



**ON-LINE ΣΥΣΤΗΜΑ ΕΥΡΕΣΗΣ ΚΑΙ
ΔΙΑΧΕΙΡΙΣΗΣ ΒΙΒΛΙΟΓΡΑΦΙΚΩΝ ΑΝΑΦΟΡΩΝ**

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΑΤΡΙΒΗ

ΤΟΥ

Γεράσιμου Κ. Χατζηδαμιανού (ΑΕΜ: 203)

Εκπονήθηκε υπό την επίβλεψη του Καθηγητή

Χούστη Ηλία

Εξεταστική Επιτροπή

Επ. Καθηγητής Βερούκιος Βασίλειος

Επ. Καθηγητής Μποζάνης Παναγιώτης

ΒΟΛΟΣ, ΙΟΥΝΙΟΣ 2009

Στην Όλγα,
Με την ελπίδα ότι μαζί μου,
θα είναι πάντα ευτυχισμένη...

Πρόλογος

Η παρούσα μεταπτυχιακή διατριβή ασχολείται με το πρόβλημα της εύρεσης, καταγραφής και διαχείρισης των βιβλιογραφικών αναφορών (citations) που, σε πρακτικό επίπεδο, αντιμετωπίζουν σε κάποιο σημείο της καριέρας τους όλοι οι ασχολούμενοι με την έρευνα (ακαδημαϊκοί, ερευνητές, κτλ). Αλλά και σε θεωρητικό επίπεδο, το πρόβλημα αποτελεί ένα ενδιαφέρον επιστημονικά αντικείμενο, κυρίως λόγω της απουσίας ενός κοινού προτύπου σύνταξης.

Αυτό το δύσκολο πρόβλημα αυτόματης αναζήτησης και καταχώρησης των βιβλιογραφικών αναφορών έχει απασχολήσει πλήθος ερευνητών αλλά και εταιρίες, με αποτέλεσμα να έχουν αναπτυχθεί αρκετά συστήματα, κυρίως για ερευνητικούς σκοπούς, όπως τα Google Scholar, το CiteSeer, το Libra, το Scirus, το ParaCite και DBLP.

Το σύστημα που αναπτύχθηκε στο πλαίσιο της παρούσας μεταπτυχιακής διατριβής, αξιοποιεί τη μηχανή αναζήτησης του Google Scholar για να συλλέξει αποτελέσματα πιθανών βιβλιογραφικών αναφορών και δίνει τη δυνατότητα στο χρήστη να διαχειριστεί τα αποτελέσματα αυτά σύμφωνα με την προσωπική του κρίση, ενώ παράλληλα κάνει και τις δικές του προτάσεις με βάση τις καταχωρημένες δημοσιεύσεις του χρήστη.

Η παρούσα διατριβή εκπονήθηκε στα πλαίσια του Μεταπτυχιακού Προγράμματος Σπουδών του Τμήματος Μηχανικών Η/Υ, Τηλεπικοινωνιών και Δικτύων του Πανεπιστημίου Θεσσαλίας υπό την επίβλεψη του κ. Καθηγητή Χούστη Ηλία.

Σε αυτό το σημείο θα ήθελα να εκφράσω τις θερμές ευχαριστίες μου στον επιβλέποντα Καθηγητή του Τμήματος Μηχανικών Η/Υ, Τηλεπικοινωνιών και Δικτύων του Πανεπιστημίου Θεσσαλίας κ. Χούστη Ηλία και τον διδάσκοντα (βάσει ΠΔ407/80) στο ΤΜΗΥΤΔ Κόκκορα Φώτη για την εμπιστοσύνη που έδειξαν στο πρόσωπό μου με την ανάθεση του θέματος και την πολύτιμη βοήθεια που μου παρείχαν καθ' όλη τη διάρκεια της διατριβής τόσο σε ουσιαστικά τμήματα υλοποίησης όσο και σε διαδικαστικά. Τέλος, θα ήθελα να απευθύνω ένα μεγάλο ευχαριστώ στον Ντονά Κώστα για το υπέρ-πολύτιμο εργαλείο ΔΕΙΧΤο που ανέπτυξε και τις συμβουλές του σε θέματα σχετικά με αυτό.

Γεράσιμος Χατζηδαμιανός

22/06/2009

Περιεχόμενα

ΠΡΟΛΟΓΟΣ	V
ΠΕΡΙΕΧΟΜΕΝΑ	VII
1 ΕΙΣΑΓΩΓΗ	1
2 ΒΙΒΛΙΟΓΡΑΦΙΚΕΣ ΑΝΑΦΟΡΕΣ	5
2.1 Η ΣΗΜΑΣΙΑ ΤΩΝ ΒΙΒΛΙΟΓΡΑΦΙΚΩΝ ΑΝΑΦΟΡΩΝ	5
2.2 ΜΗΧΑΝΕΣ ΑΝΑΖΗΤΗΣΗΣ ΑΝΑΦΟΡΩΝ ΣΤΟ WEB	6
2.2.1 Γενικές Μηχανές Αναζήτησης	6
2.2.2 Εξειδικευμένες Μηχανές Αναζήτησης	7
2.2.3 Σύγκριση Εξειδικευμένων Μηχανών Αναζήτησης	17
2.3 ΠΕΡΙΓΡΑΦΗ ΤΟΥ ΠΡΟΒΛΗΜΑΤΟΣ.....	18
2.3.1 Πρόβλημα ταυτοποίησης αναφορών (CM Problem).....	18
2.3.2 Προβλήματα διάκρισης συγγραφέων (MC & SC Problems)	20
2.4 ΠΕΡΙΛΗΨΗ ΠΡΟΤΕΙΝΟΜΕΝΗΣ ΛΥΣΗ	23
2.5 ΆΛΛΑ ΣΥΣΤΗΜΑΤΑ ΔΙΑΧΕΙΡΙΣΗΣ ΑΝΑΦΟΡΩΝ.....	25
2.5.1 EndNote.....	25
2.5.2 Zotero.....	26
2.5.3 RefWorks.....	27
2.5.4 Publish or Perish.....	28
2.5.5 Σύγκριση Συστημάτων Βιβλιογραφικών Αναφορών.....	30
2.5.6 Σύστημα Εύρεσης και Διαχείρισης Βιβλιογραφικών Αναφορών.....	33
3 ΑΝΑΛΥΣΗ ΑΠΑΙΤΗΣΕΩΝ.....	35
3.1 ΛΕΙΤΟΥΡΓΙΚΕΣ ΑΠΑΙΤΗΣΕΙΣ	35
3.1.1 Γενικές	36
3.1.2 Διαχείρισης αποτελεσμάτων αναζήτησης	36
3.1.3 Διαχείρισης Δημοσιεύσεων.....	37
3.1.4 Διαχείρισης Ετεροαναφορών.....	37
3.1.5 Δημιουργίας Αναφορών	38
3.2 ΜΗ ΛΕΙΤΟΥΡΓΙΚΕΣ ΑΠΑΙΤΗΣΕΙΣ.....	38
3.2.1 Απαιτήσεις Απόδοσης	39
3.2.2 Απαιτήσεις Ασφάλειας	39
4 ΣΧΕΔΙΑΣΗ ΣΥΣΤΗΜΑΤΟΣ	41

4.1	ΔΙΑΓΡΑΜΜΑ ΠΕΡΙΠΤΩΣΕΩΝ ΧΡΗΣΗΣ.....	41
4.2	ΔΙΑΓΡΑΜΜΑΤΑ ΔΡΑΣΤΗΡΙΟΤΗΤΩΝ.....	42
	4.2.1 Διαχείριση αποτελεσμάτων αναζήτησης.....	42
	4.2.2 Διαχείριση δημοσιεύσεων χρήστη.....	45
	4.2.3 Διαχείριση ετεροαναφορών (citations).....	47
	4.2.4 Προβολή αναφορών εκτύπωσης.....	50
	4.2.5 Επεξεργασία στοιχείων χρήστη.....	53
4.3	ΣΧΗΜΑ ΒΑΣΗΣ ΔΕΔΟΜΕΝΩΝ.....	55
	4.3.1 Πίνακας ‘User’.....	56
	4.3.2 Πίνακας ‘Results’.....	56
	4.3.3 Πίνακας ‘Publications’.....	58
	4.3.4 Πίνακας ‘Citations’.....	59
	4.3.5 Πίνακας ‘CitationPublicationMM’.....	59
4.4	ΠΕΡΙΒΑΛΛΟΝ ΛΕΙΤΟΥΡΓΙΑΣ.....	60
	4.4.1 Βάση δεδομένων.....	60
	4.4.2 Γλώσσα προγραμματισμού.....	61
	4.4.3 Λειτουργικό σύστημα.....	61
5	ΥΛΟΠΟΙΗΣΗ ΣΥΣΤΗΜΑΤΟΣ.....	63
5.1	ΠΕΡΙΒΑΛΛΟΝ ΥΛΟΠΟΙΗΣΗΣ.....	63
	5.1.1 jQuery JavaScript Library [37].....	63
	5.1.2 Yii Web Programming Framework [36].....	64
	5.1.3 Eclipse Development Platform [35].....	65
5.2	ΒΑΣΙΚΕΣ ΛΕΙΤΟΥΡΓΙΕΣ.....	66
	5.2.1 Αναζήτηση, δημιουργία και διαχείριση ετεροαναφορών.....	66
	5.2.2 Υπολογισμός «προτάσεων συστήματος».....	68
	5.2.3 Δημιουργία και διαχείριση προσωπικών δημοσιεύσεων του χρήστη.....	71
	5.2.4 Συσχέτιση των προσωπικών δημοσιεύσεων με τις ετεροαναφορές.....	71
5.3	ΠΡΟΒΛΗΜΑΤΑ ΥΛΟΠΟΙΗΣΗΣ.....	72
5.4	ΣΕΝΑΡΙΑ ΧΡΗΣΗΣ.....	72
	5.4.1 Σενάριο 1: Αναζήτηση βιβλιογραφικών αναφορών και δημιουργία ετεροαναφορών.....	72
	5.4.2 Σενάριο 2: Διαχείριση δημοσιεύσεων και ετεροαναφορών.....	76
6	ΣΥΜΠΕΡΑΣΜΑΤΑ & ΕΠΕΚΤΑΣΕΙΣ.....	81
6.1	ΣΥΜΠΕΡΑΣΜΑΤΑ.....	81
6.2	ΕΠΕΚΤΑΣΕΙΣ – ΜΕΛΛΟΝΤΙΚΗ ΕΡΓΑΣΙΑ.....	82
	ΒΙΒΛΙΟΓΡΑΦΙΑ.....	87

1 Εισαγωγή

Οι αλματώδεις ρυθμοί με τους οποίους προοδεύει η επιστήμη σήμερα, καθώς και οι ολοένα αυξανόμενοι ερευνητικοί τομείς που προκύπτουν, έχουν σαν αποτέλεσμα την παραγωγή ενός ασύλληπτα μεγάλου αριθμού δημοσιεύσεων σε παγκόσμιο επίπεδο. Η ανάγκη για άμεση επικοινωνία και διάχυση της παραγόμενης γνώσης επέβαλλε αρκετούς βασικούς κανόνες στη διαδικασία συγγραφής και παρουσίασης των ερευνητικών πονημάτων και με τη βοήθεια της τεχνολογίας επιλύθηκαν πλήθος σημαντικών σχετικών προβλημάτων.

Η ανάγκη για διατήρηση της συμβατότητας με τον παλιό τρόπο συγγραφής ερευνητικού έργου, η έλλειψη αυστηρών προτύπων σύνταξης βιβλιογραφικών αναφορών, καθώς και η συμβολή του ανθρώπινου παράγοντα σε ορθογραφικά λάθη είχαν σαν αποτέλεσμα τη δημιουργία σημαντικού αριθμού προβλημάτων ιδιαίτερα όσον αφορά στην ταυτοποίηση αναφορών (Citation Matching problem) και τη διάκριση των συγγραφέων (Mixed Citation και Split Citation problems). Για μια λεπτομερέστερη ανάλυση των προβλημάτων αυτών, ο αναγνώστης παραπέμπεται στις παραγράφους 2.3.1 και 2.3.2 αντίστοιχα. Χάρη στην εξέλιξη των αλγορίθμων έχει επιτευχθεί σημαντική πρόοδος στην επίλυση των παραπάνω προβλημάτων, ωστόσο ακόμη δεν είναι δυνατή η ασφαλής εύρεση βιβλιογραφικών αναφορών χωρίς την ανθρώπινη συνδρομή.

Τα παραπάνω λήφθηκαν υπόψη κατά τη σχεδίαση του συστήματος που παρουσιάζεται σε αυτή τη διατριβή, στόχος του οποίου είναι η διευκόλυνση του ερευνητή στην αναζήτηση ετεροαναφορών στο ερευνητικό του έργο και η οργάνωση των προσωπικών του δημοσιεύσεων. Για την αναζήτηση των ετεροαναφορών το σύστημα παρέχει την απαραίτητη υποδομή η οποία επιτρέπει τη χρήση της εξειδικευμένης μηχανής αναζήτησης Google Scholar για τη λήψη των αποτελεσμάτων αναζήτησης. Η αξιοποίηση των αποτελεσμάτων αναζήτησης γίνεται χάρη σε μια φιλική διεπαφή η οποία επιτρέπει την κατηγοριοποίηση αυτών, την εισαγωγή σχολίων και την απ' ευθείας δημιουργία αρχείου ετεροαναφορών. Ιδιαίτερης μνείας αξίζει το γεγονός ότι το σύστημα έχει τη δυνατότητα να «θυμάται» την κατάσταση επεξεργασίας των αποτελεσμάτων διευκολύνοντας με αυτόν τον τρόπο τόσο την διαδικασία επεξεργασίας του συνήθως μεγάλου πλήθους

αυτών όσο και τη μη εμφάνιση ήδη επεξεργασμένων αποτελεσμάτων. Επιπλέον, το σύστημα περιλαμβάνει αλγόριθμο ανάλυσης των αποτελεσμάτων του Google Scholar σε σχέση με το καταγεγραμμένο συγγραφικό έργο του χρήστη, που οδηγεί σε πρόταση για πιθανή ταξινόμηση.

Το κείμενο της διατριβής ακολουθεί τη δομή μιας επιστημονικής αναφοράς. Έτσι, αρχικά παρουσιάζεται η ανάγκη για την δημιουργία ενός συστήματος εύρεσης και διαχείρισης βιβλιογραφικών αναφορών, περιγράφοντας το πρόβλημα και την υφιστάμενη κατάσταση. Στη συνέχεια δίνεται μια συνοπτική περιγραφή του συστήματος που προτείνεται ενώ παρουσιάζονται και σχετικές εργασίες τρίτων. Τα υπόλοιπα κεφάλαια ασχολούνται με περισσότερο τεχνικά ζητήματα, όπως ανάλυση απαιτήσεων, σχεδίαση του συστήματος και υλοποίηση αυτού. Πιο αναλυτικά, το περιεχόμενο των κεφαλαίων έχει ως εξής:

- ***Κεφάλαιο 2. Βιβλιογραφικές Αναφορές***

Στην παράγραφο 2.1 περιγράφεται η σημασία των βιβλιογραφικών αναφορών και οι λόγοι που τις καθιστούν τόσο σημαντικές. Η παρουσίαση της υφιστάμενης κατάστασης πραγματοποιείται στις παραγράφους 2.2.1, 2.2.2 και 2.2.3 όπου γίνεται μια απλή αναφορά στα δύο είδη μηχανών αναζήτησης, γενικές και εξειδικευμένες και εξετάζονται οι πέντε πιο σημαντικές εξειδικευμένες μηχανές αναζήτησης: Scirus, Google Scholar, Microsoft Libra, CiteSeer^x και DBLP. Η παράγραφος 2.3 περιέχει την περιγραφή των τριών πιο σημαντικών προβλημάτων που συναντώνται στην ταυτοποίηση και διάκριση βιβλιογραφικών αναφορών ενώ στην ενότητα 2.4 παρουσιάζεται η προτεινόμενη λύση στο γενικότερο ζήτημα της εύρεσης και διαχείρισης ετεροαναφορών. Το κεφάλαιο κλείνει με την αναφορά σε σημαντικές εμπορικές και μη εφαρμογές, οι οποίες σχετίζονται με κάποιο τρόπο με το σύστημα που αναπτύχθηκε στην εργασία αυτή.

- ***Κεφάλαιο 3. Ανάλυση Απαιτήσεων***

Το κεφάλαιο 3 δίνει μια λεπτομερή περιγραφή των απαιτήσεων των χρηστών από ένα σύστημα σαν το προτεινόμενο. Χωρίζεται σε λειτουργικές (παράγραφος 3.1) και μη λειτουργικές απαιτήσεις (παράγραφος 3.2) και η μελέτη τους μπορεί να δώσει στον αναγνώστη όλες εκείνες τις απαραίτητες πληροφορίες, οι οποίες θα τον βοηθήσουν να κατανοήσει πλήρως τη λειτουργικότητα του συστήματος.

- **Κεφάλαιο 4. Σχεδίαση Συστήματος**

Στο κεφάλαιο 4 παρουσιάζονται τα διαγράμματα που δημιουργήθηκαν σε UML κατά τη σχεδίαση του συστήματος. Πιο συγκεκριμένα, η ενότητα 4.1 περιλαμβάνει το διάγραμμα περιπτώσεων χρήσης ή αλλιώς Use Case Model όπου παρουσιάζονται οι υπηρεσίες που έχει στη διάθεσή του ο χρήστης. Οι υπηρεσίες αυτές αναλύονται στη συνέχεια σε διαγράμματα δραστηριοτήτων (παράγραφος 4.2) όπου απεικονίζεται τόσο με γραφικό τρόπο όσο και με λεκτικό η ροή λειτουργιών του συστήματος. Στην ενότητα 4.3 περιγράφεται αναλυτικά το σχήμα της βάσης δεδομένων που αναπτύχθηκε και παρατίθεται η επεξήγηση όλων των πεδίων για καθένα πίνακα ξεχωριστά. Τέλος, η παράγραφος 4.4 εστιάζει στο περιβάλλον όπου θα λειτουργεί το σύστημα και εξηγεί το σκεπτικό πίσω από την επιλογή συστήματος διαχείρισης βάσεων δεδομένων, web server, γλώσσας προγραμματισμού και λειτουργικού συστήματος.

- **Κεφάλαιο 5. Υλοποίηση Συστήματος**

Το κεφάλαιο 5 ασχολείται με το θέμα της υλοποίησης του συστήματος και τον τρόπο με τον οποίο επιτελούνται συγκεκριμένες βασικές λειτουργίες, παραθέτοντας τα δύο συνηθέστερα σενάρια χρήσης του συστήματος. Πιο συγκεκριμένα, στην ενότητα 5.1 παρουσιάζεται το περιβάλλον στο οποίο υλοποιήθηκε το σύστημα και το οποίο αποτελείται από το Eclipse IDE, το Yii Framework και τη βιβλιοθήκη jQuery για JavaScript. Στην ενότητα 5.2 παρουσιάζονται οι βασικές λειτουργίες του συστήματος όπως αναζήτηση βιβλιογραφικών αναφορών, υπολογισμός προτάσεων κατηγοριοποίησης αποτελεσμάτων του συστήματος, διαχείριση δημοσιεύσεων και ετεροαναφορών καθώς και συσχέτιση αυτών. Η ενότητα 5.3 αναφέρεται στην αντιμετώπιση του προβλήματος λήψης δεδομένων από το Google Scholar. Το κεφάλαιο κλείνει με δύο πραγματικά σενάρια χρήσης συνοδευόμενα από εικόνες του συστήματος σε λειτουργία.

- **Κεφάλαιο 6. Συμπεράσματα & Επεκτάσεις**

Στο τελευταίο κεφάλαιο παρατίθενται τα συμπεράσματα στο οποία καταλήγει η διατριβή και επιπλέον γίνεται μια εκτενής αναφορά στις πιθανές επεκτάσεις και βελτιώσεις που μπορεί να γίνουν στο σύστημα.

2 Βιβλιογραφικές Αναφορές

Σε έναν κόσμο όπου υπέρογκα ποσά επενδύονται στους τομείς έρευνας και ανάπτυξης, είναι προφανής η ανάγκη για διάχυση, διαμοιρασμό και εύκολη εύρεση της γνώσης που παράγεται με φρενήρεις ρυθμούς, σε παγκόσμιο επίπεδο. Πλήθος σημαντικών εργαλείων έχουν αναπτυχθεί με στόχο τη εξυπηρέτηση των ερευνητικών αναγκών και την οργάνωση της γνώσης σε πηγές που να εντοπίζονται εύκολα.

2.1 Η Σημασία των Βιβλιογραφικών Αναφορών

Οι βιβλιογραφικές αναφορές κατέχουν εξέχουσα θέση σε όλα σχεδόν τα επιστημονικά έργα. Κανένα έργο, κανενός ερευνητή δεν μπορεί να υπάρξει από μόνο του χωρίς τις αναφορές, ή με άλλα λόγια την παραπομπή του αναγνώστη σε άλλα παρεμφερή έργα, όπως βιβλία, άρθρα, ιστοσελίδες, συνέδρια κ.α.

Οι βιβλιογραφικές αναφορές δίνουν τη δυνατότητα στον αναγνώστη να κατανοήσει καλύτερα **το κείμενο** το οποίο μελέτα, παρέχοντάς του πηγές στις οποίες μπορεί να ανατρέξει για περισσότερες λεπτομέρειες. Επιπλέον του δίνουν τη δυνατότητα να ενημερωθεί καλύτερα για **το ερευνητικό πεδίο** το οποίο μελετά καθώς μέσα από τη μελέτη των αναφερόμενων πηγών του δίνεται η δυνατότητα να γνωρίσει άλλους τρόπους αντιμετώπισης του ίδιου προβλήματος, ερευνητές που ασχολούνται με το ίδιο αντικείμενο, καθώς και γενικότερες πληροφορίες σχετικά με την κατεύθυνση των ερευνών.

Η αξία των βιβλιογραφικών αναφορών δεν περιορίζεται στην παροχή πηγών στον αναγνώστη για περαιτέρω ανάγνωση, αλλά αποτελεί και μια μορφή αναγνώρισης και αξιολόγησης του ερευνητικού έργου των ακαδημαϊκών που αναφέρονται. Πιο συγκεκριμένα, όσο πιο αξιολογη είναι μια ερευνητική δουλειά, τόσο πιο συχνά αποτελεί αντικείμενο αναφοράς από έργα άλλων επιστημόνων του χώρου. Είναι σαφές ότι η αναζήτηση του αντίκτυπου που έχει το έργο ενός ερευνητή στην επιστημονική κοινότητα αποτελεί πάγια ανάγκη για τον ίδιο και η διαδικασία της αναζήτησης αυτής είναι γνωστή και ως ανάλυση ετεροαναφορών (citation analysis).

Όπως γίνεται εύκολα κατανοητό, το έργο της ανάλυσης των ετεροαναφορών ήταν μια ιδιαίτερα επίπονη διαδικασία στο παρελθόν όπου είτε υπήρχε παντελής έλλειψη του

Internet ως μέσο, είτε δεν είχαν δημιουργηθεί ακόμη οι απαραίτητες υπηρεσίες που θα μπορούσαν να βοηθήσουν τον ερευνητή σε ένα τέτοιο έργο. Με την πάροδο των χρόνων και χάρη στη ραγδαία εξέλιξη του Παγκόσμιου Ιστού, η διαδικασία αναζήτησης σχετικής βιβλιογραφίας και αναφορών απλοποιήθηκε, χωρίς όμως να έχουν εξαλειφθεί πλήρως ορισμένα βασικά προβλήματα.

2.2 Μηχανές Αναζήτησης Αναφορών στο Web

Στο επίπεδο του Παγκόσμιου Ιστού, οι επιλογές που έχει κάποιος ερευνητής όταν επιθυμεί να αναζητήσει και να βρει τα ερευνητικά έργα που έχουν αναφορές σε δικές του δημοσιεύσεις, είναι δύο:

1. οι μηχανές αναζήτησης γενικού σκοπού, όπως Google, Microsoft Bing, Yahoo κ.α.
2. οι εξειδικευμένες μηχανές αναζήτησης, όπως Google Scholar, Microsoft Libra, CiteSeer^x, Scirus, ParaCite, DBLP κ.α.

Στη συνέχεια της ενότητας γίνεται μια απλή αναφορά στον τρόπο με τον οποίο μπορούν να γίνουν αναζητήσεις βιβλιογραφικών αναφορών με τη χρήση γενικών μηχανών αναζήτησης και παρουσιάζονται αναλυτικά οι πιο σημαντικές εξειδικευμένες μηχανές αναζήτησης για βιβλιογραφικές αναφορές. Τέλος, επιχειρείται μια σύγκριση των παρουσιασθέντων μηχανών αναζήτησης με στόχο την ανάδειξη των πλεονεκτημάτων και των μειονεκτημάτων της καθεμιάς.

2.2.1 Γενικές Μηχανές Αναζήτησης

Η μέθοδος εύρεσης πληροφορίας με τη χρήση γενικών μηχανών αναζήτησης είναι πολύ αποτελεσματική όταν η πληροφορία που αναζητείται είναι γενική ενώ δεν αποδίδει το ίδιο καλά όταν η πληροφορία που αναζητείται είναι εξειδικευμένη.

Ένας από τους λόγους είναι ότι τα επιστρεφόμενα αποτελέσματα συνήθως είναι υπερβολικά πολλά για να εξεταστούν ένα προς ένα, με τα περισσότερα από αυτά να είναι από λίγο σχετικά έως εντελώς άσχετα με το θέμα και ως εκ τούτου η χρήσιμη πληροφορία να χάνεται μέσα στην λιγότερο χρήσιμη ή και εντελώς ανούσια.

Ένας άλλος λόγος είναι ότι οι μηχανές αναζήτησης προβάλλουν τα αποτελέσματα με σειρά που αυτές θεωρούν ότι θα ενδιέφερε περισσότερο το χρήστη. Έτσι προτεραιότητα παίρνουν τα αποτελέσματα που είναι στην ίδια γλώσσα με την τοποθεσία του

χρήστη, μετά έρχονται τα αποτελέσματα που είναι στην αγγλική και στη συνέχεια στις υπόλοιπες γλώσσες. Ενώ η πρωτοβουλία αυτή που παίρνει το σύστημα είναι ιδιαίτερα χρήσιμη σε γενικού τύπου αναζητήσεις, σε εξειδικευμένου τύπου αναζητήσεις όπως βιβλιογραφικές αναφορές, μπορεί κάλλιστα να οδηγήσει σε απόκρυψη χρήσιμων πληροφοριών.

Η αναζήτηση βιβλιογραφικών αναφορών με τη χρήση γενικών μηχανών παρουσιάζει επίσης τη δυσκολία ότι πρέπει το «κλειδί» (keyword) που θα χρησιμοποιηθεί για την αναζήτηση να είναι βέλτιστο όσον αφορά στην ποιότητα των αποτελεσμάτων. Με άλλα λόγια, κάποιος ακαδημαϊκός, χρησιμοποιώντας ένα σύνηθες κλειδί για να αναζητήσει βιβλιογραφικές αναφορές σε έργα του, όπως π.χ. το ονοματεπώνυμό του, θα λάβει συνήθως ένα μεγάλο πλήθος αποτελεσμάτων που θα αφορούν στο πρόσωπό του, σε συνώνυμα πρόσωπα καθώς και σε απλές αναφορές στον ίδιο ή και σε ταυτόνομα πρόσωπα.

Γίνεται λοιπόν εύκολα κατανοητό ότι ο χρόνος που απαιτείται από τον χρήστη προκειμένου να ξεκαθαρίσει αυτό το τεράστιο πλήθος αποτελεσμάτων είναι πολύ μεγάλος και στις περισσότερες των περιπτώσεων η όλη διαδικασία δεν είναι εφικτή. Επιπρόσθετα, το γεγονός του ότι σε επόμενη αναζήτηση η μηχανή αναζήτησης θα συμπεριλάβει και εγγραφές που ο χρήστης έχει ήδη ξεδιαλύνει, αποθαρρύνει τους χρήστες από την αναζήτηση βιβλιογραφικών αναφορών μέσω γενικών μηχανών αναζήτησης.

2.2.2 Εξειδικευμένες Μηχανές Αναζήτησης

Σε αντίθεση με τις γενικές μηχανές αναζήτησης, οι εξειδικευμένες μηχανές παρέχουν την πληροφορία που ο χρήστης επιθυμεί να λάβει πιο αξιόπιστα και με λιγότερο κόπο. Με τον όρο «εξειδικευμένες μηχανές» αναφερόμαστε σε εκείνες τις μηχανές αναζήτησης οι οποίες παρέχουν υποδομές δεικτοδότησης και αναζήτησης για πληροφορίες σχετικές με επιστημονικά έργα. Σε αρκετές περιπτώσεις ορισμένες μηχανές παρέχουν ακόμη μεγαλύτερη εξειδίκευση προσφέροντας υπηρεσίες μόνο για συγκεκριμένους επιστημονικούς τομείς.

Στη συνέχεια της ενότητας θα παρουσιαστούν εν συντομία ορισμένες από τις πιο δημοφιλείς εξειδικευμένες μηχανές αναζήτησης για βιβλιογραφικές αναφορές ανεξάρτητα με τους επιστημονικούς τομείς στους οποίους εξειδικεύονται.

Scirus

Η μηχανή αναζήτησης Scirus, είναι ένα από τα πιο πλήρη εργαλεία επιστημονικής αναζήτησης στον παγκόσμιο ιστό. Η βάση δεικτοδότησης του συγκεκριμένου εργαλείου αποτελείται από 485 εκατομμύρια συγγραφικά έργα και η ιδιαιτερότητά του είναι ότι επιτρέπει στους χρήστες να πραγματοποιούν αναζητήσεις τόσο μέσα στα άρθρα που έχουν δημοσιευτεί όσο και στις προσωπικές ιστοσελίδες των επιστημόνων, στις ιστοσελίδες μαθημάτων, σε προεκτυπωμένο υλικό που βρίσκεται δε διάφορους διακομιστές, σε αρχεία πατεντών κ.α.



Εικόνα 1. Αρχική σελίδα του Scirus (<http://www.scirus.com/>)

Το σύστημα Scirus σχεδιάστηκε για χρήση από ερευνητές και επιστήμονες οποιουδήποτε κλάδου και έχει αποδειχτεί ιδιαίτερα επιτυχημένο στον εντοπισμό επιστημονικών αποτελεσμάτων στον παγκόσμιο ιστό, δεδομένου ότι ψηφίστηκε ως η Καλύτερη Εξειδικευμένη Μηχανή Αναζήτησης (Best Specialty Search Engine) το 2001 και το 2002 και ως το Καλύτερο Ευρετήριο ή Μηχανή Αναζήτησης (Best Directory or Search Engine Website) τις χρονιές 2004, 2005 και 2006.

Το σύστημα Scirus τέθηκε σε λειτουργία για πρώτη φορά το 2001 από τον μεγαλύτερο εκδοτικό οίκο ιατρικών και επιστημονικών συγγραμμάτων Elsevier. Οι υπηρεσίες του διατίθενται δωρεάν στο κοινό, ενώ ορισμένα αποτελέσματα απαιτούν την ύπαρξη συνδρομής σε πακέτα ψηφιακών βιβλιοθηκών.

Το όνομα Scirus προήλθε από τον έλληνα προφήτη «Σκύρο», ο οποίος σύμφωνα με το διήγημα «*Η περιγραφή της Ελλάδας*» του Πausανία, πολέμησε στο πλευρό των Ελευσίνιων εναντίον του Ηράκλειτου και πέθανε στη μάχη αφού προηγουμένως έχτισε ένα ναό προς τιμή της Αθηνάς Σκύρου. Σύμφωνα με τους δημιουργούς του συστήματος, επιλέχθηκε το συγκεκριμένο όνομα, καθώς, με τον ίδιο τρόπο που οι προφήτες κρίνουν από τα σημάδια το τι πρόκειται να ακολουθήσει, έτσι και η μηχανή αναζήτησης Scirus θα μπορεί να υποστηρίζει ενεργά τον ρόλο των επιστημόνων ως προφήτες.

Μερικές από τις πιο σημαντικές υπηρεσίες που παρέχει το σύστημα Scirus είναι οι εξής:

- Φιλτράρισμα των μη-επιστημονικών αποτελεσμάτων, π.χ. η αναζήτηση για REM δε θα επιστρέψει το γνωστό μουσικό συγκρότημα αλλά πληροφορίες σχετικές με το φαινόμενο REM κατά τη διάρκεια του ύπνου κ.α.
- Εύρεση και υποβληθέντων άρθρων τα οποία είναι στο στάδιο του ελέγχου (peer review) και είναι συνήθως αόρατα από τις άλλες μηχανές αναζήτησης
- Εύρεση της σχετικότερης δυνατής πληροφορίας χάρη στην αναζήτηση στον πιο πλήρη συνδυασμό πηγών πληροφορίας στον παγκόσμιο ιστό
- Δυνατότητα αναζήτησης σε ένα μεγάλο εύρος επιστημονικών πεδίων όπως υγεία, φυσικές και κοινωνικές επιστήμες κ.α.
- Περιορισμός της αναζήτησης με βάση κάποιον συγκεκριμένο συγγραφέα, άρθρο ή δημοσίευση
- Περιορισμός των αποτελεσμάτων σε ένα συγκεκριμένο εύρος ημερομηνιών
- Εύρεση επιστημονικών συνεδρίων, περιλήψεων και πατεντών
- Προσαρμογή και αποθήκευση των αποτελεσμάτων αναζήτησης

Google Scholar

Το σύστημα Google Scholar είναι το εργαλείο αναζήτησης, που δημιούργησε για τους ακαδημαϊκούς, η μεγαλύτερη και ισχυρότερη μηχανή αναζήτησης αυτή τη στιγμή στον κόσμο, η Google. Το Google Scholar προέκυψε από μια συζήτηση που είχαν οι επιστήμονες πληροφορικής Alex Verstak και Anurag Acharya, οι οποίοι εκείνη την περίοδο εργαζόταν στη δημιουργία του κυρίως ευρετηρίου της Google. Η πρώτη Beta έκδοση του Google Scholar έγινε διαθέσιμη στο κοινό το Νοέμβριο του 2004 κατά την οποία

δεικτοδοτήθηκαν πλήρη κείμενα ακαδημαϊκών δημοσιεύσεων διαφόρων φόρμα και επιστημονικών πεδίων.

Σήμερα, το Google Scholar δίνει τη δυνατότητα στους χρήστες του να αναζητήσουν πληροφορίες στον παγκόσμιο ιστό από ένα μεγάλο πλήθος επιστημονικών περιοδικών, περιλήψεων, άρθρων προς δημοσίευση (peer-reviewed articles), διπλωματικών εργασιών, διατριβών, βιβλίων, παρουσιάσεων PowerPoint και τεχνικών αναφορών. Οι πηγές του είναι κυρίως τα πανεπιστήμια, τα ακαδημαϊκά ινστιτούτα, οι επαγγελματικές κοινότητες, οι ερευνητικές ομάδες και οι βάσεις δεδομένων προτυπωμένων άρθρων (preprint repositories) σε ολόκληρο τον κόσμο.



Εικόνα 2. Αρχική σελίδα του Google Scholar (<http://scholar.google.gr/>)

Τα άρθρα που προβάλλονται στα αποτελέσματα του Google Scholar ταξινομούνται με τρόπο παρόμοιο με αυτόν που χρησιμοποιούν οι ερευνητές. Πιο συγκεκριμένα, δίνεται βαρύτητα στο πλήρες κείμενο του άρθρου, στον συγγραφέα, τη δημοσίευση στην οποία εμφανίζεται το συγκεκριμένο άρθρο καθώς και το πόσο συχνά αναφέρεται το συγκεκριμένο άρθρο στην υπόλοιπη ακαδημαϊκή βιβλιογραφία. Τα πιο σχετικά αποτελέσματα εμφανίζονται στην αρχική σελίδα.

Το 2006 σε απάντηση στη δημιουργία της ανταγωνιστικής υπηρεσίας Windows Live Academic Search από τη Microsoft, προστέθηκε η δυνατότητα εισαγωγής αναφορών μέσω προγραμμάτων διαχείρισης αναφορών όπως τα RefWorks, RefMan, EndNote, και BibTeX, ενώ το 2007 ανακοινώθηκε από τον Acharya η έναρξη ενός

προγράμματος ψηφιοποίησης και φιλοξενίας των άρθρων περιοδικών κατόπιν συμφωνίας με τους εκδότες τους.

Ορισμένα από τα χαρακτηριστικά του Google Scholar που το έκαναν ιδιαίτερα δημοφιλές, είναι:

- η ταχύτητα στην εμφάνιση αποτελεσμάτων
- η δυνατότητα υποβολής σύνθετων ερωτημάτων με βάση το όνομα του συγγραφέα, τον τίτλο της δημοσίευσης και τη χρονολογία αυτής
- η απλή διεπαφή με την οποία είναι ήδη εξοικειωμένοι όλοι οι χρήστες του Google παγκοσμίως
- η διασύνδεση με ψηφιακές βιβλιοθήκες, οι οποίες επιτρέπουν την προβολή των άρθρων με ή χωρίς συνδρομή
- η μεγαλύτερη πληρότητα σε δεικτοδοτημένο περιεχόμενο σε σχέση με τις άλλες μηχανές αναζήτησης
- η ύπαρξη δεσμών που οδηγούν απευθείας στο πλήρες άρθρο σε περίπτωση που αυτό είναι διαθέσιμο χωρίς την απαίτηση συνδρομής, όπως για παράδειγμα στην ιστοσελίδα του συγγραφέα
- το πλήθος των πληροφοριών που παρέχονται στα αποτελέσματα αναζήτησης, όπως τίτλος άρθρου, συγγραφείς, συνέδριο και τύπος δημοσίευσης, έτος, περίληψη, δυνατότητα πλοήγησης σε σχετικά άρθρα κ.α.

Microsoft Libra

Η μηχανή αναζήτησης Libra της Microsoft, για ακαδημαϊκά συγγράμματα, εξειδικεύεται στο ερευνητικό πεδίο της επιστήμης των υπολογιστών και αναπτύχθηκε από την Microsoft Research Asia. Για την εύρεση των αποτελεσμάτων, χρησιμοποιεί τη μέθοδο της αντικειμενοστραφούς κάθετης αναζήτησης κατά την οποία οι δημοσιεύσεις (papers), οι συγγραφείς, τα συνέδρια και οι διάφορες περιοδικές εκδόσεις (journals) θεωρούνται διαφορετικοί τύποι αντικειμένων. Ο βαθμός σχετικότητας του κάθε αντικειμένου με το «κλειδί» αναζήτησης υπολογίζεται λαμβάνοντας υπόψη όλες τις πληροφορίες που έχουν συλλεχθεί σχετικά με τα χαρακτηριστικά του καθώς και με τη σχέση του με άλλα αντικείμενα.

Η βάση δεδομένων του συστήματος Libra περιλαμβάνει ακαδημαϊκά άρθρα τα οποία δημοσιεύτηκαν σε περιοδικές επιστημονικές εκδόσεις (journals) ή εργασίες. Μέχρι

τον Μάιο του 2007 περιελάμβανε περισσότερα από 1.8 εκατομμύρια άρθρα και 1.3 εκατομμύρια συγγραφείς, ενώ σήμερα το σύστημα εκτελεί αναζητήσεις σε πάνω από 3 εκατομμύρια δημοσιεύσεις.



Εικόνα 3. Η αρχική σελίδα του Libra (<http://libra.msra.cn/>)

Ορισμένες από τις σημαντικότερες υπηρεσίες που παρέχει το Libra είναι:

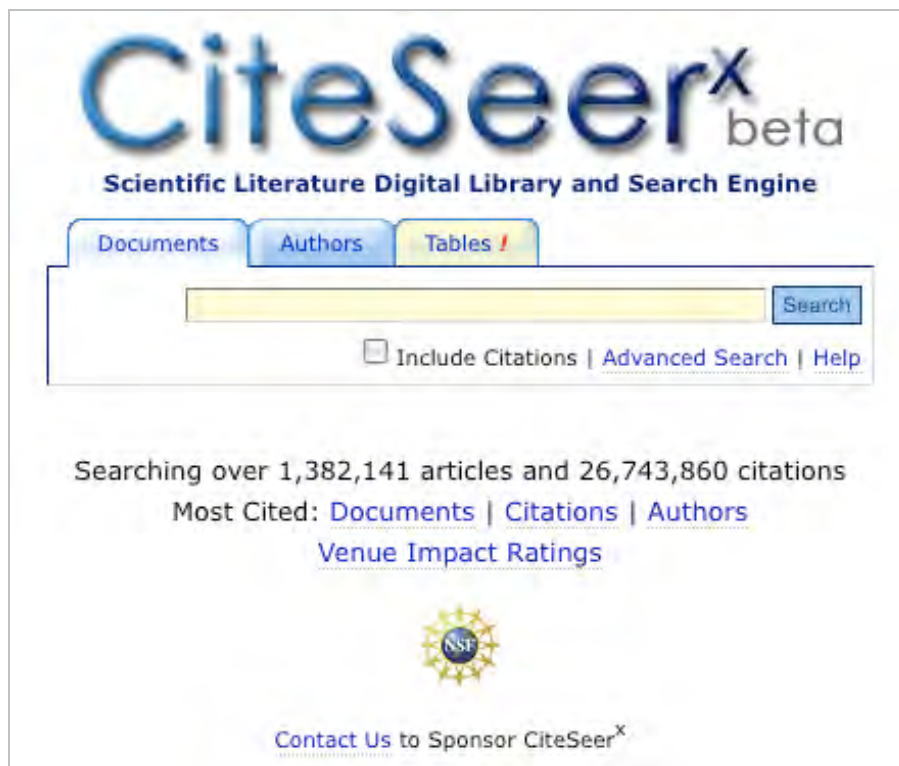
- Αναζήτηση κορυφαίων επιστημόνων, συνεδρίων και περιοδικών σε ένα συγκεκριμένο επιστημονικό πεδίο
- Πληροφορίες σχετικά με την ανάπτυξη και εξέλιξη των ερευνητικών κοινοτήτων
- Εύρεση των κορυφαίων δημοσιεύσεων
- Αναγνώριση των ανερχόμενων αστεριών ή των πιο σημαντικών θεμάτων στον τομέα ενδιαφέροντος του χρήστη

CiteSeer^x

Είναι μία από τις πιο γνωστές μηχανές αναζήτησης με εξειδίκευση στον επιστημονικό τομέα των υπολογιστών και την επιστήμη της πληροφορικής. Το CiteSeer^x αποτελεί τον απόγονο του CiteSeer, το οποίο αναπτύχθηκε το 1997 από το ινστιτούτο ερευνών NEC, Princeton, New Jersey από τους Steve Lawrence, Lee Giles και Kurt Bollacker. Το CiteSeer ήταν η πρώτη ψηφιακή βιβλιοθήκη και μηχανή αναζήτησης, η

οποία παρείχε αυτόματη δεικτοδότηση βιβλιογραφικών αναφορών και σύνδεση αυτών, τεχνικές υπηρεσίες και λογισμικό προς χρήση από άλλες ψηφιακές βιβλιοθήκες, καθώς και πληροφορίες όπως αλγόριθμοι, δεδομένα και μετα-δεδομένα.

Ο στόχος του CiteSeer ήταν να βελτιώσει τόσο τη διάχυση της γνώσης όσο και την πρόσβαση στην ακαδημαϊκή και επιστημονική βιβλιογραφία γι' αυτό και λειτουργεί ως μη κερδοσκοπικός οργανισμός, ο οποίος παρέχει τις υπηρεσίες του δωρεάν, όντας ενεργό μέλος του κινήματος ανοιχτής πρόσβασης (open access movement).



Εικόνα 4. Αρχική σελίδα του CiteSeer^x (<http://citeseerx.ist.psu.edu/>)

Το όνομα CiteSeer προέκυψε από κάποιο λογοπαίγνιο το οποίο έχει δύο πιθανές εκδοχές. Σύμφωνα με την πρώτη προήλθε από τη λέξη «*sightseer*» η οποία ερμηνεύεται ως ένας τουρίστας που βλέπει τα αξιοθέατα, οπότε παρόμοια η λέξη «*cite seer*» αναφέρεται σε έναν ερευνητή που βλέπει τις αναφορές κειμένων. Σύμφωνα με τη δεύτερη εκδοχή, «*seer*» είναι ο προφήτης και «*cite seer*» είναι ο προφήτης των αναφορών. Από πρακτικής απόψεως, το CiteSeer δίνει τη δυνατότητα στο χρήστη να υποβάλλει ερωτήματα με βάση τις βιβλιογραφικές ετεροαναφορές (citations) και επιστρέφει τα αποτελέσματα ταξινομημένα σύμφωνα με την απήχηση που αυτά έχουν στην επιστημονική κοινότητα.

Το CiteSeer παρέδωσε τη σκυτάλη στην επόμενη γενιά, το CiteSeer^x περίπου στο 2005 όποτε και σταμάτησαν οι αναβαθμίσεις του CiteSeer καθώς ήταν αδύνατο να εξελιχθεί λόγω των περιορισμών που επέβαλε η αρχιτεκτονική του συστήματος. Σήμερα εξακολουθεί να έχει ένα αντιπροσωπευτικό δείγμα ερευνητικών συγγραμμάτων αλλά δεν είναι ιδιαίτερα εκτενές, καθώς έχει πρόσβαση μόνο στα δημοσίως διαθέσιμα άρθρα τα οποία συνήθως βρίσκονται στην ιστοσελίδα του συγγραφέα ή έχουν υποβληθεί από τον ίδιο το συγγραφέα.

Οι πιο σημαντικές υπηρεσίες που παρέχει το σύστημα CiteSeer^x είναι οι ακόλουθες:

- Αυτόματη δεικτοδότηση (indexing) επιστημονικών άρθρων, που βρίσκονται σε μορφή Post-Script και PDF ελεύθερα στον Παγκόσμιο Ιστό
- Στατιστικά αναφορών και συνδέσμους προς σχετικά κείμενα για τα άρθρα που βρίσκονται στη βάση δεδομένων
- Προβολή του περιεχομένου των αναφορών σε ένα άρθρο, επιτρέποντας στους ενδιαφερόμενους να δουν γρήγορα και εύκολα τι λένε οι άλλοι για κάποιο άρθρο
- Εντοπισμός σχετικών κειμένων χρησιμοποιώντας μετρικές βασισμένες στις λέξεις και τις αναφορές
- Εμφάνιση ποσοστού ομοιότητας των κειμένων
- Δυνατότητα αναζήτησης με βάση τα αρχικά του συγγραφέα
- Εξαγωγή μετα-δεδομένων για όλα τα κείμενα με αυτόματο τρόπο
- Αυτόματος υπολογισμός της συνεισφοράς ενός άρθρου στην επιστημονική κοινότητα (automatic acknowledgement indexing)

DBLP

Με το ακρωνύμιό του να μεταφράζεται ως Digital Bibliography & Library Project, το DBLP είναι ένα σύστημα το οποίο παρέχει βιβλιογραφικές πληροφορίες αποκλειστικά για την επιστήμη της πληροφορικής. Αποτελεί το πιο παλιό σύστημα καθώς δημιουργήθηκε στη δεκαετία του 1980 όπου αρχικά φιλοξενούνταν ως ένας μικρός διακομιστής στο πανεπιστήμιο της Τριέρης στη Γερμανία (Universität Trier) και εξυπηρετούσε μια μικρή, τοπική κοινωνία επιστημόνων. Στην πορεία το μικρό αυτό project γιγαντώθηκε με αποτέλεσμα να περιλαμβάνει σήμερα περισσότερα από 1.200.000 άρθρα σχετικά με την επιστήμη των υπολογιστών και να λειτουργεί και πέντε mirror sites.

dblp.uni-trier.de
Computer Science
Bibliography

Universität Trier

The DBLP Computer Science Bibliography

maintained by [Michael Ley](#)

Welcome to **DBLP**. This server provides bibliographic information on major **computer science journals and proceedings**. DBLP indexes more than one million articles and contains more than 10000 links to home pages of computer scientists.

Search for a **person**: [->Help](#)

General search on DBLP: [CompleteSearch](#) - [Faceted search](#)

Conferences & Workshops: [A](#) - [B](#) - [C](#) - [D](#) - [E](#) - [F](#) - [G](#) - [H](#) - [I](#) - [J](#) - [K](#) - [L](#) - [M](#) - [N](#) - [O](#) - [P](#) - [Q](#) - [R](#) - [S](#) - [T](#) - [U](#) - [V](#) - [W](#) - [X](#) - [Y](#) - [Z](#)

Journals: [A](#) - [B](#) - [C](#) - [D](#) - [E](#) - [F](#) - [G](#) - [H](#) - [I](#) - [J](#) - [K](#) - [L](#) - [M](#) - [N](#) - [O](#) - [P](#) - [Q](#) - [R](#) - [S](#) - [T](#) - [U](#) - [V](#) - [W](#) - [X](#) - [Y](#) - [Z](#)

DBLP is available from several hosts: [Trier I](#) - [Trier II](#) - [ACM SIGMOD](#) - [VLDB](#) - [SunSITE CE](#)

Your Contribution ... : Please read the [FAQ Pages](#). Send an e-mail to ley@uni-trier.de, if ...

- ... you notice errors or omissions. I try to unify different spellings of names, often this isn't possible without your help.
- ... you are listed, but a link to your home page is missing or broken.
- ... if you edit a journal or proceedings and want this publication to be included into the DBLP server.

Disclaimer: All information on the DBLP server is provided in good faith, but we take no responsibility for the accuracy of the information. Please read [the copyright page](#).

[Home](#) | [Conferences](#) | [Journals](#) | [Series](#) | [FAQ](#) — Search: [Faceted](#) | [Complete](#) | [Author](#)


Copyright © Thu Jan 15 11:09:38 2009 by [Michael Ley](#) (ley@uni-trier.de)

Εικόνα 5. Η αρχική σελίδα του συστήματος DBLP (<http://dblp.uni-trier.de/>)

Οι πληροφορίες που δεικτοδοτεί το DBLP πηγάζουν από το VLDB, τα IEEE Transactions και ACM Transactions καθώς και άρθρα από εργασίες συνεδρίων. Ιδιαίτερης μνείας χρήζει το γεγονός ότι για τη συντήρηση/διαχείριση του συστήματος υπεύθυνος είναι αποκλειστικά ο Michael Ley, ο οποίος μάλιστα βραβεύτηκε για το έργο του τρεις φορές. Το 1997, από το ACM SIGMOD (Service Award) και το VLDB Endowment (Special Recognition Award) και το 2003 πάλι από το ACM SIGMOD (Contribution Award).

Στα πλεονεκτήματα του DBLP περιλαμβάνεται η δυνατότητα φυλλομέτρησης όλων των συνεδρίων (conferences) και των περιοδικών που έχουν δεικτοδοτηθεί από το σύστημα και η δυνατότητα εξαγωγής των αποτελεσμάτων σε μορφή BibTeX. Οι αναφορές σε δεικτοδοτημένα συγγράμματα από συνέδρια, περιοδικά, σειρές και βιβλία παρέχονται στη μορφή καταλόγου (Εικόνα 6) και σε μερικές περιπτώσεις καλύπτονται χρονολογικά πολύ παλιές εκδόσεις για τον χώρο της επιστήμης των υπολογιστών, όπως για παράδειγμα το περιοδικό Communications of the ACM από το 1958 και τα συνέδρια SIGMOD και Very Large Data Bases (VLDB) από το 1975.

dblp.uni-trier.de
Computer Science
Bibliography



The DBLP Computer Science Bibliography

maintained by [Michael Ley](#) - [Welcome](#) - [FAQ](#)

DBLP is available from several hosts: [Trier I](#) - [Trier II](#) - [ACM SIGMOD](#) - [VLDB](#) - [SunSITE CE](#)

Search

- [Author](#)
- [Faceted search](#) (L3S Research Center, U. Hannover)
- [CompleteSearch](#) (Holger Bast, Max Planck Institut f. Inf.)

Bibliographies

- **Conferences:** [SIGMOD](#), [VLDB](#), [PODS](#), [ER](#), [EDBT](#), [ICDE](#), [POPL](#), ...
- **Journals:** [CACM](#), [TODS](#), [TOIS](#), [TOPLAS](#), [DKE](#), [VLDB J.](#), [Inf. Systems](#), [TPLP](#), [TCS](#), ...
- **Series:** [LNCS/LNAI](#), [IFIP](#)
- **Books:** [Reference](#) - [Collections](#) - [DB Textbooks](#)
- **By Subject:** [Database Systems](#), [Logic Prog.](#), [IR](#), ...

Full Text: [ACM SIGMOD Anthology](#)

Links

- **Computer Science Organizations:** [ACM](#) ([DL](#) / [SIGMOD](#) / [SIGIR](#)), [IEEE Computer Society](#) ([DL](#)), [IEEE Xplore](#), [IFIP](#), ...
- **Related Services:** [Google Scholar](#), [CiteSeer](#)/[CiteSeerX](#), [CS BibTeX](#) ([DBLP](#)), [io-port.net](#), [CoRR](#), [HAL](#), [NZ-DL](#), [Zentralblatt MATH](#), [MathSciNet](#), [Erdős Number Proj.](#), [Math Genealogy Proj.](#), [BibSonomy](#), [CiteULike](#), [Libra](#), [Arnetminer](#), ...

Εικόνα 6. Σελίδα καταλόγου DBLP (<http://dblp.uni-trier.de/db/>)

Στα μειονεκτήματα του DBLP περιλαμβάνονται το όχι ιδιαίτερα εύχρηστο και λειτουργικό περιβάλλον του συστήματος, η έλλειψη προχωρημένης αναζήτησης καθώς και το γεγονός ότι δεν παρέχεται δυνατότητα για αναζήτηση με βάση άλλα στοιχεία π.χ. keywords, το κείμενο ή τον τίτλο του άρθρου, παρά μόνο με βάση τον συγγραφέα. Πλέον μπορεί κάποιος να αναζητήσει δημοσιεύσεις βάση λέξεων κλειδιών που υπάρχουν στο τίτλο και το κείμενο των δημοσιεύσεων, χάρη σε mirrors όπως CompleteSearch (<http://dblp.mpi-inf.mpg.de/dblp-mirror/index.php>) και Faceted search (<http://dblp.l3s.de/>) οι οποίοι συγχρονίζονται σε καθημερινή βάση με το DBLP.

Συνοψίζοντας, οι υπηρεσίες που προσφέρονται από το DBLP είναι:

- μια πλήρης βάση δεδομένων βιβλιογραφικών αναφορών σχετικές με την επιστήμη των υπολογιστών
- εμφάνιση των βιβλιογραφικών αναφορών σε μορφή BibTeX
- δεικτοδοτημένους καταλόγους σημαντικών/ιστορικών εντύπων
- παροχή συνδέσμου προς το πλήρες κείμενο
- αναζήτηση κατά την πληκτρολόγηση μέσω του mirror CompleteSearch

2.2.3 Σύγκριση Εξειδικευμένων Μηχανών Αναζήτησης

Σε αυτή την ενότητα επιχειρείται μια σύγκριση των εξειδικευμένων – σε ακαδημαϊκά συγγράμματα – μηχανών αναζήτησης, προκειμένου να αποκτήσει ο αναγνώστης μια πληρέστερη εικόνα σχετικά με τις προσφερόμενες υπηρεσίες και τον ειδικότερο τομέα εξειδίκευσης των μηχανών.

Ο Πίνακας 1 παρουσιάζει στοιχεία σχετικά με τα επιστημονικά πεδία για τα οποία περιλαμβάνει πληροφορίες η κάθε μία από τις εξεταζόμενες μηχανές αναζήτησης. Εύκολα παρατηρεί κανείς πως οι μηχανές Scirus και Google Scholar προσφέρουν δυνατότητα αναζήτησης σε όλα τα επιστημονικά πεδία, οι Microsoft Libra και CiteSeer^x στα πεδία των θετικών επιστημών, ενώ η DBLP μόνο στην επιστήμη των υπολογιστών.

Μηχανή αναζήτησης	Όλα τα επιστημονικά πεδία	Φυσικές επιστήμες, μαθηματικά και επιστήμη των υπολογιστών	Επιστήμη των υπολογιστών
Scirus	X		
Google Scholar	X		
Microsoft Libra		X	
CiteSeer ^x		X	
DBLP			X

Πίνακας 1. Σχέση μηχανών αναζήτησης και επιστημονικών πεδίων

Μηχανή Αναζήτησης	Μέγεθος Βάσης Δεδομένων	Περιεχόμενο Ευρητήριου	Έτος Δημιουργίας
Scirus	485 εκατομμύρια συγγραφικά έργα	Πλήρες κείμενο, Βιβλιογραφικές αναφορές	2001
Google Scholar	500 εκατομμύρια κείμενα	Πλήρες κείμενο	2004
Microsoft Libra	3 εκατομμύρια δημοσιεύσεις	Πλήρες κείμενο	-
CiteSeer ^x	1,3 εκατομμύρια άρθρα 26,7 εκατομμύρια αναφορές	Πλήρες κείμενο, Βιβλιογραφικές αναφορές	1997
DBLP	1,2 εκατομμύρια αναφορές	Βιβλιογραφικές αναφορές	Δεκαετία 1980

Πίνακας 2. Σύγκριση των μηχανών αναζήτησης με βάση το μέγεθος και το είδος των εγγραφών, καθώς και το έτος δημιουργίας της κάθε μηχανής.

Ο Πίνακας 2 παρέχει πληροφορίες σχετικά με το πλήθος και το είδος των εγγραφών που υπάρχουν στη βάση δεδομένων του κάθε συστήματος καθώς και για το έτος δημιουργίας του κάθε ενός. Παρατηρούμε λοιπόν ότι η DBLP είναι η παλαιότερη μηχανή

αναζήτησης και ως εκ τούτου περιέχει τις περισσότερες βιβλιογραφικές αναφορές στο επιστημονικό πεδίο της επιστήμης των υπολογιστών. Τα συστήματα Scirus και Google Scholar έχουν παραπλήσιο μέγεθος και επιτρέπουν την αναζήτηση στο πλήρες κείμενο του εγγράφου καθώς αυτό δεικτοδοτείται ολόκληρο.

2.3 Περιγραφή του Προβλήματος

Είναι προφανές ότι ο όγκος και ο βαθμός εξειδίκευσης της πληροφορίας που διατίθεται ελεύθερα στον παγκόσμιο ιστό σχετικά με το αντικείμενο των ακαδημαϊκών δημοσιεύσεων και των βιβλιογραφικών αναφορών, απαιτεί υποδομές και αυτοματισμούς για τη διαχείρισή του. Στις περιπτώσεις των μηχανών αναζήτησης Scirus και Google Scholar, οι προσφερόμενες πληροφορίες είναι τόσο πλήρεις και καλά δομημένες που ευνοούν την ανάπτυξη συστημάτων όπως αυτό της παρούσης διατριβής.

Το πρόβλημα που καλείται να λύσει κάθε ερευνητής όταν για παράδειγμα επιχειρεί τη σύνταξη ενός βιογραφικού, είναι η καταγραφή και οργάνωση της πληροφορίας σχετικά με το ερευνητικό του έργο και την απήχηση που αυτό έχει στην επιστημονική κοινότητα. Στην πλειοψηφία των περιπτώσεων δεν υπάρχει πλήρης εικόνα και έλεγχος των αναφορών που έχουν γίνει στα έργα ενός ερευνητή καθώς δεν υπάρχει "το απόλυτο" σύστημα αναζήτησης οργάνωσης και διαχείρισης αυτής της πληροφορίας.

Η εύρεση των ετεροαναφορών (citations) για συγκεκριμένη ακαδημαϊκή δημοσίευση και η προσεκτική και οργανωμένη καταγραφή τους είναι ένα δύσκολο και χρονοβόρο εγχείρημα. Το πρόβλημα γίνεται εντονότερο, όταν το πλήθος των υπό εξέταση έργων είναι μεγάλο. Ως τυπικό παράδειγμα όγκου δουλειάς που απαιτείται, ένας σχετικά νέος ερευνητής με 20 δημοσιεύσεις για παράδειγμα, χρειάζεται μερικές ώρες σύνθετης αναζήτησης και πλοήγησης στον Παγκόσμιο Ιστό για να βρει όλες τις λεπτομέρειες από τις περισσότερες αναφορές τρίτων στο έργο του.

Στη συνέχεια αναφέρονται ορισμένα από τα προβλήματα που εμφανίζονται κατά την αναζήτηση βιβλιογραφικών αναφορών.

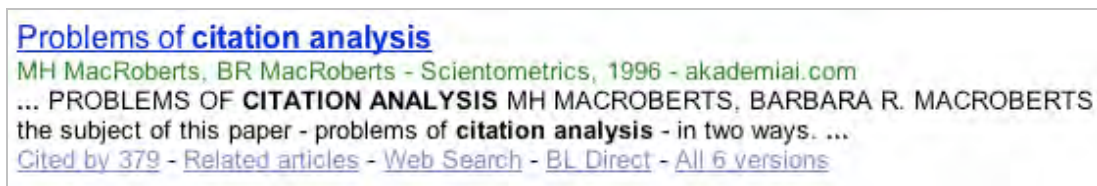
2.3.1 Πρόβλημα ταυτοποίησης αναφορών (CM Problem)

Το πρόβλημα ταυτοποίησης αναφορών (Citation Matching Problem), δημιουργείται από την έλλειψη ενός τυποποιημένου (standard) προτύπου συγγραφής αναφορών. Λόγω της έλλειψης του εν λόγω προτύπου, ο κάθε συγγραφέας συντάσσει τη βιβλιογραφία

στις δημοσιεύσεις του με τρόπο διαφορετικό, είτε αλλάζοντας τη σειρά με την οποία εμφανίζονται τα πεδία της αναφοράς (π.χ. τίτλος, χρονολογία δημοσίευσης, ονόματα των συγγραφέων κλπ.), είτε παραλείποντας ορισμένα πεδία, είτε ακόμη παραθέτοντας τους τίτλους περιοδικών ή συνεδρίων σε σύντμηση. Η πιθανή ύπαρξη ορθογραφικών λαθών εντείνει το πρόβλημα με αποτέλεσμα αναφορές στο ίδιο έργο να μην μπορούν να ταυτοποιηθούν αυτόματα με την σωστή δημοσίευση.

Ορισμένες εξειδικευμένες μηχανές αναζήτησης όπως το CiteSeer και το Google Scholar αντιμετωπίζουν σε ικανοποιητικό βαθμό το πρόβλημα ταυτοποίησης αναφορών, χρησιμοποιώντας άπληστους (greedy) αλγορίθμους ομαδοποίησης, με τους οποίους μετράται η ομοιότητα του κειμένου. Φυσικά στον χρήστη παρουσιάζεται μόνο το αποτέλεσμα της σωστής αναφοράς και μια επιλογή για εμφάνιση των υπόλοιπων εκδόσεων.

Για παράδειγμα, η αναζήτηση στο Google Scholar με keyword «*citation analysis*» επιστρέφει, μεταξύ άλλων και το αποτέλεσμα που εμφανίζεται στην Εικόνα 7 όπου στο δεξί μέρος παρατηρούμε πως υπάρχει ένας δεσμός «**All 6 versions**».



Εικόνα 7. Αποτέλεσμα αναζήτησης με keyword «*citation analysis*» στο Google Scholar

Ακολουθώντας το δεσμό αυτό βλέπουμε την ίδια δημοσίευση με ορισμένες διαφορές στα πεδία τους, οι οποίες αναγνωρίστηκαν από το Google Scholar ως ίδιες (Εικόνα 8).

[Problems of citation analysis](#)

MH MacRoberts, BR MacRoberts - *Scientometrics*, 1996 - [akademai.com](#)
We have approached the subject of this paper - problems of citation analysis - in two ways. In the first section, we list the traditionally recognized problems of citation analysis and briefly summarize empirical research on each. In ...
[Cited by 379](#) - [Related articles](#) - [Web Search](#) - [BL Direct](#)

[Problems of Citation Analysis: A Critical Review.](#)

MH MacRoberts, BR MacRoberts - *Journal of the American Society for Information Science*, 1989 - [eric.ed.gov](#)
EJ401129 - Problems of Citation Analysis: A Critical Review.
[Cached](#) - [Web Search](#)

[Problems of citation analysis: a critical review](#)

MH MACROBERTS, BR MACROBERTS - *Journal of the American Society for Information Science*, 1989 - [cat](#)
Problems of citation analysis: a critical review. MH MACROBERTS, BR MACROBERTS *Journal of the American Society for Information Science* 40:55, 342-349, John Wiley & Sons, 1989. Une analyse critique ...
[Web Search](#)

[Problems of citations analysis](#)

MH MACROBERTS, BR MACROBERTS - *Scientometrics(Print)*, 1996 - [cat.inist.fr](#)
Problems of citations analysis. MH MACROBERTS, BR MACROBERTS *Scientometrics(Print)* 36:33, 435-444, Kluwer, 1996. Critique de l'analyse des citations: cette analyse ne reflète pas l'influence réelle d'un auteur sur un scientifique donné. ...
[Web Search](#)

[Problems of citation analysis: A critical review](#)

MH MacRoberts, BR MacRoberts - *Journal of the American Society for Information Science*, 1989 - [interscienc](#)
Problems of Citation Analysis: A Critical Review ... Michael H. MacRoberts and Barbara Ft. MacRoberts 740 Columbia, Shreveport, Louisiana 71704 ... We review the problems of citation analysis. Most of them have either not been ...
[Web Search](#)

[Problems of citation analysis](#)

MH MacRoberts, BR MacRoberts - *Scientometrics*, 1996 - [Springer](#)
We have approached the subject of this paper - problems of citation analysis - in two ways. In the first section, we list the traditionally recognized problems of citation analysis and briefly summarize empirical research on each. In ...
[Web Search](#)

Εικόνα 8. Οι έξι διαφορετικές εκδόσεις της ίδιας δημοσίευσης στο Google Scholar

2.3.2 Προβλήματα διάκρισης συγγραφέων (MC & SC Problems)

Το πρόβλημα διάκρισης συγγραφέων αφορά σε λάθη που οφείλονται στη δυσκολία αντιστοίχισης των συγγραφέων με τις δημοσιεύσεις τους και χωρίζεται σε δύο παρόμοια υποπροβλήματα. Το πρώτο ονομάζεται mixed citation problem και αφορά σε εκείνες τις περιπτώσεις όπου οι αναφορές ενός συγγραφέα συγχέονται με τις αναφορές ενός άλλου και θεωρούνται μόνο του ενός, ενώ το δεύτερο ονομάζεται split citation problem και είναι ακριβώς το αντίθετο, δηλαδή αναφορές σε συγγράμματα του ίδιου συγγραφέα εμφανίζονται σαν να είναι αναφορές σε συγγράμματα κάποιου άλλου.

Mixed Citation Problem

Το mixed citation problem παρατηρείται κυρίως σε περιπτώσεις όπου οι συγγραφείς είναι συνονόματοι, έχουν όμοια ονόματα, εισάγονται λάθος στο σύστημα ή έχουν τα ίδια αρχικά. Έτσι, σε περίπτωση που ένας εκ των δύο συγγραφέων έχει πολύ πιο πλούσιο συγγραφικό έργο σε σύγκριση με τον άλλο, δημιουργείται σύγχυση και καθίσταται αδύνατη η ασφαλής αναγνώριση της απήχησης των έργων του καθενός.

Το πρόβλημα επιδεικνύεται καλύτερα με τη βοήθεια του συστήματος DBLP. Αναζητώντας έναν συγγραφέα με κοινό όνομα, όπως για παράδειγμα «**A. Smith**», επιστρέφονται οι τέσσερις εγγραφές που παρουσιάζονται στην Εικόνα 9. Ενώ θα περίμενε κανείς οι εγγραφές να αντιστοιχούν στον ίδιο ακαδημαϊκό, στη συγκεκριμένη περίπτωση οι τέσσερις εγγραφές αντιστοιχούν σε τρεις διαφορετικούς συγγραφείς με το ίδιο όνομα.

2007	
4	EE C. G. Kilsby, P. D. Jones, A. Burton, A. C. Ford, H. J. Fowler, C. Harpham, P. James, A. Smith, R. L. Wilby: A daily use in climate change studies. <i>Environmental Modelling and Software</i> 22(12): 1705-1719 (2007)
2006	
3	EE T. J. Jones, Carl Allen Reidsema, A. Smith: Automated Feature Recognition System for supporting conceptual engine <i>Journal</i> 10(6): 477-492 (2006)
2005	
2	EE T. J. Jones, Carl Allen Reidsema, A. Smith: Supporting Design for Manufacture Through Neutral Files and Feature R <i>2005</i> : 496-502
1982	
1	A. Smith, D. H. Bell: A Multiprogramming System for a Microcomputer. <i>Softw., Pract. Exper.</i> 12(5): 405-410 (1982)

Εικόνα 9. Αποτέλεσμα αναζήτησης στο DBLP με keyword «A. Smith»

Γίνεται λοιπόν εύκολα κατανοητό, ότι το συγκεκριμένο πρόβλημα έχει ιδιαίτερα μεγάλη επίπτωση στην αναγνώριση της συνεισφοράς του ακαδημαϊκού καθώς αυτή δεν μπορεί να μετρηθεί με ακρίβεια. Αυτό που καθιστά το συγκεκριμένο πρόβλημα ακόμη πιο δύσκολο, είναι το γεγονός ότι η λύση του δεν είναι καθόλου προφανής ακόμα και για τον άνθρωπο. Πώς είναι δυνατόν να καταλάβει κάποιος ότι γίνεται αναφορά σε περισσότερα από ένα άτομα όταν έχουν το ίδιο όνομα; Η λογική απάντηση σε αυτό το ερώτημα, είναι η εξέταση επιπλέον πληροφορίας για τη λήψη της απόφασης. Με άλλα λόγια, αυτό που θα έκανε κάποιος για να δει αν πρόκειται για ετεροπροσωπία, είναι η αναζήτηση στοιχείων που θα μπορούσαν να χαρακτηρίζουν μοναδικά τον συγγραφέα. Τέτοια στοιχεία θα μπορούσαν να είναι:

- η θέση εργασίας αν πρόκειται για ερευνητή κάποιας εταιρίας ή το πανεπιστήμιο αν πρόκειται για ακαδημαϊκό
- η λίστα με τους υπόλοιπους συγγραφείς, οι οποίοι συνετέλεσαν στη συγγραφή των δημοσιεύσεων
- συνηθισμένες λέξεις κλειδιά που πιθανώς χρησιμοποιούνται στα άρθρα

Οι Dongwon Lee et al. και Tan Y. F. et al. προτείνουν ([16], [17]) δύο εξαιρετικούς αλγορίθμους, οι οποίοι θα μπορούσαν να χρησιμοποιηθούν για την επίλυση του Mixed Citation προβλήματος με μηχανιστικό τρόπο. Η ανάλυση των αλγορίθμων αυτών ξεφεύγει από τους σκοπούς της παρούσης διατριβής. Ο αναγνώστης που ενδιαφέρεται να μελετήσει σε μεγαλύτερο βάθος τους συγκεκριμένους αλγορίθμους παραπέμπεται στα σχετικά άρθρα.

Split Citation Problem

Το Split Citation Problem δημιουργείται από τη μη προτυποποιημένη εισαγωγή των ονομάτων των συγγραφέων στις ψηφιακές βιβλιοθήκες. Πιο συγκεκριμένα, είναι πιθανό το όνομα ενός συγγραφέα να εμφανιστεί σε πολλές παραλλαγές, όπως για παράδειγμα ολόκληρο το ονοματεπώνυμο, μόνο τα αρχικά, κάποιος συνδυασμός αρχικών και επωνύμου κ.α. Όπως και στα προηγούμενα προβλήματα, έτσι και σε αυτό, το πρόβλημα εντείνεται από τις περιπτώσεις λάθους εισαγωγής στοιχείων. Σε κάθε περίπτωση, το αποτέλεσμα είναι ότι το σύστημα θεωρεί την κάθε εγγραφή σαν ξεχωριστό συγγραφέα με άμεση συνέπεια στην εύρεση των αναφορών που αντιστοιχούν σε συγκεκριμένο συγγραφέα.

Πρόκειται για συνηθισμένο πρόβλημα των ψηφιακών βιβλιοθηκών και η σοβαρότητα του έγκειται στο γεγονός ότι οι αναφορές, που κανονικά θα έπρεπε να αφορούν σε ένα συγκεκριμένο άτομο, διασπώνται σε αναφορές σε έργα διαφορετικών ατόμων με αποτέλεσμα το μέγεθος της συνεισφοράς του καθενός στην επιστημονική κοινότητα να μην μπορεί να υπολογιστεί ακριβώς και να εμφανίζεται κατά πολύ μειωμένο. Όπως είναι σαφές, όσο πιο πολλές παραλλαγές υπάρχουν για το όνομα του συγκεκριμένου ατόμου, τόσο πιο έντονο είναι και το πρόβλημα.

Επιπλέον, σε περίπτωση που ένας αναγνώστης αναζητά δημοσιεύσεις αυτού του ατόμου, τότε στα αποτελέσματα δεν θα επιστραφούν όλες οι δημοσιεύσεις αλλά μόνο αυτές που αντιστοιχούν στο ακριβές όνομα με το οποίο έγινε η αναζήτηση. Το σύστημα

θεωρεί ότι οι υπόλοιπες δημοσιεύσεις ανήκουν σε άλλους συγγραφείς και δεν τις επιστρέφει με αποτέλεσμα να χάνεται σημαντική πληροφορία για τον αναγνώστη.

Μια απλή επίδειξη του προβλήματος μπορεί να γίνει εύκολα αν κάποιος αναζητήσει στο Google Scholar δημοσιεύσεις του Καθηγητή κ. Χούστη Ηλία με keyword «**Houstis**». Θα παρατηρήσει ότι επιστρέφονται άρθρα του συγκεκριμένου ακαδημαϊκού με το όνομά του να έχει μία από τις παρακάτω εκδοχές:

- Elias N. Houstis
- E. Houstis
- Elias Houstis
- E. N. Houstis

Η επίλυση του εν λόγω προβλήματος μπορεί να επιτευχθεί, ως ένα βαθμό, εξετάζοντας ενδελεχώς τις υπόλοιπες πληροφορίες που μπορούν να αντληθούν από μια βιβλιογραφική αναφορά. Τέτοιες πληροφορίες είναι τα ονόματα των υπόλοιπων συγγραφέων, οι τίτλοι των άρθρων καθώς και τα ονόματα των συνεδρίων ή περιοδικών όπου αυτά δημοσιεύτηκαν. Χρησιμοποιώντας διάφορες ευφυείς ευριστικές τεχνικές και αλγορίθμους μέτρησης αποστάσεων συμβολοσειρών (π.χ. Jaro, Levenstein, Q-gram) είναι δυνατόν να ανιχνευθεί η ομοιότητα ανάμεσα σε δημοσιεύσεις του ίδιου συγγραφέα και με αυτόν τον τρόπο να παρουσιαστούν χωρίς την ύπαρξη του Split Citation Problem. Ο ενδιαφερόμενος αναγνώστης μπορεί να αναζητήσει περαιτέρω πληροφορίες στις εργασίες [16], [17], [18] και [19].

2.4 Περίληψη Προτεινόμενης Λύση

Έχοντας αναφερθεί εκτενώς στην αναγκαιότητα της ύπαρξης ενός συστήματος διαχείρισης βιβλιογραφικών αναφορών (ενότητα 2.1), έχοντας παρουσιάσει την τρέχουσα κατάσταση αναφορικά με το αντικείμενο της αναζήτησης και εύρεσης βιβλιογραφικών αναφορών (ενότητα 2.2) και έχοντας εκθέσει τα σημαντικότερα προβλήματα που αντιμετωπίζει η αυτόματη αναζήτηση τους (ενότητα 2.3), αντιλαμβάνεται κανείς ότι το πρόβλημα δεν είναι άμεσα επιλύσιμο. Επιπλέον, οι μη πλήρεις λύσεις πολλές φορές δεν είναι αποδεκτές καθώς μια-δυο μόνο επιπλέον αναφορές μπορεί για παράδειγμα να επηρεάσουν σημαντικά την κρίση ενός ερευνητή ή ακαδημαϊκού.

Στο πλαίσιο της εργασίας κατασκευάστηκε ένα «*On-line σύστημα Εύρεσης και Διαχείρισης Βιβλιογραφικών Αναφορών*» του οποίου στόχος είναι:

- να αξιοποιεί την αδιαμφισβήτητη ισχύ των εξειδικευμένων μηχανών αναζήτησης
- να επιτρέπει την οργάνωση και διαχείριση τόσο των δημοσιεύσεων (publications) ενός ερευνητή, όσο και των αναφορών (citations) που γίνονται σε αυτές από άλλους
- να επιτρέπει την αποσπασματική και σε επεισόδια διαχείριση των αποτελεσμάτων των μηχανών αναζήτησης
- να μπορεί κατά το δυνατό να επεξεργάζεται αυτόματα τα αποτελέσματα των μηχανών αναζήτησης

Το «*Citations Management System*», όπως είναι η αγγλική ονομασία της εφαρμογής, αποτελεί μια προσπάθεια οργάνωσης των δημοσιεύσεων των ερευνητών καθώς και των αναφορών σε αυτές. Παρέχει τη δυνατότητα αναζήτησης δημοσιεύσεων και βιβλιογραφικών αναφορών με την απ' ευθείας υποβολή ερωτημάτων στην εξειδικευμένη μηχανή αναζήτησης Google Scholar, όπως επίσης τη δημιουργία και διαχείριση προσωπικού αρχείου με δημοσιεύσεις (publications) και αναφορές σε αυτές (citations). Τα αποτελέσματα που επιστρέφονται από την υποβολή των ερωτημάτων, παρουσιάζονται με αρθρωτό τρόπο στον χρήστη και είναι στη διάθεσή του προς αξιολόγηση και ταξινόμηση χωρίς κανέναν χρονικό περιορισμό. Με αυτόν τον τρόπο μπορεί ο χρήστης να επεξεργάζεται τα αποτελέσματα όποτε αυτός θεωρεί βολικό. Επιπρόσθετα, αποτελέσματα τα οποία έχουν ήδη ταξινομηθεί από τον χρήστη δεν του εμφανίζονται ως «μη επεξεργασμένα» εκ νέου, με αποτέλεσμα να μειώνεται αισθητά ο χρόνος που απαιτείται για την επεξεργασία αποτελεσμάτων που προκύπτουν σε μελλοντικές αναζητήσεις. Όσο περισσότερο ο χρήστης χρησιμοποιεί το σύστημα, τόσο μειώνεται ο χρόνος που απαιτούν οι μελλοντικές συνεδρίες χρήσης με σκοπό την διατήρηση ενός ενημερωμένου αρχείου αναφορών. Το σύστημα παρέχει επίσης σε κάθε αποτέλεσμα μια «πρόταση ταξινόμησης» (classification suggestion) προκειμένου να βοηθήσει τον χρήστη στην ταξινόμηση των αποτελεσμάτων.

Αξίζει σε αυτό το σημείο να σημειωθεί ότι το εν λόγω σύστημα δεν αποτελεί κάποιου είδους μετρητή της απήχησης ενός ερευνητή, ούτε μπορεί να χρησιμοποιηθεί από τα εκπαιδευτικά ιδρύματα για αξιολόγηση των επιδόσεων των ακαδημαϊκών τους καθώς δεν αποτελεί αντικείμενο της παρούσης μελέτης ο υπολογισμός και παρουσίαση των εν λόγω στατιστικών, μετρικών και συμπερασμάτων.

Πριν την εκτενή παρουσίαση του συστήματος, θεωρείται χρήσιμη μια ανασκόπηση των υπαρχόντων συστημάτων διαχείρισης βιβλιογραφικών αναφορών, ώστε να γίνουν πιο σαφείς οι διαφορές στις παρεχόμενες υπηρεσίες και δυνατότητες.

2.5 Άλλα Συστήματα Διαχείρισης Αναφορών

Καθώς το πρόβλημα της εύρεσης και οργάνωσης βιβλιογραφικών αναφορών δεν είναι καινούριο, θεωρείται επιβεβλημένο να γίνει μια συνοπτική αναφορά σε παρόμοια, υπάρχοντα συστήματα. Στη συνέχεια παρουσιάζονται τα συστήματα EndNote, Zotero, RefWorks και το αρκετά διαφορετικό Publish or Perish. Επιπλέον γίνεται και σύγκριση των δυνατοτήτων τους.

Τέλος, αναφορά γίνεται και σε Σύστημα Εύρεσης και Διαχείρισης Βιβλιογραφικών Αναφορών σε μορφή desktop Java εφαρμογής που αναπτύχθηκε στο πλαίσιο της πτυχιακής εργασίας της Εύας Ζιάκα. Το on-line σύστημα διαχείρισης βιβλιογραφικών αναφορών της παρούσας διατριβής αποτελεί μια ριζικά ανανεωμένη, εκ του μηδενός επανακωδικοποιημένη και δραματικά βελτιωμένη εκδοχή της παραπάνω εφαρμογής.

2.5.1 EndNote

Το σύστημα EndNote της Thomson Scientific είναι η παλαιότερη desktop εφαρμογή οργάνωσης βιβλιογραφικών αναφορών καθώς η πρώτη της έκδοση δημιουργήθηκε το 1988. Με την πάροδο των χρόνων εξελίχθηκε σε ένα πλήρες εμπορικό πακέτο λογισμικού και βρίσκεται στην έκδοση 12 (X2) τόσο για λειτουργικά συστήματα Windows, όσο και για Mac OS.

Πρόκειται για το γνωστότερο εμπορικό λογισμικό διαχείρισης βιβλιογραφικών αναφορών καθώς χρησιμοποιείται από εκατομμύρια ερευνητές, ακαδημαϊκούς, μαθητές και βιβλιοθηκονόμους ανά τον κόσμο. Μερικές από τις υπηρεσίες που προσφέρει είναι η αναζήτηση σε online βιβλιογραφικές βάσεις δεδομένων, οργάνωση των βιβλιογραφικών αναφορών, των εικόνων και των αρχείων PDF του χρήστη, καθώς και η δημιουργία λιστών βιβλιογραφικών αναφορών και εικόνων αυτόματα.

Μια συνοπτική λίστα των λειτουργιών που προσφέρονται από το EndNote παρουσιάζεται παρακάτω:

- Εισαγωγή βιβλιογραφικών αναφορών από online πηγές, αρχείο ή πληκτρολόγηση στη βάση δεδομένων του EndNote

- Επιλογή του τύπου αναφοράς (βιβλίο, άρθρο εφημερίδας/περιοδικού, συνέδριο κλπ.) και των πεδίων (συγγραφέας, τίτλος, έτος, ISBN, περίληψη κ.α.)
- Αναζήτηση δημοσιεύσεων σε βιβλιογραφικές βάσεις δεδομένων (π.χ. Web of Science, Ovid, PubMed, Βιβλιοθήκη του Αμερικανικού Κογκρέσου) στο Internet μέσα από το ίδιο το πρόγραμμα
- Οργάνωση των αναφορών, εικόνων, PDF και άλλων αρχείων
- Δυνατότητα συνεργασίας με άλλους συγγραφείς με τη χρήση του εργαλείου EndNote Web
- Ενσωμάτωση με τον κειμενογράφο Microsoft Word
- Επιλογή μορφοποίησης βιβλιογραφικών αναφορών από μια λίστα 3.600 στυλ για όλους τους επιστημονικούς τομείς (π.χ. APA, Chicago/Turabian, Harvard, MLA)
- Δυνατότητα εξαγωγής βιβλιογραφικών αναφορών σε απλά αρχεία κειμένου, πλούσιου κειμένου (RTF), HTML και XML
- Υποστήριξη μορφοποίησης OpenDocument (ODT)

2.5.2 Zotero

Το Zotero είναι ένα ελεύθερο λογισμικό ανοιχτού κώδικα (open source) το οποίο επιτρέπει στους χρήστες του τη συλλογή, διαχείριση και αναζήτηση αναφορών μέσω του browser. Αναπτύχθηκε από το Κέντρο Ιστορίας και Νέων Μέσων (Center for History and New Media) του πανεπιστημίου George Mason το 2006, έχει τη μορφή πρόσθετου (add-on) το οποίο εγκαθίσταται αποκλειστικά στον browser Firefox του ιδρύματος Mozilla και στόχος του είναι να αντικαταστήσει τα παραδοσιακά «offline» προγράμματα διαχείρισης αναφορών. Το όνομά του προέκυψε από το Αλβανικό ρήμα «zotero» που σημαίνει «διαπρέπω».

Οι λειτουργίες, που παρέχονται από το Zotero, το καθιστούν περισσότερο ένα γενικό λογισμικό διαχείρισης αναφορών από οποιαδήποτε πηγή παρά ένα εξειδικευμένο λογισμικό διαχείρισης βιβλιογραφικών αναφορών για χρήση σε άρθρα και εκθέσεις. Το Zotero είναι ιδιαίτερα χρήσιμο καθώς στις μεγαλύτερες ερευνητικές ιστοσελίδες όπως Google Scholar, Google Books, Amazon.com ακόμη και Wikipedia εντοπίζει αυτόματα πότε ένα βιβλίο, άρθρο ή άλλη πηγή προβάλλεται και με ένα κλικ του ποντικιού μπορεί να αποθηκεύσει σε κάποιο τοπικό αρχείο την πλήρη βιβλιογραφική αναφορά. Στην πε-

ρίπτωση όπου η πηγή είναι ένα online άρθρο ή ιστοσελίδα, το Zotero μπορεί επίσης να αποθηκεύσει και ένα τοπικό αντίγραφο της. Τέλος, οι χρήστες του συστήματος μπορούν να προσθέσουν τις σημειώσεις, τις ετικέτες και τα μετα-δεδομένα που αυτοί επιθυμούν μέσω της διεπαφής που παρέχεται μέσα στον browser.

Συνοψίζοντας, οι λειτουργίες που παρέχονται από το Zotero είναι:

- Αυτόματη αναγνώριση και «σύλληψη» αναφορών
- Αποθήκευση αρχείων PDF, εικόνων και ιστοσελίδων
- Δημιουργία αναφορών από το Microsoft Word και το OpenOffice
- Μεγάλο εύρος επιλογών εισαγωγής/εξαγωγής δεδομένων
- Πρόσβαση στη βιβλιοθήκη του χρήστη από οποιοδήποτε σημείο
- Χρήση χιλιάδων στυλ βιβλιογραφικών αναφορών
- Αναζήτηση στο κείμενο των PDF και των σημειώσεων του χρήστη
- Προχωρημένα εργαλεία αναζήτησης και εξόρυξης δεδομένων

2.5.3 RefWorks

Το RefWorks είναι ένα web-based εμπορικό λογισμικό διαχείρισης αναφορών το οποίο επιτρέπει στους χρήστες του τη διαχείριση αναφορών, την αναζήτηση βιβλιογραφικής πληροφορίας και τον σχεδιασμό λιστών αναφορών για χρήση σε κείμενα. Δημιουργήθηκε το 2001 από την εταιρία ProQuest και χρησιμοποιείται καθημερινά από χιλιάδες ερευνητές σε περισσότερους από 900 οργανισμούς παγκοσμίως.

Η δημοτικότητα του συστήματος οφείλεται σε ένα μεγάλο βαθμό στην υποστήριξη εκατοντάδων online βάσεων δεδομένων και στυλ βιβλιογραφικών αναφορών καθώς επίσης και στη συνεργασία με τους πιο διακεκριμένους παροχείς υπηρεσιών πληροφοριών όπως ProQuest, BioOne, EBSCO, Elsevier, HighWire, H.W. Wilson, ISI, OCLC, Ovid και Serial Solutions.

Ένας άλλος λόγος της μεγάλης απήχησης του συστήματος, είναι το γεγονός ότι πρόκειται για μία από τις σημαντικότερες web-based εφαρμογές του είδους. Έτσι, προσφέρονται σημαντικές online υπηρεσίες και δυνατότητες για τους ερευνητές όπως η χρήση των αναφορών τους από οποιοδήποτε σημείο κι αν βρίσκονται, το γεγονός ότι το σύστημα δουλεύει το ίδιο καλά σε όλα τα λειτουργικά συστήματα, η δυνατότητα διαμοιρασμού των αναφορών μεταξύ άλλων χρηστών του συστήματος κ.α.

Συνοψίζοντας τα παραπάνω σε μορφή λίστας, οι λειτουργίες που παρέχονται από το σύστημα RefWorks είναι:

- Αξιοποίηση της λίστας βιβλιογραφικών αναφορών του χρήστη από οποιοδήποτε σημείο και οποιοδήποτε λειτουργικό σύστημα (web-based εφαρμογή)
- Διαμοιρασμός αναφορών ανάμεσα σε χρήστες του συστήματος
- Δυνατότητα δημιουργίας προσωπικών βάσεων δεδομένων και χρήση σε διάφορες ερευνητικές δραστηριότητες
- Ευκολία στη συντήρηση και αναβάθμιση του συστήματος καθώς ελέγχεται κεντρικά
- Εισαγωγή αναφορών στο σύστημα από αρχεία κειμένου ή online βιβλιογραφικές βάσεις δεδομένων
- Δημιουργία βιβλιογραφιών και φορμάρισμα αυτών για χρήση σε οποιοδήποτε κείμενο
- Διαχείριση ειδοποιήσεων μέσω ενός RSS feed reader για απ' ευθείας είσοδο δεδομένων
- Λειτουργία αναζήτησης στη βάση δεδομένων των βιβλιογραφικών αναφορών του συστήματος

2.5.4 Publish or Perish

Το λογισμικό "Publish or Perish" δεν μπορεί να θεωρηθεί σύστημα διαχείρισης αναφορών, αλλά εύρεσης της απήχησης του έργου ενός ερευνητή στην παγκόσμια κοινότητα ή με άλλα λόγια λογισμικό ανάλυσης ετεροαναφορών (citation analysis software program). Το όνομα προέκυψε περί το 1950 από τον γενετιστή του Πανεπιστημίου Columbia, Dr. Kimball C. Atwood III και αναφέρεται στη διαρκή πίεση που δέχονται οι ερευνητές για νέες δημοσιεύσεις με στόχο την προώθηση ή διατήρηση της ακαδημαϊκής τους καριέρας. Αναπτύχθηκε από την ακαδημαϊκό Anne-Wil Harzing, καθηγήτρια του Πανεπιστημίου της Μελβούρνης και μέχρι σήμερα διατίθεται δωρεάν σε δύο εκδόσεις. Μία για το λειτουργικό σύστημα Windows και μία για το Linux.

Χρησιμοποιώντας δεδομένα από το Google Scholar, το Publish or Perish λαμβάνει τις «ακατέργαστες» βιβλιογραφικές αναφορές, τις αναλύει και υπολογίζει πλήθος στατιστικών και μετρικών. Ορισμένες από αυτές τις μετρικές αναφέρονται παρακάτω:

- Συνολικός αριθμός δημοσιεύσεων

Εικόνα 10. Διεπαφή του λογισμικού Publish or Perish

- Συνολικός αριθμός ετεροαναφορών
- Μέσος όρος ετεροαναφορών ανά δημοσίευση
- Μέσος όρος ετεροαναφορών ανά συγγραφέα
- Μέσος όρος δημοσιεύσεων ανά συγγραφέα
- Μέσος όρος ετεροαναφορών ανά έτος
- Δείκτης h του Hirsch [29] (δείκτης μέτρησης της απήχησης ενός ακαδημαϊκού ο οποίος συνδυάζει την ποσότητα με την ποιότητα)
- Δείκτης g του Egghe [30] (βελτίωση του δείκτη h δίνοντας μεγαλύτερη βαρύτητα στα άρθρα που έχουν πολλές ετεροαναφορές)
- Ρυθμός ετεροαναφορών σταθμισμένος με βάση το έτος (age-weighted citation rate)
- Ανάλυση του αριθμού των συγγραφέων ανά δημοσίευση
- Αντιγραφή αποτελεσμάτων στο πρόχειρο ή εξαγωγή τους σε πλήθος διαφορετικών format

2.5.5 Σύγκριση Συστημάτων Βιβλιογραφικών Αναφορών

Στην παράγραφο αυτή, επιχειρείται μια σύγκριση των συστημάτων βιβλιογραφικών αναφορών με βάση τα γενικά χαρακτηριστικά τους (Πίνακας 3), τα υποστηριζόμενα λειτουργικά συστήματα (Πίνακας 4), το format των αρχείων εξόδου και εισόδου (Πίνακας 5 και Πίνακας 6 αντίστοιχα) καθώς και τα υποστηριζόμενα στυλ βιβλιογραφίας που παρέχονται. Εδώ θα πρέπει να σημειωθεί ότι οι πληροφορίες που παρουσιάζονται είναι αυτές που κρίθηκαν πιο σημαντικές και ο αναγνώστης παραπέμπεται στις προηγούμενες παραγράφους για μια πιο πλήρη περιγραφή του κάθε συστήματος, καθώς επίσης και στις επίσημες ιστοσελίδες των προγραμμάτων αυτών [20], [22], [24], [26].

Ο Πίνακας 3 παρουσιάζει γενικές πληροφορίες σχετικά με τον δημιουργό του λογισμικού, το έτος δημιουργίας, το κόστος αγοράς, εφόσον υφίσταται και το αν είναι ανοιχτού κώδικα ή όχι. Εκεί παρατηρούμε ότι το EndNote είναι το παλαιότερο (1988) και το ακριβότερο, ενώ το Zotero, το οποίο είναι και το μοναδικό λογισμικό ανοιχτού κώδικα είναι το πλέον πρόσφατο (2006). Αξίζει να σημειωθεί ότι οι δυνατότητες που προσφέρονται από το κάθε σύστημα διαφοροποιούνται σημαντικά με αποτέλεσμα η απ' ευθείας σύγκριση τιμών να μην μπορεί να θεωρηθεί ασφαλής.

Λογισμικό	Δημιουργός	Πρώτη έκδοση	Κόστος αγοράς	Open Source
EndNote	Thomson Corporation	1988	US\$299.95	Όχι
Zotero	Center for History and New Media	2006	Δωρεάν	Ναι
RefWorks	RefWorks	2001	US\$100 / χρόνο	Όχι
Publish or Perish	Anne-Wil Harzing	-	Δωρεάν	Όχι

Πίνακας 3. Συστήματα βιβλιογραφικών αναφορών – γενικές πληροφορίες

Μια σημαντική πληροφορία θεωρείται το ποια λειτουργικά συστήματα υποστηρίζονται από το κάθε σύστημα, δεδομένου ότι ολοένα και περισσότεροι χρήστες εγκαταλείπουν το πιο διαδεδομένο λειτουργικό σύστημα (τα Windows) και μεταβαίνουν σε άλλα που καλύπτουν επαρκέστερα τις ανάγκες τους. Είναι προφανές ότι συστήματα όπως τα Zotero και RefWorks τα οποία βασίζονται στην ύπαρξη browser (του Firefox συγκεκριμένα για το Zotero), έχουν το πλεονέκτημα ότι μπορούν να χρησιμοποιηθούν σε οποιοδήποτε λειτουργικό σύστημα, σε αντίθεση με τα EndNote και Publish or Perish για τα οποία απαιτούνται ειδικές εκδόσεις.

Λογισμικό	Windows	Mac OS X	Linux	BSD	Unix
EndNote	Ναι	Ναι	Όχι	Όχι	Όχι
Zotero	Ναι	Ναι	Ναι	Ναι	Ναι
RefWorks	-	-	-	-	-
Publish or Perish	Ναι	Όχι	Ναι	Όχι	Όχι

Πίνακας 4. Συστήματα βιβλιογραφικών αναφορών – υποστηριζόμενα λειτουργικά συστήματα

Αυτό στο οποίο προσδοκά κάθε χρήστης ενός συστήματος βιβλιογραφικών αναφορών, είναι να μπορεί τόσο να εξάγει όσο και να εισάγει δεδομένα σε αυτό με τον πιο βολικό γι' αυτόν τρόπο. Έτσι, η αυτόματη εξαγωγή και εισαγωγή δεδομένων με τη χρήση ειδικά διαμορφωμένων αρχείων αποτελεί προϋπόθεση για την ευρεία αποδοχή του συστήματος. Οι παρακάτω δύο πίνακες (Πίνακας 5 και Πίνακας 6) απεικονίζουν τις δυνατότητες του κάθε λογισμικού αναφορικά με την εξαγωγή ή ανάγνωση συγκεκριμένων τύπων αρχείων αντίστοιχα. Εδώ αξίζει να σημειωθεί ότι το πρόγραμμα που υποστηρίζει τους περισσότερους τύπους αρχείων **εξόδου** είναι το EndNote ενώ αντίθετα αυτό που υποστηρίζει τους περισσότερους τύπους αρχείων **εισόδου** είναι το RefWorks.

Λογισμικό	BibTeX	Endnote/Refer/BibIX	Medline	RIS	Άλλο
EndNote	Ναι	Ναι	Ναι	Ναι	Διάφορα
Zotero	Ναι	Ναι	Όχι	Ναι	RDF, Wikipedia citation templates
RefWorks	Ναι	Όχι	Όχι	Ναι	διάφορα
Publish or Perish	Ναι	Ναι	Όχι	Ναι	CSV

Πίνακας 5. Συστήματα βιβλιογραφικών αναφορών – format αρχείων εξόδου

Λογισμικό	BibTeX	Copac	CSA	Endnote/Refer/BibIX	ISI	Medline	PubMed
EndNote	Όχι	Όχι	Ναι	Ναι	Ναι	Ναι	Ναι
Zotero	Ναι	Όχι	Όχι	Ναι	Μερικώς	Όχι	Μερικώς
RefWorks	Ναι	Ναι	Ναι	Ναι	Ναι	-	-
Publish or Perish	-	-	-	-	-	-	-

Πίνακας 6. Συστήματα βιβλιογραφικών αναφορών – format αρχείων εισόδου

Εντύπωση προκαλεί το γεγονός ότι το EndNote, το παλαιότερο κι ένα από τα δημοφιλέστερα προγράμματα της αγοράς, δεν υποστηρίζει εγγενώς τον πιο διαδεδομένο τύ-

πο αρχείου βιβλιογραφικών αναφορών, το BibTeX. Προκειμένου να καλυφθεί αυτό το κενό, έχουν αναπτυχθεί μέθοδοι κατά τις οποίες BibTeX αρχεία μπορούν να μετατραπούν σε τύπου Refer, τα οποία διαβάζονται από το EndNote, αλλά όπως είναι σαφές κάτι τέτοιο εισάγει ένα ακόμη βήμα επεξεργασίας, αυξάνοντας την πολυπλοκότητα του λογισμικού.

Σημαντικό είναι επίσης να αναφερθεί και το γεγονός ότι το λογισμικό Publish or Perish δεν υποστηρίζει τη διαδικασία εισαγωγής δεδομένων από αρχείο καθώς όπως έχει αναφερθεί και στην παράγραφο 2.5.4, το εν λόγω λογισμικό δεν προσφέρει υπηρεσίες οργάνωσης βιβλιογραφικών αναφορών, αλλά μέτρησης της απήχησης του έργου ενός ερευνητή. Παρέχοντας υποστήριξη πολλών τύπων αρχείων εξόδου δίνεται η δυνατότητα στο χρήστη να αξιοποιήσει τα αποτελέσματα του συγκεκριμένου λογισμικού σε συστήματα οργάνωσης και διαχείρισης βιβλιογραφικών αναφορών.

Τέλος, ο Πίνακας 7 παρουσιάζει στοιχεία σχετικά με την υποστήριξη των πιο διαδεδομένων συντακτικών στυλ βιβλιογραφίας. Ο αναγνώστης που δεν είναι εξοικειωμένος με τις ονομασίες των στυλ αυτών μπορεί να ανατρέξει στον πίνακα 8. Όπως γίνεται εύκολα κατανοητό, όλα τα συστήματα υποστηρίζουν όλες τις διαφορετικές μορφοποιήσεις σύνταξης βιβλιογραφιών με την εξαίρεση του Publish or Perish το οποίο δεν παρέχει αυτή τη δυνατότητα.

Λογισμικό	APA	Chicago/Turabian	Harvard	MLA	Άλλο
EndNote	Ναι	Ναι	Ναι	Ναι	Διάφορα
Zotero	Ναι	Ναι	Ναι	Ναι	Διάφορα
RefWorks	Ναι	Ναι	Ναι	Ναι	Διάφορα
Publish or Perish	-	-	-	-	-

Πίνακας 7. Συστήματα βιβλιογραφικών αναφορών – υποστηριζόμενο στυλ βιβλιογραφίας

APA	Gray, H. (1910). <i>Anatomy, descriptive and applied</i> (18th ed.). Philadelphia: Lea & Febiger.
Chicago/Turabian	Gray, Henry. <i>Anatomy, Descriptive and Applied</i> . 18th ed. Philadelphia: Lea & Febiger, 1910.
Harvard	Gray, H. 1910, <i>Anatomy, descriptive and applied</i> , 18th ed., Lea & Febiger, Philadelphia.
MLA	Gray, Henry. <i>Anatomy, Descriptive and Applied</i> . 18th ed. Philadelphia: Lea & Febiger, 1910.

Πίνακας 8. Τα πιο δημοφιλή συντακτικά στυλ βιβλιογραφίας

2.5.6 Σύστημα Εύρεσης και Διαχείρισης Βιβλιογραφικών Αναφορών

Το σύστημα αυτό εκπονήθηκε στα πλαίσια της πτυχιακής εργασίας της προπτυχιακής φοιτήτριας του Τμήματος Πληροφορικής της Σχολής Θετικών Επιστημών του Αριστοτελείου Πανεπιστημίου Θεσσαλονίκης, Εύας Ζιάκα. Παρέχει τη δυνατότητα υποβολής ερωτημάτων στο Google Scholar και ταξινόμησης των ληφθέντων αποτελεσμάτων και αποτελεί σημαντική απόδειξη του ότι ένα τέτοιο σύστημα είναι εφικτό.

Παρόλα αυτά, δεν μπορεί να θεωρηθεί συγκρίσιμο των EndNote, Zotero, RefWorks και Publish or Perish καθώς:

- πρόκειται για μία desktop εφαρμογή, της οποίας η ανάπτυξη έχει γίνει σε Java με Microsoft Access βάση δεδομένων
- μπορεί να εκτελεστεί αποκλειστικά στο λειτουργικό σύστημα Windows
- δεν υποστηρίζει την εισαγωγή ή εξαγωγή δεδομένων βιβλιογραφικών αναφορών μέσω κάποιου αρχείου παρά μόνο χειροκίνητα
- αποτελεί περισσότερο ένα σύστημα το οποίο αναπτύχθηκε με στόχο την «απόδειξη της ιδέας» (proof of concept) παρά ένα λογισμικό προς χρήση από την ενδιαφερόμενη κοινότητα – ως αποτέλεσμα δεν ακολουθεί πάντα τους διαδεδομένους κανόνες χρηστικότητας των γραφικών διεπαφών, είναι αρκετά δύσχρηστο και γενικά απέχει κατά πολύ από το να χαρακτηριστεί "πρακτικά χρήσιμη" εφαρμογή
- στην περίοδο που δοκιμάστηκε, ο μηχανισμός λήψης δεδομένων από το Google Scholar είχε τεθεί μόνιμα εκτός λειτουργίας εξαιτίας μεταβολών στον τρόπο έκδοσης των αποτελεσμάτων αναζήτησης από το Google Scholar.

3 Ανάλυση Απαιτήσεων

Στο προηγούμενο κεφάλαιο αναλύθηκε η τρέχουσα κατάσταση στον τομέα της διαχείρισης βιβλιογραφικών αναφορών και των ετεροαναφορών σε αυτές. Αυτό που εύκολα παρατηρεί κανείς είναι η απουσία ενός συστήματος το οποίο θα συγκεντρώνει τόσο δυνατότητες διαχείρισης βιβλιογραφικών αναφορών για χρήση σε βιβλιογραφία πονημάτων, βιογραφικό κ.α. όσο και δυνατότητες διατήρησης στοιχείων απήχησης του δημοσιευμένου έργου ενός ακαδημαϊκού, όπως π.χ. ο αριθμός των ετεροαναφορών (citations) που έχει κάποια δημοσίευσή του. Το σύστημα που αναπτύχθηκε έχει σαν στόχο να καλύψει το προαναφερθέν κενό, προσφέροντας υπηρεσίες και στα δύο πεδία, δηλαδή οργάνωση των βιβλιογραφικών αναφορών ενός ακαδημαϊκού και παρακολούθηση των ετεροαναφορών σε αυτές.

Σε αυτό το κεφάλαιο θα πραγματοποιηθεί ανάλυση των απαιτήσεων του συστήματος. Αυτές διαχωρίζονται σε λειτουργικές και σχετίζονται με τις υπηρεσίες που θα πρέπει να παρέχονται από το σύστημα (παράγραφος 3.1) και μη λειτουργικές, οι οποίες αναφέρονται κυρίως στους περιορισμούς και τις απαιτήσεις εκείνες που θα καθιστούσαν το σύστημα αποδεκτό από τους χρήστες (παράγραφος 3.2).

3.1 Λειτουργικές Απαιτήσεις

Οι απαιτήσεις των χρηστών από ένα σύστημα πολλές φορές είναι ποικίλες, συχνά αντικρουόμενες και ασαφείς. Στην περίπτωση του συγκεκριμένου συστήματος διαχείρισης βιβλιογραφικών αναφορών, το πλήθος των απαιτήσεων περιορίστηκε από το γεγονός ότι σχεδιάστηκε για ένα συγκεκριμένο τύπο χρηστών. Ο τύπος αυτός είναι οι ερευνητές ή ακαδημαϊκοί, οι οποίοι διαθέτουν επιστημονικό συγγραφικό έργο και έχουν την ανάγκη να οργανώσουν την πληροφορία αυτή με τρόπο που να καθίσταται πιο εύκολη η αξιοποίησή της. Στη συνέχεια της παραγράφου παρουσιάζονται οι λειτουργικές απαιτήσεις (functional requirements) του συστήματος χωρισμένες σε γενικές απαιτήσεις, απαιτήσεις σχετικές με τη διαχείριση προσωπικών δημοσιεύσεων, τη διαχείριση ετεροαναφορών και την προβολή αναφορών εκτύπωσης.

3.1.1 Γενικές

Οι γενικές απαιτήσεις περιγράφουν το γενικότερο πλαίσιο λειτουργίας του συστήματος και σύμφωνα με αυτές το σύστημα θα πρέπει:

1. να είναι ανεξάρτητο λειτουργικού συστήματος και προσβάσιμο από οποιοδήποτε σημείο (web-based εφαρμογή)
2. να παρέχει μια υποδομή δημιουργίας λογαριασμών και αναγνώρισης χρηστών με τη χρήση κωδικού
3. να παρέχει στο χρήστη δυνατότητα αναζήτησης βιβλιογραφικών αναφορών μέσω της εξειδικευμένης μηχανής αναζήτησης Google Scholar
4. να επιτρέπει τη διαχείριση εγγραφών προσωπικών δημοσιεύσεων (publications)
5. να επιτρέπει τη διαχείριση εγγραφών αναφορών σε προσωπικές δημοσιεύσεις (citations)
6. να παρουσιάζει λίστες των δημοσιεύσεων και των ετεροαναφορών σε μορφή κατάλληλη για χρήση σε βιβλιογραφία πονημάτων, βιογραφικό κ.α. (reports)
7. να βασίζεται σε πακέτα λογισμικού τα οποία είναι δωρεάν και κατά προτίμηση ανοιχτού κώδικα καθώς με αυτόν τον τρόπο δε θα είναι απαραίτητη η χρέωση των χρηστών για την παροχή της υπηρεσίας

3.1.2 Διαχείρισης αποτελεσμάτων αναζήτησης

Ένα από τα σημαντικότερα τμήματα του συστήματος είναι η αναζήτηση αποτελεσμάτων στην πληρέστερη μηχανή αναζήτησης επιστημονικών συγγραμμάτων (Google Scholar) και η παροχή της δυνατότητας επεξεργασίας αυτών. Για την παροχή της λειτουργικότητας αυτής το σύστημα θα πρέπει:

1. να εμφανίζει τα αποτελέσματα της αναζήτησης με τρόπο φιλικό προς τον χρήστη
2. να δίνει τη δυνατότητα κατηγοριοποίησης των αποτελεσμάτων σε «Ιδίες δημοσιεύσεις», «Πραγματικές αναφορές», «Ψευδείς αναφορές» και «Άγνωστο»
3. να προτείνει μια πιθανή κατηγοριοποίηση, η οποία θα βοηθά τον χρήστη στη λήψη της τελικής απόφασης
4. να δίνει τη δυνατότητα διαγραφής αποτελεσμάτων αναζήτησης
5. να επιτρέπει την αναίρεση της κατηγοριοποίησης ή της διαγραφής

6. να «θυμάται» τις αλλαγές του χρήστη προκειμένου αυτός να είναι σε θέση να συνεχίσει την επεξεργασία από το σημείο στο οποίο σταμάτησε
7. να μην εμφανίζει αποτελέσματα τα οποία έχουν ήδη κατηγοριοποιηθεί σε επόμενη αναζήτηση
8. να επιτρέπει την εισαγωγή, ενημέρωση και διαγραφή σχολίων για κάθε αποτέλεσμα ξεχωριστά
9. να παρέχει τη δυνατότητα απ' ευθείας δημιουργίας εγγραφής «Πραγματικής αναφοράς» (True citation) με χρήση στοιχείων που υπάρχουν στα αποτελέσματα

3.1.3 Διαχείρισης Δημοσιεύσεων

Με δεδομένο ότι το σύστημα θα διαχειρίζεται αναφορές αποκλειστικά σε δημοσιεύσεις (Publications) του ίδιου του χρήστη, γίνεται εύκολα κατανοητή η σημασία της παροχής ενός άρτιου μηχανισμού διαχείρισης αυτών. Έτσι, το σύστημα θα πρέπει:

1. να επιτρέπει τη δημιουργία/ενημέρωση/διαγραφή εγγραφής προσωπικής δημοσίευσης
2. να διαθέτει όλα τα απαραίτητα πεδία για την εισαγωγή της πληροφορίας που ορίζονται στις βιβλιογραφικές αναφορές (τίτλος, συγγραφείς, συνέδριο / περιοδικό / βιβλίο δημοσίευσης, σελίδες, τεύχος, μήνας, έτος, πόλη κλπ.)
3. να παρέχει τη δυνατότητα στον χρήστη να κατηγοριοποιήσει τις εγγραφές του σύμφωνα με τον τύπο της δημοσίευσης (Journal, Conference Proceedings, Workshop Proceedings, Book, Book chapter κλπ.)
4. να παρέχει τη δυνατότητα προσωπικής ταξινόμησης των εγγραφών
5. να δίνει τη δυνατότητα στο χρήστη να εισάγει με αυτόματο τρόπο πλήθος εγγραφών προσωπικών δημοσιεύσεων με τη χρήση BibTeX αρχείου
6. να επιτρέπει την προβολή του BibTeX αρχείου εφόσον αυτό βρίσκεται ήδη στον διακομιστή χωρίς την ανάγκη για εκ νέου φόρτωση (upload)

3.1.4 Διαχείρισης Ετεροαναφορών

Η αρχική ανάγκη που οδήγησε στη σύλληψη του συστήματος, αφορούσε στην εύρεση και συγκέντρωση των ετεροαναφορών (citations) και στην καταγραφή των σχέσεών τους με τις δημοσιεύσεις του χρήστη. Στο συγκεκριμένο λοιπόν, κομμάτι του συστήματος είναι απαραίτητο:

1. να επιτρέπεται η δημιουργία/ενημέρωση/διαγραφή ετεροαναφοράς
2. να διατίθενται όλα τα απαραίτητα πεδία για την εισαγωγή της πληροφορίας που συναντάται στις βιβλιογραφικές αναφορές (τίτλος, συγγραφείς, συνέδριο / περιοδικό / βιβλίο δημοσίευσης, σελίδες, τεύχος, μήνας, έτος, πόλη κλπ.)
3. να παρέχεται η δυνατότητα κατηγοριοποίησης των εγγραφών σύμφωνα με τον τύπο της δημοσίευσης (Journal, Conference Proceedings, Workshop Proceedings, Book, Book chapter κλπ.)
4. να παρέχεται η δυνατότητα προσωπικής ταξινόμησης των εγγραφών
5. να δίνεται η δυνατότητα στο χρήστη να εισάγει με αυτόματο τρόπο πλήθος εγγραφών προσωπικών δημοσιεύσεων με τη χρήση BibTeX αρχείου
6. να επιτρέπεται η προβολή του BibTeX αρχείου εφόσον αυτό βρίσκεται ήδη στον διακομιστή χωρίς την ανάγκη για εκ νέου φόρτωση (upload)

3.1.5 Δημιουργίας Αναφορών

Για την αξιοποίηση των δεδομένων που δημιουργεί και συντηρεί ο κάθε χρήστης θα πρέπει το σύστημα να παράγει έναν αριθμό αναφορών (Reports). Πιο συγκεκριμένα θα πρέπει:

1. να προβάλλεται η λίστα με τις δημοσιεύσεις του χρήστη ανά έτος με φθίνουσα ταξινόμηση ώστε να μπορεί να αντιγραφεί και να επικολληθεί σε κάποιον κειμενογράφο
2. να προβάλλεται η λίστα των δημοσιεύσεων του χρήστη κατηγοριοποιημένες σύμφωνα με τον τύπο των δημοσιεύσεων (βλ. 3.1.3, σημείο 3) και ταξινομημένη με βάση το αριθμό προσωπικής ταξινόμησης (βλ. 3.1.3, σημείο 4)
3. να προβάλλεται η λίστα των δημοσιεύσεων μαζί με τον αριθμό των ετεροαναφορών που έχει η κάθε μία, ταξινομημένη με φθίνουσα σειρά πλήθους ετεροαναφορών

3.2 Μη Λειτουργικές Απαιτήσεις

Ενώ οι λειτουργικές απαιτήσεις προσδίδουν στο σύστημα τον χαρακτήρα που αυτό θα έχει καθώς καλύπτουν τις ανάγκες των χρηστών, οι μη λειτουργικές είναι αυτές οι ο-

ποιές κρίνουν την αποδοχή ή όχι του συστήματος από τους χρήστες για τους οποίους δημιουργήθηκε.

3.2.1 Απαιτήσεις Απόδοσης

Σε θέματα απόδοσης (Performance) του συστήματος είναι σημαντικό:

1. να έχει μικρό χρόνο απόκρισης στη φόρτωση και παρουσίαση των δεδομένων, όπου αυτό είναι εφικτό και όπου δεν είναι να παρουσιάζει μήνυμα αναμονής στον χρήστη
2. να διατηρεί τις εγγραφές μέχρι ο χρήστης να τις ταξινομήσει ή διαγράψει

3.2.2 Απαιτήσεις Ασφάλειας

Σε θέματα ασφαλείας το σύστημα θα πρέπει:

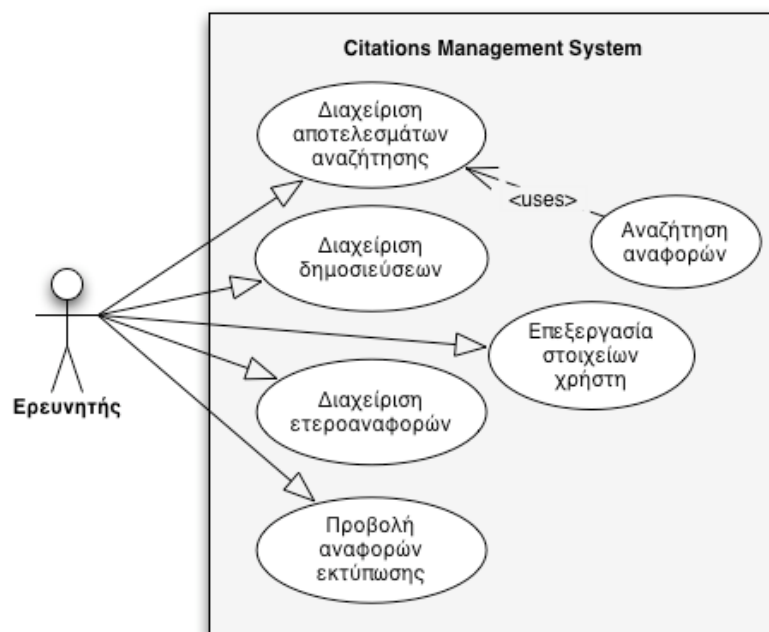
1. να απαγορεύει την προβολή και τροποποίηση των εγγραφών όταν ο χρήστης δεν έχει πιστοποιηθεί με τον προσωπικό κωδικό του
2. να προειδοποιεί τον χρήστη όταν αυτός επιχειρεί τροποποίηση/διαγραφή εγγραφών, οι οποίες συνδέονται με άλλες
3. να εξασφαλίζει ότι τα προσωπικά δεδομένα θα είναι διαθέσιμα αποκλειστικά στον χρήστη στον οποίο ανήκουν
4. να χρησιμοποιεί συστήματα κρυπτογράφησης στην αποθήκευση ευαίσθητων δεδομένων π.χ. κωδικών χρήσης
5. να τηρεί αρχείο καταγραφής συμβάντων για την αντιμετώπιση κακόβουλων ενεργειών

4 Σχεδίαση Συστήματος

Στο κεφάλαιο 3 δόθηκε μια αναλυτική περιγραφή των απαιτήσεων των χρηστών από το σύστημα διαχείρισης αναφορών ενώ σε αυτό το κεφάλαιο θα παρουσιαστεί η σχεδίαση του συστήματος σε συνδυασμό με τις τεχνικές αποφάσεις που ελήφθησαν. Αρχικά θα γίνει η περιγραφή των απαιτήσεων με τη χρήση διαγραμμάτων περιπτώσεων χρήσης της UML. Στη συνέχεια θα παρουσιαστεί το σχήμα της βάσης δεδομένων που δημιουργήθηκε για την αποθήκευση των δεδομένων.

4.1 Διάγραμμα Περιπτώσεων Χρήσης

Στόχος των διαγραμμάτων περιπτώσεων χρήσης (Use Case Model) είναι η απεικόνιση και κατανόηση της λειτουργικότητας, συμπεριφοράς και αλληλεπίδρασης του συστήματος με τους εξωτερικούς χρήστες. Με άλλα λόγια περιγράφουν τον τρόπο με τον οποίο θα συμπεριφέρεται το σύστημα προκειμένου να ικανοποιεί τις απαιτήσεις των χρηστών και απεικονίζουν ουσιαστικά την απαιτούμενη συμπεριφορά του συστήματος από την οπτική του χρήστη.



Εικόνα 11. Διάγραμμα περιπτώσεων χρήσης

Στην Εικόνα 11 παρουσιάζεται το διάγραμμα των υπηρεσιών που θα προσφέρονται από το σύστημα. Οι υπηρεσίες αυτές είναι:

1. Διαχείριση αποτελεσμάτων αναζήτησης
2. Διαχείριση δημοσιεύσεων χρήστη
3. Διαχείριση ετεροαναφορών
4. Προβολή αναφορών εκτύπωσης
5. Επεξεργασία στοιχείων χρήστη

Στην παράγραφο 4.2 παρουσιάζονται τα διαγράμματα δραστηριοτήτων της κάθε μιας από τις παραπάνω υπηρεσίες.

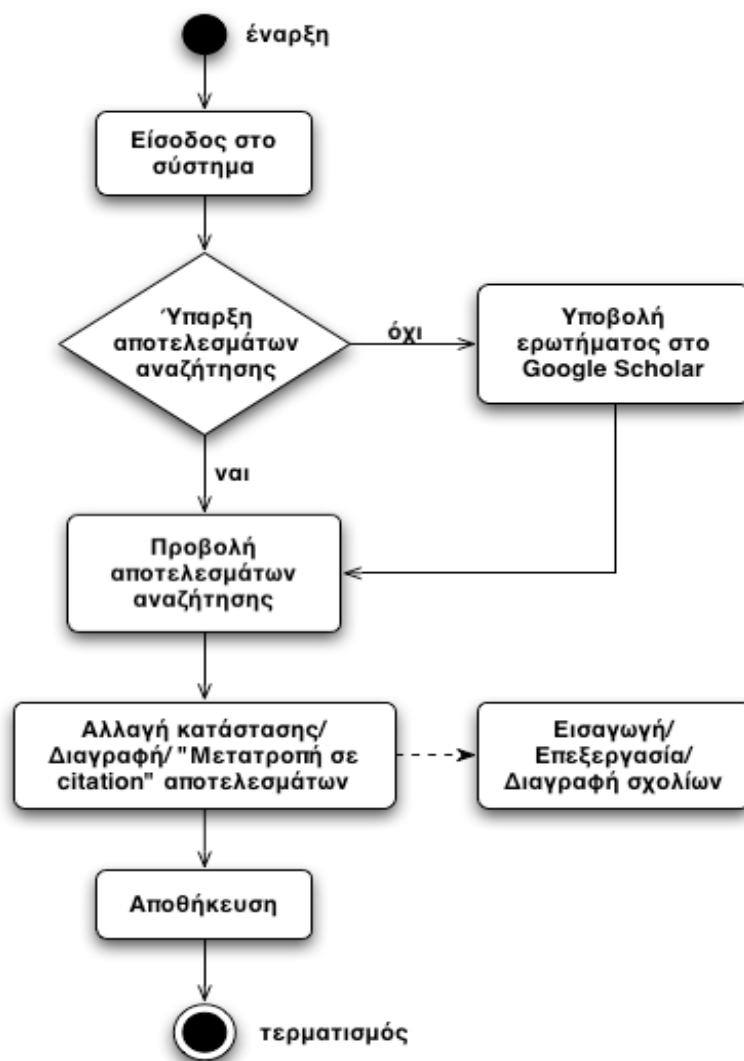
4.2 Διαγράμματα Δραστηριοτήτων

Τα Διαγράμματα Δραστηριοτήτων (Activity Diagrams) δείχνουν με αφηρημένο τρόπο το τι μπορεί να κάνει ο χρήστης του εν λόγω συστήματος.

4.2.1 Διαχείριση αποτελεσμάτων αναζήτησης

Τα αποτελέσματα αναζήτησης προκύπτουν από την υποβολή ερωτήματος στο Google Scholar κι αφορούν αποκλειστικά σε επιστημονικές δημοσιεύσεις. Το προτεινόμενο «κλειδί» αναζήτησης είναι το ονοματεπώνυμο το οποίο χρησιμοποιεί ο ερευνητής στις δημοσιεύσεις του, δεδομένου ότι σκοπός της αναζήτησης είναι να βρεθούν οι αναφορές που γίνονται σε δημοσιεύσεις του ιδίου.

Στη συνέχεια παρουσιάζεται το διάγραμμα δραστηριοτήτων της συγκεκριμένης υπηρεσίας (Εικόνα 12) καθώς και η επεξήγηση αυτού (Πίνακας 9).



Εικόνα 12. Διάγραμμα δραστηριοτήτων υπηρεσίας Διαχείρισης αποτελεσμάτων αναζήτησης

Τίτλος: Διαχείριση αποτελεσμάτων αναζήτησης		ID: 1	Επίπεδο σπουδαιότητας: Υψηλό
Αποδέκτης (Actor): Ερευνητής ή ακαδημαϊκός ο οποίος έχει παρουσιάσει ερευνητικό έργο, απόρροια του οποίου είναι ένας αριθμός δημοσιεύσεων		Τύπος υπηρεσίας: Λεπτομερής, πραγματική	
Περιγραφή: Η υπηρεσία αυτή εκτελείται όταν ερευνητής ή ακαδημαϊκός, χρήστης του συστήματος, επιθυμεί να διαχειριστεί τα αποτελέσματα αναζήτησης αναφορών σε επιστημονικά του συγγράμματα.			
Εκροή: Κατηγοριοποίηση αποτελεσμάτων αναζήτησης			
Εμπλεκόμενοι/ρόλοι (Stakeholders):			
A/A	Εμπλεκόμενος	Περιγραφή	Ρόλος
1.	Ερευνητής ή ακαδημαϊκός	Επιθυμεί την εύρεση αναφορών για τις δημοσιεύσεις του.	Χρήστης συστήματος
2.	ΠΣ	Προσφέρει τη διεπαφή για αναζήτηση και κατηγοριοποίηση των βιβλιογραφικών αναφορών του χρήστη.	Πληροφοριακό σύστημα

<p>Αφορμή/Περιοδικότητα (Trigger): Η εκτέλεση της υπηρεσίας διαχείρισης αποτελεσμάτων πραγματοποιείται κατόπιν επιθυμίας του χρήστη στη χρονική στιγμή που εκείνος επιθυμεί και μπορεί να επαναλαμβάνεται οποτεδήποτε αυτός θέλει. Η αναζήτηση των αναφορών γίνεται επίσης κατά βούληση με προτεινόμενη περιοδικότητα 1-2 μήνες.</p>
<p>Σχέσεις: Συσχέτιση (Association): Περίληψη (Include): Επέκταση (Extend): Αναζήτηση αναφορών Γενίκευση (Generalization):</p>
<p>Κανονική ροή (Normal flow):</p> <ol style="list-style-type: none"> 1. Ο Ερευνητής ή ακαδημαϊκός χρησιμοποιεί το όνομα χρήστη και το συνθηματικό του για να πιστοποιηθεί από το σύστημα. Σε περίπτωση που δεν έχει δημιουργήσει όνομα χρήστη και συνθηματικό για το σύστημα εκτελείται η υπό-ροή S-1. 2. Προβολή αποτελεσμάτων αναζήτησης. 3. Επιλογή κατηγορίας αποτελεσμάτων εκ των «New Results», «Own Publications», «True Citation», «False Citations», «Unknown». 4. Ανάθεση διαφορετικής κατηγορίας από την τρέχουσα. 5. Αποθήκευση αλλαγής.
<p>Υπό-ροές (Subflows): S-1: Εγγραφή στο σύστημα και δημιουργία λογαριασμού</p> <ol style="list-style-type: none"> 1. Ο Ερευνητής ή ακαδημαϊκός εκτελεί τη λειτουργία δημιουργίας λογαριασμού. 2. Συμπληρώνει τα στοιχεία που ζητούνται στην φόρμα που εμφανίζεται και την υποβάλλει. 3. Το όνομα χρήστη και ο κωδικός έχουν δημιουργηθεί και μπορούν να χρησιμοποιηθούν για την είσοδο στο σύστημα.
<p>Ροές κατά εξαίρεση (Exceptional flows):</p> <p>2α. Ο χρήστης χρησιμοποιεί την υπηρεσία για πρώτη φορά, οπότε δεν υπάρχουν διαθέσιμα αποτελέσματα αναζήτησης και υποβάλλει ερώτημα στο Google Scholar προκειμένου να του επιστραφούν κάποια.</p> <p>3α. Ο χρήστης χρησιμοποιεί τη λειτουργία εισαγωγής σχολίου (<i>Add a comment</i>) και μετά επιλέγει αποθήκευση (<i>Save</i>).</p> <p>4α. Ο χρήστης επιλέγει τη διαγραφή του αποτελέσματος επιλέγοντας τη λειτουργία «Delete Record!».</p> <p>4β. Ο χρήστης δημιουργεί εγγραφή ετεροαναφοράς, επιλέγοντας τη λειτουργία «Save as citation».</p>
<p>Προαπαιτούμενα (Preconditions):</p> <ol style="list-style-type: none"> 1. Ο Ερευνητής ή ακαδημαϊκός να διαθέτει δημοσιεύσεις στις οποίες να υπάρχουν αναφορές τρίτων.
<p>Ισχύοντα μετά την εκτέλεση/Αποτέλεσμα (Post-conditions): Ο ερευνητής ή ακαδημαϊκός, διαθέτει μια κατηγοριοποιημένη λίστα αποτελεσμάτων του Google Scholar σχετικά με τις δημοσιεύσεις του.</p>

Πίνακας 9. Περιγραφή υπηρεσίας διαχείρισης αποτελεσμάτων αναζήτησης

4.2.2 Διαχείριση δημοσιεύσεων χρήστη

Η υπηρεσία διαχείρισης δημοσιεύσεων χρήστη επιτρέπει τη δημιουργία, επεξεργασία, διαγραφή και εισαγωγή μέσω αρχείου BibTeX, των δημοσιεύσεων του χρήστη. Επιπλέον, η συγκεκριμένη υπηρεσία παρέχει πλήθος προβολών, όπως: Προβολή λίστας, Προβολή λεπτομερειών, Προβολή επεξεργασίας κ.α.

Στη συνέχεια παρουσιάζεται το διάγραμμα δραστηριοτήτων της υπηρεσίας (Εικόνα 13) καθώς και η επεξήγηση αυτού (Πίνακας 10).



Εικόνα 13. Διάγραμμα δραστηριοτήτων υπηρεσίας Διαχείρισης δημοσιεύσεων χρήστη

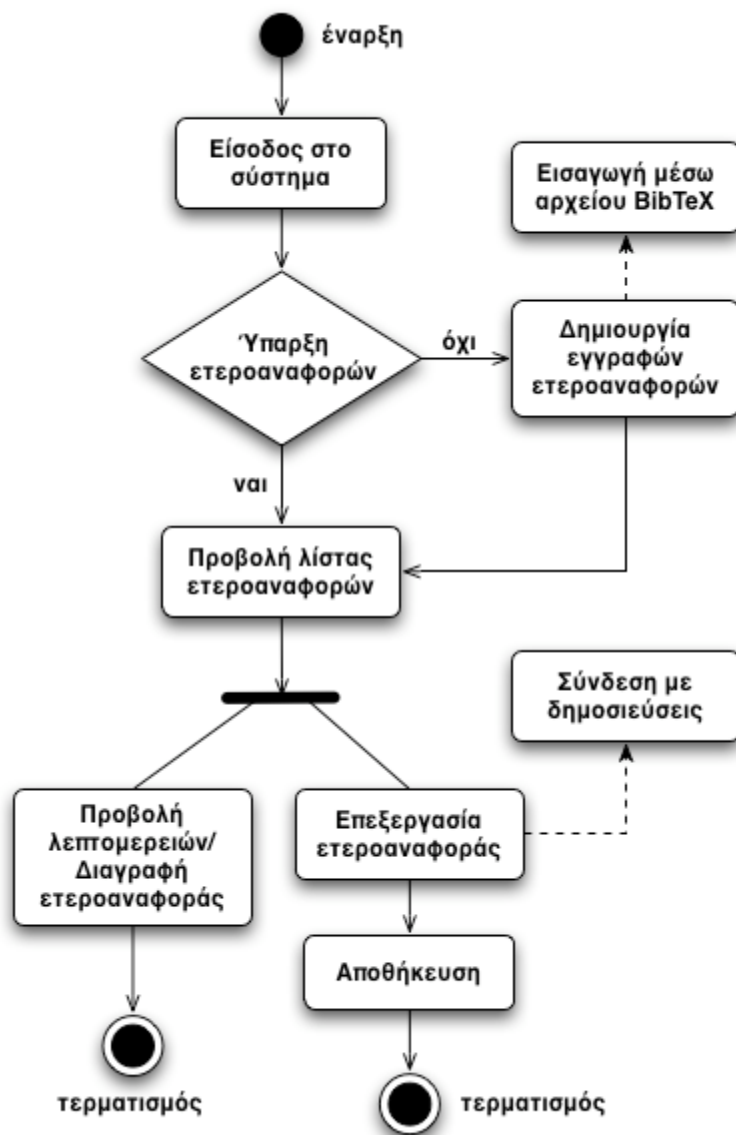
Τίτλος: Διαχείριση δημοσιεύσεων χρήστη	ID: 2	Επίπεδο σπουδαιότητας: Υψηλό	
Αποδέκτης (Actor): Ερευνητής ή ακαδημαϊκός ο οποίος έχει παρουσιάσει ερευνητικό έργο, απόρροια του οποίου είναι ένας αριθμός δημοσιεύσεων.		Τύπος υπηρεσίας: Λεπτομερής, πραγματική	
Περιγραφή: Η υπηρεσία αυτή εκτελείται όταν ερευνητής ή ακαδημαϊκός, χρήστης του συστήματος, επιθυμεί τη δημιουργία εγγραφών βιβλιογραφικών αναφορών σχετικές με τις δημοσιεύσεις του.			
Εκροή: Λίστα βιβλιογραφικών αναφορών δημοσιεύσεων του χρήστη.			
Εμπλεκόμενοι/ρόλοι (Stakeholders):			
A/A	Εμπλεκόμενος	Περιγραφή	Ρόλος
1.	Ερευνητής ή ακαδημαϊκός	Επιθυμεί τη δημιουργία βιβλιογραφικών αναφορών για τις δημοσιεύσεις του.	Χρήστης συστήματος
2.	ΠΣ	Προσφέρει τη διεπαφή για δημιουργία, επεξεργασία και διαγραφή των βιβλιογραφικών αναφορών του χρήστη.	Πληροφοριακό σύστημα
Αφορμή/Περιοδικότητα (Trigger): Η εκτέλεση της υπηρεσίας διαχείρισης δημοσιεύσεων πραγματοποιείται κατόπιν επιθυμίας του χρήστη στη χρονική στιγμή που εκείνος επιθυμεί και μπορεί να επαναλαμβάνεται οποτεδήποτε προκύπτουν νέες δημοσιεύσεις.			
Σχέσεις: Συσχέτιση (Association): Περίληψη (Include): Επέκταση (Extend): Εισαγωγή εγγραφών μέσω αρχείου BibTeX Γενίκευση (Generalization):			
Κανονική ροή (Normal flow):			
<ol style="list-style-type: none"> Ο Ερευνητής ή ακαδημαϊκός χρησιμοποιεί το όνομα χρήστη και το συνθηματικό του για να πιστοποιηθεί από το σύστημα. Σε περίπτωση που δεν έχει δημιουργήσει όνομα χρήστη και συνθηματικό για το σύστημα εκτελείται η υπό-ροή S-1. Ο Ερευνητής ή ακαδημαϊκός επιλέγει την προβολή της λίστας δημοσιεύσεων. Σε περίπτωση που η λίστα δημοσιεύσεων είναι κενή εκτελείται η υπό-ροή S-2 για τη δημιουργία δημοσίευσης με τη χρήση φόρμας ή η υπό-ροή S-3 για την εισαγωγή δημοσιεύσεων από αρχείο BibTeX. Επιλογή επεξεργασίας δημοσίευσης. Σε περίπτωση που επιθυμείται η προβολή ή διαγραφή της δημοσίευσης επιλέγεται η υπό-ροή S-4 ή S-5 αντίστοιχα. Τροποποίηση στοιχείων. Αποθήκευση αλλαγών. 			

<p>Υπό-ροές (Subflows):</p> <p>S-1: Εγγραφή στο σύστημα και δημιουργία λογαριασμού</p> <ol style="list-style-type: none"> 1. Ο χρήστης εκτελεί τη λειτουργία δημιουργίας λογαριασμού. 2. Συμπληρώνει τα στοιχεία που ζητούνται στην φόρμα που εμφανίζεται και την υποβάλλει. 3. Το όνομα χρήστη και ο κωδικός έχουν δημιουργηθεί και μπορούν να χρησιμοποιηθούν για την είσοδο στο σύστημα. <p>S-2: Δημιουργία εγγραφών δημοσιεύσεων</p> <ol style="list-style-type: none"> 1. Ο χρήστης εκτελεί τη λειτουργία δημιουργίας εγγραφής δημοσίευσης. 2. Συμπληρώνει τα στοιχεία της φόρμας που εμφανίζεται και την υποβάλλει. <p>S-3: Εισαγωγή δημοσιεύσεων από αρχείο BibTeX</p> <ol style="list-style-type: none"> 1. Ο χρήστης εκτελεί τη λειτουργία φόρτωσης αρχείου BibTeX. 2. Στη λίστα με τις εγγραφές που βρέθηκαν μέσα στο αρχείο, επιλέγει αυτές που επιθυμεί να εισάγει. 3. Ενεργοποιεί τη διαδικασία εισαγωγής. <p>S-4: Προβολή λεπτομερειών δημοσιεύσεων</p> <ol style="list-style-type: none"> 1. Ο χρήστης εκτελεί τη λειτουργία προβολής λεπτομερειών επιλέγοντας τη λειτουργία «Details». 2. Εμφανίζεται ένας πίνακας με όλα τα στοιχεία της δημοσίευσης που έχουν εισαχθεί. <p>S-5: Διαγραφή δημοσίευσης</p> <ol style="list-style-type: none"> 1. Ο χρήστης επιλέγει τη λειτουργία «Delete». 2. Απαντάει θετικά στο μήνυμα επιβεβαίωσης. 3. Εμφανίζεται η λίστα με τις υπόλοιπες δημοσιεύσεις.
<p>Ροές κατά εξαίρεση (Exceptional flows):</p> <p>S-3α. Ο χρήστης έχει ήδη φορτώσει κάποιο αρχείο μορφής BibTeX και επιθυμεί να εισάγει κάποιες από τις εγγραφές που υπάρχουν σε αυτό και επιλέγει τη λειτουργία «View uploaded file entries».</p>
<p>Προαπαιτούμενα (Preconditions):</p> <ol style="list-style-type: none"> 1. Ο Χρήστης να διαθέτει δημοσιεύσεις σε περιοδικά, συνέδρια, βιβλία κλπ.
<p>Ισχύοντα μετά την εκτέλεση/Αποτέλεσμα (Post-conditions): Ο ερευνητής ή ακαδημαϊκός, διαθέτει μια λίστα με τις δημοσιεύσεις του, τις οποίες θα μπορεί στη συνέχεια να συνδέσει με ετεροαναφορές (citations).</p>

Πίνακας 10. Περιγραφή υπηρεσίας διαχείρισης δημοσιεύσεων

4.2.3 Διαχείριση ετεροαναφορών (citations)

Η υπηρεσία διαχείρισης ετεροαναφορών αποτελεί μία από τις σημαντικότερες λειτουργίες του συστήματος καθώς χάρη σε αυτήν δημιουργούνται οι συσχετίσεις ανάμεσα στις δημοσιεύσεις και τις ετεροαναφορές καθιστώντας τη δημιουργία των αναφορών εκτύπωσης δυνατή.



Εικόνα 14. Διάγραμμα δραστηριοτήτων υπηρεσίας Διαχείρισης ετεροαναφορών

Ο Πίνακας 11 παρουσιάζει την επεξήγηση του διαγράμματος δραστηριοτήτων της υπηρεσίας, όπως αυτό εμφανίζεται στην Εικόνα 16.

Τίτλος: Διαχείριση ετεροαναφορών χρήστη	ID: 3	Επίπεδο σπουδαιότητας: Υψηλό
Αποδέκτης (Actor): Ερευνητής ή ακαδημαϊκός ο οποίος έχει παρουσιάσει ερευνητικό έργο, απόρροια του οποίου είναι ένας αριθμός δημοσιεύσεων.		Τύπος υπηρεσίας: Λεπτομερής, πραγματική
Περιγραφή: Η υπηρεσία αυτή εκτελείται όταν ερευνητής ή ακαδημαϊκός, χρήστης του συστήματος, επιθυμεί τη δημιουργία εγγραφών ετεροαναφορών προς μία ή περισσότερες δημοσιεύσεις του.		
Εκροή: Λίστα ετεροαναφορών συνδεδεμένων με τις δημοσιεύσεις του χρήστη.		

Εμπλεκόμενοι/ρόλοι (Stakeholders):

A/A	Εμπλεκόμενος	Περιγραφή	Ρόλος
1.	Ερευνητής ή ακαδημαϊκός	Επιθυμεί τη δημιουργία ετεροαναφορών για τις δημοσιεύσεις του.	Χρήστης συστήματος
2.	ΠΣ	Προσφέρει τη διεπαφή για δημιουργία, επεξεργασία και διαγραφή των ετεροαναφορών σε δημοσιεύσεις του χρήστη.	Πληροφοριακό σύστημα

Αφορμή/Περιοδικότητα (Trigger):

Η εκτέλεση της υπηρεσίας διαχείρισης ετεροαναφορών πραγματοποιείται κατόπιν επιθυμίας του χρήστη στη χρονική στιγμή που εκείνος επιθυμεί και επαναλαμβάνεται συνήθως όταν προκύψει νέα ετεροαναφορά από την επεξεργασία των αποτελεσμάτων αναζήτησης ή από κάποια άλλη πηγή.

Σχέσεις:

Συσχέτιση (Association):

Περίληψη (Include):

Επέκταση (Extend): Εισαγωγή εγγραφών μέσω αρχείου BibTeX

Γενίκευση (Generalization):

Κανονική ροή (Normal flow):

1. Ο Ερευνητής ή ακαδημαϊκός χρησιμοποιεί το όνομα χρήστη και το συνθηματικό του για να πιστοποιηθεί από το σύστημα.
Σε περίπτωση που δεν έχει δημιουργήσει όνομα χρήστη και συνθηματικό για το σύστημα εκτελείται η υπό-ροή S-1.
2. Ο Ερευνητής ή ακαδημαϊκός επιλέγει την προβολή της λίστας ετεροαναφορών.
Σε περίπτωση που η λίστα ετεροαναφορών είναι κενή εκτελείται η υπό-ροή S-2 για τη δημιουργία ετεροαναφοράς με τη χρήση φόρμας ή η υπό-ροή S-3 για την εισαγωγή ετεροαναφορών από αρχείο BibTeX.
3. Επιλογή επεξεργασίας ετεροαναφοράς.
Σε περίπτωση που επιθυμείται η προβολή ή διαγραφή της ετεροαναφοράς επιλέγεται η υπό-ροή S-4 ή S-5 αντίστοιχα.
4. Τροποποίηση στοιχείων και επιλογή των δημοσιεύσεων που αναφέρονται.
5. Αποθήκευση αλλαγών.

Υπό-ροές (Subflows):**S-1:** Εγγραφή στο σύστημα και δημιουργία λογαριασμού

1. Ο χρήστης εκτελεί τη λειτουργία δημιουργίας λογαριασμού.
2. Συμπληρώνει τα στοιχεία που ζητούνται στην φόρμα που εμφανίζεται και την υποβάλλει.
3. Το όνομα χρήστη και ο κωδικός έχουν δημιουργηθεί και μπορούν να χρησιμοποιηθούν για την είσοδο στο σύστημα.

S-2: Δημιουργία εγγραφών ετεροαναφορών

1. Ο χρήστης εκτελεί τη λειτουργία δημιουργίας εγγραφής ετεροαναφοράς.
2. Συμπληρώνει τα στοιχεία της φόρμας που εμφανίζεται.
3. Επιλέγει τις δημοσιεύσεις του που βρίσκονται στην ετεροαναφορά και υποβάλλει τη συμπληρωμένη φόρμα.

S-3: Εισαγωγή δημοσιεύσεων από αρχείο BibTeX

1. Ο χρήστης εκτελεί τη λειτουργία φόρτωσης αρχείου BibTeX.
2. Στη λίστα με τις εγγραφές που βρέθηκαν μέσα στο αρχείο, επιλέγει αυτές που επιθυμεί να εισάγει.
3. Ενεργοποιεί τη διαδικασία εισαγωγής.
4. Επεξεργάζεται τις ετεροαναφορές που εισήγαγε προκειμένου να επιλέξει τις δημοσιεύσεις στις οποίες αναφέρονται.
5. Αποθηκεύει τις αλλαγές.

S-4: Προβολή λεπτομερειών ετεροαναφορών

1. Ο χρήστης εκτελεί τη λειτουργία προβολής λεπτομερειών επιλέγοντας τη λειτουργία «Details».
2. Εμφανίζεται ένας πίνακας με όλα τα στοιχεία της ετεροαναφοράς που έχουν εισαχθεί.

S-5: Διαγραφή ετεροαναφοράς

1. Ο χρήστης επιλέγει τη λειτουργία «Delete».
2. Απαντάει θετικά στο μήνυμα επιβεβαίωσης.
3. Εμφανίζεται η λίστα με τις υπόλοιπες ετεροαναφορές.

Ροές κατά εξαίρεση (Exceptional flows):

- S-3α.** Ο χρήστης έχει ήδη φορτώσει κάποιο αρχείο μορφής BibTeX και επιθυμεί να εισάγει κάποιες από τις εγγραφές που υπάρχουν σε αυτό και επιλέγει τη λειτουργία «View uploaded file entries».

Προαπαιτούμενα (Preconditions):

1. Η ύπαρξη ετεροαναφορών στις δημοσιεύσεις του χρήστη.

Ισχύοντα μετά την εκτέλεση/Αποτέλεσμα (Post-conditions): Ο ερευνητής ή ακαδημαϊκός, διαθέτει μια λίστα με τις ετεροαναφορές σε δημοσιεύσεις του, χάρη στις οποίες μπορεί να προβάλει αναφορές εκτύπωσης.

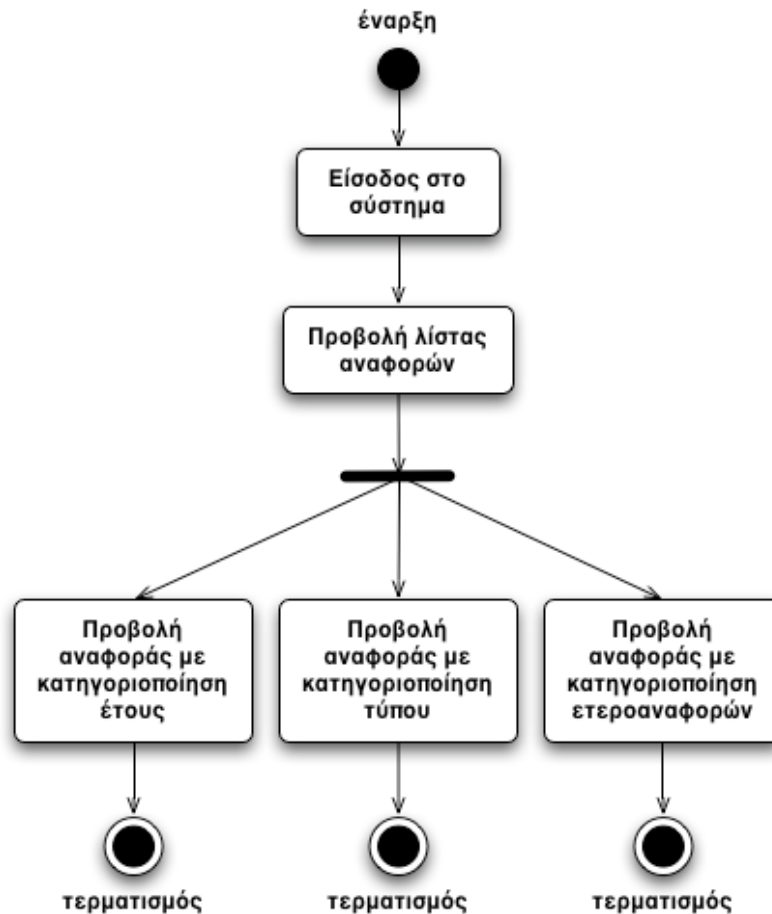
Πίνακας 11. Περιγραφή υπηρεσίας διαχείρισης ετεροαναφορών

4.2.4 Προβολή αναφορών εκτύπωσης

Οι αναφορές εκτύπωσης δίνουν τη δυνατότητα στο χρήστη να αξιοποιήσει την πληροφορία που έχει εισάγει στο σύστημα καθώς του επιτρέπουν να αντιγράψει και να επι-

κολλήσει τις αναφορές αυτές σε κάποιον κειμενογράφο ή απλά να δει πληροφορίες σχετικά με το πλήθος των δημοσιεύσεων ανά έτος, τον τύπο των δημοσιεύσεων του καθώς και για το πλήθος των ετεροαναφορών ανά δημοσίευση.

Στη συνέχεια παρουσιάζεται το διάγραμμα δραστηριοτήτων της υπηρεσίας (Εικόνα 15) καθώς και η επεξήγηση αυτού (Πίνακας 12).



Εικόνα 15. Διάγραμμα δραστηριοτήτων υπηρεσίας Προβολής αναφορών εκτύπωσης

Τίτλος: Προβολή αναφορών εκτύπωσης	ID: 4	Επίπεδο σπουδαιότητας: Υψηλό
Αποδέκτης (Actor): Ερευνητής ή ακαδημαϊκός ο οποίος έχει παρουσιάσει ερευνητικό έργο, απόρροια του οποίου είναι ένας αριθμός δημοσιεύσεων.		Τύπος υπηρεσίας: Λεπτομερής, πραγματική
Περιγραφή: Η υπηρεσία αυτή εκτελείται όταν ερευνητής ή ακαδημαϊκός, χρήστης του συστήματος, επιθυμεί την προβολή αναφορών εκτύπωσης σχετικές με τις δημοσιεύσεις του και τις ετεροαναφορές σε αυτές.		
Εκροή: Αναφορές εκτύπωσης.		

Εμπλεκόμενοι/ρόλοι (Stakeholders):			
A/A	Εμπλεκόμενος	Περιγραφή	Ρόλος
1.	Ερευνητής ή ακαδημαϊκός	Επιθυμεί την προβολή των αναφορών εκτύπωσης των δεδομένων του.	Χρήστης συστήματος
2.	ΠΣ	Προσφέρει τη διεπαφή για την προβολή των αναφορών εκτύπωσης που σχετίζονται με δημοσιεύσεις του χρήστη.	Πληροφοριακό σύστημα
Αφορμή/Περιοδικότητα (Trigger): Η εκτέλεση της υπηρεσίας προβολής αναφορών εκτύπωσης πραγματοποιείται κατόπιν επιθυμίας του χρήστη στη χρονική στιγμή που εκείνος επιθυμεί και μπορεί να επαναληφθεί οποτεδήποτε ο ίδιος το θελήσει.			
Σχέσεις: Συσχέτιση (Association): Περίληψη (Include): Επέκταση (Extend): Γενίκευση (Generalization):			
Κανονική ροή (Normal flow): 1. Ο Ερευνητής ή ακαδημαϊκός χρησιμοποιεί το όνομα χρήστη και το συνθηματικό του για να πιστοποιηθεί από το σύστημα. Σε περίπτωση που δεν έχει δημιουργήσει όνομα χρήστη και συνθηματικό για το σύστημα εκτελείται η υπό-ροή S-1. 2. Ο Ερευνητής ή ακαδημαϊκός επιλέγει την λειτουργία προβολής των αναφορών εκτύπωσης. 3. Το σύστημα εμφανίζει τις δημοσιεύσεις του χρήστη κατηγοριοποιημένες με βάση τη χρονιά δημοσίευσης. Σε περίπτωση που επιθυμείται η κατηγοριοποίηση με βάση τον τύπο της δημοσίευσης επιλέγεται η υπό-ροή S-2. Σε περίπτωση που επιθυμείται η κατηγοριοποίηση με βάση τον αριθμό ετεροαναφορών της κάθε δημοσίευσης εκτελείται η υπό-ροή S-3.			
Υπό-ροές (Subflows): S-1: Εγγραφή στο σύστημα και δημιουργία λογαριασμού 1. Ο χρήστης εκτελεί τη λειτουργία δημιουργίας λογαριασμού. 2. Συμπληρώνει τα στοιχεία που ζητούνται στην φόρμα που εμφανίζεται και την υποβάλλει. 3. Το όνομα χρήστη και ο κωδικός έχουν δημιουργηθεί και μπορούν να χρησιμοποιηθούν για την είσοδο στο σύστημα. S-2: Προβολή αναφοράς εκτύπωσης με βάση τον τύπο δημοσίευσης 1. Ο χρήστης εκτελεί τη λειτουργία προβολής αναφοράς εκτύπωσης «Per Type». S-3: Προβολή αναφοράς εκτύπωσης με βάση τον αριθμό ετεροαναφορών 1. Ο χρήστης εκτελεί τη λειτουργία προβολής αναφοράς εκτύπωσης «Per Citation»..			
Ροές κατά εξαίρεση (Exceptional flows):			
Προαπαιτούμενα (Preconditions): 1. Η ύπαρξη εγγραφών δημοσιεύσεων του χρήστη και ετεροαναφορών σε αυτές.			
Ισχύοντα μετά την εκτέλεση/Αποτέλεσμα (Post-conditions): Ο ερευνητής ή ακαδημαϊκός, διαθέτει αναφορές των δημοσιεύσεών του κατηγοριοποιημένες με βάση το έτος συγγραφής, το τύπο δημοσίευσης ή τον αριθμό των ετεροαναφορών σε αυτές.			

Πίνακας 12. Περιγραφή υπηρεσίας Προβολής αναφορών εκτύπωσης

4.2.5 Επεξεργασία στοιχείων χρήστη

Με την εγγραφή του στο σύστημα ο κάθε χρήστης εισάγει ορισμένα προσωπικά στοιχεία όπως τίτλος, ονοματεπώνυμο, email, όνομα όπως χρησιμοποιείται στις δημοσιεύσεις και κωδικό πρόσβασης. Η υπηρεσία επεξεργασίας στοιχείων του επιτρέπει να τροποποιήσει τα στοιχεία αυτά σε περίπτωση που κάτι τέτοιο κριθεί απαραίτητο.

Στη συνέχεια παρουσιάζεται το διάγραμμα δραστηριοτήτων της υπηρεσίας (Εικόνα 16) καθώς και η επεξήγηση αυτού (Πίνακας 13).



Εικόνα 16. Διάγραμμα δραστηριοτήτων υπηρεσίας Επεξεργασίας στοιχείων χρήστη

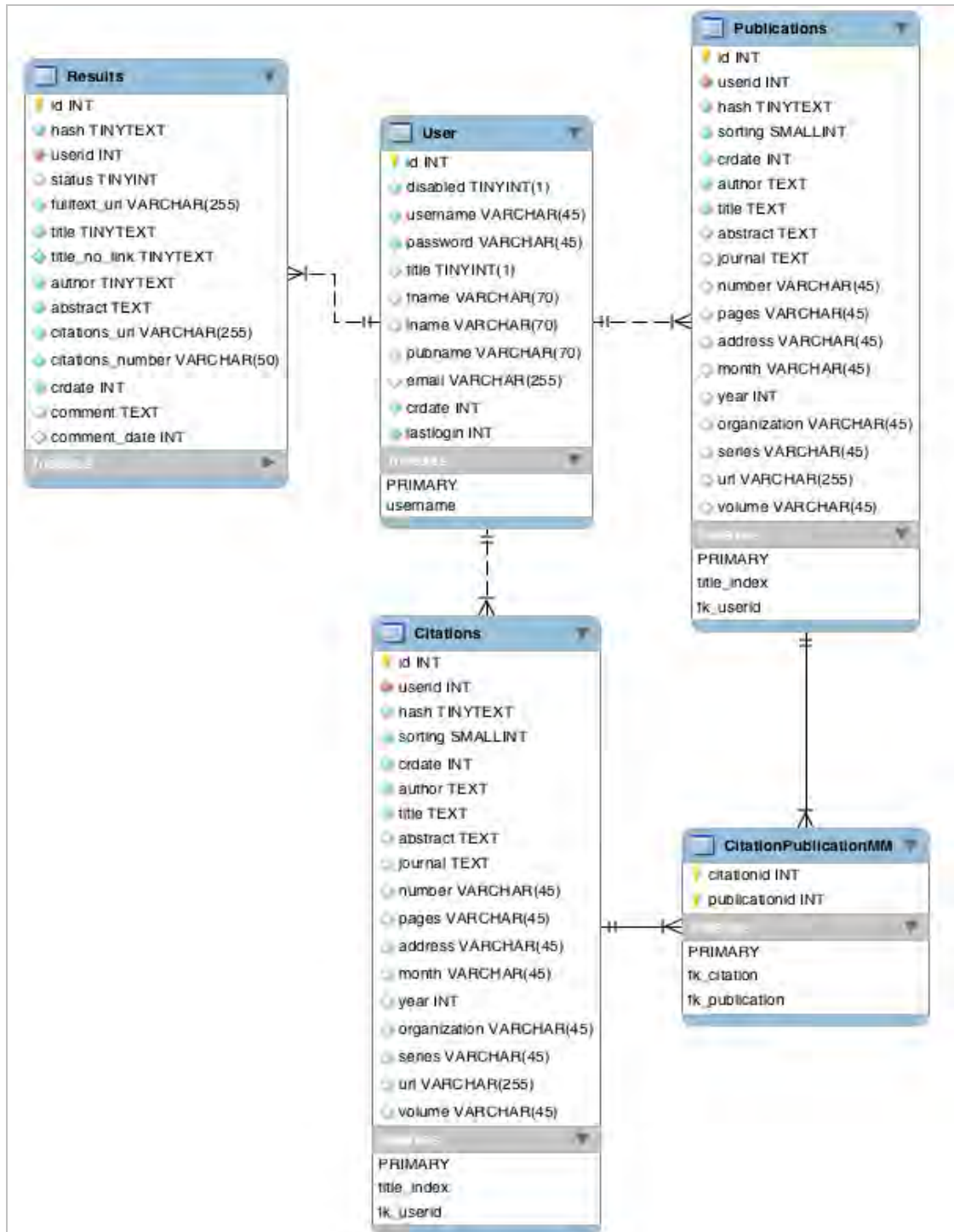
Τίτλος: Επεξεργασία στοιχείων χρήστη	ID: 5	Επίπεδο σπουδαιότητας: Υψηλό
Αποδέκτης (Actor): Ερευνητής ή ακαδημαϊκός ο οποίος επιθυμεί να τροποποιήσει τα στοιχεία του που είναι καταχωρημένα στο σύστημα.	Τύπος υπηρεσίας: Λεπτομερής, πραγματική	
Περιγραφή: Η υπηρεσία αυτή εκτελείται όταν ερευνητής ή ακαδημαϊκός, χρήστης του συστήματος, επιθυμεί την τροποποίηση των στοιχείων του που είναι καταχωρημένα στο σύστημα.		

Εκροή: Τροποποιημένα στοιχεία λογαριασμού χρήστη.			
Εμπλεκόμενοι/ρόλοι (Stakeholders):			
A/A	Εμπλεκόμενος	Περιγραφή	Ρόλος
1.	Ερευνητής ή ακαδημαϊκός	Επιθυμεί την τροποποίηση των στοιχείων του λογαριασμού του που είναι καταχωρημένα στο σύστημα.	Χρήστης συστήματος
2.	ΠΣ	Προσφέρει τη διεπαφή για την τροποποίηση των στοιχείων του χρήστη.	Πληροφοριακό σύστημα
Αφορμή/Περιοδικότητα (Trigger): Η εκτέλεση της υπηρεσίας επεξεργασίας στοιχείων χρήστη πραγματοποιείται κατόπιν επιθυμίας του χρήστη στη χρονική στιγμή που εκείνος επιθυμεί και μπορεί να επαναληφθεί οποτεδήποτε ο ίδιος το θελήσει.			
Σχέσεις: Συσχέτιση (Association): Περίληψη (Include): Επέκταση (Extend): Γενίκευση (Generalization):			
Κανονική ροή (Normal flow): <ol style="list-style-type: none"> Ο Ερευνητής ή ακαδημαϊκός χρησιμοποιεί το όνομα χρήστη και το συνθηματικό του για να πιστοποιηθεί από το σύστημα. Σε περίπτωση που δεν έχει δημιουργήσει όνομα χρήστη και συνθηματικό για το σύστημα εκτελείται η υπό-ροή S-1. Ο Ερευνητής ή ακαδημαϊκός επιλέγει την λειτουργία επεξεργασίας στοιχείων χρήστη (Edit Account). Το σύστημα εμφανίζει μια φόρμα με τα στοιχεία του χρήστη όπως αυτά είναι καταχωρημένα στο σύστημα. Ο χρήστης τροποποιεί τα στοιχεία του όπως επιθυμεί και υποβάλλει τη φόρμα προκειμένου να αποθηκευτούν τα νέα στοιχεία. 			
Υπό-ροές (Subflows): S-1: Εγγραφή στο σύστημα και δημιουργία λογαριασμού <ol style="list-style-type: none"> Ο χρήστης εκτελεί τη λειτουργία δημιουργίας λογαριασμού. Συμπληρώνει τα στοιχεία που ζητούνται στην φόρμα που εμφανίζεται και την υποβάλλει. Το όνομα χρήστη και ο κωδικός έχουν δημιουργηθεί και μπορούν να χρησιμοποιηθούν για την είσοδο στο σύστημα. 			
Ροές κατά εξαίρεση (Exceptional flows): 4α. Κάποιο από τα στοιχεία που έχει πληκτρολογήσει ο χρήστης είναι λανθασμένο και η φόρμα δεν υποβάλλεται αν δεν συμπληρωθεί σωστά.			
Προαπαιτούμενα (Preconditions): <ol style="list-style-type: none"> Ο Ερευνητής ή ακαδημαϊκός να διαθέτει ενεργό λογαριασμό στο σύστημα. 			
Ισχύοντα μετά την εκτέλεση/Αποτέλεσμα (Post-conditions): Ο ερευνητής ή ακαδημαϊκός, διαθέτει αναφορές των δημοσιεύσεών του κατηγοριοποιημένες με βάση το έτος συγγραφής, το τύπο δημοσίευσης ή τον αριθμό των ετεροαναφορών σε αυτές.			

Πίνακας 13. Περιγραφή υπηρεσίας Επεξεργασίας στοιχείων χρήστη

4.3 Σχήμα Βάσης Δεδομένων

Το βασικότερο ίσως δομικό στοιχείο του συστήματος αποτελεί η Βάση Δεδομένων αυτού (Database Schema). Στην Εικόνα 17 παρουσιάζεται γραφικά το σχήμα της και όπως φαίνεται, αποτελείται από 5 πίνακες ενώ οι επόμενες παράγραφοι περιέχουν λεπτομερή περιγραφή των πεδίων τους.



Εικόνα 17. Σχήμα Βάσης Δεδομένων

4.3.1 Πίνακας ‘User’

Στον πίνακα ‘User’ αποθηκεύονται πληροφορίες σχετικά με τους χρήστες του συστήματος. Τα στοιχεία που δίνει κατά την εγγραφή του ο ερευνητής ή ακαδημαϊκός, αποθηκεύονται σε αυτόν τον πίνακα ο οποίος περιέχει τα παρακάτω πεδία:

1. **id**: αύξων αριθμός, ο οποίος χρησιμοποιείται από το σύστημα για την μονοσήμαντη αναφορά στον χρήστη
2. **username**: το όνομα χρήστη, το οποίο απαιτείται από το σύστημα κατά την ταυτοποίηση του χρήστη
3. **password**: ο κωδικός χρήστη, ο οποίος αποθηκεύεται κωδικοποιημένος κατά md5 στη βάση και χρησιμοποιείται για την ταυτοποίηση των στοιχείων εισόδου στο σύστημα
4. **title**: ο τίτλος του χρήστη, δηλαδή Mr., Mrs, Prof. Dr. κλπ.
5. **fname**: το μικρό όνομα του χρήστη
6. **lname**: το επίθετο του χρήστη
7. **pubname**: το ονοματεπώνυμο του χρήστη όπως αυτό εμφανίζεται στις δημοσιεύσεις του
8. **email**: ο λογαριασμός ηλεκτρονικού ταχυδρομείου του χρήστη
9. **crdate**: η ημερομηνία και ώρα δημιουργίας του λογαριασμού σε timestamp
10. **lastlogin**: η ημερομηνία και ώρα της τελευταίας σύνδεσης του χρήστη με το σύστημα σε timestamp

Αξίζει να σημειωθεί το γεγονός ότι πέραν του πεδίου ‘id’ το οποίο αποτελεί το πρωτεύον κλειδί και είναι ούτως ή άλλως δεικτοδοτημένο, το πεδίο ‘username’ έχει επίσης δεικτοδοτηθεί προκειμένου να επιταχυνθεί η διαδικασία των αναζητήσεων με βάση το συγκεκριμένο πεδίο στο σύστημα. Αυτό κρίθηκε απαραίτητο για το λόγο ότι κατά την εγγραφή του χρήστη ελέγχεται κατά πόσον το επιθυμητό ‘username’ βρίσκεται ήδη στη βάση επειδή κάποιος άλλος χρήστης το έχει χρησιμοποιήσει ή όχι.

4.3.2 Πίνακας ‘Results’

Ο πίνακας ‘Results’ περιλαμβάνει τα αποτελέσματα που προκύπτουν από την αναζήτηση στο Google Scholar. Τα πεδία που περιέχει αντιπροσωπεύουν κατά κύριο λόγο τα

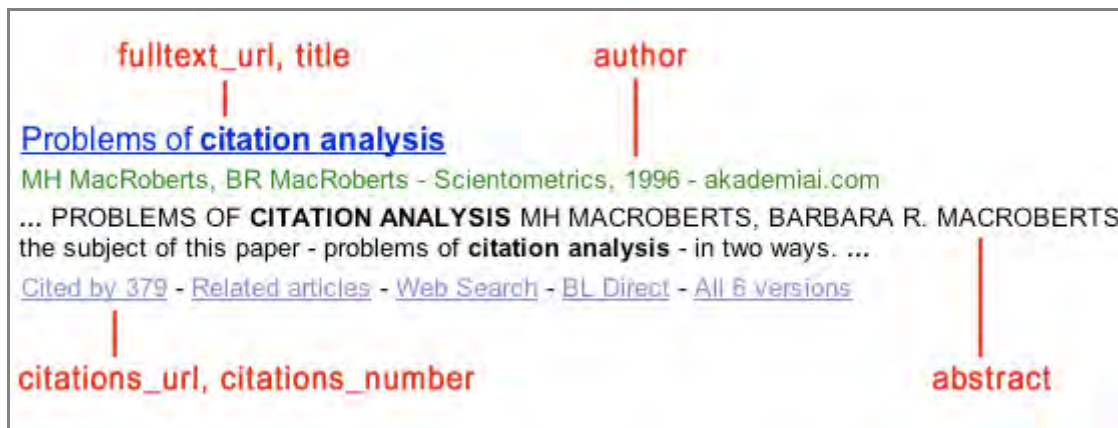
στοιχεία που παρουσιάζονται σε κάθε ένα από τα αποτελέσματα μιας αναζήτησης και είναι:

1. **hash**: ο τίτλος του αποτελέσματος σε hash, md5 μορφή
2. **userid**: το id του χρήστη που πραγματοποίησε την αναζήτηση στο Google Scholar
3. **status**: η κατάσταση του αποτελέσματος (New, True Citation, False Citation, Unknown κλπ.)
4. **related_to**: το id του αποτελέσματος στο οποίο το τρέχον αποτέλεσμα έχει κατηγοριοποιηθεί ως ετεροαναφορά (Citation) από το Google Scholar
5. **suggestion**: το πλήρες κείμενο της πρότασης του συστήματος σχετικά με την κατηγοριοποίηση του αποτελέσματος
6. **fulltext_url**: το url του αποτελέσματος
7. **title**: ο τίτλος του αποτελέσματος
8. **author**: οι συγγραφείς και το συνέδριο όπου έγινε η δημοσίευση
9. **abstract**: η περιγραφή της δημοσίευσης όπως αυτή εμφανίζεται στο αποτέλεσμα
10. **citations_url**: το url που οδηγεί σε πιθανά citations της δημοσίευσης
11. **citations_number**: ο αριθμός των πιθανών citation της δημοσίευσης του αποτελέσματος
12. **crdate**: η ημερομηνία που δημιουργήθηκε η εγγραφή
13. **comment**: το σχόλιο του χρήστη (αν υπάρχει)
14. **comment_date**: η ημερομηνία που δημιουργήθηκε ή ενημερώθηκε το σχόλιο του χρήστη

Η πληροφορία που υπάρχει στον πίνακα 'Results' χρησιμοποιείται τόσο για την εμφάνιση των αποτελεσμάτων στον χρήστη και την απόδοση της δυνατότητας κατηγοριοποίησης αυτών, όσο και για την διατήρηση της κατάστασης επεξεργασίας. Συγκεκριμένα, το πεδίο hash χρησιμοποιείται για τον έλεγχο ύπαρξης του αποτελέσματος πριν από την αποθήκευσή του έτσι ώστε να εξασφαλίζεται ότι τα ήδη επεξεργασμένα αποτελέσματα δεν θα εμφανίζονται ξανά, ενώ το πεδίο suggestion διατηρεί το περιεχόμενο της

πρότασης όπως αυτή υπολογίζεται από το σύστημα προκειμένου να μην εισάγεται επιπλέον καθυστέρηση κατά την εμφάνιση των αποτελεσμάτων.

Στην Εικόνα 18 παρουσιάζεται ένα αποτέλεσμα αναζήτησης του Google Scholar και η αντιστοίχιση των πληροφοριών που παρέχονται σε αυτό με τα πεδία της βάσης δεδομένων του συστήματος.



Εικόνα 18. Αντιστοίχιση αποτελέσματος Google Scholar και πεδίων της Βάσης Δεδομένων

4.3.3 Πίνακας 'Publications'

Στον πίνακα 'Publications' αποθηκεύονται οι πληροφορίες που αντιστοιχούν στις δημοσιεύσεις του χρήστη. Προκειμένου να διατηρηθεί η συμβατότητα με τα αρχεία BibTeX, δημιουργήθηκαν τα ίδια ακριβώς πεδία, όπως αυτά εμφανίζονται σε ένα τέτοιο αρχείο. Παρακάτω παρατίθεται η λίστα των πεδίων:

1. **userid**: το id του χρήστη που είναι συνδεδεμένος στο σύστημα και επιθυμεί να δημιουργήσει εγγραφές δημοσιεύσεων
2. **crdate**: η ημερομηνία δημιουργίας της εγγραφής σε timestamp
3. **sorting**: πεδίο το οποίο χρησιμοποιείται για την προσωπική ταξινόμηση των δημοσιεύσεων
4. **type**: ο τύπος της δημοσίευσης (Book, Book Chapter, Journal, Conference, Workshop κλπ.)
5. **author**: οι συγγραφείς της δημοσίευσης
6. **title**: ο τίτλος της δημοσίευσης
7. **abstract**: μια περίληψη του κειμένου που δημοσιεύτηκε

8. **journal**: το περιοδικό στο οποίο δημοσιεύτηκε
9. **number**: ο αριθμός του περιοδικού όπου δημοσιεύτηκε
10. **pages**: οι σελίδες στις οποίες βρίσκεται το δημοσιευμένο κείμενο
11. **address**: η πόλη στην οποία έγινε η δημοσίευση
12. **month**: ο μήνας δημοσίευσης
13. **year**: το έτος δημοσίευσης
14. **volume**: το τεύχος του περιοδικού όπου βρίσκεται το κείμενο
15. **keywords**: λέξεις κλειδιά του κειμένου
16. **pdf**: το url για το κείμενο σε μορφή pdf
17. **booktitle**: ο τίτλος του βιβλίου
18. **editor**: το όνομα του εκδότη
19. **url**: το url της δημοσίευσης
20. **doi**: ο μοναδικός αριθμός σήμανσης βιβλιογραφικών αναφορών
21. πλήθος BibTeX πεδίων όπως: organization, series, bibtexEntryType, bibtexCitation, institution, publisher, bibtex_type, note

4.3.4 Πίνακας ‘Citations’

Ο πίνακας ‘Citations’ αποθηκεύει πληροφορίες σχετικά με τις δημοσιεύσεις τρίτων οι οποίες αναφέρονται σε δημοσιεύσεις του χρήστη. Για τον λόγο ότι αποθηκεύει ακριβώς τον ίδιο τύπο δεδομένων με τον πίνακα ‘Publications’ έχει ακριβώς τα ίδια πεδία. Ο αναγνώστης παραπέμπεται στην παράγραφο 4.3.3 για περισσότερες πληροφορίες σχετικά με τα πεδία αυτά.

4.3.5 Πίνακας ‘CitationPublicationMM’

Όπως αναφέρθηκε και παραπάνω, ο πίνακας ‘Citations’ αποθηκεύει τα δεδομένα των δημοσιεύσεων τρίτων, οι οποίες αναφέρονται σε δημοσιεύσεις του χρήστη ή αλλιώς, στα δεδομένα του πίνακα ‘Publications’. Προκειμένου να είναι δυνατή η συσχέτιση των ετεροαναφορών (Citations) τρίτων με τις δημοσιεύσεις (Publications) του χρήστη, κρίθηκε απαραίτητη η δημιουργία ενός πίνακα στον οποίο θα αποθηκεύεται αποκλειστικά η πληροφορία αυτή.

Ο πίνακας αυτός ονομάζεται ‘CitationPublicationMM’ με το MM να προκύπτει από τα αρχικά της φράσης «Many to Many» η οποία να αναφέρεται στη σχέση μεταξύ των

δεδομένων των πινάκων 'Citations' και 'Publications'. Πιο συγκεκριμένα, αυτό σημαίνει ότι ένα Citation μπορεί να συσχετίζεται με πολλά Publications και το αντίθετο, δηλαδή το κάθε Publication να περιέχει πολλά Citations. Τα πεδία του πίνακα αυτού είναι μόνο δύο και είναι τα εξής:

1. **citationid**: τα id εγγραφών του πίνακα 'Citations'
2. **publicationid**: τα id εγγραφών του πίνακα 'Publications'

4.4 Περιβάλλον λειτουργίας

Προκειμένου να καλυφθεί η απαίτηση χρήσης του συστήματος σε οποιονδήποτε υπολογιστή ανεξάρτητα από το λειτουργικό του σύστημα και να υπάρχει πρόσβαση σε αυτό ανεξάρτητα από την τοποθεσία, αποφασίστηκε ότι το σύστημα θα πρέπει να αναπτυχθεί ως εφαρμογή βασισμένη στο διαδίκτυο (web-based εφαρμογή). Αυτό δίνει τη δυνατότητα εκτέλεσης του συστήματος από τον browser, λογισμικό το οποίο βρίσκεται σε όλα τα λειτουργικά συστήματα.

4.4.1 Βάση δεδομένων

Η απαίτηση ενός εύχρηστου τρόπου δημιουργίας, τροποποίησης, αποθήκευσης και ανάγνωσης των δεδομένων οδήγησε στην απόφαση χρησιμοποίησης ενός συστήματος διαχείρισης σχεσιακών βάσεων δεδομένων (RDBMS) το οποίο θα τηρεί όλα τα δεδομένα του συστήματος, όπως λογαριασμοί χρηστών, αποτελέσματα αναζητήσεων, δημοσιεύσεις, ετεροαναφορές κ.α. Το RDBMS σύστημα που επιλέχθηκε είναι το MySQL της Sun Microsystems και οι λόγοι που οδήγησαν σε αυτή την επιλογή είναι:

- η ωριμότητα και σταθερότητα του συστήματος (η πρώτη έκδοση βγήκε το Μάιο του 1994 ενώ η τρέχουσα έκδοση είναι η 5.1.34)
- η δυνατότητα υποστήριξης πολύ μεγάλων και ιδιαίτερα σύνθετων σχημάτων βάσεων δεδομένων
- η ύπαρξη εκδόσεων του συστήματος για όλα τα λειτουργικά συστήματα
- το γεγονός ότι δεν απαιτείται πολυδάπανη αγορά του συστήματος δεδομένου ότι είναι ανοιχτού κώδικα
- η μεγάλη απήχηση στον κόσμο των προγραμματιστών, γεγονός το οποίο διευκολύνει την ανάπτυξη εφαρμογών δεδομένου ότι το μεγαλύτερο μέρος των προ-

βλημάτων που μπορεί να συναντήσει κανείς έχουν ήδη αντιμετωπιστεί από άλλους προγραμματιστές

4.4.2 Γλώσσα προγραμματισμού

Θεωρώντας δεδομένη τη χρησιμοποίηση του RDBMS συστήματος MySQL, είναι σαφές πως οι σελίδες της εφαρμογής θα πρέπει να παράγονται δυναμικά. Με άλλα λόγια η εφαρμογή θα πρέπει να αναπτυχθεί είτε σε ASP.net, είτε σε JSP (JavaServer Pages), είτε σε PHP. Η επιλογή της PHP θεωρήθηκε ιδανική, καθώς:

- υποστηρίζει εγγενώς την ιδανικότερη δυνατή συνεργασία με τη MySQL
- είναι η συνηθέστερη γλώσσα προγραμματισμού στο web με αποτέλεσμα να υπάρχουν πάρα πολλές σχετικές πηγές πληροφόρησης
- υποστηρίζεται τόσο από τον Apache web server (linux) όσο και από τον IIS (windows)
- είναι ανοιχτού κώδικα με αποτέλεσμα να διατίθεται δωρεάν
- ο γράφων έχει εμπειρία στην ανάπτυξη λογισμικού με τη χρήση της συγκεκριμένης γλώσσας

4.4.3 Λειτουργικό σύστημα

Παρόλο που δεν υπάρχει κάποιος περιορισμός στο λειτουργικό σύστημα του διακομιστή όπου θα εγκατασταθεί η εφαρμογή, η πλέον ενδεδειγμένη και δοκιμασμένη λύση είναι η χρήση κάποιου συστήματος Unix, π.χ. Linux. Αυτό κυρίως λόγω της πολύ καλής συνεργασίας ανάμεσα στον Apache, τον MySQL και την PHP, αλλά και των καλύτερων δυνατοτήτων ελέγχου που δίνει στον διαχειριστή η χρήση των *nix λειτουργικών συστημάτων. Αξίζει να σημειωθεί ότι τόσο ο Apache, όσο ο MySQL και η PHP δημιουργήθηκαν αρχικά για χρήση σε συστήματα Unix/Linux.

Ανακεφαλαιώνοντας, το περιβάλλον στο οποίο θα λειτουργεί η εφαρμογή περιλαμβάνει το λειτουργικό σύστημα Linux, τον web server Apache, τη βάση δεδομένων MySQL και τη γλώσσα προγραμματισμού PHP. Αυτός ο συνδυασμός είναι γνωστός και με το ακρωνύμιο LAMP με τα αρχικά να αντιστοιχούν στις λέξεις Linux, Apache, MySQL/mSQL και Perl/PHP/Python. [34].

5 Υλοποίηση Συστήματος

Έχοντας περιγράψει τη δομή του λογισμικού που αναπτύχθηκε στο κεφάλαιο 4, στο τρέχον κεφάλαιο παρουσιάζεται αρχικά το περιβάλλον στο οποίο αναπτύχθηκε το σύστημα (παράγραφος 5.1), ο τρόπος λειτουργίας του συστήματος όπως αυτός υλοποιήθηκε, ή με άλλα λόγια οι εργασίες τις οποίες σχεδιάστηκε να επιτελεί (παράγραφος 5.2), τα προβλήματα που προέκυψαν κατά την ανάπτυξή του και τις λύσεις που δόθηκαν σε αυτά (παράγραφος 5.3), ενώ στην τελευταία παράγραφο του κεφαλαίου παρουσιάζονται τα δύο πιθανότερα σενάρια χρήσης του συστήματος (παράγραφος 5.4).

5.1 Περιβάλλον Υλοποίησης

Έχοντας αναλύσει τους λόγους οι οποίοι οδήγησαν στις επιλογές του καταλληλότερου περιβάλλοντος λειτουργίας του συστήματος στην παράγραφο 4.4 σε αυτήν την παράγραφο θα παρουσιαστεί το περιβάλλον, το οποίο χρησιμοποιήθηκε για την υλοποίηση του.

5.1.1 jQuery JavaScript Library [37]

Η προσεκτική ανάλυση και σχεδίαση της εφαρμογής οδήγησε στην απόφαση χρήσης κώδικα JavaScript προκειμένου να διευκολυνθούν ορισμένες από τις λειτουργίες που καλείται να επιτελέσει ο χρήστης. Πιο συγκεκριμένα, τα σημεία στα οποία κρίθηκε απαραίτητη η χρήση κώδικα JavaScript είναι:

- στη διαδικασία δημιουργίας, επεξεργασίας και διαγραφής σχολίων στο πλαίσιο της αξιολόγησης των αποτελεσμάτων
- στην επιλογή εγγραφών από το αρχείο BibTeX προκειμένου να εισαχθούν ως δημοσιεύσεις ή ετεροαναφορές
- στη μορφοποίηση της σχεδίασης της διεπαφής (χρήση δεσμών ως πλήκτρα υποβολής φορμών κ.α.)

Η πιο διαδεδομένη και δυναμικά αναπτυσσόμενη βιβλιοθήκη για την ανάπτυξη κώδικα σε JavaScript σήμερα είναι η jQuery με πλήθος υποστηρικτών και με παροχή μιας ιδιαίτερα ευρείας υποστήριξης στους προγραμματιστές μέσα από αμέτρητα εγχειρίδια

λειτουργίας και παραδειγμάτων ελεύθερα προσβάσιμα στο Internet. Αξιοσημείωτο είναι το γεγονός ότι παρόλο που ο γράφων δεν διέθετε προηγούμενη εμπειρία στη χρήση της συγκεκριμένης βιβλιοθήκης, οι πληροφορίες που υπάρχουν ελεύθερα στο Internet ήταν ικανές να του δώσουν τις απαραίτητες γνώσεις σε πολύ σύντομο χρονικό διάστημα. Επίσης ένα σημαντικό πλεονέκτημα του jQuery είναι ότι παρέχεται με τη μορφή λογισμικού ανοιχτού κώδικα. Πιο συγκεκριμένα, τα πλεονεκτήματα του jQuery είναι:

- η ταχύτητα και το μικρό του μέγεθος
- η υποστήριξη event handling, animating και Ajax
- ότι παρουσιάζει ιδιαίτερα μικρό κόστος σε όγκο (19KB)
- ότι είναι συμβατό με όλους τους Browser και
- ότι υποστηρίζει το πρότυπο CSS3

5.1.2 Yii Web Programming Framework [36]

Είναι ασφαλές να ειπωθεί ότι μετά από αρκετά χρόνια παλινδρόμησης ανάμεσα σε διαφορετικές υποδομές και γλώσσες προγραμματισμού για ανάπτυξη διαδικτυακών εφαρμογών ο συνδυασμός της PHP με την MySQL έχουν καταφέρει να κατακτήσουν το μεγαλύτερο μέρος της παγκόσμιας αγοράς. Οι λόγοι που συνετέλεσαν σε αυτό είναι πολλοί και ένας από τους πιο σημαντικούς είναι η ύπαρξη υποδομών οι οποίες διευκολύνουν σε μεγάλο βαθμό την ανάπτυξη συστημάτων μεγάλης κλίμακας.

Μια τέτοια υποδομή είναι το Yii Web Programming Framework, το οποίο αποτελεί μια υψηλής απόδοσης, αρθρωτή PHP υποδομή για την ανάπτυξη web εφαρμογών. Η υψηλή απόδοση επιτυγχάνεται χάρη στη μέγιστη δυνατή επαναχρησιμοποίηση τμημάτων κώδικα γεγονός το οποίο δύναται να επιταχύνει και τη διαδικασία ανάπτυξης. Το όνομα Yii (προφέρεται ως Yee ή [ji:]) προκύπτει από τις αγγλικές λέξεις easY, efficient και extensible ή αλλιώς εύκολο, αποδοτικό και επεκτάσιμο. Η επιλογή του Yii ως συστήματος ανάπτυξης κρίθηκε η καλύτερη δυνατή επειδή:

- είναι εύκολο στην εκμάθηση και την χρήση του, καθώς τα μόνα απαραίτητα εφόδια είναι η γνώση της γλώσσας προγραμματισμού PHP και η γνώση αντικειμενοστραφούς προγραμματισμού
- είναι πάρα πολύ γρήγορο δεδομένου ότι ο φόρτος που προκύπτει από τις εφαρμογές που αναπτύσσονται πάνω του είναι αμελητέος

- έχει πολύ μεγάλες δυνατότητες επεκτασιμότητας και επαναχρησιμοποίησης του κώδικα καθώς λόγω της αντικειμενοστραφούς δημιουργίας του το κάθε συστατικό είναι αυτόνομο και παραμετροποιήσιμο
- διαθέτει ένα πολύ πλούσιο σετ από χαρακτηριστικά, μεταξύ των οποίων MVC, DAO/ActiveRecord, χρήση θεμάτων και μεταγλώττισης, τα οποία το καθιστούν ιδιαίτερα ικανό στην ανάπτυξη των σημερινών Web 2.0 εφαρμογών
- διαθέτει πολύ ενημερωμένη και λεπτομερή βιβλιογραφία στην οποία περιλαμβάνεται όλη η πληροφορία που απαιτείται για τη γρήγορη εκμάθηση του
- είναι προσεκτικά σχεδιασμένο από την αρχή μέχρι το τέλος για την ανάπτυξη απαιτητικών web εφαρμογών κατά την ανάπτυξη του οποίου αναλύθηκε πλήθος παρόμοιων συστημάτων και τέλος
- διατίθεται υπό τη νέα BSD άδεια σύμφωνα με την οποία είναι δωρεάν τόσο οικονομικά όσο και νομικά η χρήση του Yii για την ανάπτυξη είτε εμπορικών εφαρμογών, είτε εφαρμογών ανοιχτού κώδικα

5.1.3 Eclipse Development Platform [35]

Το Eclipse αποτελεί μια πολυγλωσσική πλατφόρμα ανάπτυξης λογισμικού η οποία περιλαμβάνει ένα ενσωματωμένο περιβάλλον ανάπτυξης (IDE - Integrated Development Environment) και ένα σύστημα επεκτάσεων (plug-in) χάρη στο οποίο μπορούν να πολλαπλασιαστούν οι δυνατότητές του. Είναι γραμμένο σε Java και χάρη σε plugins επιτρέπει την ανάπτυξη εφαρμογών σε πάρα πολλές γλώσσες προγραμματισμού όπως C, C++, COBOL, Python, Perl, PHP και άλλες.

Το περιβάλλον ανάπτυξης για PHP περιλαμβάνει:

- editor για τη συγγραφή κώδικα
- debugger
- φυλλομετρητή κλάσεων (class browser)
- επιθεωρητή αντικειμένων (object inspector)
- ιεραρχικό διάγραμμα κλάσεων (class hierarchy diagram)
- τοπικό ιστορικό αλλαγών

Η επιλογή του Eclipse ως πλατφόρμα ανάπτυξης του PHP κώδικα βασίστηκε κυρίως στο γεγονός ότι ο γράφοντας διαθέτει εμπειρία ετών στο συγκεκριμένο σύστημα παρά σε κάποια ιδιαίτερα χαρακτηριστικά του που το έκαναν καλύτερο του ανταγωνιστι-

κού NetBeans. Σε κάθε περίπτωση, ορισμένα από τα σημαντικά χαρακτηριστικά του είναι ότι διατίθεται δωρεάν, ιδιαίτερα απλό στη χρήση, υποστηρίζει όλα τα λειτουργικά συστήματα και δεν παρουσιάζει προβλήματα κατά τη λειτουργία του.

5.2 Βασικές Λειτουργίες

Όπως γίνεται εύκολα αντιληπτό από το κεφάλαιο 3, στο οποίο περιγράφεται η ανάλυση των απαιτήσεων του συστήματος, οι βασικές λειτουργίες που αυτό σχεδιάστηκε να επιτελεί, είναι:

- η αναζήτηση, δημιουργία και διαχείριση βιβλιογραφικών ετεροαναφορών
- ο υπολογισμός των «προτάσεων συστήματος» (system suggestions)
- η δημιουργία και διαχείριση αρχείου προσωπικών δημοσιεύσεων του χρήστη ως αυτόνομες εγγραφές και
- η συσχέτιση των προσωπικών δημοσιεύσεων με τις ετεροαναφορές

5.2.1 Αναζήτηση, δημιουργία και διαχείριση ετεροαναφορών

Στόχος του συστήματος είναι η παροχή μιας υποδομής εύρεσης των ετεροαναφορών στις δημοσιεύσεις ενός ερευνητή και η συσχέτιση αυτών. Προκειμένου να καταστεί αυτό εφικτό, δίνεται η δυνατότητα αναζήτησης βιβλιογραφικών αναφορών με τη χρήση του Google Scholar. Ο λόγος για τον οποίο επιλέχθηκε η συγκεκριμένη μηχανή αναζήτησης είναι γιατί παρουσιάζει τα αποτελέσματα σε μια αρκετά δομημένη μορφή, γεγονός το οποίο βοηθάει στην επεξεργασία. Η διαδικασία αναζήτησης πραγματοποιείται σε δύο επίπεδα και στην συνέχεια παρουσιάζεται με ένα παράδειγμα.

Έστω ένας ερευνητής ο οποίος αναζητά τυχόν αναφορές που έχουν γίνει στα έργα του. Κάνοντας μια αναζήτηση στο Google Scholar με το όνομά του θα διαπιστώσει ότι τα επιστρεφόμενα αποτελέσματα περιλαμβάνουν στο μεγαλύτερο μέρος τους δικές του δημοσιεύσεις (Εικόνα 19) ενώ σε ένα μικρότερο ποσοστό περιλαμβάνουν αποτελέσματα τα οποία θεωρούνται ετεροαναφορές (citations) από το Google Scholar (Εικόνα 20).



Εικόνα 19. Παράδειγμα τυπικού αποτελέσματος «πρώτου επιπέδου» στο Google Scholar

[CITATION] 3DHYP: Implementation of a hypertext system using PROLOG
F Kokkoras - Diplomathesis (in Greek), 1992
[Cited by 3](#) - [Related articles](#) - [Web Search](#)

Εικόνα 20. Παράδειγμα αποτελέσματος αναφοράς στο Google Scholar

Οι παραπάνω δύο τύποι αποτελεσμάτων θεωρούνται από το σύστημα ως αποτελέσματα «πρώτου επιπέδου» και στην πρώτη αναζήτηση που πραγματοποιεί ο χρήστης με την εγγραφή του στο σύστημα παρουσιάζονται ως «Νέα αποτελέσματα». Προκειμένου όμως να παρουσιαστούν το δυνατό περισσότερα σχετικά αποτελέσματα στον χρήστη, το σύστημα προχωράει ένα επίπεδο πιο μέσα, ακολουθώντας τα links «**Cited by #**», τα οποία – σύμφωνα πάντα με το Google Scholar – παρουσιάζουν μεγαλύτερο ενδιαφέρον για την αναζήτηση του χρήστη καθώς πρόκειται για **Citations**, δηλαδή δημοσιεύσεις τρίτων οι οποίες περιλαμβάνουν με τη μορφή αναφοράς (reference) τη δημοσίευση του χρήστη στην οποία ανήκουν. Τα αποτελέσματα που προκύπτουν ακολουθώντας τα «**Cited by #**» links ονομάζονται αποτελέσματα «δεύτερου επιπέδου» και επίσης παρουσιάζονται ως «Νέα αποτελέσματα» την πρώτη φορά που χρησιμοποιείται το σύστημα.

Ο χρήστης στη συνέχεια έχει τη δυνατότητα να επεξεργαστεί όλα τα αποτελέσματα που του παρουσιάζονται προσθέτοντας σχόλια, διαγράφοντας τα ή κατηγοριοποιώντας τα ανάλογα με το αν πρόκειται για:

- Ιδία Δημοσίευση (Own Publication),
- Πραγματική Ετεροαναφορά (True Citation),
- Ψευδής Ετεροαναφορά (False Citation) ή
- Άγνωστο (Unknown)

Ένα από τα ζητήματα που έπρεπε να αντιμετωπιστεί προκειμένου να καταστεί δυνατή η λήψη και παρουσίαση των αποτελεσμάτων του Google Scholar ήταν η απουσία παροχής προγραμματιστικής διεπαφής ή αλλιώς API (Application Programming Interface) από την Google. Έπρεπε λοιπόν να βρεθεί κάποιος τρόπος να ληφθούν και να αποθηκευτούν τα αποτελέσματα της αναζήτησης στη βάση δεδομένων του συστήματος χωρίς την ύπαρξη κάποιου feed με τη βοήθεια του API. Η λύση αυτή βρέθηκε στο εργαλείο ΔΕΙΧΤο [32] μια σύντομη περιγραφή του οποίου, επιχειρείται στη συνέχεια.

Περιγραφή του εργαλείου «ΔΕΙΧΤο»

Το ΔΕΙΧΤο είναι ένα πολύ χρήσιμο εργαλείο για την αυτόματη εξαγωγή δεδομένων που βρίσκονται σε ιστοσελίδες του Διαδικτύου. Δημιουργεί ένα δένδρο για την σελίδα που εξετάζεται κάθε φορά, το οποίο αναπαριστά την ιεραρχία των ετικετών HTML (DOM μοντέλο). Σε σελίδες, οι οποίες περιέχουν τα δεδομένα τους πολύ καλά δομημένα, η εξαγωγή της πληροφορίας γίνεται με πολύ μεγάλη ακρίβεια. Τα αποτελέσματα μιας αναζήτησης στο Google Scholar ανήκουν σε αυτή την κατηγορία σελίδων και ως εκ τούτου η εξαγωγή επιτυγχάνεται απόλυτα.

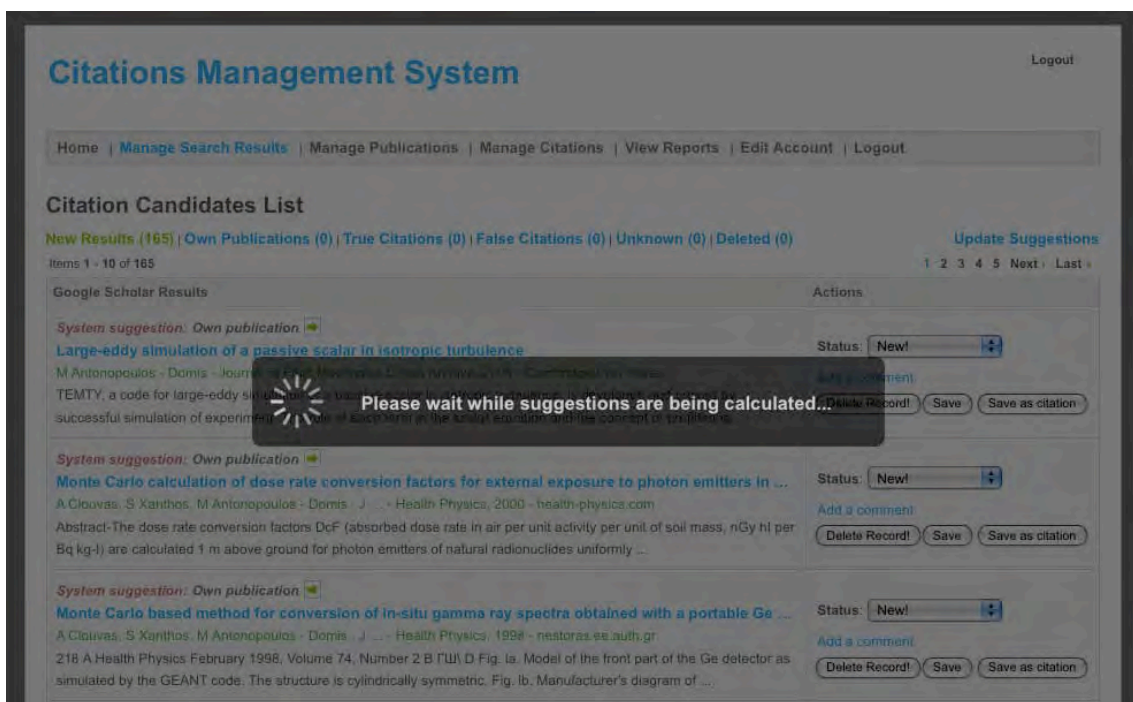
Προκειμένου το ΔΕΙΧΤο να γνωρίζει το ποια ακριβώς δεδομένα θα εξαχθούν πρέπει να δημιουργηθούν οι κανόνες εξαγωγής (ή αλλιώς wrapper), πράγμα το οποίο επιτυγχάνεται εύκολα μέσω ενός γραφικού περιβάλλοντος που έχει δημιουργηθεί γι' αυτόν το σκοπό. Ορισμένες από τις πλέον χρήσιμες δυνατότητες του ΔΕΙΧΤο είναι η εξαγωγή των αποτελεσμάτων σε αρχείο xml, η αυτόματη συνέχιση λήψης αποτελεσμάτων όταν αυτά βρίσκονται σε επόμενη σελίδα ακολουθώντας τον σύνδεσμο «Next» μέχρι το βάθος που ορίζεται από τον χρήστη, η προσομοίωση συμπεριφοράς κανονικού browser για την αποφυγή απαγόρευσης από το site του οποίου τα δεδομένα εξάγει κ.α.

5.2.2 Υπολογισμός «προτάσεων συστήματος»

Με στόχο την απλοποίηση της πιο επίπονης διαδικασίας, αυτής της κατηγοριοποίησης των αποτελεσμάτων σε Own Publication, True Citation, False Citation ή Unknown, το σύστημα χρησιμοποιεί έναν αλγόριθμο προκειμένου να υπολογίσει μια πρόταση (suggestion) σχετικά με το είδος του κάθε αποτελέσματος. Πιο συγκεκριμένα, πραγματοποιεί αναζήτηση του τίτλου κάθε αποτελέσματος αναζήτησης στις δημοσιεύσεις του χρήστη επιχειρώντας την ανεύρεση πιθανών ταυτίσεων (matches). Πιο αναλυτικά τα βήματα που ακολουθούνται στον αλγόριθμο είναι:

1. αφαίρεση των χαρακτήρων «...» από το τέλος του τίτλου του αποτελέσματος, καθώς και των κενών χαρακτήρων τόσο στην αρχή όσο και στο τέλος αυτού
2. μέτρηση του πλήθους χαρακτήρων που έχουν απομείνει και εφαρμογή του μήκους αυτού στους τίτλους όλων των δημοσιεύσεων του χρήστη, με αποκοπή των υπόλοιπων χαρακτήρων
3. υπολογισμός της απόστασης Levenshtein (Levenshtein Distance - LD) μεταξύ των τίτλων

4. σύγκριση της απόστασης Levenshtein με σταθμισμένη τιμή ανάλογης με το αρχικό μήκος του τίτλου προκειμένου να εξαχθούν συμπεράσματα σχετικά με την ταύτιση ή όχι του αποτελέσματος
5. υπολογισμός της πρότασης με βάση τη συσχέτιση του υπό εξέταση αποτελέσματος με άλλα αποτελέσματα αναζήτησης π.χ. αν ένα αποτέλεσμα δηλώνεται ως citation ενός άλλου στο Google Scholar, αυτό λαμβάνεται υπόψη στον υπολογισμό της «πρότασης του συστήματος»

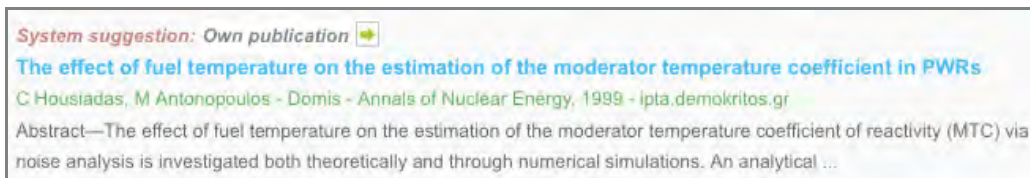



Εικόνα 21. Υπολογισμός «προτάσεων συστήματος»

Το γεγονός ότι εξετάζονται ένα προς ένα τα αποτελέσματα που έχουν προκύψει από το Google Scholar με κάθε μία από τις δημοσιεύσεις που έχει καταχωρήσει στο σύστημα ο χρήστης, αυξάνουν εκθετικά την πολυπλοκότητα του αλγορίθμου και ως εκ τούτου και τον χρόνο απόκρισης του συστήματος. Προκειμένου ο χρήστης να μην τερματίσει άθελά του τη διαδικασία υπολογισμού το παράθυρο της εφαρμογής σκιάζεται και απενεργοποιείται, ενώ πάνω του εμφανίζεται το μήνυμα «Please wait while suggestions are being calculated...» (βλ. Εικόνα 21). Επίσης, έχει προβλεφθεί η ύπαρξη ενός πληκτρού ανανέωσης των προτάσεων προκειμένου να μπορεί ο χρήστης να τις ενημερώσει όταν αυτός κρίνει απαραίτητο (π.χ. μετά την προσθήκη νέων δημοσιεύσεων). Όπως φαίνεται και στις εικόνες που ακολουθούν, οι προτάσεις του συστήματος εμφανίζονται με κόκκινο χρώμα και στις περιπτώσεις εκείνες όπου έχει βρεθεί συσχέτιση με κάποια

από τις δημοσιεύσεις του χρήστη, εμφανίζεται ένας δεσμός (link) προς τη δημοσίευση αυτή. Οι πιθανές περιπτώσεις προτάσεων είναι τέσσερις και είναι οι ακόλουθες:

1. **Own publication:** Ο τίτλος του αποτελέσματος ταυτίζεται με κάποια από τις δημοσιεύσεις, ενώ το αποτέλεσμα δε σχετίζεται με κάποιο άλλο. Δίπλα στη φράση «Own publication», εμφανίζεται ένα πράσινο βέλος, το οποίο σηματοδοτεί την ύπαρξη δεσμού (link) προς τη δημοσίευση (βλ. παρακάτω εικόνα).

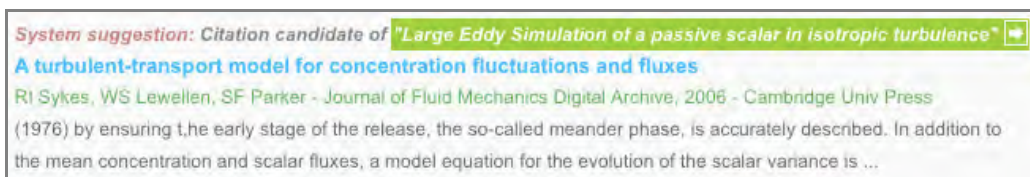



System suggestion: Own publication 

The effect of fuel temperature on the estimation of the moderator temperature coefficient in PWRs
C Housiadas, M Antonopoulos - Domis - Annals of Nuclear Energy, 1999 - ipta.demokritos.gr

Abstract—The effect of fuel temperature on the estimation of the moderator temperature coefficient of reactivity (MTC) via noise analysis is investigated both theoretically and through numerical simulations. An analytical ...

2. **Citation candidate of «publication_title»:** Το αποτέλεσμα σχετίζεται με κάποιο άλλο, το οποίο έχει ταυτιστεί με συγκεκριμένη δημοσίευση η οποία βρίσκεται στη βάση. Έτσι, αποτελεί υποψήφια ετεροαναφορά της δημοσίευσης με την οποία ταυτίστηκε το σχετικό αποτέλεσμα αναζήτησης.

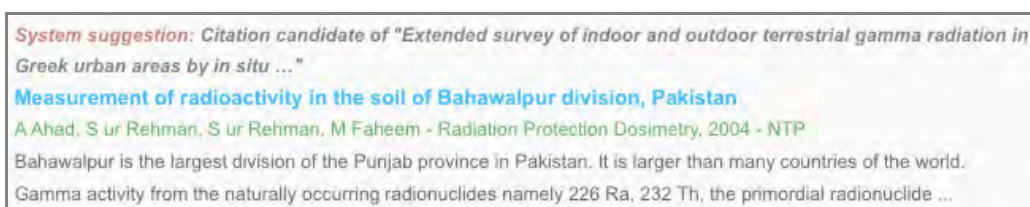


System suggestion: Citation candidate of "Large Eddy Simulation of a passive scalar in isotropic turbulence" 

A turbulent-transport model for concentration fluctuations and fluxes
Ri Sykes, WS Lewellen, SF Parker - Journal of Fluid Mechanics Digital Archive, 2006 - Cambridge Univ Press

(1976) by ensuring the early stage of the release, the so-called meander phase, is accurately described. In addition to the mean concentration and scalar fluxes, a model equation for the evolution of the scalar variance is ...

3. **Citation candidate of «result_title»:** Το αποτέλεσμα σχετίζεται με κάποιο άλλο, το οποίο όμως δεν ήταν δυνατόν να ταυτιστεί με κάποια δημοσίευση από αυτές που είναι ήδη καταχωρημένες στη βάση.



System suggestion: Citation candidate of "Extended survey of indoor and outdoor terrestrial gamma radiation in Greek urban areas by in situ ..."

Measurement of radioactivity in the soil of Bahawalpur division, Pakistan
A Ahad, S ur Rehman, S ur Rehman, M Faheem - Radiation Protection Dosimetry, 2004 - NTP

Bahawalpur is the largest division of the Punjab province in Pakistan. It is larger than many countries of the world. Gamma activity from the naturally occurring radionuclides namely ^{226}Ra , ^{232}Th , the primordial radionuclide ...

4. **No matching was possible:** Το αποτέλεσμα δε σχετίζεται με κανένα άλλο και ούτε ήταν δυνατός ο εντοπισμός του στον πίνακα των δημοσιεύσεων, επομένως το σύστημα αδυνατεί να πραγματοποιήσει κάποια πρόταση με την πληροφορία που έχει.

System suggestion: No matching with existing publications was possible.

Moderator temperature coefficient of reactivity in Pressurized Water Reactors: theoretical ...

M Antonopoulos - Domis , C Housiadas - Nuclear science and engineering, 1999 - cat.inist.fr

Estimation of the moderator temperature coefficient(MTC) of reactivity of pressurized water reactors by noise analysis is investigated, both theoretically and through numerical simulations. It is found that using local neutron ...

5.2.3 Δημιουργία και διαχείριση προσωπικών δημοσιεύσεων του χρήστη

Προκειμένου να συσχετιστούν οι ετεροαναφορές, που βρίσκονται μέσω του Google Scholar και αποθηκεύονται στη βάση του συστήματος χάρη στη χρήση του ΔΕΙΧΤο είναι απαραίτητη η τήρηση ενός αρχείου προσωπικών δημοσιεύσεων. Έτσι, το σύστημα δίνει τη δυνατότητα στον χρήστη είτε να δημιουργήσει τις εγγραφές χειροκίνητα μία μία είτε να χρησιμοποιήσει την αυτόματη εισαγωγή μέσω αρχείου BibTeX. Η κάθε εγγραφή προσωπικής δημοσίευσης (Publication) περιλαμβάνει πλήθος πεδίων, τα περισσότερα εκ των οποίων αποτελούν standards βιβλιογραφικών αναφορών. Ιδιαίτερη προσοχή έχει δοθεί στη δημιουργία βοηθητικών πεδίων για προσωπική ταξινόμηση και κατηγοριοποίηση των αναφορών. Ο αναγνώστης παραπέμπεται στην παράγραφο 4.3.3 για περισσότερες λεπτομέρειες.

5.2.4 Συσχέτιση των προσωπικών δημοσιεύσεων με τις ετεροαναφορές

Η συσχέτιση των προσωπικών δημοσιεύσεων με τις ετεροαναφορές γίνεται αποκλειστικά μέσω της φόρμας των ετεροαναφορών (citations) όπου παρουσιάζεται στον χρήστη μια λίστα με τις δημοσιεύσεις του, μέσω της οποίας μπορεί να επιλέξει τις αναφερόμενες δημοσιεύσεις. Ο συσχετισμός αυτός αποθηκεύεται στον πίνακα 'CitationPublicationMM' και επιτρέπει την δημιουργία αναφορών εκτύπωσης.

Αξίζει να σημειωθεί το γεγονός ότι το σύστημα υποστηρίζει τη δυνατότητα δημιουργίας εγγραφών ετεροαναφορών τόσο χειροκίνητα με τη συμπλήρωση μιας φόρμας όσο και αυτόματα με τη χρήση αρχείων BibTeX. Επίσης κατά τη δημιουργία ετεροαναφοράς απ' ευθείας από κάποιο αποτέλεσμα του Google Scholar το σύστημα συμπληρώνει αυτόματα τα πεδία εκείνα για τα οποία έχει δεδομένα, όπως author, abstract, title κλπ.

5.3 Προβλήματα Υλοποίησης

Το Google Scholar θέτει έναν περιορισμό ως προς τον αριθμό των ερωτημάτων που μπορεί να κάνει ένας χρήστης μέσα σε ένα χρονικό διάστημα. Αν ξεπεραστεί αυτός ο αριθμός τότε απαγορεύεται στον χρήστη να κάνει άλλα ερωτήματα. Αυτό συμβαίνει για την προστασία του Google από προγράμματα που στέλνουν συνεχώς ερωτήματα προς αυτό με αποτέλεσμα να καθιστούν την υπηρεσία πιο αργή ή και να τη θέτουν εκτός λειτουργίας (Denial Of Service Attacks – DoS Attacks).

Για την αντιμετώπιση του παραπάνω προβλήματος, το ΔΕΙΧΤο προσποιείται ότι είναι κανονικός browser και επιπλέον εισάγει μια χρονική καθυστέρηση μεταξύ των διαδοχικών ερωτημάτων. Αυτό όμως έχει ως αποτέλεσμα τη σημαντική αύξηση της διάρκειας λήψης των αποτελεσμάτων. Το πρόβλημα μπορεί να παρακαμφθεί άμεσα, με το να καταστεί εφικτή η επίσημη λήψη αποτελεσμάτων από το Google Scholar (απαιτεί διαπραγμάτευση με τη Google) ή έμμεσα, μεταφέροντας το τμήμα λήψης αποτελεσμάτων στον τελικό χρήστη, ο οποίος στη συνέχεια θα "ανεβάζει" τα αποτελέσματα που του επέστρεψε το Google Scholar στο σύστημα (ως XML αρχείο).

5.4 Σενάρια Χρήσης

Το σύστημα δημιουργήθηκε για να διευκολύνει την αναζήτηση βιβλιογραφικών αναφορών μέσω Google Scholar και την οργάνωση αυτών με βάση τις δημοσιεύσεις του χρήστη. Υπό αυτό το πρίσμα προκύπτουν τα παρακάτω δύο σημαντικά σενάρια χρήσης για το σύστημα:

1. Χρήση της πλήρους λειτουργικότητας του συστήματος, δηλαδή αναζήτηση και διαχείριση βιβλιογραφικών αναφορών
2. Χρήση του συστήματος για απλή οργάνωση των δημοσιεύσεων και των ετεροαναφορών

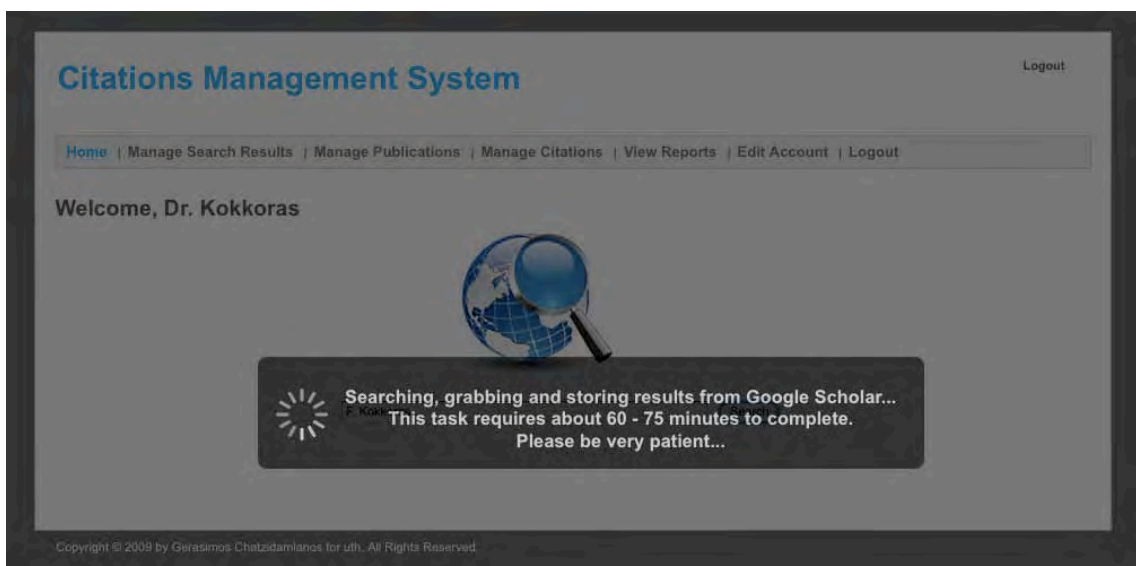
5.4.1 Σενάριο 1: Αναζήτηση βιβλιογραφικών αναφορών και δημιουργία ετεροαναφορών

Σε αυτό το σενάριο ο χρήστης του συστήματος αναζητεί βιβλιογραφικές αναφορές στο Google Scholar μέσω της διεπαφής της εφαρμογής και τις κατηγοριοποιεί ανάλογα, προσθέτει σχόλια σε αυτές ή τις διαγράφει. Στην Εικόνα 22 παρουσιάζεται η διεπαφή αναζήτησης η οποία αποτελεί και τη αρχική οθόνη μετά την εισαγωγή του χρήστη στο σύστημα.



Εικόνα 22. Αρχική οθόνη συστήματος

Το πεδίο αναζήτησης είναι προσυμπληρωμένο με το όνομα, που έχει δηλώσει ο χρήστης ότι εμφανίζεται στις δημοσιεύσεις του. Φυσικά είναι επεξεργάσιμο και δίνεται η δυνατότητα αναζήτησης και με άλλη λέξη κλειδί. Δεδομένου του περιορισμού στις αναζητήσεις μέσω Google Scholar (βλ. παράγραφο **Σφάλμα! Το αρχείο προέλευσης της αναφοράς δεν βρέθηκε.**) η διαδικασία της αναζήτησης και εξαγωγής των αποτελεσμάτων διαρκεί αρκετή ώρα (κατά μέσο όρο 60 - 75 λεπτά) και συνιστάται υπομονή ή ενασχόληση με κάτι άλλο από τον χρήστη (Εικόνα 23).

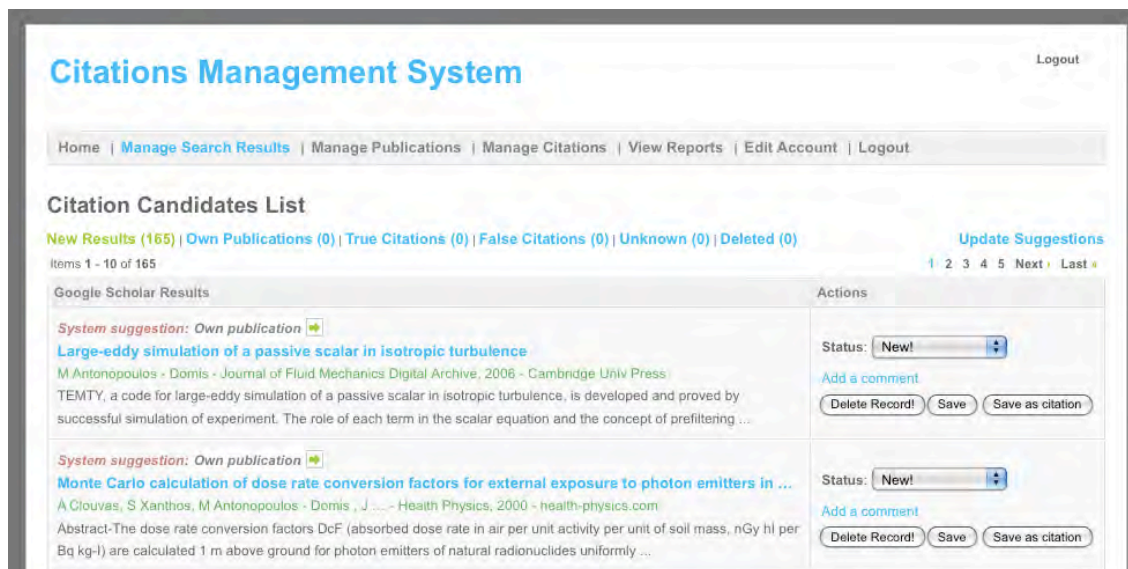


Εικόνα 23. Μήνυμα αναμονής αναζήτησης και εξαγωγής αποτελεσμάτων

Μετά το πέρας της αναζήτησης εμφανίζονται στον χρήστη τα αποτελέσματα αυτής (βλ. Εικόνα 24). Η μορφή με την οποία παρουσιάζονται είναι παρόμοια αυτής με την

οποία εμφανίζονται στο Google Scholar και αυτό έγινε με σκοπό να διατηρηθεί η φιλικότητα εμφάνισης των αποτελεσμάτων, εμφάνιση με την οποία είναι ήδη εξοικειωμένοι οι περισσότεροι χρήστες. Με την κόκκινη γραμματοσειρά ξεχωρίζει η «πρόταση» του συστήματος, η οποία μπορεί να βοηθήσει τον χρήστη να κατηγοριοποιήσει ευκολότερα τα αποτελέσματα της αναζήτησης.

Μετά την πρώτη αναζήτηση όλα τα αποτελέσματα εμφανίζονται ως νέα (**Status: New!**), ενώ σε μελλοντικές αναζητήσεις τα αποτελέσματα που έχουν ήδη κατηγοριοποιηθεί από το χρήστη διατηρούν την κατηγοριοποίησή τους. Αυτό είναι ιδιαίτερα σημαντικό αφ' ενός διότι επιτρέπει στο χρήστη να επεξεργάζεται τα αποτελέσματα όποτε αυτός θελήσει και αφ' ετέρου διότι οι μελλοντικές αναζητήσεις θα συνοδεύονται από πολύ λιγότερο φόρτο εργασίας.



The screenshot displays the 'Citations Management System' interface. At the top, there is a navigation bar with links for Home, Manage Search Results, Manage Publications, Manage Citations, View Reports, Edit Account, and Logout. Below this is a 'Citation Candidates List' section with statistics for New Results (165), Own Publications (0), True Citations (0), False Citations (0), Unknown (0), and Deleted (0). The main content area shows two entries from Google Scholar. Each entry includes a 'System suggestion: Own publication' label, a title, author information, and a brief abstract. To the right of each entry is an 'Actions' column with a 'Status' dropdown menu (set to 'New!'), an 'Add a comment' link, and buttons for 'Delete Record!', 'Save', and 'Save as citation'.

Εικόνα 24. Αποτελέσματα αναζήτησης

Στη στήλη «**Actions**» ο χρήστης μπορεί να επιλέξει από μια σειρά ενεργειών για το αποτέλεσμα στο οποίο εμφανίζεται. Μπορεί να αλλάξει την κατάσταση του αποτελέσματος από τον επιλογέα «**Status**», να εισάγει κάποιο σχόλιο όταν δεν υπάρχει κανόνας κλικ στο δεσμό «**Add a comment**» ή να διαγράψει/επεξεργαστεί το υπάρχον επιλέγοντας τους δεσμούς «**Remove comment**» και «**Edit comment**» αντίστοιχα, να διαγράψει το αποτέλεσμα με το πλήκτρο «**Delete Record!**», να αποθηκεύσει τις αλλαγές στην κατάσταση και στα σχόλια με το πλήκτρο «**Save**» και τέλος να δημιουργήσει εγγραφή απ' ευθείας στον πίνακα των Citations με το πλήκτρο «**Save as citation**».

Ενεργοποιώντας την λειτουργία «**Save as citation**», εμφανίζεται ένα μήνυμα το οποίο ενημερώνει τον χρήστη ότι η κατηγοριοποίηση του αποτελέσματος θα αλλάξει σε «**True Citation**» και ότι θα δημιουργηθεί μια νέα εγγραφή στον πίνακα των Citations (βλ. Εικόνα 25).



Εικόνα 25. Μήνυμα επιβεβαίωσης αποθήκευσης αποτελέσματος ως citation

New Citation

[Citations List](#)

Cited publication(s)	- Not Related -
Author(s) *	F Kokkoras , H Jiang, I Vlahavas, AK Elmagam, EN ... - Multimedia Systems, 2002 - Springer
Title *	Smart VideoText: a video data model based on conceptual graphs
Doi	
Bibtex Citation	
Type	--Not Defined--
Personal Ordering	9999
Journal/Conference	
Booktitle	
Number	
Volume	
Year	
Pages	
City	
Month	
Abstract	Abstract. An intelligent annotation-based video data model called Smart VideoText is introduced. It utilizes the conceptual graph knowledge representation formalism to capture the semantic associations among the ...

Εικόνα 26. Φόρμα δημιουργίας ετεροαναφοράς

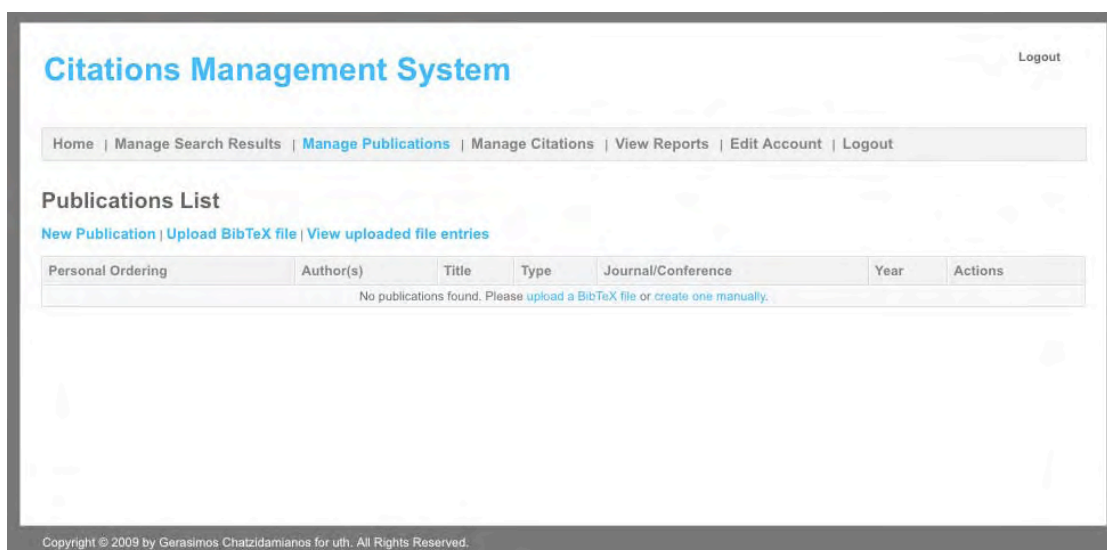
Αφού ο χρήστης επιβεβαιώσει ότι όντως επιθυμεί τη δημιουργία νέας εγγραφής και δεν πάτησε κατά λάθος το συγκεκριμένο πλήκτρο, εμφανίζεται η φόρμα δημιουργίας ετεροαναφοράς με ορισμένα από τα στοιχεία της ήδη συμπληρωμένα (Εικόνα 26). Για περισσότερες πληροφορίες σχετικά με τη δημιουργία και διαχείριση των ετεροαναφορών ο αναγνώστης παραπέμπεται στο σενάριο 2, στην παράγραφο 5.4.2.

5.4.2 Σενάριο 2: Διαχείριση δημοσιεύσεων και ετεροαναφορών

Στο σενάριο 1 (παράγραφος 5.4.1) παρουσιάστηκε ο τρόπος με τον οποίο το σύστημα επιτρέπει στον χρήστη να αναζητήσει ετεροαναφορές στο Google Scholar, να επεξεργαστεί τα αποτελέσματα που του παρουσιάζονται κατηγοριοποιώντας τα κατάλληλα, καθώς και ένας από τους τρόπους δημιουργίας ετεροαναφορών.

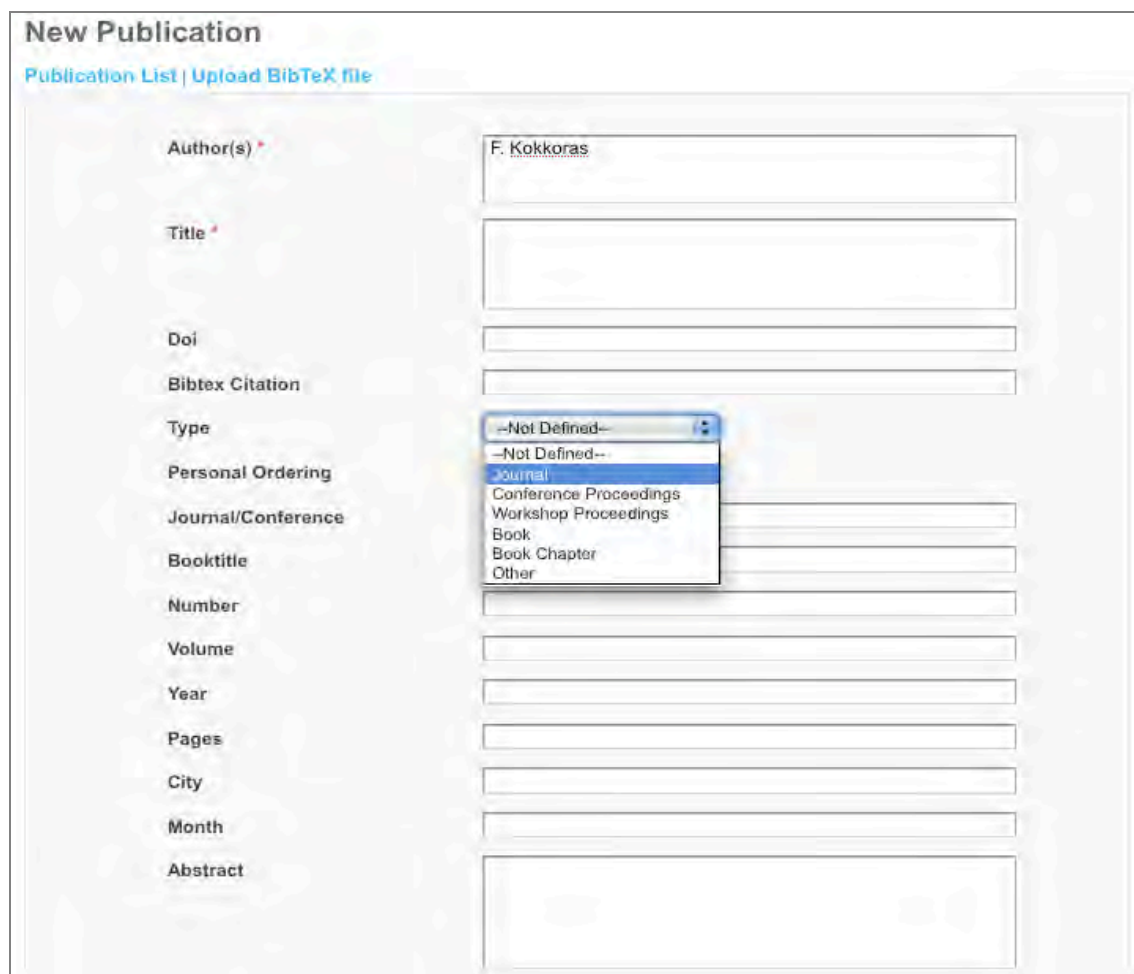
Στο τρέχον σενάριο θα παρουσιαστεί η δυνατότητα χρήσης του συστήματος ανεξάρτητα από τη λειτουργία αναζήτησης ετεροαναφορών, καθώς εδώ θεωρείται δεδομένη η γνώση της συσχέτισης των δημοσιεύσεων του χρήστη με τις ετεροαναφορές. Πιο συγκεκριμένα, στο σενάριο αυτό θεωρείται ότι ο χρήστης ενημερώνεται για την ύπαρξη ετεροαναφορών στα έργα του μέσω άλλων μεθόδων πέρα από την αναζήτηση μέσω Internet, όπως ενημέρωση από τον εκδότη ή τον συγγραφέα της ετεροαναφοράς, ενημέρωση από κάποιον συνάδελφο κ.α. Για να έχει νόημα η δημιουργία ετεροαναφορών είναι πρωτίστως απαραίτητη η δημιουργία δημοσιεύσεων.

Ανάλογα με τη μορφή που έχει ο χρήστης στα χέρια του τα δεδομένα, δηλαδή αρχείο BibTeX ή απλό κείμενο θα πρέπει να επιλέξει τη δημιουργία νέας δημοσίευσης «χειροκίνητα» ή την αυτόματη δημιουργία μέσω αρχείου.



Εικόνα 27. Αυτόματη ή χειροκίνητη δημιουργία δημοσίευσης

Την πρώτη φορά που χρησιμοποιεί κάποιος τη διαχείριση δημοσιεύσεων ο αντίστοιχος πίνακας είναι κενός και το σύστημα εμφανίζει το μήνυμα «*No publications found. Please upload a BibTeX file or create one manually.*» (Εικόνα 27). Αν επιλέξει τη χειροκίνητη δημιουργία θα εμφανιστεί μια φόρμα με όλα τα πεδία μιας βιβλιογραφικής αναφοράς στην οποία το μοναδικό συμπληρωμένο πεδίο είναι το όνομα του χρήστη ως author (Εικόνα 28), όπως αυτό έχει δηλωθεί κατά την εγγραφή.



Εικόνα 28. Φόρμα δημιουργίας δημοσίευσης

Στην περίπτωση αυτόματης δημιουργίας μέσω αρχείου BibTeX ο χρήστης πρέπει αρχικά να κάνει κλικ στην επιλογή «**Upload BibTeX File**» που βρίσκεται στο μενού διαχείρισης δημοσιεύσεων, στη συνέχεια να επιλέξει το αρχείο που περιέχει τις αναφορές οι οποίες τον ενδιαφέρουν μέσω της λειτουργίας «**Browse...**» και να πατήσει το πλήκτρο «**submit**» (Εικόνα 29) προκειμένου να μεταβεί στη λίστα επιλογής των δημοσιεύσεων που θα εισαχθούν (Εικόνα 31).

Εικόνα 29. Φόρμα επιλογής αρχείου BibTeX

Η λίστα εισαγωγής των δημοσιεύσεων του αρχείου BibTeX δίνει τη δυνατότητα στον χρήστη να επιλέξει τις δημοσιεύσεις που επιθυμεί να εισάγει επιλέγοντας τη λειτουργία «**Import Selected(#)**». Σε περίπτωση που επιθυμεί να εισάγει το σύνολο των εγγραφών, ανεξαρτήτως επιλογής, μπορεί να επιλέξει τη λειτουργία «**Import All**». Μετά την εισαγωγή το σύστημα μεταφέρει τον χρήστη αυτόματα στη λίστα των δημοσιεύσεων όπου μπορεί να δει λεπτομέρειες γι' αυτές και να τις επεξεργαστεί κατά το δοκούν (Εικόνα 32).

Η διαδικασία δημιουργίας ετεροαναφορών χειροκίνητα και μέσω αρχείου είναι ακριβώς η ίδια με αυτή των δημοσιεύσεων – δεδομένου ότι οι ετεροαναφορές είναι κι αυτές δημοσιεύσεις τρίτων – και γι' αυτόν το λόγο δε θα παρουσιαστεί στο παρόν σενάριο. Το μόνο σημείο διαφοροποίησης είναι ότι στις εγγραφές των ετεροαναφορών, δίνεται η δυνατότητα συσχέτισης τους με τις εγγραφές των δημοσιεύσεων. Αυτό φαίνεται ξεκάθαρα στην Εικόνα 30, όπου στο πεδίο «**Cited publication(s)**» δίνεται η δυνατότητα επιλογής δημοσιεύσεων από αυτές που είναι καταχωρημένες στο σύστημα.

Εικόνα 30. Συσχέτιση ετεροαναφορών με δημοσιεύσεις

Import BibTeX file entries

Select All | Select None

Import Selected (4) | Import All (ignores selection)

Items 1 - 20 of 35

1 2 Next >

<input checked="" type="checkbox"/>	[BergD:2006] Simon K. Johnston and Daniel Berg, "Method and Apparatus for Dynamic Script Execution in an Interactive Development Environment", 2006
<input checked="" type="checkbox"/>	[Brown2005] A. W. Brown and M. Delbaere and P. Eeles and S. Johnston and R. Weaver, "Realizing service-oriented solutions with the IBM Rational Software Development Platform", IBM Systems Journal, 2005
<input type="checkbox"/>	[Brown2006] Alan Brown and Sridhar Iyengar and Simon Johnston, "A Rational Approach to Model-Driven Development", IBM Systems Journal, July/September 2006
<input type="checkbox"/>	[Brown2006A] Alan Brown and Simon Johnston, "A Model-Driven Development Approach to Creating Service-Oriented Solutions", 11 2006
<input type="checkbox"/>	[BrownAW:dW:2003] Alan Brown and Simon Johnston and Kevin Kelly, "Using Service-Oriented Architecture and Component-Based Development to build Web Service Applications", 2003
<input checked="" type="checkbox"/>	[BKP:ibm:2005] Alan W. Brown and Simon K. Johnston and Grant Larsen and Jim Palistrant, "SOA Development Using the IBM Rational Software Development Platform: A Practical Guide", September 2005
<input type="checkbox"/>	[JohnstonSK:RA:1999] Magnus Christerson and Simon Johnston, "Business Modeling and E-Commerce", Rose Architect, 1999
<input checked="" type="checkbox"/>	[JohnstonSK:dW:2006a] Simon Johnston, "Developing a language feature for Rational Software Architect", March 2006
<input type="checkbox"/>	[Johnston2006A] Simon Johnston, "UML 2.0 Profile for Software Services", April 2006
<input type="checkbox"/>	[JohnstonSK:dW:2006b] Simon Johnston, "Modeling Web Services, Part 2: Modeling and generating WSDL", April 2006
<input type="checkbox"/>	[JohnstonSK:tre:2005] Simon Johnston, "Modeling Service-Oriented Solutions", The Rational Edge, July 2005
<input type="checkbox"/>	[JohnstonSK:dW:2005b] Simon Johnston, "Modeling Web Services, Part 1 - XML Schema", November 2005
<input type="checkbox"/>	[JohnstonSK:dW:2005a] Simon Johnston, "UML Profile for Software Services, RSA Plug-In", May 2005
<input type="checkbox"/>	[JohnstonSK:dW:2005] Simon Johnston, "UML 2.0 Profile for Software Services", April 2005
<input type="checkbox"/>	[JohnstonSK:dW:2004a] Simon Johnston, "Rational UML Profile for Business Modeling", 2004
<input type="checkbox"/>	[JohnstonSK:dW:2004] Simon Johnston, "Modeling security concerns in service-oriented architectures", 2004
<input type="checkbox"/>	[JohnstonSK:IBM] Simon Johnston, "IBM developerWorks Blog", 2003-
<input type="checkbox"/>	[JohnstonSK:MedBiq:2003] Simon Johnston, "MedBiquitous Software Architecture Document", November 2003
<input type="checkbox"/>	[JohnstonSK:Ratl:2003] Simon Johnston, "UML Profile for Web Services", April 2003
<input type="checkbox"/>	[JohnstonSK:dW:2002] Simon Johnston, "Automatic WinForm: A .NET Pattern for Creating a Visual C# WinForm in Rational XDE Release 2", 2002

Items 1 - 20 of 35

1 2 Next >

Εικόνα 31. Λίστα επιλογής δημοσιεύσεων προς εισαγωγή

Publications List

[New Publication](#) | [Upload BibTeX file](#) | [View uploaded file entries](#)

Items 1 - 10 of 35

1 2 3 4 Next >

Personal Ordering	Author(s)	Title	Type	Journal/Conference	Year	Actions
-	Simon K. Johnston and Daniel Berg	Method and Apparatus for Dynamic Script Execution in an Interactive Development Environment	--Not Defined--		2006	Details Edit Delete
-	Alan Brown and Sridhar Iyengar and Simon Johnston	A Rational Approach to Model-Driven Development	--Not Defined--	IBM Systems Journal	2006	Details Edit Delete
-	Alan Brown and Simon Johnston	A Model-Driven Development Approach to Creating Service-Oriented Solutions	--Not Defined--		2006	Details Edit Delete
-	Simon Johnston	Developing a language feature for Rational Software Architect	--Not Defined--		2006	Details Edit Delete
-	Simon Johnston	UML 2.0 Profile for Software Services	--Not Defined--		2006	Details Edit Delete
-	Simon Johnston	Modeling Web Services, Part 2: Modeling and generating WSDL	--Not Defined--		2006	Details Edit Delete
-	Simon Johnston and Ali Arsanjani and John Smith	IBM RUP for Service-Oriented Modeling and Architecture V2.4	--Not Defined--		2006	Details Edit Delete
-	Simon Johnston and Pablo Irassar and Martin Nally	BOM2P Core Specification	--Not Defined--		2006	Details Edit Delete
-	A. W. Brown and M. Delbaere and P. Eeles and S. Johnston and R. Weaver	Realizing service-oriented solutions with the IBM Rational Software Development Platform	--Not Defined--	IBM Systems Journal	2005	Details Edit Delete
-	Alan W. Brown and Simon K. Johnston and Grant Larsen and Jim Palistrant	SOA Development Using the IBM Rational Software Development Platform: A Practical Guide	--Not Defined--		2005	Details Edit Delete

Items 1 - 10 of 35

1 2 3 4 Next >

Εικόνα 32. Λίστα δημοσιεύσεων

6 Συμπεράσματα & Επεκτάσεις

6.1 Συμπεράσματα

Στην παρούσα μεταπτυχιακή διατριβή παρουσιάστηκε μια λίγο διαφορετική από τις τρέχουσες και ενδιαφέρουσα – κατά τη γνώμη μας – προσέγγιση του προβλήματος εύρεσης και οργάνωσης ετεροαναφορών σε έργα ενός ερευνητή.

Αρχικά, προκειμένου να γίνει απόλυτα διακριτή η ανάγκη δημιουργίας του συστήματος όπως αυτό προτείνεται στη διατριβή, παρουσιάστηκε ο ρόλος των βιβλιογραφικών αναφορών και η σημασία που αυτές έχουν για τον ερευνητή και την ερευνητική διαδικασία εν γένει. Στη συνέχεια δόθηκε μια σύντομη περιγραφή της τρέχουσας κατάστασης στην αναζήτηση βιβλιογραφικών αναφορών, εστιάζοντας στις δυσκολίες που αυτή συνεπάγεται για τον ερευνητή. Έχοντας περιγράψει την τρέχουσα κατάσταση πραγματοποιείται η περιγραφή του προβλήματος προκειμένου να καταστεί πιο σαφής η λύση, η οποία και προτείνεται στη συνέχεια. Η ανάγκη δημιουργίας του συστήματος διαγράφεται εντονότερα, όταν παρουσιάζονται τα συστήματα που υπάρχουν ήδη και σχετίζονται με το προτεινόμενο, καθώς εμφανίζουν ομοιότητες σε ορισμένα σημεία αλλά και ουσιαστικές διαφορές.

Οι διαφορές με τις υπόλοιπες προσεγγίσεις βρίσκονται κατά κύριο λόγο στο ότι κανένα άλλο σύστημα δεν προσφέρει ταυτόχρονα υπηρεσίες αναζήτησης και διαχείρισης στοχευμένες σε προσωπικές δημοσιεύσεις και στο ότι τα περισσότερα (εξαιρείται το RefWorks) είναι desktop εφαρμογές ή plugin σε browser. Συγκεκριμένα, τα EndNote, Zotero και RefWorks αποτελούν περισσότερο εφαρμογές βάσης δεδομένων βιβλιογραφικών αναφορών, που σαν στόχο έχουν τη συγκέντρωση βιβλιογραφικών αναφορών χειροκίνητα προς βοήθεια του ερευνητή στο συγγραφικό του έργο και λιγότερο λογισμικό αναζήτησης και οργάνωσης προσωπικών δημοσιεύσεων και ετεροαναφορών σε αυτές με στόχο την παρακολούθηση της προσωπικής προόδου και απήχησης (impact). Στις διαφορές αυτές οφείλεται και η ονομασία του λογισμικού «On-line Σύστημα Εύρεσης και Διαχείρισης Βιβλιογραφικών Αναφορών» ή στα Αγγλικά «Citations Management System».

Το Publish or Perish από την άλλη, δεν διαθέτει καμία λειτουργία δημιουργίας/επεξεργασίας και αποθήκευσης αναφορών, αλλά εστιάζει αποκλειστικά στην απήχηση του ερευνητικού έργου κάποιου ερευνητή και στην πλήρως αυτόματη και συνεπώς ιδιαίτερα επισφαλή μέτρηση των ετεροαναφορών των δημοσιεύσεών του με τη χρήση του Google Scholar. Η πλήρως αυτόματη μέτρηση των ετεροαναφορών, όπως αυτή παρέχεται από το Publish or Perish θεωρείται επισφαλής καθώς συνήθως πέραν των πραγματικών ετεροαναφορών τρίτων, προσμετρούνται και οι ετεροαναφορές που έχουν γίνει από έργα του ιδίου, με αποτέλεσμα ο αριθμός που προκύπτει να μην ανταποκρίνεται πλήρως στην πραγματικότητα.

Ένα άλλο σημαντικό στοιχείο του «Citations Management System» είναι η ύπαρξη μνήμης αναφορικά με την επεξεργασία των αποτελεσμάτων που προκύπτουν από τις αναζητήσεις. Έχοντας διεκπεραιώσει το βαρύ φορτίο της αρχικής εκκαθάρισης των αποτελεσμάτων, ο χρήστης δεν είναι αναγκασμένος να ξανά-ασχοληθεί με αυτά τα αποτελέσματα στο μέλλον και ως εκ τούτου, ο εντοπισμός νέων ετεροαναφορών πραγματοποιείται πολύ ευκολότερα. Το σύστημα διευκολύνει την αρχική εκκαθάριση των αποτελεσμάτων αναζήτησης υπολογίζοντας και παρουσιάζοντας μια πρόταση σχετικά με το είδος του αποτελέσματος που βρέθηκε, αξιοποιώντας με αυτόν τον τρόπο τόσο την πληροφορία που παρέχεται από το Google Scholar, όσο και την πληροφορία των δημοσιεύσεων που έχει ήδη αποθηκεύσει στο σύστημα ο χρήστης.

Αξίζει να σημειωθεί πως το πρόβλημα της αναζήτησης αναφορών αποτελεί ένα πολύπλοκο πρόβλημα με πολλές πτυχές το οποίο απασχολεί για μεγάλο χρονικό διάστημα τον ερευνητικό κόσμο. Είναι σαφές ότι με τη διάδοση του Παγκόσμιου Ιστού έχουν σημειωθεί αλματώδεις πρόοδοι στο θέμα, παρόλα αυτά η δημιουργία ενός συστήματος που θα το επιλύει σε ικανοποιητικό βαθμό αυτοματοποιώντας τη διαδικασία απαιτεί πολύ χρόνο, αρκετούς οικονομικούς πόρους και απρόσκοπτη συνεργασία "ξένων" μεταξύ τους συστημάτων που πολλές φορές ανήκουν σε διαφορετικούς φορείς που έχουν για τα συστήματά τους διαφορετικά πλάνα.

6.2 Επεκτάσεις – Μελλοντική Εργασία

Όπως είναι φυσικό και κυρίως λόγω χρονικών περιορισμών, το σύστημα που αναπτύχθηκε επιδέχεται βελτιώσεις και επεκτάσεις. Αυτές που κρίνονται ότι θα μπορούσαν να καταστήσουν την εφαρμογή ακόμη πιο χρήσιμη και λειτουργική, παρουσιάζονται στη συνέχεια.

Κλιμάκωση υπηρεσιών

Μια προσθήκη που θα μπορούσε να προσθέσει στη λειτουργικότητα του συστήματος είναι η δυνατότητα κλιμάκωσης των προσφερόμενων υπηρεσιών (scalability). Στη μορφή που είναι τώρα η εφαρμογή και λόγω των περιορισμών του Google Scholar, δεν είναι δυνατή η ταυτόχρονη αναζήτηση από περισσότερους του ενός χρήστες. Αυτό συμβαίνει για το λόγο ότι το ΔΕiXTo δεν έχει τη δυνατότητα να εκτελείται ταυτόχρονα από πολλούς χρήστες. Με δεδομένο ότι μια τυπική εκτέλεση μπορεί να διαρκεί και 10 λεπτά (ή περισσότερο), γίνεται εύκολα αντιληπτό πως η κλιμάκωση αποτελεί πρώτη προτεραιότητα για την ευρεία χρήση του συστήματος.

Το πρόβλημα μπορεί να λυθεί έμμεσα βέβαια (όπως ήδη αναφέρθηκε) με εκτέλεση του ΔΕiXTo στον Η/Υ του χρήστη και upload των αποτελεσμάτων στην εφαρμογή.

Χρήση και άλλων εξειδικευμένων μηχανών αναζήτησης

Η αξιοποίηση των αποτελεσμάτων αναζήτησης και άλλων εξειδικευμένων στις βιβλιογραφικές αναφορές μηχανών αναζήτησης θα επέτρεπε στον χρήστη να βρίσκει ετεροαναφορές στα έργα του με τη βοήθεια και άλλων μηχανών αναζήτησης. Αυτό τεχνικά είναι εφικτό δημιουργώντας νέους κανόνες εξαγωγής (extraction rules) για το ΔΕiXTo, ώστε να μπορούν να αντληθούν τα συγκεκριμένα αποτελέσματα.

Upload αρχείων pdf

Μια λειτουργία η οποία θα ήταν χρήσιμη σε πλήθος ερευνητών είναι η δυνατότητα αποθήκευσης στο σύστημα των δημοσιεύσεων σε pdf μορφή. Έτσι, ο κάθε χρήστης θα μπορούσε να έχει το πλήρες κείμενο τόσο των προσωπικών του δημοσιεύσεων όσο και των ετεροαναφορών σε μορφή pdf και να είναι διαθέσιμες από οποιοδήποτε σημείο οποτεδήποτε ο ίδιος το θελήσει χωρίς να είναι υποχρεωμένος να μεταφέρει συνεχώς τα εν λόγω αρχεία μαζί του.

Μια επιπλέον και ιδιαίτερα σημαντική λειτουργικότητα που θα μπορούσε να προσδώσει στο σύστημα η αποθήκευση αρχείων pdf είναι η πλήρης δεικτοδότηση του κειμένου αυτών και συνεπακόλουθα η δυνατότητα αναζήτησης σε αυτά.

Τα παραπάνω βέβαια συγκρούονται με ζητήματα πνευματικών δικαιωμάτων των δημοσιεύσεων καθώς αυτά ανήκουν κατά κύριο λόγο σε εκδοτικούς οίκους. Θα μπορούσε ίσως να γίνει για καθαρά προσωπική χρήση και υπηρεσίες αναζήτησης.

Υποστήριξη και άλλων κωδικοποιήσεων βιβλικών αναφορών

Στην παρούσα έκδοση, το σύστημα υποστηρίζει την αυτόματη δημιουργία εγγραφών δημοσιεύσεων και ετεροαναφορών μόνο με τη χρήση αρχείων BibTeX. Προκειμένου να βελτιωθεί η διαλειτουργικότητα του συστήματος με άλλα πακέτα λογισμικού θα ήταν χρήσιμη η υποστήριξη και άλλων μορφών αρχείων εισόδου όπως EndNote, Medline, RIS, PubMed κ.α. Αυτό θα επέτρεπε τη διείσδυση σε μεγαλύτερο μέρος χρηστών καθώς θα εξασφάλιζε τη συμβατότητα με άλλα παρεμφερή συστήματα όπως τα EndNote, Zotero, RefWorks κλπ.

Εξαγωγή λίστας δημοσιεύσεων σε αρχείο

Η χρήση των αρχείων BibTeX στην τρέχουσα έκδοση της εφαρμογής, περιορίζεται στη δημιουργία ετεροαναφορών και δημοσιεύσεων. Προκειμένου όμως το σύστημα να αποτελέσει μια ολοκληρωμένη πρόταση οργάνωσης βιβλιογραφικών αναφορών θα πρέπει να δίνει τη δυνατότητα στο χρήστη να εξάγει τη λίστα των δημοσιεύσεων σε αρχεία κατάλληλου τύπου για είσοδο σε άλλα συστήματα. Αρχεία τέτοιας μορφής θα μπορούσαν να είναι τα BibTeX, EndNote, Medline, RIS, PubMed κ.α.

Εμφάνιση στατιστικών στοιχείων

Κατά γενική ομολογία, τα στατιστικά στοιχεία είναι απαραίτητα για την μελέτη διαφόρων φαινομένων σε βάθος χρόνου. Ένα από τα στατιστικά στοιχεία που θα μπορούσαν να βοηθήσουν τον ερευνητή είναι η παρακολούθηση της δημοτικότητας των δημοσιεύσεων του. Έτσι το σύστημα θα μπορούσε να παράγει διαγράμματα στα οποία θα φαίνεται το πόσο δημοφιλής ήταν μία δημοσίευση στην πάροδο των χρόνων. Άλλα χρήσιμα στατιστικά στοιχεία για τον ερευνητή είναι:

- Ο συνολικός αριθμός δημοσιεύσεων και ετεροαναφορών
- Ο μέσος όρος ετεροαναφορών ανά δημοσίευση
- Ο μέσος όρος ετεροαναφορών ανά έτος
- Ο δείκτης h του Hirsch [29] και
- Ο δείκτης g του Egghe [30]

Προγραμματισμός αυτόματης αναζήτησης

Μια λειτουργία η οποία θα μπορούσε να βελτιώσει σε μεγάλο βαθμό την ευχρηστία της εφαρμογής είναι η δυνατότητα προγραμματισμού αναζήτησης βιβλιογραφικών αναφορών. Θα μπορούσε για παράδειγμα να δίνεται η δυνατότητα στους χρήστες να προγραμματίζουν την αναζήτηση σε ώρες που τους βολεύει καλύτερα επιτρέποντάς τους να μη σπαταλούν χρόνο αναμένοντας τα αποτελέσματα της. Επίσης μια τέτοια δυνατότητα θα ήταν χρήσιμη και για την παρακολούθηση των ετεροαναφορών καθώς το σύστημα θα μπορούσε να προγραμματιστεί έτσι ώστε να ελέγχει ανά τακτά χρονικά διαστήματα την ύπαρξη νέων αποτελεσμάτων.

Δυνατότητα δημιουργίας προσωπικών αναφορών εκτύπωσης

Ένα σημείο στο οποίο θα μπορούσαν να γίνουν αρκετές βελτιώσεις έτσι ώστε το σύστημα να ταιριάζει απόλυτα στο προφίλ του κάθε χρήστη, είναι η δυνατότητα δημιουργίας προσωπικών αναφορών εκτύπωσης. Αυτό σημαίνει ότι ο χρήστης θα μπορεί να επιλέγει τα πεδία που θα εμφανίζονται στις αναφορές, τον τρόπο ταξινόμησης των εγγραφών καθώς και το στυλ σύνταξης βιβλιογραφικών αναφορών. Μερικά από τα πιο διαδεδομένα στυλ, τα οποία υποστηρίζονται από τα περισσότερα εμπορικά προγράμματα είναι τα APA, Chicago/Turabian, Harvard, MLA κ.α.

Βιβλιογραφία

- [1] Metasearch Engines on Wikipedia:
http://en.wikipedia.org/wiki/Metasearch_engine
- [2] CiteSeer on Wikipedia: <http://en.wikipedia.org/wiki/CiteSeer>
- [3] CiteSeer^x on Wikipedia: <http://en.wikipedia.org/wiki/CiteSeerX>
- [4] CiteSeer^x home page: <http://citeseerx.ist.psu.edu/>
- [5] Scirus on Wikipedia: <http://en.wikipedia.org/wiki/Scirus>
- [6] Scirus home page: <http://www.scirus.com/>
- [7] Gale Reference Reviews on Scirus:
<http://www.gale.cengage.com/reference/peter/200806/scirus.htm>
- [8] Elsevier, the world's largest publisher of medical and scientific literature:
<http://en.wikipedia.org/wiki/Elsevier>
- [9] Microsoft Libra on Wikipedia:
[http://en.wikipedia.org/wiki/Libra_\(Academic_Search\)](http://en.wikipedia.org/wiki/Libra_(Academic_Search))
- [10] Microsoft Libra home page: <http://libra.msra.cn/>
- [11] Google Scholar on Wikipedia: http://en.wikipedia.org/wiki/Google_Scholar
- [12] Google Scholar home page: <http://scholar.google.gr/>
- [13] Alireza Noruzi, "Google Scholar: The New Generation of Citation Indexes", *Libri*, 2005, vol. 55, pp. 170–180
- [14] DBLP home page: <http://dblp.uni-trier.de/>
- [15] DBLP on Wikipedia: <http://en.wikipedia.org/wiki/DBLP>
- [16] Dongwon Lee , Byung-Won On , Jaewoo Kang , Sanghyun Park, " Effective and scalable solutions for mixed and split citation problems in digital libraries", *Proceedings of the 2nd international workshop on Information quality in information systems*, 2005.
- [17] Tan, Y. F., Kan, M.-Y., Lee, D., "Search Engine Driven Author Disambiguation", *Proceedings of the ACM/IEEE-CS Joint Conference on Digital Libraries*, 2006.

- [18] A. Monge and C. Elkan, "The field matching problem: Algorithms and applications", In: Proc. 2nd Internat. Conference Knowledge Discovery and Data Mining, 1996.
- [19] Pasula, H., Marthi, B., Milch, B., Russell, S., and Shpitser I., "Identity uncertainty and citation matching", In Advances in Neural Processing Systems 15. Vancouver, British Columbia: MIT Press, 2002.
- [20] EndNote home page: <http://www.endnote.com/>
- [21] EndNote on Wikipedia: <http://en.wikipedia.org/wiki/EndNote>
- [22] Zotero home page: <http://www.zotero.org/>
- [23] Zotero on Wikipedia: <http://en.wikipedia.org/wiki/Zotero>
- [24] RefWorks home page: <http://www.refworks.com/>
- [25] RefWorks on Wikipedia: <http://en.wikipedia.org/wiki/RefWorks>
- [26] Publish or Perish home page: <http://www.harzing.com/pop.htm>
- [27] Publish or Perish on Wikipedia: http://en.wikipedia.org/wiki/Publish_or_perish
- [28] Reference management software comparison on Wikipedia:
http://en.wikipedia.org/wiki/Comparison_of_reference_management_software
- [29] J.E. Hirsch, "An index to quantify an individual's scientific research output", arXiv:physics/0508025 v5 29, Sep 2005
- [30] Egghe, "Theory and practice of the g-index", Scientometrics, Vol. 69, No 1 (2006), pp. 131-152.
- [31] Alan Dennis, Barbara Haley Wixon, David Tegarden "System analysis and design with UML version 2.0 - Second Edition", Willey International Edition, 2005.
- [32] ΔEiXTo home page <http://deixto.csd.auth.gr/>
- [33] MySQL home page <http://www.mysql.com/>
- [34] LAMP on Wikipedia: [http://en.wikipedia.org/wiki/LAMP_\(software_bundle\)](http://en.wikipedia.org/wiki/LAMP_(software_bundle))
- [35] Eclipse open development platform home page: <http://www.eclipse.org/>
- [36] Yii Web Programming Framework: <http://www.yiiframework.com/>
- [37] jQuery JavaScript Library: <http://jquery.com/>
- [38] Ζιάκα Εύα, Σύστημα Εύρεσης και Διαχείρισης Βιβλιογραφικών Αναφορών, Πτυχιακή Εργασία, Τμήμα Πληροφορικής ΑΠΘ, 2008.