

Ψηφιοποίηση του περιοδικού «Σπουδαί» του Πανεπιστημίου Πειραιά: το έργο μέχρι σήμερα

Όλγα Κολοβιστέα

Αρχειονόμος-Βιβλιοθηκονόμος
Βιβλιοθήκη Πανεπιστημίου Πειραιώς
olgak@unipi.gr

Περίληψη

Το έργο ψηφιοποίησης του περιοδικού «Σπουδαί» άρχισε να υλοποιείται τον Ιανουάριο του 2007 από τη Βιβλιοθήκη του Πανεπιστημίου Πειραιώς (www.lib.unipi.gr). Στόχος του έργου είναι η δημιουργία ενός ψηφιακού αποθετηρίου, όπου αποθηκεύονται και διατηρούνται τα σαρωμένα πρωτότυπα των άρθρων του περιοδικού, προκειμένου να είναι διαθέσιμα προς αναζήτηση και ανάκτηση από το ευρύ κοινό, επιστημονικό ή μη. Το παρόν άρθρο περιγράφει το έργο που έχει πραγματοποιηθεί μέχρι σήμερα και θέτει ορισμένα ζητήματα που αφορούν το μέλλον του περιοδικού, αλλά και των έργων ψηφιοποίησης εν γένει.

Λέξεις-κλειδιά: ψηφιοποίηση, αποθετήριο

Εισαγωγή

Τη στιγμή που το Google προσφέρει στους εκδότες δωρεάν ψηφιοποίηση των παλιών τευχών των περιοδικών τους (Grogg and Ashmore, 2007) και τις συζητήσεις στους βιβλιοθηκονομικούς κύκλους μονοπωλούν θέματα όπως η «ψηφιοποίηση» και η «ανοιχτή πρόσβαση», οι ελληνικές ακαδημαϊκές βιβλιοθήκες δραστηριοποιούνται και σε αυτό το πεδίο (Τσιμπόγλου κ.ά, 2006). Ενδεικτικά αναφέρονται η Βιβλιοθήκη του Πανεπιστημίου Μακεδονίας που προχώρησε στην ψηφιοποίηση του περιοδικού «ΤΟΠΟΣ: επιθεώρηση αστικών και περιφερειακών μελετών», καθώς και η Βιβλιοθήκη του Πανεπιστημίου Πατρών που δημιούργησε την συλλογή «Κοσμόπολις» με περιοδικά του Ελληνικού Λογοτεχνικού και Ιστορικού Αρχείου (Ε.Λ.Ι.Α.). Η ανάδειξη ανάλογων έργων με δημοσιεύματα στον τύπο (Μαγκλίνης και Σελλά, 2007) υπογραμμίζει τη σημασία τους τόσο για τον ακαδημαϊκό κόσμο και δη τους ερευνητές, όσο και για το ευρύτερο κοινωνικό σύνολο και τη λεγόμενη «δημοκρατία της γνώσης».

Σε αυτό το πλαίσιο η Βιβλιοθήκη του Πανεπιστημίου Πειραιώς αποφάσισε τον Ιανουάριο του 2007 την ψηφιοποίηση του περιοδικού «Σπουδαί», το οποίο εκδίδεται από το 1950 από το Πανεπιστήμιο Πειραιώς. Η θεματολογία του περιοδικού καλύπτει τομείς συναφείς με την Οικονομική Επιστήμη, καθώς και τη Διοίκηση Επιχειρήσεων. Το έργο της ψηφιοποίησης σχεδιάστηκε από μια βιβλιοθηκονόμο και έναν πληροφορικό, ενώ καθ' όλη τη διάρκεια υλοποίησής του έχει την υποστήριξη του Τμήματος Μηχανοργάνωσης της Βιβλιοθήκης του Πανεπιστημίου Πειραιώς. Τα μέχρι στιγμής αποτελέσματα της ψηφιοποίησης είναι διαθέσιμα μέσω του ιστοχώρου της Βιβλιοθήκης του Πανεπιστημίου Πειραιώς (<http://digilib.lib.unipi.gr/spoudai>) και η πρόσβαση είναι ανοιχτή για κάθε ενδιαφερόμενο.

Σχεδιασμός του έργου και στόχοι

Το έργο εντάσσεται στο ΕΠΕΑΕΚ II «Ενίσχυση και εμπλουτισμός Βιβλιοθήκης του Πανεπιστημίου Πειραιώς».

Οι λόγοι που οδήγησαν στην απόφαση ψηφιοποίησης του περιοδικού συνοψίζονται στους εξής:

- *εκπαιδευτικοί*, καθώς το περιοδικό περιέχει επιστημονικό υλικό χρήσιμο για την εκπαιδευτική λειτουργία των διαφόρων ακαδημαϊκών τμημάτων του Πανεπιστημίου Πειραιώς,

- *ιστορικοί*, αφού το περιοδικό μετρά περισσότερα από 50 χρόνια παρουσίας (1950 - σήμερα) με βράβευση από την Ακαδημία Αθηνών,
- *ανοιχτής και εύκολης πρόσβασης*, ώστε το πλήρες κείμενο του συνόλου του περιοδικού να είναι διαθέσιμο 24/7 από την ιστοσελίδα της Βιβλιοθήκης του Πανεπιστημίου για τα μέλη της ακαδημαϊκής κοινότητας, αλλά και για οποιονδήποτε ενδιαφερόμενο. Επιπλέον το πλήρες κείμενο των άρθρων είναι εύκολα και γρήγορα αναζητήσιμο μέσω των λειτουργιών που παρέχει το αποθετήριο,
- *διατήρησης*, δια μέσου της λειτουργίας ενός ψηφιακού αποθετηρίου, όπου θα παραμένει το υλικό ώστε να εξασφαλιστεί η μακροχρόνια διατήρησή του, ακόμη και σε περίπτωση καταστροφής του φυσικού (έντυπου) μέσου. Αυτή η παράμετρος είναι ιδιαίτερα σημαντική για παλαιότερα τεύχη του περιοδικού που είναι δυσεύρετα και σπάνια.

Κύριο στόχο του έργου αποτέλεσε η αύξηση της προσβασιμότητας στο περιοδικό και η παράλληλη λειτουργία με την έντυπη έκδοσή του.

Το κοινό στο οποίο στοχεύει το συγκεκριμένο έργο περιλαμβάνει:

- την ακαδημαϊκή κοινότητα του Πανεπιστημίου Πειραιά για την εξυπηρέτηση των πληροφοριακών αναγκών της (πρώτο επίπεδο),
- την ελληνική ακαδημαϊκή κοινότητα εν γένει (δεύτερο επίπεδο),
- τους χρήστες του Διαδικτύου (τρίτο επίπεδο).

Η παραπάνω ομαδοποίηση καθόρισε τα χαρακτηριστικά της διεπαφής που δημιουργήθηκε με έμφαση στο κοινό πρώτου επιπέδου. Η ακαδημαϊκή κοινότητα και οι ερευνητές είναι άτομα που επιθυμούν σε βάθος πληροφόρηση για την υποστήριξη της έρευνάς τους. Σημαντικά στοιχεία για την ψηφιοποίηση αποτελούν τόσο η ανάλυση της σάρωσης, όσο και η πρόσβαση, η ανακάλυψη και η ανάκτηση. Έχοντας τα παραπάνω υπόψη, η διεπαφή σχεδιάστηκε ώστε να είναι λιτή και κατανοητή.

Αξίζει να σημειωθεί ακόμη ότι το περιοδικό «Σπουδαί» βρίσκεται στα ράφια της Βιβλιοθήκης του Πανεπιστημίου Πειραιώς, αλλά και στο αρχείο της, σε πολλαπλά αντίτυπα, με εξαίρεση κάποια παλαιά τεύχη με ένα μόνο αντίτυπο.

Επιπλέον η φυσική κατάσταση της συλλογής είναι μέτρια έως πολύ καλή. Ενδεχομένως να χρειαστεί η αφαίρεση της βιβλιοδεσίας από παλαιότερους τόμους, οι οποίοι δεν έχουν σαρωθεί ακόμη, ώστε να μην καταστραφούν περαιτέρω.

Τέλος το σύνολο της συλλογής είναι σε έντυπη μορφή, ενώ τα μελλοντικά τεύχη θα ενσωματώνονται άμεσα στην Ψηφιακή Βιβλιοθήκη στο αντίστοιχο μορφότυπο (.pdf) όπως τα παρέχει η εκδοτική εταιρεία.

Διαδικασία ψηφιοποίησης

Τα στάδια ψηφιοποίησης του περιοδικού «Σπουδαί» περιλαμβάνουν:

- ❖ τη σάρωση («σκανάρισμα») των τευχών σε υψηλή ανάλυση (600 dpi) και την αποθήκευσή τους ως εικόνων τύπου JPEG,
- ❖ την οπτική αναγνώριση χαρακτήρων (OCR) στις σελίδες κειμένου. Αξίζει να σημειωθεί ότι τα γραφήματα και οι πίνακες που συνοδεύουν τα άρθρα έχουν διατηρήσει το μορφότυπο JPEG. Όλο το υπόλοιπο κείμενο είναι αναζητήσιμο και ανακτήσιμο λέξη προς λέξη (full text),
- ❖ τη διόρθωση του κειμένου, η οποία γίνεται ηλεκτρονικά με την υποστήριξη του λεξικού που διατρέχει το αντίστοιχο πρόγραμμα και εμπλουτίζεται καθημερινά. Ωστόσο οι διορθώσεις πραγματοποιούνται χειροκίνητα, αφού κάθε λέξη ελέγχεται από τον βιβλιοθηκονόμο,
- ❖ την ενσωμάτωση στην πλατφόρμα Dspace. Το Dspace είναι ένα σύστημα ψηφιακής βιβλιοθήκης που αναπτύχθηκε με τη συνεργασία των βιβλιοθηκών του Massachusetts Institute of Technology (MIT) και της εταιρείας Hewlett-Packard (HP).

- ❖ την περιγραφή του σαρωμένου υλικού που γίνεται μέσα στο ψηφιακό αποθετήριο Dspace με την εμπλουτισμένη έκδοση του προτύπου Dublin Core. Η περιγραφή γίνεται σε επίπεδο αντικειμένου (δηλαδή άρθρου),
- ❖ τον ποιοτικό έλεγχο του παραγόμενου αποτελέσματος.

Τα βασικά χαρακτηριστικά του λογισμικού Dspace που οδήγησαν στην υιοθέτησή του φαίνονται παρακάτω.

- **Υποστηριζόμενα πρότυπα διαλειτουργικότητας**

Το DSpace υποστηρίζει το Open Archives Initiative's Protocol for Metadata Harvesting (OAI-PMH) v2.0 και λειτουργεί ως data provider. Το Dspace δεν υποστηρίζει το πρότυπο Z39.50 αλλά υποστηρίζει το πρότυπο SRW / SRU (Search Retrieve Web Service / Search and Retrieve URL Service), το οποίο ουσιαστικά αποτελεί την εξέλιξη του Z39.50 και επιτρέπει στους χρήστες την έρευνα και την ανάκτηση αποτελεσμάτων από απομακρυσμένες βάσεις δεδομένων. Υποστηρίζεται το πρότυπο METS για την περιγραφή των ψηφιακών αντικειμένων και των μεταδεδομένων τους καθώς και το OpenURL Standard.

- **Υποστήριξη πολλαπλών γλωσσών και ιδιαίτερα Ελληνικών**

Σε επίπεδο μεταδεδομένων χρησιμοποιεί την κωδικοποίηση Unicode για την αποθήκευση των μεταδεδομένων. Σε επίπεδο ψηφιακού υλικού υποστηρίζεται το πρότυπο Unicode, οπότε το σύστημα μπορεί να υποστηρίξει υλικό σε διάφορες γλώσσες και αλφάβητα.

- **Τεχνολογίες υλοποίησης**

Το Dspace έχει δοκιμαστεί σε διάφορες πλατφόρμες Unix και Linux, καθώς επίσης σε MacOS X και Windows. Το Dspace έχει υλοποιηθεί σε γλώσσα Java. Μπορεί να χρησιμοποιηθεί με τις βάσεις δεδομένων PostgreSQL και Oracle. Μπορεί να χρησιμοποιηθεί οποιοσδήποτε web server. Το σύστημα χρησιμοποιεί τη δωρεάν (freeware) μηχανή αναζήτησης Lucene.

- **Δυνατότητα παραμετροποίησης του περιβάλλοντος**

Το Dspace είναι επεκτάσιμο καθώς πρόκειται για λογισμικό ανοικτού κώδικα. Όμως οι δυνατότητες παραμετροποίησης που προσφέρει είναι ελάχιστες, καθώς παρέχει έτοιμο περιβάλλον (μέσω του παγκόσμιου ιστού) για τη διαχείριση των συλλογών, των ψηφιακών αντικειμένων και χρηστών, το οποίο δεν μπορεί να παραμετροποιηθεί. Κάποιες δυνατότητες δίνονται από το σύστημα (μέσω API) για τη δημιουργία εφαρμογών που αφορούν στην καταχώρηση, επεξεργασία και παρουσίαση (μέσω ανάλογης διεπαφής χρήστη) των ψηφιακών αντικειμένων. Επιπλέον αλλαγές στο ήδη υπάρχον περιβάλλον συνεπάγονται επεμβάσεις σε επίπεδο κώδικα.

- **Δυνατότητα υποστήριξης / λήψης βοήθειας**

Υπάρχει αρκετά καλή τεκμηρίωση με εγχειρίδια (τεχνικά και εγκατάστασης) καθώς επίσης mailing lists, λίστες συζητήσεων και bug track/feature request system από το Sourceforge.

Η επιλογή ενός λογισμικού ανοιχτής πρόσβασης (Dspace) αποτέλεσε μια απόφαση για το Τμήμα Μηχανοργάνωσης του Πανεπιστημίου Πειραιώς που εργάστηκε για την παραμετροποίηση του κώδικα και την προσαρμογή του στις επιμέρους ανάγκες της συλλογής άρθρων του «Σπουδαι» που φιλοξενεί σήμερα. Προηγήθηκε η αξιολόγηση ανάλογων εργαλείων, όπως το Greenstone και το Eprints.

Η επιλογή που αναβάθμισε ποιοτικά το παραγόμενο αποτέλεσμα για το συγκεκριμένο έργο ψηφιοποίησης ήταν η χρήση ενός συστήματος οπτικής αναγνώρισης χαρακτήρων, προκειμένου τα τελικά αρχεία να περιλαμβάνουν κείμενο και όχι εικόνα. Αυτή η επιλογή αύξησε ιδιαίτερα το κόστος του έργου, κυρίως σε όρους χρόνου και ανθρώπινης παρέμβασης, αλλά έχει αξιόλογα αποτελέσματα στο επίπεδο της αναζήτησης και της ανάκτησης.

Χρήση προτύπων

Ο σχεδιασμός του έργου της ψηφιοποίησης στηρίχθηκε σε κείμενα καλών πρακτικών, αλλά και σε ευρέως αποδεκτά πρότυπα, προκειμένου να εξασφαλιστεί η διαλειτουργικότητα και η διατήρηση του υλικού.

Οι βασικές αρχές της σάρωσης στηρίχθηκαν στα κείμενα "Scanning Guidelines and Recommendations" του Missouri Digitization Project Committee (2001). Έτσι λήφθηκαν υπόψη όλοι οι παράγοντες που επηρεάζουν την ποιότητα της σάρωσης (bitmap, ανάλυση) και χρησιμοποιήθηκαν κατάλληλοι σαρωτές βαριάς χρήσης. Ακόμη χρησιμοποιήθηκε το πρότυπο JPEG για την συμπίεση των εικόνων, καθώς αυτό είναι καταλληλότερο για την παρουσίαση εικόνων μέσω του διαδικτύου.

Για τα μεταδεδομένα περιγραφής του σαρωμένου υλικού χρησιμοποιήθηκαν στοιχεία του προτύπου Dublin Core. Γενικά τα μεταδεδομένα που χρησιμοποιούνται από το Dspace ανήκουν σε τρεις κατηγορίες:

- περιγραφικά μεταδεδομένα,
- διαχειριστικά μεταδεδομένα, όπως μεταδεδομένα διατήρησης, προέλευσης και πρόσβασης,
- δομικά μεταδεδομένα, δηλαδή πληροφορίες για την παρουσίαση των άρθρων.

Τα πλέον συχνά χρησιμοποιούμενα μεταδεδομένα κατά τη δημιουργία εγγραφών στην εφαρμογή του «Σπουδαί» φαίνονται στον πίνακα που ακολουθεί:

Πεδίο DublinCore	Περιγραφή
dc.contributor.author	Συγγραφέας – πνευματική υπευθυνότητα
dc.date.accessioned	Ημερομηνία εισαγωγής του άρθρου
dc.date.available	Ημερομηνία διάθεσης του άρθρου
dc.date.issued	Ημερομηνία δημοσίευσης σε συμβατική μορφή
dc.identifier.uri	URI
dc.description	Περιγραφή (συνήθως περιλαμβάνει ταξινόμηση του άρθρου κατά JEL)
dc.description.abstract	Περίληψη
dc.language.iso	Γλώσσα του κειμένου κωδικοποιημένη κατά ISO
dc.publisher	Εκδότης
dc.subject	Θεματική περιγραφή με λέξεις-κλειδιά στα αγγλικά και στα ελληνικά
dc.title	Τίτλος
dc.type	Τύπος τεκμηρίου
dc.corporate.name	Συλλογικό όργανο το οποίο συνδέεται με την παραγωγή ή/και δημοσίευση ή/και διάθεση του άρθρου

Η θεματική περιγραφή γίνεται μέσα από λέξεις-κλειδιά και το ελεγχόμενο λεξιλόγιο που επιλέχθηκε είναι ο θησαυρός EUROVOC που παρέχεται από την Ευρωπαϊκή Ένωση στην ελληνική και στην αγγλική γλώσσα.

Ένας παράπλευρος στόχος που θα επιτευχθεί μέσα από τη διαδικασία ψηφιοποίησης είναι η διατήρηση του σαρωμένου υλικού στο χρόνο. Οι δύο μέθοδοι διατήρησης που εφαρμόζονται είναι η απομακρυσμένη αποθήκευση του υλικού σε CD-ROM και σε ηλεκτρονικό υπολογιστή (εκτός του server), καθώς και η χρήση κοινώς αποδεκτών και ευρέως υποστηριζόμενων προτύπων, όπως αναπτύχθηκαν στο παρόν κεφάλαιο.

Το έργο σήμερα

Το έργο ψηφιοποίησης του περιοδικού «Σπουδαί» έχει μέχρι στιγμής καλύψει τα τελευταία 20 χρόνια ύπαρξής του (1987-2006) παρακολουθώντας συγχρόνως τα τρέχοντα τεύχη και είναι διαθέσιμο μέσω της ιστοσελίδας του Πανεπιστημίου Πειραιώς (<http://digilib.lib.unipi.gr/spoudai>).

Το περιοδικό «Σπουδαί» στην ηλεκτρονική μορφή του έχει καταχωρηθεί στη μηχανή αναζήτησης ελληνικών ψηφιακών βιβλιοθηκών ανοικτής πρόσβασης "open archives".

Το «Σπουδαί» στο μέλλον

Όπως συμβαίνει σήμερα με τις περισσότερες περιοδικές εκδόσεις, οι συντάκτες και η επιτροπή του περιοδικού «Σπουδαί» θα κληθούν να αντιμετωπίσουν το ερώτημα σχετικά με το ποια μορφή θα έχει στο μέλλον. Τα σενάρια είναι τρία: έντυπη μορφή, ψηφιακή (online) μορφή ή παράλληλη ύπαρξη ψηφιακής και συμβατικής μορφής. Το αποθετήριο που δημιουργήθηκε με την ευκαιρία της ψηφιοποίησης εξασφαλίζει ότι και τα μελλοντικά τεύχη του περιοδικού θα ενσωματώνονται σε αυτό προκειμένου να διατηρηθούν στο διηνεκές και να είναι προσβάσιμα στο ευρύ κοινό. Επομένως μένει να αποφασιστεί η συνέχιση ή μη της έντυπης μορφής του.

Πάντως αξίζει να τονιστεί ότι η ψηφιακή υπόσταση ενός ακαδημαϊκού περιοδικού διευρύνει το κοινό του και ενδεχομένως προσελκύει αξιόλογους συγγραφείς. Επομένως το αξίωμα σύμφωνα με το οποίο εξασφαλίζοντας την ποιότητα των συγγραφέων, εξασφαλίζεται και η ποιότητα των αναγνωστών, βρίσκει εφαρμογή με την προσέλκυση ευρύτερου αναγνωστικού κοινού και σε αυτή την περίπτωση.

Συμπεράσματα

Τα έργα ψηφιοποίησης πληθαίνουν στον ελληνικό χώρο και γίνεται πλέον δυνατό, ακόμη και για μικρότερες σε μέγεθος βιβλιοθήκες, να αποκτήσουν μια ψηφιακή συλλογή του υλικού τους το οποίο θεωρούν ιδιαίτερης αξίας. Το κόστος έχει μειωθεί και η τεχνογνωσία ανθρώπων και οργανισμών ολοένα και εμπλουτίζεται από πολύτιμες εμπειρίες.

Η ψηφιοποίηση του περιοδικού «Σπουδαί» από τη Βιβλιοθήκη του Πανεπιστημίου Πειραιώς αποτέλεσε μια πολύτιμη εμπειρία για το προσωπικό που ενεπλάκη. Είναι γεγονός πως το άμεσο αποτέλεσμα του έργου βρίσκει καθημερινή ανταπόκριση στο κοινό της Βιβλιοθήκης χωρίς τη χρονοβόρα διαδικασία της φυλλομέτρησης τευχών. Η διατήρηση και η πρόσβαση αποτέλεσαν τους δύο κεντρικούς άξονες γύρω από τους οποίους περιστράφηκε το έργο. Η εργασία ενός έτους είχε ως αποτέλεσμα την επιτάχυνση των διαδικασιών και τη μείωση του συνολικού απαιτούμενου χρόνου για την προσθήκη ενός άρθρου στο ψηφιακό αποθετήριο.

Οι στόχοι που έχουν τεθεί για το 2008 είναι αφενός η προώθηση του έργου στο επιστημονικό και βιβλιοθηκονομικό κοινό, η βελτίωση της διεπαφής προκειμένου να γίνει πιο φιλική και η ολοκλήρωση της ψηφιοποίησης των παλιότερων τευχών.

Ευχαριστίες

Ευχαριστώ τους συναδέλφους κκ. Ανθή Κατσιρικού, Αναστάσιο Σαββόπουλο, Ρόζα Μαυροπόδη και Πηνελόπη Κατσιγιάννη για την παρακίνηση και τη στήριξη.

Βιβλιογραφία

Grogg, J. and B. Ashmore (2007), "Google book search libraries and their digital copies", *Searcher*, vol.15, no.4, διαθέσιμο στο: http://www.infotoday.com/searcher/apr07/Grogg_Ashmore.shtml, ανακτήθηκε στις 7-1-2008.

Μαγκλίνης, Η. και Ο. Σελλά, (2007), «Σπάνιοι θησαυροί σώζονται σε ψηφιακή μορφή», *Η Καθημερινή*, 14 Ιανουαρίου.

Τσιμπούλου, Φ. κ.ά. (2006), «Πρακτικές ψηφιοποίησης αρχείων και ανάδυση νέων πεδίων συνεργασίας Ακαδημαϊκών Βιβλιοθηκών», στο Τσάκωνας, Γ., *Proceedings 15^ο Πανελλήνιο Συνέδριο Ακαδημαϊκών Βιβλιοθηκών*, Πάτρα.