



**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

Τίτλος Εργασίας

**Ζητήματα Ψηφιακής Διατήρησης Ηλεκτρονικών Ακαδημαϊκών
Συγγραμμάτων – Η περίπτωση της Συλλογής Κάλλιπος**

Μαρία Τσάμη (ΑΜ: mslam206682012)

Επιβλέπων: Δημήτριος Κουής

Αθήνα, Αύγουστος 2022

Επιτροπή Εξέτασης

1. Δημήτριος Κουής

2. Σαράντος Καπιδάκης

3. Σπύρος Ζερβός

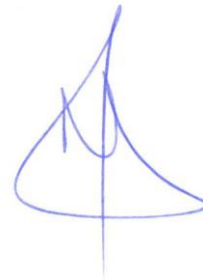
ΔΗΛΩΣΗ ΣΥΓΓΡΑΦΕΑ ΜΕΤΑΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ

Η κάτωθι υπογεγραμμένη Τσάμη Μαρία του Αλεξάνδρου, με αριθμό μητρώου 206682012, φοιτήτρια του Προγράμματος Μεταπτυχιακών Σπουδών «Διαχείριση Πληροφοριών σε Βιβλιοθήκες, Αρχεία και Μουσεία» του Τμήματος Βιβλιοθηκονομίας και Συστημάτων Πληροφόρησης της Σχολής Διοικητικών, Οικονομικών και Κοινωνικών Επιστημών του Πανεπιστημίου Δυτικής Αττικής, δηλώνω ότι:

«Είμαι συγγραφέας αυτής της μεταπτυχιακής εργασίας και ότι κάθε βοήθεια την οποία είχα για την προετοιμασία της, είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης, οι όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών ή λέξεων, είτε ακριβώς είτε παραφρασμένες, αναφέρονται στο σύνολό τους, με πλήρη αναφορά στους συγγραφείς, τον εκδοτικό οίκο ή το περιοδικό, συμπεριλαμβανομένων και των πηγών που ενδεχομένως χρησιμοποιήθηκαν από το διαδίκτυο. Επίσης, βεβαιώνω ότι αυτή η εργασία έχει συγγραφεί από μένα αποκλειστικά και αποτελεί προϊόν πνευματικής ιδιοκτησίας τόσο δικής μου, όσο και του Ιδρύματος.

Παράβαση της ανωτέρω ακαδημαϊκής μου ευθύνης αποτελεί ουσιώδη λόγο για την ανάκληση του πτυχίου μου».

Η Δηλούσα



Ευχαριστίες – Αφιερώσεις

Θα ήθελα να ευχαριστήσω όλους τους κοντινούς μου ανθρώπους για την στήριξη και την εμπύχωση που κατέδειξαν προς το πρόσωπό μου κατά τη διάρκεια των σπουδών μου και ειδικότερα την περίοδο της εκπόνησης της παρούσας εργασίας. Περισσότερο όμως, οφείλω ένα μεγάλο ευχαριστώ στον επιβλέπων καθηγητή της διπλωματικής μου εργασίας, κύριο Κουή, για τις γνώσεις που μου μετέδωσε, την καθοδήγηση που με υπομονή υπέδειξε και την πολύτιμη βοήθειά του για την ολοκλήρωση της μελέτης.

Ημερομηνία

7.9.2022

Συγγραφέας

Μαρία Τσάμη

Περίληψη στα ελληνικά

Η ψηφιακή διατήρηση είναι μία πολύπλοκη διαδικασία, η οποία αποτελείται από μία σειρά διεργασιών, απαραίτητη για την εξασφάλιση της πρόσβασης στο ψηφιακό περιεχόμενο στο διηνεκές. Καθώς κάθε τύπος υλικού, που πρέπει να διατηρηθεί, έχει τις ανάλογες ιδιαιτερότητες, ανακύπτουν διάφορα ζητήματα συντήρησης, με αποτέλεσμα να απαιτούνται διαφορετικές στρατηγικές για κάθε περίπτωση. Ειδικότερα, στην παρούσα μελέτη, εξετάζεται η περίπτωση της ψηφιακής διατήρησης της συλλογής του Αποθετηρίου «Κάλλιπος», με έμφαση τα ηλεκτρονικά ακαδημαϊκά συγγράμματα, τα οποία αποτελούν μία σύνθετη κατηγορία υλικού, καθώς ενσωματώνουν πολύπλοκο περιεχόμενο, όπως είναι τα διαδραστικά στοιχεία, οι εξισώσεις, το πολυμεσικό υλικό κ.ά. Στόχος της εργασίας είναι η μελέτη της στοιχειοθεσίας της συλλογής του «Κάλλιπου» και στη συνέχεια η ανάπτυξη ενός μοντέλου κοστολόγησης για την υλοποίηση της ψηφιακής συντήρησης του υλικού. Τα αποτελέσματα έδειξαν ότι αυτό που κοστίζει και απαιτεί περισσότερο χρόνο επεξεργασίας είναι η μορφή του αρχείου και το πλήθος των στοιχείων που περιέχει. Οι στρατηγικές διατήρησης βασίζονται σε αυτά τα στοιχεία για τον σχεδιασμό ενός μοντέλου συντήρησης του υλικού, καθώς υπάρχουν διάφορα μορφότυπα μέσα στη συλλογή και διαφορετικά επίπεδα δυσκολίας για συγκεκριμένα είδη ηλεκτρονικών ακαδημαϊκών συγγραμμάτων, όπως είναι τα διαδραστικά βιβλία. Τα πιο σημαντικά βήματα για τον σχεδιασμό μιας στρατηγικής διατήρησης είναι η ανάλυση των αναγκών διατήρησης του οργανισμού, οι μορφές αρχείων που υπάρχουν, τα τρέχοντα πρότυπα μεταδεδομένων που χρησιμοποιούνται, τα εργαλεία που επιλέγονται για τη δημιουργία των εγχειριδίων και οι κίνδυνοι που μπορεί να προκύψουν σχετικά με τη πρόσβαση.

Λέξεις Κλειδιά: Ψηφιακή διατήρηση, Στρατηγικές Διατήρησης, Ηλεκτρονικά ακαδημαϊκά συγγράμματα, Αποθετήριο, Ηλεκτρονικές εκδόσεις, Μορφότυπα

Περίληψη στα αγγλικά

Digital preservation is a complex process, consisting of a series of processes, necessary to ensure access to digital content in perpetuity. As each type of material to be preserved has its own particularities, different issues arise, requiring different strategies for each case. In particular, in the present study, the case of the digital preservation of the collection of the "Kallipos" Repository is examined, with an emphasis on electronic academic writings, which are a complex category of material, as they incorporate complex content, such as interactive elements, equations, multimedia material etc. The aim of the thesis is the study of the typesetting of the "Kallipos" collection and then the development of a costing model for the implementation of the digital preservation of the material. The results showed that what costs and requires more processing time is the file format and the number of elements it contains. Preservation strategies rely on these elements to design a material preservation model, as there are different formats within the collection and different levels of difficulty for specific types of e-academic books, such as interactive books. The most important steps in planning a preservation strategy are the analysis of the organization's retention needs, the existing file formats, the current metadata standards in use, the tools chosen to create the books, and the risks that may arise regarding the access.

Keywords: Digital preservation, Preservation strategies, E-textbooks, Repository, Electronic publications, File formats

Πίνακας περιεχομένων

ΕΠΙΤΡΟΠΗ ΕΞΕΤΑΣΗΣ	II
ΔΗΛΩΣΗ ΣΥΓΓΡΑΦΕΑ ΜΕΤΑΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ	III
ΕΥΧΑΡΙΣΤΙΕΣ – ΑΦΙΕΡΩΣΕΙΣ	IV
ΠΕΡΙΛΗΨΗ ΣΤΑ ΕΛΛΗΝΙΚΑ	V
ΠΕΡΙΛΗΨΗ ΣΤΑ ΑΓΓΛΙΚΑ	VI
ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ	VII
ΠΙΝΑΚΑΣ ΣΧΗΜΑΤΩΝ	X
ΠΙΝΑΚΑΣ ΠΙΝΑΚΩΝ	XII
ΚΕΦΑΛΑΙΟ 1. ΕΙΣΑΓΩΓΗ	13
1.1 ΠΛΑΙΣΙΟ, ΣΚΟΠΟΣ ΚΑΙ ΣΤΟΧΟΙ ΤΗΣ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ	13
1.2 ΕΡΕΥΝΗΤΙΚΕΣ ΥΠΟΘΕΣΕΙΣ	13
1.3 ΜΕΘΟΔΟΛΟΓΙΑ	14
1.4 ΠΕΡΙΟΡΙΣΜΟΙ	15
1.5 ΟΡΙΣΜΟΙ	16
1.6 ΟΡΓΑΝΩΣΗ ΚΕΦΑΛΑΙΩΝ Η ΔΙΑΡΘΡΩΣΗ ΤΗΣ ΕΡΓΑΣΙΑΣ	16
ΚΕΦΑΛΑΙΟ 2. ΘΕΩΡΗΤΙΚΟ ΜΕΡΟΣ – ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΕΡΕΥΝΑ – ΣΧΕΤΙΚΕΣ ΠΡΟΣΠΑΘΕΙΕΣ	18
2.1 ΗΛΕΚΤΡΟΝΙΚΑ ΑΚΑΔΗΜΑΪΚΑ ΣΥΓΓΡΑΜΜΑΤΑ	18
2.2 ΕΙΔΗ / ΟΡΙΣΜΟΙ ΗΛΕΚΤΡΟΝΙΚΩΝ ΒΙΒΛΙΩΝ ΚΑΙ ΠΡΟΚΛΗΣΕΙΣ ΩΣ ΠΡΟΣ ΤΗ ΔΙΑΤΗΡΗΣΗ ΤΟΥΣ	18
2.2.1 Ηλεκτρονικά βιβλία με ενσωματωμένα στοιχεία και περιεχόμενο από το διαδίκτυο	18
2.2.2 Σχολικά βιβλία	19
2.2.3 Βιβλία αναφοράς	20
2.2.4 Ψηφιοποιημένα βιβλία	20
2.3 ΕΚΔΟΣΕΙΣ ΨΗΦΙΑΚΩΝ ΒΙΒΛΙΟΘΗΚΩΝ	21
2.3.1 Αυτοέκδοση	21
2.4 ΤΕΧΝΟΛΟΓΙΑ ΚΑΙ ΠΡΟΤΥΠΑ	22
2.5 ΜΟΡΦΟΤΥΠΑ ΓΙΑ ΤΗΝ ΔΗΜΟΣΙΕΥΣΗ ΗΛΕΚΤΡΟΝΙΚΩΝ ΒΙΒΛΙΩΝ	22
2.5.1 PDF	22
2.5.2 HTML	23

2.5.3	EPUB	24
2.5.4	MOBI	24
2.6	ΕΡΓΑΛΕΙΑ ΣΤΟΙΧΕΙΟΘΕΣΙΑΣ ΚΕΙΜΕΝΟΥ ΓΙΑ ΕΝΤΥΠΕΣ ΚΑΙ ΨΗΦΙΑΚΕΣ ΕΚΔΟΣΕΙΣ	26
2.6.1	Adobe In Design	26
2.6.2	LATEX	27
2.6.3	DocBook	28
2.6.4	Pressbooks	29
2.6.5	OpenStax	29
2.7	ΣΤΡΑΤΗΓΙΚΕΣ ΔΙΑΤΗΡΗΣΗΣ	30
2.7.1	Προσομοίωση	30
2.7.2	Αντίστροφη μηχανική	32
2.7.3	Μετανάστευση	32
2.8	OAIS REFERENCE MODEL	33
2.8.1	Το περιβάλλον του OAIS	34
2.8.2	Πακέτο Πληροφοριών (Information Package)	35
2.8.3	Το λειτουργικό μοντέλο (Functional model)	37
2.9	ΠΡΟΤΥΠΑ ΔΙΑΤΗΡΗΣΗΣ ΜΕΤΑΔΕΔΟΜΕΝΩΝ	39
2.9.1	PREMIS	39
2.9.2	METS	41
2.10	ΙΔΡΥΜΑΤΙΚΑ ΑΠΟΘΕΤΗΡΙΑ ΚΑΙ ΨΗΦΙΑΚΗ ΔΙΑΤΗΡΗΣΗ	42
2.11	Η ΠΕΡΙΠΤΩΣΗ ΤΗΣ ΣΥΛΛΟΓΗΣ ΚΑΛΛΙΠΟΣ	43
ΚΕΦΑΛΑΙΟ 3. ΜΕΘΟΔΟΛΟΓΙΑ Η/ΚΑΙ ΣΧΕΔΙΟ ΕΡΓΑΣΙΩΝ– ΥΛΟΠΟΙΗΣΗ – ΕΦΑΡΜΟΓΗ		45
3.1	ΔΙΑΔΙΚΑΣΙΑ ΣΥΓΓΡΑΦΗΣ ΚΑΙ ΑΝΑΝΕΩΣΗΣ ΣΥΓΓΡΑΜΜΑΤΩΝ ΣΥΛΛΟΓΗΣ ΚΑΛΛΙΠΟΥ	45
3.2	ΜΕΘΟΔΟΛΟΓΙΑ ΠΡΟΣΔΙΟΡΙΣΜΟΥ ΑΝΘΡΩΠΟΠΡΟΣΠΑΘΕΙΑΣ - ΚΟΣΤΟΥΣ	49
3.3	ΥΠΟΛΟΓΙΣΜΟΣ ΑΝΘΡΩΠΟΠΡΟΣΠΑΘΕΙΑΣ ΓΙΑ ΚΑΘΕ ΚΑΤΗΓΟΡΙΑ ΥΛΙΚΟΥ	53
3.3.1	Χρόνος εργασίας για εικόνες / σχήματα	53
3.3.2	Χρόνος εργασίας για πίνακες	54
3.3.3	Χρόνος εργασίας για εξισώσεις	54
3.3.4	Χρόνος εργασίας για βιβλιογραφικές αναφορές	55
3.3.5	Ώρες και ημέρες	56
3.4	ΜΗΤΡΑ ΚΟΣΤΟΥΣ ΚΑΙ ΠΑΡΑΔΕΙΓΜΑ	57
3.4.1	Βιβλίο «Πληροφοριακά Συστήματα Διοίκησης»	59
ΚΕΦΑΛΑΙΟ 4. ΑΠΟΤΕΛΕΣΜΑΤΑ – ΕΥΡΗΜΑΤΑ / ΕΠΙΤΕΥΓΜΑΤΑ		61
4.1	ΑΝΑΛΥΤΙΚΗ ΠΑΡΟΥΣΙΑΣΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ	61
4.2	ΚΥΡΙΟΤΕΡΑ ΕΥΡΗΜΑΤΑ/ ΑΠΟΤΕΛΕΣΜΑΤΑ	67
4.3	ΕΠΙΠΛΕΟΝ ΚΕΝΤΡΑ ΕΠΙΒΑΡΥΝΣΗΣ	74

ΚΕΦΑΛΑΙΟ 5. ΣΥΖΗΤΗΣΗ – ΣΥΜΠΕΡΑΣΜΑΤΑ – ΜΕΛΛΟΝΤΙΚΕΣ ΕΠΕΚΤΑΣΕΙΣ.....	76
5.1 ΑΝΑΚΕΦΑΛΑΙΩΣΗ.....	76
5.2 ΣΥΖΗΤΗΣΗ / ΣΥΜΠΕΡΑΣΜΑΤΑ	77
5.3 ΑΞΙΟΠΟΙΗΣΗ / ΠΡΑΚΤΙΚΕΣ ΠΡΟΕΚΤΑΣΕΙΣ ΤΗΣ ΕΡΕΥΝΑΣ	78
5.4 ΜΕΛΛΟΝΤΙΚΕΣ ΕΠΕΚΤΑΣΕΙΣ / ΠΡΑΚΤΙΚΕΣ ΠΡΟΕΚΤΑΣΕΙΣ ΤΗΣ ΈΡΕΥΝΑΣ.....	79
ΞΕΝΗ ΒΙΒΛΙΟΓΡΑΦΙΑ.....	81
ΠΡΟΣΘΕΤΗ ΒΙΒΛΙΟΓΡΑΦΙΑ	84

Πίνακας Σχημάτων

Εικόνα 1: Περιβάλλον του OAIS.....	34
Εικόνα 2: Πακέτο πληροφοριών του OAIS	36
Εικόνα 3: Το λειτουργικό μοντέλο του OAIS (Functional model)	37
Εικόνα 4: Πρότυπο διατήρησης δεδομένων PREMIS	40
Εικόνα 5: Ροή διαδικασίας συγγραφής ενός συγγράμματος	45
Εικόνα 6. Βασικές ροές εργασίας για την παραγωγή ενός συγγράμματος στον Κάλλιπο.	48
Εικόνα 7. Ανανέωση συγγράμματος.....	49
Εικόνα 8. Μεθοδολογία για τη δημιουργία μοντέλου προσδιορισμού ανθρωποπροσπάθειας – κόστους – Εφαρμογή μοντέλου	50
Εικόνα 9: Μήτρα κόστους.....	57
Εικόνα 10: Παράδειγμα βιβλίου από την συλλογή του Κάλλιπου	59
Εικόνα 11: Σύνολο βιβλίων ανά μορφή του τελικού αρχείου.....	62
Εικόνα 12: Κατηγοριοποίηση βιβλίων σε σχέση με τον αριθμό εικόνων/ σχημάτων	62
Εικόνα 13: Κατηγοριοποίηση βιβλίων σε σχέση με τον αριθμό πινάκων	63
Εικόνα 14: Κατηγοριοποίηση σε σχέση με τον αριθμό εξισώσεων.....	64
Εικόνα 15: Κατηγοριοποίηση βιβλίων σε σχέση με τον αριθμό βιβλιογραφικών αναφορών.....	64
Εικόνα 16: Ποσότητα βιβλίων που περιέχουν ή όχι βίντεο.....	65
Εικόνα 17: Ποσότητα βιβλίων που περιέχουν ή όχι ήχο.....	65
Εικόνα 18: Ποσότητα βιβλίων που περιέχουν ή όχι διαδραστικά στοιχεία	66
Εικόνα 19: Κατηγοριοποίηση βιβλίων σε σχέση με τον αριθμό σελίδων που περιέχουν	66
Εικόνα 20: Κατηγοριοποίηση βιβλίων σε σχέση με τον αριθμό των ημερών επεξεργασίας.....	68
Εικόνα 21: Συνολικός χρόνος επεξεργασίας για τα αρχεία σε μορφή LATEX.....	69
Εικόνα 22: Συνολικός χρόνος επεξεργασίας για τα αρχεία σε μορφή Word.....	70
Εικόνα 23: Συνολικός χρόνος επεξεργασίας για τα αρχεία σε μορφή Adobe In Design με Word	71
Εικόνα 24: Συνολικός χρόνος επεξεργασίας για τα αρχεία σε μορφή Adobe In Design χωρίς Word	72
Εικόνα 25: Συνολικός χρόνος επεξεργασίας για τα αρχεία σε άλλη μορφή.....	73
Εικόνα 26: Σύνολο ωρών επεξεργασίας για κάθε μορφότυπο.....	74

Εικόνα 27: Σύνολο ημερών επεξεργασίας για κάθε μορφότυπο 74

Πίνακας Πινάκων

Πίνακας 1: Υπολογισμός κατά προσέγγιση της συχνότητας εμφάνισης των στοιχείων	52
Πίνακας 2: Εκτίμηση χρόνου σε λεπτά.....	53
Πίνακας 3: Χρόνος απαιτούμενης προσπάθειας / ανά κατηγορία υλικού.....	53
Πίνακας 4: Μέση τιμή ανά κατηγορία συχνότητας στοιχείων.....	53
Πίνακας 5: Χρόνος εργασίας που απαιτείται για εικόνες / σχήματα	53
Πίνακας 6: Χρόνος εργασίας που απαιτείται για τους πίνακες	54
Πίνακας 7: Χρόνος εργασίας που απαιτείται για τις εξισώσεις.....	54
Πίνακας 8: Χρόνος εργασίας που απαιτείται για τις βιβλιογραφικές αναφορές	55
Πίνακας 9: Χρονική επιβάρυνση λοιπών παραγόντων.....	56
Πίνακας 10: Ομαδοποίηση στοιχείων για την τροφοδότηση του μοντέλου.....	56
Πίνακας 11: Ανάκτηση τιμών από την μήτρα κόστους.....	59
Πίνακας 12: Ανάλυση υπολογισμού εργατωρών	59
Πίνακας 13: Σύνολο ημερών επεξεργασίας των βιβλίων	67
Πίνακας 14: Συνολικό κόστος μελέτης.....	68

Κεφάλαιο 1. Εισαγωγή

1.1 Πλαίσιο, σκοπός και στόχοι της διπλωματικής εργασίας

Η παρούσα διπλωματική εργασία ερευνά το πεδίο της ψηφιακής διατήρησης των ηλεκτρονικών βιβλίων, και ειδικότερα των ηλεκτρονικών ακαδημαϊκών συγγραμμάτων, με σκοπό τη μελέτη της στοιχειοθεσίας των βιβλίων για τον υπολογισμό του κόστους που απαιτείται για τη διατήρησή τους.

Οι βασικοί στόχοι της διπλωματικής εργασίας είναι, αρχικά, η κατανόηση των αναγκών και των απαιτήσεων που επιβάλλει η έννοια της ψηφιακής διατήρησης σε σχέση με τα σύνθετα ψηφιακά αντικείμενα, όπως είναι τα ηλεκτρονικά επιστημονικά βιβλία. Παράλληλα, στόχος είναι η καταγραφή των βασικών σημείων που πρέπει να διαθέτει μία στρατηγική μακροχρόνιας ψηφιακής διατήρησης μιας συλλογής ηλεκτρονικών ακαδημαϊκών συγγραμμάτων και στη συνέχεια, η εφαρμογή των αποτελεσμάτων της υπό εκπόνηση στρατηγικής σε μία πραγματική συλλογή ηλεκτρονικών ακαδημαϊκών συγγραμμάτων, όπως είναι αυτή του Κάλλιπου (περίπου 520 βιβλία), με στόχο τη δημιουργία μιας οικονομοτεχνικής μελέτης για την ψηφιακή τους διατήρηση.

1.2 Ερευνητικές υποθέσεις

Ο σκοπός της παρούσας έρευνας ήταν να διερευνηθεί η μέθοδος για τον υπολογισμό του κόστους επεξεργασίας σχετικά με την ανανέωση της συλλογής της πρώτης φάσης του Αποθετηρίου «Κάλλιπος». Η μελέτη επικεντρώνεται στην καταγραφή της στοιχειοθεσίας και των χαρακτηριστικών των ηλεκτρονικών ακαδημαϊκών βιβλίων της συλλογής, με στόχο τον υπολογισμό των εργατωρών που απαιτούνται για την επεξεργασία και κατ' επέκταση το κόστος που δημιουργείται για την υλοποίηση της ψηφιακής διατήρησής τους.

Σε συνάρτηση με τον σκοπό της εργασίας αναδύθηκαν διάφορα ερευνητικά ερωτήματα. Το πρώτο ερώτημα αφορούσε τους βασικούς τύπους και τα εργαλεία ανάπτυξης των ηλεκτρονικών (ψηφιακών) συγγραμμάτων, δηλαδή ποια είδη τύπων και εργαλείων υπάρχουν, με εστίαση στα ακαδημαϊκά συγγράμματα και την πολυπλοκότητα που παρουσιάζουν.

Το δεύτερο ερώτημα που αναλύεται στην παρούσα εργασία είναι το πως εφαρμόζονται οι τεχνικές ψηφιακής διατήρησης στα ψηφιακά συγγράμματα και αν υπάρχουν ειδικά ζητήματα σε σχέση με τα ακαδημαϊκά συγγράμματα. Τα ψηφιακά

συγγράμματα εμπίπτουν στην κατηγορία του ψηφιακού περιεχομένου, οπότε ακολουθούν τις πολιτικές ψηφιακής διατήρησης, όπως είναι η μετανάστευση, προσομοίωση κ.ά., έχουν όμως ορισμένες ιδιαιτερότητες. Η διατήρηση του πρωτογενούς υλικού είναι ένα σύνθετο θέμα και πρέπει να αντιμετωπιστεί για τις περιπτώσεις αναθεώρησης / επικαιροποίησης και επαναχρησιμοποίησης του.

Το τρίτο ερευνητικό ερώτημα ανήκει στη μελέτη της συλλογής Κάλλιπος και στην ανάπτυξη ενός μοντέλου κοστολόγησης για την επικαιροποίησή της. Η συλλογή της πρώτης φάσης του Κάλλιπου περιείχε ορισμένες ιδιαιτερότητες, όπως ελλιπή στοιχεία ή αρχεία, τα οποία δυσχέραναν τον ακριβή υπολογισμό του κόστους, για αυτό η υλοποίηση της αναθεώρησης/ επικαιροποίησης του υλικού έχει πραγματοποιηθεί κατά προσέγγιση.

Από τα ευρήματα που ανέκυψαν, συμπεραίνουμε ότι υπάρχουν διάφορες στρατηγικές διατήρησης, αλλά για την ψηφιακή διατήρηση των συγγραμμάτων δεν ακολουθείται μία συγκεκριμένη πολιτική από όλους τους πληροφοριακούς οργανισμούς. Για την αναθεώρηση / επικαιροποίηση και επαναχρησιμοποίηση του υλικού, το ζητούμενο είναι η διατήρηση του πρωτογενούς υλικού σύνθετων συγγραμμάτων, ενώ παράλληλα προκύπτει και η απαίτηση για την ταυτόχρονη συνύπαρξη πολλαπλών μορφοτύπων δημοσίευσης για το ίδιο βιβλίο.

Στο τελικό στάδιο της εκπόνησης της εργασίας, το μοντέλο κοστολόγησης ανέδειξε σημαντικά αποτελέσματα. Υπολογίστηκε ο μέσος όρος, ο μέγιστος και ο ελάχιστος των ωρών και των ημερών επεξεργασίας που απαιτείται για την υλοποίηση της ανανέωσης του υλικού και από αυτά τα στοιχεία προέκυψε και το κόστος επεξεργασίας της συλλογής του Κάλλιπου για τις τρεις περιπτώσεις.

1.3 Μεθοδολογία

Η μεθοδολογία που ακολουθήθηκε στην παρούσα εργασία ήταν, αρχικά, η βιβλιογραφική ανασκόπηση στα θέματα της ψηφιακής διατήρησης των ηλεκτρονικών βιβλίων και τις πρακτικές που ακολουθούνται παγκοσμίως από τους οργανισμούς διατήρησης. Στη συνέχεια, πραγματοποιήθηκε έρευνα και αναλύθηκαν τα εργαλεία και τα μορφότυπα που χρησιμοποιούνται για τη δημιουργία των ηλεκτρονικών βιβλίων. Αφότου ολοκληρώθηκε η ανασκόπηση της διεθνούς βιβλιογραφίας σχετικά με τις τεχνολογίες και τα πρότυπα που χρησιμοποιούνται για την επεξεργασία των ηλεκτρονικών βιβλίων, ακολούθησε η μελέτη του υλικού της συλλογής ηλεκτρονικών ακαδημαϊκών συγγραμμάτων του αποθετηρίου «Κάλλιπος» και πραγματοποιήθηκε η αναλυτική καταγραφή της σύνθεσής τους. Με βάση την αναλυτική καταγραφή των βιβλίων και το πλήθος των στοιχείων που προέκυψαν από

αυτή, δημιουργήθηκε ένα μοντέλο για την αποτίμηση / εκτίμηση των εργασιών που απαιτούνται για την ψηφιακή διατήρηση του υλικού. Επομένως, το επόμενο βήμα ήταν η εφαρμογή του μοντέλου σε πραγματικά δεδομένα, δηλαδή στα βιβλία της συλλογής του αποθετηρίου του «Κάλλιπου», με έμφαση στον υπολογισμό του κόστους για την υλοποίηση της ψηφιακής τους διατήρησης. Μετά την ολοκλήρωση της εφαρμογής του μοντέλου στα βιβλία της συλλογής του «Κάλλιπου», εξήχθησαν ορισμένα χρήσιμα συμπεράσματα και στατιστικά, σχετικά με το ποια στοιχεία κοστίζουν περισσότερο και απαιτούν περισσότερο χρόνο επεξεργασίας. Παράλληλα, υπολογίστηκε ο συνολικός χρόνος και το συνολικό κόστος της ανθρωποπροσπάθειας για την υλοποίηση της ψηφιακής διατήρησης της συλλογής.

1.4 Περιορισμοί

Σύμφωνα με το βασικό σκοπό της εργασίας και με βάση τους στόχους της, προέκυψαν περιορισμοί, οι οποίοι αποτέλεσαν και διαμόρφωσαν το πλαίσιο της μελέτης.

Αρχικά, η μελέτη περιορίστηκε στο πεδίο του Αποθετηρίου «Κάλλιπος» και ειδικότερα στα βιβλία που εμπεριέχονταν στη συλλογή της πρώτης φάσης, δηλαδή τα ακαδημαϊκά συγγράμματα που εισήχθησαν κατά την έναρξη της λειτουργίας του «Κάλλιπου». Η συγκεκριμένη συλλογή επιλέχθηκε ως κέντρο μελέτης και έρευνας στη παρούσα εργασία, λόγω της χρονολογίας κατά την οποία δημιουργήθηκαν και εισήχθησαν στο Αποθετήριο.

Ένας δεύτερος περιορισμός που παρατηρήθηκε και έπαιξε σημαντικό ρόλο στη μελέτη ήταν η επικέντρωση στον τομέα των ηλεκτρονικών ακαδημαϊκών συγγραμμάτων, καθώς το σύνολο της συλλογής απαρτίζεται μόνο από το συγκεκριμένο είδος βιβλίων.

Παράλληλα, εντοπίστηκαν επιπλέον περιορισμοί κατά τη διαδικασία καταγραφής της στοιχειοθεσίας των βιβλίων. Πιο συγκεκριμένα, από ορισμένα συγγράμματα απουσίαζε το πρωτογενές υλικό, δηλαδή είτε δεν υπήρχε το αρχικό κείμενο του συγγραφέα είτε το επιμελημένο κείμενο του γραφίστα, ή υπήρχαν περιορισμένα στοιχεία, όπως για παράδειγμα περιέχονταν στο φάκελο μόνο οι εικόνες του τεκμηρίου. Τα προαναφερθέντα θεωρήθηκαν ελλείμματα, καθώς δυσχέραιναν τον υπολογισμό του κόστους, και γι' αυτό τον λόγο αφαιρέθηκαν από το μοντέλο κοστολόγησης.

Συμπερασματικά, από την σφαιρική κατανόηση των περιορισμών ο αναγνώστης αποκτά πληρέστερη αντίληψη των αποτελεσμάτων και των αντίκτυπων στο συγκεκριμένο ερευνητικό πεδίο.

1.5 Ορισμοί

Ακολουθούν ορισμοί στο κυρίως σώμα της εργασίας (Κεφάλαιο 2). Σκοπός αυτής της επιλογής ήταν η ενσωμάτωση των ορισμών στο αντίστοιχο θεματικό μέρος. Με αυτόν τον τρόπο επιτυγχάνεται μία πληρέστερη εικόνα της θεωρίας του κάθε κεφαλαίου με την αντίστοιχη θεματική ανάλυση. Μέσω της ταυτόχρονης αναφοράς ορισμών και της ανάλυσής τους ο αναγνώστης μπορεί πιο εύκολα να αντιληφθεί τόσο τον σκοπό όσο και τους στόχους της έρευνας.

1.6 Οργάνωση Κεφαλαίων ή Διάρθρωση της Εργασίας

Η διάρθρωση της εργασίας ακολουθεί τη δομή της μεθοδολογίας που αναπτύχθηκε στο κεφάλαιο 1.3. Στο κεφάλαιο 2, αναπτύσσεται ο ορισμός και τα είδη των ηλεκτρονικών βιβλίων και οι προκλήσεις που μπορεί να ανακύψουν σχετικά με την ψηφιακή διατήρηση του κάθε είδους βιβλίου. Αναλύονται τα μορφότυπα των ηλεκτρονικών βιβλίων και οι τεχνολογίες που απαιτούνται για τη δημιουργία, την επεξεργασία, την ανάγνωση και τη δημοσίευσή τους. Παράλληλα, αναφέρονται οι πιο διαδεδομένες στρατηγικές διατήρησης που έχουν εφαρμοστεί και αναλύεται το εννοιολογικό μοντέλο OAI Reference Model, το οποίο θεωρείται ως το καταλληλότερο εργαλείο για την ψηφιακή διατήρηση. Στο τέλος του δεύτερου κεφαλαίου, περιγράφεται η συλλογή του αποθετηρίου «Κάλλιπος», πως ξεκίνησε η Δράση «Ελληνικά Ακαδημαϊκά Ηλεκτρονικά Συγγράμματα και Βοηθήματα», οι εκδόσεις που περιλαμβάνονται στο αποθετήριο, αλλά και τα χαρακτηριστικά της συλλογής της πρώτης φάσης του «Κάλλιπου».

Στο κεφάλαιο 3, αναπτύσσεται η διαδικασία που ακολουθείται για τη συγγραφή ή / και την ανανέωση των συγγραμμάτων της συλλογής του «Κάλλιπου» και στη συνέχεια αφού καταγράφονται όλα τα βιβλία και η σύνθεσή τους, πραγματοποιείται μία διερεύνηση γύρω από τα κέντρα κόστους, δηλαδή ποια στοιχεία επιβαρύνουν περισσότερο την κοστολόγηση της επιμέλειας των βιβλίων. Για αυτόν τον σκοπό, δημιουργείται ένα μοντέλο (μήτρα κόστους), με εκτιμήσεις τριών γραφιστών σχετικά με τον χρόνο που απαιτείται για την επεξεργασία κάθε στοιχείου και ανάλογα με την πυκνότητα που συναντώνται μέσα στο κείμενο.

Από το μοντέλο προκύπτουν, στο κεφάλαιο 4, στατιστικά στοιχεία και αποτελέσματα σχετικά με τον χρόνο που απαιτεί η επεξεργασία κάθε μορφοτύπου, των εικόνων και των σχημάτων, των πινάκων, των εξισώσεων, των βιβλιογραφικών αναφορών, των αρχείων βίντεο και ήχου, και των διαδραστικών στοιχείων. Μέσω των γραφημάτων που αποτυπώνονται σε αυτό το κεφάλαιο, απεικονίζονται συγκεντρωτικά τα βιβλία

ανάλογα με το πλήθος των στοιχείων που περιέχουν και περιγράφεται ο επιμερισμός των ωρών επεξεργασίας ανά μορφότυπο. Σύμφωνα με τα ευρήματα που προέκυψαν, υπολογίστηκε η κατά προσέγγιση αμοιβή του ανθρώπινου δυναμικού που θα συντελέσει τις διεργασίες και έτσι ανέκυψε το συνολικό κόστος της μελέτης.

Στο τελευταίο κεφάλαιο, ανακεφαλαιώνεται η ροή των εργασιών που συντελέστηκαν στην έρευνα και παρουσιάζονται τα κυριότερα ευρήματα της μελέτης. Επίσης, αναλύονται οι δυνατότητες αξιοποίησης της παρούσας εργασίας και προτείνονται μελλοντικές προεκτάσεις για έρευνα.

Κεφάλαιο 2. Θεωρητικό μέρος – Βιβλιογραφική έρευνα – Σχετικές προσπάθειες

Σε αυτό το κεφάλαιο παρουσιάζονται τα είδη των ηλεκτρονικών ακαδημαϊκών συγγραμμάτων και οι τεχνολογίες με τα πρότυπα που χρησιμοποιούνται για τη δημιουργία τους. Γίνεται αναφορά στις στρατηγικές ψηφιακής διατήρησης των ηλεκτρονικών βιβλίων και αναλύεται το εννοιολογικό μοντέλο OAIIS και τα πρότυπα διατήρησης μεταδεδομένων PREMIS και METS. Στο τελευταίο μέρος του κεφαλαίου παρουσιάζεται η συλλογή του Αποθετηρίου «Κάλλιπος».

2.1 Ηλεκτρονικά ακαδημαϊκά συγγράμματα

Με την ανάπτυξη των Τεχνολογιών Πληροφορικής και Επικοινωνιών (ΤΠΕ), η χρήση των ηλεκτρονικών εγχειριδίων (e-textbooks) έχει διευρυνθεί στον χώρο των Πανεπιστημίων, και έχει μειωθεί αισθητά η χρήση των έντυπων συγγραμμάτων. Τα πλεονεκτήματα των ηλεκτρονικών εγχειριδίων είναι πολυδιάστατα σε σχέση με τα αναλογικά βιβλία, διότι ενσωματώνουν πολυμέσα, όπως βίντεο, αρχεία ήχου, κινούμενες εικόνες, ασκήσεις και διαδραστικό υλικό (Lee, H. J., Messom, C., Yau, K. A., 2012). Παράλληλα, παρέχουν προηγμένες λειτουργίες όπως η σελιδοποίηση, η αναζήτηση εντός του κειμένου, η ευρετηρίαση, η αντιγραφή και η επικόλληση, καθώς και δυνατότητα συνεχούς ενημέρωσης του περιεχομένου τους. Η διαδραστικότητα και η χρηστικότητα που προσφέρουν, βοηθούν στη μαθησιακή διαδικασία και εξυπηρετούν στην αναζήτηση πηγών μέσω του διαδικτύου, καθώς στην πλειοψηφία τους τα βιβλία, διατίθενται με καθεστώς ανοιχτής πρόσβασης.

2.2 Είδη / ορισμοί ηλεκτρονικών βιβλίων και προκλήσεις ως προς τη διατήρησή τους

Υπάρχουν διάφορα είδη ηλεκτρονικών βιβλίων, κατηγοριοποιημένα ανάλογα με το περιεχόμενό τους, όπως για παράδειγμα:

2.2.1 Ηλεκτρονικά βιβλία με ενσωματωμένα στοιχεία και περιεχόμενο από το διαδίκτυο

Πολλές μορφές ηλεκτρονικών βιβλίων υποστηρίζουν ενσωματωμένα αρχεία ήχου ή βίντεο, ενσωματωμένες γραμματοσειρές, υπερσυνδέσμους και άλλους τύπους ενσωματωμένου περιεχομένου. Για τη βιωσιμότητα αυτού του ψηφιακού περιεχομένου χρειάζεται, σύμφωνα με τη Βιβλιοθήκη του Κογκρέσου, η «μορφή περιτυλίγματος» (wrapper format)

(Kirchhoff & Morrissey, 2014). Η πλήρης διατήρηση τέτοιων εμπλουτισμένων ηλεκτρονικών βιβλίων συνεπάγεται, την προσπέλαση της δομής των δεδομένων του εγγράφου, τον εντοπισμό και την αναγνώριση ενσωματωμένων αντικειμένων, με σκοπό τον σχεδιασμό κατάλληλων ενεργειών διατήρησης των στοιχείων αυτών, μέσα σε αυτή τη «μορφή περιτυλίγματος» και σε συνάρτηση με τη διατήρηση του «πνευματικού περιεχομένου», δηλαδή του κειμένου (Kirchhoff & Morrissey, 2014). Ταυτόχρονα, πρέπει να εξετάζονται οι περιπτώσεις όπου μπορεί να ισχύουν διαφορετικά δικαιώματα πνευματικής ιδιοκτησίας για τα ενσωματωμένα αντικείμενα, από τα πνευματικά δικαιώματα που σχετίζονται με το κείμενο (Kirchhoff & Morrissey, 2014).

Αποτελεί θέμα πολιτικής του κάθε οργανισμού διατήρησης το αν θα ακολουθήσει όλη την προαναφερθείσα διαδικασία συντήρησης όλων των ενσωματωμένων αρχείων ενός ηλεκτρονικού βιβλίου. Οι οργανισμοί διατήρησης οφείλουν να εξετάσουν προσεκτικά τι είναι τεχνικά εφικτό, τι περιλαμβάνει μια αποδεκτή υποβολή και ποια επίπεδα διατήρησης θα υποστηρίξουν. Για την αποφυγή επιπλέον ζητημάτων διατήρησης, οι δημιουργοί περιεχομένου θα πρέπει να συμμορφώνονται με τις οδηγίες, που περιορίζουν τους τύπους μορφής που μπορούν να ενσωματωθούν σε ένα ηλεκτρονικό βιβλίο, καθώς επίσης θα πρέπει να καθορίσουν εάν το δικαίωμά τους να διατηρήσουν ένα ηλεκτρονικό βιβλίο περιλαμβάνει το δικαίωμα διατήρησης αντικειμένων που είναι ενσωματωμένα σε αυτό (Kirchhoff & Morrissey, 2014).

Παράλληλα, σε περιπτώσεις που τα ηλεκτρονικά συγγράμματα ενδέχεται να περιέχουν υπερσυνδέσμους, οι οποίοι οδηγούν σε στοιχεία εκτός του ίδιου του βιβλίου, ο οργανισμός διατήρησης οφείλει να προσδιορίσει την πολιτική του σχετικά με τη διατήρηση των συνδεδεμένων στοιχείων, συμπεριλαμβανομένων των αποφάσεων που προτίθεται να λάβει, όσον αφορά τη διατήρηση της μονιμότητας αυτών των συνδέσμων, και σε τι ποσοστό αυτού του περιεχομένου έχει τη δικαιοδοσία να αναλάβει τη διατήρησή του (Kirchhoff & Morrissey, 2014).

2.2.2 Σχολικά βιβλία

Τα σχολικά βιβλία αποτελούν μια αναδυόμενη κατηγορία επεκτάσιμων ηλεκτρονικών βιβλίων, που μπορεί να περιέχουν τόσο ενσωματωμένο όσο και συνδεδεμένο περιεχόμενο, καθώς και προσωπικούς σχολιασμούς από κάθε αναγνώστη. Αποτελούν παράδειγμα των δυνητικά ανατρεπτικών επιπτώσεων των ηλεκτρονικών βιβλίων στον ακαδημαϊκό χώρο και στην παιδαγωγική γενικότερα (Kirchhoff & Morrissey, 2014).

Οι πρωτογενώς ψηφιακές εκδόσεις των σχολικών εγχειριδίων υπερτερούν των ψηφιοποιημένων έντυπων βιβλίων. Συχνά περιλαμβάνουν ενότητες που μπορούν να

ενσωματωθούν σε ένα σύστημα διαχείρισης μαθημάτων, ένα μαζικό ανοιχτό διαδικτυακό μάθημα (MOOC) ή άλλα μαθήματα ηλεκτρονικής μάθησης. Υπόκεινται στις προκλήσεις διατήρησης, που προαναφέρθηκαν, για τα ηλεκτρονικά βιβλία με ενσωματωμένα στοιχεία και συνδεδεμένο περιεχόμενο, ενώ παρουσιάζουν πρόσθετες προκλήσεις για τη διατήρησή τους, από τους οργανισμούς διατήρησης, οι οποίοι πρέπει να καθορίσουν σε ποιο βαθμό θα πρέπει να διατηρηθούν οι προσωποποιημένες προσαρμογές για τον αναγνώστη (Kirchhoff & Morrissey, 2014).

2.2.3 Βιβλία αναφοράς

Η μορφή των ηλεκτρονικών βιβλίων αναφοράς ποικίλλει. Υπάρχει η περίπτωση να δημοσιευθούν ως ένα εμπορικό ηλεκτρονικό βιβλίο σε μία συμβατική μορφή. Ωστόσο, συχνά αποτελούν διαδικτυακές βάσεις δεδομένων, με πληροφορίες που ενημερώνονται συνεχώς, όπως για παράδειγμα τα λεξικά, οι εγκυκλοπαίδειες, οι παραπομπές ιατρικής κ.ά. Και αυτό το περιεχόμενο επιβάλλεται να διατηρηθεί, όμως εάν δεν είναι δημοσιευμένο σε μία συμβατική μορφή ηλεκτρονικού βιβλίου, θα απαιτήσει ένα εντελώς διαφορετικό σχήμα διατήρησης, που μοιάζει περισσότερο με τη διατήρηση σχεσιακών ή/και άλλων βάσεων δεδομένων (Kirchhoff & Morrissey, 2014).

2.2.4 Ψηφιοποιημένα βιβλία

Με την ανάπτυξη των νέων τεχνολογιών, η ψηφιοποίηση των έντυπων πηγών και βιβλίων κατέστη αναγκαία, με σκοπό τα βιβλία να είναι εύκολα προσβάσιμα και αναζητήσιμα σε ψηφιακή μορφή. Βιβλιοθήκες σε όλο τον κόσμο, ανέλαβαν μεγάλα ή μικρότερα έργα ψηφιοποίησης του υλικού τους, όπως η Βρετανική Βιβλιοθήκη, η Εθνική Γαλλική Βιβλιοθήκη και η Εθνική Βιβλιοθήκη της Ολλανδίας, καθώς πραγματοποιήθηκαν και άλλα προγράμματα ψηφιοποίησης, όπως το πρόγραμμα της Google για την ψηφιοποίηση ολόκληρων βιβλιοθηκών. Άλλες εθνικές βιβλιοθήκες ανέλαβαν την ψηφιοποίηση ιστορικού υλικού σε συνεργασία με εμπορικούς φορείς. Οι συνεργασίες αυτές πραγματοποιήθηκαν με στόχο οι εμπορικοί φορείς (όπως για παράδειγμα η ProQuest) να παρέχουν πρόσβαση στο περιεχόμενο, το οποίο ψηφιοποίησαν οι βιβλιοθήκες (Kirchhoff & Morrissey, 2014).

Στις περιπτώσεις των συνεργασιών των βιβλιοθηκών με άλλους φορείς, εκτός των προαναφερθέντων προκλήσεων, προκύπτουν επιπλέον ζητήματα ως προς τα δικαιώματα και τις ευθύνες διατήρησης. Σε ορισμένες περιπτώσεις, προσδιορίζονται ρητές πολιτικές για τη διατήρηση, όπως για παράδειγμα το μέρος στο οποίο θα αποθηκευτεί το υλικό. Παράλληλα, με την αλλαγή της μορφής του αναλογικού υλικού σε ψηφιακό, εγείρονται ζητήματα διασφάλισης της ποιότητας του τεκμηρίου (π.χ. ανεπαρκή μεταδεδομένα,

ελαττωματικές σελίδες και εικόνες), καθώς και ζητήματα πνευματικής ιδιοκτησίας και χρήσης των συγκεκριμένων τεκμηρίων, διότι αλλάζει η μορφή τους και επηρεάζεται η διανομή και η πρόσβαση σε αυτά (Kirchhoff & Morrissey, 2014).

2.3 Εκδόσεις ψηφιακών βιβλιοθηκών

Οι ψηφιακές βιβλιοθήκες και τα αποθετήρια έχουν εντάξει, με αυξανόμενο ρυθμό, υπηρεσίες εκδόσεως ηλεκτρονικών βιβλίων στις πολιτικές τους. Πολλά ιδρυματικά αποθετήρια έχουν αναπτύξει τις δικές τους εκδόσεις, καθιστώντας ευκολότερη τη διανομή και τη διάδοση των νέων συγγραμμάτων στην κοινότητα του Πανεπιστημίου, μέσω της διάθεσης του ψηφιακού περιεχομένου, με καθεστώς ανοικτής πρόσβασης.

Παρόλα αυτά, δημιουργούνται ζητήματα σχετικά με τις γνώσεις των βιβλιοθηκονόμων στο θέμα της διαχείρισης των ηλεκτρονικών εκδόσεων. Απαιτούνται ειδικές γνώσεις του προσωπικού και συνεργασία με εξειδικευμένους εξωτερικούς συνεργάτες (π.χ. γραφίστες) για την ολοκλήρωση των διεργασιών της έκδοσης ενός βιβλίου. Η έκδοση ενός βιβλίου περνά από διάφορα στάδια έως την τελική του δημοσίευση, όπως είναι η γλωσσική επιμέλεια, η αξιολόγηση από τους ειδικούς για την ποιότητα του κειμένου, η γραφιστική επιμέλεια, η καταχώρηση του ISBN κ.ά.

Παράλληλα με την ενασχόλησή τους στον χώρο των εκδόσεων, οι πληροφοριακοί οργανισμοί οφείλουν να λάβουν υπόψη και τη μακροπρόθεσμη διατήρηση του υλικού που παράγεται (Moulaison & Sandy, 2016).

2.3.1 Αυτοέκδοση

Ένας άλλος τομέας που γνωρίζει ευρεία ανάπτυξη είναι το πεδίο της αυτοέκδοσης (self-publishing). Πολλοί ανεξάρτητοι συγγραφείς προτιμούν να δημοσιεύουν τα συγγράμματά τους με τη βοήθεια εφαρμογών και εργαλείων, τα οποία βρίσκονται διαθέσιμα στο διαδίκτυο. Αυτή η επιλογή, τους δίνει τη δυνατότητα να έχουν τον πλήρη έλεγχο κατά τη διάρκεια όλων των διεργασιών της έκδοσης μέχρι την τελική δημοσίευση του τεκμηρίου (Kumar & Dhanya, 2021). Ταυτόχρονα αποτελεί ένα οικονομικότερο μέσο έκδοσης βιβλίων, διότι τα περισσότερα εργαλεία διατίθενται δωρεάν, μέσα από πλατφόρμες και λογισμικά ανοικτής πρόσβασης.

Σύμφωνα με μία έρευνα που πραγματοποιήθηκε στις Ηνωμένες Πολιτείες Αμερικής τον Φεβρουάριο του 2015 στο Midwinter Forum της American Library Association (ALA), με το ερώτημα «Τι χρειάζεται να γνωρίζουν περισσότερο για την αυτοέκδοση οι βιβλιοθήκες;», οι ακαδημαϊκές βιβλιοθήκες που είχαν επιτυχημένη πορεία στις εκδόσεις απάντησαν ότι θα ήθελαν να γνωρίζουν περισσότερα για τα επιχειρηματικά μοντέλα σχετικά με την

αυτοέκδοση, όπως επίσης και για την ευρετηρίαση, την ανάθεση DOI και το μάρκετινγκ (Moulaison & Le Beau, 2015). Τρεις βιβλιοθήκες με λιγότερη πείρα στον τομέα των εκδόσεων ανέφεραν την ανάγκη να γνωρίσουν τις βέλτιστες πρακτικές που έχουν εφαρμοσθεί σε άλλους οργανισμούς (Moulaison & Le Beau, 2015). Οι συμμετέχοντες στην έρευνα, προερχόμενοι από βιβλιοθήκες χωρίς εμπειρία στις εκδόσεις, εξέφρασαν την επιθυμία τους να μάθουν με ποιους τρόπους μπορεί να χρηματοδοτηθεί ένα τέτοιο πρόγραμμα, το όφελος που θα έχει η βιβλιοθήκη έναντι αυτής της επένδυσης, τη διασφάλιση της ποιότητας του υλικού και την ανάγκη για τη διάθεση περισσότερων πόρων σε πρωτοβουλίες των βιβλιοθηκών (Moulaison & Le Beau, 2015).

2.4 Τεχνολογία και Πρότυπα

Για τις ανάγκες της επιμέλειας και της διατήρησης των ηλεκτρονικών βιβλίων, είναι απαραίτητο να αναλυθούν οι πιο κοινές μορφές και πρότυπα που χρησιμοποιούνται στη δημοσίευση και στη διανομή τους. Υπάρχουν δύο κατηγορίες ηλεκτρονικών βιβλίων, τα βιβλία σταθερής διάταξης, τα οποία αφορούν κυρίως την ψηφιοποιημένη μορφή των έντυπων βιβλίων και τα βιβλία που διαθέτουν ανανεώσιμη μορφή ως προς το περιεχόμενό τους. Η δομή του κειμένου, στα βιβλία σταθερής διάταξης, παραμένει αμετάβλητη σε όλες τις συσκευές ανάγνωσης ηλεκτρονικών βιβλίων, αντίθετα τα βιβλία με ανανεώσιμο περιεχόμενο προσαρμόζουν την εμφάνισή τους ανάλογα με την μορφή απεικόνισης της εκάστοτε συσκευής (Kirchhoff, A. & Morrissey, S., 2014). Ως παράδειγμα κειμένου σταθερής διάταξης αποτελεί το μορφότυπο PDF, ενώ κείμενο με ανανεώσιμο περιεχόμενο συναντάται συνήθως σε μορφή HTML, EPUB κ.ά. Τα κείμενα που περιέχουν ανανεώσιμο περιεχόμενο βασίζονται κυρίως στην γλώσσα σήμανσης XML ή στην γλώσσα σήμανσης υπερκειμένου XHTML, η οποία βασίζεται στην XML¹.

2.5 Μορφότυπα για την δημοσίευση ηλεκτρονικών βιβλίων

2.5.1 PDF

Το μορφότυπο PDF σχεδιάστηκε αρχικά για τη διατήρηση της σταθερής διάταξης της δομής των σελίδων, με σκοπό την συμβατή απόδοση του περιεχομένου σε οποιαδήποτε συσκευή ή υπολογιστή (Kirchhoff, A. & Morrissey, S., 2014). Αναπτύσσεται από την εταιρεία Adobe Acrobat και είναι λογισμικό ανοιχτού κώδικα. Η πιο πρόσφατη έκδοσή του είναι η 2.0. Έχει τυποποιηθεί ως πρότυπο από την ISO (International Standard Organization) και αποτελεί την πιο διαδεδομένη μορφή δημοσίευσης ηλεκτρονικών βιβλίων. Εκτός από κείμενο, έχει

¹ <https://www.w3.org/TR/xhtml1/>

τη δυνατότητα να περιέχει εικόνες, υπερσυνδέσμους, πολυμέσα, συνημμένα αρχεία κ.ά. Βασίζεται στην γλώσσα περιγραφής σελίδων PostScript, με αποτέλεσμα κάθε αρχείο PDF να περιέχει μία πλήρης περιγραφή της σταθερής διάταξης του εγγράφου, περιλαμβανομένων του κειμένου, των γραμματοσειρών και των γραφικών. Επίσης, ενσωματώνει χαρακτήρες ASCII 7-bit, εκτός από ορισμένα στοιχεία που μπορεί να έχουν δυαδικό περιεχόμενο. Διατηρεί μία λογική δομή εγγράφου και σημασιολογικού περιεχομένου, μέσω των ετικετών, οι οποίες βοηθούν στην αξιόπιστη εξαγωγή και στην προσβασιμότητα του κειμένου.

Με το εργαλείο Adobe Acrobat DC δίνεται η δυνατότητα δημιουργίας, επεξεργασίας, σχολιασμού, κοινοποίησης και μετατροπής ενός αρχείου². Παρόλο που διατίθεται λογισμικό δημιουργίας αρχείων PDF, σε πολλές περιπτώσεις προτιμάται έγγραφα άλλης μορφής να μετατρέπονται σε PDF, από το να δημιουργηθούν εξ αρχής σε αυτή τη μορφή³.

2.5.2 HTML

Η HTML αποτελεί την κύρια γλώσσα σήμανσης για τις ιστοσελίδες, καθώς διαθέτει τα βασικά στοιχεία δόμησης της εμφάνισής τους. Καθορίζει τη βασική διάταξη του περιεχομένου του κειμένου, όπως τίτλους, παραγράφους, πίνακες κ.ά μέσω των ετικετών σήμανσης. Για τον εμπλουτισμό της παρουσίασης του κειμένου ενσωματώνει γλώσσες όπως η Javascript, η οποία μετατρέπει τις ιστοσελίδες από στατικές σε διαδραστικές, καθώς και εργαλεία μορφοποίησης CSS (Cascading Style sheets) για τον ορισμό της διάταξης του συνόλου του κειμένου⁴. Τα παραπάνω στοιχεία καθιστούν την HTML κατάλληλη για δημοσίευση ηλεκτρονικών βιβλίων με ανανεώσιμο περιεχόμενο. Η τελευταία έκδοση της HTML είναι η HTML5, η οποία παρουσιάζει μια σειρά από νέα στοιχεία που είναι χρήσιμα για τη δημιουργία ανανεώσιμης μορφής εγγράφων.

Τα βασικά χαρακτηριστικά της σήμανσης της γλώσσας HTML είναι τα στοιχεία ή οι ιδιότητες τους (τύποι δεδομένων βασισμένοι σε χαρακτήρες), οι αναφορές χαρακτήρων, οι αναφορές οντοτήτων και η δήλωση τύπου εγγράφου (Document Type Declaration). Τα στοιχεία της HTML αποτελούνται από τις ετικέτες έναρξης και τερματισμού, περιλαμβάνοντας ανάμεσά τους τις ιδιότητες του κειμένου (αναγνωριστικό, τίτλος, γλώσσα κ.ά) και το περιεχόμενο του κειμένου των ετικετών. Οι αναφορές οντοτήτων και χαρακτήρων διασαφηνίζουν την χρήση των διαφόρων συμβόλων και καθορίζουν την

² https://www.adobe.com/gr_en/acrobat/features.html#share

³ <https://docs.fileformat.com/pdf/>

⁴ <https://el.wikipedia.org/wiki/HTML>

ερμηνεία τους, ώστε να γίνονται αντιληπτά κατά την ανάγνωση του εγγράφου από τους διακομιστές.

2.5.3 EPUB

Το EPUB αποτελεί ένα μορφότυπο δημοσίευσης ηλεκτρονικών βιβλίων, διαμορφωμένο για ανάγνωση σε ειδικά σχεδιασμένες συσκευές και εφαρμογές. Η πιο πρόσφατη έκδοση του είναι η 3.2. Η δημιουργία ενός αρχείου EPUB μοιάζει αρκετά με τη δομή ενός ιστότοπου, καθώς βασίζεται στο σχήμα XML. Περιέχει τη γλώσσα σήμανσης HTML και το εργαλείο μορφοποίησης CSS, στοιχεία που θα μπορούσαν να το χαρακτηρίσουν ως ένα μέσο αναπαράστασης και κωδικοποίησης δομημένου και σημασιολογικά βελτιωμένου περιεχομένου ιστοσελίδων σε ενιαία μορφή ενός αρχείου⁵. Παράλληλα, με την ενσωμάτωση της τελευταίας έκδοσης HTML5 οι εκδόσεις βιβλίων σε μορφή EPUB μπορούν να περιέχουν βίντεο, ήχο και διαδραστικά στοιχεία⁶.

Υποστηρίζεται από πολλές συσκευές ανάγνωσης, ανοιχτού ή κλειστού λογισμικού, όπως Apple Books, Caliber, Kobo, Google Play Books κ.ά. Στο αρχείο EPUB μπορεί να εμπεριέχεται η διαχείριση ψηφιακών δικαιωμάτων (Digital Rights Management), επιτρέποντας στην ευχέρεια του εκδότη να επιλέξει το σχήμα δικαιωμάτων που επιθυμεί, καθώς δεν απαιτείται κάποιο συγκεκριμένο σχήμα. Αυτό βέβαια μπορεί να επηρεάσει την προβολή του αρχείου στις συσκευές ανάγνωσης.

Επιπλέον, υπάρχει μια παραλλαγή του EPUB, η μορφή iBooks. Το iBooks είναι η μορφή που παράγεται από την εφαρμογή της Apple, iBooks Author. Αν και βασίζεται στο EPUB, έχει επεκτάσεις που το καθιστούν αποκλειστικό και εμποδίζουν την ανάγνωση ή την επεξεργασία του σε εφαρμογές εκτός του iBooks και του iBooks Author (Kirchhoff, A. & Morrissey, S., 2014).

2.5.4 MOBI

Παρομοίως, ένα μορφότυπο για συσκευές ανάγνωσης ηλεκτρονικών βιβλίων και συσκευές Palm είναι το MOBI. Περιέχει σελιδοδείκτες, σημειώσεις, διορθώσεις και γλώσσα JavaScript. Στο αρχείο του ηλεκτρονικού βιβλίου, μπορεί να περιέχεται η διαχείριση ψηφιακών δικαιωμάτων (DRM) ή η προστασία των πνευματικών δικαιωμάτων, για την αποτροπή αντιγραφής και παράνομης προβολής. Υποστηρίζει ανανεώσιμο περιεχόμενο, το οποίο περιλαμβάνεται ως XHTML και παράλληλα περιλαμβάνει CSS για την απόδοση του περιεχομένου XHTML. Διανέμεται ως συμπιεσμένο αρχείο και μπορεί να περιέχει

⁵<https://web.archive.org/web/20120415004531/http://idpf.org/epub/30/spec/epub30-overview.html>

⁶ <http://www.epubzone.org/epub-3-overview/understanding-epub-3/>

προηγμένα στοιχεία ελέγχου πλοήγησης, ενώ υποστηρίζει και ευρετηρίαση. Τα KF841 και AZW42 της Amazon είναι δύο παραλλαγές του MOBI ειδικά για τη συσκευή ανάγνωσης ηλεκτρονικών βιβλίων Kindle της Amazon. Οι τρέχουσες μορφές Kindle (AZW3, KF8 και KFX) βασίζονται στο MOBI, η οποία αποτελεί μια ιδιόκτητη μορφή που χρησιμοποιείται αποκλειστικά σε συσκευές Kindle. Η μορφή AZW δεν επιτρέπει την JavaScript, ενώ το KF8, η πιο πρόσφατη μορφή Kindle, το επιτρέπει.

Είναι σημαντικό να αναφερθεί ότι οι συσκευές και οι εφαρμογές Kindle δεν αποδίδουν βιβλία σε μορφή EPUB, ενώ τα Nook, iBooks και Kobo δεν αποδίδουν βιβλία σε μορφή MOBI (ή KF8 ή AZW). Οι τελικοί χρήστες μπορούν να χρησιμοποιήσουν μια εφαρμογή υπολογιστή όπως το Calibre για μετατροπή από μια μορφή σε άλλη.

Οι συνηθέστερες μορφές ανάγνωσης που παραδίδουν οι πλατφόρμες προς τον τελικό χρήστη είναι η HTML και το PDF. Στα συστήματα των εκδοτών συνήθως τα βιβλία κωδικοποιούνται σε μορφή XML και στη συνέχεια μετατρέπονται, για λόγους απεικόνισης, σε HTML με τη χρήση CSS, το οποίο χρησιμοποιείται για τη μορφοποίηση του HTML εγγράφου. Για τη σήμανση των ηλεκτρονικών βιβλίων, πολλοί εκδότες χρησιμοποιούν είτε το NCBI Book Tag Set είτε μια έκδοση ή παραλλαγή της δικής τους ιδιόκτητης, εσωτερικής μορφής XML.

Παρόλο που τα επιστημονικά βιβλία των εκδοτών είναι διαθέσιμα στην Amazon βασισμένα στο μορφότυπο MOBI, οι εκδότες τείνουν να παραδίδουν την μορφή XML ή PDF των ηλεκτρονικών βιβλίων στους οργανισμούς συντήρησης. Ο τρόπος παράδοσης των επιστημονικών βιβλίων στους οργανισμούς διατήρησης πραγματοποιείται με τον διαχωρισμό του κάθε βιβλίου σε κεφάλαια. Τα κεφάλαια διατίθενται σε μορφή PDF και το αρχείο PDF κάθε κεφαλαίου αναφέρεται από την σήμανση XML. Πολλές φορές οι εκδότες παραδίδουν ολόκληρο το βιβλίο σε ένα αρχείο. Σε αυτή την περίπτωση τα συνοδευτικά μεταδεδομένα των βιβλίων παρέχονται ως XML σε μορφότυπο ONIX.

Λίγα επιστημονικά βιβλία παραδίδονται στους οργανισμούς διατήρησης διαμορφωμένα σε XML. Η XML που παρέχεται με την εικόνα PDF του περιεχομένου περιέχει βιβλιογραφικά μεταδεδομένα μόνο για τις επικεφαλίδες. Σε αντίθεση, το περιεχόμενο που παρέχεται σε μορφές που έχουν σχεδιαστεί για ειδικές συσκευές και εφαρμογές ανάγνωσης ηλεκτρονικών βιβλίων, όπως το MOBI και το EPUB, βασίζεται στη γλώσσα πλήρους κειμένου XHTML (Kirchhoff, A. & Morrissey, S., 2014).

2.6 Εργαλεία στοιχειοθεσίας κειμένου για έντυπες και ψηφιακές εκδόσεις

2.6.1 Adobe In Design

Το λογισμικό Adobe In Design αποτελεί μία επαγγελματική εφαρμογή η οποία ειδικεύεται στο σχεδιασμό της διάταξης των κειμένων τόσο για έντυπες όσο και για ηλεκτρονικές εκδόσεις. Χρησιμοποιείται ευρέως από επαγγελματίες στον χώρο των εκδόσεων, των περιοδικών, των εφημερίδων και της διαφήμισης. Προσφέρει ακριβή έλεγχο κατά τη διαδικασία δημιουργίας των εγγράφων, ενσωματώνοντας πληθώρα εργαλείων μέσω ενός εύχρηστου περιβάλλοντος σχεδίασης. Το Adobe In Design επιτρέπει την ενοποίηση με εργαλεία όπως το Photoshop, το Illustrator, το Acrobat DC Pro, το InCopy και το Flash Professional, με τα οποία μοιράζεται ένα κοινό περιβάλλον εργασίας χρήστη, καθιστώντας δυνατή τη μετάβαση αρχείων από πρόγραμμα σε πρόγραμμα, διατηρώντας παράλληλα αυτούσια τα αρχεία κινουμένων σχεδίων, βίντεο ή ήχου.

Ένα επιπλέον χαρακτηριστικό του Adobe In Design είναι η δυνατότητα επέκτασής του με τη χρήση άλλων τεχνολογιών όπως είναι το ActionScript, το C++, το JavaScript, το AppleScript και το VBScript. Παράλληλα, οι προγραμματιστές μπορούν να διανείμουν ή και να πουλήσουν τις επεκτάσεις τους σε ολόκληρη την κοινότητα του λογισμικού μέσω του Adobe Exchange, το οποίο είναι ενσωματωμένο με το In Design⁷.

Η τελευταία έκδοση του Adobe In Design είναι η 17.0 και κυκλοφόρησε τον Οκτώβριο του 2021. Διαθέτει μία σειρά εργαλείων σχετικά με την επεξεργασία και δημιουργία στοιχείων σελίδας καθώς και με την επιλογή τύπου, σχημάτων και γραμμών. Το μορφότυπο του αρχείου που δημιουργείται από το Adobe In Design έχει συνήθως κατάληξη .indd, αλλά εμφανίζεται και με τις καταλήξεις .indl, .indt, .indb κ.ά. Η μορφή εξαγωγής των αρχείων μπορεί να πραγματοποιηθεί σε pdf, eps (Encapsulated PostScript), jpg (JPEG), xml (Extensible Markup Language), EPUB, HTML κ.ά.

Η εξαγωγή σε μορφή PDF αποτελεί μία αξιόπιστη επιλογή, διότι το αρχείο μπορεί να κοινοποιηθεί, να προβληθεί και να εκτυπωθεί από οποιονδήποτε διαθέτει το δωρεάν λογισμικό Adobe Reader⁸. Όταν εξάγεται ένα αρχείο In Design σε PDF διατηρεί τα στοιχεία πλοήγησης, όπως τον πίνακα περιεχομένων και το ευρετήριο, καθώς επίσης και λειτουργίες διαδραστικότητας, όπως είναι οι υπερσύνδεσμοι ή τα πολυμέσα. Παράλληλα, το Adobe PDF διαθέτει προεπιλογές οι οποίες βοηθούν στη διαδικασία δημιουργίας του αρχείου PDF. Οι συγκεκριμένες προεπιλογές παρέχουν υψηλής ποιότητας αρχεία (High quality print),

⁷ https://helpx.adobe.com/gr_en/indesign/faq.html.

⁸ https://helpx.adobe.com/gr_en/indesign/using/exporting-publishing-pdf.html.

διατηρούν όλα τα δεδομένα που έχουν δημιουργηθεί μέσω του εργαλείου Illustrator χωρίς απώλεια δεδομένων (Illustrator default), δημιουργεί αρχεία κατάλληλα για προβολή και εκτύπωση των εικόνων που περιέχουν (Acrobat only) καθώς χρησιμοποιείται και για μακροχρόνια διατήρηση ηλεκτρονικών εγγράφων (PDF / A-1b 2005).

Επιπροσθέτως, το In Design παρέχει εξαγωγή αρχείου σε μορφή EPUB για τους χρήστες των ομώνυμων συσκευών ανάγνωσης. Υποστηρίζει τη ροή ανανεώσιμου περιεχομένου δίνοντας τη δυνατότητα βελτιστοποίησης του τεκμηρίου και της προβολής του κειμένου. Επίσης, διαθέτει και την επιλογή σταθερής διάταξης κειμένου, η οποία θεωρείται καταλληλότερη για έγγραφα που περιλαμβάνουν μεγάλο αριθμό γραφικών, αρχείων ήχου και βίντεο⁹. Κατά τη διάρκεια εξαγωγής του EPUB δημιουργείται για κάθε έγγραφο του βιβλίου ένα CSS (Cascading Style Sheets), το οποίο ορίζει τα στυλ του κειμένου, τα οποία είναι μοναδικά, και στο τέλος παράγεται ένα συνολικό CSS για ολόκληρο το βιβλίο.

Παράλληλα, μία άλλη επιλογή εξαγωγής περιεχομένου είναι η μορφή HTML, κατάλληλη για άμεση δημοσίευση στο διαδίκτυο. Η μορφή HTML διατηρεί όλα τα στυλ που δημιουργήθηκαν κατά την επεξεργασία του εγγράφου χρησιμοποιώντας το στυλ CSS. Το In Design εξάγει όλα τα στοιχεία που έχουν μορφοποιηθεί μέσα στο κείμενο (πίνακες, γραφικά, βίντεο, κείμενο, λίστες, παραπομπές κ.ά) εκτός από ετικέτες XML, σελιδοδείκτες ή δείκτες ευρετηρίου¹⁰.

2.6.2 LATEX

Το LATEX είναι ένα σύστημα λογισμικού προετοιμασίας εγγράφων για στοιχειοθεσία υψηλής ποιότητας¹¹. Χρησιμοποιείται ευρέως για τη δημιουργία τεχνικών και επιστημονικών εγγράφων, αλλά και γενικότερα για οποιαδήποτε μορφή δημοσίευσης. Δεν αποτελεί επεξεργαστή κειμένου, αντιθέτως αναφέρεται στη γλώσσα την οποία είναι γραμμένα τα έγγραφα. Η τελευταία αναθεωρημένη έκδοση είναι η LATEX3 και κυκλοφόρησε το 2020. Είναι ένα λογισμικό ανοιχτού κώδικα και διανέμεται με τη δημόσια άδεια LATEX Project.

Τα αρχεία LATEX έχουν τη δυνατότητα να ανοίξουν με οποιοδήποτε πρόγραμμα επεξεργασίας κειμένου και μπορούν να διανεμηθούν μετατρέποντας το αρχείο σε μορφή RTF (Rich Text Format, .rtf) ή XML. Επίσης, μπορεί να αποδοθεί σε αρχείο PDF χρησιμοποιώντας την επέκταση pdfLaTeX. Τα αρχεία LATEX που περιέχουν κείμενο Unicode

⁹ https://helpx.adobe.com/gr_en/indesign/using/export-content-epub-cc.html.

¹⁰ https://helpx.adobe.com/gr_en/indesign/using/export-content-html-cc.html.

¹¹ <https://www.latex-project.org/about/>.

μπορούν να τεθούν υπό επεξεργασία σε μορφή PDF με τις επεκτάσεις TeX, XeLaTeX και LuaLaTeX.

2.6.3 DocBook

Το DocBook είναι μια σημασιολογική γλώσσα για τεχνική τεκμηρίωση, η οποία επιτρέπει στους χρήστες να δημιουργούν περιεχόμενο εγγράφου¹². Εστιάζει στη δομή του περιεχομένου του εγγράφου και όχι στην εμφάνισή του. Το κύριο χαρακτηριστικό του αποτελεί η εύκολη μετατροπή του αρχείου σε μορφές όπως HTML, XHTML, EPUB, PDF κ.ά χωρίς να χρειαστούν αλλαγές στην πηγή του κώδικα.

Σύμφωνα με τον Walsh (2010), το DocBook παρέχει ένα σύστημα για τη σύνταξη δομημένων εγγράφων το οποίο είναι βασισμένο πάνω στα πρότυπα του σχήματος XML. Υποστηρίζεται από πληθώρα εργαλείων, ανοιχτού κώδικα ή κώδικα εμπορικής χρήσης, με πιο πρόσφατη την έκδοση 5.0. Χρησιμοποιεί τη γλώσσα σχήματος RELAX NG με ενσωμάτωση κανόνων Schematron.

Το πρώτο βήμα που ακολουθείται για τη δημιουργία ενός βιβλίου με γλώσσα DocBook είναι η δήλωση XML, η οποία προσδιορίζει τις πτυχές του εγγράφου. Αυτό εξυπηρετεί στην διασφάλιση της διατήρησης των χαρακτηριστικών του βιβλίου σε περίπτωση μελλοντικών αλλαγών ή μεταφοράς του αρχείου σε άλλη μορφή. Η δήλωση XML περιέχει την έκδοση που έχει χρησιμοποιηθεί και τον χαρακτηρισμό της κωδικοποίησης που είναι γραμμένο το έγγραφο. Σε περίπτωση δηλώσεων εντός του εγγράφου χρησιμοποιείται το DTD (Document Type Declaration), το οποίο διαχωρίζει το έγγραφο σε εσωτερικά σημεία του κειμένου, για παράδειγμα σε οντότητες. Οι συγκεκριμένες δηλώσεις αποτελούν ένα εσωτερικό υποσύνολο δηλώσεων.

Αφού ολοκληρωθεί η διαδικασία δημιουργίας του εγγράφου, το βήμα που ακολουθεί είναι η επικύρωση της σήμανσης του τεκμηρίου. Το σχήμα DocBook αποτελεί μία ακριβής περιγραφή της εγκυρότητας της σύνθεσης του εγγράφου, δηλαδή της σειράς των στοιχείων και το περιεχόμενό τους (Walsh, 2010). Με την χρήση του σχήματος RELAX NG η επικύρωση υποστηρίζεται σε περιβάλλοντα επεξεργασίας όπως το <oxygen/>, XMLmind και Emacs (με λειτουργία nxml-mode) (Walsh, 2010).

Με την ολοκλήρωση της επικύρωσης του εγγράφου έπεται το βήμα της διαδικασίας της δημοσίευσής του, είτε σε έντυπη είτε σε ηλεκτρονική μορφή. Για την επίτευξη της δημοσίευσης του εγγράφου, απαραίτητη είναι η χρήση εργαλείων μορφοποίησης (stylesheet), όπως είναι το CSS, το XSL και το XQuery. Τα πιο κοινά πρότυπα που

¹² <https://en.wikipedia.org/wiki/DocBook>.

εφαρμόζονται για αρχεία του DocBook είναι οι νεότερες μορφές του XSL, το XSLT 1.0 και το XSLT 2.0, τα οποία διαθέτουν παρόμοιες τεχνολογίες και μεγάλο εύρος νέων χαρακτηριστικών. Αυτά τα εργαλεία μορφοποίησης βοηθούν στην μετατροπή του DocBook εγγράφου σε ποικιλία μορφών δημοσίευσης, συμπεριλαμβανομένων των HTML, EPUB, XSL-FO, Javahelp και HTML Help.

2.6.4 Pressbooks

Το Pressbooks είναι ένα λογισμικό διαχείρισης περιεχομένου ανοιχτού κώδικα, το οποίο έχει σχεδιαστεί για τη δημιουργία βιβλίων, βασισμένο στην πλατφόρμα WordPress. Το περιεχόμενο που φιλοξενείται στο Pressbooks εμφανίζεται κεφάλαιο προς κεφάλαιο, κυρίως σε μορφή HTML, με δυνατότητα εξαγωγής του αρχείου σε PDF, EPUB, EPUB3, XHTML, MOBI, WXR και / ή ODT (Cullier et al, 2016).

Πολλοί συγγραφείς διαθέτουν δημόσια στο κοινό το αρχείο τους και στη μορφή του Pressbooks (WXR), αλλά για την επεξεργασία και την πρόσβαση σε αυτό το αρχείο απαιτείται εγκατάσταση ή απόκτηση συνδρομής στο Open Textbook Network, μια διαδικτυακή κοινότητα ιδρυμάτων ανώτατης εκπαίδευσης που προάγει τη πρόσβαση και τη χρήση ανοιχτών εγχειριδίων (Cullier et al, 2016).

Τα αρχεία WXR, όπως και τα αρχεία Word, EPUB, HTML, .txt και OpenOffice (.odt), μπορούν να εισαχθούν και να επεξεργαστούν απευθείας στο Pressbooks. Όμως, για να καταστεί δυνατή η εισαγωγή και η δυνατότητα επεξεργασίας ενός αρχείου, θα πρέπει το έγγραφο να διαθέτει συγκεκριμένη και καθορισμένη διάταξη κειμένου (π.χ. διαχωρισμός των κεφαλαίων).

2.6.5 OpenStax

Το OpenStax είναι ένα εργαλείο εκπαιδευτικής τεχνολογίας, το οποίο έχει ως αντικείμενο τη δημιουργία εγχειριδίων ανοιχτής πρόσβασης, τα οποία είναι αξιολογημένα από ειδικούς. Αυτή η πρωτοβουλία ανήκει στο Πανεπιστήμιο Rice και ξεκίνησε το 2012¹³. Όλα τα βιβλία που δημιουργούνται σε αυτήν την πλατφόρμα χρησιμοποιούνται σύμφωνα με τις άδειες Creative Commons (Cullier et al, 2016).

Το εργαλείο επεξεργασίας για τα αρχεία OpenStax (CNX) είναι το Legacy και διατίθεται δωρεάν μέσω του διαδικτύου. Μέσω του Legacy, τα αρχεία μπορούν να εξαχθούν σε μορφή HTML, CNXML, PDF, EPUB. Στο εργαλείο OpenStax μπορούν να εισαχθούν και αρχεία Pressbooks (XML ή WXR) ή EPUB, τα οποία μπορούν να επεξεργασθούν μέσω του διαδικτυακού εργαλείου Pressbooks. Από τα αρχεία XML και

¹³ <https://openstax.org>

WXR μπορούν να εξαχθούν οι μορφές HTML, PDF, EPUB, MOBI, ODT. Από την επεξεργασία του αρχείου EPUB μπορούν να εξαχθούν αρχεία HTML, PDF, EPUB και WXR. Επιπλέον, η μορφή EPUB μπορεί να υποστεί επεξεργασία και μέσω εργαλείων υπολογιστή, όπως είναι το Sigil και το Calibre, όπου το αρχείο εξάγεται μόνο σε μορφή EPUB (Cullier et al, 2016).

2.7 Στρατηγικές Διατήρησης

Οι πληροφοριακοί οργανισμοί ανέκαθεν αποτελούσαν τον παράγοντα-κλειδί στη διαφύλαξη και την διάχυση της γνώσης. Στόχος της πληροφορίας είναι να παράγεται, διανέμεται, χρησιμοποιείται και συνεπώς να αποθηκεύεται για μελλοντική χρήση. Στη σύγχρονη εποχή, με την ανάπτυξη των νέων τεχνολογιών, και ιδιαίτερα του Διαδικτύου, οι ρυθμοί παραγωγής των πληροφοριών είναι ανυπολόγιστοι. Το μεγαλύτερο μέρος του υλικού παράγεται πλέον μόνο ή και σε ψηφιακή μορφή, και αυτό καθιστά την ανάγκη για τη διατήρηση του ψηφιακού αλλά και του αναλογικού υλικού επιτακτική. Στόχος των πληροφοριακών οργανισμών αποτελεί η διατήρηση και η εξασφάλιση μακροχρόνιας πρόσβασης στο υλικό.

Πρώτα απ' όλα, πρέπει να διευκρινιστούν τα ιδιαίτερα χαρακτηριστικά των ψηφιακών τεκμηρίων, σε σύγκριση με τα φυσικά τεκμήρια, πάνω στα οποία βασίζεται ο σχεδιασμός των στρατηγικών διατήρησης. Πρώτο χαρακτηριστικό είναι η μορφή και το περιεχόμενο των τεκμηρίων, διότι εκτός από την πολυπλοκότητα της επεξεργασίας κάθε μορφοτύπου, το τεκμήριο μπορεί να περιέχει εικόνες, βίντεο, αρχεία ήχου ή άλλα πολυμέσα εκτός από κείμενο. Επίσης, η πρόσβαση στα τεκμήρια συνδέεται άμεσα με το μέσο το οποίο φέρει την πληροφορία, καθώς είναι πολύ σημαντικό να μην έχει παλαιωθεί ώστε να είναι δυνατή η ανάγνωση του περιεχομένου και η μελλοντική μεταφορά του σε κάποιο άλλο μέσο. Παράλληλα, ιδιαίτερο ρόλο διαδραματίζει και το λογισμικό, το οποίο χρειάζεται για την αναγνώριση του μέσου και την προβολή του περιεχομένου του (Borghoff, Rodig, Schmitz & Scheffczyk, 2006). Συνεπώς, το ψηφιακό υλικό καθίσταται ευαίσθητο με την πάροδο του χρόνου, καθώς εξαρτάται άμεσα με την τεχνολογία και την εξέλιξη της. Στις τρεις επόμενες ενότητες αναλύονται οι πιο διαδεδομένες πολιτικές και μέθοδοι διατήρησης των ψηφιακών αντικειμένων.

2.7.1 Προσομοίωση

Η προσομοίωση είναι η διαδικασία δημιουργίας εικονικού περιβάλλοντος, το οποίο μιμείται την αρχική πλατφόρμα, με την ενεργοποίηση του δικαιώματος εκτέλεσης του πρωτότυπου λογισμικού. Σε αυτή την πρακτική ο ψηφιακός πόρος διατηρείται αναλλοίωτος, διασφαλίζοντας ότι το λογισμικό που το διαβάζει λειτουργεί όπως το αρχικό.

Αυτό επιτυγχάνεται με την συντήρηση του υπάρχοντος τεχνολογικού εξοπλισμού, σε συνδυασμό πολλές φορές με τη διαδικασία της προσομοίωσης. Το ψηφιακό τεκμήριο σε ένα αποθετήριο ερμηνεύεται από μια εφαρμογή λογισμικού, το οποίο τρέχει πάνω σε ένα λειτουργικό σύστημα, που ελέγχεται από ένα σύνολο υλικού (hardware). Η προσομοίωση πραγματοποιείται προσθέτοντας σε αυτή την στοίβα ένα επιπλέον εξάρτημα αφαίρεσης μεταξύ του λογισμικού και του λειτουργικού συστήματος, διασφαλίζοντας την εκτέλεση του σε όσο το δυνατόν περισσότερες πλατφόρμες. Ένα παράδειγμα εξομοιωτή hardware είναι το Dioscuri, το οποίο είναι γραμμένο σε γλώσσα Java και μπορεί να μεταφερθεί σε οποιοδήποτε υπολογιστή υποστηρίζει Java Virtual Machine (JVM). Παράλληλα, είναι συμβατό με εξαρτήματα υλικού (components), τα οποία προσομοιώνονται από ένα υποκατάστατο λογισμικού (module). Συνδυάζοντας μερικές συμβατές μονάδες, ο χρήστης μπορεί να διαμορφώσει οποιοδήποτε σύστημα υπολογιστή, έχοντας τη δυνατότητα να προσθέσει στο μέλλον αναβαθμισμένα εξαρτήματα, ώστε να μπορέσει ο εξομοιωτής να τα τρέξει.

Πλεονεκτήματα:

- Αναπαράσταση της λειτουργικότητας του αρχικού συστήματος όσο το δυνατόν πλησιέστερα στο καινούριο εικονικό περιβάλλον.
- Η διαδικασία εξομοίωσης εστιάζει στην αλλαγή του περιβάλλοντος, κι όχι στο ψηφιακό τεκμήριο.
- Από άποψη προϋπολογισμού καθίσταται οικονομική μέθοδος, καθώς εφαρμόζεται σε μεγάλες συλλογές και πραγματοποιείται μόνο μία φορά.
- Έχει τη δυνατότητα να αφαιρέσει τον ψηφιακό πόρο από την παλαιά τεχνολογία και να τον εισάγει στην καινούρια πλατφόρμα.

Μειονεκτήματα:

- Η δημιουργία γλώσσας περιγραφής υλικού (hardware) με τη βοήθεια εξειδικευμένων εξαρτημάτων για την ακριβή αναπαράσταση του συστήματος είναι πολύπλοκη και με μεγάλο κόστος.
- Σε πολλές περιπτώσεις χρησιμοποιούνται εμπορικές εφαρμογές και λογισμικό, επομένως προκύπτει ζήτημα πνευματικών δικαιωμάτων.
- Απαιτείται από τους χρήστες να κατανοήσουν ένα σύστημα χωρίς την κατάλληλη εμπειρία.
- Χρειάζονται πόροι και έμπειροι προγραμματιστές υπολογιστών για την καταγραφή για τον κώδικα του εξομοιωτή.

2.7.2 Αντίστροφη μηχανική

Μία άλλη προσέγγιση ψηφιακής διατήρησης, η οποία ερευνάται από την εταιρεία IBM, είναι η αντίστροφη μηχανική. Για τον λόγο αυτό αναπτύχθηκε η εικονική μηχανή Universal Virtual Computer (UVC), μια στρατηγική διατήρησης ψηφιακών πόρων σχεδιασμένη για αρχειακές διεργασίες, η οποία επιτρέπει τόσο την μετανάστευση όσο και την προσομοίωση υλικού σε μία ουδέτερη μορφή γλώσσας, όπως η XML. Παρέχει τη δυνατότητα εφαρμογής εξομοιωτών, όπως το Dioscuri, όμως ως επί το πλείστον χρησιμοποιείται μέσω προσαρμοσμένων εφαρμογών αποκωδικοποίησης μορφοτύπων, τα οποία εκτελούνται κατευθείαν στο UVC. Στη συνέχεια, με τη βοήθεια του σχήματος δεδομένων Logical Data Scheme (LDS), οι πληροφορίες αποθηκεύονται σε μια εσωτερική αναπαράσταση, περιγράφοντας τον τρόπο που κωδικοποιείται η μορφή του τεκμηρίου, καθορίζοντας παράλληλα μεθόδους πρόσβασης στην πληροφορία. Τη στιγμή που γράφεται το σχήμα δεδομένων, μία μονάδα λογισμικού πρέπει να δημιουργείται με σκοπό την αποκωδικοποίηση της μορφοποίησης και την έξοδο της πληροφορίας. Μελλοντικά, οι προδιαγραφές μπορούν να χρησιμοποιηθούν για την ανοικοδόμηση του UVC για οποιαδήποτε τρέχουσα πλατφόρμα, ενώ το LDS μπορεί να χρησιμοποιηθεί για την κατασκευή ενός προγράμματος επαναφοράς που χρησιμοποιεί τη μονάδα UVC για να διαβάσει το ψηφιακό αντικείμενο και στη συνέχεια είτε αποδίδει το αντικείμενο είτε γράφει τις πληροφορίες σε μια νέα μορφή αρχείου.

2.7.3 Μετανάστευση

Η στρατηγική μετανάστευσης αντιπροσωπεύει τη μεταφορά του υλικού σε καινούρια και αναβαθμισμένα συστήματα, είτε αυτό αφορά το μορφότυπο των τεκμηρίων είτε το λειτουργικό σύστημα, με στόχο να παραμένουν οι πόροι προσβάσιμοι και λειτουργικοί. Λόγω του συνεχώς αυξανόμενου αριθμού μορφών κειμένου, αρχείων εικόνας και ήχου, βάσεων δεδομένων και σύνθετων εγγράφων, δεν μπορεί να εφαρμοσθεί μία ενιαία στρατηγική μετανάστευσης για όλες τις περιπτώσεις υλικού αρχείων. Παράλληλα, η εξέλιξη της τεχνολογίας και των λειτουργικών συστημάτων καθιστά τη μετανάστευση μία περιοδική στρατηγική, καθώς χρειάζεται συνεχής ανανέωση των μέσων μεταφοράς του υλικού. Οι βασικοί στόχοι της συγκεκριμένης στρατηγικής είναι:

- Η διατήρηση των δεδομένων χωρίς απώλεια της πληροφορίας.
- Διαφύλαξη από μη εξουσιοδοτημένη πρόσβαση.
- Συνεχής πρόσβαση στο υλικό.
- Εκτέλεση σε λειτουργικά συστήματα.

Πλεονεκτήματα:

- Είναι κατάλληλη στρατηγική για έγγραφα κειμένου.
- Αξιόπιστη για τη διατήρηση ψηφιακού περιεχομένου.
- Διατίθεται εύκολα το λογισμικό μετατροπής.

Μειονεκτήματα:

- Πολλά και διαφορετικά μορφότυπα πηγών.
- Διαρκής μετανάστευση των αρχείων λόγω της τεχνολογικής ανάπτυξης.
- Απώλεια πληροφοριών κατά τη διάρκεια της μετανάστευσης.
- Κίνδυνος απώλειας της αυθεντικότητας των αρχείων.
- Εξειδικευμένα εργαλεία μετατροπής της μορφής του υλικού.

2.8 OAIS Reference Model

Το μοντέλο αναφοράς OAIS (Open Archival Information System) είναι ένα πρότυπο του Διεθνούς Οργανισμού Τυποποίησης (ISO) και χρησιμοποιείται για την μακροπρόθεσμη διατήρηση ψηφιακών πληροφοριών. Η ιδέα για τη δημιουργία του μοντέλου αναφοράς OAIS ξεκίνησε το 1995 όταν ο Διεθνής Οργανισμός Τυποποίησης (ISO) ανέθεσε στη Συμβουλευτική Επιτροπή για Συστήματα Διαστημικών Δεδομένων (CCSDS) την ανάπτυξη αυτού του προτύπου. Η πρώτη έκδοση ολοκληρώθηκε τον Μάιο του 1999, η οποία αναθεωρήθηκε πολλές φορές, μέχρι την καθιέρωση της έκδοσης του 2005 ως πρότυπο ISO.

Ο σκοπός της δημιουργίας του OAIS προήλθε από τις ανάγκες των οργανισμών να διατηρήσουν μακροπρόθεσμα τις πληροφορίες και να τις διαθέσουν σε μία Καθορισμένη Κοινότητα¹⁴. Ο όρος «μακροπρόθεσμα» δεν μπορεί να προσδιορισθεί χρονικά, διότι καθορίζεται από τους ρυθμούς ανάπτυξης της τεχνολογίας και τις ανάγκες τις εκάστοτε κοινότητας.

«Ως αμιγώς εννοιολογικό πλαίσιο, το μοντέλο OAIS δεν απαιτεί τη χρήση κάποιας συγκεκριμένης υπολογιστικής πλατφόρμας, περιβάλλοντος συστήματος, παραδείγματος σχεδιασμού συστήματος, μεθοδολογία ανάπτυξης συστήματος, σύστημα διαχείρισης βάσεων δεδομένων, πρότυπο σχεδίασης βάσης δεδομένων, γλώσσα ορισμού δεδομένων, γλώσσα εντολών, διεπαφή συστήματος, διεπαφή χρήστη, τεχνολογία ή πολυμέσα ώστε ένα αρχείο να είναι συμβατό. Στόχος του είναι να θέσει τα πρότυπα για τις δραστηριότητες που εμπλέκονται στη διατήρηση ενός ψηφιακού αρχείου και όχι τη μέθοδο για την πραγματοποίηση αυτών των δραστηριοτήτων»¹⁵.

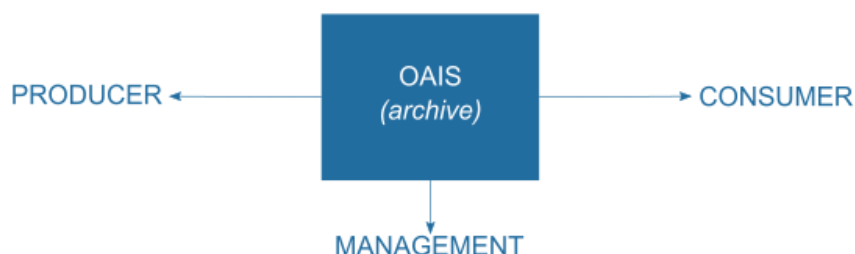
¹⁴ <http://www.oais.info>

¹⁵ https://en.wikipedia.org/wiki/Open_Archival_Information_System

Το Ανοιχτό Αρχειακό Πληροφοριακό Σύστημα (OAIS) αποτελεί ένα Αρχείο, το οποίο οργανώνεται από έναν οργανισμό, σε συνδυασμό με το ανθρώπινο δυναμικό και τα συστήματα που έχουν αναλάβει την ευθύνη της ψηφιακής διατήρησης των πληροφοριών του οργανισμού. Η βαρύτητα δίνεται στο σύνολο των δραστηριοτήτων που απαιτούνται, κι αυτό διαχωρίζει το πρότυπο από άλλες χρήσεις του όρου «αρχείο». Ονομάζεται «Ανοιχτό», καθώς αναφέρεται στο τρόπο που αναπτύσσεται η σύστασή του, κι όχι στο ότι μπορεί να χρησιμοποιείται ελεύθερα χωρίς περιορισμούς.

2.8.1 Το περιβάλλον του OAIS

Το αρχείο OAIS περιβάλλεται από τρεις οντότητες, τους Παραγωγούς (Producers), τους Καταναλωτές (Consumers) και τη Διαχείριση (Management) (βλέπε Εικόνα 1).



Εικόνα 1: Περιβάλλον του OAIS

(https://en.wikipedia.org/wiki/Open_Archival_Information_System, 2022)

Ο ρόλος του **Παραγωγού** αναφέρεται στα άτομα ή τα συστήματα πελατών που παραχωρούν τις πληροφορίες που πρέπει να διατηρηθούν.

Οι **Καταναλωτές** αποτελούνται από άτομα ή συστήματα πελατών, τα οποία αλληλοεπιδρούν με τις υπηρεσίες του OAIS για να βρουν και να ανακτήσουν τις διατηρημένες πληροφορίες που επιθυμούν. Είναι γνωστοί και ως Καθορισμένη Κοινότητα (Designated Community), οι οποίοι θα πρέπει να έχουν τη δυνατότητα να κατανοούν τις διατηρούμενες πληροφορίες.

Ο ρόλος της **Διαχείρισης** εφαρμόζεται από τα άτομα που δημιουργούν και καθορίζουν το πλαίσιο των πολιτικών και των διαδικασιών που θα ακολουθηθούν στο εννοιολογικό μοντέλο OAIS.

Οι παραπάνω οντότητες μπορούν να αλληλοεπιδράσουν μεταξύ τους και η καθεμία να παίξει έναν διαφορετικό ρόλο σε ένα άλλο Αρχείο OAIS. Για παράδειγμα, ένα Αρχείο OAIS μπορεί να αναλάβει το ρόλο του Καταναλωτή σε ένα άλλο Αρχείο, δηλαδή να εξάγει πληροφορίες, όταν το πρώτο Αρχείο, του Καταναλωτή, χρειαστεί να βασιστεί στο δεύτερο για έναν τύπο πληροφοριών που δεν χρησιμοποιεί συχνά και δεν επιθυμεί να διατηρήσει στο δικό του περιβάλλον.

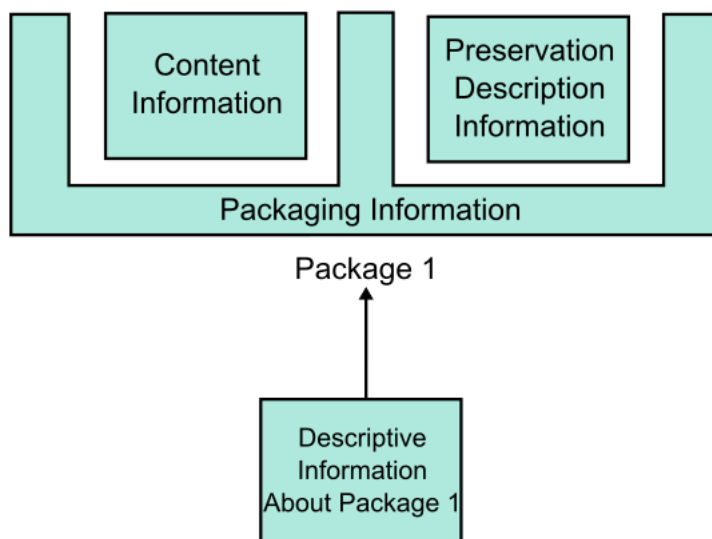
2.8.2 Πακέτο Πληροφοριών (Information Package)

Το μοντέλο OAIS διαμορφώνει ένα μοντέλο Πακέτων Πληροφοριών (βλέπε Εικόνα 2). Τα φυσικά ή ψηφιακά στοιχεία, που διατηρούνται στο Αρχείο και φέρουν πληροφορία, ονομάζονται αντικείμενα δεδομένων. Κάθε πληροφορία, η οποία εισέρχεται στο Αρχείο OAIS από έναν Παραγωγό και κάθε πληροφορία που διανέμεται στον Καταναλωτή, πρέπει να ερμηνεύεται και να γίνεται κατανοητή, είτε μέσω μιας εγκατεστημένης Γνωσιακής Βάσης, είτε μέσω συμπληρωματικών «Πληροφοριών Αναπαράστασης» που περιέχει το αντικείμενο δεδομένων. Έτσι, δημιουργείται το Πακέτο Πληροφοριών, το οποίο περιλαμβάνει τα εξής τέσσερα αντικείμενα πληροφορίας:

- **Πληροφορίες Περιεχομένου (Content Information):** αποτελούν τη βασική πληροφορία του αντικειμένου που πρέπει να διατηρηθεί. Ενσωματώνει το περιεχόμενο του αντικειμένου δεδομένων και τη σχετιζόμενη «Πληροφορία Αναπαράστασης», η οποία είναι αναγκαία για την ερμηνεία της πληροφορίας από την Καθορισμένη Κοινότητα ή τον Καταναλωτή.
- **Πληροφορίες Περιγραφής Διατήρησης (Preservation Description Information):** είναι αναγκαίες για τη διατήρηση των Πληροφοριών Περιεχομένου, για την προφύλαξη της σαφήνειάς τους και την κατανόηση του αρχικού περιβάλλοντος στο οποίο δημιουργήθηκαν. Υπάρχουν πέντε τύποι πληροφοριών που πρέπει να διαφυλαχθούν από το αρχικό αντικείμενο:
 - Η **προέλευση**, η οποία περιγράφει την πηγή των Πληροφοριών Περιεχομένου, ποιος είχε την αρχική επιμέλεια της δημιουργίας του τεκμηρίου και την ιστορία του.
 - Το **πλαίσιο**, όπου αναλύεται ο τρόπος που σχετίζονται οι Πληροφορίες Περιεχομένου με άλλες πληροφορίες εκτός του Πακέτου Πληροφοριών.
 - Η **αναφορά**, η οποία προσδιορίζει ένα ή περισσότερα μοναδικά αναγνωριστικά του τεκμηρίου, όπως το ISBN ενός βιβλίου, ένα handle κ.ά.
 - Η **σταθερότητα**, η οποία προστατεύει τις Πληροφορίες Περιεχομένου μέσω ενός τείχους προστασίας από μη τεκμηριωμένες τροποποιήσεις.
 - Τα **δικαιώματα πρόσβασης**, τα οποία προσδιορίζουν τους όρους χρήσης, πρόσβασης, διανομής και διατήρησης των Πληροφοριών Περιεχομένου.
- **Πληροφορίες Συσκευασίας (Packaging Information):** είναι οι πληροφορίες, οι οποίες ορίζουν και συσχετίζουν τις Πληροφορίες Περιεχομένου και τις Πληροφορίες Περιγραφής Διατήρησης. Αυτές, για παράδειγμα, μπορεί να

περιέχουν πληροφορίες για τον τόμο του τεκμηρίου ή κατάλογο πληροφοριών που περιέχονται σε αυτό.

- **Περιγραφικές Πληροφορίες (Descriptive Information):** αυτές οι πληροφορίες προσδιορίζουν το πακέτο που περιέχει τις Πληροφορίες Περιεχομένου που έχουν τεθεί σε αναζήτηση.



Εικόνα 2: Πακέτο πληροφοριών του OAIS

(https://en.wikipedia.org/wiki/Open_Archival_Information_System, 2022)

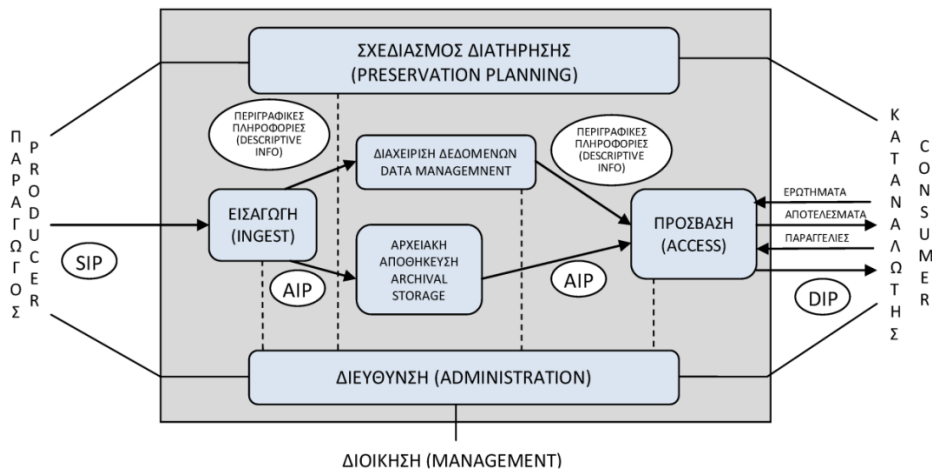
Υπάρχουν τρεις τύποι Πληροφοριακών Πακέτων:

- **Πακέτο Πληροφοριών Υποβολής (Submission Information Package),** είναι το πακέτο που εναποτίθεται στον OAIS από τον Παραγωγό. Περιέχει ορισμένες Πληροφορίες Περιεχομένου και Περιγραφής Διατήρησης.
- **Πακέτο Αρχειακών Πληροφοριών (Archival Information Package),** το οποίο δημιουργείται μέσα από το Πακέτο Πληροφοριών Υποβολής και μπορούν να παραχθούν ένα ή περισσότερα πακέτα για διατήρηση. Περιέχει το σύνολο των Πληροφοριών Περιγραφής Διατήρησης, οι οποίες σχετίζονται με τις Πληροφορίες Περιεχομένου. Σε αυτό το πακέτο, οι Πληροφορίες Συσκευασίας εναρμονίζονται με τα ενσωματωμένα πρότυπα του Παραγωγού, και ενδέχεται να αλλάζουν από τη διαχείριση του OAIS.
- **Πακέτο Πληροφοριών Διάδοσης (Dissemination Information Package),** περιλαμβάνει το σύνολο ή ένα μέρος του Πακέτου Πληροφοριών Αρχείων που παρέχονται στον Καταναλωτή, έπειτα από το αίτημά του. Υπάρχει πιθανότητα να περιέχει Πληροφορίες Περιγραφής Διατήρησης, αλλά σίγουρα θα περιλαμβάνει

Πληροφορίες Συσκευασίας κάποιων στοιχείων του τεκμηρίου, ώστε να διακρίνονται με σαφήνεια οι πληροφορίες που ζητήθηκαν.

2.8.3 Το λειτουργικό μοντέλο (Functional model)

Το λειτουργικό μοντέλο του OAIS (βλέπε **Σφάλμα! Το αρχείο προέλευσης της αναφοράς δεν βρέθηκε.**) αποτυπώνει τις κύριες ροές πληροφοριών και τις σχετιζόμενες διεπαφές των λειτουργικών οντοτήτων. Τις λειτουργικές οντότητες αποτελούν η Απορρόφηση (Ingest), η Διαχείριση των Δεδομένων (Data Management), η Αποθήκευση των Αρχείων (Archival Storage), ο Σχεδιασμός Διατήρησης (Preservation Planning), η Διοίκηση (Administration), και η Πρόσβαση (Access). Τα βέλη συνδέουν τις οντότητες και προσδιορίζουν τα κανάλια επικοινωνίας από τα οποία περνάει η πληροφορία.



Εικόνα 3: Το λειτουργικό μοντέλο του OAIS (Functional model)

(Ζερβός, 2015)

Οι ρόλοι των λειτουργικών οντοτήτων είναι οι εξής:

- **Εισαγωγή (Ingest):** παρέχει τις υπηρεσίες και τις λειτουργίες για την αποδοχή των Πακέτων Πληροφοριών Υποβολής (SIP) που υποβάλλονται από τους Παραγωγούς και προετοιμάζει το περιεχόμενο για διαχείριση και αποθήκευση. Οι λειτουργίες που διενεργούνται περιλαμβάνουν, τη λήψη των Πακέτων Πληροφοριών Υποβολής (SIP), όπου διασφαλίζει την ποιότητά τους, δημιουργεί το Πακέτο Αρχειακών Πληροφοριών (AIP), όπου συμμορφώνεται με τα πρότυπα μορφοποίησης και τεκμηρίωσης των δεδομένων του αρχείου, εξάγει Περιγραφικές Πληροφορίες (Descriptive Information) από τα Πακέτα Αρχειακών Πληροφοριών (AIP) και συντονίζει τις ενημερώσεις στις λειτουργικές οντότητες Διαχείρισης των Δεδομένων (Data Management) και Αποθήκευσης των Αρχείων (Archival Storage).

- **Διαχείριση Δεδομένων (Data Management):** παρέχει τις υπηρεσίες και τις λειτουργίες για τη συμπλήρωση, τη διατήρηση και την πρόσβαση στις Περιγραφικές Πληροφορίες (Descriptive Information), οι οποίες προσδιορίζουν και τεκμηριώνουν τα Αρχεία και διατηρούν τα πρότυπα μορφοποίησης και τεκμηρίωσης των δεδομένων τους, τα οποία χρησιμοποιούνται για τη διαχείριση των Αρχείων. Οι λειτουργίες αυτές περιλαμβάνουν, τη διαχείριση των λειτουργιών της βάσης δεδομένων του Αρχείου, την εκτέλεση ενημερώσεων σε αυτήν, την εκπλήρωση και απάντηση αιτημάτων στα δεδομένα διαχείρισης και τη δημιουργία αναφορών μέσω αυτών των απαντήσεων.
- **Αρχειακή Αποθήκευση (Archival Storage):** παρέχει τις υπηρεσίες και τις λειτουργίες για την αποθήκευση, τη συντήρηση και την ανάκτηση των Πακέτων Αρχειακών Πληροφοριών (AIP). Οι λειτουργίες αυτές περιλαμβάνουν, τη λήψη των Πακέτων Αρχειακών Πληροφοριών (AIP) από την Εισαγωγή (Ingest), την προσθήκη τους σε ένα μόνιμο χώρο αποθήκευσης, τη διαχείριση της ιεραρχίας της αποθήκευσής τους, την ανανέωση των μέσων στα οποία αποθηκεύονται, την εκτέλεση τακτικών ελέγχων για τυχόν σφάλματα, την παροχή δυνατοτήτων ανάκτησης από καταστροφές και την προετοιμασία των Πακέτων Αρχειακών Πληροφοριών (AIP) να εκπληρώσουν αιτήματα.
- **Σχεδιασμός Διατήρησης (Preservation Planning):** παρέχει τις υπηρεσίες και τις λειτουργίες για την παρακολούθηση του περιβάλλοντος του OAIS, παρέχοντας σχέδια διατήρησης ώστε να διασφαλίσει την πρόσβαση στις πληροφορίες που είναι αποθηκευμένες στο εννοιολογικό μοντέλο OAIS, ακόμα κι αν το αρχικό υπολογιστικό περιβάλλον του Αρχείου καταστεί απαρχαιωμένο. Οι λειτουργίες του Σχεδιασμού Διατήρησης περιλαμβάνουν, την αξιολόγηση του περιεχομένου του Αρχείου, την περιοδική ενημέρωση των αρχειακών πληροφοριών, τη σύσταση της μετανάστευσης του Αρχείου αν απαιτείται, την ανάπτυξη προτάσεων για τα πρότυπα και τις πολιτικές που ήδη ακολουθούνται, την παροχή αναφορών ανάλυσης κινδύνου, καθώς και την παρακολούθηση αλλαγών στο τεχνολογικό περιβάλλον. Επίσης, έχει τη δυνατότητα να αναπτύξει λεπτομερή σχέδια μετανάστευσης του Αρχείου, να δημιουργήσει πρότυπα λογισμικού και στη συνέχεια να τα εφαρμόσει μέσω δοκιμών στα σχέδια μετανάστευσης της Διοίκησης (Administration).
- **Διοίκηση (Administration):** παρέχει υπηρεσίες και λειτουργίες για τη συνολική λειτουργία του συστήματος Αρχαιοθήκης. Οι λειτουργίες αυτές περιλαμβάνουν,

την αναζήτηση και τη διαπραγμάτευση συμφωνιών υποβολής με τους Παραγωγούς, τον έλεγχο των υποβολών για τη διασφάλιση της πληρότητας των προτύπων Αρχαιοθέτησης και τη διατήρηση της μορφοποίησης του λογισμικού και του υλικού συστήματος. Παράλληλα, διαθέτει μηχανικές συστημικές λειτουργίες για την παρακολούθηση και τη βελτίωση των λειτουργιών του Αρχείου, καθώς και για την απογραφή ή τη μετεγκατάσταση / ενημέρωση του περιεχομένου του. Επίσης, η Διοίκηση είναι υπεύθυνη για την καθιέρωση και τη διατήρηση προτύπων και πολιτικών, ενώ παρέχει τη δυνατότητα ενεργοποίησης παλαιότερων αιτημάτων.

- **Πρόσβαση (Access):** παρέχει τις υπηρεσίες και τις λειτουργίες που βοηθούν τους Καταναλωτές να προσδιορίσουν την ύπαρξη, την περιγραφή, την τοποθεσία και τη διαθεσιμότητα των πληροφοριών που είναι αποθηκευμένες στο μοντέλο OAIS, επιτρέποντάς τους να αιτηθούν και να λάβουν πακέτα πληροφοριών. Οι λειτουργίες της Πρόσβασης περιλαμβάνουν, την επικοινωνία με τους Καταναλωτές για τη λήψη αιτημάτων, την εφαρμογή ελέγχων για τον περιορισμό της πρόσβασης σε ειδικά προστατευμένες πληροφορίες, τον συντονισμό των εργασιών για την εκτέλεση των αιτημάτων και την ολοκλήρωση και παράδοση των απαντήσεων προς τους Καταναλωτές.

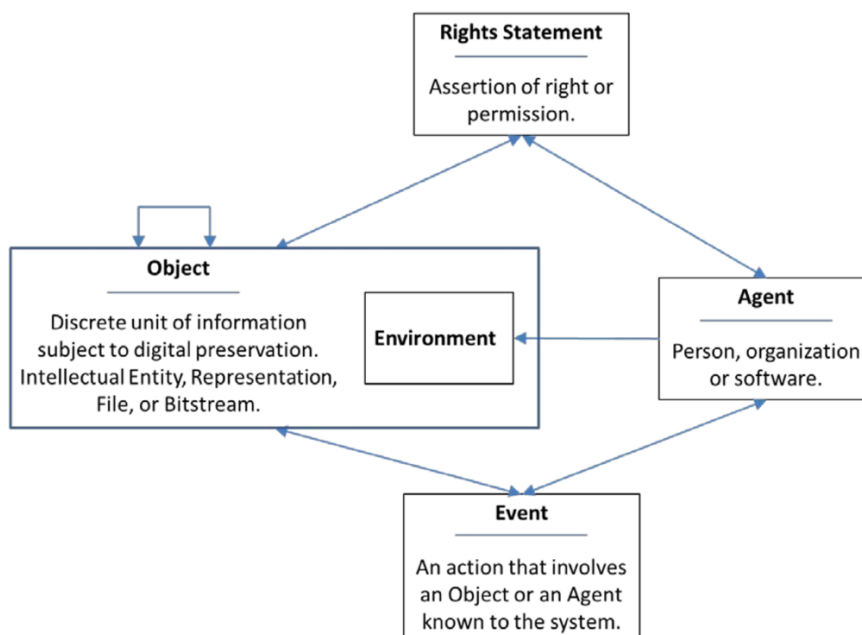
2.9 Πρότυπα διατήρησης μεταδεδομένων

2.9.1 PREMIS

Για την μακροπρόθεσμη διατήρηση των ψηφιακών αντικειμένων, εκτός από το ίδιο το τεκμήριο και το περιεχόμενό του, είναι απαραίτητη και η διατήρηση των πληροφοριών που προσδιορίζουν το τεκμήριο, δηλαδή τα μεταδεδομένα του, ώστε να εξασφαλίζεται η αναζήτηση, ο εντοπισμός και η πρόσβαση σε αυτό.

Το πιο διαδεδομένο πρότυπο για τη διατήρηση των μεταδεδομένων είναι το PREMIS (PREservation Metadata: Implementation Strategies) και η πιο πρόσφατη έκδοση του είναι η 3.0. Το PREMIS αποτελεί ένα λεξικό δεδομένων, το οποίο ορίζεται από σημασιολογικές ενότητες. Κάθε σημασιολογική ενότητα αντιστοιχίζεται σε μία οντότητα, η οποία είναι δομημένη μέσα σε ένα μοντέλο δεδομένων, οπότε η σημασιολογική ενότητα μπορεί να χαρακτηριστεί ως ιδιότητα της οντότητας¹⁶. Οι τέσσερις οντότητες που ορίζει το μοντέλο είναι τα Αντικείμενα (Object), τα Γεγονότα (Events), οι Πράκτορες (Agents) και τα Δικαιώματα (Rights) (βλέπε Εικόνα 4).

¹⁶ <https://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf>



Εικόνα 4: Πρότυπο διατήρησης δεδομένων PREMIS

(PREMIS committee, 2015)

Όπως φαίνεται στο σχήμα, οι οντότητες αποτυπώνονται σε πλαίσια και οι σχέσεις μεταξύ των οντοτήτων προσδιορίζονται από βέλη. Όταν τα βέλη είναι αμφίδρομα, τότε κάθε τύπος οντότητας περιέχει μία σημασιολογική ενότητα που του επιτρέπει να συνδέεται με τον άλλο (PREMIS committee, 2015). Για παράδειγμα, η οντότητα Γεγονότων περιλαμβάνει μία σημασιολογική ενότητα που καταγράφει πληροφορίες σχετικά με έναν Πράκτορα, και αντίστοιχα η οντότητα Πρακτόρων περιλαμβάνει μία σημασιολογική ενότητα που καταγράφει πληροφορίες σχετικά με ένα Γεγονός.

Στην οντότητα Αντικειμένων, όπου το βελάκι έχει αμφίδρομη πορεία προς την ίδια οντότητα, υποδεικνύει ότι οι σημασιολογικές ενότητες υποστηρίζουν την καταγραφή σχέσεων μεταξύ των Αντικειμένων. Σύμφωνα με την Εικόνα 4, καμία άλλη οντότητα δεν μπορεί να συσχετιστεί με τον εαυτό της. Οι οντότητες του μοντέλου ορίζονται ως εξής:

- **Αντικείμενο (Object):** είναι η μονάδα πληροφοριών που υπόκειται σε ψηφιακή διατήρηση.
- **Περιβάλλον (Environment):** είναι η τεχνολογία που χρησιμοποιείται (το λογισμικό ή το υλικό) που υποστηρίζει το Αντικείμενο. Το Περιβάλλον μπορεί να αποτυπωθεί ως το μέρος, μέσα στο αποθετήριο, όπου διατηρείται το Αντικείμενο (π.χ. αρχείο).
- **Γεγονός (Events):** είναι η ενέργεια που περιλαμβάνει ή επηρεάζει τουλάχιστον ένα Αντικείμενο ή Πράκτορα που βρίσκεται στο αποθετήριο.

- **Πράκτορας (Agents):** είναι άτομο, οργανισμός ή σύστημα λογισμικού που σχετίζεται με τα Γεγονότα ενός Αντικείμενου, ή με τα Δικαιώματα που συνδέονται με ένα Αντικείμενο.
- **Δήλωση Δικαιωμάτων (Rights):** είναι δήλωση ενός ή περισσότερων Δικαιωμάτων ή αδειών που σχετίζονται με ένα Αντικείμενο και / ή Πράκτορα.

2.9.2 METS

Το Πρότυπο Κωδικοποίησης και Μεταφοράς Μεταδεδομένων (Metadata Encoding and Transmission Standard, METS) είναι μία προδιαγραφή κωδικοποίησης και μετάδοσης δεδομένων, εκφρασμένη σε σχήμα XML, η οποία παρέχει τα μέσα για τη μεταφορά των μεταδεδομένων, που είναι απαραίτητα τόσο για τη διαχείριση των ψηφιακών αντικειμένων όσο και για την ανταλλαγή τους μεταξύ των αποθετηρίων. Ανάλογα με τη χρήση του, ένα έγγραφο METS θα μπορούσε να χρησιμοποιηθεί στο ρόλο του Πακέτου πληροφοριών Υποβολής (SIP), του Πακέτου Πληροφοριών Αρχαιοθήκης (AIP) ή του Πακέτου Πληροφοριών Διάδοσης (DIP) εντός του μοντέλου αναφοράς OAI¹⁷. Σχεδιάστηκε για να επιτρέπει την συνεργασία για την ανάπτυξη εργαλείων και υπηρεσιών διαχείρισης πληροφοριών και να διευκολύνει τη διαλειτουργικότητα με στόχο την ανταλλαγή ψηφιακού υλικού μεταξύ των οργανισμών¹⁸.

Ένα έγγραφο METS αποτελείται από επτά ενότητες, οι οποίες περιέχουν γνωρίσματα και στοιχεία για την περιγραφή τους. Οι ενότητες είναι οι εξής:

- **Επικεφαλίδα METS:** περιλαμβάνει πληροφορίες σχετικά με το ίδιο το έγγραφο, όπως δημιουργό, εκδότη, ημερομηνία δημιουργίας ή τροποποίησης, αναγνωριστικά ή συντελεστές.
- **Περιγραφικά Μεταδεδομένα:** αυτή η ενότητα οδηγεί σε περιγραφικά δεδομένα εντός ή εκτός του εγγράφου METS, όπου μπορεί να συνδέεται με μία εγγραφή MARC από ένα κατάλογο μιας βιβλιοθήκης ή μία περιγραφή EAD (πρότυπο αρχειακής περιγραφής) από ένα διακομιστή, ενώ υπάρχει πιθανότητα να περιέχει και τις δύο περιπτώσεις.
- **Μεταδεδομένα Διαχείρισης:** παρέχει πληροφορίες σχετικά με τον τρόπο δημιουργίας και αποθήκευσης των αρχείων, με τα δικαιώματα πνευματικής ιδιοκτησίας και με την προέλευση του ψηφιακού αντικείμενου. Και εδώ, όπως με

¹⁷ <https://www.loc.gov/standards/mets/METSOverview.v2.html>

¹⁸ <https://www.loc.gov/standards/mets/METSPrimer.pdf>

τα Περιγραφικά Μεταδεδομένα, οι πληροφορίες μπορεί να συνδέονται με δεδομένα εντός ή εκτός του εγγράφου METS.

- **Ενότητα Φακέλου:** σε αυτή την ενότητα παρατίθενται οι φακέλοι που περιλαμβάνουν περιεχόμενο σχετικά με τις ηλεκτρονικές εκδόσεις του ψηφιακού αντικειμένου (pdf, html κ.ά). Σε αυτό τον φάκελο υποδιαιρούνται τα αρχεία των μορφών των ηλεκτρονικών εκδόσεων του τεκμηρίου.
- **Δομικός Χάρτης:** περιέχει την περιγραφή της ιεραρχικής δομής του εγγράφου. Ο Δομικός Χάρτης συνδέει τα στοιχεία της ιεραρχικής δομής με αρχεία περιεχομένου και με μεταδεδομένα που σχετίζονται με κάθε στοιχείο της δομής.
- **Δομικές Συνδέσεις:** σε αυτή την ενότητα οι δημιουργοί έχουν τη δυνατότητα να καταγράφουν την ύπαρξη υπερσυνδέσμων μεταξύ των στοιχείων του Δομικού Χάρτη, το οποίο εξυπηρετεί την αρχειοθέτηση των ιστοσελίδων.
- **Ενότητα Συμπεριφοράς:** η λειτουργία της ενότητας συμπεριφοράς είναι να συσχετίσει εκτελέσιμες συμπεριφορές στο περιεχόμενο του εγγράφου METS. Περιέχει ένα στοιχείο ορισμού διεπαφής, το οποίο εκπροσωπεί ένα σύνολο συμπεριφορών, οι οποίες αντιπροσωπεύονται από μία συγκεκριμένη ενότητα συμπεριφοράς.

2.10 Ιδρυματικά Αποθετήρια και Ψηφιακή Διατήρηση

Σύμφωνα με τον Lynch (2003), ένα Ιδρυματικό Αποθετήριο είναι «ένα σύνολο υπηρεσιών που προσφέρει ένα Πανεπιστήμιο στα μέλη της κοινότητάς του για τη διαχείριση και τη διάδοση ψηφιακού υλικού που δημιουργείται από το ίδρυμα και τα μέλη της κοινότητάς του. Είναι ουσιαστικά μια οργανωτική δέσμευση για τη διαχείριση αυτών των ψηφιακών υλικών, συμπεριλαμβανομένης της μακροπρόθεσμης διατήρησης, όπου χρειάζεται, καθώς και της οργάνωσης και της πρόσβασης ή διανομής».

Οι διαχειριστές των ιδρυματικών αποθετηρίων ενώ υποστηρίζουν τη δέσμευσή τους ως προς τη μακροπρόθεσμη διατήρηση του υλικού και τη πρόσβαση σε αυτό, δεν υπάρχουν αρκετά στοιχεία σχετικά με το πώς ενσωματώνουν στρατηγικές διατήρησης στα αποθετήρια τους (Barrueco & Termens, 2022). Οι περισσότεροι οργανισμοί, από την αρχική δημιουργία των αποθετηρίων, εστίαζαν το ενδιαφέρον τους στον εμπλουτισμό των συλλογών τους και στη διαλειτουργικότητα των μεταδεδομένων τους (Cho, 2014). Τα τελευταία χρόνια, παρατηρήθηκε η στροφή των διαχειριστών των αποθετηρίων προς την αξιολόγηση και την εφαρμογή επιχειρηματικών μοντέλων σχετικά με την ανάπτυξη των συλλογών (Barrueco & Termens, 2022).

Σύμφωνα με την Αμερικανική Ένωση Βιβλιοθηκών (όπως αναφέρεται στο Barrueco & Termens, 2022), τα τρία βασικά χαρακτηριστικά της ψηφιακής διατήρησης είναι οι πολιτικές, οι στρατηγικές και οι δράσεις που υιοθετούν. Οι πολιτικές αντικατοπτρίζουν τη δέσμευση των ιδρυμάτων ως προς τη μακροχρόνια διατήρηση του υλικού τους, ενώ οι στρατηγικές και οι δράσεις αποδεικνύουν την πρακτική εφαρμογή των πολιτικών στη ροή εργασιών ενός αποθετηρίου (Barrueco & Termens, 2022). Παράλληλα, σημαντικό στοιχείο ως προς την αξιοπιστία ενός αποθετηρίου, αποτελεί η χρήση καθιερωμένων προτύπων στις συλλογές του.

Συμπερασματικά, ο σχεδιασμός και η υλοποίηση μίας στρατηγικής ψηφιακής διατήρησης θα πρέπει να εφαρμόζεται από το αρχικό στάδιο της δημιουργίας ενός ψηφιακού αντικειμένου και να βασίζεται στη διατήρηση των μεταδεδομένων του, στα μορφότυπα των τεκμηρίων, και παράλληλα στην τεχνολογία, τα συστήματα και τα πρότυπα που χρησιμοποιεί το αποθετήριο (Barrueco & Termens, 2022).

Στην Ινδονησία πραγματοποιήθηκε έρευνα σε δύο Πανεπιστημιακές βιβλιοθήκες σχετικά με το ποιες στρατηγικές ψηφιακής διατήρησης χρησιμοποιούνται. Η έρευνα απέδειξε ότι, και στους δύο οργανισμούς κάθε μήνα διατηρούνται αντίγραφα ασφαλείας των συλλογών σε έναν ειδικό υπολογιστή ή σε εξωτερικά μέσα, που χρησιμοποιούνται μόνο για την αποθήκευση των αρχείων, και σε έναν διακομιστή που βρίσκεται σε έναν άλλον χώρο από αυτόν της βιβλιοθήκης (Rachman, 2019).

Σε μία άλλη μελέτη με τη συμμετοχή 24 ιδρυματικών αποθετηρίων της Αφρικής, αποδείχτηκε ότι η πλειοψηφία των αποθετηρίων χρησιμοποιούν την στρατηγική της μετανάστευσης, τα 7 από τα 24 αποθετήρια προτιμούν την ενθυλάκωση ή την ανανέωση του υλικού, ενώ μόλις τα 6 αποθετήρια ανέφεραν την προσομοίωση (Anyaoaku et al., 2019).

Αντιθέτως, σε έρευνα που πραγματοποιήθηκε στους ιστοτόπους ιδρυματικών αποθετηρίων της Βραζιλίας αποκαλύφθηκε ότι δεν υπάρχουν δημοσιευμένες πολιτικές ψηφιακής διατήρησης στα Πανεπιστήμια της χώρας, παρόλο που ορισμένα από αυτά δήλωσαν την πρόθεσή τους να διατηρήσουν την συλλογή των αποθετηρίων τους (da Silva & Borges, 2017).

2.11 Η περίπτωση της Συλλογής Κάλλιπος

Το ψηφιακό Αποθετήριο Κάλλιπος δημιουργήθηκε από τη Δράση «Ελληνικά Ακαδημαϊκά Ηλεκτρονικά Συγγράμματα και Βοηθήματα», η οποία ξεκίνησε το 2013 και ολοκληρώθηκε το 2015. Ο απώτερος σκοπός της Δράσης ήταν η δημιουργία ανοιχτών ηλεκτρονικών ακαδημαϊκών συγγραμμάτων με ταυτόχρονη επιδίωξη τη συστηματική καταγραφή, οργάνωση και μακροχρόνια διατήρηση συγγραμμάτων και μαθησιακών αντικειμένων, που

έχει ήδη παραγάγει και που θα παράγει στο μέλλον η ελληνική Ακαδημαϊκή Κοινότητα (Κουής, Χριστάκη, Κουτσιλέου, κ.ά., 2016, σ. 17). Αποτελεί «μια πρωτοποριακή, ολοκληρωμένη Δράση του Συνδέσμου Ελληνικών Ακαδημαϊκών Βιβλιοθηκών (ΣΕΑΒ), σε συνεργασία με το Εθνικό Δίκτυο Έρευνας και Τεχνολογίας (ΕΔΕΤ), που ήδη, με την ολοκλήρωση του πιλοτικού σταδίου της, έχει παραγάγει περισσότερα από 520 ακαδημαϊκά ηλεκτρονικά συγγράμματα, σύγχρονα, αξιολογημένα, ελληνόγλωσσα και δημιουργημένα από το ένα τρίτο, περίπου, των Ελλήνων πανεπιστημιακών» (Κουής, Χριστάκη, Κουτσιλέου, κ.ά., 2016, σ. 12).

Οι εκδόσεις της Δράσης «Κάλλιπος» φιλοξενούν πέντε συλλογές κατηγοριοποιημένες ανά γνωστικό πεδίο:

- Ανθρωπιστικές και Νομικές Επιστήμες, Επιστήμες Τεχνών και Γραμμάτων
- Επιστήμες Μηχανικών και Πληροφορική
- Οικονομικές, Πολιτικές, Κοινωνικές και Γεωπονικές Επιστήμες
- Ιατρικές Επιστήμες και Επιστήμες της Ζωής
- Φυσικές Επιστήμες.

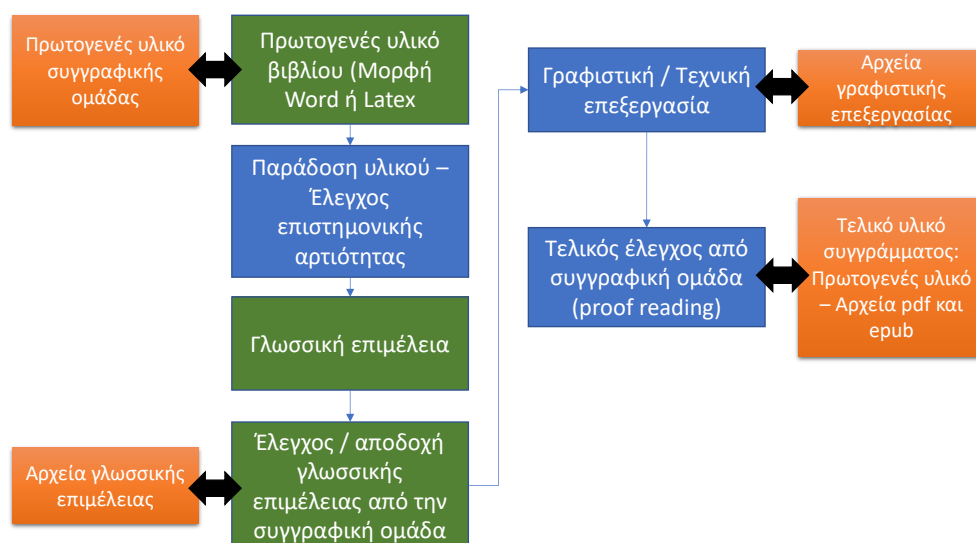
Η πρώτη φάση του Κάλλιπου εμπεριείχε συνολικά 517 επιστημονικά συγγράμματα. Η σύνθεση των βιβλίων ποίκιλλε ανάλογα με το επιστημονικό πεδίο που διαπραγματευόταν. Τα τεκμήρια που είχαν αντικείμενο τα μαθηματικά, την φυσική, την χημεία και την μηχανική περιείχαν πολλές εξισώσεις και πίνακες, ενώ παράλληλα αποτελούνταν και από διαδραστικό υλικό, όπως ασκήσεις, υπερσυνδέσεις με ιστοτόπους του διαδικτύου κ.ά. Σε αντίθεση, τα πολιτικά, τα κοινωνικά και όσα εμπίπτουν στις ανθρωπιστικές επιστήμες αποτελούνταν κυρίως από απλό κείμενο και σε ορισμένες περιπτώσεις από εικόνες ή πίνακες. Βιβλία που ανήκαν στον τομέα της πληροφορικής, του προγραμματισμού, της τεχνολογίας, της ιατρικής ή των γλωσσικών επιστημών συμπεριλάμβαναν στο περιεχόμενο τους και αρχεία βίντεο ή ήχου. Το πρωτογενές υλικό των συγγραμμάτων στην πλειοψηφία του ήταν σε μορφή MSWord με ή χωρίς Adobe InDesign ή σε Latex. Σε λίγες περιπτώσεις δεν υπήρχε πρωτογενές αρχείο ή υπήρχε άλλη μορφή, όπως MD Markup Language. Το τελικό αρχείο που διατίθεται μέσω του Αποθετηρίου για τους τελικούς χρήστες είναι κατά κύριο λόγο σε μορφή PDF, ενώ ορισμένα τεκμήρια διαθέτουν και μορφή EPUB ή HTML.

Κεφάλαιο 3. Μεθοδολογία ή/και Σχέδιο εργασιών–Υλοποίηση – Εφαρμογή

Το κεφάλαιο αυτό ασχολείται με τις ροές εργασιών που ακολουθούνται για τη συγγραφή και την ανανέωση ηλεκτρονικών συγγραμμάτων. Στη συνέχεια αναλύεται η μεθοδολογία που εφαρμόστηκε για τον προσδιορισμό της απαιτούμενης ανθρωποπροσπάθειας για την επιμέλεια των βιβλίων της συλλογής του «Κάλλιπου» και διερευνώνται τα στοιχεία που κοστίζουν για την υλοποίηση της ανανέωσης της συλλογής.

3.1 Διαδικασία συγγραφής και ανανέωσης συγγραμμάτων συλλογής Κάλλιπου

Η αρχική διαδικασία της δημιουργίας ενός συγγράμματος ξεκινάει από τον συγγραφέα του τεκμηρίου (βλέπε ροή εργασίας στην Εικόνα 5). Αρχικά, οι συγγραφείς προετοιμάζουν το σύγγραμμα τους σε ένα αρχείο Word (.docX) ή LATEX, σύμφωνα με τις οδηγίες μορφοποίησης που έχουν λάβει από την ομάδα του Κάλλιπου. Αφού ολοκληρωθεί η διαδικασία της συγγραφής το κείμενο στέλνεται για να ελεγχθεί και να αξιολογηθεί από σχετικούς με το θέμα ακαδημαϊκούς – έλεγχος επιστημονικής αρτιότητας. Στη συνέχεια και αφού ολοκληρωθεί η διαδικασία της συγγραφής και αξιολόγησης, το κείμενο παραδίδεται στους γλωσσικούς επιμελητές για τυχόν διορθώσεις. Στο στάδιο της γλωσσικής επιμέλειας δίνεται βαρύτητα στην ορθογραφία, την στίξη, την γραμματική και την σύνταξη του περιεχομένου, καθώς επίσης ελέγχεται η γλώσσα των επιστημονικών όρων και η δομή της βιβλιογραφίας.



Εικόνα 5: Ροή διαδικασίας συγγραφής ενός συγγράμματος

Στη συνέχεια, και αφού το γλωσσικά επιμελημένο αρχείο ελεγχθεί από τη συγγραφική ομάδα τίθεται σε γραφιστική / τεχνική επεξεργασία από τους γραφίστες. Η γραφιστική και τεχνική επεξεργασία αφορά την εικαστική μορφή και τα «τεχνικά» μέρη του βιβλίου, όπως είναι η μορφοποίηση του κειμένου, τα λογότυπα, οι εικόνες, τα σχήματα, οι πίνακες, οι εξισώσεις, τα βίντεο, τα ηχητικά αρχεία και το διαδραστικό υλικό. Για τη γραφιστική επιμέλεια απαιτείται η τήρηση της μορφοποίησης του κειμένου με βάση το πρότυπο συγγραφής του Κάλλιπου. Οφείλεται να τηρείται το μέγεθος της σελίδας, τα περιθώρια, το μέγεθος της γραμματοσειράς, το διάστιχο κ.ά. Παράλληλα, συστήνεται η εξαγωγή του αρχείου σε μορφή EPUB. Για την σωστή εξαγωγή του αρχείου είναι απαραίτητο τα στοιχεία του κειμένου να είναι χαρακτηρισμένα / επισημασμένα (marked up). Σύμφωνα με το πρότυπο συγγραφής, το κείμενο πρέπει να έχει ίδιο στυλ παραγράφων και χαρακτήρων, ώστε να αποτυπώνονται τα αντίστοιχα στυλ στον κώδικα HTML του EPUB, με συνέπεια να εξάγεται «καθαρός κώδικας». Αυτό είναι αρκετά βοηθητικό και για τους προγραμματιστές, καθώς σε περιπτώσεις μελλοντικών σφαλμάτων θα έχουν τη δυνατότητα να επέμβουν άμεσα στον κώδικα.

Μετά και τη γραφιστική επεξεργασία το κείμενο ελέγχεται για ακόμα μια φορά από τη συγγραφική ομάδα (proof reading) και στη συνέχεια παραδίδεται για ανάρτηση στο Αποθετήριο του Κάλλιπου.

Από την παραπάνω περιγραφή γίνεται κατανοητό ότι κατά τη διαδικασία δημιουργίας ενός συγγράμματος προκύπτουν μια σειρά από αρχεία που αντικατοπτρίζουν την κατάσταση του συγγράμματος σε διαφορετικές χρονικές στιγμές. Ενδεικτικά, το αρχικό σύγγραμμα αποτελεί το πρωτογενές υλικό της συγγραφικής ομάδας, το οποίο στη συνέχεια δίδεται στους γλωσσικούς επιμελητές και προκύπτει το γλωσσικά επιμελημένο αρχείο. Μετέπειτα, αφού ολοκληρωθεί η γραφιστική επιμέλεια, το αρχείο διαμορφώνεται σε μορφή επεξεργασίας από τον γραφίστα, ακολουθεί ο τελικός έλεγχος του συγγράμματος, όπου προκύπτει το τελικό αρχείο στο ανάλογο μορφότυπο κατά περίπτωση, και στο τέλος παράγεται η δημοσιεύσιμη μορφή σε PDF ή EPUB. Οι μορφές των αρχείων αυτών ανταποκρίνονται σε διαφορετικές ανάγκες κάθε φορά, ανάλογα σε ποιον απευθύνονται. Για παράδειγμα οι συγγραφείς προτιμούν τη μορφή σε Microsoft Word διότι είναι πιο εξοικειωμένοι σε αυτό το εργαλείο. Αντίστοιχα, οι γραφίστες «δουλεύουν» για την παραγωγή του τελικού συγγράμματος με εργαλεία όπως το Adobe In Design γιατί προσφέρει αυξημένες δυνατότητες διαμόρφωσης της εμφάνισης για δημοσίευση (desktop publishing).

Συγκεκριμένα στο παρακάτω σχήμα (Εικόνα 6) αποτυπώνεται η ροή των εργασιών τριών περιπτώσεων από τη διαδικασία δημιουργίας του συγγράμματος έως την τελική / δημοσιεύσιμη μορφή του. Υπάρχουν τρεις μορφές αρχείων που μπορεί να προκύψουν στο στάδιο της γραφιστικής επιμέλειας. Στην πρώτη περίπτωση το πρωτογενές υλικό του βιβλίου δημιουργείται σε μορφή Microsoft Word. Οι εργασίες της αξιολόγησης από τους ειδικούς και της γλωσσικής επιμέλειας διενεργούνται πάνω στα αρχεία Microsoft Word και διατηρούν την ίδια μορφή. Στο στάδιο της γραφιστικής επιμέλειας, οι γραφίστες εργάζονται και διαμορφώνουν το στυλ και τη δομή του κειμένου στη μορφή του Microsoft Word. Τέλος, το τελικό αρχείο αλλάζει μορφή και μετατρέπεται σε Adobe PDF.

Στην δεύτερη περίπτωση, έχουμε παρόμοια διαδικασία ροής εργασιών του εγγράφου σε μορφή Microsoft Word, έως την ανάληψη της επεξεργασίας από τους γραφίστες. Σε αυτή την περίπτωση, το πρωτογενές υλικό εισάγεται στο λογισμικό Adobe In Design και διαμορφώνεται μέσα από εξειδικευμένα εργαλεία εμφάνισης κειμένου. Η επεξεργασία με το συγκεκριμένο εργαλείο βελτιώνει τη διάταξη και τη σχεδίαση του κειμένου, αναπαριστά γραφικά υψηλής ποιότητας και προετοιμάζει το περιεχόμενο για το στάδιο της τελικής μορφής προς δημοσίευση σε Adobe PDF και EPUB.

Στην τρίτη περίπτωση, το πρωτογενές υλικό του βιβλίου έχει παραχθεί από τον συγγραφέα σε μορφή LATEX, το οποίο είναι μία γλώσσα δημιουργίας εγγράφων. Ειδικότερα, δεν αποτελεί εργαλείο επεξεργασίας κειμένου, αλλά προετοιμάζει το έγγραφο μέσω των ετικετών σήμανσης του περιεχομένου, ώστε να υποβληθεί προς επεξεργασία (The Latex project, χ.χ.). Συνεπώς, καθώς παρουσιάζει μεγάλες τεχνικές δυσκολίες η πρόσβαση στο κείμενο στην μορφή LATEX, από τους αξιολογητές και τους υπευθύνους της γλωσσικής επιμέλειας, το τεκμήριο εξάγεται σε μορφή Adobe PDF με σκοπό την ηλεκτρονική ανάγνωσή του και την περαιτέρω επεξεργασία του. Οι όποιες αλλαγές συμβούν στο κείμενο (είτε λόγω της αξιολόγησης ή της γλωσσικής επιμέλειας), αυτές σημειώνονται στο αρχείο PDF και εν συνεχεία δίδονται στην συγγραφική ομάδα για να προστεθούν στα πρωτογενή αρχεία LATEX. Σχετικά με την γραφιστική επιμέλεια του αρχείου, το βιβλίο παράγεται απευθείας σε PDF από τη συγγραφική ομάδα και παραδίδεται για δημοσίευση.

Πρωτογενές υλικό βιβλίου	Αξιολόγηση από ειδικούς	Γλωσσική επιμέλεια	Γραφιστική επιμέλεια	Τελική μορφή
Microsoft Word	Microsoft Word	Microsoft Word	Microsoft Word	Adobe PDF
Microsoft Word	Microsoft Word	Microsoft Word	Adobe In Design	Adobe PDF και Epub
LATEX	Adobe PDF	Adobe PDF	LATEX	Adobe PDF

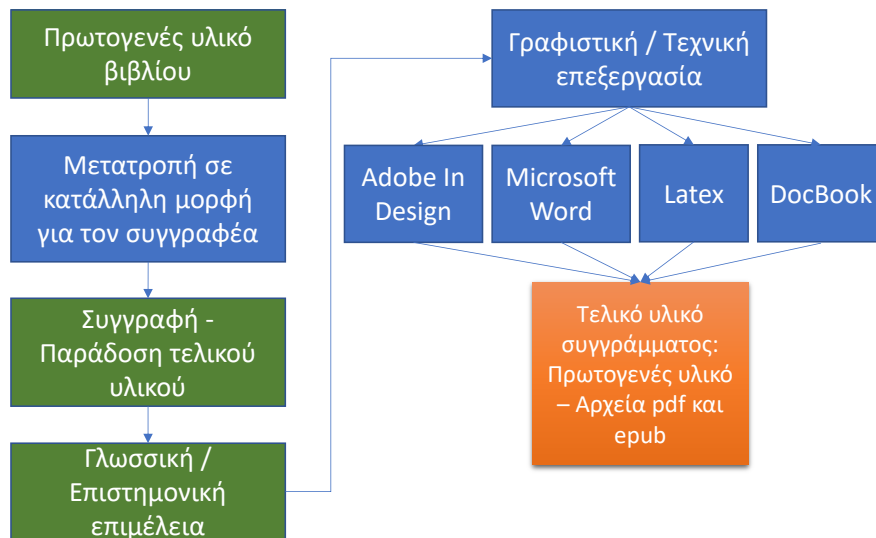
Εικόνα 6. Βασικές ροές εργασίας για την παραγωγή ενός συγγράμματος στον Κάλλιπο.

Στην περίπτωση που είναι απαραίτητη / επιθυμητή η ανανέωση / επικαιροποίηση των συγγραμμάτων, το πρωτογενές υλικό των βιβλίων θα πρέπει να επιστραφεί στους συγγραφείς για τυχόν αλλαγές, σε μορφή κατάλληλη προς επεξεργασία (βλέπε Εικόνα 7). Στην πρώτη περίπτωση, η διαδικασία επιστροφής του πρωτογενούς υλικού διενεργείται χωρίς ιδιαίτερες δυσκολίες, καθώς σε όλες τις φάσεις επεξεργασίας η μορφή του κειμένου παρέμεινε η ίδια με την μορφή που δημιούργησε ο συγγραφέας.

Αντιθέτως, η δεύτερη περίπτωση ροής εργασιών παρουσιάζει αυξημένη πολυπλοκότητα σε ότι αφορά την επιστροφή του πρωτογενούς υλικού στους συγγραφείς, διότι οι τελευταίοι δεν μπορούν να επεξεργαστούν το αρχείο σε μορφή Adobe In Design. Αυτό που επιβάλλεται να γίνει από τη γραφιστική ομάδα είναι να ανατρέξει στους φακέλους του πρωτογενούς υλικού του συγγράμματος και να μετατρέψει / εξάγει το περιεχόμενο του Adobe In Design σε αρχείο σε μορφή Microsoft Word. Εναλλακτικά, θα μπορούσε να δοθεί στην συγγραφική ομάδα το αρχείο μετά την γλωσσική επιμέλεια, το οποίο όμως ενδέχεται να μην περιέχει τις τροποποιήσεις που πραγματοποιήθηκαν στα επόμενα βήματα.

Η τρίτη περίπτωση αποτελεί μία σχετικά διαχειρίσιμη κατάσταση, γιατί το γραφιστικά επιμελημένο αρχείο είναι στην ίδια μορφή με αυτή που παρέδωσαν οι συγγραφείς, δηλαδή σε LATEX, επομένως καθίσταται εύχρηστο προς εκείνους για αλλαγές.

Οι συγγραφείς αφού επανεξετάσουν το αρχικό κείμενο μπορούν να προβούν σε επικαιροποίηση στοιχείων και δεδομένων ή αναθεωρήσεις απόψεων και συμπερασμάτων του συγγράμματος.

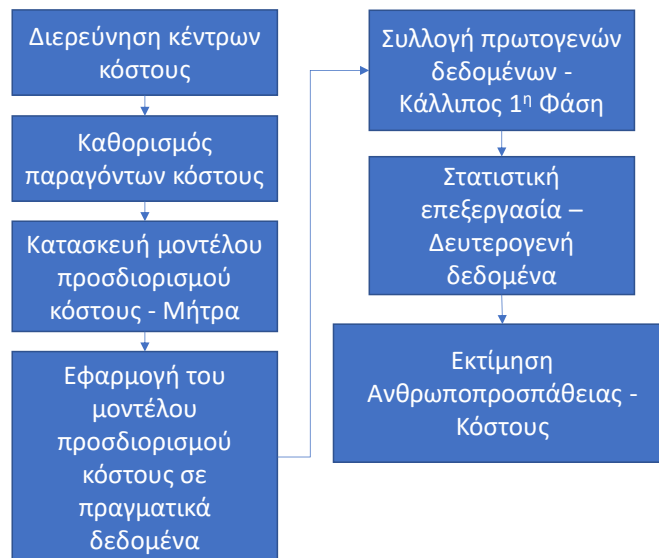


Εικόνα 7. Ανανέωση συγγράμματος

Σημειώνεται, ότι στην νέα φάση του Κάλλιπου, προστέθηκε μια νέα επιλογή για την ανάπτυξη των βιβλίων που είναι το DocBook. Σύμφωνα με τον Walsh (2010), το DocBook παρέχει ένα σύστημα για τη σύνταξη δομημένων εγγράφων το οποίο είναι βασισμένο πάνω στα πρότυπα του σχήματος XML. Υποστηρίζεται από πληθώρα εργαλείων, ανοιχτού κώδικα ή κώδικα εμπορικής χρήσης, με πιο πρόσφατη την έκδοση 5.0. Περισσότερες λεπτομέρειες παρουσιάστηκαν στο Κεφάλαιο 2.

3.2 Μεθοδολογία προσδιορισμού ανθρωποπροσπάθειας - κόστους

Όπως έχει ήδη ειπωθεί, βασικό ζητούμενο στην παρούσα πτυχιακή εργασία είναι να βρεθεί ένα μοντέλο / μια μεθοδολογία για τον προσδιορισμό της απαιτούμενης ανθρωποπροσπάθειας (επομένως και του κόστους) που δημιουργούν οι γραφιστικές εργασίες τόσο για την αρχική παραγωγή ενός συγγράμματος όσο και για την ανανέωσή του. Τα βασικά βήματα που ακολουθήθηκαν αποτυπώνονται στο σχήμα που ακολουθεί (Εικόνα 8).



Εικόνα 8. Μεθοδολογία για τη δημιουργία μοντέλου προσδιορισμού ανθρωποπροσπάθειας – κόστους – Εφαρμογή μοντέλου

Για τη μελέτη και την κοστολόγηση της ανανέωσης των βιβλίων της πρώτης φάσης του Κάλλιπου καθοριστικό ρόλο είχε η διερεύνηση των κέντρων κόστους. Ως κέντρα κόστους ορίζονται, οι περιοχές / στοιχεία που απαιτούν την περισσότερη εργασία εκ μέρους των γραφιστών και κατ' επέκταση δημιουργούν το μεγαλύτερο κόστος. Όπως έχει ήδη αναφερθεί η εκτίμηση του κόστους ανανέωσης όλης ή μέρος της συλλογής του Κάλλιπου (1η φάση) είναι απαραίτητο βήμα για τον προγραμματισμό των εργασιών και τον καθορισμό του προϋπολογισμού ενός μελλοντικού έργου.

Αρχικά, πραγματοποιήθηκε η αναλυτική καταγραφή των στοιχείων που περιέχονται στα βιβλία, η οποία είναι πρωταρχικής σημασίας για τον προσδιορισμό του υλικού που πρέπει να επεξεργασθεί και για τον καθορισμό του μεγέθους της ανθρωποπροσπάθειας που θα χρειαστεί. Σε πρώτο στάδιο, συγκεντρώθηκε όλο το πρωτογενές υλικό και εφαρμόστηκε η αναλυτική καταγραφή των στοιχείων που περιλαμβάνονται στα βιβλία της 1ης φάσης του Κάλλιπου. Η δομή που ακολουθήθηκε για τη διερεύνηση αποτελούνταν από:

- Το αναγνωριστικό **ID** του κάθε τεκμηρίου,
- Τον **φάκελο** στον οποίο περιέχεται το υλικό κάθε συγγράμματος,
- Τον **τίτλο** του βιβλίου,
- Το **σύνδεσμο (link)** που οδηγεί στην εγγραφή του ηλεκτρονικού βιβλίου στο αποθετήριο Κάλλιπος,
- Τις **σελίδες** που αποτελείται το βιβλίο, δίνοντας μια πρώτη εικόνα του μεγέθους του,

- Τις μορφές που μπορεί να έχει δημοσιευθεί, δηλαδή αν διατίθεται σε μορφή **HTML ή EPUB** μέσω του αποθετηρίου, οι οποίες διευκολύνουν την ανάγνωση του κειμένου σε διάφορες συσκευές,
- Το **σύνολο των λέξεων** του κειμένου,
- Το **σύνολο των χαρακτήρων χωρίς τα κενά**,
- Τον **αριθμό των κεφαλαίων** που απαρτίζουν το βιβλίο,
- Την **μορφή** του αρχείου που έχει δημιουργηθεί από τον **συγγραφέα**, δηλαδή το μορφότυπο του πρωτογενούς υλικού πριν την επεξεργασία της επιμέλειας (π.χ. MSWord, LATEX, Adobe In Design με Microsoft Word ή χωρίς αρχείο Microsoft Word και σπάνια σε κάποια άλλη μορφή),
- Την **μορφή** του αρχείου που έχει διαμορφωθεί από τον **γραφίστα** (με ή χωρίς Adobe Indesign, LATEX),
- Το **σύνολο των εικόνων και των σχημάτων** που περιέχονται στο βιβλίο,
- Το **σύνολο των πινάκων**,
- Το **σύνολο των εξισώσεων**,
- Το **σύνολο των βίντεο**,
- Το **σύνολο των αρχείων ήχου**,
- Τα **είδη των διαδραστικών στοιχείων** που περιέχονται (π.χ. Υπερσύνδεσμοι, JavaScript, γλώσσα HTML, κ.ά)
- Τον **χαρακτηρισμό δυσκολίας** του κάθε τεκμηρίου, ο οποίος διαχωρίστηκε σε τρία επίπεδα, απλό/ σύνθετο/ περίπλοκο,
- Και τέλος το πεδίο **παρατηρήσεις**, όπου περιγράφονται τυχόν ιδιαιτερότητες των βιβλίων.

Σχετικά με το σύνολο των εικόνων / σχημάτων / πινάκων/ εξισώσεων / βίντεο / αρχείων ήχου, δεν κατέστη δυνατή, σε ορισμένα βιβλία, η συγκέντρωση του ακριβούς αριθμού, εφόσον δεν τα περιείχε ο φάκελος του υλικού (ήταν απευθείας ενσωματωμένα στο αρχείο Microsoft Word ή LATEX). Η πρακτική που ακολουθήθηκε σε αυτές τις περιπτώσεις, καθώς δεν ήταν εφικτό να υπολογιστούν με ακρίβεια τα στοιχεία, ήταν ο υπολογισμός κατά προσέγγιση της συχνότητας που εμφανίζονταν στο σύγγραμμα ανά 100 σελίδες. Στον Πίνακα 1, παρουσιάζονται οι βαθμίδες που διαχωρίστηκαν τα στοιχεία σύμφωνα με τη συχνότητα εμφάνισής τους μέσα στο κείμενο.

Σχήματα / Εικόνες	Αριθμός (ανά 100 σελίδες)	Πίνακες	Αριθμός	Εξισώσεις	Αριθμός	Βιβλιογραφία	Αριθμός
Καθόλου - Λίγο	1 ως 5	Καθόλου - Λίγο	1 ως 5	Καθόλου - Λίγο	1 ως 10	Καθόλου - Λίγο	ως 30
Αρκετές	6 ως 10	Αρκετούς	6 ως 10	Αρκετές	11 ως 30	Αρκετές	31 ως 80
Πολλές	11 ως 40	Πολλούς	11 ως 40	Πολλές	31 ως 100	Πολλές	81 ως 150
Πάρα πολλές	41 ως 70	Πάρα πολλούς	41 ως 70	Πάρα πολλές	101 και άνω	Πάρα πολλές	151 και άνω

Πίνακας 1: Υπολογισμός κατά προσέγγιση της συχνότητας εμφάνισης των στοιχείων

Σε δεύτερο στάδιο, επιλέχθηκαν οι παράγοντες των κέντρων κόστους. Μελετήθηκε ποια στοιχεία απαιτούν περισσότερη επεξεργασία με βάση τον χρόνο και επακολούθως κοστίζουν περισσότερο επιβαρύνοντας τον τελικό προϋπολογισμό. Οι παράγοντες που καθορίστηκαν για την κοστολόγηση ήταν:

- Σχήματα/ Εικόνες
- Πίνακες
- Εξισώσεις
- Βιβλιογραφικές αναφορές
- Σελίδες κειμένου
- Βίντεο
- Ήχος
- Διαδραστικό υλικό
- Μορφότυπο

Πρωταρχικά, για την δημιουργία του μοντέλου ζητήθηκε η εκτίμηση τριών διαφορετικών γραφιστών σχετικά με το χρόνο (σε λεπτά) που απαιτείται για την επεξεργασία των στοιχείων ανά μονάδα και συγκεκριμένα:

- ανά εικόνα / σχήμα,
- ανά πίνακα,
- ανά εξίσωση,
- ανά βιβλιογραφική αναφορά,
- και για την επεξεργασία 100 σελίδων απλού κειμένου.

Οι τιμές που συγκεντρώθηκαν από τους τρεις γραφίστες ανά στοιχείο, σε λεπτά, ήταν οι εξής:

	Γραφίστας 1	Γραφίστας 2	Γραφίστας 3
Ανά εικόνα / σχήμα	1,50	1,50	2,00
Ανά πίνακα	8,00	4,00	8,00
Ανά εξίσωση	1,50	1,00	1,50
Ανά βιβλιογραφία	1,00	0,60	0,50

Ανά 100 σελίδες κειμένου	300	200	350
--------------------------	-----	-----	-----

Πίνακας 2: Εκτίμηση χρόνου σε λεπτά

Αφού συγκεντρώθηκαν οι εκτιμήσεις χρόνου επεξεργασίας των γραφιστών, υπολογίστηκαν τρεις βαθμίδες χρόνου ανά περίπτωση, δηλαδή ο μέσος όρος, ο μέγιστος και ο ελάχιστος χρόνος επεξεργασίας των στοιχείων. Οι τιμές αποτυπώνονται στον παρακάτω πίνακα:

Χρόνος απαιτούμενης προσπάθειας / ανά κατηγορία υλικού	Χρόνος ανά εικόνα	Χρόνος ανά πίνακα	Χρόνος ανά εξίσωση	Χρόνος ανά αναφορά	Χρόνος ανά 100 σελίδες κειμένου
Μέσος όρος	1,67	6,67	1,33	0,70	285
Μέγιστη τιμή	2,00	8,00	1,50	1,00	350
Ελάχιστη τιμή	1,50	4,00	1,00	0,50	200

Πίνακας 3: Χρόνος απαιτούμενης προσπάθειας / ανά κατηγορία υλικού

Καθώς οι προσεγγιστικές τιμές στοιχείων όπως οι εικόνες, οι πίνακες κ.λπ. ανά 100 σελίδες αφορούσαν εύρη τιμών, ήταν απαραίτητος ο προσδιορισμός μιας μέσης τιμής ανά περίπτωση. Για την διευκόλυνση των υπολογισμών αποφασίστηκαν οι αντιστοιχίσεις του πίνακα που ακολουθεί.

Σχήματα / Εικόνες	Μέση τιμή	Πίνακες	Μέση τιμή	Εξισώσεις	Αριθμός	Βιβλιογραφία	Αριθμός
Καθόλου - Λίγο	3	Καθόλου - Λίγο	3	Καθόλου - Λίγο	6	Καθόλου - Λίγο	25
Αρκετές	8	Αρκετούς	8	Αρκετές	25	Αρκετές	70
Πολλές	30	Πολλούς	30	Πολλές	80	Πολλές	140
Πάρα πολλές	60	Πάρα πολλούς	60	Πάρα πολλές	150	Πάρα πολλές	200

Πίνακας 4: Μέση τιμή ανά κατηγορία συχνότητας στοιχείων

3.3 Υπολογισμός ανθρωποπροσπάθειας για κάθε κατηγορία υλικού

Στους πίνακες που ακολουθούν παρατίθενται οι υπολογισμοί με βάση τους οποίους προκύπτουν οι τιμές ανθρωποπροσπάθειας για κάθε κατηγορία υλικού.

3.3.1 Χρόνος εργασίας για εικόνες / σχήματα

Συχνότητα εμφάνισης	Μέση τιμή συχνότητας εμφάνισης	Μέσος όρος διάρκειας (1,67)	Μέγιστη τιμή	Ελάχιστη Τιμή
Ποιοτικός χαρακτηρισμός	Αριθμός εικόνων	1,67*	2,00	1,50
Καθόλου - Λίγο	3	5,01	6,00	4,50
Αρκετές	8	13,36	16,00	12,00
Πολλές	30	50,10	60,00	45,00
Πάρα πολλές	60	100,20	120,00	90,00

Πίνακας 5: Χρόνος εργασίας που απαιτείται για εικόνες / σχήματα

*Παράδειγμα υπολογισμού: υπολογίστηκε ένας μέσος αριθμός εικόνων / σχημάτων ανά κατηγορία επί της υπολογισμένης μέσης τιμής της γραφιστικής επεξεργασίας.

- Καθόλου – Λίγο: 3 (εικόνες / σχήματα) * 1,67 (λεπτά) = 5,01 (λεπτά)
- Αρκετές: 8 (εικόνες / σχήματα) * 1,67 (λεπτά) = 13,36 (λεπτά)
- Πολλές: 30 (εικόνες / σχήματα) * 1,67 (λεπτά) = 50,10 (λεπτά)
- Πάρα πολλές: 60 (εικόνες / σχήματα) * 1,67 (λεπτά) = 100,20 (λεπτά)

3.3.2 Χρόνος εργασίας για πίνακες

Συχνότητα εμφάνισης	Μέση τιμή συχνότητας εμφάνισης	Μέσος όρος διάρκειας (6,67)	Μέγιστη τιμή	Ελάχιστη Τιμή
Ποιοτικός χαρακτηρισμός	Αριθμός πινάκων	6,67*	8,00	4,00
Καθόλου - Λίγο	3	20,01	24,00	12,00
Αρκετούς	8	53,36	64,00	32,00
Πολλούς	30	200,10	240,00	120,00
Πάρα πολλούς	60	400,20	480,00	240,00

Πίνακας 6: Χρόνος εργασίας που απαιτείται για τους πίνακες

*Παράδειγμα υπολογισμού: υπολογίστηκε ένας μέσος αριθμός πινάκων ανά κατηγορία επί της υπολογισμένης μέσης τιμής της γραφιστικής επεξεργασίας.

- Καθόλου – Λίγο: 3 (πίνακες) * 6,67 (λεπτά) = 20,01 (λεπτά)
- Αρκετούς: 8 (πίνακες) * 6,67 (λεπτά) = 53,36 (λεπτά)
- Πολλούς: 30 (πίνακες) * 6,67 (λεπτά) = 200,10 (λεπτά)
- Πάρα πολλούς: 60 (πίνακες) * 6,67 (λεπτά) = 400,20 (λεπτά)

3.3.3 Χρόνος εργασίας για εξισώσεις

Συχνότητα εμφάνισης	Μέση τιμή συχνότητας εμφάνισης	Μέσος όρος διάρκειας (1,33)	Μέγιστη τιμή	Ελάχιστη Τιμή
Ποιοτικός χαρακτηρισμός	Αριθμός εξισώσεων	1,33*	1,50	1,00
Καθόλου - Λίγο	6	7,98	9,00	6,00
Αρκετές	25	33,25	37,50	25,00
Πολλές	80	106,40	120,00	80,00
Πάρα πολλές	150	199,50	225,00	150,00

Πίνακας 7: Χρόνος εργασίας που απαιτείται για τις εξισώσεις

*Παράδειγμα υπολογισμού: υπολογίστηκε ένας μέσος αριθμός εξισώσεων ανά κατηγορία επί της υπολογισμένης μέσης τιμής της γραφιστικής επεξεργασίας.

- Καθόλου – Λίγο: 6 (εξισώσεις) * 1,33 (λεπτά) = 7,98 (λεπτά)
- Αρκετές: 25 (εξισώσεις) * 1,33 (λεπτά) = 33,25 (λεπτά)
- Πολλές: 80 (εξισώσεις) * 1,33 (λεπτά) = 106,40 (λεπτά)
- Πάρα πολλές: 150 (εξισώσεις) * 1,33 (λεπτά) = 199,50 (λεπτά)

3.3.4 Χρόνος εργασίας για βιβλιογραφικές αναφορές

Συχνότητα εμφάνισης	Μέση τιμή συχνότητας εμφάνισης	Μέσος όρος διάρκειας (0,70)	Μέγιστη τιμή	Ελάχιστη Τιμή
Ποιοτικός χαρακτηρισμός	Αριθμός βιβλιογραφικών αναφορών	0,70*	1,00	0,50
Καθόλου - Λίγο	25	17,50	25,00	12,50
Αρκετές	70	49,00	70,00	35,00
Πολλές	140	98,00	140,00	70,00
Πάρα πολλές	200	140,00	200,00	100,00

Πίνακας 8: Χρόνος εργασίας που απαιτείται για τις βιβλιογραφικές αναφορές

*Παράδειγμα υπολογισμού: υπολογίστηκε ένας μέσος αριθμός βιβλιογραφικών αναφορών ανά κατηγορία επί της υπολογισμένης μέσης τιμής της γραφιστικής επεξεργασίας.

- Καθόλου – Λίγο: 25 (βιβλιογραφικές αναφορές) * 0,70 (λεπτά) = 17,50 (λεπτά)
- Αρκετές: 70 (βιβλιογραφικές αναφορές) * 0,70 (λεπτά) = 49,00 (λεπτά)
- Πολλές: 140 (βιβλιογραφικές αναφορές) * 0,70 (λεπτά) = 98,00 (λεπτά)
- Πάρα πολλές: 200 (βιβλιογραφικές αναφορές) * 0,70 (λεπτά) = 140,00 (λεπτά).

Αφότου ολοκληρώθηκε η δημιουργία του μοντέλου σχετικά με τον χρόνο επεξεργασίας και τη συχνότητα των στοιχείων μέσα στο κείμενο, διερευνήθηκαν επιπλέον παράγοντες που απαιτούσαν επιπλέον προσπάθεια και επομένως επιβάρυναν το τελικό κόστος. Τα ποσοστά εξήχθησαν προσεγγιστικά, με βάση την πολυπλοκότητα της επεξεργασίας τους (βλέπε Πίνακας 9). Ενδεικτικά, το αρχείο Word, με μηδενικό ποσοστό δυσκολίας, καθίσταται αρκετά εύχρηστο και μπορεί εύκολα να υποστεί οποιαδήποτε αλλαγή ή επεξεργασία, αντιθέτως το αρχείο Latex παρουσιάζει προβλήματα και δυσλειτουργίες κατά την χρήση του, γι' αυτό και το ποσοστό δυσκολίας της επεξεργασίας του ανέρχεται σε 50%. Τα αρχεία Adobe In Design με αρχείο Word (20%) και χωρίς αρχείο Word (30%) έχουν απόκλιση ποσοστού δυσκολίας 10%, διότι όταν δεν διατίθεται Word η επεξεργασία του τεκμηρίου καθίσταται δυσκολότερη καθώς το Adobe In Design δεν είναι εύκολα διαχειρίσιμο για τους συγγραφείς. Τα βίντεο και το διαδραστικό υλικό υπολογίστηκαν με βαθμό δυσκολίας 15% και τα αρχεία ήχου με 20%, διότι απαιτούν επιπλέον χρόνο επεξεργασίας. Τα ποσοστά των πρόσθετων παραγόντων αποτυπώνονται στον επόμενο πίνακα:

Μορφότυπο / Τελικό αρχείο				Λοιποί παράγοντες		
Word	Adobe In Design-Χωρίς Word	Adobe In Design-Me Word	Latex	Βίντεο	Ήχος	Διαδραστικό υλικό

0%	30%	20%	50%	15%	20%	15%
----	-----	-----	-----	-----	-----	-----

Πίνακας 9: Χρονική επιβάρυνση λοιπών παραγόντων

3.3.5 Ώρες και ημέρες

Η διαδικασία του υπολογισμού τόσο των ημερών όσο και των ωρών της επεξεργασίας κάθε βιβλίου, προέκυψε από την εισαγωγή των δεδομένων στο μοντέλο. Η συγκεκριμένη διαδικασία ρυθμίστηκε βάσει μίας σειράς συναρτήσεων, συμπεριλαμβανόμενων όλων των προαναφερθέντων δεδομένων, με απώτερο σκοπό την εξαγωγή αποτελεσμάτων του συνολικού χρόνου (σε εργατώρες). Η συνάρτηση απαρτίστηκε από τα πεδία και τις τιμές της μήτρας κόστους, και ανάλογα με τα δεδομένα που εισάγονταν στο μοντέλο εξαγόταν ο απαιτούμενος χρόνος.

Αφού ολοκληρώθηκε η μήτρα προσδιορισμού του κόστους, το μοντέλο τροφοδοτήθηκε με το σύνολο των βιβλίων της 1^{ης} φάσης του Αποθετηρίου «Κάλλιπος», σύμφωνα με τις τιμές που αποτυπώνονται στον Πίνακας 10:

Τελικό αρχείο	Word	AdobeInDesign- Xword	AdobeInDesign- Mword	Latex	Άλλο
Βίντεο	Ναι			Όχι	
Ήχος	Ναι			Όχι	
Διάδραση	Ναι			Όχι	
Εικόνες/ Σχήματα	Καθόλου -Λίγο	Αρκετές		Πολλές	Πάρα πολλές
Πίνακες	Καθόλου -Λίγο	Αρκετούς		Πολλούς	Πάρα πολλούς
Εξισώσεις	Καθόλου -Λίγο	Αρκετές		Πολλές	Πάρα πολλές
Βιβλιογραφία	Καθόλου -Λίγο	Αρκετές		Πολλές	Πάρα πολλές

Πίνακας 10: Ομαδοποίηση στοιχείων για την τροφοδότηση του μοντέλου

3.4 Μήτρα κόστους και παράδειγμα

Η τελική διαμόρφωση της μήτρας κόστους απέκτησε την παρακάτω διάρθρωση:

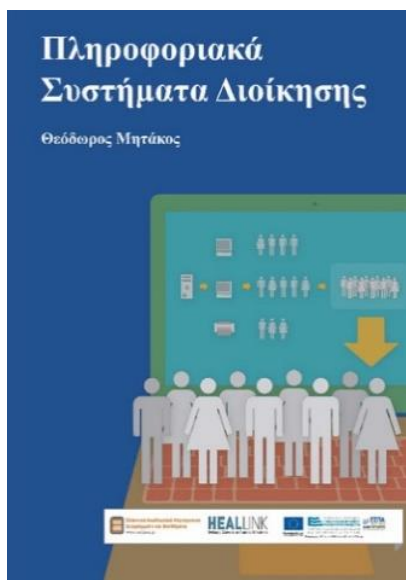
Μέση τιμή	Σχήματα / Εικόνες	Αριθμός (ανά 100 σελίδες)	Χρόνος σε λεπτά	Πίνακες	Αριθμός (ανά 100 σελίδες)	Χρόνος σε λεπτά	Εξισώσεις	Αριθμός (ανά 100 σελίδες)	Χρόνος σε λεπτά	Βιβλιογραφία	Αριθμός (ανά 100 σελίδες)	Χρόνος σε λεπτά
	Καθόλου - Λίγο	1 ως 5	5,01	Καθόλου - Λίγο	1 ως 5	20,01	Καθόλου - Λίγο	1 ως 10	7,98	Καθόλου - Λίγο	ως 30	17,50
	Αρκετές	6 ως 10	13,36	Αρκετούς	6 ως 10	53,36	Αρκετές	11 ως 30	33,25	Αρκετές	31 ως 80	49,00
	Πολλές	11 ως 40	50,10	Πολλούς	11 ως 40	200,10	Πολλές	31 ως 100	106,40	Πολλές	81 ως 150	98,00
	Πάρα πολλές	41 ως 70	100,20	Πάρα πολλούς	41 ως 70	400,20	Πάρα πολλές	101 και άνω	199,50	Πάρα πολλές	151 και άνω	140,00
Μέγιστη τιμή	Σχήματα / Εικόνες	Αριθμός (ανά 100 σελίδες)	Χρόνος σε λεπτά	Πίνακες	Αριθμός (ανά 100 σελίδες)	Χρόνος σε λεπτά	Εξισώσεις	Αριθμός (ανά 100 σελίδες)	Χρόνος σε λεπτά	Βιβλιογραφία	Αριθμός (ανά 100 σελίδες)	Χρόνος σε λεπτά
	Καθόλου - Λίγο	1 ως 5	6,00	Καθόλου - Λίγο	1 ως 5	24,00	Καθόλου - Λίγο	1 ως 10	9,00	Καθόλου - Λίγο	ως 30	25,00
	Αρκετές	6 ως 10	16,00	Αρκετούς	6 ως 10	64,00	Αρκετές	11 ως 30	37,50	Αρκετές	31 ως 80	70,00
	Πολλές	11 ως 40	60,00	Πολλούς	11 ως 40	240,00	Πολλές	31 ως 100	120,00	Πολλές	81 ως 150	140,00
	Πάρα πολλές	41 ως 70	120,00	Πάρα πολλούς	41 ως 70	480,00	Πάρα πολλές	101 και άνω	225,00	Πάρα πολλές	151 και άνω	200,00
Ελάχιστη τιμή	Σχήματα / Εικόνες	Αριθμός (ανά 100 σελίδες)	Χρόνος σε λεπτά	Πίνακες	Αριθμός (ανά 100 σελίδες)	Χρόνος σε λεπτά	Εξισώσεις	Αριθμός (ανά 100 σελίδες)	Χρόνος σε λεπτά	Βιβλιογραφία	Αριθμός (ανά 100 σελίδες)	Χρόνος σε λεπτά
	Καθόλου - Λίγο	1 ως 5	4,50	Καθόλου - Λίγο	1 ως 5	12,00	Καθόλου - Λίγο	1 ως 10	6,00	Καθόλου - Λίγο	ως 30	12,50
	Αρκετές	6 ως 10	12,00	Αρκετούς	6 ως 10	32,00	Αρκετές	11 ως 30	25,00	Αρκετές	31 ως 80	35,00
	Πολλές	11 ως 40	45,00	Πολλούς	11 ως 40	120,00	Πολλές	31 ως 100	80,00	Πολλές	81 ως 150	70,00
	Πάρα πολλές	41 ως 70	90,00	Πάρα πολλούς	41 ως 70	240,00	Πάρα πολλές	101 και άνω	150,00	Πάρα πολλές	151 και άνω	100,00

Εικόνα 9: Μήτρα κόστους

Ακολουθεί ένα παράδειγμα ανάλυσης του χρόνου επεξεργασίας του βιβλίου «Πληροφορικά Συστήματα Διοίκησης» της συλλογής του Κάλλιπου, σε ημέρες και ώρες.

3.4.1 Βιβλίο «Πληροφοριακά Συστήματα Διοίκησης»

Το βιβλίο «Πληροφοριακά συστήματα διοίκησης» κατά τη διάρκεια της καταγραφής του περιεχομένου του βρέθηκε ότι περιείχε πολλές εικόνες / σχήματα, λίγους έως καθόλου πίνακες και εξισώσεις και αρκετές βιβλιογραφικές αναφορές.



Εικόνα 10: Παράδειγμα βιβλίου από την συλλογή του Κάλλιπου

Στον Πίνακα 11 απεικονίζονται οι τιμές που ανακτήθηκαν (χρόνος σε λεπτά) από την μήτρα κόστους, σύμφωνα με τα πρωτογενή δεδομένα που εισήχθησαν στο μοντέλο για το συγκεκριμένο βιβλίο. Για τον υπολογισμό του μέσου όρου αθροίστηκαν το σύνολο των λεπτών με την τιμή αναφοράς για 100 σελίδες ανά κείμενο, πολλαπλασιάστηκαν επί των σελίδων του συγγράμματος και διαιρέθηκαν διά 100 που είναι η τιμή αναφοράς των σελίδων. Με την ίδια διαδικασία μελετήθηκε ο μέγιστος και ο ελάχιστος χρόνος (βλέπε Πίνακα 12).

	Σχήματα / Εικόνες	Συνολικός χρόνος σε λεπτά	Πίνακες	Συνολικός χρόνος σε λεπτά	Εξισώσεις	Συνολικός χρόνος σε λεπτά	Βιβλιογραφία	Συνολικός χρόνος σε λεπτά
Μ. όρος	Πολλές	50,10	Καθόλου - Λίγο	20,01	Καθόλου - Λίγο	7,98	Αρκετές	49,00
Μέγιστο	Πολλές	60,00	Καθόλου - Λίγο	24,00	Καθόλου - Λίγο	9,00	Αρκετές	70,00
Ελάχιστο	Πολλές	45,00	Καθόλου - Λίγο	12,00	Καθόλου - Λίγο	6,00	Αρκετές	35,00

Πίνακας 11: Ανάκτηση τιμών από την μήτρα κόστους

	Μέσος όρος	Μέγιστος χρόνος	Ελάχιστος χρόνος
Συνολικός χρόνος	1162,09	1446,66	840,36
Σε ώρες	19,37	24,11	14,01
Σε ημέρες	2,42	3,01	1,75
Σελίδες	282		

Πίνακας 12: Ανάλυση υπολογισμού εργατωρών

Υπολογισμός ωρών και ημερών επεξεργασίας του βιβλίου

Μέσος όρος

$285+50,10+20,01+7,98+49,00=412,09$ (λεπτά) * 282 (σελίδες)= 116.209,38 ÷ 100 (τιμή αναφοράς)=1162,09 συνολικός χρόνος σε λεπτά

Ώρες: $1162,09 \div 60= 19,37$ ώρες

Ημέρες: $19,37 \div 8= 2,42$ ημέρες

Μέγιστος χρόνος

$350+60,00+24,00+9,00+70,00= 513$ (λεπτά) * 282 (σελίδες)= 144.666 ÷ 100 (τιμή αναφοράς)= 1446,66 μέγιστος χρόνος σε λεπτά

Ώρες: $1446,66 \div 60= 24,11$ ώρες

Ημέρες: $24,11 \div 8= 3,01$ ημέρες

Ελάχιστος χρόνος

$200+45,00+12,00+6,00+35,00= 298$ (λεπτά) * 282 (σελίδες)= 84.036÷100 (τιμή αναφοράς)= 840,36 ελάχιστος χρόνος σε λεπτά

Ώρες: $840,36 \div 60= 14,01$ ώρες

Ημέρες: $14,01 \div 8= 1,75$ ημέρες

Αποτέλεσμα:

Ο μέσος όρος επεξεργασίας που χρειάζεται το συγκεκριμένο βιβλίο είναι **19,37 ώρες** (ή **2,42 ημέρες**), ο μέγιστος χρόνος επεξεργασίας είναι **24,11 ώρες**, (ή **3,01 ημέρες**) και ο ελάχιστος χρόνος επεξεργασίας υπολογίστηκε σε **14,01 ώρες** (ή **1,75 ημέρες**).

Επιπλέον, στον χρόνο επεξεργασίας πρέπει να προστεθεί αθροιστικά το ποσοστό της χρονικής επιβάρυνσης 20% για το μορφότυπο του αρχείου που είναι σε μορφή Adobe In Design με Word και το ποσοστό 35% επειδή περιέχει βίντεο (20%) και διαδραστικό υλικό (15%), συνολικά ένα 55% ποσοστό χρονικής επιβάρυνσης.

Οπότε μετά και τη χρονική επιβάρυνση λόγω των υπολοίπων παραγόντων ο συνολικός χρόνος επεξεργασίας υπολογίζεται ως εξής:

- μέσος όρος επεξεργασίας : 30,02 ώρες ή 3,75 μέρες
- μέγιστος χρόνος επεξεργασίας : 37,37 ώρες ή 4,67 μέρες
- ελάχιστος χρόνος επεξεργασίας : 21,71 ώρες ή 2,71 μέρες

Κεφάλαιο 4. Αποτελέσματα – Ευρήματα / Επιτεύγματα

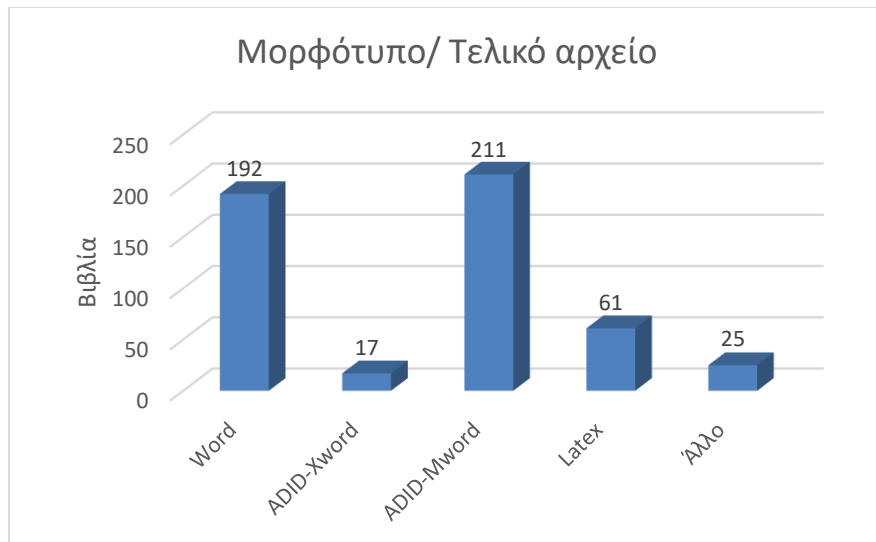
Σε αυτό το κεφάλαιο πραγματοποιείται η παρουσίαση των δευτερογενών δεδομένων που προέκυψαν από τη μήτρα κόστους και αναλύονται όλα τα στατιστικά στοιχεία σχετικά με τη στοιχειοθεσία των ηλεκτρονικών συγγραμμάτων της συλλογής του «Κάλλιπου». Τέλος, παρουσιάζεται ο υπολογισμός του χρόνου επεξεργασίας, για όλη τη συλλογή της πρώτης φάσης του αποθετηρίου «Κάλλιπος», και του συνολικού κόστους για την υλοποίηση της ψηφιακής διατήρησης της συλλογής.

4.1 Αναλυτική παρουσίαση αποτελεσμάτων

Από την διαδικασία της καταγραφής της στοιχειοθεσίας των ηλεκτρονικών συγγραμμάτων και βάσει του μοντέλου της μήτρας κόστους προέκυψαν ορισμένα δευτερογενή δεδομένα. Από τα 517 ηλεκτρονικά συγγράμματα, τα 11 ήταν ανολοκλήρωτα, συνεπώς δεν ήταν αναρτημένα στο Αποθετήριο του Κάλλιπου, για αυτό δεν συμπεριλήφθηκαν στην εξαγωγή αποτελεσμάτων, οπότε το σύνολο των βιβλίων ανήλθε στα 506 βιβλία.

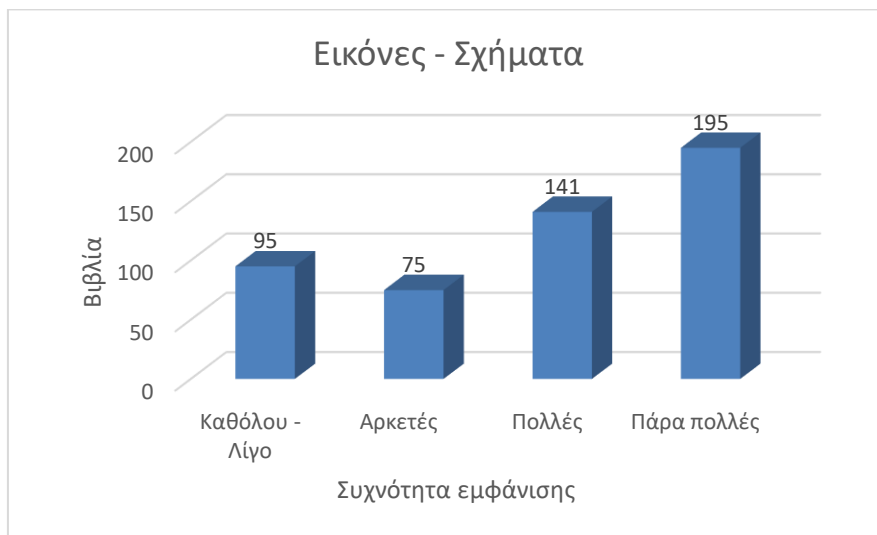
Στα παρακάτω γραφήματα αποτυπώνεται μια επισκόπηση των δευτερογενών δεδομένων, τα οποία προέρχονται από την αναλυτική καταγραφή της 1^{ης} συλλογής του Κάλλιπου.

Τα μορφότυπα των τελικών αρχείων που ανακτήθηκαν από τους φακέλους ομαδοποιήθηκαν σε πέντε κατηγορίες (βλέπε Εικόνα 11). Τα 211 από τα 506 βιβλία (ποσοστό 41,7 %) ήταν επεξεργασμένα με το λογισμικό Adobe In Design και παράλληλα περιείχαν την αρχική μορφή του πρωτογενούς υλικού σε Microsoft Word. Σε ποσοστό 37,94%, δηλαδή τα 192 από τα 506 βιβλία, περιήλθαν τα αρχεία που συγγράφηκαν σε μορφή Word. Τα βιβλία που ήταν γραμμένα σε μορφή Latex ανήλθαν σε 61 από 506 βιβλία, κατείχαν δηλαδή το 12,6% του συνόλου. Τέλος, 25 αρχεία είτε ήταν σε άλλη μορφή είτε δεν υπήρχε διαθέσιμο το πρωτογενές υλικό τους.



Εικόνα 11: Σύνολο βιβλίων ανά μορφή του τελικού αρχείου

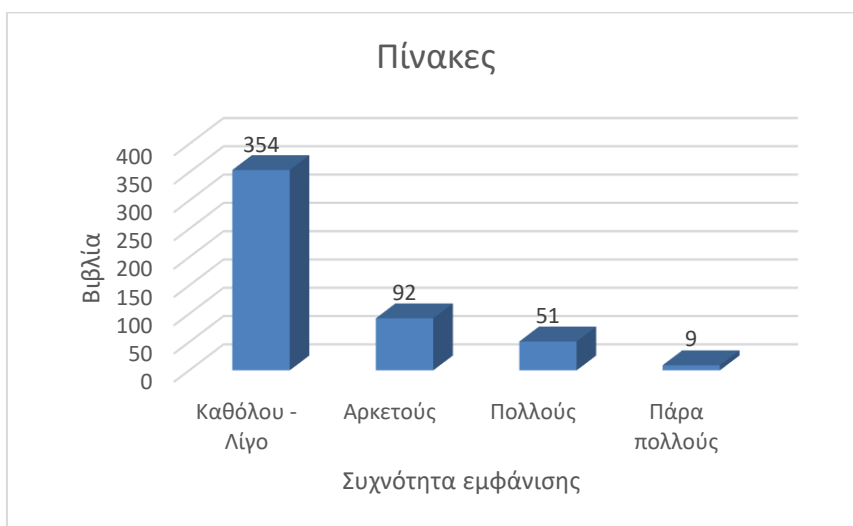
Οι εικόνες και τα σχήματα διαχωρίστηκαν σε τέσσερα επίπεδα (βλέπε Εικόνα 12), ανάλογα με την συχνότητα που παρουσιάζονται μέσα στο κείμενο (όπως παρουσιάστηκαν στο κεφάλαιο 3.2). Τα 195 από τα 506 βιβλία, με ποσοστό 38,54%, περιείχαν «Πάρα πολλές» εικόνες και / ή σχήματα, καθώς και τα 141 από τα 506 βιβλία, με ποσοστό 27,87%, εμπειρείχαν «Πολλές» εικόνες και / ή σχήματα. Σε αντίθεση, τα 95 από τα 506 βιβλία, με ποσοστό 18,77%, περιλάμβαναν «Ελάχιστες» εικόνες και σε ορισμένες περιπτώσεις «Καθόλου». Τέλος, σε ποσοστό 14,82%, δηλαδή τα 75 από τα 506 βιβλία, ενσωμάτωναν «Αρκετές» εικόνες και / ή σχήματα.



Εικόνα 12: Κατηγοριοποίηση βιβλίων σε σχέση με τον αριθμό εικόνων/ σχημάτων

Παρομοίως, τα βιβλία κατηγοριοποιήθηκαν σε τέσσερα επίπεδα (βλέπε Εικόνα 13), ανάλογα με την πυκνότητα εμφάνισης των πινάκων στο κείμενο. Σε μεγάλο ποσοστό της τάξεως 69,96%, δηλαδή τα 354 από τα 506 βιβλία, περιελάμβαναν «Ελάχιστους» έως

«Καθόλου» πίνακες στο περιεχόμενό τους. Με μεγάλη διαφορά, το αμέσως επόμενο μεγαλύτερο ποσοστό ανήλθε στο 18,18% και αφορούσε τα 92 από τα 506 βιβλία, τα οποία περιείχαν «Αρκετούς» πίνακες. Μικρότερα ποσοστά εμφανίζονται για τα επόμενα δύο επίπεδα. Συγκεκριμένα μόλις τα 51 από τα 506 βιβλία ανήκαν στο επίπεδο «Πολλούς πίνακες», δηλαδή ποσοστό 10,08%. Τέλος, μόνο 9 από τα 506 βιβλία ή ποσοστό 1,78% περιείχαν «Πάρα πολλούς» πίνακες.



Εικόνα 13: Κατηγοριοποίηση βιβλίων σε σχέση με τον αριθμό πινάκων

Με την ίδια λογική ποσοτικοποιήθηκαν και οι εξισώσεις (βλέπε Εικόνα 14). Και εδώ παρατηρήθηκε ένα μεγάλο ποσοστό, της τάξεως του 60,47%, δηλαδή τα 306 από τα 506 βιβλία, να μην διαθέτει «Καθόλου» ή να έχει «Ελάχιστες» εξισώσεις. Αντιθέτως, μόλις τα 31 από τα 506 βιβλία, με ποσοστό 6,13%, περιείχαν «Αρκετές» εξισώσεις και ομοίως με μικρό ποσοστό 9,09%, μόλις τα 46 από τα 506 βιβλία περιλάμβαναν «Πολλές» εξισώσεις. Με σχετικά μεγαλύτερο ποσοστό σε σχέση με τα δύο προηγούμενα επίπεδα δηλαδή 24,31%, αναδείχθηκαν τα 123 από τα 506 βιβλία με «Πάρα πολλές» εξισώσεις.



Εικόνα 14: Κατηγοριοποίηση σε σχέση με τον αριθμό εξισώσεων

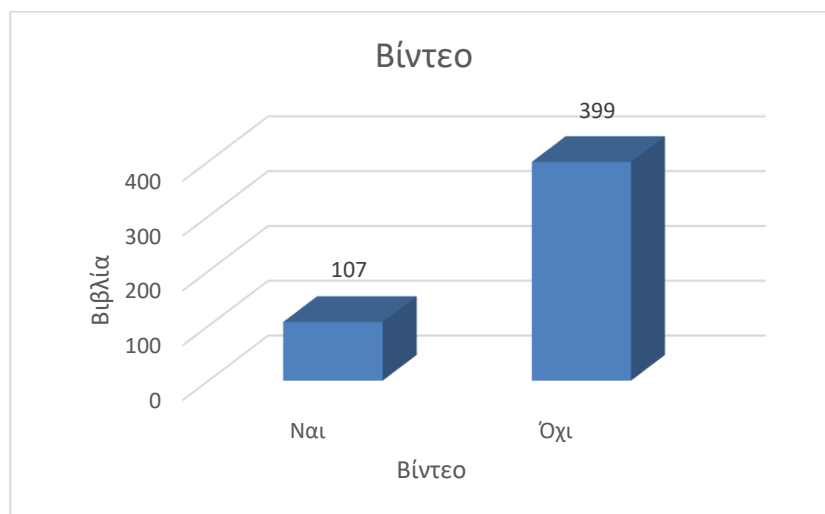
Αντίστοιχα, και οι βιβλιογραφικές αναφορές ομαδοποιήθηκαν σε τέσσερα επίπεδα (βλέπε Εικόνα 15). Εδώ παρατηρούνται μικρότερες διαφορές ανάμεσα στα επίπεδα συχνότητας εμφάνισης των βιβλιογραφικών αναφορών, με μεγαλύτερο ποσοστό 30,83%, δηλαδή τα 156 από τα 506 βιβλία, να διαθέτουν «Πάρα πολλές» βιβλιογραφικές αναφορές. «Καθόλου ή λίγο» περιείχαν τα 128 από τα 506 βιβλία, με ποσοστό 25,30%, ενώ «Αρκετές» βιβλιογραφικές αναφορές περιελάμβαναν τα 118 από τα 506 βιβλία (23,32%). Τέλος, «Πολλές» βιβλιογραφικές αναφορές κατείχαν τα 104 από τα 506 βιβλία (20,55%).



Εικόνα 15: Κατηγοριοποίηση βιβλίων σε σχέση με τον αριθμό βιβλιογραφικών αναφορών

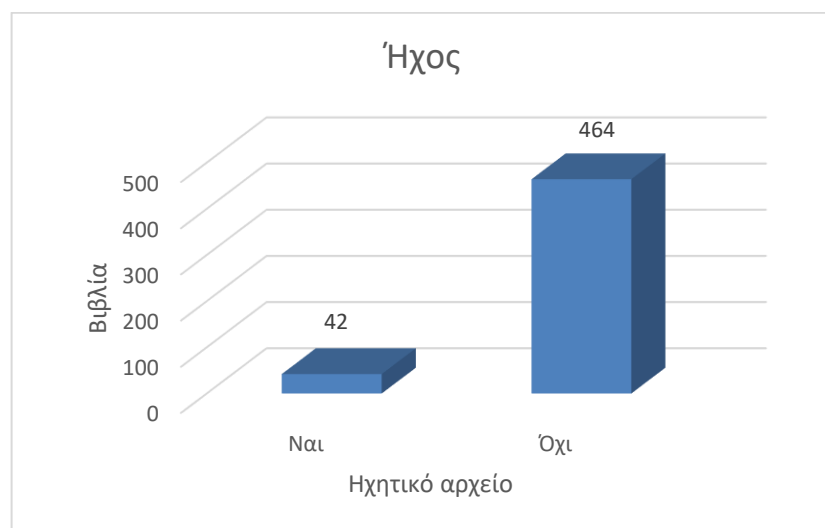
Στη συνέχεια, συγκεντρώθηκαν τα σύνολα των βιβλίων που περιείχαν, ή όχι, βίντεο (βλέπε Εικόνα 16). Η ομαδοποίηση των βίντεο πραγματοποιήθηκε με την δυαδική επιλογή Ναι / Όχι. Τα δεδομένα που εξήχθησαν απέδειξαν ότι τα 399 από τα 506 βιβλία, στο

συντριπτικό ποσοστό του 78,85%, δεν περιλάμβαναν αρχεία βίντεο, σε αντίθεση με τα 107 από τα 506 βιβλία, με ποσοστό 21,15%, τα οποία διέθεταν βίντεο στο περιεχόμενό τους.



Εικόνα 16: Ποσότητα βιβλίων που περιέχουν ή όχι βίντεο

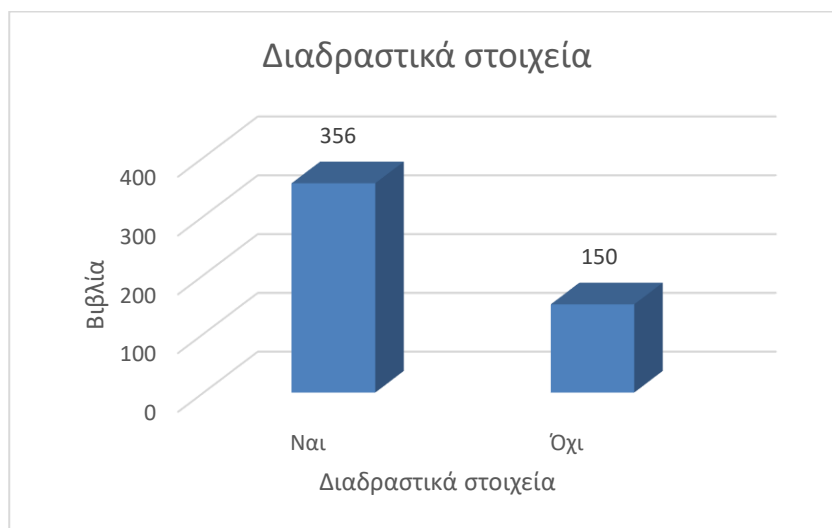
Η ίδια διαδικασία ακολουθήθηκε και με τα αρχεία ήχου (βλέπε Εικόνα 17). Η κατηγοριοποίηση έγινε με βάση το δυαδικό διαχωρισμό σε Ναι / Όχι. Τα 464 από τα 506 βιβλία, με ποσοστό 91,70%, δεν περιλάμβαναν αρχεία ήχου στο κείμενό τους, ενώ τα 42 από τα 506 βιβλία, με ποσοστό 8,30% διέθεταν.



Εικόνα 17: Ποσότητα βιβλίων που περιέχουν ή όχι ήχο

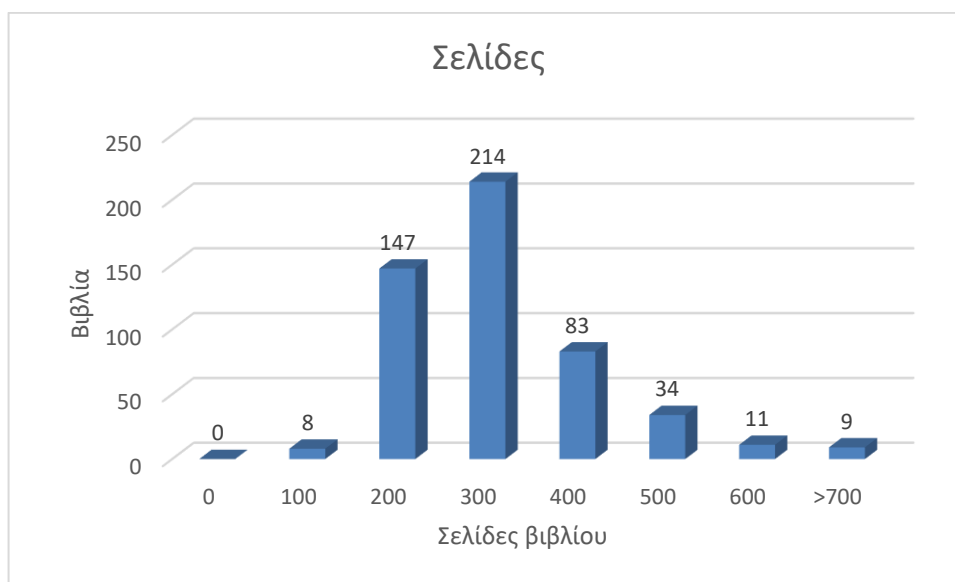
Τα διαδραστικά στοιχεία περιείχαν πολλές μορφές αρχείων, όμως για την απλοποίηση της εξαγωγής των συμπερασμάτων προτιμήθηκε η δυαδική επιλογή Ναι / Όχι, σχετικά με την ύπαρξή τους (ή όχι) μέσα στο κείμενο. Σύμφωνα με το γράφημα (βλέπε Εικόνα 18), τα 356 από τα 506 βιβλία, με ποσοστό 70,36%, διέθεταν διαδραστικά στοιχεία,

σε σύγκριση με τα 150 από τα 506 βιβλία, ποσοστό 29,64%, τα οποία δεν περιελάμβαναν κάποιο διαδραστικό στοιχείο στο περιεχόμενό τους.



Εικόνα 18: Ποσότητα βιβλίων που περιέχουν ή όχι διαδραστικά στοιχεία

Επιπροσθέτως, εξήχθησαν επιπλέον δεδομένα σχετικά με το σύνολο των σελίδων που απαρτιζόταν κάθε βιβλίο. Στο παρακάτω γράφημα (βλέπε Εικόνα 19) αποτυπώνεται το σύνολο των βιβλίων σε σχέση με την κατηγορία (ανά εκατοντάδα) σελίδων όπου ανήκει (περιέχει). Για παράδειγμα, 147 βιβλία περιείχαν έως 200 σελίδες, ενώ 214 βιβλία διέθεταν μέχρι 300 σελίδες. Με περισσότερες από 700 σελίδες καταγράφηκαν μόνο 9 βιβλία.



Εικόνα 19: Κατηγοριοποίηση βιβλίων σε σχέση με τον αριθμό σελίδων που περιέχουν

4.2 Κυριότερα Ευρήματα/ Αποτελέσματα

Σε αυτό το κεφάλαιο παρουσιάζονται τα αποτελέσματα που προέκυψαν, αφότου ολοκληρώθηκε η αναλυτική καταγραφή της σύνθεσης των βιβλίων της 1^{ης} φάσης του Κάλλιπου και πραγματοποιήθηκε η επεξεργασία και η ανάλυση των δεδομένων τους.

Σύμφωνα με τα δεδομένα που εισήχθησαν στο μοντέλο της μήτρα κόστους, για κάθε ένα βιβλίο ξεχωριστά, προέκυψαν εκτιμήσεις χρόνου σχετικά με την επιμέλεια και την επεξεργασία των βιβλίων. Το μοντέλο είχε σχεδιαστεί με στόχο να υπολογίσει για κάθε τεκμήριο τον μέσο όρο, τον μέγιστο και τον ελάχιστο χρόνο εργατωρών, για την γραφιστική επιμέλεια των κειμένων, σε ώρες και ημέρες.

Με την ολοκλήρωση της συμπλήρωσης του μοντέλου, συγκεντρώθηκε ο μέσος όρος ημερών, σχετικά με την διάρκεια επεξεργασίας του κάθε βιβλίου. Στον παρακάτω πίνακα (βλέπε Πίνακας 13), καταγράφεται το άθροισμα των ημερών για την επεξεργασία του συνόλου των βιβλίων της συλλογής. Ο μέσος όρος ημερών που απαιτείται για την επιμέλεια των βιβλίων αθροίστηκε σε 2.187,47 ημέρες, ή αλλιώς 109,4 μήνες. Ο μέγιστος χρόνος υπολογίστηκε σε 2.707,00 ημέρες ή 135,4 μήνες. Αντίστοιχα, ο ελάχιστος απαιτούμενος αριθμός ημερών ανήλθε σε 1.576,02 ημέρες ή 78,8 μήνες.

Σύνολο			
	Μέσος όρος	Μέγιστος αριθμός	Ελάχιστος αριθμός
Ημέρες	2187,47	2707,00	1576,02
Μήνες	109,4	135,4	78,8
Χρόνια	9,94	12,3	7,16

Πίνακας 13: Σύνολο ημερών επεξεργασίας των βιβλίων

Στη συνέχεια, για την μελέτη του κόστους της ψηφιακής διατήρησης της 1^{ης} φάσης της συλλογής του Κάλλιπου, υπολογίστηκε η κατά προσέγγιση αμοιβή του ανθρωπίνου δυναμικού, που θα συντελέσει τις διεργασίες, με βάση το σύνολο των ημερών και κατ' επέκταση των μηνών που απαιτούνται για την συντήρηση. Η μέση τιμή χρονικής διάρκειας, στους 109,4 μήνες, πολλαπλασιάστηκε επί 3.000 € (αμοιβή ανά μήνα). Το ποσό ανήλθε σε 328.121,15 € για το σύνολο της επεξεργασίας των βιβλίων της συλλογής. Το κόστος ανά βιβλίο μελετήθηκε κατά προσέγγιση σε 648,46 €.

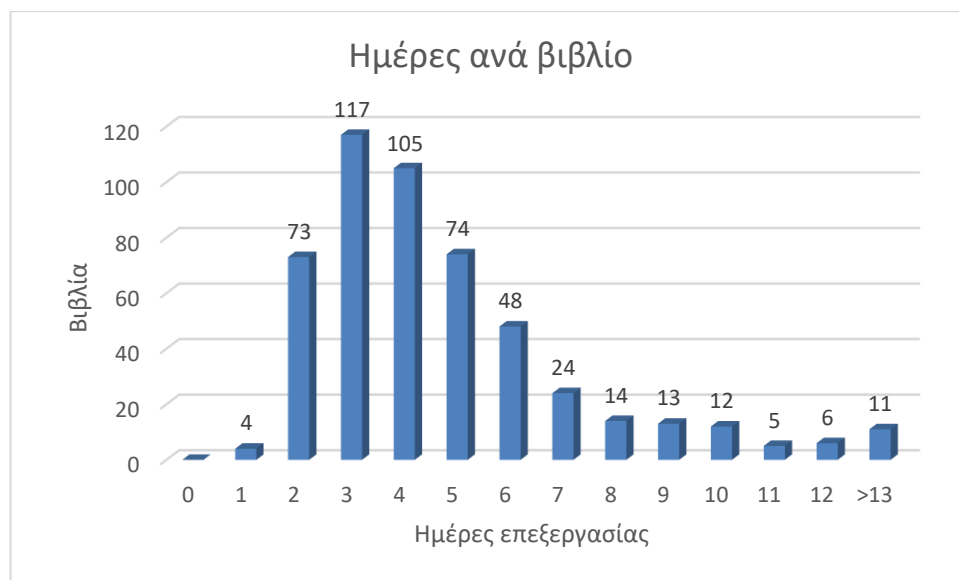
Παρόμοια, σύμφωνα με το μέγιστο αριθμό ημερών και μηνών, το συνολικό κόστος ανέρχεται σε 406.200,00 € και η τιμή ανά μονάδα τεκμηρίου υπολογίζεται σε 802,76 €. Για τον ελάχιστο αριθμό απαιτούμενων ημερών το κόστος μειώνεται στις 236.400,00 €, με τιμή ανά βιβλίο στα 467,19 € (βλέπε Πίνακας 14).

Σύνολο			
	Μέσος όρος	Μέγιστος αριθμός	Ελάχιστος αριθμός

Ημέρες	2187,47	2707,00	1576,02
Μήνες	109,4	135,4	78,8
Συνολικό κόστος	328.121,15 €	406.200,00 €	236.400,00 €
Κόστος ανά βιβλίο	648,46 €	802,76 €	467,19 €

Πίνακας 14: Συνολικό κόστος μελέτης

Παράλληλα, δημιουργήθηκε ένα γράφημα, αναφορικά με τις ημέρες που απαιτούνται για την ολοκλήρωση της επεξεργασίας κάθε βιβλίου (βλέπε Εικόνα 20). Στην παρακάτω εικόνα αποτυπώνεται η κατηγοριοποίηση των συνόλων των βιβλίων σύμφωνα με τις ημέρες που απαιτούνται για την επιμέλειά τους. Ενδεικτικά, 117 βιβλία χρειάζονται, κατά μέσο όρο, 3 μέρες προς επεξεργασία το καθένα, ενώ 24 βιβλία απαιτούν 7 ημέρες επεξεργασίας το καθένα.

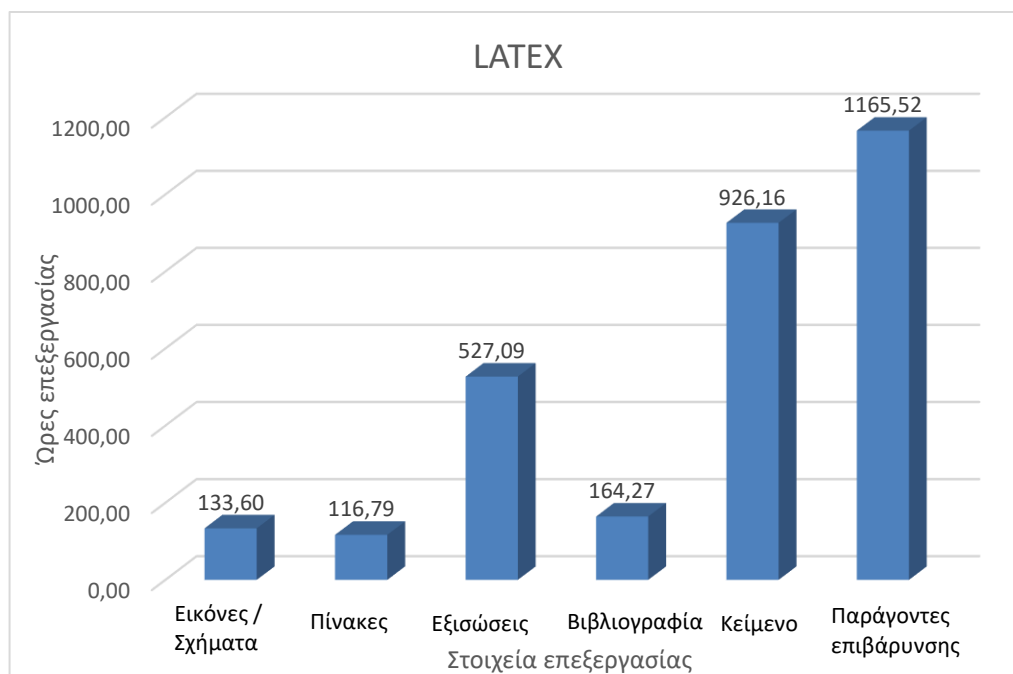


Εικόνα 20: Κατηγοριοποίηση βιβλίων σε σχέση με τον αριθμό των ημερών επεξεργασίας

Στη συνέχεια, πραγματοποιήθηκε μία ανάλυση σχετικά με τον χρόνο που απαιτείται για τον χρόνο επεξεργασίας που χρειάζονται τα στοιχεία όλων των μορφοτύπων.

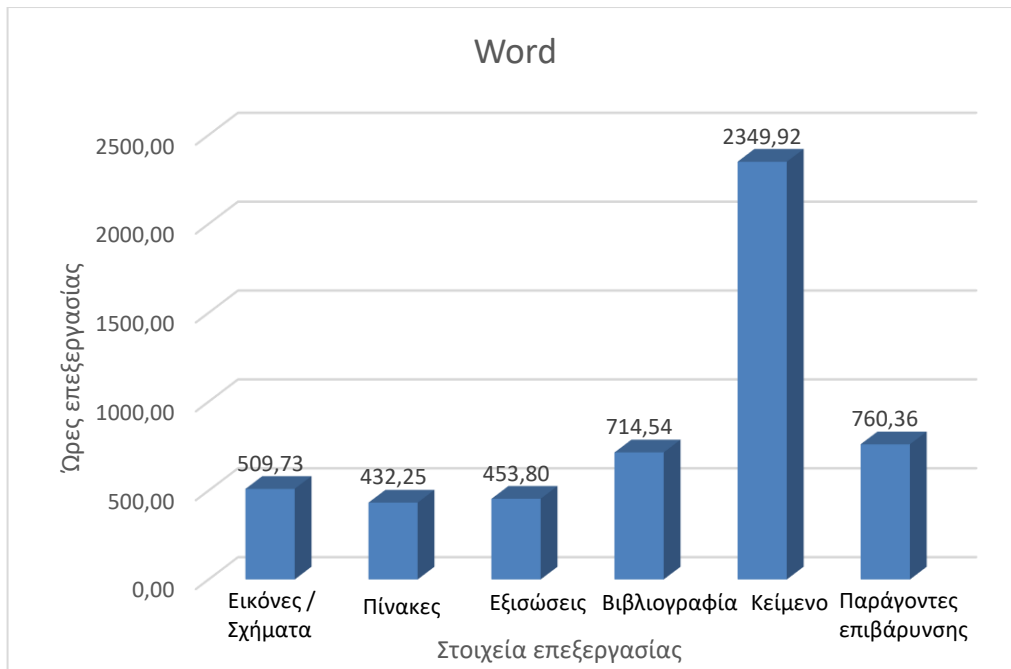
Το σύνολο των αρχείων που είναι σε μορφή LATEX ανέρχονται σε 61 βιβλία. Στο παρακάτω γράφημα (βλέπε Εικόνα 21), αποτυπώνονται οι ώρες επεξεργασίας για το σύνολο των στοιχείων που περιέχονται στα αρχεία LATEX, διαχωρισμένα ανά στοιχείο επεξεργασίας. Ο χρόνος του συνόλου των εικόνων και των σχημάτων ανέρχεται σε 133,60 ώρες επεξεργασίας, δηλαδή 16,7 ημέρες, για τους πίνακες χρειάζονται 116,79 ώρες επεξεργασίας, δηλαδή 14,59 ημέρες, για τις εξισώσεις απαιτούνται 527,09 ώρες επεξεργασίας, δηλαδή 65,88 ημέρες, για την βιβλιογραφία χρειάζεται 164,27 ώρες επεξεργασίας, δηλαδή 20,53 ημέρες, για το κείμενο 926,16 ώρες επεξεργασίας, δηλαδή 115,77 ημέρες και για τους παράγοντες επιβάρυνσης (βίντεο, αρχεία ήχου, διαδραστικό

υλικό) 1165,52 ώρες επεξεργασίας, δηλαδή 145,69 ημέρες. Συνολικά, ο μέσος όρος επεξεργασίας αθροίστηκε σε 3033,43 ώρες, άρα 379,17 ημέρες εργασίας για την επεξεργασία των αρχείων LATEX.



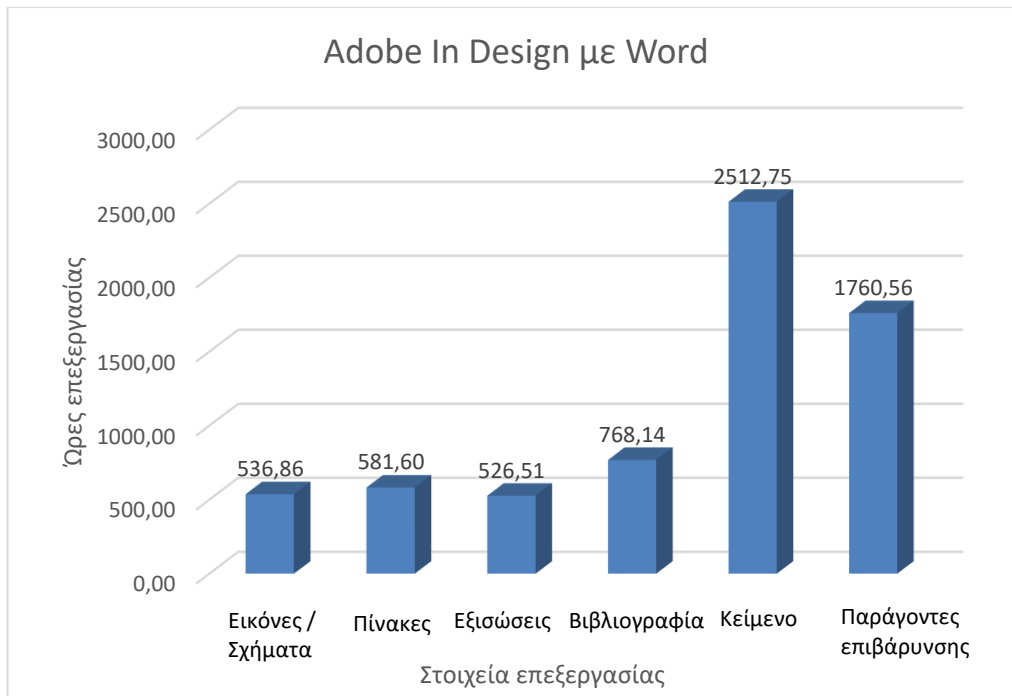
Εικόνα 21: Συνολικός χρόνος επεξεργασίας για τα αρχεία σε μορφή LATEX

Το σύνολο των αρχείων που είναι σε μορφή Word αθροίστηκαν σε 192 βιβλία. Στο επόμενο γράφημα (βλέπε Εικόνα 22), αποτυπώνονται οι ώρες επεξεργασίας για το σύνολο των στοιχείων που περιέχονται στα αρχεία Word, διαχωρισμένα ανά στοιχείο επεξεργασίας. Ο χρόνος του συνόλου των εικόνων και των σχημάτων ανέρχεται σε 509,73 ώρες επεξεργασίας, δηλαδή 63,71 ημέρες, για τους πίνακες χρειάζονται 432,25 ώρες επεξεργασίας, δηλαδή 54,03 ημέρες, για τις εξισώσεις απαιτούνται 453,80 ώρες επεξεργασίας, δηλαδή 56,72 ημέρες, για την βιβλιογραφία χρειάζεται 714,54 ώρες επεξεργασίας, δηλαδή 89,31 ημέρες, για το κείμενο 2349,92 ώρες επεξεργασίας, δηλαδή 293,74 ημέρες και για τους παράγοντες επιβάρυνσης (βίντεο, αρχεία ήχου, διαδραστικό υλικό) 760,36 ώρες επεξεργασίας, δηλαδή 95,04 ημέρες. Συνολικά, ο μέσος όρος επεξεργασίας αθροίστηκε σε 5220,6 ώρες, άρα 652,57 ημέρες εργασίας για την επεξεργασία των αρχείων Word.



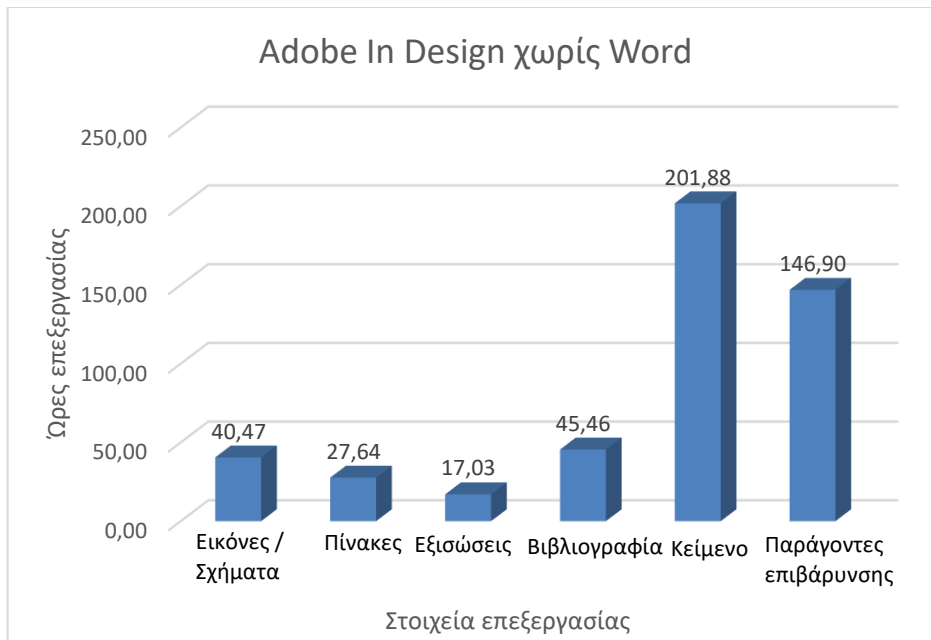
Εικόνα 22: Συνολικός χρόνος επεξεργασίας για τα αρχεία σε μορφή Word

Το σύνολο των αρχείων που είναι σε μορφή Adobe In Design με Word αθροίστηκαν σε 211 βιβλία. Στο επόμενο γράφημα (βλέπε Εικόνα 23), αποτυπώνονται οι ώρες επεξεργασίας για το σύνολο των στοιχείων που περιέχονται στα αρχεία Adobe In Design με Word, διαχωρισμένα ανά στοιχείο επεξεργασίας. Ο χρόνος του συνόλου των εικόνων και των σχημάτων ανέρχεται σε 536,86 ώρες επεξεργασίας, δηλαδή 67,10 ημέρες, για τους πίνακες χρειάζονται 581,60 ώρες επεξεργασίας, δηλαδή 72,7 ημέρες, για τις εξισώσεις απαιτούνται 526,51 ώρες επεξεργασίας, δηλαδή 65,81 ημέρες, για την βιβλιογραφία χρειάζεται 768,14 ώρες επεξεργασίας, δηλαδή 96,01 ημέρες, για το κείμενο 2512,75 ώρες επεξεργασίας, δηλαδή 314,09 ημέρες και για τους παράγοντες επιβάρυνσης (βίντεο, αρχεία ήχου, διαδραστικό υλικό) 1760,56 ώρες επεξεργασίας, δηλαδή 220,07 ημέρες. Συνολικά, ο μέσος όρος επεξεργασίας αθροίστηκε σε 6686,42 ώρες, άρα 835,80 ημέρες εργασίας για την επεξεργασία των αρχείων Adobe In Design με Word.



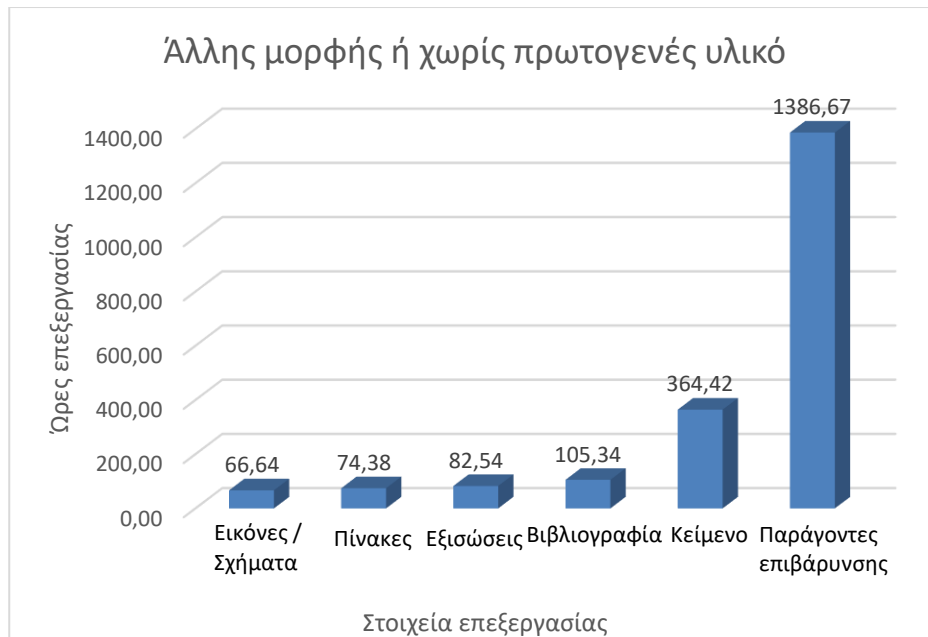
Εικόνα 23: Συνολικός χρόνος επεξεργασίας για τα αρχεία σε μορφή Adobe In Design με Word

Το σύνολο των αρχείων που είναι σε μορφή Adobe In Design χωρίς Word αθροίστηκαν σε 17 βιβλία. Στο επόμενο γράφημα (βλέπε Εικόνα 24), αποτυπώνονται οι ώρες επεξεργασίας για το σύνολο των στοιχείων που περιέχονται στα αρχεία Adobe In Design χωρίς Word, διαχωρισμένα ανά στοιχείο επεξεργασίας. Ο χρόνος του συνόλου των εικόνων και των σχημάτων ανέρχεται σε 40,47 ώρες επεξεργασίας, δηλαδή 5,05 ημέρες, για τους πίνακες χρειάζονται 27,64 ώρες επεξεργασίας, δηλαδή 3,45 ημέρες, για τις εξισώσεις απαιτούνται 17,03 ώρες επεξεργασίας, δηλαδή 2,12 ημέρες, για την βιβλιογραφία χρειάζεται 45,46 ώρες επεξεργασίας, δηλαδή 5,68 ημέρες, για το κείμενο 201,88 ώρες επεξεργασίας, δηλαδή 25,23 ημέρες και για τους παράγοντες επιβάρυνσης (βίντεο, αρχεία ήχου, διαδραστικό υλικό) 146,90 ώρες επεξεργασίας, δηλαδή 18,36 ημέρες. Συνολικά, ο μέσος όρος επεξεργασίας αθροίστηκε σε 479,38 ώρες, άρα 59,92 ημέρες εργασίας για την επεξεργασία των αρχείων Adobe In Design χωρίς Word.



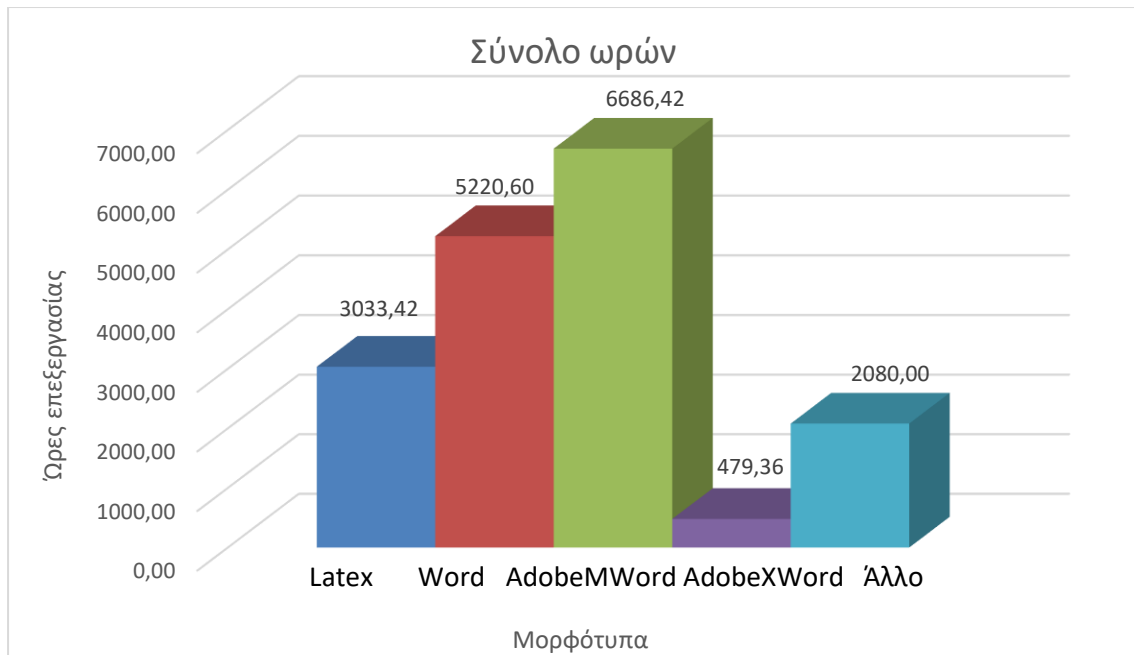
Εικόνα 24: Συνολικός χρόνος επεξεργασίας για τα αρχεία σε μορφή Adobe In Design χωρίς Word

Το σύνολο των αρχείων που ήταν σε άλλη μορφή ή δεν περιείχαν πρωτογενές υλικό ανήλθαν σε 25 βιβλία. Στο επόμενο γράφημα (βλέπε Εικόνα 25), αποτυπώνονται οι ώρες επεξεργασίας για το σύνολο των στοιχείων που περιέχονται στα αρχεία με άλλη μορφή, διαχωρισμένα ανά στοιχείο επεξεργασίας. Ο χρόνος του συνόλου των εικόνων και των σχημάτων ανέρχεται σε 66,64 ώρες επεξεργασίας, δηλαδή 8,36 ημέρες, για τους πίνακες χρειάζονται 74,38 ώρες επεξεργασίας, δηλαδή 9,29 ημέρες, για τις εξισώσεις απαιτούνται 82,54 ώρες επεξεργασίας, δηλαδή 10,31 ημέρες, για την βιβλιογραφία χρειάζεται 105,34 ώρες επεξεργασίας, δηλαδή 13,16 ημέρες, για το κείμενο 364,42 ώρες επεξεργασίας, δηλαδή 45,55 ημέρες και για τους παράγοντες επιβάρυνσης (βίντεο, αρχεία ήχου, διαδραστικό υλικό) 1386,67 ώρες επεξεργασίας, δηλαδή 173,33 ημέρες. Συνολικά, ο μέσος όρος επεξεργασίας αθροίστηκε σε 2080 ώρες, άρα 260 ημέρες εργασίας για την επεξεργασία των αρχείων Adobe In Design χωρίς Word.

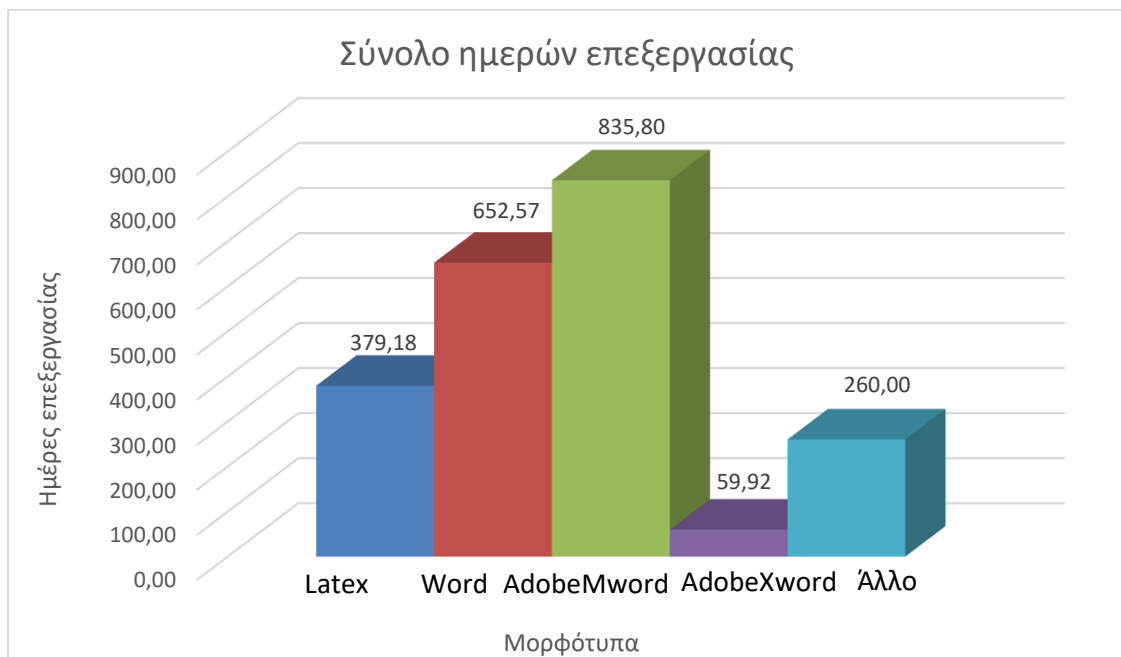


Εικόνα 25: Συνολικός χρόνος επεξεργασίας για τα αρχεία σε άλλη μορφή

Συνοπτικά, ακολουθούν δύο συγκριτικά γραφήματα, όπου αποτυπώνεται πρωτίστως το σύνολο των ωρών (βλέπε Εικόνα 26) και ακολούθως το σύνολο των ημερών επεξεργασίας (βλέπε Εικόνα 27) της συλλογής της πρώτης φάσης του Κάλλιπου, για κάθε μορφότυπο αρχείου που περιέχεται στη συλλογή. Ο συνολικός χρόνος επεξεργασίας για τα αρχεία LATEX υπολογίστηκε σε 3.033,42 ώρες, δηλαδή σε 379,18 ημέρες, για τα αρχεία Word σε 5.220,60 ώρες, δηλαδή σε 652,57 ημέρες, για τα αρχεία Adobe με Word σε 6.686,42 ώρες, δηλαδή σε 835,80 ημέρες, για τα αρχεία Adobe χωρίς Word σε 479,36 ώρες, δηλαδή σε 59,92 ημέρες και για τα αρχεία σε άλλο μορφότυπο ή που δεν περιείχαν πρωτογενές υλικό, καθώς υπάρχει μεγαλύτερος βαθμός δυσκολίας, ο χρόνος επεξεργασίας που απαιτείται ανέρχεται σε 2.080 ώρες, δηλαδή 260 ημέρες.



Εικόνα 26: Σύνολο ωρών επεξεργασίας για κάθε μορφότυπο



Εικόνα 27: Σύνολο ημερών επεξεργασίας για κάθε μορφότυπο

4.3 Επιπλέον κέντρα επιβάρυνσης

Μετά την αναλυτική περιγραφή και τον υπολογισμό των βασικών κέντρων κόστους για τη διατήρηση των ηλεκτρονικών συγγραμμάτων, πρέπει να αναφερθούν ορισμένα επιπλέον κόστη επιβάρυνσης που προκύπτουν αμέσως μετά τη δημιουργία ή την ανανέωση του υλικού.

Αρχικά, η δημιουργία του εξωφύλλου απαιτεί επιπλέον χρόνο και κόστος, καθώς χρειάζεται σχεδιασμό και γραφιστική επιμέλεια. Οι γραφίστες σε συνεννόηση με τους συγγραφείς επιλέγουν την εικόνα και τα χρώματα που ταιριάζουν στο βιβλίο, την γραμματοσειρά που θα έχει ο τίτλος και τα ονόματα της συγγραφικής ομάδας. Γενικότερα, ο γραφίστας επιμελείται τη δομή του εξωφύλλου και του οπισθόφυλλου, τα οποία μπορεί να περιέχουν, εκτός των προαναφερθέντων στοιχείων (εικόνα, τίτλος, όνομα συγγραφέα), το λογότυπο του εκδοτικού οίκου, ονοματεπώνυμο επιμελητή ή άλλες πληροφορίες σχετικά με την κυκλοφορία του βιβλίου. Στο οπισθόφυλλο περιλαμβάνονται η περίληψη, σε κάποιες περιπτώσεις κριτικές για το σύγγραμμα και το ISBN.

Το ISBN αποτελεί άλλο ένα κέντρο επιβάρυνσης του κόστους, διότι απαιτείται η καταχώρηση του στους βιβλιογραφικούς καταλόγους, με σκοπό τη διακίνηση και την εμπορία του βιβλίου. Για την καταχώρηση του ISBN ακολουθείται μία διαδικασία μέσω αίτησης στην Εθνική Βιβλιοθήκη της Ελλάδος για την χορήγηση του διεθνούς μοναδικού αριθμού του βιβλίου¹⁹.

Στις περιπτώσεις ανανέωσης του συγγράμματος, όπου το αρχείο έχει επιστραφεί στον συγγραφέα για τυχόν αλλαγές και διορθώσεις, είναι απαραίτητη η επεξεργασία των τελικών αρχείων από τους επιμελητές, ώστε να δοθεί η έγκριση για τη δημοσίευση του ηλεκτρονικού συγγράμματος.

Στο στάδιο πριν τη δημοσίευση, καθώς μπορεί να έχουν προκύψει αλλαγές και τροποποιήσεις μετά την ανανέωση του υλικού, πρέπει να ελεγχθούν τα μεταδεδομένα του βιβλίου ώστε να πραγματοποιηθεί η ανάρτηση στο αποθετήριο.

Τα συγκεκριμένα κόστη μπορεί ανά βιβλίο να μην κατέχουν μεγάλο ποσοστό στο κόστος επεξεργασίας, αλλά συνολικά για όλα τα βιβλία της συλλογής της πρώτης φάσης του Κάλλιπου, που ανέρχονται σε 506, απαιτούν επιπλέον χρόνο και κατ' επέκταση επιβαρύνουν το συνολικό κόστος.

¹⁹ <https://www.nlg.gr/static-page/isbn/>

Κεφάλαιο 5. Συζήτηση – Συμπεράσματα – Μελλοντικές επεκτάσεις

5.1 Ανακεφαλαίωση

Οι στόχοι της παρούσας διπλωματικής εργασίας ήταν η κατανόηση των αναγκών και των απαιτήσεων της ψηφιακής διατήρησης, σε σχέση με τα σύνθετα ψηφιακά αντικείμενα, όπως είναι τα ηλεκτρονικά επιστημονικά βιβλία, η στρατηγική μακροχρόνιας ψηφιακής διατήρησης μιας συλλογής, που θα πρέπει να ακολουθηθεί, και η εφαρμογή των αποτελεσμάτων της στρατηγικής σε μία πραγματική συλλογή ηλεκτρονικών ακαδημαϊκών συγγραμμάτων, στην συγκεκριμένη περίπτωση της συλλογής του «Κάλλιπου».

Η έρευνα ξεκίνησε από τη μελέτη του υλικού της πρώτης φάσης της συλλογής του αποθετηρίου «Κάλλιπος» και συνεχίστηκε με την αναλυτική καταγραφή της σύνθεσης των ηλεκτρονικών ακαδημαϊκών συγγραμμάτων που περιέχει η συλλογή, ώστε να συγκεντρωθεί το πλήθος των στοιχείων και των χαρακτηριστικών για την αποτίμηση του όγκου του περιεχομένου των βιβλίων. Με βάση τη συγκέντρωση του συνόλου των στοιχείων των συγγραμμάτων δημιουργήθηκε ένα μοντέλο (μήτρα κόστους) για την εκτίμηση των εργασιών που απαιτούνται για την ανανέωση του υλικού, με έμφαση στον υπολογισμό του κόστους και του χρόνου επεξεργασίας των βιβλίων. Το μοντέλο τροφοδοτήθηκε με τα παραπάνω στοιχεία και απέδωσε αρχικά την χρονική διάρκεια που χρειάζεται κάθε βιβλίο ξεχωριστά (μέσος όρος / μέγιστος / ελάχιστος χρόνος επεξεργασίας) και μέσω αυτού εξήχθησαν συμπεράσματα και στατιστικά στοιχεία σχετικά με το συνολικό χρόνο, και κατ' επέκταση του κόστους που απαιτείται, για την υλοποίηση της ψηφιακής διατήρησης της συλλογής.

Τα αποτελέσματα έδειξαν αρχικά ότι, από τα 517 ηλεκτρονικά συγγράμματα τα 11 ήταν ανολοκλήρωτα, συνεπώς δεν ήταν αναρτημένα στο αποθετήριο και δεν συμπεριλήφθηκαν στην εξαγωγή των αποτελεσμάτων, οπότε το σύνολο των βιβλίων ανήλθε σε 506 βιβλία. Το μορφότυπο των τελικών αρχείων που συναντήθηκε πιο συχνά κατά την καταγραφή των ηλεκτρονικών βιβλίων, δηλαδή στα 211 από τα 506 βιβλία, ήταν το Adobe In Design συμπεριλαμβανομένου του πρωτογενούς αρχείου του συγγραφέα σε μορφή Word. Τα 192 βιβλία ήταν σε μορφή Word και σε μορφή LATEX είχαν δημιουργηθεί 61 από τα 506 βιβλία, ενώ μόλις 17 βιβλία ήταν επεξεργασμένα με Adobe In Design χωρίς να περιλαμβάνουν το πρωτογενές κείμενο.

Η πλειοψηφία των ηλεκτρονικών συγγραμμάτων περιείχε πάρα πολλές εικόνες και σχήματα, ενώ οι πίνακες σχεδόν απουσιάζουν στα 354 από τα 506 βιβλία με συχνότητα εμφάνισης ανά 100 σελίδες λίγο ή καθόλου. Παρόμοια εικόνα εμφανίζεται και στις εξισώσεις, καθώς τα 306 από τα 506 βιβλία περιέχουν λίγες ή καθόλου εξισώσεις. Οι βιβλιογραφικές αναφορές παρουσίασαν μικρές διαφορές στα επίπεδα συχνότητας της εμφάνισής τους, τα περισσότερα συγγράμματα όμως διέθεταν πάρα πολλές βιβλιογραφικές αναφορές. Τα αρχεία βίντεο και ήχου εμφανίστηκαν σε λίγα βιβλία, διότι τα 107 από τα 506 περιείχαν βίντεο, και μόλις τα 42 από τα 506 βιβλία περιείχαν αρχεία ήχου. Διαδραστικά στοιχεία περιλάμβανε η πλειοψηφία των τεκμηρίων, δηλαδή τα 356 βιβλία. Ο αριθμός των σελίδων που περιείχαν κατά μέσο όρο τα συγγράμματα ήταν ανάμεσα σε 200 με 300 σελίδες, ενώ λίγα περιείχαν πάνω από 500 σελίδες.

Ο μέσος όρος ημερών που υπολογίστηκε για την επιμέλεια και την επεξεργασία της συλλογής ήταν 2.187,47 ημέρες, δηλαδή 109,4 μήνες, ο μέγιστος χρόνος αθροίστηκε σε 2.707 ημέρες, δηλαδή 135,4 μήνες και ο ελάχιστος χρόνος ανήλθε σε 1.576,02 ημέρες, δηλαδή σε 78,8 μήνες.

Σύμφωνα με αυτά τα δεδομένα υπολογίστηκε, προσεγγιστικά, το κόστος που απαιτείται για την υλοποίηση της επιμέλειας και της επεξεργασίας του συνόλου των βιβλίων της συλλογής. Για τον μέσο όρο χρονικής επεξεργασίας το συνολικό κόστος ανήλθε σε 328.121,15 €, με κόστος ανά βιβλίο περίπου στα 648,46 €. Σύμφωνα με τον μέγιστο χρόνο επεξεργασίας το συνολικό κόστος υπολογίστηκε στα 406.200,00 €, με κόστος ανά βιβλίο στα 802,76 €, ενώ για τον ελάχιστο χρόνο απαιτούμενων ημερών το συνολικό κόστος μειώθηκε στα 236.400,00 €, με κόστος ανά βιβλίο στα 467,19 €.

5.2 Συζήτηση / Συμπεράσματα

Το πρώτο βήμα για την ψηφιακή διατήρηση ηλεκτρονικών βιβλίων ενός Αποθετηρίου οφείλει να είναι η συλλογή ενός δείγματος βιβλίων, με σκοπό να αναλυθούν οι ανάγκες διατήρησης του οργανισμού, οι κοινές μορφές αρχείων που υπάρχουν, τα τρέχοντα πρότυπα μεταδεδομένων που χρησιμοποιούνται, τα εργαλεία που χρησιμοποιούνται για τη δημιουργία των εγχειριδίων και οι κίνδυνοι που μπορεί να προκύψουν σχετικά με τη πρόσβαση.

Αυτή η προσέγγιση οδηγεί στην αναζήτηση μοντέλων διατήρησης καθώς και στον εντοπισμό λύσεων για τις διαφορετικές ανάγκες που αναδύονται για συγκεκριμένα είδη βιβλίων, όπως είναι τα ηλεκτρονικά ακαδημαϊκά συγγράμματα. Για παράδειγμα, ένα διαδραστικό ηλεκτρονικό εγχειρίδιο με τρισδιάστατα αντικείμενα και βίντεο απαιτεί

διαφορετικό σχεδιασμό διατήρησης από ένα στατικό ηλεκτρονικό βιβλίο που περιέχει κυρίως κείμενο.

Παράλληλα, όπως συνέβη και στην περίπτωση της συλλογής του «Κάλλιπου», ορισμένα βιβλία δεν περιείχαν την πρωτογενή μορφή του υλικού τους ή είχαν μόνο την επεξεργασμένη μορφή των γραφιστών, περιπτώσεις που δυσκολεύουν τη διαδικασία ανανέωσης του υλικού και κατ' επέκταση τη συντήρησή του.

5.3 Αξιοποίηση / Πρακτικές προεκτάσεις της έρευνας

Η μελέτη που πραγματοποιήθηκε, μέσω της αναλυτικής καταγραφής της στοιχειοθεσίας και των χαρακτηριστικών των ηλεκτρονικών βιβλίων της πρώτης συλλογής του Κάλλιπου, αποτελεί το πρώτο βήμα για τον σχεδιασμό της ψηφιακής διατήρησης των ηλεκτρονικών εγχειριδίων. Με γνώμονα το μοντέλο κοστολόγησης που δημιουργήθηκε, δίνεται η δυνατότητα εκτίμησης του κόστους και του χρόνου επεξεργασίας κάθε βιβλίου ξεχωριστά, και παράλληλα υπολογίζεται το κόστος μιας ολόκληρης συλλογής.

Η παρούσα μελέτη αποτελεί ένα εργαλείο για τη δημιουργία μιας πρότασης προς τους αρμόδιους δημόσιους οργανισμούς, με στόχο την αρωγή τους, μέσω χρηματοδοτήσεων, για την υλοποίηση της ανανέωσης της συλλογής του Κάλλιπου. Η εφαρμογή μιας στρατηγικής διατήρησης θα βελτιώσει και θα επικαιροποιήσει το υλικό της συλλογής του Αποθετηρίου καθώς επίσης θα εξάγει ανανεωμένα μορφότυπα υλικού, τα οποία θα βοηθήσουν στις μελλοντικές απαιτήσεις της διατήρησής τους.

Παράλληλα, είναι ένα σημαντικό βοήθημα ως προς την απεμπλοκή από «κλειδωμένες» μορφές αρχείων, με στόχο μία ενιαία μορφή τύπου, όπως είναι το DocBook, όπου θα είναι εφικτή η επαναχρησιμοποίηση των αρχείων σε περιπτώσεις ανανέωσης ή επικαιροποίησης του υλικού. Με αυτόν τον τρόπο δεν θα χρειαστεί να περαστούν ξανά εξισώσεις ή πίνακες στο ανανεωμένο αρχείο, καθώς θα υπάρχει ήδη το κυρίως σώμα του διαμορφωμένου / επιμελημένου κειμένου. Αυτό είναι ένα πολύ σημαντικό βήμα, καθώς αποφεύγεται μεγάλος όγκος δουλειάς για τους γραφίστες, διότι μπορούν να πραγματοποιούν τις αλλαγές χωρίς να χρειάζεται να επεξεργαστούν από την αρχή ολόκληρο το κείμενο. Συμπερασματικά, αναδύονται μεγάλα οικονομικά οφέλη από την συγκεκριμένη διαδικασία, καθώς δεν δημιουργείται επιπλέον κόστος συντήρησης.

5.4 Μελλοντικές επεκτάσεις / Πρακτικές Προεκτάσεις της Έρευνας

Με βάση την πρακτική εφαρμογή της παρούσας μελέτης αναδύονται ανάγκες για περαιτέρω έρευνα, οι οποίες μπορούν να αποτελέσουν συνέχεια καθώς και να ενισχύσουν σημαντικά το πεδίο της ψηφιακής διατήρησης.

Στην παρούσα έρευνα εξετάζονται ο χρόνος και το κόστος επεξεργασίας για τις μορφές αρχείων Word, Latex και Adobe In Design που εμπεριέχονται ήδη στο Αποθετήριο Κάλλιπος. Βασική μελλοντική επέκταση θα μπορούσε να αποτελέσει η αναπροσαρμογή του μοντέλου κοστολόγησης με βάση τη μετάπτωση των δεδομένων στη μορφή DocBook, με στόχο την εξαγωγή ακριβέστερων αποτελεσμάτων. Σε αυτό θα βοηθήσει η συμμετοχή των συντελεστών τεχνικής επεξεργασίας, μέσω της εκτίμησής τους σχετικά με τον χρόνο ανθρωποπροσπάθειας που απαιτείται για τη μετατροπή / αναθεώρηση των αρχείων σε DocBook.

Μία πιο γενική επέκταση έρευνας είναι η διαμόρφωση κανόνων και βέλτιστων πρακτικών. Λαμβάνοντας υπόψη τους περιορισμούς και την αξιοποίηση προεκτάσεων της μελέτης, δύναται να δημιουργηθούν πολιτικές και πρακτικές για όλες τις ροές εργασιών μέχρι το τελικό στάδιο της δημοσίευσης των συγγραμμάτων. Αυτές οι πρακτικές θα μπορούσαν να αφορούν τις διαδικασίες συγγραφής και γλωσσικής επιμέλειας κειμένου, γραφιστικής επιμέλειας και διαμόρφωση του τελικού αρχείου προς δημοσίευση. Παράλληλα, θα μπορούσαν να έχουν ως έναυσμα τις τάσεις νέων τεχνολογιών και συστημάτων πληροφόρησης, μελετώντας ζητήματα που προκύπτουν από τη σύγχρονη βιβλιογραφία.

Ως απόρροια των παραπάνω το πλαίσιο συνεργασίας για τη ψηφιακή διατήρηση δεν θα μπορούσε να απέχει από το σχεδιασμό μιας στρατηγικής. Το πλαίσιο συνεργασίας θα μπορούσε να περιλαμβάνει οργανισμούς, συστήματα, ανθρώπινο δυναμικό και κοινούς στόχους. Με την εφαρμογή όλων των προαναφερθέντων επιτυγχάνεται σε ένα μεγάλο βαθμό η διαλειτουργικότητα, η οποία οδηγεί στην διάχυση της γνώσης και της πληροφόρησης.

Μία τέταρτη μελλοντική επέκταση θα μπορούσε να περιλαμβάνει τη γνώμη των άμεσα εμπλεκόμενων συντελεστών, συγγραφέων και γραφιστών, σχετικά με την εφαρμογή και τη χρήση νέων εργαλείων συγγραφής και επιμέλειας κειμένων. Αυτό κατ' επέκταση θα μπορούσε να οδηγήσει σε ένα πρόγραμμα κατάρτισης για την επιμόρφωση και την απόκτηση νέων δεξιοτήτων.

Τέλος, τα βήματα της μεθοδολογίας και το μοντέλο κοστολόγησης που υλοποιήθηκαν, δύναται να εφαρμοστούν και στις μετέπειτα προσθήκες ηλεκτρονικών ακαδημαϊκών συγγραμμάτων που περιλαμβάνει το Αποθετήριο «Κάλλιπος».

Μελλοντικά προβλήματα αναπόφευκτα θα προκύψουν, αλλά θα υπάρχουν θεμελιώδεις βάσεις για την επίλυσή τους. Στόχος θα πρέπει να είναι η, χωρίς προβλήματα, παραγωγική διαδικασία ως το τελικό αποτέλεσμα.

Ξένη Βιβλιογραφία

- Adobe (n.d.). About Adobe PDF. https://helpx.adobe.com/gr_en/indesign/using/exporting-publishing-pdf.html (Ανακτήθηκε 25.7.2022).
- Adobe (n.d.). Edit PDF files. https://www.adobe.com/gr_en/acrobat/how-to/pdf-editor-pdf-files.html (Ανακτήθηκε 25.7.2022).
- Anyaoku, E.N., Echedom, A.U.N. and Baro, E.E. (2019), "Digital preservation practices in university libraries: An investigation of institutional repositories in Africa", *Digital Library Perspectives*, Vol. 35 No. 1, pp. 41-64. <https://doi.org/10.1108/DLP-10-2017-0041>.
- Ball, A. (2010). *Preservation and Curation in Institutional Repositories (version 1.3)*. Edinburgh, UK: Digital Curation Centre.
- Barrueco, J.M. and Termens, M. (2022). *Digital preservation in institutional repositories: a systematic literature review*. *Digital Library Perspectives*, Vol. 38 No. 2, pp. 161-174. <https://doi.org/10.1108/DLP-02-2021-0011>.
- Borghoff, U. M., Rödiger, P., Schmitz, L., & Scheffczyk, J. (2006). Long-term preservation of digital documents. Berlin Heidelberg: Springer.
- Chen, L. (2019). Retailers' Differentiation Strategy and Pricing in the Rental Market of Digital Content: A case of E-Textbooks. *Journal of theoretical and applied electronic commerce research*, 61-75.
- Cullier, C., Hofer, A., Johnson, A., Labadorf, K., Lauristen, K., Potter, P., Saunders, R., and Walz, A. (2016) *Modifying an Open Textbook: What You Need to Know.*, MN: Open Textbook Network, Minneapolis, 9.
- Cho, J. (2014), "Intellectual structure of the institutional repository field: a co-word analysis", *Journal of Information Science*, Vol. 40 No. 3, pp. 386-397, doi: 10.1177/0165551514524686.
- Dappert, A., Squire Guenther, R., & Peyrard, S. (2016). *Digital Preservation Metadata for Practitioners: Implementing PREMIS*. Cham, Switzerland: Springer. DOI 10.1007/978-3-319-43763-7.
- da Silva Júnior, L.P. and Manuel Borges, M. (2017), "Digital preservation policies of the institutional repositories at Brazilian Federal Universities", *The Electronic Library*, Vol. 35 No. 2, pp. 311-321. <https://doi.org/10.1108/EL-09-2015-0170>.
- EPUB zone (n.d.). Understanding EPUB 3. <http://www.epubzone.org/epub-3-overview/understanding-epub-3/> (Ανακτήθηκε 22.7.2022)
- Fileformat. (n.d). *PDF File Format: What is a PDF File?*. <https://docs.fileformat.com/pdf/> (Ανακτήθηκε 22.7.2022).
- Hockx-Yu, H. (2006). Digital preservation in the context of institutional repositories. *Program*, 232-243
- Kirchhoff, A. & Morrissey, S. (2014). Preserving eBooks. DPC Technology Watch Report. DOI: <http://dx.doi.org/10.7207/twr14-01>.
- Koutsileou, S., Kouis, D., & Mitrou, N. (2018). "Kallipos", the first open academic textbooks initiative during the years of crisis in Greece and its sustainable continuation.
- Kumar, Vimal and Dhanya, P. (2021). Open Access and self-publishing in the academic communication: a case study. *Library Philosophy and Practice*, <http://eprints.rclis.org/42191/>
- Lee, H. J., Messom, C., & Yau, K. A. (2012). e-Textbooks: Types, characteristics and open issues. *Journal of Computer*, 4(9), 2151-9617.

- Lynch, C.A. (2003), "Institutional repositories: essential infrastructure for scholarship in the digital age", *ARL: A Bimonthly Report*, Vol. 226, available at: www.cni.org/wp-content/uploads/2003/02/arl-br-226-Lynch-IRs-2003.pdf
- METS (2022). <https://www.loc.gov/standards/mets/METSOverview.v2.html> (Ανακτήθηκε 2.8.2022).
- Moulaison, H. L. & Le Beau, C. (2015). US Libraries as Publishers: Status and Concerns. In IFLA WLIC: Acquisition and Collection Development. Cape Town, South Africa: IFLA Library. <http://library.ifla.org/id/eprint/1159/>
- Moulaison, H. L., & Million, A. J. (2015). E-publishing in libraries: The [digital] preservation imperative. *OCLC Systems & Services: International digital library perspectives*. <https://www.emerald.com/insight/content/doi/10.1108/OCLC-02-2014-0009/full/html#idm45353793932176>.
- PREMIS (2022). <https://www.loc.gov/standards/premis/> (Ανακτήθηκε 2.8.2022).
- Rachman, Y. B. (2019). Digital preservation strategies: A case study of two academic libraries in Indonesia. *Library Philosophy and Practice*, 1-9.
- Walsh, N., (2010). *DocBook: the definitive guide*. United States: O'Reilly Media, Inc.
- Walters, W. H. (2013). E-books in academic libraries: challenges for acquisition and collection management. *portal: Libraries and the Academy*, 13(2), 187-211.
- Wikipedia. (n.d.). HTML. <https://el.wikipedia.org/wiki/HTML> (Ανακτήθηκε 22.7.2022)
- Zervos, S. (2015). Conservation and preservation of paper, books and archival materials [Undergraduate textbook]. Kallipos, Open Academic Editions. <http://hdl.handle.net/11419/63>.

Ελληνική Βιβλιογραφία

- Εθνική Βιβλιοθήκη της Ελλάδος. <https://www.nlg.gr/static-page/isbn/> (Ανακτήθηκε 1.8.2022).
- Ελληνικά Ακαδημαϊκά Ηλεκτρονικά Συγγράμματα και Βοηθήματα (2013). <https://www.kallipos.gr/el/> (Ανακτήθηκε 25.7.2022).
- Ζερβός, Σ. (2015). Συντήρηση και διατήρηση χαρτιού, βιβλίων και αρχειακού υλικού (2015) [Προπτυχιακό εγχειρίδιο]. Κάλλιπος, Ανοικτές Ακαδημαϊκές Εκδόσεις.
- Κάλλιπος, Ανοικτές ακαδημαϊκές εκδόσεις (2015). <https://repository.kallipos.gr>. (Ανακτήθηκε 25.7.2022).
- Κουής, Δ., Χριστάκη, Σ., Κουτσιλέου, Σ., Αναστασοπούλου, Χ., Βεράνης, Γ., Βράττη, Α., ... & Ταμπακοπούλου, Ρ. (2016). «Κάλλιπος»-Το μεγαλύτερο ψηφιακό Αποθετήριο Ακαδημαϊκών Ηλεκτρονικών Συγγραμμάτων Ανοικτής Πρόσβασης. *HealJournal*, 2(1), 10-28.

Πρόσθετη Βιβλιογραφία

- Hare S. & Sullivan, M. (2020). A Qualitative Study on the Digital Preservation of OER. *Libraries and the Academy*, 4 (20), pp. 749-773. doi: 10.1353/pla.2020.0036.
- Castagné, M. (2013). Consider the Source: The Value of Source Code to Digital Preservation Strategies. *School of Information Student Research Journal*, 2(2). doi: 10.31979/2575-2499.020205.
- Chowdhury, G. (2010). From digital libraries to digital preservation research: the importance of users and context. *Journal of documentation*.
- Shimray, S. R. & Ramaiah, C. K. (2018). Digital Preservation Strategies: An overview. *Recent Advances in Information Technology (READIT-2018)* (pp.46-54). Kalpakam, Tamilnadu: IGCAR.
- Seamus Ross (2012) Digital Preservation, Archival Science and Methodological Foundations for Digital Libraries, *New Review of Information Networking*, 17(1), pp. 43-68, doi: 10.1080/13614576.2012.679446.