

Λογισμικό Ψηφιακών Βιβλιοθηκών Greenstone



Μάθημα: Ψηφιακές Βιβλιοθήκες

Αλέξανδρος Βαρβέρης

E-mail: avarver@law.uoa.gr

Αθήνα, Ιανουάριος 2004

ΠΕΡΙΕΧΟΜΕΝΑ

Περιεχόμενα	Σελ.
Γενικά.....	3
Λειτουργία του Συστήματος.....	5
Αρχεία Συλλογών.....	8
Αρχεία και Κατάλογοι	9
Εισαγόμενα Έγγραφα	9
Πως δουλεύει	9
Είσοδος στο σύστημα	10
Ο Συλλέκτης	10
Βήματα για το χτίσιμο της συλλογής.....	11
Πληροφορίες συλλογής	11
Κείμενα Πηγής	12
Διαμόρφωσης της συλλογής	13
Χτίσιμο της συλλογής	14
Εξέταση της συλλογής	15
Μεταδεδομένα Εγγράφων	15
Μεταδεδομένα Αρχείων	16
Plugins	17
Ταξινομητές	20
Διαμόρφωση της διεπαφής	22
Συμπεράσματα	23
Αναφορές	25

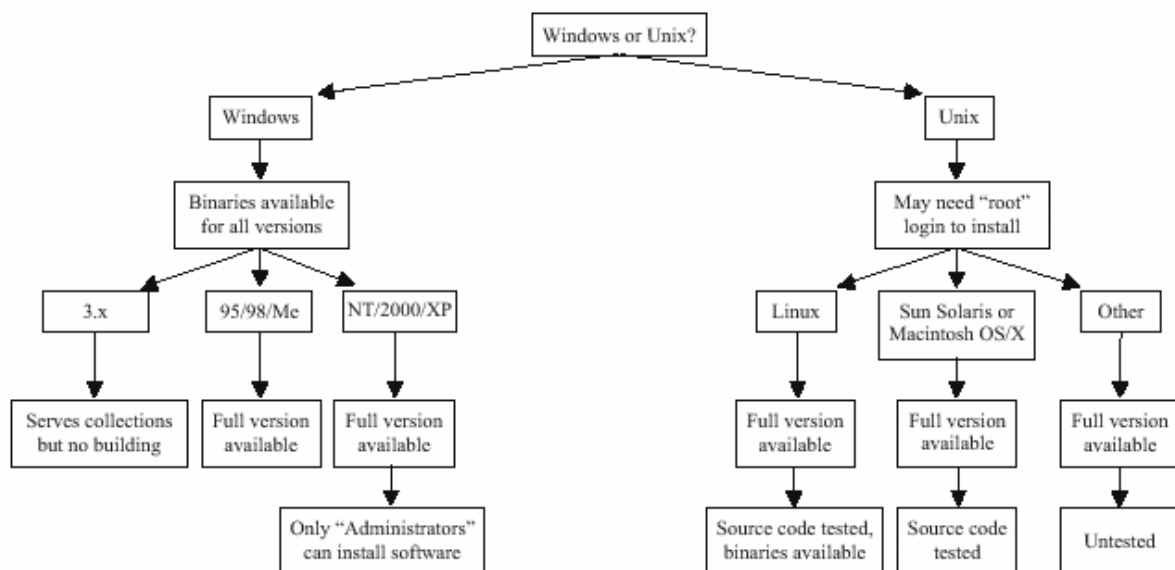
ΓΕΝΙΚΑ

Οι ψηφιακές βιβλιοθήκες αποθηκεύουν και χειρίζονται αποτελεσματικά μεγάλες συλλογές ηλεκτρονικού υλικού κάθε μορφής. Έρευνα στις ψηφιακές βιβλιοθήκες σημαίνει την έρευνα στα δικτυακά συστήματα πληροφοριών που επικεντρώνεται στο πώς να αναπτυχθεί η απαραίτητη υποδομή για τον μαζικό και αποτελεσματικό χειρισμό της πληροφορίας μέσω του δικτύου. Τα βασικά τεχνολογικά προβλήματα που προκύπτουν και οφείλουν να λυθούν αποτελεσματικά είναι η αναζήτηση και η παρουσίαση των επιθυμητών επιλογών από μεγάλες συλλογές.

Τα portals και τα intranets, είναι οι πρόδρομοι των σημερινών βιβλιοθηκών αν και υπάρχει ακόμα ανάγκη για επιπλέον έρευνα στον τομέα. Από την άλλη πλευρά, το σύστημα Greenstone είναι μια πλατφόρμα λογισμικού για την οικοδόμηση και τη διανομή συλλογών ψηφιακών βιβλιοθηκών που παρέχει έναν νέο τρόπο για την οργάνωση της πληροφορίας και της έκδοσής της στο διαδίκτυο ή σε CD-ROM.

Το Greenstone παράγεται από το Humanity Libraries Project της Νέας Ζηλανδίας του Πανεπιστημίου του Waikato. Είναι λογισμικό ανοικτού πηγαίου κώδικα(open source), που παρέχεται στο διαδίκτυο στη διεύθυνση <http://www.greenstone.org/>. Σημειώνεται ότι, σε συνεργασία με την UNESCO και το Humanity Libraries Project, το Greenstone εκδίδεται και σε CD-ROM που περιλαμβάνει την τεκμηρίωσή του στα Αγγλικά, Γαλλικά, και Ισπανικά.

Το Greenstone δουλεύει σε διαφορετικές πλατφόρμες όπως φαίνεται στο Σχήμα 1. Διάφορες εκδόσεις του λογισμικού αυτού είναι διαθέσιμες τόσο για πλατφόρμες σε Windows, όσο και σε Unix και το περιβάλλον του χρήστη μπορεί να είναι ένας web browser (Mozilla, Netscape Navigator, Internet Explorer v4 και πάνω).



Σχήμα 1

Για να ικανοποιήσει τα αυξημένα αιτήματα των ψηφιακών βιβλιοθηκών, το Greenstone είναι δημόσιο και επεκτάσιμο (extended). Βασισμένο στην φιλοσοφία του ανοικτού κώδικα λογισμικού <http://www.gnu.org/copyleft/gpl.html> (open source), οι χρήστες έχουν τη δυνατότητα να το τροποποιήσουν κατά βούληση σύμφωνα με τις ανάγκες τους και τους στόχους τους.

Το λογισμικό Greenstone χρησιμοποιείται από χώρες όπως ο Καναδάς, η Γερμανία, η Νέα Ζηλανδία, η Ρουμανία, η Αγγλία, και οι ΗΠΑ, καθώς και από ανθρωπιστικές αντιπροσωπείες με κύριο στόχο να γίνει εκμετάλλευση και προώθηση αυτού του λογισμικού σε υπανάπτυκτες χώρες. Σημειώνεται ότι, οι συλλογές του παραπάνω λογισμικού περιλαμβάνουν πλήθος από άρθρα εφημερίδων ως και τεχνικά έγγραφα και από εκπαιδευτικά περιοδικά έως παραδοσιακά τραγούδια. Με τη χρήση του συγκεκριμένου λογισμικού μια ψηφιακή βιβλιοθήκη μπορεί να περιλάβει πολλές διαφορετικές συλλογές που η κάθε μια από αυτές να οργανώνεται διαφορετικά, αν και υπάρχει μια ισχυρή οικογενειακή ομοιότητα στο πώς παρουσιάζονται οι συλλογές.

Οι ψηφιακές βιβλιοθήκες ξεχωρίζουν από το World Wide Web και η ουσιαστική διαφορά τους είναι στην επιλογή και την οργάνωση. Όμως, δεν είναι γενικά διακριτές από έναν ιστοχώρο, διότι όλες οι υπάρχουσες ψηφιακές βιβλιοθήκες φανερώνονται ως ιστοσελίδες. Το να καταστήσει κανείς τις πληροφορίες διαθέσιμες χρησιμοποιώντας το Greenstone, απέχει πολύ από μια απλή τοποθέτηση αυτών στο web. Το πλεονέκτημα

είναι ότι η συλλογή γίνεται συντηρήσιμη και εξερευνήσιμη. Κάθε συλλογή, πριν από την παρουσίαση της, υποβάλλεται σε μια διαδικασία 'χτισίματος' (building) που, μόλις ολοκληρωθεί η φάση παραμετροποίησης, γίνεται απολύτως αυτόματη. Αυτή η διαδικασία δημιουργεί όλες τις δομές για την πρόσβαση στην συλλογή. Η έρευνα (search) είναι βασισμένη στα διάφορα ευρετήρια (indexes), ενώ το ξεφύλλισμα (browsing) βασίζεται σε μεταδεδομένα (metadata). Οι δομές υποστήριξης και για τις δύο διαδικασίες (search & browsing) δημιουργούνται κατά τη διάρκεια της λειτουργίας οικοδόμησης. Όταν εμφανίζεται νέο υλικό για τη συλλογή, μπορεί αυτό να ενσωματωθεί πλήρως με τη διαδικασία επανοικοδόμησης (rebuilding).

Το διεθνές σύνολο χαρακτήρων unicode χρησιμοποιείται στο Greenstone, έτσι ώστε τα έγγραφα και οι διεπαφές (interfaces) να μπορούν να γραφτούν σε οποιαδήποτε γλώσσα. Οι συλλογές που έχουν παραχθεί μέχρι τώρα είναι στα Αγγλικά, Γαλλικά, Ισπανικά, Γερμανικά, Κινέζικα, Αραβικά και Ελληνικά.

Ο ιστοχώρος NZDL (<http://www.nzdl.org>) παρέχει πολυάριθμα παραδείγματα συλλογών. Οι συλλογές μπορούν να περιέχουν κείμενο, εικόνες, ακόμη και video clips. Παρέχεται επίσης έκδοση της διεπαφής (interface) με μόνο κείμενο για να εξυπηρετήσει χρήστες με οπτική αναπηρία. Χρησιμοποιείται ακόμη και τεχνολογία συμπίεσης για να εξασφαλιστεί η καλύτερη χρήση των αποθηκευτικών μέσων.

ΛΕΙΤΟΥΡΓΙΑ ΤΟΥ ΣΥΣΤΗΜΑΤΟΣ

Το σύστημα περιλαμβάνει μια 'διοικητική' (administrative) λειτουργία με την οποία συγκεκριμένοι χρήστες μπορούν να εξετάσουν τη σύνθεση όλων των συλλογών, να προστατεύσουν έγγραφα έτσι ώστε να μπορούν να προσπελαστούν μόνο από εγγεγραμμένους χρήστες με τη χρήση ενός κωδικού πρόσβασης.

Αν και οι συλλογές σχεδιάζονται πρωτίστως για την πρόσβαση μέσω του διαδικτύου, μπορεί αυτές να είναι διαθέσιμες ακριβώς στην ίδια μορφή και σε CD-ROM και κάθε μία από τις δυο περιπτώσεις να προσεγγίζεται μέσω ενός οποιουδήποτε web browser. Τα CD-ROM που είναι φτιαγμένα με το λογισμικό ψηφιακών βιβλιοθηκών μπορούν να λειτουργήσουν σε ένα αυτόνομο PC κάτω από Windows 3.x, 95, 98 και NT και η αλληλεπίδραση με το χρήστη να είναι ίδια με την πρόσβαση στη συλλογή μέσω web και επιπρόσθετα τα αποτελέσματα λαμβάνονται γρηγορότερα.

Οι συλλογές πληροφοριών που χτίζονται από το Greenstone συνδυάζουν τις εκτενείς εγκαταστάσεις αναζήτησης ολοκληρωμένου κειμένου με το ξεφύλλισμα των ευρετηρίων

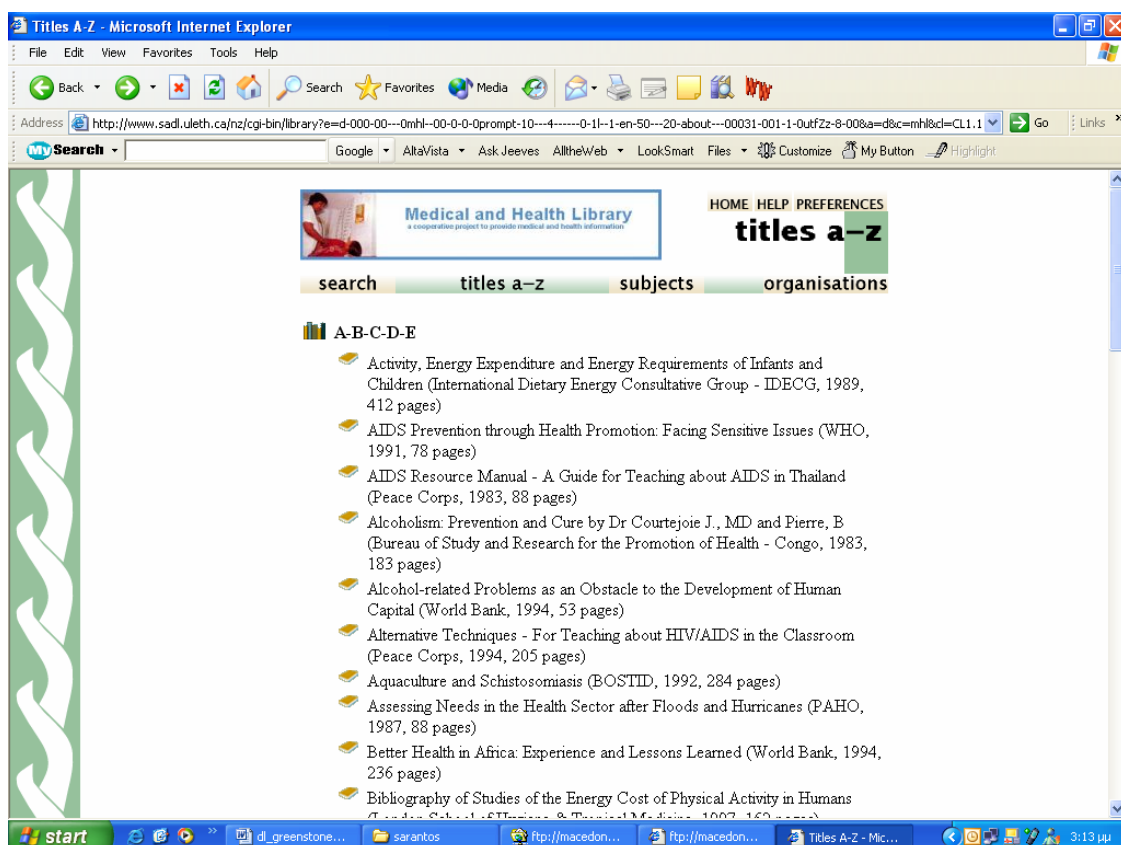
να είναι βασισμένα σε διαφορετικούς τύπους μεταδεδομένων. Υπάρχουν διάφοροι τρόποι για τους χρήστες ώστε να βρουν τις πληροφορίες, αν και διαφέρουν μεταξύ των συλλογών ανάλογα με τα διαθέσιμα μεταδεδομένα και το σχεδιασμό της συλλογής. Χαρακτηριστικά, μπορούμε να ψάξουμε για ιδιαίτερες λέξεις που εμφανίζονται σε ένα κείμενο, ή μέσα σε ένα τμήμα ενός εγγράφου, ή μέσα σε έναν τίτλο. Μπορούμε να βρούμε βιαστικά τα έγγραφα από τον *τίτλο* και χτυπώντας στην επιδειχθείσα εικόνα βιβλίων να τα διαβάσουμε. Τέλος, μπορούμε να βρούμε γρήγορα έγγραφα με βάση το *θέμα*. Τα θέματα αντιπροσωπεύονται από τα 'ράφια' και 'χτυπώντας' σε ένα 'ράφι' εμφανίζονται βιβλία σχετικά με το θέμα. Όπου απαιτείται, τα έγγραφα παρουσιάζονται πλήρη με έναν πίνακα περιεχομένων (που κατασκευάζεται αυτόματα) και κάνοντας 'κλικ' σε ένα κεφάλαιο ή μια υποενότητα μπορούμε να τα 'ανοίξουμε', να επεκτείνουμε τον πλήρη πίνακα περιεχομένων, ή να επεκτείνουμε το πλήρες έγγραφο.

Οι συλλογές μπορεί να περιλαμβάνουν πολλά έγγραφα, χιλιάδες ή ακόμα και εκατομμύρια. Κάθε έγγραφο μπορεί να οργανωθεί ιεραρχικά σε *τμήματα*, *υποενότητες*, κ.λ.π., ανάλογα με τις απαιτήσεις του χρήστη.

Με τη χρήση του λογισμικού ψηφιακών βιβλιοθηκών μπορούμε επίσης να προσθέσουμε νέο υλικό χωρίς να χρειάζεται να το ενσωματώσουμε με το χέρι ή να μορφοποιήσουμε το περιεχόμενό του. Μόλις προστεθεί το νέο υλικό πρέπει αμέσως να γίνει χρήσιμο και λειτουργικό για τη βιβλιοθήκη. Αυτό που μας επιτρέπει να ενσωματώνουμε υλικό στην υπάρχουσα δομή, χωρίς καμία χειρωνακτική παρέμβαση, είναι τα *μεταδεδομένα*.

Τα μεταδεδομένα όπως π.χ. ο συντάκτης, τίτλος, ημερομηνία, λέξεις κλειδιά, και τα λοιπά, μπορούν να συνδεθούν με τα έγγραφα, ή με μεμονωμένα τμήματα των εγγράφων. Αυτό αποτελεί την πρώτη ύλη για τους δείκτες.

Το λογισμικό ψηφιακών βιβλιοθηκών περιλαμβάνει διάφορες συλλογές. Μια αρχική σελίδα μας επιτρέπει να επιλέξουμε μια συλλογή, και κάθε συλλογή έχει δικιά της σελίδα που μας δίνει πληροφορίες για το πώς η συλλογή οργανώνεται.



Σχήμα 2: Οργάνωση μιας συλλογής

Σε όλες τις εικόνες του σχήματος 2 μπορεί κανείς να κάνει ένα 'κλικ'. Εκείνες οι εικόνες που είναι στην κορυφή της σελίδας μας επιστρέφουν στην αρχική σελίδα, παρέχουν κείμενο βοήθειας, και μας επιτρέπουν να θέσουμε το περιβάλλον που επιθυμούμε για την έρευνα της προτίμησης μας

Η συλλογή (Σχήμα 2) έχει τέσσερα ευρετήρια μεταδεδομένων. Μπορούμε να ψάξουμε για ιδιαίτερες λέξεις που εμφανίζονται σε κείμενο από τη σελίδα 'αναζήτησης'. Η διεπαφή αναζήτησης είναι απλή και κατάλληλη για μη πεπειραμένους χρήστες. Μπορούμε να έχουμε πρόσβαση στις δημοσιεύσεις από το *θέμα* με 'κλικ' στο κουμπί *θέματα*, που παρουσιάζει έναν κατάλογο θεμάτων. Μπορούμε να έχουμε πρόσβαση στις δημοσιεύσεις από τον τίτλο, με 'κλικ' στο κουμπί των *τίτλων AZ*, που παρουσιάζει έναν κατάλογο βιβλίων κατά αλφαβητική σειρά. Μπορούμε επίσης να έχουμε πρόσβαση στις δημοσιεύσεις από την *οργάνωση* που παρουσιάζει έναν κατάλογο οργανώσεων.

search	Search for particular words
subjects	Access publications by subject
titles a–z	Access publications by title
organization	Access publications by organization
how to	<u>Access publications by “how to” listing</u>

Σχήμα 3: Τι σημαίνουν οι εικόνες στην κορυφή κάθε σελίδας

	Click on a book icon to read the corresponding book
	Click on a bookshelf icon to look at books on that subject
	View this document
	Open this folder and view contents
	Click on this icon to close the book
	Click on this icon to close the folder
	Click on the arrow to go on to the next section ...
	... or back to the previous section
DETACH	Open this page in a new window
EXPAND CONTENTS	Expand table of contents
EXPAND TEXT	Display all text
HIGH-LIGHTING	Highlight search terms

Σχήμα 4: Τι σημαίνουν οι εικόνες στην μπάρα αναζήτησης/ξεφυλλίσματος

ΑΡΧΕΙΑ ΣΥΛΛΟΓΩΝ

Όταν μια νέα συλλογή δημιουργείται ή καινούργιο υλικό προστίθεται σε μια υπάρχουσα συλλογή, το υλικό εισάγεται στο σύστημα μέσω μιας διαδικασίας γνωστής ως ‘εισαγωγή’ (import). Αυτό περιλαμβάνει τη μετατροπή των εγγράφων σε μια απλής μορφής γλώσσα τύπου HTML γνωστή ως GML (Greenstone Mark-up Language), η οποία περιλαμβάνει τα οποιαδήποτε μεταδεδομένα που συνδέονται με το έγγραφο.

ΑΡΧΕΙΑ ΚΑΙ ΚΑΤΑΛΟΓΟΙ

Υπάρχει ένας χωριστός κατάλογος (directory) για κάθε συλλογή στο δίσκο του υπολογιστή που είναι εγκατεστημένο το λογισμικό ψηφιακών βιβλιοθηκών, η οποία περιέχει πέντε υποκαταλόγους (subdirectories). Οι υποκατάλογοι αυτοί περιλαμβάνουν την αρχική πρώτη ύλη (import), τα αρχεία GML που δημιουργούνται από αυτήν (archives), την τελική συλλογή όπως την βλέπουν οι χρήστες (index), ένας κατάλογος που χρησιμοποιείται κατά τη διάρκεια της διαδικασίας οικοδόμησης (building), και ένας για οποιαδήποτε ενισχυτικά αρχεία (etc) συμπεριλαμβανομένου του αρχείου διαμόρφωσης (configuration file) που ελέγχει τη διαδικασία δημιουργίας της συλλογής.

ΕΙΣΑΓΟΜΕΝΑ ΈΓΓΡΑΦΑ

Προκειμένου να προσδιοριστούν τα έγγραφα εσωτερικά, ένα μοναδικό προσδιοριστικό αντικείμενο (Unique Object Identifier, OID) ορίζεται σε κάθε αρχικό έγγραφο (source) όταν εισάγεται και αποθηκεύεται ως μεταδεδομένα μέσα σε εκείνο το έγγραφο. Είναι σημαντικό ότι τα OIDs εμμένουν σε όλη τη διαδικασία του χτισίματος έτσι ώστε το ιστορικό αναζήτησης ενός χρήστη να είναι ανεπηρέαστο από την κάθε επανοικοδόμηση της συλλογής.

ΠΩΣ ΔΟΥΛΕΥΕΙ

Το αρχικό υλικό που βρίσκεται στον κατάλογο *εισαγωγών* (*import directory*) μπορεί να είναι σε οποιαδήποτε μορφή και χρειάζονται ειδικά plugins για να επεξεργαστούν τον κάθε τύπο διαμόρφωσης. Τα plugins που χρησιμοποιεί κάθε συλλογή πρέπει να διευκρινίζονται στο αρχείο διαμόρφωσης (configuration file) της κάθε συλλογής. Το πρόγραμμα *εισαγωγών* διαβάζει τον κατάλογο με τα plugins και περνά κάθε αρχείο από plugin σε plugin με σκοπό να βρει αυτό που μπορεί να το επεξεργαστεί. Κατά την ενημέρωση μιας υπάρχουσας συλλογής, όλα τα plugins που είναι απαραίτητα να επεξεργαστούν το νέο υλικό πρέπει να έχουν ήδη διευκρινιστεί στο αρχείο διαμόρφωσης.

Στη φάση οικοδόμησης δημιουργούνται οι δείκτες και για την έρευνα και για το ξεφύλλισμα.

ΕΙΣΟΔΟΣ ΣΤΟ ΣΥΣΤΗΜΑ

Για να δει κανείς μια συλλογή που έχει δημιουργηθεί με το λογισμικό ψηφιακών βιβλιοθηκών πρέπει να συνδεθεί στον εξυπηρετητή που φιλοξενεί τη συλλογή και με κάποιο Internet browser να δει τα περιεχόμενά της. Αντίστοιχα υπάρχει η δυνατότητα να συνδεθεί κάποιος εγγεγραμμένος χρήστης στο σύστημα από μια απομακρυσμένη τοποθεσία και να χτίσει ή να μετατρέψει και γιατί όχι να διαγράψει μια συλλογή.

Ο ΣΥΛΛΕΚΤΗΣ

Ο συλλέκτης (collector) είναι μια λειτουργία του συστήματος που μας βοηθά να δημιουργήσουμε νέες συλλογές, να τροποποιήσουμε ή να προσθέσουμε υλικό σε υπάρχουσες, ή να διαγράψουμε συλλογές. Για να γίνει αυτό καθοδηγούμαστε μέσω μιας ακολουθίας ιστοσελίδων που ζητούν τις πληροφορίες που απαιτούνται.

Η δομή κάθε συλλογής καθορίζεται από τον συλλέκτη (collector). Αυτός αναλαμβάνει τη διευκρίνιση του σχήματος (ή των σχημάτων) των εγγράφων πηγής, την απόφαση του πώς να επιδείξει τα έγγραφα, τον καθορισμό των μεταδεδομένων, ποιες δυνατότητες ξεφυλλίσματος της συλλογής θα υπάρχουν, πώς θα φαίνονται τα αποτελέσματα αναζήτησης και πως αυτά θα επιδεικνύονται.

Ο 'συλλέκτης' είναι ένα διαλογικό υποσύστημα για τη διαχείριση και την πρόσβαση των συλλογών και μπορεί να χρησιμοποιηθεί για,

- να δημιουργήσει μια νέα συλλογή με ίδια δομή με μια ήδη υπάρχουσα,
- να δημιουργήσει μια νέα συλλογή με μια διαφορετική δομή,
- να προσθέσει το νέο υλικό σε μια υπάρχουσα συλλογή,
- να τροποποιήσει τη δομή μιας υπάρχουσας συλλογής,
- να διαγράψει μια συλλογή,
- να γράψει μια υπάρχουσα συλλογή σε CD-ROM.

ΒΗΜΑΤΑ ΓΙΑ ΤΟ ΧΤΙΣΙΜΟ ΤΗΣ ΣΥΛΛΟΓΗΣ

Με την ολοκλήρωση της σύνδεσης, μια νέα σελίδα εμφανίζεται στην οποία φαίνεται η σειρά των βημάτων που χρειάζονται για το 'χτίσιμο' μιας συλλογής:

1. Πληροφορίες συλλογής (collection information),
2. Στοιχεία πηγής (source data),
3. Διαμόρφωση της συλλογής (configuring the collection),
4. Οικοδόμηση της συλλογής (building the collection),
5. Εξέταση της συλλογής (viewing the collection).

Το πρώτο βήμα είναι να δώσουμε όνομα και σχετικές πληροφορίες για τη συλλογή. Το δεύτερο βήμα καθορίζει από που θα προέλθουν τα στοιχεία πηγής. Το τρίτο βήμα αφορά στη διαμόρφωση της συλλογής, κατά την οποία απαιτείται ιδιαίτερη γνώση του συστήματος και είναι πραγματικά για προχωρημένους χρήστες. Στο τέταρτο βήμα γίνεται όλη η υπολογιστική εργασία. Το σύστημα δημιουργεί όλους τους απαραίτητους δείκτες (indexes) και συγκεντρώνει τις πληροφορίες που απαιτούνται για να κάνουν τη συλλογή να λειτουργήσει. Τέλος, το πέμπτο βήμα είναι για να δούμε ότι η συλλογή έχει δημιουργηθεί.

Τα βήματα εμφανίζονται ως γραμμική ακολουθία κουμπιών στο κάτω σημείο της σελίδας του collector έτσι ώστε οι χρήστες να μπορούν να παρακολουθήσουν πού βρίσκονται. Τα κουμπιά αλλάζουν χρώμα για να απεικονίζουν το τρέχον σημείο.

ΠΛΗΡΟΦΟΡΙΕΣ ΣΥΛΛΟΓΗΣ

Στο πρώτο βήμα, 'πληροφορίες συλλογής' (*collection information*), πρέπει να εισαχθούν κάποιες πληροφορίες για τη νέα συλλογή όπως:

- Τίτλος (title),
- Διεύθυνση ηλεκτρονικού ταχυδρομείου (contact e-mail address) και,
- Σύντομη περιγραφή (brief description).

Ο τίτλος είναι μια σύντομη φράση που χρησιμοποιείται για να προσδιορίσει τη συλλογή. Η διεύθυνση ηλεκτρονικού ταχυδρομείου χρησιμοποιείται σε περίπτωση που οποιαδήποτε προβλήματα προκύψουν με τη συλλογή να αποσταλεί e-mail με διαγνωστική αναφορά. Επίσης ο χρήστης εισάγει ένα κείμενο που αναφέρει το τι

περιέχει η συλλογή. Το κείμενο αυτό εμφανίζεται στην πρώτη σελίδα όταν δημιουργηθεί η συλλογή.

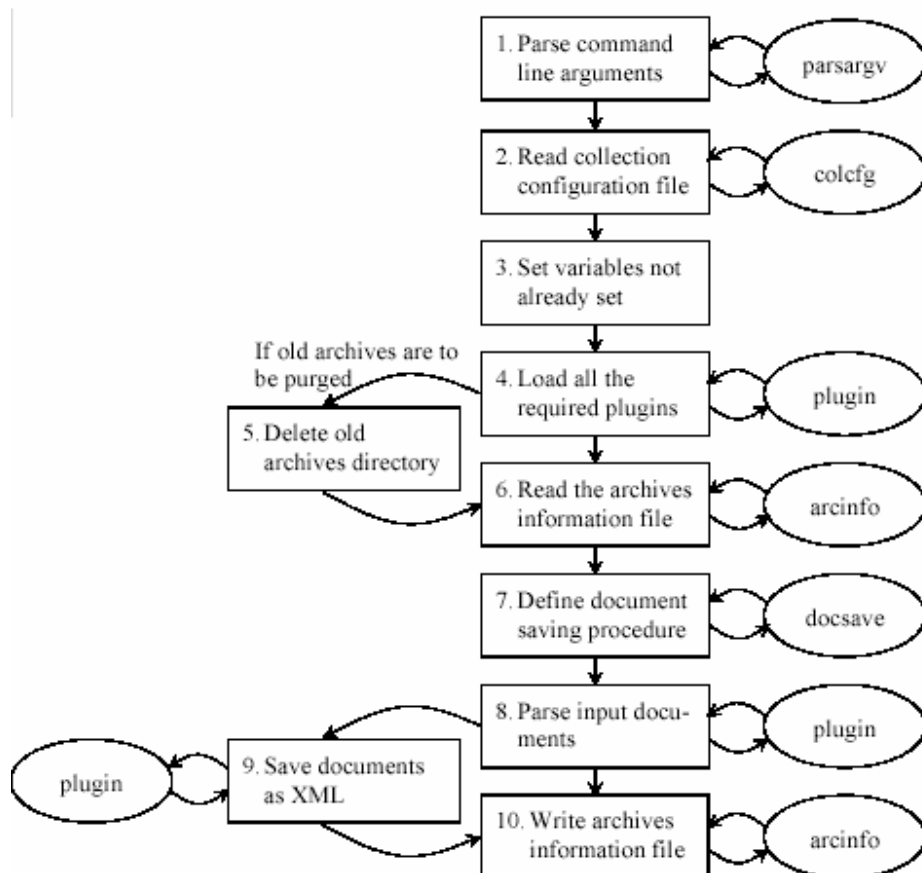
ΚΕΙΜΕΝΑ ΠΗΓΗΣ

Μετά από τη διαδικασία συλλογής πληροφοριών ο χρήστης διευκρινίζει ποια είναι και που είναι τα κείμενα πηγής (source data) που περιλαμβάνει η συλλογή. Η συλλογή μπορεί να είναι είτε νέα, απολύτως βασισμένη σε νέα δομή ή ανανέωση μιας ήδη υπάρχουσας.

Εάν επιλέξουμε τη νέα δομή, η νέα συλλογή μπορεί να περιέχει έγγραφα HTML (αρχεία που έχουν κατάληξη .html, .htm), ή plain text έγγραφα (αρχεία που έχουν κατάληξη σε .txt, .text), έγγραφα Microsoft Word (αρχεία με κατάληξη σε .doc), Έγγραφα PDF (αρχεία που έχουν κατάληξη σε .pdf) ή έγγραφα ηλεκτρονικού ταχυδρομείου (αρχεία που έχουν κατάληξη σε .elm, .email). Στην περίπτωση που βασίζεται τη νέα συλλογή σε μια υπάρχουσα, τα αρχεία στη νέα συλλογή πρέπει να είναι πανομοιότυπα με εκείνα που χρησιμοποιήθηκαν για να χτιστεί η προηγούμενη συλλογή.

Στις γραμμές επιλογής καθορίζουμε πού βρίσκονται τα έγγραφα πηγής: μέχρι τρεις χωριστές πηγές εισαγωγής μπορούν να διευκρινιστούν. Εάν χρειάζονται περισσότερες γίνεται 'κλικ' στο κουμπί 'MORE SOURCES',

- ένα όνομα καταλόγου στον εξυπηρετητή που φιλοξενεί το Greenstone (αρχίζοντας με "to file://"),
- μια διεύθυνση αρχίζοντας με "to HTTP://" για αρχεία που μεταφορτώνονται από το web,
- μια διεύθυνση αρχίζοντας με "to FTP://" για αρχεία που μεταφορτώνονται χρησιμοποιώντας το πρωτόκολλο FTP.



Σχήμα 5: Βήματα στη διαδικασία εισαγωγών

ΔΙΑΜΟΡΦΩΣΗΣ ΤΗΣ ΣΥΛΛΟΓΗΣ

Η κατασκευή και η παρουσίαση όλων των συλλογών ελέγχονται από τις προδιαγραφές που έχουμε ορίσει στο ειδικό αρχείο διαμόρφωσης της κάθε συλλογής (configuration file). Προχωρημένοι χρήστες μπορούν να χρησιμοποιήσουν αυτήν τη σελίδα για να αλλάξουν τις ρυθμίσεις διαμόρφωσης. Οι περισσότεροι, όμως μπορούν να προχωρήσουν άμεσα στο τελικό στάδιο.

creator	davidb@cs.waikato.ac.nz	1
maintainer	davidb@cs.waikato.ac.nz	2
public	True	3
		4
indexes	document:text	5
defaultindex	document:text	6
plugins	GMLPlug TEXTPlug ArcPlug RecPlug	7
		8
classify	AZList metadata=Title	9
		10
collectionmeta	collectionname "generic text collection"	11
collectionmeta	.document:text "documents"	12

creator	davidb@cs.waikato.ac.nz	1
maintainer	davidb@cs.waikato.ac.nz	2
public	True	3
		4
indexes	document:text document:From	5
defaultindex	document:text	6
plugins	GMLPlug EMAILPlug ArcPlug RecPlug	7
		8
classify	AZList metadata=Title	9
classify	DateList	10
		11
collectionmeta	collectionname "Email messages"	12
collectionmeta	.document:text "documents"	13
collectionmeta	.document:From "email senders"	14
		15
format	QueryResults \	16
	<td>[link][icon]/link</td><td>[Title]</td><td>[Author]</td>	17

Σχήμα 6: Παράδειγμα αρχείου διαμόρφωσης (configuration file)

ΧΤΙΣΙΜΟ ΤΗΣ ΣΥΛΛΟΓΗΣ

Μέχρι αυτή τη φάση, οι απαντήσεις στους διάλογους έχουν καταγραφεί σε ένα προσωρινό αρχείο. Στο στάδιο οικοδόμησης πραγματοποιείται η όλη διαδικασία. Κατά τη διάρκεια της οικοδόμησης, κατασκευάζονται οι δείκτες (indexes) για το ξεφύλλισμα (browsing) και για την έρευνα σύμφωνα με τις οδηγίες στο αρχείο διαμόρφωσης της συλλογής.

Η διαδικασία οικοδόμησης παίρνει κάποιο χρόνο ανάλογα με το μέγεθος της συλλογής, από κάποια λεπτά μέχρι μερικές ώρες, ανάλογα και με την ταχύτητα του υπολογιστή μας. Μερικές πολύ μεγάλες συλλογές παίρνουν μια ημέρα ή και περισσότερες για να χτιστούν.

Όταν φθάσουμε σ' αυτό το στάδιο, μια γραμμή στο κάτω σημείο της σελίδας, δίνει τις πληροφορίες για τον τρόπο με τον οποίο η λειτουργία προχωρά. Η ενημέρωση γίνεται κάθε πέντε δευτερόλεπτα.

Μπορούμε να σταματήσουμε τη διαδικασία οικοδόμησης οποιαδήποτε στιγμή με 'κλικ' στο 'stop building'.

ΕΞΕΤΑΣΗ ΤΗΣ ΣΥΛΛΟΓΗΣ (VIEWING THE COLLECTION)

Όταν το χτίσιμο της συλλογής έχει ολοκληρωθεί εμφανίζεται στο κάτω σημείο της σελίδας το κουμπί 'view collection' ενεργό. Κάνοντας 'κλικ' στο 'view collection' ο χρήστης μεταφέρεται άμεσα στην πρόσφατα χτισμένη συλλογή.

Τέλος, υπάρχει η δυνατότητα αποστολής e-mail στη διεύθυνση ηλεκτρονικού ταχυδρομείου που ορίσαμε όταν γινόταν η δημιουργία της συλλογής, και στο διαχειριστή του συστήματος, όταν μια συλλογή δημιουργηθεί.

ΜΕΤΑΔΕΔΟΜΕΝΑ ΕΓΓΡΑΦΩΝ

Τα μεταδεδομένα είναι περιγραφικές πληροφορίες όπως: συντάκτης, τίτλος, ημερομηνία, λέξεις κλειδιά, κ.λ.π, οι οποίες συνδέονται με ένα έγγραφο. Τα μεταδεδομένα αποθηκεύονται με τα έγγραφα.

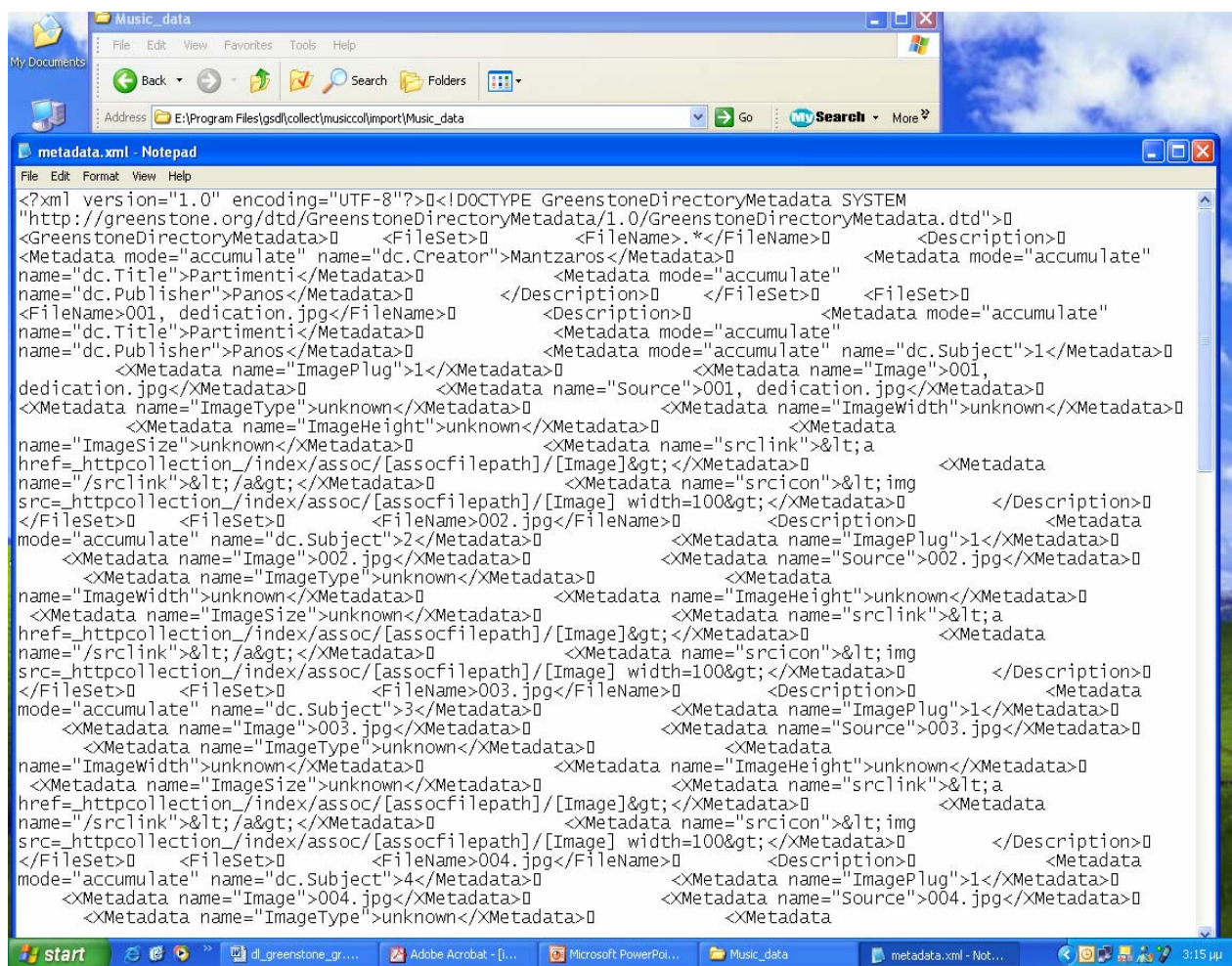
Το πρότυπο μεταδεδομένων του Dublin Core χρησιμοποιείται για τον καθορισμό των τύπων μεταδεδομένων που παρουσιάζεται στο Σχήμα 7.

Name	Metadata subtag	Definition
*Title	<i>Title</i>	A name given to the resource
*Creator	<i>Creator</i>	An entity primarily responsible for making the content of the resource
*Subject and keywords	<i>Subject</i>	The topic of the content of the resource
*Description	<i>Description</i>	An account of the content of the resource
*Publisher	<i>Publisher</i>	An entity responsible for making the resource available
Contributor	<i>Contributor</i>	An entity responsible for making contributions to the content of the resource
*Date	<i>Date</i>	The date that the resource was published or some other important date associated with the resource.
Resource type	<i>Type</i>	The nature or genre of the content of the resource
Format	<i>Format</i>	The physical or digital manifestation of the resource
*Resource identifier	<i>Identifier</i>	An unambiguous reference to the resource within a given context: this is the object identifier or OID
*Source	<i>Source</i>	A reference to a resource from which the present resource is derived
*Language	<i>Language</i>	A language of the intellectual content of the resource
Relation	<i>Relation</i>	A reference to a related resource
*Coverage	<i>Coverage</i>	The extent or scope of the content of the resource
Rights management	<i>Rights</i>	Information about rights held in and over the resource

Σχήμα 7: Πρότυπο μεταδεδομένων Dublin Core

ΜΕΤΑΔΕΔΟΜΕΝΑ ΑΡΧΕΙΩΝ

Τα μεταδεδομένα για όλα τα έγγραφα παρέχονται σε ένα ενιαίο αρχείο *metadata.xml* στον κατάλογο εισαγωγών(import) όπως φαίνονται στο Σχήμα 8.



Σχήμα 8: Αρχείο *metadata.xml*

Αυτό το απόσπασμα παρουσιάζει την αρχή και μέρος του αρχείου, μεταδεδομένων για μερικά από τα στοιχεία της συλλογής. Καθορίζει τα μεταδεδομένα Τίτλου, Θεμάτων, Παραγωγού, Εκδότη.

PLUGINS

ΕΙΣΑΓΩΓΗ

Ευλύγιστα συστήματα ψηφιακών βιβλιοθηκών πρέπει να είναι σε θέση να δεχτούν έγγραφα και μεταδεδομένα σε πολλές διαφορετικές μορφές. Τα έγγραφα μπορεί να είναι

διαθέσιμα σε μορφές προσανατολισμένες στον παγκόσμιο ιστό όπως HTML και XML, ή σε σχήματα επεξεργαστών κειμένου όπως Microsoft Word (.doc, .rtf), ή σε γλώσσες περιγραφής σελίδων όπως PostScript (.ps) ή .pdf. Τα μεταδεδομένα μπορούν επίσης να είναι διαθέσιμα με πολλές διαφορετικές μορφές όπως ενσωματωμένα στα έγγραφα, σε αρχεία μεταδεδομένων, στους υπολογισμούς σε λογιστικά φύλλα (spread sheets), ή μπορεί να εξαχθούν από τα ίδια τα έγγραφα.

Έγγραφα και μεταδεδομένα εισάγονται χρησιμοποιώντας ειδικό λογισμικό (software modules) που ονομάζονται 'plugins'. Αυτά μετατρέπουν τα έγγραφα και τα μεταδεδομένα σε μια ομοιόμορφη XML παρουσίαση που χρησιμοποιούνται έπειτα σε όλο το σύστημα. Υπάρχουν plugins για τυποποιημένα σχήματα εγγράφων και τυποποιημένα σχήματα μεταδεδομένων. Τα plugins χρησιμοποιούνται ακόμη και για να ανέβουν την ιεραρχία καταλόγου στην οποία τα έγγραφα πηγής αποθηκεύονται - αυτό μας δίνει την δυνατότητα αντιστοίχισης μεταδεδομένων με τα κείμενα αναφοράς βασισμένα στο όνομα αρχείου. Νέα plugins μπορούν να δημιουργηθούν για να καλύψουν τις ανάγκες νέων κειμένων και τύπων μεταδεδομένων.

PLUGINS ΔΙΑΣΩΛΗΝΩΣΗ (PIPELINE)

Οι συλλογές Greenstone μπορεί να περιλαμβάνουν έγγραφα και μεταδεδομένα με διαφορετικά σχήματα, και για κάθε σχήμα πρέπει να υπάρξει ένα plugin που να μπορεί να το επεξεργαστεί. Η ιεράρχηση των plugins καθορίζεται από τη σειρά στην οποία παρατίθενται στο αρχείο διαμόρφωσης της συλλογής. Εννοιολογικά, υπάρχουν τρεις τύποι plugins: *δομικό (structural)*, *σήμανσης (markup)* και *εξαγωγής (extraction)*.

Η διαδικασία εισαγωγής ξεκινά με την προσθήκη του ονόματος του καταλόγου που περιέχει τα έγγραφα πηγής. Το όνομα αυτό περνά από plugin σε plugin έως ότου ένα από αυτά να δώσει σήμα ότι μπορεί να επεξεργαστεί το αρχείο. Σε κανονικές συνθήκες, οποιοδήποτε όνομα αρχείου που ονομάζει έναν κατάλογο υποβάλλεται σε επεξεργασία από ένα 'δομικό' αποκαλούμενο plugin το RecPlug (Recursive Plugin) που απαριθμεί όλα τα αρχεία στον ονομασμένο κατάλογο και τροφοδοτεί τα ονόματά τους, ένα προς ένα στη σωλήνωση. Γενικά αυτός ο κατάλογος περιλαμβάνει τους περαιτέρω υποκαταλόγους (sub-directories), τα οποία θα υποβληθούν σε επεξεργασία με τον ίδιο τρόπο. Το RecPlug είναι ο κανονικός τρόπος για να εξετάσει τη δομή καταλόγου που συνδέεται με μια συλλογή. Εντούτοις, μπορεί εύκολα να προσαρμοστεί σε ιδιαίτερες

καταστάσεις, όπως όταν ορισμένοι τύποι αρχείων ή κατάλογοι χρειάζονται ειδική επεξεργασία, ή τα ονόματα των καταλόγων είναι σημαντικά για λόγους ανάθεσης μεταδεδομένων.

Τα έγγραφα της Microsoft Word, για να πάρουμε ένα παράδειγμα ενός τύπου εγγράφων, υποβάλλονται σε επεξεργασία XML από ένα "markup" plugin αποκαλούμενο WORDPlug. Το κείμενο και τα μεταδεδομένα που προσδιορίζονται από το WORDPlug περνούν σε μια δεύτερη σωλήνωση 'extraction' plugins, που καθορίζεται πάλι από το αρχείο διαμόρφωσης της συλλογής. Τα plugins αυτά εξάγουν τα μεταδεδομένα από το αρχείο, όπως τα αρκτικόλεξα, οι ημερομηνίες και οι διευθύνσεις ηλεκτρονικού ταχυδρομείου.

Εάν ένα στοιχείο φθάνει στο τέλος της σωλήνωσης χωρίς επεξεργασία από οποιοδήποτε plugin, το σύστημα παράγει ένα μήνυμα και πηγαίνει στο επόμενο στοιχείο. Η διαδικασία σταματά όταν η σειρά αναμονής είναι κενή.

Document Processing Plugins

Ο παρακάτω πίνακας δείχνει τα plugins που χρησιμοποιούνται για τις ευρέως χρησιμοποιούμενες μορφές εγγράφων. Το TEXTPlug ερμηνεύει ένα αρχείο κειμένου ως απλό έγγραφο. Το HTMLPlug χειρίζεται αρχεία HTML. Εξάγει τα μεταδεδομένα *τίτλου* βασισμένα στο HTML <title> tag μπορούν επίσης να εξαχθούν άλλα μεταδεδομένα σύμφωνα με την σύνταξη του HTML metatag. Αναλύει επίσης και επεξεργάζεται οποιουσδήποτε συνδέσμους (links) που περιέχει το αρχείο. Σύνδεσμοι με άλλα αρχεία στη συλλογή αντικαθίστανται από αναφορές στα αντίστοιχα έγγραφα μέσα στην ψηφιακή βιβλιοθήκη.

Το WORDPlug εισάγει έγγραφα του Microsoft Word. Υπάρχουν πολλές παραλλαγές σ' αυτή τη μορφή εγγράφων, ακόμα και προγράμματα της Microsoft έχουν μερικές φορές αποτυχίες μετατροπής. Το PDFPlug εισάγει έγγραφα σε PDF μορφή. Και τα δύο αυτά plugins χρησιμοποιούν ανεξάρτητα προγράμματα για να μετατρέψουν τα πηγαία αρχεία σε έγγραφα HTML.

Τα PSPlug εισάγει έγγραφα σε PostScript μορφή. Λειτουργεί καλύτερα εάν υπάρχει ήδη εγκατεστημένο στον υπολογιστή ένα πρόγραμμα μετατροπής, αλλά αν δεν υπάρχει το σύστημα κατ' ανάγκη καταλήγει σε έναν απλό αλγόριθμο εξαγωγής κειμένων.

EMAILPlug εισάγει αρχεία που περιέχουν e-mail, και εξετάζουν μορφές αρχείων που χρησιμοποιούνται από το Netscape, το Eudora. Κάθε έγγραφο πηγής εξετάζεται για να δει εάν περιέχει ένα ηλεκτρονικό ταχυδρομείο, ή αρκετοί που ενώνονται μαζί σε ένα αρχείο.

ZIPPlug εισάγει συμπιεσμένα αρχεία. Στηρίζεται σε προγράμματα που είναι εγκατεστημένα στον υπολογιστή μας.

<i>TEXTPlug</i>	Plain text.
<i>HTMLPlug</i>	HTML, replacing hyperlinks appropriately.
<i>WordPlug</i>	Microsoft Word documents.
<i>PDFPlug</i>	PDF documents.
<i>PSPlug</i>	PostScript documents.
<i>EMAILPlug</i>	E-mail messages, recognizing author, subject, date, etc.
<i>BibTexPlug</i>	Bibliography files in <i>BibTex</i> format.
<i>ReferPlug</i>	Bibliography files in <i>refer</i> format.
<i>SRCPlug</i>	Source code files.
<i>ImagePlug</i>	Image files for creating a library of images.
<i>SplitPlug</i>	Splits a document file into parts.
<i>ZIPPlug</i>	Uncompresses files.
<i>BookPlug</i>	Specially marked-up HTML.
<i>GBPlug</i>	Project Gutenberg E-text.
<i>TCCPlug</i>	E-mail documents from Computists' Weekly.
<i>PrePlug</i>	HTML output from the PRESCRIPT program.

Σχήμα 9: Greenstone plugins

ΤΑΞΙΝΟΜΗΤΕΣ

Οι ταξινομητές (Classifiers) χρησιμοποιούνται για να δημιουργήσουν τους δείκτες ξεφυλλίσματος (browsing index) μιας συλλογής.

	Argument	Purpose
<i>Hierarchy</i>	<i>hfile</i>	Hierarchical classification
	<i>metadata</i>	Classification file
	<i>sort</i>	Metadata element to test against <i>hfile</i> identifier
	<i>buttonname</i>	Metadata element used to sort documents within leaves (defaults to <i>Title</i>)
<i>List</i>		Name of the button used to access this classifier (defaults to value of <i>metadata</i> argument)
	<i>metadata</i>	Alphabetic list of documents
	<i>buttonname</i>	Include documents containing this metadata element
<i>SectionList</i>		Name of button used to access this classifier (defaults to value of <i>metadata</i> argument)
<i>AZList</i>		List of sections in documents
	<i>metadata</i>	List of documents split into alphabetical ranges
	<i>buttonname</i>	Include all documents containing this metadata element
<i>AZSectionList</i>		Name of button used to access this classifier (defaults to value of <i>metadata</i> argument)
<i>DateList</i>		Like <i>AZList</i> but includes every section of the document
		Similar to <i>AZList</i> but sorted by date

Σχήμα 10: Ταξινομητές Greenstone

Οι ταξινομητές, όπως και τα plugins ορίζονται στο αρχείο διαμόρφωσης μιας συλλογής. Για καθέναν υπάρχει μια γραμμή που ξεκινά από μια λέξη κλειδί που *ταξινομεί* και ακολουθείται από το όνομα του ταξινομητή και τις οποίες επιλογές αυτός παίρνει. Το βασικό αρχείο της διαμόρφωσης μιας συλλογής περιλαμβάνει τη γραμμή '*classify AZList –metadata Title*', που δημιουργεί έναν αλφαβητικό κατάλογο τίτλων, τους ταξινομεί και τους διαχωρίζει σε αλφαβητικές σειρές.

Ένας απλούστερος ταξινομητής, αποκαλούμενος '*List*', δημιουργεί έναν ταξινομημένο κατάλογο ενός δεδομένου στοιχείου μεταδεδομένων και τον επιδεικνύει χωρίς οποιεσδήποτε αλφαβητικές υποενότητες.

Το *DateList*, παράγει έναν κατάλογο επιλογής σειρών ημερομηνίας. Άλλοι ταξινομητές παράγουν τις δομές ξεφυλλίσματος που είναι ρητά ιεραρχικές. Οι ιεραρχικές ταξινομήσεις είναι χρήσιμες για τις θεματικές ταξινομήσεις και τις υποκατηγορίες τους.

Όλοι οι ταξινομητές παράγουν μια ιεραρχική δομή που χρησιμοποιείται για να φτιαχτούν οι δείκτες ξεφυλλίσματος. Τα χαμηλότερα επίπεδα της ιεραρχίας είναι συνήθως έγγραφα, αλλά σε μερικούς ταξινομητές είναι τμήματα. Οι εσωτερικοί κόμβοι της

ιεραρχίας είναι είτε *Vlist*, *Hlist*, είτε *Datelist*. Ένα *Vlist* είναι ένας κατάλογος στοιχείων που επιδεικνύεται κάθετα (vertical) κάτω στην σελίδα. Ένα *Hlist* επιδεικνύεται οριζόντια.

Οι γραμμές που χρησιμοποιούνται στα αρχεία διαμόρφωσης των συλλογών για να διευκρινίσουν τους ταξινομητές περιέχουν έναν όρο *μεταδεδομένων* που προσδιορίζει τα μεταδεδομένα σύμφωνα με τον τρόπο που τα έγγραφα είναι ταξινομημένα και σε σειρά. Οποιοδήποτε έγγραφο στη συλλογή που δεν καθορίζεται από αυτά τα μεταδεδομένα θα παραλειφθεί από τον ταξινομητή.

LIST CLASSIFIERS

Οι διάφορες κατηγορίες των ταξινομητών καταλόγων (list classifiers) παρουσιάζονται παρακάτω:

- *SectionList*,
- *AZList*,
- *AZSectionList*,
- *DateList*

ΔΙΑΜΟΡΦΩΣΗ ΤΗΣ ΔΙΕΠΑΦΗΣ

Οι ιστοσελίδες που βλέπουμε κατά τη χρησιμοποίηση του Greenstone δεν είναι στατικές αλλά δυναμικές, δηλαδή δεν είναι κάπου αποθηκευμένες αλλά παράγονται δυναμικά σύμφωνα με τη ζήτηση. Η εμφάνιση πολλών σημείων στις ιστοσελίδες ελέγχεται χρησιμοποιώντας κάποια 'format strings'. Τα 'format strings' βρίσκονται στο αρχείο διαμόρφωσης της συλλογής. Υπάρχουν δύο διαφορετικά είδη στοιχείων στις σελίδες που ελέγχονται από τα 'format strings'. Το πρώτο περιλαμβάνει τα στοιχεία στη σελίδα που παρουσιάζουν τα έγγραφα ή μέρη των εγγράφων. Το δεύτερο περιλαμβάνει τους καταλόγους που παράγονται από τους ταξινομητές και τις αναζητήσεις. Όλα τα 'format strings' ερμηνεύονται στο χρόνο που εμφανίζονται οι σελίδες. Δεδομένου ότι εφαρμόζονται μόλις σώσουμε τις οποιεσδήποτε αλλαγές στο αρχείο διαμόρφωσης *collect.cfg*, ο πειραματισμός με τα 'format strings' είναι γρήγορος και εύκολος.

Ο παρακάτω πίνακας παρουσιάζει τις δηλώσεις σχήματος που έχουν επιπτώσεις στο πως φαίνονται τα έγγραφα.

<i>format DocumentImages true/false</i>	If <i>true</i> , display a cover image at the top left of the document page (default <i>false</i>).
<i>format DocumentHeading formatstring</i>	If <i>DocumentImages</i> is <i>false</i> , the format string controls how the document header shown at the top left of the document page looks (default <i>[Title]</i>).
<i>format DocumentContents true/false</i>	Display table of contents (if document is hierarchical), or next/previous section arrows and “page k of n” text (if not).
<i>format DocumentButtons string</i>	Controls the buttons that are displayed on a document page (default <i>Detach Highlight</i>).
<i>format DocumentText formatstring</i>	Format of the text to be displayed on a document page: default <pre><center><table width=537> <tr><td>[Text]</td></tr> </table></center></pre>
<i>format DocumentArrowsBottom true/false</i>	Display next/previous section arrows at bottom of document page (default <i>true</i>).
<i>format DocumentUseHTML true/false</i>	If <i>true</i> , each document is displayed inside a separate frame. The Preferences page will also change slightly, adding options applicable to a collection of HTML documents, including the ability to go directly to the original source document (anywhere on the Web) rather than to the Greenstone copy.

Σχήμα 11: Επιλογές Διαμόρφωσης

FORMATTING GREENSTONE LISTS

Τα ‘format strings’ που ελέγχουν πώς φαίνονται οι κατάλογοι μπορούν να εφαρμοστούν σε διαφορετικά επίπεδα της δομής επίδειξης. Μπορούν να αλλάξουν όλους τους καταλόγους ενός ορισμένου τύπου μέσα σε μια συλλογή (π.χ. *DateList*), ή όλα τα μέρη ενός καταλόγου (π.χ. όλες οι καταχωρήσεις στον κατάλογο *search list*), ή συγκεκριμένα μέρη ενός ορισμένου καταλόγου (π.χ. το κάθετο μέρος καταλόγων ενός ταξινομητή *AZList* στον τίτλο).

ΣΥΜΠΕΡΑΣΜΑΤΑ

Το Greenstone είναι ένα πλούσιο σε περιεχόμενο λογισμικό για συλλογές ψηφιακών βιβλιοθηκών. Χτίζει τις δομές δεδομένων για την έρευνα και το ξεφύλλισμα από το υλικό που παρέχεται, χωρίς να χρειάζεται την οποιαδήποτε επεξεργασία με το χέρι. Η διαδικασία ελέγχεται από ένα αρχείο διαμόρφωσης, και μόλις δημιουργηθεί μια συλλογή, νέο υλικό μπορεί να προστεθεί εντελώς αυτόματα. Το ξεφύλλισμα είναι βασισμένο στα μεταδεδομένα Dublin Core.

Νέες συλλογές μπορούν να αναπτυχθούν εύκολα, ιδιαίτερα εάν έχουν ομοιότητες με υπάρχουσες. Επεκτασιμότητα επιτυγχάνεται μέσω του λογισμικού των 'plugins' που μπορεί να γραφτεί για να προσαρμοστεί σε έγγραφα, και μεταδεδομένα, με διαφορετική μορφή. Πρότυπα plugins υπάρχουν για πολλούς τύπους εγγράφων ενώ νέα γράφονται σχετικά εύκολα. Το ξεφύλλισμα (browsing) ελέγχεται από τους ταξινομητές 'classifiers' που επεξεργάζονται τα μεταδεδομένα στις δομές ξεφυλλίσματος (κατά ημερομηνία, αλφαβητικά, ιεραρχικά, κ.λπ).

Εντούτοις, η ισχυρότερη υποστήριξη για την επεκτασιμότητα επιτυγχάνεται όχι με τεχνικά μέσα αλλά με το να καταστήσει τον πηγαίο κώδικα ελεύθερα διαθέσιμο με δημόσια άδεια GNU. Μόνο μέσω μιας διεθνούς συνεταιριστικής προσπάθειας το ψηφιακό λογισμικό βιβλιοθηκών θα γίνει αρκετά περιεκτικό για να ικανοποιήσει τις παγκόσμιες ανάγκες με την αφθονία και την ευελιξία που οι χρήστες αξίζουν να έχουν.

ΑΝΑΦΟΡΕΣ

1. 'Greenstone Digital Library Developer's Guide', D. Bainbridge, D. McKay and I.H. Witten, March 2003
2. 'Greenstone Digital Library User's Guide', I.H. Witten and St. Boddie, March 2003
3. 'Greenstone: A Comprehensive Open-Source Digital Library Software System', I.H. Witten. R.J. McNab, St. Boddie, D. Bainbridge, 1999.
4. 'Defining Collections in Distributed Digital Libraries', C. Lagoze, and D. Fielding, *D-Lib Magazine*, Nov. 1998.
5. 'Inside Greenstone Collections', I.H. Witten, D. Bainbridge, St. Boddie, J.D. Kathy, J.R. McPherson, June 2003.
6. 'Greenstone FAQ' <http://www.greenstone.org/cgi-bin/library?e=p-en-home-utfZz-8&a=p&p=faq>, 2003.
7. 'New Zealand Digital Library', <http://www.sadl.uleth.ca/nz/cgi-bin/library>, 2003.
8. 'A Time for Digital Libraries', G. Goodal <http://www.deregulo.com/facetation>, June 2003.