



ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

**ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ**

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

**Διαδραστική Οπτικοποίηση Δεδομένων Ερευνητικής
Δραστηριότητας του Τμήματος Πληροφορικής και
Τηλεπικοινωνιών του Εθνικού και Καποδιστριακού
Πανεπιστημίου Αθηνών**

Παναγιώτης Θ. Ξένος

Επιβλέπουσα: Μαρία Ρούσσου, Επίκουρη Καθηγήτρια

**ΑΘΗΝΑ
ΑΥΓΟΥΣΤΟΣ 2020**

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

Διαδραστική Οπτικοποίηση Δεδομένων Ερευνητικής Δραστηριότητας του Τμήματος
Πληροφορικής και Τηλεπικοινωνιών του Εθνικού Καποδιστριακού Πανεπιστημίου
Αθηνών

Παναγιώτης Θ. Ξένος

A.M.: 1115201200131

Επιβλέπουσα: Μαρία Ρούσσου, Επίκουρη Καθηγήτρια

ΠΕΡΙΛΗΨΗ

Στο χώρο της έρευνας, η πληθώρα δημοσιεύσεων ερευνητικού περιεχομένου δυσχεραίνει την αναζήτηση της απαραίτητης πληροφορίας από τους ερευνητές. Η ανάγκη για πιο γρήγορους και αποδοτικούς τρόπους αναζήτησης της απαιτούμενης πληροφορίας φέρνει στο προσκήνιο νέους τρόπους αναζήτησης και αλληλεπίδρασης με τον χρήστη, όπως οι διαδραστικές οπτικοποιήσεις. Τα τελευταία χρόνια, έχει υπάρξει ραγδαία ανάπτυξη στις μεθόδους αναζήτησης και οπτικοποίησης ερευνητικών δημοσιεύσεων. Παρόλα αυτά, μέχρι σήμερα, δεν υπάρχει κάποια διαδραστική οπτικοποίηση η οποία να αναδεικνύει την ερευνητική δραστηριότητα του Τμήματος Πληροφορικής και Τηλεπικοινωνιών του Εθνικού Καποδιστριακού Πανεπιστημίου Αθηνών.

Η συγκεκριμένη πτυχιακή εργασία, έρχεται να καλύψει το κενό αυτό προτείνοντας μία εναλλακτική διαδραστική οπτικοποίηση από την οποία ο χρήστης μπορεί να έχει πρόσβαση στην ερευνητική δραστηριότητα του παραπάνω Τμήματος. Η οπτικοποίηση που προτείνεται, πέρα από εφαρμογή αναζήτησης ερευνητικών δημοσιεύσεων, μπορεί να χρησιμοποιηθεί προκειμένου να δώσει μία πιο σφαιρική εικόνα των πραγμάτων αναφορικά με τις ερευνητικές δημοσιεύσεις του εν λόγω Τμήματος.

Σε τέτοιου είδους εφαρμογές, η συλλογή των δεδομένων αποτελεί μία μεγάλη πρόκληση. Στη συγκεκριμένη πτυχιακή εργασία το πρόβλημα αυτό αντιμετωπίζεται, χρησιμοποιώντας έναν Crawler γραμμένο σε Python για την αυτόματη συλλογή των δεδομένων.

Ακολουθώντας μια μεθοδολογία επαναληπτικού (ελικοειδούς) σχεδιασμού, ανάπτυξης και αξιολόγησης, στην εργασία αυτή εξετάζεται η ευχρηστία και η εμπειρία χρήσης της προτεινόμενης οπτικοποίησης, καθώς και κατά πόσο αποδοτική είναι για την αναζήτηση και διαχείριση τέτοιου είδους δεδομένων. Κατά τη διάρκεια της σχεδίασης και της υλοποίησης έλαβαν χώρα τρεις (3) αξιολογήσεις, στις οποίες συμμετέχουν φοιτητές και ερευνητές του συγκεκριμένου Τμήματος. Τα αποτελέσματα της αξιολόγησης καθοδήγησαν τις σχεδιαστικές αποφάσεις και τελικά την υλοποίηση της προτεινόμενης οπτικοποίησης.

ΘΕΜΑΤΙΚΗ ΠΕΡΙΟΧΗ: Διαδραστική Οπτικοποίηση Δεδομένων

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: Σχεδίαση Οπτικοποίησης, Tableau, Αξιολόγηση ευχρηστίας, Διαδραστική, Οπτικοποίηση, D3.js, Javascript, Python, Crawler, Selenium.

ABSTRACT

In the field of research, the vast amount of research publications complicates the acquisition and the extraction of the necessary information by the researchers. The need for faster and more efficient ways to search for information brings to the fore new ways of presenting and interacting with data about published research articles, such as interactive visualizations. In recent years, there has been a rapid growth regarding the methods of searching for and visualizing research publications. However, to date, there is no interactive visualization that highlights the research activity of the Department of Informatics and Telecommunications of the National Kapodistrian University of Athens (DIT).

This thesis is proposing an alternative interactive visualization in which the user can access data related to the research activity of DIT. The proposed visualization along with a search application for the research publications provide a more comprehensive picture of the details of the research publications of DIT.

After taking into consideration the fact that data collection is a big challenge in such applications, in this thesis this problem is tackled by using a Crawler written in Python for automatic data collection.

Using an iterative design, development and evaluation methodology, this thesis examines the user-friendliness and user-experience of the proposed visualization and its efficiency in researching and managing such data. In the phase of designing and implementing the visualization three evaluations took place in which students and researchers of the DIT participated. The results of this evaluation guided the design decisions and finally the implementation of the proposed visualization.

SUBJECT AREA: Interactive Data Visualization

KEYWORDS: Visualization Design, Tableau, Usability Evaluation, Interactive, Visualization, D3.js, Javascript, Python, Crawler, Selenium.

ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να ευχαριστήσω ιδιαίτερα την επιβλέπουσα Επίκουρη Καθηγήτρια Μαρία Ρούσσου για την αδιάλειπτη υποστήριξη, την καθοδήγηση και την ενθάρρυνση που μου παρείχε κατά την διάρκεια της εκπόνησης της πτυχιακής εργασίας. Επίσης, θα ήθελα να την ευχαριστήσω για τη συμβολή της στην οργάνωση της διαδικασίας της αξιολόγησης, καθώς και για τη συνδρομή της κατά το σχεδιασμό του θέματος της εργασίας.

Στη συνέχεια, θα ήθελα να ευχαριστήσω την κ. Αθανασία Κολοβού, μέλος του Εργαστηριακού Διδακτικού Προσωπικού του Τμήματος για την υποστήριξη και την τεχνική βοήθεια που μου παρείχε πάνω στις τεχνολογίες που έπρεπε να μάθω για την υλοποίηση της εργασίας.

Επιπλέον, θα ήθελα να ευχαριστήσω τόσο τους φοιτητές όσο και τους ερευνητές που συμμετείχαν στη διαδικασία της αξιολόγησης, για το ενδιαφέρον που επέδειξαν, τον χρόνο που αφιέρωσαν και την πολύτιμη ανατροφοδότηση που παρείχαν.

Τέλος, θα ήθελα να ευχαριστήσω την οικογένειά μου, τους φίλους και την κοπέλα μου για την υπομονή τους, τη συνεχή ενθάρρυνση, καθώς και την καθημερινή τους συμπαράσταση και υποστήριξη καθ' όλη τη διάρκεια της συγγραφής της παρούσας εργασίας.

ΠΕΡΙΕΧΟΜΕΝΑ

ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ	8
ΚΑΤΑΛΟΓΟΣ ΕΙΚΟΝΩΝ	9
ΠΡΟΛΟΓΟΣ	12
1. ΕΙΣΑΓΩΓΗ	13
1.1 Η Ανάγκη για Οπτικοποίηση Δεδομένων	13
1.2 Το Αντικείμενο της Εργασίας	13
1.3 Το Πρόβλημα	14
1.4 Η Συμβολή της Εργασίας	15
1.5 Δομή Εργασίας	15
2. ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ	17
2.1 Η Οπτικοποίηση Δεδομένων	17
2.2 Ευχρηστία	18
2.3 Η Έννοια της Σχεδίασης Διαδραστικών Εφαρμογών	19
2.4 Μοντέλα Σχεδίασης Συστημάτων Λογισμικού	19
2.4.1 Σχεδίαση με επίκεντρο τον χρήστη (User-Centred Interactive Design)	20
2.4.2 Ελικοειδές μοντέλο (Spiral Model)	21
2.5 Πρότυπα Χαμηλής και Υψηλής Πιστότητας	22
2.6 Η Σημασία της Αξιολόγησης	22
2.6.1 Παρατήρηση του χρήστη	22
2.6.2 Μέτρηση της απόδοσης του χρήστη	23
2.6.3 Ερωτηματολόγια	23
3. ΙΣΤΟΡΙΚΗ ΕΞΕΛΙΞΗ	25
3.1 Οι Πρώτες Οπτικοποιήσεις	25
3.1.1 Dynamic Queries (1993 University of Maryland UIS Broadcast)	25
3.1.2 Gapminder Tools (2015)	27
3.1.3 Nobel Laureates Visualization (2018)	28
3.2 Παραδείγματα Οπτικοποιήσεων Ερευνητικών Δημοσιεύσεων	29
3.2.1 Google Scholar (Released in beta in November 2004)	29
3.2.2 Scopus (Launched in 2004)	30
3.3 Συμπεράσματα	32
4. Η ΣΥΛΛΟΓΗ ΤΩΝ ΔΕΔΟΜΕΝΩΝ	33
4.1 Τα Δεδομένα	33
4.1 Web Crawler (Spider)	34
4.1 Πρώτη Υλοποίηση Web Crawler	35

4.1	Δεύτερη Υλοποίηση Web Crawler	35
5.	Ο ΣΧΕΔΙΑΣΜΟΣ	39
5.1	Μεθοδολογία Σχεδιασμού	39
5.2	Σύνταξη Προδιαγραφών και Απαιτήσεων	39
5.3	Ο Πρώτος Σχεδιασμός της Οπτικοποίησης Χαμηλής Πιστότητας	40
5.4	Πρώτη Διαμορφωτική Αξιολόγηση - Αποτελέσματα	43
5.5	Ο Δεύτερος Σχεδιασμός της Οπτικοποίησης Υψηλής Πιστότητας	44
5.5.1	Τι Είναι το (Tableau)	44
5.5.2	Ο Σχεδιασμός στο Tableau - Φάση Φόρτωσης Δεδομένων	44
5.5.3	Ο Σχεδιασμός στο Tableau - Φάση Δημιουργίας Διαγραμμάτων	46
5.6	Δεύτερη Διαμορφωτική Αξιολόγηση με Χρήστες - Αποτελέσματα	49
5.7	Συμπεράσματα – Τελικός Σχεδιασμός	52
6.	ΥΛΟΠΟΙΗΣΗ	57
6.1	Τα Εργαλεία που Χρησιμοποιήθηκαν	57
6.2	Η Δημιουργία της Τελικής Οπτικοποίησης	58
6.3	Συμπερασματική Αξιολόγηση - Αποτελέσματα	59
7.	ΣΥΜΠΕΡΑΣΜΑΤΑ	63
7.1	Ανασκόπηση Όλης της Διαδικασίας	63
7.2	Τελικά Συμπεράσματα	63
8.	ΜΕΛΛΟΝΤΙΚΕΣ ΕΠΕΚΤΑΣΕΙΣ	65
	ΠΙΝΑΚΑΣ ΟΡΟΛΟΓΙΑΣ	67
	ΣΥΝΤΜΗΣΕΙΣ – ΑΡΚΤΙΚΟΛΕΞΑ – ΑΚΡΩΝΥΜΙΑ	68
	ΠΑΡΑΡΤΗΜΑ Ι	69
	ΠΑΡΑΡΤΗΜΑ ΙΙ	83
	ΑΝΑΦΟΡΕΣ	86

ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ

Σχήμα 1: Σχήμα ελικοειδούς μοντέλου.....	41
--	----

ΚΑΤΑΛΟΓΟΣ ΕΙΚΟΝΩΝ

Εικόνα 1: Απόσπασμα διαδραστικής οπτικοποίησης Dynamic Home Finder....σελ.	28
Εικόνα 2: Απόσπασμα οπτικοποίησης του Garpinder.....σελ.	29
Εικόνα 3: Οπτικοποίηση απονομών των βραβείων Νόμπελ που έχουν πραγματοποιηθεί μέχρι το 2018.....σελ.	31
Εικόνα 4: Απόσπασμα από το Google Scholar. Ραβδόγραμμα παραθέσεων ανά έτος για έναν συγκεκριμένο συγγραφέα.....σελ.	32
Εικόνα 5: Απόσπασμα από το Scopus. Διάγραμμα αριθμού εγγραφών ανά έτος.....σελ.	33
Εικόνα 6: Απόσπασμα από το Scopus. Διάγραμμα αριθμού εγγραφών ανά ερευνητική ομάδα.....σελ.	33
Εικόνα 7: Στιγμιότυπο οθόνης από την εκτέλεση του Crawler.....σελ.	40
Εικόνα 8: Ραβδόγραμμα συνολικού αριθμού δημοσιεύσεων ανά έτος - Πρότυπο Χαμηλής Πιστότητας.....σελ.	43
Εικόνα 9: Ραβδόγραμμα συνολικού αριθμού δημοσιεύσεων ανά ερευνητικό εργαστήριο - Πρότυπο Χαμηλής Πιστότητας.....σελ.	44
Εικόνα 10: Διάγραμμα σμήνους φυσαλίδων χρόνου ανά ερευνητικά εργαστήρια, με κάθε φυσαλίδα να είναι μια ερευνητική δημοσίευση - Πρότυπο Χαμηλής Πιστότητας.....σελ.	45
Εικόνα 11: Φόρτωση του αρχείου δεδομένων (outfile.csv) στο Tableau.....σελ.	47
Εικόνα 12: Δημιουργία ραβδογράμματος ερευνητικών δημοσιεύσεων του (DIT) ανά έτος, στο Tableau (χρωματική απεικόνιση ανάλογα με το είδος της δημοσίευσης).....σελ.	49
Εικόνα 13: Δημιουργία ραβδογράμματος συνολικών ερευνητικών δημοσιεύσεων ανά ερευνητικό εργαστήριο του (DIT), στο Tableau (χρωματική απεικόνιση ανάλογα με το είδος της δημοσίευσης).....σελ.	50
Εικόνα 14: Δημιουργία διαγράμματος φυσαλίδων ερευνητικών δημοσιεύσεων του (DIT) ανά έτος και ανά ερευνητικό εργαστήριο, στο Tableau (χρωματική απεικόνιση ανάλογα με το είδος της δημοσίευσης).....σελ.	51
Εικόνα 15: Στιγμιότυπο οθόνης από την δεύτερη διαμορφωτική αξιολόγηση...σελ.	53
Εικόνα 16: Απαντήσεις σχετικά με την κατανόηση των τριών διαγραμμάτων από τους χρήστες.....σελ.	54
Εικόνα 17: Διάγραμμα χρόνων ολοκλήρωσης σεναρίου χρήσης για 15 συμμετέχοντες στην δεύτερη αξιολόγηση.....σελ.	55
Εικόνα 18: Διάγραμμα πίτας. Απαντήσεις στην ερώτηση σχετικά με την χρήση των ραβδογραμμάτων ως φίλτρα.....σελ.	57
Εικόνα 19: Στιγμιότυπο οθόνης από την τελική οπτικοποίηση.....σελ.	60

Εικόνα 20: Διάγραμμα χρονών ολοκλήρωσης σεναρίου χρήσης για 10 συμμετέχοντες στην τρίτη αξιολόγηση.....	σελ. 62
Εικόνα 21: Στιγμιότυπο οθόνης από την χρήση φίλτρων στην τελική υλοποίηση της οπτικοποίησης - χρήση της καρτέλας «Δημοσιεύσεις ανά όνομα ερευνητικού εργαστηρίου και έτος».....	σελ. 64
Εικόνα 22: Στιγμιότυπο οθόνης από την χρήση φίλτρων στην τελική υλοποίηση της οπτικοποίησης - χρήση της καρτέλας «Όλες οι δημοσιεύσεις συνολικά».....	σελ. 64
Εικόνα 23: Παράδειγμα Εγγραφής Αρχείου Bibtex.....	σελ. 70
Εικόνα 24: Παράδειγμα Πρώτου Αρχείου Δεδομένων JSON.....	σελ. 70
Εικόνα 25: Παράδειγμα Αρχείου Δεδομένων CSV.....	σελ. 71
Εικόνα 26: Παράδειγμα Configuration File όπου ο Crawler παίρνει σαν όρισμα.....	σελ. 71
Εικόνα 27: Στιγμιότυπο οθόνης από την δεύτερη διαμορφωτική αξιολόγηση (εναλλακτική έκδοση πρωτότυπου υψηλής πιστότητας, που παρουσιάστηκε σε μερικούς χρήστες, με τα ραβδογράμματα από πάνω από το διάγραμμα φυσαλίδων).....	σελ. 72
Εικόνα 28: Στιγμιότυπο οθόνης από πρωτότυπο υψηλής πιστότητας που παρουσιάστηκε στους χρήστες στην δεύτερη διαμορφωτική αξιολόγηση – χρήση ραβδογραμμάτων σαν φίλτρα.....	σελ. 72
Εικόνα 29: Στιγμιότυπο οθόνης από πρωτότυπο υψηλής πιστότητας που παρουσιάστηκε στους χρήστες στην δεύτερη διαμορφωτική αξιολόγηση – χρήση φίλτρου επιλογής συγκεκριμένου είδους δημοσίευσης.....	σελ. 73
Εικόνα 30: Στιγμιότυπο οθόνης από την τελική οπτικοποίηση - χρήση της καρτέλας «Όλες οι δημοσιεύσεις συνολικά».....	σελ. 73
Εικόνα 31: Στιγμιότυπο οθόνης από την τελική οπτικοποίηση - χρήση της καρτέλας «Δημοσιεύσεις ανά είδος και έτος».....	σελ. 74
Εικόνα 32: Στιγμιότυπο οθόνης από την χρήση του φίλτρου «Επέλεξε Συγγραφέα» στην τελική οπτικοποίηση - χρήση της καρτέλας «Δημοσιεύσεις ανά είδος και έτος».....	σελ. 74
Εικόνα 33: Στιγμιότυπο οθόνης ερωτηματολογίου που χρησιμοποιήθηκε κατά την δεύτερη διαμορφωτική αξιολόγηση (1 από 9).....	σελ. 76
Εικόνα 34: Στιγμιότυπο οθόνης ερωτηματολογίου που χρησιμοποιήθηκε κατά την δεύτερη διαμορφωτική αξιολόγηση (2 από 9).....	σελ. 77
Εικόνα 35: Στιγμιότυπο οθόνης ερωτηματολογίου που χρησιμοποιήθηκε κατά την δεύτερη διαμορφωτική αξιολόγηση (3 από 9).....	σελ. 78
Εικόνα 36: Στιγμιότυπο οθόνης ερωτηματολογίου που χρησιμοποιήθηκε κατά την δεύτερη διαμορφωτική αξιολόγηση (4 από 9).....	σελ. 79
Εικόνα 37: Στιγμιότυπο οθόνης ερωτηματολογίου που χρησιμοποιήθηκε κατά την δεύτερη διαμορφωτική αξιολόγηση (5 από 9).....	σελ. 80

Εικόνα 38: Στιγμιότυπο οθόνης ερωτηματολογίου που χρησιμοποιήθηκε κατά την δεύτερη διαμορφωτική αξιολόγηση (6 από 9).....σελ. 81	81
Εικόνα 39: Στιγμιότυπο οθόνης ερωτηματολογίου που χρησιμοποιήθηκε κατά την δεύτερη διαμορφωτική αξιολόγηση (7 από 9).....σελ. 82	82
Εικόνα 40: Στιγμιότυπο οθόνης ερωτηματολογίου που χρησιμοποιήθηκε κατά την δεύτερη διαμορφωτική αξιολόγηση (8 από 9).....σελ. 82	82
Εικόνα 41: Στιγμιότυπο οθόνης ερωτηματολογίου που χρησιμοποιήθηκε κατά την δεύτερη διαμορφωτική αξιολόγηση (9 από 9).....σελ. 83	83
Εικόνα 42: Οπτικοποίηση Απονομών των Βραβείων Νόμπελ που Έχουν Πραγματοποιηθεί Μέχρι Σήμερα.....σελ. 84	84
Εικόνα 43: Dynamic Queries Home Finder - παράδειγμα χρήσης φίλτρου A & Bσελ. 84	84
Εικόνα 44: Στιγμιότυπο οθόνης από Dynamic Queries Home Finder - παράδειγμα οπτικοποίησης με μπάρα όταν δεν είναι διαθέσιμα γεωγραφικά δεδομένα.....σελ. 85	85
Εικόνα 45: Στιγμιότυπο οθόνης από το Scopus - παράδειγμα διαγράμματος πίτας εγγράφων ανά θεματική περιοχή.....σελ. 85	85
Εικόνα 46: Στιγμιότυπο οθόνης από το Scopus - παράδειγμα διαγράμματος εγγράφων ανά είδος.....σελ. 86	86
Εικόνα 47: Στιγμιότυπο οθόνης από το Scopus - παράδειγμα διαγράμματος εγγράφων ανά χρόνο χωρισμένα κατά προέλευση.....σελ. 86	86

ΠΡΟΛΟΓΟΣ

Η πτυχιακή εργασία που ακολουθεί εκπονήθηκε στο Τμήμα Πληροφορικής και Τηλεπικοινωνιών του Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών και έχει θέμα τη διαδραστική οπτικοποίηση δεδομένων ερευνητικής δραστηριότητας του Τμήματος αυτού. Επιβλέπουσα ήταν η Επίκουρη Καθηγήτρια Μαρία Ρούσσου.

1. ΕΙΣΑΓΩΓΗ

1.1. Η Ανάγκη για Οπτικοποίηση Δεδομένων

Στις ημέρες μας η χρήση του υπολογιστή και του κινητού και κατ' επέκταση η χρήση της οθόνης είναι καθημερινή και αποτελεί αναπόσπαστο κομμάτι της ζωής μας. Η επικοινωνία του ανθρώπου με τα ως άνω υπολογιστικά συστήματα γίνεται μέσω της οθόνης, του πληκτρολογίου, του ποντικιού ή κάποιων άλλων περιφερειακών συσκευών. Εφόσον η χρήση των συστημάτων αυτών είναι καθημερινή και απαραίτητη, οι χρήστες έχουν ανάγκη για γρήγορη εύρεση της απαιτούμενης πληροφορίας και άμεση επίτευξη των λειτουργιών που έχουν να φέρουν εις πέρας. Ως εκ τούτου, υπάρχει ανάγκη για εύχρηστες και αποδοτικές εφαρμογές που θα δίνουν στον χρήστη την ευελιξία και τον έλεγχο που χρειάζεται ανάλογα με τις ανάγκες του.

Ένας από τους πιο συχνούς και αποδοτικούς τρόπους μετάδοσης πληροφορίας είναι η οπτικοποίηση αυτής και η παρουσίασή της μέσω της οθόνης, δηλαδή χρησιμοποιώντας την εικόνα. Πολλές εταιρείες ανά τον κόσμο έχουν χρησιμοποιήσει τη μέθοδο της οπτικοποίησης στις εφαρμογές τους, προκειμένου να παρουσιάσουν πολύπλοκα δεδομένα με έναν πιο απλό και κατανοητό τρόπο από το να τα παρουσιάζουν σε πίνακες από δεδομένα. Μια οπτική και διαδραστική απεικόνιση δεδομένων είναι πιο ευχάριστη και εύχρηστη για τον χρήστη, ο οποίος τις περισσότερες φορές αποφεύγει να διαβάσει και να καταβάλει ιδιαίτερη προσπάθεια για να αλληλεπιδράσει με την εφαρμογή. Ένας ακόμα λόγος που η οπτικοποίηση είναι ένας καλός τρόπος παρουσίασης δεδομένων είναι ότι ο άνθρωπος αντιλαμβάνεται καλύτερα διαισθητικά τα σύμβολα, τις εικόνες και τα σχέδια από τα γράμματα και τους αριθμούς. Δεν είναι τυχαίο που υπάρχει η λαϊκή παροιμία «μια εικόνα χίλιες λέξεις».

Εν κατακλείδι, όταν υπάρχουν στη διάθεσή μας πολλά και πολύπλοκα δεδομένα ο καλύτερος τρόπος για να τα απλουστεύσουμε και να τα παρουσιάσουμε σε ένα χρήστη, ο οποίος δεν επιθυμεί να καταβάλει χρόνο και προσπάθεια, ώστε να τα κατανοήσει, είναι να τα οπτικοποιήσουμε.

1.2. Το Αντικείμενο της Εργασίας

Το αντικείμενο της εργασίας αφορά τη διαδραστική οπτικοποίηση δεδομένων. Συγκεκριμένα, παρουσιάζεται ο σχεδιασμός και η υλοποίηση μιας εφαρμογής οπτικοποίησης, η οποία επιτρέπει στον χρήστη να δει τα δεδομένα με διάφορους τρόπους ανάλογα με την πληροφορία που επιθυμεί να αντλήσει κάθε φορά. Τα δεδομένα είναι ερευνητικές δημοσιεύσεις (του Τμήματος Πληροφορικής και Τηλεπικοινωνιών του Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών) που υπάρχουν στο Google Scholar¹, μια από τις πιο ευρέως διαδεδομένες ιστοσελίδες για την εύρεση επιστημονικών δημοσιεύσεων. Σε αυτήν την εργασία δίνεται έμφαση στο σχεδιασμό, ώστε η εφαρμογή να είναι φιλική και εύχρηστη για τον χρήστη, συμπεριλαμβανομένων των ατόμων με δυσκολία διάκρισης χρωμάτων (π.χ. αχρωματοψία). Τέλος, αναδεικνύει μια διαφορετική οπτική των δεδομένων ανάλογα

¹ <https://scholar.google.com/>

με τις ανάγκες που υπάρχουν, δίνοντας στον χρήστη τη δυνατότητα να αλλάζει την οπτική αυτή με τη χρήση φίλτρων.

Το πρόβλημα που βλέπουμε παραπάνω δεν απουσιάζει από τον κόσμο της επιστήμης και της έρευνας. Με την όλο και αυξανόμενη επιστημονική πρόοδο, τα ερευνητικά κέντρα και τα πανεπιστήμια εκδίδουν συνεχώς δημοσιεύσεις με τις τελευταίες εξελίξεις στον τομέα τους και αυξάνουν σημαντικά τον όγκο των διαθέσιμων επιστημονικών δεδομένων. Ως εκ τούτου, οι ερευνητές πρέπει να καταναλώνουν πολύ χρόνο στην αναζήτηση δεδομένων, προκειμένου να διενεργήσουν μια απλή έρευνα.

Σύμφωνα με έρευνα που πραγματοποιήθηκε τον Ιούλιο του 2010 σε Πανεπιστήμιο της Οτάβα [1], ο αριθμός των ενεργών ακαδημαϊκών περιοδικών αυξάνεται σταθερά από το 1665. Το 1665 δημοσιεύθηκε στη Γαλλία το πρώτο μοντέλο του σύγχρονου περιοδικού. Το συγκεκριμένο περιοδικό ακολουθήθηκε από τις Φιλοσοφικές συζητήσεις, ενώ δημοσιεύθηκε από τη Βασιλική κοινότητα στο Λονδίνο αργότερα το ίδιο έτος. Από εκείνη την περίοδο μέχρι το έτος 2009, ο συνολικός αριθμός επιστημονικών δημοσιεύσεων έφτασε τα πενήντα (50) εκατομμύρια. Κάθε χρόνο δημοσιεύονται περίπου 2,5 εκατομμύρια νέα επιστημονικά έγγραφα.

Με βάση τα παραπάνω, είναι κατανοητό ότι το να ανεύρει κάποιος την πληροφορία που αναζητά, σχετικά με μια επιστημονική δημοσίευση ή κάποιον επιστημονικό τομέα, γίνεται όλο ένα και δυσκολότερο. Επομένως, δημιουργείται ανάγκη για πιο αποδοτικούς τρόπους προσέγγισης αυτών των δεδομένων. Με αφορμή την ανάγκη αυτή, γεννήθηκε η ιδέα της εν λόγω εργασίας. Σε μια τέτοια συνεχώς εξελισσόμενη επιστημονική πραγματικότητα, η εργασία αυτή καλείται να δώσει μια διαφορετική προσέγγιση στον τρόπο που αναζητούμε και επεξεργαζόμαστε το μεγάλο αυτόν όγκο επιστημονικών δεδομένων.

1.3. Το Πρόβλημα

Στην Ελλάδα και συγκεκριμένα στο Εθνικό Καποδιστριακό Πανεπιστήμιο, υπάρχει μεγάλη ερευνητική δραστηριότητα. Προς το παρόν δεν υπάρχει κάποιος εύκολος και διαδραστικός τρόπος, όπως μέσα από μια οπτικοποίηση, προκειμένου να αναζητήσει κάποιος ερευνητικές δημοσιεύσεις του παραπάνω τμήματος. Η αναζήτηση επιστημονικού περιεχομένου γίνεται κυρίως μέσα από εφαρμογές όπως το Scholar και το Scopus, όπου κάθε ερευνητής διαθέτει ένα λογαριασμό και προσθέτει την ερευνητική του δραστηριότητα, ώστε να είναι προσβάσιμη από άλλους ερευνητές, καθώς και οποιοδήποτε άλλο ενδιαφερόμενο. Στη συνέχεια της πτυχιακής εργασίας θα εξετάσουμε αναλυτικά τις εφαρμογές αυτές. Η αναπαράσταση των δεδομένων σε αυτές τις εφαρμογές γίνεται κυρίως σε μορφή λίστας και η εύρεση της δημοσίευσης γίνεται με μία εξελιγμένη και αρκετά εύχρηστη διεπαφή αναζήτησης.

Η παραπάνω μέθοδος με τη μορφή λίστας, δεν είναι τόσο αποδοτική και εύχρηστη, σε αντίθεση με μια διαδραστική εφαρμογή οπτικοποίησης που παρέχει στο χρήστη μια συνολική εικόνα και μια καλύτερη αντίληψη για το πού θα βρει την πληροφορία που αναζητά.

1.4. Η Συμβολή της Εργασίας

Στην ήδη υπάρχουσα κατάσταση που η προβολή των ερευνητικών δημοσιεύσεων γίνεται σε μορφή λίστας, η παρούσα εργασία έρχεται να προτείνει μια διαφορετική οπτική στην παρουσίαση αυτού του είδους δεδομένων. Η διαδραστική οπτικοποίηση των δεδομένων είναι αυτός ο διαφορετικός τρόπος παρουσίασης. Όπως αναφέρθηκε και παραπάνω μια εφαρμογή οπτικοποίησης, η οποία δεν είναι στατική αλλά διαδραστική, είναι πιο αποδοτική, καθώς ο χρήστης έχει τη δυνατότητα να αντιδρά δυναμικά πάνω στην αναπαράσταση της οπτικοποίησης κατά την διάρκεια εκτέλεσης της εφαρμογής. Η αλληλεπίδραση του χρήστη με την οπτικοποίηση τον βοηθάει να «εμπλακεί» με τον τρόπο προβολής των δεδομένων και κατά αυτόν τον τρόπο να τα κατανοήσει καλύτερα. Επίσης, βοηθάει στην προβολή των δεδομένων από πολλές οπτικές γωνίες ανάλογα με τον στόχο που ο χρήστης επιθυμεί να επιτύχει κάθε φορά. Τέλος, ορισμένοι χρήστες επιθυμούν να δουν μια γενική εικόνα των δεδομένων, ενώ άλλοι προτιμούν να εμβαθύνουν περισσότερο σε αυτά. Η διαδραστική οπτικοποίηση μπορεί να παρέχει και την ανωτέρω δυνατότητα και επιλογή ανάλογα με τις ανάγκες του κάθε χρήστη. Επομένως, για όλους τους παραπάνω λόγους οι διαδραστικές οπτικοποιήσεις αποτελούν ιδανική λύση.

1.5. Δομή Εργασίας

Κεφάλαιο 2^ο (ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ): Στο δεύτερο κεφάλαιο, αναπτύσσονται τα θεωρητικά θέματα που είναι αναγκαία για την καλύτερη κατανόηση της παρούσας εργασίας. Εξετάζονται θέματα όπως η οπτικοποίηση δεδομένων, τα μοντέλα σχεδιασμού λογισμικού, η ευχρηστία και η αξιολόγηση.

Κεφάλαιο 3^ο (ΙΣΤΟΡΙΚΗ ΕΞΕΛΙΞΗ): Στο τρίτο κεφάλαιο, πραγματοποιείται μια ιστορική αναδρομή στην εξέλιξη των οπτικοποιήσεων μέχρι σήμερα. Στη συνέχεια, εξετάζονται οπτικοποιήσεις από διάφορες γνωστές εφαρμογές από τις οποίες έχουν αντληθεί στοιχεία για την παρούσα εργασία. Αναλύεται με ποιόν τρόπο αναπαριστούν αυτές οι εφαρμογές τα δεδομένα και πόσο αποδοτικός είναι αυτός για τον χρήστη.

Κεφάλαιο 4^ο (Η ΣΥΛΛΟΓΗ ΤΩΝ ΔΕΔΟΜΕΝΩΝ): Στο τέταρτο κεφάλαιο, παρουσιάζεται η διαδικασία απόκτησης των δεδομένων. Εξετάζονται αναλυτικά τα εργαλεία που χρησιμοποιήθηκαν και αναλύεται όλη η διαδικασία που ακολουθήθηκε για τη συλλογή των δεδομένων, καθώς και η τροποποίηση τους, ώστε να γίνουν αυτά διαχειρίσιμα.

Κεφάλαιο 5^ο (Ο ΣΧΕΔΙΑΣΜΟΣ): Στο πέμπτο κεφάλαιο, αναλύεται η διαδικασία που πραγματοποιήθηκε για τον σχεδιασμό της διαδραστικής οπτικοποίησης που προτείνεται. Επιπλέον, εισάγεται ένα εργαλείο σχεδιασμού, που χρησιμοποιήθηκε σε αυτήν την εργασία, το «Tableau». Τέλος, παρουσιάζονται οι δύο πρώτες διαμορφωτικές αξιολογήσεις, καθώς και τα συμπεράσματα τα οποία εξήχθησαν από αυτές.

Κεφάλαιο 6^ο (ΥΛΟΠΟΙΗΣΗ): Στο έκτο κεφάλαιο, παρουσιάζονται τα εργαλεία που χρησιμοποιήθηκαν για την υλοποίηση και αναλύονται διεξοδικά όλα τα βήματα που έγιναν για την υλοποίηση της τελικής οπτικοποίησης. Τέλος, δίνονται τα αποτελέσματα και τα συμπεράσματα από την τελευταία συμπερασματική αξιολόγηση.

Κεφάλαιο 7^ο (ΣΥΜΠΕΡΑΣΜΑΤΑ): Στο έβδομο κεφάλαιο, παρουσιάζονται τα συμπεράσματα που προέκυψαν από την εκπόνηση της εργασίας αυτής.

Κεφάλαιο 8^ο (ΜΕΛΛΟΝΤΙΚΕΣ ΕΠΕΚΤΑΣΕΙΣ): Στο όγδοο κεφάλαιο, δίνονται ορισμένες μελλοντικές επεκτάσεις, οι οποίες θα μπορούσαν να λάβουν χώρα για την ενδεχόμενη βελτίωση της παρούσας διαδραστικής οπτικοποίησης, ωστόσο δεν πραγματοποιήθηκαν στα πλαίσια της συγκεκριμένης πτυχιακής εργασίας.

2. ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ

2.1. Η Οπτικοποίηση Δεδομένων

Πριν από την παράθεση του τρόπου εκπόνησης της παρούσας εργασίας και για την καλύτερη κατανόηση αυτής, αξίζει να αναλυθούν ορισμένες έννοιες οι οποίες θα χρησιμοποιηθούν στη συνέχεια. Μία από τις έννοιες αυτές είναι η διαδραστική οπτικοποίηση δεδομένων. Παρακάτω δίνεται ένας πιο ακριβής ορισμός της διαδραστικής οπτικοποίησης δεδομένων, καθώς και τα προβλήματα που παρουσιάζονται συνήθως κατά τη δημιουργία μιας διαδραστικής οπτικοποίησης δεδομένων.

Η οπτικοποίηση δεδομένων αφορά στην οπτική αναπαράσταση αφηρημένων πληροφοριών, προκειμένου να γίνονται αντιληπτά στον αποδέκτη, με στόχο την παρουσίαση ή τη μελέτη τους. Περιλαμβάνει την κατασκευή εικόνων και διαγραμμάτων που επιδιώκουν την επικοινωνία των δεδομένων στον αποδέκτη με έναν πιο κατανοητό τρόπο. Δηλαδή, προάγει την ευχρηστία στην επαφή του αποδέκτη με τα δεδομένα. Η οπτικοποίηση παρέχει μια γρήγορη και διαισθητική ανάλυση πολύπλοκων και μεγάλων σε όγκο δεδομένων με την χρήση της εικόνας, ενώ παράλληλα ενισχύει το μήνυμα που πρέπει να τονιστεί και βοηθάει το χρήστη να κατανοήσει εύκολα τα δεδομένα. Επιπροσθέτως, παρέχεται στο χρήστη ένας τρόπος να αναλύσει τις πληροφορίες καθώς και να ανακαλύψει χαρακτηριστικά και διασυνδέσεις που διαφορετικά θα παρέμεναν στην αφάνεια. Η προσέγγιση αυτή είναι πολύτιμη στην οικοδόμηση της γνώσης, καθώς θέτει κρίσιμα ερωτήματα και ενισχύει την δημιουργική σκέψη. Φέρνει στην επιφάνεια θέματα και έννοιες με αποτέλεσμα την κατανόηση και την ολοκληρωμένη επεξεργασία τεράστιων ποσοτήτων πληροφορίας. Περαιτέρω, η οπτικοποίηση έχει υπάρξει ένας αποτελεσματικός τρόπος για μετάδοση τόσο αφηρημένων όσο και συγκεκριμένων ιδεών εδώ και πολλούς αιώνες, καθόσον έχει χρησιμοποιηθεί σε χάρτες, επιστημονικά σχέδια και διαγράμματα δεδομένων [30]. Οι περισσότερες από τις έννοιες και αρχές με τις οποίες αναπτύχθηκαν οι εικόνες αυτές χρησιμοποιούνται ακόμα και τώρα. Η οπτικοποίηση στις μέρες μας αναπτύσσεται ταχύτατα στην εκπαίδευση, τη βιομηχανία, την ιατρική, καθώς και σε πληθώρα άλλων κλάδων.

Συγκεκριμένα, λόγω της εξέλιξης των υπολογιστών τα τελευταία έτη, πραγματοποιείται μια διαφορετική προσέγγιση στην οπτικοποίηση δεδομένων, ιδίως αυτή της διαδραστικής οπτικοποίησης. Η διαδραστική οπτικοποίηση ορίζεται ως η οπτικοποίηση που επιτρέπει σε ένα χρήστη κάνοντας άμεσες ενέργειες να επηρεάζει τα στοιχεία που απεικονίζονται από μία ή πολλαπλές γραφικές αναπαραστάσεις. Με τη διαδραστική οπτικοποίηση, παρέχεται στο χρήστη η δυνατότητα να αλλάζει με τη χρήση του υπολογιστή τον τρόπο με το οποίο παρουσιάζονται τα δεδομένα σε πραγματικό χρόνο [29]. Σε αντίθεση με ένα στατικό διάγραμμα, η διαδραστική οπτικοποίηση προσφέρει στο χρήστη τη δυνατότητα να διακρίνει την πολυμορφικότητα των δεδομένων, δηλαδή να τα παρατηρήσει από πολλές πλευρές. Ο χρήστης έχει μια πιο σφαιρική και διαισθητική εικόνα των δεδομένων, χωρίς να χρειάζεται να τα διαβάσει, με αποτέλεσμα να μπορεί να εξάγει συμπεράσματα πιο γρήγορα και πιο στοχευμένα σύμφωνα με τις απαιτήσεις του. Στο πλαίσιο της εν λόγω πτυχιακής εξετάζονται οπτικοποιήσεις αυτού του είδους.

Ωστόσο, υπάρχουν και αρκετά προβλήματα που πρέπει κάποιος να αναλογιστεί, εφόσον επιθυμεί να δημιουργήσει μια οπτικοποίηση. Καταρχήν, εάν τα στοιχεία στα

οποία βασίζεται η οπτικοποίηση δεν είναι αξιόπιστα, η πληροφορία που εξάγεται είναι ελαττωματική. Στην περίπτωση όμως που τα δεδομένα είναι αξιόπιστα και η οπτικοποίηση έχει δημιουργηθεί με κακότεχνο και πρόχειρο τρόπο μπορεί η τελευταία είτε να παρουσιάσει πληροφορίες με χαμηλό ενδιαφέρον, είτε να υπερβάλλει στη σπουδαιότητα κάποιων χαρακτηριστικών οδηγώντας σε εσφαλμένα συμπεράσματα. Ακολουθώντας, η πλειονότητα των ανθρώπων δεν είναι εξοικειωμένη με τις οπτικές αναπαραστάσεις και αδυνατεί να εκμεταλλευτεί τις δυνατότητες που αυτές παρέχουν, με αποτέλεσμα η οπτικοποίηση να μην πετυχαίνει τον σκοπό της, δηλαδή να βοηθάει το χρήστη να κατανοεί, να επεξεργάζεται και να αναλύει πολύπλοκα και μεγάλα σε όγκο δεδομένα. Τέλος, υπάρχει και ένα ποσοστό ανθρώπων με διαταραχές όρασης ή αναγνώρισης των αντικειμένων με χαρακτηριστικό το παράδειγμα της αχρωματοψίας, που έχουν ως αποτέλεσμα την αδυναμία αντίληψης κάποιων μορφών οπτικοποίησης.

2.2. Ευχρηστία

Στην παρούσα εργασία ο όρος ευχρηστία αναφέρεται αρκετές φορές, συνεπώς είναι επιτακτική ανάγκη να δοθεί ένας πιο σαφής ορισμός της εν λόγω έννοιας.

Ο όρος ευχρηστία αφορά την ιδιότητα ενός χρησιμοποιήσιμου συστήματος ή συσκευής να καθιστά εύκολη τη χρήση του. Στο γνωστικό πεδίο της επικοινωνίας χρήστη-υπολογιστή ο όρος ευχρηστία (usability) αναφέρεται σε μία βασική ποιοτική παράμετρο ενός αλληλεπιδραστικού υπολογιστικού συστήματος. Ορισμούς ευχρηστίας περιέχουν διεθνή πρότυπα που αφορούν ποιότητα λογισμικού (ISO/IEC 9126) και επικοινωνία χρήστη-υπολογιστή (ISO/DIS 9241-11). Η ευχρηστία είναι εν γένει αναγκαία ιδιότητα ενός συστήματος, συσκευής, προϊόντος ή υπηρεσίας και σχετίζεται με την εμπειρία χρήσης του από τους χρήστες.

Σύμφωνα με το πρότυπο ISO 9241-11, ευχρηστία είναι η δυνατότητα ενός προϊόντος/συστήματος ή υπηρεσίας, που χρησιμοποιείται από ορισμένους χρήστες με καθορισμένους στόχους και υπό συγκεκριμένες συνθήκες χρήσης, να παρέχει αποτελεσματικότητα, αποδοτικότητα και να ικανοποιεί τους χρήστες του.

Σύμφωνα με τον Nielsen [2] η ευχρηστία του συστήματος αναλύεται στις εξής παραμέτρους:

- ευκολία εκμάθησης,
- υψηλή απόδοση εκτέλεσης έργου,
- χαμηλή συχνότητα σφαλμάτων χρήστη,
- ευκολία κατανόησης και αποστήθισης της χρήσης του και
- υποκειμενική ικανοποίηση χρήστη.

Όπως είναι επομένως προφανές, ο βαθμός ευχρηστίας εξαρτάται από τη διασύνδεση χρήστη τόσο ως προς λογισμικό (π.χ. το GUI ενός λειτουργικού συστήματος), όσο και ως προς το απαιτούμενο υλικό (π.χ. περιφερειακές συσκευές εισόδου / εξόδου, όπως ποντίκι, πληκτρολόγιο, οθόνη κλπ).

2.3. Η Έννοια της Σχεδίασης Διαδραστικών Εφαρμογών

Για τη δημιουργία μιας αποδοτικής και εύχρηστης οπτικοποίησης και εν γένει μιας διαδραστικής εφαρμογής είναι σημαντικό να αποτυπωθεί με σαφήνεια ο στόχος της οπτικοποίησης αυτής, καθώς και να πραγματοποιηθεί σωστή σχεδίαση (Design) διαδραστικής εφαρμογής, ώστε η τελευταία να ικανοποιεί με τον καλύτερο δυνατό τρόπο τις ανάγκες των χρηστών. Για το λόγο αυτό, αξίζει να διευκρινιστεί η σημασία της έννοιας «σχεδίαση διαδραστικής εφαρμογής».

Η σχεδίαση είναι η προδιαγραφή ενός αντικειμένου, όπως εκφράζεται από μια ενδιαμέση αναπαράσταση (σχέδιο), το οποίο προσπαθεί να ικανοποιήσει έναν συγκεκριμένο σκοπό, σε ένα ορισμένο πλαίσιο και το οποίο αποτελείται από βασικά επιμέρους τμήματα, ικανοποιεί ορισμένες απαιτήσεις και υπόκειται σε περιορισμούς. Η σχεδίαση αφορά στο «πώς τα πράγματα θα πρέπει να είναι», σε αντίθεση με τις φυσικές επιστήμες που μελετούν το «πώς τα πράγματα είναι».

Ο Herbert A. Simon όρισε αρχικά το 1969 την επιστήμη της σχεδίασης (design science) ως «ένα σώμα από συγκεκριμένες, αναλυτικές, μερικώς προτυποποιημένες, μερικώς εμπειρικές, ικανές να διδαχθούν, οδηγίες για τη σχεδιαστική διαδικασία (design process)» [3]. Τέσσερα έτη αργότερα ο Robert McKim (1973)[4] εισάγει τον όρο της μηχανικής του σχεδιασμού (design engineering) και το 1979 ο Archer κάνει λόγο για την ύπαρξη ενός ιδιαίτερου σχεδιαστικού τρόπου σκέψης, ο οποίος διαφέρει τόσο από τον επιστημονικό όσο και από το λογικό τρόπο σκέψης, αλλά παράλληλα είναι τόσο ισχυρός όσο οι επιστημονικές και λογικές μέθοδοι όταν εφαρμόζονται στα δικά τους είδη προβλημάτων [5].

Η σχεδίαση διαδραστικών συστημάτων (interactive systems design) είναι μία διαδικασία που παράγει όχι μόνο το σχέδιο, αλλά και το ίδιο το τελικό προϊόν. Για τα προβλήματα σχεδίασης δεν υπάρχει μία και μοναδική σωστή λύση, ενώ είναι ιδιαίτερα απαιτητικά, καθώς ο σχεδιαστής καλείται να πάρει αποφάσεις χωρίς να έχει στη διάθεση του το σύνολο της πληροφορίας. Η σχεδίαση διαδραστικών συστημάτων αποτελεί μία διαδικασία επίλυσης προβλημάτων, η οποία είναι κατευθυνόμενη από τους στόχους των χρηστών και από τον τρόπο που οι χρήστες προτίθενται να χρησιμοποιήσουν τη λύση του προβλήματος.

2.4. Μοντέλα Σχεδίασης Συστημάτων Λογισμικού

Η τεχνολογία λογισμικού έχει αναπτύξει διαδικασίες που επιτρέπουν τη συνεχή ανάπτυξη προϊόντων λογισμικού και οι οποίες περιγράφουν τις σημαντικές φάσεις ανάπτυξης, γνωστές ως μοντέλα ανάπτυξης ή ως κύκλος ζωής συστημάτων λογισμικού. Τα μοντέλα αυτά ενισχύουν την κατανόηση της διαδικασίας ανάπτυξης, και καθοδηγούν τον τρόπο σκέψης, δράσης και συνεργασίας της ομάδας σχεδιασμού. Ο ειδικός επικοινωνίας χρήστη - υπολογιστή πρέπει να έχει υπόψιν του ένα εύρος από τα ως άνω μοντέλα, ούτως ώστε να σχεδιάζει κατάλληλα τη διαδικασία ανάπτυξης, όπως άλλωστε προτείνει και το μοντέλο ανθρωποκεντρικού σχεδιασμού κατά ISO 9241-210:2010. Στην συνέχεια παρουσιάζονται κάποια από τα μοντέλα αυτά.

2.4.1. Σχεδίαση με επίκεντρο τον χρήστη (User-Centred Design)

Σύμφωνα με το πρότυπο ISO 9241-210:2010, η διαδικασία ανάπτυξης διαδραστικών συστημάτων περιλαμβάνει τα εξής βήματα:

- σχεδιασμός της ανθρωποκεντρικής διαδικασίας,
- προσδιορισμός του πλαισίου χρήσης,
- προσδιορισμός απαιτήσεων χρηστών και οργανωτικών απαιτήσεων,
- παραγωγή σχεδιαστικών λύσεων και
- αξιολόγηση σχεδίασης βάσει απαιτήσεων. Η διαδικασία ανάπτυξης εκτελείται επαναληπτικά.

Το πρότυπο αυτό προτείνει τις παρακάτω έξι βασικές αρχές προκειμένου να διασφαλίζεται ο ανθρωποκεντρικός σχεδιασμός ενός διαδραστικού συστήματος:

1. Ο σχεδιασμός βασίζεται σε μια σαφή κατανόηση των χρηστών, των εργασιών και του περιβάλλοντος χρήσης.
2. Οι χρήστες εμπλέκονται σε όλη τη διάρκεια του σχεδιασμού και της ανάπτυξης.
3. Ο σχεδιασμός οδηγείται και βελτιώνεται βάσει ανθρωποκεντρικής αξιολόγησης.
4. Η διαδικασία είναι επαναληπτική.
5. Ο σχεδιασμός αφορά το σύνολο της εμπειρίας του χρήστη.
6. Η ομάδα σχεδιασμού περιλαμβάνει διεπιστημονικές δεξιότητες και οπτικές.

2.4.2. Ελικοειδές μοντέλο (Spiral Model)

Το ελικοειδές μοντέλο είναι ουσιαστικά ένα Iterative UCD model, συνεπώς ανήκει στο σχεδιασμό με επίκεντρο τον χρήστη που αναλύεται παραπάνω. Πιο συγκεκριμένα, το ελικοειδές μοντέλο (spiral model) [8], γνωστό και ως σπειροειδές μοντέλο ή μοντέλο εξελικτικής ανάπτυξης αντιμετωπίζει τα προβλήματα του μοντέλου καταρράκτη. Το μοντέλο αυτό, το οποίο χρησιμοποιείται στην παρούσα εργασία και αναφέρεται στο κεφάλαιο του σχεδιασμού, παρουσιάζει την ανάπτυξη του συστήματος ως μία εξελικτική διαδικασία διαδοχικών βελτιώσεων ενός αρχικού πρωτοτύπου. Κάθε φάση του ελικοειδούς μοντέλου περιλαμβάνει ανάλυση και συγγραφή ή βελτίωση των απαιτήσεων, σχεδιασμό και τέλος ανάπτυξη ενός πρωτοτύπου του τελικού προϊόντος, το οποίο σε κάθε διαδοχική σπείρα είναι πιο λεπτομερές και ακριβές. Τα διαδοχικά πρωτότυπα που αναπτύσσονται, ελέγχονται και αξιολογούνται με βάση τις αρχικές απαιτήσεις. Τα αποτελέσματα της αξιολόγησης αυτής αποτελούν κάθε φορά τη βάση για τον προγραμματισμό του επόμενου κύκλου ανάπτυξης.

Το ελικοειδές μοντέλο ταιριάζει ιδιαίτερα στις αντικειμενοστραφείς μεθοδολογίες ανάλυσης και σχεδιασμού [9] που έχουν αναπτυχθεί και χρησιμοποιούνται ευρύτατα τα τελευταία χρόνια. Σύμφωνα με τις μεθοδολογίες αυτές, κατά το στάδιο της ανάπτυξης δημιουργούνται μοντέλα των χρηστών και των εργασιών τους που παίρνουν τη μορφή αντικειμένων. Στη συνέχεια προσδιορίζονται οι συμπεριφορές

των αντικειμένων και οι μεταξύ τους σχέσεις. Με βάση αυτά τα στοιχεία σχεδιάζονται με λεπτομέρεια τα αντικείμενα του συστήματος και της διεπαφής με τα οποία αλληλεπιδρά ο χρήστης. Τα αντικείμενα αυτά δε, ενσωματώνονται σε πρωτότυπα, τα οποία αξιολογούνται. Η ύπαρξη πρωτοτύπων στα πρώτα στάδια της διαδικασίας ανάπτυξης, επιτρέπει στους σχεδιαστές να παρουσιάσουν στους χρήστες τμήματα του τελικού προϊόντος, προκειμένου να καταγράψουν και να μετρήσουν την αντίδρασή τους. Η πληροφορία αυτή επιτρέπει στη συνέχεια την κατάλληλη προσαρμογή των χαρακτηριστικών του συστήματος στις απαιτήσεις των χρηστών. Σύμφωνα με τις αρχές του ανθρωποκεντρικού σχεδιασμού, οι χρήστες του συστήματος πρέπει να καταγραφούν και τα χαρακτηριστικά τους να αναλυθούν κατά τις φάσεις συγκέντρωσης απαιτήσεων. Οι χρήστες όμως θα πρέπει στη συνέχεια να εμπλακούν στην αξιολόγηση του συστήματος που σε μια εξελικτική διαδικασία είναι επαναλαμβανόμενη και συνεχής.

2.5. Πρωτότυπα Χαμηλής και Υψηλής Πιστότητας

Το πρωτότυπο είναι ένα προσχέδιο με βάση το οποίο πρέπει να σχεδιαστεί η τελική εφαρμογή ή το τελικό σύστημα και στην εν λόγω περίπτωση η οπτικοποίηση. Είναι δηλαδή ένα σχεδιάγραμμα, στο οποίο σκιαγραφείται η μορφή της τελικής οπτικοποίησης. Τα πρωτότυπα που υπάρχουν μπορεί να είναι χαμηλής πιστότητας, τα οποία δεν είναι κατασκευασμένα με αυστηρό τρόπο και είναι λιγότερο λεπτομερή και τα πρωτότυπα υψηλής πιστότητας, τα οποία παρουσιάζουν επακριβώς τη μορφή της τελικής οπτικοποίησης.

Ο σχεδιασμός των πρωτοτύπων περιλαμβάνει πολλές ενδιάμεσες φάσεις και οι σχεδιαστές πρέπει να αποφασίσουν για το στυλ αλληλεπίδρασης, τις συσκευές αλληλεπίδρασης, καθώς και τη μεταφορά (αναλογία) που υπερισχύει και που θα χρησιμοποιηθεί στο σύστημα. Ακολουθώντας, οι σχεδιαστές θα πρέπει να λάβουν υπόψη, προκειμένου να προχωρήσουν στο σχεδιασμό πρωτοτύπου, τη ροή του διαλόγου, το μοντέλο πλοήγησης και το εικονικό περιεχόμενο και τη διάταξη των οθονών.

2.6. Η Σημασία της Αξιολόγησης

Η αξιολόγηση συνιστά μια σημαντική και αυτονόητη απαίτηση για την ανάπτυξη αξιόλογου διαδραστικού λογισμικού. Σύμφωνα με τους Lewis και Rieman (1993) [10], το βασικότερο κριτήριο αξιολόγησης ποιότητας μιας διεπαφής είναι η αποτελεσματική ή όχι χρήση της από πραγματικούς χρήστες.

Υπάρχουν πολλές μέθοδοι αξιολόγησης για την εκτίμηση της ευχρηστίας μιας εφαρμογής. Στην παρούσα εργασία χρησιμοποιούνται μέθοδοι αξιολόγησης με τη συμμετοχή τελικών χρηστών. Συγκεκριμένα, οι μέθοδοι που χρησιμοποιούνται είναι η παρατήρηση του χρήστη και η μέτρηση απόδοσής του, προσπαθώντας να εκτελέσει στοιχειώδεις εργασίες (Σενάρια χρήσης), καθώς και η συμπλήρωση ερωτηματολογίων από τους χρήστες, αφού πρώτα έχουν χρησιμοποιήσει την υπό κατασκευή οπτικοποίηση.

2.6.1. Παρατήρηση του χρήστη

Οι μέθοδοι παρατήρησης του χρήστη (user observation) περιλαμβάνουν έναν ερευνητή, ο οποίος παρατηρεί τους χρήστες ενόσω εργάζονται και κρατάει σημειώσεις για τη δραστηριότητα που πραγματοποιείται. Η μέθοδος αυτή είναι χρήσιμη τόσο στα αρχικά στάδια της έρευνας και της καταγραφής προδιαγραφών χρήστη για την απόκτηση ποιοτικών δεδομένων, όσο και για τη μελέτη των υφιστάμενων εργασιών και διαδικασιών. Η παρατήρηση μπορεί να είναι είτε άμεση, όπου ο ερευνητής είναι παρών κατά την εκτέλεση της δραστηριότητας, είτε έμμεση όπου η δραστηριότητα παρατηρείται με άλλα μέσα, όπως μέσω της εγγραφής βίντεο ή ηχογράφησης. Επίσης, μπορεί να γίνεται είτε στο φυσικό χώρο που λαμβάνει χώρα η δραστηριότητα είτε σε συνθήκες εργαστηρίου. Ακόμη, η παρατήρηση μπορεί να διαρκεί λίγες ώρες ή πολλές ημέρες. Τέλος, μπορεί να αφορά στην παρατήρηση μεμονωμένων ατόμων ή ομάδων. Η Συνεργατική Ανακάλυψη (co-discovery learning) αποτελεί ένα παράδειγμα μεθόδου, όπου ο ερευνητής παρατηρεί δύο άτομα που εξερευνούν μαζί μία ιδέα ή ένα προϊόν και συζητούν για αυτό [11].

Η παρατήρηση του χρήστη προσφέρει πληθώρα καταγραφή πληροφοριών για την υπό μελέτη δραστηριότητα του χρήστη, καθώς η παρατήρηση δεν στηρίζεται σε διαδικασίες της μνήμης του συμμετέχοντα και επιτρέπει στον ερευνητή να διαπιστώσει τι πραγματικά κάνουν οι χρήστες στο πλαίσιο της υπό μελέτη δραστηριότητας. Συγκεκριμένα, επιτρέπει την ακριβή καταγραφή των βημάτων και αποφάσεων των συμμετεχόντων, καθώς και των εκφράσεων προσώπου και σώματος αυτών.

Από την άλλη πλευρά, η παρατήρηση του χρήστη είναι μια χρονοβόρα διαδικασία (επεξεργασία και ανάλυση των σημειώσεων, βίντεο, ηχογραφήσεων του παρατηρητή), με αποτέλεσμα να επικεντρώνεται συνήθως σε ένα σχετικά μικρό δείγμα συμμετεχόντων. Επιπλέον, η παρατήρηση μπορεί να δημιουργήσει πρόβλημα συγκέντρωσης στους συμμετέχοντες και είναι πιθανόν να τροποποιούν τη συμπεριφορά τους λόγω της παρουσίας του παρατηρητή.

2.6.2. Μέτρηση της απόδοσης του χρήστη

Η μέτρηση της απόδοσης έχει σκοπό να πιστοποιήσει εάν η ευχρηστία επιτυγχάνεται ή όχι. Η απόδοση των χρηστών μετράται συνήθως υποβάλλοντας μια ομάδα χρηστών σε εκτέλεση σεναρίων χρήσης και συλλέγοντας στοιχεία που μπορεί να αφορούν σε σφάλματα που διαπιστώνονται ή στους χρόνους εκτέλεσης των εργασιών. Οι δοκιμές πραγματοποιούνται συνήθως σε ένα εργαστήριο ευχρηστίας λογισμικού, το οποίο έχει τον απαραίτητο εξοπλισμό για καταγραφή των ενεργειών των χρηστών (καταγραφή ήχου, εικόνας, καταγραφή οθόνης υπολογιστή κατά τη διάρκεια της αλληλεπίδρασης) και ταυτόχρονα προσομοιώνει τον τυπικό χώρο χρήσης του λογισμικού ελαχιστοποιώντας την αίσθηση παρακολούθησης στον χρήστη. Με μια τέτοια διαδικασία ελέγχου, μπορούν να εντοπιστούν πολλά προβλήματα ευχρηστίας.

Ένα πλεονέκτημα της ως άνω μεθόδου αξιολόγησης είναι η συγκέντρωση μετρήσιμων χρόνων και ποσοστών που καθιστά τη σύγκριση εναλλακτικών σχεδιασμών εύκολη. Από την άλλη πλευρά, αρνητικός παράγοντας στη διεξαγωγή των εργαστηριακών δοκιμών κατά τα αρχικά στάδια του σχεδιασμού είναι ότι δεν υπάρχει σημαντικός αριθμός υποψήφιων αντιπροσωπευτικών χρηστών, δηλαδή

χρηστών που να ανταποκρίνονται στα πραγματικά χαρακτηριστικά των ανθρώπων που χρησιμοποιούν το σύστημα. Το συγκεκριμένο πρόβλημα μπορεί να οδηγήσει σε μεθοδολογικές αδυναμίες στον προγραμματισμό, την εγκυρότητα και την αξιοπιστία των αποκτηθέντων μετρήσεων [12].

2.6.3. Ερωτηματολόγια

Μία έρευνα με ερωτηματολόγιο (survey) αποτελεί μια χρήσιμη προσέγγιση, όταν ο στόχος είναι να διερευνηθούν ζητήματα σχετικά με πιθανές ανησυχίες-παρανοήσεις των χρηστών και την υποκειμενική τους ικανοποίηση [13]. Τα ερωτηματολόγια περιλαμβάνουν την κατασκευή, διανομή και ανάλυση ενός συνόλου ερωτήσεων που μπορούν να διανεμηθούν εύκολα σε πολλούς χρήστες. Επιπλέον, αποτελούν μια σχετικά ανέξοδη μέθοδο ερευνών. Τυπικά περιλαμβάνουν ένα σύνολο «κλειστών» ερωτήσεων με προκαθορισμένες απαντήσεις σε διάφορες μορφές (π.χ. πενταβάθμια κλίμακα) και ένα σύνολο «ανοικτών» ερωτήσεων που απαιτούν ελεύθερη διατύπωση από το συμμετέχοντα. Οι συμμετέχοντες είναι δυνατόν να συμπληρώνουν το ερωτηματολόγιο σε έντυπη μορφή, σε ηλεκτρονική μορφή μέσω email ή διαδικτυακών φορμών.

Οι έρευνες αυτές μπορούν να βοηθήσουν στον καθορισμό των χαρακτηριστικών και αναγκών του χρήστη, των πρακτικών εργασίας, των στάσεων του χρήστη απέναντι στο υπό ανάπτυξη σύστημα κ.λ.π.. Παρόλα αυτά, είναι συχνά δύσκολο να συγκεντρωθούν αντικειμενικά αποτελέσματα κατά την αξιοποίηση των ερωτηματολογίων, δεδομένου ότι οι απαντήσεις των χρηστών είναι βασισμένες σε αυτό που έχουν την αίσθηση ότι κάνουν και όχι σε αυτό που κάνουν πραγματικά [14].

3. ΙΣΤΟΡΙΚΗ ΕΞΕΛΙΞΗ

3.1. Οι Πρώτες Οπτικοποιήσεις

Η οπτικοποίηση πληροφορίας δεν είναι σημερινή ανακάλυψη και έχει τις ρίζες της σε παλαιότερες ιστορικές περιόδους. Χαρακτηριστικά παραδείγματα είναι οι αστρονομικοί χάρτες, η χαρτογράφηση περιοχών, η αναπαράσταση στατιστικών δεδομένων, καθώς επίσης και οι επινοήσεις γραφικών παραστάσεων και η πύκνωση της χρήσης, παραγωγής και αναπαραγωγής τους που κάνουν την εμφάνισή τους κατά τον 18^ο αιώνα [15].

Από τα τέλη του 18ου αιώνα μέχρι τις αρχές του 19ου αιώνα κάνουν την εμφάνισή τους απεικονίσεις δεδομένων με τη χρήση διανυσματικών γραφικών, οι οποίες χρησιμοποιούνται ακόμα και στις μέρες μας. Γραμμικά διαγράμματα, διαγράμματα χρονικής σειράς (time-series plots), περιγράμματα (contour plots), στατιστικά γραφικά με ράβδους ή και πίτες, είναι μερικά χαρακτηριστικά παραδείγματα των απεικονίσεων αυτών [16]. Τον 19ο αιώνα δε, λόγω της ανάγκης διαχείρισης μεγάλου όγκου αριθμητικών δεδομένων, η οποία καθιστά αναγκαία την απεικόνιση των δεδομένων αυτών, διεισδύει σε τομείς όπως η πολεοδομία, η αστική ανάπτυξη, το εμπόριο, οι μεταφορές, η βιομηχανία και η επιδημιολογία [17][18].

Σήμερα, η χρήση οπτικοποιήσεων είναι αναπόσπαστο κομμάτι του Data Science (επιστήμης των δεδομένων). Πολλές ιστοσελίδες και γενικότερα εφαρμογές χρησιμοποιούν το συγκεκριμένο μέσο, προκειμένου να παρουσιάζουν τα δεδομένα τους στους χρήστες με έναν πιο ευχάριστο και κατανοητό τρόπο. Στη συνέχεια του κεφαλαίου εξετάζονται ορισμένες οπτικοποιήσεις που αξίζει να αναφερθούν, είτε διότι έχουν ιστορική σημασία για την ανάδειξη της πορείας της οπτικοποίησης, είτε διότι έχουν άμεση σχέση με το θέμα και τα δεδομένα που πραγματεύεται η πτυχιακή εργασία.

3.1.1. Dynamic Queries (1993 University of Maryland UIS Broadcast)

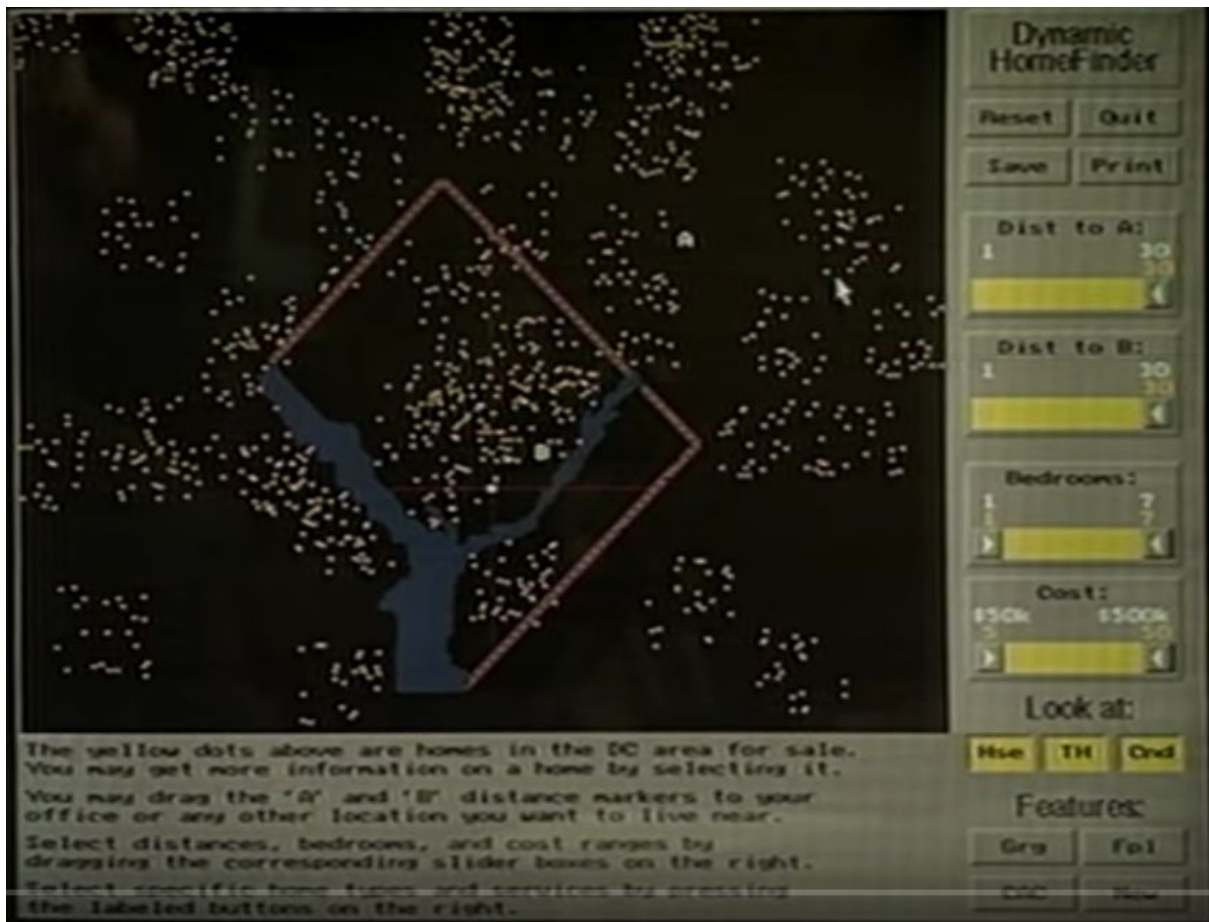
Μία από τις πρώτες διαδραστικές οπτικοποιήσεις δεδομένων με προηγμένα φίλτρα στο χώρο των ηλεκτρονικών υπολογιστών είναι τα Dynamic Queries [19].

Ένα χαρακτηριστικό παράδειγμα των Dynamic Queries είναι το “Dynamic Home Finder”. Το “Dynamic Home Finder” είναι μία διαδραστική οπτικοποίηση, στην οποία οι χρήστες μπορούν να αναζητήσουν κατοικίες που διατίθενται προς πώληση από μία βάση δεδομένων και να επιλέξουν αυτήν που τους ταιριάζει απομονώνοντας όσες κατοικίες ικανοποιούν τις ανάγκες τους. Η οπτικοποίηση αποτελείται από ένα χάρτη, στον οποίο οι κατοικίες που είναι διαθέσιμες αναπαριστώνται με μικρά λευκά στίγματα. Στο χάρτη υπάρχουν τα σημεία A και B, τα οποία ο χρήστης μπορεί να μετακινήσει μέσα σε αυτόν χρησιμοποιώντας τον κέρσορα του. Στα δεξιά της οπτικοποίησης υπάρχουν διάφορα φίλτρα, τα οποία κυρίως είναι “slider bar” (ρυθμιζόμενο φίλτρο κύλισης με μπάρα).

Τα πρώτα φίλτρα αφορούν στην απόσταση από τα σημεία A και B, δηλαδή το πόσο πολύ θα απέχει η κατοικία που θα εμφανίζεται στην αναζήτηση από το κάθε σημείο. Ο χρήστης της εφαρμογής αυτής μπορεί να ορίσει μία συγκεκριμένη απόσταση από το σημείο A και μία διαφορετική, ωστόσο και πάλι συγκεκριμένη απόσταση από το σημείο B και να βρει όλες τις κατοικίες, οι οποίες ικανοποιούν και τους δύο

περιορισμούς ταυτόχρονα. Ακόμη, υπάρχει ένα φίλτρο για τον αριθμό των δωματίων που θα έχει η κατοικία, καθώς επίσης και ένα για την τιμή πώλησης της κατοικίας. Τέλος, υπάρχει δυνατότητα φιλτραρίσματος με βάση τον τύπο του οικήματος, όπως για παράδειγμα «Απλό Σπίτι», «Σπίτι σε Πόλη» ή «Σπίτι σε Συγκρότημα», καθώς και με βάση επιπρόσθετων χαρακτηριστικών του ακινήτου.

Στην περίπτωση που ο χάρτης της περιοχής δεν είναι διαθέσιμος, η συγκεκριμένη εφαρμογή μπορεί να χρησιμοποιηθεί και με έναν εναλλακτικό τρόπο παρουσιάζοντας τις διαθέσιμες κατοικίες αναζήτησης σε μία λίστα. Σε αυτόν τον τρόπο παρουσίασης ενσωματώνεται μία μπάρα στο κάτω μέρος της εφαρμογής, η οποία δείχνει πόσες από τις κατοικίες ικανοποιούν τους περιορισμούς που έχουν οριστεί στα φίλτρα. Η συγκεκριμένη προσέγγιση παρουσιάζει ενδιαφέρον, καθώς όταν η απεικόνιση με χάρτη δεν είναι δυνατή, προσφέρει μία εναλλακτική οπτικοποίηση με τη χρήση του ραβδόγραμματος στο κάτω μέρος της εφαρμογής. Σε αντίθεση με μία απλή λίστα το ραβδόγραμμα παρέχει μία αποτελεσματικότερη κατανόηση των δεδομένων.



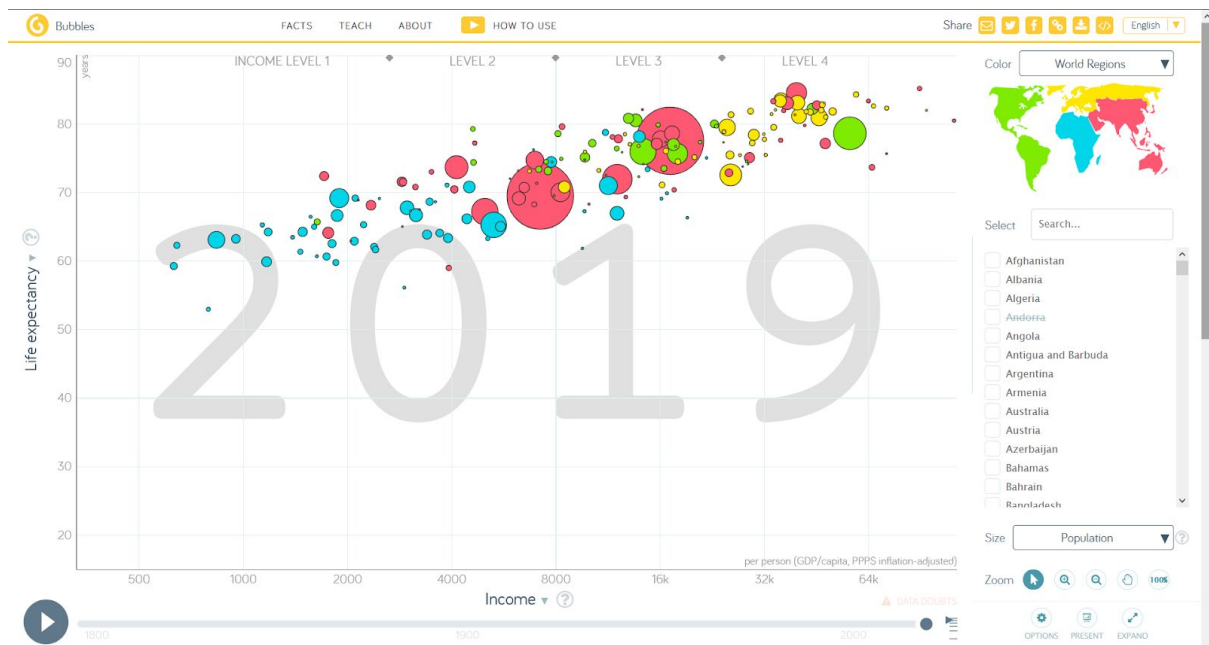
Εικόνα 1: Απόσπασμα διαδραστικής οπτικοποίησης Dynamic Home Finder, πηγή: <https://www.youtube.com/watch?v=5X8XY9430fM>, επιπλέον στιγμιότυπα οθόνης παρατίθενται στο Παράρτημα II

3.1.2. Gapminder Tools (2015)

Ένα χαρακτηριστικό παράδειγμα οπτικοποίησης, η οποία αναπαριστά πολυμορφικά δεδομένα με έναν πολύ αποτελεσματικό τρόπο είναι οι οπτικοποιήσεις του Gapminder².

Το Gapminder είναι μια μη κερδοσκοπική επιχείρηση εγγεγραμμένη στη Στοκχόλμη της Σουηδίας, η οποία προωθεί τη βιώσιμη παγκόσμια ανάπτυξη και την επίτευξη των αναπτυξιακών στόχων με αυξημένη χρήση και κατανόηση των στατιστικών και άλλων πληροφοριών σχετικά με την κοινωνική, οικονομική και περιβαλλοντική ανάπτυξη σε τοπικό, εθνικό και παγκόσμιο επίπεδο [22]. Το Gapminder ιδρύθηκε το 2005 από την Ola Rosling, την Anna Rosling Rönnlund και τον Hans Rosling [23]. Το όνομα Gapminder προήλθε από τα προειδοποιητικά μηνύματα "Mind the Gap" στο London Underground [24].

Το Gapminder παρέχει ένα περιβάλλον στο οποίο ο χρήστης μπορεί να δημιουργήσει μία δική του εκδοχή της οπτικοποίησης, η οποία δημιουργείται χρησιμοποιώντας τις ρυθμίσεις που βρίσκονται στα δεξιά. Με τη χρήση των δεδομένων που παρέχει το Gapminder, ο χρήστης μπορεί να αλλάξει τα χρώματα, οτιδήποτε δείχνουν οι άξονες ακόμα και τον τύπο του διαγράμματος και να δημιουργήσει μία δική του διαδραστική οπτικοποίηση. Στη συνέχεια, παρατίθεται εικόνα αποσπάσματος από την πιο χαρακτηριστική οπτικοποίηση του Gapminder, η οποία αναλύεται παρακάτω.



Εικόνα 2: Απόσπασμα οπτικοποίησης του Gapminder, πηγή:
<https://www.gapminder.org/tools/>

Στην παραπάνω εικόνα ο κατακόρυφος άξονας αντιπροσωπεύει το προσδόκιμο ζωής, ενώ ο οριζόντιος άξονας αντιπροσωπεύει το κατά κεφαλήν εισόδημα. Στο εσωτερικό του διαγράμματος απεικονίζονται με τη μορφή κύκλων οι χώρες. Το μέγεθος του κύκλου σχετίζεται με τον πληθυσμό κάθε χώρας, ενώ το χρώμα

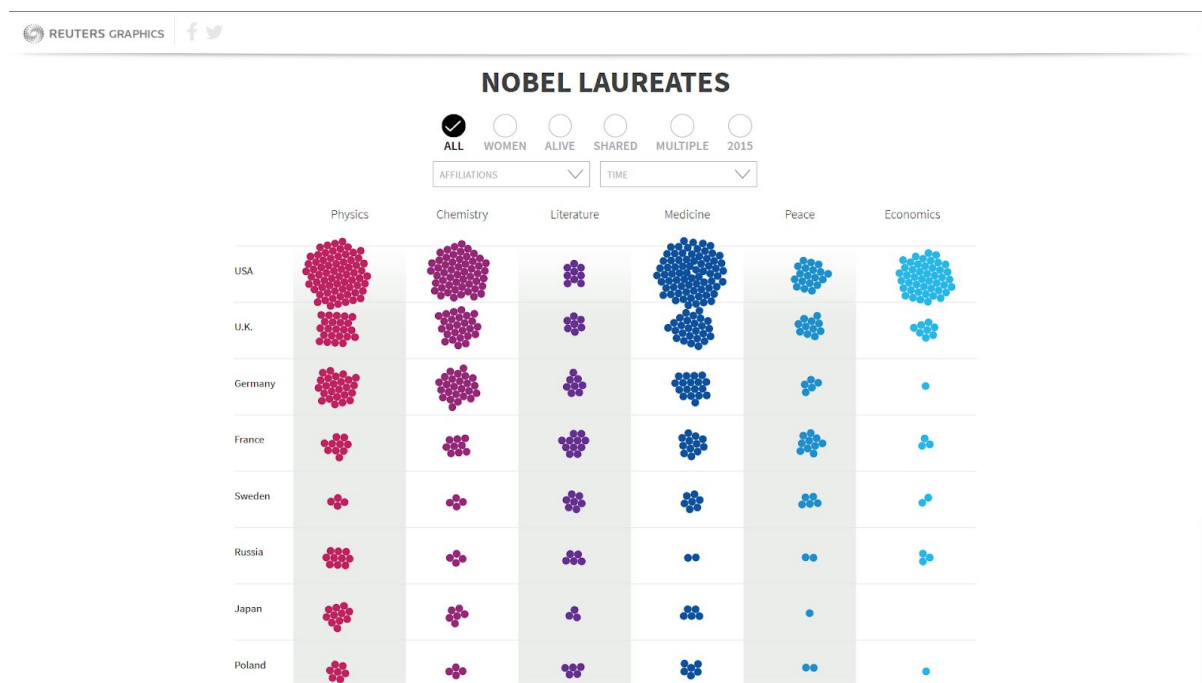
² <https://www.gapminder.org/tools>

αντιπροσωπεύει την Ήπειρο που ανήκει η κάθε χώρα. Επομένως, η κάτω αριστερή πλευρά του διαγράμματος παρουσιάζει τις φτωχές και με χαμηλό προσδόκιμο ζωής χώρες, ενώ η πάνω δεξιά πλευρά του διαγράμματος δείχνει τις πλούσιες και με υψηλό προσδόκιμο ζωής χώρες. Επιπρόσθετα, στο εσωτερικό του διαγράμματος με μεγάλα γκρι γράμματα φαίνεται το εκάστοτε έτος που αντιστοιχεί στα δεδομένα που εμφανίζονται στο διάγραμμα. Τα έτη μεταβάλλονται μετακινώντας την μπάρα στο κάτω μέρος της οπτικοποίησης ή εναλλακτικά πατώντας το «play» όπου και θα γίνει μία διαδοχική αλλαγή όλων των ημερομηνιών από την παλαιότερη μέχρι την πιο πρόσφατη που υπάρχει στα δεδομένα. Δεδομένου ότι ο χρήστης μπορεί να μεταβάλλει το συγκεκριμένο διάγραμμα, υπάρχουν πολλές παραλλαγές της συγκεκριμένης οπτικοποίησης.

Η διαδραστική οπτικοποίηση του Gapminder που παρουσιάστηκε παραπάνω αποτελεί έναν πολύ αποτελεσματικό τρόπο για την ανάλυση των συγκεκριμένων δεδομένων. Όπως αναφέρεται και στα επόμενα κεφάλαια, το βασικό διάγραμμα το οποίο υλοποιείται στη συγκεκριμένη πτυχιακή εργασία είναι ένα διάγραμμα φυσαλίδων και έτσι κρίθηκε αναγκαία η αναφορά του παραπάνω διαγράμματος.

3.1.3. Nobel Laureates Visualization (2018)

Ένα παράδειγμα οπτικοποίησης που αξίζει να σημειωθεί, είναι αυτή που παρουσιάζει τις απονομές των βραβείων Νόμπελ που έχουν πραγματοποιηθεί μέχρι σήμερα. Σε αυτήν την οπτικοποίηση παρουσιάζεται το κάθε βραβείο που έχει αποδοθεί με μια φυσαλίδα (διάγραμμα φυσαλίδων). Ο χρήστης μπορεί να επιλέξει, από ένα μενού επιλογών (dropdown menu), που παρουσιάζεται στον οριζόντιο άξονα, είτε τον επιστημονικό κλάδο στον οποίο ανήκει η ανακάλυψη που βραβεύτηκε, είτε την χρονολογία στην οποία αποδόθηκε το βραβείο. Στον κατακόρυφο άξονα, ο χρήστης μπορεί να επιλέξει με έναν παρόμοιο τρόπο, εάν θα προβάλλεται η χώρα του βραβευμένου ή το πανεπιστήμιο στο οποίο ανήκει η ερευνητική ομάδα που συγκαταλέγεται ο βραβευμένος. Η συγκεκριμένη οπτικοποίηση διαθέτει και διάφορα άλλα φίλτρα με τα οποία μπορεί ο χρήστης να φιλτράρει τα δεδομένα, ώστε να γίνονται ευδιάκριτα μονό αυτά τα οποία εκείνος επιθυμεί. Τα φίλτρα αυτά αφορούν, για παράδειγμα, στον προσδιορισμό του φύλου του βραβευμένου, εάν αυτός είναι εν ζωή ή όχι, εάν το βραβείο μοιράστηκε σε περισσότερα από ένα άτομα, καθώς και το φίλτρο που αφορά στη χρονολογία. Οι φυσαλίδες μετακινούνται στις αντίστοιχες θέσεις, εξαφανίζονται ή εμφανίζονται ανάλογα με τις επιλογές των φίλτρων. Παρακάτω φαίνεται μια εικόνα από την εν λόγω οπτικοποίηση. Επίσης, στο Παράρτημα II παρατίθεται ακόμη μια εικόνα που δείχνει την χρήση των φίλτρων σε αυτήν την οπτικοποίηση.



Εικόνα 3: Οπτικοποίηση απονομών των βραβείων Νόμπελ που έχουν πραγματοποιηθεί μέχρι το 2018. (πηγή: <http://graphics.thomsonreuters.com/15/nobels>)

3.2. Παραδείγματα Οπτικοποιήσεων Ερευνητικών Δημοσιεύσεων

Στο κομμάτι της παρουσίασης των ερευνητικών δημοσιεύσεων, που είναι και το θέμα της παρούσας εργασίας, υπάρχουν διάφορες οπτικοποιήσεις που παρουσιάζονται στην συνέχεια.

3.2.1. Google Scholar (Released in beta in November 2004)

Το Google Scholar (ή αλλιώς Μελετητής Google) είναι μια ελεύθερα προσβάσιμη μηχανή αναζήτησης ιστού που ευρετηριάζει το πλήρες κείμενο ή τα μεταδεδομένα της ακαδημαϊκής λογοτεχνίας σε μια σειρά μορφών και επιστημονικών εκδόσεων. Κυκλοφόρησε σε έκδοση «beta» τον Νοέμβριο του 2004. Το ευρετήριο του Google Scholar περιλαμβάνει τα περισσότερα διαδικτυακά ακαδημαϊκά περιοδικά και βιβλία, επιστημονικές δημοσιεύσεις, διατριβές, διατριβές, περιλήψεις, τεχνικές εκθέσεις και άλλη επιστημονική βιβλιογραφία, συμπεριλαμβανομένων γνωμοδοτήσεων και διπλωμάτων ευρεσιτεχνίας [25]. Παρόλο που η Google δεν δημοσιεύει το μέγεθος της βάσης δεδομένων του Google Scholar, υπολογίστηκε ότι περιέχει περίπου τριακόσια ογδόντα εννέα εκατομμύρια έγγραφα, συμπεριλαμβανομένων άρθρων, παραθέσεων και διπλωμάτων ευρεσιτεχνίας, καθιστώντας την τη μεγαλύτερη ακαδημαϊκή μηχανή αναζήτησης στον κόσμο τον Ιανουάριο του 2018 [26]. Τον Μάιο του 2014, το μέγεθος υπολογίστηκε σε εκατόν εξήντα εκατομμύρια έγγραφα [27]. Μια προηγούμενη στατιστική εκτίμηση που δημοσιεύθηκε στο PLOS ONE χρησιμοποιώντας μια μέθοδο Mark and recapture υπολόγισε περίπου το 80-90% κάλυψης όλων των άρθρων που δημοσιεύθηκαν στα Αγγλικά με εκτίμηση εκατό εκατομμυρίων [20]. Αυτή η εκτίμηση καθόρισε επίσης πόσα έγγραφα ήταν ελεύθερα διαθέσιμα στον Ιστό.

Όσον αφορά την οπτικοποίηση, το Google Scholar διαθέτει ένα ραβδόγραμμα, μία σχετικά απλή οπτικοποίηση. Το συγκεκριμένο ραβδόγραμμα κάνει την εμφάνισή του στο προφίλ ενός συγγραφέα και αναπαριστά τις αναφορές του εκάστοτε συγγραφέα ανά έτος (Citations per year). Παρότι το Google Scholar είναι μία ευρέως γνωστή πλατφόρμα αναζήτησης ερευνητικών δημοσιεύσεων και χρησιμοποιείται καθημερινά από πολλούς χρήστες, δεν χρησιμοποιεί την οπτικοποίηση ως βασικό μέσο για να παρουσιάσει τα δεδομένα, τα οποία διαθέτει. Αντιθέτως, χρησιμοποιεί κυρίως τον παραδοσιακό τρόπο απεικόνισης (παρουσίαση δέκα ανά σελίδα αποτελεσμάτων υπό μορφή λεπτομερούς λίστας) που χρησιμοποιεί γενικά η Google στην αναζήτηση της.



Εικόνα 4: Απόσπασμα από το Google Scholar. Ραβδόγραμμα παραθέσεων ανά έτος για έναν συγκεκριμένο συγγραφέα, πηγή: <https://scholar.google.com/>

3.2.2. Scopus (Launched in 2004)

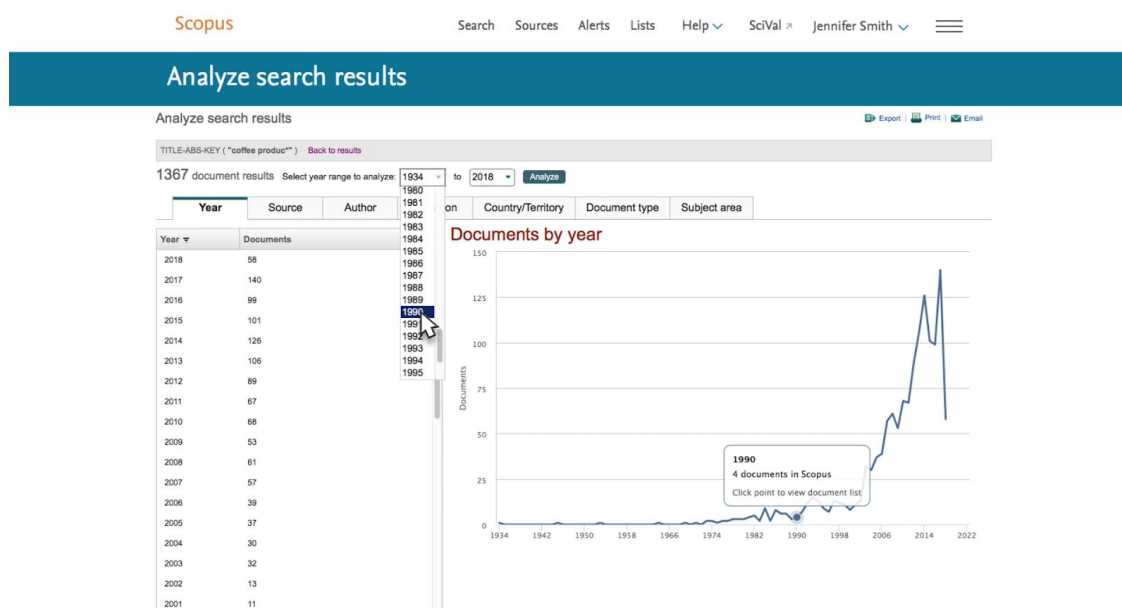
Ένα χαρακτηριστικό παράδειγμα είναι τα διαγράμματα που έχει η πλατφόρμα Scopus³ για να οπτικοποιεί τις ερευνητικές δημοσιεύσεις. Το Scopus είναι η βάση δεδομένων ερευνητικών δημοσιεύσεων και παρέχει περιλήψεις (abstracts) και παραπομπές (citations) για κάθε ερευνητική δημοσίευση που έχει στην βάση δεδομένων του. Το Scopus κυκλοφόρησε το 2004 από την Elsevier⁴. Το Scopus καλύπτει σχεδόν 36.377 τίτλους από περίπου 11.678 εκδότες, εκ των οποίων 34.346 είναι περιοδικά με κριτές σε κορυφαία θεματικά πεδία: βιοεπιστήμες, κοινωνικές επιστήμες, φυσικές επιστήμες και επιστήμες υγείας. Καλύπτει τρεις τύπους πηγών: σειρές βιβλίων, περιοδικά και εμπορικά περιοδικά. Όλα τα περιοδικά που καλύπτονται στη βάση δεδομένων Scopus, ανεξάρτητα από το ποια δημοσιεύονται, ελέγχονται κάθε χρόνο για να διασφαλιστεί ότι διατηρούνται υψηλά πρότυπα ποιότητας. Οι αναζητήσεις στο Scopus περιλαμβάνουν επίσης αναζητήσεις βάσεων δεδομένων διπλωμάτων ευρεσιτεχνίας. Το Scopus δίνει τέσσερις τύπους ποιοτικού

³ <https://www.scopus.com/>

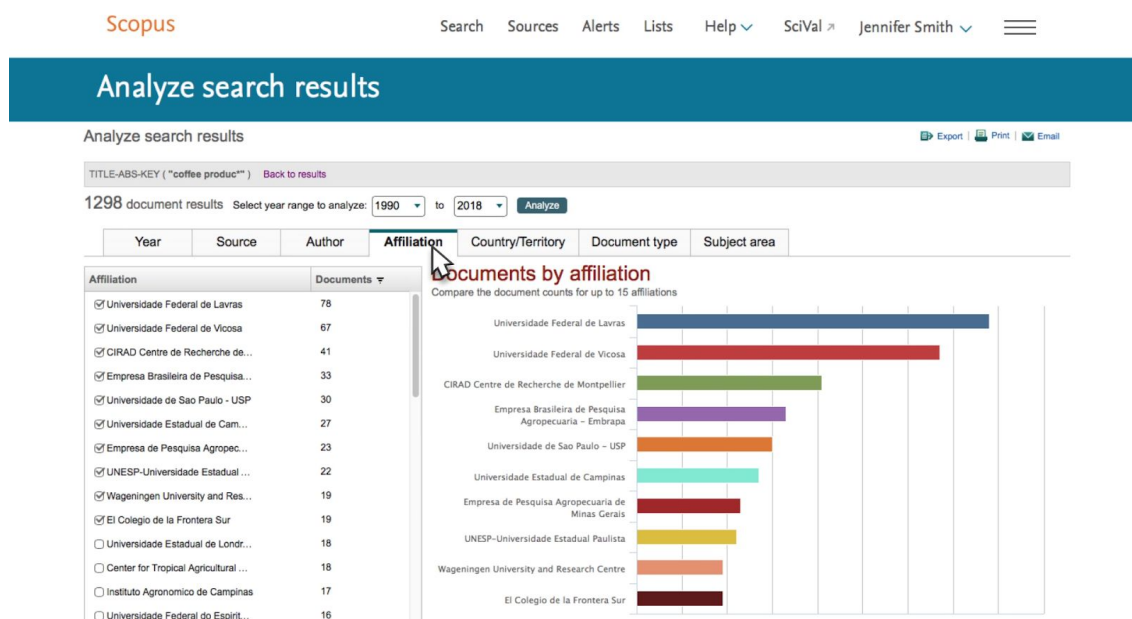
⁴ <https://www.elsevier.com/>

μέτρου για κάθε τίτλο. Αυτά είναι h-Index, CiteScore, SJR (SCImago Journal Rank) και SNIP (Source Normalized Impact per Paper).

Μία σημαντική παρατήρηση που πρέπει να αναφερθεί είναι ότι το Scopus παρέχει μία διαδραστική διεπαφή στην οποία ο χρήστης χρησιμοποιώντας φίλτρα μπορεί να επιλέξει ποια δεδομένα επιθυμεί να βλέπει. Επίσης, παρέχει στους χρήστες διάφορες οπτικοποιήσεις στις οποίες οπτικοποιεί τα δεδομένα αυτά κάνοντας με αυτόν τον τρόπο ολόκληρη την πλατφόρμα πιο αποδοτική και εύχρηστη για αναζήτηση ερευνητικών δημοσιεύσεων. Στην «Εικόνα 5» και «Εικόνα 6» φαίνονται δύο παραδείγματα από τις εν λόγω οπτικοποιήσεις. Περισσότερες οπτικοποιήσεις από την πλατφόρμα του Scopus παρατίθενται στο Παράρτημα II.



Εικόνα 5: Απόσπασμα από το Scopus. Διάγραμμα αριθμού εγγραφών ανά έτος, πηγή: <https://www.scopus.com>



Εικόνα 6: Απόσπασμα από το Scopus. Διάγραμμα αριθμού εγγραφών ανά ερευνητική ομάδα, πηγή: <https://www.scopus.com>

3.3. Συμπεράσματα

Παραπάνω παρατίθενται διάφορες οπτικοποιήσεις και αναλύεται ο τρόπος με τον οποίο διαχειρίζονται και παρουσιάζουν την πληροφορία. Εξετάζονται δεδομένα που έχουν σχέση με αυτά που οπτικοποιούνται στη συγκεκριμένη εργασία, καθώς και οι τρόποι με τους οποίους τα δεδομένα οπτικοποιούνται. Από τα ως άνω παραδείγματα διαπιστώνεται ότι η χρήση οπτικοποίησης για τέτοιες πληροφορίες (δεδομένα που διαθέτουν ονόματα, ημερομηνίες και κατηγορίες κάθε είδους) είναι αναγκαία και κάνει την κατανόηση των δεδομένων ευκολότερη. Επίσης, η χρήση οπτικοποίησης στις εν λόγω περιπτώσεις προσφέρει μεγάλη ευελιξία στον τρόπο προβολής των δεδομένων δίνοντας στον χρήστη μια πιο αφηρημένη οπτική, προκειμένου να έχει τη δυνατότητα να οδηγηθεί σε συμπεράσματα. Προς το παρόν, δεν υπάρχει κάποια ενιαία οπτικοποίηση που να δείχνει όλη την ερευνητική δραστηριότητα του Τμήματος Πληροφορικής και Τηλεπικοινωνιών του ΕΚΠΑ. Ακόμα και σε κορυφαίες πλατφόρμες, όπως παραδείγματος χάριν το Scholar και το Scopus που αναφέρονται παραπάνω, η πληροφορία που οπτικοποιείται είναι μεμονωμένη και τόσο διάχυτη που αδυνατεί ο οποιοσδήποτε να εξάγει άμεσα ολοκληρωμένα και εμπειριστατωμένα συμπεράσματα. Στη συνέχεια, παρουσιάζεται ο σχεδιασμός και η υλοποίηση της διαδραστικής οπτικοποίησης που πραγματεύεται η εργασία και εμπεριέχει δεδομένα από όλα τα ερευνητικά εργαστήρια του Τμήματος Πληροφορικής και Τηλεπικοινωνιών του ΕΚΠΑ.

4. Η ΣΥΛΛΟΓΗ ΤΩΝ ΔΕΔΟΜΕΝΩΝ

4.1. Τα Δεδομένα

Τα δεδομένα που χρειάζονται για τη συγκεκριμένη οπτικοποίηση δεν υπάρχουν σε μια ενιαία βάση δεδομένων του τμήματος αλλά στην ιστοσελίδα εκάστου ερευνητικού εργαστηρίου που έχει αναρτήσει τις δικές του ερευνητικές δημοσιεύσεις. Επομένως, η συλλογή του δεδομένων από τις επιμέρους ιστοσελίδες δεν είναι δυνατή. Μία εναλλακτική προσέγγιση για την επίλυση του ανωτέρω προβλήματος, είναι η χρήση της διαθέσιμης πληροφορίας που μπορεί να αντληθεί από το Google Scholar⁵.

Τα δεδομένα συγκεντρώνονται από το Google Scholar με την βοήθεια ενός προγράμματος τύπου «Αυτόματου Συλλέκτη Δεδομένων Ιστοσελίδων» (Web Crawler ή αλλιώς Spider) (βλ. ΠΙΝΑΚΑΣ ΟΡΟΛΟΓΙΑΣ). Ο συγκεκριμένος τύπος προγράμματος αναλύεται παρακάτω.

Τα δεδομένα αποτελούν ένα από τα σημαντικότερα στοιχεία μιας οπτικοποίησης δεδομένων. Όπως αναφέρεται και παραπάνω, βασικό χαρακτηριστικό μιας οπτικοποίησης είναι η εύκολη πρόσβαση στα δεδομένα. Μία οπτικοποίηση δεν έχει νόημα εάν ο χρήστης αδυνατεί να προσπελάσει, να φιλτράρει και να προσαρμόσει τον τρόπο με τον οποίο φαίνονται τα δεδομένα, ώστε να πετύχει το στόχο του. Με άλλα λόγια, σε μία οπτικοποίηση ο χρήστης πρέπει να μπορεί να αλληλεπιδρά με τα δεδομένα.

Τα δεδομένα που χρησιμοποιούνται στην εργασία αφορούν στην δραστηριότητα των ερευνητικών εργαστηρίων του Τμήματος Πληροφορικής και Τηλεπικοινωνιών του ΕΚΠΑ. Δηλαδή τα δεδομένα που μας απασχολούν είναι ερευνητικές δημοσιεύσεις που πραγματοποιήθηκαν από τα ερευνητικά εργαστήρια από την ίδρυση του Τμήματος (1989) μέχρι και σήμερα (2020). Τα δεδομένα περιορίζονται στις δημοσιεύσεις δεκαεπτά (17) καθηγητών του εν λόγω τμήματος, οι οποίοι διαθέτουν προφίλ στο Google Scholar. Οι καθηγητές που δεν διαθέτουν προφίλ, καθώς και τα υπόλοιπα μέλη των ερευνητικών εργαστηρίων που δεν τυγχάνουν καθηγητές του Τμήματος εξαιρούνται από τα δεδομένα στα πλαίσια αυτής της εργασίας, ώστε να μπορούν αυτά να είναι διαχειρίσιμα. Παρόλα αυτά, ακολουθώντας την ίδια μεθοδολογία μπορούν σταδιακά να προστεθούν όλα τα μέλη των ερευνητικών εργαστηρίων. Στο σημείο αυτό αξίζει να σημειωθεί ότι ορισμένοι καθηγητές ανήκουν σε περισσότερα από ένα ερευνητικά εργαστήρια, γεγονός που δημιουργεί πρόβλημα στον τρόπο παρουσίασης των δεδομένων από την οπτικοποίηση. Κατωτέρω θα εξεταστεί πιο αναλυτικά το συγκεκριμένο πρόβλημα. Το τελικό αρχείο δεδομένων που παρουσιάζεται μέσα από αυτή την οπτικοποίηση αποτελείται από τρεις χιλιάδες επτακόσιες τρεις (3.703) ερευνητικές δημοσιεύσεις.

Στην παρούσα εργασία δίνεται μεγάλη έμφαση στην αλληλεπίδραση του χρήστη με τα δεδομένα, ενώ ακόμη του δίνεται η δυνατότητα να δει ολόκληρο το πλήθος δεδομένων που διατίθενται, να διαβάσει την ερευνητική δημοσίευση όταν ο σύνδεσμος είναι διαθέσιμος, καθώς και να περιηγηθεί στον ιστοχώρο από τον οποίο προέρχονται αυτά τα δεδομένα (Google Scholar).

Η αρχική μορφή των δεδομένων είναι σε μορφή bibtex (παράδειγμα στο Παράρτημα Ι). Τα δεδομένα λαμβάνονται («κατεβαίνουν») από το Google Scholar στην ως άνω

⁵ <https://scholar.google.com/>

μορφή, μέσω ενός προγράμματος τύπου Web Crawler που αναλύεται κάτωθι. Στη συνέχεια, χρησιμοποιώντας ένα πρόγραμμα σε Python μετατρέπονται σε μορφή csv (παράδειγμα στο Παράρτημα Ι). Με αυτό το αρχείο σε μορφή csv δημιουργείται και η δεύτερη διαμορφωτική αξιολόγηση που εξετάζεται στο επόμενο κεφάλαιο.

Το πρώτο αρχείο δεδομένων αποτελείται από δεκαοκτώ (18) στήλες (παράδειγμα του αρχείου στο Παράρτημα Ι) και τα βασικά πεδία που αφορούν την οπτικοποίηση αυτής της εργασίας είναι:

- id (Μοναδικός αριθμός εγγραφής)
- title (Τίτλος ερευνητικής δημοσίευσης)
- author (Συγγραφέας/-είς)
- year (Έτος δημοσίευσης)
- type (Τύπος δημοσίευσης)
- ResearchLab (Ερευνητικό εργαστήριο που έκανε την δημοσίευση)

Το τελικό αρχείο δεδομένων προέρχεται από έναν Crawler που «κατεβάζει» τα δεδομένα σε JSON μορφή, όπως θα εξεταστεί και στη συνέχεια του κεφαλαίου. Σε επόμενη φάση, χρησιμοποιώντας ένα πρόγραμμα γραμμένο σε Python, το JSON αρχείο μετατρέπεται σε αρχείο csv που περιέχει τα ίδια δεδομένα.

Ορισμένες από τις στήλες του αρχείου csv στο τελικό αρχείο δεδομένων έχουν πλέον ελληνικό όνομα όπως «Συγγραφείς» και «Ημερομηνία δημοσίευσης», ενώ το είδος του αρχείου αποτυπώνεται πλέον στην ελληνική γλώσσα (Βιβλίο, Επιστημονικό άρθρο, Διάσκεψη, Αναφορά, Ευρεσιτεχνία, Διάφορα).

4.2. Web Crawler (Spider)

Ένας Web Crawler ή αλλιώς Spider ή Spider Bot είναι ένα Internet Bot που συστηματικά περιηγείται στο World Wide Web, συνήθως για λόγους εύρεσης δεδομένων (Web Indexing). Ο τρόπος συλλογής των δεδομένων από τον Crawler είναι αυτόματος, ακολουθεί τη σύνδεση που υπάρχει σε μια ιστοσελίδα και συλλέγει τα δεδομένα που υπάρχουν αποθηκεύοντας αυτά, έως ότου ολοκληρώσει τη συλλογή των δεδομένων ή φτάσει σε αδιέξοδο. Στην τελευταία περίπτωση, συνεχίζει με μια νέα περιήγηση.

Οι Crawlers χρησιμοποιούνται κυρίως από τις διάφορες μηχανές αναζήτησης, προκειμένου να συλλέξουν στοιχεία για τη βάση δεδομένων τους, ώστε να εμφανίζουν ταχύτερα και πιο σχετικά αποτελέσματα αναζήτησης. Στην πραγματικότητα, όταν ένας χρήστης εισάγει κάποιον όρο αναζήτησης οι μηχανές αυτές δεν αναζητούν σε ολόκληρο το διαδίκτυο, αλλά εξετάζουν την βάση δεδομένων των ιστοσελίδων που έχουν συλλέξει οι Crawlers τους.

Ακόμη, υπάρχουν και μηχανισμοί για τις ιστοσελίδες που δεν επιθυμούν να ανιχνευθούν από κάποιο Crawler. Για παράδειγμα, μια ιστοσελίδα μπορεί να ζητήσει από τον Crawler να προσθέσει στο ευρετήριο του μόνο τμήματα αυτής ή ακόμη και κανένα τμήμα της.

Στην παρούσα εργασία κατασκευάζεται ένας Web Crawler για την αυτόματη συλλογή δεδομένων από την ιστοσελίδα Google Scholar. Η συγκεκριμένη διαδικασία κατασκευής του Web Crawler περιγράφεται ενδελεχώς παρακάτω.

4.3. Πρώτη Υλοποίηση Web Crawler

Όπως αναφέρεται και παραπάνω, τα δεδομένα που χρησιμοποιούνται στην εργασία συγκεντρώνονται από το Google Scholar με την βοήθεια ενός προγράμματος τύπου Web Crawler. Η δημιουργία του προγράμματος αυτού πραγματοποιείται τμηματικά, ώστε να φτάσει στην επιθυμητή μορφή. Για το στάδιο του σχεδιασμού δεν χρειάζονται πραγματικά δεδομένα. Κατόπιν της συλλογής ενός συγκεκριμένου μέρους των δεδομένων, δημιουργούνται τυχαία δεδομένα, χρησιμοποιώντας τα ήδη υπάρχοντα για να συμπληρώσουν τα απαιτούμενα για την οπτικοποίηση δεδομένα, χρησιμοποιώντας ένα πρόγραμμα γραμμένο σε Python.

Το πρόγραμμα που προαναφέρθηκε υλοποιείται με την βοήθεια μιας έτοιμης βιβλιοθήκης σε Python⁶ που βρέθηκε online. Ο κώδικας προσαρμόζεται αναλόγως, προκειμένου να πραγματοποιεί λήψη (να «κατεβάζει») πάνω από δέκα (10) εγγραφές αυτόματα -που είναι το ποσοτικό όριο του προγράμματος- και στην συνέχεια να τις συγκεντρώνει σε ένα αρχείο τύπου csv. Μετά από έναν αριθμό δεδομένων, η συγκεκριμένη εφαρμογή αναγκάζεται να σταματήσει την συλλογή των δεδομένων από το λογισμικό που περιορίζει την μη ανθρώπινη δραστηριότητα ονόματι reCAPTCHA⁷ (βλ. ΣΥΝΤΜΗΣΕΙΣ- ΑΡΚΤΙΚΟΛΕΞΑ- ΑΚΡΩΝΥΜΙΑ). Το συγκεκριμένο πρόβλημα αντιμετωπίζεται μετατρέποντας τον Crawler, ώστε να συλλέγει δεδομένα με μικρότερη ταχύτητα χρησιμοποιώντας την συνάρτηση «sleep()» της βιβλιοθήκης «sys» της Python. Ο Crawler ρυθμίστηκε να σταματάει την λήψη (με χρήση sleep) για μια μέρα (24 ώρες) και έπειτα να συνεχίζει.

Ένα πρόβλημα που δημιουργείται κατά τη διάρκεια συλλογής των δεδομένων είναι ότι όταν η εφαρμογή Google Scholar ενημερώνεται (Update), ο Web Crawler παύει να λειτουργεί. Για το λόγο αυτό, ο Web Crawler υλοποιήθηκε για δεύτερη φορά.

4.4. Δεύτερη Υλοποίηση Web Crawler

Κατά την δεύτερη προσέγγιση της υλοποίησης του Web Crawler, χρησιμοποιείται μια βιβλιοθήκη της Python, το Selenium⁸ και συγκεκριμένα το Selenium WebDriver, που αναλύεται παρακάτω.

Το Selenium είναι εν γένει μια εφαρμογή η οποία αυτοματοποιεί τους Web Browser (τα προγράμματα περιήγησης). Κατά κύριο λόγο αυτοματοποιεί εφαρμογές ιστού για σκοπούς δοκιμών, αλλά επεκτείνεται και σε άλλες χρήσεις.

Το Selenium WebDriver λειτουργεί χρησιμοποιώντας έναν Webdriver (πρόγραμμα οδήγησης περιηγητή), ώστε να διαχειρίζεται τον Web Browser και να περιηγείται αυτόματα στο Internet. Στα πλαίσια της πτυχιακής εργασίας γίνεται χρήση του

⁶ <https://github.com/pablooliveira/scholar.py>

⁷ <https://www.google.com/recaptcha/about/>

⁸ <https://selenium-python.readthedocs.io/>

«chromedriver.exe»⁹ ως Webdriver. Το Webdriver χρησιμοποιεί τον Web Browser όπως κάνει και ένας πραγματικός χρήστης, μέσω μιας διεπαφής API, με την διαφορά ότι το πρώτο το κάνει αυτόματα, είτε τοπικά είτε σε απομακρυσμένα μηχανήματα. Το WebDriver είναι απλό, καθώς έχει σχεδιαστεί ως μια συνοπτική και συμπαγής διεπαφή προγραμματισμού. Μέσω μιας απλής ρύθμισης, το WebDriver μπορεί να χρησιμοποιηθεί με όλα τα μεγάλα προγράμματα περιήγησης. Αυτοματοποιεί πραγματικές αλληλεπιδράσεις χρηστών σε Firefox, Safari, Edge, Chrome, Internet Explorer και άλλα.

Οι εντολές του Selenium που χρησιμοποιούνται για την υλοποίηση του web crawler είναι οι εξής:

- **webdriver.Chrome(«Η θέση του chromedriver.exe στον σκληρό δίσκο»);**

Η συγκεκριμένη εντολή αρχικοποιεί τον webdriver.

- **driver.get(SEARCH_LINK);**

Με αυτήν την εντολή ο webdriver θα προσπελάσει το link που δίνεται ως παράμετρος (SEARCH_LINK) και στη συνέχεια θα επιστρέψει την απάντηση (HTML response) σε ένα αίτημα (HTML request) που λαμβάνει ένας web browser όταν προσπαθεί να προσπελάσει το συγκεκριμένο link (SEARCH_LINK).

- **driver.find_elements_by_id();**

Επιστρέφει ένα στοιχείο (DOM element) από το html response σύμφωνα με το «element id». Στη συγκεκριμένη περίπτωση σαν όρισμα δίνεται το όνομα του συγγραφέα.

- **driver.execute_script("return document.documentElement.outerHTML");**

Η συγκεκριμένη εντολή εκτελεί ένα script στον Webdriver (return document.documentElement.outerHTML), το οποίο αποθηκεύει το HTML από το HTML response σε μία μεταβλητή ώστε να είναι δυνατή η περαιτέρω ανάλυση του.

- **driver.find_elements_by_class_name();**

Επιστρέφει ένα στοιχείο (DOM element) από το html response σύμφωνα με το «element class name». Στη συγκεκριμένη περίπτωση σαν όρισμα δίνεται το όνομα του συγγραφέα.

- **driver.quit();**

Η συγκεκριμένη εντολή κλείνει τον webdriver.

⁹ <https://chromedriver.chromium.org/>

Ο Crawler κατασκευάζεται ώστε να παίρνει σαν Input (είσοδο) αρχικά το όνομα του αρχείου εξόδου (βλ. παρακάτω «outPutFile.json») που θα αποθηκευτούν τα δεδομένα και στην συνέχεια κάποια Configuration Files (αρχεία διαμόρφωσης). Με αυτόν τον τρόπο, μετά το πέρας της εκτέλεσης του προγράμματος στο αρχείο εξόδου αποθηκεύονται τα απαιτούμενα δεδομένα, τα οποία αναλύονται στην πρώτη υποενότητα του κεφαλαίου, σε μορφή JSON.

Η εκτέλεση του Crawler γίνεται από την γραμμή εντολών των Windows και η σύνταξη της εντολής εκτέλεσης είναι η εξής:

- **python crawler.py outPutFile.json ConfigFile1.txt ConfigFile2.txt ConfigFile3.txt ...**

Η μορφή των Configuration Files (ConfigFile1.txt, ConfigFile2.txt, ConfigFile3.txt κ.λ.π.) παρατίθεται στο Παράρτημα Ι.

Η μορφή του JSON αρχείο εξόδου (outPutFile.json) παρατίθεται επίσης στο Παράρτημα Ι.

Ο συγκεκριμένος Crawler λειτουργεί ως εξής. Αρχικά «διαβάζει» το πρώτο Configuration File, στο οποίο αντιστοιχεί το προφίλ ενός συγγραφέα - καθηγητή στο Google Scholar. Στη συνέχεια, χρησιμοποιώντας το Webdriver, ελέγχει τον Web Browser όπως θα έκανε ένας πραγματικός χρήστης και πλοηγείται στο συγκεκριμένο προφίλ αποθηκεύοντας διαδοχικά κάθε μία δημοσίευση, η οποία παρατίθεται στο συγκεκριμένο προφίλ. Μόλις ολοκληρωθούν οι δημοσιεύσεις του συγκεκριμένου προφίλ «διαβάζει» το επόμενο Configuration File και επαναλαμβάνει τη διαδικασία για τον επόμενο προφίλ.

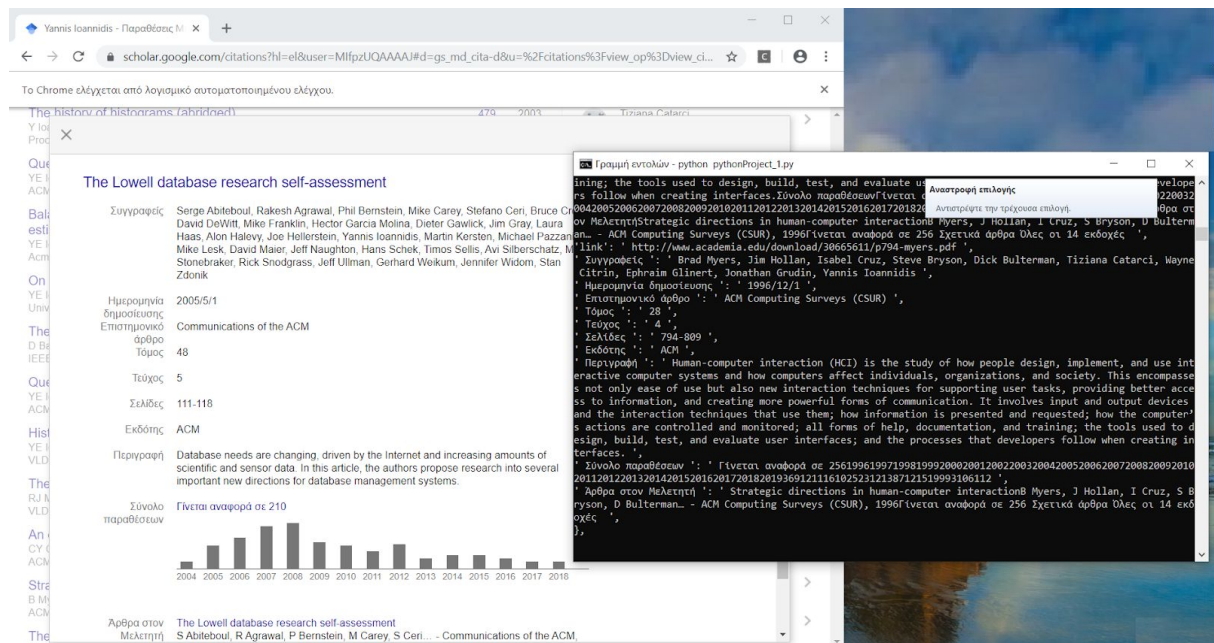
Μετά το πέρας της συγκεκριμένης εκτέλεσης, τα δεδομένα από μορφή JSON μετατρέπονται σε μορφή csv, ώστε να είναι έτοιμα για την αξιολόγηση, καθώς και τη διαχείριση τους από την τελική οπτικοποίηση. Για αυτήν τη μετατροπή χρησιμοποιήθηκε ένα πρόγραμμα σε Python, το οποίο εκτελείται στη γραμμή εντολών των Windows ακολούθως:

- **python jsonToCsv.py outFile.json**

Το συγκεκριμένο πρόγραμμα δίνει ως έξοδο ένα αρχείο με το ίδιο όνομα του αρχείου εισόδου, αλλά με διαφορετική κατάληξη, παραδείγματος χάρη «outFile.csv».

Στην εικόνα παρακάτω παρουσιάζεται ένα παράδειγμα εκτέλεσης της τελευταίας υλοποίησης του Crawler. Ακόμη, ο κώδικας της τελευταίας υλοποίησης του Crawler μαζί με το Python Script, το οποίο κάνει μετατροπή από αρχείο τύπου JSON σε CSV, βρίσκονται στην ακόλουθη διεύθυνση σε ένα Git Repository¹⁰.

¹⁰ <https://github.com/xenospan/webcrawler/>



Εικόνα 7: Στιγμιότυπο οθόνης από την εκτέλεση του Crawler.

Μετά από την ολοκλήρωση της συλλογής των δεδομένων, το επόμενο βήμα είναι ο σχεδιασμός της οπτικοποίησης. Για το σχεδιασμό πρέπει να καθορίζονται το αντικείμενο της απεικόνισης, ο τρόπος της απεικόνισης και η αλληλεπίδραση του χρήστη μαζί της. Τα στοιχεία αυτά δεν καθορίστηκαν πλήρως από τον πρώτο σχεδιασμό, αλλά σταδιακά και κατόπιν αλληλεπιδράσεων σχεδιασμών και αξιολογήσεων, χρησιμοποιώντας το ελικοειδές μοντέλο που αναφέρθηκε παραπάνω.

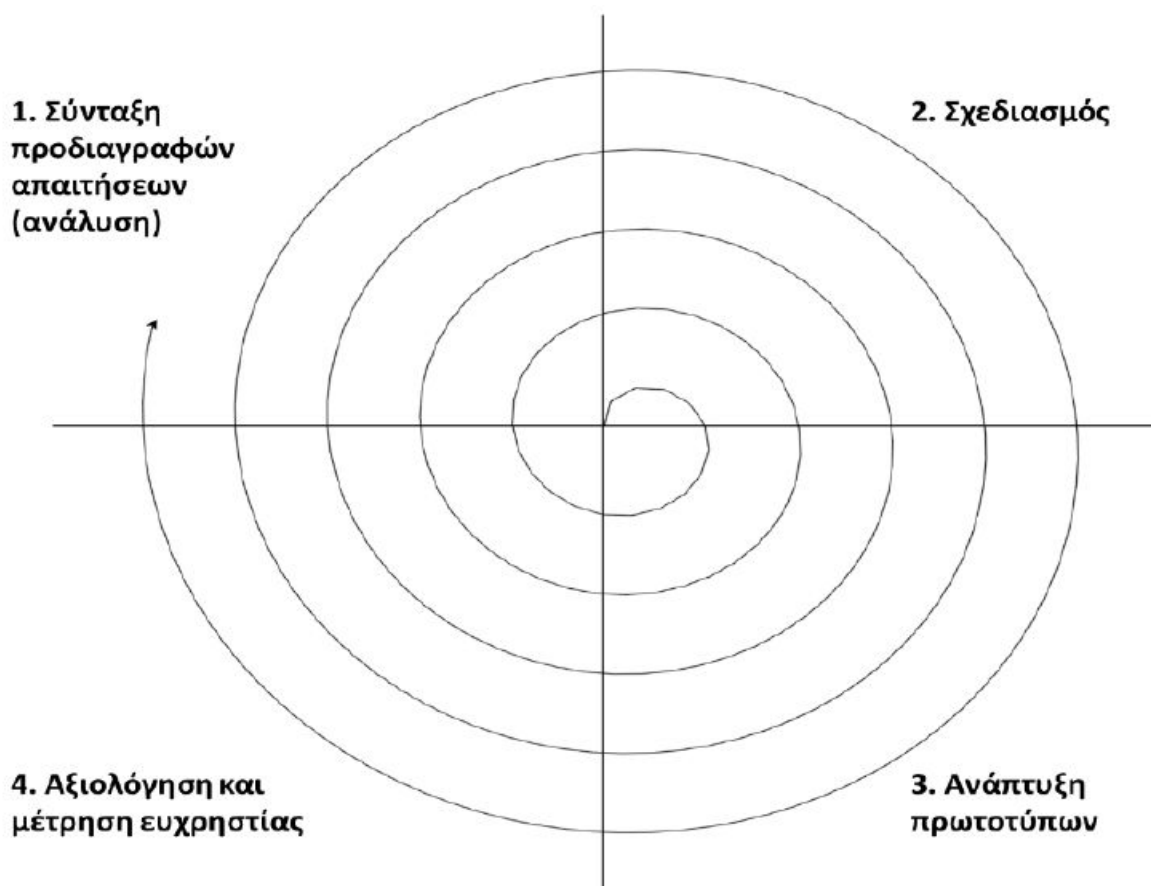
5. Ο ΣΧΕΔΙΑΣΜΟΣ

5.1. Μεθοδολογία Σχεδιασμού

Στην παρούσα εργασία, ακολουθείται η επαναληπτική προσέγγιση (iterative design) ή αλλιώς το Ελικοειδές Μοντέλο σχεδιασμού [8]. Η επαναληπτική προσέγγιση υλοποίησης εφαρμογών χρησιμοποιεί τα στάδια του σχεδιασμού, της υλοποίησης και της αξιολόγησης επαναληπτικά μέχρις ότου η υλοποίηση φτάσει σε ένα επιθυμητό αποτέλεσμα, σύμφωνα με τα συμπεράσματα της αξιολόγησης. Στο δεύτερο κεφάλαιο μαζί με άλλα είδη σχεδιαστικών μοντέλων, αναλύεται και το παρόν μοντέλο σχεδιασμού.

Αρχικά λαμβάνει χώρα ο σχεδιασμός της οπτικοποίησης και κατόπιν πραγματοποιείται μια πρώτη διαμορφωτική αξιολόγηση στην οποία τίθεται η βάση για την πρώτη υλοποίηση. Το επόμενο βήμα είναι η πρώτη υλοποίηση της οπτικοποίησης και στη συνέχεια ακολουθείται μια αξιολόγηση με χρήστες, προκειμένου να διαπιστωθεί κατά πόσο η τρέχουσα υλοποίηση είναι αποδοτική και εύχρηστη. Τέλος, προκειμένου να επιτευχθεί το επιθυμητό αποτέλεσμα, χρειάστηκε ακόμα μια διορθωτική υλοποίηση και μια εκ νέου συμπερασματική αξιολόγηση, στην οποία κρίθηκε ότι ο στόχος έχει επιτευχθεί.

Στο σχήμα παρακάτω βλέπουμε τα στάδια του ελικοειδούς μοντέλου.



Σχήμα 1: Σχήμα ελικοειδούς μοντέλου [28].

5.2. Σύνταξη Προδιαγραφών και Απαιτήσεων

Μια από τις πρώτες ενέργειες που πραγματοποιούνται είναι η σύνταξη προδιαγραφών και απαιτήσεων της οπτικοποίησης. Στην οπτικοποίηση αυτή δίνεται έμφαση στην ανάδειξη της δραστηριότητας του Τμήματος, δηλαδή με μια μόνο εικόνα να έχει τη δυνατότητα ο χρήστης να λάβει γνώση για το τί είδους ερευνητικές δημοσιεύσεις πραγματοποιεί το Τμήμα Πληροφορικής και Τηλεπικοινωνιών του ΕΚΠΑ, πάνω σε ποιά ερευνητικά πεδία και από ποιά ερευνητικά εργαστήρια. Επιπροσθέτως, δίνεται στον χρήστη και η δυνατότητα να δει όλα τα δεδομένα που είναι διαθέσιμα για κάθε δημοσίευση, εφόσον το επιθυμεί. Δηλαδή, ο πρωταρχικός στόχος είναι να δημιουργηθούν διαγράμματα, τα οποία δίνουν μια γενική πρώτη εικόνα για την ερευνητική δραστηριότητα του Τμήματος. Επιπλέον, σαν δεύτερη απαίτηση, ορίζεται η δυνατότητα προβολής των αναλυτικών δεδομένων από τον χρήστη. Για το λόγο αυτό, είναι αναγκαία η χρήση ενός διαγράμματος που εξυπηρετεί και αυτήν την απαίτηση. Τέλος, για την διευκόλυνση αναζήτησης κρίνεται απαραίτητη η χρήση φίλτρων, προκειμένου ο χρήστης να αναζητά την πληροφορία που θέλει και να αλλάζει τον τρόπο προβολής των δεδομένων μέσα από την οπτικοποίηση αυτή. Συνεπώς, χρειάζεται να δημιουργηθεί μια διαδραστική οπτικοποίηση, ώστε ο χρήστης να είναι σε θέση να διαμορφώνει τον τρόπο προβολής των δεδομένων ανάλογα με τις ανάγκες του.

Επομένως οι προδιαγραφές και οι απαιτήσεις εν συντομία είναι:

- Ανάδειξη της ερευνητικής δραστηριότητας του Τμήματος ανά έτος
- Ανάδειξη της ερευνητικής δραστηριότητας ανά ερευνητικό εργαστήριο, ερευνητικό τομέα ή θεματική περιοχή.
- Δυνατότητα αναλυτικής προβολής δεδομένων με την χρήση φίλτρων.
- Δυνατότητα να προβολής της εκάστοτε ερευνητικής δημοσίευσης.

Η λύση που προτείνεται, προκειμένου να ικανοποιεί με τον πιο αποτελεσματικό και εύχρηστο τρόπο όλα τα παραπάνω αναλύεται διεξοδικά στη συνέχεια.

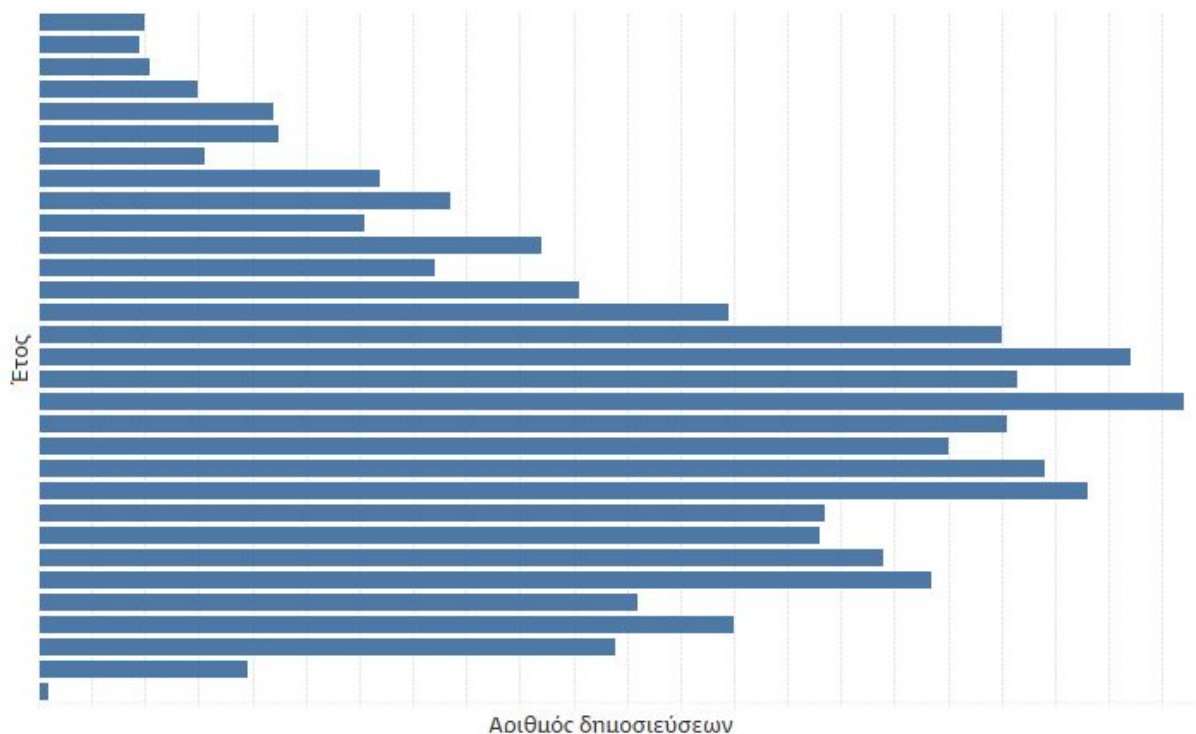
5.3. Ο Πρώτος Σχεδιασμός της Οπτικοποίησης Χαμηλής Πιστότητας

Για τον αρχικό σχεδιασμό δημιουργήθηκε ένα πρότυπο σε χαρτί και μετά μεταφέρθηκε στην εφαρμογή Paint των Windows. Αυτό το πρωτότυπο είναι αναγκαίο, προκειμένου να μεταδώσει την αρχική ιδέα και να κριθεί κατά πόσο είναι αποδεκτό, ώστε να δώσει λύση στο πρόβλημα και να μπορούμε να προχωρήσουμε σε περαιτέρω σχεδιασμό.

Τα διαγράμματα που σχεδιάστηκαν στο πρωτότυπο χαμηλής πιστότητας και χρησιμοποιήθηκαν για την οπτικοποίηση μέχρι τέλους είναι τρία:

1. Ραβδόγραμμα συνολικού αριθμού δημοσιεύσεων ανά έτος.

Σε αυτό το διάγραμμα απεικονίζεται ο συνολικός αριθμός δημοσιεύσεων ολόκληρου του Τμήματος ανά έτος. Οι ράβδοι είναι οριζόντιες και δεν περιέχουν χρωματική κωδικοποίηση, όπως θα παρατηρήσουμε ότι έχει το τελευταίο διάγραμμα της οπτικοποίησης. Η περίοδος που οπτικοποιείται ξεκινάει από το έτος 1989 και έπειτα, δηλαδή από την ίδρυση του Τμήματος έως σήμερα.



Εικόνα 8: Ραβδόγραμμα συνολικού αριθμού δημοσιεύσεων ανά έτος - Πρωτότυπο Χαμηλής Πιστότητας

2. Ραβδόγραμμα συνολικού αριθμού δημοσιεύσεων ανά ερευνητικό εργαστήριο.

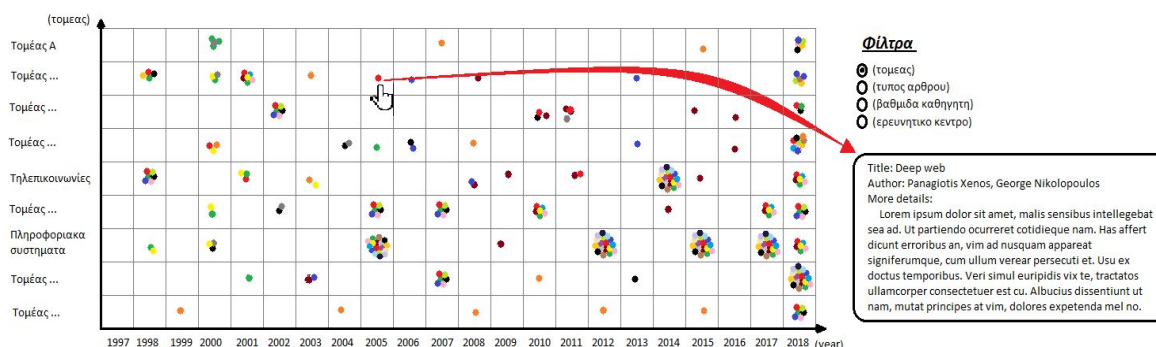
Σε αυτό το διάγραμμα φαίνεται ο συνολικός αριθμός δημοσιεύσεων όλου του Τμήματος ανά ερευνητικό εργαστήριο του Τμήματος. Οι ράβδοι είναι επίσης οριζόντιες και δεν περιέχουν χρωματική κωδικοποίηση, όπως θα δούμε παρακάτω ότι έχει το τελευταίο διάγραμμα της οπτικοποίησης.



Εικόνα 9: Ραβδόγραμμα συνολικού αριθμού δημοσιεύσεων ανά ερευνητικό εργαστήριο - Πρωτότυπο Χαμηλής Πιστότητας

3. Διάγραμμα σμήνους φυσαλίδων χρόνου ανά ερευνητικά εργαστήρια, με κάθε φυσαλίδα να είναι μια ερευνητική δημοσίευση.

Οι φυσαλίδες δημιουργούν σμήνος ανάλογα με το έτος που δημοσιεύθηκαν και το ερευνητικό εργαστήριο που ανήκουν. Έτσι δημιουργείται ένα πλέγμα με σμήνος από φυσαλίδες όπως στην εικόνα παρακάτω (Εικόνα 2). Τα χρώματα δείχνουν τον τύπο της ερευνητικής δημοσίευσης. Οι τύποι που μπορεί να έχει μια δημοσίευση είναι οι εξής: Άρθρο (article), Βιβλίο (book), Βιβλιάριο (booklet), Κομμάτι από βιβλίο (Inbook), Κομμάτι από συλλογή (Incollection), Πρακτικά Συνεδρίου (Inproceedings), Εγχειρίδιο (Manual), Μεταπτυχιακή εργασία (Mastersthesis), Διάφορα (Misc), Διδακτορική διατριβή (Phdthesis), Πρακτικά (Proceedings), Τεχνολογική αναφορά (Techreport), Αδημοσίευτο (Unpublished) [21].



Εικόνα 10: Διάγραμμα σμήνους φυσαλίδων χρόνου ανά ερευνητικά εργαστήρια, με κάθε φυσαλίδα να είναι μια ερευνητική δημοσίευση - Πρωτότυπο Χαμηλής Πιστότητας

5.4. Πρώτη Διαμορφωτική Αξιολόγηση - Αποτελέσματα

Η πρώτη απόφαση που λαμβάνεται στην εργασία είναι το πρωτότυπο χαμηλής πιστότητας της οπτικοποίησης. Το πρώτο διάγραμμα γίνεται στην εφαρμογή Paint των Windows, κατόπιν της έρευνας που πραγματοποιήθηκε σχετικά με τους πιο αποδοτικούς τρόπους οπτικοποίησης και παρουσιάζεται παραπάνω. Η αρχική ιδέα περιέχει δύο ραβδογράμματα και ένα διάγραμμα φυσαλίδων. Η ιδέα αυτή αναλύεται διεξοδικά στην πρώτη αξιολόγηση της εργασίας, που πραγματοποιήθηκε με την επιβλέπουσα Επίκουρη Καθηγήτρια Μαρία Ρούσσου. Η αξιολόγηση αυτή έχει ως στόχο να αποτελέσει τη βάση για την επόμενη αξιολόγηση με χρήστες που πραγματοποιείται στη συνέχεια αυτής της εργασίας.

Τα βασικότερα προβλήματα που παρατηρήθηκαν σε αυτήν την αξιολόγηση είναι:

- Η έλλειψη αλληλεπίδρασης των διαγραμμάτων με τον χρήστη καθιστώντας τα μη ρεαλιστικά, καθώς ο χρήστης πρέπει να υποθέτει τις αλλαγές που λαμβάνουν χώρα ανάλογα με τις ενέργειές του (οι αλλαγές γίνονται μέσω του προφορικού λόγου από το άτομο που διάγει την αξιολόγηση).
- Η αλλαγή της σημασίας των αξόνων στα ραβδογράμματα. Δηλαδή σε ορισμένα διαγράμματα ο οριζόντιος άξονας εμφανίζει τα έτη των δημοσιεύσεων και σε άλλα εμφανίζει τον συνολικό αριθμό των δημοσιεύσεων για έναν ερευνητικό τομέα. Αυτό δημιουργεί σύγχυση και έλλειψη συνέπειας, δυσχεραίνοντας την άμεση κατανόηση των διαγραμμάτων.

Για την αντιμετώπιση των παραπάνω προβλημάτων, χρησιμοποιείται το Tableau¹¹, μια εφαρμογή ανάλυσης και οπτικοποίησης δεδομένων που παρουσιάζεται παρακάτω.

Ο σχεδιασμός στο Tableau γίνεται τμηματικά και τροποποιείται πολλές φορές μέχρι να φτάσει στο τελικό στάδιο που χρησιμοποιείται για τη δεύτερη αξιολόγηση της εργασίας, ήτοι την αξιολόγηση με χρήστες. Με βάση τα παραπάνω συμπεράσματα η νέα οπτικοποίηση πρέπει να είναι διαδραστική και ταυτόχρονα απλή, ώστε ο

¹¹ <https://www.tableau.com/>

χρηστής να «παίρνει μια ιδέα» της πληροφορίας που είναι διαθέσιμη χωρίς να υποχρεούται να τη δει αναλυτικά.

5.5. Ο Δεύτερος Σχεδιασμός της Οπτικοποίησης Υψηλής Πιστότητας

Όπως αναφέρθηκε προηγουμένως, το πρωτότυπο χαμηλής πιστότητας είναι αναγκαίο για τη διαδικασία σχεδιασμού, ώστε να καθοριστούν πλήρως οι προδιαγραφές και οι απαιτήσεις για την υπό κατασκευή οπτικοποίηση. Παρόλα αυτά, η ανάγκη για δημιουργία ενός πρωτοτύπου πιο κοντά στο τελικό αποτέλεσμα είναι επιτακτική, καθώς το πρωτότυπο χαμηλής πιστότητας δεν είναι επαρκές. Συνεπώς, το επόμενο πρωτότυπο, υψηλής πιστότητας, είναι το επόμενο βήμα που ακολουθείται σε αυτήν την εργασία. Το πρωτότυπο αυτό δημιουργείται σε μια πλατφόρμα ανάλυσης δεδομένων και σχεδιασμού διαδραστικών οπτικοποιήσεων με την ονομασία «Tableau» και η οποία παρουσιάζεται αναλυτικά παρακάτω.

5.5.1. Τι Είναι το Tableau

Το Tableau είναι μια πλατφόρμα ανάλυσης δεδομένων. Είναι ένα εμπορικό εργαλείο κατάλληλο για μεμονωμένους αναλυτές, για οργανισμούς και επιχειρήσεις, καθώς και για ανάλυση δεδομένων και οπτικοποίηση αυτών με σκοπό την πιο έξυπνη και γρήγορη κατανόηση αυτών από τους χρήστες.

Η συγκεκριμένη πλατφόρμα είναι κατάλληλη για επαγγελματική χρήση, λόγω της μεγάλης ευελιξίας διαχείρισης των δεδομένων και της δυνατότητας που δίνει στον χρήστη να οπτικοποιήσει τα δεδομένα με μια πληθώρα διαγραμμάτων. Ο χρήστης μπορεί να επεξεργάζεται ευρύ όγκο δεδομένων (στατικών ή και διαρκώς μεταβαλλόμενων) και να δημιουργεί τη διαδραστική οπτικοποίηση που επιθυμεί, ώστε να διευκολύνει την κατανόηση και την παρατήρηση αυτών από διάφορες οπτικές γωνίες. Με την χρήση του Tableau επιτυγχάνεται η παρουσίαση των δεδομένων με έναν οπτικό και διαδραστικό τρόπο, προκειμένου ο χρήστης να μπορεί να αναγνωρίσει πρότυπα ή μοτίβα στα δεδομένα.

Τα προϊόντα Tableau μπορούν να περιγράψουν τις σχεσιακές βάσεις δεδομένων (Relational Databases), τους κύβους επεξεργασίας ηλεκτρονικών αναλύσεων (Online Analytical Processing Cubes), τις βάσεις δεδομένων στο σύννεφο (Cloud Databases) και τα υπολογιστικά φύλλα (Spreadsheets) για την παραγωγή οπτικοποιήσεων δεδομένων τύπου γραφικών παραστάσεων. Τα προϊόντα μπορούν επίσης να εξάγουν, να αποθηκεύουν και να ανακτούν δεδομένα από μία μηχανή δεδομένων εντός μνήμης (In-memory Data Engine).

5.5.2. Ο Σχεδιασμός στο Tableau - Φάση Φόρτωσης Δεδομένων

Η σχεδίαση στο Tableau ξεκινάει εισάγοντας τα δεδομένα που συνελέγησαν από το Scholar στην εφαρμογή του Tableau, ώστε να είναι διαθέσιμα στο πρόγραμμα για επεξεργασία. Αυτή η λειτουργία γίνεται στην καρτέλα του Tableau «Data Source» (κάτω αριστερά). Τα δεδομένα περιλαμβάνονται σε ένα αρχείο csv (βλ. ακρωνύμια), το οποίο και φορτώνεται στην εφαρμογή του Tableau.

FileData Server Window Help

outputfile

outputfile

Text file

Files

Use Data Interpreter

Data interpreter might be able to clean your Text file workbook.

outputfile.tsv

New Union

outputfile

Connection

LiveExtract

Filters

0Add

Sort fields

Modified

Show analyses

Show hidden fields

1,000

rows

#	Abc	Abc	Abc	Abc	Abc	Abc	Abc	Abc	Abc	Abc	Abc	Abc	Abc	Abc	Abc	Abc	Abc	Abc	Abc	Abc	Abc	Abc	Abc	Abc	Abc	Abc	Abc	Abc	Abc	Abc	
outputfile.tsv	outputfile.tsv	outputfile.tsv	outputfile.tsv	outputfile.tsv	outputfile.tsv	outputfile.tsv	outputfile.tsv	outputfile.tsv	outputfile.tsv	outputfile.tsv	outputfile.tsv	outputfile.tsv	outputfile.tsv	outputfile.tsv	outputfile.tsv	outputfile.tsv	outputfile.tsv	outputfile.tsv	outputfile.tsv	outputfile.tsv	outputfile.tsv	outputfile.tsv	outputfile.tsv	outputfile.tsv	outputfile.tsv	outputfile.tsv	outputfile.tsv	outputfile.tsv	outputfile.tsv	outputfile.tsv	outputfile.tsv
F1	publisher	school	ResearchLabL...	title	journal	author	author - Split 1	author - Split 2	author - Split 3	author - Split 4																					
0	IEEE	null	http://www.mad...	Indexing animat...	IEEE Transaction...	Kollios George a...	Kollios George	Tsotras Vassilis J.	Gunopulos Dimit...	Delis Alex																					
1	null	null	http://www.mad...	GPS-Free node lo...	null	Alkan H("useyin ...	Alkan H("useyin ...	Kriakov Vassil	Br("onnimann H...	Delis Alex																					
2	null	null	http://www.mad...	Performance and...	null	Delis Alex and Ro...	Delis Alex	Roussopoulos Nick																							
3	null	null	http://www.mad...	Flexible use of cl...	null	Tsakalozos Konst...	Tsakalozos Konst...	Kilapi Herald	Sitaridi Eva	Roussopoulos M...																					
4	Citeseer	University of Cali...	http://www.mad...	The ADMS projec...	null	Roussopoulos Ni...	Roussopoulos Nick	Chen Chungmin M	Kelley Stephen	Delis Alex																					
5	null	null	http://www.mad...	Reaching availab...	null	Verroios Vasilis a...	Verroios Vasilis	Efstathiou Vasilis	Delis Alex																						
6	IEEE	null	http://www.mad...	Performance co...	IEEE Transaction...	Delis Alex and Ro...	Delis Alex	Roussopoulos Nick																							
7	American Societ...	null	http://www.mad...	Automatic fire-c...	Journal of compu...	Delis Ethymios A ...	Delis Ethymios A	Delis Alex																							
8	null	null	http://www.mad...	Nefeli: Hint-base...	null	Tsakalozos Konst...	Tsakalozos Konst...	Roussopoulos M...	Floros Vangelis	Delis Alex																					
9	null	null	http://www.mad...	Another outlier b...	null	Deligiannakis An...	Deligiannakis An...	Kotidis Yannis	Vassalos Vasilis	Stoumpos Vassilis																					
10	IEEE	null	http://www.mad...	Reaching availab...	IEEE Transaction...	Delis Alex and Ro...	Delis Alex	Roussopoulos Nick																							

Data Source

Sheet1Sheet2Sheet3Dashboard2Dashboard3Story1

Εικόνα 11: Φόρτωση του αρχείου δεδομένων (outfile.csv) στο Tableau.

Έπειτα από τη φόρτωση του αρχείου δεδομένων στο Tableau, γίνεται επεξεργασία των δεδομένων, προκειμένου να έχουν την επιθυμητή μορφή. Ορισμένα από τα δεδομένα που δεν είναι στη σωστή μορφή είναι τα εξής:

- Η στήλη έτος - η οποία περιλαμβάνει έτη είτε σε μορφή συμβολοσειράς, η οποία δεν είναι διαχειρίσιμη, αφού είναι αδύνατο να ταξινομηθεί, είτε περιέχει άκυρους τύπους δεδομένων, όπως παραδείγματος χάριν το «null» (βλ. πίνακα ορολογίας) και πρέπει να καθαριστεί.
- Η στήλη των συγγραφέων - η οποία περιλαμβάνει όλους τους συγγραφείς μαζί, χωρισμένους από τους λατινικούς χαρακτήρες «and» και αυτό δυσκολεύει την επεξεργασία των δεδομένων.
- Ορισμένα δεδομένα τα οποία παρουσιάζουν άκυρους χαρακτήρες.
- Δεν εμφανίζεται στήλη με τον αριθμό των δημοσιεύσεων ανά έτος ή ανά ερευνητικό εργαστήριο.

Για τους παραπάνω λόγους είναι αναγκαία η δημιουργία επιπλέον στηλών μέσα στο Tableau.

Μία επιπλέον στήλη που δημιουργείται και προστίθεται στο αρχείο δεδομένων είναι η στήλη για τις χρονολογίες. Από τον αλφαριθμητικό τύπο της μεταβλητής χρονολογιών γίνεται μετατροπή σε αριθμητικό και στις περιπτώσεις που η χρονολογία δεν είναι διαθέσιμη και υπάρχει «null», τοποθετείται στο κελί ο αριθμός 0 (μηδέν). Επίσης, μερικές ημερομηνίες έχουν την μορφή «2020.0» και μετατρέπονται στην μορφή «2020». Τέλος, οι ημερομηνίες που είναι πριν από το έτος 1989, δηλαδή την ίδρυση του Τμήματος Πληροφορικής και Τηλεπικοινωνιών, φιλτράρονται, ώστε να μην παρουσιάζονται στην οπτικοποίηση. Κατά αυτόν τον τρόπο γίνεται αργότερα πιο εύκολος ο διαχωρισμός και η ταξινόμηση των χρονολογιών στο διάγραμμα.

Στην πορεία, προστίθενται δύο επιπλέον μεταβλητές, οι οποίες περιέχουν τον συνολικό αριθμό των δημοσιεύσεων για κάθε χρονολογία και για κάθε ερευνητικό εργαστήριο, αντίστοιχα. Αυτή η πληροφορία δεν είναι διαθέσιμη άμεσα στο αρχείο δεδομένων, αλλά χρειάζεται για να δημιουργηθούν δύο από τα βασικά διαγράμματα της οπτικοποίησης.

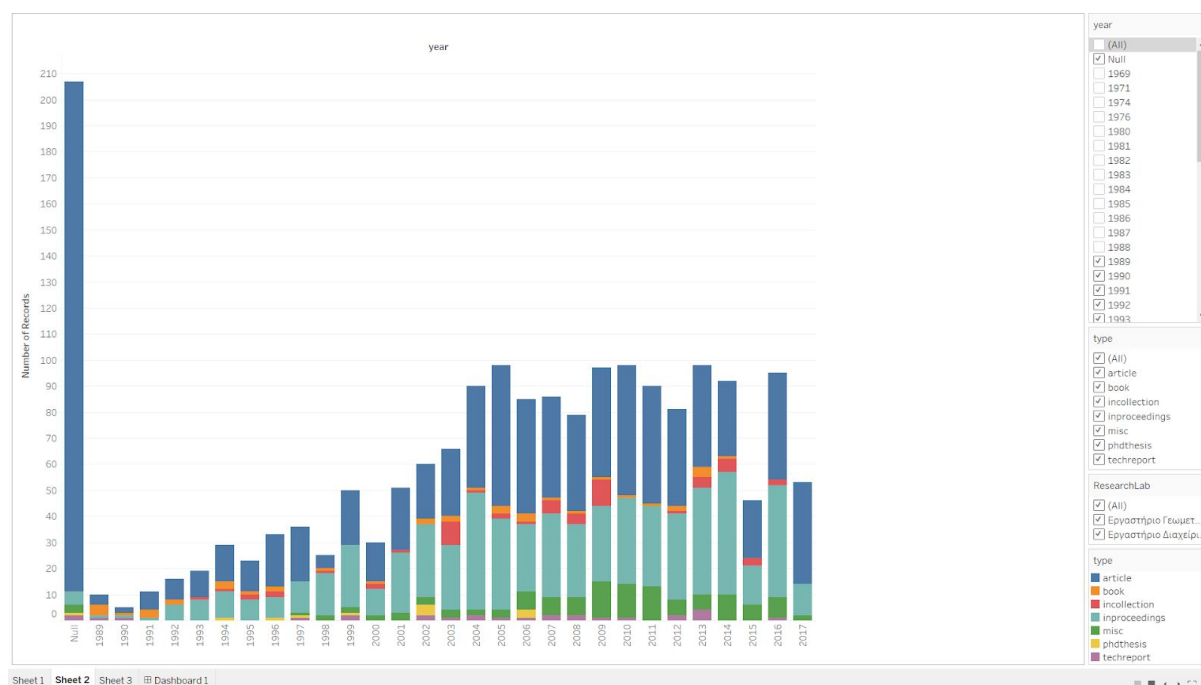
5.5.3. Ο Σχεδιασμός στο Tableau - Φάση Δημιουργίας Διαγραμμάτων

Στην υποενότητα αυτή αναλύεται η δημιουργία διαγραμμάτων υψηλής πιστότητας στο Tableau.

Μετά τη φόρτωση των δεδομένων στο Tableau και τη δημιουργία των επιπλέον μεταβλητών όπως εξετάζεται και παραπάνω, με την επιλογή “New WorkSheet” δημιουργείται ένα νέο Sheet και ξεκινάει η δημιουργία των οπτικοποιήσεων. Στη φάση του σχεδιασμού, δημιουργούνται τρία Sheets και ένα Dashboard, το οποίο ενσωματώνει τα ως άνω τρία Sheets, τα οποία είναι συνδεδεμένα σε μία ενιαία οπτικοποίηση και αλληλεπιδρούν μεταξύ τους με τη χρήση φίλτρων. Στο εν λόγω στάδιο επανασχεδιάστηκαν τα ίδια διαγράμματα που δημιουργήθηκαν κατά τον πρώτο σχεδιασμό χαμηλής πιστότητας και τροποποιήθηκαν σύμφωνα με την πρώτη διαμορφωτική αξιολόγηση. Σε αυτό το σημείο αξίζει να σημειωθεί, ότι λόγω του σημαντικού χρόνου που χρειαζόταν ο Crawler, προκειμένου να κατεβάσει τα δεδομένα, ο κώδικας του οποίου παράλληλα έπρεπε να τροποποιηθεί, ώστε να επιφέρει τα επιδιωκόμενα αποτελέσματα της εργασίας, υπήρχε μία συνεχιζόμενη προσθήκη δεδομένων από την αρχή της εργασίας μέχρι το τέλος αυτής. Το γεγονός αυτό δεν δημιούργησε πρόβλημα στο σχεδιασμό της εφαρμογής, ωστόσο σε κάθε μεταβατικό στάδιο έπρεπε να γίνουν οι απαραίτητες τροποποιήσεις, ώστε να συμβαδίζουν οι σχεδιαστικές απαιτήσεις με τον όγκο των δεδομένων της εκάστοτε χρονικής στιγμής.

Αρχικά δημιουργήθηκαν τα δύο bar charts. Με τη χρήση του Tableau, η δημιουργία των δύο Bar Chart είναι αρκετά εύκολη, καθώς όπως προαναφέρεται, το Tableau παρέχει ένα γραφικό και διαδραστικό περιβάλλον στο οποίο μερικά “Drag and Drop” είναι αρκετά, ώστε να δώσουν το επιθυμητό αποτέλεσμα.

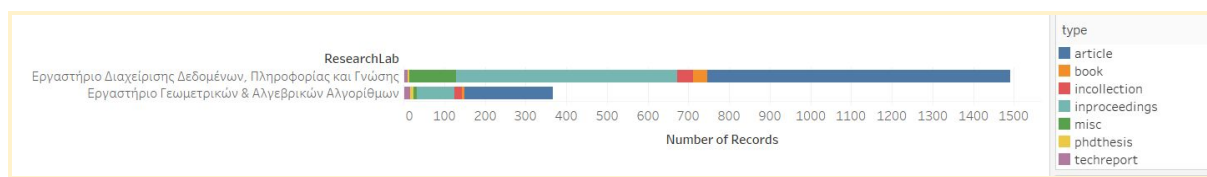
Το πρώτο Bar Chart που σχεδιάστηκε στο Tableau παρουσιάζει το συνολικό αριθμό των δημοσιεύσεων από τα ερευνητικά εργαστήρια του Τμήματος Πληροφορικής και Τηλεπικοινωνιών σε συνάρτηση με το χρόνο. Για τη δημιουργία του συγκεκριμένου διαγράμματος επιλέχθηκε η μεταβλητή “year” (δηλαδή η χρονιά που δημοσιεύτηκε η εκάστοτε δημοσίευση) από το πεδίο Dimensions στα αριστερά και έγινε Drag and drop επάνω στο Columns (Στήλες). Ακολούθως, με τον ίδιο τρόπο επιλέχθηκε η μεταβλητή “number of records” ως Rows (γραμμές). Η μεταβλητή number of records ορίστηκε ως “Measure(Sum)” και “Continuous”, ρυθμίσεις που πρέπει να οριστούν για αυτή τη μεταβλητή, ώστε να λειτουργήσει σωστά η οπτικοποίηση. Επιπλέον, η μεταβλητή type έγινε drag and drop πάνω στο Color (το οποίο βρίσκεται στα αριστερά μέσα στην κατηγορία Marks). Τέλος, επιλέχθηκαν οι μεταβλητές “ResearchLab” και “type” για φίλτρα (filters) τύπου “Multiple Values List”. Η κάτωθι εικόνα απεικονίζει το διάγραμμα που περιγράφεται ανωτέρω.



Εικόνα 12: Δημιουργία ραβδογράμματος ερευνητικών δημοσιεύσεων του (DIT) ανά έτος, στο Tableau (χρωματική απεικόνιση ανάλογα με το είδος της δημοσίευσης)

Σε αυτόν το σχεδιασμό ο χρόνος απεικονίζεται σε έτη, στον κάθετο άξονα εμφανίζεται ο αριθμός των δημοσιεύσεων και η χρωματική απεικόνιση συμβολίζει το type (είδος) της ερευνητικής δημοσίευσης (το τι είδους μπορεί να είναι μία ερευνητική δημοσίευση, αναλύεται παραπάνω κατά την εξέταση του πρώτου σχεδιασμού της οπτικοποίησης).

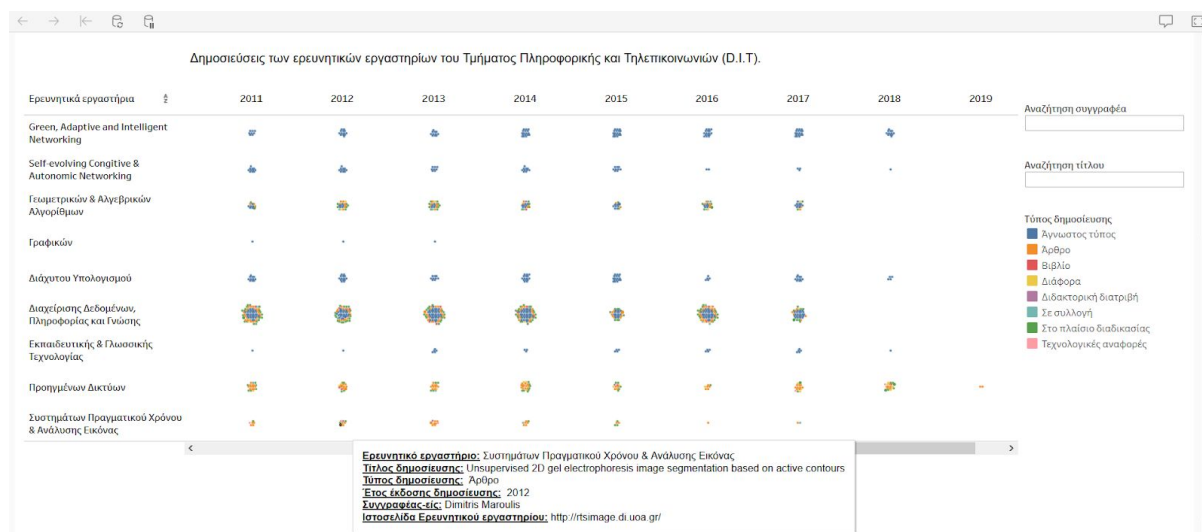
Το δεύτερο Bar Chart που σχεδιάστηκε στο Tableau απεικονίζει τον συνολικό αριθμό των δημοσιεύσεων για κάθε ένα από τα ερευνητικά εργαστήρια του Τμήματος Πληροφορικής και Τηλεπικοινωνιών. Για τη δημιουργία αυτού του διαγράμματος επιλέχθηκε η μεταβλητή "ResearchLab" (δηλαδή το ερευνητικό εργαστήριο από το οποίο έγινε η εκάστοτε δημοσίευση) από το πεδίο dimensions στα αριστερά και έγινε drag and drop επάνω στο Rows (γραμμές). Στη συνέχεια, με τον ίδιο τρόπο επιλέχθηκε η μεταβλητή "number of records" σαν Columns (Στήλες). Η μεταβλητή "number of records" ορίστηκε ως "Measure(Sum)" και "Continuous", ρυθμίσεις που πρέπει να οριστούν για αυτήν τη μεταβλητή, ώστε να λειτουργήσει σωστά η οπτικοποίηση. Τέλος, και σε αυτό το διάγραμμα όπως και στο προηγούμενο, η μεταβλητή "type" έγινε drag and drop πάνω στο Color (το οποίο βρίσκεται στα αριστερά μέσα στην κατηγορία Marks) και επιλέχθηκαν οι μεταβλητές "ResearchLab" και "type" για φίλτρα (filters) τύπου "Multiple Values List". Η κάτωθι εικόνα απεικονίζει το διάγραμμα που περιγράφεται ανωτέρω.



Εικόνα 13: Δημιουργία ραβδογράμματος συνολικών ερευνητικών δημοσιεύσεων ανά ερευνητικό εργαστήριο του (DIT), στο Tableau (χρωματική απεικόνιση ανάλογα με το είδος της δημοσίευσης)

Στον οριζόντιο άξονα απεικονίζονται δύο από τα εργαστήρια του Τμήματος και συγκεκριμένα τα εργαστήρια για τα οποία υπήρχαν επαρκή δεδομένα από τον Crawler εκείνη τη χρονική στιγμή, ενώ στον κάθετο άξονα απεικονίζεται ο αριθμός των δημοσιεύσεων. Η χρωματική απεικόνιση συμβολίζει το type (είδος) της ερευνητικής δημοσίευσης (το είδος της ερευνητικής δημοσίευσης αναλύεται παραπάνω κατά την εξέταση του πρώτου σχεδιασμού της οπτικοποίησης).

Στη συνέχεια το επόμενο βήμα είναι η δημιουργία του διαγράμματος φυσαλίδων στο Tableau, ώστε να αποτυπώνεται όσο πιο πιστά και πιο ρεαλιστικά είναι εφικτό το συγκεκριμένο διάγραμμα (διάγραμμα υψηλής πιστότητας) σε σχέση με την πρώτη αξιολόγηση στην οποία δεν υπάρχει διάδραση. Το διάγραμμα φυσαλίδων αποτελείται από σμήνη φυσαλίδων όπου κάθε φυσαλίδα αντιπροσωπεύει μία ερευνητική δημοσίευση και τα οποία διατάσσονται ανάλογα με το ερευνητικό εργαστήριο από το οποίο δημοσιεύθηκαν και με την τεχνολογία στην οποία δημοσιεύθηκαν. Ο οριζόντιος άξονας αντιπροσωπεύει το χρόνο και είναι σε χρόνια, ενώ ο κατακόρυφος άξονας εμφανίζει τα ερευνητικά εργαστήρια του Τμήματος Πληροφορικής και Τηλεπικοινωνιών. Στην εικόνα 14 παρακάτω οπτικοποιούνται μόνο δύο ερευνητικά εργαστήρια για τα οποία, όπως αναφέρεται παραπάνω, υπήρχαν επαρκή δεδομένα από τον Crawler μέχρι εκείνη τη χρονική στιγμή. Για τη δημιουργία αυτού του διαγράμματος πραγματοποιήθηκαν τα ακόλουθα βήματα. Με τη χρήση drag and drop η μεταβλητή year επιλέχθηκε ως Columns, ενώ η μεταβλητή ResearchLab ως Rows. Στο πεδίο Marks (στα αριστερά στην εφαρμογή του Tableau), οι παρακάτω μεταβλητές έχουν την ακόλουθη αντιστοιχία: η μεταβλητή “type” επιλέχθηκε ως Color, η μεταβλητή “title” επιλέχθηκε ως Detail, η μεταβλητή “Number of Records” επιλέχθηκε ως Size (με ρυθμίσεις “Measure(Sum)” και “Continuous” όπως και παραπάνω). Επίσης, τα απαιτούμενα πεδία όπως όλα τα παραπάνω, καθώς και η μεταβλητή “author” επιλέχθηκαν ως tooltip, δηλαδή να εμφανίζονται με το hover του ποντικιού πάνω στην κάθε φυσαλίδα. Τέλος, τοποθετήθηκαν ορισμένα φίλτρα, ώστε οι χρήστες να έχουν τη δυνατότητα να ελέγχουν το εμφανιζόμενο αποτέλεσμα και να αλληλεπιδρούν με την οπτικοποίηση. Παρακάτω απεικονίζεται η εν λόγω οπτικοποίηση κατά τη φάση της δημιουργίας της.



Εικόνα 14: Δημιουργία διαγράμματος φυσαλίδων ερευνητικών δημοσιεύσεων του (DIT) ανά έτος και ανά ερευνητικό εργαστήριο, στο Tableau (χρωματική απεικόνιση ανάλογα με το είδος της δημοσίευσης)

5.6. Δεύτερη Διαμορφωτική Αξιολόγηση με Χρήστες - Αποτελέσματα

Όπως εξετάζεται παραπάνω στο κεφάλαιο 2.6. «Η σημασία της αξιολόγησης», καταλήγουμε στο συμπέρασμα ότι οποιαδήποτε εφαρμογή κατά τη διάρκεια της σχεδιάσής της χρειάζεται να υποβληθεί αρκετές φορές στη διαδικασία της αξιολόγησης, προκειμένου να φτάσει στο επιθυμητό αποτέλεσμα. Στην προηγούμενη και πρώτη διαμορφωτική αξιολόγηση παρουσιάζονται αρχικώς τα δύο βασικότερα προβλήματα του συγκεκριμένου σχεδιασμού και κατόπιν οι λύσεις αυτών, οι οποίες οδηγούν στην αλλαγή της μορφής του σχεδιασμού. Υπάρχει όμως επιτακτική ανάγκη για μία δεύτερη διαδικασία αξιολόγησης με μεγαλύτερο αριθμό συμμετεχόντων, ώστε να εντοπίζονται τυχόν προβλήματα, προκειμένου να αποκατασταθούν πριν από την υλοποίηση της οπτικοποίησης. Για το λόγο αυτό, πραγματοποιήθηκε μία αξιολόγηση με τη συμμετοχή δεκαπέντε (15) ατόμων - συμμετεχόντων, πέντε (5) γυναικών και δέκα (10) άντρων, η διεξαγωγή της οποίας με τα ανάλογα συμπεράσματα αναλύεται κατωτέρω.

Για την προετοιμασία της δεύτερης διαμορφωτικής αξιολόγησης με χρήστες πραγματοποιήθηκαν τα ακόλουθα βήματα. Αρχικά, μία από τις πρώτες ενέργειες που έπρεπε να γίνουν ήταν η δημιουργία μιας φόρμας συγκατάθεσης προς συμπλήρωση από τους συμμετέχοντες, η οποία θα ενημερώνει τους τελευταίους για την πρόθεση χρήσης τεχνητών μέσων (ηχογράφησης/βιντεοσκόπησης) και θα ζητά τη συγκατάθεσή τους. Ο λόγος έγκειται στο γεγονός ότι για τη διαδικασία αξιολόγησης κρίνεται σκόπιμη η ηχογράφηση των συμμετεχόντων, με την οποία θα καταγράφονται οι απαντήσεις των τελευταίων, ούτως ώστε ο ερευνητής να έχει τη δυνατότητα να ανατρέχει σε αυτές σε οποιαδήποτε χρονική στιγμή. Συγκεκριμένα, στη δημιουργηθείσα φόρμα συγκατάθεσης αναφέρεται ο σκοπός της συγκεκριμένης έρευνας, αναλύεται ο τρόπος διεξαγωγής της έρευνας, γίνεται ενημέρωση για το ενδεχόμενο της ηχογράφησης προς διευκόλυνση της διεξαγωγής της έρευνας και ότι είναι εθελοντική, ενώ οι συμμετέχοντες μπορούν να αποχωρήσουν οποιαδήποτε στιγμή αυτοί επιθυμούν χωρίς καμία οφειλή ή υποχρέωση. Επίσης, ζητείται από τους συμμετέχοντες η συγκατάθεση τους για τη συμμετοχή στην έρευνα. Τέλος, όπως

αναφέρεται και στη φόρμα συγκατάθεσης η συγκεκριμένη έρευνα συμμορφώνεται με τους κανονισμούς της Ευρωπαϊκής Ένωσης για τη Γενική Προστασία Δεδομένων (GDPR) και μόνο ο ερευνητής έχει πρόσβαση στα δεδομένα και διατηρείται η ανωνυμία των προσωπικών στοιχείων των συμμετεχόντων. Στο Παράρτημα I παρατίθεται η φόρμα συγκατάθεσης που χρησιμοποιήθηκε.

Εν συνεχεία, το επόμενο βήμα ήταν η προετοιμασία της διαδικασίας αξιολόγησης, καθώς και η δημιουργία των σεναρίων χρήσης τα οποία οι χρήστες θα εκτελέσουν. Όπως αναλύεται στο κεφάλαιο 2 (ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ), τα σενάρια χρήσης προσομοιώνουν τον τρόπο χρήσης της εκάστοτε εφαρμογής και βοηθούν στη διαπίστωση τυχόν προβλημάτων. Τα σενάρια χρήσης τα οποία δημιουργήθηκαν είναι τα ακόλουθα:

Σενάριο χρήσης 1: Πόσες είναι οι συνολικές δημοσιεύσεις του Τμήματος Πληροφορικής και Τηλεπικοινωνιών του ΕΚΠΑ σε ένα συγκεκριμένο έτος, όπως για παράδειγμα το έτος “2006”.

Σενάριο χρήσης 2: Ποιες είναι οι δημοσιεύσεις ενός συγκεκριμένου συγγραφέα, όπως για παράδειγμα “Μαρία Ρούσσου”.

Σενάριο χρήσης 3: Αναζήτηση με βάση τον τίτλο δημοσίευσης, όπως για παράδειγμα “Learning by doing and learning through play: an exploration of interactivity in virtual environments for children”.

Σενάριο χρήσης 4: Σε ένα συγκεκριμένο έτος και για ένα συγκεκριμένο εργαστήριο να βρεθούν πόσες δημοσιεύσεις συγκεκριμένου είδους υπάρχουν, όπως για παράδειγμα “Το 2006, από το εργαστήριο MaDgIK, πόσες βιβλιογραφικές αναφορές δημοσιεύτηκαν;”

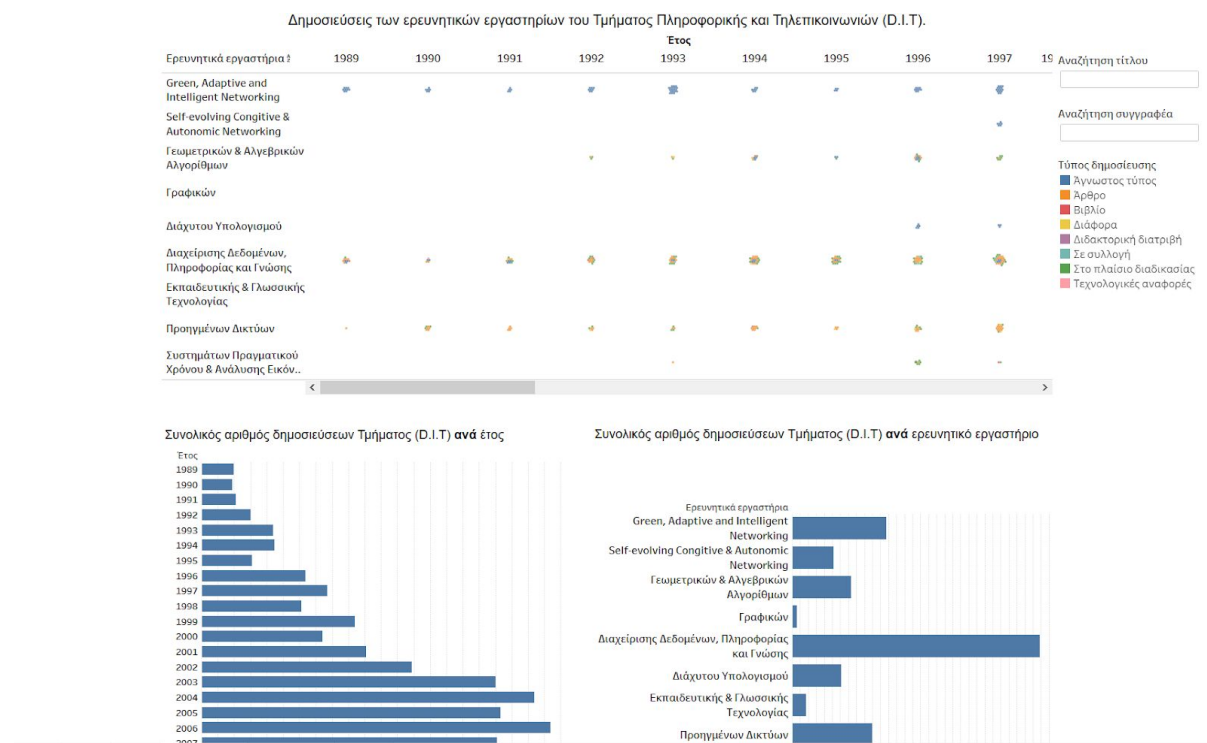
Σενάριο χρήσης 5: Αναζήτηση λέξης κλειδί, όπως για παράδειγμα “Database”.

Ένα ακόμα σημαντικό βήμα που έγινε για την προετοιμασία της δεύτερης διαμορφωτικής αξιολόγησης με χρήστες ήταν η δημιουργία ερωτηματολογίων. Στο Παράρτημα I παρατίθεται το ερωτηματολόγιο που χρησιμοποιήθηκε. Για την ταχύτερη και πιο αποδοτική συλλογή αποτελεσμάτων κατά τη διάρκεια της αξιολόγησης, δημιουργήθηκαν ερωτηματολόγια τα οποία σχεδιάστηκαν με τρόπο ώστε να συμπληρώνονται από τον ερευνητή που διεξάγει την αξιολόγηση. Συνεπώς τα ερωτηματολόγια συμπληρώθηκαν με βάση τις απαντήσεις και τις ενέργειες των συμμετεχόντων κατά τη διάρκεια εκτέλεσης των σεναρίων χρήσης.

Το τελευταίο βήμα της προετοιμασίας της δεύτερης διαμορφωτικής αξιολόγησης με χρήστες ήταν η επιλογή του χώρου και του τρόπου συνάντησης των συμμετεχόντων με τους ερευνητές. Συγκεκριμένα, ως τόπος επιλέχθηκε το εργαστήριο Linux του Τμήματος Πληροφορικής και Τηλεπικοινωνιών του Εθνικού Καποδιστριακού Πανεπιστημίου Αθηνών. Για την οργάνωση των συναντήσεων χρησιμοποιήθηκε η εφαρμογή Doodle, στην οποία οι συμμετέχοντες επέλεξαν την ώρα και τη μέρα που επιθυμούσαν προκειμένου να λάβουν μέρος στην προαναφερθείσα αξιολόγηση.

Η διαδικασία αξιολόγησης, στην οποία οι 15 συμμετέχοντες προσήλθαν την ίδια ημέρα, ξεκίνησε με μία εισαγωγή στο θέμα και δόθηκε το πλαίσιο μέσα στο οποίο παρουσιάστηκε η συγκεκριμένη διαδραστική οπτικοποίηση. Συγκεκριμένα, αναφέρθηκε ότι η οπτικοποίηση αυτή αναδεικνύει την ερευνητική δραστηριότητα του

Τμήματος Πληροφορικής και Τηλεπικοινωνιών του ΕΚΠΑ και θα βρίσκεται στην ιστοσελίδα του Τμήματος αυτού.



Εικόνα 15: Στιγμιότυπο οθόνης από την δεύτερη διαμορφωτική αξιολόγηση, περισσότερα στιγμιότυπα οθόνης παρατίθενται στο Παράρτημα Ι

Κατόπιν, ζητήθηκε από τους συμμετέχοντες καθ' όλη τη διάρκεια της αξιολόγησης να προσπαθήσουν να εξηγήσουν τί θεωρούν ότι απεικονίζεται στην παρούσα οπτικοποίηση και να εκφράζουν οποιαδήποτε απορία σχετικά με τον τρόπο που λειτουργεί αυτή. Με αυτό τον τρόπο συμπεραίνεται εάν η οπτικοποίηση είναι κατανοητή στους χρήστες χωρίς να χρειάζεται να τους δοθούν περαιτέρω εξηγήσεις. Στην παραπάνω εικόνα φαίνεται το πρωτότυπο υψηλής πιστότητας που χρησιμοποιήθηκε κατά την δεύτερη διαμορφωτική αξιολόγηση. Στο Παράρτημα Ι παρατίθενται επιπλέον εικόνες. Στη συνέχεια, ζητήθηκε από τους χρήστες να εκτελέσουν κάποια σενάρια χρήσης, τα οποία αναλύονται παραπάνω, και παρατηρήθηκε ο χρόνος που χρειάζονται, προκειμένου να εκτελέσουν κάθε ενέργεια καθώς και εάν σε κάποιο σημείο παρουσιάστηκε κάποια δυσκολία. Τέλος, με βάση τις απαντήσεις και τις ενέργειες των χρηστών συμπληρώθηκε ένα ερωτηματολόγιο από το οποίο προκύπτει η ευχρηστία της συγκεκριμένης διαδραστικής οπτικοποίησης, καθώς και τυχόν αδυναμίες που διαθέτει. Παρακάτω, αναλύονται τα αποτελέσματα των ερωτηματολογίων αυτών, καθώς και τα συμπεράσματα που προκύπτουν από αυτά.

Μετά το πέρας της αξιολόγησης μία από τις πρώτες ενέργειες που έγιναν ήταν η συλλογή και ανάλυση των ερωτηματολογίων, με τη βοήθεια των “φορμών Google” στις οποίες δημιουργήθηκαν τα ερωτηματολόγια και στη συνέχεια αναλύθηκαν χρησιμοποιώντας τη διεπαφή ανάλυσης απαντήσεων με διαγράμματα που παρέχει η Google. Το ερωτηματολόγιο που χρησιμοποιήθηκε παρατίθεται ολόκληρο στο (Παράρτημα Ι).

Τέλος, μία σημαντική ενέργεια που πραγματοποιήθηκε μετά την αξιολόγηση, ήταν η απομαγνητοφώνηση και ανάλυση των ηχογραφήσεων καθενός από τους συμμετέχοντες, προκειμένου να γίνουν αντιληπτά τα υπέρ και τα κατά του συγκεκριμένου σχεδιασμού και να διορθωθούν τυχόν λάθη και παραλείψεις που αναδείχθηκαν κατά τη διάρκεια της αξιολόγησης.

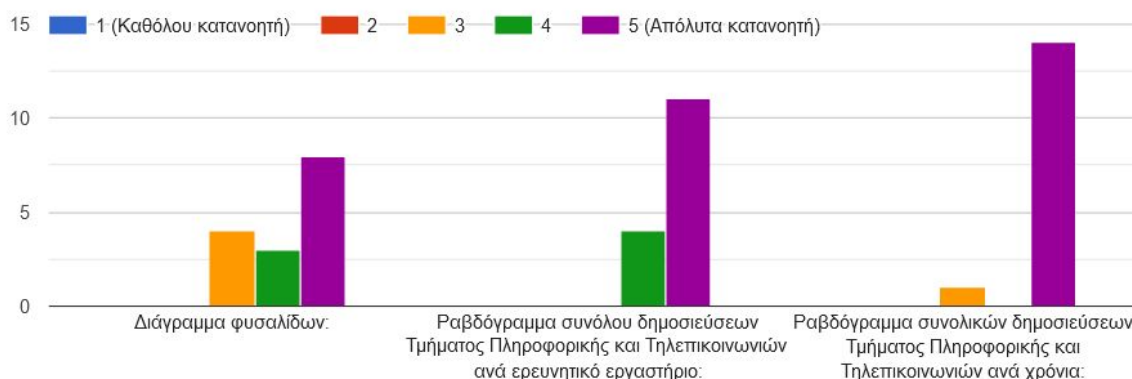
Κατά τη διάρκεια της συγκεκριμένης αξιολόγησης δεν υπήρξαν ιδιαίτερα προβλήματα, ωστόσο μία δυσκολία, η οποία αντιμετωπίστηκε, ήταν η καταγραφή, κατηγοριοποίηση και ιεράρχηση των προβλημάτων που αναφέρουν οι χρήστες κατά τη διάρκεια της αξιολόγησης. Κατά την ιεράρχηση των προβλημάτων έπρεπε να διαπιστωθούν τα προβλήματα που χρήζουν άμεσης επίλυσης, εκείνα που κρίνεται χρήσιμο να διορθωθούν και εκείνα που είναι ήσσονος σημασίας.

Παρακάτω παρατίθενται οι απαντήσεις των χρηστών στο σχετικό ερωτηματολόγιο, καθώς και οι παρατηρήσεις που διαπιστώθηκαν αναφορικά με τη δυσκολία των συμμετεχόντων κατά τη διάρκεια της αξιολόγησης.

5.7. Συμπεράσματα – Τελικός Σχεδιασμός

Αρχικά έπρεπε να διαπιστωθεί κατά πόσο οι συμμετέχοντες κατανόησαν τα διαγράμματα που παρουσιάστηκαν σε αυτούς. Στο παρακάτω διάγραμμα παρουσιάζεται με βαθμολογία από τον αριθμό ένα (1)-(καθόλου κατανοητό) έως τον αριθμό πέντε (5)-(απόλυτα κατανοητό) ο βαθμός κατανόησης του κάθε ραβδόγραμματος από τους συμμετέχοντες στην έρευνα.

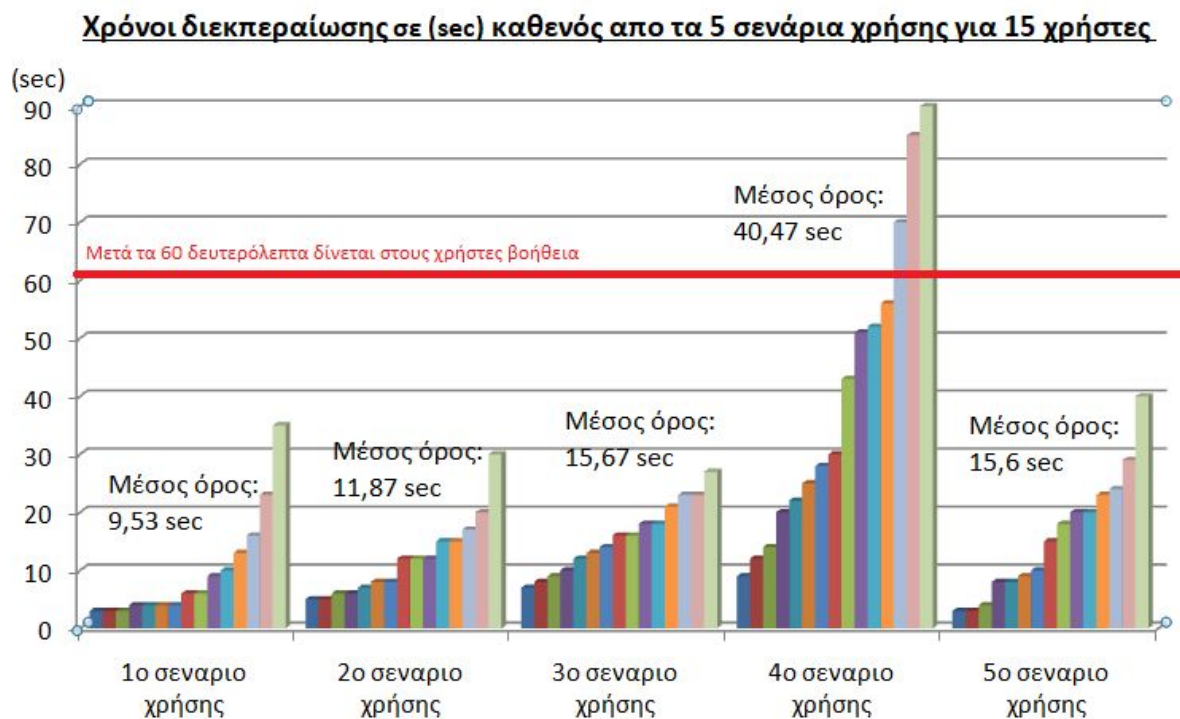
Πόσο κατανοητό είναι κάθε διάγραμμα ξεχωριστά; Βαθμολογήστε από 1-5.



Εικόνα 16: Απαντήσεις σχετικά με την κατανόηση των τριών διαγραμμάτων από τους χρήστες

Από το παραπάνω διάγραμμα συμπεραίνεται ότι το ραβδόγραμμα συνολικών δημοσιεύσεων του Τμήματος Πληροφορικής και Τηλεπικοινωνιών ανά έτος είναι το πιο κατανοητό από τους χρήστες, ενώ αντίθετα, το διάγραμμα για το οποίο υπάρχει μεγαλύτερη σύγχυση από την πλευρά των χρηστών είναι το διάγραμμα φυσαλίδων.

Εν συνεχεία, καταγράφηκαν οι χρόνοι διεκπεραίωσης καθενός από τα σενάρια χρήσης κατά τη διάρκεια της αξιολόγησης και παρατίθενται στο παρακάτω διάγραμμα.



Εικόνα 17: Διάγραμμα χρόνων ολοκλήρωσης σεναρίου χρήσης για 15 συμμετέχοντες στην δεύτερη αξιολόγηση

Από τις μετρήσεις του χρόνου διεκπεραίωσης των σεναρίων χρήσης από τους χρήστες, προκύπτει ότι οι χρήστες κατορθώνουν να ολοκληρώσουν τα σενάρια χρήσης χωρίς να αντιμετωπίσουν ιδιαίτερη δυσκολία. Παρόλα αυτά στο τέταρτο σενάριο χρήσης, τρεις (3) χρήστες δεν καταφέρνουν να ολοκληρώσουν το ζητηθέν μέσα στο χρόνο των εξήντα (60) δευτερολέπτων, ο οποίος σύμφωνα με τον σχεδιασμό κρίθηκε επαρκής, και συνεπώς τους δόθηκε μία υπόδειξη ώστε να συνεχίσουν. Ωστόσο, το τέταρτο σενάριο χρήσης είναι το πιο απαιτητικό και ορισμένα προβλήματα που προκύπτουν, οδηγούν στην αδυναμία διεκπεραίωσής του από τους τρεις (3) χρήστες. Τα εν λόγω προβλήματα εξετάζονται παρακάτω.

Μετά το πέρας της αξιολόγησης και σύμφωνα με τα αποτελέσματά της, γίνονται οι απαραίτητες διορθώσεις στο σχεδιασμό της οπτικοποίησης, ώστε να ανταποκρίνεται στις ανάγκες των χρηστών. Τα βασικότερα προβλήματα που παρουσιάζονται κατά τη διάρκεια της αξιολόγησης μαζί με τις αλλαγές που γίνονται για την επίλυσή τους, χωρίζονται σε τρεις βασικές κατηγορίες:

1) Προβλήματα κατανόησης ή δημιουργία σύγχυσης.

Το πρώτο πρόβλημα που προκύπτει και δημιουργεί σύγχυση στους συμμετέχοντες είναι ότι το διάγραμμα φυσαλίδων ξεκινάει από τα παλαιότερα έτη και οι συμμετέχοντες πρέπει να κάνουν «Scroll», ώστε να φτάσουν στα πιο πρόσφατα έτη. Συνεπώς κατά την υλοποίηση της οπτικοποίησης σε JavaScript το συγκεκριμένο πρόβλημα επιλύεται με το διάγραμμα να ξεκινάει από τα πιο πρόσφατα έτη.

Ένα ακόμη πρόβλημα που προκύπτει κατά τη διάρκεια της αξιολόγησης είναι η ασυνέπεια που υπάρχει κατά τη σύνδεση συγγραφέων και ερευνητικών εργαστηρίων όσον αφορά στις δημοσιεύσεις που έχουν εκδοθεί. Όπως αναφέρεται παραπάνω,

ορισμένοι καθηγητές-συγγραφείς ανήκουν σε περισσότερα από ένα ερευνητικά εργαστήρια, καθιστώντας δύσκολη τη σύνδεση των δημοσιεύσεων τους με ένα μόνο από αυτά τα εργαστήρια. Η λύση σε αυτό το πρόβλημα δίνεται, εμφανίζοντας τις ερευνητικές δημοσιεύσεις κάθε καθηγητή-συγγραφέα σε όλα τα ερευνητικά εργαστήρια στα οποία ανήκει.

Ακολουθώντας, το γεγονός ότι ο χρήστης αδυνατεί να «κλικάρει» πάνω σε κάθε ερευνητικό εργαστήριο και να μεταβεί στην ιστοσελίδα του επιλεγόμενου εργαστηρίου αποτελεί πρόβλημα, καθώς ο χρήστης δεν έχει τη δυνατότητα να κατανοήσει ποιο είναι το εργαστήριο στο οποίο αναφέρεται η οπτικοποίηση. Για το λόγο αυτό εισάγεται ένα link για κάθε ερευνητικό εργαστήριο, το οποίο ανακατευθύνει τον χρήστη στην εκάστοτε ιστοσελίδα.

Προβλήματα ήσσονος σημασίας, τα οποία ωστόσο πρέπει να αναφερθούν καθώς εντοπίζονται από τρεις (3) συμμετέχοντες εκ των δεκαπέντε (15) συμμετεχόντων, είναι ότι επιθυμούν τα φίλτρα να είναι στα αριστερά και τα ραβδογράμματα να ορίζονται σε Stacked Bar Chart διαχωρίζοντας κάθε είδος δημοσίευσης, με τον ίδιο τρόπο όπως χωρίζονται στο διάγραμμα φουσαλίδων.

Το γεγονός δε, ότι εμφανίζεται μια στήλη με όλες τις δημοσιεύσεις, οι οποίες δεν έχουν ορισμένες ημερομηνίες δημοσίευσης μπερδεύουν τους συμμετέχοντες.

Τέλος, η μίξη Ελληνικών και Αγγλικών στην οπτικοποίηση είναι ακόμη ένα πρόβλημα, το οποίο προκαλεί σύγχυση στους χρήστες. Το πρόβλημα αυτό δημιουργήθηκε κατά τη συλλογή των δεδομένων. Ορισμένα από τα δεδομένα συλλέχθηκαν στην ελληνική γλώσσα, ενώ άλλα δεδομένα συλλέχθηκαν στην αγγλική γλώσσα, καθώς δεν ήταν διαθέσιμη η ελληνική εκδοχή τους. Το παρόν πρόβλημα λύθηκε μεταφράζοντας τους Ελληνικούς όρους σε Αγγλικούς και τους Αγγλικούς σε Ελληνικούς, διατηρώντας αυστηρά μόνο την ελληνική γλώσσα στην τελική οπτικοποίηση.

2) Προβλήματα αναφορικά με τα φίλτρα και τον τρόπο φιλτραρίσματος - αναζήτησης των δεδομένων.

Αναφορικά με τον τρόπο αναζήτησης των δεδομένων, ένα πρόβλημα το οποίο προέκυψε κατά τη διάρκεια της αξιολόγησης ήταν η δυσκολία των συμμετεχόντων να διακρίνουν τις φουσαλίδες στο διάγραμμα φουσαλίδων. Το συγκεκριμένο πρόβλημα επιλύεται με την εισαγωγή της λειτουργίας «zoom», δηλαδή το πρόγραμμα να μπορεί να κάνει «zoom in» και «zoom out» στα δεδομένα, ώστε οι φουσαλίδες να φαίνονται μεγαλύτερες σε περίπτωση που αυτό απαιτείται.

Ένα πρόβλημα το οποίο παρατηρήθηκε ήταν ότι οι χρήστες αδυνατούσαν να βρουν το συνολικό αριθμό των δημοσιεύσεων για μια συγκεκριμένη αναζήτηση ικανοποίησης περιορισμών με τη χρήση φίλτρων. Το συγκεκριμένο πρόβλημα παρατηρήθηκε μόνο στο σενάριο χρήσης 4 (τέσσερα), καθώς μόνον αυτό απαιτούσε ακριβή αριθμό δημοσιεύσεων που τηρούν κάποια συγκεκριμένα κριτήρια. Για την επίλυση αυτού του προβλήματος εισήχθη ένας αριθμός κάτω από τα φίλτρα, ο οποίος παρουσιάζει πόσες εγγραφές ικανοποιούν τους περιορισμούς των φίλτρων που έχουν οριστεί από τον χρήστη.

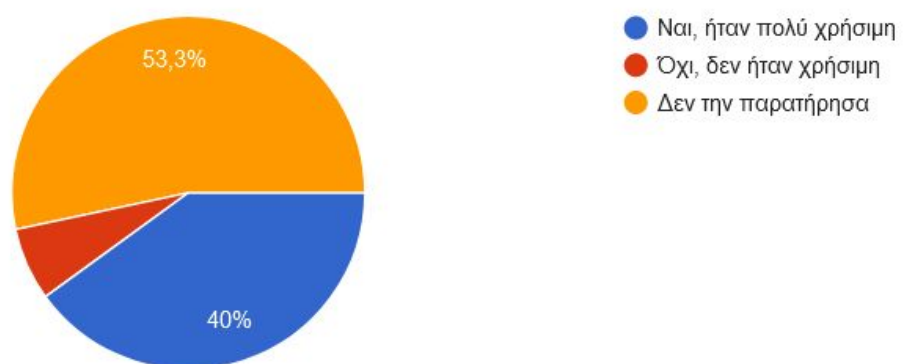
Κατά την αξιολόγηση παρατηρήθηκε ότι κάποιοι από τους συμμετέχοντες μεταβαίνοντας από τα ένα σενάριο χρήσης στο επόμενο δεν ακύρωναν τα ήδη επιλεγμένα φίλτρα. Αυτό οφείλεται στην μη ορατότητα της κατάστασης του

συστήματος (Βλ. κεφάλαιο 2: ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ) δηλαδή στο ότι δεν ήταν ευδιάκριτα τα φίλτρα τα οποία είχαν οριστεί. Το συγκεκριμένο πρόβλημα επιλύθηκε κάνοντας πιο ευδιάκριτα τα φίλτρα τα οποία είχαν οριστεί και εισάγοντας ένα κουμπί Reset (κουμπί ακύρωσης όλων των φίλτρων), ώστε να ακυρώνονται όλα τα επιλεγμένα φίλτρα.

Το κυριότερο πρόβλημα της συγκεκριμένης κατηγορίας είναι ότι το μεγαλύτερο ποσοστό των χρηστών της οπτικοποίησης δεν παρατηρούν τη λειτουργία φιλτραρίσματος του διαγράμματος φουσαλίδων, χρησιμοποιώντας τα δύο ραβδογράμματα. Κατά τη λειτουργία αυτή, ο χρήστης μπορεί να κλικάρει σε μία ράβδο ανά ραβδόγραμμα («Ραβδόγραμμα Συνόλου Δημοσιεύσεων Τμήματος Πληροφορικής και Τηλεπικοινωνιών ανά Ερευνητικό Εργαστήριο» και «Ραβδόγραμμα Συνολικών Δημοσιεύσεων Τμήματος Πληροφορικής και Τηλεπικοινωνιών ανά χρόνια») και να φιλτράρει τα δεδομένα που εμφανίζονται στο διάγραμμα φουσαλίδων. Παραδείγματος χάριν, επιλέγοντας τη ράβδο που αντιπροσωπεύει το έτος 2006, στο Ραβδόγραμμα Συνολικών Δημοσιεύσεων Τμήματος Πληροφορικής και Τηλεπικοινωνιών ανά έτος, φιλτράρονται τα αποτελέσματα στο διάγραμμα φουσαλίδων, ούτως ώστε να εμφανίζει μόνο δημοσιεύσεις για το έτος 2006. Αντίστοιχα μπορεί να επιλεχθεί μία ράβδος, δηλαδή ένα εργαστήριο, από το Ραβδόγραμμα Συνόλου Δημοσιεύσεων Τμήματος Πληροφορικής και Τηλεπικοινωνιών ανά Ερευνητικό Εργαστήριο και να φιλτράρει το διάγραμμα φουσαλίδων, ούτως ώστε να εμφανίζει δημοσιεύσεις μόνο για το επιλεχθέν συγκεκριμένο ερευνητικό εργαστήριο. Συνεπώς, δεδομένων των αποτελεσμάτων της αξιολόγησης, αποφασίζεται να παραλειφθεί η παραπάνω δυνατότητα από την οπτικοποίηση και το φιλτράρισμα του διαγράμματος φουσαλίδων να γίνεται αποκλειστικά από τα φίλτρα, τα οποία το πλαισιώνουν.

Χρησιμοποίησες τη λειτουργία στην οποία κλικάρεις τα δύο ραβδογράμματα για να επηρεάσεις το περιεχόμενο του κυρίως διαγράμματος;

15 απαντήσεις



Εικόνα 18: Διάγραμμα πίτας. Απαντήσεις στην ερώτηση σχετικά με την χρήση των ραβδογραμμάτων ως φίλτρα.

3) Βελτιώσεις ή προτάσεις από την πλευρά των χρηστών.

Παρακάτω παρουσιάζονται επιγραμματικά ορισμένες προτάσεις από την πλευρά των χρηστών επί των δυσκολιών που αντιμετώπισαν κατά τη διάρκεια της αξιολόγησης, οι οποίες αν και δεν αποτέλεσαν ουσιαστικά προβλήματα, λαμβάνονται υπόψη, διότι ενδέχεται να βελτιώσουν την εμπειρία χρήσης της συγκεκριμένης οπτικοποίησης.

- Να γίνεται μετάβαση στην ιστοσελίδα των ερευνητικών εργαστηρίων κάνοντας «Click» επάνω σε αυτά.
- Να γίνει προσθήκη ενός «Dropdown», προκειμένου να γίνεται επιλογή του συγγραφέα κατά την αναζήτηση.
- Να αναγράφονται οι συνολικές δημοσιεύσεις κάθε σμήνους φυσαλίδων κατά το φιλτράρισμα, ώστε να εμφανίζονται οι συνολικές δημοσιεύσεις κάθε ενός καθηγητή ή ερευνητικού εργαστηρίου ανά έτος.
- Να προστεθεί ένας πίνακας στο κάτω μέρος της οπτικοποίησης, με όλα τα δεδομένα, τα οποία υπάρχουν στη βάση, ώστε να είναι δυνατή η προσπέλαση ή αντιγραφή τους από τους χρήστες. Ακόμη να υπάρχει ένας γρήγορος τρόπος να αντιγράφουν όλα αυτά τα δεδομένα με τη χρήση ενός ή πολλαπλών κουμπιών «copy».

6. ΥΛΟΠΟΙΗΣΗ

6.1. Τα Εργαλεία που Χρησιμοποιήθηκαν

Η τελική οπτικοποίηση είναι γραμμένη σε JavaScript και συγκεκριμένα σε μία βιβλιοθήκη της JavaScript, την D3 και εκτελείται σε έναν Xampp Apache server.

- D3.js (Javascript,CSS,HTML5,SVG):

Το D3.js (επίσης γνωστό ως D3) είναι μια βιβλιοθήκη JavaScript για την παραγωγή δυναμικών, διαδραστικών οπτικοποιήσεων δεδομένων σε προγράμματα περιήγησης ιστού (web browsers). Χρησιμοποιεί πρότυπα Scalable Vector Graphics (SVG), HTML5 και Cascading Style Sheets (CSS). Είναι ο διάδοχος του παλαιότερου πλαισίου Protovis. Η ανάπτυξή του σημειώθηκε το 2011, καθώς η έκδοση 2.0.0 κυκλοφόρησε τον Αύγουστο του 2011.

- Xampp Apache server:

Το XAMPP είναι ένα δωρεάν πακέτο λύσεων διακομιστή ιστού ανοικτού κώδικα (open-source cross-platform web server solution stack) που αναπτύχθηκε από την Apache Friends και που αποτελείται κυρίως από τον διακομιστή Apache HTTP, τη βάση δεδομένων MariaDB και διερμηνείς (interpreters) για σενάρια (scripts) γραμμένα στις γλώσσες προγραμματισμού PHP και Perl. Δεδομένου ότι οι περισσότερες πραγματικές εφαρμογές διακομιστή ιστού χρησιμοποιούν τα ίδια στοιχεία με το XAMPP, καθιστά δυνατή τη μετάβαση από έναν τοπικό δοκιμαστικό διακομιστή σε έναν ζωντανό διακομιστή. Η ευκολία ανάπτυξης του XAMPP σημαίνει ότι μια στοίβα WAMP ή LAMP μπορεί να εγκατασταθεί γρήγορα και απλά σε ένα λειτουργικό σύστημα από έναν προγραμματιστή, με το πλεονέκτημα ότι μπορούν να προστεθούν γνωστές εφαρμογές σαν πρόσθετα, όπως το WordPress και το Joomla.

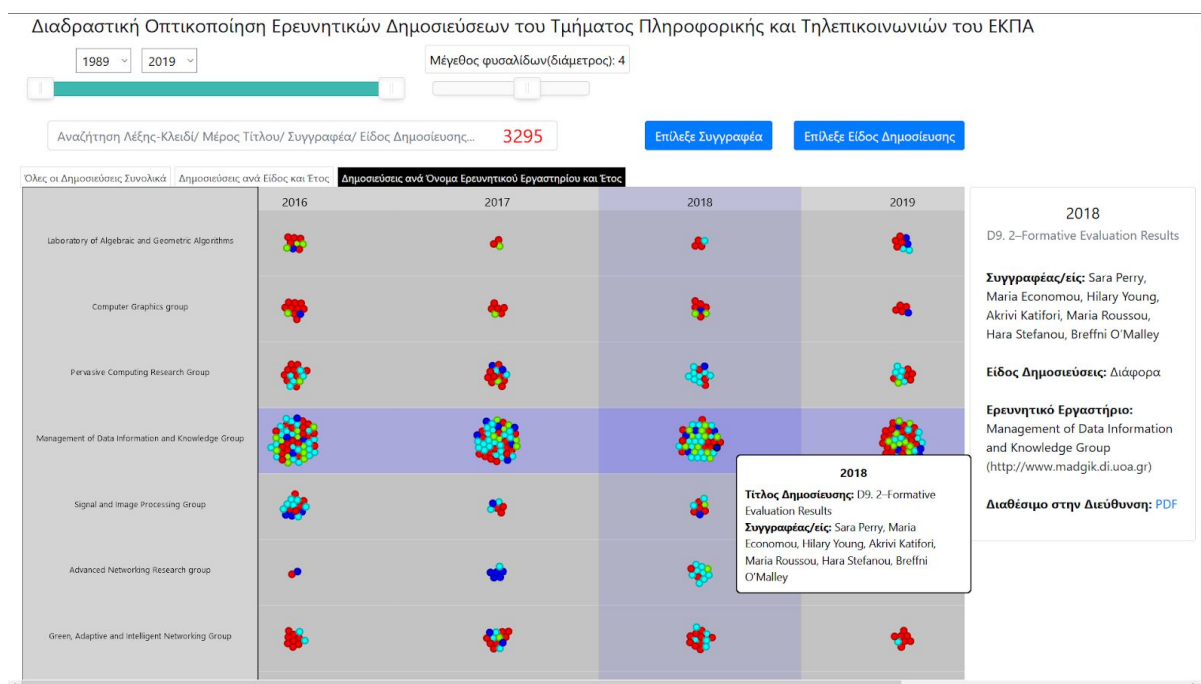
Πριν από την εισαγωγή του αρχείου δεδομένων στην τελική οπτικοποίηση έπρεπε το αρχείο να «καθαριστεί» και να τροποποιηθεί κατάλληλα, ώστε να μπορεί να «διαβαστεί» από το πρόγραμμα JavaScript με το οποίο ενεγράφη η τελική οπτικοποίηση. Για την επεξεργασία και τον «καθαρισμό» των δεδομένων, τα οποία συλλέχθηκαν από τον Crawler, χρησιμοποιήθηκαν ορισμένες από τις λειτουργίες του Tableau, οι δυνατότητες που παρέχει το Excel, καθώς και ένα πρόγραμμα γραμμένο σε Python.

Λόγω αναβάθμισης της ιστοσελίδας Google Scholar κατά τη διάρκεια εκπόνησης της εργασίας, χρησιμοποιούνται διαφορετικής δομής δεδομένα στο σχεδιασμό της οπτικοποίησης και διαφορετικά δεδομένα στην υλοποίησή της. Τα βασικά πεδία των δεδομένων δεν αλλάζουν, αλλάζει όμως ο τρόπος συλλογής τους και ο τύπος των αρχείων στα οποία αποθηκεύονται. Συνεπώς, δεν δημιουργείται πρόβλημα στις σχεδιαστικές αποφάσεις που έχουν γίνει, απλώς η υλοποίηση προσαρμόζεται εκ νέου. Τα παραπάνω προγράμματα βοήθησαν στην ενοποίηση όλων των δεδομένων, προκειμένου να μπορεί να δουλεύει η τελική οπτικοποίηση.

6.2. Η Δημιουργία της Τελικής Οπτικοποίησης

Η διαδικασία υλοποίησης της τελικής διαδραστικής οπτικοποίησης ξεκινάει αρχικά αναζητώντας τις τεχνολογίες και τα εργαλεία που είναι απαραίτητα για την επίτευξη παρόμοιου αποτελέσματος με αυτό που έχει προκύψει από το σχεδιασμό. Η

δέσμευση που υπάρχει είναι ότι η οπτικοποίηση πρέπει να υλοποιηθεί σε web περιβάλλον, ώστε να είναι Portable και να μπορεί εύκολα να παρουσιαστεί σε πολλούς χρήστες ταυτόχρονα και να υποβληθεί σε μία τελική συμπερασματική αξιολόγηση. Η εύρεση μιας βιβλιοθήκης που να διατηρεί όλα τα παραπάνω και να υλοποιεί διάγραμμα φυσαλίδων για κατηγορικά δεδομένα ήταν αρκετά δύσκολη. Το ζητούμενο δηλαδή είναι να υλοποιηθεί ένα διάγραμμα φυσαλίδων, το οποίο αποτελείται από σμήνη φυσαλίδων χωρίς η μία φυσαλίδα να επικαλύπτει την άλλη και χωρίς να υπάρχουν ενδιάμεσες τιμές, με άλλα λόγια ένα διάγραμμα που να είναι πολύ κοντά με αυτό που αναλύεται κατά το σχεδιασμό της συγκεκριμένης οπτικοποίησης χρησιμοποιώντας το Tableau. Τη λύση σε όλα τα παραπάνω έδωσε η βιβλιοθήκη της JavaScript D3.js, η οποία περιγράφεται παραπάνω.



Εικόνα 19: Στιγμιότυπο οθόνης από την τελική οπτικοποίηση, πηγή: στο Git Pages: <https://xenospan.github.io/Interactive-Visualization-of-DIT-Research/>

Η διαδραστική οπτικοποίηση που υλοποιείται στη συγκεκριμένη πτυχιακή εργασία αποτελείται από ένα διάγραμμα φυσαλίδων. Οι φυσαλίδες συγκεντρώνονται γύρω από κάποια κέντρα, τα οποία μεταβάλλονται με τη χρήση φίλτρων. Στο επάνω μέρος της οπτικοποίησης υπάρχουν τρεις καρτέλες και ανάλογα με την επιλεγμένη καρτέλα εμφανίζεται, είτε ένα διάγραμμα φυσαλίδων με τις δημοσιεύσεις ανά όνομα ερευνητικού εργαστηρίου και έτος, είτε ένα διάγραμμα φυσαλίδων με τις δημοσιεύσεις ανά είδος και έτος, ή ένα διάγραμμα φυσαλίδων με όλες τις δημοσιεύσεις συνολικά. Στον οριζόντιο άξονα, ο οποίος αποτυπώνεται στο επάνω μέρος του διαγράμματος αναγράφονται τα έτη δημοσίευσης μιας έρευνας. Στον κατακόρυφο άξονα αναγράφονται τα ερευνητικά εργαστήρια στα οποία ανήκει η κάθε ερευνητική δημοσίευση ή με την επιλογή της αντίστοιχης καρτέλας το είδος (type) των ερευνητικών δημοσιεύσεων. Η χρωματική απεικόνιση αντιπροσωπεύει το είδος (type) της κάθε ερευνητικής δημοσίευσης (το υπόμνημα βρίσκεται στο κάτω μέρος του διαγράμματος). Με την επιλογή «Όλες οι δημοσιεύσεις συνολικά» από την

καρτέλα επιλογών στο πάνω μέρος, οι φυσαλίδες συγκεντρώνονται σε ένα ενιαίο σμήνος. Η συγκεκριμένη επιλογή βοηθάει στην αναζήτηση μιας δημοσίευσης όταν η ημερομηνία δημοσίευσης είναι άγνωστη.

Κάνοντας «hover» επάνω σε μία φυσαλίδα εμφανίζονται σε ένα «Tooltip» το έτος, ο τίτλος, καθώς και οι συγγραφείς μιας ερευνητικής δημοσίευσης. Επιπλέον, κάνοντας «κλικ» επάνω σε μία φυσαλίδα στη δεξιά πλευρά της οπτικοποίησης εμφανίζεται μία καρτέλα με πιο αναλυτικά δεδομένα, όπως το είδος δημοσίευσης, το ερευνητικό εργαστήριο από το οποίο εκδόθηκε η δημοσίευση και σε περίπτωση που είναι διαθέσιμο παρατίθεται και σχετικό PDF.

Η οπτικοποίηση διαθέτει φίλτρα, τα οποία μεταβάλλουν τα δεδομένα που παρουσιάζονται. Μετά από κάθε αναζήτηση εμφανίζεται με κόκκινα νούμερα στο επάνω μέρος της οπτικοποίησης ο αριθμός των δημοσιεύσεων που ικανοποιούν τους περιορισμούς των φίλτρων. Υπάρχουν τριών ειδών φίλτρα, τα «Drop Down» φίλτρα από τα οποία ο χρήστης μπορεί να επιλέξει μία ανάμεσα σε πολλές διαθέσιμες επιλογές, τα «Slide Bar» φίλτρα στο οποίο ο χρήστης σέρνοντας μία μπάρα μπορεί να ρυθμίσει διάφορες τιμές και τέλος, το «Search Bar» φίλτρο, το οποίο είναι μία μπάρα αναζήτησης στην οποία ο χρήστης μπορεί να αναζητά λέξεις-κλειδιά.

Ο κώδικας της ανωτέρω διαδραστικής οπτικοποίησης παρατίθεται σε ένα Git Repository¹², ο οποίος δημιουργήθηκε για τους σκοπούς της εργασίας.

6.3. Συμπερασματική Αξιολόγηση - Αποτελέσματα

Μετά το πέρας της υλοποίησης της διαδραστικής οπτικοποίησης και δεδομένου του γεγονότος ότι αρκετές σχεδιαστικές επιλογές μεταβλήθηκαν κατά τη διάρκεια της υλοποίησης, κυρίως για τεχνικούς λόγους, έπρεπε να διαπιστωθεί ο βαθμός ευχρηστίας της τελικής οπτικοποίησης. Για το σκοπό αυτό πραγματοποιήθηκε μία συμπερασματική αξιολόγηση με χρήστες. Στη συγκεκριμένη αξιολόγηση έλαβαν μέρος δέκα (10) άτομα, τα οποία συμμετείχαν στην αξιολόγηση εξ' αποστάσεως¹³ μέσω της εφαρμογής Skype. Στη συγκεκριμένη αξιολόγηση, αξιολογήθηκε αυτόνομα μόνο η διαδραστική οπτικοποίηση του διαγράμματος φυσαλίδων. Η παρουσίαση της διαδραστικής οπτικοποίησης στους συμμετέχοντες πραγματοποιήθηκε με τη βοήθεια του Github Pages¹⁴, μία διαδικτυακή πλατφόρμα, η οποία παρέχει τη δυνατότητα στους χρήστες να «ανεβάζουν» μία web εφαρμογή, την οποία έχουν κατασκευάσει οι ίδιοι και αυτή να εκτελείται και να είναι προσβάσιμη από άλλους χρήστες, απλώς πατώντας ένα link.

Η διαδικασία που ακολουθήθηκε σε αυτή την συμπερασματική αξιολόγηση ήταν παρόμοια με τις προηγούμενες αξιολογήσεις. Αρχικά, οι συμμετέχοντες κλήθηκαν να εξηγήσουν τι βλέπουν στη συγκεκριμένη οπτικοποίηση. Με αυτό τον τρόπο διαπιστώθηκε κατά πόσο ξεκάθαρος είναι ο στόχος της συγκεκριμένης οπτικοποίησης, καθώς επίσης και τα επιμέρους στοιχεία της. Στη συνέχεια, ζητήθηκε από τους συμμετέχοντες να πραγματοποιήσουν ορισμένα σενάρια χρήσης. Τα

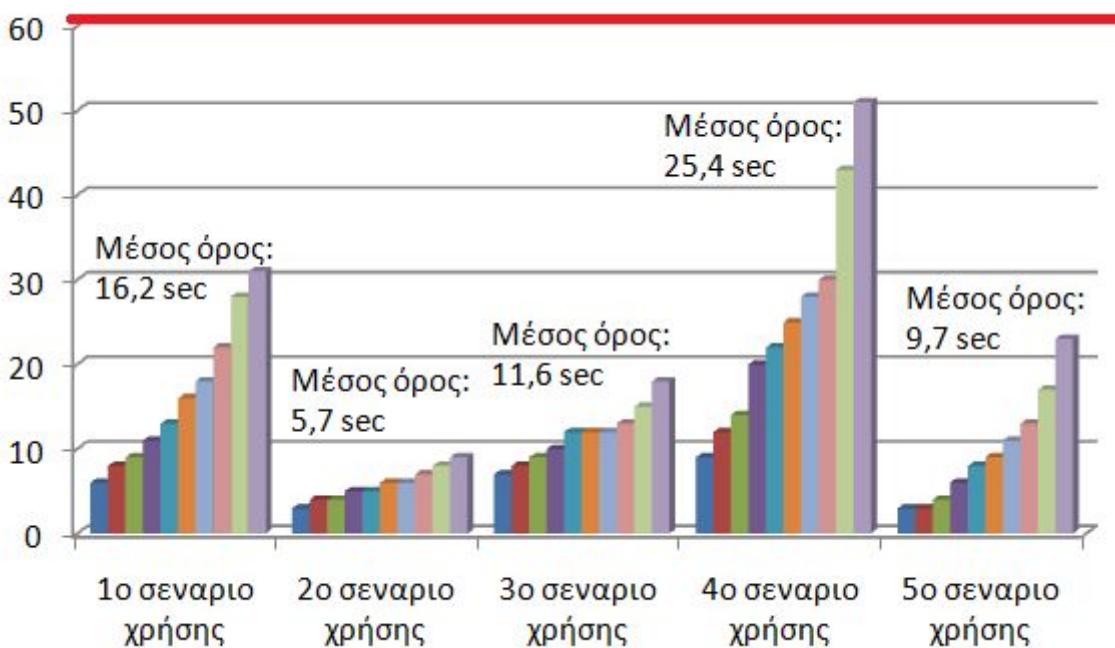
¹² <https://github.com/xenospan/Interactive-Visualization-of-DIT-Research>

¹³ Εξαιτίας της πανδημίας του κορονοϊού (COVID-19) δεν ήταν δυνατή η κατ' ιδίαν συνάντηση με τους συμμετέχοντες της αξιολόγησης.

¹⁴ <https://pages.github.com/>

σενάρια χρήσης, τα οποία πραγματοποιήσαν οι συμμετέχοντες, ήταν τα ίδια με τα σενάρια χρήσης των προηγούμενων αξιολογήσεων, ώστε να μπορεί να είναι δυνατή η σύγκριση. Τέλος, ζητήθηκε από τους χρήστες να εκφράσουν τις δυσκολίες που συνάντησαν (εάν αυτές υπήρχαν) κατά την αλληλεπίδραση τους με τη συγκεκριμένη διαδραστική οπτικοποίηση και να επισημάνουν τυχόν παρατηρήσεις που έχουν. Με αυτό το τρόπο συλλέχθηκαν αρκετά ποιοτικά δεδομένα, τα οποία αναλύονται στη συνέχεια.

Από τη συγκεκριμένη συμπερασματική αξιολόγηση φαίνεται ότι οι ανάγκες των χρηστών ικανοποιούνται κατά μεγάλο ποσοστό και δεν διαπιστώνονται σημαντικά προβλήματα κατά την εκτέλεση των σεναρίων χρήσης από τους χρήστες. Οι συμμετέχοντες ολοκλήρωσαν εξίσου γρήγορα και τα σενάρια χρήσης, τα οποία στοχεύουν στην ευρύτερη κατανόηση της οπτικοποίησης από πλευράς τους, καθώς επίσης και τα σενάρια χρήσης, τα οποία στοχεύουν στη αποδοτική χρήση φίλτρων και στην κατανόηση του τρόπου αλληλεπίδρασης με την οπτικοποίηση. Παρακάτω απεικονίζεται το αντίστοιχο διάγραμμα με τους χρόνους εκτέλεσης σε κάθε σενάριο χρήσης από δέκα συμμετέχοντες. Επάνω από κάθε ομάδα χρόνων (sec) αναγράφεται ένας αριθμός, ο οποίος αντιστοιχεί στο μέσο όρο το χρόνων (sec) για το συγκεκριμένο σενάριο χρήσης.



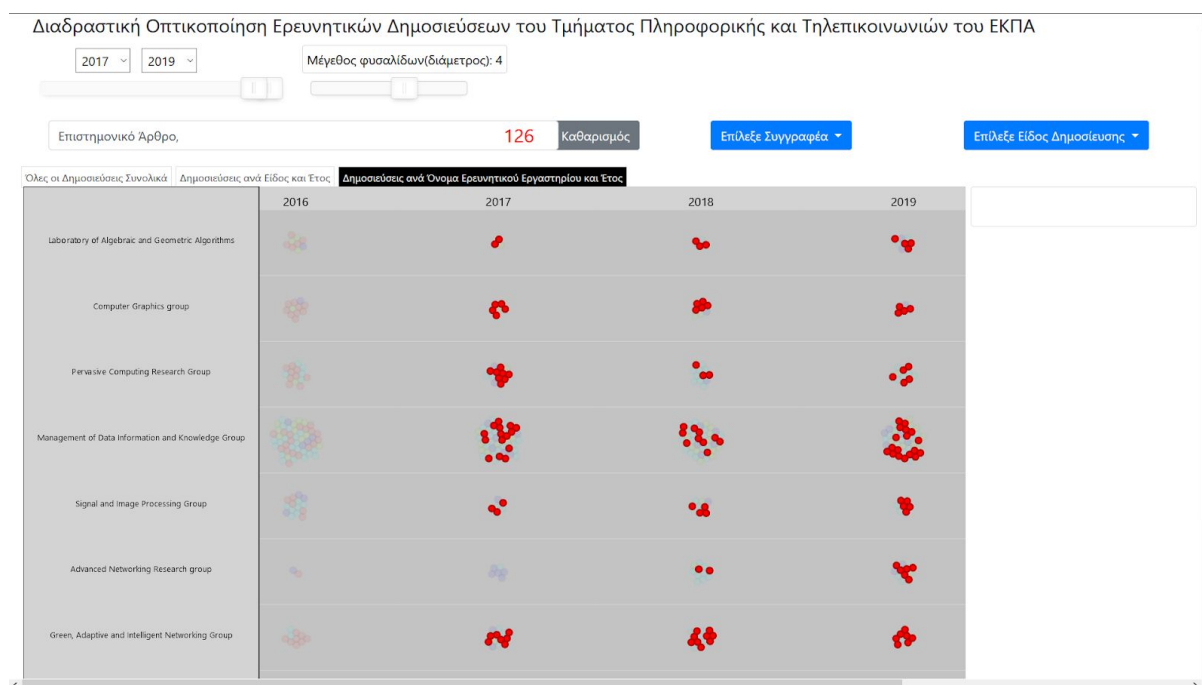
Εικόνα 20: Διάγραμμα χρόνων ολοκλήρωσης σεναρίου χρήσης για 10 συμμετέχοντες στην τρίτη αξιολόγηση

Από την παραπάνω εικόνα παρατηρούμε ότι οι μέσοι χρόνοι περάτωσης των σεναρίων χρήσης από τους συμμετέχοντες της συμπερασματικής αξιολόγησης έχουν βελτιωθεί αισθητά. Εξαίρεση αποτελεί το πρώτο σενάριο χρήσης στο οποίο οι χρήστες έπρεπε να εντοπίσουν τον αριθμό όλων των δημοσιεύσεων του Τμήματος Πληροφορικής και Τηλεπικοινωνιών του ΕΚΠΑ ένα συγκεκριμένο έτος. Η ζητούμενη πληροφορία εντοπίστηκε από τους χρήστες με τη χρήση των ραβδογραμμάτων που

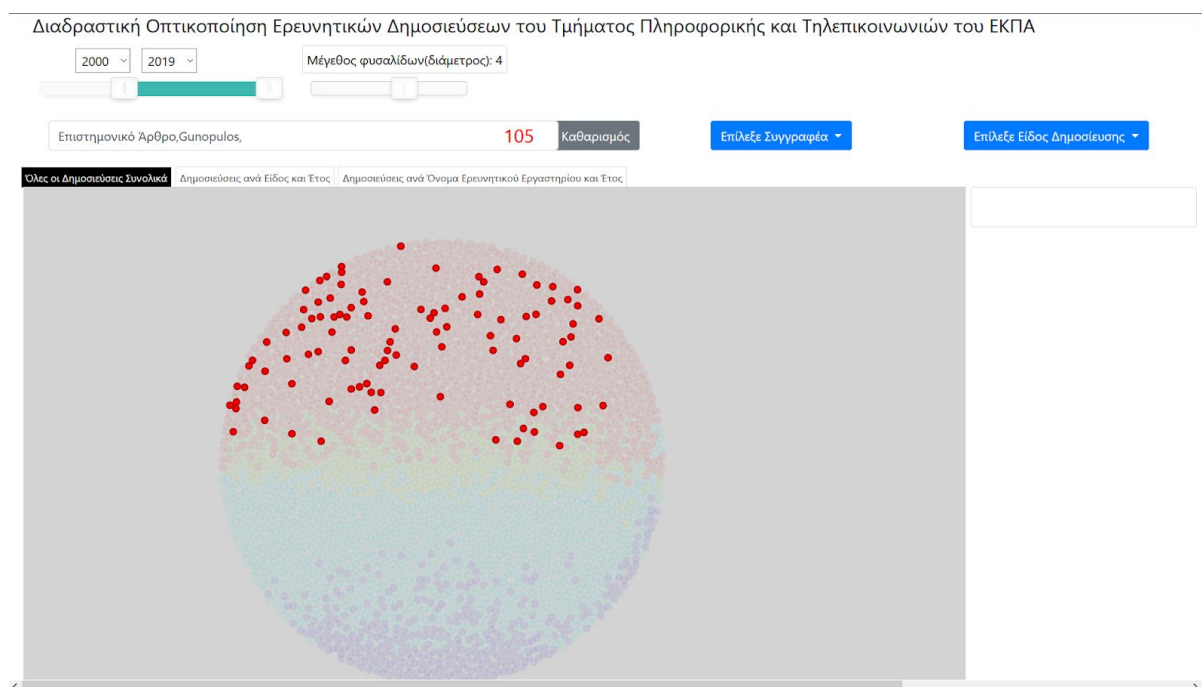
υπήρξαν στη δεύτερη διαμορφωτική αξιολόγηση και επομένως αυτό το σενάριο χρήσης υλοποιήθηκε σε μικρότερο χρόνο στη συγκεκριμένη περίπτωση.

Παρακάτω παρατίθενται επιγραμματικά οι παρατηρήσεις των χρηστών στη συγκεκριμένη συμπερασματική αξιολόγηση:

- Το γεγονός ότι η κάρτα στα δεξιά ήταν κενή πριν επιλεγεί κάποια φυσαλίδα μπερδεύει τους χρήστες, καθώς είχαν την εντύπωση ότι επρόκειτο για σφάλμα της εφαρμογής. Για την αντιμετώπιση του προβλήματος προστέθηκε μέσα στην κάρτα ένα κείμενο, το οποίο ενθαρρύνει το χρήστη να πατήσει επάνω σε μία σελίδα για να δει περισσότερα δεδομένα για τη συγκεκριμένη δημοσίευση.
- Παράλληλα, η προσθήκη αριθμού μέσα στην μπάρα αναζήτησης που δείχνει πόσες δημοσιεύσεις ικανοποιούν τους περιορισμούς των φίλτρων, διόρθωσε ένα από τα προβλήματα που παρατηρήθηκαν όταν κάποιοι χρήστες δεν αντιλήφθηκαν την αλλαγή της οπτικοποίησης μετά την χρήση ενός φίλτρου.
- Ένα πρόβλημα, το οποίο παρατηρήθηκε κατά τη διάρκεια της συμπερασματικής αξιολόγησης ήταν ότι οι χρήστες προσπαθώντας να πλοηγηθούν στη σελίδα προς τα κάτω, εκ παραδρομής πραγματοποιούσαν «zoom out», καθώς το ζουμ εντός του διαγράμματος της οπτικοποίησης γίνεται με τη χρήση της ροδέλας. Το συγκεκριμένο πρόβλημα λύθηκε περιορίζοντας τα σημεία που έχει τη δυνατότητα ο χρήστης να κάνει «zoom», καθώς επίσης και ο βαθμός στον οποίο μπορεί να κάνει «zoom in» ή «zoom out».
- Ένα πρόβλημα, το οποίο παρατηρήθηκε από ορισμένους χρήστες στην τελευταία αξιολόγηση ήταν ότι η οπτικοποίηση παρουσίαζε μια μικρή καθυστέρηση στην κίνηση των φυσαλίδων (Bubble Animation). Το συγκεκριμένο πρόβλημα δεν διορθώθηκε στα πλαίσια της συγκεκριμένης πτυχιακής εργασίας, εντούτοις θα ήταν ενδιαφέρον να ανευρεθούν οι λόγοι για τους οποίους συμβαίνει αυτό και να επιλυθούν σε μία μελλοντική επέκταση.
- Η έλλειψη κουμπιού καθαρισμού της μπάρας αναζήτησης, ήταν ένα απλό πρόβλημα που λύθηκε με αρκετή ευκολία προσθέτοντας το συγκεκριμένο κουμπί.
- Η απουσία κάποιας ένδειξης για το ποιά φίλτρα έχουν οριστεί προκάλεσε σύγχυση στους συμμετέχοντες. Για την επίλυση του προβλήματος αυτού τα Dropdown φίλτρα μετατράπηκαν κατάλληλα ώστε χρησιμοποιώντας τα να αναγράφεται η επιλεγμένη τιμή του φίλτρου στην μπάρα αναζήτησης.
- Τέλος, ένα από τα βασικότερα προβλήματα ήταν το γεγονός ότι δεν εμφανιζόταν τότε έχουν εφαρμοστεί φίλτρα και τότε όχι, καθώς οι φυσαλίδες που δεν ικανοποιούσαν τους περιορισμούς των φίλτρων εξαφανίζονταν τελείως προκαλώντας σύγχυση. Για την επίλυση του προβλήματος, η οπτικοποίηση τροποποιήθηκε κατάλληλα, ούτως ώστε οι φυσαλίδες οι οποίες δεν ικανοποιούσαν τους περιορισμούς να μην εξαφανίζονταν τελείως, αντιθέτως να «θόλωναν» (εντολή CSS «opacity: 0.07») και η λειτουργία «tooltip» να μην ήταν πλέον διαθέσιμη για αυτά.



Εικόνα 21: Στιγμιότυπο οθόνης από την χρήση φίλτρων στην τελική υλοποίηση της οπτικοποίησης - χρήση της καρτέλας «Δημοσιεύσεις ανά όνομα ερευνητικού εργαστηρίου και έτος»



Εικόνα 22: Στιγμιότυπο οθόνης από την χρήση φίλτρων στην τελική υλοποίηση της οπτικοποίησης - χρήση της καρτέλας «Όλες οι δημοσιεύσεις συνολικά», περισσότερα στιγμιότυπα οθόνης παρατίθενται στο Παράρτημα Ι

7. ΣΥΜΠΕΡΑΣΜΑΤΑ

7.1. Ανασκόπηση Όλης της Διαδικασίας

Στα κεφάλαια που προηγήθηκαν πραγματοποιήθηκε μία ιστορική αναδρομή στο χώρο των οπτικοποιήσεων και εξετάστηκαν οι τρόποι με τους οποίους διάφορες εφαρμογές διαδραστικής οπτικοποίησης χειρίζονται και οπτικοποιούν τα δεδομένα. Στη συνέχεια, αναλύθηκαν η δομή και ο τύπος των δεδομένων, τα οποία οπτικοποιούνται στην εργασία, ενώ παράλληλα εξετάστηκε και η μέθοδος που ακολουθήθηκε για την απόκτηση αυτών των δεδομένων. Επίσης, καθορίστηκαν οι απαιτήσεις και οι προδιαγραφές της υπό κατασκευή οπτικοποίησης και ορίστηκε με αυστηρό τρόπο ο σχεδιασμός της με τη χρήση προτύπων χαμηλής και υψηλής πιστότητας. Στη συνέχεια, πραγματοποιήθηκε μία αξιολόγηση από την οποία προήλθαν σημαντικά δεδομένα για την πρόληψη και την αποφυγή σφαλμάτων, καθώς και για τη βελτίωση της εφαρμογής της συγκεκριμένης οπτικοποίησης. Έπειτα από δύο διαμορφωτικές αξιολογήσεις, πραγματοποιήθηκε η υλοποίηση της εφαρμογής, η οποία υποβλήθηκε σε ακόμη μία αξιολόγηση, τη συμπερασματική αξιολόγηση. Τα αποτελέσματα της τελευταίας αξιολόγησης ανέδειξαν κάποια προβλήματα, τα οποία αναλύονται στα ανωτέρω κεφάλαια, και τα οποία ωστόσο επιλύθηκαν, προκειμένου να καταλήξουμε στην τελική υλοποίηση της διαδραστικής οπτικοποίησης.

7.2. Τελικά Συμπεράσματα

Τα συμπεράσματα από την παραπάνω διαδικασία είναι τα εξής.

Η αναζήτηση και απόκτηση των δεδομένων είναι μία αρκετά δύσκολη και χρονοβόρα διαδικασία, καθώς τα δεδομένα δεν βρίσκονται σε κάποιο Server του Τμήματος Πληροφορικής και Τηλεπικοινωνιών του ΕΚΠΑ και δεν υπάρχει άμεση πρόσβαση σε αυτά. Το κατέβασμα των δεδομένων με ένα χειροκίνητο ή οποιονδήποτε άλλο μη αυτόματο τρόπο έχει ως αρνητικό χαρακτηριστικό ότι πρέπει να προστεθούν χειροκίνητα και όλα τα μελλοντικά δεδομένα. Η επιλογή ενός Crawler για τη συλλογή των δεδομένων είναι μία απόφαση, η οποία έχει ως στόχο την αυτοματοποίηση της διαδικασίας συλλογής δεδομένων, ώστε να είναι δυνατή και η χρήση του στο μέλλον.

Κατά τη διάρκεια του σχεδιασμού της οπτικοποίησης ένα σημαντικό εργαλείο είναι το Tableau. Το Tableau ενισχύει ιδιαίτερα τη διαδικασία σχεδιασμού, καθώς έχει τη δυνατότητα να μεταβάλλει εύκολα το σχέδιο της οπτικοποίησης κατά τη διάρκεια της αξιολόγησης, ώστε να εξετάζονται παράλληλα πολλές και διάφορες εκδοχές της ίδιας οπτικοποίησης. Όταν εντοπίζεται κάποιο πρόβλημα από έναν χρήστη, ο ερευνητής μπορεί να μετατρέψει το σχεδιασμό της οπτικοποίησης ακριβώς εκείνη τη στιγμή και να δει εάν η συγκεκριμένη σχεδιαστική αλλαγή λύνει το πρόβλημα. Επομένως, σε κάθε σχεδιασμό ένα πρόγραμμα το οποίο μπορεί δυναμικά να αλλάζει και να μεταβάλλει το σχεδιασμό μιας οπτικοποίησης εύκολα και γρήγορα είναι αναγκαίο. Με αυτό τον τρόπο συγκεντρώνονται περισσότερες πληροφορίες από τα άτομα που συμμετέχουν σε μία αξιολόγηση και κερδίζεται αρκετός χρόνος κατά τη διαδικασία του σχεδιασμού.

Από τα συμπεράσματα και των τριών αξιολογήσεων που έλαβαν χώρα στην πτυχιακή εργασία είναι εμφανές ότι το διάγραμμα φυσαλίδων και η επιλογή του για τα συγκεκριμένα δεδομένα είναι πρωτότυπη και επομένως προκαλεί έκπληξη σε

ορισμένους χρήστες. Αυτό αποδεικνύεται από το γεγονός ότι το συγκεκριμένο διάγραμμα είναι αυτό με τις πιο αντιφατικές αξιολογήσεις από τους χρήστες. Δηλαδή, υπάρχουν χρήστες που τους άρεσε η ιδέα του εν λόγω διαγράμματος, ενώ υπάρχουν χρήστες που βρήκαν άστοχη την επιλογή αυτού του διαγράμματος. Παρόλα αυτά, οι περισσότεροι χρήστες κατόρθωσαν να βρουν τις αιτηθείσες από τα σενάρια χρήσης πληροφορίες και από το ανωτέρω γεγονός συμπεραίνεται ότι η συγκεκριμένη οπτικοποίηση είναι εύχρηστη σε ικανοποιητικό βαθμό. Όλα τα παραπάνω έρχονται σε συμφωνία με το γεγονός ότι κάποιοι χρήστες προτείνουν την ύπαρξη ενός πίνακα με τα δεδομένα κάτω από την οπτικοποίηση, ώστε να μπορεί ο χρήστης, χρησιμοποιώντας την αναζήτηση του Web Browser (ctrl+f) να αναζητήσει την πληροφορία που επιθυμεί. Ο συγκεκριμένος τρόπος αναζήτησης είναι ο τρόπος με τον οποίο έχουν «συνηθίσει» οι εν λόγω χρήστες να αναζητούν μία ερευνητική δημοσίευση. Αυτό οφείλεται κυρίως στην υπάρχουσα κατάσταση αναζήτησης ερευνητικών δημοσιεύσεων που έχει αναλυθεί στα προηγούμενα κεφάλαια. Συνεπώς, παρότι η διαδραστική οπτικοποίηση που προτείνεται στη εργασία είναι πιο αποδοτική και ευχάριστη, οι χρήστες τείνουν να αναζητούν την πληροφορία που χρειάζονται με τον τρόπο που έχουν «συνηθίσει».

Ένα επιπλέον συμπέρασμα που προκύπτει από τη διαδικασία αξιολόγησης και υλοποίησης της πτυχιακής εργασίας είναι ότι οι απαιτήσεις και οι ανάγκες των χρηστών συνεχώς αλλάζουν. Όπως διαπιστώθηκε από τις αξιολογήσεις οι ανάγκες των χρηστών ικανοποιούνται και οι περισσότεροι κατορθώνουν να εντοπίσουν την πληροφορία που τους ενδιαφέρει. Εντούτοις, οι ανάγκες αυτές μεταβάλλονται συνεχώς και για το λόγο αυτό η συγκεκριμένη οπτικοποίηση πρέπει να μεταβάλλεται διαρκώς, ώστε να παραμένει αποδοτική.

8. ΜΕΛΛΟΝΤΙΚΕΣ ΕΠΕΚΤΑΣΕΙΣ

Η διαδραστική οπτικοποίηση που σχεδιάστηκε στην πτυχιακή εργασία μετά από τρεις (3) κύκλους αξιολόγησης και επανασχεδιασμού με βάση το ελικοειδές μοντέλο σχεδιασμού, βρίσκεται σε αρκετά καλό επίπεδο ευχρηστίας και εμπειρίας χρήσης. Βέβαια, όσο περισσότεροι κύκλοι αξιολόγησης και επανασχεδιασμού πραγματοποιούνται σε μία εφαρμογή τόσο καλύτερο είναι το αποτέλεσμα που προκύπτει. Στα πλαίσια της παρούσας πτυχιακής εργασίας τρεις (3) αξιολογήσεις και επανασχεδιασμοί κρίθηκαν επαρκείς για μία εύχρηστη και αποδοτική διαδραστική οπτικοποίηση. Ωστόσο, κάποιες μελλοντικές επεκτάσεις - βελτιώσεις, οι οποίες μπορούν να λάβουν χώρα χρησιμοποιώντας το ίδιο μοντέλο σχεδιασμού, αναλύονται τη συνέχεια.

Καταρχήν, μία μελλοντική επέκταση που πιθανολογείται στη βελτίωση της εμπειρίας χρήσης είναι η προσθήκη ενός πιο εξελιγμένου συστήματος αναζήτησης. Ένα σύστημα αναζήτησης το οποίο θα αναζητά πολλαπλές λέξεις-κλειδιά και όχι μόνο μία λέξη-κλειδί είναι μία καλή μελλοντική επέκταση.

Ένα πρόβλημα που παρατηρείται από ορισμένους χρήστες στην τελευταία αξιολόγηση και αναφέρεται στο κεφάλαιο της υλοποίησης είναι ότι η οπτικοποίηση παρουσιάζει μια μικρή καθυστέρηση στην κίνηση των φυσαλίδων (Bubble Animation). Πιθανή βελτίωση είναι η βελτιστοποίηση του κώδικα που αφορά στις επιδόσεις, ώστε να μπορεί να τρέχει ομαλά και σε υπολογιστικά συστήματα με μικρότερες επιδόσεις. Το παραπάνω μπορεί να υλοποιηθεί με αρκετούς τρόπους παρόλα αυτά μία λύση που αξίζει να αναφερθεί είναι η ενσωμάτωση δομών στον κώδικα, οι οποίες αποθηκεύουν τη θέση των φυσαλίδων του εκάστοτε αντικειμένου που επιθυμούμε να έχει κίνηση και δεν υπολογίζεται κάθε φορά στα συστήματα στα οποία υπάρχει ανεπάρκεια πόρων. Συνεπώς, σε αυτά τα συστήματα προβάλλεται μία εναλλακτική εκδοχή της οπτικοποίησης, η οποία δεν θα έχει τόσο παραστατική κίνηση (animation) αλλά θα αποτελείται από τις τελικές θέσεις των αντικειμένων που έχουν υπολογιστεί και αποθηκευτεί.

Η επιλογή ενός Crawler για το συστηματικό κατέβασμα των δεδομένων στο μέλλον ήταν μία κρίσιμη απόφαση που λήφθηκε κατά τη διάρκεια εκπόνησης της εργασίας. Παρόλο που ο Crawler κατεβάζει αυτόματα δεδομένα, η εκτέλεσή του γίνεται χειροκίνητα. Η μελλοντική επέκταση στο ανωτέρω ζήτημα είναι η αυτόματη εκτέλεση του Crawler περιοδικά και η αυτόματη ενσωμάτωση των νέων δεδομένων στη διαδραστική οπτικοποίηση που κατασκευάστηκε.

Μία μελλοντική επέκταση, η οποία συνδέεται άρρηκτα με την προηγούμενη επέκταση που αναφέρεται είναι η επέκταση των επιλογών που είναι διαθέσιμες στα φίλτρα επιλογής συγγραφέα. Τα δεδομένα περιορίζονται στις δημοσιεύσεις δεκαεπτά (17) καθηγητών του Τμήματος Πληροφορικής και Τηλεπικοινωνιών του ΕΚΠΑ που διαθέτουν προφίλ στο Google Scholar. Συνεπώς, το φίλτρο επιλογής συγγραφέα έχει και αυτό δεκαεπτά (17) επιλογές. Στην περίπτωση προσθήκης περισσότερων δεδομένων από άλλους συγγραφείς πρέπει να προστεθούν και οι συγκεκριμένοι συγγραφείς στο φίλτρο επιλογής συγγραφέα. Παρόλα αυτά, μία ενδιαφέρουσα προσθήκη είναι η ανάλυση των δεδομένων και η συλλογή όλων των συγγραφέων που έχουν συνεισφέρει σε κάθε δημοσίευση που υπάρχει στα εν λόγω δεδομένα και η προσθήκη αυτών στο φίλτρο επιλογής συγγραφέα με αυτόματο τρόπο.

Εν κατακλείδι, από τις ως άνω αναφερόμενες επεκτάσεις αποδεικνύεται ολοφάνερα ότι η διαδικασία βελτίωσης μιας εφαρμογής δεν παύει ποτέ και υπάρχει πάντα το περιθώριο για περαιτέρω βελτίωση. Στα πλαίσια της παρούσας πτυχιακής εργασίας οι παραπάνω προεκτάσεις δεν υλοποιήθηκαν, ωστόσο παρουσιάζει έντονο ενδιαφέρον η μελλοντική υλοποίησή τους και η παρατήρηση των δυνατοτήτων αναζήτησης και ευχρηστίας της συγκεκριμένης διαδραστικής οπτικοποίησης με δεδομένα που θα προέρχονται από περισσότερα Τμήματα του ΕΚΠΑ ή και από περισσότερες Σχολές.

ΠΙΝΑΚΑΣ ΟΡΟΛΟΓΙΑΣ

Ξενόγλωσσος όρος	Ελληνικός Όρος
null	Κενή μεταβλητή
in-memory data engine	Μηχανή δεδομένων εντός μνήμης
Web Crawler (spider)	Αυτόματος Συλλέκτης Δεδομένων Ιστοσελίδων
python	Γλώσσα προγραμματισμού
HTML request	Με βάση το πρωτόκολλο «http» ένας Web Browser στέλνει ένα αίτημα (HTML request) προσπαθώντας να προσπελάσει έναν ιστοχώρο.
HTML response	Με βάση το πρωτόκολλο «http» ένας Web Browser λαμβάνει μια απάντηση (HTML response) σε ένα αίτημα που έχει κάνει (HTML request) προσπαθώντας να προσπελάσει έναν ιστοχώρο.

ΣΥΝΤΜΗΣΕΙΣ – ΑΡΚΤΙΚΟΛΕΞΑ – ΑΚΡΩΝΥΜΙΑ

ΕΚΠΑ	Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών
CSV	Αρχείο δεδομένων που χωρίζεται με κόμμα (Comma-separated file). Παρατίθεται παράδειγμα στο Παράρτημα Ι.
Bibtex	Τύπος αρχείου δεδομένων με συγκεκριμένο συντακτικό. Παρατίθεται παράδειγμα στο Παράρτημα Ι.
reCAPTCHA	Αυτοματοποιημένο τεστ διαχωρισμού ανθρώπων από προγράμματα (Completely Automated Public Turing test to tell Computers and Humans Apart)
JSON	Αρχείο δεδομένων με συγκεκριμένη μορφή. Παρατίθεται παράδειγμα στο Παράρτημα Ι.
DIT	Τμήμα Πληροφορικής και Τηλεπικοινωνιών του Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών (Department of Informatics and Telecommunications)

ΠΑΡΑΡΤΗΜΑ Ι

```
@article{roussou2004learning,
  title={Learning by doing and learning through play: an exploration of interactivity in virtual environments for children},
  author={Roussou, Maria},
  journal={Computers in Entertainment (CIE)},
  volume={2},
  number={1},
  pages={10--10},
  year={2004},
  publisher={ACM New York, NY, USA}
}
```

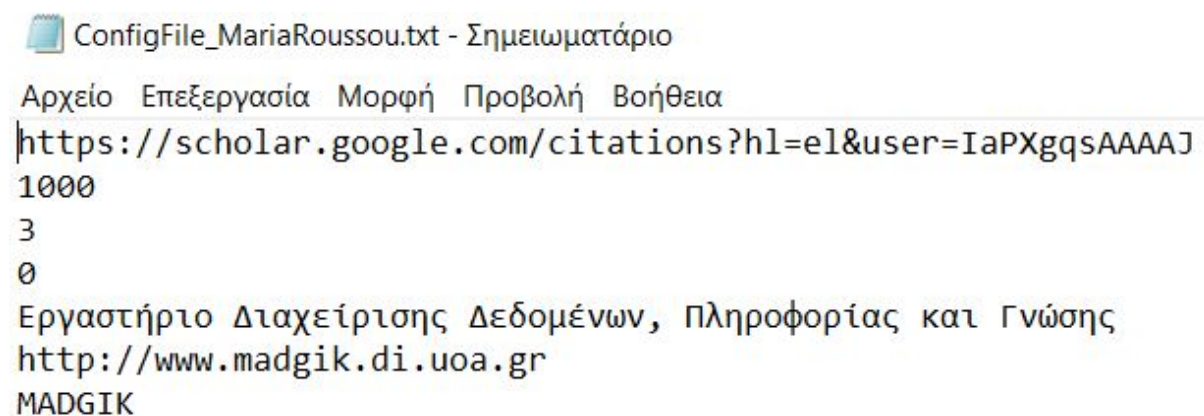
Εικόνα 23: Παράδειγμα Εγγραφής Αρχείου Bibtex

```
{
  #
  [
  {
    "link": "https://www.cs.cornell.edu/johannes/papers/1998/sigmod1998-clique.pdf",
    "title": "Automatic subspace clustering of high dimensional data for data mining applications",
    "year": "1998",
    "lab": "Εργαστήριο Διαχείρισης Δεδομένων, Πληροφορίας και Γνώσης",
    "lablink": "http://www.madgik.di.uoa.gr",
    "labShortName": "MADGIK",
    "Συγγραφείς": "Rakesh Agrawal, Johannes Gehrke, Dimitrios Gunopulos, Prabhakar Raghavan",
    "Ημερομηνία δημοσίευσης": "1998/6/1",
    "Βιβλίο": "Proceedings of the 1998 ACM SIGMOD international conference on Management of data",
    "Σελίδες": "94-105",
    "Περιγραφή": "Data mining applications place special requirements on clustering algorithms including: the ability to find clusters embedded in high dimensional data",
    "Σύνολο παραθέσεων": "Γίνεται αναφορά σε 3648",
    "Άρθρα στον Μελετητή": "Automatic subspace clustering of high dimensional data for data mining applicationsR Agrawal, J Gehrke, D Gunopulos",
    "dit_author": "Dimitrios Gunopulos",
    "data_from": "https://scholar.google.com/citations?hl=el&user=PHJokEQAAAAJ",
    "local_id": "1"
  },
  {
    "link": "Link unavailable.",
    "title": "Data classification: algorithms and applications",
    "year": "2014",
    "lab": "Εργαστήριο Διαχείρισης Δεδομένων, Πληροφορίας και Γνώσης",
    "lablink": "http://www.madgik.di.uoa.gr",
    "labShortName": "MADGIK",
    "Συγγραφείς": "Charu C Aggarwal",
    "Ημερομηνία δημοσίευσης": "2014/7/25",
    "Εκδότης": "CRC press",
    "Περιγραφή": "Comprehensive Coverage of the Entire Area of Classification Research on the problem of classification tends to be fragmented",
    "Σύνολο παραθέσεων": "Γίνεται αναφορά σε 1550",
    "Άρθρα στον Μελετητή": "Data classification: algorithms and applicationsCC Aggarwal - 2014Γίνεται αναφορά σε 1550 Σχετικά άρθρα όλες οι 7",
    "dit_author": "Dimitrios Gunopulos",
    "data_from": "https://scholar.google.com/citations?hl=el&user=PHJokEQAAAAJ",
    "local_id": "2"
  }
]
```

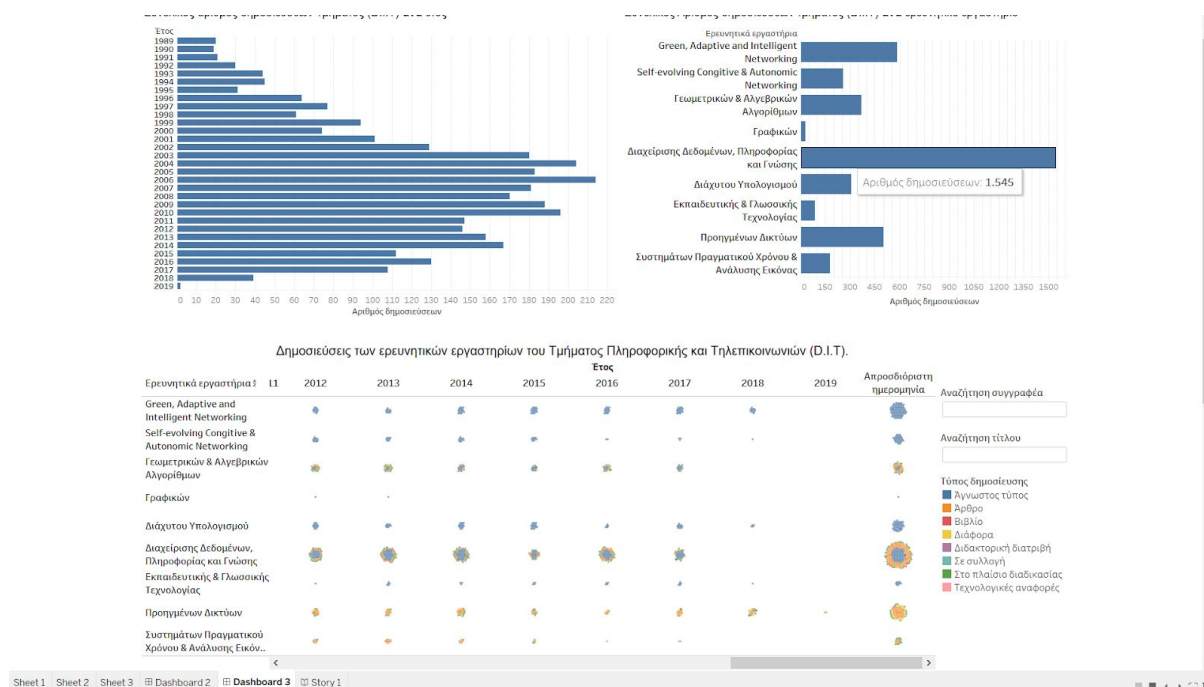
Εικόνα 24: Παράδειγμα Πρώτου Αρχείου Δεδομένων JSON

	C	D	E	F	G
1	title	year	lab	url	author
2	Automatic subspace clustering of high dimensional data for data mining applications	1998	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK Rakesh Agrawal, Johannes Gehrke, Dimitrios Gunopulos, Prabhakar Raghu
3	Data classification: algorithms and applications	2014	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK Charu C Aggarwal
4	Discovering similar multidimensional trajectories	2002	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK Michail Vlachos, George Kollios, Dimitrios Gunopulos
5	Mining process models from workflow logs	1998	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK Rakesh Agrawal, Dimitrios Gunopulos, Frank Leymann
6	Data streams: models and algorithms	2007	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK Charu C Aggarwal
7	Constraint-based rule mining in large, dense databases	2000	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK Roberto J Bayardo, Rakesh Agrawal, Dimitrios Gunopulos
8	A local search mechanism for peer-to-peer networks	2002	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK Vana Kalogeraki, Dimitrios Gunopulos, Demetrios Zemanipour-Yazti
9	On indexing mobile objects	1999	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK George Kollios, Dimitrios Gunopulos, Vassilis J Tsotras
10	Online outlier detection in sensor data using non-parametric models	2006	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK Sharmila Subramaniam, Themis Palpanas, Dimitris Papadopoulos, Vana Kalogeraki
11	Indexing multi-dimensional time-series with support for multiple distance measures	2003	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK Michail Vlachos, Marios Hadjieleftheriou, Dimitrios Gunopulos, Eamonn Keogh
12	Finding similar time series	1997	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK Gautam Das, Dimitrios Gunopulos, Heikki Mannila
13	Locally adaptive metric nearest-neighbor classification	2002	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK Carlotta Domeniconi, Jing Peng, Dimitrios Gunopulos
14	Automatic subspace clustering of high dimensional data	2005	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK Rakesh Agrawal, Johannes Gehrke, Dimitrios Gunopulos, Prabhakar Raghu
15	Identifying similarities, periodicities and bursts for online search queries	2004	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK Michail Vlachos, Christopher Meek, Zografoula Vagena, Dimitrios Gunopulos
16	Indexing large human-motion databases	2004	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK Eamonn Keogh, Themistoklis Palpanas, Victor B Zordan, Dimitrios Gunopulos
17	Data mining, hypergraph transversals, and machine learning	1997	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK Dimitrios Gunopulos, Heikki Mannila, Roni Khardon, Hannu Toivonen
18	Locally adaptive metrics for clustering high dimensional data	2007	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK Carlotta Domeniconi, Dimitrios Gunopulos, Sheng Ma, Bojun Yan, Munia A
19	Prospective customer selection using customer and market reference data	2000	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK Dimitrios Gunopulos, George Kollios, Vassilis J Tsotras, Carlotta Domeniconi
20	Approximating multi-dimensional aggregate range queries over real attributes	2003	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK Dimitrios Gunopulos, Roni Khardon, Heikki Mannila, Sangeer Saluja, Hannu Toivonen
21	Discovering all most specific sentences	2003	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK George Kollios, Dimitrios Gunopulos, Nick Koudas, Stefan Berchtold
22	Efficient based sampling for approximate clustering and outlier detection in large data sets	2003	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK Themistoklis Palpanas, Dimitris Papadopoulos, Vana Kalogeraki, Dimitrios Gunopulos
23	Distributed deviation detection in sensor networks	2004	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK Jessica Lin, Michail Vlachos, Eamonn Keogh, Dimitrios Gunopulos
24	Iterative incremental clustering of time series	2010	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK Theodoros Lappas, Evimaria Terzi, Dimitrios Gunopulos, Heikki Mannila
25	Finding effectors in social networks	2001	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK Carlotta Domeniconi, Dimitrios Gunopulos
26	Incremental support vector machine construction	2006	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK Gautam Das, Dimitrios Gunopulos, Nick Koudas, Dimitris Tsirigianis
27	Answering top-k queries using views	2002	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK Michail Vlachos, Carlotta Domeniconi, Dimitrios Gunopulos, George Kollios
28	Non-linear dimensionality reduction techniques for classification and visualization	2002	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK Marios Hadjieleftheriou, George Kollios, Vassilis J Tsotras, Dimitrios Gunopulos
29	Efficient indexing of spatiotemporal objects	2005	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK Demetrios Zemanipour-Yazti, Song Lin, Vana Kalogeraki, Dimitrios Gunopulos
30	MicroHash: An Efficient Index Structure for Flash-Based Sensor Devices.	2004	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK Themistoklis Palpanas, Michail Vlachos, Eamonn Keogh, Dimitrios Gunopulos
31	Online amnesic approximation of streaming time series	2003	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK Michail Vlachos, Jessica Lin, Eamonn Keogh, Dimitrios Gunopulos
32	A wavelet-based anytime algorithm for k-means clustering of time series	2004	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK Carlotta Domeniconi, Dimitris Papadopoulos, Dimitrios Gunopulos, Sheng Ma
33	Subspace clustering of high dimensional data	2005	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK Chaitan An Ratanamahatana, Jessica Lin, Dimitrios Gunopulos, Eamonn Keogh
34	Mining time series data	1997	Εργαστήριο Διαχείρισης Δεδομένων	http://www.madgik.di.uoa.gr	MADGIK Béla Bollobás, Gautam Das, Dimitrios Gunopulos, Heikki Mannila
35	Time-series similarity problems and well-separated geometric sets				

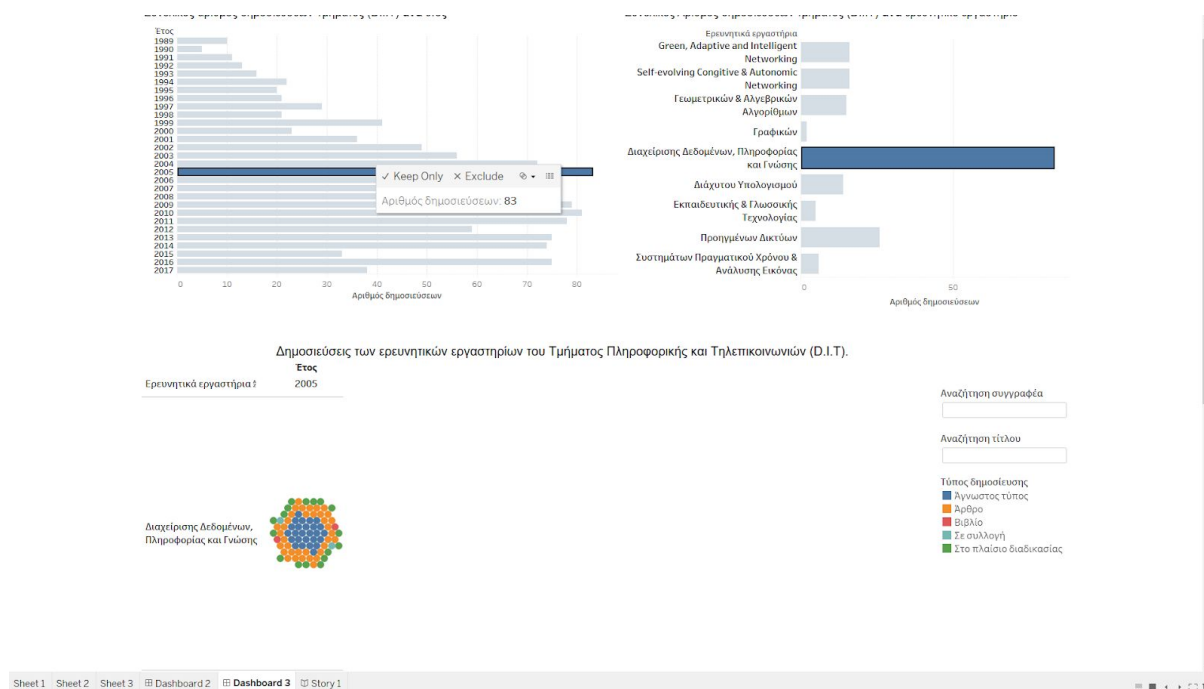
Εικόνα 25: Παράδειγμα Αρχείου Δεδομένων CSV



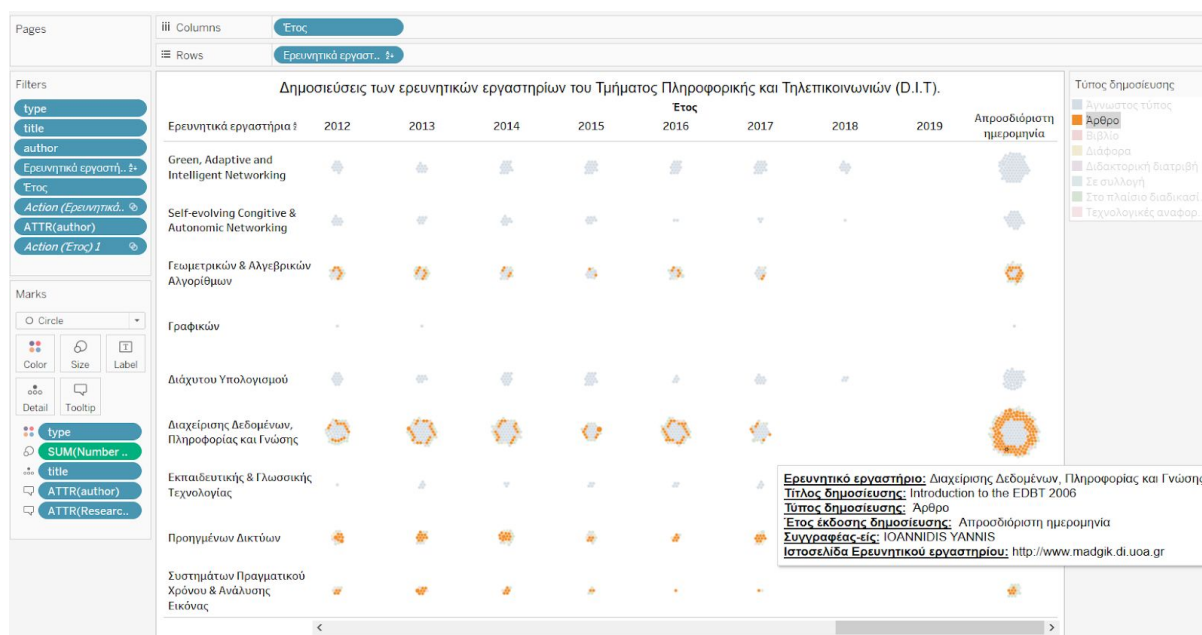
Εικόνα 26: Παράδειγμα Configuration File όπου ο Crawler παίρνει σαν όρισμα



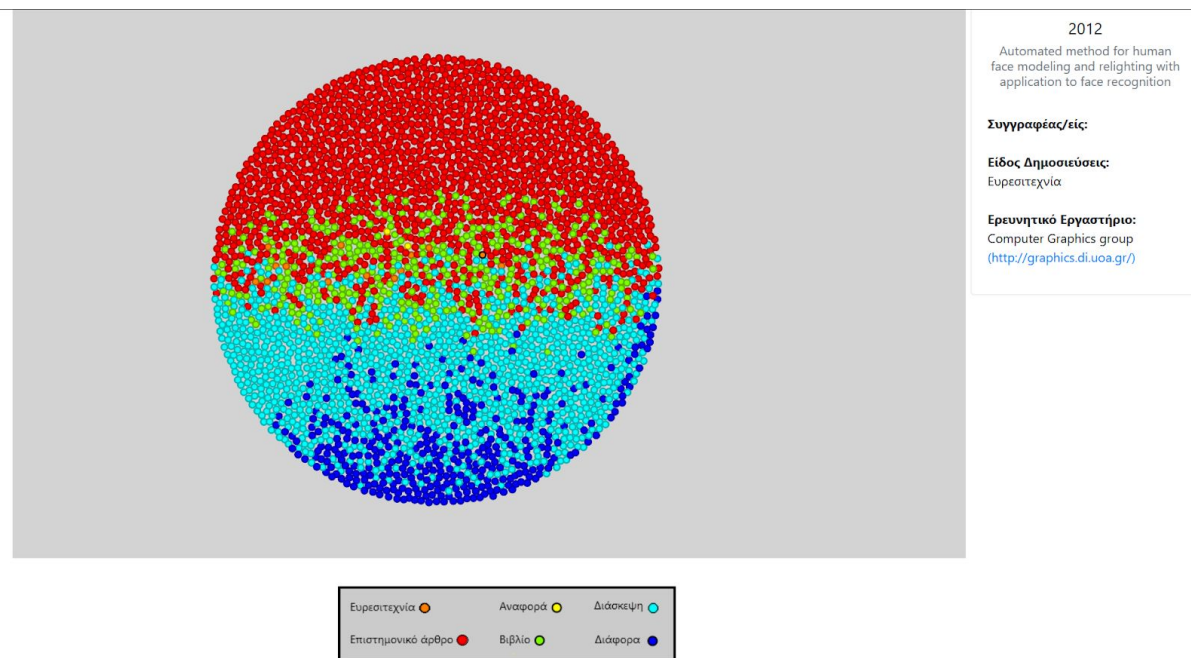
Εικόνα 27: Στιγμιότυπο οθόνης από την δεύτερη διαμορφωτική αξιολόγηση (εναλλακτική έκδοση πρωτοτύπου υψηλής πιστότητας, που παρουσιάστηκε σε μερικούς χρήστες, με τα ραβδογράμματα από πάνω από το διάγραμμα φυσαλίδων)



Εικόνα 28: Στιγμιότυπο οθόνης από πρωτότυπο υψηλής πιστότητας που παρουσιάστηκε στους χρήστες στην δεύτερη διαμορφωτική αξιολόγηση – χρήση ραβδογραμμάτων σαν φίλτρα



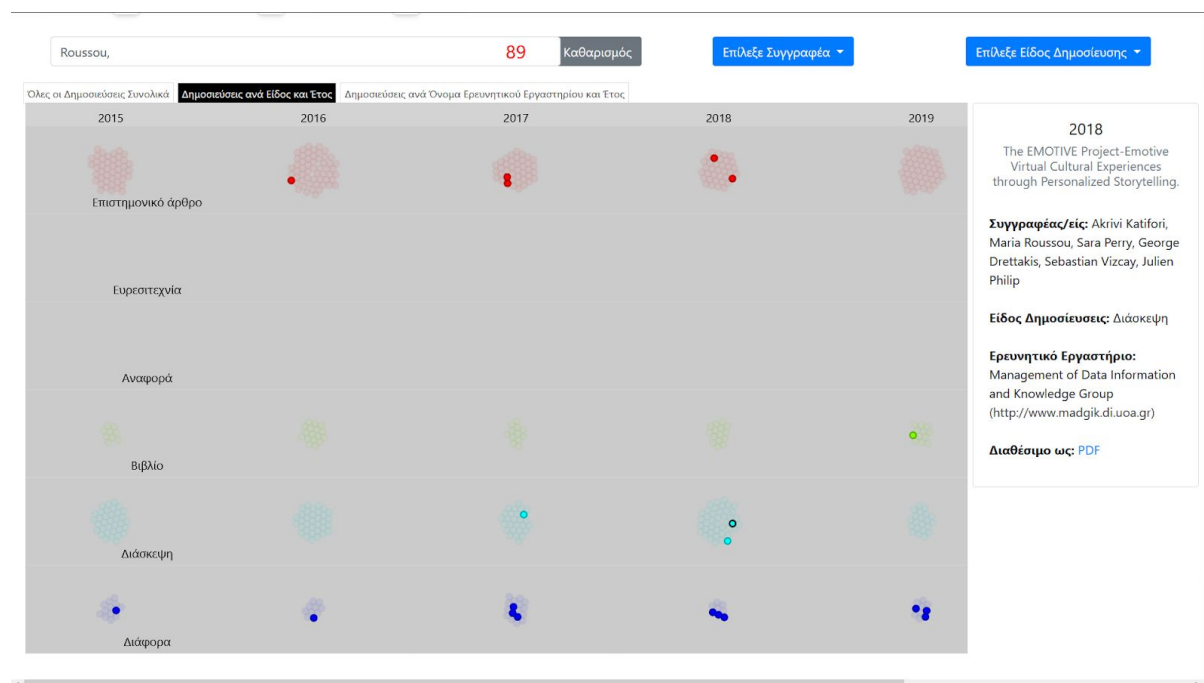
Εικόνα 29: Στιγμιότυπο οθόνης από πρωτότυπο υψηλής πιστότητας που παρουσιάστηκε στους χρήστες στην δεύτερη διαμορφωτική αξιολόγηση – χρήση φίλτρου επιλογής συγκεκριμένου είδους δημοσίευσης



Εικόνα 30: Στιγμιότυπο οθόνης από την τελική οπτικοποίηση - χρήση της καρτέλας «Όλες οι δημοσιεύσεις συνολικά»



Εικόνα 31: Στιγμιότυπο οθόνης από την τελική οπτικοποίηση - χρήση της καρτέλας «Δημοσιεύσεις ανά είδος και έτος» σε «zoom out»



Εικόνα 32: Στιγμιότυπο οθόνης από την χρήση του φίλτρου «Επέλεξε Συγγραφέα» στην τελική οπτικοποίηση - χρήση της καρτέλας «Δημοσιεύσεις ανά είδος και έτος»

ΦΟΡΜΑ ΣΥΓΚΑΤΑΘΕΣΗΣ

ΣΥΜΜΕΤΟΧΗ ΣΤΗΝ ΕΡΕΥΝΑ ΑΞΙΟΛΟΓΗΣΗΣ ΤΗΣ ΥΠΟ ΔΙΑΜΟΡΦΩΣΗ ΔΙΑΔΡΑΣΤΙΚΗΣ ΟΠΤΙΚΟΠΟΙΗΣΗΣ ΤΙΣ ΕΡΕΥΝΗΤΙΚΗΣ ΔΡΑΣΤΗΡΙΟΤΗΤΑΣ ΤΟΥ ΤΜΗΜΑΤΟΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ ΤΟΥ ΕΚΠΑ (DIT)

Η παρούσα έρευνα διεξάγεται από τον προπτυχιακό φοιτητή Ξένο Παναγιώτη στα πλαίσια της πτυχιακής του εργασίας με επιβλέπουσα την Επίκουρη Καθηγήτρια Μαρία Ρούσσου. Ο σκοπός είναι η δοκιμή και βελτίωση της υπό διαμόρφωσης διαδραστικής οπτικοποίησης δεδομένων ερευνητικής δραστηριότητας Τμήματος (DIT), που θα βρίσκεται στην καινούργια ιστοσελίδα του Τμήματος Πληροφορικής και Τηλεπικοινωνιών. Η συγκεκριμένη εφαρμογή θα αναδεικνύει την έρευνα του Τμήματος Πληροφορικής και Τηλεπικοινωνιών. Η έρευνα αυτή θα συμβάλει στην αξιολόγηση της ευχρηστίας της διαδραστικής οπτικοποίησης και τη διασφάλιση της ικανοποίησης των απαιτήσεων των χρηστών της.

Η διαδικασία της αξιολόγησης είναι η εξής: αρχικά θα σας ζητηθεί να εκτελέσετε κάποια σενάρια εργασιών. Κατά τη διάρκεια της έρευνας, οι ερευνητές θα ρωτούν ερωτήσεις και θα κρατούν σημειώσεις. Θα σας ζητηθεί να εξηγήσετε τις ενέργειες σας και τις αποφάσεις που παίρνετε ενόσω αλληλεπιδράτε με την εφαρμογή. Στη συνέχεια θα σας ζητηθεί να συμπληρώσετε ένα ερωτηματολόγιο σχετικά με την εμπειρία από την χρήση της οπτικοποίησης.

Στη διάρκεια της αξιολόγησης θα υπάρχει η πιθανότητα ηχογράφησης και λήψης φωτογραφιών. Το υλικό θα χρησιμοποιηθεί αποκλειστικά και μόνο για την ανάλυση των δεδομένων της παρούσας έρευνας και θα καταστραφεί μετά το πέρας αυτής.

Θα θέλαμε να φωτογραφίσουμε κάποιες από τις στιγμές της έρευνας για χρήση στην ανάλυση και στο κείμενο της πτυχιακής εργασίας.

Μπορείτε να αποσυρθείτε οποιαδήποτε στιγμή από την έρευνα χωρίς καμία οφειλή ή υποχρέωση.

Δέχομαι να συμμετάσχω στην έρευνα	
Έχω πληροφορηθεί επαρκώς για τους στόχους και τη διαδικασία της συγκεκριμένης έρευνας	
Κατανοώ ότι η συμμετοχή μου στην έρευνα είναι εθελοντική και ότι δεν υπάρχει χρηματική αμοιβή ή υποχρέωση οποιασδήποτε μορφής	
Συμφωνώ στην ηχογράφησή μου κατά τη διάρκεια της έρευνας	
Συμφωνώ στη φωτογράφησή μου κατά τη διάρκεια της έρευνας	
Συμφωνώ στην ανώνυμη δημοσίευση φωτογραφιών μου στην πτυχιακή εργασία	

/ /2018

ΟΝΟΜΑΤΕΠΩΝΥΜΟ

ΥΠΟΓΡΑΦΗ

ΗΜΕΡΟΜΗΝΙΑ

Η υπογραφή σας υποδηλώνει συγκατάθεση στους παραπάνω όρους που αφορούν στην εθελοντική συμμετοχή στην παρούσα έρευνα. Ως συμμετέχοντες στην έρευνα έχετε ενημερωθεί για το δικαίωμα πρόσβασης στα προσωπικά σας δεδομένα σύμφωνα με όσα ορίζονται στις σχετικές διατάξεις του Ν. 2472/97 για την προστασία του ατόμου από την επεξεργασία δεδομένων προσωπικού χαρακτήρα.

Σας ενημερώνουμε ότι για την αρχειοθέτηση των στοιχείων σας (ονοματεπώνυμο και διεύθυνση ηλεκτρονικού ταχυδρομείου) τηρούμε όλα όσα προβλέπει ο νέος Κανονισμός της Ευρωπαϊκής Ένωσης για την Γενική Προστασία Δεδομένων (GDPR) και κανείς άλλος δεν έχει πρόσβαση στα στοιχεία αυτά. Δεν μοιραζόμαστε ή εμπορεύμαστε ποτέ τα προσωπικά σας δεδομένα με τρίτους φορείς.

Ερωτηματολόγιο Δεύτερης Αξιολόγησης (Google Forms)

Προσωπικά στοιχεία

Κωδικός συμμετέχοντος (ID)

Η απάντησή σας

Επάγγελμα

Η απάντησή σας

Φίλο

☐ Άντρας

☐ Γυναίκα

Ηλικία

Η απάντησή σας

Πίσω Επόμενο

Σελίδα 2 από 5

Εικόνα 33: Στιγμιότυπο οθόνης ερωτηματολογίου που χρησιμοποιήθηκε κατά την δεύτερη διαμορφωτική αξιολόγηση (1 από 9)

Ερωτήσεις σχετικά με την εμπειρία περιήγησης

Πόσο κατανοητό είναι κάθε διάγραμμα ξεχωριστά; Βαθμολογήστε από 1-5.

	1 (Καθόλου κατανοητή)	2	3	4	5 (Απόλυτα κατανοητή)
Διάγραμμα φυσαλίδων:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ραβδόγραμμα συνόλου δημοσιεύσεων Τμήματος Πληροφορικής και Τηλεπικοινωνιών ανά ερευνητικό εργαστήριο:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ραβδόγραμμα συνολικών δημοσιεύσεων Τμήματος Πληροφορικής και Τηλεπικοινωνιών ανά χρόνια:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Παρατηρήσεις

Η απάντησή σας

Εικόνα 34: Στιγμιότυπο οθόνης ερωτηματολογίου που χρησιμοποιήθηκε κατά την δεύτερη διαμορφωτική αξιολόγηση (2 από 9)

Πόσο κατανοητή είναι η παρουσία διαδραστική οπτικοποίηση στο σύνολο της;
Βαθμολογήστε από 1-5.

1

2

3

4

5

Καθόλου κατανοητή

☐

☐

☐

☐

☐

Απόλυτα κατανοητή

Παρατηρήσεις

Η απάντησή σας

Πίσω

Επόμενο

Σελίδα 3 από 5

Εικόνα 35: Στιγμιότυπο οθόνης ερωτηματολογίου που χρησιμοποιήθηκε κατά την δεύτερη διαμορφωτική αξιολόγηση (3 από 9)

Σενάρια χρήσης

Σενάριο χρήσης 1: Πόσες είναι οι συνολικές δημοσιεύσεις του Τμήματος Πληροφορικής και Τηλεπικοινωνιών του ΕΚΠΑ σε ένα συγκεκριμένο έτος. Παραδείγματος χάριν το 2006; (ώρα διεκπεραίωσης σε δευτερόλεπτα)

Η απάντησή σας _____

Παρατηρήσεις σεναρίου χρήσης 1

Η απάντησή σας _____

Σενάριο χρήσης 2: Ποιες είναι οι δημοσιεύσεις ενός συγκεκριμένου συγγραφέα; Παραδείγματος χάριν “Μαρία Ρούσσου”. (ώρα διεκπεραίωσης σε δευτερόλεπτα)

Η απάντησή σας _____

Παρατηρήσεις σεναρίου χρήσης 2

Η απάντησή σας _____

Εικόνα 36: Στιγμιότυπο οθόνης ερωτηματολογίου που χρησιμοποιήθηκε κατά την δεύτερη διαμορφωτική αξιολόγηση (4 από 9)

Σενάριο χρήσης 3: Αναζήτηση με βάση τίτλο δημοσιεύσεις. Παραδείγματος χάριν “Learning by doing and learning through play: an exploration of interactivity in virtual environments for children”. (ώρα διεκπεραίωσης σε δευτερόλεπτα)

Η απάντησή σας

Παρατηρήσεις σεναρίου χρήσης 3

Η απάντησή σας

Σενάριο χρήσης 4: Σε ένα συγκεκριμένο έτος και για ένα συγκεκριμένο εργαστήριο να βρεθούν πόσες δημοσιεύσεις συγκεκριμένου είδους υπάρχουν. Παραδείγματος χάριν “Το 2006, από το εργαστήριο Magic, πόσες τεχνολογικές αναφορές δημοσιεύτηκαν;” (ώρα διεκπεραίωσης σε δευτερόλεπτα)

Η απάντησή σας

Παρατηρήσεις σεναρίου χρήσης 4

Η απάντησή σας

Εικόνα 37: Στιγμιότυπο οθόνης ερωτηματολογίου που χρησιμοποιήθηκε κατά την δεύτερη διαμορφωτική αξιολόγηση (5 από 9)

Σενάριο χρήσης 5: Αναζήτηση η λέξης κλειδί. Παραδείγματος χάριν “Database”.
(ώρα διεκπεραίωσης σε δευτερόλεπτα)

Η απάντησή σας

Παρατηρήσεις σεναρίου χρήσης 5

Η απάντησή σας

Γενικές παρατηρήσεις

Η απάντησή σας

[Πίσω](#)[Επόμενο](#)

Σελίδα 4 από 5

Εικόνα 38: Στιγμιότυπο οθόνης ερωτηματολογίου που χρησιμοποιήθηκε κατά την δεύτερη διαμορφωτική αξιολόγηση (6 από 9)

Ερωτήσεις σχετικά με τα φίλτρα και την ευχρηστία - εμπειρία χρήσης της διαδραστικής οπτικοποίησης

Σχετικά με την συνολικότερη εμπειρία χρήσης της διαδραστικής οπτικοποίησης

	1	2	3	4	5	
Καθόλου καλή εμπειρία χρήσης	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Πολύ καλή εμπειρία χρήσης

Παρατηρήσεις

Η απάντησή σας

Εικόνα 39: Στιγμιότυπο οθόνης ερωτηματολογίου που χρησιμοποιήθηκε κατά την δεύτερη διαμορφωτική αξιολόγηση (7 από 9)

Πόσο βοηθητικά ήταν τα φίλτρα στην επίτευξη αυτών που σας ζητηθήκαν;

	1	2	3	4	5	
Καθολου βοηθητικά	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Πολυ βοηθητικά

Παρατηρήσεις

Η απάντησή σας

Εικόνα 40: Στιγμιότυπο οθόνης ερωτηματολογίου που χρησιμοποιήθηκε κατά την δεύτερη διαμορφωτική αξιολόγηση (8 από 9)

Χρησιμοποίησες τη λειτουργία στην οποία κλικάρεις τα δύο ραβδογράμματα για να επηρεάσεις το περιεχόμενο του κυρίως διαγράμματος;

☐ Ναι, ήταν πολύ χρήσιμη

☐ Όχι, δεν ήταν χρήσιμη

☐ Δεν την παρατήρησα

Παρατηρήσεις

Η απάντησή σας

Υπάρχει κάτι που σε μπέρδεψε ή κάτι που θα ήθελες να ήταν διαφορετικό; (γράψε παρακάτω)

Η απάντησή σας

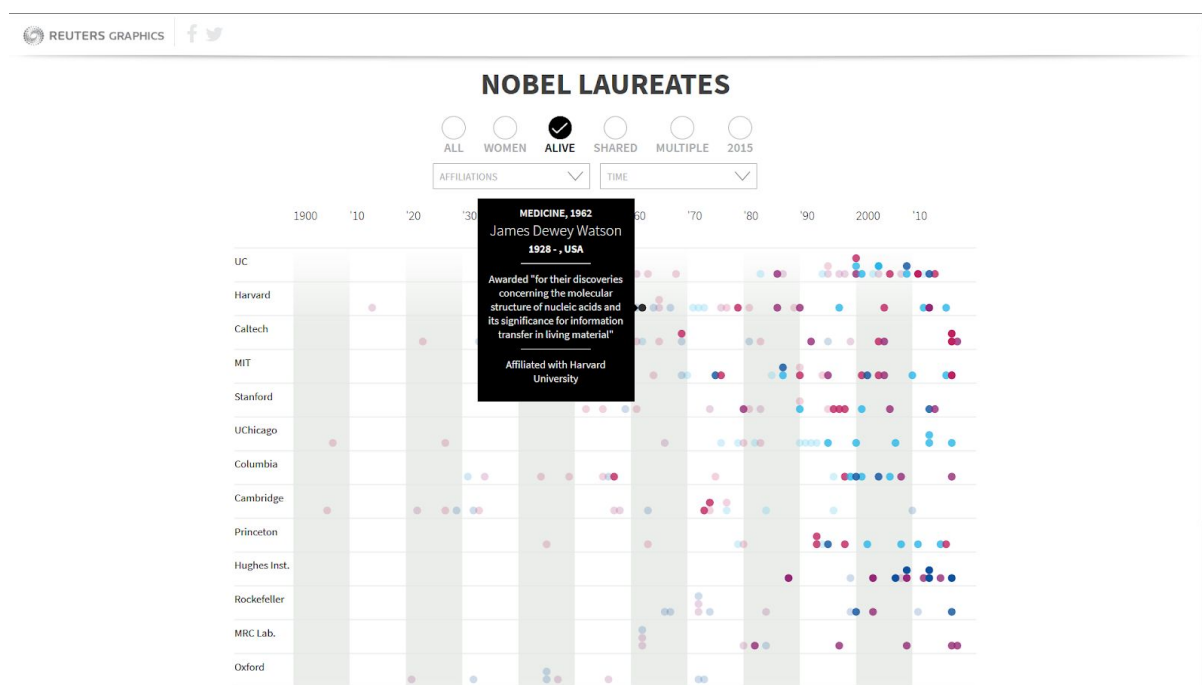
Πίσω

Υποβολή

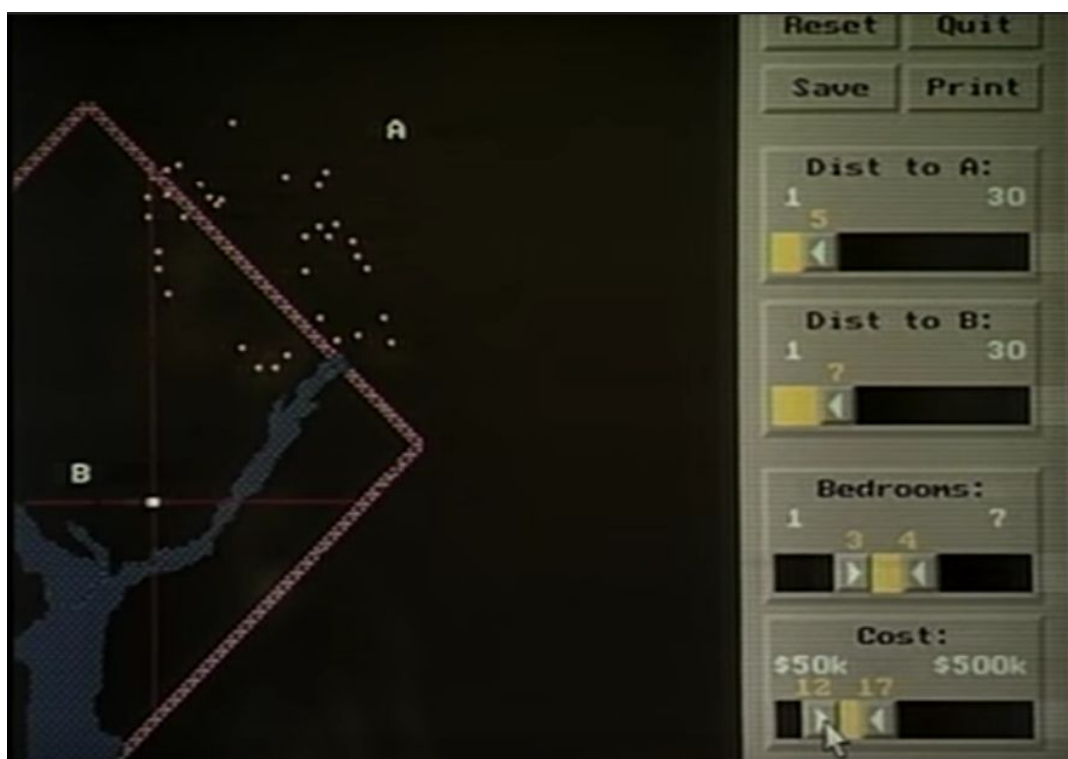
Σελίδα 5 από 5

Εικόνα 41: Στιγμιότυπο οθόνης ερωτηματολογίου που χρησιμοποιήθηκε κατά την δεύτερη διαμορφωτική αξιολόγηση (9 από 9)

ΠΑΡΑΡΤΗΜΑ II



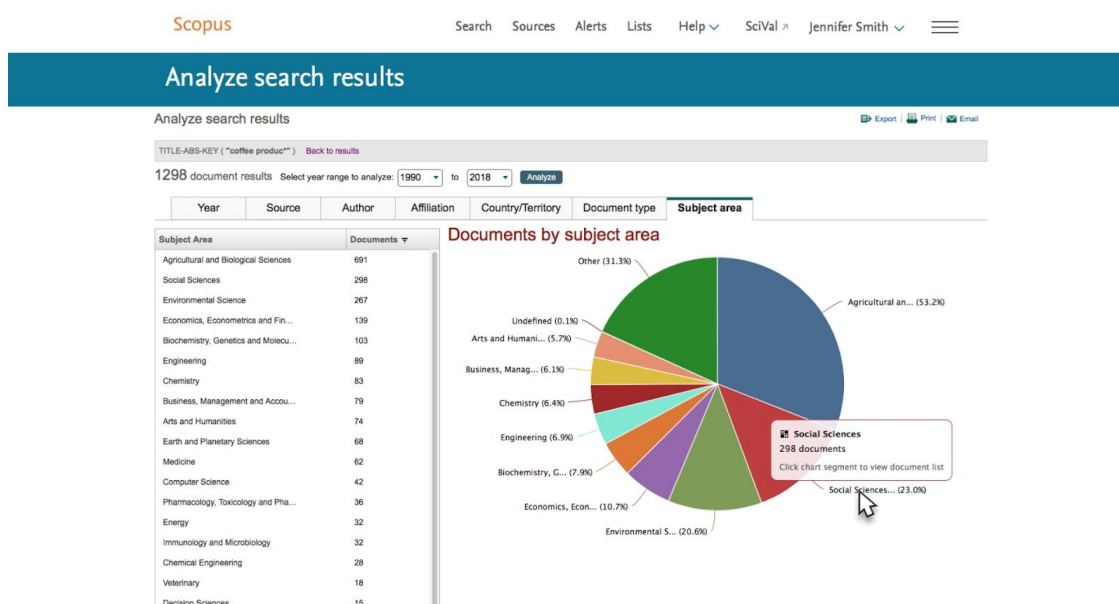
Εικόνα 42: Οπτικοποίηση Απονομών των Βραβείων Νόμπελ που Έχουν Πραγματοποιηθεί Μέχρι Σήμερα



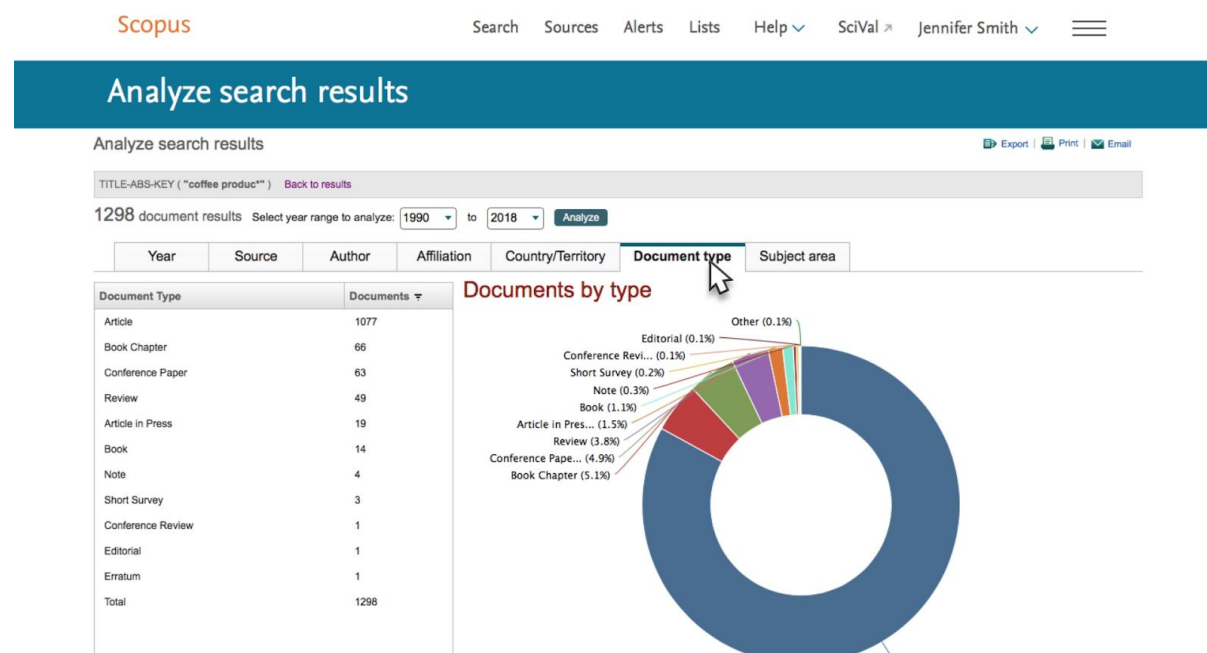
Εικόνα 43: Dynamic Queries Home Finder - παράδειγμα χρήσης φίλτρου A & B



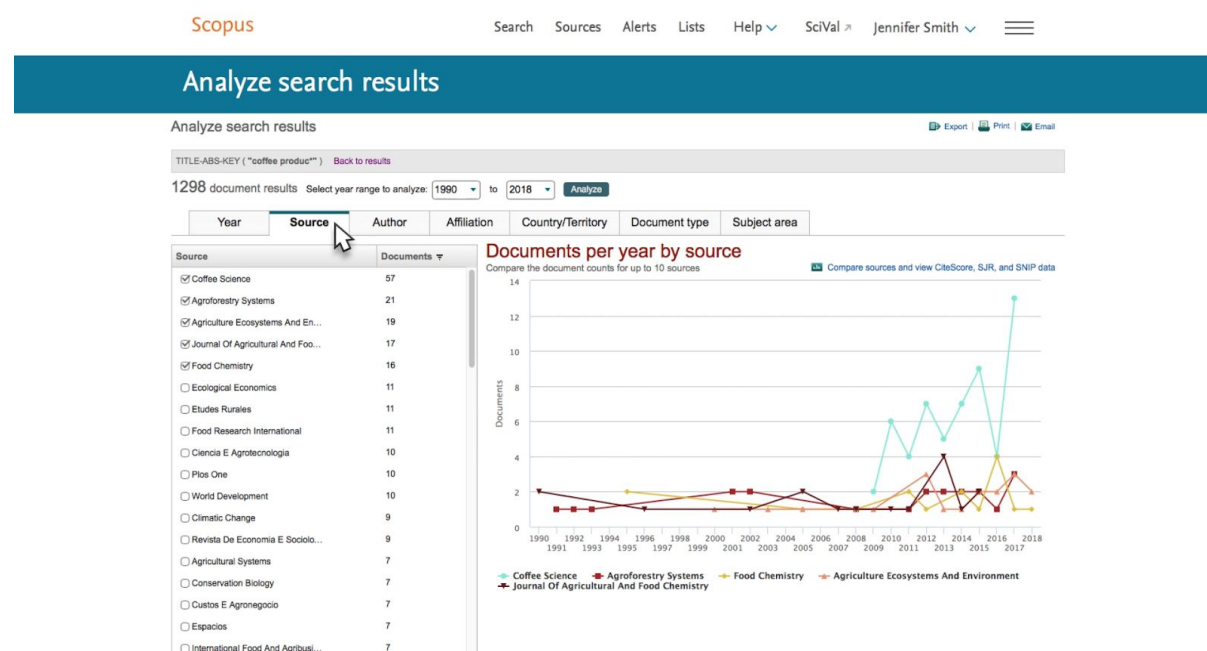
Εικόνα 44: Στιγμιότυπο οθόνης από Dynamic Queries Home Finder - παράδειγμα οπτικοποίησης με μπάρα όταν δεν είναι διαθέσιμα γεωγραφικά δεδομένα



Εικόνα 45: Στιγμιότυπο οθόνης από το Scopus - παράδειγμα διαγράμματος πίτας εγγράφων ανά θεματική περιοχή



Εικόνα 46: Στιγμιότυπο οθόνης από το Scopus - παράδειγμα διαγράμματος εγγράφων ανά είδος



Εικόνα 47: Στιγμιότυπο οθόνης από το Scopus - παράδειγμα διαγράμματος εγγράφων ανά χρόνο χωρισμένα κατά προέλευση

ΑΝΑΦΟΡΕΣ

- [1] A. E. Jinha, «Article 50 million: an estimate of the number of scholarly articles in existence», Learn. Publ., vol. 23, no. 3, pp. 258–263, Jul. 2010.
- [2] Nielsen J., «Usability Engineering», San Diego Acad. Press, 1993.
- [3] Simon HA., «The sciences of the artificial.», Cambridge, 1969.
- [4] McKim, Robert H., «Experiences in visual thinking.», 1972.
- [5] Archer B., «Design as a discipline. Design studies.», 1979 Jul
- [6] K. Petersen, C. Wohlin, and D. Baca, «The waterfall model in large-scale development», in Lecture Notes in Business Information Processing, 2009, διαθέσιμο: https://link.springer.com/chapter/10.1007/978-3-642-02152-7_29
- [7] Hix D, Hartson HR. «Developing user interfaces: Ensuring usability through product and process.», New York, J. Wiley & Sons, 1993.
- [8] Boehm, Barry W., «A Spiral Model of Software Development and Enhancement», IEEE Computer, 1988.
- [9] Booch G. «Object-oriented Analysis and Design With Applications Sec.», Redwood City t Benjamin/Cummings, 1994.
- [10] Lewis C, Rieman J. «Task-centered user interface design. A practical introduction.», 1993, διαθέσιμο: <http://web.cs.dal.ca/~jamie/TCUID/tcuid.pdf.gz>
- [11] J. Nielsen and Y. Kagamihara, «The regulation of presynaptic inhibition during co - contraction of antagonistic muscles in man.», J. Physiol., 1993.
- [12] Holleran PA., «A methodological note on pitfalls in usability testing. Behaviour & Information Technology.», 1991.
- [13] J. NIELSEN, «The Usability Engineering Lifecycle», in Usability Engineering, Elsevier, 1993, pp. 71–114.
- [14] Cairns P. και Cox A.L., «Research Methods for Human–Computer Interaction.» 2008.
- [15] Michael Friendly, Chun-houh Chen, Wolfgang Karl Härdle και Antony Unwin, «A brief history of data visualization», Handbook of Data Visualization, Βερολίνο 2008.
- [16] Manovich, «What is visualization?», The Journal of the Initiative for Digital Humanities, 2010, διαθέσιμο: <https://journals.tdl.org/paj/index.php/paj/article/view/19/58>
- [17] Friendly, «A brief history of data visualization», σ. 14, 2001.
- [18] Friendly και Daniel J. Denis, «The roots and branches of statistical graphics», Journal de la Société Française de Statistique, 141/4, σ. 51–60, 2001.
- [19] B. Shneiderman, «Dynamic queries for visual information seeking», in IEEE Software, vol. 11, no. 6, pp. 70–77, Nov. 1994, διαθέσιμο: <https://www.cs.umd.edu/~ben/papers/Shneiderman1994Dynamic.pdf>
- [20] Madian Khabisa, C. Lee Giles, «The Number of Scholarly Documents on the Public Web», PLOS ONE, 2014, διαθέσιμο: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0093949>
- [21] «BibTeX Entry and Field Types» Ιαν. 2020, <http://bib-it.sourceforge.net/help/fieldsAndEntryTypes.php> [Προσπελάστηκε 26/1/2020]
- [22] «Gapminder about: Our mission» Αυγ. 2020, <https://www.gapminder.org/about-gapminder/our-mission/> [Προσπελάστηκε 30/7/2020]
- [23] "Gapminder about: History" Αυγ. 2020, <https://www.gapminder.org/about-gapminder/history/> [Προσπελάστηκε 30/7/2020]
- [24] Roberts, Sam, «Hans Rosling, Swedish Doctor and Pop-Star Statistician, Dies at 68», The New York Times, ISSN 0362-4331,

www.nytimes.com/2017/02/09/world/europe/hans-rosling-dead-statistician.html [Προσπελάστηκε 30/7/2020]

[25] «Search Tips: Content Coverage», Google Scholar, [Προσπελάστηκε 30/7/2020]

[26] Gusenbauer, Michael, «Google Scholar to overshadow them all? Comparing the sizes of 12 academic search engines and bibliographic databases», *Scientometrics*, 2019, διαθέσιμο: <https://epub.jku.at/urn:nbn:at:at-ubl:3-289>

[27] Orduña-Malea, E., Ayllón, J. M., Martín-Martín, A., & Delgado López-Cózar, E., «Methods for estimating the size of Google Scholar. *Scientometrics*», 2015, διαθέσιμο: <https://arxiv.org/ftp/arxiv/papers/1506/1506.03009.pdf>

[28] Νικόλαος Αβούρης, Χρήστος Κατσάνος, Νικόλαος Τσέλιος, Κωνσταντίνος Μουστάκας, «Εισαγωγή στην Αλληλεπίδραση Ανθρώπου-Υπολογιστή», Εκδόσεις Πανεπιστημίου Πατρών, 2015, σελ. 220.

[29] Spence Robert, «Information Visualization: Design for Interaction (2nd ed.)», Prentice Hall, 2007.

[30] Efrossini Gkountava, «Οπτικοποίηση δεδομένων (data visualization)», 2013, διαθέσιμο: https://www.researchgate.net/publication/236199811_Optikopoiese_dedomenon_data_visualization