

ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ

**ΣΥΓΓΡΑΦΗ ΚΑΙ ΠΑΡΟΥΣΙΑΣΗ ΕΓΓΡΑΦΩΝ ΣΕ
ΕΞΑΤΟΜΙΚΕΥΜΕΝΑ ΣΥΣΤΗΜΑΤΑ
ΝΕΩΝ ΚΑΤ' ΑΠΑΙΤΗΣΗ**

ΜΙΧΑΗΛ Α. ΠΕΤΡΟΠΟΥΛΟΣ

Διπλωματική Εργασία που υποβλήθηκε
στα πλαίσια των απαιτήσεων
για την απόκτηση διπλώματος στο

ΤΜΗΜΑ ΗΛΕΚΤΡΟΝΙΚΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

Αύγουστος 1998

ΠΕΡΙΛΗΨΗ

Η παρούσα αναφορά περιγράφει το σχεδιασμό και την υλοποίηση των υποσυστημάτων συγγραφής άρθρων και παρουσίασης νέων του εξατομικευμένου συστήματος νέων κατ' απαίτηση "Hypermedia Custom News System". Σκοπός της ανάπτυξης των υποσυστημάτων αυτών είναι η δημιουργία ενός ολοκληρωμένου περιβάλλοντος συγγραφής και διαχείρισης εγγράφων για τους αρθρογράφους, καθώς και η παροχή σύγχρονων υπηρεσιών νέων στους τελικούς χρήστες του συστήματος μέσω των τεχνολογιών εξατομικευμένων νέων και νέων κατ' απαίτηση, που απαιτούν τα σημερινά πληροφοριακά συστήματα στο διαδίκτυο.

Το υποσύστημα συγγραφής άρθρων ενσωματώνει στο μοντέλο εγγράφου υποσύνολο αυτού της HTML, παρέχοντας αυξημένες δυνατότητες μορφοποίησης εγγράφου, διαχείρισης εγγράφων και πολυμέσων του συστήματος, καθώς και υποστήριξη συγγραφής και συσχέτισης μεταφραζόμενων άρθρων σε πολλαπλές γλώσσες. Επίσης προσαρμόστηκε σε αυτό υποσύστημα υποβολής ερωτήσεων στο αρχείο του συστήματος και υποσύστημα κατηγοριοποίησης εγγράφων και πολυμέσων [Καρασ98]. Το υποσύστημα παρουσίασης νέων, πέρα από τις βασικές υπηρεσίες νέων (παρουσίαση άρθρων, πλοήγηση, αναζήτηση), φροντίζει για την παρουσίαση των εξατομικευμένων νέων και των νέων κατ' απαίτηση στους τελικούς χρήστες, έτσι όπως αυτά προκύπτουν από τις αποθηκευμένες προτιμήσεις αυτών στα προσωπικά τους διαγράμματα. Οι υπηρεσίες αυτές μπορούν να παρουσιαστούν επίσης σε πολλαπλές γλώσσες.

Ιδιαίτερη έμφαση δόθηκε στην υποστήριξη αυθαίρετου αριθμού γλωσσών που αποτελεί ένα από τα σημαντικότερα χαρακτηριστικά του συστήματος. Συγκεκριμένα, παρέχεται η δυνατότητα αύξησης του αριθμού γλωσσών που υποστηρίζει το σύστημα χωρίς να χρειάζεται καμία αλλαγή σε κάποιο από τα υποσυστήματά του.

Αφιέρωση

Στους γονείς και την αδελφή μου

ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να ευχαριστήσω τον καθηγητή κ. Σταύρο Χριστοδουλάκη για την επίβλεψη του στην ολοκλήρωση αυτής της διπλωματικής εργασίας και για τις εμπειρίες που μου πρόσφερε στα πλαίσια της εργασίας μου στο Εργαστήριο Διανεμημένων Πληροφοριακών Συστημάτων και Εφαρμογών Πολυμέσων (MU.S.I.C.) του τμήματος Η.Μ.Μ.Υ. του Πολυτεχνείου Κρήτης.

Θα ήθελα να ευχαριστήσω προκαταβολικά τον καθηγητή κ. Βασίλειο Διγαλάκη και τον καθηγητή κ. Παναγιώτη Τριανταφύλλου για το χρόνο που διέθεσαν για την ανάγνωση της παρούσας διπλωματικής εργασίας και τις τυχόν παρατηρήσεις τους.

Επίσης ευχαριστώ τον Μανόλη Φραγκονικολάκη και τον Άλκη Σερβετά για την επίβλεψη και τη βοήθειά τους στο σχεδιασμό τμημάτων της διπλωματικής εργασίας.

Ιδιαίτερα θα ήθελα να ευχαριστήσω τον Στέφανο Καρασαββίδη για τη αποδοτική συνεργασία και συμβίωση στον ίδιο χώρο εργασίας, καθώς και τον Γιώργο Τερεζάκη για τη συμβολή του στην υλοποίηση τμημάτων της διπλωματικής εργασίας. Τέλος, θα ήθελα να ευχαριστήσω όλους τους εργαζόμενους του MU.S.I.C. για την επαγγελματική και ηθική υποστήριξή τους στα χρόνια της συνεργασίας μας.

Μιχάλης Πετρόπουλος

Χανιά, Αύγουστος 1998

ΠΕΡΙΕΧΟΜΕΝΑ

1	ΕΙΣΑΓΩΓΗ	10
1.1	Αναγκαιότητα	12
1.2	Το σύστημα “Hypermedia Custom News System”	15
1.3	Στόχοι της διπλωματικής εργασίας	18
1.4	Δομή της διπλωματικής εργασίας	20
2	ΕΠΙΣΚΟΠΗΣΗ ΥΠΑΡΧΟΝΤΩΝ ΣΥΣΤΗΜΑΤΩΝ ΝΕΩΝ	22
2.1	Εισαγωγή στα Συστήματα Νέων	22
2.2	Κατηγορίες Συστημάτων Νέων	23
2.2.1	Συμβατικά Συστήματα Νέων	24
2.2.2	Εξατομικευμένα Συστήματα Νέων	24
2.2.3	Εξατομικευμένα Συστήματα Νέων Κατ’ απαίτηση	25
2.3	Υπάρχοντα Συστήματα Νέων	26
2.4	Ανακεφαλαίωση	29
3	ΕΠΙΣΚΟΠΗΣΗ ΑΡΧΙΤΕΚΤΟΝΙΚΗΣ ΣΥΣΤΗΜΑΤΟΣ	30
3.1	Σύστημα Παραγωγής Άρθρων	32
3.1.1	Μηχανισμός Ασφαλείας	32
3.1.2	Υποσύστημα Συγγραφής Άρθρων	32
3.1.3	Υποσύστημα Κατηγοριοποίησης	33
3.1.4	Υποσύστημα Υποβολής Ερωτήσεων	33
3.2	Σύστημα Εξυπηρετητών	34
3.2.1	Εξυπηρετητής Βάσης Δεδομένων	34
3.2.2	Εξυπηρετητής Παγκόσμιου Ιστού	35
3.2.3	Εξυπηρετητής Μεταφοράς Αρχείων	35
3.3	Σύστημα Υπηρεσιών Νέων	35
3.3.1	Υποσύστημα Παρουσίασης Νέων	36
3.3.2	Υποσύστημα Υποβολής Ερωτήσεων	36
3.3.3	Υποσύστημα Εξατομικεύσεων	37
3.3.4	Υποσύστημα Νέων Κατ’ Απαίτηση	37
3.4	Ανακεφαλαίωση	38
4	Η ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ	39
4.1	Ανάλυση Απαιτήσεων	39
4.2	Διάγραμμα Οντοτήτων – Σχέσεων	41
4.3	Σχεσιακό Μοντέλο Βάσης Δεδομένων	42
4.4	Ανακεφαλαίωση	44
5	ΥΠΟΣΥΣΤΗΜΑ ΣΥΓΓΡΑΦΗΣ ΑΡΘΡΩΝ	45
5.1	Δομή Εγγράφου	45
5.2	Δομή Μέσου	49
5.3	Αρχιτεκτονική	50
5.4	Μηχανισμοί Συγγραφής Άρθρου	52
5.5	Μηχανισμός Παρουσίασης Αποτελεσμάτων Ερώτησης	53
5.6	Μηχανισμός Εισαγωγής Πολυμέσων	56
5.7	Μηχανισμοί Μορφοποίησης Άρθρου	57
5.7.1	Εισαγωγή Διασυνδέσεων εντός Κειμένου	57

5.7.2	Εισαγωγή Αντικειμένων εντός Κειμένου	60
5.7.3	Εισαγωγή Συσχετίσεων Άρθρου	62
5.8	Μηχανισμός Προεπισκόπησης Άρθρων	63
5.9	Μηχανισμοί Διαχείρισης Άρθρων	64
5.10	Ανακεφαλαίωση	67
6	ΥΠΟΣΥΣΤΗΜΑ ΠΑΡΟΥΣΙΑΣΗΣ ΝΕΩΝ	68
6.1	Υπηρεσίες Νέων	68
6.2	Αρχιτεκτονική	70
6.3	Μηχανισμός Παρουσίασης Αποτελεσμάτων Ερώτησης	70
6.4	Μηχανισμοί Πλοήγησης	73
6.5	Μηχανισμός Παρουσίασης Διαγραμμάτων	74
6.6	Μηχανισμός Παρουσίασης Νέων Κατ' Απαίτηση	76
6.7	Μηχανισμός Παρουσίασης Εγγράφων	77
6.8	Ανακεφαλαίωση	79
7	ΥΠΟΣΤΗΡΙΞΗ ΠΟΛΛΑΠΛΩΝ ΓΛΩΣΣΩΝ	80
7.1	Εισαγωγή	80
7.2	Σχεσιακό Μοντέλο Βάσης Δεδομένων	81
7.2.1	Το Σύστημα Νέων HyNoDe	81
7.2.2	Το Σύστημα Νέων “Hypermedia Custom News System”	82
7.3	Αναπαράσταση και χειρισμός κειμένου	86
7.3.1	Κανονισμός ASCII	86
7.3.2	Κανονισμός ISO-8859-x	87
7.3.3	Κανονισμός UNICODE	88
7.3.4	Αποθήκευση και ανάκτηση κειμένου με JDBC	88
7.4	Παρουσίαση πολυγλωσσικού κειμένου	93
7.5	Ανακεφαλαίωση	94
8	ΤΕΧΝΟΛΟΓΙΑ ΥΛΟΠΟΙΗΣΗΣ	95
8.1	Microsoft SQL Server	95
8.2	Φυλλομετρητές, HTTP και FTP πρωτόκολλα και εξυπηρετητές	96
8.3	Java	97
8.3.1	Java Applets	99
8.3.2	Java Foundation Classes (JFC)	100
8.3.3	Java Database Connectivity (JDBC)	101
8.3.4	Java Servlets	102
9	ΑΝΑΚΕΦΑΛΑΙΩΣΗ ΚΑΙ ΜΕΛΛΟΝΤΙΚΕΣ ΕΠΕΚΤΑΣΕΙΣ	104
9.1	Ανακεφαλαίωση	104
9.2	Μελλοντικές Επεκτάσεις	105
A	ΔΙΕΥΘΥΝΣΕΙΣ ΣΥΣΤΗΜΑΤΩΝ ΝΕΩΝ ΣΤΟ ΔΙΑΔΙΚΤΥΟ	108
	ΓΛΩΣΣΑΡΙ	109
	ΒΙΒΛΙΟΓΡΑΦΙΑ	110

ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ

Σχήμα 1-1	Γενική Αρχιτεκτονική του συστήματος εξατομικευμένων νέων κατ' απαίτηση "Hypermedia Custom News System"	16
Σχήμα 2-1	Σελίδα παρουσίασης εξατομικευμένων νέων του συστήματος "CNN Custom News"	27
Σχήμα 2-2	Ο μηχανισμός ειδοποίησης τελικού χρήστη του συστήματος "MSNBC"	28
Σχήμα 3-1	Αρχιτεκτονική του συστήματος νέων "Hypermedia Custom News System"	31
Σχήμα 4-1	Διάγραμμα οντοτήτων – σχέσεων του συστήματος νέων "Hypermedia Custom News System"	42
Σχήμα 4-2	Σχεσιακό μοντέλο της βάσης δεδομένων του συστήματος νέων "Hypermedia Custom News System"	43
Σχήμα 5-1	Δομή της οντότητας Εγγράφου	46
Σχήμα 5-2	Οντότητες Μορφοποίησης Εγγράφου	47
Σχήμα 5-3	Δομή της οντότητας Μέσου	50
Σχήμα 5-4	Αρχιτεκτονική του Υποσυστήματος Συγγραφής Άρθρων του Συστήματος Παραγωγής Νέων	51
Σχήμα 5-5	Δομές των οντοτήτων Επισήμανσης και Ενότητας	53
Σχήμα 5-6	Οι Μηχανισμοί Συγγραφής Άρθρου στο Υποσύστημα Συγγραφής Άρθρων	53
Σχήμα 5-7	Δομή Αποτελέσματος Εγγράφου	54
Σχήμα 5-8	Δομή Αποτελέσματος Μέσου	55
Σχήμα 5-9	Δομή Απόλυτης Internet Διεύθυνσης	58
Σχήμα 5-10	Οι Μηχανισμοί Μορφοποίησης Άρθρου στο Υποσύστημα Συγγραφής Άρθρων	63
Σχήμα 5-11	Οι Μηχανισμοί Διαχείρισης Άρθρων στο Υποσύστημα Συγγραφής Άρθρων	66
Σχήμα 6-1	Αρχιτεκτονική του Υποσυστήματος Παρουσίασης Νέων του Συστήματος Υπηρεσιών Νέων	71
Σχήμα 6-2	Δομή Αποτελέσματος Εγγράφου	72
Σχήμα 6-3	Η αρχική σελίδα του συστήματος "Hypermedia Custom News System"	74
Σχήμα 6-4	Ο Μηχανισμός Παρουσίασης Διαγραμμάτων του Υποσυστήματος Παρουσίασης Νέων	75
Σχήμα 6-5	Ο Μηχανισμός Παρουσίασης Εγγράφων του Υποσυστήματος Παρουσίασης Νέων	78

Σχήμα 7-1	Μορφή κειμένου 8-bit στην πλευρά της Βάσης Δεδομένων, και μορφή 16-bit στην πλευρά των Java εφαρμογών.....	90
Σχήμα 7-2	Προσαρμογή της μετατροπής 8-bit χαρακτήρων σε 16-bit, έτσι ώστε να παραμετροποιείται η κωδικοποίηση.....	92
Σχήμα 7-3	Η διαφορά αντιστοίχισης 8-bit χαρακτήρων σε 16-bit με χρήση του iso-8859-1 (Latin-1), και κάποιας άλλης κωδικοσελίδας...	92
Σχήμα 8-1	Τα βήματα που ακολουθούνται για την εκτέλεση ενός Java προγράμματος	99
Σχήμα 8-2	Ο ρόλος του JDBC στη σύνδεση με Βάσεις Δεδομένων και η διαφορά ανάμεσα στους οδηγούς τύπου 3 και 4.....	101
Σχήμα 8-3	Η αλληλεπίδραση των servlet με τη Βάση Δεδομένων και τους φυλλομετρητές	102

ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

Πίνακας 5-1	Δυνατότητες Διαχείρισης Διασυνδέσεων Εγγράφου.....	59
Πίνακας 5-2	Δυνατότητες Διαχείρισης Αντικειμένων εντός Κειμένου.....	61
Πίνακας 5-3	Δυνατότητες Διαχείρισης Συσχετίσεων Εγγράφου.....	62
Πίνακας 7-1	Τμήμα πίνακα αναπαράστασης λέξεων – κλειδιών στη βάση δεδομένων του συστήματος HyNoDe	81
Πίνακας 7-2	Τμήμα πίνακα αναπαράστασης εγγράφων στη βάση δεδομένων του συστήματος HyNoDe	82
Πίνακας 7-3	Πίνακας αναπαράστασης υποστηριζόμενων γλωσσών στη βάση δεδομένων του “Hypermedia Custom News System” ...	83
Πίνακας 7-4	Περιεχόμενα του πίνακα “Languages” για τις γλώσσες που υποστηρίζονται αυτή τη στιγμή από το “Hypermedia Custom News System”	83
Πίνακας 7-5	Πίνακας αναπαράστασης λέξεων – κλειδιών στη βάση δεδομένων του “Hypermedia Custom News System” ...	84
Πίνακας 7-6	Ο πίνακας “Translations” που συσχετίζει μεταφράσεις εγγράφων	85

1 ΕΙΣΑΓΩΓΗ

Τα τελευταία χρόνια η ανάπτυξη του διαδικτύου ήταν, όπως αναμενόταν, ραγδαία. Στην ωρίμανση του και στην καθιέρωση αυτού στην καθημερινή ζωή σημαντικό ρόλο έπαιξε η ανάγκη του απλού πολίτη πλέον, και όχι μόνο του εξειδικευμένου χρήστη, για πληρέστερη και αμεσότερη ενημέρωση. Ανάγκη την οποία καλούνται να καλύψουν τα συστήματα νέων.

Πρώτοι απ' όλους, στην τεχνολογία των συστημάτων νέων επένδυσαν οι αρμόδιοι για την πληροφόρηση του κοινού, τα Μ.Μ.Ε. Εφημερίδες, περιοδικά, ραδιοφωνικοί και τηλεοπτικοί σταθμοί και γενικότερα όλες οι παραδοσιακές μορφές ενημέρωσης έκαναν την εμφάνιση τους στο διαδίκτυο στα μέσα της δεκαετίας του 90, υπακούοντας στις εξελικτικές τάσεις της εποχής. Ακολούθησαν και άλλου τύπου εταιρείες ή οργανισμοί που απλά επιθυμούσαν να έχουν αμεσότερη επικοινωνία με τους πελάτες τους και το κοινό.

Τα πρώτα συστήματα νέων που έκαναν την εμφάνιση τους στο διαδίκτυο χαρακτηρίζονταν από έλλειψη δυναμικά μεταβαλλόμενου περιεχομένου, εσωτερικής οργάνωσης, εύκολης αλληλεπίδρασης με τους αρθρογράφους, ενώ ο τελικός χρήστης είχε μόνο τη δυνατότητα απλής πλοήγησης. Με λίγα λόγια, η διαδικασία συγγραφής και έκδοσης νέων στο διαδίκτυο απέιχε πολύ από τις συνηθισμένες μεθόδους παραγωγής νέων στα παραδοσιακά μέσα. Αυτές οι ανάγκες καλύπτονται με την ταυτόχρονη ανάπτυξη της τεχνολογίας σε διάφορους τομείς.

Το αυξημένο μέγεθος της πληροφορίας αναλαμβάνουν πλέον να οργανώσουν τα μοντέρνα συστήματα διαχείρισης βάσεων δεδομένων, που και αυτά πέρασαν με τη σειρά τους σε μια νέα φάση εξέλιξης, ακολουθώντας νέες κατευθύνσεις έρευνας και ανάπτυξης. Ενσωμάτωσαν μηχανισμούς επικοινωνίας με εξυπηρετητές παγκόσμιου ιστού και φυλλομετρητές για τη πλήρη εκμετάλλευση των δυνατοτήτων τους για την

οργάνωση, αποθήκευση, ανάκτηση και έκδοση των παραγόμενων νέων, και πλέον αποτελούν τη βάση ανάπτυξης συστήματος πληροφόρησης στον παγκόσμιο ιστό.

Η γλώσσα προγραμματισμού Java, από την άλλη, παρέχοντας ανεξαρτησία πλατφόρμας, επέτρεψε την ανάπτυξη εφαρμογών ικανών να προσφέρουν αυξημένες δυνατότητες, τόσο στους αρθρογράφους όσο και στους τελικούς χρήστες.

Ξεπερνώντας έτσι το αρχικό στάδιο του πειραματισμού με το διαδίκτυο, τα σημερινά συστήματα νέων προσπαθούν να προσαρμοστούν όσο το δυνατό καλύτερα στις ανάγκες των χρηστών, αλλά και των ανθρώπων που εργάζονται για την παραγωγή της πληροφορίας. Φθάνοντας έτσι στα τωρινά θέματα έρευνας και ανάπτυξης, αιχμή αποτελεί η τεχνολογία **εξατομίκευσης των νέων (custom news)**, καθώς και η τεχνολογία **νέων κατ' απαίτηση (news on demand)**. Η πρώτη έχει να κάνει με το είδος της πληροφορίας που θα παρουσιάζεται στον χρήστη, ενώ η δεύτερη έχει να κάνει με τον τρόπο που ο χρήστης θα βλέπει την πληροφορία αυτή.

Κατά την τεχνολογία εξατομίκευσης των νέων ο τελικός χρήστης έχει τη δυνατότητα να εξειδικεύει το περιεχόμενο της παρεχόμενης πληροφορίας σύμφωνα με τα δικά του ενδιαφέροντα επιλέγοντας τα κατάλληλα υποσύνολα του μοντέλου κατηγοριοποίησης που χρησιμοποιεί το σύστημα νέων και καθορίζοντας την περίοδο ανανέωσης των νέων. Με τον τρόπο αυτό προσδιορίζει ακριβέστερα τις ειδήσεις που θέλει να του παρουσιάζονται από το σύστημα κάθε φορά που επιθυμεί να ενημερωθεί από αυτό.

Κατά την τεχνολογία νέων κατ' απαίτηση ο τελικός χρήστης μπορεί να καθορίσει το χρόνο και τον τρόπο με τον οποίο θα παραλαμβάνει τα παραγόμενα νέα. Χρησιμοποιώντας κάποιο μηχανισμό ειδοποίησης έχει τη δυνατότητα να βρίσκεται χρονικά πλέον πολύ κοντά στην παραγωγή των νέων. Επίσης, μέσω κατάλληλου εργαλείου μπορεί να οργανώνει την παρουσίαση των νέων με τρόπο που αυτός επιθυμεί.

Η παρούσα διπλωματική διατριβή εκπονήθηκε στα πλαίσια της ανάπτυξης του συστήματος νέων **“Hypermedia Custom News System”**. Ένα σύστημα το οποίο εσωκλείνει την τεχνολογία εξατομίκευσης των νέων και την τεχνολογία των νέων

κατ' απαίτηση και έχει ως στόχο να ενσωματωθεί σε προγράμματα του Εργαστηρίου Διανεμημένων Πληροφοριακών Συστημάτων και Εφαρμογών Πολυμέσων (MU.S.I.C.) του τμήματος Η.Μ.Μ.Υ. του Πολυτεχνείου Κρήτης. Ειδικότερα, η διατριβή επικεντρώνεται στην ανάπτυξη του **Υποσυστήματος Συγγραφής Άρθρων (News Authoring System)** καθώς και στο **Υποσύστημα Παρουσίασης Νέων (News Presentation System)** αυτών στο διαδίκτυο.

Τέλος, για την εκπόνηση της διπλωματικής διατριβής το περιβάλλον και η υποστήριξη από αυτό μπορεί να θεωρηθούν ιδανικά, αφού ήδη υπάρχει σημαντική εμπειρία από συναδέλφους, οι οποίοι έχουν συμμετάσχει σε ευρωπαϊκά ερευνητικά προγράμματα σχετικά με το αντικείμενο. Η εμπειρία αυτή αποτελεί βάση πάνω στη οποία στηρίζεται η ανάπτυξη και υλοποίηση της παρούσας διατριβής. Η τεχνογνωσία που έχει αποκτηθεί και προτείνεται για τη λύση ορισμένων προβλημάτων που παρουσιάζει η δημιουργία ενός τέτοιου συστήματος είναι γνώση ιδιαίτερα χρήσιμη και ικανή να μειώσει το χρόνο υλοποίησης και να αυξήσει την ποιότητα του τελικού αποτελέσματος.

1.1 Αναγκαιότητα

Πολλά από τα προγράμματα που αναλαμβάνει το MU.S.I.C. κινούνται στο χώρο του τουρισμού και του πολιτισμού και έχουν ανάγκη από συστήματα νέων, τα οποία πρέπει να καλύπτουν όλες τις βασικές ανάγκες για παραγωγή και έκδοση νέων στο διαδίκτυο και να υλοποιούν τεχνολογίες αιχμής επιθυμητές από τους τελικούς χρήστες. Επίσης, από τη φύση των προγραμμάτων επιβάλλεται η υποστήριξη αυθαίρετου αριθμού γλωσσών. Η ανάγκη αυτή δεν μπορούσε να καλυφθεί με την αγορά ενός έτοιμου προϊόντος, γιατί οι δυνατότητες που προσφέρουν τα συστήματα αυτά είναι πολύ περιορισμένες. Συστήματα με αυξημένες δυνατότητες υπάρχουν στο διαδίκτυο, αλλά έχουν αναπτυχθεί κυρίως για συγκεκριμένα ειδησεογραφικά πρακτορεία και σε συνεργασία με εταιρείες ανάπτυξης λογισμικού.

Το MU.S.I.C. έχει επίσης συμβάλλει στην ανάπτυξη ενός συστήματος εξατομικευμένων νέων κατ' απαίτηση, στα πλαίσια του ευρωπαϊκού ερευνητικού προγράμματος HyNoDe, αλλά χωρίς να έχει δικαίωμα χρήσης του όλου συστήματος.

Οπότε η ανάπτυξη ενός συστήματος νέων από το ίδιο το MU.S.I.C. κρίθηκε επιβεβλημένη. Αυτή ήταν η αναγκαιότητα και ο σκοπός της δημιουργίας του συστήματος νέων **“Hypermedia Custom News System”**.

Σε ένα τέτοιο σύστημα το υποσύστημα συγγραφής άρθρων αποτελεί ένα από τα βασικότερα τμήματα. Μέσω αυτού τροφοδοτείται το σύστημα με πληροφορία, η οποία πρέπει να είναι όσο το δυνατόν πληρέστερη. Η εκμάθηση και χρήση του οφείλει να είναι εύκολη ακόμα και για αρθρογράφους οι οποίοι δεν έχουν μεγάλη πείρα με υπολογιστές. Αναπτύσσεται σύμφωνα με την μοντελοποίηση που γίνεται για κάθε οντότητα του συστήματος και το σχεδιασμό της βάσης δεδομένων, και πρέπει να καλύπτει τουλάχιστον τις ελάχιστες απαιτήσεις των αρθρογράφων για μορφοποίηση κειμένου και εισαγωγή αντικειμένων μέσα στο κείμενο (π.χ. hyperlinks, εικόνες). Συγκεκριμένα καλείται να καλύψει πλήρως τις ακόλουθες απαιτήσεις:

- Διαδικασία συγγραφής άρθρων από οποιοδήποτε σημείο του διαδικτύου, εύκολα και γρήγορα, η οποία να είναι όσο το δυνατόν πανομοιότυπη με την παραδοσιακή που ήδη χρησιμοποιούν οι αρθρογράφοι. Η παραπάνω διαδικασία πρέπει να μπορεί να εκτελείται ταυτόχρονα από πολλούς αρθρογράφους σε πολλούς διαφορετικούς σταθμούς εργασίας.
- Δυνατότητα συγγραφής από αρθρογράφους που εργάζονται σε διαφορετικούς οργανισμούς ή είναι ανεξάρτητοι, εξασφαλίζοντας παράλληλα και την ασφάλεια του περιεχομένου των άρθρων τους. Με αυτόν τον τρόπο οι τελικοί χρήστες έχουν πολύπλευρη ενημέρωση χρησιμοποιώντας ένα σύστημα.
- Δυνατότητες μορφοποίησης εγγράφου στους αρθρογράφους παρόμοιες με αυτές που παρέχουν οι κλασικοί επεξεργαστές κειμένου (Text editors) και οι επεξεργαστές HTML σελίδων (HTML editors).
- Υποστήριξη αυθαίρετου αριθμού γλωσσών, αφού στο σύστημα θα έχουν πρόσβαση χρήστες πολλών εθνικοτήτων.

- Δυνατότητα αναζήτησης από τον αρθρογράφο, μέσω ενός κατάλληλα προσαρμοσμένου υποσυστήματος, στο αρχείο του συστήματος ώστε να μπορεί να συνδέσει τις ιστορίες μεταξύ τους ή να συμπληρώσει την ιστορία του με πολυμέσα (π.χ. εικόνες, ήχους, video) που ήδη υπάρχουν.
- Δυνατότητα κατηγοριοποίησης και χαρακτηρισμού του άρθρου, μέσω ενός κατάλληλα προσαρμοσμένου υποσυστήματος. Ιδιαίτερα σημαντική διαδικασία, αφού σε αυτήν στηρίζονται οι μηχανισμοί αναζήτησης, πλοήγησης και διαχείρισης των διαγραμμάτων (profiles) των τελικών χρηστών στο σύστημα.

Ένα άλλο κρίσιμο υποσύστημα ενός συστήματος νέων είναι αυτό της παρουσίασης της εξαγόμενης, από το υποσύστημα συγγραφής άρθρων, πληροφορίας στους τελικούς χρήστες. Πρέπει να αντικατοπτρίζει την οργάνωση που διατηρείται στη βάση δεδομένων και να είναι όσο το δυνατόν πιο εύκολη και βολική στη χρήση για τον τελικό χρήστη. Η ανάπτυξη μηχανισμών παρουσίασης άρθρων, αποτελεσμάτων αναζήτησης, πλοήγησης γενικά μέσα στο σύστημα, παρουσίασης προσωπικού διαγράμματος τελικού χρήστη (profile) και νέων κατ' απαίτηση, κρίνονται αναγκαίες. Ειδικότερα, τα παρακάτω τέθηκαν ως ελάχιστες απαιτήσεις για την κάλυψη των βασικότερων αναγκών του συστήματος:

- Ανάπτυξη μηχανισμού παρουσίασης άρθρου. Πρέπει να περιλαμβάνει όλα τα χαρακτηριστικά που καθορίστηκαν κατά τη συγγραφή του άρθρου και που θα φανούν αναγκαία στον τελικό χρήστη.
- Ανάπτυξη μηχανισμού παρουσίασης αποτελεσμάτων αναζήτησης. Τα στοιχεία που πλαισιώνουν κάθε άρθρο του αποτελέσματος πρέπει να είναι αρκετά ώστε ο τελικός χρήστης να αναγνωρίζει εύκολα και γρήγορα αυτά που τον ενδιαφέρουν.
- Ανάπτυξη μηχανισμού πλοήγησης στο μοντέλο κατηγοριοποίησης που υλοποιεί το σύστημα. Ο τελικός χρήστης πρέπει να έχει τη δυνατότητα να διασχίζει τις κατηγορίες του συστήματος εύκολα και γρήγορα χωρίς να χάνεται.

- Ανάπτυξη μηχανισμού παρουσίασης διαγράμματος τελικού χρήστη και νέων κατ' απαίτηση. Σε αυτή την περίπτωση το σύστημα πρέπει να συμπεριφέρεται σύμφωνα με τις προτιμήσεις του τελικού χρήστη, όπως αυτές έχουν καταγραφεί στο προσωπικό του διάγραμμα. Δηλαδή όλοι οι παραπάνω μηχανισμοί λειτουργούν χρησιμοποιώντας το διάγραμμα ως φίλτρο.
- Όλοι οι παραπάνω μηχανισμοί και γενικότερα όλη η παρουσίαση πρέπει να υποστηρίζει πολλαπλές γλώσσες. Ο χρήστης πρέπει να έχει τη δυνατότητα ανάγνωσης και αναζήτησης άρθρων σε όλες τις γλώσσες που υποστηρίζονται από το σύστημα, χωρίς να χρειάζεται να μεταπηδά σε άλλη γλωσσική έκδοση του συστήματος.
- Τα κύρια χαρακτηριστικά της παρουσίασης του συστήματος, όπως λογότυπα, banners, headers και footers, πρέπει να μεταβάλλονται με εύκολο τρόπο, έτσι ώστε το σύστημα να μπορεί να προσαρμόζεται σε διαφορετικού τύπου περιεχόμενο (π.χ. Αθλητισμός, Τουρισμός, Πολιτισμός).

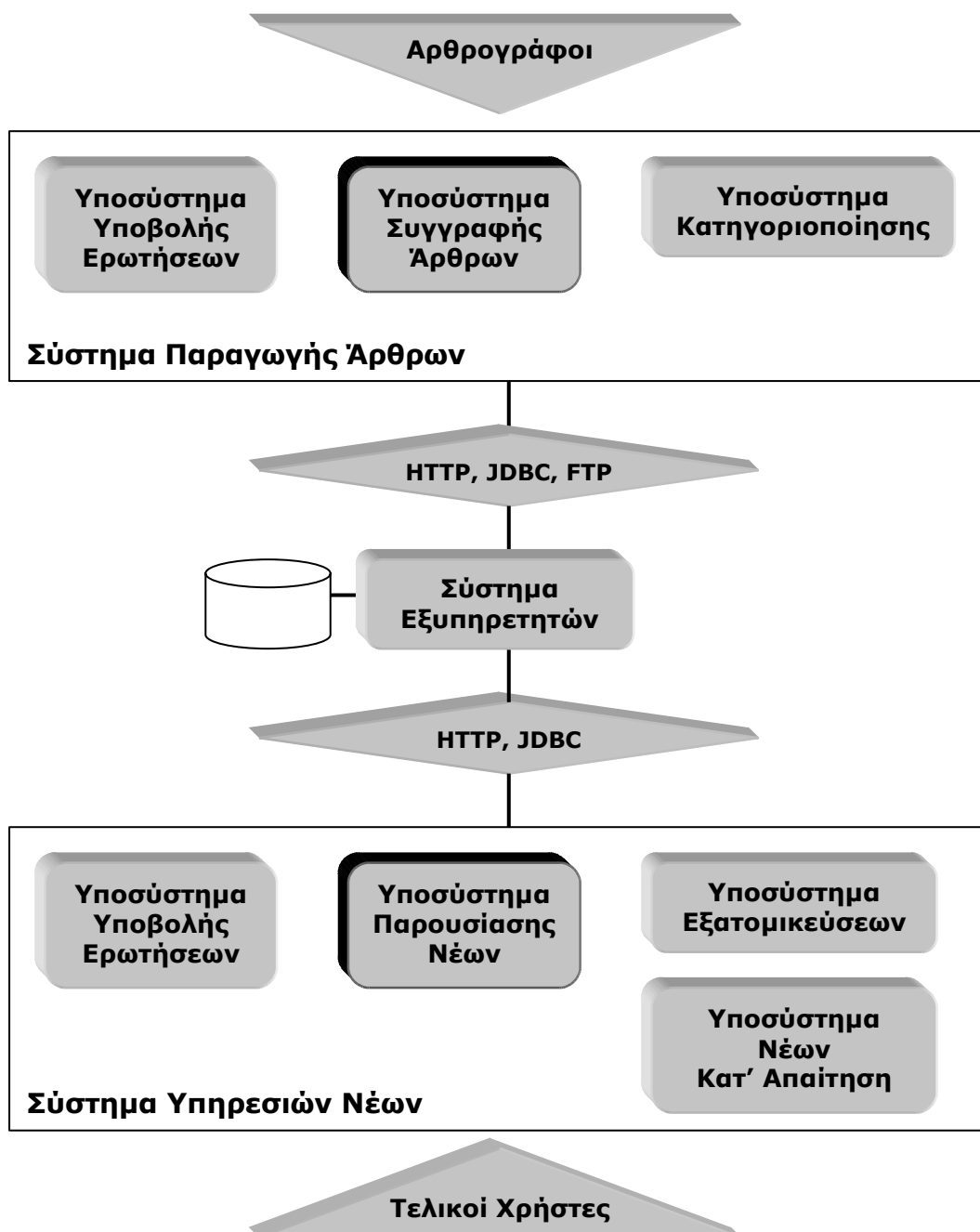
Οι παραπάνω διαπιστώσεις υπογραμμίζουν την αναγκαιότητα του συστήματος νέων **“Hypermedia Custom News System”**, και παρουσιάζουν τη χρησιμότητα και συνεισφορά του. Ιδιαίτερα τον καιρό αυτό όπου η δραστηριότητα που παρατηρείται στον τομέα συστημάτων νέων, καθώς και η μελλοντικά διαγραφόμενη εξάλειψη των έντυπων μέσων ενημέρωσης, καθιστούν αναγκαία και πολύτιμη την απόκτηση της τεχνογνωσίας που απαιτεί ο σχεδιασμός και η υλοποίηση ενός τέτοιου συστήματος.

1.2 Το σύστημα **“Hypermedia Custom News System”**

Το Εργαστήριο Διανεμημένων Πληροφοριακών Συστημάτων και Εφαρμογών Πολυμέσων (MU.S.I.C.) συμμετέχει σε ερευνητικά προγράμματα της Ευρωπαϊκής Ένωσης με κυριότερα πεδία δράσης τον τουρισμό και τον πολιτισμό. Η ανάπτυξη των προγραμμάτων αυτών περιλαμβάνει και συστήματα νέων για την ενημέρωση των χρηστών τους.

Στόχος λοιπόν του συστήματος **“Hypermedia Custom News System”** είναι η δημιουργία ενός ολοκληρωμένου συστήματος νέων που θα ικανοποιεί τις κοινές

απαιτήσεις των προγραμμάτων αυτών και θα υλοποιεί τις τελευταίες καινοτομίες στον τομέα των συστημάτων νέων. Θα απευθύνεται στο σύγχρονο και απαιτητικό χρήστη του διαδικτύου και στους ανθρώπους που εργάζονται για τη δημιουργία των νέων. Επίσης η υποστήριξη πολλαπλών γλωσσών είναι πολύ σημαντική, αφού το κοινό του συστήματος αναμένεται να προέρχεται από πολλές διαφορετικές χώρες.



Σχήμα 1-1 Γενική Αρχιτεκτονική του συστήματος εξατομικευμένων νέων κατ' απαίτηση "Hypermedia Custom News System"

Στο Σχήμα 1-1 απεικονίζεται γραφικά, σε υψηλό επίπεδο αφαίρεσης, η αρχιτεκτονική του συστήματος “**Hypermedia Custom News System**”. Χωρίζεται σε τρία κύρια τμήματα, τα οποία περιγράφονται περιληπτικά παρακάτω.

- Το **Σύστημα Παραγωγής Άρθρων** είναι υπεύθυνο για την παραγωγή των νέων και περιλαμβάνει τρία υποσυστήματα. Μέσω του *Υποσυστήματος Συγγραφής Άρθρων* οι αρθρογράφοι έχουν τη δυνατότητα δημιουργίας, μορφοποίησης και διαχείρισης των άρθρων τους. Το *Υποσύστημα Κατηγοριοποίησης* τους παρέχει τη δυνατότητα κατάταξης των άρθρων και πολυμέσων στη μοντέλο κατηγοριοποίησης του συστήματος, ενώ το *Υποσύστημα Υποβολής Ερωτήσεων* τους βοηθά να εντοπίσουν υπάρχοντα άρθρα ή πολυμέσα του συστήματος για τον εμπλουτισμό του παραγόμενου άρθρου τους.
- Το **Σύστημα Εξυπηρετητών** αποτελείται από ένα σύνολο εξυπηρετητών αναγκαίων για τη σωστή και καλή λειτουργία του συστήματος. Συγκεκριμένα, περιλαμβάνει έναν εξυπηρετητή βάσης δεδομένων, του οποίου ο ρόλος είναι η οργανωμένη αποθήκευση, ταξινόμηση και ανάκτηση των δεδομένων, έναν εξυπηρετητή παγκόσμιου ιστού και έναν εξυπηρετητή μεταφοράς αρχείων. Με αυτούς αλληλεπιδρά το Σύστημα Παραγωγής Άρθρων για τη δημιουργία, αποθήκευση και ανάκτηση των άρθρων που δημιουργούνται από τους Αρθρογράφους, αλλά και το Σύστημα Υπηρεσιών Νέων για την ανάκτηση δεδομένων αναγκαίων για την ικανοποίηση των αιτήσεων των Τελικών Χρηστών. Τα δύο αυτά συστήματα επικοινωνούν με το Σύστημα Εξυπηρετητών μέσω των πρωτοκόλλων HTTP (Hypertext Transfer Protocol), JDBC (Java Database Connectivity) και FTP (File Transfer Protocol).
- Το **Σύστημα Υπηρεσιών Νέων** είναι υπεύθυνο για την παροχή υπηρεσιών νέων στους τελικούς χρήστες και περιλαμβάνει και αυτό τρία υποσυστήματα. Μέσω του *Υποσυστήματος Παρουσίασης Νέων* οι Τελικοί Χρήστες έχουν τη δυνατότητα πλοήγησης στο μοντέλο κατηγοριοποίησης του συστήματος και επισκόπησης των άρθρων που περιέχονται σε αυτό. Το *Υποσύστημα Υποβολής Ερωτήσεων* τους επιτρέπει να αναζητήσουν άρθρα του συστήματος βάσει παραμέτρων που

περιγράφουν τα ενδιαφέροντα τους. Το υποσύστημα αυτό είναι ουσιαστικά κοινό με αυτό του Συστήματος Παραγωγής Άρθρων, αλλά διαφέρει ο τρόπος που εκτελεί τις διεργασίες του. Το *Υποσύστημα Εξατομικεύσεων* είναι υπεύθυνο για τη διαχείριση των διαγραμμάτων που δημιουργούν οι Τελικοί Χρήστες και συνεργάζεται με το Υποσύστημα Παρουσίασης Νέων και το Υποσύστημα Υποβολής Ερωτήσεων για την εφαρμογή των προτιμήσεων που έχουν αποθηκευτεί στα διαγράμματα. Τέλος, το *Υποσύστημα Νέων Κατ' Απαίτηση* είναι υπεύθυνο για την ειδοποίηση του τελικού χρήστη, μέσω ειδικής εφαρμογής που τρέχει στο σταθμό εργασίας του, για την έκδοση νέων άρθρων που ταιριάζουν με το διάγραμμά του.

1.3 Στόχοι της διπλωματικής εργασίας

Στόχος της παρούσας διατριβής είναι η δημιουργία ενός ολοκληρωμένου περιβάλλοντος συγγραφής και διαχείρισης εγγράφων, καθώς επίσης και η υλοποίηση σύγχρονων υπηρεσιών νέων στους τελικούς χρήστες του συστήματος νέων **“Hypermedia Custom News System”** μέσω των τεχνολογιών εξατομικευμένων νέων και νέων κατ' απαίτηση.

Το σύστημα παραγωγής άρθρων, πέρα από την εισαγωγή βασικών χαρακτηριστικών, παρέχει τη δυνατότητα μορφοποίησης εγγράφων με στοιχεία παρμένα από το μοντέλο αναπαράστασης της HTML (ενότητες κειμένου, διασυνδέσεις, συσχετίσεις, αντικείμενα εντός κειμένου). Επίσης προσφέρει ολοκληρωμένη διαχείριση των εγγράφων και των πολυμέσων του συστήματος (εισαγωγή, τροποποίηση, μετάφραση, διαγραφή). Το υποσύστημα παρουσίασης νέων, πέρα από τις βασικές λειτουργίες (παρουσίαση άρθρων, πλοήγηση, αναζήτηση), φροντίζει για την παρουσίαση των εξατομικευμένων νέων και των νέων κατ' απαίτηση στους τελικούς χρήστες, έτσι όπως αυτά προκύπτουν από τις αποθηκευμένες προτιμήσεις αυτών στα προσωπικά τους διαγράμματα.

Φροντίζοντας αρχικά να ικανοποιούνται οι βασικές προδιαγραφές του συστήματος, ιδιαίτερη έμφαση δόθηκε στην υποστήριξη αυθαίρετου αριθμού γλωσσών. Το σύστημα παρέχει τη δυνατότητα αύξησης του αριθμού γλωσσών που υποστηρίζει

χωρίς να χρειάζεται καμία αλλαγή σε κάποιο από τα υποσυστήματά του, καθώς ο αριθμός αυτός υπολογίζεται δυναμικά. Μόλις προστεθεί μία επιπλέον γλώσσα, οι χρήστες του συστήματος μπορούν απευθείας να εισάγουν ή να επισκοπήσουν άρθρα στη γλώσσα αυτή.

Συγκεκριμένα η υλοποίηση του υποσυστήματος συγγραφής άρθρων χωρίστηκε στις εξής ενότητες:

- Καθορισμός μοντέλου αναπαράστασης εγγράφου.
- Σχεδιασμός της βάσης δεδομένων εγγράφων σύμφωνα με το παραπάνω μοντέλο.
- Υλοποίηση μηχανισμού μορφοποίησης κειμένου.
- Καθορισμός μοντέλου αναπαράστασης μέσου.
- Σχεδιασμός της βάσης δεδομένων πολυμέσων σύμφωνα με το παραπάνω μοντέλο.
- Υλοποίηση μηχανισμού εισαγωγής νέων πολυμέσων στο σύστημα.
- Υλοποίηση μηχανισμού εισαγωγής πολυμέσων μέσα στο κείμενο και καθορισμού των χαρακτηριστικών τους.
- Προσαρμογή υποσυστήματος υποβολής ερωτήσεων για αναζήτηση άρθρων και πολυμέσων στο ήδη υπάρχον αρχείο του συστήματος (βλ. και [Καρασ98]).
- Μηχανισμός παρουσίασης των αποτελεσμάτων του παραπάνω υποσυστήματος μέσα στο υποσύστημα συγγραφής άρθρων.
- Υλοποίηση μηχανισμού συσχέτισης άρθρων.
- Προσαρμογή υποσυστήματος κατηγοριοποίησης και χαρακτηρισμού άρθρου και μέσου (βλ. και [Καρασ98]).
- Υλοποίηση μηχανισμού συσχέτισης μεταφραζόμενων άρθρων.

- Υλοποίηση μηχανισμού αποθήκευσης, ανάκτησης και διαγραφής άρθρου από τη βάση δεδομένων.
- Υλοποίηση μηχανισμού ενημέρωσης στατιστικής βάσης δεδομένων κατά την εισαγωγή, ενημέρωση ή διαγραφή άρθρου (βλ. και [Καρασ98]).
- Υλοποίηση της διαδικασίας συγγραφής άρθρων με τρόπο τέτοιο ώστε αυτή να εκτελείται ταυτόχρονα από πολλούς αρθρογράφους και από οποιοδήποτε σημείο του διαδικτύου.
- Ελαχιστοποίηση των διαφορών μεταξύ της παραπάνω διαδικασίας και της παραδοσιακής που ήδη χρησιμοποιούν οι αρθρογράφοι.
- Καθορισμός και υλοποίηση πολιτικής ασφαλείας για τους αρθρογράφους.
- Υλοποίηση μηχανισμού υποστήριξης αυθαίρετου αριθμού γλωσσών.

Η υλοποίηση του υποσυστήματος παρουσίασης νέων χωρίστηκε στις εξής ενότητες:

- Υλοποίηση μηχανισμού ανάκτησης και παρουσίασης άρθρου.
- Υλοποίηση μηχανισμού πλοήγησης του τελικού χρήστη στο μοντέλο κατηγοριοποίησης που υλοποιεί το σύστημα.
- Υλοποίηση μηχανισμού παρουσίασης αποτελεσμάτων αναζήτησης.
- Υλοποίηση μηχανισμού παρουσίασης προσωπικού διαγράμματος τελικού χρήστη.
- Υλοποίηση μηχανισμού παρουσίασης νέων κατ' απαίτηση.
- Υλοποίηση μηχανισμού υποστήριξη αυθαίρετου αριθμού γλωσσών.

1.4 Δομή της διπλωματικής εργασίας

Στο κεφάλαιο 2 της διατριβής αναφέρονται αναλυτικά οι βασικές αρχές των σημερινών συστημάτων νέων και η κατηγοριοποίηση που γίνεται σύμφωνα με αυτές.

Επίσης, γίνεται μια επισκόπηση των υπαρχόντων συστημάτων νέων στο διαδίκτυο και των τεχνολογιών που υποστηρίζουν, τόσο για το διεθνή όσο και για τον ελληνικό χώρο.

Το κεφάλαιο 3 αποτελεί αναλυτική περιγραφή της αρχιτεκτονικής του συστήματος. Περιγράφει τους μηχανισμούς και τα τμήματα που περιλαμβάνει κάθε υποσύστημα, τη λειτουργικότητα τους και τον τρόπο αλληλεπίδρασης τους.

Στο κεφάλαιο 4 αναλύεται η ανάπτυξη του σχεσιακού μοντέλου βάσης δεδομένων που χρησιμοποιείται για την υποστήριξη του συστήματος και υλοποιείται από κατάλληλο Σύστημα Διαχείρισης Βάσεων Δεδομένων.

Στο κεφάλαιο 5 περιγράφεται το Υποσύστημα Συγγραφής Άρθρων που αποτελεί και το ένα από τα δύο μεγάλα τμήματα της παρούσας διατριβής. Περιγράφονται όλοι οι μηχανισμοί που αυτό περιέχει, καθώς και η λειτουργικότητα αυτών.

Το κεφάλαιο 6 αναφέρεται στο άλλο μεγάλο τμήμα της διατριβής, το Υποσύστημα Παρουσίασης Νέων. Αναλύονται όλοι οι μηχανισμοί που αναπτύχθηκαν για την παρουσίαση των άρθρων, για την πλοήγηση των τελικών χρηστών στο σύστημα, για την παρουσίαση αποτελεσμάτων αναζήτησης, για την παρουσίαση των προσωπικών διαγραμμάτων και των νέων κατ' απαίτηση.

Στο κεφάλαιο 7 περιγράφεται ο μηχανισμός υποστήριξης πολλαπλών γλωσσών ο οποίος προστέθηκε τόσο στο εργαλείο συγγραφείς άρθρων όσο και στους μηχανισμούς παρουσίασης νέων και αποτελεί ίσως τη σημαντικότερη καινοτομία του συστήματος.

Στο κεφάλαιο 8 καταγράφεται η τεχνολογία που χρησιμοποιήθηκε για την υλοποίηση των τμημάτων του συστήματος.

Το κεφάλαιο 9 αποτελεί σύντομη ανακεφαλαίωση της παρούσας διπλωματικής και αναφέρεται στις μελλοντικές επεκτάσεις που μπορούν να ενσωματωθούν στο σύστημα.

2 ΕΠΙΣΚΟΠΗΣΗ ΥΠΑΡΧΟΝΤΩΝ ΣΥΣΤΗΜΑΤΩΝ ΝΕΩΝ

Στο κεφάλαιο αυτό αναφέρονται συνοπτικά οι βασικές αρχές που διέπουν τα σημερινά συστήματα νέων στο διαδίκτυο και οι βασικές κατηγορίες στις οποίες διαχωρίζονται. Στη συνέχεια περιγράφονται τα πιο δημοφιλή από τα υπάρχοντα συστήματα νέων, τα οποία επηρέασαν την εκπόνηση της παρούσας διατριβής. Συγκεκριμένα καταγράφονται οι δυνατότητες που παρέχουν στους χρήστες τους και σε ποια κατηγορία κατατάσσονται σύμφωνα με αυτές. Τα συστήματα αυτά προέρχονται τόσο από το διεθνές όσο και από τον ελληνικό χώρο.

2.1 Εισαγωγή στα Συστήματα Νέων

Τα συστήματα νέων στο διαδίκτυο είναι μια νέα μορφή ενημέρωσης, που σκοπό έχει την αμεσότερη και πιο ολοκληρωμένη ενημέρωση του κοινού από οποιαδήποτε άλλη μορφή πληροφόρησης (π.χ. τηλεόραση, ραδιόφωνο, εφημερίδες, περιοδικά). Η πρόσβαση στα συστήματα αυτά είναι δυνατή από οποιοδήποτε σημείο του κόσμου οποιαδήποτε στιγμή, διευρύνοντας έτσι το κοινό των παραδοσιακών μέσων και καθιστώντας δυνατή την ενημέρωση χρηστών απομακρυσμένων από την πηγή παραγωγής των ειδήσεων, χωρίς κανένα χρονικό περιορισμό. Επίσης η δυνατότητα οργάνωσης της πληροφορίας στα συστήματα αυτά επέτρεψε τη δημιουργία ειδησεογραφικών αρχείων στα οποία ο χρήστης μπορεί να ανατρέξει οποιαδήποτε στιγμή.

Ικανοποιώντας αρχικά τις παραπάνω αρχές, τα συστήματα νέων συνεχίζουν να ενσωματώνουν νέες τεχνολογίες και δυνατότητες, αφού με την πάροδο του χρόνου σημειώθηκε σημαντική πρόοδος και στον τομέα των δικτύων υπολογιστών και των τηλεπικοινωνιών. Μηχανισμοί αναζήτησης και κατάταξης των αποτελεσμάτων εφαρμόζονται στα ειδησεογραφικά αρχεία έτσι ώστε ο χρήστης να εντοπίζει την πληροφορία που τον ενδιαφέρει πολύ πιο γρήγορα. Τα άρθρα εμπλουτίζονται με πολυμέσα (π.χ. ήχος, video) και συσχετίζονται με άλλα άρθρα με σκοπό την

πληρέστερη κάλυψη μιας είδησης. Ο χρήστης έχει τη δυνατότητα αλληλεπίδρασης με το σύστημα και καθορισμού των προσωπικών του προτιμήσεων (personalization), με σκοπό τη μείωση του χρόνου πλοήγησης και αναζήτησης πληροφορίας μέσα στο σύστημα. Δυνατή είναι επίσης και η κατ' απαίτηση (on demand) ενημέρωση των χρηστών με τη χρήση μηχανισμών ειδοποίησης και αποστολής δεδομένων που έχουν αναπτυχθεί γι' αυτό το σκοπό.

Από τη στιγμή που οι τελικοί χρήστες αλληλεπιδρούν με το σύστημα, νέες δυνατότητες αποκτούν και οι εταιρείες που παρέχουν τα νέα. Στατιστικές μελέτες και δειγματοληψίες πάνω στις προτιμήσεις των χρηστών είναι εφικτές και μπορούν να αποδειχθούν ιδιαίτερα κερδοφόρες κυρίως για διαφημιστικούς σκοπούς. Επίσης δυνατή είναι και η εφαρμογή διαδικασιών χρέωσης των τελικών χρηστών για την πρόσβαση στο σύστημα.

Όπως βλέπουμε από τα παραπάνω, τα συστήματα νέων αυξάνουν συνεχώς τις παρεχόμενες υπηρεσίες τους. Και θα μπορούσαν να το κάνουν με γρηγορότερο ρυθμό αν δεν υπήρχε ο περιορισμός ταχύτητας μετάδοσης δεδομένων πάνω από το διαδίκτυο. Ένας άλλος περιορισμός είναι ότι τα συστήματα νέων δεν αποτελούν ακόμα σίγουρη επένδυση, αφού η απαιτούμενη τεχνογνωσία για την ανάπτυξη τους στοιχίζει ακριβά, αλλά και το ευρύ κοινό δεν έχει εξοικειωθεί ακόμα με την ιδέα της ηλεκτρονικής ενημέρωσης. Τέλος, περιοριστική είναι και η έλλειψη προστασίας των δικαιωμάτων αναπαραγωγής των νέων και γενικότερα η ελλιπής ασφάλεια στο διαδίκτυο.

2.2 Κατηγορίες Συστημάτων Νέων

Σύμφωνα με τα χαρακτηριστικά που αναφέρθηκαν παραπάνω τα σημερινά συστήματα νέων χωρίζονται κυρίως σε τρεις κατηγορίες, οι οποίες προκύπτουν από τις δυνατότητες που παρέχουν στους αρθρογράφους και στους τελικούς χρήστες. Τα συμβατικά συστήματα νέων, τα εξατομικευμένα συστήματα νέων και τα εξατομικευμένα συστήματα νέων κατ' απαίτηση είναι οι τρεις αυτές κατηγορίες. Η κάθε μία από αυτές περιγράφεται λεπτομερέστερα παρακάτω.

2.2.1 Συμβατικά Συστήματα Νέων

Τα συστήματα που κατατάσσονται σ' αυτή την κατηγορία διαθέτουν ειδησεογραφικό αρχείο οργανωμένο στη δευτερεύουσα μνήμη και όχι σε βάση δεδομένων. Οι δυνατότητες των εργαλείων συγγραφής είναι αρκετά περιορισμένες και δεν υποστηρίζεται η ταυτόχρονη πρόσβαση από πολλούς αρθρογράφους. Το μοντέλο αναπαράστασης εγγράφου περιέχει μόνο τα κυριότερα χαρακτηριστικά ενός άρθρου, συνήθως σε μορφή απλού κειμένου. Η χωρική και χρονική κατηγοριοποίηση που ακολουθεί το σύστημα υλοποιείται και αυτή στη δευτερεύουσα μνήμη σε μορφή φακέλων (directories). Η παρουσίαση των άρθρων και της δομής κατηγοριοποίησης τους πραγματοποιείται μέσω στατικών HTML σελίδων, γεγονός που καθιστά δυνατή την εφαρμογή έτοιμου μηχανισμού αναζήτησης, ο οποίος όμως κοστίζει ακριβά και πιθανότατα να μην ικανοποιεί πλήρως τις απαιτήσεις των χρηστών του συστήματος. Ο τελικός χρήστης για να εντοπίσει το άρθρο που επιθυμεί πρέπει να διασχίσει τη δομή κατηγοριοποίησης του συστήματος και να το επιλέξει από μια λίστα υπαρχόντων.

Τα συστήματα αυτά αποτελούν συνήθως μεταφορά των νέων ενός άλλου μέσου ενημέρωσης (π.χ. εφημερίδας) σε ηλεκτρονική μορφή. Ο ρυθμός ανανέωσης των νέων είναι αργός, ενώ δεν παρέχονται εξατομικευμένες υπηρεσίες στους τελικούς χρήστες. Τέλος οι περιορισμένες δυνατότητες που παρέχουν τα συστήματα αυτά στον τελικό χρήστη καθιστούν την εμπορική εκμετάλλευση σχεδόν αδύνατη.

2.2.2 Εξατομικευμένα Συστήματα Νέων

Σ' αυτή την κατηγορία τα συστήματα μπορούν να χαρακτηριστούν ως σύγχρονα αφού υλοποιούν τις περισσότερες από τις τεχνολογίες αιχμής σήμερα. Όχι μόνο έγγραφα αλλά και πολυμέσα οργανώνονται, αποθηκεύονται και ταξινομούνται σε εξυπηρετητή βάσης δεδομένων. Τα εργαλεία συγγραφής έχουν τη δυνατότητα εισαγωγής και διαχείρισης πολυμέσων μέσα στα άρθρα, εμπλουτίζοντας κατά πολύ το μοντέλο αναπαράστασης εγγράφου. Διαθέτουν μηχανισμούς ταξινόμησης και αναζήτησης των νέων, παρέχοντας στους αρθρογράφους τη δυνατότητα κατηγοριοποίησης, ανεύρεσης και συσχέτισης των άρθρων μεταξύ τους. Επίσης το

μοντέλο κατηγοριοποίησης άρθρων και πολυμέσων είναι μεταβλητό επιτρέποντας έτσι την επέκταση του εύρους της πληροφορίας που καλύπτει το σύστημα.

Την παρουσίαση του συστήματος αναλαμβάνουν να εκτελέσουν δυναμικά προς τον τελικό χρήστη κατάλληλοι μηχανισμοί παρουσίασης. Πέρα από τη δυνατότητα αναζήτησης και κατάταξης εγγράφων στον τελικό χρήστη παρέχεται μηχανισμός αποθήκευσης των προσωπικών του προτιμήσεων (personalization), μειώνοντας έτσι κατά πολύ τον απαιτούμενο χρόνο πλοήγησης του στο σύστημα για την ανεύρεση της επιθυμητής πληροφορίας. Ο τελικός χρήστης όταν μπαίνει στο σύστημα βλέπει τα αποτελέσματα του φιλτραρίσματος που έγινε πάνω στα δεδομένα του συστήματος σύμφωνα με τις δικές του προτιμήσεις.

Τα παραπάνω συστήματα, έχουν τη δυνατότητα ταυτόχρονης πρόσβασης παραπάνω του ενός αρθρογράφου, ακόμα και παραπάνω του ενός ειδησεογραφικού οργανισμού, αυξάνοντας έτσι το ρυθμό ανανέωσης των ειδήσεων δραστικά σε σχέση με τα συμβατικά συστήματα νέων. Τέλος λόγω των αυξημένων δυνατοτήτων που παρέχουν στους τελικούς χρήστες υπάρχουν περιθώρια και για εμπορική εκμετάλλευση.

2.2.3 Εξατομικευμένα Συστήματα Νέων Κατ' απαίτηση

Τα εξατομικευμένα συστήματα νέων μπορεί να έχουν υψηλό ρυθμό ανανέωσης των νέων, αλλά δεν παραδίδουν τη νέα πληροφορία στους τελικούς χρήστες μόλις αυτή είναι διαθέσιμη. Αυτή την ανάγκη έρχεται να καλύψει η τεχνολογία νέων κατ' απαίτηση (news on demand), που αποτελεί σήμερα αιχμή στην εξέλιξη των συστημάτων νέων. Ο τελικός χρήστης, αφού ορίσει τις προσωπικές του προτιμήσεις, μπορεί να απαιτήσει από το σύστημα να λαμβάνει ή να ειδοποιείται για τα εισερχόμενα νέα που καλύπτουν τα ενδιαφέροντα του μέσω κάποιου κατάλληλου μηχανισμού. Με αυτόν τον τρόπο ο τελικός χρήστης βρίσκεται πολύ κοντά, χρονικά, στην παραγωγή των νέων. Επιπλέον, στην περίπτωση που ο ρυθμός ανανέωσης του συστήματος είναι αρκετά γρήγορος, μπορούμε να πούμε ότι παρέχεται ενημέρωση σχεδόν σε πραγματικό χρόνο.

Οι μηχανισμοί ειδοποίησης που παρέχονται στους τελικούς χρήστες είναι κυρίως δύο ειδών. Είτε μέσω ηλεκτρονικού ταχυδρομείου (e-mail), όπου το σύστημα στέλνει στο

χρήστη ολόκληρο το άρθρο ή μια περίληψη μαζί με τη διεύθυνση όπου μπορεί να βρει το πλήρες άρθρο, είτε με τη χρήση ειδικών εφαρμογών ειδοποίησης που ο τελικός χρήστης πρέπει να εγκαταστήσει στο μηχανήμά του. Ο τελευταίος αυτός μηχανισμός παρέχει αρκετή ευελιξία στον χρήστη, είναι πλήρως παραμετρικός και είναι ιδιαίτερα απλός και βολικός στη χρήση.

Τέτοιου είδους συστήματα προορίζονται κυρίως για εμπορική εκμετάλλευση διότι έχουν τη δυνατότητα παροχής εξειδικευμένων υπηρεσιών ακόμα και για κρίσιμες εφαρμογές (π.χ. χρηματιστήριο, καιρός κλπ). Τα συστήματα αυτά έχουν ήδη αρχίσει να κάνουν την εμφάνισή τους στο διαδίκτυο και αναμένεται να φανούν αναγκαία από όλο και περισσότερους χρήστες τα επόμενα χρόνια.

2.3 Υπάρχοντα Συστήματα Νέων

Η σχεδίαση και υλοποίηση του συστήματος “Hypermedia Custom News System” βασίστηκε αρχικά στην εμπειρία συναδέλφων, οι οποίοι είχαν συμμετάσχει σε ευρωπαϊκά ερευνητικά προγράμματα σχετικά με το αντικείμενο, και ειδικότερα στο πρόγραμμα **HyNoDe** (Hypermedia News On Demand). Το πρόγραμμα προέβλεπε την κατασκευή συστήματος εξατομικευμένων νέων κατ’ απαίτηση, στα πλαίσια του οποίου το MU.S.I.C. είχε αναλάβει το σχεδιασμό και την ανάπτυξη του εξυπηρετητή πολυμέσων και την υλοποίηση των μηχανισμών επικοινωνίας με αυτόν. Τα παραδοτέα του HyNoDe ήταν ιδιαίτερα πολύτιμο υλικό για την αρχική σχεδίαση και μοντελοποίηση του συστήματος, από τη στιγμή που περιέγραφαν τις προδιαγραφές όλων των τμημάτων του προγράμματος. Αλλά και στη συνέχεια, κατά την πιλοτική φάση λειτουργίας του προγράμματος, πολλές ιδέες πάρθηκαν για τον τρόπο παρουσίασης του συστήματος στον τελικό χρήστη.

Δύο άλλα μεγάλα διεθνή και πρωτοποριακά συστήματα νέων που επηρέασαν την πορεία ανάπτυξης της παρούσας διατριβής ήταν το “CNN Interactive” και το “MSNBC”. Και τα δύο αναπτύχθηκαν στις Ηνωμένες Πολιτείες και καλύπτουν ειδησεογραφικά όλο τον κόσμο.

Το “**CNN Interactive**” αποτελεί τη δικτυακή μορφή ενημέρωσης του γνωστού αμερικανικού τηλεοπτικού καναλιού. Πέρα από το κυρίως σύστημα περιέχει και

τέσσερα υποσυστήματα, τα τρία από τα οποία αφορούν πολιτικές (CNN/TIME Allpolitics), οικονομικές (CNN finance) και αθλητικές (CNN Sports Illustrated) ειδήσεις, ενώ το τέταρτο έχει να κάνει με τη διαχείριση των προσωπικών προτιμήσεων των τελικών χρηστών και ονομάζεται “CNN Custom News”. Αναπτύχθηκε σε συνεργασία με την Oracle και υλοποιεί την τεχνολογία εξατομικευμένων νέων. Επίσης υποστηρίζει και την τεχνολογία νέων κατ’ απαίτηση υλοποιώντας όμως μόνο το μηχανισμό παράδοσης νέων μέσω ηλεκτρονικού ταχυδρομείου.



Σχήμα 2-1 Σελίδα παρουσίασης εξατομικευμένων νέων του συστήματος “CNN Custom News

Το “**MSNBC**” προέκυψε από τη συνεργασία της Microsoft και του αμερικανικού δικτύου NBC και είναι και αυτό εξατομικευμένο σύστημα νέων κατ’ απαίτηση. Πέρα από το μηχανισμό παράδοσης νέων μέσω ηλεκτρονικού ταχυδρομείου, το σύστημα παρέχει και ειδικό μηχανισμό ειδοποίησης του τελικού χρήστη, ο οποίος εγκαθίσταται στο σταθμό εργασίας του και των ενημερώνει ανά τακτά χρονικά διαστήματα.



Σχήμα 2-2 Ο μηχανισμός ειδοποίησης τελικού χρήστη του συστήματος “MSNBC”

Στον ελληνικό χώρο το μόνο σύστημα που ξεφεύγει από τα συμβατικά είναι αυτό της Ναυτεμπορικής, το οποίο υποστηρίζει εξατομικευμένα νέα και περιλαμβάνει μηχανισμό παραλαβής νέων μέσω ηλεκτρονικού ταχυδρομείου. Επίσης το σύστημα αθλητικής ενημέρωσης “Sportline” διαθέτει πολύ καλή παρουσίαση νέων, ενσωματώνοντας πολυμέσα μέσα στο κείμενο των άρθρων. Από τα συμβατικά συστήματα νέων ξεχωρίζουν το “Μακεδονικό Πρακτορείο Ειδήσεων” και το “Flash 9.61 Interactive”. Και τα δύο παρέχουν πλήρη ενημέρωση και καλαίσθητη παρουσίαση.

Από τα συστήματα που αναφέρθηκαν παραπάνω, κανένα δεν έχει τη δυνατότητα να υποστηρίξει αυθαίρετο αριθμό γλωσσών. Τα διεθνή χρησιμοποιούν μόνο αγγλικά, ενώ τα ελληνικά έχουν κυρίως δύο διαφορετικές εκδόσεις σε αγγλικά και ελληνικά.

2.4 Ανακεφαλαίωση

Όπως φάνηκε σε αυτό το κεφάλαιο τα συστήματα νέων στο διαδίκτυο επικεντρώνουν το ενδιαφέρον, τόσο των τελικών χρηστών όσο και των ειδησεογραφικών οργανισμών, όλο και περισσότερο με την πάροδο του χρόνου. Οι προοπτικές εξέλιξης τους είναι επίσης πολύ καλές, καθώς υπάρχει αρκετό επιστημονικό και αγοραστικό ενδιαφέρον προς αυτή την κατεύθυνση. Οι διευθύνσεις των συστημάτων που αναφέρθηκαν εδώ, μαζί με κάποια άλλα που εξετάστηκαν συμπληρωματικά, μπορούν να βρεθούν στο παράρτημα Α της παρούσας διατριβής.

Στο επόμενο κεφάλαιο θα εξεταστεί η αρχιτεκτονική του συστήματος “Hypermedia Custom News System” με μεγαλύτερη λεπτομέρεια, επικεντρώνοντας κυρίως στο εργαλείο συγγραφής άρθρων και στους μηχανισμούς παρουσίασης νέων.

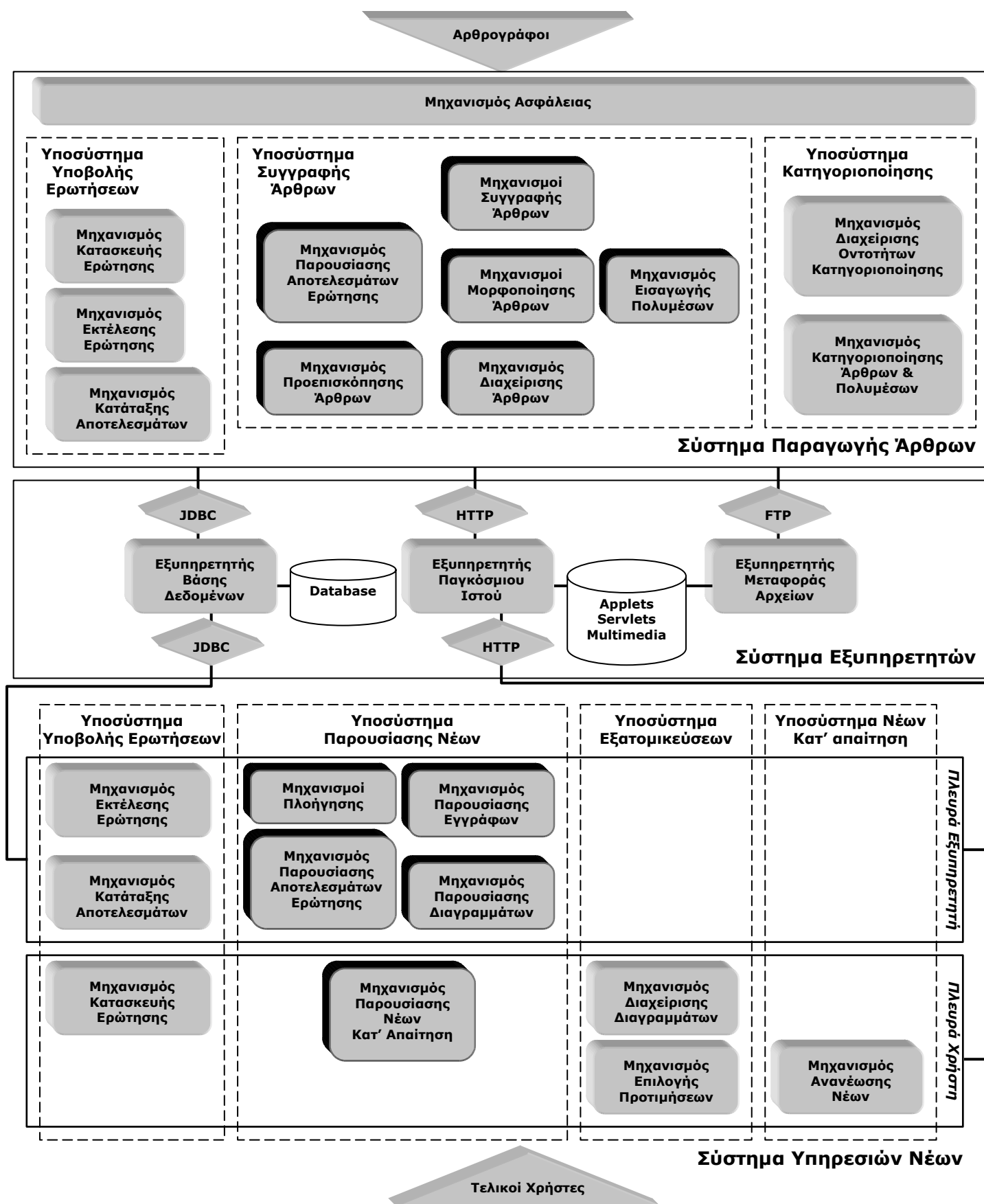
3 ΕΠΙΣΚΟΠΗΣΗ ΑΡΧΙΤΕΚΤΟΝΙΚΗΣ ΣΥΣΤΗΜΑΤΟΣ

Στο παρόν κεφάλαιο περιγράφεται η πλήρης αρχιτεκτονική του συστήματος “Hypermedia Custom News System” παρουσιάζοντας αναλυτικότερα το Σχήμα 1-1 του πρώτου κεφαλαίου. Κάθε υποσύστημα και επιμέρους τμήμα του συστήματος μοντελοποιήθηκε με σαφή δομή και συμπεριφορά, ώστε να καλύπτει με τον καλύτερο δυνατό τρόπο τις βασικές ανάγκες ενός εξατομικευμένου συστήματος νέων κατ’ απαίτηση, έτσι όπως αυτές αναφέρθηκαν στην παράγραφο 1.1.

Κύριο μέλημα της διαδικασίας σχεδιασμού ήταν η αρχιτεκτονική να χαρακτηρίζεται ως γενική, ώστε να υπάρχει η δυνατότητα εύκολης προσαρμογής της σε συστήματα με διαφορετικού είδους περιεχόμενο, αλλά και ανοικτή, ώστε να μπορεί να ενσωματώσει εξειδικευμένες επεκτάσεις, επιθυμητές ανά περίπτωση. Εξίσου σημαντικό χαρακτηριστικό της αρχιτεκτονικής είναι και το γεγονός ότι το σύστημα έχει τη δυνατότητα υποστήριξης αυθαίρετου αριθμού γλωσσών, κάτι που μέχρι σήμερα τουλάχιστον δεν έχει παρατηρηθεί σε άλλο σύστημα νέων.

Στο Σχήμα 3-1 της επόμενης σελίδας βλέπουμε ολοκληρωμένα την αρχιτεκτονική του συστήματος. Παραθέτονται γραφικά όλα τα υποσυστήματα και τμήματα του συστήματος που υλοποιήθηκαν, τόσο από την παρούσα διατριβή όσο και από άλλες (βλ. [Καρασ98] και [Σκον98]). Επίσης διακρίνονται οι χρήστες του συστήματος και πώς αυτοί αλληλεπιδρούν με το σύστημα, καθώς επίσης και τα πρωτόκολλα επικοινωνίας που χρησιμοποιούνται για τη σύνδεση των κυριοτέρων τμημάτων.

Στις παραγράφους που ακολουθούν περιγράφονται αναλυτικά όλα τα υποσυστήματα, και τα τμήματα αυτών, που διακρίνονται στο Σχήμα 3-1, καθώς επίσης και ο τρόπος που αλληλεπιδρούν μεταξύ τους, δίνοντας μεγαλύτερη έμφαση σε αυτά που υλοποιήθηκαν στα πλαίσια της παρούσας διατριβής.



Σχήμα 3-1 Αρχιτεκτονική του συστήματος νέων "Hypermedia Custom News System"

3.1 Σύστημα Παραγωγής Άρθρων

Το σύστημα παραγωγής άρθρων είναι υπεύθυνο για την παραγωγή και διαχείριση των άρθρων και αποτελείται από τρία κύρια υποσυστήματα και από ένα μηχανισμό ασφαλείας. Όλα τα τμήματα του τρέχουν στο σταθμό εργασίας του αρθρογράφου, ο οποίος έχει τη δυνατότητα πρόσβασης στο σύστημα μέσω του διαδικτύου. Στη συνέχεια περιγράφονται οι ανάγκες που καλύπτει κάθε υποσύστημα και μια σύντομη αναφορά στη λειτουργικότητα των τμημάτων τους.

3.1.1 Μηχανισμός Ασφαλείας

Ο Μηχανισμός Ασφαλείας επιτρέπει στον αρθρογράφο να κάνει χρήση του συστήματος αφού πρώτα επαληθεύσει την ταυτότητα του. Κάθε αρθρογράφος εφοδιάζεται με ένα όνομα και έναν κωδικό τα οποία του ζητούνται στην αρχή κάθε συνόδου του με το σύστημα. Ο μηχανισμός αυτός επικοινωνεί με όλα τα υποσυστήματα του συστήματος παραγωγής άρθρων για την κοινοποίηση της ταυτότητας του συνδεδεμένου χρήστη, έτσι ώστε να μπορεί το σύστημα να εγγυηθεί την ακεραιότητα της συγγραφικής εργασίας των αρθρογράφων.

3.1.2 Υποσύστημα Συγγραφής Άρθρων

Το Υποσύστημα Συγγραφής Άρθρων αποτελείται από έξι κύριους μηχανισμούς, όπου ο καθένας καλύπτει συγκεκριμένες ανάγκες στη δημιουργία ενός άρθρου. Αρχικά ο *Μηχανισμός Συγγραφής Άρθρων* επιτρέπει στον αρθρογράφο να εισάγει τα βασικά χαρακτηριστικά ενός άρθρου (τίτλος, υπότιτλος, κυρίως κείμενο) και να τονίσει τα σημεία του κειμένου που επιθυμεί κάνοντας χρήση βασικών μεθόδων επισήμανσης (έντονα, πλάγια, υπογραμμισμένα γράμματα). Ο *Μηχανισμός Μορφοποίησης Άρθρων* παρέχει τη δυνατότητα ενσωμάτωσης πολυμέσων στο κείμενο του άρθρου, εισαγωγής συνδέσμων (hyperlinks) και συσχετισμού του με άλλα άρθρα. Τα πολυμέσα και τα άρθρα αυτά ο αρθρογράφος έχει τη δυνατότητα να τα εντοπίσει μέσω του Υποσυστήματος Υποβολής Ερωτήσεων και να τα επιλέξει μέσω του *Μηχανισμού Παρουσίασης Αποτελεσμάτων Ερώτησης*. Αν ο αρθρογράφος επιθυμεί να εισάγει ένα νέο μέσο στο σύστημα και να το ενσωματώσει στο άρθρο που δημιουργεί, μπορεί να το πράξει κάνοντας χρήση του *Μηχανισμού Εισαγωγής Πολυμέσων*.

Ο *Μηχανισμός Προεπισκόπησης Άρθρων και Πολυμέσων* παρέχει τη δυνατότητα προβολής των πολυμέσων στον αρθρογράφο και τη δυνατότητα προεπισκόπησης του άρθρου που δημιουργείται πριν αυτό εκδοθεί και αποθηκευτεί στη βάση δεδομένων. Τέλος ο *Μηχανισμός Διαχείρισης Άρθρων* είναι υπεύθυνος για τη μεταφορά και αποθήκευση ενός νέου άρθρου από το σύστημα παραγωγής άρθρων στη βάση δεδομένων, όπως επίσης και για την ανάκτηση ενός υπάρχοντος άρθρου αν ο αρθρογράφος επιθυμεί να το τροποποιήσει.

3.1.3 Υποσύστημα Κατηγοριοποίησης

Το Υποσύστημα Κατηγοριοποίησης (βλ. και [Καρασ98]) είναι ενσωματωμένο στο Υποσύστημα Συγγραφής Άρθρων και είναι υπεύθυνο για την διαχείριση του μοντέλου κατηγοριοποίησης που ακολουθεί το σύστημα καθώς και για την κατάταξη σύμφωνα με αυτό των άρθρων και των πολυμέσων που εισάγονται στο σύστημα. Τα επίπεδα κατηγοριοποίησης του συστήματος είναι δύο, κατηγορίες και υποκατηγορίες, ενώ διατηρείται και ιεραρχία λέξεων – κλειδιών αυθαίρετου αριθμού επιπέδων, που χρησιμοποιούνται για τον χαρακτηρισμό άρθρων και πολυμέσων.

Ο *Μηχανισμός Διαχείρισης Οντοτήτων Κατηγοριοποίησης* επιτρέπει στον αρθρογράφο την εισαγωγή, μετατροπή ή διαγραφή κατηγοριών, υποκατηγοριών ή λέξεων – κλειδιών, έτσι ώστε να μπορεί αυτός να κατατάσσει τα άρθρα και τα πολυμέσα που εισάγει με μεγαλύτερη ακρίβεια και σαφήνεια. Ο *Μηχανισμός Κατηγοριοποίησης Άρθρων και Πολυμέσων* παρέχει τη δυνατότητα αντιστοίχισης άρθρων και πολυμέσων με λέξεις – κλειδιά και κατάταξης αυτών σε κατηγορίες και υποκατηγορίες. Η τελευταία διαδικασία είναι ιδιαίτερα σημαντική, γιατί αν δεν εκτελεστεί με προσοχή από τον αρθρογράφο υπάρχει η πιθανότητα αδυναμίας εύρεσης του εισαχθέντος άρθρου ή μέσου στο σύστημα με τη χρήση του Υποσυστήματος Υποβολής Ερωτήσεων.

3.1.4 Υποσύστημα Υποβολής Ερωτήσεων

Το Υποσύστημα Υποβολής Ερωτήσεων (βλ. και [Καρασ98]) είναι επίσης ενσωματωμένο στο Υποσύστημα Συγγραφής Άρθρων και παρέχει τη δυνατότητα αναζήτησης άρθρων και πολυμέσων που υπάρχουν στο σύστημα. Μέσω του

Μηχανισμού Κατασκευής Ερώτησης ο αρθρογράφος επιλέγει τις λέξεις – κλειδιά και τις κατηγορίες και υποκατηγορίες όπου πιστεύει ότι κατατάσσονται τα άρθρα ή πολυμέσα που επιθυμεί να εντοπίσει. Ο *Μηχανισμός Εκτέλεσης Ερώτησης* εξάγει τα αποτελέσματα της λογικής ερώτησης που απευθύνεται στη βάση δεδομένων, και ο *Μηχανισμός Κατάταξης Αποτελεσμάτων* (βλ. και [Σκον98]) κατατάσσει τα αποτελέσματα σύμφωνα με τη σχετικότητα τους με τις αρχικές παραμέτρους της ερώτησης. Τέλος αυτά παραδίδονται στο Μηχανισμό Παρουσίασης Αποτελεσμάτων Ερώτησης του Υποσυστήματος Συγγραφής Άρθρων και παρουσιάζονται στον αρθρογράφο.

3.2 Σύστημα Εξυπηρετητών

Το σύστημα εξυπηρετητών αποτελείται από τρεις εξυπηρετητές, ο κάθε ένας από τους οποίους μπορεί να βρίσκεται σε οποιοδήποτε σημείο του διαδικτύου, και οι οποίοι εξυπηρετούν αιτήσεις που δέχονται τόσο από το Σύστημα Παραγωγής Άρθρων όσο και από το Σύστημα Υπηρεσιών Νέων στους τελικούς χρήστες. Ο ρόλος του κάθε ενός από αυτούς περιγράφεται αναλυτικά παρακάτω.

3.2.1 Εξυπηρετητής Βάσης Δεδομένων

Ο Εξυπηρετητής Βάσης Δεδομένων είναι ίσως το σημαντικότερο τμήμα του συστήματος καθώς είναι υπεύθυνο για την διασφάλιση των παρεχόμενων υπηρεσιών και για τη συνέπεια του συστήματος. Αυτός υλοποιεί το διάγραμμα οντοτήτων – σχέσεων που σχεδιάστηκε για τις ανάγκες του συστήματος. Εκεί αποθηκεύονται σε δομημένη μορφή τα σημαντικότερα δεδομένα του συστήματος και από αυτόν εξυπηρετούνται οι αιτήσεις αποθήκευσης και ανάκτησης δεδομένων που δέχεται από όλα τα υποσυστήματα. Στην ουσία είναι ένα ολοκληρωμένο σύστημα διαχείρισης σχεσιακών βάσεων δεδομένων, τα κύρια χαρακτηριστικά του οποίου είναι η επεξεργασία (processing) και βελτιστοποίηση (optimization) των ερωτήσεων ανάκτησης, οι συνδιαλλαγές (transactions), ο ταυτοχρονισμός (concurrency) και η ανάνηψη (recovery). Τα υποσυστήματα επικοινωνούν με τον Εξυπηρετητή Βάσης Δεδομένων μέσω του πρωτοκόλλου JDBC (Java DataBase Connectivity), της μεθόδου δηλαδή σύνδεσης των Java Applets με βάσεις δεδομένων.

3.2.2 Εξυπηρετητής Παγκόσμιου Ιστού

Ο Εξυπηρετητής Παγκόσμιου Ιστού είναι υπεύθυνος για την παροχή των υπηρεσιών νέων στους τελικούς χρήστες, ενεργοποιώντας τους κατάλληλους μηχανισμούς παρουσίασης μέσω των Servlets που τρέχουν στην πλευρά του εξυπηρετητή. Επίσης παρέχει τις απαραίτητες Java κλάσεις στους χρήστες του συστήματος, έτσι ώστε αυτοί να έχουν πρόσβαση στα υποσυστήματα, που τρέχουν στο σταθμό εργασίας τους, μέσω των Java Applets. Οι αρθρογράφοι και οι τελικοί χρήστες επικοινωνούν με τον Εξυπηρετητή Παγκόσμιου Ιστού κάνοντας χρήση του πρωτοκόλλου μεταφοράς υπερκειμένου HTTP (HyperText Transfer Protocol) που υλοποιούν οι φυλλομετρητές παγκόσμιου ιστού (Web Browsers).

3.2.3 Εξυπηρετητής Μεταφοράς Αρχείων

Ο Εξυπηρετητής Μεταφοράς Αρχείων είναι υπεύθυνος για την υποστήριξη του Μηχανισμού Εισαγωγής Πολυμέσων, μέσω του οποίου οι αρθρογράφοι εισάγουν νέα πολυμέσα στο σύστημα. Συγκεκριμένα, κατά την ολοκλήρωση της διαδικασίας εισαγωγής ενός νέου μέσου στο σύστημα αποστέλλεται αίτηση μεταφοράς του περιχομένου του στον Εξυπηρετητή Μεταφοράς Αρχείων και τότε αρχίζει η μεταφορά του σε αυτόν. Μόλις αυτή ολοκληρωθεί, το μέσο έχει αποθηκευτεί στη δευτερεύουσα μνήμη και η διαδικασία εισαγωγής του θεωρείται επιτυχημένη. Ο Μηχανισμός Εισαγωγής Πολυμέσων επικοινωνεί με τον Εξυπηρετητή Μεταφοράς Αρχείων με το πρωτόκολλο μεταφοράς αρχείων FTP (File Transfer Protocol).

3.3 Σύστημα Υπηρεσιών Νέων

Το σύστημα αυτό είναι υπεύθυνο για την παροχή των υπηρεσιών του συστήματος στους τελικούς χρήστες και αποτελείται από τέσσερα κύρια υποσυστήματα. Κάποια τμήματα των υποσυστημάτων αυτών τρέχουν στο σταθμό εργασίας του τελικού χρήστη, ο οποίος έχει τη δυνατότητα πρόσβασης στο σύστημα μέσω του διαδικτύου, ενώ άλλα τρέχουν στην πλευρά του συστήματος εξυπηρετητών. Στη συνέχεια περιγράφονται οι ανάγκες που καλύπτει κάθε υποσύστημα και μια σύντομη αναφορά στη λειτουργικότητα των τμημάτων τους.

3.3.1 Υποσύστημα Παρουσίασης Νέων

Το Υποσύστημα Παρουσίασης Νέων αναλαμβάνει την παρουσίαση του συστήματος στον τελικό χρήστη με τον καλύτερο δυνατό τρόπο, έτσι ώστε η αλληλεπίδραση αυτών των δύο να είναι όσο πιο φιλική και εύκολη γίνεται. Συγκεκριμένα καθορίζει το περιεχόμενο και κατασκευάζει τις HTML σελίδες που παρουσιάζονται στον τελικό χρήστη ανάλογα με τις αιτήσεις που αυτός υποβάλλει μέσω του εξυπηρετητή παγκόσμιου ιστού. Αποτελείται από πέντε κύριους μηχανισμούς, όπου ο καθένας καλύπτει συγκεκριμένες ανάγκες παρουσίασης.

Ο *Μηχανισμός Πλοήγησης* επιτρέπει στον τελικό χρήστη την επισκόπηση των κατηγοριών και υποκατηγοριών του συστήματος, καθώς και μια σύντομη περιγραφή των νεότερων άρθρων που αυτές περιέχουν. Ο *Μηχανισμός Παρουσίασης Αποτελεσμάτων Ερώτησης* αναλαμβάνει την παρουσίαση σύντομων περιγραφών των άρθρων που προκύπτουν από τις ερωτήσεις που ο τελικός χρήστης υποβάλλει στο σύστημα μέσω του Υποσυστήματος Υποβολής Ερωτήσεων.

Ο *Μηχανισμός Παρουσίασης Διαγραμμάτων* εξασφαλίζει ότι η ενημέρωση των τελικών χρηστών από το σύστημα φιλτράρεται σύμφωνα με τις προσωπικές τους προτιμήσεις, τις οποίες ο τελικός χρήστης έχει δηλώσει και αποθηκεύσει μέσω του Υποσυστήματος Εξατομικεύσεων. Ο *Μηχανισμός Παρουσίασης Νέων Κατ' Απαίτηση* μοιάζει αρκετά με τον Μηχανισμό Παρουσίασης Διαγραμμάτων, μόνο που αυτός ενεργοποιείται στην πλευρά του τελικού χρήστη, και φροντίζει για την περιληπτική παρουσίαση σε αυτόν μόνο των νεότερων άρθρων που τον ενδιαφέρουν. Τέλος, ο *Μηχανισμός Παρουσίασης Εγγράφων* πραγματοποιεί την ολοκληρωμένη παρουσίαση των χαρακτηριστικών ενός άρθρου. Οι παραπάνω μηχανισμοί ενεργοποιούνται στην πλευρά του συστήματος εξυπηρετητών, όπως φαίνεται και στο Σχήμα 3-1.

3.3.2 Υποσύστημα Υποβολής Ερωτήσεων

Το Υποσύστημα Υποβολής Ερωτήσεων (βλ. και [Καρασ98]) παρέχει τη δυνατότητα αναζήτησης άρθρων του συστήματος από τους τελικούς χρήστες και είναι παρόμοιο με αυτό που παρουσιάστηκε στο σύστημα παραγωγής άρθρων. Μόνο που εδώ όλοι οι μηχανισμοί του δεν τρέχουν στην πλευρά του τελικού χρήστη, όπως φαίνεται και στο

Σχήμα 3-1. Συγκεκριμένα ο *Μηχανισμός Κατασκευής Ερώτησης* είναι ο μόνος που τρέχει στην πλευρά του χρήστη και του επιτρέπει να επιλέξει τις λέξεις – κλειδιά και τις κατηγορίες και υποκατηγορίες όπου πιστεύει ότι κατατάσσονται τα άρθρα που επιθυμεί να εντοπίσει.

Ο *Μηχανισμός Εκτέλεσης Ερώτησης* και ο *Μηχανισμός Κατάταξης Αποτελεσμάτων* ενεργοποιούνται στην πλευρά του συστήματος εξυπηρετητών και εξάγουν καταταγμένα τα άρθρα που προέκυψαν από την ερώτηση. Το γεγονός αυτό βελτιώνει κατά πολύ την απόδοση του συστήματος, αφού η επεξεργασία και η εκτέλεση της ερώτησης γίνεται τοπικά. Στη συνέχεια τα εξαγόμενα αποτελέσματα παραδίδονται στο *Μηχανισμό Παρουσίασης Αποτελεσμάτων Ερώτησης*, που βρίσκεται επίσης στην πλευρά του συστήματος εξυπηρετητών, και ο οποίος παράγει τις HTML σελίδες που παρουσιάζονται στον τελικό χρήστη.

3.3.3 Υποσύστημα Εξατομικεύσεων

Το Υποσύστημα Εξατομικεύσεων προσδίδει στο σύστημα το χαρακτηρισμό του εξατομικευμένου, υλοποιώντας την αντίστοιχη τεχνολογία. Αποτελείται από δύο κύριους μηχανισμούς, όπου ο καθένας καλύπτει συγκεκριμένες ανάγκες εξατομίκευσης των νέων. Ο *Μηχανισμός Διαχείρισης Διαγραμμάτων* επιτρέπει στον τελικό χρήστη τη δημιουργία ενός νέου ή τη μετατροπή του υπάρχοντος διαγράμματος του. Ενώ μέσω του *Μηχανισμού Επιλογής Προτιμήσεων* ο χρήστης επιλέγει τις κατηγορίες και υποκατηγορίες, καθώς και τις λέξεις – κλειδιά, που πιστεύει ότι περιέχουν και χαρακτηρίζουν νέα τα οποία καλύπτουν τα προσωπικά του ενδιαφέροντα. Και οι δύο παραπάνω μηχανισμοί ενεργοποιούνται στην πλευρά του τελικού χρήστη, όπως φαίνεται και στο Σχήμα 3-1.

3.3.4 Υποσύστημα Νέων Κατ' Απαίτηση

Το Υποσύστημα Νέων Κατ' Απαίτηση προσδίδει στο σύστημα το χαρακτηρισμό του κατ' απαίτηση, και αποτελείται από ένα μηχανισμό. Ο *Μηχανισμός Ανανέωσης Νέων* ενεργοποιείται στην πλευρά του τελικού χρήστη και περιοδικά συνδέεται με το σύστημα για την αναζήτηση νεοεκδοθέντων άρθρων που ταιριάζουν με το διάγραμμά του. Σε περίπτωση ύπαρξης τέτοιων άρθρων ειδοποιεί ανάλογα τον τελικό χρήστη.

3.4 Ανακεφαλαίωση

Όπως φάνηκε σε αυτό το κεφάλαιο, η αρχιτεκτονική του συστήματος “Hypermedia Custom News System” καλύπτει όλες τις βασικές ανάγκες ενός συστήματος εξατομικευμένων νέων κατ’ απαίτηση και παρέχει όλη τη λειτουργικότητα που ο τελικός χρήστης μπορεί να ζητήσει από ένα τέτοιο σύστημα. Επίσης, φαίνονται και τα τμήματα που υλοποιεί η παρούσα διατριβή και η σημαντικότητα τους μέσα στο σύστημα. Συνοψίζοντας, τα κυριότερα πλεονεκτήματα της αρχιτεκτονικής του συστήματος είναι τα παρακάτω:

- Είναι γενική ώστε να προσαρμόζεται εύκολα σε συστήματα με διαφορετικού είδους περιεχόμενο.
- Είναι ανοικτή ώστε να ενσωματώνει εύκολα υπηρεσίες και λειτουργικότητες απαραίτητες σε εξειδικευμένα συστήματα.
- Υποστηρίζει αυθαίρετο αριθμό γλωσσών.
- Παρέχει ένα ολοκληρωμένο σύστημα παραγωγής νέων στους αρθρογράφους, με πολλές δυνατότητες μορφοποίησης και διαχείρισης των άρθρων τους.
- Υλοποιεί ένα εξελιγμένο σύστημα κατηγοριοποίησης και αναζήτησης άρθρων και πολυμέσων.
- Διαθέτει κατάλληλους μηχανισμούς παρουσίασης για την πληρέστερη και ευκολότερη ενημέρωση των τελικών χρηστών από το σύστημα.
- Υποστηρίζει τις τεχνολογίες εξατομίκευσης των νέων και των νέων κατ’ απαίτηση με τη δημιουργία και παρουσίαση διαγραμμάτων από και προς τους τελικούς χρήστες αντίστοιχα.

Στο επόμενο κεφάλαιο θα εξεταστεί η βάση δεδομένων που προκύπτει από την παραπάνω αρχιτεκτονική και υποστηρίζει όλες τις διεργασίες των υποσυστημάτων και τμημάτων του συστήματος.

4 Η ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ

Στο κεφάλαιο αυτό παρουσιάζεται η πορεία σχεδιασμού και ανάπτυξης της βάσης δεδομένων του συστήματος “Hypermedia Custom News System”. Μέσω της ανάλυσης απαιτήσεων καθορίζονται οι οντότητες του συστήματος και οι σχέσεις μεταξύ τους, έτσι όπως αυτές προκύπτουν από την αρχιτεκτονική που παρουσιάστηκε στο προηγούμενο κεφάλαιο, και με σκοπό να υποστηρίξουν όλα τα υποσυστήματα και τμήματα του συστήματος. Από αυτή προκύπτει το διάγραμμα οντοτήτων – σχέσεων, από το οποίο καθορίζεται με τη σειρά του το σχεσιακό μοντέλο της βάσης δεδομένων του συστήματος που υλοποιείται στον εξυπηρετητή βάσεων δεδομένων.

4.1 Ανάλυση Απαιτήσεων

Παρακάτω παρουσιάζονται οι οντότητες που κρίθηκαν αναγκαίες και καθορίστηκαν βάσει της αρχιτεκτονικής του συστήματος.

- **Η οντότητα γλωσσών (languages entity)** είναι αυτή που περιέχει την περιγραφή και την κωδικοποίηση των γλωσσών που υποστηρίζει το σύστημα και είναι το πρώτο και σημαντικότερο βήμα για την υποστήριξη αυθαίρετου αριθμού γλωσσών.
- **Η οντότητα αρθρογράφων (authors entity)** κρατά πληροφορίες για τους αρθρογράφους που αναγνωρίζει το σύστημα και αποτελεί τη βάση ανάπτυξης του μηχανισμού ασφαλείας.
- **Η οντότητα κατηγοριών (categories entity)** περιέχει τα δύο επίπεδα κατηγοριοποίησης που ακολουθεί το σύστημα, δηλαδή βασικές κατηγορίες και υποκατηγορίες αυτών. Αποτελεί τη μία από τις δύο οντότητες υποστήριξης του υποσυστήματος κατηγοριοποίησης.

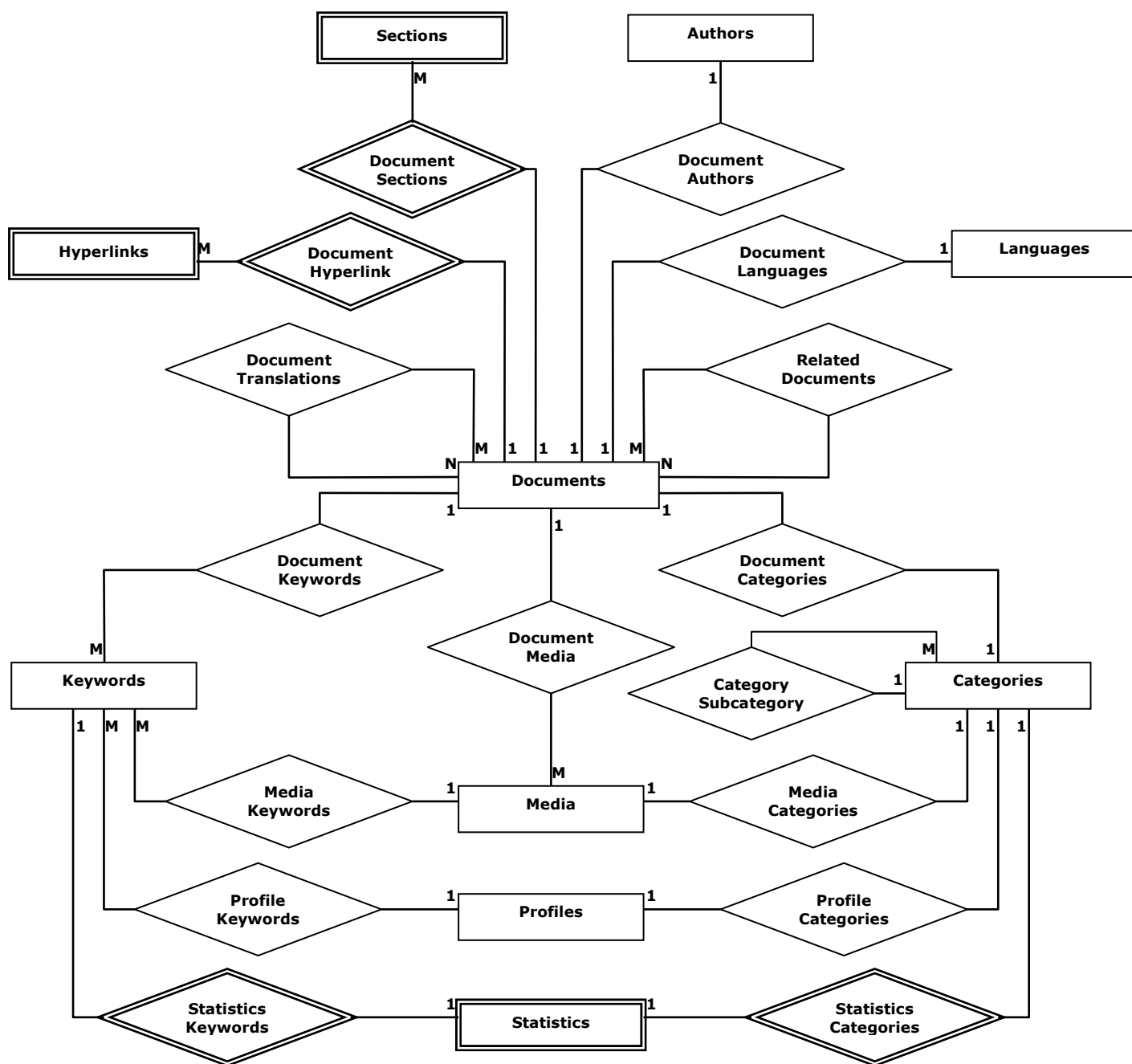
- **Η οντότητα λέξεων - κλειδιών (keywords entity)** είναι η δεύτερη οντότητα υποστήριξης του υποσυστήματος κατηγοριοποίησης και εσωκλείνει την ιεραρχία των λέξεων - κλειδιών που διαθέτει το σύστημα για τον χαρακτηρισμό των άρθρων και των πολυμέσων.
- **Η οντότητα στατιστικών (statistics entity)** κρατά στατιστικά στοιχεία για τον αριθμό εγγράφων ανά κατηγορία, υποκατηγορία και λέξη – κλειδί και έχει να κάνει με το μοντέλο υπολογισμού σχετικότητας που υλοποιεί το σύστημα (βλ. και [Σκον98]). Για τις κατηγορίες και λέξεις – κλειδιά που υπάρχουν στο σύστημα η οντότητα ενημερώνεται μέσω δυο σχέσεων με τις αντίστοιχες οντότητες.
- **Η οντότητα πολυμέσων (media entity)** μπορεί να περιγράψει τρία είδη πολυμέσων, εικόνες, ήχους και video. Περιέχει κάθε χαρακτηριστικό αναγκαίο και χρήσιμο για την αποθήκευση και παρουσίαση του, είτε αυτό είναι ενσωματωμένο σε κάποιο άρθρο είτε αυτό προβάλλεται ανεξάρτητα. Συνδέεται με τις οντότητες κατηγοριών και λέξεων – κλειδιών έτσι ώστε κάθε μέσο να κατηγοριοποιείται μέσα στο σύστημα. Αυτός ο συσχετισμός είναι αναγκαίος για την αναζήτηση και τον εντοπισμό τους από το υποσύστημα υποβολής ερωτήσεων.
- **Η οντότητα εγγράφων (documents entity)** είναι η σημαντικότερη οντότητα του συστήματος και σχετίζεται σχεδόν με όλες τις υπόλοιπες οντότητες. Περιέχει όλα τα αναγκαία χαρακτηριστικά για τη διαχείριση και την παρουσίαση ενός εγγράφου. Συνδέεται με την οντότητα γλωσσών για τον καθορισμό της γλώσσας στην οποία είναι γραμμένο και με την οντότητα αρθρογράφων για τον καθορισμό του συγγραφέα του. Επίσης και αυτή η οντότητα συνδέεται με τις οντότητες κατηγοριών και λέξεων – κλειδιών ώστε κάθε έγγραφο να κατηγοριοποιείται μέσα στο σύστημα. Το υποσύστημα υποβολής ερωτήσεων είναι άλλωστε υπεύθυνο για την αναζήτηση και πολυμέσων και εγγράφων. Μέσω δύο σχέσεων με τον εαυτό της η οντότητα μπορεί να περιγράψει τα άρθρα που συσχετίζονται με κάθε έγγραφο καθώς και τις μεταφράσεις του κάθε άρθρου. Τέλος υπάρχει σύνδεση της οντότητας αυτής με την οντότητα πολυμέσων για τον καθορισμό των πολυμέσων που περιέχει κάθε έγγραφο και για τον τρόπο που παρουσιάζονται μέσα σε αυτό.

Τόσο το υποσύστημα συγγραφής άρθρων όσο και το υποσύστημα παρουσίασης νέων στηρίζονται κατά πρώτο λόγο σε αυτή την οντότητα, και κατά δεύτερο στην οντότητα πολυμέσων.

- **Η οντότητα ενότητων (sections entity)** περιέχει τις ενότητες στις οποίες μπορεί να χωρίζεται ένα έγγραφο και συνδέεται με την οντότητα εγγράφων για το σκοπό αυτό.
- **Η οντότητα διασυνδέσεων (hyperlinks entity)** περιέχουν τις διασυνδέσεις που μπορεί να περιέχονται μέσα στο κείμενο ενός εγγράφου, και παραπέμπουν τον αναγνώστη σε κάποια διεύθυνση στο διαδίκτυο ή σε κάποιο μέσο. Συνδέεται και αυτή με την οντότητα εγγράφων.
- **Η οντότητα διαγραμμάτων (profiles entity)** αποτελεί τα προσωπικά διαγράμματα των τελικών χρηστών. Αυτά αποτελούνται από τις προσωπικές πληροφορίες του χρήστη και από την περιγραφή των ενδιαφερόντων του. Τις προτιμήσεις του ο χρήστης τις εκφράζει επιλέγοντας κατηγορίες και λέξεις – κλειδιά, οπότε η οντότητα αυτή πρέπει να συνδέεται με τις οντότητες κατηγοριών και λέξεων – κλειδιών. Τα υποσυστήματα εξατομικεύσεων και νέων κατ' απαίτηση στηρίζονται αποκλειστικά στην οντότητα που μόλις περιγράψαμε.

4.2 Διάγραμμα Οντοτήτων – Σχέσεων

Το διάγραμμα οντοτήτων – σχέσεων που προκύπτει από την παραπάνω ανάλυση απαιτήσεων φαίνεται στο Σχήμα 4-1 της επόμενης σελίδας. Στο σχήμα αυτό φαίνονται όλες οι οντότητες που περιγράψαμε παραπάνω, και οι οποίες υποστηρίζουν όλα τα υποσυστήματα και τους μηχανισμούς αυτών όπως απεικονίζονται στην αρχιτεκτονική του συστήματος. Επίσης φαίνονται οι σχέσεις που συνδέουν τις οντότητες μεταξύ τους και ο τύπος της κάθε σχέσης.



Σχήμα 4-1 Διάγραμμα οντοτήτων – σχέσεων του συστήματος νέων
“Hypermedia Custom News System”

4.3 Σχεσιακό Μοντέλο Βάσης Δεδομένων

Από το παραπάνω διάγραμμα οντοτήτων – σχέσεων εξάγεται το σχεσιακό μοντέλο της βάσης δεδομένων του συστήματος που φαίνεται στο Σχήμα 4-2. Διακρίνονται οι πίνακες που προέκυψαν από τις οντότητες και από τις σχέσεις αυτών, μαζί με όλα τα πεδία τους.

```

    erDiagram
        Document_Hyperlinks ||--o{ Documents : "has"
        Authors ||--o{ Documents : "has"
        Document_Sections ||--o{ Documents : "has"
        Languages ||--o{ Documents : "has"
        Translations ||--o{ Documents : "has"
        Documents ||--o{ Document_Keywords : "has"
        Documents ||--o{ Document_Related : "has"
        Documents ||--o{ Document_Media : "has"
        Documents ||--o{ Categories : "has"
        Media ||--o{ Categories : "has"
        Media ||--o{ Media_Keywords : "has"
        Media ||--o{ Profiles : "has"
        Profiles ||--o{ Profile_Keywords : "has"
        Profiles ||--o{ Profile_Categories : "has"
        Profiles ||--o{ Statistics : "has"
        Keywords ||--o{ Document_Keywords : "has"
        Keywords ||--o{ Profile_Keywords : "has"
        Keywords ||--o{ Statistics : "has"
        Document_Keywords ||--o{ Statistics : "has"
        Profile_Keywords ||--o{ Statistics : "has"
        Document_Related ||--o{ Statistics : "has"
        Categories ||--o{ Statistics : "has"
  
```

Document_Hyperlinks

- document_id
- hyperlink_id
- offset_start
- offset_end
- url
- type
- title

Authors

- id
- name
- login
- password

Document_Sections

- document_id
- offset_start
- offset_end

Languages

- id
- name
- rank
- ie_encoding
- nc_encoding

Translations

- Li_document_id

Documents

- id
- title
- abstract
- language
- text
- attributes
- category_id
- subcategory_id
- date
- author

Document_Keywords

- document_id
- keyword_id
- weight
- language

Document_Related

- document1_id
- document2_id
- url
- type
- title

Document_Media

- document_id
- media_id
- offset
- caption
- url
- type
- alignment

Categories

- id
- parent_id
- Li_description
- Li_counter_documents

Media

- id
- date
- url
- type
- width
- height
- size
- category_id
- subcategory_id
- Li_description

Media_Keywords

- media_id
- keyword_id
- weight

Keywords

- id
- parent_id
- Li_description

Profile_Keywords

- keyword_id
- profile_id

Profiles

- id
- first_name
- last_name
- address
- city
- zip
- country
- email
- tel
- fax
- login
- password
- dates
- ordering
- kboolean
- language_id

Profile_Categories

- category_id
- profile_id

Statistics

- subcategory_id
- category_id
- keyword_id
- Li_counter_documents

4.4 Ανακεφαλαίωση

Όπως φάνηκε σε αυτό το κεφάλαιο, ο σχεδιασμός της βάσης δεδομένων του συστήματος “Hypermedia Custom News System” πραγματοποιήθηκε με τέτοιο τρόπο ώστε να υποστηρίζει πλήρως τις απαιτήσεις της αρχιτεκτονικής του συστήματος. Το διάγραμμα οντοτήτων – σχέσεων είναι πλήρες, απλό και κατανοητό με αποτέλεσμα το σχεσιακό μοντέλο που προκύπτει από αυτό να κληρονομεί τα ίδια χαρακτηριστικά. Ο σχεδιασμός αυτός της επιτρέπει επίσης να είναι ευέλικτη σε αλλαγές και σε προσθήκες άλλων οντοτήτων και σχέσεων στο μέλλον.

Στο επόμενο κεφάλαιο περιγράφεται το υποσύστημα συγγραφής άρθρων, ένα από τα δύο μεγάλα τμήματα της παρούσας διατριβής. Αναλύονται οι βασικές δομές που χρησιμοποιεί, καθώς και οι μηχανισμοί και η λειτουργικότητα τους. Το άλλο μεγάλο τμήμα της διατριβής είναι το υποσύστημα παρουσίασης νέων και περιγράφεται στο κεφάλαιο 6.

5 ΥΠΟΣΥΣΤΗΜΑ ΣΥΓΓΡΑΦΗΣ ΑΡΘΡΩΝ

Στο κεφάλαιο αυτό περιγράφεται λεπτομερειακά το Υποσύστημα Συγγραφής Άρθρων που αποτελεί τμήμα του Συστήματος Παραγωγής Άρθρων. Αρχικά παρουσιάζονται οι δομές που αναπτύχθηκαν για την αναπαράσταση του εγγράφου και του μέσου, τα οποία είναι και οι κεντρικές οντότητες του συστήματος. Στη συνέχεια παρουσιάζεται αναλυτικότερα το τμήμα της αρχιτεκτονικής του συστήματος που αντιστοιχεί στο Υποσύστημα Συγγραφής Άρθρων, επιδεικνύοντας τις διεργασίες που περιέχει κάθε μηχανισμός του.

Ακολούθως, περιγράφονται λεπτομερειακά κάθε μία από τις διεργασίες αυτές και πως αλληλεπιδρούν μεταξύ τους για τη συγγραφή ενός άρθρου. Επίσης αναλύονται οι δομές που αναπτύχθηκαν για τις ανάγκες κάθε διεργασίας και πως αυτές συμβάλλουν στη συγκρότηση της δομής ενός εγγράφου.

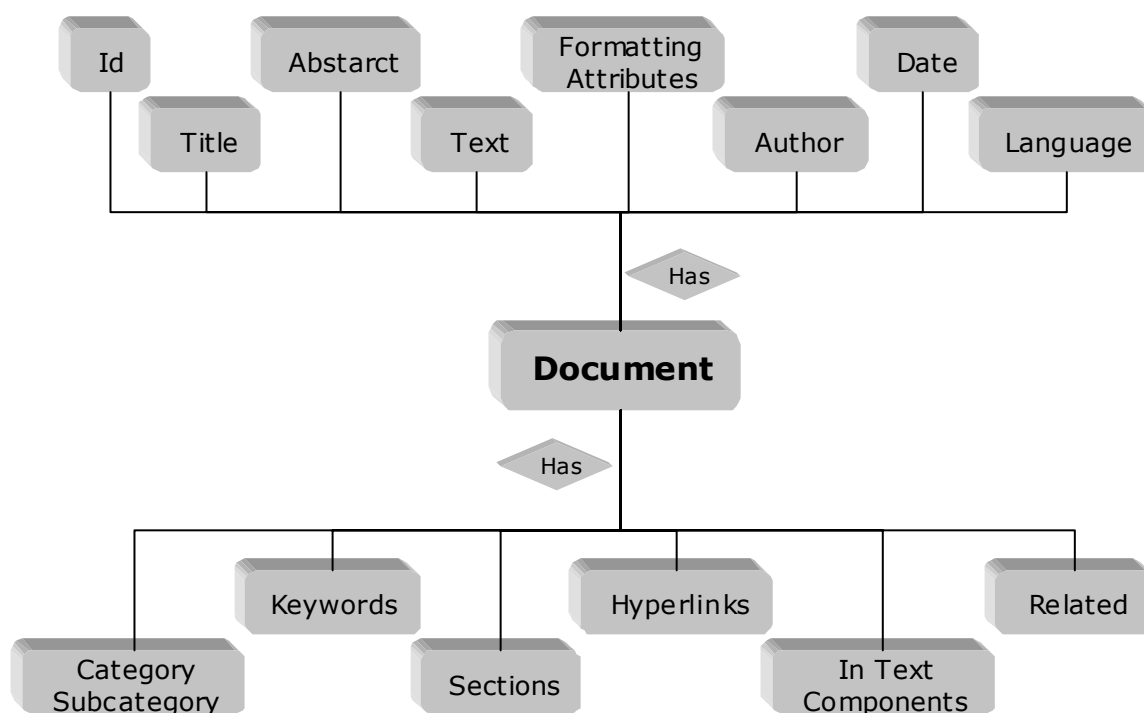
5.1 Δομή Εγγράφου

Το έγγραφο αποτελεί τη σημαντικότερη οντότητα του συστήματος. Σε αυτή βασίστηκε η ανάπτυξη του Υποσυστήματος Συγγραφής Άρθρων και ο σχεδιασμός του σχεσιακού μοντέλου βάσης δεδομένων που αναφέρεται στο τέταρτο κεφάλαιο. Οι σχέσεις του τελευταίου υλοποιούν τη δομή του εγγράφου έτσι όπως αυτή φαίνεται στο Σχήμα 5-2 της επόμενης σελίδας. Η δομή αυτή περιέχει υποσύνολο του μοντέλου αναπαράστασης εγγράφου που χρησιμοποιεί η γλώσσα HTML, το οποίο περιγράφεται αναλυτικά στις επόμενες παραγράφους.

Κάθε έγγραφο αναγνωρίζεται από έναν μοναδικό κωδικό που του δίδεται κατά την έκδοση και αποθήκευση του στη βάση δεδομένων. Άλλα αναγνωριστικά χαρακτηριστικά του εγγράφου είναι ο αρθρογράφος που το δημιούργησε, η ημερομηνία και ώρα έκδοσης ή τελευταίας τροποποίησης του και η γλώσσα στην

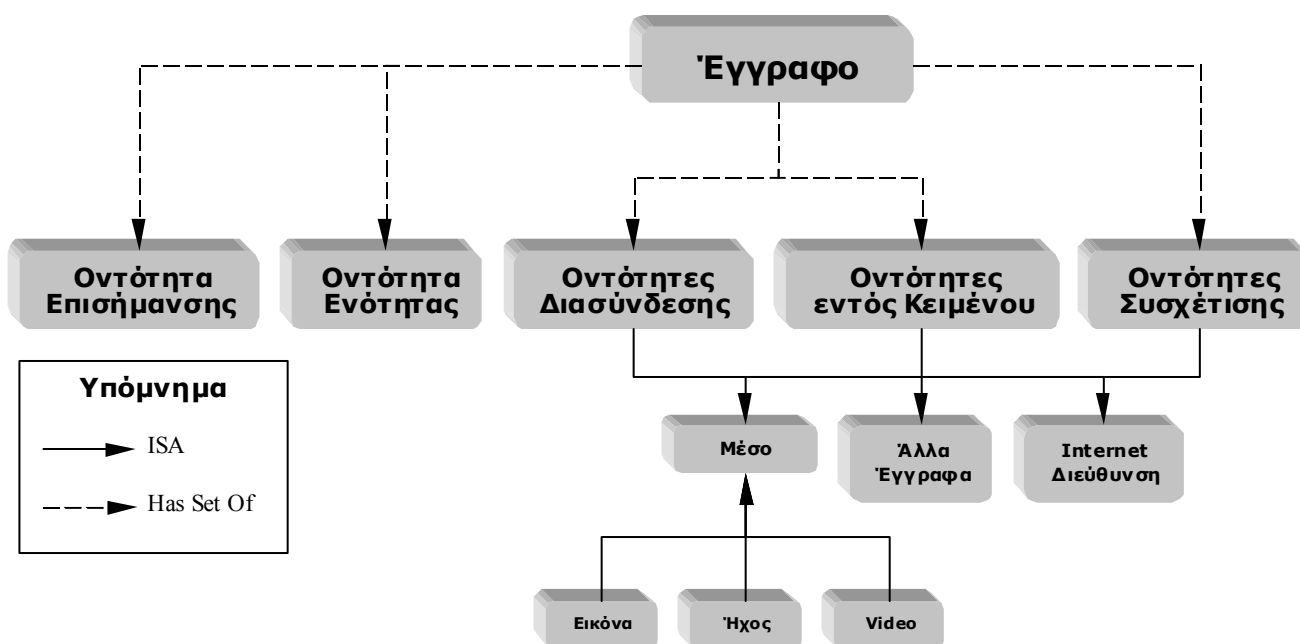
οποία είναι γραμμένο. Το τελευταίο χαρακτηριστικό είναι ιδιαίτερα σημαντικό, γιατί εξασφαλίζει τη σωστή παρουσίαση του εγγράφου στον τελικό χρήστη.

Το έγγραφο περιέχει επίσης και δύο χαρακτηριστικά για την κατηγοριοποίηση του μέσα στο σύστημα. Το ένα είναι η κατηγορία και η υποκατηγορία στην οποία ανήκει, και το δεύτερο είναι ένα σύνολο από λέξεις – κλειδιά που ο αρθρογράφος προσδίδει στο έγγραφο ως αντιπροσωπευτικές για το περιεχόμενό του. Αυτά τα χαρακτηριστικά είναι σημαντικά για τον εύκολο και έγκυρο εντοπισμό του εγγράφου, γιατί βάσει αυτών κυρίως ο αρθρογράφος και ο τελικός χρήστης αναζητούν ένα έγγραφο μέσα στο σύστημα. Όλα τα παραπάνω χαρακτηριστικά, μαζί με τον τίτλο, είναι τα ελάχιστα που μπορεί να έχει κάθε έγγραφο και είναι υποχρεωτικά. Άλλα, επιπλέον χαρακτηριστικά του εγγράφου σε μορφή κειμένου, εκτός του τίτλου, είναι η περίληψη του περιεχομένου του και το κυρίως κείμενο του.



Σχήμα 5-1 Δομή της οντότητας Εγγράφου

Εκτός από τα παραπάνω βασικά του χαρακτηριστικά, ένα έγγραφο μπορεί να περιέχει και άλλες οντότητες, οι οποίες έχουν ως στόχο τη μορφοποίηση του. Με την ενσωμάτωση αυτών το έγγραφο εμπλουτίζεται και η παροχή πληροφορίας στον τελικό χρήστη γίνεται πληρέστερη και πιο λειτουργική. Οι οντότητες αυτές, καθώς και τι τύπου μπορεί να είναι η κάθε μία, διακρίνονται στο Σχήμα 5-3 και περιγράφονται αναλυτικά στη συνέχεια.



Σχήμα 5-2 Οντότητες Μορφοποίησης Εγγράφου

- Το κυρίως κείμενο του έγγραφου μπορεί να περιέχει ένα σύνολο από οντότητες επισημάνσεων σε διάφορα σημεία του, οι οποίες τονίζουν κάποιες φράσεις ανάλογα με τη σημασία που θέλει να τους δώσει ο αρθρογράφος (π.χ. δηλώσεις προσώπων, συμπέρασμα παραγράφων κλπ). Οι επισημάνσεις είναι τριών τύπων, έντονοι, πλάγιοι και υπογραμμισμένοι χαρακτήρες, και είναι παρόμοιες με αυτές που παρέχουν οι κλασικοί επεξεργαστές κειμένου και οι επεξεργαστές HTML εγγράφων.
- Επιπλέον των επισημάνσεων το κυρίως κείμενο του εγγράφου μπορεί να περιέχει ένα σύνολο από οντότητες ενότητων, οι οποίες το διαχωρίζουν σε θεματικές

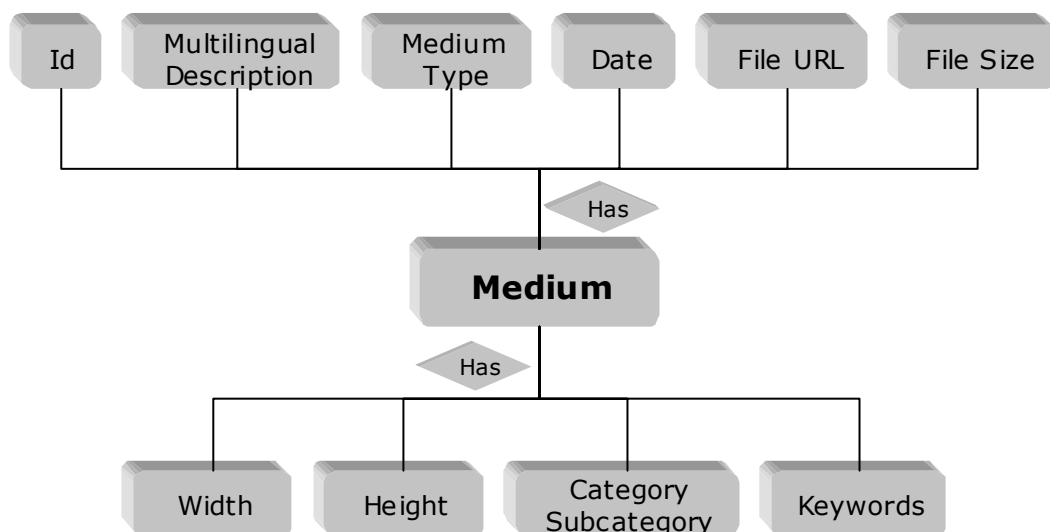
ενότητες με δικό τους τίτλο η κάθε μία. Ιδιαίτερα χρήσιμη οντότητα αν το κείμενο του άρθρου είναι μεγάλο σε μέγεθος, γιατί ο τελικός χρήστης έχει τη δυνατότητα άμεσης μετακίνησης από το ένα τμήμα στο άλλο.

- Επίσης υπάρχει η δυνατότητα εισαγωγής οντοτήτων διασύνδεσης (hyperlinks) στο έγγραφο, οι οποίες ενσωματώνονται πάνω σε φράσεις του κυρίως κειμένου και επισημαίνονται κατάλληλα. Η δομή των οντοτήτων αυτών έχει παρθεί από το μοντέλο αναπαράστασης εγγράφων της HTML. Μέσω αυτών ο τελικός χρήστης μπορεί να μεταπηδήσει σε στοιχεία που συμπληρώνουν πληροφοριακά το έγγραφο που διαβάζει. Τα στοιχεία αυτά μπορεί να είναι πολυμέσα, όπως εικόνα, ήχος ή video, άλλα έγγραφα του συστήματος ή μία Internet διεύθυνση (URL) κάποιας ιστοσελίδας (Web Page).
- Υπάρχει επίσης η δυνατότητα ενσωμάτωσης στο έγγραφο συνόλου οντοτήτων, οι οποίες επιδεικνύουν το περιεχόμενο τους μέσα στο κείμενο και περιβάλλονται από αυτό. Οι οντότητες αυτές μπορούν να συνοδεύονται από κάποιο σχόλιο (π.χ. εικόνα με λεζάντα), ενώ μπορεί να καθοριστεί και η οριζόντια ευθυγράμμιση τους (left or right alignment). Οι οντότητες αυτές, όπως και στην περίπτωση των οντοτήτων διασύνδεσης, μπορεί να είναι πολυμέσα, άλλα έγγραφα του συστήματος ή μία απόλυτη Internet διεύθυνση.
- Τέλος, το έγγραφο μπορεί να περιέχει ένα σύνολο από οντότητες συσχετίσεων. Αυτές οι οντότητες τοποθετούνται στο τέλος του εγγράφου και η σημασία τους είναι παρόμοια με αυτή των οντοτήτων διασύνδεσης. Μέσω αυτών δηλαδή ο τελικός χρήστης αποκτά συνολική εικόνα για το θέμα που περιγράφει το έγγραφο που διαβάζει. Και αυτές οι οντότητες μπορεί να είναι πολυμέσα, άλλα έγγραφα του συστήματος ή μία απόλυτη Internet διεύθυνση.

Στη δομή του εγγράφου, το πεδίο των χαρακτηριστικών μορφοποίησης (formatting attributes) κρατά πληροφορίες για τη θέση μέσα στο κυρίως κείμενο των οντοτήτων ενότητας, επισήμανσης, διασύνδεσης, καθώς και για τη θέση των οντοτήτων που παρουσιάζουν το περιεχόμενο τους εντός κειμένου. Αυτές οι οντότητες προβλέπονται από το μοντέλο αναπαράστασης εγγράφου της HTML που βλέπει ο τελικός χρήστης.

5.2 Δομή Μέσου

Το μέσο είναι η δεύτερη σημαντική οντότητα του συστήματος και αναπτύχθηκε για την αναπαράσταση των σημαντικότερων τύπων πολυμέσων, όπως εικόνες, ήχοι και video. Το σχεσιακό μοντέλο βάσης δεδομένων του συστήματος υποστηρίζει τη διαχείριση τέτοιων οντοτήτων και οι σχέσεις του υλοποιούν τη δομή του μέσου έτσι όπως αυτή φαίνεται στο Σχήμα 5-4 που ακολουθεί.



Σχήμα 5-3 Δομή της οντότητας Μέσου

Οι τύποι πολυμέσων που αναπαριστά η παραπάνω δομή είναι εικόνα, ήχος και video. Πιο συγκεκριμένα, το σύστημα υποστηρίζει τα αρχεία εικόνων τύπου GIF και JPG, τα αρχεία ήχων WAV και MP3 και τα αρχεία video τύπου AVI και MPG. Ειδικότερα τα αρχεία τύπου JPG, MP3 και MPG χρησιμοποιούνται ευρύτατα στο διαδίκτυο κυρίως λόγω του μικρού μεγέθους τους, που προκύπτει εφαρμόζοντας τις τελευταίες μεθόδους κωδικοποίησης, και βοηθά στη γρήγορη μεταφορά τους ακόμα και πάνω από δίκτυα χαμηλής ταχύτητας.

Κάθε μέσο αναγνωρίζεται επίσης από έναν μοναδικό κωδικό που του δίδεται μετά την εισαγωγή του στη βάση δεδομένων και τη μεταφορά του περιεχομένου του στον εξυπηρετητή μεταφοράς αρχείων. Επίσης περιλαμβάνει περιγραφή του περιεχομένου του σε όλες τις γλώσσες που υποστηρίζει το σύστημα, οι οποίες θέτονται κατά την

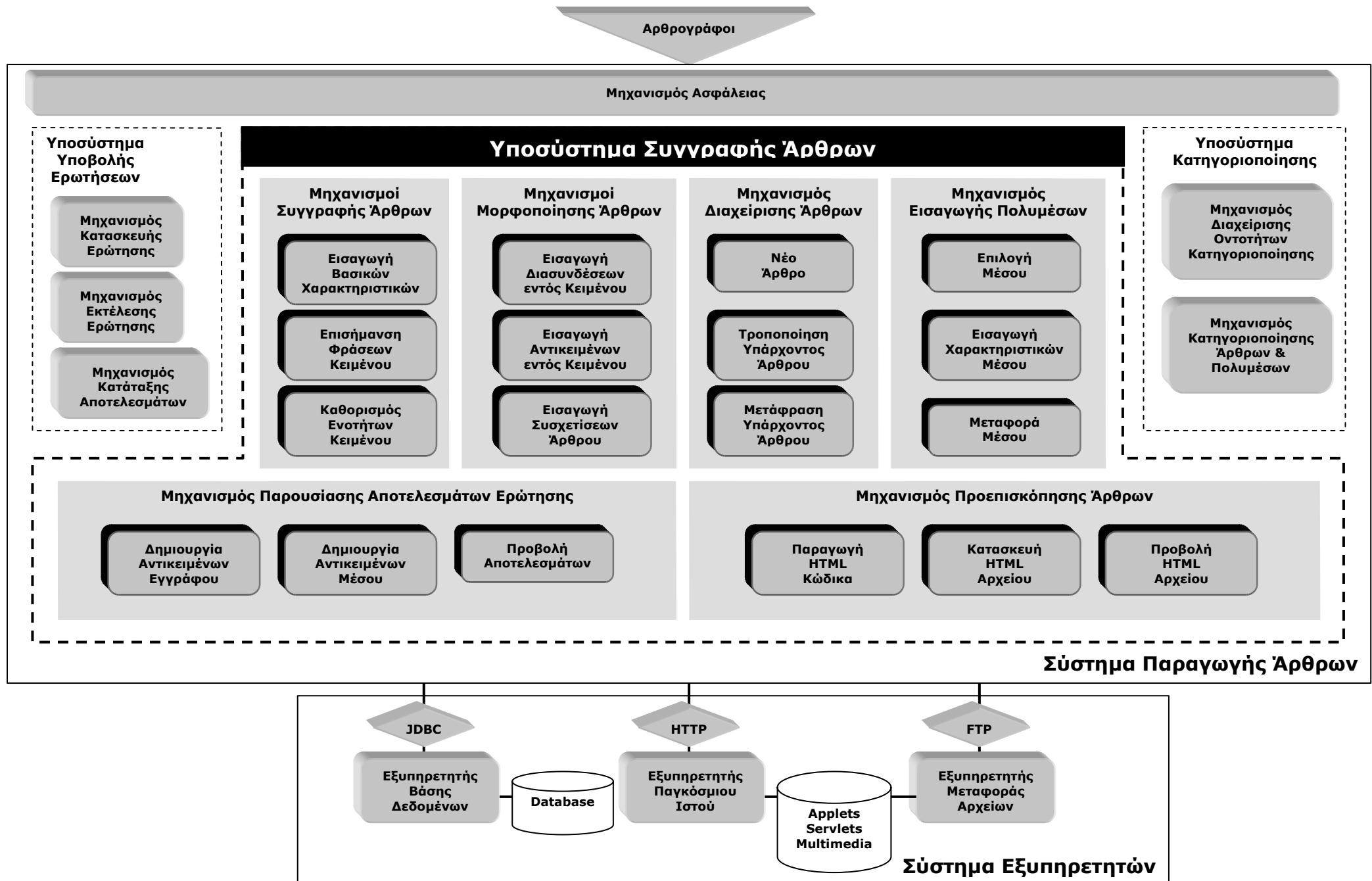
εισαγωγή του μέσου στο σύστημα ή τροποποιώντας το αργότερα. Το γεγονός αυτό προσφέρει τη δυνατότητα εκμετάλλευσης του μέσου από όλους τους αρθρογράφους του συστήματος, ανεξαρτήτως γλώσσας.

Άλλα αναγνωριστικά χαρακτηριστικά του μέσου είναι η ημερομηνία και ώρα εισαγωγής του στο σύστημα, το μέγεθος του αρχείου του, καθώς και το πλάτος και ύψος αυτού αν πρόκειται για εικόνα. Επίσης περιλαμβάνει και το όνομα του αρχείου του στη δευτερεύουσα μνήμη, το οποίο προκύπτει από το μοναδικό κωδικό του και τον τύπο του αρχείου (π.χ. 24.mpg). Τα πολυμέσα του συστήματος τοποθετούνται στον εξυπηρετητή μεταφοράς αρχείων και διαχωρίζονται στη δευτερεύουσα μνήμη σε φακέλους (directories) ανάλογα με τον τύπο τους (picture, audio, video).

Το μέσο κατηγοριοποιείται μέσα στο σύστημα με τον ίδιο τρόπο που κατηγοριοποιείται και ένα έγγραφο. Περιέχει δηλαδή την κατηγορία και την υποκατηγορία στην οποία ανήκει, και ένα σύνολο από λέξεις – κλειδιά που ο αρθρογράφος προσδίδει στο μέσο ως αντιπροσωπευτικές για το περιεχόμενό του. Αυτά τα χαρακτηριστικά είναι εξίσου σημαντικά με τα αντίστοιχα του εγγράφου, γιατί οι αρθρογράφοι αναζητούν ένα μέσο μέσα στο σύστημα με τον ίδιο τρόπο που αναζητούν και ένα έγγραφο. Όλα τα παραπάνω χαρακτηριστικά του μέσου είναι υποχρεωτικά.

5.3 Αρχιτεκτονική

Το Σχήμα 5-1 της επόμενης σελίδας αποτελεί τμήμα του σχήματος που παρουσιάζει την αρχιτεκτονική του συστήματος και αναλύει σε βάθος το Υποσύστημα Συγγραφής Άρθρων. Συγκεκριμένα φαίνονται όλες οι διεργασίες των μηχανισμών που περιέχει το συγκεκριμένο υποσύστημα, καθώς και τα υποσυστήματα από τα οποία πλαισιώνεται για την παραγωγή των άρθρων. Επίσης διακρίνεται το Σύστημα Εξυπηρετητών με το οποίο αλληλεπιδρούν και τα τρία υποσυστήματα του Συστήματος Παραγωγής Άρθρων. Στη συνέχεια περιγράφεται αναλυτικά η λειτουργικότητα των διεργασιών κάθε μηχανισμού, οι δομές που αναπτύχθηκαν για τις ανάγκες τους, καθώς και ο τρόπος με τον οποίο συμβάλλει η κάθε μία στη συγκρότηση της δομής ενός εγγράφου.



Σχήμα 5-4 Αρχιτεκτονική του Υποσυστήματος Συγγραφής Άρθρων του Συστήματος Παραγωγής Νέων

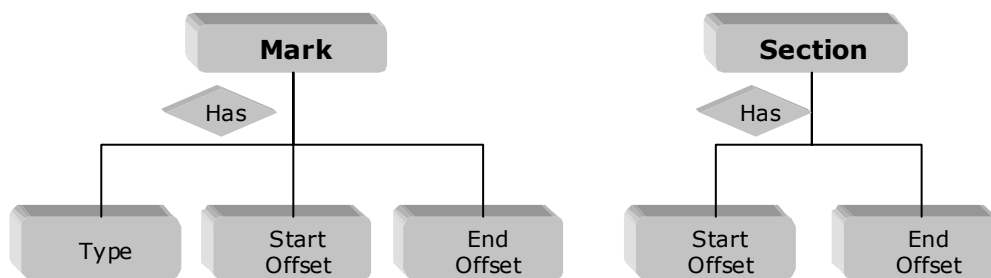
5.4 Μηχανισμοί Συγγραφής Άρθρου

Οι Μηχανισμοί Συγγραφής Άρθρου ομαδοποιούν τρεις διαδικασίες, οι οποίες έχουν να κάνουν με την εισαγωγή των χαρακτηριστικών σε μορφή κειμένου ενός εγγράφου. Μέσω της διαδικασίας Εισαγωγής Βασικών Χαρακτηριστικών του εγγράφου, ο αρθρογράφος εισάγει τον τίτλο, την περίληψη και το κυρίως κείμενο του άρθρου. Επίσης, καθορίζει τη γλώσσα στην οποία γράφει το άρθρο, ενώ αυτόματα εισάγονται η ταυτότητα του αρθρογράφου και η ημερομηνία και ώρα συγγραφής του άρθρου.

Στη συνέχεια με τη διαδικασία Επισήμανσης Φράσεων Κειμένου ο αρθρογράφος εφαρμόζει σε φράσεις του κυρίως κειμένου κλασσικές μεθόδους επισήμανσης, οι οποίες είναι τριών τύπων και παρέχουν τη δυνατότητα μετατροπής χαρακτήρων σε έντονους, πλάγιους και υπογραμμισμένους. Κάθε φορά που εφαρμόζεται μια τέτοια μέθοδος μία οντότητα επισήμανσης δημιουργείται και προστίθεται στον αντίστοιχο πίνακα της δομής του εγγράφου. Η οντότητα επισήμανσης περιέχει τον τύπο της μεθόδου επισήμανσης, τη θέση από την αρχή του κειμένου που η επισήμανση ξεκινά και τη θέση που τερματίζεται, όπως φαίνεται και στο Σχήμα 5-5.

Επίσης, με τη διαδικασία Καθορισμού Ενοτήτων Κειμένου ο αρθρογράφος μπορεί να χωρίσει το κυρίως κείμενο σε θεματικές ενότητες. Μια ενότητα ορίζεται επιλέγοντας ο αρθρογράφος το κείμενο που επιθυμεί ως τίτλο της ενότητας και εφαρμόζοντας σε αυτό μια προκαθορισμένη από το σύστημα μέθοδο επισήμανσης. Κάθε φορά που ορίζεται μία τέτοια ενότητα δημιουργείται μια οντότητα ενότητας που προστίθεται και αυτή στον αντίστοιχο πίνακα της δομής εγγράφου. Η δομή της οντότητας ενότητας περιέχει μόνο τα όρια του τίτλου της μέσα στο κυρίως κείμενο του εγγράφου όπως φαίνεται στο Σχήμα 5-5. Ο αρθρογράφος από τη στιγμή που θα ορίσει μία ενότητα δεν μπορεί να τροποποιήσει το κείμενο του τίτλου της. Πρέπει πρώτα να αφαιρέσει την ενότητα, να τροποποιήσει το κείμενο και μετά να ορίσει ξανά την ενότητα. Κάτι τέτοιο δεν ισχύει για τη διαδικασία Επισήμανσης Φράσεων Κειμένου.

Ολοκληρώνοντας, στο Σχήμα 5-6 φαίνονται οι Μηχανισμοί Συγγραφής Άρθρου μέσα στο Υποσύστημα Συγγραφής Άρθρων και τα αποτελέσματα της εφαρμογής τους από τους αρθρογράφους.



Σχήμα 5-5 Δομές των οντοτήτων Επισήμανσης και Ενότητας



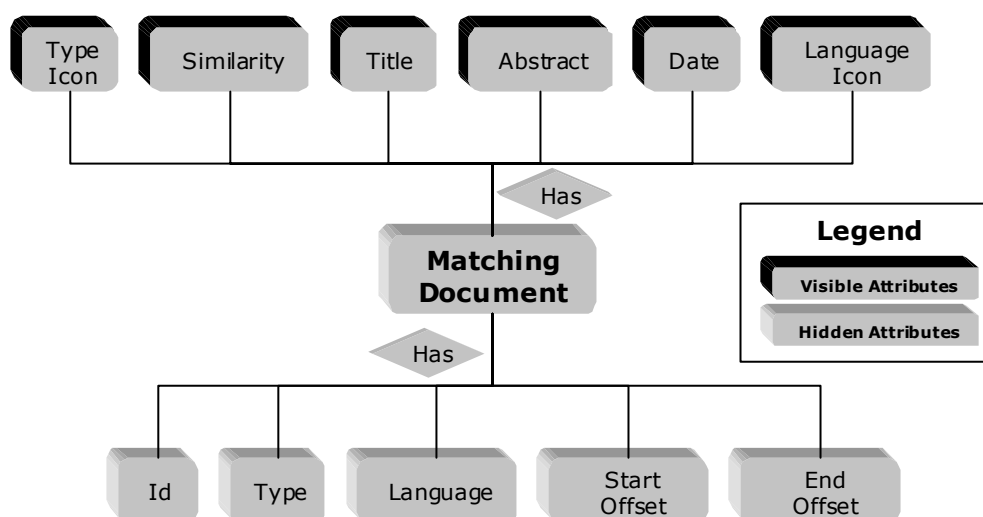
Σχήμα 5-6 Οι Μηχανισμοί Συγγραφής Άρθρου στο Υποσύστημα Συγγραφής Άρθρων

5.5 Μηχανισμός Παρουσίασης Αποτελεσμάτων Ερώτησης

Πριν προχωρήσουμε στην περιγραφή των Μηχανισμών Μορφοποίησης Άρθρου, θα περιγράψουμε το Μηχανισμό Παρουσίασης Αποτελεσμάτων Ερώτησης, γιατί επηρεάζει άμεσα τον τρόπο με τον οποίο μορφοποιείται ένα άρθρο.

Ο αρθρογράφος μπορεί να υποβάλει μία ερώτηση στο σύστημα μέσω του Υποσυστήματος Υποβολής Ερωτήσεων (βλ. και [Καρασ98]) αναζητώντας έγγραφα και πολυμέσα που υπάρχουν στο αρχείο και τον ενδιαφέρουν, αφού επιλέξει τις κατάλληλες κατηγορίες και λέξεις – κλειδιά. Το υποσύστημα κατόπιν επιστέφει μια λίστα με τους μοναδικούς κωδικούς των εγγράφων και των πολυμέσων που βρήκε ότι ταιριάζουν με τις παραμέτρους που έθεσε ο αρθρογράφος, η οποία είναι διατεταγμένη σύμφωνα με τη σχετικότητα τους. Τα αποτελέσματα αυτά πρέπει να έχουν μια συγκεκριμένη δομή ώστε να μπορούν να ενσωματωθούν στη δομή του εγγράφου ως διασυνδέσεις, συσχετίσεις και αντικείμενα εντός κειμένου.

Κατά τη διεργασία Δημιουργίας Αντικειμένων Εγγράφου, για κάθε έγγραφο, που περιέχεται στα αποτελέσματα της ερώτησης, δημιουργείται η δομή που φαίνεται στο Σχήμα 5-7. Τα χαρακτηριστικά που είναι εμφανή στον αρθρογράφο είναι το εικονίδιο που προδίδει τον τύπο του αποτελέσματος, η σχετικότητα του με τις παραμέτρους της ερώτησης, ο τίτλος του εγγράφου, η περίληψη του, η ημερομηνία έκδοσης του και το εικονίδιο που προσδιορίζει τη γλώσσα του εγγράφου.



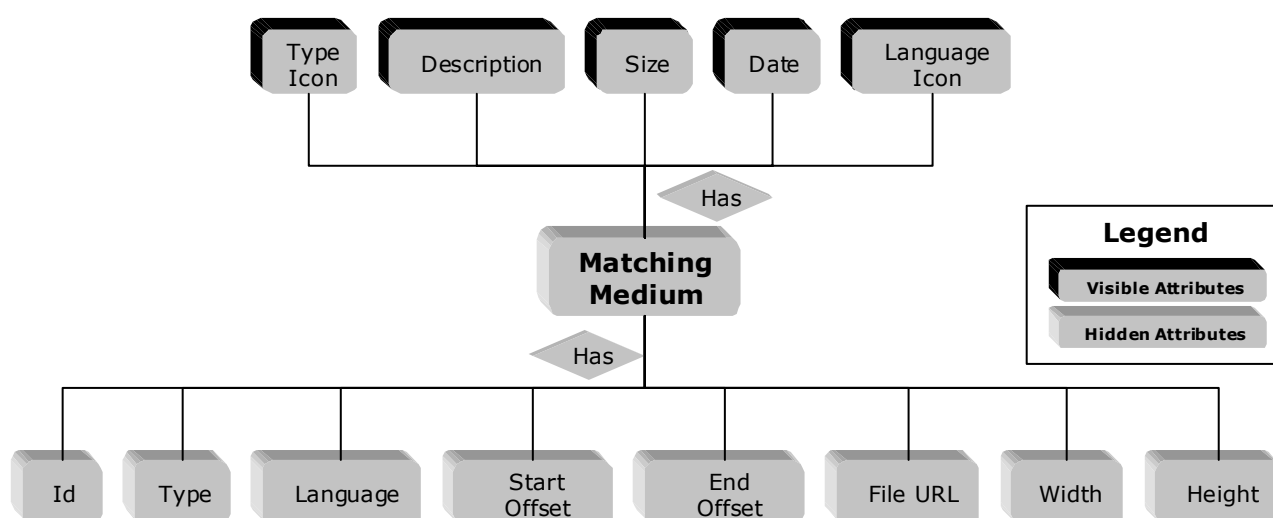
Σχήμα 5-7 Δομή Αποτελέσματος Εγγράφου

Τα χαρακτηριστικά που δε βλέπει ο αρθρογράφος, και είναι απαραίτητα για τη διαχείριση της δομής, είναι ο μοναδικός κωδικός του εγγράφου, ο τύπος του

αποτελέσματος και η γλώσσα στην οποία είναι γραμμένο. Η δομή περιέχει επίσης δύο χαρακτηριστικά (Start Offset, End Offset) που χρησιμοποιούνται σε περίπτωση που το έγγραφο ενσωματωθεί στο κείμενο του άρθρου ως διασύνδεση ή αντικείμενο εντός κειμένου.

Όπως και παραπάνω, έτσι και κατά τη διαδικασία Δημιουργίας Αντικειμένων Μέσου, για κάθε μέσο, που περιέχεται στα αποτελέσματα της ερώτησης, δημιουργείται η δομή που φαίνεται στο Σχήμα 5-8. Τα χαρακτηριστικά που είναι εμφανή στον αρθρογράφο είναι κοινά με αυτά του εγγράφου, μόνο που για τα πολυμέσα δεν υπολογίζεται σχετικότητα, και αντί του τίτλου και της περίληψης εμφανίζονται η περιγραφή και το μέγεθος του μέσου.

Τα χαρακτηριστικά που δε βλέπει ο αρθρογράφος, είναι επίσης κοινά με αυτά του εγγράφου, αλλά και μερικά ακόμα. Όπως το όνομα του αρχείου του μέσου στη δευτερεύουσα μνήμη και το πλάτος και ύψος του, αν το μέσο είναι εικόνα.



Σχήμα 5-8 Δομή Αποτελέσματος Μέσου

Τέλος, κατά τη διαδικασία Προβολής Αποτελεσμάτων τα έγγραφα και τα πολυμέσα παρουσιάζονται στον αρθρογράφο σε κατάλληλο πίνακα αποτελεσμάτων, όπου διακρίνονται όλα τα εμφανή χαρακτηριστικά τους. Αν κάποια από αυτά κρίνει ότι δεν τον ενδιαφέρουν μπορεί να τα διαγράψει από τον πίνακα αυτό των αποτελεσμάτων.

5.6 Μηχανισμός Εισαγωγής Πολυμέσων

Σε περίπτωση που ο αρθρογράφος έχει στη διάθεση του πολυμέσα τα οποία δεν βρίσκονται στο αρχείο του συστήματος, έχει τη δυνατότητα να τα προσθέσει σε αυτό μέσω του Μηχανισμού Εισαγωγής Πολυμέσων. Ο μηχανισμός αυτός δημιουργεί μία νέα δομή μέσου, όπως αυτή φαίνεται στο Σχήμα 5-4 παραπάνω, την οποία συμπληρώνουν οι τρεις διαδικασίες του μηχανισμού που περιγράφονται αναλυτικά στη συνέχεια.

Αρχικά, μέσω της διαδικασίας Επιλογής Μέσου, ο αρθρογράφος επιλέγει το αρχείο του μέσου που επιθυμεί να εισάγει στο σύστημα και το οποίο βρίσκεται στη δευτερεύουσα μνήμη του δικού του σταθμού εργασίας. Αυτή η διαδικασία ορίζει στη δομή τον τύπο του μέσου, το μέγεθος του αρχείου, το όνομα που έχει το αρχείο στη δευτερεύουσα μνήμη του σταθμού εργασίας του αρθρογράφου και το ύψος και το πλάτος του σε περίπτωση που το μέσο είναι εικόνα.

Στη συνέχεια, μέσω της διαδικασίας Εισαγωγής Χαρακτηριστικών Μέσου, ο αρθρογράφος έχει τη δυνατότητα να εισάγει την περιγραφή του μέσου σε όλες τις γλώσσες που υποστηρίζει το σύστημα. Αν δε γνωρίζει όλες τις μεταφράσεις της περιγραφής, έχει τη δυνατότητα να τις συμπληρώσει αργότερα μέσω κατάλληλης διαδικασίας. Επίσης οφείλει να κατηγοριοποιήσει το μέσο, μέσω του Υποσυστήματος Κατηγοριοποίησης, προσάπτοντάς του τις κατάλληλες λέξεις – κλειδιά και κατατάσσοντάς το στην κατάλληλη κατηγορία και υποκατηγορία του συστήματος. Η διαδικασία αυτή συμπληρώνει τα χαρακτηριστικά στη δομή του μέσου που αφορούν την περιγραφή του και την κατηγοριοποίηση του.

Τέλος, μέσω της διαδικασίας Μεταφοράς Μέσου, η δομή του μέσου με την τωρινή μορφή της αποθηκεύεται στη βάση δεδομένων. Η ενέργεια αυτή δίδει στο μέσο ημερομηνία και ώρα εισαγωγής στο σύστημα και έναν μοναδικό κωδικό, ο οποίος θα είναι και το νέο όνομα του αρχείου του μέσου, όταν αυτό μεταφερθεί στον Εξυπηρετητή Μεταφοράς Αρχείων μέσω του πρωτοκόλλου FTP. Όταν ολοκληρωθεί η μεταφορά, το όνομα του αρχείου του μέσου που έχει αποθηκευτεί στη βάση δεδομένων τροποποιείται κατάλληλα. Επιπλέον, με μία μικρή τροποποίηση, η δομή

του μέσου προστίθεται στον πίνακα αποτελεσμάτων, έτσι ώστε ο αρθρογράφος να έχει τη δυνατότητα άμεσης χρήσης του, χωρίς να χρειαστεί να το αναζητήσει στο σύστημα. Αν η μεταφορά του αρχείου διακοπεί για κάποιο λόγο, σβήνεται και η εισαγωγή που έχει γίνει στη βάση δεδομένων, και η διαδικασία εισαγωγής του μέσου στο σύστημα ακυρώνεται.

5.7 Μηχανισμοί Μορφοποίησης Άρθρου

Έχοντας αναζητήσει έγγραφα και πολυμέσα από το αρχείο του συστήματος ή έχοντας εισάγει νέα πολυμέσα, ο αρθρογράφος έχει τη δυνατότητα μορφοποίησης του άρθρου ακολουθώντας τις παρακάτω τρεις διαδικασίες των Μηχανισμών Μορφοποίησης Άρθρου.

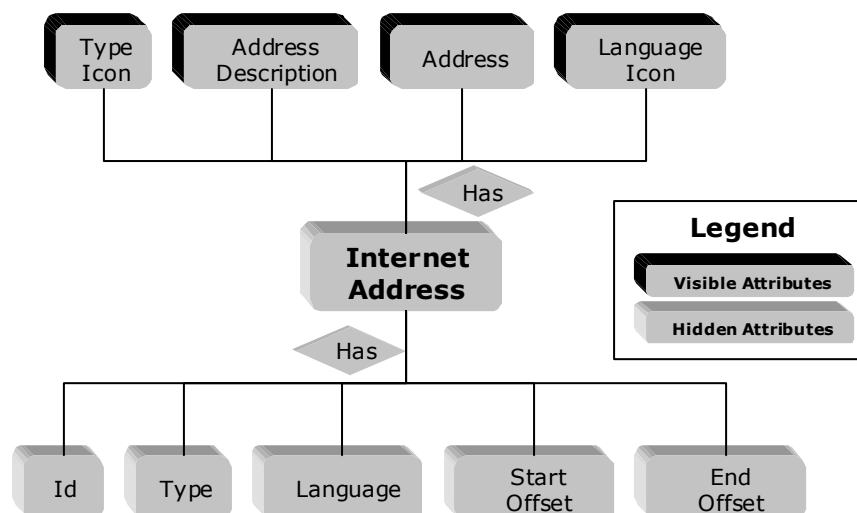
5.7.1 Εισαγωγή Διασυνδέσεων εντός Κειμένου

Μέσω της διαδικασίας Εισαγωγής Διασυνδέσεων εντός Κειμένου ο αρθρογράφος έχει τη δυνατότητα να ενσωματώσει σε λέξεις και φράσεις του κυρίως κειμένου ένα σύνολο διασυνδέσεων (hyperlinks) προς άλλα έγγραφα, πολυμέσα ή και προς μία απόλυτη Internet διεύθυνση, όπως φαίνεται και στο Σχήμα 5-3. Για να εισάγει μία διασύνδεση, ο αρθρογράφος πρέπει κατ' αρχήν να επιλέξει το τμήμα του κειμένου στο οποίο θέλει να ενσωματώσει τη διασύνδεση. Η συνέχεια εξαρτάται από το είδος του αντικειμένου προς το οποίο θέλει να δείχνει η διασύνδεση.

Αν η διασύνδεση επιθυμεί να είναι προς ένα άλλο έγγραφο του συστήματος, επιλέγει το έγγραφο αυτό από τον πίνακα αποτελεσμάτων και ενεργοποιεί τη διαδικασία Εισαγωγής Διασύνδεσης εντός Κειμένου. Η διαδικασία αυτή δημιουργεί ένα αντίγραφο της δομής του αποτελέσματος εγγράφου (Σχήμα 5-7) που έχει επιλεγεί και συμπληρώνει τα χαρακτηριστικά που δεν φαίνονται στον αρθρογράφο και αφορούν τις θέσεις του κειμένου όπου αρχίζει και τελειώνει η διασύνδεση (Start Offset, End Offset). Τέλος, το αντίγραφο της δομής προστίθεται στον αντίστοιχο πίνακα διασυνδέσεων που κρατά η δομή του εγγράφου και ο οποίος εμφανίζεται στον αρθρογράφο, ενώ το τμήμα του κειμένου στο οποίο ενσωματώθηκε η διασύνδεση επισημαίνεται κατάλληλα (Foreground Blue).

Με παρόμοιο τρόπο, αν η διασύνδεση είναι προς ένα μέσο του συστήματος, αυτό πρέπει να επιλεγεί από τον πίνακα αποτελεσμάτων και κατόπιν να ενεργοποιηθεί η διαδικασία Εισαγωγής Διασύνδεσης εντός Κειμένου, η οποία δημιουργεί ένα αντίγραφο της δομής του αποτελέσματος μέσου (Σχήμα 5-8). Πάνω σε αυτή τη δομή, εκτός από τα χαρακτηριστικά που συμπληρώνει το σύστημα και αφορούν τα όρια της διασύνδεσης μέσα στο κείμενο, ο αρθρογράφος μπορεί να τροποποιήσει και την περιγραφή που υπάρχει ήδη στη δομή του μέσου και να εισάγει μία καταλληλότερη. Το αντίγραφο της δομής αυτής προστίθεται επίσης στον αντίστοιχο πίνακα διασυνδέσεων του εγγράφου και το τμήμα του κειμένου επίσης επισημαίνεται κατάλληλα.

Μία διασύνδεση μπορεί να οριστεί, εκτός προς ένα έγγραφο ή μέσο, και προς μία απόλυτη Internet διεύθυνση. Ο αρθρογράφος σε αυτή την περίπτωση καλείται να εισάγει τη διεύθυνση μαζί με το πρωτόκολλο επικοινωνίας (π.χ. <http://www.tuc.gr/>) και μία περιγραφή για το περιεχόμενο της διεύθυνσης αυτής (π.χ. Κεντρική Σελίδα Πολυτεχνείου Κρήτης). Για κάθε τέτοια διασύνδεση δημιουργείται μία δομή απόλυτης Internet διεύθυνσης, όπως αυτή φαίνεται στο Σχήμα 5-9, και η οποία είναι παρόμοια με τις δομές αποτελέσματος των σχημάτων 5-7 και 5-8 που παρουσιάστηκαν παραπάνω. Τα ορατά χαρακτηριστικά της δομής είναι το εικονίδιο



Σχήμα 5-9 Δομή Απόλυτης Internet Διεύθυνσης

που προσδιορίζει τον τύπο της διασύνδεσης, η περιγραφή της διεύθυνσης που εισήγαγε ο αρθρογράφος, η διεύθυνση η ίδια και το εικονίδιο της γλώσσας στην οποία είναι γραμμένη η περιγραφή της διεύθυνσης. Τα υπόλοιπα χαρακτηριστικά είναι κοινά με αυτά της δομής αποτελέσματος εγγράφου. Αφού δημιουργηθεί, λοιπόν, αυτή η δομή προστίθεται στον πίνακα διασυνδέσεων του εγγράφου και το αντίστοιχο τμήμα του κειμένου επισημαίνεται κατάλληλα.

Ο πίνακας διασυνδέσεων του εγγράφου είναι παρόμοιος με τον πίνακα αποτελεσμάτων ερώτησης, αλλά επιπλέον προσφέρει στον αρθρογράφο δυνατότητες διαχείρισης των διασυνδέσεων που περιέχει, οι οποίες φαίνονται στον Πίνακα 5-1.

Hyperlink Handling
Show Hyperlink
Edit Hyperlink
Move Hyperlink
Delete Hyperlink
View Hyperlink

Πίνακας 5-1 Δυνατότητες Διαχείρισης Διασυνδέσεων Εγγράφου

Ο αρθρογράφος με την εντολή “Show Hyperlink” μπορεί να δει μία διασύνδεση του πίνακα διασυνδέσεων που ακριβώς έχει τοποθετηθεί μέσα στο κείμενο. Ιδιαίτερα χρήσιμη λειτουργικότητα αν το κείμενο του άρθρου περιέχει πολλές διασυνδέσεις. Αυτή η λειτουργικότητα όμως παρέχεται και αντίστροφα. Δηλαδή, ο αρθρογράφος έχει τη δυνατότητα να εντοπίσει στον πίνακα διασυνδέσεων μία διασύνδεση που βλέπει μέσα στο κείμενο, τοποθετώντας το δρομέα μέσα σε μία διασύνδεση και ενεργοποιώντας την αντίστροφη εντολή “Show Hyperlink”. Η εντολή “Edit Hyperlink” επιτρέπει στον αρθρογράφο να τροποποιήσει τις παραμέτρους που έθεσε κατά την εισαγωγή των διασυνδέσεων (π.χ. περιγραφή μέσου). Επιλέγοντας ένα τμήμα του κειμένου και κάνοντας χρήση της εντολής “Move Hyperlink” ο αρθρογράφος μπορεί να μετακινήσει μία διασύνδεση, χωρίς να χρειαστεί να σβήσει την παλιά και να δημιουργήσει άλλη καινούργια. Με την εντολή “Delete Hyperlink” παρέχεται η δυνατότητα διαγραφής μίας διασύνδεσης, η οποία επίσης δουλεύει και

αντίστροφα. Δηλαδή, μπορεί ο αρθρογράφος να επιλέξει τμήμα του κειμένου που περιέχει διασυνδέσεις και εφαρμόζοντας την αντίστροφη εντολή “Delete Hyperlink” να διαγράψει όλες τις διασυνδέσεις που περιέχονται στο επιλεγμένο τμήμα. Τέλος, με την εντολή “View Hyperlink” βλέπει που δείχνει η διασύνδεση σε νέο παράθυρο φυλλομετρητή (New Browser Window).

Όπως στην περίπτωση των ενοτήτων, έτσι και εδώ, από τη στιγμή που ο αρθρογράφος θα ορίσει μία διασύνδεση δεν μπορεί να τροποποιήσει το κείμενο στο οποίο έχει ενσωματωθεί. Πρέπει πρώτα να αφαιρέσει ή να μετακινήσει τη διασύνδεση, να τροποποιήσει το κείμενο και μετά να ορίσει ξανά τη διασύνδεση.

5.7.2 Εισαγωγή Αντικειμένων εντός Κειμένου

Παρόμοια με τη διαδικασία Εισαγωγής Διασυνδέσεων, μέσω της διαδικασίας Εισαγωγής Αντικειμένων εντός Κειμένου ο αρθρογράφος έχει τη δυνατότητα να τοποθετήσει σε συγκεκριμένες θέσεις του κυρίως κειμένου του άρθρου ένα σύνολο αντικειμένων (in text components). Τα αντικείμενα αυτά μπορεί να είναι έγγραφα, πολυμέσα ή και απόλυτες Internet διευθύνσεις, όπως φαίνεται επίσης και στο Σχήμα 5-3. Για να εισάγει ο αρθρογράφος ένα αντικείμενο σε μία συγκεκριμένη θέση του κειμένου, πρέπει κατ’ αρχήν να τοποθετήσει το δρομέα (cursor) στη θέση αυτή.

Αν το αντικείμενο επιθυμεί να είναι ένα άλλο έγγραφο του συστήματος, επιλέγει το έγγραφο αυτό είτε από τον πίνακα αποτελεσμάτων είτε από τον πίνακα διασυνδέσεων του άρθρου και ενεργοποιεί τη διαδικασία Εισαγωγής Αντικειμένου εντός Κειμένου. Η διαδικασία αυτή είναι παρόμοια με αυτή της Εισαγωγής Διασύνδεσης, έχοντας όμως μερικές διαφορές. Στο αντίγραφο της δομής που δημιουργείται τα χαρακτηριστικά που αφορούν τα όρια της διασύνδεσης μέσα στο κείμενο χρησιμοποιούνται διαφορετικά. Το πρώτο χαρακτηριστικό (Start Offset) κρατά τη θέση μέσα στο κείμενο στην οποία εισάγεται το αντικείμενο, ενώ το δεύτερο (End Offset) κρατά την οριζόντια ευθυγράμμιση του αντικειμένου (left or right alignment), για την οποία έχει ερωτηθεί και απαντήσει ο αρθρογράφος κατά την εισαγωγή του αντικειμένου. Τέλος, το αντίγραφο της δομής προστίθεται επίσης στον αντίστοιχο πίνακα αντικειμένων εντός κειμένου που κρατά η δομή του εγγράφου και στη θέση

του κειμένου που εισήχθηκε το αντικείμενο τοποθετείται ένα εικονίδιο που προσδιορίζει τον τύπο του αντικειμένου.

Οι παραπάνω διαφοροποιήσεις από τη διαδικασία Εισαγωγής Διασυνδέσεων ισχύουν και για την περίπτωση που το αντικείμενο που τοποθετείται μέσα στο κείμενο είναι κάποιο μέσο του συστήματος ή κάποια απόλυτη Internet διεύθυνση.

Ο πίνακας αντικειμένων εντός κειμένου του εγγράφου είναι επίσης παρόμοιος με τον πίνακα διασυνδέσεων, αλλά προσφέρει στον αρθρογράφο μία επιπλέον δυνατότητα όπως φαίνονται και στον Πίνακα 5-2 που ακολουθεί.

In Text Components Handling
Show Component
Edit Component
Move Component
Delete Component
View Component
Change Alignment

Πίνακας 5-2 Δυνατότητες Διαχείρισης Αντικειμένων εντός Κειμένου

Ο αρθρογράφος με την εντολή “Edit Component” μπορεί να αλλάξει επιπλέον των άλλων χαρακτηριστικών και την οριζόντια ευθυγράμμιση του αντικειμένου από αριστερή σε δεξιά ή αντίστροφα. Την ευθυγράμμιση μπορεί επίσης να την αλλάξει χωρίς να χρειάζεται να εμφανισθούν όλες τις παράμετροι του αντικειμένου σε κάποιο διάλογο, κάνοντας χρήση της εντολής “Change Alignment”.

Όπως ειπώθηκε και παραπάνω, κατά την εισαγωγή αντικειμένου, μέσα στο κείμενο εισάγεται μόνο το εικονίδιο που προσδίδει τον τύπο του αντικειμένου. Για να δει ο αρθρογράφος πως το περιεχόμενο του αντικειμένου εμφανίζεται μέσα στο κείμενο του εγγράφου πρέπει να προεπισκοπίσει το άρθρο. Αυτή η διαδικασία περιγράφεται σε επόμενη παράγραφο στη συνέχεια.

5.7.3 Εισαγωγή Συσχετίσεων Άρθρου

Η διαδικασία Εισαγωγής Συσχετίσεων Άρθρου είναι πολύ απλούστερη από τις προηγούμενες δύο. Μέσω αυτής ο αρθρογράφος έχει τη δυνατότητα να συσχετίσει το άρθρο που δημιουργεί με άλλα έγγραφα, πολυμέσα ή απόλυτες Internet διευθύνσεις. Η διαφορά των συσχετίσεων από τις διασυνδέσεις είναι ότι οι τελευταίες ενσωματώνονται σε λέξεις και φράσεις του κειμένου, ενώ οι συσχετίσεις παρουσιάζονται ομαδοποιημένες στο τέλος του άρθρου. Μέσω αυτών ο αρθρογράφος προσφέρει ολοκληρωμένη κάλυψη του θέματος του άρθρου του, όπως και στην περίπτωση των διασυνδέσεων.

Αν η συσχέτιση που επιθυμεί να εισάγει ο αρθρογράφος είναι ένα άλλο έγγραφο ή μέσο του συστήματος ή κάποια απόλυτη Internet διεύθυνση, απλά κάνει την επιλογή του είτε από τον πίνακα αποτελεσμάτων είτε από τον πίνακα διασυνδέσεων είτε από τον πίνακα αντικειμένων του άρθρου και ενεργοποιεί τη διαδικασία Εισαγωγής Συσχέτισης Άρθρου. Η διαδικασία αυτή δημιουργεί ένα αντίγραφο της δομής, ή δημιουργεί μία νέα αν πρόκειται για νέα Internet διεύθυνση, και το προσθέτει στον αντίστοιχο πίνακα συσχετίσεων που κρατά η δομή του εγγράφου.

Ο πίνακας συσχετίσεων του εγγράφου είναι και αυτός όμοιος με τον πίνακα διασυνδέσεων, αλλά οι δυνατότητες που χρειάζεται να παρέχει στον αρθρογράφο είναι λιγότερες όπως φαίνεται και στον Πίνακα 5-3 που ακολουθεί.

Related Information Handling
Edit Related
Delete Related
View Related

Πίνακας 5-3 Δυνατότητες Διαχείρισης Συσχετίσεων Εγγράφου

Ανακεφαλαιώνοντας, στο Σχήμα 5-10 που ακολουθεί φαίνονται οι παραπάνω Μηχανισμοί Μορφοποίησης Άρθρου μέσα στο Υποσύστημα Συγγραφής Άρθρων με τη μορφή που αυτοί παρέχονται στους αρθρογράφους.



Σχήμα 5-10 Οι Μηχανισμοί Μορφοποίησης Άρθρου στο Υποσύστημα Συγγραφής Άρθρων

5.8 Μηχανισμός Προεπισκόπησης Άρθρων

Πριν τη χρήση του Μηχανισμού Προεπισκόπησης Άρθρων ή των Μηχανισμών Διαχείρισης Άρθρων, ο αρθρογράφος μέσω του Υποσυστήματος Κατηγοριοποίησης (βλ. και [Καρασ98]) οφείλει να κατηγοριοποιήσει το άρθρο, επιλέγοντας τις κατάλληλες λέξεις – κλειδιά και την κατάλληλη κατηγορία και υποκατηγορία του συστήματος για να το χαρακτηρίσει. Έτσι συμπληρώνονται και τα χαρακτηριστικά στη δομή του εγγράφου που αφορούν την κατηγοριοποίηση του με δομές που παρέχει το παραπάνω υποσύστημα. Συγκεκριμένα, στη δομή του εγγράφου ενσωματώνονται η κατηγορία και η υποκατηγορία, στις οποίες κατατάχθηκε το άρθρο, και ένας πίνακας με δομές που αντιπροσωπεύουν τις λέξεις – κλειδιά και το βάρος της κάθε μίας έτσι όπως τα έθεσε ο αρθρογράφος.

Ο Μηχανισμός Προεπισκόπησης Άρθρων παρουσιάζει στον αρθρογράφο το άρθρο που δημιουργεί έτσι όπως θα το βλέπει ο τελικός χρήστης. Με τη διαφορά ότι το άρθρο δεν έχει ακόμα εκδοθεί και αποθηκευτεί στη βάση δεδομένων. Έτσι ο αρθρογράφος έχει τη δυνατότητα να ελέγξει τη μορφοποίηση εγγράφου και κυρίως τον τρόπο με τον οποίο εμφανίζονται τα αντικείμενα εντός του κειμένου. Ο μηχανισμός αυτός αποτελείται από τρεις διαδικασίες, η λειτουργικότητα των οποίων παρουσιάζεται στη συνέχεια.

Μέσω της διαδικασίας Παραγωγής HTML Κώδικα, από τη δομή του εγγράφου που έχει δημιουργήσει ο αρθρογράφος εξάγεται ο κατάλληλος HTML κώδικας για την παρουσίαση του άρθρου. Ο τρόπος με τον οποίο γίνεται αυτό είναι ίδιος με αυτόν που χρησιμοποιείται και κατά την παρουσίαση ενός άρθρου στον τελικό χρήστη και περιγράφεται αναλυτικά στο επόμενο κεφάλαιο.

Στη συνέχεια, μέσω της διαδικασίας Κατασκευής HTML Αρχείου, δημιουργείται στον Εξυπηρετητή Μεταφοράς Αρχείων ένα HTML αρχείο του οποίου το όνομα είναι ο κωδικός του αρθρογράφου που αναγνωρίζει το σύστημα. Σε αυτό το αρχείο μεταφέρεται, με τη χρήση του πρωτοκόλλου FTP, και αποθηκεύεται ο κώδικας που παράχθηκε με την προηγούμενη διαδικασία.

Τέλος, η διαδικασία Προβολής HTML Αρχείου ανοίγει ένα νέο παράθυρο φυλλομετρητή, στο οποίο εμφανίζεται το HTML αρχείο που μόλις δημιουργήθηκε.

5.9 Μηχανισμοί Διαχείρισης Άρθρων

Ο αρθρογράφος διαχειρίζεται τα έγγραφα του μέσω των Μηχανισμών Διαχείρισης Άρθρων, οι οποίοι αποτελούνται από τρεις κύριες διαδικασίες. Οι διαδικασίες αυτές είναι υπεύθυνες για την σωστή αποθήκευση και ανάκτηση των στοιχείων που συγκροτούν τη δομή ενός εγγράφου προς και από τον Εξυπηρετητή Βάσης Δεδομένων του συστήματος. Από τη στιγμή που ο αρθρογράφος συνδέεται επιτυχώς με το Σύστημα Παραγωγής Άρθρων έχει τη δυνατότητα να ξεκινήσει μία από τις τρεις παραπάνω διαδικασίες, κάθε μία από τις οποίες θέτει το σύστημα σε κάποια κατάσταση.

Αν ο αρθρογράφος επιλέξει τη συγγραφή ενός νέου άρθρου, τότε ενεργοποιείται η διαδικασία Νέου Άρθρου, η οποία θέτει το Σύστημα Παραγωγής Άρθρων στην αντίστοιχη κατάσταση. Η διαδικασία αυτή δημιουργεί μία νέα δομή εγγράφου, η οποία αρχικά είναι κενή και συμπληρώνεται με τους μηχανισμούς και τις διαδικασίες τους που είδαμε παραπάνω. Μετά το τέλος της δημιουργίας του νέου εγγράφου ο αρθρογράφος εκδίδει το άρθρο στο σύστημα, το οποίο το αποθηκεύει στον Εξυπηρετητή Βάσης Δεδομένων. Στη συνέχεια, η δομή του νέου εγγράφου που μόλις εκδόθηκε συνεχίζει να αναπαριστάται από το Σύστημα Παραγωγής Άρθρων και ο αρθρογράφος έχει τη δυνατότητα τροποποίησης του σύμφωνα με την παρακάτω διαδικασία.

Η διαδικασία Τροποποίησης Υπάρχοντος Άρθρου ενεργοποιείται με δύο τρόπους, θέτοντας βέβαια το Σύστημα Παραγωγής Άρθρων στην αντίστοιχη κατάσταση. Είτε ο αρθρογράφος υποβάλει μία ερώτηση στο σύστημα, εντοπίζει ένα έγγραφο που επιθυμεί να τροποποιήσει και ενεργοποιεί τη διαδικασία, είτε αυτή ενεργοποιείται αυτόματα αμέσως μετά την έκδοση ενός νέου άρθρου. Και στις δύο περιπτώσεις, η διαδικασία δημιουργεί μία δομή εγγράφου που συμπληρώνεται με τα στοιχεία που έχουν αποθηκευτεί στον Εξυπηρετητή Βάσης Δεδομένων και την οποία αναπαριστά το Σύστημα Παραγωγής Άρθρων. Επίσης δημιουργείται και ένα αντίγραφο της δομής αυτής που χρησιμεύει κατά την έκδοση της τροποποιημένης δομής του εγγράφου. Ο αρθρογράφος επεμβαίνει πάνω στο υπάρχον έγγραφο, το τροποποιεί και στη συνέχεια το εκδίδει στο σύστημα. Κατά την αποθήκευση του τροποποιημένου εγγράφου, το σύστημα βασίζεται στο αντίγραφο της δομής για να διαγράψει τα παλαιά στοιχεία στον Εξυπηρετητή Βάσης Δεδομένων και στη συνέχεια αποθηκεύει την τροποποιημένη δομή με τον ίδιο τρόπο που αποθηκεύει και ένα νέο άρθρο.

Αν ο αρθρογράφος επιθυμεί να συγγράψει ένα ήδη υπάρχον άρθρο σε διαφορετική γλώσσα, τότε υποβάλλοντας ερώτηση στο σύστημα, εντοπίζει το έγγραφο που θέλει να μεταφράσει και ενεργοποιεί τη διαδικασία Μετάφρασης Υπάρχοντος Άρθρου. Η διαδικασία αυτή δημιουργεί μία δομή εγγράφου που συμπληρώνεται με τα στοιχεία του αρχικού άρθρου και αναπαρίσταται από το Σύστημα Παραγωγής Άρθρων. Με αυτόν τον τρόπο ο αρθρογράφος έχει τη δυνατότητα να μεταφράζει σταδιακά το

άρθρο που του παρουσιάζεται. Επιπλέον, η διαδικασία κρατά τον μοναδικό κωδικό του εγγράφου που ο αρθρογράφος αρχικά επέλεξε να μεταφράσει. Κατά την έκδοση του μεταφρασμένου εγγράφου, η μόνη διαφορά από αυτή του νέου εγγράφου είναι ότι στον Εξυπηρετητή Βάσης Δεδομένων ενημερώνεται και ο αντίστοιχος πίνακας που συσχετίζει τις μεταφράσεις των εγγράφων, όπως θα δούμε και σε επόμενο κεφάλαιο.

Τέλος, στο Σχήμα 5-11 που ακολουθεί φαίνονται οι Μηχανισμοί Διαχείρισης Άρθρων μέσα στο Υποσύστημα Συγγραφής Άρθρων με τη μορφή που αυτοί παρέχονται στους αρθρογράφους.



Σχήμα 5-11 Οι Μηχανισμοί Διαχείρισης Άρθρων στο Υποσύστημα Συγγραφής Άρθρων

5.10 Ανακεφαλαίωση

Σε αυτό το κεφάλαιο παρουσιάστηκε αναλυτικά το Υποσύστημα Συγγραφής Άρθρων. Αρχικά, περιγράφηκε η δομή του εγγράφου που χρησιμοποιεί το σύστημα, η οποία περιέχει ένα υποσύνολο του μοντέλου αναπαράστασης εγγράφων της γλώσσας HTML. Αυτό φάνηκε στη συνέχεια, καθώς είδαμε πώς εφαρμόζονται στο έγγραφο κλασσικές μέθοδοι επισήμανσης και πώς ενσωματώνονται διασυνδέσεις (hyperlinks), συσχετίσεις (related information) και αντικείμενα εντός του κειμένου του (in text components).

Στη συνέχεια παρουσιάστηκαν με λεπτομέρεια όλοι οι μηχανισμοί και οι διαδικασίες του υποσυστήματος, που καλούνται να καλύψουν τις ανάγκες ενός απαιτητικού αρθρογράφου για ένα πλήρες και ολοκληρωμένο περιβάλλον συγγραφής άρθρων.

Στο επόμενο κεφάλαιο περιγράφεται το Υποσύστημα Παρουσίασης Νέων, το οποίο αποτελεί το άλλο μεγάλο τμήμα της παρούσας διατριβής. Αναλύονται οι υπηρεσίες νέων που παρέχει το σύστημα στους τελικούς χρήστες, και πώς αυτές παρουσιάζονται μέσω των μηχανισμών του υποσυστήματος.

6 ΥΠΟΣΥΣΤΗΜΑ ΠΑΡΟΥΣΙΑΣΗΣ ΝΕΩΝ

Σε αυτό το κεφάλαιο περιγράφεται αναλυτικά το Υποσύστημα Παρουσίασης Νέων που αποτελεί τμήμα του Συστήματος Υπηρεσιών Νέων, και είναι υπεύθυνο για την παρουσίαση των υπηρεσιών που παρέχει το σύστημα στους τελικούς χρήστες. Αρχικά, καταγράφονται οι υπηρεσίες νέων που παρέχει το σύστημα και οι ανάγκες που εμφανίζουν αυτές για παρουσίαση. Στη συνέχεια παρουσιάζεται αναλυτικότερα το τμήμα της αρχιτεκτονικής του συστήματος που αντιστοιχεί στο Υποσύστημα Παρουσίασης Νέων, επιδεικνύοντας τις διεργασίες που περιέχει κάθε μηχανισμός του. Τέλος, περιγράφεται λεπτομερειακά κάθε μία από τις διεργασίες αυτές και επιδεικνύεται ο τρόπος με τον οποίο παρουσιάζουν τις υπηρεσίες νέων.

6.1 Υπηρεσίες Νέων

Οι πέντε βασικές υπηρεσίες που παρέχει το σύστημα νέων “Hypermedia Custom News System” στους τελικούς χρήστες είναι οι εξής:

- *Αναζήτηση στο περιεχόμενο του συστήματος.* Ο τελικός χρήστης μπορεί να υποβάλει μία ερώτηση στο σύστημα, μέσω του Υποσυστήματος Υποβολής Ερωτήσεων, αναζητώντας έγγραφα με συγκεκριμένα χαρακτηριστικά. Το σύστημα θα του επιστρέψει μία λίστα από έγγραφα, διατεταγμένη σύμφωνα με τη σχετικότητα τους ως προς τις παραμέτρους της ερώτησης που έθεσε αρχικά ο τελικός χρήστης.
- *Πλοήγηση στο περιεχόμενο του συστήματος.* Ο τελικός χρήστης έχει τη δυνατότητα άποψης του μοντέλου κατηγοριοποίησης που ακολουθεί το σύστημα και ειδικότερα των κατηγοριών σύμφωνα με τις οποίες κατατάσσονται τα άρθρα. Επίσης, επιλέγοντας μία από τις κατηγορίες του συστήματος, μπορεί να δει τα άρθρα που αυτή περιέχει.

- *Εξατομικευμένα Νέα.* Ο τελικός χρήστης έχει τη δυνατότητα να δημιουργήσει ένα προσωπικό διάγραμμα, καταγράφοντας και αποθηκεύοντας τις προτιμήσεις του. Συγκεκριμένα, επιλέγει τις κατηγορίες, υποκατηγορίες και λέξεις – κλειδιά του συστήματος που πιστεύει ότι χαρακτηρίζουν τα άρθρα που τον ενδιαφέρουν. Στη συνέχεια, επιλέγει τη γλώσσα στην οποία θέλει να του παρουσιάζονται τα άρθρα, τον τρόπο ταξινόμησής τους (ημερομηνία, σχετικότητα) και το χρονικό διάστημα που αυτά θα καλύπτουν (τελευταίες μία, τρεις, επτά ή τριάντα ημέρες). Τέλος, εφοδιάζεται με ένα όνομα χρήστη και έναν κωδικό, με τα οποία μπορεί ο χρήστης να συνδεθεί με το σύστημα και να ενημερωθεί μόνο για τα νέα που ταιριάζουν με τις προτιμήσεις του.
- *Νέα κατ' απαίτηση.* Ο τελικός χρήστης μπορεί να προμηθευτεί από το σύστημα μέσω του διαδικτύου και να εγκαταστήσει στο σταθμό εργασίας του μία μικρή εφαρμογή, η οποία λειτουργεί ως μηχανισμός ειδοποίησης. Ειδικότερα, μέσω της εφαρμογής αυτής ο τελικός χρήστης μπορεί να συνδεθεί με το σύστημα, δίνοντας το όνομα και τον κωδικό πρόσβασης του που έχει ορίσει κατά τη δημιουργία του προσωπικού διαγράμματός του. Στη συνέχεια, μπορεί να ορίσει ένα χρονικό διάστημα σύμφωνα με το οποίο έχει τη δυνατότητα να ενημερώνεται περιοδικά από το σύστημα για την έκδοση νέων άρθρων. Τα νέα αυτά άρθρα ταιριάζουν επίσης με τις προτιμήσεις που έχει καταγράψει ο χρήστης στο διάγραμμά του.
- *Επισκόπηση Άρθρων Συστήματος.* Η πιο βασική υπηρεσία του συστήματος είναι αυτή που δίνει τη δυνατότητα στον τελικό χρήστη της επισκόπησης των άρθρων του συστήματος. Μπορούμε να πούμε ότι αυτή η υπηρεσία είναι το κατώτερο επίπεδο πλοήγησης του χρήστη στο σύστημα, αλλά το αναφέρουμε ως ξεχωριστή υπηρεσία λόγω της σημαντικότητάς της.

Οι ανάγκες παρουσίασης που εμφανίζουν οι παραπάνω υπηρεσίες επηρέασαν καθοριστικά το σχεδιασμό της αρχιτεκτονικής του Υποσυστήματος Παρουσίασης Νέων, η οποία επιδεικνύεται στην επόμενη παράγραφο.

6.2 Αρχιτεκτονική

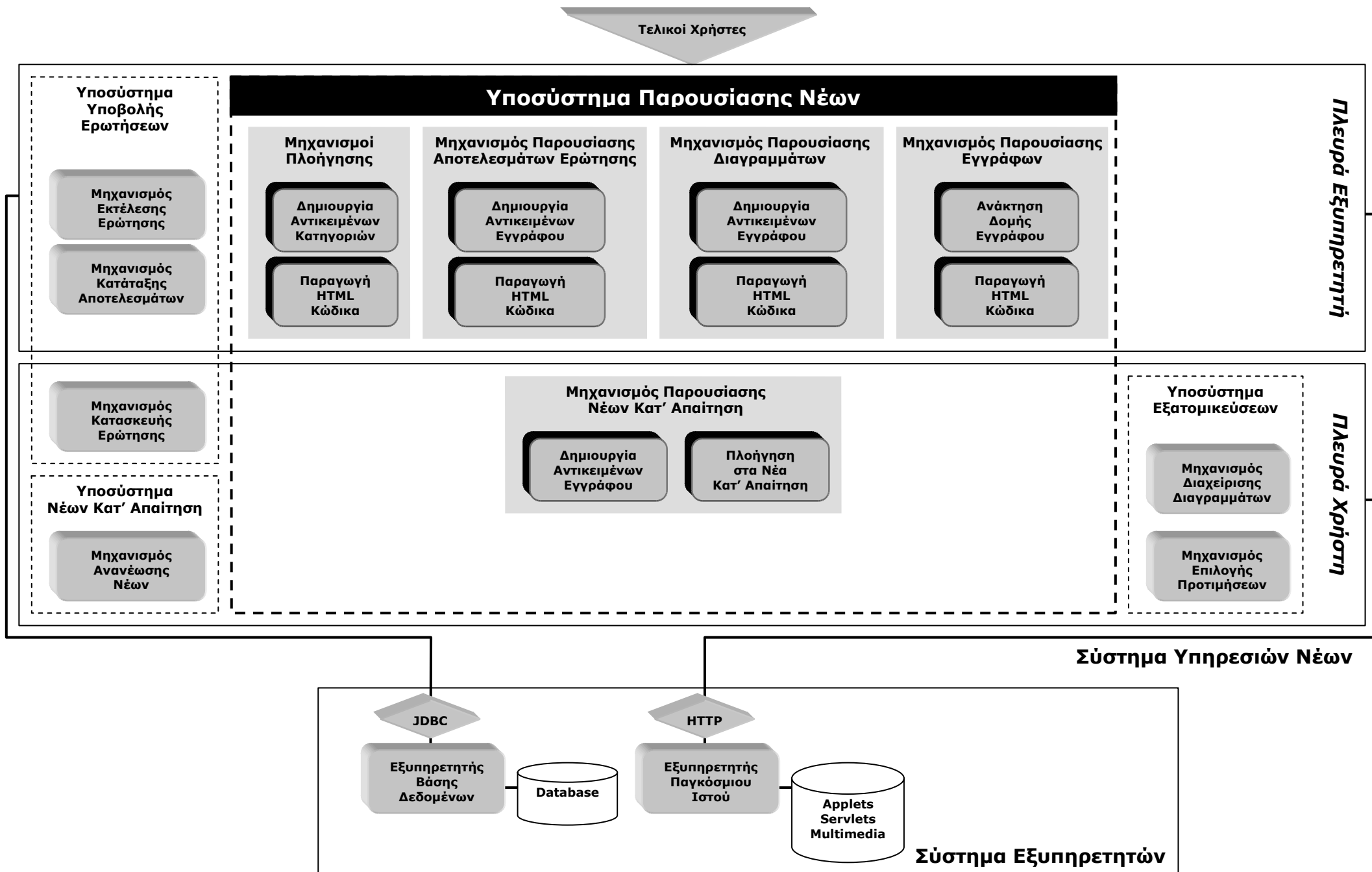
Το Σχήμα 6-1 της επόμενης σελίδας αποτελεί τμήμα του σχήματος που παρουσιάζει την αρχιτεκτονική του συστήματος και αναλύει σε βάθος το Υποσύστημα Παρουσίασης Νέων. Συγκεκριμένα φαίνονται όλες οι διεργασίες των μηχανισμών που περιέχει το συγκεκριμένο υποσύστημα, καθώς και τα υποσυστήματα με τα οποία επικοινωνεί για την παρουσίαση των υπηρεσιών του συστήματος.

Επίσης φαίνεται για κάθε μηχανισμό των υποσυστημάτων του Συστήματος Υπηρεσιών Νέων αν αυτός ενεργοποιείται στην πλευρά του τελικού χρήστη ή στην πλευρά του εξυπηρετητή. Στη συνέχεια περιγράφεται αναλυτικά η λειτουργικότητα των διεργασιών κάθε μηχανισμού και ο τρόπος με τον οποίο καλύπτουν τις ανάγκες παρουσίασης των υπηρεσιών.

6.3 Μηχανισμός Παρουσίασης Αποτελεσμάτων Ερώτησης

Ο Μηχανισμός Παρουσίασης Αποτελεσμάτων Ερώτησης είναι ο πιο σημαντικός μηχανισμός παρουσίασης, γιατί δεν εξυπηρετεί μόνο την υπηρεσία αναζήτησης του περιεχομένου του συστήματος, αλλά και τις υπηρεσίες πλοήγησης στο περιεχόμενο του συστήματος, εξατομικευμένων νέων και νέων κατ' απαίτηση.

Ο τελικός χρήστης υποβάλει μία ερώτηση στο σύστημα μέσω του Υποσυστήματος Υποβολής Ερωτήσεων (βλ. και [Καρασ98]) αναζητώντας έγγραφα που υπάρχουν στο αρχείο και τον ενδιαφέρουν, αφού επιλέξει τις κατάλληλες κατηγορίες, υποκατηγορίες και λέξεις – κλειδιά. Το υποσύστημα κατόπιν επιστέφει μια λίστα με τους μοναδικούς κωδικούς των εγγράφων που βρήκε ότι ταιριάζουν με τις παραμέτρους που έθεσε ο χρήστης, η οποία είναι διατεταγμένη σύμφωνα με τη σχετικότητα τους. Τα αποτελέσματα αυτά πρέπει να παρουσιαστούν στον τελικό χρήστη με μια συγκεκριμένη δομή ώστε να του δίνουν τη μέγιστη δυνατή πληροφορία στον ελάχιστο δυνατό χώρο.



Σύστημα Υπηρεσιών Νέων

JDBC

Εξυπηρετητής Βάσης Δεδομένων

Database

HTTP

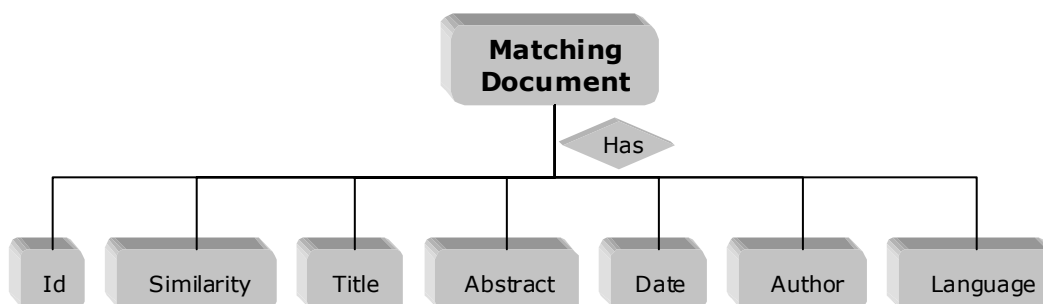
Εξυπηρετητής Παγκόσμιου Ιστού

Applets
Servlets
Multimedia

Σύστημα Εξυπηρετητών

Σχήμα 6-1 Αρχιτεκτονική του Υποσυστήματος Παρουσίασης Νέων του Συστήματος Υπηρεσιών Νέων

Την παραπάνω ανάγκη έρχεται να καλύψει ο Μηχανισμός Παρουσίασης Αποτελεσμάτων Ερώτησης. Αρχικά, ενεργοποιείται στην πλευρά του εξυπηρετητή η διεργασία Δημιουργίας Αντικειμένων Εγγράφου, η οποία για κάθε έγγραφο που περιέχεται στα αποτελέσματα της ερώτησης, δημιουργεί μία δομή έτσι όπως αυτή φαίνεται στο Σχήμα 6-2. Τα χαρακτηριστικά των δομών αυτών συμπληρώνονται ανακτώντας τα κατάλληλα δεδομένα από τον Εξυπηρετητή Βάσης Δεδομένων και συγκεκριμένα το μοναδικό κωδικό του κάθε εγγράφου, τη σχετικότητα του με τις παραμέτρους της ερώτησης, τον τίτλο του, την περίληψή του, την ημερομηνία έκδοσης του, τον αρθρογράφο που το δημιούργησε και τη γλώσσα στην οποία είναι γραμμένο. Όλα αυτά τα χαρακτηριστικά είναι εμφανή στον τελικό χρήστη πλην του μοναδικού κωδικού του εγγράφου, ο οποίος χρησιμεύει στη δημιουργία διασύνδεσης προς την πλήρη παρουσίαση του εγγράφου.



Σχήμα 6-2 Δομή Αποτελέσματος Εγγράφου

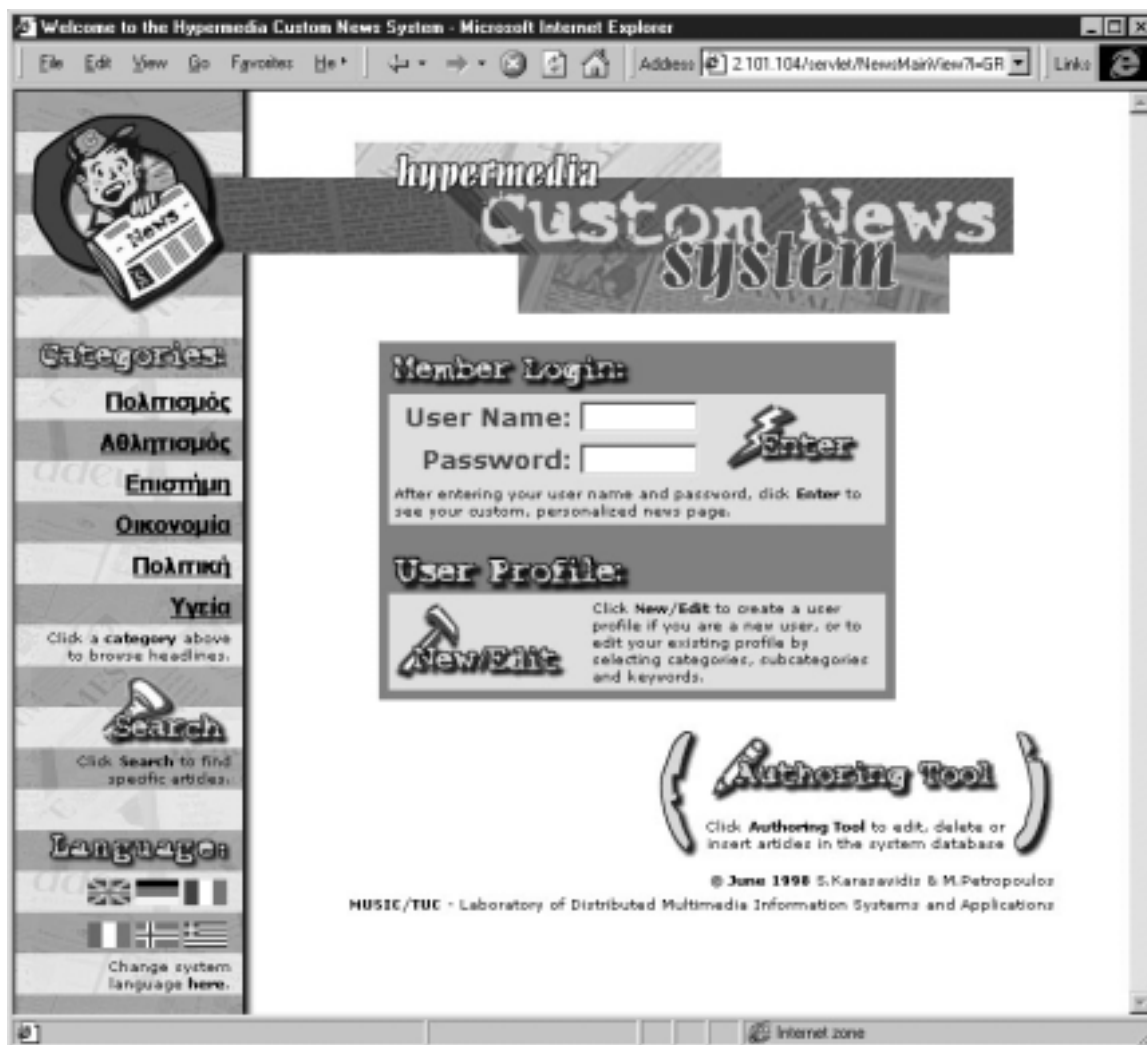
Στη συνέχεια, κατά τη διαδικασία Παραγωγής HTML Κώδικα, που ενεργοποιείται επίσης στην πλευρά του εξυπηρετητή, οι παραπάνω δομές συμπληρώνουν ένα HTML πρότυπο, το οποίο έχει σχεδιαστεί κατάλληλα για τις ανάγκες παρουσίασης αποτελεσμάτων ερώτησης. Συγκεκριμένα, για κάθε έγγραφο εμφανίζεται ο τίτλος του, στον οποίο έχει ενσωματωθεί διασύνδεση προς την πλήρη παρουσίασή του, η περίληψή του και σε μία γραμμή όλα τα υπόλοιπα εμφανή χαρακτηριστικά του. Ο τελικός κώδικας που προκύπτει αποστέλλεται στο φυλλομετρητή του τελικού χρήστη και έτσι παρουσιάζονται σε αυτόν τα αποτελέσματα της ερώτησης που υπέβαλε στο σύστημα.

6.4 Μηχανισμοί Πλοήγησης

Όταν ένας τελικός χρήστης συνδέεται με το σύστημα, στην πλευρά του εξυπηρετητή ενεργοποιείται ο Μηχανισμός Πλοήγησης του συστήματος. Αρχικά, μέσω της διαδικασίας Δημιουργίας Αντικειμένων Κατηγοριών, ο μηχανισμός αυτός ανακτά από τον Εξυπηρετητή Βάσης Δεδομένων τις κατηγορίες του συστήματος και για κάθε μία από αυτές δημιουργεί μία δομή που περιέχει τον μοναδικό κωδικό της κατηγορίας και το όνομα αυτής. Στη συνέχεια, η διαδικασία Παραγωγής HTML Κώδικα ενσωματώνει τις δομές αυτές σε κατάλληλο HTML πρότυπο, όπου εμφανίζεται το όνομα κάθε κατηγορίας, στο οποίο έχει ενσωματωθεί μία διασύνδεση προς το περιεχόμενο της. Ο τελικός κώδικας αποστέλλεται στο φυλλομετρητή του τελικού χρήστη, ο οποίος έχει έτσι τη δυνατότητα άποψης του ανώτερου επιπέδου του μοντέλου κατηγοριοποίησης που ακολουθεί το σύστημα. Ο μηχανισμός αυτός ενεργοποιείται συμπληρωματικά σε όλες τις σελίδες του συστήματος, έτσι ώστε από κάθε σημείο του να είναι δυνατή η πλοήγηση.

Στη συνέχεια, ο τελικός χρήστης μπορεί να ακολουθήσει τις διασυνδέσεις που έχουν ενσωματωθεί στα ονόματα των κατηγοριών και να επισκοπήσει τα άρθρα που έχουν καταταχθεί σε κάθε μία από αυτές. Αυτή η περίπτωση αντιμετωπίζεται από το σύστημα σαν μία προκαθορισμένη ερώτηση που υποβάλει ο χρήστης, οπότε ενεργοποιείται το Υποσύστημα Υποβολής Ερωτήσεων και στη συνέχεια ο Μηχανισμός Παρουσίασης Αποτελεσμάτων Ερώτησης. Ακολουθούνται όλες οι διαδικασίες που αναλύθηκαν στην προηγούμενη παράγραφο, με τη μόνη διαφορά ότι για τα έγγραφα που προκύπτουν δεν υπολογίζεται η σχετικότητα τους με τις παραμέτρους της ερώτησης, οπότε αυτά διατάσσονται σύμφωνα με την ημερομηνία έκδοσής τους.

Ολοκληρώνοντας, στο Σχήμα 6-3 που ακολουθεί φαίνεται η αρχική σελίδα του συστήματος “Hypermedia Custom News System” με τη μορφή που αυτή εμφανίζεται στους τελικούς χρήστες, όπου διακρίνονται οι Μηχανισμοί Πλοήγησης του Υποσυστήματος Παρουσίασης Νέων που παρέχονται σε αυτούς.



Σχήμα 6-3 Η αρχική σελίδα του συστήματος “Hypermedia Custom News System”

6.5 Μηχανισμός Παρουσίασης Διαγραμμάτων

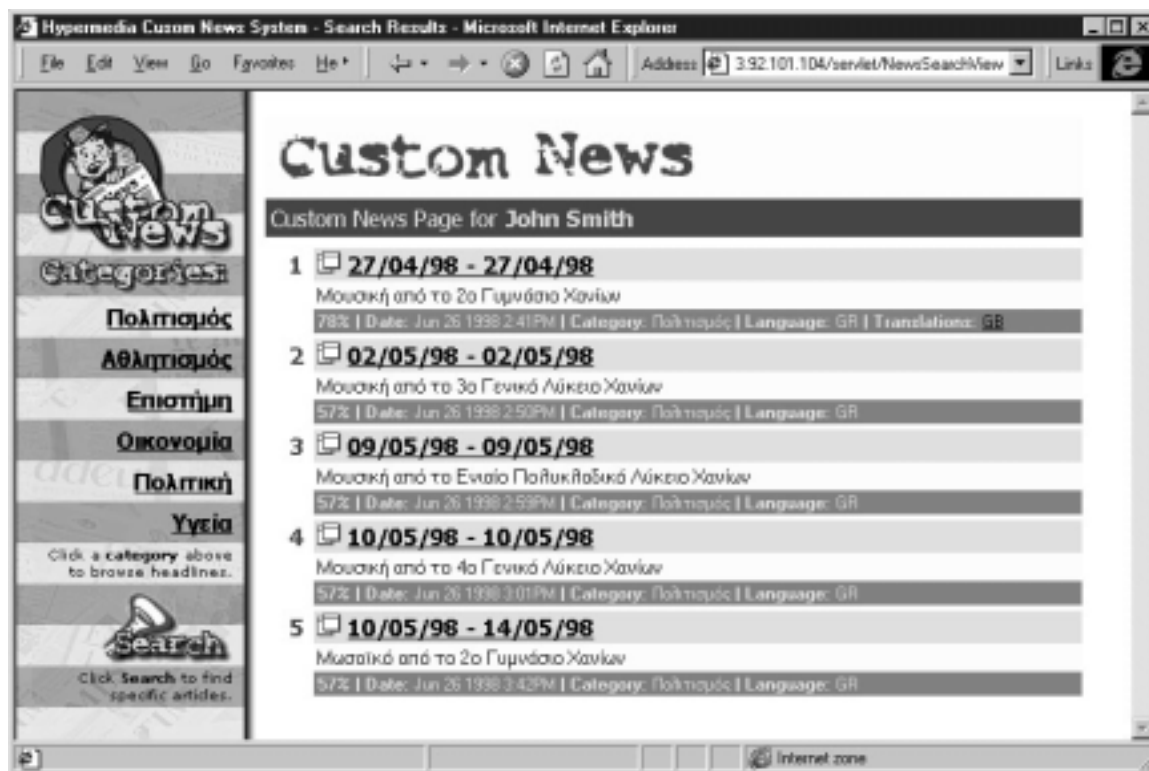
Ο Μηχανισμός Παρουσίασης Διαγραμμάτων ενεργοποιείται όταν ο τελικός χρήστης συνδεθεί με το σύστημα και εισάγει το όνομα και τον κωδικό που όρισε κατά τη δημιουργία του προσωπικού του διαγράμματος, με σκοπό την επισκόπηση των νέων που ταιριάζουν στις προτιμήσεις που έχει αποθηκεύσει σε αυτό. Με τον τρόπο αυτό εντοπίζεται το διάγραμμά του στον Εξυπηρετητή Βάσης Δεδομένων, απ’ όπου ανακτώνται και οι προτιμήσεις που έχει εκφράσει, βάσει των οποίων θα επιλεγούν τα άρθρα που θα του παρουσιάσει το σύστημα.

Και αυτή η περίπτωση αντιμετωπίζεται από το σύστημα σαν μία προκαθορισμένη ερώτηση, γιατί με το να εκφράσει ο τελικός χρήστης τις προσωπικές του προτιμήσεις κατά τη δημιουργία του προσωπικού του διαγράμματος, στην ουσία κατασκευάζει μία ερώτηση, η οποία εκτελείται κάθε φορά που αυτός συνδέεται στο σύστημα χρησιμοποιώντας τα αναγνωριστικά του, χωρίς να χρειάζεται να σχηματίσει την

ερώτηση αυτή ξανά. Οπότε ενεργοποιείται και πάλι ο Μηχανισμός Παρουσίασης Αποτελεσμάτων Ερώτησης και ακολουθούνται όλες οι διαδικασίες που αναλύθηκαν στην αντίστοιχη παράγραφο.

Κατά τη δημιουργία του διαγράμματός του ο τελικός χρήστης μπορεί να ορίσει και το χρονικό διάστημα που θα καλύπτουν τα νέα που θα του παρουσιάζονται, το οποίο μπορεί να είναι οι τελευταίες τριάντα, επτά, τρεις ή μία ημέρες. Επιλέγει επίσης τον τρόπο με τον οποίο αυτά θα διατάσσονται, διαλέγοντας μεταξύ της σχετικότητας αυτών με τις παραμέτρους της ερώτησης και της ημερομηνίας έκδοσής τους. Τέλος, αν ο τελικός χρήστης επισκοπήσει πλήρως ένα έγγραφο που περιέχεται στο διάγραμμά του, τότε αυτό προστίθεται στη λίστα των εγγράφων που δεν θα του παρουσιαστούν ξανά.

Τέλος, στο Σχήμα 6-4 που ακολουθεί φαίνεται το αποτέλεσμα του Μηχανισμού Παρουσίασης Διαγραμμάτων του Υποσυστήματος Παρουσίασης Νέων με τη μορφή που αυτό εμφανίζεται στους τελικούς χρήστες.



Σχήμα 6-4 Ο Μηχανισμός Παρουσίασης Διαγραμμάτων του Υποσυστήματος Παρουσίασης Νέων

6.6 Μηχανισμός Παρουσίασης Νέων Κατ' Απαίτηση

Ο Μηχανισμός Παρουσίασης Νέων Κατ' Απαίτηση ενεργοποιείται περιοδικά στην εφαρμογή ειδοποίησης που εγκαθιστά ο τελικός χρήστης στο σταθμό εργασίας του. Η περίοδος αυτή εξαρτάται από το χρονικό διάστημα ειδοποίησης για την έκδοση νέων άρθρων που αυτός έχει ορίσει μέσω της εφαρμογής.

Κάθε φορά που παρέρχεται το τέλος μίας περιόδου, η εφαρμογή ειδοποίησης υποβάλλει στο σύστημα την ερώτηση που ο τελικός χρήστης έχει αποθηκεύσει στο προσωπικό του διάγραμμα, προσθέτοντας όμως σε αυτή ακόμα μία παράμετρο. Η παράμετρος αυτή είναι η τελευταία ημερομηνία και ώρα που ο συγκεκριμένος χρήστης ενημερώθηκε από την εφαρμογή ειδοποίησης για την έκδοση νέων άρθρων και η οποία είναι επίσης αποθηκευμένη στο προσωπικό διάγραμμά του. Τα έγγραφα που προκύπτουν είναι διατεταγμένα σύμφωνα με την ημερομηνία και ώρα έκδοσής τους και για κάθε ένα από αυτά η διαδικασία Δημιουργίας Αντικειμένων Εγγράφου κατασκευάζει και συμπληρώνει μία δομή που περιέχει μόνο τον μοναδικό κωδικό του εγγράφου, τον τίτλο του και την περίληψή του.

Στη συνέχεια, η διαδικασία Πλοήγησης στα Νέα Κατ' Απαίτηση επιτρέπει στον τελικό χρήστη να διασχίσει τη λίστα των νέων άρθρων, βλέποντας για κάθε ένα από αυτά τον τίτλο, την περίληψή του και μία διασύνδεση προς την πλήρη παρουσίασή του. Παρέχεται επίσης και η δυνατότητα διαγραφής ενός ή όλων των νέων άρθρων από τη σχετική λίστα, καθώς τα νέα άρθρα που προκύπτουν με το πέρας της επόμενης περιόδου προσθέτονται στα ήδη υπάρχοντα χωρίς να τα διαγράφουν. Η λίστα των νέων άρθρων διαγράφεται επίσης αν ο χρήστης κλείσει την εφαρμογή ειδοποίησης. Τα έγγραφα τώρα που διαγράφονται από τη λίστα νέων άρθρων δε σημαίνει ότι ο τελικός χρήστης δε θα τα δει την επόμενη φορά που θα επισκοπήσει το διάγραμμά του μέσα από το σύστημα. Αυτά απλά διαγράφονται από τη λίστα των άρθρων για τα οποία πρέπει αυτός να ειδοποιηθεί. Από την άλλη όμως, αν ο χρήστης ακολουθήσει τη διασύνδεση που συνοδεύει ένα έγγραφο και επισκοπήσει την πλήρη παρουσίασή του, τότε το έγγραφο αυτό δε θα είναι μέσα στα αποτελέσματα του διαγράμματός του.

6.7 Μηχανισμός Παρουσίασης Εγγράφων

Ο Μηχανισμός Παρουσίασης Εγγράφων υλοποιεί την πιο βασική υπηρεσία του συστήματος, αυτή της πλήρους παρουσίασης ενός εγγράφου προς τον τελικό χρήστη. Ενεργοποιείται από όλους τους μηχανισμούς που είδαμε παραπάνω και περιέχει δύο διαδικασίες για τους σκοπούς του.

Αρχικά, η διαδικασία Ανάκτησης Δομής Εγγράφου δημιουργεί μία δομή εγγράφου όμοια με αυτή που φαίνεται στο Σχήμα 5-1 του προηγούμενου κεφαλαίου, η οποία συμπληρώνεται με τα στοιχεία που έχουν αποθηκευτεί στον Εξυπηρετητή Βάσης Δεδομένων. Αυτή περιέχει όλη την πληροφορία που πλαισιώνει ένα έγγραφο και είναι χρήσιμη για τον τελικό χρήστη.

Στη συνέχεια, η διαδικασία Παραγωγής HTML Κώδικα εξάγει το περιεχόμενο της ιστοσελίδας που θα παρουσιαστεί στον τελικό χρήστη βάσει ενός HTML προτύπου. Εκτός από τα απλά χαρακτηριστικά του εγγράφου, στο HTML πρότυπο πρέπει να τοποθετηθούν και οι οντότητες που ενσωματώνονται στο έγγραφο, οι οποίες φαίνονται στο Σχήμα 5-2 του προηγούμενου κεφαλαίου. Εκεί είχαμε αναφέρει ότι το πεδίο των χαρακτηριστικών μορφοποίησης (formatting attributes) κρατά πληροφορίες για τη θέση μέσα στο κυρίως κείμενο του εγγράφου των οντοτήτων ενότητας, επισημάνσης, διασύνδεσης, καθώς και για τη θέση των οντοτήτων που παρουσιάζουν το περιεχόμενο τους εντός κειμένου. Οι θέσεις στο πεδίο αυτό διασταυρώνονται με τις θέσεις που έχουν αποθηκευτεί μαζί με κάθε ενσωματωμένη οντότητα κατά την έκδοση του άρθρου και έτσι προσδιορίζεται η θέση της κάθε οντότητας στην τελική παρουσίαση του εγγράφου.

Τέλος, στο Σχήμα 6-5 που ακολουθεί φαίνεται το αποτέλεσμα του Μηχανισμού Παρουσίασης Εγγράφων με τη μορφή που αυτό εμφανίζεται στους τελικούς χρήστες, όπου διακρίνονται όλα τα στοιχεία της δομής του εγγράφου.

Hypermedia Custom News System - News - Microsoft Internet Explorer

File Edit View Go Favorites Help Address http://193.92.101.104/service/News Links



Categories:

- Cultural
- Sports
- Science
- Economy
- Politics
- Health

Click a category above to browse headlines.

Search

Click Search to find specific articles.

Yeltsin under fire as Russia's financial crisis escalates

Article Info

Category: Russia

Author: Michael Petropoulos

Date: Aug 28 1998 4:36AM

(Greece)

Language: 


abstract: Central Bank halts dollar sales through weekend

Pressure mounted on President Boris Yeltsin to resign Thursday as parliament leaders proposed Soviet-style measures to contain Russia's escalating financial crisis, including currency controls, fixed prices and state ownership. Yeltsin spokesman Sergei Yastrzhembsky issued a vehement denial of resignation rumors, calling them "absolutely groundless, idle inventions." But the speculation illustrated the growing frustration that for all his bluster and bravado, Yeltsin has so far failed to contain the crisis. To that end, it was clear that the Kremlin is now willing to consider solutions previously rejected -- steps the West is likely to see as moving backward from the free market.



Customers line up outside a Moscow bank

Price controls possible



Inflation is causing price increases

Alexander Shokhin, the head of a parliament faction close to Chernomyrdin, said the government may consider suspending the ruble's convertibility and instituting price controls to cope with the crisis. Shokhin also said some major commercial banks may be nationalized. "Now we can't rely on liberal monetary approaches," he said. The financial crisis so far has had little effect on ordinary Russians, who don't own stocks or bonds or mutual funds. The anxiety some are beginning to feel comes largely from fear of inflation, which is expected to rise in the wake of the government's decision to devalue the ruble. "Prices on most things haven't risen much yet, but we are buying more just in case," said Galina Filatova, 51, who runs a small candy shop.

Related Information

Documents

- Public worries in rural Russia - Aug 28 1998 4:18AM

Media

- Inflation is causing price increases - Aug 28 1998 3:39AM - 20KB
- Russia's Economy - Aug 28 1998 4:03AM - 81KB
- Yeltsin's Message - Aug 28 1998 4:00AM - 23KB

Sites

- Russian News Agency

Internet zone

Σχήμα 6-5 Ο Μηχανισμός Παρουσίασης Εγγράφων του Υποσυστήματος Παρουσίασης Νέων

6.8 Ανακεφαλαίωση

Σε αυτό το κεφάλαιο παρουσιάστηκε αναλυτικά το Υποσύστημα Παρουσίασης Νέων. Αρχικά, περιγράφηκαν οι υπηρεσίες νέων που παρέχει το σύστημα “Hypermedia Custom News System”, και οι οποίες ευελπιστούν να καλύψουν τις ανάγκες των τελικών χρηστών. Στη συνέχεια παρουσιάστηκαν με λεπτομέρεια όλοι οι μηχανισμοί και οι διαδικασίες του υποσυστήματος, που καλούνται να υλοποιήσουν τις υπηρεσίες αυτές. Επίσης, φάνηκε πως η δομή του εγγράφου, που χρησιμοποιεί το σύστημα και περιέχει ένα υποσύνολο του μοντέλου αναπαράστασης εγγράφων της γλώσσας HTML, περνά από το Υποσύστημα Συγγραφής Άρθρων και παρουσιάζεται στον τελικό χρήστη σε μορφή απλής ιστοσελίδας.

Στο επόμενο κεφάλαιο περιγράφεται ο τρόπος με τον οποίο το σύστημα υποστηρίζει πολλαπλές γλώσσες, και είναι το χαρακτηριστικό το οποίο αποτελεί ίσως τη μεγαλύτερη καινοτομία του συστήματος νέων “Hypermedia Custom News System”.

7 ΥΠΟΣΤΗΡΙΞΗ ΠΟΛΛΑΠΛΩΝ ΓΛΩΣΣΩΝ

Σε αυτό το κεφάλαιο παρουσιάζεται η υποστήριξη πολλαπλών γλωσσών από το σύστημα. Αρχικά, καταγράφονται τα προβλήματα που προέκυψαν στην προσπάθεια υλοποίησης του χαρακτηριστικού αυτού και στη συνέχεια αναλύονται οι μέθοδοι που ακολουθήθηκαν για την αντιμετώπισή τους. Επίσης, περιγράφεται η εφαρμογή του χαρακτηριστικού αυτού στο Υποσύστημα Συγγραφής Άρθρων και στο Υποσύστημα Παρουσίασης Νέων.

7.1 Εισαγωγή

Για το σύστημα η υποστήριξη πολλαπλών γλωσσών μεταφράζεται κατά κύριο λόγο ως η δυνατότητα συγγραφής και παρουσίασης άρθρων σε παραπάνω από μία γλώσσες. Κάτι τέτοιο βέβαια συνεπάγεται και τη δυνατότητα αποθήκευσης του μοντέλου κατηγοριοποίησης, καθώς και της περιγραφής των μέσων στον Εξυπηρετητή Βάσης Δεδομένων επίσης σε παραπάνω από μία γλώσσες.

Πιο συγκεκριμένα, ως στόχος του συστήματος τέθηκε η υποστήριξη περιεχομένου σε πολλαπλές γλώσσες με τέτοιο τρόπο ώστε οι μηχανισμοί που το αποτελούν να είναι ικανοί να αλληλεπιδράσουν με τον Εξυπηρετητή Βάσης Δεδομένων, χωρίς να γνωρίζουν τον ακριβή αριθμό των γλωσσών που υποστηρίζει το σύστημα, ούτε να διαθέτουν λεπτομέρειες για αυτές. Το γεγονός αυτό οδηγεί σε πλήρη ανεξαρτησία των μηχανισμών από τις υποστηριζόμενες γλώσσες, αυξάνοντας έτσι στο μέγιστο την ευελιξία του συστήματος ως προς την προσαρμογή του στην προσθαφαίρεση γλωσσών.

Τα προβλήματα που παρουσιάστηκαν στην προσπάθεια υποστήριξης πολλαπλών γλωσσών εντοπίζονται σε τέσσερα διαφορετικά επίπεδα:

1. Στο σχεδιασμό του σχεσιακού μοντέλου βάσης δεδομένων του συστήματος.

2. Στον τρόπο με τον οποίο ο Εξυπηρετητής Βάσης Δεδομένων χειρίζεται δεδομένα κειμένου σε συνάρτηση με το λειτουργικό σύστημα στο οποίο τρέχει.
3. Στους μηχανισμούς που αποθηκεύουν και ανακτούν δεδομένα κειμένου προς και από τον Εξυπηρετητή Βάσης Δεδομένων.
4. Στην τελική παρουσίαση των εγγράφων σε συνάρτηση με το λειτουργικό σύστημα και το φυλλομετρητή που χρησιμοποιεί ο τελικός χρήστης.

Κάθε ένα από τα παραπάνω ζητήματα απαιτεί ξεχωριστή αντιμετώπιση ώστε να ελαχιστοποιηθούν οι εξαρτήσεις από τις ιδιαιτερότητες των εξυπηρετητών, των τελικών χρηστών και των λειτουργικών συστημάτων.

7.2 Σχισιακό Μοντέλο Βάσης Δεδομένων

Στην παράγραφο αυτή αναλύεται η επίδραση της υποστήριξης πολλαπλών γλωσσών στο σχεδιασμό του σχεσιακού μοντέλου βάσης δεδομένων του συστήματος νέων HyNoDe και του συστήματος “Hypermedia Custom News System”.

7.2.1 Το Σύστημα Νέων HyNoDe

Το σύστημα νέων που αναπτύχθηκε στα πλαίσια του προγράμματος HyNoDe είχε ως προδιαγραφή την υποστήριξη περισσότερων από μία γλωσσών. Η αρχική επιλογή εστιάστηκε στην υποστήριξη ελληνικών, αγγλικών και ιταλικών εγγράφων, γεγονός που αποτυπώθηκε και στο σχεδιασμό της Βάσης Δεδομένων του HyNoDe. Έτσι οι πίνακες που περιγράφουν τις λέξεις κλειδιά και τα έγγραφα στο HyNoDe έχουν τα χαρακτηριστικά που φαίνονται παρακάτω.

Πεδίο	Τύπος
KeyInBasicLanguage	Varchar
ParentKeyword	Varchar
English	Varchar
Greek	Varchar
Italian	Varchar

Πίνακας 7-1 Τμήμα πίνακα αναπαράστασης λέξεων – κλειδιών στη βάση δεδομένων του συστήματος HyNoDe

Πεδίο	Τύπος
StoryIdentifier	VarChar
Title	VarChar
Abstract	VarChar
Category	VarChar
SubCategory	VarChar
Language	VarChar

Πίνακας 7-2 Τμήμα πίνακα αναπαράστασης εγγράφων στη βάση δεδομένων του συστήματος HyNoDe

Όπως φαίνεται στον Πίνακα 7-1, για τις λέξεις – κλειδιά υπάρχει ένα πεδίο με τη μετάφραση της λέξης για κάθε γλώσσα που υποστηρίζει το σύστημα. Όμοια στον Πίνακα 7-2, η γλώσσα του εγγράφου αποθηκεύεται στο πεδίο “Language”, και οι πιθανές τιμές είναι είτε “English”, είτε “Italian” είτε “Greek”.

Το πρόβλημα που προκύπτει με την παραπάνω προσέγγιση είναι ότι οι μηχανισμοί που έχουν πρόσβαση στη βάση δεδομένων πρέπει να γνωρίζουν τον αριθμό των γλωσσών που υποστηρίζει το σύστημα, καθώς και την ακριβή ονομασία των πεδίων των πινάκων ώστε να μπορούν να προσπελάσουν τα δεδομένα. Επίσης, για να προστεθεί μία επιπλέον γλώσσα, θα πρέπει όλες οι μηχανισμοί που χρησιμοποιούν τη βάση δεδομένων να ειδοποιηθούν για την ύπαρξη της επιπλέον γλώσσας, καθώς και για την ακριβή ονομασία των διάφορων πεδίων στους πίνακες της βάσης. Τέλος, τα έγγραφα είναι ανεξάρτητα μεταξύ τους και δεν υπάρχει η έννοια της μετάφρασης ενός εγγράφου από μία υποστηριζόμενη γλώσσα σε κάποια άλλη.

7.2.2 Το Σύστημα Νέων “Hypermedia Custom News System”

Τα προβλήματα που παρουσιάζονται για την υποστήριξη πολλαπλών γλωσσών στο επίπεδο της σχεδίασης της βάσης δεδομένων του συστήματος HyNoDe αντιμετωπίστηκαν στα πλαίσια ανάπτυξης τους παρόντος συστήματος νέων.

7.2.2.1 Γενικό μοντέλο για την υποστήριξη πολλαπλών γλωσσών

Για να ανεξαρτητοποιηθεί η ονοματολογία των πεδίων των πινάκων της βάσης δεδομένων από το πλήθος των γλωσσών που υποστηρίζει το σύστημα, προστέθηκε

ένας επιπλέον πίνακας. Ο πίνακας αυτός ονομάζεται “Languages”, όπως φαίνεται και στον Πίνακα 7-3, και περιγράφει τις γλώσσες που υποστηρίζονται από το σύστημα. Για κάθε γλώσσα περιέχει εκτός από το πλήρες όνομά της, και ένα ακρωνύμιο το οποίο χρησιμοποιείται στην ονομασία των πεδίων των υπολοίπων πινάκων, όπου απαιτείται η ύπαρξη πεδίων για κάθε υποστηριζόμενη γλώσσα (π.χ. όνομα κατηγορίας). Με τον τρόπο αυτό οι μηχανισμοί που προσπελαίνουν τη βάση δεδομένων γνωρίζουν τον αριθμό των υποστηριζόμενων γλωσσών καθώς και την ακριβή ονομασία των πολυγλωσσικών πεδίων των υπόλοιπων πινάκων. Το πεδίο “Rank” χρησιμοποιείται για τη διατεταγμένη παρουσίαση των υποστηριζόμενων γλωσσών στις εφαρμογές, σύμφωνα με τις εκτιμήσεις που έγιναν όσον αφορά τους τελικούς χρήστες του συστήματος. Τα πεδία “ie_encoding” και “nc_encoding” θα περιγραφούν σε επόμενη παράγραφο.

Languages	Περιγραφή
Id	ακρωνύμιο γλώσσας
Name	πλήρες όνομα
Rank	σειρά προτεραιότητας εμφάνισης
ie_encoding	κωδικοποίηση για IE
nc_encoding	κωδικοποίησης για NC

Πίνακας 7-3 Πίνακας αναπαράστασης υποστηριζόμενων γλωσσών στη βάση δεδομένων του “Hypermedia Custom News System”

Το σύστημα αυτή τη στιγμή υποστηρίζει έξι γλώσσες, και τα περιεχόμενα του πίνακα “Languages” είναι τα παρακάτω:

Id	Name	Rank	ie_encoding	nc_encoding
GR	Greek	3	cp1253	iso-8859-7
N	Norwegian	6	cp1252	iso-8859-1
F	French	4	cp1252	iso-8859-1
D	German	2	cp1252	iso-8859-1
GB	English	1	cp1252	iso-8859-1
IT	Italian	5	cp1252	iso-8859-1

Πίνακας 7-4 Περιεχόμενα του πίνακα “Languages” για τις γλώσσες που υποστηρίζονται αυτή τη στιγμή από το “Hypermedia Custom News System”

Σύμφωνα με τα παραπάνω, ο πίνακας που αναπαριστά τις λέξεις – κλειδιά στη βάση δεδομένων έχει για κάθε υποστηριζόμενη γλώσσα και διαφορετικό πεδίο, και η μορφή του φαίνεται στον παρακάτω πίνακα:

Keywords	Περιγραφή
Id	Κωδικός λέξης κλειδί
Parent_id	Κωδικός πατέρα
GR_description	Ελληνική μετάφραση
F_description	Γαλλική μετάφραση
D_description	Γερμανική μετάφραση
IT_description	Ιταλική μετάφραση
GB_description	Αγγλική μετάφραση
N_description	Νορβηγική μετάφραση

Πίνακας 7-5 Πίνακας αναπαράστασης λέξεων – κλειδιών στη βάση δεδομένων του “Hypermedia Custom News System”

7.2.2.2 Υποστήριξη επιπλέον γλωσσών

Η υποστήριξη μίας επιπλέον γλώσσας απαιτεί την πρόσθεση της αντίστοιχης εγγραφής στον πίνακα “Languages”, η οποία θα περιέχει το όνομα της γλώσσας, το ακρωνύμιό της, τη σειρά εμφάνισής της, καθώς και τα υπόλοιπα αναγκαία στοιχεία. Παράλληλα χρειάζεται να προστεθεί σε κάθε σχέση που έχει πεδία για κάθε υποστηριζόμενη γλώσσα, και ένα πεδίο για τη νέα γλώσσα. Το όνομα αυτού του πεδίου θα είναι το ακρωνύμιο της γλώσσας ακολουθούμενο από το βασικό όνομα του πεδίου, όπως περιγράφηκε στην προηγούμενη παράγραφο.

Η παραπάνω διαδικασία μπορεί να αυτοματοποιηθεί πλήρως κάνοντας χρήση κατάλληλου μηχανισμού trigger, έτσι ώστε με την εισαγωγή μίας εγγραφής στον πίνακα “Languages” να δημιουργούνται αυτόματα τα νέα πεδία στους διάφορους πίνακες.

7.2.2.3 Υποστήριξη μεταφράσεων εγγράφων

Με το σχεδιασμό του σχεσιακού μοντέλου βάσης δεδομένων που περιγράψαμε παραπάνω υπάρχει η δυνατότητα για τη συγγραφή εγγράφων σε περισσότερες από

μία γλώσσες. Τα έγγραφα αυτά όμως δε σχετίζονται μεταξύ τους, ακόμα και αν αποτελούν μετάφραση του ίδιου κειμένου. Υπάρχει, δηλαδή, η περίπτωση ένας τελικός χρήστης, αναζητώντας έγγραφα στο αρχείο του συστήματος, να πάρει ως αποτέλεσμα έγγραφα τα οποία είναι γραμμένα σε διαφορετική γλώσσα, αλλά αναφέρονται στο ίδιο περιεχόμενο.

Από τα παραπάνω προκύπτει η ανάγκη ύπαρξης κάποιας συσχέτισης μεταξύ των εγγράφων που αποτελούν το ένα μετάφραση του άλλου. Για την αντιμετώπιση αυτού του προβλήματος, στο σχεσιακό μοντέλο του συστήματος ένας επιπλέον πίνακας, ο οποίος συσχετίζει αυτά τα έγγραφα. Η δομή του πίνακα για τις έξι γλώσσες που υποστηρίζονται αυτή τη στιγμή από το σύστημα, φαίνεται στον παρακάτω πίνακα.

Translations	Περιγραφή
GR	κωδικός εγγράφου της ελληνικής μετάφρασης
N	κωδικός εγγράφου της νορβηγικής μετάφρασης
F	κωδικός εγγράφου της γαλλικής μετάφρασης
D	Κωδικός εγγράφου της γερμανικής μετάφρασης
GB	Κωδικός εγγράφου της αγγλικής μετάφρασης
IT	Κωδικός εγγράφου της ιταλικής μετάφρασης

Πίνακας 7-6 Ο πίνακας “Translations” που συσχετίζει μεταφράσεις εγγράφων

Κάθε εγγραφή του πίνακα “Translations” αντιπροσωπεύει ένα έγγραφο του οποίου το περιεχόμενο είναι διαθέσιμο σε κάποιο αριθμό γλωσσών. Όπως φαίνεται από τον πίνακα, ακολουθήθηκε και εδώ η ονοματολογία των πεδίων που περιγράφηκε και στις προηγούμενες παραγράφους, ώστε να μην απαιτείται από τις εφαρμογές η ακριβή γνώση των ονομάτων των πεδίων.

Έχοντας τώρα ως αποτέλεσμα κάποιας αναζήτησης ένα σύνολο κωδικών εγγράφων *S*, τα οποία αντιστοιχούν σε διάφορες γλώσσες, οι κωδικοί ομαδοποιούνται ανά μεταφρασμένα έγγραφα με την ακόλουθη SQL ερώτηση.

```
SELECT DISTINCT L1, L2, ..., Ln
FROM Translations
```

WHERE (*L1* IN *S*) OR (*L2* IN *S*) OR ... OR (*Ln* IN *S*)

όπου *L1* έως *Ln* είναι τα ακρωνύμια των υποστηριζόμενων γλωσσών.

Από τις εγγραφές που επιστρέφονται, στον τελικό χρήστη παρουσιάζεται μόνο τα έγγραφα στην γλώσσα που επιθυμεί η τελικός χρήστης, ενώ μπορεί να δοθεί η δυνατότητα να παρουσιαστούν και οι μεταφράσεις αυτών.

7.3 Αναπαράσταση και χειρισμός κειμένου

Η αποθήκευση, αναπαράσταση και ο χειρισμός δεδομένων κειμένου σε υπολογιστές βασίζεται σε μία πληθώρα κανονισμών από οργανισμούς τυποποίησης. Μερικά από τα θέματα που καλύπτουν οι κανονισμοί είναι τα παρακάτω:

- Καθορισμός ελάχιστου δυνατού αριθμού στοιχείων κειμένου προς κωδικοποίηση
- Ανάθεση μοναδικού αριθμητικού κωδικού σε κάθε στοιχείο
- Παροχή βασικών κανόνων για την κωδικοποίηση και ερμηνεία κειμένων έτσι ώστε να είναι δυνατή η ανάγνωση και η επεξεργασία αυτών από εφαρμογές

Οι κανονισμοί αυτοί συνήθως δεν ορίζουν τη μορφή που θα έχουν τα στοιχεία του κειμένου (glyphs) όταν αυτά προβάλλονται στην οθόνη ή εκτυπώνονται σε χαρτί. Οι κανονισμοί ορίζουν με ποιον τρόπο πρέπει να ερμηνεύονται οι χαρακτήρες και όχι πως αυτοί πρέπει να ζωγραφιστούν στην οθόνη. Αυτήν την εργασία την αναλαμβάνουν τα υποσυστήματα του κάθε λειτουργικού συστήματος.

7.3.1 Κανονισμός ASCII

Η τυποποίηση κατά ASCII (American Standard Code for Information Interchange) αναπτύχθηκε από το American National Standards Institute (ANSI) για να ορίσει τον τρόπο με τον οποίο υπολογιστές διαβάζουν και γράφουν χαρακτήρες κειμένου. Το ASCII είναι ένας κωδικός που χρησιμοποιεί 7-bit για την κωδικοποίηση των χαρακτήρων, προσφέροντας έτσι ένα σύνολο από 128 χαρακτήρες. Σε αυτούς συμπεριλαμβάνονται γράμματα, αριθμοί, σημεία στίξης και κωδικοί ελέγχου (για παράδειγμα ο χαρακτήρας που συμβολίζει το τέλος μιας γραμμής).

Οι 128 δυνατοί χαρακτήρες καλύπτονται από το λατινικό αλφάβητο, τους αριθμούς, τα βασικά σημεία στίξης και τους χαρακτήρες ελέγχου. Οι χαρακτήρες αυτοί μπορεί για τις αγγλόφωνες χώρες να μην αποτελεί περιορισμό, αλλά για ένα διεθνές περιβάλλον όπως είναι το διαδίκτυο αποτελεί τροχοπέδη για την χρήση εφαρμογών που απευθύνονται σε ένα ευρύτερο κοινό.

7.3.2 Κανονισμός ISO-8859-x

Η απαίτηση για υποστήριξη χαρακτήρων εκτός των αυστηρά λατινικών χαρακτήρων, οδήγησε στην επέκταση του 7-bit κωδικού που χρησιμοποιεί η τυποποίηση κατά ASCII σε 8-bit. Αυτό διπλασιάζει τους διαθέσιμους χαρακτήρες σε 256 δίνοντας έτσι χώρο για την κάλυψη ενός ευρύτερου φάσματος χαρακτήρων. Στις επιπλέον 128 θέσεις μπορούν να καλυφθούν όλοι οι ειδικοί χαρακτήρες με τα σημεία στίξης που έχουν χώρες όπως η Γαλλία, η Γερμανία και οι Σκανδιναβικές χώρες. Δεν υπάρχει όμως αρκετός χώρος για αλφάβητα που μοιάζουν ελάχιστα με το λατινικό αλφάβητο, όπως είναι το ελληνικό ή το ρωσικό αλφάβητο.

Για να καλυφθεί το παραπάνω κενό επινοήθηκε η έννοια της κωδικοσελίδας (code page). Όλες οι κωδικοσελίδες έχουν κοινούς χαρακτήρες από τους κωδικούς 0 έως 127 ενώ διαφέρουν στους υπόλοιπους 128 χαρακτήρες. Αυτό επιτρέπει με μία κωδικοσελίδα να αναπαρασταθεί το πλήρες λατινικό αλφάβητο μαζί με κάποιους άλλους χαρακτήρες που βρίσκονται στις θέσεις 128 έως 255. Τέτοια σύνολα κωδικοσελίδων έχουν αναπτύξει πολλοί οργανισμοί. Γνωστά σύνολα είναι αυτά της IBM (π.χ. Codepage 437 για ελληνικά), της Microsoft (Windows-cp1253 για ελληνικά) και ο Διεθνής Οργανισμός Τυποποίησης (iso-8859-7 για ελληνικά).

Ο τελευταίος έχει καθιερωθεί τα τελευταία χρόνια, έχοντας υποστήριξη από πολλούς κατασκευαστές υπολογιστικών συστημάτων και εφαρμογών. Μερικές από τις πιο γνωστές κωδικοσελίδες είναι η iso-8859-1 ή Latin-1 που περιλαμβάνει τους περισσότερους χαρακτήρες που βασίζονται στο λατινικό αλφάβητο, συμπεριλαμβανομένου χαρακτήρες του γαλλικού και του γερμανικού αλφαβήτου, και των αλφαβήτων σκανδιναβικών χώρων, η κωδικοσελίδα iso-8859-7 που περιλαμβάνει το ελληνικό αλφάβητο και η κωδικοσελίδα iso-8859-5 για το κυριλλικό αλφάβητο.

7.3.3 Κανονισμός UNICODE

Όπως φάνηκε στο προηγούμενο υποκεφάλαιο, οι κωδικοσελίδες αντιμετώπισαν μεν το πρόβλημα της εσωτερικής αναπαράστασης χαρακτήρων πέραν των λατινικών, αλλά παρόλα αυτά δεν έλυσαν το πρόβλημα της αποθήκευσης κειμένου. Για παράδειγμα, εάν κάποια εφαρμογή διαβάσει από κάποιο αρχείο έναν χαρακτήρα (8-bit) που αντιστοιχεί στον κωδικό 0xD1, δεν γνωρίζει σε ποια κωδικοσελίδα αντιστοιχεί αυτός ο χαρακτήρας. Στην περίπτωση της κωδικοσελίδας iso-8859-1 (Latin-1) πρόκειται για το χαρακτήρα *Ñ* ενώ στην περίπτωση της κωδικοσελίδας iso-8859-7 (Greek) για το χαρακτήρα *Ρ*. Επιπλέον αυτού, η εξάπλωση των ηλεκτρονικών υπολογιστών και σε ευρύτερα στρώματα του πληθυσμού και κυρίως η εξάπλωση του διαδικτύου δημιούργησαν την ανάγκη για μεγαλύτερη απλοποίηση στην αναπαράσταση και διαχείριση κειμένου. Έτσι προωθήθηκαν κανονισμοί όπως το UNICODE που θα περιγραφεί στις παρακάτω παραγράφους.

Το UNICODE αποτελεί έναν ολοκληρωμένο κανονισμό για την κωδικοποίηση χαρακτήρων, για την αναπαράσταση κειμένου και την επεξεργασία του από ηλεκτρονικούς υπολογιστές. Προσφέρει έναν συνεπή (consistent) τρόπο για την κωδικοποίηση πολυγλωσσικού απλού κειμένου και φέρνει έτσι τάξη στο χάος που υπάρχει μέχρι σήμερα και δυσκολεύει την ανταλλαγή αρχείων κειμένου ανά τον κόσμο. Ο σχεδιασμός του UNICODE βασίστηκε στην απλότητα και συνέπεια του ASCII κανονισμού, αλλά υπερβαίνει τον περιορισμό της κωδικοποίησης λατινικών μόνο χαρακτήρων. Χρησιμοποιεί αντί των 7-bit 16-bit για την κωδικοποίηση και παρέχει έτσι τη δυνατότητα για αναπαράσταση για περισσότερους από 65000 χαρακτήρες.

7.3.4 Αποθήκευση και ανάκτηση κειμένου με JDBC

Τα περισσότερα λειτουργικά συστήματα χρησιμοποιούν κωδικοποίηση των χαρακτήρων κατά ASCII ενώ για την αναπαράσταση μη λατινικών χαρακτήρων, χρησιμοποιούν είτε κάποιο διεθνές πρότυπο όπως το iso-8859-x είτε κάποιο πρότυπο που έχει αναπτυχθεί ειδικά για μια συγκεκριμένη πλατφόρμα.

Έτσι για παράδειγμα το λειτουργικό σύστημα Microsoft Windows 95 χρησιμοποιεί το δικό του πρότυπο κωδικοσελίδων με ονομασία “windows cp125x”, ενώ τα Microsoft Windows NT χρησιμοποιούν Unicode κωδικοποίηση για την αναπαράσταση κειμένων. Σε κάθε περίπτωση, πέραν του βασικού τρόπου που χρησιμοποιεί κάθε λειτουργικό σύστημα, είναι προετοιμασμένο και για άλλες κωδικοποιήσεις.

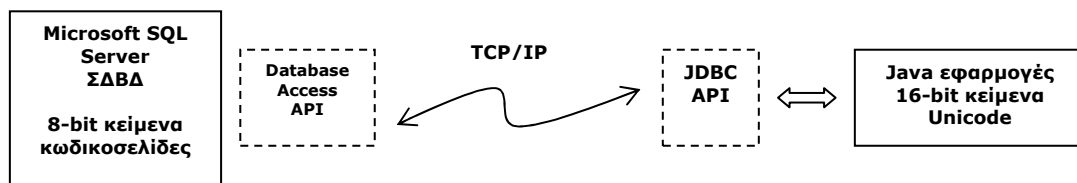
Πέραν της κωδικοποίησης χαρακτήρων που χρησιμοποιούν τα λειτουργικά συστήματα εσωτερικά για την αναπαράσταση κειμένων, προκύπτουν άλλα δύο ζητήματα που επηρεάζουν το σωστό χειρισμό κειμένου.

1. Η αποθήκευση κειμένου σε αρχεία
2. Η υποστήριξη κειμένου από τις εφαρμογές

Το πρώτο σημείο αναφέρεται στον τρόπο με τον οποίο τα κείμενα αποθηκεύονται στη δευτερεύουσα μνήμη, ενώ το δεύτερο σημείο αναφέρεται στις διάφορες εφαρμογές που καλούνται να χειριστούν τα κείμενα.

Παρόλο που το λειτουργικό σύστημα του Εξυπηρετητή Βάσεων Δεδομένων είναι τα Microsoft Windows NT, τα οποία είναι ικανά να χειριστούν Unicode χαρακτήρες, ο Microsoft SQL Server, δηλαδή το Εξυπηρετητής Βάσης Δεδομένων που επιλέχτηκε, αποθηκεύει δεδομένα κειμένου χρησιμοποιώντας 8-bit ανά χαρακτήρα. Επίσης, οι πελάτες της Βάσης Δεδομένων, που είναι κυρίως Java εφαρμογές, χρησιμοποιούν εσωτερικά Unicode χαρακτήρες για την αναπαράσταση κειμένου, και αυτό σε ένα περιβάλλον λειτουργικού συστήματος το οποίο δεν είναι προκαθορισμένο. Τέλος, το JDBC (Java Database Connectivity) που χρησιμοποιείται για την πρόσβαση στη βάση δεδομένων, δεν παρέχει λειτουργικότητα επιρροής στην πρόσβαση σε κείμενα και την μετατροπή τους σε 16-bit Unicode Java χαρακτήρες.

Δεδομένου του παραπάνω γεγονότος, η αξίωση για την υποστήριξη πολλών γλωσσών ταυτόχρονα στην ίδια εφαρμογή, συναντάει σχεδόν ανυπέρβλητα προβλήματα, όταν απαιτείται, εκτός της εισαγωγής του κειμένου στην εφαρμογή, και η αποθήκευση και ανάκτησή του.



Σχήμα 7-1 Μορφή κειμένου 8-bit στην πλευρά της Βάσης Δεδομένων, και μορφή 16-bit στην πλευρά των Java εφαρμογών

Το πρόβλημα έγκειται στο ότι

1. οι εφαρμογές Java δεν γνωρίζουν σε ποια γλώσσα είναι γραμμένα τα κείμενα τα οποία ανακτώνται ώστε να χρησιμοποιήσουν τον κατάλληλο πίνακα αντιστοίχισης για την μετατροπή των χαρακτήρων, και
2. γνωρίζοντας τη γλώσσα του υπό ανάκτηση κειμένου, οι άμεσοι μέθοδοι που προσφέρει το JDBC για την ανάκτηση κειμένου δε είναι παραμετροποιημένοι ως προς κάποιες μεθόδους αντιστοίχισης χαρακτήρων.

Όπως αναφέρθηκε στις προηγούμενες παραγράφους, οι εφαρμογές γνωρίζουν τον αριθμό των υποστηριζόμενων γλωσσών, και επίσης γνωρίζουν σε ποια γλώσσα αντιστοιχεί κάποιο πολυγλωσσικό πεδίο ενός πίνακα.

Η παραδοχή που έγινε για την αντιμετώπιση του πρώτου σκέλους του προβλήματος είναι ότι έχοντας ένα πεδίο της βάσης δεδομένων το οποίο αντιστοιχεί σε κάποια συγκεκριμένη γλώσσα, τα κείμενα που αποθηκεύονται σε αυτό το πεδίο, είναι όλα και αποκλειστικά γραμμένα σε αυτήν τη γλώσσα. Αυτή η παραδοχή δεν επιτρέπει για παράδειγμα την εισαγωγή μιας γαλλικής λέξης (μία λέξη που να χρησιμοποιεί ειδικούς γαλλικούς χαρακτήρες) σε ένα πεδίο της βάσης δεδομένων που αντιστοιχεί στην ελληνική γλώσσα.

Με την παραπάνω παραδοχή, και έχοντας υπ' όψιν τον πίνακα "Languages" της Βάσης Δεδομένων από τις προηγούμενες παραγράφους, η οποία περιγράφει τις υποστηριζόμενες γλώσσες του συστήματος, οι εφαρμογές είναι σε θέση να γνωρίζουν τη γλώσσα του υπό ανάκτηση κειμένου.

Η κωδικοσελίδα που πρέπει να χρησιμοποιηθεί από τις εφαρμογές για τη μετατροπή των 8-bit χαρακτήρων σε 16-bit χαρακτήρες, αποθηκεύεται στον πίνακα “Languages” στα πεδία “nc_encoding” και “ie_encoding” όπως φαίνεται στους Πίνακες 7-3 και 7-4. Το πεδίο “nc_encoding” αποθηκεύει την κωδικοσελίδα στην οποία αντιστοιχεί η συγκεκριμένη γλώσσα, ενώ το πεδίο “ie_encoding” χρησιμοποιείται για τον ίδιο σκοπό ειδικά όμως όταν το Java Applet τρέχει στο περιβάλλον του φυλλομετρητή της Microsoft (Internet Explorer) επειδή η υλοποίηση της Java της Microsoft δεν δέχεται τους “iso-8859-x” κανονισμούς όταν πρόκειται να μετατρέψει 8-bit σε 16-bit χαρακτήρες.

Για να αντιμετωπιστεί το δεύτερο σκέλος του προβλήματος, δηλαδή την μη υποστήριξη επιλογής κωδικοποίησης κατά την ανάκτηση κειμένων από το JDBC, χρειάστηκε να παρακαμφθούν οι μέθοδοι που προσφέρει το JDBC για την ανάκτηση κειμένου, και να αντικατασταθούν από μεθόδους στις οποίες μπορεί να γίνει μετατροπή του κειμένου από κωδικοποίηση σε κωδικοποίηση έτσι ώστε να συνυπολογίζεται στη μετατροπή και η γλώσσα στην οποία βρίσκεται το κείμενο.

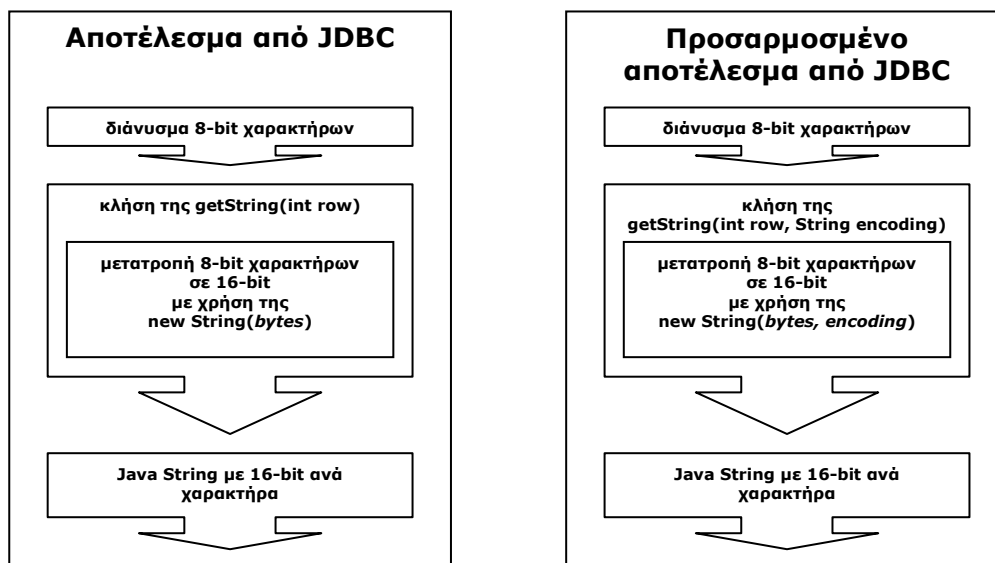
7.3.4.1 Ανάκτηση δεδομένων κειμένου με JDBC

Έχοντας το αποτέλεσμα μιας ερώτησης στη βάση δεδομένων, το JDBC API προσφέρει τη μέθοδο “getString(int row)” για να επιστρέψει το κείμενο του πεδίου “row”. Για να υποστηριχθεί ο μηχανισμός για την μετατροπή των χαρακτήρων, δημιουργήθηκε μία ξεχωριστή Java κλάση, η οποία χρησιμοποιεί το αποτέλεσμα μιας ερώτησης αλλά επιτρέπει τη χρησιμοποίηση μεθόδων όπως

```
getString(int row, String encoding)
```

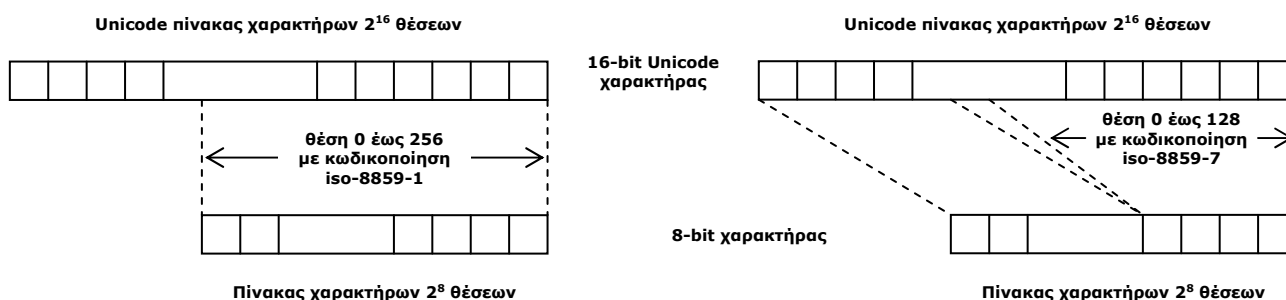
όπου “encoding” το όνομα της κωδικοσελίδας του κειμένου (για παράδειγμα iso-8859-7 για ελληνικά).

Στο Σχήμα 7-2 που ακολουθεί φαίνονται τα βήματα που γίνονται για τη σωστή μετατροπή των κειμένων.



Σχήμα 7-2 Προσαρμογή της μετατροπής 8-bit χαρακτήρων σε 16-bit, έτσι ώστε να παραμετροποιείται η κωδικοποίηση

Η αντιστοίχιση που γίνεται καλώντας τη μέθοδο “new String(bytes, encoding)” φαίνεται στο ακόλουθο Σχήμα 7-3.



Σχήμα 7-3 Η διαφορά αντιστοίχισης 8-bit χαρακτήρων σε 16-bit με χρήση του iso-8859-1 (Latin-1), και κάποιας άλλης κωδικοσελίδας

7.3.4.2 Ενημέρωση δεδομένων κειμένου με JDBC

Όσον αφορά την αντίστροφη διαδικασία, δηλαδή την ενημέρωση υπαρχόντων κειμένων ή την εισαγωγή νέων, χρησιμοποιούνται οι μηχανισμοί που προσφέρει το JDBC για την εκτέλεση SQL ερωτήσεων.

Για την εισαγωγή μιας νέας εγγραφής ή την ενημέρωση κάποιας άλλης, χρησιμοποιούνται οι γνωστές εντολές της SQL INSERT και UPDATE. Σε αυτές τις εκφράσεις υπάρχουν και οι τιμές των πεδίων που πρέπει να ενημερωθούν. Οι SQL

εκφράσεις είναι σε μορφή Java 16-bit Strings, οπότε προκύπτει πάλι το ερώτημα της μετατροπής των τιμών των πεδίων (16-bit Strings) σε χαρακτήρες που θα αποθηκευτούν στη Βάση Δεδομένων (8-bit χαρακτήρες).

Παρόλο που η αντιστοίχιση χαρακτήρων Unicode σε iso-8859-x είναι μονοσήμαντη, και δε θα έπρεπε να απαιτείται ιδιαίτερη μέριμνα, στην πράξη η εκτέλεση μιας τέτοιας SQL έκφρασης (που περιέχει χαρακτήρες πέραν των λατινικών) δεν οδηγεί σε σωστά αποτελέσματα. Το πρόβλημα έγκειται κάπου ανάμεσα στον οδηγό πρόσβασης που χρησιμοποιείται από το JDBC για την πρόσβαση στον Microsoft SQL Server, και στον ίδιο τον SQL Server. Αντί να γίνει απλή αντιστοίχιση των Unicode χαρακτήρων σε 8-bit, φαίνεται να κόβονται απλώς τα περισσότερα σημαντικά 8-bit από τους χαρακτήρες και να χρησιμοποιούνται οι υπόλοιποι ως έχουν.

Επειδή δεν υπάρχει πρόσβαση στον οδηγό του JDBC, οι SQL εκφράσεις υπόκεινται σε ειδική μεταχείριση πριν την εκτέλεσή τους. Έτσι, έχοντας το String της SQL έκφρασης, πρώτα μετατρέπεται σε διάνυσμα από bytes, χρησιμοποιώντας την κωδικοσελίδα της γλώσσας που μας ενδιαφέρει, και επανασηματίζεται από τα bytes το SQL String, χρησιμοποιώντας την κωδικοσελίδα Latin-1 (δηλαδή καμία αλλαγή). Μετά από αυτή τη διαδικασία, τα τμήματα της SQL έκφρασης έχουν μείνει αναλλοίωτα, επειδή περιέχουν μόνο λατινικούς χαρακτήρες, ενώ τα τμήματα που έχουν τις τιμές των πεδίων είναι μετασχηματισμένα με σωστό τρόπο ώστε όταν κοπούν τα πάνω 8-bit των χαρακτήρων για να εκτελεστεί η έκφραση, οι τιμές να βρίσκονται στη σωστή τους μορφή.

Πάντως πρέπει να τονιστεί ότι στην περίπτωση που ο κατασκευαστής του οδηγού αναπτύξει μία έκδοση που να μεριμνά για πολυγλωσσικές τιμές, η παραπάνω διαδικασία θα πρέπει να προσαρμοστεί κατάλληλα.

7.4 Παρουσίαση πολυγλωσσικού κειμένου

Έχοντας αποθηκεύσει τα κείμενα των εγγράφων στη Βάση Δεδομένων με τη σωστή μορφή, η παρουσίασή τους δεν συναντάει κάποια ιδιαίτερη δυσκολία. Την παρουσίαση των εγγράφων την αναλαμβάνει, όπως είδαμε στο προηγούμενο κεφάλαιο, το Υποσύστημα Παρουσίασης Νέων, το οποίο διαβάζει από τη Βάση

Δεδομένων όλες τις σχετικές με ένα έγγραφο πληροφορίες και κατασκευάζει τις κατάλληλες HTML σελίδες.

Κάθε φυλλομετρητής που θα κλιθεί να παρουσιάσει την HTML σελίδα, έχει τους δικούς του μηχανισμούς για την επιλογή της κωδικοσελίδας στην οποία θα εμφανίσει τα κείμενα. Συνήθως υπάρχει μία επιλογή στο φυλλομετρητή που καθορίζει σε ποια κωδικοσελίδα ο χρήστης επιθυμεί να δει τα κείμενα. Πέραν αυτού όμως, μπορεί να καθοριστεί από το ίδιο το HTML έγγραφο η κατάλληλη κωδικοσελίδα, εισάγοντας στην αρχή του εγγράφου τις κατάλληλες HTML δομές. Έτσι εισάγοντας τη δομή

```
<META HTTP-EQUIV="Content-Type" CONTENT="text/html; charset=ISO-8859-7">
```

στην αρχή του HTML εγγράφου, ο φυλλομετρητής γνωρίζει ότι πρόκειται να δείξει ένα HTML έγγραφο, χρησιμοποιώντας την κωδικοσελίδα iso-8859-7. Με αυτόν τον τρόπο απαλλάσσεται ο χρήστης από την χειροκίνητη επιλογή της κωδικοσελίδας.

7.5 Ανακεφαλαίωση

Στο κεφάλαιο παρουσιάστηκε η έννοια του πολυγλωσσικού περιεχομένου που υποστηρίζει το σύστημα νέων “Hypermedia Custom News System” και τα πλεονεκτήματά του έναντι του συστήματος HyNoDe. Έγινε μία σύντομη αναφορά στους τρόπους κωδικοποίησης δεδομένων κειμένου και τη δυσκολία που υπάρχει στην ταυτόχρονη χρήση και χειρισμό κειμένων πολλών γλωσσών.

Προτάθηκε μία λύση που μοντελοποιεί τις υποστηριζόμενες γλώσσες στο σύστημα και αναπτύχθηκαν οι μέθοδοι που χρησιμοποιήθηκαν για την συνεπή ανάκτηση και αποθήκευση πολυγλωσσικών κειμένων σε Βάση Δεδομένων με τη χρησιμοποίηση της Java και του JDBC.

Παρόλο που η λύση που προτάθηκε δεν παρουσιάζει προβλήματα, ο χειρισμός πολυγλωσσικού κειμένου παραμένει ακόμα ένα ανοιχτό θέμα. Η υποστήριξη από μεριάς Βάσεων Δεδομένων του κανονισμού Unicode, είναι ικανή να λύσει τα περισσότερα προβλήματα που συναντώνται σήμερα, έχοντας όμως και την κατάλληλη υποστήριξη από μεριάς εφαρμογών και λειτουργικών συστημάτων.

8 ΤΕΧΝΟΛΟΓΙΑ ΥΛΟΠΟΙΗΣΗΣ

Το σύστημα “Hypermedia Custom News System” έχει υλοποιηθεί χρησιμοποιώντας μια σειρά από διαφορετικές τεχνολογίες και στηρίζεται σε ένα σύνολο εξυπηρετητών, με κομβικό σημείο τον εξυπηρετητή Βάσεων Δεδομένων.

Κύριο μέλημα κατά διάρκεια του σχεδιασμού του συστήματος, ήταν το σύνολο των τεχνολογιών που επρόκειτο να χρησιμοποιηθούν, να βασίζεται σε διεθνή και ανοιχτά πρότυπα, με στόχο αφενός την όσο γίνεται μεγαλύτερη ανεξάρτηση από κατασκευαστές λογισμικού και αφετέρου την ελαχιστοποίηση, από πλευράς τελικών χρηστών, της ανάγκης εγκατάστασης ειδικού λογισμικού προκειμένου να χρησιμοποιήσουν το σύστημα.

Στην πλευρά του τελικού χρήστη, οι παραπάνω στόχοι οδήγησαν στους γνωστούς φυλλομετρητές ως βάση πάνω στην οποία προσφέρονται οι υπηρεσίες προς τους χρήστες. Επίσης, ως κεντρική γλώσσα προγραμματισμού των εφαρμογών, τόσο στην πλευρά του εξυπηρετητή όσο και στην πλευρά του τελικού χρήστη, επιλέχθηκε η Java, ενώ για την επικοινωνία με τα διάφορα υποσυστήματα επιλέχθηκαν τα γνωστά πρωτόκολλα HTTP και FTP.

Στις επόμενες παραγράφους ακολουθεί μία σύντομη περιγραφή των προϊόντων και τεχνολογιών που χρησιμοποιήθηκαν στην ανάπτυξη του συστήματος.

8.1 Microsoft SQL Server

Ο Εξυπηρετητής Βάσης Δεδομένων είναι ίσως το σημαντικότερο τμήμα του συστήματος καθώς είναι υπεύθυνο για την διασφάλιση των παρεχόμενων υπηρεσιών και για τη συνέπεια του συστήματος. Ο Microsoft SQL Server στην έκδοση 6.5 που επιλέχθηκε, είναι ένα ολοκληρωμένο σύστημα διαχείρισης σχεσιακών βάσεων δεδομένων, τα κύρια χαρακτηριστικά του οποίου είναι η επεξεργασία (processing) και

βελτιστοποίηση (optimization) των ερωτήσεων ανάκτησης, οι συνδιαλλαγές (transactions), ο ταυτοχρονισμός (concurrency) και η ανάνηψη (recovery).

Στο παρόν σύστημα, πέραν των τυποποιημένων λειτουργιών για πρόσβαση σε ένα Σχεσιακό Σύστημα Διαχείρισης Βάσεων Δεδομένων, δεν χρησιμοποιήθηκαν κάποια ιδιαίτερα χαρακτηριστικά του συγκεκριμένου προϊόντος της Microsoft, με αποτέλεσμα το όλο σύστημα να μπορεί να μεταφερθεί εύκολα σε οποιοδήποτε άλλο ΣΣΔΒΔ.

8.2 Φυλλομετρητές, HTTP και FTP πρωτόκολλα και εξυπηρετητές

Το διαδίκτυο έχει ταυτιστεί με την έννοια των φυλλομετρητών και την χρήση αυτών ως βάση για κάθε είδους υπηρεσία που προσφέρεται μέσω του διαδικτύου.

Η επιλογή των φυλλομετρητών ως περιβάλλον για την χρήση του “Hypermedia Custom News System” αποτέλεσε φυσική εξέλιξη κατά τη διάρκεια σχεδιασμού του συστήματος. Η ευελιξία που προσφέρει η HTML στην παρουσίαση πληροφορίας, σε συνδυασμό με το καλαίσθητο αποτέλεσμα που μπορεί να επιτευχθεί, κατέστησαν το δίδυμο φυλλομετρητή-HTML ιδανικό συνδυασμό όσον αφορά τη διάθεση των εγγράφων στους τελικούς χρήστες.

Η δυνατότητα που προσφέρουν οι σημερινής τεχνολογίας φυλλομετρητές για την ενσωμάτωση, σε κοινές σελίδες HTML, και εφαρμογών οι οποίες είναι ικανές να αντεπεξέλθουν και σε πιο απαιτητικές εργασίες, οδήγησαν στην επιλογή των φυλλομετρητών ως περιβάλλον εργασίας ακόμα και για τις διαδικασίες της συγγραφής των άρθρων, αλλά και για την αναζήτηση εγγράφων από τους τελικούς χρήστες.

Τα πρωτόκολλα επικοινωνίας HTTP (Hypertext Transfer Protocol) και FTP (File Transfer Protocol) που είναι υπεύθυνα για την μεταφορά HTML σελίδων και αρχείων αντίστοιχα, και υποστηρίζονται από όλους τους σημερινούς φυλλομετρητές, χρησιμοποιήθηκαν και για τις ανάγκες της συγγραφής και παρουσίασης των εγγράφων. Για την υποστήριξη αυτών των πρωτοκόλλων απαιτείται η εγκατάσταση αντίστοιχου εξυπηρετητή, η οποία όμως δεν παρουσιάζει κάποιο ιδιαίτερο πρόβλημα,

επειδή είναι ευρέως διαδεδομένοι και δεν περιορίζονται από συγκεκριμένα λειτουργικά συστήματα.

Τέλος, η ευρεία διάδοση των φυλλομετρητών στους χρήστες υπολογιστών, εγγυάται τη δυνατότητα πρόσβασης στο σύστημα από όλους τους εν δυνάμει χρήστες του διαδικτύου.

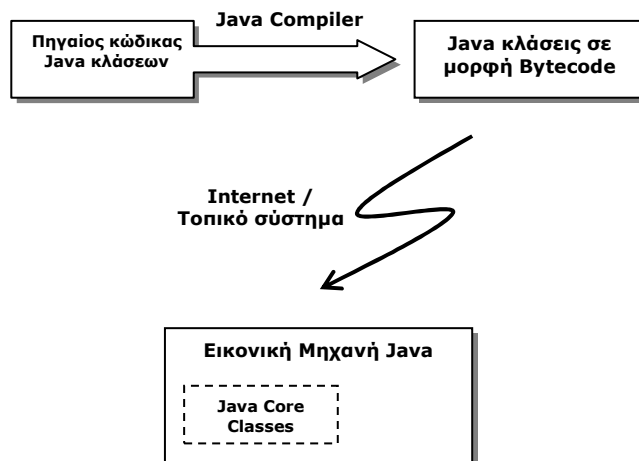
8.3 Java

Η γλώσσα προγραμματισμού Java θα μπορεί να χαρακτηριστεί ως απλή, αντικειμενοστρεφής, ενήμερη για δίκτυα (network-savvy), διερμηνόμενη (interpreted), στιβαρή (robust), ασφαλής, ουδέτερη αρχιτεκτονικής, μεταφερόμενη (portable), υψηλής ταχύτητας, πολυνηματική (multithreaded) και δυναμική. Για κάθε έναν από τους προηγούμενους χαρακτηρισμούς, θα δοθεί μία μικρή εξήγηση στις ακόλουθες παραγράφους.

- **Απλή**, διότι παρόλο που μοιάζει πολύ με τη C++ έχουν απαλειφθεί πολλές ιδιομορφίες και χαρακτηριστικά της C++, που συνήθως δε χρησιμοποιούνται αλλά προκαλούν σύγχυση στον προγραμματιστή. Για παράδειγμα δεν υπάρχει υπερφόρτωση τελεστών (operator overloading) και πολλαπλή κληρονομικότητα (multiple inheritance). Επιπλέον έχει προστεθεί η αυτόματη συλλογή αχρείαστης μνήμης (automatic garbage collection), απαλείφοντας έτσι ένα από τα σημαντικότερα προβλήματα που συναντάει ο προγραμματιστής της C ή της C++.
- **Αντικειμενοστρεφής**, χρησιμοποιώντας ουσιαστικά το ίδιο μοντέλο που χρησιμοποιεί η C++ για τον αντικειμενοστραφή χαρακτηρισμό της.
- **Ενήμερη για δίκτυα**, έχοντας σχεδιαστεί με δεδομένο τη χρησιμοποίηση των εφαρμογών σε δικτυακά περιβάλλοντα. Η γλώσσα έχει ενσωματωμένες πολλές λειτουργίες και ρουτίνες για επικοινωνία σε δίκτυα υπολογιστών, καθιστώντας έτσι την πρόσβαση σε υπηρεσίες του δικτύου το ίδιο εύκολη με την πρόσβαση σε υπηρεσίες του τοπικού υπολογιστή.
- **Διερμηνόμενη**, διότι μεταφράζεται σε γλώσσα μηχανής τη στιγμή της εκτέλεσης.

- **Στιβαρή**, διότι προνοεί για τον έλεγχο λαθών από τα πρώτα στάδια της ανάπτυξης μέχρι και το στάδιο της εκτέλεσης της εφαρμογής. Δεν παρέχει αριθμητικές πράξεις σε δείκτες μνήμης και παρέχει απόλυτη ασφάλεια όσον αφορά την πρόσβαση σε μνήμη εκτός του ιδίου προγράμματος. Επίσης η γλώσσα θέτει αυστηρά όρια στην αλλαγή τύπου (casting).
- **Ασφαλής**, περιορίζοντας τις εφαρμογές που τρέχουν σε περιβάλλοντα δικτύων από την πρόσβαση σε δεδομένα του υπολογιστή χωρίς την άδεια του χρήστη.
- **Ουδέτερη αρχιτεκτονικής και μεταφερόμενη**, διότι δεν εξαρτάται από ιδιαιτερότητες των CPU ή των λειτουργικών συστημάτων παρέχοντας έναν πλήρη ορισμένο τρόπο μεταφοράς και εκτέλεσης των εφαρμογών.
- **Υψηλής ταχύτητας**, σε σχέση με άλλες ερμηνευόμενες γλώσσες, επειδή η μεταφερόμενη δομή του προγράμματος (bytecodes) είναι ήδη πολύ κοντά σε γλώσσα μηχανής.
- **Πολυνηματική**, παρέχοντας τη δυνατότητα για εκτέλεση πολλών διεργασιών «ταυτόχρονα» στην ίδια εφαρμογή εξασφαλίζοντας παράλληλα την αποφυγή ταυτόχρονης πρόσβασης σε πόρους του συστήματος με αναπτυγμένους τρόπους συγχρονισμού.
- **Δυναμική**, διότι φορτώνει και ενώνει τις χρησιμοποιούμενες κλάσεις την ώρα που θα χρειαστεί η πρόσβαση σε αυτές. Επίσης παρέχει τρόπους για έλεγχο του τύπου των αντικειμένων την ώρα της εκτέλεσης (run time type checking).

Σε αντίθεση με παραδοσιακές γλώσσες προγραμματισμού, το αποτέλεσμα της μεταγλώττισης ενός προγράμματος Java, δεν είναι κώδικας μηχανής αλλά ένα ενδιάμεσο στάδιο που ονομάζεται Bytecode. Την ερμηνεία των Bytecodes και την εκτέλεση του κώδικα την αναλαμβάνει μία «Εικονική Μηχανή Java» (Java Virtual Machine). Τέτοιες μηχανές υπάρχουν διαθέσιμες από πολλούς κατασκευαστές, ενώ είναι ενσωματωμένες και στους γνωστούς φυλλομετρητές, δίνοντάς τους τη δυνατότητα εκτέλεσης προγραμμάτων Java που μεταφέρονται μέσα από το δίκτυο.



Σχήμα 8-1 Τα βήματα που ακολουθούνται για την εκτέλεση ενός Java προγράμματος

Παρόλο που υπήρχαν από την αρχή της εμφάνισης της Java μεγάλες προσδοκίες από την αγορά για την εκπλήρωση της υπόσχεσης για ανεξαρτησία από την αρχιτεκτονική των υπολογιστών και από τα λειτουργικά συστήματα, η Java, για μεγάλο χρονικό διάστημα, δεν μπόρεσε να αντεπεξέλθει σε αυτές τις προσδοκίες, εμφανίζοντας ασυμβατότητες σε διαφορετικά υπολογιστικά συστήματα, ακόμα και μεταξύ διαφορετικών υλοποιήσεων στον ίδιο τον υπολογιστή. Σήμερα όμως η γλώσσα έχει φτάσει σε μεγάλο σημείο ωριμότητας και παρέχει πλέον ικανές συνθήκες για να μπορεί να χρησιμοποιηθεί για την ανάπτυξη σοβαρών εφαρμογών. Επίσης η Java αποκτά όλο και μεγαλύτερη σημασία όσον αφορά τη χρήση της και στην πλευρά του εξυπηρετητή.

8.3.1 Java Applets

Τα Applets είναι προγράμματα που φορτώνονται δυναμικά από έναν HTTP εξυπηρετητή και εκτελούνται στο περιβάλλον ενός φυλλομετρητή. Σε αντίθεση με απλές εφαρμογές Java, τα Applets έχουν περιορισμένη πρόσβαση στο σύστημα στο οποίο εκτελούνται, εξασφαλίζοντας έτσι την ασφάλεια του τελικού χρήστη από ανεπιθύμητες ενέργειες, όπως την υποκλοπή προσωπικών πληροφοριών που βρίσκονται στον υπολογιστή του, ή την κακοήθη πρόκληση βλαβών (ιοί υπολογιστών). Με την έγκριση του χρήστη, μπορεί ακόμα και σε ένα Applet να δοθεί πρόσβαση σε επιπλέον τμήματα του υπολογιστή.

Τα Applets, πέραν του περιορισμού στην πρόσβαση σε όλα τα τμήματα του υπολογιστή που εκτελούνται, έχουν πλήρες δυνατότητες, όπως κάθε άλλη γλώσσα προγραμματισμού, ακόμα και για την εκτέλεση πολύπλοκων υπολογισμών στην πλευρά του πελάτη. Τα Applets ενσωματώνονται σε απλές HTML σελίδες και φορτώνονται δυναμικά μαζί με τα κείμενα και τις εικόνες της σελίδας. Αυτή η δυναμική επιτρέπει αλλαγή του κώδικα του Applet χωρίς να το καταλαβαίνει ο χρήστης. Με αυτόν τον τρόπο μπορούν να γίνουν διορθώσεις και επεκτάσεις στα προγράμματα με διαφανή (transparent) για τον τελικό χρήστη τρόπο.

8.3.2 *Java Foundation Classes (JFC)*

Οι Java Foundation Classes (JFC) αποτελούν μία συλλογή κλάσεων που αναπτύχθηκαν από κοινού από την Netscape και την Sun. Μερικά από τα χαρακτηριστικά της συλλογής είναι οι λειτουργίες για 2-D γραφικά, λειτουργίες για χρήση προγραμμάτων από άτομα με ειδικές ανάγκες (accessibility API), προσαρμοζόμενο τρόπο εμφάνισης και συμπεριφοράς (pluggable look & feel), λειτουργικότητα «σύρε και άφησε» (drag & drop) και τα συστατικά Swing (Swing components).

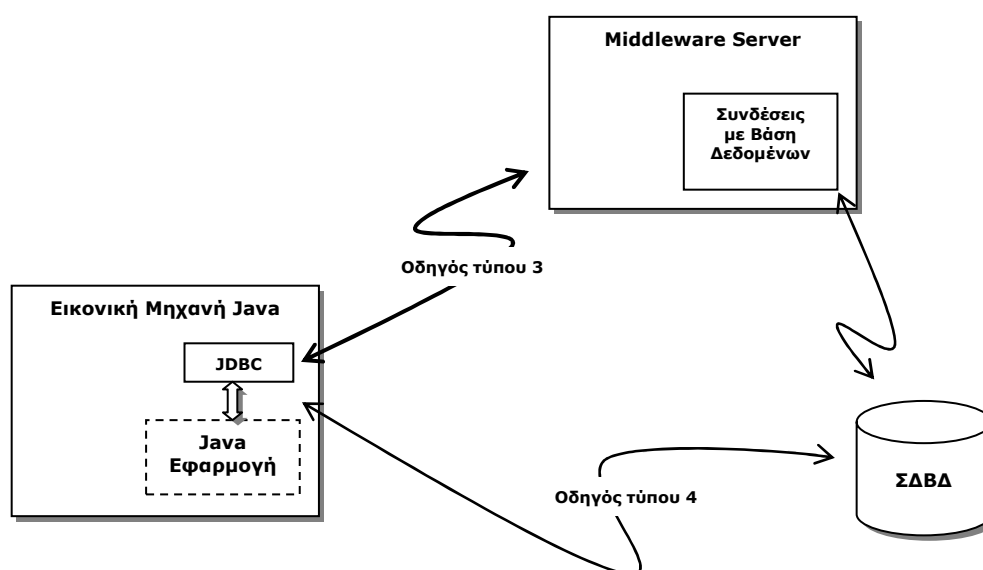
Τα συστατικά Swing, σε συνδυασμό με τον προσαρμοζόμενο τρόπο εμφάνισης, αποτελούν την κύρια καινοτομία της JFC. Πρόκειται ουσιαστικά για μία εκτενή συλλογή από έτοιμα συστατικά για την επικοινωνία ανθρώπου – υπολογιστή (πίνακες, δέντρα, λίστες) που διευκολύνουν στην δημιουργία αποδοτικών και ισχυρών μηχανισμών για την αλληλεπίδραση με το χρήστη (User Interfaces), πράγμα που ως τώρα δεν ήταν δυνατόν να επιτευχθεί με τις βασικές λειτουργίες της Java. Λόγω της τεχνολογίας που χρησιμοποιούν, εξασφαλίζεται ίδια εμφάνιση των προγραμμάτων, ανεξαρτήτως της πλατφόρμας στην οποία εκτελούνται.

Για να χρησιμοποιηθεί η JFC, αυτή τη στιγμή πρέπει είτε να εγκατασταθούν οι κλάσεις στον υπολογιστή που πρόκειται να τρέξουν τα προγράμματα, είτε να μεταφερθούν μαζί με τις υπόλοιπες κλάσεις των προγραμμάτων. Επειδή όμως οι JFC κλάσεις έχουν ενσωματωθεί στην επόμενη έκδοση της Java ως κυρίως κλάσεις (Core Java Classes), αναμένεται η JFC να είναι συμπαγές συστατικό των «Εικονικών

Μηχανών Java». Η επόμενη έκδοση της Java αναμένεται να εκδοθεί στα τέλη του καλοκαιριού του '98.

8.3.3 Java Database Connectivity (JDBC)

Το JDBC API ορίζει Java κλάσεις για την αναπαράσταση συνδέσεων σε Σχεσιακές Βάσεις Δεδομένων, δηλώσεων SQL, αποτελεσμάτων ερωτήσεων κλπ. Επιτρέπει στον προγραμματιστή να εκτελεί ερωτήσεις SQL και να επεξεργάζεται τα αποτελέσματα αυτών. Το JDBC API είναι μέρος των κυρίων κλάσεων της Java (Java Core Classes).



Σχήμα 8-2 Ο ρόλος του JDBC στη σύνδεση με Βάσεις Δεδομένων και η διαφορά ανάμεσα στους οδηγούς τύπου 3 και 4

Υπάρχουν τέσσερις τύποι οδηγών που υλοποιούν το JDBC API:

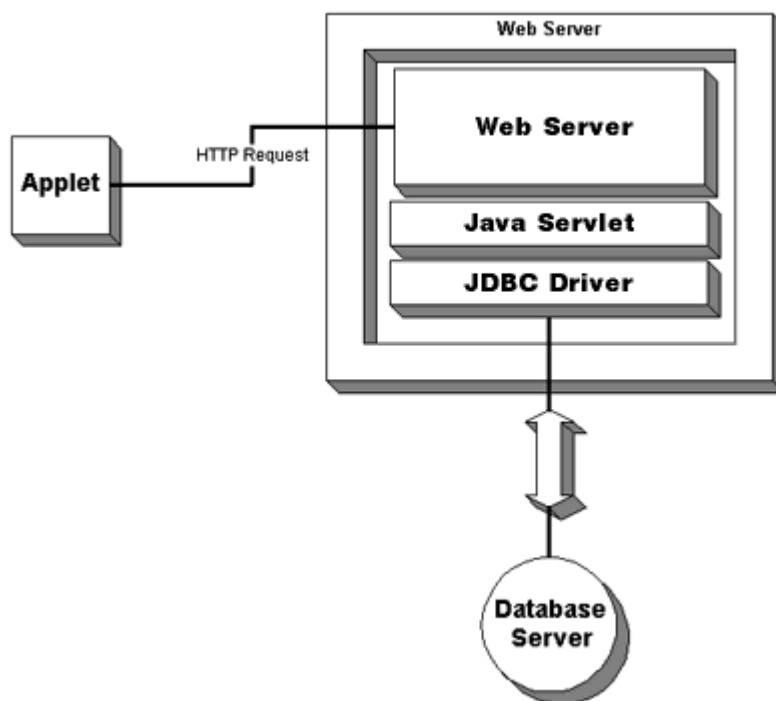
1. Γέφυρα ανάμεσα σε JDBC και ODBC δίνοντας έτσι πρόσβαση σε σχεδόν όλες τις βάσεις δεδομένων μέσω του γνωστού και διαδεδομένου πρωτοκόλλου ODBC.
2. Οδηγοί που σε κάποιο τμήμα δεν είναι ανεξάρτητοι της αρχιτεκτονικής.
3. Οδηγοί γραμμένοι σε Java που προσφέρονται από κατασκευαστές ενδιάμεσων εξυπηρετητών (middleware servers), οι οποίοι αναλαμβάνουν τη σύνδεση με πολλές Βάσεις Δεδομένων και δρουν ως ενδιάμεσοι σταθμοί ανάμεσα στους πελάτες και τους εξυπηρετητές για την διοχέτευση αιτήσεων.

4. Οδηγοί γραμμένοι σε Java που υλοποιούν το πρωτόκολλο επικοινωνίας του Συστήματος Διαχείρισης Βάσεων Δεδομένων και συνδέονται απευθείας σε αυτόν.

Από τους παραπάνω τύπους οδηγών μόνο οι δύο τελευταίοι είναι τελείως ανεξάρτητοι από την πλατφόρμα στην οποία εκτελούνται οι εφαρμογές Java και είναι έτσι ικανές να υποστηρίζουν εφαρμογές που επικοινωνούν μέσω του διαδικτύου από ακαθόριστου τύπου μηχανήματα. Οι οδηγοί τύπου 3 έχουν ως κύριο πλεονέκτημα την εξοικονόμηση αδειών για σύνδεση στα ΣΔΒΔ, επειδή μπορούν με μία σύνδεση να εξυπηρετήσουν «ταυτόχρονα» πολλές αιτήσεις.

8.3.4 Java Servlets

Τα servlets είναι προγράμματα γραμμένα σε Java τα οποία εκτελούνται στην πλευρά του εξυπηρετητή HTTP. Η λειτουργικότητά τους είναι παρόμοια με αυτή των γνωστών CGI προγραμμάτων, δηλαδή προγραμμάτων που εκτελούνται από τον HTTP εξυπηρετητή κατόπιν κάποιας αίτησης από ένα φυλλομετρητή, και το αποτέλεσμα τους επιστρέφεται στο φυλλομετρητή.



Σχήμα 8-3 Η αλληλεπίδραση των servlet με τη Βάση Δεδομένων και τους φυλλομετρητές

Τα πλεονεκτήματα των servlets έναντι των CGI προγραμμάτων, είναι κυρίως ότι όλες οι αιτήσεις σε ένα servlet εξυπηρετούνται από την ίδια διεργασία, δημιουργώντας κάθε φορά και ένα νέο thread. Δηλαδή δεν απαιτείται η δημιουργία μιας νέας διεργασίας με κάθε αίτηση στο servlet οπότε εξαλείφεται ένα μεγάλο πρόβλημα απόδοσης που υπάρχει στα παραδοσιακά CGI. Επιπλέον αυτού, στα servlets παρέχεται η δυνατότητα ύπαρξης κοινών μεταβλητών από περισσότερες αιτήσεις στο ίδιο servlet, το οποίο εξυπηρετεί κυρίως εφαρμογές που απαιτούν πρόσβαση σε Βάση Δεδομένων, οπότε μπορούν να χρησιμοποιούν τη ίδια σύνδεση στη Βάση χωρίς να απαιτείται επανασύνδεση με τον Εξυπηρετητή Βάσεων Δεδομένων κάθε φορά που ενεργοποιείται το servlet.

9 ΑΝΑΚΕΦΑΛΑΙΩΣΗ ΚΑΙ ΜΕΛΛΟΝΤΙΚΕΣ ΕΠΕΚΤΑΣΕΙΣ

9.1 Ανακεφαλαίωση

Στην παρούσα διατριβή υλοποιήθηκαν τα υποσυστήματα συγγραφής άρθρων και παρουσίασης νέων, τα οποία αποτελούν δύο από τα σημαντικότερα τμήματα του εξατομικευμένου συστήματος νέων κατ' απαίτηση "Hypermedia Custom News System". Σκοπός της ανάπτυξης των υποσυστημάτων αυτών είναι η δημιουργία ενός ολοκληρωμένου περιβάλλοντος συγγραφής και διαχείρισης εγγράφων για τους αρθρογράφους, καθώς επίσης και η παροχή σύγχρονων υπηρεσιών νέων στους τελικούς χρήστες του συστήματος μέσω των τεχνολογιών εξατομικευμένων νέων και νέων κατ' απαίτηση.

Τα σημαντικότερα χαρακτηριστικά του υποσυστήματος συγγραφής άρθρων είναι οι αυξημένες δυνατότητες μορφοποίησης εγγράφου, η διαχείριση εγγράφων και πολυμέσων του συστήματος, καθώς και η υποστήριξη συγγραφής και συσχέτισης μεταφραζόμενων άρθρων σε πολλαπλές γλώσσες. Το υποσύστημα παρουσίασης νέων, από την άλλη, φροντίζει για την παρουσίαση των σύγχρονων υπηρεσιών εξατομικευμένων νέων και των νέων κατ' απαίτηση στους τελικούς χρήστες.

Συγκεκριμένα, για το υποσύστημα συγγραφής άρθρων σχεδιάστηκαν και αναπτύχθηκαν σε ικανοποιητικό βαθμό οι εξής δυνατότητες:

- Ενσωμάτωση στο μοντέλο εγγράφου υποσυνόλου του μοντέλου της HTML.
- Μορφοποίηση εγγράφου με τη χρήση μεθόδων επισήμανσης, με την εισαγωγή ενοτήτων, διασυνδέσεων και αντικειμένων εντός κειμένου και με τον ορισμό συσχετίσεων.
- Διαχείριση αρχείου εγγράφων και πολυμέσων του συστήματος.

- Υποστήριξη συγγραφής και συσχέτισης άρθρων σε πολλαπλές γλώσσες.
- Προσαρμογή υποσυστήματος υποβολής ερωτήσεων και υποσυστήματος κατηγοριοποίησης (βλ. και [Καρασ98]).
- Πολιτική ασφαλείας για τους αρθρογράφους του συστήματος.

Η ανάπτυξη του υποσυστήματος παρουσίασης νέων προσανατολίστηκε στην κάλυψη των υπηρεσιών νέων που παρέχει το σύστημα υλοποιώντας τις ακόλουθες δυνατότητες:

- Παρουσίαση εξατομικευμένων νέων.
- Παρουσίαση νέων κατ' απαίτηση.
- Παρουσίαση εγγράφου.
- Πλοήγηση τελικού χρήστη στο μοντέλο κατηγοριοποίησης του συστήματος.
- Παρουσίαση αποτελεσμάτων αναζήτησης.
- Παρουσίαση σε πολλαπλές γλώσσες.

Ιδιαίτερη έμφαση δόθηκε στην υποστήριξη αυθαίρετου αριθμού γλωσσών και συγκεκριμένα στη δυνατότητα αύξησης του αριθμού γλωσσών που υποστηρίζει το σύστημα χωρίς να χρειάζεται καμία αλλαγή σε κάποιο από τα υποσυστήματά του.

9.2 Μελλοντικές Επεκτάσεις

Η αρχιτεκτονική του συστήματος σχεδιάστηκε με τέτοιο τρόπο ώστε να αφήνει μεγάλα περιθώρια βελτίωσης και περαιτέρω ανάπτυξης των δύο υποσυστημάτων. Για το υποσύστημα συγγραφής άρθρων τα παρακάτω θεωρούνται εφικτοί στόχοι στο εγγύς μέλλον:

- Ενσωμάτωση στο μοντέλο εγγράφου μεγαλύτερου υποσυνόλου του μοντέλου της HTML. Για παράδειγμα, δυνατότητα τοποθέτησης στο κείμενο του εγγράφου πινάκων με στοιχεία παρμένα από το αρχείο πολυμέσων του συστήματος.
- Υποστήριξη και άλλων τύπων πολυμέσων, όπως αρχεία τρισδιάστατων αναπαραστάσεων.
- Υποστήριξη αυθαίρετων τύπων αντικειμένων. Μπορεί για ειδικές εφαρμογές να υπάρχει η ανάγκη ενσωμάτωσης στα έγγραφα εξειδικευμένων αντικειμένων, όπως αρχιτεκτονικά σχέδια, οικονομικές μελέτες, κλπ.
- Δυνατότητα απόδοσης αυθαίρετου αριθμού χαρακτηριστικών στα αντικείμενα του συστήματος.
- Προεπισκόπηση εγγράφων χρησιμοποιώντας το HotJava Browser Bean. Κάτι τέτοιο θα βελτιώνει το χρόνο και θα διευκόλυνε τη δημιουργία προεπισκόπησης.
- Αυτόματος τρόπος εντοπισμού και βαθμολόγησης των λέξεων – κλειδιών στο κείμενο του εγγράφου. Με τον τρόπο αυτό θα γινόταν καλύτερη αξιολόγηση των βαρών των λέξεων – κλειδιών. Επίσης η δουλειά των αρθρογράφων θα γινόταν πολύ πιο εύκολη γιατί δεν θα αναγκάζονταν να κατατάσσουν οι ίδιοι το έγγραφό τους στο μοντέλο κατηγοριοποίησης του συστήματος.

Επίσης, και το υποσύστημα παρουσίασης νέων επιδέχεται πολλές βελτιώσεις ανάλογα με την ποιότητα των υπηρεσιών που επιθυμούμε να προσφέρουμε. Τα παρακάτω θεωρούνται απαραίτητες βελτιώσεις για την εμπορική χρήση του συστήματος:

- Τμηματοποίηση αποτελεσμάτων ερώτησης, έτσι ώστε να μην παρουσιάζονται στον τελικό χρήστη μεγάλες λίστες αποτελεσμάτων αναζήτησης.
- Δυνατότητα επισκόπησης κυριοτέρων τίτλων ειδήσεων στην αρχική σελίδα του συστήματος.

- Μέθοδος εξαγωγής σημαντικότερων νέων συνολικά για τον σύστημα, αλλά και για κάθε κατηγορία ξεχωριστά. Κάτι τέτοιο θα μπορούσε να επιτευχθεί βαθμολογώντας τις κατηγορίες και υποκατηγορίες ανάλογα με τη σημαντικότητά τους μέσα στο σύστημα.
- Παρουσίαση διαφημίσεων στις σελίδες των χρηστών ανάλογα με τις προτιμήσεις που έχουν εκφράσει στο διάγραμμά τους. Εύκολα υλοποιήσιμο αν θεωρήσουμε τις διαφημίσεις ως καταταγμένα αντικείμενα του συστήματος.
- Καταγραφή επιλογών των τελικών χρηστών και αυτόματη δημιουργία προτάσεων για την αλλαγή των διαγραμμάτων τους.

Α ΔΙΕΥΘΥΝΣΕΙΣ ΣΥΣΤΗΜΑΤΩΝ ΝΕΩΝ ΣΤΟ ΔΙΑΔΙΚΤΥΟ

Στη συνέχεια παρατίθενται οι διευθύνσεις τόσο των διεθνών όσο και των ελληνικών συστημάτων νέων στο διαδίκτυο, τα οποία εξετάστηκαν κατά την ανάπτυξη του συστήματος “Hypermedia Custom News System”:

- Hypermedia News On Demand - <http://hynode.intranet.gr/hyn/>
- CNN Custom News - http://customnews.cnn.com/cnews/pna_auth.welcome
- MSNBC - <http://www.msnbc.com/news/default.asp>
- NSF Custom News Service - <http://www.nsf.gov/home/cns/start.htm>
- Infocity - <http://www.mclink.it/n/infocity/english.htm>
- Chubu Weekly's On Line - <http://www.eal.or.jp/CW/>
- Naftemporiki - <http://www.naftemporiki.gr/>
- Macedonian Press Agency - <http://www.mpa.gr/>
- Sportline - <http://www.sport.gr/>
- Flash 9.61 Interactive - <http://www.flash.gr/>
- Enet Information Services (Ελευθεροτυπία) - <http://www.enet.gr/>
- Athens News Agency - <http://www.ana.gr/>

Διάγραμμα Profile

Διαδίκτυο Internet

Διασύνδεση Interface

Δρομέας Cursor

Εξατομικευμένα Νέα Personalized News

Ιστοσελίδα Web Page

Κείμενο Text

Λέξεις – Κλειδιά Keywords

Μορφοποίηση Format

Νέα κατ' απαίτηση News On-demand

Παγκόσμιος Ιστός World Wide Web

Πρότυπο Template

Τελικός Χρήστης End User

Φυλλομετρητής Browser

Χαρακτηριστικά Attributes

Hypermedia News On Demand HyNoDe

HyperText Markup Language HTML

Universal Resource Locator URL

BIBΛΙΟΓΡΑΦΙΑ

[Cass97] Cassady-Dorion Luke, Industrial Strength Java, New Riders Publishing, 1997.

[Hyn1] INTRACOM, HyNoDe, Specification of the HyNoDe Initial Prototype, 1997.

[Hyn2] INTRACOM, HyNoDe, Preliminary market survey for the NoD application, 1996.

[Hyn3] INTRACOM, HyNoDe, Definition of NoD Scenarios, 1996.

[Hyn4] INTRACOM, HyNoDe, Identification of HyNoDe Actors, 1996.

[Hyn5] IBM, HyNoDe, Specification of an HTML-based News Document structure, 1996.

[Hyn6] ETNOTEAM, HyNoDe, Browsing tool adapted to the document structure, 1997.

[Hyn7] COSI, HyNoDe, Specification of the News Authoring Tool (HTML based), 1997.

[Java1] Java Technology - <http://www.javasoft.com/>

[Java2] Java Developer Connection - <http://developer.javasoft.com/developer/>

[O'Neil94] O'Neil Patrick, Database, Principles, Programming, Performance, Morgan Kaufmann Publishers, 1994.

[Sol97] Solomon David S., Microsoft SQL Server 6.5 Unleashed, HW Sams, 1997.

[Souk] Soukup P., Inside Microsoft SQL Server 6.5, Microsoft Press, 1997.

[Swi1] The Swing Connection - <http://java.sun.com/products/jfc/tsc/>

[Καρασ98] Καρασαββίδης Στέφανος, Σχεδιασμός και Υλοποίηση Υποσυστήματος Κατηγοριοποίησης και Αναζήτησης Εγγράφων, και Υποσυστήματος Διαχείρισης Διαγραμμάτων σε Εξατομικευμένα Συστήματα Νέων Κατ' Απαίτηση, Διπλωματική εργασία που υποβλήθηκε στο Τμήμα Ηλεκτρονικών Μηχανικών και Μηχανικών Υπολογιστών του Πολυτεχνείου Κρήτης, 1998.

[Σκον98] Σκόνδρας Παναγιώτης, Σχεδιασμός και Υλοποίηση Μηχανισμών Προώθησης και Ανέλκυσης (Push & Pull) σε Σύστημα Νέων Κατ' Απαίτηση, Διπλωματική εργασία που υποβλήθηκε στο Τμήμα Ηλεκτρονικών Μηχανικών και Μηχανικών Υπολογιστών του Πολυτεχνείου Κρήτης, 1998.