

Ατομική Διπλωματική Εργασία

**ΑΥΤΟΜΑΤΟΠΟΙΗΜΕΝΗ ΜΕΘΟΔΟΣ ΕΞΟΡΥΞΗΣ
ΜΟΝΤΕΛΟΥ ΠΡΟΣΩΠΙΚΟΤΗΤΑΣ ΜΕΣΩ ΤΟΥ ΚΟΙΝΩΝΙΚΟΥ
ΔΙΚΤΥΟΥ FACEBOOK**

Κωνσταντίνος Ηροδότου

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΥΠΡΟΥ



ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ

Μάιος 2014

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΥΠΡΟΥ

ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ

Αυτοματοποιημένη Μέθοδος Εξόρυξης Μοντέλου Προσωπικότητας μέσω του Κοινωνικού Δικτύου Facebook

Κωνσταντίνος Ηροδότου

Επιβλέπων Καθηγητής

Δρ. Γιώργος Σαμάρας

Η Ατομική Διπλωματική Εργασία υποβλήθηκε προς μερική εκπλήρωση των απαιτήσεων
απόκτησης του πτυχίου Πληροφορικής του Τμήματος Πληροφορικής του Πανεπιστημίου
Κύπρου

Μάιος 2014

Ευχαριστίες

Θα ήθελα να ευχαριστήσω τον επιβλέπων καθηγητή της διπλωματικής μου εργασίας Καθ. Γιώργο Σαμάρα ο οποίος μου έδωσε την ευκαιρία να συνεργαστούμε και μου έδειξε εμπιστοσύνη να εκτελέσω την έρευνα που υλοποιείται στην παρούσα διπλωματική εργασία. Επίσης θα ήθελα να ευχαριστήσω τον διδακτορικό φοιτητή Μάριο Belk, για την βοήθεια και καθοδήγηση που μου πρόσφερε κατά τη διάρκεια της εκπόνησης της παρούσας διπλωματικής εργασίας.

Οφείλω όμως το μεγαλύτερο ευχαριστώ στη Δρ. Στυλιανή Κλεάνθους η οποία μου πρόσφερε τεράστια στήριξη και εμπιστοσύνη. Την ευχαριστώ για τις απεριόριστες ώρες βοήθειας και καθοδήγησης καθώς επίσης και αμέτρητες καρποφόρες συζητήσεις που βοήθησαν στην εξέλιξη της δημιουργίας της ερευνητικής προσέγγισης. Η εμπιστοσύνη που μου έδειξε έπαιξε καθοριστικό ρόλο για την ώθηση που χρειαζόμουν να ολοκληρώσω την παρούσα διπλωματική εργασία. Η συμβολή της ήταν καθοριστική στην ολοκλήρωση του έργου, προσφέροντας αρκετό από τον προσωπικό της χρόνο.

Ακόμη ένα ευχαριστώ θα ήθελα να πω στους συμφοιτητές μου Χρίστο Πιτσιλλίδη, Αντρέα Χ' Δημήτρη και Αργύρη Κωνσταντινίδη για τη στήριξη και τη άψογη συνεργασία που είχαμε κατά την εκπόνηση των διπλωματικών μας.

Τέλος θα ήθελα να ευχαριστήσω τους γονείς μου Μάριο και Μαρία για συνεχή ενθάρρυνση, την υποστήριξη και κατανόηση που μου έδειχναν τα τελευταία 4 χρόνια.

Περίληψη

Τα τελευταία 5 χρόνια έχει παρατηρηθεί ραγδαία ανάπτυξη στη χρήση των κοινωνικών δικτύων. Πλέον στις μέρες μας, πολύ λίγα άτομα που χρησιμοποιούν το διαδίκτυο δεν έχουν τουλάχιστο ένα προφίλ σε κάποιο μέσο κοινωνικής δικτύωσης. Αυτό έχει ως αποτέλεσμα την δημιουργία τεράστιας ποσότητας δεδομένων δραστηριότητας, τα οποία μένουν αναξιοποίητα. Το Facebook, με το οποίο ασχολείται και η παρούσα διπλωματική εργασία, είναι το πιο διαδεδομένο κοινωνικό δίκτυο στις μέρες μας με περισσότερους από ένα εκατομμύριο χρήστες εκ των οποίων το 48% χρησιμοποιούν το κοινωνικό δίκτυο καθημερινά. Τα τελευταία 5 περίπου χρόνια ο κλάδος της πληροφορικής έρχεται για να δώσει έξυπνες και αυτοματοποιημένες μεθόδους εξαγωγής πληροφοριών από τα κοινωνικά δίκτυα. Έρευνες στο κλάδο της ψυχολογίας αποκάλυψαν ότι οι διαστάσεις της προσωπικότητας του “Big Five” είναι καλύτερος δείκτης της χρήσης του διαδικτύου παρά τα γνωστικά χαρακτηριστικά. Έτσι οδεύσαμε προς αυτή την κατεύθυνση.

Μέχρι σήμερα οι προσπάθειες που έγιναν από το κλάδο της πληροφορικής είχαν μελετήσει κάποια από τα δεδομένα δραστηριότητας των χρηστών και πως αυτά συσχετίζονται με τις διαστάσεις προσωπικότητας του “Big Five”. Η δική μας προσέγγιση διαφέρει από τις μέχρι τώρα ως εξής: (i) μελετούσε περισσότερα και πιο διαφορετικά δεδομένα δραστηριότητας, (ii) υλοποιούσε αυτοματοποιημένη δημιουργία ενός μοντέλου προσωπικότητας βάσει των διαστάσεων του “Big Five”, (iii) σύγκριση του μοντέλου αυτού μαζί με ένα πιο ακριβή μοντέλο που προέρχεται από ειδικό ερωτηματολόγιο καθώς επίσης και (iv) παροχή εξατομικευμένων πληροφοριών και εισηγήσεων στο χρήστη βάση των δεδομένων που μας παρέχει.

Η αξιολόγηση που εκτελέστηκε στην έρευνα της παρούσας διπλωματικής εργασίας, είχε ως σκοπό να εξακριβώσει κατά πόσο είναι δυνατό παίρνοντας δεδομένα δραστηριότητας ενός χρήστη από το κοινωνικό δίκτυο Facebook, να εξαχθούν συμπεράσματα ως προς τα χαρακτηριστικά της προσωπικότητας του. Τα αποτελέσματα έδειξαν ότι όντως μπορούμε να εξάγουμε σημαντικά συμπεράσματα για μεγάλο ποσοστό των διαστάσεων προσωπικότητας του “Big Five”. Επιπλέον η έρευνα αυτή συνεισφέρει στην ερευνητική κοινότητα με τα εξής: αποδείχτηκε ότι με τα δεδομένα που μαζεύτηκαν από το Facebook μπορεί να αναπτυχθεί ένα μοντέλο χρήστη, το οποίο να εφαρμοστεί για παροχή εξατομικευμένων πληροφοριών και συστάσεων.

Περιεχόμενα

Περιεχόμενα	iii
ΚΕΦΑΛΑΙΟ 1 ΕΙΣΑΓΩΓΗ	1
1.1 Ορισμός Προβλήματος	1
1.2 Σκοπός Διπλωματικής.....	2
1.3 Περιγραφή Κεφαλαίων	3
ΚΕΦΑΛΑΙΟ 2 ΣΧΕΤΙΚΗ ΒΙΒΛΙΟΓΡΑΦΙΑ.....	4
2.1 Εισαγωγή.....	4
2.2 Γνωρίσματα Προσωπικότητας του “Big Five” μέσα από την Δραστηριότητα του Χρήστη σε ένα Κοινωνικό Δίκτυο	5
2.2.1 Γιατί Γνωρίσματα Προσωπικότητας;	5
2.2.2 Τα Γνωρίσματα Προσωπικότητας του “Big Five”	6
2.2.3 Πως συσχετίζονται τα Γνωρίσματα Προσωπικότητας με τις δραστηριότητες του Facebook	8
2.3 Εξαγωγή και ανάλυση των δεδομένων του Facebook τα οποία σχετίζονται με τα Γνωρίσματα του “Big Five”	10
2.3.1 Ημι-αυτοματοποιημένες Τεχνικές Συλλογής Δεδομένων	11
2.3.2 Αυτοματοποιημένες Τεχνικές Συλλογής Δεδομένων	14
2.3.3 Η προσέγγιση που ακολουθεί η εργασία	15
2.4 PersonaWeb	16
2.5 Επίλογος.....	16
ΚΕΦΑΛΑΙΟ 3 ΥΠΟΛΟΓΙΣΤΙΚΟ ΠΛΑΙΣΙΟ (ΑΡΧΙΤΕΚΤΟΝΙΚΗ)	18
3.1 Εισαγωγή.....	18
3.2 Το Κοινωνικό Δίκτυο Facebook.....	19
3.3 Μοντελοποίηση Χρήστη	19

3.4 Υπολογιστικό Πλαίσιο Μοντελοποίησης Χρήστη	20
3.4.1 Περιγραφή Μηχανισμού Εξόρυξης Πληροφοριών του Χρήστη από το Facebook	21
3.4.2 Μοντέλο Χρήστη	24
3.4.3 Εφαρμογή Μοντέλου Προσωπικότητας για την Παροχή Πρόσθετων Πληροφοριών στο Χρήστη ...	26
3.5 Τεχνολογίες	27
3.6 Εργαλεία ανάπτυξης λογισμικού	29
3.7 Επίλογος.....	30
 ΚΕΦΑΛΑΙΟ 4 ΥΛΟΠΟΙΗΣΗ.....	 31
4.1 Εισαγωγή.....	32
4.2 Απαιτήσεις του Συστήματος	32
4.3 Προδιαγραφές	33
4.3.1 Προδιαγραφές από την Μεριά του Εξυπηρετητή	33
4.3.2 Προδιαγραφές από την Μεριά του Χρήστη	34
4.4 Προαπαιτήσεις του Συστήματος.....	35
4.5 Τεχνική Υλοποίηση του Συστήματος.....	35
4.5.1 Συλλογή Δεδομένων.....	36
4.5.2 Επεξεργασία των Δεδομένων	41
4.6 Επίλογος.....	56
 ΚΕΦΑΛΑΙΟ 5 ΕΦΑΡΜΟΓΗ PERSONA WEB ΣΤΟ FACEBOOK.....	 57
5.1 Εισαγωγή.....	57
5.2 Δημιουργία Εφαρμογής	57
5.3 Παρουσίαση Λειτουργιών και Πληροφοριών στο Χρήστη	60
5.4 Εξωτερικές Λειτουργίες Εφαρμογής	70
5.5 Επίλογος.....	72
 ΚΕΦΑΛΑΙΟ 6 ΑΞΙΟΛΟΓΗΣΗ.....	 73

6.1 Εισαγωγή.....	73
6.2 Σκοπός Αξιολόγησης.....	74
6.3 Περιγραφή Διαδικασίας Αξιολόγησης.....	75
6.3.1 Συγκέντρωση του Δείγματος.....	75
6.4 Περιγραφή Δείγματος	76
6.5 Πειραματική Διάταξη	78
6.6 Ανάλυση Δεδομένων	79
6.7 Επίλογος.....	83
ΚΕΦΑΛΑΙΟ 7 ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΜΕΛΛΟΝΤΙΚΕΣ ΠΡΟΟΠΤΙΚΕΣ	85
7.1 Συμπεράσματα	85
7.2 Περιορισμοί.....	86
7.3 Μελλοντικές Προτάσεις.....	88
ΒΙΒΛΙΟΓΡΑΦΙΑ.....	90
ΠΑΡΑΡΤΗΜΑ Α.....	Α-1
ΠΑΡΑΡΤΗΜΑ Β.....	Β-1
ΠΑΡΑΡΤΗΜΑ Γ.....	Γ-1
ΠΑΡΑΡΤΗΜΑ Δ.....	Δ-1

Σχήματα

Σχήμα 3-1 Υπολογιστικό Πλαίσιο	21
Σχήμα 4-1 Διάγραμμα επικοινωνίας της εφαρμογής	36
Σχήμα 4-2 Σχηματική περιγραφή της βάσης δεδομένων	38
Σχήμα 6-1 Δημογραφικά Στοιχεία Συμμετεχόντων – Φύλο	77
Σχήμα 6-2 Δημογραφικά Συμμετεχόντων - Ηλικίες	78

Πίνακες

Πίνακας 3-1 - Πίνακες από το Facebook που χρησιμοποιήθηκαν για τις ανάγκες της εργασίας	23
Πίνακας 3-2 Δεδομένα που αποτελούν το μοντέλο χρήστη	26
Πίνακας 4-1 Διαχωρισμός Δεδομένων στις Διαστάσεις της Προσωπικότητας	52
Πίνακας 6-1 Αποτελέσματα στατιστικής ανάλυσης - Εξωστρέφεια	80
Πίνακας 6-2 Αποτελέσματα στατιστικής ανάλυσης - Τερπνότητα	81
Πίνακας 6-3 Αποτελέσματα στατιστικής ανάλυσης - Ευσυνειδησία	81
Πίνακας 6-4 Αποτελέσματα στατιστικής ανάλυσης - Νεύρωση	82
Πίνακας 6-5 Αποτελέσματα στατιστικής ανάλυσης - Διαφάνεια	82

Εικόνες

Εικόνα 3-1 Παράδειγμα FQL με κώδικα PHP	28
Εικόνα 5-1 Δημιουργία νέας εφαρμογής στο Facebook.....	58
Εικόνα 5-2 Ανάκτηση Μοναδικών Κωδικών Εφαρμογής.....	59
Εικόνα 5-3 Συμπλήρωση στοιχείων Εφαρμογής – Α Μέρος.....	59
Εικόνα 5-4 Συμπλήρωση στοιχείων Εφαρμογής - Β Μερους.....	60
Εικόνα 5-5 Παράθυρο ταυτοποίησης του χρήστη με το λογαριασμό του στο Facebook.....	61
Εικόνα 5-6 Συμβόλαιο Εφαρμογής - Χρήστη	61
Εικόνα 5-7 Αναδυόμενο παράθυρο όπου ο χρήστης μπορεί να κοινοποιήσει τα αποτελέσματα του	62
Εικόνα 5-8 Αριστερά: Φόρμα παρουσίασης προσωπικών στοιχείων χρήστη. Δεξιά: Φόρμα όπου ο χρήστης μπορεί να μεταβεί στην ιστοσελίδα του PersonaWeb.....	63
Εικόνα 5-9 Συγκεντρωτικές πληροφορίες σε σχέση με τα likes του χρήστη σε αντικείμενα στο Facebook.....	64
Εικόνα 5-10 Συγκεντρωτικές πληροφορίες σε σχέση με τις σελίδες στο Facebook και ανάλογα με τον τύπο τους, τις οποίες ο χρήστης έκανε like	65
Εικόνα 5-11 Ο χρήστης μπορεί να δει πληροφορίες σε σχέση με τον ιδιοκτήτη των περισσότερων αντικειμένων των οποίων έκανε like	65
Εικόνα 5-12 Ο χρήστης μπορεί να δει τον ιδιοκτήτη των περισσότερων αντικειμένων στα οποία είναι tagged ο ίδιος.....	66
Εικόνα 5-13 Πληροφορίες σχετικά με τις εκδηλώσεις που παρευρέθηκε ο χρήστης.....	66
Εικόνα 5-14 Μοντέλο προσωπικότητας χρήστη που δημιουργείται βάσει την δραστηριότητα και τα δεδομένα του στο Facebook	67
Εικόνα 5-15 Συγκριτικά αποτελέσματα απο το μοντέλο προσωπικότητας χρήστη και των αποτελεσμάτων απο το ερωτηματολόγιο "Big Five"	68
Εικόνα 5-16 Πληροφορίες για κάθε μία από τις πέντε διαστάσεις του μοντέλου προσωπικότητας του “Big Five”	69
Εικόνα 5-17 Εικόνες προφίλ των χρηστών που ο χρήστης έχει ομοιότητα ως προς τη δραστηριότητα του στο Facebook.....	70
Εικόνα 5-18 Συστάσεις σε μια σειρά από σελίδες οι οποίες μπορεί να ενδιαφέρουν τον χρήστη	70
Εικόνα 5-19 Φόρμα Εισόδου μέσω PersonaWeb	71
Εικόνα 5-20 Σύνδεσμος επιστροφής στην εφαρμογή	72

Κεφάλαιο 1

Εισαγωγή

1.1 Ορισμός Προβλήματος	1
1.2 Σκοπός Διπλωματικής.....	2
1.3 Περιγραφή Κεφαλαίων	3

1.1 Ορισμός Προβλήματος

Τα κοινωνικά μέσα δικτύωσης έχουν αναπτυχθεί ραγδαία τα τελευταία 5 χρόνια. Όλο και περισσότεροι άνθρωποι αποκτούν προφίλ τουλάχιστο σε ένα κοινωνικό δίκτυο όπως το Twitter ή το Facebook. Ο όγκος των δεδομένων που δημιουργείτε είναι τεράστιος και επιστήμονες από διάφορες ειδικότητες αποκτούν ενδιαφέρον προς την μελέτη και ανάλυση αυτών των δεδομένων ώστε να μπορέσουν να δώσουν κάποιο νόημα και να αντλήσουν συμπεράσματα σχετικά με την αλληλεπίδραση, τόσο των ανθρώπων μεταξύ τους, μέσω των κοινωνικών δικτύων, αλλά και με τα συστήματα αυτά.

Συγκεκριμένα, στο τομέα της ψυχολογίας οι ερευνητές προσπαθούν τα τελευταία χρόνια να κατανοήσουν πως η προσωπικότητα ενός χρήστη μπορεί να αντικατοπτρίζεται στην δραστηριότητα του στα κοινωνικά αυτά δίκτυα. Οι προσεγγίσεις αυτές επικεντρώνονται κυρίως στη χρήση ερωτηματολογίων όπου ο χρήστης θα πρέπει να συμπληρώσει στοιχεία σχετικά με τις δραστηριότητες του. Η μεθοδολογία αυτή βασίζεται στην προσωπική άποψη του χρήστη με αποτέλεσμα οι απαντήσεις που δίνει να είναι προκατειλημμένες και τα αποτελέσματα που αντλούνται να μην αντικατοπτρίζουν την πραγματικότητα.

Τα τελευταία πέντε περίπου χρόνια ο κλάδος της πληροφορικής άρχισε να συνεισφέρει στο τομέα αυτό με την δημιουργία έξυπνων και αυτοματοποιημένων μεθόδων εξόρυξης δεδομένων από τα κοινωνικά δίκτυα. Με τον τρόπο αυτό μπορούμε να καταλήξουμε σε συμπεράσματα για τον χρήστη μέσω της δραστηριότητας του χωρίς όμως να περιμένουμε ο ίδιος να συνεισφέρει με κάποιο τρόπο. Έτσι, η εμπειρία του χρήστη δεν αλλοιώνετε και την ίδια στιγμή οι ερευνητές μπορούν να προσφέρουν στο χρήστη υποστήριξη ως προς το έργο τους.

Μελετώντας την έρευνα που έχει εκπονηθεί στον τομέα αυτό, και παρουσιάζεται στο Κεφάλαιο 2, παρατηρήθηκε ότι οι μέχρι τώρα έρευνες αγνοούν πολλά από τα δεδομένα που δημιουργούνται από την καθημερινή δραστηριότητα του χρήστη στα κοινωνικά δίκτυα και θα μπορούσαν να εκμεταλλευτούν ως προς την παροχή υποστήριξης προς αυτόν.

1.2 Σκοπός Διπλωματικής

Σε αυτά τα πλαίσια η παρούσα διπλωματική εργασία καλείτε να καλύψει το κενό που υπάρχει στο τομέα αυτό με τη δημιουργία μιας εφαρμογής στο Facebook. Μέσω αυτής της εφαρμογής τα δεδομένα του χρήστη ανακτώνται και επεξεργάζονται με σκοπό την δημιουργία ενός μοντέλου χρήστη και την εφαρμογή του για παροχή εξατομικευμένων πληροφοριών στο χρήστη και εξαγωγή ενός μοντέλου προσωπικότητας με βάση τις διαστάσεις του “Big Five”.

Κατά συνέπεια τα ερευνητικά ερωτήματα που έρχεται να απαντήσει αυτή η μελέτη είναι τα ακόλουθα:

- i. Πως μπορούμε να αναπτύξουμε ένα μοντέλο χρήστη χρησιμοποιώντας δεδομένα δραστηριότητας από το Facebook;
- ii. Εφαρμόζοντας το μοντέλο χρήστη πώς μπορούμε να προσφέρουμε εξατομικευμένες πληροφορίες στο χρήστη ως προς την δραστηριότητα του στο Facebook;
- iii. Πως μπορούμε μέσω των δεδομένων δραστηριότητας στο Facebook να συμπεράνουμε την προσωπικότητα του χρήστη με βάση το μοντέλο “Big Five”;

1.3 Περιγραφή Κεφαλαίων

Αφού ορίσαμε το σκοπό αυτής της εργασίας και κλείνοντας την εισαγωγή ακολουθεί μια σύντομη περιγραφή των επόμενων κεφαλαίων.

Στο Κεφάλαιο 2 θα αναλύσουμε τα χαρακτηριστικά του μοντέλου προσωπικότητας “Big Five”, καθώς και κάποιες από τις τεχνικές που ήδη χρησιμοποιούνται για επίλυση του προβλήματος αυτού από συναφείς μελέτες.

Στη συνέχεια, στο Κεφάλαιο 3 περιγράφεται το υπολογιστικό πλαίσιο το οποίο ακολουθήσαμε για την επίτευξη του στόχου μας. Στο Κεφάλαιο 4 αναφέρονται οι απαιτήσεις, οι προαπαιτήσεις και οι προδιαγραφές του συστήματος και γίνεται μια πλήρη ανάλυση της τεχνικής υλοποίησης που εκτελέστηκε στα πλαίσια της παρούσας διπλωματικής εργασίας. Ακόμη στο Κεφάλαιο 5 αναφέρονται τα βήματα ανάπτυξης μιας εφαρμογής στο Facebook και εξηγούνται οι λειτουργίες που μπορεί να εκτελέσει ο χρήστης.

Ακολούθως, στο Κεφάλαιο 6 επεξηγεται ο τρόπος με τον οποίο εκτελέστηκε η πειραματική μελέτη καθώς επίσης παρουσιάζονται και αναλύονται τα αποτελέσματα αυτής. Τέλος, στο Κεφάλαιο 7 αναφέρονται κάποιοι από τους περιορισμούς υλοποίησης που συναντήσαμε, μελλοντική δουλειά η οποία μπορεί να αναπτυχθεί καθώς και τα συμπεράσματα τα οποία πήραμε από την μελέτη μας.

Κεφάλαιο 2

Σχετική Βιβλιογραφία

2.1 Εισαγωγή.....	4
2.2 Γνωρίσματα Προσωπικότητας του “Big Five” μέσα από την Δραστηριότητα του Χρήστη σε ένα Κοινωνικό Δίκτυο.....	5
2.2.1 Γιατί Γνωρίσματα Προσωπικότητας;.....	5
2.2.2 Τα Γνωρίσματα Προσωπικότητας του “Big Five”	6
2.2.3 Πως συσχετίζονται τα Γνωρίσματα Προσωπικότητας με τις δραστηριότητες του Facebook	8
2.3 Εξαγωγή και ανάλυση των δεδομένων του Facebook τα οποία σχετίζονται με τα Γνωρίσματα του “Big Five”	10
2.3.1 Ημι-αυτοματοποιημένες Τεχνικές Συλλογής Δεδομένων	11
2.3.2 Αυτοματοποιημένες Τεχνικές Συλλογής Δεδομένων	14
2.3.3 Η προσέγγιση που ακολουθεί η εργασία	15
2.4 PersonaWeb	16
2.5 Επίλογος.....	16

2.1 Εισαγωγή

Σε αυτό το κεφάλαιο θα αναλυθεί το ερευνητικό υπόβαθρο που έγινε η βάση πίσω από την ιδέα της παρούσας διπλωματικής εργασίας. Αρχικά, θα περιγράψουμε τι είναι τα γνωρίσματα προσωπικότητας του τεστ ψυχολογίας “Big Five” και πως μπορούν να εντοπιστούν μέσα από την δραστηριότητα του χρήστη σε ένα κοινωνικό δίκτυο. Στη συνέχεια θα αναφερθούμε σε παρόμοιες μελέτες οι οποίες προσπάθησαν αρχικά να αναλύσουν κάποια ακατέργαστα δεδομένα από το Facebook σε σχέση με τα αποτελέσματα του τεστ προσωπικότητας “Big Five”.

Ακολούθως, θα παρουσιαστεί η προσέγγιση που ακολουθήθηκε σε αυτή την εργασία και πως διαφέρει από τις προσεγγίσεις που αναφέρθηκαν στο τομέα αυτό.

2.2 Γνωρίσματα Προσωπικότητας του “Big Five” μέσα από την Δραστηριότητα του Χρήστη σε ένα Κοινωνικό Δίκτυο

Μέσα από τη δραστηριότητα του χρήστη σε ένα κοινωνικό δίκτυο και συγκεκριμένα στο Facebook μπορούμε να συμπεράνουμε το τύπο της προσωπικότητας που διαθέτει. Στο υποκεφάλαιο αυτό θα αναφερθούμε στα συγκεκριμένα γνωρίσματα του “Big Five” και πως αυτά συνδέονται με δραστηριότητες του χρήστη στο Facebook.

2.2.1 Γιατί Γνωρίσματα Προσωπικότητας;

Προτού αναφερθούμε στα γνωρίσματα προσωπικότητας ατόμων μέσα από τον φακό του Big Five πρέπει να αναφερθεί ότι η επιλογή των γνωρισμάτων προσωπικότητας δεν ήταν η αρχική επιλογή προσέγγισης που θα ακολουθούσαμε για ερμηνεία των ακατέργαστων δεδομένων τα οποία θα μπορούσε να προσφέρει το Facebook. Αρχική ιδέα ήταν η επεξεργασία των δεδομένων της δραστηριότητας του χρήστη Facebook με σκοπό να ερμηνεύσουμε τα γνωστικά του χαρακτηριστικά. Ακολουθώντας το ερευνητικό έργο των McElroy et al. [1], τα αποτελέσματα των οποίων αποδεικνύουν, μετά από έρευνα, πως οι γενικοί παράγοντες προσωπικότητας προβλέπουν πτυχές της χρήσης του διαδικτύου από ένα άτομο, σε αντίθεση με τα γνωστικά χαρακτηριστικά που όπως έχει αποδειχτεί δεν είναι το κατάλληλο μέσο για μελέτη. Τα γνωστικά στυλ αποτυγχάνουν να εξηγήσουν σημαντικές ποσότητες διακύμανσης πέρα από τις μεταβλητές ελέγχου του άγχους, της αυτό-αποτελεσματικότητας και του φύλου.

Επιπρόσθετα, οι Moore και McElroy [2], μελετούν κατά πόσο η προσωπικότητα επηρεάζει τις δραστηριότητες που κάνουν οι χρήστες στο κοινωνικό δίκτυο σε σχέση με το τι δεδομένα παρουσιάζουν για εκείνους και τους φίλους τους. Επίσης εξετάζουν τις συνέπειες που έχει η προσωπικότητα ενός ατόμου, σε σχέση με την αίσθηση λύπης και της χρήσης που κάνει στο κοινωνικό αυτό δίκτυο. Σε αυτή τη μελέτη υποστηρίζεται πως η προσωπικότητα είναι πολύ

καλύτερη όσο αφορά την πρόβλεψη της χρήσης του διαδικτύου παρά τα γνωστικά στυλ. Επιπλέον αποτελέσματα από τις πιο πάνω μελέτες θα αναφερθούν στη συνέχεια.

2.2.2 Τα Γνωρίσματα Προσωπικότητας του “Big Five”

Στον τομέα της ψυχολογίας, τα χαρακτηριστικά της προσωπικότητας “Big Five” είναι πέντε γενικοί τομείς, ή θέτοντας το καλύτερα, πέντε διαστάσεις που χρησιμοποιούνται για να περιγράψουν την ανθρώπινη προσωπικότητα [3]. Οι πέντε μεγάλες αυτές διαστάσεις όπως έχουν διατυπωθεί από τον Goldberg [3] είναι η διαφάνεια (openness), ευσυνειδησία (conscientiousness), εξωστρέφεια (extraversion), τερπνότητα (agreeableness), και νεύρωση (neuroticism).

Το μοντέλο του “Big Five” έχει την ιδιότητα να αντιπροσωπεύει διαφορετικά χαρακτηριστικά της προσωπικότητας χωρίς επικάλυψη. Αυτή η δομή με πέντε παράγοντες έχει εξεταστεί σε ένα ευρύ φάσμα συμμετεχόντων διαφόρων ηλικιών και από διάφορους πολιτισμούς. Το μοντέλο “Big Five” ορίστηκε από διάφορες ανεξάρτητες ομάδες ερευνητών. Αρχικά μελέτησαν διάφορα γνωστά χαρακτηριστικά της προσωπικότητας και στη συνέχεια αναλύοντας και αποκτώντας εκατοντάδες μετρήσεις από τους παράγοντες των χαρακτηριστικών αυτών κατέληξαν σε μια συγκεκριμένη δομή.

Το αρχικό μοντέλο προωθήθηκε από τον Ernest Tupes και Raymond Christal το 1961 [4], αλλά δεν κατάφεραν να πείσουν το ακαδημαϊκό κοινό μέχρι το 1980. Το 1990, ο JM Digman εξέλιξε το μοντέλο πέντε παραγόντων της προσωπικότητας, το οποίο ο Luis Goldberg [3] επέκτεινε στη σημερινή γνωστή δομή. Αυτοί οι πέντε πρωταρχικοί τομείς περιέχουν τα πιο γνωστά χαρακτηριστικά της προσωπικότητας και παρέχουν ένα πλούσιο και ολοκληρωμένο εννοιολογικό πλαίσιο για την ενσωμάτωση όλων των πορισμάτων της έρευνας και της θεωρίας στην ψυχολογία στο τομέα της προσωπικότητας.

Οι πέντε διαστάσεις του μοντέλου “Big Five” ερμηνεύονται ως εξής:

Εξωστρέφεια (Extraversion): Είναι “η πράξη, η κατάσταση, ή η συνήθεια του να ασχολείται κάποιος κατά κύριο λόγο με την απόκτηση και την ικανοποίηση από αυτό που είναι έξω από τον εαυτό του”. Οι εξωστρεφείς τείνουν να απολαμβάνουν ανθρώπινες αλληλεπιδράσεις και να είναι ενθουσιώδεις, ομιλητικοί, δυναμικοί, και αγελαίοι.

Ενεργοποιούνται και ευδοκιμούν όταν είναι γύρω από άλλους ανθρώπους. Παίρνουν ευχαρίστηση σε δραστηριότητες που αφορούν μεγάλες κοινωνικές εκδηλώσεις, όπως πάρτι, δραστηριότητες της κοινότητας, δημόσιες διαδηλώσεις, και ίσως των επιχειρήσεων ή των πολιτικών ομάδων. Ένα εξωστρεφές άτομο είναι πιο επιρρεπές στην πλήξη όταν είναι μόνος τους.

Τερπνότητα (Agreeableness): Είναι ένα χαρακτηριστικό της προσωπικότητας που εκδηλώνεται σε επιμέρους χαρακτηριστικά των συμπεριφορών των ατόμων που διαθέτουν αυτό το χαρακτηριστικό όπως για παράδειγμα άτομα που θεωρούνται ως ευγενικά, συμπονετικά, συνεργάσιμα, ζεστά και διακριτικά. Στη σύγχρονη ψυχολογία προσωπικότητας, η τερπνότητα αντανakλά τις ατομικές διαφορές στη συνεργασία και την κοινωνική αρμονία. Οι άνθρωποι που παρουσιάζουν υψηλές επιδόσεις σε αυτήν την διάσταση τείνουν να πιστεύουν ότι οι περισσότεροι άνθρωποι είναι τίμιοι και αξιόπιστοι.

Ευσυνειδησία (Conscientiousness): Είναι το χαρακτηριστικό της προσωπικότητας που ορίζεται ως το να είσαι εξονυχιστικός, προσεκτικός, ή σε εγρήγορση. Υπονοεί την επιθυμία να κάνει μια εργασία καλά. Ευσυνείδητοι άνθρωποι είναι αποτελεσματικοί και οργανωμένοι σε αντίθεση με χαλαροί και απρόσεκτοι. Εκθέτουν την τάση να δείχνουν αυτοπειθαρχία, να ενεργούν σύμφωνα με το νόμο, και στοχεύουν για επιτεύγματα. Εμφανίζουν προγραμματισμένη και όχι αυθόρμητη συμπεριφορά και είναι γενικά οργανωμένοι και αξιόπιστοι. Εκδηλώνουν χαρακτηριστικές συμπεριφορές όπως είναι το να είναι τακτοποιημένοι, και συστηματικοί και επίσης έχουν στοιχεία όπως η σύνεση, η πληρότητα και η διαβούλευση (η τάση να σκέφτεται προσεκτικά πριν να ενεργήσει). Οι έχοντας ευσυνειδησία είναι άνθρωποι οι οποίοι, όπως θα λέγαμε παραδοσιακά, «έχοντα χαρακτήρα». Ευσυνείδητα άτομα γενικά εργάζονται σκληρά και είναι αξιόπιστα. Όταν βρεθούν σε μια ακραία κατάσταση μπορεί επίσης να ονομαστούν «εργασιομανείς», τελειομανείς, και με ψυχαναγκαστική συμπεριφορά προς το άτομο τους. Οι άνθρωποι που σκοράρουν χαμηλά στην ευσυνειδησία τείνουν να είναι πιο χαλαροί, με λιγότερο προσανατολισμό σε στόχους, και δεν οδηγούνται από την επιτυχία. Επίσης είναι πιο πιθανό να εμπλακούν σε αντικοινωνική και εγκληματική συμπεριφορά από κάποιο που σκοράρει ψηλά σε αυτή.

Νεύρωση (Neuroticism): Είναι η τάση να βιώνει κάποιος έντονα αρνητικά συναισθήματα, όπως ο θυμός, το άγχος ή η κατάθλιψη. Μπορεί να ονομαστεί και Συναισθηματική Αστάθεια ή

αντίστροφα αναφέρεται ως συναισθηματική σταθερότητα. Σύμφωνα με τη θεωρία προσωπικότητας του Eysenck, η νευρώση είναι αλληλένδετη με χαμηλή ανοχή στο άγχος ή σε αρνητικά ερεθίσματα. Εκείνοι που παρουσιάζουν υψηλές επιδόσεις σε νευρώση είναι συναισθηματικά αντιδραστικοί και ευάλωτοι στο στρες. Είναι πιθανό να ερμηνεύουν συνηθισμένες καταστάσεις ως απειλές, καθώς και μικρές απογοητεύσεις σε απελπιστικά δύσκολες. Οι αρνητικές συναισθηματικές αντιδράσεις τους έχουν την τάση να επιμένουν για ασυνήθιστα μεγάλο χρονικό διάστημα, πράγμα που σημαίνει ότι είναι συχνά σε κακή διάθεση. Αυτά τα προβλήματα στην ρύθμιση των συναισθημάτων μπορούν να μειώσουν την ικανότητα ενός ατόμου να σκέφτεται καθαρά, να παίρνουν αποφάσεις και να αντιμετωπίσουν αποτελεσματικά το άγχος. Η έλλειψη ικανοποίησης για τα επιτεύγματά της ζωής ενός ατόμου μπορούν να συσχετίζονται με υψηλές βαθμολογίες στη νευρώση και αυξάνει τις πιθανότητες ενός ατόμου να πέσουν στην κλινική κατάθλιψη.

Διαφάνεια (Openness to experience): Άτομα που σκοράρουν ψηλά στη Διαφάνεια έχουν μια γενική εκτίμηση για τέχνη, συγκίνηση, περιπέτεια, ασυνήθιστες ιδέες, φαντασία, περιέργεια, και για ποικιλία εμπειριών. Οι άνθρωποι που είναι ανοιχτοί στην εμπειρία είναι διανοητικά περίεργοι, συγκινούνται από τη τέχνη, και είναι ευαίσθητοι στην ομορφιά. Σε σύγκριση με κλειστούς ανθρώπους, τείνουν να είναι πιο δημιουργικοί και να έχουν καλύτερη επίγνωση των συναισθημάτων τους. Ωστόσο, ένα άτομο που ανήκει σε αυτή τη κατηγορία μπορεί να ενδιαφέρεται να μάθει και να εξερευνήσει νέους πολιτισμούς αλλά, να μην τον ενδιαφέρουν θέματα όπως η τέχνη και η ποίηση. Άλλο ένα χαρακτηριστικό των ατόμων αυτών είναι πως διαθέτουν ένα νοητό μηχανισμό που τους επιτρέπει να σκέφτονται με σύμβολα ή γενικά πιο αφηρημένα. Οι άνθρωποι με χαμηλή βαθμολογία στη Διαφάνεια, τείνουν να έχουν πιο συμβατικά, παραδοσιακά ενδιαφέροντα. Προτιμούν το απλό, το άμεσο και το προφανές παρά το πολύπλοκο και το ασαφές.

2.2.3 Πως συσχετίζονται τα Γνωρίσματα Προσωπικότητας με τις δραστηριότητες του Facebook

Ο σκοπός της εργασίας αυτής είναι να μελετήσει κατά πόσο μπορούμε από τις δραστηριότητες του χρήστη στο Facebook να εξάγουμε κάποια συμπεράσματα σε σχέση με την προσωπικότητα του. Κατά συνέπεια, έγινε μελέτη κάποιων προϋπάρχων ερευνών, οι οποίες

προσπαθούσαν να συνδέσουν κάποιες από τις δραστηριότητες που μπορεί να εκτελέσει και τα αντικείμενα που μπορεί να αλληλεπιδράσει ένας χρήστης στο Facebook, σε σχέση με τα γνωρίσματα προσωπικότητας του “Big Five” τα οποία αναλύσαμε πιο πάνω. Τα αποτελέσματα είναι τα εξής:

Εξωστρέφεια (Extraversion): Τα άτομα αυτής της κατηγορίας έχουν την τάση να περνούν περισσότερες ώρες μέσα σε κοινωνικά δίκτυα και προτιμούν να έχουν ένα μεγαλύτερο αποτύπωμα γι’ αυτό και στο Facebook ανήκουν σε παραπάνω groups και έχουν περισσότερους φίλους. Ακόμη τείνουν να κάνουν περισσότερες συνομιλίες με φίλους τους στο Facebook, ήταν πιο πιθανό να μεταδίδουν τις δραστηριότητες και τις εκδηλώσεις τους καθώς και να δημοσιεύουν περισσότερες φωτογραφίες. Άτομα αυτού του τύπου ανήκουν σε περισσότερα groups και έχουν μεγαλύτερο αριθμό φίλων. Επίσης, κάποιος εξωστρεφής χρήστης του Facebook θα δημοσιεύει περισσότερες καταστάσεις του εαυτού του και θα κάνει πολλά like σε δημοσιεύσεις των φίλων του στο Facebook [5].

Τερπνότητα (Agreeableness): Για αυτή την διάσταση της προσωπικότητας, οι [2] παρατήρησαν πως άτομα που σκοράρουν ψηλά σε αυτή προτιμούν να επικοινωνούν περισσότερο με προσωπικά μηνύματα στο Facebook με τους φίλους τους και να ασχολούνται περισσότερο με διαδικτυακά παιχνίδια που προσφέρει το Facebook. Επίσης αναφέρεται ότι άτομα που εμπίπτουν σε αυτή την κατηγορία δεν χρησιμοποιούν το κοινωνικό δίκτυο για μεγάλο χρονικό διάστημα και αποφεύγουν να κάνουν δημοσιεύσεις για φίλους τους. Επιπρόσθετα, με βάση τους [5] άτομα που χαρακτηρίζονται από τερπνότητα τείνουν να μην χρησιμοποιούν πολύ την λειτουργία like του Facebook.

Ευσυνειδησία (Conscientiousness): Οι Moore και Kelly [2] καταγράφουν ότι άτομα με ψηλή ευσυνειδησία, επειδή είναι πιο προσηλωμένα σε στόχους, η δραστηριότητα τους στο Facebook θα είναι μειωμένη. Δηλαδή, θα χρησιμοποιούν λιγότερη ώρα το Facebook, θα έχουν λιγότερους φίλους και θα δημοσιεύουν λιγότερες φωτογραφίες και καταστάσεις. Οι [6] από την δική τους έρευνα εντόπισαν σε αντίθεση ότι άτομα που ανήκουν σε αυτή τη κατηγορία θα χρησιμοποιούν λιγότερο την λειτουργία like, θα ανήκουν σε λίγες ομάδες (groups) και θα δημοσιεύουν αρκετές φωτογραφίες.

Νεύρωση (Neuroticism): Άτομα που χαρακτηρίζονται από νεύρωση, λόγω των αρνητικών συναισθημάτων που υποφέρουν, τείνουν να είναι πιο συχνόι χρήστες του Facebook λόγω του ότι

θέλουν να ελέγχουν της πληροφορίες που αφορούν εκείνους και το περιβάλλον τους [2]. Έτσι και οι πιο συχνή τους δραστηριότητα είναι να δημοσιεύουν τις πληροφορίες, ή της καταστάσεις, που θέλουν, ενώ ταυτόχρονα θα αποφεύγουν να δημοσιεύουν φωτογραφίες. Άτομα που ανήκουν στην κατηγορία της νεύρωσης τείνουν να έχουν λιγότερους φίλους στο Facebook αλλά, ταυτόχρονα, να χρησιμοποιούν συχνά την λειτουργία like σε δημοσιεύσεις αυτών [6].

Διαφάνεια (Openness to experience): Άτομα που διακρίνονται από διαφάνεια, λόγο της δεκτικότητας που διαθέτουν για νέες εμπειρίες θα τείνουν να χρησιμοποιούν το κοινωνικό δίκτυο περισσότερη ώρα για να αποκτήσουν νέες εμπειρίες και θα δημοσιεύουν συχνότερα της δικές τους καινούργιες καταστάσεις [2]. Επίσης, τείνουν να χρησιμοποιούν συχνά τη λειτουργία like σε ότι θεωρούν ενδιαφέρον, θα δημοσιεύουν πολύ τις καταστάσεις τους και θα εντάσσονται σε πολλές ομάδες [6].

2.3 Εξαγωγή και ανάλυση των δεδομένων του Facebook τα οποία σχετίζονται με τα Γνωρίσματα του “Big Five”

Σε αυτό το υποκεφάλαιο θα αναφέρουμε κάποιους από τους βασικότερους τρόπους προσέγγισης άλλων ερευνητικών ομάδων καθώς και πως διαφέρει στην προσέγγιση της η παρούσα διπλωματική εργασία.

Λόγω της ραγδαίας ανάπτυξης που έχει το Facebook, σε σχέση με άλλα κοινωνικά δίκτυα, και λόγω της τεράστιας ποσότητας πληροφοριών που διαθέτει πλέον για τους χρήστες του, πολλές ερευνητικές ομάδες έχουν προσπαθήσει να επεξεργαστούν αυτά τα δεδομένα και να εξαγάγουν από αυτά συμπεράσματα σε σχέση με την προσωπικότητα ή άλλους τομείς της ψυχολογίας και κοινωνιολογίας. Για την επίτευξη αυτού του στόχου, οι διάφορες ερευνητικές ομάδες χρησιμοποιούν 2 κύριους τρόπους προσέγγισης. Οι δύο αυτοί τρόποι είναι α) Ημι-αυτοματοποιημένα - χρησιμοποιώντας ανίχνευση δεδομένων σε δημόσια δεδομένα και ερωτηματολόγια για τη συλλογή των δεδομένων του Facebook και β) Αυτοματοποιημένα - με την δημιουργία μιας εφαρμογής στο Facebook, η οποία μετά από την άδεια του χρήστη μαζεύει δεδομένα από τη βάση δεδομένων που διαθέτει στους προγραμματιστές του το Facebook.

Επίσης ακόμη μια μέθοδος που χρησιμοποίησαν κάποιες από τις ερευνητικές ομάδες που συναντήσαμε στην έρευνα μας όπως οι [5, 7] , χρησιμοποίησαν έτοιμα δεδομένα για χρήστες τα οποία διαθέτονται από το myPersonality project [8].

myPersonality project: Το ερευνητικό έργο myPersonality αναπτύχθηκε σαν μια εφαρμογή στο Facebook η οποία υποχρέωνε στην αρχή τους χρήστες να συμπληρώσουν το “Big Five” ερωτηματολόγιο και ακολούθως για όσους χρήστες έδιναν την άδεια τους, μάζευε κάποια δεδομένα από το λογαριασμό τους στο Facebook. Τέθηκε σε λειτουργία τον Ιούνιο του 2007 και ίσχυε μέχρι το 2011. Το ελάττωμα αυτής της βάσης δεδομένων είναι πως ένα μεγάλο ποσοστό των χρηστών δεν έχουν δώσει την άδεια για όλα τα στοιχεία και δεν μπορούν να χρησιμοποιηθούν για δημιουργία συγκριτικών αποτελεσμάτων. Επίσης τα δεδομένα τα οποία δημιουργούνται και αποθηκεύονται στη βάση δεδομένων του Facebook το 2007 είναι εκθετικά λιγότερα από τα σημερινά δεδομένα. Αυτό γιατί με το πέρασμα του χρόνου ο κόσμος χρησιμοποιεί περισσότερο τα κοινωνικά δίκτυα και αφού δεν έχουν περάσει και πολλά χρόνια που το διαδίκτυο έγινε τόσο εύκολα προσβάσιμο. Επιπρόσθετα, το Facebook άρχισε να αποκτά τον τεράστιο αριθμό χρηστών του από το 2008 και μετά, έτσι πολλές από τις μετρικές τις οποίες συλλέχτηκαν το 2007 δεν θα αντιπροσωπεύει τις σημερινές μετρικές που δημιουργεί ο μέσος χρήστης του Facebook..

2.3.1 Ημι-αυτοματοποιημένες Τεχνικές Συλλογής Δεδομένων

Το πλεονέκτημα της μεθόδου αυτής είναι ότι για να συλλέξει τα δεδομένα από τους χρήστες δεν χρειάζεται καμία ενέργεια από αυτούς. Σε αντίθεση, το κύριο μειονέκτημα είναι πως τα ανοικτά δεδομένα τα οποία οι ερευνητές μπορούν να συλλέξουν χωρίς την έγκριση του χρήστη, μειώνονται κατά πολύ με τη πάροδο του χρόνου, αφού το Facebook εφαρμόζει όλο και πιο περιοριστικές πολιτικές προστασίας των ιδιωτικών δεδομένων. Επίσης, τα αποτελέσματα από ερωτηματολόγια είναι προκατειλημμένα ως προς την άποψη που έχουν οι χρήστες για τον εαυτό τους. Συνήθως αυτή την προσέγγιση χρησιμοποιούν ερευνητικές ομάδες του τομέα της ψυχολογίας όπου το τεχνολογικό τους υπόβαθρο δεν τους βοηθά για κάτι πιο αυτοματοποιημένο.

Χρησιμοποιώντας αυτή την προσέγγιση, οι Randall et al. [9] μάζεψαν από τα ανοικτά δεδομένα των χρηστών οι οποίοι έλαβαν μέρος στη μελέτη τους, 32 αριθμητικά, και

δημογραφικά δεδομένα, από διάφορα μέρη του προφίλ τους. Παραδείγματα αυτών των δεδομένων περιλαμβάνουν την ηλικία, το φύλο, την περιοχή διαμονής, το μήκος της βιογραφίας, καθώς και των αγαπημένων φράσεων, την κατάσταση της σχέσης τους και τον αριθμό των φίλων, των φωτογραφιών, των ενδιαφερόντων, και τα σχόλια που παρέχονται από το χρήστη. Κάποια από τα προβλήματα που συνάντησαν κατά τη συλλογή των δεδομένων είναι πως κάποια από τα ανοικτά δεδομένα που έχουν οι χρήστες έχουν ελεύθερη μορφή εισόδου, δηλαδή οι χρήστες είναι ελεύθεροι να εισάγουν οποιαδήποτε συμβολοσειρά κειμένου σε αυτόν τον πεδίο, και ως εκ τούτου η ανάλυση τέτοιων πεδίων παρουσιάζει πρόκληση. Παράλληλα, μαζεύοντας πεδία κειμένου, όπως περιγραφές φωτογραφιών και τις δημοσιεύσεις τους, χρησιμοποίησαν έτοιμο λογισμικό γλωσσικής έρευνας και καταμέτρησης λέξεων το οποίο, τους προσέφερε δεδομένα όπως: θετικά και αρνητικά συναισθήματα, περιεχόμενο, εάν υπάρχει αυτό-αναφορά, βασικά θέματα του κειμένου, στυλ γραφής και πολλά άλλα. Σε αυτή τη μελέτη έχει χρησιμοποιήσει ένα σύνολο δεδομένων αποτελούμενο από 527 άτομα. Οι χρήστες της μελέτης δεν έπαιρναν πίσω κάποιο αποτέλεσμα από τα δεδομένα που έδιναν. Η σημασία των αποτελεσμάτων της μελέτης αυτής είναι ότι κατά την εξέταση μιας μεγάλης ομάδας χρηστών, η αυτόματη ανάλυση μπορεί να χρησιμοποιηθεί για τον εντοπισμό εκείνων με συγκεκριμένα γνωρίσματα προσωπικότητας. Επίσης, αυτή η έρευνα δείχνει ότι ακόμη και φαινομενικά ‘αθώες’ περιγραφές εικόνων και πληροφορίες κατάστασης μπορούν να χρησιμοποιηθούν για να διακρίνουν κάποιες πληροφορίες που οι χρήστες μπορεί να προτιμούν να παραμένουν κρυμμένες.

Οι Bachrach et al. [5] χρησιμοποίησαν τα έτοιμα δεδομένα από το myPersonality project από τα οποία επέλεξαν χρήστες οι οποίοι είχαν συμπληρώσει το ερωτηματολόγιο “Big Five” και είχαν δώσει στο πρόγραμμα τα δεδομένα τους σε σχέση με: τον αριθμό των φίλων τους, τον αριθμό των ομάδων (Groups) στις οποίες ανήκαν, στον αριθμό των likes τους σε δημοσιεύσεις, τον αριθμό των φωτογραφιών και καταστάσεων τους και τον αριθμό των περιπτώσεων που σημείωσαν (tag) τον χρήστη οι φίλοι του σε φωτογραφίες. Οι χρήστες της μελέτης ήταν εικονικοί αφού δεν αλληλεπιδρούσαν απευθείας με την ερευνητική ομάδα. Ο στόχος της μελέτης αυτής ήταν να εντοπίσει σε ποιά από τα δεδομένα τα οποία συνέλεξαν υπήρχε θετική ή αρνητική συσχέτιση με τις 5 διαστάσεις της προσωπικότητας που έδινε το ερωτηματολόγιο “Big Five”. Κάποιες σημαντικές συσχετίσεις τις οποίες εντόπισαν ήταν ότι η Εξωστρέφεια συσχετίζεται θετικά με τους αριθμούς των φίλων, των ομάδων (Groups), των like και των καταστάσεων που

δημοσιεύουν οι χρήστες αυτής της κατηγορίας, η Τερπνότητα συσχετίζεται αρνητικά με τον αριθμό των likes, η Νεύρωση συσχετίζεται θετικά με τον αριθμό των likes αλλά αρνητικά με τον αριθμό των φίλων, η Ευσυνειδησία συσχετίζεται θετικά με τον αριθμό των likes και αρνητικά με τον αριθμό των ομάδων και των φωτογραφιών του χρήστη και τέλος η Διαφάνεια συσχετίζεται θετικά με τον αριθμό των likes, των καταστάσεων που δημοσιεύει ο χρήστης και των αριθμό των ομάδων που ανήκει.

Ακόμα μία μελέτη που χρησιμοποίησε αυτή την μέθοδο είναι η [10], κατά την οποία η ανάκτηση των δεδομένων χωρίζεται σε 2 φάσεις. Στη πρώτη φάση, όλοι οι συμμετέχοντες κλήθηκαν να συμπληρώσουν το NEO-PI-R, προκειμένου να εκτιμήσουν την προσωπικότητα κατά μήκος των περιοχών του μοντέλου του “Big Five”. Η δεύτερη φάση της μελέτης ήταν η απόσπαση των δημόσιων δεδομένων των χρηστών στο Facebook. Με βάση τη ρύθμιση του προφίλ των χρηστών γενικά στο Facebook, δημιουργήθηκαν τέσσερις διαστάσεις των χρηστών, δηλαδή: Βασικές πληροφορίες, Προσωπικές πληροφορίες, στοιχεία Επικοινωνίας και Εκπαίδευσης, και πληροφορίες Εργασίας. Οι βασικές πληροφορίες περιλαμβάνουν πληροφορίες που οι χρήστες ανέβασαν στο Facebook σχετικά με το φύλο, την ημερομηνία γέννησης, τη πατρίδα, τα μέλη της οικογένειας, την οικογενειακή κατάσταση (μόνος, αρραβωνιασμένος, παντρεμένος, κ.λπ.), το είδος της σχέσης που ο χρήστης αναζητά (φιλία, ραντεβού, τυχερά παιχνίδια, κ.λπ.), τις πολιτικές και θρησκευτικές απόψεις του. Οι προσωπικές πληροφορίες περιλαμβάνουν δεδομένα όπως τις δραστηριότητες, τα ενδιαφέροντα, την αγαπημένη μουσική, τις αγαπημένες τηλεοπτικές εκπομπές, τις αγαπημένες ταινίες, αγαπημένα βιβλία, τα αγαπημένα αποσπάσματα, και το κομμάτι «Για Μένα». Για τα στοιχεία επικοινωνίας πληροφορίες που περιλαμβάνονται όνομα, ηλεκτρονική διεύθυνση, κινητό τηλέφωνο, σταθερό τηλέφωνο, τη διεύθυνση, πόλη / χωριό, τη γειτονιά, τον ταχυδρομικό κώδικα και την ιστοσελίδα. Για την εκπαίδευση και την εργασία κάποιες πληροφορίες που περιλαμβάνονται είναι πληροφορίες σχετικά με το πανεπιστήμιο, την εστίαση, το βαθμό, το γυμνάσιο, εργοδότες, περιγραφή θέσεων εργασίας, πόλη / χωριό και τη χρονική περίοδο. Η έρευνα έγινε σε 237 φοιτητές ενός Ισραηλινού πανεπιστημίου και η ανάκτηση των δεδομένων χρειαζόταν 30 με 45 λεπτά για κάθε άτομο. Σε αυτή τη μελέτη όλοι οι συμμετέχοντες έλαβαν πιστωτικές μονάδες σε εισαγωγικά μαθήματα για τη συμμετοχή τους στη μελέτη αλλά δεν έπαιρναν ανατροφοδότηση για τα δεδομένα που έδιναν. Αυτή η μελέτη έδειξε ότι η προσωπικότητα συσχετίζεται πολύ με τη χρήση των κοινωνικών δικτύων.

2.3.2 Αυτοματοποιημένες Τεχνικές Συλλογής Δεδομένων

Κατά τις αυτοματοποιημένες μεθόδους η βασική ιδέα είναι να δημιουργηθεί μια εφαρμογή στο Facebook η οποία θα ζητά από το χρήστη έγκριση για πρόσβαση στα δεδομένα του. Σε αυτή την περίπτωση όμως για να ενθαρρυνθεί ο χρήστης να παρέχει τα δεδομένα του είναι αναγκαίο η εφαρμογή να του παρουσιάζει κάποια αποτελέσματα για τον ίδιο ή που να είναι σχετικά με τα δεδομένα του. Τα πλεονεκτήματα αυτής της προσέγγισης είναι πως εκτός από το ότι η εξόρυξη των δεδομένων γίνεται αυτοματοποιημένα, με την άδεια του χρήστη μπορούν οι προγραμματιστές των εφαρμογών του Facebook να πάρουν περισσότερα και με μεγαλύτερη πιθανότητα ερευνητικής χρησιμότητας δεδομένα εκτός από τα δημόσια δεδομένα του προφίλ του χρήστη. Ακόμη αφού μια εφαρμογή στο Facebook είναι προσιτή σε όλους τους χρήστες του Facebook τότε χρησιμοποιώντας την λειτουργία της κοινοποίησης, οι αρχικοί χρήστες του συστήματος μπορεί παθητικά να πείσουν τον κύκλο τους να χρησιμοποιήσει την εφαρμογή και να συνεχιστεί αυτό αναδρομικά.

Αυτό το είδος προσέγγισης χρησιμοποιούν οι [7], οι οποίοι δημιούργησαν μια εφαρμογή η οποία στην αρχή υποχρέωνε τους χρήστες της να απαντούν το ερωτηματολόγιο ZKPQ-60-cc για να προβλέψουν την προσωπικότητα του χρήστη. Στη συνέχεια ζητούσαν τα δικαιώματα του χρήστη για να αποκτήσουν πρόσβαση στα δεδομένα του «τοίχου» (wall) του, να μπορούν να επεξεργαστούν τα δεδομένα του ακόμη και αν αυτός δεν είναι ενωμένος στο διαδίκτυο και επίσης την δυνατότητα να κοινοποιούν τα αποτελέσματα του τεστ προσωπικότητας στο Facebook του χρήστη. Ακολούθως, παρουσίαζαν στο χρήστη τα αποτελέσματα του τεστ που συμπλήρωσε υποχρεωτικά στην αρχή, παρόμοιους χρήστες που έχουν ήδη χρησιμοποιήσει την εφαρμογή, να συγκρίνουν τα αποτελέσματα τους με τα αποτελέσματα των φίλων τους, αν έχουν συμπληρώσει και αυτοί το τεστ, και μπορούσαν επίσης να προσκαλέσουν τους φίλους τους να χρησιμοποιήσουν την εφαρμογή. Με βάση τα δεδομένα που συγκέντρωσαν, χρησιμοποιήθηκαν διάφορες τεχνικές μηχανικής μάθησης, με σκοπό την αναζήτηση για μοτίβα αλληλεπίδρασης με τα αντικείμενα και τις δραστηριότητες του Facebook, που παρέχουν πληροφορίες σχετικά με τους χρήστες. Αυτοί οι ταξινομητές είναι σε θέση να προβλέπουν την προσωπικότητα του χρήστη ξεκινώντας από τις παραμέτρους που σχετίζονται με την αλληλεπίδραση των χρηστών, όπως ο αριθμός των φίλων τους ή ο αριθμός των δημοσιεύσεων στον «τοίχο» τους. Τα αποτελέσματα δείχνουν ότι οι ταξινομητές έχουν ένα υψηλό επίπεδο ακρίβειας, καθιστώντας την

προτεινόμενη προσέγγιση μια αξιόπιστη μέθοδο για την πρόβλεψη της προσωπικότητας του χρήστη.

Μία ακόμα μελέτη που χρησιμοποίησε αυτή τη τεχνική είναι η [6], κατά την οποία δημιουργήθηκε μια εφαρμογή στο Facebook με δύο λειτουργίες. Πρώτον, στους χρήστες χορηγήθηκε ένα ερωτηματολόγιο 45 ερωτήσεων - μια έκδοση του τεστ προσωπικότητας “Big Five”. Την ίδια στιγμή, συγκεντρώναν επίσης όλες τις πληροφορίες σχετικά με το προφίλ των χρηστών που ήταν στη διάθεση των εφαρμογών του Facebook, δηλαδή τις βασικές πληροφορίες του προφίλ των χρηστών. Η συγκεκριμένη ερευνητική ομάδα συγκεντρώσε στο σύνολο 161 στατιστικές. Η εφαρμογή που δημιουργήθηκε δεν έδινε στους χρήστες κάποια ανατροφοδότηση. Η πληροφορία που έχουν συλλέξει δείχνουν ότι τα χαρακτηριστικά της προσωπικότητας του “Big Five” μπορούν να προβλεφθούν από δεδομένα στο Facebook.

2.3.3 Η προσέγγιση που ακολουθεί η εργασία

Η προσέγγιση της παρούσας διπλωματικής εργασίας μοιάζει περισσότερο με αυτή των [7], δηλαδή, χρησιμοποιούμε μια αυτοματοποιημένη τεχνική, δημιουργώντας μια εφαρμογή στο Facebook για την εξόρυξη δεδομένων των χρηστών της μελέτης μας. Οι ομοιότητες των εφαρμογών μας είναι στο ότι και στις δύο περιπτώσεις παρουσιάζονται τα αποτελέσματα των τεστ προσωπικότητας στους χρήστες, δίνετε στους χρήστες η δυνατότητα κοινοποίησης των αποτελεσμάτων τους στο Facebook και επίσης τους παρουσιάζονται χρήστες που έχουν ήδη χρησιμοποιήσει την εφαρμογή και έχουν επιτύχει παρόμοια αποτελέσματα με αυτούς. Οι διαφορές που έχουν οι εφαρμογές μας είναι ότι η εφαρμογή που δημιουργήσαμε στα πλαίσια της παρούσας διπλωματικής εργασίας προχωρά την έρευνα το τομέα αυτό ένα βήμα παραπάνω και προσπαθεί να δώσει περισσότερη ικανοποίηση στο χρήστη. Αυτό γιατί η εφαρμογή ζητά από το χρήστη πιο εξειδικευμένα και περισσότερα στοιχεία του σε σχέση με την δραστηριότητα του στο Facebook, που μετά από την επεξεργασία που εκτελείται, παρουσιάζουμε στο χρήστη μια υποθετική προσέγγιση των αποτελεσμάτων των διαστάσεων της προσωπικότητας του “Big Five”. Τα αποτελέσματα αυτά μπορούν να συγκριθούν με τα αποτελέσματα του τεστ προσωπικότητας, μόνο αν το έχει συμπληρώσει ο χρήστης. Στη δική μας περίπτωση δεν υποχρεώνουμε το χρήστη να συμπληρώσει το ερωτηματολόγιο προσωπικότητας “Big Five”. Ακόμη παρουσιάζονται στον χρήστη συγκεντρωτικές γραφικές παραστάσεις, σε σχέση με τα

likes του σε δημοσιεύσεις και σε σχέση με το τύπο των διαφόρων σελίδων που δημιουργούνται τώρα στο Facebook. Η εφαρμογή μας επιπρόσθετα, παρουσιάζει στο χρήστη συστάσεις σε σελίδες του Facebook ανάλογα με τον τύπο σελίδας που προτιμά και δεν έχει ακόμη κάνει like. Τέλος, δίνουμε την δυνατότητα στους χρήστες μας να μεταβούν στην σελίδα του PersonaWeb και να εκτελέσουν διάφορα τεστ τα οποία θα τους πληροφορήσουν σχετικά με χαρακτηριστικά του εαυτού τους όπως γνωστικά χαρακτηριστικά και άλλα.

2.4 PersonaWeb¹

Το έργο PersonaWeb εστιάζει στην υλοποίηση και αξιολόγηση ενός δυναμικού Web-based πλαισίου για την επίτευξη δυναμικής προσαρμογής και εξατομίκευσης ενός γενικού περιεχομένου eServices και βασίζεται σε ανθρώπινους παράγοντες. Πιο συγκεκριμένα, δεδομένου ότι οι χρήστες έχουν διαφοροποιημένες αντιληπτικές προτιμήσεις, το ίδιο περιεχόμενο ή υπηρεσία που παρέχεται από έναν οργανισμό, θα πρέπει να ανακατασκευαστεί και να παρουσιαστεί στους χρήστες με βάση των γνωστικών και άλλων χαρακτηριστικών. Με αυτό τον τρόπο, θα αυξήσει την αφομοίωση πληροφοριών, η ακρίβεια στις δραστηριότητες αναζήτησης γνωστικών στόχων και, κατά συνέπεια, να ενισχύσει την ικανοποίηση και την αποδοχή των υπηρεσιών στο διαδίκτυο.

2.5 Επίλογος

Πιο πάνω είδαμε πως η βιβλιογραφία αναφέρεται σε θέματα προσωπικότητας σε σχέση με την χρήση εργαλείων κοινωνικών δικτύων και συγκεκριμένα το Facebook. Έχουμε αναφερθεί στην θεωρία του Big Five σαν μέθοδο κατηγοριοποίησης της προσωπικότητας ενός ατόμου, και πως οι ενέργειες του μέσα από το Facebook θα μπορούσαν να μας δώσουν περισσότερες πληροφορίες σχετικά με την προσωπικότητα του. Στο επόμενο κεφάλαιο θα αναπτυχθεί το

¹ PersonaWeb Project: <https://personaweb.cs.ucy.ac.cy/>

πλαίσιο που ακολουθήσαμε σε αυτή την διπλωματική εργασία για να συλλέξουμε δεδομένα, να κτίσουμε ένα μοντέλο χρήστη και να εφαρμόσουμε αυτό το μοντέλο ώστε να δώσουμε κάποιες χρήσιμες και εξατομικευμένες πληροφορίες στο χρήστη.

Κεφάλαιο 3

Υπολογιστικό Πλαίσιο (Αρχιτεκτονική)

3.1 Εισαγωγή.....	18
3.2 Το Κοινωνικό Δίκτυο Facebook.....	19
3.3 Μοντελοποίηση Χρήστη	19
3.4 Υπολογιστικό Πλαίσιο Μοντελοποίησης Χρήστη	20
3.4.1 Περιγραφή Μηχανισμού Εξόρυξης Πληροφοριών του Χρήστη από το Facebook	21
3.4.2 Μοντέλο Χρήστη	24
3.4.3 Εφαρμογή Μοντέλου Προσωπικότητας για την Παροχή Πρόσθετων Πληροφοριών στο Χρήστη	26
3.5 Τεχνολογίες	27
3.6 Εργαλεία ανάπτυξης λογισμικού.....	29
3.7 Επίλογος.....	30

3.1 Εισαγωγή

Στο προηγούμενο κεφάλαιο είδαμε μια ανασκόπηση της ερευνητικής περιοχής που συνάδει με τα ερευνητικά ερωτήματα της παρούσας διπλωματικής. Σε αυτό το κεφάλαιο θα παρουσιαστεί μια γενική περιγραφή του κοινωνικού δικτύου Facebook, η μοντελοποίηση του χρήστη, μια γενική επεξήγηση του υπολογιστικού πλαισίου του οποίου δημιουργήθηκε για τη μοντελοποίηση του χρήστη και οι τεχνολογίες οι οποίες χρησιμοποιήθηκαν για την επίτευξη αυτού του συστήματος.

3.2 Το Κοινωνικό Δίκτυο Facebook

Το Facebook είναι αυτή τη στιγμή η δημοφιλέστερη ιστοσελίδα κοινωνικής δικτύωσης παγκόσμια με 1,310,000,000² μηνιαίους χρήστες, με το 48% από αυτούς να ενώνονται καθημερινά στο προφίλ τους. Κατά συνέπεια, ο όγκος των δεδομένων που δημιουργούνται καθημερινά μέσω της αλληλεπίδρασης, τόσο των χρηστών μεταξύ τους, όσο και με το ίδιο το σύστημα είναι τεράστιος. Όπως έχουμε δει στο Κεφάλαιο 2 πολλή ερευνητές βλέπουν τον όγκο αυτών των δεδομένων σαν πηγή ερευνητικών ερωτημάτων που μπορούν να απαντηθούν με τη σωστή επεξεργασία και προσθέτοντας νόημα σε αυτά.

Μέσω του Facebook ο χρήστης μπορεί να δημιουργήσει ενώσεις με άλλους χρήστες (friends) να δηλώσει αν κάτι που οι φίλοι του ή ο ίδιος ανεβάζουν στο προσωπικό τους χώρο (wall) του αρέσει (Like) ή που βρίσκεται μια συγκεκριμένη στιγμή με άλλους φίλους καθώς και τι διάθεση έχει. Επίσης, ο χρήστης μπορεί να ανεβάσει φωτογραφίες και βίντεο του εαυτού του ή κάποιου άλλου και να προσθέσει πληροφορίες (tags) σχετικά με τα άτομα που βρίσκονται σε αυτά.

Τα δεδομένα αυτά που δημιουργούνται φυλάγονται από το Facebook σε πίνακες στην βάση δεδομένων τους, από την οποία με συγκεκριμένα ερωτήματα (queries) μπορούμε να εξάγουμε λογικά δεδομένα για ένα ή περισσότερους χρήστες. Περισσότερες πληροφορίες σχετικά με την επεξεργασία δεδομένων θα δούμε στο υποκεφάλαιο 4.5.2. Με αυτό τον τρόπο πολλοί ερευνητές χρησιμοποιούν την ελεύθερη πρόσβαση σε αυτά τα δεδομένα για σκοπούς έρευνας.

3.3 Μοντελοποίηση Χρήστη

Η ερευνητική περιοχή της μοντελοποίησης χρήστη ασχολείται με την δημιουργία ενός μοντέλου, αναπαριστάμενο από ένα αριθμό παραμέτρων οι οποίες αντιπροσωπεύουν το χρήστη κάποιου συγκεκριμένου συστήματος, ιστοσελίδας παιχνιδιού κτλ. Οι παράμετροι που θα

²Τα στατιστικά στοιχεία σχετικά με το Facebook είναι αναρτημένα στον εξής σύνδεσμο <http://www.statisticbrain.com/facebook-statistics/>

χρησιμοποιηθούν εξαρτώνται πάντα από την εφαρμογή και το σύστημα στο οποίο θα χρησιμοποιηθεί το μοντέλο αυτό. Ο σκοπός του σχηματισμού ενός μοντέλου για το χρήστη κάποιου συστήματος είναι η μετέπειτα εφαρμογή του ώστε να παρέχει το σύστημα στο χρήστη μια πιο εξατομικευμένη ή/και προσαρμοσμένη εμπειρία αλληλεπίδρασης με αυτό [8]. Κατά συνέπεια, ο χρήστης θα μπορεί να χειρίζεται καλύτερα ή να παίρνει πιο χρήσιμες πληροφορίες από το σύστημα και θα είναι πιο ευχαριστημένος με την χρήση του.

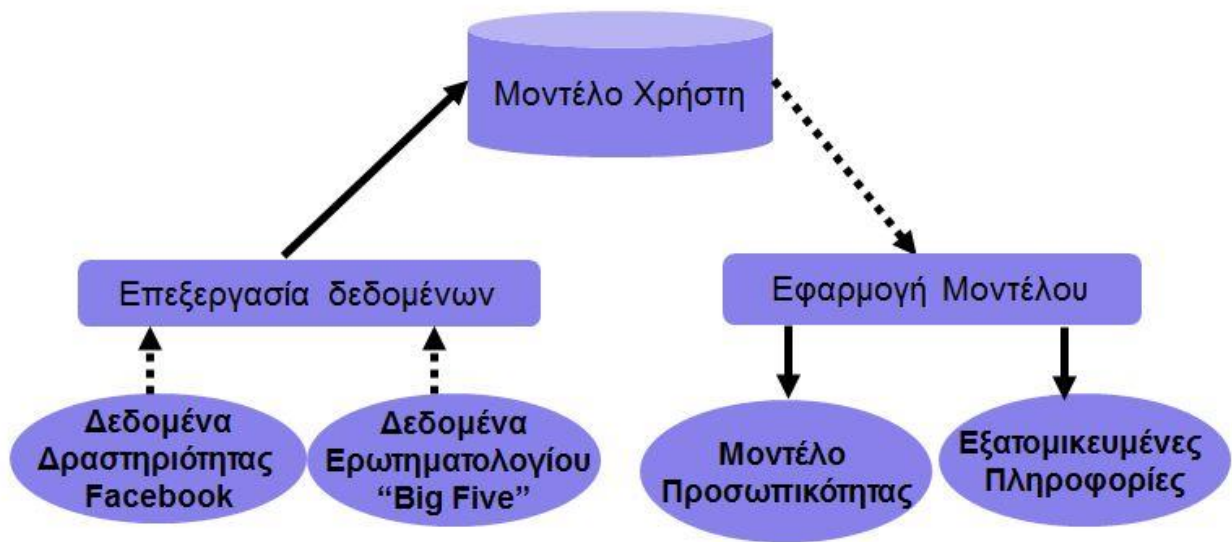
Το μοντέλο χρήστη είναι το πιο σημαντικό κομμάτι μεγάλου αριθμού ιστοσελίδων και άλλων διαδικτυακών εφαρμογών. Χρησιμοποιώντας τις πληροφορίες που αποθηκεύουν σχετικά με το χρήστη, προσπαθούν να προσαρμόσουν το υλικό τους ανάλογα και σε κάποιες ακόμα περιπτώσεις να προσφέρουν εισηγήσεις του τύπου ‘Τι θα άρεσε στο χρήστη αυτό περισσότερο’[9]. Τέτοιο παράδειγμα αποτελεί η Amazon³, μια εταιρεία με χρόνια δραστηριότητα στις πωλήσεις μέσω διαδικτύου.

Στο επόμενο υποκεφάλαιο θα περιγραφεί συνοπτικά η διαδικασία μοντελοποιήσεως των χρηστών όπως έχει αναπτυχθεί στην παρούσα διπλωματική εργασία.

3.4 Υπολογιστικό Πλαίσιο Μοντελοποίησης Χρήστη

Η κύρια υπόθεση της παρούσας διπλωματικής εργασίας είναι: χρησιμοποιώντας επεξεργασμένα δεδομένα χρηστών από το Facebook, κατά πόσο μπορεί να υπολογιστούν χαρακτηριστικά της προσωπικότητας του κάθε χρήστη. Βασισμένοι στην πιο πάνω υπόθεση και προσθέτοντας στις μελέτες που συζητήθηκαν στο υποκεφάλαιο 2.3, παρουσιάζεται, σε αυτό το κεφάλαιο, ένα υπολογιστικό πλαίσιο (Σχήμα 3-1) για την αυτοματοποιημένη δημιουργία ενός μοντέλου χρήστη.

³ www.amazon.com



Σχήμα 3-1 Υπολογιστικό Πλαίσιο

Το πλαίσιο αυτό χωρίζεται σε δύο φάσεις: Την εξόρυξη και επεξεργασία δεδομένων ενός χρήστη του Facebook με σκοπό την δημιουργία του μοντέλου χρήστη

Την εφαρμογή του μοντέλου χρήστη για την σχηματισμό της διάστασης του “Big Five” στην οποία ο χρήστης ανήκει και για την μορφοποίηση εξατομικευμένων συστάσεων σε σχέση με άλλους χρήστες και σελίδες του Facebook.

Σκοπός των ακόλουθων ενοτήτων είναι να παρέχουν μια γενική εικόνα του μηχανισμού για τον υπολογισμό του μοντέλου προσωπικότητας του χρήστη (υποκεφάλαιο 3.4.1), τα στοιχεία του προτεινόμενου μοντέλου (υποκεφάλαιο 3.4.2) και πως αυτό το μοντέλο θα εφαρμοστεί για την απόκτηση γνώσης για την προσωπικότητα του χρήστη και την παροχή εξατομικευμένων πληροφοριών στο χρήστη αυτό (υποκεφάλαιο 3.4.3).

3.4.1 Περιγραφή Μηχανισμού Εξόρυξης Πληροφοριών του Χρήστη από το Facebook

Η εξαγωγή του μοντέλου προσωπικότητας του χρήστη περιλαμβάνει (i) την ανάκτηση των δεδομένων του χρήστη από τους πίνακες της βάσεις δεδομένων που διαθέτει το Facebook για τους προγραμματιστές του, (ii) το ξεκαθάρισμα των ουσιαστικών στοιχείων και την αποθήκευσή τους στη δική μας βάση δεδομένων και (iii) την επεξεργασία των δεδομένων αυτών για την

δημιουργία του μοντέλου προσωπικότητας του χρήστη και παροχή εξατομικευμένων πληροφοριών. Το πλαίσιο μας χρησιμοποιεί για εισαγωγή στο σύστημα δημιουργίας του μοντέλου χρήστη, τα δεδομένα δραστηριότητας τα οποία μπορούμε να ανακτήσουμε με την άδεια του χρήστη. Τα δεδομένα δραστηριότητας τα οποία εξάγονται για το χρήστη, (συμπεριλαμβανόμενου γενικές δημόσια διαθέσιμες πληροφορίες για το χρήστη, αντικείμενα τα οποία κοινοποίησε ο ίδιος αλλά και δεδομένα τα οποία φανερώουν διαδικτυακή δραστηριότητα, μέσω Facebook, μεταξύ του χρήστη για τον οποίο δημιουργούμε το μοντέλο και άλλων χρηστών), είναι απλά ένα κομμάτι των δεδομένων των οποίων διαθέτει το Facebook στους προγραμματιστές του. Έτσι, με τις ανάλογες προσαρμογές στο μοντέλο απόκτησης των δεδομένων, η εφαρμογή της παρούσας διπλωματικής εργασίας θα μπορούσε να ενημερωθεί σε ένα πιο ολοκληρωμένο μοντέλο προσωπικότητας χρήστη, εάν νέα δεδομένα εμφανίσουν συσχέτιση με ενδείξεις προσωπικότητας.

Τα δεδομένα τα οποία αποφασίσαμε να ανακτήσουμε για το χρήστη με σκοπό να δημιουργήσουμε το μοντέλο προσωπικότητας του, βρίσκονταν διασκορπισμένα σε μια γκάμα πινάκων της βάσης δεδομένων του Facebook. Οι πίνακες αυτοί [10] δεν είναι διαχωρισμένοι με ένα μοναδικό τρόπο ευρετηρίασης δηλαδή, ένα υποσύνολο των πινάκων του Facebook έχει ως ευρετήριο πεδίο τον μοναδικό κωδικό κάποιου χρήστη, το οποίο ήταν και το επιθυμητό μας ευρετήριο. Η άλλη περίπτωση είναι πως το ευρετήριο σε κάποιους από αυτούς τους πίνακες είναι ο κωδικός συγκεκριμένης δημοσίευσης (και όχι του ιδιοκτήτη της που θα ήταν κάποιος χρήστης ή σελίδα) είτε επιπρόσθετα κάποιο άλλο πεδίο το οποίο χρησιμοποιείτε πολύ περιορισμένα και συνήθως μόνο από το ίδιο το Facebook. Αυτό μπορεί να αιτιολογηθεί με το ότι το Facebook κρατά δεδομένα από την στιγμή που πρωτοεμφανίστηκε το 2006 όταν πρώτο έγινε διαθέσιμο στο ευρύ κοινό κι έτσι με κάθε καινούργια λειτουργία που δημιουργεί και διαθέτει για τους χρήστες του βρίσκει ένα τέτοιο τρόπο ώστε να μην χρειαστεί αναδιοργάνωση της βάσης κάθε φορά που γίνεται μια αλλαγή ή πρόσθεση λειτουργίας.

Οι πίνακες με τους οποίους εργαστήκαμε στη παρούσα διπλωματική εργασία φαίνονται στον Πίνακας 3-1

Όνομα πίνακα Facebook σε FQL	Περιγραφή και Σκοπός
album	Συλλογή φωτογραφιών και βίντεο όπως αντιπροσωπεύεται

checkin	Επιστρέφει πληροφορίες σε σχέση με τα checkin που κάνει ο χρήστης σε ένα γεωγραφικό σημείο
event	Επιστρέφει πληροφορίες σε σχέση με μια διοργάνωση
event_member	Περιέχει μια λίστα με άτομα που παρευρέθηκαν σε μια διοργάνωση
friend	Επιστρέφει μια λίστα με τους φίλους ενός χρήστη
like	Επιστρέφει τις ταυτότητες των χρηστών που έκαναν like σε ένα αντικείμενο
link	Τα links που ανέβασε ο χρήστης
page	Επιστρέφει πληροφορίες σε σχέση με μια σελίδα
page_fan	Πληροφορίες σχετικά με το αν σε κάποιο χρήστη αρέσει μια σελίδα
photo	Επιστρέφει πληροφορίες για μια φωτογραφία
photo_tag	Tag σε μια φωτογραφία
profile_pic	Φωτογραφίες προφίλ κάποιου χρήστη
status	Τα μηνύματα 'status' ενός χρήστη
stream	Επιστρέφει μια σειρά από δημοσιεύσεις
user	Πληροφορίες σχετικά με το προφίλ ενός χρήστη
video	Πληροφορίες σχετικά με βίντεο
video_tag	Tag σε ένα βίντεο

Πίνακας 3-1 - Πίνακες από το Facebook που χρησιμοποιήθηκαν για τις ανάγκες της εργασίας

Από αυτούς τους πίνακες επιλέγηκαν πεδία τα οποία είχαν σχέση με τον τύπο των αντικειμένων, των ιδιοτήτων τους και συγκεντρωτικές μετρικές για το πλήθος των αντικειμένων που αντιστοιχούσαν στο συγκεκριμένο χρήστη. Η επεξεργασία που χρειαζόνταν μερικά από αυτά τα δεδομένα δεν ήταν πάντα στο ίδιο βήμα της απόκτησης τους. Αξίζει να αναφέρουμε ότι η επεξεργασία κάποιων δεδομένων βρισκόταν στο τρόπο με τον οποίο θα τα ζητούσαμε από τη βάση δεδομένων του Facebook ενώ από την άλλη πλευρά, κάποια από αυτά τα δεδομένα χρειαζόνταν επεξεργασία στην εξόρυξη ουσιώδη δεδομένων από αυτά πριν να μπορούμε να τα αποθηκεύσουμε στη δική μας βάση δεδομένων και να τα χρησιμοποιήσουμε για την δημιουργία του μοντέλου προσωπικότητας του χρήστη.

Η διαδικασία ανάκτησης των δεδομένων περιλαμβάνει τρία απλά βήματα. 1) ο χρήστης παραχωρεί τα δικαιώματα σε σχέση με τα δεδομένα τα οποία θα ανακτηθούν από το προφίλ του ή από τη βάση δεδομένων του Facebook που σχετίζονταν με αυτόν. Αξιοσημείωτο είναι το ότι ενώ τα δεδομένα της δραστηριότητας του χρήστη στο κοινωνικό δίκτυο τα ανακτούσαμε με ξεχωριστά καλέσματα προς τη βάση δεδομένων του Facebook, παρόλα αυτά, τα δικαιώματα που χρειαζόμασταν από το χρήστη τα ζητούσαμε από την αρχή ομαδοποιημένα και όχι κατά την ώρα του καλέσματος του συγκεκριμένου ερωτήματος προς τη βάση δεδομένων. 2) Ακολουθώντας, για κάθε ομάδα αντικειμένων που θα ανακτούσαμε για επεξεργασία και εξόρυξη ουσιαστική δεδομένων, χρησιμοποιούσαμε ένα ειδικευμένο ερώτημα βάσης δεδομένων με τη χρήση της εξειδικευμένης διεπαφής προγραμματισμού εφαρμογών που προσφέρει προς τους επεξεργαστές του το Facebook. Εδώ θα ήθελα να αναφέρω ότι για την επιτυχή εκτέλεση της ερώτησης χρειαζόταν το διακριτικό πρόσβασης του χρήστη, που δημιουργούταν με την παροχή της άδειας του χρήστη και με τον μυστικό κωδικό της συγκεκριμένης εφαρμογής, αυτής που δημιουργήσαμε σε αυτή τη περίπτωση, για να δεχτεί το Facebook να στείλει πίσω αποτέλεσμα. 3) Στη συνέχεια αφού παίρναμε ως απάντηση κάθε ερωτήματος από το Facebook τα δεδομένα αυτά, εκτελούνταν η επεξεργασία των δεδομένων και τέλος αποθηκεύονταν οι πληροφορίες που εξορύξαμε, στη δική μας βάση δεδομένων.

3.4.2 Μοντέλο Χρήστη

Το μοντέλο προσωπικότητας αντιπροσωπεύει τα ποσοστά τα οποία σκοράρει ένας χρήστης στις 5 διαστάσεις της προσωπικότητας του “Big Five”. Ωστόσο τα δεδομένα τα οποία χρησιμοποιούμε για να δημιουργήσουμε αυτό το μοντέλο είναι τα επεξεργασμένα δεδομένα τα οποία εξορύξαμε από το προφίλ του χρήστη και αυτά είναι που χρησιμοποιούνται για να δώσουν νόημα στα πέντε εκείνα ποσοστά των διαστάσεων της προσωπικότητας. Εν τέλει, αποφασίσαμε να χρησιμοποιήσουμε τα εξής στοιχεία από το προφίλ του χρήστη: τους φίλους του, τους ενεργούς του φίλους (δηλαδή αυτούς που έχουν φανερή δραστηριότητα με το χρήστη μας στο κοινωνικό δίκτυο Facebook), τα likes του σε δημοσιεύσεις, των τύπο των δημοσιεύσεων που του αρέσουν, τον δημιουργό και ιδιοκτήτη των δημοσιεύσεων αυτών, τις σελίδες που του αρέσουν, τους τύπους των σελίδων που του αρέσουν, τις δημοσιεύσεις τις οποίες ο χρήστης ήταν tagged, τις φωτογραφίες του, τα βίντεο του, τα check-ins του, τα check-ins που έκαναν οι φίλοι του μαζί

με το χρήστη μας, τα events που παραβρέθηκε, τα events που δημιούργησε, τις δημοσιεύσεις του (τα status του όπως αποκαλούνται από την κοινότητα του Facebook), τους συνδέσμους τους οποίους δημοσίευσε, την πληροφορία αν ο χρήστης έκανε tag τον εαυτό του σε κάποια δημοσίευση είτε δική του είτε όχι, το ποσοστό των φωτογραφιών προφίλ του σε σχέση με τον αριθμό όλων των φωτογραφιών και το ποσοστό του αγαπημένου του τύπου σελίδας σε σχέση με τον αριθμό όλων των σελίδων που του αρέσουν. Τα ονόματα που χαρακτηρίζουν τα πιο πάνω φαίνονται στον Πίνακας 3-2.

Δεδομένα	Περιγραφή
user_friends	φίλοι του χρήστη
active_friends	οι ενεργοί του φίλοι (δηλαδή αυτοί που έχουν φανερή δραστηριότητα με το χρήστη μας στο κοινωνικό δίκτυο Facebook)
posts_likes	τα likes του σε δημοσιεύσεις
posts_type_likes	οι τύποι των δημοσιεύσεων που του αρέσουν
post_creator	δημιουργός δημοσίευσης
post_owner	ιδιοκτήτης των δημοσιεύσεων
pages_liked	τις σελίδες που του αρέσουν
pages_liked_type	οι τύποι των σελίδων που του αρέσουν
posts_user_tagged	δημοσιεύσεις στις οποίες ο χρήστης ήταν tagged
user_photos	φωτογραφίες του χρήστη
user_video	τα βίντεο του χρήστη
user_check-ins	τα check-ins του χρήστη
user_friends_checkins	τα check-ins που έκαναν φίλοι του μαζί με το χρήστη
user_attended_events	τα events που παραβρέθηκε ο χρήστης
user_created_events	τα events που δημιούργησε ο χρήστης
user_status	τις δημοσιεύσεις του - status
user_links	τους συνδέσμους τους οποίους δημοσίευσε
user_self_tag	την πληροφορία αν ο χρήστης έκανε tag τον εαυτό του σε κάποια δημοσίευση είτε δική του είτε όχι
profile_pic_per	το ποσοστό των φωτογραφιών προφίλ του χρήστη σε σχέση με τον αριθμό όλων των φωτογραφιών

favourite_page_per	το ποσοστό του αγαπημένου τύπου σελίδας σε σχέση με τον αριθμό όλων των σελίδων που αρέσουν στο χρήστη
--------------------	--

Πίνακας 3-2 Δεδομένα που αποτελούν το μοντέλο χρήστη

3.4.3 Εφαρμογή Μοντέλου Προσωπικότητας για την Παροχή Πρόσθετων Πληροφοριών στο Χρήστη

Η εφαρμογή του μοντέλου για κάποιο χρήστη μας δίνει μια εικόνα της δραστηριότητας του μέσα στο κοινωνικό δίκτυο Facebook και επίσης μας παρουσιάζει την ευκαιρία να τον ενημερώσουμε για την προσωπικότητα του και να του παρουσιάσουμε συστάσεις σε σχέση με άλλους υπάρχον χρήστες του συστήματος μας. Τα αποτελέσματα που παρουσιάζονται χωρίζονται σε 2 κατηγορίες, (i) αποτελέσματα τα οποία προκύπτουν απλά από τα δεδομένα του στο Facebook και (ii) αποτελέσματα τα οποία δημιουργούμε βάσει του μοντέλου χρήστη που δημιουργούμε πιο πάνω.

Τα αποτελέσματα δραστηριότητας του χρήστη, προκύπτουν συμμαζεύοντας τα δεδομένα της δραστηριότητας του στο κοινωνικό δίκτυο Facebook. Η παρουσίαση γίνεται ανάλογα με το είδος του αποτελέσματος σε μια μορφή προσιτή προς τον χρήστη, όπως για παράδειγμα σε μορφή εικόνας ή γραφικής παράστασης. Συνοπτικά, παρουσιάζεται στο χρήστη μια γραφική παράσταση με τον τύπο των αντικειμένων του Facebook στα οποία έχει κάνει like. Ο τύπος του αντικειμένου που του άρεσε μπορεί να είναι ένας από τους εξής: φωτογραφία, άλμπουμ, εκδήλωση, ομάδα, σημείωση, σύνδεση, βίντεο, εφαρμογή, κατάσταση, check-in, κριτική, σχόλια. Επιπλέον, παρουσιάζεται ακόμα μια γραφική παράσταση που παρουσιάζει συγκεντρωτικά τους τύπους των σελίδων τις οποίες ο χρήστης έκανε like στο Facebook και ακολουθεί τις δημοσιεύσεις που κοινοποιούν. Παράλληλα, παρουσιάζονται στο χρήστη πληροφορίες σχετικά με: τον ιδιοκτήτη των αντικειμένων στα οποία ο χρήστης έχει κάνει τα περισσότερα likes και του ιδιοκτήτη των αντικειμένων στα οποία ο χρήστης είναι περισσότερες φορές tagged. Ακόμη ένα στοιχείο που μπορεί να δει ο χρήστης από τα στοιχεία δραστηριότητας του είναι τον αριθμό των εκδηλώσεων που δήλωσε στο Facebook ότι θα παραστεί και επίσης τις εικόνες που καθόρισαν οι δημιουργοί των τελευταίων 3 εκδηλώσεων για αυτές, στις οποίες έχει παραστεί.

Σχετικά με τα αποτελέσματα τα οποία δημιουργούμε βάσει του μοντέλου χρήστη, ο χρήστης μπορεί να δει σε μια γραφική παράσταση τα αποτελέσματα που σκόραρε βάση του αλγόριθμου που υλοποιήθηκε για την παρούσα διπλωματική εργασία, στις πέντε διαστάσεις της προσωπικότητας του “Big Five”. Θα ήταν παράληψη να αναφέρω ότι αν ο χρήστης επιλέξει να απαντήσει το ερωτηματολόγιο της προσωπικότητας που του προσφέρει η εφαρμογή μας για μια πιο ακριβή περιγραφή της προσωπικότητας του, τότε του δίνετε το δικαίωμα να συγκρίνει τα αποτελέσματα των δύο αυτών διαφορετικών προσεγγίσεων. Στη συνέχεια ο χρήστης μπορεί να δει μια λίστα, με μέγιστο αριθμό 5 άτομα, τα οποία έχουν παρόμοια συμπεριφορά και δεδομένα δραστηριότητας στο κοινωνικό δίκτυο Facebook με αυτόν. Κλείνοντας, δίνουμε την ευκαιρία στο χρήστη να δει κάποιες σελίδες στο Facebook οι οποίες είναι του αγαπημένου του τύπου σελίδας και δεν τις έχει κάνει like ακόμη. Πιο λεπτομερής περιγραφή θα δοθεί στο Κεφάλαιο 4

3.5 Τεχνολογίες

FQL: Η Facebook Query Language [10], επιτρέπει στους προγραμματιστές του Facebook να χρησιμοποιήσουν μια διεπαφή παρόμοιου τύπου με την SQL για την αναζήτηση των στοιχείων που εκτίθενται από την Διεπαφή Προγραμματισμού της Εφαρμογής Γραφήματος του Facebook (Facebook’s Graph API). Παρέχει ορισμένες προηγμένες λειτουργίες οι οποίες δεν είναι διαθέσιμες στο Graph API, όπως η χρήση των αποτελεσμάτων ενός ερωτήματος της βάσης δεδομένων σε μια άλλη. Τα ερωτήματα είναι της μορφής `SELECT [πεδία] FROM [πίνακας] WHERE [όρους]`. Σε αντίθεση με την SQL, η FQL στο σημείο FROM του κάθε ερωτήματος, μπορεί να περιέχει μόνο ένα πίνακα. Μια από τις, μειωμένου αριθμού, λέξεις κλειδιά που διαθέτει η FQL είναι το IN και μπορεί να χρησιμοποιηθεί στα σημεία SELECT και WHERE, ώστε να δημιουργηθούν υποερωτήματα. Τα υποερωτήματα δεν μπορούν να κάνουν αναφορά σε μεταβλητές στο πεδίο εφαρμογής του εξωτερικού ερωτήματος. Τα ερωτήματα στη βάση δεδομένων του Facebook θα πρέπει επίσης να είναι ευρετηριοποίηση, πράγμα που σημαίνει ότι τουλάχιστο μία από ιδιότητες των ερωτημάτων που χρησιμοποιούνται στο σημείο WHERE πρέπει να είναι σημειωμένη ως ευρετηριοποίησης, ή αλλιώς ως κλειδί του συγκεκριμένου πίνακα στον οποίο απευθύνετε το ερώτημα.

Η FQL, όπως αναφέραμε και πιο πάνω, παρέχει μειωμένο αριθμό λειτουργιών. Μπορεί να χειριστεί απλές μαθηματικές πράξεις, βασικούς τελεστές Boolean και OXI λογικούς τελεστές, ούτε σημεία ORDER BY και LIMIT όπως υπάρχουν στην SQL. Το σημείο ORDER BY μπορεί να περιέχει μόνο ένα πίνακα. Για οποιαδήποτε ερώτημα που παίρνει μια ταυτότητα χρήστη (uid), μπορούμε να περάσουμε την συνάρτηση me() για να επιστρέψει το συνδεδεμένο χρήστη.

```
$fql = "SELECT name, url, pic FROM profile WHERE id IN
      (SELECT uid2 FROM friend WHERE uid1 = me() )";

$user_profile = $facebook->api( array(
    'method' => 'fql.query',
    'query' => $fql,
) );
```

Εικόνα 3-1 Παράδειγμα FQL με κώδικα PHP

jQuery: Είναι μια cross-browser JavaScript βιβλιοθήκη η οποία παρέχει την δυνατότητα ενσωμάτωσης script κώδικα στον html κώδικα με σκοπό την απλοποίηση του client-side scripting. Μερικές από τις λειτουργίες του είναι η εύκολη διαχείριση των DOM elements, η δημιουργία animations, διαχείριση events (e.g. Click event) ή ακόμα και η ανάπτυξη AJAX εφαρμογών. Στην δική μου εφαρμογή γίνεται συχνή χρήση της τεχνολογίας AJAX για το λόγο ότι επιτρέπει την αλλαγή συγκεκριμένης περιοχής του περιεχομένου (αποφεύγοντας τα postbacks), και παράλληλα αυξάνοντας το user experience.

PHP: Είναι μια scripting γλώσσα από την πλευρά του διακομιστή για την ανάπτυξη ιστοσελίδων, αλλά χρησιμοποιείται επίσης ως μια γενικής χρήσης γλώσσα προγραμματισμού. Από τον Ιανουάριο του 2013, η PHP έχει εγκατασταθεί σε περισσότερους από 240 εκατομμύρια ιστοσελίδες και 2,1 εκατομμύρια διακομιστές δικτύου. Δημιουργήθηκε από Rasmus Lerdorf το 1995.

PHP κώδικας ερμηνεύεται από έναν διακομιστή δικτύου με μια μονάδα επεξεργαστή PHP, η οποία παράγει την προκύπτουσα ιστοσελίδα. Εντολές PHP μπορεί να ενσωματωθούν άμεσα σε ένα έγγραφο HTML κώδικα και όχι καλώντας ένα εξωτερικό αρχείο για την επεξεργασία δεδομένων. Η PHP είναι ελεύθερο λογισμικό, που διατίθεται βάσει της PHP License. Η PHP έχει μεταφερθεί σε μεγάλο βαθμό και μπορεί να αναπτυχθεί σε περισσότερους διακομιστές δικτύων (σχεδόν σε κάθε λειτουργικό σύστημα και πλατφόρμα) δωρεάν.

Highcharts JS⁴: Η Highsoft AS είναι η εταιρία υπεύθυνη για την δημιουργία αυτής της έτοιμης βιβλιοθήκης δημιουργίας γραφικών παραστάσεων στη γλώσσα προγραμματισμού JavaScript. Κυκλοφόρησε στα τέλη του 2009. Το Highcharts έχει γίνει γρήγορα ένας από τους ηγέτες της βιομηχανίας σε πρότυπα συμβατά, με τη χρήση JavaScript, γραφήματα.

IBM SPSS: Είναι ένα πακέτο λογισμικού που χρησιμοποιείται για στατιστική ανάλυση. Αρχικά παραγόταν από την εταιρία SPSS A.E., που στη συνέχεια εξαγοράστηκε από την IBM το 2009. Το όνομα του λογισμικού σημαίνει Στατιστικό Πακέτο για τις Κοινωνικές Επιστήμες (Statistical Package for the Social Sciences - SPSS), αντιπροσωπεύοντας την αρχική αγορά, αν και το λογισμικό είναι πλέον δημοφιλές και σε άλλους τομείς, συμπεριλαμβανομένων των επιστημών υγείας και των επιστημών εμπορίας.

Το SPSS είναι ένα ευρέως χρησιμοποιούμενο πρόγραμμα για τη στατιστική ανάλυση στις κοινωνικές επιστήμες. Επίσης, χρησιμοποιείται από τους ερευνητές στο τομέα της αγοράς, τους ερευνητές στο τομέα της υγείας, εταιρείες που ασχολούνται κυρίως με έρευνες, κυβερνήσεις, για την εκπαίδευση ερευνητών, σε οργανώσεις εμπορίας, για εξόρυξη δεδομένων, και άλλα. Το αρχικό εγχειρίδιο του SPSS (Nie, Bent & Hull, 1970) έχει περιγραφεί ως ένα από τα "πιο σημαντικά βιβλία της κοινωνιολογίας" για το λόγο ότι επιτρέπει σε απλούς ερευνητές να κάνουν τη δική τους στατιστική ανάλυση. Εκτός από την στατιστική ανάλυση, χρησιμοποιείτε για τη διαχείριση δεδομένων και την τεκμηρίωση των δεδομένων, λειτουργίες οι οποίες είναι χαρακτηριστικά του βασικού λογισμικού.

3.6 Εργαλεία ανάπτυξης λογισμικού

Adobe Dreamweaver: Είναι μια εφαρμογή για τον σχεδιασμό ιστοσελίδων και την ανάπτυξη εφαρμογών που παρέχουν ένα οπτικό συντάχτη WYSIWYG, (κοινώς αναφέρεται ως το Design view), και ένα πρόγραμμα επεξεργασίας κώδικα με πρότυπα χαρακτηριστικά, όπως επισήμανση σύνταξης, ολοκλήρωση κώδικα και κατάρρευση κώδικα, καθώς και πιο εξελιγμένα

⁴ <http://www.highcharts.com/>

χαρακτηριστικά, όπως σύνταξη σε πραγματικό χρόνο και τον έλεγχο κώδικα ενδοσκόπησης για τη δημιουργία του κώδικα υπόδειξης για να βοηθήσει τον χρήστη στη σύνταξη κώδικα. Το Dreamweaver διαθέτει ένα ολοκληρωμένο πρόγραμμα περιήγησης για την προεπισκόπηση ανεπτυγμένων ιστοσελίδων στο παράθυρο προεπισκόπησης του προγράμματος του, καθώς επιτρέπει επίσης το περιεχόμενο να είναι ανοικτό σε τοπικά εγκατεστημένα προγράμματα περιήγησης στο διαδίκτυο. Ο πίνακας συμπεριφορών επιτρέπει τη χρήση των βασικών JavaScript χωρίς κωδικοποίηση της γνώσης, και ενσωμάτωση με το πλαίσιο της Adobe Spry Ajax που προσφέρει εύκολη πρόσβαση στο να δημιουργείται δυναμικά το περιεχόμενο και οι διασυνδέσεις.

MySQL Server: Είναι το δεύτερο πιο ευρέως χρησιμοποιούμενο ανοικτού κώδικα σύστημα διαχείρισης σχεσιακών βάσεων δεδομένων του κόσμου (RDBMS). Πήρε το όνομά της από την κόρη του συνιδρυτή Michael Widenius, την My. Η φράση SQL σημαίνει Structured Query Language. Το σχέδιο ανάπτυξης MySQL έχει κάνει τον πηγαίο κώδικα του διαθέσιμο υπό τους όρους της GNU (General Public License), καθώς και κάτω από μια ποικιλία ιδιοκτητών συμφωνιών. Ανήκει στην εταιρία Oracle Corporation.

Η MySQL είναι μια δημοφιλής επιλογή βάσης δεδομένων για χρήση σε εφαρμογές διαδικτύου, και αποτελεί βασικό συστατικό στοιχείο της ευρέως χρησιμοποιούμενης LAMP, στοίβας εφαρμογών διαδικτύου ανοικτού κώδικα. LAMP είναι ακρωνύμιο για το "Linux, Apache, MySQL, Perl / PHP / Python." Τα ελεύθερα έργα λογισμικού ανοικτού κώδικα που απαιτούν ένα πλήρως εξοπλισμένο σύστημα διαχείρισης βάσεων δεδομένων συχνά χρησιμοποιούν την MySQL.

3.7 Επίλογος

Στο κεφάλαιο αυτό περιγράφηκαν συνοπτικά τα πιο σημαντικά μέρη του πλαισίου που ακολουθήθηκε για τον ορισμό και ανάπτυξη της προτεινόμενης εφαρμογής. Στο κεφάλαιο που ακολουθεί θα αναλυθούν με λεπτομέρεια τα τεχνικά μέρη υλοποίησης του συστήματος και πως αυτά «επικοινωνούν» με τη βάση δεδομένων του Facebook.

Κεφάλαιο 4

Υλοποίηση

4.1 Εισαγωγή.....	32
4.2 Απαιτήσεις του Συστήματος	32
4.3 Προδιαγραφές	33
4.3.1 Προδιαγραφές από την Μεριά του Εξυπηρετητή.....	33
4.3.2 Προδιαγραφές από την Μεριά του Χρήστη.....	34
4.4 Προαπαιτήσεις του Συστήματος.....	35
4.5 Τεχνική Υλοποίηση του Συστήματος.....	35
4.5.1 Συλλογή Δεδομένων.....	36
Βάση Δεδομένων.....	37
Συλλογή δεδομένων δραστηριότητας του χρήστη	39
Συλλογής δεδομένων προσωπικότητας του χρήστη από το ερωτηματολόγιο.....	40
4.5.2 Επεξεργασία των Δεδομένων	41
Εξόρυξη Βασικών Δεδομένων από Facebook	41
Σχηματισμός Μοντέλου Χρήστη.....	49
Εξαγωγή Όμοιων Φίλων (Deriving Similar Friends).....	53
Συστάσεις ‘Ενδιαφέρον’ σελίδων στο Facebook (Recommendation of Interesting Facebook Pages)	54
Συγκεντρωτική Παρουσίαση Στοιχείων Δραστηριότητας στο Facebook	55
4.6 Επίλογος	56

4.1 Εισαγωγή

Σε αυτό το κεφάλαιο θα αναλύσουμε αρχικά τις απαιτήσεις, τις προαπαιτήσεις και τις προδιαγραφές της εφαρμογής που υλοποιήσαμε για την παρούσα διπλωματική εργασία. Ακολούθως, θα επεξηγήσουμε αναλυτικά τις τεχνικές λεπτομέρειες της υλοποίησης της εφαρμογής.

4.2 Απαιτήσεις του Συστήματος

Ο πρωτεύον στόχος της εφαρμογής, η οποία δημιουργήθηκε στα πλαίσια της παρούσας ατομικής διπλωματικής εργασίας, ήταν να δημιουργεί, σε γενικές γραμμές, ένα συμβόλαιο με τον χρήστη που προέβλεπε την ανάκτηση των προσωπικών του δεδομένων, τα οποία είναι αναγκαία για την λειτουργία της εφαρμογής, με σκοπό την ανάλυση τους και την παρουσίαση των αποτελεσμάτων. Δεδομένου ότι ο χρήστης θα αποδεχτεί τους όρους της εφαρμογής, θα συμφωνήσει δηλαδή να παραθέσει συγκεκριμένα προσωπικά του δεδομένα τα οποία ζητά η εφαρμογή, τότε στη συνέχεια, ξεκινά η εξαγωγή των πληροφοριών από την βάση δεδομένων του Facebook. Αυτό θα εκπληρωθεί με χρήση συγκεκριμένων ερωτημάτων προς την βάση δεδομένων του Facebook, τα οποία βασίζονται στο χρήστη και τις δραστηριότητες του στο κοινωνικό δίκτυο. Αφού ολοκληρωθεί με επιτυχία η ανάκτηση των δεδομένων, η εφαρμογή θα επεξεργαστεί τα δεδομένα αυτά και θα αποθηκεύσει τις ουσιώδεις πληροφορίες που δημιουργεί στην δική μας βάση δεδομένων.

Δεδομένου όμως ότι μια εφαρμογή στο Facebook αλληλεπιδρά απευθείας με το χρήστη, δεν θα ήταν προσιτό για τον χρήστη να δίνει απλά τα δεδομένα του και να μην παίρνει αποτέλεσμα για αυτή του την ενέργεια. Έτσι ο δεύτερος στόχος της εφαρμογής ήταν με εύχρηστο και ευπαρουσίαστο τρόπο να μπορεί ο χρήστης να αλληλεπιδράσει με τα αποτελέσματα του καθώς και να τα μοιραστεί με τους φίλους του, αφού βρισκόμαστε στο κοινωνικό δίκτυο Facebook, το οποίο ονομάζει τις σχέσεις των χρηστών του ως φιλίες.

4.3 Προδιαγραφές

Για να μπορεί η εφαρμογή να ενεργοποιηθεί και να μπορεί να ολοκληρώσει τον σκοπό της ανά το παγκόσμιο, δηλαδή να μπορεί οποιοσδήποτε άνθρωπος στο κόσμο να γίνει χρήστης του συστήματος μας και να την χρησιμοποιήσει από τον προσωπικό του υπολογιστή, θα πρέπει να ισχύουν κάποιες απαραίτητες τεχνικές προϋποθέσεις. Οι προδιαγραφές της εφαρμογής μας χωρίζονται σε 2 κατηγορίες, σε αυτές που πρέπει να ισχύουν από την πλευρά του εξυπηρετητή και αυτές από την πλευρά του χρήστη.

4.3.1 Προδιαγραφές από την Μεριά του Εξυπηρετητή

Εξυπηρετητής Διαδικτύου: Η εφαρμογή μας είναι της μορφής ιστοσελίδας η οποία χρησιμοποιεί, κατά κύριο της μέσο, την γλώσσα προγραμματισμού PHP για την συλλογή των δεδομένων των χρηστών που την χρησιμοποιούν, την επεξεργασία των προαναφερθέντων δεδομένων και εν τέλει, την δυναμική αναπαράσταση εξατομικευμένων δεδομένων στον χρήστη. Από τα προαναφερθέντα, καταλήγουμε στο συμπέρασμα ότι, απαραίτητη προδιαγραφή του εξυπηρετητή διαδικτύου που θα χρησιμοποιήσουμε, είναι να υποστηρίζει την εκτέλεση PHP κώδικα για μια ιστοσελίδα την οποία φιλοξενεί.

Βάση Δεδομένων: Η εφαρμογή μας, όπως έχουμε ήδη πολλάκις αναφέρει, έχει ως βασικό σκοπό να συλλέξει δεδομένα από το χρήστη για επεξεργασία και στη συνέχεια τη μελέτη. Τα δεδομένα αυτά θα προέρχονται από διαφορετικούς χρήστες οι οποίοι το μόνο κοινό που απαιτείτε να έχουν είναι το να είναι χρήστες του Facebook. Έτσι απαραίτητη προδιαγραφή του συστήματος που δημιουργούμε είναι να υποστηρίζεται από μια βάση δεδομένων στην οποία θα αποθηκεύονται και θα οργανώνονται τα δεδομένα με σκοπό την αποτελεσματική χρησιμοποίηση τους στη συνέχεια για την μελέτη που θα ακολουθήσει.

Ταυτοποίηση Εφαρμογής στο Facebook: Για να μπορεί να ολοκληρώσει η εφαρμογή μας την συλλογή των δεδομένων δραστηριότητας ενός χρήστη του Facebook, πρέπει να κατέχει τα απαραίτητα δικαιώματα ώστε αποδεχτεί η βάση δεδομένων του Facebook να επιστρέψει τα αποτελέσματα των ερωτήσεων που της παραθέτει η εφαρμογή. Εκτός από τα δικαιώματα του χρήστη, τα οποία θα εξηγήσουμε καλύτερα στο υποκεφάλαιο 4.3.2 χρειαζόμαστε επίσης και

δικαιώματα εγγεγραμμένης εφαρμογής στο Facebook. Τα δικαιώματα αυτά μπορούν να αποκτηθούν μόνο από κάποιο εγγεγραμμένο προγραμματιστή του Facebook, ο οποίος έχει ήδη γνώση των πολιτικών προστασίας προσωπικών δεδομένων του συγκεκριμένου κοινωνικού δικτύου και ο οποίος είναι υπεύθυνος στο να συγχρονίσει τον εξυπηρετητή δικτύου ο οποίος φιλοξενεί την εφαρμογή με το κέντρο εφαρμογών το οποίο διαθέτει το Facebook. Με την επιτυχία αυτής της ενέργειας ο προγραμματιστής δημιουργεί για την εφαρμογή μια μοναδική ταυτότητα και ένα μοναδικό κωδικό ο οποίος είναι απαραίτητος στη χρήση των ερωτημάτων της βάσης δεδομένων του Facebook. Εδώ περιέχονται τα μοναδικά δικαιώματα που εγγράφονται στο όνομα της εφαρμογής. Θα ήταν παράληψη να αναφέρουμε ότι, με την εκτέλεση αυτής της εισαγωγικής διαδικασίας της εφαρμογής σαν εγγεγραμμένη εφαρμογή του Facebook, ο προγραμματιστής δίνει στην εφαρμογή και μια μοναδική διεύθυνση σαν πόρο του Παγκόσμιου Ιστού εντός της διεύθυνσης της πλατφόρμας εφαρμογών του κοινωνικού δικτύου Facebook. Η διεύθυνση εντός του Facebook της εφαρμογής της παρούσας διπλωματικής εργασίας είναι https://apps.facebook.com/fb_cognitive_char/.

4.3.2 Προδιαγραφές από την Μεριά του Χρήστη

Διακριτικό Πρόσβασης: Το κοινωνικό δίκτυο του Facebook διαχειρίζεται τεράστια ποσότητα προσωπικών δεδομένων για τα εκατομμύρια χρήστες που το χρησιμοποιούν. Είναι λογικό ότι θα έχει σύνθετες πολιτικές προστασίας της ιδιωτικής ζωής των χρηστών του και εκτεταμένο πρόγραμμα ασφαλείας προς τα δεδομένα δραστηριότητας τους σε αυτό. Για τον λόγο αυτό είναι απαραίτητο για κάθε εφαρμογή στο Facebook, η οποία εκμεταλλεύεται κάποιες από τις πληροφορίες του χρήστη, εκτός από το να εγκριθεί από το Facebook, να δημιουργήσει ένα συμβόλαιο με κάθε χρήστη που έχει πρόθεση να την χρησιμοποιήσει. Το συμβόλαιο αυτό θα δημιουργήσει ένα μοναδικό διακριτικό πρόσβασης το οποίο θα δίνεται στην εφαρμογή, έμμεσα, από τον χρήστη, για να μπορεί η εφαρμογή να ολοκληρώσει επιτυχώς τα ερωτήματα στη βάση δεδομένων του Facebook. Η εφαρμογή δημιουργεί ένα μοναδικό διακριτικό για κάθε ένα από τους χρήστες της. Αξιοσημείωτο θα ήταν να αναφέρουμε ότι οποιοσδήποτε χρήστης του Facebook, για κάθε ξεχωριστή εφαρμογή την οποία δείχνει πρόθεση να χρησιμοποιήσει στο Facebook, πρέπει να δημιουργεί ένα ξεχωριστό διακριτικό πρόσβασης, χωρίς να παίζει ρόλο το γεγονός ότι κάποιες από τις εφαρμογές αυτές ζητάνε από τον χρήστη τα ίδια ακριβώς δεδομένα.

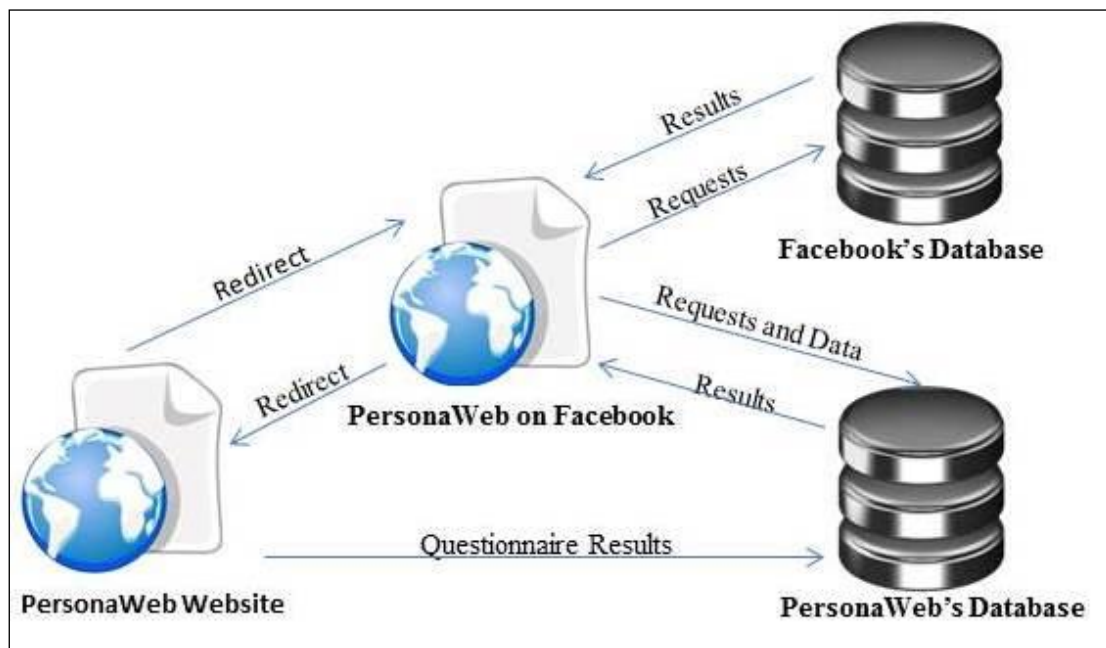
Καταληκτικά, η προδιαγραφή η οποία χρειάζεται από την μεριά του χρήστη, είναι η παραχώρηση του διακριτικού πρόσβασης στην εφαρμογή μας για εξόρυξη των δεδομένων του από τη βάση δεδομένων του Facebook.

4.4 Προαπαιτήσεις του Συστήματος

Ο χρήστης πρέπει να τηρεί κάποιες προαπαιτήσεις για να μπορεί να αλληλεπιδράσει με το σύστημα μας. Τεχνικές προαπαιτήσεις είναι να διαθέτει ο χρήστης ένα υπολογιστή με δυνατότητα πρόσβασης στο διαδίκτυο. Επίσης, οι λειτουργικές προαπαιτήσεις του συστήματος είναι πως πρέπει το λογισμικό του χρήστη να υλοποιεί οποιαδήποτε τεχνολογία δικτύου όπως είναι η Wi-Fi (IEEE 802.11) ή το Ethernet (IEEE 802.3) και επιπρόσθετα να μπορεί να έχει πρόσβαση στο διαδίκτυο μέσω οποιουδήποτε φυλλομετρητή.

4.5 Τεχνική Υλοποίηση του Συστήματος

Σε αυτό το υποκεφάλαιο θα εξηγήσουμε αναλυτικά την τεχνική υλοποίηση των τριών βασικών συστατικών του συστήματος που υλοποιήθηκε, στα πλαίσια της παρούσας διπλωματικής εργασίας. Στο Σχήμα 4-1 φαίνεται η βασική επικοινωνία της εφαρμογής με: ι) τη βάση δεδομένων του Facebook, από όπου θα εξάγουμε τα δεδομένα σχετικά με τον χρήστη, ιι) την ιστοσελίδα του PersonaWeb, όπου ο χρήστης θα μπορεί να συμπληρώσει το ερωτηματολόγιο προσωπικότητας με βάση το μοντέλο του “Big Five” και ιιι) με τη βάση δεδομένων PersonaWeb όπου θα αποθηκεύονται οι πληροφορίες σχετικά με τους χρήστες μας και τα αποτελέσματα του ερωτηματολογίου που συμπληρώνουν. Στα υποκεφάλαια που ακολουθούν θα αναλυθούν με λεπτομέρεια οι διαδικασίες αυτές.



Σχήμα 4-1 Διάγραμμα επικοινωνίας της εφαρμογής

4.5.1 Συλλογή Δεδομένων

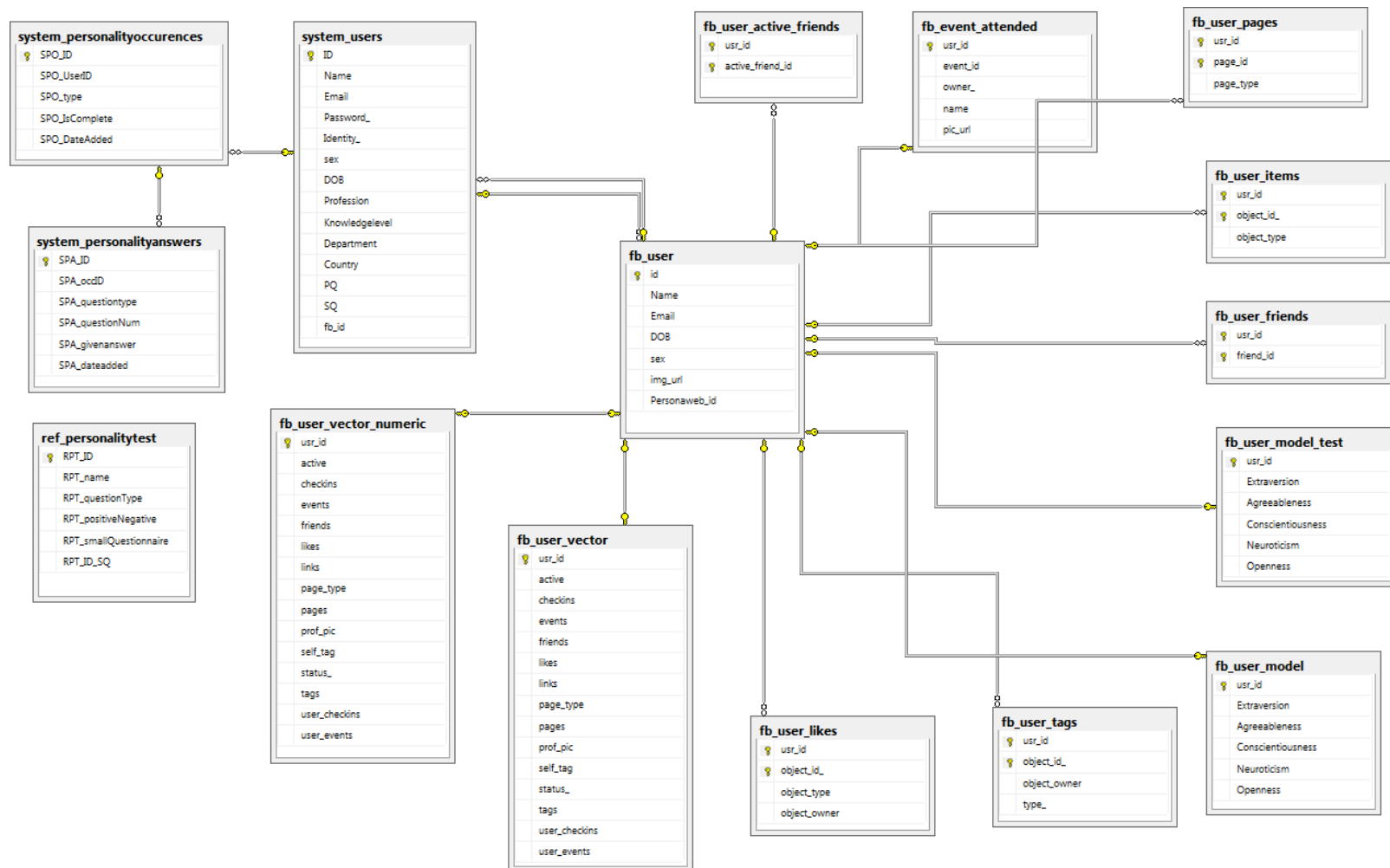
Για να εκπληρώσει τον στόχο της, η εφαρμογή που δημιουργήσαμε, πρώτο και βασικό κομμάτι το οποίο πρέπει να υλοποιηθεί είναι η συλλογή των δεδομένων του χρήστη. Η συλλογή των δεδομένων χωρίζεται σε δύο ξεχωριστές φάσεις. Πρωταρχική και πιο σημαντική από αυτές είναι η ανάκτηση των δεδομένων δραστηριότητας του χρήστη από την βάση δεδομένων του Facebook (Σχήμα 4-1). Το δεύτερο κομμάτι είναι η συλλογή των απαντήσεων του χρήστη στο ερωτηματολόγιο προσωπικότητας του “Big Five” το οποίο μας δίνει και τα ακριβής σκορ του χρήστη στις πέντε διαστάσεις της προσωπικότητας (Big Five). Απαραίτητη όμως προϋπόθεση για να μπορούμε να συλλέξουμε τα δεδομένα αυτά ήταν η δημιουργία της δικής μας βάσης δεδομένων η οποία θα κρατούσε τα στοιχεία αυτά για την μελέτη που θα εκτελεστεί και επιπρόσθετα, για την σύγκριση των χρηστών που εκτελείτε στην παρουσίαση των αποτελεσμάτων.

Βάση Δεδομένων

Η βάση δεδομένων που υλοποιούμε στην παρούσα διπλωματική εργασία φιλοξενείται στο Πανεπιστήμιο Κύπρου. Υλοποιεί γλώσσα MySQL και τα δεδομένα μας αποθηκεύονται στο κομμάτι που παραχωρεί το Πανεπιστήμιο μας, για το ερευνητικό έργο **PersonaWeb** (βλέπε υποκεφάλαιο 2.4). Στη βάση δεδομένων μας αποθηκεύουμε τις βασικές πληροφορίες εγγραφής του χρήστη, τα δεδομένα δραστηριότητας που δημιουργεί στο κοινωνικό δίκτυο του Facebook καθώς επίσης και τα αποτελέσματα των ερωτηματολογίων και τεστ προσωπικότητας τα οποία ολοκληρώνει στην ιστοσελίδα του PersonaWeb.

Για την δημιουργία των πινάκων ακολουθήθηκε μια μεθοδολογία κατασκευής με σκοπό την γρήγορη πρόσβαση στους πίνακες για διάβαση ή ενημέρωση των στοιχείων που κρατούν. Επίσης αφού δεν γνωρίζουμε την ποσότητα των δεδομένων που έχει κάποιος χρήστης διαθέσιμη, έπρεπε τα καλέσματα προς την βάση για αποθήκευση των δεδομένων αυτών να είναι όσο πιο μικρά και γρήγορα θα μπορούσαν αφού η γλώσσα προγραμματισμού PHP εκτελεί τα ερωτήματα εισαγωγής σειριακά (ένα – ένα).

Οι πίνακες μας ενώνονται κυρίως με τη χρήση της μοναδικής ταυτότητας του χρήστη την οποία δημιουργεί το Facebook για τον κάθε ένα, κατά την δημιουργία του λογαριασμού του στο κοινωνικό δίκτυο. Η ταυτότητα αυτή δεν διαλέχτηκε τυχαία για πρωτεύον κλειδί των πινάκων των δεδομένων μας, αφού και οι πίνακες στη βάση δεδομένων του Facebook χρησιμοποιούν την ίδια αυτή ταυτότητα για να αποθηκεύσουν τα δεδομένα δραστηριότητας του χρήστη κατά την δημιουργία τους. Αξίζει επίσης να αναφέρουμε ότι κάθε αντικείμενο δραστηριότητας που δημιουργείται στο Facebook, χαρακτηρίζεται από ένα μοναδικό αριθμός ως ταυτότητα. Στις περιπτώσεις που θέλουμε να αποφύγουμε την δημιουργία διπλότυπων εγγραφών στη βάση δεδομένων μας κατά την ενημέρωση του προφίλ του χρήστη, αυτή η ταυτότητα αντικειμένων χρησιμοποιείται σε κάποιους από τους πίνακες μας για κλειδί (μαζί με την ταυτότητα του χρήστη με τον οποίο συσχετίζεται). Στο Σχήμα 4-2 παρουσιάζεται το διάγραμμα της βάσης δεδομένων που δημιουργήθηκε για αυτή την εφαρμογή στα πλαίσια της παρούσας διπλωματικής εργασίας. Η δημιουργία του σχήματος της βάσης δεδομένων βρίσκεται στο Παράρτημα Α.



Σχήμα 4-2 Σχηματική περιγραφή της βάσης δεδομένων

Συλλογή δεδομένων δραστηριότητας του χρήστη

Αφού εκτελεστεί ο κώδικας ο οποίος εμφανίζει στο χρήστη το ειδικό συμβόλαιο, το οποίο πρέπει να αποδεχτεί για να μπορεί να γίνει χρήστης του συστήματος μας, τότε ξεκινά η διαδικασία ανάκτησης των δεδομένων δραστηριότητας του χρήστη. Όπως είχαμε προαναφέρει στο υποκεφάλαιο 4.3.2, για να μπορεί η εφαρμογή να ζητήσει μέσω ερωτημάτων στη βάση δεδομένων του Facebook δεδομένα για κάποιο συγκεκριμένο χρήστη, πρέπει να έχει στη διάθεση της το ειδικό διακριτικό πρόσβασης του συγκεκριμένου χρήστη. Η φάση της συλλογής των δεδομένων προϋποθέτει ότι τα απαραίτητα διακριτικά για την πρόσβαση στη βάση δεδομένων του Facebook ισχύουν και για τις δύο μεριές (εξυπηρετητής – χρήστης). Θα ήταν απαραίτητο να αναφέρουμε εδώ, ότι η διαδικασία που θα επεξηγηθεί περεταίρω, είναι η κοινή διαδικασία που εκτελείτο για κάθε ερώτημα το οποίο απευθυνόταν στην βάση δεδομένων του κοινωνικού δικτύου Facebook.

Η διαδικασία εκτέλεσης ενός ερωτήματος ξεκινά, γράφοντας το ερώτημα σαν μια συμβολοσειρά (`string`). Άξιο αναφοράς είναι το γεγονός ότι, για το σκοπό της σύνταξης των ερωτημάτων, εξατομικευμένων για το χρήστη, χρησιμοποιείτε η διαθέσιμη συνάρτηση της FQL γλώσσας “`me ()`” η οποία επιστρέφει την μοναδική ταυτότητα του τρέχων χρήστη, του οποίου χρησιμοποιεί την εφαρμογή μας. Στη συνέχεια, εκκινείτε μια ρήτρα «δοκίμασε – σύλλαβε» (`try – catch clause`), στην οποία εκτελείται το κάλεσμα του ερωτήματος στη βάση δεδομένων του Facebook ως αρχικό βήμα. Στην περίπτωση που επιτύχει το κάλεσμα του ερωτήματος, τότε το Facebook θα μας επιστρέψει την απάντηση με τα δεδομένα που ζητήσαμε σε μορφή αντικειμένου JSON. Από την άλλη, αν αποτύχει το κάλεσμα του ερωτήματος τότε θα εκτελεστεί ο κώδικας στη ρήτρα του «σύλλαβε» (`catch clause`) ο οποίος τυπώνει το λάθος στο αρχείο καταγραφής σφαλμάτων στον εξυπηρετητή για να μπορέσουν σε μελλοντικό στάδιο οι προγραμματιστές του συστήματος να διορθώσουν την οποιαδήποτε βλάβη υπάρχει. Θα ήταν παράληψη να μην αναφέρουμε εδώ, ότι μέσα στην ρήτρα του «δοκίμασε - σύλλαβε» εκτελείτε και η επιμέρους επεξεργασία του JSON αντικειμένου, για το συγκεκριμένο κάλεσμα σε κάθε ρήτρα, καθώς και η αποθήκευση των ουσιώδη πληροφοριών στην δική μας βάση δεδομένων. Η επεξεργασία αυτή των δεδομένων θα επεξηγηθεί περεταίρω στο υποκεφάλαιο 4.5.2. Σε καμία

περίπτωση δεν ευθύνεται η επεξεργασία ή η αποθήκευση στη δική μας βάση δεδομένων για την αποτυχημένη εκτέλεση μιας ρήτηρας «δοκίμασε - σύλλαβε».

Συλλογής δεδομένων προσωπικότητας του χρήστη από το ερωτηματολόγιο

Με την επιτυχή φόρτωση της εφαρμογής και με την ολοκλήρωση της συλλογής των δεδομένων δραστηριότητας του τρέχων χρήστη της εφαρμογής, όπως επεξηγούμε πιο πάνω, ακολουθεί η εκτέλεση του ερωτηματολογίου του τεστ προσωπικότητας “Big Five”. Για την εκκίνηση αυτής της διαδικασίας, καταρχάς ο χρήστης πρέπει να κάνει κλικ στο κουμπί **“Take it now!”**. Με το πάτημα του κουμπιού, εκτελείτε ο κώδικας ο οποίος θα εμφανίσει στο χρήστη ένα υπόμνημα μπροστά από τα υπόλοιπα δεδομένα. Εκεί ο χρήστης έχει τις επιλογές να εκτελέσει την σύντομη εκδοχή του “Big Five” τεστ προσωπικότητας, την εκτεταμένη του εκδοχή και επίσης να ακυρώσει αυτή του την ενέργεια και να επιστρέψει στην παρουσίαση των αποτελεσμάτων του από τα δεδομένα που συλλέξαμε και επεξεργαστήκαμε γι’ αυτόν. Εάν επιλέξει να εκτελέσει μια από τις δύο εκδοχές του τεστ προσωπικότητας τότε, η εφαρμογή παραπέμπει στην ιστοσελίδα του PersonaWeb η οποία φιλοξενεί το συγκεκριμένο ερωτηματολόγιο, καθώς και άλλα τεστ προσωπικότητας και τεστ γνωστικών χαρακτηριστικών, τα οποία έχει την επιλογή ο χρήστης να εκτελέσει για να ενημερωθεί για τα χαρακτηριστικά του εαυτού του. Αφού ολοκληρώσει την εκτέλεση του τεστ που έχει επιλέξει τότε το σύστημα ξεκινά μια διαδικασία αποθηκεύσεις των απαντήσεων που ο χρήστης έδωσε στο τεστ. Οι απαντήσεις για κάθε ερώτημα του παρουσιάζονται σαν μια ομάδα κουμπιών επιλογής, από τα οποία μπορεί να επιλέξει μόνο ένα για κάθε απάντηση. Εδώ θα ήταν παράληψη να μην αναφέρουμε, ότι για να μπορεί να γίνει η διόρθωση του τεστ και να παραχθούν τα συγκεντρωτικά αποτελέσματα του χρήστη για τις πέντε διαστάσεις της προσωπικότητας του “Big Five”, υπάρχει ήδη δημιουργημένος ένας πίνακας στη βάση δεδομένων του συστήματος μας, ο οποίος περιέχει κάθε ερώτηση και στοιχεία για αυτή όπως (i) για πια διάσταση της προσωπικότητας απευθύνεται, (ii) αν είναι θετικού ή αρνητικού βάρους προς αυτή τη διάσταση καθώς και (iii) σε πια από τις δύο εκδοχές (σύντομη ή εκτεταμένη) του τεστ ανήκει η συγκεκριμένη ερώτηση. Επιπρόσθετα, για τις απαντήσεις του χρήστη δημιουργούμε και ένα μοναδικό διακριτικό για το συγκεκριμένο περιστατικό εκτέλεσης του τεστ και αυτό γιατί αφού δίνουμε στο χρήστη την δυνατότητα να ξαναεκτελέσει το τεστ, τότε υπάρχει η πιθανότητα να υπάρχουν γι’ αυτών περισσότερες από μία

εκτελέσεις του συγκεκριμένου τεστ. Με την ολοκλήρωση του τεστ, εάν είναι χρήστης του PersonaWeb προήλθε από το Facebook τότε θα τον παραπέμψει πίσω στην εφαρμογή στην οποία μπορεί να δει τις διαφορές μεταξύ των μοντέλων που δημιουργούνται για αυτών (το μοντέλο που εξήγαγε η εφαρμογή και αυτό που έδωσε το ερωτηματολόγιο). Περισσότερες λεπτομέρειες για αυτό θα παρουσιαστούν στο υποκεφάλαιο 5.3 .

4.5.2 Επεξεργασία των Δεδομένων

Σε αυτό το υποκεφάλαιο θα επεξηγηθεί η ανάλυση η οποία γίνεται σε κάθε ένα από τα δεδομένα τα οποία ανακτούμε για το χρήστη από τα ερωτήματα στη βάση δεδομένων του Facebook, Επίσης θα περιγραφούν η επεξεργασία και ο αλγόριθμος πίσω από την δημιουργία του μοντέλου του χρήστη και την παρουσίαση των ουσιώδη δεδομένων που δημιουργούνται από την εφαρμογή.

Όπως έχουμε προαναφέρει στο υποκεφάλαιο 4.5.1, η επεξεργασία των επιμέρους δεδομένων τα οποία ανακτούμε από την βάση δεδομένων του Facebook, γίνεται για το κάθε ένα, μέσα στην ρήτρη «δοκίμασε – σύλλαβε» την οποία χρησιμοποιούμε για το κάλεσμα του εξειδικευμένου ερωτήματος για τη συγκεκριμένη πληροφορία. Στην συνέχεια, θα αναλυθεί η ξεχωριστή επεξεργασία που γίνεται για την ανάκτηση κάθε μιας από τις πληροφορίες που συλλέγουμε.

Εξόρυξη Βασικών Δεδομένων από Facebook

Ξεκινώντας, πρωτεύοντα ρόλο έχει **το ερώτημα το οποίο επιστρέφει τα αρχικά δεδομένα του χρήστη** τα οποία είναι (i) η μοναδική του ταυτότητα του προφίλ του στο Facebook, (ii) το όνομα του, (iii) η διεύθυνση του ηλεκτρονικού του ταχυδρομείου, (iv) η ημερομηνία γέννησης του καθώς και (v) η εικόνα προφίλ την οποία έχει στο Facebook. Τα δεδομένα αυτά του παρουσιάζονται σε φόρμα πάνω αριστερά στην εφαρμογή (Εικόνα 5-8) με σκοπό να κερδίσουμε την εμπιστοσύνη του χρήστη και να του δώσουμε την δυνατότητα να επιβεβαιώσει ότι τα ακόλουθα δεδομένα που θα δημιουργηθούν, του ανήκουν και είναι αληθείς. Η σημαντικότητα αυτού του ερωτήματος καθορίζεται από το ότι οι πληροφορίες που συλλέγονται είναι αυτές που θα χρησιμοποιηθούν για την εγγραφή του τρέχον χρήστη στο σύστημα μας. Ακολουθώντας, εκτελείτε ένα ερώτημα προς τη δική μας βάση δεδομένων στην οποία κοιτάζουμε αν υπάρχει

ήδη ο χρήστης εγγεγραμμένος στο σύστημα. Αν όντως δεν είναι χρήστης που χρησιμοποιεί την εφαρμογή για πρώτη φορά, τότε ενημερώνουμε τα δεδομένα που αναφέρονται πιο πάνω με σκοπό, το σύστημα να είναι ενημερωμένο με τα πιο πρόσφατα του δεδομένα (π.χ. αν έχει αλλάξει κάποια από τα δεδομένα αυτά, όπως για παράδειγμα την φωτογραφία προφίλ του). Από την άλλη, αν ο χρήστης χρησιμοποιεί την εφαρμογή για πρώτη φορά, τότε τα δεδομένα αυτά θα χρησιμοποιηθούν για να δημιουργήσουν μια νέα εγγραφή στη βάση δεδομένων μας για το χρήστη. Στη συνέχεια, θα εγγραφεί και σαν χρήστης του συστήματος PersonaWeb για να του ανατεθεί μια ξεχωριστή ταυτότητα με την οποία θα μπορεί να εκτελεί ερωτηματολόγια και τεστ. Αφού έχουμε τώρα και τις δύο μοναδικές ταυτότητες του χρήστη, τότε γίνεται συγχρονισμός των δύο αυτών εγγραφών, ενημερώνοντας η μία την άλλη για την αντίστοιχη μοναδική ταυτότητα. Εδώ πρέπει να αναφερθεί, ο λόγος που γίνεται αυτός ο συγχρονισμός: Τα ερωτήματα προς τη βάση δεδομένων του Facebook και τα ερωτήματα για τα σκορ στο τεστ προσωπικότητας του χρήστη, βρίσκονται σε δύο ξεχωριστά συστήματα και χρησιμοποιεί το καθένα τη δική του ταυτότητα για το χρήστη, ώστε να θέσει τη μοναδικότητα των δεδομένων του στη βάση δεδομένων του αντίστοιχου συστήματος.

Ακολουθεί το **ερώτημα για την ανακτήση των 'friends' του χρήστη στο Facebook**. Όταν επιστραφεί η απάντηση από τη βάση δεδομένων του Facebook τότε, εντοπίζουμε τη μοναδική ταυτότητα του φίλου του τρέχων χρήστη και δημιουργούμε για κάθε έναν, μια εγγραφή στον πίνακα ``fb_user_friends`` με τις δύο αυτές ταυτότητες. Επιπρόσθετα βρίσκουμε τον συνολικό αριθμό φίλων που έχει ο χρήστης, τον οποίο θα χρησιμοποιήσουμε στη συνέχεια κατά την παραγωγή του μοντέλου του χρήστη.

Στη συνέχεια εκτελείτε το **ερώτημα ανάκτησης των 'likes' του χρήστη** σε διάφορες δημοσιεύσεις στο Facebook. Παίρνοντας την απάντηση με τα σχετικά δεδομένα από το Facebook τότε εντοπίζουμε για κάθε εγγραφή του πίνακα του οποίου μας επιστρέφεται το μοναδικό αριθμό του αντικειμένου στο οποίο έκανε like ο τρέχων χρήστης καθώς και τον τύπο του αντικειμένου αυτού. Αποθηκεύουμε τα δεδομένα αυτά με πρωτεύον κλειδί την ταυτότητα του χρήστη και τον αριθμό του αντικειμένου. Αυτό γίνεται για δύο λόγους: α) για να αποφύγουμε την δημιουργία διπλότυπων εγγραφών στη δική μας βάση δεδομένων και β) καθορίζουμε κάθε εγγραφή στο πίνακα ``fb_user_likes`` σαν ένα πραγματικό like. Ακολούθως καταμετρούμε το μέγεθος του αποτελέσματος που μας επιστράφηκε για να δούμε τον συνολικό

αριθμό likes που έχει κάνει ο χρήστης. Αξίζει εδώ να αναφέρουμε ότι ο μέγιστος αριθμός δεδομένων που κρατά στο συγκεκριμένο πίνακα το Facebook για ένα χρήστη είναι το 1000.

Συνεχίζοντας την ανάκτηση των δεδομένων για τα likes του χρήστη, όταν αποθηκευτούν επιτυχώς τα στοιχεία από το προηγούμενο ερώτημα τότε καλείτε μια συνάρτηση η οποία με τη σειρά της θα δημιουργήσει **οκτώ ερωτήματα προς τη βάση δεδομένων του Facebook για την ανάκτηση των ιδιοκτητών ή δημιουργών (owners ή creators) των αντικειμένων που έκανε like ο τρέχων χρήστης**. Κάθε ένα από αυτά τα ερωτήματα απευθύνετε σε ένα ξεχωριστό πίνακα της βάσης δεδομένων του Facebook. Ο λόγος είναι γιατί το Facebook οργανώνει τα αντικείμενα του, βάσει του τύπου τους σε ξεχωριστούς πίνακες. Έτσι για να εκτελεστεί το κάλεσμα του ερωτήματος για κάθε ένα από τους διάφορους τύπους, πρέπει καταρχάς να διαβαστούν από την δική μας βάση δεδομένων τα likes του τρέχων χρήστη σε αντικείμενα του συγκεκριμένου τύπου. Αργότερα, θα συγκεντρωθούν οι μοναδικοί αριθμοί, αριθμητικοί κωδικοί των αντικειμένων τους οποίους συλλέξαμε με το προηγούμενο ερώτημα και θα συμμαζευτούν σε μια συμβολοσειρά. Αυτή η συμβολοσειρά θα χρησιμοποιηθεί στο σημείο WHERE του ερωτήματος που θα δημιουργηθεί για τον πίνακα του συγκεκριμένου τύπου αντικειμένου. Το αποτέλεσμα του επιτυχές καλέσματος του ερωτήματος θα μας επιστρέψει ένα αντικείμενο με τον κωδικό αριθμό του αντικειμένου καθώς και την ταυτότητα του ιδιοκτήτη ή δημιουργού του αντικειμένου αυτού. Έχοντας τώρα τα δύο προαναφερθέντα δεδομένα και την ταυτότητα του χρήστη, τότε μέσα στην συγκεκριμένη ρήτρα «δοκίμασε – σύλλαβε» που δημιουργείται για κάθε διαφορετικού τύπου αντικείμενα, γίνετε η ενημέρωση των εγγραφών της δικής μας βάσης δεδομένων του πίνακα ``fb_user_likes``.

Στη συνέχεια θα εκτελεστεί το **ερώτημα για την ανάκτηση των σελίδων του Facebook που αρέσουν του χρήστη**. Το ερώτημα αυτό είναι πιο σύνθετο από τα υπόλοιπα ερωτήματα που έχουμε αναφέρει μέχρι τώρα για το λόγο ότι πρέπει ταυτόχρονα να ανακτήσουμε δεδομένα από ένα πίνακα του Facebook και ταυτόχρονα να τα χρησιμοποιήσουμε σε δεύτερο πίνακα του Facebook για να μας επιστραφεί στο τέλος ένα αντικείμενο με τα δεδομένα που θέλουμε. Αυτό το σύνθετο ερώτημα επιτυγχάνεται με την εξής τεχνική: Μέσα στο κύριο ερώτημα (το οποίο θέλουμε να μας επιστρέψει την ταυτότητα και τον τύπο της σελίδας του Facebook που αρέσουν του χρήστη) στο σημείο WHERE αυτού του ερωτήματος χρησιμοποιούμε την εντολή IN. Η εντολή IN μας επιτρέπει να εκκινούμε ένα εσωτερικό ερώτημα σε άλλο πίνακα της βάσης

δεδομένων του Facebook και τα αποτελέσματα αυτού να τα χρησιμοποιήσουμε στο εξωτερικό ερώτημα. Το εσωτερικό ερώτημα που χρησιμοποιούμε σε αυτή την περίπτωση, μας επιστρέφει ποιες σελίδες του Facebook ο τρέχων χρήστης έκανε like. Από το αντικείμενο το οποίο παίρνουμε ως απάντηση, δημιουργούμε μια εγγραφή στο πίνακα ``fb_user_pages`` στην βάση δεδομένων μας, με την ταυτότητα του χρήστη, την ταυτότητα της σελίδας και τον τύπο της. Επιπρόσθετα κρατούμε τον συνολικό αριθμό σελίδων που αρέσουν στο χρήστη.

Συνεχίζουμε την ανάκτηση των δεδομένων του τρέχων χρήστη, **συλλέγοντας αντικείμενα στα οποία είναι tagged**. Για αυτό το είδος δεδομένων δημιουργήσαμε συγκεκριμένο πίνακα τον ``fb_user_tags``. Όπως και προηγουμένως, πρέπει να ψάξουμε για αυτά τα δεδομένα σε ξεχωριστούς πίνακες της βάσης δεδομένων του Facebook. Ξεκινάμε με τα tags του χρήστη τα οποία έγιναν σε βίντεο στο Facebook. Παρόμοια με πριν, χρησιμοποιούμε ένα σύνθετο ερώτημα για αυτή τη πληροφορία για το λόγο ότι παίρνουμε από το εσωτερικό ερώτημα τους μοναδικούς κωδικούς του κάθε βίντεο που συσχετίζεται με το χρήστη και χρησιμοποιούμε τις πληροφορίες αυτές για να πάρουμε ως απάντηση το κάθε βίντεο με την μοναδική ταυτότητα του ιδιοκτήτη του. Εντοπίζοντας αυτές τις πληροφορίες από το αντικείμενο που μας επιστρέφεται, δημιουργούμε στο προαναφερόμενο πίνακα μια εγγραφή καθορίζοντας τον τύπο του συγκεκριμένου αντικειμένου στην εγγραφή. Στη συνέχεια εκτελούμε ακριβώς την ίδια διαδικασία για αντικείμενα τύπου εικόνες. Και για αυτά πρέπει να δημιουργηθεί ένα σύνθετο ερώτημα με αντίστοιχα δεδομένα το καθένα όπως και για τα βίντεο και επίσης αποθηκεύονται στον ίδιο πίνακα όπως τα προηγούμενα αφού αφορούν τον ίδιο τύπο πληροφορίας, δηλαδή, τα tags.

Για να ολοκληρώσουμε την **συλλογή δεδομένων για τα tags του χρήστη** στο Facebook, εκτελούμε πανομοιότυπο σύνθετο ερώτημα για την ανάκτηση των αντικειμένων τύπου **check-ins**, για τα οποία το εσωτερικό query χρησιμοποιεί την ταυτότητα του χρήστη μέσω της συνάρτησης `me()`, για να εντοπίσει αν βρίσκεται στη λίστα με τις ταυτότητες που έχει φυλαγμένο η εγγραφή του κάθε check-in στη βάση δεδομένων του Facebook. Αυτή η πληροφορία σε συνεργασία με τον δημιουργό του check-in θα μας επιστρέψουν τα check-in στα οποία συμμετείχε ο χρήστης με κάποιους φίλους του. Να αναφέρω εδώ, ότι οι ταυτότητες των φίλων του χρήστη που χρησιμοποιεί το πιο πάνω ερώτημα, συλλέγονται και χρησιμοποιούνται από ένα εσωτερικό ερώτημα στο πίνακα που κρατά τις απαραίτητες

πληροφορίες για τους φίλους του τρέχων χρήστη. Τέλος θα ενημερώσουμε τον πίνακα της βάσης δεδομένων μας που δημιουργήσαμε για τα tags του χρήστη, προσθέτοντας τις εγγραφές για τον χρήστη σε αντικείμενα τύπου `check-in` που μόλις συλλέξαμε.

Θα ήταν βασικό εδώ να αναφέρουμε, ότι για τα προηγούμενα τρία ερωτήματα κρατούμε τον συνολικό τους αριθμό και στο τέλος προσθέτουμε αυτούς τους τρεις αριθμούς για να έχουμε συγκεντρωτικά πόσα tags έχει ο τρέχων να αναφέρουμε, το οποίο θα χρησιμοποιηθεί μετά κατά την δημιουργία του μοντέλου του χρήστη.

Ακολούθως **θα συλλέξουμε δεδομένα για τα check-ins τα οποία δημιούργησε ο τρέχων χρήστης**. Εκτελούμε ένα απλό ερώτημα στη βάση δεδομένων του Facebook ζητώντας τα `check-ins` τα οποία έχουν καταγραφμένο σαν δημιουργό τους τον χρήστη μας, χρησιμοποιώντας την συνάρτηση `me()`. Ο λόγος που αυτό το ερώτημα είναι απλό είναι γιατί δεδομένα όπως δημιουργό (`creator`) και τύπο αντικειμένου (`object_type`) που ζητούσαμε πριν, στη συγκεκριμένη αυτή περίπτωση, τα γνωρίζουμε από την αρχή. Δημιουργούμε έτσι εγγραφές με την ταυτότητα του χρήστη, τον μοναδικό κωδικό του αντικειμένου και τον τύπο του στον ειδικό πίνακα που δημιουργήσαμε για τα αντικείμενα των χρηστών μας, τον ``fb_user_items``.

Στη συνέχεια θα εκτελέσουμε **ερωτήματα για να εξορύξουμε τις συσχετίσεις του τρέχων χρήστη με τις εκδηλώσεις** οι οποίες δημιουργούνται στο Facebook. Για την απόκτηση των πληροφοριών αυτών χωρίσαμε την ανάκτηση των δεδομένων σε δύο κατηγορίες. Τις εκδηλώσεις στις οποίες ο χρήστης παραβρέθηκε και τις εκδηλώσεις τις οποίες δημιούργησε ο ίδιος. Για τις εκδηλώσεις τις οποίες παραβρέθηκε, χρειαζόμαστε μια προεπεξεργασία στην οποία, πρέπει να ανακτήσουμε την τρέχων ώρα και ημερομηνία στην οποία εκτελείτε ο κώδικας και να την μετατρέψουμε στην μορφή UNIX Timestamp.

UNIX Timestamp: Είναι ένας τρόπος για παρακολούθηση του χρόνου ως τρέχων σύνολο δευτερολέπτων. Είναι μια αριθμητική μορφή παρουσίας της ημερομηνίας και ώρας την συγκεκριμένη στιγμή που μεταφράζεται σε δευτερόλεπτα που πέρασαν από την Unix Epoch, η οποία είναι η 1^η Ιανουαρίου 1970 σε UTC (Coordinated Universal Time). Θα πρέπει επίσης να επισημανθεί ότι αυτό το σημείο στο χρόνο, τεχνικά, δεν αλλάζει όπου κι αν βρίσκονται στον κόσμο. Αυτό είναι πολύ χρήσιμο για τα συστήματα πληροφορικής, για τον εντοπισμό και τη

διαλογή επικαιροποιημένων πληροφοριών, δηλαδή πληροφοριών που συσχετίζονται με χρόνο, σε δυναμικές και κατανεμημένες εφαρμογές στο διαδίκτυο.

Αφού γίνει η προεπεξεργασία που αναφέραμε, τότε δημιουργούμε ένα ερώτημα προς τη βάση δεδομένων του Facebook δηλώνοντας ότι θέλουμε να μας επιστρέψει στοιχεία για εκδηλώσεις τις οποίες είχαν στην λίστα καλεσμένων τους τον τρέχων χρήστη, ο συγκεκριμένος χρήστης δήλωσε ότι θα παραστεί σε αυτήν και η εκδήλωση συνέβηκε πριν από την τρέχων ώρα την οποία εκτελείτε το ερώτημα. Αξίζει εδώ να αναφέρουμε ότι αν παραβλέπαμε την προεπεξεργασία την οποία εξηγούμε πιο πάνω, τότε το ερώτημα θα μας επέστρεφε μόνο τις εκδηλώσεις τις οποίες ο χρήστης δήλωσε ότι θα πάει αλλά δεν έχουν λάβει χώρα ακόμη. Αφού επιστρέψει η βάση δεδομένων του Facebook το αντικείμενο με την απάντηση που ζητούμε, τότε δημιουργούμε εγγραφές σε ένα πίνακα στη βάση δεδομένων μας στον οποίο κρατούμε μόνο πληροφορίες για τις εκδηλώσεις τις οποίες παραβρέθηκαν οι χρήστες του συστήματος μας. Οι πληροφορίες οι οποίες διαλέγουμε από την απάντηση είναι (i) ο μοναδικός κωδικός του αντικειμένου, (ii) ο δημιουργός της εκδήλωσης, (iii) την εικόνα την οποία ανέθεσε ο δημιουργός για την συγκεκριμένη εκδήλωση καθώς και (iv) το όνομα της εκδήλωσης. Εν τέλει, κρατούσαμε το συνολικό αριθμό των εκδηλώσεων που παραβρέθηκε.

Για τις εκδηλώσεις τις οποίες δημιούργησε ο ίδιος, το ερώτημα που δημιουργείτε είναι πολύ πιο απλοϊκό αφού χρειάζεται την ταυτότητα του χρήστη ως δημιουργό της εκδήλωσης και παίρνουμε τα αποτελέσματα που θέλουμε. Η εγγραφή που δημιουργείτε με το μοναδικό κωδικό της κάθε εκδήλωσης του χρήστη, αποθηκεύετε στο πίνακα ``fb_user_items`` τον οποίο χρησιμοποιήσαμε και προηγουμένως.

Για να ολοκληρώσουμε την σύναξη των δεδομένων για τον πίνακα ``fb_user_items``, χρειαζόμασταν επίσης **τις φωτογραφίες και τα βίντεο τα οποία ανέβασε στο Facebook, της καταστάσεις τις οποίες έγραψε και κοινοποίησε στο κοινωνικό δίκτυο ο χρήστης καθώς επίσης και τους συνδέσμους του διαδικτύου τους οποίους κοινοποίησε στον «τοίχο» του στο Facebook.**

Και για τα τέσσερα αυτά ξεχωριστά δεδομένα χρησιμοποιούμε την ίδια απλοϊκή μεθοδολογία, για απόκτηση αυτών των δεδομένων. Στους τέσσερεις ανάλογους πίνακες της βάσης δεδομένων του Facebook εκτελείτε ερώτημα θέτοντας στο σημείο WHERE ως ιδιοκτήτη ή δημιουργό την ταυτότητα του τρέχων χρήστη μας, χρησιμοποιώντας την προαναφερθέντα

συνάρτηση της FQL, με (). Για κάθε ένα αντικείμενο που μας επιστρεφόταν ως αποτέλεσμα του αντίστοιχου ερωτήματος, δημιουργούσαμε εγγραφές στον πίνακα `fb\_user\_items`, χρησιμοποιώντας την ταυτότητα του χρήστη, τον μοναδικό κωδικό του κάθε αντικειμένου, καθώς και τον αντίστοιχο τύπο των αντικειμένων τον οποίο γνωρίζαμε εμείς για κάθε εκτέλεση ερωτήματος από τις τέσσερις ξεχωριστές εκτελέσεις.

Στη συνέχεια εκτελείτε το **ερώτημα το οποίο θα μας δώσει τις πληροφορίες για τους ενεργούς φίλους του χρήστη**, βάση της δραστηριότητας τους στο κοινωνικό δίκτυο. Για την επιτυχή ανάκτηση αυτής της ουσιώδους πληροφορίας του χρήστη, χρειάστηκε προεργασία πριν την εκτέλεση του ερωτήματος στη βάση δεδομένων του Facebook, καθώς και μια εκτεταμένη επεξεργασία των δεδομένων που μας επιστρέφονται ως αποτέλεσμα.

Η κύρια υπόθεση που ακολουθήσαμε για να ανακτήσουμε την πληροφορία των ενεργών φίλων ενός χρήστη στο Facebook ήταν η εξής: Για να θεωρηθεί κάποιος χρήστης, ο οποίος είναι φίλος του τρέχων χρήστη του συστήματος μας, ως ενεργός του φίλος, τότε πρέπει να έχει δημοσιεύσει ή μοιραστεί δημοσίευση⁵ στον «τοίχο» του χρήστη μας τουλάχιστο τέσσερις φορές εντός ενός χρόνου. Ο λόγος που θεωρούμε κάποιον που έχει κάνει πάνω από τέσσερις δημοσιεύσεις, είναι για να αποφύγουμε φίλους του χρήστη μας οι οποίοι ίσως του ευχήθηκαν για τα γενέθλια ή την ονομαστική του εορτή κατά την διάρκεια ενός χρόνου. Η προσέγγιση αυτή είναι παρόμοια με το τι ακολούθησαν οι [7].

Βάσει της πιο πάνω υπόθεσης, η προεπεξεργασία που χρειαζόμασταν για να εκτελέσουμε το ερώτημα στη βάση δεδομένων του Facebook ήταν να δημιουργήσουμε σε μορφή UNIX Timestamp, που εξηγήσαμε πιο πάνω, για ένα χρόνο πίσω από την τρέχουσα ώρα και ημερομηνία. Αφού από προηγούμενη ανάκτηση δεδομένων δημιουργήσαμε την τρέχων ώρα και ημερομηνία στη μορφή που θέλαμε, τότε απόμενε να αφαιρέσουμε από αυτή τα δευτερόλεπτα τα οποία εμπεριέχονται σε ένα χρόνο. Με την νέα τιμή που δημιουργούμε, εκτελούμε το ερώτημα

⁵ Για να μοιραστείς μια δημοσίευση σου με κάποιο φίλο σου στο Facebook, πρέπει είτε να τον αναφέρεις μέσα στο κείμενο της δημοσίευσης, είτε να τον κάνεις tag στη συγκεκριμένη δημοσίευση είτε να δηλώσεις ότι είσαι μαζί του την ώρα που την δημιουργείς και κοινοποιείς.

και ζητούμε όλες τις δημοσιεύσεις οι οποίες, έγιναν στο «τοίχο» του τρέχων χρήστη μας, από κάποιο άλλο εκτός από τον ίδιο και να έχουν δημιουργηθεί μέσα σε ένα χρόνο απόσταση από την τρέχων ώρα και ημερομηνία.

Παραλαμβάνοντας το αντικείμενο με τα αποτελέσματα του ερωτήματος που θέλαμε, ακολουθεί μια επεξεργασία των δεδομένων για οργάνωση τους κατά χρήστη έτσι ώστε να μπορούν να δεχτούν καταμέτρηση και να μας παράξουν τις πληροφορίες που θέλουμε. Για αυτή την επεξεργασία χρειαζόμαστε δύο δυναμικούς, προσωρινούς πίνακες της γλώσσας προγραμματισμού PHP. Ο ένας θα κρατά τις ταυτότητες των χρηστών που δημιούργησαν την κοινοποίηση στο «τοίχο» του τρέχων χρήστη μας και ο άλλος θα χρησιμοποιεί τις ταυτότητες αυτές για θέσεις, ή καλύτερα, ευρετήρια του πίνακα και θα κρατά για το καθένα ποσοτικά, το σύνολο των δημοσιεύσεων που έκανε. Για κάθε εγγραφή που διαβάζουμε από το αντικείμενο του αποτελέσματος, τότε ελέγχουμε εάν υπάρχει στον πίνακα που αναφέραμε ο οποίος κρατά της ταυτότητες. Εάν δεν υπάρχει τότε τον προσθέτουμε στο πίνακα καθώς επίσης και δημιουργούμε θέση στον άλλο πίνακα με την ταυτότητα του και την αρχικοποιούμε με την τιμή ένα με τη σκέψη ότι αφού είναι η πρώτη φορά που τον συναντούμε τότε είναι και η πρώτη δημοσίευση που κάνει στο «τοίχο» του τρέχων χρήστη. Σε αντίθετη περίπτωση, εάν ήδη υπάρχει στο πίνακα με τις ταυτότητες τότε βρίσκουμε την θέση που κρατούμε, στο πίνακα με τις τιμές, με την ταυτότητα του και προσθέτουμε ένα στην τιμή που ήδη υπάρχει στην συγκεκριμένη θέση.

Στη συνέχεια και όταν έχουμε τελειώσει με το αποτέλεσμα του αντικειμένου που λάβαμε ως απάντηση, προχωρούμε χρησιμοποιώντας μόνο τους δύο δυναμικούς πίνακες που μόλις συμπληρώσαμε. Βρίσκουμε το μέγεθός του πίνακα με τις ταυτότητες, και αφού όπως εξηγήσαμε πιο πάνω προσθέτουμε θέση στον δεύτερο πίνακα μόνο αν θα εισάγουμε τιμή στο πρώτο, τότε το μέγεθος των δύο αυτών πινάκων είναι κοινό. Για κάθε μια από αυτές τις θέσεις ελέγχουμε: αν η τιμή η οποία έχουν, μετά το πέρας της ανάλυσης των δεδομένων του αποτελέσματος, είναι μεγαλύτερη από 4, τότε δημιουργούμε μια εγγραφή στη βάση δεδομένων μας, στον πίνακα τον οποίο κρατάμε τους ενεργούς φίλους των χρηστών μας, τον ``fb_user_active_friends``. Επιπρόσθετα αυξάνουμε ένα μετρητή ο οποίος στο τέλος, θα κρατά συγκεντρωτικά τον συνολικό αριθμό των ενεργών φίλων του τρέχων χρήστη μας.

Συνεχίζοντας, εκτελούμε **ερώτημα για την ανάκτηση της πληροφορίας αν ο χρήστης έκανε tag τον εαυτό του σε κάποιο αντικείμενο ή δημοσίευση**. Ένα απλό ερώτημα προς την

βάση δεδομένων δημιουργείτε το οποίο ζητά τους μοναδικούς κωδικούς των αντικειμένων όπου ο χρήστης μας είναι tagged και το άτομο που τον έκανε είναι ο ίδιος. Για αυτή την πληροφορία δεν δημιουργούμε καμία εγγραφή στη βάση δεδομένων μας. Απλά, κρατάμε τον συνολικό αριθμό αυτών των δεδομένων.

Τέλος, το τελευταίο ερώτημα που εκτελούμε είναι για να μάθουμε το μέγεθος του άλμπουμ του χρήστη με τις φωτογραφίες προφίλ του. Όπως και για το προηγούμενο ερώτημα, δεν δημιουργούμε εγγραφές στη βάση δεδομένων μας, αλλά κρατούμε το μέγεθος του άλμπουμ αυτού το οποίο θα χρησιμοποιηθεί στην συνέχεια, κατά την δημιουργία του μοντέλου του χρήστη, για να βρούμε το ποσοστό των φωτογραφιών προφίλ σε σχέση με όλες τις φωτογραφίες του χρήστη.

Τα ερωτήματα προς τη βάση δεδομένων του Facebook βρίσκονται συγκεντρωμένα στο Παράρτημα Δ.

Σχηματισμός Μοντέλου Χρήστη

Αφού ολοκληρωθεί η συλλογή όλων των δεδομένων του χρήστη και η επεξεργασία των επιμέρους πληροφοριών, ακολουθεί η δημιουργία του μοντέλου προσωπικότητας του. Για την ολοκλήρωση αυτού του κομματιού δημιουργούμε δύο διανύσματα για το χρήστη. Το πρώτο διάνυσμα περιέχει τις αριθμητικές στατιστικές τις οποίες δημιουργήσαμε και συλλέξαμε κατά την ανάκτηση των δεδομένων του χρήστη. Αυτό το διάνυσμα μπορεί να χρησιμοποιηθεί αργότερα για μια συγκεντρωτική μελέτη των δεδομένων του συστήματος. Για την δημιουργία του δεύτερου διανύσματος, χρησιμοποιείτε το πρώτο διάνυσμα που αναφέραμε και συγκρίνονται οι τιμές του με κάποια όρια (thresholds) τα οποία θέσαμε για τα δεδομένα μας. Το συγκεκριμένο διάνυσμα είναι το κύριο μέσω το οποίο θα χρησιμοποιηθεί και στη συνέχεια για τον υπολογισμό, της κύριας υπόθεσης της προσωπικότητας, που κάνουμε στην παρούσα διπλωματική εργασία, καθώς επίσης και για την σύγκριση των χρηστών του συστήματος μας για τον εντοπισμό ομοιότητας μεταξύ τους. Αυτά θα αναλυθούν περεταίρω στη συνέχεια.

Τα όρια (thresholds) τα οποία αναφέραμε είναι συγκεκριμένες τιμές οι οποίες χρησιμοποιούνται για να συγκρίνουμε τα συγκεντρωτικά δεδομένα τα οποία αποτελούν το πρώτο διάνυσμα του χρήστη και να δημιουργούν νέες τιμές. Οι νέες τιμές τις οποίες

προσπαθούμε να δημιουργήσουμε κυμαίνονται μεταξύ του ένα και του πέντε για κάθε στοιχείο μιμούμενοι τις τιμές που παίρνουν οι απαντήσεις που μπορεί να δώσει ο χρήστης σε ένα ερωτηματολόγιο προσωπικότητας. Για κάθε δεδομένο θέταμε τέσσερις τιμές για όρια και ακολούθως, ελέγχουμε αν η τιμή που είχαμε δημιουργήσει, για το συγκεκριμένο δεδομένο, κατά την εκτέλεση της ανάκτησης των δεδομένων, ήταν συγκριτικά είτε μικρότερη του κατώτατου ορίου, είτε κάπου ανάμεσα στα όρια, είτε ακόμη, μεγαλύτερη και από το ανώτατο όριο.

Για την δημιουργία και τον καθορισμό των ορίων αυτών, τα οποία ήταν μοναδικά για κάθε είδος δεδομένου που χρησιμοποιούσαμε, εκτελέσαμε μια σύντομη προμελέτη, μελετώντας τα δεδομένα από πέντε εθελοντές φοιτητές. Οι φοιτητές αυτοί ήταν διαφορετικών κλάδων πρόσφεραν πληροφορίες ώστε να καθοριστεί ο όγκος της δραστηριότητας τους στο κοινωνικό δίκτυο Facebook. Οι φοιτητές που επιλέχτηκαν είχαν, όσο το δυνατό, μεγαλύτερη διαφορά στη δραστηριότητα τους. Έτσι, προσπαθήσαμε να βασίσουμε τις τιμές που καθορίσαμε για τα όρια, στις τιμές που αντιπροσώπευαν οι διαφορές των δεδομένων, τον πέντε αυτών χρηστών.

Οι τιμές των ορίων για το κάθε δεδομένο μπορούν να παρατηρηθούν στο Παράρτημα Β.

Αφού ολοκληρωθούν οι συγκρίσεις του πρώτου διανύσματος με τα όρια τους και δημιουργηθεί επιτυχώς το δεύτερο διάνυσμα, υπολογίζεται στη συνέχεια το μοντέλο προσωπικότητας του “Big Five”, βασισμένο στα δεδομένα του Facebook που συλλέξαμε και επεξεργαστήκαμε. Για να επιτύχουμε την δημιουργία του μοντέλου προσωπικότητας του χρήστη αναθέσαμε σε κάθε μια από τις πέντε διαστάσεις προσωπικότητας του “Big Five” σχετικά δεδομένα του Facebook που συλλέξαμε για το χρήστη. Επίσης για κάθε στοιχείο σε κάποια διάσταση, αναθέταμε αν ήταν θετικής ή αρνητικής σημασίας για τη διάσταση αυτή καθώς και τη βαρύτητα το οποίο είχε το συγκεκριμένο δεδομένο για την συγκεκριμένη διάσταση. Αξίζει εδώ να αναφέρουμε ότι, κάποια από τα δεδομένα χαρακτήριζαν παραπάνω από μία διαστάσεις. Ωστόσο, δεν ήταν αναγκαίο να τις καθορίζουν με το ίδιο βάρος σημασίας ή να είχαν την ίδια θετική ή αρνητική σημασία για όλες τις διαστάσεις. Αντίθετος παρατηρήθηκε ότι συνήθως για κάθε διάσταση την οποία καθόριζε κάποιο δεδομένο, είχε διαφορετικές τιμές σημασίας σε αυτές.

Εδώ πρέπει να ειπωθεί ότι, οι κατατάξεις των δεδομένων, σε θετικής ή αρνητικής σημασίας, για τις διαστάσεις της προσωπικότητας (Πίνακας 4-1), βασίστηκε στις συναφείς εργασίες τις οποίες μελετήσαμε στο Κεφάλαιο 2, καθώς επίσης και στα δεδομένα των πέντε φοιτητών, τα

οποία χρησιμοποιήσαμε στην προμελέτη της εργασίας μας, και των οποίων γνωρίζαμε τα μοντέλα προσωπικότητας.

Σημαντικό είναι να εξηγήσουμε τι εννοούμε με τους όρους θετικής (+) και αρνητικής (-) σημασίας για κάποια διάσταση. Όταν κάποιο δεδομένο είναι θετικής σημασίας για μια διάσταση, αυτό υπονοεί ότι όσο μεγαλύτερη είναι η τιμή του δεδομένου τόσο το καλύτερο, δηλαδή, τόσο περισσότερη θα είναι η βαθμολογία του σε αυτή τη διάσταση. Για παράδειγμα: στην διάσταση της Εξωστρέφειας, όσο μεγαλύτερος ο συνολικός αριθμός φίλων που έχει ένας χρήστης, τόσο πιο μεγάλη σημασία θα αποκτά η διάσταση αυτή προς το μοντέλο προσωπικότητας του χρήστη. Έτσι, αφού οι τιμές που χρησιμοποιούμε κυμαίνονται από το ένα μέχρι το πέντε, όσο πιο κοντά στο πέντε είναι η τιμή τόσο το καλύτερο. Αντιθέτως, όταν αναφέρουμε ότι κάποιο δεδομένο είναι αρνητικής σημασίας για μια διάσταση, τότε όσο μεγαλύτερη είναι η τιμή που έχει τόσο λιγότερο θα είναι το σκορ του για τη συγκεκριμένη διάσταση.

Διαχωρισμός Δεδομένων στις Διαστάσεις της Προσωπικότητας		
Γνωρίσματα	+	-
Εξωστρέφεια (Extraversion)	• Συνολικός Αριθμός Φίλων • Συνολικός Αριθμός Καταστάσεων που Δημοσίευσε • Συνολικός Αριθμός των likes που πραγματοποίησε • Συνολικός Αριθμός των Check-ins που δημιούργησε • Συνολικός Αριθμός των Εκδηλώσεων που δημιούργησε • Συνολικός Αριθμός των κοινοποιήσεων στις οποίες έκανε tag τον εαυτό του (self-tagging) • Ποσοστό των εικόνων προφίλ του σε σχέση με το σύνολο των φωτογραφιών του	
Τερπνότητα (Agreeableness)	• Συνολικός Αριθμός των κοινοποιήσεων στις οποίες είναι tagged • Συνολικός Αριθμός των	• Συνολικός Αριθμός των likes που πραγματοποίησε • Συνολικός Αριθμός

	Συνδέσμων που μοιράστηκε στο Facebook	των Ενεργών του Φίλων • Συνολικός Αριθμός των Check-ins στα οποία συμμετέχει
Ευσυνειδησία (Conscientiousness)	• Ποσοστό του Αγαπημένου Τύπου Σελίδας του Χρήστη σε Σχέση με τον Συνολικό Αριθμό Σελίδων που έκανε like	• Συνολικός Αριθμός των Ενεργών του Φίλων • Συνολικός Αριθμός Καταστάσεων που Δημοσίευσε • Συνολικός Αριθμός των likes που πραγματοποίησε • Συνολικός Αριθμός των Εκδηλώσεων που παραβρέθηκε • Συνολικός Αριθμός των κοινοποιήσεων στις οποίες είναι tagged • Συνολικός Αριθμός των Συνδέσμων που μοιράστηκε στο Facebook
Νεύρωση (Neuroticism)	• Συνολικός Αριθμός των likes που πραγματοποίησε • Συνολικός Αριθμός Σελίδων στο Facebook που έκανε like	• Συνολικός Αριθμός Φίλων • Ποσοστό των εικόνων προφίλ του σε σχέση με το σύνολο των φωτογραφιών του
Διαφάνεια (Openness to experience)	• Συνολικός Αριθμός Καταστάσεων που Δημοσίευσε • Συνολικός Αριθμός των likes που πραγματοποίησε • Συνολικός Αριθμός των Εκδηλώσεων που παραβρέθηκε • Συνολικός Αριθμός Σελίδων στο Facebook που έκανε like	• Ποσοστό του Αγαπημένου Τύπου Σελίδας του Χρήστη σε Σχέση με τον Συνολικό Αριθμό Σελίδων που έκανε like

Πίνακας 4-1 Διαχωρισμός Δεδομένων στις Διαστάσεις της Προσωπικότητας

Οι τιμές σημασίας και τα βάρη των δεδομένων για κάθε διάσταση, μπορούν να παρατηρηθούν στο Παράρτημα Δ.

Εξαγωγή Όμοιων Φίλων (Deriving Similar Friends)

Στο δεύτερο διάνυσμα το οποίο δημιουργούμε, βασίζεται ακόμη ένα μεγάλο κομμάτι της παρούσας διπλωματικής εργασίας, η σύγκριση και εύρεση όμοιων χρηστών του συστήματος μας για παρουσίαση στον τρέχων χρήστη. Για την εύρεση ομοιότητας μεταξύ των χρηστών μας, χρησιμοποιούμε την Συνημιτονοειδή Ομοιότητα (Cosine Similarity).

Συνημιτονοειδής Ομοιότητα: Είναι ένα μέτρο ομοιότητας μεταξύ δύο διανυσμάτων ενός εσωτερικού χώρου (ενός προϊόντος), που μετρά το συνημίτονο της γωνίας μεταξύ των δύο συγκεκριμένων διανυσμάτων. Το συνημίτονο των 0° είναι 1, και για οποιαδήποτε άλλη γωνία είναι λιγότερο από 1. Πρόκειται για μια κρίση προσανατολισμού παρά μεγέθους, δηλαδή, δύο διανύσματα με τον ίδιο προσανατολισμό έχουν Συνημιτονοειδή Ομοιότητα 1, δύο διανύσματα στους 90° έχουν μια ομοιότητα 0, και δύο διανύσματα διαμετρικά αντίθετα, έχουν ομοιότητα -1, ανεξαρτήτως του μεγέθους τους. Η Συνημιτονοειδής Ομοιότητα χρησιμοποιείται ιδιαίτερα σε θετικό χώρο, δηλαδή για διανύσματα χωρίς αρνητικές τιμές, όπου τα αποτελέσματα είναι τακτοποιημένα και οριοθετούνται στο $[0,1]$. Η τεχνική αυτή χρησιμοποιείται για τη μέτρηση της συνοχής εντός των ομάδων (clusters) στον τομέα της εξόρυξης δεδομένων (Data Mining).

$$similarity = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}}$$

Η Συνημιτονοειδής Ομοιότητα δύο διανυσμάτων A και B είναι ο λόγος του αθροίσματος των γινομένων των αντίστοιχων στοιχείων των δύο διανυσμάτων με το γινόμενο των ριζών των αθροισμάτων των στοιχείων στο τετράγωνο των δύο διανυσμάτων.

Έχοντας σαν βάση το διάνυσμα του χρήστη, εκτελούμε ένα ερώτημα στην δική μας βάση δεδομένων, το οποίο θα μας επιστρέψει τα διανύσματα των υπόλοιπων προϋπαρχόντων χρηστών του συστήματος. Ακολουθώντας, για κάθε ένα από αυτά τα διανύσματα, υπολογίζουμε την Συνημιτονοειδή Ομοιότητα τους με το διάνυσμα του τρέχων χρήστη μας. Αν η τιμή της ομοιότητας είναι μεγαλύτερη του 0.8, τότε προσθέτουμε τον χρήστη εκείνο σε ένα δυναμικό πίνακα. Ας αναφέρουμε εδώ ότι, η τιμή 0.8 επιλέχθηκε μετά από μελέτη δοκιμών έτσι ώστε να

επιλεγεί μία τιμή τέτοια που να καθορίζει ότι, όντως οι χρήστες μας είναι αρκετά όμοιοι, αλλά επίσης να μπορεί να δώσει αποτελέσματα στα πρώτα στάδια της χρήσης της εφαρμογής μας, όπου οι χρήστες που θα την χρησιμοποιούσαν θα ήταν σημαντικά λίγοι (cold start problem of recommender systems⁶).

Αφού ολοκληρωθεί η σύγκριση μεταξύ του τρέχων χρήστη μας, με όλους τους υπόλοιπους χρήστες του συστήματος και συγκεντρωθούν στον δυναμικό πίνακα που αναφέραμε προηγουμένως, τότε διαλέγονται από αυτό το πίνακα πέντε τυχαίοι χρήστες και τους παρουσιάζουμε με την τρέχων εικόνα προφίλ που έχουν και με το όνομα τους. Αξίζει εδώ να αναφέρουμε ότι, εκτελούμε για κάθε μια από τις φωτογραφίες προφίλ των χρηστών αυτών ένα ερώτημα στη βάση δεδομένων του Facebook για να ανακτήσουμε προσωρινά μια μεγαλύτερου μεγέθους φωτογραφία από αυτή που φυλάξαμε κατά την συλλογή των δεδομένων. Κάνοντας κλικ στην εικόνα ή στο όνομα τους, θα ανοίξει σε διπλανό παράθυρο στον φυλλομετρητή του τρέχων χρήστη το προφίλ του χρήστη που επέλεξε. Σε περίπτωση που δεν υπάρχουν όμοιοι χρήστες στο σύστημα με τον τρέχων χρήστη, του παρουσιάζουμε ένα μήνυμα να τον ενημερώσουμε για το γεγονός αυτό και επίσης τον προτρέπουμε να μοιραστεί αυτή του την εμπειρία με τους φίλους του στο Facebook για να χρησιμοποιήσουν το σύστημα μας και τότε ίσως να αλλάξει το αποτέλεσμα αυτό.

Συστάσεις ‘Ενδιαφέρον’ σελίδων στο Facebook (Recommendation of Interesting Facebook Pages)

Φτάνοντας αισίως σε αυτό το σημείο της εκτέλεσης, έχουμε στη διάθεση μας όλα τα δεδομένα του χρήστη, το μοντέλο του και ένα δυναμικό πίνακα ο οποίος περιέχει μια λίστα με όμοιους χρήστες με τον τρέχων χρήστη μας. Χρησιμοποιώντας τα δεδομένα τα οποία συλλέξαμε για τις σελίδες του Facebook, που αρέσουν στον χρήστη μας, σε συνδυασμό με τους όμοιους του χρήστες, θα παρουσιάσουμε στον χρήστη μια λίστα με σελίδες, του αγαπημένου του τύπου σελίδας, τις οποίες δεν έχει κάνει ακόμη like.

⁶ Cold Start Problem: http://en.wikipedia.org/wiki/Cold_start

Η επεξεργασία η οποία εκτελείτε για την επίτευξη αυτής της ενέργειας, αρχίζει με την εξόρυξη από τη βάση δεδομένων μας, του αγαπημένου τύπου σελίδας (με βάση του σε ποιο τύπο σελίδας κάνει like πιο πολύ) του Facebook για το χρήστη μας. Αφού αποκτήσουμε αυτή την πληροφορία την χρησιμοποιούμε για το επόμενο ερώτημα το οποίο θα μας επιστρέψει τους μοναδικούς κωδικούς των σελίδων που έχει ήδη κάνει like. Το αποτέλεσμα αυτού του ερωτήματος θα αποθηκευτεί σε ένα δυναμικό πίνακα της γλώσσας προγραμματισμού PHP. Αφού ολοκληρωθούν αυτά τα ερωτήματα, μαζεύουμε της ταυτότητες των χρηστών οι οποίοι είναι όμοιοι με τον τρέχων χρήστη μας, τους οποίους εντοπίσαμε και αποθηκεύσαμε σε προηγούμενο στάδιο της επεξεργασίας μας. Χρησιμοποιώντας αυτές τις δύο πληροφορίες, δημιουργούμε ένα ερώτημα στη βάση δεδομένων μας το οποίο θα μας επιστρέψει τους κωδικούς των σελίδων, του αγαπημένου τύπου σελίδας, που έχουν κάνει like όλοι οι όμοιοι χρήστες του τρέχων χρήστη μας. Το αποτέλεσμα αυτού του ερωτήματος θα αποθηκευτεί παρομοίως σε ένα ξεχωριστό δυναμικό πίνακα.

Για κάθε ένα από τα στοιχεία που βρίσκονται στο πίνακα με τις σελίδες των όμοιων χρηστών, ελέγχεται εάν υπάρχει στο πίνακα με τις σελίδες που ήδη αρέσουν στο χρήστη. Εάν υπάρχει στον πίνακα, τότε συνεχίζουμε την εκτέλεση με την επόμενη σελίδα. Εάν όμως δεν υπάρχει, τότε παρουσιάζεται στον χρήστη με την εικόνα προφίλ και το όνομα της. Η εκτέλεση αυτού του βρόγχου θα συνεχιστεί είτε μέχρι ελεγχτούν όλες οι σελίδες των όμοιων χρηστών, είτε μέχρι να παρουσιαστούν πέντε διαφορετικές σελίδες στο χρήστη. Να αναφέρουμε εδώ ότι εάν ο τρέχων χρήστης μας αποφασίσει να κάνει κλικ σε μια από τις σελίδες που του παρουσιάζουμε, τότε θα ανοίξει άλλο παράθυρο στο φυλλομετρητή του με την σελίδα στο Facebook. Επίσης, τα δεδομένα για τις σελίδες που του παρουσιάζουμε, συγκεντρώνονται μέσω νέων ερωτημάτων στη βάση δεδομένων του Facebook.

Συγκεντρωτική Παρουσίαση Στοιχείων Δραστηριότητας στο Facebook

Τέλος θα αναφέρουμε για το αρχικό κομμάτι δεδομένων που παρουσιάζουμε στον χρήστη, το οποίο έχει σχέση με την συγκεντρωτική ανάλυση των δεδομένων που ανακτήσαμε στην αρχή από το προφίλ του και όχι με το μοντέλο του. Το κομμάτι αυτό περιέχει δύο γραφικές παραστάσεις καθώς και τρεις ομάδες στοιχείων.

Η πρώτη γραφική παράσταση παρουσιάζει τα ποσοστά των likes του χρήστη, σε κάθε τύπο αντικειμένου σε σχέση με τον συνολικό αριθμό των likes του που κρατούνται αποθηκευμένα στη βάση δεδομένων μας. Να αναφέρουμε εδώ ότι αν δεν είναι η πρώτη φορά που επισκέπτεται ο χρήστης την εφαρμογή μας ο συνολικός αριθμός των likes του μπορεί να είναι μεγαλύτερος του 1000 το οποίο είναι και ο μέγιστος αριθμός likes, που μπορεί να μας επιστρέψει η βάση δεδομένων του Facebook. Αυτό συμβαίνει, λόγω της ενημέρωσης που εκτελείτε στη βάση δεδομένων μας για τον χρήστη κάθε φορά που εισέρχεται στην εφαρμογή μας. Η δεύτερη γραφική παράσταση παρουσιάζει τα ποσοστά του κάθε τύπου σελίδας που έχει κάνει like ο χρήστης σε σχέση με τον συνολικό αριθμό σελίδων που έκανε like.

Οι τρεις ομάδες στοιχείων που παρουσιάζουμε στο χρήστη είναι: (i) τον ιδιοκτήτη των περισσότερων αντικειμένων των οποίων έκανε like ο χρήστης, (ii) τον ιδιοκτήτη των περισσότερων αντικειμένων στα οποία είναι tagged ο χρήστης και (iii) τον συνολικό αριθμό εκδηλώσεων στις οποίες παραβρέθηκε καθώς και τις τελευταίες 3 στις οποίες παραβρέθηκε. Τα δεδομένα αυτά υπάρχουν ήδη αποθηκευμένα στη βάση δεδομένων μας και με την εκτέλεση των κατάλληλων ερωτημάτων παίρνουμε αυτές τις πληροφορίες και τις παρουσιάζουμε στο χρήστη. Η παρουσίαση αυτών των πληροφοριών γίνεται με την εικόνα προφίλ των δύο ατόμων, τις εικόνες που ανέθεσε στις εκδηλώσεις ο δημιουργός τους, καθώς επίσης και με τα ονόματά τους.

4.6 Επίλογος

Στο κεφάλαιο αυτό παρουσιάστηκαν οι απαιτήσεις του συστήματος καθώς και οι προδιαγραφές τόσο από την πλευρά του χρήστη όσο και από την πλευρά του εξυπηρετητή. Ακολούθως ορίστηκαν οι προαπαιτήσεις του συστήματος και παρουσιάστηκε αναλυτική περιγραφή των πιο σημαντικών λεπτομερειών υλοποίησης. Το επόμενο κεφάλαιο παρουσιάζει την δημιουργία της εφαρμογής PersonaWeb στο Facebook.

Κεφάλαιο 5

Εφαρμογή PersonaWeb στο Facebook

5.1 Εισαγωγή.....	57
5.2 Δημιουργία Εφαρμογής	57
5.3 Παρουσίαση Λειτουργιών και Πληροφοριών στο Χρήστη	60
5.4 Εξωτερικές Λειτουργίες Εφαρμογής	70
5.5 Επίλογος	72

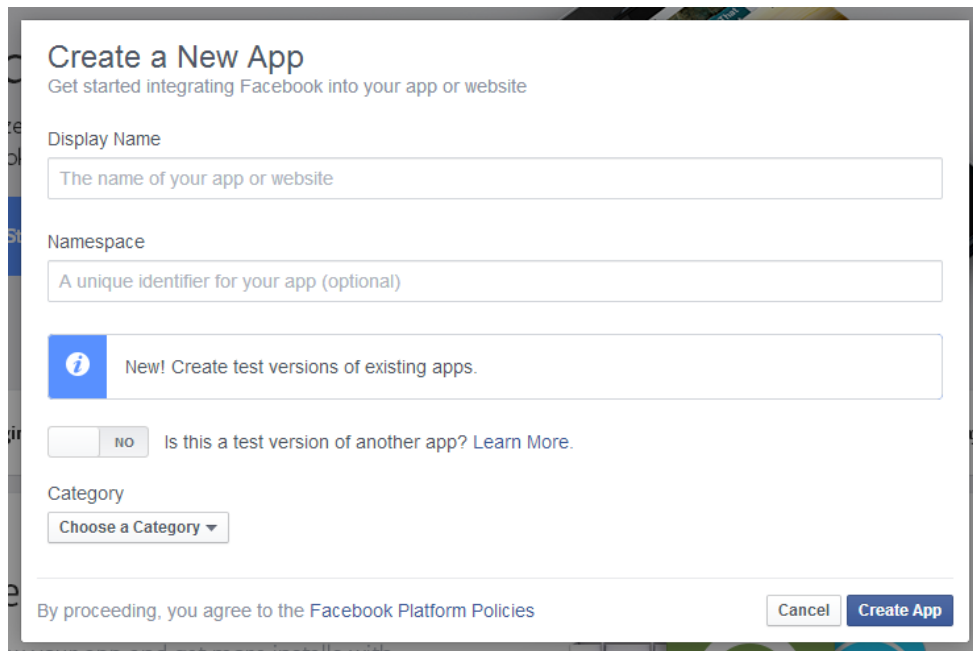
5.1 Εισαγωγή

Σε αυτό το κεφάλαιο θα παρουσιάσουμε τα βασικά βήματα της ανάπτυξης μια εφαρμογής στο Facebook, καθώς επίσης όλες τις λειτουργίες τις οποίες διαθέταμε στους χρήστες μας για να αλληλεπιδράσουν με την εφαρμογή μας.

5.2 Δημιουργία Εφαρμογής

Αρχικά για να μπορούμε να επικοινωνήσουμε με τη βάση δεδομένων του Facebook, χρειαζόμασταν να εισάκουμε την εφαρμογή μας στο κοινωνικό δίκτυο Facebook έτσι ώστε να χρησιμοποιήσουμε έτοιμες λειτουργίες που παρέχει στους προγραμματιστές του. Οι λειτουργίες αυτές ήταν απαραίτητες για την ανάκτηση δεδομένων από τη βάση δεδομένων του Facebook καθώς επίσης και για τη δημιουργία συμβολαίου με το χρήστη για να τον ενημερώσουμε ποια δεδομένα του επιθυμούμε να εξάγουμε και ακολούθως να μας επιτρέψει πρόσβαση σε αυτά.

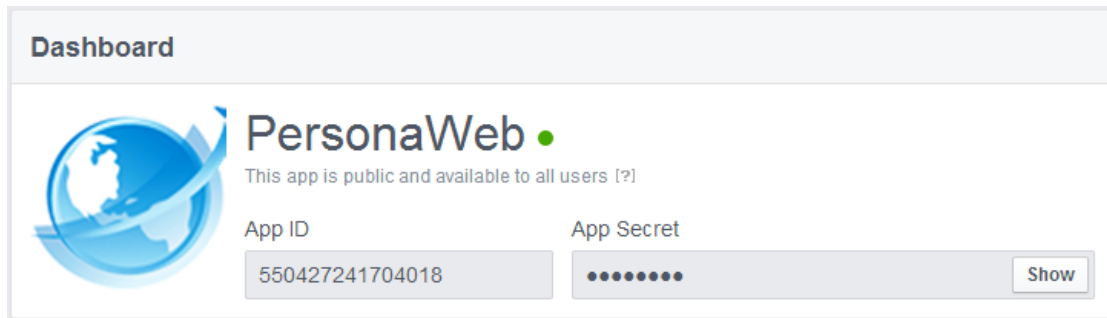
Για την εισαγωγή μιας εφαρμογής στην ειδική πλατφόρμα που παρέχει το Facebook, χρειάζονται κάποια βασικά βήματα από τους δημιουργούς της εφαρμογής. Αρχικά ο προγραμματιστής πρέπει να δώσει στην εφαρμογή ένα μοναδικό όνομα το οποίο δεν έχει καμία άλλη εφαρμογή που υπάρχει στο Facebook. Κατά την εκτέλεση της ίδιας φάσης, ο προγραμματιστής μπορεί προαιρετικά να εισάγει και ένα μοναδικό διαδικτυακό όνομα χώρου (internet namespace), το οποίο χρησιμοποιείτε για την δημιουργία ενός μοναδικού συνδέσμου στο διαδίκτυο για την εφαρμογή μέσα στο Facebook. Ο σύνδεσμος αυτός θα είναι της μορφής <https://apps.facebook.com/<namespace>>. Επιπρόσθετα ο προγραμματιστής μπορεί να καθορίσει στο ίδιο σημείο τον τύπο της εφαρμογής από μία προκαθορισμένη λίστα τύπων που παρέχει το Facebook. Μια εφαρμογή μπορεί να είναι παιχνίδι, εκπαιδευτική κ.τ.λ. Η φάση αυτή φαίνεται στην Εικόνα 5-1.



Εικόνα 5-1 Δημιουργία νέας εφαρμογής στο Facebook

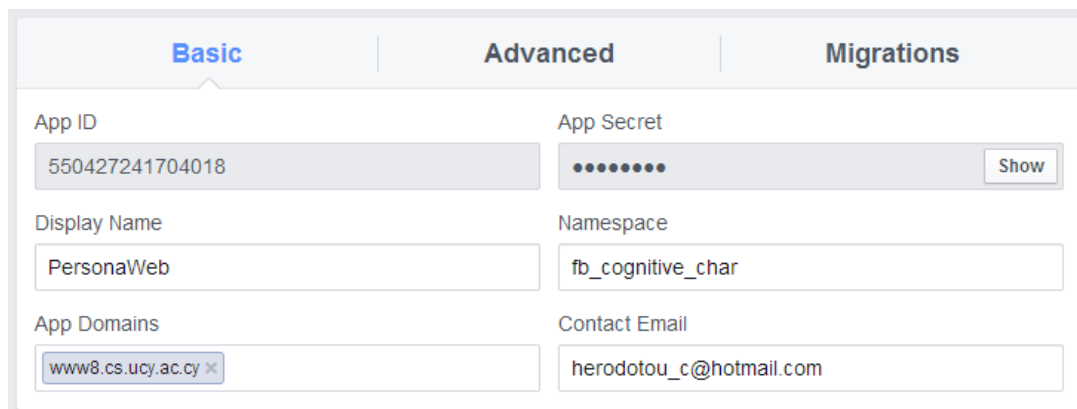
Με το πέρας αυτής της πρώτης φάσης το Facebook δημιουργεί δύο μοναδικούς αριθμούς οι οποίοι καθορίζουν την εφαρμογή. Ο ένας είναι ο μοναδικός κωδικός της εφαρμογής και ο άλλος είναι το μυστικό της εφαρμογής, το οποίο χρησιμοποιείται για την επικοινωνία της εφαρμογής με τη βάση δεδομένων. Το μυστικό αυτό γίνεται γνωστό μόνο στους δημιουργούς της εφαρμογής και υπάρχει κίνδυνος κλοπής δικαιωμάτων αν αυτός ο κωδικός κλαπεί από κάποιο

άλλο άτομο. Στην Εικόνα 5-2 παρουσιάζεται ο τρόπος με τον οποίο δίνει το Facebook αυτές τις πληροφορίες στους προγραμματιστές της εφαρμογής.



Εικόνα 5-2 Ανάκτηση Μοναδικών Κωδικών Εφαρμογής

Ακολουθώντας, για να μπορέσει η εφαρμογή στο Facebook να αντικατοπτρίζει την εφαρμογή που δημιουργείται, πρέπει να δοθούν οι σύνδεσμοι του βασικού τομέα που φιλοξενείται η εφαρμογή καθώς επίσης και ο σύνδεσμος ο οποίος θα ανοίγει μέσα στο Facebook ως αρχική σελίδα της εφαρμογής. Τα δεδομένα και οι επιλογές αυτές που πρέπει να μελετήσει και να καθορίσει ο προγραμματιστής στη σελίδα ελέγχου της εφαρμογής του, φαίνονται στις Εικόνα 5-3 και Εικόνα 5-4.



Εικόνα 5-3 Συμπλήρωση στοιχείων Εφαρμογής – Α Μέρος

App on Facebook [X]

Canvas Page

☐ NO Unity Integration
 "Yes" to use the Facebook Unity SDK

Canvas URL

Secure Canvas URL

☐ NO Canvas Fixed Width
 "Yes" sets canvas width to 760 px.

☐ NO Canvas Fixed Height
 "Yes" allows setting fixed height.

Εικόνα 5-4 Συμπλήρωση στοιχείων Εφαρμογής - Β Μερους

5.3 Παρουσίαση Λειτουργιών και Πληροφοριών στο Χρήστη

Η εφαρμογή μας είναι κομμάτι του κέντρου εφαρμογών του Facebook και χρησιμοποιεί βασικές συναρτήσεις τις οποίες παραθέτει το Facebook στους προγραμματιστές του για να υλοποιήσουν στις εφαρμογές του. Χρησιμοποιώντας συγκεκριμένη συνάρτηση του Facebook μπορούμε να εντοπίσουμε αν ο τρέχων χρήστης είναι συνδεδεμένος με το Facebook. Αν όχι τότε πριν ξεκινήσει η εφαρμογή ζητούμε από τον χρήστη να συνδεθεί με το Facebook παρουσιάζοντας τη σχετική σελίδα σύνδεσης του Facebook για καταχώριση των διακριτικών του χρήστη, όπως φαίνεται στην Εικόνα 5-5. Ακολούθως, αν είναι η πρώτη φορά που ο χρήστης θα χρησιμοποιήσει την εφαρμογή μας, τότε εμφανίζεται ένας διάλογος παρουσιάζοντας το συμβόλαιο που πρέπει να αποδεχτεί ο χρήστης για να μπορέσει να προχωρήσει με την εφαρμογή, Εικόνα 5-6. Εδώ πρέπει να αναφερθεί ότι, εάν ο χρήστης αρνηθεί να μας δώσει τα δικαιώματα να ανακτήσουμε όλα τα δεδομένα τα οποία ζητάμε από αυτόν, τότε δεν μπορεί να προχωρήσει στην εκκίνηση και είσοδο της εφαρμογής μας.

Facebook Login

Email or Phone:

Password:

☒ Keep me logged in

[Log In](#) or [Sign up for Facebook](#)

[Forgot your password?](#)

[English \(US\)](#)
[Ελληνικά](#)
[Español](#)
[Português \(Brasil\)](#)
[Français \(France\)](#)
[Deutsch](#)
[Italiano](#)
[العربية](#)
[हिन्दी](#)
[中文\(简体\)](#)
[...](#)

Εικόνα 5-5 Παράθυρο ταυτοποίησης του χρήστη με το λογαριασμό του στο Facebook




PersonaWeb will receive the following info:
 your public profile, friend list, email address,
 custom friends lists, News Feed, birthday,
 status updates, check-ins, events, Photos,
 videos and likes and your friends' status
 updates, check-ins, events, photos and videos.

 This does not let the app post to Facebook.

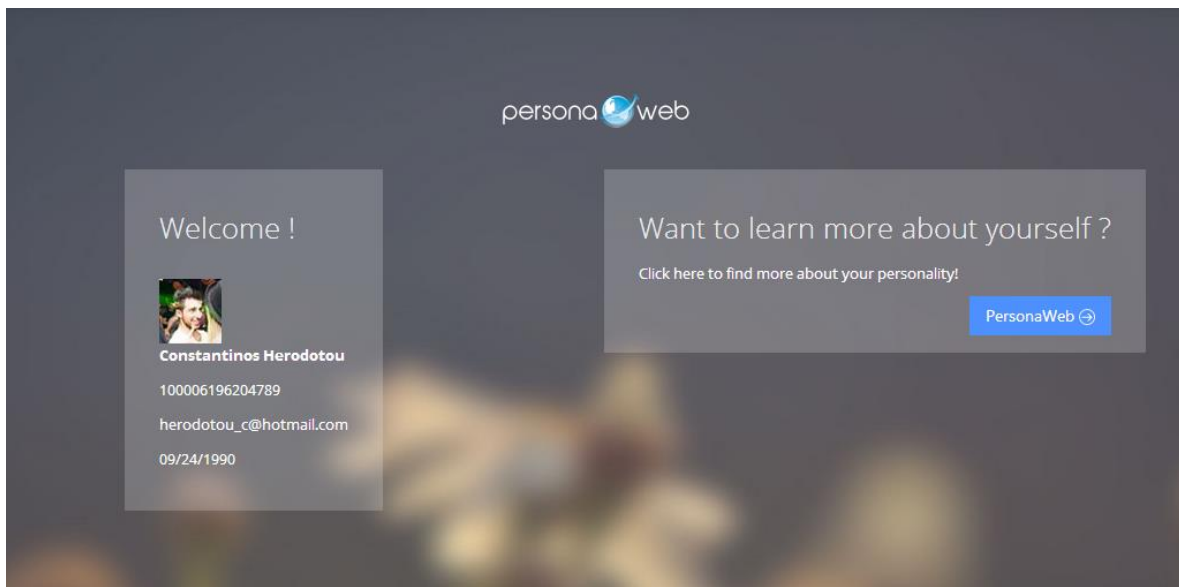
[Privacy policy](#)
[Cancel](#)
[OK](#)

Εικόνα 5-6 Συμβόλαιο Εφαρμογής - Χρήστη

Στη συνέχεια, με την ολοκλήρωση του κώδικα ανάκτησης των πληροφοριών του χρήστη, εμφανίζουμε σε αυτόν μια φόρμα στην πάνω αριστερή γωνία, η οποία περιέχει κάποια από τα βασικά του δεδομένα Εικόνα 5-8. Ένα αναδυόμενο παράθυρο, όπως φαίνεται στην Εικόνα 5-7, θα του προτείνει να μοιραστεί το αποτέλεσμα της προσωπικότητας του με τους φίλους του στο Facebook μέσω μιας κοινοποίησης. Στην πάνω δεξιά γωνία (Εικόνα 5-8), υπάρχει μια άλλη φόρμα η οποία περιέχει ένα κουμπί το οποίο εάν κάνει κλικ σε αυτό, ο χρήστης, τον ανακατευθύνει στην ιστοσελίδα του PersonaWeb όπου εκεί μπορεί να χρησιμοποιήσει όλα τα διαθέσιμα τεστ και ερωτηματολόγια για να ενημερωθεί για την προσωπικότητα του και τα γνωστικά του χαρακτηριστικά.

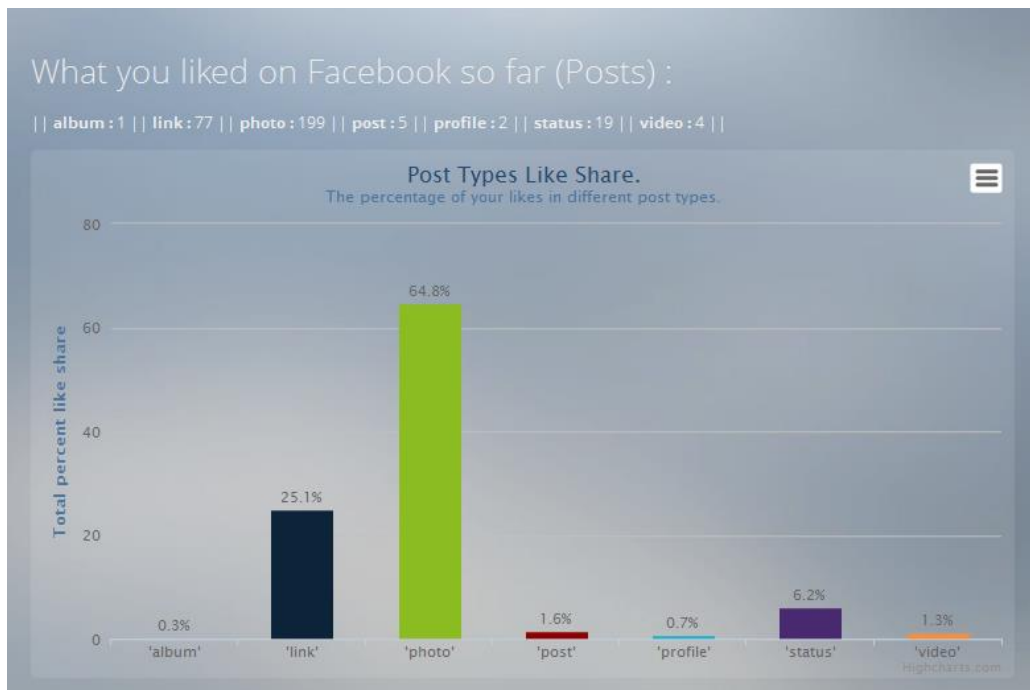


Εικόνα 5-7 Αναδυόμενο παράθυρο όπου ο χρήστης μπορεί να κοινοποιήσει τα αποτελέσματα του



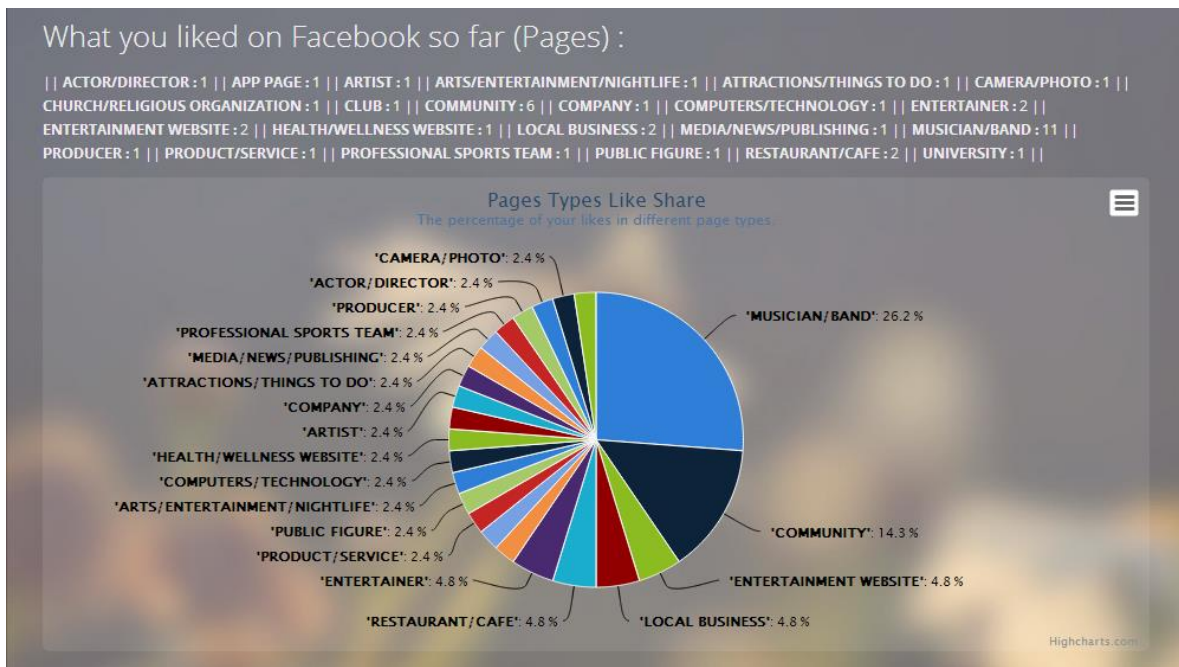
Εικόνα 5-8 Αριστερά: Φόρμα παρουσίασης προσωπικών στοιχείων χρήστη. Δεξιά: Φόρμα όπου ο χρήστης μπορεί να μεταβεί στην ιστοσελίδα του PersonaWeb

Προχωρώντας πιο κάτω στην εφαρμογή μας, ο χρήστης θα παρατηρήσει μία γραφική παράσταση Εικόνα 5-9, η οποία του παρουσιάζει συγκεντρωτικές πληροφορίες σε σχέση με τα likes του σε αντικείμενα στο Facebook.

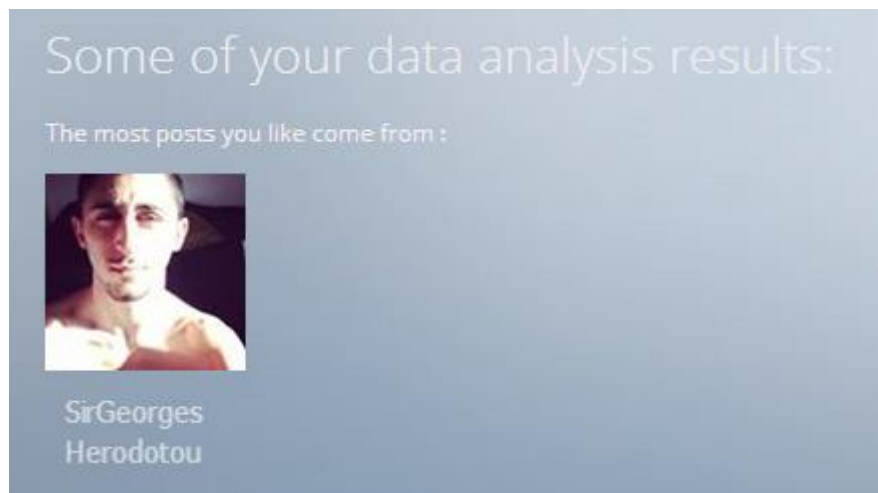


Εικόνα 5-9 Συγκεντρωτικές πληροφορίες σε σχέση με τα likes του χρήστη σε αντικείμενα στο Facebook

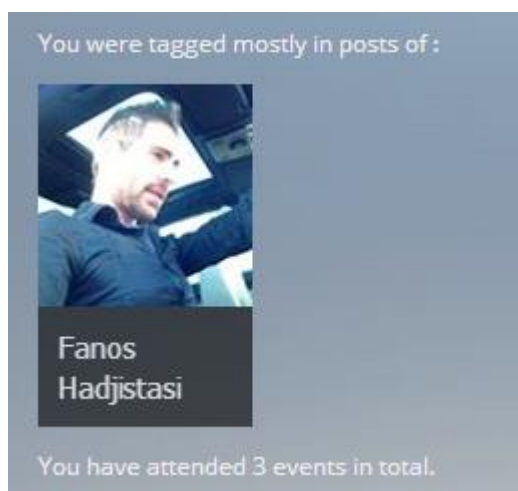
Πιο κάτω υπάρχει ακόμα μια γραφική παράσταση η οποία του παρουσιάζει συγκεντρωτικές πληροφορίες για τις σελίδες του Facebook σε σχέση με τον τύπο τους, τις οποίες έκανε like (Εικόνα 5-10). Συνεχίζοντας θα δει κάποια από τα στοιχεία του που συλλέξαμε για τους φίλους του και για τις εκδηλώσεις στις οποίες παραβρέθηκε Εικόνα 5-13. Τα στοιχεία για τους φίλους του σε σχέση με τα δεδομένα του που μπορεί να δει είναι τον ιδιοκτήτη των περισσότερων αντικειμένων των οποίων έκανε like (Εικόνα 5-11) καθώς επίσης και τον ιδιοκτήτη των περισσότερων αντικειμένων στα οποία είναι tagged ο ίδιος Εικόνα 5-12.



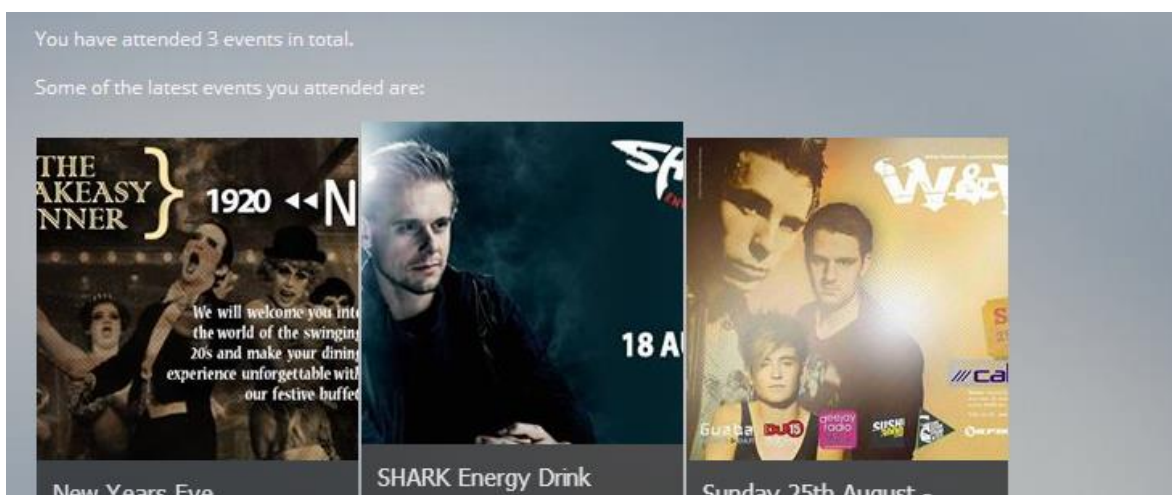
Εικόνα 5-10 Συγκεντρωτικές πληροφορίες σε σχέση με τις σελίδες στο Facebook και ανάλογα με τον τύπο τους, τις οποίες ο χρήστης έκανε like



Εικόνα 5-11 Ο χρήστης μπορεί να δει πληροφορίες σε σχέση με τον ιδιοκτήτη των περισσότερων αντικειμένων των οποίων έκανε like

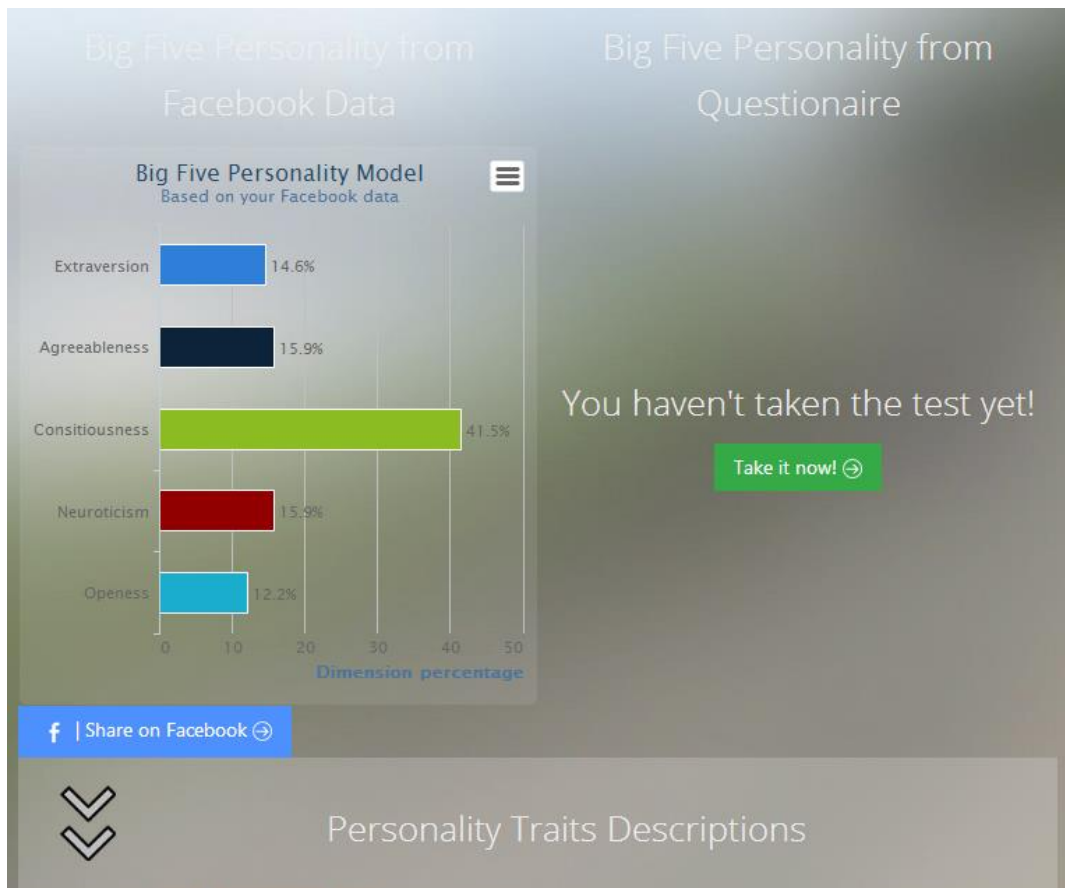


Εικόνα 5-12 Ο χρήστης μπορεί να δει τον ιδιοκτήτη των περισσότερων αντικειμένων στα οποία είναι tagged ο ίδιος



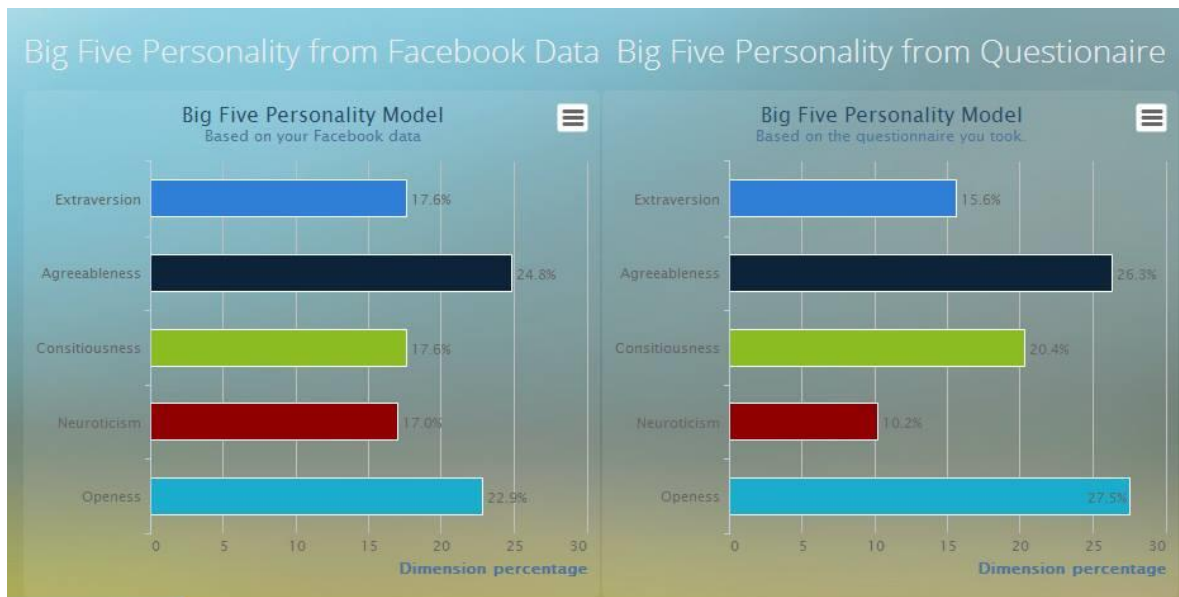
Εικόνα 5-13 Πληροφορίες σχετικά με τις εκδηλώσεις που παρευρέθηκε ο χρήστης

Το επόμενο κομμάτι με πληροφορίες και επιλογές που παρουσιάζεται του χρήστη, παράγεται βάσει του μοντέλου προσωπικότητας που δημιουργείται για αυτόν, βασισμένο στην δραστηριότητα και τα δεδομένα του, στο Facebook. Πρωτεύον και βασικό στοιχείο που του παρουσιάζουμε είναι το ίδιο το μοντέλο της προσωπικότητας του που δημιουργήσαμε, ως μια γραφική παράσταση η οποία περιέχει το ποσοστό που σκόραρε ο χρήστης σε κάθε μια από τις πέντε διαστάσεις της προσωπικότητας. Αυτό φαίνεται στην Εικόνα 5-14.



Εικόνα 5-14 Μοντέλο προσωπικότητας χρήστη που δημιουργείται βάσει την δραστηριότητα και τα δεδομένα του στο Facebook

Σε αυτό το σημείο, αυτό που παρουσιάζεται στον χρήστη διαφοροποιείται από το αν έχει συμπληρώσει το ερωτηματολόγιο της προσωπικότητας. Η διαφοροποίηση είναι η εξής: αν ο χρήστης έχει ήδη αποτελέσματα της προσωπικότητας του και από το ερωτηματολόγιο, τότε δίπλα από την γραφική παράσταση του μοντέλου προσωπικότητας του, μπορεί να δει ακόμη μία γραφική παράσταση, του ίδιου ακριβώς τύπου, με τα αποτελέσματα του ερωτηματολογίου και μπορεί έτσι να συγκρίνει και ο ίδιος την διαφορά των δύο μοντέλων του (Εικόνα 5-15).

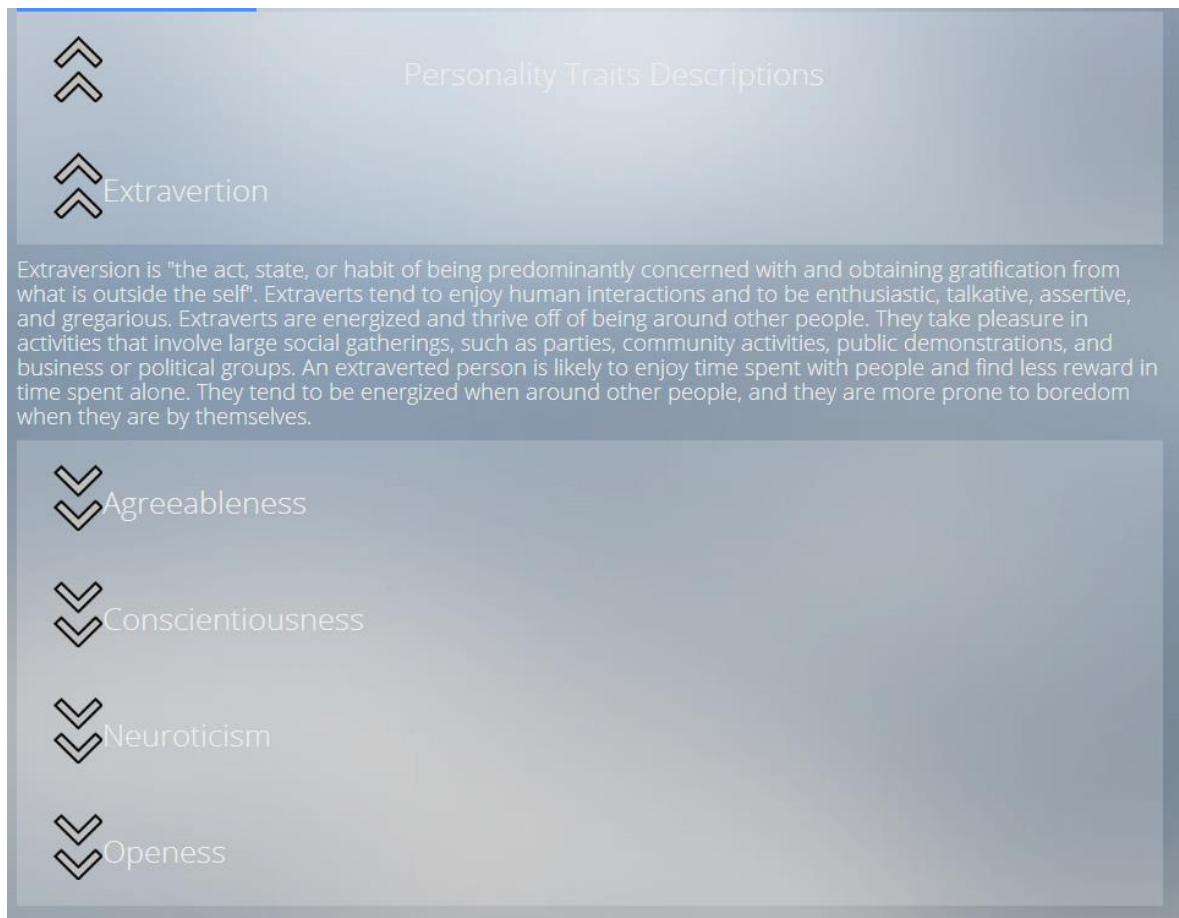


Εικόνα 5-15 Συγκριτικά αποτελέσματα απο το μοντέλο προσωπικότητας χρήστη και των αποτελεσμάτων απο το ερωτηματολόγιο "Big Five"

Εάν όμως, δεν έχει συμπληρώσει το ερωτηματολόγιο ακόμη, τότε δίπλα από την γραφική του μοντέλου προσωπικότητας του, θα παρατηρήσει ένα κουμπί το οποίο, όταν ο χρήστης κάνει κλικ σε αυτό θα εμφανιστεί μπροστά από την εφαρμογή ένα υπόδειγμα από το οποίο μπορεί να διαλέξει ποια εκδοχή του ερωτηματολογίου προσωπικότητας θέλει να εκτελέσει μέσω του PersonaWeb (Εικόνα 5-14). Οι επιλογές που έχει είναι η σύντομη και η εκτεταμένη εκδοχή του ερωτηματολογίου.

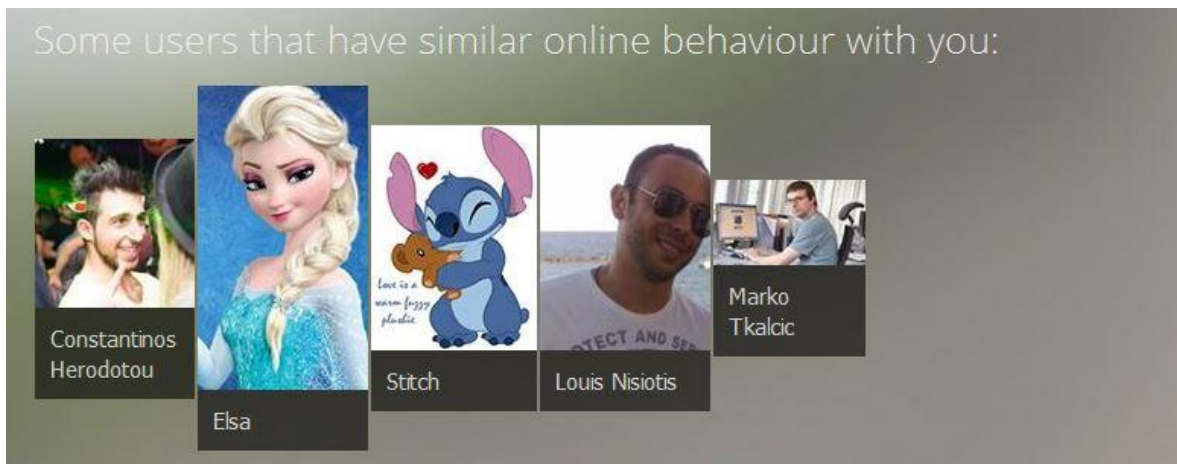
Ακριβώς κάτω από τις γραφικές παραστάσεις προσωπικότητας υπάρχει ένα κουμπί με το οποίο ο χρήστης μπορεί να ανακαλέσει όποτε θέλει το αναδυόμενο παράθυρο που αναφέραμε και πιο πάνω, για να μοιραστεί στο Facebook μέσω μιας κοινοποίησης το σημαντικότερο αποτέλεσμα του μοντέλου προσωπικότητας του.

Στη συνέχεια ο χρήστης, θα παρατηρήσει τον τίτλο ενός συμμαζεμένου τμήματος, στο οποία αν κάνει κλικ θα του εμφανιστούν πληροφορίες για κάθε μία από τις πέντε διαστάσεις του μοντέλου προσωπικότητας του "Big Five". Τα τμήματα αυτά με τις εξηγήσεις, εναλλάσσουν από κλειστά σε ανοιχτά με κάθε κλικ σε αυτά, όπως φαίνεται στην Εικόνα 5-16.



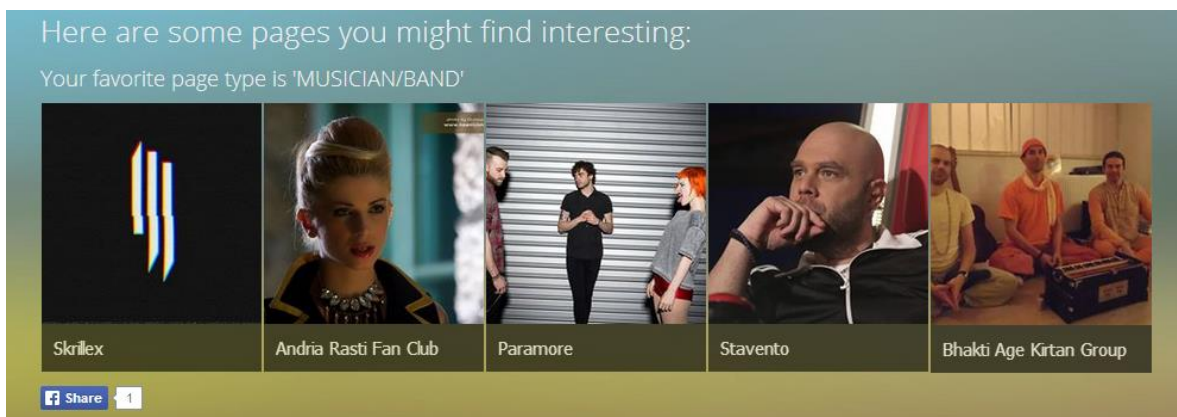
Εικόνα 5-16 Πληροφορίες για κάθε μία από τις πέντε διαστάσεις του μοντέλου προσωπικότητας του “Big Five”

Το επόμενο κομμάτι εμφανίζεται ανάλογα με το αν το μοντέλο προσωπικότητας του χρήστη, μοιάζει με τα μοντέλα άλλων χρηστών του συστήματος μας. Εάν όχι τότε ο χρήστης θα δει μόνο το γεγονός ότι δεν υπάρχουν πανομοιότυποι του χρήστες και θα τον προτρέπουμε να καλέσει τους φίλους του να χρησιμοποιήσουν το σύστημα μας. Εάν όμως υπάρχουν όμοιοι χρήστες με αυτόν, τότε στη συνέχεια θα μπορεί να παρατηρήσει τις εικόνες προφίλ των χρηστών αυτών (Εικόνα 5-17), καθώς επίσης και μια σειρά από σελίδες οι οποίες μπορεί να τον ενδιαφέρουν (Εικόνα 5-18).



Εικόνα 5-17 Εικόνες προφίλ των χρηστών που ο χρήστης έχει ομοιότητα ως προς τη δραστηριότητα του στο Facebook

Τέλος υπάρχει το κουμπί κοινοποίησης, το οποίο μπορεί να χρησιμοποιήσει ο χρήστης για να μοιραστεί την εφαρμογή μας στον δικό του «τοίχο» ή κάποιου φίλου του, ή ακόμη σε κάποια από τις ομάδες που ανήκει στο Facebook (Εικόνα 5-18).

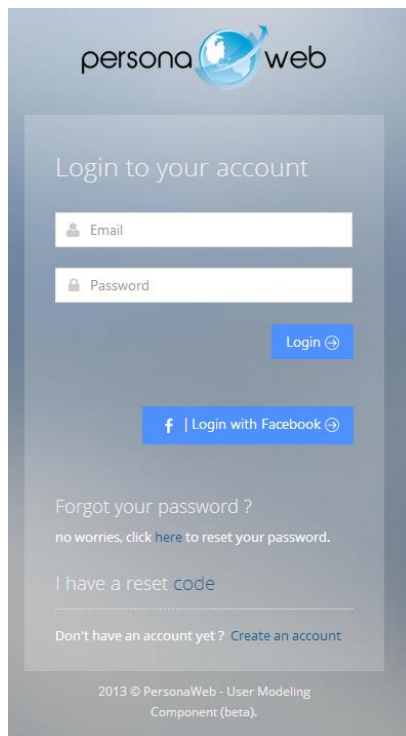


Εικόνα 5-18 Συστάσεις σε μια σειρά από σελίδες οι οποίες μπορεί να ενδιαφέρουν τον χρήστη

5.4 Εξωτερικές Λειτουργίες Εφαρμογής

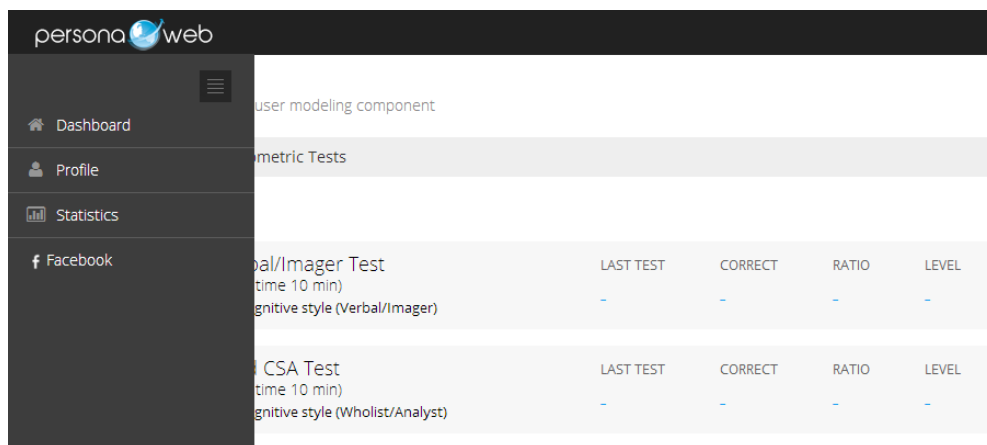
Επίσης δώσαμε την ευκαιρία σε χρήστες του PersonaWeb project, να ενωθούν στο σύστημα μέσω του Facebook και τις εφαρμογές μας, σε αντίθεση με το να δημιουργήσουν ένα καινούργιο προφίλ στην σελίδα σύνδεσης του PersonaWeb. Δημιουργείτε αυτοματοποιημένα ένα προφίλ για αυτούς βάσει των δεδομένων τους στο Facebook. Αφού επιλέξουν να ενωθούν μέσω της

εφαρμογής τότε τους εμφανίζετε το παράθυρο στο οποίο καθορίζεται το συμβόλαιο με την εφαρμογή (Εικόνα 5-6) και ακολούθως τους μεταφέρει στην εφαρμογή απευθείας για να εκκινήσει έτσι τη λειτουργία ανάκτησης των δεδομένων δραστηριότητας τους στο κοινωνικό δίκτυο. Το κουμπί στην σελίδα σύνδεσης του PersonaWeb που εισήχθηκε για τους χρήστες φαίνεται στην Εικόνα 5-19 στην κάτω δεξιά γωνία της φόρμας εισόδου.



Εικόνα 5-19 Φόρμα Εισόδου μέσω PersonaWeb

Ακόμη, οι χρήστες του συστήματος μπορούν να επιλέξουν να μεταφερθούν στην ιστοσελίδα του PersonaWeb για να εκτελέσουν περισσότερα τεστ και ερωτηματολόγια για την προσωπικότητα τους και τα χαρακτηριστικά τους. Μια διευκόλυνση που προσφέρθηκε στους χρήστες σε αυτό το σημείο είναι πως στο αριστερά πτυσσόμενο μενού επιλογών προστέθηκε η επιλογή “Facebook” έτσι ώστε να μπορεί να μεταφερθεί από όπου και να βρισκόταν στη σελίδα του PersonaWeb πίσω στην εφαρμογή. Απαραίτητη προϋπόθεση για να εμφανιστεί η επιλογή αυτή στο μενού ήταν να είχαν μεταφερθεί στην σελίδα του PersonaWeb αρχικά από την εφαρμογή μας. Η επιλογή αυτή μπορεί να παρατηρηθεί στην Εικόνα 5-20.



Εικόνα 5-20 Σύνδεσμος επιστροφής στην εφαρμογή

5.5 Επίλογος

Στο κεφάλαιο αυτό έγινε μια εκτεταμένη περιγραφή σχετικά με τις τεχνικές υλοποίησης της Facebook εφαρμογής που αναπτύχθηκε σε αυτή τη διπλωματική εργασία. Έγινε αναφορά στις τεχνικές εξόρυξης και επεξεργασίας δεδομένων από το Facebook και πως αυτά χρησιμοποιήθηκαν για να υλοποιηθούν βασικά μέρη του μοντέλου χρήστη και εξαγωγή των χαρακτηριστικών προσωπικότητας ενός χρήστη μέσω της δραστηριότητας του στο Facebook. Επίσης στο υποκεφάλαιο 5.3 παρουσιάστηκαν οι βασικές λειτουργίες που ο χρήστης μας μπορεί να εκτελέσει μέσω της εφαρμογής αλλά και οι χρήσιμες εξατομικευμένες πληροφορίες που του προσφέρονται σε αντάλλαγμα με τα δεδομένα που μας πρόσφερε για μελέτη των χαρακτηριστικών προσωπικότητας του.

Στο επόμενο κεφάλαιο θα παρουσιάσουμε την μελέτη αξιολόγησης της τεχνικής εξαγωγής χαρακτηριστικών προσωπικότητας που περιγράφηκε πιο πάνω.

Κεφάλαιο 6

Αξιολόγηση

6.1 Εισαγωγή.....	73
6.2 Σκοπός Αξιολόγησης.....	74
6.3 Περιγραφή Διαδικασίας Αξιολόγησης.....	75
6.3.1 Συγκέντρωση του Δείγματος.....	75
6.4 Περιγραφή Δείγματος	76
6.5 Πειραματική Διάταξη	78
6.6 Ανάλυση Δεδομένων	79
6.7 Επίλογος.....	83

6.1 Εισαγωγή

Στο κεφάλαιο αυτό θα εξηγηθούν ο σκοπός της αξιολόγησης που εκτελέστηκε, όπως επίσης και η λογική πίσω από την δημιουργία του πειράματος αυτού. Ακολούθως θα γίνει μια περιγραφή της διαδικασίας εκτέλεσης της αξιολόγησης μας, η οποία θα περιέχει την συγκέντρωση του δείγματος, τα βήματα του πειράματος και την δραστηριότητα του συστήματος κατά την εκτέλεση αυτού. Στη συνέχεια θα περιγράψουμε τα χαρακτηριστικά του δείγματος που συγκεντρώσαμε και τέλος θα ακολουθήσει μια στατιστική ανάλυση των αποτελεσμάτων των οποίων δημιουργηθήκαν από την μελέτη μας.

6.2 Σκοπός Αξιολόγησης

Όπως έχουμε προαναφέρει στο Κεφάλαιο 2, ερευνητικές ομάδες του κλάδου της ψυχολογίας εκτελούν μια συνεχή και εκτεταμένη μελέτη της προσωπικότητας των ατόμων και των γνωστικών τους χαρακτηριστικών. Με την εξέλιξη της τεχνολογίας τα τελευταία 20 χρόνια και την εισαγωγή των υολογιστών και του διαδικτύου στην καθημερινή ρουτίνα των ανθρώπων, οι έρευνες τους έπρεπε να εξελιχτούν έτσι ώστε να μπορούν να αναγνωρίζουν τα χαρακτηριστικά ενός ατόμου μέσω της δραστηριότητας που δημιουργεί αυτό στο διαδίκτυο. Μεγάλη πρόκληση προσφέρει η εξόρυξη αυτών των δεδομένων από την δραστηριότητα ενός ατόμου στα κοινωνικά δίκτυα, που είναι ένας φλέγων τομέας έρευνας στη σημερινή κοινωνία.

Από μελέτες που έχουν ήδη εκτελεστεί, οι ερευνητές παρατήρησαν ότι για την δραστηριότητα που δημιουργεί ένα άτομο σε κάποιο κοινωνικό δίκτυο, η προσωπικότητα του είναι καλύτερη ένδειξη από τα γνωστικά του χαρακτηριστικά κι αυτό είναι προφανές για το λόγο ότι ο κόσμος χρησιμοποιεί τα κοινωνικά δίκτυα για προσωπική ικανοποίηση και χωρίς να δίνει στις κινήσεις που κάνει μέσα σε αυτά πολλή σκέψη. Έτσι, αν με βάση τα δεδομένα που δημιουργούνται σε αυτά τα περιβάλλοντα εξάγουμε κάποια συμπεράσματα σχετικά με τη προσωπικότητα των χρηστών υποθέτουμε ότι θα είναι αρκετά ακριβής. Ήδη ερευνητές στο τομέα της πληροφορικής έχουν αρχίσει να υλοποιούν αυτή την ιδέα, όπου με υπολογιστικά μέσα γίνεται αναγνώριση της προσωπικότητας ενός ατόμου.

Στη μελέτη που παρουσιάζεται εδώ, ασχοληθήκαμε με άτομα του κοινωνικού δικτύου Facebook το οποίο είναι και το πιο γρήγορα αναδυόμενο από τα σημερινά κοινωνικά δίκτυα. Μετά από την δημιουργία μιας εφαρμογής με την οποία θα μπορούσαμε να υλοποιήσουμε την δική μας ιδέα μέσα στο κοινωνικό δίκτυο Facebook (Κεφάλαιο 4), προσπαθήσαμε να προσεγγίσουμε το θέμα, βασιζόμενοι σε μελέτες (Κεφάλαιο 2) και συμβάλλοντας ερευνητικά στο τομέα αυτό.

Με βάση την βιβλιογραφική ανάλυση που είχαμε δει στο Κεφάλαιο 2, μέσα από μελέτες που εκπονήθηκαν στο τομέα της ψυχολογίας, καταλήξαμε στη συλλογή δεδομένων από το κοινωνικό δίκτυο Facebook, χρησιμοποιώντας τις τεχνικές που περιγράψαμε στο Κεφάλαιο 4, και σε σχέση πάντοτε με τα χαρακτηριστικά προσωπικότητας του μοντέλου “Big Five”. Αφού έχουμε αναλύσει τα δεδομένα αυτά το επόμενο στάδιο είναι η αξιολόγηση της προσέγγισης αυτής. Ο

σκοπός της αξιολόγησης είναι να ελέγξουμε κατά πόσο μπορούμε να εξάγουμε χαρακτηριστικά της προσωπικότητας χρήστη από τα δεδομένα δραστηριότητας του στο κοινωνικό δίκτυο Facebook.

Η βάση της αξιολόγησης μας, θεωρείτε το ερωτηματολόγιο που έχει υλοποιηθεί στα πλαίσια του PersonaWeb project, το οποίο αποτελεί αξιόπιστο μέσω αξιολόγησης της προσωπικότητας ενός ατόμου με βάση το “Big Five” μοντέλο προσωπικότητας.

6.3 Περιγραφή Διαδικασίας Αξιολόγησης

Σε αυτό το υποκεφάλαιο θα δούμε τους τρόπους με τους οποίους συγκεντρώθηκε το δείγμα των ατόμων που συμμετείχαν στην πειραματική αξιολόγηση, την διαδικασία που ακολούθησαν για να εκτελέσουν το πείραμα καθώς επίσης και την δραστηριότητα του συστήματος μας κατά την εκτέλεση του πειράματος.

6.3.1 Συγκέντρωση του Δείγματος

Η συλλογή δείγματος σε μια μελέτη, συνήθως είναι μια χρονοβόρα διαδικασία η οποία καθυστερεί τους ερευνητές να καταλήξουν στα επιθυμητά αποτελέσματα και ακόμη τους περιορίζει στις αλλαγές που μπορεί να αναγκαστούν να κάνουν λόγω της επανάληψης της διαδικασίας. Λόγοι για τους οποίους είναι πιθανό να συμβαίνει αυτό είναι ότι αρχικά μια μελέτη γίνεται σε ελεγχόμενο περιβάλλον από τους ερευνητές. Στη δική μας περίπτωση όμως, το πείραμα χρησιμοποιούσε μια διαδικτυακή εφαρμογή η οποία γινόταν διαθέσιμη μέσω του κοινωνικού δικτύου Facebook. Έτσι, μπορέσαμε με εναλλακτικούς τρόπους να επισπεύσουμε αυτή τη διαδικασία. Οι τρόποι αυτοί είναι οι εξής:

Λόγο της φύσης της εφαρμογής μας, η οποία ήταν μια εφαρμογή στο κοινωνικό δίκτυο Facebook, χρησιμοποιήσαμε βασικές λειτουργίες οι οποίες είναι διαθέσιμες σε αυτό, για να επιτύχουμε το σκοπό μας. Συγκεκριμένα χρησιμοποιήσαμε αυτή την δυνατότητα με τους εξής δύο τρόπους; Πρώτο, επιδιώξαμε να προτρέψουμε το χρήστη να μοιραστεί μέσω μιας κοινοποίησης στο δικό του «τοίχο», τα αποτελέσματα του μοντέλου προσωπικότητας του. Η μέθοδος αυτή ενεργοποιούνταν με δύο ενέργειες. Αρχικά με την είσοδο του χρήστη στην

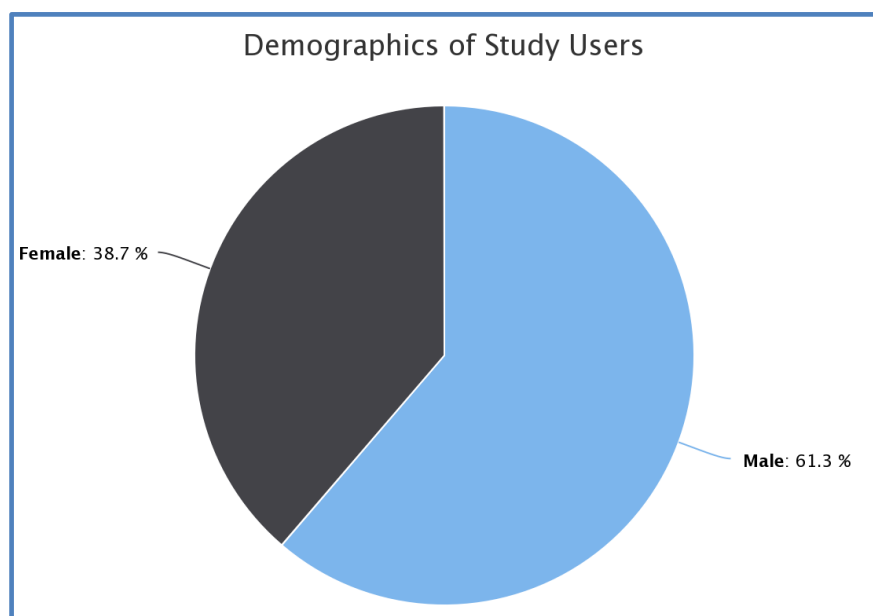
εφαρμογή μας, ένα υπόδειγμα εμφανιζόταν σε αυτόν το οποίο του πρότεινε να δημοσιεύσει τα αποτελέσματα (Εικόνα 5-7) και ακολούθως, αν ο χρήστης επιθυμούσε να μοιραστεί τα αποτελέσματα του, τότε με το πάτημα ενός κουμπιού που βρισκόταν κάτω από την παρουσίαση του μοντέλου του (Εικόνα 5-14), το υπόδειγμα που προαναφέραμε ξαναεμφανιζόταν σε αυτόν. Ο δεύτερος τρόπος ήταν με την χρήση ενός κουμπιού στο τέλος της εφαρμογής μας, ο χρήστης μπορούσε να μοιραστεί την εμπειρία του, στον «τοίχο» του, ή στον «τοίχο» κάποιου φίλου του ή ακόμη σε μια ομάδα στο Facebook στην οποία ανήκει (Εικόνα 5-18). Το κουμπί αυτό προσφερόταν από το Facebook και φέρει το σήμα κατατεθέν της εταιρίας Facebook το οποίο χρησιμοποιούν οι ίδιοι για τις κοινοποιήσεις τους.

Μια άλλη μέθοδος η οποία χρησιμοποιήσαμε είναι αυτή του “Φίλος ενός Φίλου” (Friend of a Friend). Μέσω των διαθέσιμων δυνατοτήτων που προσφέρει το Facebook στους χρήστες του, χρησιμοποιήσαμε προσωπικά μηνύματα (Facebook Messenger) για να προτρέψουμε τους φίλους μας στο Facebook να χρησιμοποιήσουν την εφαρμογή και αυτοί με τη σειρά τους προωθούσαν το μήνυμα στους δικούς τους φίλους.

Ένας διαφορετικός αλλά ευρέως διαδεδομένος τρόπος συλλογής δείγματος στο τομέα της αλληλεπίδρασης ανθρώπου υπολογιστή, είναι η διάδοση μιας ερευνητικής μελέτης ή πειράματος μέσω μιας παγκόσμιας λίστας αλληλογραφίας (International Mailing List). Στη περίπτωση της παρούσας διπλωματικής εργασίας, ένα ηλεκτρονικό μήνυμα με σαφείς οδηγίες ως προς τον σκοπό της μελέτης μας αλλά και τις ενέργειες που θα έπρεπε να κάνει ο χρήστης, προωθήθηκε σε δύο λίστες αλληλογραφίας: (i) um@unito.it και (ii) bcshci@jiscmail.ac.uk. Με τον τρόπο αυτό επιτύχαμε να εμπλουτίσουμε το δείγμα μας με χρήστες πολλαπλών δημογραφικών χαρακτηριστικών.

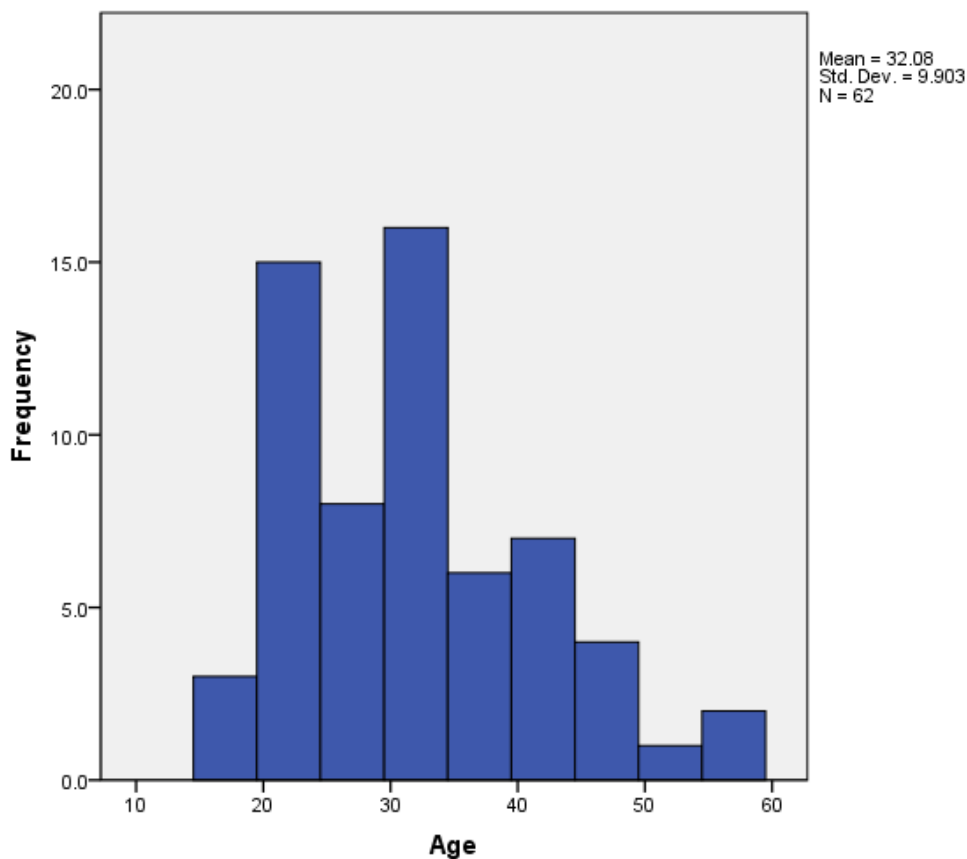
6.4 Περιγραφή Δείγματος

Αφού στάλθηκαν τα μηνύματα, τόσο μέσω Facebook όσο και μέσω ηλεκτρονικού ταχυδρομείου, επισκέφθηκαν την εφαρμογή 75 χρήστες του κοινωνικού δικτύου Facebook σε διάρκεια 8 ημερών. Από τους 75 χρήστες μόνο οι 62 είχαν ολοκληρώσει την συμπλήρωση του ερωτηματολογίου προσωπικότητας μέσω της ιστοσελίδας του PersonaWeb και έτσι αφαιρέθηκαν από τη βάση δεδομένων.



Σχήμα 6-1Δημογραφικά Στοιχεία Συμμετεχόντων – Φύλο

Τα δημογραφικά στοιχεία των συμμετεχόντων είχαν ως εξής: Από τους 62 συνολικά χρήστες του Facebook που συμμετείχαν στην έρευνα μας οι 38 ήταν άνδρες και 24 γυναίκες. Γραφική αναπαράσταση των ποσοστών μπορεί να παρατηρηθεί στο Σχήμα 6-1. Επίσης οι ηλικίες τους κυμαίνονταν από 17 μέχρι 59 ετών Σχήμα 6-2. Η επιτυχία του εύρους της διαφορετικότητας των χρηστών του συστήματος οφείλεται κυρίως στην χρήση των παγκόσμιων λιστών αλληλογραφίας για τη συλλογή του δείγματος μας.



Σχήμα 6-2 Δημογραφικά Συμμετεχόντων - Ηλικίες

6.5 Πειραματική Διάταξη

Για να συμμετάσχουν οι χρήστες στο πείραμα, λάμβαναν ένα μήνυμα με οδηγίες είτε στον Messenger στο Facebook είτε στο ηλεκτρονικό τους ταχυδρομείο. Το μήνυμα που λάμβαναν τους καθοδηγούσε για να κάνουν τα εξής:

Αρχικά ακολουθούσαν τον σύνδεσμο ο οποίος τους οδηγούσε στην εφαρμογή. Αφού ήταν καινούργιοι χρήστες του συστήματος το πρώτο βήμα ήταν να δεχτούν το συμβόλαιο (Εικόνα 4-2) που καθόριζε τα δικαιώματα στην εφαρμογή για να ανακτήσει συγκεκριμένα δεδομένα δραστηριότητας του χρήστη. Αφού η εφαρμογή ανακτούσε τα απαραίτητα δεδομένα τότε εμφάνιζε στο χρήστη εξατομικευμένες πληροφορίες και εξήγαγε το μοντέλου προσωπικότητας με βάση τα δεδομένα δραστηριότητας του.

Ακολουθώς, ο χρήστης πατούσε το κουμπί δίπλα από το μοντέλο προσωπικότητας του (Εικόνα 4-10) το οποίο τον οδηγούσε στην σελίδα του PersonaWeb, στη συγκεκριμένη ιστοσελίδα που περιείχε το ερωτηματολόγιο προσωπικότητας. Τέλος, αφού ολοκλήρωνε το ερωτηματολόγιο, μεταφερόταν αυτόματα πίσω στην εφαρμογή έτσι ώστε να μπορεί και ο ίδιος να δει τα δύο μοντέλα προσωπικότητας που του δημιουργούνταν (το ένα από τα δεδομένα του στο Facebook και το άλλο από τις απαντήσεις του στο ερωτηματολόγιο) (Εικόνα 4-11).

6.6 Ανάλυση Δεδομένων

Σκοπός της πειραματικής μελέτης που εκτελέστηκε ήταν να δούμε κατά πόσο μπορούμε να εξάγουμε χαρακτηριστικά της προσωπικότητας χρήστη από τα δεδομένα δραστηριότητας του στο κοινωνικό δίκτυο Facebook. Βάση για σύγκριση των αποτελεσμάτων μας θεωρήθηκε το ερωτηματολόγιο προσωπικότητας που είχε υλοποιηθεί στα πλαίσια του PersonaWeb project, το οποίο απαντούσαν οι χρήστες οι οποίοι θέλησαν να συμμετάσχουν στο πείραμα μας.

Η προσέγγιση μας ήταν να συγκρίνουμε τα μοντέλα προσωπικότητας μεταξύ τους, δηλαδή, το μοντέλο προσωπικότητας που δημιουργούσαμε αυτόματα με τα δεδομένα προσωπικότητας, με το μοντέλο που προέκυπτε από τις απαντήσεις του ερωτηματολογίου. Για να μπορέσουμε να εκτελέσουμε αυτή τη σύγκριση, έπρεπε να ελέγχουμε τις αντίστοιχες διαστάσεις των δύο μοντέλων μεταξύ τους, αφού δεν οριζόταν κάποιο αποτέλεσμα ώστε να συγκρίνουμε το σύνολο των πέντε διαστάσεων του ενός μοντέλου με του άλλου.

Η προσέγγιση που ακολουθήσαμε είχε ως εξής: Να διαπιστωθεί η συσχέτιση μεταξύ κάθε μιας από τις πέντε διαστάσεις μαζί με τις αντίστοιχες πέντε διαστάσεις του ακριβές μοντέλου που δημιουργούταν από τις απαντήσεις του ερωτηματολογίου. Έτσι, θα μπορούσαμε να συμπεράνουμε ότι το μοντέλο προσωπικότητας που δημιουργούμε με τα δεδομένα δραστηριότητας από το Facebook είναι εξίσου αξιόπιστο.

Τα δεδομένα μας ήταν ποσοστιαία και το δείγμα μας αποτελείτο από 62 τελικά συμμετέχοντες. Χρησιμοποιήσαμε Pearson Correlation για να συγκρίνουμε τις μεταβλητές μας, που σε αυτή τη περίπτωση ήταν οι διαστάσεις της προσωπικότητας.

Pearson Correlation: Είναι μια στατιστική δοκιμή η οποία δημιουργεί ένα συντελεστή που ονομάζεται συντελεστής συσχέτισης Pearson, και συμβολίζεται με r . Ο συντελεστής αυτός μετρά τη δύναμη και την κατεύθυνση μιας γραμμικής σχέσης μεταξύ δύο συνεχών μεταβλητών. Η τιμή του r μπορεί να κυμαίνεται από -1 για μια τέλεια αρνητική γραμμική σχέση μέχρι +1 για μια τέλεια θετική γραμμική σχέση. Η τιμή 0 δείχνει ότι δεν υπάρχει σχέση μεταξύ των δύο μεταβλητών.

Απαραίτητη προϋπόθεση για να μπορεί να εκτελεστεί αυτή η δοκιμή είναι να υπάρχουν δύο μεταβλητές οι οποίες να χαρακτηρίζονται από τα εξής στοιχεία: (i) να είναι συνεχείς και (ii) σε ζεύγος, δηλαδή, σε κάθε περίπτωση να υπάρχουν 2 τιμές (μία για κάθε μεταβλητή).

Θα παρατηρήσετε στους πίνακες που ακολουθούν, ότι τα αποτελέσματα που προκύπτουν είναι σε ένα δισδιάστατο πίνακα με ίδιες γραμμές και στήλες έτσι ώστε κάθε αποτέλεσμα επαναλαμβάνεται. Ωστόσο, τα κελιά που θα μας απασχολήσουν θα είναι μόνο εκείνα που περιέχουν τη συσχέτιση Pearson, δηλαδή το πάνω δεξιά ή το κάτω αριστερά. Τα αποτελέσματα τα οποία πήραμε αφού εκτελέσαμε την δοκιμή Pearson στο δείγμα μας ήταν τα εξής για κάθε διάσταση:

Στον παρακάτω πίνακα παρουσιάζονται τα αποτελέσματα για την συσχέτιση των διαστάσεων της Εξωστρέφειας.

Correlations			
		FbExtraversion	TestExtraversion
FbExtraversion	Pearson Correlation	1	.259*
	Sig. (2-tailed)		.042
	N	62	62
TestExtraversion	Pearson Correlation	.259*	1
	Sig. (2-tailed)	.042	
	N	62	62

Πίνακας 6-1 Αποτελέσματα στατιστικής ανάλυσης - Εξωστρέφεια

Όπως μπορούμε να παρατηρήσουμε από τον πιο πάνω πίνακα, τα αποτελέσματα έδειξαν μια μέτρια θετική συσχέτιση μεταξύ των δύο διαστάσεων της Εξωστρέφειας στα δύο μοντέλα προσωπικότητας, με τιμή $r = 0.259$.

Στον παρακάτω πίνακα παρουσιάζονται τα αποτελέσματα για την συσχέτιση των διαστάσεων της Τερπνότητας.

Correlations			
		FbAgreeableness	TestAgreeableness
FbAgreeableness	Pearson Correlation	1	.032
	Sig. (2-tailed)		.805
	N	62	62
TestAgreeableness	Pearson Correlation	.032	1
	Sig. (2-tailed)	.805	
	N	62	62

Πίνακας 6-2 Αποτελέσματα στατιστικής ανάλυσης - Τερπνότητα

Ο πιο πάνω πίνακα δείχνει ότι υπάρχει κάποια θετική συσχέτιση μεταξύ των διαστάσεων της Τερπνότητας των δύο μοντέλων, με τιμή $r = 0.32$.

Στον παρακάτω πίνακα παρουσιάζονται τα αποτελέσματα για την συσχέτιση των διαστάσεων της Ευσυνειδησίας.

Correlations			
		FbConscientiousness	TestConscientiousness
FbConscientiousness	Pearson Correlation	1	.281 [*]
	Sig. (2-tailed)		.027
	N	62	62
TestConscientiousness	Pearson Correlation	.281 [*]	1
	Sig. (2-tailed)	.027	
	N	62	62

Πίνακας 6-3 Αποτελέσματα στατιστικής ανάλυσης - Ευσυνειδησία

Όπως μπορούμε να παρατηρήσουμε από τον πιο πάνω πίνακα, τα αποτελέσματα έδειξαν μια μέτρια θετική συσχέτιση μεταξύ των δύο διαστάσεων της Ευσυνειδησίας των δύο μοντέλων, με τιμή $r = 0.281$.

Στον πίνακα που ακολουθεί, παρουσιάζονται τα αποτελέσματα για την συσχέτιση των διαστάσεων της Νεύρωσης.

Correlations			
		FbNeuroticism	TestNeuroticism
FbNeuroticism	Pearson Correlation	1	-.010
	Sig. (2-tailed)		.937
	N	62	62
TestNeuroticism	Pearson Correlation	-.010	1
	Sig. (2-tailed)	.937	
	N	62	62

Πίνακας 6-4 Αποτελέσματα στατιστικής ανάλυσης - Νεύρωση

Από τον πιο πάνω πίνακα παρατηρείται ότι, μεταξύ των διαστάσεων της Νεύρωσης των δύο μοντέλων υπάρχει αρνητική συσχέτιση μεταξύ τους, με τιμή $r = -0.010$.

Στον πίνακα που ακολουθεί, παρουσιάζονται τα αποτελέσματα για την συσχέτιση των διαστάσεων της Διαφάνειας.

Correlations			
		FbOpenness	TestOpenness
FbOpenness	Pearson Correlation	1	-.161
	Sig. (2-tailed)		.210
	N	62	62
TestOpenness	Pearson Correlation	-.161	1
	Sig. (2-tailed)	.210	
	N	62	62

Πίνακας 6-5 Αποτελέσματα στατιστικής ανάλυσης - Διαφάνεια

Από τον πιο πάνω πίνακα παρατηρείται ότι, μεταξύ των διαστάσεων της Διαφάνειας των δύο μοντέλων υπάρχει αρνητική συσχέτιση μεταξύ τους, με τιμή $r = -0.161$.

Συνοπτικά, τα αποτελέσματα που δημιουργήθηκαν με τη σύγκριση των διαστάσεων των δύο μοντέλων προσωπικότητας είναι αναμενόμενα. Μπορούμε από τα δεδομένα δραστηριότητας του Facebook να εξάγουμε τις διαστάσεις του Extraversion, του Conscientiousness και του Agreeableness με σχετική ακρίβεια. Χρειάζονται όμως περαιτέρω βελτίωση οι προβλέψεις των διαστάσεων του Neuroticism και του Openness.

Τα αποτελέσματα που είχαμε σε σχέση με τις διαστάσεις προσωπικότητας συνάδουν με τα αποτελέσματα ερευνών τόσο από το κλάδο της ψυχολογίας, όσο και με αυτά που δημοσιεύτηκαν από το κλάδο της πληροφορικής και αναφέρθηκαν στο Κεφάλαιο 2. Οι περιορισμοί τόσο στην εξόρυξη συγκεκριμένων δεδομένων από το Facebook, όσο και στις μελέτες που αναφέρονται στα δεδομένα που σχετίζονται με κάθε διάσταση είχαν σαν αποτέλεσμα τις αρνητικές συσχετίσεις στις διαστάσεις του Neuroticism και του Openess. Περαιτέρω μελέτη χρειάζεται ως προς το ποιά δεδομένα δραστηριότητας από το κοινωνικό δίκτυο συσχετίζονται και πώς (αρνητικά ή θετικά) με τις διαστάσεις αυτές.

Σε γενικές γραμμές οι χρήστες που είχαν επαφή με την εφαρμογή δήλωσαν ότι τα αποτελέσματα ήταν ικανοποιητικά και αντικατόπτριζαν την προσωπικότητα τους. Κάποια σχόλια που είχαμε είναι: «Λίγο πολύ γνωρίζω τον εαυτό μου, και συμφωνώ με βάση του τι πιστεύει το μοντέλο προσωπικότητας από το Facebook για μένα». Επίσης, «Πολύ καλή δουλειά στην εξόρυξη του μοντέλου από το Facebook. Τώρα προχωρώ στο πράσινο κουμπί για το ερωτηματολόγιο».

6.7 Επίλογος

Εν κατακλείδι, στο κεφάλαιο αυτό αναλύουμε και περιγράφουμε όλες τις λεπτομέρειες σχετικά με την αξιολόγηση της μελέτης που παρουσιάστηκε σε αυτή την διπλωματική εργασία. Τα αποτελέσματα έδειξαν ότι μέσα από τα δεδομένα δραστηριότητας ενός χρήστη του κοινωνικού δικτύου Facebook, μπορούμε να πάρουμε σημαντικές πληροφορίες τόσο σχετικά με τη δραστηριότητα του χρήστη όσο και ως προς την προσωπικότητα του. Μπορέσαμε να παρουσιάσουμε στον χρήστη εξατομικευμένες πληροφορίες σχετικά με χρήστες όμοιους με αυτόν, σελίδες στο κοινωνικό δίκτυο που πιθανόν να τον ενδιαφέρουν, αλλά και συγκεντρωτικά στοιχεία σε σχέση με την δραστηριότητα του.

Τα πιο πάνω, ήταν αποτέλεσμα της δημιουργίας ενός μοντέλου χρήστη, το οποίο εφαρμόσαμε ώστε να μπορέσουμε να πάρουμε τις σχετικές πληροφορίες και παράλληλα να δημιουργήσουμε ένα μοντέλο της προσωπικότητας του.

Τα αποτελέσματα της συγκριτικής μελέτης ανάμεσα στο μοντέλο προσωπικότητας χρήστη όπως προέκυψε από τα δεδομένα του Facebook σε σχέση με το μοντέλο προσωπικότητας που

έδωσε το ερωτηματολόγιο για τον χρήστη, έδειξαν ότι για τρείς από τις πέντε διαστάσεις προσωπικότητας του Big Five η συσχέτιση ήταν στατιστικά σημαντική, σε αντίθεση με τις άλλες δύο διαστάσεις όπου χρειάζεται περεταίρω μελέτη.

Κεφάλαιο 7

Συμπεράσματα και Μελλοντικές Προοπτικές

7.1 Συμπεράσματα.....	85
7.2 Περιορισμοί.....	86
7.3 Μελλοντικές Προτάσεις.....	88

7.1 Συμπεράσματα

Τα τελευταία 5 χρόνια έχει παρατηρηθεί ραγδαία ανάπτυξη στη χρήση των κοινωνικών δικτύων. Πλέον στις μέρες μας, πολύ λίγα άτομα που χρησιμοποιούν το διαδίκτυο δεν έχουν τουλάχιστο ένα προφίλ σε κάποιο μέσο κοινωνικής δικτύωσης. Αυτό έχει ως αποτέλεσμα την δημιουργία τεράστιας ποσότητας δεδομένων δραστηριότητας, τα οποία μένουν αναξιοποίητα.

Επιπρόσθετα, με το πέρασ του καιρού, όσο η τεχνολογία αναπτύσσεται και εισέρχεται περαιτέρω στην καθημερινότητα μας, τόσο περισσότερες είναι και οι λειτουργίες οι οποίες αυτοματοποιούνται στη ζωή μας. Ο κλάδος της ψυχολογίας, καθώς και ο τομέας της αλληλεπίδρασης ανθρώπου υπολογιστή, συνεχώς προσπαθούν να εξάγουν συμπεράσματα για ένα άτομο ως προς τη συμπεριφορά του, τα γνωστικά του χαρακτηριστικά, την προσωπικότητα του, τις προτιμήσεις του καθώς επίσης και των προβλεπόμενων αντιδράσεων του. Κύριος τρόπος ανάκτησης των δεδομένων τους μέχρι σήμερα ήταν με τη χρήση τεστ και ερωτηματολογίων τα οποία είναι χρονοβόρα, κουραστικά και στη τελική τα αποτελέσματα τους είναι προκατειλημμένα ως προς την υποσυνείδητη άποψη που έχουν οι χρήστες για τους εαυτούς τους. Έτσι πρέπει να δημιουργηθούν νέοι τρόποι, αυτοματοποιημένοι, με τους οποίους οι

πληροφορίες αυτές να μπορούν να ανακτηθούν από ήδη υπάρχοντα δεδομένα και δραστηριότητες του ατόμου.

Συνδυάζοντας αυτές τις δύο πληροφορίες, η παρούσα διπλωματική εργασία έρχεται να μελετήσει ένα καινοτόμο τρόπο με τον οποίο ερευνητές και οργανισμοί στο Facebook μπορούν να εξάγουν κάποιες από τις πληροφορίες για τους χρήστες του κοινωνικού δικτύου.

Στα πλαίσια της παρούσας διπλωματικής εργασίας, εκτελέσαμε μια έρευνα, η οποία αρχικά μελετούσε μια πληθώρα αναξιοποίητων μέχρι στιγμής δεδομένων δραστηριότητας του κοινωνικού δικτύου Facebook. Σκοπός της έρευνας αυτής ήταν να δώσει νόημα σε αυτά τα δεδομένα και να τα αξιοποιήσει έτσι ώστε να εξάγουμε πληροφορίες για την προσωπικότητα του χρήστη στον οποίο ανήκαν τα δεδομένα.

Έτσι αποφασίσαμε να δημιουργήσουμε μια εφαρμογή στο κοινωνικό δίκτυο Facebook. Η συγκεκριμένη εφαρμογή, μετά από την έγκριση του χρήστη που τη χρησιμοποιεί, ανακτά τα δεδομένα δραστηριότητας του και τα επεξεργάζεται. Επίσης επιστρέφει στο χρήστη αναλυτικά κάποια στοιχεία για αυτόν, ως ανατροφοδότηση για την συλλογή των δεδομένων του. Τέλος εκτελέστηκε μια στατιστική ανάλυση για την μελέτη του κατά πόσο είναι το μοντέλο προσωπικότητας που εξάγουμε ακριβές.

Με αυτή την έρευνα έχουμε καταφέρει και αποδείξει πως όντως είναι δυνατό να δημιουργήσουμε ένα μοντέλο χρήστη χρησιμοποιώντας δεδομένα δραστηριότητας από το κοινωνικό δίκτυο Facebook. Ακολούθως αποδείξαμε ότι εφαρμόζοντας αυτό το μοντέλο μπορούμε να προσφέρουμε εξατομικευμένες πληροφορίες στο χρήστη ως προς την δραστηριότητα του στο Facebook. Και, εν κατακλείδι, μπορέσαμε να δείξουμε πώς με τα δεδομένα δραστηριότητας που εξάγουμε, μπορούμε να αναγνωρίσουμε κάποιες από τις διαστάσεις προσωπικότητας του “Big Five” με σχετική ακρίβεια. Για τις υπόλοιπες χρειάζεται περισσότερη έρευνα ως προς το ποια επιπρόσθετα δεδομένα πρέπει να συλλεχθούν.

7.2 Περιορισμοί

Κατά την εξέλιξη της παρούσας ερευνητικής εργασίας, προέκυψαν ορισμένοι περιορισμοί οι οποίοι θα συζητηθούν σε αυτό το υποκεφάλαιο.

Καταρχάς, απαραίτητο μέρος της υλοποίησης που δημιουργήσαμε, ήταν η επικοινωνία με το κοινωνικό δίκτυο Facebook. Το Facebook, είναι το πιο διαδεδομένο κοινωνικό δίκτυο στις μέρες μας και συνεχώς προσθέτει νέες λειτουργίες και διευκολύνσεις τόσο για τους χρήστες του, όσο και για τους προγραμματιστές του. Κατά την δημιουργία της εφαρμογής στο Facebook χρησιμοποιούσαμε την ειδική γλώσσα ερωτημάτων FQL. Για την εκτέλεση της FQL οριζόταν συγκεκριμένη μέθοδος από το Facebook, η οποία στη συνέχεια αναβαθμίστηκε και άλλαξε εντελώς. Ο νέος τρόπος όντως ήταν πιο εύκολος στη χρήση, αλλά είχε ως συνέπεια η δουλειά που είχε υλοποιηθεί μέχρι εκείνη τη στιγμή να προσαρμοστεί από την αρχή.

Επίσης, λόγω του ότι το Facebook διαχειρίζεται ένα απεριόριστο αριθμό προσωπικών δεδομένων των χρηστών του, έχει ένα μεγάλο αριθμό περιοριστικών πολιτικών προστασίας οι οποίες μας εμπόδισαν με δύο τρόπους. Πρώτα απ' όλα, δεν θέλαμε να εισβάλουμε σε μεγάλο βαθμό στην προσωπική ζωή του χρήστη, γιατί αυτό θα μείωνε κατά πολύ τους χρήστες οι οποίοι θα δέχονταν να χρησιμοποιήσουν την εφαρμογή μας. Έτσι αποφύγαμε να συλλέξουμε δεδομένα όπως τα προσωπικά μηνύματα που αποστέλλουν οι χρήστες του Facebook με τους φίλους τους, παρά το ότι εκδήλωναν μεγάλο ερευνητικό ενδιαφέρον ως δεδομένα. Επιπρόσθετα, το Facebook έχει πολιτικές που επιτρέπουν τη χρήση ορισμένων από τους πίνακες τους για συγκεκριμένη χρήση μόνο. Το γεγονός αυτό μας απότρεψε να χρησιμοποιήσουμε κάποια από τα δεδομένα με τον τρόπο που επιθυμούσαμε για να εξάγουμε τις πληροφορίες που θέλαμε.

Από τη συνεχή μας αλληλεπίδραση με τη βάση του Facebook παρατηρήσαμε ότι αν χρειαζόταν αρκετός χρόνος ώστε να εκτελεστεί ένα ερώτημα προς τη βάση από την εφαρμογή, η διαδικασία τερματιζόταν χωρίς αποτέλεσμα. Αυτό μας περιόρισε στο να χρησιμοποιούμε ερωτήματα τα οποία επέστρεφαν πάντα αποτέλεσμα για να μην δημιουργηθούν ελλιπής μοντέλα για κάποιους από τους χρήστες μας.

Στην προσπάθεια να εξάγουμε όσο το δυνατόν περισσότερα δεδομένα για ένα χρήστη, (π.χ. τα τελευταία 2500 likes που έκανε ο χρήστης), παρατηρήθηκε ότι ορισμένοι από τους πίνακες της βάσης δεδομένων του Facebook κρατούν μόνο τις τελευταίες 1000 δραστηριότητες που είχε ο χρήστης στο κοινωνικό δίκτυο (όπως για παράδειγμα ο πίνακας των `likes`). Παρόλο που για τους προγραμματιστές του Facebook η πληροφορία μπορεί να αντλείτε με κάποιο άλλο τρόπο, για τους εξωτερικούς προγραμματιστές, οι οποίοι έχουν περιορισμένα δικαιώματα, πολλά από τα δεδομένα του χρήστη ήταν περιορισμένης εμβέλειας.

Λόγω του ότι ένα μεγάλο ποσοστό των χρηστών μας ήταν ελληνόφωνοι χρήστες του Facebook και λόγω του ότι ένα μεγάλο ποσοστό των ελληνόφωνων χρηστών του διαδικτύου χρησιμοποιούν λατινικούς χαρακτήρες για να γράφουν σε ελληνική διάλεκτο, δεν μπορούσαμε να χρησιμοποιήσουμε ανάλυση κειμένου ώστε να επεξεργαστούμε πληροφορίες όπως το περιεχόμενο των καταστάσεων που δημοσιεύει ή τη περιγραφή που δίνει ένας χρήστης σε φωτογραφίες του.

Επιπλέον, κατά την παρουσίαση των όμοιων χρηστών του συστήματος στον τρέχον χρήστη (Υποκεφάλαιο) χρησιμοποιούμε την Συνημιτονοειδή Ομοιότητα για εξαγωγή αποτελεσμάτων. Γνωρίζουμε ότι με ομαδοποίηση (clustering) των χρηστών μας θα μπορούσαμε να είχαμε πιο ακριβείς προβλέψεις, όμως λόγω του ότι κάθε φορά που ο χρήστης προσπαθούσε να εισέλθει στην εφαρμογή μας θα εκτελείτο ο κώδικας της ομαδοποίησης, τότε θα είχαμε καθυστέρηση στο να φορτώσει το σύστημα και οι χρήστες θα αποχωρούσαν από την εφαρμογή πριν καν ανοίξει.

Κατά την συλλογή του δείγματος των χρηστών που θα εκτελούσαν το πείραμα μας, ανακαλύψαμε ότι οι χρήστες μας δεν μπορούσαν να εισέλθουν στην εφαρμογή μας, μέσω των εφαρμογών Facebook που διατίθενται σε κινητές συσκευές. Αυτό μείωνε τον αριθμό των υποψήφιων χρηστών οι οποίοι επέλεγαν να μας δώσουν τα δεδομένα τους αλλά μπορούσαν να χρησιμοποιήσουν το Facebook μόνο από τις κινητές τους συσκευές.

Τέλος, το ερωτηματολόγιο προσωπικότητας το οποίο χρησιμοποιούσαμε σαν βάση για την στατιστική ανάλυση των αποτελεσμάτων μας, δεν ήταν υποχρεωτικό να γίνει από τους χρήστες μας. Γι' αυτό και στο υποκεφάλαιο 5.κάτι μπορούμε να παρατηρήσουμε ότι είχαμε περισσότερους χρήστες που εισήλθαν στην εφαρμογή αλλά επέλεξαν να μην ολοκληρώσουν την διαδικασία και έτσι να μην περιληφθούν στο δείγμα μελέτης που είχαμε.

7.3 Μελλοντικές Προτάσεις

Η προσέγγιση και η υλοποίηση που ακολουθήθηκε στη παρούσα διπλωματική εργασία είναι καινοτόμα. Αυτό σημαίνει ότι υπάρχουν βελτιστοποιήσεις και ιδέες επέκτασης οι οποίες μπορούν να υλοποιηθούν πάνω στο σύστημα που δημιουργήθηκε για την βελτίωση των αποτελεσμάτων της έρευνας. Πιο κάτω θα θίξουμε κάποιες από τις μελλοντικές εργασίες, η οποίες προκύπτουν από την δική μας μελέτη:

Όπως είχαμε δει στο υποκεφάλαιο 5.κάτι τα αποτελέσματα των διαστάσεων της Νεύρωσης και της Διαφάνειας χρειάζονται περαιτέρω βελτίωση ώστε να μπορέσουμε να διακρίνουμε συσχέτιση μεταξύ αυτών και των πιο ακριβή μετρήσεων που παίρναμε από το ερωτηματολόγιο της ψυχολογίας. Αυτό θα μπορούσε να επιτευχθεί με την επέκταση των δεδομένων δραστηριότητας τα οποία ανακτούμε από τους χρήστες του Facebook. Επίσης, περαιτέρω μελέτη χρειάζεται ως προς τις τιμές των ορίων και των βαρών σημασίας που καθορίσαμε για καλύτερα αποτελέσματα.

Όπως έχουμε αναφέρει και πιο πάνω και στο Υποκεφάλαιο 4.κάτι, η εύρεση των όμοιων χρηστών γίνεται με τη χρήση της Συνημιτονοειδής Ομοιότητας αντί με κάποια μέθοδο ομαδοποίησης (όπως παραδείγματος χάρη, η k-means ομαδοποίηση). Μια καλή επέκταση της μεθόδου αυτής θα ήταν να εκτελεστεί ομαδοποίηση όχι σε πραγματικό χρόνο και οι ομάδες που θα δημιουργηθούν να αποθηκευτούν έτσι ώστε με τον ερχομό ενός νέου χρήστη στο σύστημα απλά να ελέγχεται σε πια ομάδα ανήκει και να επιλέγονται χρήστες από αυτή την ομάδα. Αυτό όμως επιβάλλει τον κώδικα της ομαδοποίησης να εκτελείτε τακτικά ανά διαστήματα για το λόγο ότι μπορεί να αλλάξουν οι ομάδες υπό την επήρεια περισσότερων χρηστών.

Κατά την εκτέλεση του πειράματος μας, παρατηρήσαμε ότι η χρήστες μας δεν είχαν την ευχέρεια να εισέρθουν στην εφαρμογή μας από τις εφαρμογές Facebook που υπάρχουν σε κινητές συσκευές. Για τη λύση του προβλήματος αυτού χρειάζεται να επεκταθεί (να υλοποιηθεί) η εφαρμογή μας και σε λειτουργικά συστήματα ios και android για να μπορεί να υποστηρίζεται όλες τις πλατφόρμες του Facebook.

Τέλος, μια ενδιαφέρον ερευνητική πρόταση προκύπτει από το γεγονός ότι τα δεδομένα που συλλέγουμε για αντικείμενα του κοινωνικού δικτύου Facebook που συσχετίζονται οι χρήστες μας, αποθηκεύονται αναλυτικά στη βάση δεδομένων του Πανεπιστημίου Κύπρου. Η ιδέα που προκύπτει είναι να δημιουργηθεί ένας γράφος των αντικειμένων των χρηστών του συστήματος για εντοπισμό συσχέτισης μεταξύ των αντικειμένων του Facebook αντί των χρηστών του.

Βιβλιογραφία

- [1] C.J. McElroy, R.A. Hendrickson, M.A. Townsend and M.S. DeMarie, "Dispositional factors in internet use," MIS Quarterly, vol. 31, pp. 809-820, 2007.
- [2] K. Moore and J.C. McElroy, "The influence of personality on Facebook usage, wall postings, and regret," Comput.Hum.Behav., vol. 28, pp. 267-274, 2012.
- [3] L.R. Goldberg, "An alternative "description of personality": the big-five factor structure," J.Pers.Soc.Psychol., vol. 59, pp. 1216-1229, Dec. 1990.
- [4] E.C. Tupes and R.E. Christal, "Recurrent Personality Factors Based on Trait Ratings," J.Pers., vol. 60, pp. 225 - 251, 1992.
- [5] Y. Bachrach, M. Kosinski, T. Graepel, P. Kohli and D. Stillwell, "Proceedings of the 3rd Annual ACM Web Science Conference on - WebSci '12; Personality and patterns of Facebook usage," pp. 24 - 32, 2012.
- [6] J. Golbeck, C. Robles and K. Turner, "Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems - CHI EA '11; Predicting personality with social media," pp. 253, 2011.
- [7] A. Ortigosa, R.M. Carro and J.I. Quiroga, "Predicting user personality by mining social interactions in Facebook," Journal of Computer and System Sciences, vol. 80, pp. 57 - 71, 2014.
- [8] P. Brusilovsky, A. Kobsa and W. Nejdl, "The Adaptive Web," vol. 4321, 2007.

- [9] P. Lops, M. de Gemmis and G. Semeraro, "Content-based recommender systems: State of the art and trends," in *Recommender Systems Handbook*, Springer, 2011, pp. 73-105.
- [10] "FQL Overview," vol. 2014, .

Παράρτημα Α.

Δημιουργία των πινάκων της βάσης δεδομένων

```
-- phpMyAdmin SQL Dump
-- version 3.5.8.2
-- http://www.phpmyadmin.net
--
-- Host: dbserver.in.cs.ucy.ac.cy
-- Generation Time: May 28, 2014 at 02:58 PM
-- Server version: 5.0.95-log
-- PHP Version: 5.3.3

SET SQL_MODE="NO_AUTO_VALUE_ON_ZERO";
SET time_zone = "+00:00";

/*!40101 SET @OLD_CHARACTER_SET_CLIENT=@@CHARACTER_SET_CLIENT */;
/*!40101 SET @OLD_CHARACTER_SET_RESULTS=@@CHARACTER_SET_RESULTS */;
/*!40101 SET @OLD_COLLATION_CONNECTION=@@COLLATION_CONNECTION */;
/*!40101 SET NAMES utf8 */;

--
-- Database: `personaweb`
--
-- -----
--
-- Table structure for table `fb_event_attended`
--
CREATE TABLE IF NOT EXISTS `fb_event_attended` (
  `usr_id` varchar(20) NOT NULL,
  `event_id` varchar(25) NOT NULL,
  `owner` varchar(20) NOT NULL,
  `name` varchar(100) NOT NULL,
  `pic_url` varchar(200) NOT NULL,
```

```

PRIMARY KEY (`usr_id`,`event_id`)
) ENGINE=MyISAM DEFAULT CHARSET=latin1;

```

--

-- Table structure for table `fb\_user`

--

```

CREATE TABLE IF NOT EXISTS `fb_user` (
  `id` varchar(25) NOT NULL,
  `Name` varchar(50) NOT NULL,
  `Email` varchar(50) NOT NULL,
  `DOB` varchar(15) NOT NULL,
  `sex` varchar(6) NOT NULL,
  `img_url` varchar(100) NOT NULL,
  `Personaweb_id` int(11) NOT NULL,
  PRIMARY KEY (`id`)
) ENGINE=MyISAM DEFAULT CHARSET=latin1;

```

--

-- Table structure for table `fb\_user\_active\_friends`

--

```

CREATE TABLE IF NOT EXISTS `fb_user_active_friends` (
  `usr_id` varchar(25) NOT NULL,
  `active_friend_id` varchar(25) NOT NULL,
  PRIMARY KEY (`usr_id`,`active_friend_id`)
) ENGINE=MyISAM DEFAULT CHARSET=latin1;

```

--

-- Table structure for table `fb\_user\_friends`

--

```
CREATE TABLE IF NOT EXISTS `fb_user_friends` (
  `usr_id` varchar(20) NOT NULL,
  `friend_id` varchar(20) NOT NULL,
  PRIMARY KEY (`usr_id`,`friend_id`)
) ENGINE=MyISAM DEFAULT CHARSET=latin1;
```

```
-- -----
```

```
--
```

```
-- Table structure for table `fb_user_items`
```

```
--
```

```
CREATE TABLE IF NOT EXISTS `fb_user_items` (
  `usr_id` varchar(20) NOT NULL,
  `object_id` varchar(20) NOT NULL,
  `object_type` varchar(8) NOT NULL,
  PRIMARY KEY (`usr_id`,`object_id`)
) ENGINE=MyISAM DEFAULT CHARSET=latin1;
```

```
-- -----
```

```
--
```

```
-- Table structure for table `fb_user_likes`
```

```
--
```

```
CREATE TABLE IF NOT EXISTS `fb_user_likes` (
  `usr_id` varchar(20) NOT NULL,
  `object_id` varchar(20) NOT NULL,
  `object_type` varchar(20) NOT NULL,
  `object_owner` varchar(20) NOT NULL,
  PRIMARY KEY (`usr_id`,`object_id`)
) ENGINE=MyISAM DEFAULT CHARSET=latin1;
```

```
-- -----
```

```
--
```

```
-- Table structure for table `fb_user_model`
```

```
--
```

```
CREATE TABLE IF NOT EXISTS `fb_user_model` (  
  `usr_id` varchar(25) NOT NULL,  
  `Extraversion` float NOT NULL,  
  `Agreeableness` float NOT NULL,  
  `Conscientiousness` float NOT NULL,  
  `Neuroticism` float NOT NULL,  
  `Openness` float NOT NULL,  
  PRIMARY KEY (`usr_id`)  
) ENGINE=MyISAM DEFAULT CHARSET=utf8;
```

```
-----
```

```
--
```

```
-- Table structure for table `fb_user_model_test`
```

```
--
```

```
CREATE TABLE IF NOT EXISTS `fb_user_model_test` (  
  `usr_id` varchar(25) NOT NULL,  
  `Extraversion` float NOT NULL,  
  `Agreeableness` float NOT NULL,  
  `Conscientiousness` float NOT NULL,  
  `Neuroticism` float NOT NULL,  
  `Openness` float NOT NULL,  
  PRIMARY KEY (`usr_id`)  
) ENGINE=MyISAM DEFAULT CHARSET=utf8;
```

```
-----
```

```
--
```

```
-- Table structure for table `fb_user_msgs`
```

```
--
```

```
CREATE TABLE IF NOT EXISTS `fb_user_msgs` (  
  `usr_id` varchar(20) NOT NULL,
```

```

`author_id` varchar(20) NOT NULL,
`created_time` int(32) unsigned NOT NULL,
`thread_id` varchar(25) NOT NULL
) ENGINE=MyISAM DEFAULT CHARSET=latin1;

```

--

-- Table structure for table `fb\_user\_pages`

--

```

CREATE TABLE IF NOT EXISTS `fb_user_pages` (
  `usr_id` varchar(20) NOT NULL,
  `page_id` varchar(20) NOT NULL,
  `page_type` varchar(50) NOT NULL,
  PRIMARY KEY (`usr_id`,`page_id`)
) ENGINE=MyISAM DEFAULT CHARSET=latin1;

```

--

-- Table structure for table `fb\_user\_tags`

--

```

CREATE TABLE IF NOT EXISTS `fb_user_tags` (
  `usr_id` varchar(20) NOT NULL,
  `object_id` varchar(25) NOT NULL,
  `object_owner` varchar(20) NOT NULL,
  `type` varchar(8) NOT NULL,
  PRIMARY KEY (`usr_id`,`object_id`)
) ENGINE=MyISAM DEFAULT CHARSET=latin1;

```

--

-- Table structure for table `fb\_user\_vector`

--

```

CREATE TABLE IF NOT EXISTS `fb_user_vector` (
  `usr_id` varchar(25) NOT NULL,
  `active` int(11) NOT NULL,
  `checkins` int(11) NOT NULL,
  `events` int(11) NOT NULL,
  `friends` int(11) NOT NULL,
  `likes` int(11) NOT NULL,
  `links` int(11) NOT NULL,
  `page_type` int(11) NOT NULL,
  `pages` int(11) NOT NULL,
  `prof_pic` int(11) NOT NULL,
  `self_tag` int(11) NOT NULL,
  `status` int(11) NOT NULL,
  `tags` int(11) NOT NULL,
  `user_checkins` int(11) NOT NULL,
  `user_events` int(11) NOT NULL,
  PRIMARY KEY (`usr_id`)
) ENGINE=MyISAM DEFAULT CHARSET=latin1;

```

```

--
-- Table structure for table `fb_user_vector_numeric`
--

```

```

CREATE TABLE IF NOT EXISTS `fb_user_vector_numeric` (
  `usr_id` varchar(25) NOT NULL,
  `active` int(11) NOT NULL,
  `checkins` int(11) NOT NULL,
  `events` int(11) NOT NULL,
  `friends` int(11) NOT NULL,
  `likes` int(11) NOT NULL,
  `links` int(11) NOT NULL,
  `page_type` float NOT NULL,
  `pages` int(11) NOT NULL,
  `prof_pic` float NOT NULL,

```

```

`self_tag` int(11) NOT NULL,
`status` int(11) NOT NULL,
`tags` int(11) NOT NULL,
`user_checkins` int(11) NOT NULL,
`user_events` int(11) NOT NULL,
PRIMARY KEY (`usr_id`)
) ENGINE=MyISAM DEFAULT CHARSET=latin1;

-----

--
-- Table structure for table `system_personalityanswers`
--

CREATE TABLE IF NOT EXISTS `system_personalityanswers` (
  `SPA_ID` int(11) NOT NULL auto_increment,
  `SPA_occID` int(11) NOT NULL,
  `SPA_questiontype` int(11) NOT NULL,
  `SPA_questionNum` int(11) NOT NULL,
  `SPA_gIVENANSWER` int(11) NOT NULL,
  `SPA_dateadded` timestamp NOT NULL default CURRENT_TIMESTAMP,
  PRIMARY KEY (`SPA_ID`)
) ENGINE=MyISAM DEFAULT CHARSET=latin1 AUTO_INCREMENT=5751 ;

-----

--
-- Table structure for table `system_personalityoccurences`
--

CREATE TABLE IF NOT EXISTS `system_personalityoccurences` (
  `SPO_ID` int(11) NOT NULL auto_increment,
  `SPO_UserID` int(11) NOT NULL,
  `SPO_type` int(11) NOT NULL,
  `SPO_IsComplete` int(11) NOT NULL default '0',
  `SPO_DateAdded` timestamp NOT NULL default CURRENT_TIMESTAMP,
  PRIMARY KEY (`SPO_ID`)

```

```
) ENGINE=MyISAM  DEFAULT CHARSET=latin1 AUTO_INCREMENT=157 ;
```

```
-- -----
```

```
--
```

```
-- Table structure for table `system_users`
```

```
--
```

```
CREATE TABLE IF NOT EXISTS `system_users` (  
  `ID` int(11) NOT NULL auto_increment,  
  `Name` varchar(50) NOT NULL,  
  `Email` varchar(50) NOT NULL,  
  `Password` varchar(12) NOT NULL,  
  `Identity` varchar(30) character set utf8 collate utf8_unicode_ci NOT NULL,  
  `sex` varchar(1) NOT NULL,  
  `DOB` varchar(15) NOT NULL,  
  `Profession` varchar(20) NOT NULL,  
  `Knowledgelevel` varchar(20) NOT NULL,  
  `Department` varchar(20) NOT NULL,  
  `Country` varchar(10) NOT NULL,  
  `PQ` varchar(1) NOT NULL,  
  `SQ` varchar(1) NOT NULL,  
  `fb_id` varchar(50) NOT NULL,  
  PRIMARY KEY (`ID`)  
) ENGINE=InnoDB  DEFAULT CHARSET=latin1 AUTO_INCREMENT=10149 ;
```

```
-- -----
```

```
--
```

```
-- Table structure for table `ref_personalitytest`
```

```
--
```

```
CREATE TABLE IF NOT EXISTS `ref_personalitytest` (  
  `RPT_ID` int(11) NOT NULL,  
  `RPT_name` varchar(200) character set utf8 collate utf8_unicode_ci NOT NULL,  
  `RPT_questionType` int(11) NOT NULL,
```

```

`RPT_positiveNegative` int(11) NOT NULL,
`RPT_smallQuestionnaire` int(11) NOT NULL,
`RPT_ID_SQ` int(11) NOT NULL,
PRIMARY KEY (`RPT_ID`)
) ENGINE=MyISAM DEFAULT CHARSET=latin1;

--
-- Dumping data for table `ref_personalitytest`
--

INSERT INTO `ref_personalitytest` (`RPT_ID`, `RPT_name`, `RPT_questionType`,
`RPT_positiveNegative`, `RPT_smallQuestionnaire`, `RPT_ID_SQ`) VALUES
(1, 'Am the life of the party.', 1, 1, 1, 1),
(2, 'Insult people.', 2, 0, 1, 12),
(3, 'Am always prepared.', 3, 1, 1, 3),
(4, 'Get stressed out easily.', 4, 0, 1, 4),
(5, 'Have a rich vocabulary.', 5, 1, 1, 5),
(6, 'Often feel uncomfortable around others.', 1, 0, 0, 0),
(7, 'Am interested in people.', 2, 1, 1, 7),
(8, 'Leave my belongings around.', 3, 0, 1, 8),
(9, 'Am relaxed most of the time.', 4, 1, 1, 9),
(10, 'Have difficulty understanding abstract ideas.', 5, 0, 1, 10),
(11, 'Feel comfortable around people.', 1, 1, 1, 11),
(12, 'Am not interested in other people's problems.', 2, 0, 1, 22),
(13, 'Pay attention to details.', 3, 1, 1, 13),
(14, 'Worry about things.', 4, 0, 1, 14),
(15, 'Have a vivid imagination.', 5, 1, 1, 15),
(16, 'Keep in the background.', 1, 0, 1, 16),
(17, 'Sympathize with others' feelings.', 2, 1, 1, 17),
(18, 'Make a mess of things.', 3, 0, 1, 18),
(19, 'Seldom feel blue.', 4, 1, 1, 19),
(20, 'Am not interested in abstract ideas.', 5, 0, 1, 20),
(21, 'Start conversations.', 1, 1, 1, 21),
(22, 'Feel little concern for others.', 2, 0, 1, 2),
(32, 'Am not really interested in others.', 2, 0, 1, 32),
(23, 'Get chores done right away.', 3, 1, 1, 23),
(24, 'Am easily disturbed.', 4, 0, 1, 24),
(25, 'Have excellent ideas.', 5, 1, 1, 25),

```

(26, 'Have little to say.', 1, 0, 1, 26),
(27, 'Have a soft heart.', 2, 1, 1, 27),
(28, 'Often forget to put things back in their proper place.', 3, 0, 1, 28),
(29, 'Am not easily bothered by things.', 4, 1, 0, 0),
(30, 'Do not have a good imagination.', 5, 0, 1, 30),
(31, 'Talk to a lot of different people at parties.', 1, 1, 1, 31),
(33, 'Like order.', 3, 1, 1, 33),
(34, 'Get upset easily.', 4, 0, 1, 29),
(35, 'Am quick to understand things.', 5, 1, 1, 35),
(36, 'Don''t like to draw attention to myself.', 1, 0, 1, 36),
(37, 'Take time out for others.', 2, 1, 1, 37),
(38, 'Shirk my duties.', 3, 0, 1, 38),
(39, 'Rarely get irritated.', 4, 1, 0, 0),
(40, 'Try to avoid complex people.', 5, 0, 0, 0),
(41, 'Don''t mind being the center of attention.', 1, 1, 1, 41),
(42, 'Am hard to get to know.', 2, 0, 0, 0),
(43, 'Follow a schedule.', 3, 1, 1, 43),
(44, 'Change my mood a lot.', 4, 0, 1, 34),
(45, 'Use difficult words.', 5, 1, 1, 40),
(46, 'Am quiet around strangers.', 1, 0, 1, 46),
(47, 'Feel others'' emotions.', 2, 1, 1, 42),
(48, 'Neglect my duties.', 3, 0, 0, 0),
(49, 'Seldom get mad.', 4, 1, 0, 0),
(50, 'Have difficulty imagining things.', 5, 0, 0, 0),
(51, 'Make friends easily.', 1, 1, 0, 0),
(52, 'Am indifferent to the feelings of others.', 2, 0, 0, 0),
(53, 'Am exacting in my work.', 3, 1, 1, 48),
(54, 'Have frequent mood swings.', 4, 0, 1, 39),
(55, 'Spend time reflecting on things.', 5, 1, 1, 45),
(56, 'Find it difficult to approach others.', 1, 0, 0, 0),
(57, 'Make people feel at ease.', 2, 1, 1, 47),
(58, 'Waste my time.', 3, 0, 0, 0),
(59, 'Get irritated easily.', 4, 0, 1, 44),
(60, 'Avoid difficult reading material.', 5, 0, 0, 0),
(61, 'Take charge.', 1, 1, 0, 0),
(62, 'Inquire about others'' well-being.', 2, 1, 0, 0),
(63, 'Do things according to a plan.', 3, 1, 0, 0),

(64, 'Often feel blue.', 4, 0, 1, 49),
(65, 'Am full of ideas.', 5, 1, 1, 50),
(66, 'Don''t talk a lot.', 1, 0, 1, 6),
(67, 'Know how to comfort others.', 2, 1, 0, 0),
(68, 'Do things in a half-way manner.', 3, 0, 0, 0),
(69, 'Get angry easily.', 4, 0, 0, 0),
(70, 'Will not probe deeply into a subject.', 5, 0, 0, 0),
(71, 'Know how to captivate people.', 1, 1, 0, 0),
(72, 'Love children.', 2, 1, 0, 0),
(73, 'Continue until everything is perfect.', 3, 1, 0, 0),
(74, 'Panic easily.', 4, 0, 0, 0),
(75, 'Carry the conversation to a higher level.', 5, 1, 0, 0),
(76, 'Bottle up my feelings.', 1, 0, 0, 0),
(77, 'Am on good terms with nearly everyone.', 2, 1, 0, 0),
(78, 'Find it difficult to get down to work.', 3, 0, 0, 0),
(79, 'Feel threatened easily.', 4, 0, 0, 0),
(80, 'Catch on to things quickly.', 5, 1, 0, 0),
(81, 'Feel at ease with people.', 1, 1, 0, 0),
(82, 'Have a good word for everyone.', 2, 1, 0, 0),
(83, 'Make plans and stick to them.', 3, 1, 0, 0),
(84, 'Get overwhelmed by emotions.', 4, 0, 0, 0),
(85, 'Can handle a lot of information.', 5, 1, 0, 0),
(86, 'Am a very private person.', 1, 0, 0, 0),
(87, 'Show my gratitude.', 2, 1, 0, 0),
(88, 'Leave a mess in my room.', 3, 0, 0, 0),
(89, 'Take offense easily.', 4, 0, 0, 0),
(90, 'Am good at many things.', 5, 1, 0, 0),
(91, 'Wait for others to lead the way.', 1, 0, 0, 0),
(92, 'Think of others first.', 2, 1, 0, 0),
(93, 'Love order and regularity.', 3, 1, 0, 0),
(94, 'Get caught up in my problems.', 4, 0, 0, 0),
(95, 'Love to read challenging material.', 5, 1, 0, 0),
(96, 'Am skilled in handling social situations.', 1, 1, 0, 0),
(97, 'Love to help others.', 2, 1, 0, 0),
(98, 'Like to tidy up.', 3, 1, 0, 0),
(99, 'Grumble about things.', 4, 0, 0, 0),
(100, 'Love to think up new ways of doing things.', 5, 1, 0, 0);

```
/*!40101 SET CHARACTER_SET_CLIENT=@OLD_CHARACTER_SET_CLIENT */;  
/*!40101 SET CHARACTER_SET_RESULTS=@OLD_CHARACTER_SET_RESULTS */;  
/*!40101 SET COLLATION_CONNECTION=@OLD_COLLATION_CONNECTION */;
```

Παράρτημα Β.

Τιμές ορίων (thresholds)

```
// number of likes
$thresh_likes = array(500,750,1000,1250);
// number of pages
$thresh_pages = array(50,150,350,500);
// number of friends
$thresh_friends = array(250,500,750,1000);
// active friends
$thresh_active = array(3,7,15,30);
// self tag count
$thresh_self_tag = array(2,5,10,20);
// number of events
$thresh_events = array(10,30,50,80);
// number of user events
$thresh_u_events = array(1,3,7,10);
// number of check-ins
$thresh_checkins = array(50,100,200,300);
// number of user check-ins
$thresh_u_checkins = array(25,50,75,100);
// number of user's status
$thresh_status = array(25,50,75,100);
// number of user's links
$thresh_links = array(50,150,350,500);
// number of tags
$thresh_tags = array(50,90,140,200);
// percentage of favorite page type
$thresh_page_type = array(0.05,0.2,0.35,0.50);
// percentage of profile pictures
$thresh_prof_pic = array(0.1,0.3,0.5,0.8);
```

Δημιουργία μοντέλου χρήστη βάση των ορίων

```
// check if user exists
$check_user = "SELECT * FROM `fb_user_vector` WHERE `usr_id` =
'$usr_id'";

$res1 = mysqli_query($db,$check_user);
if($row = mysqli_fetch_assoc($res1)){
    // user already has a vector - update values
    $update_vector = "UPDATE `fb_user_vector` SET `active` =
'$per_active', `checkins` = '$per_checkins', `events` = '$per_events', `friends` = '$per_friends',
```

```

`likes` = '$per_likes', `links` = '$per_links', `page_type` = '$per_page_type', `pages` =
'$per_pages', `prof_pic` = '$per_prof_pif', `self_tag` = '$per_self', `status` = '$per_status', `tags` =
'$per_tags', `user_checkins` = '$per_u_checkins', `user_events` = '$per_u_events' WHERE
`usr_id` = '$usr_id';

        //echo $update_vector;
        $vector_run = mysqli_query($db,$update_vector);
        $update_vector2 = "UPDATE `fb_user_vector_numeric` SET
`active` = '$num_active', `checkins` = '$num_checkins', `events` = '$num_events_att', `friends` =
'$num_friends', `likes` = '$num_likes', `links` = '$num_links', `page_type` =
'$num_page_type_ratio', `pages` = '$num_pages', `prof_pic` = '$user_prof_pic_ratio', `self_tag`
= '$user_auto_tag', `status` = '$num_status', `tags` = '$user_tags', `user_checkins` =
'$num_u_checkins', `user_events` = '$num_u_events' WHERE `usr_id` = '$usr_id';
        $vector_run = mysqli_query($db,$update_vector2);
    }else{
        // add user
        $add_vector = "INSERT INTO `fb_user_vector`(`usr_id`, `active`,
`checkins`, `events`, `friends`, `likes`, `links`, `page_type`, `pages`, `prof_pic`, `self_tag`,
`status`, `tags`, `user_checkins`, `user_events`) VALUES ('$usr_id', '$per_active',
'$per_checkins', '$per_events', '$per_friends', '$per_likes', '$per_links', '$per_page_type',
'$per_pages', '$per_prof_pif', '$per_self', '$per_status', '$per_tags', '$per_u_checkins',
'$per_u_events')";

        //echo $add_vector;
        $vector_run = mysqli_query($db,$add_vector);
        $add_vector2 = "INSERT INTO
`fb_user_vector_numeric`(`usr_id`, `active`, `checkins`, `events`, `friends`, `likes`, `links`,
`page_type`, `pages`, `prof_pic`, `self_tag`, `status`, `tags`, `user_checkins`, `user_events`)
VALUES ('$usr_id', '$num_active', '$num_checkins', '$num_events_att', '$num_friends',
'$num_likes', '$num_links', '$num_page_type_ratio', '$num_pages', '$user_prof_pic_ratio',
'$user_auto_tag', '$num_status', '$user_tags', '$num_u_checkins', '$num_u_events')";
        $vector_run = mysqli_query($db,$add_vector2);
    }
}

```

Παράρτημα Γ.

Τιμές βαρών σημασίας των δεδομένων δραστηριότητας σε κάθε διάσταση

```
// *** personality traits weights ***  
// extraversion weights  
$extr_friends = 0.1;  
$extr_likes = 0.1;  
$extr_u_checkins = 0.2;  
$extr_status = 0.2;  
$extr_prof = 0.1;  
$extr_u_events = 0.1;  
$extr_self_tag = 0.2;  
  
// agreeable weights  
$agr_tags = 0.2;  
$agr_likes = 0.1;  
$agr_active_friends = 0.3;  
$agr_checkins = 0.2;  
$agr_links = 0.2;  
  
// conscientiousness weights  
$cont_active_friends = 0.2;  
$cont_status = 0.1;  
$cont_likes = 0.1;  
$cont_events = 0.1;  
$cont_tags = 0.2;  
$cont_page_types = 0.1;  
$cont_links = 0.1;  
$cont_pages = 0.1;  
  
// neuroticism weights  
$neuro_likes = 0.3;  
$neuro_friends = 0.2;  
$neuro_prof = 0.2;  
$neuro_pages = 0.3;  
  
// openness weights  
$open_status = 0.3;  
$open_likes = 0.1;  
$open_events = 0.2;  
$open_pages = 0.1;  
$open_page_types = 0.3;
```

Δημιουργία μοντέλου προσωπικότητας με τη χρήση των βαρών σημασίας καθώς επίσης και το χώρισμα των δεδομένων που γίνεται σε κάθε διάσταση

```
// counter variables for the big 5 traits
$extravert = 0;
$agreeable = 0;
$conscientious = 0;
$neurotism = 0;
$open = 0;

$query = "SELECT * FROM `fb_user_vector` WHERE `usr_id` = '$usr_id'";
$run = mysqli_query($db,$query);
if($row = mysqli_fetch_assoc($run)){
    // *** vector value fetch ***
    $per_active = $row['active'];
    $per_checkins = $row['checkins'];
    $per_events = $row['events'];
    $per_friends = $row['friends'];
    $per_likes = $row['likes'];
    $per_links = $row['links'];
    $per_page_type = $row['page_type'];
    $per_pages = $row['pages'];
    $per_prof_pif = $row['prof_pic'];
    $per_self = $row['self_tag'];
    $per_status = $row['status'];
    $per_tags = $row['tags'];
    $per_u_checkins = $row['user_checkins'];
    $per_u_events = $row['user_events'];
    /*
    // *** calculate big 5 personality traits without weights ***
    // *** extraversion calculation ***
    // number of friends is positively correlated
    $extravert += $per_friends;
    // number of statuses is positively correlated
    $extravert += $per_status;
    // number of likes is positively correlated
    $extravert += $per_likes;
    // number of user check-ins is positively correlated
    $extravert += $per_u_checkins;
    // number of user events is positively correlated
    $extravert += $per_u_events;
    // self tagging is positively correlated
    $extravert += $per_self;
    // number of profile picture ratio is positively correlated
    $extravert += $per_prof_pif;
```

```

// *** agreeableness calculation ***
// number of tags is positively correlated
$agreeable += $per_tags;
// number of likes is negatively correlated
$agreeable += (5 - $per_likes + 1);
// number of active friends is negatively correlated
$agreeable += (5 - $per_active + 1);
// number of checkins is negatively correlated
$agreeable += (5 - $per_checkins + 1);
// number of links is positively correlated
$agreeable += $per_links;

// *** conscientiousness calculation ***
// number of active friends is negatively correlated
$conscientious += (5 - $per_active + 1);
// number of statuses is negatively correlated
$conscientious += (5 - $per_status + 1);
// number of likes is negatively correlated
$conscientious += (5 - $per_likes + 1);
// number of events attended is negatively correlated
$conscientious += (5 - $per_events + 1);
// number of tags is negatively correlated
$conscientious += (5 - $per_tags + 1);
// number of favorite page type ratio is positively correlated
$conscientious += $per_page_type;
// number of links is negatively correlated
$conscientious += (5 - $per_links + 1);
// number of pages is negatively correlated
$conscientious += (5 - $per_pages + 1);

// *** neuroticism calculation ***
// number of likes is positively correlated
$neurotism += $per_likes;
// number of friends is negatively correlated
$neurotism += (5 - $per_friends + 1);
// number of profile pictures ratio is negatively correlated
$neurotism += (5 - $per_prof_pif + 1);
// number of pages is positively correlated
$neurotism += $per_pages;

// *** openness calculation ***
// number of statuses is positively correlated
$open += $per_status;
// number of likes is positively correlated
$open += $per_likes;

```

```

// number of events attended is positively correlated
$open += $per_events;
// number of pages is positively correlated
$open += $per_pages;
// number of favorite page type ratio is negatively correlated
$open += (5 - $per_page_type + 1);

// *** calculate the percentages ***
$total = $extravert + $agreeable + $consientious + $neurotism + $open;
$extravert = $extravert * 100 / $total;
$agreeable = $agreeable * 100 / $total;
$consientious = $consientious * 100 / $total;
$neurotism = $neurotism * 100 / $total;
$open = $open * 100 / $total;

// *** calculate big 5 personality traits with weights ***
// *** extraversion calculation ***
// number of friends is positively correlated
$extravert += $extr_friends * $per_friends;
// number of statuses is positively correlated
$extravert += $extr_status * $per_status;
// number of likes is positively correlated
$extravert += $extr_likes * $per_likes;
// number of user check-ins is positively correlated
$extravert += $extr_u_checkins * $per_u_checkins;
// number of user events is positively correlated
$extravert += $extr_u_events * $per_u_events;
// self tagging is positively correlated
$extravert += $extr_self_tag * $per_self;
// number of profile picture ratio is positively correlated
$extravert += $extr_prof * $per_prof_pif;

// *** agreeableness calculation ***
// number of tags is positively correlated
$agreeable += $agr_tags * $per_tags;
// number of likes is negatively correlated
$agreeable += $agr_likes * (5 - $per_likes + 1);
// number of active friends is negatively correlated
$agreeable += $agr_active_friends * (5 - $per_active + 1);
// number of checkins is negatively correlated
$agreeable += $agr_checkins * (5 - $per_checkins + 1);
// number of links is positively correlated
$agreeable += $agr_links * $per_links;

// *** consientiousness calculation ***
// number of active friends is negatively correlated

```

```

$consientious += $cont_active_friends * (5 - $per_active + 1);
// number of statuses is negatively correlated
$consientious += $cont_status * (5 - $per_status + 1);
// number of likes is negatively correlated
$consientious += $cont_likes * (5 - $per_likes + 1);
// number of events attended is negatively correlated
$consientious += $cont_events * (5 - $per_events + 1);
// number of tags is negatively correlated
$consientious += $cont_tags * (5 - $per_tags + 1);
// number of favorite page type ratio is positively correlated
$consientious += $cont_page_types * $per_page_type;
// number of links is negatively correlated
$consientious += $cont_links * (5 - $per_links + 1);
// number of pages is negatively correlated
$consientious += $cont_pages * (5 - $per_pages + 1);

// *** neuroticism calculation ***
// number of likes is positively correlated
$neurotism += $neuro_likes * $per_likes;
// number of friends is negatively correlated
$neurotism += $neuro_friends * (5 - $per_friends + 1);
// number of profile pictures ratio is negatively correlated
$neurotism += $neuro_prof * (5 - $per_prof_pif + 1);
// number of pages is positively correlated
$neurotism += $neuro_pages * $per_pages;

// *** openness calculation ***
// number of statuses is positively correlated
$open += $open_status * $per_status;
// number of likes is positively correlated
$open += $open_likes * $per_likes;
// number of events attended is positively correlated
$open += $open_events * $per_events;
// number of pages is positively correlated
$open += $open_pages * $per_pages;
// number of favorite page type ratio is negatively correlated
$open += $open_page_types * (5 - $per_page_type + 1);

// *** calculate the percentages ***
$total = $extravert + $agreeable + $consientious + $neurotism + $open;
$extravert = $extravert * 100 / $total;
$agreeable = $agreeable * 100 / $total;
$consientious = $consientious * 100 / $total;
$neurotism = $neurotism * 100 / $total;
$open = $open * 100 / $total;
*/

```

```

// *** calculate big 5 personality traits with weights and max 1 for each trait ***
// *** extraversion calculation ***
// number of friends is positively correlated
$extravert += $extr_friends * ($per_friends / 5.0);
// number of statuses is positively correlated
$extravert += $extr_status * ($per_status / 5.0);
// number of likes is positively correlated
$extravert += $extr_likes * ($per_likes / 5.0);
// number of user check-ins is positively correlated
$extravert += $extr_u_checkins * ($per_u_checkins / 5.0);
// number of user events is positively correlated
$extravert += $extr_u_events * ($per_u_events / 5.0);
// self tagging is positively correlated
$extravert += $extr_self_tag * ($per_self / 5.0);
// number of profile picture ratio is positively correlated
$extravert += $extr_prof * ($per_prof_pif / 5.0);

// *** agreeableness calculation ***
// number of tags is positively correlated
$agreeable += $agr_tags * ($per_tags / 5.0);
// number of likes is negatively correlated
$agreeable += $agr_likes * ((5 - $per_likes + 1) / 5.0);
// number of active friends is negatively correlated
$agreeable += $agr_active_friends * ((5 - $per_active + 1) / 5.0);
// number of checkins is negatively correlated
$agreeable += $agr_checkins * ((5 - $per_checkins + 1) / 5.0);
// number of links is positively correlated
$agreeable += $agr_links * ($per_links / 5.0);

// *** conscientiousness calculation ***
// number of active friends is negatively correlated
$conscientious += $cont_active_friends * ((5 - $per_active + 1) / 5.0);
// number of statuses is negatively correlated
$conscientious += $cont_status * ((5 - $per_status + 1) / 5.0);
// number of likes is negatively correlated
$conscientious += $cont_likes * ((5 - $per_likes + 1) / 5.0);
// number of events attended is negatively correlated
$conscientious += $cont_events * ((5 - $per_events + 1) / 5.0);
// number of tags is negatively correlated
$conscientious += $cont_tags * ((5 - $per_tags + 1) / 5.0);
// number of favorite page type ratio is positively correlated
$conscientious += $cont_page_types * ($per_page_type / 5.0);
// number of links is negatively correlated
$conscientious += $cont_links * ((5 - $per_links + 1) / 5.0);
// number of pages is negatively correlated
$conscientious += $cont_pages * ((5 - $per_pages + 1) / 5.0);

```

```

// *** neuroticism calculation ***
// number of likes is positively correlated
$neurotism += $neuro_likes * ($per_likes / 5.0);
// number of friends is negatively correlated
$neurotism += $neuro_friends * ((5 - $per_friends + 1) / 5.0);
// number of profile pictures ratio is negatively correlated
$neurotism += $neuro_prof * ((5 - $per_prof_pif + 1) / 5.0);
// number of pages is positively correlated
$neurotism += $neuro_pages * ($per_pages / 5.0);

// *** openness calculation ***
// number of statuses is positively correlated
$open += $open_status * ($per_status / 5.0);
// number of likes is positively correlated
$open += $open_likes * ($per_likes / 5.0);
// number of events attended is positively correlated
$open += $open_events * ($per_events / 5.0);
// number of pages is positively correlated
$open += $open_pages * ($per_pages / 5.0);
// number of favorite page type ratio is negatively correlated
$open += $open_page_types * ((5 - $per_page_type + 1) / 5.0);

// *** calculate the percentages ***
$total = $extravert + $agreeable + $conscientious + $neurotism + $open;
$extravert = $extravert * 100 / $total;
$agreeable = $agreeable * 100 / $total;
$conscientious = $conscientious * 100 / $total;
$neurotism = $neurotism * 100 / $total;
$open = $open * 100 / $total;

if($extravert > $agreeable){
    if($extravert > $conscientious){
        if($extravert > $neurotism){
            if($extravert > $open){
                $stop = "Extravert";
            }else{
                $stop = "Open";
            }
        }else{
            if($neurotism > $open){
                $stop = "Neurotism";
            }else{
                $stop = "Open";
            }
        }
    }
}

```

```
}else{
```

```

    }
}

// *** check if the user is in the system and insert or update him***
$query2 = "SELECT * FROM `fb_user_model` WHERE `usr_id` = '$usr_id'";
$run2 = mysqli_query($db,$query2);
if($row2 = mysqli_fetch_assoc($run2)){
    $update_query = "UPDATE `fb_user_model` SET `Extraversion` = '$extravert',
`Agreeableness` = '$agreeable', `Conscientiousness` = '$conscientious', `Neuroticism` =
'$neurotism', `Openness` = '$open' WHERE `usr_id` = '$usr_id'";
    $exec = mysqli_query($db,$update_query);
}else{
    $add_query = "INSERT INTO `fb_user_model`(`usr_id`, `Extraversion`,
`Agreeableness`, `Conscientiousness`, `Neuroticism`, `Openness`) VALUES ('$usr_id',
'$extravert', '$agreeable', '$conscientious', '$neurotism', '$open')";
    $exec = mysqli_query($db,$add_query);
}

```

Παράρτημα Δ.

Ερωτήματα προς τη βάση δεδομένων του Facebook

<pre>SELECT uid, name, email, birthday_date, sex, pic_small FROM user WHERE uid = me()</pre>	Ερώτημα το οποίο επιστρέφει τα αρχικά δεδομένα του χρήστη για την δημιουργία του λογαριασμού του στην εφαρμογή και στο PersonaWeb.
<pre>SELECT uid2 FROM friend WHERE uid1 = me()</pre>	Ερώτημα για την ανακτήσει των 'friends' του χρήστη στο Facebook.
<pre>SELECT object_id, object_type FROM like WHERE user_id=me() LIMIT 5000</pre>	Ερώτημα ανάκτησης των 'likes' του χρήστη.
<pre>SELECT object_id,owner FROM album WHERE ('.\$albums.')</pre>	Ερωτήματα για την ανάκτηση των ιδιοκτητών (owners) των αντικειμένων που έκανε like ο τρέχων χρήστης του τύπου albums. Το αντικείμενο \$albums περιέχει τους κωδικούς των αντικειμένων τύπου album που έκανε like ο τρέχων χρήστης.
<pre>SELECT object_id,owner FROM photo WHERE ('.\$photos.')</pre> <pre>LIMIT 1000</pre>	Ερωτήματα για την ανάκτηση των ιδιοκτητών (owners) των αντικειμένων που έκανε like ο τρέχων χρήστης του τύπου photos. Το αντικείμενο \$photos περιέχει τους κωδικούς των αντικειμένων τύπου photo που έκανε like ο τρέχων χρήστης.
<pre>SELECT checkin_id, author_uid FROM checkin WHERE ('.\$checkins.')</pre>	Ερωτήματα για την ανάκτηση των δημιουργών (creators) των αντικειμένων που έκανε like ο τρέχων χρήστης του τύπου check-ins. Το αντικείμενο \$checkins περιέχει τους κωδικούς των αντικειμένων τύπου check-in που έκανε like ο τρέχων χρήστης.
<pre>SELECT vid,owner FROM video WHERE ('.\$videos.')</pre>	Ερωτήματα για την ανάκτηση των ιδιοκτητών (owners) των αντικειμένων που έκανε like ο τρέχων χρήστης του τύπου videos. Το αντικείμενο \$videos περιέχει τους κωδικούς των αντικειμένων τύπου video που έκανε like ο τρέχων χρήστης.
<pre>SELECT status_id,uid FROM status WHERE ('.\$statuss.')</pre>	Ερωτήματα για την ανάκτηση των ιδιοκτητών (owners) των αντικειμένων που έκανε like ο τρέχων χρήστης του τύπου

	status. Το αντικείμενο \$statuss περιέχει τους κωδικούς των αντικειμένων τύπου status που έκανε like ο τρέχον χρήστης.
SELECT link_id,owner FROM link WHERE ('.\$links.')	Ερωτήματα για την ανάκτηση των ιδιοκτητών (owners) των αντικειμένων που έκανε like ο τρέχων χρήστης του τύπου links. Το αντικείμενο \$links περιέχει τους κωδικούς των αντικειμένων τύπου link που έκανε like ο τρέχων χρήστης.
SELECT eid,creator FROM event WHERE ('.\$events.')	Ερωτήματα για την ανάκτηση των δημιουργών (creators) των αντικειμένων που έκανε like ο τρέχων χρήστης του τύπου events. Το αντικείμενο \$events περιέχει τους κωδικούς των αντικειμένων τύπου event που έκανε like ο τρέχων χρήστης.
SELECT pic FROM user WHERE uid = '\$like_owner_id'	Ερώτημα το οποίο επιστρέφει την εικόνα προφίλ του χρήστη στον οποίο ο τρέχων χρήστης έκανε τα περισσότερα likes. Η μεταβλητή \$like_owner_id κρατά το μοναδικό κωδικό του χρήστη στον οποίο ο τρέχων χρήστης έκανε τα περισσότερα likes.
SELECT page_id, type FROM page WHERE page_id IN(SELECT page_id,profile_section,type FROM page_fan WHERE uid = me())	Ερώτημα για την ανάκτηση των σελίδων του Facebook που αρέσουν του χρήστη. Αρχικά πρέπει να πάρουμε βασικές πληροφορίες για το χρήστη από το πίνακα page_fan για να χρησιμοποιήσουμε στο ερώτημα μας.
SELECT owner, vid FROM video WHERE vid IN (SELECT vid FROM video_tag WHERE subject = me())	Ερώτημα το οποίο επιστρέφει τα βίντεο και τον ιδιοκτήτη τους στα οποία είναι tagged ο τρέχων χρήστης.
SELECT owner, object_id FROM photo WHERE object_id IN (SELECT object_id FROM photo_tag WHERE subject = me() LIMIT 1000)	Ερώτημα το οποίο επιστρέφει τις φωτογραφίες και τον ιδιοκτήτη τους στα οποία είναι tagged ο τρέχων χρήστης.
SELECT checkin_id,author_uid FROM checkin WHERE (me() IN tagged_uids) AND (author_uid IN (SELECT uid2 FROM friend WHERE uid1 = me()) OR author uid = me())	Ερώτημα για την ανάκτηση των αντικειμένων τύπου check-ins στα οποία είναι tagged ο τρέχων χρήστης.
SELECT checkin_id	Ερώτημα για τη συλλογή των δεδομένων

FROM checkin WHERE author_uid = me()	των check-ins τα οποία δημιούργησε ο τρέχων χρήστης.
SELECT eid, creator, pic_big, name FROM event WHERE eid IN (SELECT eid FROM event_member WHERE uid = me() AND rsvp_status = 'attending' AND start_time < '\$time')	Ερώτημα για τη συλλογή των δεδομένων των εκδηλώσεων στις οποίες παραβρέθηκε ο χρήστης. Η συσχέτιση των χρηστών με τα events βρίσκεται στο πίνακα `event_member` τον οποίο πρέπει να χρησιμοποιήσουμε για να πάρουμε στοιχεία για τις συγκεκριμένες εκδηλώσεις που θέλουμε. Η μεταβλητή \$time κρατά την τρέχον ώρα και ημερομηνία που εκτελείται το ερώτημα σε μορφή Unix Timestamp.
SELECT eid FROM event WHERE creator = me()	Ερώτημα για τη συλλογή των δεδομένων των εκδηλώσεων τα οποία δημιούργησε ο τρέχων χρήστης.
SELECT link_id FROM link WHERE owner = me()	Ερώτημα για τη συλλογή των συνδέσμων του διαδικτύου τους οποίους κοινοποίησε ο τρέχων χρήστης.
SELECT object_id FROM photo WHERE owner = me()	Ερώτημα για τη συλλογή των φωτογραφιών που ανέβασε στο Facebook ο τρέχων χρήστης.
SELECT vid FROM video WHERE owner = me()	Ερώτημα για τη συλλογή των βίντεο που ανέβασε στο Facebook ο τρέχων χρήστης.
SELECT status_id FROM status WHERE uid = me()	Ερώτημα για τη συλλογή των καταστάσεων που κοινοποίησε ο τρέχων χρήστης.
SELECT post_id, actor_id, created_time FROM stream WHERE source_id = me() AND filter_key = 'others' AND actor_id != me() AND created_time > '\$q_time' LIMIT 1000	Ερώτημα το οποίο θα μας δώσει τις πληροφορίες για τους ενεργούς φίλους του χρήστη. Η μεταβλητή \$q_time κρατά την ώρα και ημερομηνία ένα χρόνο πριν από την στιγμή που εκτελέστηκε το ερώτημα.
SELECT post_id FROM stream_tag WHERE target_id = me() AND actor_id = me()	Ερώτημα για την ανάκτηση της πληροφορίας αν ο χρήστης έκανε tag τον εαυτό του σε κάποιο αντικείμενο ή δημοσίευση.
SELECT object_id, size FROM album WHERE owner = me() AND name = 'Profile Pictures'	Ερώτημα για να μάθουμε το μέγεθος του άλμπουμ του χρήστη με τις φωτογραφίες προφίλ του.
SELECT pic FROM user WHERE uid = '\$tag_owner_id'	Ερώτημα το οποίο επιστρέφει την εικόνα προφίλ του χρήστη στον οποίο ο τρέχων χρήστης είχε τα περισσότερα tags. Η μεταβλητή \$tag_owner_id κρατά το μοναδικό κωδικό του χρήστη στον οποίο ο

	τρέχον χρήστης είχε τα περισσότερα tags.
SELECT pic FROM user WHERE uid = '\$temp_user'	Ερώτημα το οποίο επιστρέφει την εικόνα προφίλ του χρήστη τον οποίο θα δείξουμε στο τρέχον χρήστης ότι έχουν όμοια δραστηριότητα στο κοινωνικό δίκτυο Facebook. Η μεταβλητή \$temp_user κρατά το μοναδικό κωδικό του χρήστη τον οποίο θα δείξουμε στο τρέχον χρήστης ότι έχουν όμοια δραστηριότητα στο κοινωνικό δίκτυο Facebook.
SELECT page_id, name, page_url, pic_big FROM page WHERE ('\$suggestion_url.')	Ερώτημα το οποίο επιστρέφει στοιχεία για τις σελίδες στο Facebook τις οποίες θα προτείνουμε στον τρέχον χρήστη. Η μεταβλητή \$suggestion_url κρατά τους μοναδικούς κωδικούς των σελίδων τις οποίες θα προτείνουμε στο τρέχον χρήστη.