα αυτά είναι πολύ βολικό αν το εκάστοτε μοντέλο γνωρίζει πώς να αναγνώσει και να αποθηκεύσει τον εαυτό του στο μέσο μόνιμης αποθήκευσης της εφαρμογής. Μία υλοποίηση της συγκεκριμένης ιδέας επιτυγχάνεται με το πρότυπο σχεδίασης Active Record.

4.1.2 View

Προβάλλει ένα μοντέλο σε μορφή κατανοητή από το χρήστη και προσαρμοσμένη στο μέσο απεικόνισης. Συνήθως, παρέχονται διαφορετικά views για κάθε μοντέλο τα οποία χρησιμοποιούνται ανάλογα την εκάστοτε ανάγκη παρουσίασης της πληροφορίας του μοντέλου. Συνηθισμένο παράδειγμα είναι η ύπαρξη ενός view με συνοπτικές πληροφορίες για το μοντέλο και ενός δεύτερου view, το οποίο απεικονίζει αναλυτικά όλα τα στοιχεία του μοντέλου. Εκτός από τη διάκριση στον τρόπο με τον οποίο απεικονίζεται το μοντέλο, τα views διακρίνονται και με βάση το μέσο στο οποίο είναι σχεδιασμένα να προβάλλονται. Πολλές φορές ένα μεγάλο σύστημα παρέχει στους χρήστες του πολλαπλές δυνατότητες πρόσβασης και απεικόνισης της πληροφορίας, π.χ. μέσω του web browser, μέσω μίας ανεξάρτητης γραφικής εφαρμογής, μέσω μίας ανεξάρτητης εφαρμογής κειμένου ή μέσω εκτυπωμένης πληροφορίας. Σε κάθε διαφορετικό μέσο απεικόνισης αντιστοιχεί και ένα διαφορετικό view, το οποίο λαμβάνει υπόψη του τις ιδιαιτερότητες του μέσου ώστε να εμφανίσει με το βέλτιστο δυνατό τρόπο τα στοιχεία του μοντέλου, π.χ. η προβολή ενός γραφήματος στην κονσόλα του συστήματος ή η προβολή μίας έγχρωμης εικόνας σε μία ασπρόμαυρη εκτύπωση είναι περιπτώσεις που απαιτούν ιδιαίτερο χειρισμό από το αντίστοιχο view.

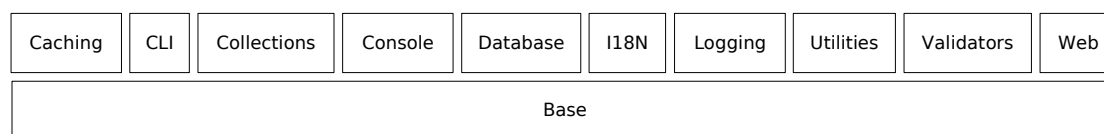
4.1.3 Controller

Λαμβάνει τα μηνύματα και τις εντολές εισόδου από το χρήστη και αποκρίνεται ενεργώντας πάνω στα μοντέλα ή / και εμφανίζοντας κάποιο συγκεκριμένο view. Ο ελεγκτής αποτελεί το συνδετικό κρίκο του συστήματος και είναι υπεύθυνος για τη διεκπεραίωση των εντολών του χρήστη της εφαρμογής. Η συνήθης λειτουργία του είναι η αναγνώριση των εντολών του χρή-

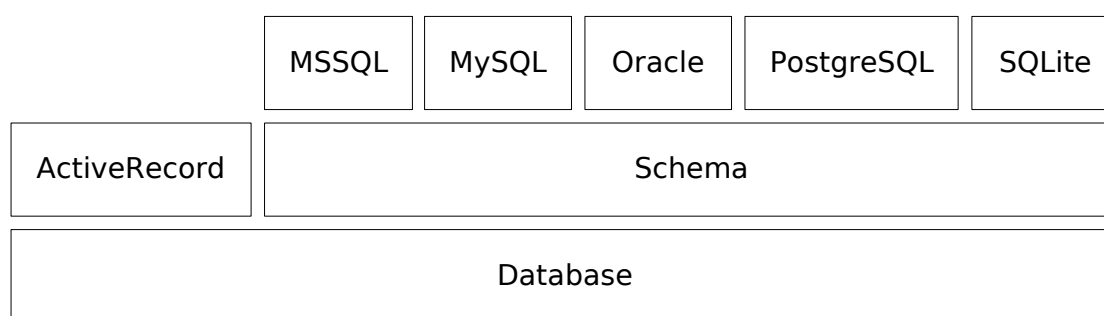
στη, η άντληση και επεξεργασία της πληροφορίας από τα μοντέλα βάσει των εντολών του χρήστη και η παρουσίαση της ζητούμενης πληροφορίας μέσω του κατάλληλου view.

4.2 Yii PHP framework

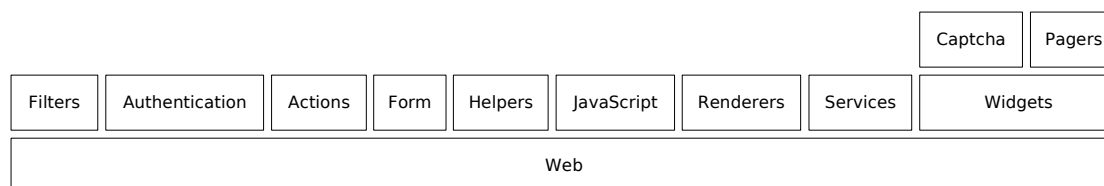
Για τη σχεδίαση της εφαρμογής με το αρχιτεκτονικό πρότυπο Model – View – Controller χρησιμοποιήθηκε το Yii PHP framework, το οποίο διευκολύνει την ανάπτυξη εφαρμογών σε PHP ακολουθώντας το συγκεκριμένο αρχιτεκτονικό πρότυπο. Η επιλογή του συγκεκριμένου framework βασίστηκε στην πληθώρα δυνατοτήτων, οι οποίες παρέχονται από αυτό. Οι ενότητες που ακολουθούν περιγράφουν τις δυνατότητες και τα χαρακτηριστικά του Yii PHP framework. Τα διαγράμματα που παρατίθενται στη συνέχεια παρουσιάζουν τα στρώματα στα οποία οργανώνεται ο κώδικας του Yii PHP framework. Στο πρώτο σχήμα φαίνεται η βάση, η οποία περιλαμβάνει όλες τις απαραίτητες λειτουργίες για τον εύρυθμο συντονισμό των υπολοίπων τμημάτων του framework. Πάνω από τη βάση βρίσκονται μία σειρά από υποσυστήματα, εκ των οποίων το υποσύστημα της διαχείρισης βάσεων δεδομένων και το υποσύστημα λειτουργιών σχετικών με το Web αναλύονται περαιτέρω σε ξεχωριστά διαγράμματα που ακολουθούν.



Εικόνα 13: Σχηματικό διάγραμμα της αρχιτεκτονικής δομής του Yii PHP framework και των συστημάτων που το αποτελούν.



Εικόνα 14: Σχηματικό διάγραμμα της αρχιτεκτονικής δομής του συστήματος διαχείρισης βάσεων δεδομένων του Yii PHP framework.

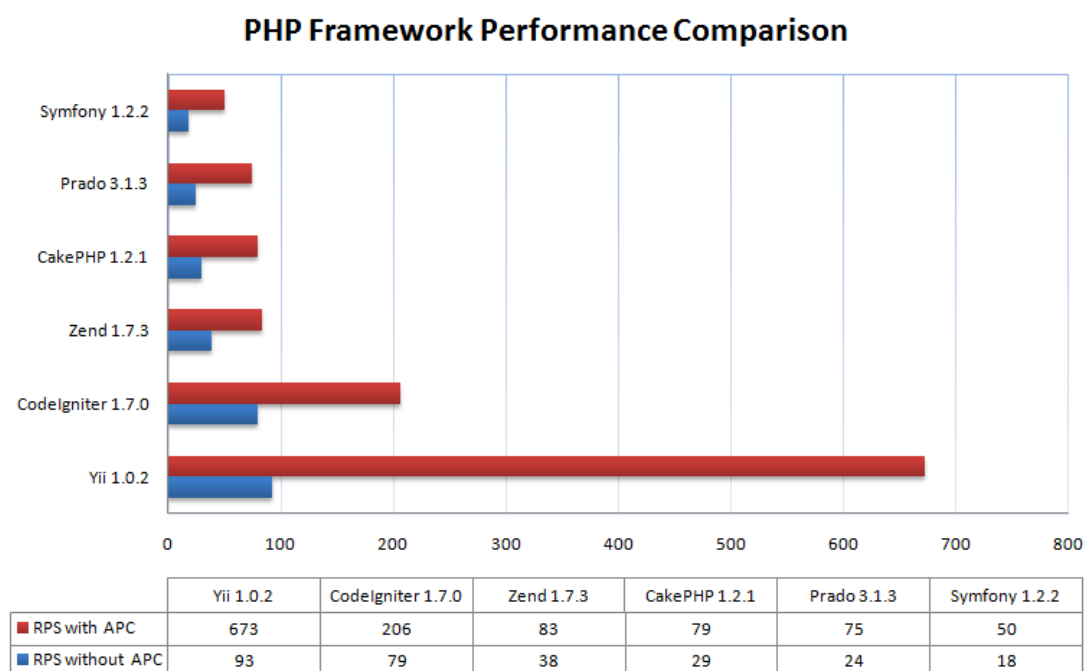


Εικόνα 15: Σχηματικό διάγραμμα της αρχιτεκτονικής δομής του συστήματος δημιουργίας δικτυακών εφαρμογών του Yii PHP framework.

4.2.1 Ταχύτητα

Το Yii PHP framework παρέχει αξιόλογη ταχύτητα στην εκτέλεση της εφαρμογής μέσω της τεχνικής της σκληρής φόρτωσης των συστατικών του. Αυτό σημαίνει ότι φορτώνει τα απαραίτητα συστατικά για την εκτέλεση της εφαρμογής καθώς αυτά απαιτούνται σε αντίθεση με τα υπόλοιπα PHP frameworks, τα οποία για λόγους ευκολίας φορτώνουν στη μνήμη όλα τα συστατικά πριν την εξυπηρέτηση του κάθε HTTP αιτήματος προσθέτοντας με αυτόν τον τρόπο μία αναπόφευκτη χρονική καθυστέρηση στην HTTP απάντηση.

Στο σχήμα που ακολουθεί παρουσιάζεται μία συγκριτική αναπαράσταση της ταχύτητας ορισμένων δημοφιλών PHP frameworks. Η μονάδα μέτρησης είναι ο αριθμός αιτημάτων ανά δευτερόλεπτο (RPS) που μπορεί να εξυπηρετήσει το κάθε framework. Επιπλέον, λαμβάνεται υπόψη και η συνεισφορά του Alternative PHP Cache (APC), το οποίο είναι ένα σύστημα βελτιστοποίησης του ενδιάμεσου κώδικα της PHP, το οποίο αποθηκεύει στη μνήμη του διακομιστή το μεταγλωττισμένο κώδικα και τα απαραίτητα δεδομένα έτσι ώστε να επιταχύνεται η εκτέλεση των εφαρμογών PHP. Για κάθε PHP framework η άνω ράβδος δηλώνει τον αριθμό αιτημάτων ανά δευτερόλεπτο με χρήση του APC και η κάτω ράβδος δηλώνει το RPS δίχως τη χρήση APC.



Εικόνα 16: Γραφική παράσταση των επιδόσεων σε αιτήματα ανά δευτερόλεπτο (RPS) έξι διαδομένων PHP frameworks.

4.2.2 Ποιότητα κώδικα

Το Yii PHP framework είναι γραμμένο εξολοκλήρου αξιοποιώντας τις αντικειμενοστρεφείς δυνατότητες της PHP με αποτέλεσμα να αποτελείται από ευανάγνωστο κώδικα, ο οποίος είναι εύκολος στη συντήρηση. Με αυτό τον τρόπο υπάρχει η εγγύηση ότι τυχόν προβλήματα που ανακαλύπτονται στο framework είναι εύκολο να επιλυθούν σε σύντομο χρονικό διάστημα. Η χρήση των αντικειμενοστραφών χαρακτηριστικών της PHP συνεπάγεται την ανάγκη χρήσης της έκδοσης 5 ή μεταγενέστερης αυτής, από την οποία και έπειτα υποστηρίζονται πλήρως τα συγκεκριμένα χαρακτηριστικά.

4.2.3 Data Access Objects και σχεδιαστικό πρότυπο Active Record

Η χρήση των Data Access Objects βασίζεται στην παροχή ενός κοινού interface για την πρόσβαση στους μηχανισμούς μίας βάσης δεδομένων. Η χρήση των συγκεκριμένων αντικειμένων διαχωρίζει την υλοποίηση του τρόπου επικοινωνίας με τη βάση από τη χρήση των δεδομένων που περιέχονται σε αυτή. Με αυτόν τον τρόπο, ο κώδικας των εφαρμογών που χρησιμοποιεί τη συγκεκριμένη τεχνική είναι εφαρμόσιμος σε οποιοδήποτε βάση δεδομένων υποστηρίζεται από τα Data Access Objects δίχως περαιτέρω αλλαγές.

Το Yii PHP framework παρέχει μία πλήρη μηχανή πρόσβασης σε σχεσιακές βάσεις δεδομένων ενώ ταυτόχρονα προσφέρει εύχρηστη πρόσβαση στα δεδομένα με χρήση αντικειμένων. Κατά αυτόν τον τρόπο ο προγραμματιστής δε χρειάζεται να γράψει κώδικα SQL και η εφαρμογή του είναι εκ των πραγμάτων ανεξάρτητη από τη σχεσιακή βάση δεδομένων. Παρόλα αυτά σε συγκεκριμένες περιπτώσεις, όπως π.χ. όταν υπάρχει ανάγκη βελτίωσης της απόδοσης των SQL ερωτημάτων, το Yii PHP framework δίνει τη δυνατότητα παρέμβασης και παραμετροποίησης του μηχανισμού πρόσβασης στις σχεσιακές βάσεις δεδομένων.

Επιπλέον, το Yii PHP framework υλοποιεί το σχεδιαστικό πρότυπο Active Record, σύμφωνα με το οποίο κάθε εγγραφή που βρίσκεται στη βάση δεδομένων αντιμετωπίζεται από τον κώδικα της εφαρμογής ως ένα ξεχωριστό αντικείμενο, το οποίο διαθέτει μία σειρά από μεθόδους και ιδιότητες. Οι μέθοδοι αντιστοιχούν στις λειτουργίες εισαγωγής, ενημέρωσης και διαγραφής της αντίστοιχης εγγραφής ενώ οι ιδιότητες αντιστοιχούν στις αντίστοιχες στήλες της εγγραφής. Υλοποιώντας το συγκεκριμένο πρότυπο, ο προγραμματιστής έχει την ευχέρεια να προσπελαύνει τα δεδομένα με ένα κοινό μηχανισμό, είτε αυτά βρίσκονται αποθηκευμένα στη βάση δεδομένων είτε όχι.

4.2.4 Ενσωμάτωση του jQuery

Το Yii PHP framework συνεργάζεται με τη βιβλιοθήκη JavaScript jQuery (<http://jquery.com>), έτσι ώστε να παρέχει τη δυνατότητα ανάπτυξης σύγχρονων ιστοσελίδων που εκμεταλλεύο-

νται τα χαρακτηριστικά των browser τελευταίας γενιάς., όπως είναι η ανάπτυξη δυναμικών σελίδων που ανταποκρίνονται άμεσα στις ενέργειες του χρήστη μέσω της τεχνικής AJAX.

4.2.5 Πιστοποίηση χρηστών

Το Yii PHP framework παρέχει ένα ολοκληρωμένο σύστημα εισόδου και ταυτοποίησης χρηστών στην εφαρμογή, δίνοντας τη δυνατότητα στον προγραμματιστή να αναπτύξει με ευκολία εφαρμογές που παρέχουν υπηρεσίες διαμορφωμένες στον εκάστοτε χρήστη.

4.2.6 Εισαγωγή δεδομένων από HTML φόρμες

Το Yii PHP framework έρχεται με πλήρη υποστήριξη εισαγωγής και επαλήθευσης δεδομένων από HTML φόρμες, προσφέροντας απλότητα στη χρήση και μέγιστη ασφάλεια για το σύστημα. Επιπλέον, παρέχει μία σειρά από βοηθητικές μεθόδους και αντικείμενα που σκοπό έχουν να προσφέρουν έτοιμες λύσεις στα συχνότερα προβλήματα εισαγωγής δεδομένων από το χρήστη.

4.2.7 Δυνατότητες παραμετροποίησης της εμφάνισης της εφαρμογής

Το Yii PHP framework παρέχει δυνατότητες αλλαγής στην εμφάνιση της εφαρμογής χωρίς να είναι απαραίτητη κάποια τροποποίηση στον κώδικα αυτής. Επιπλέον, οι ξεχωριστές εικαστικές διαμορφώσεις του συστήματος μπορούν να συνδεθούν με ξεχωριστούς λογαριασμούς χρηστών έτσι ώστε ο κάθε χρήστης να χρησιμοποιεί την εφαρμογή με τα χρώματα και το στυλ της επιθυμίας του. Παρέχονται, συγκεκριμένα, οι δυνατότητες χρήσης themes, τα οποία δρουν στα διαθέσιμα views της εφαρμογής και skins, τα οποία δρουν στα διαθέσιμα widgets της εφαρμογής. Οι συγκεκριμένες δυνατότητες επιτρέπουν ένα μεγάλο βαθμό ευελιξίας στη διαμόρφωση της εμφάνισης της εφαρμογής.

4.2.8 Υποστήριξη για web services

Το Yii PHP framework έρχεται εξοπλισμένο με υποστήριξη για την εύκολη δημιουργία web services. Υποστηρίζει τη γλώσσα περιγραφής Web Service Definition Language (<http://www.w3.org/TR/wsdl>) πετυχαίνοντας με αυτόν τον τρόπο μέγιστη λειτουργικότητα των υπό ανάπτυξη υπηρεσιών. Αν και στο πρώτο πλάνο της σχεδίασης της εφαρμογής δεν υπήρχε η ανάγκη παράλληλης ανάπτυξης αντίστοιχου web service, κρίθηκε χρήσιμη η δυνατότητα εύκολης μελλοντικής επέκτασης στο συγκεκριμένο αντικείμενο.

4.2.9 Υποστήριξη πολλαπλών γλωσσών

Το Yii PHP framework υποστηρίζει μια σειρά από τεχνικές που διευκολύνουν την ανάπτυξη δικτυακών εφαρμογών σε πολλαπλές γλώσσες. Συγκεκριμένα, έχει τη δυνατότητα να υπο-

στηρίζει τη μετάφραση όλων των λέξεων και φράσεων που χρησιμοποιεί η εφαρμογή σε ένα απεριόριστο πλήθος γλωσσών. Υποστηρίζει την ανάγνωση και εμφάνιση ημερομηνιών στην τοπική μορφή αλλά και τη μορφοποίηση αριθμών σύμφωνα με τα τοπικά πρότυπα. Οι συγκεκριμένες δυνατότητες μπορούν να εφαρμοστούν σε κάθε λογαριασμό χρήστη ξεχωριστά με αποτέλεσμα να δημιουργούνται εφαρμογές που έχουν την ευχέρεια να απευθυνθούν στο παγκόσμιο κοινό.

4.2.10 Δυνατότητες caching πολλαπλών επιπέδων

Το Yii PHP framework έχει πολλαπλές δυνατότητες αξιοποίησης τεχνικών caching. Μπορεί να εφαρμόσει caching στα επίπεδα των δεδομένων, των στατικών σελίδων, των τμημάτων σελίδων ή και ακόμα των δυναμικών σελίδων. Για την αποθήκευση δύναται να χρησιμοποιήσει μία πληθώρα από αποθηκευτικά μέσα, όπως αρχεία ή / και βάσεις δεδομένων.

4.2.11 Διαχείριση σφαλμάτων και καταγραφή συμβάντων

Το Yii PHP framework διαχειρίζεται τα σφάλματα με το βέλτιστο δυνατό τρόπο δίνοντας στον προγραμματιστή όλες τις απαραίτητες τεχνικές πληροφορίες για να εντοπίσει και να επιδιορθώσει τις δυσλειτουργίες. Σε κάθε περίπτωση αστοχίας παρέχονται πληροφορίες σχετικές με το σφάλμα, το stack trace αλλά και τμήμα του κώδικα που προκάλεσε το πρόβλημα. Εκτός αυτού, υπάρχει η δυνατότητα αξιοποίησης ενός δυνατού μηχανισμού καταγραφής συμβάντων, ο οποίος μπορεί να υποστηρίξει κατηγοριοποίηση των συμβάντων και πολλαπλές οδεύσεις για την παρουσίαση ή αποθήκευσή τους.

4.2.12 Ασφάλεια

Το Yii PHP framework υλοποιεί όλες τις σύγχρονες τεχνικές πρόληψης και αποτροπής δικτυακών επιθέσεων. Συγκεκριμένα, υλοποιεί τεχνικές αποτροπής του Cross Site Scripting, του Cross Site Request Forgery και τις αθέμιτης τροποποίησης των cookies.

4.2.13 Συμβατότητα με πρότυπα

Ο κώδικας που παράγεται από το Yii PHP framework σε επίπεδο HTML είναι πλήρως συμβατός με το πρότυπο XHTML εξασφαλίζοντας τη λειτουργία της εφαρμογής σε όλους τους σύγχρονους browser.

4.2.14 Εργαλεία αυτόματης δημιουργίας κώδικα

Το Yii PHP framework παρέχει μία σειρά από αυτόματα εργαλεία δημιουργίας κώδικα που διευκολύνουν την προγραμματιστική διαδικασία. Συγκεκριμένα, έχει τη δυνατότητα να κατασκευάσει αυτόματα απλά interfaces για τη δημιουργία και επεξεργασία μοντέλων, ενώ επί-

σης έχει τη δυνατότητα να δημιουργήσει αυτόματα τον κώδικα των μοντέλων διαβάζοντας τη βάση δεδομένων και αντλώντας τις απαραίτητες μετα-πληροφορίες, όπως είναι το όνομα των πινάκων, τα ονόματα και ο τύπος των πεδίων αλλά και η ύπαρξη ξένων κλειδίων.

4.2.15 Τεκμηρίωση

Το Yii PHP framework συνοδεύεται από άψογη τεκμηρίωση, η οποία βρίσκεται τόσο σε επίπεδο κώδικα σε μορφή rhdos όσο και σε επίπεδο ξεχωριστής τεκμηρίωσης σε μορφή HTML αλλά και PDF. Με αυτόν τον τρόπο η ανάπτυξη εφαρμογών πάνω στη συγκεκριμένη πλατφόρμα διευκολύνεται μιας και ο προγραμματιστής έχει στη διάθεσή του όλες τις απαραίτητες πληροφορίες με την απαιτούμενη λεπτομέρεια ώστε να υλοποιήσει το σύστημα σύμφωνα με τις προδιαγραφές που έχουν τεθεί.

4.2.16 Δυνατότητα επεκτάσεων

Το Yii PHP framework έχει τη δυνατότητα να δεχτεί επεκτάσεις με αποτέλεσμα η λίστα των χαρακτηριστικών του να είναι διαρκώς αυξανόμενη. Η συγγραφή επεκτάσεων είναι μία καλά τεκμηριωμένη διαδικασία με αποτέλεσμα να υπάρχει ήδη ένας εκτενής κατάλογος από επεκτάσεις που έχουν υλοποιηθεί για την πλατφόρμα. Η χρήση των επεκτάσεων είναι εξίσου εύκολη με αποτέλεσμα ο προγραμματιστής να έχει στη διάθεσή του μία μεγάλη συλλογή από έτοιμες λύσεις, οι οποίες σε γενικές γραμμές καλύπτουν τα συνηθέστερα ζητήματα ανάπτυξης δικτυακών εφαρμογών.

4.2.17 Ενεργή ανάπτυξη και υποστήριξη

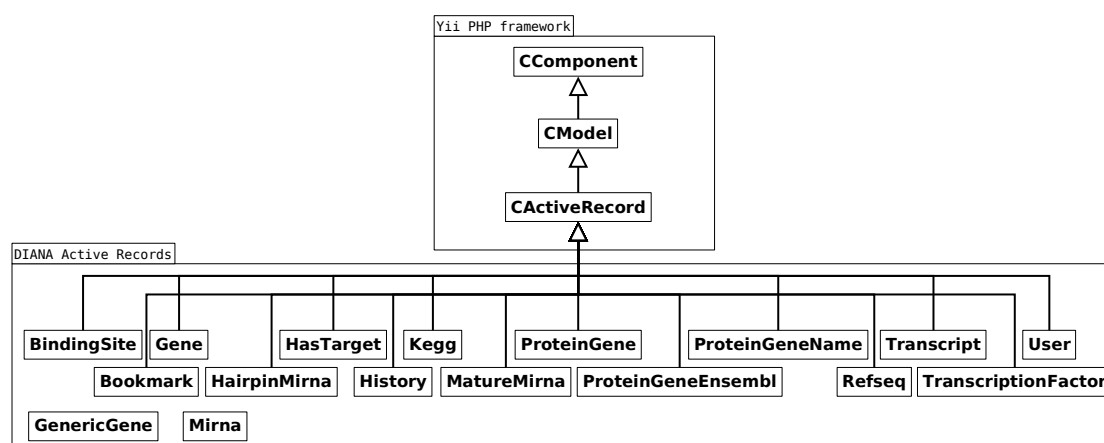
Το Yii PHP framework γνωρίζει ιδιαίτερη ανάπτυξη προβαίνοντας σε συχνές ανανεώσεις, οι οποίες στοχεύουν τόσο στη διόρθωση τυχόν δυσλειτουργιών όσο και στην προσθήκη νέων χαρακτηριστικών. Διαθέτει μία ομάδα ανάπτυξης, η οποία ανταποκρίνεται καθημερινά στα ερωτήματα των χρηστών, οι οποίοι έχουν στη διάθεσή τους ένα εξειδικευμένο forum συζητήσεων (<http://www.yiiframework.com/forum/>) και ένα σύστημα issue tracking για την αντιμετώπιση των προβλημάτων τους (<http://code.google.com/p/yii/issues/list>). Η βάση ενεργών χρηστών αυξάνεται σταδιακά, όπως αυτό αποτυπώνεται και από την αυξανόμενη δραστηριότητα στο αντίστοιχο forum συζητήσεων.

4.3 Περιγραφή Κλάσεων

Στην παρούσα ενότητα τεκμηριώνονται οι κλάσεις του συστήματος με συνοπτικό και σαφή τρόπο. Για κάθε κλάση προσδιορίζεται η χρήση της στο πλαίσιο του συστήματος και παρουσιάζονται οι σημαντικότερες μέθοδοι αυτών.

4.3.1 Κλάσεις *Active Record*

Οι κλάσεις που ακολουθούν το σχεδιαστικό πρότυπο *Active Record* χρησιμοποιούνται για τη σύνδεση της εφαρμογής με τη σχεσιακή βάση δεδομένων τηρώντας το αντικειμενοστραφές μοντέλο προγραμματισμού. Ουσιαστικά, υπάρχει μία κλάση ανά πίνακα της βάσης που ενθυλακώνει τις βασικές λειτουργίες που μπορούν να πραγματοποιηθούν στο συγκεκριμένο πίνακα, όπως είναι η αναζήτηση, η προσθήκη και η διαγραφή γραμμών. Επιπλέον, υπάρχουν και επιπρόσθετες μέθοδοι που υλοποιούν συγκεκριμένες χρήσιμες ενέργειες όπου αυτό κρίνεται απαραίτητο.



Εικόνα 17: Διάγραμμα κλάσεων για τις κλάσεις *Active Record* του συστήματος.

4.3.1.1 Κλάση *BindingSite*

Μέθοδος `getRnaHybridGraphic`

Χρησιμοποιείται για την εύκολη ανάγνωση του γραφικού που αναπαριστά το σημείο πρόσδεσης ενός βιολογικού στόχου. Το γραφικό αποτελείται από 4 γραμμές κειμένου και επιστρέφεται ως πίνακας για καλύτερη διαχείριση.

4.3.1.2 Κλάση *Bookmark*

Μέθοδος `putFilter`

Χρησιμοποιείται για την αποθήκευση σε μία αγαπημένη ιστοσελίδα των επιλογών του χρήστη όσον αφορά τους προβαλλόμενους βιολογικούς στόχους.

Μέθοδος `getFilter`

Χρησιμοποιείται για την ανάγνωση από τη βάση δεδομένων των προβαλλόμενων βιολογικών στόχων μίας αγαπημένης ιστοσελίδας.

4.3.1.3 Κλάση Gene

Μέθοδος getTransformedDescription

Μετασχηματίζει την περιγραφή ενός γονιδίου με τέτοιο τρόπο ώστε να μπορεί να χρησιμοποιηθεί για αναζήτηση στη δικτυακή βάση δεδομένων δημοσιεύσεων βιοιατρικής PubMed.

4.3.1.4 Κλάση HairpinMirna

Η παρούσα κλάση αναπαριστά ένα hairpin microRNA.

4.3.1.5 Κλάση HasTarget

Μέθοδος filterByScore

Προσθέτει ένα κριτήριο σε μία αναζήτηση βιολογικών στόχων, το οποίο απομονώνει και επιστρέφει αυτούς που επιτυγχάνουν βαθμολογία ανώτερη της δοθείσας.

Μέθοδος filterByMatureMirna

Προσθέτει ένα κριτήριο σε μία αναζήτηση βιολογικών στόχων, το οποίο απομονώνει και επιστρέφει αυτούς που αντιστοιχούν σε συγκεκριμένα microRNA.

Μέθοδος filterByTranscripts

Προσθέτει ένα κριτήριο σε μία αναζήτηση βιολογικών στόχων, το οποίο απομονώνει και επιστρέφει αυτούς που αντιστοιχούν σε συγκεκριμένα transcripts, τα οποία μπορεί να είναι περισσότερα του ενός σε πλήθος.

Μέθοδος filterByKeggPath

Προσθέτει ένα κριτήριο σε μία αναζήτηση βιολογικών στόχων, το οποίο απομονώνει και επιστρέφει αυτούς που αντιστοιχούν σε γονίδια που ταιριάζουν στο δοθέν KEGG pathway.

Μέθοδος withBindingSitesCount

Προσθέτει ένα κριτήριο σε μία αναζήτηση βιολογικών στόχων, το οποίο επιστρέφει το πλήθος των σημείων πρόσδεσης για κάθε βιολογικό στόχο. Χρησιμοποιείται για τον αρχικό υπολογισμό του πλήθους των σημείων πρόσδεσης που εντοπίζονται από μία αναζήτηση.

Μέθοδος getAlsoFound

Επιστρέφει σε μορφή πίνακα τα πεδία also found των βιολογικών στόχων. Αυτά τα πεδία δηλώνουν το γεγονός ότι ένας συγκεκριμένος στόχος έχει βρεθεί και από κάποιον άλλο ανεξάρτητο αλγόριθμο. Συγκεκριμένα, λαμβάνονται υπόψη οι αλγόριθμοι PicTar και TargetScan.

4.3.1.6 Κλάση History

Μέθοδος add

Προσθέτει μία εγγραφή στον κατάλογο του ιστορικού για τον τρέχοντα συνδεδεμένο χρήστη και την τρέχουσα επιχειρούμενη αναζήτηση.

4.3.1.7 Κλάση Kegg

Η συγκεκριμένη κλάση αναπαριστά ένα KEGG pathway.

4.3.1.8 Κλάση MatureMirna

Μέθοδος getTransformedNameToIHopQuery

Μετασχηματίζει το όνομα ενός microRNA με τέτοιο τρόπο ώστε να μπορεί να χρησιμοποιηθεί για αναζήτηση στη δικτυακή βάση δεδομένων iHOP.

Μέθοδος getTransformedName

Μετασχηματίζει το όνομα ενός microRNA με τέτοιο τρόπο ώστε να μπορεί να χρησιμοποιηθεί για αναζήτηση στη δικτυακή βάση δεδομένων δημοσιεύσεων βιοιατρικής PubMed.

Μέθοδος getHairpinNames

Επιστρέφει έναν πίνακα με όλα τα ονόματα των hairpins ενός microRNA.

Μέθοδος getExpressionLevelValue

Επιστρέφει την τιμή ενός expression level του microRNA. Υπάρχουν 240 expression levels σε κάθε microRNA και για να διευκολυνθεί η προσπέλασή τους παρέχεται η συγκεκριμένη μέθοδος.

Μέθοδος getExpressionLevelColor

Επιστρέφει το χρώμα που αντιστοιχεί σε ένα expression level του microRNA. Η αντιστοίχιση της τιμής στο χρώμα γίνεται μέσω μίας συνάρτησης βαθμιαίας μεταβολής του χρώματος από άσπρο σε πορτοκαλί και τελικά σε κόκκινο.

4.3.1.9 Κλάση ProteinGene

Μέθοδος getNames

Επιστρέφει έναν πίνακα με όλα τα ονόματα ενός γονιδίου.

Μέθοδος getRefseqs

Επιστρέφει έναν πίνακα με όλα τα ονόματα των refseq ενός γονιδίου.

Μέθοδος getTranscripts

Επιστρέφει έναν πίνακα με όλα τα ονόματα των transcripts ενός γονιδίου.

Μέθοδος getTranscriptIds

Επιστρέφει έναν πίνακα με όλα τα αναγνωριστικά των transcripts ενός γονιδίου.

4.3.1.10 Κλάση *ProteinGeneName*

Η παρούσα κλάση αναπαριστά ένα από τα ονόματα ενός συγκεκριμένου γονιδίου.

4.3.1.11 Κλάση *Refseq*

Η συγκεκριμένη κλάση αναπαριστά ένα από τα refseqs ενός γονιδίου.

4.3.1.12 Κλάση *Transcript*

Η παρούσα κλάση αναπαριστά ένα από τα transcripts ενός συγκεκριμένου γονιδίου.

4.3.1.13 Κλάση *TranscriptionFactor*

Μέθοδος suggest

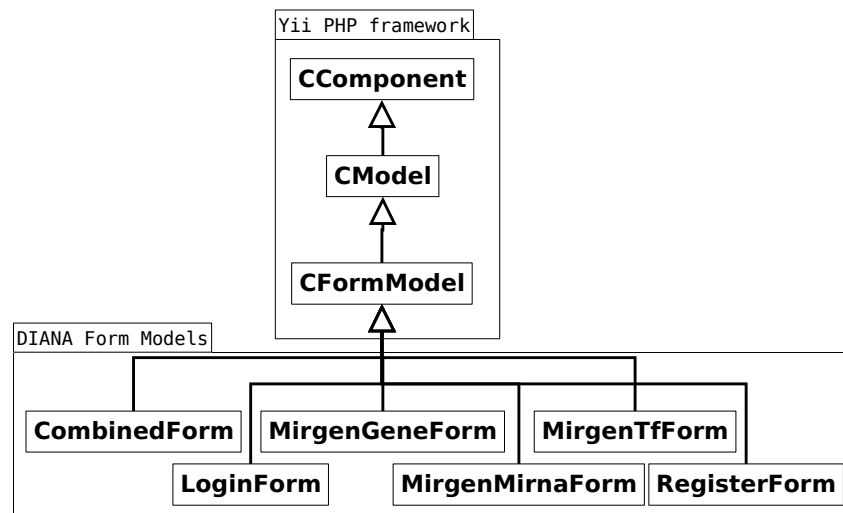
Επιστρέφει μία σειρά από προτάσεις εφόσον ο όρος αναζήτησης ενός transcription factor δεν εντοπίζει μία μοναδική εγγραφή από τον αντίστοιχο πίνακα.

4.3.1.14 Κλάση *User*

Η συγκεκριμένη κλάση αναπαριστά έναν χρήστη του συστήματος και χρησιμοποιείται σε όλα τα υποσυστήματα που προσφέρουν προσωπικές υπηρεσίες ούτως ώστε να επιτυγχάνεται ο προσδιορισμός του χρήστη που προσπελαύνει την εκάστοτε υπηρεσία.

4.3.2 Κλάσεις *Form Model*

Οι συγκεκριμένες κλάσεις χρησιμοποιούνται για την αναπαράσταση των δεδομένων που αντιστοιχούν στις φόρμες HTML. Χρησιμοποιούνται από το Yii PHP framework για την ανάγνωση των δεδομένων από τις συγκεκριμένες φόρμες και τη μετέπειτα επεξεργασία τους. Έχουν συγκεκριμένες δυνατότητες που προσφέρουν ασφάλεια, αξιοπιστία και ευκολία στο χειρισμό δεδομένων που προέρχονται από το χρήστη.



Εικόνα 18: Διάγραμμα κλάσεων για τις κλάσεις Form Model του συστήματος.

4.3.2.1 Κλάση *CombinedForm*

Η συγκεκριμένη κλάση χρησιμοποιείται για την ανάγνωση των όρων αναζήτησης βιολογικών στόχων, την πραγματοποίηση της αναζήτησης και την παρουσίαση των αποτελεσμάτων στο χρήστη.

4.3.2.2 Κλάση *MirgenMirnaForm*

Η συγκεκριμένη κλάση χρησιμοποιείται για την αναζήτηση πληροφοριών ενός microRNA στην εφαρμογή mirGen. Αναλαμβάνει τις διαδικασίες από την ανάγνωση του όρου αναζήτησης από το χρήστη έως την απεικόνιση των αποτελεσμάτων.

4.3.2.3 Κλάση *MirgenTfForm*

Η συγκεκριμένη κλάση χρησιμοποιείται για την αναζήτηση πληροφοριών ενός transcription factor στην εφαρμογή mirGen. Αναλαμβάνει τις διαδικασίες από την ανάγνωση του όρου αναζήτησης από το χρήστη έως την απεικόνιση των αποτελεσμάτων.

4.3.2.4 Κλάση *LoginForm*

Μέθοδος *authenticate*

Ελέγχει αν τα στοιχεία πρόσβασης που έχουν εισαχθεί στη φόρμα αντιστοιχούν σε υπάρχοντα χρήστη του συστήματος. Σε περίπτωση που το όνομα χρήστη και ο κωδικός πρόσβασης είναι έγκυρα, τότε το σύστημα προχωράει στην ταυτοποίηση του χρήστη μέσω ενός session cookie.

4.3.2.5 Κλάση *RegisterForm*

Μέθοδος *uniqueUsername*

Ελέγχει αν το ζητούμενο όνομα χρήστη με το οποίο επιχειρεί να γραφτεί ένας νέος χρήστης στο σύστημα είναι ελεύθερο για χρήση. Σε περίπτωση που το όνομα χρήστη είναι δεσμευμένο τότε εμφανίζεται κατάλληλο μήνυμα ενημέρωσης προς το χρήστη ούτως ώστε να επιλέξει ένα διαφορετικό όνομα χρήστη.

Μέθοδος afterValidate

Καταχωρεί τα στοιχεία πρόσβασης που επιθυμεί να χρησιμοποιεί ο νέος χρήστης στο σύστημα. Τα στοιχεία καταχωρούνται στη βάση δεδομένων και, πλέον, δίνεται η δυνατότητα στο νέο χρήστη να συνδέεται στο σύστημα και να χρησιμοποιεί τις προσωπικές υπηρεσίες που επιθυμεί.

4.3.3 Άλλα μοντέλα

4.3.3.1 Κλάση GenericGene

Επειδή ένα γονίδιο μπορεί να αναπαρασταθεί με πολλούς τρόπους, κρίθηκε αναγκαία η δημιουργία ενός μετα-μοντέλου που να χρησιμοποιείται για την περιγραφή ενός γονιδίου σε υψηλότερο επίπεδο αφαίρεσης σε σχέση με τα μοντέλα Transcript και ProteinGene.

Μέθοδος suggest

Επιστρέφει προτάσεις για αναζήτηση γονιδίου με βάση το όνομα. . Οι προτάσεις αναζητούνται στα εξής πεδία:

- Πεδία ensemble_id των γονιδίων
- Ονόματα των γονιδίων
- Πεδία transcript_id των transcripts
- Πεδία refseq_id των refseqs

Μέθοδος suggestByDescription

Επιστρέφει προτάσεις για αναζήτηση γονιδίου με βάση την περιγραφή.

Μέθοδος getName

Επιστρέφει το χαρακτηριστικό όνομα ενός γονιδίου. Πρώτα προτιμάτε το ensemble του γονιδίου (αν υπάρχει) αλλιώς προτιμάτε το όνομα του transcript.

4.3.3.2 Κλάση Mirna

Για τους ίδιους λόγους όπως και στην περίπτωση των γονιδίων υπάρχει ανάγκη δημιουργίας ενός μετα-μοντέλου για την αναπαράσταση των microRNA σε υψηλότερο αφαιρετικό επίπεδο σε σχέση με τα μοντέλα MatureMirna και HairpinMirna.

Μέθοδος suggest

Επιστρέφει προτάσεις για αναζήτηση microRNA με βάση το όνομα. Οι προτάσεις αναζητούνται στα εξής πεδία:

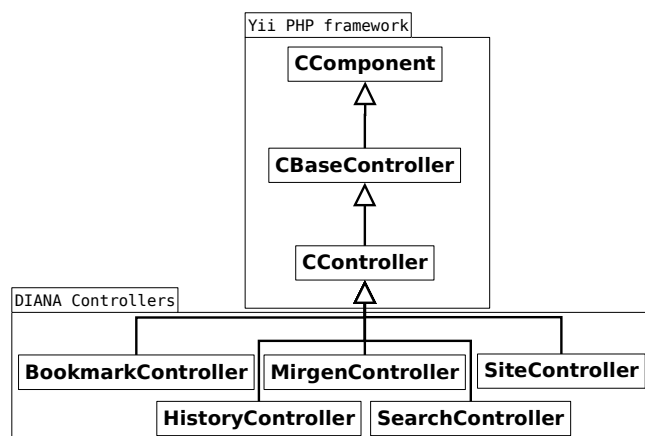
- Ονόματα των mature microRNA
- Πεδία mimat των mature microRNA
- Ονόματα των hairpin microRNA
- Πεδία mima_id των hairpin microRNA

Μέθοδος getName

Επιστρέφει το χαρακτηριστικό όνομα ενός microRNA. Πρώτα προτιμάτε το όνομα του mature microRNA (αν υπάρχει) αλλιώς προτιμάτε το όνομα του hairpin microRNA.

4.3.4 Ελεγκτές

Οι ελεγκτές είναι μέρος του αρχιτεκτονικού προτύπου Model – View – Controller και αποτελούν τη ραχοκοκαλιά του συστήματος.



Εικόνα 19: Διάγραμμα κλάσεων για τους ελεγκτές του συστήματος.

4.3.4.1 Κλάση BookmarkController

Η συγκεκριμένη κλάση υποστηρίζει την εκτέλεση των απαραίτητων ενεργειών για τη διαχείριση των αγαπημένων αναζητήσεων. Οι ενέργειες που υποστηρίζονται είναι οι εξής:

- Προσθήκη αγαπημένης αναζήτησης
- Διαγραφή αγαπημένης αναζήτησης
- Τροποποίηση υπάρχουσας αγαπημένης αναζήτησης
- Προβολή καταλόγου αγαπημένων αναζητήσεων για τον τρέχοντα χρήστη

4.3.4.2 Κλάση *HistoryController*

Η συγκεκριμένη κλάση υποστηρίζει τις ενέργειες για τη διαχείριση του ιστορικού από την πλευρά του εκάστοτε χρήστη. Ο κάθε χρήστης έχει τη δυνατότητα να επιλέξει να καθαρίσει το ιστορικό των αναζητήσεών του. Η εντολή για τον καθαρισμό του ιστορικού δίνεται μέσω της συγκεκριμένης κλάσης.

4.3.4.3 Κλάση *MirgenController*

Η συγκεκριμένη κλάση υποστηρίζει τη λειτουργία της εφαρμογής mirGen. Υποστηρίζει την εύρεση πληροφοριών για microRNA και transcription factor.

4.3.4.4 Κλάση *SearchController*

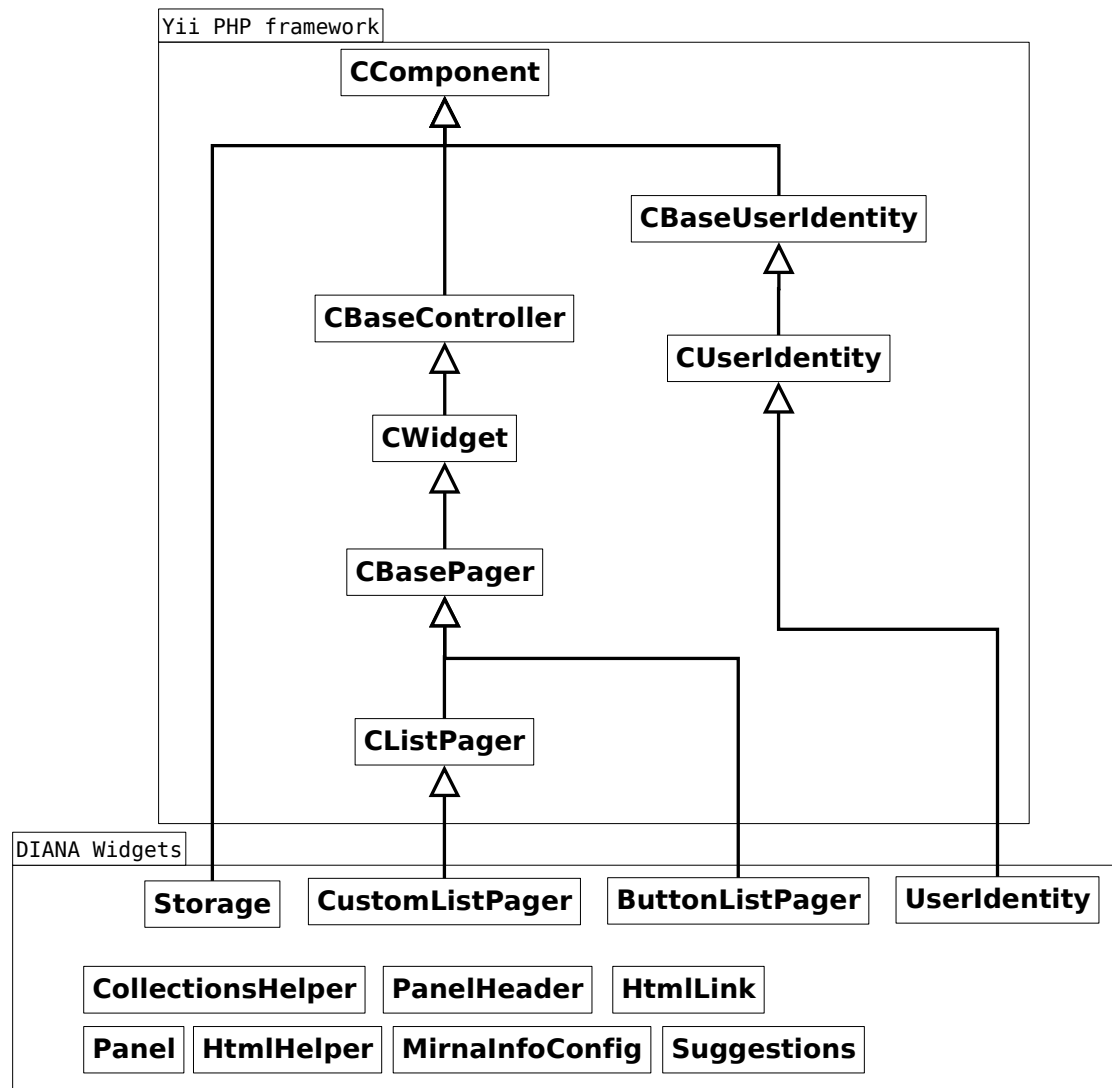
Η συγκεκριμένη κλάση υποστηρίζει τη λειτουργία της αναζήτησης βιολογικών στόχων. Υποστηρίζει μία λειτουργία σύνθετης αναζήτησης, η οποία προσαρμόζεται ανάλογα με το είδος και το πλήθος των δεδομένων που εισάγει ο χρήστης ώστε να δώσει τα αναμενόμενα αποτελέσματα κάθε φορά.

4.3.4.5 Κλάση *SiteController*

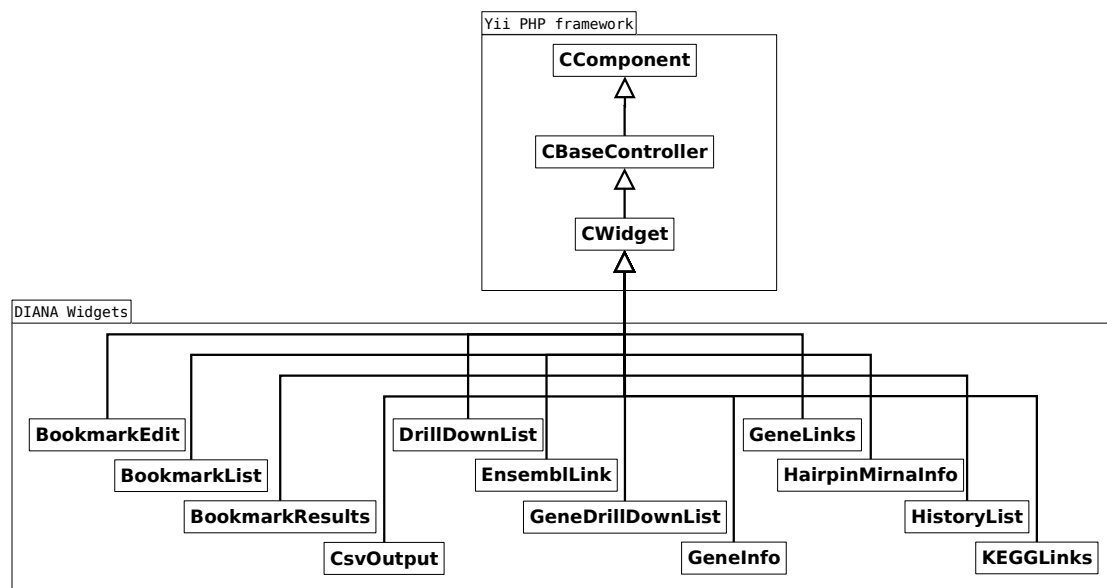
Η συγκεκριμένη κλάση υλοποιεί τις υποστηρικτές λειτουργίες της δικτυακής εφαρμογής. Μέσω της συγκεκριμένης κλάσης πραγματοποιείται η είσοδος, η έξοδος και η εγγραφή χρηστών στο σύστημα. Επιπλέον, υπάρχει δυνατότητα για προβολή στοιχείων σχετικών με το σύστημα όπως είναι:

- Το διαθέσιμο λογισμικό
- Οι διαθέσιμες βάσεις δεδομένων με πληροφορίες βιολογικού ενδιαφέροντος
- Τα μέλη του εργαστηρίου που συνέβαλαν στη δημιουργία της εφαρμογής
- Οι σχετικές με το αντικείμενο δημοσιεύσεις
- Η βοήθεια του συστήματος

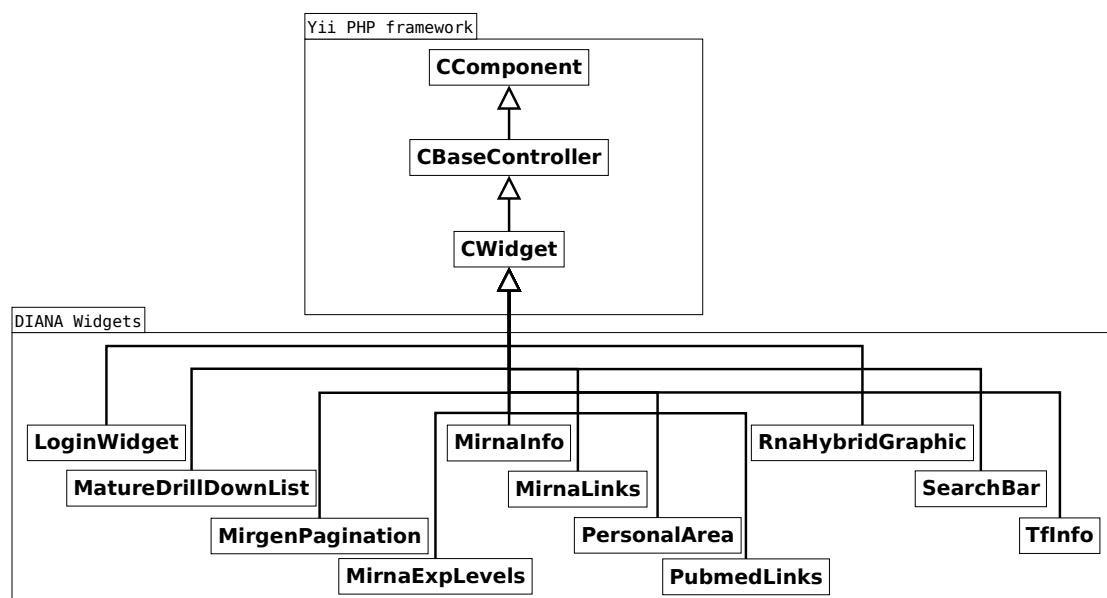
4.3.5 Widgets



Εικόνα 20: Διάγραμμα κλάσεων για τα widgets του συστήματος.



Εικόνα 21: Διάγραμμα κλάσεων για τα widgets του συστήματος.



Εικόνα 22: Διάγραμμα κλάσεων για τα widgets του συστήματος.

4.3.5.1 Κλάση CollectionsHelper

Μέθοδος getAttributeFromCollection

Επιστρέφει σε μορφή πίνακα ένα δεδομένο attribute από μία συλλογή αντικειμένων. Χρησιμοποιείται για την απομόνωση στοιχείων από συλλογές Active Records, π.χ. το όνομα όλων των hairpins.

4.3.5.2 Κλάση Panel

Αναπαριστά ένα panel του στοιχείου drill down list.

Μέθοδος createPanel

Κατασκευάζει και επιστρέφει ένα εσωτερικό panel.

Μέθοδος setHeader

Ορίζει την επικεφαλίδα του panel.

Μέθοδος hasCheckbox

Ελέγχει αναδρομικά αν το συγκεκριμένο panel ή κάποιο από τα εσωτερικά του έχουν ορατό checkbox.

Μέθοδος hasToggle

Ελέγχει αναδρομικά αν το συγκεκριμένο panel ή κάποιο από τα εσωτερικά του έχουν ορατή επιλογή για την εμφάνιση ή απόκρυψή τους.

4.3.5.3 Κλάση PanelHeader

Αναπαριστά μία επικεφαλίδα που χρησιμοποιείται στο στοιχείο drill down list.

Μέθοδος __construct

Κατασκευάζει μία καινούρια επικεφαλίδα. Κατά την κατασκευή ελέγχεται αν το συνολικό πλάτος της επικεφαλίδας, το οποίο ορίζεται από τα επιμέρους πλάτη των στηλών που περιλαμβάνονται ξεπερνάει ένα δεδομένο μέγιστο όριο. Αν υπάρχει υπέρβαση, τότε προκαλείται σφάλμα εκτέλεσης ούτως ώστε ο προγραμματιστής να διορθώσει τις διαστάσεις της επικεφαλίδας.

Μέθοδος __toString

Επιστρέφει σε μορφή HTML την επικεφαλίδα για την άμεση προσθήκη αυτής στη διαμορφούμενη σελίδα που προορίζεται για εμφάνιση στον browser του χρήστη.

Μέθοδος getWidth

Επιστρέφει το πλάτος μίας δοσμένης στήλης. Χρησιμοποιείται για την ευθυγραμμισμένη παρουσίαση των στηλών της επικεφαλίδας και των στηλών των περιεχομένων του panel.

4.3.5.4 Κλάση HtmlHelper

Μέθοδος start_table

Επιστρέφει κώδικα HTML για το άνοιγμα ενός πίνακα που προσφέρει κατάλληλη μορφοποίηση στα περιεχόμενά του. Χρησιμοποιείται για την ομαδοποίηση της πληροφορίας στη σελίδα που εμφανίζεται στον browser του χρήστη.

Μέθοδος close_table

Αντίστοιχα όπως και παραπάνω, επιστρέφει κώδικα HTML για το κλείσιμο του πίνακα.

Μέθοδος `activeTextFieldWithHint`

Επιστρέφει ένα πεδίο εισαγωγής κειμένου, το οποίο έχει το πρόσθετο χαρακτηριστικό να παρουσιάζει μία προτροπή προς το χρήστη. Η προτροπή έχει ως στόχο την παροχή κατεύθυνσης προς το χρήστη σχετικά με τα στοιχεία που οφείλει να εισάγει στο πεδίο κειμένου. Η προτροπή εξαφανίζεται αυτόματα όταν ο χρήστης επιλέξει το στοιχείο και εμφανίζεται αυτόματα όταν ο χρήστης απομακρυνθεί από αυτό.

4.3.5.5 Κλάση `HtmlLink`

Κατασκευάζει έναν σύνδεσμο HTML, ο οποίος έχει τη δυνατότητα να εμφανίζει το στόχο του είτε σε νέο παράθυρο (καρτέλα για τους μοντέρνους browser) είτε με τη μορφή αναδυόμενου παραθύρου (popup window). Για τη λειτουργία του συνδέσμου χρησιμοποιείται η βιβλιοθήκη JQuery.

4.3.5.6 Κλάση `MirnaInfoConfig`

Περιέχει τα ονόματα των expression levels που παρουσιάζονται ως πληροφορία μαζί με τα microRNA. Περιλαμβάνει δύο πίνακες, όπου ο πρώτος έχει τα ονόματα για τα microRNA που αντιστοιχούν στον άνθρωπο και ο δεύτερος έχει τα ονόματα που αντιστοιχούν στο ποντίκι.

4.3.5.7 Κλάση `Suggestions`**Μέθοδος `suggest`**

Επιστρέφει προτάσεις για ένα όρο αναζήτησης ακολουθώντας τις προδιαγραφές που δίνονται ως παράμετρος.

Μέθοδος `uniqueSuggestion`

Ελέγχει αν υπάρχει μοναδική πρόταση για δοσμένο όρο αναζήτησης και την επιστρέφει.

Μέθοδος `multipleSuggestions`

Επιστρέφει πολλαπλές προτάσεις για δοσμένο όρο αναζήτησης ταξινομημένες σύμφωνα με το βαθμό συνάφειάς τους.

Μέθοδος `cmp_suggestions_by_score`

Χρησιμοποιείται από την παραπάνω μέθοδο για τη σύγκριση και ταξινόμηση των προτάσεων σύμφωνα με το βαθμό συνάφειάς τους.

4.3.5.8 Κλάση `Storage`

Η παρούσα κλάση χρησιμοποιείται για τη μεταφορά πληροφορίας μεταξύ διαφόρων υποσυστημάτων της εφαρμογής. Η συγκεκριμένη πληροφορία έχει σύντομη διάρκεια ζωής, συγκεκρι-

κριμένα είναι χρήσιμη μόνο κατά τη διάρκεια επεξεργασίας του αιτήματος HTTP, επομένως δε χρειάζεται να αποθηκευτεί στο session. Γι' αυτό το λόγο αποθηκεύεται σε ένα προσωρινό αντικείμενο στη μνήμη του συστήματος και διαγράφεται αυτομάτως όταν λήξει η επεξεργασία του HTTP ερωτήματος και επιστραφεί το αποτέλεσμα στον browser του χρήστη.

4.3.5.9 Κλάση *UserIdentity*

Μέθοδος *authenticate*

Ελέγχει αν η προσπάθεια σύνδεσης ενός χρήστη είναι έγκυρη. Ο έλεγχος πραγματοποιείται αναζητώντας τον εκάστοτε χρήστη στη βάση δεδομένων και επαληθεύοντας τη σωστή εισαγωγή του κωδικού πρόσβασής του.

Μέθοδος *getId*

Επιστρέφει ένα μοναδικό αναγνωριστικό που προσδιορίζει τον εκάστοτε συνδεδεμένο χρήστη.

4.3.5.10 Κλάση *ButtonListPager*

Αναπαριστά ένα στοιχείο επιλογής σελίδων, το οποίο περιλαμβάνει μία αναδυόμενη λίστα με τις διαθέσιμες σελίδες καθώς και δύο πλήκτρα μετάβασης στην προηγούμενη και στην επόμενη σελίδα από την τρέχουσα.

4.3.5.11 Κλάση *CustomListPager*

Υλοποιεί την αναδυόμενη λίστα του στοιχείου επιλογής σελίδων. Κρίθηκε απαραίτητη η συγκεκριμένη υλοποίηση ούτως ώστε να προσπεραστούν ορισμένες αδυναμίες της εσωτερικής υλοποίησης που παρέχεται από το Yii PHP framework.

4.3.5.12 Κλάση *BookmarkEdit*

Η παρούσα κλάση υλοποιεί ένα στοιχείου ελέγχου, το οποίο εμφανίζεται στον προσωπικό χώρο του χρήστη και επιτρέπει την επεξεργασία μίας αγαπημένης αναζήτησης.

4.3.5.13 Κλάση *BookmarkList*

Η συγκεκριμένη κλάση υλοποιεί ένα στοιχείο προβολής, το οποίο εμφανίζεται στον προσωπικό χώρο του χρήστη και συνοψίζει όλες τις αγαπημένες αναζητήσεις αυτού. Επιπλέον, υπάρχουν σύνδεσμοι για την προβολή, τροποποίηση και διαγραφή κάθε αγαπημένης αναζήτησης.

4.3.5.14 Κλάση *BookmarkResults*

Η παρούσα κλάση αναπαριστά ένα στοιχείο ελέγχου, το οποίο ελέγχει αν η εμφανιζόμενη σελίδα αποτελεσμάτων δεν υπάρχει αποθηκευμένη στις αγαπημένες αναζητήσεις του εκάστοτε χρήστη και εμφανίζει κατάλληλο σύνδεσμο για την προσθήκη της. Σε αντίθετη περίπτωση, εμφανίζει ενημερωτικό μήνυμα. Το συγκεκριμένο στοιχείο ελέγχου βρίσκεται στον προσωπικό χώρο του χρήστη.

4.3.5.15 Κλάση *CsvOutput*

Η συγκεκριμένη κλάση παράγει ένα αρχείο Comma Separated Values από έναν πίνακα με δεδομένα και το επιστρέφει στον browser του χρήστη. Στο αρχείο δίνεται μοναδικό όνομα αποτελούμενο από τον όρο αναζήτησης και την τρέχουσα ημερομηνία.

4.3.5.16 Κλάση *DrillDownList*

Η παρούσα κλάση υλοποιεί το στοιχείο παρουσίασης drill down list, το οποίο εμφανίζει δεδομένα σε δομημένη μορφή.

4.3.5.17 Κλάση *EnsemblLink*

Η συγκεκριμένη κλάση υλοποιεί το σύνδεσμο προς τη δικτυακή βάση δεδομένων Ensembl.

4.3.5.18 Κλάση *GeneDrillDownList*

Η παρούσα κλάση εξειδικεύει το στοιχείο drill down list για την παρουσίαση των αποτελεσμάτων σε μία αναζήτηση με βάση ένα γονίδιο.

4.3.5.19 Κλάση *GeneInfo*

Η συγκεκριμένη κλάση εμφανίζει το στοιχείο προβολής που περιλαμβάνει διάφορες πληροφορίες για ένα γονίδιο, όπως είναι το όνομά του, η περιγραφή του, τα KEGG pathways που αντιστοιχούν σε αυτό, κλπ.

4.3.5.20 Κλάση *GeneLinks*

Η παρούσα κλάση εμφανίζει τους συνδέσμους προς τις δικτυακές βάσεις δεδομένων UniProt και iHOP.

4.3.5.21 Κλάση *HairpinMirnaInfo*

Η συγκεκριμένη κλάση εμφανίζει το στοιχείο προβολής που περιλαμβάνει διάφορες πληροφορίες για ένα hairpin microRNA, όπως είναι το όνομά του, η περιοχή του και τα SNPs που αντιστοιχούν σε αυτό.

4.3.5.22 Κλάση *HistoryList*

Η παρούσα κλάση εμφανίζει σε συνοπτική μορφή τον κατάλογο των πρόσφατων αναζητήσεων του χρήστη. Οι αναζητήσεις εμφανίζονται σε αντίστροφη χρονολογική σειρά πραγματοποίησης και παρουσιάζονται στον προσωπικό χώρο του χρήστη.

4.3.5.23 Κλάση *KEGGLinks*

Η συγκεκριμένη κλάση εμφανίζει το σύνδεσμο προς τη δικτυακή βάση δεδομένων GenomeNet.

4.3.5.24 Κλάση *LoginWidget*

Η παρούσα κλάση υλοποιεί το στοιχείο ελέγχου μέσω του οποίου επιτυγχάνεται η είσοδος στο σύστημα των ανώνυμων χρηστών. Η είσοδος και ταυτοποίηση πραγματοποιείται μέσω του ονόματος χρήστη και του κωδικού πρόσβασης που εισάγονται.

4.3.5.25 Κλάση *MatureDrillDownList*

Η συγκεκριμένη κλάση εξειδικεύει το στοιχείο drill down list για την παρουσίαση των αποτελεσμάτων σε μία αναζήτηση με βάση ένα microRNA.

4.3.5.26 Κλάση *MirgenPagination*

Η παρούσα κλάση υλοποιεί το στοιχείο αλλαγής σελίδων για την παρουσίαση αποτελεσμάτων της εφαρμογής mirGen.

4.3.5.27 Κλάση *MirnaExpLevels*

Η συγκεκριμένη κλάση εμφανίζει τα expression levels ενός δεδομένου microRNA. Η εμφάνιση πραγματοποιείται μέσω ενός χρωματικού χάρτη, ο οποίος δίνει περισσότερες λεπτομέρειες στο χρήστη όταν αυτός περάσει το δείκτη του ποντικιού του πάνω από τα στοιχεία του χάρτη.

4.3.5.28 *Κλάση MirnaInfo*

Η παρούσα κλάση εμφανίζει το στοιχείο προβολής που περιλαμβάνει διάφορες πληροφορίες για ένα microRNA, όπως είναι το όνομά του, η περιγραφή του, η ακολουθία του καθώς και τα hairpin microRNAs που αντιστοιχούν σε αυτό.

4.3.5.29 *Κλάση MirnaLinks*

Η συγκεκριμένη κλάση εμφανίζει τους συνδέσμους προς τις δικτυακές βάσεις δεδομένων iHOP και miRBase.

4.3.5.30 *Κλάση PersonalArea*

Η παρούσα κλάση υλοποιεί τον προσωπικό χώρο του χρήστη, ο οποίος εμφανίζεται στο δεξί μέρος της σελίδας της εφαρμογής. Για τους χρήστες που δεν έχουν συνδεθεί εμφανίζεται το στοιχείο ελέγχου σύνδεσης, ενώ για τους συνδεδεμένους χρήστες εμφανίζονται οι αγαπημένες αναζητήσεις τους καθώς και το ιστορικό των πρόσφατων αναζητήσεών τους.

4.3.5.31 *Κλάση PubmedLinks*

Η συγκεκριμένη κλάση εμφανίζει τους συνδέσμους προς τη δικτυακή βάση δεδομένων PubMed.

4.3.5.32 *Κλάση RnaHybridGraphic*

Η παρούσα κλάση εμφανίζει το γραφικό που αναπαριστά το σημείο πρόσδεσης ενός βιολογικού στόχου. Εμφανίζονται κατάλληλα χρωματισμένες αλληλουχίες βάσεων RNA.

4.3.5.33 *Κλάση SearchBar*

Η συγκεκριμένη κλάση υλοποιεί το στοιχείο ελέγχου μέσω του οποίου έχει τη δυνατότητα ο χρήστης να μεταβάλλει και να εξειδικεύσει τους όρους της αναζήτησής του. Παρέχεται η δυνατότητα ορισμού του ελάχιστου βαθμού των εμφανιζόμενων βιολογικών στόχων καθώς και η προβολή βιολογικών στόχων που αντιστοιχούν σε δοσμένα KEGG pathways. Το συγκεκριμένο στοιχείο ελέγχου εμφανίζεται σε κάθε σελίδα αποτελεσμάτων.

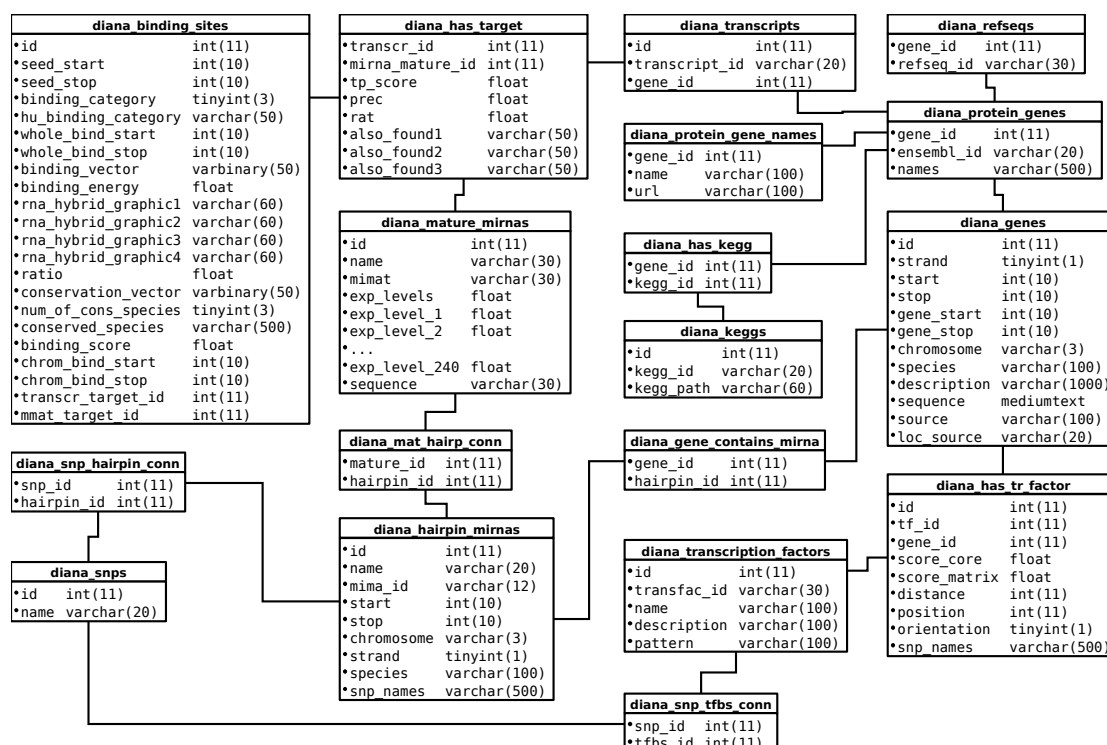
4.3.5.34 *Κλάση TfInfo*

Η παρούσα κλάση εμφανίζει το στοιχείο προβολής που περιλαμβάνει διάφορες πληροφορίες για ένα transcription factor, όπως είναι το όνομά του, κλπ.

4.4 Βάση Δεδομένων

Στο προηγούμενο κεφάλαιο της παρούσας διπλωματικής εργασίας παρουσιάστηκε το μοντέλο οντοτήτων συσχετίσεων, το οποίο χρησιμοποιήθηκε για την αναπαράσταση της πληροφορίας που διαχειρίζεται το σύστημα. Ακολουθώντας το συγκεκριμένο μοντέλο έγινε ο σχεδιασμός του αντίστοιχου σχήματος της βάσης δεδομένων, το οποίο περιλαμβάνει τους πίνακες και τις σχέσεις μεταξύ αυτών. Το τελικό αποτέλεσμα της σχεδίασης της βάσης δεδομένων φαίνεται σχηματικά παρακάτω. Αξίζουν να σημειωθούν ορισμένα σημεία που ελήφθησαν υπόψη κατά τη σχεδίαση του σχήματος:

- Κάθε πίνακας έχει πρωτεύον κλειδί ώστε να μεγιστοποιείται η ταχύτητα τυχαίας προσπέλασης των εγγραφών του.
- Το πρωτεύον κλειδί κάθε πίνακα έχει ληφθεί μέριμνα να έχει το ελάχιστο δυνατό μήκος ώστε το αντίστοιχο ευρετήριο που δημιουργείται από το σύστημα της βάσης δεδομένων να είναι όσο το δυνατόν περισσότερο αποδοτικό.
- Το μήκος κάθε πεδίου των πινάκων έχει επιλεγεί με γνώμονα τόσο τη δυνατότητα εισαγωγής δεδομένων δίχως περιορισμούς όσο και την εξοικονόμηση χώρου αλλά και την αύξηση των επιδόσεων.



Εικόνα 23: Σχηματική αναπαράσταση των πινάκων και των συσχετίσεων μεταξύ τους που χρησιμοποιούνται για την υποστήριξη του συστήματος της παρούσας διπλωματικής εργασίας.

Στη συνέχεια παρουσιάζονται και αναλύονται επιγραμματικά οι πίνακες του σχήματος της βάσης δεδομένων.

diana_binding_sites: Αναπαριστά ένα σημείο πρόσδεσης ενός microRNA πάνω σε ένα γονίδιο. Περιλαμβάνονται διάφορα χαρακτηριστικά μεγέθη με βιολογικό ενδιαφέρον.

diana_genes: Αναπαριστά ένα γονίδιο. Περιλαμβάνεται μία περιγραφή και η γονιδιακή ακολουθία.

diana_hairpin_mirnas: Αναπαριστά ένα hairpin microRNA, από το οποίο δημιουργούνται τα mature microRNA. Περιλαμβάνεται το όνομα και βιολογικά χαρακτηριστικά.

diana_has_target: Αναπαριστά ένα βιολογικό στόχο, δηλαδή την ύπαρξη σχέσης μεταξύ ενός microRNA και ενός γονιδίου. Περιλαμβάνονται ο βαθμός του βιολογικού στόχου και οι αλγόριθμοι που εντοπίζουν το συγκεκριμένο στόχο.

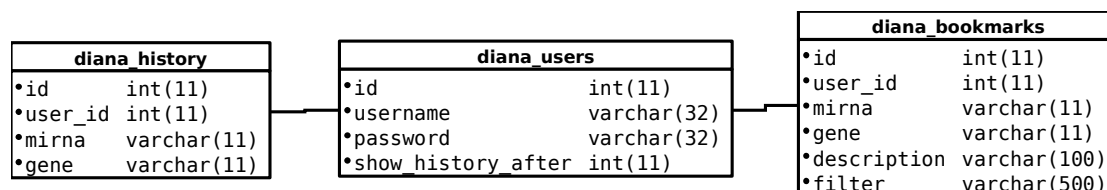
diana_keggs: Αναπαριστά ένα KEGG pathway.

diana_mature_mirnas: Αναπαριστά ένα mature microRNA. Περιλαμβάνεται το όνομα και ένα σύνολο από επίπεδα έκφρασης.

diana_protein_gene_names: Αναπαριστά ένα όνομα ενός γονιδίου. Κάθε γονίδιο δύναται να έχει ένα σύνολο από διαφορετικά ονόματα.

diana_protein_genes: Αναπαριστά ένα γονίδιο που κωδικοποιεί μία πρωτεΐνη. Είναι εξειδίκευση του γονιδίου. Περιλαμβάνεται το αναγνωριστικό για τη δικτυακή βάση δεδομένων Ensembl.

diana_refseqs: Αναπαριστά μία εγγραφή στη δικτυακή βάση δεδομένων RefSeq. Κάθε γονίδιο μπορεί να διαθέτει ένα σύνολο από εγγραφές στη συγκεκριμένη βάση δεδομένων.



Εικόνα 24: Σχηματική αναπαράσταση των πινάκων και των συσχετίσεων μεταξύ τους που χρησιμοποιούνται για την παροχή προσωπικών υπηρεσιών στο σύστημα της παρούσας διπλωματικής εργασίας.

Οι πίνακες που χρησιμοποιούνται για την παροχή των προσωπικών υπηρεσιών είναι οι ακόλουθοι:

diana_users: Αποθηκεύει τους εγγεγραμμένους χρήστες στο σύστημα.

diana_history: Αποθηκεύει το ιστορικό των αναζητήσεων που πραγματοποιούν οι χρήστες του συστήματος.

diana_bookmarks: Αποθηκεύει τις αγαπημένες αναζητήσεις των χρηστών του συστήματος. Κάθε αγαπημένη αναζήτηση έχει μία περιγραφή και τη δυνατότητα περιορισμού των εμφανιζόμενων αποτελεσμάτων.

5

Υλοποίηση

Στο παρόν κεφάλαιο παρουσιάζονται θέματα σχετικά με την υλοποίηση του συστήματος της διπλωματικής και αναλύονται ζητήματα που χρήζουν περαιτέρω σχολιασμού.

5.1 Επιλογή μηχανής αποθήκευσης

Το σύστημα σχεσιακής βάσης δεδομένων MySQL έρχεται εξοπλισμένο με μία ποικιλία από μηχανές αποθήκευσης των δεδομένων. Κάθε μηχανή αποθήκευσης έχει συγκεκριμένα προτερήματα αλλά και αδυναμίες. Ο διαχειριστής της βάσης δεδομένων καλείται, συνήθως, να επιλέξει την κατάλληλη μηχανή λαμβάνοντας υπόψη τις ανάγκες της εφαρμογής ως προς την πρόσβαση στα δεδομένα. Συνοπτικά, η MySQL υποστηρίζει τις ακόλουθες μηχανές αποθήκευσης δεδομένων:

- **MyISAM:** Είναι η προκαθορισμένη μηχανή αποθήκευσης και είναι αυτή που χρησιμοποιείται περισσότερο σε εφαρμογές του παγκόσμιου ιστού.
- **InnoDB:** Είναι μία μηχανή αποθήκευσης που υποστηρίζει τις ιδιότητες ACID (atomicity, consistency, isolation, durability) και επομένως διασφαλίζει την αξιόπιστη διαχείριση των transactions. Επιπλέον, υποστηρίζει τις λειτουργίες commit και rollback ενώ έχει την ικανότητα να εξασφαλίσει την ακεραιότητα των δεδομένων ακόμη και μετά από απρόσμενο τερματισμό της βάσης δεδομένων. Τέλος, υποστηρίζει την τοποθέτηση ξένων κλειδιών μεταξύ των πινάκων της βάσης δεδομένων.

- **Memory:** Αποθηκεύει όλα τα δεδομένα στη μνήμη RAM του συστήματος για πολύ γρήγορη προσπέλαση σε αυτά.
- **Merge:** Επιτρέπει την προσπέλαση ενός αριθμού από MyISAM πίνακες ως ένα αντικείμενο. Είναι χρήσιμη σε περιβάλλοντα διαχείρισης μεγάλου όγκου δεδομένων.
- **Archive:** Επιτρέπει την αποδοτική πρόσβαση σε μεγάλο όγκο από δεδομένα που όμως προσπελάζονται με εξαιρετικά μικρή συχνότητα. Συνιστάται για την αποθήκευση ιστορικών δεδομένων ή δεδομένων από καταγραφές ασφαλείας.
- **Federated:** Επιτρέπει τη λογική σύνδεση πολλαπλών MySQL διακομιστών. Είναι χρήσιμη σε περιβάλλοντα κατακευματισμένων βάσεων δεδομένων.
- **CSV:** Επιτρέπει την αποθήκευση των δεδομένων της βάσης σε απλά αρχεία κειμένου διαχωρισμένων με ερωτηματικό. Συνιστάται όταν είναι απαραίτητη η μεταφορά των δεδομένων από και προς εφαρμογές που διαχειρίζονται τέτοιου είδους αρχεία.
- **Blackhole:** Η συγκεκριμένη μηχανή αποθήκευσης δέχεται δεδομένα χωρίς να τα αποθηκεύει. Η ανάκτηση δεδομένων από αυτήν επιστρέφει πάντα το κενό σύνολο. Χρησιμοποιείται σπάνια και μόνο σε ειδικές περιπτώσεις.

Από τις παραπάνω μηχανές αποθήκευσης το μεγαλύτερο ενδιαφέρον για χρήση στο σύστημα της διπλωματικής εργασίας παρουσίασαν η MyISAM και η InnoDB. Στην αρχική σχεδίαση της εφαρμογής είχε επιλεγεί η χρήση της μηχανής αποθήκευσης InnoDB κυρίως εξαιτίας της δυνατότητας που παρείχε για χρήση ξένων κλειδιών στους πίνακες. Με αυτόν τον τρόπο εξασφαλιζόταν η εγκυρότητα των δεδομένων της εφαρμογής. Παρόλα αυτά στις πρώτες μετρήσεις επιδόσεων πριν την τελική απόφαση παρατηρήθηκε διαφορά υπέρ της μηχανής αποθήκευσης MyISAM. Είναι γνωστό, άλλωστε, πως η μηχανή αποθήκευσης InnoDB υστερεί σε σχέση με τη μηχανή MyISAM ως αναφορά την ταχύτητα ανάγνωσης των δεδομένων. Το γεγονός αυτό σε συνδυασμό με τη δυνατότητα της μηχανής αποθήκευσης MyISAM να υποστηρίζει full text indexes οδήγησε στην τελική επιλογή αυτής για χρήση από τη βάση δεδομένων του συστήματος. Η χρήση των full text indexes ήταν βασική απαίτηση έτσι ώστε να βελτιωθεί η απόδοση ορισμένων ερωτημάτων προς τη βάση. Επιπλέον, από την ανάλυση απαιτήσεων για το σχεδιασμό του συστήματος της παρούσας διπλωματικής εργασίας είχε διασφαλιστεί ότι τα δεδομένα προς αποθήκευση στη βάση δεδομένων δεν πρόκειται να αλλάζουν συχνά και, επομένως, ο έλεγχος εγκυρότητας που εξασφαλίζεται από την ύπαρξη ξένων κλειδιών της μηχανής InnoDB μπορεί να περιοριστεί, χωρίς κίνδυνο, στην αρχική φόρτωση της βάσης δεδομένων. Επομένως, τα δεδομένα προς εισαγωγή στο σύστημα τοποθετούνται αρχικά σε πίνακες InnoDB ούτως ώστε να εξασφαλιστεί η εγκυρότητά τους και στη συνέχεια μεταφέρονται σε πίνακες MyISAM ώστε να βελτιωθεί η απόδοση του συστήματος.

5.2 Βελτιστοποίηση των ερωτημάτων στη βάση δεδομένων

Με βάση την ανάλυση απαιτήσεων του συστήματος γίνεται σαφές πως υπάρχει η ανάγκη το σύστημα να αποκρίνεται στις εντολές του χρήστη μέσα σε ένα εύλογο χρονικό διάστημα. Σε αντίθετη περίπτωση, η μεγάλη χρονική καθυστέρηση στην εύρεση και παρουσίαση των αποτελεσμάτων στο χρήστη δημιουργεί άσχημη εντύπωση και πιθανόν είναι σημάδι λανθασμένης σχεδίασης και υλοποίησης του συστήματος. Σε αυτό το πλαίσιο κρίθηκε αναγκαία η βελτιστοποίηση των SQL ερωτημάτων που εκτελούνται στη σχεσιακή βάση δεδομένων για κάθε αναζήτηση του χρήστη. Τα συγκεκριμένα ερωτήματα δημιουργούνται από το Yii PHP framework με χρήση της μηχανής Object Relational Mapping που διαθέτει. Ορισμένες φορές τα παραγόμενα ερωτήματα θυσιάζουν την απόδοση για χάρη της γενικότητας. Για το λόγο αυτό πολλές φορές κρίνεται σκόπιμη η χειροκίνητη βελτιστοποίηση των ερωτημάτων.

Το πρώτο βήμα για τη βελτιστοποίηση των SQL ερωτημάτων του συστήματος είναι η καταγραφή τους. Η συγκεκριμένη διαδικασία κατέστη δυνατή χάρη στο σύστημα καταγραφής που προσφέρει το Yii PHP framework και το οποίο αποθηκεύει σε κυλιόμενα αρχεία καταγραφής όλα τα ερωτήματα που εκτελούνται στη σχεσιακή βάση δεδομένων. Με κατάλληλη χρήση εργαλείων επεξεργασίας αρχείων κειμένου δημιουργήθηκε μία λίστα που περιλάμβανε το σύνολο των SQL ερωτημάτων.

Στη συνέχεια εντοπίστηκαν με προσεκτική μελέτη της λίστας εκείνα τα ερωτήματα που ήταν πιθανότερο να απαιτούν και τον περισσότερο χρόνο επεξεργασίας. Πρόκειται κυρίως για ερωτήματα που περιλαμβάνουν joins πινάκων.

Εντύπωση προκάλεσε το γεγονός ότι το Yii PHP framework επέλεγε ως τρόπο σύνδεσης των πινάκων το left outer join. Το αποτέλεσμα ενός left outer join μεταξύ δύο πινάκων, έστω A left outer join B, έχει τα εξής χαρακτηριστικά:

- Για κάθε συνδυασμό γραμμών των πινάκων A και B για τον οποίο ισχύει η συνθήκη του συνδέσμου, προστίθεται στο τελικό αποτέλεσμα μία γραμμή που περιλαμβάνει τιμές από τις δύο γραμμές των πινάκων A και B που ταιριάζουν. Η συμπεριφορά αυτή είναι ταυτόσημη με τη συμπεριφορά του inner join.
- Επιπλέον, για κάθε γραμμή του πίνακα A για την οποία δεν υπάρχει γραμμή στον πίνακα B που να ταιριάζει μαζί της, προστίθεται στο αποτέλεσμα μία γραμμή που περιλαμβάνει τις τιμές της γραμμής του πίνακα A και την τιμή NULL στις θέσεις που έπρεπε να καταλαμβάνουν οι τιμές του πίνακα B. Με άλλα λόγια, το left outer join μεταξύ των πινάκων A και B περιλαμβάνει υποχρεωτικά όλες τις εγγραφές του πίνακα A (αριστερός πίνακας στο σύνδεσμο).

Για την καλύτερη κατανόηση της παραπάνω λειτουργίας δίνεται το ακόλουθο παράδειγμα εφαρμογής ενός συνδέσμου left outer join:

Έστω δύο πίνακες με τα εξής δεδομένα:

LastName	DepartmentID
Rafferty	31
Jones	33
Steinberg	33
Robinson	34
Smith	34
John	999

Πίνακας 1: Πίνακας Employee.

DepartmentID	DepartmentName
31	Sales
33	Engineering
34	Clerical
35	Marketing

Πίνακας 2: Πίνακας Department.

Στους συγκεκριμένους πίνακες εκτελείται το ακόλουθο ερώτημα:

```
SELECT *
FROM employee e LEFT OUTER JOIN department d
ON e.DepartmentID = d.DepartmentID
```

Ερώτημα 1: Ερώτημα συνδέσμου left outer join.

Η εκτέλεση του παραπάνω ερωτήματος εφαρμόζει ένα σύνδεσμο left outer join μεταξύ των δύο πινάκων. Το τελικό αποτέλεσμα είναι ο ακόλουθος πίνακας:

e.LastName	e.DepartmentID	d.DepartmentName	d.DepartmentID
Jones	33	Engineering	33
Rafferty	31	Sales	31
Robinson	34	Clerical	34
Smith	34	Clerical	34
John	999	NULL	NULL
Steinberg	33	Engineering	33

Πίνακας 3: Το αποτέλεσμα της εφαρμογής ενός συνδέσμου left outer join στους πίνακες Employee και Department στο κοινό πεδίο τους.

Παρατηρείται το γεγονός ότι η εγγραφή John του πίνακα Employee, για την οποία δεν υπάρχει αντίστοιχη εγγραφή στον πίνακα Department, εμφανίζεται στο αποτέλεσμα του συνδέσμου left outer join έχοντας στα πεδία που προέρχονται από τον πίνακα Department κενές τιμές.

Η χρήση του left outer join δεν ήταν αναγκαία για τη σωστή και απρόσκοπτη εκτέλεση των ερωτημάτων, επομένως κρίθηκε σκόπιμη η αλλαγή των συνδέσμων στην κλασική τεχνική του inner join. Όπως αποδείχτηκε με πειραματική διαδικασία, η συγκεκριμένη αλλαγή των συνδέσμων βελτίωσε κατά ένα μεγάλο ποσοστό την ταχύτητα και αποτελεσματικότητα των εκτελούμενων SQL ερωτημάτων.

Μελετώντας εκ των υστέρων την περιγραφή εκτέλεσης των ερωτημάτων όπως δίνεται από τη μηχανή εκτέλεσης της σχεσιακής βάσης δεδομένων MySQL προέκυψε ότι η μετατροπή των συνδέσμων κατά το συγκεκριμένο τρόπο οδήγησε σε ανακατάταξη της σειράς ανάγνωσης

των πινάκων. Συγκεκριμένα, η αφαίρεση της απαίτησης που εισήγαγε το left outer join να περιλαμβάνονται όλες ανεξαιρέτως οι εγγραφές του αριστερού πίνακα διευκόλυνε τη μηχανή SQL ερωτημάτων να εφαρμόσει επιπλέον βελτιστοποιήσεις στον τρόπο εκτέλεσης των ερωτημάτων, οι οποίες βελτιστοποιήσεις δεν εφαρμόζονταν μέχρι εκείνη τη στιγμή λόγω της ύπαρξης του left outer join.

5.3 Full text search

Στο πλαίσιο της βελτίωσης της απόδοσης του συστήματος κομβικό σημείο κατείχε η βελτιστοποίηση της αναζήτησης με βάση την περιγραφή ενός γονιδίου και η αναζήτηση με βάση το όνομα ενός KEGG pathway. Και στις δύο περιπτώσεις τα αντίστοιχα πεδία στα οποία γίνεται η αναζήτηση περιέχουν ελεύθερο κείμενο. Ενδεικτικά, η περιγραφή ενός γονιδίου είναι “t-cell receptor delta d gene segment” και το όνομα ενός KEGG pathway είναι “fatty acid elongation in mitochondria”, οπότε ο χρήστης του συστήματος επιθυμεί να αναζητήσει τις συγκεκριμένες εγγραφές εισάγοντας μία από τις περισσότερες λέξεις που περιλαμβάνονται σε αυτές. Στις περιπτώσεις όπου είναι επιθυμητή η αναζήτηση σε πεδία τέτοιου τύπου ενδείκνυται η χρήση ειδικών ευρετηρίων για αναζήτηση κειμένου. Τα συγκεκριμένα ευρετήρια αναλύουν τις λέξεις που εντοπίζουν στα υπό εξέταση πεδία και τις αποθηκεύουν με τέτοιο τρόπο ώστε, πλέον, η αναζήτηση με βάση κάποιες λέξεις να πραγματοποιείται ταχύτερα. Ένα ακόμη πλεονέκτημα από την χρήση των συγκεκριμένων ευρετηρίων είναι πως παρέχουν σε κάθε αναζήτηση που πραγματοποιείται και ένα βαθμό σχετικότητας για τα αποτελέσματα που επιστρέφονται. Ο βαθμός σχετικότητας υπολογίζεται με βάση τις λέξεις κλειδιά που αναζητεί ο χρήστης και τη συχνότητα εμφάνισής τους στο σώμα της αναζήτησης.

Μερικά χαρακτηριστικά από την αναζήτηση πλήρους κειμένου όπως υλοποιείται στη σχεσιακή βάση δεδομένων MySQL είναι τα εξής:

- Υπάρχει ελάχιστο όριο στο μήκος των λέξεων που μπορούν να αναζητηθούν. Συγκεκριμένα, αν μία λέξη που αναζητείται είναι κάτω του ορίου τότε θεωρείται πως δε συνεισφέρει στην αναζήτηση και αγνοείται.
- Υπάρχει μία λίστα από κοινές λέξεις της αγγλικής γλώσσας, οι οποίες εφόσον αναζητηθούν αγνοούνται διότι δεν έχουν σημαντική συνεισφορά στην εύρεση αποτελεσμάτων.

5.4 Ευρετήρια

Το ευρετήριο είναι μία δομή δεδομένων που χρησιμοποιείται ευρύτατα στις βάσεις δεδομένων με σκοπό τη βελτίωση της ταχύτητας πρόσβασης στα δεδομένα των πινάκων. Η αποθήκευση των συγκεκριμένων δομών δεδομένων συνήθως απαιτεί πολύ λιγότερο χώρο από τα

ίδια τα δεδομένα με αποτέλεσμα να είναι οικονομική η τοποθέτησή τους στην κύρια μνήμη του συστήματος αυξάνοντας με αυτό τον τρόπο περαιτέρω την αποδοτικότητά τους. Ένα ευρετήριο μπορεί να χρησιμοποιηθεί για την εύρεση συγκεκριμένων εγγραφών με τυχαία προσπέλαση ή και ακόμα την εύρεση ενός εύρους ταξινομημένων εγγραφών. Κατά την κατασκευή του ευρετηρίου ο διαχειριστής της βάσης δεδομένων επιλέγει μία ή περισσότερες στήλες του αντίστοιχου πίνακα για τις οποίες μπορεί να χρησιμοποιηθεί το υπό κατασκευή ευρετήριο.

Η χρησιμότητα των ευρετηρίων ελαττώνεται όταν η βάση δεδομένων περιέχει δεδομένα που είτε αλλάζουν είτε προστίθενται νέα με συχνό ρυθμό. Αυτό συμβαίνει διότι οι μεταβολές και οι προσθήκες γραμμών σε ένα πίνακα με ευρετήρια είναι υπολογιστικά ακριβότερες από τις αντίστοιχες ενέργειες σε ένα πίνακα δίχως ευρετήρια, μιας και για κάθε μεταβολή των γραμμών του πίνακα οφείλουν να ενημερώνονται και τα αντίστοιχα ευρετήρια.

Το σύστημα σχεσιακών βάσεων δεδομένων MySQL υποστηρίζει τρία διαφορετικά είδη ευρετηρίων. Πρόκειται για ευρετήρια B-tree, ευρετήρια R-tree και ευρετήρια hash. Τα τελευταία υποστηρίζονται μόνο από τη μηχανή αποθήκευσης Memory, δηλαδή υποστηρίζονται μόνο για πίνακες που αποθηκεύονται στη μνήμη RAM του συστήματος επομένως δε μπορούν να εφαρμοστούν στην περίπτωση του συστήματος της παρούσας διπλωματικής εργασίας. Τα ευρετήρια R-tree χρησιμοποιούνται κατά κόρον για την ευρετηριοποίηση γεωγραφικών δεδομένων μιας και έχουν την ικανότητα να επιταχύνουν ερωτήματα της μορφής: «βρες όλα τα σημεία ενδιαφέροντος που βρίσκονται σε απόσταση μικρότερη των δύο χιλιομέτρων από μία δεδομένη θέση». Τα ευρετήρια B-tree είναι μία γενίκευση των δυαδικών δέντρων αναζήτησης που προσφέρουν βελτιστοποιημένη αναζήτηση, εισαγωγή, διαγραφή και σειριακή προσπέλαση των δεδομένων. Είναι ιδανικά για χρήση από συστήματα που διαβάζουν και γράφουν μεγάλους όγκους δεδομένων, όπως είναι οι βάσεις δεδομένων και τα συστήματα αρχείων.

Στο σύστημα που αναπτύχθηκε για την παρούσα διπλωματική, η βάση δεδομένων χρησιμοποιήθηκε κατά κύριο λόγο για την ανάγνωση δεδομένων. Γι' αυτό λόγο κρίθηκε απαραίτητη η σωστή και προσεκτική υλοποίηση μίας σειράς από ευρετήρια στους πίνακες της βάσης δεδομένων του συστήματος. Η τακτική που χρησιμοποιήθηκε για την επιλογή των πινάκων και των συγκεκριμένων στηλών αυτών που επρόκειτο να τεθούν υπό ευρετηριοποίηση ήταν η ακόλουθη:

- Ευρετήρια δημιουργήθηκαν στις στήλες των πινάκων που χρησιμοποιούνται στην εκτέλεση των προτάσεων where του συνόλου των ερωτημάτων SQL προς τη βάση δεδομένων.
- Επίσης, ευρετήρια δημιουργήθηκαν στις στήλες που χρησιμοποιούνται στις ενώσεις (join) μεταξύ πινάκων. Συνήθως, οι ενώσεις μεταξύ πινάκων πραγματοποιούνται με-

ταξύ του κύριου κλειδιού (primary key) ενός πίνακα και του ξένου κλειδιού (foreign key) ενός δεύτερου πίνακα. Το κύριο κλειδί στις βάσεις δεδομένων είναι εξ ορισμού ευρετήριο. Δεν ισχύει το ίδιο και για τα ξένα κλειδιά, όπου είναι ευθύνη του διαχειριστή της βάσης δεδομένων να επιλέξει αν θα τα ορίσει ως ευρετήρια ή όχι.

- Ιδιαίτερη προσοχή χρειάστηκε στην κατασκευή ευρετηρίων που περιλαμβάνουν περισσότερες από μία στήλες ενός πίνακα. Στα συγκεκριμένα ευρετήρια καθοριστικό ρόλο στην απόδοση διαδραματίζει η σειρά με την οποία τοποθετούνται στο ευρετήριο οι επιλεγμένες στήλες. Επιπλέον, η σωστή τοποθέτηση των στηλών στο ευρετήριο δύναται να επιτρέψει τη χρήση του ευρετηρίου τόσο για συνδυαστικές αναζητήσεις πάνω σε όλες τις στήλες του ευρετηρίου όσο και σε μεμονωμένες αναζητήσεις πάνω σε μία από τις στήλες αυτού. Με αυτόν τον τρόπο, γίνεται οικονομία χώρου έχοντας ένα ευρετήριο, το οποίο μπορεί να χρησιμοποιηθεί με πολλαπλούς τρόπους.

Από τα διαθέσιμα είδη ευρετηρίων που παρέχει το σύστημα διαχείρισης σχεσιακών βάσεων δεδομένων MySQL, τα B-trees αποτελούν την πιο διαδεδομένη λύση έχοντας εφαρμογή στις περισσότερες των περιπτώσεων. Το σύστημα της παρούσας διπλωματικής εργασίας χρησιμοποιεί B-trees για την υλοποίηση όλων των απαραίτητων ευρετηρίων με εξαίρεση τα ευρετήρια full text, τα οποία υλοποιούνται με το δικό τους ξεχωριστό μηχανισμό.

5.5 Μηχανή προτάσεων για τις αναζητήσεις των χρηστών του συστήματος

Πολλές φορές συμβαίνει ο χρήστης του συστήματος να δίνει στην αναζήτησή του ένα όρο (είτε microRNA είτε γονίδιο) που δεν ταιριάζει απόλυτα με κάποιο υπάρχον όνομα στη βάση δεδομένων. Σε αυτή την περίπτωση είναι επιθυμητή η παροχή από το σύστημα προτάσεων προς το χρήστη ούτως ώστε να ολοκληρώσει την αναζήτησή του με επιτυχία. Οι προτάσεις αυτές οφείλουν να είναι σχετικές και παραπλήσιες με τον αρχικό όρο αναζήτησης. Για την πληρέστερη κατανόηση της συγκεκριμένης λειτουργίας παρατίθεται το ακόλουθο παράδειγμα:

Έστω ότι υπάρχουν τα εξής ονόματα microRNA στον αντίστοιχο πίνακα της βάσης δεδομένων του συστήματος:

hsa-let-7a
hsa-let-7b
hsa-let-7c
mmu-let-7a
mmu-let-7b
mmu-let-7c

Πίνακας 4: Ονόματα microRNA.

Εφόσον ο χρήστης πραγματοποιήσει μία αναζήτηση με βάση το microRNA δίνοντας ως όρο “let-7a”, θα λάβει μία ταξινομημένη λίστα με προτάσεις που έχει ως εξής:

hsa-let-7a
mmu-let-7a
hsa-let-7b
mmu-let-7b
hsa-let-7c
mmu-let-7c

Πίνακας 5: Προτάσεις για αναζήτηση με βάση τον όρο “let-7a”.

Η ταξινόμηση των προτάσεων βασίζεται στο κριτήριο απόστασης Levenshtein, έτσι ώστε οι πρώτες προτάσεις να έχουν και τη μεγαλύτερη συνάφεια με τον αρχικό όρο αναζήτησης. Το συγκεκριμένο κριτήριο ορίζεται ως ο ελάχιστος αριθμός αλλαγών για το μετασχηματισμό ενός αρχικού κειμένου σε ένα τελικό. Μετασχηματισμός ορίζεται η προσθήκη, η αφαίρεση ή / και η αντικατάσταση ενός χαρακτήρα. Το κριτήριο Levenshtein χρησιμοποιείται από τα συστήματα λογισμικού για να αντιμετωπίσει τυχόν λάθη στην πληκτρολόγηση από τους χρήστες. Για την υποστήριξη της δυνατότητας παροχής προτάσεων για τις αναζητήσεις αναπτύχθηκε μία μηχανή δημιουργίας προτάσεων με βάση τους όρους αναζήτησης του χρήστη. Η υλοποίηση της συγκεκριμένης μηχανής βρίσκεται στην κλάση Suggestions του συστήματος. Ακολουθεί η περιγραφή της υλοποίησης:

Μέθοδος suggest

Παράμετροι

- **\$term:** Εδώ δίνεται ο όρος της αναζήτησης του χρήστη.
- **\$suggest_specs:** Είναι ένας πίνακας από προδιαγραφές για τη κατασκευή των προτάσεων. Κάθε στοιχείο του πίνακα έχει την εξής δομή:

```
Array('prefix'=>..., 'field'=>..., 'sql'=>..., 'fmt'=>..., 'args'=>...)
```

Το στοιχείο sql είναι το ερώτημα που εκτελείται στη βάση για την άντληση όλων των στοιχείων πάνω στα οποία θα γίνει προσπάθεια να ταιριάζει η παράμετρος \$term. Πρέπει να υπάρχει μία στήλη ονομασμένη id στο συγκεκριμένο ερώτημα.

Το στοιχείο field είναι το όνομα της στήλης του προηγούμενου ερωτήματος με τα δεδομένα της οποίας θα ταιριάζει η παράμετρος \$term.

Το στοιχείο fmt είναι το κείμενο που εμφανίζεται στο χρήστη ως πρόταση.

Το στοιχείο args είναι ένας πίνακας που χρησιμοποιείται για τη μορφοποίηση του κειμένου από τη συνάρτηση της PHP vsprintf.

Το στοιχείο prefix είναι ένα μοναδικό πρόθεμα που συνδυάζεται με τη στήλη id του ερωτήματος ούτως ώστε να δημιουργηθεί μία μοναδική ταυτότητα για την παρεχόμενη από το σύστημα πρόταση.

Επιστρεφόμενη τιμή

Ένας πίνακας με τις προτάσεις. Κάθε στοιχείο του πίνακα έχει την εξής δομή:

```
array('id'=>..., 'score'=>..., 'label'=>...)
```

Το στοιχείο id είναι το μοναδικό αναγνωριστικό της πρότασης.

Το στοιχείο score είναι ο βαθμός συνάφειας της πρότασης με τον όρο αναζήτησης του χρήστη.

Το στοιχείο label είναι η ετικέτα της πρότασης που εμφανίζεται στο χρήστη.

Αλγόριθμος

- Ελέγχεται αν το \$term εμφανίζεται ακριβώς σε κάποιο πεδίο field από τις προδιαγραφές. Αν ναι, τότε δε χρειάζεται να δοθούν περαιτέρω προτάσεις και επιστρέφεται άμεσα το στοιχείο που βρέθηκε.
- Σε αντίθετη περίπτωση επιστρέφεται ένας πίνακας προτάσεων.

Μέθοδος uniqueSuggestion

Παράμετροι

Όπως και στη μέθοδο suggest.

Επιστρεφόμενη τιμή

Όπως και στη μέθοδο suggest. Επιπλέον, η τιμή false αν υπάρχουν πολλαπλές προτάσεις.

Αλγόριθμος

- Για κάθε προδιαγραφή, εκτελείται το ερώτημα SQL προσθέτοντας μία πρόταση where για τον ακριβή εντοπισμό του \$term.
- Αν επιστραφεί μία μόνο γραμμή, τότε η μέθοδος επιστρέφει με επιτυχία το στοιχείο που βρέθηκε.

Μέθοδος multipleSuggestions

Παράμετροι

Όπως και στη μέθοδο suggest.

Επιστρεφόμενη τιμή

Όπως και στη μέθοδο suggest.

Αλγόριθμος

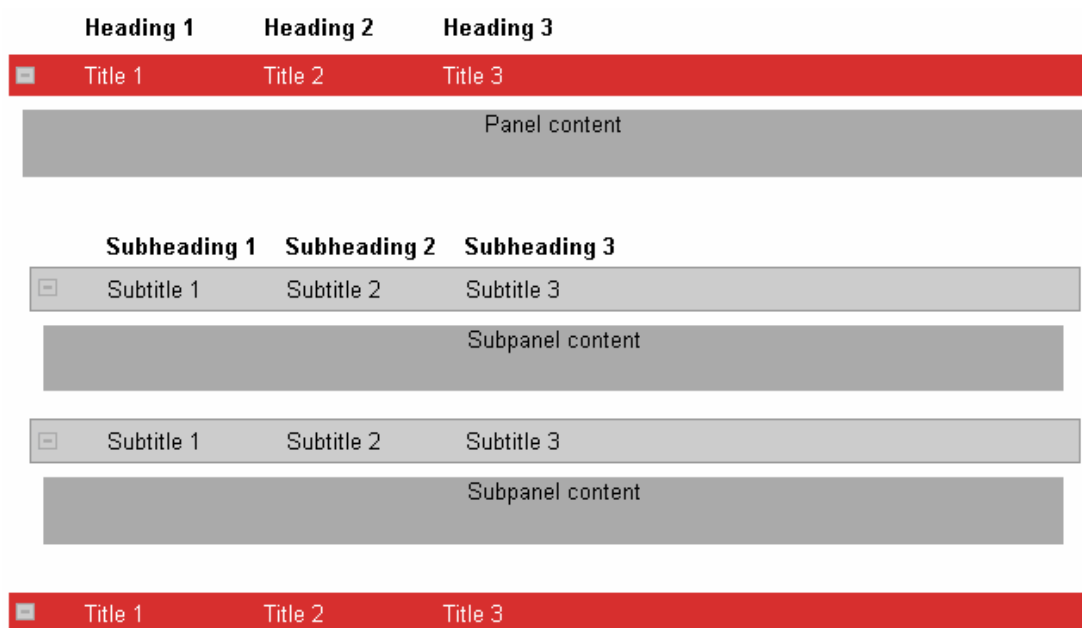
- Για κάθε προδιαγραφή, εκτελείται το ερώτημα SQL.
- Για κάθε γραμμή από τα αποτελέσματα του ερωτήματος, υπολογίζεται ο βαθμός συνάφειας με το \$term. Αυτό γίνεται είτε διαβάζοντας τη στήλη score (αν υπάρχει) της

γραμμής, είτε υπολογίζοντας την απόσταση Levenshtein μεταξύ της στήλης field και του \$term.

- Στη συνέχεια κατασκευάζεται το μήνυμα της πρότασης από τα πεδία fmt και args.
- Αφού επεξεργαστούν όλες οι προδιαγραφές, ταξινομούνται οι προτάσεις με το βαθμό συνάφειας και επιστρέφονται οι τριάντα με τον υψηλότερο βαθμό.

5.6 Επαναχρησιμοποιήσιμο στοιχείο λίστας με λεπτομέρειες

Ένα από τα μοναδικά χαρακτηριστικά της εφαρμογής είναι η χρήση μίας εξελιγμένης λίστας για την παρουσίαση των αποτελεσμάτων. Η συγκεκριμένη λίστα παρέχει τη δυνατότητα αρχικά συνοπτική προβολής των αποτελεσμάτων ανά βιολογικό στόχο. Όταν ο χρήστης, το επιθυμεί μπορεί να προβάλει με ένα κλικ περισσότερες λεπτομέρειες για το στόχο της επιλογής του. Για κάθε στόχο πέρα από τις γενικές πληροφορίες παρέχονται συνοπτικές πληροφορίες για τα σημεία πρόσδεσης πάνω στο βιολογικό στόχο. Ο χρήστης έχει εκ νέου τη δυνατότητα να προβάλει περισσότερες πληροφορίες για κάποιο από αυτά κάνοντας ένα απλό κλικ. Η φύση της λίστας αυτής που αξιοποιεί εις διπλούν ένα αναδιδόμενο interface επικοινωνίας με το χρήστη αυξάνει τη χρηστικότητα και την αποτελεσματικότητα της εφαρμογής.



Εικόνα 25: Παράδειγμα του στοιχείου λίστας. Εμφανίζονται δύο panels, το πρώτο αναπτυγμένο και το δεύτερο συνεπτυγμένο. Το πρώτο panel περιέχει δύο αναπτυγμένα sub panels. Διακρίνονται οι επικεφαλίδες, οι τίτλοι και οι χώροι για τα περιεχόμενα των panels και των sub panels.

Τα παραπάνω χαρακτηριστικά της λίστας οδήγησαν στη δημιουργία ενός ανεξάρτητου και ευέλικτου συστατικού στοιχείου που σκοπό έχει να καταστήσει απλή στον προγραμματιστή

τη χρήση της λίστας στις εφαρμογές που επιθυμεί. Η υλοποίηση της λίστας βρίσκεται στο αρχείο DrillDownList.php του συστήματος. Ακολουθεί η περιγραφή της υλοποίησης:

Κλάση DrillDownList

Περιγραφή

Περιγράφει και διαχειρίζεται τη λίστα.

Μέθοδοι

- init
- run

Μέθοδος init

Αλγόριθμος

- Δημιουργεί ένα root panel το οποίο στην ουσία είναι ο περιέκτης όλης της λίστας. Δεν έχει δυνατότητα απόκρυψης ούτε εμφανίζεται πλάι του checkbox.
- Φορτώνει στη σελίδα που χρησιμοποιεί τη λίστα τα απαραίτητα JavaScript και Cascading Style Sheets αρχεία.

Μέθοδος run

Αλγόριθμος

- Σχεδιάζει το root panel στη σελίδα HTML που ετοιμάζεται για εμφάνιση στον browser του χρήστη.

Κλάση Panel

Περιγραφή

Περιγράφει ένα panel της λίστας. Το panel έχει επικεφαλίδα, περιεχόμενο, εσωτερικά panels, μπορεί να έχει τη δυνατότητα απόκρυψης και μπορεί να έχει ενσωματωμένο checkbox για να είναι επιλέξιμο. Το ιδιαίτερο χαρακτηριστικό των panel είναι η δυνατότητα εμφωλευμένης τοποθέτησής τους.

Κλάση PanelHeader

Περιγραφή

Περιγράφει μία επικεφαλίδα ενός panel. Έχει την επιπρόσθετη δυνατότητα να διαχειρίζεται το πλάτος των στηλών. Σε περίπτωση που το συνολικό πλάτος των στηλών ξεπερνάει το μέ-

γιστο δυνατό, τότε εμφανίζεται μήνυμα λάθους στη σελίδα ώστε ο προγραμματιστής να μερμνήσει κατάλληλα.

5.7 Πλατφόρμες και προγραμματιστικά εργαλεία

Στην παρούσα ενότητα αναφέρονται τα προγραμματιστικά εργαλεία και οι πλατφόρμες ανάπτυξης λογισμικού που χρησιμοποιήθηκαν κατά τη διάρκεια της κατασκευής του συστήματος της διπλωματικής εργασίας. Η εφαρμογή που αναπτύχθηκε αποτελεί μία δικτυακή εφαρμογή, επομένως και τα εργαλεία που παρουσιάζονται περιστρέφονται γύρω από την παροχή υπηρεσιών στον παγκόσμιο ιστό.

5.7.1 Apache HTTP Server

Ο Apache HTTP Server είναι ένας διακομιστής ιστοσελίδων ο οποίος διαδραμάτισε προεξέχοντα ρόλο στην αρχική ανάπτυξη του παγκόσμιου ιστού. Μέχρι σήμερα, είναι ο πιο διαδεδομένος διακομιστής με περισσότερα από 100 εκατομμύρια ιστοσελίδες να βασίζονται τη λειτουργία τους σε αυτόν. Τα χαρακτηριστικά που διαθέτει του επιτρέπουν να ανταγωνίζεται τους υπόλοιπους διαθέσιμους διακομιστές τόσο σε όρους παρεχόμενων λειτουργιών όσο και σε απόδοση. Ο Apache διατίθεται για μία μεγάλη πληθώρα λειτουργικών συστημάτων, όπως Windows, Linux, FreeBSD, κλπ.. Παρόλα αυτά στις περισσότερες των περιπτώσεων προτιμάται η λειτουργία του σε συνδυασμό με το λειτουργικό σύστημα Linux. Αποτελεί ανοικτό λογισμικό υπό την άδεια Apache License και η ανάπτυξη και διαχείρισή του πραγματοποιείται από μία ανοικτή κοινότητα προγραμματιστών υπό την αιγίδα του Apache Software Foundation.

Ο Apache HTTP Server υποστηρίζει ένα ευρύ σύνολο χαρακτηριστικών, πολλά από τα οποία υλοποιούνται ως ξεχωριστά modules τα οποία και επεκτείνουν το βασικό σύνολο λειτουργιών του διακομιστή. Ενδεικτικά αναφέρονται οι εξής δυνατότητες:

- Υποστήριξη γλωσσών προγραμματισμού στην πλευρά του διακομιστή, όπως είναι PHP, Perl, Python, κλπ.
- Δυνατότητα ταυτοποίησης χρηστών μέσω διάφορων διαδεδομένων συστημάτων ταυτοποίησης.
- Υποστήριξη ασφαλούς μεταφοράς της πληροφορίας μέσω του πρωτοκόλλου TLS.
- Συμπίεση των αποστελλόμενων δεδομένων για βελτίωση των επιδόσεων του διακομιστή.

Αν και ο βασικός σχεδιαστικός στόχος του Apache δεν ήταν η επίτευξη των μέγιστων επιδόσεων, παρόλα αυτά εξακολουθεί να είναι ένας από τους πιο γρήγορους διακομιστές που κυκλοφορούν σήμερα. Έχει τη δυνατότητα μέσω του αντίστοιχου module να υποστηρίξει δια-

φορετικές αρχιτεκτονικές υπολογιστικών συστημάτων με αποτέλεσμα να εκμεταλλεύεται με το μέγιστο δυνατό τρόπο την εκάστοτε υπολογιστική υποδομή.

Έχοντας υπόψη το σύνολο των παραπάνω αναφερόμενων πληροφοριών για τον Apache HTTP Server πραγματοποιήθηκε η επιλογή του ως διακομιστή για τη δικτυακή εφαρμογή που αναπτύχθηκε στα πλαίσια της παρούσας διπλωματικής εργασίας.

5.7.2 *MySQL*

Η MySQL είναι ένα διαδεδомένο σύστημα σχεσιακής βάσης δεδομένων που αριθμεί σήμερα περισσότερες από 6 εκατομμύρια εγκαταστάσεις παγκοσμίως. Λειτουργεί ως server και, επομένως, μπορεί να εξυπηρετήσει την πρόσβαση πολλαπλών χρηστών σε πολλαπλές βάσεις δεδομένων ταυτόχρονα. Είναι ανοικτό λογισμικό μιας και διανέμεται υπό την άδεια GNU General Public License. Ανήκει στην κερδοσκοπική εταιρία MySQL AB, η οποία πλέον είναι θυγατρική εταιρία της Sun Microsystems, η οποία με τη σειρά της βρίσκεται στη διαδικασία απορρόφησης από την Oracle Corporation. Η MySQL χρησιμοποιείται πολύ συχνά σε έργα ανοικτού λογισμικού και, συνήθως, συνδυάζεται με το λειτουργικό σύστημα Linux και το διακομιστή ιστοσελίδων Apache HTTP Server. Ένας σημαντικός αριθμός από μεγάλης κλίμακας διαδικτυακών προϊόντων χρησιμοποιεί τη MySQL, όπως Wikipedia, Google και Facebook.

Η MySQL είναι σχεδιασμένη ώστε να υποστηρίζει ένα μεγάλο αριθμό από λειτουργικά συστήματα περιλαμβανομένων των Linux, Windows, NetBSD, OpenBSD, κλπ.. Επιπλέον, διατίθενται βιβλιοθήκες πρόσβασης στο σύστημα της βάσης δεδομένων για τις περισσότερες γνωστές γλώσσες προγραμματισμού. Τα χαρακτηριστικά που διαθέτει καλύπτουν ένα ευρύ φάσμα εφαρμογών. Ενδεικτικά παρατίθενται τα εξής:

- Υποστήριξη του ANSI SQL 99.
- Υποστήριξη για stored procedures, cursors και triggers.
- Υποστήριξη για updatable views.
- Παρέχει πολλαπλές ανεξάρτητες μηχανές αποθήκευσης δεδομένων, όπως MyISAM, InnoDB, κλπ.. Η κάθε μηχανή είναι βελτιστοποιημένη για συγκεκριμένη χρήση.
- Υποστήριξη Full-text indexing.

Για όλους τους παραπάνω λόγους επιλέχθηκε να χρησιμοποιηθεί η MySQL ως το σύστημα σχεσιακής βάσης δεδομένων το οποίο ανέλαβε το ρόλο να υποστηρίξει τη διαχείριση των δεδομένων της παρούσας εφαρμογής.

5.7.3 *PHP*

Η PHP είναι μία ευρέως διαδεδομένη scripting γλώσσα προγραμματισμού, η οποία ειδικεύεται και χρησιμοποιείται κατά κόρον στην ανάπτυξη εφαρμογών για τον παγκόσμιο ιστό. Είναι συμβατή με μία ευρεία συλλογή από λειτουργικά συστήματα και διακομιστές ιστοσελίδων με αποτέλεσμα να χρησιμοποιείται σε περισσότερες από 20 εκατομμύρια ιστοσελίδες παγκοσμίως. Αποτελεί ανοικτό λογισμικό και διανέμεται υπό την άδεια PHP License. Αξίζει να σημειωθεί ότι, αν και ο αρχικός στόχος της γλώσσας ήταν αποκλειστικά η υποστήριξη δημιουργίας δυναμικών ιστοσελίδων, σήμερα μπορεί να χρησιμοποιηθεί για τη συγγραφή ανεξάρτητων εφαρμογών είτε γραμμής εντολών είτε με γραφικό interface.

Από την έκδοση 4 και έπειτα, τα αρχεία πηγαίου κώδικα μεταγλωττίζονται αυτόματα σε bytecode και εκτελούνται από την αφηρημένη μηχανή Zend Engine με αποτέλεσμα αυξημένες επιδόσεις σε σχέσεις με τις προγενέστερες εκδόσεις. Παράλληλα, υπάρχουν λύσεις σε επίπεδο λογισμικού που αξιοποιούν ένα σύνολο από τεχνικές για να επιτύχουν ακόμη μεγαλύτερες επιδόσεις όταν αυτό κρίνεται αναγκαίο για την εύρυθμη εξυπηρέτηση του συνόλου των χρηστών της δικτυακής εφαρμογής.

Άλλο ένα σημαντικό χαρακτηριστικό της γλώσσας προγραμματισμού PHP είναι ότι από την έκδοση 5 και έπειτα υποστηρίζει πλήρως το αντικειμενοστραφές μοντέλο προγραμματισμού με αποτέλεσμα ο προγραμματιστής να έχει την ευχέρεια να αξιοποιήσει σύγχρονες τεχνικές ανάπτυξης λογισμικού για την συγγραφή της εκάστοτε εφαρμογής.

Στη βάση όλων των παραπάνω χαρακτηριστικών που διαθέτει η γλώσσα προγραμματισμού PHP επιλέχθηκε ως η γλώσσα ανάπτυξης του συστήματος της παρούσας διπλωματικής εργασίας. Επιπλέον, κρίθηκε ως θετικό στοιχείο η δυνατότητα άριστης συνεργασίας τόσο με τον Apache HTTP Server όσο και με το σύστημα σχεσιακής βάσης δεδομένων MySQL. Άλλος ένα παράγοντας που συντέλεσε στην απόφαση για χρήση της γλώσσας προγραμματισμού PHP ήταν η ύπαρξη μίας βάσης κώδικα, ο οποίος ήταν σε PHP, και υποστήριξε τη λειτουργία της προηγούμενης έκδοσης της εφαρμογής.

5.7.4 *PHPUnit*

Το PHPUnit είναι ένα σύστημα που παρέχει την υποδομή για πραγματοποίηση αυτοματοποιημένου ελέγχου ενός συστήματος λογισμικού. Συγκεκριμένα, υποστηρίζει τη μεθοδολογία ελέγχου λογισμικού unit testing. Η συγκεκριμένη μεθοδολογία είναι ευρέως διαδομένη και στενά συνδεδεμένη με διάφορες τεχνικές σύγχρονης ανάπτυξης λογισμικού και επιλέχθηκε για τον έλεγχο και την πιστοποίηση της ορθής λειτουργίας του συστήματος που αναπτύχθηκε στα πλαίσια της παρούσας διπλωματικής εργασίας.

Το RHPUnit ανήκει στην ευρύτερη οικογένεια συστημάτων ελέγχου λογισμικού xUnit. Αυτή η οικογένεια αποτελείται από συστήματα ελέγχου λογισμικού τα οποία οδηγούνται από κώδικα ειδικής μορφής τον οποίο οφείλει να παρέχει ο προγραμματιστής. Η αρχική σχεδίαση ενός τέτοιου συστήματος ήταν το σύστημα SUnit και είχε πραγματοποιηθεί για να υποστηρίξει τον έλεγχο προγραμμάτων γραμμένων στη γλώσσα προγραμματισμού Smalltalk. Έκτοτε παρόμοια συστήματα έχουν αναπτυχθεί για τις περισσότερες σύγχρονες γλώσσες προγραμματισμού.

Η λειτουργία του RHPUnit βασίζεται στις εξής έννοιες:

- **Test fixture:** είναι οι προϋποθέσεις ή οι συνθήκες κάτω από τις οποίες πρέπει να εκτελεστεί ένας δεδομένος έλεγχος ώστε να πετύχει.
- **Test suite:** είναι μία συλλογή από ξεχωριστούς ελέγχους που μοιράζονται το ίδιο test fixture. Η σειρά εκτέλεσης των ελέγχων δεν πρέπει να επηρεάζει το αποτέλεσμα του ελέγχου.
- **Assertion:** είναι ένα σημείο ελέγχου το οποίο όταν εκτελεστεί κατά τη διάρκεια του ελέγχου επαληθεύει την ορθή συμπεριφορά του κώδικα. Η επαλήθευση πραγματοποιείται συγκρίνοντας την κατάσταση του κώδικα που βρίσκεται υπό έλεγχο με την αναμενόμενη κατάσταση όπως αυτή έχει οριστεί από τις προδιαγραφές του κώδικα. Συνήθως, η αποτυχία ενός σημείου ελέγχου σημαίνει τη διακοπή του αντίστοιχου test suite μέσω κάποιας εξαίρεσης ή κάποιου άλλου παρόμοιου μηχανισμού.

Η χρήση του RHPUnit πραγματοποιείται συνήθως ανά τακτά χρονικά διαστήματα ώστε να επαληθεύεται η ορθή λειτουργία όλων των υποσυστημάτων της εφαρμογής κατά τη διάρκεια ανάπτυξής της. Αν κάποιος έλεγχος αστοχήσει, τότε δίνεται στον προγραμματιστή η δυνατότητα να ερευνήσει άμεσα τις αιτίες και, μάλιστα, να συνδυάσει την αστοχία με τυχόντες αλλαγές ή προσθήκες που πραγματοποιούνται εκείνη τη χρονική περίοδο στο σύστημα ούτως ώστε να μπορέσει να εντοπίσει το πρόβλημα και να προβεί στην επίλυσή του ταχύτερα. Σχετικά με τη χρήση του RHPUnit για τον έλεγχο του συστήματος της παρούσας διπλωματικής εργασίας γίνεται εκτενής αναφορά στο 7^ο κεφάλαιο, όπου περιγράφονται με λεπτομέρεια οι τεχνικές και οι μεθοδολογίες που ακολουθήθηκαν κατά τον έλεγχο του συστήματος.

6

Έλεγχος

Στο παρόν κεφάλαιο παρουσιάζονται οι τεχνικές και οι μεθοδολογίες που χρησιμοποιήθηκαν για τον έλεγχο της ορθής λειτουργίας του συστήματος. Καθότι το σύστημα αποτελείται από πολλά υποσυστήματα τα οποία βρίσκονται σε συνεχή αλληλεπίδραση κρίθηκε απαραίτητος ο συνεχόμενος και ενδεδειγμένος έλεγχος του συστήματος με χρήση μοντέρνων τεχνικών, όπως είναι το unit testing. Η χρήση των συγκεκριμένων τεχνικών κατέστη δυνατή χάρη και στην ανάπτυξη της εφαρμογής βάσει του αρχιτεκτονικού προτύπου Model – View – Controller, το οποίο αποσυμπλέκει τη λογική της εφαρμογής από την παρουσίαση. Έτσι καθίσταται ευκολότερος ο έλεγχος της ορθής λογικής του συστήματος.

6.1 Test fixture

Το test fixture αναφέρεται στη σταθερή κατάσταση που χρησιμοποιείται ως βάση για την εκτέλεση όλων των ελέγχων του συστήματος. Ο λόγος ύπαρξης ενός test fixture είναι για την διασφάλιση της ύπαρξης ενός γνωστού και καλά προκαθορισμένου περιβάλλοντος στο οποίο μπορεί να εκτελεστεί ένας έλεγχος του οποίου το αποτέλεσμα να είναι προβλέψιμο και επαναλαμβανόμενο.

Η ανάγκη ύπαρξης των test fixtures προκύπτει από την ανάγκη αυτοματοποίησης των ελέγχων. Κάθε έλεγχος που πραγματοποιείται στο σύστημα λειτουργεί στη βάση τριών βημάτων:

- Πραγματοποίηση ενέργειας στο σύστημα
- Παρατήρηση της κατάστασης του συστήματος

- Σύγκριση με αναμενόμενη κατάσταση

Το ζητούμενο είναι η τελική κατάσταση του συστήματος να συμπίπτει με την αναμενόμενη ώστε ο έλεγχος να κριθεί επιτυχημένος. Για την επίτευξη αυτού του στόχου είναι προϋπόθεση η εκ των προτέρων τοποθέτηση του συστήματος σε μία γνωστή αρχική κατάσταση. Αυτή τη ανάγκη καλύπτουν τα test fixtures. Στη συνέχεια δίνονται ορισμένα παραδείγματα χρήσης των test fixtures για τον ορισμό της αρχικής κατάστασης συστημάτων λογισμικού:

- Πριν την εκτέλεση ενός συνόλου από ελέγχους φορτώνεται μία βάση δεδομένων (ή κάποιοι πίνακες αυτής) με γνωστά και προκαθορισμένα δεδομένα, π.χ. αν πρόκειται να πραγματοποιηθεί ο έλεγχος ενός αλγορίθμου που υπολογίζει το μέσο κόστος των προϊόντων ενός ηλεκτρονικού καταστήματος, φορτώνεται ο πίνακας των προϊόντων με ένα σύνολο από γνωστά αντικείμενα, των οποίων είναι γνωστό το μέσο κόστος. Αν κανείς επιχειρήσει να πραγματοποιήσει τον έλεγχο χρησιμοποιώντας τα πραγματικά δεδομένα όπως αυτά βρίσκονται ανά πάσα στιγμή αποθηκευμένα στη βάση δεδομένων του συστήματος, δε θα είχε τη δυνατότητα να επαληθεύσει την ορθή εκτέλεση του ελέγχου μιας και το μέσο κόστος των προϊόντων δεν είναι σταθερό και, κατά πάσα πιθανότητα, αλλάζει ανάμεσα σε διαδοχικούς ελέγχους.
- Διαγραφή αρχείων δεδομένων από το δίσκο, π.χ. για τον έλεγχο ενός συστήματος το οποίο διατηρεί καταγραφές σε αρχεία δεδομένων απαιτείται ο καθαρισμός των αρχείων ούτως ώστε μετά την εκτέλεση του ελέγχου τα αρχεία δεδομένων να περιέχουν μόνο εγγραφές, οι οποίες προστέθηκαν κατά τη διάρκεια εκτέλεσής του. Σε αντίθετη περίπτωση, τα αρχεία δεδομένων, ενδεχομένως, περιέχουν εγγραφές από την προγενέστερη λειτουργία του συστήματος καθιστώντας τη διαδικασία του ελέγχου δυσκολότερη μιας και πρέπει να ξεχωρίσει τις παλιές από τις νέες εγγραφές.
- Αρχικοποίηση αρχείων δεδομένων στο δίσκο με γνωστά δεδομένα, π.χ. για τον έλεγχο ενός συστήματος που αξιοποιεί στοιχεία από αρχεία δεδομένων του συστήματος.
- Αντιγραφή προκαθορισμένων αρχείων δεδομένων στην κατάλληλη τοποθεσία του συστήματος αρχείων

Για τον έλεγχο του συστήματος της διπλωματικής εργασίας χρησιμοποιήθηκαν test fixtures για τη φόρτωση των πινάκων της σχεσιακής βάσης δεδομένων με προκαθορισμένα δεδομένα ούτως ώστε η εκτέλεση των unit tests και των system tests να παρήγαγε ένα γνωστό και αναμενόμενο αποτέλεσμα. Η χρήση και η εφαρμογή των test fixtures είναι ιδιαίτερα απλή στο περιβάλλον του Yii PHP framework, το οποίο δίνει τη δυνατότητα φόρτωσης δεδομένων στους πίνακες της σχεσιακής βάσης δεδομένων μέσω κώδικα PHP.

6.2 Unit testing

Το unit testing είναι μία μεθοδολογία επαλήθευσης και πιστοποίησης της ορθής λειτουργίας του λογισμικού. Σύμφωνα με αυτή τη μεθοδολογία το υπό έλεγχο λογισμικό χωρίζεται σε units, τα οποία είναι τα μικρότερα δυνατά κομμάτια του λογισμικού που μπορούν να ελεγχθούν ανεξάρτητα από το υπόλοιπο λογισμικό. Συνήθως, τέτοια κομμάτια λογισμικού είναι οι συναρτήσεις στις γλώσσες προστακτικού προγραμματισμού και οι κλάσεις στις γλώσσες αντικειμενοστραφούς προγραμματισμού.

Ίδανικά, ένα unit test οφείλει να είναι ανεξάρτητο από τα υπόλοιπα και να ελέγχει εξ' ολοκλήρου το κομμάτι κώδικα που του αντιστοιχεί. Πολλές φορές για να επιτευχθεί η απαιτούμενη ανεξαρτησία του κώδικα χρησιμοποιούνται διάφορες τεχνικές, όπως είναι τα fakes objects, mock objects και stub objects.

Στόχος του unit testing είναι να απομονώνει κάθε ξεχωριστό μέρος του λογισμικού και να δείχνει ότι τα ανεξάρτητα τμήματα αυτού λειτουργούν σύμφωνα με τις προδιαγραφές τους. Ένα unit test είναι ένας αυστηρός ορισμός – συμβόλαιο για τη συμπεριφορά που οφείλει να επιδεικνύει ένα συγκεκριμένο κομμάτι κώδικα. Η χρήση unit testing σε ένα μεγάλο σύστημα συνήθως βοηθάει στην ανάδειξη των προβλημάτων της εφαρμογής στα αρχικά στάδια της ανάπτυξης αυτής.

Το unit testing διευκολύνει την αναδιοργάνωση του κώδικα στην περίπτωση που αυτό κριθεί απαραίτητο στο μέλλον. Πολλές φορές η αναδιοργάνωση του κώδικα κρίνεται απαραίτητη για λόγους διευκόλυνσης της συντήρησης ή αύξησης της απόδοσης και συνήθως δε συνοδεύεται από αλλαγές στις λειτουργικές απαιτήσεις του κώδικα. Αυτό σημαίνει πως ένας εξωτερικός παρατηρητής του συστήματος δεν πρέπει να δει ή να εντοπίσει κάποια διαφορά στη λειτουργία αυτού. Όταν το unit testing χρησιμοποιείται για το συγκεκριμένο σκοπό ονομάζεται regression testing. Regression testing καλείται η τεχνική σύμφωνα με την οποία ο προγραμματιστής φροντίζει να καλύψει τον κώδικα που γράφει με unit tests, επομένως αν στο μέλλον κάποια αλλαγή στον κώδικα προκαλέσει κάποιο σφάλμα θα γίνει άμεσα αντιληπτό χάρη στην ύπαρξη των unit tests.

Το unit testing διευκολύνει την απρόσκοπτη συναρμολόγηση των διαφορετικών υποσυστημάτων. Αυτό επιτυγχάνεται διότι η χρήση των unit tests οδηγεί στον εντοπισμό των σφαλμάτων στα υποσυστήματα πριν αυτά συνδυαστούν σε πιο πολύπλοκες διαρρυθμίσεις, οπότε και είναι δυσκολότερος ο αποτελεσματικός και γρήγορος εντοπισμός σφαλμάτων.

Το unit testing προσφέρει την πληρέστερη και πιο ενημερωμένη τεκμηρίωση του συστήματος. Πολλές φορές η γραπτή τεκμηρίωση του κώδικα είναι καταδικασμένη να μένει ξεπερασμένη διότι δεν ανανεώνεται πάντα μαζί με τον κώδικα. Σε αυτές τις περιπτώσεις ο προγραμματιστής μπορεί να καταφύγει στην ανάγνωση των unit tests για να κατανοήσει πώς λειτουρ-

γεί ή πώς χρησιμοποιείται ένα συγκεκριμένο κομμάτι κώδικα. Επιπλέον, τα unit tests τεκμηριώνουν τόσο τη σωστή και ενδεδειγμένη χρήση του κώδικα όσο και τη χρήση που οδηγεί σε αναμενόμενα σφάλματα οπότε ο προγραμματιστής μπορεί να αποκτήσει μία σφαιρική κατανόηση του υποσυστήματος που μελετάει και θέλει ενδεχομένως να χρησιμοποιήσει βλέποντας τόσο παραδείγματα ορθής χρήσης όσο και παραδείγματα λανθασμένης χρήσης.

Το unit testing μπορεί να χρησιμοποιηθεί επιπρόσθετα για τον αρχικό σχεδιασμό του υπό ανάπτυξη συστήματος. Η συγκεκριμένη χρήση είναι συνηθισμένη στην test-driven μεθοδολογία ανάπτυξης λογισμικού. Σύμφωνα με αυτή τη μεθοδολογία η συγγραφή των unit tests προηγείται της συγγραφής του κώδικα της εφαρμογής και εξυπηρετεί το σκοπό τόσο της εκ των προτέρων σχεδίασης του συστήματος όσο και της εκ των υστέρων επαλήθευσης των αρχικών προδιαγραφών. Η χρήση unit tests για τον αρχικό σχεδιασμό ενός συστήματος λογισμικού έναντι των πιο κοινών μεθόδων, όπως είναι τα διαγράμματα UML, προσφέρει το πλεονέκτημα ότι αν ο προγραμματιστής παρεκκλίνει από την αρχική σχεδίαση τότε το unit test θα αποτύχει με αποτέλεσμα να εντοπίζονται άμεσα τυχόν αποκλίσεις από την αρχική σχεδίαση του συστήματος.

Παρόλα τα πλεονεκτήματα που προσφέρει η χρήση των unit tests στην ανάπτυξη σύγχρονου λογισμικού, υπάρχουν ορισμένα σημεία τα οποία πρέπει να έχει κανείς υπόψη.

Η χρήση των unit tests είναι εξαιρετικά σπάνιο να καλύπτει όλα τα πιθανά μονοπάτια εκτέλεσης ενός προγράμματος, το οποίο έχει τουλάχιστον μέτρια πολυπλοκότητα. Σε αυτή την περίπτωση είναι σαφές ότι η τεχνική των unit tests πρέπει να χρησιμοποιείται συμπληρωματικά με άλλες τεχνικές ελέγχου έτσι ώστε να εξασφαλίζεται ο βέλτιστος έλεγχος του συστήματος. Σε τελική ανάλυση, η χρήση των unit tests μπορεί να αποδείξει την ύπαρξη σφαλμάτων σε ένα πρόγραμμα αλλά δε μπορεί σε καμία περίπτωση να αποδείξει την απουσία σφαλμάτων, δηλαδή την ορθότητα του λογισμικού.

Επιπλέον η χρήση των unit tests προσθέτει έναν επιπλέον φόρτο εργασίας στους προγραμματιστές του συστήματος, οι οποίοι είναι επιφορτισμένοι με το έργο της παράλληλης συγγραφής κώδικα τόσο για την ίδια την εφαρμογή όσο και για τα unit tests αυτής. Οφείλει, λοιπόν, να υπάρχει μία ισορροπία ανάμεσα στις δύο αυτές δραστηριότητες έτσι ώστε το τελικό αποτέλεσμα να ενισχύει την παραγωγικότητα της ομάδας των προγραμματιστών. Σε καμία περίπτωση, η συγγραφή των unit tests δεν πρέπει να παρακωλύει την ομαλή ανάπτυξη και υλοποίηση του συστήματος.

Τέλος, η χρήση της μεθοδολογίας των unit tests οφείλει να συνδυάζεται με τη χρήση ενός συστήματος source version control έτσι ώστε πιθανές αποτυχίες των unit tests να μπορούν γρήγορα να επιλύονται αναζητώντας το ιστορικό των αλλαγών του προβληματικού κώδικα ή / και του κώδικα του εν λόγω unit test.

Για τον έλεγχο του συστήματος που αναπτύχθηκε στα πλαίσια της διπλωματικής εργασίας συντάχθηκε ένα σύνολο από unit tests που καλύπτουν περισσότερα από 300 διαφορετικά σημεία ελέγχου. Μέσω αυτής της τεχνικής πραγματοποιήθηκε ο έλεγχος και η πιστοποίηση της καλής λειτουργίας των μοντέλων (models), των ελεγκτών (controllers) και των συστατικών (components) του συστήματος. Στη συνέχεια, δίνονται ορισμένα παραδείγματα ελέγχων που ακολουθούν τη συγκεκριμένη τεχνική.

6.2.1 Παράδειγμα unit testing #1

Κάθε microRNA έχει μία λίστα από επίπεδα έκφρασης, όπου το κάθε επίπεδο έκφρασης αντιστοιχεί σε έναν θετικό πραγματικό αριθμό. Για την καλύτερη παρουσίαση των επιπέδων έκφρασης, ο συγκεκριμένος αριθμός αντιστοιχίζεται σε ένα χρώμα, το οποίο μεταβάλλεται ομοιόμορφα από λευκό σε πορτοκαλί και, τέλος, σε κόκκινο, όσο αυξάνεται το επίπεδο έκφρασης. Ο έλεγχος της μεθόδου απεικόνισης του επιπέδου έκφρασης σε χρώμα είναι ο ακόλουθος:

```
$this->helperGetExpressionLevelColor(0, '#ffffff');  
$this->helperGetExpressionLevelColor(5, '#ffff00');  
$this->helperGetExpressionLevelColor(200, '#ff0000');
```

Κώδικας 1: Unit test για τον έλεγχο της μεθόδου απεικόνισης του επιπέδου έκφρασης ενός microRNA σε χρώμα.

Ο παραπάνω έλεγχος επαληθεύει ότι η τιμή 0 αντιστοιχίζεται στο λευκό χρώμα (σε κωδικοποίηση #RRGGBB), η τιμή 5 αντιστοιχίζεται στο κίτρινο χρώμα και η τιμή 200 αντιστοιχίζεται στο κόκκινο χρώμα.

6.2.2 Παράδειγμα unit testing #2

Για την εγγραφή ενός χρήστη στο σύστημα είναι απαραίτητη η διεύθυνση ηλεκτρονικού ταχυδρομείου. Κατά τη διάρκεια εγγραφής πραγματοποιείται ένας τυπικός έλεγχος εγκυρότητας της συγκεκριμένης διεύθυνσης. Ο κώδικας που αναλαμβάνει να ελέγξει τη συγκεκριμένα δικλείδα ασφαλείας του συστήματος είναι ο ακόλουθος:

```
$f = new RegisterForm;  
$f->username = 'xyz';  
$f->password = '123';  
$f->password_repeat = '123';  
$this->helperTestEmail($f, NULL, 'Email cannot be blank.');
```

Κώδικας 2: Unit test για τον έλεγχο του συστήματος επαλήθευσης και εγκυρότητας της διεύθυνσης ηλεκτρονικού ταχυδρομείου για την εγγραφή νέων χρηστών στο σύστημα.

6.2.3 Παράδειγμα unit testing #3

Όταν ο χρήστης πραγματοποιήσει μία αναζήτηση, της οποίας ο όρος δεν προσδιορίζει με μοναδικό τρόπο ένα microRNA ή ένα γονίδιο, τότε το σύστημα δίνει στο χρήστη μία σειρά από προτάσεις από τις οποίες καλείται να επιλέξει αυτή που επιθυμεί. Στην καρδιά του συγκεκριμένου υποσυστήματος βρίσκεται μία δομή δεδομένων, η οποία δέχεται στοιχεία και κρατάει τα κορυφαία N από αυτά, σύμφωνα με μία συνάρτηση ταξινόμησης που δίνει ο προγραμματιστής. Για τον έλεγχο της συγκεκριμένης δομής δεδομένων χρησιμοποιείται το ακόλουθο unit test:

```
$this->top = new TopNArray(3, new Comparer());  
$this->assertSame(array(), $this->top->get_array());  
$this->putAndAssert(10, array(10));  
$this->putAndAssert(20, array(20,10));  
$this->putAndAssert(30, array(30,20,10));  
$this->putAndAssert(0, array(30,20,10));  
$this->putAndAssert(15, array(30,20,15));
```

Κώδικας 3: Unit test για τον έλεγχο της δομής δεδομένων top n array.

Ο παραπάνω έλεγχος προσθέτει σταδιακά στοιχεία στη δομή δεδομένων και επαληθεύει διαρκώς τη σωστή κατάσταση αυτής.

6.3 System testing

Το system testing είναι η διαδικασία ελέγχου που συνήθως ακολουθεί τον έλεγχο με τα unit tests. Σε αντίθεση με το unit testing, όπου ελέγχονται όλα τα συστατικά του συστήματος ανεξάρτητα το ένα από το άλλο, στο system testing γίνεται έλεγχος στο πλήρες ολοκληρωμένο σύστημα. Στόχος του ελέγχου είναι να πιστοποιήσει πως το σύστημα ακολουθεί τις προδιαγραφές που είχαν τεθεί κατά το σχεδιασμό του. Το system testing λειτουργεί ως έλεγχος μαύρου κουτιού, το οποίο σημαίνει ότι δε χρειάζεται να λαμβάνει υπόψη του τις λεπτομέρειες υλοποίησης του συστήματος.

Στη συγκεκριμένη εφαρμογή εφαρμόστηκε μία αυτοματοποιημένη διαδικασία system testing. Σύμφωνα με αυτή τη διαδικασία ορίστηκε ένας αριθμός από σενάρια χρήσης – που κυμαίνονται από πολύ απλά έως αρκετά σύνθετα – τα οποία στη συνέχεια εκτελέστηκαν με αυτόματο τρόπο στη δικτυακή εφαρμογή και τα αποτελέσματά τους αποθηκεύτηκαν σε αρχεία. Στη συνέχεια, ήταν απαραίτητος ο ανθρώπινος παράγοντας για να εξασφαλίσει ότι τα αρχεία αποτελεσμάτων ήταν σωστά και αναμενόμενα για κάθε σενάριο που εκτελέστηκε. Εφόσον τα αρχεία κρίθηκαν αναμενόμενα, αποθηκεύτηκαν πλέον σε μόνιμη θέση με σκοπό να συγκρίνονται με τα νέα αρχεία που προκύπτουν από μεταγενέστερα system tests και να επαληθεύουν την ορθή ανταπόκριση του συστήματος στα σενάρια που έχουν οριστεί γι' αυτό το σκοπό.

Ο αυτοματοποιημένος έλεγχος πραγματοποιήθηκε με το εργαλείο wget (<http://www.gnu.org/software/wget/>), το οποίο μπορεί να εκτελεί HTTP ερωτήματα σε κάποιον web server και να

αποθηκεύει την απάντηση. Το συγκεκριμένο εργαλείο προσφέρει σημαντικές δυνατότητες, οι οποίες ήταν απαραίτητες για τον πλήρη έλεγχο όλων των χαρακτηριστικών του συστήματος:

- Έχει τη δυνατότητα να εκτελεστεί αυτόματα, χωρίς ανθρώπινη παρέμβαση, μέσω ενός κατάλληλου command line script.
- Έχει τη δυνατότητα να αποθηκεύει την απάντηση του web server σε αρχείο με όνομα που ορίζει ο προγραμματιστής.
- Έχει τη δυνατότητα να ορίζει τις παραμέτρους GET του HTTP ερωτήματος με απλό τρόπο.
- Έχει τη δυνατότητα να αποθηκεύει τα HTTP cookies σε κατάλληλο αρχείο κειμένου, από το οποίο μετέπειτα δύναται να τα αναγνώσει και να τα στείλει στο web server με ένα νέο HTTP ερώτημα. Αυτή η δυνατότητα κρίθηκε απαραίτητη ώστε να λειτουργήσουν τα system tests που αφορούν τις εξατομικευμένες λειτουργίες του συστήματος, όπως είναι η είσοδος του χρήστη, η χρήση των αγαπημένων αναζητήσεων και η χρήση του ιστορικού των αναζητήσεων.
- Έχει τη δυνατότητα να αποθηκεύει εκτός από τα persistent HTTP cookies και τα session HTTP cookies, τα οποία είναι αυτά που χρησιμοποιούνται κατά την πιστοποίηση χρηστών στο σύστημα. Τα session HTTP cookies έχουν διάρκεια μίας συνεδρίας, το οποίο σημαίνει ότι καταργούνται όταν ο χρήστης τερματίσει τον browser του.
- Έχει τη δυνατότητα να εκτελεί HTTP ερωτήματα τύπου POST και να ορίζει με εύκολο τρόπο τα δεδομένα που αποστέλλονται μαζί με το ερώτημα. Αυτή η δυνατότητα κρίθηκε αναγκαία για την αποστολή των στοιχείων του χρήστη κατά τη διάρκεια της ταυτοποίησής του. Το όνομα χρήστη και ο κωδικός πρόσβασης αποστέλλονται στο web server μέσω POST για μέγιστη ασφάλεια.
- Έχει τη δυνατότητα να εκτελεί HTTP ερωτήματα αγνοώντας την απάντηση του web server.

Μετά την εκτέλεση του αυτοματοποιημένου command line script που εκτελεί όλα τα σενάρια και αποθηκεύει τις απαντήσεις του web server γινόταν η σύγκριση των αποτελεσμάτων με τα αναμενόμενες απαντήσεις με χρήση του εργαλείου diff. Το εργαλείο diff έχει τη δυνατότητα να συγκρίνει δύο λίστες αρχείων και να δώσει ως έξοδο τις διαφορές που παρατηρούνται μεταξύ των αντίστοιχων αρχείων. Ο προγραμματιστής ήταν υπεύθυνος, πλέον, να κρίνει αν οι διαφορές που εμφανίζονται είναι επιθυμητές ή αν πρόκειται για κάποιο σφάλμα που εισήχθη στο σύστημα και πρέπει να εντοπιστεί και να αντιμετωπιστεί.

Για τον ενδελεχή και πλήρη έλεγχο του συστήματος σχεδιάστηκε μία σειρά από 57 σενάρια ελέγχου. Το αποτέλεσμα κάθε σεναρίου ελέγχου συγκρίνεται με το αναμενόμενο κάθε φορά που εκτελούνται τα system tests. Τα σενάρια ελέγχου είναι τα ακόλουθα:

1. Παρουσίαση της αρχικής σελίδας.
2. Προβολή αποτελεσμάτων για ένα microRNA δίνοντας το ακριβές όνομά του.
3. Προβολή προτάσεων για ένα microRNA.
4. Προβολή αποτελεσμάτων για ένα microRNA ακολουθώντας μία από τις προτάσεις του συστήματος.
5. Προβολή αποτελεσμάτων για ένα γονίδιο δίνοντας το ακριβές όνομά του.
6. Προβολή προτάσεων για ένα γονίδιο.
7. Προβολή αποτελεσμάτων για ένα γονίδιο ακολουθώντας μία από τις προτάσεις του συστήματος.
8. Προβολή αποτελεσμάτων για ένα microRNA και ένα γονίδιο συνδυαστικά ακολουθώντας μία από τις προτάσεις του συστήματος.
9. Προβολή προτάσεων για ένα microRNA και ένα γονίδιο στην ίδια σελίδα.
10. Προβολή αποτελεσμάτων για ένα microRNA και ένα γονίδιο συνδυαστικά δίνοντας τα ακριβή ονόματά τους.
11. Προβολή προτάσεων για ένα γονίδιο, όταν ο χρήστης έχει δώσει μοναδικό όνομα microRNA και διαφορετικό όνομα γονιδίου.
12. Προβολή προτάσεων για ένα microRNA, όταν ο χρήστης έχει δώσει διαφορετικό όνομα microRNA και μοναδικό όνομα γονιδίου.
13. Αναζήτηση με κενά πεδία, η οποία πρέπει να ανακατευθύνει το χρήστη της εφαρμογής στην αρχική σελίδα.
14. Αναζήτηση με άκυρο αναγνωριστικό microRNA.
15. Αναζήτηση με άκυρο αναγνωριστικό γονιδίου.
16. Αναζήτηση με άκυρα αναγνωριστικά microRNA ή / και γονιδίου. Και στις τρεις παραπάνω περιπτώσεις η εφαρμογή πρέπει να ανακατευθύνει το χρήστη στην αρχική σελίδα.
17. Προβολή αποτελεσμάτων για αναζήτηση με βάση το microRNA όταν δε βρεθούν βιολογικοί στόχοι.
18. Προβολή αποτελεσμάτων για αναζήτηση με βάση το γονίδιο όταν δε βρεθούν βιολογικοί στόχοι.
19. Προβολή αποτελεσμάτων για αναζήτηση με βάση το microRNA και το γονίδιο όταν δε βρεθούν βιολογικοί στόχοι.
20. Προβολή αποτελεσμάτων για αναζήτηση με βάση το microRNA όταν τα αποτελέσματα είναι περισσότερα από μία σελίδα.

21. Μετακίνηση στη δεύτερη σελίδα των αποτελεσμάτων του προηγούμενου σεναρίου ελέγχου.
22. Προβολή αποτελεσμάτων για αναζήτηση με βάση το microRNA και, επιπλέον, περιορισμό με βάση το score των αποτελεσμάτων.
23. Προβολή αποτελεσμάτων για αναζήτηση με βάση το microRNA και, επιπλέον, περιορισμό των αποτελεσμάτων με βάση το KEGG pathway των αποτελεσμάτων.
24. Όπως και στο προηγούμενο σενάριο ελέγχου όταν τα αποτελέσματα ξεπερνούν τη μία σελίδα.
25. Μετακίνηση στη δεύτερη σελίδα των αποτελεσμάτων του προηγούμενου σεναρίου ελέγχου.
26. Προβολή αποτελεσμάτων για αναζήτηση με βάση το microRNA και, επιπλέον, περιορισμό με βάση το KEGG pathway των αποτελεσμάτων όταν δε βρεθούν βιολογικοί στόχοι που να ικανοποιούν και τα δύο κριτήρια.
27. Προβολή προτάσεων για γονίδιο όταν ο χρήστης εισάγει διαφορετικό όνομα refseq γονιδίου.
28. Προβολή αποτελεσμάτων για αναζήτηση με βάση το refseq γονιδίου όταν ο χρήστης εισάγει μοναδικό όνομα.
29. Εξαγωγή και μεταφόρτωση αρχείου αποτελεσμάτων σε μορφή Comma Separated Values για αναζήτηση με βάση το microRNA.
30. Εξαγωγή και μεταφόρτωση αρχείου αποτελεσμάτων σε μορφή Comma Separated Values για αναζήτηση με βάση το γονίδιο.
31. Προβολή προτάσεων για αναζήτηση με βάση την περιγραφή γονιδίου.
32. Προβολή προτάσεων για αναζήτηση με βάση την περιγραφή γονιδίου, όταν η περιγραφή που δίνεται δεν ταιριάζει με κάποια γνωστή περιγραφή.
33. Προβολή προτάσεων για αναζήτηση με βάση το όνομα γονιδίου και την περιγραφή γονιδίου. Σε αυτή την περίπτωση, το σύστημα πρέπει να αγνοήσει την περιγραφή του γονιδίου που δίνεται και να δώσει προτάσεις μόνο με βάση το όνομα του γονιδίου.

Για την πραγματοποίηση των παρακάτω σεναρίων ελέγχου συνδέονται στο σύστημα δύο διαφορετικοί χρήστες, έστω ο χρήστης A και ο χρήστης B.

34. Προβολή αρχικής σελίδας για το συνδεδεμένο στο σύστημα χρήστη A.
35. Προβολή αποτελεσμάτων για αναζήτηση με βάση το microRNA όταν ο χρήστης A δεν έχει αποθηκεύσει τη συγκεκριμένη αναζήτηση στις αγαπημένες του.
36. Προβολή λίστας αγαπημένων αναζητήσεων για το χρήστη A.

37. Προσθήκη αναζήτησης στη λίστα αγαπημένων του χρήστη A και προβολή των αποτελεσμάτων της αναζήτησης για τον ίδιο χρήστη.
38. Προβολή αποτελεσμάτων για αναζήτηση με διαφορετικό microRNA για το χρήστη A.
39. Προβολή αποτελεσμάτων για την αναζήτηση που προστέθηκε στη λίστα αγαπημένων του χρήστη A από το χρήστη B.
40. Προσθήκη ακόμη μίας αναζήτησης στη λίστα αγαπημένων του χρήστη A και προβολή της λίστας αγαπημένων για τον ίδιο χρήστη.
41. Προβολή λίστας αγαπημένων για το χρήστη B.
42. Προβολή σελίδας επεξεργασίας αγαπημένης αναζήτησης για το χρήστη A.
43. Αλλαγή της περιγραφής μίας αγαπημένης αναζήτησης από το χρήστη A και προβολή της λίστας αγαπημένων αναζητήσεων για τον ίδιο χρήστη.
44. Διαγραφή μίας αγαπημένης αναζήτησης από το χρήστη A και προβολή της λίστας αγαπημένων αναζητήσεων για τον ίδιο χρήστη.

45. Προβολή αποτελεσμάτων για αναζήτηση με βάση το αναγνωριστικό ενός hairpin microRNA.
46. Προβολή αποτελεσμάτων για αναζήτηση με βάση το μοναδικό όνομα ενός hairpin microRNA.
47. Προβολή αποτελεσμάτων για αναζήτηση με βάση το αναγνωριστικό ενός microRNA στην εφαρμογή mirGen.
48. Προβολή αποτελεσμάτων για αναζήτηση με βάση το αναγνωριστικό ενός transcription factor στην εφαρμογή mirGen.

49. Καθαρισμός του ιστορικού αναζητήσεων για το χρήστη A.
50. Επεξεργασία αγαπημένης αναζήτησης από το χρήστη A όταν τα αποτελέσματά της ξεπερνούν τη μία σελίδα.
51. Αλλαγή της περιγραφής και ορισμός ενός βιολογικού στόχου για το παραπάνω σενάριο ελέγχου.
52. Μετάβαση στη δεύτερη σελίδα για το παραπάνω σενάριο ελέγχου.
53. Εκ νέου αλλαγή της περιγραφής και ορισμός ενός ακόμη βιολογικού στόχου για το παραπάνω σενάριο ελέγχου.
54. Επιστροφή στην πρώτη σελίδα των αποτελεσμάτων για το παραπάνω σενάριο ελέγχου.

55. Επεξεργασία αγαπημένης αναζήτησης με βάση το γονίδιο.
56. Προβολή αγαπημένης αναζήτησης με βάση το microRNA με επιλεγμένο έναν αγαπημένο βιολογικό στόχο.
57. Προβολή αγαπημένης αναζήτησης με βάση το microRNA με επιλεγμένους δύο αγαπημένους βιολογικούς στόχους.

7

Επίλογος

Στο τελευταίο κεφάλαιο συνοψίζεται η παρούσα διπλωματική εργασία. Παρουσιάζονται επιγραμματικά οι στόχοι που επιτεύχθηκαν και γίνεται σύγκριση με τις αρχικές προσδοκίες ούτως ώστε να προσδιοριστεί ο βαθμός επιτυχίας της ανάπτυξης του συστήματος. Επιπλέον, τεκμηριώνεται μία σειρά από συμπεράσματα τα οποία προέκυψαν μέσα από τη διαδικασία της ανάπτυξης του λογισμικού της διπλωματικής εργασίας. Τέλος, πραγματοποιείται μία σύντομη αναφορά στα θέματα που ήταν εκτός του αρχικού σκοπού της εργασίας και, επομένως, αφήνονται ανοικτά στο υπάρχον σύστημα. Τέτοια θέματα μπορούν να καλυφθούν μετέπειτα με μία σειρά από μελλοντικές επεκτάσεις στο σύστημα, για τις οποίες δίνεται ένα σύνολο από στόχους και προδιαγραφές και ορίζονται οι στόχοι που μένουν να επιτευχθούν.

7.1 Σύνοψη και συμπεράσματα

Στην παρούσα διπλωματική εργασία αναπτύχθηκε ένα ολοκληρωμένο σύστημα έρευνας και αναζήτησης πληροφορίας με βιολογικό ενδιαφέρον. Το κοινό στο οποίο απευθύνεται η εφαρμογή αποτελείται από ερευνητές του επιστημονικού πεδίου της βιολογίας. Η ανάπτυξη της εφαρμογής είναι μία έμπρακτη απόδειξη της άριστης συνεργασίας του χώρου της βιολογίας με το χώρο της πληροφορικής και συγκεκριμένα το χώρο της ανάπτυξης λογισμικού. Τα προβλήματα που καλούνται να αντιμετωπίσουν οι βιολόγοι αποτελούν προκλήσεις όταν ζητείται να υλοποιηθούν ως σύστημα λογισμικού μιας και περιλαμβάνουν, συνήθως, μεγάλο όγκο δεδομένων ενώ συχνά υπάρχουν αυστηρές απαιτήσεις για την ταχύτητα απόκρισης του προ-

ϊόντος αλλά και για τον τρόπο παρουσίασης των αποτελεσμάτων. Από την άλλη πλευρά, η ανάπτυξη συστημάτων λογισμικού για την επεξεργασία των βιολογικών δεδομένων δίνει μία ώθηση στη σχετική έρευνα μιας και παρέχει τα απαραίτητα εκείνα εργαλεία που επιτρέπουν σε έναν ερευνητή να πραγματοποιήσει με εύκολο και γρήγορο τρόπο μία σειρά από μελέτες ή πειράματα τα οποία θα ήταν αδύνατο να ολοκληρωθούν δίχως τη συνδρομή του εκάστοτε λογισμικού.

Η εφαρμογή που αναπτύχθηκε καλείται να συνδράμει στην έρευνα των microRNA παρέχοντας μία πλήρη βάση δεδομένων και τον τρόπο πρόσβασης σε αυτήν. Η βάση δεδομένων περιλαμβάνει ένα πλήθος από σημεία πρόσδεσης των μορίων microRNA πάνω στα μόρια mRNA. Τα σημεία πρόσδεσης έχουν προκύψει μέσω ενός αλγόριθμου προσομοίωσης της βιολογικής διαδικασίας σύνδεσης των συγκεκριμένων μορίων. Ο αλγόριθμος έχει αναπτυχθεί από την ομάδα του Ερευνητικού Κέντρου Βιοϊατρικών Επιστημών «Αλέξανδρος Φλέμινγκ». Ο χρήστης έχει τις δυνατότητες απλής και σύνθετης αναζήτησης στα δεδομένα ούτως ώστε να μπορεί αποτελεσματικά να εντοπίσει τη ζητούμενη πληροφορία. Συγκεκριμένα, παρέχεται η δυνατότητα αναζήτησης και προβολής των βιολογικών στόχων ενός microRNA, ενός γονιδίου αλλά και ενός συνδυασμού των δύο τελευταίων. Η μελέτη των βιολογικών στόχων των μορίων microRNA διαδραματίζει καθοριστικό ρόλο στην κατανόηση της λειτουργίας των οργανισμών αλλά και στην αντιμετώπιση ασθενειών και παθήσεων, όπως έχει αποδείξει η σύγχρονη μελέτη των εν λόγω μορίων. Επιπλέον, έχουν υλοποιηθεί μία σειρά από διαφορετικούς τρόπους παρουσίασης των αποτελεσμάτων έτσι ώστε αυτά να είναι άμεσα κατανοητά και οικεία προς τον χρήστη που υποβάλει τα ερωτήματά του. Τέλος, παρέχεται ένα σύνολο από προσωπικές υπηρεσίες στις οποίες έχει πρόσβαση ο κάθε χρήστης του συστήματος με στόχο τη βελτίωση της εμπειρίας αυτού από τη χρήση του συστήματος.

Έχοντας ως βάση την προηγούμενη έκδοση της εφαρμογής, η παρούσα διπλωματική εργασία κλήθηκε να βελτιώσει τις επιδόσεις και να εμπλουτίσει τις δυνατότητες αυτής. Οι στόχοι που επιτεύχθηκαν κατά την ανάπτυξη της νέας έκδοσης του συστήματος παρουσιάζονται επιγραμματικά παρακάτω:

- Σχεδίαση και υλοποίηση εκ νέου της σχεσιακής βάσης δεδομένων η οποία περιλαμβάνει όλες τις πληροφορίες βιολογικού ενδιαφέροντος που διαχειρίζεται η εφαρμογή. Η σχεδίαση αφορούσε τη συγγραφή ενός νέου σχήματος για τη βάση δεδομένων το οποίο είχε στόχο τη βελτίωση της απόδοσης, την ορθότερη αναπαράσταση των δεδομένων καθώς και τη δυνατότητα μελλοντικών επεκτάσεων. Στο κομμάτι της υλοποίησης ελήφθησαν υπόψη παράγοντες που βελτιώνουν την απόδοση της βάσης δεδομένων όπως είναι η σωστή εφαρμογή των ευρετηρίων.
- Εκ νέου συγγραφή του κώδικα της εφαρμογής σύμφωνα με το αρχιτεκτονικό πρότυπο Model – View – Controller έτσι ώστε ο κώδικας να είναι δομημένος, να παρέχει

ευκολία στον έλεγχο του και να είναι, όσο το δυνατόν περισσότερο, επεκτάσιμος σύμφωνα με τις μελλοντικές ανάγκες που πιθανόν προκύπτουν.

- Βελτίωση της παρουσίασης της εφαρμογής απλοποιώντας το user interface και χρησιμοποιώντας χειριστήρια ελέγχου με ομοιόμορφη εμφάνιση αλλά και λειτουργίες. Με αυτόν τον τρόπο ο χρήστης της εφαρμογής εξοικειώνεται γρηγορότερα με το περιβάλλον λειτουργίας εφόσον δεν υπάρχουν δεκάδες διαφορετικά χειριστήρια ελέγχου, τα οποία πρέπει να συνηθίσει ξεχωριστά.
- Παροχή ενός συνόλου από ελέγχους αξιοπιστίας και ορθής λειτουργίας του συστήματος. Οι έλεγχοι ακολουθούν τη σύγχρονη μεθοδολογία του unit testing και του system testing. Παρέχουν σχεδόν πλήρη κάλυψη του κώδικα και έλεγχο τόσο στους μεμονωμένους αλγορίθμους που χρησιμοποιούνται όσο και τη μεταξύ τους σύνδεση και αρμονική λειτουργία. Το σύνολο των ελέγχων είναι αυτοματοποιημένο επιτρέποντας με αυτόν τον τρόπο ανά πάσα στιγμή τον έλεγχο και την επαλήθευση της ορθής εκτέλεσης του συστήματος.
- Προσθήκη προσωπικών υπηρεσιών για κάθε χρήστη της εφαρμογής. Οι προσωπικές υπηρεσίες παρέχονται στο πλαίσιο της εξυπηρέτησης των διαφορετικών αναγκών των χρηστών. Έχει υλοποιηθεί σύστημα διαχείριση λογαριασμών, με το οποίο ο κάθε χρήστης μπορεί να αποκτήσει εξατομικευμένη πρόσβαση στο σύστημα. Οι υπηρεσίες που παρέχονται αφορούν την αποθήκευση και επεξεργασία των αναζητήσεων του χρήστη καθώς και την παροχή ιστορικής καταγραφής των πρόσφατων αναζητήσεών του.
- Κεντρική καταγραφή των αναζητήσεων που πραγματοποιούνται στο σύστημα από όλους τους χρήστες. Στη συγκεκριμένη καταγραφή έχει πρόσβαση μόνο ο διαχειριστής του συστήματος και παρέχεται για λόγους επεξεργασίας και ανάλυσης του τρόπου χρήσης της εφαρμογής αλλά και στατιστικής αξιοποίησης του συνόλου των αναζητήσεων.

Κατά τη διάρκεια ανάπτυξης του συνόλου των δυνατοτήτων του συστήματος παρουσιάστηκαν και αντιμετωπίστηκαν διάφορα ζητήματα, τα οποία αφορούσαν είτε την απόδοση του συστήματος γενικότερα και της βάσης δεδομένων ειδικότερα είτε τη μορφή παρουσίασης των αποτελεσμάτων, η οποία έπρεπε να είναι άμεσα κατανοητή στον ερευνητή – βιολόγο που προτίθεται να χρησιμοποιήσει την εφαρμογή της παρούσας διπλωματικής εργασίας.

Παρόλα τα ζητήματα που προέκυπταν, ο ρυθμός ανάπτυξης του συστήματος παρέμεινε υψηλός και προς αυτή την κατεύθυνση συνέδραμε η σωστή αρχιτεκτονική σχεδίαση της εφαρμογής. Συγκεκριμένα, η χρήση του αρχιτεκτονικού προτύπου Model – View – Controller αποδείχτηκε άριστη επιλογή για ένα σύστημα με τη δεδομένη πολυπλοκότητα μιας και δίνει τη δυνατότητα στον προγραμματιστή να αναπτύσσει την εφαρμογή μεθοδικά χωρίς πολύπλοκες

αλληλεξαρτήσεις στον κώδικά της. Επιπλέον, η συχνή χρήση αυτόνομων συστατικών στοιχείων για τη δόμηση της εφαρμογής επέτρεψε την αποδοτική επαναχρησιμοποίηση του κώδικα δίνοντας εύκολα λύση σε προβλήματα που εμφανίζονταν πολλαπλές φορές.

Αξιολογή ήταν, επίσης, και η συνεισφορά στην ανάπτυξη του λογισμικού η συγγραφή και χρήση των αυτοματοποιημένων ελέγχων, μέσω των οποίων γίνονταν αμέσως αντιληπτά τα προβλήματα της εφαρμογής προτού ακόμη υπάρξει η ευκαιρία να περάσει το σύστημα από ανθρώπινο έλεγχο. Στην πορεία, άλλωστε, οι αυτοματοποιημένοι έλεγχοι αυξήθηκαν ώστε να καλύπτουν όλες τις δυνατές περιπτώσεις χρήσης του συστήματος με αποτέλεσμα να είναι εξαιρετικά δύσκολο και χρονοβόρο για έναν χρήστη να πραγματοποιήσει με ακρίβεια το ίδιο σύνολο ελέγχων.

Τέλος, ο σωστός σχεδιασμός του μοντέλου οντοτήτων – συσχετίσεων και ο ακόλουθος σωστός σχεδιασμός του σχεσιακού σχήματος της βάσης δεδομένων της εφαρμογής έδωσε τη λύση στα προβλήματα που παρουσίαζε η προγενέστερη έκδοση της εφαρμογής. Συγκεκριμένα, αντιμετωπίστηκαν προβλήματα στην παρουσίαση των δεδομένων και στη σελιδοποίηση των αποτελεσμάτων αλλάζοντας τον τρόπο αναπαράστασης των βιολογικών στόχων και των σημείων πρόσδεσης με τέτοιο τρόπο ώστε να ευνοούνται τα εκτελούμενα ερωτήματα. Επίσης βελτιώθηκε η ταχύτητα απόκρισης του συστήματος μιας και ο νέος σχεδιασμός της βάσης δεδομένων επέτρεψε τη χρήση αποδοτικότερων ευρετηρίων.

Εν κατακλείδι, η ενασχόληση με τη σχεδίαση και υλοποίηση του συστήματος της παρούσας διπλωματικής εργασίας βελτίωσε σε πολλά σημεία την προγενέστερη έκδοση της εφαρμογής και έθεσε τις βάσεις για τις μελλοντικές επεκτάσεις αυτής. Ταυτόχρονα, μελετήθηκαν μία σειρά από σύγχρονες μεθοδολογίες ανάπτυξης λογισμικού οι οποίες και εφαρμόστηκαν στην πράξη δίνοντας ώθηση στην ανάπτυξη της εφαρμογής.

7.2 Μελλοντικές επεκτάσεις

Η εφαρμογή που αναπτύχθηκε στο πλαίσιο της παρούσας διπλωματικής εργασίας καθοδηγήθηκε από τις διαρκείς ανάγκες των ερευνητών – βιολόγων, τις οποίες και προσπαθεί να καλύψει με το βέλτιστο δυνατό τρόπο. Οι εφαρμογές που αναλαμβάνουν να εξυπηρετήσουν τις καθημερινές απαιτήσεις του απλού χρήστη οφείλουν να βρίσκονται σε μία διαρκή προσπάθεια ανταπόκρισης στις νέες προκλήσεις, όπως αυτές προσδιορίζονται από τις ανακατατάξεις των προσδοκιών του χρήστη από το σύστημα λογισμικού που χρησιμοποιεί. Στο πλαίσιο, επομένως, της διαρκούς βελτίωσης της εφαρμογής έχουν ήδη εντοπιστεί ορισμένα σημεία που φιλοδοξούν να ανήκουν σε μία από τις επικείμενες μελλοντικές επεκτάσεις του συστήματος. Οι συγκεκριμένες προοπτικές επέκτασης της εφαρμογής αναφέρονται συνοπτικά παρακάτω.

Μία επιδίωξη της εφαρμογής είναι η προσθήκη ενός υποσυστήματος διαχείρισης, στο οποίο θα έχουν πρόσβαση μέλη από την ομάδα διαχείρισης του συστήματος και θα υπάρχει η δυνατότητα πρόσβασης τόσο στα στοιχεία των χρηστών όσο και στα αρχεία καταγραφών που διατηρεί η εφαρμογή για τις αναζητήσεις που πραγματοποιούνται μέσω αυτής. Στόχος του υποσυστήματος είναι να προσφέρει έναν απλό και εύχρηστο μηχανισμό για τη διαχείριση του όλου συστήματος αλλά και την άντληση των αρχείων καταγραφής έτσι ώστε μέσω της επεξεργασίας αυτών να προκύπτουν χρήσιμα στατιστικά δεδομένα. Προς το παρόν, η διαχείριση της εφαρμογής δύναται να πραγματοποιηθεί με τη συνδρομή εξωτερικών εργαλείων και χρήση της γλώσσας SQL.

Ένα δεύτερο σημείο στο οποίο θα μπορούσε να βελτιωθεί η εφαρμογή της διπλωματικής εργασίας είναι ο εμπλουτισμός του συστήματος καταγραφής με περισσότερες δυνατότητες έτσι ώστε να είναι δυνατή η εξαγωγή περισσότερων συμπερασμάτων από τα αρχεία καταγραφής. Προς το παρόν, καταγράφονται οι όροι των αναζητήσεων που πραγματοποιούν οι χρήστες της εφαρμογής. Η επιδίωξη είναι να καταγράφεται και ο τρόπος με τον οποίο χειρίζονται τη σελίδα με τα αποτελέσματα. Αυτή η καταγραφή περιλαμβάνει τους βιολογικούς στόχους που επιλέγει ο χρήστης να προβάλλει, τους συνδέσμους προς εξωτερικές σελίδες που ενδέχεται να χρησιμοποιήσει, το πλήθος των σελίδων που επιλέγει να προσπελάσει και τα φίλτρα (ελάχιστος βαθμός ή / και KEGG pathway) που επιλέγει να εφαρμόσει στα αποτελέσματα.

Στο πλαίσιο της παροχής καλύτερων υπηρεσιών εξατομίκευσης εντάσσεται μία ακόμη μελλοντική επέκταση της εφαρμογής. Πρόκειται για τη δημιουργία ενός υποσυστήματος, το οποίο θα αναλαμβάνει να επεξεργαστεί το αρχείο καταγραφής ενός χρήστη του συστήματος, να εντοπίσει τα πιθανά ενδιαφέροντά του και να προβεί στη δημιουργία ενός καταλόγου από προτάσεις προς το συγκεκριμένο χρήστη. Στόχος του συγκεκριμένου υποσυστήματος είναι να βοηθήσει τον ερευνητή – βιολόγο να βρίσκει γρήγορα και εύκολα τις πληροφορίες που αναζητά. Η δημιουργία του συγκεκριμένου υποσυστήματος στην παρούσα εφαρμογή έχει σημαντική πρακτική αξία μιας και η εφαρμογή έχει μεγάλη χρησιμότητα στην ερευνητική κοινότητα και οι ερευνητές – βιολόγοι που την αξιοποιούν αφήνουν καθημερινά σε αυτήν το ψηφιακό ίχνος των ερευνητικών τους δραστηριοτήτων, το οποίο και μπορεί να αξιοποιηθεί για να προσφερθούν αρτιότερες υπηρεσίες προς αυτούς. Ένα παράδειγμα προτάσεων που μπορούν να προσφερθούν από το συγκεκριμένο υποσύστημα είναι οι προτεινόμενες εργασίες προς ανάγνωση οι οποίες ταιριάζουν στην ερευνητική περιοχή του χρήστη όπως αυτή προκύπτει από το ιστορικό των αναζητήσεών του. Είναι προφανές πως με τη σωστή αξιοποίηση των αρχείων καταγραφής μπορούν να δημιουργηθούν νέες υπηρεσίες με πρόσθετη αξία για τον τελικό χρήστη.

Τέλος, πρέπει να αναφερθεί ότι οι προβλέψεις των στόχων που παρουσιάζονται από το σύστημα που αναπτύχθηκε στην παρούσα διπλωματική εργασία προέρχονται από την εφαρμογή

ενός αλγορίθμου που έχει αναπτυχθεί από την ερευνητική ομάδα του Ερευνητικού Κέντρου Βιοϊατρικών Επιστημών «Αλέξανδρος Φλέμινγκ», με την οποία άλλωστε υπήρξε και άριστη συνεργασία κατά τη διάρκεια ανάπτυξης του συστήματος. Ο συγκεκριμένος αλγόριθμος έχει τη δυνατότητα πρόβλεψης στόχων παρακάμπτοντας την ανάγκη πραγματοποίησης χρονοβόρων και δαπανηρών πειραμάτων επιτρέποντας με αυτόν τον τρόπο στην ερευνητική κοινότητα των βιολόγων να έχει πρόσβαση σε πολύτιμες πληροφορίες, οι οποίες ενδεχομένως μπορούν να συνδράμουν στην αντιμετώπιση ασθενειών. Προσπάθειες για την ανάπτυξη παρόμοιων αλγορίθμων έχουν πραγματοποιηθεί από αρκετές ερευνητικές ομάδες ανά τον κόσμο που ασχολούνται με τη μελέτη των μορίων microRNA. Ιδιαίτερο ενδιαφέρον, επομένως, παρουσιάζει για την ερευνητική κοινότητα των βιολόγων η συγκριτική παρουσίαση των στόχων όπως προκύπτουν από την εφαρμογή διαφορετικών αλγορίθμων. Στο υπάρχον σύστημα, η συγκεκριμένη ανάγκη καλύπτεται από το πεδίο “also predicted by”, το οποίο χαρακτηρίζει ένα βιολογικό στόχο και δηλώνει ένα σύνολο από αλγορίθμους, οι οποίοι προβλέπουν το συγκεκριμένο στόχο. Σε αυτή την κατεύθυνση, επιθυμητή και χρήσιμη δυνατότητα για τους ερευνητές – βιολόγους είναι η αναζήτηση στόχων που προβλέπονται από έναν δεδομένο ελάχιστο αριθμό αλγορίθμων. Για αυτό το λόγο, μελλοντική επιδίωξη με στόχο τη βελτίωση του συστήματος της παρούσας διπλωματικής εργασίας είναι η μελέτη της συγκεκριμένης λειτουργικότητας και η προσπάθεια για μία πιο πλήρη και ολοκληρωμένη ενσωμάτωση αυτής στο σύστημα.

8

Βιβλιογραφία

- [AMP+09] Alexiou, Panagiotis and Maragkakis, Manolis and Papadopoulos, Giorgos L. and Reczko, Martin and Hatzigeorgiou, Artemis G., Lost in translation: an assessment and perspective for computational microRNA target identification, *Bioinformatics*, 25(23), 2009
- [Apache] Apache HTTP Server Version 2.2 Documentation, <http://httpd.apache.org/docs/2.2/en/>
- [AVG+09] Alexiou, Panagiotis and Vergoulis, Thanasis and Gleditsch, Martin and Prekas, George and Dalamagas, Theodore and Megraw, Molly and Grosse, Ivo and Sellis, Timos and Hatzigeorgiou, Artemis G., miRGen 2.0: a database of microRNA genomic information and regulation, *Nucl. Acids Res.*, 2009
- [Bec02] Kent Beck, *Test Driven Development: By Example*, Addison-Wesley Longman, 2002
- [Lei09] Agnes Leitner, *MicroRNA target prediction*, Graz University of Technology, 2009
- [MAP+09] Maragkakis, Manolis and Alexiou, Panagiotis and Papadopoulos, Giorgio and Reczko, Martin and Dalamagas, Theodore and Giannopoulos, George and Goumas, George and Koukis, Evangelos and Kourtis, Kornilios and Simossis, Victor and Sethupathy, Praveen and Vergoulis, Thanasis and

- Koziris, Nectarios and Sellis, Timos and Tsanakas, Panagiotis and Hatzigeorgiou, Artemis, Accurate microRNA target prediction correlates with protein repression levels, BMC Bioinformatics, 10, 2009
- [MRS+09] Maragkakis, M. and Reczko, M. and Simossis, V. A. and Alexiou, P. and Papadopoulos, G. L. and Dalamagas, T. and Giannopoulos, G. and Goumas, G. and Koukis, E. and Kourtis, K. and Vergoulis, T. and Koziris, N. and Sellis, T. and Tsanakas, P. and Hatzigeorgiou, A. G., DIANA-microT web server: elucidating microRNA functions through target prediction, Nucl. Acids Res., 37, 2009
- [MSC+06] Megraw, Molly and Sethupathy, Praveen and Corda, Benoit and Hatzigeorgiou, Artemis G., miRGen: a database for the study of animal microRNA genomic organization and function, Nucl. Acids Res., 2006
- [Mysql] MySQL 5.1 Reference Manual, <http://dev.mysql.com/doc/refman/5.1/en/index.html>
- [Phpuni] PHPUnit Manual, <http://www.phpunit.de/manual/3.4/en/index.html>
- [Php] PHP Manual, <http://www.php.net/manual/en/>