

ΙΟΝΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ

ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ



Πτυχιακή Εργασία

Ανάπτυξη Εφαρμογής Καταγραφής Ερευνητικής
Δραστηριότητας Τμήματος Πληροφορικής Ιονίου
Πανεπιστημίου

Κοντίνης Ιωάννης

Επιβλέπων: Αυλωνίτης Μάρκος

Κέρκυρα, Φεβρουάριος 2019

Επιβλέπων Καθηγητής

Αυλωνίτης Μάρκος

Επίκουρος Καθηγητής, Ιόνιο Πανεπιστήμιο, Τμήμα Πληροφορικής

Τριμελής Επιτροπή

Αυλωνίτης Μάρκος

Επίκουρος Καθηγητής, Ιόνιο Πανεπιστήμιο, Τμήμα Πληροφορικής

Στεφανιδάκης Μιχαήλ

Επίκουρος Καθηγητής, Ιόνιο Πανεπιστήμιο, Τμήμα Πληροφορικής

Πανάρετος Αλέξανδρος

Εργαστηριακό Διδακτικό Προσωπικό (ΕΔΙΠ), Ιόνιο Πανεπιστήμιο, Τμήμα Πληροφορικής

Ευχαριστίες

Η παρούσα πτυχιακή εργασία δεν θα είχε ολοκληρωθεί χωρίς τη βοήθεια και τη συμπαράσταση ορισμένων ανθρώπων. Ως ελάχιστη δυνατή μνεία, οφείλω να ευχαριστήσω θερμά όλους όσους συνέβαλαν στην εκπόνηση της και ιδιαίτερα:

- Τον κύριο *Αλέξανδρο Πανάρετο* μέλος του Εργαστηριακού και Διδακτικού Προσωπικού (ΕΔΙΠ), του Τμήματος Πληροφορικής του Ιονίου Πανεπιστημίου, για τις πολύτιμες συμβουλές του και τη σημαντική καθοδήγηση που μου παρείχε σε όλη τη διάρκεια της ενασχόλησής μου με την πτυχιακή εργασία.
- Τον κύριο *Μάρκο Αυλωνίτη* Επίκουρο Καθηγητή του Τμήματος Πληροφορικής του Ιονίου Πανεπιστημίου για την εμπιστοσύνη που μου έδειξε, δεχόμενος να είναι ο επιβλέπων της πτυχιακής εργασίας.
- Τον κύριο *Μιχάλη Στεφανιδάκη* Επίκουρο Καθηγητή του Τμήματος Πληροφορικής του Ιονίου Πανεπιστημίου και μέλος της τριμελούς επιτροπής στον οποίο οφείλω τις γνώσεις που αποκόμισα από την παρακολούθηση των μαθημάτων του σχετικά με την ανάλυση δεδομένων.
- Τον εκλεκτό φίλο *Σταμάτη Πουλιάση*, Διδάκτωρ και ερευνητή Μαθηματικών, για τις πολύ εποικοδομητικές συζητήσεις μας μέσα από τις οποίες κατάλαβα πως πραγματοποιείται η διαδικασία δημοσίευσης μιας ερευνητικής εργασίας.
- *Την οικογένειά μου* για τη στήριξη και τη συμπαράσταση που μου παρείχε σε όλη τη διάρκεια των σπουδών μου.

Κέρκυρα, Φεβρουάριος 2019
Γιάννης Κοντίνης

Περιεχόμενα

1	Εισαγωγή	1
1.1	Σχετικά με την πτυχιακή εργασία	1
1.2	Η δομή του κειμένου	2
2	Βιβλιομετρικά δεδομένα - Βιβλιομετρικοί δείκτες	5
2.1	Βάσεις δεδομένων επιστημονικών δημοσιεύσεων	5
2.2	Δείκτες βιβλιομετρικής ανάλυσης	8
2.3	Επιλογή διαδικτυακού τόπου άντλησης δεδομένων	12
3	Η βάση δεδομένων της εφαρμογής	15
3.1	Ο καθορισμός των απαιτήσεων	15
3.2	Το Διάγραμμα Οντοτήτων – Συσχετίσεων	17
3.3	Το σχεσιακό σχήμα της βάσης δεδομένων	23
3.4	Κατασκευή της βάσης δεδομένων	28
3.5	Δημιουργία πινάκων της βάσης δεδομένων	30
3.6	Σύνδεση πινάκων και επιβολή ακεραιότητας αναφοράς	34
4	Η ανάλυση (parsing) των δεδομένων	39
4.1	Αρχεία csv	39
4.2	Το αρχείο scopus.csv	39
4.3	Η Ανάλυση (parsing) του αρχείου scopus.csv	44
4.4	Προβλήματα και λύσεις κατά την ανάλυση	53
5	Παρουσίαση της εφαρμογής	59
5.1	Παρουσίαση του τμήματος για τον διαχειριστή του συστήματος	59
5.2	Παρουσίαση του τμήματος για τους χρήστες του συστήματος .	70
5.2.1	Ιστοσελίδα All Research	72
5.2.2	Ιστοσελίδα Professors	80
5.2.3	Ιστοσελίδα All years	85
5.2.4	Ιστοσελίδα Timeline	86
5.2.5	Ιστοσελίδα Journals	88
5.2.6	Ιστοσελίδα Conferences	91

6	Μελλοντικές επεκτάσεις και σύνοψη	93
6.1	Μελλοντικές επεκτάσεις	93
6.2	Σύνοψη και τελικός απολογισμός	94
	A' Το αρχείο Database.php	97
	Βιβλιογραφία	103

Κατάλογος Εικόνων

2.1	Γραφικός υπολογισμός του h – index	11
3.1	Το Διάγραμμα Οντοτήτων – Συσχετίσεων	18
3.2	Οι σχέσεις που προκύπτουν απο την μετατροπή των κύριων και δευτερευουσών τάξεων σε σχεσιακό σχήμα	24
3.3	Οι σχέσεις που προκύπτουν απο τις ισχυρές οντότητες	24
3.4	Οι σχέσεις που προκύπτουν απο τα γνωρίσματα πολλαπλών τιμών	25
3.5	Η σχέση Persons μετά την προσθήκη ξένου κλειδιού	25
3.6	Οι σχέσεις Journal_articles, Conference_articles και Confe- rences όπως προκύπτουν από την εφαρμογή του τέταρτου βήματος	27
3.7	Η σχέση Authorships που προκύπτει από την εφαρμογή του πέμπτου βήματος	27
3.8	Το τελικό σχεσιακό σχήμα της βάσης δεδομένων	28
3.9	Ο πίνακας ελέγχου του XAMPP με τον Apache και την My- SQL σε λειτουργία	28
3.10	Η αρχική οθόνη του phpMyAdmin	29
3.11	Το μενού της αρχικής οθόνης του phpMyAdmin	30
3.12	Δημιουργία της βάσης δεδομένων	30
3.13	Δημιουργία πίνακα	30
3.14	Ορισμός πεδίων πίνακα	31
3.15	Ο ορισμός των πεδίων του πίνακα Administrator	32
3.16	Ο ορισμός των πεδίων του πίνακα Authorships	32
3.17	Ο ορισμός των πεδίων του πίνακα Books	32
3.18	Ο ορισμός των πεδίων του πίνακα Book_chapters	32
3.19	Ο ορισμός των πεδίων του πίνακα Conferences	32
3.20	Ο ορισμός των πεδίων του πίνακα Conference_articles	33
3.21	Ο ορισμός των πεδίων του πίνακα Journals	33
3.22	Ο ορισμός των πεδίων του πίνακα Journal_articles	33
3.23	Ο ορισμός των πεδίων του πίνακα Persons	33
3.24	Ο ορισμός των πεδίων του πίνακα Proceedings	33
3.25	Ο ορισμός των πεδίων του πίνακα Publications	33
3.26	Δημιουργία ευρετηρίου στον πίνακα prof_names	34
3.27	Ο ορισμός των επιλόγων του ευρετηρίου	34

3.28 Δήλωση της αναφοράς του ξένου κλειδιού	35
3.29 Οι πίνακες της βάσης δεδομένων και οι συνδέσεις μεταξύ τους.	37
4.1 Ένα αρχείο scopus.csv	40
4.2 Ένα τμήμα των δεδομένων του αρχείου scopus.csv σε μορφή διασυνδεδεμένου πίνακα	48
5.1 Η φόρμα εισόδου του διαχειριστή στο σύστημα	60
5.2 Η αρχική σελίδα του διαχειριστή	60
5.3 Η ιστοσελίδα αναζήτησης συγγραφέα στο Scopus	61
5.4 Τα αποτελέσματα της αναζήτησης	62
5.5 Ο σύνδεσμος που οδηγεί στην εξαγωγή των αποτελεσμάτων	62
5.6 Η επιλογή των πεδίων που θα περιέχει το αρχείο csv.	63
5.7 Η εξαγωγή των αποτελεσμάτων σε αρχείο csv	63
5.8 Αποθήκευση του αρχείου csv	64
5.9 Η φόρμα εγγραφής νέου μέλους	64
5.10 Πληροφορίες για έναν ερευνητή στο Scopus	65
5.11 Μέρος του πίνακα με τα καταχωρημένα μέλη ΔΕΠ	65
5.12 Ο πίνακας με τα ονόματα του μέλους ΔΕΠ χωρίς δεδομένα	66
5.13 Προσθήκη ονόματος	66
5.14 Ο πίνακας με τα ονόματα του μέλους ΔΕΠ	67
5.15 Το παράθυρο επιβεβαίωσης εκκίνησης της ανάλυσης	67
5.16 Η μπάρα προόδου της διαδικασίας	68
5.17 Το μήνυμα στο τέλος της διαδικασίας	68
5.18 Ο πίνακας με τις δημοσιεύσεις ενός μέλους ΔΕΠ	69
5.19 Επιλογή των δημοσιεύσεων προς διαγραφή	69
5.20 Το παράθυρο επιβεβαίωσης της διαγραφής	70
5.21 Πραγματοποίηση της διαγραφής	70
5.22 Η αρχική σελίδα του τμήματος των χρηστών του συστήματος	71
5.23 Πληροφορίες για τον αριθμό των δημοσιεύσεων ανα τύπο	72
5.24 Ο πίνακας με τα στοιχειά όλων των δημοσιεύσεων	73
5.25 Αναζήτηση αποτελεσμάτων	74
5.26 Εμφάνιση επιπλέον πληροφοριών με επέκταση της γραμμής	75
5.27 Το πλήθος δημοσιεύσεων κάθε μέλους ΔΕΠ	76
5.28 Γράφημα συνολικού αριθμού αναφορών ανα μέλος	77
5.29 Γράφημα μέσου αριθμού αναφορών ανα μέλος	78
5.30 Γράφημα του h-index όλων των μελών	78
5.31 Το δίκτυο των συνεργασιών μεταξύ των μελών	79
5.32 Ένα τμήμα της ιστοσελίδας Professors	80
5.33 Συνοπτικές πληροφορίες ερευνητικού προφίλ	81
5.34 Ο πίνακας όλων των δημοσιεύσεων ενός μέλους	81
5.35 Το γράφημα υπολογισμού του h-index	82
5.36 Ο πίνακας με τους συνεργαζόμενους συγγραφείς	83
5.37 Οι κοινές δημοσιεύσεις δυο ερευνητών	84

5.38 Δημοσιεύσεις ενός συγκεκριμένου έτους	84
5.39 Γράφημα αριθμού δημοσιεύσεων ενός μέλους ανά έτος	85
5.40 Ένα τμήμα της ιστοσελίδας All years	86
5.41 Ο πίνακας της ιστοσελίδας Timeline για το έτος 2017	87
5.42 Γράφημα πλήθους δημοσιεύσεων ανά έτος	88
5.43 Ο πίνακας της ιστοσελίδας Journals	89
5.44 Τα άρθρα που έχουν δημοσιευτεί σε ένα περιοδικό	90
5.45 Επιπλέον πληροφορίες με την επέκταση της γραμμής	90

Κατάλογος Πινάκων

3.1	Οι επιλογές των ξένων κλειδιών του πίνακα Authorsips	36
3.2	Οι επιλογές των ξένων κλειδιών του πίνακα Books	36
3.3	Οι επιλογές των ξένων κλειδιών του πίνακα Book_chapters .	36
3.4	Οι επιλογές των ξένων κλειδιών του πίνακα Conferences . . .	36
3.5	Οι επιλογές ξένων κλειδιών του πίνακα Conference_articles .	36
3.6	Οι επιλογές των ξένων κλειδιών του πίνακα Journal_articles .	37
3.7	Οι επιλογές των ξένων κλειδιών του πίνακα Persons	37
3.8	Οι επιλογές των ξένων κλειδιών του πίνακα prof_names . . .	37
4.1	Τα πεδία της κεφαλίδας του αρχείου scopus.csv που χρησιμο- ποιούνται στην εφαρμογή	45

Κατάλογος Τμημάτων Κώδικα

4.1	parceCSV.php γραμμές 1-8	45
4.2	parceCSV.php γραμμές 9-31	45
4.3	parceCSV.php γραμμές 70-77	47
4.4	parceCSV.php γραμμές 78-189	48
4.5	parceCSV.php γραμμές 190-200	52
4.6	parceCSV.php γραμμές 32-69	55

Κεφάλαιο 1

Εισαγωγή

1.1 Σχετικά με την πτυχιακή εργασία

Η πτυχιακή εργασία για την οποία γίνεται λόγος σε αυτό το κείμενο πραγματοποιήθηκε στα πλαίσια εκπόνησης υποχρεωτικής πτυχιακής εργασίας του Τμήματος Πληροφορικής του Ιονίου Πανεπιστημίου. Μέσα από πληθώρα διαθέσιμων θεωρητικών (βιβλιογραφικές εργασίες) και πρακτικών θεμάτων (εργασίες ανάπτυξης), επιλέχτηκε το αντικείμενο της πτυχιακής εργασίας να είναι η ανάπτυξη μιας διαδικτυακής εφαρμογής η οποία παρουσιάζει την ερευνητική δραστηριότητα των μελών ΔΕΠ του Τμήματος Πληροφορικής του Ιονίου Πανεπιστημίου.

Ένας σημαντικός λόγος που οδήγησε στην συγκεκριμένη επιλογή ήταν η ανάγκη της δημιουργίας ενός συστήματος συγκέντρωσης του ερευνητικού έργου του Τμήματος Πληροφορικής το οποίο θα είναι δυνατόν να χρησιμοποιηθεί μελλοντικά ως ένα εργαλείο αξιολόγησης αλλά και για την εξαγωγή στατιστικών στοιχείων και συμπερασμάτων που αφορούν την έρευνα που έχει συντελεστεί στο Τμήμα και στο Ιόνιο Πανεπιστήμιο. Επίσης η διαδικτυακή εφαρμογή που υλοποιήθηκε για τις ανάγκες αυτής της πτυχιακής εργασίας έχει ως στόχο να καλύψει, εν μέρει, ένα κενό το οποίο πιστεύουμε ότι υπάρχει σχετικά με τέτοιου είδους συστήματα στο Ιόνιο Πανεπιστήμιο, αλλά και να αποτελέσει την απαρχή της δημιουργίας, στο μέλλον, ενός ολοκληρωμένου συστήματος παρουσίασης και διαχείρισης της ερευνητικής δραστηριότητας του Ιδρύματος. Επιπλέον ο όγκος, αλλά και η δυσκολία του συγκεκριμένου εγχειρήματος σε σχέση με ό,τι άλλο είχε επιχειρηθεί σε αντίστοιχα ζητήματα μέχρι τώρα ήταν από μόνα τους μία πρόκληση.

Εδώ πρέπει να τονιστεί ότι ο πραγματικός σκοπός ανάπτυξης αυτής της εφαρμογής ως αντικείμενο εκπόνησης πτυχιακής εργασίας δεν είναι ασφαλώς η δημιουργία ενός συστήματος πλούσιων δυνατοτήτων απαλλαγμένου από λάθη αλλά η ευκαιρία που παρουσιάζεται για ενασχόληση σε βάθος με τις βασικότερες πτυχές μιας διαδικασίας δημιουργίας ολοκληρωμένης εφαρμογής που θα προσφέρει σημαντική εμπειρία σε μελλοντικές προσπάθειες. Έτσι, μέσα α-

πό την προσπάθεια αυτή δόθηκε ευκαιρία για βαθύτερη κατανόηση, αλλά και απόκτηση νέων γνώσεων πάνω σε πρακτικά ζητήματα του κλάδου της Πληροφορικής όπως τα επόμενα:

- Η κατασκευή ιστοσελίδων και βάσεων δεδομένων, καθώς και η μεταξύ τους επικοινωνία.
- Η υλοποίηση εφαρμογών ιστού με γλώσσες προγραμματισμού όπως η PHP, η MySQL, η HTML, η CSS και η JavaScript.
- Η χρήση του framework Bootstrap, που συνδυάζει τις τρεις τελευταίες γλώσσες προγραμματισμού για την ανάπτυξη ευέλικτων διδακτικών τοποθεσιών.
- Η τεχνολογία AJAX που συντελεί στην δημιουργία πιο γρήγορων, διαδραστικών και φιλικών προς τον χρήστη εφαρμογών ιστού.
- Η ενασχόληση με τις JavaScript βιβλιοθήκες οπτικοποίησης δεδομένων d3.js και Google Charts.

1.2 Η δομή του κειμένου

Στο κεφάλαιο 2 γίνεται αναφορά στους διαδικτυακούς τόπους και τις βάσεις δεδομένων που περιέχουν πληροφορίες σχετικές με την ερευνητική δραστηριότητα ατόμων και οργανισμών. Επιπλέον, αναφέρονται οι σημαντικότεροι δείκτες ανάλυσης και αξιολόγησης των δεδομένων ερευνητικής δραστηριότητας. Ακόμη παρουσιάζονται οι λόγοι οι οποίοι οδήγησαν στην επιλογή ενός συγκεκριμένου διαδικτυακού τόπου για την άντληση των δεδομένων που θα επεξεργάζεται η εφαρμογή.

Στο κεφάλαιο 3 γίνεται ο καθορισμός των απαιτήσεων της εφαρμογής και αναλύεται ο σχεδιασμός του Διαγράμματος Οντοτήτων – Συσχετίσεων από το οποίο κατασκευάζεται το σχεσιακό σχήμα της βάσης δεδομένων που θα χρησιμοποιεί η εφαρμογή. Εν συνεχεία, με βάση το σχεσιακό σχήμα παρουσιάζεται αναλυτικά ο τρόπος κατασκευής της βάσης δεδομένων στο περιβάλλον phpMyAdmin του πακέτου προγραμμάτων λογισμικού XAMPP. Το τέλος του κεφαλαίου αναφέρεται στην δημιουργία των πινάκων της βάσης, την μεταξύ τους σύνδεση και την επιβολή ακεραιότητας αναφορών.

Στο κεφάλαιο 4 επεξηγείται αναλυτικά ο κώδικας PHP με την βοήθεια του οποίου έγινε ανάλυση των δεδομένων που συλλέχτηκαν σε αρχεία csv από το διαδίκτυο και η εισαγωγή τους στην βάση δεδομένων της εφαρμογής. Επιπλέον γίνεται λόγος για τα προβλήματα που πρόεκυψαν κατά την διάρκεια της υλοποίησης του κώδικα αυτού, καθώς και για τους τρόπους που ξεπεράστηκαν προγραμματιστικά τα προβλήματα αυτά.

Στο κεφάλαιο 5 γίνεται παρουσίαση καθενός από τα δυο τμήματα από τα οποία αποτελείται η εφαρμογή καθώς και των λειτουργιών που τα τμήματα αυτά επιτελούν.

Το κεφάλαιο 6 αναφέρεται στις μελλοντικές επεκτάσεις που προτείνονται για την αναβάθμιση της εφαρμογής και την ποιοτικότερη λειτουργικότητα της. Επίσης πραγματοποιείται η σύνοψη και απολογισμός της προσπάθειας από την εκπόνηση της πτυχιακής εργασίας.

Η στοιχειοθεσία του κειμένου υλοποιήθηκε με το σύστημα ηλεκτρικής τυπογραφίας L^AT_EX.

Το τμήμα της εφαρμογής που αναπτύχτηκε για την παρουσίαση της ερευνητικής δραστηριότητας των μελών ΔΕΠ του Τμήματος Προφορικής του Ιονίου Πανεπιστημίου βρίσκεται στην ηλεκτρονική διεύθυνση:

<https://diionioresearch.000webhostapp.com/index.php>

Ενώ το τμήμα για το διαχειριστή της βάσης δεδομένων της εφαρμογής υπάρχει στην διεύθυνση:

<https://diionioresearch.000webhostapp.com/admin/>

Τέλος μπορεί κάποιος να βρει ολόκληρο τον κώδικα της εφαρμογής στο διαδικτυακό τόπο του GitHub ακολουθώντας τον παρακάτω σύνδεσμο:

<https://github.com/p09kont/graduate-project>

Κεφάλαιο 2

Βιβλιομετρικά δεδομένα - Βιβλιομετρικοί δείκτες

Σκοπός της πτυχιακής εργασίας, είναι ο σχεδιασμός και η ανάπτυξη μιας διαδικτυακής εφαρμογής παρουσίασης της ερευνητικής δραστηριότητας των μελών ΔΕΠ του Τμήματος Πληροφορικής του Ιόνιου Πανεπιστημίου. Ως εκ τούτου είναι απαραίτητο η εφαρμογή να επεξεργάζεται, να αναλύει και να παρουσιάζει δεδομένα ερευνητικών δημοσιεύσεων.

Στο κεφάλαιο αυτό πραγματοποιείται αναφορά στους διαδικτυακούς τρόπους και τις βάσεις δεδομένων που περιέχουν πληροφορίες σχετικές με την ερευνητική δραστηριότητα ατόμων και οργανισμών, καθώς και για τους τρόπους, που τα δεδομένα αυτά μπορούν να χρησιμοποιηθούν από άλλες εφαρμογές. Επιπλέον αναφέρονται δείκτες ανάλυσης και αξιολόγησης των δεδομένων ερευνητικής δραστηριότητας. Τέλος παρουσιάζονται οι λόγοι οι οποίοι οδήγησαν στην επιλογή ενός συγκεκριμένου διαδικτυακού τύπου για την άντληση των δεδομένων που θα επεξεργάζεται η εφαρμογή.

2.1 Βάσεις δεδομένων επιστημονικών δημοσιεύσεων

Η επιστημονική έρευνα αποτελεί ένα από τα κυριότερα εργαλεία ανάπτυξης και προόδου μιας κοινωνίας. Ως εκ τούτου η καταγραφή, η εξαγωγή συμπερασμάτων και η μελέτη των αποτελεσμάτων της έρευνας που πραγματοποιούνται διάφοροι οργανισμοί ανά τον κόσμο θα μπορούσε να μας δώσει πολύτιμη γνώση για το ποιές ερευνητικές κοινότητες προοδεύουν, πως εξελίσσεται η επιστημονική παραγωγή σε κάθε χώρα, και πως συνεργάζονται οι οργανισμοί σε παγκόσμιο επίπεδο. Κάτι τέτοιο αποτελεί θεμελιώδη παράγοντα για το σχεδιασμό πολιτικών έρευνας και καινοτομίας, που θα μπορούσαν να συμβάλουν αποφασιστικά στη διαμόρφωση νέων αναπτυξιακών μοντέλων, θέτοντας τη γνώση ως ανταγωνιστικό πλεονέκτημα.

Η μελέτη των αποτελεσμάτων είναι επίσης εξαιρετικά χρήσιμη για τους ερευνητές ώστε να μπορέσουν να επεκτείνουν τη γνώση τους και να σχηματίσουν ένα δίκτυο γνώσης που θα οδηγεί σε νέα ερευνητικά αποτελέσματα.

Όσον αφορά τα πανεπιστήμια, η μελέτη των ερευνητικών δραστηριοτήτων και των αποτελεσμάτων της που πραγματοποιείται στα πλαίσια του κάθε ιδρύματος, δίνει τη δυνατότητα στο να αναδείξουν και να προβάλλουν τα επιστημονικά πεδία στα οποία διαθέτουν πραγματογνωμοσύνη, αλλά και να δουν τα πεδία στα οποία εμφανίζουν αδυναμίες, ώστε να επιδιώξουν βελτίωση. Επίσης, δίνει τη δυνατότητα να προβάλλουν το έργο τους σε άλλα πανεπιστήμια, ιδρύματα ή ερευνητές και μέσω συσχέτισης των ερευνών να επιδιώξουν συνεργασίες. Τέλος τους παρέχει τη δυνατότητα να προβάλλουν τα γενικότερα αποτελέσματα και την υπάρχουσα γνώση και τεχνογνωσία στο ευρύ κοινό.

Το ερώτημα που γεννάται είναι πώς και από πού θα αντλήσουμε την πληροφορία που θα μας οδηγήσει στα παραπάνω συμπεράσματα. Τα αποτελέσματα μιας έρευνας που πραγματοποιεί μια ομάδα επιστημόνων δημοσιεύονται, μετά από συνήθως αυστηρή κρίση, ως άρθρα σε κάποιο επιστημονικό περιοδικό (ή πρακτικά συνεδρίων). Οι δημοσιεύσεις σε επιστημονικά περιοδικά, το κυριότερο μέσο για τη διάχυση των ερευνητικών αποτελεσμάτων στην επιστημονική κοινότητα, αποτελούν μια σημαντική πηγή για την καταγραφή και αποτίμηση του ερευνητικού έργου. Σκοπός μας λοιπόν είναι να συγκεντρώσουμε δεδομένα που προκύπτουν από τα παραπάνω άρθρα. Για να το κάνουμε αυτό, μπορούμε να χρησιμοποιήσουμε κάποια μηχανή αναζήτησης ερευνητικών δραστηριοτήτων ή κάποια διεθνή βιβλιογραφική βάση. Στη συνέχεια γίνεται αναφορά σε κάποιες από τις πιο γνωστές μηχανές αναζήτησης ερευνητικών δραστηριοτήτων.

Scopus

Το σύστημα **Scopus** (<https://www.scopus.com/home.uri>) της εταιρίας Elsevier αποτελεί μία από τις μεγαλύτερες βάσεις δεδομένων για ερευνητικές δραστηριότητες. Περιλαμβάνει πληροφορίες για επιστημονικές δημοσιεύσεις και τις αναφορές σε αυτές. Οι παλαιότερες βιβλιογραφικές εγγραφές ξεκινούν το 1966, ενώ οι αναφορές σε δημοσιεύσεις καταγράφονται μετά το 1995. Τα δεδομένα περιέχουν αναλυτικά στοιχεία για άρθρα, συγγραφείς και οργανισμούς, ενώ η εισαγωγή των εκδόσεων πραγματοποιείται μετά από ποιοτική αξιολόγηση βάσει κριτηρίων. Στη βάση δεδομένων του Scopus εμπεριέχονται περίπου 49 εκατομμύρια εγγραφές. Περιέχει επίσης έξυπνα εργαλεία για παρατήρηση, ανάλυση και οπτικοποίηση των αποτελεσμάτων της αναζήτησης. Στο νεότερο σύστημα Scopus, ευρετηριάζονται πάνω από 18.500 τίτλοι επιστημονικών περιοδικών, πρακτικά συνεδρίων και βιβλία, οι οποίοι διευρύνονται συνεχώς.

Web Of Science

Το σύστημα **Web Of Science** (www.webofknowledge.com) της εταιρίας Thomson Reuters παρέχει πρόσβαση σε βάσεις δεδομένων που καλύπτουν τα επιστημονικά περιοδικά με υψηλή επιρροή παγκοσμίως. Είναι η παλαιότερη βάση δεδομένων επιστημονικών δημοσιεύσεων με βιβλιογραφικές εγγραφές και ανα-

φορές που χρονολογούνται, για ορισμένους επιστημονικούς κλάδους, από το 1900. Αντλεί δεδομένα από περισσότερα από 12.000 περιοδικά τα οποία υπόκεινται σε αξιολόγηση κριτών (peer-review) ενώ από το 1990 προστέθηκε και η καταγραφή ορισμένων πρακτικών συνεδρίων. Περιλαμβάνονται αναλυτικά μεταδεδομένα για άρθρα, συγγραφείς και ερευνητικούς οργανισμούς. Σημαντικό πλεονέκτημα του Web of Science θεωρείται η αξιοπιστία του, αποτέλεσμα της αυστηρής αξιολόγησης των εκδόσεων και κυρίως των περιοδικών που εισάγονται στο σύστημα βάσει συγκεκριμένων κριτηρίων, μεταξύ των οποίων και η επιστημονική τους απήχηση.

Google Scholar

Το σύστημα *Google Scholar* (<https://scholar.google.gr/>) κυκλοφόρησε σε δοκιμαστική έκδοση το 2004 και γρήγορα έγινε ένα από τα πιο δημοφιλή συστήματα του είδους του. Οι λόγοι για την προτίμηση των χρηστών ήταν η απλότητα της χρήσης του και η δομή του συστήματος που είναι παρόμοια με την απλή μηχανή αναζήτησης της Google. Ένα ακόμη θετικό στοιχείο του Google Scholar είναι, ότι δεν επικεντρώνεται μόνο πάνω σε ένα επιστημονικό πεδίο, αλλά προσπαθεί να συλλέξει πληροφορίες για όλα τα επιστημονικά πεδία. Το σύστημα αντλεί δεδομένα από επιστημονικά άρθρα, διατριβές, εκθέσεις, περιλήψεις που μπορεί να βρίσκονται σε ιστότοπους ακαδημαϊκών δημοσιεύσεων, σε online αποθήκες, σε ιστοσελίδες πανεπιστημίων και σε άλλους διαδικτυακούς τόπους. Στο Google Scholar μία αναζήτηση γίνεται όπως και στην κανονική αναζήτηση της Google αλλά τα αποτελέσματα αφορούν επιστημονικές δραστηριότητες και δείκτες μέτρησης τους. Το Google Scholar περιλαμβάνει έναν τεράστιο αριθμό πηγών, πολλές από τις οποίες αφορούν δημοσιεύματα «γκρίζας βιβλιογραφίας» (π.χ. ακαδημαϊκές εργασίες, διατριβές, τεχνικές εκθέσεις, παραδοτέα ερευνητικών έργων). Παρά το μεγάλο εύρος σε περιεχόμενο, τα μεταδεδομένα των εγγραφών είναι περιορισμένα και απουσιάζουν κρίσιμα στοιχεία για τη γεωγραφική προέλευση ή τη διεύθυνση των οργανισμών στους οποίους ανήκουν οι συγγραφείς. Επιπλέον, τα κριτήρια εισαγωγής των πηγών του περιεχομένου είναι περιορισμένα, ενώ δεν διατίθενται ακριβή στοιχεία για τη γεωγραφική ή τη θεματική κάλυψη που παρέχει το σύστημα.

DBLP

Το σύστημα *DBLP* (Digital Bibliography & Library Project, <https://dblp.uni-trier.de>) δημιουργήθηκε τη δεκαετία του 80 από το πανεπιστήμιο του Trier της Γερμανίας και ξεκίνησε αρχικά σαν μία βάση δεδομένων και μία ιστοσελίδα όπου αποθηκεύονταν βιβλιογραφίες σχετικές με το λογικό προγραμματισμό. Εκείνη τη περίοδο DBLP σήμαινε DataBase systems & Logic Programming. Βέβαια με τα χρόνια η λίστα με τις βιβλιογραφίες έγινε αποθήκη δεδομένων και σήμερα διαθέτει πάνω από 1,2 εκατομμύρια άρθρα πάνω στην επιστήμη των υπολογιστών. Γενικά το DBLP διαθέτει ένα αρκετά δύσχερο περιβάλλον εργασίας, εμφανίζοντας τα δεδομένα με πεπαλαιωμένο τρόπο, χρησιμοποιώντας απλούς πίνακες HTML και λίστες, με σχεδόν καθόλου γραφικά.

Παρ' όλα αυτά δεν παύει να είναι μια εξαιρετική βάση δεδομένων που αυξάνει συνεχώς τον αριθμό των αποθηκευμένων πληροφοριών της.

Microsoft Academic Search

Το *Microsoft Academic Search* (<https://academic.microsoft.com/>), ήταν η προσπάθεια της Microsoft να εισχωρήσει και σε αυτόν τον τομέα της πληροφορικής. Αναπτύχθηκε από την ερευνητική ομάδα της Microsoft στην Ασία και η βάση δεδομένων της βασιζόταν σε πληροφορίες σχετικά με έγγραφα συνεδρίων και σε δημοσιεύσεις που έχουν γίνει σε επιστημονικά περιοδικά. Το 2008 η βάση δεδομένων της μηχανής αναζήτησης σταμάτησε να ανανεώνεται και συγχωνεύτηκε με την απλή αναζήτηση της Microsoft, κάνοντας μέχρι τότε αναζήτηση σε πάνω από 3 εκατομμύρια έγγραφα. Η μηχανή αναζήτησης της Microsoft παρέχει μια πολύ καλή κατηγοριοποίηση των δημοσιεύσεων όπως επίσης και έναν εξαιρετικά καλαίσθητο τρόπο παρουσίασης των αποτελεσμάτων μέσα από πλούσιες και διαδραστικές οπτικοποιήσεις δεδομένων. Ένα πολύ καλό στοιχείο της μηχανής αναζήτησης είναι ότι διαθέτει μία λίστα με τα πιο υψηλά καταταγμένα έγγραφα, συγγραφείς, συνέδρια και περιοδικά, για πολλά διαφορετικά πεδία της επιστήμης της πληροφορικής όπως, θεωρία αλγορίθμων, βιοπληροφορική, λειτουργικά συστήματα, πολυμέσα, μηχανική λογισμικού και άλλα πολλά.

2.2 Δείκτες βιβλιομετρικής ανάλυσης

Σύμφωνα με το Εθνικό Ίδρυμα Ερευνών και το Εθνικό Κέντρο Τεκμηρίωσης αντικείμενο της βιβλιομετρικής ανάλυσης είναι η καταγραφή και επεξεργασία δεδομένων που σχετίζονται με τις επιστημονικές δημοσιεύσεις και η εξαγωγή των σχετικών «βιβλιομετρικών δεικτών», όπως ο αριθμός των δημοσιεύσεων, οι αναφορές σε αυτές από άλλες δημοσιεύσεις (citations), ο συσχετισμός τους με συγκεκριμένους φορείς, επιστημονικά πεδία, κ.λπ.

Οι βιβλιομετρικοί δείκτες μέτρησης της ερευνητικής δραστηριότητας συμβάλλουν στη διαμόρφωση μιας κατά το δυνατόν αντικειμενικά μετρήσιμης εικόνας των ερευνητικών συστημάτων. Χρησιμοποιούνται για την αξιολόγηση ερευνητικών οργανισμών, ομάδων και ερευνητών, καταγράφουν τα ερευνητικά πεδία στα οποία δραστηριοποιείται η επιστημονική κοινότητα, αποτυπώνουν τα νέα ερευνητικά πεδία που αναδύονται καθώς και τα επιστημονικά δίκτυα που δημιουργούνται για την υλοποίηση κοινών ερευνητικών στόχων.

Η βιβλιομετρική ανάλυση αποτελεί σημαντικό εργαλείο για τον προσδιορισμό χαρακτηριστικών και τάσεων της ερευνητικής παραγωγής σε επίπεδο οργανισμού, χώρας ή ευρύτερου συνόλου χωρών, την εκτίμηση της απήχησης του επιστημονικού έργου, την αξιολόγηση της ερευνητικής δραστηριότητας και την ανάδειξη εθνικών και πολυεθνικών δικτύων μεταξύ επιστημόνων και επιστημονικών κλάδων. Οι μελέτες που στηρίζονται στη βιβλιομετρική ανάλυση αυξάνονται συνεχώς τα τελευταία χρόνια στο διεθνές χώρο. Οι βιβλιομετρι-

κοί δείκτες χρησιμοποιούνται για την αξιολόγηση ερευνητικών συστημάτων ή οργανισμών και συνεισφέρουν στη διαμόρφωση εθνικών πολιτικών για την έρευνα.

Στη βιβλιογραφία αναφέρονται προβλήματα που σχετίζονται με τον «καθαρισμό» των πρωτογενών δεδομένων και την ταυτοποίηση των δημοσιεύσεων καθώς και περιορισμοί που πρέπει να λαμβάνονται υπόψη στον υπολογισμό και τη χρήση του αριθμού των αναφορών, δεδομένου ότι ο αριθμός των αναφορών που λαμβάνουν οι δημοσιεύσεις επηρεάζεται σημαντικά από παραμέτρους που παρατίθενται στη συνέχεια:

- **Επιστημονικά πεδία:** υπάρχουν σημαντικές διαφοροποιήσεις στην πρακτική δημοσίευσης ερευνητικών εργασιών, στον χρόνο απαξίωσης των ερευνητικών αποτελεσμάτων και στις πρακτικές αναφορών σε δημοσιεύσεις (citation patterns) ανάλογα με τους επιστημονικούς τομείς. Το γεγονός αυτό καθιστά αδύνατη την άμεση σύγκριση δεικτών μεταξύ διαφορετικών ερευνητικών πεδίων, συχνά και μεταξύ επιμέρους περιοχών του ίδιου ευρύτερου επιστημονικού πεδίου. Για παράδειγμα, στην Ιατρική και στη Μοριακή Βιολογία, παρατηρείται υψηλότερη παραγωγικότητα ως προς την ποσότητα των επιστημονικών άρθρων που δημοσιεύονται ανά έτος. Συνήθως, ο αριθμός αναφορών των δημοσιεύσεων αυτών αυξάνεται και φτάνει στο μέγιστο σε μικρό χρονικό διάστημα από τη δημοσίευσή τους. Σε αντίθεση, στις κοινωνικές επιστήμες οι ρυθμοί δημοσίευσης είναι πολύ πιο αργοί, ενώ καταγράφονται αναφορές για μεγάλο χρονικό διάστημα μετά τη δημοσίευσή τους. Στις ανθρωπιστικές επιστήμες, το μεγαλύτερο μέρος των επιστημονικών δημοσιεύσεων είναι βιβλία, μονογραφίες και άρθρα που δημοσιεύονται σε επιστημονικά περιοδικά εθνικής εμβέλειας. Σε άλλες επιστημονικές περιοχές (π.χ. σε ορισμένους κλάδους της Πληροφορικής) είναι κοινή πρακτική η δημοσίευση σημαντικών ερευνητικών αποτελεσμάτων σε συνέδρια, χωρίς να ακολουθεί απαραίτητα αντίστοιχη δημοσίευση σε επιστημονικά περιοδικά. Και στις δύο περιπτώσεις, δημοσιεύσεις υψηλής επιστημονικής αξίας, όπως και μέρος των αναφορών τους, κατά κανόνα δεν καταμετρώνται στις βάσεις δεδομένων με αποτέλεσμα οι βιβλιομετρικοί δείκτες στις αντίστοιχες επιστημονικές περιοχές να μην αποτυπώνουν πλήρως την πραγματικότητα.
- **Χρονικό διάστημα ανάλυσης αναφορών:** ο αριθμός των αναφορών σε μια επιστημονική εργασία εξαρτάται από το χρονικό διάστημα που έχει παρέλθει μετά τη δημοσίευσή της. Συνήθως οι παλαιότερες δημοσιεύσεις έχουν και περισσότερες αναφορές, χωρίς αυτό να συνδέεται πάντα με την απήχηση τους στην επιστημονική κοινότητα. Για την εξομάλυνση των διαφορών που οφείλονται αποκλειστικά στη φυσιολογική αύξηση του αριθμού των αναφορών σε παλαιότερες δημοσιεύσεις, ορίζονται συγκεκριμένα χρονικά διαστήματα για τη μέτρηση των αναφορών σε μία δημοσίευση.

- **Είδος επιστημονικών δημοσιεύσεων:** ο αριθμός των αναφορών επηρεάζεται επίσης σημαντικά από τον τύπο της δημοσίευσης. Τα άρθρα ανασκόπησης (review) έχουν περισσότερες αναφορές από τα απλά επιστημονικά άρθρα. Για τον λόγο αυτό, ο προσδιορισμός διαφορετικών βιβλιομετρικών δεικτών ανάλογα με το είδος της δημοσίευσης, είναι απαραίτητος σε περιπτώσεις αξιολόγησης του επιστημονικού έργου μιας ερευνητικής ομάδας ή μεμονωμένων ερευνητών.

Οι παραπάνω προβληματισμοί δεν αναιρούν τη σημασία των βιβλιομετρικών δεικτών ως πολύτιμη πηγή δεδομένων και, όπως άλλωστε ισχύει με την ερμηνεία των περισσότερων δεικτών, μπορούν να ξεπεραστούν όταν οι βιβλιομετρικοί δείκτες ερμηνευτούν στο σωστό πλαίσιο.

Υπάρχει μεγάλη πληθώρα βιβλιομετρικών δεικτών που μπορούν να χρησιμοποιηθούν σε βιβλιομετρικές αναλύσεις. Ο καθένας από αυτούς έχει τον δικό του ρόλο και σημασία. Εδώ θα αναφερθούν μόνο οι βιβλιομετρικοί δείκτες οι οποίοι χρησιμοποιήθηκαν στην εφαρμογή που αναπτύχτηκε.

Ο αριθμός των δημοσιεύσεων

Είναι ο αριθμός των δημοσιεύσεων (*Number Publications*) που αντιστοιχεί στον κάθε ερευνητή. Στην εφάρμογή μια δημοσίευση συγκαταλέγεται σε κάποια από τις επόμενες κατηγορίες: Άρθρο σε επιστημονικό περιοδικό, άρθρο σε συνέδριο, κεφάλαιο σε βιβλίο και βιβλίο.

Αριθμός αναφορών από άλλες δημοσιεύσεις

Ο αριθμός των αναφορών από άλλες δημοσιεύσεις (*Number of Citations*) είναι ένας δείκτης που εκφράζει πόσο σημαντική είναι μία δημοσίευση αφού δείχνει τον αριθμό των δημοσιεύσεων που αναφέρονται σε αυτήν.

Αριθμός αναφορών ανά δημοσίευση

Ο αριθμός αναφορών ανά δημοσίευση (*Number of Citations per Document*) δείχνει τον μέσο όρο των αναφορών που έχει λάβει μια δημοσίευση και υπολογίζεται σύμφωνα με τον μαθηματικό τύπο:

$$\text{Μέσος αριθμός αναφορών} = \frac{\text{Συνολικός αριθμός αναφορών}}{\text{Συνολικός αριθμός δημοσιεύσεων}}$$

H – index

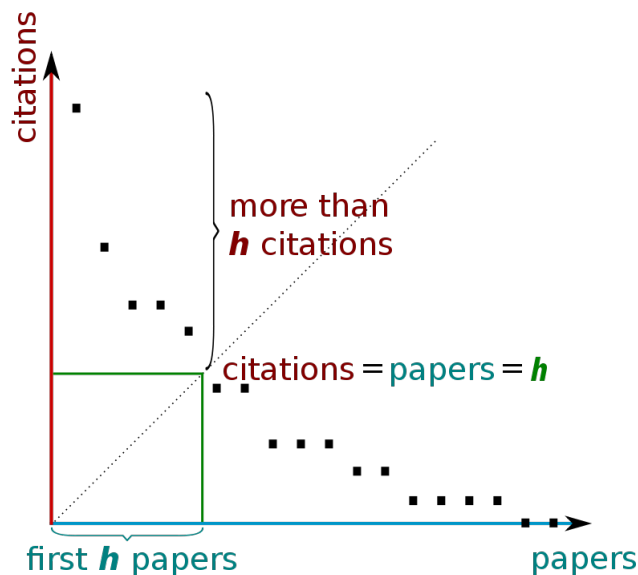
Ο δείκτης επιστημονικής ποιότητας h , ή απλά δείκτης h ($h - index$) είναι ένας αριθμητικός παράγοντας που στοχεύει στη μέτρηση της ποιότητας των ακαδημαϊκών δημοσιεύσεων ενός επιστήμονα. Έπινοήθηκε το 2005 από τον Jorge Hirsch. Σύμφωνα με τον ορισμό που έδωσε ο Hirsch ένας συγγραφέας εμφανίζει δείκτη h όταν h από τις N εργασίες του έχουν τουλάχιστον h αναφορές η καθεμία, ενώ οι υπόλοιπες ($N - h$) εργασίες του έχουν λιγότερες από h αναφορές η καθεμία. Παράδειγμα: Αν μια δοσμένη ομάδα εργασιών ή οι δοσμένες εργασίες κάποιου συγγραφέα εμφανίζουν $h = 12$, αυτό σημαίνει ότι σε αυτή

την ομάδα των εργασιών υπάρχουν 12 εργασίες που έχουν τουλάχιστον 12 (ή και περισσότερες) αναφορές η καθεμία. Δηλαδή, ο h – index εξάγεται πρακτικά με κατάταξη της δοσμένης ομάδας των εργασιών κατά φθίνουσα σειρά ξεκινώντας από αυτή που έχει τις περισσότερες αναφορές. Καθώς προχωράμε προς τα κάτω, στο σημείο που ο αριθμός των αναφορών που έχει λάβει κάποια εργασία γίνει μικρότερος από τον αύξοντα αριθμό της εργασίας, τραβάμε μία γραμμή και λαμβάνουμε ως h – index τον αμέσως προηγούμενο αύξοντα αριθμό. Εξ' ορισμού η τιμή του h – index εξαρτάται από τη βάση δεδομένων στην οποία θα γίνει η αναζήτηση.

Σύμφωνα με την διαδικτυακή εγκυκλοπαίδεια *Wikipedia* (<https://www.wikipedia.org/>) ο h – index ενός ερευνητή μπορεί να υπολογιστεί μαθηματικά ως εξής: Έστω f η συνάρτηση που αντιστοιχεί στον αριθμό των αναφορών για κάθε δημοσίευση και N το πλήθος των δημοσιεύσεων του ερευνητή. Αν ταξινομήσουμε τη συνάρτηση f σε φθίνουσα σειρά από τη μεγαλύτερη τιμή στη χαμηλότερη, μπορούμε να υπολογίσουμε τον δείκτη h από τον τύπο:

$$h\text{-index}(f) = \max_i \min(f(i), i) \quad i = 1, 2, \dots, N$$

Η εικόνα 2.1, που λήφθηκε από τον ιστότοπο της *Wikipedia* στην διεύθυνση <https://en.wikipedia.org/wiki/H-index>, παρουσιάζει ένα τρόπο γραφικού υπολογισμού του h – index.



Εικόνα 2.1: Γραφικός υπολογισμός του h – index

2.3 Επιλογή διαδικτυακού τόπου άντλησης δεδομένων

Για να είναι δυνατόν η εφαρμογή, που αναπτύχθηκε στα πλαίσια της πτυχιακής εργασίας, να λειτουργήσει αποτελεσματικά και να επιτύχει τους στόχους για τους οποίους υλοποιήθηκε είναι απαραίτητη η δημιουργία μιας βάσης δεδομένων από την οποία η εφαρμογή θα δέχεται και θα επεξεργάζεται δεδομένα. Η ανάγκη αξιοπιστίας των εν λόγω δεδομένων επιβάλλει αυτά να συλλεχτούν από ένα σύστημα που ευρετηριάζει δεδομένα ερευνητικών δημοσιεύσεων. Από τα συστήματα που παρουσιάζονται στην ενότητα 2.1 εκείνο το οποίο επιλέχτηκε για να γίνει συλλογή των δεδομένων που θα αποθηκευτούν στη βάση της εφαρμογής είναι το Scopus. Στη συνέχεια παρουσιάζονται οι λόγοι που οδήγησαν στην συγκεκριμένη επιλογή.

Αναμφίβολα από τα συστήματα για τα οποία γίνεται λόγος στην ενότητα 2.1 το δημοφιλέστερο αλλά και αυτό το οποίο περιέχει τον μεγαλύτερο όγκο δεδομένων είναι το Google Scholar. Με βάση τα παραπάνω χαρακτηριστικά, το σύστημα Google Scholar, παρά τον τεράστιο αριθμό πηγών που περιλαμβάνει, και της μεγάλης δημοφιλίας του, δεν είναι κατάλληλο για βιβλιομετρικές αναλύσεις, λόγω της έλλειψης μεταδεδομένων που απαιτούνται για την ταυτοποίηση των δημοσιεύσεων και της απουσίας κριτηρίων που διασφαλίζουν την ποιότητα των δημοσιεύσεων που περιλαμβάνονται στο σύστημα.

Επιπλέον από ευρεία αναζήτηση που πραγματοποιήθηκε στο διαδίκτυο, διαπιστώθηκε ότι δεν υπάρχει κάποιο εργαλείο το οποίο να είναι δυνατόν να χρησιμοποιηθεί ώστε να μετατραπούν τα δεδομένα των ιστοσελίδων του Google Scholar σε μορφή στην οποία να μπορούν να εισαχθούν σε μια βάση δεδομένων. Όσον αφορά εφαρμογές λογισμικού που αναχτούν και αναλύουν τα δεδομένα βιβλιογραφικών αναφορών από το Google Scholar με πιο χαρακτηριστική το λογισμικό Publish or Perish (<https://harzing.com/resources/publish-or-perish>), αυτές είναι δυνατόν να συλλέξουν μόνο τις δημοσιεύσεις όσων ερευνητών έχουν δημιουργήσει προφίλ στο Google Scholar.

Για τους λόγους που προαναφέρθηκαν αλλά και εξαιτίας της επιθυμίας η εφαρμογή να είναι απαλλαγμένη από τη χρησιμοποίηση κάποιου βοηθητικού λογισμικού, το οποίο ίσως να απέτρεπε πολλούς από τη χρήση της, δεν έγινε επιλογή του συστήματος Google Scholar για την άντληση των δεδομένων.

Το σύστημα Microsoft Academic Search δεν παρέχει καθόλου μεταδεδομένα των βιβλιογραφικών αναφορών που ευρετηριάζει. Έκτος αυτού δεν βρεθήκαν και εφαρμογές λογισμικού που να είναι δυνατόν να ανακτήσουν δεδομένα από τις ιστοσελίδες του. Επομένως είναι προφανές ότι δεν επιλέχτηκε.

Το DBLP ευρετηριάζει δεδομένα βιβλιογραφικών δημοσιεύσεων σε περιοδικά και συνέδρια που αφορούν επιστημονικά πεδία σχετικά μόνο με τον τομέα της Πληροφορικής. Επιπλέον παρέχει μεταδεδομένα για όλες ή μεμονωμένες έγγραφες της βάσης του σε μορφή XML. Το γεγονός αυτό καθίστα αρκετά εύκολη την συλλογή αυτών των δεδομένων προγραμματιστικά και την εισαγωγή

τους σε βάση δεδομένων. Αυτό μπορεί να επιτευχθεί με την χρήση βιβλιοθηκών που παρέχουν εντολές για την ανάλυση (parsing) αρχείων XML σε γλώσσες προγραμματισμού όπως η JAVA και η PHP.

Τον πρώτο καιρό της ενασχόλησης με την πτυχιακή εργασία προτιμήθηκε να γίνει συλλογή των δεδομένων επιστημονικών δημοσιεύσεων των μελών ΔΕΠ του Τμήματος Πληροφορικής από τα αρχεία XML του DBLP. Η συλλογή έγινε προγραμματιστικά σε γλώσσα PHP με την χρήση εντολών της βιβλιοθήκης ανάλυσης δεδομένων SAX (Simple API for XML).

Ο SAX είναι ένας σειριακής πρόσβασης parser ο οποίος παρέχει ένα μηχανισμό για την ανάγνωση δεδομένων από ένα έγγραφο XML που είναι μια εναλλακτική λύση σε σχέση με το Document Object Model (DOM). Το βασικό πλεονέκτημα του SAX είναι ότι δε χρειάζεται πρώτα να αποθηκεύσει άρα και να διαβάσει ολόκληρο το XML αρχείο όπως κάνει ένας DOM parser, που δεσμεύει μνήμη για αποθήκευση του αρχείου σε μορφή δέντρου. Όταν το DOM λειτουργεί στο έγγραφο ως σύνολο, ο SAX λειτουργεί σε κάθε κομμάτι του XML διαδοχικά.

Επιπλέον ο SAX είναι event-driven API δηλαδή ο χρήστης ορίζει μια σειρά από μεθόδους επανάκλησης που θα κληθούν όταν συμβαίνουν συγκεκριμένα γεγονότα κατά τη διάρκεια της ανάλυσης. Κατά συνέπεια με αυτές τις μεθόδους ο χρήστης μπορεί να επιλέξει ό,τι πληροφορίες θέλει από το αρχείο XML.

Το εγχείρημα συλλογής δεδομένων από τα XML αρχεία του συστήματος DBLP έφτασε ως το σημείο κατασκευής της βάσης διαδομένων και της εισαγωγής όλων των στοιχείων που είχαν συλλεχτεί στην βάση αυτή. Το επόμενο βήμα ήταν η δημιουργία των ιστοσελίδων της εφαρμογής που θα παρουσίαζαν τα δεδομένα αυτά. Σε εκείνο το σημείο όμως και σε συνεννόηση με την τριμελή επιτροπή η προσπάθεια αυτή εγκαταλείφτηκε.

Ένας λόγος για τον οποίο εγκαταλείφτηκε αυτή η προσπάθεια οφείλεται στο ότι το DBLP περιέχει δεδομένα δημοσιεύσεων που σχετίζονται μόνο με το επιστημονικό πεδίο του κλάδου της Πληροφορικής. Το γεγονός αυτό έχει ως αποτέλεσμα να μην υπάρχουν στοιχεία για δημοσιεύσεις ερευνητών σε άλλα επιστημονικά πεδία και ως εκ τούτου, να καθίσταται αδύνατη μια σαφή εικόνα για την ποσότητα των ερευνητικών δημοσιεύσεων ενός μέλους ΔΕΠ.

Εξίσου σημαντικός λόγος ήταν και το γεγονός ότι στα δεδομένα του DBLP δεν περιέχονται πληροφορίες για τον αριθμό των αναφορών που έχει λάβει μια δημοσίευση από άλλες. Αυτό έχει ως αποτέλεσμα να είναι αδύνατον να υπολογιστεί ο h – index ενός ερευνητή και επομένως να μην μπορεί να γίνει ποιοτική αξιολόγηση της ερευνητικής του δραστηριότητας.

Όσον αφορά τα συστήματα Web of Science και Scopus, και τα δύο διασφαλίζουν τη διάθεση αναλυτικών μεταδεδομένων και την ποιότητα των δημοσιεύσεων που περιλαμβάνουν και είναι αυτά που χρησιμοποιούνται διεθνώς για την υλοποίηση βιβλιομετρικών μελετών. Οι βάσεις δεδομένων των δύο αυτών συστημάτων επικαιροποιούνται σε συνεχή βάση, διευρύνουν τον αριθμό και το εύρος των επιστημονικών περιοδικών που ευρετηριάζουν ενώ εμπλουτίζουν τις πηγές δεδομένων συμπεριλαμβάνοντας πρακτικά συνεδρίων, βιβλία, μονογρα-

φίες κ.λπ. Ωστόσο και οι δύο βάσεις δεν επιτυγχάνουν να παρέχουν πλήρη κάλυψη των επιστημονικών δημοσιεύσεων, τόσο γεωγραφικά (αναφέρεται ότι το Web of Science έχει καλύτερη κάλυψη των αγγλόφωνων περιοδικών και κυρίως των ΗΠΑ ενώ το Scopus έχει καλύτερη κάλυψη των ευρωπαϊκών περιοδικών) όσο και θεματικά στα διάφορα επιστημονικά πεδία (αναφέρεται ότι το Web of Science υπερτερεί στην κάλυψη των επιστημονικών περιοχών των Φυσικών επιστημών ενώ το Scopus των Ιατρικών επιστημών).

Η δομή και ο τρόπος με τον οποίο έχει υλοποιηθεί το σύστημα Web of Science καθίστα ιδιαίτερα χρονοβόρα την ανάλυση και την εισαγωγή των δεδομένων του σε βάση δεδομένων. Έτσι σύμφωνα με τους λογούς που αναφέρθηκαν επιλέχτηκε να γίνει συλλογή των δεδομένων της εφαρμογής από μεταδεδομένα που περιέχονται σε αρχεία της μορφής csv (comma separated values) τα οποία παράγονται από το σύστημα του Scopus.

Στο κεφάλαιο 4 γίνεται εκτενής αναφορά στον κώδικα σε γλώσσά PHP με τον οποίο προγραμματίστηκε η ανάλυση των αρχείων csv που παράγονται από το Scopus και η εισαγωγή των εγγραφών τους στη βάση δεδομένων της εφαρμογής. Ο τρόπος με τον οποίο μπορεί να παραχθεί από το σύστημα Scopus ένα αρχείο csv που περιέχει δεδομένα δημοσιεύσεων ενός ερευνητή παρουσιάζεται στην ενότητα 5.1.

Κεφάλαιο 3

Η βάση δεδομένων της εφαρμογής

3.1 Ο καθορισμός των απαιτήσεων

Στόχος της πτυχιακής εργασίας είναι ο σχεδιασμός και η ανάπτυξη μιας διαδικτυακής εφαρμογής παρουσίασης της ερευνητικής δραστηριότητας των μελών ΔΕΠ του Τμήματος Πληροφορικής του Ιονίου Πανεπιστημίου. Ο καθορισμός των απαιτήσεων έγινε σε συνεννόηση την τριμελή επιτροπή, αλλά και από τη μελέτη του αρχείου scopus.csv, το οποίο βρίσκεται στον δικομιστή του Scopus. Το αρχείο αυτό περιέχει δεδομένα για τις ερευνητικές δημοσιεύσεις κάθε μέλους ΔΕΠ και χρησιμοποιείται για την ανάλυση των εγγραφών του και την εισαγωγή τους στη βάση δεδομένων της εφαρμογής.

Σύμφωνα με τον καθορισμό των απαιτήσεων η εφαρμογή θα αποτελείται από δυο τμήματα (modules), το τμήμα για τον διαχειριστή (administrator) και το τμήμα των χρηστών. Στη συνέχεια γίνεται περιγραφή της λειτουργίας αυτών των τμημάτων.

Τμήμα διαχειριστή

Ο διαχειριστής του συστήματος, μέσω μιας διεπαφής ιστού (web interface) θα έχει την δυνατότητα αλληλεπίδρασης με τη βάση δεδομένων. Συγκεκριμένα ο διαχειριστής θα μπορεί να κάνει τα εξής:

- Να εισάγει νέα μέλη ΔΕΠ στη βάση δεδομένων.
- Να τροποποιήσει τα στοιχεία τους.
- Να διαγράψει ένα μέλος ΔΕΠ από τη βάση δεδομένων, στην περίπτωση που αυτό έχει αποχωρίσει από το δυναμικό του Τμήματος.
- Να κατεβάσει από τον δικομιστή του Scopus τα αρχεία scopus.csv που αντιστοιχούν σε κάθε μέλος ΔΕΠ και να τα ανεβάσει στο δικομιστή της εφαρμογής.

- Να ξεκινήσει τη διαδικασία ανάλυσης του αρχείου scopus.csv για κάθε μέλος ΔΕΠ, η οποία έχει ως αποτέλεσμα την εισαγωγή των δημοσιεύσεων του συγκεκριμένου μέλους στη βάση δεδομένων.
- Να διαγράφει από τη βάση δεδομένων όλες ή ορισμένες από τις δημοσιεύσεις ενός συγκεκριμένου μέλους ΔΕΠ.

Τμήμα χρηστών συστήματος

Οι χρήστες του συστήματος θα έχουν την δυνατότητα πληροφόρησης για την έρευνα που έχει παραχθεί από κάθε μέλος ΔΕΠ ξεχωριστά, αλλά και από ολόκληρο το Τμήμα γενικότερα.

Σχετικά με το σύνολο της έρευνας που έχει παραχθεί από ολόκληρο το Τμήμα κάθε χρήστης θα μπορεί να κάνει τα εξής:

- Θα έχει τη δυνατότητα να δει τους τίτλους όλων των δημοσιεύσεων που έχουν γίνει από όλα τα μέλη ΔΕΠ και να τις ταξινομήσει, φθίνουσα ή αύξουσα, ανά όνομα, έτος, τύπο και αριθμό αναφορών.
- Θα μπορεί να πληροφορηθεί τον αριθμό των δημοσιεύσεων ανά τύπο, αλλά και συνολικά που αντιστοιχούν σε κάθε μέλος ΔΕΠ.
- Να δει γραφήματα που απεικονίζουν, για όλα τα μέλη, βιβλιοδετικούς δείκτες όπως ο συνολικός αριθμός των αναφορών, ο μέσος όρος των αναφορών και ο h – index.
- Μέσω ενός διαδραστικού γραφήματος αναπαράστασης δικτύου συνεργασιών, να λάβει πληροφορίες για τις συνεργασίες που έχουν γίνει μεταξύ των μελών στην συγγραφή δημοσιεύσεων.
- Θα έχει τη δυνατότητα να πλοηγηθεί από το πρώτο έως το τελευταίο έτος έρευνας του τμήματος και να ταξινομήσει ανά είδος και συνολικό αριθμό δημοσιεύσεων την ερευνητική δραστηριότητα κάθε μέλους, ώστε να είναι σε θέση να πληροφορηθεί ποιο μέλος παρήγαγε περισσότερη έρευνα το εκάστοτε έτος.
- Να δει τα ονόματα των περιοδικών και των συνεδρίων, καθώς και το πλήθος των άρθρων που έχουν δημοσιευτεί σε κάθε περιοδικό ή στα πρακτικά του συνεδρίου.

Σχετικά με την έρευνα που έχει παραχθεί από κάποιο συγκεκριμένο μέλος ΔΕΠ κάθε χρήστης θα μπορεί να κάνει τα εξής:

- Θα είναι σε θέση να δει τους τίτλους των δημοσιεύσεων που αντιστοιχούν σε κάθε μέλος ΔΕΠ και να τις ταξινομήσει φθίνουσα ή αύξουσα, ανά τίτλο, τύπο, έτος και αριθμό αναφορών.
- Να πληροφορηθεί ποιός είναι ο h – index κάθε μέλους ΔΕΠ.

- Να ενημερωθεί για τη χρονολογική εξέλιξη της έρευνας του συγκεκριμένου μέλους ΔΕΠ μέσω ενός χρονολογικού γραφήματος.
- Να δει τον αριθμό του είδους των δημοσιεύσεων για κάθε έτος έρευνας ενός συγκεκριμένου μέλους ΔΕΠ, και να λάβει γνώση για το ποιο είναι το έτος που παράχθηκε περισσότερη έρευνα από το μέλος αυτό.
- Να πληροφορηθεί με ποιούς άλλους ερευνητές έχει συνεργαστεί ένα μέλος ΔΕΠ για την συγγραφή δημοσιεύσεων, καθώς και το πλήθος, αλλά και τα στοιχεία των δημοσιεύσεων που έχει κάνει το συγκεκριμένο μέλος ΔΕΠ με τον εκάστοτε ερευνητή.

3.2 Το Διάγραμμα Οντοτήτων – Συσχετίσεων

Για τις απαιτήσεις της εφαρμογής είναι αναγκαία η κατασκευή μιας βάσης δεδομένων, στην οποία θα αποθηκεύονται όλα τα απαραίτητα στοιχεία. Για το σκοπό αυτό, πρέπει μετά τον καθορισμό των απαιτήσεων, να σχεδιαστεί το Διάγραμμα Οντοτήτων – Συσχετίσεων. Το διάγραμμα αυτό βοηθάει στην καλύτερη οργάνωση και καταγραφή του όγκου των πληροφοριών που λαμβάνονται από τον καθορισμό των απαιτήσεων. Ακόμα συμβάλει στο να αποσαφηνιστούν και να εντοπιστούν όλες οι φυσικές και τεχνητές οντότητες που θα αλληλεπιδρούν μεταξύ τους, καθώς επίσης να οριστούν και όλες οι συσχετίσεις μεταξύ αυτών. Επιπλέον διευκολύνει στην δημιουργία του σχεσιακού σχήματος της βάσης δεδομένων.

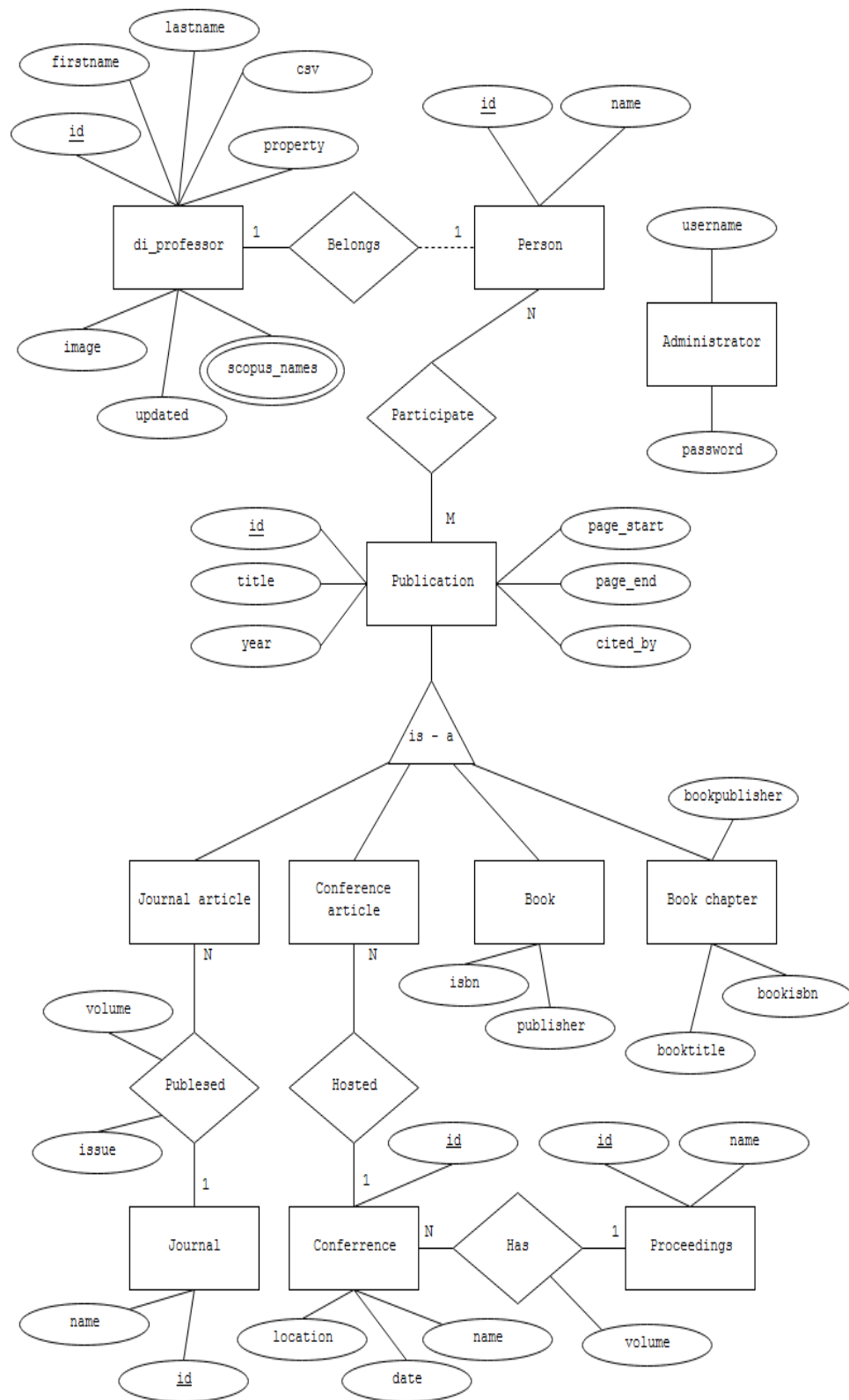
Το Διάγραμμα Οντοτήτων – Συσχετίσεων (ΔΟΣ), περιέχει τις έννοιες *οντότητα*, *γνώρισμα* και *συσχέτιση* οι οποίες συμβολίζονται με γεωμετρικά σχήματα που συνδέονται μεταξύ τους με γραμμές. Στην συνέχεια δίνεται μια περιγραφή των εννοιών αυτών.

Οντότητα είναι κάθε φυσικό ή νοητό αντικείμενο του πραγματικού κόσμου, για το οποίο θέλουμε να αποθηκεύσουμε δεδομένα στη βάση δεδομένων. Σε ένα ΔΟΣ κάθε οντότητα συμβολίζεται με ένα παραλληλόγραμμο μέσα στο οποίο αναγράφεται το όνομα της οντότητας στον ενικό αριθμό.

Οι οντότητες έχουν ιδιότητες. Οι ιδιότητες αυτές περιγράφουν πλήρως τις οντότητες και ονομάζονται γνωρίσματα (attributes). Τα γνωρίσματα των οντοτήτων αντιστοιχούν σε συγκεκριμένα δεδομένα που θέλουμε να αποθηκεύσουμε στη βάση δεδομένων. Κάθε γνώρισμα αναπαριστάται γραφικά με μία έλλειψη, η οποία προσαρτάται στην οντότητα στην οποία ανήκει.

Οι οντότητες συσχετίζονται μεταξύ τους. Η συσχέτιση μεταξύ δυο οντοτήτων μπορεί να δείχνει συνεργασία ή αλληλεπίδραση ή άλλου είδους σχέση. Η συσχέτιση συμβολίζεται με ένα ρόμβο, ο οποίος συνδέει τις οντότητες που συσχετίζονται. Σε κάθε συσχέτιση δίνεται μια ονομασία, η οποία εμφανίζεται μέσα στο ρόμβο και συνήθως είναι ένα ρήμα.

Από τον καθορισμό των απαιτήσεων που έγινε στην προηγούμενη ενότητα, προκύπτει το Διάγραμμα Οντοτήτων – Συσχετίσεων της επόμενης εικόνας.



Εικόνα 3.1: Το Διάγραμμα Οντοτήτων – Συσχετίσεων

Στο διάγραμμα της εικόνας 3.1 υπάρχει η οντότητα `di_professors` στην οποία έχουν εκχωρηθεί τα γνωρίσματα που ταυτοποιούν κάθε μέλος ΔΕΠ αλλά και όσα είναι απαραίτητα για την λειτουργία της εφαρμογής. Τα γνωρίσματα αυτά είναι:

- **id**
Μοναδικό προσδιοριστικό κάθε μέλους ΔΕΠ (πρωτεύων κλειδί).
- **firstname**
Το όνομα του μέλους ΔΕΠ.
- **lastname**
Το επώνυμο του μέλους ΔΕΠ.
- **property**
Η βαθμίδα του μέλους ΔΕΠ (π.χ. Professor, Associate Professor, Assistant Professor κλπ).
- **csv**
Το όνομα του csv αρχείου που βρίσκεται στον διακομιστή της εφαρμογής και στο οποίο αναγράφονται τα στοιχεία των δημοσιεύσεων του συγκεκριμένου μέλους ΔΕΠ.
- **image**
Το όνομα του αρχείου εικόνας στον διακομιστή της εφαρμογής, η οποία απεικονίζει το συγκεκριμένο μέλος ΔΕΠ.
- **updated**
Μια χρονοσφραγίδα στην οποία αποθηκεύεται η ημερομηνία και η ώρα της τελευταίας ενημέρωσης των δημοσιεύσεων του συγκεκριμένου μέλους ΔΕΠ.
- **names**
Χαρακτηριστικό πολλαπλών τιμών το οποίο περιέχει το όνομα, ή τα ονόματα, του συγκεκριμένου μέλους ΔΕΠ όπως αυτά αναγράφονται στο csv αρχείο. Το χαρακτηριστικό αυτό είναι πολλαπλών τιμών, διότι κάποιο μέλος ΔΕΠ μπορεί σε ορισμένες δημοσιεύσεις να έχει γράψει, το πατρώνυμο εκτός από το όνομα και το επώνυμο του, ενώ σε κάποιες άλλες να έχει γράψει μόνο το όνομα και το επώνυμο του. Η μορφή με την οποία αναγράφονται τα ονόματα στο αρχείο csv περιγράφεται στην σελίδα 40.

Στην οντότητα `Person` αποθηκεύονται τα ονόματα των συγγραφέων κάθε δημοσίευσης όπως αυτά αναγράφονται στο αρχείο csv. Στην οντότητα αυτή έχουν εκχωρηθεί τα επόμενα χαρακτηριστικά.

- **id**
Μοναδικό προσδιοριστικό κάθε συγγραφέα (πρωτεύων κλειδί)

- **name**

Το όνομα του συγγραφέα.

Η οντότητα `di_professor` συνδέεται με την οντότητα `Person` μέσω της συσχέτισης `Belongs` με λόγο πολλαπλότητας «ένα προς ένα» (1:1) αφού κάθε μέλος ΔΕΠ ανήκει στους συγγραφείς, αλλά και κάποιος από τους συγγραφείς μιας δημοσίευσης μπορεί να είναι μέλος ΔΕΠ. Η διακεκομμένη γραμμή με την όποια η συσχέτιση `Belongs` συνδέεται με την οντότητα `Person` υποδηλώνει ότι η συσχέτιση είναι μερικής συμμετοχής από το μέρος του `Person`. Αυτό σημαίνει ότι ορισμένοι (όχι όλοι) οι αντιπρόσωποι της οντότητας `Person` συμμετέχουν στην συσχέτιση, αφού ενδέχεται στην συγγραφή μιας δημοσίευσης να έχουν λάβει μέρος και άτομα που δεν είναι μέλη ΔΕΠ.

Η οντότητα `Publication` χρησιμοποιείται για να αποθηκευτούν τα στοιχεία κάθε δημοσίευσης στην βάση δεδομένων. Τα χαρακτηριστικά της οντότητας αυτής είναι:

- **id**

Μοναδικό προσδιοριστικό κάθε δημοσίευσης (πρωτεύων κλειδί).

- **title**

Ο τίτλος που έδωσαν οι συγγραφείς στη δημοσίευση.

- **year**

Το έτος που έγινε η δημοσίευση.

- **page_start**

Η σελίδα του περιοδικού, βιβλίου ή πρακτικών συνεδρίου από την οποία ξεκινάει η συγκεκριμένη δημοσίευση.

- **page_end**

Η σελίδα του περιοδικού, βιβλίου ή πρακτικών συνεδρίου στην οποία τελειώνει η συγκεκριμένη δημοσίευση.

- **cited_by**

Ο αριθμός των αναφορών που έχει λάβει από άλλες δημοσιεύσεις.

Η οντότητα `Publication` συνδέεται με την οντότητα `Person` μέσω της συσχέτισης `Participate` με λόγο πολλαπλότητας «πολλά προς πολλά» (N:M) αφού κάθε άτομο μπορεί να έχει συμμετοχή σε πολλές δημοσιεύσεις, αλλά και κάθε δημοσίευση να έχει πολλούς συγγραφείς.

Κάθε δημοσίευση αποτελείται από τις κατηγορίες `Journal article`, `Conference article`, `Book` ή `Book chapter`. Προκύπτουν έτσι οι οντότητές `tuypu is-a`, `Journal article`, `Conference article`, `Book chapter` και `Book`, που είναι υποσύνολα της οντότητας `Publication` και επομένως κληρονομούν τα χαρακτηριστικά της, αφού κάθε `Journal article`, `Conference article`, `Book` και `Book chapter` είναι ένα `Publication`.

Καθεμία από τις οντότητες αυτές, έκτος από τα γνωρίσματα που κληρονομεί από την *Publication*, μπορεί να έχει τα δικά της χαρακτηριστικά, που είναι απαραίτητο να αποθηκευτούν στην βάση δεδομένων, ή να συμμετέχει σε επιπλέον συσχετίσεις και για αυτό το λόγο έγιναν ξεχωριστές οντότητες.

Στην οντότητα *Book* αποθηκεύονται τα στοιχεία των βιβλίων στα οποία κάποιο μέλος ΔΕΠ έχει λάβει μέρος στην συγγραφή τους. Η οντότητα αυτή, όπως αναφέρθηκε, κληρονομεί όλα τα χαρακτηριστικά της οντότητας *Publication* και έχει δυο επιπλέον δικά της χαρακτηριστικά που είναι:

- **isbn**
Ο διεθνής αναγνωριστικός αριθμός εμπορικού βιβλίου.
- **publisher**
Ο εκδότης του βιβλίου.

Η οντότητα *Book chapter* δημιουργήθηκε για να αποθηκευτούν δεδομένα δημοσιεύσεων που είναι κεφάλαια σε βιβλία. Εκτός από τα χαρακτηριστικά που κληρονομεί από την οντότητα *Publication* έχει τα επόμενα τρία ακόμη γνωρίσματα:

- **bookisbn**
Ο διεθνής αναγνωριστικός αριθμός εμπορικού βιβλίου στο οποίο το συγκεκριμένο κεφάλαιο ανήκει.
- **bookpublisher**
Ο εκδότης του βιβλίου.
- **booktitle**
Ο τίτλος του βιβλίου.

Στην οντότητα *Journal article* αποθηκεύονται δεδομένα άρθρων σε επιστημονικά περιοδικά. Η οντότητα αυτή κληρονομεί τα γνωρίσματα της οντότητας *Publication*, δεν έχει όμως δικά της ξεχωριστά χαρακτηριστικά. Συνδέεται με την οντότητα *Journal* μέσω της συσχέτισης *Published* με λόγω πολλαπλότητας «ένα προς πολλά» (1:N), αφού έγινε η παραδοχή ότι ένα άρθρο δημοσιεύεται σε ένα μόνο περιοδικό, αλλά στο ίδιο περιοδικό είναι δυνατόν να δημοσιευτούν πολλά άρθρα.

Η συσχέτιση *Published* έχει τα επόμενα δυο γνωρίσματα:

- **volume**
Ο τόμος του περιοδικού στο οποίο το συγκεκριμένο άρθρο δημοσιεύεται.
- **issue**
Το τεύχος του περιοδικού στο οποίο το άρθρο έχει δημοσιευτεί.

Τα γνωρίσματα *volume* και *issue* είναι σύγχρονως γνωρίσματα και του *Journal* και του *Journal article*. Αυτό γίνεται περισσότερο σαφές από

το γεγονός ότι ένα άρθρο δημοσιεύεται σε ένα συγκεκριμένο τεύχος που μπορεί να ανήκει σε κάποιον τόμο ενός περιοδικού. Εξάλλου για να μπορέσουμε να προσδιορίσουμε το άρθρο χρειάζεται να γνωρίζουμε σε πιο τεύχος και σε ποιον τόμο του περιοδικού έχει δημοσιευτεί. Γίνεται λοιπόν φανερό ότι τα γνωρίσματα αυτά δεν ανήκουν σε μια οντότητα αλλά μοιράζονται στις δύο οντότητες, επομένως τα γνωρίσματα αυτά είναι της συσχέτισης που υπάρχει ανάμεσα στις δυο οντότητες.

Στην οντότητα *Journal* έχουν εκχωρηθεί τα χαρακτηριστικά που απαιτούνται για να αποθηκευτούν τα δεδομένα των περιοδικών στα οποία δημοσιεύονται επιστημονικά άρθρα. Τα χαρακτηριστικά αυτά είναι:

- **id**
Μοναδικό προσδιοριστικό κάθε περιοδικού (πρωτεύον κλειδί).
- **name**
Το όνομα του περιοδικού.

Η οντότητα *Conference article* χρησιμοποιείται για να αποθηκευτούν δεδομένα επιστημονικών άρθρων που παρουσιάζονται σε συνέδρια και δημοσιεύονται στα πρακτικά τους. Και αυτή η οντότητα δεν έχει δικά της γνωρίσματα αλλά κληρονομεί τα χαρακτηριστικά της οντότητας *Publication*. Επιπλέον συνδέεται με την οντότητα *Conference* μέσω της συσχέτισης *Hosted* η οποία έχει λόγο πολλαπλότητας «πολλά προς ένα», εξαιτίας της θεώρησης ότι ένα άρθρο παρουσιάζεται σε ένα μόνο συνέδριο, αλλά στο ίδιο συνέδριο μπορούν να παρουσιαστούν πολλά άρθρα.

Η οντότητα *Conference* χρησιμοποιείται για να αποθηκευτούν δεδομένα συνεδρίων στα οποία παρουσιάζονται επιστημονικές δημοσιεύσεις. Τα χαρακτηριστικά που εκχωρούνται στην οντότητα αυτή είναι τα επόμενα:

- **id**
Μοναδικό προσδιοριστικό κάθε συνεδρίου (πρωτεύον κλειδί).
- **name**
Το όνομα του συνεδρίου.
- **date**
Η ημερομηνία διεξαγωγής του συνεδρίου.
- **location**
Η τοποθεσία διεξαγωγής του συνεδρίου.

Στην οντότητα *Proceedings* εκχωρούνται τα χαρακτηριστικά που είναι απαραίτητα για να αποθηκευθούν τα στοιχεία των πρακτικών των συνεδρίων στα οποία δημοσιεύονται τα άρθρα που παρουσιάζονται σε αυτά. Τα χαρακτηριστικά της εν λόγω οντότητας είναι:

- **id**
Μοναδικό προσδιοριστικό κάθε πρακτικού (πρωτεύον κλειδί).

- **name**

Το όνομα των πρακτικών.

Η οντότητα **Conference** συνδέεται με την οντότητα **Proceedings** μέσω της συσχέτισης **Has** με λόγω πολλαπλότητας «πολλά προς ένα», διότι σε περιδικά πρακτικών συνεδρίων με το ίδιο όνομα μπορούν να δημοσιευτούν άρθρα που έχουν παρουσιαστεί σε διαφορετικά συνέδρια. Η συσχέτιση **Has** έχει μόνο ένα χαρακτηριστικό που είναι το εξής:

- **volume**

Ο τόμος των πρακτικών.

Ομοίως με τη συσχέτιση **Published**, το γνώρισμα **volume** είναι σύγχρονος γνώρισμα και του **Conference** και του **Proceedings**, διότι για να προσδιοριστεί ένα άρθρο δημοσιευμένο σε πρακτικά συνεδρίου, χρειάζεται να γνωρίζουμε σε ποιο τόμο των πρακτικών έχει δημοσιευτεί.

Τέλος στο Διάγραμμα Οντοτήτων – Συσχετίσεων περιέχεται και η οντότητα **Administrator** η οποία χρησιμοποιείται για να αποθηκευθούν στη βάση δεδομένων το όνομα χρήστη και ο κωδικός πρόσβασης των διαχειριστών της. Η οντότητα αυτή δεν συνδέεται με καμιά άλλη αφού δημιουργήθηκε καθαρά για τεχνικούς λόγους. Τα χαρακτηριστικά της οντότητας αυτής είναι:

- **username**

Το όνομα χρήστη του διαχειριστή (πρωτεύον κλειδί).

- **password**

Ο κωδικός πρόσβασης του διαχειριστή.

3.3 Το σχεσιακό σχήμα της βάσης δεδομένων

Για να υλοποιηθεί μια βάση δεδομένων σε ηλεκτρονικό υπολογιστή απαιτείται η μετατροπή του Διαγράμματος Οντοτήτων – Συσχετίσεων που την περιγράφει, σε σχεσιακό σχήμα.

Το σχεσιακό μοντέλο δεδομένων είναι το πρωτεύον μοντέλο δεδομένων, το οποίο χρησιμοποιείται ευρέως σε όλο τον κόσμο για αποθήκευση και επεξεργασία δεδομένων. Αυτό το μοντέλο είναι απλό και διαθέτει όλες τις ιδιότητες και τις δυνατότητες που απαιτούνται για την επεξεργασία δεδομένων με αποδοτικότητα αποθήκευσης. Η απλότητα του οφείλεται στη μια και μοναδική δομή που υιοθέτει για την αποθήκευση δεδομένων η οποία δεν είναι άλλη από τον πίνακα.

Στην ενότητα αυτή παρουσιάζεται αναλυτικά ο τρόπος με τον οποίο το Διάγραμμα Οντοτήτων – Συσχετίσεων της σελίδας 18 μετατρέπεται σε σχήμα σχεσιακής βάσης δεδομένων.

Η μετατροπή ενός ΔΟΣ σε σχήμα σχεσιακής βάσης δεδομένων γίνεται σε δύο στάδια. Κατά το πρώτο στάδιο μετατρέπονται σε σχήμα σχεσιακής βάσης

δεδομένων οι κύριες και δευτερεύουσες τάξεις του ΔΟΣ, ενώ στο δεύτερο στάδιο μετατρέπονται σε σχήμα σχεσιακής βάσης δεδομένων τα υπόλοιπα στοιχεία του ΔΟΣ, δηλαδή οι ισχυρές και οι ασθενής οντότητες, τα γνωρίσματα πολλαπλών τιμών και οι συσχετίσεις.

Κατά την μετατροπή των κύριων και δευτερευουσών τάξεων του ΔΟΣ, που περιγράφεται στην προηγούμενη ενότητα, ακολουθήθηκε η προσέγγιση να δημιουργηθεί μια σχέση για κάθε κύρια τάξη και μια σχέση για κάθε δευτερεύουσα τάξη. Κάθε σχέση που δημιουργείται έχει ως χαρακτηριστικά τα γνωρίσματα της τάξης στην οποία αντιστοιχεί. Στις σχέσεις που αντιστοιχούν σε δευτερεύουσες τάξεις προστίθεται ως πρωτεύον κλειδί το πρωτεύον κλειδί της σχέσης που αντιστοιχεί στην κύρια τάξη που είναι γονέας τους δηλαδή το `id`, το οποίο σε κάθε δευτερεύουσα τάξη μετονομάζεται σε `pub_id`.

Έτσι, κατά την εφαρμογή του πρώτου σταδίου, για το ΔΟΣ της εικόνας 3.1 προκύπτουν οι σχέσεις (πίνακες) με τα χαρακτηριστικά που φαίνονται παρακάτω.

Publications (<u>id</u> , title, year, page_start, page_end, cited_by) Journal_articles (<u>pub_id</u>) Conference_articles (<u>pub_id</u>) Books (<u>pub_id</u> , isbn, publisher) Book_chapters (<u>pub_id</u> , booktitle, bookisbn, bookpublisher)
--

Εικόνα 3.2: Οι σχέσεις που προκύπτουν από την μετατροπή των κύριων και δευτερευουσών τάξεων σε σχεσιακό σχήμα

Η εφαρμογή του δεύτερου σταδίου γίνεται σε πέντε βήματα. Στο πρώτο βήμα δημιουργείται μια σχέση για κάθε ισχυρή οντότητα του ΔΟΣ. Τα χαρακτηριστικά της σχέσης αυτής είναι τα απλά χαρακτηριστικά της οντότητας. Το πρωτεύον κλειδί της σχέσης είναι το πρωτεύον γνώρισμα της οντότητας. Το ΔΟΣ της εικόνας 3.1 περιέχει τις ισχυρές οντότητες `di_professor`, `Person`, `Administrator`, `Journal`, `Conference` και `Proceedings` επομένως δημιουργούνται οι επόμενες σχέσεις:

di_professors (<u>id</u> , firstname, lastname, property, csv, image, updated) Persons (<u>id</u> , name) Administrator (<u>username</u> , password) Journals (<u>id</u> , name) Conferences (<u>id</u> , name, date, location) Proceedings (<u>id</u> , name)

Εικόνα 3.3: Οι σχέσεις που προκύπτουν από τις ισχυρές οντότητες

Στο δεύτερο βήμα μετατρέπονται τα γνωρίσματα πολλαπλών τιμών του ΔΟΣ σε σχέσεις του σχεσιακού μοντέλου. Με την διαδικασία αυτή δημιουργε-

ίται μια νέα σχέση για κάθε γνώρισμα πολλαπλών τιμών του ΔΟΣ. Τα χαρακτηριστικά της νέας σχέσης είναι το γνώρισμα πολλαπλών τιμών και το πρωτεύον γνώρισμα της οντότητας στην οποία ανήκει το γνώρισμα πολλαπλών τιμών. Το πρωτεύον κλειδί της νέας σχέσης απαρτίζεται από όλα τα χαρακτηριστικά της σχέσης που δημιουργήθηκε στο βήμα αυτό. Το χαρακτηριστικό της σχέσης που αντιστοιχεί στο πρωτεύον γνώρισμα της οντότητας που έχει το χαρακτηριστικό πολλαπλών τιμών ορίζεται ξένο κλειδί.

Στο ΔΟΣ της εικόνας 3.1 υπάρχει μόνο ένα γνώρισμα πολλαπλών τιμών το οποίο είναι το γνώρισμα `scopus_names` της οντότητας `di_professor`. Έτσι δημιουργείται η σχέση `prof_names` που έχει ως χαρακτηριστικά το γνώρισμα πολλαπλών τιμών `name` και το πρωτεύον κλειδί της οντότητας `di_professors` (που είναι το `id`) όπου, για λόγους αποφυγής σύγχυσης, στη σχέση αυτή του δόθηκε η ονομασία `prof_id`. Το πρωτεύον κλειδί της σχέσης `prof_names` είναι σύνθετο και απαρτίζεται από όλα τα χαρακτηριστικά της σχέσης. Τέλος το `prof_id` της σχέσης `prof_names` είναι ταυτόχρονα ξένο κλειδί και αναφέρεται στο πρωτεύον κλειδί της σχέσης `di_professors` (δηλαδή στο `id`). Η σχέση `prof_names` και τα χαρακτηριστικά της φαίνονται στην επόμενη εικόνα.

prof_names (<u>prof_id</u> , <u>name</u>)
--

Εικόνα 3.4: Οι σχέσεις που προκύπτουν από τα γνωρίσματα πολλαπλών τιμών

Το τρίτο βήμα αναπαριστάνει τις ένα προς ένα συσχετίσεις του ΔΟΣ ως ξένα κλειδιά τα οποία προστίθενται ως χαρακτηριστικά σε υπάρχουσες σχέσεις της σχεσιακής βάσης δεδομένων. Προστίθεται έτσι ένα ξένο κλειδί για κάθε 1:1 συσχέτιση. Ως ξένο κλειδί επιλέγεται το πρωτεύον γνώρισμα μιας από τις οντότητες που συμμετέχει στην 1:1 συσχέτιση. Αυτό προστίθεται στη σχέση που αντιστοιχεί στην άλλη οντότητα που συμμετέχει στην 1:1 συσχέτιση. Στην ίδια σχέση προστίθεται ως χαρακτηριστικό και κάθε γνώρισμα της συσχέτισης.

Η μόνη 1:1 συσχέτιση του ΔΟΣ της εικόνας 3.1 είναι η συσχέτιση `Belongs` στην οποία συμμετέχουν οι οντότητες `Person` και `di_professor`. Επιλέγεται η οντότητα `di_professor` η οποία έχει ως πρωτεύον γνώρισμα το `id`. Αυτό το γνώρισμα προστίθεται ως ξένο κλειδί στη σχέση που αντιστοιχεί στην οντότητα `Person`. Έτσι η σχέση `Persons` εμπλουτίζεται με ένα πρόσθετο χαρακτηριστικό που είναι το `id` του μέλους ΔΕΠ και για αυτόν το λόγο στο χαρακτηριστικό αυτό δίνεται η ονομασία `prof_id`. Σύμφωνα με τα προηγούμενα, η σχέση `Persons` της εικόνας 3.3 παίρνει τώρα τη μορφή που φαίνεται παρακάτω.

Persons (<u>id</u> , <u>name</u> , <u>prof_id</u>)

Εικόνα 3.5: Η σχέση `Persons` μετά την προσθήκη ξένου κλειδιού

Η σχέση `Belongs` είναι μερικής συμμετοχής από το μέρος του `Persons`. Αυτό σημαίνει ότι ορισμένα (όχι όλα) τα μέλη της σχέσης `Persons` ανήκουν

στη σχέση `di_professors`. Άρα όσα μέλη της σχέσης `Persons` δεν ανήκουν στη σχέση `di_professors` θα έχουν στο χαρακτηριστικό `prof_id` την τιμή `NULL`. Επομένως η σχέση `Persons` θα έχει πολλές πλειάδες οι οποίες θα έχουν την τιμή `NULL`, στο χαρακτηριστικό `prof_id`.

Σε πολλές περιπτώσεις το να περιέχει μια σχέση μεγάλο αριθμό πλειάδων οι οποίες θα έχουν την τιμή `NULL` σε κάποιο χαρακτηριστικό, δεν είναι καλή πρακτική για την δημιουργία μιας βάσης δεδομένων. Στην προκειμένη περίπτωση όμως η συγκεκριμένη λύση είναι η μόνη δυνατή. Αυτό οφείλεται στο γεγονός ότι ορισμένα μέλη ΔΕΠ αναγράφονται στο αρχείο `csv` με δύο η περισσότερα διαφορετικά ονόματα και επομένως αυτός είναι ο μόνος τρόπος με τον οποίο θα μπορούσε να γίνει σαφές αν ένα μέλος της σχέσης `Persons` ανήκει και στη σχέση `di_professors`. Δηλαδή να είναι σε θέση η εφαρμογή να προσδιορίσει αν κάποιος από τους συγγραφείς μιας δημοσίευσης είναι ή όχι μέλος ΔΕΠ του Τμήματος Πληροφορικής του Ιονίου Πανεπιστημίου.

Το τέταρτο βήμα αναπαριστά τις 1:N συσχετίσεις του ΔΟΣ ως ξένα κλειδιά που προστίθενται σε ήδη υπάρχουσες σχέσεις της σχεσιακής βάσης δεδομένων. Με την διαδικασία αυτή προστίθεται ένα ξένο κλειδί για κάθε 1:N συσχέτιση. Ως ξένο κλειδί επιλέγεται το γνώρισμα της οντότητας που συμμετέχει στη συσχέτιση από το μέρος του 1. Αυτό προστίθεται ως ξένο κλειδί στη σχέση που αντιστοιχεί στην οντότητα που συμμετέχει στη συσχέτιση από το μέρος του N. Στην ίδια σχέση προστίθεται κάθε γνώρισμα της συσχέτισης.

Στο ΔΟΣ υπάρχουν οι εξής 1:N συσχετίσεις `Published`, `Hosted` και `Has` (Εικόνα 3.1). Για κάθε μία από αυτές εφαρμόζεται το τέταρτο βήμα ως ακολούθως:

- Στη συσχέτιση `Published` συμμετέχουν οι οντότητες `Journal_article` από το μέρος του N και `Journal` από το μέρος του 1. Επιλέγεται η οντότητα που συμμετέχει από το μέρος του 1, δηλαδή η οντότητα `Journal`. Το πρωτεύον γνώρισμα της οντότητας `Journal` είναι το `id`. Αυτό προστίθεται ως ξένο κλειδί στη σχέση `Journal_articles` (ως `journal_id`) αφού αυτή είναι η σχέση η οποία αντιστοιχεί στην οντότητα που συμμετέχει στη συσχέτιση `Published` από μέρος του N. Επιπλέον στη σχέση `Journal_articles` προστίθενται και τα γνωρίσματα της συσχέτισης `Published` δηλαδή τα `volume` και `issue`.
- Στη συσχέτιση `Hosted` συμμετέχουν οι οντότητες `Conference_article` από το μέρος του N και `Conference` από το μέρος του 1. Επιλέγεται η οντότητα που συμμετέχει από το μέρος του 1, δηλαδή η οντότητα `Conference`. Το πρωτεύον γνώρισμα της οντότητας `Conference` είναι το `id`. Αυτό προστίθεται ως ξένο κλειδί στη σχέση `Conference_articles` (ως `conf_id`) αφού αυτή είναι η σχέση η οποία αντιστοιχεί στην οντότητα που συμμετέχει στη συσχέτιση `Hosted` από μέρος του N.
- Στη συσχέτιση `Has` συμμετέχουν οι οντότητες `Conference` από το μέρος του N και `Proceedings` από το μέρος του 1. Επιλέγεται η οντότητα

που συμμετέχει από το μέρος του 1, δηλαδή η οντότητα Proceedings. Το πρωτεύον γνώρισμα της οντότητας Proceedings είναι το `id`. Αυτό προστίθεται ως ξένο κλειδί στη σχέση Conferences (ως `proc_id`) αφού αυτή είναι η σχέση η οποία αντιστοιχεί στην οντότητα που συμμετέχει στη συσχέτιση Has από μέρος του N. Επιπλέον στην σχέση Conferences προστίθενται και το γνώρισμα της συσχέτισης Has δηλαδή το `volume`.

Στην εικόνα 3.6 παρουσιάζονται οι σχέσεις `Journal_Articles`, `Conference_Articles` και `Conferences`, όπως αυτές προκύπτει από την εφαρμογή της προηγούμενης διαδικασίας.

Journal_articles (<u>pub_id</u> , volume, issue, journal_id) Conference_articles (<u>pub_id</u> , conf_id) Conferences (<u>id</u> , name, date, location, volume, proc_id)
--

Εικόνα 3.6: Οι σχέσεις `Journal_articles`, `Conference_articles` και `Conferences` όπως προκύπτουν από την εφαρμογή του τέταρτου βήματος

Το πέμπτο βήμα αναπαριστά κάθε N:N συσχέτιση του ΔΟΣ ως μια νέα σχέση που προστίθεται στη σχεσιακή βάση δεδομένων. Κατά την εφαρμογή του βήματος αυτού προστίθεται μια σχέση για κάθε N:N συσχέτιση. Η νέα σχέση περιλαμβάνει ως χαρακτηριστικά τα πρωτεύοντα κλειδιά των οντοτήτων που συμμετέχουν στη N:N συσχέτιση. Αυτά αποτελούν και το σύνθετο πρωτεύον κλειδί της σχέσης. Στη σχέση αυτή προστίθεται και ως χαρακτηριστικό κάθε γνώρισμα της συσχέτισης.

Η μοναδική N:N συσχέτιση του ΔΟΣ της εικόνας 3.1 είναι η συσχέτιση Participate. Για τη συσχέτιση αυτή δημιουργείται μια νέα σχέση στη σχεσιακή βάση δεδομένων και της δίνεται το όνομα Authorships. Η συσχέτιση Participate συσχετίζει την οντότητα Publication (που έχει πρωτεύον γνώρισμα το `id`) με την οντότητα Person (που και αυτή έχει πρωτεύον γνώρισμα το `id`). Επομένως στη σχέση Authorships προστίθενται ως χαρακτηριστικά αυτά τα πρωτεύοντα ορίσματα με ονομασία `pub_id` και `pers_id` αντίστοιχα. Τα δυο αυτά χαρακτηριστικά μαζί αποτελούν το σύνθετο πρωτεύον κλειδί της σχέσης Authorships. Καθένα από αυτά τα χαρακτηριστικά είναι από μόνο του ξένο κλειδί. Το χαρακτηριστικό `pub_id` αναφέρεται στο χαρακτηριστικό `id` της σχέσης Publications ενώ το χαρακτηριστικό `pers_id` αναφέρεται στο χαρακτηριστικό `id` της σχέσης Persons.

Authorships (<u>pub_id</u> , <u>pers_id</u>)

Εικόνα 3.7: Η σχέση `Authorships` που προκύπτει από την εφαρμογή του πέμπτου βήματος

Με την εφαρμογή του πέμπτου βήματος ολοκληρώνεται η διαδικασία μετατροπής του ΔΟΣ σε σχήμα σχεσιακής βάσης δεδομένων. Προκύπτει έτσι το

τελικό σχήμα που φαίνεται στην επόμενη εικόνα.

```

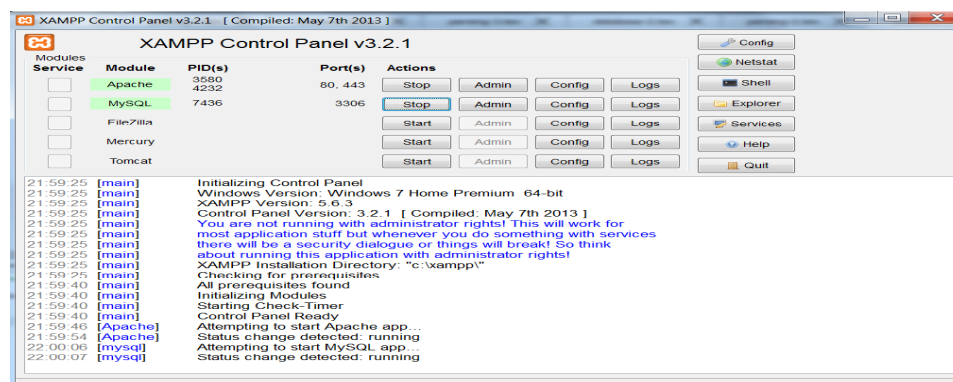
Publications (id, title, year, page_start, page_end, cited_by)
Journal_articles (pub_id, volume, issue, journal_id)
Conference_articles (pub_id, conf_id)
Books (pub_id, isbn, publisher)
Book_chapters (pub_id, booktitle, bookisbn, bookpublisher)
di_professors (id, firstname, lastname, property, csv, image, updated)
Persons (id, name, prof_id)
Administrator (username, password)
Journals (id, name)
Conferences (id, name, date, location, volume, proc_id)
Proceedings (id, name)
prof_names (prof_id, name)
Authorships (pub_id, pers_id)

```

Εικόνα 3.8: Το τελικό σχεσιακό σχήμα της βάσης δεδομένων

3.4 Κατασκευή της βάσης δεδομένων

Στην ενότητα αυτή παρουσιάζεται ο τρόπος κατασκευής της βάσης δεδομένων σύμφωνα με το σχεσιακό σχήμα που φαίνεται στην εικόνα 3.8. Η κατασκευή της βάσης δεδομένων θα γίνει στο περιβάλλον phpMyAdmin του πακέτου προγραμμάτων λογισμικού XAMPP.



Εικόνα 3.9: Ο πίνακας ελέγχου του XAMPP με τον Apache και την MySQL σε λειτουργία

Το XAMPP

Το XAMPP είναι ένα πακέτο προγραμμάτων ελεύθερου λογισμικού, ανοικτού κώδικα και ανεξαρτήτου πλατφόρμας το οποίο περιέχει τον εξυπηρετητή ιστοσελίδων http Apache, που μπορεί να εξυπηρετεί δυναμικές ιστοσελίδες τε-

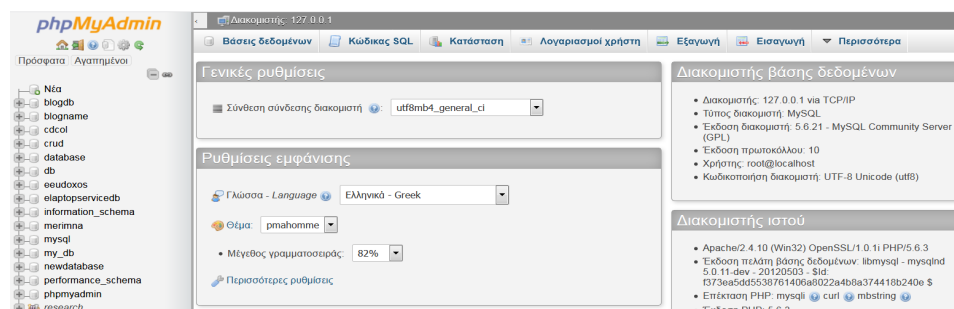
χνολογίας PHP/MySQL, την βάση δεδομένων MySQL και ένα διερμηνέα για κώδικα γραμμένο σε γλώσσες προγραμματισμού PHP και Perl. Το XAMPP τρέχει σε Microsoft Windows, Linux, Solaris, και Mac OS X και χρησιμοποιείται ως πλατφόρμα για τη σχεδίαση και ανάπτυξη ιστοσελίδων με τεχνολογίες όπως PHP, JSP και Servlets. Επίσης το XAMPP μπορεί να χρησιμοποιηθεί ως εργαλείο ανάπτυξης και δοκιμής ιστοσελίδων τοπικά στον υπολογιστή χωρίς να είναι απαραίτητη η σύνδεση στο διαδίκτυο. Στην πράξη το XAMPP ορισμένες φορές χρησιμοποιείται και για τη φιλοξενία ιστοσελίδων.

Η MySQL

Η MySQL είναι το πιο δημοφιλές σύστημα διαχείρισης σχεσιακών βάσεων δεδομένων ανοιχτού κώδικα, και διατίθεται δωρεάν για πολλά λειτουργικά συστήματα. Η MySQL εκτελείται ως διεργασία του λειτουργικού συστήματος που αναμένει εντολές και αιτήματα από το χρήστη. Για αυτό τον λόγο αναφέρεται και ως εξυπηρετητής ή διακομιστής (server). Ο χρήστης για να επικοινωνήσει με την MySQL, δηλαδή να υποβάλει σε αυτήν τα αιτήματα του και να παραλάβει της αποκρίσεις της, χρησιμοποιεί προγράμματα τα οποία συνεργάζονται με την MySQL και είναι γνωστά ως πελάτες.

Το phpMyAdmin

Το phpMyAdmin είναι ένα ανοιχτού κώδικα λογισμικό γραμμένο κυρίως στην γλώσσά PHP. Συνεργάζεται με την MySQL ως πρόγραμμα πελάτης και χρησιμοποιείται ως εργαλείο διαχείρισης. Το phpMyAdmin καλείται μέσω ενός φυλλομετρητή ιστού, πληκτρολογώντας στο πεδίο κειμένου την διεύθυνση του διακομιστή ακολουθούμενη από το λεκτικό «phpMyAdmin». Στην επόμενη εικόνα φαίνεται η αρχική οθόνη του phpMyAdmin.



Εικόνα 3.10: Η αρχική οθόνη του phpMyAdmin

Δημιουργία της βάσης δεδομένων

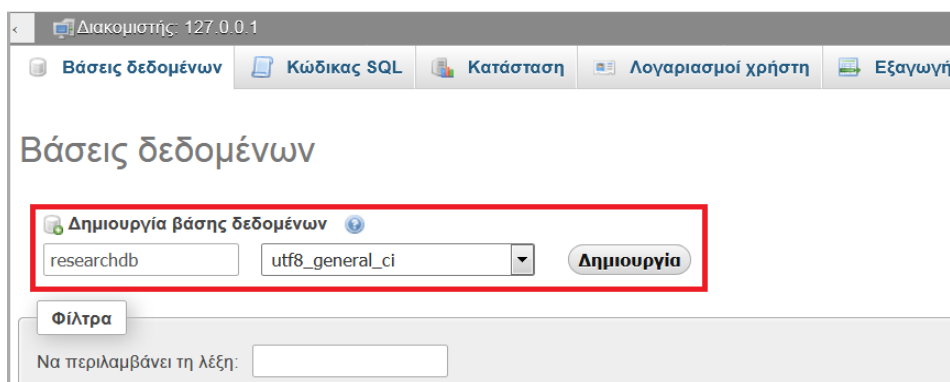
Για την δημιουργία μιας νέας βάσης δεδομένων στο phpMyAdmin δίνουμε την επιλογή «Βάσεις Δεδομένων» από το μενού της αρχικής οθόνης.

Στο παράθυρο που εμφανίζεται (Εικόνα 3.12) πληκτρολογούμε στο πλαίσιο κειμένου του όνομα της βάσης δεδομένων που θα δημιουργήσουμε (researchdb), επιλέγουμε τη σύνθεση (utf8mb4_general_ci) και πατάμε το κουμπί «Δη-



Εικόνα 3.11: Το μενού της αρχικής οθόνης του *phpMyAdmin*

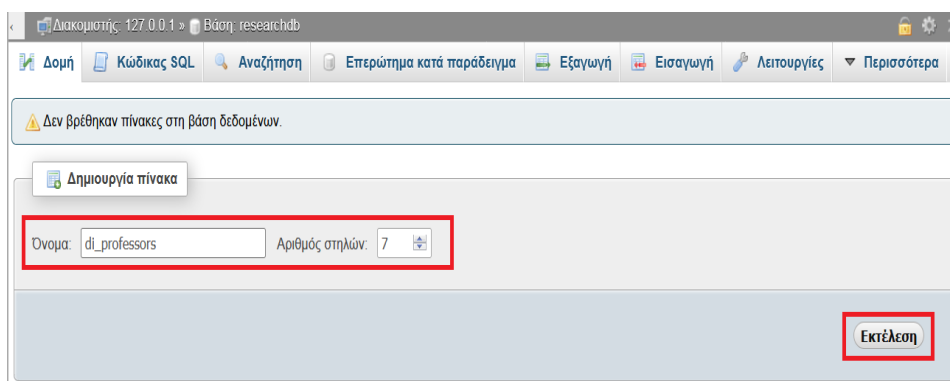
μιουργία». Η σύνθεση `utf8mb4_general_ci` υποστηρίζει ελληνικούς και αγγλικούς χαρακτήρες και δεν κάνει διάκριση πεζών κεφαλαίων (το `ci` σημαίνει *case insensitive*).



Εικόνα 3.12: Δημιουργία της βάσης δεδομένων

3.5 Δημιουργία πινάκων της βάσης δεδομένων

Για να δημιουργήσουμε σε μια βάση δεδομένων έναν νέο πίνακα, επιλέγουμε τη βάση δεδομένων και οδηγούμαστε στην οθόνη δημιουργίας πίνακα (Εικόνα 3.13). Στην οθόνη αυτή δηλώνουμε το όνομα του νέου πίνακα (στο παράδειγμα `di_professors`), τον αριθμό των στηλών του πίνακα (στο παράδειγμα 7) και πατάμε το κουμπί «Εκτέλεση».



Εικόνα 3.13: Δημιουργία πίνακα

Στη συνέχεια εμφανίζεται το παράθυρο ορισμού των στηλών στο οποίο περιγράφεται με λεπτομέρεια κάθε στήλη του πίνακα δηλώνοντας υποχρεωτικά τουλάχιστον το όνομα της στήλης και τον τύπο των δεδομένων της. Έτσι για τον πίνακα `di_professors`, σύμφωνα με το τελικό σχεσιακό σχήμα της εικόνας 3.8, δημιουργούνται οι στήλες `id`, `firstname`, `lastname`, `property`, `csv`, `image`, και `updated`. Για καθεμία από τις προηγούμενες στήλες επιλέγουμε από την επόμενη οθόνη τα χαρακτηριστικά που της αντιστοιχούν και πατάμε το κουμπί αποθήκευση για να δημιουργηθεί ο πίνακας στη βάση δεδομένων.

The screenshot shows the 'Create Table' dialog in phpMyAdmin for a table named 'di_professors'. The dialog has several tabs at the top: 'Περιήγηση', 'Δομή', 'Κώδικας SQL', 'Αναζήτηση', 'Προσθήκη', 'Εξαγωγή', 'Εισαγωγή', 'Διακρίματα', 'Λειτουργίες', 'Παρακολούθηση', and 'Δείκτες'. The 'Δομή' (Structure) tab is active. Below the tabs, there's a section for 'Ονομασία πίνακα:' (Table name) with 'di_professors' and 'Προσθήκη' (Add) button. The main area is a table with columns: 'id', 'firstname', 'lastname', 'property', 'csv', 'image', and 'updated'. Each row has dropdowns for 'Τύπος' (Type), 'Μήκος/Τιμές' (Length/Values), 'Προεπιλογή' (Default), 'Σύνθεση' (Collation), 'Χαρακτηριστικά' (Flags), 'Κενό' (Empty), 'Ευρετήριο' (Index), and 'A_I' (Auto Increment). The 'id' row has 'INT' type, length '11', 'PRIMARY' checked, and 'A_I' checked. The 'updated' row has 'TIMESTAMP' type and 'A_I' checked. At the bottom, there's a section for 'Σχόλια πίνακα:' (Table comments), 'Σύνθεση:' (Collation), and 'Μηχανή αποθήκευσης:' (Storage engine) set to 'InnoDB'. There are also 'Προεπισκόπηση SQL' (Preview SQL) and 'Αποθήκευση' (Save) buttons at the bottom right.

Εικόνα 3.14: Ορισμός πεδίων πίνακα

Δήλωση πρωτεύοντος κλειδιού

Για να δηλώσουμε μια στήλη του πίνακα ως πρωτεύον κλειδί επιλέγουμε στο ευρετήριο από την λίστα την τιμή `PRIMARY`. Σε περίπτωση που το πρωτεύον κλειδί είναι σύνθετο επιλέγουμε την τιμή `PRIMARY` σε όλα τα πεδία που συμμετέχουν στο πρωτεύον κλειδί.

Δήλωση πεδίου αυτόματης αρίθμησης

Ορισμένες φορές είναι βολικό κάποια πεδία ενός πίνακα να παίρνουν αυτόματα τιμές. Οι τιμές που δίνονται σε αυτά τα πεδία είναι διαδοχικοί ακέραιοι αριθμοί. Για να γίνει αυτόματη απόδοση μιας ακέραιας τιμής σε ένα πεδίο ενεργοποιούμε την επιλογή `A_I` (`Auto Increment`) που εμφανίζεται στο παράθυρο ορισμού πεδίων.

Απόδοση της τιμής `NULL`

Στην `phpMyAdmin` η προεπιλεγμένη τιμή για τα πεδία ενός πίνακα είναι να μην παίρνουν την τιμή `NULL`. Αν ο χρήστης επιθυμεί να επιτρέπεται σε ένα πεδίο να δέχεται την τιμή `NULL` το δηλώνει στη στήλη «Κενό».

Σύμφωνα με την προηγούμενη διαδικασία δημιουργούνται και οι υπόλοιποι πίνακες της βάσης δεδομένων. Ορίζονται οι στήλες τους καθώς επίσης και τα χαρακτηριστικά που σε καθεμία από αυτές αντιστοιχούν, τα όποια φαίνονται στις εικόνες που ακολουθούν.

#	Όνομα	Τύπος	Σύνθεση	Χαρακτηριστικά	Κενό	Προεπιλογή	Σχόλια	Πρόσθετα
1	username 🗝️	varchar(100)	utf8mb4_general_ci		Όχι	Καμία		
2	password	varchar(100)	utf8mb4_general_ci		Όχι	Καμία		

Εικόνα 3.15: Ο ορισμός των πεδίων του πίνακα *Administrator*

#	Όνομα	Τύπος	Σύνθεση	Χαρακτηριστικά	Κενό	Προεπιλογή	Σχόλια	Πρόσθετα
1	pub_id 🗝️🔑	int(11)			Όχι	Καμία		
2	pers_id 🗝️🔑	int(11)			Όχι	Καμία		

Εικόνα 3.16: Ο ορισμός των πεδίων του πίνακα *Authorships*

#	Όνομα	Τύπος	Σύνθεση	Χαρακτηριστικά	Κενό	Προεπιλογή	Σχόλια	Πρόσθετα
1	pub_id 🗝️🔑	int(11)			Όχι	Καμία		
2	isbn	varchar(20)	utf8mb4_general_ci		Ναι	NULL		
3	publisher	varchar(60)	utf8mb4_general_ci		Ναι	NULL		

Εικόνα 3.17: Ο ορισμός των πεδίων του πίνακα *Books*

#	Όνομα	Τύπος	Σύνθεση	Χαρακτηριστικά	Κενό	Προεπιλογή	Σχόλια	Πρόσθετα
1	pub_id 🗝️🔑	int(11)			Όχι	Καμία		
2	booktitle	varchar(500)	utf8_general_ci		Ναι	NULL		
3	bookisbn	varchar(20)	utf8mb4_general_ci		Ναι	NULL		
4	bookpublisher	varchar(60)	utf8mb4_general_ci		Ναι	NULL		

Εικόνα 3.18: Ο ορισμός των πεδίων του πίνακα *Book_chapters*

#	Όνομα	Τύπος	Σύνθεση	Χαρακτηριστικά	Κενό	Προεπιλογή	Σχόλια	Πρόσθετα
1	id 🗝️	int(11)			Όχι	Καμία		AUTO_INCREMENT
2	name	varchar(500)	utf8mb4_general_ci		Όχι	Καμία		
3	date	varchar(100)	utf8mb4_general_ci		Ναι	NULL		
4	location	varchar(100)	utf8mb4_general_ci		Ναι	NULL		
5	proc_id 🗝️🔑	int(11)			Ναι	NULL		
6	volume	varchar(10)	utf8mb4_general_ci		Ναι	NULL		

Εικόνα 3.19: Ο ορισμός των πεδίων του πίνακα *Conferences*

#	Όνομα	Τύπος	Σύνθεση	Χαρακτηριστικά	Κενό	Προεπιλογή	Σχόλια	Πρόσθετα
1	pub_id 🗑️	int(11)			Όχι	Καμία		
2	conf_id 🗑️	int(11)			Ναι	NULL		

Εικόνα 3.20: Ο ορισμός των πεδίων του πίνακα *Conference_articles*

#	Όνομα	Τύπος	Σύνθεση	Χαρακτηριστικά	Κενό	Προεπιλογή	Σχόλια	Πρόσθετα
1	id 🗑️	int(11)			Όχι	Καμία		AUTO_INCREMENT
2	name	varchar(300)	utf8mb4_general_ci		Όχι	Καμία		

Εικόνα 3.21: Ο ορισμός των πεδίων του πίνακα *Journals*

#	Όνομα	Τύπος	Σύνθεση	Χαρακτηριστικά	Κενό	Προεπιλογή	Σχόλια	Πρόσθετα
1	pub_id 🗑️	int(11)			Όχι	Καμία		
2	volume	varchar(10)	utf8mb4_general_ci		Ναι	NULL		
3	issue	varchar(10)	utf8mb4_general_ci		Ναι	NULL		
4	journal_id 🗑️	int(11)			Ναι	NULL		

Εικόνα 3.22: Ο ορισμός των πεδίων του πίνακα *Journal_articles*

#	Όνομα	Τύπος	Σύνθεση	Χαρακτηριστικά	Κενό	Προεπιλογή	Σχόλια	Πρόσθετα
1	id 🗑️	int(11)			Όχι	Καμία		AUTO_INCREMENT
2	name	varchar(80)	utf8mb4_general_ci		Όχι	Καμία		
3	prof_id 🗑️	int(11)			Ναι	NULL		

Εικόνα 3.23: Ο ορισμός των πεδίων του πίνακα *Persons*

#	Όνομα	Τύπος	Σύνθεση	Χαρακτηριστικά	Κενό	Προεπιλογή	Σχόλια	Πρόσθετα
1	id 🗑️	int(11)			Όχι	Καμία		AUTO_INCREMENT
2	name	varchar(300)	utf8mb4_general_ci		Ναι	NULL		

Εικόνα 3.24: Ο ορισμός των πεδίων του πίνακα *Proceedings*

#	Όνομα	Τύπος	Σύνθεση	Χαρακτηριστικά	Κενό	Προεπιλογή	Σχόλια	Πρόσθετα
1	id 🗑️	int(11)			Όχι	Καμία		AUTO_INCREMENT
2	title	varchar(500)	utf8mb4_general_ci		Όχι	Καμία		
3	year	int(4)			Όχι	Καμία		
4	page_start	varchar(5)	utf8mb4_general_ci		Ναι	NULL		
5	page_end	varchar(5)	utf8mb4_general_ci		Ναι	NULL		
6	cited_by	int(5)			Όχι	Καμία		

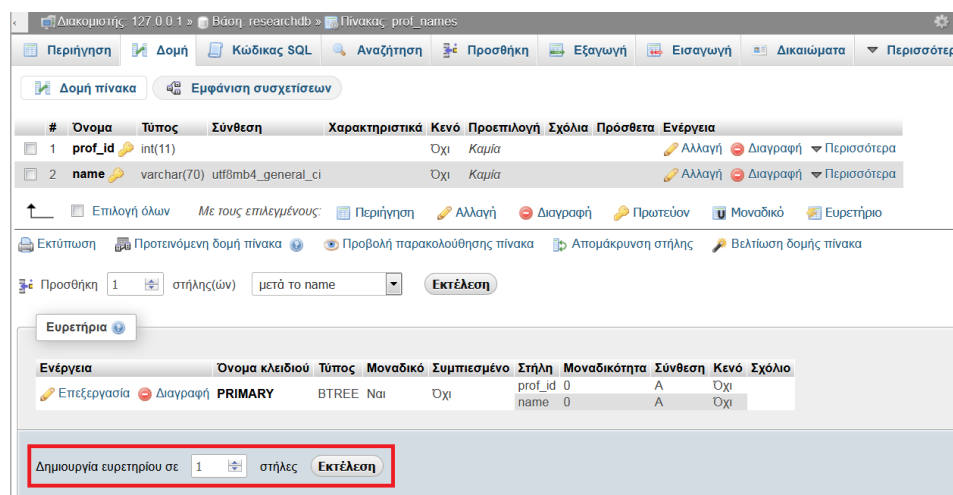
Εικόνα 3.25: Ο ορισμός των πεδίων του πίνακα *Publications*

3.6 Σύνδεση πινάκων και επιβολή ακεραιότητας αναφοράς

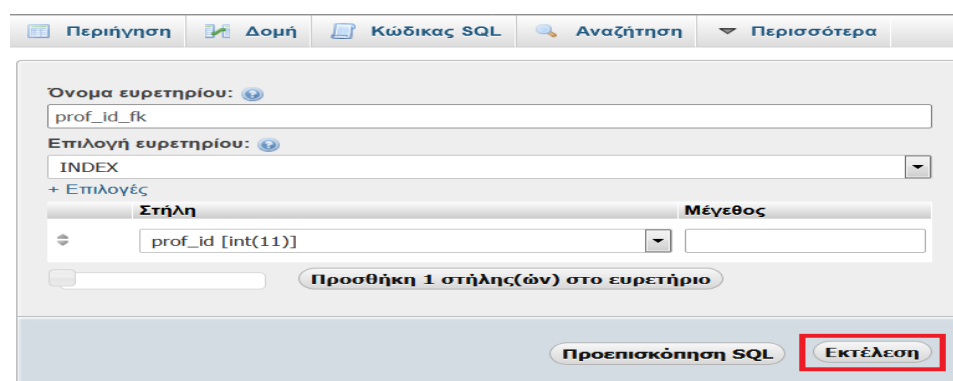
Για να δηλώσουμε τον τρόπο με τον οποίο συνδέονται οι πίνακες μιας βάσης δεδομένων μέσω ξένων κλειδιών στο phpMyAdmin πρέπει να ορίσουμε για κάθε ξένο κλειδί ένα ευρετήριο (index) και στη συνέχεια να δηλώσουμε τη στήλη στην οποία αναφέρεται το ξένο κλειδί. Η διαδικασία αυτή περιγράφεται αναλυτικά στο επόμενο παράδειγμα.

Συνδεση των πινάκων di_professors και prof_names

Στο παράδειγμα αυτό παρουσιάζονται τα βήματα που πρέπει να εκτελεστούν για να οριστεί το πεδίο prof_id του πίνακα prof_names ως ξένο κλειδί το οποίο αναφέρεται στο πεδίο id του πίνακα di_professors.



Εικόνα 3.26: Δημιουργία ευρετηρίου στον πίνακα prof_names



Εικόνα 3.27: Ο ορισμός των επιλογών του ευρετηρίου

Σύμφωνα με την προηγούμενη διαδικασία ορίζονται ξένα κλειδιά και στους υπόλοιπους πίνακες της βάσης δεδομένων. Οι επιλογές που δίνονται σε αυτά τα ξένα κλειδιά φαίνονται στους πίνακες που ακολουθούν.

Ξένα κλειδιά του πίνακα Authorships				
Στήλη	Αναφορά		Περιορισμοί	
	Πίνακας	Στήλη	On delete	On update
pers_id	Persons	id	Cascade	Cascade
pub_id	Publications	id	Cascade	Cascade

Πίνακας 3.1: Οι επιλογές των ξένων κλειδιών του πίνακα *Authorships*

Ξένα κλειδιά του πίνακα Books				
Στήλη	Αναφορά		Περιορισμοί	
	Πίνακας	Στήλη	On delete	On update
pub_id	Publications	id	Cascade	Cascade

Πίνακας 3.2: Οι επιλογές των ξένων κλειδιών του πίνακα *Books*

Ξένα κλειδιά του πίνακα Book_chapters				
Στήλη	Αναφορά		Περιορισμοί	
	Πίνακας	Στήλη	On delete	On update
pub_id	Publications	id	Cascade	Cascade

Πίνακας 3.3: Οι επιλογές των ξένων κλειδιών του πίνακα *Book_chapters*

Ξένα κλειδιά του πίνακα Conferences				
Στήλη	Αναφορά		Περιορισμοί	
	Πίνακας	Στήλη	On delete	On update
proc_id	Proceedings	id	Set null	Cascade

Πίνακας 3.4: Οι επιλογές των ξένων κλειδιών του πίνακα *Conferences*

Ξένα κλειδιά του πίνακα Conference_articles				
Στήλη	Αναφορά		Περιορισμοί	
	Πίνακας	Στήλη	On delete	On update
conf_id	Conferences	id	Set null	Cascade
pub_id	Publications	id	Cascade	Cascade

Πίνακας 3.5: Οι επιλογές ξένων κλειδιών του πίνακα *Conference_articles*

Ξένα κλειδιά του πίνακα Journal_articles				
	Αναφορά		Περιορισμοί	
Στήλη	Πίνακας	Στήλη	On delete	On update
journal_id	Journals	id	Set null	Cascade
pub_id	Publications	id	Cascade	Cascade

Πίνακας 3.6: Οι επιλογές των ξένων κλειδιών του πίνακα Journal_articles

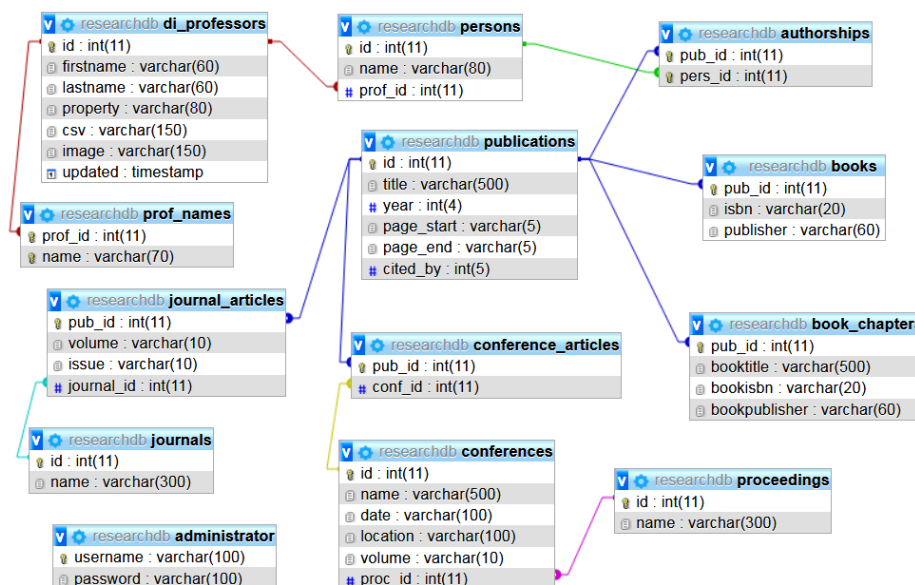
Ξένα κλειδιά του πίνακα Persons				
	Αναφορά		Περιορισμοί	
Στήλη	Πίνακας	Στήλη	On delete	On update
prof_id	di_professors	id	Set null	Cascade

Πίνακας 3.7: Οι επιλογές των ξένων κλειδιών του πίνακα Persons

Ξένα κλειδιά του πίνακα prof_names				
	Αναφορά		Περιορισμοί	
Στήλη	Πίνακας	Στήλη	On delete	On update
prof_id	di_professors	id	Cascade	Cascade

Πίνακας 3.8: Οι επιλογές των ξένων κλειδιών του πίνακα prof_names

Στην επόμενη εικόνα φαίνονται όλοι οι πίνακες της βάσης δεδομένων και οι συνδέσεις μεταξύ τους μέσω των ξένων κλειδιών.



Εικόνα 3.29: Οι πίνακες της βάσης δεδομένων και οι συνδέσεις μεταξύ τους.

Κεφάλαιο 4

Η ανάλυση (parsing) των δεδομένων

4.1 Αρχεία csv

Στην Πληροφορική ένα αρχείο με επέκταση .csv (comma separated values) αναπαριστά δεδομένα πίνακα τα οποία έχουν αποθηκευτεί ως απλό κείμενο διαχωρισμένα με κόμματα. Τα csv αρχεία συνήθως περιέχουν επικεφαλίδα η οποία αντιπροσωπεύει την επικεφαλίδα του πίνακα. Τα πεδία της επικεφαλίδας περιγράφουν το είδος των δεδομένων που είναι αποθηκευμένα στις επόμενες γραμμές. Κάθε επόμενη γραμμή του αρχείου είναι ένα σύνολο δεδομένων. Καθένα από αυτά τα σύνολα δεδομένων αποτελείται από ένα ή περισσότερα πεδία, επίσης χωρισμένα με κόμμα.

Η μορφή αρχείου csv δεν είναι πλήρως τυποποιημένη. Σε απλές υλοποιήσεις csv αρχείων δεν επιτρέπονται τιμές πεδίων που περιέχουν κόμμα ή άλλους ειδικούς χαρακτήρες, όπως ο χαρακτήρας αλλαγής γραμμής. Στις πιο σύνθετες υλοποιήσεις αυτό μπορεί να συμβαίνει, αρκεί οι συγκεκριμένοι χαρακτήρες να είναι ενσωματωμένοι σε δίπλα εισαγωγικά.

Τα csv αρχεία χρησιμεύουν επειδή καθιστούν εύκολη την εισαγωγή των δεδομένων σε οποιοδήποτε πρόγραμμα υπολογιστικών φύλλων ή βάση δεδομένων ανεξάρτητα από το είδος του υπολογιστή ή του προγράμματος λογισμικού που χρησιμοποιούνται.

4.2 Το αρχείο scopus.csv

Στον δικτυακό τόπο του Scopus επιτρέπεται σε έναν χρήστη να εξάγει σε αρχείο, το οποίο μπορεί να αποθηκεύσει στον υπολογιστή του, όλες, ή ορισμένες, από τις δημοσιεύσεις ενός ερευνητή που είναι αποθηκευμένες στη βάση δεδομένων του υπολογιστικού συστήματος. Η εξαγωγή αυτή μπορεί να γίνει σε αρχείο απλού κειμένου, σε αρχείο BibTeX, αλλά και σε αρχείο csv με την ονομασία scopus.csv.

Το αρχείο `scopus.csv` δημιουργείται αυτόματα από το σύστημα, όταν ο χρήστης βρίσκεται στην σελίδα που παρουσιάζει τις δημοσιεύσεις ενός ερευνητή και κάνει κλικ στην επιλογή `Export all` που είναι τοποθετημένη ακριβώς πάνω από την λίστα των δημοσιεύσεων (εικόνα 5.5, σελίδα 62) και στη συνέχεια δώσει την επιλογή `csv` από τις διαθέσιμες μορφές αρχείων που θα εμφανιστούν (εικόνα 5.6, σελίδα 63). Το σύστημα δίνει τη δυνατότητα στον χρήστη να επιλέξει ποια πεδία επιθυμεί να περιέχονται στο αρχείο.

```
1 Authors,Author Ids,Title,Year,Source title,Volume,Issue,Art. No.,Page start,Page end,Page count,Cited by,DOI,Link,Affiliations,Authors with affiliations
2 "Arvanitakis K., Avlonitis M., Papadimitriou E.",57191263146;6506118499;7005909662,"Introducing stochastic recurrence interval to classification algo
3 "Trokidis S.G., Marmarokopoulos K., Avlonitis M.",57203158001;57188846150;6506118499,"Investigation of possible landslide precursor activity in a small
4 "Marmarokopoulos K., Doukakis D., Frantziakou G., Avlonitis M.",57188846150;57201183782;7003281756;6506118499,"Leak detection in plastic water supply
5 "Kalamitanos R., Avlonitis M.",36634300700;6506118499,"Microclimates and their Stochastic Effect on Olive Fruit Fly Evolution: Modeling and Simulat
6 "Arvanitakis K., Kalamitanos R., Avlonitis M.",57191263146;36634300700;6506118499,"Self-optimized computational method calculating robust b values f
7 "Marmarokopoulos K., Keramidis K., Karydis I., Avlonitis M.",36634300700;2309084000;36950839300;6506118499,"Treating stochasticity of olive-fruit fly
8 "Efremidis G., Avlonitis M., Konstantinidis A., Aifantis E.C.",12796270600;6506118499;7004510031;55412103000,"A statistical study of precursor activi
9 "Karydis I., Sioutas S., Avlonitis M., Mylonas P., Kanavas A.",36950839300;16550651500;6506118499;660329346;55520653600,"A survey on big data and ci
10 "Spandidou A., Avlonitis I., Kalamitanos R., Avlonitis M.",55557431500;57192378178;57192373917;6506118499,"Linking population models with public h
11 "Avlonitis M.",6506118499,"On the problem of early detection of users interaction outbreaks via stochastic differential models",2016,"Engineering App
12 "Kalamitanos R., Avlonitis M., Stravrovaidis S.",36634300700;6506118499;2439947900,"Complex networks and simulation strategies: An application to o
13 "Marmarokopoulos K., Efremidis G., Avlonitis M.",57188846150;12796270600;6506118499,"Geophone networks and environmental studies: Application to landsl
14 "Arvanitakis K., Avlonitis M.",57191263146;6506118499,"Identifying asperity patterns via machine learning algorithms",2016,"IFIP Advances in Informa
15 "Avlonitis M., Efremidis G., Konstantinidis A., Aifantis E.C.",6506118499;12796270600;7004510031;55412103000,"The sliding surface roughness effect o
16 "Kalamitanos R., Keramidis K., Avlonitis M., Karydis I.",36634300700;2309084000;6506118499;36950839300,"Environmental impact on predicting olive f
17 "Kalamitanos R., Avlonitis M.",36634300700;6506118499,"The role of olive trees distribution and fruit bearing in olive fruit fly infestation",2015,"
18 "Avlonitis M., Kalaitzidou K.",6506118499;6506810443,"Estimating the real contact area between sliding surfaces by means of a modified OFC model",20
19 "Efremidis G., Avlonitis M., Konstantinidis A., Aifantis E.C.",12796270600;6506118499;7004510031;55412103000,"Realistic spring block models for earth
20 "Avlonitis M., Karydis I., Sioutas S.",6506118499;36950839300;16550651500,"Early prediction in collective intelligence on video users' activity",201
21 "Avlonitis M., Chorianopoulos K.",6506118499;20336592900,"Video pulses: User-based modeling of interesting video segments",2014,"Advances in Multimed
22 "Avlonitis M., Papadopoulos G.A.",6506118499;2445838700,"Foreshocks and b Value: Bridging Macroscopic Observations to Source Mechanical Considerati
23 "Avlonitis M., Kalaitzidou K., Streator J.",6506118499;6506810443;6701756549,"Investigation of friction statistics and real contact area by means of
24 "Karydis I., Avlonitis M., Chorianopoulos K., Sioutas S.",36950839300;6506118499;20336592900;16550651500,"Identifying important segments in videos: f
25 "Spedicato L., Armeni I., Giannoccaro N.I., Avlonitis M., Papavasiliopoulos S.",24313574700;56352831100;9638641900;6506118499;6505564608,"A dynamic ide
26 "Spandidou A., Karydis I., Avlonitis M.",55557431500;36950839300;6506118499,"Mimicking real users' interactions on web videos through a controlled
27 "Naples E., Avlonitis M., Kotsanikolaou P., Stefanidakis M.",6507186367;6506118499;55887494200;16425034600,"Toward early warning against Internet voi
28 "Karydis I., Avlonitis M., Sioutas S.",36950839300;6506118499;16550651500,"Collective intelligence in video user's activity",2012,"IFIP Advances in I
```

Εικόνα 4.1: Ένα αρχείο `scopus.csv`

Στο αρχείο `scopus.csv` αποθηκεύονται δεδομένα ερευνητικών εργασιών που έχουν δημοσιευτεί σε επιστημονικά περιοδικά, πρακτικά συνεδρίων και βιβλία. Το εν λόγω αρχείο περιέχει επικεφαλίδες που αποτελείται από πεδία το πλήθος των οποίων ποικίλει ανάλογα με τις επιλογές που έδωσε ο χρήστης. Το σύνολο των δεδομένων κάθε πεδίου περιέχεται μέσα σε διπλά εισαγωγικά (π.χ. "Arvanitakis K., Avlonitis M., Papadimitriou E.") ενώ, όπως φαίνεται και στην εικόνα 4.1, τα στοιχεία του συνόλου ξεχωρίζονται μεταξύ του με κόμμα. Επίσης με κόμμα ξεχωρίζονται και τα σύνολα δεδομένων.

Ο μέγιστος αριθμός των κεφαλίδων, ανά και των αντιστοιχών πεδίων που μπορεί να περιέχει το αρχείο `scopus.csv` είναι 44. Φυσικά αν ο χρήστης δεν επιλέξει, σύμφωνα με την διαδικασία που περιγράφεται πιο πάνω, να μην εξάγει όλα τα επιτρεπόμενα δυνατά πεδία ο αριθμός αυτός θα είναι μικρότερος από 44 και καθορίζεται αποκλειστικά από την επιλογή του χρήστη.

Στη συνέχεια δίνεται μια συνοπτική περιγραφή όλων των πεδίων επικεφαλίδας, που μπορεί να περιέχει το αρχείο `scopus.csv`:

1. Authors

Τα ονόματα των συγγραφέων της δημοσίευσης. Τα ονόματα αυτά έχουν καταχωρηθεί στο αρχείο με τις εξής μορφές.

- Επώνυμο, κενό, πρώτο γράμμα ονόματος, τελεία.
- Επώνυμο, κενό, πρώτο γράμμα ονόματος, τελεία, πρώτο γράμμα πατρωνύμου, τελεία.

2. Author Ids

Κωδικός αριθμός που προσδιορίζει μοναδικά κάθε συγγραφέα στη βάση δεδομένων του Scopus.

3. Title

Ο τίτλος της δημοσίευσης.

4. Year

Το έτος που δημοσιεύτηκε.

5. Source title

Το όνομα του περιοδικού, των πρακτικών συνεδρίου ή του βιβλίου που δημοσιεύτηκε.

6. Volume

Ο τόμος του περιοδικού, των πρακτικών συνεδρίου ή του βιβλίου.

7. Issue

Το τεύχος του περιοδικού, των πρακτικών συνεδρίου ή του βιβλίου.

8. Art. No.

Ο κωδικός αριθμός που το περιοδικό ή τα πρακτικά συνεδρίου δίνουν στη δημοσίευση.

9. Page start

Ο αριθμός της σελίδας του περιοδικού, των πρακτικών συνεδρίου ή του βιβλίου από την οποία ξεκάνει η δημοσίευση.

10. Page end

Ο αριθμός της σελίδας του περιοδικού, των πρακτικών συνεδρίου ή του βιβλίου στην οποία τελειώνει η δημοσίευση.

11. Page count

Το πλήθος των σελίδων. Στις περισσότερες των περιπτώσεων το πεδίο αυτό δεν περιέχει κάποια τιμή. Έξαλλου το πλήθος των σελίδων μπορεί εύκολα να προσδιοριστεί με αφαίρεση των δυο προηγούμενων πεδίων (Page end – Page start).

12. Cited by

Ο αριθμός των αναφορών (citations) που έχει λάβει η συγκεκριμένη δημοσίευση.

13. DOI

Προέρχεται από τα αρχικά των λέξεων Digital Object Identifier. Είναι μια σειρά από αριθμούς, γράμματα και σύμβολα που χρησιμοποιούνται για τη μόνιμη αναγνώριση ενός άρθρου ή εγγράφου και τη σύνδεσή του στον ιστό. Ένα DOI θα βοηθήσει τον αναγνώστη να εντοπίσει εύκολα ένα έγγραφο από την παραπομπή του.

14. Link

Σύνδεσμος προς ιστοσελίδα του δικτυακού τόπου του Scopus η οποία περιέχει κάποιες βασικές πληροφορίες για τη δημοσίευση, όπως την περίληψη της, τους συγγραφείς της κλπ.

15. Affiliations

Τα εκπαιδευτικά ιδρύματα ή οι οργανισμοί στους οποίους ανήκουν οι συγγραφείς της δημοσίευσης.

16. Authors with affiliations

Αντιστοιχίζει κάθε συγγραφέα στο δικό του ίδρυμα ή οργανισμό.

17. Abstract

Η περίληψη της δημοσίευσης.

18. Author Keywords

Οι λέξεις κλειδιά που έχουν δώσει οι συγγραφείς στη δημοσίευση. Συνήθως έχουν να κάνουν με το επιστημονικό θέμα στο οποίο η δημοσίευση αναφέρεται.

19. Index Keywords

Λέξεις κλειδιά που κάνουν ευκολότερη και πιο γρήγορη την αναζήτηση στη βάση δεδομένων του Scopus.

20. Molecular Sequence Numbers

Αριθμοί μοριακών ακολουθιών. Χρησιμοποιείται συνήθως για δημοσιεύσεις στον κλάδο της Ιατρικής ή της Βιολογίας.

21. Chemicals/CAS

Ο αριθμός μητρώου CAS (CAS registry number). Είναι ένας μοναδικός αριθμός καταγραφής και ταυτοποίησης χημικών στοιχείων, χημικών ενώσεων, πολυμερών, βιολογικών σκευασμάτων, μειγμάτων και κραμάτων.

22. Tradenames

Είναι απλώς η επίσημη εμπορική ονομασία κάτω από την οποία ένα άτομο ή μια εταιρεία επιλέγει να ασκεί δραστηριότητές.

23. Manufacturers

Τα ονόματα των κατασκευαστών. Χρησιμοποιείται συνήθως για δημοσιεύσεις στο επιστημονικό πεδίο της Μηχανικής.

24. Funding Details

Λεπτομέρειες για τους χρηματοδότες του έργου, αν υπάρχουν.

25. Funding Text 1

Το μέρος του κειμένου της δημοσίευσης όπου οι συγγραφείς αποδίδουν ευχαριστίες σε όσους βοήθησαν στην ολοκλήρωση της.

26. References

Αναφορές σε άλλες δημοσιεύσεις.

27. Correspondence Address

Διεύθυνση αλληλογραφίας.

28. Editors

Οι συντάκτες του περιοδικού, ή των πρακτικών του συνεδρίου στο οποίο δημοσιεύεται.

29. Sponsors

Οι χορηγοί της δημοσίευσης.

30. Publisher

Ο εκδότης του περιοδικού, των πρακτικών συνεδρίου ή του βιβλίου στο οποίο φιλοξενείται η δημοσίευση.

31. Conference name

Το όνομα του συνεδρίου στο οποίο παρουσιάζεται η δημοσίευση.

32. Conference date

Η ημερομηνία διεξαγωγής του συνεδρίου.

33. Conference location

Ο τόπος διεξαγωγής του συνεδρίου.

34. Conference code

Ένα κωδικός αριθμός ο οποίος προσδιορίζει μοναδικά κάθε συνέδριο.

35. ISSN

Προέρχεται από τα αρχικά των λέξεων International Standard Serial Number και είναι ένας οκταψήφιος αριθμός ο οποίος χρησιμοποιείται για τη μοναδική αναγνώριση μιας δημοσίευσης. Το ISSN είναι ιδιαίτερα χρήσιμο για τη διάκριση μεταξύ δημοσιεύσεων με τον ίδιο τίτλο.

36. ISBN

Προέρχεται από τα αρχικά των λέξεων International Standard Book Number και είναι ένα μοναδικό αναγνωριστικό εμπορικού βιβλίου.

37. CODEN

Είναι ένας αλφαριθμητικός βιβλιογραφικός κώδικας έξι χαρακτήρων ο οποίος παρέχει συνοπτική, μοναδική και ξεκάθαρη αναγνώριση των τίτλων των περιοδικών και των δημοσιεύσεων από όλους τους τομείς.

38. PubMed ID

Είναι ο μοναδικός αριθμός αναγνωριστικού που χρησιμοποιείται στην PubMed. Η PubMed είναι μια μηχανή αναζήτησης με δωρεάν πρόσβαση κυρίως στη βάση δεδομένων MEDLINE με αναφορές και περιλήψεις για τις Βιοεπιστήμες και τη Βιοϊατρική.

39. Language of Original Document

Η γλώσσά του πρωτότυπου κειμένου της δημοσίευσης.

40. Abbreviated Source Title

Συντομογραφία του ονόματος του περιοδικού, των πρακτικών συνεδρίου ή του βιβλίου που έγινε η δημοσίευση.

41. Document Type

Το είδος του έγγραφου. Αποτελείται από τις κατηγορίες Conference Paper, Article, Book Chapter, Review, Editorial, Book και Article in Press.

42. Access Type

Ο τύπος πρόσβασης της δημοσίευσης. Αν είναι δηλαδή ανοιχτής πρόσβασης ή όχι. Η ανοιχτή πρόσβαση (Open Access) αναφέρεται σε ερευνητικά αποτελέσματα που διανέμονται ηλεκτρονικά και χωρίς κόστος ενδεχομένως με την προσθήκη δημιουργικής άδεια κοινής χρήσης (Creative Commons).

43. Source

Η προέλευση των δεδομένων.

44. EID

Ο κωδικός αναφοράς της επιστημονικής εργασίας στη βάση δεδομένων του Scopus.

Στο σημείο αυτό αξίζει να σημειωθεί ότι ανάλογα με το είδος της δημοσίευσης που καταγράφεται σε κάθε γραμμή του αρχείου, ορισμένα πεδία μπορεί να μην έχουν καθόλου δεδομένα δηλαδή να είναι κενά. Για παράδειγμα αν η δημοσίευση είναι του τύπου (Document Type) Article, τότε τα πεδία που αντιστοιχούν στις επικεφαλίδες Conference name, Conference date, Conference location, και Conference code θα έχουν ως τιμή την κενή συμβολοσειρά ("") αφού πρόκειται για άρθρο το οποίο δημοσιεύεται σε περιοδικό, και δεν παρουσιάζετε σε κάποιο συνέδριο.

4.3 Η Ανάλυση (parsing) του αρχείου scopus.csv

Στην ενότητα αυτή παρουσιάζεται ο κώδικας σε γλώσσα PHP με τον οποίο έγινε η ανάλυση του αρχείου scopus.csv και η εισαγωγή των δεδομένων του στη βάση δεδομένων της εφαρμογής.

Από τον καθορισμό των απαιτήσεων της εφαρμογής προέκυψε ότι δεν χρειάζονται όλα τα δεδομένα που αντιστοιχούν και στα 44 πεδία της κεφαλίδας του αρχείου που αναφέρθηκαν στην προηγούμενη ενότητα αλλά μόνο ορισμένα από αυτά. Τα απαραίτητα δεδομένα είναι όσα αντιστοιχούν στα πεδία της κεφαλίδας που υπάρχουν στον πίνακα 4.1.

Title	Authors	Year
Cited by	Document Type	Source title
Page start	Page end	ISBN
Publisher	Volume	Issue
Conference name	Conference date	Conference location
Source title		

Πίνακας 4.1: Τα πεδία της κεφαλίδας του αρχείου *scopus.csv* που χρησιμοποιούνται στην εφαρμογή

Τα τμήματα κώδικα που ακολουθούν αποτελούν μέρη το αρχείου *parseCSV.php* στο οποίο υλοποιείται ο αλγόριθμος εισαγωγής των δεδομένων του αρχείου *scopus.csv* στη βάση της εφαρμογής.

Τμήμα κώδικα 4.1: *parceCSV.php* γραμμές 1-8

```

1 <?php
2 session_start();
3 include("../Database.php");
4 $conn = connectToDB();
5 $id = $_POST["user_id"];
6 $array = array();
7 $sql = "SELECT firstname, lastname, csv FROM di_professors WHERE id = $id";
8 $result = mysqli_query($conn, $sql);

```

Στη γραμμή 3 με την εντολή `include` γίνεται φόρτωση των συναρτήσεων που περιέχονται στο αρχείο *Database.php* ώστε να μπορούν να χρησιμοποιηθούν στο *parseCSV.php*. Το αρχείο *Database.php* περιέχει βοηθητικές συναρτήσεις για τη σύνδεση με τη βάση δεδομένων και την εισαγωγή εγγγραφών σε αυτήν. Ο κώδικας με τον οποίο υλοποιούνται οι συγκεκριμένες συναρτήσεις βρίσκεται στο παράρτημα.

Στη γραμμή 4 εγκαθιδρύεται η σύνδεση με τη βάση δεδομένων. Στη γραμμή 6 δηλώνεται ένας πίνακας στον οποίο θα αποθηκευτούν τα δεδομένα που είναι απαραίτητα για να λειτουργήσει σωστά η progress bar που δείχνει το ποσοστό ολοκλήρωσης της διαδικασίας (εικόνα 5.16, σελίδα 68). Η διαδικασία εισαγωγής των δεδομένων ερευνητικών δημοσιεύσεων των μελών ΔΕΠ του Τμήματος Πληροφορικής του Ιονίου Πανεπιστημίου, γίνεται ξεχωριστά για κάθε μέλος. Έτσι στις γραμμές 7 και 8 ορίζεται και εκτελείται ένα ερώτημα προς τη βάση δεδομένων με το οποίο αναζητούνται στον πίνακα *di_professors* στοιχεία του μέλους ΔΕΠ που είναι απαραίτητα για να λάβουν χώρα κατάλληλοι έλεγχοι και να προχωρήσει η διαδικασία.

Τμήμα κώδικα 4.2: *parceCSV.php* γραμμές 9-31

```

9 if (mysqli_num_rows($result) > 0) {
10     while ($row = mysqli_fetch_assoc($result)) {
11         $firstname = $row["firstname"];
12         $lastname = $row["lastname"];
13         $file = $row["csv"];
14         $checking = "SELECT id FROM prof_names WHERE prof_id = $id";
15         $checkingResult = mysqli_query($conn, $checking);

```

```

16     if (mysqli_num_rows($checkingResult) > 0) {
17         $array["check"] = true;
18         //read csv file as a string
19         $csv = file_get_contents("../resources/csv/$file", FILE_USE_INCLUDE_PATH)
20             ↪ ;
21         //convert string to array
22         $lines = explode("\n", $csv);
23         $linesLength = count($lines);
24         //remove the last element of lines if empty because causes problems to
25         ↪ parsing
26         if ($lines[$linesLength - 1] == "") {
27             array_pop($lines);
28         }
29         //remove the first element from the array. This element is the header of
30         ↪ CSV file
31         $head = str_getcsv(array_shift($lines), ",", "'");
32         $head[0] = "Authors";
33         // count how many fields has the header of the file
34         $headLength = count($head);

```

Αν τα στοιχεία αυτά υπάρχουν στην βάση δεδομένων τότε εκτελείται, νέο ερώτημα (γραμμές 14,15) το οποίο αναζητά αν έχουν αποθηκευτεί στη βάση δεδομένων τα ονόματα με τα οποία το μέλος ΔΕΠ είναι καταχωρημένο στο αρχείο scorpus.csv.

Αν στον πίνακα `prof_names` της βάσης δεδομένων υπάρχουν έγγραφες που αντιστοιχούν στο μέλος ΔΕΠ για το οποίο γίνεται η διαδικασία, εκτελείται η εντολή `file_get_contents` (γραμμή 19) η οποία διαβάζει τα περιεχόμενα του αρχείου csv που αντιστοιχεί στο συγκεκριμένο μέλος ΔΕΠ και τα αποθηκεύει στην μεταβλητή `$csv` τύπου συμβολοσειράς (string). Στη γραμμή 21 χρησιμοποιείται η συνάρτηση `explode` της PHP η οποία μετατρέπει τη συμβολοσειρά `$csv` σε πίνακα χρησιμοποιώντας ως διαχωριστικό τον χαρακτήρα αλλαγής γραμμής `\n`. Τα δεδομένα αυτά εκχωρούνται στη μεταβλητή `$lines`. Με τον τρόπο αυτόν κάθε γραμμή του csv αρχείου είναι μια γραμμή του πίνακα `$lines`. Επίσης υπολογίζεται και αποθηκεύεται στη μεταβλητή `$linesLength` το μέγεθος του πίνακα `$lines` (γραμμή 22).

Η συνάρτηση `array_pop` της PHP διαγράφει το τελευταίο στοιχείο ενός πίνακα. Η προαναφερθείσα συνάρτηση, χρησιμοποιείται στη γραμμή 24 για να διαγράψει το τελευταίο στοιχείο του πίνακα `$lines`, αν αυτό είναι η κενή συμβολοσειρά. Η διαδικασία αυτή είναι απαραίτητη, διότι διαπιστώθηκε πως αν το τελευταίο στοιχείο του πίνακα είναι κενή συμβολοσειρά, τότε προκαλείται πρόβλημα στο parsing.

Στη γραμμή 28 χρησιμοποιείται η συνάρτηση `array_shift` η οποία αφαιρεί και επιστρέφει την πρώτη γραμμή του πίνακα που δέχεται ως όρισμα. Έτσι λοιπόν αφαιρείται η πρώτη γραμμή του πίνακα `$lines`, η οποία είναι ουσιαστικά η επικεφαλίδα του αρχείου csv. Η γραμμή αυτή εν συνεχεία δίνεται ως πρώτο όρισμα στη συνάρτηση, `str_getcsv` η οποία αυτό που κάνει είναι να επιστρέψει έναν πίνακα όπου κάθε γραμμή περιέχει τα πεδία της συμβολοσειράς αρχείου csv του πρώτου ορίσματος της χρησιμοποιώντας ως διαχωριστικά τους χαρακτήρες που παίρνει στο δεύτερο και το τρίτο της όρισμα. Με τον τρόπο αυτό ορίζεται ο πίνακας `$head` ο οποίος περιέχει σε κάθε γραμμή του κι από

ένα πεδίο της κεφαλίδας που αρχείου csv.

Το πρώτο στοιχείο του πίνακα `$head ($head[0])` πρέπει, όπως αναφέρθηκε και στην σελίδα 40, να έχει τιμή `Authors`. Σε δοκιμές που έγιναν, για κάποιο άγνωστο έως τώρα λόγο, η τιμή αυτή τυπωνόταν ως ``Authors`. Το γεγονός αυτό δεν προκαλεί πρόβλημα στο parsing αρκεί στην συνέχεια του προγράμματος όπου χρησιμοποιείται αυτή η τιμή, να γράφεται κατά αυτόν τον τρόπο. Όμως για λόγους ευκολίας, αλλά και για να αποφευχθεί η σύγχυση που μπορεί να προκληθεί κατά την δοκιμή του κώδικα, από λάθη που είναι δυνατόν να συμβούν στην πληκτρολόγηση, εκχωρήθηκε στην μεταβλητή `$head[0]` η τιμή `Authors` (γραμμή 31).

Το τμήμα κώδικα των γραμμών 32 έως 69 χρησιμοποιείται για να αντιμετωπιστούν κάποια προβλήματα που παρουσιάστηκαν στο parsing από τη χρήση της συνάρτησης `array_combine` στη γραμμή 72. Για τις εντολές των γραμμών αυτών θα γίνει λόγος στην επόμενη ενότητα.

Τμήμα κώδικα 4.3: *parceCSV.php* γραμμές 70-77

```

70     $data = array();
71     foreach ($lines as $line) {
72         $data[] = array_combine($head, str_getcsv($line));
73     }
74     $publicationsCounter = 0;
75     $updatedCitationsCounter = 0;
76     $length = count($data);
77     $loop = 0;
```

Η παρουσίαση του κώδικα συνεχίζεται από την γραμμή 70 όπου δηλώνεται η μεταβλητή τύπου πίνακα `$data`. Στον πίνακα αυτόν θα αποθηκευτούν όλα τα στοιχεία που μετέπειτα θα εισαχθούν στη βάση δεδομένων.

Οι διασυνδεδεμένοι πίνακες (associative arrays), είναι πίνακες στους οποίους το ευρετήριο (δείκτης) είναι μια συμβολοσειρά (string), αντί για αριθμός, και αποθηκεύουν τα δεδομένα σε συνδυασμούς ζευγών κλειδιού και τιμής (key, value pairs). Η συνάρτηση `array_combine` της PHP δέχεται ως ορίσματα δυο πίνακες και δημιουργεί ένα νέο διασυνδεδεμένο πίνακα που έχει ως κλειδιά τα στοιχεία του πρώτου ορίσματος και ως τιμές τα στοιχεία του δεύτερου. Απαραίτητη προϋπόθεση για να λειτουργήσει είναι οι δυο πίνακες που χρησιμοποιούνται ως ορίσματα να έχουν το ίδιο μέγεθος.

Στη γραμμή 72, η συνάρτηση `array_combine` παίρνει ορίσματα τον πίνακα `$head` που περιέχει τα πεδία τις επικεφαλίδας του csv αρχείου, και κάθε γραμμή του πίνακα `$lines` διοχετευμένη στη συνάρτηση `str_getcsv`. Ο πίνακας `$lines` μετά την αφαίρεση της επικεφαλίδας, περιέχει πλέον μόνο τα δεδομένα του csv αρχείου. Με τη συνάρτηση `str_getcsv`, και μέσω της δομής επανάληψης, κάθε γραμμή του πίνακα `$lines` μετατρέπεται σε ένα πίνακα ο οποίος περιέχει τα δεδομένα που αντιστοιχούν σε κάθε πεδίο της επικεφαλίδας του αρχείου csv. Έτσι κάθε γραμμή του πίνακα `$data` είναι ένας διασυνδεδεμένος πίνακας που έχει ως κλειδιά τα πεδία τις επικεφαλίδας και ως τιμές τα αντίστοιχα δεδομένα.

Για παράδειγμα τα δεδομένα του αρχείου csv της εικόνας 4.1 στη σελίδα 40 έχουν τώρα αποθηκευτεί στον πίνακα \$data με την εξής μορφή:

```
$data[0] =>
("Authors" => "Arvanitakis K., Avlonitis M., Papadimitriou E."
 "Author Ids" => "57191263146;6506118499;7005909662;", ...)
$data[1] =>
("Authors" => "Krokidis S.G., Marmarokopos K., Avlonitis M."
 "Author Ids" => "57203158801;57188846150;6506118499;", ...)
```

Εικόνα 4.2: Ένα τμήμα των δεδομένων του αρχείου *scopus.csv* σε μορφή διασυνδεδεμένου πίνακα

Στις γραμμές 74 έως 77 δηλώνονται οι μεταβλητές \$publicationsCounter, \$updatedCitationsCounter, \$length, \$loop οι οποίες χρειάζονται για να λειτουργήσει η γραμμή προόδου που εμφανίζει το ποσοστό ολοκλήρωσης της διαδικασίας, αλλά και μηνύματα προς τον χρήστη για το πλήθος των εγγγραφών που έγιναν στη βάση δεδομένων και πόσες αναφορές των δημοσιεύσεων έχουν ενημερωθεί.

Τμήμα κώδικα 4.4: *parceCSV.php* γραμμές 78-189

```
78     foreach ($data as $key => $value) {
79         //defanie helpul variables
80         $title = $data[$key]["Title"];
81         $authors = $data[$key]["Authors"];
82         $year = $data[$key]["Year"];
83         $citedBy = $data[$key]["Cited by"];
84         $listOfAuthors = (explode(",", $authors));
85         $documentType = $data[$key]["Document Type"];
86         $source = $data[$key]["Source title"];
87         $pageStart = $data[$key]["Page start"];
88         $pageEnd = $data[$key]["Page end"];
89         $allIsbns = explode(";", $data[$key]["ISBN"]);
90         $isbn = $allIsbns[0];
91         $publisher = $data[$key]["Publisher"];
92         $volume = $data[$key]["Volume"];
93         $issue = $data[$key]["Issue"];
94         $confName = $data[$key]["Conference name"];
95         $confDate = $data[$key]["Conference date"];
96         $confLocation = $data[$key]["Conference location"];
97         $procName = $data[$key]["Source title"];
98         $procVolume = $data[$key]["Volume"];
99         //Start inserting to DB
100        $hasInserted = publicationHasInserted($conn, $title);
101        if (!$hasInserted) {
102            switch ($documentType) {
103                case "Article":
104                case "Article in Press":
105                case "Review":
106                    $pub_id = insertIntoPublications($conn, $title, $year,
107                        ↪ $pageStart, $pageEnd, $citedBy);
108                    if ($pub_id != null) {
109                        $publicationsCounter++;
110                        insertIntoPersons($conn, $listOfAuthors);
111                        insertIntoAuthorships($conn, $pub_id, $listOfAuthors);
```

```

111         insertIntoJournals($conn, $source);
112         insertIntoJournalArticles($conn, $pub_id, $volume,
113             ↪ $issue, $source);
114     }
115     break;
116 case "Conference Paper":
117     $pub_id = insertIntoPublications($conn, $title, $year,
118         ↪ $pageStart, $pageEnd, $citedBy);
119     if ($pub_id != null) {
120         $publicationsCounter++;
121         insertIntoPersons($conn, $listOfAuthors);
122         insertIntoAuthorships($conn, $pub_id, $listOfAuthors);
123         if (empty($confName)) {
124             insertIntoJournals($conn, $source);
125             insertIntoJournalArticles($conn, $pub_id, $volume,
126                 ↪ $issue, $source);
127         } else {
128             insertIntoProccendings($conn, $procName);
129             insertIntoConferences($conn, $confName, $confDate,
130                 ↪ $confLocation, $procVolume, $procName);
131             insertIntoConferenceArticles($conn, $pub_id,
132                 ↪ $confName, $confDate);
133         }
134     }
135     break;
136 case "Editorial":
137     if (!empty($pageStart)) {
138         $pub_id = insertIntoPublications($conn, $title, $year,
139             ↪ $pageStart, $pageEnd, $citedBy);
140         if ($pub_id != null) {
141             $publicationsCounter++;
142             insertIntoPersons($conn, $listOfAuthors);
143             insertIntoAuthorships($conn, $pub_id, $listOfAuthors
144                 ↪ );
145             if (empty($confName)) {
146                 insertIntoJournals($conn, $source);
147                 insertIntoJournalArticles($conn, $pub_id,
148                     ↪ $volume, $issue, $source);
149             } else {
150                 insertIntoProccendings($conn, $procName);
151                 insertIntoConferences($conn, $confName,
152                     ↪ $confDate, $confLocation, $procVolume,
153                     ↪ $procName);
154                 insertIntoConferenceArticles($conn, $pub_id,
155                     ↪ $confName, $confDate);
156             }
157         }
158     }
159     break;
160 case "Book":
161     $pub_id = insertIntoPublications($conn, $title, $year,
162         ↪ $pageStart, $pageEnd, $citedBy);
163     if ($pub_id != null) {
164         $publicationsCounter++;
165         insertIntoPersons($conn, $listOfAuthors);
166         insertIntoAuthorships($conn, $pub_id, $listOfAuthors);
167         insertIntoBooks($conn, $pub_id, $isbn, $publisher);
168     }
169     break;
170 case "Book Chapter":
171     $pub_id = insertIntoPublications($conn, $title, $year,
172         ↪ $pageStart, $pageEnd, $citedBy);

```

```

160         if ($pub_id != null) {
161             $publicationsCounter++;
162             insertIntoPersons($conn, $listOfAuthors);
163             insertIntoAuthorships($conn, $pub_id, $listOfAuthors);
164             insertIntoBookChapters($conn, $pub_id, $source, $isbn,
                                   ↪ $publisher);
165         }
166         break;
167     default:
168         break;
169     } //end switch
170 } else {
171     updatePersonIfIsProfessorAndHasNullProfidInPersonsTable($conn,
                                   ↪ $listOfAuthors);
172     $citationsHasUpdated = updateCitations($conn, $title, $citedBy);
173     if ($citationsHasUpdated) {
174         $updatedCitationsCounter++;
175     }
176 }
177 $loop++;
178 $percent = intval($loop / $length * 100);
179 $array["percent"] = $percent;
180 $array["message"] = $loop . "of $length row(s) processed.";
181 $array["inserted"] = $publicationsCounter;
182 $array["updated"] = $updatedCitationsCounter;
183 file_put_contents("tmp/" . session_id() . ".txt", json_encode($array))
184     ↪ ;
185     usleep(100000);
186 } //end foreach
187 $update = "UPDATE di_professors SET updated=NOW() WHERE id = $id";
188 if (!mysqli_query($conn, $update)) {
189     echo mysqli_error($conn);
190 }

```

Οι εντολές που υπάρχουν στο τμήμα κώδικα των γραμμών 78 έως 189, πραγματοποιούν εγγραφή και ενημέρωση των δεδομένων στη βάση. Αυτό γίνεται μέσω της δομής επανάληψης `foreach`, η οποία διατρέχει όλα τα στοιχεία του πίνακα `$data`.

Στις γραμμές 80 έως 98 ορίζονται κάποιες βοηθητικές μεταβλητές που κάνουν ευκολότερη την πληκτρολόγηση του κώδικα. Στη γραμμή 84 η συνάρτηση `explode` μετατρέπει τα στοιχεία της συμβολοσειράς `$authors`, που δέχεται ως όρισμα σε πίνακα, χρησιμοποιώντας ως διαχωριστικό της συμβολοσειράς το χαρακτήρα `','`. Ο πίνακας αυτός αποθηκεύεται στη μεταβλητή `$listOfAuthors`. Αντίστοιχη χρήση της συνάρτησης `explode` λαμβάνει χώρα στην γραμμή 89, όπου αυτή τη φορά μετατρέπει σε πίνακα τη συμβολοσειρά που περιέχει τους αριθμούς ISBN της δημοσίευσης και τον αποθηκεύει στην μεταβλητή με όνομα `$allIsbns`.

Μια ερευνητική εργασία μπορεί να έχει δημοσιευτεί σε διάφορα βιβλία και επομένως να έχει περισσότερους από έναν αριθμούς ISBN. Με τον τρόπο που δομήθηκε η βάση δεδομένων της εφαρμογής όμως αποθηκεύεται σε αυτή, μόνο ένας αριθμός ISBN, για αυτόν το λόγο στην γραμμή 90 ορίζεται η μεταβλητή `$isbn` στην οποία εκχωρείται ως τιμή μόνο το πρώτο στοιχείο (`$allIsbns[0]`) του πίνακα `$allIsbns`, που είναι τελικά αυτός ο αριθμός ISBN ο οποίος στη συνέχεια θα αποθηκευτεί στη βάση δεδομένων.

Στη γραμμή 100 εκχωρείται στη μεταβλητή `$hasInserted`, το αποτέλεσμα της συνάρτησης `publicationHasInserted` που ορίζεται στο ενσωματωμένο (γραμμή 3) αρχείο `Database.php`. Η συνάρτηση αυτή ελέγχει αν έχει αποθηκευτεί η όχι στη βάση δεδομένων η δημοσίευση της οποίας τον τίτλο δέχεται ως όρισμα και επιστέφει `true` η `false` αντίστοιχα.

Στην περίπτωση που τα στοιχεία της δημοσίευσης δεν είναι αποθηκευμένα στη βάση δεδομένων εκτελείται το τμήμα κώδικα των γραμμών 102 έως 169, το οποίο περιέχει τις εντολές οι οποίες είναι απαραίτητες για την εισαγωγή των στοιχείων της δημοσίευσης στη βάση δεδομένων. Στο τμήμα κώδικα αυτό θα αναφερθούμε παρακάτω.

Εάν τα στοιχεία της δημοσίευσης είναι καταχωρημένα στη βάση δεδομένων εκτελείται ο κώδικας των γραμμών 171-175. Στη γραμμή 171 καλείται η συνάρτηση `updatePersonIfIsProfessorAndHasNullProfidInPersonsTable` του ενσωματωμένου αρχείου `Database.php`. Η συνάρτηση αυτή δέχεται ως όρισμα τον πίνακα `$listOfAuthors` που περιέχει τα ονόματα των συγγράφων της δημοσίευσης και αν πληρούνται οι συνθήκες που ορίζονται στο σώμα της, θέτει στο πεδίο `prof_Id` του πίνακα `persons` την τιμή που έχει το πεδίο `prof_Id` από πίνακα `prof_names` της βάσης.

Στη συνέχεια (γραμμή 172) καλείται η συνάρτηση `updateCitations` του ενσωματωμένου αρχείου η οποία, αν πληρούνται οι συνθήκες, ενημερώνει στη βάση δεδομένων το πεδίο που περιέχει τον αριθμό των αναφορών που έχει λάβει η δημοσίευση τον τίτλο της οποίας δέχεται ως όρισμα, και επιστέφει την τιμή `true` η `false` αντίστοιχα, ανάλογα με το αν το πεδίο ενημερώθηκε ή όχι. Η τιμή αυτή εκχωρείται στη μεταβλητή `$citationsHasUpdated`. Ακολούθως γίνεται έλεγχος αν η μεταβλητή `$citationsHasUpdated` έχει την τιμή `true` και αν ναι αυξάνεται κατά ένα η τιμή της μεταβλητής `$updatedCitationsCounter`, που δείχνει σε πόσες δημοσιεύσεις έχει γίνει ανανέωση των αναφορών (γραμμές 173, 174).

Αν τα στοιχεία της δημοσίευσης δεν έχουν καταχωρηθεί στη βάση δεδομένων, τότε όπως προαναφέρθηκε, εκτελείται το τμήμα κώδικα των γραμμών 102 έως 169. Στην γραμμή 102 μέσω της δομής `switch` επιλέγεται ο τύπος της δημοσίευσης που είναι αποθηκευμένος στη μεταβλητή `$documentType`. Αν ο τύπος της δημοσίευσης είναι ένας από τους `Article`, `Article in Press`, `Review`, εκτελούνται οι εντολές των γραμμών 106 έως 112 και τα εισάγονται τα στοιχεία της δημοσίευσης στον πίνακα `journal_articles` της βάσης δεδομένων καθώς και όλα τα απαραίτητα στοιχεία σε άλλους πίνακες. Η διαδικασία αυτή γίνεται με τη βοήθεια των συναρτήσεων που υπάρχουν στο ενσωματωμένο αρχείο `Database.php`. Επιπλέον αυξάνεται κατά ένα η τιμή της μεταβλητής `$publicationsCounter` που δείχνει τον αριθμό των δημοσιεύσεων που έχουν εισαχθεί στη βάση δεδομένων.

Η ίδια ακριβώς διαδικασία γίνεται και για τις άλλες περιπτώσεις του τύπου της δημοσίευσης. Αν η δημοσίευση είναι `Conference Paper` εκτελείται το τμήμα των γραμμών 116 έως 129, αν είναι `Editorial` το τμήμα 132 έως 147, αν είναι `Book` το τμήμα 150 έως 156 και τέλος αν είναι `Book Chapter`, το

τμήμα 159 έως 165. Σε όλες τις περιπτώσεις αυξάνεται η τιμή της μεταβλητής `$publicationsCounter` κατά ένα.

Αν ο τύπος της δημοσίευσης είναι `Editorial`, και η μεταβλητή `$confName`, που αποθηκεύεται το όνομα του συνεδρίου, είναι η κενή συμβολοσειρά, τότε η δημοσίευση θεωρείται ότι είναι `Journal Article`, αλλιώς `Conference Paper` και τα στοιχεία της καταχωρούνται στους αντίστοιχους πίνακες ανάλογα με την περίπτωση.

Στις γραμμές 176 έως 184 του κώδικα περιέχονται εντολές για να τη λειτουργία της γραμμής προόδου (progress bar) που δείχνει το ποσοστό ολοκλήρωσης της διαδικασίας εγγραφής και ανανέωσης των δεδομένων στη βάση. Συγκεκριμένα η μεταβλητή `$loop` είναι ένας μετρητής που δείχνει τον αριθμό των φόρων που εκτελείται η επαναληπτική διαδικασία προσπέλασης των γραμμών του πίνακα `$data`. Κάθε φορά που ολοκληρώνεται μια επανάληψη, η τιμή αυτού του μετρητή αυξάνεται κατά ένα (γραμμή 177). Στη μεταβλητή `$percent` αποθηκεύεται το ποσοστό ολοκλήρωσης της διαδικασίας, το οποίο μπορεί να υπολογιστεί από τη διαίρεση της `$loop` με τη μεταβλητή `$length`, και πολλαπλασιασμό με το 100 (γραμμή 178). Η μεταβλητή `$length` εκφράζει το μέγεθος του πίνακα `$data`, δηλαδή το συνολικό αριθμό των επαναλήψεων που πρέπει να γίνουν για να προσπελαστούν όλα του τα στοιχεία. Οι τιμές των μετρητών `$percent`, `$publicationsCounter` και `$updatedCitationsCounter` αποθηκεύονται στο διασυνδεδεμένο πίνακα `$array` (γραμμές 179-182). Η συνάρτηση `json_encode` δέχεται ως όρισμα τον πίνακα `$array` και τον μετατρέπει σε κωδικοποιημένη μορφή JSON. Στη συνέχεια το αποτέλεσμα της κωδικοποίησης γράφεται σε ένα προσωρινό αρχείο με την εντολή, `file_put_contents` (γραμμή 183). Τα δεδομένα του αρχείου αυτού, με την τεχνολογία AJAX και εντολές jQuery που περιέχονται σε αλλά αρχεία εμφανίζουν στον browser τα μηνύματα που πληροφορούν τον χρήστη για την πρόοδο της διαδικασίας και τον αριθμό των εγγραφών και των ανανεώσεων που έγιναν στην βάση δεδομένων.

Αφού ολοκληρωθεί η επαναληπτική διαδικασία προσπέλασης των γραμμών του πίνακα `$data`, στις γραμμές 186 και 187 ορίζεται και εκτελείται ερώτημα προς την βάση δεδομένων το οποίο πραγματοποιεί την ενημέρωση του πεδίου `updated` του πίνακα `di_professors` της βάσης με τη χρονοσφραγίδα της τρέχουσας ημερομηνίας και ώρας. Αυτό γίνεται με τη χρήση της συνάρτησης `NOW()` της MySQL. Η διαδικασία αυτή είναι απαραίτητη ώστε να υπάρχει στη βάση δεδομένων της εφαρμογής μια καταγραφή για το πότε έγινε τελευταία φόρα καταχώριση νέων και ενημέρωση των υπάρχοντων δημοσιεύσεων κάθε μέλους ΔΕΠ.

Τμήμα κώδικα 4.5: *parceCSV.php* γραμμές 190-200

```

190     } else {
191         $array["check"] = false;
192         $array["percent"] = 100;
193         $array["message"] = "<p>There is not inserted Scopus names for
        ↳ $firstname $lastname. </p>"

```

```

194         . "<p>For insert or update data correctly required inserted at
           ↳ least one Scopus name for $firstname $lastname. </p>"
195         . "<p>Please click \"Names\" button and insert at least one Scopus
           ↳ name for $firstname $lastname. </p>";
196         file_put_contents("tmp/" . session_id() . ".txt", json_encode($array));
197     }
198 } //end while
199 } //end if
200 mysqli_close($conn);

```

Το τμήμα κώδικα των γραμμών 190 έως 197 εκτελείται, αν η συνθήκη της γραμμής 16 είναι ψευδής, δηλαδή να δεν υπάρχουν καταχωρημένα στον πίνακα `prof_names` της βάσης δεδομένων τα ονόματα με τα οποία αναγράφεται στο CSV αρχείο το μέλος ΔΕΠ για το οποίο γίνεται η διαδικασία. Τότε εμφανίζεται ένα μήνυμα που ενημερώνει τον διαχειριστή του συστήματος για το γεγονός αυτό και τον προτρέπει να προβεί στην καταχώριση τους. Τέλος ο κώδικας της ανάλυσης του αρχείου CSV ολοκληρώνεται με την εντολή `mysqli_close($conn)` της γραμμής 200, η οποία τερματίζει τη σύνδεση με τη βάση δεδομένων.

Για όσα αναφέρθηκαν στην ενότητα αυτή, σχετικά με το χειρισμό του CSV αρχείου από την PHP, αλλά και την αποθήκευση των δεδομένων του σε διασυνδεδεμένο πίνακα με τη βοήθεια της συνάρτησης `array_combine`, αντλήθηκαν ιδέες και πληροφορίες από τις παρακάτω ιστοσελίδες:

- <http://azuliadesigns.com/php-function-import-csv-array/>
- <http://www.webslesson.info/2016/10/import-csv-file-data-into-mysql-database-using-php-ajax.html>
- <https://magp.ie/2011/01/28/convert-csv-data-into-an-associative-array/>

4.4 Προβλήματα και λύσεις κατά την ανάλυση

Στην ενότητα αυτή γίνεται λόγος για τα προβλήματα που παρουσιάστηκαν κατά την ανάλυση του αρχείου `scopus.csv`, αλλά και για τους τρόπους αντιμετώπισης αυτών των προβλημάτων οι οποίοι υλοποιήθηκαν προγραμματιστικά στις γραμμές 32 έως 69 του αρχείου `parseCSV.php`.

Τα προβλήματα αυτά προκλήθηκαν από λάθη στην κωδικοποίηση του αρχείου `scopus.csv` που αντιστοιχούσε σε ορισμένα μέλη ΔΕΠ. Η εμφάνιση αυτών των λαθών είχε ως αποτέλεσμα τη μη σωστή λειτουργία της συνάρτησης `array_combine` στη γραμμή 72.

Η συνάρτηση `array_combine` δέχεται ως οράματα δυο πίνακες και δημιουργεί έναν νέο διασυνδεδεμένο πίνακα που περιέχει ως κλειδιά τα στοιχεία του πρώτου ορίσματος και ως τιμές τα στοιχεία του δεύτερου. Είναι προφανές ότι για να λειτουργήσει σωστά η συνάρτηση αυτή πρέπει και οι δυο πίνακες,

που δέχεται ως ορίσματα, να έχουν το ίδιο μήκος διαφορετικά δημιουργείται πρόβλημα στην αντιστοίχιση των κλειδιών με τις τιμές.

Ο πίνακας `$head` περιέχει σε κάθε του γραμμή κι από ένα πεδίο της επικεφαλίδας του αρχείου `scopus.csv` και επομένως έχει μήκος 44. Ο πίνακας `$lines`, περιέχει σε μορφή συμβολοσειράς κάθε μια γραμμή του αρχείου `scopus.csv`, έκτος της επικεφαλίδας. Κάθε μια από αυτές τις γραμμές αντιστοιχεί στα δεδομένα κάθε δημοσίευσης που έχει κάνει το μέλος ΔΕΠ στο οποίο ανήκει το αρχείο.

Σε κάθε γραμμή του πίνακα `$lines` περιέχονται 44 πεδία μέσα σε δίπλα εισαγωγικά διαχωρισμένα μεταξύ τους με κόμμα, όσα είναι δηλαδή και τα πεδία της επικεφαλίδας. Τα δεδομένα αυτά είναι για παράδειγμα στην μορφή "Arvanitakis K., Avlonitis M., Papadimitriou E.", "57191263146; 6506118499; 7005909662;", ... Το πρώτο από τα πεδία αυτά αντιστοιχεί στο πεδίο με όνομα Authors της επικεφαλίδας, το δεύτερο αντιστοιχεί στο πεδίο με όνομα Authors Ids κ.ο.κ.

Η συνάρτηση `str_getcsv` της PHP αναλύει μια συμβολοσειρά από πεδία σε μορφή CSV και επιστρέφει έναν πίνακα που περιέχει αυτά τα πεδία. Η συνάρτηση αυτή μπορεί να δεχτεί τρία ορίσματα, ένα υποχρεωτικό και αλλά δυο προαιρετικά. Το πρώτο από τα ορίσματα αυτά που είναι και υποχρεωτικό, πρέπει να είναι η συμβολοσειρά η οποία θα αναλυθεί. Το δεύτερο είναι ένας χαρακτήρας που καθορίζει το διαχωριστικό των πεδίων της συμβολοσειράς. Το όρισμα αυτό είναι προαιρετικό και αν δεν του δοθεί κάποια τιμή, η συνάρτηση θέτει αυτόματα την προεπιλεγμένη που είναι το κόμμα. Το τρίτο είναι ο χαρακτήρας μέσα στον οποίο περικλείονται τα πεδία. Το όρισμα αυτό είναι επίσης προαιρετικό και η προεπιλεγμένη τιμή είναι τα δίπλα εισαγωγικά.

Στη γραμμή 72 του αρχείου `parseCSV.php`, μέσω μιας επαναληπτικής διαδικασίας, κάθε γραμμή του πίνακα `$lines` δίνεται όρισμα στη συνάρτηση `str_getcsv` και μετατρέπεται σε πίνακα. Ο πίνακας αυτός μαζί με τον πίνακα `$head`, δίνονται ως ορίσματα στη συνάρτηση `array_combine` η οποία αποθηκεύει τα δεδομένα του αρχείου `scopus.csv` στο διασυνδεδεμένο πίνακα `$data`, με την μορφή που φαίνεται στην εικόνα 4.2.

Το σημαντικότερο πρόβλημα που προέκυψε οφείλεται στο ότι το όνομα ενός συνεδρίου περιέχει δίπλα εισαγωγικά, με συνέπεια η συνάρτηση `str_getcsv` να αντιλαμβάνεται τα εισαγωγικά αυτά ως ένα νέο πεδίο. Στην περίπτωση αυτή υπάρχουν γραμμές στον πίνακα που δημιουργεί η συνάρτηση οι οποίες έχουν περισσότερα από 44 πεδία, με αποτέλεσμα όταν συνδυαστούν οι γραμμές αυτές με τα στοιχεία του πίνακα `$head` μέσω της συνάρτησης `array_combine`, να προκαλείται πρόβλημα στην αντιστοίχιση.

Συγκεκριμένα υπάρχει συνέδριο με όνομα *2010 IEEE International Symposium on "A World of Wireless, Mobile and Multimedia Networks"*, *WoW-MoM 2010* και τα πρακτικά αυτού του συνεδρίου που έχουν τίτλο *2010 IEEE International Symposium on "A World of Wireless, Mobile and Multimedia Networks"*, *WoW-MoM 2010 - Digital Proceedings*. Για να μπορέσει η συνάρτηση `str_getcsv` να αναλύσει τα διπλά εισαγωγικά χωρίς να δημιουργήσει νέο

πεδίο πρέπει αυτά να είναι γραμμένα δυο φορές δηλαδή ως εξής: *2010 IEEE International Symposium on "A World of Wireless, Mobile and Multimedia Networks"*, *WoWMoM 2010 - Digital Proceedings*.

Στο πεδίο που αντιστοιχεί σε αυτό με ονομασία Source title της επικεφαλίδας, δηλαδή στον τίτλο των πρακτικών, τα διπλά εισαγωγικά είναι γραμμένα δυο φορές οπότε δεν δημιουργείται πρόβλημα. Στο πεδίο όμως που αντιστοιχεί στο Conference name, δηλαδή στο όνομα του συνεδρίου, τα διπλά εισαγωγικά είναι γραμμένα μία φορά, οπότε η συνάρτηση `str_getcsv` δημιουργεί δυο επιπλέον πεδία.

Τμήμα κώδικα 4.6: *parceCSV.php* γραμμές 32-69

```

32     foreach ($lines as $key => $line) {
33         if (count(str_getcsv($line, ",", '"')) == $headLength + 2) {
34             $lineForFixing = $line;
35             $regexForDetectBrokenPart = '/\s"[^"]+.?"/';
36             $regexForDetectTowExtraFilds = '/",{3}"/';
37             if (preg_match($regexForDetectBrokenPart, $lineForFixing, $match))
38                 ↪ {
39                 $brokenPart = $match[0];
39                 $fixedPart = str_replace('"', '"', $brokenPart);
40                 $fixedLine = preg_replace($regexForDetectBrokenPart, $fixedPart
41                     ↪ , $lineForFixing);
42                 $lines[$key] = $fixedLine;
42             } else if (preg_match($regexForDetectTowExtraFilds, $lineForFixing
43                 ↪ , $match)) {
43                 $extraCommas = $match[0];
44                 $oneComma = '"', '"';
45                 $fixedLine = preg_replace($regexForDetectTowExtraFilds,
46                     ↪ $oneComma, $lineForFixing);
46                 $lines[$key] = $fixedLine;
47             }
48         }
49         if (count(str_getcsv($line, ",", '"')) == $headLength + 1) {
50             $lineForFixing = $line;
51             $regexForDetectTowExtraFilds = '/",{2}"/';
52             if (preg_match($regexForDetectTowExtraFilds, $lineForFixing,
53                 ↪ $match)) {
53                 $extraCommas = $match[0];
54                 $oneComma = '"', '"';
55                 $fixedLine = preg_replace($regexForDetectTowExtraFilds,
56                     ↪ $oneComma, $lineForFixing);
56                 $lines[$key] = $fixedLine;
57             }
58         }
59         if (count(str_getcsv($line, ",", '"')) == $headLength - 1) {
60             $lineForFixing = $line;
61             $regexForDetectTowExtraFilds = '/",{6}"/';
62             if (preg_match($regexForDetectTowExtraFilds, $lineForFixing,
63                 ↪ $match)) {
63                 $extraCommas = $match[0];
64                 $oneComma = '"', '"', '"', '"', '"', '"';
65                 $fixedLine = preg_replace($regexForDetectTowExtraFilds,
66                     ↪ $oneComma, $lineForFixing);
66                 $lines[$key] = $fixedLine;
67             }
68         }
69     }

```

Η αντιμετώπιση του προβλήματος αυτού, καθώς και άλλων που θα αναφερθούν στη συνέχεια, υλοποιείται προγραμματιστικά με τη χρήση κανονικών εκφράσεων (regular expressions) στις γραμμές 32 έως 69 του αρχείου `parse-CSV.php`.

Στη γραμμή 32 ορίζεται μια επαναληπτική διαδικασία που θα προσπελάσει όλα τα στοιχεία του πίνακα `$lines`. Στη συνέχεια (γραμμή 33) ελέγχεται αν το μέγεθος του πίνακα που δημιουργεί η συνάρτηση `str_getcsv`, όταν της δοθεί ως όρισμα η τρέχουσα γραμμή του πίνακα `$lines`, είναι κατά δυο μονάδες μεγαλύτερο από το μέγεθος του πίνακα `$head`. Αν συμβαίνει αυτό, τα δεδομένα της τρέχουσας γραμμής του πίνακα `$lines` πρέπει να υποβληθούν σε επεξεργασία η οποία να αντιμετωπίζει το πρόβλημα. Για τον λόγο αυτό εκχωρούνται στη μεταβλητή `$lineForFixing` (γραμμή 34) και ακολούθως ορίζεται η κανονική έκφραση `/\s"[^"]".+?"/` που εκχωρείται στη μεταβλητή `$regexForDetectBrokenPart`. Αυτή η κανονική έκφραση θα αναζητήσει για χαρακτήρα κενού διαστήματος τον οποίο ακολουθούν δίπλα εισαγωγικά και μετά οποιοσδήποτε άλλος χαρακτήρας εκτός διπλών εισαγωγικών μια ή περισσότερες φορές (γραμμή 35).

Στην γραμμή 37 με την συνάρτηση `preg_match` ελέγχεται αν η κανονική έκφραση βρίσκει ταιρίασμα. Αν ναι, τότε το τμήμα της συμβολοσειράς `$lineForFixing` που περιέχει το ταιρίασμα εκχωρείται στη μεταβλητή `$brokenPart`. Στη συνέχεια σε οποιοδήποτε σημείο της `$brokenPart` υπάρχει ο χαρακτήρας των διπλών εισαγωγικών, αντικαθίσταται από δυο χαρακτήρες διπλών εισαγωγικών με τη χρήση της συνάρτησης `str_replace`, και η συμβολοσειρά που δημιουργείται εκχωρείται στη μεταβλητή `$fixedPart`. Η `$fixedPart` μαζί με την κανονική έκφραση `$regexForDetectBrokenPart` και την `$lineForFixing` δίνονται ορίσματα στη συνάρτηση `preg_replace` η οποία αντικαθιστά το τμήμα της `$lineForFixing` που ταιριάζει με την κανονική έκφραση και εκχωρεί το αποτέλεσμα, στη μεταβλητή `$fixedLine`. Η `$fixedLine` εκχωρείται στη θέση του πίνακα `$lines` που υπήρχε η συμβολοσειρά η οποία δημιουργούσε το πρόβλημα.

Ένα άλλο πρόβλημα που πρόέκυψε ήταν ότι σε ορισμένες γραμμές του αρχείου `scopus.csv` μετά το πεδίο που αντιστοιχεί στο Funding Text 1 της επικεφαλίδας, εμφανίζεται ο χαρακτήρας κόμμα τρεις επιπλέον φορές από ότι θα έπρεπε. Αυτό έχει ως αποτέλεσμα η συνάρτηση `str_getcsv` να δημιουργεί δυο πεδία περισσότερα από όσα έχει η επικεφαλίδα και επομένως η συνάρτηση `array_combine` να μην δίνει τα σωστά αποτελέσματα. Το πρόβλημα αυτό διορθώνεται στο τμήμα κώδικα των γραμμών 42 έως 46 όπου χρησιμοποιείται η κανονική έκφραση που ορίστηκε στη γραμμή 35, η οποία βρίσκει και αντικαθιστά τους τρεις αυτούς χαρακτήρες κόμμα από έναν.

Επίσης σε ορισμένες γραμμές του αρχείου `scopus.csv` μετά το πεδίο που αντιστοιχεί στο Funding Text 1 της επικεφαλίδας, εμφανίζεται ο χαρακτήρας κόμμα δυο επιπλέον φορές από ότι θα έπρεπε. Αυτό έχει ως αποτέλεσμα η συνάρτηση `str_getcsv` να δημιουργεί ένα πεδίο περισσότερο από όσα έχει η επικεφαλίδα. Για την αντιμετώπιση του προβλήματος αυτού εφαρμόζεται στις

γραμμές 49 έως 56 ακριβώς η ίδια διαδικασία για να αντικατασταθούν αυτή τη φορά οι δυο χαρακτήρες κόμμα από έναν.

Επιπλέον σε ορισμένες γραμμές του αρχείου scopus.csv αυτή τη φορά πριν το πεδίο που αντιστοιχεί στο Funding Text 1 της επικεφαλίδας εμφανίζεται ο χαρακτήρας κόμμα 6 συνεχόμενες φορές. Αυτό έχει ως αποτέλεσμα η συνάρτηση `str_getcsv` να δημιουργεί ένα πεδίο λιγότερο από όσα έχει η επικεφαλίδα. Θα επέτρεπε δηλαδή ο χαρακτήρας κόμμα να εμφανίζεται 7 συνεχόμενες φορές αντί για 6. Το πρόβλημα αυτό αντιμετωπίζεται στις γραμμές 59 έως 66 του αρχείου `parseCSV.php` με την ίδια ακριβώς διαδικασία, όπως αυτήν που περιγράφηκε πιο πάνω, και έτσι αντικαθίστούνται οι 6 συνεχόμενοι χαρακτήρες κόμμα από 7.

Αξίζει να σημειωθεί ότι τα προβλήματα με τους συνεχόμενους χαρακτήρες κόμμα που αναφέρθηκαν εδώ προέκυψαν από λάθη στην κωδικοποίηση του αρχείου scopus.csv στην έκδοση εκείνη κατά την οποία υλοποιήθηκε ο κώδικας. Αργότερα διαπιστώθηκε ότι σε επόμενες εκδόσεις του αρχείου τα λάθη αυτά είχαν διορθωθεί, από το ίδιο το Scopus.

Πληροφορίες σχετικά με τη χρήση κανονικών εκφράσεων στην PHP αντλήθηκαν από την ιστοσελίδα: <https://phpenthusiast.com/blog/php-regular-expressions>. Επίσης πολύ χρήσιμη αποδείχθηκε μια online εφαρμογή η οποία βοηθάει στον έλεγχο και τον εντοπισμό σφαλμάτων στη χρήση κανονικών εκφράσεων. Η εφαρμογή αυτή βρίσκεται στη διεύθυνση: <https://regex101.com/>.

Κεφάλαιο 5

Παρουσίαση της εφαρμογής

Στο κεφάλαιο αυτό παρουσιάζονται με τη βοήθεια στιγμιότυπων οθόνης (screenshots) όλες οι λειτουργίες τις οποίες εκτελεί η εφαρμογή που υλοποιήθηκε στα πλαίσια εκπόνησης της πτυχιακής εργασίας και για την οποία γίνεται λόγος σε αυτό το κείμενο.

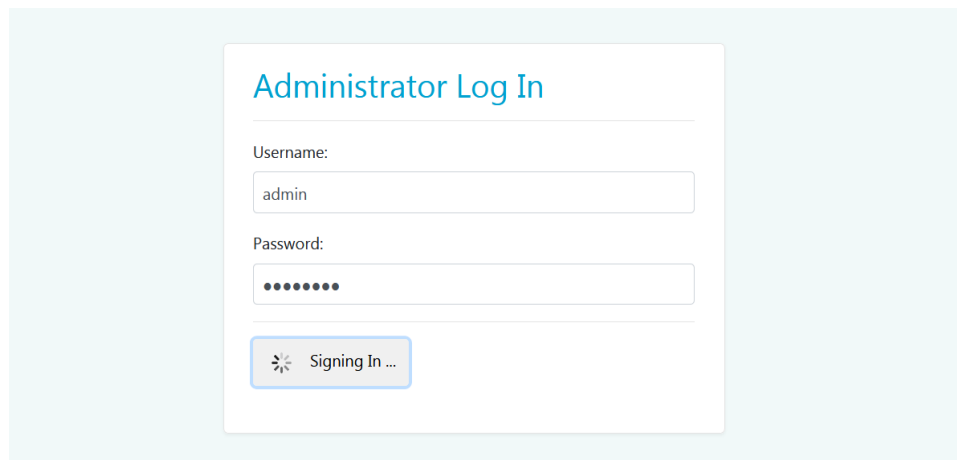
Όπως αναφέρθηκε και σε προηγούμενα κεφάλαια, η εφαρμογή έχει ως αντικείμενο της την παρουσίαση της ερευνητικής δραστηριότητας των μελλών ΔΕΠ του Τμήματος Πληροφορικής του Ιονίου Πανεπιστημίου. Η εφαρμογή υλοποιήθηκε με δεδομένα που αντλήθηκαν από αρχεία με επέκταση .csv, τα οποία βρίσκονται στον διαδικτυακό τόπο του Scopus, και αποτελείται από δυο τμήματα (modules), το τμήμα του διαχειριστή και το τμήμα των χρηστών του συστήματος.

5.1 Παρουσίαση του τμήματος για τον διαχειριστή του συστήματος

Στην παρούσα ενότητα παρουσιάζονται όλες οι λειτουργίες που έχει την δυνατότητα να εκτελέσει ο διαχειριστής του συστήματος. Οι λειτουργίες αυτές έχουν να κάνουν κυρίως με την αλληλεπίδραση του διαχειριστή με τη βάση δεδομένων, δηλαδή την εισαγωγή, την τροποποίηση, την ενημέρωση και τη διαγραφή των πληροφοριών που υπάρχουν στη βάση. Όλες αυτές οι ενέργειες εκτελούνται στον φυλλομετρητή (browser) μέσω μιας διεπαφής ιστού (web interface).

Στην εικόνα [5.1](#) φαίνεται η σελίδα εισόδου στο τμήμα της εφαρμογής που δημιουργήθηκε για τον διαχειριστή του συστήματος. Η σελίδα αυτή περιέχει τη φόρμα εισαγωγής του ονόματος χρήστη (username) και του κωδικού πρόσβασης (password) που είναι απαραίτητα για την είσοδο του διαχειριστή στο σύστημα.

Αφού δοθούν σωστά το όνομα χρήστη και ο κωδικός πρόσβασης εμφανίζεται η αρχική σελίδα του τμήματός του διαχειριστή. Ένα μέρος της σελίδας αυτής βλέπουμε στην εικόνα [5.2](#). Όπως μπορούμε να παρατηρήσουμε, η σελίδα

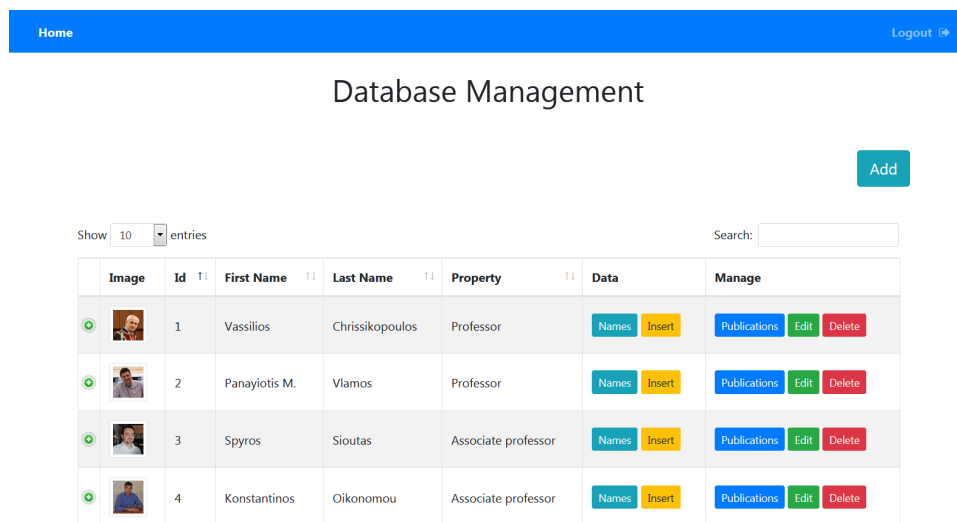


The image shows a login form titled "Administrator Log In". It has two input fields: "Username:" with the value "admin" and "Password:" with masked characters. Below the fields is a button labeled "Signing In ..." with a small icon.

Εικόνα 5.1: Η φόρμα εισόδου του διαχειριστή στο σύστημα

αυτή αποτελείται από το μενού πλοήγησης και έναν πίνακα ο οποίος περιέχει τα στοιχεία των μελών ΔΕΠ του Τμήματος Πληροφορικής του Ιονίου Πανεπιστημίου, που είναι αποθηκευμένα μέχρι αυτή τη στιγμή στη βάση δεδομένων της εφαρμογής. Στην εικόνα φαίνεται μόνο ένα μέρος του πίνακα αυτού.

Επιπλέον στη σελίδα υπάρχει ένα πτυσσόμενο μενού από το οποίο ο χρήστης μπορεί να επιλέξει τον αριθμό των εγγγραφών που επιθυμεί να βλέπει σελιδοποιημένες. Μπορούμε επίσης να διακρίνουμε ένα πλαίσιο κειμένου αναζήτησης των εγγγραφών αλλά και το κουμπί «Add» με το οποίο ο διαχειριστής έχει την δυνατότητα να προσθέσει ένα νέο μέλος ΔΕΠ στη βάση δεδομένων.



The image shows the "Database Management" interface. It has a blue header bar with "Home" and "Logout" links. Below the header is a table with 8 columns: Image, Id, First Name, Last Name, Property, Data, and Manage. There are 4 rows of data. Above the table is a search bar and a dropdown menu for "Show 10 entries". To the right of the table is an "Add" button.

Image	Id	First Name	Last Name	Property	Data	Manage
	1	Vassilios	Chrissikopoulos	Professor	Names Insert	Publications Edit Delete
	2	Panayiotis M.	Vlamos	Professor	Names Insert	Publications Edit Delete
	3	Spyros	Sioutas	Associate professor	Names Insert	Publications Edit Delete
	4	Konstantinos	Oikonomou	Associate professor	Names Insert	Publications Edit Delete

Εικόνα 5.2: Η αρχική σελίδα του διαχειριστή

Για κάθε μέλος ΔΕΠ στην στήλη «Manage» του πίνακα υπάρχουν τα κου-

μπιά με ονομασίες «Publications», «Edit» και «Delete». Με το πάτημα του κουμπιού «Publications» ο διαχειριστής του συστήματος μεταβαίνει στη σελίδα που περιχέει τις δημοσιεύσεις του συγκεκριμένου μέλους ΔΕΠ όπου εκεί έχει τη δυνατότητα να τις επεξεργαστεί, με τρόπο που θα αναφερθεί στη συνέχεια. Το κουμπί «Edit» παρέχει στον διαχειριστή την δυνατότητα να επεξεργαστεί τα στοιχεία του συγκεκριμένου μέλους ΔΕΠ. Πατώντας το κουμπί «Delete» εμφανίζεται ένα πλαίσιο διαλόγου, το οποίο ζητάει από τον διαχειριστή αν επιθυμεί να διαγράψει ή όχι το συγκεκριμένο μέλος ΔΕΠ από την βάση δεδομένων.

Στη στήλη «Data» του πίνακα υπάρχουν για κάθε μέλος ΔΕΠ τα κουμπιά με ονομασίες «Names» και «Insert». Με το πάτημα του κουμπιού «Insert» η εφαρμογή παίρνει εντολή να ξεκινήσει, για το αντίστοιχο μέλος ΔΕΠ, την ανάλυση του αρχείου csv, η οποία περιγράφεται στο προηγούμενο κεφάλαιο, και έχει ως αποτέλεσμα την εισαγωγή όλων των δημοσιεύσεων του συγκεκριμένου μέλους ΔΕΠ στη βάση δεδομένων. Για να γίνει με επιτυχία η ανάλυση του αρχείου csv, πρέπει να καταχωρηθούν στη βάση δεδομένων τα ονόματα με τα οποία το συγκεκριμένο μέλος ΔΕΠ αναγράφεται στο αρχείο csv. Με το κουμπί «Names» δίνεται στο διαχειριστή η δυνατότητα να εισάγει στη βάση δεδομένων της εφαρμογής τα ονόματα με τα οποία το συγκεκριμένο μέλος ΔΕΠ είναι καταχωρημένο στο αρχείο csv.

Στη συνέχεια γίνεται αναλυτική περιγραφή της διαδικασίας καταχώρισης ενός νέου μέλους ΔΕΠ και εισαγωγής των δημοσιεύσεων στις οποίες έχει λάβει μέρος στη βάση δεδομένων της εφαρμογής. Όπως αναφέρθηκε και σε προηγούμενα κεφάλαια οι δημοσιεύσεις ενός μέλους ΔΕΠ θα ληφθούν από τον διαδικτυακό τόπο του Scopus σε ένα αρχείο csv, το οποίο στη συνέχεια θα αναλυθεί από την εφαρμογή, και θα εισαχθούν οι εγγραφές του στη βάση δεδομένων, με τον τρόπο που περιγράφηκε στην παράγραφο 4.3.

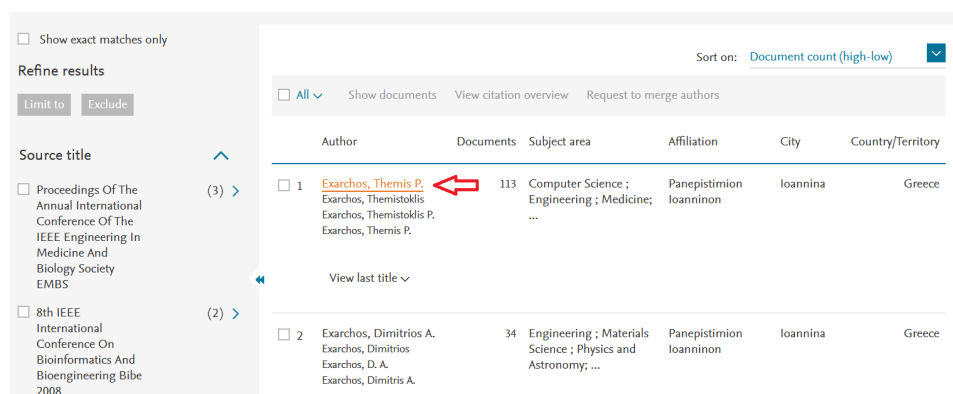
Εικόνα 5.3: Η ιστοσελίδα αναζήτησης συγγραφέα στο Scopus

Το νέο μέλος ΔΕΠ που θα καταχωρηθεί στην βάση δεδομένων είναι ο Επίκουρος Καθηγητής κύριος Θεμιστοκλής Έξαρχος. Αρχικά θα ληφθεί από

το διαδικτυακό τόπο του Scopus, και θα αποθηκευτεί τοπικά στον υπολογιστή, το αρχείο csv που περιέχει τις δημοσιεύσεις που έχει λάβει μέρος το εν λόγω μέλος ΔΕΠ. Το αρχείο αυτό στη συνέχεια, θα ανέβει στον διακομιστή της εφαρμογής ώστε να μπορεί να αναλυθεί.

Για τη λήψη του αρχείου csv με τις δημοσιεύσεις του κύριου Έξαρχου μεταβαίνουμε στην αρχική σελίδα του Scopus, ένα μέρος της οποίας φαίνεται στην εικόνα 5.3. Επιλέγουμε την καρτέλα «Authors» και στο πλαίσιο κειμένου με ετικέτα «Author last name» γράφουμε «Exarchos», στη συνέχεια πατάμε το κουμπί «Search».

Εμφανίζεται τότε η ιστοσελίδα με τα αποτελέσματα της αναζήτησης όλων των συγγραφέων που είναι καταχωρημένοι στη βάση δεδομένων του Scopus με το επώνυμο «Έξαρχος». Ένα τμήμα της ιστοσελίδας αυτής παρουσιάζεται στην εικόνα 5.4. Από τα αποτελέσματα αυτά το πρώτο είναι εκείνο που θέλουμε στην προκειμένη περίπτωση.



The screenshot shows the Scopus search results interface. On the left, there are filters for 'Source title' with two options: 'Proceedings Of The Annual International Conference Of The IEEE Engineering In Medicine And Biology Society EMBS' (3 results) and '8th IEEE International Conference On Bioinformatics And Bioengineering Bibe 2008' (2 results). The main table displays search results sorted by 'Document count (high-low)'. The first result is for 'Exarchos, Themis P.' with 113 documents. A red arrow points to this result. The second result is for 'Exarchos, Dimitrios A.' with 34 documents. The table has columns for Author, Documents, Subject area, Affiliation, City, and Country/Territory.

Author	Documents	Subject area	Affiliation	City	Country/Territory
1 Exarchos, Themis P. Exarchos, Themistoklis Exarchos, Themistoklis P. Exarchos, Themis P.	113	Computer Science ; Engineering ; Medicine; ...	Panepistimion Ioanninon	Ioannina	Greece
2 Exarchos, Dimitrios A. Exarchos, Dimitrios Exarchos, D. A. Exarchos, Dimitris A.	34	Engineering ; Materials Science ; Physics and Astronomy; ...	Panepistimion Ioanninon	Ioannina	Greece

Εικόνα 5.4: Τα αποτελέσματα της αναζήτησης

Ακολουθούμε το σύνδεσμο και οδηγούμαστε στην ιστοσελίδα που περιέχει όλες τις δημοσιεύσεις στη συγγραφή των οποίων έχει λάβει μέρος ο κύριος Έξαρχος. Προχωρώντας προς τα κάτω στη σελίδα πατάμε το σύνδεσμο «Export all» που φαίνεται στην εικόνα 5.5.



The screenshot shows the Scopus author profile for 'Themis P. Exarchos'. At the top, it displays statistics: '113 Documents', 'Cited by 1007 documents', '150 co-authors', and 'Author history'. Below these are links: 'View all in search results format >', 'Export all' (highlighted with a red box), 'Add all to list', 'Set document alert', and 'Set document feed'. The 'Sort on:' dropdown is set to 'Date (newest)'. The main table shows a list of documents with columns: Document title, Authors, Year, Source, and Cited by. The first document listed is 'Towards the Establishment of a Biomedical Ontology for the Primary Sjögren's Syndrome' by 'Pezoulas, V.C., Exarchos, T.P., Andronikou, V., (...), De Vita, S., Fotiadis, D.I.' published in 2018 in the 'Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS'.

Document title	Authors	Year	Source	Cited by
Towards the Establishment of a Biomedical Ontology for the Primary Sjögren's Syndrome	Pezoulas, V.C., Exarchos, T.P., Andronikou, V., (...), De Vita, S., Fotiadis, D.I.	2018	Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS 2018-July, 8513349, pp. 4089-4092	0

Εικόνα 5.5: Ο σύνδεσμος που οδηγεί στην εξαγωγή των αποτελεσμάτων

Στο παράθυρο που αναδύεται (εικόνα 5.6), δίνεται η δυνατότητα σε έναν

χρήστη του Scopus να εξάγει τις δημοσιεύσεις ενός ερευνητή σε διάφορες μορφές αρχείων, μεταξύ των οποίων και σε αρχεία τύπου csv. Από τις διαθέσιμες επιλογές στο τμήμα παράθυρου με ονομασία «Select your method of export» διαλέγουμε τη csv, ενώ από το τμήμα «What information do you want to export?» επιλέγουμε όλες τις δυνατές επιλογές που είναι διαθέσιμες και στη συνέχεια κάνουμε κλικ στο κουμπί Export.

Export document settings

You have chosen to export 113 documents

Select your method of export

☐ MENDELEY
 ☐ RefWorks
 ☐ RIS Format EndNote, Reference Manager
 ☒ CSV Excel
 ☐ BibTeX
 ☐ Plain Text ASCII in HTML

What information do you want to export?

Citation information	Bibliographical information	Abstract & keywords	Funding details	Other information
☒ Author(s)	☒ Affiliations	☒ Abstract	☒ Number	☒ Tradenames & manufacturers
☒ Author(s) ID	☒ Serial identifiers (e.g. ISSN)	☒ Author keywords	☒ Acronym	☒ Accession numbers & chemicals
☒ Document title	☒ PubMed ID	☒ Index keywords	☒ Sponsor	☒ Conference information
☒ Year	☒ Publisher		☒ Funding text	☒ Include references
☒ EID	☒ Editor(s)			
☒ Source title	☒ Language of original document			
☒ volume, issue, pages	☒ Correspondence address			
☒ Citation count	☒ Abbreviated source title			
☒ Source & document type				
☒ Publication Stage				
☒ DOI				
☒ Access Type				

Cancel Export

Εικόνα 5.6: Η επιλογή των πεδίων που θα περιέχει το αρχείο csv.

Με την εκτέλεση της προηγούμενης ενέργειας εμφανίζεται το παράθυρο που παρουσιάζεται στην εικόνα 5.7. Επιλέγουμε αποθήκευση αρχείου και πατάμε το κουμπί OK.

Ανοίγματος scopus.csv

Επιλέξτε να ανοίξετε:

scopus.csv

που είναι: Αρχείο τιμών διαχωρισμένων με κόμμα του Microsoft Office
από: <https://www.scopus.com>

Τι να κάνει το Firefox με αυτό το αρχείο:

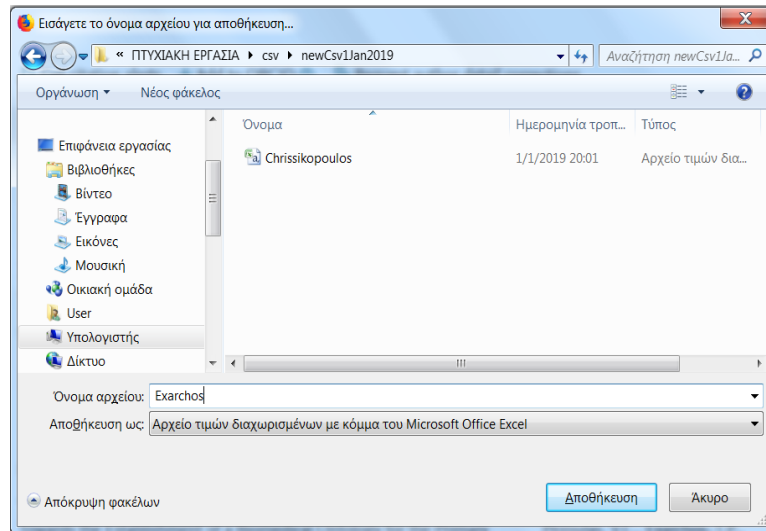
☐ Ανοίγματος με Microsoft Office Excel (προεπιλογή)
 ☒ Αποθήκευση αρχείου

☐ Να γίνεται αυτόματα από εδώ και πέρα για αρχεία αυτού του είδους.

OK Ακύρωση

Εικόνα 5.7: Η εξαγωγή των αποτελεσμάτων σε αρχείο csv

Στο επόμενο παράθυρο, (εικόνα 5.8) επιλέγουμε το φάκελο που θέλουμε να αποθηκευτεί το αρχείο, γράφουμε το όνομα του αρχείου και πατάμε αποθήκευση.



Εικόνα 5.8: Αποθήκευση του αρχείου csv

Στη συνέχεια γίνεται προσθήκη των στοιχείων του κυρίου Έξαρχου στη βάση δεδομένων της εφαρμογής. Στην αρχική σελίδα του διαχειριστή του συστήματος που φαίνεται στην εικόνα 5.2 πατάμε το κουμπί Add και εμφανίζεται τότε το παράθυρο της εικόνας 5.9 στο οποίο υπάρχει η φόρμα εισαγωγής των στοιχείων.

Add User

Enter First Name

Themistoklis

Enter Last Name

Exarchos

Enter Property

Assistant Professor

Upload Person's CSV file

Aναζήτηση...

Exarchos.csv

Select Person's Image

Aναζήτηση...

16.jpg

Add

Close

Εικόνα 5.9: Η φόρμα εγγραφής νέου μέλους

Όπως φαίνεται στην εικόνα αναγράφουμε το όνομα, το επώνυμο και τη βαθμίδα που κατέχει στο Τμήμα το νέο μέλος ΔΕΠ, στη συνέχεια επιλέγουμε το αρχείο csv και τη φωτογραφία που του αντιστοιχούν. Τέλος πατάμε το κουμπί Add για να καταχωρηθούν τα στοιχεία αυτά στη βάση δεδομένων. Η νέα εγγραφή με τα στοιχεία που καταχωρήθηκαν από την προηγούμενη διαδικασία υπάρχει στο τέλος του πίνακα με τα μέλη ΔΕΠ, μέρος του οποίου εμφανίζεται στην εικόνα 5.2.

Εικόνα 5.10: Πληροφορίες για έναν ερευνητή στο Scopus

Στη συνέχεια θα εισαχθούν στη βάση δεδομένων τα ονόματα με τα οποία ο κύριος Έξαρχος αναγράφεται στο αρχείο csv. Τα ονόματα αυτά μπορούν να βρεθούν στην ιστοσελίδα του Scopus η οποία περιέχει τις δημοσιεύσεις του κυρίου Έξαρχου, στο σημείο που φαίνεται στην εικόνα 5.10. Από τα ονόματα αυτά χρειάζονται εκείνα που βρίσκονται μέσα στο κόκκινο πλαίσιο αφού, όπως αναφέρθηκε και στην ενότητα 4.2 (σελίδα 40), μόνο με αυτές τις δυο μορφές τα ονόματα αναγράφονται στο αρχείο csv.

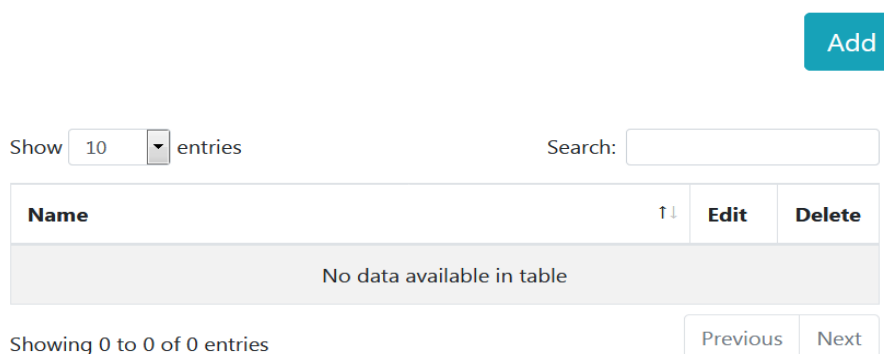
	11	Markos	Avlonitis	Assistant Professor	Names	Insert	Publications	Edit	Delete
	12	Michalis	Stefanidakis	Assistant Professor	Names	Insert	Publications	Edit	Delete
	13	Konstantinos	Chorianopoulos	Assistant Professor	Names	Insert	Publications	Edit	Delete
	14	Katia	Kermanidis	Assistant Professor	Names	Insert	Publications	Edit	Delete
	15	Aggeliki	Tsohou	Assistant Professor	Names	Insert	Publications	Edit	Delete
	19	Themistoklis	Exarchos	Assistant Professor	Names	Insert	Publications	Edit	Delete

Εικόνα 5.11: Μέρος του πίνακα με τα καταχωρημένα μέλη ΔΕΠ

Για να γίνει η εισαγωγή των ονομάτων αυτών στη βάση δεδομένων, μεταβαίνουμε στην αρχική σελίδα του διαχειριστή, στο σημείο το οποίο βρίσκεται

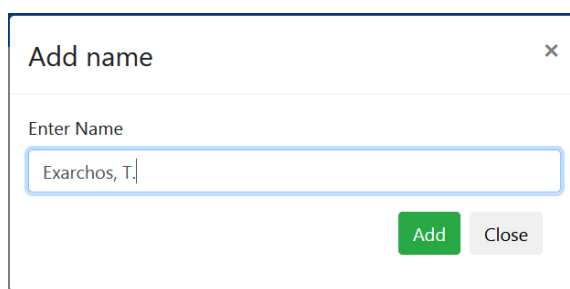
ο πίνακας που περιέχει τα καταχωρημένα στη βάση μέλη ΔΕΠ. Στη συνέχεια, πατάμε το κουμπί με ονομασία Names που αντιστοιχεί στον κύριο Έξαρχο, και το οποίο είναι αυτό που βρίσκεται στο εσωτερικό του κόκκινου πλαισίου της εικόνας 5.11.

Με το πάτημα του κουμπιού Names πλοηγούμαστε στη σελίδα διαχείρισης των ονομάτων που αναγράφονται στο αρχείο csv το οποίο αντιστοιχεί στο συγκεκριμένο μέλους ΔΕΠ. Όπως φαίνεται και στην εικόνα 5.12, μέχρι στιγμής δεν υπάρχουν καταχωρήσεις.



Εικόνα 5.12: Ο πίνακας με τα ονόματα του μέλους ΔΕΠ χωρίς δεδομένα

Για να προσθέσουμε ένα νέο όνομα πατάμε το κουμπί Add. Εμφανίζεται τότε το παράθυρο της εικόνας 5.13, όπου στο πλαίσιο κειμένου γραφούμε Exarchos, T. και πατάμε το κουμπί Add, ώστε να προστεθεί αυτή η εγγραφή στη βάση δεδομένων. Στη συνέχεια επαναλαμβάνουμε την προαναφερθείσα διαδικασία, μόνο που τώρα στο πλαίσιο κειμένου γραφούμε Exarchos, T.P. ώστε να γίνει η καταχώριση και της δεύτερης μορφής του ονόματος στη βάση δεδομένων.



Εικόνα 5.13: Προσθήκη ονόματος

Στην εικόνα 5.14 φαίνεται ότι οι δυο προηγούμενες μορφές του ονόματος έχουν πλέον καταχωρηθεί στη βάση δεδομένων και είναι διαθέσιμες για επεξεργασία, σε περίπτωση λάθους, με το κουμπί Edit ή για διαγραφή, με το κουμπί Delete.

Show entries Search:

Name	↑↓	Edit	Delete
Exarchos T.		<button>Edit</button>	<button>Delete</button>
Exarchos T.P.		<button>Edit</button>	<button>Delete</button>

Showing 1 to 2 of 2 entries

Previous **1** Next

Εικόνα 5.14: Ο πίνακας με τα ονόματα του μέλους ΔΕΠ

Μετά την καταχώριση στη βάση δεδομένων των ονομάτων με τα οποία το νέο μέλος ΔΕΠ εμφανίζεται στο αρχείο csv, μπορεί να ξεκινήσει η διαδικασία ανάλυσης του αρχείου από την εφαρμογή. Η διαδικασία αυτή έχει ως αποτέλεσμα την εισαγωγή των δημοσιεύσεων του νέου μέλους ΔΕΠ στη βάση δεδομένων. Επιπλέον αν κάποιες από τις δημοσιεύσεις αυτές υπάρχουν ήδη στη βάση δεδομένων, η διαδικασία ανάλυσης ενημερώνει της αναφορές (citations) που έχουν λάβει οι δημοσιεύσεις αυτές, στην περίπτωση που είναι διαφορετικές από τις προϋπάρχουσες. Ο κώδικας σε γλώσσά PHP με τον οποίο υλοποιείται η ανάλυση του αρχείου csv, περιγράφεται στην ενότητα 4.3.

Για να ξεκινήσει η διαδικασία ανάλυσης του αρχείου csv που αντιστοιχεί στον κύριο Έξαρχο μεταβαίνουμε στην αρχική σελίδα του διαχειριστή. Βρίσκουμε τη γραμμή του πίνακα στην οποία αναγράφονται τα στοιχεία του εν λόγω μέλους ΔΕΠ και πατάμε το κουμπί Insert (εικόνα 5.11).

Εμφανίζεται τότε το παράθυρο της εικόνας 5.15 στο οποίο ζητείται αν επιθυμούμε να ξεκινήσει η ανάλυση ή όχι. Πατάμε το κουμπί YES και έτσι αρχίζει η εισαγωγή και η ενημέρωση των δεδομένων των δημοσιεύσεων στις οποίες έχει λάβει μέρος στη συγγραφή τους ο κύριος Έξαρχος.

Insert or update data for Themistoklis Exarchos

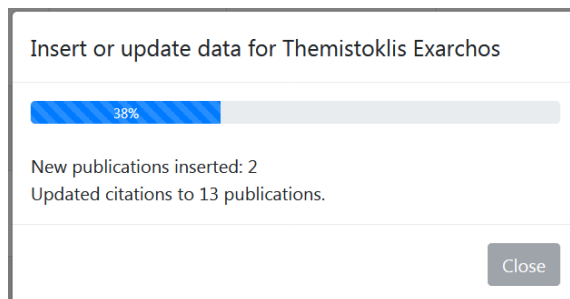
Do you want to insert or update data for **Themistoklis Exarchos** ?

YES NO

Εικόνα 5.15: Το παράθυρο επιβεβαίωσης εκκίνησης της ανάλυσης

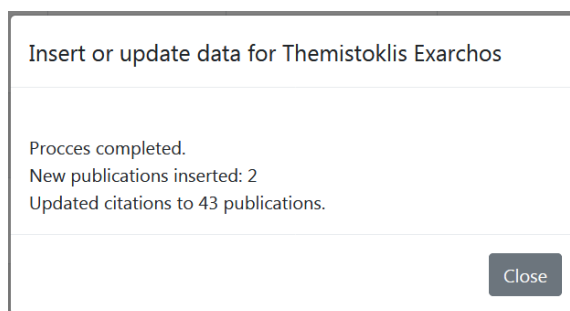
Στην εικόνα 5.16 φαίνεται η μπάρα που δείχνει το ποσοστό ολοκλήρωσης της διαδικασίας ανάλυσης το οποίο στην προκειμένη περίπτωση είναι 38% και μέχρι εκείνο το σημείο έχουν εισαχθεί στη βάση δεδομένων δυο νέες δημοσιεύ-

ύσεις ενώ έχουν ανανεωθεί οι αναφορές από δεκατρείς.



Εικόνα 5.16: Η μπάρα προόδου της διαδικασίας

Όταν ολοκληρωθεί η διαδικασία στο παράθυρο εμφανίζεται το μήνυμα της εικόνας 5.17 το οποίο μας ενημερώνει ότι εισήχθησαν στη βάση δεδομένων δύο νέες δημοσιεύσεις ενώ ενημερώθηκαν οι αναφορές από 43 που ήδη προϋπήρχαν.



Εικόνα 5.17: Το μήνυμα στο τέλος της διαδικασίας

Μπορούμε να δούμε όλες τις δημοσιεύσεις του κύριου Έξαρχου πατώντας το κουμπί Publications του κόκκινου πλαισίου της εικόνας 5.11.

Η παρουσίαση του τμήματος της εφαρμογής για το διαχειριστή του συστήματος συνεχίζεται με τη διαδικασία διαγράψης από τη βάση δεδομένων των δημοσιεύσεων που δεν ανήκουν σε κάποιο συγκεκριμένο μέλος ΔΕΠ. Η ανάγκη διαγραφής προέρχεται κυρίως από το γεγονός ότι διαπιστώθηκε πως σε ορισμένα αρχεία csv υπάρχουν δημοσιεύσεις που δεν αντιστοιχούν στο συγκεκριμένο μέλος ΔΕΠ. Το λάθος αυτό προκύπτει από τον τρόπο με τον οποίο το σύστημα του Scopus δημιουργεί τα αρχεία csv και οφείλεται κυρίως σε λόγους συνωνυμίας των ερευνητών.

Χαρακτηριστικό είναι το παράδειγμα των δημοσιεύσεων του μέλους ΔΕΠ κύριου *Μιχάλη Στεφανιδάκη* όπου στις δημοσιεύσεις του υπάρχουν και ορισμένες που σχετίζονται με τις επίσημες της Βιολογίας και της Ιατρικής. Διαπιστώθηκε ότι οι δημοσιεύσεις αυτές δεν ανήκαν στο εν λόγω μέλος ΔΕΠ και το λάθος αυτό οφείλεται σε συνωνυμία με τον ερευνητή κύριο *Μιχάλη Στεφανιδάκη*, ο οποίος εργάζεται στο Πανεπιστήμιο του Ελσίνκι. Στη συνέχεια θα πραγματοποιηθεί διαγραφή των δημοσιεύσεων αυτών.

Πατώντας το κουμπί Publications στη γραμμή του πίνακα που αντιστοιχεί στον κύριο Σταφανιδάκη (εικόνα 5.11) μεταβαίνουμε στην σελίδα με τις δημοσιεύσεις του εν λόγω μέλους ΔΕΠ, ένα μέρος της οποίας φαίνεται στην εικόνα 5.18 όπου οι δημοσιεύσεις εμφανίζονται σελιδοποιημένες και ταξινομημένες σε φθίνουσα σειρά με βάση τον αριθμό των αναφορών.

Publicatins of Michalis Stefanidakis

Show entries Search:

	Title	Year	Type	Cited by	
	Cell-surface association between matrix metalloproteinases and integrins: Role of the complexes in leukocyte migration and cancer progression	2006	Journal Article	120	
	Activation of NMDA receptors promotes dendritic spine development through MMP-mediated ICAM-5 cleavage	2007	Journal Article	113	
	α -Chain phosphorylation of the human leukocyte CD11b/CD18 (Mac-1) integrin is pivotal for integrin activation to bind ICAMs and leukocyte extravasation	2006	Journal Article	63	
	Identification of a Negatively Charged Peptide Motif within the Catalytic Domain of Progelatinases That Mediates Binding to Leukocyte β 2 Integrins	2003	Journal Article	55	
	Intracellular and cell surface localization of a complex between α M β 2 integrin and promatrix metalloproteinase-9 progelatinase in neutrophils	2004	Journal Article	44	

Εικόνα 5.18: Ο πίνακας με τις δημοσιεύσεις ενός μέλους ΔΕΠ

Μπορούμε να επιλέξουμε κάθε δημοσίευση πατώντας το μέσα στο τετραγώνάκι που της αντιστοιχεί, στο δεξιό μέρος του πίνακα. Πατώντας το τετραγώνάκι της επικεφαλίδας επιλέγονται όλες οι δημοσιεύσεις. Από τις δημοσιεύσεις αυτές οι πρώτες 11 είναι εκείνες που δεν ανήκουν στον κύριο Σταφανιδάκη. Για να τις διαγράψουμε, τις επιλέγουμε και πατάμε το κουμπί Delete που βρίσκεται στο τέλος της σελίδας, όπως φαίνεται και στην επόμενη εικόνα.

	Characterization of ICAM-4 binding to the I domains of the CD11a/CD18 and CD11b/CD18 leukocyte integrins	2003	Journal Article	36	☒
	Endothelial CD47 interaction with SIRP(gamma) is required for human T-cell transendothelial migration under shear flow conditions in vitro	2008	Journal Article	35	☒
	Peptide-mediated delivery of therapeutic and imaging agents into mammalian cells	2004	Journal Article	35	☒
	Endothelial CD47 promotes vascular endothelial-cadherin tyrosine phosphorylation and participates in T cell recruitment at sites of inflammation in vivo	2012	Journal Article	18	☒

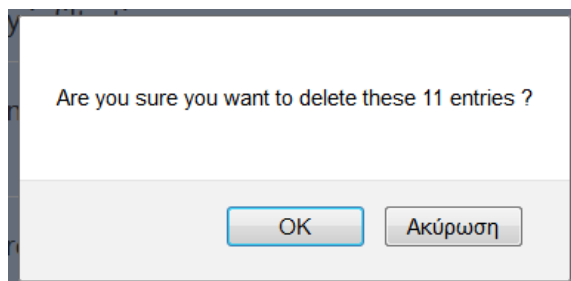
Showing 1 to 10 of 45 entries 11 rows selected

Previous **1** 2 3 4 5 Next

Delete

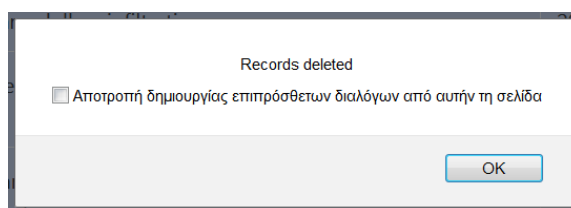
Εικόνα 5.19: Επιλογή των δημοσιεύσεων προς διαγραφή

Εμφανίζεται, τότε το επόμενο παράθυρο στο οποίο η εφαρμογή ζητάει αν πράγματι είναι επιθυμητή η διαγραφή αυτών των 11 δημοσιεύσεων.



Εικόνα 5.20: Το παράθυρο επιβεβαίωσης της διαγραφής

Κάνουμε κλικ στο κουμπί OK για να πραγματοποιηθεί η διαγραφή, και τότε εμφανίζεται το παράθυρο της εικόνας 5.21 που μας ενημερώνει ότι η διαγραφή πραγματοποιήθηκε.



Εικόνα 5.21: Πραγματοποίηση της διαγραφής

5.2 Παρουσίαση του τμήματος για τους χρήστες του συστήματος

Στην ενότητα αυτή παρουσιάζονται όλες οι λειτουργίες της εφαρμογής που είναι δυνατόν να γίνουν όταν κάποιος χρησιμοποιεί το τμήμα χρηστών αυτής. Οι λειτουργίες αυτές έχουν να κάνουν κυρίως με την ποσοτική και ποιοτική παρουσίαση της έρευνας που έχει συντελεστεί συνολικά από το Τμήμα Πληροφορικής του Ιονίου Πανεπιστημίου, αλλά και από κάθε μέλος του ξεχωριστά, στο πέρασμα των ετών.

Στην εικόνα 5.22 φαίνεται η αρχική σελίδα του τμήματος των χρηστών της εφαρμογής. Όπως μπορεί κάποιος να διακρίνει στη σελίδα αυτή υπάρχουν σύνδεσμοι προς τους διαδικτυακούς τόπους του Ιονίου Πανεπιστημίου, του Τμήματος Πληροφορικής του Ιδρύματος αυτού, καθώς επίσης και προς το διαδικτυακό τόπο του Scopus, από τον οποίο λήφθηκαν τα δεδομένα. Επιπλέον υπάρχουν έξι σύνδεσμοι που οδηγούν προς εσωτερικές σελίδες της εφαρμογής οι οποίες είναι οι επόμενες:



Εικόνα 5.22: Η αρχική σελίδα του τμήματος των χρηστών του συστήματος

- **All Research**

Με τη χρήση πινάκων και γραφημάτων γίνεται παρουσίαση συγκεντρωτικών στοιχείων όλης της έρευνας που έχει πειραχθεί από το Τμήμα Πληροφορικής.

- **Professors**

Περιέχει μια λίστα με όλα τα μέλη ΔΕΠ του Τμήματος Πληροφορικής του Ιονίου Πανεπιστημίου, τα οποία μέχρι αυτή τη στιγμή είναι καταχωρημένα στη βάση δεδομένων της εφαρμογής. Εν συνεχεία ο χρήστης ακολουθώντας κατάλληλο σύνδεσμο έχει τη δυνατότητα να πλοηγηθεί στην σελίδα που περιέχει πληροφορίες σχετικά με την ερευνητική δραστηριότητα ενός συγκεκριμένου μέλους.

- **All years**

Στη σελίδα αυτή υπάρχουν συγκεντρωτικοί πίνακες που δείχνουν τον αριθμό των δημοσιεύσεων, τις οποίες έχει πραγματοποιήσει ένα συγκεκριμένο μέλος ΔΕΠ ανά έτος, σε όλη τη διάρκεια της ερευνητικής του δραστηριότητας.

- **Timeline**

Η σελίδα αυτή περιέχει μια χρονολογική γραμμή (timeline) η οποία δίνει τη δυνατότητα στο χρήστη να πλοηγηθεί από το πρώτο έως το τελευταίο έτος έρευνας του Τμήματος και να δει έναν πίνακα με τον αριθμό των δημοσιεύσεων κάθε μέλους ΔΕΠ, σε κάθε έτος έρευνας. Επιπλέον ένα χρονολογικό γραφήματά, απεικονίζει την εξέλιξη της ερευνητικής δραστηριότητας του Τμήματος ανά έτος.

- **Journals**

Στη σελίδα αυτή παρουσιάζονται σε έναν πίνακα τα ονόματα των περιοδικών στα όποια έχουν δημοσιευτεί τα επιστημονικά άρθρα όλων των μελών ΔΕΠ του τμήματος, καθώς και ο αριθμός των άρθρων που έχουν δημοσιευτεί σε κάθε περιοδικό.

- **Conferences**

Στη σελίδα αυτή παρουσιάζονται σε έναν πίνακα τα ονόματα των συνεδρίων στα οποία έχουν παρουσιαστεί άρθρα όλων των μελών ΔΕΠ του τμήματος, καθώς και ο αριθμός των άρθρων που έχουν παρουσιαστεί σε κάθε συνέδριο.

Στη συνέχεια παρουσιάζεται εκτενέστερα καθεμία από τις προηγούμενες έξι σελίδες της εφαρμογής. Στο σημείο αυτό θα πρέπει να τονιστεί ότι όλα τα δεδομένα τα οποία περιέχονται στις ιστοσελίδες αυτές προκύπτουν από ερωτήματα (queries) του έχουν γίνει στη βάση δεδομένων με εντολές γλωσσάς SQL ενσωματωμένες σε κώδικα PHP. Τα δεδομένα αυτά είναι δυναμικά, δηλαδή αλλάζουν κάθε φορά που πραγματοποιείται εισαγωγή, ενημέρωση, τροποποίηση ή διαγραφή εγγγραφών στη βάση δεδομένων από το διαχειριστή του συστήματος.

5.2.1 Ιστοσελίδα All Research

Στη σελίδα All Research περιέχονται πληροφορίες που αφορούν ολόκληρη την έρευνα που έχει πειραχθεί από τα μέλη ΔΕΠ του Τμήματος Πληροφορικής του Ιονίου Πανεπιστημίου, που μέχρι αυτή τη στιγμή είναι καταχωρημένα στη βάση δεδομένων της εφαρμογής.

Document Types	
Document Types	Number
Journal Articles	372
Conference Articles	475
Book Chapters	27
Books	3
Total	877

Εικόνα 5.23: Πληροφορίες για τον αριθμό των δημοσιεύσεων ανα τύπο

Στην εικόνα 5.23 φαίνεται το πάνω μέρος της σελίδας, στο οποίο διακρίνεται το μενού πλοήγησης προς τις άλλες σελίδες της εφαρμογής. Επιπλέον υπάρχει

ένας πίνακας ο οποίος αναγράφει τον αριθμό του είδους κάθε δημοσίευσης (άρθρα σε επιστημονικά περιοδικά, άρθρα σε συνέδρια, βιβλία και κεφάλαια σε βιβλία) που έχουν γίνει από όλα τα μέλη ΔΕΠ. Στην τελευταία γραμμή του πίνακα μπορούμε να ενημερωθούμε για το συνολικό αριθμό των δημοσιεύσεων που υπάρχουν στη βάση δεδομένων.

Προχωρώντας προς τα κάτω στη σελίδα θα βρούμε τις καρτέλες με ονομασία All Publications, Journal Articles, Conference Articles, Book Chapters και Books. Σε καθεμία από τις καρτέλες αυτές περιέχονται σελιδοποιημένα σε πίνακα δεδομένα που αντιστοιχούν στις δημοσιεύσεις κάθε κατηγορίας. Για τις καρτέλες Journal Articles, Conference Articles, Book Chapters και Books, οι κεφαλίδες των πινάκων αποτελούνται από τα πεδία Title, Year και Cited by. Κάνοντας κλικ σε καθεμία από τις κεφαλίδες αυτές είναι δυνατόν να ταξινομήσουμε τα δεδομένα κάθε πίνακα κατά τίτλο, έτος και αριθμό αναφορών σε αύξουσα ή φθίνουσα σειρά.

Publications

All PublicationsJournal ArticlesConference ArticlesBook ChaptersBooks

Show10▼entriesSearch:

	Title	Year	Type	Cited by
🕒	Machine learning applications in cancer prognosis and prediction	2015	Journal Article	219
🕒	Adaptive probabilistic search for peer-to-peer networks	2003	Conference Article	193
🕒	A research framework for analysing eBusiness models	2004	Journal Article	181
🕒	Integrating GIS, GPS and GSM technologies for the effective management of ambulances	2001	Journal Article	92
🕒	Automated diagnosis of coronary artery disease based on data mining and fuzzy modeling	2008	Journal Article	90
🕒	A methodology for the automated creation of fuzzy expert systems for ischaemic and arrhythmic beat classification based on a set of rules obtained by a decision tree	2007	Journal Article	84
🕒	Semantic image segmentation and object labeling	2007	Journal Article	78
🕒	Secure transactions with mobile agents in hostile environments	2000	Conference Article	75
🕒	Parametric person identification from the EEG using computational geometry	1999	Conference Article	68
🕒	An association rule mining-based methodology for automated detection of ischemic ECG beats	2006	Journal Article	60
	Title	Year	Type	Cited by

Showing 1 to 10 of 877 entries

Previous

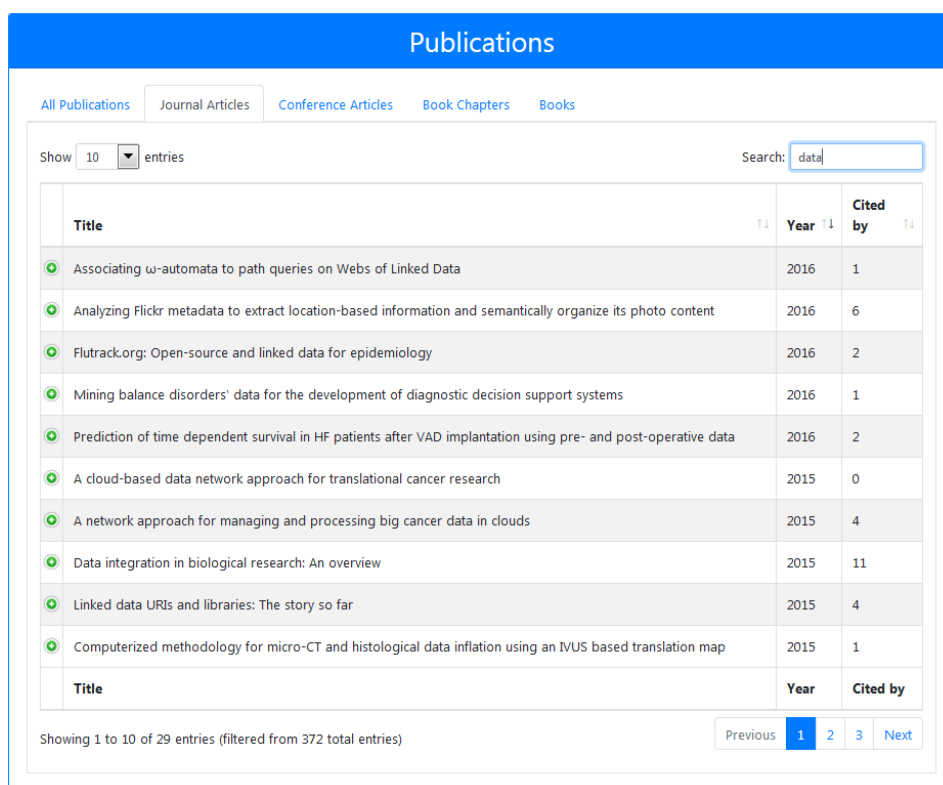
12345...88Next

Εικόνα 5.24: Ο πίνακας με τα στοιχεία όλων των δημοσιεύσεων

Οπώς φαίνεται στην εικόνα 5.24 η καρτέλα All Publications περιέχει πίνακα με τα δεδομένα όλων των δημοσιεύσεων που υπάρχουν στη βάση, δίχως αυτά

να είναι διαχωρισμένα ανάλογα με το είδος της δημοσίευσης. Για το λόγο αυτό στην κεφαλίδα του πίνακα βρίσκεται ένα ακόμα πεδίο με όνομα *Type*, το οποίο υπάρχει για να ξεχωρίσει κάθε δημοσίευση στον τύπο της. Όπως και στους πίνακες των άλλων καρτελών, έτσι και σε αυτόν τα δεδομένα κάθε στήλης μπορούν να ταξινομηθούν σε αύξουσα ή φθίνουσα σειρά.

Επίσης μπορούμε να διακρίνουμε, ότι για τα δεδομένα του πίνακα της καρτέλας αυτής, αλλά και των άλλων καρτελών, υπάρχει η δυνατότητα αναζήτησης των εγγραφών πληκτρολογώντας την επιθυμητή λέξη ή φάση στο πεδίο *Search*. Επιπλέον υπάρχει η δυνατότητα να γίνει σελιδοποίηση των αποτελεσμάτων ανάμεσα σε 10 (προεπιλογή) 25, 50 και 100, εγγραφές, επιλέγοντας το αντίστοιχο νούμερο από το πτυσσόμενο μενού. Στην εικόνα 5.25 φαίνονται τα 10 πρώτα, από τα 29 συνολικά, αποτελέσματα αναζήτησης για την λέξη «data» στην καρτέλα *Journal Articles*.



Publications			
All Publications Journal Articles Conference Articles Book Chapters Books			
Show 10 entries		Search: data	
Title	Year	Cited by	
Associating ω-automata to path queries on Webs of Linked Data	2016	1	
Analyzing Flickr metadata to extract location-based information and semantically organize its photo content	2016	6	
Flutrack.org: Open-source and linked data for epidemiology	2016	2	
Mining balance disorders' data for the development of diagnostic decision support systems	2016	1	
Prediction of time dependent survival in HF patients after VAD implantation using pre- and post-operative data	2016	2	
A cloud-based data network approach for translational cancer research	2015	0	
A network approach for managing and processing big cancer data in clouds	2015	4	
Data integration in biological research: An overview	2015	11	
Linked data URIs and libraries: The story so far	2015	4	
Computerized methodology for micro-CT and histological data inflation using an IVUS based translation map	2015	1	
Title	Year	Cited by	

Showing 1 to 10 of 29 entries (filtered from 372 total entries)

Previous 1 2 3 Next

Εικόνα 5.25: Αναζήτηση αποτελεσμάτων

Πατώντας το σύμβολο + που βρίσκεται μέσα σε έναν κύκλο με πράσινο φόντο στην πρώτη στήλη του πίνακα, μπορούμε να δούμε επιπλέον πληροφορίες για τη συγκεκριμένη δημοσίευση όπως, για παράδειγμα, τους συγγραφείς της, τον τύπο της (*Journal article*, *Conference article*, *Book chapter*, *Book*) και κάποια αλλά επιπλέον στοιχεία ανάλογα με τον τύπο της δημοσίευσης αυτής.

Στην εικόνα 5.26 βλέπουμε τη μορφή που παίρνει ο πίνακας αμέσως μετά

το πάτημα του σύμβολου +. Μπορούμε να διακρίνουμε ότι ο πίνακας έχει επεκταθεί ώστε να υπάρχουν τώρα πρόσθετα στοιχεία για τη συγκεκριμένη δημοσίευση, όπως το όνομα του συνεδρίου στο οποίο παρουσιάστηκε, καθώς επίσης την τοποθεσία και την ημερομηνία διεξαγωγής του. Επιπλέον μπορούμε να δούμε τον τίτλο και τις σελίδες των πρωκτικών που δημοσιεύτηκε, αλλά και τους συγγραφείς της. Για όσους από τους συγγραφείς είναι μέλη ΔΕΠ του Τμήματος Πληροφορικής, υπάρχει σύνδεσμος προς τη σελίδα της εφαρμογής που αντιστοιχεί στο προφίλ τους. Ακόμα παρατηρούμε ότι το + έχει τώρα μετατραπεί σε – το οποίο βρίσκεται μέσα σε έναν κύκλο με κόκκινο φόντο. Αν πατήσουμε το σύμβολο – ο πίνακας θα επανέλθει στην προηγούμενη του μορφή.

Publications

All PublicationsJournal ArticlesConference ArticlesBook ChaptersBooks

Show10▼ entries

Search: data

Title	Year	Cited by
 H 2RDF: Adaptive query processing on RDF data in the cloud	2012	49
Authors: Tsoumakos D., Konstantinou I., Koziris N., Papailiou N.		
Type: Conference Article		
Conference: 21st Annual Conference on World Wide Web, WWW'12		
Date: 16 April 2012 through 20 April 2012		
Location: Lyon		
Proceedings: WWW'12 - Proceedings of the 21st Annual Conference on World Wide Web Companion		
Pages: 397 - 400		
 Data mining for blood glucose prediction and knowledge discovery in diabetic patients: The METABO diabetes modeling and management system	2009	40

Εικόνα 5.26: Εμφάνιση επιπλέον πληροφοριών με επέκταση της γραμμής

Κατεβαίνοντας προς τα κάτω στη σελίδα υπάρχει ο πίνακας της εικόνας 5.27. Ο πίνακας αυτός περιέχει στην πρώτη στήλη το επώνυμο όλων των μελών ΔΕΠ που είναι καταχωρημένα στη βάση δεδομένων. Στις επόμενες στήλες του πίνακα εμφανίζεται ο αριθμός των άρθρων σε περιοδικά, των άρθρων σε συνεδρία, των κεφαλαίων σε βιβλία και των βιβλίων που κάθε ένα μέλος ΔΕΠ έχει λάβει μέρος στη συγγραφή τους. Στην τελευταία στήλη του πίνακα αθροίζονται οι προηγούμενοι τέσσερις αριθμοί και προκύπτει έτσι ο συνολικός αριθμός των δημοσιεύσεων που έχει κάθε μέλος ΔΕΠ. Επιπλέον και σε αυτόν τον πίνακα υπάρχει η δυνατότητα ταξινόμησης κάθε στήλης σε αύξουσα ή φθίνουσα σειρά, πατώντας στη επικεφαλίδα της. Η λειτουργία αυτή μας επιτρέπει πολύ εύκολα να δούμε ποιο είναι το μέλος ΔΕΠ που έχει τις

περισσότερες δημοσιεύσεις ανά κατηγορία αλλά και συνολικά.

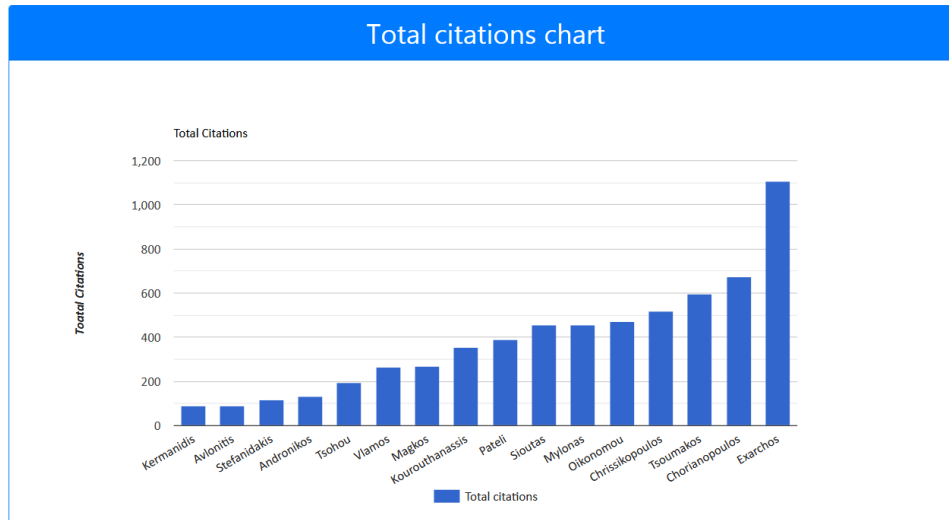
Faculty research					
Professor	Journal Articles	Conference Articles	Book Chapters	Books	Total
Exarchos	54	51	4	1	110
Sioutas	48	59	2	1	110
Chorianopoulos	32	55	1	1	89
Oikonomou	35	46	0	0	81
Mylonas	29	48	2	0	79
Vlamos	37	36	6	0	79
Tsoumakos	11	54	0	0	65
Kermanidis	20	31	3	0	54
Chrissikopoulos	21	24	2	0	47
Magkos	15	25	2	0	42
Andronikos	16	24	0	0	40
Avlonitis	21	15	0	0	36
Kourouthanassis	19	14	2	0	35
Stefanidakis	15	17	2	0	34
Pateli	16	12	2	0	30
Tsohou	14	15	1	0	30

Εικόνα 5.27: Το πλήθος δημοσιεύσεων κάθε μέλους ΔΕΠ

Στην υλοποίηση των πινάκων που παρουσιάζονται εδώ, αλλά και για όλους τους πίνακες που υπάρχουν στην εφαρμογή, χρησιμοποιήθηκε το jQuery plug-in με ονομασία DataTables. Το plug-in αυτό δίνει την δυνατότητα να δημιουργηθούν πίνακες με εντολές HTML που να περιέχουν προηγμένους ελέγχους αλληλεπίδρασης με την χρήση της τεχνολογίας AJAX, όπως ταξινόμηση, αναζήτηση και σελιδοποίηση των αποτελεσμάτων με εύκολο και γρήγορο τρόπο. Πληροφορίες σχετικά με την χρήση του plug-in αυτού μπορούν να ληφθούν από τον ιστότοπο <https://datatables.net>, αλλά και από πολλές άλλες ιστοσελίδες που είναι δυνατόν να προκύψουν από αναζήτηση για το συγκεκριμένο plug-in στο διαδίκτυο.

Από τον πίνακα της εικόνας 5.27 μπορούν να εξαχθούν μόνο ποσοτικά συμπεράσματα για την ερευνητική δραστηριότητα κάθε μέλους ΔΕΠ. Κρίθηκε όμως αναγκαίο να υπάρχουν δεδομένα που να αξιολογούν και ποιοτικά, στο βαθμό που αυτό μπορεί να είναι αντικειμενικό, την ερευνητική δραστηριότητα των μελών ΔΕΠ. Για το λόγο αυτό προχωρώντας προς τα κάτω στην ιστοσελίδα θα συναντήσουμε γραφήματα που αναπαριστούν βιβλιομετρικούς

δείκτες αξιολόγησης ερευνητικής δραστηριότητας. Έτσι στη συνέχεια υπάρχει το γράφημα που απεικονίζει με μορφή ραβδόγραμματος το συνολικό αριθμό αναφορών που έχουν λάβει οι δημοσιεύσεις κάθε μέλους ΔΕΠ. Όπως μπορούμε να παρατηρήσουμε και στην εικόνα 5.28 τα μέλη έχουν διαταχθεί σε αύξουσα σειρά σύμφωνα με το συνολικό αριθμό αναφορών ώστε να είναι ευκολότερη η εξαγωγή συμπερασμάτων.



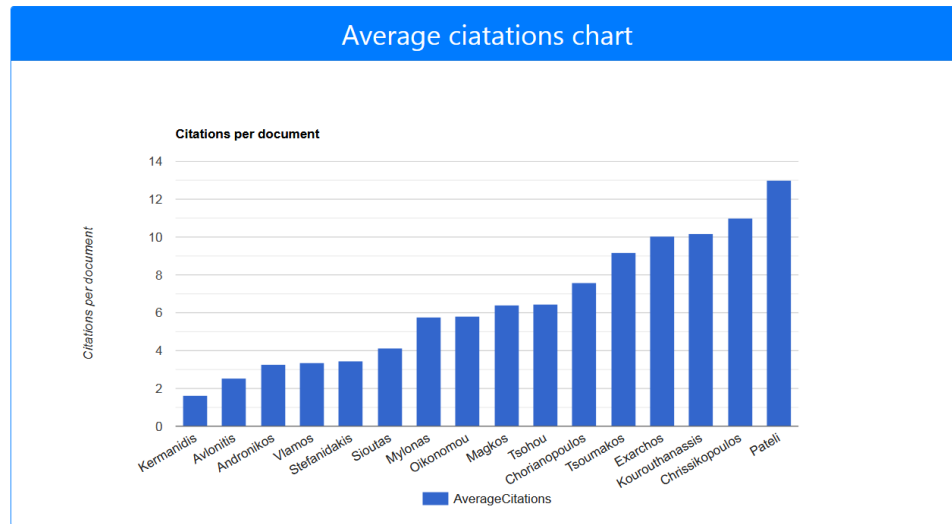
Εικόνα 5.28: Γράφημα συνολικού αριθμού αναφορών ανα μέλος

Ανάλογα με το επιστημονικό πεδίο όσο περισσότερα χρόνια έρευνας ή όσο περισσότερες δημοσιεύσεις έχει ένας ερευνητής είναι πολλές φορές λογικό να διαθέτει και μεγαλύτερο αριθμό συνολικών αναφορών. Για να είναι περισσότερο αντικειμενικά τα ποιοτικά χαρακτηριστικά της ερευνητικής δραστηριότητας υπολογίζεται και παρουσιάζεται σε ραβδόγραμμα, το οποίο παρατίθεται στη συνέχεια της ιστοσελίδας, ο μέσος αριθμός αναφορών που έχει λάβει κάθε μέλος ΔΕΠ. Ο αριθμός αυτός υπολογίζεται σύμφωνα με τον τύπο:

$$\text{Μέσος αριθμός αναφορών} = \frac{\text{Συνολικός αριθμός αναφορών}}{\text{Συνολικός αριθμός δημοσιεύσεων}}$$

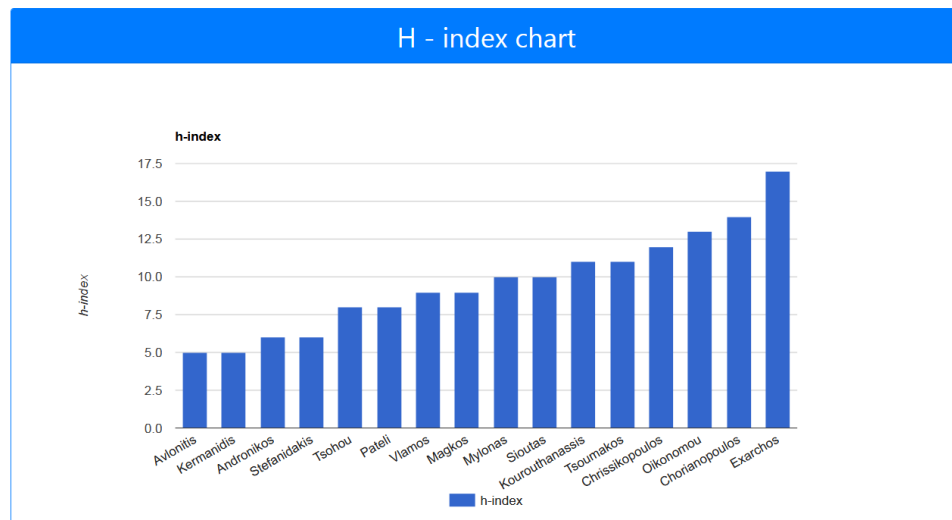
Στην εικόνα 5.29 φαίνεται το ραβδόγραμμα που βρίσκεται στην ιστοσελίδα και παρουσιάζει για κάθε μέλος ΔΕΠ το μέσο αριθμό των αναφορών που έχουν λάβει οι δημοσιεύσεις στην συγγραφή των οποίων έχει πάρει μέρος. Όπως και στο προηγούμενο ραβδόγραμμα, τα μέλη ΔΕΠ είναι διατεταγμένα σε αύξουσα σειρά με βάση το μέσο αριθμό των αναφορών που έχει το καθένα.

Ένας ευρέως διαδεδομένος και διεθνώς αναγνωρισμένος βιβλιομετρικός δείκτης ποιοτικής αξιολόγησης της ερευνητικής δραστηριότητας ενός επιστήμονα είναι ο h – index. Στην διεύθυνση <https://www.programcreek.com/2014/05/leetcode-h-index-java/> υπάρχει σε γλωσσά προγραμματισμού Java μια υλοποίηση του αλγορίθμου υπολογισμού του h – index. Για να είναι συμβατός



Εικόνα 5.29: Γράφημα μέσου αριθμού αναφορών ανα μέλος

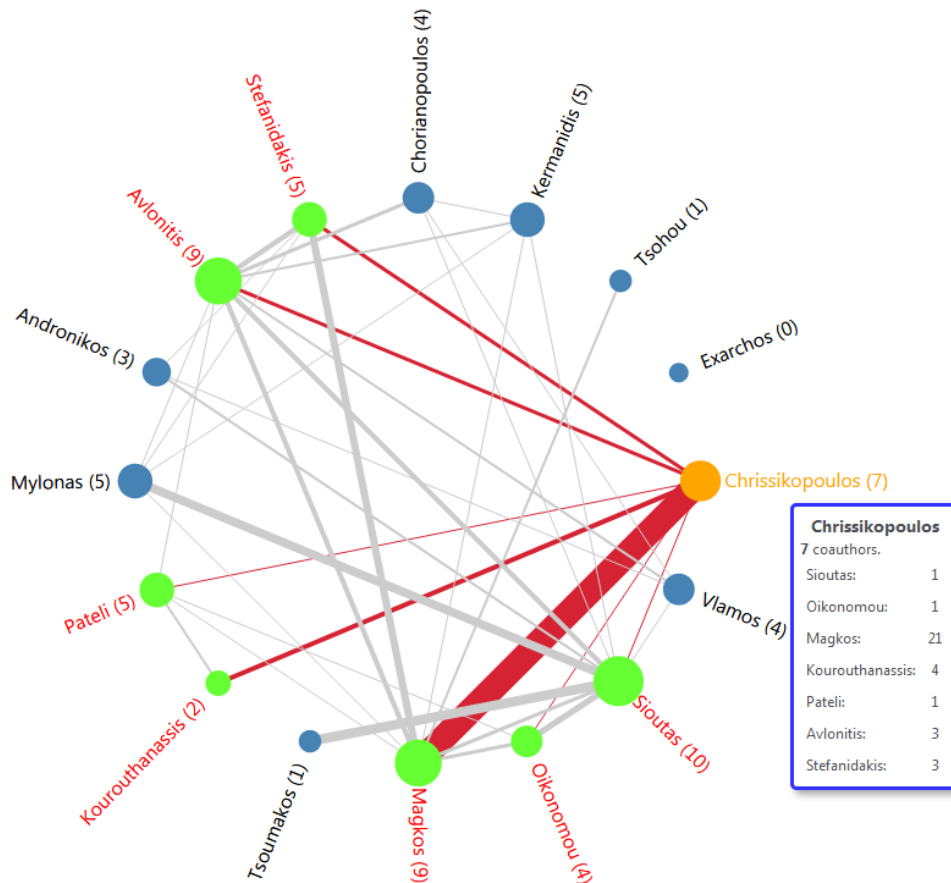
με τις ανάγκες της εφαρμογής ο κώδικας Java μετατράπηκε στις αντίστοιχες εντολές PHP ώστε να υπολογιστεί ο h - index καθενός από τα μέλη ΔΕΠ. Τα αποτελέσματα των υπολογισμών αυτών οπτικοποιήθηκαν ταξινομημένα σε αύξουσα σειρά ως προς τον h - index με την βοήθεια ενός ραβδογράμματος το οποίο φαίνεται συνεχίζοντας την κύλιση προς τα κάτω στην ιστοσελίδα και είναι αυτό που παρουσιάζεται στην εικόνα 5.30.



Εικόνα 5.30: Γράφημα του h -index όλων των μελών

Τα τρία προηγούμενα ραβδογράμματα δημιουργήθηκαν με την javascript βιβλιοθήκη απεικόνισης δεδομένων Google Charts. Πληροφορίες για τη

χρήση της βιβλιοθήκης αυτής υπάρχουν στη διεύθυνση <https://developers.google.com/chart/>.



Εικόνα 5.31: Το δίκτυο των συνεργασιών μεταξύ των μελών

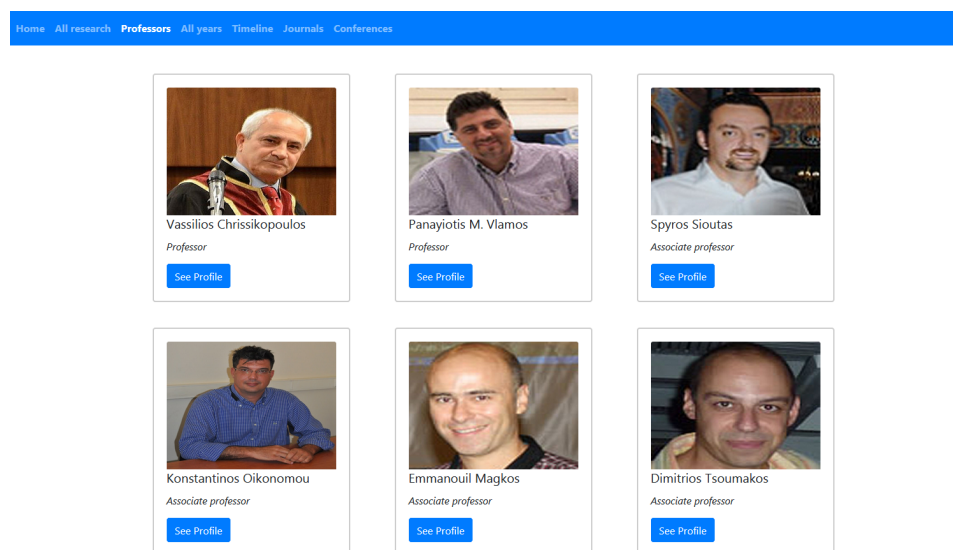
Στο τέλος της σελίδας All Research υπάρχει ένα διαδραστικό γράφημα στο οποίο απεικονίζεται το δίκτυο των συνεργασιών που έχουν λάβει χώρα μεταξύ των μελών ΔΕΠ ως προς τη συγγραφή δημοσιεύσεων. Το γράφημα αυτό παρουσιάζεται στην εικόνα 5.31. Κάθε κόμβος του γραφήματος αντιστοιχεί και σε ένα μέλος ΔΕΠ. Με όσα περισσότερα μέλη ΔΕΠ έχει συνεργαστεί ένα συγκεκριμένο μέλος, τόσο μεγαλύτερο είναι το μέγεθος του κόμβου που του αντιστοιχεί. Οι ακμές που συνδέουν δυο κόμβους εκφράζουν συνεργασία μεταξύ των μελών αυτών. Σε όσο πιο πολλές δημοσιεύσεις έχουν συνεργαστεί δυο μέλη, τόσο μεγαλύτερο είναι το πάχος της ακμής που τα συνδέει. Περνώντας το ποντίκι πάνω από κάθε κόμβο αυτός αλλάζει χρώμα από μπλε σε πορτοκάλι. Οι κόμβοι με τους οποίους συνδέεται γίνονται από μπλε πράσινοι και οι ακμές σύνδεσης λαμβάνουν κόκκινο χρώμα. Επιπλέον εμφανίζεται ένα πλαίσιο που περιέχει πληροφορίες για τον αριθμό των κοινών δημοσιεύσεων που έχει το

συγκεκριμένο μέλος με τα μέλη που συνδέεται. Αν το ποντίκι περάσει πάνω από μια ακμή, τότε αυτή παίρνει κόκκινο χρώμα ενώ οι κόμβοι που συνδέει γίνονται πράσινοι. Επίσης εμφανίζεται ένα πλαίσιο που ενημερώνει για τον αριθμό των κοινών δημοσιεύσεων που έχουν τα μέλη τα όποια συνδέονται με την ακμή αυτή.

Με το γράφημα αυτό είναι δυνατόν να λάβουμε εύκολα και γρήγορα σημαντικές πληροφορίες που αφορούν τις συνεργασίες σχετικά με τη συγγραφή επιστημονικών δημοσιεύσεων μεταξύ των μελών ΔΕΠ του Τμήματος Πληροφορικής του Ιονίου Πανεπιστημίου. Το γράφημα υλοποιήθηκε με την javascript βιβλιοθήκη οπτικοποίησης δεδομένων d3.js (<https://d3js.org/>).

5.2.2 Ιστοσελίδα Professors

Η σελίδα Professors αποτελείται από μία διάταξη που περιέχει σε μορφή κάρτας τη φωτογραφία το ονοματεπώνυμο και τη βαθμίδα κάθε μέλους ΔΕΠ που είναι καταχωρημένο στη βάση δεδομένων. Ένα μέρος της ιστοσελίδας φαίνεται στην εικόνα 5.32.



Εικόνα 5.32: Ένα τμήμα της ιστοσελίδας Professors

Όπως μπορεί κανείς να παρατηρήσει σε κάθε κάρτα υπάρχει το κουμπί See Profile το πάτημα του οποίου οδηγεί στην ιστοσελίδα με το προφίλ του μέλους που απεικονίζεται στην κάρτα. Η ιστοσελίδα αυτή περιέχει πληροφορίες για την ερευνητική δραστηριότητα του συγκεκριμένου μέλους.

Στη συνέχεια παρουσιάζεται η ιστοσελίδα με το ερευνητικό προφίλ του καθηγητή κυρίου Παναγιώτη Βλάμου η οποία εμφανίζεται με το πάτημα του κουμπιού See Profile της κάρτας του.

Στην εικόνα 5.33 φαίνεται το επάνω μέρος της ιστοσελίδας με το ερευνητικό

Home All research **Professors** All years Timeline Journals Conferences



Panayiotis M. Vlamos

Professor

79 Publications

9 H - Index

63 Co-authors

Updated: 06 September 2018

Εικόνα 5.33: Συνοπτικές πληροφορίες ερευνητικού προφίλ

προφίλ του κυρίου Βλάμου, όπου μπορούμε να παρατηρήσουμε ένα πλαίσιο που περιέχει τη φωτογραφία, το ονοματεπώνυμο και τη βαθμίδα του εν λόγω μέλους ΔΕΠ. Επιπλέον εντός του πλαισίου αυτού υπάρχουν πληροφορίες για τον αριθμό των δημοσιεύσεων, τον h – index και το πλήθος των ερευνητών που έχει συνεργαστεί για τη συγγραφή των δημοσιεύσεων. Στο κάτω μέρος του πλαισίου υπάρχει η ημερομηνία τελευταίας ενημέρωσης των δημοσιεύσεων του στη βάση δεδομένων της εφαρμογής.

Publications of Panayiotis M. Vlamos

All Publications **Journal Articles** Conference Articles Book Chapters Books

Show 10 entries Search:

Title	Year	Type	Cited by
IoT Applications with 5G Connectivity in Medical Tourism Sector Management: Third-Party Service Scenarios	2017	Book Chapter	0
Mobile Applications Improve Quality of Life on Citizens with Disorientation: The 'NeverLost App' Paradigm	2017	Book Chapter	0
Molecular chaperones in neurodegenerative diseases: A short review	2017	Book Chapter	0
EEG analysis of the neurofeedback training effect in algorithmic thinking	2017	Journal Article	0
Formal models of biological systems	2017	Journal Article	0
The effectiveness of neurofeedback training in algorithmic thinking skills enhancement	2017	Journal Article	0
Chronic lymphocytic leukemia patient clustering based on somatic hypermutation (SHM) analysis	2017	Journal Article	0
Integrating omic technologies in Alzheimer's disease	2017	Book Chapter	0
Scheduling and modeling a cognitive assessment guide for screening AD by primary care physicians	2017	Book Chapter	0
Associating ω -automata to path queries on Webs of Linked Data	2016	Journal Article	1

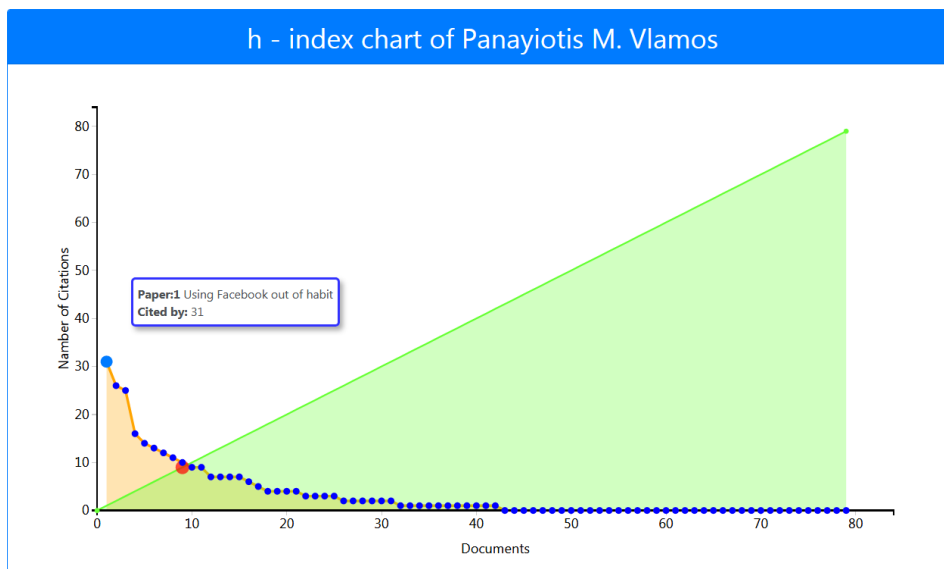
Showing 1 to 10 of 79 entries

Previous 1 2 3 4 5 ... 8 Next

Εικόνα 5.34: Ο πίνακας όλων των δημοσιεύσεων ενός μέλους

Προχωρώντας προς τα κάτω στη ιστοσελίδα θα βρούμε τις καρτέλες με ονομασία All Publications, Journal Articles, Conference Articles, Book Chapters και Books. Σε καθεμία από τις καρτέλες αυτές περιέχονται σελιδοποιημένα σε πίνακα τα δεδομένα των δημοσιεύσεων του κυρίου Βλάμου που αντιστοιχούν στις προηγούμενες κατηγορίες. Στην εικόνα 5.34 παρουσιάζονται τα πρώτα δέκα αποτελέσματα της σελιδοποίησης όλων των δημοσιεύσεων του εν λόγω μέλους ταξινομημένα σε φθίνουσα σειρά ως προς το έτος δημοσίευσης.

Συνεχίζοντας την κύλιση της ιστοσελίδας εμφανίζεται το γράφημα της εικόνας 5.35. Το γράφημα αυτό υπάρχει για να επαληθεύσει το σωστό υπολογισμό από την εφαρμογή του h - index για τον συγκεκριμένο ερευνητή. Οι τιμές στον οριζόντιο άξονα αντιπροσωπεύουν τον αριθμό των δημοσιεύσεων του ερευνητή, ενώ οι τιμές του κατακόρυφου άξονα αντιπροσωπεύουν τον αριθμό των αναφορών που έχει λάβει κάθε δημοσίευση. Οι δημοσιεύσεις ταξινομούνται σε φθίνουσα σειρά ως προς τον αριθμό των αναφορών. Έτσι κάθε δημοσίευση απεικονίζεται στο γράφημα με ένα σημείο. Ενώνοντας τα σημεία αυτά προκύπτει η καμπύλη των δημοσιεύσεων. Ο h - index είναι το σημείο τομής της καμπύλης αυτής με τη διχοτόμο ($y = x$) της γωνιάς των αξόνων. Το γράφημα υλοποιήθηκε με την χρήση της javascript βιβλιοθήκης οπτικοποίησης δεδομένων d3.js.



Εικόνα 5.35: Το γράφημα υπολογισμού του h -index

Στη συνέχεια της ιστοσελίδας περιέχονται καρτέλες στις οποίες εμφανίζονται πίνακες με τα ονόματα των ερευνητών με τους οποίους ο κύριος Βλάμος έχει συνεργαστεί για τη συγγραφή δημοσιεύσεων. Οι καρτέλες φαίνονται στην εικόνα 5.36. Η πρώτη καρτέλα με τίτλο All Co-authors περιέχει πίνακα με όλους τους ερευνητές, στη δεύτερη καρτέλα (Affiliated Co-authors) ο πίνακας περιέχει τα ονόματα όσων ερευνητών είναι μέλη ΔΕΠ του Τμήματος Πληρο-

φορικής και έχουν καταχωρηθεί στη βάση δεδομένων της εφαρμογής, ενώ ο πίνακας της τρίτης καρτέλας (Non Affiliated Co-authors) περιέχει όσους ερευνητές δεν είναι μέλη ΔΕΠ του Τμήματος. Για τους ερευνητές που είναι μέλη ΔΕΠ του τμήματος και έχουν καταχωρηθεί στη βάση δεδομένων, το όνομα τους εμφανίζεται με τη μορφή συνδέσμου ο οποίος οδηγεί στη σελίδα με το δικό τους ερευνητικό προφίλ.

Co-authors of Panayiotis M. Vlamos	
All Co-authors Affiliated Co-authors Non Affiliated Co-authors	
Show 10 entries	Search:
With	Publications
Psiha M.	4
Tzanetis D.E.	3
Bobori C.	3
Sararu C.	3
Floros A.	3
Avlonitis	2
Doukakis S.	2
Polychronidou E.	2
Avramouli A.	2
Alexiou A.T.	2
With	Publications
Showing 11 to 20 of 63 entries	
Previous 1 2 3 4 5 6 7 Next	

Εικόνα 5.36: Ο πίνακας με τους συνεργαζόμενους συγγραφείς

Επιπλέον στη δεύτερη στήλη όλων των πινάκων εμφανίζεται με μορφή συνδέσμου ο αριθμός των δημοσιεύσεων στη συγγραφή των οποίων ο κύριος Βλάμος έχει συνεργαστεί με τον εκάστοτε ερευνητή. Ακολουθώντας το σύνδεσμο αυτό οδηγούμαστε σε μία σελίδα που περιέχει πίνακα με τον τίτλο, το έτος συγγραφής, τον τύπο, και τον αριθμό των αναφορών που οι κοινές δημοσιεύσεις έχουν λάβει. Στην εικόνα 5.37 φαίνεται ο πίνακας που παρουσιάζει τις κοινές δημοσιεύσεις του κύριου Βλάμου με τον κύριο *Αυλωνίτη*.

Προχωρώντας προς τα κάτω την ιστοσελίδα που αντιστοιχεί στο ερευνητικό προφίλ του κυρίου Βλάμου εμφανίζεται ένα χρονοδιάγραμμα (timeline) με το οποίο μπορούμε να πλοηγηθούμε από το πρώτο έως και το τελευταίο έτος κατά το οποίο ο εν λόγω ερευνητής έχει δημοσιεύσεις. Για κάθε έτος εμφανίζεται πίνακας που περιέχει τον τίτλο των δημοσιεύσεων που έγιναν το έτος αυτό. Επίσης στον πίνακα εμφανίζονται ο τύπος κάθε δημοσίευσης και ο αριθμός των αναφορών που αυτή έχει λάβει.

Panayiotis M. Vlamos and Markos Avlonitis			
Show	10	entries	Search:
Title	Year	Type	Cited by
Frequency-domain stochastic error concealment for wireless audio applications	2008	Journal Article	0
Stochastic packet reconstruction for subjectively improved audio delivery over WLANs	2007	Conference Article	1
Title	Year	Type	Cited by
Showing 1 to 2 of 2 entries			
Previous		1	Next

Εικόνα 5.37: Οι κοινές δημοσιεύσεις δυο ερευνητών

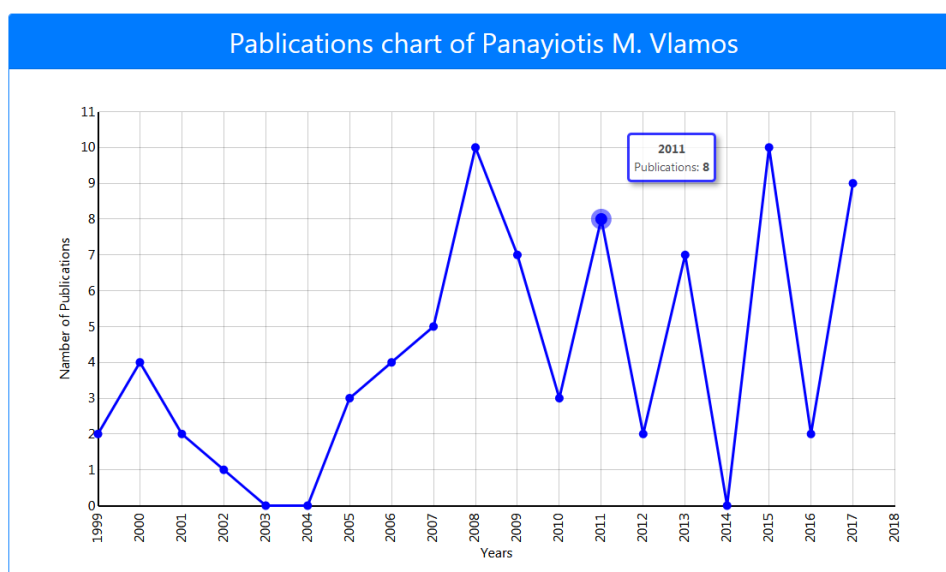
Για την υλοποίηση του χρονοδιαγράμματος αυτού και τη λειτουργία της πλοήγησης από έτος σε έτος χρησιμοποιήθηκε έτοιμος κώδικας γραμμένος σε γλώσσες προγραμματισμού HTML, css και javaScript, με τη χρήση της βιβλιοθήκης της jQuery. Ο κώδικας αυτός λήφθηκε από τη διαδικτυακή διεύθυνση <https://codyhouse.co/gem/horizontal-timeline> και τροποποιήθηκε ώστε οι λειτουργίες που επιτελεί να είναι συμβατές με τις ανάγκες και την εμφάνιση της εφαρμογής. Στην εικόνα 5.38 φαίνονται οι δημοσιεύσεις του κυρίου Βλάμου κατά το έτος 2016, έτσι όπως παρουσιάζονται μετά από την πλοήγηση του χρονοδιαγράμματος στο έτος αυτό.

Research Timeline of Panayiotis M. Vlamos			
2005 2006 2007 2008 2009 2010 2011 2012 2013 2015 2016 2017			
2016			
Show	10	entries	Search:
Title	Year	Type	Cited by
Associating ω -automata to path queries on Webs of Linked Data	2016	Journal Article	1
Molecular basis of Huntington's disease and brain imaging evidence	2016	Conference Article	2
Title	Year	Type	Cited by
Showing 1 to 2 of 2 entries			
Previous		1	Next

Εικόνα 5.38: Δημοσιεύσεις ενός συγκεκριμένου έτους

Στο τέλος της ιστοσελίδας του ερευνητικού προφίλ του κυρίου Βλάμου υπάρχει ένα χρονολογικό διάγραμμα στο οποίο απεικονίζεται η διαχρονική εξέλιξη της πορείας του αριθμού των δημοσιεύσεων του εν λόγω ερευνητή ανά έτος. Το διάγραμμα αυτό υλοποιήθηκε με την javaScript βιβλιοθήκη οπτικοποίησης δεδομένων d3.js και παρουσιάζεται στην εικόνα 5.39. Οι τιμές του

οριζόντιου άξονα αναπαριστούν τα έτη από το πρώτο μέχρι και το τελευταίο που υπάρχουν δημοσιεύσεις του κύριου Βλάμου καταχωρημένες στη βάση δεδομένων. Οι τιμές του κατακόρυφου άξονα αναπαριστούν τον αριθμό των δημοσιεύσεων σε κάθε έτος. Από την ένωση των σημείων που προκύπτουν δημιουργείται μια τεθλασμένη γραμμή η οποία δίνει άμεσες πληροφορίες για τη χρονολογική εξέλιξη του αριθμού των δημοσιεύσεων αλλά και σε ποια έτη έχει παραχθεί περισσότερη έρευνα.



Εικόνα 5.39: Γράφημα αριθμού δημοσιεύσεων ενός μέλους ανά έτος

5.2.3 Ιστοσελίδα All years

Η παρουσίαση του τμήματος της εφαρμογής που αφορά τους χρήστες του συστήματος συνεχίζεται με την ιστοσελίδα All years. Η ιστοσελίδα αυτή έχει σκοπό να παρουσιάσει για κάθε μέλος ΔΕΠ το πλήθος των δημοσιεύσεων που έχουν γίνει σε όλα τα χρόνια της έρευνας του χωρίς να είναι απαραίτητο κάποιος να ανατρέξει στη σελίδα με το ερευνητικό προφίλ κάθε μέλους.

Οι πίνακες που περιέχονται στη σελίδα αυτή αποτελούνται από έξι στήλες. Στην πρώτη στήλη αναγράφονται όλα τα έτη έρευνας του μέλους ΔΕΠ στο οποίο αντιστοιχεί ο πίνακας. Στη δεύτερη στήλη υπάρχει ο αριθμός των άρθρων σε περιοδικά που έχουν δημοσιευτεί σε κάθε έτος από το συγκεκριμένο μέλος ΔΕΠ. Η τρίτη στήλη του πίνακα περιέχει το πλήθος των άρθρων σε συνέδρια ενώ η τέταρτη και η πέμπτη τον αριθμό των κεφαλαίων σε βιβλία και τον αριθμό των βιβλίων αντίστοιχα που το συγκεκριμένο μέλος ΔΕΠ έχει δημοσιεύσει σε κάθε έτος. Σε κάθε κελί της τελευταίας στήλης του πίνακα αθροίζονται οι τέσσερις προηγούμενοι αριθμοί και προκύπτει έτσι ο συνολικός αριθμός των δημοσιεύσεων του μέλους για κάθε έτος.

Σε όλες τις στήλες υπάρχει η δυνατότητα ταξινόμησης των δεδομένων σε αύξουσα ή φθίνουσα σειρά. Αυτό γίνεται με το πάτημα στην επικεφαλίδα του πίνακα που αντιστοιχεί στη στήλη που θέλουμε να ταξινομηθεί.

Year	Journal Papers	Conference Papers	Book Chapters	Books	Total
2017	3	2	0	0	5
2016	0	6	0	0	6
2015	3	1	0	0	4
2014	4	1	0	0	5
2013	2	0	0	0	2
2012	0	2	0	0	2
2009	1	2	0	0	3
2008	2	0	0	0	2
2007	2	1	0	0	3
2006	1	0	0	0	1
2005	1	0	0	0	1
2000	1	0	0	0	1
1998	1	0	0	0	1
TOTAL	21	15	0	0	36

Εικόνα 5.40: Ένα τμήμα της ιστοσελίδας *All years*

Η λειτουργία της ταξινόμησης βοηθάει αποτελεσματικά στο να εξάγουμε άμεσα συμπεράσματα σε ποιο έτος ένα μέλος ΔΕΠ είχε τις περισσότερες δημοσιεύσεις ανά κατηγορία άλλα και συνολικά. Με τον τρόπο αυτό μπορούμε εύκολα να διαπιστώσουμε ποιο ήταν το έτος εκείνο κατά το οποίο ένα μέλος παρήγαγε περισσότερη έρευνα.

Στην εικόνα 5.40 φαίνεται ένα μέρος της ιστοσελίδας *All years* με τον πίνακα που αντιστοιχεί στις δημοσιεύσεις του επίκουρου καθηγητή κυρίου Μάρκου Αυλωνίτη. Προχωρώντας προς τα κάτω ή προς τα πάνω την ιστοσελίδα εμφανίζονται οι πίνακες που αντιστοιχούν στα υπόλοιπα μέλη ΔΕΠ που είναι καταχωρημένα στη βάση δεδομένων. Αν θέλουμε να μεταβούμε κατευθείαν στον πίνακα ενός συγκεκριμένου μέλους, μπορούμε να το πραγματοποιήσουμε πατώντας το σύνδεσμο με το επώνυμο του που βρίσκεται στο αριστερό πλαίσιο μενού της σελίδας.

5.2.4 Ιστοσελίδα *Timeline*

Στην ιστοσελίδα *Timeline* υπάρχει ένα χρονοδιάγραμμα αντίστοιχο με αυτό της σελίδας με το ερευνητικό προφίλ των μελών ΔΕΠ, μόνο που τώρα το χρονοδιάγραμμα αυτό αφορά όλη την έρευνα του Τμήματος. Έτσι μπορούμε να πλοηγηθούμε από το πρώτο έως και το τελευταίο έτος στο οποίο όλα τα μέλη ΔΕΠ που είναι καταχωρημένα στη βάση δεδομένων έχουν κάποια δημοσίευση.

Κατά την πλοήγηση του χρονοδιαγράμματος από χρόνο σε χρόνο εμφανίζετε για κάθε έτος ένας πίνακας ο οποίος περιέχει στην πρώτη στήλη το επώνυμο όλων των μελών ΔΕΠ που είναι καταχωρημένα στη βάση δεδομένων.

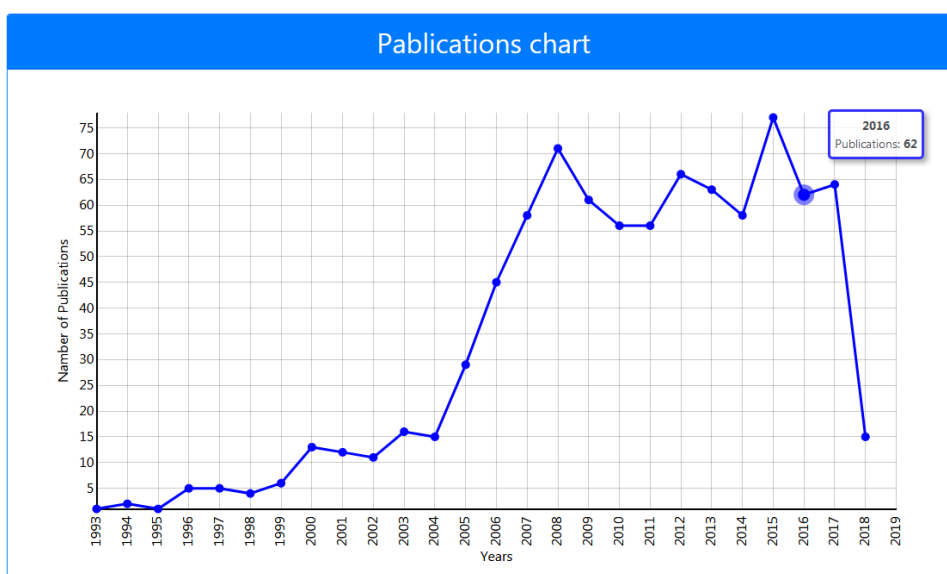
Στις επόμενες στήλες του πίνακα εμφανίζεται ο αριθμός των άρθρων σε περιοδικά, των άρθρων σε συνεδρία, των κεφαλαίων σε βιβλία και των βιβλίων που κάθε ένα μέλος ΔΕΠ έχει λάβει μέρος στη συγγραφή τους το συγκεκριμένο έτος. Στην τελευταία στήλη του πίνακα αθροίζονται οι προηγούμενοι τέσσερεις αριθμοί και προκύπτει έτσι ο συνολικός αριθμός των δημοσιεύσεων που έχει κάθε μέλος ΔΕΠ στο έτος αυτό. Επιπλέον, δίνεται η δυνατότητα ταξινόμησης κάθε στήλης σε αύξουσα ή φθίνουσα σειρά, πατώντας στην επικεφαλίδα της. Η λειτουργία αυτή μας επιτρέπει πολύ εύκολα να δούμε ποιο είναι το μέλος ΔΕΠ που έχει τις περισσότερες δημοσιεύσεις ανά κατηγορία αλλά και συνολικά για το συγκεκριμένο έτος. Στην εικόνα 5.41 φαίνεται ο πίνακας για το έτος 2017.



Εικόνα 5.41: Ο πίνακας της ιστοσελίδας *Timeline* για το έτος 2017

Αν προχωρήσουμε προς τα κάτω στην ιστοσελίδα παρουσιάζεται ένα χρονολογικό διάγραμμα αντίστοιχο με αυτό της σελίδας με το ερευνητικό προφίλ των μελών ΔΕΠ, μόνο που τώρα το διάγραμμα αυτό απεικονίζει τη διαχρο-

νική εξέλιξη της πορείας του αριθμού των δημοσιεύσεων συνολικά από όλα τα μέλη ΔΕΠ του Τμήματος που είναι καταχωρημένα στη βάση δεδομένων. Το διάγραμμα αυτό υλοποιήθηκε με την javascript βιβλιοθήκη οπτικοποίησης δεδομένων d3.js και παρουσιάζεται στην εικόνα 5.42. Οι τιμές του οριζόντιου άξονα αναπαριστούν τα έτη από το πρώτο μέχρι και το τελευταίο που υπάρχουν δημοσιεύσεις. Οι τιμές του κατακόρυφου άξονα αναπαριστούν τον αριθμό των δημοσιεύσεων σε κάθε έτος. Από την ένωση των σημείων που προκύπτουν δημιουργείται μια τεθλασμένη γραμμή η οποία δίνει άμεσες πληροφορίες για τη χρονολογική εξέλιξη του αριθμού των δημοσιεύσεων αλλά και σε ποια έτη έχει παραχθεί περισσότερη έρευνα.



Εικόνα 5.42: Γράφημα πλήθους δημοσιεύσεων ανά έτος

5.2.5 Ιστοσελίδα Journals

Στην ιστοσελίδα journals περιέχονται σελιδοποιημένα σε πίνακες όλα τα επιστημονικά περιοδικά στα οποία έχουν γίνει δημοσιεύσεις από όσα μέλη ΔΕΠ είναι καταχωρημένα στη βάση δεδομένων. Στην πρώτη στήλη του πίνακα αναγράφεται το όνομα του περιοδικού, ενώ στη δεύτερη στήλη υπάρχει σε μορφή συνδέσμου ο αριθμός των άρθρων της βάσης δεδομένων που είναι δημοσιευμένα στο περιοδικό. Στον πίνακα υπάρχει δυνατότητα ταξινόμησης των δεδομένων κάθε στήλης σε αύξουσα ή φθίνουσα σειρά με το πάτημα στην επικεφαλίδα της. Ο πίνακας αυτός παρουσιάζεται στην εικόνα 5.43. Επιπλέον η σελίδα παρέχει λειτουργία αναζήτησης με πληκτρολόγηση φράσης σε πεδίο κειμένου.

Ακολουθώντας το σύνδεσμο με τον αριθμό των άρθρων που έχουν δημοσιευτεί σε ένα περιοδικό οδηγούμαστε στην ιστοσελίδα που περιέχει πληροφορίες για τα άρθρα που υπάρχουν στην βάση δεδομένων και έχουν δημοσιευτεί

Journals	
Show 10 entries	Search:
Journal name ↑↓	Articles ↑↓
Advances in Experimental Medicine and Biology	20
Multimedia Tools and Applications	9
Engineering Applications of Artificial Intelligence	7
Computer Communications	7
Computers in Biology and Medicine	7
Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference	5
Neurocomputing	5
IFIP Advances in Information and Communication Technology	5
Journal of Universal Computer Science	4
Journal of Parallel and Distributed Computing	4
Journal name	Articles
Showing 1 to 10 of 222 entries	
Previous 1 2 3 4 5 ... 23 Next	

Εικόνα 5.43: Ο πίνακας της ιστοσείδας Journals

στο συγκεκριμένο περιοδικό. Οι πληροφορίες αυτές παρουσιάζονται σελιδοποιημένες σε πίνακες όπου στην πρώτη στήλη τους αναγράφεται ο τίτλος του άρθρου, στη δεύτερη το έτος δημοσίευσης και στην τρίτη ο αριθμός των αναφορών που το άρθρο αυτό έχει λάβει. Κάθε στήλη έχει τη δυνατότητα ταξινόμησης των δεδομένων της σε αύξουσα ή φθίνουσα σειρά με πάτημα στην επικεφαλίδα της. Επιπλέον και σε αυτήν τη σελίδα υπάρχει η δυνατότητα της αναζήτησης των δεδομένων. Στην εικόνα 5.44 φαίνεται η σελίδα που περιέχει τους τίτλους όλων των άρθρων που έχουν δημοσιευτεί στο περιοδικό με ονομασία «Journal of Parallel and Distributed Computing».

Πατώντας το σύμβολο + που βρίσκεται μέσα σε έναν κύκλο με πράσινο φόντο στην πρώτη στήλη του πίνακα, μπορούμε να δούμε επιπλέον πληροφορίες για τη συγκεκριμένη δημοσίευση όπως, για παράδειγμα, τους συγγραφείς της, τον τύπο της, τον τόμο και το τεύχος του περιοδικού που είναι δημοσιευμένη, καθώς και τις σελίδες στις οποίες βρίσκεται. Για όσους από τους συγγραφείς είναι μέλη ΔΕΠ και έχουν καταχωρηθεί στη βάση δεδομένων, υπάρχει σύνδεσμος προς τη σελίδα της εφαρμογής που αντιστοιχεί στο προφίλ τους. Στην εικόνα 5.45 βλέπουμε τη μορφή που παίρνει ο πίνακας αμέσως μετά το πάτημα του συμβόλου +. Ακόμα παρατηρούμε ότι το + έχει τώρα μετατραπεί σε – το οποίο βρίσκεται μέσα σε έναν κύκλο με κόκκινο φόντο. Αν πατήσουμε το σύμβολο – ο πίνακας θα επανέλθει στην προηγούμενη του μορφή δηλαδή αυτή με την οποία εμφανίζεται στην εικόνα 5.44.

Journal

• **Name:** Journal of Parallel and Distributed Computing

Articles			
Show 10 entries		Search:	
	Title	Year	Cited by
✓	Optimal Scheduling for UET/UET-UCT Generalizedn-Dimensional Grid Task Graphs	1999	13
✓	Enhancing self-scheduling algorithms via synchronization and weighting	2008	10
✓	Brown Dwarf: A fully-distributed, fault-tolerant data warehousing system	2011	4
✓	Online querying of d-dimensional hierarchies	2011	4
	Title	Year	Cited by
Showing 1 to 4 of 4 entries			
Previous		1	Next

Εικόνα 5.44: Τα άρθρα που έχουν δημοσιευτεί σε ένα περιοδικό

Articles			
Show 10 entries		Search:	
	Title	Year	Cited by
✓	Optimal Scheduling for UET/UET-UCT Generalizedn-Dimensional Grid Task Graphs	1999	13
✓	Enhancing self-scheduling algorithms via synchronization and weighting	2008	10
✓	Brown Dwarf: A fully-distributed, fault-tolerant data warehousing system	2011	4
✗	Online querying of d-dimensional hierarchies	2011	4
Authors: Tsoumakos D., Koziris N., Doka K. Type: Journal Article Journal: Journal of Parallel and Distributed Computing Volume: 71 Issue: 3 Pages: 424 - 437			
	Title	Year	Cited by
Showing 1 to 4 of 4 entries			
Previous		1	Next

Εικόνα 5.45: Επιπλέον πληροφορίες με την επέκταση της γραμμής

5.2.6 Ιστοσελίδα Conferences

Στην ιστοσελίδα Conferences περιέχονται σελιδοποιημένα σε πίνακες όλα τα συνέδρια στα οποία έχουν παρουσιαστεί δημοσιεύσεις από όσα μέλη ΔΕΠ είναι καταχωρημένα στη βάση δεδομένων. Τα δεδομένα αυτά παρουσιάζονται με παρόμοιο ακριβώς τρόπο όπως αυτόν της σελίδας Journals. Στην πρώτη στήλη του πίνακα αναγράφεται το όνομα του συνεδρίου, ενώ στη δεύτερη στήλη υπάρχει σε μορφή συνδέσμου ο αριθμός των άρθρων που έχουν παρουσιαστεί στο συνέδριο αυτό και δημοσιευτεί στα πρακτικά του.

Ακολουθώντας, και εδώ, το σύνδεσμο με τον αριθμό των άρθρων που έχουν παρουσιαστεί σε ένα συνέδριο οδηγούμαστε στην ιστοσελίδα που περιέχει πληροφορίες για τα άρθρα που υπάρχουν στη βάση δεδομένων και έχουν παρουσιαστεί στο συγκεκριμένο συνέδριο. Οι πληροφορίες αυτές παρουσιάζονται σελιδοποιημένες σε πίνακες όπου στην πρώτη στήλη τους αναγράφεται ο τίτλος του άρθρου, στη δεύτερη το έτος δημοσίευσης και στην τρίτη ο αριθμός των αναφορών που το άρθρο αυτό έχει λάβει. Κάθε στήλη έχει τη δυνατότητα ταξινόμησης των δεδομένων της σε αύξουσα ή φθίνουσα σειρά με πάτημα στην επικεφαλίδα της.

Κεφάλαιο 6

Μελλοντικές επεκτάσεις και σύνοψη

6.1 Μελλοντικές επεκτάσεις

Η εφαρμογή που αναπτύχθηκε στα πλαίσια της εκπόνησης πτυχιακής εργασίας για το Τμήμα Πληροφορική του Ιόνιου Πανεπιστήμιου, αποτελεί ένα ανθρώπινο δημιούργημα και συνεπώς δεν είναι τέλεια ούτε απαλλαγμένη από λάθη και παραλήψεις. Εξάλλου η υλοποίηση ενός ολοκληρωμένου συστήματος παρουσίασης και αξιολόγησης της ερευνητικής δραστηριότητας του Τμήματος Πληροφορικής, έχει απαιτήσεις που ξεφεύγουν από τα πλαίσια μιας πτυχιακής εργασίας προπτυχιακού επιπέδου.

Στη συνέχεια αναφέρονται κάποιες προτάσεις οι οποίες θα μπορούσαν μελλοντικά να υλοποιηθούν, ώστε να βελτιώσουν την εφαρμογή έχοντας ως στόχο να την αναβαθμίσουν και να την καταστήσουν περισσότερο λειτουργική.

Ένα στοιχείο το οποίο θα μπορούσε να προσδώσει στην εφαρμογή πολύ μεγαλύτερη λειτουργικότητα είναι ταυτόχρονα με την καταχώριση ενός νέου μέλους ΔΕΠ να γίνεται αυτόματα και η λήψη του αρχείου csv, από τον διακομιστή του Scopus, το οποίο περιέχει τις δημοσιεύσεις του συγκεκριμένου μέλους ΔΕΠ. Χωρίς δηλαδή να είναι απαραίτητη η μεσολάβηση του διαχειριστή για να ληφθεί το αρχείο και να αποθηκευτεί στο διακομιστή της εφαρμογής.

Επίσης πολύ σημαντικό θα ήταν να βρεθεί τρόπος ώστε η εφαρμογή έκτος από το Scopus, να συλλέγει δεδομένα και από άλλες βάσεις δεδομένων επιστημονικών δημοσιεύσεων, όπως το Google Scholar ή το Web of Science. Έτσι θα υπάρχει μια πληρέστερη εικόνα για την ερευνητική δραστηριότητα ενός μέλους ΔΕΠ αφού όσες δημοσιεύσεις του δεν διατίθενται στο Scopus θα μπορούν να συλλεχτούν από άλλες βάσεις. Επιπλέον θα είναι δυνατόν να γίνει και σύγκριση των δεδομένων που διατίθενται από κάθε βιβλιογραφική βάση.

Ένα ακόμα εξίσου σημαντικό στοιχείο θα ήταν να μπορεί κάθε μέλος ΔΕΠ να παρέμβει στο ερευνητικό του προφίλ και να διορθώσει τυχόν λανθασμένα δεδομένα. Για παράδειγμα να είναι δυνατόν να διαγράψει δημοσιεύσεις οι οποίες

δεν του αντιστοιχούν και του έχουν λανθασμένα αποδοθεί. Επιπλέον καλό θα ήταν να έχει την δυνατότητα να εισάγει δεδομένα δημοσιεύσεων του, που δεν υπάρχουν στη βάση της εφαρμογής αλλά και να επικαιροποιήσει τον αριθμό των αναφορών που οι δημοσιεύσεις του έχουν λάβει, στην περίπτωση κατά την οποία αυτός έχει αλλάξει.

Μια ακόμη λειτουργία που θα μπορούσε μελλοντικά να παρέχει η εφαρμογή θα ήταν να δίνει την δυνατότητα σε κάθε μέλος ΔΕΠ να εξάγει την λίστα των δημοσιεύσεων του σε διάφορες μορφές αρχείων μεταξύ των οποίων και σε αρχεία BibTeX, ώστε να εξυπηρετούν στη συγγραφή βιογραφικού σημειώματος.

Ιδιαίτερα σημαντικό ρόλο στην αποτελεσματικότερη παρουσίαση των δεδομένων από την εφαρμογή θα έχει η ύπαρξη μιας ιστοσελίδας σύγκρισης, μέσα από πίνακες και διαδραστικά γραφήματα, της ερευνητικής δραστηριότητας μεταξύ δυο μελών ΔΕΠ. Αυτή είναι μια ακόμα πρόταση μελλοντικής επέκτασης της εφαρμογής.

Όσον αφορά τις συνεργασίες, μια πρόταση είναι να βρεθεί ένα τρόπος αξιόπιστης αντιστοίχισης όλων των ερευνητών που είναι καταχωρημένοι στη βάση δεδομένων της εφαρμογής με τα ερευνητικά ιδρύματα και τους οργανισμούς στους οποίους ανήκουν. Με την ενέργεια αυτή θα είναι δυνατόν να εξαχθούν συμπεράσματα για τις συνεργασίες που αναπτύσσονται ανάμεσα στο Τμήμα Πληροφορικής και σε άλλα Ιδρύματα, δείχνοντας παράλληλα τον βαθμό της εξωστρέφειας του κάθε μέλους άλλα και του Τμήματος γενικότερα.

6.2 Σύνοψη και τελικός απολογισμός

Η διαδικτυακή εφαρμογή που υλοποιήθηκε για τις ανάγκες της πτυχιακής εργασίας πιστεύουμε ότι καλύπτει ένα μέρος, του κενού το οποίο υπάρχει σχετικά με τέτοιου είδους συστήματα στο Ιόνιο Πανεπιστήμιο, αλλά και ότι θα αποτελέσει την απαρχή της δημιουργίας, στο μέλλον, ενός ολοκληρωμένου συστήματος παρουσίασης και διαχείρισης της ερευνητικής δραστηριότητας ολόκληρου του Ιδρύματος.

Θα ήταν προφανώς αυταπάτη ο ισχυρισμός ότι δημιουργήθηκε ένα λογισμικό απαλλαγμένο από προγραμματιστικά λάθη που να προσφέρει την καλύτερη δυνατή εμπειρία στους χρήστες του και μάλιστα στην πρώτη σοβαρή προσπάθεια δημιουργίας λογισμικού τέτοιου μεγέθους. Υπάρχουν σίγουρα δυσλειτουργίες, λάθη που δεν ανακαλύφθηκαν ακόμα, και πολλά περιθώρια για βελτίωση γι αυτό και είναι επιθυμητή οποιαδήποτε εποικοδομητική κριτική καθώς και προτάσεις για προσθήκες και βελτιώσεις.

Τελικά όμως ο πραγματικός στόχος του συγκεκριμένου εγχειρήματος επιτεύχθηκε καθώς αποκτήθηκε σημαντική προγραμματιστική εμπειρία πάνω στο διαδικτυακό προγραμματισμό και στις τεχνολογίες που χρησιμοποιούνται για τη δημιουργία διαδικτυακών εφαρμογών. Επιπλέον η εφαρμογή αυτή είναι πολύ πιθανό να χρησιμοποιηθεί για το σκοπό που δημιουργήθηκε, ξεφεύγοντας έτσι από το να χαρακτηριστεί η δημιουργία της ως απλά μία αναγκαία διαδικασία

για την απόκτηση ενός πτυχίου τριτοβάθμιας εκπαίδευσης.

Θα επιφέρει μεγάλη προσωπική ικανοποίηση και ιδιαίτερη χαρά το γεγονός κάποιος να θελήσει κάποτε να στηριχτεί στην ιδέα που παρουσιάστηκε εδώ με σκοπό την αναβάθμιση της εφαρμογής και τη δημιουργία ενός ολοκληρωμένου συστήματος παρουσίασης της ερευνητικής δραστηριότητας τμημάτων και οργανισμών.

Τέλος ο δημιουργός της εφαρμογής θεωρεί ότι θα αποτελέσει μεγάλη τιμή για αυτόν να διοχετεύσει την εμπειρία και την τεχνογνωσία που αποκόμισε κατά τη διάρκεια της υλοποίησης της και να παρέχει οποιαδήποτε βοήθεια σε κάποιον που θα θελήσει στο μέλλον να την επεκτείνει. Για οποιαδήποτε πληροφορία σχετικά με αυτό μπορεί όποιος επιθυμεί, να επικοινωνήσει μαζί του μέσω ηλεκτρονικού ταχυδρομείου στις διευθύνσεις p09kont@ionio.gr και p09kont@gmail.com.

Παράρτημα Α΄

Το αρχείο Database.php

```
1 <?php
2
3 function connectToDB() {
4     $host = "localhost";
5     $username = "root";
6     $password = "";
7     $database = "research_db_test";
8     $connection = mysqli_connect($host, $username, $password, $database);
9     if (mysqli_connect_errno()) {
10         echo "Failure in connection to Database" . mysqli_connect_error() . "<br/>";
11     }
12     mysqli_set_charset($connection, "utf8");
13     return $connection;
14 }
15
16 function publicationHasInserted($link, $title) {
17     $theTitle = mysqli_real_escape_string($link, trim($title));
18     $select = "SELECT id FROM publications WHERE title = '$theTitle'";
19     $result = mysqli_query($link, $select);
20
21     if(mysqli_num_rows($result) > 0){
22         return true;
23     }
24     else{
25         return false;
26     }
27 }
28
29 function selectIdFromPublications($link, $title) {
30     $theTitle = mysqli_real_escape_string($link, trim($title));
31     $select = "SELECT id FROM publications WHERE title = '$theTitle'";
32     $result = mysqli_query($link, $select);
33     return mysqli_num_rows($result);
34 }
35
36 function insertIntoPersons($link, $listOfAuthors) {
37     foreach ($listOfAuthors as $a) {
38         $trimAuthor = trim($a);
39         $author = mysqli_real_escape_string($link, $trimAuthor);
40         $select = "SELECT id FROM persons WHERE name = '$author'";
41         $result = mysqli_query($link, $select);
42         if (mysqli_num_rows($result) == 0) {
43             $ionianId = selectProfId($link, $author);
```

```

44     $prof_id = !is_null($ionianId) ? $ionianId : "NULL";
45     $insertPerson = "INSERT INTO persons (name, prof_id) VALUES ('$author' ,
        ↪ $prof_id)";
46     $r = mysqli_query($link, $insertPerson);
47     if (!$r) {
48         echo mysqli_error($link);
49     }
50 }
51 }
52 }
53
54 function selectProfId($link, $name) {
55     $select = "SELECT prof_id FROM prof_names WHERE name = '$name' ";
56     $result = mysqli_query($link, $select);
57     if (mysqli_num_rows($result) == 0) {
58         return null;
59     } else {
60         $row = mysqli_fetch_assoc($result);
61         $id = $row["prof_id"];
62         return $id;
63     }
64 }
65
66 function insertIntoPublications($link, $title, $year, $pageStart, $pageEnd, $citedBy
        ↪ ) {
67     $theTitle = mysqli_real_escape_string($link, trim($title));
68     $theYear = mysqli_real_escape_string($link, trim($year));
69     $firstPage = mysqli_real_escape_string($link, trim($pageStart));
70     $lastPage = mysqli_real_escape_string($link, trim($pageEnd));
71     $cited = mysqli_real_escape_string($link, trim($citedBy));
72     $citations = empty($cited) ? 0 : $cited;
73     $page_start = empty($firstPage) ? "NULL" : "'$firstPage'";
74     $page_end = empty($lastPage) ? "NULL" : "'$lastPage'";
75     $insertPublications = "INSERT INTO publications (title, year, page_start,
        ↪ page_end, cited_by)"
76         . "VALUES ('$theTitle', '$theYear', $page_start, $page_end, $citations)";
77     if (mysqli_query($link, $insertPublications)) {
78         $last_id = mysqli_insert_id($link);
79         return $last_id;
80     } else {
81         echo mysqli_error($link);
82         return null;
83     }
84 }
85
86 function insertIntoBooks($link, $id, $isbn, $publisher) {
87     $theIsbn = mysqli_real_escape_string($link, trim($isbn));
88     $thePublisher = mysqli_real_escape_string($link, trim($publisher));
89     $ISBN = empty($theIsbn) ? "NULL" : "'$theIsbn'";
90     $insertBook = "INSERT INTO books (pub_id, isbn, publisher) "
91         . "VALUES ($id, $ISBN, '$thePublisher')";
92     $result = mysqli_query($link, $insertBook);
93     if (!$result) {
94         echo mysqli_error($link);
95     }
96 }
97
98 function insertIntoAuthorships($link, $pub_id, $listOfAuthors) {
99     foreach ($listOfAuthors as $a) {
100         $trimAuthor = trim($a);
101         $author = mysqli_real_escape_string($link, $trimAuthor);
102         $select = "SELECT id FROM persons WHERE name = '$author'";

```

```

103     $result = mysqli_query($link, $select);
104     $row = mysqli_fetch_assoc($result);
105     $pers_id = $row["id"];
106     $insertAuthorship = "INSERT INTO authorships (pub_id, pers_id) "
107         . "VALUES($pub_id, $pers_id)";
108     $r = mysqli_query($link, $insertAuthorship);
109     if (!$r) {
110         echo mysqli_error($link);
111     }
112 }
113 }
114
115 function insertIntoBookChapters($link, $pub_id, $bookTitle, $bookIsbn, $bookPublisher
    ↪ ) {
116     $theBookTitle = mysqli_real_escape_string($link, trim($bookTitle));
117     $theBookIsbn = mysqli_real_escape_string($link, trim($bookIsbn));
118     $theBookPublisher = mysqli_real_escape_string($link, trim($bookPublisher));
119     $bookISBN = empty($theBookIsbn) ? "NULL" : "'$theBookIsbn'";
120     $insertBook = "INSERT INTO book_chapters (pub_id, booktitle, bookisbn,
    ↪ bookpublisher) "
121         . "VALUES ($pub_id, '$theBookTitle', $bookISBN , '$theBookPublisher')";
122     $result = mysqli_query($link, $insertBook);
123     if (!$result) {
124         echo mysqli_error($link);
125     }
126 }
127
128 function insertIntoJournalArticles($link, $pub_id, $volume, $issue, $journalName) {
129     $theVolume = mysqli_real_escape_string($link, trim($volume));
130     $theIssue = mysqli_real_escape_string($link, trim($issue));
131     $articleVolume = empty($theVolume) ? "NULL" : "'$theVolume'";
132     $articleIssue = empty($theIssue) ? "NULL" : "'$theIssue'";
133     $select = "SELECT id FROM journals WHERE name = '$journalName'";
134     $result = mysqli_query($link, $select);
135     $row = mysqli_fetch_assoc($result);
136     $journal_id = $row["id"];
137     $insertJournalArticle = "INSERT INTO journal_articles (pub_id, volume, issue,
    ↪ journal_id) "
138         . "VALUES ($pub_id, $articleVolume, $articleIssue, $journal_id)";
139     $r = mysqli_query($link, $insertJournalArticle);
140     if (!$r) {
141         echo mysqli_error($link);
142     }
143 }
144
145 function insertIntoJournals($link, $name) {
146     $theName = mysqli_real_escape_string($link, trim($name));
147     $select = "SELECT id FROM journals WHERE name = '$theName'";
148     $result = mysqli_query($link, $select);
149     if (mysqli_num_rows($result) == 0) {
150         $insertJournal = "INSERT INTO journals (name) VALUES ('$theName')";
151         $r = mysqli_query($link, $insertJournal);
152         if (!$r) {
153             echo mysqli_error($link);
154         }
155     }
156 }
157
158 function insertIntoProccendings($link, $name) {
159     $theName = mysqli_real_escape_string($link, trim($name));
160     $select = "SELECT id FROM proccendings WHERE name = '$theName'";
161     $result = mysqli_query($link, $select);

```

```

162     if (mysqli_num_rows($result) == 0) {
163         $insertProc = "INSERT INTO proccendings (name) VALUES ('$theName')";
164         $r = mysqli_query($link, $insertProc);
165         if (!$r) {
166             echo mysqli_error($link);
167         }
168     }
169 }
170
171 function insertIntoConferences($link, $name, $date, $location, $volume, $procName) {
172     $theName = mysqli_real_escape_string($link, trim($name));
173     $theDate = mysqli_real_escape_string($link, trim($date));
174     $theLocation = mysqli_real_escape_string($link, trim($location));
175     $theVolume = mysqli_real_escape_string($link, trim($volume));
176     $theProcName = mysqli_real_escape_string($link, trim($procName));
177     $confDate = empty($theDate) ? "NULL" : "'$theDate'";
178     $confLocation = empty($theLocation) ? "NULL" : "'$theLocation'";
179     $procVolume = empty($theVolume) ? "NULL" : "'$theVolume'";
180     $selectConf = "SELECT id FROM conferences WHERE name = '$theName' AND date = '
        ↳ $theDate'";
181     $resultConf = mysqli_query($link, $selectConf);
182     if (mysqli_num_rows($resultConf) == 0) {
183         $select = "SELECT id FROM proccendings WHERE name = '$theProcName'";
184         $result = mysqli_query($link, $select);
185         $row = mysqli_fetch_assoc($result);
186         $proc_id = $row["id"];
187         $insertConference = "INSERT INTO conferences (name, date, location, proc_id,
        ↳ volume) "
188             . "VALUES ('$theName', $confDate, $confLocation, $proc_id, $procVolume
        ↳ )";
189         $r = mysqli_query($link, $insertConference);
190         if (!$r) {
191             echo mysqli_error($link);
192         }
193     }
194 }
195
196 function insertIntoConferenceArticles($link, $pub_id, $confName, $confDate) {
197     $theConfName = mysqli_real_escape_string($link, trim($confName));
198     $theConfDate = mysqli_real_escape_string($link, trim($confDate));
199     $date = empty($theConfDate) ? "NULL" : "'$theConfDate'";
200     $select = "SELECT id FROM conferences WHERE name = '$theConfName' AND date =
        ↳ $date";
201     $result = mysqli_query($link, $select);
202     if (!$result) {
203         echo mysqli_error($link);
204     } else {
205         $row = mysqli_fetch_assoc($result);
206         $conf_id = $row["id"];
207         $insertConferenceArticle = "INSERT INTO conference_articles (pub_id, conf_id)
        ↳ "
208             . "VALUES ($pub_id, $conf_id)";
209         $r = mysqli_query($link, $insertConferenceArticle);
210         if (!$r) {
211             echo mysqli_error($link);
212         }
213     }
214 }
215
216 function isJournalArticle($link, $id) {
217     $select = "SELECT pub_id FROM journal_articles WHERE pub_id = $id";
218     $result = mysqli_query($link, $select);

```

```
219     if(mysqli_num_rows($result) > 0 ){
220         return true;
221     }
222     else{
223         return false;
224     }
225 }
226
227 function updatePersonIfIsProfessorAndHasNullProfidInPersonsTable($link,
    ↪ $listOfAuthors) {
228     foreach ($listOfAuthors as $a) {
229         $trimAuthor = trim($a);
230         $author = mysqli_real_escape_string($link, $trimAuthor);
231         $select = "SELECT id, prof_id FROM persons WHERE name = '$author'";
232         $result = mysqli_query($link, $select);
233         if (mysqli_num_rows($result) > 0) {
234             $row = mysqli_fetch_assoc($result);
235             $id = $row["id"];
236             $profId = $row["prof_id"];
237             $ionianId = selectProfId($link, $author);
238             if(!is_null($ionianId) && ($ionianId != $profId)) {
239                 $update = "UPDATE persons SET prof_id = $ionianId WHERE id = $id";
240                 if(!mysqli_query($link, $update)){
241                     echo mysqli_error($link);
242                 }
243             }
244         }
245     }
246 }
247
248 function updateCitations($link, $title, $citedBy) {
249     $theTitle = mysqli_real_escape_string($link, trim($title));
250     $select = "SELECT id, cited_by FROM publications WHERE title = '$theTitle'";
251     $result = mysqli_query($link, $select);
252     if (mysqli_num_rows($result) == 1) {
253         $row = mysqli_fetch_array($result);
254         if ($row["cited_by"] < $citedBy) {
255             $id = $row["id"];
256             $update = "UPDATE publications SET cited_by = $citedBy WHERE id = $id";
257             if (mysqli_query($link, $update)) {
258                 return true;
259             } else {
260                 echo mysqli_error($link);
261                 return false;
262             }
263         } else {
264             return false;
265         }
266     }
267 }
```


Βιβλιογραφία

- [1] *Σχισιακές Βάσεις Δεδομένων*, Ευάγγελος Κεχρής, 2η Έκδοση 2015
- [2] *Συστήματα Βάσεων Δεδομένων*, Silberschatz, Korth, Sudarsan, 4η Έκδοση 2002
- [3] *L^AT_EX*, Απόστολος Συρόπουλος, 1998
- [4] *Ψηφιακή τυπογραφία με το X_YL^AT_EX*, Απόστολος Συρόπουλος, 2010
- [5] *A relational database for bibliometric analysis*, Nicolai Mallig, Άρθρο, Karlsruhe, 2010
- [6] *Βιβλιομετρική Ανάλυση: Μεθοδολογική Προσέγγιση ΕΚΤ*, Εθνικό Κέντρο Τεκμηρίωσης, Σαχίνη Ε., Μάλλιου Ν., Χούσος Ν., Άρθρο, 3η έκδοση, Δεκέμβριος 2014
- [7] *Εισαγωγή στις PHP 6 & MYSQL 5*, Lary Ullman, Εκδόσεις Κλειδάριθμος, 2009
- [8] <https://www.wikipedia.org/>
- [9] <https://en.wikipedia.org/wiki/H-index>
- [10] <http://azuliadesigns.com/php-function-import-csv-array/>
- [11] <http://www.webslesson.info/2016/10/import-csv-file-data-into-mysql-database-using-php-ajax.html>
- [12] <https://magp.ie/2011/01/28/convert-csv-data-into-an-associative-array/>
- [13] <https://phpenthusiast.com/blog/php-regular-expressions>
- [14] <https://regex101.com/>
- [15] <https://datatables.net>
- [16] <https://www.programcreek.com/2014/05/leetcode-h-index-java/>

- [17] [*https://developers.google.com/chart/*](https://developers.google.com/chart/)
- [18] [*https://d3js.org/*](https://d3js.org/)
- [19] [*https://codyhouse.co/gem/horizontal-timeline*](https://codyhouse.co/gem/horizontal-timeline)
- [20] [*http://w3shaman.com/article/php-ajax-progress-bar*](http://w3shaman.com/article/php-ajax-progress-bar)
- [21] [*https://bl.ocks.org/mohdsanadzakirizvi/6fc325042ce110e1afc1a7124d087130*](https://bl.ocks.org/mohdsanadzakirizvi/6fc325042ce110e1afc1a7124d087130)
- [22] [*http://bl.ocks.org/vlandham/5087480*](http://bl.ocks.org/vlandham/5087480)
- [23] [*https://www.webslesson.info/2017/01/php-pdo-ajax-crud-with-data-tables-and-bootstrap-modals.html*](https://www.webslesson.info/2017/01/php-pdo-ajax-crud-with-data-tables-and-bootstrap-modals.html)
- [24] [*https://www.webslesson.info/2016/04/ajax-delete-multiple-data-with-checkboxes-in-php-jquery-mysql.html*](https://www.webslesson.info/2016/04/ajax-delete-multiple-data-with-checkboxes-in-php-jquery-mysql.html)
- [25] [*https://www.gyrocode.com/projects/jquery-datatables-checkboxes/*](https://www.gyrocode.com/projects/jquery-datatables-checkboxes/)
- [26] [*https://www.w3schools.com/html/default.asp*](https://www.w3schools.com/html/default.asp)
- [27] [*https://www.w3schools.com/css/default.asp*](https://www.w3schools.com/css/default.asp)
- [28] [*https://www.w3schools.com/js/default.asp*](https://www.w3schools.com/js/default.asp)
- [29] [*https://www.w3schools.com/sql/default.asp*](https://www.w3schools.com/sql/default.asp)
- [30] [*https://www.w3schools.com/php/default.asp*](https://www.w3schools.com/php/default.asp)
- [31] [*https://www.w3schools.com/bootstrap4/default.asp*](https://www.w3schools.com/bootstrap4/default.asp)
- [32] [*https://getbootstrap.com/docs/4.0/components/card/*](https://getbootstrap.com/docs/4.0/components/card/)
- [33] [*https://www.overleaf.com/learn*](https://www.overleaf.com/learn)