

Σχεδιασμός του Ολοκληρωμένου Συστήματος Ψηφιακής Βιβλιοθήκης
του Πανεπιστημίου Αθηνών

Design of the Integrated Digital Library System of the University of
Athens

Κ. Βίγλας, Ε. Λουρδή, Μ. Νικολαΐδη, Γ. Πυρουνάκης, Κ. Σαΐδης
Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Υπολογιστικό Κέντρο Βιβλιοθηκών

C. Viglas, I. Lourdi, M. Nikolaidou G. Pyrounakis, K. Saidis
National and Kapodistrian University of Athens, Libraries Computer Center

Περίληψη

Στο πλαίσιο της παρούσας εισήγησης περιγράφεται η αρχιτεκτονική του Συστήματος Ψηφιακής Βιβλιοθήκης (ΣΨΒ) του Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών (ΕΚΠΑ), που αναπτύσσεται από το Υπολογιστικό Κέντρο Βιβλιοθηκών (ΥΚΒ). Οι ψηφιακές συλλογές του ΕΚΠΑ παρουσιάζουν ιδιαίτερο πολιτιστικό και επιστημονικό ενδιαφέρον και χαρακτηρίζονται από μεγάλη ετερογένεια. Αποτελούνται από διαφορετικού τύπου υλικό (έντυπα και χειρόγραφα έγγραφα, πάπυροι, εικόνες και φωτογραφίες, ηχητικά ντοκουμέντα, βίντεο, μουσειακά αντικείμενα καθώς και διαφόρων ειδών χειροτεχνήματα) το οποίο είναι είτε αμιγώς ψηφιακό είτε ψηφιοποιημένο. Ταυτόχρονα, είναι επιθυμητό η πρόσβαση και πλοήγηση στο περιεχόμενο των συλλογών να γίνεται με εύκολο και προσιτό τρόπο εξυπηρετώντας αποτελεσματικά τους διαφορετικούς στόχους κάθε κατηγορίας χρηστών. Για λόγους διαλειτουργικότητας, ολοκλήρωσης και αποδοτικότερης διαχείρισης των ψηφιακών συλλογών αποφασίστηκε να υλοποιηθεί ένα κεντρικό σύστημα Ψηφιακής Βιβλιοθήκης στο οποίο αποθηκεύονται όλες οι συλλογές. Η διαχείριση τους πραγματοποιείται κεντρικά, με ενιαίο τρόπο, διατηρώντας όμως τα ιδιαίτερα χαρακτηριστικά τους.

Αρχικά παρουσιάζονται οι προδιαγραφές και οι απαιτήσεις των ψηφιακών συλλογών, οι οποίες θεωρήθηκαν σημαντικότερες σύμφωνα με τις ανάγκες του ΕΚΠΑ. Με βάση αυτές τις προδιαγραφές, επελέγησαν τρία σύγχρονα ΣΨΒ, ευρέως διαδεδομένα και διαθέσιμα υπό καθεστώς ανοιχτού κώδικα: το DSpace, το Fedora και το Greenstone. Μέσα από την αξιολόγηση των συστημάτων αυτών προέκυψε η ανάγκη της περαιτέρω μελέτης του DSpace και του Fedora, καθώς κάλυπταν σε σχέση με το Greenstone τις σημαντικότερες απαιτήσεις των ψηφιακών συλλογών του ΕΚΠΑ. Η περαιτέρω σύγκριση των δύο συστημάτων βασίστηκε στη συλλογή Λαογραφίας, μία από τις απαιτητικότερες συλλογές του ΕΚΠΑ, που παρουσιάζει ιδιαίτερα σύνθετη δομή και ετερογένεια. Συγκεκριμένα, μελετήθηκε σε πιλοτικό επίπεδο η πληρότητα στην κάλυψη των ιδιαίτερων αναγκών της συλλογής Λαογραφίας από τα δύο συστήματα και το κόστος υλοποίησης που απαιτείται από το καθένα. Η μελέτη, σε συνδυασμό με τις γενικότερες απαιτήσεις και προδιαγραφές για το ΣΨΒ του ΕΚΠΑ, κατέληξε στην επιλογή του Fedora, λόγω της ευελιξίας και τις επεκτασιμότητας που το χαρακτηρίζουν.

Η παρούσα εισήγηση ολοκληρώνεται με την παρουσίαση της αρχιτεκτονικής του υπό ανάπτυξη συστήματος, επεκτείνοντας τη βασική λειτουργικότητα του Fedora.

Λέξεις κλειδιά: Αρχιτεκτονική συστήματος ψηφιακής βιβλιοθήκης, διαχείριση ετερογενών συλλογών, Fedora, DSpace.

Abstract

This paper presents the architecture of the Digital Library System (DLS) of the University of Athens (UoA), developed by Libraries Computer Center (LCC). UoA digital collections exhibit exceptional cultural and scientific interest, and are characterized as highly heterogeneous. They comprise of different types of material (printed documents, manuscripts, scrolls, pictures, photographs, audio tracks, videos, museum pieces, as well as various artifacts) which are either born digital or digitized. At the same time, it is desirable to access and browse through the content of the digital collections in such a manner, to cater efficiently for the demands of different user groups. To sustain interoperability, integration and efficient administration of the digital collections, it was decided to implement an integrated DLS to support all the collections. Collection management is performed in a uniform manner, while at the same time their individual characteristics are fully supported.

Most important specifications and requirements for all the collections are presented. Based on these specifications, three DLSs were chosen that are contemporary, well known, and distributed under an Open Source License: DSpace, Fedora, and Greenstone. The evaluation of these systems, led us to further investigate DSpace and Fedora since they provided a more complete coverage of the requirements of UoA's digital collections than Greenstone. This thorough comparison was based on the Folklore Collection, one of the most demanding of collections in terms of complexity of structure and variety of form. A pilot study was carried out regarding coverage of the collection's special requirements and cost of implementation for each of the two systems. The aforementioned study in combination with the general requirements and specifications of UoA's DLS, pointed out to choose Fedora over the "competition" mainly due to its advanced flexibility and expandability. The paper concludes with the presentation of the architecture of the DSL under development which extends Fedora's basic functionality.

Keywords: Digital library architecture, heterogeneous collection management, Fedora, DSpace.

Εισαγωγή

Το Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών (ΕΚΠΑ) ξεκίνησε ένα πρόγραμμα ανάπτυξης Ψηφιακής Βιβλιοθήκης για την υποδοχή εννέα διαφορετικών συλλογών με σκοπό την βελτίωση και τον εμπλουτισμό των εκπαιδευτικών του υπηρεσιών. Οι συλλογές αυτές χαρακτηρίζονται από μεγάλη ετερογένεια και περιλαμβάνουν ποικίλο υλικό, όπως κείμενα, παπύρους, φωτογραφίες, κοσμήματα, έργα λαογραφικής παράδοσης, πετρώματα, ακουστικό υλικό, αρχειακό υλικό κ.α. Το επιστημονικό και πολιτιστικό ενδιαφέρον που παρουσιάζουν είναι αρκετά μεγάλο, ώστε να υπαγορεύει τη διατήρησή τους και την προβολή του περιεχομένου αυτών στο

κοινό μέσω του διαδικτύου. Οι συλλογές που θα ψηφιοποιηθούν προέρχονται από διαφορετικά τμήματα του Πανεπιστημίου (Κλασικής Φιλολογίας, Γεωλογίας, Θεατρικών Σπουδών κ.α.) γεγονός που από τη μία εξηγεί την ποικιλία και ετερογένεια του υλικού και από την άλλη καθιστά ιδιαίτερα απαιτητική τη διαδικασία της ψηφιοποίησης και της περιγραφής του.

Το Υπολογιστικό Κέντρο Βιβλιοθηκών (ΥΚΒ) του ΕΚΠΑ έχει σχεδιάσει και αναπτύσσει ένα ενιαίο Σύστημα Ψηφιακής Βιβλιοθήκης (ΣΨΒ), για να υποστηρίξει όλες τις ψηφιακές συλλογές του Πανεπιστημίου Αθηνών, κυρίως για λόγους διαλειτουργικότητας, ολοκλήρωσης και αποδοτικότερης διαχείρισης τους. Το σύστημα έχει σχεδιαστεί σύμφωνα με τις απαιτήσεις του υλικού και τις προδιαγραφές που τέθηκαν ώστε η πρόσβαση και η πλοήγηση στο περιεχόμενο των συλλογών να γίνεται με εύκολο και προσιτό τρόπο για κάθε ενδιαφερόμενο, εξυπηρετώντας αποτελεσματικά τους διαφορετικούς στόχους κάθε κατηγορίας χρηστών.

Στόχος της εισήγησης αυτής είναι να παρουσιάσει την αρχιτεκτονική του ΣΨΒ και να τεκμηριώσει τις σημαντικές επιλογές που έγιναν κατά το σχεδιασμό του. Αρχικά παρουσιάζονται τα βασικά χαρακτηριστικά των συλλογών του Πανεπιστημίου Αθηνών και δίνεται ιδιαίτερη βάση στα χαρακτηριστικά της Συλλογής Λαογραφίας, η οποία είναι μία από τις πιο απαιτητικές συλλογές του ΕΚΠΑ. Ακολουθεί η αναφορά των προδιαγραφών του ΣΨΒ του Πανεπιστημίου Αθηνών, όπως αυτές καθορίστηκαν από το ΥΚΒ, με κύρια βάση τη συλλογή της Λαογραφίας. Στη συνέχεια, γίνεται μια σύντομη παρουσίαση των τριών συστημάτων Ψηφιακής βιβλιοθήκης που επιλέχθηκαν και συγκρίθηκαν με βάση τις προδιαγραφές αυτές και τεκμηριώνεται η επιλογή του Fedora για την ανάπτυξη του ΣΨΒ του ΕΚΠΑ. Η εισήγηση ολοκληρώνεται με την παρουσίαση της αρχιτεκτονικής του συστήματος ακολουθούμενης από τα σχετικά συμπεράσματα.

Χαρακτηριστικά των συλλογών του ΕΚΠΑ

Γενικά

Σε γενικές γραμμές οι συλλογές που θα ψηφιοποιηθούν και θα υποστηρίζονται από το ΣΨΒ χαρακτηρίζονται από τα παρακάτω σημεία:

- Στην πλειοψηφία τους, οι συλλογές δεν έχουν καταγραφεί ούτε και καταχωρηθεί σε κάποιο ηλεκτρονικό σύστημα. Το γεγονός αυτό έχει ως αποτέλεσμα το υλικό να είναι δύσκολα προσιτό από το ευρύ κοινό, εφόσον η χρήση αυτού γίνεται μόνο επιτόπου και συνήθως μετά από σχετική άδεια από τον υπεύθυνο φορέα της κάθε συλλογής.
- Οι συλλογές του ΕΚΠΑ χαρακτηρίζονται από ευρεία *ανομοιογένεια* και μεγάλη ποικιλία υλικού. Όπως ειπώθηκε, πρόκειται για συλλογές που αποτελούνται από διάφορα είδη υλικού: έντυπα και χειρόγραφα έγγραφα, παπύρους, εικόνες και φωτογραφίες, ηχητικά ντοκουμέντα, βίντεο, μουσειακά αντικείμενα, διαφόρων ειδών χειροτεχνήματα κ.α. Έτσι, το υλικό, σε ότι αφορά στην ψηφιακή του μορφή, μπορούμε να πούμε ότι θα είναι είτε αμιγώς ψηφιακό (born digital) είτε ψηφιοποιημένο (digitized).
- Επίσης, οι περισσότερες συλλογές χαρακτηρίζονται από *πολυπλοκότητα* στη δομή. Η κάθε συλλογή έχει τη δική της εσωτερική δομή και ιεραρχία, ενώ διαμορφώνονται διαφορετικές συσχετίσεις μεταξύ των αντικειμένων της.

- Λόγω του ότι οι συλλογές προέρχονται από διαφορετικά τμήματα του ΕΚΠΑ είναι εύλογο να καλύπτουν πολλές κατηγορίες χρηστών με ξεχωριστά ενδιαφέροντα και απαιτήσεις.
- Τέλος η ξεχωριστή φύση και αξία κάθε πηγής ορίζει κάθε συλλογή να διέπεται από διαφορετικές πολιτικές πρόσβασης και χρήσης του περιεχομένου αυτής, που πολλές φορές προκύπτει και από το πλαίσιο του κανονισμού λειτουργίας του Τμήματος στο οποίο ανήκει.

Στη συνέχεια ακολουθεί μια εκτενής παρουσίαση της συλλογής Λαογραφίας, που φαίνεται να υπερκαλύπτει σε απαιτήσεις τις υπόλοιπες συλλογές και ουσιαστικά χρησιμοποιήθηκε ως πρότυπη συλλογή τόσο στον καθορισμό των προδιαγραφών του ΣΨΒ όσο και στο σχεδιασμό του.

Συλλογή Λαογραφίας

Δομή και χαρακτηριστικά

Η συλλογή Λαογραφίας αποτελεί πολύτιμη πηγή μελέτης, εξερεύνησης και αξιολόγησης της παράδοσης της Ελλάδας, δεδομένου ότι αποτελείται από ανόθευτα και πηγαία εγχώρια εθνοχαρακτηριστικά διαφόρων χωριών και περιοχών της χώρας. Η συλλογή απαρτίζεται από ποικίλα είδη υλικού, ενώ το περιεχόμενό της παραμένει στην πλειοψηφία του μη καταγεγραμμένο σε κάποιο υπολογιστικό σύστημα. Το γεγονός αυτό έχει ως αποτέλεσμα η πρόσβαση σε αυτό να γίνεται μόνο με επιτόπια επίσκεψη και έτσι να αποκλείεται ένας μεγάλος αριθμός χρηστών, που θα μπορούσαν να βρουν πολύτιμα στοιχεία μελέτης και έρευνας.

Η συλλογή Λαογραφίας αποτελείται από δύο κύρια μέρη: α) εργασίες φοιτητών που έχουν εκπονηθεί ύστερα από επίσκεψη σε διάφορα μέρη της Ελλάδας και β) αντικείμενα που εκτίθενται στη βιβλιοθήκη του Τομέα Λαογραφίας του Τμήματος Κλασικής Φιλολογίας. Οι εργασίες αποτελούνται κυρίως από γραπτό κείμενο, το οποίο συνήθως συνοδεύεται από χάρτες ή φωτογραφίες της περιοχής που επισυνάπτονται καθώς και από μικρά αντικείμενα όπως είδη υφάσματος, σπόρους κ.α. Υπάρχουν, επίσης, περιπτώσεις όπου οι εργασίες συνοδεύονται από κασέτες που περιέχουν ηχογραφήσεις και συνεντεύξεις. Αυτά τα ιδιαίτερα στοιχεία που αποτελούν μια εργασία θα πρέπει να αντιμετωπιστούν ως ξεχωριστά αντικείμενα τα οποία ανήκουν σε συγκεκριμένες υπο-συλλογές. Αυτές είναι:

- Υπο-συλλογή εργασιών. Η κάθε εργασία αποτελείται από συγκεκριμένα κεφάλαια και υποκεφάλαια και τον πίνακα περιεχομένων, καθώς και από φωτογραφίες και αντικείμενα προερχόμενα από την περιοχή την οποία έχει ως θέμα μελέτης.
- Υπο-συλλογή φωτογραφιών. Αποτελείται από τις φωτογραφίες που βρίσκονται στις εργασίες των φοιτητών
- Υπο-συλλογή ηχητικού υλικού. Αποτελείται από κασέτες με συνεντεύξεις, τραγούδια, μύθους κ.α. τα οποία σχετίζονται με τις εργασίες των φοιτητών.
- Υπο-συλλογή αντικειμένων που εκτίθενται στη βιβλιοθήκη.

Απαιτήσεις συλλογής

Οι απαιτήσεις που υπαγορεύονται από την ιδιαίτερη φύση και την ετερογένεια του υλικού της συλλογής, με κύριο γνώμονα τη διευκόλυνση του χρήστη της, είναι οι παρακάτω:

Α) *Περιγραφή σε επίπεδο συλλογής*: Είναι σημαντικό για το χρήστη της συλλογής να διαθέτει τη γενική περιγραφή της συλλογής, ώστε να είναι ευκολότερη η αναζήτηση και ανάκτηση πληροφοριών για το περιεχόμενο της (Dempsey 1999). Ο χρήστης με αυτό τον τρόπο μπορεί να δει τι περιέχει η συλλογή, να τη συγκρίνει με άλλες και να αποφασίσει αν ενδιαφέρεται περαιτέρω για τα αντικείμενα αυτής. Η περιγραφή της συλλογής θα πρέπει να περιέχει στοιχεία για τα θέματα που πραγματεύεται, το μέγεθος αυτής, το σκοπό αυτής κ.α.

Β) *Υποστήριξη υποσυλλογών*: Λόγω της μεγάλης ποικιλίας του υλικού και των πολύπλοκων συσχετίσεων των πηγών λαογραφίας θα πρέπει να γίνει διαχωρισμός της συλλογής σε υπο-συλλογές με βάση διάφορα κριτήρια. Με αυτό τον τρόπο η πλοήγηση στο περιεχόμενο της συλλογής γίνεται πιο εύκολη από το ένα επίπεδο στο άλλο και παράλληλα ορίζονται τα στοιχεία που κληρονομούνται από τη συλλογή στις υπο-συλλογές και στα αντικείμενα.

Γ) *Παρουσίαση σύνθετων αντικειμένων*: Στην τωρινή κατάσταση ο χρήστης είναι υποχρεωμένος να διαβάζει ολόκληρη την εργασία προκειμένου να βρει τις πληροφορίες που θέλει. Βασικός στόχος στη σχεδίαση του ΣΨΒ είναι η ελαχιστοποίηση αυτού του φαινομένου, όπου αυτό είναι εφικτό, για τη «ψηφιακή μορφή» της συλλογής. Δεδομένου ότι η συλλογή θα καταχωρηθεί στο σύστημα με τη μορφή εικόνων και όχι κειμένου, ο χρήστης δεν θα έχει τη δυνατότητα ελεύθερης αναζήτησης κειμένου (free text search). Για το λόγο αυτό, θα πρέπει η παρουσίαση κάθε εργασίας / αντικειμένου της συλλογής να εμπλουτιστεί με όσο το δυνατόν περισσότερη πληροφορία για τη δομή και τα περιεχόμενά της. Κάθε εργασία, επομένως, θεωρείται ένα σύνθετο αντικείμενο αποτελούμενο από διάφορα επιμέρους τμήματα (κεφάλαια, υποκεφάλαια, φωτογραφίες κλπ) τα οποία σχετίζονται μεταξύ τους με κάποιον κανόνα (δομή ιεραρχείας, σελίδα – φωτογραφία κλπ). Κάθε ένα από αυτά τα αντικείμενα θα πρέπει να περιγράφεται και να χαρακτηρίζεται ξεχωριστά με τα κατάλληλα μεταδεδομένα ώστε ο χρήστης να παραπέμπεται απευθείας σε αυτό που ψάχνει.

Δ) *Περιγραφή συσχετίσεων*: Είναι απαραίτητο να περιγραφούν όλες οι συσχετίσεις που πιθανόν υφίστανται ανάμεσα στα αντικείμενα μιας συλλογής αλλά και συγκριτικά με τις υπόλοιπες συλλογές.

Ε) *Υποστήριξη διαφορών ειδών μεταδεδομένων*: Λόγω της ετερογένειας της συλλογής, η περιγραφή του υλικού θα πρέπει να γίνεται βάσει διαφορών σχημάτων μεταδεδομένων, ανάλογα με την κάθε περίπτωση. Για παράδειγμα θα πρέπει να χρησιμοποιηθεί το Dublin Core (Dublin Core Metadata Initiative 2003) για τη γενική περιγραφή των πηγών αλλά και το LOM (IEEE 2002) για την περιγραφή του εκπαιδευτικού χαρακτήρα αυτών. Επίσης, θα μπορεί να γίνει χρήση μεταδεδομένων για την περιγραφή των φωτογραφιών ή των αντικειμένων και για την περιγραφή των θεμάτων ή των τεχνικών διαδικασιών ψηφιοποίησης. Είναι σημαντικό να αποδοθεί πλήρως και με κάθε λεπτομέρεια το περιεχόμενο των πηγών και οτιδήποτε έχει σχέση με τη διατήρησή τους χωρίς να υπάρξει απώλεια πληροφορίας.

Προδιαγραφές του Συστήματος Ψηφιακής Βιβλιοθήκης του ΕΚΠΑ

Με βάση τα χαρακτηριστικά των συλλογών που αναφέρθηκαν παραπάνω και χρησιμοποιώντας τη συλλογή Λαογραφίας ως πρότυπο, οριοθετήθηκαν οι απαιτήσεις και τέθηκαν οι βασικές προδιαγραφές του ΣΨΒ του Πανεπιστημίου Αθηνών.

Το ΣΨΒ πρέπει να υποστηρίζει πολλαπλές ψηφιακές συλλογές, που αποτελούνται από διαφορετικές κατηγορίες αντικειμένων, με ετερογενή και ανομοιόμορφα χαρακτηριστικά. Παρόλα αυτά, το ΣΨΒ πρέπει να παρέχει μια ενιαία διεπαφή προς τους τελικούς χρήστες (που είναι η ευρύτερη ακαδημαϊκή κοινότητα), μέσα από τον παγκόσμιο ιστό. Το σύστημα, επομένως, πρέπει να είναι παραμετροποιήσιμο σε μεγάλο βαθμό για να μπορεί να προσαρμοστεί στις ανάγκες της εκάστοτε συλλογής, υπάρχουσας ή μελλοντικής. Βασική απαίτηση είναι να υποστηρίζει ανοικτά πρότυπα σε όλους τους τομείς, έτσι ώστε να εξασφαλίζεται η βιωσιμότητα του, η δυνατότητα αναβάθμισης του από το ΥΚΒ, η μελλοντική μετάπτωση του σε άλλο σύστημα και η συνεργασία του με άλλες εφαρμογές λογισμικού ή ανάλογα συστήματα ψηφιακής βιβλιοθήκης.

Πρέπει να υποστηρίζεται οποιοδήποτε πρότυπο μεταδεδομένων ή επεκτάσεις του, καθώς και τοπικά σύνολα μεταδεδομένων που μπορεί να είναι απαραίτητα για μια συλλογή. Ένα βασικό πρότυπο που υποστηρίζεται από την πλειοψηφία των ΣΨΒ και είναι επιθυμητό ως βασικό σύνολο μεταδεδομένων για τις ψηφιακές συλλογές του ΕΚΠΑ είναι το Dublin Core. Επίσης απαιτείται η δυνατότητα περιγραφής της συλλογής με κάποιο πρότυπο μεταδεδομένων, όπως είναι το Dublin Core Collection Description Application Profile (2003).

Είναι επιθυμητό, το ΣΨΒ του ΕΚΠΑ να είναι όσο το δυνατό πιο συμβατό με τα υπάρχοντα πρότυπα διαλειτουργικότητας μεταξύ διαφορετικών Ψηφιακών Βιβλιοθηκών. Υπάρχει η προοπτική συνεργασίας της Ψηφιακής Βιβλιοθήκης του ΕΚΠΑ με εκείνες άλλων Ακαδημαϊκών Ιδρυμάτων, όπως ήδη συμβαίνει και με ανάλογες ψηφιακές υπηρεσίες των βιβλιοθηκών. Γι' αυτό το λόγο, πρέπει να υποστηριχθεί το πρότυπο Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) (Lagoze 2001), που επιτρέπει το διαμοιρασμό των μεταδεδομένων μεταξύ Ψηφιακών Βιβλιοθηκών.

Το ψηφιακό υλικό πρέπει να αποθηκεύεται σε διαφορετικές μορφές, ανάλογα με τη χρήση για την οποία προορίζεται: υλικό υψηλής ποιότητας για λόγους εσωτερικής χρήσης από το Πανεπιστήμιο και χαμηλής ποιότητας για εύκολη πρόσβαση από το διαδίκτυο (McCray 2001). Ένας επιπλέον λόγος, για την ύπαρξη πολλών μορφών ψηφιακού υλικού, είναι η επίτευξη της διατηρησιμότητας (preservation). Αυτό σημαίνει ότι για κάθε ψηφιακό υλικό πρέπει να εξασφαλίζεται η διάρκεια του στο χρόνο και αυτό επιτυγχάνεται με την αποθήκευση του σε μια μορφή, η οποία έχει πολλές πιθανότητες να παραμείνει ενεργή στο μέλλον. Η συγκεκριμένη μορφή μπορεί να μη χρησιμεύει για την παρουσίαση του ψηφιακού υλικού στους χρήστες της Ψηφιακής Βιβλιοθήκης, αλλά να αποτελεί το πρωτότυπο για τη δημιουργία άλλων μορφών. Συνήθως για το σκοπό αυτό χρησιμοποιούνται τύποι αρχείων που αποτελούν ανοικτά πρότυπα, παρέχουν υψηλή ποιότητα του ψηφιακού υλικού, είναι απλά στην υλοποίηση τους και έχουν αποδείξει την αντοχή τους στο χρόνο. Το ψηφιακό υλικό, που περιλαμβάνουν οι συλλογές του Πανεπιστημίου Αθηνών, αποτελείται από αρχεία που ανήκουν σε ανοικτά πρότυπα μορφών αποθήκευσης. Πιο συγκεκριμένα, για την αποθήκευση φωτογραφιών πρόκειται να χρησιμοποιηθούν οι μορφές αρχείων TIFF για υψηλή ποιότητα και JPEG για χαμηλότερη. Για έντυπα έγγραφα, που θα υποστούν ψηφιοποίηση, θα χρησιμοποιηθούν οι τύποι αρχείων TIFF και JPEG, ενώ για αμιγώς ψηφιακά έγγραφα ο τύπος αρχείου PDF. Για την

αποθήκευση ηχητικού υλικού υψηλής ποιότητας θα χρησιμοποιηθεί ο τύπος αρχείου WAV, ενώ για χαμηλότερη ποιότητα και χρήση από το Διαδίκτυο ο τύπος αρχείου MP3. Τέλος, για video θα χρησιμοποιηθεί η κωδικοποίηση MPEG-1 και MPEG-2.

Πρέπει να υποστηρίζονται διαφορετικά επίπεδα ασφάλειας και πιστοποίησης των χρηστών, που καθορίζουν την πρόσβαση και τις ενέργειες που μπορεί να πραγματοποιήσει ένας χρήστης. Πιο συγκεκριμένα, πρέπει να ορίζονται δικαιώματα σε επίπεδο ψηφιακού αντικειμένου, συλλογής και ολόκληρης της Ψηφιακής Βιβλιοθήκης. Στο κάθε επίπεδο μπορεί να οριστεί ποιοι χρήστες ή ομάδες χρηστών έχουν πρόσβαση ή δεν επιτρέπεται να έχουν πρόσβαση.

Το ΣΨΒ πρέπει να υποστηρίζει αρχιτεκτονική υπηρεσιών τριών επιπέδων: αποθήκευσης, διαχείρισης και παρουσίασης. Κάθε επίπεδο διαθέτει σαφώς διαχωρισμένες υπηρεσίες και ανάλογη προγραμματιστική διεπαφή (API), για την επικοινωνία με τα γειτονικά επίπεδα και τις εξωτερικές εφαρμογές. Η διεπαφή χρήστη πρέπει να είναι ανεξάρτητη από τη λειτουργικότητα της εφαρμογής, έτσι ώστε να είναι προσιτή η ολοκλήρωση του ΣΨΒ με τον υπάρχον δικτυακό τόπο των βιβλιοθηκών.

Το ΣΨΒ πρέπει οπωσδήποτε να υποστηρίζει πολλαπλές γλώσσες με βασική την ελληνική. Για να υποστηρίζει ένα σύστημα την ελληνική γλώσσα, πρέπει καταρχήν να δέχεται τη χρήση ελληνικών χαρακτήρων στα μεταδεδομένα. Έπειτα, πρέπει να λειτουργεί σωστά η αναζήτηση μεταδεδομένων που περιέχουν ελληνικούς χαρακτήρες. Τέλος, πρέπει οι πληροφορίες που προέρχονται από τα μεταδεδομένα και περιέχουν ελληνικούς χαρακτήρες, να μπορούν να παρουσιαστούν σωστά, μέσω της διεπαφής, στον τελικό χρήστη. Εκτός αυτών, είναι επίσης απαραίτητο να παρέχεται η δυνατότητα υλοποίησης πολύγλωσσης διεπαφής. Η ιδανική λύση, που προτείνεται για την πλήρη υποστήριξη της ελληνικής αλλά και των υπόλοιπων γλωσσών σε διεθνές επίπεδο, είναι η χρήση του προτύπου Unicode, στην κωδικοποίηση των μεταδεδομένων.

Το ΣΨΒ πρέπει να είναι ευέλικτο και επεκτάσιμο, έτσι ώστε να επιτρέπει την προσθήκη επιπλέον λειτουργικότητας και την παραμετροποίηση του με προκαθορισμένο τρόπο. Ένα βασικό χαρακτηριστικό που προκύπτει από αυτή την απαίτηση είναι να διατίθεται ως λογισμικό ανοικτού κώδικα (Open Source Software 2004), για να υπάρχει η δυνατότητα της ελεύθερης επέκτασης και προσαρμογής του. Ως προτιμότερη γλώσσα προγραμματισμού για την ανάπτυξη του ΣΨΒ, είναι η Java, λόγω της ευκολίας ολοκλήρωσης του με τα ήδη υπάρχοντα συστήματα του ΥΚΒ.

Σύγκριση Συστημάτων Ψηφιακής Βιβλιοθήκης

Το θέμα της επιλογής του κατάλληλου συστήματος για την υποδοχή και υποστήριξη των συλλογών του ΕΚΠΑ είναι θεμελιώδους σημασίας για την επιτυχία της Ψηφιακής Βιβλιοθήκης. Δυστυχώς, οι απαιτήσεις των συλλογών του Παν/μίου, όπως προδιαγράφηκαν παραπάνω, δεν καλύπτονται από κανένα υπάρχον ΣΨΒ. Υπάρχουν διάφορα συστήματα που με τον ένα ή τον άλλο τρόπο καλύπτουν μικρό μέρος των προδιαγραφών σε σχετικά ικανοποιητικό επίπεδο, αλλά κανένα τους δεν φαίνεται να ικανοποιεί όλες τις απαιτήσεις. Για το λόγο αυτό, δίνεται μεγάλη βάση στην επεκτασιμότητα και ευελιξία που προσφέρει το κάθε σύστημα, καθώς είναι βέβαιο ότι μεγάλο μέρος της απαιτούμενης λειτουργικότητας θα υλοποιηθεί από το ΥΚΒ. Τέλος, ένα ακόμα σημαντικό σημείο της διαδικασίας επιλογής ΣΨΒ, είναι η απαίτηση το σύστημα να διατίθεται ως λογισμικό ανοιχτού κώδικα (Open Source Software),

καθώς σε οποιαδήποτε άλλη περίπτωση, δεν προσφέρονται, ουσιαστικά, δυνατότητες επέκτασης.

Σύμφωνα με τη λίστα συστημάτων Ψηφιακής Βιβλιοθήκης που προτείνει το Open Society Institute (2004) επιλέγονται δύο συστήματα, τα οποία καλύπτουν τις περισσότερες προδιαγραφές που έχουν τεθεί: το DSpace (DSpace Federation 2004) (Smith 2003) και το Fedora (Fedora Project 2004) (Staples 2003). Παράλληλα, αξιολογήθηκε και το σύστημα Greenstone (Greenstone DL Software 2004) (Witten 2000), το οποίο είναι ευρέως διαδεδομένο και χρησιμοποιείται από αρκετούς οργανισμούς.

Greenstone

Το Greenstone είναι ένα σύστημα για την κατασκευή και διάθεση ψηφιακών συλλογών. Έχει δημιουργηθεί από το University of Waikato στο πλαίσιο του New Zealand Digital Library Project. Το Greenstone είναι διαθέσιμο ως εφαρμογή ανοικτού κώδικα, μέσω του GNU General Public Licence, για λειτουργικά συστήματα Unix/Linux και Windows και έχει υλοποιηθεί σε γλώσσα C και Perl. Βασική οντότητα στο Greenstone είναι η συλλογή. Κάθε συλλογή διαθέτει ορισμένα χαρακτηριστικά, όπως είναι το ψηφιακό υλικό που μπορεί να περιέχει, τα ευρετήρια που παρέχει και στοιχεία για την παρουσίαση του υλικού. Η εισαγωγή του ψηφιακού υλικού στο σύστημα, γίνεται με τη χρήση ειδικών φίλτρων (plugins), που μετατρέπουν το εισαγόμενο αρχείο σε ένα έγγραφο που διαθέτει τη μορφή Greenstone Archive Format. Αν το ψηφιακό υλικό περιέχει αρχεία εικόνων ή οτιδήποτε άλλο εκτός κειμένου, τότε τα αρχεία αυτά εισάγονται ως είναι και δημιουργούνται σύνδεσμοι από το έγγραφο σε αυτά. Το βασικό πρότυπο μεταδεδομένων που χρησιμοποιείται είναι το Dublin Core.

Λόγω των ελλείψεων του Greenstone σε βασικά χαρακτηριστικά, το συγκεκριμένο ΣΨΒ κρίνεται ανεπαρκές, για την υποστήριξη των ψηφιακών συλλογών του ΕΚΠΑ. Συγκεκριμένα, το Greenstone δεν προβλέπει την υποστήριξη λειτουργιών για την επίτευξη της διατηρησιμότητας του ψηφιακού υλικού, αλλά αντίθετα όλο το υλικό αποθηκεύεται στη μορφή του Greenstone Archive Format. Επίσης, το Greenstone δεν διαθέτει επαρκείς μεθόδους για τη διαχείριση των μεταδεδομένων και του ψηφιακού υλικού. Δεν υποστηρίζει το πρωτόκολλο διαλειτουργικότητας OAI-PMH, που χρησιμοποιείται από τα περισσότερα ΣΨΒ, αλλά βασίζεται στο παλαιότερο Z39.50. Τέλος, το πιο σημαντικό μειονέκτημα του Greenstone, είναι οι ελάχιστες δυνατότητες παραμετροποίησης και η ελλιπής προγραμματιστική διεπαφή που παρέχει.

Επιλογή Συστήματος Ψηφιακής Βιβλιοθήκης

Η αξιολόγηση των δύο συστημάτων, του DSpace και του Fedora, βασίστηκε στα χαρακτηριστικά της συλλογής Λαογραφίας. Συγκεκριμένα, μελετήθηκε κατά πόσο τα συστήματα αυτά είναι σε θέση να υποστηρίξουν τη συγκεκριμένη συλλογή «out of the box», χωρίς δηλαδή την παραμικρή επέκταση. Παρ' όλο που το DSpace φαίνεται να προσφέρει περισσότερα λειτουργικά χαρακτηριστικά, το κόστος (από άποψη υλοποίησης και μελλοντικής συντήρησης) που απαιτείται για την πλήρη υποστήριξη όλων των συλλογών σε αυτό είναι τελικά κατά πολύ μεγαλύτερο από αυτό που απαιτείται από το Fedora.

DSpace

Το DSpace είναι ένα σύστημα ΨΒ για Πανεπιστήμια, που αναπτύχθηκε με τη συνεργασία του MIT και της εταιρίας HP. Βασικός στόχος του DSpace είναι η παροχή υπηρεσιών προς τα μέλη της πανεπιστημιακής κοινότητας, για τη διαχείριση και διάθεση του ψηφιακού υλικού που παράγεται εντός του Πανεπιστημίου. Η πρώτη έκδοση του λογισμικού ήταν διαθέσιμη το Νοέμβριο 2002, σύμφωνα με τους όρους που διέπει το BSD open source licence. Έχει υλοποιηθεί σε γλώσσα Java και κυκλοφορεί για πλατφόρμες Unix/Linux και Windows. Το DSpace βασίζεται στην αρχιτεκτονική των τριών διακριτών επιπέδων: εφαρμογής, λογικών διαδικασιών και αποθήκευσης. Κάθε ψηφιακό αντικείμενο (item στην ορολογία του Dspace) συνοδεύεται από ξεχωριστή Dublin Core εγγραφή και το περιεχόμενο του ορίζεται ως ένα σύνολο από bundles, που περιέχουν bitstreams. Τα bitstreams αντιστοιχούν σε αρχεία ενώ τα bundles χρησιμοποιούνται για την ομαδοποίηση τους. Κάθε bitstream αντιστοιχεί σε ένα bitstream format, το οποίο συνιστά έναν μοναδικό και συνεπή τρόπο για την αναπαράσταση των τύπων αρχείων (file formats).

Το DSpace μπορεί να χρησιμοποιηθεί για τη δημιουργία μιας ψηφιακής βιβλιοθήκης αποτελούμενης από ανεξάρτητα ψηφιακά έγγραφα. Παρέχει βασικές, χρηστικές και αποτελεσματικές λειτουργίες, όπως είναι η διαχείριση χρηστών, η διαχείριση απλών ροές εργασίας, η διατήρηση και ευρετηρίαση / αναζήτηση υλικού. Αυτές οι ευκολίες, σίγουρα επιταχύνουν και διευκολύνουν την ανάπτυξη μιας Ψηφιακής Βιβλιοθήκης, αλλά, το σύστημα δεν προβλέπει καθόλου ψηφιοποιημένο περιεχόμενο, διασυνδέσεις και συσχετίσεις αντικειμένων ούτε και πρόσθετα σχήματα μεταδεδομένων σε ικανοποιητικό επίπεδο σαφήνειας και λεπτομέρειας. Επιπρόσθετα, οι δυνατότητες παραμετροποίησης που προσφέρει περιορίζονται κυρίως σε α) αισθητικά χαρακτηριστικά διεπαφής χρήσης, β) συνολική (για όλες τις συλλογές) τροποποίηση του σχήματος μεταδεδομένων και γ) καθορισμό ροών εργασίας αποτελούμενες, το πολύ, από τρία βήματα.

Παρ' όλο που, θεωρητικά, οι απαιτούμενες προσθήκες και τροποποιήσεις είναι δυνατό να γίνουν, το βασικό πρόβλημα είναι ότι το DSpace βασίζεται σε μια σχεσιακή βάση δεδομένων για την αποθήκευση της πληροφορίας. Έτσι, παρ' όλο που είναι δυνατό να τροποποιηθεί το σχήμα της βάσης δεδομένων, με την προσθήκη των απαιτούμενων πεδίων ή / και πινάκων, παρουσιάζονται δύο σημαντικά ζητήματα, τα οποία, σε βάθος χρόνου, καθιστούν μία τέτοια ενέργεια αναποτελεσματική. Κατά πρώτο λόγο, η τροποποίηση του σχήματος της βάσης δεδομένων απαιτεί, εύλογα, τη συγγραφική τροποποίηση (σε επίπεδο πηγαίου κώδικα) των βασικών «συστατικών» ή «υποσυστημάτων» του DSpace. Για να πραγματοποιηθούν δραστικές αλλαγές στη λειτουργικότητα του συστήματος απαιτούνται τροποποιήσεις τόσο στο σχήμα της βάσης δεδομένων όσο και στον αντίστοιχο πηγαίο κώδικα. Κατά δεύτερο λόγο, οι αλλαγές αυτές θα καταργήσουν τη συμβατότητα του συστήματος με μελλοντικές εκδόσεις του DSpace αλλά και με τις υπόλοιπες Ψηφιακές Βιβλιοθήκες που βασίζονται σε αυτό, περιορίζοντας, ουσιαστικά, στο ελάχιστο τη δυνατότητα εκμετάλλευσης πιθανών μελλοντικών βελτιώσεων και επεκτάσεων.

Το βασικό, επομένως, σημείο, το οποίο καθιστά το DSpace ανεπαρκές για τις ανάγκες του ΕΚΠΑ είναι ότι η υπάρχουσα αρχιτεκτονική του θα εμποδίσει παρά θα ωφελήσει στην ανάπτυξη των χαρακτηριστικών που έχουν προδιαγραφεί, καθώς αυτά δεν συμπεριλαμβάνονταν στον αρχικό του σχεδιασμό.

Fedora

Το σύστημα Ψηφιακής Βιβλιοθήκης Fedora βασίζεται στην αρχιτεκτονική Flexible Extensible Digital Object and Repository Architecture (FEDORA) (Payette 1998). Η ανάπτυξη του ξεκίνησε το 2001 από τα πανεπιστήμια University of Virginia και Cornell University. Η υλοποίηση του λογισμικού έχει πραγματοποιηθεί σε γλώσσα Java και είναι διαθέσιμο σε πλατφόρμες Windows και Unix/Linux. Τα ψηφιακά αντικείμενα (digital objects), η βασική οντότητα του Fedora, είναι βασισμένα στη δομή που ορίζει το πρότυπο Metadata Encoding and Transmission Standard (METS) (2004). Το ψηφιακό αντικείμενο αποτελείται από έναν μοναδικό προσδιοριστή (PID), τα μεταδεδομένα συστήματος, ένα ή περισσότερα datastream (που αποτελούν τα μεταδεδομένα ή το ψηφιακό υλικό) και τα disseminator που συνδέουν τα datastream με κάποια συμπεριφορά. Υπάρχουν τριών ειδών ψηφιακά αντικείμενα: τα αντικείμενα δεδομένων (data objects), τα αντικείμενα καθορισμού συμπεριφοράς (behavior definition objects) και τα αντικείμενα μηχανισμού συμπεριφοράς (behavior mechanism objects). Τα αντικείμενα δεδομένων αποτελούν τις οντότητες που περιέχουν το ψηφιακό υλικό και τα μεταδεδομένα ενώ τα υπόλοιπα δυο καθορίζουν και υλοποιούν τις μεθόδους για την παρουσίαση ή τη μετατροπή των datastream ενός αντικειμένου δεδομένων.

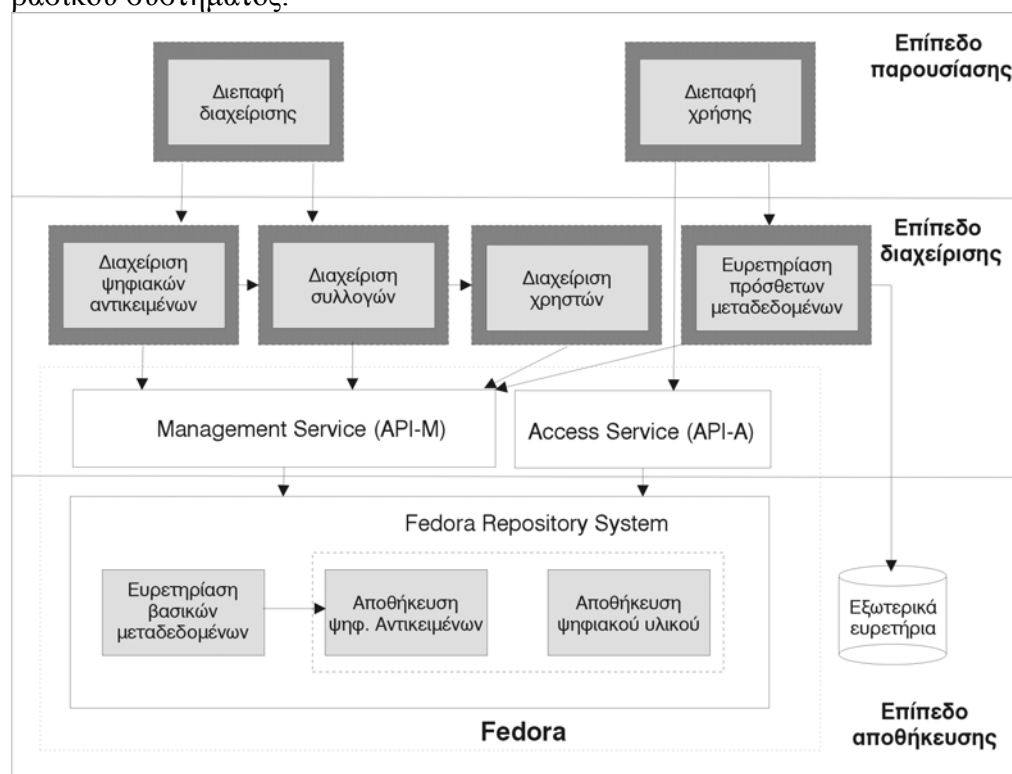
Το Fedora παρουσιάζει την πιο ευέλικτη αρχιτεκτονική από τα τρία συστήματα και το μοντέλο ψηφιακών αντικειμένων που υποστηρίζει προσφέρει τεράστιες δυνατότητες επέκτασης. Αυτός είναι και ο σημαντικότερος λόγος επιλογής του. Πιο συγκεκριμένα, οι περισσότερες επεκτάσεις βασίζονται στα datastreams από τα οποία απαρτίζεται κάθε ψηφιακό αντικείμενο του Fedora. Έτσι, παρέχεται η δυνατότητα υποστήριξης σύνθετων αντικειμένων καθώς και η ενσωμάτωση ετερογενών και ποικίλων μεταδεδομένων (περιγραφικά, τεχνικά, διαχειριστικά ή δομικά). Επιπλέον, με τη χρήση συμπεριφορών στα ψηφιακά αντικείμενα είναι δυνατό να υλοποιηθούν συσχετίσεις μεταξύ αντικειμένων με δυναμικό και ενιαίο τρόπο. Υποστηρίζεται, επίσης, η δυνατότητα χρήσης προτύπων (templates) για τη δημιουργία, επεξεργασία και παρουσίαση των ψηφιακών αντικειμένων. Τέλος, οι διαφορετικές συλλογές και υπο-συλλογές μπορούν να αναπαρασταθούν ως κανονικά ψηφιακά αντικείμενα, προσφέροντας ένα ενιαίο και κοινό σχήμα αναπαράστασης στο προγραμματιστικό επίπεδο.

Η ευέλικτη αρχιτεκτονική του Fedora παρέχει τη δυνατότητα ανάπτυξης πρόσθετης λειτουργικότητας, η οποία ούτε περιορίζει αλλά ούτε και περιορίζεται από την υπάρχουσα διάρθρωση του συστήματος, καθώς κάθε «υποσύστημα» μπορεί να λειτουργεί ξεχωριστά και απομονωμένα από τα υπόλοιπα, διατηρώντας όμως την ικανότητα επικοινωνίας μέσω του βασικού Fedora API. Έτσι, επιτυγχάνεται η ανάπτυξη ενός ΣΨΒ «πάνω από» το Fedora, το οποίο μπορεί να λειτουργεί ανεξάρτητα σε ότι αφορά στην επιχειρησιακή του λογική, τα μεταδεδομένα που χρησιμοποιεί και τη σημασία που αυτά έχουν, παραμένοντας συμβατό με τις νέες εκδόσεις του Fedora.

Αρχιτεκτονική συστήματος ΨΒ / Επεκτάσεις του Fedora

Η αρχιτεκτονική του ΣΨΒ του Πανεπιστημίου Αθηνών (Pygounakis 2004), διακρίνεται στο Σχήμα 1. Στο τελικό ΣΨΒ το Fedora παρέχει υπηρεσίες, που αντιστοιχούν στο επίπεδο αποθήκευσης και ένα ποσοστό του επιπέδου διαχείρισης. Υπάρχουν τέσσερις βασικές υπηρεσίες, που πρέπει να υλοποιηθούν σε επίπεδο

διαχείρισης προκειμένου να δίδεται η πλήρης απαιτούμενη λειτουργικότητα προς το επίπεδο παρουσίασης: η διαχείριση ψηφιακών αντικειμένων, η διαχείριση συλλογών, η διαχείριση χρηστών και η ευρετηρίαση/αναζήτηση των πρόσθετων μεταδεδομένων, που δεν καλύπτεται άμεσα από το Fedora. Οι τέσσερις αυτές υπηρεσίες επικοινωνούν άμεσα με το API διαχείρισης (API-M) του Fedora. Με αυτό τον τρόπο, τυχόν αλλαγές που προκύπτουν στο εσωτερικό σύστημα του Fedora σε νεότερες εκδόσεις, δεν επηρεάζουν τις υπηρεσίες του επιπέδου διαχείρισης, εφόσον δεν μεταβληθούν οι μέθοδοι του API. Αυτή είναι μια βασική αιτία για την επιλογή χρήσης του Fedora, ως βασικού συστήματος.



Σχήμα 1: Η προτεινόμενη αρχιτεκτονική του ΣΨΒ του ΕΚΠΑ

Στη συνέχεια αναφέρεται η λειτουργικότητα που παρέχεται από κάθε υπηρεσία του επιπέδου διαχείρισης:

Διαχείριση συλλογών

Δεδομένου ότι είναι αδύνατον να αναπτύσσονται νέες εφαρμογές όποτε πρόκειται να προστεθεί μια συλλογή στο ΣΨΒ, αναζητείται μια λύση που να αντιμετωπίζει το πρόβλημα με ευελιξία και ελαχιστοποίηση κόστους. Η λύση που υιοθετήθηκε για τη συγκεκριμένη περίπτωση είναι ο παραμετρικός ορισμός των συλλογών, με σκοπό την ανάπτυξη εφαρμογών, που εκτελούνται δυναμικά, με βάση τα χαρακτηριστικά που έχουν οριστεί σε κάθε συλλογή. Για την πλήρη περιγραφή μιας συλλογής, μπορεί να οριστεί ένα XML σχήμα, όπου να καθορίζονται οι ιδιαιτερότητες και τα χαρακτηριστικά της.

Το κυριότερο χαρακτηριστικό της συλλογής, που αφορά την παραμετρική υλοποίηση εφαρμογών, είναι τα πρότυπα (templates) που απαιτούνται για κάθε τύπο ψηφιακών αντικειμένων. Σε μια συλλογή μπορούν να οριστούν διαφορετικοί τύποι ψηφιακών αντικειμένων, ειδικότερα όταν η συλλογή αποτελείται από σύνθετα αντικείμενα. Για

παράδειγμα, στην περίπτωση των λαογραφικών εργασιών, έχουμε τρεις τύπους ψηφιακών αντικειμένων: εργασία, κεφάλαιο εργασίας και σελίδα. Κάθε τύπος ψηφιακού αντικειμένου διαθέτει ένα πρότυπο, σύμφωνα με το οποίο δημιουργείται κάθε νέο ψηφιακό αντικείμενο της συλλογής. Στο πρότυπο ορίζονται τα πεδία μεταδεδομένων, οι τύποι του ψηφιακού υλικού που το αποτελούν, οι επιτρεπτές συσχετίσεις με άλλα ψηφιακά αντικείμενα και οι μέθοδοι που καθορίζουν τη συμπεριφορά του ψηφιακού αντικειμένου.

Κάθε συλλογή αναπαρίσταται ως ένα ψηφιακό αντικείμενο του Fedora, που περιέχει ένα `datastream` για τα περιγραφικά και διαχειριστικά μεταδεδομένα της συλλογής και ένα `datastream` για τα πρότυπα των ψηφιακών αντικειμένων που δημιουργούνται στο πλαίσιο της συλλογής. Τα `datastream` αποθηκεύονται ως XML ενσωματωμένη στο ψηφιακό αντικείμενο.

Η υπηρεσία διαχείρισης συλλογών παρέχει μεθόδους για τη δημιουργία συλλογών και υποσυλλογών, όπως και την προσθήκη και ανάκτηση των μεταδεδομένων της συλλογής. Επίσης, παρέχει μεθόδους για τη δημιουργία και επεξεργασία του προτύπου ενός συγκεκριμένου τύπου ψηφιακού αντικειμένου, που διαθέτει η συλλογή. Για την ανάκτηση των επιμέρους χαρακτηριστικών, που περιέχονται στο πρότυπο (πεδία μεταδεδομένων, αρχεία, συσχετίσεις, συμπεριφορές), διατίθενται επιπλέον μέθοδοι, οι οποίες είναι απαραίτητες κυρίως για τη δυναμική λειτουργία των διεπαφών διαχείρισης και χρήσης. Τέλος, είναι απαραίτητες μέθοδοι για τον καθορισμό και την ανάκτηση των χαρακτηριστικών παρουσίασης της συλλογής.

Διαχείριση ψηφιακών αντικειμένων

Η συγκεκριμένη υπηρεσία επεκτείνει τις μεθόδους διαχείρισης ψηφιακών αντικειμένων, που παρέχει το Fedora μέσω του API-M. Προσφέρει μεθόδους για τη δημιουργία και διαγραφή ψηφιακών αντικειμένων, προσθήκη, διαγραφή και επεξεργασία των `datastream` και `disseminator`. Η κυριότερη διαφορά από το API-M είναι ότι οι μέθοδοι αυτές λαμβάνουν υπόψη τα πρότυπα ψηφιακών αντικειμένων προκειμένου να δημιουργήσουν νέα `datastream` ή `disseminator`. Η διαχείριση ψηφιακών αντικειμένων επικοινωνεί με τη διαχείριση συλλογών μέσω των ανάλογων μεθόδων.

Επιπλέον, για την αναπαράσταση της λογικής δομής ενός σύνθετου αντικειμένου, αλλά και τη συσχέτιση ανεξάρτητων ψηφιακών αντικειμένων που ανήκουν στην ίδια ή σε διαφορετικές συλλογές, είναι απαραίτητη η ύπαρξη συσχετίσεων. Για την υλοποίηση των συσχετίσεων χρησιμοποιείται σε κάθε ψηφιακό αντικείμενο ένα επιπλέον `datastream` που περιέχει δομικά μεταδεδομένα της μορφής:

```
<relation type='relation_type'>destinationPID</relation>
```

Η τιμή `'relation_type'` προσδιορίζει τον τύπο συσχέτισης μεταξύ του συγκεκριμένου ψηφιακού αντικειμένου και εκείνου με μοναδικό προσδιοριστή `destinationPID`. Οι επιτρεπτές συσχετίσεις για ένα ψηφιακό αντικείμενο έχουν καθοριστεί στο ανάλογο πρότυπο, σε επίπεδο συλλογής. Για παράδειγμα οι τύποι συσχετίσεων μιας λαογραφικής εργασίας είναι: `'previous'`, `'next'` για την πλοήγηση μεταξύ των σελίδων, `'chapter'` και `'main'` για τον καθορισμό του κεφαλαίου και της κύριας εργασίας αντίστοιχα.

Διαχείριση χρηστών

Η διαχείριση χρηστών παρέχει μεθόδους για τη δημιουργία και επεξεργασία χρηστών και ομάδων χρηστών, εφόσον μέχρι τη συγκεκριμένη έκδοση το Fedora δεν προσφέρει ανάλογες δυνατότητες. Παράλληλα, προσφέρει μεθόδους για τον καθορισμό δικαιωμάτων των χρηστών για κάθε συλλογή, όπως και την ανάκτηση πληροφοριών για τα δικαιώματα κάθε χρήστη.

Ευρετηρίαση και αναζήτηση πρόσθετων μεταδεδομένων

Εφόσον το Fedora υποστηρίζει την ευρετηρίαση και αναζήτηση μόνο των μεταδεδομένων που ανήκουν στο πρότυπο Dublin Core, πρέπει να παρέχεται μια εξωτερική υπηρεσία για την ευρετηρίαση και αναζήτηση των πρόσθετων μεταδεδομένων κάθε συλλογής. Οι μέθοδοι, που προσφέρει η συγκεκριμένη υπηρεσία, αφορούν τη δημιουργία ευρετηρίου ανά συλλογή, όπως και την ενημέρωση του με τα μεταδεδομένα συγκεκριμένου ψηφιακού αντικείμενου. Επίσης, διαθέτει μεθόδους για την απλή ή σύνθετη αναζήτηση λέξεων στα επιλεγμένα ευρετήρια. Δύο εφαρμογές ανοικτού κώδικα που μπορούν να χρησιμοποιηθούν για το συγκεκριμένο σκοπό είναι το Jakarta Lucene (2004) και το Apache Xindice (2004).

Στο επίπεδο παρουσίασης, πρόκειται να αναπτυχθούν οι διεπαφές διαχείρισης και χρήσης. Η διεπαφή διαχείρισης αποτελείται από την εφαρμογή ψηφιοποίησης και καταλογογράφησης, καθώς και την εφαρμογή διαχείρισης συλλογών. Η διεπαφή διαχείρισης επικοινωνεί με τις υπηρεσίες διαχείρισης ψηφιακών αντικείμενων και συλλογών στο επίπεδο διαχείρισης. Η διεπαφή χρήσης παρέχει πρόσβαση στον τελικό χρήστη μέσω του παγκοσμίου ιστού, όπου του δίνεται η δυνατότητα να πραγματοποιήσει αναζητήσεις, να πλοηγηθεί στα αντικείμενα των συλλογών και να επεξεργαστεί το ψηφιακό υλικό. Η διεπαφή χρήσης επικοινωνεί με την υπηρεσία ευρετηρίασης/αναζήτησης των πρόσθετων μεταδεδομένων και με το API πρόσβασης (API-A) του Fedora. Το API-A προσφέρει μεθόδους για την αναζήτηση στα βασικά μεταδεδομένα (Dublin Core) και την ανάκτηση των συμπεριφορών, που διαθέτει κάθε ψηφιακό αντικείμενο, προκειμένου να παρουσιαστεί το ψηφιακό υλικό και να χρησιμοποιηθούν τα μεταδεδομένα.

Επίλογος / Μελλοντικά Σχέδια

Στην παρούσα εισήγηση έγινε μια περιληπτική παρουσίαση του σχεδιασμού του ΣΨΒ, το οποίο αναπτύσσεται για να φιλοξενήσει τις ψηφιακές συλλογές του ΕΚΠΑ. Το σύστημα αυτό σχεδιάστηκε με στόχο να εξυπηρετήσει ετερογενείς ψηφιακές συλλογές που προέρχονται από διάφορες θεματικές περιοχές και έχουν επιστημονικό ή πολιτιστικό ενδιαφέρον για την ακαδημαϊκή κοινότητα. Η ανάπτυξη του έχει βασιστεί στο σύστημα ανοικτού κώδικα Fedora, το οποίο παρέχει τη βασική λειτουργικότητα μιας Ψηφιακής Βιβλιοθήκης, ενώ ταυτόχρονα επιτρέπει την εύκολη παραμετροποίηση και επέκταση του. Η αρχιτεκτονική που προτάθηκε στην εισήγηση, επεκτείνει το Fedora, με την υλοποίηση επιπλέον υπηρεσιών που καλύπτουν απόλυτα τις ιδιαιτερότητες που παρουσιάζουν οι ψηφιακές συλλογές του ΕΚΠΑ.

Αυτό το διάστημα πραγματοποιείται η ανάπτυξη των βασικών τμημάτων του ΣΨΒ του Πανεπιστημίου Αθηνών. Παράλληλα καθορίζονται οι μέθοδοι ψηφιοποίησης που

πρόκειται να χρησιμοποιηθούν για κάθε συλλογή, έτσι ώστε να ανταποκρίνονται στις ιδιαιτερότητες του κάθε υλικού. Στη συνέχεια θα υλοποιηθούν τα τμήματα που προσδίδουν επιπλέον λειτουργικότητα στο σύστημα, όπως είναι η ευρετηρίαση των πρόσθετων μεταδεδομένων, η διεπαφή χρήσης για τον τελικό χρήστη καθώς και η ανάπτυξη ροών εργασίας. Στόχος μας είναι να προστίθενται συνεχώς νέες λειτουργίες και ευκολίες στο σύστημα, ώστε να αποτελέσει μια σύγχρονη Ψηφιακή Βιβλιοθήκη, τόσο για τους χρήστες του ΕΚΠΑ όσο και για την ευρύτερη ακαδημαϊκή κοινότητα.

Αναφορές

- Apache Xindice (2004) [<http://xml.apache.org/xindice/>]
- Dempsey, L. (1999) “Scientific, Industrial, and Cultural Heritage: a shared approach.” Ariadne Issue 22 [<http://www.ariadne.ac.uk/issue22/>]
- DSpace Federation (2004) [<http://www.dspace.org/>]
- Dublin Core Metadata Initiative: Dublin core Metadata Element Set, Version 1.1: Reference Description (2003) [<http://www.dublincore.org/documents/dces/>]
- Dublin Core Collection Description Application Profile (2003) [<http://www.ukoln.ac.uk/metadata/dcmi/collection-application-profile/2003-08-25/>]
- Fedora Project (2004) [<http://www.fedora.info/>]
- Greenstone Digital Library Software (2004) [<http://www.greenstone.org/>]
- IEEE (2002) IEEE P1484.12.1/D6.4 Draft Standard for Learning Object Metadata, [http://ltsc.ieee.org/doc/wg12/LOM_WD6_4.pdf]
- Jakarta Lucene (2004) [<http://jakarta.apache.org/lucene/>]
- Lagoze, C. and Van de Sompel, H. (2001) “The Open Archives Initiative: Building a low-barrier interoperability framework.” JCDL '01
- McCray, A. T. and Gallagher, E. (2001) “Principles for Digital Library development.” Communications of the ACM, 44(5), pp. 49-54
- Metadata Encoding and Transmission Standard (2004) [<http://www.loc.gov/standards/mets/>]
- Open Society Institute (2004) A Guide to Institutional Repository Software. 2nd Edition
- Open Source Software (2004) [<http://www.opensource.org/>]
- Payette S. and Lagoze C. (1998) “Flexible and Extensible Digital Object and Repository Architecture (FEDORA).” ECDL '98, LNCS 1513, pp. 41-59
- Pyrounakis, G., Saidis, K., Nikolaidou, M. and Lourdi, I. (2004) “Designing an Integrated Digital Library Framework to support Multiple Heterogeneous Collections.” ECDL 2004
- Smith, M. et al. (2003) “DSpace: An Open Source Dynamic Digital Repository.” D-Lib Magazine, 9(1)
- Staples, T., Wayland, R. and Payette S. (2003) “The Fedora Project: An Open-source Digital Object Repository Management System.” D-Lib Magazine, 9(4)
- Witten, I. H., McNab, R. J., Boddie, S. J. and Bainbridge, D. (2000) “Greenstone: A Comprehensive Open-Source Digital Library Software System.” Proceedings of the fifth ACM conference on Digital libraries