



**ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ:
«ΔΙΑΧΕΙΡΙΣΗ ΠΛΗΡΟΦΟΡΙΩΝ ΣΕ ΒΙΒΛΙΟΘΗΚΕΣ, ΑΡΧΕΙΑ, ΜΟΥΣΕΙΑ»**

**ΤΜΗΜΑ ΑΡΧΕΙΟΝΟΜΙΑΣ, ΒΙΒΛΙΟΘΗΚΟΝΟΜΙΑΣ ΚΑΙ ΣΥΣΤΗΜΑΤΩΝ ΠΛΗΡΟΦΟΡΗΣΗΣ
ΣΧΟΛΗ ΔΙΟΙΚΗΤΙΚΩΝ, ΟΙΚΟΝΟΜΙΚΩΝ ΚΑΙ ΚΟΙΝΩΝΙΚΩΝ ΕΠΙΣΤΗΜΩΝ**

**DEPARTMENT OF ARCHIVAL, LIBRARY AND INFORMATION STUDIES
SCHOOL OF MANAGEMENT, ECONOMICS AND SOCIAL SCIENCES**

Διπλωματική Εργασία

**Αξιολόγηση προτύπων μεταδεδομένων ψηφιακών
αποθετηρίων**

Συγγραφέας

Μαρίλια Αποστολάκη (ΑΜ: 186682022)

Επιβλέπων: Αλέξανδρος Κουλούρης

Αθήνα, Οκτώβριος 2020

Ευχαριστίες – Αφιερώσεις

Ολοκληρώνοντας την παρούσα διπλωματική εργασία θα ήθελα να ευχαριστήσω τον καθηγητή κ. Κουλούρη Αλέξανδρο του Τμήματος Αρχειονομίας, Βιβλιοθηκονομίας και Συστημάτων Πληροφόρησης του Πανεπιστημίου Δυτικής Αττικής αφενός για την ευκαιρία που μου έδωσε να ασχοληθώ με το συγκεκριμένο θέμα και αφετέρου για την πολύτιμη καθοδήγηση και βοήθεια που μου προσέφερε καθ' όλη τη διάρκεια εκπόνησης της διπλωματικής αυτής εργασίας. Η εμπιστοσύνη και η υπομονή που έδειξε στο πρόσωπό μου ήταν πολύ σημαντικές για εμένα. Κλείνοντας, ένα τεράστιο ευχαριστώ στην αδελφή μου Ελεάννα και τον φίλο μου Κώστα για τη βοήθεια και τη στήριξη τους καθώς και σε όλη την οικογένειά μου, στην οποία και αφιερώνω την εργασία αυτή, για την αγάπη, τη στήριξη και τη βοήθεια που μου προσφέρει με κάθε δυνατό τρόπο και ειδικότερα στα παιδάκια μου Αννούλα και Ανδρέα.

14 Οκτωβρίου 2020

Μαρίλια Αποστολάκη

Περίληψη στα ελληνικά

Είναι γεγονός ότι η διενέργεια έρευνας, όπως επίσης η δημοσιοποίηση αυτής και των αποτελεσμάτων αυτής στην επιστημονική κοινότητα οδεύει στην εξ ολοκλήρου πραγματοποίησή της μέσω του διαδικτύου. Σε αυτήν την απαιτητική περίοδο οι βιβλιοθήκες και άλλοι ερευνητικοί φορείς αναπτύσσουν υπηρεσίες, όπως τα ψηφιακά αποθετήρια (ΨΑ), για να συμβάλλουν στην αύξηση της προβολής και αναγνώρισης του έργου των δημιουργών και της ερευνητικής δραστηριότητας των φορέων τους. Ωστόσο, έρευνες δείχνουν ότι η ευρετηρίαση του περιεχομένου των ΨΑ δεν πραγματοποιείται ικανοποιητικά από τις μηχανές αναζήτησης. Εκτιμάται ότι το πρόβλημα αυτό οφείλεται μεταξύ άλλων στην ποιότητα των μεταδεδομένων, όπως και στο πρότυπο μεταδεδομένων που έχει υιοθετηθεί από το εκάστοτε ΨΑ.

Αυτή η εργασία στοχεύει να ελέγξει αν και κατά πόσο το πρότυπο μεταδεδομένων (Dublin Core), που χρησιμοποιείται από την πλειοψηφία των ΨΑ, επηρεάζει την ανακάλυψη του περιεχομένου τους από τις μηχανές αναζήτησης όπως το Google Scholar (GS). Επίσης, αν επιτυγχάνεται μεγαλύτερη ευρετηρίαση του περιεχομένου των ΨΑ που περιγράφεται με το ίδιο πρότυπο, αν ληφθούν κάποια μέτρα από πλευράς ΨΑ. Τέλος, εξετάζεται αν η παροχή με έγκυρο τρόπο συγκεκριμένων στοιχείων μεταδεδομένων επηρεάζει την ανακάλυψη κι ευρετηρίαση των πόρων των ΨΑ από τις δυναμικές μηχανές αναζήτησης όπως το GS.

Για αυτή τη διπλωματική εργασία διεξήχθη η παρούσα ποσοτική έρευνα, με συμμετοχή 407 ΨΑ, 57 ελληνικών και 350 του εξωτερικού. Δημιουργήθηκαν δύο ερωτηματολόγια (σε ελληνικά και αγγλικά), τα οποία διανεμήθηκαν ηλεκτρονικά. Απαντήθηκαν 98 ερωτηματολόγια (ποσοστό 25%). Τα ευρήματα της έρευνας δεν επαληθεύουν στατιστική συσχέτιση του προτύπου μεταδεδομένων με το ποσοστό ευρετηρίασης, ούτε τη συσχέτιση των βελτιωτικών ενεργειών που προτείνονται με το ποσοστό ευρετηρίασης των πόρων των ΨΑ. Ωστόσο, επιβεβαιώνεται η στατιστική συσχέτιση μεταξύ των δύο μεταβλητών (Συγγραφέας και Λέξη-Κλειδί) με το ποσοστό ευρετηρίασης των πόρων ενός ΨΑ από το GS. Επομένως η διάθεση με έγκυρο τρόπο και ενδεχομένως με τη χρήση προτύπων όπως το Dublin Core τουλάχιστον συγκεκριμένων στοιχείων μεταδεδομένων, αν όχι όλων, θεωρείται σημαντική για την ανακάλυψη των ψηφιακών πόρων από τις δυναμικές μηχανές αναζήτησης όπως το GS.

Λέξεις Κλειδιά: Dublin Core, Μεταδεδομένα, Μελετητής Google, Ψηφιακά Αποθετήρια, Ποσοστό Ευρετηρίασης, Ορατότητα, Πρότυπα Μεταδεδομένων, Τυποποίηση, Ανοικτή Πρόσβαση, Ανακάλυψη, Ανάκτηση

Περίληψη στα αγγλικά

It is a fact that conducting research nowadays, as well as publishing it and its results to the scientific community, means that it is carried out fully on the internet. That is why libraries and other research institutions are developing services, such as Digital Repositories (DR), in order to have their contents indexed by search engines, to achieve the visibility, recognition and promotion of the work of the authors and the research activity of their institutions. However, research shows that the content of DR is not indexed in a satisfactory way by search engines such as Google Scholar (GS).

It is estimated though that this problem is due, among other things, to the quality of metadata, as well as to the metadata standard adopted by the majority of DR. This work aims to test whether and to what extent the metadata standard (Dublin Core) which used by the majority of DR affects the discovery of their content by search engines such as GS. Also, if a higher index of the content of the DR described by the same standard is achieved, if specific action is taken by the DR. Furthermore, whether the valid provision of specific metadata elements affects the discovery and indexing of DR's resources by dynamic search engines such as GS.

For the purpose of this master thesis a quantitative research was carried out, with a participation of 407 DR (selected with specific criteria), 57 Greek and 350 from abroad. Two questionnaires were created (in Greek and English), which were distributed electronically. 98 questionnaires were answered (25%). The research findings do not verify a statistical correlation of the metadata standard with the indexing rate, nor the correlation of the proposed improvement actions with the indexing rate of the resources of the DR. However, a statistical correlation is confirmed between the two variables (Author and Keyword) with the GS's resource index rate. Therefore, validating, possibly using a metadata standard such as Dublin Core, even for specific metadata elements, if not for all, is considered important for the discovery of digital resources by dynamic search engines such as GS.

Keywords: Dublin Core, Metadata, Google Scholar, Repositories, Indexing ratio, Visibility, Metadata Standards, Open access, Standardization, Discovery, Retrieval

Πίνακας περιεχομένων

ΕΥΧΑΡΙΣΤΙΕΣ – ΑΦΙΕΡΩΣΕΙΣ	III
ΠΕΡΙΛΗΨΗ ΣΤΑ ΕΛΛΗΝΙΚΑ	IV
ΠΕΡΙΛΗΨΗ ΣΤΑ ΑΓΓΛΙΚΑ	VI
ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ	VIII
ΠΙΝΑΚΑΣ ΣΧΗΜΑΤΩΝ	X
ΠΙΝΑΚΑΣ ΓΡΑΦΗΜΑΤΩΝ	XI
ΠΙΝΑΚΑΣ ΠΙΝΑΚΩΝ	XII
ΚΕΦΑΛΑΙΟ 1. ΕΙΣΑΓΩΓΗ	1
1.1 ΠΛΑΙΣΙΟ, ΣΚΟΠΟΣ ΚΑΙ ΣΤΟΧΟΙ ΤΗΣ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ	2
1.2 ΕΡΕΥΝΗΤΙΚΕΣ ΥΠΟΘΕΣΕΙΣ	3
1.3 ΜΕΘΟΔΟΛΟΓΙΑ	3
1.4 ΠΕΡΙΟΡΙΣΜΟΙ	4
1.5 ΟΡΙΣΜΟΙ	5
1.5.1 <i>Μεταδεδομένα</i>	5
1.5.2 <i>Διαλειτουργικότητα</i>	7
1.5.3 <i>Πρότυπα και σχήματα μεταδεδομένων</i>	8
1.5.4 <i>Ψηφιακά αποθετήρια</i>	9
1.5.5 <i>Google Scholar</i>	10
1.5.6 <i>Ανοικτή Πρόσβαση</i>	10
1.5.7 <i>Βιβλιομετρικοί δείκτες – H-index και i10-index</i>	10
1.6 ΟΡΓΑΝΩΣΗ ΚΕΦΑΛΑΙΩΝ	11
ΚΕΦΑΛΑΙΟ 2. ΘΕΩΡΗΤΙΚΟ ΜΕΡΟΣ – ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΕΡΕΥΝΑ – ΣΧΕΤΙΚΕΣ ΠΡΟΣΠΑΘΕΙΕΣ	13
2.1 ΘΕΩΡΗΤΙΚΟ ΜΕΡΟΣ – ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΕΡΕΥΝΑ	13
2.1.1 <i>Dublin Core - Πλεονεκτήματα και μειονεκτήματα</i>	16
2.1.2 <i>Google Scholar – Πλεονεκτήματα και μειονεκτήματα</i>	21
2.1.3 <i>Google Scholar – Μεταδεδομένα</i>	25
2.2 ΣΧΕΤΙΚΕΣ ΠΡΟΣΠΑΘΕΙΕΣ – ΕΡΕΥΝΕΣ	28
ΚΕΦΑΛΑΙΟ 3. ΜΕΘΟΔΟΛΟΓΙΑ	31

3.1	ΣΤΟΧΟΣ ΤΗΣ ΈΡΕΥΝΑΣ ΚΑΙ ΜΕΘΟΔΟΛΟΓΙΑ.....	31
3.2	ΣΥΓΚΡΟΤΗΣΗ ΔΕΙΓΜΑΤΟΣ & ΣΥΛΛΟΓΗ ΔΕΔΟΜΕΝΩΝ.....	31
3.3	[ΜΕΤΡΗΣΗ ΠΑΡΑΓΟΝΤΩΝ &] ΚΑΤΑΣΚΕΥΗ ΕΡΕΥΝΗΤΙΚΟΥ ΕΡΓΑΛΕΙΟΥ.....	33
3.4	ΜΕΘΟΔΟΙ ΣΤΑΤΙΣΤΙΚΗΣ ΑΝΑΛΥΣΗΣ	34
ΚΕΦΑΛΑΙΟ 4. ΑΠΟΤΕΛΕΣΜΑΤΑ – ΕΥΡΗΜΑΤΑ / ΕΠΙΤΕΥΓΜΑΤΑ		38
4.1	ΑΝΑΛΥΤΙΚΗ ΠΑΡΟΥΣΙΑΣΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ.....	38
4.2	ΚΥΡΙΟΤΕΡΑ ΕΥΡΗΜΑΤΑ / ΑΠΟΤΕΛΕΣΜΑΤΑ.....	58
ΚΕΦΑΛΑΙΟ 5. ΜΕΛΛΟΝΤΙΚΕΣ ΕΠΕΚΤΑΣΕΙΣ		67
5.1	ΑΝΑΚΕΦΑΛΑΙΩΣΗ.....	67
5.2	ΣΥΖΗΤΗΣΗ / ΣΥΜΠΕΡΑΣΜΑΤΑ	69
5.3	ΑΞΙΟΠΟΙΗΣΗ / ΠΡΑΚΤΙΚΕΣ ΠΡΟΕΚΤΑΣΕΙΣ ΤΗΣ ΕΡΕΥΝΑΣ	72
5.4	ΜΕΛΛΟΝΤΙΚΕΣ ΕΠΕΚΤΑΣΕΙΣ / ΠΡΑΚΤΙΚΕΣ ΠΡΟΕΚΤΑΣΕΙΣ ΤΗΣ ΈΡΕΥΝΑΣ.....	74
ΒΙΒΛΙΟΓΡΑΦΙΚΕΣ ΑΝΑΦΟΡΕΣ.....		75
ΠΑΡΑΡΤΗΜΑ Α'– ΣΧΕΔΙΟ ΕΡΩΤΗΜΑΤΟΛΟΓΙΟΥ (ΕΛΛΗΝΙΚΗ ΕΚΔ.).....		82
ΠΑΡΑΡΤΗΜΑ Β'– ΣΧΕΔΙΟ ΕΡΩΤΗΜΑΤΟΛΟΓΙΟΥ (ΑΓΓΛΙΚΗ ΕΚΔ.).....		84
ΠΑΡΑΡΤΗΜΑ Γ' - ΠΙΝΑΚΑΣ ΣΥΜΜΕΤΕΧΟΝΤΩΝ		86

Πίνακας Σχημάτων

Εικόνα 1. Φύλλο εργασίας στο περιβάλλον SPSS	34
Εικόνα 2. Παραμετροποίηση μεταβλητών.....	35
Εικόνα 3. Πίνακας Συχνότητων.....	36
Εικόνα 4. Διαδικασία Crosstabs	36

Πίνακας Γραφημάτων

Γράφημα 1. Ηλικιακή ομάδα.....	39
Γράφημα 2. Φύλο.....	39
Γράφημα 3. Ιδιότητα/θέση στο Φορέα.....	40
Γράφημα 4. Χρησιμοποιείτε την πιο ενημερωμένη (τελευταία) έκδοση του λογισμικού του ΨΑ σας ή κάποια παλαιότερη; * Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ΨΑ σας από τις μηχανές αναζήτησης όπως το GS;.....	44
Γράφημα 5. Θεωρείτε σημαντική την παρακολούθηση και αφομοίωση των εξελίξεων και των ενημερώσεων του προτύπου μεταδεδομένων που χρησιμοποιείται από το ΨΑ σας;...	46
Γράφημα 6. Θεωρείτε ότι η τυποποίηση των μεταδεδομένων είναι σημαντική για την ανάκτηση των ψηφιακών πόρων;.....	48
Γράφημα 7. Θεωρείτε ότι η ανακάλυψη των ψηφιακών πόρων του αποθετηρίου σας εξυπηρετείται με επάρκεια από τις μηχανές αναζήτησης όπως το Google Scholar;	49
Γράφημα 8. Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ΨΑ σας από τις μηχανές αναζήτησης όπως το GS;.....	51
Γράφημα 9. Θεωρείτε ότι το υλικό του αποθετηρίου σας δεν χρειάζεται να ευρετηριάζεται από τις μηχανές αναζήτησης όπως το Google Scholar;	52
Γράφημα 10. Θεωρείτε ότι το πρότυπο μεταδεδομένων παίζει σημαντικό ρόλο στην ευρετηρίαση των ψηφιακών πόρων του αποθετηρίου σας από τις μηχανές ερευνητικής αναζήτησης όπως το Google Scholar	54
Γράφημα 11. Σύμφωνα με ποιο πρότυπο μεταδεδομένων περιγράφονται οι ψηφιακοί πόροι στο αποθετήριό σας * Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ψηφιακού αποθετηρίου σας από τις μηχανές αναζήτησης όπως το Google Scholar	60
Γράφημα 12. Ενέργειες στις οποίες έχετε προχωρήσει για την αύξηση της ευρετηρίασης του περιεχομένου του ΨΑ σας από τις μηχανές αναζήτησης όπως το GS * Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ΨΑ σας από τις μηχανές αναζήτησης όπως το GS	62
Γράφημα 13. Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ψηφιακού αποθετηρίου σας από τις μηχανές αναζήτησης όπως το Google Scholar * Συγγραφέας.....	65
Γράφημα 14. Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ψηφιακού αποθετηρίου σας από τις μηχανές αναζήτησης όπως το Google Scholar * Λέξεις-Κλειδιά ..	66

Πίνακας Πινάκων

Πίνακας 1. Συντομογραφίες	xv
Πίνακας 2. Τι είδους ψηφιακό αποθετήριο διαθέτει ο φορέας σας;	41
Πίνακας 3. Είδη υλικού που περιλαμβάνονται στο ψηφιακό αποθετήριο	41
Πίνακας 4. Πως γίνεται η απόθεση των ψηφιακών πόρων στο αποθετήριό σας;	42
Πίνακας 5. Ποιο λογισμικό έχετε χρησιμοποιήσει για την υλοποίηση του ψηφιακού αποθετηρίου σας;	43
Πίνακας 6. Σύμφωνα με ποιο πρότυπο μεταδεδομένων περιγράφονται οι ψηφιακοί πόροι του αποθετηρίου σας;	45
Πίνακας 7. Απαιτούμενα στοιχεία μεταδεδομένων που καταχωρίζονται στο ΨΑ	45
Πίνακας 8. Θεωρείτε σημαντική την παρακολούθηση και αφομοίωση των εξελίξεων και των ενημερώσεων (νέες εκδόσεις) του προτύπου μεταδεδομένων που χρησιμοποιείται από το ΨΑ σας;	46
Πίνακας 9. Θεωρείτε ότι η τυποποίηση των μεταδεδομένων είναι σημαντική για την ανάκτηση των ψηφιακών πόρων;	47
Πίνακας 10. Ποια δήλωση από τις ακόλουθες περιγράφει καλύτερα τον τρόπο της θεματικής ευρετηρίασης του ψηφιακού αποθετηρίου σας;	48
Πίνακας 11. Θεωρείτε ότι η ανακάλυψη των ψηφιακών πόρων του αποθετηρίου σας εξυπηρετείται με επάρκεια από τις μηχανές αναζήτησης όπως το Google Scholar;	49
Πίνακας 12. Είναι το ποσοστό ευρετηρίασης του περιεχομένου του ΨΑ σας από τις μηχανές αναζήτησης όπως το GS;	50
Πίνακας 13. Είστε ικανοποιημένοι από τον βαθμό ευρετηρίασης του περιεχομένου του αποθετηρίου σας από το Google Scholar;	51
Πίνακας 14. Έχετε ελέγξει τον βαθμό ικανοποίησης σε σχέση με την ορατότητα, τις προβολές και τις λήψεις του περιεχομένου του αποθετηρίου σας από μηχανές αναζήτησης όπως το Google Scholar;	52
Πίνακας 15. Θεωρείτε ότι το υλικό του αποθετηρίου σας δεν χρειάζεται να ευρετηριάζεται από τις μηχανές αναζήτησης όπως το Google Scholar	52
Πίνακας 16. Θεωρείτε ότι το πρότυπο μεταδεδομένων παίζει σημαντικό ρόλο στην ευρετηρίαση των ψηφιακών πόρων του αποθετηρίου σας από τις μηχανές ερευνητικής αναζήτησης όπως το Google Scholar	53

Πίνακας 17. Έχετε προχωρήσει σε κατάλληλες ενέργειες για την επίτευξη αύξησης της ευρετηρίασης του περιεχομένου του ψηφιακού αποθετηρίου σας από μηχανές αναζήτησης όπως το Google Scholar;	54
Πίνακας 18. Παρακαλώ κατονομάστε τις ενέργειες στις οποίες έχετε προχωρήσει για την αύξηση της ευρετηρίασης του περιεχομένου του ψηφιακού σας αποθετηρίου από τις μηχανές αναζήτησης όπως το Google Scholar.....	55
Πίνακας 19. Ποια προβλήματα θεωρείτε ότι υφίστανται και επηρεάζουν την ευρετηρίαση του περιεχομένου του αποθετηρίου σας από τις μηχανές αναζήτησης όπως το Google Scholar;.....	57
Πίνακας 20. Σύμφωνα με ποιο πρότυπο μεταδεδομένων περιγράφονται οι ψηφιακοί πόροι στο αποθετήριό σας * Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ψηφιακού αποθετηρίου σας από τις μηχανές αναζήτησης όπως το Google Scholar; Crosstabulation.....	59
Πίνακας 21. Σύμφωνα με ποιο πρότυπο μεταδεδομένων περιγράφονται οι ψηφιακοί πόροι στο αποθετήριό σας * Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ψηφιακού αποθετηρίου σας από τις μηχανές αναζήτησης όπως το Google Scholar; Pearson's Chi-Square Tests.....	60
Πίνακας 22. Ενέργειες στις οποίες έχετε προχωρήσει για την αύξηση της ευρετηρίασης του περιεχομένου του ΨΑ σας από τις μηχανές αναζήτησης όπως το GS * Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ΨΑ σας από τις μηχανές αναζήτησης όπως το GS; Crosstabulation.....	61
Πίνακας 23. Ενέργειες στις οποίες έχετε προχωρήσει για την αύξηση της ευρετηρίασης του περιεχομένου του ΨΑ σας από τις μηχανές αναζήτησης όπως το GS * Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ΨΑ σας από τις μηχανές αναζήτησης όπως το GS; Pearson's Chi-Square Tests.....	62
Πίνακας 24. Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ψηφιακού αποθετηρίου σας από τις μηχανές αναζήτησης όπως το Google Scholar * Συγγραφέας; Crosstabulation.....	64
Πίνακας 25. Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ψηφιακού αποθετηρίου σας από τις μηχανές αναζήτησης όπως το Google Scholar * Συγγραφέας; Pearson's Chi-Square Tests.....	64
Πίνακας 26. Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ψηφιακού αποθετηρίου σας από τις μηχανές αναζήτησης όπως το Google Scholar * Λέξεις-Κλειδιά; Crosstabulation.....	65

Πίνακας 27. Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ψηφιακού αποθετηρίου σας από τις μηχανές αναζήτησης όπως το Google Scholar * Λέξεις-Κλειδιά; Pearson's Chi-Square Tests.....	66
Πίνακας 28. Πίνακας συμμετεχόντων	86

Πίνακας 1. Συντομογραφίες

DC	Dublin Core
GS	Google Scholar
GSACT	Google Scholar Author Citation Tracker
GSC	Google Scholar Citations
GSM	Google Scholar Metrics for Publications
IR	Institutional Repository
RDF	Resource Description Framework
XML	eXtensible Markup Language
ΑΠ	Ανοικτή Πρόσβαση
ΙΑ	Ιδρυματικά Αποθετήρια
ΠΜΣ	Πρόγραμμα Μεταπτυχιακών Σπουδών
ΨΑ	Ψηφιακά Αποθετήρια

Κεφάλαιο 1. Εισαγωγή

Οι ταχύτατα εξελισσόμενες Τεχνολογίες των Επικοινωνιών και του Παγκόσμιου Ιστού έχουν προκαλέσει πολλές αλλαγές στον τρόπο δημιουργίας και παροχής της επιστημονικής γνώσης. Οι επιστήμονες χρησιμοποιούν και βασίζονται όλο και πιο πολύ στη χρήση ηλεκτρονικών μέσων για την πραγματοποίηση των ερευνών τους, όπως και για την δημοσιοποίηση των αποτελεσμάτων αυτών. Υπολογίζεται ότι σύντομα η έρευνα όπως και η κοινοποίησή της στην ερευνητική κοινότητα θα γίνεται κυρίως μέσω του διαδικτύου, με αποτέλεσμα την επίτευξη πολλών ωφελειών σε επίπεδο ανάπτυξης συνεργασιών, αναγνωρισιμότητας, διαχείρισης εργασιών, προσέλκυση και εξασφάλιση χρηματοδότησης της ερευνητικής εργασίας κ.λπ. Σε ένα τέτοιο περιβάλλον, οι βιβλιοθήκες όπως και άλλοι φορείς έχουν αναπτύξει υπηρεσίες, όπως τα Ψηφιακά Αποθετήρια, που συμβάλλουν στην αύξηση της προβολής και της αναγνώρισης του έργου των δημιουργών. Με αυτόν τον τρόπο υποστηρίζεται παράλληλα το κίνημα της Ανοικτής Πρόσβασης (ΑΠ), καθώς παρέχεται ελεύθερη πρόσβαση σε πληθώρα επιστημονικών εργασιών, ενισχύοντας τη δυνατότητα ανακάλυψης και ανάκτησης της πληροφορίας. Αντίστοιχα, έχουν αναπτυχθεί μηχανές ερευνητικής αναζήτησης όπως το Google Scholar (GS), οι οποίες λειτουργούν ως συσσωρευτές επιστημονικής γνώσης. Οι μηχανές αυτές προσφέρουν ένα ευρύ φάσμα πρόσβασης σε πολλά αποθετήρια, απαλλάσσοντας έτσι τους ίδιους τους χρήστες από την μεμονωμένη αναζήτηση σε κάθε αποθετήριο χωριστά. Η αυξημένη πρόσβαση και η ορατότητα οδηγεί σε απόκτηση περισσότερων ετεροαναφορών, οι οποίες παραμένουν ως η αδιαμφισβήτητη απόδειξη της ακαδημαϊκής/επιστημονικής αναγνώρισης και προαγωγής (Shanks & Arlitsch, 2016). Παρόλο που η αυξανόμενη έμφαση σε εναλλακτικές μετρικές (altmetrics) μπορεί να μεταβάλλει αργά τον ορισμό της ακαδημαϊκής απήχησης, οι τρέχουσες απαιτήσεις εξακολουθούν να επικεντρώνονται στη χρηματοδότηση της έρευνας και οι αναφορές (citations) ως μετρήσιμα στοιχεία παίζουν πολύ σημαντικό ρόλο στον αντίκτυπο αυτό (Ebadi & Schiffauerova, 2015). Η χρηματοδότηση των ερευνητών και οι αναφορές που λαμβάνουν οι συγγραφείς επιστημονικών εργασιών οδηγούν επίσης σε αύξηση της καλής φήμης των πανεπιστημίων και άνοδό τους στη διεθνή κατάταξη. Επομένως, επηρεάζεται θετικά ο αριθμός εγγραφής νέων φοιτητών στα πανεπιστήμια και κατά συνέπεια και η οικονομική υγεία και ευημερία των ιδρυμάτων αυτών. Τα ψηφιακά αποθετήρια κάθε είδους αντιμετωπίζουν λοιπόν μια κοινή πρόκληση, πρέπει το περιεχόμενό τους να ανακαλύπτεται και να ανακτάται από τους ενδιαφερόμενους χρήστες,

ανάμεσα στον τεράστιο όγκο πληροφοριών που διατίθεται στο Διαδίκτυο. Αυτό σημαίνει ότι τα τεκμήρια ενός αποθετηρίου πρέπει να συμπεριλαμβάνονται στα ευρετήρια των μεγάλων και δυναμικών μηχανών αναζήτησης, εκεί όπου η μεγάλη πλειοψηφία των χρηστών ξεκινά την έρευνά του (Arlitsch & O'Brien, 2012). Παρ' όλα αυτά, έρευνες έχουν δείξει ότι η ευρετηρίαση του περιεχομένου των ψηφιακών αποθετηρίων δεν πραγματοποιείται σε ικανοποιητικό βαθμό από τις μηχανές ερευνητικής αναζήτησης. Εκτιμάται ότι το πρόβλημα αυτό οφείλεται μεταξύ άλλων στην ποιότητα των μεταδεδομένων, όπως και στο πρότυπο μεταδεδομένων που έχει υιοθετηθεί από το εκάστοτε ψηφιακό αποθετήριο, το οποίο πιθανώς δεν πληροί τις απαιτήσεις σε μεταδεδομένα που έχουν τεθεί από τις μηχανές ερευνητικής αναζήτησης (Arlitsch & O'Brien, 2012).

1.1 Πλαίσιο, σκοπός και στόχοι της διπλωματικής εργασίας

Ο στόχος αυτής της διπλωματικής εργασίας είναι η αξιολόγηση των προτύπων μεταδεδομένων που χρησιμοποιούνται από τα ελληνικά και διεθνή αποθετήρια, όπως επίσης των προτύπων μεταδεδομένων που χρησιμοποιούνται από μηχανές ερευνητικής αναζήτησης, όπως το GS. Κύριος στόχος των μεταδεδομένων είναι να παράγουν πληροφορίες σε δομημένη μορφή, επιτρέποντας στις μηχανές καθώς και στους ανθρώπους να τις εντοπίζουν και να τις χρησιμοποιούν. Επομένως, η σημασία και η αξία τους είναι πολύ σημαντική, καθώς ο όγκος της πληροφορίας που συγκεντρώνεται καθημερινά με τη χρήση του Διαδικτύου και του Παγκόσμιου Ιστού είναι τεράστιος. Θεωρείται σημαντικό λοιπόν να εντοπιστεί το ποσοστό ευρετηρίασης των ψηφιακών τεκμηρίων που περιγράφονται για παράδειγμα με το Dublin Core (DC) σε δομημένα ψηφιακά αποθετήρια, με ψηφιακά τεκμήρια όπως αυτά περιγράφονται από μηχανές αναζήτησης όπως το GS κάνοντας χρήση άλλων προτύπων μεταδεδομένων. Σκοπός είναι να αναδειχθούν τα πλεονεκτήματα και τα μειονεκτήματα που εντοπίζονται στους δύο αυτούς τρόπους περιγραφής κι οργάνωσης ψηφιακών τεκμηρίων. Μέσα από τα αποτελέσματά της η διπλωματική αυτή φιλοδοξεί να προτείνει την καλύτερη δυνατή λύση, αναφορικά με τα πεδία μεταδεδομένων η συμπλήρωση των οποίων θεωρείται πολύ σημαντική και που πρέπει να χρησιμοποιούνται με συνέπεια από τα ψηφιακά αποθετήρια γνώσης. Τα βασικά οφέλη που προκύπτουν από αυτή τη διαδικασία, εκτός της ορθής περιγραφής των τεκμηρίων, σχετίζονται άμεσα με την αύξηση των δεικτών ευρετηρίασης των ψηφιακών πόρων από τις σύγχρονες και δυναμικές μηχανές ερευνητικής αναζήτησης όπως το GS. Με τον τρόπο αυτό επιτυγχάνεται η μεγαλύτερη δυνατή ορατότητα (visibility) κι ανάκτηση

(retrieval) των τεκμηρίων, η αύξηση του δείκτη απήχησης (impact factor) και η επίτευξη μεγαλύτερης αναγνώρισης των ερευνητών, μέσω της λήψης περισσότερων ετεροαναφορών από την ερευνητική κοινότητα στην οποία ανήκουν, η βελτίωση πολλές φορές της σειράς κατάταξης (ranking) των ίδιων των Πανεπιστημιακών/Ερευνητικών Φορέων που οι ερευνητές αντιπροσωπεύουν και τέλος η προώθηση της ΑΠ και της εξέλιξης της γνώσης.

1.2 Ερευνητικές υποθέσεις

Τα μεταδεδομένα στα ψηφιακά αποθετήρια παίζουν καθοριστικό ρόλο. Μπορεί όμως το πρότυπο μεταδεδομένων που χρησιμοποιεί ένα ψηφιακό αποθετήριο να επηρεάζει την ανακάλυψη των ψηφιακών πόρων που περιέχονται σε αυτό, από τις σύγχρονες και δυναμικές μηχανές ερευνητικής αναζήτησης όπως το GS;

Το Dublin Core είναι το πρότυπο και σχήμα συνάμα μεταδεδομένων που χρησιμοποιείται από την πλειοψηφία των ψηφιακών αποθετηρίων σε όλο τον κόσμο με τα πλεονεκτήματα και τα μειονεκτήματά του. Μπορούμε όμως να πετύχουμε μεγαλύτερη ευρετηρίαση και ορατότητα του ψηφιακού περιεχομένου που περιγράφεται με το συγκεκριμένο πρότυπο από τις μηχανές αναζήτησης όπως το GS;

Μήπως η διάθεση/παροχή με έγκυρο τρόπο συγκεκριμένων στοιχείων μεταδεδομένων επηρεάζει την ανακάλυψη και ευρετηρίαση των ψηφιακών πόρων των αποθετηρίων από τις δυναμικές μηχανές αναζήτησης όπως το GS;

1.3 Μεθοδολογία

Η μεθοδολογία έρευνας που ακολουθήθηκε σε αυτή τη διπλωματική εργασία είναι η ποσοτική έρευνα με αποστολή ερωτηματολογίων. Για τον σκοπό αυτό δημιουργήθηκαν δύο ερωτηματολόγια (σε ελληνική και αγγλική έκδοση), τα οποία απεστάλησαν στους υπεύθυνους/απασχολούμενους σε ψηφιακά αποθετήρια διαφόρων φορέων σε Ελλάδα και εξωτερικό. Τα ερωτηματολόγια αποτελούνταν από 23 ερωτήσεις, οι οποίες βασίστηκαν στη βιβλιογραφική επισκόπηση, προκειμένου να συσχετιστούν με τα ευρήματα της έρευνας. Τα ερωτηματολόγια απεστάλησαν σε ένα μέγεθος δείγματος (n) τετρακοσίων επτά (407) ψηφιακών αποθετηρίων σε Ελλάδα και εξωτερικό, γνωρίζοντας ότι ο συνολικός αριθμός (πληθυσμός) των ψηφιακών αποθετηρίων ξεπερνά παγκοσμίως τις τρεις χιλιάδες (3000). Η επιλογή τους έγινε τυχαία, επιλέγοντας από την ιστοσελίδα Ranking Web of Repositories (<https://repositories.webometrics.info/en/node/30>) 9th ed. April 2020 τριακόσια πενήντα (350) ψηφιακά αποθετήρια από το εξωτερικό και αντίστοιχα από τις ιστοσελίδες Open DOAR (Directory of Open Access Repositories) - Greece

(https://v2.sherpa.ac.uk/view/repository_by_country/Greece.html) και Open DOAR (Directory of Open Access Repositories) - Cyprus (https://v2.sherpa.ac.uk/view/repository_by_country/Cyprus.html) επιλέγοντας πενήντα επτά (57) ελληνικά και κυπριακά ψηφιακά αποθετήρια. Κατά την επιλογή των συμμετεχόντων στην έρευνα αφενός δόθηκε έμφαση στην ελληνική πραγματικότητα και αφετέρου λάβαμε υπόψη, για τα διεθνή κυρίως αποθετήρια, την δυναμική τους σε σχέση με τα τεκμήρια που ευρετηριάζονται στο Google Scholar (τον αριθμό κατάταξής τους), τη γεωγραφική τους θέση, τον αντίκτυπό τους αναφορικά με την επιστημονική και εκδοτική τους παραγωγή. Επίσης έγινε μία προσπάθεια επιλογής και ανάδειξης ψηφιακών αποθετηρίων που τυγχάνουν υψηλής, μέτριας και χαμηλής ευρετηρίασης στο Google Scholar, έτσι ώστε να διερευνηθούν οι πρακτικές που ακολουθούνται, το πρότυπο και το λογισμικό των αποθετηρίων αυτών και πως μπορεί να συσχετίζονται μεταξύ τους. Στην εν λόγω έρευνα συμμετείχαν με την απάντησή τους ενενήντα οκτώ (98) από το σύνολο των τετρακοσίων επτά (407) ΨΑ, εκ των οποίων τα είκοσι εννέα (29) αφορούν ελληνικά και κυπριακά ΨΑ και εξήντα εννέα (69) αφορούν ΨΑ του εξωτερικού. Για την στατιστική ανάλυση των δεδομένων χρησιμοποιήθηκε το λογισμικό στατιστικής ανάλυσης “IBM SPSS STATISTICS” version 20. Για την εκτίμηση των πιθανών σχέσεων μεταξύ των μεταβλητών της έρευνας, χρησιμοποιήθηκε ο συντελεστής Pearson Chi-square (χ^2).

1.4 Περιορισμοί

Η ευρεία χρήση των ηλεκτρονικών μέσων και του διαδικτύου γενικότερα τόσο στην πραγματοποίηση ερευνών, όσο και στην δημοσιοποίηση των αποτελεσμάτων αυτών στην ερευνητική κοινότητα παρουσιάζει ιδιαίτερο ενδιαφέρον τα τελευταία χρόνια. Για τον λόγο μεγάλη έμφαση έχει δοθεί από βιβλιοθήκες και ερευνητικούς οργανισμούς στην ευρετηρίαση του περιεχομένου των ψηφιακών αποθετηρίων τους από τις μηχανές επιστημονικής/ακαδημαϊκής αναζήτησης, με σκοπό την αύξηση της προβολής και αναγνώρισης της δραστηριότητάς τους – κυρίως εκδοτικής. Στη διπλωματική αυτή εργασία έγινε προσπάθεια εντοπισμού ενός αντιπροσωπευτικού αριθμού ελληνικών και διεθνών ψηφιακών αποθετηρίων για συμμετοχή τους στην έρευνά μας, έτσι ώστε να συλλέξουμε απαντήσεις σχετικές με τον τρόπο δόμησης και λειτουργίας των αποθετηρίων και να εισπράξουμε πιθανώς αποτελέσματα που σχετίζονται με καλύτερες πρακτικές κυρίως σε επίπεδο καταχώρησης μεταδεδομένων και των προτύπων αυτών. Ωστόσο, λόγω του μεγάλου αριθμού ψηφιακών αποθετηρίων ανά τον κόσμο, ο αριθμός των συμμετεχόντων είναι αρκετά περιορισμένος με μόλις 350 αποθετήρια του εξωτερικού και 57 ελληνικά και

κυπριακά. Στο σημείο αυτό είναι σημαντικό να σημειωθεί ότι η ελληνική εκπαιδευτική πραγματικότητα με τη συγχώνευση των Τεχνολογικών Εκπαιδευτικών Ιδρυμάτων με τα Ανώτατα Εκπαιδευτικά Ιδρύματα και τη δημιουργία ενιαίων ΨΑ για το συγχωνευμένο πλέον ίδρυμα, πιθανώς ευθύνεται για το γεγονός ότι αρκετά από τα e-mails συμμετοχής στην έρευνά μας που απεστάλησαν σε ελληνικά ΨΑ δεν παραδόθηκαν ποτέ στους παραλήπτες τους. Λαμβάνοντας λοιπόν υπόψη την ανάγκη για διεξαγωγή της έρευνας σε συγκεκριμένη χρονική περίοδο, τη χρήση συγκεκριμένης μεθοδολογίας έρευνας, καθώς και τον περιορισμένο αριθμό του δείγματος ο οποίος διαμορφώθηκε τυχαία, θεωρούμε ότι δεν επιτρέπεται η γενίκευση των αποτελεσμάτων της έρευνας σε όλο τον πληθυσμό των ΨΑ. Παρ' όλα αυτά, στη συγκεκριμένη έρευνα παρουσιάζονται σημαντικά αποτελέσματα αποτυπώνοντας τις τάσεις, τις πρακτικές, τις αδυναμίες και τις ελλείψεις που εντοπίζονται στον χώρο των ΨΑ και που σχετίζονται με τη χρήση των προτύπων μεταδεδομένων και την ευρετηρίαση του υλικού τους από τις μηχανές επιστημονικής αναζήτησης όπως το GS.

1.5 Ορισμοί

Παρακάτω ακολουθούν οι ορισμοί των κυριότερων όρων που θα αναφερθούν και θα αναλυθούν στην εργασία.

1.5.1 Μεταδεδομένα

Τα μεταδεδομένα είναι δεδομένα τα οποία περιγράφουν άλλα δεδομένα, είναι παντού και καθιστούν τις πληροφορίες χρήσιμες. Η αποτελεσματική διαχείριση των πληροφοριών εξαρτάται από τα μεταδεδομένα, που θεωρούνται ως η βάση για την επιτυχή διάδοση των πληροφοριών αυτών. Τα μεταδεδομένα βοηθούν στον τρόπο με τον οποίο ένας ψηφιακός πληροφοριακός πόρος μπορεί να περιγραφεί, να εντοπιστεί και να ανακαλυφθεί. Επομένως, τα μεταδεδομένα είναι κρίσιμα, καθώς δίνουν νόημα και πλαίσιο στον πόρο που περιγράφουν. Χωρίς αυτά ένας πόρος θα μπορούσε να παραμείνει αταυτοποίητος χωρίς να εντοπιστεί ποτέ (Simons & Richardson, 2013). Τα μεταδεδομένα είναι δεδομένα τα οποία μπορούν να γίνουν κατανοητά από τα διάφορα λογισμικά/μηχανές (Berners-Lee, 1997). Επομένως, μπορούν να χρησιμοποιηθούν για να κάνουν τη ζωή μας ευκολότερη, για να μπορούμε να κάνουμε τα πάντα πιο ομαλά και πιο γρήγορα. Έχουν επίσης καθορισμένη σημασιολογία και δομή. Ο στόχος των μεταδεδομένων είναι να παράγουν πληροφορίες σε δομημένη μορφή, επιτρέποντας σε μηχανές και ανθρώπους να τις εντοπίζουν, να τις ανακαλύπτουν και να τις χρησιμοποιούν. Η δημιουργία αρχείου μεταδεδομένων αποτελεί σημαντική προϋπόθεση για την αποτελεσματική διαχείριση πληροφοριών. Ωστόσο, η

δημιουργία μεταδεδομένων είναι μία χρονοβόρα διαδικασία, η ποιότητα του αποτελέσματος της οποίας εξαρτάται από τις ικανότητες όσων εμπλέκονται σε αυτήν (Mondal, Meenachi, & Sai Baba, 2018). Παραδοσιακά, οι βιβλιοθήκες, τα ευρετήρια, οι ειδικοί και οι συγγραφείς παράγουν με μη αυτοματοποιημένο τρόπο (manually) αρχεία μεταδεδομένων. Στο ψηφιακό περιεχόμενο, τα μεταδεδομένα είναι δομημένα και κωδικοποιημένα δεδομένα τα οποία περιγράφουν χαρακτηριστικά πληροφοριακών οντοτήτων, επιδιώκοντας την ταύτιση, αναγνώριση, ανακάλυψη, αξιολόγηση και διαχείριση των οντοτήτων που περιγράφονται (Κυριάκη-Μάνεση & Κουλούρης, 2015). Διαφορετικές κατηγορίες ατόμων μπορούν να δημιουργήσουν αρχεία μεταδεδομένων χρησιμοποιώντας διάφορα εργαλεία και τεχνικές. Οι δημιουργοί μεταδεδομένων μπορούν να διαχωριστούν ως εξής: i) Επαγγελματίες δημιουργοί μεταδεδομένων, ii) Δημιουργοί τεχνικών μεταδεδομένων (μεταδεδομένων που σχετίζονται δηλαδή με τις τεχνικές ιδιότητες ενός ψηφιακού αρχείου ή με συγκεκριμένα περιβάλλοντα υλικού και λογισμικού που απαιτούνται για την απόδοση ή την επεξεργασία ψηφιακών πληροφοριών κ.λπ.), iii) Δημιουργοί περιεχομένου και iv) Δημιουργοί θεματικών ευρετηρίων και εξειδικευμένων όρων, καθώς θεωρούνται ειδικοί σε μία θεματική ενότητα (subject or community creators) (Greenberg, 2003). Οι πληροφορίες που γίνονται όλο και περισσότερο διαθέσιμες σε ψηφιακή μορφή, απαιτούν υιοθέτηση νέων προσεγγίσεων για τη δημιουργία μεταδεδομένων και είναι επιθυμητό να υπάρχουν μηχανικές (αυτοματοποιημένες) μέθοδοι. Η απόκτηση μεταδεδομένων μπορεί να γίνει με δύο μεθόδους (Tonkin & Muller, 2008), είτε με αυτόματη εξαγωγή τους από το pre-print, είτε με την εισαγωγή τους με παραδοσιακό τρόπο (μη αυτόματο) από τον συγγραφέα ή τον βιβλιοθηκονόμο ψηφιακών συλλογών. Κάθε μέθοδος δημιουργίας μεταδεδομένων έχει οφέλη και μειονεκτήματα. Η χειροκίνητη δημιουργία (manually), για παράδειγμα, μπορεί να προσαρμοστεί καλύτερα σε έναν μεμονωμένο πληροφοριακό πόρο. Από την άλλη είναι μια δαπανηρή διαδικασία, όσον αφορά τους μισθούς του προσωπικού, την εμπειρογνομosύνη και τον χρόνο. Η αυτόματη δημιουργία μπορεί να χρησιμοποιηθεί για τη λήψη μεταδεδομένων κατά τη δημιουργία πληροφοριακών πόρων, καθώς και κατά τη διάρκεια ζωής του πόρου, αλλά μπορεί να συνεπάγεται αρχικό κόστος για την αγορά λογισμικού. Καμία από τις μεθόδους δεν είναι εντελώς απαλλαγμένη από σφάλματα, καθώς οι άνθρωποι κάνουν λάθη και τα μηχανήματα ενδέχεται να έχουν λανθασμένες ρυθμίσεις ή να μην λειτουργούν σωστά για κάποιο λόγο (Simons & Richardson, 2013). Τα αρχεία μεταδεδομένων που δημιουργούνται από διαφορετικές οντότητες μπορούν ευρέως να ταξινομηθούν σε χειροκίνητα (manually)/μη αυτόματο τρόπο, σε ημιαυτόματο και αυτόματο τρόπο. Τα μεταδεδομένα προέρχονται και

αφορούν αποκλειστικά τον πληροφοριακό πόρο που περιγράφουν. Η προτυποποίηση των μεταδεδομένων στοχεύει στην αξιοπιστία και στην ομοιομορφία των εγγραφών, έτσι ώστε να αυξάνεται/ισχυροποιείται η διαλειτουργικότητα μεταξύ των συλλογών και των συστημάτων (Κυριάκη-Μάνεση & Κουλούρης, 2015). Για την επίτευξη της τυποποίησης και της ομοιομορφίας στην περιγραφή ενός πληροφοριακού πόρου χρησιμοποιούνται οδηγίες, κανόνες και πρότυπα (π.χ. AACR2, RDA, DC κ.λπ.). Τα μεταδεδομένα διακρίνονται σε περιγραφικά, δομικά και διαχειριστικά. Οι τρεις αυτές βασικές ομάδες μεταδεδομένων αποτυπώνουν διαφορετικά στοιχεία από έναν πληροφοριακό πόρο. Τα περιγραφικά μεταδεδομένα, χρησιμοποιούνται για την περιγραφή ενός πόρου με σκοπό τον εντοπισμό και την ανακάλυψή του (π.χ. τίτλος, συγγραφέας, κ.λπ.). Τα δομικά μεταδεδομένα χρησιμοποιούνται για να αποτυπώσουν τα μέρη από τα οποία συντίθεται ένας πόρος (π.χ. τα κεφάλαια που απαρτίζουν ένα βιβλίο). Τα διαχειριστικά μεταδεδομένα χρησιμεύουν στη διαχείριση ενός πληροφοριακού πόρου (π.χ. αποτύπωση των αδειών χρήσης ενός πόρου) (Simons & Richardson, 2013).

1.5.2 Διαλειτουργικότητα

Η χρησιμότητα των μεταδεδομένων και η αξιοποίηση πρότυπων σχημάτων μεταδεδομένων αποσκοπεί στη διαλειτουργικότητα μεταξύ των πληροφοριακών συστημάτων. Η επίτευξη επικοινωνίας και ανταλλαγής πληροφοριών, ακόμη και μεταξύ πληροφοριακών συστημάτων που δεν μοιράζονται κοινά λειτουργικά συστήματα, γίνεται χάριν στην εφαρμογή κοινών ανοικτών προτύπων για τη δημιουργία των μεταδεδομένων τους. Ως εκ τούτου, γίνεται επιτακτική η χρήση ανοικτών προτύπων δημιουργίας μεταδεδομένων, τα οποία έχουν σχεδιαστεί με τέτοιο τρόπο ώστε να μπορούν να επικοινωνούν με τους περισσότερους, αν όχι όλους τους γνωστούς τύπους λειτουργικών συστημάτων, όπως και με διάφορους τύπους φυλλομετρητών (browsers) στον Παγκόσμιο Ιστό (Chrome, Firefox, Internet Explorer, Safari, κ.λπ.). Για την περιγραφή των πληροφοριακών πόρων πρέπει κάθε πεδίο μεταδεδομένων να δηλώνει με ακρίβεια το είδος της πληροφορίας που μεταφέρει (πχ. ημερομηνία έκδοσης). Γίνεται λοιπόν αντιστοίχιση κάθε πεδίου με ένα στοιχείο μεταδεδομένων. Χρησιμοποιούμε επίσης ανοικτά διαλειτουργικά πρότυπα, τα οποία επιτρέπουν τη διασύνδεση της κάθε διαδικασίας τεκμηρίωσης ψηφιακών ή ψηφιοποιημένων τεκμηρίων με άλλα συστήματα αναζήτησης, παρέχοντας παράλληλα τη δυνατότητα συγκομιδής (harvesting) μεταδεδομένων από άλλους οργανισμούς (Κυριάκη-Μάνεση & Κουλούρης, 2015).

1.5.3 Πρότυπα και σχήματα μεταδεδομένων

Εφαρμογή των μεταδεδομένων μπορεί να γίνει τόσο σε φυσικά όσο και σε ψηφιακά τεκμήρια. Θα ήταν πολύ δύσκολο όμως για τους χρήστες να εντοπίσουν και να ερμηνεύσουν τις πληροφορίες που αναζητούν, αν κάθε δημιουργός μεταδεδομένων χρησιμοποιούσε τον δικό του μοναδικό τρόπο για να περιγράψει τους πόρους αυτούς. Αυτός είναι ο λόγος για τον οποίο τα σχήματα και τα πρότυπα μεταδεδομένων έχουν ζωτική σημασία. Ένα σχήμα μεταδεδομένων αποτελείται από ένα σύνολο στοιχείων (elements) που αντιστοιχούν σε συγκεκριμένα πεδία. Περιέχει επίσης οδηγίες για την έννοια και το περιεχόμενο που πρέπει να καταγράφεται σε κάθε πεδίο, εξασφαλίζοντας με αυτόν τον τρόπο ότι σε κάθε πεδίο θα καταχωρίζεται συγκεκριμένη πληροφορία, αποτρέποντας έτσι τη δημιουργία αναντιστοιχιών ως προς την έννοια και το περιεχόμενο των μεταδεδομένων. Αυτός είναι ένας τρόπος επίτευξης της διαλειτουργικότητας (Κυριάκη-Μάνεση & Κουλούρης, 2015). Στην περιγραφή ενός πόρου, επιλέγονται τα στοιχεία του σχήματος και ορίζονται τα χαρακτηριστικά τους (οι τιμές), για παράδειγμα το στοιχείο «Εκδότης» συμπληρώνεται με μία τιμή όπως «Taylor & Francis» (Simons & Richardson, 2013). Τα σχήματα περιλαμβάνουν κι άλλες σημαντικές πληροφορίες, όπως ορισμούς που υποδεικνύουν ποια στοιχεία είναι υποχρεωτικά και ποια προαιρετικά, αν κάποια είναι επαναλαμβανόμενα ή όχι, αν πρέπει να βρίσκονται σε μια συγκεκριμένη σειρά, πώς κωδικοποιούνται αυτά και τα χαρακτηριστικά τους για τον Ιστό - για παράδειγμα, σε γλώσσα XML (eXtensible Markup Language) και RDF (Resource Description Framework). Ένα μόνο σχήμα μεταδεδομένων δεν είναι κατάλληλο για να περιγράψει όλους τους τύπους πόρων. Διάφοροι οργανισμοί και επιστημονικές κοινότητες ανέπτυξαν σχήματα μεταδεδομένων για τη διαχείριση και την περιγραφή των συλλογών τους. Ορισμένα σχήματα, συνήθως κατόπιν αιτήματος των οργανισμών διαχείρισής τους, έχουν επίσης γίνει πρότυπα, τα οποία έχουν εγκριθεί τυπικά και από οργανισμούς τυποποίησης, όπως ο Information Standards Organization (ISO) και ο National Information Standards Organization (NISO). Τα πρότυπα μεταδεδομένων είναι ένα μέσο καθορισμού μιας κοινής αντίληψης ή χρήσης ενός κοινού συνόλου προσδιοριστών ή λεξιλογίου, για τη διευκόλυνση τόσο της ανακάλυψης των πόρων όσο και της ανταλλαγής δεδομένων μεταξύ των συστημάτων (Simons & Richardson, 2013). Υπάρχουν λοιπόν πρότυπα μεταδεδομένων που εξυπηρετούν τη βιβλιογραφική περιγραφή, πρότυπα αρχειακής περιγραφής, μουσειακών τεκμηρίων κ.λπ. Από τα πιο δημοφιλή πρότυπα και σχήματα μεταδεδομένων είναι το Dublin Core (DC) το οποίο και θα εξετάσουμε λεπτομερώς παρακάτω.

1.5.4 Ψηφιακά αποθετήρια

Τα Ψηφιακά αποθετήρια (ΨΑ) αποτελούν ουσιαστικά την επέκταση των ψηφιακών βιβλιοθηκών. Ορίζονται ως βάσεις δεδομένων που συλλέγουν και αποθηκεύουν ψηφιοποιημένα ή πρωτογενώς ψηφιακά τεκμήρια, καθιστώντας τα ελεύθερα προσιτά στο κοινό. Συνήθως λειτουργούν με τεχνολογίες διακομιστή Ιστού (web), χρησιμοποιούν περιγραφικά μεταδεδομένα και επειδή υποστηρίζουν την ανοικτή πρόσβαση, όλοι επωφελούνται από την επιτυχή συγκομιδή και ευρετηρίαση του περιεχομένου τους από τις μηχανές αναζήτησης στο διαδίκτυο (Arlitsch & O'Brien, 2012). Είναι σημεία στα οποία κατατίθεται και φυλάσσεται η ερευνητική παραγωγή των ακαδημαϊκών ιδρυμάτων με σκοπό την διατήρησή της εις το διηνεκές, την τεκμηρίωσή της, την ανακάλυψη και την επαναχρησιμοποίησή της (Κυριάκη-Μάνεση & Κουλούρης, 2015). Τα αποθετήρια διακρίνονται σε ιδρυματικά, ερευνητικά, εθνικά και θεματικά (Armbruster & Romary, 2010). Τα ιδρυματικά αποθετήρια εξαρτώνται από κάποιο ακαδημαϊκό ίδρυμα, τα ερευνητικά αποθετήρια είναι συνήθως επιχορηγούμενα από ερευνητικούς οργανισμούς, τα εθνικά αποθετήρια συλλέγουν επίσης την ακαδημαϊκή παραγωγή μιας χώρας με μια πιο γενική άποψη όμως, σκοπεύοντας στην υποστήριξη της εκπαιδευτικής διαδικασίας και τα θεματικά αποθετήρια επικεντρώνονται γύρω από ένα γνωστικό αντικείμενο. Τα ιδρυματικά αποθετήρια συλλέγουν και διαχειρίζονται την πνευματική παραγωγή που παράγεται από τα μέλη της κοινότητας των ακαδημαϊκών ιδρυμάτων ή επιστημονικών κλάδων έρευνας, ενώ τα θεματικά αποθετήρια συλλέγουν υλικό που σχετίζεται με ένα συγκεκριμένο επιστημονικό τομέα, χωρίς όμως περιορισμό στον φορέα προέλευσης του υλικού αυτού. Τα αποθετήρια βασίζονται αποκλειστικά σε πρότυπα για την δημιουργία των μεταδεδομένων τους και τεκμηριώνουν το σύνολο των τεκμηρίων τους. Η υποβολή του ψηφιακού περιεχομένου γίνεται με την μέθοδο της αυτό-απόθεσης από τους ίδιους τους δημιουργούς, διαδραματίζοντας έτσι ενεργό ρόλο στη δημοσιοποίηση και τεκμηρίωση του έργου τους. Με τον τρόπο αυτό πραγματοποιείται συλλογή της γκρίζας βιβλιογραφίας, κάτι που ήταν ιδιαίτερα δύσκολο να επιτευχθεί στο παρελθόν. Με την εφαρμογή προτύπων μεταδεδομένων και τη δημιουργία ειδικών πεδίων εξασφαλίζεται η ποιότητα του περιεχομένου του ψηφιακού αποθετηρίου και πραγματοποιείται η διαχείριση σε θέματα αδειών χρήσης και πνευματικής ιδιοκτησίας. Παρέχεται η δυνατότητα απομακρυσμένης πρόσβασης στα πλήρη κείμενα, στηρίζοντας με τον τρόπο αυτό το κίνημα της ανοικτής πρόσβασης. Η διαλειτουργικότητα σε όλα τα επίπεδα (σχήματα και πρότυπα μεταδεδομένων, πληροφοριακά συστήματα, πρωτόκολλα επικοινωνίας και ανταλλαγής δεδομένων όπως το OAI (Open Archives Initiative) και το Z39.50) παρέχει δυνατότητες

συγκομιδής και αύξησης του υλικού των αποθετηρίων, αυξάνοντας παράλληλα τη χρήση τους. Αυτό έχει ως αποτέλεσμα την ανακάλυψη και την ανάκτηση μεγάλου όγκου ψηφιακού περιεχομένου χωρίς την μεσολάβηση των εμπορικών βάσεων.

1.5.5 Google Scholar

Το Google Scholar είναι μια μηχανή αναζήτησης του Παγκόσμιου Ιστού με ελεύθερη πρόσβαση, η οποία ευρετηριάζει το πλήρες κείμενο ή τα μεταδεδομένα της ακαδημαϊκής και επιστημονικής βιβλιογραφίας. Περιλαμβάνει άρθρα από τα περισσότερα επιστημονικά περιοδικά και βιβλία, δημοσιεύσεις από πρακτικά συνεδρίων, επιστημονικές εργασίες και διατριβές, προ-εκδόσεις άρθρων (pre-prints), περιλήψεις, τεχνικές εκθέσεις, δικαστικές αποφάσεις, διπλώματα ευρεσιτεχνίας κ.λπ. ("Google Scholar," n.d.).

1.5.6 Ανοικτή Πρόσβαση

Η Ανοικτή Πρόσβαση (ΑΠ) είναι ένα μοντέλο δημοσίευσης και διανομής που καθιστά το ακαδημαϊκό και ερευνητικό περιεχόμενο άμεσα ελεύθερα διαθέσιμο στο κοινό χωρίς κόστος ή άλλα εμπόδια πρόσβασης. Το κίνημα της ανοιχτής πρόσβασης ξεκίνησε για να βελτιώσει την πρόσβαση στην έρευνα που χρηματοδοτείται από δημόσιους πόρους και για να βοηθήσει τις βιβλιοθήκες να αντιμετωπίσουν τον αμείωτο πληθωρισμό στις τιμές των συνδρομών των περιοδικών στοχεύοντας στην αξιολόγηση της επιστημονικής βιβλιογραφίας από ομότιμους επιστήμονες ("Open access," n.d.). Ιστορικά, αυτό βασιζόταν κυρίως σε έντυπα επιστημονικά περιοδικά τα οποία κάλυπταν τις δαπάνες δημοσίευσής τους μέσω των συνδρομών ή ακόμα και σε ηλεκτρονικά περιοδικά μέσω των χρεώσεων ανά προβολή των άρθρων τους, συνδρομή κ.λπ. Αξιοποιώντας όμως τη δύναμη του Διαδικτύου και με τη συγκατάθεση του δημιουργού (κατόχου των πνευματικών δικαιωμάτων), η ΑΠ βελτιώνει την ποιότητα της έρευνας, επιταχύνει την ανακάλυψη και τροφοδοτεί την καινοτομία (Suber, 2004).

1.5.7 Βιβλιομετρικοί δείκτες – H-index και i10-index

Οι βιβλιομετρικοί δείκτες απαντούν σε σημαντικά ερωτήματα όπως ποιες είναι οι ερευνητικές επιδόσεις, ποια η κατάταξη σύμφωνα με τις αναφορές που λαμβάνει μία επιστημονική εργασία, ποιος είναι ο βαθμός επιρροής και η αναγνώριση της έρευνας ενός επιστήμονα/ερευνητή. Επίσης, οι δείκτες αυτοί λαμβάνονται υπόψη, ώστε να εκτιμηθεί η επίδοση/απήχηση ενός ερευνητικού/ακαδημαϊκού οργανισμού και να καθοριστεί ο μελλοντικός τρόπος δράσης του, επιδιώκοντας την προαγωγή της ερευνητικής του δραστηριότητας, όπως και πιθανή χρηματοδότησή της. Ο δείκτης h-index επιχειρεί να

μετρήσει τόσο την παραγωγικότητα, όσο και την επιρροή των αναφορών σε δημοσιεύσεις ενός επιστήμονα. Ο δείκτης βασίζεται στο σύνολο των δημοσιεύσεων ενός ερευνητή και στον αριθμό των ετεροαναφορών που έχουν λάβει οι δημοσιεύσεις αυτές, σε δημοσιεύσεις τρίτων (ομότιμων επιστημόνων και μη) . Είναι μια μέτρηση που στοχεύει στην περιγραφή της επιστημονικής παραγωγικότητας και της απήχησης ενός ερευνητή στην επιστημονική κοινότητα (Hirsch, 2005; Jacsó, 2012a). Ο ερευνητικός δείκτης είναι η μέτρηση του αριθμού των πολύ σημαντικών εργασιών που έχει δημοσιεύσει ένας ερευνητής. Όσο μεγαλύτερος είναι ο αριθμός των σημαντικών εργασιών, τόσο υψηλότερος είναι ο δείκτης h. Για τον υπολογισμό του, απαιτούνται μόνο δύο πληροφορίες: ο συνολικός αριθμός δημοσιευμένων εργασιών ενός ερευνητή (Np) και ο αριθμός των ετεροαναφορών (Nc) που έλαβε η εκάστοτε εργασία. Ο δείκτης h καθορίζεται από τις πόσες φορές οι δημοσιεύσεις (Np) ενός ερευνητή έχουν λάβει τουλάχιστον h ετεροαναφορές η κάθε μία, εντός ενός χρονικού διαστήματος στο οποίο γίνεται η ανάλυση (Aithal, 2017; Παπαβλασόπουλος, 2015). Ο ίδιος δείκτης μπορεί επίσης να χρησιμοποιηθεί για να μετρήσει την ποιότητα ενός επιστημονικού περιοδικού, καθώς και μιας ομάδας επιστημόνων, ενός Πανεπιστημιακού Τμήματος, ενός Πανεπιστημίου ή ακόμα και μιας ολόκληρης χώρας ή ηπείρου, όσον αφορά την παραγωγικότητά της στην υψηλής ποιότητας έρευνα (Aithal, 2017; Jacsó, 2012a; Παπαβλασόπουλος, 2015). Αντίστοιχα ένας ακόμη βιβλιομετρικός δείκτης είναι και ο i10-index, ο οποίος παρέχεται δωρεάν από το GS (Aithal, 2017; Jacsó, 2012a; Παπαβλασόπουλος, 2015). Ο δείκτης αυτός μας δείχνει πόσες επιστημονικές δημοσιεύσεις ενός επιστήμονα έχουν λάβει τουλάχιστον δέκα (10) ετεροαναφορές. Είναι ένας δείκτης από τον οποίο μπορούμε να αποκτήσουμε μια αίσθηση για την παραγωγικότητα και την απήχηση ενός συγκεκριμένου συγγραφέα (Aithal, 2017; Jacsó, 2012a; Kozak & Bornmann, 2012).

1.6 Οργάνωση Κεφαλαίων

Το 1^ο κεφάλαιο αποτελεί την εισαγωγή αυτής της διπλωματικής εργασίας. Στην εισαγωγή παρουσιάζονται ο σκοπός και οι στόχοι που καθορίστηκαν σε ένα αρχικό στάδιο, οι ερευνητικές υποθέσεις που τέθηκαν, η αιτιολόγηση του θέματος, οι περιορισμοί που εντοπίστηκαν στο πλαίσιο της έρευνας, όπως και οι ορισμοί των κυριότερων όρων που αναφέρονται και αναλύονται στην εργασία.

Στο 2^ο κεφάλαιο γίνεται η παράθεση της βιβλιογραφίας που χρησιμοποιήθηκε και αποτελεί το θεωρητικό μέρος της μελέτης. Η βιβλιογραφική επισκόπηση πραγματοποιήθηκε σε 2 στάδια. Στο πρώτο στάδιο εξετάζεται η συμπεριφορά των ερευνητών ως προς τους τρόπους

αναζήτησης και εντοπισμού των ψηφιακών πόρων, η σημασία και ο ρόλος των ψηφιακών αποθετηρίων ως μέσο ανάδειξης και προβολής της επιστημονικής κι εκδοτικής παραγωγής των ερευνητικών/ακαδημαϊκών οργανισμών, όπως επίσης η ανάδειξη των μηχανών επιστημονικής αναζήτησης ως μέσο ανακάλυψης των βιβλιογραφικών πηγών. Οι λόγοι που οδηγούν στη συνεχή αύξηση χρήσης των μηχανών αναζήτησης και οι λόγοι από τους οποίους μπορεί να επηρεάζεται η ευρετηρίαση του περιεχομένου των ψηφιακών αποθετηρίων από τις μηχανές αναζήτησης όπως το GS. Στο δεύτερο στάδιο γίνεται αξιολόγηση των δυνατοτήτων που παρέχονται από τα πρότυπα μεταδεδομένων τόσο στην τυποποίηση, όσο και στην ευρετηρίαση και ανάκτηση της πληροφορίας τόσο από τα ψηφιακά αποθετήρια, όσο και από τις μηχανές αναζήτησης, παρουσιάζοντας τα πλεονεκτήματα και τα μειονεκτήματα του περισσότερο δημοφιλούς προτύπου μεταδεδομένων ψηφιακών αποθετηρίων (Dublin Core), όπως και των σχημάτων μεταδεδομένων που προτείνονται από τις μηχανές αναζήτησης και συγκεκριμένα από το Google Scholar.

Στο 3^ο κεφάλαιο παρουσιάζεται αναλυτικά η μεθοδολογία που ακολουθήθηκε. Συγκεκριμένα γίνεται αναφορά στον τρόπο συλλογής και καταχώρησης των δεδομένων, όπως επίσης στους πίνακες, στα γραφήματα και στα διαγράμματα που περιλαμβάνονται στη μελέτη, στις στατιστικές μεθόδους ανάλυσης, καθώς και στον τρόπο παρουσίασης των συμπερασμάτων μετά την επεξεργασία των στοιχείων.

Στο 4^ο κεφάλαιο αναπτύσσονται τα ευρήματα και η ανάλυση των δεδομένων αυτής της διπλωματικής εργασίας.

Στο 5^ο και τελευταίο κεφάλαιο περιλαμβάνεται η συζήτηση της μεταπτυχιακής εργασίας αναφορικά με τα ευρήματα, τα οποία σχετίζονται με τη θεωρία και με την προηγούμενη έρευνα. Επιπλέον, γίνεται συζήτηση για την αξιοποίηση της έρευνας όπως και για πιθανή μελλοντική έρευνα. Έπειτα ακολουθούν η Βιβλιογραφία και τα Παραρτήματα.

Κεφάλαιο 2. Θεωρητικό μέρος – Βιβλιογραφική έρευνα – Σχετικές προσπάθειες

Στη βιβλιογραφική επισκόπηση γίνεται προσπάθεια εντοπισμού κι ανάλυσης των πλεονεκτημάτων και μειονεκτημάτων των προτύπων μεταδεδομένων που χρησιμοποιούν τα περισσότερα ΨΑ και πιο συγκεκριμένα του Dublin Core και αυτών που προτείνει και χρησιμοποιεί το Google Scholar. Γίνεται προσπάθεια να εντοπιστούν και να γίνουν κατανοητές οι αδυναμίες που υπάρχουν και που εμποδίζουν την ευρετηρίαση και τεκμηρίωση του επιστημονικού περιεχομένου των ψηφιακών αποθετηρίων από τις δυναμικές μηχανές ερευνητικής αναζήτησης όπως το GS.

2.1 Θεωρητικό μέρος – Βιβλιογραφική έρευνα

Στο πλαίσιο της διάδοσης και της αξιολόγησης της επιστημονικής εκδοτικής παραγωγής και γνώσης γενικότερα, σε συνδυασμό με την άνοδο του κινήματος της Ανοικτής Πρόσβασης παρατηρείται αύξηση στη δημιουργία αποθετηρίων κυρίως ιδρυματικών. Τα περισσότερα πανεπιστήμια σε όλο τον κόσμο έχουν δημιουργήσει το δικό τους ιδρυματικό αποθετήριο με σκοπό την παροχή υπηρεσιών ανοικτής πρόσβασης στην επιστημονική γνώση που παράγουν, τη διάσωση και διάδοση της ακαδημαϊκής γνώσης και την αντανάκλαση των δραστηριοτήτων κάθε ιδρύματος. Τα αποθετήρια αυτά διαδραματίζουν παράλληλα καθοριστικό ρόλο στη δημιουργία της διαδικτυακής ψηφιακής ταυτότητας του ιδρύματος που αντιπροσωπεύουν (I. Aguillo, 2009). Υπάρχουν επίσης εμπειρικά στοιχεία που αποδεικνύουν ότι η αυτό-αρχαιοθέτηση, η οποία στα αποθετήρια παίζει κεντρικό ρόλο, έχει ήδη γίνει ο κύριος μοχλός πίσω από την ΑΠ (Archambault et al., 2013). Έχοντας υπόψη τη σημασία της αυτό-αρχαιοθέτησης για την παραγωγή περιεχομένου ΑΠ, η παρουσία και η ορατότητα του περιεχομένου που φιλοξενείται από τα ψηφιακά αποθετήρια στον Ιστό (και ειδικά στις μηχανές αναζήτησης όπως το GS), αποτελούν απαραίτητες προϋποθέσεις για τη διασφάλιση, ότι η ΑΠ είναι πραγματικά αποτελεσματική για την ερευνητική κοινότητα. Οι συνεχείς εξελίξεις στις τεχνολογίες των επικοινωνιών και του Παγκόσμιου Ιστού παράλληλα με το κίνημα της ΑΠ, έχουν οδηγήσει σε αλλαγή της συμπεριφοράς των ερευνητών όλων των βαθμίδων, με αποτέλεσμα την αύξηση χρήσης των μηχανών αναζήτησης στη διαδικασία ανακάλυψης και ανάκτησης επιστημονικού περιεχομένου. Επομένως, έχει δοθεί

μεγάλη βαρύτητα στην εξέλιξη των μηχανών αναζήτησης του Διαδικτύου, με πρώτη σε προτίμηση των χρηστών το GS (Orduña-Malea & Delgado López-Cózar, 2015; Rieger, 2009). Βλέπουμε λοιπόν από τα μέσα της δεκαετίας του 2000 φοιτητές να χρησιμοποιούν ακόμη και σε ποσοστό που άγγιζε το 89% τις μηχανές αναζήτησης, για την αναζήτηση και ανάκτηση επιστημονικής βιβλιογραφίας, ενώ αντίστοιχα σημειώνεται μόλις ένα 2% των φοιτητών της ίδιας έρευνας, να διενεργεί την έρευνά του ξεκινώντας την αναζήτησή του από τις ιστοσελίδες του ιδρύματός του (De Rosa et al., 2005). Η επικαιροποίηση της ίδιας έρευνας το 2010 διαπιστώνει ακόμα μεγαλύτερη διεύρυνση του προβλήματος αυτού. Η κατάσταση για τους ιστοτόπους των βιβλιοθηκών έχει επιδεινωθεί ακόμη περισσότερο, καθώς δεν εντοπίστηκε ούτε ένας από τους ερωτώμενους, ο οποίος να ξεκίνησε την έρευνά του χρησιμοποιώντας τις ιστοσελίδες της βιβλιοθήκης του ιδρύματός του ή κάποιου άλλου ιδρύματος (De Rosa et al., 2010). Αυτό είναι μία συνέπεια που προέρχεται και από την αδυναμία των χρηστών στην πραγματοποίηση αναζητήσεων μέσω των αποθετηρίων, λόγω της μικρής εμπειρίας τους στην αναζήτηση υλικού ή και ελλιπούς γνώσης ως προς τη χρήση των αποθετηρίων και των πληροφοριακών πηγών γενικότερα (Arlitsch & O'Brien, 2012). Η κατάσταση που περιγράφεται δεν είναι καθόλου ευχάριστη για τις βιβλιοθήκες και τις υπηρεσίες τους. Ακόμη κι αν μέσω ερευνών επιβεβαιώνεται η έλλειψη αξιοπιστίας και ακρίβειας στα αποτελέσματα που εντοπίζονται μέσω μιας μηχανής αναζήτησης, οι φοιτητές συνεχίζουν να πραγματοποιούν την έρευνά τους χρησιμοποιώντας τον πιο εύκολο και πιο βολικό για αυτούς τρόπο. Ωστόσο, δεν είναι μόνο οι φοιτητές που δείχνουν να αγνοούν τους ιστοτόπους των βιβλιοθηκών και να απευθύνονται κατευθείαν στο Διαδίκτυο και στις μηχανές αναζήτησης για να πραγματοποιήσουν την έρευνά τους. Έρευνα που πραγματοποιήθηκε το 2010 μεταξύ ενεργών ερευνητών, διδασκόντων σε τέσσερα μεγάλα πανεπιστήμια, έφερε στο φως ότι ακόμα και οι ερευνητές θεωρούν την Google και την αντίστοιχη μηχανή επιστημονικής αναζήτησής της το GS ως εκπληκτικά αποτελεσματικά εργαλεία για την κάλυψη των αναγκών τους σε ανάκτηση πληροφοριών και δέχονται τα αποτελέσματα των μηχανών αυτών ως «αρκετά καλά» σε πολλές περιπτώσεις (Kroll & Forsman, 2010). Επίσης σε έρευνα που διενεργήθηκε στα πλαίσια του Ithaka Faculty Survey το 2009, διαπιστώθηκε ότι η ανακάλυψη επιστημονικών εργασιών μέσω Google και GS βρίσκεται στην τρίτη θέση (Schonfeld & Housewright, 2010). Είναι λοιπόν γεγονός ότι δίνεται μεγάλη προσοχή και σημασία στην ευρετηρίαση του επιστημονικού και ακαδημαϊκού έργου από τέτοιες μηχανές αναζήτησης, καθώς οι επιστήμονες τις χρησιμοποιούν και βασίζονται όλο και πιο πολύ στη χρήση τους. Παρ' όλα αυτά, οι συγγραφείς έχουν διαπιστώσει, ότι κάνοντας αποκλειστική χρήση του GS για την

ανακάλυψη και ανάκτηση ακαδημαϊκού περιεχομένου, μεγάλο μέρος της επιστημονικής βιβλιογραφίας από τα αποθετήρια δεν ανακτάται, μένει στην αφάνεια (Arlitsch & O'Brien, 2012; Orduña-Malea & Delgado López-Cózar, 2015). Σε αυτές ακριβώς τις περιπτώσεις τα αποθετήρια διαδραματίζουν καθοριστικό ρόλο, στην αποτροπή δηλαδή της απώλειας της ακαδημαϊκής προσπάθειας. Το πρόβλημα αυτό παρατηρείται και μέσω των ερευνών που σχετίζονται με την βελτιστοποίηση των μηχανών αναζήτησης (Search Engine Optimization - SEO), στις οποίες εντοπίζεται η ύπαρξη ιστοσελίδων που δεν ευρετηριάζονται, ή που δεν παρουσιάζονται σε ικανοποιητική σειρά εμφάνισης στα αποτελέσματα μιας αναζήτησης στο Διαδίκτυο. Μέσω των αναζητήσεων που πραγματοποιούνται στο Διαδίκτυο, αυξάνεται ο αριθμός επισκεψιμότητας των ιστοτόπων. Τι γίνεται όμως με τις τοποθεσίες (sites) που δεν μπορούν εύκολα να εντοπιστούν από μία μηχανή αναζήτησης στο Διαδίκτυο; Το πρόβλημα που επισημαίνεται είναι αρκετά σημαντικό. Η ανεπαρκής ευρετηρίαση από το GS των τεκμηρίων που αποτελούν μέρος των συλλογών ψηφιακών αποθετηρίων κορυφαίων ιδρυμάτων, δημιουργεί σοβαρή ανησυχία για την κοινότητα της ΑΠ (I. F. Aguillo, 2018). Η κατάσταση καταδεικνύεται με σαφήνεια από τον χαμηλό αριθμό κατάταξης των αποθετηρίων μεγάλων και σημαντικών πανεπιστημίων, όπως αυτό του Κέμπριτζ (Cambridge University Institutional Repository) και της Οξφόρδης (Oxford University Research Archive), τα οποία περιλαμβάνονται στη βάση δεδομένων της Google Scholar (<http://repositories.webometrics.info/en/transparent>). Λαμβάνοντας λοιπόν υπόψη την ευρεία χρήση των μηχανών αναζήτησης ως σημείο εκκίνησης για την πραγματοποίηση αναζητήσεων ακαδημαϊκού περιεχομένου και το γεγονός ότι στις περισσότερες περιπτώσεις ο πόρος αποκτάται απευθείας, χωρίς να διέλθει από τον ιστότοπο του αποθετηρίου, είναι ζωτικής σημασίας να αναπροσαρμοστεί ο τρόπος διάθεσης του περιεχομένου των ψηφιακών αποθετηρίων, όπως και ο τρόπος παροχής των μεταδεδομένων αυτού, ώστε να ευρετηριάζονται επιτυχώς από τις μηχανές αναζήτησης (Orduña-Malea & Delgado López-Cózar, 2015). Ορισμένα προβλήματα ευρετηρίασης των τεκμηρίων των ΨΑ από το GS συνοψίζονται στα εξής: i. οι διακομιστές Web ενδέχεται να μην έχουν ρυθμιστεί σωστά, ώστε να προσκαλούν προγράμματα ανίχνευσης (crawlers) μηχανών αναζήτησης και η απόδοση της ταχύτητας του διακομιστή μπορεί να είναι απαράδεκτα αργή, ii. τα μεταδεδομένα συχνά δεν είναι μοναδικά ή δομημένα ως αναγνωρίσιμες οντολογίες, και σε ορισμένες περιπτώσεις οι μηχανές αναζήτησης ενδέχεται να μην δέχονται το σχήμα μεταδεδομένων που χρησιμοποιείται από το αποθετήριο, iii. το λογισμικό του αποθετηρίου ενδέχεται να παρουσιάζει έναν αδιαπέραστο λαβύρινθο για τα προγράμματα ανίχνευσης, iv. οι πολιτικές και οι πρακτικές των μηχανών αναζήτησης

αλλάζουν και πρέπει να παρακολουθούνται ανάλογα και ν. οι μηχανές αναζήτησης ενδέχεται να μην υποστηρίζουν αποδεκτά πρότυπα και πρωτόκολλα της βιβλιοθηκονομικής κοινότητας (όπως OAI-PMH) (Arlitsch, Kenning, O'Brien, 2013). Στην παρούσα εργασία γίνεται προσπάθεια να εξεταστεί το πόσο επηρεάζει το πρότυπο μεταδεδομένων και οι πρακτικές που ακολουθούνται από τα ψηφιακά αποθετήρια ως προς αυτά, τον βαθμό ευρετηρίασης των αποθετηρίων από τις μηχανές επιστημονικής αναζήτησης. Συγκεκριμένα εξετάζονται τα πλεονεκτήματα και τα μειονεκτήματα του πιο διαδεδομένου σε χρήση σχήματος μεταδεδομένων σε ψηφιακά αποθετήρια, δηλαδή του Dublin Core (Arlitsch & O'Brien, 2012; Palmer, Zavalina, & Mustafoff, 2007; Park & Childress, 2009; Simons & Richardson, 2013), σε σχέση με τα σχήματα μεταδεδομένων (δηλαδή την εξαγωγή βιβλιογραφικών δεδομένων σε ετικέτες <meta> HTML και συγκεκριμένα σε ετικέτες Highwire Press, Eprints, BE Press και PRISM) (Google Scholar, 2010) που προτείνονται από την κυρίαρχη μηχανή ερευνητικής αναζήτησης, δηλαδή το GS (Orduña-Malea & Delgado López-Cózar, 2015; Rieger, 2009).

2.1.1 Dublin Core - Πλεονεκτήματα και μειονεκτήματα

Το Dublin Core είναι ένα από τα πιο ευρέως χρησιμοποιούμενα πρότυπα μεταδεδομένων ψηφιακών αποθετηρίων. Αναπτύχθηκε με πρωτοβουλία της OCLC για την περιγραφή κυρίως ηλεκτρονικών και διαδικτυακών πόρων. Από την ίδρυσή του το 1995 στο Δουβλίνο του Οχάιο, το σύστημα μεταδεδομένων DC έχει αποκτήσει αξιοπιστία και εγκρίθηκε ως πρότυπο NISO (Z39.85) το 2001 (National Information Standards Organization, 2001) και το 2003 ως πρότυπο ISO 15836 (Dublin Core Metadata Initiative, n.d.-b) με πολλές ενημερώσεις μέχρι σήμερα. Το πρότυπο συντηρείται και ενημερώνεται από το Dublin Core Metadata Initiative (Dublin Core Metadata Initiative, n.d.-c), οργανισμός ο οποίος είναι υπεύθυνος για την ανάπτυξη και τη βελτίωσή του (Kurtz, 2010; Κυριάκη-Μάνεση & Κουλούρης, 2015). Ο βασικός άξονας του προτύπου (Simple) Dublin Core αποτελείται από 15 στοιχεία (elements), τα οποία είναι όλα προαιρετικά και επαναλαμβανόμενα και τα οποία έχουν τέτοια σημασιολογία, ώστε να εφαρμόζονται σε όλα τα είδη υλικού (πόρους του διαδικτύου, όπως βίντεο, εικόνες, ιστοσελίδες κλπ., καθώς και φυσικούς πόρους όπως βιβλία, CD, ή έργα τέχνης). Το εξειδικευμένο (Qualified) Dublin Core παρέχει τη δυνατότητα περαιτέρω εξειδίκευσης στα στοιχεία αυτά (με προσθήκη 7 ακόμη στοιχείων), βοηθώντας με τον τρόπο αυτό στον καθορισμό της σημασιολογίας των στοιχείων και επιτρέποντας μεγαλύτερη λεπτομέρεια στην περιγραφή, με στόχο την ακριβέστερη αναζήτηση και ανάκτηση των ψηφιακών πόρων. Παρ' όλα αυτά μία από τις πιο βασικές αρχές του Dublin

Core δεν είναι η πληροφοριακή ανάλυση με εξαντλητική λεπτομέρεια στην περιγραφή, αλλά η δημιουργία σημείων πρόσβασης στην πληροφορία χρησιμοποιώντας βασικές και κοινώς παραδεκτές έννοιες. Από την άποψη αυτή το Dublin Core είναι ένα πρότυπο που στηρίζει τις απαιτήσεις της εκτεταμένης χρήσης και την πρόσβαση έναντι της ανάλυσης. Το σύστημα μεταδεδομένων DC αναπτύχθηκε κυρίως για έναν πολύ σημαντικό σκοπό - για να διευκολύνει την περιγραφή πληροφοριακών πόρων και την ανταλλαγή τους μεταξύ διαφορετικών τομέων, συλλογών και οργανισμών (Κυριάκη-Μάνεση & Κουλούρης, 2015). Η απλότητα και η ευελιξία στην εφαρμογή του προτύπου, επιτρέπει την ευκολότερη καταλογογράφηση του υλικού, με μικρότερη ανάλυση βέβαια, περιορίζοντας όμως τα λάθη και τις περιγραφικές ασυνέπειες. Με τον τρόπο αυτό δίνεται μία λύση έναντι του μεγάλου όγκου συσσωρευμένων ακαταλογογράφητων πληροφοριακών πόρων και έναντι του υψηλού κόστους της καταλογογράφησης γενικότερα (Καπιδάκης, Λαζαρίνης, & Τοράκη, 2015). Κατά συνέπεια, η σημασιολογία DC παρέχει μία εύκολη και απλή διαδικασία ως προς την περιγραφή των πόρων, επιτρέποντας σε συγγραφείς και δημιουργούς εν γένει, χωρίς κάποιο ιδιαίτερο εκπαιδευτικό και επαγγελματικό υπόβαθρο στην καταλογογράφηση, να το χρησιμοποιήσουν για να περιγράψουν το έργο τους (Park & Childress, 2009; Simons & Richardson, 2013). Φυσικά για την επίτευξη της διαλειτουργικότητας και για την δημιουργία «καλών/έγκυρων» μεταδεδομένων είναι απαραίτητο να ακολουθούνται με προσοχή οι οδηγίες που δίνονται από το πρότυπο σχετικά με το περιεχόμενο κάθε στοιχείου και να δηλώνονται όλα τα πρότυπα και οι κανόνες που εφαρμόστηκαν στην εγγραφή του κάθε φορά (π.χ. AACR2, RDA, LCSH, κ.λπ.). Επειδή η τυποποίηση είναι πολύ σημαντική ιδιαίτερα στη συμπλήρωση στοιχείων που σχετίζονται με την θεματική περιγραφή των πόρων, αποδίδοντας μεγάλα οφέλη κυρίως στην ανάκτηση του υλικού, το DC ενσωματώνει πρότυπα. Επιτρέπει τη χρήση δηλαδή ελεγχόμενων λεξιλογίων ή αρχείων καθιερωμένου τύπου τα οποία μπορούν να χρησιμοποιηθούν για τη συμπλήρωση διαφόρων στοιχείων του DC (π.χ. Library of Congress Subject Headings, Θεματικές Επικεφαλίδες της ΕΒΕ κ.λπ.) (Simons & Richardson, 2013; Κυριάκη-Μάνεση & Κουλούρης, 2015). Η ποιότητα των μεταδεδομένων αντικατοπτρίζει τον βαθμό στον οποίο τα εν λόγω μεταδεδομένα εκτελούν τις βασικές βιβλιογραφικές λειτουργίες της ανακάλυψης και της χρήσης. Υπό αυτήν την έννοια, ο κύριος στόχος των μεταδεδομένων σχετίζεται σε έναν μεγάλο βαθμό με τον κύριο στόχο των παραδοσιακών ηλεκτρονικών καταλόγων βιβλιοθηκών και βάσεων δεδομένων, δηλαδή στην εύρεση, αναγνώριση, ταυτοποίηση, επιλογή και ανάκτηση των πόρων (Park & Childress, 2009). Το Dublin Core δεν παρέχει λεπτομερείς κανόνες καταλογογράφησης, με χρήση συγκεκριμένης

στήξης ή κανόνες που σχετίζονται με τον τρόπο καταχώρησης των τιμών που δηλώνονται στα στοιχεία του. Κανόνες στους οποίους βασίζονται άλλα πρότυπα, για να ορίσουν για παράδειγμα πώς το όνομα ενός φυσικού προσώπου με μνεία ευθύνης (συγγραφέας) καταχωρίζεται πάντα με το μικρό όνομα να έπεται του επιθέτου, χωριζόμενα από ένα κόμμα. Εντούτοις αφήνει στους τοπικούς κανόνες καταλογογράφησης, τους περιορισμούς αυτούς όπου χρειάζονται. Όμως, το Dublin Core επιτρέπει την επεκτασιμότητα σε άλλες κατηγορίες πόρων μέσω της εξειδίκευσης και της δημιουργίας προφίλ εφαρμογών (application profiles), τα οποία μπορούν να εφαρμόσουν συνδυαστικά περισσότερα του ενός προτύπου, ικανοποιώντας με αυτόν τον τρόπο τις ιδιαιτερότητες του υλικού και τις ανάγκες των συλλογών σε τοπικό επίπεδο (Simons & Richardson, 2013; Καπιδάκης et al., 2015; Κυριάκη-Μάνεση & Κουλούρης, 2015). Το DC είναι επίσης η ελάχιστη απαίτηση για την τήρηση του πρωτοκόλλου Open Archives Initiative for Metadata Harvesting (OAI-PMH) (“OAI-PMH Implementation Guidelines,” n.d.) και ως εκ τούτου έχει εφαρμοστεί ευρέως σε αποθετήρια με σκοπό την κοινή χρήση μεταδεδομένων από άλλους παρόχους υπηρεσιών πληροφόρησης. Ωστόσο, το DC δεν χρειάζεται να επιλεγεί ως η εγγενής μορφή μεταδεδομένων για εγγραφές που διατηρούνται στο αποθετήριο, εφόσον η εγγενής μορφή μπορεί να αντιστοιχιστεί σε DC (mapping conversion to DC) για σκοπούς συλλογής μέσω OAI-PMH. Αυτό διευκολύνει ιδιαίτερα στις περιπτώσεις που είναι ήδη καταχωρημένο το υλικό σε κάποιο άλλο περισσότερο αναλυτικό σχήμα (Simons & Richardson, 2013; Κυριάκη-Μάνεση & Κουλούρης, 2015). Η διεθνοποίηση του DC είναι κάτι ακόμα που αντιμετωπίζεται από το Dublin Core International, μεταφράζοντάς το σε αρκετές γλώσσες ενισχύοντας με αυτόν τον τρόπο τη διάδοσή του και εξυπηρετώντας τις απαιτήσεις της τοπικής γλώσσας και της πολυγλωσσικής περιγραφής. Η προσαρμογή του γίνεται χωρίς τεχνικές αλλαγές, διατηρώντας δηλαδή τα μηχανικά αναγνώσιμα σύμβολα των στοιχείων του κοινά για όλους ανεξαρτήτως γλώσσας, διευκολύνοντας με αυτόν τον τρόπο τη διαλειτουργικότητα. Το Dublin Core είναι ένα από τα πιο συνηθισμένα πρότυπα σε συστήματα ψηφιακών βιβλιοθηκών και αποθετηρίων για την περιγραφή των ψηφιακών τους πόρων, εξαιτίας κυρίως της απλότητας και της ευκολίας χρήσης του (Simons & Richardson, 2013; Καπιδάκης et al., 2015). Δημοφιλείς εφαρμογές λογισμικού ψηφιακών βιβλιοθηκών, με πολλές εγκαταστάσεις σε όλο τον κόσμο, όπως τα greenstone (“Greenstone Digital Library Software,” n.d.), dspace (“DSpace - A Turnkey Institutional Repository Application,” n.d.), fedora (“Fedora - The Flexible, Modular, Open-Source Repository Platform,” n.d.), omeka (“Omeka,” n.d.) κ.λπ., χρησιμοποιούν το DC ως συνιστώμενο σχήμα μεταδεδομένων, καλύπτοντας την πλειοψηφία των εγκατεστημένων ψηφιακών βιβλιοθηκών, τόσο διεθνές

όσο και σε τοπικό επίπεδο. Η ευρεία χρήση του Dublin Core αποδεικνύεται και από τον βαθμό επιρροής που έχει σε άλλα σχήματα μεταδεδομένων τα οποία βασίστηκαν σε αυτό (Καπιδάκης et al., 2015). Επίσης το DC προετοιμάζει τις ιστοσελίδες για ένα μέλλον στο οποίο οι μηχανές αναζήτησης παίζουν πολύ σημαντικό ρόλο. Βοηθά στην προετοιμασία των ιστοσελίδων για την εξέλιξη του σημασιολογικού ιστού - έναν κόσμο στον οποίο ορίζεται η σημασιολογία των πληροφοριών και των υπηρεσιών στον Ιστό, επιτρέποντάς του να αναγνωρίζει και να ικανοποιεί αιτήματα που αφορούν το περιεχόμενο από φυσικά πρόσωπα και μηχανές. Υποστηρίζει ανοικτά πρότυπα και δημιουργεί δυναμική για αυτά.

Συνοψίζοντας, τα πλεονεκτήματα που προκύπτουν από την επιλογή ενός καθιερωμένου προτύπου μεταδεδομένων όπως το DC είναι τα εξής:

- i. Ευρέως χρησιμοποιούμενο διεθνές πρότυπο
- ii. Πολύ απλό στο σύνολο των στοιχείων του
- iii. Επεκτάσιμο
- iv. Συμβατό με τις απαιτήσεις του πρωτοκόλλου συγκομιδής OAI-PMH για τη συλλογή μεταδεδομένων
- v. Με διαθέσιμη την κωδικοποίησή του για τον Ιστό, σε γλώσσες XML (eXtensible Markup Language) και RDF (Resource Description Framework)
- vi. Υποστηρίζει ανοικτά πρότυπα και δημιουργεί δυναμική για αυτά
- vii. Υποστηρίζεται και ενημερώνεται από τον Dublin Core Metadata Initiative, ο οποίος είναι υπεύθυνος για την ανάπτυξη και τη βελτίωσή του
- viii. Παρέχει αντιστοιχίσεις (mapping) από και προς άλλα σχήματα μεταδεδομένων
- ix. Προωθεί την ανάκτηση της πληροφορίας από τις μηχανές αναζήτησης
- x. Παρέχεται σε αρκετές γλώσσες (διεθνοποίηση)

Παρά τα οφέλη και τα πλεονεκτήματα του προτύπου μεταδεδομένων Dublin Core, διατυπώνονται διάφορες απόψεις σχετικά με την αποδοτικότητά του στην επίτευξη της ανακάλυψης και ανάκτησης των ψηφιακών πόρων από τις δυναμικές μηχανές αναζήτησης του Παγκόσμιου Ιστού, κάτι που έχει μεγάλη αξία και σημασία, ιδιαίτερα όταν πρόκειται για την ανακάλυψη και ανάκτηση της επιστημονικής παραγωγής και γνώσης. Αποτελέσματα ερευνών αποδεικνύουν ότι η περιγραφή ιστοσελίδων και ψηφιακών πόρων βάση του προτύπου DC, δεν επηρεάζει σημαντικά την επίδοση της ανακάλυψης και ανάκτησής τους από τις μηχανές αναζήτησης του Διαδικτύου (Safari, 2005). Ωστόσο, η ανακάλυψη πόρων είναι αδύνατη χωρίς την περιγραφή τους. Η επαρκής και ποιοτική περιγραφή των ψηφιακών πόρων διασφαλίζει την αποτελεσματική ανακάλυψή τους (Arlitsch & O'Brien, 2012; Park & Tosaka, 2010; Safari, 2005). Πιστεύεται λοιπόν ότι η βελτίωση της ανακάλυψης

πόρων στον Ιστό, έγκειται στη χρήση μεταδεδομένων. Για να μπορεί όμως το σχήμα μεταδεδομένων να παίζει ουσιαστικό ρόλο στην ανακάλυψη των ψηφιακών πόρων, πρέπει να είναι ευρέως αποδεκτό και να εφαρμοστεί τόσο από τους παρόχους περιεχομένου όσο και από τις υπηρεσίες ευρετηρίασης ιστού με συστηματικό τρόπο (Safari, 2005). Το DC αν και είναι ένα ευρέως διαδεδομένο πρότυπο μεταδεδομένων, που χρησιμοποιείται από μεγάλο αριθμό ψηφιακών βιβλιοθηκών και αποθετηρίων (Arlitsch & O'Brien, 2012; Palmer et al., 2007; Park & Childress, 2009; Simons & Richardson, 2013), δεν δείχνει να απολαμβάνει στον ίδιο βαθμό την αποδοχή των σχεδιαστών μηχανών αναζήτησης. Μελετώντας λόγω χάρη τις οδηγίες του Google Scholar "Inclusion Guidelines for Webmasters" (Google Scholar, 2010), γίνεται εμφανής η σύσταση χρήσης μετα-ετικετών (meta-tags) περιγραφής DC ως την έσχατη επιλογή. Αυτό πιθανώς να οφείλεται στο γεγονός ότι το DC δεν ερμηνεύεται με σαφήνεια (Arlitsch & O'Brien, 2012; Park & Childress, 2009) και δεν περιλαμβάνει επαρκή και συγκεκριμένα πεδία για την καταχώρηση όλης της πληροφορίας (δεν παρέχονται συγκεκριμένα πεδία ώστε να διακρίνονται σαφώς οι τύποι ακαδημαϊκής εργασίας (προ-έκδοση (pre-print), κεφάλαιο βιβλίου, εργασία από συμμετοχή σε συνέδριο ή διατριβή, κ.λπ.). Τα μεταδεδομένα που αφορούσαν τις πληροφορίες παραπομπής σε τίτλο περιοδικού (την πηγή δηλαδή προέλευσης του άρθρου), τον αριθμό τόμου και έκδοσης και τον αριθμό σελίδων του άρθρου συνήθως εισάγονταν σε ένα μόνο στοιχείο (element), όπως π.χ. dcterms:relation ή dcterms:source στο απλό DC, και δεν υπήρχε καθορισμένη μορφή ή συνέπεια στον τρόπο καταχώρησής τους. Αυτό είχε ως αποτέλεσμα να καθίσταται δύσκολη η ανάλυση και η ευρετηρίαση των βιβλιογραφικών δεδομένων ενός αποθετηρίου από τις μηχανές αναζήτησης όπως το GS. Ωστόσο, μελετώντας τις ενημερωμένες εκδόσεις του DC βλέπουμε, ότι έχει υπάρξει βελτίωση ως προς αυτό το σημείο με την εισαγωγή του στοιχείου dcterms:bibliographicCitation είτε σε μορφή απλού κειμένου (text format), είτε σε δομημένη μορφή (structured format), η οποία είναι προτιμώμενη από τις μηχανές αναζήτησης γενικότερα (Dublin Core Metadata Initiative, n.d.-a). Επομένως το πρόβλημα μετατίθεται στην μη υιοθέτηση των πιο πρόσφατων ενημερώσεων των εξειδικευτών DC (Qualifiers) από τα ίδια τα ψηφιακά αποθετήρια, εντείνοντας τη δυσκολία εντοπισμού και ευρετηρίασης του υλικού τους από τις μηχανές αναζήτησης (Arlitsch & O'Brien, 2012). Εν ολίγοις, οι βιβλιοθήκες και οι οργανισμοί πληροφόρησης εν γένει δεν επικεντρώνονται αρκετά στο να κάνουν τα μεταδεδομένα τους αναγνώσιμα από τις μηχανές. Υπάρχουν επίσης έρευνες, από τις οποίες προκύπτει ότι το σχήμα μεταδεδομένων αυτό καθαυτό δεν παίζει τόσο σημαντικό ρόλο στην ανάκτηση των πληροφοριακών πόρων των ψηφιακών αποθετηρίων από τις μηχανές

αναζήτησης, όσο ορισμένα στοιχεία μεταδεδομένων, και πιο συγκεκριμένα τα στοιχεία τίτλος (title), περιγραφή (description), θέμα (subject) (Yang, 2016). Ανεξάρτητα λοιπόν από το πόσο πολύ επηρεάζει το πρότυπο μεταδεδομένων, ή η διάθεση με έγκυρο τρόπο συγκεκριμένων στοιχείων μεταδεδομένων την ανακάλυψη των ψηφιακών πόρων από τις δυναμικές μηχανές αναζήτησης, έρευνες αποδεικνύουν ότι είναι σημαντικό να προστίθεται/ενσωματώνεται στο HTML κάθε ιστοσελίδας και κάθε ψηφιακού πόρου ένα μικρό μπλοκ κώδικα μεταδεδομένων DC –μια διαδικασία η οποία μπορεί εύκολα να αυτοματοποιηθεί χρησιμοποιώντας τα μεταδεδομένα HTML για να δημιουργηθούν τα μεταδεδομένα του Dublin Core- επιστρέφοντας με αυτόν τον τρόπο καλύτερα αποτελέσματα, πιο γρήγορα και πιο ψηλά στις λίστες αποτελεσμάτων των μηχανών ερευνητικής αναζήτησης όπως το GS, από ιστοσελίδες ή πόρους που δεν το παρέχουν (Arlitsch & O'Brien, 2012; Hoenisch, 2009; Zhang & Dimitroff, 2004).

2.1.2 Google Scholar – Πλεονεκτήματα και μειονεκτήματα

Το Google Scholar (GS) είναι μια μηχανή επιστημονικής αναζήτησης που δημιουργήθηκε από την Google Inc. (τώρα Alphabet) στις 18 Νοεμβρίου 2004. Ο κύριος στόχος του είναι να προσφέρει δωρεάν τον πιο ολοκληρωμένο κατάλογο ερευνητικών εργασιών που είναι διαθέσιμες στον Ιστό. Μέσω μαζικής χρήσης ανιχνευτών και εργαλείων συλλογής, εξάγει μεταδεδομένα από τις πιο σχετικές και αναγνωρισμένες επιστημονικές πηγές στο διαδίκτυο, καθιστώντας το μια από τις πιο ενημερωμένες και εξαντλητικές υπηρεσίες επιστημονικών πληροφοριών στον Ιστό (Ortega, 2014). Παρέχει έναν απλό τρόπο ευρείας αναζήτησης επιστημονικής βιβλιογραφίας, επιτρέποντας με αυτόν τον τρόπο στους χρήστες του να εντοπίζουν σχετικές εργασίες σε όλο τον κόσμο της επιστημονικής έρευνας, σε πολύ μικρό χρόνο και σε πολλές περιπτώσεις με απευθείας ανάκτηση του πλήρους κειμένου των τεκμηρίων αυτών (Hightower & Caldwell, 2010). Το γεγονός αυτό εκτιμάται ιδιαίτερα από τους χρήστες των βιβλιοθηκών, καθώς με μία μόνο αναζήτηση οδηγούνται στην ψηφιακή εκδοχή πλήρους κειμένου του πόρου που τους ενδιαφέρει, δεδομένου ότι πολλές βιβλιοθήκες δεν υποστηρίζονται από μηχανές ενιαίας αναζήτησης πηγών (federated search engine), επιβάλλοντας κατά κάποιο τρόπο στους χρήστες τους την πραγματοποίηση της ίδιας έρευνας για κάθε βάση δεδομένων στην οποία διατηρεί συνδρομή η βιβλιοθήκη, κάτι που από μόνο του δεν αντανάκλα την επικρατούσα προσέγγιση ως προς την εξυπηρέτηση των αναγκών των χρηστών (Jacsó, 2005). Το γεγονός ότι αναπτύχθηκε κάτω από την ομπρέλα μιας εταιρείας όπως η Google, και χρησιμοποίησε μέρος της τεχνολογίας της, οδήγησε σε άμεση αποδοχή από ένα σημαντικό μέρος του ακαδημαϊκού εκδοτικού κόσμου

και από ορισμένους επαγγελματίες και ερευνητές (Orduña-Malea, Martín-Martín, & Delgado López-Cózar, 2017; Ortega, 2014). Επίσης διαθέτει μία εύχρηστη διεπαφή χρηστών και ανταποκρίνεται πολύ γρήγορα στα ερωτήματα που δέχεται, παρέχοντας γρήγορα αποτελέσματα, εξοικονομώντας έτσι πολύτιμο χρόνο ειδικά σε σύγκριση με παρόμοιες αναζητήσεις σε άλλες βάσεις δεδομένων (όπως π.χ. το PubMed) (Giustini & Kamel Boulos, 2013). Ο τρόπος λειτουργίας του είναι παρόμοιος με αυτόν της βασικής μηχανής αναζήτησης της Google, καθώς είναι ένα σύστημα που βασίζεται στην παροχή των καλύτερων δυνατών αποτελεσμάτων σε ερωτήματα χρηστών που έχουν εισαχθεί σε ένα πλαίσιο απλής αναζήτησης (Ortega, 2014). Στην περίπτωση του GS, επιστρέφει αποτελέσματα για εκατομμύρια ακαδημαϊκά/επιστημονικά έργα (περιλήψεις, άρθρα, διατριβές, βιβλία, κεφάλαια βιβλίων, εργασίες συνεδρίων, τεχνικές εκθέσεις ή τα σχέδιά τους, προ-εκδόσεις, διπλώματα ευρεσιτεχνίας, γκρίζα βιβλιογραφία, κ.λπ.) και γενικά οτιδήποτε οι ανιχνευτές (crawlers) του GS εντοπίζουν αυτόματα στον ακαδημαϊκό χώρο, από ακαδημαϊκούς εκδότες, πανεπιστήμια, επιστημονικές και επαγγελματικές εταιρείες, έως και κάθε ιστότοπο που περιέχει ακαδημαϊκό υλικό (Orduña-Malea, Martín-Martín, Ayllón, & Delgado López-Cózar, 2016). Όπως και με την Google, τα αποτελέσματα που ανακτώνται για ένα συγκεκριμένο ερώτημα ταξινομούνται από έναν αλγόριθμο, ο οποίος λαμβάνει υπόψη του ένα μεγάλο αριθμό μεταβλητών, όπως τον τόπο έκδοσης/δημοσίευσης της εργασίας, το όνομα του συγγραφέα, τη συχνότητα αναφοράς του έργου σε άλλες επιστημονικές εργασίες, κ.λπ., χωρίς όμως να γίνεται γνωστή η ακριβής λειτουργία αυτού του αλγορίθμου και η βαρύτητα κάθε μεταβλητής. Ωστόσο, αρκετές εμπειρικές μελέτες έχουν δείξει ότι ο αριθμός των παραπομπών που έλαβε μία επιστημονική εργασία είναι ένας από τους βασικούς παράγοντες κατάταξης της στα αποτελέσματα του GS (Beel & Gipp, 2009; Martín-Martín, Orduña-Malea, Ayllón, & Delgado López-Cózar, 2014; Martín-Martín, Orduña-Malea, Harzing, & Delgado López-Cózar, 2017). Ένα ακόμη βασικό χαρακτηριστικό του GS είναι ότι η όλη διαδικασία ανίχνευσης, ανάλυσης βιβλιογραφικών μεταδεδομένων, εξαγωγής βιβλιογραφικών αναφορών κ.λπ. γίνεται με πλήρως αυτοματοποιημένο τρόπο (Orduña-Malea et al., 2017). Μεγάλο του πλεονέκτημα είναι η πραγματοποίηση αναζητήσεων στα ευρετήρια που δημιουργούνται από το πλήρες κείμενο του αρχικού πόρου ή μέρους αυτού, ακόμη κι αν η πρόσβαση περιορίζεται μόνο σε ένα μικρό απόσπασμα του πλήρους κειμένου, και όχι αποκλειστικά από τις βιβλιογραφικές εγγραφές, τις περιλήψεις και τους θεματικούς όρους του πόρου (όταν και εφόσον παρέχονται) (Jacsó, 2005). Με το βλέμμα στραμμένο προς την ικανοποίηση της ανάγκης της επιστημονικής κοινότητας για αξιολόγηση της έρευνας τόσο φυσικών προσώπων όσο και

ακαδημαϊκών ιδρυμάτων – ερευνητικών φορέων, το GS εισήγαγε δύο νέες υπηρεσίες, το Google Scholar Metrics for Publications (GSM) τον Απρίλιο του 2012 (Jacsó, 2012b) και το Google Scholar Author Citation Tracker (GSACT) το 2011 (Jacsó, 2012a). Το GSM παρέχει χρήσιμες πληροφορίες σχετικά με τη χρήση και τη σημασία των ακριβών περιοδικών και θα μπορούσε να διαδραματίσει σημαντικό ρόλο στην επιλογή υλικού και στη διαχείριση των συλλογών γενικότερα στο πλαίσιο της ανάπτυξης των συλλογών της βιβλιοθήκης ενός φορέα/ακαδημαϊκού ιδρύματος κ.λπ., δεδομένου ότι τα κόστη συνδρομών των επιστημονικών περιοδικών συνεχίζουν να αυξάνονται, ενώ οι προϋπολογισμοί των βιβλιοθηκών μειώνονται σημαντικά. Ακόμη οι αξιολογήσεις της παραγωγικότητας της έρευνας και της επιρροής των μελών των πανεπιστημιακών σχολών και της ακαδημαϊκής κοινότητας γενικότερα, που στοχεύουν σε προσωπική εξέλιξη, αναγνώριση, υποτροφίες, χορηγίες, κ.λπ. βασίζονται κυρίως στην επιστημονική απήχηση και φήμη των περιοδικών εκδόσεων και πρακτικών συνεδρίων στα οποία οι δημιουργοί συμμετέχουν με ανακοινώσεις/δημοσιεύσεις. Αντίστοιχα το GSC επιτρέπει στους εγγεγραμμένους χρήστες να δημιουργούν και να επεξεργάζονται το επιστημονικό προφίλ τους και κάποιους βιβλιομετρικούς δείκτες δωρεάν, όπως τον h-index και τον i10-index (Παπαβλασόπουλος, 2015). Η νέα ενότητα προσφέρει επίσης μερικές βασικές επιλογές, όπως ταξινόμηση των αποτελεσμάτων κατά το έτος δημοσίευσής τους, τον τίτλο και τις παραπομπές που έχουν λάβει. Παρατηρήθηκε επίσης αξιοσημείωτη εκκαθάριση εγγραφών της βάσης του GS, καθώς οι εγγεγραμμένοι στην υπηρεσία αυτή χρήστες μπορούν να διορθώσουν σφάλματα που εντοπίστηκαν στα ακαδημαϊκά προφίλ τους, να προσθέσουν εγγραφές που δεν συλλέχθηκαν από τα προγράμματα ανίχνευσης του GS, να συγχωνεύσουν τυχόν διπλοεγγραφές και να συγκεντρώσουν τις αναφορές τους κάτω από τη σωστή καταχώρηση φυσικού προσώπου. Η διαδικασία υπολογισμού και εξαγωγής των μετρικών ενημερώνεται αυτόματα, κάθε φορά που οι ανιχνευτές του GS εντοπίζουν νέες παραπομπές για το έργο ενός συγκεκριμένου συγγραφέα (Shanks & Arlitsch, 2016). Μειονέκτημα παραμένει η συμπερίληψη των αυτοαναφορών (Belew, 2005; Majeed et al., 2019; Σελίμης & Σιμιακάκης, 2016), προσθέτοντας θόρυβο στα βιβλιομετρικά αποτελέσματα συγγραφέων και τίτλων. Παρ' όλο που τα οφέλη είναι αρκετά, οφείλουμε να αναφέρουμε ότι και οι δύο αυτές υπηρεσίες του GS χρήζουν ακόμη βελτιώσεων, ώστε να μπορούν να αποδώσουν παρέχοντας καλύτερα και περισσότερο αξιόπιστα αποτελέσματα, εξυπηρετώντας με τον τρόπο αυτόν τον σκοπό της δημιουργίας τους (Jacsó, 2012a, 2012b). Το GS είναι μία μηχανή ερευνητικής αναζήτησης η οποία σε γενικές γραμμές χαίρει της χρήσης και της ικανοποίησης των χρηστών της (Rieger, 2009). Μια ιδιαιτερότητα της παρουσίασης

ακαδημαϊκών εργασιών στο GS είναι η απευθείας παροχή του συνδέσμου του ζητούμενου PDF αρχείου. Αυτό είναι βέβαια πολύ βολικό για τους χρήστες, καθώς τους οδηγεί απευθείας στο περιεχόμενο που αναζητούν, αγνοώντας όμως οποιοδήποτε περιεχόμενο το οποίο μπορεί να διατίθεται από την HTML εμφάνιση του αποθετηρίου. Με άλλα λόγια, τα μεταδεδομένα, τα ιδρυματικά λογότυπα και άλλες πληροφορίες που εμφανίζονται κανονικά στους χρήστες στην HTML εμφάνιση του αποθετηρίου χάνονται, εκτός εάν αυτά εισαχθούν στο ίδιο το PDF. Η συγκεκριμένη πρακτική μπορεί να επηρεάσει την αναφορά στατιστικών ως προς τις επισκέψεις (visitation statistics) μέσω των λογισμικών ανάλυσης ιστοτόπων, που χρησιμοποιούν ετικέτες σελίδων (page tagging). Για παράδειγμα, το Google Analytics απαιτεί έναν κώδικα παρακολούθησης που έχει εισαχθεί στο HTML κάθε σελίδας ενός δεδομένου ιστοτόπου για τη συλλογή στατιστικών στοιχείων. Κάθε φορά που αυτή η σελίδα εμφανίζεται σε ένα πρόγραμμα περιήγησης ιστού, μετράται ως επίσκεψη από το Google Analytics. Με το διαχωρισμό όμως του αρχείου PDF από την εμφάνιση HTML, δεν επιτρέπεται η μέτρηση της επίσκεψης, καθώς ο απαιτούμενος κώδικας παρακολούθησης δεν εκτελείται όταν το PDF καλείται απευθείας. Αυτό το πρόβλημα μπορεί να ξεπεραστεί εάν εκτελέσει ο διακομιστής ιστού ένα σενάριο (script) PHP που να περιέχει τον κώδικα παρακολούθησης πριν από την προβολή του ζητούμενου PDF. Ωστόσο, θεωρείται δύσκολο να πραγματοποιηθεί κάτι τέτοιο από πολλούς διαχειριστές αποθετηρίων, καθώς είναι πιθανό να αγνοούν την παράμετρο αυτή, ή ακόμη είναι πιθανό να αγνοούν το γεγονός ότι τα στατιστικά στοιχεία επισκέψεων και λήψεων του αποθετηρίου τους ενδέχεται να μην συμπεριλαμβάνονται στις μετρήσεις, ως αποτέλεσμα της πρακτικής εμφάνισης στοιχείων του GS (Arlitsch & O'Brien, 2012). Κάνοντας επίσης αναζήτηση στο GS σύμφωνα με τον τελεστή "ιστότοπος" (site operator), πρέπει να γνωρίζουμε ότι δεν επιστρέφονται στα αποτελέσματα όλα τα τεκμήρια που έχει ευρετηριάσει το GS από ένα αποθετήριο, πράγμα που σημαίνει ότι δεν είναι πολύ εξαντλητικό. Ακόμη, το σύστημα ομαδοποίησης των πολλαπλών εκδόσεων μίας εργασίας λειτουργεί με τέτοιο τρόπο, ώστε μόνο μια έκδοση να λαμβάνεται ως η κύρια. Αυτή η διαδικασία γίνεται αυτόματα, παρόλο που οι συγγραφείς μπορούν να επιλέξουν και χειροκίνητα ακόμη ποια θεωρούν την κύρια έκδοση του έργου τους. Η εντολή "site" επιστρέφει μόνο δεδομένα που σχετίζονται με την κύρια έκδοση. Στην περίπτωση που μία εργασία φιλοξενείται σε διαφορετικές πλατφόρμες (π.χ. περιοδικό και αποθετήριο), εάν έχει θεωρηθεί ως κύρια έκδοση αυτή που δημοσιεύτηκε στο περιοδικό, ο τελεστής "ιστότοπος" που εφαρμόζεται στο αποθετήριο δεν θα μετρήσει το τεκμήριο και αντίστροφα, ακόμη κι αν η εργασία ευρετηριάζεται κανονικά και στις δύο πλατφόρμες (Arlitsch & O'Brien, 2012; Orduña-Malea & Delgado López-Cózar, 2015). Ένα σημαντικό

μειονέκτημα του GS είναι η μυστικότητα που διατηρείται γύρω από το επιστημονικό πεδίο που καλύπτεται και του αριθμού των πληροφοριακών πόρων που ευρετηριάζονται από αυτό (Gusenbauer, 2019; Jacsó, 2012c; Orduña-Malea et al., 2016). Ωστόσο συγκρίνοντάς το με το Web of Science και την Scopus, βάσεις δεδομένων οι οποίες παρέχουν αξιόπιστα και ακριβή δεδομένα για τα περιοδικά που καλύπτουν, το GS παρέχει πολύ καλύτερη κάλυψη της έρευνας στα πεδία των κοινωνικών και ανθρωπιστικών επιστημών, όπου εκείνες υστερούν. Γενικά φαίνεται το GS να παρέχει αντίστοιχα ποσοστά κάλυψης σε όλα τα επιστημονικά πεδία, παρ' όλο που αναφέρονται σημαντικά προβλήματα αξιοπιστίας και εγκυρότητας των δεδομένων που προσφέρει (Mingers & Meyer, 2017). Πολλές προσπάθειες έχουν γίνει μέχρι σήμερα προτείνοντας διάφορες μεθόδους τόσο για την εκτίμηση του τρέχοντος μεγέθους του GS, όσο και προσδιορισμού της εγκυρότητας, της ακρίβειας και της αξιοπιστίας του. Υπάρχουν ωστόσο διάφοροι παράγοντες που δεν επιτρέπουν την ακριβή εκτίμηση του όγκου των πληροφοριακών πόρων που ευρετηριάζονται από αυτό. Σε κάθε περίπτωση λοιπόν θα πρέπει να ακολουθήσει περαιτέρω έρευνα (Gusenbauer, 2019).

2.1.3 Google Scholar – Μεταδεδομένα

Οι βιβλιοθηκονόμοι πιστεύουν πολύ στα πρότυπα. Κατά τη δημιουργία των ψηφιακών αποθετηρίων τους λοιπόν τα ακολούθησαν πιστά για τη σάρωση, τη δημιουργία μεταδεδομένων, τη συλλογή (harvesting) και τις υπηρεσίες Ιστού μεταξύ άλλων. Από τις μηχανές αναζήτησης ωστόσο, δεν απαιτείται η τήρηση προτύπων. Η συλλογή μεταδεδομένων στο GS δεν πραγματοποιείται μέσω χρήσης του πρωτοκόλλου OAI-PMH, καθώς έχει πάψει να το υποστηρίζει από το 2008 (Mueller, 2008). Για τον σκοπό αυτό χρησιμοποιεί προγράμματα ανίχνευσης, τα οποία ακολουθούν τους συνδέσμους κάθε ψηφιακού αντικείμενου, μια διαδικασία που διευκολύνεται σε μεγάλο βαθμό από την υποβολή χαρτών ιστοτόπων (sitemap), που λειτουργούν ως επίσημες προσκλήσεις και οδηγοί, αποκαλύπτοντας την προτιμώμενη διεύθυνση URL για καθένα από αυτά τα ψηφιακά αντικείμενα. Οι ανιχνευτές "συγκομίζουν" μεταδεδομένα και άλλες πληροφορίες σχετικές με τους πόρους και τα αποστέλλουν στη μηχανή αναζήτησης. Εκεί αναλύονται από αλγορίθμους, οι οποίοι έχοντας λάβει υπόψη τους πολλούς παράγοντες, αποφασίζουν εάν θα προσθέσουν τα μεταδεδομένα αυτά στο ευρετήριο της μηχανής αναζήτησης (Arlitsch & O'Brien, 2012). Ωστόσο οι προγραμματιστές του GS επέλεξαν να αγνοήσουν τα ελεύθερα διαθέσιμα μεταδεδομένα που προσφέρονται από τους ακαδημαϊκούς εκδότες και από άλλες υπηρεσίες ευρετηρίασης, αφήνοντας τους ανεπαρκώς εκπαιδευμένους αναλυτές του

(parsers) να τα ανιχνεύσουν και να τα αναγνωρίσουν, χωρίς να εξασφαλίζεται επιτυχημένο αποτέλεσμα (Jacsó, 2009b, 2010). Μία τέτοια πρακτική λειτούργησε αρκετά καλά στην περίπτωση της βασικής μηχανής αναζήτησης της Google, ως προς την ευρετηρίαση του μη δομημένου Ιστού, εστιάζοντας περισσότερο στον προσδιορισμό της διεύθυνσης URL και του τίτλου της ιστοσελίδας από τις εντολές HTML, κατά την ευρετηρίαση των σελίδων χωρίς ουσιαστικά την ύπαρξη μεταδεδομένων, και πριν από την εισαγωγή του προτύπου Dublin Core (Jacsó, 2010). Δυστυχώς η πρακτική αυτή δεν έφερε ανάλογα αποτελέσματα κατά την ευρετηρίαση των πόρων των ψηφιακών αποθετηρίων από το GS, αν και έχουν πραγματοποιηθεί αρκετές βελτιώσεις στις επιδόσεις του μέχρι σήμερα (Jacsó, 2009b, 2010, 2011). Αποτέλεσμα αυτής της διαδικασίας είναι η δημιουργία τεράστιων σφαλμάτων κατά τον εντοπισμό και την ανάλυση των διαφόρων στοιχείων μιας βιβλιογραφικής εγγραφής. Από τις χαρακτηριστικές συνέπειες αυτού, ήταν η ανακάλυψη ότι ο συγγραφέας "I Introduction" ήταν ο πιο παραγωγικός σύμφωνα με το GS, με περισσότερες από 40.000 δημοσιεύσεις (Jacsó, 2006) και ότι ο "F Password" είχε τις περισσότερες αναφορές (Jacsó, 2008b, 2012c). Η λανθασμένη λειτουργία των αναλυτών του GS έχει οδηγήσει σε δημιουργία λανθασμένου έτους δημοσίευσης των πόρων είτε χρησιμοποιώντας τμήμα της διεθνούς αριθμοδότησης της περιοδικής έκδοσης (ISSN) (Jacsó, 2008a), είτε κάνοντας χρήση της αρίθμησης των τόμων και των τευχών των περιοδικών εκδόσεων, ή ακόμη και μέρους τηλεφωνικών αριθμών (Jacsó, 2010). Επίσης κάνοντας λανθασμένη χρήση στοιχείων από τις επιλογές μενού (menu options) των ιστοσελίδων, από τις επικεφαλίδες ενοτήτων και λογοτύπων από τίτλους περιοδικών έχει οδηγήσει σε δημιουργία διάφορων ονομάτων μη υπαρκτών συγγραφέων (Jacsó, 2009a). Έχουν σημειωθεί επίσης περιπτώσεις στις οποίες οι αναλυτές του GS μπερδεύουν ονόματα συγγραφέων με άλλο περιεχόμενο (Jacsó, 2008a), ή δεν μπορούν να εντοπίσουν δημοσιεύσεις στις οποίες το όνομα του συγγραφέα περιλαμβάνει ακολουθίες χαρακτήρων που δεν ανήκουν σε ένα παραδοσιακό στοιχειοθετημένο είδος π.χ. περιπτώσεις που ο συγγραφέας έχει χρησιμοποιήσει το LaTeX για να συντάξει την μελέτη του (Harzing & van der Wal, 2008). Προβλήματα επίσης έχουν παρατηρηθεί ακόμη και στην κωδικοποίηση χαρακτήρων, όπως τους ASCII, όπου μπορεί να έχουν δημιουργήσει ή να δημιουργήσουν προβλήματα και σφάλματα (ασυνέπειες στα ονόματα των συγγραφέων και εσφαλμένη απόδοση παραπομπών), ειδικά σε συγγραφείς των οποίων τα ονόματα είναι γραμμένα σε μη λατινικό αλφάβητο (Belew, 2005). Αυτά κι ακόμη περισσότερα προβλήματα έχουν εντοπιστεί κατά καιρούς και σχετίζονται άμεσα με τα μεταδεδομένα, τη διαχείρισή τους και κυρίως την έλλειψη ποιοτικού ελέγχου από την πλευρά του GS (Jacsó, 2010). Αποτέλεσμα όλων αυτών των σφαλμάτων είναι η δημιουργία

σύγχυσης κατά την ανακάλυψη των πόρων και η διαστρέβλωση των βιβλιομετρικών δεικτών σε επίπεδο φυσικών προσώπων, συλλογικών οργάνων και περιοδικών εκδόσεων (Jacsó, 2012c). Τα σφάλματα ανάλυσης θεωρούνται από τα πιο σημαντικά, καθώς η εμφάνισή τους προκαλεί αλυσιδωτή αντίδραση, που μπορεί να δημιουργήσει και να μεταδώσει νέα σφάλματα σε άλλες επιστημονικές εργασίες και σε εξαιρετικά μεγάλη κλίμακα. Ως εκ τούτου, μια εφαρμογή μπορεί να προσδιορίσει τα διάφορα μέρη μιας βιβλιογραφικής εγγραφής (συγγραφέας, τίτλος, πηγή, τόμος, αριθμός, σελιδοποίηση) τόσο ενός παρατιθέμενου έργου (μεταδεδομένα), όσο και ενός παρατιθέμενου έργου (βιβλιογραφική αναφορά που περιέχεται στην ενότητα βιβλιογραφίας ενός ακαδημαϊκού έργου) (Orduña-Malea et al., 2017). Εάν παρεχόταν στους ανιχνευτές του GS μια αποτελεσματική μέθοδος πρόσβασης στα επιστημονικά έργα και προσφέρονταν ακριβείς βιβλιογραφικές πληροφορίες στις ετικέτες κεφαλίδας της σελίδας HTML κάθε πόρου, με τη χρήση αποδεκτών σχημάτων μεταδεδομένων, τα προβλήματα αυτά θα αντιμετωπίζονταν σε πολύ μεγάλο βαθμό. Τα προτεινόμενα πρότυπα μεταδεδομένων από το GS είναι τα Highwire Press, Eprints, BE Press και PRISM (Google Scholar, 2010). Το Highwire Press αναπτύχθηκε από Τμήμα του Πανεπιστημίου του Στάνφορντ. Είναι ένα σχήμα για περιγραφή άρθρων περιοδικών, του οποίου τις ετικέτες περιγραφής επέκτεινε το GS για να καλύψει επιπλέον ακαδημαϊκούς τύπους έργων, όπως έγγραφα εργασίας, διατριβές, χειρόγραφα, εργασίες συνεδρίων, βιβλία και κεφάλαια βιβλίων και έχει καλή εφαρμογή (Arlitsch & O'Brien, 2012; Verrelli, 2018). Σε γενικές γραμμές η εφαρμογή του είναι αρκετά ικανοποιητική. Η χρήση μετα-ετικετών Eprints περιορίζεται σε υλοποιήσεις λογισμικού αποθετηρίων Eprints. Υπάρχουν πολύ λίγα χαρακτηριστικά ονομάτων (name attributes) στο χώρο ονομάτων (name space) Eprints - προφανώς μόνο δέκα. Επίσης, δεν εντοπίζεται επίσημο πρότυπο (Verrelli, 2018). Αντίστοιχα και οι μετα-ετικέτες BE Press χρησιμοποιούνται σε υλοποιήσεις λογισμικού αποθετηρίων του Berkley Electronic Press SelectedWorks (Digital Commons). Τα μεταδεδομένα BE Press δεν φαίνεται να χρησιμοποιούνται πολύ τακτικά. Συνήθως, τα χαρακτηριστικά ονομάτων στο χώρο ονομάτων BE Press είναι πρακτικά αντίγραφα από αυτά του χώρου ονομάτων citation (Highwire Press), με χρήση του προθέματος bepress. Η διαφορά απλά έγκειται στην ύπαρξη λιγότερων χαρακτηριστικών από αυτά του Highwire Press. Αντί να επεκτείνει το περιεχόμενο του χώρου ονομάτων BE Press, φαίνεται ότι η BE Press είναι ανοιχτή στην υιοθέτηση ετικετών ή / και συμβάσεων από άλλα συστήματα μεταδεδομένων, όπως το Dublin Core, το OpenURL, το PubMed (Verrelli, 2018). Η χρήση μετα-ετικετών PRISM γίνεται από εκδοτικούς οίκους που χρησιμοποιούν ήδη το λογισμικό PRISM στη ροή

εργασίας τους. Το PRISM αναμένεται να χρησιμοποιηθεί με ονόματα μεταδεδομένων του Dublin Core στο DC name space. Αυτό είναι ιδιαίτερα προφανές παρατηρώντας ότι το PRISM χρησιμοποιεί το DC.creator για να παρέχει μεταδεδομένα που προσδιορίζουν έναν δημιουργό, αντί να εισάγει ένα ανταγωνιστικό χαρακτηριστικό ονόματος (name attribute). Φαίνεται ότι το PRISM έχει σχεδιαστεί για να συμπληρώνει το Dublin Core και όχι για να το αντικαθιστά. Το PRISM παρέχει επίσης χρήση ελεγχόμενων λεξιλογίων (Verrelli, 2018). Θεωρητικά η μετατροπή μεταδεδομένων σε σχήματα Eprints, PRISM και BE Press είναι πιθανό να έχει επίσης θετικό αποτέλεσμα, αν και αυτός ο ισχυρισμός απαιτεί επιπρόσθετες δοκιμές (Arlitsch & O'Brien, 2012). Ωστόσο, όπως ήδη αναφέρθηκε δεν απευθύνονται σε μεγάλο αριθμό αποθετηρίων, καθώς η χρήση τους περιορίζεται σε υλοποιήσεις συγκεκριμένων λογισμικών. Το Dublin Core αν και είναι το πιο διαδεδομένο πρότυπο ψηφιακών αποθετηρίων (Arlitsch & O'Brien, 2012; Palmer et al., 2007; Park & Childress, 2009; Simons & Richardson, 2013), στις οδηγίες του Google Scholar "Inclusion Guidelines for Webmasters" (Google Scholar, 2010), γίνεται εμφανής η σύσταση χρήσης μετα-ετικετών περιγραφής DC ως έσχατη επιλογή. Παρ' όλα αυτά η ενσωμάτωση στην σελίδα HTML κάθε πόρου, ενός μικρού μπλοκ κώδικα σύμφωνα με τις ενημερωμένες εκδόσεις του DC, πιθανώς να προσέφερε μία καλύτερη λύση, στην περιγραφή των πόρων με κατάλληλο τρόπο (χρήση ελεγχόμενων λεξιλογίων, κ.λπ.) και είναι κάτι που πρέπει να ελεγχθεί και να δοκιμαστεί. Τα μεταδεδομένα με αυτόν τον τρόπο θα γίνονταν αναγνώσιμα και ανιχνεύσιμα από τις μηχανές αναζήτησης όπως το GS, απαλλαγμένα από τα περιττά λάθη και οι συγγραφείς δεν θα στερούνταν της αναγνώρισης και θα απολάμβαναν τη θέση που τους αξίζει στους βιβλιομετρικούς δείκτες. Για να επιτευχθεί κάτι τέτοιο είναι απαραίτητη η δημιουργία σαφών οδηγιών για την εφαρμογή των στοιχείων DC και την επιλογή εκείνων των στοιχείων που πρέπει απαραίτητα να συμπληρώνονται, ώστε η καταχώρησή τους από όλους να γίνεται με ακρίβεια και συνέπεια (Ho & Stokes, 2019). Σε κάθε περίπτωση πρέπει οπωσδήποτε να πραγματοποιηθούν επιπρόσθετες δοκιμές για ενδεχόμενη καλύτερη απόδοση των ενημερωμένων μετα-ετικετών DC. Ακόμη κι αν τα αποτελέσματα δεν αποδώσουν τα μέγιστα, η μετατροπή τους π.χ. σε μετα-ετικέτες Highwire Press θα μπορούσε να επιτευχθεί σχετικά εύκολα και να έχουμε επιτέλους τα ικανοποιητικά αποτελέσματα που αναζητάμε.

2.2 Σχετικές προσπάθειες – έρευνες

Οι σχετικές εργασίες που παρουσιάζονται σε αυτήν την ενότητα αφορούν κυρίως στον εντοπισμό του προβλήματος και στην προτεινόμενη αντιμετώπισή του. Η επιτακτική ανάγκη

για αύξηση της ορατότητας, της ανακάλυψης και της ανάκτησης του περιεχομένου των ψηφιακών αποθετηρίων από τις δυναμικές μηχανές επιστημονικής αναζήτησης όπως το GS οδήγησε σε προσπάθειες ταυτοποίησης του προβλήματος. Παρατηρείται λοιπόν η πραγματοποίηση εργασιών βελτίωσης και αύξησης της ορατότητας του υλικού αποθετηρίων από το GS, τόσο σε διεθνές όσο και σε τοπικό επίπεδο. Μέσα από τις εργασίες αυτές ενισχύεται η αρχική υπόθεση, ότι η μη συμμόρφωση του προτύπου μεταδεδομένων που χρησιμοποιείται από τα περισσότερα ψηφιακά αποθετήρια (Dublin Core), με τα προτεινόμενα από το GS πρότυπα (Highwire Press, Eprints, BE Press και PRISM) (Acharya, 2015; Google Scholar, 2010), δυσχεραίνει την ευρετηρίαση του περιεχομένου των ψηφιακών αποθετηρίων από τις μηχανές ερευνητικής αναζήτησης. Πραγματοποιώντας αντιστοίχιση των στοιχείων του DC σε ετικέτες Highwire Press (προτεινόμενο από το GS πρότυπο), ώστε να αποδίδονται τα μεταδεδομένα ταυτόχρονα και στα δύο πρότυπα, τα αποθετήρια οδηγήθηκαν σε αυξημένη ευρετηρίαση του περιεχομένου τους από το GS. Επίσης, διατήρησαν την ομοιογένεια, την τυποποίηση και την ποιότητα των μεταδεδομένων τους, χρησιμοποιώντας ελεγχόμενα λεξιλόγια για την τεκμηρίωση του υλικού τους όπως και προηγουμένως (LCSH, λέξεις κλειδιά, Dublin Core κ.λπ.), χωρίς να χάσουν τη λειτουργικότητά τους με άλλα συστήματα. Θα ήταν καλό να αναφερθεί ότι παρέχονται οι αντιστοιχίσεις (mappings) σε μεταδεδομένα Highwire Press για επτά διαφορετικά είδη υλικού (προ-εκδόσεις, άρθρα περιοδικών, βιβλία, κεφάλαια βιβλίων, διδακτορικές διατριβές, διπλωματικές εργασίες και έγγραφα εργασίας (working papers)), τα οποία είναι πράγματι πολύ βοηθητικά (Arlitsch & O'Brien, 2012). Μέσω των εργασιών αυτών γίνεται αντιληπτό ότι για την επίτευξη ευρείας προβολής των αποθετηρίων και ικανοποιητικής ευρετηρίασης του περιεχομένου τους από τις μηχανές αναζήτησης, απαιτείται η παροχή στους ανιχνευτές των μηχανών αναζήτησης αποτελεσματικότερων μεθόδων πρόσβασης στο ακαδημαϊκό περιεχόμενο και περισσότερο αποδεκτών προτύπων μεταδεδομένων, τα οποία να προσφέρουν ακριβείς βιβλιογραφικές πληροφορίες στις ετικέτες κεφαλίδας της σελίδας HTML κάθε πόρου. Σε κάθε περίπτωση είναι απαραίτητη περαιτέρω έρευνα και δοκιμή, καθώς η επιστήμη της πληροφόρησης εξελίσσεται και τα ίδια τα πρότυπα ενημερώνονται συνεχώς παρέχοντας ενδεχομένως άλλες δυνατότητες και μεγαλύτερη ευελιξία. Επίσης, δεν αποκλείεται η ύπαρξη και άλλων παραγόντων οι οποίοι μπορεί να εμποδίζουν την αύξηση της ορατότητας των ψηφιακών πόρων των αποθετηρίων από τις μηχανές αναζήτησης και οι οποίοι θα πρέπει εξίσου να αντιμετωπιστούν από τα ψηφιακά αποθετήρια (Acharya, 2015; Arlitsch & O'Brien, 2012; Google Scholar, 2010; Orduña-Malea & Delgado López-Cózar, 2015). Οι δύο εργασίες που παρατίθενται είναι οι εξής: «Η μετεξέλιξη ενός αποθετηρίου

κοινωνικών επιστημών: το παράδειγμα της Πάνδημου σε περιβάλλον Fedora» (Καμάτσος, Σωτηροπούλου, Θεοφιλάτου, Μπογιατζή, & Κακάλη, 2013) και «Invisible Institutional Repositories : addressing the low indexing ratios of IRs in Google Scholar» (Arlitsch & O'Brien, 2012).

Κεφάλαιο 3. Μεθοδολογία

3.1 Στόχος της Έρευνας και Μεθοδολογία

Ο στόχος αυτής της διπλωματικής εργασίας είναι να εξετάσει εάν το πρότυπο μεταδεδομένων επηρεάζει τον βαθμό ευρετηρίασης και ανάκτησης του περιεχομένου των ΨΑ από τις μηχανές αναζήτησης όπως το GS ή όχι, κι αν ναι με ποιες άλλες μεταβλητές συσχετίζεται. Επιπροσθέτως, η μελέτη αυτή έχει στόχο να αναδείξει τις τάσεις και τις πρακτικές δόμησης των ΨΑ παγκοσμίως κυρίως ως προς την επιλογή του προτύπου μεταδεδομένων και των πεδίων που θεωρούνται περισσότερο σημαντικά για περιγραφή, τις αδυναμίες που τυχόν υπάρχουν, ή τις κατάλληλες ενέργειες στις οποίες έχουν ήδη προβεί οι επαγγελματίες πληροφόρησης που εργάζονται σε αυτά, ώστε να πετύχουν καλύτερα αποτελέσματα στην ευρετηρίαση και ορατότητα των πόρων που περιέχονται σε αυτά. Σε αυτήν την ενότητα παρουσιάζεται η μεθοδολογία έρευνας που ακολουθήθηκε για να οδηγηθούμε στα ευρήματα που ακολουθούν στο επόμενο κεφάλαιο. Η μελέτη αυτή είναι μία ποσοτική έρευνα, επομένως αποβλέπει στην ανάλυση των δεδομένων χρησιμοποιώντας αριθμητικά στοιχεία με σκοπό την επιβεβαίωση ή την απόρριψη των ερευνητικών υποθέσεων. Έχει ως στόχο τη γενίκευση, καθώς επιτρέπει τη συσχέτιση δύο ή περισσότερων χαρακτηριστικών/μεταβλητών για μεγάλο αριθμό δείγματος με σκοπό την ανάδειξη γενικών τάσεων.

3.2 Συγκρότηση Δείγματος & Συλλογή Δεδομένων

Κατά την προετοιμασία της έρευνας ο καθορισμός του πληθυσμού και η επιλογή του ερευνητικού δείγματος ήταν μια βασική και δύσκολη συνάμα διαδικασία. Επειδή όμως ο πληθυσμός των ΨΑ παγκοσμίως είναι πολύ μεγάλος (περισσότερα από 3000 σε όλο τον κόσμο), σκεφτήκαμε να επιλέξουμε ένα τυχαίο δείγμα σε έναν αριθμό της τάξης των περίπου τετρακοσίων (400) ΨΑ. Επιλέξαμε λοιπόν τριακόσια πενήντα (350) ΨΑ του εξωτερικού, πραγματοποιώντας τυχαία επιλογή από την ιστοσελίδα Ranking Web of Repositories (<https://repositories.webometrics.info/en/node/30>) 9th ed. April 2020 και αντίστοιχα πενήντα επτά (57) ελληνικά και κυπριακά ΨΑ από τις ιστοσελίδες Open DOAR (Directory of Open Access Repositories) - Greece (https://v2.sherpa.ac.uk/view/repository_by_country/Greece.html) και Open DOAR

(Directory of Open Access Repositories) - Cyprus (https://v2.sherpa.ac.uk/view/repository_by_country/Cyprus.html). Ο συνολικός αριθμός δείγματος (n) ανήλθε στα τετρακόσια επτά (407) ΨΑ. Η επιλογή αυτή έγινε δίνοντας έμφαση στην ελληνική πραγματικότητα και λαμβάνοντας υπόψη, για τα διεθνή κυρίως αποθετήρια, την δυναμική τους σε σχέση με τα τεκμήρια που ευρετηριάζονται στο Google Scholar (τον αριθμό κατάταξής τους), τη γεωγραφική τους θέση και τον αντίκτυπό τους αναφορικά με την επιστημονική και εκδοτική παραγωγή των φορέων που αντιπροσωπεύουν. Επομένως, έγινε μία προσπάθεια επιλογής και ανάδειξης ψηφιακών αποθετηρίων που τυγχάνουν υψηλής, μέτριας και χαμηλής ευρετηρίασης στο Google Scholar, έτσι ώστε να διερευνηθούν οι πρακτικές που ακολουθούνται, το πρότυπο μεταδεδομένων, το λογισμικό των αποθετηρίων αυτών, τυχόν αδυναμίες, όπως επίσης οι πιθανές συσχετίσεις που θα προέκυψαν. Το συγκεκριμένο δείγμα μας επέτρεψε να αξιολογήσουμε και να εξετάσουμε τις απόψεις των ατόμων που δημιουργούν, δομούν με πληροφορία και εργάζονται καθημερινά σε αρκετά ΨΑ. Μέσω του δείγματος αυτού μπορέσαμε να εξετάσουμε τις τάσεις, τα πρότυπα, τις πρακτικές που ακολουθούνται, τα προβλήματα που αντιμετωπίζονται και τις ενέργειες που προσπαθεί κάθε επαγγελματίας σε ΨΑ να πραγματοποιήσει, ώστε να επιτύχει μεγαλύτερη προβολή της δραστηριότητας του φορέα που αντιπροσωπεύει και εν γένει μεγαλύτερη προβολή του περιεχομένου του ΨΑ του στην επιστημονική κοινότητα. Μετά την επιλογή του δείγματος, προχωρήσαμε σε εντοπισμό των διευθύνσεων ηλεκτρονικού ταχυδρομείου από τις επίσημες ιστοσελίδες των ΨΑ (επικοινωνία/contact), με σκοπό να διανείμουμε ηλεκτρονικά τα ερωτηματολόγια. Τα ερωτηματολόγια αποτελούνταν από 23 ερωτήσεις, οι οποίες βασίστηκαν στη βιβλιογραφική επισκόπηση (βλ. Παραρτήματα Α' και Β'). Η αποστολή των ερωτηματολογίων έγινε σε στάδια. Η πρώτη αποστολή απευθυνόμενη στα ελληνικά αποθετήρια έγινε στις 15/7/2020, με δεύτερη αποστολή ως «ευγενική υπενθύμιση» στις 27/7/2020. Αντίστοιχα, η πρώτη αποστολή στα ΨΑ του εξωτερικού έγινε στις 23/7/2020 και συνεχίστηκε τμηματικά καθ' όλο τον μήνα Αύγουστο, με τελική αποστολή μηνύματος «ευγενικής υπενθύμισης» στις 24/8/2020 και καταληκτική ημερομηνία τις 8/9/2020. Από το σύνολο των μηνυμάτων ηλεκτρονικού ταχυδρομείου που απεστάλησαν τριάντα τέσσερα (34) μηνύματα δεν παραδόθηκαν ποτέ. Παρ' όλα αυτά στο σημείο αυτό θεωρώ ότι πρέπει να αναφερθεί ότι ενώ τα ερωτηματολόγια ήταν ανώνυμα, πολλοί ήταν οι συνάδελφοι, Έλληνες και ξένοι, οι οποίοι απέστειλαν απαντητικό email, με το οποίο ενημέρωναν για τη συμμετοχή τους στη συγκεκριμένη έρευνα. Επίσης, ήταν αρκετοί οι συμμετέχοντες κυρίως από το εξωτερικό, οι οποίοι επικοινωνήσαν μαζί μου και πρότειναν κάποια μικρή προσθήκη (επιπλέον επιλογή)

ως προτεινόμενη απάντηση, που δεν είχε συμπεριληφθεί από την αρχή στο ερωτηματολόγιο και που η προσθήκη της θεωρήθηκε σημαντική. Ο συνολικός αριθμός απαντημένων ερωτηματολογίων είναι ενενήντα οκτώ (98), με τα είκοσι εννέα (29) να αντιστοιχούν σε ελληνικά και κυπριακά ΨΑ και τα εξήντα εννέα (69) να αντιστοιχούν σε ΨΑ του εξωτερικού. Τέλος, είναι γνωστό ότι τα δημογραφικά στοιχεία σε μία έρευνα μας επιτρέπουν να αποκτήσουμε μία εικόνα για τα άτομα που καλούνται να συμμετέχουν σε αυτήν. Στην παρούσα έρευνα τα δημογραφικά στοιχεία μας βοήθησαν να αντιληφθούμε ορισμένα στοιχεία της προσωπικότητας των ερωτώμενων που σχετίζονται με την επαγγελματική τους εμπειρία και θέση, την εξοικείωσή τους με τις νέες τεχνολογίες και πιθανώς τη διάθεσή τους για εξέλιξη και αναβάθμιση των παρεχόμενων από αυτούς υπηρεσιών. Ωστόσο, οι ερωτήσεις αυτές αποτελούν ένα μικρό μόνο κομμάτι της έρευνας και διατυπώθηκαν με έναν αρκετά προσεκτικό τρόπο ώστε να εξασφαλιστεί η διακριτικότητα. Οι ερωτήσεις αυτές αφορούν το φύλο, την ηλικιακή ομάδα και τη θέση εργασίας στον εκάστοτε φορέα.

3.3 [Μέτρηση Παραγόντων &] Κατασκευή Ερευνητικού

Εργαλείου

Στη μελέτη αυτή έγινε συλλογή πρωτογενών δεδομένων αναφορικά με τα πρότυπα μεταδεδομένων που χρησιμοποιούνται από τα ψηφιακά αποθετήρια, καθώς και πολλές άλλες μεταβλητές οι οποίες θεωρήθηκαν ότι μπορεί να σχετίζονται με τον βαθμό ευρετηρίασης των πόρων των ΨΑ από τις μηχανές ερευνητικής αναζήτησης όπως το GS. Για τον λόγο αυτό δημιουργήθηκαν δύο ερωτηματολόγια σε ελληνική και αγγλική έκδοση αντίστοιχα, χρησιμοποιώντας την πλατφόρμα Google Forms, αποτελούμενα από 23 ερωτήσεις/στοιχεία. Επίσης κρίθηκε απαραίτητη η σύνταξη συνοδευτικής επιστολής (σε ελληνικά και αγγλικά), με την οποία παρείχαμε πληροφορίες σχετικές με την ταυτοποίηση του ερευνητή, την ταυτοποίηση του Εκπαιδευτικού Ιδρύματος και του ΠΜΣ για το οποίο πραγματοποιείται η έρευνα, όπως επίσης πληροφορίες για τον επιβλέποντα καθηγητή και βέβαια τις πληροφορίες που σχετίζονται με τον σκοπό της έρευνας και την αξία της συμμετοχής των παραληπτών. Ο σύνδεσμος κάθε ερωτηματολογίου ενσωματώθηκε στην αντίστοιχη συνοδευτική επιστολή και διανεμήθηκε ηλεκτρονικά με e-mail σε Ελλάδα και εξωτερικό, αφού πρώτα διαβεβαιώσαμε τους παραλήπτες των επιστολών ότι τα δεδομένα που θα συγκεντρωθούν είναι εμπιστευτικά και τα αποτελέσματα θα χρησιμοποιηθούν αυστηρά στα πλαίσια της στατιστικής ανάλυσης της έρευνας. Οι ερωτήσεις του

ερωτηματολογίου προήλθαν έπειτα από τη μελέτη της βιβλιογραφικής επισκόπησης και παρουσιάζονται στα παραρτήματα Α' και Β' της παρούσας μεταπτυχιακής εργασίας.

3.4 Μέθοδοι Στατιστικής Ανάλυσης

Στο αρχικό στάδιο της ανάλυσης δεδομένων πραγματοποιούνται περιγραφικές στατιστικές για διερευνητικούς σκοπούς. Οι περιγραφικές στατιστικές περιλαμβάνουν τα μέτρα θέσης, όπως η μέση τιμή, η συχνότητα, η διάμεσος και η επικρατούσα τιμή και μέτρα διασποράς όπως η διακύμανση, το εύρος και η τυπική απόκλιση ανάλογα με το είδος των μεταβλητών. Για τη διεξαγωγή της παρούσας έρευνας συγκροτήθηκαν δύο ερωτηματολόγια (σε ελληνικά και αγγλικά), τα οποία διανεμήθηκαν εντός των καλοκαιρινών μηνών σε ένα συγκεκριμένο δείγμα ΨΑ. Η επεξεργασία των απαντήσεων που παραλήφθηκαν πραγματοποιήθηκε κάνοντας χρήση του λογισμικού στατιστικής ανάλυσης “IBM SPSS STATISTICS” version 20. Προτιμήθηκε η δημιουργία ενός ενιαίου φύλλου εργασίας στο εν λόγω λογισμικό (Εικόνα 1) στο οποίο καταχωρήθηκαν οι απαντήσεις και από τα δύο ερωτηματολόγια (δεδομένα). Κάθε στήλη αντιστοιχεί σε μία μεταβλητή/χαρακτηριστικό που εξετάζεται και κάθε σειρά αντιστοιχεί σε ένα απαντημένο ερωτηματολόγιο.

Εικόνα 1. Φύλλο εργασίας στο περιβάλλον SPSS

The screenshot shows the IBM SPSS Statistics software interface. The menu bar includes File, Edit, View, Data, Transform, Analyze, Direct Marketing, Graphs, Utilities, Add-ons, Window, and Help. The toolbar contains various icons for file operations and data manipulation. The main window displays a data view table with 26 rows and 20 columns. The columns are labeled with variable names in Greek, such as Gender, Age, ThesiStoFore, IdromatikaAp, ThematikaAp, EirenitikoOrg, Allo, Vnifia, KafalaiaVnifw, Periodika, AthraPeriod, ProEkturpse, ErgasiesSun, GkizelVniogi, Optikoakoust, Architektonik, Fotografies, and Arxia. The data in the table consists of categorical responses like 'Διαγορισ...', 'Απαντήθηκε', and 'Δεν Απαν...'. The status bar at the bottom indicates 'IBM SPSS Statistics Processor is ready'.

Οι ερωτήσεις διακρίθηκαν i. σε ερωτήσεις πολλαπλής επιλογής (multiple choice) με αποδεκτή μία μοναδική απάντηση, ii. σε ερωτήσεις πολλαπλής επιλογής (multiple choice) με αποδεκτές μίας ή περισσότερων από τις προτεινόμενες απαντήσεις και iii. σε ερωτήσεις που αντιστοιχούν σε τετράβαθμη κλίμακα Likert. Προτιμήθηκε η τετράβαθμη κλίμακα Likert, για να υπάρχει μία πιο ξεκάθαρη προσωπική τοποθέτηση (θετική ή αρνητική) των ερωτώμενων. Για την καταχώρηση των μεταβλητών που αφορούν τις ερωτήσεις στις οποίες γίνονται αποδεκτές περισσότερες της μίας από τις προτεινόμενες απαντήσεις, δημιουργήθηκαν μεταβλητές όσες και οι ερωτήσεις αυτές, όπου καθορίστηκαν οι τιμές 0 και 1 αντίστοιχα (συγκεκριμένα όπου 0= δεν απαντήθηκε και όπου 1= απαντήθηκε). Οι ερωτήσεις αυτές έπειτα αναλύθηκαν χρησιμοποιώντας την εντολή Ανάλυση πολλαπλών απαντήσεων (Multiple Response). Οι στήλες που αφορούν όλα τα χαρακτηριστικά/μεταβλητές παραμετροποιήθηκαν ανάλογα με το είδος της προτεινόμενης απάντησης σε ονομαστικές (nominal), διατάξιμες (ordinal) και συνεχής (scale) μεταβλητές (Σφάλμα! Το αρχείο προέλευσης της αναφοράς δεν βρέθηκε.).

Εικόνα 2. Παραμετροποίηση μεταβλητών

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	Gender	Numeric	8	2	Φύλο	{1,00, Γυναι...	None	8	Right	Nominal	Input
2	Age	Numeric	8	2	Ηλικία	{1,00, Κάτω...	None	8	Right	Scale	Input
3	ThesiStoFor...	Numeric	8	2	Ποια είναι η θέ...	{1,00, Βιβλι...	None	8	Right	Ordinal	Input
4	IdrumatikoA...	Numeric	8	2	Ιδρυματικό Απο...	{,00, Δεν Α...	None	8	Right	Nominal	Input
5	ThematikoA...	Numeric	8	2	Θεματικό Αποθ...	{,00, Δεν Α...	None	8	Right	Nominal	Input
6	EreunitikoO...	Numeric	8	2	Ερευνητικό/Όργ...	{,00, Δεν Α...	None	8	Right	Nominal	Input
7	Allo	Numeric	8	2	Άλλο Αποθετήριο	{,00, Δεν Α...	None	8	Right	Nominal	Input
8	Vivlia	Numeric	8	2	Βιβλία	{,00, Δεν Α...	None	8	Right	Nominal	Input
9	KefalaiaVivli...	Numeric	8	2	Κεφάλαια Βιβλί...	{,00, Δεν Α...	None	8	Right	Nominal	Input
10	Periodika	Numeric	8	2	Περιοδικά	{,00, Δεν Α...	None	8	Right	Nominal	Input
11	ArthraPerio...	Numeric	8	2	Άρθρα Περιοδικ...	{,00, Δεν Α...	None	8	Right	Nominal	Input
12	ProEktupws...	Numeric	8	2	Προεκτυπώσεις	{,00, Δεν Α...	None	8	Right	Nominal	Input
13	ErgasiesSu...	Numeric	8	2	Εργασίες Συνεδ...	{,00, Δεν Α...	None	8	Right	Nominal	Input
14	GkrizaVivlio...	Numeric	8	2	Γκριζα Βιβλιογ...	{,00, Δεν Α...	None	8	Right	Nominal	Input
15	Optikoakou...	Numeric	8	2	Οπτικοακουστι...	{,00, Δεν Α...	None	8	Right	Nominal	Input
16	Architektoni...	Numeric	8	2	Αρχιτεκτονικά Σ...	{,00, Δεν Α...	None	8	Right	Nominal	Input
17	Fwtografies	Numeric	8	2	Φωτογραφίες	{,00, Δεν απ...	None	8	Right	Nominal	Input
18	Arxeiakoyliko	Numeric	8	2	Αρχαιακό Υλικό	{,00, Δεν Α...	None	8	Right	Nominal	Input
19	PwsGinetai...	Numeric	8	2	Πως γίνεται η α...	{1,00, Με α...	None	8	Right	Ordinal	Input
20	Poiologismi...	Numeric	8	2	Ποιο λογισμικό ...	{1,00, Dspa...	None	8	Right	Ordinal	Input
21	Xrismopoiei...	Numeric	8	2	Χρησιμοποιείτε ...	{1,00, Ναι τ...	None	8	Right	Nominal	Input
22	SumfwnaMe...	Numeric	8	2	Σύμφωνα με πο...	{1,00, Dubli...	None	8	Right	Ordinal	Input
23	Titlos	Numeric	8	2	Τίτλος	{,00, Δεν Α...	None	8	Right	Nominal	Input
24	Suggrafeas	Numeric	8	2	Συγγραφέας	{,00, Δεν Α...	None	8	Right	Nominal	Input
25	Ekdotis	Numeric	8	2	Εκδότης	{,00, Δεν Α...	None	8	Right	Nominal	Input
26	XronologiaE...	Numeric	8	2	Χρονολογία Έκ...	{,00, Δεν Α...	None	8	Right	Nominal	Input
27	Thema	Numeric	8	2	Θέμα	{,00, Δεν Α...	None	8	Right	Nominal	Input
28	LexeisKleidia	Numeric	8	2	Λέξεις-Κλειδιά	{,00, Δεν Α...	None	8	Right	Nominal	Input
29	Glwssa	Numeric	8	2	Γλώσσα	{,00, Δεν Α...	None	8	Right	Nominal	Input

Για την εκτίμηση του ποσοστού απάντησης κάθε μεταβλητής εκτελέστηκε η διαδικασία του Πίνακα συχνοτήτων (Frequency table) (

Εικόνα 3).

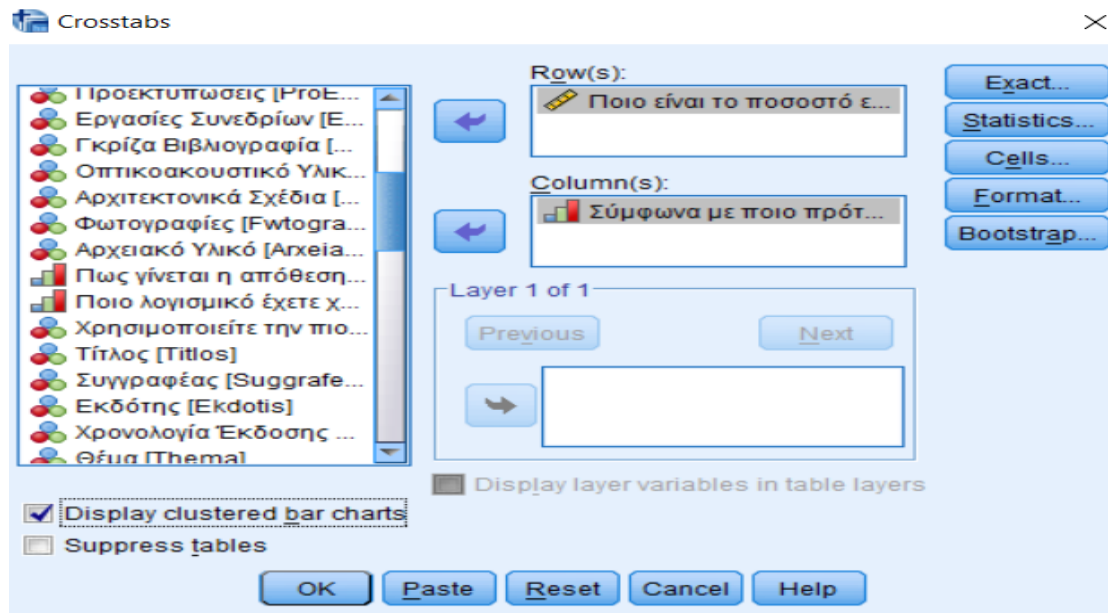
Εικόνα 3. Πίνακας Συχνοτήτων

Statistics											
		Ποια είναι η θέση σας στον φορέα	Πως γίνεται η απόθεση των ψηφιακών αποθετηρίοσας;	Ποιο λογισμικό έχετε χρησιμοποιήσει για την υλοποίηση του ψηφιακού αποθετηρίουσας;	Σύμφωνα με ποιο πρότυπο μεταδομένων περιγράφονται οι ψηφιακοί πόροι στο αποθετηρίοσας;	Θεωρείτε σημαντική την παρακολούθηση και αφομοίωση των εξελίξεων και των ενδημιώσεων (νέες εκδόσεις) του προτύπου μεταδομένων που χρησιμοποιείται από το ψηφιακό αποθετήριο;	Θεωρείτε ότι η τυποποίηση των μεταδομένων είναι σημαντική για την ανάκτηση των ψηφιακών πόρων;	Ποια δήλωση από τις ακόλουθες περιγράφει καλύτερα τον τρόπο της θεματικής ευρετηρίασης του ψηφιακού αποθετηρίουσας;	Θεωρείτε ότι η ανακάλυψη των ψηφιακών πόρων του αποθετηρίουσας εξυπηρετείται με επάρκεια από τις μηχανές αναζήτησης όπως το Google Scholar;	Είστε ικανοποιημένοι από τον βαθμό ευρετηρίασης του περιχομένου του αποθετηρίουσας από το Google Scholar;	Θεωρείτε ότι το πρότυπο μεταδομένων παίζει σημαντικό ρόλο στην ευρετηρίαση των ψηφιακών πόρων του αποθετηρίουσας από τις μηχανές ερευνητικής αναζήτησης όπως το Google Scholar;
N	Valid	87	98	98	98	98	98	98	98	98	95
	Missing	11	0	0	0	0	0	0	0	0	3
	Mean	2,7356	3,9796	4,0510	4,0306	3,5306	3,7347	2,5000	3,1327	3,0102	3,3684
	Median	3,0000	4,0000	1,5000	2,0000	4,0000	4,0000	2,0000	3,0000	3,0000	3,0000
	Mode	3,00	5,00	1,00	2,00	4,00	4,00	2,00	3,00	3,00	4,00
	Std. Deviation	,95795	,98420	4,33726	3,55068	,67662	,60156	1,12378	,69812	,66573	,70034

Frequency Table

Επίσης, για τον εντοπισμό και την ανάδειξη πιθανών σχέσεων μεταξύ των μεταβλητών (στη συγκεκριμένη περίπτωση πρόκειται κυρίως για ποιοτικές μεταβλητές (ονομαστικές και διατάξιμες)) διενεργήθηκε η διαδικασία Crosstabs (Εικόνα 4).

Εικόνα 4. Διαδικασία Crosstabs



Για να πραγματοποιηθεί ο έλεγχος ύπαρξης πιθανής συσχέτισης μεταξύ δύο ή περισσότερων κατηγορικών μεταβλητών (μη διατάξιμων (nominal)), χρησιμοποιήθηκε ο συντελεστής ανεξαρτησίας Pearson Chi-square (χ^2) (Ρούσσος & Ευσταθίου, 2008). Ο συντελεστής Pearson Chi-square (χ^2) είναι ένας έλεγχος υποθέσεων όπου η μηδενική υπόθεση H_0 δηλώνει ότι οι μεταβλητές είναι ανεξάρτητες. Για να έχουμε συσχέτιση μεταξύ των μεταβλητών πρέπει το p-value (Sig.) να είναι μικρότερο από 5% (p-value (Asymptotic significance) <0.05). Όσο πιο μεγάλο είναι το χ^2 , τόσο πιο σίγουρη είναι η εξάρτηση. Για να υπολογιστεί το χ^2 στο SPSS, από την επιλογή Statistics επιλέγουμε το Chi-Square. Το Asymptotic significance δηλώνει ότι ένα αποτέλεσμα είναι στατιστικά σημαντικό και συμβολίζεται με το p-value.

Κεφάλαιο 4. Αποτελέσματα – Ευρήματα /

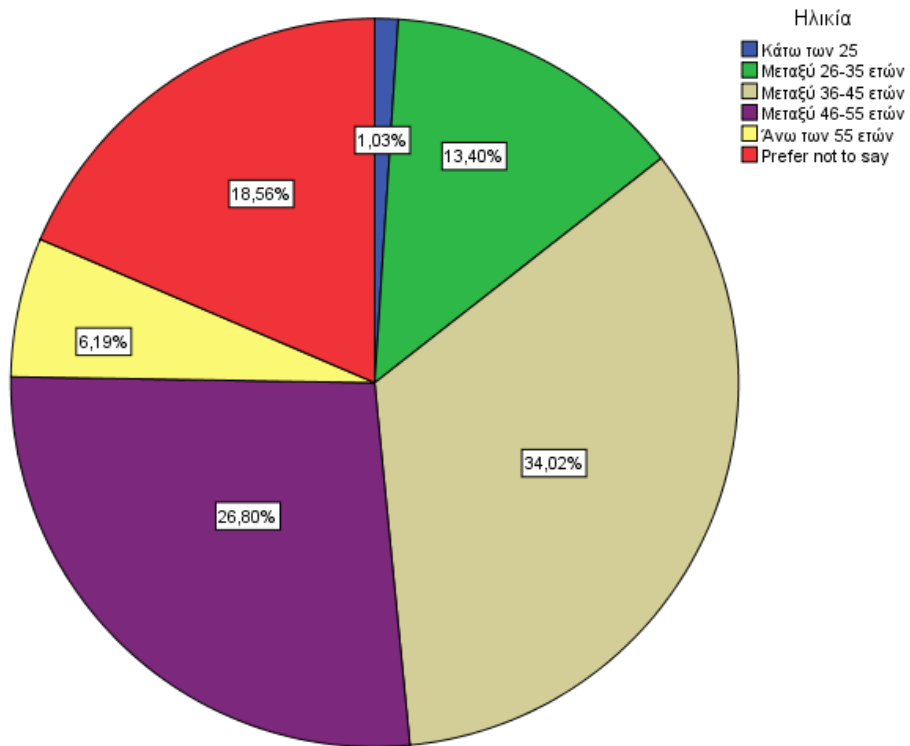
Επιτεύγματα

4.1 Αναλυτική παρουσίαση αποτελεσμάτων

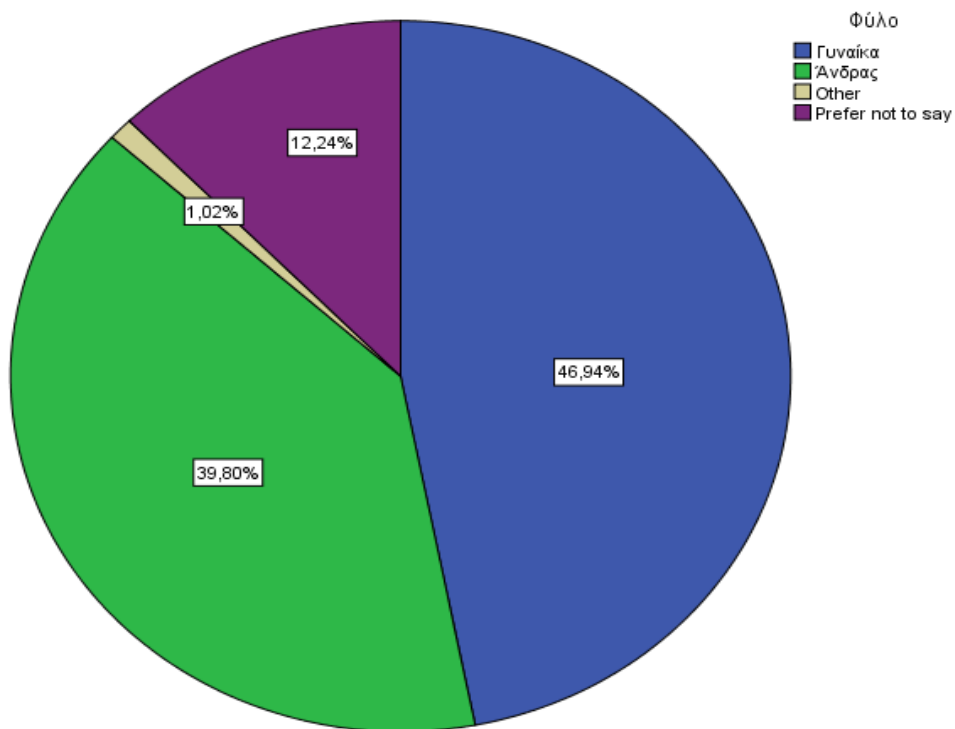
Σε αυτό το κεφάλαιο πραγματοποιείται η αναλυτική παρουσίαση των αποτελεσμάτων της έρευνας. Στη συνέχεια, στο κεφάλαιο 4.2, παρουσιάζονται οι αναλύσεις συσχέτισης με σκοπό να εξεταστούν οι πιθανές σχέσεις μεταξύ των προτύπων μεταδεδομένων, των πεδίων μεταδεδομένων που περιγράφονται και που πιθανώς επηρεάζουν το ποσοστό ευρητηρίασης των τεκμηρίων που περιέχονται στα ΨΑ από τις μηχανές αναζήτησης όπως το GS.

Αναφορικά με τις δημογραφικές ερωτήσεις της έρευνας και από τις απαντήσεις που παραλήφθηκαν, προκύπτει ότι οι περισσότεροι από τους ερωτώμενους ανήκουν στην ηλικιακή ομάδα μεταξύ των 36-45 ετών, ακολουθούν όσοι ανήκουν στην ηλικιακή ομάδα μεταξύ των 46-55 ετών, και τελευταίοι έρχονται όσοι ανήκουν στην ηλικιακή ομάδα των κάτω των 25 ετών (Γράφημα 2) ενώ στην πλειοψηφία τους είναι γυναίκες (Γράφημα 2). Η ηλικιακή κατάσταση του δείγματος μαζί με την εμπειρία και τη διάθεση για εξέλιξη και ανάδειξη των δραστηριοτήτων των φορέων που αντιπροσωπεύει το δείγμα, θεωρείται πως διαδραματίζει ήδη και θα συνεχίσει να διαδραματίζει σημαντικό ρόλο ως προς την εξέλιξη και ανάπτυξη ποιοτικών υπηρεσιών στα ΨΑ, καθώς πρόκειται για νέους σε ηλικία και παραγωγικούς επαγγελματίες και προφανώς με εμπειρία στο χώρο. Επαγγελματίες που ενδεχομένως με την κατάλληλη μέριμνα από τους φορείς που αντιπροσωπεύουν και εξυπηρετούν (όπως π.χ. ενημερωμένο λογισμικό ΨΑ, συμμετοχή σε εκπαιδευτικά σεμινάρια/προγράμματα, εξασφάλιση ικανοποιητικού αριθμού εξειδικευμένου προσωπικού ακόμα κι εξασφάλιση κονδυλίων για περαιτέρω έρευνα & ανάπτυξη) θα μπορέσουν να αποδώσουν καλύτερα προσφέροντας ακόμα πιο ποιοτικές υπηρεσίες.

Γράφημα 1. Ηλικιακή ομάδα

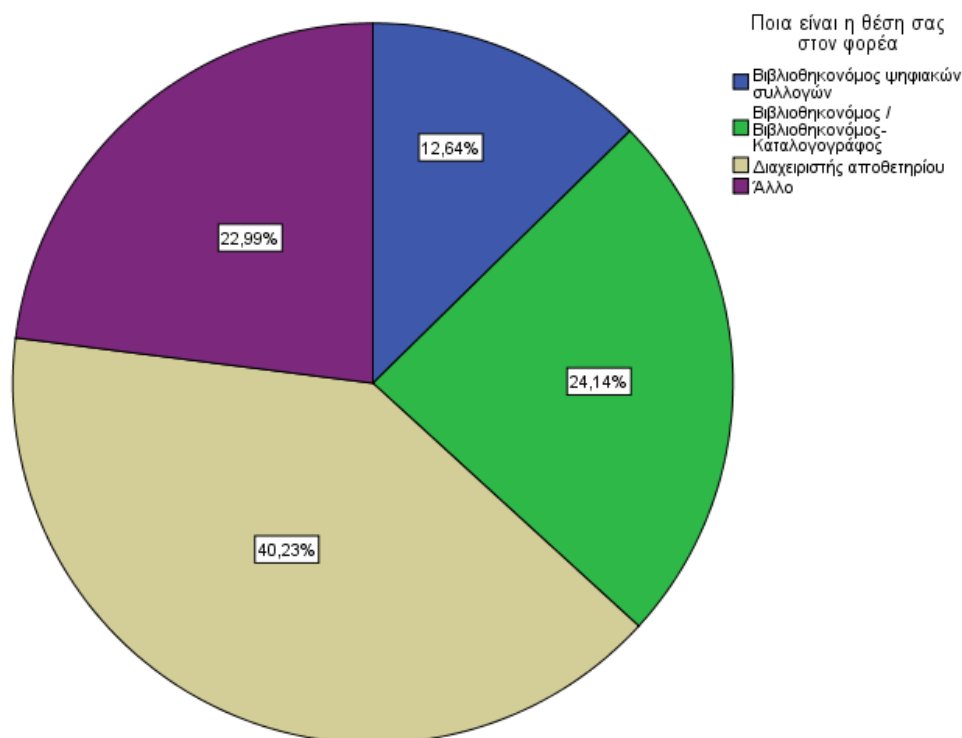


Γράφημα 2. Φύλο



Η ιδιότητα/θέση των συμμετεχόντων στον φορέα εργασίας τους παρουσιάζει επίσης ενδιαφέρον. Όπως φαίνεται και στο γράφημα που ακολουθεί (Γράφημα 3) το μεγαλύτερο ποσοστό (40,23%) αντιστοιχεί σε διαχειριστές ΨΑ, ένα αρκετά μεγάλο ποσοστό (24,14%) είναι βιβλιοθηκονόμοι (συμπεριλαμβανομένων και των καταλογογράφων), όπως επίσης και βιβλιοθηκονόμοι που ειδικεύονται στη διαχείριση των ψηφιακών συλλογών (12,64%). Τέλος, ένα σημαντικό ποσοστό (22,99%) δηλώνει μία διαφορετική από τις προτεινόμενες ιδιότητα (επιλογή άλλο), μεταξύ των οποίων συμπεριλαμβάνονται οι εξής: διευθυντές Βιβλιοθηκών/Ερευνητικών Φορέων, υπεύθυνοι επιστημονικής επικοινωνίας, ερευνητές, μηχανικοί πληροφοριακών συστημάτων, πληροφορικοί, υπεύθυνοι εκδοτικών τμημάτων, κ.λπ. Γίνεται λοιπόν αντιληπτό ότι υπάρχει ένα μεγάλο εύρος επαγγελματιών διαφορετικών ειδικοτήτων και με διαφορετικό υπόβαθρο, οι οποίοι ασχολούνται κατά βάση με τη δόμηση και τη λειτουργία των ΨΑ ανά τον κόσμο. Η διάκριση αυτή θα μπορούσε λοιπόν να δηλώνει την λήψη διαφορετικών αποφάσεων σχετικά με τις πολιτικές δόμησης των ΨΑ, την επιλογή προτύπου μεταδεδομένων, τον τρόπο περιγραφής των πόρων, την επιλογή των σημείων πρόσβασης στον ψηφιακό πόρο (στοιχείων μεταδεδομένων που συμπληρώνονται), ακόμα και με τις προτεραιότητες που τίθενται ως προς τους τρόπους επικοινωνίας και ανοίγματος/προβολής της επιστημονικής και εκδοτικής δραστηριότητας του φορέα που αντιπροσωπεύουν στην επιστημονική/ακαδημαϊκή κοινότητα.

Γράφημα 3. Ιδιότητα/θέση στο Φορέα



Η πλειοψηφία των συμμετεχόντων της παρούσας έρευνας εργάζεται, όπως μπορούμε να δούμε και στον πίνακα που ακολουθεί, σε Ιδρυματικά ΨΑ με ποσοστό 80,6%, ενώ τα Θεματικά και τα Ερευνητικά Αποθετήρια έρχονται στην δεύτερη θέση αμφότερα, διατηρώντας το ίδιο ποσοστό της τάξης του 8,3%. Η ερώτηση βέβαια έδινε τη δυνατότητα στους συμμετέχοντες να κάνουν αποδεκτή περισσότερες της μίας απαντήσεις, καθώς δεν ήταν λίγοι οι φορείς που διέθεταν περισσότερα του ενός ΨΑ (εξού και ο μεγαλύτερος αριθμός των 98 αναμενόμενων απαντήσεων).

Πίνακας 2. Τι είδους ψηφιακό αποθετήριο διαθέτει ο φορέας σας;

	Απαντήσεις		Ποσοστό επί των συμμετεχόντων
	N	Ποσοστό	
Ιδρυματικό Αποθετήριο	87	80,6%	88,8%
Είδος ΨΑ του Φορέα Θεματικό Αποθετήριο	9	8,3%	9,2%
απασχόλησης Ερευνητικό/Οργανισμού Αποθετήριο	9	8,3%	9,2%
Άλλο Αποθετήριο	3	2,8%	3,1%
Σύνολο	108	100,0%	110,2%

Η ερώτηση που αφορά στα είδη υλικού που περιέχονται στις συλλογές των ΨΑ έρχεται να αποδείξει το προφανές, ότι δηλαδή το περιεχόμενο των ΨΑ είναι πολύ πλούσιο (Πίνακας 3), συμπεριλαμβάνοντας υλικό (μη εμπορικό) που δύσκολα εντοπίζεται οπουδήποτε αλλού. Η ευρετηρίαση του υλικού αυτού από τις μηχανές ερευνητικής αναζήτησης όπως το GS κρίνεται ιδιαίτερα σημαντική, καθώς σχετίζεται άμεσα με την ανακάλυψη και τη χρήση του.

Πίνακας 3. Είδη υλικού που περιλαμβάνονται στο ψηφιακό αποθετήριο

	Απαντήσεις		Ποσοστό επί των συμμετεχόντων
	N	Ποσοστό	
Βιβλία	78	11,5%	79,6%
Κεφάλαια Βιβλίων	72	10,6%	73,5%
Περιοδικά	43	6,3%	43,9%
Άρθρα Περιοδικών	81	11,9%	82,7%
Είδη υλικού στα ΨΑ Προ-εκδόσεις	57	8,4%	58,2%
Εργασίες Συνεδρίων	83	12,2%	84,7%
Γκρίζα Βιβλιογραφία (Διπλωματικές εργασίες/Διατριβές)	82	12,1%	83,7%

	Οπτικοακουστικό Υλικό	50	7,4%	51,0%
	Αρχιτεκτονικά Σχέδια	12	1,8%	12,2%
	Φωτογραφίες	40	5,9%	40,8%
	Αρχειακό Υλικό	32	4,7%	32,7%
	Άλλο Υλικό	22	3,2%	22,4%
	Πολυμέσα	27	4,0%	27,6%
Σύνολο		679	100,0%	692,9%

Η επιλογή «Άλλο υλικό» συμπληρώθηκε από τους συμμετέχοντες στην έρευνα με τα εξής είδη υλικού: τεχνικές εκθέσεις και αναφορές, αφίσες, άρθρα εφημερίδων, μαθησιακά αντικείμενα, δεδομένα έρευνας (research data), πηγαίο κώδικα (Source Code), διοικητικά έγγραφα, πατέντες, συναυλίες, ηχογραφήσεις, θεατρικές παραστάσεις, μουσικές συνθέσεις, εικαστικές εκθέσεις, έργα εικαστικών τεχνών, κ.ά.

Αναφορικά με τους τρόπους με τους οποίους γίνεται η απόθεση των ψηφιακών πόρων στα ΨΑ η πλειοψηφία των συμμετεχόντων (38,8%) έδωσε την απάντηση «Άλλο» (Πίνακας 4), με δεύτερη στην προτίμησή τους την επιλογή «Απόθεση από τον βιβλιοθηκονόμο ψηφιακών συλλογών». Πέραν λοιπόν των προτεινόμενων απαντήσεων (i. Με αυτοαρχειοθέτηση (8.2%), ii. Με αυτοαρχειοθέτηση με συμβολή και ποιοτικό έλεγχο από τον βιβλιοθηκονόμο ψηφιακών συλλογών (24,5%), iii. Από τον βιβλιοθηκονόμο ψηφιακών συλλογών (28,6%)), οι συμμετέχοντες συμπλήρωσαν στην επιλογή «Άλλο» τη συνδυαστική χρήση πολλών από τους ήδη αναφερόμενους τρόπους απόθεσης των πόρων στα ΨΑ. Η διαχείριση ωστόσο της αρχειοθέτησης τουλάχιστον με συμβολή από εξειδικευμένο επαγγελματία (βιβλιοθηκονόμο) δείχνει την τάση των ΨΑ για έλεγχο των μεταδεδομένων πριν την οριστική υποβολή των πόρων στο ΨΑ. Με μία δεύτερη δηλαδή ανάγνωση θα μπορούσε κανείς να εντοπίσει την βούληση από πλευράς ΨΑ για εξασφάλιση ενδεχομένως καλύτερων, πιο ποιοτικών μεταδεδομένων.

Πίνακας 4. Πως γίνεται η απόθεση των ψηφιακών πόρων στο αποθετήριό σας;

	Συχνότητα	Ποσοστό	Έγκυρο Ποσοστό	Αθροιστικό Ποσοστό
Με αυτοαρχειοθέτηση	8	8,2	8,2	8,2
Με αυτοαρχειοθέτηση με συμβολή και ποιοτικό έλεγχο από τον βιβλιοθηκονόμο ψηφιακών συλλογών	24	24,5	24,5	32,7

Από τον βιβλιοθηκονόμο ψηφιακών συλλογών	28	28,6	28,6	61,2
Άλλο	38	38,8	38,8	100,0
Σύνολο	98	100,0	100,0	

Όσον αφορά στην απάντηση των συμμετεχόντων σχετικά με το λογισμικό που έχει χρησιμοποιηθεί για την υλοποίηση του εκάστοτε ψηφιακού αποθετηρίου (Πίνακας 5), πρώτο σε επιλογή και με διαφορά έρχεται το ανοικτό λογισμικό DSpace με ποσοστό 50%, με το λογισμικό Eprints να ακολουθεί με ποσοστό 18,4%. Επίσης η επιλογή «Άλλο» έχει συμπληρωθεί με τη χρήση των εξής λογισμικών: WEKO, HAL CCSD, Orus, Samvera/Hyrax, συνδυασμό Esploro και DSpace, συνδυασμό OJS και Pressbooks, συνδυασμό Fedora και Nioivity, LibreCat και PURE, Keystone, κάποιο άλλο ιδιοπαγές λογισμικό βασισμένο στο Fedora, Libsolution, Υπηρεσίες όπως το μοντέλο SaaS που αναπτύσσει το EKT, DIAS (Digital Information Archival System), Museum Plus, κ.ά.

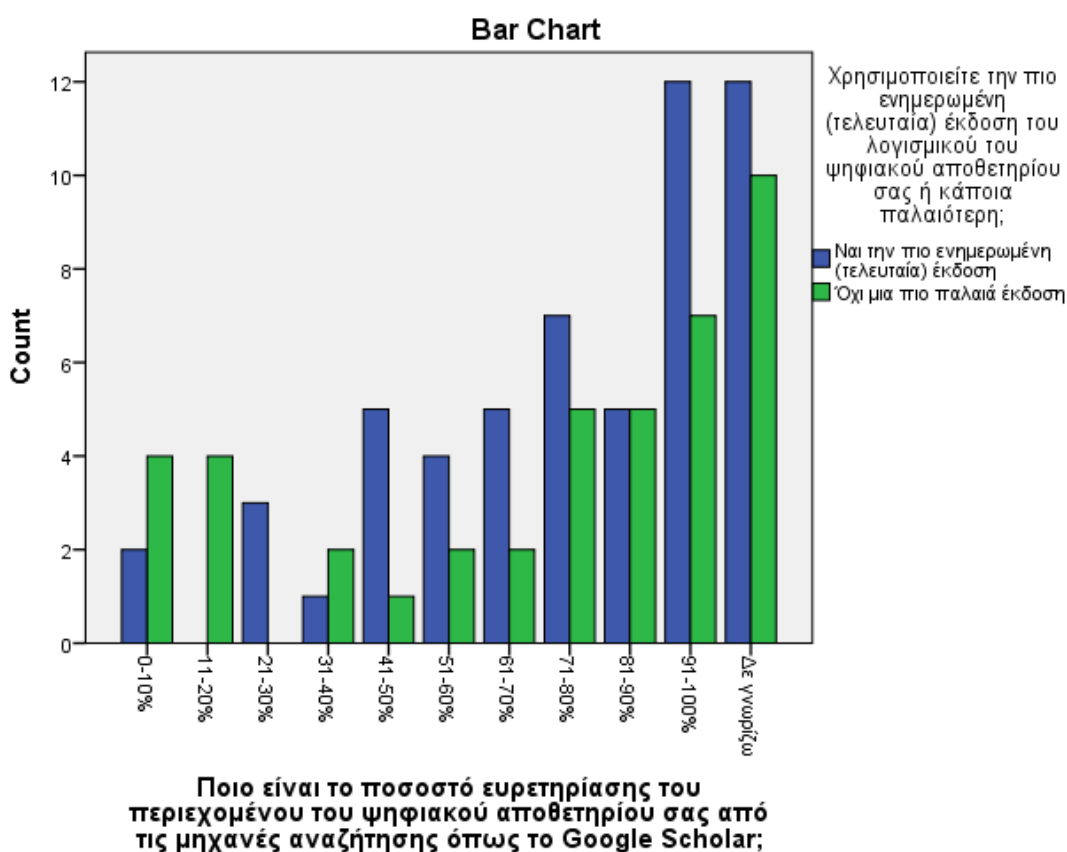
Πίνακας 5. Ποιο λογισμικό έχετε χρησιμοποιήσει για την υλοποίηση του ψηφιακού αποθετηρίου σας;

	Συχνότητα	Ποσοστό	Έγκυρο Ποσοστό	Αθροιστικό Ποσοστό
DSpace	49	50,0	50,0	50,0
Fedora	2	2,0	2,0	52,0
Eprints	18	18,4	18,4	70,4
Digital Commons	6	6,1	6,1	76,5
CDSware	1	1,0	1,0	77,6
OJS	1	1,0	1,0	78,6
Islandora	1	1,0	1,0	79,6
Άλλο	20	20,4	20,4	100,0
Σύνολο	98	100,0	100,0	

Συμπληρωματική επίσης είναι η ερώτηση για τη χρήση ή μη της τελευταίας (πιο νέας) έκδοσης του λογισμικού με το οποίο γίνεται η υλοποίηση του ΨΑ. Μπορούμε λοιπόν να αντιληφθούμε ότι ίσως δεν έχει δοθεί η δέουσα προσοχή ως προς αυτήν τη μεταβλητή από τους υπευθύνους των ΨΑ, καθώς όπως αναφέρεται και στις οδηγίες του Google Scholar “Inclusion Guidelines for Webmasters” (Google Scholar, 2010), γίνεται εμφανής η σύσταση χρήσης της τελευταίας (πιο ενημερωμένης) έκδοσης συγκεκριμένα του λογισμικού DSpace. Στο γράφημα που ακολουθεί (Γράφημα 4) αποτυπώνεται η χρήση της τελευταίας ή μη

έκδοσης λογισμικού ΨΑ ως προς το ποσοστό ευρετηρίασής του από το GS. Όπως μπορούμε να διακρίνουμε στις περιπτώσεις στις οποίες χρησιμοποιείται η τελευταία έκδοση του λογισμικού του ΨΑ, τυχαίνει, τις περισσότερες φορές, να αποδίδονται και τα υψηλότερα ποσοστά ευρετηρίασης.

Γράφημα 4. Χρησιμοποιείτε την πιο ενημερωμένη (τελευταία) έκδοση του λογισμικού του ΨΑ σας ή κάποια παλαιότερη; * Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ΨΑ σας από τις μηχανές αναζήτησης όπως το GS;



Σχετικά με το πρότυπο μεταδεδομένων σύμφωνα με το οποίο περιγράφονται οι ψηφιακοί πόροι στα ΨΑ που συμμετείχαν στην έρευνα, πρώτο σε επιλογή έρχεται το εξειδικευμένο (Qualified) Dublin Core σε ποσοστό 54,1% και αμέσως μετά ακολουθεί η επιλογή «Άλλο» σε ποσοστό 17,3%, ενώ το μη εξειδικευμένο Dublin Core (Unqualified) έρχεται στην τρίτη θέση με ποσοστό 10,2%. Τα πρότυπα μεταδεδομένων δε, που προτείνονται από τις μηχανές αναζήτησης όπως το GS, αναφέρονται σε πολύ χαμηλά ποσοστά στην παρούσα μελέτη (Highwire Press tags (1.0%), Eprints tags (7.1%), BE Press tags (2.0%)) (Πίνακας 6). Η διαπίστωση αυτή μας κάνει να επιδιώκουμε ακόμη περισσότερο την εύρεση λύσης ως προς τη βελτίωση της ευρετηρίασης των ΨΑ που χρησιμοποιούν το πρότυπο Dublin Core για την περιγραφή των πόρων τους από τις μηχανές αναζήτησης όπως το GS.

Πίνακας 6. Σύμφωνα με ποιο πρότυπο μεταδεδομένων περιγράφονται οι ψηφιακοί πόροι του αποθετηρίου σας;

	Συχνότητα	Ποσοστό	Έγκυρο Ποσοστό	Αθροιστικό Ποσοστό
Dublin Core Unqualified	10	10,2	10,2	10,2
Dublin Core Qualified	53	54,1	54,1	64,3
Marc21	5	5,1	5,1	69,4
Highwire Press tags	1	1,0	1,0	70,4
Eprints tags	7	7,1	7,1	77,6
BE Press tags	2	2,0	2,0	79,6
MODS	2	2,0	2,0	81,6
TEI	1	1,0	1,0	82,7
Άλλο	17	17,3	17,3	100,0
Σύνολο	98	100,0	100,0	

Τα στοιχεία μεταδεδομένων τα οποία συμπληρώνονται και θεωρούνται απαραίτητα για την περιγραφή των πόρων των ΨΑ παρουσιάζονται στον πίνακα που ακολουθεί (Πίνακας 7), στον οποίο γίνεται εμφανής η συχνότητα χρήσης και η βαρύτητα περιγραφής που δίνεται σε κάθε στοιχείο. Πρώτο σε συμπλήρωση έρχεται το στοιχείο Τίτλος με ποσοστό 13,3%, έπεται το στοιχείο Συγγραφέας με ποσοστό 12,5% και ακολουθούν με μικρότερα ποσοστά συμπλήρωσης τα στοιχεία Χρονολογία έκδοσης/δημοσίευσης με ποσοστό 11,7%, Εκδότης με ποσοστό 8,0%, Γλώσσα με ποσοστό 7,9%, Θέμα με ποσοστό 7,5%, Λέξη-κλειδί με ποσοστό 6,9%, Αριθμός τόμου με ποσοστό 5,8%, Παραπομπή, Αριθμός τεύχους και Αριθμός σελίδων με ίδιο ποσοστό 5,7%, Αναγνωριστικό με ποσοστό 5,5% και τέλος το «Άλλο στοιχείο» με ποσοστό 3,7%. Σχετικά με τη συμπλήρωση του «Άλλου» στοιχείου έχουν αναφερθεί τα εξής: μοναδικός αριθμός κλειδί, περίληψη, μέλη επιτροπών, χωρική κάλυψη, βιβλιογραφική αναφορά στο υποβληθέν τεκμήριο (π.χ. IEEE style), είδος τεκμηρίου, άδειες χρήσης, κατάσταση δημοσίευσης, κ.λπ.

Πίνακας 7. Απαιτούμενα στοιχεία μεταδεδομένων που καταχωρίζονται στο ΨΑ

		Απαντήσεις		Αθροιστικό Ποσοστό
		N	Ποσοστό	
Απαιτούμενα στοιχεία μεταδεδομένων τα οποία καταχωρίζονται στο ΨΑ σας	Τίτλος	96	13,3%	98,0%
	Συγγραφέας	90	12,5%	91,8%
	Εκδότης	58	8,0%	59,2%
	Χρονολογία Έκδοσης/ Δημοσίευσης	84	11,7%	85,7%

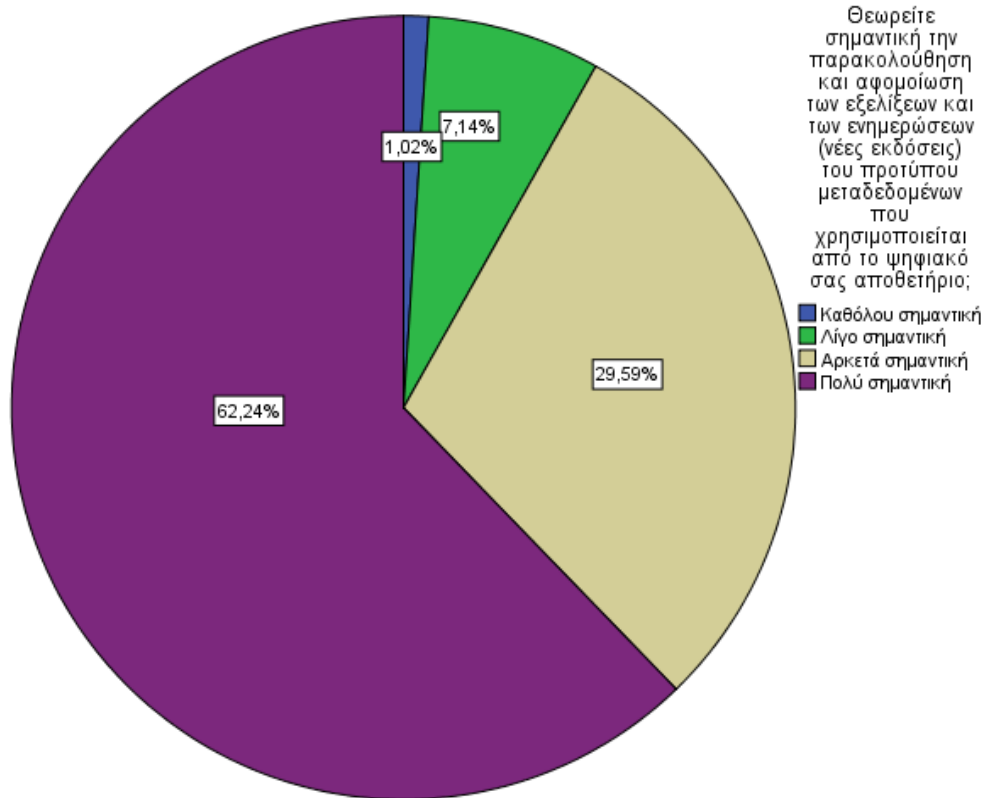
Θέμα	54	7,5%	55,1%
Λέξεις-Κλειδιά	50	6,9%	51,0%
Γλώσσα	57	7,9%	58,2%
Παραπομπή	41	5,7%	41,8%
Αριθμός Τόμου	42	5,8%	42,9%
Αριθμός Τεύχους	41	5,7%	41,8%
Αριθμός Σελίδων	41	5,7%	41,8%
Αναγνωριστικό ISBN_ISSN_DOI	40	5,5%	40,8%
Άλλο Στοιχείο	27	3,7%	27,6%
Σύνολο	721	100,0%	735,7%

Η άποψη ότι η παρακολούθηση και αφομοίωση των εξελίξεων και των ενημερώσεων (νέες εκδόσεις) του προτύπου μεταδεδομένων που χρησιμοποιείται από κάθε ΨΑ θεωρείται πολύ σημαντική διακρίνεται εύκολα, καθώς έχει απαντηθεί θετικά ως «Πολύ σημαντική» και «Αρκετά σημαντική» σε ποσοστό 62,2% και 29,6% αντίστοιχα (Πίνακας 8 και Γράφημα 5).

Πίνακας 8. Θεωρείτε σημαντική την παρακολούθηση και αφομοίωση των εξελίξεων και των ενημερώσεων (νέες εκδόσεις) του προτύπου μεταδεδομένων που χρησιμοποιείται από το ΨΑ σας;

	Συχνότητα	Ποσοστό %	Έγκυρο Ποσοστό	Αθροιστικό Ποσοστό
Καθόλου σημαντική	1	1,0	1,0	1,0
Λίγο σημαντική	7	7,1	7,1	8,2
Αρκετά σημαντική	29	29,6	29,6	37,8
Πολύ σημαντική	61	62,2	62,2	100,0
Σύνολο	98	100,0	100,0	

Γράφημα 5. Θεωρείτε σημαντική την παρακολούθηση και αφομοίωση των εξελίξεων και των ενημερώσεων του προτύπου μεταδεδομένων που χρησιμοποιείται από το ΨΑ σας;



Αντίστοιχα, καθίσταται σαφές ότι η τυποποίηση των μεταδεδομένων θεωρείται πολύ σημαντική για την ανάκτηση των ψηφιακών πόρων των ΨΑ (Πίνακας 9 και Γράφημα 6). Η άποψη αυτή ενισχύεται παράλληλα από το γεγονός ότι είναι αρκετά αυξημένο το ποσοστό του τρόπου περιγραφής της θεματικής ευρετηρίασης του περιεχομένου των ΨΑ με χρήση λέξεων-κλειδιών ή ταξινομήσεων σύμφωνα με τυποποιημένα συστήματα στη γλώσσα της χώρας έκδοσης και/ή στα αγγλικά (Στον πίνακα που ακολουθεί μπορούμε να αντλήσουμε πληροφορία για τον τρόπο με τον οποίο γίνεται η θεματική ευρετηρίαση των τεκμηρίων στα ΨΑ και όπως βλέπουμε η επικρατέστερη δήλωση είναι η θεματική ευρετηρίαση με «Μη ελεγχόμενες λέξεις-κλειδιά [στη γλώσσα της χώρας έκδοσης και/ή στα αγγλικά]» με ποσοστό 38,8% και με την δήλωση ευρετηρίαση με «Λέξεις-κλειδιά ή ταξινομήσεις σύμφωνα με τυποποιημένα συστήματα [στη γλώσσα της χώρας έκδοσης και/ή στα αγγλικά]» να ακολουθεί με ποσοστό 29,6% (Πίνακας 10) και τις υπόλοιπες επιλογές να έπονται.

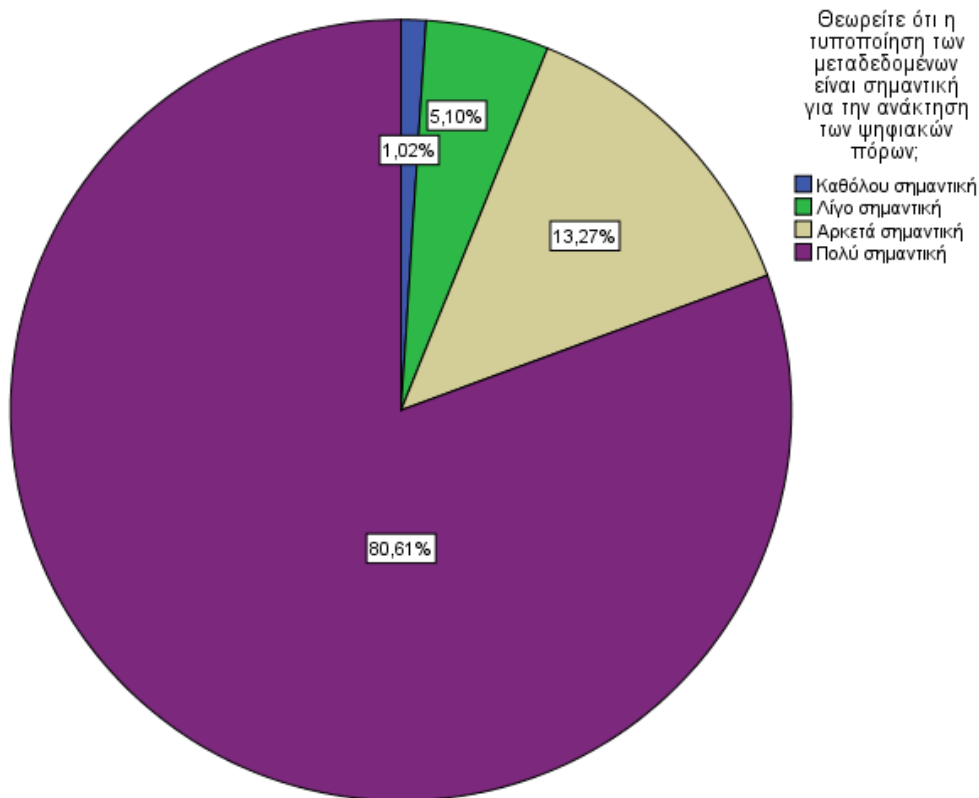
Πίνακας 10).

Πίνακας 9. Θεωρείτε ότι η τυποποίηση των μεταδεδομένων είναι σημαντική για την ανάκτηση των ψηφιακών πόρων;

	Συχνότητα	Ποσοστό	Έγκυρο Ποσοστό	Αθροιστικό Ποσοστό

Καθόλου σημαντική	1	1,0	1,0	1,0
Λίγο σημαντική	5	5,1	5,1	6,1
Αρκετά σημαντική	13	13,3	13,3	19,4
Πολύ σημαντική	79	80,6	80,6	100,0
Σύνολο	98	100,0	100,0	

Γράφημα 6. Θεωρείτε ότι η τυποποίηση των μεταδεδομένων είναι σημαντική για την ανάκτηση των ψηφιακών πόρων;



Στον πίνακα που ακολουθεί μπορούμε να αντλήσουμε πληροφορία για τον τρόπο με τον οποίο γίνεται η θεματική ευρετηρίαση των τεκμηρίων στα ΨΑ και όπως βλέπουμε η επικρατέστερη δήλωση είναι η θεματική ευρετηρίαση με «Μη ελεγχόμενες λέξεις-κλειδιά [στη γλώσσα της χώρας έκδοσης και/ή στα αγγλικά]» με ποσοστό 38,8% και με την δήλωση ευρετηρίαση με «Λέξεις-κλειδιά ή ταξινομήσεις σύμφωνα με τυποποιημένα συστήματα [στη γλώσσα της χώρας έκδοσης και/ή στα αγγλικά]» να ακολουθεί με ποσοστό 29,6% (Πίνακας 10) και τις υπόλοιπες επιλογές να έπονται.

Πίνακας 10. Ποια δήλωση από τις ακόλουθες περιγράφει καλύτερα τον τρόπο της θεματικής ευρετηρίασης του ψηφιακού αποθετηρίου σας;

	Συχνότητα	Ποσοστό	Έγκυρο Ποσοστό	Αθροιστικό Ποσοστό

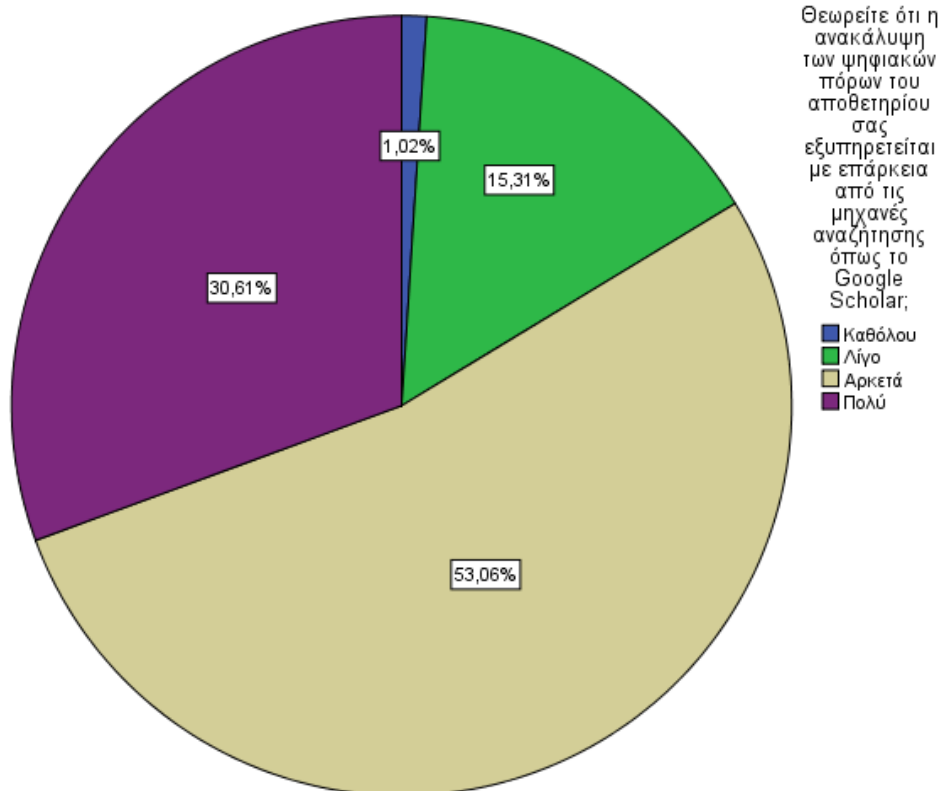
Μη ελεγχόμενες λέξεις-κλειδιά [στη γλώσσα της χώρας έκδοσης]	20	20,4	20,4	20,4
Μη ελεγχόμενες λέξεις-κλειδιά [στη γλώσσα της χώρας έκδοσης και/ή στα αγγλικά]	38	38,8	38,8	59,2
Λέξεις-κλειδιά ή ταξινομήσεις σύμφωνα με τυποποιημένα συστήματα [ελεγχόμενα λεξιλόγια στη γλώσσα της χώρας έκδοσης]	11	11,2	11,2	70,4
Λέξεις-κλειδιά ή ταξινομήσεις σύμφωνα με τυποποιημένα συστήματα [στη γλώσσα της χώρας έκδοσης και/ή στα αγγλικά]	29	29,6	29,6	100,0
Σύνολο	98	100,0	100,0	

Στην ερώτηση σχετικά με την άποψη των συμμετεχόντων για το αν η ανακάλυψη των ψηφιακών πόρων που περιέχονται στο αποθετήριο τους εξυπηρετείται με επάρκεια από τις μηχανές αναζήτησης όπως το GS, η πλειοψηφία απάντησε «Αρκετά» σε ποσοστό 53,1% και «Πολύ» σε ποσοστό 30,6%, δείχνοντας με αυτόν τον τρόπο την ικανοποίησή τους ως προς αυτήν την παράμετρο (Πίνακας 11 και Γράφημα 7).

Πίνακας 11. Θεωρείτε ότι η ανακάλυψη των ψηφιακών πόρων του αποθετηρίου σας εξυπηρετείται με επάρκεια από τις μηχανές αναζήτησης όπως το Google Scholar;

	Συχνότητα	Ποσοστό	Έγκυρο Ποσοστό	Αθροιστικό Ποσοστό
Καθόλου	1	1,0	1,0	1,0
Λίγο	15	15,3	15,3	16,3
Αρκετά	52	53,1	53,1	69,4
Πολύ	30	30,6	30,6	100,0
Σύνολο	98	100,0	100,0	

Γράφημα 7. Θεωρείτε ότι η ανακάλυψη των ψηφιακών πόρων του αποθετηρίου σας εξυπηρετείται με επάρκεια από τις μηχανές αναζήτησης όπως το Google Scholar;



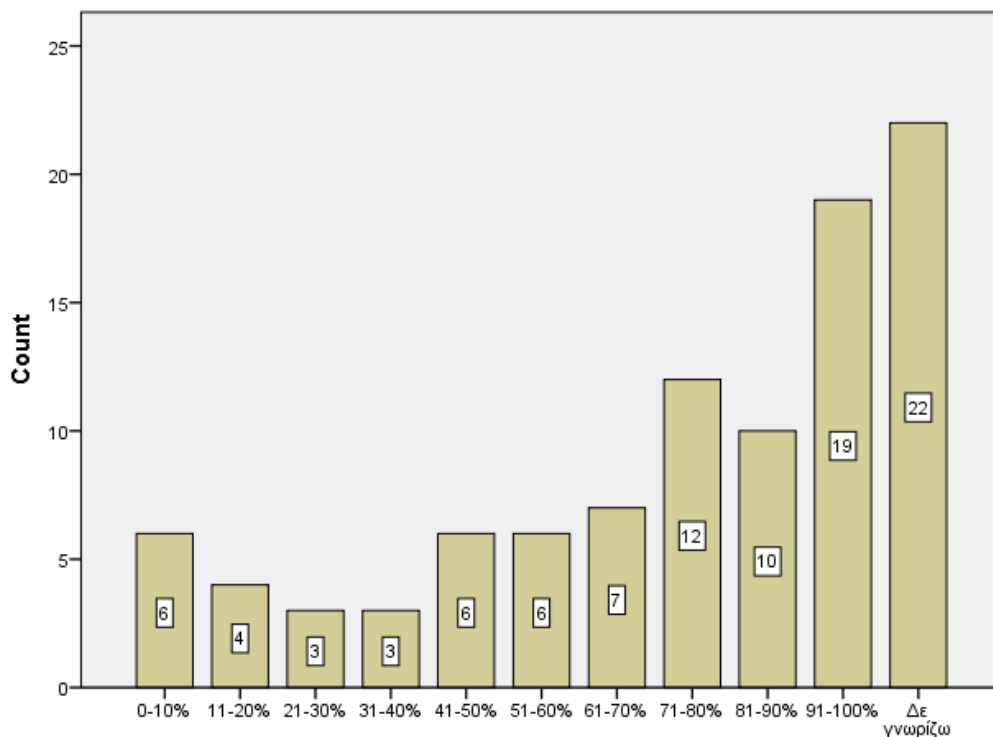
Ωστόσο, παρατηρώντας το γράφημα που απεικονίζει το ποσοστό ευρετηρίασης των ΨΑ του εκάστοτε φορέα από τις μηχανές ερευνητικής αναζήτησης όπως το GS (Γράφημα 8), γίνεται αντιληπτό ότι η πλειοψηφία των συμμετεχόντων στην έρευνα αδυνατεί να απαντήσει στην ερώτηση δηλώνοντας άγνοια. Αυτή η παράμετρος δημιουργεί καλώς ή κακώς σύγχυση σχετικά με το πόσο σίγουροι είναι οι συμμετέχοντες στις απαντήσεις που δίνουν, στο κατά πόσο έχουν ελέγξει το ποσοστό ευρετηρίασης του περιεχομένου του ΨΑ τους από το GS, ή ακόμα στο αν γίνονται πλήρως κατανοητές οι ερωτήσεις. Παρ' όλα αυτά, κατά δήλωση των συμμετεχόντων στην συγκεκριμένη έρευνα, τα ποσοστά ευρετηρίασης του περιεχομένου των ΨΑ από τις μηχανές αναζήτησης είναι ικανοποιητικά (από 80% και πάνω) για το 30% περίπου των συμμετεχόντων (Πίνακας 12 και Γράφημα 8).

Πίνακας 12. είναι το ποσοστό ευρετηρίασης του περιεχομένου του ΨΑ σας από τις μηχανές αναζήτησης όπως το GS;

	Συχνότητα	Ποσοστό	Έγκυρο Ποσοστό	Αθροιστικό Ποσοστό
0-10%	6	6,1	6,1	6,1
11-20%	4	4,1	4,1	10,2
21-30%	3	3,1	3,1	13,3
31-40%	3	3,1	3,1	16,3

41-50%	6	6,1	6,1	22,4
51-60%	6	6,1	6,1	28,6
61-70%	7	7,1	7,1	35,7
71-80%	12	12,2	12,2	48,0
81-90%	10	10,2	10,2	58,2
91-100%	19	19,4	19,4	77,6
Δε γνωρίζω	22	22,4	22,4	100,0
Σύνολο	98	100,0	100,0	

Γράφημα 8. Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ΨΑ σας από τις μηχανές αναζήτησης όπως το GS;



Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ψηφιακού αποθετηρίου σας από τις μηχανές αναζήτησης όπως το Google Scholar;

Στην ερώτηση σχετικά με την ικανοποίηση από τον βαθμό ευρετηρίασης του περιεχομένου των ΨΑ από τη μηχανή αναζήτησης Google Scholar η πλειοψηφία των ερωτώμενων δηλώνει ικανοποιημένη. Παραμένει όμως δυσαρεστημένο ένα σημαντικό ποσοστό (περίπου το 15% του συνόλου των συμμετεχόντων) (Πίνακας 13) .

Πίνακας 13. Είστε ικανοποιημένοι από τον βαθμό ευρετηρίασης του περιεχομένου του αποθετηρίου σας από το Google Scholar;

	Συχνότητα	Ποσοστό	Έγκυρο Ποσοστό	Αθροιστικό Ποσοστό

Πολύ δυσαρεστημένοι	3	3,1	3,1	3,1
Δυσάρεστημένοι	12	12,2	12,2	15,3
Ικανοποιημένοι	64	65,3	65,3	80,6
Πολύ ικανοποιημένοι	19	19,4	19,4	100,0
Σύνολο	98	100,0	100,0	

Μία ακόμη παράμετρος που ερευνήθηκε είναι αν έχει ελεγχθεί ο βαθμός ικανοποίησης των ερωτώμενων σε σχέση με την ορατότητα, τις προβολές και τις λήψεις του περιεχομένου του αποθετηρίου τους από μηχανές αναζήτησης όπως το Google Scholar. Η πλειοψηφία απάντησε καταφατικά με ποσοστό 59,2% (Πίνακας 14).

Πίνακας 14. Έχετε ελέγξει τον βαθμό ικανοποίησης σε σχέση με την ορατότητα, τις προβολές και τις λήψεις του περιεχομένου του αποθετηρίου σας από μηχανές αναζήτησης όπως το Google Scholar;

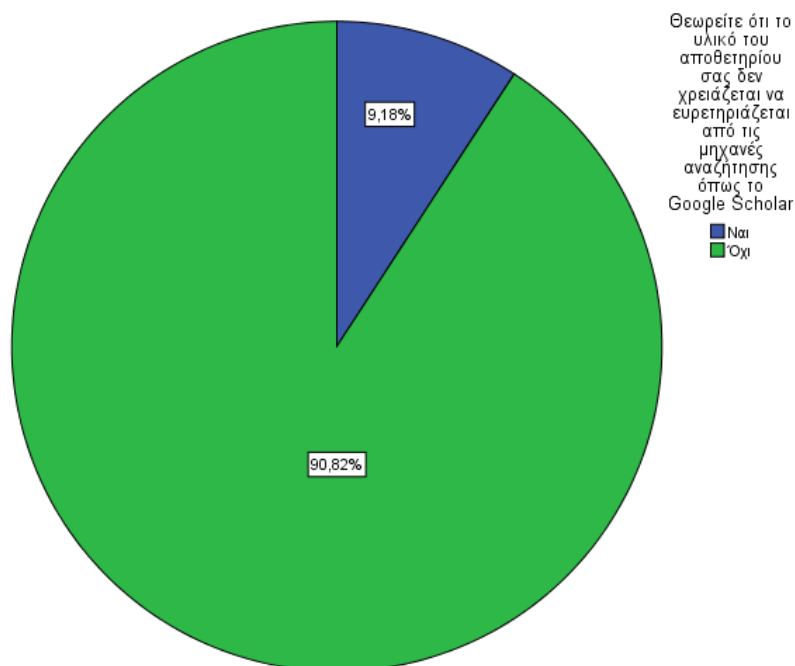
	Συχνότητα	Ποσοστό	Έγκυρο Ποσοστό	Αθροιστικό Ποσοστό
Ναι	58	59,2	59,2	59,2
Όχι	40	40,8	40,8	100,0
Σύνολο	98	100,0	100,0	

Είναι πολύ σημαντικό επίσης το γεγονός ότι οι περισσότεροι συμμετέχοντες αναγνωρίζουν την αναγκαιότητα ευρετηρίασης του περιεχομένου του ΨΑ τους από τις μηχανές αναζήτησης όπως το GS σε ποσοστό 90,82% (Πίνακας 15 και Γράφημα 9).

Πίνακας 15. Θεωρείτε ότι το υλικό του αποθετηρίου σας δεν χρειάζεται να ευρετηριάζεται από τις μηχανές αναζήτησης όπως το Google Scholar

	Συχνότητα	Ποσοστό	Έγκυρο Ποσοστό	Αθροιστικό Ποσοστό
Ναι	9	9,2	9,2	9,2
Όχι	89	90,8	90,8	100,0
Σύνολο	98	100,0	100,0	

Γράφημα 9. Θεωρείτε ότι το υλικό του αποθετηρίου σας δεν χρειάζεται να ευρετηριάζεται από τις μηχανές αναζήτησης όπως το Google Scholar;

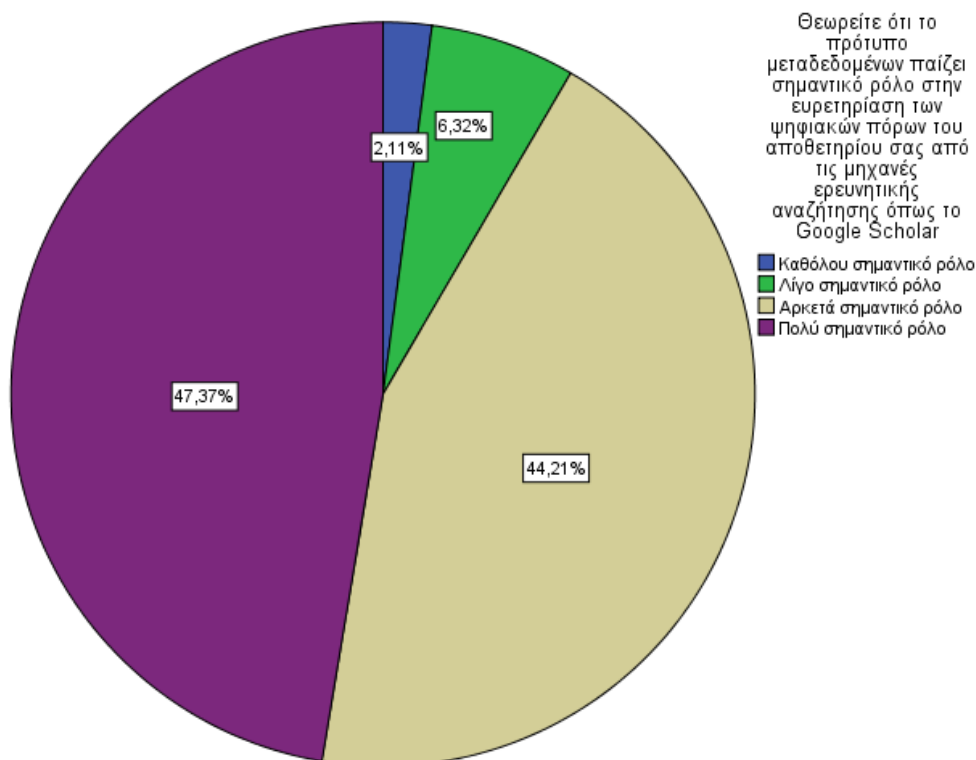


Επίσης, οι συμμετέχοντες θεωρούν ότι το πρότυπο μεταδεδομένων παίζει πολύ σημαντικό ρόλο στην ευρετηρίαση των πόρων των ΨΑ από τις μηχανές αναζήτησης όπως το GS (Πίνακας 16 και Γράφημα 10). Η άποψη αυτή υποστηρίζεται πολύ θερμά, καθώς η πλειονότητα των συμμετεχόντων της έρευνας είναι βιβλιοθηκονόμοι/επαγγελματίες της πληροφόρησης, οι οποίοι στηρίζουν πολύ τα πρότυπα και τα ακολούθησαν πιστά κατά τη δημιουργία των ΨΑ τους, τόσο για τη δημιουργία και διάθεση των μεταδεδομένων, όσο και για τη συλλογή τους (harvesting).

Πίνακας 16. Θεωρείτε ότι το πρότυπο μεταδεδομένων παίζει σημαντικό ρόλο στην ευρετηρίαση των ψηφιακών πόρων του αποθετηρίου σας από τις μηχανές ερευνητικής αναζήτησης όπως το Google Scholar

	Συχνότητα	Ποσοστό	Έγκυρο Ποσοστό	Αθροιστικό Ποσοστό
Καθόλου σημαντικό ρόλο	2	2,0	2,1	2,1
Λίγο σημαντικό ρόλο	6	6,1	6,3	8,4
Αρκετά σημαντικό ρόλο	42	42,9	44,2	52,6
Πολύ σημαντικό ρόλο	45	45,9	47,4	100,0
Σύνολο	95	96,9	100,0	
Ελλείποντα	3	3,1		
Σύνολο	98	100,0		

Γράφημα 10. Θεωρείτε ότι το πρότυπο μεταδεδομένων παίζει σημαντικό ρόλο στην ευρετηρίαση των ψηφιακών πόρων του αποθετηρίου σας από τις μηχανές ερευνητικής αναζήτησης όπως το Google Scholar



Στο ερώτημα αν έχουν προχωρήσει σε κατάλληλες ενέργειες για την επίτευξη αύξησης της ευρετηρίασης του περιεχομένου του ΨΑ κάθε φορά από τις μηχανές αναζήτησης όπως το GS, οι ερωτώμενοι απάντησαν με ποσοστό 62,2% ότι έχουν προχωρήσει σε κατάλληλες ενέργειες για την αύξηση της ευρετηρίασης του αποθετηρίου τους (Πίνακας 17).

Πίνακας 17. Έχετε προχωρήσει σε κατάλληλες ενέργειες για την επίτευξη αύξησης της ευρετηρίασης του περιεχομένου του ψηφιακού αποθετηρίου σας από μηχανές αναζήτησης όπως το Google Scholar;

	Συχνότητα	Ποσοστό	Έγκυρο Ποσοστό	Αθροιστικό Ποσοστό
Ναι	61	62,2	62,2	62,2
Όχι	37	37,8	37,8	100,0
Σύνολο	98	100,0	100,0	

Αναφορικά με τις ενέργειες στις οποίες έχει προχωρήσει κάθε φορέας για την αύξηση της ευρετηρίασης του ΨΑ του μπορούμε να παρατηρήσουμε το εξής (Πίνακας 18): το 15,2% των ερωτώμενων έχουν προχωρήσει σε μετατροπή του προτύπου μεταδεδομένων σε ένα

περισσότερο συμβατό πρότυπο με τις μηχανές αναζήτησης (πχ. Highwire Press tags, Eprints tags, BE Press tags, PRISM tags), το 13,5% των ερωτώμενων έχει προχωρήσει σε ενσωμάτωση ενός μικρού block κώδικα του προτύπου μεταδεδομένων του ΨΑ σας στο HTML της ιστοσελίδας κάθε ψηφιακού πόρου σε μορφή μετα-ετικετών (meta tags), το 20,2% των ερωτώμενων έχει προχωρήσει σε συμμόρφωση με τις οδηγίες συμπερίληψης των μηχανών αναζήτησης (π.χ. Inclusion Guidelines for Webmasters), το 10,7% των ερωτώμενων έχει προχωρήσει σε παρακολούθηση των ενημερώσεων του προτύπου μεταδεδομένων του ΨΑ τους και αφομοίωση των προτεινόμενων αλλαγών, το 16,3% έχει προχωρήσει σε εγκατάσταση της τελευταίας, πιο ενημερωμένης έκδοσης του λογισμικού του ΨΑ τους κι ένα πολύ μικρό ποσοστό της τάξης του 5,6% δηλώνει ότι έχει προβεί σε κάποια άλλη ενέργεια πλην των προτεινόμενων, όπου κυρίως εδώ γίνεται αναφορά στην αυξημένη χρήση λέξεων-κλειδιών και στην βελτιστοποίηση του χάρτη του ιστοτόπου του ΨΑ στο Google. Τέλος, ένα σημαντικό ποσοστό της τάξης των 18,5% δηλώνει πως δεν έχει προχωρήσει σε καμία ενέργεια για την αύξηση της ορατότητας του ΨΑ τους. Βάσει λοιπόν των απαντήσεων που παραλήφθηκαν, γίνεται σαφές ότι η ευρετηρίαση και η ορατότητα του περιεχομένου των ΨΑ από τις μηχανές αναζήτησης είναι ένα ζήτημα που απασχολεί αρκετά σε γενικές γραμμές.

Πίνακας 18. Παρακαλώ κατονομάστε τις ενέργειες στις οποίες έχετε προχωρήσει για την αύξηση της ευρετηρίασης του περιεχομένου του ψηφιακού σας αποθετηρίου από τις μηχανές αναζήτησης όπως το Google Scholar

	Απαντήσεις		Ποσοστό επί των συμμετεχόντων
	N	Ποσοστό	
Ενέργειες για αύξηση της ευρετηρίασης των ΨΑ από τις μηχανές αναζήτησης όπως το GS	27	15,2%	27,6%
	24	13,5%	24,5%

Σε συμμόρφωση με τις οδηγίες συμπερίληψης των μηχανών αναζήτησης (π.χ. Inclusion Guidelines for Webmasters);	36	20,2%	36,7%
Σε παρακολούθηση των ενημερώσεων του προτύπου μεταδεδομένων του ΨΑ σας και αφομοίωση των προτεινόμενων αλλαγών;	19	10,7%	19,4%
Σε εγκατάσταση της τελευταίας, πιο ενημερωμένης έκδοσης του λογισμικού του ΨΑ του Φορέα σας;	29	16,3%	29,6%
Δεν έχω προχωρήσει σε καμία ενέργεια για την αύξηση της ορατότητας του περιεχομένου του αποθετηρίου μου από τις μηχανές αναζήτησης	33	18,5%	33,7%
Άλλη ενέργεια;	10	5,6%	10,2%
Σύνολο	178	100,0%	181,6%

Το τελευταίο ερώτημα της έρευνας σχετίζεται με την ανάδειξη των προβλημάτων που θεωρούνται ότι υφίστανται και επηρεάζουν την ευρετηρίαση του περιεχομένου του ΨΑ κάθε φορέα από τις μηχανές αναζήτησης όπως το GS (Πίνακας 19). Όπως παρατηρείται από τον πίνακα που ακολουθεί το μεγαλύτερο ποσοστό (30,1%) δηλώνει ότι δεν αντιμετωπίζει κανένα απολύτως πρόβλημα ως προς την ευρετηρίαση του περιεχομένου του ΨΑ του από το GS. Το γεγονός αυτό ενδεχομένως εν μέρει να δικαιολογείται, καθώς πράγματι αρκετοί από τους συμμετέχοντες στην έρευνα ανήκουν στην κατηγορία «υψηλής» ευρετηρίασης από το GS (αυτό μπορεί εύκολα κανείς να το αντιληφθεί και από τον Πίνακας 12 και από τη σειρά κατάταξης των ΨΑ από το Ranking Web of Repositories – Παράρτημα Γ΄ της παρούσας Πίνακας 28). Υπάρχει επίσης η πιθανότητα να έχει γίνει κάποια παρανόηση στην αντίληψη της ερώτησης. Ουσιαστικό πάντως παραμένει το γεγονός ότι καταγράφονται οι δυσκολίες/τα προβλήματα που αντιμετωπίζουν τα ΨΑ σε σχέση με την ευρετηρίασή τους από τις μηχανές αναζήτησης, έτσι ώστε να ληφθούν τα απαραίτητα μέτρα που θα

οδηγήσουν στην εξάλειψή τους. Κατά δήλωση λοιπόν των συμμετεχόντων στην έρευνα τα προβλήματα που εντοπίζονται είναι τα εξής: i. Συμβατότητας του προτύπου μεταδεδομένων του ΨΑ με τις απαιτήσεις των μηχανών αναζήτησης σε ποσοστό 12,8%, ii. Δυσκολία ενσωμάτωσης στο HTML των ιστοσελίδων του ψηφιακού αποθετηρίου των μεταδεδομένων των ψηφιακών πόρων σε μορφή μετα-ετικετών (meta-tags) σε ποσοστό 8,3%, iii. Δυσκολία στην αντιστοίχιση (mapping) του προτύπου μεταδεδομένων που χρησιμοποιείται από το ΨΑ σε κάποιο άλλο πρότυπο περισσότερο συμβατό με τη μηχανή αναζήτησης σε ποσοστό 11,3%, iv. Μη ύπαρξης σαφών οδηγιών καταλογογράφησης του υλικού του ΨΑ σε ποσοστό 6,0%, v. Έλλειψης εξειδικευμένου προσωπικού, τόσο για την επιλογή των θεματικών όρων, όσο και για την επιβεβαίωση της ορθότητας των στοιχείων των μεταδεδομένων πριν την απόθεση των ψηφιακών πόρων στο ΨΑ σε ποσοστό 16,5% και vi. «Άλλο πρόβλημα» σε ποσοστό 15,0%. Στην τελευταία αυτή επιλογή «Άλλο» σημειώνονται κυρίως τα παρακάτω: ο χρόνος καθυστέρησης της ευρετηρίασης νέων πόρων από το GS αποτελεί πρόβλημα, γίνεται αναφορά σε λανθασμένη διαμόρφωση του GS με αποτέλεσμα την ανάγκη των ΨΑ για επικοινωνία με άτομο τεχνικής υποστήριξης στο GS. Επίσης αναφέρθηκαν προβλήματα σχετικά με ανεπαρκείς συνδέσμους προς το περιεχόμενο των ΨΑ, με αποτέλεσμα τη χαμηλή κατάταξη των πόρων των ΨΑ στα αποτελέσματα της αναζήτησης στο GS. Επιπλέον αναφέρθηκε το γεγονός ότι το GS ευρετηριάζει εκ νέου κάθε 6 μήνες, επομένως το περιεχόμενο που αλλάζει στο ΨΑ παραμένει «στάσιμο» στο GS για το διάστημα αυτό. Τέλος υπήρξαν ορισμένοι εκ των συμμετεχόντων οι οποίοι δήλωσαν αδυναμία απάντησης λόγω έλλειψης τεχνικών γνώσεων.

Πίνακας 19. Ποια προβλήματα θεωρείτε ότι υφίστανται και επηρεάζουν την ευρετηρίαση του περιεχομένου του αποθετηρίου σας από τις μηχανές αναζήτησης όπως το Google Scholar;

		Απαντήσεις		Ποσοστό επί των συμμετεχόντων
		N	Ποσοστό	
Προβλήματα που επηρεάζουν την ευρετηρίαση του περιεχομένου του ΨΑ από τις μηχανές	Συμβατότητας του προτύπου μεταδεδομένων του αποθετηρίου σας με τις απαιτήσεις των μηχανών αναζήτησης;	17	12,8%	17,5%

αναζήτησης όπως το GS	Δυσκολία ενσωμάτωσης στο HTML των ιστοσελίδων του ψηφιακού αποθετηρίου σας των μεταδεδομένων των ψηφιακών πόρων σε μορφή μετα-ετικετών (meta-tags);	11	8,3%	11,3%
	Δυσκολία στην αντιστοίχιση (mapping) του προτύπου μεταδεδομένων που χρησιμοποιείται από το αποθετήριό σας σε κάποιο άλλο πρότυπο περισσότερο συμβατό με τη μηχανή αναζήτησης;	15	11,3%	15,5%
	Μη ύπαρξης σαφών οδηγιών καταλογογράφησης του υλικού του ΨΑ σας;	8	6,0%	8,2%
	Έλλειψη εξειδικευμένου προσωπικού, τόσο για την επιλογή των θεματικών όρων, όσο και για την επιβεβαίωση της ορθότητας των στοιχείων των μεταδεδομένων πριν την απόθεση των ψηφιακών πόρων στο αποθετήριο;	22	16,5%	22,7%
	Δεν υφίσταται κανένα πρόβλημα	40	30,1%	41,2%
	Άλλο πρόβλημα	20	15,0%	20,6%
Σύνολο		133	100,0%	137,1%

4.2 Κυριότερα ευρήματα / αποτελέσματα

Στην ενότητα παρουσιάζονται τα σημαντικότερα ευρήματα της έρευνας. Για τον εντοπισμό και την παρουσίασή τους πραγματοποιήθηκε ο έλεγχος ύπαρξης πιθανής συσχέτισης μεταξύ δύο ή περισσότερων μεταβλητών, έτσι ώστε να ελεγχθεί αν επαληθεύονται ή όχι οι

ερευνητικές υποθέσεις που τέθηκαν. Στη συγκεκριμένη έρευνα επειδή οι μεταβλητές είναι κατηγορικές/ονομαστικές (nominal), χρησιμοποιήθηκε ο συντελεστής ανεξαρτησίας Pearson Chi-square (χ^2) (Ρούσσος & Ευσταθίου, 2008). Ο συντελεστής Pearson Chi-square (χ^2) είναι ένας έλεγχος υποθέσεων όπου η μηδενική υπόθεση H_0 δηλώνει ότι οι μεταβλητές είναι ανεξάρτητες. Για να έχουμε συσχέτιση μεταξύ των μεταβλητών πρέπει το p-value (Sig.) να είναι μικρότερο από 5% (p-value (Asymptotic significance)<0.05). Όσο μεγαλύτερο είναι το χ^2 , τόσο πιο σίγουρη είναι η εξάρτηση (Ζαφειρόπουλος, 2015).

Αρχικά λοιπόν έγινε έλεγχος συσχέτισης κάνοντας χρήση της εντολής Crosstabs (Πίνακας 20) μεταξύ του προτύπου μεταδεδομένων σύμφωνα με το οποίο περιγράφονται οι ψηφιακοί πόροι στο αποθετήριο κάθε φορέα σε σχέση με το ποσοστό ευρετηρίασης του περιεχομένου του ΨΑ από τις μηχανές αναζήτησης όπως το GS. Από τον έλεγχο αυτό προέκυψαν τα εξής συμπεράσματα: εφόσον το p-value (sig) είναι μεγαλύτερο από 5% (Sig.=0.548>0,05) (Πίνακας 21) τότε ισχύει η H_0 . Αυτό πρακτικά σημαίνει ότι οι δύο μεταβλητές είναι ανεξάρτητες. Επομένως, με βάση τα αποτελέσματα της διαδικασίας χ^2 δεν φαίνεται να υπάρχει στατιστικά μία συσχέτιση μεταξύ των δύο αυτών μεταβλητών. Παρ' όλα αυτά παρατηρώντας το Γράφημα 11 που ακολουθεί, γίνεται αντιληπτό ότι ΨΑ που έχουν υιοθετήσει το πρότυπο Dublin Core για την περιγραφή των πόρων τους, έχουν παραδόξως δηλώσει υψηλότερα ποσοστά ευρετηρίασης του περιεχομένου τους από το GS, σε αντίθεση με ΨΑ που δηλώνουν ότι κάνουν χρήση ενός περισσότερου συμβατού/προτεινόμενου προτύπου από το GS. Η διαπίστωση αυτή είναι πολύ ευχάριστη και ανέλπιστα παράλληλα, και θεωρώ ότι χρήζει περαιτέρω έρευνας, δεδομένου ότι στις οδηγίες του Google Scholar "Inclusion Guidelines for Webmasters" (Google Scholar, 2010), γίνεται εμφανής η σύσταση χρήσης μετα-ετικετών (meta-tags) περιγραφής DC ως την έσχατη επιλογή.

Πίνακας 20. Σύμφωνα με ποιο πρότυπο μεταδεδομένων περιγράφονται οι ψηφιακοί πόροι στο αποθετήριό σας * Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ψηφιακού αποθετηρίου σας από τις μηχανές αναζήτησης όπως το Google Scholar; Crosstabulation

Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ψηφιακού αποθετηρίου σας από τις μηχανές αναζήτησης όπως το Google Scholar;											Σύνολο ο
0-10%	11-20%	21-30%	31-40%	41-50%	51-60%	61-70%	71-80%	81-90%	91-100%	Δε γνωρίζω	

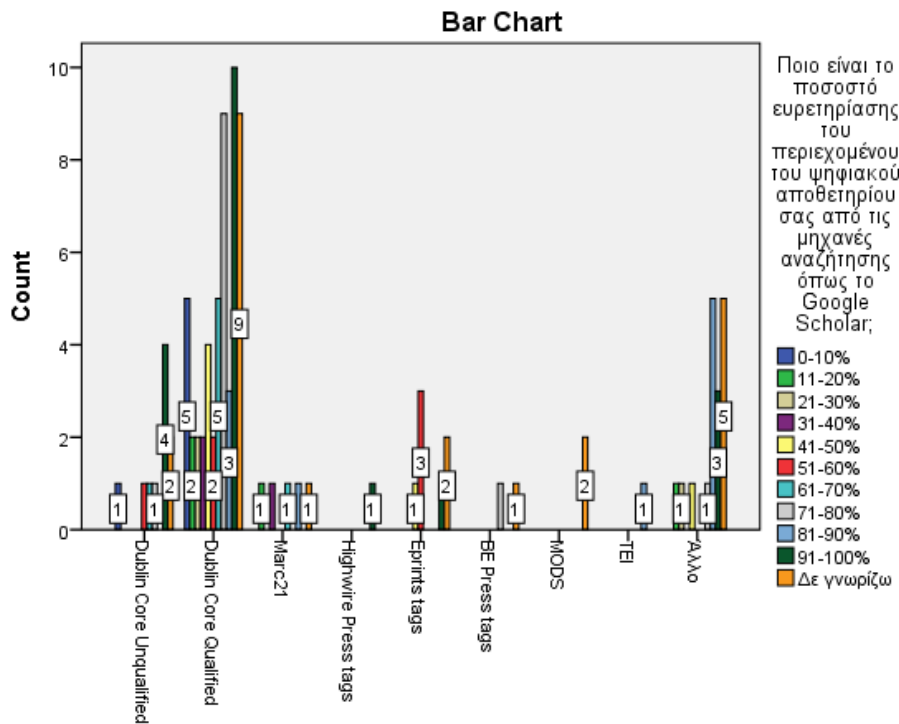
Σύμφωνα με ποιο πρότυπο μεταδεδομένων περιγράφονται οι ψηφιακοί πόροι στο αποθετήριο σας;	Dublin Core Unqualified	1	0	0	0	0	1	1	1	0	4	2	10
	Dublin Core Qualified	5	2	2	2	4	2	5	9	3	10	9	53
	Marc21	0	1	0	1	0	0	1	0	1	0	1	5
	Highwire Press tags	0	0	0	0	0	0	0	0	0	1	0	1
	Eprints tags	0	0	0	0	1	3	0	0	0	1	2	7
	BE Press tags	0	0	0	0	0	0	0	1	0	0	1	2
	MODS	0	0	0	0	0	0	0	0	0	0	2	2
	TEI	0	0	0	0	0	0	0	0	1	0	0	1
	Άλλο	0	1	1	0	1	0	0	1	5	3	5	17
Σύνολο	6	4	3	3	6	6	7	12	10	19	22	98	

Πίνακας 21. Σύμφωνα με ποιο πρότυπο μεταδεδομένων περιγράφονται οι ψηφιακοί πόροι στο αποθετήριό σας * Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ψηφιακού αποθετηρίου σας από τις μηχανές αναζήτησης όπως το Google Scholar; Pearson's Chi-Square Tests

	Value	Df	Asymp. Sig. (2-sided)
Pearson Chi-Square	77,829 ^a	80	,548
Likelihood Ratio	67,829	80	,832
Linear-by-Linear Association	2,975	1	,085
N of Valid Cases	98		

Γράφημα 11. Σύμφωνα με ποιο πρότυπο μεταδεδομένων περιγράφονται οι ψηφιακοί πόροι στο αποθετήριό σας * Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ψηφιακού

αποθετηρίου σας από τις μηχανές αναζήτησης όπως το Google Scholar



Σύμφωνα με ποιο πρότυπο μεταδεδομένων περιγράφονται οι ψηφιακοί πόροι στο αποθετήρι...

Στη συνέχεια διενεργήθηκε ακόμη ένας έλεγχος ανεξαρτησίας Pearson Chi-square (χ^2) και έγινε χρήση ξανά της εντολής Crosstabs (Πίνακας 22 Πίνακας 20) για να ερευνηθεί το ενδεχόμενο ύπαρξης συσχέτισης μεταξύ των ενεργειών, στις οποίες έχουν προχωρήσει οι υπεύθυνοι των ΨΑ για την επίτευξη αύξησης της ευρετηρίασης του περιεχομένου των ΨΑ τους από τις μηχανές αναζήτησης όπως το Google Scholar σε σχέση με το ποσοστό ευρετηρίασης του περιεχομένου του ΨΑ κάθε φορά από τις μηχανές αναζήτησης όπως το GS. Από τον έλεγχο και αυτής της πιθανής συσχέτισης προέκυψε ότι το p-value (sig) είναι μεγαλύτερο από 5% (Sig.=0.212>0,05) (Πίνακας 23). Επομένως, και εδώ ισχύει η H_0 . Άρα οι δύο μεταβλητές είναι ανεξάρτητες και με βάση τα αποτελέσματα της διαδικασίας χ^2 δεν φαίνεται να υπάρχει στατιστικά μία συσχέτιση μεταξύ των δύο αυτών μεταβλητών. Ωστόσο, παρατηρείται ότι τα ΨΑ που έχουν προχωρήσει σε ορισμένες ενέργειες, έχουν πράγματι επιτύχει υψηλότερα ποσοστά ευρετηρίασης από τις μηχανές αναζήτησης όπως το GS (Γράφημα 12).

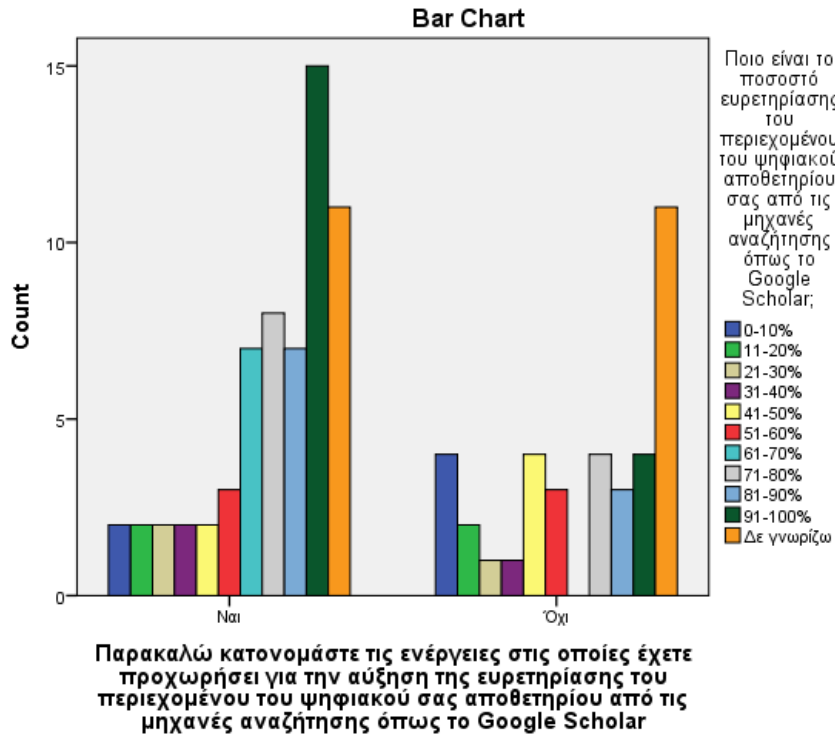
Πίνακας 22. Ενέργειες στις οποίες έχετε προχωρήσει για την αύξηση της ευρετηρίασης του περιεχομένου του ΨΑ σας από τις μηχανές αναζήτησης όπως το GS * Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ΨΑ σας από τις μηχανές αναζήτησης όπως το GS; Crosstabulation

		Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ψηφιακού αποθετηρίου σας από τις μηχανές αναζήτησης όπως το Google Scholar;											Σύνολο
		0-10 %	11-20%	21-30%	31-40%	41-50%	51-60%	61-70%	71-80%	81-90%	91-100%	Δε γνωρίζω	
Παρακαλώ κατονομάστε τις ενέργειες στις οποίες έχετε προχωρήσει για την αύξηση της ευρετηρίασης του περιεχομένου του ψηφιακού αποθετηρίου από τις μηχανές αναζήτησης όπως το Google Scholar	Ναι	2	2	2	2	2	3	7	8	7	15	11	61
	Όχι	4	2	1	1	4	3	0	4	3	4	11	37
Σύνολο		6	4	3	3	6	6	7	12	10	19	22	98

Πίνακας 23. Ενέργειες στις οποίες έχετε προχωρήσει για την αύξηση της ευρετηρίασης του περιεχομένου του ΨΑ σας από τις μηχανές αναζήτησης όπως το GS * Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ΨΑ σας από τις μηχανές αναζήτησης όπως το GS; Pearson's Chi-Square Tests

	Value	Df	Asymp. Sig. (2-sided)
Pearson Chi-Square	13,217 ^a	10	,212
Likelihood Ratio	15,593	10	,112
Linear-by-Linear Association	1,306	1	,253
N of Valid Cases	98		

Γράφημα 12. Ενέργειες στις οποίες έχετε προχωρήσει για την αύξηση της ευρετηρίασης του περιεχομένου του ΨΑ σας από τις μηχανές αναζήτησης όπως το GS * Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ΨΑ σας από τις μηχανές αναζήτησης όπως το GS



Τέλος, διενεργήθηκε ακόμη ένας έλεγχος ανεξαρτησίας με τον ίδιο τρόπο, εστιάζοντας αυτή τη φορά στα στοιχεία/πεδία μεταδεδομένων που συμπληρώνονται κατά την περιγραφή των πόρων των ΨΑ, για να ερευνηθεί το ενδεχόμενο ύπαρξης συσχέτισης μεταξύ αυτών σε σχέση με το ποσοστό ευρετηρίασης του περιεχομένου του ΨΑ κάθε φορά από τις μηχανές αναζήτησης όπως το GS (Πίνακας 24 και Γράφημα 13 κι αντίστοιχα Πίνακας 26 και Γράφημα 14). Ο έλεγχος πραγματοποιήθηκε για κάθε στοιχείο/πεδίο μεταδεδομένων ξεχωριστά (ως ξεχωριστή μεταβλητή) και προέκυψε ότι το p -value (sig) είναι μικρότερο από 5% για την περίπτωση του στοιχείου του Συγγραφέα (Sig.=0,026<0,05) (Πίνακας 25) και την περίπτωση του στοιχείου Λέξη-Κλειδί (Sig.=0,043<0,05) (Πίνακας 27). Επομένως, ισχύει η H1, ότι δηλαδή οι δυο μεταβλητές που εξετάζουμε είναι εξαρτημένες. Επομένως, με βάση τα αποτελέσματα της διαδικασίας χ^2 φαίνεται να υπάρχει στατιστικά μία σημαντική συσχέτιση μεταξύ των δύο αυτών μεταβλητών (Συγγραφέας και Λέξη-Κλειδί) με το ποσοστό ευρετηρίασης των πόρων ενός ΨΑ από το GS. Με την επαλήθευση αυτής της υπόθεσης επιβεβαιώνονται οι έρευνες, από τις οποίες προκύπτει ότι ορισμένα στοιχεία μεταδεδομένων παίζουν πολύ σημαντικό ρόλο στην ανάκτηση των πόρων των ΨΑ από τις μηχανές αναζήτησης. Επομένως, η διάθεση με έγκυρο τρόπο συγκεκριμένων στοιχείων μεταδεδομένων και δει κάνοντας χρήση ενός προτύπου μεταδεδομένων όπως το Dublin Core, θεωρείται σημαντική για την ανακάλυψη των ψηφιακών πόρων από τις δυναμικές μηχανές αναζήτησης όπως το GS.

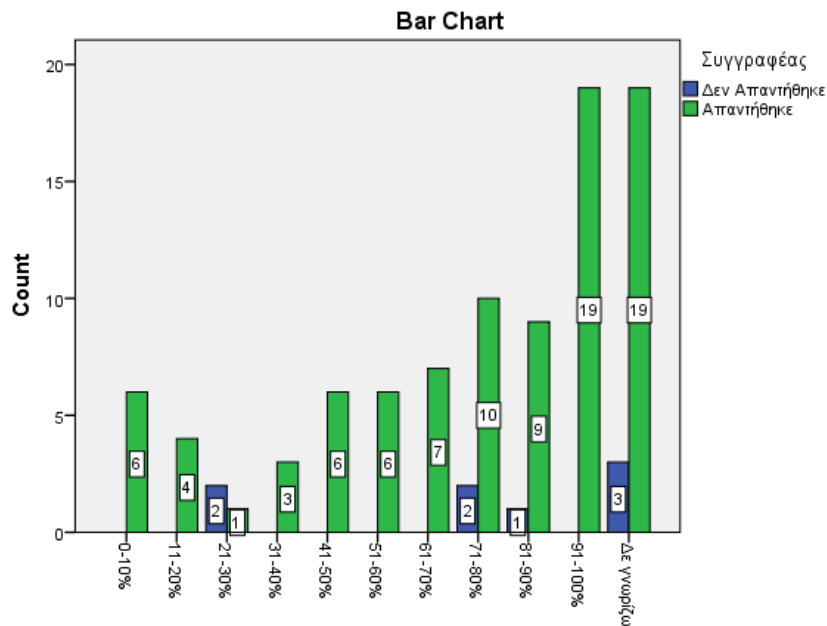
Πίνακας 24. Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ψηφιακού αποθετηρίου σας από τις μηχανές αναζήτησης όπως το Google Scholar * Συγγραφέας; Crosstabulation

	Συγγραφέας		Σύνολο
	Δεν Απαντήθηκε	Απαντήθηκε	
0-10%	0	6	6
11-20%	0	4	4
21-30%	2	1	3
31-40%	0	3	3
41-50%	0	6	6
51-60%	0	6	6
61-70%	0	7	7
71-80%	2	10	12
81-90%	1	9	10
91-100%	0	19	19
Δε γνωρίζω	3	19	22
Σύνολο	8	90	98

Πίνακας 25. Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ψηφιακού αποθετηρίου σας από τις μηχανές αναζήτησης όπως το Google Scholar * Συγγραφέας; Pearson's Chi-Square Tests

	Value	Df	Asymp. Sig. (2-sided)
Pearson Chi-Square	20,311 ^a	10	,026
Likelihood Ratio	16,757	10	,080
Linear-by-Linear Association	,024	1	,876
N of Valid Cases	98		

Γράφημα 13. Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ψηφιακού αποθετηρίου σας από τις μηχανές αναζήτησης όπως το Google Scholar * Συγγραφείς



Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ψηφιακού αποθετηρίου σας από τις μηχανές αναζήτησης όπως το Google Scholar;

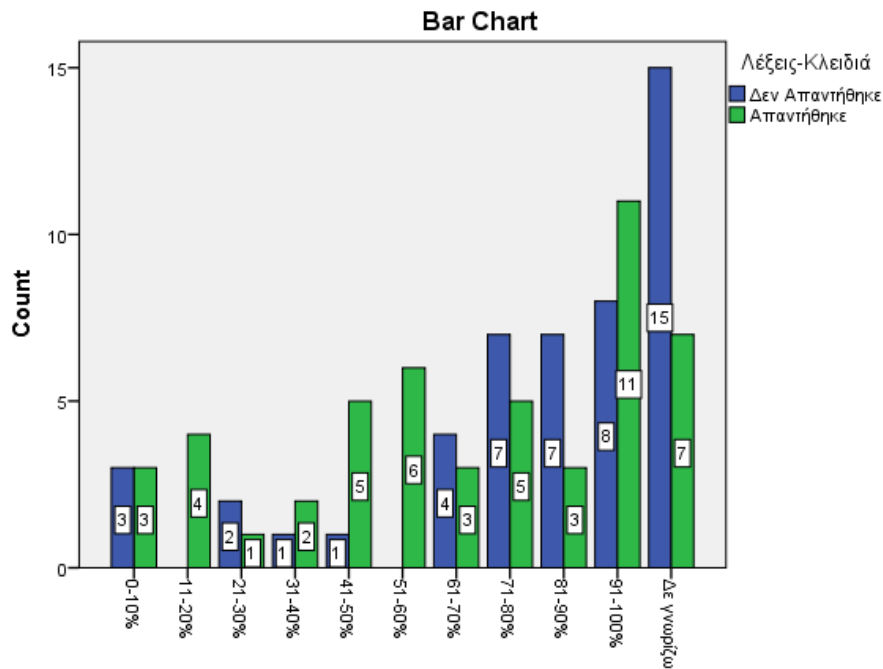
Πίνακας 26. Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ψηφιακού αποθετηρίου σας από τις μηχανές αναζήτησης όπως το Google Scholar * Λέξεις-Κλειδιά; Crosstabulation

	Λέξεις-Κλειδιά		Σύνολο
	Δεν Απαντήθηκε	Απαντήθηκε	
0-10%	3	3	6
11-20%	0	4	4
21-30%	2	1	3
31-40%	1	2	3
41-50%	1	5	6
51-60%	0	6	6
61-70%	4	3	7
71-80%	7	5	12
81-90%	7	3	10
91-100%	8	11	19
Δε γνωρίζω	15	7	22
Σύνολο	48	50	98

Πίνακας 27. Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ψηφιακού αποθετηρίου σας από τις μηχανές αναζήτησης όπως το Google Scholar * Λέξεις-Κλειδιά; Pearson's Chi-Square Tests

	Value	Df	Asymp. Sig. (2-sided)
Pearson Chi-Square	18,759 ^a	10	,043
Likelihood Ratio	22,989	10	,011
Linear-by-Linear Association	4,890	1	,027
N of Valid Cases	98		

Γράφημα 14. Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ψηφιακού αποθετηρίου σας από τις μηχανές αναζήτησης όπως το Google Scholar * Λέξεις-Κλειδιά



Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ψηφιακού αποθετηρίου σας από τις μηχανές αναζήτησης όπως το Google Scholar;

Κεφάλαιο 5. Μελλοντικές επεκτάσεις

Σε αυτό το τελευταίο κεφάλαιο θα γίνει μία συνοπτική ανακεφαλαίωση, συζήτηση επί των σημαντικότερων αποτελεσμάτων της παρούσας διπλωματικής εργασίας.

5.1 Ανακεφαλαίωση

Ανακεφαλαιώνοντας, θα πρέπει αρχικά να θυμηθούμε τον λόγο που μας κέντρισε τόσο πολύ το ενδιαφέρον αυτή η μελέτη. Είναι γεγονός ότι η διενέργεια έρευνας, όπως και η δημοσιοποίηση αυτής και των αποτελεσμάτων της στην επιστημονική κοινότητα οδεύει στην εξ ολοκλήρου πραγματοποίησή της μέσω του διαδικτύου. Σε αυτήν την πολύ απαιτητική και ανταγωνιστική περίοδο οι βιβλιοθήκες και άλλοι ερευνητικοί φορείς αναπτύσσουν υπηρεσίες, όπως τα Ψηφιακά Αποθετήρια, με σκοπό την επίτευξη της ανοικτότητας, της προβολής των ερευνητικών και εκδοτικών δραστηριοτήτων τους, την ανάπτυξη και την προσέλκυση συνεργασιών, την αναγνώριση ακόμα και την εξασφάλιση χρηματοδότησης για την διεύρυνση των ερευνητικών εργασιών/δραστηριοτήτων του εκάστοτε επιστημονικού φορέα. Αντίστοιχα, μηχανές ερευνητικής αναζήτησης όπως το Google Scholar έχουν αναπτυχθεί και λειτουργούν ως συσσωρευτές επιστημονικής γνώσης, προσφέροντας πρόσβαση σε περιεχόμενο ευρέος φάσματος ψηφιακών συλλογών και αποθετηρίων. Ο τρόπος με τον οποίο γίνεται η αναζήτηση και η ανακάλυψη των ψηφιακών πόρων εντός του GS είναι εξαιρετικά εύκολος και βολικός, καθώς λειτουργεί με τον ίδιο τρόπο που λειτουργούν οι βασικές μηχανές αναζήτησης. Προσφέρει λοιπόν ευκολία και ταχύτητα συνάμα στον χρήστη, αφού τον απαλλάσσει από την μεμονωμένη αναζήτηση σε κάθε αποθετήριο ξεχωριστά. Ο ανταγωνισμός όμως μεταξύ των ερευνητών είναι μεγάλος και τα οφέλη της πρόσβασης και ορατότητας για την επιστημονική κοινότητα πολλά, με πρώτο από όλα την απόκτηση περισσότερων αναφορών, τεκμήριο αδιαμφισβήτητο για την απόδειξη της ακαδημαϊκής/επιστημονικής αναγνώρισης και προαγωγής (Shanks & Arlitsch, 2016). Παρ' όλο λοιπόν που το ακαδημαϊκό/ερευνητικό περιβάλλον έχει γίνει πολύ ανταγωνιστικό, το διαδίκτυο και οι μηχανές αναζήτησης κερδίζουν συνεχώς έδαφος ως προς τη διάθεση μεγάλου όγκου επιστημονικής γνώσης και τη συνεχή αύξηση της χρήσης τους. Τα ΨΑ, από την άλλη πλευρά, ενώ διαθέτουν πλούσιο μη εμπορικό υλικό, αρκετά δυσεύρετο και ιδιαίτερο, δεν έχουν καταφέρει να επιτύχουν την ικανοποιητική ευρετηρίασή του από τις μηχανές αναζήτησης. Η αύξηση της ευρετηρίασης του υλικού των

ΨΑ από τις μηχανές αναζήτησης θεωρείται ζωτικής σημασίας, καθώς η χρήση τους στη διεξαγωγή κάθε είδους έρευνας θεωρείται ουσιαστικά μονόδρομος, εφόσον είναι πολλοί οι ερευνητές οι οποίοι ξεκινούν και ολοκληρώνουν την έρευνά τους πραγματοποιώντας τις αναζητήσεις τους στη μηχανή αναζήτησης και μόνο (Arlitsch & O'Brien, 2012; De Rosa et al., 2010). Η ανάγκη λοιπόν για ευρετηρίαση και προβολή των επιστημονικών εργασιών των ερευνητών που περιέχονται σε ΨΑ από τις μηχανές ερευνητικής αναζήτησης είναι μεγάλη. Με πολλούς επιστήμονες να εκτιμούν το πρόβλημα της μη ικανοποιητικής ευρετηρίασης να οφείλεται μεταξύ άλλων στην ποιότητα των μεταδεδομένων, όπως και στο πρότυπο μεταδεδομένων που έχει υιοθετηθεί από το εκάστοτε ψηφιακό αποθετήριο, το οποίο πιθανώς δεν πληροί τις απαιτήσεις σε μεταδεδομένα που έχουν τεθεί από τις μηχανές ερευνητικής αναζήτησης (Arlitsch & O'Brien, 2012). Με γνώμονα λοιπόν αυτές τις υποθέσεις δομήθηκαν και οι ερευνητικές υποθέσεις της παρούσας διπλωματικής εργασίας. Με μεγάλο ενδιαφέρον έγινε η μελέτη της βιβλιογραφίας και η συλλογή χρήσιμων πληροφοριών, ώστε να γίνει κατανοητός ο τρόπος λειτουργίας των ΨΑ αφενός, των μηχανών αναζήτησης και συγκεκριμένα του GS αφετέρου. Έγινε μία προσπάθεια εντοπισμού των πλεονεκτημάτων και των μειονεκτημάτων των προτύπων μεταδεδομένων και πιο συγκεκριμένα του πλέον διαδεδομένου στα ΨΑ προτύπου του DC και ανάδειξης των οφελών του. Αντίστοιχα έγινε εντοπισμός των πλεονεκτημάτων και των μειονεκτημάτων του GS και των προσφερόμενων υπηρεσιών του, σε μια προσπάθεια να αντιληφθούμε που υστερεί και αν μπορεί να λειτουργήσει συμπληρωματικά η πιο «παραδοσιακή» βιβλιοθηκονομία, αυτή που αγαπά και στηρίζεται στα πρότυπα με στόχο να βγούμε όλοι κερδισμένοι στο τέλος. Τέθηκαν λοιπόν τρεις (3) ερευνητικές υποθέσεις γύρω από τις οποίες εστίασε τόσο η βιβλιογραφική ανασκόπηση, όσο και η έρευνα που διενεργήθηκε σε δεύτερη φάση. Οι ερευνητικές υποθέσεις που τέθηκαν είναι οι εξής: i. Μπορεί το πρότυπο μεταδεδομένων που χρησιμοποιεί ένα ΨΑ να επηρεάζει την ανακάλυψη των ψηφιακών πόρων που περιέχονται σε αυτό, από τις σύγχρονες και δυναμικές μηχανές ερευνητικής αναζήτησης; ii. Υπάρχει τρόπος βελτίωσης του προτύπου DC και αύξησης της συμβατότητας των στοιχείων/πεδίων περιγραφής του σχήματος αυτού, τα οποία χρησιμοποιούνται από τα περισσότερα ψηφιακά αποθετήρια, έτσι ώστε να επιτυγχάνεται μεγαλύτερη ορατότητα και ανάκτηση του ψηφιακού περιεχομένου από μηχανές ερευνητικής αναζήτησης όπως το GS; iii. Μήπως η διάθεση/παροχή με έγκυρο τρόπο συγκεκριμένων στοιχείων μεταδεδομένων επηρεάζει την ανακάλυψη και ευρετηρίαση των ψηφιακών πόρων των αποθετηρίων από τις δυναμικές μηχανές αναζήτησης όπως το GS; Έγινε λοιπόν μία ποσοτική έρευνα, με διανομή ερωτηματολογίων και λήψη πρωτογενών δεδομένων/απαντήσεων. Ο στόχος της

έρευνας είναι να παρέχει χρήσιμη πληροφορία στους επαγγελματίες πληροφόρησης και όχι μόνο, που εργάζονται και δομούν τα ΨΑ με πληροφορία, για την επίτευξη καλύτερων αποτελεσμάτων ως προς την προβολή και ανάκτηση του περιεχομένου των ΨΑ, μέσω της χρήσης μηχανών επιστημονικής αναζήτησης όπως το GS. Επιπροσθέτως, ένας ακόμη τελικός στόχος της έρευνας αυτής είναι η βελτίωση των παρεχόμενων υπηρεσιών προς το κοινό των ΨΑ, μέσω του ανοίγματος και της δημοσιοποίησης των προβληματισμών που δημιουργήθηκαν προς την επιστημονική κοινότητα και την αναμονή για πιθανή επίλυσή τους.

5.2 Συζήτηση / Συμπεράσματα

Η ερευνητική αυτή εργασία επικεντρώθηκε στην ανάδειξη του προβλήματος που παρουσιάζεται, αναφορικά δηλαδή με τον χαμηλό βαθμό ευρετηρίασης από τις μηχανές επιστημονικής/ακαδημαϊκής αναζήτησης όπως το GS, μεγάλου μέρους της εκδοτικής δραστηριότητας που περιέχεται σε ΨΑ, προβάλλοντας τα πλεονεκτήματα και τα μειονεκτήματα κάθε υπηρεσίας και την ανάγκη για άμεση λύση του. Επειδή το ζήτημα είναι επιτακτικό και με πολλές προεκτάσεις, χρησιμοποιήθηκε προηγούμενη γνώση ως υπόβαθρο (βιβλιογραφική επισκόπηση), για να κατανοήσουμε τις πτυχές του και να δημιουργήσουμε έναν τρόπο προσέγγισης και διερεύνησής του, με τον οποίο θα μπορούσαμε να αντλήσουμε χρήσιμα συμπεράσματα. Ακολουθήθηκε λοιπόν η μεθοδολογία της ποσοτικής έρευνας, η οποία συμπεριλάμβανε τη δημιουργία ενός ερωτηματολογίου με πολλά ερωτήματα σε δύο γλώσσες (ελληνικά και αγγλικά), έχοντας σκοπό να καλύψει όσες περισσότερες πτυχές του θέματος γίνεται, και να εξηγηρηθεί στο μέτρο του δυνατού τις ερευνητικές υποθέσεις. Τα ερωτηματολόγια διανεμήθηκαν ηλεκτρονικά εντός των καλοκαιρινών μηνών του τρέχοντος έτους, σε τετρακόσια επτά (407) ψηφιακά αποθετήρια στην Ελλάδα και στο εξωτερικό. Η διαφοροποίηση της μελέτης αυτής σε σύγκριση με προηγούμενες σημαντικές προσεγγίσεις είναι ότι απευθύνθηκε σε ένα αρκετά μεγάλο αριθμό δείγματος ξεπερνώντας τις γεωγραφικές, γλωσσικές, οικονομικές ή και άλλες τυχόν διακρίσεις, σε μία προσπάθεια κάλυψης όσο το δυνατό διαφορετικότερων μα και τόσο ίδιων τελικά ΨΑ. Σκοπός της έρευνας είναι να μάθουμε πως λειτουργούν τα ΨΑ από τη μία άκρη της γης στην άλλη, ποιες πρακτικές ακολουθούνται κυρίως αναφορικά με τη χρήση προτύπων, την καταχώρηση μεταδεδομένων και την θεματική ευρετηρίαση, αν απασχολεί το ζήτημα της ευρετηρίασης του περιεχομένου των ΨΑ από τις σύγχρονες και δυναμικές μηχανές αναζήτησης όπως το GS, αν έχουν προχωρήσει ήδη σε κάποια ενέργεια για βελτίωση και αύξηση της ευρετηρίασης του υλικού τους και τι προβλήματα

ενδεχομένως αντιμετωπίζουν, τα οποία δεν τους επιτρέπουν να επιτύχουν το ποσοστό ευρετηρίασης και ορατότητας του περιεχομένου του ΨΑ τους από το GS. Επίσης, προτείνεται η λήψη συγκεκριμένων μέτρων, με τα οποία θεωρείται ότι θα υπάρξει θετικό αποτέλεσμα (Πίνακας 23). Αναφορικά με τα αποτελέσματα της έρευνας, πραγματοποιήθηκε έλεγχος ανεξαρτησίας Pearson's Chi-Square (X^2), μία εφαρμογή που μας βοήθησε να εντοπίσουμε την ύπαρξη ή μη πιθανής συσχέτισης μεταξύ των μεταβλητών που ερευνηθήκαν. Οι μεταβλητές αυτές στόχευαν στις ερευνητικές υποθέσεις, επομένως στο αν και κατά πόσο τελικά επηρεάζει η μία μεταβλητή την άλλη. Βάση λοιπόν της στατιστικής ανάλυσης συμπεραίνουμε ότι τελικά δεν υπάρχει σημαντική συσχέτιση μεταξύ του προτύπου μεταδεδομένων με το οποίο περιγράφονται οι πόροι ενός ΨΑ και του ποσοστού ευρετηρίασης των πόρων αυτών από τις μηχανές αναζήτησης όπως το GS. Παρ' όλα αυτά η παρατήρηση του Γράφημα 11 παρουσιάζει ενδιαφέρον, καθώς γίνεται αντιληπτό ότι ΨΑ που έχουν υιοθετήσει και χρησιμοποιούν το πρότυπο Dublin Core για την περιγραφή των πόρων τους, έχουν παραδόξως δηλώσει υψηλότερα ποσοστά ευρετηρίασης του περιεχομένου τους από το GS, σε αντίθεση με άλλα ΨΑ που ενώ δηλώνουν ότι κάνουν χρήση ενός περισσότερο συμβατού/προτεινόμενου προτύπου από το GS, εμφανίζουν χαμηλότερα ποσοστά ευρετηρίασης των πόρων τους από την ίδια μηχανή αναζήτησης. Η διαπίστωση αυτή είναι ανέλπιστα ευχάριστη, κρίνεται όμως απαραίτητη η περαιτέρω διερεύνησή της, δεδομένου ότι τα αποτελέσματα της έρευνάς μας δείχνουν ότι το πρότυπο μεταδεδομένων είναι ανεξάρτητο από το ποσοστό ευρετηρίασης των πόρων από το GS, εν αντιθέσει δηλαδή με τις οδηγίες του GS "Inclusion Guidelines for Webmasters" (Google Scholar, 2010), όπου γίνεται σαφής η σύσταση χρήσης μετα-ετικετών (meta-tags) περιγραφής DC ως την τελευταία επιλογή. Επιπροσθέτως, διενεργήθηκε έλεγχος ανεξαρτησίας Pearson's Chi-Square (X^2) από τον οποίο δεν προέκυψε η ύπαρξη συσχέτισης μεταξύ των μεταβλητών ποσοστό ευρετηρίασης του περιεχομένου του ΨΑ από τις μηχανές αναζήτησης όπως το Google Scholar και ενεργειών στις οποίες έχουν προχωρήσει τα ΨΑ για την αύξηση της ευρετηρίασης του περιεχομένου τους από τις μηχανές αναζήτησης όπως το GS. Παρατηρείται ωστόσο, ότι τα ΨΑ τα οποία έχουν προχωρήσει σε συγκεκριμένες ενέργειες για αύξηση της ευρετηρίασης του περιεχομένου τους από το GS δηλώνουν υψηλότερα ποσοστά ευρετηρίασης από άλλα ΨΑ, τα οποία δεν έχουν προβεί ακόμη σε καμία ενέργεια (Γράφημα 12). Τέλος διενεργήθηκε ακόμη ένας έλεγχος ανεξαρτησίας με τον ίδιο τρόπο, εστιάζοντας αυτή τη φορά στα στοιχεία/πεδία μεταδεδομένων που συμπληρώνονται κατά την περιγραφή των πόρων των ΨΑ, για να ερευνηθεί το ενδεχόμενο ύπαρξης συσχέτισης μεταξύ αυτών των Στοιχείων μεταδεδομένων σε σχέση με το Ποσοστό

ευρετηρίασης του περιεχομένου του ΨΑ κάθε φορέα από τις μηχανές αναζήτησης όπως το GS (Πίνακας 24 & Γράφημα 13 και Πίνακας 26 & Γράφημα 14). Ο έλεγχος πραγματοποιήθηκε για κάθε στοιχείο/πεδίο μεταδεδομένων ξεχωριστά (ως ξεχωριστή μεταβλητή) και προέκυψε ότι το p-value (sig) είναι μικρότερο από 5% για την περίπτωση του πεδίου Συγγραφέας (Sig.=0,026<0,05) (Πίνακας 25) και την περίπτωση του πεδίου Λέξη-Κλειδί (Sig.=0,043<0,05) (Πίνακας 27). Επομένως, ισχύει ότι οι δυο μεταβλητές που εξετάζουμε είναι εξαρτημένες. Με βάση λοιπόν τα αποτελέσματα της διαδικασίας χ^2 φαίνεται να υπάρχει στατιστικά μία σημαντική συσχέτιση μεταξύ των δύο αυτών μεταβλητών Συγγραφέας και Λέξη-Κλειδί με το ποσοστό ευρετηρίασης των πόρων ενός ΨΑ από το GS. Με την επαλήθευση αυτής της υπόθεσης επιβεβαιώνονται και προηγούμενες έρευνες, από τις οποίες προέκυπτε ότι ορισμένα στοιχεία μεταδεδομένων παίζουν πολύ σημαντικό ρόλο στην ανάκτηση των πληροφοριακών πόρων των ΨΑ από τις μηχανές αναζήτησης. Επομένως, η περιγραφή και διάθεση με έγκυρο τρόπο τουλάχιστον ορισμένων συγκεκριμένων στοιχείων μεταδεδομένων, αν όχι του συνόλου των μεταδεδομένων, θεωρείται σημαντική για την ανακάλυψη των ψηφιακών πόρων από τις δυναμικές μηχανές αναζήτησης όπως το GS. Η ορθή περιγραφή των ψηφιακών πόρων που περιέχονται σε δομημένα ψηφιακά αποθετήρια, βασίζεται ωστόσο στη χρήση προτύπων και κυρίως όπως ήδη έχει παρατηρηθεί στη χρήση του πιο διαδεδομένου προτύπου μεταδεδομένων σε ψηφιακά αποθετήρια, δηλαδή του Dublin Core (Arlitsch & O'Brien, 2012; Palmer, Zavalina, & Mustafoff, 2007; Park & Childress, 2009; Simons & Richardson, 2013). Η συλλογή των πρωτογενών δεδομένων της έρευνας αποκάλυψε επίσης αρκετές παρανοήσεις και ελλείψεις, η ύπαρξη των οποίων δεν θεωρούταν πολύ πιθανή προηγουμένως. Οι ελλείψεις αυτές σχετίζονται κυρίως με την άγνοια ορισμένων εκ των συμμετεχόντων της έρευνας περί ευρετηρίασης των ΨΑ από τις μηχανές αναζήτησης όπως το GS, όπως επίσης με την άγνοια ορισμένων από τους συμμετέχοντες σε βιβλιοθηκονομικά θέματα. Η δυσκολία εντοπίστηκε κυρίως όταν κλήθηκαν να απαντήσουν στην ερώτηση π.χ. που αφορά στη διενέργεια κατάλληλων ενεργειών για την αύξηση της ευρετηρίασης των πόρων των ΨΑ από τις μηχανές αναζήτησης (π.χ. Μετατροπή του προτύπου μεταδεδομένων σε ένα άλλο περισσότερο συμβατό με τις μηχανές αναζήτησης, Συμμόρφωση με τις οδηγίες συμπερίληψης από τις μηχανές αναζήτησης (π.χ. Inclusion Guidelines for Webmasters), Ενσωμάτωση ενός μικρού block κώδικα του προτύπου μεταδεδομένων του αποθετηρίου σας στο HTML της ιστοσελίδας κάθε ψηφιακού πόρου σε μορφή μετα-ετικετών (meta tags), κ.λπ.). Υπήρξαν ερωτηματολόγια στα οποία απαντήθηκαν με ασυνέπεια ορισμένες συναφείς ερωτήσεις δημιουργώντας σύγχυση (π.χ. στην ερώτηση «Είστε ικανοποιημένοι

από τον βαθμό ευρετηρίασης του περιεχομένου του ΨΑ σας από τις μηχανές αναζήτησης όπως το GS» μπορεί να έδιναν καταφατική απάντηση «Ναι» και στην επόμενη ερώτηση που ζητούσε να επιλέξουν «Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ψηφιακού αποθετηρίου σας από τις μηχανές αναζήτησης όπως το Google Scholar;» να έδιναν την απάντηση «Δεν γνωρίζω»). Η ασυνέπεια αυτή δηλώνει πολλές φορές άγνοια ή ότι ίσως δεν έχει δοθεί η δέουσα προσοχή σε αυτό το τόσο σημαντικό και καίριο θέμα. Εν μέρει η ασυνέπεια αυτή και η αδυναμία απάντησης θα μπορούσε να δικαιολογηθεί, πρώτον γιατί οι συμμετέχοντες δεν προέρχονται αποκλειστικά από τον κλάδο των βιβλιοθηκονόμων/επιστημόνων πληροφόρησης, επομένως δεν υπάρχει πλήρης και ξεκάθαρη εξειδίκευση και κατά δεύτερον όπως έχει ήδη αναφερθεί η σύσταση του δείγματος ήταν τέτοια, με ΨΑ δηλαδή που τυγχάνουν υψηλής, μέτριας και χαμηλής ευρετηρίασης στο GS, με σκοπό να αποκαλυφθούν τα προβλήματα και οι αδυναμίες.

5.3 Αξιοποίηση / Πρακτικές προεκτάσεις της έρευνας

Η ανάγκη για ανοικτή πρόσβαση, προβολή και ανάδειξη του μόχθου και της έρευνας κάθε ερευνητή/επιστήμονα, σε συνδυασμό με την ανάγκη για αναγνώριση τόσο των δημιουργών όσο και των φορέων που αντιπροσωπεύουν, έχει οδηγήσει τις βιβλιοθήκες και τους οργανισμούς πληροφόρησης γενικότερα στη δημιουργία ΨΑ ως μέσο δημοσιοποίησης της ερευνητικής τους δραστηριότητας. Ωστόσο, η ραγδαία ανάπτυξη της τεχνολογίας και των υπηρεσιών διαδικτύου ευνοούν τη χρήση του Ιστού, πολύ περισσότερο δε παρέχοντας δωρεάν εργαλεία ερευνητικής αναζήτησης, τα οποία προσφέρουν εύκολη και γρήγορη πλοήγηση και ανακάλυψη τεράστιου αριθμού πηγών πλήρους κειμένου στους χρήστες του και χρήσιμες μετρικές με το πάτημα ενός κουμπιού. Πάνω σε αυτή τη βάση και δεδομένου ότι δεν υπάρχει ακριβής γνώση ούτε του επιστημονικού πεδίου που καλύπτεται από το GS ούτε του αριθμού των πληροφοριακών πόρων που ευρετηριάζονται από αυτό (Gusenbauer, 2019; Jacsó, 2012c; Orduña-Malea et al., 2016), πρέπει οι βιβλιοθήκες και οι οργανισμοί πληροφόρησης να βρουν τρόπους να ευρετηριάζεται το περιεχόμενο των ΨΑ τους στον μέγιστο βαθμό κυρίως από την πρώτη σε προτίμηση των χρηστών μηχανή αναζήτησης όπως το GS (Orduña-Malea & Delgado López-Cózar, 2015; Rieger, 2009). Είναι εξαιρετικά σημαντικό το υλικό των ΨΑ να λαμβάνει την προβολή που του αξίζει, έτσι ώστε να αναδεικνύεται το έργο των δημιουργών και των οργανισμών/ιδρυμάτων γενικότερα. Με την ευρετηρίαση του υλικού των ΨΑ από τις μηχανές αναζήτησης όπως το GS δημιουργούνται καλύτερες προοπτικές και για τους δημιουργούς και για τα ιδρύματα που αυτοί εκπροσωπούν. Δίνεται λοιπόν η δυνατότητα στους δημιουργούς για επίτευξη

μεγαλύτερης αναγνώρισης και φήμης, λήψης υποτροφιών και χορηγιών έρευνας, επίτευξη εξέλιξης, κύρους σε προσωπικό και επιστημονικό επίπεδο. Είναι επίσης εξαιρετικά σημαντικό οι βιβλιοθήκες και οι πληροφοριακοί οργανισμοί μέσω των ΨΑ τους να κινητοποιηθούν, ώστε να παρέχουν καλύτερες υπηρεσίες, προωθώντας πάντα την προαγωγή κι εξέλιξη της γνώσης κι ενδεχομένως να κερδίσουν την αναγνώριση του κοινού, σε αντίθεση με την επικρατούσα κατάσταση (De Rosa et al., 2010). Με την παρούσα μελέτη γίνεται μία προσπάθεια ανάδειξης του προβλήματος, ευαισθητοποίησης των επαγγελματιών πληροφόρησης και των απασχολούμενων στα ΨΑ γενικότερα, για τις δυνατότητες και τις προεκτάσεις που μπορεί να λάβει η δουλειά τους αν κινηθούν προς την κατεύθυνση λήψης μέτρων και επίλυσης των προβλημάτων που καταγράφονται. Η στατιστική ανάλυση της παρούσας έρευνας επαληθεύει μόνο τη μία εκ των τριών υποθέσεων. Συγκεκριμένα την υπόθεση ότι η περιγραφή και παροχή με έγκυρο τρόπο συγκεκριμένων στοιχείων μεταδεδομένων (στοιχείο Συγγραφέας και στοιχείο Λέξη-κλειδί), επηρεάζει την ανακάλυψη και ευρετηρίαση των ψηφιακών πόρων των αποθετηρίων από τις δυναμικές μηχανές αναζήτησης όπως το GS. Έχοντας επισημάνει λοιπόν την ύπαρξη μεγάλου αριθμού σφαλμάτων στα μεταδεδομένα που υπάρχουν στη μηχανή αναζήτησης GS, μπορούμε να αντιληφθούμε την αναγκαιότητα διάθεσης ποιοτικών μεταδεδομένων, δίνοντας βαρύτητα στην τυποποίηση και στην χρήση προτύπων όπως το Dublin Core, με τα οποία μπορούμε να περιγράψουμε με συνέπεια και αξιοπιστία τους πόρους των ΨΑ. Παρά του γεγονότος ότι δεν επαληθεύτηκαν οι πρώτες δύο υποθέσεις της έρευνας, θεωρούμε ότι δίνονται τα ερεθίσματα για περαιτέρω δοκιμές. Όπως ήδη προσέξαμε αν και δεν εντοπίζεται συσχέτιση ούτε του προτύπου μεταδεδομένων με το ποσοστό ευρετηρίασης του ΨΑ από το GS, ούτε συσχετίζονται οι ληφθείσες βελτιωτικές ενέργειες με το ποσοστό ευρετηρίασης του περιεχομένου του ΨΑ από τις μηχανές αναζήτησης, πάλι έχουμε καταγράψει καλύτερα ποσοστά ευρετηρίασης στις περιπτώσεις που έχουν ληφθεί μέτρα, όπως επίσης αυξημένα ποσοστά ευρετηρίασης έχουν και πάρα πολλά ψηφιακά αποθετήρια που χρησιμοποιούν εντούτοις ως πρότυπο μεταδεδομένων τους το DC και όχι κάποιο από τα πλέον προτεινόμενα από το GS σχήματα. Τέλος, εκτιμάται ότι αξιοποιώντας και εφαρμόζοντας τις προτεινόμενες ενέργειες βελτίωσης, θα προκύψουν καλύτερα αποτελέσματα ευρετηρίασης του περιεχομένου των ΨΑ από το GS.

5.4 Μελλοντικές επεκτάσεις / Πρακτικές Προεκτάσεις της Έρευνας

Είναι γεγονός ότι οι περιορισμοί μίας έρευνας δημιουργούν τις κατ' εξοχήν ιδανικές συγκυρίες/προτάσεις για περαιτέρω έρευνα. Στη συγκεκριμένη περίπτωση όπως είναι φυσικό, το δείγμα παίζει πολύ σημαντικό ρόλο, επομένως η επιλογή άλλου δείγματος, ενδεχομένως περισσότερο διευρυμένου ή επιλεγμένου με διαφορετικά κριτήρια, θα μπορούσε να αποδώσει και διαφορετικά αποτελέσματα. Πιθανώς να επέτρεπε και τη γενίκευση των αποτελεσμάτων της έρευνας. Επίσης, η διενέργεια της έρευνας κάνοντας χρήση ενός άλλου ερευνητικού εργαλείου ή και συνδυασμού δύο εργαλείων (μετρικές και ερωτηματολόγιο), πιθανώς να προσέφερε νέα αποτελέσματα και διαφορετικές προσεγγίσεις. Αντίστοιχα, η χρονική περίοδος που διεξήχθη η έρευνα, έχοντας λάβει υπόψη συγκεκριμένης βιβλιογραφίας είναι σημαντική, κι επειδή η επιστήμη εξελίσσεται δεν θα μπορούμε να γνωρίζουμε εκ των προτέρων την εξέλιξη μιας τέτοιας έρευνας.

Βιβλιογραφικές Αναφορές

- Acharya, A. (2015). Indexing Repositories : Pitfalls & Best Practices. *International Conference on Open Repositories (10th : 2015 : Indianapolis, Ind.)*. Retrieved from <http://www.or2015.net/wp-content/uploads/2015/06/or-2015-anurag-google-scholar.pdf>
- Aguillo, I. (2009). Measuring the Institution's Footprint in the Web. *Library Hi Tech, Vol. 27*(No. 4), 540–556. <https://doi.org/https://doi.org/10.1108/073788309>
- Aguillo, I. F. (2018). Altmetrics of the Open Access Institutional Repositories: A Webometrics Approach. *23rd International Conference on Science and Technology Indicators (STI 2018), September 12–14, 2018, Leiden*, 159–169. Retrieved from <https://hdl.handle.net/1887/64521>
- Aithal, P. S. (2017). Comparative study of various research indices used to measure quality of research publications. *International Journal of Applied and Advanced Scientific Research (IJASR) Impact Factor: 5.255*, 2(1), 81–89. Retrieved from www.dvpublication.com
- Archambault, E., Amyot, D., Deschamps, P., Nicol, A., Rebout, L., & Roberge, G. (2013). *Proportion of Open Access Peer-Reviewed Papers at the European and World Levels-2004-2011*. <https://doi.org/10.1371/journal.pone.0011273>
- Arlitsch, Kenning, O'Brien, P. S. (2013). Improving the Visibility and Use of Digital Repositories through SEO A LITA Guide. In *LITA Guide*. Retrieved from <https://eds.b.ebscohost.com/eds/ebookviewer/ebook/bmxlYmtfXzU3ODU1MV9fQU41?sid=8bdec2b7-f908-411c-8767-d26e5c3d66fd@sessionmgr102&vid=0&format=EB&rid=1>
- Arlitsch, K., & O'Brien, P. S. (2012). Invisible institutional repositories: Addressing the low indexing ratios of IRs in Google. *Library Hi Tech*, 30(1), 60–81. <https://doi.org/10.1108/07378831211213210>
- Armbruster, C., & Romary, L. (2010). Comparing Repository Types-Challenges and barriers for subject-based repositories, research repositories, national repository systems and institutional repositories in serving scholarly communication. *International Journal of Digital Library Systems*, 1(4), 61–73. <https://doi.org/10.4018/jdls.2010100104>

- Beel, J., & Gipp, B. (2009). Google scholar's ranking algorithm: The impact of articles' age (an empirical study). *ITNG 2009 - 6th International Conference on Information Technology: New Generations*, (April), 160–164. <https://doi.org/10.1109/ITNG.2009.317>
- Belew, R. K. (2005). *Scientific impact quantity and quality: Analysis of two sources of bibliographic data*.
- Berners-Lee, T. (1997). Axioms of Web Architecture: Metadata. Retrieved January 22, 2020, from <https://www.w3.org/DesignIssues/Metadata.html>
- De Rosa, C., Cantrell, J., Carlson, M., Gallagher, P., Hawk, J., & Sturtz, C. (2010). *Perceptions of Libraries, 2010: Context and Community: A REPORT TO THE OCLC MEMBERSHIP* (B. Gauder, Ed.). Retrieved from https://www.oclc.org/content/dam/oclc/reports/2010perceptions/2010perceptions_all_singlepage.pdf
- De Rosa, C., Cantrell, J., Cellentani, D., Hawk, J., Jenkins, L., & Wilson, A. (2005). *Perceptions of Libraries and Information Resources: A Report to the OCLC Membership*. Dublin, Ohio: OCLC Online Computer Library Center, Inc.
- DSpace - A Turnkey Institutional Repository Application. (n.d.). Retrieved May 31, 2020, from <https://duraspace.org/dspace/>
- Dublin Core Metadata Initiative. (n.d.-a). DCMI: Guidelines for Encoding Bibliographic Citation Information in Dublin Core Metadata. Retrieved December 8, 2019, from <https://www.dublincore.org/specifications/dublin-core/dc-citation-guidelines/>
- Dublin Core Metadata Initiative. (n.d.-b). Dublin Core™ Metadata Initiative: ISO. Retrieved June 2, 2020, from <https://www.dublincore.org/collaborations/iso/>
- Dublin Core Metadata Initiative. (n.d.-c). Dublin Core™ Metadata Initiative (DCMI). Retrieved April 30, 2020, from <https://dublincore.org/>
- Ebadi, A., & Schiffauerova, A. (2015). How to receive more funding for your research? get connected to the right people. *PLoS ONE*, 10(7). <https://doi.org/10.1371/journal.pone.0133061>
- Fedora - The Flexible, Modular, Open-Source Repository Platform. (n.d.). Retrieved May 31, 2020, from <https://duraspace.org/fedora/>
- Giustini, D., & Kamel Boulos, M. N. (2013). Google Scholar is not enough to be used alone for systematic reviews. *Online Journal of Public Health Informatics*, 5(2), 1–9. <https://doi.org/10.5210/ojphi.v5i2.4623>

- Google Scholar. (n.d.). Retrieved January 24, 2020, from In Wikipedia website:
https://en.wikipedia.org/wiki/Google_Scholar
- Google Scholar. (2010). Inclusion Guidelines for Webmasters. Retrieved December 8, 2019,
from <https://scholar.google.com/intl/en/scholar/inclusion.html>
- Greenberg, J. (2003). Metadata and the World Wide Web. In M. A. Drake (Ed.), *Encyclopedia of Library and Information Science* (2nd ed., pp. 1876–1888). <https://doi.org/DOI:10.1081/E-ELIS120008663>
- Greenstone Digital Library Software. (n.d.). Retrieved May 31, 2020, from
<http://www.greenstone.org/>
- Gusenbauer, M. (2019). Google Scholar to overshadow them all? Comparing the sizes of 12 academic search engines and bibliographic databases. *Scientometrics*, *118*(1), 177–214. <https://doi.org/10.1007/s11192-018-2958-5>
- Harzing, A. W. K., & van der Wal, R. (2008). Google Scholar as a new source for citation analysis. *Ethics in Science and Environmental Politics*, *8*(1), 61–73. <https://doi.org/10.3354/ese00076>
- Hightower, C., & Caldwell, C. (2010). Shifting Sands: Science Researchers on Google Scholar, Web of Science, and PubMed, with Implications for Library Collections Budgets. *Issues in Science and Technology Librarianship (Online Journal)*, (Fall). <https://doi.org/10.5062/F4V40S4J>
- Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(46), 16569–16572. Retrieved from www.pnas.org/cgi/doi/10.1073/pnas.0507655102
- Ho, J., & Stokes, C. (2019). Core Metadata Element Recommendations for Institutional Repositories at Texas A&M University Libraries. *Journal of Library Metadata*, *19*(3–4), 1–28. <https://doi.org/10.1080/19386389.2019.1651499>
- Hoenisch, S. (2009). SEO with Dublin Core. Retrieved April 22, 2020, from <https://www.criticism.com/seo/dublin-core-metadata.php>
- Jacsó, P. (2005). Google scholar: The pros and the cons. *Online Information Review*, *29*(2), 208–214. <https://doi.org/10.1108/14684520510598066>
- Jacsó, P. (2006). Deflated, inflated and phantom citation counts. *Online Information Review*, *30*(3), 297–309. <https://doi.org/10.1108/14684520610675816>
- Jacsó, P. (2008a). Google Scholar revisited. *Online Information Review*, *32*(1), 102–114.

<https://doi.org/10.1108/14684520810866010>

- Jacsó, P. (2008b). The pros and cons of computing the h-index using Google Scholar. *Online Information Review*, 32(3), 437–452. <https://doi.org/10.1108/14684520810889718>
- Jacsó, P. (2009a). Calculating the h-index and other bibliometric and scientometric indicators from Google Scholar with the Publish or Perish software. *Online Information Review*, 33(6), 1189–1200. <https://doi.org/10.1108/14684520911011070>
- Jacsó, P. (2009b). Google Scholar's ghost authors. *Library Journal*, 26–27. Retrieved from <http://www.urayasu.meikai.ac.jp/japanese/meikainihongo/14/matsuda.pdf>
- Jacsó, P. (2010). Metadata mega mess in Google Scholar. *Online Information Review*, 34(1), 175–191. <https://doi.org/10.1108/14684521011024191>
- Jacsó, P. (2011). Google Scholar duped and deduped - The aura of “robometrics.” *Online Information Review*, 35(1), 154–160. <https://doi.org/10.1108/14684521111113632>
- Jacsó, P. (2012a). Google scholar author citation tracker: Is it too little, too late? *Online Information Review*, 36(1), 126–141. <https://doi.org/10.1108/14684521211209581>
- Jacsó, P. (2012b). Google Scholar Metrics for Publications: The software and content features of a new open access bibliometric service. *Online Information Review*, 36(4), 604–619. <https://doi.org/10.1108/14684521211254121>
- Jacsó, P. (2012c). Using Google Scholar for journal impact factors and the h-index in nationwide publishing assessments in academia-siren songs and air-raid sirens. *Online Information Review*, 36(3), 462–478. <https://doi.org/10.1108/14684521211241503>
- Kozak, M., & Bornmann, L. (2012). A New Family of Cumulative Indexes for Measuring Scientific Performance. *PLoS ONE*, 7(10), 1–5. <https://doi.org/10.1371/journal.pone.0047679>
- Kroll, S., & Forsman, R. (2010). *A Slice of Research Life : Information Support for Research in the United States Please direct correspondence to :* (June), 1–24.
- Kurtz, M. (2010). Dublin core, dspace, and a brief analysis of three university repositories. *Information Technology and Libraries*, 29(1), 40–46. <https://doi.org/10.6017/ital.v29i1.3157>
- Majeed, F., Shafiq, M., Ali, A., Hassan, M. A., Abbas, S. A., Alzahrani, M. E., ... Irshad, A. (2019). Self-Citation Analysis on Google Scholar Dataset for H-Index Corrections. *IEEE Access*, 7, 126025–126036. <https://doi.org/10.1109/ACCESS.2019.2938657>
- Martín-Martín, A., Orduña-Malea, E., Ayllón, J. M., & Delgado López-Cózar, E. (2014). *Does*

- Google Scholar contain all highly cited documents (1950-2013)?* (No. 19). Retrieved from <http://arxiv.org/abs/1410.8464>
- Martín-Martín, A., Orduña-Malea, E., Harzing, A. W., & Delgado López-Cózar, E. (2017). Can we use Google Scholar to identify highly-cited documents? *Journal of Informetrics*, *11*(1), 152–163. <https://doi.org/10.1016/j.joi.2016.11.008>
- Mingers, J., & Meyer, M. (2017). Normalizing Google Scholar data for use in research evaluation. *Scientometrics*, *112*(2), 1111–1121. <https://doi.org/10.1007/s11192-017-2415-x>
- Mondal, T., Meenachi, N. M., & Sai Baba, M. (2018). Metadata Creation Methods: A Study. *Journal of Advances in Library and Information Science*, *7*(1), 177–182. Retrieved from <http://jhove.openpreservation.org/>
- Mueller, J. (2008). Retiring support for OAI-PMH in Sitemaps Wednesday, April 23, 2008. Retrieved June 15, 2020, from Google Webmaster Central Blog website: <https://webmasters.googleblog.com/2008/04/retiring-support-for-oai-pmh-in.html>
- National Information Standards Organization. (2001). *The Dublin Core Metadata Element Set*. Retrieved from www.niso.org
- OAI-PMH Implementation Guidelines. (n.d.). Retrieved May 31, 2020, from <https://www.openarchives.org/OAI/2.0/guidelines.htm>
- Omeka. (n.d.). Retrieved May 14, 2020, from <https://omeka.org/>
- Open access. (n.d.). Retrieved February 17, 2020, from In Wikipedia website: https://en.wikipedia.org/wiki/Open_access
- Orduña-Malea, E., & Delgado López-Cózar, E. (2015). The dark side of open access in google and google scholar: The case of latin-american repositories. *Scientometrics*, *102*(1), 829–846. <https://doi.org/10.1007/s11192-014-1369-5>
- Orduña-Malea, E., Martín-Martín, A., Ayllón, J. M., & Delgado López-Cózar, E. (2016). *La revolución Google Scholar: Destapando la caja de Pandora académica*. Retrieved from https://www.researchgate.net/publication/309741937_La_revolucion_Google_Scholar_Destapando_la_caja_de_Pandora_academica
- Orduña-Malea, E., Martín-Martín, A., & Delgado López-Cózar, E. (2017). Google Scholar as a source for scholarly evaluation: A bibliographic review of database errors. *Revista Española de Documentación Científica*, *40*(4), 1–33. <https://doi.org/10.3989/redc.2017.4.1500>

- Ortega, J. L. (2014). Academic Search Engines: A Quantitative Outlook. In *Academic Search Engines: A Quantitative Outlook*. <https://doi.org/doi.org/10.1016/C2013-0-23226-8>
- Palmer, C. L., Zavalina, O. L., & Mustafoff, M. (2007). Trends in metadata practices: A longitudinal study of collection federation. *Proceedings of the ACM International Conference on Digital Libraries*, (c), 386–395. <https://doi.org/10.1145/1255175.1255251>
- Park, J. R., & Childress, E. (2009). Dublin Core metadata semantics: An analysis of the perspectives of information professionals. *Journal of Information Science*, 35(6), 727–739. <https://doi.org/10.1177/0165551509337871>
- Park, J. R., & Tosaka, Y. (2010). Metadata creation practices in digital repositories and collections: Schemata, selection criteria, and interoperability. *Information Technology and Libraries*, 29(3), 104–116. <https://doi.org/10.6017/ital.v29i3.3136>
- Rieger, O. Y. (2009). Search engine use behavior of students and faculty: User perceptions and implications for future research. *First Monday*, 14(12). <https://doi.org/10.5210/fm.v14i12.2716>
- Safari, M. (2005). Search Engines and Resource Discovery on the Web: Is Dublin Core an Impact Factor? *Webology*, 2(2). Retrieved from <http://webology.org/2005/v2n2/a13.html>
- Schonfeld, R. C., & Housewright, R. (2010). *Faculty Survey 2009 : Key Strategic Insights for Libraries , Publishers , and Societies*. <https://doi.org/https://doi.org/10.18665/sr.22364>
- Shanks, J., & Arlitsch, K. (2016). Making Sense of Researcher Services. *Journal of Library Administration*, 56(3), 295–316. <https://doi.org/10.1080/01930826.2016.1146534>
- Simons, N., & Richardson, J. (2013). *New content in digital repositories : the changing research landscape*. Philadelphia, PA: Chandos Publishing.
- Suber, P. (2004). A very brief introduction to Open Access. Retrieved February 12, 2020, from <https://legacy.earlham.edu/~peters/fos/brief.htm>
- Tonkin, E., & Muller, H. L. (2008). Keyword and metadata extraction from pre-prints. *Open Scholarship: Authority, Community, and Sustainability in the Age of Web 2.0 - Proceedings of the 12th International Conference on Electronic Publishing, ELPUB 2008*, (January 2008), 30–44.
- Verrelli, D. I. (2018). Metadata tags for academic publications. Retrieved June 19, 2020, from Division One, Academic and Language Services website: <http://div.div1.com.au/div->

thoughts/div-commentaries/66-div-commentary-metadata

- Yang, L. (2016). Metadata Effectiveness in Internet Discovery: An Analysis of Digital Collection Metadata Elements and Internet Search Engine Keywords. *College & Research Libraries*, 77(1), 7–19. Retrieved from <https://crl.acrl.org/index.php/crl/article/view/16487/17933>
- Zhang, J., & Dimitroff, A. (2004). Internet search engines' response to metadata Dublin Core implementation. *Journal of Information Science*, 30(4), 310–320. <https://doi.org/10.1177/0165551504045851>
- Ζαφειρόπουλος, Κ. (2015). *Βοηθητικές σημειώσεις για SPSS. Στατιστική ανάλυση με χρήση Η/Υ*. Retrieved from <http://opencourses.uom.gr/assets/site/content/courses/72/Notes-SPSS.pdf>
- Καμάτσος, Π., Σωτηροπούλου, Α., Θεοφιλάτου, Ά., Μπογιατζή, Ε., & Κακάλη, Κ. (2013). *Η μετεξέλιξη ενός αποθετηρίου κοινωνικών επιστημών: το παράδειγμα της Πάνδημου σε περιβάλλον Fedora*. 1–16. Retrieved from http://eprints.rclis.org/22803/1/pandemos_paper%283%29.pdf
- Καπιδάκης, Σ., Λαζαρίνης, Φ., & Τοράκη, Κ. (2015). *Θέματα Βιβλιοθηκονομίας και Επιστήμης των Πληροφοριών*. Retrieved from <http://hdl.handle.net/11419/1674>
- Κυριάκη-Μάνεση, Δ., & Κουλούρης, Α. (2015). *Διαχείριση Ψηφιακού Περιεχομένου*. Retrieved from https://repository.kallipos.gr/bitstream/11419/2496/1/9558_master_document.pdf
- Παπαβλασόπουλος, Σ. (2015). *Βιβλιομετρία*. <https://doi.org/http://hdl.handle.net/11419/4755>
- Ρούσσο, Π., & Ευσταθίου, Γ. (2008). *Σύντομο εγχειρίδιο SPSS 16.0*. Retrieved from http://old.psych.uoa.gr/~roussosp//stats/Manual_SPSS16.pdf
- Σελίμης, Γ., & Σιμακάκης, Δ. (2016). *Αξιολόγηση της αξιοπιστίας του Google Scholar* (Πανεπιστήμιο Πελοποννήσου, Σχολή Οικονομίας, Διοίκησης και Πληροφορικής). Retrieved from <http://gav.uop.gr/docs/thesis/BSc2016SelimisSimiakakis.pdf>

Παράρτημα Α' – Σχέδιο Ερωτηματολογίου (ελληνική έκδ.)

Ερώτηση 1^η : Φύλο

Ερώτηση 2^η: Ηλικία

Ερώτηση 3^η: Ποια η θέση σας στον Φορέα;

Ερώτηση 4^η: Τι είδους ψηφιακό αποθετήριο διαθέτει ο φορέας σας;

Ερώτηση 5^η: Ποια είδη υλικού περιλαμβάνονται στο αποθετήριό σας;

Ερώτηση 6^η: Πως γίνεται η απόθεση των ψηφιακών πόρων στο αποθετήριό σας;

Ερώτηση 7^η: Ποιο λογισμικό έχετε χρησιμοποιήσει για την υλοποίηση του αποθετηρίου σας;

Ερώτηση 8^η: Χρησιμοποιείτε την πιο ενημερωμένη (τελευταία) έκδοση του λογισμικού του ψηφιακού αποθετηρίου σας ή κάποια παλαιότερη;]

Ερώτηση 9^η: Σύμφωνα με ποιο πρότυπο περιγράφονται οι ψηφιακοί πόροι στο αποθετήριό σας;

Ερώτηση 10^η: Ποια είναι τα απαιτούμενα στοιχεία μεταδεδομένων που καταχωρίζονται στο αποθετήριό σας;

Ερώτηση 11^η: Είναι προκαθορισμένα ή κρίνονται κατά περίπτωση;

Ερώτηση 12^η: Θεωρείτε σημαντική την παρακολούθηση και αφομοίωση των εξελίξεων και των ενημερώσεων (νέες εκδόσεις) του προτύπου μεταδεδομένων που χρησιμοποιείται από το ψηφιακό σας αποθετήριο;

Ερώτηση 13^η: Θεωρείτε ότι η τυποποίηση των μεταδεδομένων είναι σημαντική για την ανάκτηση των ψηφιακών πόρων;

Ερώτηση 14^η: Ποια δήλωση από τις ακόλουθες περιγράφει καλύτερα τον τρόπο της θεματικής ευρετηρίασης του ψηφιακού αποθετηρίου σας;

Ερώτηση 15^η: Θεωρείτε ότι η ανακάλυψη των ψηφιακών πόρων του αποθετηρίου σας εξυπηρετείται με επάρκεια από τις μηχανές αναζήτησης όπως το Google Scholar;

Ερώτηση 16^η: Έχετε ελέγξει τον βαθμό ικανοποίησης σε σχέση με την ορατότητα, τις προβολές και τις λήψεις του περιεχομένου του αποθετηρίου σας από μηχανές αναζήτησης όπως το Google Scholar;

Ερώτηση 17^η: Είστε ικανοποιημένοι από τον βαθμό ευρετηρίασης του περιεχομένου του αποθετηρίου σας από το Google Scholar;

Ερώτηση 18^η: Θεωρείτε ότι το υλικό του αποθετηρίου σας δεν χρειάζεται να ευρετηριάζεται από τις μηχανές αναζήτησης όπως το Google Scholar;

Ερώτηση 19^η: Ποιο είναι το ποσοστό ευρετηρίασης του περιεχομένου του ψηφιακού αποθετηρίου σας από τις μηχανές αναζήτησης όπως το Google Scholar;

Ερώτηση 20^η: Θεωρείτε ότι το πρότυπο μεταδεδομένων παίζει σημαντικό ρόλο στην ευρετηρίαση των ψηφιακών πόρων του αποθετηρίου σας από τις μηχανές ερευνητικής αναζήτησης όπως το Google Scholar;

Ερώτηση 21^η: Έχετε προχωρήσει σε κατάλληλες ενέργειες για την επίτευξη αύξησης της ευρετηρίασης του περιεχομένου του ψηφιακού αποθετηρίου σας από μηχανές αναζήτησης όπως το Google Scholar;

Ερώτηση 22^η: Παρακαλώ κατονομάστε τις ενέργειες στις οποίες έχετε προχωρήσει για την αύξηση της ευρετηρίασης του περιεχομένου του ψηφιακού σας αποθετηρίου από τις μηχανές αναζήτησης όπως το Google Scholar.

Ερώτηση 23^η: Ποια προβλήματα θεωρείτε ότι υφίστανται και επηρεάζουν την ευρετηρίαση του περιεχομένου του αποθετηρίου σας από τις μηχανές αναζήτησης όπως το Google Scholar;

Παράρτημα Β' – Σχέδιο Ερωτηματολογίου (αγγλική έκδ.)

Question 1st: Gender

Question 2nd: Age range

Question 3rd: Position in your organization

Question 4th: Your organization's type of digital repository

Question 5th: Material types presented in the digital repository at your institution/organization?

Question 6th: How are the documents deposited in your repository?

Question 7th: Repository software

Question 8th: Do you use the latest version of your repository software?

Question 9th: Metadata standards employed in your digital repository?

Question 10th: Which are the minimum requirements of metadata elements within your digital repository?

Question 11th: Are they predetermined or judged on a case-by-case basis?

Question 12th: Do you believe that is important to keep up with developments relatively to the updates of the metadata standard used by your digital repository?

Question 13th: Do you think that standardization of metadata is important for the discovery of digital resources?

Question 14th: Which of the following statements best describes the subject indexing method of your digital repository?

Question 15th: Do you think the discovery of digital resources of your repository is adequately served by search engines, like Google Scholar?

Question 16th: Do you know the degree of satisfaction regarding the visibility of the content of your digital repository?

Question 17th: Are you satisfied by the rate of content indexing of your repository by Google Scholar?

Question 18th: Do you feel that your content does not need to be indexed by search engines?

Question 19th: Indexing ratio of your digital repository in Google Scholar

Question 20th: Do you think that the metadata standard plays an important role in indexing the items that are part of your digital repository collections, from search engines such as Google Scholar?

Question 21st: Have you taken the appropriate action to enhance the indexing rate of your digital repository content from search engines such as Google Scholar?

Question 22nd: Please name the specific action you have taken to enhance the indexing rate of your repository content from search engines such as Google Scholar

Question 23rd: Which problems have you detected in relation to the indexing rate of your digital repository in Google Scholar?

Παράρτημα Γ' - Πίνακας Συμμετεχόντων

Πίνακας 28. Πίνακας συμμετεχόντων

Ελληνικά Ψηφιακά Αποθετήρια			
A/A	Σειρά κατάταξης του αποθετηρίου στο Ranking Web of Repositories (9th ed. Apr. 2020)	Τίτλος αποθετηρίου	Google Scholar records
1	Εκτός κατάταξης	arothesis - Ιδρυματικό αποθετήριο Ελληνικού Ανοικτού Πανεπιστημίου	
2	Εκτός κατάταξης	Arothesis - Ιδρυματικό αποθετήριο Ελληνικού Μεσογειακού Πανεπιστημίου (πρώην ΤΕΙ Κρήτης)	
3	2675	E-Locus - Ιδρυματικό αποθετήριο Πανεπιστημίου Κρήτης	19
4	Εκτός κατάταξης	Oeconomicus - Ψηφιακό αποθετήριο Κέντρου Προγραμματισμού και Οικονομικών Ερευνών (ΚΕΠΕ)	
5	Εκτός κατάταξης	RUOMO - Πανεπιστήμιο Μακεδονία	
6	Εκτός κατάταξης	Αθηνά - Ιδρυματικό αποθετήριο πρώην ΤΕΙ Δυτικής Ελλάδας	
7	Εκτός κατάταξης	Ακρόπολις - Αποθετήριο εκπαιδευτικού περιεχομένου για την Ακρόπολη	
8	Εκτός κατάταξης	Αμητός - Ιδρυματικό αποθετήριο Πανεπιστημίου Πελοποννήσου	

9	Εκτός κατάταξης	Απόθεσις - Ιδρυματικό αποθετήριο πρώην ΤΕΙ Κεντρικής Μακεδονίας	
10	Εκτός κατάταξης	Αποθετήριο Ακαδημίας Αθηνών	
11	2369	Αποθετήριο ΕΛ/ΛΑΚ	173
12	Εκτός κατάταξης	Αρχιπέλαγος - Πανεπιστήμιο Αιγαίου	
13	Εκτός κατάταξης	Ασκληπιός- Ιδρυματικό Αποθετήριο Επιστημών Υγείας Ασκληπιείου Βούλας	
14	Εκτός κατάταξης	Γνώσις - Ιδρυματικό αποθετήριο Πανεπιστημίου Κύπρου	
15	393	Διώνη - Ιδρυματικό αποθετήριο Πανεπιστημίου Πειραιώς	10900
16	Εκτός κατάταξης	Εθνικό Αρχείο Διδακτορικών Διατριβών - Εθνικό Κέντρο Τεκμηρίωσης	
17	Εκτός κατάταξης	Ελλάνικος - Πανεπιστήμιο Αιγαίου	
18	Εκτός κατάταξης	Ερευνητικό Αποθετήριο της Παν/πολης 2 Πανεπιστημίου Δυτικής Αττικής	
19	Εκτός κατάταξης	Ερευνητικό Αποθετήριο πρώην ΤΕΙ Ηπείρου (CRIS)	
20	Εκτός κατάταξης	Εστία - Ιδρυματικό αποθετήριο Χαροκοπείου Πανεπιστημίου	
21	Εκτός κατάταξης	Εύρηκα - Ιδρυματικό αποθετήριο πρώην ΑΤΕΙ Θεσσαλονίκης	
22	Εκτός κατάταξης	Θεόδωρος Τριβέλλας - Ιδρυματικό αποθετήριο πρώην ΤΕΙ Ανατολικής Μακεδονίας και Θράκης	
23	Εκτός κατάταξης	Ιδρυματικό αποθετήριο	

		Δημοκριτείου Πανεπιστημίου Θράκης	
24	1466	Ιδρυματικό αποθετήριο Διεθνούς Πανεπιστημίου Ελλάδος	1320
25	307	Ιδρυματικό Αποθετήριο Εθνικού Μετσόβιου Πολυτεχνείου	13900
26	1687	Ιδρυματικό Αποθετήριο Πανεπιστημίου Θεσσαλίας	862
27	Εκτός κατάταξης	Ιδρυματικό Αποθετήριο Πολυτεχνείου Κρήτης	
28	Εκτός κατάταξης	Ιδρυματικό Αποθετήριο πρώην ΤΕΙ Ηπείρου	
29	Εκτός κατάταξης	Ιδρυματικό Αποθετήριο Ωκεανός - ΕΛΚΕΘΕ	
30	449	Ιδρυματικό Καταθετήριο @νάκτησις (Ανάκτησις) - Πρώην ΤΕΙ Δυτικής Μακεδονίας	9340
31	Εκτός κατάταξης	Ιδρυματικό Καταθετήριο Επιστημονικών Εργασιών (ΙΚΕΕ) - Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης	
32	Εκτός κατάταξης	Ιδρυματικό Καταθετήριο Επιστημονικών Εργασιών (ΙΚΕΕ) - Πανεπιστήμιο Δυτικής Μακεδονίας	
33	Εκτός κατάταξης	Κάλλιπος - Ψηφιακό Αποθετήριο Ανοικτής Πρόσβασης (Επιστημονικά συγγράμματα)	
34	934	Κτίσις - Ιδρυματικό καταθετήριο Τεχνολογικού Πανεπιστημίου Κύπρου (ΤΕΠΑΚ)	3390
35	Εκτός κατάταξης	Κυψέλη - Ιδρυματικό Αποθετήριο Ανοικτού	

		Πανεπιστημίου Κύπρου	
36	Εκτός κατάταξης	Κυψέλη - Πανεπιστήμιο Κρήτης	
37	2773	ΛΗΚΥΘΟΣ - Ιδρυματικό αποθετήριο και Ψηφιακή Βιβλιοθήκη Πανεπιστημίου Κύπρου	0
38	Εκτός κατάταξης	Μέδουσα - Δημόσια Κεντρική Βιβλιοθήκη Βέροιας	
39	398	Νημερτής - Πανεπιστήμιο Πατρών	10700
40	Εκτός κατάταξης	Ολυμπιάς - Ιδρυματικό αποθετήριο Πανεπιστημίου Ιωαννίνων	
41	1408	Πάνδημος - Ακαδημαϊκό καταθετήριο Παντείου Πανεπιστημίου Κοινωνικών και Πολιτικών Επιστημών	1480
42	Εκτός κατάταξης	Πέργαμος - Ιδρυματικό αποθετήριο Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών	
43	Εκτός κατάταξης	Προθήκη - Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης	
44	Εκτός κατάταξης	Πυξίδα - Ιδρυματικό αποθετήριο Οικονομικού Πανεπιστημίου Αθηνών	
45	Εκτός κατάταξης	Σοφία - Τμήμα Μουσικών Σπουδών Αριστοτελείου Πανεπιστημίου Θεσσαλονίκης	
46	Εκτός κατάταξης	Υπατία - Ιδρυματικό Αποθετήριο της Παν/πολης 1 Πανεπιστημίου Δυτικής Αττικής	
47	Εκτός κατάταξης	Ψηφιακή Βιβλιοθήκη - Δημόσια Κεντρική Βιβλιοθήκη Σερρών	

48	Εκτός κατάταξης	Ψηφιακή Βιβλιοθήκη -Δημόσια Κεντρική Βιβλιοθήκη Λεβαδείας	
49	Εκτός κατάταξης	Ψηφιακό αποθετήριο Αρχείων Σύγχρονης Κοινωνικής Ιστορίας (ΑΣΚΙ)	
50	310	Ψηφιακό Αποθετήριο Γεωπονικού Πανεπιστημίου Αθηνών	13800
51	Εκτός κατάταξης	Ψηφιακό Αποθετήριο Ιδρύματος Ευγενίδου	
52	Εκτός κατάταξης	Ψηφιακό Αποθετήριο Ιονίου Πανεπιστημίου	
53	Εκτός κατάταξης	Ψηφιακό Αποθετήριο ΤΕΙ Ιονίων Νήσων	
54	Εκτός κατάταξης	Ψηφιακό Αποθετήριο της Κεντρικής Βιβλιοθήκης του ΕΜΠ	
55	Εκτός κατάταξης	Ψηφιακό Αποθετήριο του Ελληνικού Κέντρου Βιοτόπων-Υγροτόπων (ΕΚΒΥ)	
56	467	Ψηφίδα - Πανεπιστήμιο Μακεδονία	8920
57	Εκτός κατάταξης	Ωκεανίς - Ιδρυματικό Αποθετήριο της Παν/πολης 2 Πανεπιστημίου Δυτικής Αττικής	
58	Εκτός κατάταξης	Ωνάσειος Ψηφιακή Βιβλιοθήκη	
Διεθνή ψηφιακά αποθετήρια			
A/A	Σειρά κατάταξης του αποθετηρίου στο Ranking Web of Repositories (9th ed. Apr. 2020)	Τίτλος αποθετηρίου	Google Scholar records
59	842	Academic Repository Poltava University of Economics and Trade	4060

60	327	Academic Works Graduate Center CUNY	13100
61	148	Açık Arşiv Marmara Üniversitesi	25000
62	299	ADDI Repositorio Institucional Universidad del País Vasco	14100
63	313	Adelaide Research & Scholarship University of Adelaide	13600
64	21	Airlangga University Repository	74000
65	726	AKU Institutional Repository Aga Khan University	4960
66	2773	Aligarh Muslim University Social Sciences Cybrary	0
67	1015	American University in Cairo Digital Archive and Research Repository	2910
68	Εκτός κατάταξης	Angelo State University Digital Repository	
69	203	Archive of European Integration	19500
70	1654	Athabasca University Auspace	910
71	1677	Aurora Health Care Digital Repository	873
72	156	Australian National University Open Research	24100
73	273	Ball State University Institutional Repository Cardinal Scholar	15200
74	1638	Baskent University Institutional Repository	938
75	Εκτός κατάταξης	Birmingham City University Open Access Repository (BCU)	Birmingham City University Open Access Repository
76	11	Belarusian State University Digital Library	134000
77	381	Belarusian State University of Culture and Arts Repository	11200

78	Εκτός κατάταξης	Bennington College Digital Repository	
79	1900	Berman Jewish Policy Archive Stanford University	567
80	458	Bern Open Repository and Information System BORIS	9040
81	2740	Biblioteca Digital de la Universidad Católica Argentina	3
82	2773	Biblioteca Digital do Desenvolvimento	0
83	648	Biblioteca Digital Zamorano	5810
84	177	Biblos-e Archivo Universidad Autónoma de Madrid	22000
85	1691	Bowdoin Digital Commons Bowdoin College Research	850
86	739	Bridgewater State University Virtual Commons	4900
87	758	Brock University Digital Repository	4740
88	973	Bryant University DigitalCommons	3130
89	368	Butler University Digital Commons	11600
90	604	Cadmus European University Institute Institutional Repository	6440
91	299	Cal Poly DigitalCommons	14100
92	1687	California State University Chico Digital Repository	862
93	876	California State University Sacramento Digital Repository	3790
94	480	CalTech Electronic Theses and Dissertations	8610
95	1711	Cambridge University Engineering Department	816

		Publications Database	
96	240	Cambridge University Institutional Repository	16900
97	913	Cardiff Metropolitan University's Repository	3510
98	221	Carolina Digital Repository University of North Carolina at Chapel Hill	18100
99	1703	Central Electrochemical Research Institute Repository	831
100	1212	Central Food Technological Research Institute Institutional Repository	2040
101	411	Central Marine Fisheries Research Institute Institutional Repository	10200
102	12	CERN Document Server	128000
103	1127	Chicago Unbound University of Chicago Law School Research	2370
104	1191	Cogprints Cognitive Sciences ePrint Archive	2120
105	535	Colecciones Digitales Uniminuto Corporacion Universitaria Minuto de Dios	7450
106	278	Columbia University Academic Commons	14900
107	443	Core Scholar Wright State University Research	9430
108	1897	Creative Matter Skidmore College	571
109	2773	Cronfa Swansea University	0
110	1470	CrossAsia Repository	1310
111	483	CU Scholar University of Colorado Boulder	8530

112	1256	Daffodil International University Repository	1870
113	268	DalSpace Dalhousie University Institutional Repository	15300
114	2720	Danish Research Centre for Magnetic Resonance Repository	9
115	858	DataSpace at Princeton University	3900
116	86	Delft University of Technology Repository	35100
117	1712	Depositum Université du Québec Abitibi Témiscamingue et Cégep	813
118	691	Dépôt institutionnel Université du Québec à Trois-Rivières	5330
119	1643	Dickinson Scholar Dickinson College Research	928
120	148	DIGIBUG Repositorio Institucional Universidad de Granada	25000
121	2106	Digital Commons CIIS California Institute of Integral Studies	368
122	284	Digital Commons Georgia Southern University	14600
123	1246	Digital Commons Kent State University Libraries	1910
124	1651	Digital Commons OWU Ohio Wesleyan University Research	914
125	1880	Digital Commons Rockefeller University	591
126	1871	Digital Commons Wofford Wofford College Research	600
127	1116	Digital Education Resource Archive DERA	2400
128	851	Digital Howard Howard	3940

		University Research	
129	145	Digital Kenyon Kenyon College Research	25400
130	1870	Digital Library of Nara Institute of Science and Technology	602
131	74	Digital Library Tomsk State University	38400
132	141	Digital Library University of Maribor DKUM	25700
133	34	Digital Repository Iowa State University	60300
134	542	DROPS Dagstuhl Research Online Publication Server Schloss Dagstuhl Leibniz-Zentrum für Informatik	7310
135	385	DUGiDocs Universitat de Girona	11100
136	939	Durham Research Online University of Durham	3350
137	311	E LIS ePrints in Library and Information Science	13700
138	995	École Polytechnique de Montréal PolyPublie	3030
139	162	École Polytechnique Federale de Lausanne Infoscience	23500
140	120	Education and Research Archive University of Alberta	28200
141	833	EdUHK Research Repository	4160
142	2773	Electronic Environmental Resources Library eERL	0
143	284	Electronic Library of Belarusian State Technological University	14600
144	1662	Electronic Library Universität Osnabrück	896
145	356	Electronic Publication Alfred	12000

		Wegener Institute for Polar and Marine Research	
146	1634	Elektronische Hochschulschriften Universität Trier	950
147	411	eLibrary National Mining University of Ukraine	10200
148	1615	Epoka University Repository	988
149	565	Epsilon Open Archive Swedish University of Agricultural Sciences	6940
150	340	Escuela Politécnica Nacional Repositorio Digital	12600
151	282	Etheses of Maulana Malik Ibrahim State Islamic University	14700
152	1656	Euskal Memoria Digitala	905
153	597	Exeter University Open Research Exeter ORE	6530
154	1693	Faculty of Food Technology in Osijek Repository	847
155	563	Florida International University DigitalCommons	7000
156	394	Florida State University DigiNole Commons Repository	10800
157	1569	Forsvarets høgskole FHS Brage	1080
158	1889	Fred Hutchinson Cancer Research Center Authors	578
159	416	Funes repositorio Digital de Documentos en Educación Matemática	10100
160	113	Georgia Tech's Institutional Repository	29400
161	1645	Gifu Pharmaceutical University Research Information Repository	927
162	303	Goce Delcev University	14000

		Academic Repository	
163	107	Goteborg University Open Archive	30600
164	348	Grand Valley State University ScholarWorks	12200
165	189	Gredos Repositorio Documental Universidad de Salamanca	20900
166	503	Hacettepe University Open Archive	8090
167	1889	HAL Clermont Université Repository	578
168	1835	HAL École Nationale Supérieure des Mines Saint Etienne	646
169	1065	HAL Institut National de Physique Nucleaire et de Physique des Particules	2670
170	129	HAL Institut National de Recherche en Informatique et en Automatique Archive Ouverte	27600
171	1569	HAL Université de Bourgogne	1080
172	540	HAL Université Toulouse III Paul Sabatier	7330
173	438	Harvard University Digital Access to Scholarship	9600
174	385	Helvia Repositorio Institucional Universidad de Córdoba	11100
175	714	Høgskulen på Vestlandet HVL Open	5120
176	1921	Hollins Digital Commons Hollins University Research	546
177	222	ELAKPI, the Electronic Archive of Igor Sikorsky Kyiv Polytechnic Institute	17900
178	315	Humboldt Universitat zu Berlin	13500

		Publikationsserver	
179	731	Ilia State University Repository	4940
180	981	Indian Institute of Science Bangalore Electronic Theses and Dissertations	3110
181	727	Indian Institute of Technology Delhi Institutional Repository	4950
182	584	Indiana University Purdue University Indianapolis Digital Archive	6730
183	174	Indiana University Scholarworks	22500
184	708	InK Institutional Knowledge Singapore Management University	5160
185	443	Institutional Repository der Universität Konstanz	9430
186	1275	Institutional Repository ISI Surakarta	1800
187	1630	International Research Center for Japanese Studies Repository	960
188	159	IRIS Institutional Research Information System Università degli Studi di Ferrara	23700
189	601	I-Shou University Institutional Repository	6490
190	1915	ISRZ Repository Institute for Social Research at Zagreb	550
191	280	ITU Academic Open Archive Istanbul Technical University	14800
192	600	Johns Hopkins University Repository JScholarship	6500
193	950	Jomo Kenyatta University of Agriculture and Technology Digital Repository	3290

194	171	Kansas State University Research Exchange	22600
195	1175	King Fahd University of Petroleum and Minerals Institutional Repository	2170
196	893	Kingston University London Institutional Repository	3670
197	1070	Kochi University Digital Repository for Academic Resources	2620
198	389	Kwansei Gakuin University Repository	11000
199	9	Kyoto University Research Information Repository	140000
200	1637	Landmark University Repository	939
201	1706	Landspítali University Hospital Research Archive	826
202	1198	Leeds Beckett University Repository	2090
203	89	Leiden University Digital Repository	34800
204	870	LSE Theses Online London School of Economics and Political Science	3810
205	268	Ludwig Maximilians Universitat Munchen Dissertationen	15300
206	20	Lund University Publications	75800
207	992	Macquarie University ResearchOnline	3040
208	236	MacSphere McMaster University Institutional Repository	17100
209	656	Mashhad University of Medical Sciences Repository	5720
210	362	Massey University Research	11800

		Online	
211	737	Maxwell Pontificia Universidade Católica do Rio de Janeiro	4910
212	1681	Memoria Repósitorio Institucional Instituto Federal do Rio Grande do Norte	872
213	434	Memorial University Research Repository	9730
214	29	MIT Institutional Repository	63900
215	878	Morehead State University ScholarWorks	3780
216	122	Mountain Scholar Digital Collections of Colorado and Wyoming	28100
217	135	Munich Personal Repec Archive	26600
218	335	Murdoch University Murdoch Research Repository	12800
219	1370	Musi Charitas Catholic University Repository	1590
220	310	National Agricultural Library Digital Repository	13800
221	1661	National Centre for Research Methods Repository	897
222	26	National Chiao Tung University Institutional Repository	67200
223	1735	National Louis University Research	785
224	179	National Taiwan Normal University Institutional Repository	21600
225	1652	Netherlands Aerospace Centre Reports Repository	913
226	1129	New Bulgarian University Scholar	2360

		Electronic Repository	
227	919	Newcastle University eTheses	3480
228	1009	Nicolaus Copernicus University Repository	2950
229	1677	Nihon University Repository	873
230	1632	Norges Bank Vitenarkiv	954
231	1619	Norges idrettshøgskole Brage NiH	982
232	1860	Norges musikkhøgskole NMH Brage	612
233	1847	Norsk Utenrikspolitisk Institutt NUPI Research Online	625
234	1853	OAK Open Access Archive Novartis	618
235	79	OAK Trust Texas A&M University Repository	37400
236	926	OceanDocs	3430
237	66	Ohio State University Knowledge Bank	41700
238	2773	Open Archive Università degli Studi di Napoli L'Orientale	0
239	Εκτός κατάταξης	OpenBU Repository - IR of the Boston University	
240	1116	OPUS Open Portal to University Scholarship Governors State University	2400
241	1620	Osaka Shoin Women's University Repository	979
242	153	Oxford University Research Archive	24600
243	1648	Özyeğin University eResearch	920
244	1621	Parkland College SPARK Scholarship	978

245	1624	Penn Law Legal Scholarship Repository University of Pennsylvania Law School Research	973
246	284	Pennsylvania State University Electronic Theses and Dissertation Archive	14600
247	317	Petra Christian University Scientific Repository	13400
248	1105	Philosophy of Science Archive	2440
249	1919	Pilot Scholars University of Portland Research	547
250	1685	Plymouth Marine Science Electronic Archive PlyMSEA	863
251	1873	Politihøgskolen PIA	599
252	2773	Polytechnic in Pula Digital Repository	0
253	369	Pontificia Universidad Católica del Perú Repositorio de Tesis	11500
254	2773	Production Scientifique de l'Université M'hamed Bougara Boumerdès	0
255	1732	Publication Server of Furtwangen University of Applied Sciences	787
256	1625	Publication Server Wuppertal Institute for Climate Environment and Energy	965
257	1459	Publications of the German Soil Science Society	1330
258	1894	Publikationsserver der TU Clausthal	574
259	546	Queen's University Qspace	7270

260	1580	Queensland Department of Agriculture and Fisheries archive of scientific and research publications	1050
261	100	Queensland University of Technology Institutional Repository	31600
262	73	REAL Repository Hungarian Academy of Science Library	38500
263	262	REFUBIUM Institutional Repository Freie Universität Berlin	15600
264	1865	Repositori Institucional de la Uvic Universitat de Vic	607
265	165	Repositori Institucional Universitat Jaume I	23400
266	1910	Repositório Aberto Instituto Superior Miguel Torga	552
267	63	Repositorio Académico de la Universidad de Chile	43500
268	535	Repositório Científico do Instituto Politécnico de Lisboa	7450
269	146	Repositorio Digital CEPAL	25100
270	8	Repositório Digital Universidade Federal do Rio Grande do Sul LUME	159000
271	2104	Repositorio Escuela Colombiana de Ingeniería	369
272	252	Repositorio Institucional de la Universidad de El Salvador	16000
273	345	Repositorio Institucional de la Universidad de La Laguna	12500
274	1617	Repositorio Institucional Fundación Nacional para el	985

		Desarrollo	
275	1550	Repositorio Institucional UCA Universidad Centroamericana José Simeón Cañas	1110
276	27	Repositório Institucional UNESP Universidade Estadual Paulista Júlio de Mesquita Filho	64600
277	284	Repositorio Institucional Universidad Autónoma del Estado de México	14600
278	1131	Repositorio Institucional Universidad de Medellín	2350
279	1059	Repositorio Institucional Universidad Francisco Gavidia	2710
280	255	Repositorio Institucional Universidad Santo Tomás	15900
281	1163	Repositorio Minedu Ministerio de Educación Perú	2210
282	1594	Repositorio UPTC Universidad Pedagógica y Tecnológica de Colombia	1030
283	1246	Repository Centrum Wiskunde & Informatica Amsterdam	1910
284	74	Repository of Thesis Tomas Bata University	38400
285	1636	Repository Universitas Islam Riau	948
286	136	Repository Universitas Lampung	26500
287	945	Research Repository University College Dublin	3310
288	1590	RIDAA Repositorio Institucional Digital Universidad Nacional de Quilmes	1040

289	335	RIUMA Repositorio Institucional Universidad de Málaga	12800
290	1576	Rochester Institute of Technology Digital Archive	1060
291	1077	Roger Williams University DigitalCommons	2610
292	317	Rutgers Community Repository Rutgers University	13400
293	1015	Sacred Heart University Research	2910
294	2773	Scholar University of Cincinnati	0
295	1883	Scholarly Commons George Washington University Law School Research	583
296	411	Scholarly Commons University of the Pacific	10200
297	348	Scholarship Western University of Western Ontario	12200
298	454	ScholarWorks UARK University of Arkansas	9120
299	2773	Scottish Institute for Research in Economics Repository	0
300	46	Seoul National University's Institutional Repository S-Space	54400
301	291	Serveur Académique Lausannois	14500
302	1857	SerWisS Server für wissenschaftliche Schriften der FH Hannover	615
303	704	Sheffield Hallam University Research Archive	5200
304	111	Siberian Federal University Digital Repository	30100
305	348	Smithsonian Institution Digital	12200

		Repository	
306	862	South Carolina State Library Digital Collections	3880
307	425	South Dakota State University Open PRAIRIE Repository	9850
308	230	Southern Illinois University Carbondale Repository OpenSIUC	17200
309	1182	St Catherine University Research SOPHIA	2160
310	1101	St. Mary's University Institutional Repository	2470
311	1914	STAIN Parepare Repository	551
312	203	STARS University of Central Florida	19500
313	219	Stellenbosch University Scholar Repository	18200
314	190	Sudan University for Science and Technology Institutional Repository	20700
315	340	Swedish Technical University of Agricultural Sciences Epsilon Undergraduate Theses	12600
316	201	Swedish University of Agricultural Sciences Epsilon Open Archive	19700
317	1211	Swinburne University of Technology Research Bank	2050
318	122	Technical University of Denmark Online Research Database in Technology	28100
319	721	Technical University of Moldova Institutional Repository	4990

320	103	Texas ScholarWorks University of Texas at Austin	31300
321	617	The Lincoln Repository University of Lincoln UK	6300
322	1580	The Scholars Repository LLU Loma Linda University Research	1050
323	356	TigerPrints Clemson University Research	12000
324	762	Trinity College Digital Repository Trinity College Research	4690
325	916	Trinity College University of Dublin Research Archive	3500
326	143	UCL Discovery University College London Institutional Repository	25600
327	369	Universidad Nacional de Colombia Repositorio Institucional	11500
328	23	Università degli Studi di Milano Institutional Archive of Research	70300
329	249	Università di Milano Bicocca Open Archive	16200
330	104	Università di Pisa Electronic Theses	31200
331	1835	Università di Pisa Institutional Repository	646
332	618	Universitas Terbuka Repository	6290
333	59	Universitat Autònoma de Barcelona Dipòsit Digital de Documents	45800
334	1187	Universitat Kassel Institutional Repository	2140
335	717	Universität Leipzig Dokumentenserver Fakultät für	5100

		Mathematik und Informatik	
336	109	Universität Wien Diplomarbeiten und Dissertationen	30300
337	2720	Universität Wien Phaidra	9
338	1136	Universiti Kuala Lumpur Institutional Repository	2320
339	337	Universiti Putra Malaysia Institutional Repository	12700
340	2773	University College Dublin Digital Library	0
341	323	University of Auckland Researchspace	13200
342	1877	University of Benghazi Digital Repository	592
343	1727	University of Birmingham e- Papers	794
344	527	University of Birmingham eTheses	7630
345	2773	University of Biskra Repository	0
346	13	University of California eScholarship Repository	111000
347	317	University of Canterbury Research Repository	13400
348	1910	University of Dubrovnik Digital Repository	552
349	402	University of Edinburgh Research Archive	10600
350	540	University of Ghana Digital Collection UGSpace	7330
351	78	University of Ghent Institutional Archive	37500
352	273	University of Glasgow Theses	15200
353	635	University of Hawaii at Manoa	6030

		eVols	
354	151	University of Hawai'i at Manoa ScholarSpace	24700
355	2773	University of Helsinki Ethesis	0
356	171	University of Helsinki HELDA	22600
357	155	University of Iowa Research Online	24500
358	257	University of Johannesburg Research Output	15800
359	2773	University of Jordan Repository	0
360	196	University of Kentucky Uknowledge	20100
361	69	University of Latvia E-resource Repository	40200
362	2773	University of Leicester Research Archive	0
363	1138	University of Limerick Institutional Repository	2310
364	291	University of Lodz Repository / Repozytorium Uniwersytetu łódzkiego	14500
365	480	University of Malaya Students' Repository	8610
366	182	University of Manitoba Mspace	21400
367	215	University of Maryland Digital Repository DRUM	18400
368	1885	University of Miami School of Law Institutional Repository	582
369	82	University of Minnesota Digital Conservancy	36600
370	193	University of Missouri Repository Mospace	20500
371	57	University of Nebraska Lincoln	47200

		DigitalCommons	
372	321	University of Nevada Las Vegas Repository	13300
373	332	University of New Hampshire Scholars' Repository	12900
374	769	University of New Orleans ScholarWorks	4650
375	1580	University of North Alabama Institutional Repository	1050
376	2773	University of North Texas Health Science Center Scholarly Repository	0
377	1858	University of Northern British Columbia Digital Institutional Repository	613
378	398	University of Oregon Scholars' Bank	10700
379	389	University of Pittsburgh d- Scholarship	11000
380	124	University of Pretoria Institutional Repository	28000
381	331	University of Saint Gallen Forschungsplattform Alexandria	13000
382	240	University of South Florida Tampa Scholar Commons	16900
383	455	University of Strathclyde Institutional Repository	9090
384	37	University of Sumatera Utara Repository	58300
385	230	University of Surabaya Repository	17200
386	295	University of Sydney E- Scholarship	14400

387	385	University of Tasmania Institutional Repository	11100
388	2773	University of Tennessee at Chattanooga Digital Collections	0
389	634	University of Texas Arlington Institutional Repository	6060
390	950	University of the Arts London Research Online	3290
391	124	University of Tokyo Repository	28000
392	168	University of Toronto T Space	23000
393	265	University of West Bohemia Digital Library	15400
394	259	University of Wisconsin Digital Repository	15700
395	131	University of Wollongong Research Online	27200
396	1622	University of Zagreb Medical School Repozitorij	975
397	139	University of Zurich Open Repository and Archive	26400
398	572	UR Scholarship Repository University of Richmond Research	6840
399	115	Utah State University Digital Commons	29200
400	17	Utrecht University Repository	85300
401	1569	Vilnius Gediminas Technical University Institutional Repository	1080
402	389	Virginia Henderson Global Nursing e-Repository	11000
403	989	Virginia Tech University Digital Library and Archives	3060
404	930	Warwick Research Archive Portal	3420

405	1842	WestCollections Western Connecticut State University Research	638
406	184	Western Michigan University ScholarWorks	21300
407	1714	Whitworth University Digital Commons	805
408	1594	Yamagata University Repository	1030