



# Τεχνολογικό Εκπαιδευτικό Ίδρυμα Κρήτης

Σχολή Τεχνολογικών Εφαρμογών

Τμήμα Μηχανικών Πληροφορικής

## Πτυχιακή Εργασία

**Τίτλος: “Διαδικτυακό σύστημα συγκέντρωσης και αναζήτησης προϊόντων ηλεκτρονικών καταστημάτων”**

**Αθανάσιος Νόκας (Α.Μ. 548)**

**Νικόλαος Θεοδωρής (Α.Μ. 1060)**

**Επιβλέπων Καθηγητής: κ. Κωστής Αϊβαλής**

**Επιτροπή Αξιολόγησης: κ. Κωστής Αϊβαλής**

**κ. Αθανάσιος Μαλάμος**

**κ. Δημήτριος Καραγιαννάκης**

**Ημερομηνία Παρουσίασης: Πέμπτη, 4 Ιουνίου 2015**



## Ευχαριστίες

Θα θέλαμε να ευχαριστήσουμε τον επιβλέποντα καθηγητή μας κ. Κωστή Αϊβαλή για την πολύτιμη βοήθεια του, τόσο κατά την ανάθεση του θέματος και την δυνατότητα ενασχόλησης μας με αυτό, όσο και καθ' όλη την διάρκεια εκπόνησης της παρούσας πτυχιακής εργασίας.



## Abstract

The purpose of the present paper is to develop and implement a web application suitable for gathering products from several e-shops, which subsequently will be divided into categories in order to provide ease of search to end users. Users using the specific application will be able to roam among a variety of products, use search criteria, as well as comparing the products' prices. This will be achieved by implementing the application on object-oriented PHP.

In this context, the web application will be based on LAMP Server by using Linux-Apache-MySQL-PHP software package, as well as Java environment (JRE & JDK). Given the high capabilities of the application, regarding the great volume of data it will be called to host, we will also develop a full text search/indexing procedure, based on SOLR and Jetty, which will furthermore offer real time search capacity.

In addition, the web application will be widely ported. This will offer the possibility of roaming through the application at ease, regardless of the using device, whether it is a personal computer or whichever smart device. Finally, in order to ensure future expansion the entire application, as well as the web server will be packaged in a Virtual Machine.



## Σύνοψη

Με την παρούσα πτυχιακή εργασία έγινε μια προσπάθεια ανάπτυξης και υλοποίησης μιας διαδικτυακής εφαρμογής, η οποία θα συγκεντρώνει προϊόντα από ηλεκτρονικά καταστήματα (e-shops) και στη συνέχεια θα τα κατηγοριοποιεί, με σκοπό την αναζήτηση από το χρήστη. Ο χρήστης κάνοντας χρήση της web εφαρμογής θα έχει την δυνατότητα περιήγησης ανάμεσα σε πληθώρα προϊόντων, αναζήτησης τους βάσει κριτηρίων, καθώς και σύγκριση των τιμών τους, κάτι το οποίο θα επιτυγχάνεται μέσω υλοποίησης της εφαρμογής πάνω σε αντικειμενοστραφή PHP.

Στο πλαίσιο αυτό, η Web εφαρμογή θα βασιστεί σε LAMP server με χρήση Linux-Apache-MySQL-PHP, καθώς και περιβάλλον Java. Δεδομένης της υψηλής δυναμικής της εφαρμογής, ως προς την φιλοξενία μεγάλου όγκου δεδομένων θα υλοποιηθεί διαδικασία αναζήτησης ευρετηρίου, βασισμένη σε SOLR και Jetty, η οποία κατ' επέκταση θα επιτρέπει και την αναζήτηση σε πραγματικό χρόνο.

Επιπρόσθετα, η web εφαρμογή θα παρουσιάζει υψηλή προσαρμοστικότητα ως προς την συσκευή χρήσης, έτσι ώστε ο χρήστης να έχει την δυνατότητα περιήγησης με την ίδια ευκολία, είτε από τον προσωπικό του υπολογιστή, είτε από οποιαδήποτε smart συσκευή. Τέλος λαμβάνοντας υπόψη και την δυνατότητα περαιτέρω μελλοντικής επέκτασης, το σύνολο της εφαρμογής, καθώς και ο web server θα εισαχθούν σε εικονική μηχανή (Virtual Machine).





## Πίνακας περιεχομένων

Ευχαριστίες .....	iii
Abstract .....	v
Σύνοψη.....	vii
Πίνακας περιεχομένων .....	ix
Πίνακας Εικόνων .....	xiii
Κεφάλαιο Πρώτο: Εισαγωγή .....	1
1.1 Περίληψη .....	1
1.2 Κίνητρο για την Διεξαγωγή της Εργασίας.....	1
1.3 Σκοπός και Στόχοι Εργασίας .....	1
1.4 Δομή Εργασίας .....	2
Κεφάλαιο Δεύτερο: Μεθοδολογία Υλοποίησης.....	3
2.1 Μέθοδος Ανάλυσης & Ανάπτυξης Πτυχιακής .....	3
2.1.1 Τεχνολογίες που χρησιμοποιήθηκαν .....	3
Κεφάλαιο Τρίτο: Συστήματα Διαχείρισης Διαδικτυακού Περιεχομένου (WCMS).....	5
3.1 Εννοιολογικός Προσδιορισμός .....	5
3.2 Δομή ενός Συστήματος Διαχείρισης Διαδικτυακού Περιεχομένου .....	5
3.3 Στόχοι ενός Συστήματος Διαχείρισης Διαδικτυακού Περιεχομένου .....	6
3.4 Είδη Συστημάτων Διαχείρισης Διαδικτυακού Περιεχομένου .....	7
3.5 Συστήματα Διαχείρισης Διαδικτυακού Περιεχομένου Ανοιχτού Κώδικα .....	7
3.5.1 Πλεονεκτήματα .....	8
3.5.2 Μειονεκτήματα .....	9
3.5.3 Οφέλη από τη συμμετοχή στην ανάπτυξη open source λογισμικού .....	9
3.6 Πλαίσιο Διαχείρισης Περιεχομένου (Content Management Framework – CMF).....	10
Κεφάλαιο Τέταρτο: Drupal .....	11
4.1 Ορισμός.....	11
4.2 Ιστορική εξέλιξη .....	11
4.3 Διάγραμμα Ροής “The Drupal flow” .....	12
4.4 Drupal Core .....	13
4.5 Modules.....	14
4.6 Themes .....	14
4.7 Κριτική Διερεύνηση Drupal .....	14
4.7.1 Βασικά Πλεονεκτήματα Drupal .....	15

4.7.2 Drupal vs. Joomla .....	15
4.7.3 Drupal vs. Wordpress .....	16
4.8 Λόγοι επιλογής του Drupal .....	16
Κεφάλαιο Πέμπτο: LAMP (Software Bundle) .....	19
5.1 Εννοιολογικός Προσδιορισμός .....	19
5.2 Linux .....	20
5.2.1 Ορισμός .....	20
5.2.2 Ιστορική αναδρομή .....	21
5.2.3 Χαρακτηριστικά .....	21
5.2.4 Διανομές Linux .....	21
5.3 Apache Web Server .....	22
5.3.1 Εισαγωγικά στοιχεία .....	22
5.3.2 Χαρακτηριστικά .....	23
5.4 MySQL .....	23
5.4.1 Ορισμός .....	23
5.4.2 Πλεονεκτήματα .....	24
5.5 PHP .....	25
5.5.1 Ορισμός .....	25
5.5.2 Ιστορική εξέλιξη .....	25
5.5.3 Χαρακτηριστικά και τρόπος λειτουργίας .....	25
5.5.4 Πλεονεκτήματα .....	27
Κεφάλαιο Έκτο: Άλλες Τεχνολογίες .....	29
6.1 XML .....	29
6.1.1 Εισαγωγικά στοιχεία .....	29
6.1.2 Ορισμός XML .....	29
6.1.3 Ιστορική εξέλιξη .....	30
6.1.4 Πλεονεκτήματα – Μειονεκτήματα XML .....	31
6.2 RSS .....	32
6.2.1 Ορισμός .....	32
6.2.2 Ιστορική εξέλιξη .....	32
6.2.3 Πλεονεκτήματα .....	33
6.2.4 RSS Readers .....	34
6.3 Java .....	34
6.3.1 Ορισμός .....	34

6.3.2 Java Virtual Machine .....	34
6.4 Apache SOLR .....	35
6.4.1 Ορισμός .....	35
6.4.2 Ιστορικά στοιχεία .....	35
6.4.3 Βασικά χαρακτηριστικά.....	36
6.5 Jetty.....	36
6.5.1 Ορισμός.....	36
6.6 Apache Subversion .....	37
6.6.1 Εισαγωγικά στοιχεία .....	37
6.6.2 Ιστορικά στοιχεία .....	37
6.7 Responsive Web Design .....	37
6.7.1 Ορισμός.....	37
6.7.2 Πλεονεκτήματα .....	38
6.8 XPath .....	38
6.8.1 Εννοιολογικός Προσδιορισμός .....	38
6.9 RegEx.....	39
6.9.1 Εννοιολογικός Προσδιορισμός .....	39
6.10 Style Sheet Language.....	40
6.10.1 Εννοιολογικός προσδιορισμός .....	40
6.10.2 CSS (Cascading Style Sheet) .....	40
Κεφάλαιο Έβδομο: Υλοποίηση .....	43
7.1 Διάφορες Απαιτούμενες Εγκαταστάσεις στο server .....	43
7.1.1. Εγκατάσταση Samba .....	43
7.1.2. Εγκατάσταση Drush.....	43
7.1.3 Εγκατάσταση Compass.....	44
7.1.4 Εγκατάσταση SOLR .....	44
7.1.5 Δημιουργία SVN .....	44
7.2 Modules που χρησιμοποιήθηκαν.....	45
7.2.1 Έτοιμα Modules.....	45
7.2.2 Modules που αναπτύχθηκαν.....	47
7.3 Παρουσίαση χρήσης της Εφαρμογής .....	48
7.4 Παρουσίαση περιβάλλοντος διαχείρισης της εφαρμογής .....	52
7.4.1 Παρουσίαση του menu “Content” .....	52
7.4.2 Εισαγωγή νέου καταστήματος .....	53

Κεφάλαιο Όγδοο: Αποτελέσματα .....	65
8.1 Συμπεράσματα .....	65
8.2 Μελλοντική Εργασία και Επεκτάσεις .....	65
Βιβλιογραφία .....	67
Ηλεκτρονικές Πηγές .....	67
Παράρτημα .....	69

## Πίνακας Εικόνων

ΕΙΚΟΝΑ 1: ΤΡΟΠΟΣ ΛΕΙΤΟΥΡΓΙΑΣ CMS .....	5
ΕΙΚΟΝΑ 2: ΛΟΓΟΤΥΠΟ DRUPAL .....	11
ΕΙΚΟΝΑ 3: "THE DRUPAL FLOW" .....	12
ΕΙΚΟΝΑ 4: ΛΟΓΟΤΥΠΙΑ ΔΗΜΟΦΙΛΕΣΤΕΡΩΝ CMS .....	15
ΕΙΚΟΝΑ 5: ΛΟΓΟΤΥΠΟ LAMP.....	19
ΕΙΚΟΝΑ 6: ΛΟΓΟΤΥΠΟ LINUX .....	20
ΕΙΚΟΝΑ 7: ΛΟΓΟΤΥΠΟ ΑΡΑΧΗ .....	22
ΕΙΚΟΝΑ 8: ΛΟΓΟΤΥΠΟ MYSQL .....	23
ΕΙΚΟΝΑ 9: ΛΟΓΟΤΥΠΟ PHP .....	25
ΕΙΚΟΝΑ 10: ΤΡΟΠΟΣ ΛΕΙΤΟΥΡΓΙΑΣ PHP .....	26
ΕΙΚΟΝΑ 11: ΤΡΟΠΟΣ ΛΕΙΤΟΥΡΓΙΑΣ XML .....	30
ΕΙΚΟΝΑ 12: ΛΟΓΟΤΥΠΟ RSS .....	32
ΕΙΚΟΝΑ 13: ΤΑΧΥΤΗΤΑ ΑΝΑΖΗΤΗΣΗΣ SOLR vs MySQL .....	44
ΕΙΚΟΝΑ 14: ΠΑΡΟΥΣΙΑΣΗ ΤΗΣ ΚΑΡΤΕΛΑΣ "RECENT" .....	48
ΕΙΚΟΝΑ 15: ΔΥΝΑΤΟΤΗΤΑ "MOUSE OVER" .....	49
ΕΙΚΟΝΑ 16: ΠΑΡΟΥΣΙΑΣΗ ΤΗΣ ΚΑΡΤΕΛΑΣ "SEARCH" .....	50
ΕΙΚΟΝΑ 17: ΠΑΡΟΥΣΙΑΣΗ ΤΩΝ ΠΡΟΣΘΕΤΩΝ ΦΙΛΤΡΩΝ.....	51
ΕΙΚΟΝΑ 18: ΠΑΡΟΥΣΙΑΣΗ ΤΗΣ ΚΑΡΤΕΛΑΣ "FEATURED" .....	52
ΕΙΚΟΝΑ 19: ΠΑΡΟΥΣΙΑΣΗ ΤΟΥ MENU "CONTENT".....	53
ΕΙΚΟΝΑ 20: ΠΡΟΣΘΗΚΗ CONTENT TYPE .....	54
ΕΙΚΟΝΑ 21: ΔΗΜΙΟΥΡΓΙΑ FEED IMPORTER.....	55
ΕΙΚΟΝΑ 22: ΣΥΝΔΕΣΗ FEED IMPORTER ΜΕ ΤΟ ΑΝΤΙΣΤΟΙΧΟ CONTENT TYPE .....	55
ΕΙΚΟΝΑ 23: ΕΠΙΛΟΓΗ ΤΟΥ ΤΡΟΠΟΥ ΕΙΣΑΓΩΓΗΣ ΤΩΝ ΠΡΟΪΟΝΤΩΝ.....	56
ΕΙΚΟΝΑ 24: ΠΡΟΣΘΕΤΕΣ ΡΥΘΜΙΣΕΙΣ ΓΙΑ ΤΗΝ ΕΙΣΑΓΩΓΗ ΠΡΟΪΟΝΤΩΝ .....	57
ΕΙΚΟΝΑ 25: ΕΠΙΛΟΓΗ ΤΡΟΠΟΥ ΑΝΑΛΥΣΗΣ ΔΕΔΟΜΕΝΩΝ .....	58
ΕΙΚΟΝΑ 26: ΕΠΙΛΟΓΗ ΤΡΟΠΟΥ ΕΠΕΞΕΡΓΑΣΙΑΣ ΤΩΝ ΔΕΔΟΜΕΝΩΝ.....	58
ΕΙΚΟΝΑ 27: ΡΥΘΜΙΣΕΙΣ ΓΙΑ ΤΗΝ ΕΠΕΞΕΡΓΑΣΙΑ ΤΩΝ ΔΕΔΟΜΕΝΩΝ .....	59
ΕΙΚΟΝΑ 28: ΡΥΘΜΙΣΕΙΣ ΑΡΧΙΚΩΝ ΠΛΗΡΟΦΟΡΙΩΝ .....	59
ΕΙΚΟΝΑ 29: ΣΥΜΠΛΗΡΩΣΗ ΤΩΝ ΧΡΑΤΗ.....	60
ΕΙΚΟΝΑ 30: ΜΟΡΦΟΠΟΙΗΣΗ ΔΕΔΟΜΕΝΩΝ .....	61
ΕΙΚΟΝΑ 31: ΔΗΜΙΟΥΡΓΙΑ CONTENT .....	62
ΕΙΚΟΝΑ 32: CRAWLING CONFIGURATION .....	63



## Κεφάλαιο Πρώτο: Εισαγωγή

Ακολουθεί μια σύντομη περιγραφή της εργασίας και των στόχων της, καθώς και του τρόπου υλοποίησης της.

### 1.1 Περίληψη

Σκοπός της παρούσας εργασίας ήταν η υλοποίηση μιας web εφαρμογής, η οποία θα συγκεντρώνει προϊόντα από ηλεκτρονικά καταστήματα (e-shops) και θα τα κατηγοριοποιεί με σκοπό την αναζήτηση τους από τον χρήστη. Με την χρήση της εφαρμογής ο χρήστης θα έχει την δυνατότητα να αναζητήσει προϊόντα επιλέγοντας ο ίδιος τα κριτήρια αναζήτησης, να συγκρίνει τιμές και να επιλέξει το πιο οικονομικά συμφέρον. Πιο συγκεκριμένα η υλοποίηση έγινε σε τρία στάδια:

Στο πρώτο στάδιο έγινε η εγκατάσταση εξυπηρετητή (server) πάνω σε Linux-Apache-MySQL-PHP και η εγκατάσταση περιβάλλοντος Java (JRE και JDK). Στην συνέχεια έγινε σχεδιασμός και υλοποίηση της web εφαρμογής με την χρήση αντικειμενοστραφούς PHP, ενώ στο παράλληλα υλοποιήθηκε διαδικασία αναζήτησης ευρετηρίου (full text search/indexing) με την χρήση SOLR και Jetty

Βασικό χαρακτηριστικό της εφαρμογής είναι η δυνατότητα που παρέχει στον χρήστη να πραγματοποιεί αναζήτηση ανάμεσα σε πληθώρα προϊόντων σε πραγματικό χρόνο, προσδίδοντας της με αυτόν τον τρόπο μεγαλύτερη χρηστική αξία. Επιπλέον, αξίζει να σημειωθεί, η προσαρμοστικότητα του περιβάλλοντος της εφαρμογής ως προς την συσκευή χρήσης από τον τελικό χρήστη.

### 1.2 Κίνητρο για την Διεξαγωγή της Εργασίας

Κίνητρο για την διεξαγωγή της εργασίας αποτέλεσε η συνεχώς αυξανόμενη χρήση του διαδικτύου για την αγορά προϊόντων τεχνολογίας. Καθημερινά χιλιάδες καταναλωτές επισκέπτονται ηλεκτρονικά καταστήματα συγκρίνοντας τιμές προϊόντων με στόχο την εύρεση καλύτερης τιμής. Παράλληλα πρόκειται για μια αγορά άκρως ανταγωνιστικής με την συνεχή είσοδο νέων καταστημάτων στο κόσμο του ηλεκτρονικού εμπορίου. Με βάση τα παραπάνω η εφαρμογή καθίσταται ένα πολύτιμο εργαλείο στη διάθεση του καταναλωτή.

### 1.3 Σκοπός και Στόχοι Εργασίας

Σκοπός της εργασίας ήταν να δημιουργηθεί μια web εφαρμογή που θα βοηθήσει τον τελικό χρήστη στην γρήγορη αναζήτηση του οικονομικότερου προϊόντος. Στα πλαίσια επίτευξης αυτού του σκοπού, η εφαρμογή υλοποιήθηκε με γνώμονα τη φιλικότητα προς τον χρήστη, παρέχοντας αυξημένες δυνατότητες στην αναζήτηση και παρουσίαση των προϊόντων, στον μικρότερο δυνατό χρόνο.

## 1.4 Δομή Εργασίας

Η δομή της εργασίας αναλύεται ως εξής:

Μετά τη παράθεση κάποιων εισαγωγικών στοιχείων σχετικά με το αντικείμενο της πτυχιακής εργασίας, των στόχων και των κινήτρων διεξαγωγής της, στο επόμενο κεφάλαιο ακολουθεί ένα γενικό πλάνο της μεθοδολογίας που υιοθετήθηκε για την ανάπτυξη της web εφαρμογής, καθώς και των σημαντικότερων τεχνολογιών που χρησιμοποιήθηκαν για αυτό το σκοπό.

Στο τρίτο κεφάλαιο γίνεται μια εισαγωγή στα συστήματα διαχείρισης περιεχομένου (CMS), ενώ αναλύεται διεξοδικότερα η δομή και ο τρόπος λειτουργίας ενός συστήματος διαχείρισης διαδικτυακού περιεχομένου (WCMS).

Το τέταρτο κεφάλαιο είναι αφιερωμένο στο λογισμικό Drupal, αναλύοντας τα πλεονεκτήματα του έναντι άλλων αντίστοιχων λογισμικών και τους λόγους επιλογής του για την υλοποίηση της παρούσας εφαρμογής. Στη συνέχεια παρουσιάζονται τα modules και theme που υλοποιήθηκαν και χρησιμοποιήθηκαν.

Στο πέμπτο κεφάλαιο γίνεται διεξοδική αναφορά στο λογισμικό πακέτο LAMP, καθώς και τα επιμέρους στοιχεία του, Linux-Apache-MySQL-PHP.

Στο έκτο κεφάλαιο γίνεται ανάλυση των βασικότερων τεχνολογιών που χρησιμοποιήθηκαν για να διασφαλιστεί η εύρυθμη λειτουργία της web εφαρμογής. Ενδεικτικά αναφέρουμε τη χρήση Java, XML, Responsive Web Design, RSS, καθώς και RegEx, XPath.

Στο κεφάλαιο που ακολουθεί, έβδομο, γίνεται παρουσίαση με εκτενή ανάλυση του τρόπου ανάπτυξης και υλοποίησης της εφαρμογής.

Τέλος, το όγδοο και τελευταίο κεφάλαιο παρουσιάζει τα αποτελέσματα της πτυχιακής και τις δυνατότητες περαιτέρω μελλοντικής επέκτασης.



## Κεφάλαιο Δεύτερο: Μεθοδολογία Υλοποίησης

Το αντικείμενο της παρούσας πτυχιακή εργασίας, όπως προαναφέρθηκε, είναι η υλοποίηση μιας διαδικτυακής εφαρμογής, η οποία θα συγκεντρώνει προϊόντα από ηλεκτρονικά καταστήματα και θα τα κατηγοριοποιεί με σκοπό την αναζήτησή τους από το χρήστη. Βασικό κομμάτι πάνω στο οποίο στηρίχθηκε η υλοποίηση και ανάπτυξη της εφαρμογής αποτέλεσε το πακέτο λογισμικού LAMP, καθώς και η εγκατάσταση περιβάλλοντος Java. Στη συνέχεια, γίνεται αναφορά του τρόπου υλοποίησης της εφαρμογής και των βασικών τεχνολογιών που χρησιμοποιήθηκαν.

### 2.1 Μέθοδος Ανάλυσης & Ανάπτυξης Πτυχιακής

Αρχικά εγκαταστάθηκε το πακέτο λογισμικού LAMP. Πιο συγκεκριμένα, χρησιμοποιήθηκε ένας virtual server πάνω στον οποίο έγινε εγκατάσταση του Linux διανομής xubuntu 13.4, του Apache web server έκδοσης 2, της βάσης δεδομένων MySQL έκδοσης 5.5 και της γλώσσας προγραμματισμού PHP έκδοσης 5.5.

Στη συνέχεια προχωρήσαμε στην εγκατάσταση:

- Java JRE και Java JDK
- SOLR πάνω σε Jetty
- Drupal
- Modules

#### 2.1.1 Τεχνολογίες που χρησιμοποιήθηκαν

Στις κυριότερες τεχνολογίες και software που χρησιμοποιήθηκαν για την ανάπτυξη της web εφαρμογής συγκαταλέγονται επιγραμματικά τα εξής:

- *LAMP*. Πακέτο λογισμικού για την κατασκευή δυναμικών ιστοσελίδων, με συστατικά Linux-Apache-MySQL-PHP.
- *Java*. Αντικειμενοστραφής γλώσσα προγραμματισμού.
- *Jetty*. Ένας αμιγώς βασιζόμενος σε Java HTTP Server και Java Servlet Container, ο οποίος χρησιμοποιείται για τη διασύνδεση μεταξύ υπολογιστικών συστημάτων, που αποτελούν μέρος μεγαλύτερων λογισμικών πλαισίων.
- *Apache SOLR*. Ο Apache SOLR είναι μια πολύ γρήγορη πλατφόρμα αναζήτησης ανοιχτού κώδικα γραμμένη σε Java, ενώ μπορεί να λειτουργεί ως αυτόνομος server αναζήτησης πλήρους κειμένου χρησιμοποιώντας ένα servlet Container, όπως Apache Tomcat ή Jetty.
- *Drupal*. Ανοιχτού κώδικα αρθρωτό σύστημα διαχείρισης περιεχομένου (Content Management System, CMS), γραμμένο στη γλώσσα προγραμματισμού PHP, το οποίο διανέμεται σύμφωνα με τη Γενική Άδεια Δημόσιας Χρήσης GNU.
- *XML*. Γλώσσα προγραμματισμού, η οποία χρησιμοποιείται για την περιγραφή μιας σημειακής γλώσσας, όπως HTML.
- *RSS*. Αποτελεί μια προτυποποιημένη μέθοδο ανταλλαγής ψηφιακού πληροφοριακού περιεχομένου στο διαδίκτυο, μέσω των ροών RSS (RSSFeeds), η οποία στηρίζεται στην γλώσσα σήμανσης XML.
- *Responsive Web Design*. Διαδικασία σχεδιασμού και κατασκευής ιστοσελίδων, η οποία στοχεύει στην παροχή της βέλτιστης εμπειρίας θέασης, μέσω της εύκολης ανάγνωσης και

πλοήγησης, με την ελάχιστη δυνατή αλλαγή μεγέθους, μετακίνησης και scrolling σε ένα ευρύ φάσμα συσκευών.

- *XPath*. Γλώσσα ερωτημάτων για την επιλογή «κόμβων» (“nodes”) σε ένα XML έγγραφο.
- *RegEx*. Γλώσσα που έχει σχεδιασθεί ειδικά για την αυτοματοποίηση της επεξεργασίας κειμένου, γενικών αρχείων κειμένου, ειδικών μορφών κειμένου, αλλά και συμβολοσειρών εισόδου, με τον πιο περιεκτικό και ευέλικτο τρόπο.

## Κεφάλαιο Τρίτο: Συστήματα Διαχείρισης Διαδικτυακού Περιεχομένου (WCMS)

### 3.1 Εννοιολογικός Προσδιορισμός

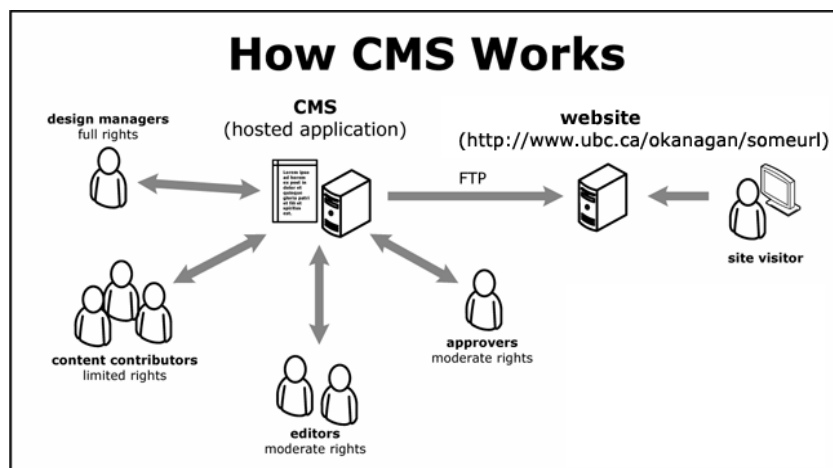
Για πολλά χρόνια, η δυνατότητα μιας επιχείρησης να δημιουργήσει ένα website και κυρίως να το διατηρήσει ενημερωμένο υπήρξε άρρηκτα συνδεδεμένη με τη σύναψη συμβολαίου με μια εταιρία παροχής υπηρεσιών συντήρησης. Αξίζει να σημειωθεί ότι η παραπάνω επιλογή συνεπαγόταν και την ανάληψη σημαντικού κόστους από την πλευρά της επιχείρησης, ενώ ο βαθμός δυσκολίας αυξανόταν με την αύξηση του όγκου της πληροφορίας, η οποία έπρεπε να ενσωματωθεί, καθώς και με την αυξανόμενη ανάγκη ύπαρξης υψηλού ρυθμού ανανέωσης αυτής της πληροφορίας.

Εναλλακτική λύση στο παραπάνω πρόβλημα εμφανίστηκε μόλις πριν λίγα χρόνια, με την εμφάνιση των Συστημάτων Διαχείρισης Περιεχομένου. Το σύνολο των εφαρμογών που υποστηρίζει ένα Σύστημα Διαχείρισης Περιεχομένου (Content Management System – CMS), στηρίζεται σε ένα κοινό σημείο, την διαχείριση περιεχομένου. Με το όρο περιεχόμενο (Content) αναφερόμαστε σε κάθε είδος ψηφιακής πληροφορίας, όπως ηλεκτρονικά έγγραφα, εικόνες, γραφικά, video, java animation, flash animation, αρχεία ήχου, οτιδήποτε δηλαδή μπορεί να διαχειριστεί σε ηλεκτρονική μορφή.

Ένα CMS δύναται να πάρει διάφορες μορφές ανάλογα με τον τρόπο εφαρμογής του. Η μορφή που θα μας απασχολήσει στη συνέχεια είναι αυτή του Συστήματος Διαχείρισης Διαδικτυακού Περιεχομένου (WCMS), το οποίο αφορά στην διάθεση και διαχείριση περιεχομένου στο διαδίκτυο (internet), μέσω ενός εταιρικού ιστοχώρου, καθώς και στη διάθεση περιεχομένου στο εταιρικό ενδοδίκτυο (intranet).

Ένα WCMS είναι ένα εργαλείο που επιτρέπει, στην πλειονότητα τεχνικά εξειδικευμένου ή μη προσωπικού, την δημιουργία, επεξεργασία, διαχείριση και τέλος δημοσίευση ποικίλου περιεχομένου, ενώ παράλληλα περιορίζεται από ένα σύνολο κανόνων, διαδικασιών και ροής εργασιών, με στόχο την εξασφάλιση της συνάφειας και της αξιοπιστίας του ηλεκτρονικού περιεχομένου.

### 3.2 Δομή ενός Συστήματος Διαχείρισης Διαδικτυακού Περιεχομένου



Εικόνα 1: Τρόπος Λειτουργίας CMS

Η λειτουργία ενός WCMS κατηγοριοποιείται στα εξής στάδια:

- *Δημιουργία Περιεχομένου:* Τα WCMS στηρίζονται σε ένα εύχρηστο περιβάλλον συγγραφής, όπως το Word, για τη δημιουργία περιεχομένου νέων σελίδων ή την ανανέωση του υπάρχοντος περιεχομένου, χωρίς να απαιτείται τεχνική εξειδίκευση, όπως η γνώση HTML. Παράλληλα, τα περισσότερα WCMS παρέχουν τη δυνατότητα ανανέωσης του περιεχομένου από οποιαδήποτε συσκευή έχει πρόσβαση στο διαδίκτυο, απλά μέσω ενός web browser.
- *Διαχείριση Περιεχομένου:* Βασικό κομμάτι ενός WCMS είναι η ύπαρξη μιας βάσης δεδομένων, στην οποία αρχειοθετείται το σύνολο των δεδομένων που αφορούν τον ιστοχώρο, ενώ παράλληλα παρέχει και τη δυνατότητα παρακολούθησης των αλλαγών που έχει υποστεί μια σελίδα, το ποιος άλλαξε τι και πότε, καθώς επίσης εξασφαλίζει ότι κάθε χρήστης μπορεί να αλλάξει τον τομέα ευθύνης του στον ιστοχώρο.

Σημαντικό στοιχείο στο στάδιο αυτό είναι και οι δυνατότητες ροής εργασίας που προσφέρει ένα WCMS. Σε κάθε βήμα της διαχείρισης του περιεχομένου, το σύστημα ειδοποιεί τους χρήστες που εμπλέκονται, επιτρέποντας έτσι τη συμμετοχή περισσότερων συγγραφέων στη διαχείριση του ιστοχώρου, διατηρώντας πάντα υψηλά επίπεδα ελέγχου της ποιότητας, ακρίβειας και συνοχής της πληροφορίας.

- *Δημοσίευση:* Το τελικό περιεχόμενο μπορεί να δημοσιευτεί είτε στον ιστοχώρο, είτε στο ενδοδίκτυο της επιχείρησης. Παρέχονται εξειδικευμένες μηχανές δημοσίευσης, οι οποίες φροντίζουν για την εναρμόνιση του περιεχομένου με την εμφάνιση και τη δομή του ιστοχώρου αυτόματα μόλις δημοσιευτεί.
- *Παρουσίαση:* Τα WCMS προσφέρουν μια μεγάλη ποικιλία δυνατοτήτων, ώστε να καταστεί ο ιστοχώρος δυναμικός και διαδραστικός. Για παράδειγμα, δημιουργούν αυτόματα την πλοήγηση του ιστοχώρου, διαβάζοντας τη δομή κατευθείαν από τη βάση δεδομένων ή διευκολύνουν την υποστήριξη πολλαπλών browsers.

### 3.3 Στόχοι ενός Συστήματος Διαχείρισης Διαδικτυακού Περιεχομένου

Ένα WCMS θα πρέπει να καλύπτει τους εξής στόχους για να καταστεί λειτουργικό:

- Να υποστηρίζει την αποκεντρωμένη συγγραφή, έτσι ο κάθε χρήστης να μπορεί να συντάσσει και να ανανεώνει το δικό του περιεχόμενο χωρίς να καλεί το τμήμα IT, ενώ θα πρέπει να αποστέλλεται αυτόματα και στους υπεύθυνους για την έγκρισή της δημοσίευσής του.
- Να συμβάλλει στην απλούστευση της διαδικασίας συγγραφής και ανανέωσης του περιεχομένου του ιστοχώρου.
- Να παρέχει υψηλό επίπεδο ασφάλειας, καθορίζοντας για παράδειγμα, ποιος μπορεί να δημοσιεύσει τι.
- Να παρέχει υψηλή ευελιξία, μέσω της εύκολης ανανέωσης όλων των σελίδων, ώστε να αφομοιώνεται άμεσα μια αλλαγή στο εταιρικό προφίλ.
- Να συμβάλλει στην ενσωμάτωση ακριβέστερης πληροφόρησης και να περιορίζει το φαινόμενο της εμφάνισης διπλής εγγραφής μιας πληροφορίας.
- Να βελτιώνει την εμπειρία πλοήγησης στον ιστοχώρο, τόσο για τους συγγραφείς περιεχομένου, όσο και για τους απλούς επισκέπτες.
- Να παρέχει ένα πλούσιο διαδραστικό περιβάλλον για τον πελάτη.

- Να έχει δυνατότητα διευκόλυνσης στην επικοινωνία του marketing.
- Να παρέχει υποστήριξη λειτουργιών e-commerce.
- Να συμβάλλει στη μείωση του κόστους συντήρησης, ανάπτυξης, δημοσίευσης και υποστήριξης πελατών.

### 3.4 Είδη Συστημάτων Διαχείρισης Διαδικτυακού Περιεχομένου

Μπορούμε να διακρίνουμε τρεις βασικές κατηγορίες WCMS ανάλογα με κάποια βασικά χαρακτηριστικά τους:

Με βάση το που βρίσκεται ο χώρος αποθήκευσης και διαχείρισης της βάσης δεδομένων του WCMS σε:

- **Application Service Provider (ASP)**, δηλαδή συστήματα Υποστήριξης Παρόχου Υπηρεσίας, όπου ο κατασκευαστής φιλοξενεί όλα τα δεδομένα και το λογισμικό στους server της εταιρίας του. Με αυτόν τον τρόπο δεν αναλαμβάνονται κόστη για την αγορά software και hardware για την υποστήριξη του συστήματος, καθώς επίσης μειώνονται και οι δαπάνες σε τεχνικούς πόρους.
- **Licensed**, δηλαδή συστήματα με παροχή άδειας, όπου ο ρόλος του παρόχου τελειώνει με την πώληση του προϊόντος, δηλαδή την παροχή της άδειας χρήσης. Στη συνέχεια ο χρήστης είναι υπεύθυνος για την εγκατάσταση, ρύθμιση και συντήρηση του WCMS. Αυτό το είδος WCMS ενδείκνυται για οργανισμούς, οι οποίοι διατηρούν ήδη στο χώρο τους κάποιο είδος παρόμοιας υπηρεσίας, όπως σύστημα Διαχείρισης Εξυπηρέτησης Πελατών (CRM).

Με βάση το είδος του παρόχου σε:

- **Commercial**, το λογισμικό παρέχεται από κερδοσκοπικές ή μη εταιρίες. Πωλείται και υποστηρίζεται τεχνικά από αυτές.
- **Open Source**. Στην περίπτωση WCMS ανοικτού κώδικα, το λογισμικό δημιουργείται και συντηρείται από έναν ανεπίσημο, ανιδιοτελή συνεργάτη μια κοινότητας χρηστών. Στη συνέχεια το λογισμικό αυτό διανέμεται δωρεάν στα μέλη της κοινότητας.

### 3.5 Συστήματα Διαχείρισης Διαδικτυακού Περιεχομένου Ανοιχτού Κώδικα

Σήμερα, τα Συστήματα Διαχείρισης Διαδικτυακού Περιεχομένου ανοικτού κώδικα (open-source WCMS) έχουν εξελιχθεί πλέον αρκετά, ώστε να μπορούν να ανταγωνισθούν επάξια τα εμπορικά πακέτα. Προς αυτή την κατεύθυνση συνέβαλε η ανάπτυξη του διαδικτύου, η καθιέρωση του Apache και της Linux στην αγορά, αλλά και η υποστήριξη από μεγάλους προμηθευτές, όπως η IBM.

Τα WCMS ανοικτού κώδικα χρησιμοποιούν την άδεια GNU GPL. Πρόκειται για μια άδεια copyleft, σύμφωνα με την οποία οποιοδήποτε άτομο λάβει αντίγραφο του κώδικα που διέπεται από αυτήν την άδεια αποκτά το δικαίωμα ελεύθερης αναπαραγωγής, προσαρμογής και διανομής του κώδικα, υπό την προϋπόθεση οι αντιγραφές και προσαρμογές που θα προκύψουν να διέπονται από την ίδια άδεια.

### 3.5.1 Πλεονεκτήματα

Σημαντικά πλεονεκτήματα ενός open source WCMS είναι:

- **Χαμηλό κόστος.** Τα συστήματα ανοιχτού κώδικα διατίθενται δωρεάν. Παράλληλα, ενδεχόμενη τεχνική υποστήριξή τους προσφέρεται φθηνότερα από την αντίστοιχη τεχνική υποστήριξη εμπορικών πακέτων. Το χαμηλό κόστος τα καθιστά ιδιαίτερα συμβατά για μικρούς ιστοχώρους, μη κερδοσκοπικούς οργανισμούς και κυβερνητικά τμήματα.
- **Δυνατότητα ανάληψης υψηλότερου κόστους υπηρεσιών,** δεδομένου ότι το λογισμικό παρέχεται δωρεάν. Τέτοιες υπηρεσίες θα μπορούσαν να αφορούν στην παραμετροποίηση του λογισμικού, την ενσωμάτωση του συστήματος με άλλα επιχειρησιακά συστήματα, την ανάπτυξη πλαισίων και σελίδων μορφοποίησης, την επιθεώρηση και ανακατασκευή περιεχομένου, τη διενέργεια τεστ χρηστικότητα, την εκπαίδευση των συγγραφέων και χρηστών.
- **Ευκολία παραμετροποίησης.** Σε αντίθεση με τα εμπορικά πακέτα, όπου ο κώδικας είναι μυστικός, ως πηγή ανταγωνιστικού πλεονεκτήματος, στα open source όλος ο κώδικας είναι εξ ορισμού προσβάσιμος. Αυτό επιτρέπει στους προγραμματιστές επέμβουν στο σύστημα και να το προσαρμόσουν με τέτοιο τρόπο ώστε να καλύπτει εξατομικευμένες ανάγκες
- **Ανοιχτές πλατφόρμες.** Η ανάπτυξη των WCMS ανοιχτού κώδικα με τη χρήση ανοιχτών εργαλείων (open tools), όπως οι γλώσσες PHP, Perl, Java και Unix, επιτρέπουν την μείωση της δέσμευσης με συγκεκριμένους προμηθευτές και πακέτα λογισμικού.
- **Κοινή πηγή άντλησης πόρων μεταξύ οργανισμών.** Δεδομένου του ότι πρόκειται για ανοιχτό κώδικα κάποιοι οργανισμοί, όπως το δημόσιο, θα μπορούσαν να συγκεντρώσουν τους πόρους τους και να μοιραστούν τα έξοδα παραμετροποίησης, εφαρμόζοντας τις λύσεις που υιοθετούνται από έναν δημόσιο οργανισμό σε κάποιον άλλο.
- **Εύκολη ενσωμάτωση.** Η ενσωμάτωση των WCMS με άλλα λογισμικά καθίσταται πολύ εύκολη μέσω του συνδυασμού των ανοιχτών πλατφόρμων και της απλότητας της παραμετροποίησης. Επίσης, πολλές νέες τεχνολογίες που έχουν αναπτυχθεί με τη λογική της ενσωμάτωσης, όπως η XML, αποτελούν μέρος του εξοπλισμού πολλών πακέτων ανοιχτού κώδικα.
- **Υποστήριξη από την κοινότητα.** Σε αντίθεση με τα εμπορικά πακέτα, τα WCMS ανοιχτού κώδικα υποστηρίζονται από μια ενεργή κοινότητα χιλιάδων προγραμματιστών.
- **Δυνατότητα δοκιμής πριν την αγορά.** Στα λογισμικά ανοιχτού κώδικα, ο ενδιαφερόμενος έχει την δυνατότητα να κατεβάσει το λογισμικό, να το αξιολογήσει και στην συνέχεια να αποφασίσει αν καλύπτει τις ανάγκες του ή όχι.
- **Άμεση επίλυση προβλημάτων.** Στην περίπτωση χρήσης ενός εμπορικού πακέτου η διαδικασία για τη διόρθωση κάποιου σφάλματος που ενδεχομένως να εμφανιστεί περιλαμβάνει την ειδοποίηση του προμηθευτή και την αναμονή μέχρι να αντιμετωπιστεί. Το πιο δυσοίωνα σενάριο θα μπορούσε να είναι η αναμονή για τη διάθεση μιας νέας έκδοσης που θα αντιμετωπίζει το σφάλμα, η οποία μπορεί να διαρκέσει και ένα χρόνο.

Με τη χρήση ενός WCMS ανοιχτού κώδικα η επίλυση του προβλήματος μπορεί να γίνει είτε με την αναφορά του προβλήματος στην κοινότητα, η οποία με τη σειρά της ανακατασκευάζει ένα patch αντιμετώπισης μέσα σε μερικές μέρες, είτε με τη λύση του προβλήματος από τον ίδιο τον ενδιαφερόμενο. Ας μην ξεχνάμε ότι παρέχεται πλήρης πρόσβαση στον πηγαίο κώδικα, αρκεί λοιπόν ο ενδιαφερόμενος να έχει τις απαιτούμενες τεχνικές γνώσεις για την επίλυση του προβλήματος.

- **Μελλοντική εξασφάλιση.** Ο υψηλός κορεσμός της αγοράς εμπορικών πακέτων διαχείρισης περιεχομένων εκτιμάται από τους ειδικούς ότι τα επόμενα χρόνια θα οδηγήσει σε μια γενική ανακατάταξη μεταξύ των ανταγωνιστικών επιχειρήσεων, η οποία και θα οδηγήσει σε επικράτηση λίγων, μεγάλων προμηθευτών. Στην περίπτωση παύσης λειτουργίας ενός προμηθευτή του αντίστοιχου πακέτου λογισμικού, παύει και η τεχνική υποστήριξη προς τον πελάτη. Το πρόβλημα αυτό εξαλείφεται με τη χρήση πακέτων ανοιχτού κώδικα. Ακόμη και αν το λογισμικό απορριφθεί, ο κώδικας παραμένει διαθέσιμος, ώστε να παρέχεται η δυνατότητα στο χρήστη να συνεχίσει ο ίδιος τις απαραίτητες αναβαθμίσεις. Συμπερασματικά, θα λέγαμε ότι τα open source WCMS είναι εξασφαλισμένα σε βάθος χρόνου απέναντι στην αβεβαιότητα της αγοράς.

### 3.5.2 Μειονεκτήματα

Τα διάφορα προβλήματα με τα οποία μπορεί να έρθει αντιμέτωπος ένας χρήστης Συστήματος Διαχείρισης Περιεχομένου ανοιχτού κώδικα είναι:

- Η δωρεάν διάθεση δε συνεπάγεται και την ανάληψη μηδενικού κόστους. Ένα μεγάλο μέρος του κόστους που αφορά στη χρήση ενός Συστήματος Διαχείρισης Περιεχομένου πηγάζει από την ίδια την εφαρμογή του συστήματος και τις παραμετροποιήσεις του που ακολουθούν.
- Πιθανή αδυναμία υποστήριξης σε επίπεδο επιχείρησης. Τα περισσότερα open source συστήματα είναι πιθανό να μην μπορούν να καλύψουν κάποιες λειτουργίες απαραίτητες για επιχειρήσεις. Για το λόγο αυτό στοχεύουν σε μικρού έως μετρίου μεγέθους εφαρμογές.
- Απουσία εμπορικής υποστήριξης. Δεν παρέχουν εμπορική υποστήριξη και εγγυήσεις όσον αφορά την υποστήριξη, δεδομένου του ότι στηρίζονται στην πρωτοβουλία μιας διαδικτυακής κοινότητας.
- Λιγότερο εξελιγμένα. Σε σχέση με τα εμπορικά πακέτα, τα οποία λόγω του υψηλού ανταγωνισμού στην αγορά, αναπτύσσουν υψηλούς ρυθμούς καινοτομίας, τα λογισμικά ανοιχτού κώδικα είναι πιθανό να μην μπορούν να ακολουθήσουν.
- Μικρή χρηστικότητα. Τα πακέτα ανοιχτού κώδικα συνήθως επικεντρώνονται στην τεχνική αρχιτεκτονική και στην ομάδα χαρακτηριστικών με αποτέλεσμα να υπολείπονται σε χρηστικότητα.
- Έλλειψη τεκμηρίωσης. Δεδομένου του ότι η αποτελεσματική τεκμηρίωση απαιτεί υψηλούς πόρους και ικανούς συγγραφείς, είναι πιθανό τα open source WCMS να υπολείπονται τεκμηρίωσης και υποστηρικτών πληροφοριών, αφού βασίζονται σε διαδικτυακές κοινότητες.
- Κίνδυνος υπερεπένδυσης. Παρέχοντας τη δυνατότητα της εύκολης παραμετροποίησης, είναι πιθανό να υπάρξει υπερεπένδυση κατά την ανάπτυξη του προϊόντος ακολουθώντας μια στρατηγική υπερβολικής παραμετροποίησης.

### 3.5.3 Οφέλη από τη συμμετοχή στην ανάπτυξη open source λογισμικού

Ο προσπορισμός κέρδους δεν αποτελεί βασικό στόχο της συμμετοχής στην ανάπτυξη λογισμικού ανοικτού κώδικα. Αυτό βέβαια δεν αποκλείει τελικά την πιθανότητα η διαδικασία αυτή να αποβεί κερδοφόρα. Βασικοί λόγοι να συμμετάσχει κανείς στην ανάπτυξη ενός open source λογισμικού, είναι:

- Η αποκτηθείσα εμπειρία από τη συμμετοχή στο project, αλλά και από την συνεργασία με άλλους προγραμματιστές.
- Η φήμη που δύναται να αποκτήσει κανείς συμμετέχοντας σε αυτή τη διαδικασία, στην περίπτωση που το λογισμικό γίνει αποδεκτό από το ευρύ κοινό και καταστεί επιτυχημένο.
- Η ταχύτερη ανάπτυξη ενός λογισμικού, χρήστης του οποίου ενδεχομένως να είναι και ο ίδιος ο συμμετέχων στην ανάπτυξη του.
- Η ηθική ικανοποίηση, ότι συνεισφέρει κανείς χωρίς προσωπικό όφελος.

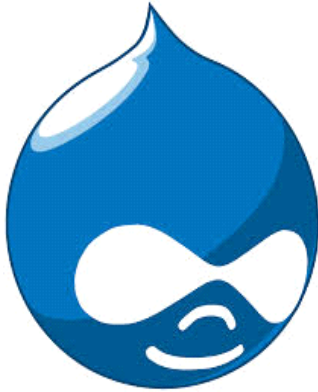
Σε περίπτωση επαναπροσδιορισμού των στόχων, το λογισμικό μπορεί πάντα να πωληθεί σε μορφή έτοιμη προς χρήση, με τον πηγαίο κώδικα, καθώς και με την παροχή υπηρεσιών που συνδέονται με το λογισμικό.

### **3.6 Πλαίσιο Διαχείρισης Περιεχομένου (Content Management Framework – CMF)**

Το πλαίσιο διαχείρισης περιεχομένου είναι ένας διαμεσολαβητής προγραμματισμού εφαρμογών, που συμβάλει στη δημιουργία ενός προσαρμοσμένου συστήματος διαχείρισης δεδομένων. Το CMF είναι ένα πλαίσιο που παρέχει στους προγραμματιστές εργαλεία, όπως η ροή εργασίας, η προσωποποίηση και η κατηγοριοποίηση, για να φτιάξουν ένα προϊόν.



## Κεφάλαιο Τέταρτο: Drupal



Εικόνα 2: Λογότυπο Drupal

### 4.1 Ορισμός

Το αποτελεσματικό web design προϋποθέτει την εξισορρόπηση μεταξύ ευελιξίας και απλότητας. Ένα υπεραπλουστευμένο σύστημα είναι πιθανό να μη μπορεί να χρησιμοποιηθεί για την εξυπηρέτηση πολλαπλών σκοπών, παρά μόνο ενός, ενώ αντίστοιχα ένα ιδιαίτερα ευέλικτο σύστημα ενδεχομένως να απαιτεί εξειδικευμένες γνώσεις, οι οποίες είναι δύσκολο να αποκτηθούν από ένα νέο χρήστη.

Το Drupal ως ένα ανοιχτού κώδικα αρθρωτό σύστημα διαχείρισης περιεχομένου (Content Management System, CMS), γραμμένο στη γλώσσα προγραμματισμού PHP, το οποίο διανέμεται σύμφωνα με τη Γενική Άδεια Δημόσιας Χρήσης GNU (GNU - General Public License), επιδιώκει να συνδυάσει τους δύο παραπάνω αντικρουόμενους στόχους. Παρέχει στους χρήστες τα εργαλεία που χρειάζονται για να δημιουργήσουν μια εξατομικευμένη λύση διαχείρισης δεδομένων, έχοντας παράλληλα δημιουργήσει ένα βασικό πακέτο πάνω στο οποίο θα στηριχθούν, με τις περισσότερες εργασίες να γίνονται με ελάχιστο έως καθόλου προγραμματισμό. Είναι δομημένο με τέτοιο τρόπο, ώστε να επιτρέπει στο διαχειριστή την αποτελεσματική οργάνωση του περιεχομένου, την προσαρμογή της εμφάνισης, την αυτοματοποίηση διαχειριστικών εργασιών, καθώς και την διαχείριση των επισκεπτών.

Το Drupal μπορεί να εκτελεστεί σε οποιαδήποτε πλατφόρμα, αρκεί να πληρούνται οι εξής δύο προϋποθέσεις:

- Να υποστηρίζεται ένας web server, ο οποίος να έχει τη δυνατότητα να υποστηρίξει τη γλώσσα προγραμματισμού PHP έκδοσης 5.2 ή νεότερη, όπως ο Apache, IIS, Lighttpd ή nginx και
- Να υπάρχει μια βάση δεδομένων, απαραίτητη για την αποθήκευση περιεχομένου και ρυθμίσεων, όπως η MySQL, SQLite ή PostgreSQL.

### 4.2 Ιστορική εξέλιξη

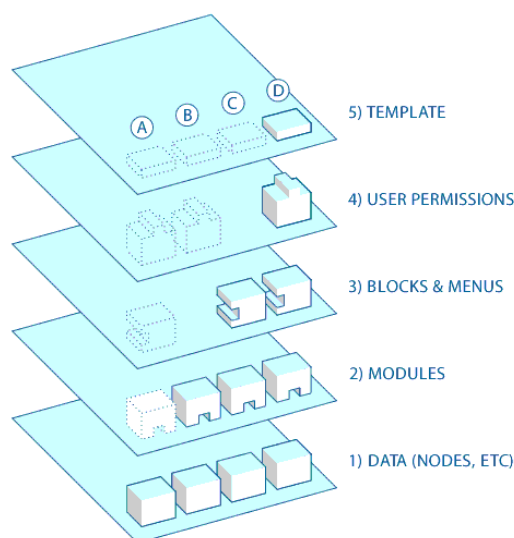
Δημιουργός του Drupal υπήρξε ο Dries Buytaert. Όντας φοιτητής Ολλανδικού Πανεπιστήμιου ασχολήθηκε με το project της δημιουργίας ενός συστήματος διαμοιρασμού νέων και γεγονότων με

τους φίλους του, γνωστό ως σύστημα πίνακα ανακοινώσεων (BBS, Bulletin Board System). Το project αυτό κατέληξε να μετατραπεί σε εγχείρημα ανοιχτού κώδικα το 2001, το οποίο υιοθετήθηκε άμεσα και ένθερμα από την κοινότητα, καθιστώντας το σήμερα μια από τις πλέον δυναμικές και πλούσιες σε δυνατότητες πλατφόρμες στο διαδίκτυο.

Αξίζει να σημειωθεί ότι από το Μάιο του 2006 μέχρι τον Απρίλιο του 2007, το επίσημο site του Drupal κατέγραψε περισσότερα από 600.000 downloads από χρήστες. Σήμερα είναι διαθέσιμη η έκδοση 7, ενώ σύντομα αναμένεται η διάθεση της νέας, βελτιωμένης έκδοσης Drupal 8, με περισσότερα από 200 νέα χαρακτηριστικά, συμπεριλαμβανομένων νέων δυνατοτήτων εμφάνισης δεδομένων σε mobile συσκευές, δημιουργίας APIs και προσαρμογής σε πολυγλωσσικές ανάγκες.

### 4.3 Διάγραμμα Ροής “The Drupal flow”

Πριν προχωρήσουμε στην περαιτέρω ανάλυση των δομικών συστατικών του Drupal, θεωρούμε σκόπιμο να αναφερθούμε στον τρόπο ροής της πληροφορίας μεταξύ των διαφόρων στρωμάτων τους συστήματος. Υπάρχουν πέντε βασικά στρώματα, η διάταξη των οποίων παρουσιάζεται στην εικόνα που ακολουθεί:



Εικόνα 3: "The Drupal Flow"

- Στη βάση του συστήματος υπάρχει το σύνολο των κόμβων (“nodes”), που αποτελούν την «δεξαμενή» των δεδομένων. Για να παρουσιαστεί οποιαδήποτε πληροφορία στο site, θα πρέπει προηγουμένως να έχει γίνει η εισαγωγή της στο σύστημα με τη μορφή δεδομένων.
- Στο αμέσως επόμενο στρώμα υπάρχουν τα modules, δηλαδή λειτουργικά πρόσθετα, τα οποία είτε αποτελούν μέρος του πυρήνα (“core”), είτε είναι στοιχεία τα οποία έχουν δημιουργηθεί με τη συνεισφορά της κοινότητας του Drupal. Τα modules είναι σχεδιασμένα έτσι ώστε να ενισχύουν τη λειτουργικότητα του Drupal core, επιτρέποντας στο χρήστη να προσαρμόζει τα δεδομένα, να εγκαθιστά πρόσθετα ηλεκτρονικού εμπορίου, να καθορίζει την εφαρμογή φίλτρων σχετικά με τη παρουσίαση και τη διαμόρφωση του περιεχομένου.
- Στη συνέχεια υπάρχουν τα blocks και menus. Με το όρο blocks αναφερόμαστε σε εργαλεία του Drupal, τα οποία παρουσιάζουν το αποτέλεσμα κάποιου module ή δημιουργούνται για να παρουσιάσουν οτιδήποτε επιθυμεί ο χρήστης και στη συνέχεια τοποθετούνται σε οποιοδήποτε

σημείο της διάταξης του προτύπου. Τα blocks μπορούν να ρυθμιστούν ως προς τον τρόπο παρουσίασης του αποτελέσματος, καθώς και να ρυθμιστούν έτσι, ώστε να παρουσιάζονται σε συγκεκριμένες σελίδες ή για συγκεκριμένους μόνο, προκαθορισμένους χρήστες. Τα menus αποτελούν πλοηγούς του Drupal και προσδιορίζουν το περιεχόμενο που θα παρουσιάζεται σε κάθε καθορισμένη διαδρομή (σχετικό url).

- Στο επόμενο στάδιο έχουμε τις άδειες χρηστών. Στο στρώμα αυτό δημιουργούνται οι ρυθμίσεις που προσδιορίζουν τι μπορούν να δουν και να κάνουν οι διάφορες κατηγορίες χρηστών.
- Στο πέμπτο και τελευταίο στρώμα υπάρχουν οι θεματικές παραλλαγές (“themes”), δηλαδή θέματα που μπορούν να εγκατασταθούν στην πλατφόρμα, προκειμένου να προσαρμόσουν τον σχεδιασμό και τη διάταξη του site. Τα θέματα δημιουργούνται κυρίως με XHTML και CSS, αλλά και με το συνδυασμό ορισμένων μεταβλητών PHP. Επιπλέον, σε κάθε θέμα συμπεριλαμβάνεται ένα σύνολο λειτουργιών, οι οποίες μπορούν να χρησιμοποιηθούν, ώστε να παρακαμφθούν βασικές λειτουργίες των modules παρέχοντας τη δυνατότητα πλήρους ελέγχου πάνω στον τρόπο παρουσίασης του αποτελέσματος τους.

### 4.4 Drupal Core

Το Drupal αποτελείται από ένα βασικό στρώμα ή πυρήνα (“core”), το οποίο παρέχει τα βασικά του χαρακτηριστικά, καθώς και τη δυνατότητα υποστήριξης αρθρωτών μονάδων, οι οποίες ενισχύουν τη λειτουργικότητά του. Τα βασικά συστατικά του «πυρήνα» παρέχουν δυνατότητες δημιουργίας και διαχείρισης:

- Περιεχομένου.
- Upload/Download αρχείων.
- Μενού, όπως να δημιουργούνται μενού με πολλαπλά επίπεδα.
- Ταξινόμησης, όπως υποβολή, αναθεώρηση και κατηγοριοποίηση της ύλης.
- Λογαριασμών χρηστών, όπως να εγγράφονται ως χρήστες και να διαχειρίζονται το λογαριασμό τους, να τροποποιούν και να βλέπουν προσωπικά προφίλ.
- Ρόλων και αδειών, όπως να καθορίζονται ρόλοι χρηστών, με τον ορισμό αδειών για τη χρησιμοποίηση επιλεγμένων χαρακτηριστικών του site, να χρησιμοποιούνται κανόνες πρόσβασης για την άρνηση πρόσβασης στο site σε συγκεκριμένα ονόματα χρηστών, διευθύνσεις email ή διευθύνσεις IP.
- Blogs και forums, όπως να υποβάλλουν σχόλια ή να λαμβάνουν μέρος σε ομάδες συζητήσεων.
- Online δημοσκοπήσεων.
- Συνεργατικών συγγραφικών έργων.
- Ροών RSS, καθώς και να συλλέγεται περιεχόμενο από ροές RSS άλλων site.
- Των θεματικών παραλλαγών.

## 4.5 Modules

Για τις περιπτώσεις, όπου το Drupal core αδυνατεί να καλύψει υψηλότερες απαιτήσεις λειτουργικότητας η κοινότητα του Drupal έχει δημιουργήσει μια πληθώρα πρόσθετων μονάδων, modules, οι οποίες και διατίθενται δωρεάν μέσω του επίσημου site.

Τα modules καλύπτουν μια πληθώρα γενικών κατηγοριών, όπως e-commerce, χαρακτηριστικά ροής εργασιών, γκαλερί φωτογραφιών, χάρτες ιστοτόπου για το Google, αντικείμενα Amazon, διαχείριση λιστών email. Μερικά από τα πιο ευρέως χρησιμοποιούμενα modules είναι:

Ubercart: Πρόκειται για ένα πλήρως εξοπλισμένο module, που παρέχει όλους τους απαραίτητους μηχανισμούς για την πώληση προϊόντων (φυσικών και μη), για την πληρωμή μέσω πιστωτικών καρτών και για την εξέλιξη της παραγγελίας.

Views: Πρόκειται για ένα εύχρηστο module, το οποίο μπορεί να χρησιμοποιηθεί για την εξαγωγή πληροφοριών από τη βάση δεδομένων του Drupal και την άμεση παρουσίαση των αποτελεσμάτων της στο site.

Content Construction Kit (CCK): Το CCK παρέχει τη δυνατότητα δημιουργίας και διαχείρισης τυπικών τύπων περιεχομένων στο site.

Google Analytics: Παρέχεται η δυνατότητα στο χρήστη της εγκατάστασης του Google Analytics στο site, το οποίο αποτελεί μια δωρεάν υπηρεσία, η οποία δίνει τη δυνατότητα παρακολούθησης του αριθμού των επισκεπτών του site, καθώς και εξατομικευμένων στοιχείων τους, όπως από ποια χώρα είναι, ποιες σελίδες επισκέφθηκαν, πόσο χρόνο παρέμειναν στη σελίδα.

## 4.6 Themes

Οι θεματικές παραλλαγές, αποτελούν ένα εργαλείο του Drupal, με το οποίο καθορίζεται η δομή μιας ιστοσελίδας καθώς και πτυχές της εμφάνισής τους. Πιο συγκεκριμένα, ο χρήστης έχει τη δυνατότητα να επιλέξει:

- Τον αριθμό των στηλών με πληροφορίες που θα παρουσιάζονται.
- Την ύπαρξη ή μη banner στην κορυφή της σελίδας.
- Την ύπαρξη ή μη υποσέλιδου.
- Το σημείο όπου θα εμφανίζονται τα μενού πλοήγησης.
- Τα χρώματα και το φόντο της σελίδας, καθώς και
- Την εμφάνιση γραφικών, όπως το λογότυπο της εταιρίας.

## 4.7 Κριτική Διερεύνηση Drupal

Βασικό χαρακτηριστικό του Drupal είναι η ευελιξία και η προσαρμοστικότητα του, σε συνδυασμό πάντα με την εγγενή του επεκτασιμότητα, χάρη στο hook-based API του, και η ασφάλεια. Έχει υιοθετηθεί από πληθώρα οργανισμών σε παγκόσμιο επίπεδο, όπως NASA, Sony Ericsson, Linux Foundation, Nokia, Warner Bros, 20<sup>th</sup> Century Fox, Intel Corporation, ενώ το 2009 κέρδισε για τρίτη φορά το βραβείο καλύτερου CMS.

Στις βασικές δυνατότητες του Drupal συγκαταλέγονται:

- Η επεκτάσιμη αρθρωτή αρχιτεκτονική του, με πάνω από 16.000 modules και με ευέλικτο API.

- Η ύπαρξη ενός ισχυρού συστήματος ταξινόμησης περιεχομένου, με δυνατότητα ελεύθερων «ετικετών», που καθιστούν δυνατή την κατηγοριοποίηση του περιεχομένου σε πραγματικό χρόνο.
- Η δυνατότητα διαχείρισης των διαφόρων εκδόσεων περιεχομένου, καθώς και επαναφοράς αναθεωρήσεων.
- Η δυνατότητα υποστήριξης ομάδων και ρόλων χρηστών με τη διαχείριση αδειών χρήσης.
- Η δυνατότητα να υποστηρίζει πολλαπλούς ιστότοπους, με ταυτόχρονη λειτουργία ενδοδικτύου με εξωδίκτυο και ροή περιεχομένου μεταξύ τους.
- Η ισχυρή υποστήριξη SEO (Search Engine Optimization), έχοντας αναπτυχθεί από τη βάση του φιλικό προς τις μηχανές αναζήτησης.
- Η δυνατότητα πολλαπλών δομών εμφάνισης και μενού, με ενσωμάτωση αντικειμένων στο περιεχόμενο και αυτόματη διαχείριση flash, video και φωτογραφιών από διαφορετικές πηγές.
- Τέλος, μέσω ειδικού χειρισμού της λανθάνουσας μνήμης, προσφέρει έξυπνο caching, για επίτευξη μεγάλης ταχύτητας εμφάνισης των σελίδων.

#### 4.7.1 Βασικά Πλεονεκτήματα Drupal

Το Drupal αποτελεί ένα ελεύθερο CMS, χωρίς απαιτήσεις άδειας για την εγκατάσταση ή τη χρήση του, με απεριόριστη επεκτασιμότητα και ευελιξία και με μια μεγάλη κοινότητα που το υποστηρίζει και φροντίζει συνεχώς για τη εξέλιξη και ανάπτυξη του.

Παράλληλα, παρέχει μια τεράστια ποικιλία εφαρμογών για εμπορικούς, διαφημιστικούς και κοινωνικούς ιστότοπους, φιλικό προς τις μηχανές αναζήτησης, παρέχοντας υψηλό επίπεδο ασφάλειας, μέσω συνεχούς ενημέρωσης των εκδόσεων, καθώς και υψηλή αξιοπιστία στη λειτουργία του, ακόμη και όταν εξυπηρετεί χιλιάδες χρήστες ταυτόχρονα.

Τέλος, είναι ένα CMS φιλικό προς τις επιχειρήσεις, παρέχοντας δυνατότητα κρυπτογράφησης δεδομένων και αυτόματη διαχείριση δεδομένων μεγάλου όγκου.



Εικόνα 4: Λογότυπα Δημοφιλέστερων CMS

#### 4.7.2 Drupal vs. Joomla

Το Joomla αποτελεί ένα ανοιχτού κώδικα CMS, απλό στην εγκατάσταση, με έτοιμα εμπορικά θέματα, βασικό ανταγωνιστή του Drupal στον ελληνικό χώρο. Οι περισσότερες εταιρίες διαφημιστικών ιστοσελίδων στην Ελλάδα χρησιμοποιούν Joomla.

Βασικά μειονεκτήματα του Joomla είναι:

- Η αδυναμία ενσωμάτωσης διαδικτυακών εφαρμογών και υποστήριξη SEO.
- Η περιορισμένη διαχείριση πολυγλωσσικότητας και αδειών χρήσης.
- Η περιορισμένη κατηγοριοποίηση περιεχομένου.
- Η απουσία Enterprise δυνατοτήτων.

Συμπερασματικά, θα λέγαμε πως το Joomla είναι κατάλληλο για απλούστερες εγκαταστάσεις, ενώ δεν απαιτεί γνώσεις σχεδίασης και προγραμματισμού. Αντίθετα το Drupal είναι κατάλληλο για σύνθετες εγκαταστάσεις διαχείρισης περιεχομένου και κάθετες εφαρμογές, απαιτεί περισσότερο εξειδικευμένους σχεδιαστές και προγραμματιστές, ενώ παραμένει το πιο εργονομικό και εύχρηστο σύστημα για τον τελικό χρήστη, μετά από επαγγελματική προσαρμογή.

### 4.7.3 Drupal vs. Wordpress

Το Wordpress αποτελεί και αυτό ένα ανοιχτού κώδικα CMS, απλό στην εγκατάσταση, με απλές επεκτάσεις και θέματα, δωρεάν υπηρεσίες φιλοξενίας, που παρέχει όμως επαρκές SEO, σε αντίθεση με το Joomla.

Βασικά μειονεκτήματα του Wordpress είναι:

- Η ύπαρξη δύσχρηστου API, ανεπαρκές για την ανάπτυξη πολύπλοκων λύσεων.
- Προσανατολισμένο για διαχείριση ιστολογίου.
- Ακατάλληλο για κάθετες λύσεις.
- Η απουσία Enterprise δυνατοτήτων.

Συμπεραίνουμε λοιπόν ότι, ενώ το Wordpress επαρκεί για την εγκατάσταση καλαίσθητων ιστολογίων και δεν απαιτεί ιδιαίτερες γνώσεις διαδικτυακής τεχνολογίας, το Drupal πρέπει να προτιμάται αν υπάρχουν σχέδια επέκτασης και ενσωμάτωσης περισσότερων λειτουργιών στο μέλλον.

## 4.8 Λόγοι επιλογής του Drupal

- Προηγμένος έλεγχος των διευθύνσεων URL: Το Drupal σε αντίθεση με άλλα ανοιχτού κώδικα πακέτα, όπως το Wordpress παρέχει τη δυνατότητα ακριβούς ελέγχου πάνω στη δομή των διευθύνσεων URL. Σε κάθε στοιχείο του περιεχομένου (node) είναι δυνατό να δοθεί μια διεύθυνση URL (URL alias). Για παράδειγμα, το module Global Redirect, έχει τη δυνατότητα να επαναπροσανατολίζει αυτόματα 301 διευθύνσεις από το εσωτερικό του Drupal σε URL alias.
- Custom τύποι περιεχομένου και Views: Το Drupal παρέχει εξειδικευμένα modules για την δημιουργία νέων τύπων περιεχομένου, όπως το Content Construction Kit (CCK), χωρίς την ανάγκη σύνταξης νέου κώδικα. Επιπλέον, παρέχει modules για την καλύτερη εμφάνιση και παρουσίαση του περιεχομένου, όπως το Views.
- Έλεγχος αναθεώρησης: Παρέχει τη δυνατότητα κάθε μετατροπή σε μια σελίδα, να αποθηκεύεται ως μια νέα έκδοση της σελίδας. Ουσιαστικά, λοιπόν δίνει τη δυνατότητα στο χρήστη να έχει πρόσβαση ή και να επαναφέρει παλαιότερες σελίδες, όταν χρειαστεί.

- Taxonomy: Διαθέτει ένα ισχυρό σύστημα κατηγοριοποίησης, το οποίο επιτρέπει στο χρήστη να οργανώσει το περιεχόμενο όπως επιθυμεί με εύκολο και κατανοητό τρόπο.
- Διαχείριση χρηστών: Παρέχει ιδιαίτερα υψηλή λειτουργικότητα και αυτονομία, αναφορικά με τους ρόλους χρηστών και τον έλεγχο πρόσβασής τους στο περιεχόμενο του site. Έχουμε τη δυνατότητα να δημιουργήσουμε όσους ρόλους χρηστών επιθυμούμε παρέχοντας τους το κατάλληλο επίπεδο πρόσβασης στο περιεχόμενο.
- Τίτλοι σελίδων και Metatags: Με το Page title module μπορούμε να αποκτήσουμε τον έλεγχο των HTML <title> στοιχείων. Αντίστοιχα, το Metatags module μας δίνει τον έλεγχο επί των ετικετών περιγραφής των σελίδων.
- Άριστη τεκμηρίωση: Αποτελεί το πιο επαρκώς τεκμηριωμένο open source πακέτο λογισμικού, με επίσημα εγχειρίδια, tutorials, blogs, videos που ανανεώνονται συνεχώς.
- Drupal Cookbooks: Λόγω της ύπαρξης μιας μεγάλης διαδικτυακής κοινότητας, η οποία αναπτύσσει συνεχώς νέα modules, είναι σχεδόν απίθανο να μην έχει κανείς πρόσβαση σε κάποια συγκεκριμένη λειτουργία που χρειάζεται και δεν είναι διαθέσιμη στο βασικό πακέτο.
- Μεγάλη και φιλική κοινότητα: Το Drupal αποτελεί ένα από τα πιο ευρέως χρησιμοποιούμενα CMF, με μια εξαιρετικά ενεργή κοινότητα να το υποστηρίζει και ένα ιδιαίτερα ενεργό forum για τη διατύπωση ερωτημάτων και σχολίων.





## Κεφάλαιο Πέμπτο: LAMP (Software Bundle)



Εικόνα 5: Λογότυπο LAMP

### 5.1 Εννοιολογικός Προσδιορισμός

Το LAMP αποτελεί ένα πακέτο, εξ' ολοκλήρου, ελεύθερου λογισμικού, ανοιχτού κώδικα. Είναι κατάλληλο για την κατασκευή υψηλής διαθεσιμότητας δυναμικών ιστοσελίδων, ικανό να εξυπηρετήσει δεκάδες χιλιάδες αιτήσεις ταυτόχρονα. Στην ουσία αποτελεί ένα ακρόνυμο που βασίζεται στα πρώτα γράμματα από τα τέσσερα βασικά συστατικά στοιχεία του πακέτου:

- **Linux**, το λειτουργικό σύστημα
- **Apache HTTP Server**, ο web server
- **MySQL, MariaDB ή MongoDB**, το σύστημα διαχείρισης βάσεων δεδομένων
- **PHP, Perl, ή Python**, γλώσσες προγραμματισμού που χρησιμοποιούνται για την ανάπτυξη δυναμικών ιστοσελίδων και web development.

Ο ακριβής συνδυασμός λογισμικού που περιλαμβάνεται σε μια στοίβα LAMP μπορεί να παρουσιάζει μεγάλες διακυμάνσεις. Για παράδειγμα, ο Apache web server μπορεί να αντικατασταθεί από κάποιο άλλο λογισμικό web server, ενώ αντίστοιχα η PHP από κάποια άλλη γλώσσα προγραμματισμού, που χρησιμοποιείται για την ανάπτυξη δυναμικών ιστοσελίδων.

Παρόλο που οι σχεδιαστές των επιμέρους εφαρμογών κατά την δημιουργία τους δεν είχαν σαν στόχο την ενσωμάτωση των εφαρμογών τους σε ένα πακέτο λογισμικού, η φιλοσοφία ανάπτυξης αυτών, καθώς και τα εργαλεία που αναπτυχθήκαν σε στενή συνεργασία, βοήθησαν στο να λειτουργούν απρόσκοπτα μεταξύ τους.

Ο συνδυασμός του λογισμικού έχει γίνει δημοφιλής, χάρη στα επιμέρους συστατικά του, τα οποία αποτελούν δωρεάν λογισμικό ανοιχτού κώδικα, κάτι που σημαίνει ότι κάθε συστατικό μπορεί να προσαρμοστεί στο άλλο και να ικανοποιήσει στο μέγιστο δυνατό βαθμό τις προδιαγραφές του, χωρίς την παραμικρή παρέμβαση του κατασκευαστή.

Λόγω της φύσης του λογισμικού ελεύθερου και ανοιχτού κώδικα και της δυναμικής των συστατικών του, κάθε στοιχείο της στοίβας LAMP είναι πολύ καλά δοκιμασμένο, όσον αφορά τις επιδόσεις και την ασφάλεια, ενώ ταυτόχρονα ενεργή κοινότητα προγραμματιστών παρέχει την δυνατότητα για διάφορες προσαρμογές και πολύπλοκες ρυθμίσεις.

Τα περισσότερα, αν όχι όλα τα συστατικά της στοίβας LAMP υπάρχουν σε διαφορεές διανομές Linux δίνοντας την δυνατότητα σε οποιοδήποτε τελικό χρήστη να τα εγκαταστήσει, να τα ρυθμίσει και να λειτουργήσει μια στοίβα LAMP. Τέλος, αποτελεί πακέτο λογισμικού που ενέχει χαμηλό, αν όχι μηδενικό κόστος, ενώ παράλληλα παρουσιάζει σταθερή ανάπτυξη και εξέλιξη.

Το πακέτο LAMP μπορεί και συχνά συνδυάζεται με πολλά άλλα πακέτα λογισμικού ελεύθερου και ανοιχτού κώδικα, όπως για παράδειγμα:

- Netsniff-ng, για τον έλεγχο της ασφάλειας
- Snort, για την ανίχνευση εισβολής (IDS) και σύστημα αποτροπής εισβολών (IPS)
- Rrd tool, για τη δημιουργία διαγραμμάτων
- Nagios, Collectd, για την παρακολούθηση.

Ακολουθεί εκτενής ανάλυση των επιμέρους συστατικών μιας στοίβας LAMP.

## 5.2 Linux



Εικόνα 6: Λογότυπο Linux

### 5.2.1 Ορισμός

Το Linux αποτελεί ένα ανοικτού κώδικα project που παρέχεται δωρεάν σε οποιονδήποτε το επιθυμεί, δίνοντας τη δυνατότητα ελεύθερης αντιγραφής, τροποποίησης και αναδιανομής, χωρίς κανένα περιορισμό, ενώ είναι διαθέσιμο υπό τη άδεια GNU General Public License. Ειδοποιός διαφορά, λοιπόν, μεταξύ του Linux και άλλων δημοφιλών λειτουργικών συστημάτων είναι ότι ο πυρήνας του, αλλά και πολλές σημαντικές εφαρμογές του είναι ελεύθερο, ανοιχτού κώδικα λογισμικό.

Αποτελεί ένα από τα πιο χαρακτηριστικά project που οφείλουν τη δημιουργία τους στη συνεργασία μεταξύ διάφορων διαδικτυακών κοινοτήτων. Πρόκειται για ένα γενικό όρο που αναφέρεται στο σύνολο των λειτουργικών συστημάτων που βασίζονται στον πυρήνα Linux. Αξίζει να σημειωθεί, ότι ενώ η αρχιτεκτονική του Linux βασίζεται στις αρχές του λειτουργικού Unix, έχει αναπτυχθεί από μηδενική βάση και δεν περιέχει κώδικα από το Unix. Αναπτύσσεται με βάση το πρότυπο POSIX, που αποτελεί μια προσπάθεια τυποποίησης όλων των συστημάτων που βασίζονται σε Unix.

Βασικό χαρακτηριστικό του Linux είναι η δυνατότητα εγκατάστασης και λειτουργίας σε πλήθος υπολογιστικών συστημάτων, καλύπτοντας μια ευρεία γκάμα από φορητές υπολογιστικές συσκευές (smart phones) έως μεγάλα υπολογιστικά συστήματα και υπερυπολογιστές. Η χρήση του σε προσωπικούς υπολογιστές δεν είναι ιδιαίτερα διαδεδομένη, αν και τα τελευταία χρόνια παρατηρείται άνοδος, χάρη στη δημοφιλή διανομή Ubuntu, κατέχει όμως υψηλά ποσοστά χρήσης σε διακομιστές.

### 5.2.2 Ιστορική αναδρομή

Εμπνευστής και κύριος δημιουργός του Linux, στον οποίο οφείλει και την ονομασία του είναι ο Linus Torvalds, ο οποίος όντας ακόμη φοιτητής στο Πανεπιστήμιο του Ελσίνκι ασχολήθηκε για πρώτη φορά το 1991 με την ανάπτυξη ενός μη εμπορικού λειτουργικού, το οποίο θα βασιζόταν σε Unix. Στηριζόμενος κυρίως στο MINIX, λειτουργικό, το οποίο στηριζόταν επίσης σε Unix, ξεκίνησε τη δημιουργία του σημερινού πυρήνα Linux. Αρχικά, και μέχρι το Linux να φτάσει ένα ικανό στάδιο ανάπτυξης, ο Torvalds έγραφε προγράμματα τα οποία έτρεχαν και στο MINIX. Στη συνέχεια, η απόφασή του να αλλάξει την άδεια χρήσης, η οποία μέχρι τότε δεν επέτρεπε την ελεύθερη αναδιανομή του Linux για εμπορικούς σκοπούς, σε GNU GPL, έδωσε τεράστια ώθηση στο Linux, το οποίο σε σύντομο χρονικό διάστημα ξεπέρασε το MINIX.

Κομβικό σημείο στην ανάπτυξη του Linux αποτέλεσε η προσχώρηση του στο GNU, δίνοντας τη δυνατότητα σε χιλιάδες προγραμματιστές παγκοσμίως να συμβάλλουν στην ανάπτυξη και εξέλιξη του. Μέχρι σήμερα οι διανομές του Linux στηρίζονται σε ενεργές κοινότητες προγραμματιστών, εταιρίες που πωλούν εμπορικές εκδόσεις, αλλά και εταιρίες που παρέχουν τεχνική υποστήριξη για δωρεάν εκδόσεις.

Τέλος, σημαντική τα τελευταία χρόνια είναι και η συνεισφορά στη διάδοση του Linux εταιριών, όπως η IBM, η Google, η Intel, έχοντας ως στόχο την αύξηση των πωλήσεων τους σε hardware, επακόλουθο της κατοχής από το Linux σημαντικού μεριδίου αγοράς σε διακομιστές, smart phones και netbooks.

### 5.2.3 Χαρακτηριστικά

Το Linux αποτελείται από ένα αρθρωτό πυρήνα, ο οποίος είναι υπεύθυνος για την διαδικασία επεξεργασίας, τη δικτύωση, την πρόσβαση σε αρχεία συστήματος αλλά και τις περιφερειακές συσκευές, οι οδηγοί των οποίων αποτελούν μέρος του πυρήνα και εντάσσονται άμεσα σε αυτόν. Οι λειτουργίες υψηλού επιπέδου, όπως το γραφικό περιβάλλον των χρηστών, διασυνδέονται με τον πυρήνα χωρίς να αποτελούν τμήμα του και συνθέτουν έτσι ένα ολοκληρωμένο λειτουργικό σύστημα.

Από την πλευρά των χρηστών ένα σύστημα βασισμένο σε Linux μπορεί να ελεγχθεί με τους εξής τρόπους:

- μέσω μιας γραμμής εντολών
- μέσω του γραφικού περιβάλλοντος χρήστη (GUI) ή
- μέσω της χρήσης πλήκτρων ειδικής λειτουργίας που υπάρχουν στο σχετικό υλικό.

Όσον αφορά τους προσωπικούς υπολογιστές, ως προεπιλεγμένος τρόπος είναι το γραφικό περιβάλλον, με τη γραμμή εντολών να είναι διαθέσιμη μέσω τερματικού ή εικονικής κονσόλας. Από τα πιο διαδεδομένα γραφικά περιβάλλοντα χρήστη στο Linux είναι το X Window System. Το X καλύπτει μόνο λειτουργίες χαμηλού επιπέδου, παρέχει βασικά εργαλεία για τη δημιουργία, εμφάνιση, τοποθέτηση και διαχείριση γραφικών παραθύρων, ενώ υποστηρίζει και δικτυακή λειτουργία.

Οι ελλείψεις που παρουσιάζουν τα γραφικά περιβάλλοντα χρηστών καλύπτονται από τα λεγόμενα περιβάλλοντα εργασίας (Desktop Environments), όπως το GNOME, το KDE και το Xfce, τα οποία είναι και τα πιο διαδεδομένα. Τα περιβάλλοντα εργασίας παρέχουν τη δυνατότητα να υλοποιούν γραμμές εργασιών (taskbars), πίνακες εφαρμογών, screensavers.

### 5.2.4 Διανομές Linux

Μια διανομή Linux ορίζεται ως μια συλλογή από λογισμικό συστήματος και πακέτα εφαρμογών, όπως βιβλιοθήκες, εργαλεία συστήματος, η οποία διατίθεται για λήψη και εγκατάσταση μέσω μιας σύνδεσης δικτύου, κάτι που παρέχει στον τελικό χρήστη τη δυνατότητα προσαρμογής του λειτουργικού συστήματος στις ανάγκες του. Ουσιαστικά, μέσω μια διανομής Linux, εγκαθίσταται και

ρυθμίζεται ο πυρήνας, ενώ παράλληλα διασφαλίζεται η γενική ασφάλεια του συστήματος και ορίζεται το επίπεδο ενσωμάτωσης των διάφορων πακέτων λογισμικού. Για την εγκατάσταση, ενημέρωση και αφαίρεση του λογισμικού ενός συστήματος μέσω μιας ενιαίας διαδικασίας, οι διανομές Linux χρησιμοποιούν κάποιον χειριστή πακέτων, όπως ο Synaptic.

Κάθε διανομή Linux μπορεί να διακριθεί είτε σε γενικού σκοπού, είτε σε πιο εξειδικευμένη, ανάλογα με το κίνητρο δημιουργίας της (φιλική προς το χρήστη, εύκολα παραμετροποιήσιμη κ.α.). Για παράδειγμα, κάποιες διανομές έχουν ως στόχο να εξυπηρετήσουν ένα συγκεκριμένο τύπο χρηστών, να δίνουν έμφαση σε ένα διαφορετικό τομέα του συστήματος, όπως ευχρηστία, ασφάλεια, πολυγλωσσία ή ενδεχομένως να λειτουργούν μόνο σε συγκεκριμένο υλικό υπολογιστών.

Οι περισσότερες, αν όχι όλες, διανομές Linux συμπεριλαμβάνουν τον web browser Mozilla Firefox και τη σουίτα εφαρμογών γραφείου LibreOffice. Από τις πιο διαδεδομένες διανομές Linux για web server είναι το Debian και το Ubuntu, όπου μοιράζονται το 58.5% του μεριδίου αγοράς έναντι των RHEL, Fedora και CentOS, όπου μαζί μοιράζονται το 37.3 %, με βάση στοιχεία που ανακοινώθηκαν τον Οκτώβριο του 2013.

### 5.3 Apache Web Server



Εικόνα 7: Λογότυπο Apache

#### 5.3.1 Εισαγωγικά στοιχεία

Για να μπορέσει κανείς να ανεβάσει μια ιστοσελίδα στο διαδίκτυο θα πρέπει να προχωρήσει στην ενοικίαση κάποιου χώρου σε έναν web server, ο οποίος θα φιλοξενήσει την ιστοσελίδα. Με τον όρο web server software αναφερόμαστε στο λογισμικό που τρέχει σε ένα διαδικτυακό κόμβο και επιτρέπει σε άλλους υπολογιστές να αποκτούν αντίγραφα των ιστοσελίδων, που είναι αποθηκευμένες σε αυτόν. Για να καταστεί εφικτή η επικοινωνία του web server με τους άλλους υπολογιστές, παρέχεται σε κάθε server μια μοναδική διεύθυνση IP.

Ο Apache web server ξεκινώντας με την ονομασία NCSA HTTPd το 1993 από τον Robert McCool, έφτασε το 1996 να γίνει η πρώτη εναλλακτική λύση απέναντι στον Netscape Corporation server και σταδιακά να ανταγωνίζεται σε λειτουργικότητα αλλά και απόδοση όλους τους web server βασισμένους σε Unix. Σήμερα αποτελεί τον πιο αναγνωρίσιμο web server, συγκεντρώνοντας πάνω από το 60% του συνόλου των ιστοσελίδων.

Ο Apache, μέρος του Apache software foundation, αποτελεί ένα project που χρησιμοποιεί δωρεάν λογισμικό ανοιχτού κώδικα. Ένα από τα κύρια χαρακτηριστικά που βοήθησαν στην μεγάλη αναγνωρισιμότητα του είναι η μεγάλη γκάμα λειτουργικών συστημάτων για τα οποία είναι διαθέσιμος και στα οποία εγκαταλούνται τα Unix, Linux, Solaris, Novell NetWare, GNU, FreeBSD, MacOS X, Windows, OS/2, TPF και ecomStation.

### 5.3.2 Χαρακτηριστικά

Ένα από τα σημαντικότερα πλεονεκτήματα του Apache web server είναι οι δυνατότητες που μπορεί να προσφέρει, χάρη στην επεκτασιμότητα του, η οποία επιτυγχάνεται με την ενσωμάτωση modules. Προσφέρονται modules, με τα οποία:

- Υποστηρίζονται server side γλώσσες προγραμματισμού, όπως είναι η Perl, Python, Tcl και PHP.
- Παρέχονται δυνατότητες πιστοποίησης και αυθεντικοποίησης, όπως είναι τα mod\_access, mod\_auth, mod\_digest και mod\_auth\_digest, διάδοχος του mod\_digest.
- Υποστηρίζονται πρωτόκολλα Secure Sockets Layer και Transport Layer Security (mod\_ssl), μονάδα μεσολάβησης (mod\_proxy), επανεγγραφής URL (mod\_rewrite), προσαρμοσμένα αρχεία καταγραφής (mod\_log\_config), καθώς και φίλτρα (mod\_include και mod\_ext\_filter).

Επιπρόσθετα, από την μεγάλη ποικιλία των προσφερόμενων modules δεν θα μπορούσαν να παραλειφθούν modules ασφαλείας, όπως είναι το ModSecurity, το οποίο είναι ανοιχτού κώδικα και έχει την δυνατότητα να εντοπίζει και να εμποδίζει εισβολές σε διαδικτυακές εφαρμογές, αλλά και modules συμπίεσης, όπως το mod\_gzip, το οποίο εφαρμόζεται για να βοηθήσει την μείωση του μεγέθους των ιστοσελίδων που εξυπηρετούνται μέσω HTTP.

Τέλος, κάποια επιπλέον σημαντικά χαρακτηριστικά του Apache είναι η δυνατότητά του να υποστηρίζει virtual hosting, δηλαδή με μια εγκατάσταση του Apache server, μπορούν να εξυπηρετηθούν πολλαπλές σελίδες. Επιπλέον, παρέχει τη δυνατότητα υποστήριξης παραμετροποιήσιμων μηνυμάτων σφάλματος, βάσεων δεδομένων βασισμένες σε αυθεντικοποίηση DBMS, καθώς και πολλά γραφικά περιβάλλοντα χρήστη (GUIs).

## 5.4 MySQL



Εικόνα 8: Λογότυπο MySQL

### 5.4.1 Ορισμός

Πρόκειται για ένα open-source σύστημα διαχείρισης σχεσιακών βάσεων δεδομένων. Βασικές δυνατότητες που παρέχει η MySQL είναι:

- Δυνατότητα αναζήτησης, αποθήκευσης, ταξινόμησης, ομαδοποίησης και ανάκλησης δεδομένων. Για να επιτευχθεί αυτό η MySQL στηρίζεται στη γλώσσα ερωτημάτων SQL.

- Σχεσιακή: Δυνατότητα οργάνωσης των δεδομένων σε διαφορετικούς πίνακες, η συσχέτιση των οποίων γίνεται με κάποιο σαφώς ορισμένο τρόπο.
- Δυνατότητα ελέγχου της πρόσβασης στα δεδομένα, έτσι ώστε να εξασφαλίζεται η πρόσβαση από διαφορετικούς χρήστες.
- Τέλος, δυνατότητα παροχής συγκεκριμένων δικαιωμάτων πάνω στις βάσεις δεδομένων σε κάθε χρήστη.

Σήμερα, η MySQL ανήκει και χρηματοδοτείται από την Oracle Corporation (Sun Microsystems, MySQL AB), ενώ έχει στο ενεργητικό της πάνω από 11 εκατομμύρια εγκαταστάσεις. Ενδεικτικά παραδείγματα χρήσης της από διαδεδομένες διαδικτυακές υπηρεσίες είναι το Flickr, το YouTube, η Wikipedia, το Facebook, το Twitter.

#### 5.4.2 Πλεονεκτήματα

Βασικά πλεονεκτήματα της MySQL που την καθιστούν άκρως ανταγωνιστική είναι:

- Απόδοση: Είναι αρκετά γρήγορη, κάτι που έχει επιβεβαιωθεί και από δοκιμές με ανταγωνιστικά προϊόντα, οι οποίες κατέληξαν στο συμπέρασμα της υπεροχής της.
- Κόστος: Πρόκειται για ένα προϊόν ανοιχτού κώδικα, το οποίο διατίθεται δωρεάν για προσωπική χρήση. Σε περίπτωση επιλογής της για εμπορικές εφαρμογές, που δεν θα είναι ανοιχτού κώδικα, η εμπορική της άδεια διατίθεται με χαμηλό κόστος.
- Εκμάθηση: Είναι αρκετά εύκολο να εξοικειωθεί κανείς με τη χρήση της MySQL.
- Μεταφερσιμότητα: Μπορεί να χρησιμοποιηθεί σε πολλά σύγχρονα λειτουργικά συστήματα. Για παράδειγμα, είναι συμβατή με πολλές εκδόσεις των Windows, αλλά και με Unix λειτουργικά (Linux).
- Πηγαίος Κώδικας: Ως open-source λογισμικό παρέχει τη δυνατότητα σε κάθε χρήστη να αποκτήσει, να τροποποιήσει και να προσαρμόσει τον πηγαίο κώδικα στις ανάγκες του. Επιπλέον, δεν είναι απαραίτητη η αναμονή για μια νέα επίσημη έκδοση, για να καλυφθεί κάποια έλλειψη, αφού κάτι τέτοιο γίνεται πολύ εύκολα από τους χρήστες της, οι οποίοι έχουν πρόσβαση στον κώδικα.
- Τέλος, σημαντικό πλεονέκτημα είναι η σημαντική υποστήριξη που παρέχεται μέσω εγχειριδίων, tutorials και βοήθειας στην επίλυση προβλημάτων από το επίσημο site της MySQL ([www.mysql.com](http://www.mysql.com)).

## 5.5 PHP



Εικόνα 9: Λογότυπο PHP

### 5.5.1 Ορισμός

Η PHP αποτελεί μια γενικού σκοπού γλώσσα συμβάντων, ειδικά σχεδιασμένη για τη δημιουργία ιστοσελίδων δυναμικού περιεχομένου. Βασικό χαρακτηριστικό της είναι ότι πρόκειται για μια server-side γλώσσα, δηλαδή λειτουργεί στην πλευρά του διακομιστή. Μέσω ενός συμβατού διακομιστή, ο οποίος επεξεργάζεται την σελίδα PHP, παράγεται σε πραγματικό χρόνο το τελικό περιεχόμενο, το οποίο και αποστέλλεται στον web browser που χρησιμοποιεί ο επισκέπτης του site σε μορφή κώδικα HTML.

Βασική χρήση της PHP είναι η δημιουργία scripts για ιστοσελίδες. Επιπλέον, με τη βοήθεια ενός κατάλληλου μεταγλωττιστή μπορεί να χρησιμοποιηθεί και για command line-scripting αλλά και για εγγραφή client-side GUI εφαρμογών.

### 5.5.2 Ιστορική εξέλιξη

Η αρχή στην ανάπτυξη της γλώσσας PHP έγινε από τον Rasmus Lerdorf, το 1994, οποίος δημιούργησε ένα απλό script, με όνομα php.cgi, χρησιμοποιώντας τη γλώσσα Perl, με στόχο να διατηρεί μια λίστα στατιστικών στοιχείων για τα άτομα που έβλεπαν το online βιογραφικό του σημείωμα. Το script αυτό άρχισε να διαδίδεται και το 1997 έφτασε στην έκδοση 2.0, με την ονομασία PHP/FI (Personal Home Page/From Interpreter), βασιζόμενο πλέον στην γλώσσα προγραμματισμού C.

Την ίδια χρόνια, η γλώσσα ξαναγράφηκε από την αρχή από τους Andi Gutmans και Zeev Suraski, στηριζόμενοι πάντα στην PHP/FI, φθάνοντας στην έκδοση 3.0. Αργότερα, προχώρησαν στη σύσταση της εταιρία Zend, η οποία ασχολείται μέχρι και σήμερα με την ανάπτυξη και εξέλιξη της PHP. Πλέον χρησιμοποιούνται ευρέως οι εκδόσεις 4 και 5, ενώ σύντομα αναμένεται και η επίσημη διάθεση της έκδοσης 6.

### 5.5.3 Χαρακτηριστικά και τρόπος λειτουργίας

Τα βασικά χαρακτηριστικά της PHP είναι:

Αποτελεί γλώσσα script και όχι γλώσσα προγραμματισμού.

Η PHP είναι σχεδιασμένη να εκτελεί μια ενέργεια μετά από κάποιο συμβάν. Για παράδειγμα, στην περίπτωση που ο χρήστης «πατήσει» κάποιο link στην ιστοσελίδα.

Αποτελεί server-side και όχι client-side γλώσσα.

Η PHP λειτουργεί στην πλευρά του διακομιστή (server). Γίνεται εγκατάσταση της PHP στον server, στη συνέχεια τα script που είναι γραμμένα σε αυτή χρησιμοποιούν πόρους του server για την εκτέλεσή τους. Τέλος, τα αποτελέσματα της εκτέλεσης αποστέλλονται στον πελάτη (client) σε μορφή HTML.

Επιπλέον, λόγω του ότι αποτελεί μια server-side γλώσσα, δεν μπορεί να χρησιμοποιηθεί για αναδύμενα μενού, άνοιγμα νέων παραθύρων, προσθήκη εφέ, ενεργειών με τη μετακίνηση του ποντικιού. Τα παραπάνω αποτελούν δυνατότητες, οι οποίες παρέχονται από τις client-side γλώσσες, όπως η Javascript.

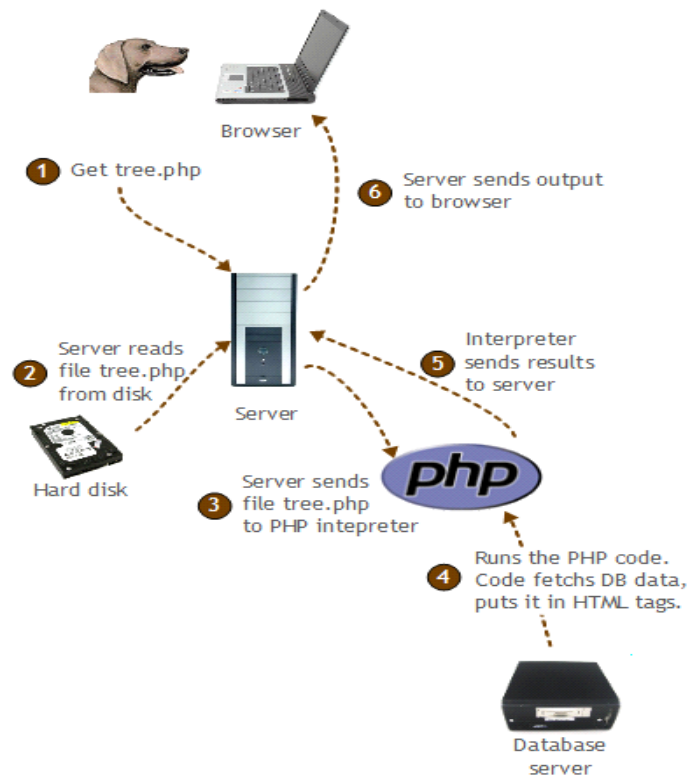
Ο κώδικας της PHP παρεμβάλλεται σε κώδικα HTML.

Αυτό γίνεται με τη κατάλληλη σήμανση στην αρχή και στο τέλος του κώδικα PHP, με τα κατάλληλα tags έναρξης – τερματισμού, τα οποία εξαρτώνται από τις ρυθμίσεις που έχουν γίνει κατά την εγκατάσταση. Επιπλέον, τα αρχεία που περιέχουν κώδικα PHP πρέπει να αποθηκευτούν με κατάλληλη κατάληξη, όπως \*.php. Ουσιαστικά λοιπόν, ο κώδικας της PHP εκτελείται ταυτόχρονα με τον κώδικα της HTML και όχι αυτόνομα.

Εύκολη συγγραφή του κώδικα.

Η διαδικασία συγγραφής του κώδικα είναι σχετικά απλή, ενώ γράφεται ταυτόχρονα με τον HTML κώδικα και ενσωματώνεται σε αυτόν. Για τη συγγραφή χρησιμοποιείται ένας απλός text-editor, παρέχοντας στον προγραμματιστή το πλεονέκτημα της μέγιστης δυνατής εποπτείας. Από την άλλη, βασικό πρόβλημα στη διαδικασία ενδεχομένως να αποτελέσει το ότι απαιτεί πολύ χρόνο, κάτι το οποίο είναι δυνατό να αντιμετωπιστεί με τη χρήση διάφορων εργαλείων, ανοιχτού κώδικα, τα οποία προσφέρουν έτοιμες standard δομές ή μπορούν να προσθέσουν χαρακτηριστικά στην σελίδα με ένα μόνο «κλικ».

Διαγραμματικά, ο τρόπος με τον οποίο ενσωματώνεται η χρήση της PHP σε μια ιστοσελίδα αναλύεται ως εξής:



Εικόνα 10: Τρόπος λειτουργίας PHP



- Ο επισκέπτης ζητάει μια ιστοσελίδα με τη διεύθυνση της, μέσω ενός web browser.
- Ο browser μεταβιβάζει το αίτημα στον server.
- Η PHP επεξεργάζεται το αρχείο που αφορά την αίτηση.
- Η PHP εκτελεί το μέρος του κώδικα που την αφορά.
- Η PHP επιστρέφει το αποτέλεσμα σε μορφή HTML.
- Ο κώδικας HTML επιστρέφεται στον browser του επισκέπτη.

#### 5.5.4 Πλεονεκτήματα

Στη συνέχεια, παρουσιάζονται τα βασικά πλεονεκτήματα που προσφέρει η χρήση της γλώσσας PHP:

- Μηδενικό κόστος: Η PHP διατίθεται δωρεάν, δεδομένου του ότι πρόκειται για μια open source γλώσσα.
- Διαθεσιμότητα Πηγαίου Κώδικα: Ο πηγαίος κώδικας της PHP είναι διαθέσιμος σε όλους για χρήση, τροποποίηση και αναδιανομή σε αντίθεση με τα εμπορικά πακέτα κλειστού κώδικα.
- Χρηματοδότηση και Υποστήριξη: Η εταιρία πίσω από την καθιέρωση της PHP, Zend, χρηματοδοτεί την ανάπτυξη της, προσφέροντας υποστήριξη και σχετικό λογισμικό.
- Διαθέσιμη σε πολλά λειτουργικά συστήματα, που σημαίνει ότι ο κώδικας δουλεύει χωρίς αλλαγές. Ενδεικτικά παραδείγματα, λειτουργικών συστημάτων στα οποία λειτουργεί η PHP είναι τα Windows, Linux, Solaris.
- Υψηλή αποδοτικότητα: Η PHP μπορεί να εξυπηρετήσει εκατομμύρια επισκέψεων σε ημερήσια βάση, με τη χρήση ενός φθηνού διακομιστή.
- Δυνατότητα διασύνδεσης με πολλά διαφορετικά συστήματα βάσεων δεδομένων: Προσφέρει δικές της συνδέσεις με συστήματα βάσεων δεδομένων, όπως MySQL, PostgreSQL, Oracle, dbm, Hyperwave. Παράλληλα, έχει ενσωματωμένη SQL διασύνδεση στο επίπεδο αρχείο SQLite, ενώ μπορεί να συνδεθεί σε οποιαδήποτε βάση έχει πρόγραμμα οδήγησης ODBC, όπως τα προϊόντα της Microsoft.
- Ενσωματωμένες βιβλιοθήκες, μέσω των οποίων προσφέρει λειτουργίες δυναμικής δημιουργίας εικόνων gif, σύνδεσης με άλλες υπηρεσίες δικτύων, αποστολής email, ανάλυσης XML, δημιουργίας εγγράφων pdf.
- Αντικειμενοστραφής υποστήριξη: Η PHP προσφέρει αντικειμενοστραφής λειτουργίες, όπως η κληρονομικότητα, οι διασυνδέσεις, οι αφηρημένες κλάσεις και μέθοδοι, οι προστατευμένες ιδιότητες, οι συναρτήσεις δημιουργίας.
- Εύκολη εκμάθηση: Πρόκειται για μια ευέλικτη και εύκολη στην εκμάθηση γλώσσα, χωρίς να απαιτεί ιδιαίτερες γνώσεις προγραμματισμού. Επιπλέον, το γεγονός ότι η σύνταξη της στηρίζεται σε άλλες γλώσσες προγραμματισμού και κυρίως τη C, επιτρέπουν την ακόμη ευκολότερη χρήση της από κάποιον γνώστη μιας γλώσσας προγραμματισμού της οικογένειας C.



## Κεφάλαιο Έκτο: Άλλες Τεχνολογίες

### 6.1 XML

#### 6.1.1 Εισαγωγικά στοιχεία

Η HTML (Hypertext Markup Language) ή στα ελληνικά Γλώσσα Σήμανσης Υπερκειμένου, αποτελεί την κύρια γλώσσα κατασκευής ιστοσελίδων. Σκοπός ενός web browser είναι να διαβάσει τα HTML documents και να συνθέσει σελίδες που να μπορεί κανείς να διαβάσει ή να ακούσει. Ο όρος markup αναφέρεται στη χρήση tags, δηλαδή αναγνωριστικών, τα οποία μαρκάρουν το κείμενο, χωρίζοντας το σε λογικά μέρη, όπως παραγράφους, πίνακες. Η HTML, ουσιαστικά γράφεται σε μορφή HTML στοιχείων, τα οποία αποτελούνται από ετικέτες έναρξης – λήξης (tags).

Παρόλο που δεν αποτελεί γλώσσα προγραμματισμού στηρίζεται σε συγκεκριμένους κανόνες σύνταξης. Ακολουθώντας αυτούς τους κανόνες κατασκευάζεται ένα HTML document. Κατά τον σχεδιασμό της επίσημης σύνταξης της HTMLDTD (HTML Document Type Definition) χρησιμοποιήθηκε η μετα-γλώσσα SGML. Με αυτόν τον τρόπο παρέχεται ένα σύστημα τυποποίησης κειμένων, το οποίο χρησιμοποιείται από τους webpage designers, στην περιγραφή της δομής και της οργάνωσης των σελίδων τους.

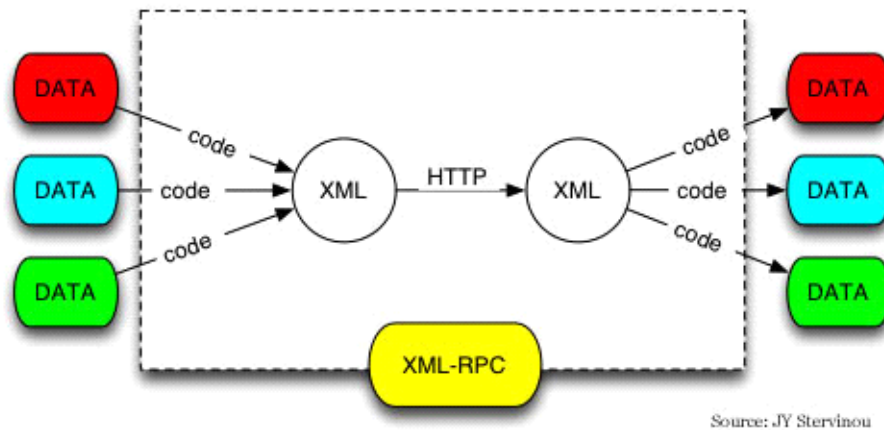
Η HTML καθιστά δυνατή την ενσωμάτωση εικόνων και άλλων αντικειμένων μέσα σε μια σελίδα καθώς και την εμφάνιση διαδραστικών φορμών. Δημιουργεί δομημένα έγγραφα, καθορίζοντας σημαντικά δομικά στοιχεία, όπως κεφαλίδες, παραγράφους, συνδέσμους. Τέλος, παρέχει την δυνατότητα ενσωμάτωσης σεναρίων σε άλλες γλώσσες, όπως η Javascript, τα οποία δύνανται να επηρεάζουν τη συμπεριφορά των HTML ιστοσελίδων.

#### 6.1.2 Ορισμός XML

Σε αντίθεση με την HTML, η XML (Extensible Markup Language) δεν είναι σημειακή γλώσσα αλλά μετά-γλώσσα, δηλαδή γλώσσα προγραμματισμού, η οποία χρησιμοποιείται για την περιγραφή μιας σημειακής γλώσσας, όπως η HTML. Η XML παρέχοντας λύση σε πολλά προβλήματα με τα οποία αναγκαζόταν να έρθει αντιμέτωπος ένας web designer, οδήγησε στην δημιουργία της ανασχεδιασμένης HTML, της XHTML, η οποία θα χρησιμοποιείται σε βάθος χρόνου, αφού προσφέρει αποτελεσματικές και δυναμικές πολυμεσικές λύσεις.

Βασικός στόχος κατά τη δημιουργία της XML ήταν η ικανοποίηση διάφορων αναγκών, μέσω της παροχής δυνατότητας στα έγγραφα να αποκτήσουν μεγαλύτερη ευελιξία και προσαρμοστικότητα στο στυλ και τη δομή από αυτό που παρείχε μέχρι τότε η HTML. Η XML, όπως έχει ήδη αναφερθεί, δεν είναι μια απλή markup γλώσσα για έγγραφα που περιέχουν δομημένες πληροφορίες, αλλά μια μετά-γλώσσα (metalinguage), δηλαδή μια γλώσσα που χρησιμοποιείται για να καθορίσει νέες markup γλώσσες. Αξίζει να σημειωθεί ότι στόχος της XML δεν ήταν να αντικαταστήσει την HTML, αλλά να τη συμπληρώσει. Ειδοποιός διαφορά των δύο είναι η εξής:

- Η HTML χρησιμοποιείται στη διατύπωση και εμφάνιση των δεδομένων, με προκαθορισμένα tags, ενώ
- Η XML αναπαριστά τη συναφή έννοια των δεδομένων και παρέχει στους χρήστες τη δυνατότητα καθορισμού των tags και των δομημένων μεταξύ τους σχέσεων.



Εικόνα 11: Τρόπος Λειτουργίας XML

Τα XML έγγραφα δεν είναι πολύπλοκα, αλλά απλά και αποτελεσματικά. Αποτελούνται από μονάδες αποθήκευσης, που παρέχουν αναλυμένες ή μη πληροφορίες και καλούνται entities (οντότητες). Οι αναλυμένες πληροφορίες αποτελούνται από characters (χαρακτήρες), οι οποίοι συνθέτουν character data και άλλους οι οποίοι συνθέτουν markup. Με τη μορφή markup γίνεται κωδικοποίηση της περιγραφής της τελικής αποθήκευσης και της λογικής δομής του εγγράφου.

Στη συνέχεια για να διαβαστεί ένα XML έγγραφο και να αποκτήσουμε πρόσβαση στη δομή και το περιεχόμενό του, χρησιμοποιείται ένα λογισμικό μοντέλο, ο επεξεργαστής XML. Ο επεξεργαστής λειτουργεί εκ μέρους μιας εφαρμογής (application), που περιγράφει την απαιτούμενη συμπεριφορά του επεξεργαστή, δηλαδή πως θα πρέπει να διαβάζονται τα XML έγγραφα και ποιες πληροφορίες θα πρέπει να παρέχονται στην εφαρμογή.

### 6.1.3 Ιστορική εξέλιξη

Η αφορμή για την ανάπτυξη της XML δόθηκε το 1990, με την αλματώδη ανάπτυξη του διαδικτύου και τον έντονο ανταγωνισμό μεταξύ των browser της Microsoft και της Netscape. Η «επίσημη», μέχρι τότε, γλώσσα κατασκευής των websites, HTML, εμφάνισε προβλήματα όπως:

- Δημιουργήθηκε η ανάγκη οι web designers να πρέπει να αναπτύξουν διαφορετική έκδοση για κάθε browser, λόγω του ότι τα ίδια κομμάτια HTML σχεδιάζονταν με διαφορετικό τρόπο.
- Αναπτύχθηκαν εργαλεία HTML από ορισμένους browsers, τα οποία δεν ήταν αναγνωρίσιμα από όλους.
- Δεδομένου του ότι η HTML παρείχε πληροφορίες μόνο για την παρουσίαση του website και όχι για το τι αφορούσε, ήταν αδύνατο να αντληθεί οποιαδήποτε τυποποιημένη πληροφορία μέσω αυτού. Το πρόβλημα ενέτεινε η εμφάνιση των μηχανών αναζήτησης.

Με στόχο την επίλυση αυτών των προβλημάτων, η Κοινοπραξία Παγκόσμιου Ιστού (W3C), το 1996, προχώρησε στην ανάπτυξη μιας γλώσσας, που θα υποσκελίζε την HTML και θα παρείχε τις εξής δυνατότητες:

- Να είναι εύχρηστη στο διαδίκτυο.
- Να μπορεί να υποστηρίξει ποικιλία εφαρμογών (για παράδειγμα βάσεις δεδομένων μηχανών αναζήτησης).
- Να είναι συμβατή με την SGML.

- Να μην υπάρχουν, αν είναι δυνατόν, προαιρετικά χαρακτηριστικά στην XML.
- Να παρέχει τη δυνατότητα εύκολης δημιουργίας προγραμμάτων που θα επεξεργάζονται XML έγγραφα.
- Τα XML έγγραφα θα πρέπει να είναι ευανάγνωστα και να δημιουργούνται εύκολα.
- Ο σχεδιασμός της θα πρέπει να είναι τυπικός και περιεκτικός.

Η XML αποτελεί σήμερα μια σταθερά για το διαδίκτυο, με την επίσημη παρουσίαση της να λαμβάνει χώρα το 1998 από τον W3C. Μέχρι σήμερα έχουν αναπτυχθεί εκατοντάδες γλώσσες που βασίζονται στην XML, όπως RRS, SOAP, ενώ και οι περισσότερες σουίτες εφαρμογών γραφείου, (Microsoft Office, OpenOffice.org, iWork της Apple) περιέχουν προεπιλεγμένες κωδικοποιήσεις βασισμένες στην XML.

### 6.1.4 Πλεονεκτήματα – Μειονεκτήματα XML

Τα πλεονεκτήματα της XML πηγάζουν από το ότι είναι απλό κείμενο και δεν ασχολείται με την επεξεργασία των δεδομένων, αλλά μόνο με τη σημασία τους. Επιπλέον μέσω βιβλιοθηκών που έχουν δημιουργηθεί ή μέσω της ίδιας της XML (XLS) μπορούμε εύκολα να επεξεργαστούμε αρχεία αυτής της μορφής.

Βασικό της πλεονέκτημα βέβαια αποτελεί ότι η XML παρέχει τη δυνατότητα σε οποιαδήποτε γλώσσα προγραμματισμού να δημιουργήσει δεδομένα σε αυτή τη μορφή και να τα διαβάσει.

Τέλος, η XML υπερτερεί έναντι παλαιότερων μορφών μετάδοσης δεδομένων, όπως με τη χρήση δομημένων κειμένων (CSV) ή τη χρήση ροής δεδομένων (CORBA), ακριβώς επειδή μπορεί να χρησιμοποιηθεί άμεσα και να διευκολύνει τη μετάδοση αντικειμενοστραφών πληροφοριών. Για παράδειγμα, το CSV αποδείχτηκε υποδεέστερο, εξαιτίας της ύπαρξης διαφορετικής κωδικοποίησης μεταξύ των δομημένων κειμένων, ενώ το CORBA, λόγω της έλλειψης ομοειδούς υποστήριξης από τις διάφορες πλατφόρμες.

Από την άλλη στα μειονεκτήματα της XML θα μπορούσαμε να αναφέρουμε τα μεγάλα αρχεία στα οποία οδηγεί η κωδικοποίηση σε XML, εξαιτίας του ότι όλα τα δεδομένα βρίσκονται μέσα σε tags. Ενδεικτικά αναφέρουμε, ότι είναι πιθανό πάνω από το 70% ενός αρχείου XML να καταλαμβάνεται από tags.

Παράλληλα, πολλές φορές, σε περιπτώσεις επιστημονικών εφαρμογών ή εφαρμογών που δίνουν μεγάλη σημασία στην απόδοση, ενδεχομένως να μην ενδείκνυται η χρήση της, λόγω της εξαιρετικά χρονοβόρας διαδικασίας που απαιτείται για την κωδικοποίηση των δεδομένων σε απλό κείμενο.

## 6.2 RSS



Εικόνα 12: Λογότυπο RSS

### 6.2.1 Ορισμός

Αν και η πλειοψηφία των ανεπτυγμένων τεχνολογιών του διαδικτύου, ανήκει στις αποκαλούμενες pull τεχνολογίες, όπου οι πληροφορίες μεταφέρονται ύστερα από αίτημα, υπάρχουν και οι τεχνολογίες push, όπου το υλικό μεταδίδεται κατά βούληση σε ένα συνδρομητή. Από τις πιο γνωστές τεχνολογίες push είναι τα RSSFeeds, τα οποία και θα αναλύσουμε στη συνέχεια.

Η τεχνολογία RSS (Rich Site Summary ή όπως παραφράζεται συχνά Really Simple Syndication) αποτελεί μια προτυποποιημένη μέθοδο ανταλλαγής ψηφιακού πληροφοριακού περιεχομένου μέσω του διαδικτύου, η οποία στηρίζεται στην γλώσσα σήμανσης XML. Ουσιαστικά, μέσω των ροών RSS (RSSFeeds), ένας χρήστης μπορεί να λαμβάνει ενημερώσεις και νέα από οποιαδήποτε ιστοσελίδα επιθυμεί και υποστηρίζει την τεχνολογία RSS, κάνοντας απλά εγγραφή στην υπηρεσία RSS της αντίστοιχης σελίδας. Οι ροές RSS είναι δυνατό να περιέχουν πλήρη δεδομένα, σύνοψη των δεδομένων, σχετικά μεταδεδομένα.

### 6.2.2 Ιστορική εξέλιξη

Η τεχνολογία RSS εμφανίστηκε το Μάρτιο του 1999, υπό την έκδοση RSS 0.9 από τη Netscape για χρήση στο portal “MyNetscape”. Τον Απρίλιο του 2001, ενώ είχε ήδη αρχίσει να χρησιμοποιείται εκτενώς παρέμεινε χωρίς «ιδιοκτήτη», λόγω της εξαγοράς της AOL και της επακόλουθης αναδιάρθρωσης της Netscape, η οποία αφαίρεσε την υποστήριξη RSS/XML από το “MyNetscape” σε επανασχεδιασμό του.

Τον Ιούλιο του 2003, ο Winer και η Userland Software απέκτησαν την κυριότητα της προδιαγραφής του RSS 2.0. Ορόσημο στην καθιέρωση του προτύπου αποτέλεσε η πρωτοβουλία της New York Times, τον Νοέμβριο του 2002, να διαθέσει στους αναγνώστες της τη δυνατότητα να εγγραφούν σε RSS newsfeeds σχετικά με διάφορα θέματα.

Έως το 2003, βασική μέθοδος διάδοσης περιεχομένου στο web ήταν το πρότυπο RSS. Εξαιτίας των ανεπαρκειών που εμφάνιζε και παράλληλα της αδυναμίας πραγματοποίησης αλλαγών απευθείας στο πρότυπο RSS 2.0, λόγω κατοχύρωσης της πνευματικής ιδιοκτησίας, δημιουργήθηκε μια νέα τάση.

Τον Ιούνιο του 2003, ο Sam Ruby δημιούργησε ένα wiki με στόχο την ανεύρεση λύσης στο πως θα μπορούσε να δομηθεί σωστά μια εγγραφή ιστορικού. Το νέο έργο βρήκε πληθώρα υποστηρικτών. Βασικοί στόχοι που ετέθησαν για το νέο πρότυπο ήταν:

- Να είναι εντελώς ανεξάρτητο από τον κατασκευαστή, ώστε να αποφευχθεί ενδεχόμενο πάγωμα στην επέκτασή του.
- Να είναι εύκολα υλοποιήσιμο.
- Να είναι ξεκάθαρα ορισμένο.

- Να είναι δυνατή η ελεύθερη επέκταση του από οποιονδήποτε.

Η πρώτη έκδοση του εν λόγω προτύπου έγινε διαθέσιμη τον Ιούλιο του 2003 με την ονομασία Atom 0.2. Η επόμενη έκδοση λίγους μήνες αργότερα (Atom 0.3) υιοθετήθηκε ευρέως από τα εργαλεία syndication και προστέθηκε σε πολλές υπηρεσίες της Google, όπως το Blogger, Google News και Gmail. Η αναγνώριση του Atom 1.0 από την IETF (Internet Engineering Task Force) πραγματοποιήθηκε τον Δεκέμβριο του 2005.

Παράλληλα, τον Ιανουάριο του 2005, οι Sean Palmer, Christopher Schmidt και Cody Woodard, προχώρησαν στη δημιουργία μιας προέκδοσης του RSS 1.1, επιδιορθώνοντας λάθη της προηγούμενης έκδοσης, αφαιρώντας άχρηστες δυνατότητες, απλοποιώντας τη σύνταξη και βελτιώνοντας την τεκμηρίωση.

Μέχρι τον Δεκέμβριο του 2005 η έκδοση RSS 1.1 είχε καταστεί ευρέως αποδεκτή. Ενδεικτικά αναφέρουμε:

- Απρίλιος 2005: Η Apple Computer διαθέτει την έκδοση 2.0 του Safari με ενσωματωμένες δυνατότητες υποστήριξης RSS Feed.
- Νοέμβριος 2005: Η Microsoft προτείνει τη δικιά της “Simple Sharing Extensions” για το RSS.
- Δεκέμβριος 2005: Η ομάδα του Internet Explorer και η ομάδα του Outlook, ανακοινώνουν ότι τα blogs τους θα υιοθετήσουν το γνωστό πορτοκαλί εικονίδιο με τα ραδιοκύματα, που χρησιμοποιούσε ο Mozilla Firefox browser, καθιστώντας το σύμβολο του RSS, αλλά και σχετικών προτύπων, όπως το Atom.
- Τέλος, το Φεβρουάριο του 2006, η Opera Software ανακοινώνει πως στην έκδοση 9 του δικού της browser θα υιοθετήσει το ίδιο εικονίδιο.

### 6.2.3 Πλεονεκτήματα

Τα βασικά πλεονεκτήματα που παρουσιάζει η χρήση των RSS Feeds είναι:

- Η απελευθέρωση της πληροφορίας και η ταχύτητα διασποράς.
- Η ευκολία στην ανεύρεση της πληροφορίας.
- Η θεματική ομαδοποίηση και εμφάνιση των δεδομένων.
- Η επιλογή της εμφανιζόμενης πληροφορίας.
- Η αποφυγή της άχρηστης και ενοχλητικής πληροφορίας (spam).
- Η δυνατότητα προσωπικής παραμετροποίησης της εμφάνισης ανάλογα με το περιεχόμενο κάθε site.
- Η αποφυγή χρονοβόρας «φόρτωσης» της ιστοσελίδας, με τον αποκλεισμό γραφικών, διαφημίσεων, flash.
- Η εξοικονόμηση χρόνου.
- Η αύξηση της επισκεψιμότητας (traffic) των ιστοσελίδων που παρέχουν RSSFeeds.
- Τέλος, η εύκολη χρήση και σε κινητά, PDA.

## 6.2.4 RSS Readers

Τα RSS Readers ή RSS Aggregator είναι προγράμματα που έχουν δημιουργηθεί, ώστε να καθίσταται δυνατή η αξιοποίηση των πληροφοριών από τον χρήστη, που του παρέχονται μέσω των RSS Feeds. Με την βοήθεια ενός εσωτερικού μηχανισμού παρακολούθησης των αλλαγών που πραγματοποιούνται στις ροές RSS, ενημερώνουν το χρήστη για τη διάθεση νέων πληροφοριών και τον προτρέπουν να τις δει. Βασικά χαρακτηριστικά των RSS Readers είναι:

- Απλή χρήση.
- Δυνατότητα στο χρήση καταχώρησης νέων RSS Feeds.
- Εύκολος εντοπισμός πληροφοριών από τα RSS Feeds που ήδη παρακολουθεί.

Σήμερα, οι πιο ευρέως διαδεδομένοι RSS Readers είναι web-based. Οι web-based readers είναι διαδικτυακές εφαρμογές, που δεν απαιτούν την εγκατάσταση κάποιου προγράμματος στον υπολογιστή του χρήστη ή την ανάγκη συγχρονισμού της λίστας των feeds που παρακολουθεί μεταξύ υπολογιστών και φορητών συσκευών τους οποίους ενδεχομένως να χρησιμοποιεί. Ο χρήστης λοιπόν μέσω των διαδικτυακών αυτών εφαρμογών έχει τη δυνατότητα να παρακολουθεί τα RSS Feeds, που έχει δηλώσει όπου και αν βρίσκεται.

Ενδεικτικό παράδειγμα αποτελεί ο Google Reader, ο οποίος ανήκει στις web-based εφαρμογές της Google. Μέσω του Goggle Reader, ο χρήστης μπορεί να παρακολουθεί τα feeds, να τα οργανώνει σε κατηγορίες, να σχολιάζει, να αξιολογεί, να τα κοινοποιεί σε φίλους του.

## 6.3 Java

### 6.3.1 Ορισμός

Η Java αποτελεί μια αντικειμενοστραφή γλώσσα προγραμματισμού. Δημιουργός της είναι η Sun Microsystems. Όλα τα εργαλεία που μπορεί να χρειαστεί κανείς κατά τη δημιουργία προγραμμάτων σε Java παρέχονται δωρεάν, από περιβάλλον ανάπτυξης μέχρι εργαλεία build.

Σημαντικό πλεονέκτημα της Java είναι η ανεξαρτησία που προσφέρει αναφορικά με το λειτουργικό σύστημα. Ένα πρόγραμμα γραμμένο σε Java είναι δυνατό να τρέχει ακριβώς το ίδιο σε Windows, Linux, Unix και Macintosh. Ο τρόπος για να επιτευχθεί αυτό ήταν η ανάπτυξη της Εικονικής Μηχανής (Virtual Machine).

### 6.3.2 Java Virtual Machine

Η εικονική μηχανή της Java είναι αυτή που καθιστά δυνατή την εκτέλεση λογισμικού γραμμένου σε Java. Ο τρόπος λειτουργίας της, αρχικά, ήταν ο εξής:

- Το πρόγραμμα γράφεται σε Java
- Το πρόγραμμα μεταγλωττίζεται μέσω του μεταγλωττιστή javac, ο οποίος παράγει ένα αριθμό από αρχεία .class (σε κώδικα byte).
- Η Java Virtual Machine, η οποία θα πρέπει να είναι ήδη εγκατεστημένη στο μηχάνημα που θέλουμε να εκτελεστεί η εφαρμογή αναλαμβάνει να διαβάσει τα αρχεία .class
- Τέλος, τα μεταφράζει σε γλώσσα μηχανής που να υποστηρίζεται από το εκάστοτε λειτουργικό σύστημα.

Βέβαια, πιο σύγχρονες εφαρμογές JVM έχουν την δυνατότητα να μεταγλωττίζουν εκ των προτέρων τμήματα bytecode απευθείας σε κώδικα μηχανής, με αποτέλεσμα η όλη διαδικασία να πραγματοποιείται με υψηλή ταχύτητα.



Τέλος, αξίζει να αναφερθεί ότι υπάρχουν διαφορετικές εκδόσεις ανάλογα με το είδος του λειτουργικού συστήματος της αρχιτεκτονικής του επεξεργαστή. Για παράδειγμα, υπάρχουν διαφορετικές εκδόσεις JVM για Windows, Linux, Unix, smart phones, playstations.

## 6.4 Apache SOLR

### 6.4.1 Ορισμός

Ο Apache SOLR είναι μια πολύ γρήγορη πλατφόρμα αναζήτησης ανοιχτού κώδικα γραμμένη σε Java και αποτελεί μέρος του Apache Lucent Project. Λειτουργεί ως αυτόνομος server αναζήτησης πλήρους κειμένου χρησιμοποιώντας ένα servlet Container, όπως Apache Tomcat ή Jetty. Ο SOLR χρησιμοποιεί την βιβλιοθήκη αναζήτησης της Lucene Java για την δημιουργία ευρετηρίου πλήρους κειμένου, ενώ παράλληλα μέσω των REST-like HTTP/XML και JSON APIs μπορεί να χρησιμοποιηθεί εύκολα από οποιαδήποτε γλώσσα προγραμματισμού.

Βασικό πλεονέκτημα του SOLR αποτελεί η άμεση προσαρμογή του σε πλήθος εφαρμογών, χωρίς την ανάγκη ύπαρξης Java κωδικοποίησης. Επιπλέον διατίθεται και plug-in για πιο προηγμένη παραμετροποίηση όταν αυτό απαιτείται.

Στα κύρια χαρακτηριστικά του συμπεριλαμβάνονται:

- Ισχυρή αναζήτηση πλήρους κειμένου.
- Πολύπλευρη έρευνα σχεδόν σε πραγματικό χρόνο ευρετηρίασης.
- Δυναμική ομαδοποίηση.
- Αναζήτηση εμπλουτισμένου κειμένου (word, pdf).

Ο SOLR είναι εξαιρετικά αξιόπιστος, επεκτάσιμος και ανεκτικός σε σφάλματα, παρέχοντας διανομές ευρετηρίασης και αντιγραφής με αυτοματοποιημένη προστασία αποτυχίας και ανάκτησης, κεντρική παραμετροποίηση και πολλά άλλα. Για τους παραπάνω λόγους, ο SOLR χρησιμοποιείται σήμερα σε αναζητήσεις και λειτουργίες πλοήγησης από τις μεγαλύτερες ιστοσελίδες στον κόσμο.

### 6.4.2 Ιστορικά στοιχεία

Ο SOLR δημιουργήθηκε από τον Yonik Seeley, στα πλαίσια ενός project για την αύξηση των δυνατοτήτων αναζήτησης του επιχειρησιακού site της CNET Networks. Τον Ιανουάριο του 2006, η CNET Networks αποφάσισε να προχωρήσει στη δωρεάν διάθεση του πηγαίου κώδικα, παραχωρώντας τα δικαιώματα στην Apache Software Foundation στα πλαίσια του έργου υψηλού επιπέδου Lucene. Όπως τα περισσότερα project που έχει αναλάβει η Apache Software Foundation, το ίδιο και ο SOLR, διήνυσε μια περίοδο «επώασης», κατά την οποία διευθετήθηκαν οργανωτικά, νομοθετικά και οικονομικά ζητήματα.

Στη συνέχεια, στις αρχές του 2007, ξεκίνησε η σταθερή του ανάπτυξη με την ενσωμάτωση συσσωρευμένων χαρακτηριστικών, που οδήγησαν στην προσέλκυση μιας μεγάλης κοινότητας χρηστών και συνεισφερόντων. Παρότι σε αρχικό ακόμη στάδιο, ο SOLR εξυπηρετούσε ήδη ιστοσελίδες με υψηλή επισκεψιμότητα.

Ακολουθεί η διάθεση αρκετών εκδόσεων (1.3/2008, 1.4/2009), μέχρι την ενοποίηση μεταξύ των δύο project SOLR και Lucene, το Μάρτιο του 2010. Το 2011 γίνεται αλλαγή στον τρόπο αρίθμησης των εκδόσεων του SOLR, έτσι ώστε να μπορούν εύκολα να αντιστοιχηθούν με την

αρίθμηση του Lucene. Η επόμενη έκδοση του SOLR είναι η 3.1, ακολουθεί η έκδοση 4.0 τον Οκτώβριο του 2012, η οποία συμπεριλαμβάνει το νέο πρόσθετο SOLR Cloud.

Τέλος, το 2013, γίνονται διαθέσιμες οι εκδόσεις 4.1 (Ιανουάριος), 4.2 (Μάρτιος), 4.2.1 (Απρίλιος), 4.3.0 (Μάιος), οι οποίες συμπεριλαμβάνουν σημαντικές βελτιώσεις των SOLR Cloud χαρακτηριστικών, όπως custom sharding.

### 6.4.3 Βασικά χαρακτηριστικά

- Χρησιμοποιεί τη βιβλιοθήκη Lucene για την αναζήτηση πλήρους κειμένου.
- Δυνατότητα πολύπλευρης πλοήγησης.
- Δυνατότητα επισήμανσης των «hits».
- Γλώσσα ερωτημάτων, που υποστηρίζει τη δομημένη αναζήτηση, καθώς και την αναζήτηση βάσει κειμένου.
- Διεπαφή με διαχειριστή HTML.
- Δυνατότητα αντιγραφής σε άλλους SOLR servers – δυνατή κλιμάκωση QPS.
- Δυνατότητα κατανεμημένης αναζήτησης μέσω Sharding – δυνατή κλιμάκωση όγκου περιεχομένου.
- Δυνατότητα επέκτασης μέσω της χρήσης προσθέτων.
- Ομαδοποίηση αποτελεσμάτων αναζήτησης με βάση το Carrot2.
- Ευελιξία ως προς τη συνάφεια, μέσω ενίσχυσης με συναρτήσεις ερωτημάτων.
- Δυνατότητα προσωρινής αποθήκευσης (caching) ερωτημάτων, φίλτρων, εγγράφων.
- Δυνατότητα ενσωμάτωσης του σε εφαρμογές Java.
- Αναζήτηση βάσει χωρών.
- Δυνατότητα αυτόματης διαχείρισης μεγάλων cluster μέσω του ZooKeeper.
- Πλήθος συναρτήσεων ερωτημάτων.
- Δημιουργία νέου στοιχείου αυτό-προτάσεων.

## 6.5 Jetty

### 6.5.1 Ορισμός

Ο Jetty είναι ένας αμιγώς βασιζόμενος σε Java HTTP Server και Java Servlet Container. Χρησιμοποιείται συχνά για τη διασύνδεση μεταξύ υπολογιστικών συστημάτων, που αποτελούν μέρος μεγαλύτερων λογισμικών πλαισίων. Έχει αναπτυχθεί από την Eclipse Foundation ως ένα ελεύθερο, ανοιχτού κώδικα project, από το 2009 οπότε και προσχώρησε σε αυτή. Χρησιμοποιείται ευρέως σε προϊόντα όπως ο Apache Active MQ, Alfresco, Apache Geronimo, Apache Maven, Apache Spark, Google App Engine, Eclipse, FUSE, Twitter's Streaming API and Zimbra, ενώ αποτελεί βασική επιλογή και σε project ανοιχτού κώδικα, όπως Lift, Eucalyptus, Red5, Hadoop and I2P. Επιπλέον, υποστηρίζει το τελευταίο Java Servlet API (με JSP υποστήριξη), καθώς και τα πρωτόκολλα SPDY και WebSocket.

Βασικά χαρακτηριστικά του είναι:

- Έχει πλήθος χαρακτηριστικών και βασίζεται σε πρότυπα.
- Αποτελεί λογισμικό ανοιχτού κώδικα.
- Ευέλικτο και επεκτάσιμο.
- Φιλικό προς επιχειρήσεις.
- Δυνατότητα ενσωμάτωσης.
- Υπό διπλή άδεια (Apache & Eclipse).

## 6.6 Apache Subversion

### 6.6.1 Εισαγωγικά στοιχεία

Ο Apache Subversion δημιουργήθηκε από την CollabNet Inc. το 2000 και αποτελεί σήμερα, ένα υψηλού επιπέδου Apache project, το οποίο δημιουργήθηκε και χρησιμοποιείται από μια παγκόσμια κοινότητα χρηστών. Πρόκειται για software versioning και revision control system, το οποίο διανέμεται δωρεάν από την Apache License. Ο Subversion χρησιμοποιείται για τη διατήρηση των τρεχόντων, αλλά και προηγούμενων εκδόσεων αρχείων, όπως πηγαίος κώδικας, ιστοσελίδες και τεκμηρίωση (documentation). Βασικός του στόχος είναι να αποτελέσει τον πιο συμβατό διάδοχο του ευρέως χρησιμοποιούμενου Concurrent Version System (CVS).

Ο Subversion έχει χρησιμοποιηθεί από την κοινότητα ελεύθερου λογισμικού σε project, όπως το Apache Software Foundation, Free Pascal, FreeBSD, GCC, Mono και SourceForge. Επιπλέον, το Google Code παρέχει δυνατότητα φιλοξενίας του Subversion στα ελεύθερου λογισμικού project της, καθώς και η CodePlex.

Παράλληλα, ο Subversion έχει υιοθετηθεί και από τον επιχειρηματικό κόσμο. Ενδεικτικά αναφέρουμε την ανακοίνωση από τη Forrester Research, το 2007, στην οποία αναγνωρίζεται η ηγετική του θέση στην κατηγορία Standalone Software Configuration Management (SCM). Αντίστοιχα, στην κατηγορία Software Configuration and Change Management (SCCM) χαρακτηρίστηκε ως δυνατός performer.

### 6.6.2 Ιστορικά στοιχεία

Το Subversion project δημιουργήθηκε το 2000 από την CollabNet, σε μια προσπάθεια συγγραφής ενός ανοιχτού κώδικα συστήματος ελέγχου εκδόσεων, το οποίο θα λειτουργούσε με τρόπο παρόμοιο με το CVS, αλλά θα προχωρούσε σε διόρθωση των σφαλμάτων και στην ενσωμάτωση χαρακτηριστικών που είχαν παραλειφθεί.

Μέχρι το 2001, η εξέλιξη του Subversion ήταν τέτοια, ώστε να είναι σε θέση να φιλοξενεί τον δικό του πηγαίο κώδικα, ενώ το Φεβρουάριο του 2004 γίνεται η διάθεση της 1.0 έκδοσής του. Τον Νοέμβριο του 2009 ακολουθεί στο Apache Incubator, με αποτέλεσμα μέχρι το Φεβρουάριο του 2010 να αποτελεί ένα υψηλού επιπέδου Apache Project.

## 6.7 Responsive Web Design

### 6.7.1 Ορισμός

Η ραγδαία αύξηση τα τελευταία χρόνια της χρήσης φορητών ηλεκτρονικών συσκευών (tablets, smartphones) για την πλοήγηση στο διαδίκτυο δημιούργησε πρόβλημα ως προς τον τρόπο εμφάνισης μιας ιστοσελίδας στις συσκευές αυτές. Με τη βασική διαφοροποίηση στην εμφάνιση μιας ιστοσελίδας να έγκειται στις διάφορες διαστάσεις των οθονών.

Στόχος του Responsive Web Design είναι η αποφυγή αυτού του προβλήματος και η παροχή της καλύτερης δυνατής εμπειρίας πλοήγησης στον χρήστη. Με τον όρο Responsive Web Design (RWD) εννοούμε τη διαδικασία σχεδιασμού και κατασκευής ιστοσελίδων, η οποία στοχεύει στην παροχή της βέλτιστης εμπειρίας θέασης, μέσω της εύκολης ανάγνωσης και πλοήγησης, με την ελάχιστη δυνατή αλλαγή μεγέθους, μετακίνησης και scrolling σε ένα ευρύ φάσμα συσκευών.

Η σχεδίαση ενός website με τη χρήση του RWD μπορεί να προσαρμόσει τη διάταξη της σελίδας στο περιβάλλον προβολής με τη χρήση:

- Ρευστών δικτύων (fluid grids), τα οποία απαιτούν τα στοιχεία μεγέθους μιας σελίδας να είναι εκφρασμένα σε σχετικές μονάδες, όπως ποσοστά και όχι σε απόλυτες, όπως pixels.
- Ευέλικτων εικόνων, οι οποίες και πάλι θα πρέπει να εκφράζονται σε σχετικές μονάδες, ώστε να αποφεύγεται το ενδεχόμενο προβολής τους εκτός πλαισίου.
- Media queries, τα οποία επιτρέπουν στην ιστοσελίδα να χρησιμοποιεί διαφορετικά CSS στυλ, ανάλογα με τα χαρακτηριστικά της συσκευής από την οποία γίνεται η πλοήγηση στη σελίδα, όπως το εύρος του browser.
- Στοιχείων server-side (RESS) σε συνδυασμό με στοιχεία client-side, έτσι ώστε να επιτυγχάνεται μεγαλύτερη ταχύτητα downloading, όταν η πρόσβαση γίνεται μέσω δικτύων κινητής τηλεφωνίας. Επιπλέον, με τον παραπάνω συνδυασμό γίνεται δυνατή η παροχή πλουσιότερης λειτουργικότητας/χρησιμότητας αποφεύγοντας τις παγίδες από την αποκλειστική χρήση device-side λύσεων.

### 6.7.2 Πλεονεκτήματα

Βασικά πλεονεκτήματα από τη χρήση του Responsive Web Design θα μπορούσαν να είναι:

- Παρέχει την καλύτερη δυνατή εμπειρία πλοήγησης από οποιαδήποτε συσκευή και αν επισκεφτούμε ένα website.
- Analytics: Παρέχει τη δυνατότητα δημιουργίας πλήρους αναφοράς σχετικά με τις επισκέψεις μιας ιστοσελίδας από διαφορετικού είδους συσκευές.
- Παρέχει τη δυνατότητα δημιουργίας κοινών συνδέσμων (links).
- SEO: Παρέχει τη δυνατότητα ένα url να συμπεριλαμβάνει όλα τα links που θέλουμε.
- Παρέχει μεγαλύτερη ευκολία συντήρησης και περιορισμό των σφαλμάτων.
- Μειώνει το χρόνο σχεδίασης της δομής ενός website για διαφορετικές συσκευές.

Συμπερασματικά, η χρήση του Responsive Web Design αποτελεί κυρίαρχη τάση στο χώρο του web design, παρόλα αυτά η σχεδίαση μιας ιστοσελίδας με την τεχνική RWD δεν αποτελεί ένα εύκολο εγχείρημα, καθώς απαιτεί γνώσεις HTML, CSS και Javascript.

## 6.8 XPath

### 6.8.1 Εννοιολογικός Προσδιορισμός

Η ανάπτυξη της XPath (XML Path Language) ξεκίνησε από την W3C (World Wide Web Consortium), το 1999, με την τελευταία έκδοσή της, την 3.0 να γίνεται διαθέσιμη τον Απρίλιο του 2014. Παρόλα αυτά, η πιο ευρέως διαδεδομένη και χρησιμοποιούμενη έκδοσή της μέχρι σήμερα δεν παύει να είναι η 1.0 (Νοέμβριος, 1999). Η XPath 1.0 εφαρμόζεται είτε αυτόνομα, μέσω ενός API από γλώσσες, όπως Java, C, Python ή JavaScript, είτε ενσωματωμένη σε γλώσσες, όπως XSLT, XProc, XML Schema ή XForms.

Πρόκειται για μια γλώσσα ερωτημάτων για την επιλογή «κόμβων» (“nodes”) σε ένα XML έγγραφο. Στηρίζεται στη αναπαράσταση της δομής του εγγράφου XML σαν ένα δένδρο με κόμβους και παρέχει τη δυνατότητα πλοήγησης μεταξύ των διάφορων κόμβων, καθώς και επιλογής τους με βάσει ποικίλα κριτήρια.

Μια έκφραση XPath, η οποία αποτελεί βασική συντακτική έννοια, είναι μια σειρά από βήματα τοποθεσίας (location steps) χωρισμένα με τον χαρακτήρα / (slash). Κάθε βήμα αποτελείται από τρία στοιχεία:

- Έναν άξονα, ο οποίος ορίζει την κατεύθυνση προς την οποία θα κινηθούμε μέσα στο δένδρο.
- Έναν έλεγχο κόμβου, με τον οποίο ελέγχεται εάν οι κόμβοι που συναντώνται κατά το μήκος ενός συγκεκριμένου άξονα θα πρέπει να επιλεγούν ή όχι για το επόμενο βήμα.
- Μηδέν ή περισσότερα κατηγορήματα, δηλαδή μια έκφραση, με μορφή ισότητας (=, !=, >, <), η οποία φιλτράρει τους επιλεγμένους κόμβους από το προηγούμενο στάδιο.

Τα αντικείμενα που είναι δυνατό να παραχθούν από την αποτίμηση μιας έκφρασης μπορεί να έχουν τους εξής τύπους:

- Node-set, δηλαδή μια αταξινόμητη συλλογή κόμβων χωρίς επαναλήψεις.
- Boolean, δηλαδή Αλήθεια ή Ψέματα
- Number, δηλαδή ένας αριθμός κινητής υποδιαστολής και
- String, δηλαδή μια ακολουθία από UCS χαρακτήρες.

Επιπλέον, αξίζει να σημειωθεί η δυνατότητα της XPath υποστήριξης XML Namespace, έτσι ώστε το όνομα ενός κόμβου να μοντελοποιείται ως ζεύγος αποτελούμενο από ένα τοπικό κομμάτι και ένα πιθανά κενό namespace URL (expanded-name).

Ένα από τα σημαντικότερα είδη εκφράσεων της XPath είναι το μονοπάτι τοποθεσίας (location path). Ουσιαστικά, ένα μονοπάτι τοποθεσίας επιλέγει το σύνολο των κόμβων, οι οποίοι σχετίζονται με τον κόμβο περιεχομένου. Το αποτέλεσμα της αποτίμησης ενός μονοπατιού τοποθεσίας είναι το σύνολο κόμβων (node-set), το οποίο περιλαμβάνει τους επιλεγμένους από το μονοπάτι τοποθεσίας κόμβους.

## 6.9 RegEx

### 6.9.1 Εννοιολογικός Προσδιορισμός

Η RegEx (Regular expression) αναφέρεται σε μια αλληλουχία χαρακτήρων, η οποία δημιουργεί ένα μοτίβο αναζήτησης, το οποίο χρησιμοποιείται κυρίως στην αντιστοίχιση string. Κάθε χαρακτήρας στη RegEx μπορεί να έχει δύο μορφές, είτε του τυπικού χαρακτήρα, όποτε εκλαμβάνεται με την κυριολεκτική του ερμηνεία, είτε ως μετα-χαρακτήρας, προσδίδοντας του μια ιδιαίτερη ερμηνεία.

Οι δύο αυτές μορφές μπορούν να συνδυαστούν για τον προσδιορισμό του κειμένου ενός δεδομένου προτύπου, είτε για την επεξεργασία μιας σειράς εμφανίσεών του, οι οποίες μπορεί να κυμαίνονται από την απόλυτη ακρίβεια έως μια γενική ομοιότητα.

Πρόκειται για μια γλώσσα που έχει σχεδιασθεί ειδικά για την αυτοματοποίηση της επεξεργασίας κειμένου, γενικών αρχείων κειμένου, ειδικών μορφών κειμένου, αλλά και συμβολοσειρών εισόδου, με τον πιο περιεκτικό και ευέλικτο τρόπο.

Η χρήση των regular expressions εμφανίστηκε το 1960 και επεκτάθηκε το 1980 με την ενοποίηση των βιομηχανικών προτύπων, όπως το ISO SGML. Το 1997 ο Rhipil Hazel ανέπτυξε την Perl Compatible Regular Expressions (PCRE), η οποία προσπάθησε να μιμηθεί την λειτουργικότητα της regular expression της Perl και χρησιμοποιείται σε πολλά σύγχρονα εργαλεία, όπως η PHP και ο Apache Http Server.

Σήμερα, η RegEx έχει καταστεί ιδιαίτερα χρήσιμη, με αποτέλεσμα την μεγάλη ανάπτυξη της, αναφορικά με βασικά αλλά και εξελιγμένα πρότυπα γραμματικής και συντακτικής μορφοποίησης. Οι επεξεργαστές RegEx, συναντώνται σε διάφορες μηχανές αναζήτησης, σε παράθυρα διαλόγου αναζήτησης-αντικατάστασης διάφορων επεξεργαστών κειμένου, καθώς και σε γραμμές εντολών text processing utilities, όπως sed και AWK.

Τέλος, πολλές γλώσσες προγραμματισμού παρέχουν δυνατότητες RegEx, είτε ενσωματωμένες, όπως η Perl, Ruby, AWK και Tcl, είτε μέσω τυπικών βιβλιοθηκών, όπως η Java, Python, C++.

## 6.10 Style Sheet Language

### 6.10.1 Εννοιολογικός προσδιορισμός

Μια γλώσσα φύλλων στυλ (style sheet language) είναι μια γλώσσα μηχανής, η οποία αφορά στη δημιουργία και παρουσίαση δομημένων εγγράφων. Ένα από τα βασικά χαρακτηριστικά της είναι η δυνατότητα που παρέχει, μέσω της δημιουργίας δομημένων εγγραφών, το περιεχόμενο να μπορεί να επαναχρησιμοποιηθεί σε διάφορα πλαίσια και να παρουσιαστεί με διάφορους τρόπους.

Στις πιο γνωστές γλώσσες φύλλων στυλ συμπεριλαμβάνονται η CSS (Cascading Style Sheet), η XSL (Extensible StyleSheet Language) και DSSSL (Documents Style Semantics and Specification Language).

Στα βασικά δομικά συστατικά μιας StyleSheet γλώσσας αναφέρονται τα εξής:

- ✓ Σύνταξη, η οποία απαιτείται ώστε να εκφραστεί το στυλ με τρόπο που να μπορεί να διαβαστεί.
- ✓ Επιλογείς, οι οποίοι καθορίζουν ποια στοιχεία θα επηρεαστούν από τον κανόνα στυλ και ποια όχι. Θα μπορούσαμε να πούμε ότι οι επιλογείς αποτελούν το συνδετικό στοιχείο μεταξύ της δομής του εγγράφου και των κανόνων των φύλλων στυλ.
- ✓ Ιδιότητες. Όλες οι γλώσσες φύλλων στυλ συνοδεύονται από ένα πλαίσιο ιδιοτήτων, στις οποίες δίνονται αντίστοιχες τιμές, με στόχο την αλλαγή του τρόπου απόδοσης ενός στοιχείου. Οι περισσότερες παρέχουν περίπου 50 ιδιότητες για να περιγράψουν τον τρόπο παρουσίασης του εγγράφου.
- ✓ Τιμές και Μονάδες. Δίνοντας μια συγκεκριμένη τιμή σε μια ιδιότητα, όπως προαναφέρθηκε, μπορούμε να επεμβαίνουμε στον τρόπο παρουσίασης ενός στοιχείου. Η τιμή μπορεί να είναι μια συμβολοσειρά, μια λέξη κλειδί, ένα νούμερο με μια μονάδα αναγνώρισης, μια λίστα ή μια έκφραση όλων των παραπάνω. Οι περισσότερες γλώσσες φύλλων στυλ παρέχουν περίπου 10 μονάδες αναφοράς.
- ✓ Μηχανισμός διάδοσης τιμών. Για να αποφευχθεί η ανάγκη προσδιορισμού τιμών για όλες τις ιδιότητες, οι γλώσσες φύλλων στυλ διαθέτουν μηχανισμούς, έτσι ώστε οι τιμές να δίδονται αυτόματα.
- ✓ Μοντέλα μορφοποίησης. Στις περισσότερες stylesheet γλώσσες συναντάται ένα οπτικό μοντέλο μορφοποίησης, το οποίο περιγράφει αρκετά λεπτομερώς τον τρόπο μορφοποίησης του κειμένου και περιεχομένου στη τελική παρουσίαση.

### 6.10.2 CSS (Cascading Style Sheet)

Από τις πιο ευρέως χρησιμοποιούμενες γλώσσες φύλλων στυλ είναι και η CSS (ή διαφορετικά Αλληλουχία Φύλλων Στυλ). Η CSS χρησιμοποιείται για τον έλεγχο της εμφάνισης ενός εγγράφου, το οποίο έχει γραφεί σε γλώσσες HTML και XHTML. Στόχος της είναι η στυλιστική διαμόρφωση μιας ιστοσελίδας, μέσω του καθορισμού διαφόρων χαρακτηριστικών, όπως χρώματα, μέγεθος, στοίχιση κ.α..

Αξίζει να σημειωθεί ότι η εφαρμογή του προτύπου CSS έγκειται στην ύπαρξη περιπτώσεων, όπου για ένα έγγραφο θα υπάρχουν περισσότερα από ένα φύλλα στυλ, που περιέχουν δηλώσεις για την εμφάνιση ενός στοιχείου. Ένα φύλλο στυλ μπορεί να έχει δημιουργηθεί από τον συγγραφέα μιας ιστοσελίδας, το χρήστη του πλοηγού ή ακόμη και τον ίδιο τον πλοηγό.

Στην περίπτωση ύπαρξης συγκρουόμενων δηλώσεων για ένα στοιχείο η CSS έχει καθορίσει μια συγκεκριμένη αλληλουχία με την οποία θα ταξινομηθούν οι δηλώσεις και θα εφαρμοστεί η πρώτη στη σειρά. Η δημιουργία αυτής της αλληλουχίας στηρίζεται στα εξής στάδια:

- ✓ Ανεύρεση όλων των δηλώσεων που αφορούν στο στοιχείο που μας ενδιαφέρει.
- ✓ Ταξινόμηση τους με βάση τη σημασία (κανονική – σημαντική) και την προέλευση (χρήστης – συγγραφέας - πλοηγός χρήστη).
- ✓ Σε περίπτωση δηλώσεων ίδιας σημασίας και προέλευσης, η ταξινόμηση γίνεται με βάση την εξειδίκευση του επιλογέα. Υπερισχύουν οι πιο εξειδικευμένοι επιλογείς έναντι των πιο γενικών.
- ✓ Τέλος, η ταξινόμηση δηλώσεων ίδιας σημασίας, προέλευσης και εξειδίκευσης γίνεται με βάση το χρόνο προσδιορισμού. Η δήλωση που προσδιορίστηκε τελευταία επικρατεί.

Ακλουθώντας τους παραπάνω κανόνες δημιουργείται μια σειρά δηλώσεων, η τελευταία της οποίας επιλέγεται για την εμφάνιση του στοιχείου.





## Κεφάλαιο Έβδομο: Υλοποίηση

### 7.1 Διάφορες Απαιτούμενες Εγκαταστάσεις στο server

Για την σωστή και απρόσκοπτη δημιουργία και εκτέλεση της εφαρμογής έγιναν κάποιες επιπλέον εγκαταστάσεις, οι οποίες περιγράφονται στις επιμέρους ενότητες.

#### 7.1.1. Εγκατάσταση Samba

Η Samba είναι μια σουίτα λογισμικού που επιτρέπει σε υπολογιστές με διαφορετικά λογισμικά, μέσω της σύνδεσής τους σε δίκτυο, να επιτυγχάνουν την μεταφορά και κοινή χρήση αρχείων και εκτυπωτών, καθώς και άλλων υπηρεσιών, χωρίς cross-platform συγκρούσεις. Στην παρούσα εργασία χρησιμοποιήθηκε για να κάνουμε κοινόχρηστα αρχεία από τη virtual machine, που αποτελεί τον server, στον τοπικό υπολογιστή. Οδηγηθήκαμε σε αυτή την επιλογή, έτσι ώστε να κρατήσουμε την μνήμη σε λογικά επίπεδα και να μην χρειαστεί να εγκαταστήσουμε IDE.

Για την εγκατάσταση χρησιμοποιούμε το terminal των Linux και εκτελούμε την εντολή `sudo apt-get install samba`. Για να μπούμε στο configuration αρχείο του Samba εκτελούμε την εντολή `sudo pico /etc/samba/smb.conf` και προσθέτουμε:

```
[aggregator]
path = /var/www/aggregator
available = yes
valid users = aggregator
read only = no
browseable = yes
public = yes
writable = yes
create mask = 0644
directory mask = 0755
```

#### 7.1.2. Εγκατάσταση Drush

Το Drush είναι μια shell-based εφαρμογή με μία μεγάλη γκάμα χαρακτηριστικών και δυνατοτήτων, η οποία χρησιμοποιείται κυρίως για τον έλεγχο, τον χειρισμό και τη διαχείριση ιστοσελίδων Drupal, καθώς και για την ενημέρωση των modules.

Για την εγκατάσταση του χρησιμοποιούμε το terminal των Linux και γράφουμε την εντολή `sudo apt-get install drush`. Η έκδοση που γίνεται εγκατάσταση είναι η 4.5 και για την αναβάθμιση του στην τελευταία έκδοση χρησιμοποιούμε την εντολή `sudo apt-get install drush /share`. Για τον έλεγχο της σωστής αναβάθμισης χρησιμοποιούμε την εντολή `drush -version`.

### 7.1.3 Εγκατάσταση Compass

Πρόκειται για ένα εργαλείο που αφορά στα φύλλα συγγραφής. Βασίζεται σε Sass, ένα στυλ γλώσσας συγγραφής, παρόμοια με τη CSS, αλλά με ένθετα κανόνες και μεταβλητές. Για την εγκατάσταση του χρησιμοποιούμε το terminal των Linux και εκτελούμε τις παρακάτω εντολές:

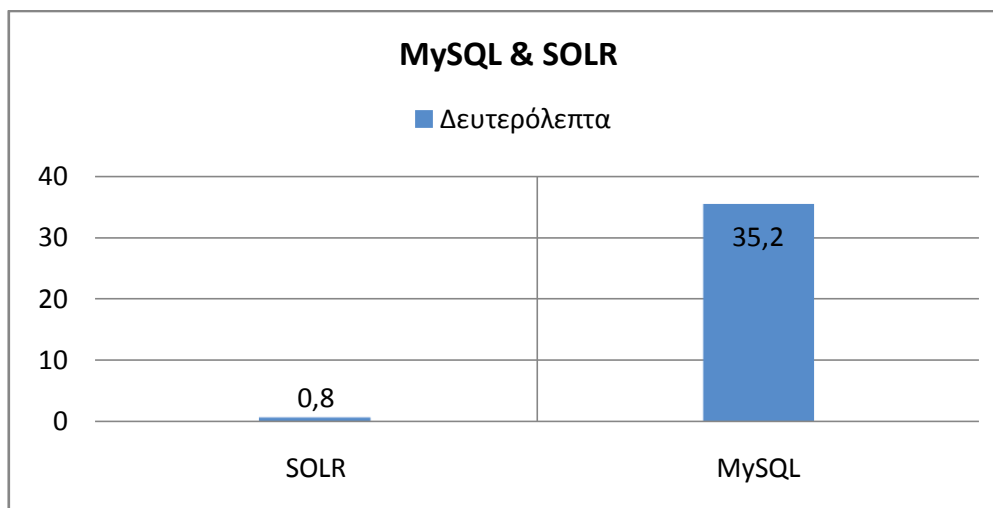
```
gem update --system
gem install compass
```

### 7.1.4 Εγκατάσταση SOLR

Ο SOLR όπως έχει αναφερθεί και στο θεωρητικό κομμάτι είναι μια πολύ γρήγορη πλατφόρμα αναζήτησης, ανοιχτού κώδικα, γραμμένη σε Java. Για τον λόγο αυτό, πριν την εγκατάσταση του θα πρέπει να έχει προηγηθεί η εγκατάσταση της Java. Χρησιμοποιούμε το terminal των Linux και εκτελούμε την εντολή `sudo apt-get install sun-java6-plugin`. Έπειτα με την εντολή `java -version` επιβεβαιώνουμε την σωστή εγκατάσταση της.

Στην συνέχεια κατεβάσαμε την έκδοση 4.8 του SOLR και αντιγράψαμε τα αρχεία ρυθμίσεων που βρίσκονται στην διαδρομή `SOLR-conf/directory/4.8` σε κατάλογο διαμόρφωσης του SOLR `“solr/aggregator/SOLR/collection1/conf/4.8”`. Για την εκτέλεση του SOLR από την διαδρομή `“$SOLR/aggregator/”` δίνουμε την εντολή `java -jar start.jar` και από την σελίδα διαχείρισης του SOLR μπορούμε να επιβεβαιώσουμε ότι λειτουργεί σωστά.

Στο σημείο αυτό αξίζει να αναφερθεί και η υπεροχή του SOLR έναντι της MySQL σε ταχύτητα αναζήτησης. Η σύγκριση έγινε ως προς την αναζήτηση λέξης – κλειδιού σε πλήθος 5.000.000 εγγραφών. Τα αποτελέσματα εμφανίζονται στο γράφημα που ακολουθεί:



Εικόνα 13: Ταχύτητα αναζήτησης SOLR vs MySQL

### 7.1.5 Δημιουργία SVN

Ο Subversion χρησιμοποιείται για τη διατήρηση των τρεχόντων, αλλά και προηγούμενων εκδόσεων αρχείων, όπως πηγαίος κώδικας, ιστοσελίδες και τεκμηρίωση. Για την δημιουργία του χρησιμοποιούμε το terminal των Linux και δίνουμε τις δυο παρακάτω εντολές:

```
-> mkdir /home/svn/aggregator
-> svnadmin create /home/svn/aggregator
```

Για την δημιουργία ενός αρχείου κωδικών για τους χρήστες της εφαρμογής δίνουμε την εντολή:

```
-> htpasswd -c /etc/subversion/passwd aggregator
```

Τέλος, για την σύνδεση του με τον Apache server δίνουμε την εντολή:

```
-> sudo pico /etc/apache2/mods-enabled/dav_svn.conf
```

και προσθέτουμε:

```
<Location /svn/eagggregator>
  DAV svn
  SVNPath /home/svn/eagggregator
  AuthType Basic
  AuthName "myproject subversion repository"
  AuthUserFile /etc/subversion/passwd
  <LimitExcept GET PROPFIND OPTIONS REPORT>
    Require valid-user
  </LimitExcept>
</Location>
```

## 7.2 Modules που χρησιμοποιήθηκαν

Για την ολοκλήρωση της εφαρμογής χρησιμοποιήσαμε κάποια έτοιμα modules πάνω σε Drupal, αλλά και κάποια modules που αναπτύξαμε οι ίδιοι για τις ανάγκες της εφαρμογής. Στη συνέχεια γίνεται περιγραφή των παραπάνω modules.

### 7.2.1 Έτοιμα Modules

Στην κατηγορία αυτή συμπεριλαμβάνονται τα εξής modules:

- ✓ **Admin\_menu.** Παρέχει ένα περιβάλλον διαχείρισης ανεξάρτητο από το επιλεγμένο θέμα. Τα διαχειριστικά link εμφανίζονται μέσω ενός CSS/JS βασιζόμενου μενού στην κορυφή όλων των σελίδων ενός site. Περιλαμβάνει πέρα από τις βασικές λειτουργίες του μενού, δυνατότητα ολοκλήρωσης εργασιών και ενεργειών, παρέχοντας έτσι εύκολη πρόσβαση σε κάθε πηγή διαχείρισης.
- ✓ **Boolean.** Παρέχει ένα “boolean” τύπο πεδίου με την επιπρόσθετη δυνατότητα επιλογής “not set”, πέρα των συνηθισμένων “true/false” ή “on/off”.
- ✓ **Captcha.** Πρόκειται για ένα τεστ με παροχή μιας απάντησης-πρόκλησης, που εμφανίζεται σε διαδικτυακές φόρμες, με σκοπό να καθορίσει εάν ο χρήστης είναι άνθρωπος. Στόχος του module είναι να αποκλείσει και να προστατεύσει την ιστοσελίδα από οποιαδήποτε προσπάθεια κακόβουλης χρήσης.
- ✓ **Compact\_forms.** Παρέχει τη δυνατότητα μιας πιο συμπαγούς εμφάνισης για τα πεδία κειμένου. Σε μια φόρμα οι ετικέτες βρίσκονται σε υπέρθεση των αντίστοιχων πεδίων. Όταν ο χρήστης εστιάζει σε ένα πεδίο η ετικέτα «σβήνει» με αργό ρυθμό, ενώ σε περίπτωση που το πεδίο παραμένει κενό η ετικέτα επανέρχεται.
- ✓ **Ctools (Chaos Tool Suite).** Πρόκειται για ένα σύνολο APIs και εργαλείων σχεδιασμένα για να βελτιώσουν την εμπειρία αυτού που ασχολείται με την ανάπτυξη της ιστοσελίδας.

- ✓ **Devel.** Παρέχει τη δυνατότητα εκτύπωσης περίληψης όλων των ερωτημάτων (query) της βάσης δεδομένων για τις αναζητήσεις που έγιναν ανά σελίδα, εμφανίζοντάς τη στο κάτω μέρος της σελίδας. Η περίληψη περιλαμβάνει πληροφορίες σχετικά με το πόσες φορές εκτελέστηκε το κάθε ερώτημα σε μια σελίδα, αλλά και για το χρόνο που διήρκησε η εκτέλεση του κάθε ερωτήματος.
- ✓ **Entity.** Το συγκεκριμένο module επεκτείνει την διαπαφή προγραμματισμού εφαρμογών (API) του πυρήνα του Drupal παρέχοντας έτσι έναν ενοποιημένο τρόπο αντιμετώπισης των διαφόρων οντοτήτων (entities), καθώς και των ιδιοτήτων τους.
- ✓ **Facetapi.** Επιτρέπει στους δημιουργούς μιας ιστοσελίδας να αναπτύξουν και να διαχειριστούν με ευκολία διασυνδέσεις πολύπλευρων αναζητήσεων.
- ✓ **Facetapi\_collapsible.** Πρόκειται για ένα πρόσθετο widget. Δίνεται η δυνατότητα στα facets να ρυθμιστούν με τέτοιο τρόπο ώστε είτε να παραμένουν ανοιχτά όταν ανοίγει ένα άλλο, είτε να κλείνουν.
- ✓ **Facetapi\_slider.** Πρόκειται για ένα widget σχετικά με τη ρύθμιση του εύρους.
- ✓ **Features.** Το συγκεκριμένο module επιτρέπει τη σύλληψη και διαχείριση των διάφορων λειτουργιών (features) του Drupal. Με τον όρο feature εννοούμε ένα σύνολο οντοτήτων του Drupal, ο συνδυασμός των οποίων εξυπηρετεί ένα συγκεκριμένο σκοπό.
- ✓ **Features\_extra.** Παρέχει τη δυνατότητα εικονικών εξαγωγών, μέσω των features διαφόρων δομικών στοιχείων κατασκευής ενός site.
- ✓ **Feeds.** Παρέχει τη δυνατότητα εισαγωγής ή συγκέντρωσης δεδομένων αναφορικά με κόμβους, χρήστες, κανόνες ταξινόμησης ή απλές εγγραφές της βάσης δεδομένων.
- ✓ **Feeds\_crawler.** Πρόκειται για ένα πρόσθετο ιδιαίτερα χρήσιμο για τη δημιουργία δυναμικών web scrapers μέσω feeds. Για παράδειγμα, προσφέρει υψηλές δυνατότητες σελιδοποίησης url patterns μέσω της χρήσης ενός αρχικού δείκτη και καθορισμού του πλήθους των μέγιστων αποτελεσμάτων. Επιπλέον, δύναται να εντοπίσει αυτοματοποιημένα το επόμενο link για RSS ιστοσελίδες.
- ✓ **Feeds\_tamper.** Πρόκειται για ένα πρόσθετο το οποίο μπορεί να τροποποιεί τα δεδομένα πριν την αποθήκευσή τους.
- ✓ **Feeds\_XPathparser.** Παρέχει δυνατότητες ανάλυσης κειμένων XML και HTML. Επιτρέπει στους δημιουργούς μιας ιστοσελίδας να πολλαπλασιάζουν τη δύναμη των feeds, έτσι ώστε να εισάγονται με μεγάλη ευκολία δεδομένα από πολύπλοκες, εξωτερικές πηγές.
- ✓ **File (Field) Paths.** Το συγκεκριμένο module αυξάνει τη βασική λειτουργικότητα του upload module με στόχο την καλύτερη ταξινόμηση των αρχείων.
- ✓ **Job\_scheduler.** Παρέχει τη δυνατότητα προγραμματισμού πραγματοποίησης μιας εργασίας σε ένα προκαθορισμένο χρονικό σημείο ή ανά δεδομένα χρονικά διαστήματα.
- ✓ **Jquery\_update.** Αναβαθμίζει την έκδοση του JQuery του πυρήνα Drupal σε νεότερη.
- ✓ **Libraries.** Ο κοινός παρονομαστής όλων των modules/προφίλ/θεμάτων του Drupal που ενσωματώνονται με εξωτερικές βιβλιοθήκες.
- ✓ **Masonry και Masonry\_views.** Παρέχει δυνατότητα ενσωμάτωσης του πρόσθετου jQuery Masonry με το Drupal.
- ✓ **Module\_filter.** Μια σελίδα με τη λίστα όλων των modules μπορεί να είναι αρκετά εκτενής, ειδικά όταν έχουμε να αντιμετωπίσουμε ένα site που φιλοξενεί μεγάλο όγκο δεδομένων. Στόχος του συγκεκριμένου module είναι η δυνατότητα να εντοπίζει κανείς το module που ψάχνει εύκολα και γρήγορα χωρίς να εξαρτάται από τα ιδιαίτερα χαρακτηριστικά της μηχανής αναζήτησης του browser.
- ✓ **Responsive\_menus.** Χρησιμοποιώντας ένα CSS/jQuery επιλογή μενού, το συγκεκριμένο module μπορεί να τροποποιεί την εμφάνιση του μενού, ώστε να είναι συμβατό και προσαρμόσιμο σε κάθε συσκευή, όπως smartphone.
- ✓ **Search\_api.** Αυτό το module παρέχει ένα πλαίσιο για την εύκολη δημιουργία αναζήτησης οποιασδήποτε οντότητας του Drupal με τη χρήση οποιασδήποτε μηχανής αναζήτησης.
- ✓ **Search\_api\_ranges.** Παρέχει δυνατότητες αναζήτησης με βάση το εύρος.
- ✓ **Search\_api\_solr.** Παρέχει ένα σύστημα υποστήριξης του SOLR για το search api module.
- ✓ **Search\_api\_sorts.** Παρέχει δυνατότητες προσαρμοσμένης ταξινόμησης.
- ✓ **Strongarm.** Παρέχει την δυνατότητα στους δημιουργούς ενός site να παρακάμπτουν τις προκαθορισμένες τιμές των μεταβλητών του πυρήνα Drupal.

- ✓ **Taxonomy\_manager.** Παρέχει ένα δυναμικό περιβάλλον διαχείρισης ταξινόμησης. Για παράδειγμα το λεξιλόγιο παρουσιάζεται σε μια δυναμική μορφή δέντρου, με ανάπτυξη ή σύμπτυξη κινούμενοι στις διάφορες κατηγορίες.
- ✓ **Token.** Παρέχει τη δυνατότητα σε άλλα modules να χρησιμοποιούν τα tokens, δηλαδή μικρά κομμάτια κειμένου που ενσωματώνονται σε μεγαλύτερα, μέσω διαδρομών.
- ✓ **Transliteration.** Με το module αυτό τροποποιούνται τα ονόματα των αρχείων που ανεβάζει κανείς με την αντικατάσταση των ανεπιθύμητων χαρακτήρων. Ουσιαστικά, προσπαθεί να μετατρέψει ένα Unicode κείμενο σε US-ASCII χαρακτήρες.
- ✓ **Ultimate\_cron.** Εκτελεί μεμονωμένες εργασίες χρησιμοποιώντας παραμετροποιήσιμους κανόνες.
- ✓ **Uuid (Universally Unique Identifier).** Με αυτό το module καθίσταται εφικτή η προσθήκη UUIDs στα διάφορα αντικείμενα του Drupal.
- ✓ **Uuid\_features.** Το UUID Features Integration module παρέχει τη δυνατότητα εξαγωγής περιεχομένου (nodes, ταξινόμηση, πεδία) σε ένα features module.
- ✓ **Views.** Παρέχει δυνατότητες τροποποίησης του προκαθορισμένου τρόπου εμφάνισης.
- ✓ **Views\_field\_view.** Παρέχει δυνατότητα ενσωμάτωσης μιας προβολής ως πεδίο της προβολής.

## 7.2.2 Modules που αναπτύχθηκαν

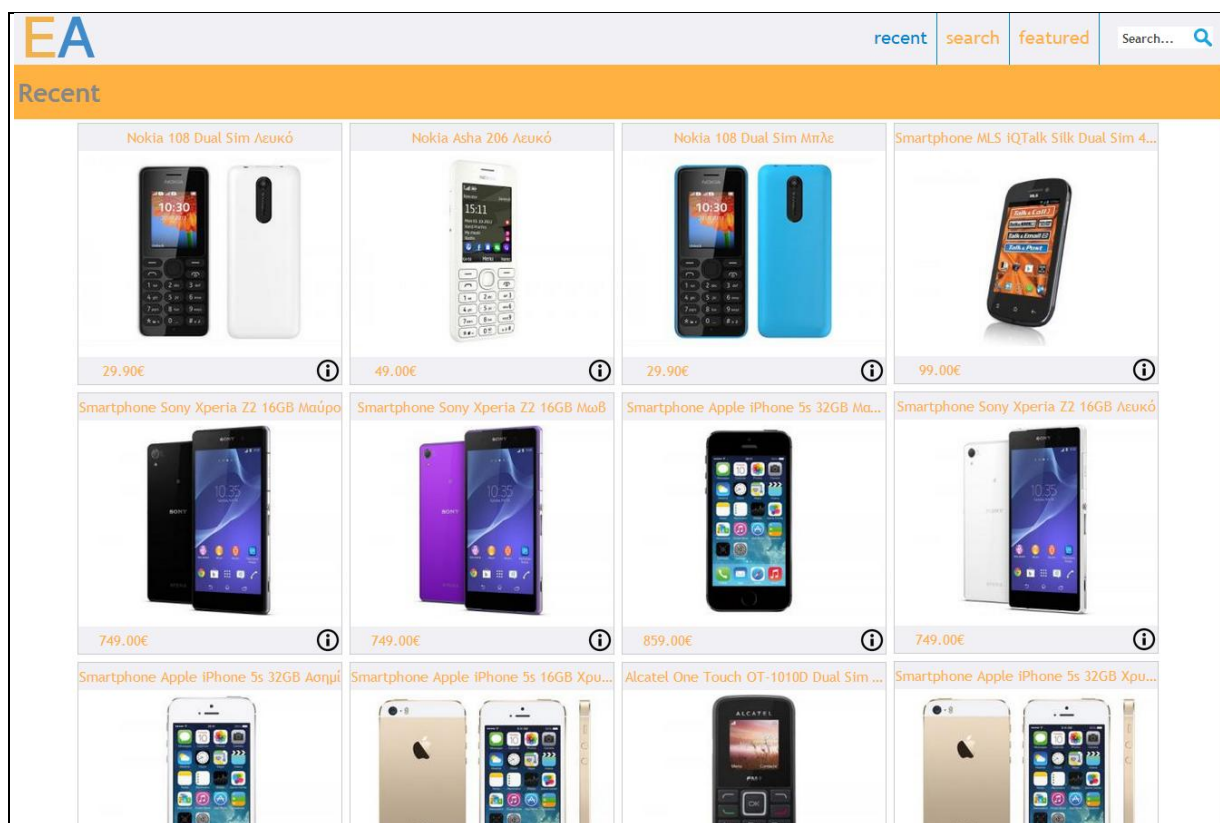
Κατά την υλοποίηση της εφαρμογής παρουσιάστηκε η ανάγκη για τη δημιουργία νέων modules που θα καλύπτουν όλες τις απαιτήσεις και τα οποία είναι τα εξής:

- ✓ **Custom\_crawler\_configuration.** Module που χρησιμοποιείται για την παραμετροποίηση του τρόπου εισαγωγής των δεδομένων από τρίτα sites, παρέχοντας ένα interface όπου ο χρήστης εισάγει τα XPath και RegEx patterns. Αυτό δημιουργεί ένα αντικείμενο που χρησιμοποιείται αργότερα από το module custom\_post\_process.
- ✓ **Custom\_general\_functions.** Γενικά functions τα οποία χρησιμοποιούνται από ένα ή περισσότερα άλλα modules, όπως για παράδειγμα trim functions, clean text κτλ.
- ✓ **Custom\_post\_process.** Το βασικότερο module του site. Περιέχει τις διαδικασίες που χρησιμοποιούνται για το crawling και την εισαγωγή των δεδομένων από τρίτα sites. Αναλαμβάνει, σύμφωνα με τις ρυθμίσεις που έχει δώσει ο διαχειριστής, να διαβάσει και να αποθηκεύσει τα προϊόντα.
- ✓ **Custom\_search\_box.** Παρέχει ένα block αναζήτησης στον χρήστη.
- ✓ **Custom\_url\_relative\_absolute.** Module βιβλιοθήκης που καλείται για τη σωστή εισαγωγή των URLs. Προσπαθεί να μετατρέψει συνδυασμούς σχετικών URLs σε απόλυτα URLs.
- ✓ **Merge\_synonyms.** Module που χρησιμοποιείται για την αυτόματη και μη διαχείριση των συνώνυμων κατηγοριών (terms) που εισάγονται από τρίτα sites, για παράδειγμα συγχώνευση των terms “Smartphones” και “Εξυπνα κινητά”.

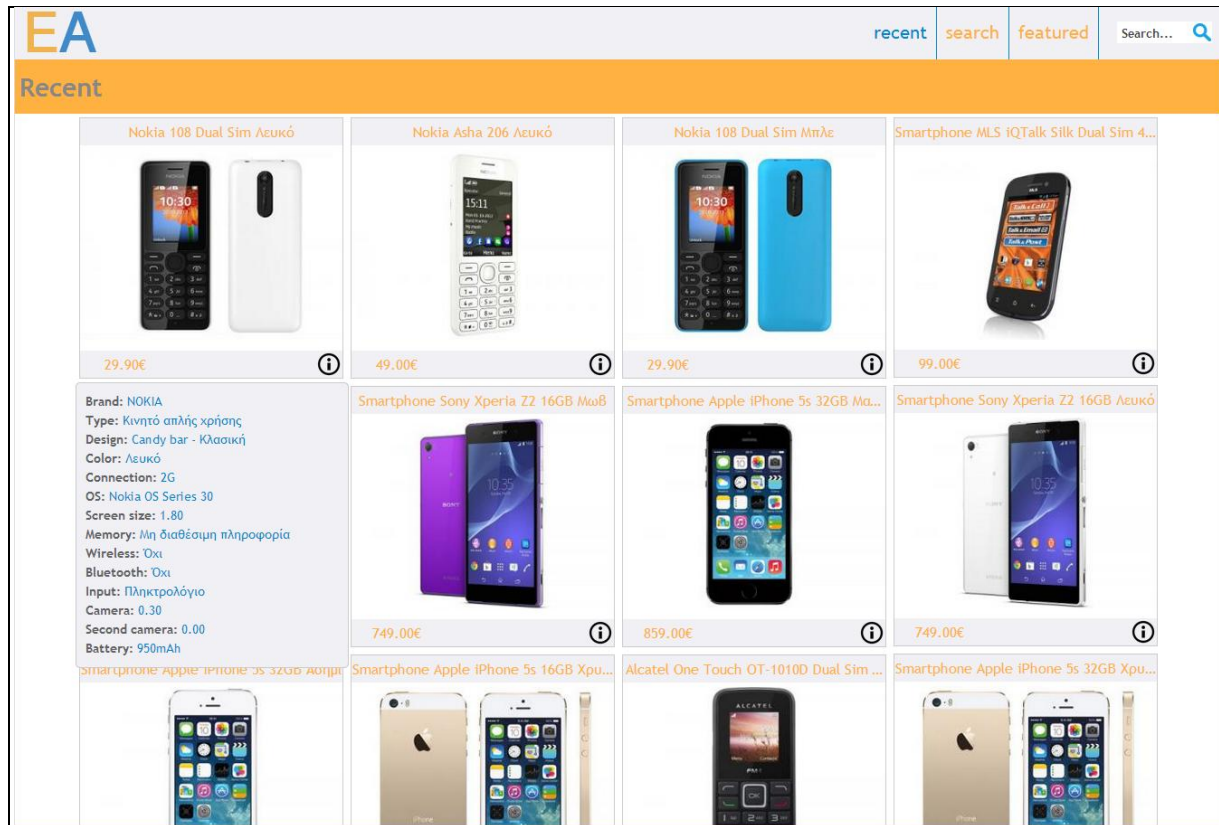
### 7.3 Παρουσίαση χρήσης της Εφαρμογής

Στην ενότητα αυτή γίνεται παρουσίαση της web εφαρμογής, την οποία έχει στην διάθεση του ο χρήστης και αποτελείται από τρεις βασικές καρτέλες:

- Τη καρτέλα **Recent**, η οποία αποτελεί την αρχική σελίδα της εφαρμογής και παρουσιάζει τα τελευταία προϊόντα που προστέθηκαν σε αυτή. Ο χρήστης έχει την δυνατότητα κάνοντας “mouse over” στην εικόνα του εκάστοτε προϊόντος να δει επιγραμματικά πληροφορίες για αυτό, ενώ επιλέγοντας την εικόνα μπορεί να μεταφερθεί αυτόματα στην σελίδα του αντίστοιχου προμηθευτή. Όπως προαναφέραμε, στα πλαίσια της πλήρους προσαρμογής της εφαρμογής ως προς την συσκευή χρήσης του χρήστη, έχει δημιουργηθεί κουμπί πληροφοριών για χρήστες smartphones και tablets, όπου δεν υπάρχει δυνατότητα “mouse over”.



Εικόνα 14: Παρουσίαση της καρτέλας "Recent"



Εικόνα 15: Δυνατότητα "mouse over"

- Τη καρτέλα **Search**, η οποία δίνει την δυνατότητα στον χρήστη να αναζητήσει προϊόντα με την χρήση συγκεκριμένων λέξεων στην περιγραφή τους και να τα ταξινομήσει με βάση την τιμή ή την επισκεψιμότητα, καθώς και να επιλέξει αν η ταξινόμηση θα είναι αύξουσα ή φθίνουσα. Επιπλέον, ο χρήστης έχει την δυνατότητα μέσα από την ίδια σελίδα και επιλέγοντας **More filters** να κάνει μια εκτενέστερη αναζήτηση χρησιμοποιώντας πολλαπλά φίλτρα, τα οποία είναι διαθέσιμα με την μορφή "facets". Τέλος, αξίζει να σημειωθεί ότι τα φίλτρα είναι δυναμικά, κάτι που σημαίνει ότι όσο ο χρήστης επιλέγει φίλτρα, οι επιλογές της αναζήτησης μειώνονται σε πραγματικό χρόνο.

The screenshot displays the 'Search' page of the EA website. At the top, there is a navigation bar with the EA logo, links for 'recent', 'search', and 'featured', and a search input field. Below this is an orange 'Search' header with a 'Home' link. A filter section includes 'More filters...', a 'keyword...' input, and sorting options: 'Sort by Price' and 'Order Asc'. An 'Apply' button is present. The main content area shows a grid of mobile phone listings:

Product Name	Price
Alcatel One Touch OT-1010D Dual Sim ...	17.90€
Samsung E1200 Λευκό	19.90€
Samsung E1200 Μαύρο	19.90€
Nokia 105 Μαύρο	19.99€
Nokia 106 Λευκό	23.90€
Nokia 106 Μαύρο	23.90€
Nokia 108 Λευκό	
Alcatel One Touch OT-1030D Dual Sim ...	
Nokia 108 Dual Sim Μαύρο	

Εικόνα 16: Παρουσίαση της καρτέλας "Search"

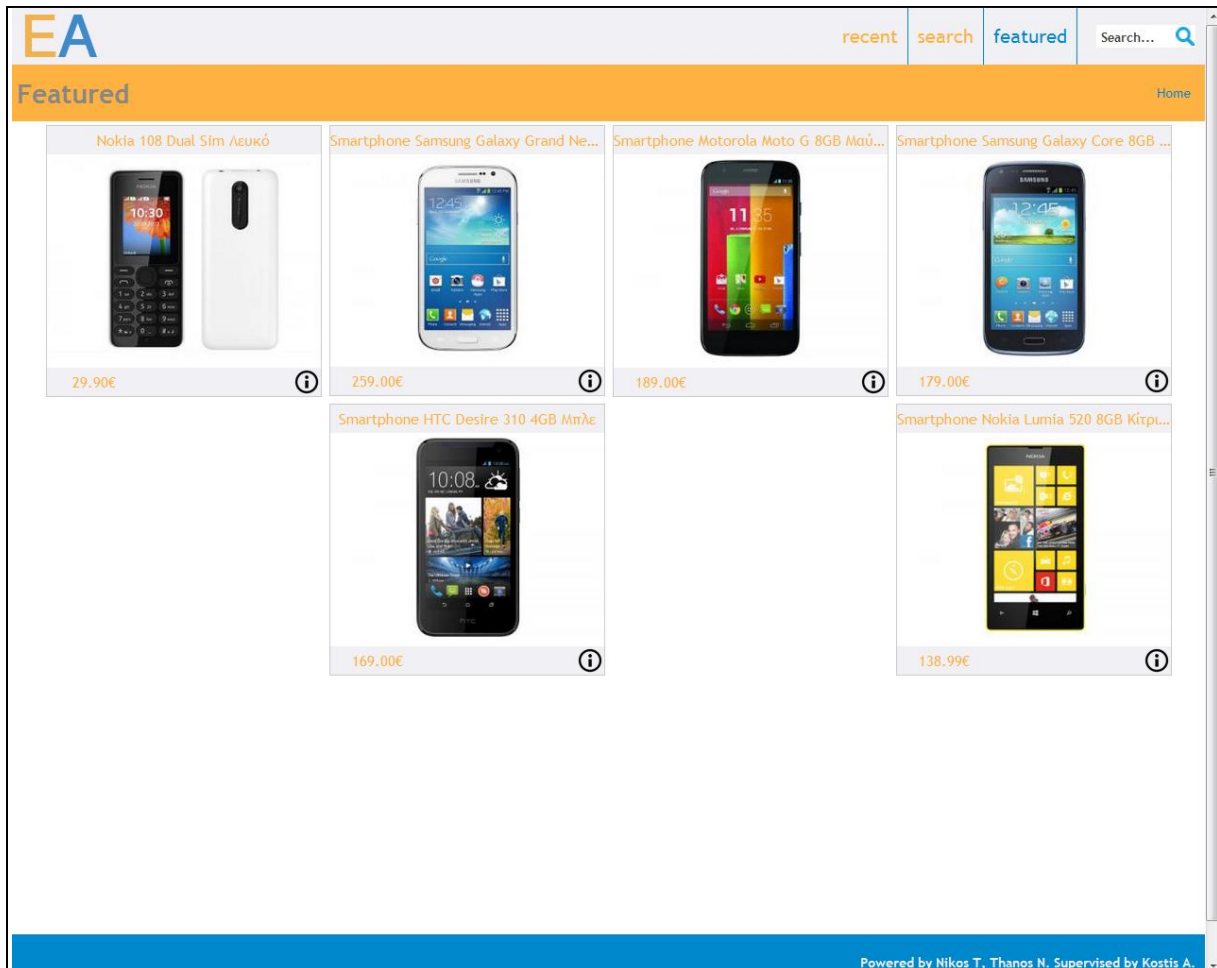


The screenshot displays a search interface for mobile phones. At the top, there is a navigation bar with the 'EA' logo, and buttons for 'recent', 'search', and 'featured'. A search bar is located on the right side of the top bar. Below the navigation bar is a 'Search' header with a 'Home' link. A 'More filters...' dropdown is visible. The main content area is divided into several filter categories:

- Price:** Price ranges from 17 € to 979 €. Includes 'From' (17) and 'To' (979) input fields, a slider, and a 'Go' button.
- Brand:**
  - NOKIA (43)
  - SAMSUNG (27)
  - SONY (25)
  - ALCATEL (19)
  - LG (11)
  - APPLE (10)
  - HUAWEI (4)
  - MLS (4)
  - HTC (3)
  - ZTE (3)
- OS:**
  - Google Android OS 4.2 (21)
  - Λειτουργικό κατασκευαστή (17)
  - Google Android OS 4.1.2 (14)
  - Google Android OS 4.1 (13)
  - Nokia OS Series 40 (11)
  - Google Android OS 4.3 (10)
  - Apple iOS 7 (9)
  - Google Android 4.4.2 (9)
  - Microsoft Windows Phone OS 8 (7)
  - Google Android OS 4.2.2 (6)
- Screen:**
  - 4.0 (45)
  - 5.0 (19)
  - 1.8 (14)
  - 3.5 (14)
  - 4.5 (14)
  - 2.4 (12)
  - 4.8 (6)
  - 4.3 (5)
  - 5.2 (4)
  - 2.0 (3)
- Design:**
  - Candy bar - Touch Screen (117)
  - Candy bar - Κλασική (28)
  - Folder type (αναδιπλούμενη) (4)
  - Slider (συρόμενη πρόσοψη) (2)
  - Candy bar (κλασική) (1)
- Color:**
  - Μαύρο (59)
  - Λευκό (48)
  - Μπλε (7)
  - Ροζ (6)
  - Κίτρινο (5)
  - Ασημί (4)
  - Πορτοκαλί (4)
  - Γκρι (3)
  - Κόκκινο (3)
  - Μωβ (3)
- Input:**
  - Οθόνη αφής (115)
- Memory:**
  - 4 GB (53)
- Bluetooth:**
  - Bluetooth 4.0 + A2DP (43)

Εικόνα 17: Παρουσίαση των πρόσθετων φίλτρων

- Η τελευταία καρτέλα είναι η **Featured**, στην οποία παρουσιάζονται προϊόντα, τα οποία έχουν προεπιλεγεί από τον κάτοχο της εφαρμογής.



Εικόνα 18: Παρουσίαση της καρτέλας "Featured"

## 7.4 Παρουσίαση περιβάλλοντος διαχείρισης της εφαρμογής

Στην παρούσα ενότητα γίνεται παρουσίαση του περιβάλλοντος διαχείρισης της web εφαρμογής, καθώς και των αντίστοιχων εργαλείων που έχει στην διάθεση του ο κάτοχος ή διαχειριστής, με σκοπό την γρήγορη και αποτελεσματική διαμόρφωση, ενημέρωση και τροποποίηση της τελικής εφαρμογής.

### 7.4.1 Παρουσίαση του menu "Content"

Αρχικά, στο menu Content, όπως παρουσιάζεται και στην παρακάτω εικόνα, ο διαχειριστής έχει τη δυνατότητα επεξεργασίας των προϊόντων. Συγκεκριμένα, αποκτά εύκολη και γρήγορη

πρόσβαση σε προϊόντα, τα οποία εισήχθησαν τελευταία και επισημαίνονται με την λέξη “new” δίπλα από την περιγραφή, μπορεί να ενημερωθεί για την ημερομηνία και ώρα που εισήχθησαν, καθώς και για το αν η εισαγωγή τους έχει γίνει αυτοματοποιημένα ή από κάποιον χρήστη της εφαρμογής. Τέλος, μπορεί να κάνει αλλαγή στα στοιχεία ή στην εικόνα ενός προϊόντος, να επιλέξει αν θα εμφανίζεται ή όχι, ακόμη και να το διαγράψει τελείως.

The screenshot displays the 'Content' management interface. At the top, there are navigation tabs: Content, Structure, Appearance, People, Modules, Configuration, and Reports. The user is logged in as 'admin'. Below the navigation, there are filter options for 'Status' and 'Type', both set to 'any'. There is also an 'Update options' section with a 'Publish selected content' dropdown and an 'Update' button. The main content area is a table with the following data:

<input type="checkbox"/>	Title	Type	Author	Status	Updated	Operations
<input type="checkbox"/>	Nokia 108 Dual Sim Λευκό <b>new</b>	Mobile phone	Anonymous	published	08/24/2014 - 20:23	<a href="#">edit</a> <a href="#">delete</a>
<input type="checkbox"/>	Smartphone Apple iPhone 5c 16GB Λευκό <b>new</b>	Mobile phone	Anonymous	published	08/24/2014 - 20:21	<a href="#">edit</a> <a href="#">delete</a>
<input type="checkbox"/>	Smartphone Apple iPhone 5c 16GB Πράσινο	Mobile phone	Anonymous	published	06/15/2014 - 20:05	<a href="#">edit</a> <a href="#">delete</a>
<input type="checkbox"/>	Smartphone Apple iPhone 5c 16GB Κίτρινο	Mobile phone	Anonymous	published	06/15/2014 - 20:05	<a href="#">edit</a> <a href="#">delete</a>
<input type="checkbox"/>	Smartphone Apple iPhone 5c 16GB Μπλε	Mobile phone	Anonymous	published	06/15/2014 - 20:05	<a href="#">edit</a> <a href="#">delete</a>
<input type="checkbox"/>	Smartphone Apple iPhone 5c 16GB Ροζ	Mobile phone	Anonymous	published	06/15/2014 - 20:05	<a href="#">edit</a> <a href="#">delete</a>
<input type="checkbox"/>	FI HTML Plaisio.gr	FI HTML Plaisio.gr	admin	not published	06/15/2014 - 17:04	<a href="#">edit</a> <a href="#">delete</a>
<input type="checkbox"/>	CC Plaisio.gr	Crawler Configuration	admin	not published	06/15/2014 - 17:04	<a href="#">edit</a> <a href="#">delete</a>
<input type="checkbox"/>	CC Public.gr	Crawler Configuration	admin	not published	06/15/2014 - 16:59	<a href="#">edit</a> <a href="#">delete</a>
<input type="checkbox"/>	FI HTML Public.gr Mobile Phones	FI HTML Public.gr	admin	not published	06/15/2014 - 16:59	<a href="#">edit</a> <a href="#">delete</a>
<input type="checkbox"/>	Nokia Asha 206 Λευκό	Mobile phone	Anonymous	published	06/14/2014 - 15:05	<a href="#">edit</a> <a href="#">delete</a>
<input type="checkbox"/>	Nokia 108 Dual Sim Μπλε	Mobile phone	Anonymous	published	06/14/2014 - 12:40	<a href="#">edit</a> <a href="#">delete</a>
<input type="checkbox"/>	Nokia 112 Dual Sim Γκρι	Mobile phone	Anonymous	published	06/14/2014 - 12:35	<a href="#">edit</a> <a href="#">delete</a>
<input type="checkbox"/>	Smartphone Samsung Galaxy Note 3 32GB Λευκό	Mobile phone	Anonymous	published	06/13/2014 - 22:12	<a href="#">edit</a> <a href="#">delete</a>
<input type="checkbox"/>	Smartphone Samsung Galaxy Note 3 32GB Μαύρο	Mobile phone	Anonymous	published	06/13/2014 - 22:12	<a href="#">edit</a> <a href="#">delete</a>
<input type="checkbox"/>	Smartphone Apple iPhone 5 64GB Λευκό	Mobile phone	Anonymous	published	06/13/2014 - 22:06	<a href="#">edit</a> <a href="#">delete</a>

Εικόνα 19: Παρουσίαση του menu "Content"

#### 7.4.2 Εισαγωγή νέου καταστήματος

Βασικό κομμάτι της παρούσας εργασίας αποτελεί ο τρόπος με τον οποίο η εφαρμογή εισάγει δεδομένα από διάφορα ηλεκτρονικά καταστήματα και ενημερώνεται αυτόματα για τις οποίες αλλαγές γίνονται σε αυτά. Στην συνέχεια ακολουθεί εκτενής παρουσίαση των σταδίων που ακολουθήσαμε για την επίτευξη αυτού του σκοπού. Η ίδια διαδικασία θα πρέπει να ακολουθείται και από το διαχειριστή της εφαρμογής, προκειμένου να εισάγονται προϊόντα από ένα νέο ηλεκτρονικό κατάστημα.

Αρχικά θα πρέπει να δημιουργηθεί ένα Content type. Από το menu Structure → Content Types επιλέγουμε Add Content type.

Content types

Individual content types can have different fields, behaviors, and permissions assigned to them.

**Name \***  
 FI HTML E-Shop.gr Machine name: fi\_html\_e\_shop\_gr [Edit]  
 The human-readable name of this content type. This text will be displayed as part of the list on the *Add new content* page. It is recommended that this name begin with a

**Description**

Describe this content type. The text will be displayed on the *Add new content* page.

**Submission form settings**  
 Title

**Publishing options**  
 Not published

**Display settings**  
 Don't display post information

**Compare revisions**

**Menu settings**

**Available menus**

- Development
- Main menu
- Management
- Navigation
- User menu

The menus available to place links in for this content type.

**Default parent item**

Choose the menu item to be the default parent for a new link in the content authoring form.

**Save content type** **Save and add fields**

Εικόνα 20: Προσθήκη Content type

Στην συνέχεια, πρέπει να δημιουργηθεί ένα Feed importer. Από το menu Structure → Feed Importer επιλέγουμε Add Importer. Δίνουμε ένα όνομα, το οποίο είναι καλό για λόγους εύκολης διαχείρισης να παραπέμπει στο ηλεκτρονικό κατάστημα από το οποίο θέλουμε να αντλήσουμε τα προϊόντα, καθώς και μια σύντομη περιγραφή, αν αυτό κρίνεται απαραίτητο. Έχοντας δημιουργήσει το Feed importer προχωράμε στη σύνδεσή του με το αντίστοιχο Content type, το οποίο έχουμε ήδη έτοιμο.

Content Structure Appearance People Modules Configuration Reports

Home » Administration » Structure » Feeds importers

## Feeds importers

**Name \***

FI HTML E-Shop.gr Machine name: fi\_html\_e\_shop\_gr [Edit]

A natural name for this configuration. Example: RSS Feed. You can always change this name later.

**Description**

A description of this configuration.

**Create**

Εικόνα 21: Δημιουργία Feed importer

Content Structure Appearance People Modules Configuration Reports

Home » Administration » Structure » Feeds importers

## FI HTML E-Shop.gr

EDIT EXPORT CLONE DELETE TAMPER

**Basic settings**

Attached to: [none] [Settings](#)

Periodic import: every 30 min

Import on submission

**Fetcher** [Change](#)

HTTP Fetcher [Settings](#)

Download content from a URL.

**Parser** [Change](#)

Common syndication parser [Settings](#)

Parse RSS and Atom feeds.

**Processor** [Change](#)

Node processor [Settings](#) [Mapping](#)

Create and update nodes.

**Basic settings**

**Name \***

FI HTML E-Shop.gr

A human readable name of this importer.

**Description**

A description of this importer.

Use standalone form

Crawler Configuration

✓ FI HTML E-Shop.gr

FI HTML Plaisio.gr

FI HTML Public.gr

Mobile phone

Every 30 min

Choose how often a source should be imported periodically. [Requires cron to be configured.](#)

Import on submission

Check if import should be started at the moment a standalone form or node form is submitted.

Process in background

For very large imports. If checked, import and delete tasks started from the web UI will be handled by a cron task in the back to be configured.

**Save**

Εικόνα 22: Σύνδεση Feed importer με το αντίστοιχο Content type

Το επόμενο βήμα είναι να δηλώσουμε στον αντίστοιχο Feed importer, που δημιουργούμε με ποιό τρόπο θέλουμε να γίνει η πρώτη εισαγωγή των προϊόντων. Στην παρούσα εργασία έχουμε επιλέξει να γίνεται από πολλαπλές διευθύνσεις διαδικτύου “URLs”. Γι’ αυτόν τον λόγο επιλεγούμε από την καρτέλα Fetcher, το Crawling fetcher και κάνοντας αποθήκευση μεταφερόμαστε στις πρόσθετες ρυθμίσεις του, όπου θα πρέπει να δηλώσουμε το πού και πώς θα βρει αυτήν την πληροφορία για να κάνει την πρώτη εισαγωγή.

Αρχικά δηλώνουμε τον αριθμό των σελίδων από τις οποίες θέλουμε να αντλήσουμε τα προϊόντα και την καθυστέρηση που θα έχει η εφαρμογή από σελίδα σε σελίδα. Στην περίπτωση μας έχουμε επιλέξει 40 σελίδες για να διασφαλίσουμε ότι θα εισαχθούν όλα τα προϊόντα και καθυστέρηση 10 seconds. Στην πραγματικότητα, ο χρόνος ανά σελίδα που χρειάζεται είναι πολύ μικρότερος, παρόλα αυτά επιλέγουμε το χρόνο των 10 seconds λαμβάνοντας υπόψη και την παράμετρο η διαδικασία να μπλοκαριστεί από κάποιο τείχος προστασίας “firewall”.

Στην συνέχεια δηλώνουμε το XPath για την επόμενη σελίδα, τον τρόπο δηλαδή με τον οποίο η εφαρμογή θα βρει τον επόμενο κόμβο μέσα στην XML, ο οποίος θα αποτελεί και την επόμενη σελίδα προϊόντων.

Έπειτα δηλώνουμε το μοτίβο σελιδοποίησης που ακολουθείται από το αντίστοιχο ηλεκτρονικό κατάστημα στις σελίδες των προϊόντων ξεχωρίζοντας τον δείκτη με βάση το βήμα που ακολουθεί και αντικαθιστώντας τον με \$index. Τέλος δηλώνουμε την αρχική τιμή του δείκτη και το βήμα ακολουθίας του.



Εικόνα 23: Επιλογή του τρόπου εισαγωγής των προϊόντων

**Number of pages**

The number of pages to fetch. 0 for unlimited

**Delay**

Number of seconds to delay in between fetches.

**Crawl on first run only**  
Only crawl on initial run. Use regular import afterward.

**Auto detect next link**  
Attempt to autodetect the next link for RSS and ATOM feeds.

**XPath selector for next link**

The XPath selector for the next link.

**URL replacement options**

**URL pattern**

A URL with the variable \$index replaced with an incrementing number. For example: http://example.com?page=\$index.

**Initial value of \$index**

The initial value of the \$index variable.

**Increment \$index by**

The increment the value of \$index variable.

Εικόνα 24: Πρόσθετες ρυθμίσεις για την εισαγωγή προϊόντων

Το επόμενο βήμα είναι να επιλέξουμε με ποιό τρόπο θα αναλυθούν τα δεδομένα που εισάγονται στην εφαρμογή και με ποιό τρόπο θα αποθηκεύονται και θα επεξεργάζονται. Επιλέγουμε από την καρτέλα Parser → HTML Parser, δηλαδή ότι η ανάλυση τους θα γίνεται από HTML και από την καρτέλα Processor → Node Processor, που σημαίνει ότι τα δεδομένα θα αποθηκεύονται και θα επεξεργάζονται σαν κόμβοι.



Select a parser		Help
<b>Common syndication parser</b> Parse XML feeds in RSS 1, RSS 2 and Atom format.	<input type="radio"/>	Select
<b>CSV parser</b> Parse data in Comma Separated Value format.	<input type="radio"/>	Select
<b>OPML parser</b> Parse OPML files.	<input type="radio"/>	Select
<b>Sitemap parser</b> Parse Sitemap XML format feeds.	<input type="radio"/>	Select
<b>XPath HTML parser</b> Parse HTML using XPath.	<input checked="" type="radio"/>	Select
<b>XPath XML parser</b> Parse XML using XPath.	<input type="radio"/>	Select

Εικόνα 25: Επιλογή τρόπου ανάλυσης δεδομένων



Select a processor		Help
<b>Node processor</b> Create and update nodes from parsed content.	<input checked="" type="radio"/>	Select
<b>Taxonomy term processor</b> Create taxonomy terms from parsed content.	<input type="radio"/>	Select
<b>User processor</b> Create users from parsed content.	<input type="radio"/>	Select

Εικόνα 26: Επιλογή τρόπου επεξεργασίας των δεδομένων

Μετά την αποθήκευση του τρόπου επεξεργασίας των δεδομένων κάνουμε κάποιες επιπλέον ρυθμίσεις, που σχετίζονται με την κατηγορία στην οποία εντάσσονται τα προϊόντα και με βάση την οποία συμπεριφέρονται. Επιλέγουμε ότι κάθε κόμβος που θα βρίσκεται κοινός θα αντικαθίσταται. Με τον τρόπο αυτό διασφαλίζουμε την αποφυγή αλλοίωσης των δεδομένων από οποιονδήποτε παράγοντα. Επιπλέον, επιλέγουμε την μορφή με την οποία θα έρθει η πληροφορία, καθώς και πότε θα λήγει ένα προϊόν. Στην συγκεκριμένη περίπτωση έχουμε θέσει το χρόνο αυτό ένα μήνα. Ο χρόνος δύναται να καθοριστεί με βάση την κατηγορία του προϊόντος, σε πόσο καιρό δηλαδή καταργείται ένα μοντέλο από την αγορά.



### Settings for Node processor

**Bundle \***  
 Mobile phone

**Update existing nodes**  
 Do not update existing nodes  
 Replace existing nodes  
 Update existing nodes  
 Existing nodes will be determined using mappings that are a "unique target".

Skip hash check  
 Force update of items even if item source data did not change.

**Text format \***  
 Plain text  
 Select the input format for the body field of the nodes to be created.

**Author**  
 anonymous  
 Select the author of the nodes to be created - leave empty to assign "anonymous".

Authorize  
 Check that the author has permission to create the node.

**Expire nodes**  
 after 1 month  
 Select after how much time nodes should be deleted. The node's published date will be used for determining the node's age, see Mapping settings.

Εικόνα 27: Ρυθμίσεις για την επεξεργασία των δεδομένων

Το επόμενο βήμα είναι να ρυθμίσουμε ποιές θα είναι οι πρώτες πληροφορίες που θα ανιχνεύονται με το XPath και στην περίπτωση μας οι πληροφορίες αυτές είναι:

- ο τίτλος του προϊόντος
- η φωτογραφία του προϊόντος
- το GUID του προϊόντος
- το site
- και αν είναι δημοσιευμένο.

### Mapping for Node processor

Define which elements of a single item of a feed (= Sources) map to which content pieces in Drupal (= Targets). Make sure that at least one definition has a *Unique target*. A unique target means that a value for a target can only occur once. E. g. only one item with the URL `http://example.com/content/1` can exist.

Source	Target	Target configuration	
+ xpathparser:0	Title	Not used as unique.	<input type="checkbox"/> Remove
+ xpathparser:1	Photo		<input type="checkbox"/> Remove
+ xpathparser:2	Link		<input type="checkbox"/> Remove
+ xpathparser:3	GUID	Not used as unique.	<input type="checkbox"/> Remove
+ xpathparser:4	Site		<input type="checkbox"/> Remove
+ xpathparser:5	Published status		<input type="checkbox"/> Remove

- Select a source -  
 An element from the feed.

- Select a target -  
 The field that stores the data.

Εικόνα 28: Ρυθμίσεις αρχικών πληροφοριών

Στη συνέχεια θα πρέπει να συμπληρώσουμε τα αντίστοιχα XPath, δηλαδή τις αντίστοιχες διαδρομές από όπου η εφαρμογή θα αντλήσει τα δεδομένα.

▼ XPath Parser Settings

**Context**

This is the base query, all other queries will run in this context.

**title**

The XPath query to run.

**field\_photo**

The XPath query to run.  
The variable *\$title* is available for replacement.

**field\_link**

The XPath query to run.  
The variables *\$title*, *\$field\_photo* are available for replacement.

**guid**

The XPath query to run.  
The variables *\$title*, *\$field\_photo*, *\$field\_link* are available for replacement.

**field\_site**

The XPath query to run.  
The variables *\$title*, *\$field\_photo*, *\$field\_link*, *\$guid* are available for replacement.

**status**

The XPath query to run.  
The variables *\$title*, *\$field\_photo*, *\$field\_link*, *\$guid*, *\$field\_site* are available for replacement.

Εικόνα 29: Συμπλήρωση των XPath

Τέλος, θα πρέπει να δηλώσουμε αντίστοιχες διαδικασίες που θα πρέπει να συμβούν στα δεδομένα πριν από την αποθήκευσή τους. Αυτό γίνεται από την καρτέλα **Tamper**, όπως φαίνεται και στην εικόνα που ακολουθεί:

Configure plugins to modify Feeds data before it gets saved. Each **mapping** can be manipulated individually.

**xpathparser:0 -> Title**

Description	Plugin	Status	Operations
+ Trim	Trim	Normal	<a href="#">Edit</a>   <a href="#">Delete</a>
<a href="#">+ Add plugin</a>			

**xpathparser:1 -> Photo**

Description	Plugin	Status	Operations
No plugins defined.			
<a href="#">+ Add plugin</a>			

**xpathparser:2 -> Link**

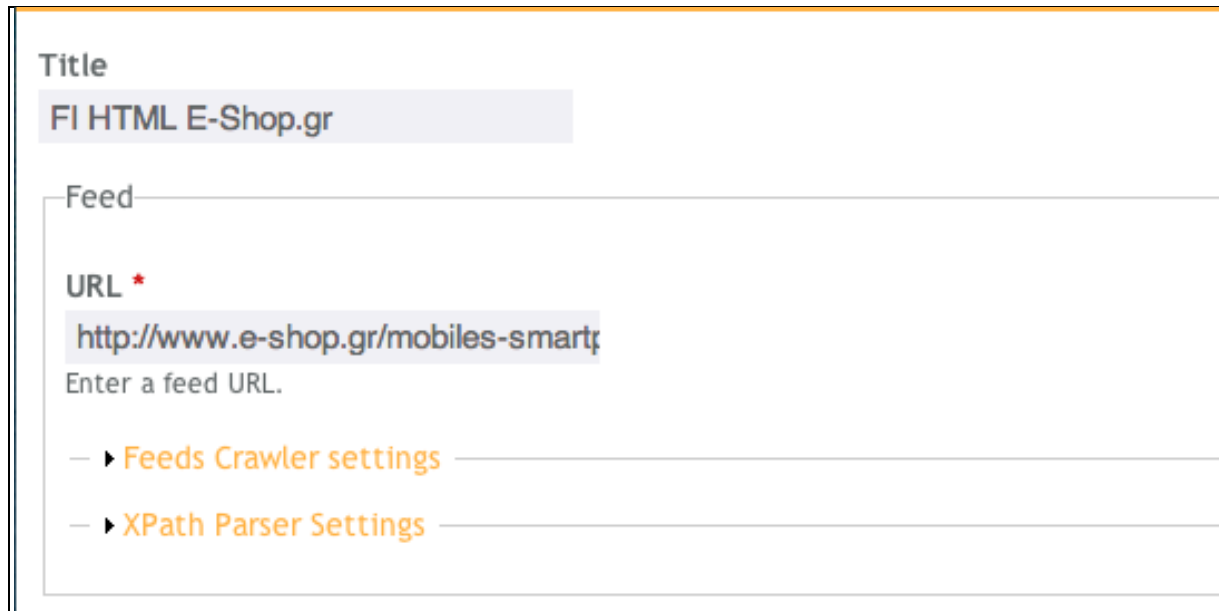
Description	Plugin	Status	Operations
+ Rewrite	Rewrite	Normal	<a href="#">Edit</a>   <a href="#">Delete</a>
<a href="#">+ Add plugin</a>			

**xpathparser:3 -> GUID**

Description	Plugin	Status	Operations
+ Rewrite	Rewrite	Normal	<a href="#">Edit</a>   <a href="#">Delete</a>
<a href="#">+ Add plugin</a>			

**Εικόνα 30: Μορφοποίηση δεδομένων**

Το τελευταίο βήμα για την ολοκλήρωση εισαγωγής ενός νέου ηλεκτρονικού καταστήματος είναι να δημιουργήσουμε ένα νέο Content, βασισμένο στο Feed Importer που χρησιμοποιήσαμε. Στην συγκεκριμένη περίπτωση, όπως φαίνεται και στην παρακάτω εικόνα, αφορά το ηλεκτρονικό κατάστημα e-shop.gr, ενώ στη συνέχεια προχωράμε στη δημιουργία του αντίστοιχου Crawling configuration.



The image shows a web form for creating content. It has a title field with the text 'FI HTML E-Shop.gr'. Below it is a 'Feed' label. The 'URL' field is marked with a red asterisk and contains the text 'http://www.e-shop.gr/mobiles-smart'. Below the URL field is the instruction 'Enter a feed URL.' and two expandable sections: 'Feeds Crawler settings' and 'XPath Parser Settings', both with orange arrow icons.

Εικόνα 31: Δημιουργία Content

Το Crawling configuration είναι το σημείο στο οποίο δηλώνουμε όλα τα RegEx και XPath, για να εισάγουμε τις επιμέρους πληροφορίες για τα προϊόντα που χρειαζόμαστε, όπως για παράδειγμα φωτογραφία, κατασκευαστή, χρώμα κλπ.

**Title \***

**Site**  
  
 Site name, as written in the content "job" to make the connection

**Active**  
 Click if the crawler configuration is active. If this remains unchecked, the job content will remain as they are

**Photo**

**Photo regex**

**1**

**Brand**

**Brand regex**

**Type**

**Type regex**

**Design**

**Design regex**

**Color**

**Color regex**

Εικόνα 32: Crawling configuration



## Κεφάλαιο Όγδοο: Αποτελέσματα

### 8.1 Συμπεράσματα

Στα πλαίσια της παρούσας εργασίας δημιουργήσαμε μια ιδιαίτερα καινοτόμα και χρηστική web εφαρμογή. Βασικός στόχος της εφαρμογής είναι η συγκέντρωση προϊόντων από διάφορα ηλεκτρονικά καταστήματα και η κατηγοριοποίηση τους με τέτοιο τρόπο, ώστε η αναζήτηση από το χρήστη να καθίσταται ιδιαίτερα εύκολη, παρέχοντας επιπρόσθετα και τη δυνατότητα αναζήτησης μέσω κριτηρίων.

Αναπτύξαμε μια εφαρμογή φιλική προς το χρήστη, με δυνατότητα άμεσης προσαρμογής στα χαρακτηριστικά οποιαδήποτε συσκευής από την οποία χρησιμοποιείται για την περιήγηση σε αυτή. Μια εφαρμογή που παρέχει στο χρήστη τη δυνατότητα παρακολούθησης των αλλαγών στα διάφορα προϊόντα σε πραγματικό χρόνο.

Κατά την ενασχόληση μας αντιληφθήκαμε την τεράστια χρησιμότητα του ελεύθερου λογισμικού και τις δυνατότητες που αυτό μπορεί να προσφέρει κατά τη δημιουργία οποιασδήποτε εφαρμογής. Το σύνολο της εφαρμογής στηρίχθηκε σε 100% ελεύθερο λογισμικό.

Επιπλέον, κατέστη σαφές το πλήθος των μονάδων που χρειάζεται μια εφαρμογή για να καταστεί λειτουργική, καθώς επίσης και η αξία της επεκτασιμότητας της. Η δυνατότητα μελλοντικής επέκτασης μιας εφαρμογής αποτελεί έναν από τους σημαντικότερους στόχους.

Τέλος, αξίζει να αναφερθεί και η τεράστια σημασία της ταχύτητας μίας εφαρμογής, όσον αφορά το πόσο εύχρηστη είναι και κατά συνέπεια το πόσο αποδεκτή θα γίνει από το ευρύ κοινό.

### 8.2 Μελλοντική Εργασία και Επεκτάσεις

Για την εξασφάλιση της δυνατότητας μελλοντικής επέκτασης της εφαρμογής έχει υλοποιηθεί σε Virtual machine. Στο πλαίσιο ενίσχυσης των δυνατοτήτων που προσφέρει στο χρήστη η εφαρμογή, θα μπορούσε να δημιουργηθεί η υποδομή για την δημιουργία forum, όπου οι χρήστες θα μπορούν να ανεβάζουν σχόλια σχετικά με την εμπειρία τους από κάποιο προϊόν ή από το κατάστημα που το έχουν προμηθευτεί.

Μια άλλη δυνατότητα επέκτασης της εφαρμογής θα μπορούσε να αφορά στην δυνατότητα πρόσβασης του χρήστη στις αποθηκευμένες αναζητήσεις του, μετά τη σύνδεση του στην εφαρμογή. Με αυτό τον τρόπο, ο χρήστης θα μπορεί, έχοντας δημιουργήσει το προσωπικό του προφίλ, να ενημερώνεται, εάν το επιθυμεί, μέσω email, για παράδειγμα, σχετικά με τις εισαγωγές νέων προϊόντων που τον αφορούν.

Τέλος, δεδομένου του ότι πρόκειται για μια scalable εφαρμογή, μια πιθανή επέκτασή της θα μπορούσε να αφορά στην κατανομή της σε περισσότερους server. Με αυτόν τον τρόπο εξασφαλίζεται η παροχή υψηλότερου επιπέδου ασφάλειας, η εξυπηρέτηση μεγαλύτερου αριθμού χρηστών, καθώς και η διαχείριση μεγαλύτερου όγκου δεδομένων και πληροφοριών.





## Βιβλιογραφία

1. Bloch Joshua, (2008), “Effective Java”, Addison-Wesley Publications.
2. Horstmann Cay & Gary Cornell, (2000), “Core Java 2: Fundamentals”, Sun Microsystems, Prentice Hall PTR, vol. 1.
3. Lerdorf Rasmus, Tatroe Kevin & Peter Maclynre, (2006), “Programming PHP”, O’Reilly Media, Inc, Sebastopol, USA
4. Naramore Elizabeth, Gerner Jason, Le Scouarnec Yann, Stolz Jeremy and Michael K. Glass, (2005), “Beginning PHP5, Apache, and MySQL Web Development”, Wiley Publishing , Inc, Indianapolis, Canada
5. Ray T. Erik, (2003), “Learning XML, 2nd Edition”, O'Reilly Media.
6. Salvador Otavio & Daiane Angolini, (2014), “Embedded Linux Development with Yocto Project”, Packt Publishing
7. Shreves, R. and B. Dunwoodie, (2010), Drupal 7 Bible, Wiley Publishing, Inc., Indianapolis, USA.
8. The Apache Software Foundation, (2010), “Apache HTTP Server 2.2 Official Documentation”, Volume 1, Fultus Cooperation, USA
9. Tomlinson Todd, (2010) Beginning Drupal 7, Apress, distributed by Springer Science & Business Media, LLC, New York, USA.
10. Travis Brian, (2011), Pro Drupal 7 for Windows Developers, Apress, distributed by Springer Science & Business Media, LLC, New York, USA.
11. Welling Luke & Laura Thomson, (2003), “PHP and MySQL Web Development”, Sams Publishing, USA.

## Ηλεκτρονικές Πηγές

1. <https://drupal.org/documentation>
2. <http://istosonline.com/drupal>
3. <http://www.cmsmatrix.org/>
4. <http://www.it.uom.gr/project/XML/Home%20Page.htm>
5. [http://www0.dmst.aueb.gr/louridas/lectures/dais/web\\_intro/ar01s21.HTML](http://www0.dmst.aueb.gr/louridas/lectures/dais/web_intro/ar01s21.HTML)
6. <http://www.mysql.com/>
7. [http://en.wikipedia.org/wiki/LAMP\\_%28software\\_bundle%29](http://en.wikipedia.org/wiki/LAMP_%28software_bundle%29)
8. <http://en.wikipedia.org/wiki/Linux>
9. [http://www.w3schools.com/XML/XML\\_what.asp](http://www.w3schools.com/XML/XML_what.asp)
10. <http://en.wikipedia.org/wiki/XML>
11. <http://en.wikipedia.org/wiki/RSS>
12. <https://www.RSS.com/>
13. [http://en.wikipedia.org/wiki/Java\\_%28programming\\_language%29](http://en.wikipedia.org/wiki/Java_%28programming_language%29)
14. <http://lucene.apache.org/solr/documentation.HTML>
15. <http://www.eclipse.org/jetty/>
16. [http://en.wikipedia.org/wiki/Jetty\\_%28web\\_server%29](http://en.wikipedia.org/wiki/Jetty_%28web_server%29)
17. [http://en.wikipedia.org/wiki/Responsive\\_web\\_design](http://en.wikipedia.org/wiki/Responsive_web_design)



## Παράρτημα

Στην συνέχεια παρατίθενται οι διαφάνειες της παρουσίασης:

# Διαδικτυακό σύστημα συγκέντρωσης και αναζήτησης προϊόντων ηλεκτρονικών καταστημάτων

Αθανάσιος Νόκας  
Νικόλαος Θεοδωρής

Επιβλέπων Καθηγητής: κ. Κωστής Αϊβαλής

## Στόχοι εργασίας

- Συγκέντρωση περιεχομένου διαφορετικών ιστότοπων
- Κοινή μορφοποίηση δεδομένων
- Ευρετηρίαση δεδομένων
- Παρουσίαση στον χρήστη

## Τεχνολογίες & εργαλεία

- LAMP stack. Ubuntu, Apache, MySQL, PHP
- Java JDK
- Apache SOLR & Jetty
- Drupal
- SVN
- VMware

## Apache SOLR

- full-text αναζήτηση
- NoSQL αποθήκευση και ανάκτηση
- Κλιμάκωση σε πολλούς εξυπηρετητές
- Ανεκτικό σε σφάλματα(Fault tolerant)
- REST διασύνδεση
- Εξαιρετικά γρήγορο!
- Δεν προτείνεται για σύνολα δεδομένων(Group by)

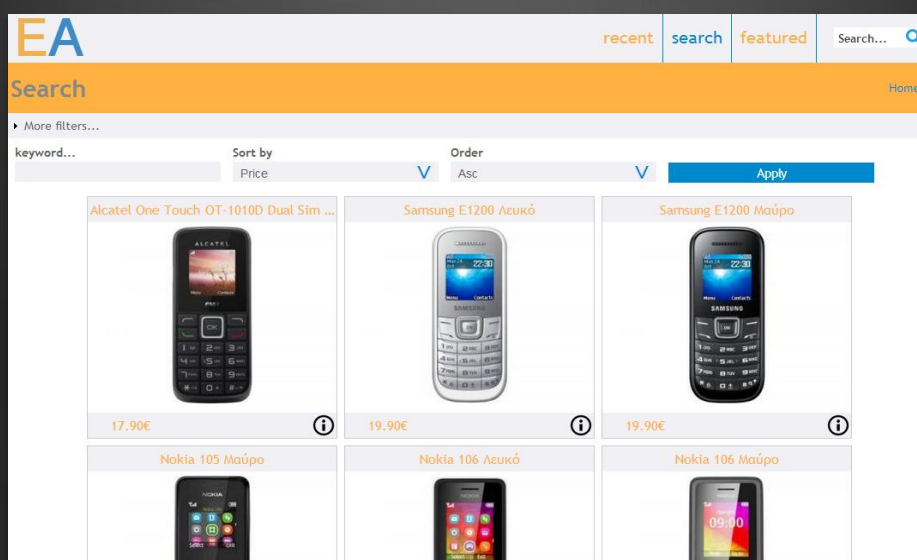
## Drupal

- Ένα από τα καλύτερα CMF για εφαρμογές επιχειρησιακού επιπέδου
- Ισχυρό framework
- Επεκτάσιμο
- Ασφαλές
- 100% Ελεύθερο Λογισμικό

## Μεθοδολογία

- Drupal για διαχείριση δικαιωμάτων, τύπων δεδομένων και διεπαφής
- XPath & RegEx για μορφοποίηση δεδομένων
- SOLR για ευρετηρίαση δεδομένων και αναζήτηση
- SVN για αποθήκευση εκδόσεων του κώδικα
- VMware για virtualization

## Ιστότοπος



## Ροή εργασιών και δεδομένων

- Σάρωση και αποθήκευση περιεχομένων ενός ιστότοπου
- Εύρεση στοιχείων προϊόντων
- Αποθήκευση στη βάση δεδομένων
- Αποθήκευση στον SOLR
- Παρουσίαση στον χρήστη

## Σημαντικές λεπτομέρειες

- Full Responsive ιστότοπος, προσαρμοσμένη διεπαφή
- Έτοιμο προς παραγωγή (Production ready)
- 100% ελεύθερο λογισμικό σε όλο το εύρος της πτυχιακής/εφαρμογής
- Εξαγωγή δεδομένων σε XML
- Προσωρινή αποθήκευση (Caching)

## Βελτιστοποίηση για κινητές συσκευές



## Ταχύτητα αναζήτησης

Αναζήτηση λέξης κλειδιού σε 5.000.000 εγγραφές

MySQL & SOLR

■ Δευτερόλεπτα





## Πιθανές επεκτάσεις

- REST εξυπηρετητής
- Forum ανταλλαγής απόψεων, σχολιασμός
- Αποθηκευμένες αναζητήσεις
- Κατανομή σε περισσότερους εξυπηρετητές

## Συμπεράσματα

- Η επεκτασιμότητα μίας εφαρμογής αποτελεί έναν από τους σημαντικότερους στόχους
- Μία εφαρμογή χρειάζεται πολλές διαφορετικές μονάδες
- Μπορείς να φτιάξεις οτιδήποτε με 100% ελεύθερο λογισμικό
- Η ταχύτητα μίας εφαρμογής είναι πολύ σημαντική στην ευχρηστία της