

Δημιουργία ψηφιακής συλλογής: καταχώρηση και βιβλιογραφικά πρότυπα συλλογών

Ανακοίνωση στην ημερίδα : «Οι ψηφιακές βιβλιοθήκες ως στοιχείο της ψηφιακής στρατηγικής της Ελλάδας», ΤΕΙ Δ. Μακεδονίας, Κοζάνη 12 Μαΐου 2006

Ανέστης Σίτας

Περίληψη

Περιγράφονται οι μεθοδολογίες, οι πρακτικές και τα εργαλεία της βιβλιοθηκονομίας που μπορούν να εφαρμοστούν για τη δημιουργία ψηφιακών συλλογών. Επισημαίνεται η χρήση προγραμμάτων ανοικτού κώδικα (open source). Αναφέρονται οι διαδικασίες επιλογής του υλικού, οι τεχνικές λεπτομέρειες διαχείρισης των ψηφιακών τεκμηρίων, τα μορφότυπα και τα πρότυπα βιβλιογραφικής περιγραφής τους. Γίνεται ιδιαίτερη μνεία στα κριτήρια επιλογής του υλικού τους, στους περιορισμούς που προκύπτουν από τα πνευματικά δικαιώματα και στις διαδικασίες απόκτησης των δικαιωμάτων χρήσης των τεκμηρίων. Γίνεται συνοπτική παρουσίαση των διαδικασιών που ακολουθήθηκαν για την υλοποίηση του έργου ψηφιοποίησης στο Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης.

Εισαγωγή

Από τους συνηθέστερους ορισμούς της βιβλιογραφίας για τον όρο «βιβλιοθήκη», είναι αυτός που την περιγράφει ως συλλογή βιβλίων και άλλου υλικού που προορίζεται για ανάγνωση, αναφορά, μελέτη και έρευνα, η οποία είναι οργανωμένη με τρόπο που να παρέχει πρόσβαση στους χρήστες και στελεχωμένη με βιβλιοθηκονόμους και άλλο προσωπικό, οι οποίοι είναι σε θέση να παρέχουν τις απαιτούμενες υπηρεσίες. Κοινό σημείο αναφοράς και χαρακτηριστικό όλων των βιβλιοθηκών ανά τους αιώνες, είναι η οριοθέτησή τους ως χώροι συγκέντρωσης, οργάνωσης και διαχείρισης υλικού καταγεγραμμένης γνώσης και πληροφοριών. Η παραδοσιακή/συμβατική βιβλιοθήκη οροθετείται από τον υλικό της συλλογής της (που είναι κυρίως έντυπο), από τον χώρο στον οποίο βρίσκεται (τον φυσικό της χώρο), τους χρήστες της (που έχουν πρόσβαση στον φυσικό χώρο της βιβλιοθήκης), από τις μορφές χρήσης της και από το εύρος των παρερχομένων υπηρεσιών.

Μία βιβλιοθήκη, που σε αντίθεση με την παραδοσιακή/συμβατική βιβλιοθήκη, το μεγαλύτερο μέρος του υλικού της βρίσκεται σε ψηφιακή μορφή, μπορεί να ονομαστεί «ψηφιακή βιβλιοθήκη». Πρέπει να διευκρινιστεί πως όταν αναφερόμαστε σε «ψηφιακό υλικό», αναφερόμαστε στο δημοσίευμα που βρίσκεται σε ψηφιακή μορφή και διακινείται, οργανώνεται, χρησιμοποιείται και αξιοποιείται αποκλειστικά στο περιβάλλον του Διαδικτύου. Η έννοια της ψηφιακής βιβλιοθήκης δεν ισοδυναμεί απλά με μια ψηφιοποιημένη συλλογή. Ουσιαστικά αφορά ένα ενιαίο περιβάλλον, μέσα στο οποίο ενώνονται συλλογές, υπηρεσίες και άνθρωποι με στόχο την υποστήριξη ολόκληρου του

κύκλου ζωής της δημιουργίας, της διάχυσης, της χρήσης και της συντήρησης των δεδομένων, της πληροφορίας και της γνώσης (Santa Fe planning workshop, 1997).

Λειτουργικά συστατικά ψηφιακής βιβλιοθήκης

Όλες οι ψηφιακές βιβλιοθήκες, δεν έχουν απαραίτητα τους ίδιους στόχους, ούτε φέρουν κοινές μορφές υλικού και οργάνωσης. Σε γενικές γραμμές όμως μπορούμε να ορίσουμε κάποια κοινά λειτουργικά συστατικά τους. Τα χαρακτηριστικά αυτά είναι τα παρακάτω (Rajashekar, 2002):

- **Επιλογή και απόκτηση:** επιλογή τεκμηρίων, ψηφιοποίηση και/ή μετατροπή στην κατάλληλη μορφή.
- **Οργάνωση:** απόδοση μεταδεδομένων σε κάθε ψηφιακό αντικείμενο.
- **Ευρετηρίαση και αποθήκευση (Indexing & storage):** ευρετηρίαση και αποθήκευση των ψηφιακών τεκμηρίων και των μεταδεδομένων τους ώστε να μπορούν να αναζητηθούν και να ανακτηθούν αποτελεσματικά.
- **Αποθετήριο (Repository):** είναι από τα βασικότερα συστατικά μιας ψηφιακής βιβλιοθήκης. Αφορά τα ψηφιακά αντικείμενα, τα μεταδεδομένα και τα ευρετήρια για την αναζήτηση και την ανάκτηση των διαθέσιμων τεκμηρίων.
- **Αναζήτηση και ανάκτηση (Search & retrieval):** το ξεφύλλισμα, η αναζήτηση, η ανάκτηση και η ανάγνωση/χρήση του περιεχομένου.
- **Ιστοσελίδα (Digital library website):** ο server που φιλοξενεί την ψηφιακή συλλογή και την παρουσιάζει στους χρήστες.
- **Συνδεσιμότητα δικτύου (Network connectivity):** ασφαλής και συνεχής σύνδεση στο Διαδίκτυο. Η πρόσβαση μπορεί να είναι ελεύθερη ή περιορισμένη.

Επιλογή υλικού προς ψηφιοποίησης σε ακαδημαϊκά περιβάλλοντα

Όταν αναφερόμαστε στην επιλογή υλικού προς ψηφιοποίηση ή στην συγκέντρωση πρωτογενούς ψηφιακού υλικού το οποίο θα διατίθεται στους χρήστες μιας βιβλιοθήκης, ο στόχος δεν είναι απλώς η επιλογή του κατάλληλου υλικού ούτε μόνο ο συνολικός αριθμός των καταχωρημένων τεκμηρίων. Σημαντικό ρόλο διαδραματίζουν δύο συνιστώσες: η διασφάλιση της οργάνωσης του υλικού αυτού και η δυνατότητα πρόσβασής του από τους χρήστες. Αυτές οι συνιστώσες επιτυγχάνονται δημιουργώντας τα κατάλληλα μεταδοδομένα και εφαρμόζοντας κατάλληλες πλατφόρμες και περιβάλλοντα εργασίας (Dempsey, 2006).

Όταν τίθεται σε εφαρμογή μια πολιτική επιλογής υλικού, για να είναι αποτελεσματική πρέπει (Digital Preservation Coalition, 2006):

- Να είναι αποδεκτή και εγκεκριμένη από τα ανώτερα όργανα του ιδρύματος.
- Η παρουσία της να είναι δυναμική σε όλο το εύρος τους ιδρύματος.
- Να αναθεωρείται σε τακτά χρονικά διαστήματα.
- Να υπάρχει διασφάλιση των οικονομικών πόρων για την βιωσιμότητα του έργου.

Ο στόχος της δημιουργίας ψηφιακών βάσεων δεδομένων, ιδιαίτερα δε όσον αφορά σε ένα εκπαιδευτικό ίδρυμα, στοχεύει: α) στην ανάδειξη των ειδικών και σπανίων συλλογών που διαθέτει (προστατεύοντας έτσι εμμέσως τα πρωτότυπα), β) στην αξιοποίηση των

αποτελεσμάτων του ερευνητικού έργου που συντελείτε και γ) στη δημιουργία υλικού που μπορεί να χρησιμοποιηθεί για την βελτίωση της εκπαιδευτικής διαδικασίας.

Επιλογή και προμήθεια λογισμικού

Για τη δημιουργία μιας ψηφιακής βάσης απαραίτητη προϋπόθεση είναι η υλοποίησή της με την εφαρμογή ενός λογισμικού. Τα λογισμικά που μπορούν να χρησιμοποιηθούν, σε ένα πρώτο επίπεδο, μπορούν να διακριθούν σε δύο κατηγορίες: στα εμπορικά και τα ελεύθερα (δωρεάν) λογισμικά.

Τα λογισμικά που διατίθενται από ιδιωτικές εταιρίες είναι πολλά και κάποια από αυτά πολύ αξιόλογα. Πλεονεκτήματα αυτής της κατηγορίας λογισμικών, είναι η υποστήριξή τους από την συγκεκριμένη εταιρία, αλλά και η δυνατότητα που προσφέρει ο πάροχός τους όσον αφορά στην εξέλιξη του προϊόντος του (στο βαθμό τουλάχιστον που αφορά σε εταιρίες που επιδιώκουν να κινούνται με αξιοπιστία στην αγορά). Από την άλλη πλευρά βέβαια, τα λογισμικά αυτά παρουσιάζουν ως βασικό μειονέκτημά τους, το κόστος αγοράς και συντήρησής τους και ασφαλώς την απόλυτη εξάρτηση του οργανισμού που τα επιλέγει, από την εταιρία διάθεσής τους.

Υπάρχει όμως και ένα άλλο είδος λογισμικών που είναι διαθέσιμα χωρίς καμία οικονομική επιβάρυνση. Πρόκειται για τα λεγόμενα λογισμικά ανοικτού κώδικα (open source) τα οποία διατίθενται δωρεάν, συνήθως από εκπαιδευτικά και ερευνητικά ιδρύματα. Τα λογισμικά αυτά, των οποίων η πρόσβαση στον κώδικά τους θεωρείται δεδομένη, αφορούν την ελευθερία των χρηστών τους (The Open Source Definition, 2006):

- να τα χρησιμοποιούν για οποιοδήποτε σκοπό θελήσουν
- να μελετούν την λειτουργία τους
- να τα προσαρμόζουν στις ιδιαίτερες ανάγκες τους
- να τα βελτιώνουν, αν χρειαστεί
- να τα αναδιανέμουν χωρίς καμία χρέωση, ώστε να συνεισφέρουν με την σειρά τους στην επιστημονική κοινότητα αλλά και στο ευρύτερο κοινωνικό σύνολο.

Επιλογή προτύπων και λογισμικού

Για να ξεκινήσει ένα έργο ψηφιοποίησης και ταυτόχρονα η επιλογή του λογισμικού που θα χρησιμοποιηθεί, απαιτείται η γνώση των διαθέσιμων προτύπων όσον αφορά:

- την οργάνωση του υλικού
- τη βιβλιογραφική περιγραφή του υλικού
- την αποθήκευση των ψηφιακών τεκμηρίων που θα δημιουργηθούν, και τέλος
- τις δυνατότητες μεταφερσιμότητας και μετατρεψιμότητας των βιβλιογραφικών εγγράφων και των ψηφιακών τεκμηρίων σε πιθανές μελλοντικές μορφές, δηλαδή τη διαλειτουργικότητα του λογισμικού και των εφαρμοζόμενων προτύπων.

Για την επιλογή του λογισμικού σε ένα έργο ψηφιοποίησης που υλοποιείται στο πλαίσιο ενός συγκεκριμένου ιδρύματος ή μιας βιβλιοθήκης, εξίσου σημαντικό ρόλο με τα εφαρμοζόμενα πρότυπα, παίζουν και οι υφιστάμενες γνώσεις, πρακτικές και πρότυπα που εφαρμόζονται από το προσωπικό της συγκεκριμένης βιβλιοθήκης. Ιδιαίτερα όσον

αφορά χρηματοδοτούμενα έργα, των οποίων η καθυστέρηση έγκρισής τους σε σχέση με τα αρχικώς προβλεπόμενα χρονοδιαγράμματα είναι αρκετά συχνό φαινόμενο, η ταχύτητα εφαρμογής τους επαφίεται και στην εκμετάλλευση των γνώσεων και των δεξιοτήτων του προσωπικού, στον βαθμό βέβαια που αυτά ταυτίζονται με την επίτευξη των στόχων της εφαρμογής των απαιτούμενων προτύπων.

Επιλογή λογισμικού στο Α.Π.Θ.

Για την υλοποίηση του έργου ψηφιοποίησης του Συστήματος Βιβλιοθηκών του Αριστοτελείου Πανεπιστημίου Θεσσαλονίκης, βασική επιδίωξη ήταν η εφαρμογή ενός λογισμικού ανοικτού κώδικα, το οποίο όμως να είναι σε θέση να διαθέτει και να ικανοποιεί συγκεκριμένα βασικά χαρακτηριστικά. Τα χαρακτηριστικά αυτά έπρεπε αφενός να καλύπτουν όλα τα απαραίτητα πρότυπα περιγραφής και αποθήκευσης των τεκμηρίων και αφετέρου να είναι όσο τον δυνατόν περισσότερο οικεία στους βιβλιοθηκονόμους και τους πληροφορικούς που θα τα εφαρμόσουν. Οι κυριότερες απαιτήσεις από ένα τέτοιο πρόγραμμα, ήταν η υποστήριξη των μορφοτύπων MARC21 (Machine Readable Cataloging) και DC (Dublin Core), καθώς και η υποστήριξη της XML (Extensible Markup Language) και του πρωτοκόλλου OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting).

Η απαίτηση της υποστήριξης του MARC21, προέρχεται από το γεγονός πως το μορφότυπο αυτό εφαρμοζόταν ήδη στο Α.Π.Θ. και παράλληλα εξασφαλίζει την διατήρηση των μεταδεδομένων και την μεταφορά τους σε πιθανά μελλοντικά λογισμικά. Όσον αφορά το DC, η υποστήριξή του κρίθηκε απαραίτητη αφού είναι το κοινό μορφότυπο που εξασφαλίζει τη διαλειτουργικότητα, την ανεύρεση και τη διαχείριση της περιγραφής των δεδομένων που απαιτούνται στο περιβάλλον του Web. Η υποστήριξη της γλώσσας XML κρίθηκε απαραίτητη, αφού ουσιαστικά είναι μια μεταγλώσσα (γλώσσα που περιγράφει άλλες γλώσσες) με την χρήση της οποίας μπορούν να οριστούν οι πληροφορίες με τρόπο περισσότερο ακριβή, ευέλικτο και προσαρμόσιμο στις ποικίλες ανάγκες και τα ποικίλα είδη τεκμηρίων που διαθέτει μια ψηφιακή συλλογή. Τέλος, επιθυμητή ήταν και η υποστήριξη του πρωτοκόλλου συγκομιδής μεταδεδομένων OAI-PMH, αφού αποτελεί το πρωτόκολλο που εξασφαλίζει το διαλειτουργικό πλαίσιο συγκομιδής των μεταδεδομένων ανεξάρτητα από την εφαρμογή στην οποία υλοποιείται.

Ουσιαστικός στόχος της διασφάλισης των προτύπων που προαναφέρθηκαν, ήταν η εξασφάλιση της δυνατότητας μεταφοράς όλων των εργασιών που γίνονται για την δημιουργία της βάσης σε μελλοντικά πρότυπα και λογισμικά, αλλά και η ουσιαστική συμμετοχή των ψηφιοποιημένων τεκμηρίων και των αντίστοιχων μεταδεδομένων τους στο πλαίσιο του Διαδικτύου. Εξ' αρχής θεωρήθηκε δεδομένο πως εάν το πρόγραμμα που θα επιλεγόταν, στην πορεία αποδειχθεί πως είναι ανεπαρκές ή εμφανιστεί ένα άλλο λογισμικό, περισσότερο σύγχρονο που να πλήρη το σύνολο των αναγκών του έργου, τότε το χρησιμοποιούμενο λογισμικό μπορεί να αλλάξει. Για αυτούς ακριβώς τους λόγους έπρεπε να υποστηρίξει όλα τα πρότυπα που εξασφαλίζουν την μεταφορά των τεκμηρίων και της βιβλιογραφικής περιγραφής τους, χωρίς καμιά απώλεια δεδομένων και με το μικρότερο δυνατό κόστος.

Σύμφωνα με τον παραπάνω σχεδιασμό, το έργο ξεκίνησε με την βιβλιογραφική επισκόπηση μεγάλου αριθμού διαθέσιμων λογισμικών. Η προσπάθεια αυτή κατέληξε στην «πρόκριση» πέντε λογισμικών, των οποίων ακολουθεί μια σύντομη περιγραφή.

Καθένα από τα λογισμικά αυτά διέθετε ορισμένες από τις επιθυμητές ιδιότητες ή πρότυπα. Διευκρινίζεται πως η επιλογή, άρα και τα αντίστοιχα χαρακτηριστικά, έγινε τον Δεκέμβριο του 2002 όταν ξεκίνησε το έργο της ψηφιοποίησης.

1. **LAURIN** (<http://laurin.uibk.ac.at/>). Πρόκειται για ένα πρόγραμμα χρηματοδοτούμενο από την Ευρωπαϊκή Ένωση και δύο υπουργεία της Αυστρίας. Αφορά περισσότερο την αυτόματη επιλογή και τη δημιουργία βάσης δεδομένων αποκομμάτων εφημερίδων. Αν και διανέμεται δωρεάν, δεν υπάγεται πλήρως στην κατηγορία των προγραμμάτων ανοικτού κώδικα. Τα βασικά χαρακτηριστικά του είναι τα εξής:
 - Υποστηρίζει αναζήτηση πλήρους κειμένου.
 - Δεν υποστηρίζει διατάξεις MARC, DC.
 - Δεν υποστηρίζει την γλώσσα XML.
 - Η υποστήριξη του παρέχεται μόνο με αμοιβή.
2. **ePrints 2** (<http://software.eprints.org/>). Το λογισμικό αυτό αναπτύσσεται από το University of Southampton της Μεγάλης Βρετανίας. Τα βασικά χαρακτηριστικά του είναι τα εξής:
 - Είναι συμβατό με το πρωτόκολλο συγκομιδής μεταδεδομένων OAI-PMH.
 - Υποστηρίζει αναζήτηση πλήρους κειμένου.
 - Δεν υποστηρίζει διατάξεις όπως MARC, DC.
 - Οι δημιουργοί του παρέχουν καλή υποστήριξη όσον αφορά την εγκατάσταση και την παραμετροποίησή του.
3. **ETD-db** (<http://scholar.lib.vt.edu/ETD-db/>). Πρόκειται για λογισμικό που αναπτύσσεται από το Virginia Tech των ΗΠΑ. Διατίθεται δωρεάν στα μέλη του NDLTD (Networked Digital Library of Theses and Dissertations). Τα βασικά χαρακτηριστικά του είναι τα εξής:
 - Δεν υποστηρίζει MARC, DC, XML.
 - Καλή τεχνική υποστήριξη.
4. **DSpace** (<http://www.dspace.org/>). Το λογισμικό αυτό δημιουργήθηκε με την συνεργασία του MIT (Massachusetts Institute of Technology) των ΗΠΑ και της Hewlett-Packard. Τα βασικά χαρακτηριστικά του είναι τα εξής:
 - Συμβατό με το πρωτόκολλο OAI-PMH.
 - Υποστηρίζει μόνο το DC.
 - Υποστηρίζει αναζήτηση πλήρους κειμένου.
5. **CDSware** (<http://cdsware.cern.ch/cdsware/>). Το λογισμικό αυτό αναπτύσσεται από το γνωστό ερευνητικό ίδρυμα που εδρεύει στην Ελβετία και στο οποίο αναπτύχθηκε το Web. Τα βασικά χαρακτηριστικά του είναι τα εξής:
 - Είναι συμβατό με το πρωτόκολλο OAI-PMH.
 - Το βασικό σχήμα μεταδεδομένων που χρησιμοποιεί είναι το MARC21.
 - Υποστηρίζει αναζήτηση πλήρους κειμένου.
 - Παρέχεται καλή τεχνική υποστήριξη από τους δημιουργούς του.

CERN Document Server Software

Μετά από την τελική αξιολόγηση όλων των λογισμικών και με βάση τα δεδομένα που υπήρχαν την συγκεκριμένη εποχή (Δεκέμβριος 2002), το λογισμικό το οποίο επιλέχθηκε από το Α.Π.Θ. είναι το **CDSware (CERN Document Server Software)**. Αυτό συγκέντρωνε τα περισσότερα επιδιωκόμενα χαρακτηριστικά, όπως αυτά τέθηκαν από τους βιβλιοθηκονόμους του συστήματος βιβλιοθηκών. Η βασική έκδοσή του είναι η 0.0.9, ενώ νεότερη διαθέσιμη είναι η 0.7.1, η οποία παρουσιάστηκε το Μάιο του 2005. Τα βασικά τεχνικά χαρακτηριστικά του είναι τα εξής:

- Λειτουργικό σύστημα: οποιοδήποτε Unix.
- Γλώσσα προγραμματισμού: Python και PHP.
- Βάση δεδομένων: MySQL.
- Web server: Apache.
- Χρησιμοποιεί δική του τεχνολογία ευρετηρίασης & μηχανή αναζήτησης.
- Όλοι οι clients HTML 4.0 και νεότεροι.

Παραμετροποίηση CDSware: δημιουργία και καταχώρηση μεταδεδομένων

Μετά από την επιλογή του λογισμικού, ακολούθησε το έργο της παραμετροποίησής του. Το CDSware χρησιμοποιεί ως δεδομένη μορφή καταχώρησης των βιβλιογραφικών εγγραφών το μορφότυπο MARC21. Αυτό συνεπάγεται την χρήση των AACR2 (Αγγλο-αμερικάνικοι Κανόνες Καταλογογράφησης). Οι βιβλιογραφικές εγγραφές σώζονται σε XML MARC και παρουσιάζονται αυτόματα σε:

- HTML format
- HTML MARC
- XML MARC
- XML DC

Η καθαρότητα ενός καταλόγου θεωρείται ο αποτελεσματικότερος τρόπος διατήρησης της ακεραιότητας, της λειτουργικότητας και της αποτελεσματικότητάς του. Ο κυριότερος τρόπος διατήρησης της καθαρότητας αυτής επιτυγχάνεται με την τήρηση των αρχείων καθιερωμένων όρων (ονόματα, θέματα, σειρές). Αν και το συγκεκριμένο σύστημα δεν διαθέτει υποσύστημα καθιερωμένων όρων, αποφασίστηκε η τήρησή τους με βάση τις αρχές και τα πρότυπα που εφαρμόζονται και στον βιβλιογραφικό κατάλογο του Συστήματος Βιβλιοθηκών.

Όσον αφορά τις καθιερώσεις των ονομάτων ελληνικών φυσικών προσώπων, αυτές βασίζονται στο «*Αρχείο καθιερωμένων επικεφαλίδων*» της Εθνικής Βιβλιοθήκης της Ελλάδος (εκδ. 2003). Τα ξενόγλωσσα ονόματα καθιερώνονται με βάση το «*Library of Congress Authorities*» (<http://authorities.loc.gov/>), η πρόσβαση στο οποίο είναι ελεύθερη. Στην περίπτωση που κάποιο όνομα δεν αναφέρεται στις παραπάνω πηγές, χρησιμοποιούνται τα αρχεία καθιερωμένων όρων που δημιουργούνται στον βιβλιογραφικό κατάλογο του Α.Π.Θ.

Όπως οι περισσότερες βιβλιοθήκες, σε διεθνές και σε εθνικό επίπεδο, οι βιβλιοθήκες του Α.Π.Θ. για την απόδοση θεματικών όρων χρησιμοποιούν τις LCSH (Θεματικές Επικεφαλίδες της Βιβλιοθήκης του Κογκρέσου) μεταφράζοντας τις στα ελληνικά. Ιδιαίτερα όσον αφορά την απόδοση θεματικών όρων στις διδακτορικές διατριβές, ακολουθείται η εξής πρακτική: οι θεματικοί όροι αποδίδονται και στην ελληνική και στην αγγλική γλώσσα. Επιπροσθέτως, οι ίδιοι οι συγγραφείς των διατριβών ορίζουν λέξεις-κλειδιά (στα ελληνικά και αγγλικά), τα οποία καταχωρούνται στο πεδίο 653 του MARC21.

Ψηφιακές Συλλογές Α.Π.Θ.

Οι προς ψηφιοποίηση συλλογές επιλέχθηκαν μετά από μελέτη που αφορούσε την διαθεσιμότητά τους και τις υπάρχουσες δυνατότητες πρόσβασης σε αυτές, αλλά και την χρησιμότητά τους όσον αφορά την χρήση τους από τους φοιτητές, το διδακτικό προσωπικό αλλά και τους ερευνητές. Στο πλαίσιο του Β' ΕΠΕΑΕΚ, Γ' ΚΠΣ, αλλά και του προγράμματος «Κοινωνία της Πληροφορίας», οι συλλογές που επιλέχθηκαν για το έργο της ψηφιοποίησης είναι οι εξής (αφορά τις συλλογές και το σύνολο των καταχωρημένων τεκμηρίων έως τον Μάιο του 2006):

- Συλλογή ψηφιοποιημένων άρθρων εφημερίδων και περιοδικών της περιόδου 1976-2000. Αναφέρονται σε εκδόσεις βιβλίων ανθρωπιστικών επιστημών (κυρίως δε λογοτεχνικών κριτικών και συνεντεύξεων συγγραφέων). Καταχωρήθηκαν 29.100 άρθρα.
- Συλλογή διδακτορικών διατριβών του Α.Π.Θ. Καταχωρήθηκαν 800 διατριβές
- Αρχαιακές συλλογές. Συνολικά έχουν καταχωρηθεί 24.070 τεκμήρια. Τα τεκμήρια αυτά ανήκουν στις παρακάτω συλλογές:
 - α) Φωτογραφικό αρχείο λαϊκής τέχνης (18-20ος αι.).
 - β) Ελληνικές εφημερίδες (πλήρη φύλλα) (περίοδος 1800-2000)
 - γ) Συλλογή Τριανταφυλλίδη
 - δ) Αρχείο Νεοελλήνων Λογοτεχνών
 - ε) Συλλογή Τρικόγλου
 - στ) Αρχαιολογικές ανασκαφές (άρθρα εφημερίδων και περιοδικών των ετών 1832-1932). Αναφέρονται σε αρχαιολογικές ανασκαφές που διενεργήθηκαν την συγκεκριμένη περίοδο.

Έως τον Μάιο του 2006 έχουν καταχωρηθεί συνολικά 54.000 εγγραφές, ενώ υπολογίζεται πως έχουν ψηφιοποιηθεί περίπου 267.000 σελίδες υλικού.

Οι παραπάνω συλλογές στην πρωτότυπη, έντυπη, μορφή τους δεν ήταν προσβάσιμες με τον απαιτούμενο τρόπο οργάνωσης στους χρήστες. Συγκεκριμένα, όσον αφορά στην ψηφιοποίηση των άρθρων από τις ελληνικές εφημερίδες και τα περιοδικά ποικίλης ύλης, ένας από τους σημαντικότερους λόγους που αποφασίστηκε η ψηφιοποίησή τους, ήταν η διάθεση σύγχρονου τεκμηριωμένου υλικού στους φιλόλογους της δευτεροβάθμιας εκπαίδευσης, απαραίτητο για την οργάνωση των μαθημάτων που αφορούν κείμενα σύγχρονων Ελλήνων λογοτεχνών, οι μελέτες για το έργο των οποίων δεν είναι ακόμη εύκολο να εντοπιστούν σε μια ολοκληρωμένη μονογραφία. Επιπλέον, η πλειονότητα των άρθρων τα οποία ψηφιοποιούνται (περίοδος 1976-2000), δεν είναι διαθέσιμα μέσω του Διαδικτύου από τους ίδιους τους εκδότες των εντύπων, αφού οι περισσότερες από αυτές

τις εφημερίδες και τα περιοδικά άρχισαν να διαθέτουν ηλεκτρονικά το υλικό τους γύρω στο έτος 2000.

Σχετικά με τις διδακτορικές διατριβές που εκπονούνται στο Α.Π.Θ., θα πρέπει να σημειωθεί πως ενώ καταχωρούνται οι βιβλιογραφικές εγγραφές όλων των διδακτορικών διατριβών που κατατέθηκαν μετά το έτος 2003, το πλήρη κείμενα είναι διαθέσιμο μόνο μετά από έγγραφη άδεια του συγγραφέα τους ή έπειτα από συγκεκριμένο χρονικό διάστημα που ορίζεται από τον ίδιο.

Πνευματικά δικαιώματα

Το πιο δυσεπίλυτο και χρονοβόρο πρόβλημα που αντιμετωπίζει η οργάνωση ενός έργου ψηφιοποίησης, είναι αυτό που αφορά τα πνευματικά δικαιώματα του υλικού που θα ψηφιοποιηθεί. Κανένα έργο ψηφιοποίησης δεν μπορεί να ξεκινήσει εάν πρώτα δεν έχει λυθεί το παραπάνω ζήτημα. Αυτός είναι ένας από τους κυριότερους λόγους που τα περισσότερα προγράμματα ψηφιοποίησης, αφορούν υλικό που βρίσκεται εκτός των χρονικών ορίων του copyright. Όσον αφορά το συγκεκριμένο έργο ψηφιοποίησης του Α.Π.Θ. και συγκεκριμένα τη συλλογή άρθρων εφημερίδων και περιοδικών ελληνικών εφημερίδων και περιοδικών ποικίλης ύλης, ζητήθηκε το δικαίωμα χρήσης τους για ψηφιοποίηση και ελεύθερη πρόσβαση από όλους τους χρήστες. Οι προσπάθειες έγιναν με ταχυδρομικές επιστολές, με ηλεκτρονικό ταχυδρομείο, με fax και με τηλεφωνική επαφή. Ο συνολικός χρόνος που διήρκησε αυτή η προσπάθεια ήταν δεκαέξι (16) μήνες. Πρέπει να σημειωθεί πως το 63,4% των εκδοτών απάντησαν θετικά στην αίτηση αυτή. Ψηφιοποιούνται μόνο τα άρθρα των εντύπων που ανταποκρίθηκαν θετικά. Συγκεκριμένα από το σύνολο των 64 τίτλων, 29.214 άρθρα (81%) ανήκουν μόνο στο 10 των εντύπων (15%), ενώ τα 5.508 άρθρα (19%) ανήκουν στους υπόλοιπους 54 τίτλους (85%).

Όσο δε αφορά την ελεύθερη διάθεση μέσω του Διαδικτύου των διδακτορικών διατριβών που εκπονούνται στο Α.Π.Θ. από το 2003 και έπειτα, το δικαίωμα χρήσης τους ζητείται με έγγραφη άδεια του συγγραφέα τους. Για τις υπόλοιπες αρχειακές συλλογές που ψηφιοποιούνται, σκανάρονται (ή δακτυλογραφείται το κείμενο) μόνο όσα τεκμήρια δεν έχουν περιορισμούς πνευματικών δικαιωμάτων.

Διάθεση υλικού και μελλοντικός σχεδιασμός

Η πρόσβαση στις ψηφιοποιημένες συλλογές του Αριστοτελείου Πανεπιστημίου Θεσσαλονίκης είναι ελεύθερη μέσα από τον δικτυακό τόπο του Συστήματος Βιβλιοθηκών «**Μέντωρ**» (<http://web.lib.auth.gr/portal/>) ή απευθείας στον δικτυακό τόπο των συλλογών αυτών, συγκεκριμένα στην διεύθυνση: <http://cds.lib.auth.gr>.

Με την εμπειρία που έχει αποκτηθεί μέχρι σήμερα, το Α.Π.Θ. διαθέτει ιδιαίτερη ευχέρεια στην υλοποίηση δράσεων ψηφιοποίησης. Ήδη προγραμματίζονται για ψηφιοποίηση και συλλογές, όπως αυτές των χειρογράφων που περιγράφονται στον κατάλογο του Λίνου Πολίτη και το αρχείο Αποδελτίωσης Τύπου με κριτικές θεατρικών έργων και παραστάσεων. Απώτεροι στόχοι είναι η δημιουργία και διάθεση υλικού στην ελληνική γλώσσα μέσα από το Διαδίκτυο, η δημιουργία υπηρεσιών ψηφιακής δημοσίευσης και τέλος η δημιουργία Ακαδημαϊκού Αποθετηρίου (Institutional Repository) του ιδρύματος (πού έχει ήδη ξεκινήσει και με την διάθεση των κειμένων των διδακτορικών διατριβών που εκπονούνται στο ίδρυμα μετά το 2003, αλλά και με την

δημιουργία ειδικής συλλογής που αφορά τα Επιστημονικά Δημοσιεύματα των μελών του Α.Π.Θ., ξεκινώντας από αυτά των μελών του Τομέα Αστροφυσικής, Αστρονομίας και Μηχανικής, του Τμήματος Φυσικής.

Βιβλιογραφία

1. CERN Document Server Software, <http://cdsware.cern.ch> [πρόσβαση 25-5-2006]
2. Dempsey, Lircan. The (Digital) Library Environment: ten years after, *Ariadne*, 46, March 2006. (<http://www.ariadne.ac.uk/issue46/dempsey/>)
3. Digital Preservation Coalition. *Decision Tree for Selection of Digital Materials for Long-term Retention*. (2006) (<http://www.dpconline.org/graphics/handbook/dec-tree.html> [πρόσβαση 25-5-2006])
4. Rajashekar, T.B. (2002) *Digital library and information services in enterprises* [<http://144.16.72.189/is214/is214-2005-01-04/topic-4.htm>]
5. Santa Fe planning workshop on distributed knowledge work environments, Digital Library Initiative, March 1997 <http://www.si.umich.edu/SantaFe/>.
6. The Open Source Definition (2006) <http://www.opensource.org/docs/definition.php>