



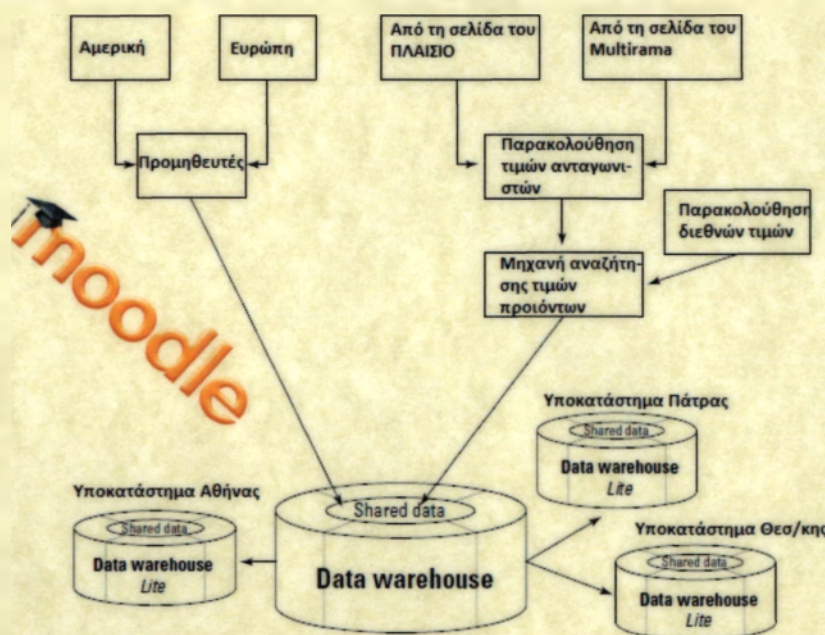
**Α.Τ.Ε.Ι. ΚΑΛΑΜΑΤΑΣ**  
**ΠΑΡΑΡΤΗΜΑ ΣΠΑΡΤΗΣ**  
**Τμήμα Τεχνολογίας Πληροφορικής**  
**και Τηλεπικοινωνιών**

## ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

**Ανάπτυξη εφαρμογής τηλεεκπαίδευσης με το περιβάλλον Moodle για τη διδασκαλία του μαθήματος**

### "Αποθήκες Δεδομένων"

[Στην πτυχιακή αυτή εργασία θα παρουσιαστεί μια μελέτη σχετικά με τα συστήματα ηλεκτρονικής μάθησης και θα χρησιμοποιηθεί το Moodle για τη δημιουργία μιας σειράς μαθημάτων για το μάθημα «Αποθήκες Δεδομένων».]



**Επιμέλεια: Γεώργιος Βάρδας**  
**Αριθμός Μητρώου : 2005023**

**Επιβλέπων καθηγητής: Νικόλαος Κούμαρης**

**Σπάρτη, Οκτώβριος 2010**



*Στους γονείς μου,  
για την αμέριστη συμπαράσταση και κατανόηση*

*Μαθαίνουμε:*

- 10% απ' όσα διαβάζουμε
- 20% απ' όσα ακούμε
- 30% απ' όσα βλέπουμε
- 50% απ' όσα βλέπουμε και ακούμε
- 70% απ' όσα συζητούμε με άλλους
- 80% απ' όσα βιώνουμε προσωπικά
- 95% απ' όσα εμείς διδάσκουμε σε κάποιον άλλον

*William Glasser*

## **Ευχαριστίες**

Ευχαριστώ θερμά τον καθηγητή μου Κούμαρη Νικόλαο, για τις διαρκείς συμβουλές, τις υποδείξεις του και την υπομονή του κατά τη διάρκεια της εκπόνησης της πτυχιακής εργασίας.

Επίσης θα ήθελα να ευχαριστήσω τους γονείς μου, που στάθηκαν δίπλα μου και με στήριξαν καθ' όλη τη διάρκεια της προσπάθειάς μου.

## **Abstract**

In this Computer Science thesis the trends, technologies, platforms and main ideas behind e-Learning and the Science of Instruction will be presented. The importance of e-Learning and the effective usage of it will be evaluated and analyzed. Physiological learning requirements will be studied and presented so that the content of a lecture can efficiently be communicated to the student.

Once the principles of e-Learning are documented, a set of ten chapters will be created regarding Data Warehousing, which is the topic of interest. Content will be positioned to a Moodle website but there will be experimentation with alternative approaches. Extra content to provide hands on experience, courseworks and demo material will be also be created. Lastly another challenge of this thesis is the collaboration between students, professors and members of the academic community.

## Πίνακας περιεχομένων

1. Εισαγωγή .....	1
2. Θεωρητικό υπόβαθρο .....	2
2.1 Μελέτες και έρευνες για την εκπαίδευση προσωπικού .....	2
2.2 Η επιστήμη της μάθησης .....	4
2.3 Οι τύποι των μαθητών.....	5
2.4 Η χρήση επιπλέον – ενδιαφέροντος υλικού.....	6
2.5 Τρόπος επικοινωνίας.....	7
2.6 Η ιστορία του moodle .....	8
2.7 Text-To-Speech.....	10
3. Παρουσίαση μαθημάτων .....	13
3.1 Μάθημα 1 <sup>ο</sup> – Εισαγωγή στο Data Warehousing.....	14
3.2 Μάθημα 2 <sup>ο</sup> – Τεχνολογίες Data Warehousing.....	22
3.3 Μάθημα 3 <sup>ο</sup> – Business Intelligence στα Data Warehouses .....	34
3.4 Μάθημα 4 <sup>ο</sup> – Πρακτική εφαρμογή Business Intelligence και Data Mining.....	46
3.5 Μάθημα 5 <sup>ο</sup> – Προετοιμασία υλοποίησης ενός Data Warehouse .....	52
3.6 Μάθημα 6 <sup>ο</sup> – Υλοποίηση και παράδοση στον πελάτη .....	59
3.7 Μάθημα 7 <sup>ο</sup> – Εξωτερικές πηγές δεδομένων .....	63
3.8 Μάθημα 8 <sup>ο</sup> – Διαχείριση ενός έργου Data Warehouse και συμβουλές.....	71
3.9 Μάθημα 9 <sup>ο</sup> – Λογισμικό για χρήση και συντήρηση Data Warehouses.....	76
3.10 Μάθημα 10 <sup>ο</sup> – Ανακοινώσεις, Επανάληψη, Συζητήσεις, Αξιολόγηση Μαθήματος .....	80

4. Δημιουργία οπτικοακουστικού υλικού.....	86
4.1 Παρουσίαση αρχείων.....	86
4.2 Αρχεία Attribute-Relation File Format.....	87
4.3 Αρχεία Audio.....	90
4.4 Αρχεία AVI.....	92
4.5 Final Compilation.....	93
4.6 Κείμενα.....	99
4.7 PowerPoint.....	100
4.8 Κείμενα Διαλόγων Video Tutorial.....	101
5. Βιβλιογραφία.....	107
APPENDIX.....	109
Questionnair module για ερωτηματολόγια.....	109
Τοποθέτηση Video στο YouTube.....	114

## Πίνακας Εικόνων

Εικόνα 2-1 Προγραμματισμός για την εκπαίδευση επαγγελματιών στην Αγγλία το 2009.....	2
Εικόνα 2-2 Επιθυμητός τρόπος παροχής εκπαίδευσης.....	3
Εικόνα 2-3 Ετήσια μεταβολή ποσοστού χρήσης εκπαιδευτικών μεθοδολογιών.....	4
Εικόνα 2-4 Στοιχεία από μελέτη σχετικά με τον τρόπο επικοινωνίας προς τον εκπαιδευόμενο .....	7
Εικόνα 2-5 Στοιχεία από μελέτη σχετικά με τη χρήση οπτικού και ακουστικού υλικού .....	8
Εικόνα 2-6. Μια χαρακτηριστική σελίδα του Moodle .....	9
Εικόνα 2-7 Το γραφικό περιβάλλον της εφαρμογής TextAloud .....	11
Εικόνα 3-1 Αρχική εισαγωγή στο μάθημα «Αποθήκες Δεδομένων».....	13
Εικόνα 3-2 Παρουσίαση πρώτου μαθήματος .....	14
Εικόνα 3-3 Παρουσίαση 2 <sup>ου</sup> μαθήματος.....	22
Εικόνα 3-4 Παρουσίαση 3 <sup>ου</sup> μαθήματος.....	35
Εικόνα 3-5 Παρουσίαση 4 <sup>ου</sup> μαθήματος.....	46
Εικόνα 3-6 Παρουσίαση 5 <sup>ου</sup> μαθήματος.....	52
Εικόνα 3-7 Παρουσίαση 6 <sup>ου</sup> μαθήματος.....	59
Εικόνα 3-8 Παρουσίαση 7 <sup>ου</sup> μαθήματος.....	63
Εικόνα 3-9 Παρουσίαση 8 <sup>ου</sup> μαθήματος.....	71
Εικόνα 3-10 Παρουσίαση 9 <sup>ου</sup> μαθήματος.....	76
Εικόνα 3-11 Παρουσίαση 10 <sup>ου</sup> μαθήματος.....	81
Εικόνα 3-12 Γενικές ερωτήσεις αξιολόγησης μαθήματος.....	82
Εικόνα 3-13 Ερωτήσεις σχετικά με γραπτές και προφορικές εργασίες.....	82



Εικόνα 3-14 Ερωτήσεις σχετικά με τον/την διδάσκοντα/ουσα .....	83
Εικόνα 3-15 Ερωτήσεις σχετικά με το φοιτητή και το κάθε μάθημα .....	83
Εικόνα 3-16 Πεδίο για ελεύθερη έκθεση ιδεών, παρατηρήσεων και σχολίων.....	84
Εικόνα 3-17 Με μπλε χρώμα εμφανίζονται οι απαντήσεις των φοιτητών .....	84
Εικόνα 3-18 Όταν παραπάνω από 1 φοιτητές απαντήσουν, τα αποτελέσματα είναι ο μέσος όρος των απαντήσεων .....	85
Εικόνα 3-19 Όταν ο φοιτητής απαντήσει μια φορά δεν δικαιούται να ξανασυμπληρώσει το ερωτηματολόγιο.....	85
Εικόνα 3-20 Μόνο εγγεγραμμένοι χρήστες στο μάθημα μπορούν να συμπληρώσουν το ερωτηματολόγιο .....	85
Εικόνα 4-1 Αρχεία για τη δημιουργία του οπτικοακουστικού υλικού .....	86
Εικόνα 4-2 Αρχείο : bank_data.arff.....	89
Εικόνα 4-3 Αρχείο bank_data_normalized.arff.....	90
Εικόνα 4-4 Αρχεία ήχου .....	90
Εικόνα 4-5 Αρχεία Video .....	92
Εικόνα 4-6 Λογισμικό καταγραφής video .....	93
Εικόνα 4-7 Εισαγωγή πολυμεσικού περιεχομένου.....	94
Εικόνα 4-8 Time Line επεξεργασίας video .....	95
Εικόνα 4-9 Δυνατότητες επεξεργασίας ήχου.....	95
Εικόνα 4-10 Διάρκεια προβολής εικόνας .....	96
Εικόνα 4-11 Εισαγωγή εφέ μετάβασης.....	96
Εικόνα 4-12 Εφέ κύβου .....	97
Εικόνα 4-13 Χρήση δεικτών (callouts) για να δοθεί έμφαση.....	98
Εικόνα 4-14 Διάρκεια δεικτών (callouts) .....	98

Εικόνα 4-15 Εμφάνιση κατάλληλων σχημάτων σε συγκεκριμένα σημεία.....	99
Εικόνα 4-16 Κείμενα εκφώνησης.....	99
Εικόνα 4-17 Κωδικοποίηση αρχείων κειμένου .....	100
Εικόνα 4-18 Δημιουργία διαφανειών .....	100
Εικόνα 4-19 Πέντε συνολικά διαφάνειες.....	101

## 1. Εισαγωγή

Σε αυτή την πτυχιακή εργασία για το τμήμα Τεχνολογίας, Πληροφορικής & Τηλεπικοινωνιών του ΤΕΙ Καλαμάτας Παραρτήματος Σπάρτης, θα παρουσιαστούν οι τάσεις, οι τεχνολογίες, η πλατφόρμα και οι βασικές ιδέες για το e-Learning καθώς και για την Επιστήμη της μάθησης. Επίσης θα αξιολογηθεί και θα αναλυθεί η σημασία της ηλεκτρονικής μάθησης και η αποτελεσματικότερη χρήση της. Επιπροσθέτως θα μελετηθούν οι μαθησιακές ανάγκες από ψυχολογικής απόψεως, ώστε το περιεχόμενο κάθε διάλεξης να παρουσιάζεται με τον πιο αποτελεσματικό τρόπο σε κάθε φοιτητή.

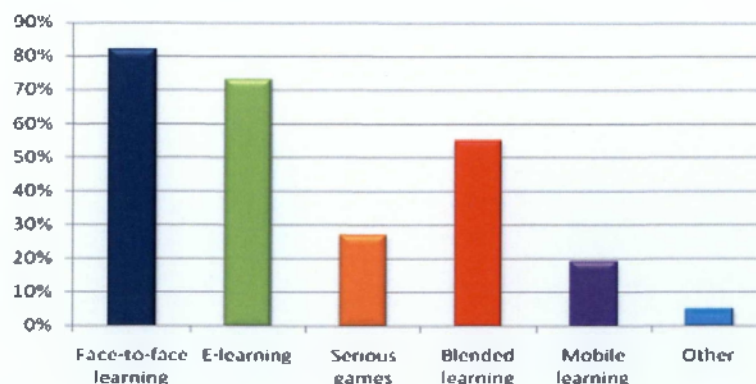
Μόλις οι αρχές της e-Learning τεκμηριωθούν, θα δημιουργηθεί ένα σύνολο δέκα κεφαλαίων που θα αφορά τις Αποθήκες Δεδομένων και το οποίο θα είναι το θέμα της μελέτης μας. Το περιεχόμενο θα τοποθετηθεί σε ένα δικτυακό τόπο Moodle, αλλά θα υπάρξει και πειραματισμός με εναλλακτικές προσεγγίσεις. Επιπλέον θα προσφερθεί περιεχόμενο για πρακτική εξάσκηση, εργασίες καθώς και άλλο υλικό. Η μεγαλύτερη πρόκληση της μελέτης αυτής είναι η δημιουργία ενός φιλικού περιβάλλοντος, όπου θα μπορεί να υπάρξει συνεργασία μεταξύ φοιτητών, καθηγητών και μελών της ακαδημαϊκής κοινότητας.

## 2. Θεωρητικό υπόβαθρο

### 2.1 Μελέτες και έρευνες για την εκπαίδευση προσωπικού

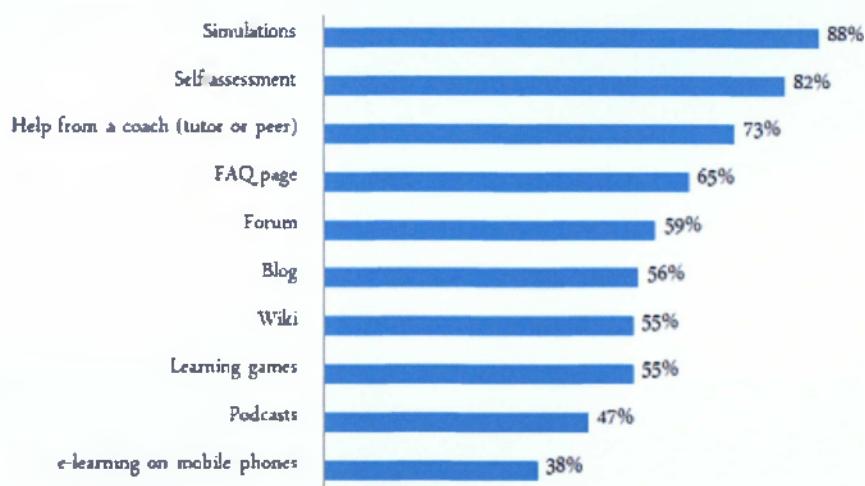
Για τον καλύτερο προσανατολισμό στα συστατικά στοιχεία που στην πραγματικότητα αφορούν την ηλεκτρονική μάθηση, είναι καλό να γίνει μια παράθεση διεθνών ερευνών όπως αυτής του Blain (2010). Στη μελέτη αυτή αναφέρεται ότι παρά την οικονομική κρίση οι διεθνείς εταιρίες θα συνεχίσουν με αμείωτο ρυθμό την επένδυση στην εκπαίδευση του προσωπικού τους. Τα στοιχεία όμως δείχνουν ότι οι εταιρίες θα στραφούν σε άλλες μορφές εκπαίδευσης και συγκεκριμένα σε τομείς e-Learning.

Το 2000 περίπου το 11,5% της εκπαίδευσης παρέχόταν μέσω συστημάτων e-Learning και το ποσοστό αυτό διαρκώς αυξανόταν, φτάνοντας το 32,6% το 2007. Από τότε υπάρχει μια μικρή μείωση των δαπανών σε συστήματα e-Learning καθώς οι εταιρίες έχοντας προσεγγίσει αρκετά το θέμα της εκπαίδευσης από πολλές πλευρές, προσπαθούν να εντοπίσουν το σωστό μείγμα όσον αφορά την καθαρή ηλεκτρονική μάθηση και την υποβοήθηση της εκπαιδευτικής διαδικασίας από πραγματικούς εκπαιδευτές. Μια κυρίαρχη τάση που παρατηρείται κυρίως στην Ευρώπη είναι η αύξηση στις επενδύσεις τόσο στο e-Learning όσο και στο Blended learning, στο συνδυασμό δηλαδή αρκετών διδακτικών μεθοδολογιών συμπεριλαμβανομένου και του e-Learning.



Εικόνα 2-1 Προγραμματισμός για την εκπαίδευση επαγγελματιών στην Αγγλία το 2009.

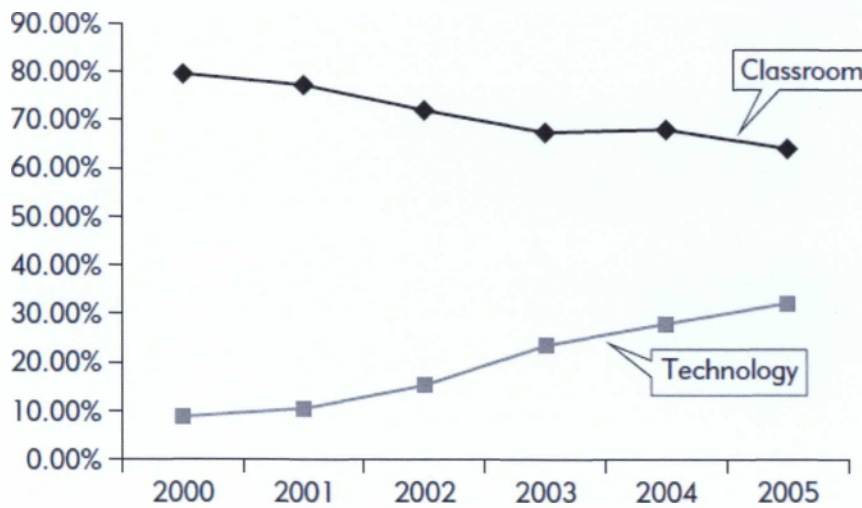
Ιδιαίτερο ενδιαφέρον έχει και η παρουσίαση στοιχείων σχετικά με τους επιθυμητούς τρόπους βελτίωσης της εκπαιδευτικής διαδικασίας. Το 88% των υπαλλήλων δηλώνουν ότι επιθυμούν να εκπαιδευτούν παρακολουθώντας πραγματικά σενάρια σχετικά με την εργασία τους, το 73% δηλώνουν ότι επιθυμούν προσωπική επαφή με τον εκπαιδευτικό, ενώ τα podcasts και η χρήση κινητών τηλεφώνων έχουν προταθεί από το 47% και 38% αντίστοιχα:



Εικόνα 2-2 Επιθυμητός τρόπος παροχής εκπαίδευσης.

Από τα παραπάνω μπορούν να εξαχθούν με ασφάλεια κάποια συμπεράσματα. Για παράδειγμα, πολλοί εκπαιδευόμενοι επιθυμούν να αξιοποιήσουν το χρόνο που χάνουν κατά τη μεταφορά τους (π.χ. στο Μετρό) μαθαίνοντας κάτι μέσω ενός Podcast ή μιας εφαρμογής από το κινητό τους. Οι περισσότεροι επιθυμούν χειροπιαστή μάθηση και όχι θεωρητική προσέγγιση, ενώ και όλες οι τεχνολογίες, όπως wikis, blogs, forums, faqs αλλά ακόμα και τα παιχνίδια μπορούν να διαδραματίσουν σημαντικό ρόλο.

Μια ακόμα παρουσίαση στατιστικών από τον Clark (2008) δείχνει τον ετήσιο ρυθμό μείωσης του ποσοστού εκπαίδευσης σε αίθουσα, σε σύγκριση με το ρυθμό αύξησης της χρήσης της τεχνολογίας για την παροχή εκπαιδευτικού περιεχομένου.



Εικόνα 2-3 Ετήσια μεταβολή ποσοστού χρήσης εκπαιδευτικών μεθοδολογιών

Σχετικά με τη διακίνηση του ηλεκτρονικού εκπαιδευτικού υλικού αξίζει να σημειωθεί ότι στην αναφορά του Galvin (2001), το 2000 το 40% του διαθέσιμου ηλεκτρονικού εκπαιδευτικού υλικού διατίθεται μέσω CD/DVD ενώ μόνο το 22% μέσω του Ιντερνέτ (με το υπόλοιπο 30% να διατίθεται μέσω εταιρικών Intranets). Αντίστοιχες μελέτες δείχνουν ότι το 2010 παραπάνω από το 90% του εκπαιδευτικού υλικού διακινείται μέσω του διαδικτύου.

## 2.2 Η επιστήμη της μάθησης

Στο κεφάλαιο αυτό θα παρουσιαστεί αρχικά ο ορισμός της **μάθησης**. Υπάρχουν τρεις κύριες τάσεις:

α) Μάθηση νοείται η ενίσχυση των σωστών απαντήσεων και η αποδυνάμωση των εσφαλμένων. Στην περίπτωση αυτή ο εκπαιδευόμενος θεωρείται ως παθητικός αποδέκτης και ο ρόλος του εκπαιδευτικού είναι να απονέμει τιμωρίες ή επιβραβεύσεις ώστε να επιτευχθεί ο στόχος της μάθησης.

β) Μάθηση νοείται η καταχώρηση νέων πληροφοριών στην μνήμη. Στην περίπτωση αυτή ο εκπαιδευόμενος προσπαθεί να καταχωρήσει όσο το δυνατόν περισσότερες πληροφορίες, ενώ ο εκπαιδευτικός καλείται να παρουσιάσει τις πληροφορίες αυτές.

γ) Μάθηση νοείται η κατανόηση του υλικού και των πληροφοριών που παρουσιάζονται στον εκπαιδευόμενο, καθώς και η νοητική αναδιοργάνωση αυτού του υλικού, αλλά και η σύνδεσή του με γνώσεις που ήδη διαθέτει ο εκπαιδευόμενος.

Σε αυτή την περίπτωση ο εκπαιδευόμενος θεωρείται ότι έχει ενεργητικό μέρος στην εκπαιδευτική διαδικασία.

Οι παραπάνω προσεγγίσεις είναι υπαρκτές από την επιστήμη της ψυχολογίας εδώ και δεκαετίες και χρησιμοποιούνται από εκπαιδευτικούς όλων των βαθμίδων της εκπαίδευσης.

### **2.3 Οι τύποι των μαθητών**

Ένα ακόμα σημαντικό κεφάλαιο σχετικά με την επιστήμη της μάθησης είναι η ύπαρξη διαφορετικών **αισθητηριακών τύπων μαθητών**. Οι άνθρωποι χρησιμοποιούν όλους τους αισθητήριους δέκτες: την ακοή, την όραση και την κιναισθησία (την κίνηση και την αφή) για να προσλάβουν πληροφορίες (Ιωάννης Βανδουλάκης). Σε κάθε άνθρωπο, μια από τις αισθήσεις αυτές είναι κυρίαρχη και καθορίζει τον τύπο του ατόμου, δηλαδή τον καλύτερο τρόπο μάθησης.

Ο ρόλος του εκπαιδευτικού και κατ' επέκταση των συστημάτων e-Learning είναι να παρουσιάζει τις πληροφορίες χρησιμοποιώντας και τους τρεις τύπους, ώστε να καλύπτονται αποτελεσματικότερα όλοι οι μαθητές. Έτσι ενώ κάθε άτομο χαρακτηρίζεται ως οπτικός, ακουστικός ή κιναισθητικός τύπος, οι άλλες δύο μέθοδοι λειτουργούν ενισχυτικά.

Για τους **οπτικούς τύπους** υπάρχουν δύο κύριες κατηγορίες: ο γλωσσικός που προτιμά τις παρουσιάσεις και γενικότερα την ανάγνωση του κειμένου και ο χωρικός που έχει καλύτερη απόδοση με τα εποπτικά υλικά (γραφικές παραστάσεις, διαγράμματα, βίντεο κτλ).

Οι **ακουστικοί τύποι** έχουν κατά κανόνα καλύτερες επιδόσεις όταν ακούν ένα μάθημα ή συζητούν με κάποιο συνάδελφο. Οι λεκτικές οδηγίες προς αυτούς είναι ιδιαίτερα αποτελεσματικές.

Οι **κιναισθητικοί τύποι** αφομοιώνουν καλύτερα με την αφή και με την κίνηση. Επιθυμούν να συμμετέχουν ενεργητικά στη διαδικασία της μάθησης, να κρατούν σημειώσεις και να πειραματίζονται άμεσα με κάθε νέα γνώση.

Στατιστικά, το μεγαλύτερο μέρος του πληθυσμού (περίπου το 80%) είναι οπτικοί τύποι, περίπου το 10% είναι ακουστικοί και 10% κιναισθητικοί (Ιωάννης Βελίκης).

Παρατηρώντας την εκπαιδευτική διαδικασία σε όλες τις βαθμίδες της (από πρωτοβάθμια μέχρι και τριτοβάθμια) μπορεί κάποιος να δει μια εύνοια προς τους ακουστικούς. Η παράδοση του μαθήματος γίνεται κατά κανόνα λεκτικά με οπτικά ερεθίσματα να παρουσιάζονται κυρίως στις θετικές επιστήμες. Τελευταία παρατηρείται διεθνώς μια τάση προς την εισαγωγή περισσότερων ερεθισμάτων στην εκπαιδευτική διαδικασία. Για παράδειγμα η χρήση των διαδραστικών πινάκων που σταδιακά εισάγονται σε δημοτικά σχολεία, γυμνάσια και λύκεια καθώς και η ευρεία χρήση των προβολέων και άλλων οπτικών μέσων δείχνουν σε ένα μέλλον, όπου ο μαθητής θα παρακολουθεί τόσο οπτικά, ακουστικά αλλά και κινητικά τον καθηγητή να παραδίδει το μάθημα του.

#### **2.4 Η χρήση επιπλέον - ενδιαφέροντος υλικού**

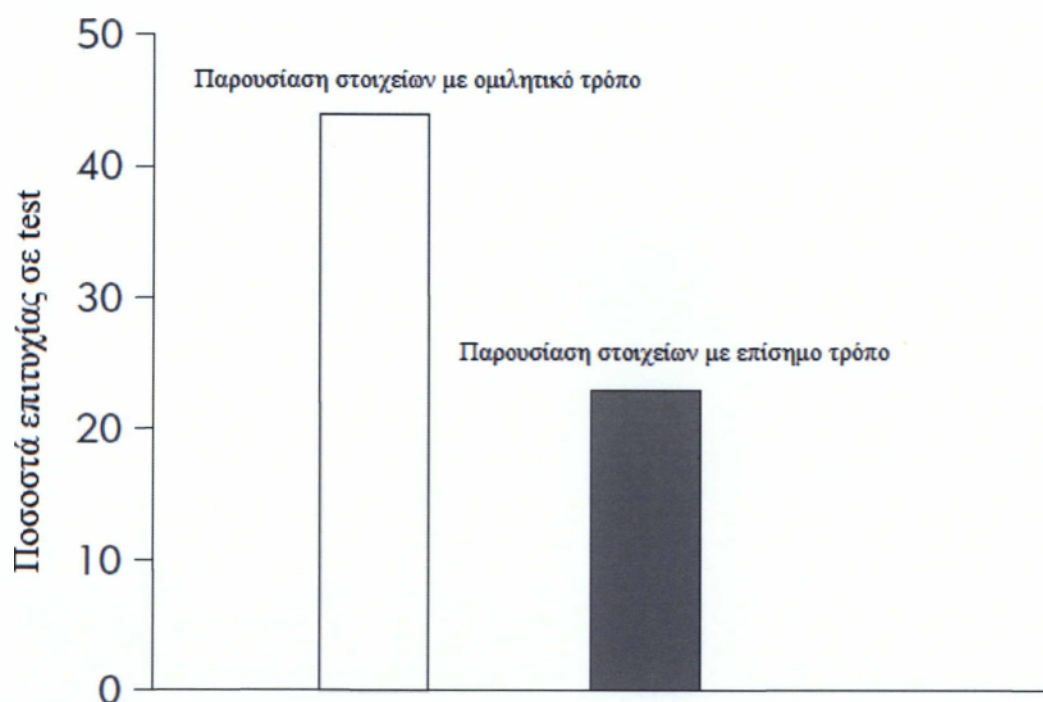
Ένα ακόμα ενδιαφέρον σημείο είναι ότι μελέτες έχουν αποδείξει πως η εισαγωγή επιπλέον πληροφοριών κατά την παρουσίαση ενός μαθήματος επιβραδύνουν τη διαδικασία της μάθησης. Συγκεκριμένα οι Mayer, Heiser, και Lonn ( 2001) έκαναν μια έρευνα κατά την οποία παρουσίασαν σε δύο ομάδες μαθητών υλικό σχετικά με τη δημιουργία κεραυνών. Στην πρώτη περίπτωση το μάθημα περιορίστηκε στην παρουσίαση των στοιχείων και των φυσικών γεγονότων. Στη δεύτερη ομάδα χρηστών παρουσιάστηκε το ίδιο ακριβώς υλικό, με τη διαφορά ότι εισήχθησαν ορισμένες επιπλέον ενδιαφέρουσες πληροφορίες, οι οποίες όμως δεν ήταν σχετικές με τη μάθηση. Για παράδειγμα, μια επιπλέον πληροφορία ήταν ότι οι παίκτες του γκολφ κινδυνεύουν αρκετά όταν παίζουν κατά τη διάρκεια βροχής λόγω πιθανής ηλεκτροπληξίας, επειδή κρατούν μεταλλικά μπαστούνια. Από παρόμοιες μελέτες αποδείχθηκε ότι η παρουσίαση άσχετων εικόνων, κειμένων και ομιλίας μειώνουν το αποτέλεσμα της εκπαίδευσης κατά 30-35%!



## 2.5 Τρόπος επικοινωνίας

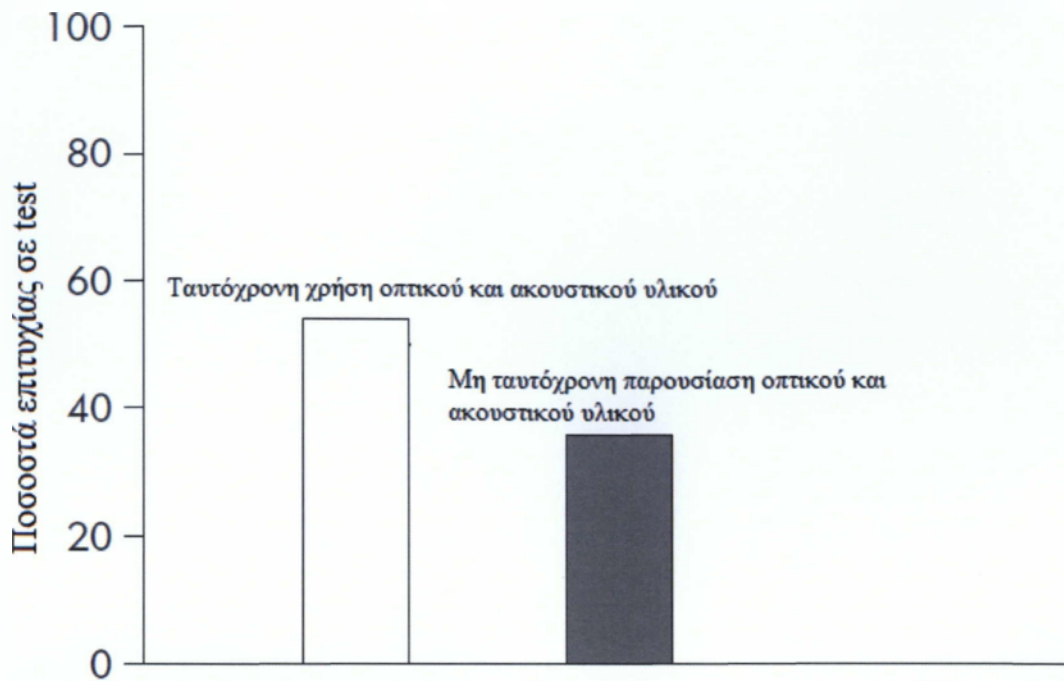
Υπάρχουν αρκετές διαφορετικές απόψεις σχετικά με το πρόσωπο στο οποίο πρέπει να μεταδίδεται η πληροφορία από τον εκπαιδευτικό προς τον εκπαιδευόμενο. Πολλοί καθηγητές προτιμούν τον τυπικό/επίσημο τόνο, ενώ άλλοι χρησιμοποιούν ανεπίσημο τρόπο μιλώντας συχνά σε πρώτο πρόσωπο, δημιουργώντας ένα διάλογο με τους ακροατές τους.

Μέσα από σειρά μελετών που παρουσιάζονται στο Moreno και Mayer (2000b, 2004) είναι εμφανές ότι η χρήση ομιλητικού τρόπου κατά την παρουσίαση στοιχείων έχει ως και δύο φορές μεγαλύτερη αποτελεσματικότητα.



Εικόνα 2-4 Στοιχεία από μελέτη σχετικά με τον τρόπο επικοινωνίας προς τον εκπαιδευόμενο

Επίσης έχουν αποδειχθεί πειραματικά όσα παρουσιάστηκαν στο κεφάλαιο 2.3 σχετικά με τους τύπους μαθητών. Η χρήση τόσο οπτικού όσο και ακουστικού υλικού λειτουργεί ενισχυτικά στη διαδικασία της μάθησης. «Η ταυτόχρονη χρήση του ακουστικού και του οπτικού υλικού μάλιστα επιτυγχάνει τα καλύτερα αποτελέσματα», Mayer και Anderson (2001):

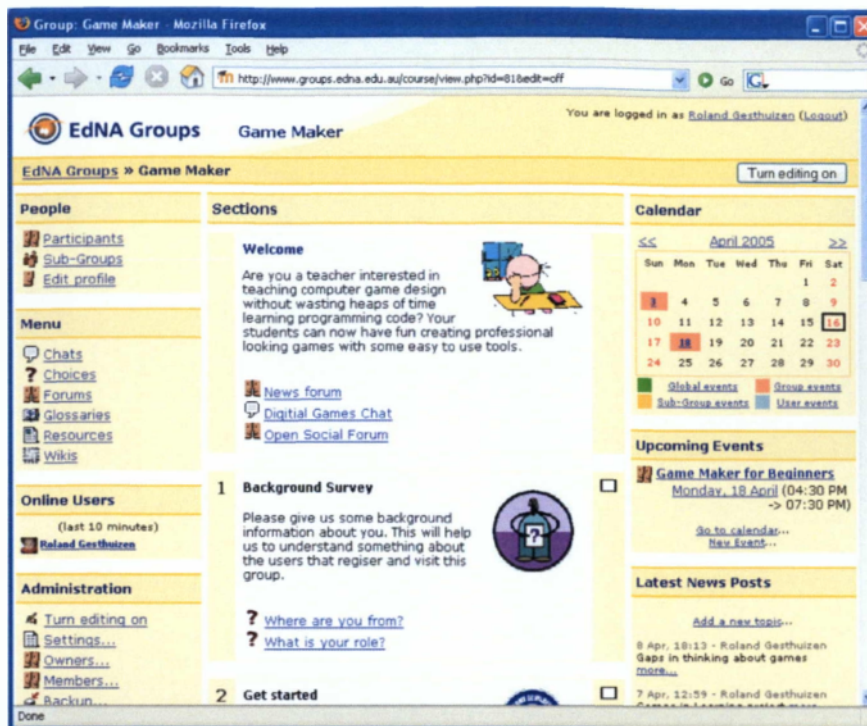


Εικόνα 2-5 Στοιχεία από μελέτη σχετικά με τη χρήση οπτικού και ακουστικού υλικού

## 2.6 Η ιστορία του moodle

Το moodle (Modular Object - Oriented Dynamic Learning Environment) αποτελεί μια ασύγχρονη πλατφόρμα τηλεκπαίδευσης που άρχισε να δημιουργείται το 1999 από τον Αυστραλό φοιτητή ελληνικής καταγωγής Martin Dougiamas, ο οποίος έκανε τη διδακτορική του εργασία. Είναι μια εφαρμογή ανοιχτού κώδικα, γραμμένη σε PHP που παρέχεται δωρεάν κάτω από την άδεια GNU.

Μπορεί να εγκατασταθεί και να λειτουργήσει σε οποιοδήποτε υπολογιστή που έχει εγκατεστημένο έναν εξυπηρετητή ιστού και μια σχεσιακή βάση δεδομένων SQL. Σκοπός του moodle είναι να παρέχει ένα περιβάλλον μέσα από το οποίο οι εκπαιδευτικοί θα μπορούν να παρουσιάζουν υλικό εκπαίδευσης και οι εκπαιδευόμενοι να το παρακολουθούν απομακρυσμένα. Σταδιακά στο πακέτο αυτό του λογισμικού προστέθηκαν αρκετές επιπλέον δυνατότητες που υπαγορεύει η παιδαγωγική και η επιστήμη της εκπαίδευσης σχετικά με την αλληλεπίδραση, μεταξύ των μαθητών και μεταξύ του καθηγητή και αυτών. Πλέον, το 2010 μπορεί να πει κανείς με ασφάλεια ότι το moodle, δεδομένου ότι είναι δωρεάν, αποτελεί την καλύτερη επιλογή για παροχή απομακρυσμένης e-Learning εκπαίδευσης. Τέλος διατίθεται σε περίπου 75 γλώσσες.



Εικόνα 2-6. Μια χαρακτηριστική σελίδα του Moodle

Η φιλοσοφία που χρησιμοποιήθηκε κατά την ανάπτυξη του Moodle είναι γνωστή ως **Social Constructionist Pedagogy**. Οι αρχές πάνω στις οποίες στηρίζεται η παιδαγωγική αυτή μεθοδολογία είναι οι παρακάτω:

1. Όλοι είμαστε υποψήφιοι εκπαιδευτικοί καθώς και μαθητές - στο πλαίσιο ενός πραγματικά συνεργατικού περιβάλλοντος.
2. Μαθαίνουμε ιδιαίτερα καλά όταν δημιουργούμε κάτι, ώστε να το παρουσιάσουμε ή να το εκφράσουμε σε κάποιον.
3. Όλοι μπορούμε να διδαχθούμε πολλά παρατηρώντας και μόνο τη δραστηριότητα των γύρω μας.
4. Με την κατανόηση του περιβάλλοντος των άλλων, μπορούμε να διδάξουμε με αποδοτικότερο τρόπο (κονστрукτιβισμός).
5. Ένα εκπαιδευτικό περιβάλλον πρέπει να είναι ευέλικτο και ευπροσάρμοστο, ώστε να μπορεί να ανταποκριθεί γρήγορα στις ανάγκες των συμμετεχόντων σε αυτό.

Γενικώς υπάρχουν δύο κατηγορίες χρηστών στο Moodle, οι εκπαιδευτικοί και οι εκπαιδευόμενοι. Όσο περισσότερο εξοικειώνεται κάποιος με το εργαλείο αυτό συνήθως περνάει από τις παρακάτω φάσεις:

1. Τοποθέτηση κειμένων και παρουσιάσεων.
2. Δημιουργία ενός φόρουμ συζητήσεων.
3. Χρήση quiz και εργασιών.
4. Ενσωμάτωση Wiki, Γλωσσαρίου εργαλείων βασιζόμενα στη βάση δεδομένων.
5. Ενεργή αξιοποίηση του φόρουμ.
6. Συνδυασμός δραστηριοτήτων σε ακολουθίες. Το αποτέλεσμα κάθε ακολουθίας χρησιμοποιείται στην επόμενη μαθησιακή δραστηριότητα.
7. Εις βάθος αναδιοργάνωση του υλικού και της μεθοδολογίας και ενσωμάτωση περισσότερων δυνατοτήτων.
8. Χρήση module για έρευνα όπως π.χ. ερωτηματολόγια, για την κατανόηση της αποτελεσματικότητας του εκπαιδευτικού υλικού και της παρουσιάσής του.
9. Χρήση εξειδικευμένων εργαλείων όπως το Workshop για την εξατομικευμένη παρακολούθηση του κάθε εκπαιδευόμενου.

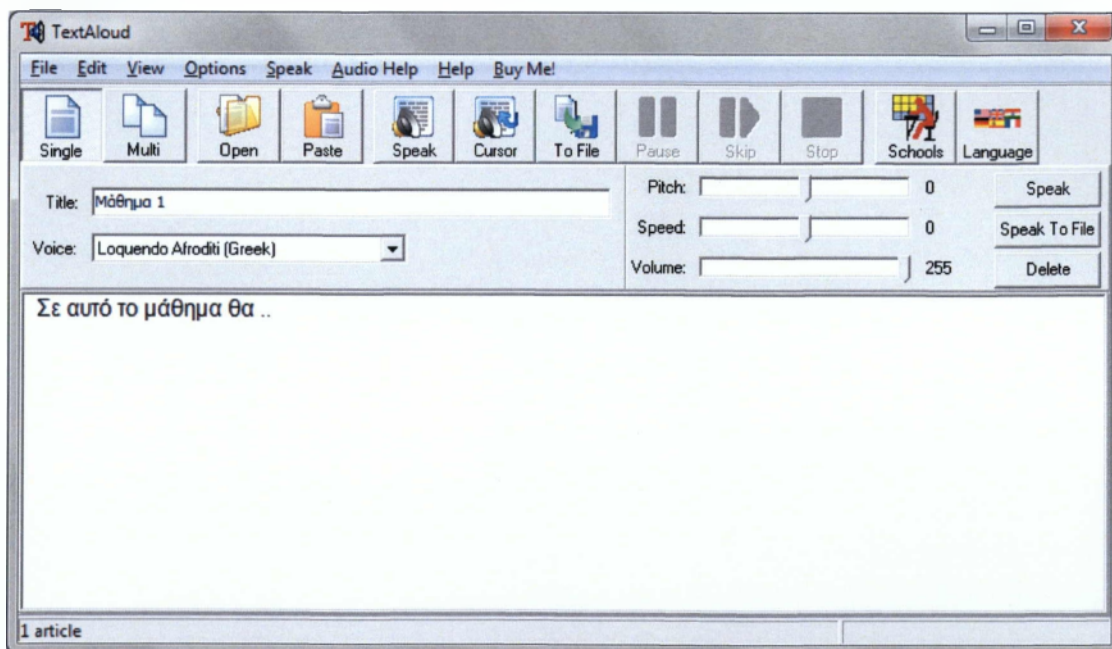
## **2.7 Text-To-Speech**

Μια ακόμα τεχνολογία που έχει βελτιωθεί σχετικά με τη δημιουργία ακουστικού υλικού είναι η TTS (Text-To-Speech), δηλαδή η δημιουργία ήχου του εκφωνητή με χρήση πρωτότυπου κειμένου. Συχνά χρησιμοποιούνται μέθοδοι ψηφιακής ηχογράφησης με χρήση μικροφώνου σε επαγγελματικά στούντιο για τη δημιουργία του ακουστικού υλικού.

Η παραπάνω διαδικασία προϋποθέτει χρόνο αλλά και κόστος. Επίσης για να γίνει ακόμα και μια μικρή διόρθωση πρέπει κάποιος να ξαναμπεί στο στούντιο. Με τη χρήση κάποιου λογισμικού τεχνολογίας Text-To-Speech, όπως του TextAloud και με τη χρήση της Ελληνικής γυναικείας φωνής «Αφροδίτη» της εταιρίας Loquendo, μπορούν να δημιουργηθούν αρκετά αξιόλογες ηχογραφήσεις.

Τα λογισμικά TTS χρησιμοποιούνται όλο και περισσότερο τελευταία στην Ελλάδα. Οι υπηρεσίες phone-banking αρκετών τραπεζών στηρίζονται πλέον σε αυτού του είδους τις εφαρμογές που αναλαμβάνουν την ανάγνωση και την εκφώνηση γραπτού κειμένου. Επίσης άλλες εφαρμογές των συστημάτων TTS είναι η παροχή βοήθειας σε άτομα με προβλήματα στην όραση ή την ομιλία, η ακουστική προσπέλαση βάσεων κειμένου και η εκφώνηση οδηγιών. Τέλος όσον αφορά τα τεχνικά χαρακτηριστικά τους, τα υπάρχοντα συστήματα TTS στηρίζονται στη μοντελοποίηση του ανθρώπινου συστήματος παραγωγής φωνής ή στη μοντελοποίηση του σήματος φωνής. Στη δεύτερη κατηγορία ανήκουν και τα συστήματα που συνθέτουν φωνή, συρράπτοντας στοιχειώδη κομμάτια της (πχ φθόγγους, συλλαβές) που ονομάζονται φωνητικές μονάδες.

Ως αποτέλεσμα η τεχνολογία αυτή έγινε έναυσμα για σχετική έρευνα. Συγκεκριμένα, μέσα από την εφαρμογή TextAloud ο χρήστης μπορεί να ρυθμίσει την ταχύτητα, την ένταση και τον τόνο της φωνής, ενώ τέλος μπορεί να εξάγει τα ηχητικά clip σε wav format.



Εικόνα 2-7 Το γραφικό περιβάλλον της εφαρμογής TextAloud

### Ορισμοί

**Virtual Classroom** ονομάζεται η υβριδική χρήση ασύγχρονης εκπαίδευσης e-Learning αλλά και επικοινωνίας, σε πραγματικό χρόνο μεταξύ εκπαιδευτικού και εκπαιδευόμενου.

### 3. Παρουσίαση μαθημάτων

Στο κεφάλαιο αυτό θα παρουσιαστούν τα δέκα μαθήματα που δημιουργήθηκαν στα πλαίσια της πτυχιακής αυτής εργασίας και θα σχολιαστούν οι επιλογές που έγιναν για την προετοιμασία των κειμένων αυτών.

Αρχικά παρουσιάζεται η περιγραφή του μαθήματος «Αποθήκες Δεδομένων», ο σκοπός του μαθήματος, ο ορισμός και κατόπιν υπάρχουν δύο επιλογές σχετικά με α) την αναλυτικότερη περιγραφή του μαθήματος (όπως π.χ. σε ποιο εξάμηνο ανήκει, ποιες σχολές αφορά, ποιος είναι ο υπεύθυνος καθηγητής και σχετικά με τα απαραίτητα εργαστήρια) β) έναν πίνακα ανακοινώσεων όπου κάποιος μπορεί να συζητήσει σχετικά με αυτές.

**Περιγραφή θέματος**


Ο τεράστιος όγκος πληροφοριών κρύβει μέσα του "γνώση" η οποία είναι ανεξερεύνητη. Συσχετίσεις μεταξύ δεδομένων, αλληλουχίες γεγονότων, προβλέψεις για στόχους και πολλά άλλα είναι θαμμένα μέσα στα δεδομένα και περιμένουν με κάποιο τρόπο να έρθουν στην επιφάνεια. Η γνώση αυτή είναι κρυμμένη γιατί τα δεδομένα εξαρχής είναι οργανωμένα σε δομές οι οποίες εξυπηρετούν συγκεκριμένους σκοπούς.


**Σκοπός του μαθήματος**

*είναι να κάνουμε το πρώτο βήμα για την διαχείριση του πρόσπου όγκου των δεδομένων ,που έχει γίνει αναπόσπαστο κομμάτι πολλών εταιριών πληροφοριακών συστημάτων διοίκησης και ακόμα περισσότερο σε συστήματα υποστήριξης αποφάσεων αφού διαχειρίζεται δισεκατομμύρια bytes δεδομένων ή ακόμα και terabytes.*

**Ορισμός**

*Ο κλασικός ορισμός μιας Αποθήκης Δεδομένων είναι:  
«μία συλλογή θεματικών, εντοπιζόμενων, μη-εμμεταβλητών, μεταβαλλόμενων με τον χρόνο δεδομένων για την υποστήριξη διοικητικών αποφάσεων»  
(W.H. Inmon. 1992)*

 Περιγραφή του μαθήματος

 Ομάδα συζητήσεων ειδήσεων

Εικόνα 3-1 Αρχική εισαγωγή στο μάθημα «Αποθήκες Δεδομένων»

Στη συνέχεια ακολουθούν τα δέκα μαθήματα:

### 3.1 Μάθημα 1<sup>ο</sup> - Εισαγωγή στο Data Warehousing

Το πρώτο μάθημα αφορά την εισαγωγή στις αποθήκες δεδομένων. Παρουσιάζονται η ιστορία τους, η χρήση τους και η δομή τους.

1



#### Μάθημα 1ο: Εισαγωγή στο Data Warehousing

Στο μάθημα αυτό θα παρουσιάσουμε τις βασικές ορολογίες σχετικά με τα data warehouses, τη σκοπιμότητα τους, τα μεγέθη τους, τη χρήση τους και άλλες εισαγωγικές έννοιες.

- 1. Δημιουργία και Ιστορία των data warehouses
- 2. Χρήση και αξιοποίηση των Data Warehouse συστημάτων
- 3. Δομή ενός Data Warehouse

Εικόνα 3-2 Παρουσίαση πρώτου μαθήματος

Το μάθημα, ακολουθώντας τις παιδαγωγικές μεθοδολογίες, παρουσιάζει τον ορισμό των αποθηκών δεδομένων και αμέσως μετά ταξινομεί τους τύπους των δεδομένων που συνήθως αποθηκεύονται σε αυτά. Παρουσιάζεται μια σειρά απαραίτητων ενεργειών για τη δημιουργία των συστημάτων αυτών και κατόπιν ξεκινά μια παρουσίαση ιστορικών στοιχείων από το μακρινό παρελθόν μέχρι σήμερα.

## ΚΕΦΑΛΑΙΟ 1

Οι αποθήκες δεδομένων (data warehouses) επικεντρώνονται στη μοντελοποίηση και ανάλυση δεδομένων για τη λήψη αποφάσεων και όχι για καθημερινές λειτουργίες ή επεξεργασία δοσοληψιών. Παρέχουν μια απλή και συνοπτική όψη των δεδομένων για συγκεκριμένα θέματα και όχι τα δεδομένα τα οποία δεν είναι χρήσιμα στη διαδικασία λήψης αποφάσεων. Κατασκευάζονται συνήθως από πολλαπλές ετερογενείς πηγές δεδομένων και ο χρονικός ορίζοντας τους είναι μεγαλύτερος από τις περισσότερες εφαρμογές που χρησιμοποιεί μια επιχείρηση.

Τα δεδομένα μιας επιχείρησης μπορούν να ταξινομηθούν σε:

- Δεδομένα που **καθημερινά προστίθενται στην επιχείρηση**, όπως αυτά που



δημιουργούνται κατά την εγγραφή ενός χρήστη στην ιστοσελίδα της επιχείρησης, κατά την πληρωμή κτλ.

- Δεδομένα που **χρησιμοποιούνται για να συνδέσουν διαφορετικές εφαρμογές** μεταξύ τους. Για παράδειγμα μια επιχείρηση χρησιμοποιεί για τη μισθοδοσία υπαλλήλων μια εφαρμογή γραμμένη σε VB.NET και το τμήμα Human Resources μια εφαρμογή σε PHP και SQL για να καταχωρεί τις άδειες των υπαλλήλων. Τα δύο συστήματα είναι φαινομενικά ασύνδετα. Πολλές φορές όμως χρησιμοποιούνται master λίστες π.χ. χρησιμοποιώντας τα Α.Φ.Μ. των υπαλλήλων ώστε τα φαινομενικά ασύνδετα δεδομένα να διασυνδεθούν μεταξύ τους.
- **Δεδομένα παρακολούθησης οικονομικών**, managerial και άλλων στοιχείων που κατασκευάζονται από υπαλλήλους, που συνήθως είναι σε μορφή παρουσιάσεων και βοηθούν στη λήψη αποφάσεων.

### **Πως δημιουργούνται τα Data Warehouses ?**

Τα βήματα που ακολουθούν οι επιχειρήσεις για τη δημιουργία ενός data warehouse είναι συνήθως τα εξής:

1. Μια ομάδα αναλυτών και προγραμματιστών μελετά τις απαιτήσεις του συστήματος.
2. Ορίζονται κάποιοι υπάλληλοι σε κάθε τμήμα που αναλαμβάνουν ως ειδικοί να εξηγούν στην ομάδα που δημιουργεί το Data Warehouse τις απαιτήσεις όσον αφορά τα δεδομένα.
3. Η ομάδα δημιουργίας μελετά το είδος, την ποσότητα των δεδομένων και τον τρόπο ανάκτησής τους.
4. Αναπτύσσονται εφαρμογές οι οποίες αναλαμβάνουν την εξόρυξη δεδομένων. Ορισμένα δεδομένα για παράδειγμα μπορεί να βρίσκονται σε λογαριασμούς email ή σε αρχεία Excel. Κατάλληλες εφαρμογές θα έχουν τη δυνατότητα να εντοπίσουν ότι πιο χρήσιμο, ώστε να μπορεί στη συνέχεια να αξιοποιηθεί.

## Η ιστορία των Data Warehouses

Ο πατέρας του marketing θεωρείται από πολλούς ο **Charles Coolidge Parlin (1872–1942)** καθώς ήταν ένας από τους πρωτοπόρους που συγκέντρωσε στοιχεία σχετικά με τους αναγνώστες της εφημερίδας The Saturday Evening Post, ώστε να διαφημίσει στα φύλλα της προϊόντα που αφορούν τους συγκεκριμένους πελάτες.

Το **1923 ο Arthur C. Nielsen** ίδρυσε την πρώτη εταιρία στην ιστορία σχετικά με το marketing.

Οι δύο αυτές προσπάθειες αποτελούν ορόσημο στη συγκέντρωση δεδομένων. Το επόμενο βήμα θεωρείται από πολλούς η **δεκαετία του 70**. Δεδομένα άρχισαν να τοποθετούνται στους πρώτους υπολογιστές της εποχής. Βέβαια οι σχεσιακές βάσεις δεδομένων που σήμερα είναι τόσο διαδεδομένες, δεν υπήρχαν. Αντιθέτως γινόταν χρήση απλών αρχείων και κάθε εταιρία είχε δικό της λογισμικό, δικό της τρόπο αναζήτησης και τοποθέτησης δεδομένων. Μάλιστα, ενώ τα δεδομένα αποθηκεύονταν αποτελεσματικά ή εξόρυξη γνώσεων όπως για παράδειγμα 'Πόσες πωλήσεις του X προϊόντος έγιναν στην Y πόλη από άνδρες μεγαλύτερους των 35 ετών' ήταν σχεδόν αδύνατο να υλοποιηθούν!

Οραματιστές της εποχής ξεκίνησαν τη δεκαετία αυτή να χρησιμοποιούν ορισμένες τεχνικές. Αντί να ενοχλούν διαρκώς το τμήμα διαχείρισης δεδομένων, τους ζητούσαν περιοδικά να κάνουν backup σε μια ταινία. Οι ταινίες αυτές με τα δεδομένα μεταφέρονταν σε άλλους υπολογιστές, όπου υπάλληλοι ασχολούνταν με την εξόρυξη των πλέον χρήσιμων δεδομένων. Τα μειονεκτήματα παρέμεναν πολλά.

Αρχικά οι ταινίες αυτές κρατούσαν περιορισμένο όγκο δεδομένων. Αναζητήσεις σε δεδομένα παρελθόντων ετών ήταν δύσκολες ή αδύνατες. Επίσης αν η εταιρία έπαιρνε backups από τους υπολογιστές κάθε πόλης ήταν πάλι δύσκολη η αναζήτηση και σύγκριση δεδομένων ανάμεσα στα καταστήματα των διαφορετικών πόλεων.

Η **δεκαετία του 80** σημαδεύτηκε από την εμφάνιση των Personal Computers. Κάθε γραφείο ή τμήμα μιας εταιρίας είχε τους δικούς της υπολογιστές και τα δικά της δεδομένα.

Τα δεδομένα άρχισαν ξαφνικά να βρίσκονται σε δεκάδες διαφορετικά σημεία αντί να είναι συγκεντρωμένα σε ένα κεντρικό mainframe.

Η λύση που προτάθηκε ήταν τα Distributed Database Management Systems (DBMS). Τα δεδομένα δηλαδή να συγκεντρώνονται από όλα τα τμήματα σε ένα εξυπηρετητή που να προσφέρει απομακρυσμένα τις υπηρεσίες του.

Τη **δεκαετία του 90** η υποστήριξη λήψης αποφάσεων δεν είχε ακόμα υλοποιηθεί με επιτυχία, παρά τις προόδους που είχαν σημειωθεί. Το hardware αλλά και το software υπήρχε. Η Client/Server αρχιτεκτονική και οι ιδέες των Bill Inmon στο βιβλίο Building the Data Warehouse (Wiley) το 1993 και του Ralph Kimball στο The Data Warehouse Toolkit (Wiley), δύο ανθρώπων που θεωρούνται ως οι πατέρες του Data Warehousing άρχισαν να ξεκαθαρίζουν το τοπίο.

Το **2000** οι εταιρίες που ασχολούνταν με το Data Warehousing ήταν κολοσσοί της βιομηχανίας λογισμικού. Εξαγορές και συγχωνεύσεις, όπως για παράδειγμα η εξαγορά της ProClarity από την Microsoft το 2006, της Hyperion από την Oracle το 2007, της Business Objects από την SAP και η συγχώνευση της IBM με την Cognos άλλαξαν δραματικά το τοπίο.

Οι παράλληλες εξελίξεις της τελευταίας αυτής δεκαετίας ήταν η ευρεία χρήση ανοιχτών προτύπων και στανταρς όπως του XML και η παρουσία ορισμένων επιχειρηματικών λύσεων όπως αυτές των εταιριών Jaspersoft (Jasper reports) και Pentaho αλλά ακόμα και open-source λύσεων.

### **Τα σημερινά συστήματα Data Warehousing**

Η ιστορική παρουσίαση που μόλις έγινε είναι ιδιαίτερα σημαντική για να κατανοήσει κανείς το εξής γεγονός. Κανένα σύγχρονο Data Warehousing προϊόν δεν μπορεί να αγνοήσει τις προσπάθειες του παρελθόντος. Αν για παράδειγμα μια τράπεζα χρησιμοποιούσε για δεκαετίες δικά της συστήματα βασισμένα σε, παρωχημένες σημερινά τεχνολογίες, όπως CORBA και COBOL δεν μπορεί σε μια στιγμή να αξιοποιήσει πλήρως σύγχρονες τεχνολογίες όπως Web 2.0 / XML και Web Services. Πρέπει οπωσδήποτε να χρησιμοποιηθούν δηλαδή διεπαφές με όλες τις τεχνολογίες

του παρελθόντος ώστε να είναι αποτελεσματικό το όποιο νέο σύστημα.

Ένα ακόμη στοιχείο που πρέπει να τονιστεί είναι ότι το Data Warehousing αποτελεί μια σύγχρονη αντιμετώπιση των τεράστιων εταιρικών δεδομένων. Η μεταφορά του συνόλου των δεδομένων μιας επιχείρησης σε ένα σύστημα δεν αποτελεί δηλαδή αυτοσκοπό. Πρέπει να γίνει η επιλογή εκείνων των δεδομένων που μπορούν να αξιοποιηθούν για τη λήψη αποφάσεων.

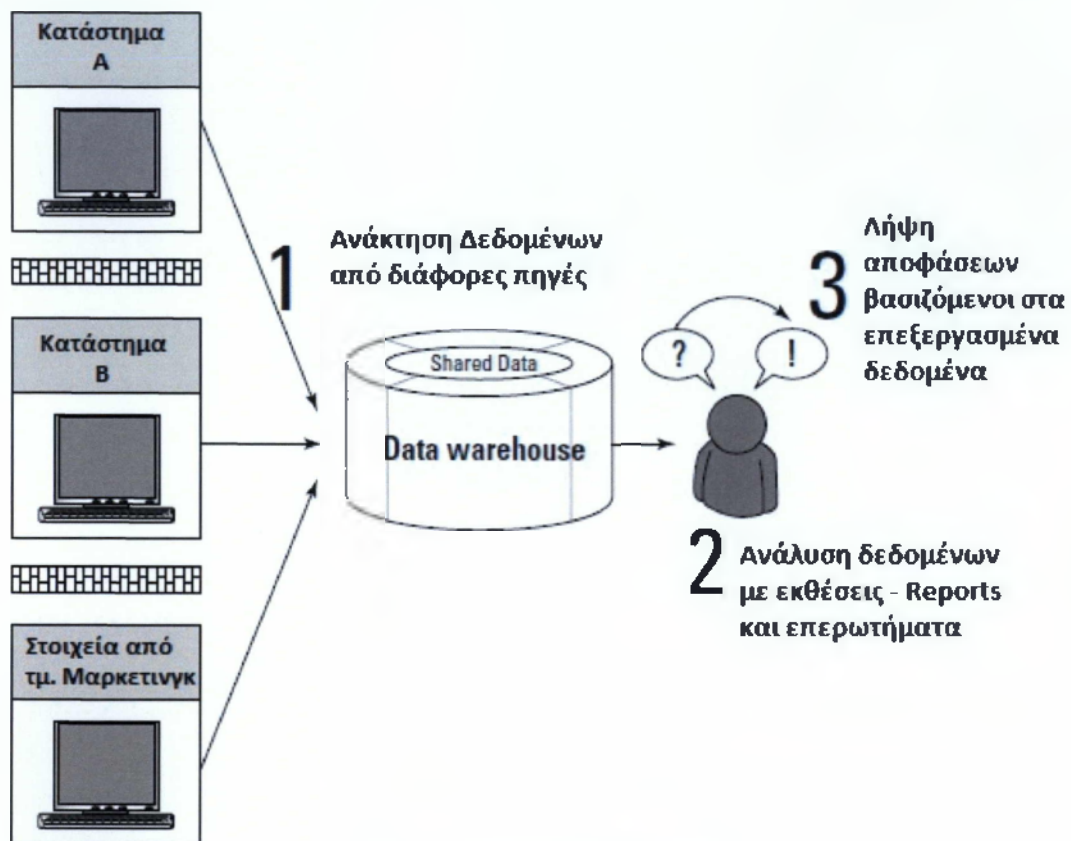
## ΚΕΦΑΛΑΙΟ 2

Στο κεφάλαιο αυτό θα παρουσιαστούν ορισμένα σενάρια χρήσης ώστε ο μαθητής να καταλάβει συγκεκριμένες διαφορές ανάμεσα στη χρήση και στην αξιοποίηση των Data Warehouse συστημάτων.

Για παράδειγμα, ο υπάλληλος **A** μιας τράπεζας έχει πρόσβαση και χρησιμοποιεί τις υπηρεσίες Data Warehousing σχετικά με την αναζήτηση πληροφοριών και τη δημιουργία ενός report. Αφού η διαδικασία ολοκληρωθεί ελέγχει το report, και το αρχειοθετεί.

Αντίστοιχος υπάλληλος **B** της ίδιας εταιρίας σε άλλη όμως περιοχή κάνει ακριβώς την ίδια διαδικασία. Μελετώντας το report που του επιστρέφει το σύστημα κυκλώνει ένα σημείο που παρουσιάζει ενδιαφέρον, δηλαδή ορισμένους αριθμούς και κρατά σημειώσεις. Ύστερα από κάποια τηλέφωνα καταλαβαίνει ότι ένα υποκατάστημα παρουσιάζει μείωση καταθέσεων πελατών. Το συμπέρασμα λοιπόν είναι ότι ο ανταγωνιστικός οργανισμός κάνει μια προσφορά που οδηγεί σε μείωση της διείσδυσης στην επιχείρηση της εταιρίας του υπαλλήλου! Τα συμπεράσματα αυτά τοποθετούνται στο σύστημα Data Warehousing και λαμβάνεται μια απόφαση ώστε να αντιμετωπιστεί το θέμα.

Η χρήση του συστήματος από τον υπάλληλο **A** είναι επιφανειακή και αναποτελεσματική. Αντιθέτως ο υπάλληλος **B** είναι αυτός που αξιοποίησε πλήρως τις δυνατότητες που του παρέχονται.



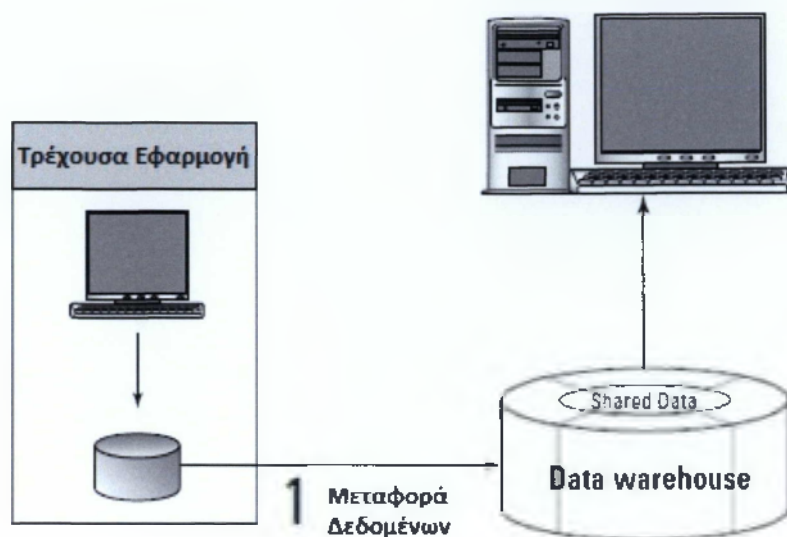
Παράδειγμα επιτυχούς χρήσης

### ΚΕΦΑΛΑΙΟ 3

Στο κεφάλαιο αυτό παρουσιάζεται η δομή ενός Data Warehouse. Η δομή αυτή διαφέρει κατά πολύ ανάμεσα σε κάθε σύστημα Data Warehouse. Το μέγεθος τους μπορεί να χαρακτηριστεί απλοϊκά ως:

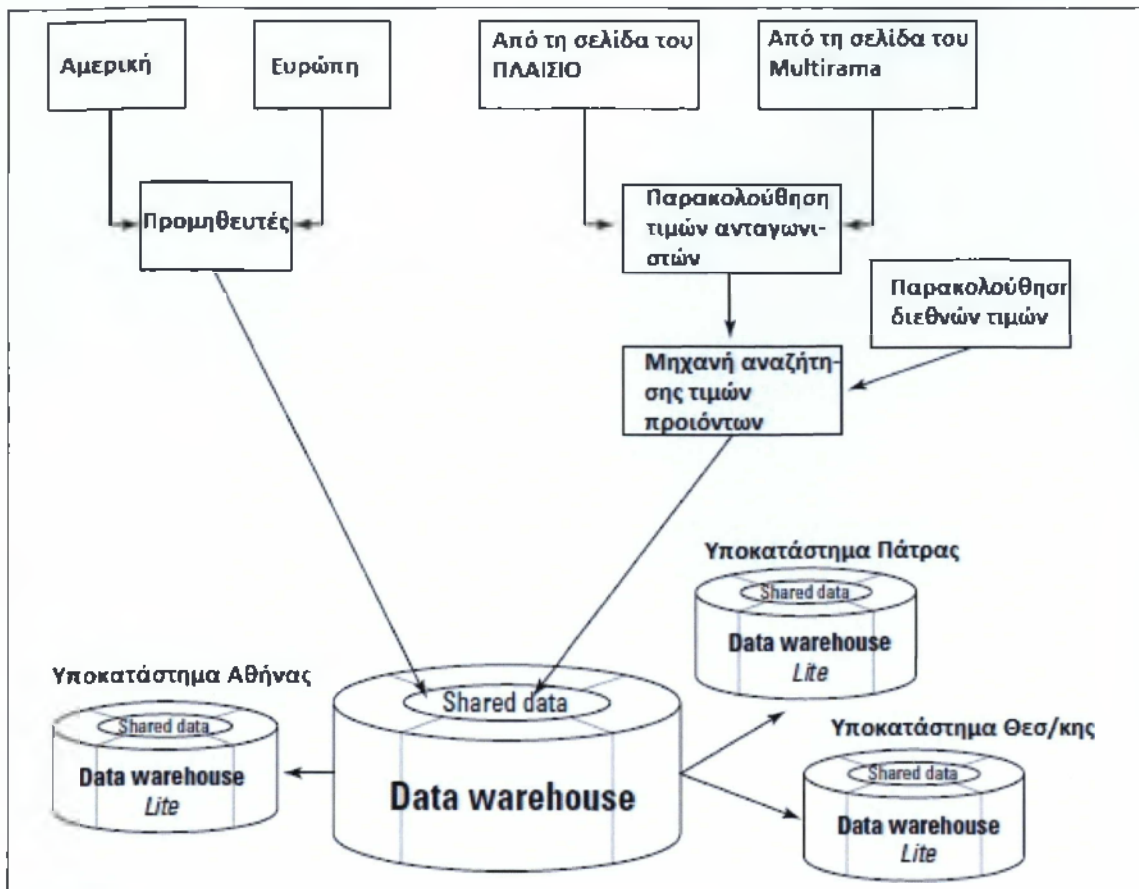
- **Lite:** Όταν μια μικρή ομάδα υλοποιεί το σύστημα χωρίς ιδιαίτερα εξειδικευμένες ανάγκες εξόρυξης δεδομένων.
- **Deluxe:** Όταν αξιοποιούνται σύνθετες τεχνολογίες και ανακτούν δεδομένα από πολλαπλές πηγές.
- **Supreme:** Όταν τα δεδομένα προέρχονται από καταναμημένες πηγές, από ιδιαίτερα συστήματα, και αναλυτικές αναζητήσεις που προϋποθέτουν

υπολογίσιμη επεξεργασία, είναι απαραίτητα για την παρουσίαση των προτεινόμενων στοιχείων για τη βοήθεια στη λήψη αποφάσεων.



Lite Αποθήκη Δεδομένων

Στην **Lite Data Warehouse** τα δεδομένα που δημιουργεί η τρέχουσα εφαρμογή που χρησιμοποιείται, ακόμα και αν είναι ένα απλό Excel αρχείο μιας δημόσιας για παράδειγμα υπηρεσίας όπως των Κέντρων Εξυπηρέτησης Πελατών μεταφέρονται σε μια σχεσιακή (SQL) βάση δεδομένων, ώστε να αξιοποιηθούν μέσω αναζητήσεων.



Delux αποθήκη δεδομένων

Στην **Delux Data Warehouse** η αρχιτεκτονική είναι πιο σύνθετη. Θεωρούμε υποθετικά το Data Warehouse ενός ηλεκτρονικού καταστήματος όπως είναι το e-shop.gr. Τα δεδομένα που καθημερινά πρέπει να προστίθενται είναι από πολλαπλές πηγές. Δηλαδή, τόσο από άλλες βάσεις δεδομένων, όσο και από το Ιντερνετ (παρακολούθηση ανταγωνισμού).

### 3.2 Μάθημα 2<sup>ο</sup> - Τεχνολογίες Data Warehousing

Το δεύτερο μάθημα αναλύει μια πιο λεπτομερή παρουσίαση μικρής κλίμακας αποθηκών δεδομένων που ονομάζονται Data Marts και αποτελούν υποσύνολα των Αποθηκών Δεδομένων, τα οποία είναι πιο εύκολο να συναντήσει κανείς στη χώρα μας. Γίνεται επίσης μια προσπάθεια να συνδεθούν οι υπάρχουσες γνώσεις του μαθητή, όπως για παράδειγμα αυτές για τις σχεσιακές βάσεις δεδομένων και τα middlewares με τις Αποθήκες Δεδομένων.

2



#### Μάθημα 2ο: Τεχνολογίες Data Warehousing

Στο μάθημα αυτό θα παρουσιάσουμε τα Data Marts και τις πιθανές αρχιτεκτονικές πηγές στην κατασκευή αυτών των πληροφοριών. Για να μη μπερδευτείτε τα Data Marts δεν είναι τίποτα άλλο, παρά μικρές εκδοχές των Data Warehouse.

Τέλος παρουσιάζεται ποιές πληροφορίες πρέπει να αποθηκευτούν σε αυτά και πώς τοποθετούνται.

- 4. Εισαγωγή στα Data Marts
- 5. Σχεσιακές Βάσεις Δεδομένων και Αποθήκες Δεδομένων
- 6. Ειδικές Β.Δ. και Αποθήκες Δεδομένων πολλών διαστάσεων (multidimensional Databases)
- 7. Middleware στα Data Warehouses

Εικόνα 3-3 Παρουσίαση 2<sup>ου</sup> μαθήματος

Στο κεφάλαιο 4 παρουσιάζονται τα Data Marts, η κατασκευή τους, η χρήση τους και το είδος των πληροφοριών που περιέχουν. Στο κεφάλαιο 5 παρουσιάζονται οι πιθανές αρχιτεκτονικές των συστημάτων αυτών καθώς και ο ρόλος των σχεσιακών βάσεων δεδομένων όπως επίσης και των συστημάτων OLAP. Στο κεφάλαιο 6 γίνεται μια αναφορά στην ιεράρχηση των δεδομένων και στα οριζόντια και κάθετα συστήματα.

#### ΚΕΦΑΛΑΙΟ 4

Τα Data Marts είναι μια λιγότερο δαπανηρή, εναλλακτική λύση αντί για μια αποθήκη δεδομένων, που απαιτούν πολύ λιγότερο χρόνο και χρήμα. Ένας σχετικά αυστηρός ορισμός είναι ότι, "αποτελεί ένα χώρο δεδομένων, δευτερεύοντα σε σχέση με τις αποθήκες δεδομένων, που περιέχει ταξινομημένο όγκο πληροφοριών".



Το Data Mart επικοινωνεί με ένα **διαμέρισμα** του χώρου των δεδομένων, που ονομάζεται θεματικός τομέας (subject area), ο οποίος δημιουργείται για χρήση από συγκεκριμένους χρήστες. Ένα Data Mart μπορεί να θεωρηθεί ως το ταξινομημένο σύνολο ή ένα υποσύνολο των συγκεντρωτικών δεδομένων. Τις περισσότερες φορές ένα τέτοιο σύνολο συνοδεύει την βάση της αποθήκης δεδομένων, αν και συνήθως εξυπηρετείται από διαφορετικό διακομιστή (server).

Για παράδειγμα σε μια εταιρεία κάθε τμήμα θα εξυπηρετείται από το δικό του data mart, δηλαδή θα υπάρχει μια δομή για το τμήμα πωλήσεων, μια για το τμήμα μηχανοργάνωσης κτλ, που όμως όλες τροφοδοτούνται από τα στοιχεία που βρίσκονται στην αποθήκη δεδομένων. Γενικά οι βάσεις δεδομένων, που υλοποιούνται με αποθήκες δεδομένων, περιέχουν εργαλεία εξόρυξης δεδομένων που εξάγουν σύνολα πληροφοριών ως απαντήσεις σε πολύ συγκεκριμένες ερωτήσεις. Τα Data Marts, ανεξάρτητα από το πλήθος τους ή τις τεχνολογίες που χρησιμοποιούν, είναι ιδιαίτερα βοηθητικά καθώς οι διαφορετικοί χρήστες μπορούν να χρησιμοποιήσουν δεδομένα που προέρχονται από μια κοινή, διαμοιραζόμενη αποθήκη δεδομένων.

Η κατασκευή ενός data mart μπορεί να γίνει με τρεις (3) διαφορετικούς τρόπους:

1. Από δεδομένα που προέρχονται εξ' ολοκλήρου από μια **αποθήκη δεδομένων**.
2. Να αποτελούν μια **ανεξάρτητη δομή** με αυτούσια δεδομένα, χωρίς διασύνδεση με κάποια άλλη δομή.
3. **Ανεξάρτητη δομή** με βλέψεις προς την **ενσωμάτωσή** της σε μια ήδη υπάρχουσα.

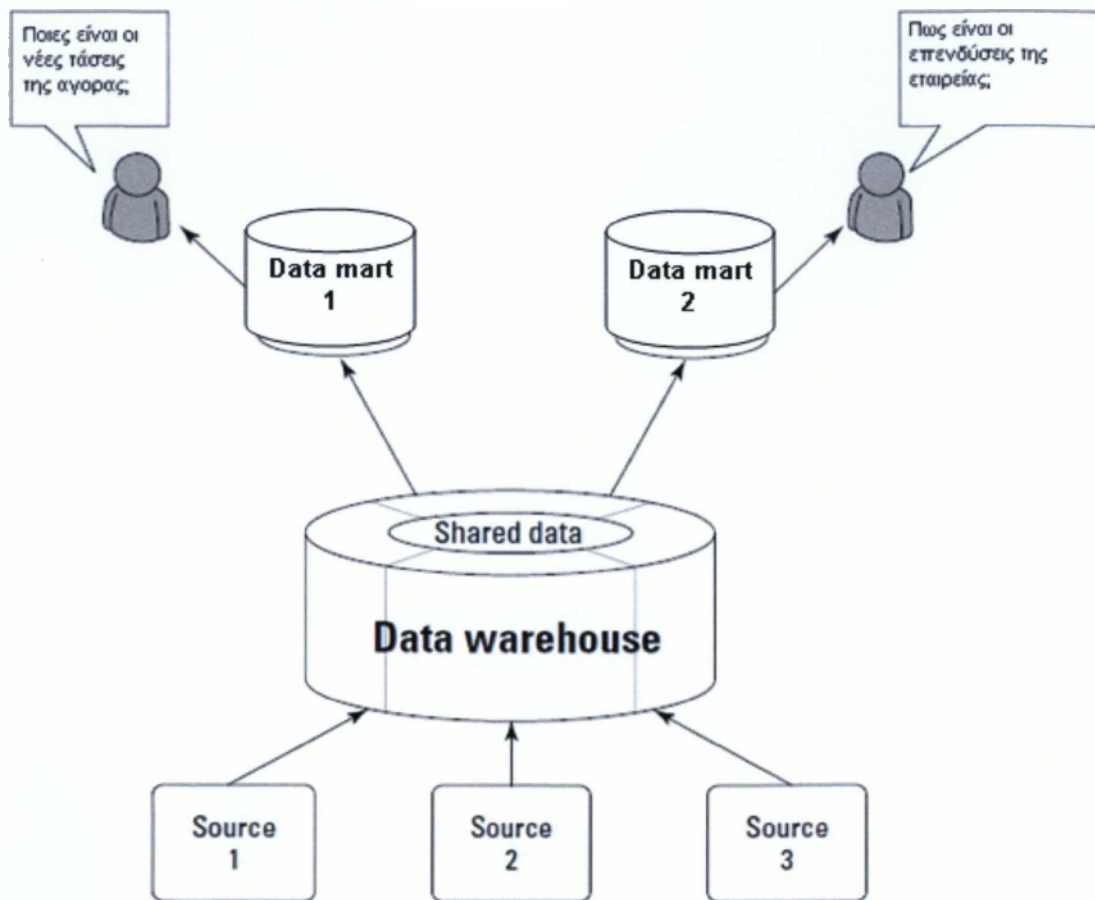
### **Δημιουργία Data Mart από ένα Data Warehouse**

**Βήμα 1ο :** Οι πηγές δεδομένων **στέλνουν** πληροφορίες στην αποθήκη δεδομένων.

**Βήμα 2ο :** Η αποθήκη δεδομένων **συλλέγει** τις πληροφορίες, τις ταξινομεί (ενοποιεί αυτές που έχουν ίδιο περιεχόμενο) και τις διανέμει στις αγορές δεδομένων. Κάθε data mart περιέχει διαφορετικά σύνολα πληροφοριών.

**Βήμα 3ο :** Οι ερωτήσεις που γίνονται στη βάση, δεν υπόκεινται σε **επεξεργασία** από τα Data Warehouses αλλά από τους «λιανέμπορους», δηλαδή τα Data Marts.

**Βήμα 4ο :** Τα Data Marts ζητούν δεδομένα από το Data Warehouse και αφού τα αποθηκεύουν, τα διαθέτουν στους πελάτες.



Δημιουργία Data Marts

Σε περίπτωση που το Data Warehouse δεν περιέχει τα δεδομένα που ζητά το Data Mart, προτείνονται δύο τρόποι αντιμετώπισης:

- Συμπλήρωση της υπολείπουσας πληροφορίας στην αποθήκη δεδομένων και ακολούθως να επιτρέπεται η συναλλαγή τους με τα data marts.
- Προσθήκη των νέων πληροφοριών απ' ευθείας στο data mart. Σε νέα επερώτηση, τα στοιχεία που συλλέγει από την αποθήκη δεδομένων και τα πρόσθετα, ενοποιούνται και δίνονται ως απάντηση στον πελάτη.

Η επιλογή κρίνεται από το είδος των πληροφοριών που προστίθενται, δηλαδή αν παραπάνω από ένα data marts ζητούν τα νέα δεδομένα, τότε είναι οικονομικότερη η πρώτη λύση. Αν μόνο ένα χρειάζεται τη νέα πληροφορία, τότε προτείνεται η δεύτερη λύση.

### **Δημιουργία Data Mart από το μηδέν**

Μέσω αυτής της τεχνικής ουσιαστικά θεωρείται ότι το Data Mart είναι μια **μινιατούρα** ενός Data Warehouse, που ικανοποιεί τους απαιτητικούς πελάτες που χρειάζονται άμεσα τις περιεχόμενες πληροφορίες. Η κατασκευή του ακολουθεί τα ίδια βήματα με εκείνα της κατασκευής μιας αποθήκης δεδομένων που συζητήθηκε σε προηγούμενο μάθημα.

- Εξαγωγή δεδομένων.
- Μετατροπή σε δεδομένα διαχωρίσιμα από την αποθήκη.
- Εξασφάλιση ποιότητας δεδομένων.
- Προσθήκη στην αποθήκη.

Γιατί να επιλέξουμε ένα Data Mart και όχι ένα Data Warehouse;

Γιατί εξοικονομείται ταχύτητα και κόστος, μειώνεται η πολυπλοκότητα (λόγω μικρότερου αριθμού δεδομένων) και συνεπώς η επικινδυνότητα. Η επιλογή εξαρτάται από τον όγκο της πληροφορίας που θέλουμε να διαχειριστούμε, τα προβλήματα που πρέπει να λυθούν και την αξία που θα έχει η ολοκληρωμένη δομή του συστήματος.

### **Δημιουργία Data Mart για μακροπρόθεσμη ενσωμάτωση**

Η δημιουργία μιας τέτοιας δομής είναι εξαιρετικά ανασφαλής και για αυτό το λόγο θεωρείται δύσκολη. Ο σχεδιασμός μπορεί να είναι ορθός και τεκμηριωμένος, όμως υπάρχει περίπτωση να μην είναι δυνατή η σύνδεση με μια αποθήκη δεδομένων.

### **Τι είδους πληροφορίες περιέχει ένα Data Mart;**

Περιέχει: πληροφορίες σχετικές με γεωγραφικές περιοχές, για παράδειγμα τις περιοχές όπου έχει υποκαταστήματα η εταιρία-πελάτης.

- Πληροφορίες σχετικές με την εταιρία, π.χ. τα προϊόντα και οι τιμές τους ή οι λογαριασμοί των πελατών για τράπεζες.
- Λειτουργικά δεδομένα για την εταιρία, όπως τα θέματα μάρκετινγκ που διαχειρίζονται όλα τα υποκαταστήματα της εταιρίας.

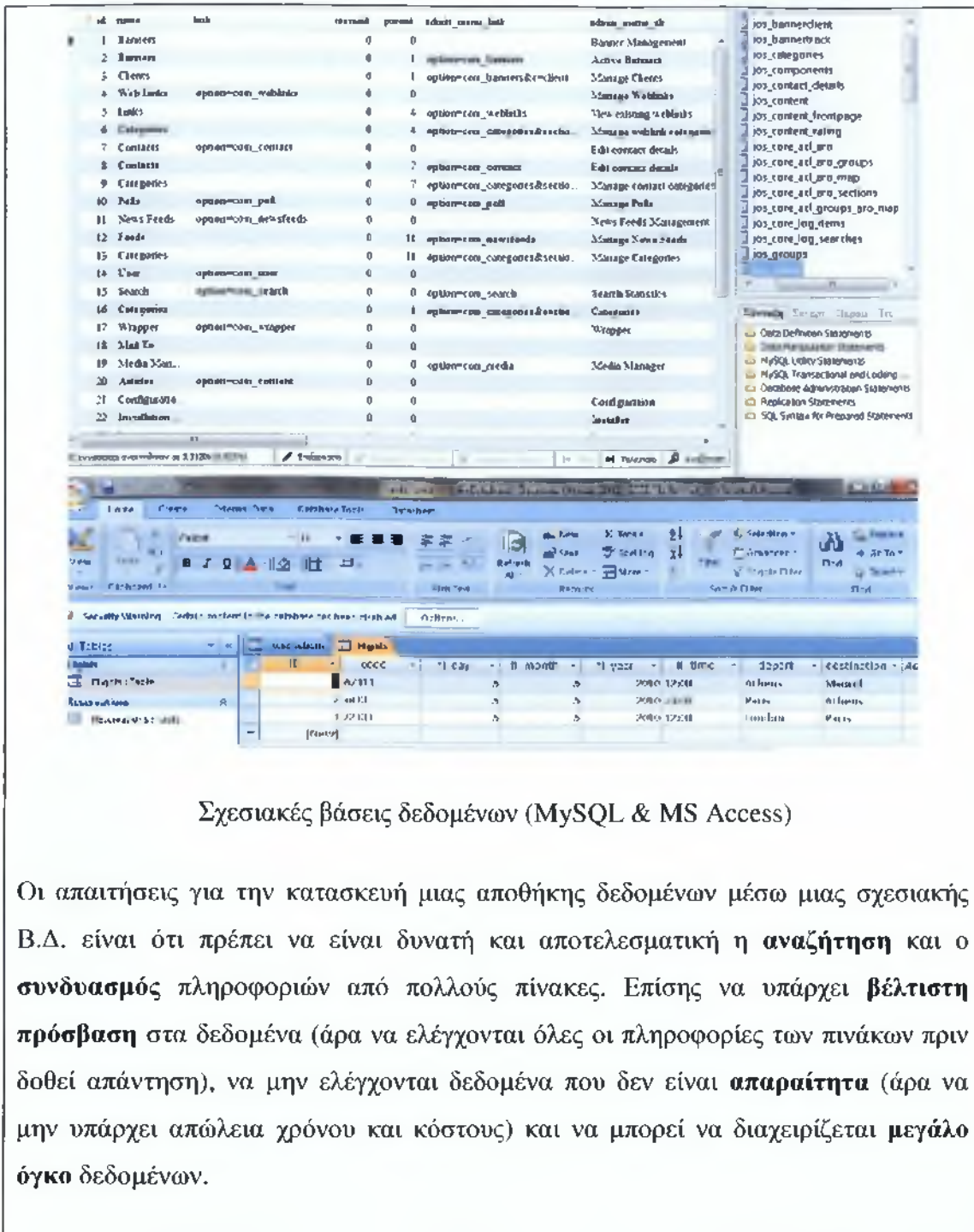
- Πληροφορίες σχετικές με την αγορά που δραστηριοποιείται η εταιρία, εξωτερικά δεδομένα από την ανάλυση της κίνησης της αγοράς, κινήσεις ανταγωνιστών κ.α.
- Απαντήσεις σε συγκεκριμένες ερωτήσεις, συνήθως επαγγελματικού χαρακτήρα.
- Οτιδήποτε θα κάνει τη λειτουργία όλης της δομής περισσότερο λειτουργική και αποτελεσματική.

#### **Μερικές Συμβουλές !**

- Ακολουθείστε μια μεθοδολογία βασισμένη σε σαφή βήματα και φάσεις του έργου (έτσι μπορεί να γίνεται καλύτερος έλεγχος σφαλμάτων).
- Καθορίστε εκ των προτέρων το χρόνο που απαιτείται για κάθε φάση.
- Μην παρεκκλίνετε σε καμία περίπτωση από το στόχο σας για να προσθέσετε κάτι νέο. Ολοκληρώστε το Data Mart και αναρωτηθείτε μετά αν χρειάζεται να γίνει κάποια προσθήκη.

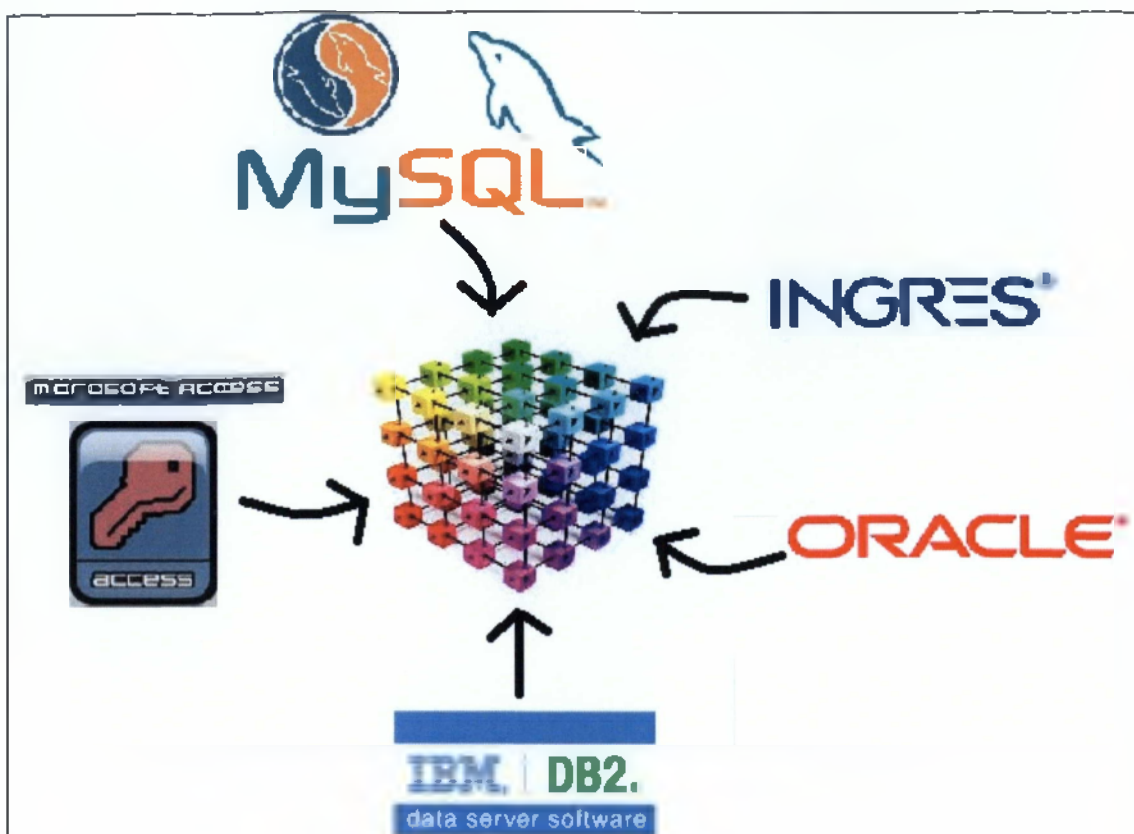
#### **ΚΕΦΑΛΑΙΟ 5**

Μια σχεσιακή Β.Δ. περιέχει την πληροφορία καταχωρημένη σε πίνακες που κάθε κελί αντιστοιχίζεται από μια σειρά και στήλη, τα οποία έχουν συγκεκριμένο χαρακτηριστικό όνομα. Οι πίνακες συσχετίζονται μεταξύ τους δημιουργώντας αντιστοιχίες και στα δεδομένα που περιέχουν. Η βασική ιδιότητα της σχεσιακής Β.Δ., είναι ότι μπορεί να απαντήσει σε πληθώρα ερωτήσεων, διότι οι πίνακες μπορούν να συνδυαστούν αναλόγως με το επιθυμητό αποτέλεσμα. Αυτά τα χαρακτηριστικά χρησιμοποιεί και μια αποθήκη δεδομένων.



Σχεσιακές βάσεις δεδομένων (MySQL & MS Access)

Οι απαιτήσεις για την κατασκευή μιας αποθήκης δεδομένων μέσω μιας σχεσιακής Β.Δ. είναι ότι πρέπει να είναι δυνατή και αποτελεσματική η αναζήτηση και ο συνδυασμός πληροφοριών από πολλούς πίνακες. Επίσης να υπάρχει βέλτιστη πρόσβαση στα δεδομένα (άρα να ελέγχονται όλες οι πληροφορίες των πινάκων πριν δοθεί απάντηση), να μην ελέγχονται δεδομένα που δεν είναι απαραίτητα (άρα να μην υπάρχει απώλεια χρόνου και κόστους) και να μπορεί να διαχειρίζεται μεγάλο όγκο δεδομένων.



Όλες οι σχεσιακές βάσεις μπορούν να τροφοδοτήσουν μια αποθήκη δεδομένων

Οι σχεσιακές Β.Δ. έχουν επίσης ένα χαρακτηριστικό το οποίο είναι μειονέκτημα για την κατασκευή Αποθηκών Δεδομένων. Οι σχεσιακές Β.Δ. είναι η κανονικοποιημένη μορφή όλων των αρχικών πινάκων, γεγονός που προσφέρει μεγάλη διευκόλυνση κατά την τροποποίηση των δεδομένων. Παρόλα αυτά σε μια αποθήκη δεδομένων η τροποποίησή τους, τουλάχιστον σε μοναδικές πλειάδες, δεν αποτελεί επιλογή. Η νέα προσέγγιση ταχύτερης απόκτησης δεδομένων σε μια Β.Δ. ονομάζεται OLAP (Online Analytical Process) και μπορεί να εφαρμοστεί στις σχεσιακές Β.Δ. (Relational OLAP ROLAP).

Μερικά προτερήματα παρουσιάζονται ακολούθως:

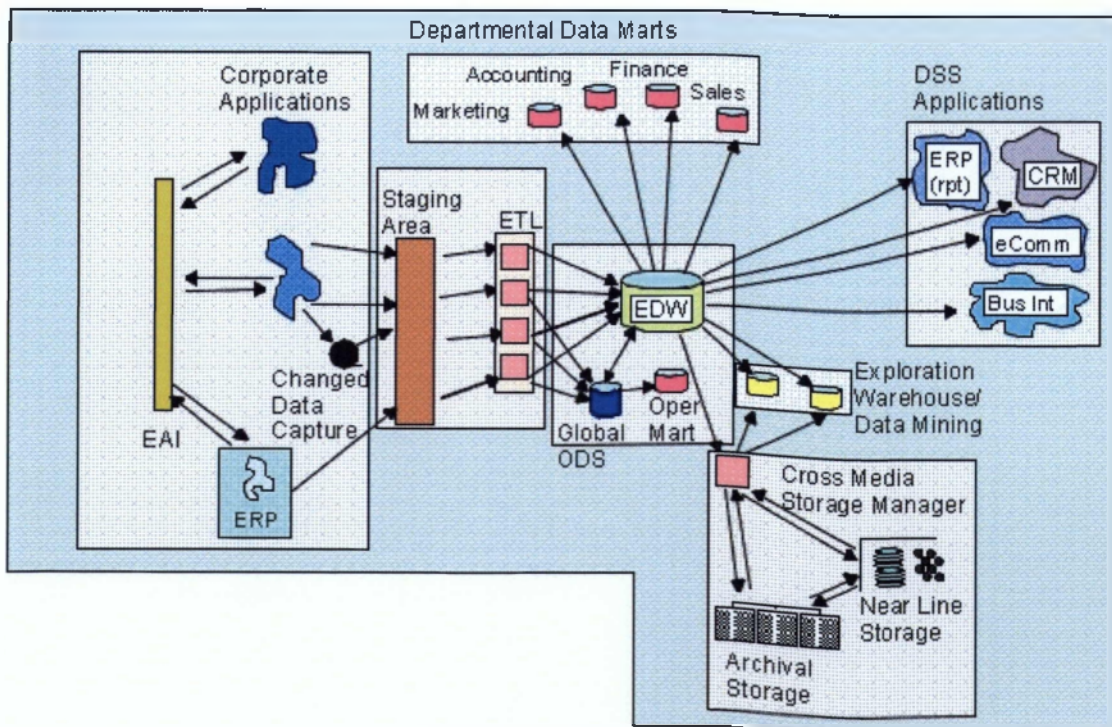
- Σύνθετα ερωτήματα για τα οποία γίνεται έλεγχος σε εκατομμύρια αρχεία.
- Σαρώσεις μεγάλης διάρκειας που μπορούν να λειτουργούν παράλληλα με τις συνηθισμένες λειτουργίες.
- Ανάγκη ιστορικών στοιχείων για την ανάλυση των νέων τάσεων.

- Μικρός χρόνος απόκρισης των Επερωτήσεων.

Άλλοι προτεινόμενοι τρόποι σχεδιασμού ενός Data Warehouse σε σχεσιακή Β.Δ.

- Corporate Information Factory (CIF): Περιέχει τρία είδη αρχιτεκτονικής δεδομένων που μεταξύ άλλων επιτρέπουν τα Data Warehouse, Data Marts, Operational Data Store.
- Star Schema: είναι το πιο απλό σχήμα οργάνωσης μιας αποθήκης δεδομένων. Το Star Schema αποτελείται από λίγους πίνακες γεγονότων (πιθανώς μόνο ενός) που συσχετίζεται με κάθε πίνακα διάστασης. Οι πίνακες που περιέχει μπορεί να μην είναι κανονικοποιημένοι.
- Data Vault: Αρχιτεκτονική που βασίζεται στην παρουσίαση των λεπτομερειών κάθε κανονικοποιημένου πίνακα, παρουσιάζοντας μοναδικούς δεσμούς, μεταξύ των ήδη υπάρχοντων και των νέων πινάκων «δορυφόρων».

### Corporate Information Factory



Ο εξοπλισμός που συνήθως συναντάται σε συστήματα CIF είναι Compaq, EMC, H-P, IBM, NCR και Sun. Τα συστήματα σχεσιακών βάσεων δεδομένων συνήθως στηρίζονται σε υλοποιήσεις IBM DB2, Oracle, MS SQL Server και Teradata.

Σχετικά με τα συστήματα CRM και ERP που συνήθως χρησιμοποιούνται από μεγάλους οργανισμούς, μπορεί κάποιος να συναντήσει τα: Ascential, Business Objects, Computer Associates, Cognos, Crystal Decisions, DataFlux, Hyperion, Informatica, Kalido, PeopleSoft, SAS και πολλά άλλα.

Για τη διαχείριση ενός τέτοιου συστήματος απαιτούνται γνώσεις σε μια πλειάδα τεχνολογιών, όπως τεχνολογίες OLTP, Data warehousing, επεξεργασία στατιστικών, data base management τεραστίων βάσεων, Archival data management, παράλληλη επεξεργασία (Parallel processing), ασφάλεια πληροφοριακών συστημάτων, διαχείριση καταναμημένων metadata (Distributed metadata management), Data quality management, Reporting κ.α.

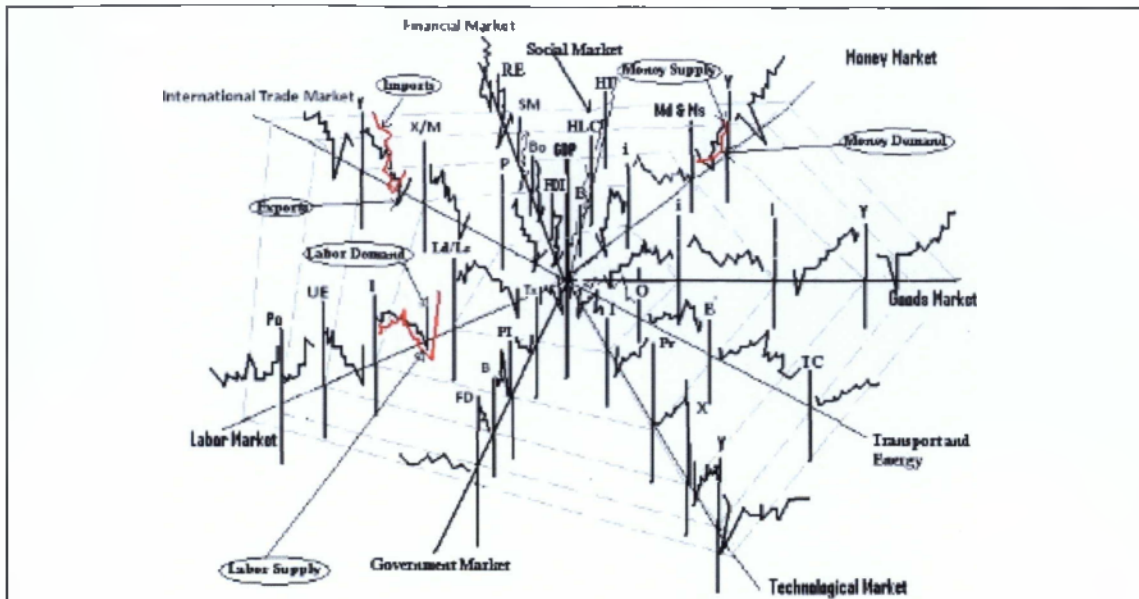
### **Μερικές Συμβουλές !**

Επιλέξτε την αρχιτεκτονική που θα χρησιμοποιήσετε προσεκτικά, σύμφωνα με τις απαιτήσεις σας και την επιλογή του εργαλείου, σταθμίζοντας όλους τους παράγοντες. Δεν υπάρχει κάποια που να θεωρείται καλύτερη από την άλλη. Με βάση τις απαιτήσεις σας και την επιλογή του εργαλείου, μια τεχνική ταιριάζει καλύτερα από όλες στις απαιτήσεις σας.

## **ΚΕΦΑΛΑΙΟ 6**

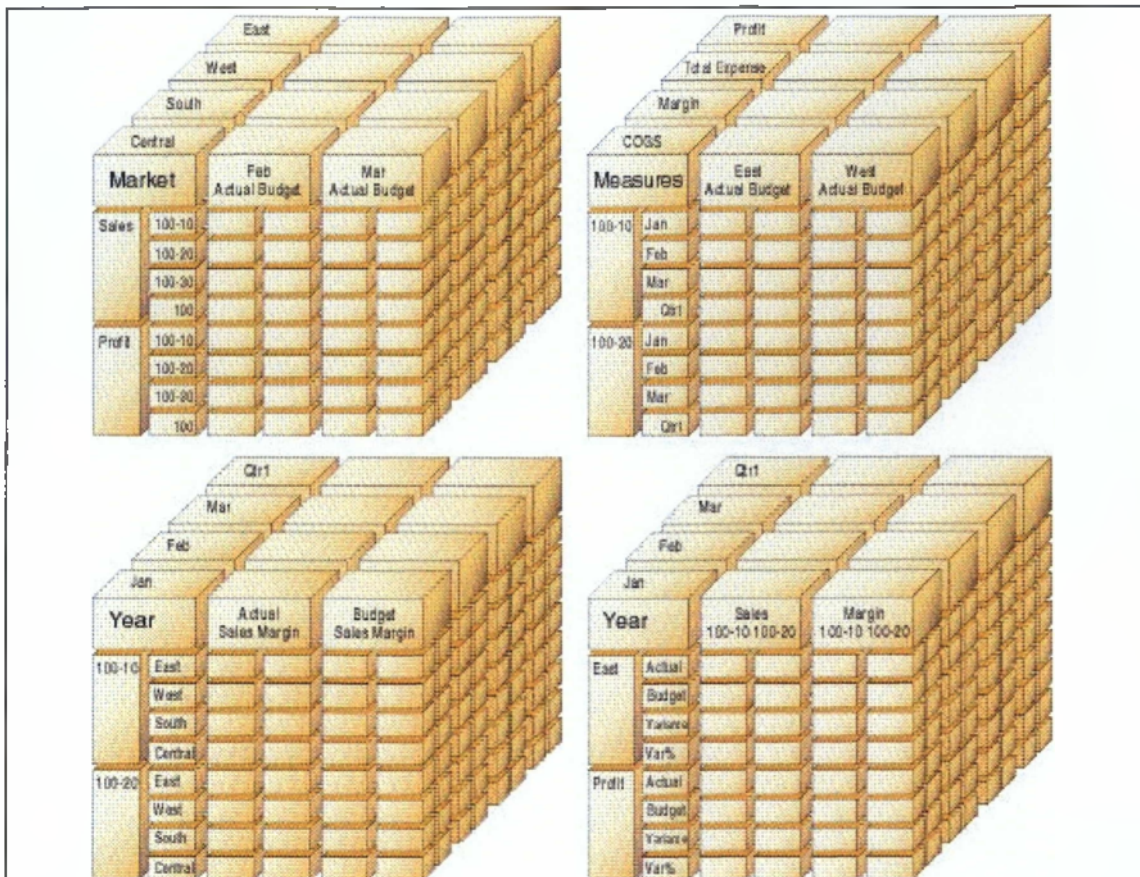
Σε αυτή την προσέγγιση η Β.Δ. θεωρείται το σύνολο των δεδομένων σε ένα χώρο πολλών διαστάσεων. Κάθε γεγονός έχει μετρήσιμες διαστάσεις, για παράδειγμα η ποσότητα αναλύεται σε ποσοστό πωλήσεων, προϋπολογισμό κ.α. Για κάθε γεγονός υπάρχει ένα σύνολο διαστάσεων που το καθορίζει. Οι διαστάσεις αποτελούν ένα αραιό σύστημα συντεταγμένων και κάθε διάσταση έχει ένα σύνολο χαρακτηριστικών. Τα χαρακτηριστικά διαχωρίζονται ανάλογα με την ιεραρχία τους, διευκολύνοντας τη βάση να απαντήσει ευκολότερα σε ερωτήσεις. Οριζόντια Συστήματα Διαχείρισης Δεδομένων είναι τα περισσότερα συστήματα διαχείρισης σχεσιακών Β.Δ. που έχουν οριζόντια μορφή, δηλαδή κάθε φορά που γίνεται μια συνδιαλλαγή, όλα τα δεδομένα τοποθετούνται ανά σειρά.





Δισδιάστατη αναπαράσταση πολυδιάστατων δεδομένων

Έχουν δημιουργηθεί πολλά προϊόντα που ακολουθούν αυτή τη προσέγγιση, ώστε να βελτιωθεί η απόδοση επεξεργασίας των ερωτήσεων. Ο στόχος αυτών των Β.Δ. είναι να αυξήσουν την ταχύτητα της λήψης αποφάσεων υπό επεξεργασία μεγάλου όγκου δεδομένων. Όλα τα **οριζόντια συστήματα** παρέχουν γρήγορο online transaction processing (OLTP) γιατί οι περισσότερες συναλλαγές προκύπτουν από τη χρήση ευρετηρίων (index). Τα ευρετήρια «δείχνουν» στην κορυφή της γραμμής ώστε να απλοποιούν και να επιταχύνουν την ανάκτηση δεδομένων από κοινές διαδρομές.



Αναπαράσταση πέντε διαστάσεων δεδομένων

Στα **κάθετα συστήματα** διαχείρισης δεδομένων η αποθήκευση των δεδομένων γίνεται ανά στήλη. Αυτή η μέθοδος αποθήκευσης επιλύει αποτελεσματικά το πρόβλημα των επερωτήσεων σε μεγάλης κλίμακας δεδομένα, αφού συνήθως ο χρήστης απλά αναζητά μερικές στήλες σε αντιπαραβολή με το πλήθος των στηλών που διαχειρίζεται το οριζόντιο σύστημα δεδομένων. Η τεχνική αυτή δεν είναι επικρατέστερη καθώς συνοδεύεται με παραπάνω δυσχέρειες λόγω των πρόσθετων τεχνολογιών που χρειάζονται για την προτυποποίησή τους.

Οι **εφαρμογές Αποθηκών Δεδομένων** (Data Warehouse Appliances) είναι ένα σύνολο ενσωματωμένων εξυπηρετητών, αποθηκευτικών και λειτουργικών συστημάτων, συστημάτων διαχείρισης Β.Δ. και λογισμικού, προεγκατεστημένο και προ- παραμετροποιημένο για τη δημιουργία αποθηκών δεδομένων. Παρέχουν λύσεις για τη διαχείριση μεγάλου όγκου δεδομένων σε αποδόσεις ελαχίστου κόστους.

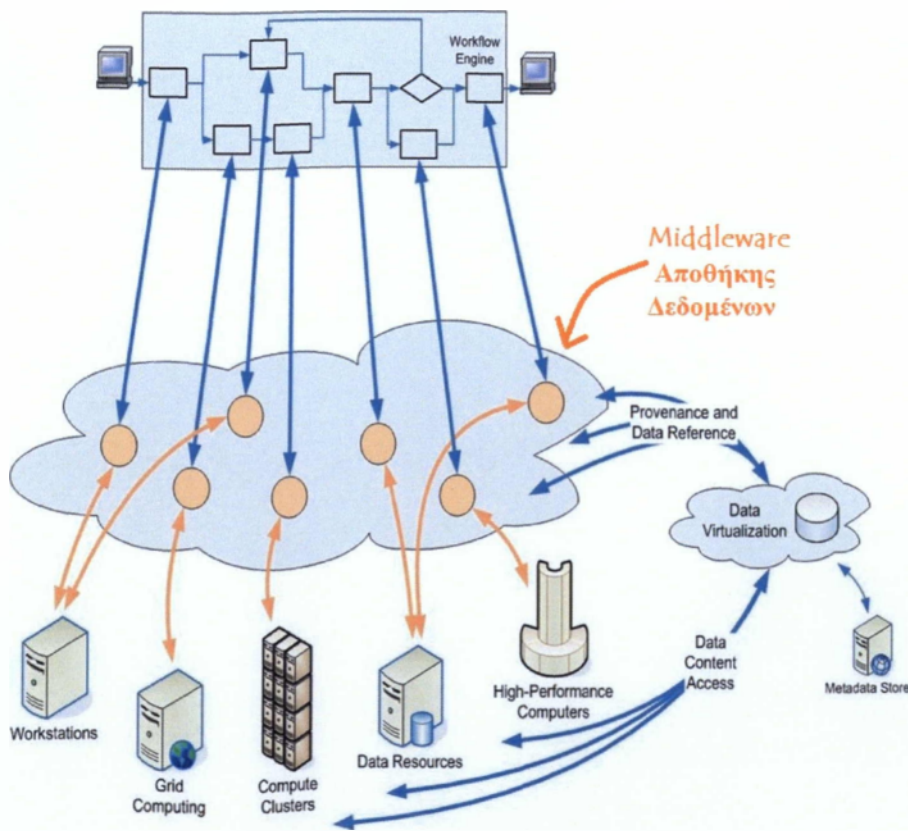
Ορισμένες από τις διαθέσιμες εφαρμογές για πολυδιάστατες βάσεις δεδομένων αποτελούν οι:

- [Microstrategy](#). Είναι εντοπιστικό ότι παρέχει clients για PDAs και συσκευές BlackBerry και iPhone.
- [RedBrick](#) της IBM.
- [EssBase](#) της Oracle.

και πολλές άλλες. Αφιερώστε λίγο χρόνο εξερευνώντας τις παραπάνω ιστοσελίδες !

## ΚΕΦΑΛΑΙΟ 7

Το λογισμικό **Middleware** αποτελεί ένα σύνολο υπηρεσιών που εκτελούν πληθώρα λειτουργιών σε ένα υπολογιστικό περιβάλλον. Οι πιο βασικές υπηρεσίες αφορούν την ασφάλεια, τη διαχείριση των συνδιαλλαγών, της ροής μηνυμάτων, των εφαρμογών, των αιτήσεων από τον Ιστό και την παροχή ευρετηρίων μονοπατιών.



Middleware σε Data Warehouses

Συγκεκριμένα για τις Αποθήκες Δεδομένων παρέχουν:

- Συλλογή και εξαγωγή δεδομένων: Πραγματοποιείται με δύο τρόπους.
  - α) Με εύρεση και εξαγωγή όλων των δεδομένων από την πηγή χωρίς να γίνεται ταξινόμηση.
  - β) Με εύρεση και εξαγωγή μόνο των δεδομένων που δεν υπάρχουν .
- Διαχείριση Ποιότητας (στο επίπεδο των συστατικών μερών): Η διόρθωση των σφαλμάτων γίνεται όσο το δυνατό νωρίτερα .
- Μεταφορά Δεδομένων (στο επίπεδο των συστατικών μερών): μεταφορά των δεδομένων μετά τη διόρθωση σε περιβάλλον που θα γίνουν παραπάνω μετασχηματισμοί.
- Χαρτογράφηση και μετασχηματισμός των δεδομένων: Ουσιαστικά είναι η ομαδοποίηση και ταξινόμηση των δεδομένων μετά τον έλεγχο ποιότητας. Συνήθως παρουσιάζονται προβλήματα ασυμβατότητας μεταξύ συστημάτων οπότε προτείνονται οι ακόλουθες λύσεις.
  - α) Σύνοψη δεδομένων
  - β) Επιλεκτική ένταξη των δεδομένων
  - γ) Σύγκλιση των δεδομένων.
- Διαχείριση ποιότητας μετά το μετασχηματισμό: Έλεγχος Σφαλμάτων.
- Μετακίνηση των δεδομένων στο περιβάλλον του Data Warehouse.
- Εισαγωγή των δεδομένων στο Warehouse: αν χρειάζεται να εισαχθούν δεδομένα μεγάλου όγκου, χρησιμοποιείστε ένα εργαλείο που απαιτεί λιγότερο χρόνο.

### **3.3 Μάθημα 3<sup>ο</sup> – Business Intelligence στα Data Warehouses**

Στο τρίτο μάθημα η στρατηγική διαφοροποιείται αρκετά. Στα πρώτα δύο μαθήματα έγινε μόνο η θεωρητική παρουσίαση των πληροφοριών. Η θεματική που παρουσιάζει το 3<sup>ο</sup> μάθημα δίνει την ευκαιρία για πρακτική εφαρμογή και παρουσίαση πακέτων λογισμικού, αποτελέσματα υποβολής ερωτήσεων και εκθέσεων, της διαδικασίας εξόρυξης δεδομένων (data mining). Το μάθημα ξεκινά με την εισαγωγή στην έννοια “Business Intelligence” και στη συνέχεια μέσω μικρών κεφαλαίων κάνει μια

εισαγωγή και παρουσίαση σε κάθε επιμέρους έννοια προσπαθώντας να τις παρουσιάσει σε όσο το δυνατόν πιο πρακτικό επίπεδο.

3



### Μάθημα 3ο: Business Intelligence στα Data Warehouse

Στο μάθημα αυτό θα παρουσιάσουμε εισαγωγικούς όρους για Business Intelligence, Data mining, δημιουργία Reports, υποβολή ερωτημάτων και OLAP διαδικασιών.

- 8. Τι είναι το 'Business Intelligence'
- 9. Υποβολή Επερωτήσεων και Εκθέσεων
- 10. Ανάλυση διαδικασιών (business analysis)
- 11. Εισαγωγή στο Data Mining
- 12. Dashboards, Scoreboards και σενάρια OLAP

Εικόνα 3-4 Παρουσίαση 3<sup>ου</sup> μαθήματος

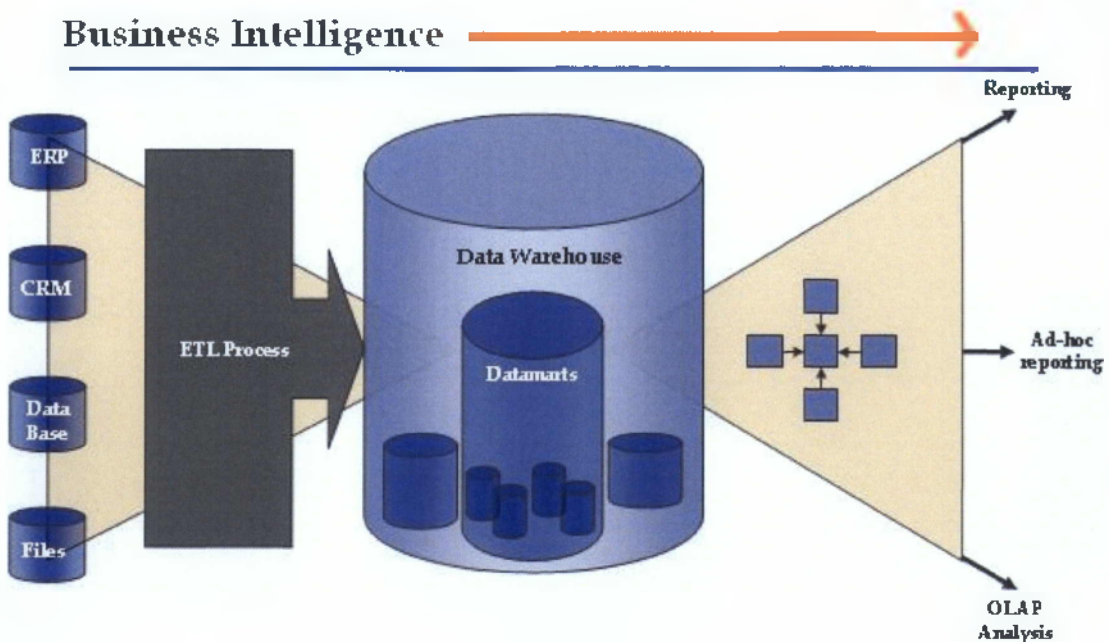
## ΚΕΦΑΛΑΙΟ 8

Ο όρος Business Intelligence περιλαμβάνει τις δεξιότητες, τις διαδικασίες, τις τεχνολογίες, τις εφαρμογές και τις πρακτικές που χρησιμοποιούνται για να ενισχύσουν τη λήψη αποφάσεων. Χωρίζεται στις εξής κατηγορίες:

1. Επερωτήσεις και υποβολή εκθέσεων: Τα δεδομένα ανακτώνται είτε , ως απαντήσεις σε συγκεκριμένες επερωτήσεις που θέτει ο πελάτης, είτε ως πρότυπες εκθέσεις.
2. Ανάλυση διαδικασιών (business analysis): ο όρος που χρησιμοποιείται για την περιγραφή απεικόνισης των δεδομένων με πολλές διαστάσεις. Αυτό διευκολύνει το χρήστη να σχεδιάσει τα δεδομένα σε σειρές και στήλες για την καλύτερη κατανόηση των σημείων σύγκλισης (OLAP).
3. Εξόρυξη δεδομένων: η τεχνική εξαγωγής προτύπων από δεδομένα. Ο στόχος της είναι η πρόβλεψη με βάση τις πιθανότητες και τις στατιστικές αναλύσεις, ή η ανακάλυψη παρέχοντας απαντήσεις ακόμα και σε περιπτώσεις που δεν τίθενται ερωτήσεις.

4. Dashboards and Scorecards: Η έννοια του dashboard (πίνακας οργάνων) απεικονίζει την παρούσα κατάσταση της πληροφορίας που είναι διαθέσιμη για τη λειτουργική αποδοτικότητα του συστήματος ενώ scorecard (πίνακας αποτελεσμάτων) θεωρείται η αποδοτικότητα του συστήματος όταν αυτό λειτουργεί υπό συγκεκριμένο στόχο και συνθήκες. Τα αποτελέσματα αυτά τροφοδοτούν το σχεδιαστή με πληροφορίες για την κατάσταση του συστήματος και προκύπτουν από τις δύο πρώτες φάσεις της διαδικασίας.

Όλα τα βήματα θα αναλυθούν πιο διεξοδικά στα κεφάλαια που ακολουθούν.



Τα παραπάνω βήματα είναι αυτά που ακολουθεί το μοντέλο του Business Intelligence. Είναι σύνηθες φαινόμενο να παρατηρούνται συμπληρωματικά όπως:

1. Στατιστική Επεξεργασία: Υπολογισμός στατιστικών μεγεθών όπως σύνολα, μέσος όρος, μέγιστες και ελάχιστες τιμές, διακύμανση και μέση τιμή. Πολύ συχνά η στατιστική λειτουργικότητα ενσωματώνεται ως χαρακτηριστική λειτουργία των εργαλείων business intelligence.
2. Γεωγραφικά Πληροφοριακά Συστήματα: Είναι οποιοδήποτε σύστημα που συλλαμβάνει, αποθηκεύει, αναλύει, διαχειρίζεται, και παρουσιάζει στοιχεία που συνδέονται με την τοποθεσία. Η διαχείριση της πληροφορίας δε γίνεται με βάση κάποια ιεραρχία.

3. Mash-ups: Η ικανότητα ανάμειξης των δεδομένων και των υπηρεσιών.

Πως συνδέεται η αρχιτεκτονική Business Intelligence με τα Data Warehouses?

Αν χρησιμοποιηθεί η αρχιτεκτονική Business Intelligence για την κατασκευή Data Warehouse, τα προτερήματα σε ότι αφορά στην λειτουργικότητα είναι τα εξής:

- Λειτουργικότητα των εξυπηρετητών (καταμερισμός διεργασιών).
- Λειτουργικότητα μέσω του Ιστού.
- Υποστήριξη απομακρυσμένων χρηστών.
- Περιβάλλον πραγματικού χρόνου.
- Τεχνολογία αντιπροσώπων (πρόσθετοι μηχανισμοί).

## ΚΕΦΑΛΑΙΟ 9

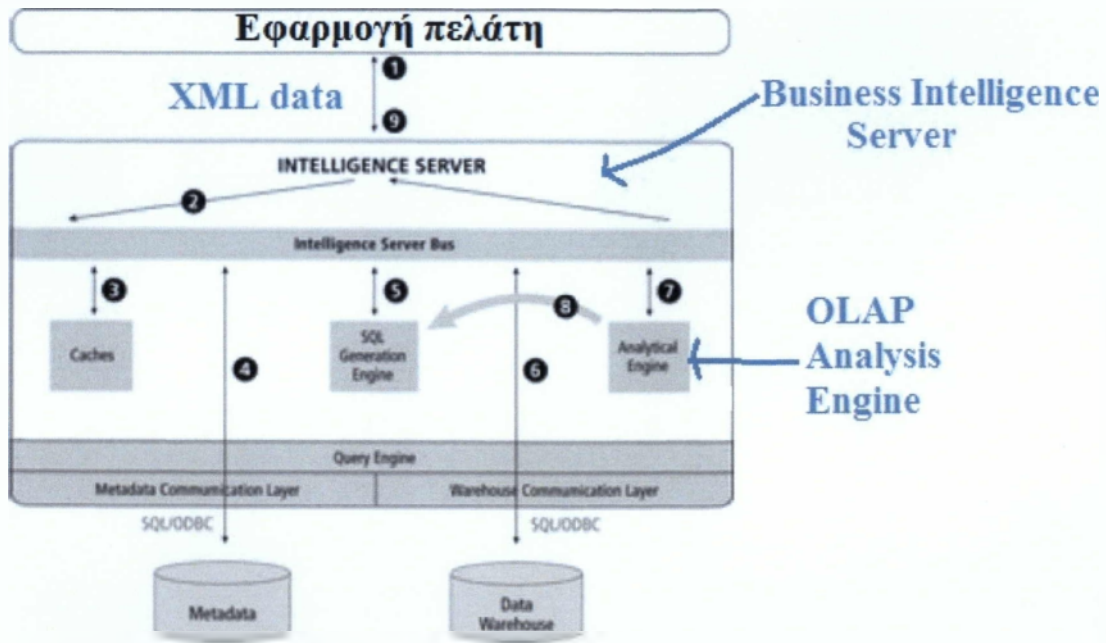
Η διαδικασία θέσης επερωτήσεων και υποβολής εκθέσεων πραγματοποιείται σε συγκεκριμένα βήματα:

1. Αίτηση για δεδομένα
2. Ανάκτηση δεδομένων
3. Διαχείριση δεδομένων (αναδιάταξη ή σύνοψη)
4. Διάταξη δεδομένων.
5. Παρουσίαση Δεδομένων

Ας δούμε την πορεία ενός ερωτήματος που ξεκινά από την εφαρμογή του πελάτη ή από ένα Web Interface. Το query μεταφέρεται ως ένα XML μήνυμα στον Intelligence Server. Το ερώτημα αρχικά κατευθύνεται στο Cache engine. Όταν ένα ερώτημα εκτελείται πολύ συχνά τότε για ταχύτερη απάντηση σε αυτό, τα απαραίτητα δεδομένα αποθηκεύονται στη μνήμη. Έτσι είναι πιθανότερο η απάντηση στο ερώτημα να έρθει ταχύτερα από εκεί. Ταυτόχρονα ο Intelligence Server επικοινωνεί με τα MetaData.

Στο βήμα (5) με τη βοήθεια των πληροφοριών από το Cache engine και τα μεταδεδομένα, δημιουργείται το SQL Query. Το ερώτημα του χρήστη συνήθως είναι λεκτικό π.χ. "Κατανάλωση ρεύματος Βοιωτία". Το SQL ερώτημα όμως θα είναι σύνθετο και μακροσκελές. Στο βήμα (6) το SQL ερώτημα εκτελείται στην αποθήκη δεδομένων. Οι πληροφορίες τροφοδοτούνται στην OLAP Analysis Engine (7) η οποία πιθανώς να δημιουργήσει επιπρόσθετα ερωτήματα (8).

Οι τελικές πληροφορίες και τα δεδομένα είναι πλέον έτοιμα να παρουσιαστούν στον πελάτη σε καθαρή μορφή, κύβους δεδομένων ή αναλυτικές εκθέσεις.



Η πορεία ενός Query σε μια Αποθήκη Δεδομένων

Οι δυνατότητες και οι λειτουργίες των εργαλείων υποβολής επερωτήσεων και εκθέσεων είναι:

- α) Υποβάλουν συχνές εκθέσεις.
- β) Δημιουργούν ταξινομημένες λίστες.
- γ) Εκτελούν επερωτήσεις συνένωσης πινάκων (cross-tabular) και παράγουν εκθέσεις.

Η σπουδαιότητα της γλώσσας SQL για την υποβολή επερωτήσεων και εκθέσεων είναι ότι έχει αναπαραστήσει έναν καθορισμένο τρόπο πρόσβασης σχεσιακών Β.Δ. και έχει υλοποιήσει πολλά εργαλεία, όπως τεχνικά εργαλεία επερωτήσεων (σύνδεση ερωτήσεων με τη βάση), εργαλεία επερωτήσεων για τους χρήστες και εργαλεία υποβολής εκθέσεων. Αυτή η λειτουργία υποστηρίζεται πλήρως και από OLAP.



## ΚΕΦΑΛΑΙΟ 10

Όταν μιλάμε για Business Analysis τότε η τεχνική OLAP θα πρέπει να είναι το πρώτο πράγμα που μας έρχεται στο μυαλό. Επίσης το OLAP είναι άρρηκτα συνδεδεμένο με την έννοια της αποθήκης δεδομένων. Σε αυτό το κεφάλαιο θα παρουσιαστούν οι τεχνικές OLAP και ο τρόπος που χρησιμοποιούνται σε ένα Data Warehouse. Οι Β.Δ. OLAP είναι οι φυσικά διαμορφωμένοι χώροι για πολυδιάστατες Β.Δ.

**MOLAP** (multidimensional OLAP): Αποθήκες δεδομένων που έχουν κατασκευαστεί σε Β.Δ. πολλών διαστάσεων, ειδικά σχεδιασμένες για να υποστηρίζουν αυτό το είδος επεξεργασίας. **ROLAP** (relational OLAP): Η χρήση σχεσιακών Β.Δ. για την υλοποίηση αποθηκών δεδομένων. **HOLAP** (hybrid OLAP): Ο συγκερασμός των δύο παραπάνω. **DOLAP** (desktop OLAP): Περιβάλλοντα στα οποία ο χρήστης είναι ο πρωταρχικός αποθηκευτικός χώρος για μια πολυδιάστατη δομή.

### Μερικές Συμβουλές !

- Αν επιλέξετε την αρχιτεκτονική του MOLAP ξανασκεφτείτε το και συγκρίνετε τις παραμέτρους με τη ROLAP ειδικά αν ο χρόνος απόκρισης του συστήματος είναι η προτεραιότητά σας.
- Μερικά πακέτα χρηματοοικονομικής ανάλυσης υλοποιούνται μόνο σε MOLAP μόνο.
- Αν η εμβάθυνση σε λεπτομέρειες είναι ο βασικός σας στόχος τότε σκεφτείτε την τεχνική ROLAP (ή την HOLAP).
- Πάντα πριν από κάθε απόφαση σκεφτείτε όλες τις παραμέτρους και τα κρυφά σημεία σε κάθε τεχνική.

Χαρακτηριστικά του OLAP Business Analysis:

**1. Drill Down:** Είναι η εμφάνιση, κατά το δυνατό, του μεγαλύτερου επιπέδου λεπτομέρειας των δεδομένων:

**Customer Listing**

Name	Category	Active	Contact Name	Telephone	Balance
Mountainview Bicycles	General Distributors	True	John Smith	+27 11 555-3039	R 821,774.88
Mountain Goals	Mountain Bike Distributors	True	Janey Smith	+27 21 555-3038	R 387,333.73
Roadbikes Unlimited	Road Bike Distributors	True	Mike Brown	+27 11 555-3030	R 1,001,445.78
Roadbikes Haven	Road Bike Distributors	True	Roger Smith	+27 11 300-7783	R 687,493.28

**Statement**

Roadbikes Unlimited  
P.O. Box 230386  
Johannesburg  
2000

Date: 31/03/2009  
Page: 1

Date	Reference	Description	Debit	Credit
27/04/2008	INV000008	Tax Invoice	R 82,840.03	
11/04/2009	INV000040	Tax Invoice	R 245,784.08	
18/05/2009	INV000047	Tax Invoice	R 28,500.08	

**Tax Invoice**

Roadbikes Unlimited  
P.O. Box 230386  
Johannesburg  
2000

Number: INV000047  
Date: 18/05/2009  
Page: 1/1  
Reference: 38/08/2009  
Overall Discount %: 0.80%

Description	Quantity	Unit Price	Disc %	Tax %	Exclusive Total	Inclusive Total
Carbon Sole Shoes	1.00	R 2,400.00	0.00%	14.00%	R 2,400.00	R 2,736.00
Carbon Road Frame	1.00	R 20,000.00	0.00%	14.00%	R 20,000.00	R 22,800.00
Bicycle Building	1.00	R 200.00	0.00%	14.00%	R 200.00	R 228.00
Time Trial Helmet	2.00	R 1,200.00	0.00%	14.00%	R 2,400.00	R 2,736.00

Total Discount: R 0.00  
Total Exclusive: R 25,000.00  
Total Tax: R 3,580.00  
Sub Total: R 28,580.00  
Total Rounding: R 0.00  
Total: R 28,580.00

Στην παραπάνω εικόνα μπορεί κανείς να παρατηρήσει στην πράξη την λειτουργία drill down. Μια τράπεζα διαθέτει πληροφορίες για μια λίστα πελατών της. Για κάθε πελάτη διαθέτει μια λίστα κινήσεων λογαριασμών και για κάθε κίνηση λογαριασμού διαθέτει πληροφορίες σχετικά με τα τιμολόγια αγοράς ή πώλησης που συνοδεύουν την κίνηση αυτή.

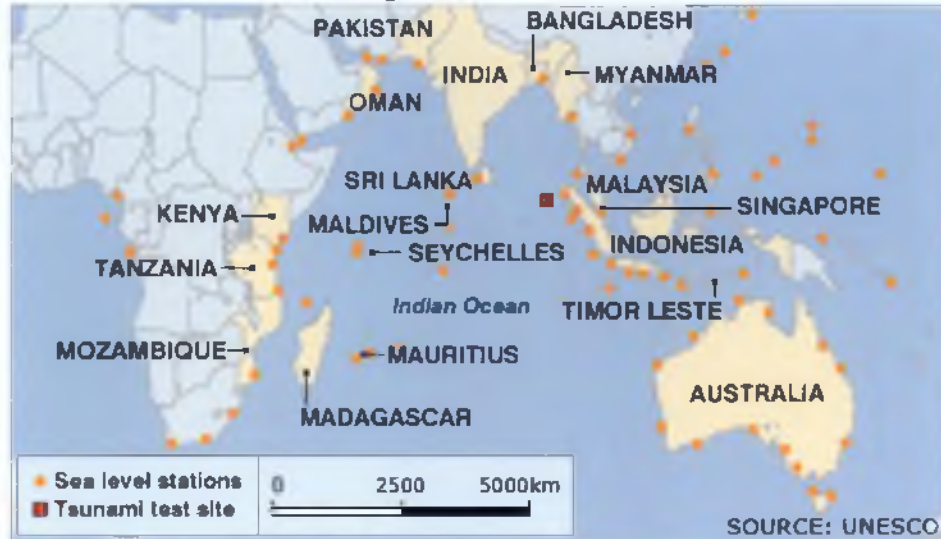
**2. Roll Up:** Είναι η ακριβώς αντίστροφη λειτουργία, όπου ο χρήστης από πιο λεπτομερή στοιχεία μπορεί να προβάλλει πιο συγκεντρωτικά.

**3. Drill across:** Όταν κινούμαστε σε διαφορετική διάσταση των δεδομένων. Για παράδειγμα κάποιος ερευνητής παρατηρεί δεδομένα με μεγάλο επίπεδο λεπτομέρειας από έναν μετεωρολογικό σταθμό. Επειδή όμως ενδιαφέρεται για τις επιπτώσεις ενός

πιθανού Tsunami, επιθυμεί να παρατηρήσει τα ίδια ακριβώς δεδομένα από έναν άλλο σταθμό σε διαφορετική γεωγραφική περιοχή.

**Early warning**

The tsunami test and monitoring stations:



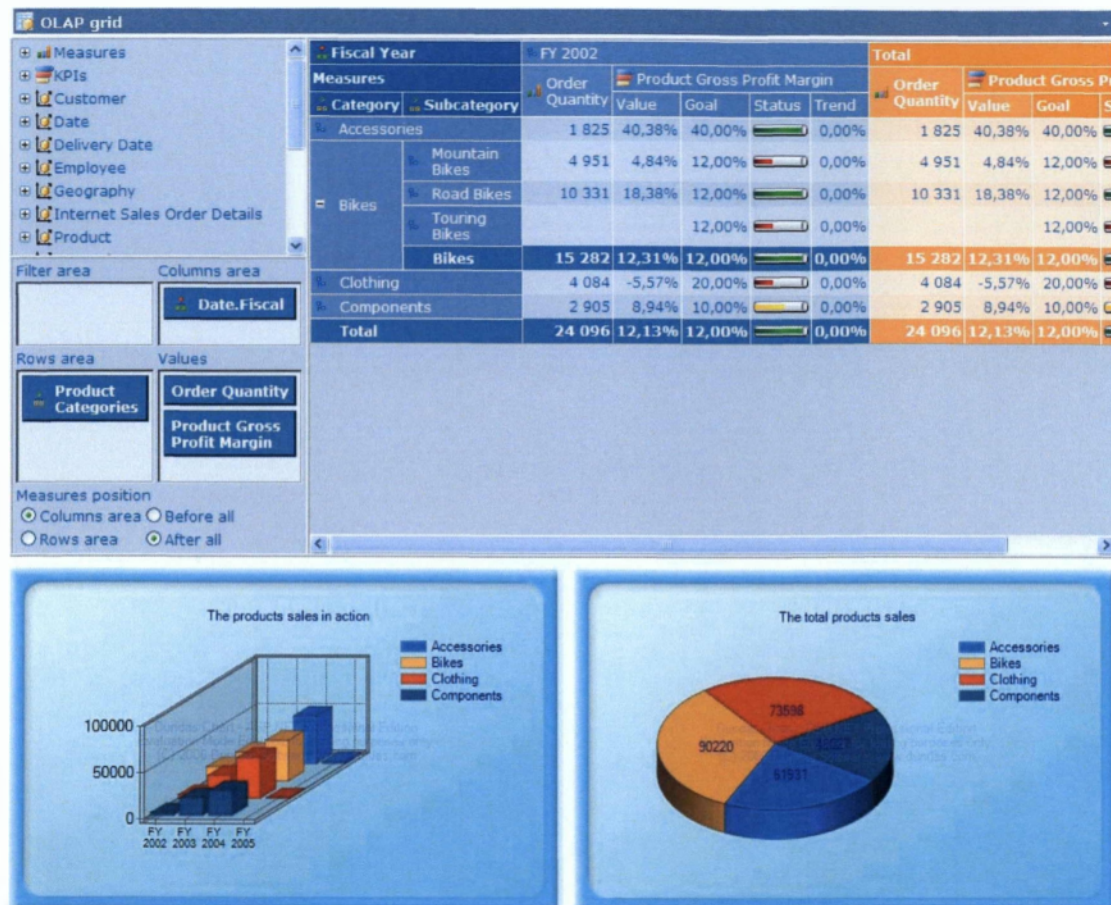
Παράδειγμα Drill-Across.

**4. Drill Through:** Είναι η εμφάνιση, κατά το δυνατόν, μεγαλύτερου επιπέδου λεπτομέρειας για τα δεδομένα.

Entry Page Title	Visits ↓	% ↓	Page Views ↓	Browse Rate ↓
Ecommerce Dashboard	1,644	49.91%	10,379	6.31
Page Item Summary	523	15.88%	3,427	6.55
Last Visitors	426	12.93%	5,871	13.78
Demographics	Areas of the World			1.05
Content	City			4.55
Navigation	Countries			15.28
Paths	Regions / State code			2.31
Campaigns	Area code (US & Canada)			9.62
Sales & Marketing	Zip code (US & Canada)			3.00
Search engines & Referrers	ISP			1.07
System	Languages			7.01
Visitor	Organizations			7.73
<b>Total</b>	Time Zones			

Παράδειγμα Drill Through

### 5. Απεικόνιση δεδομένων.



Απεικόνιση δεδομένων

### 6. Εμφωλιασμός: Η δυνατότητα εμφάνισης παραπάνω από δύο διαστάσεων δεδομένων.

Revenue as values		Camping Equipment	Golf Equipment	Mountaineering Equipment	Outdoor Protection	Personal Accessories	Products
2006	2006 Q 1	115,969,290	58,379,261	36,539,206	2,410,113	131,722,288	345,020,158
	2006 Q 2	131,594,512	62,467,714	42,683,784	2,632,786	153,205,078	392,583,874
	2006 Q 3	130,979,047	56,040,116	41,443,786	2,621,541	147,252,623	378,337,113
	2006 Q 4	122,376,880	54,485,598	40,379,738	2,694,775	162,054,064	381,991,055
<b>2006</b>	<b>500,919,729</b>	<b>231,372,689</b>	<b>161,046,514</b>	<b>10,359,215</b>	<b>594,234,053</b>	<b>1,497,932,200</b>	
2007	2007 Q 1	145,539,940	81,537,354	59,768,436	1,879,174	184,020,708	472,745,612
	2007 Q 2	153,809,380	69,081,676	60,116,560	1,887,360	194,759,998	479,654,974
	2007 Q 3	54,031,962	25,129,545	21,643,417	706,857	65,032,264	166,544,045
	2007 Q 4	0	0	0	0	0	0
<b>2007</b>	<b>353,381,282</b>	<b>175,748,575</b>	<b>141,528,413</b>	<b>4,473,391</b>	<b>443,812,970</b>	<b>1,118,944,631</b>	
<b>Years</b>	<b>1,590,730,027</b>	<b>715,631,819</b>	<b>715,631,819</b>	<b>76,002,938</b>	<b>1,586,038,235</b>	<b>4,691,531,035</b>	

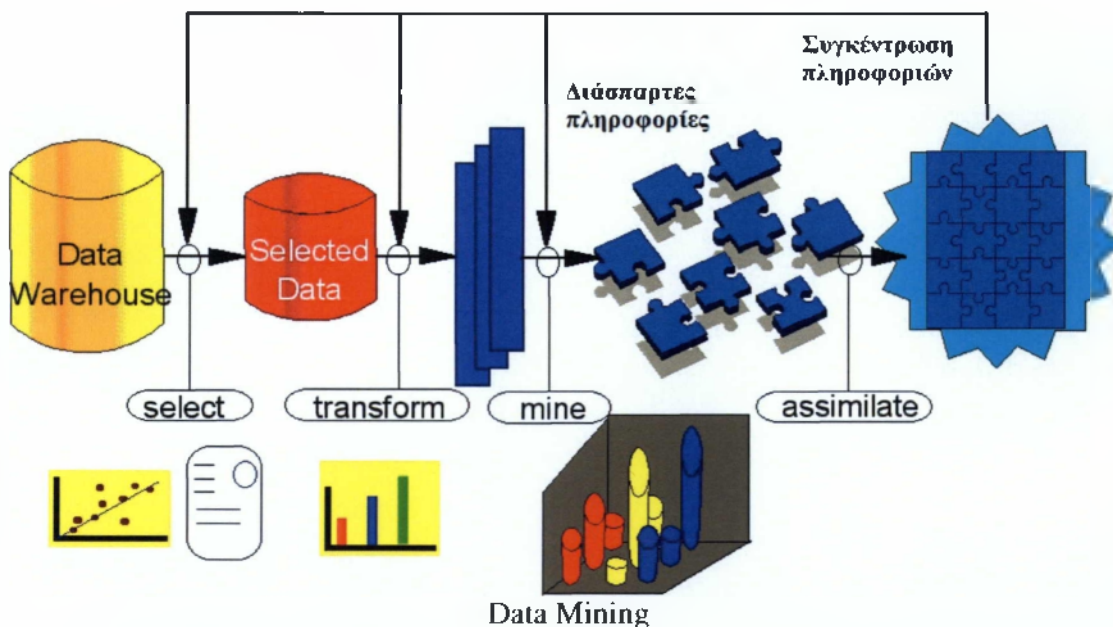
Εμφωλιασμός (nesting) δεδομένων

## ΚΕΦΑΛΑΙΟ 11

Που χρησιμοποιείται η εξόρυξη δεδομένων; Συνήθως στις παρακάτω περιπτώσεις:

- Ανίχνευση Απάτης.
- Καθορισμός Αποδοτικότητας.
- Καθορισμός κύκλου ζωής πελατών.
- Μοντελοποίηση επιχειρηματικής διεργασίας.

Και όλα αυτά γιατί μπορεί να διαχειρίζεται μεγάλο όγκο δεδομένων και μπορεί να υπολογίζει όλους τους πιθανούς συνδυασμούς μεταξύ μεγάλων υποσυνόλων αφού υποστηρίζει σύνθετους αλγόριθμους και στατιστικές τεχνικές. Πρέπει κανείς να πειραματιστεί πολύ, ώστε τα ευρήματα που θα προκύψουν από την χρήση κατάλληλων αλγορίθμων, να μπορούν να αξιοποιηθούν όπως πρέπει.

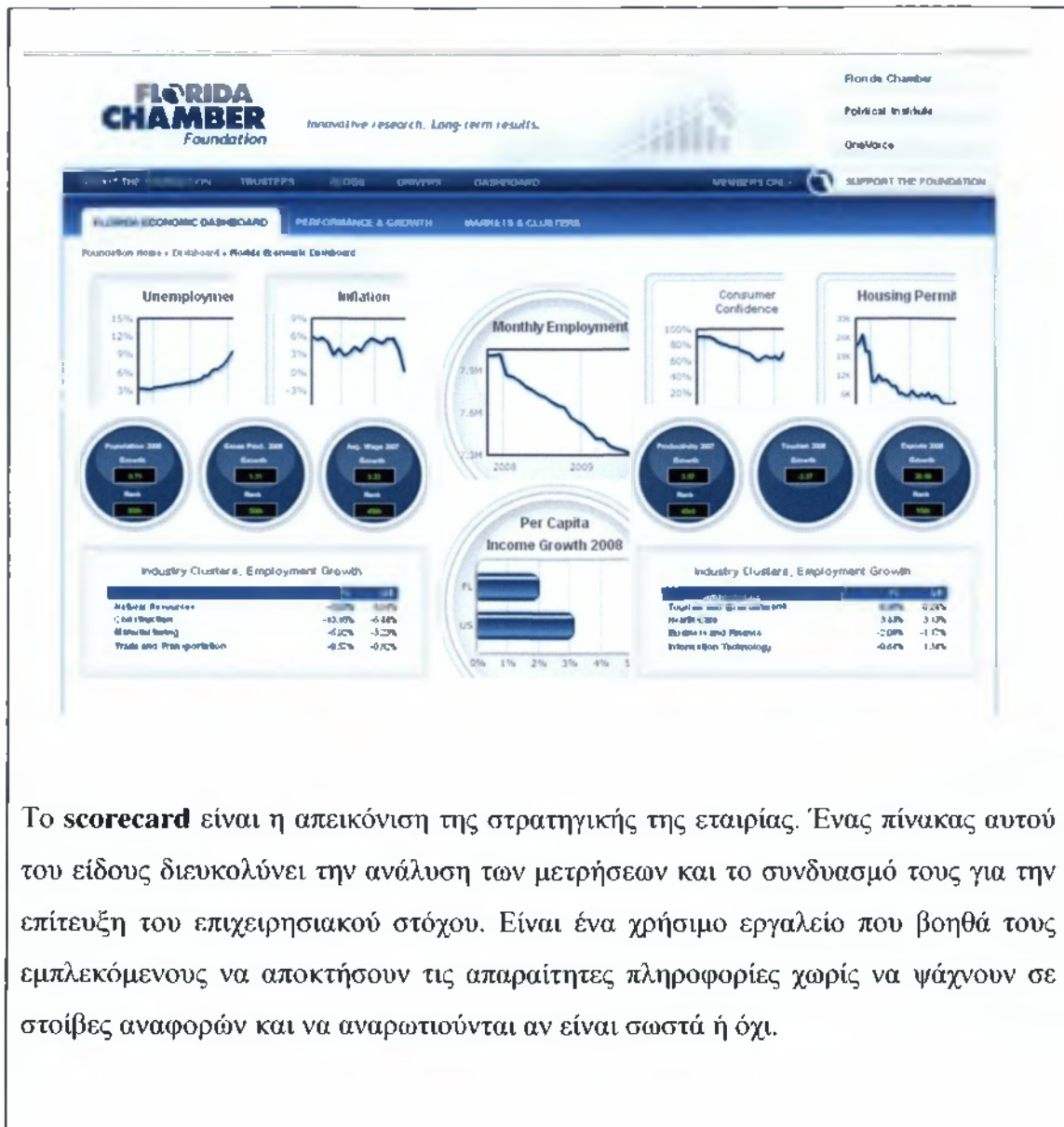


Όπως φαίνεται από το παραπάνω σχήμα αρχικά γίνεται εξαγωγή πληροφοριών (select) από μια αποθήκη δεδομένων. Τα δεδομένα αυτά μετατρέπονται / κανονικοποιούνται (transform) και κατόπιν εκτελείται κάποιος αλγόριθμος εξόρυξης δεδομένων (mine). Τα εξαγόμενα δεδομένα συγκεντρώνονται (assimilate) και παρουσιάζονται στο χρήστη. Περισσότερες πληροφορίες καθώς και ένα Video tutorial θα παρουσιαστούν στο μάθημα 4 όπου θα μπορέσετε να εκτελέσετε τεχνικές εξόρυξης δεδομένων.

## ΚΕΦΑΛΑΙΟ 12



Το **dashboard** είναι ένα σύνολο γράφων, αναφορών και δεικτών που βοηθούν στην παρακολούθηση της απόδοσης των λειτουργιών που διεξάγονται για έναν καθορισμένο στόχο. Σε ένα περιβάλλον Data Warehouse, ένας πίνακας dashboard παρουσιάζει τις αποδόσεις των λειτουργιών. Συνήθως κάθε χειριστής ενός συστήματος Data Warehouse διαθέτει τον προσωπικό του πίνακα.



Το **scorecard** είναι η απεικόνιση της στρατηγικής της εταιρίας. Ένας πίνακας αυτού του είδους διευκολύνει την ανάλυση των μετρήσεων και το συνδυασμό τους για την επίτευξη του επιχειρησιακού στόχου. Είναι ένα χρήσιμο εργαλείο που βοηθά τους εμπλεκόμενους να αποκτήσουν τις απαραίτητες πληροφορίες χωρίς να ψάχνουν σε στοιβές αναφορών και να αναρωτιούνται αν είναι σωστά ή όχι.

### 3.4 Μάθημα 4<sup>ο</sup> – Πρακτική εφαρμογή *Business Intelligence* και *Data Mining*

Το τέταρτο μάθημα διαφοροποιείται αρκετά από τα προηγούμενα και είναι απαραίτητη η συμμετοχή του εκπαιδευόμενου για το καλύτερο αποτέλεσμα της μαθησιακής διαδικασίας. Το μάθημα αποτελείται από δύο μέρη. Στο πρώτο παρουσιάζονται τα διάφορα πακέτα λογισμικού που χρησιμοποιούνται και λεπτομέρειες σχετικά με το Weka, το οποίο είναι ένα λογισμικό ανοιχτού κώδικα για εξόρυξη δεδομένων. Οδηγίες και σχετικά αρχεία με επιλεγμένα δεδομένα παρέχονται στον φοιτητή. Στο δεύτερο μέρος ο φοιτητής μπορεί να παρακολουθήσει μια δεκάλεπτη παρουσίαση – video tutorial. Σε περίπτωση που κάποιος επιθυμεί να μάθει τον τρόπο παρασκευής του οπτικοακουστικού υλικού μπορεί να κατεβάσει το σχετικό αρχείο. Ο τρόπος δημιουργίας του οπτικοακουστικού υλικού παρουσιάζεται στο κεφάλαιο 4.

4



**Μάθημα 4ο: Εργαστήριο: Πρακτική εφαρμογή *Business Intelligence* και *Data Mining***

-  ΕΡΓΑΣΤΗΡΙΟ - Από τη θεωρία στην πράξη
-  ΟΠΤΙΚΟΑΚΟΥΣΤΙΚΟ ΥΛΙΚΟ

□

Εικόνα 3-5 Παρουσίαση 4<sup>ου</sup> μαθήματος

#### Λογισμικό για Αποθήκες δεδομένων και επεξεργασία δεδομένων

Η πρόταση της Microsoft στις εταιρίες που ενδιαφέρονται για τις Αποθήκες Δεδομένων και την ανάκτηση χρήσιμων πληροφοριών από αυτές ακούει στο όνομα [Microsoft Business Intelligence](#).

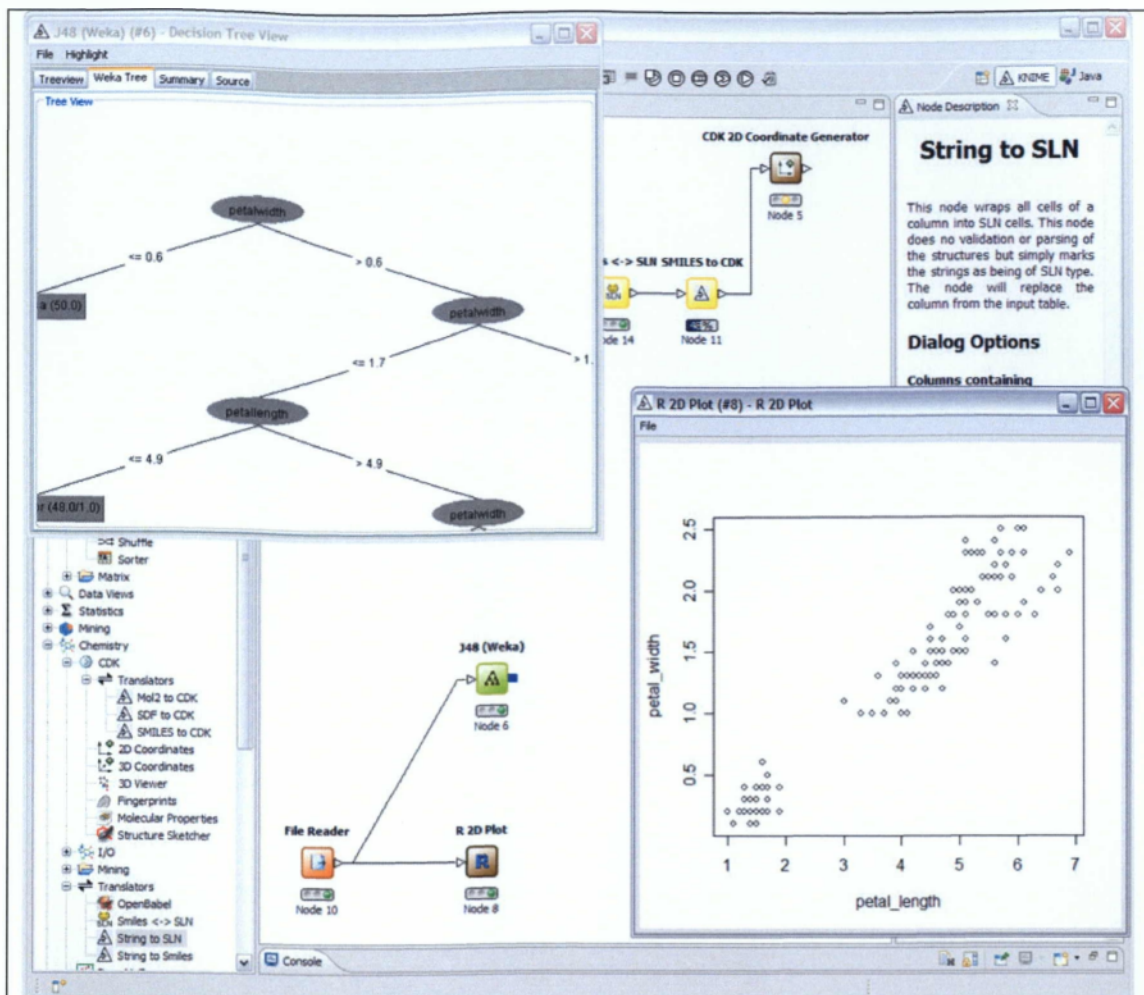




Η πρόταση της Oracle είναι μια πλατφόρμα που ονομάζεται [Oracle Business Intelligence](#).



Από τις διαθέσιμες πλατφόρμες ανοιχτού κώδικα που μπορεί κανείς να προμηθευτεί δωρεάν και να πειραματιστεί μαζί τους, ξεχωρίζει η προσπάθεια του πανεπιστημίου του Waikato που αναπτύσσει πάνω από 10 χρόνια με τη βοήθεια της κοινότητας ανοιχτού λογισμικού το εργαλείο [Weka](#).



Στη συνέχεια θα παρουσιαστεί η μεθοδολογία εγκατάστασης και εκτέλεσης του εργαλείου WEKA.

### **Βήμα 1:**

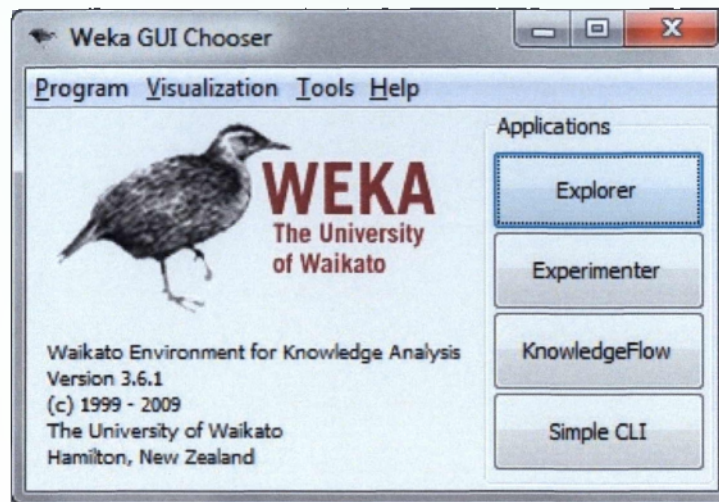
Στη σελίδα [http://www.cs.waikato.ac.nz/ml/weka/index\\_downloading.html](http://www.cs.waikato.ac.nz/ml/weka/index_downloading.html) μπορεί κάποιος να κατεβάσει την τελευταία έκδοση του Weka. Η τελευταία έκδοση αυτή τη στιγμή είναι η 3.6. Κατεβάζοντας το Stable GUI Version που περιέχει την Java Virtual Machine, η οποία απαιτεί 37 Mbytes, μπορεί κάποιος να μπει στον κόσμο του Business Intelligence και του Data Mining.

### **Βήμα 2:**

Όταν κατεβάσετε την πλατφόρμα του WEKA αποσυμπίεστε την σε έναν φάκελο π.χ. C:\Weka-3-6 και στη συνέχεια εκτελέστε το Weka. Τόσο απλά.

**Βήμα 3:**

Η παρακάτω εικόνα της εφαρμογής θα εμφανιστεί στον υπολογιστή σας.



Για να ξεκινήσουμε θα ανοίξουμε αρχικά τον Explorer. Μέσα από τον Explorer μπορούμε να γνωστοποιήσουμε στο Weka τα δεδομένα τα οποία προέρχονται από την Αποθήκη Δεδομένων μας.

**Βήμα 4:**

Από τον explorer θα ανοίξουμε το αρχείο [bank\\_data.arff](#) αλλά και το [bank\\_data\\_normalized.arff](#). Η διαφορά είναι ότι στη μια περίπτωση τα δεδομένα είναι κανονικοποιημένα ενώ στην άλλη όχι. Όπως μπορείτε να δείτε παρακάτω το αρχείο αυτό περιέχει 600 εγγραφές σχετικά με: την ηλικία, το φύλο, τον τόπο κατοικίας, το εισόδημα, την οικογενειακή κατάσταση, τον αριθμό των παιδιών, το αυτοκίνητο, αν χρησιμοποιείται λογαριασμός αποταμιεύσεων στην τράπεζα ή λογαριασμός όψεως ή υπάρχει κάποιο δάνειο ή κάποιο προσωπικό πλάνο διαχείρισης κεφαλαίων (per).

The screenshot shows the Weka Explorer interface. The 'age' attribute is selected, and its distribution is visualized as a stacked bar chart. The chart shows three categories: '0\_34' (195 instances), '35\_51' (214 instances), and '52\_max' (191 instances). The bars are stacked with blue at the bottom and red on top.

No.	Label	Count
1	0_34	195
2	35_51	214
3	52_max	191

### Βήμα 5:

Για να κατανοήσετε τη λειτουργία των διαδικασιών των τεχνικών εξόρυξης δεδομένων κατεβάστε τα παραπάνω 2 ARFF αρχεία και το λογισμικό WEKA. Στο επόμενο κεφάλαιο θα παρουσιαστεί ένα Video διάρκειας περίπου 10 λεπτών που θα περιέχει αναλυτικά τις δυνατότητες του λογισμικού.

### **ΘΕΩΡΙΑ ΓΙΑ ΤΟ WEKA**

Ενδεικτικά το WEKA μπορεί να χρησιμοποιήσει δεδομένα από πολλές πηγές. Για παράδειγμα παρουσιάζονται κάποιες από τις βάσεις δεδομένων με τις οποίες μπορεί να συνδεθεί :

- HSQLDB ( $\geq 3.4.1$ )
- MS Access ( $> 3.4.14$ ,  $> 3.5.8$ ,  $> 3.6.0$ )
- MS SQL Server 2000 ( $\geq 3.4.9$ ,  $\geq 3.5.4$ )
- MS SQL Server 2005 ( $\geq 3.4.11$ ,  $\geq 3.5.6$ )
- MySQL ( $\geq 3.4.9$ ,  $\geq 3.5.4$ )
- ODBC access via Sun's ODBC/JDBC bridge
- Oracle 10g ( $\geq 3.4.9$ ,  $\geq 3.5.4$ )
- PostgreSQL 7.4 ( $\geq 3.4.9$ ,  $\geq 3.5.4$ )

- sqlite 3.x (> 3.4.12, > 3.5.7)

Επειδή το εργαλείο WEKA είναι υλοποιημένο σε Java, χρειάζεται κάποιος να ακολουθήσει τις οδηγίες από τη σελίδα [Wiki του WEKA](#) για να κατεβάσει τον κατάλληλο JDBC Driver και να κάνει τις απαραίτητες ρυθμίσεις.

Σκοπός αυτού του ηλεκτρονικού εργαστηρίου είναι η γρήγορη επαφή σας με την τεχνολογία αυτή, επομένως τα δεδομένα θα προέλθουν από ένα αρχείο .ARFF και όχι από βάση δεδομένων.

Τα αρχεία .ARFF (Attribute-Relation File Format) αποτελούν ASCII αρχεία που ακολουθούν ορισμένους κανόνες που έχει ορίσει το WEKA και περιγράφουν το Data Set που θα χρησιμοποιηθεί. Κάθε αρχείο αποτελείται από τρία κύρια μέρη. Το **@RELATION**, το **@ATTRIBUTE** και το **@DATA**.

Η πρώτη γραμμή κάθε αρχείου ξεκινάει με:

```
@relation <ονομα-relation>
```

Στη συνέχεια ακολουθεί η δήλωση **@ATTRIBUTE**. Για κάθε στήλη από την οποία αποτελούνται τα δεδομένα απαιτείται μια αντίστοιχη περιγραφή. Αν για παράδειγμα τα δεδομένα μας αποτελούνται από τρεις αριθμητικές στήλες, τότε το 2ο μέρος του αρχείου είναι:

```
@ATTRIBUTE perigrافي_prwtis_stilis NUMERIC
@ATTRIBUTE perigrافي_deyteris_stilis NUMERIC
@ATTRIBUTE perigrافي_tritis_stilis NUMERIC
```

Επίσης ακολουθεί η δήλωση **@DATA** με τα δεδομένα.

```
@DATA
5.1,3.5,1.4
4.9,3.0,1.4
4.7,3.2,1.3
4.6,3.1,1.5
```

Τα παραπάνω αποτελούν ένα έγκυρο αρχείο .ARFF.

Ανοίγετε με το σημειωματάριο τα 2 αρχεία .ARFF που μπορείτε να κατεβάσετε από τη σελίδα αυτή για να παρατηρήσετε την εσωτερική τους δομή. Όταν είστε έτοιμοι, προχωρήστε στο επόμενο μάθημα.

### 3.5 Μάθημα 5<sup>ο</sup> - Προετοιμασία υλοποίησης ενός Data Warehouse

Στο πέμπτο μάθημα παρουσιάζονται στοιχεία σχετικά με την προετοιμασία και τη διαδικασία που πρέπει να ακολουθηθεί προτού ξεκινήσει η υλοποίηση ενός Data Warehouse Project. Από την επιλογή των μελών που θα συμμετάσχουν στην ομάδα εργασίας μέχρι τη συλλογή και ανάλυση των απαιτήσεων.

5



#### Μάθημα 5ο: Προετοιμασία υλοποίησης ενός Data Warehouse

Στο μάθημα αυτό θα παρουσιάσουμε πληροφορίες σχετικά με την απαραίτητη προετοιμασία και διαδικασία που πρέπει να ακολουθηθεί προτού ξεκινήσει η υλοποίηση ενός Data Warehouse Project. Από τη επιλογή των μελών που θα συμμετάσχουν στην ομάδα εργασίας μέχρι την συλλογή και ανάλυση απαιτήσεων.

- 13. Κοινές πρακτικές για την κατασκευή ενός Data Warehouse
- 14. Καθορισμός Ατόμων και Οργάνωση Ομάδας υλοποίηση ενός Data Warehouse
- 15. Data Warehouse - Ανάλυση Απαιτήσεων

Εικόνα 3-6 Παρουσίαση 5<sup>ου</sup> μαθήματος

## ΚΕΦΑΛΑΙΟ 13

Οργάνωση του έργου κατασκευής ενός Data Warehouse. Κοινές πρακτικές για την κατασκευή ενός Data Warehouse και οποιουδήποτε άλλου έργου:

- **Απαιτήσεις:** Η φάση κατά την οποία καθορίζεται ο στόχος του έργου και αναλύονται με λεπτομέρεια οι λειτουργίες και οι υπηρεσίες που παρέχει.
- **Σχεδιασμός:** Οι τεχνικές λεπτομέρειες για τις υπηρεσίες που αποφασίστηκαν στην προηγούμενη φάση.
- **Υλοποίηση και Έλεγχος:** Δημιουργούνται οι Β.Δ., οι εφαρμογές για να μεταφέρονται τα δεδομένα και το επίπεδο επικοινωνίας με το χρήστη.
- **Ανάπτυξη:** Οι πελάτες ξεκινούν να χρησιμοποιούν την εφαρμογή.

Οι καλές πρακτικές γενικά πρέπει να προωθούν:

- Τους συχνούς ελέγχους και την αναπροσαρμογή τους.
- Την ενίσχυση της **συνεργατικότητας**, του ομαδικού πνεύματος και της υπευθυνότητας.
- Τη **γρήγορη παράδοση** ενός λογισμικού υψηλής ποιότητας μέσω καλών μηχανικών τεχνικών.
- Την **ικανοποίηση** τόσο των αναγκών των πελατών όσο και των στόχων της επιχείρησης.

Απαιτήσεις για την κατασκευή Data Warehouse:

1. Αναγνωρίστε τις ερωτήσεις κλειδιά.
2. Καθορίστε τις συγκεκριμένες απαντήσεις.
3. Καθορίστε τα συστατικά κλειδιά του έργου.
4. Προσδιορίστε τις απαιτήσεις χρηστικότητας.
5. Τροποποιείτε τη λύση σε μορφή προτύπου.

Σχεδιασμός:

1. Αναγνωρίστε όλες τις συνιστώσες πηγές.
2. Ιεραρχήστε τις συνιστώσες πηγές.
3. Μελετήστε δείγματα δεδομένων από αυτά που έχετε εξάγει.
4. Διαμορφώστε ένα γρήγορο σχέδιο της Β.Δ.
5. Ενημερώστε τους πελάτες και αποδεχτείτε τις επιθυμητές αλλαγές.

Υλοποίηση και έλεγχος

1. Δημιουργείτε το ολοκληρωμένο σχέδιο της Β.Δ.
2. Δημιουργείτε μια ολοκληρωμένη αποθήκη δεδομένων και αντιστοιχίστε τα δεδομένα με αυτά που έχουν ληφθεί από τις πηγές.
3. Μελετήστε εκ των έσω τις πηγές δεδομένων.
4. Κατασκευάστε αλγόριθμους μετασχηματισμού.
5. Καθορίστε τα εργαλεία ή σχεδιάστε τον κώδικα για την επικοινωνία μεταξύ των στρωμάτων του λογισμικού.

6. Καθορίστε τους μηχανισμούς διαχείρισης της αποθήκης δεδομένων.
7. Διασφαλίστε την ποιότητα των δεδομένων.
8. Φορτώστε τα δεδομένα χρησιμοποιώντας τα επιλεγμένα εργαλεία.
9. Διεξάγετε πρωταρχικούς ελέγχους για την λειτουργικότητα της αρχιτεκτονικής.
10. Διεξάγετε ελέγχους από άκρο σε άκρο.
11. Διεξάγετε πρωταρχικό καθορισμό των επιπέδων σύνοψης των δεδομένων.

#### Ανάπτυξη

1. Υπολογίστε το πλήθος των δεδομένων στη Β.Δ.
2. Παρακολουθείτε τη διαδικασία φόρτωσης δεδομένων στην αποθήκη δεδομένων.
3. Προσδιορίστε στρατηγικές εκκαθάρισης και αρχειοθέτησης.
4. Καθορίστε τους μηχανισμούς παρακολούθησης αλλαγών των δεδομένων.
5. Κατασκευή ενός έργου αποθήκης δεδομένων μεγάλης κλίμακας.
6. Χρήση μιας από τις προσεγγίσεις Top-Down, Bottom- Up ή Mixed- Mode

#### Top – Down

1. Προχωρήστε με το πεδίο εφαρμογής και τις διαδικασίες της φάσης του σχεδιασμού από την πλευρά του ολικού περιβάλλοντος της αποθήκης δεδομένων, ανεξαρτήτως μεγέθους.
2. Δημιουργείτε ένα μοντέλο που περιλαμβάνει όλα τα δεδομένα σας (enterprise data warehouse model).
3. Κατακερματίστε το μοντέλο αυτό σε όσο το δυνατό μικρότερα συστατικά μέρη (μικρότερα μοντέλα).
4. Υλοποιείτε κάθε συστατικό μέρος ξεχωριστά, αλλά φροντίστε κάθε μέρος να περιέχει ένα συστατικό μοντέλο.

Η παραπάνω διαδικασία είναι πολύ συνηθισμένη και μπορεί να θεωρηθεί ως η σωστή πρακτική του «διαίρει και βασίλευε». Παρότι αυτή η τεχνική φαίνεται απόλυτα λογική, στην πραγματικότητα η υιοθέτησή της δε φέρνει πάντα τα επιθυμητά αποτελέσματα. Το μεγαλύτερο εμπόδιο είναι η δημιουργία του επιχειρησιακού μοντέλου καθώς περιέχει πολύ μεγάλη πληροφορία με αποτέλεσμα να δυσχεραίνει



την άρτια σχεδίασή του.

### Bottom –Up

Σε αυτή την παραδοχή, αφού αναγνωριστούν οι διαφορετικές θεματικές περιοχές ή τα σύνολά τους, κατασκευάζετε το καθένα ως ένα διαφορετικό έργο με μια μικρή επικάλυψη μεταξύ τους. Το μεγαλύτερο ρίσκο σε αυτή την περίπτωση είναι ότι αφού ολοκληρωθούν τα έργα υπάρχει πιθανότητα να μην είναι συμβατά μεταξύ τους, τουλάχιστον όχι για το μακρινό χρονικό ορίζοντα που θέλει ο κατασκευαστής να έχει διάρκεια το έργο του.

### Mixed- Mode

Ένας συνδυασμός των παραπάνω και μάλλον η καλύτερη πρακτική:

1. Ξεκινήστε με τον καθορισμό του σκοπού και των απαιτήσεων του έργου.
2. Καθορίστε την αρχιτεκτονική του DW σε ανεξάρτητη φάση του σχεδιασμού όπου καταγράφονται όλοι οι πόροι, και το λογισμικό το οποίο θα χρησιμοποιηθεί. Καθορίζονται επίσης οι χρήστες, οι απαιτήσεις τους και αναγνωρίζονται οι εξωτερικές πηγές που πιθανόν είναι απαραίτητες.
3. Βάλτε σε λειτουργία ένα πιλοτικό πρόγραμμα για το σχεδιασμό και την ανάπτυξη του DW κατά τη διάρκεια του οποίου θα πραγματοποιηθεί και αξιολόγηση των φάσεων που έχουν ήδη ολοκληρωθεί.
4. Συγκρίνετε την προτεινόμενη αρχιτεκτονική της βάσης με τα αποτελέσματα του πιλοτικού σχεδίου.
5. Δημιουργείστε δύο ανεξάρτητα μονοπάτια σχεδίασης και ανάπτυξης, ένα που θα καθορίζει τις δυνατότητες των υποδομών και ένα που θα καθορίζει τη λειτουργικότητα των συστατικών μερών του DW.
6. Χωρίστε τις συστατικές λειτουργίες του DW σε διαφορετικά έργα.
7. Συνεχίστε κάθε έργο ξεχωριστά για την επόμενη φάση.

**ΚΕΦΑΛΑΙΟ 14**Ρόλοι εργασίας σε μια Αποθήκη Δεδομένων:

1. **Project Manager** – Υπεύθυνος έργου: Το άτομο το οποίο πρέπει να διαθέτει επικοινωνιακές ικανότητες, να είναι συνεπής με το χρονοδιάγραμμα και παράλληλα ευέλικτος, να λαμβάνει υπεύθυνες αποφάσεις και να έχει επαρκείς γνώσεις για το αντικείμενο του DW.
2. **Technical Leader** – Υπεύθυνος τεχνικής ομάδας: το άτομο που πρέπει να διασφαλίζει ότι όλα τα συνιστώμενα μέρη λειτουργούν και να καθοδηγεί την ομάδα στην επιτυχή υλοποίηση του έργου.
3. **Chief Architect** – Υπεύθυνος αρχιτεκτονικής: διατελεί λειτουργίες σύμφωνα με τις επιχειρησιακές ανάγκες, την τεχνολογική πρόοδο και την κατευθυντήρια γραμμή της εταιρείας.
4. **Business Requirements Analyst** – Αναλυτής Λειτουργικών Απαιτήσεων: συλλέγει, ενοποιεί, οργανώνει και ιεραρχεί τις λειτουργικές ανάγκες και τα προβλήματα που παρουσιάζονται από τους μελλοντικούς χρήστες.
5. **Data Modeler** και conceptual/ logical database designer: Ο ρόλος του είναι η εννοιολογική θεμελίωση, ο σχεδιασμός της ΒΔ για την αποθήκη δεδομένων και η μοντελοποίηση των απαιτήσεων.
6. **Database Administrator** και physical database designer – Διαχειριστής ΒΔ και φυσικού σχεδιασμού: κατασκευάζει το σχήμα της ΒΔ (τη φυσική δομή), το μορφοποιεί, μελετά τη λειτουργικότητα και την απόδοση του σχήματα με την εκτέλεση συγκεκριμένων ελέγχων και εκτελεί ενέργειες δημιουργίας αντιγράφων ασφάλειας ή ανάκτησης δεδομένων.
7. **Front- end tools specialists** και developers: Οι προγραμματιστές που αξιολογούν, επιλέγουν και κατασκευάζουν τα εργαλεία και τις εφαρμογές μέσω των οποίων οι χρήστες θα έχουν πρόσβαση στο λειτουργικό περιβάλλον.
8. **Middleware Specialists**: οι προγραμματιστές που καθορίζουν τα εργαλεία των ενδιάμεσων σταδίων.
9. **Quality Assurance specialist** – Υπεύθυνος Ποιότητας: καθορίζει τις παραμέτρους, τους αλγόριθμους και τα εργαλεία/ εφαρμογές που λειτουργούν

σωστά και αποδοτικά.

10. **Source Data Analyst** – Αναλυτής Δεδομένων: Ελέγχει τα δεδομένα που εισάγονται στη ΒΔ μετά από τις εξαγωγές και τους μετασχηματισμούς στους οποίους υπόκεινται.
11. **User Community Interaction manager**: Ο υπεύθυνος επικοινωνίας με τους πελάτες συζητά τυχόν προβλήματα με τους μελλοντικούς χρήστες της ΒΔ.
12. **Technical executive sponsor**: ο ειδικός σε θέματα τεχνολογικής αρχιτεκτονικής που εκπροσωπεί την εταιρεία και το εκάστοτε έργο.
13. **User community executive sponsor**: Αντίστοιχα με το παραπάνω, που απευθύνεται στους χρήστες και εκφράζει τις απόψεις τους κατά τη διάρκεια ανάπτυξης του έργου.

Σε πολλές περιπτώσεις ένα άτομο μπορεί να έχει παραπάνω από μια αρμοδιότητες. Ο διαχωρισμός των υποχρεώσεων και των ρόλων φαίνεται στο Επιχειρησιακό Μοντέλο Λειτουργικότητας το οποίο καθορίζουν οι υπεύθυνοι κατά την ανάληψη ενός έργου και δείχνει τη λειτουργία μονάδων και τις σχέσεις μεταξύ τους, καθώς και τις αποφάσεις και τη ροή των διαδικασιών σχετικά με το κάθε τμήμα.

## ΚΕΦΑΛΑΙΟ 15

Υπάρχουν δύο προσεγγίσεις για να πραγματοποιηθεί το αρχικό στάδιο της ανάλυσης των απαιτήσεων (που έχει εξαιρετικά σημαντικό ρόλο για όλη τη συνέχεια και συνέπεια του έργου). Η μία είναι προσανατολισμένη στην τεχνολογική ανάπτυξη της αποθήκης δεδομένων, ενώ η δεύτερη στη λειτουργικότητα του σχήματος. Πιο αναλυτικά:

**Technically Driven Data Warehousing**: βασίζεται στη λήψη αποφάσεων των βασικών αρχών της εταιρείας χωρίς όμως να πραγματοποιεί ανάλυση των απαιτήσεων. Δυστυχώς λόγω μικρής ενεργοποίησης της ίδιας της εταιρείας τα έργα που σχεδιάζονται με αυτή την προσέγγιση, συνήθως δεν ολοκληρώνονται. Γενικά η προσέγγιση αυτή εστιάζει σε θεματικές περιοχές εκ των οποίων οι πιο σημαντικές είναι οι απαιτήσεις των πελατών και του προϊόντος. Στη συνέχεια σχεδιάζεται το enterprise data model που εστιάζει στις επιχειρησιακές οντότητες και βοηθά στη

συλλογή δεδομένων που δεν είχαν εντοπιστεί.

**Business Driven Business Intelligence:** Η καλύτερη μέθοδος για την ανάλυση απαιτήσεων είναι να βασιστεί ο σχεδιαστής αρχικά στις απαιτήσεις του συστήματος και της εταιρίας ακούγοντας προσεκτικά τους πελάτες. Η επιτυχία της προσέγγισης αυτής βασίζεται στην πιστή υιοθέτηση των παρακάτω αρχών. Αρχικά πρέπει να υπάρχει μια ενιαία μεθοδολογία ώστε να επιτευχθούν οι στρατηγικοί στόχοι της επιχείρησης, όπως και ένα πλάνο διαδικασιών τα αποτελέσματα του οποίου να αποδεικνύουν την επιτυχή πορεία του έργου. Βασική παραδοχή είναι ότι όλα τα δεδομένα αλληλοεπηρεάζονται, άρα υπάρχει ανάγκη για υποδομές που να μπορούν να υποστηρίξουν αυτές τις συσχετίσεις.

Όλα ξεκινούν από τη διάρθρωση των ερωτήσεων. Οι απαιτήσεις που πρέπει να πραγματοποιεί το έργο είναι είτε οι απαντήσεις, είτε ο συνδυασμός τους. Ο σχεδιαστής χρειάζεται να γνωρίζει καλά την εταιρία ώστε να κατανοήσει την επιχειρησιακή αλυσίδα αξίας.

Αφού έχει συλλεχθεί η πληροφορία πρέπει να γίνει μια **αξιολόγηση** για την ιεράρχηση των πρωταγωνιστικών ή δευτερευόντων απαιτήσεων και να καθοριστούν τα αντικείμενα/ υπηρεσίες κλειδιά.

Στη συνέχεια κατασκευάζεται ένα **επιχειρησιακό μοντέλο** που παρέχει την απεικόνιση της επιθυμητής από τους χρήστες λύσης σε συνδυασμό με τις απαντήσεις που έδωσαν στην αρχική φάση των ερωτήσεων. Το μοντέλο αυτό επιτρέπει τη δημιουργία αναλυτικής αναφοράς, ενώ περιλαμβάνει τις ιεραρχίες των δεδομένων.

Επειδή οι έγγραφες αναφορές δεν είναι αρκετές στην παρούσα φάση, πρέπει να κατασκευαστεί μια **πρότυπη εφαρμογή**, μια μικρογραφία, του τελικού συστήματος και των υπηρεσιών του. Σε όλα αυτά τα στάδια **οι πελάτες πρέπει να ενημερώνονται**, να κάνουν ελέγχους και να διορθώνουν ότι δεν τους ικανοποιεί, ώστε οι αλλαγές να γίνονται στο αρχικό στάδιο και όχι στα στάδια της υλοποίησης.

### 3.6 Μάθημα 6<sup>ο</sup> - Υλοποίηση και παράδοση στον πελάτη

Στο έκτο μάθημα παρουσιάζονται πληροφορίες σχετικά με την ανάλυση των πρωταρχικών πηγών από όπου θα προέλθουν τα δεδομένα και όλα τα απαραίτητα βήματα: Απαιτήσεις, Ανάλυση, Σχεδιασμός, Κατασκευή, Έλεγχος Ανάπτυξης, για την υλοποίηση ενός Data Warehouse.

6



#### Μάθημα 6ο: Υλοποίηση και παράδοση στον πελάτη

Στο μάθημα αυτό θα παρουσιάσουμε την ανάλυση των πρωταρχικών πηγών από όπου θα προέλθουν τα δεδομένα και όλα τα απαραίτητα βήματα: Απαιτήσεις, Ανάλυση, Σχεδιασμός, Κατασκευή, Έλεγχος Ανάπτυξης, για την υλοποίηση ενός Data Warehouse

- ☐ 16. Ανάλυση πόρων
- ☐ 17. Αρχιτεκτονική Αποθηκών Δεδομένων
- ☐ 18. Έλεγχος συστήματος & feedback

Εικόνα 3-7 Παρουσίαση 6<sup>ου</sup> μαθήματος

## ΚΕΦΑΛΑΙΟ 16

Η διαδικασία ανάλυσης των πόρων από τους οποίους συλλέγονται τα δεδομένα των αποθηκών δεδομένων έχει ως πρωταρχικό ρόλο την κατασκευή μιας αποθήκης δεδομένων καθώς όλες οι προκείμενες εξαγωγές και οι μετασχηματισμοί εξαρτώνται από το περιεχόμενό τους. Πριν από την έναρξη της έρευνας είναι πρωταρχικό μέλημα η ανίχνευση και η πλήρης κατανόηση της δομής των πηγών δεδομένων. Η κατανόηση της δομής περιλαμβάνει στοιχεία όπως:

- Οι πίνακες, η ονομασία κάθε στήλης, οι περιορισμοί της ΒΔ όπως και κάθε άλλη δήλωση DDL (Data Definition Language).
- Η δομή των αρχείων δεδομένων.
- Οι ορισμοί που ενδέχεται να είναι αποθηκευμένοι σε ένα κεντρικό ευρετήριο ή αποθετήριο δεδομένων.

Λόγω μεγάλης ποικιλίας θα πρέπει να γίνει μελέτη πολλών ειδών αρχιτεκτονικής ΒΔ και δομής αρχείων. Αν η αποθήκη δεδομένων περιέχει ιστορικά στοιχεία, είναι πολύ πιθανό να υπάρχουν αρχεία πολλών ετών αποθηκευμένα σε διαφορετικές διατάξεις. Δε χρειάζεται να γίνει μελέτη όλων των δεδομένων, αρκεί να αναγνωρίζονται αυτά που σχετίζονται με τους επιχειρησιακούς σκοπούς της ΒΔ. Η διαδικασία έχει τα εξής στάδια:

1. Διαχωρίζονται τα γεγονότα που σχετίζονται με την αποδοτικότητα της αποθήκης δεδομένων.
2. Καθορίζονται, σύμφωνα με το παραπάνω βήμα, οι πηγές δεδομένων που παρέχουν στοιχεία απαραίτητα για την κατασκευή της αποθήκης δεδομένων.
3. Ξεκινάει η επόμενη φάση του σχεδιασμού της αποθήκης δεδομένων.

Πολλά δεδομένα μπορούν να εξαλειφθούν όταν δεν είναι αναγκαίο να αναλυθούν. Συνήθως εξαλείφονται οι πίνακες που χρησιμοποιούνται για θέματα διαχείρισης συστήματος ή αυτοί που έχουν διατηρηθεί για μελλοντική χρήση.

Στην πράξη σε μερικές περιπτώσεις χρειάζεται να γίνει ανάλυση δεδομένων που δε βρίσκονται σε εμφανή σημεία. Πολλές φορές είχαν διαγραφεί δεδομένα που δεν είχαν θεωρηθεί απαραίτητα ή είχαν καταστραφεί λόγω κακής αποθήκευσης ή λόγω τροποποιήσεων της βάσης δεδομένων. Η διαδικασία αυτή μπορεί να είναι και δυσχερής ανάλογα με τη δυνατότητα ανάκτησης ή όχι των απολεσθέντων δεδομένων.

## **ΚΕΦΑΛΑΙΟ 17**

Αρχιτεκτονική μπορεί να θεωρηθεί το στυλ και η μέθοδος σχεδιασμού και κατασκευής μιας ταξινομημένης σειράς τμημάτων. Σε ότι αφορά τις αποθήκες δεδομένων, η αρχιτεκτονική τους πρέπει να διευκολύνει και να δημιουργεί πηγές δεδομένων, που είναι ακριβείς, διαμοιράσιμες και εύκολα προσβάσιμες από το κοινό. Στόχος είναι να μπορεί κάθε φορά να ολοκληρώνει όλο και πιο γρήγορα μια διεργασία. Αυτές οι απαιτήσεις μπορούν να ικανοποιηθούν αν χρησιμοποιηθεί ευπροσάρμοστη αρχιτεκτονική. Ευπροσάρμοστη σημαίνει να είναι κατάλληλη για κάθε χρήση ή σε κάθε περίπτωση.

Τα βασικά συστατικά της επιτυχούς αρχιτεκτονικής είναι:

- Οι άνθρωποι και οι ρόλοι τους: Οι ομάδες που αναλαμβάνουν την διεκπεραίωση του έργου.
- Η επαναλαμβανόμενη διαδικασία: Αφού όλες οι οντότητες του συστήματος αλληλεπιδρούν οι διαδικασίες που τις συνδέουν πρέπει επίσης να αλληλεπιδρούν.
- Η προκαθορισμένη πλατφόρμα διανομής: Πρέπει όλα τα προϊόντα που παράγονται από κάθε τμήμα να συνδέονται. Το έργο είναι η κατασκευή μίας αποθήκης δεδομένων όπου όλα τα συστατικά μέρη συνδέονται και λειτουργούν χωρίς επιπτώσεις.

Πώς γίνεται όμως η αξιολόγηση της αρχιτεκτονικής; Αρχικά πρέπει ο σχεδιαστής να κατανοήσει το αντικείμενο που κατασκευάζει και τη χρησιμότητά του.

Πρέπει να γνωρίζει τι επιθυμεί ο πελάτης – χρήστης να σχεδιάσει, ποια είναι η λειτουργικότητα του αντικειμένου και πως θα ενωθούν τα επιμέρους τμήματα ώστε το έργο να λειτουργεί σύμφωνα με τις απαιτήσεις. Έτσι, για κάθε επιμέρους τμήμα χρειάζεται να γνωρίζει πως χρησιμοποιεί κάθε τεχνική και πώς η αρχιτεκτονική της αποθήκης δεδομένων που σχεδιάζεται θα ενσωματώσει αυτή την τεχνολογία στο ίδιο το έργο και στα παράγωγά του.

Σε δεύτερο στάδιο παρουσιάζεται ο τρόπος με τον οποίο χτίζεται το έργο. Τα στάδια κατασκευής του έργου φαίνονται στην παρακάτω εικόνα και είναι τα εξής: Απαιτήσεις, Ανάλυση, Σχεδιασμός, Κατασκευή, Έλεγχος Ανάπτυξης και Παράδοση.

						User access
						Target data
	Τι υπάρχει ήδη και τι πρέπει να δημιουργηθεί;	Ποιος είναι ο καλύτερος τρόπος να δημιουργηθεί το επιθυμητό σύστημα;	Πώς το νέο σύστημα θα αξιοποιήσει τις υπάρχουσες υποδομές;	Είναι το σύστημα έτοιμο για τον τελικό χρήστη; Πως είναι η απόδοση του;	Πώς θα αξιοποιηθεί πλήρως το σύστημα ο νέος πελάτης;	Data movement & quality
Τι θέλουν οι χρήστες;						Source data
Απαιτήσεις	Ανάλυση	Σχεδιασμός	Κατασκευή	Έλεγχος	Παράδοση	

Αφού έχει αναλυθεί το αντικείμενο και ο τρόπος που κατασκευάζεται, πρέπει να διευκολυνθούν οι διαδικασίες ώστε να είναι αναγκαίος ο λιγότερο δυνατός χρόνος. Άρα οι διαδικασίες πρέπει να γίνουν όσο το δυνατό πιο απλές και να χρησιμοποιούνται εργαλεία που κάνουν τον κύκλο των διεργασιών ακόμα ταχύτερο. Συνεπώς, σκοπός είναι η αυτοματοποίηση όλων των επιμέρους διεργασιών.

Η τελευταία παράμετρος είναι ότι πρέπει πάντα η αρχιτεκτονική ενός έργου αποθήκης δεδομένων να είναι αφηρημένη, δηλαδή να είναι γενική, ώστε να είναι δυνατό να προσαρμοστεί στις τεχνολογικές εξελίξεις. Αυτή η παράμετρος είναι αλληλένδετη με την παραπάνω, καθώς έτσι θα μπορούν να χρησιμοποιούνται τα πιο εξελιγμένα εργαλεία και να αυτοματοποιείται ακόμα περισσότερο η διαδικασία κατασκευής.

## ΚΕΦΑΛΑΙΟ 18

Για την εύρυθμη λειτουργία του έργου είναι απαραίτητο για τους μελλοντικούς χρήστες να:

- Εμπλακούν σε πρώιμο στάδιο με το λειτουργικό τμήμα του έργου και να το χρησιμοποιούν πλήρως σε κανονικές συνθήκες εργασίας.
- Προτείνουν προς αλλαγή ότι δε λειτουργεί σωστά, ή πιο αργά από ότι θα έπρεπε.



- Παρέχουν πληροφορίες για τις διαφορετικές επιλογές αλληλεπίδρασης με τα χαρακτηριστικά του εργαλείου.
- Βοηθούν στην ανάλυση των δεδομένων προέλευσης και να αποκωδικοποιούν ότι δεν μπορεί να κατανοήσει ο κατασκευαστής.
- Συνεργάζονται με την ομάδα Διασφάλισης Ποιότητας ώστε να φροντίζουν οι απαιτήσεις να υλοποιούνται σωστά.
- Συμμετέχουν στις συχνές συναντήσεις του προσωπικού ώστε να συζητούν για θέματα εκπαίδευσης, υλοποίησης και τις επόμενες φάσεις της ανάπτυξης.

### 3.7 Μάθημα 7<sup>ο</sup> – Εξωτερικές πηγές δεδομένων

Στο έβδομο μάθημα παρουσιάζονται πληροφορίες σχετικά με την ανάλυση των πρωταρχικών πηγών από όπου θα προέλθουν τα δεδομένα καθώς και όλα τα απαραίτητα βήματα: Απαιτήσεις, ανάλυση, σχεδιασμός, κατασκευή, έλεγχος ανάπτυξης, για την υλοποίηση ενός Data Warehouse.

7



**Μάθημα 7ο: Εξωτερικές πηγές δεδομένων**

Στο μάθημα αυτό θα παρουσιάσουμε τις εισροές που χρησιμοποιούνται για να τροφοδοτήσουν τις αποθήκες δεδομένων και ονομάζονται Source Data Feeds, δηλαδή εξωτερικές πηγές δεδομένων.

□

📄 19. Η αλυσίδα αξίας των πληροφοριών

📄 20. Ενσωμάτωση εξωτερικών Data Warehouses

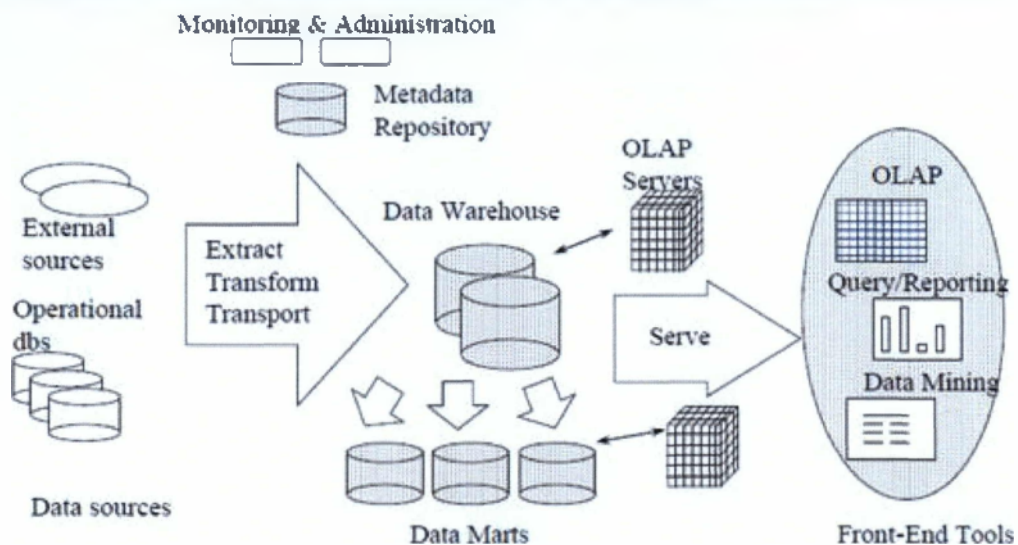
Εικόνα 3-8 Παρουσίαση 7<sup>ου</sup> μαθήματος

## ΚΕΦΑΛΑΙΟ 19

Στις προηγούμενες ενότητες μιλήσαμε για τη συλλογή δεδομένων από συγκεκριμένες δομές δεδομένων χωρίς να συμπεριλάβουμε τις εξωτερικές πηγές δεδομένων. Δυστυχώς μια αποθήκη δεδομένων θα είναι ημιτελής μέχρι να περιλαμβάνει πληροφορίες από εξωτερικές πηγές. Στο παρακάτω σχήμα φαίνεται ότι οι εξωτερικές

πηγές (external sources) είναι διαφορετικές από τις πηγές δεδομένων (είτε είναι σχεσιακές βάσεις είτε όχι). Η διαδικασία συλλογής των εξωτερικών δεδομένων έχει πολλά κοινά σημεία με τη διαδικασία συλλογής, τροποποίησης και ταξινόμησης που έχουμε ήδη περιγράψει.

## Data Warehousing Architecture



### Εξωτερικά δεδομένα

Τα εξωτερικά δεδομένα είναι απαραίτητα για ένα απλό λόγο: για να διασφαλίσετε ότι λαμβάνετε τη σωστή επιχειρησιακή απόφαση, πρέπει να έχετε δει μακροσκοπικά και μακροπρόθεσμα τον κλάδο, κάτι που συνήθως σημαίνει ότι δε μπορείτε να βρείτε όλες τις απαντήσεις στις εφαρμογές και στις βάσεις δεδομένων της εταιρίας.

Αφού εντοπίσετε όλα τα νέα δεδομένα από τις εξωτερικές πηγές, κρίνεται σκόπιμο να εφαρμόσετε κάποια κριτήρια ως προς την επιλογή τους τα οποία συνήθως σχετίζονται άμεσα με τους χρήστες. Αρχικά πρέπει να επανακαθοριστεί η λίστα των ατόμων που χρησιμοποιούν την αποθήκη δεδομένων και να καθοριστούν οι πληροφορίες που ενδέχεται να χρειαστούν. Υπάρχει κάποια πληροφορία που δε βρίσκεται στη Β.Δ του συστήματος και δεν είναι διαθέσιμη στους υπαλλήλους; Κάνοντας μια συνέντευξη με τους υπαλλήλους δημιουργείται μια λίστα με δεδομένα που χρησιμοποιούν όπως: πληροφορίες, τιμές, περιορισμούς και πληροφορίες επαφών

### Έλεγχος εξωτερικών δεδομένων

Θεωρούμε ότι έχουμε βρει τα δεδομένα που χρειαζόμαστε από εξωτερικές πηγές. Πρέπει να ελεγχθεί η ποιότητά τους με κάποιες βασικές τεχνικές. Αρχικά ελέγχουμε τα αρχεία προέλευσης των δεδομένων και από που προέρχονται. Αν είναι συνδεδεμένα και οι τιμές ενημερώνονται από αυτά τότε πρέπει να βρεθεί ένας τρόπος ώστε να ενημερώνονται και τα δεδομένα που θα μπουν στην αποθήκη δεδομένων. Επίσης αν δεν υπάρχουν τιμές ελέγχου (check values), πρέπει να τις ζητήσει ο σχεδιαστής. Έτσι θα καθοριστούν οι τιμές που είναι επιτρεπτές να έχει κάθε δεδομένο. Τέλος, κατά τη διάρκεια ελέγχου των διαδικασιών, παρακολουθούνται οι παρακάτω ιδιότητες: υπάρχουν πρωτεύοντα κλειδιά στη βάση που συνδέουν μονοσήμαντα τα δεδομένα; Είναι σωστό το εύρος των τιμών που έχουν; Λείπουν αρχεία; Γενικά ο αναλυτής πρέπει να ψάξει για παράδοξα, μπερδεμένα αποτελέσματα, ανωμαλίες, μη συμβατότητες και ότι άλλο φαίνεται περίεργο και να αναζητήσει την πηγή αυτού του δεδομένου.

Το επόμενο στάδιο αυτής της διαδικασίας είναι η οργάνωση των νέων πληροφοριών που έχουν δημιουργηθεί. Όπως έχει ήδη αναφερθεί, συνήθως γίνεται μια σύντηξη των δεδομένων αν και σε αυτή την περίπτωση μπορούν να χρησιμοποιηθούν εξωτερικοί χώροι αποθήκευσης, τα data warehouse auxiliary (μηχανισμοί που δε χρησιμοποιούνται σε καθημερινή βάση). Τα data warehouse auxiliaries περιέχουν δεδομένα που είτε μπορούν να θεωρηθούν ως μια βάση είτε ως αποθηκευτήριο πληροφοριών.

#### Πώς ενσωματώνονται τα νέα δεδομένα στο Data Warehouse?

Υπάρχουν 4 τρόποι καταχώρησης των δεδομένων:

- **1. Ολική Αντικατάσταση:** Τα νέα δεδομένα εγγράφονται πάνω στα παλιά δεδομένα. Αφού ανανεώνονται συνεχώς από τον πάροχο δεδομένων, ακόμα και τα ιστορικά δεδομένα πρέπει να είναι διαφορετικά από αυτά που εμφανίστηκαν την τελευταία φορά.
- **2. Προσάρτηση:** Τα νέα δεδομένα προσαρτώνται στα παλιά δεδομένα. Τα ιστορικά δεδομένα δεν αλλάζουν, ο πάροχος δίνει μόνο τις νέες λεπτομέρειες.

- **3. Μερική Προσάρτηση:** Παρότι απλά προστίθενται τα νέα δεδομένα, καθώς προχωρούμε στη βάση διαγράφουμε επιλεκτικά απαρχαιωμένα δεδομένα για να απελευθερωθεί χώρος για τα επόμενα. Έτσι μπορεί να διαγραφούν δεδομένα που έπρεπε να είχαν διαγραφεί καθώς μπορεί να μπερδέψουν ένα χρήστη.
- **4. Επιμέρους Ενημέρωση:** Κάθε νέο δεδομένο αντικαθιστά μια πλειάδα δεδομένων που ήδη υπάρχουν και παράλληλα εκτελείται η εντολή SQL Update.

Η απόκτηση νέων δεδομένων μπορεί να πραγματοποιηθεί επίσης με πολλούς τρόπους, είτε βρίσκοντας εξωτερικούς παράγοντες που παρέχουν τις πληροφορίες που αναζητούμε, είτε από εταιρίες που ανταλλάσσουν πληροφορίες ως προϊόντα. Όλες οι πληροφορίες που είναι διαθέσιμες είτε από ιδιώτες, είτε στον ιστό, είναι συνήθως λόγω ευγενούς ανταγωνισμού καθώς δίνουν κίνητρο για την αύξηση αποδοτικότητας και λειτουργικότητας των εταιριών. Σε κάθε περίπτωση είναι βασικό να μπορεί κάθε αλλαγή στις πληροφορίες αυτές να είναι εντοπίσιμη και ακόμα καλύτερα, προβλέψιμη.

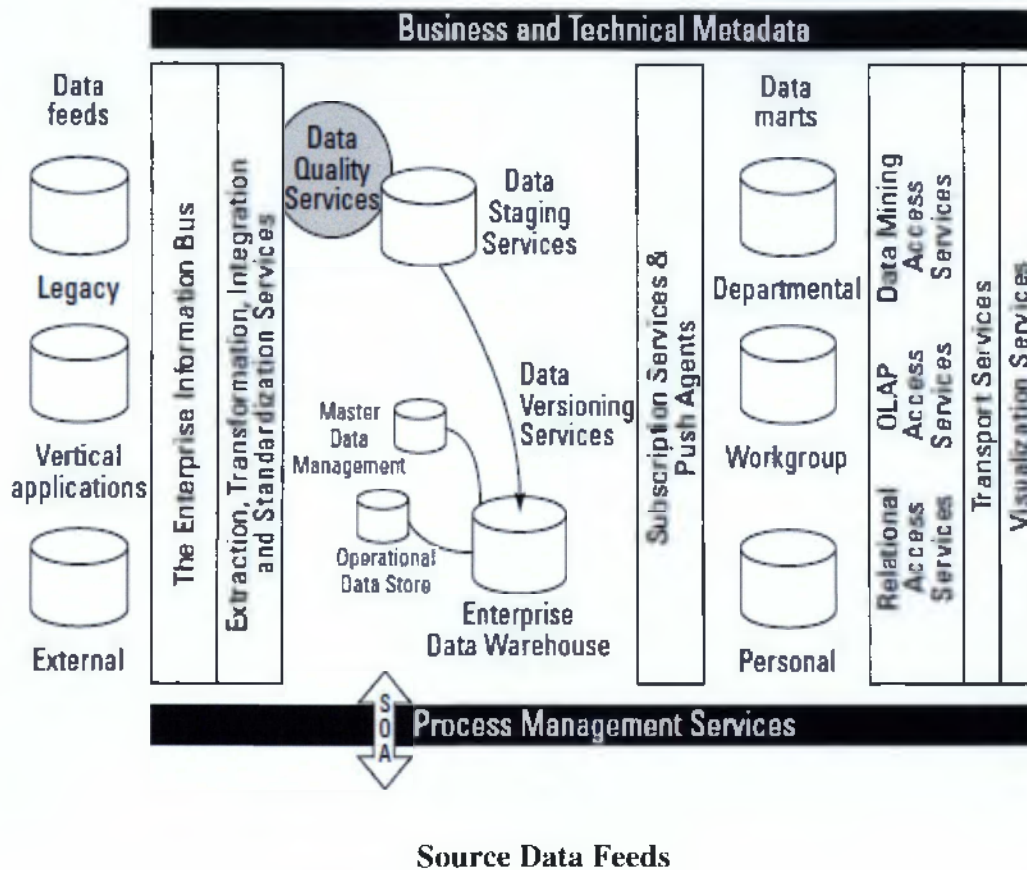
## ΚΕΦΑΛΑΙΟ 20

Η έννοια του Data Warehouse προϋποθέτει ότι υπάρχουν εξωτερικές πηγές που το τροφοδοτούν συνεχώς. Έτσι οι σχέσεις μεταξύ τους μπορούν να συνοψιστούν στα εξής: αν γίνει κάποια αλλαγή στα δεδομένα, στην αποθήκη δεδομένων πρέπει η αλλαγή (είτε είναι μετασχηματισμός, είτε τροποποίηση, είτε απλή ενημέρωση) να γίνεται και στις εξωτερικές πηγές των δεδομένων. Αλλαγές σε μια εφαρμογή που παρέχει δεδομένα στην αποθήκη δεδομένων, μπορεί να επηρεάσει το περιεχόμενό του Data Warehouse και τη διαδικασία μεταφοράς των πληροφοριών σε αυτό. Γενικά τα Data Warehouses είναι μέρη ενός συστήματος του οποίου τα τμήματα επικοινωνούν και πρέπει να εξετάζονται από άκρη σε άκρη κάθε φορά που προκύπτει μια αλλαγή.

Μια αποθήκη δεδομένων είναι εκ των πραγμάτων ένα φυσικό, καταναμημένο κατάστημα δεδομένων. Η κατανομή των πληροφοριών βοηθά στην λειτουργικότητα και χρηστικότητα του συστήματος και της εταιρίας. Το υψηλότερο επίπεδο χρησιμότητας είναι ο μεγαλύτερος στόχος και αποστολή του σχεδιαστή. Στην παρακάτω εικόνα φαίνονται τα τμήματα ενός Data Warehouse που μπορεί να

περιέχουν πηγές δεδομένων (data sources), την αποθήκη δεδομένων, ένα "κατάστημα" δεδομένων (operational data store), data marts και master data.

Στη συνέχεια θα αναλυθεί τι περιέχουν αυτές τις δομές.

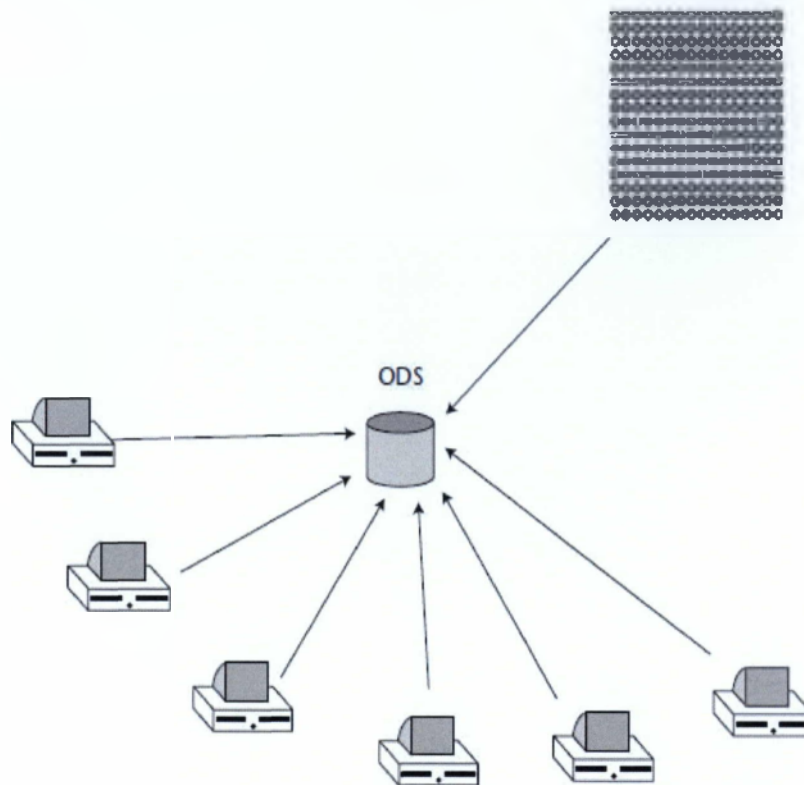


Οι εισροές που τροφοδοτούν τις αποθήκες δεδομένων ονομάζονται Source Data Feeds και είναι οι εφαρμογές που "τρέχουν" κάθε στιγμή στην επιχείρηση ή οι εξωτερικές πηγές δεδομένων. Παρότι η ομάδα που κατασκευάζει την αποθήκη δεδομένων δε σχετίζεται με την αρχιτεκτονική των Data Feeds, είναι βασικό να τη γνωρίζει. Η πιο δύσκολη επιλογή που πρέπει να κάνει ο κατασκευαστής ενός DW είναι αυτή των εξωτερικών πηγών δεδομένων, αφού αν τα δεδομένα δεν είναι καλής ποιότητας ή δεν είναι ευρέως διαθέσιμα δεν μπορούν να υποστηρίξουν ένα υψηλής ποιότητας Data Warehouse.

**Operational Data Store**

Ο όρος αυτός είναι από τους πιο βασικούς στην αρχιτεκτονική μιας αποθήκης δεδομένων και ταυτοχρόνως δύσκολος στην κατανόηση. Ουσιαστικά μπορεί να θεωρηθεί ως μια κλασική αποθήκη δεδομένων που δέχεται περιοδικά εισροές από διάφορες εξωτερικές πηγές αντικαθιστώντας τις ήδη υπάρχουσες κατά την αποθήκευση. Στη συγκεκριμένη περίπτωση είναι ένας αποθηκευτικός χώρος που επίσης παρέχει πληροφορίες στην αποθήκη δεδομένων, μέσω εφαρμογών που "ωθούν" ή δέχονται τα δεδομένα (αμοιβαία συσχέτιση), ώστε να ανανεώνεται συχνά το περιεχόμενό της. Ένας πιο εκτενής ορισμός είναι ότι ένα Operational Data Store είναι ένα πληροφοριακό περιβάλλον που κάθε στιγμή αναπαριστά την παρούσα κατάσταση του περιεχομένου του, παρότι τα δεδομένα που περιέχει μπορεί να τα διαχειρίζονται διαφορετικές και απομακρυσμένες εφαρμογές στην εταιρία. Τα ODS παρέχουν στους χρήστες τους μια πάγια, σχεδόν στιγμιαία εικόνα των διαφορετικών δεδομένων που υποστηρίζουν ένα συγκεκριμένο επιχειρησιακό θέμα.

Άρα ένα ODS δεν μπορεί να περιέχει ιστορικά δεδομένα; Εξ' ορισμού δεν μπορεί, αν περιέχει ιστορικά δεδομένα είναι αποθήκη δεδομένων, όχι ODS. Υπάρχουν περιπτώσεις κατά τις οποίες τα ιστορικά δεδομένα περιέχονται στα ODS ως περιλήψεις (summarization tables).



Μια βασική πρόκληση στην κατασκευή ενός ODS είναι οι επιχειρησιακές διεργασίες και ο τρόπος που θα αλλάξουν αν χρειαστεί. Υπάρχουν δύο προσεγγίσεις για τους χρήστες μιας τέτοιας αποθήκης δεδομένων όταν πρέπει να περιοριστούν οι δυνατότητές τους: είτε διαγράφεται τελείως ο τύπος λειτουργικότητας από την εφαρμογή που χρησιμοποιείται, είτε αποκλείεται για μερικούς χρήστες. Το δεύτερο σενάριο είναι να ενσωματωθούν στην εταιρία όλες οι διαδικασίες ανατροφοδότησης του ODS. Αυτή η διαδικασία είναι αρκετά πολύπλοκη γιατί πρέπει να καθοριστεί η αμοιβαία συσχέτιση όλων των εφαρμογών με τα δεδομένα και το ODS. Αυτή η προσέγγιση προκαλεί δικτυακή συμφόρηση και αμοιβαίες εξαρτήσεις (οι εξαρτήσεις αυξάνουν την πολυπλοκότητα του συστήματος).

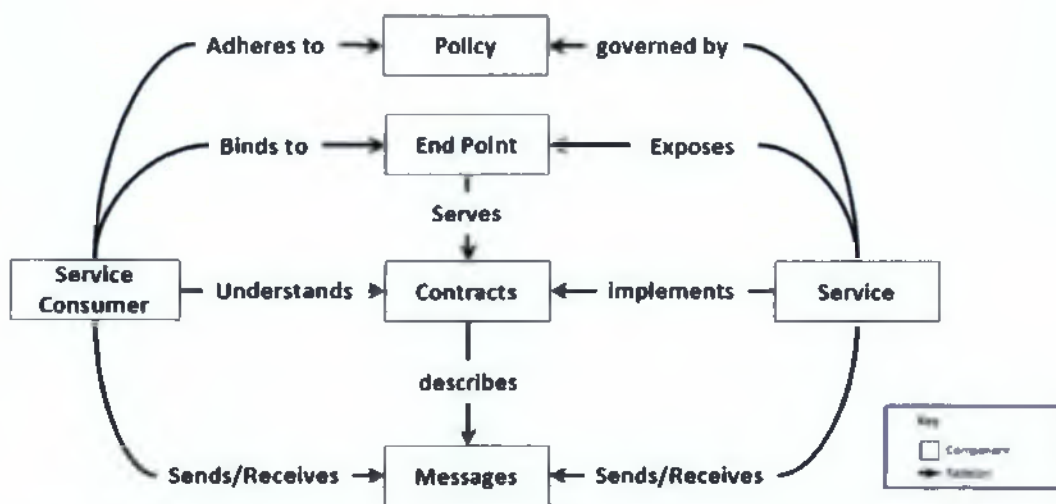
### Master Data Management

Τα δεδομένα που εισέρχονται στις αποθήκες μπορεί να προέρχονται από διαφορετικές αναφορές. Πως διασφαλίζεται η μοναδικότητά τους και η αποφυγή της άσκοπης επανάληψής τους; Σε βασικό επίπεδο, γίνεται διαχείριση των δεδομένων μέσω της αρχιτεκτονικής του Master Data Management. Αρχικά προσπαθεί να κατοχυρώσει και να εξαλείψει τα πολλαπλά αντίτυπα των κοινών αναφορών στα

δεδομένα από τις διαφορετικές εφαρμογές του συστήματος ή από μέρη του οργανισμού. Οι διαδικασίες που περιλαμβάνονται στο MDM είναι η αναγνώριση της πηγής δεδομένων, η συλλογή, η μετατροπή, η κανονικοποίηση δεδομένων, η κατανομή κανόνων, ο εντοπισμός και η διόρθωση σφαλμάτων, η σταθεροποίηση, η αποθήκευση, η διανομή και η διαχείριση των δεδομένων. Όλες αυτές οι διεργασίες χρησιμοποιούνται επίσης στη λογική εξέλιξη τροφοδοσίας των ODS. Ένα σύνολο εργαλείων έχει δημιουργηθεί ώστε να επιτρέπεται η αποτελεσματική συνδιαλλαγή των δεδομένων αναφορών με άλλα συστήματα. Έτσι υπάρχουν διεργασίες που διευκολύνουν τον εντοπισμό και τη διαγραφή διπλότυπων εγγραφών .

### Service Oriented Architecture (SOA)

Εφόσον έχετε ήδη ενσωματώσει δεδομένα μέσω των ODS και MDM, χρειάζεται ένα μέσο ανταλλαγής μηνυμάτων ή επικοινωνίας, μια αρχιτεκτονική που να επιτρέπει ακόμα και σε συστήματα που δεν κατασκευάστηκαν με αυτό το τρόπο, να μπορούν να επικοινωνούν μεταξύ τους. Μόλις μπήκατε στον κόσμο της SOA.



Η SOA είναι μια μέθοδος για ανάπτυξη και ενσωμάτωση συστημάτων στα οποία η λειτουργικότητα είναι συγκεντρωμένη στις επιμέρους διεργασίες. Επίσης περιγράφει μια υποδομή που επιτρέπει σε διαφορετικές εφαρμογές να ανταλλάσσουν δεδομένα μεταξύ τους, ενώ συμμετέχουν στην επιχειρησιακή διαδικασία. Η αρχιτεκτονική αυτή χωρίζει τις διεργασίες σε διακριτές μονάδες, ή υπηρεσίες, που είναι προσβάσιμες από ένα δίκτυο ώστε οι εφαρμογές που σχετίζονται με την επιτήρηση και την κυρίως



λειτουργία του συστήματος να συνδυάζονται και να χρησιμοποιούν ξανά αυτές τις λειτουργίες. Τελικά, αυτές οι υπηρεσίες ενσωματώνονται στην εταιρία. Επίσης επικοινωνούν μεταξύ τους μεταφέροντας δεδομένα από μια υπηρεσία σε μια άλλη, ή διαχωρίζοντας μια δραστηριότητα σε δύο ή παραπάνω υπηρεσίες. Η επικοινωνία μέσω μηνυμάτων συνήθως δεν είναι συγχρονισμένη με αυτή μεταξύ των περιβαλλόντων. Το MDM και η SOA παρέχουν την απαραίτητη τεχνολογία για την ανατροφοδότηση της αποθήκης δεδομένων που κατασκευάζετε.

### 3.8 Μάθημα 8<sup>ο</sup> - Διαχείριση ενός έργου Data Warehouse και συμβουλές

Το όγδοο μάθημα παρουσιάζει τη διαχείριση ενός έργου από τη διοίκηση της εταιρίας, την αντιμετώπιση των υπάρχουσών δομών αποθήκης δεδομένων καθώς και ορισμένες συμβουλές για την αγορά και υλοποίηση ενός Data Warehouse.

8



#### Μάθημα 8ο: Διαχείριση ενός έργου Data Warehouse και συμβουλές

Στο μάθημα αυτό θα παρουσιάσουμε τη διαχείριση ενός έργου από τη διοίκηση της εταιρίας, την αντιμετώπιση των υπάρχουσών δομών αποθήκης δεδομένων καθώς και ορισμένες συμβουλές για την αγορά και υλοποίηση ενός Data Warehouse.

Σαν εργασία για τους φοιτητές προτείνουμε την αναζήτηση πληροφοριών για το κόστος αγοράς και συντήρησης υπηρεσιών Data Warehouse από Ελληνικές και Διεθνείς εταιρίες. Το αποτέλεσμα θα προστεθούν στην ύλη του κεφαλαίου αυτού.

- ☐ 21. Πως αντιμετωπίζει η διοίκηση της εταιρίας το έργο ενός Data Warehouse;
- ☐ 22. Ήδη υπάρχουσες δομές αποθήκευσης. Αντικατάσταση ή Ανανέωση;
- ☐ 23. Μερικές Συμβουλές για την αγορά ενός Data Warehouse.
- ☐ 24. Συμβουλευτική για Data Warehousing

Εικόνα 3-9 Παρουσίαση 8<sup>ου</sup> μαθήματος

Όπως κάποιος μπορεί να δει από την παραπάνω εικόνα, προτείνεται στους φοιτητές που παρακολουθούν το μάθημα να αναζητήσουν πληροφορίες για το κόστος αγοράς

και συντήρησης υπηρεσιών Data Warehouse από Ελληνικές και Διεθνείς εταιρίες. Τα αποτελέσματα θα προστεθούν στην ύλη του κεφαλαίου αυτού.

## ΚΕΦΑΛΑΙΟ 21

Δυστυχώς οι περισσότεροι μάντζερς δεν έχουν ασχοληθεί ιδιαίτερα με τις απαιτήσεις, τις ανάγκες και τα διάφορα στάδια από τα οποία περνάει το έργο μέχρι την ολοκλήρωσή του. Σίγουρα όμως πρέπει να ενημερώνονται έτσι ώστε να μπορούν να κατανοούν και να προωθούν το έργο. Για το σκοπό αυτό φροντίστε να τους ενημερώνετε συχνά για την πρόοδο του έργου, με αναφορές, στατιστικά και οικονομικά αποτελέσματα καθώς επίσης και για τα μελλοντικά σχέδια.

Μια νέα τάση που παρατηρείται είναι οι εταιρίες του ίδιου κλάδου, όπως φαρμακευτικές εταιρίες, να συνεργάζονται μεταξύ τους ώστε να έχουν ένα ενιαίο Data Warehouse για την αύξηση των κερδών τους και τη διευκόλυνσή τους. Όποιος και αν είναι ο κλάδος ειδίκευσης και το καθεστώς της εταιρίας, κάθε οργανισμός μπορεί να συσχετισθεί σε μια συνεργασία πολλών συνεταιίρων, στην οποία η συνεργασία και ο διαμοιρασμός των πληροφοριών είναι το κλειδί της επιτυχίας. Όπως είναι λογικό ένα τέτοιο έργο είναι πολύ πιο πολύπλοκο στην κατασκευή του, από ότι ένα απλό data warehouse αφιερωμένο σε μια εταιρία. Μερικές φάσεις κατασκευής είναι ίδιες, όπως η εξαγωγή και η συλλογή δεδομένων και ο μετασχηματισμός τους. Πρέπει επίσης να αναθεωρήσετε τις απόψεις σας σε θέματα όπως: τα εταιρικά πρότυπα, το σύνολο των ευαίσθητων δεδομένων και η ιδιοκτησία τους (κάποια δεδομένα είναι ευαίσθητα για τη μία εταιρία αλλά όχι για την άλλη, άρα πρέπει να υπάρχουν κανονισμοί ώστε να μην έχει πρόσβαση σε αυτά όλο το προσωπικό), θέματα ασφάλειας, το κόστος συντήρησης, η μεθοδολογία ανάπτυξης και υλοποίησης, ο τρόπος διαχείρισης των διαφωνιών και βέβαια η ενίσχυσή του κατά τη διάρκεια της χρήσης του.

**ΚΕΦΑΛΑΙΟ 22**

Κάθε οργανισμός έχει στο πληροφοριακό του σύστημα μια δομή όπου αποθηκεύονται τα δεδομένα που χρησιμοποιεί, δηλαδή ένα σύστημα που παρέχει όλες τις απαραίτητες πληροφορίες και μερικές φορές αναλυτικές ικανότητες σε μια ομάδα ατόμων- χρηστών. Η πραγματικότητα είναι ότι αυτά τα μικρά αποθηκευτήρια, συνήθως βοηθούν ένα πολύ μικρό μέρος του πληθυσμού και δεν κατασκευάζονται με έναν καθιερωμένο τρόπο ώστε να υποστηρίζουν και να μπορούν να επεκτείνουν την αποτελεσματικότητά τους. Σε μερικές εταιρίες η δομή αυτή υπάρχει και διευκολύνει τους εργαζόμενους ενώ σε άλλες όχι.

Αρχικά, πρέπει να γίνει μια πλήρης μελέτη όλων των δομών αποθήκευσης που υπάρχουν σε όποια ηλεκτρονική μορφή και να είναι, τι περιέχουν, ποιος τα χρησιμοποιεί και γιατί. Συνήθως η διαδικασία αυτή διευκολύνεται από την παρακολούθηση και τη μελέτη των αναφορών της εταιρίας όπως επίσης και από τη συνομιλία με τους υπαλλήλους ώστε να κατανοήσουμε ποιος δημιουργεί αυτούς τους χώρους (ή ποιος έχει δικαιοδοσία να τους δημιουργήσει). Το ερωτηματολόγιο που θα χρησιμοποιηθεί κατά τη διάρκεια των συνεντεύξεων παίζει πρωταρχικό ρόλο στη συλλογή των αναγκαίων και απαραίτητων πληροφοριών. Αφού δημιουργήσετε έναν κατάλογο που περιέχει αναλυτικά όλα τα συμπεράσματά σας από τις παραπάνω διαδικασίες, πρέπει να καταγράψετε σε μορφή αναφοράς το σκοπό και το τι καλείται να αντικαταστήσει το νέο Data Warehouse που θα κατασκευάσετε. Με τον τρόπο αυτό στην πορεία του έργου σας θα μπορείτε να παρακολουθείτε την πρόοδο σε σχέση με τις προηγούμενες δομές και κατά πόσο αυτές επικαλύπτονται για να εξυπηρετούν πολλές απαιτήσεις.

**Η ώρα των αποφάσεων**

Κατά την προηγούμενη φάση ανακαλύψατε πολλούς χώρους αποθήκευσης δεδομένων και τώρα πρέπει να γίνει ο διαχωρισμός μεταξύ αυτών που θα διατηρήσετε και αυτών που θα αντικαταστήσετε. Όπως ήδη αναφέρθηκε οι περιπτώσεις είναι οι εξής:

- Διαγραφή: Η επιλογή αυτή γίνεται αφού συμφωνήσουν όλοι ομόφωνα (και επίσημα) ότι τα δεδομένα προς διαγραφή δε χρησιμοποιούνται πουθενά απολύτως.

- **Αντικατάσταση:** Για να γίνει αντικατάσταση των δεδομένων πρέπει αρχικά, να διατηρηθούν τα δεδομένα που είναι χρήσιμα στους εργαζόμενους και δε χρειάζονται αντικατάσταση, να γίνει σχεδιασμός για τη λειτουργικότητα του σχήματος και να αντικατασταθούν μόνο τα συγκεκριμένα τμήματα, διατηρώντας τα άλλα στο νέο Data Warehouse.

- **Διατήρηση για μελλοντική μετατροπή:** Οι περιπτώσεις στις οποίες κρίνεται σκόπιμη η διατήρηση των δεδομένων είναι: αν η ποιότητά τους είναι εξαιρετική, αν είναι αποθηκευμένες σε ένα αρχείο και όχι σε σχεσιακή βάση, αν τα συγκεκριμένα δεδομένα είναι πρόσφατη προσθήκη και αν χρησιμοποιούνται σχετικά συχνά. Οι μετατροπές που μπορούν να γίνουν είναι: είτε προσθήκες που σχετίζονται με τα θέματα που ήδη υπάρχουν, είτε για τη βελτιστοποίηση της δομής αποθήκευσης, είτε για την ανανέωση των δεδομένων, είτε για την προσθήκη νέων εργαλείων για τη διεύρυνση και διευκόλυνση της ανάλυσης αυτών.

Σε κάθε περίπτωση αυτό που πρέπει πάντα να υπολογίζεται ως χρυσός κανόνας κατά τη διάρκεια του σχεδιασμού και της μεταφοράς δεδομένων είναι: **Να μην υποβιβάζεται η λειτουργικότητα του συστήματος.**

## ΚΕΦΑΛΑΙΟ 23

### Μερικές Συμβουλές για την αγορά ενός Data Warehouse.

- Κάντε μελέτη όλων των πιθανών λύσεων και προϊόντων που υπάρχουν και σας εξυπηρετούν.
- Κάντε πολλές ερωτήσεις στους προμηθευτές.
- Να είστε σκεπτικιστές.
- Μη βιαστείτε να αποφασίσετε.
- Φροντίστε ώστε αυτό που θα αποφασίσετε να αγοράσετε να προσφέρει ευελιξία για μελλοντικές χρήσεις.

**ΚΕΦΑΛΑΙΟ 24**

Η συμβουλευτική είναι ένας νέος κλάδος που έχει εισχωρήσει σε όλους τους τομείς των ΙΤ συστημάτων. Το ίδιο ισχύει και για τα συστήματα Αποθηκών Δεδομένων. Σίγουρα είναι καλύτερα για έναν οργανισμό να βασίζεται μόνο στους υπαλλήλους που έχει, αλλά αν θέλει να ανήκει στην πλειοψηφία, η τύχη ενός data warehouse εξαρτάται μόνο από την αναγνώριση και τη χρήση του από έναν εξειδικευμένο σύμβουλο.

Σε κάθε περίπτωση δεν είναι όλοι αξιόλογοι και η κρίση τους μπορεί να παρουσιάζει υποκειμενικότητα. Πως μπορεί όμως κάποιος χωρίς εμπειρία και γνώσεις να επιλέξει τον κατάλληλο σύμβουλο, ώστε να είναι κατοχυρωμένος ότι οι προτάσεις του δε θα προκαλέσουν ζημία στην εταιρία;

Συνήθως πρέπει να αναζητάτε από το σύμβουλό σας τα εξής χαρακτηριστικά:

- Να έχει στο ενεργητικό του αρκετές επιτυχημένες υλοποιήσεις Data Warehousing με χρήση πληθώρας τεχνολογιών (ώστε να γνωρίζει και να προτείνει ποια θα είναι κατάλληλη σε κάθε περίπτωση).
- Μέσω της πείρας του να έχει τη διορατικότητα της επιλογής της τεχνολογίας και της αρχιτεκτονικής για τα Data Warehouses, όχι για τη χρήση συγκεκριμένων μεθόδων (επειδή πάντα το έκανε έτσι), αλλά για να έχει ευελιξία μεταξύ των προτεινόμενων λύσεων.
- Να δεσμεύεται μέσω των αποφάσεών του στην επιτυχία του έργου λαμβάνοντας πλήρη υπευθυνότητα για όλες τις διαστάσεις της υλοποίησης.

Σε κάθε περίπτωση πρέπει να υπολογίζετε πάντα το ένστικτό σας! Η τεχνολογική πρόοδος είναι ραγδαία για να περιορίζεστε μόνο στην άποψη μιας αυθεντίας που πληρεί τις παραπάνω προϋποθέσεις. Μην επαναπαύεστε.

### 3.9 Μάθημα 9<sup>ο</sup> – Λογισμικό για χρήση και συντήρηση Data Warehouses

Στο ένατο μάθημα παρουσιάζονται χρήσιμα εργαλεία για την υποστήριξη μιας Αποθήκης Δεδομένων και οι ανάγκες που αφορούν τη συντήρησή της.

9



#### Μάθημα 9ο: Λογισμικό για χρήση και συντήρηση Data Warehouses

Στο μάθημα αυτό θα παρουσιάσουμε χρήσιμα εργαλεία για την υποστήριξη μιας Data Warehouse. Επίσης θα γίνει αναφορά στις ανάγκες όσον αφορά τη συντήρηση μιας Αποθήκης Δεδομένων.

25. Θέματα συντήρησης και εργαλεία λογισμικού

26. Σημσιολογικός Ιστός

Εικόνα 3-10 Παρουσίαση 9<sup>ου</sup> μαθήματος

#### Αξιολόγηση του Λογισμικού των Data Warehouses

Σε αυτή την ενότητα παρουσιάζονται μερικές ιδέες που μπορούν να κάνουν τη διαδικασία αξιολόγησης του λογισμικού data warehousing πιο αποτελεσματική. Παρόλα αυτά δεν αποτελεί έναν πλήρη κατάλογο των καθηκόντων που πρέπει να ακολουθείται κατά τη διαδικασία αξιολόγησης λογισμικού. Αντιθέτως, αυτά είναι τα σημεία που σπάνια συζητώνται ή εφαρμόζονται, σε μια περίοδο μεγάλου ενδιαφέροντος για τις αποθήκες δεδομένων.

- Αξιολογείστε το προϊόν μόνοι σας
- Ρωτήστε αν η ήδη υπάρχουσα τεχνολογία έχει τη δυνατότητα να εξυπηρετήσει το σύστημα για το συγκεκριμένο έργο.
- Πάρτε συστάσεις για το προϊόν σας από άλλους χρήστες.
- Αν πρόκειται να παρακολουθήσετε πολλές δοκιμαστικές αναπαραστάσεις, σχεδιάστε μια υπόθεση προς μελέτη με κάθε ένα εργαλείο και σημειώστε τα συμπεράσματά σας.
- Να είστε δύσπιστοι στις απόψεις περί του λογισμικού από αυθεντίες του χώρου.

Υπάρχουν σκοπιμότητες πίσω από όλα αυτά.

- Ρωτήστε για τις ανάγκες του λογισμικού όσον αφορά τη συντήρηση.
- Κατανοήστε τα μειονεκτήματα του κάθε λογισμικού.
- Ελέγξτε την οικονομική σταθερότητα του πωλητή.
- Διαθέστε μια αντιπροσωπευτική ομάδα για να διεξάγει την αξιολόγηση.
- Αφήστε την αξιολόγηση των εργαλείων του χρήστη (user end tools) να γίνει από τους ίδιους τους χρήστες.

### **Θέματα Συντήρησης στα Συστήματα Data Warehousing**

Ένα σημαντικό θέμα στις αποθήκες δεδομένων και στα συστήματα στήριξης αποφάσεων είναι η συντήρησή τους. Σε αυτή την ενότητα παρουσιάζονται μερικά από τα προβλήματα που ενδέχεται να αντιμετωπίσετε κατά την κατασκευή αυτών των συστημάτων. Ο τρόπος που θα διευθετήσετε τα προβλήματα ποικίλει ανάλογα με το περιβάλλον. Χρειάζεται να γνωρίζετε, αφού είστε ο κατασκευαστής του συστήματος, πώς αλληλεπιδρά το σύστημά σας με τα άλλα, ώστε να σχεδιαστεί με τους καλύτερους αμυντικούς μηχανισμούς. Εάν είστε υπεύθυνος για ένα σύστημα που τροφοδοτείται από 10 πηγές, μπορεί αυτό να έχει περισσότερη έκθεση από αυτή που θα δεχόταν με το τυπικό σύστημα επεξεργασίας συναλλαγών. Αν και η χρήση της εξόρυξης των δεδομένων, του καθαρισμού, των εργαλείων φόρτωσης και των καταλόγων πληροφοριών μπορεί να διευκολύνουν σημαντικά την επιβάρυνση, πολλές αλλαγές απαιτούν ένα δίκαιο ποσό της προσπάθειας.

- Θα πρέπει να υπολογίσετε αν, τότε και πώς πρέπει να ξεκαθαρίζετε τα δεδομένα.
- Θα πρέπει να καθορίσετε ποια ερωτήματα και ποιες εκθέσεις πρέπει να είναι αυτοματοποιημένα γραπτές, καθώς και ποιες να είναι γραπτές από το χρήστη.
- Πολλές φορές θα αποθηκεύονται δεδομένα στο σύστημα μόνο και μόνο για την ικανοποίηση της αίσθησης της πληρότητας, διαδικασία που αυξάνει την πολυπλοκότητα της δομής.
- Είναι ανάγκη να υπάρχει ισορροπία μεταξύ της αποτελεσματικότητας που προσδίδουν οι δομές που κατασκευάζονται και του μεγέθους των αναγκών συντήρησης.
- Δε θα υπάρχει βοήθεια και καθοδήγηση κατά την επιλογή των εργαλείων, θα

πρέπει ο κατασκευαστής να είναι σε θέση να επιλέξει μόνος του.

- Θα πρέπει να δημιουργηθούν αποτελεσματικά μέσα μέτρησης της αποτελεσματικότητας των αλλαγών των δομών στους τελικούς χρήστες.
- Πρέπει να γίνει κατανοητή η αλληλεπίδραση του συστήματος αποθήκευσης δεδομένων και του συστήματος τροφοδοσίας - ανανέωσης περιεχομένου και πως αλληλοεπηρεάζονται σε επικείμενες αλλαγές.
- Η επιλογή των εργαλείων που θα χρησιμοποιηθούν πρέπει να γίνει μετά από διεξοδική μελέτη και δοκιμές, ώστε να βρεθεί το κατάλληλο εργαλείο που θα πληρεί τις προϋποθέσεις του έργου.
- Πριν την παράδοση του έργου και κατά τον έλεγχο του πρέπει να γίνει σωστή διαχείριση αλλαγών. Αυτό σημαίνει ότι το σύστημα πρέπει να είναι σε κάθε περίπτωση έτοιμο να αντιμετωπίσει διαφορετικές καταστάσεις των προγραμμάτων με τα οποία αλληλεπιδρά, καθώς και των επερωτημάτων που δέχεται από τους χρήστες.
- Σε ότι αφορά την αρχιτεκτονική και το σχεδιασμό της αποθήκης δεδομένων, είναι δυσκολότερη η συντήρησή τους παρά η αρχική εγκατάστασή τους.
- Οι μηχανισμοί και οι τεχνικές ασφάλειας που εφαρμόζετε στο έργο πρέπει να είναι οι κατάλληλες, δηλαδή ούτε πολύ αυστηρές ούτε πολύ χαλαρές, ώστε να μην επηρεάζεται η λειτουργικότητά του.

#### **Νέα είδη δεδομένων στην Αποθήκη Δεδομένων σας.**

Τα δεδομένα που υπάρχουν σε μια αποθήκη είναι τα δομημένα (structured) και τα αδόμητα (unstructured). Δομημένα θεωρούνται οι αριθμοί, το κείμενο, η ημερομηνία και η ώρα. Στις μέρες μας η μορφή των δεδομένων δεν είναι τόσο περιορισμένη, υπάρχουν εικόνες και πολυμεσικά αρχεία όπως ήχος και βίντεο, τα οποία ονομάζονται αδόμητα. Πλέον μέσω του Business Intelligence τα εργαλεία OLAP και όσα άλλα έχουμε αναφέρει ήδη, σχετίζονται με τη διαχείριση των αδόμητων δεδομένων. Για να γίνει περισσότερο κατανοητή η έννοια των δομημένων και αδόμητων δεδομένων θα αναφέρουμε μερικά παραδείγματα που είναι εξαιρετικά ενδιαφέροντα καθώς διευκολύνουν τους χρήστες των συστημάτων και αυξάνουν την αποδοτικότητα ολόκληρου του συστήματος.



Σκεφτείτε τα παρακάτω σενάρια: Μια φαρμακευτική εταιρία χρειάζεται να έχει πρόσβαση στα έγγραφα των εργαστηριακών ερευνών που πραγματοποιούνται εδώ και έτη και στις κλινικές δοκιμές που έχουν πραγματοποιηθεί για τη χορήγηση νέων φαρμάκων. Μια ασφαλιστική εταιρία χρειάζεται να αποθηκεύει και να μπορεί να έχει πρόσβαση σε παλαιότερα συμβόλαια, όπως και ένα τηλεφωνικό κέντρο χρειάζεται να αποθηκεύει κάθε κλήση ώστε αυτόματα να μπορεί να τις χρησιμοποιεί για ποιοτικό έλεγχο και για σωστή συμμόρφωση στους κανονισμούς. Τα παραδείγματα είναι πολλά και η βοήθεια που παρέχουν τα πολυμεσικά δεδομένα ανεκτίμητη. Αν χρησιμοποιείτε παραδοσιακά συλλογικές διεργασίες έργου, που εκτελούν δραστηριότητες όπως η διαχείριση της ροής εργασίας ή εικόνας, μπορείτε εύκολα να τις διευρύνετε, ώστε να συνδέονται με data warehousing διεργασίες που σας παρέχουν πρόσθετη αξία. Επιπροσθέτως, τα αποτελέσματα από τις επερωτήσεις που κάνατε μέχρι τώρα ως μέρος της αναλυτικής επεξεργασίας διεργασιών, μπορεί να λειτουργήσει ως ένα μέσο χρήσης πολυμεσικών δεδομένων που συμπληρώνουν τα δεδομένα που τυπικά διαχειρίζεστε.

### **Semantics - Σημασιολογικός Ιστός**

Semantics ή Σημασιολογία είναι η μελέτη της σημασίας, της έννοιας στην επικοινωνία, συμπεριλαμβανομένης της σημασίας μιας λέξης, ενός σημαδιού ή μιας πρότασης. Ουσιαστικά ο σημασιολογικός ιστός, διευκολύνει την πλοήγηση σε ότι αφορά τις εφαρμογές. Γενικά τα δύο τεχνικά χαρακτηριστικά του σημασιολογικού ιστού είναι: ένα κοινό πλαίσιο ενσωμάτωσης και συνδυασμού δεδομένων που προέρχονται από ποικίλες πηγές και μια κοινή γλώσσα για τη χαρτογράφηση των συσχετίσεων των δεδομένων- αντικειμένων. Πώς συνδέονται όμως όλα αυτά με τα data warehouses για τα οποία μιλάμε;

Μερικοί παλαιότερα προσπάθησαν να επιλύσουν το πρόβλημα των μεταδεδομένων (δεδομένα που περιέχουν δεδομένα) και της διαχείρισής τους, δημιουργώντας μια άλλη τεχνολογία, τα technical metadata repositories, που όμως απέτυχαν καθώς αποτελούν λύση απ' άκρη σε άκρη και χρειάζονται συνεχή συντήρηση, τεχνική και οικονομική. Πολλές νέες τεχνολογίες έχουν αναπτυχθεί για να βοηθήσουν σε αυτή τη περίπτωση. Για μερικές από αυτές μπορείτε να επισκεφτείτε τους παρακάτω

ιστότοπους:

- [www.balancedinsight.com](http://www.balancedinsight.com)
- [www.compositesw.com](http://www.compositesw.com)
- [www.expressor-software.com](http://www.expressor-software.com)

### **Πως θα προετοιμάσετε το έργο σας για αυτές τις τεχνολογίες;**

Για ένα έργο data warehouse υπάρχουν τρεις περιοχές που χρειάζεται να επικεντρωθεί η ομάδα κατασκευής. Το επίπεδο business intelligence semantic, η διαχείριση των business rules και πιθανόν ένας καλύτερος ορισμός των ομόσπονδων επερωτήσεων. Ουσιαστικά αυτό που πρέπει να φροντίσετε από την αρχή του έργου είναι να υπάρχει ένας υπεύθυνος, που θα μπορεί να παρέχει την πρωτεύουσα ετυμολογία για λέξεις κλειδιά που χρησιμοποιούνται στη φάση ανάλυσης των απαιτήσεων ή στην υλοποίηση του έργου. Ξεκινήστε να καθορίζετε τα semantics και ευθυγραμμίστε τα με τη ροή παραγωγής της επιχείρησής σας.

Επιπροσθέτως κρίνεται σκόπιμο να παρακολουθήσετε τη διαδικασία σχετικά με τους business rules, πώς καθορίζονται και πώς τους παράγετε. Η διαδικασία καθορισμού των κανόνων είναι παρόμοια με αυτή του καθορισμού των διεργασιών του business intelligence semantics.

### **3.10 Μάθημα 10<sup>ο</sup> - Ανακοινώσεις, Επανάληψη, Συζητήσεις, Αξιολόγηση Μαθήματος**

Στο δέκατο μάθημα θα αναφερθούμε σε συζητήσεις, επαναλήψεις, ερωτήσεις και

απαντήσεις. Οι φοιτητές θα αξιολογήσουν το μάθημα μέσα από ένα ηλεκτρονικό ερωτηματολόγιο και θα παρουσιαστούν πιθανές βελτιώσεις σχετικά με το μάθημα.

10



### Μάθημα 10ο: Ανακοινώσεις, Επανάληψη, Συζητήσεις, Αξιολόγηση Μαθήματος

Στο μάθημα αυτό θα γίνει μια επανάληψη της ύλης, θα απαντηθούν ερωτήσεις, θα αξιολογήσετε το μάθημα και θα συζητήσουμε πιθανές βελτιώσεις.

A' Exetastiki Periodos 2010

Αξιολόγηση μαθήματος

Οδηγίες για module Questionnaire

Οδηγίες δημιουργίας Video Tutorial

#### Εικόνα 3-11 Παρουσίαση 10<sup>ου</sup> μαθήματος

Στο μάθημα αυτό, το οποίο είναι το τελευταίο θα γίνουν ανακοινώσεις σχετικά με την εξεταστική περίοδο και θα δοθούν οδηγίες στους φοιτητές σχετικά με τη διαδικασία δημιουργίας τόσο του οπτικοακουστικού υλικού όσο και του ερωτηματολογίου.

Τα ερωτηματολόγια είναι σχεδιασμένα και τοποθετημένα με προκαθορισμένες ερωτήσεις και πιθανές απαντήσεις με τέτοιο τρόπο έτσι ώστε:

- α) Ο κάθε φοιτητής να μπορεί να αξιολογήσει το μάθημα 1 φορά μόνο.
- β) Ο καθηγητής να μπορεί να παρακολουθεί τη συγκεντρωτική/συνολική αξιολόγηση του συνόλου των μαθητών.
- γ) Μη εγκεκριμένοι χρήστες να μην επιτρέπεται να αξιολογήσουν το μάθημα.
- δ) Ο φοιτητής να μπορεί να σχολιάσει/γράψει παρατηρήσεις.
- ε) Τα ερωτηματολόγια να μπορούν να εκτυπωθούν για χειρόγραφη συμπλήρωση.

**Αξιολόγηση μαθημάτων**

**\*1** Ερωτήσεις σχετικά με το σύνολο του μαθήματος.

	Καθόλου	Λίγο	Μέτρια	Πολύ	Πάρα πολύ
1. Οι στόχοι του μαθήματος ήταν σαφείς;	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. Η ύλη που καλύφθηκε ανταποκρινόταν στους στόχους του μαθήματος;	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. Η ύλη που διδάχθηκε ήταν καλά οργανωμένη;	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. Το εκπαιδευτικό υλικό που χρησιμοποιήθηκε βοήθησε στην καλύτερη κατανόηση του θέματος;	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. Τα εκπαιδευτικά βοηθήματα («σύγγραμμα», σημειώσεις, πρόσθετη βιβλιογραφία) χορηγήθηκαν εγκαίρως;	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. Πόσο ικανοποιητικό βρίσκετε το κύριο βιβλίο(α) ή τις σημειώσεις;	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7. Πόσο εύκολα διαθέσιμη είναι η βιβλιογραφία στην Πανεπιστημιακή Βιβλιοθήκη;	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8. Πόσο απαραίτητα κρίνετε τα προσπατούμενα του μαθήματος;	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9. Χρήση γνώσεων από / σύνδεση με άλλα μαθήματα.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10. Πώς κρίνετε το επίπεδο δυσκολίας του μαθήματος για το έτος του;	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
11. Πώς κρίνετε τον αριθμό Διδακτικών Μονάδων σε σχέση με τον φόρτο εργασίας;	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
12. Διαφάνεια των κριτηρίων βαθμολόγησης.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Εικόνα 3-12** Γενικές ερωτήσεις αξιολόγησης μαθήματος

**\*2** Στις περιπτώσεις όπου υπήρχαν γραπτές ή/και προφορικές εργασίες:

	Καθόλου	Λίγο	Μέτρια	Πολύ	Πάρα πολύ
13. Το θέμα δόθηκε εγκαίρως;	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14. Η καταληκτική ημερομηνία για υποβολή ή παρουσίαση των εργασιών ήταν λογική;	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
15. Υπήρχε σχετικό ερευνητικό υλικό στη βιβλιοθήκη;	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
16. Υπήρχε καθοδήγηση από τον διδάσκοντα;	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
17. Τα σχόλια του διδάσκοντος ήταν επικοινωνιακά και αναλυτικά;	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
18. Δόθηκε η δυνατότητα βελτίωσης της εργασίας;	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
19. Η συγκεκριμένη εργασία σας βοήθησε να κατανοήσετε το συγκεκριμένο θέμα;	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Εικόνα 3-13** Ερωτήσεις σχετικά με γραπτές ή/και προφορικές εργασίες

\*3 Ο/Η διδάσκων/ουσα:

	Καθόλου	Λίγο	Μέτρια	Πολύ	Πάρα πολύ
20. Οργανώνει καλά την παρουσίαση της ύλης στα μαθήματα;	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
21. Επιτυγχάνει να διεγείρει το ενδιαφέρον για το αντικείμενο του μαθήματος;	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
22. Αναλύει και παρουσιάζει τις έννοιες με τρόπο απλό και ενδιαφέροντα χρησιμοποιώντας παραδείγματα;	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
23. Ενθαρρύνει τους φοιτητές να διατυπώνουν απορίες και ερωτήσεις και να για να αναπτύξουν την κρίση τους;	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
24. Ήταν συνεπής στις υποχρεώσεις του/της (παρουσία στα μαθήματα, έγκαιρη διόρθωση εργασιών ή εργαστηριακών αναφορών, ώρες συνεργασίας με τους φοιτητές);	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
25. Είναι γενικά προσιτός στους φοιτητές;	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Εικόνα 3-14 Ερωτήσεις σχετικά με τον/την διδάσκων/ουσα

\*4 Εγώ ο/η φοιτητής/τρια:

	Καθόλου	Λίγο	Μέτρια	Πολύ	Πάρα πολύ
33. Παρακολουθώ τακτικά τις διαλέξεις.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
34. Μελετώ συστηματικά την ύλη.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

\*5 Αξιολογήστε με άριστα το 10 πώς σας φάνηκε το:

	1	2	3	4	5	6	7	8	9	10
1ο μάθημα	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2ο μάθημα	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3ο μάθημα	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4ο μάθημα	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5ο μάθημα	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6ο μάθημα	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7ο μάθημα	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8ο μάθημα	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9ο μάθημα	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Εικόνα 3-15 Ερωτήσεις σχετικά με το φοιτητή και το κάθε μάθημα

\* Παρατηρήσεις και σχόλια

6

**Εικόνα 3-16** Πεδίο για ελεύθερη έκθεση ιδεών, παρατηρήσεων και σχολίων

Αφού οι φοιτητές συμπληρώσουν το παραπάνω ερωτηματολόγιο ο καθηγητής θα μπορεί να παρατηρήσει συγκεντρωτικά το μέσο όρο των απαντήσεων που έδωσαν οι φοιτητές.

**1.** Ερωτήσεις σχετικά με το σύνολο του μαθήματος.

	Average rank					
	Καθόλου	Λίγο	Μέτρια	Πολύ	Πάρα πολύ	
1. Οι στόχοι του μαθήματος ήταν σαφείς;				■		4.0
2. Η ύλη που καλύφθηκε ανταποκρινόταν στους στόχους του μαθήματος;					■	5.0
3. Η ύλη που διδάχθηκε ήταν καλά οργανωμένη;					■	5.0
4. Το εκπαιδευτικό υλικό που χρησιμοποιήθηκε βοήθησε στην καλύτερη κατανόηση του θέματος;				■		4.0
5. Το εκπαιδευτικό βοηθήματα («σύγγραμμα», σημειώσεις, πρόσθετη βιβλιογραφία) χορηγήθηκαν εγκαίρως;				■		4.0
6. Πόσο ικανοποιητικό βρίσκετε το κύριο βιβλίο(ο) ή τις σημειώσεις;				■		4.0
7. Πόσο εύκολα διαθέσιμη είναι η βιβλιογραφία στην Πανεπιστημιακή Βιβλιοθήκη;			■			3.0
8. Πόσο απαραίτητα κρίνετε τα προσαρμοσμένα του μαθήματος;			■			3.0
9. Χρήση γνώσεων από / σύνδεση με άλλα μαθήματα.			■			3.0
10. Πώς κρίνετε το επίπεδο δυσκολίας του μαθήματος για το έτος του;			■			3.0
11. Πώς κρίνετε τον αριθμό Διδακτικών Μονάδων σε σχέση με			■			3.0

**Εικόνα 3-17** Με μπλε χρώμα εμφανίζονται οι απαντήσεις των φοιτητών

**5.** Αξιολογήστε με άριστα το 10 πώς σας φάνηκε το:

	Average rank										
	1	2	3	4	5	6	7	8	9	10	
1ο μάθημα									■		9.0
2ο μάθημα								■			8.5
3ο μάθημα									■		9.0
4ο μάθημα										■	10.0
5ο μάθημα									■		9.0
6ο μάθημα								■			8.0
7ο μάθημα									■		7.5
8ο μάθημα									■		8.0
9ο μάθημα										■	7.0

**Εικόνα 3-18** Όταν παραπάνω από 1 φοιτητές απαντήσουν τα αποτελέσματα είναι ο μέσος όρος των απαντήσεων



You have already filled out this questionnaire for us. Thank you.

**Εικόνα 3-19** Όταν ο φοιτητής απαντήσει μια φορά δε δικαιούται να συμπληρώσει ξανά το ερωτηματολόγιο



Sorry, but you do not currently have permissions to do that (View a questionnaire)

[More information about this error](#)

Continue

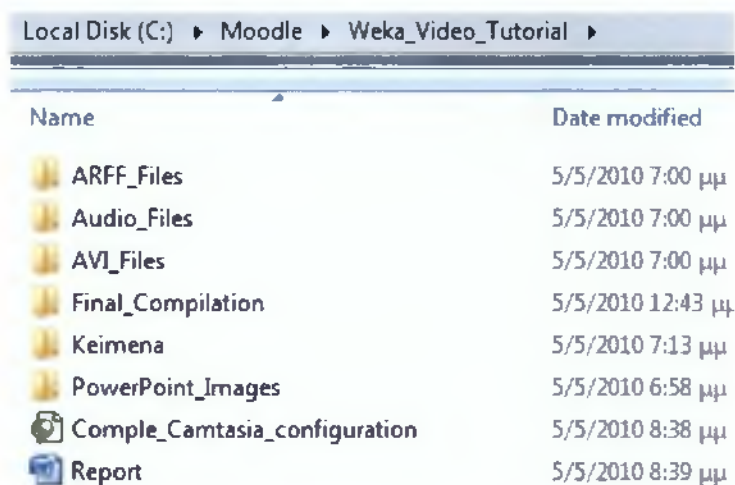
**Εικόνα 3-20** Μόνο εγγεγραμμένοι χρήστες στο μάθημα μπορούν να συμπληρώσουν το ερωτηματολόγιο

Πληροφορίες σχετικά με τη δημιουργία του παραπάνω ηλεκτρονικού ερωτηματολογίου μπορούν να βρεθούν στο APPENDIX.

## 4. Δημιουργία οπτικοακουστικού υλικού

### 4.1 Παρουσίαση αρχείων

Όλα τα αρχεία σχετικά με τη δημιουργία του Video tutorial τοποθετήθηκαν στο φάκελο C:\Moodle\Weka\_Video\_Tutorial. Υπάρχουν οι εξής φάκελοι:



Name	Date modified
ARFF_Files	5/5/2010 7:00 μμ
Audio_Files	5/5/2010 7:00 μμ
AVI_Files	5/5/2010 7:00 μμ
Final_Compilation	5/5/2010 12:43 μμ
Keimena	5/5/2010 7:13 μμ
PowerPoint_Images	5/5/2010 6:58 μμ
Comple_Camtasia_configuration	5/5/2010 8:38 μμ
Report	5/5/2010 8:39 μμ

Εικόνα 4-1 Αρχεία για τη δημιουργία του οπτικοακουστικού υλικού

- **ARFF\_Files** Στο φάκελο αυτό βρίσκονται δύο αρχεία τύπου ARFF (Attribute-Relation File Format) που είναι επίσης διαθέσιμα στη σελίδα τηλεκπαίδευσης Moodle τα οποία χρησιμοποιούνται κατά τη διάρκεια της παρουσίασης και τα οποία οι φοιτητές μπορούν να κατεβάσουν και να χρησιμοποιήσουν για να εξασκηθούν στις τεχνικές που παρουσιάζονται.
- **Audio\_Files** Στο φάκελο αυτό βρίσκονται τα αρχεία ήχου μορφής .WAV που δημιουργήθηκαν από την εφαρμογή Text-To-Speech και περιέχουν τους διαλόγους που εκφωνούνται.
- **AVI\_Files** Στο φάκελο αυτό βρίσκονται αρχεία μορφής .AVI που καταγράφουν το χειρισμό του λογισμικού που παρουσιάζεται στην παρουσίαση αυτή.



- **Final\_Compilation** Στο φάκελο αυτό βρίσκεται το τελικό αρχείο της παρουσίασης που περιέχει όλο το οπτικό και ακουστικό υλικό που δημιουργήθηκε στα πλαίσια της εργασίας.
- **Κείμενα** Στο φάκελο αυτό βρίσκονται αρχεία .TXT που έχουν χρησιμοποιηθεί για την ηλεκτρονική εκφώνηση των κειμένων.
- **PowerPoint\_Images** Στο φάκελο αυτό βρίσκεται ένα αρχείο παρουσίασης τύπου PowerPoint που περιέχει 4 διαφάνειες καθώς και τέσσερις εικόνες μορφής JPEG που περιέχουν το περιεχόμενο της κάθε διαφάνειας.

Το αρχείο **Complete\_Camtasia\_configuration** περιέχει σε μορφή XML όλες τις πληροφορίες σχετικά με την παρουσίαση του οπτικοακουστικού υλικού, την τοποθεσία των πρωτότυπων αρχείων εικόνας και ήχου καθώς και τον συγχρονισμό και τα οπτικά εφέ.

Τέλος το αρχείο **Report.doc** περιέχει τα κείμενα τα οποία θα συμπεριληφθούν σε αυτά της πτυχιακής εργασίας. Στη συνέχεια του κεφαλαίου θα παρουσιαστεί αναλυτικά η διαδικασία προετοιμασίας κάθε επιμέρους υλικού ξεχωριστά.

## **4.2 Αρχεία Attribute-Relation File Format**

Ξεκινώντας από το <http://maya.cs.depaul.edu/classes/ect584/WEKA/preprocess.html> όπου παρουσιάζεται η διαδικασία μετατροπής δεδομένων από ένα CSV (Comma Separated Values) αρχείο σε αρχεία .ARFF, εντοπίστηκαν ορισμένα δεδομένα, τα οποία είναι κατανοητά από το μέσο φοιτητή στον οποίο θα παρουσιαστεί η μελέτη αυτή.

Τα δεδομένα αποτελούνται από 600 εγγραφές σχετικά με τους πελάτες μιας τράπεζας.

Για κάθε πελάτη υπάρχουν οι παρακάτω πληροφορίες:

1. **age** Η ηλικία του πελάτη.
2. **sex** Το φύλο του πελάτη, δηλαδή η τιμή MALE ή FEMALE.
3. **region** Ο τόπος κατοικίας του πελάτη, δηλαδή μια από τις τιμές inner\_city ή rural ή suburban ή town.
4. **income** Το εισόδημα του πελάτη.
5. **married** Η οικογενειακή κατάσταση. YES αν είναι παντρεμένος, NO διαφορετικά.
6. **children** Ο αριθμός των τέκνων.
7. **car** Η ιδιοκτησία αυτοκινήτου. YES αν έχει αυτοκίνητο, NO διαφορετικά.
8. **save\_acct** Αν ο πελάτης έχει λογαριασμό αποταμίευσης. YES αν έχει NO διαφορετικά.
9. **current\_acct** Αν ο πελάτης έχει λογαριασμό όψεως. YES αν έχει NO διαφορετικά.
10. **mortgage** Αν ο πελάτης έχει κάποιο δάνειο. YES αν έχει NO διαφορετικά.
11. **pep** Αν ο πελάτης έχει προσωπικό πλάνο διαχείρισης κεφαλαίων. YES αν έχει NO διαφορετικά.

Δύο μορφές των ιδίων δεδομένων αποφασίστηκε να χρησιμοποιηθούν για την παρουσίαση αυτή. Στην **πρώτη μορφή** τα δεδομένα περιέχουν κανονικοποιημένες και μη κανονικοποιημένες τιμές. Στις μη-κανονικοποιημένες τιμές περιλαμβάνεται η ηλικία και το ετήσιο εισόδημα. Στην **δεύτερη μορφή** τα δεδομένα είναι πλήρως κανονικοποιημένα. Δηλαδή ακόμα και η ηλικία και το ετήσιο εισόδημα περιγράφονται από διακριτές τιμές και όχι αριθμητικές.

Λέγοντας **διακριτές** τιμές εννοούμε την αντικατάσταση αριθμητικών τιμών με συγκεκριμένα όρια αυτών.

Για την ηλικία για παράδειγμα τρία όρια επιλέχθηκαν:

α) έως 34 ετών.

β) από 35 μέχρι 51 ετών.

γ) από 52 και πάνω.

Η πραγματική ηλικία των πελατών αντικαταστάθηκε από μια διακριτή τιμή ανάμεσα από τις 0\_34 ή 35\_51 ή 52\_max. Παρομοίως επιλέχθηκαν τρία όρια και για το εισόδημα. Ακολουθεί η μορφή των δύο αρχείων:

```
@attribute age numeric
@attribute sex {FEMALE,MALE}
@attribute region {INNER_CITY,TOWN,RURAL,SUBURBAN}
@attribute income numeric
@attribute married {NO,YES}
@attribute children numeric
@attribute car {NO,YES}
@attribute save_act {NO,YES}
@attribute current_act {NO,YES}
@attribute mortgage {NO,YES}
@attribute pep {YES,NO}

@data
48,FEMALE,INNER_CITY,17546,NO,1,NO,NO,NO,NO,YES
40,MALE,TOWN,30085.1,YES,3,YES,NO,YES,YES,NO
```

Εικόνα 4-2 Αρχείο : bank\_data.arff

```

@attribute age {0_34,35_51,52_max}
@attribute sex {FEMALE,MALE}
@attribute region {INNER_CITY,TOWN,RURAL,SUBURBAN}
@attribute income {0_24386,24387_43758,43759_max}
@attribute married {NO,YES}
@attribute children {0,1,2,3}
@attribute car {NO,YES}
@attribute save_act {NO,YES}
@attribute current_act {NO,YES}
@attribute mortgage {NO,YES}
@attribute pep {YES,NO}

@data

35_51,FEMALE,INNER_CITY,0_24386,NO,1,NO,NO,NO,NO,YES
35_51,MALE,TOWN,24387_43758,YES,3,YES,NO,YES,YES,NO
    
```

Εικόνα 4-3 Αρχείο bank\_data\_normalized.arff

### 4.3 Αρχεία Audio

Συνολικά δημιουργήθηκαν και χρησιμοποιήθηκαν τα παρακάτω 8 ηχητικά αρχεία:

Local Disk (C:) ▶ Moodle ▶ Weka_Video_Tutorial ▶ Audio	
Name	Size
1_Εισαγωγή_1ο_Slide	694 KB
2_Εισαγωγή_2ο_Slide	241 KB
3_Μη_κανονικοποιημένα_δεδομένα_1ο_...	4.799 KB
4_Μη_κανονικοποιημένα_δεδομένα_2ο_...	2.142 KB
5_Παρουσίαση_3ου_Slide	1.314 KB
6_Παρουσίαση_4ου_Slide	571 KB
7_Κανονικοποιημένα_δεδομένα	5.837 KB
8_Επίλογος	425 KB

Εικόνα 4-4 Αρχεία ήχου

Η συνολική διάρκεια των αρχείων .WAV είναι 8 λεπτά και 32 δευτερόλεπτα. Η κωδικοποίηση είναι στα 256 kbps και το συνολικό μέγεθος 15,6 Mbytes.

Όνομα αρχείου	Διάρκεια ηχογράφησης	Μέγεθος δεδομένων
1_Εισαγωγή_1ο_Slide.wav	0:22 δευτ	693 Kbytes
2_Εισαγωγή_2ο_Slide.wav	0:07 δευτ	240 Kbytes
3_Μη_κανονικοποιημένα_δεδομένα_1ο_μέρος.wav	2:33 δευτ	4,68 Mbytes
4_Μη_κανονικοποιημένα_δεδομένα_2ο_μέρος.wav	1:08 δευτ	2,09 Mbytes
5_Παρουσίαση_3ου_Slide.wav	0:42 δευτ	1,28 Mbytes
6_Παρουσίαση_4ου_Slide.wav	0:18 δευτ	570 Kbytes
7_Κανονικοποιημένα_δεδομένα.wav	3:06 δευτ	5,69 Mbytes
8_Επίλογος.wav	0:13 δευτ	424 Kbytes

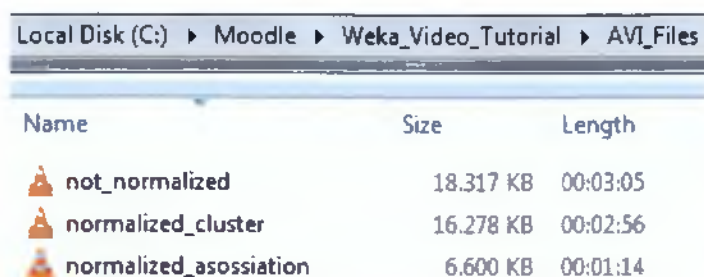
Τα παραπάνω αρχεία δημιουργήθηκαν μέσω της εφαρμογής TextAloud χρησιμοποιώντας την φωνή της «Αφροδίτης» από τη Loquento που παρουσιάστηκε στο κεφάλαιο 2.7.

Τα κείμενα των διαλόγων βρίσκονται στο APPENDIX A – «Κείμενα Διαλόγων Video Tutorial». Αν κάποιος παρατηρήσει προσεκτικά τα κείμενα αυτά μπορεί να δει ότι για να υπάρχει μεγαλύτερη επιτυχία και λόγω του ότι το TextAloud στην έκδοση που χρησιμοποιήθηκε δε μπορεί να μετατρέψει πολύ-γλωσσικό κείμενο σε φυσική ομιλία, οι αγγλικοί όροι έχουν πληκτρολογηθεί φωνητικά στην Ελληνική γλώσσα.

Για παράδειγμα για το ακουστικό αποτέλεσμα “Weka” πληκτρολογήθηκε η λέξη «Γουέεκα». Ομοίως για την λέξη “Assosiation” πληκτρολογήθηκε η λέξη «Ασοσιέϊσιον». Τέλος να σημειωθεί ότι οι κενές γραμμές που μπορεί κάποιος να δει στο APPENDIX A παίζουν το ρόλο της στιγμιαίας παύσης.

#### 4.4 Αρχεία AVI

Συνολικά δημιουργήθηκαν τρία αρχεία βίντεο μορφής .AVI συνολικού χρόνου 7 λεπτών και 15 δευτερολέπτων. Η ανάλυση των αρχείων αυτών είναι 800 επί 600 εικονοστοιχεία (pixels) και το μέσο bitrate είναι 800 Kbps.



Name	Size	Length
not_normalized	18.317 KB	00:03:05
normalized_cluster	16.278 KB	00:02:56
normalized_asossiation	6.600 KB	00:01:14

Εικόνα 4-5 Αρχεία Video

Για τη δημιουργία των παραπάνω βίντεο χρησιμοποιήθηκε η εφαρμογή Camtasia Recorder:



Εικόνα 4-6 Λογισμικό καταγραφής video

Αφού αποφασίστηκε ποια σενάρια θα εκτελεστούν στα τραπεζικά δεδομένα, εκτελέστηκε η εφαρμογή Weka. Κατόπιν μέσα από το Camtasia Recorder επιλέχθηκε η εφαρμογή Weka και ξεκίνησε η βιντεοσκόπηση. Η παραπάνω εφαρμογή καταγράφει όσα εξελίσσονται στην οθόνη του υπολογιστή συμπεριλαμβανομένου της κίνησης του ποντικιού.

#### 4.5 Final Compilation

Στο φάκελο αυτό βρίσκονται τα τελικά παραδοτέα αρχεία. Ένα αρχείο AVI που εμπεριέχει το σύνολο του οπτικού και ακουστικού υλικού καθώς και μια σελίδα HTML, ώστε να μπορεί να τοποθετηθεί το τελικό αρχείο στο διαδίκτυο. Το μέγεθος του τελικού αρχείου προσεγγίζει τα 75 Mbytes, και είναι ανάλυσης 800 επί 600 και παίζει με 10 Frames ανά δευτερόλεπτο. Βεβαίως αναλόγως των δυνατοτήτων του εξυπηρετητή Moodle ή της υπηρεσίας παροχής Video streaming (π.χ. YouTube) το τελικό βίντεο μπορεί να επανακωδικοποιηθεί σε διαφορετική ανάλυση.

Για τη δημιουργία του τελικού οπτικοακουστικού υλικού, χρησιμοποιήθηκε το λογισμικό Camtasia Studio.

Αρχικά με drag & drop τοποθετήθηκε όλο το υλικό μέσα στο περιβάλλον του λογισμικού:

Clip Bin			
Name	Size	Dimensions	Duration
<b>Image</b>			
1.jpg	110 KB (113.580 Bytes)	578 x 434	
2.jpg	79 KB (81.828 Bytes)	580 x 436	
3.jpg	106 KB (109.043 Bytes)	579 x 437	
4.jpg	83 KB (85.874 Bytes)	580 x 432	
5.jpg	80 KB (82.741 Bytes)	579 x 435	
<b>Audio</b>			
1_Εισαγωγή_1ο_Slide.wav	693 KB (709.738 Bytes)		0:00:22;05
2_Εισαγωγή_2ο_Slide.wav	240 KB (245.796 Bytes)		0:00:07;20
3_Μη κανονικοποιημένα δεδομένα_1ο_μέρος.wav	4.798 KB (4.913.162 Bytes)		0:02:33;16
5_Παρουσίαση_3ου_Slide.wav	1.313 KB (1.344.786 Bytes)		0:00:42;01
6_Παρουσίαση_4ου_Slide.wav	570 KB (583.696 Bytes)		0:00:18;07
7_Κανονικοποιημένα_δεδομένα.wav	5.836 KB (5.976.188 Bytes)		0:03:06;23
8_Επίλογος.wav	424 KB (435.038 Bytes)		0:00:13;18
<b>Video</b>			
normalized_asossiation.avi	6.599 KB (6.757.620 Bytes)	800 x 600	0:01:14;08
normalized_cluster.avi	16.277 KB (16.667.874 Bytes)	800 x 600	0:02:56;14
not_normalized.avi	18.316 KB (18.755.626 Bytes)	800 x 600	0:03:05;02

Εικόνα 4-7 Εισαγωγή πολυμεσικού περιεχομένου

Στη συνέχεια τοποθετήθηκαν διαδοχικά στην περιοχή Video του StoryBoard οι δύο πρώτες εικόνες και οι δύο πρώτες διαφάνειες της παρουσίασης που αφορούν την εισαγωγή στο μάθημα. Ακολούθησε το αρχείο βίντεο σχετικά με τα μη κανονικοποιημένα δεδομένα και κατόπιν οι διαφάνειες 3 και 4. Στη συνέχεια τοποθετήθηκαν τα δύο αρχεία βίντεο σχετικά με τα κανονικοποιημένα δεδομένα και τέλος εισήχθη η τελευταία διαφάνεια, ο επίλογος.

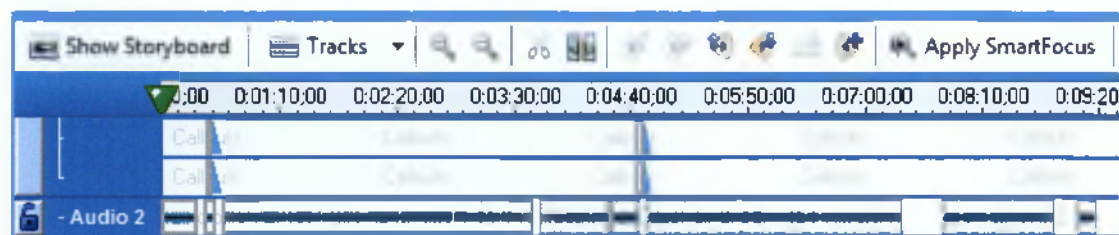




Εικόνα 4-8 Time Line επεξεργασίας video

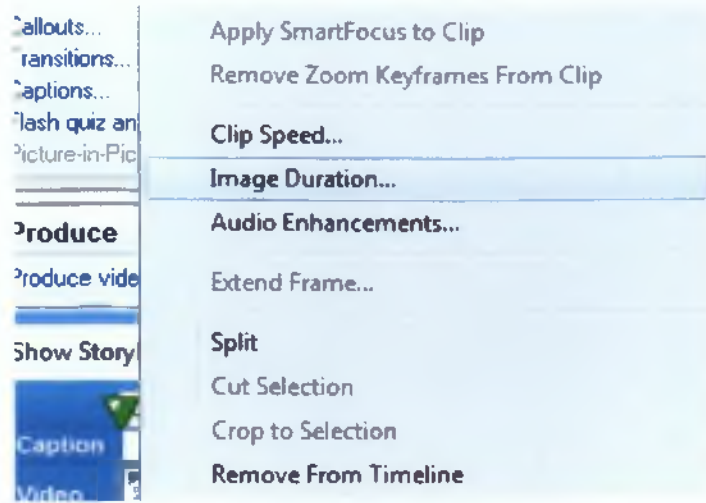
Όπως κάποιος μπορεί να παρατηρήσει, η χρονική διάρκεια του κάθε οπτικού υλικού παρουσιάζεται στην παραπάνω μπάρα.

Στη συνέχεια τοποθετήθηκαν με τη σωστή σειρά τα ηχητικά δεδομένα και η εκφώνηση του εισηγητή.



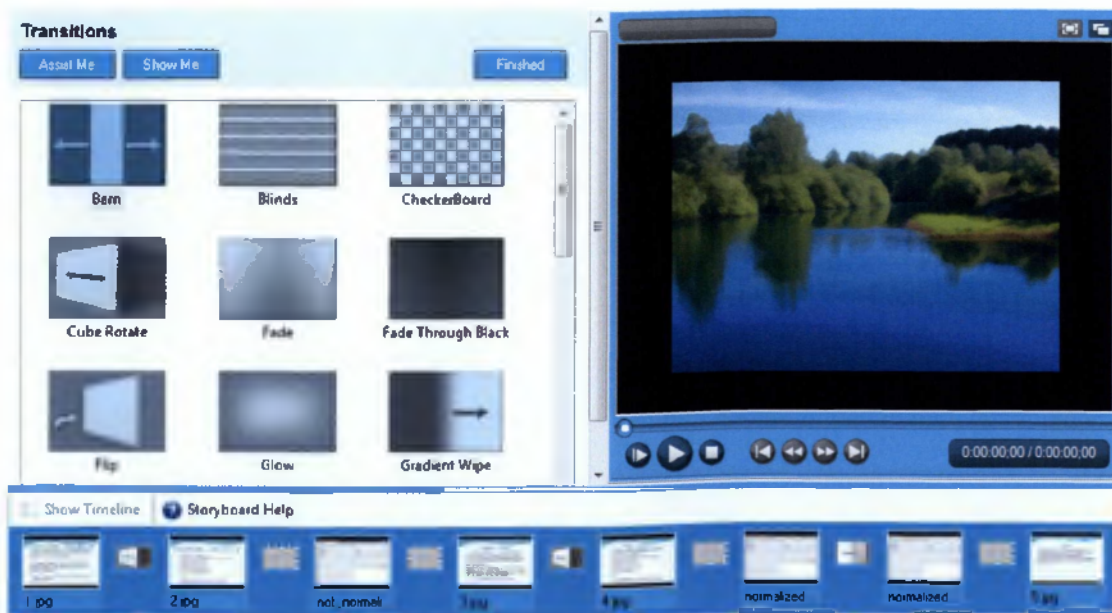
Εικόνα 4-9 Δυνατότητες επεξεργασίας ήχου

Για το συγχρονισμό της εικόνας με τον ήχο, καθορίστηκε στη συνέχεια ο χρόνος που θα προβάλλεται η κάθε εικόνα. Συγκεκριμένα η πρώτη διαφάνεια ορίστηκε να εμφανίζεται για 22 δευτερόλεπτα, επειδή αυτός είναι ο χρόνος που απαιτείται από τον εκφωνητή για να κάνει τη λεκτική εισαγωγή στο θέμα. Η διάρκεια της κάθε εικόνας μπορεί να τροποποιηθεί με δεξί κλικ πάνω της και κατόπιν -> Image Duration...



Εικόνα 4-10 Διάρκεια προβολής εικόνας

Στη συνέχεια κάνοντας εισαγωγή Transition τοποθετήθηκαν ορισμένα οπτικά εφέ στη σύνδεση του ενός οπτικού υλικού με το αμέσως επόμενο:



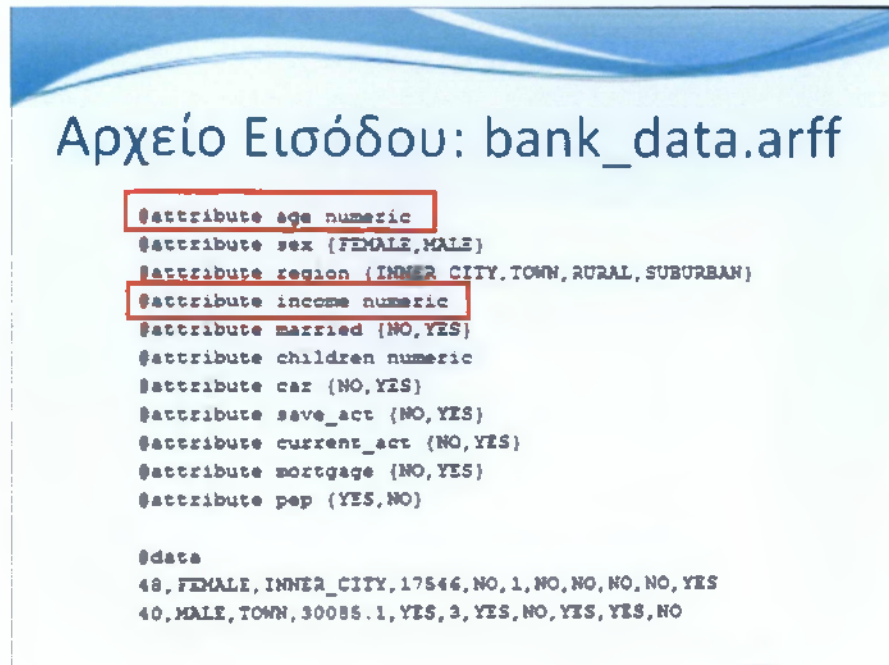
Εικόνα 4-11 Εισαγωγή εφέ μετάβασης

Το εφέ του κύβου (Cube Rotate) περιστρέφει έναν τρισδιάστατο κύβο όπου η αριστερή πλευρά του παρουσιάζει τη προηγούμενη και η δεξιά πλευρά του την επόμενη διαφάνεια.



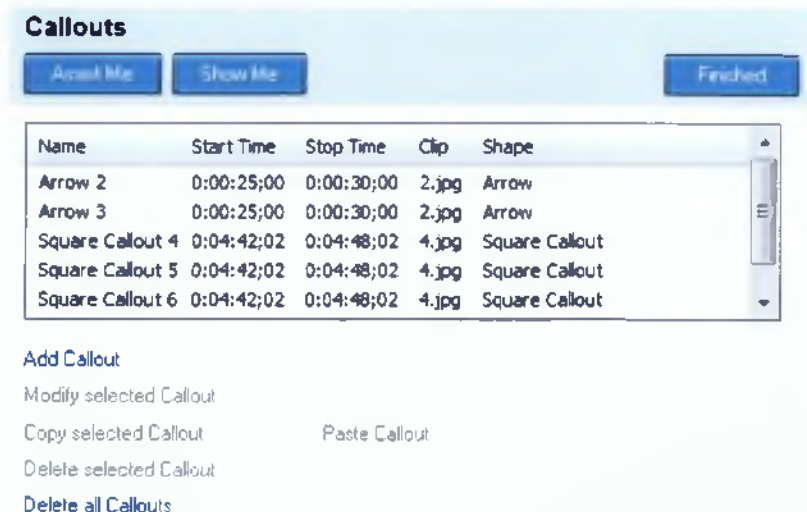
Εικόνα 4-12 Εφέ κόβου

Στη συνέχεια τοποθετήθηκαν κάποια οπτικά εφέ πάνω στις υπάρχοντες διαφάνειες. Συγκεκριμένα στην παρακάτω διαφάνεια ο ομιλητής εξηγεί ότι η ηλικία και το ετήσιο εισόδημα αποτελούνται από μη κανονικοποιημένες τιμές, δηλαδή από αριθμητικές.



Εικόνα 4-13 Χρήση δεικτών (callouts) για να δοθεί έμφαση

Για να καταλάβει καλύτερα ο εκπαιδευόμενος, τοποθετούνται δύο βέλη για να δώσουν έμφαση. Οι ετικέτες που υπάρχουν σε συγκεκριμένα σημεία της παρουσίασης και για συγκεκριμένους χρόνους ονομάζονται Callouts.



Εικόνα 4-14 Διάρκεια δεικτών (callouts)

Και χρησιμοποιήθηκαν για να δοθεί έμφαση.

## Αρχείο Εισόδου: bank\_data.arff

```

@attribute age numeric Numeric
@attribute sex {FEMALE,MALE}
@attribute region {INNER_CITY,TOWN,RURAL,SUBURBAN}
@attribute income numeric Numeric
@attribute married {NO,YES}
@attribute children numeric
@attribute car {NO,YES}
@attribute save_act {NO,YES}
@attribute current_act {NO,YES}
@attribute mortgage {NO,YES}
@attribute pep {YES,NO}










@data
49,FEMALE,INNER_CITY,17546,NO,1,NO,NO,NO,NO,YES
40,MALE,TOWN,30095.1,YES,3,YES,NO,YES,YES,NO

```

Εικόνα 4-15 Εμφάνιση κατάλληλων σχημάτων σε συγκεκριμένα σημεία

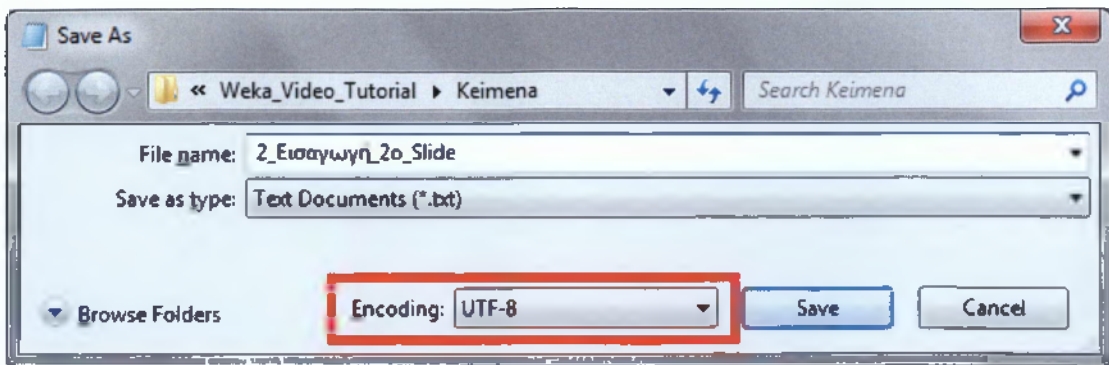
### 4.6 Κείμενα

Τα παρακάτω αρχεία περιέχουν τα επιμέρους κείμενα που εκφωνούνται κατά τη διάρκεια της παρουσίασης.

 1_Εισαγωγή_1ο_Slide	1 KB
 2_Εισαγωγή_2ο_Slide	1 KB
 3_Παρουσίαση_μη_κανονικοποιημένων_δεδομένων	6 KB
 4_Παρουσίαση_3ου_Slide	2 KB
 5_Παρουσίαση_4ου_Slide	1 KB
 6_Παρουσίαση_κανονικοποιημένων_δεδομένων_1ο_μέρος	5 KB
 7_Παρουσίαση_κανονικοποιημένων_δεδομένων_2ο_μέρος	3 KB
 8_Επίλογος	1 KB
 Όλα_τα_κείμενα	18 KB

Εικόνα 4-16 Κείμενα εκφώνησης

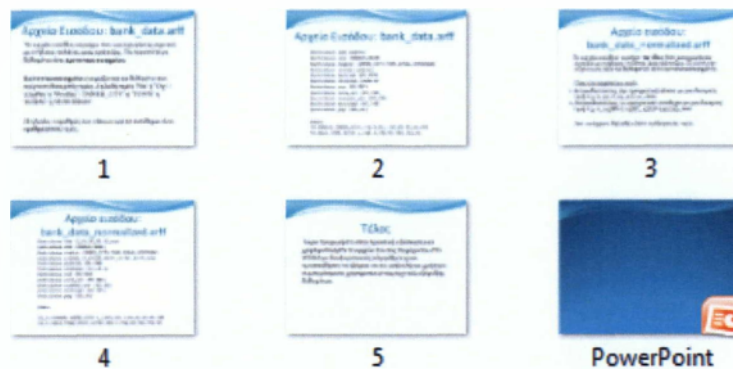
Τα αρχεία με κατάληξη .TXT έχουν αποθηκευτεί με την κωδικοποίηση UTF-8:



Εικόνα 4-17 Κωδικοποίηση αρχείων κειμένου

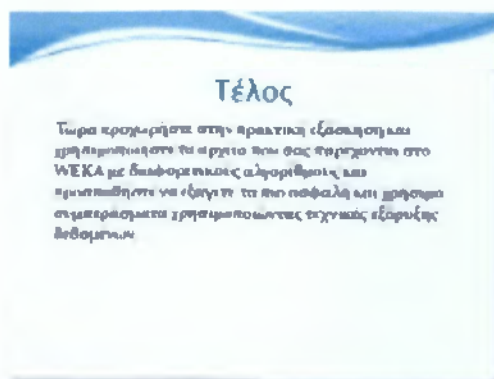
## 4.7 PowerPoint

Για τη δημιουργία των διαφανειών που χρησιμοποιούνται σε διάφορα σημεία της παρουσίασης χρησιμοποιήθηκε το PowerPoint.



Εικόνα 4-18 Δημιουργία διαφανειών

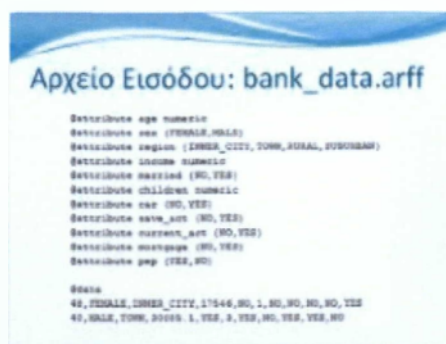
Συνολικά δημιουργήθηκαν 5 διαφάνειες, μία ως επίλογος, δύο για τα μη κανονικοποιημένα δεδομένα και δύο για τα κανονικοποιημένα:



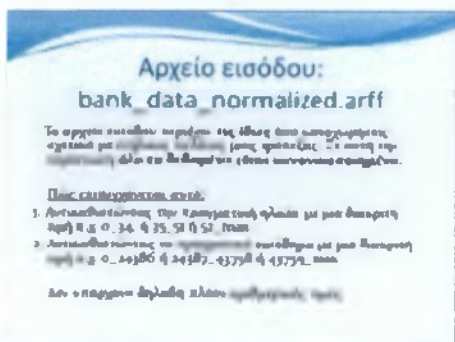
5



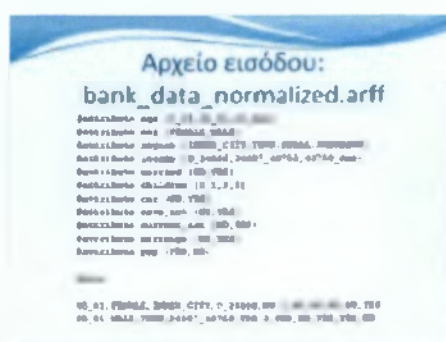
1



2



3



4

Εικόνα 4-19 Πέντε συνολικά διαφάνειες

## 4.8 Κείμενα Διαλόγων Video Tutorial

### 1 Εισαγωγή Io Slide.txt

Το αρχείο εισόδου περιέχει εξακόσιες καταχωρήσεις σχετικά με ενήλικες πελάτες μιας τράπεζας. Τα περισσότερα δεδομένα είναι κανονικοποιημένα.

Κανονικοποιημένα ονομάζονται τα δεδομένα που παίρνουν διακριτές τιμές. Δηλαδή τιμές Ναι ή Όχι, Αληθές ή Ψευδές και τα λοιπά.

Η ηλικία, ο αριθμός των τέκνων και το εισόδημα είναι αριθμητικές τιμές.

## 2 Εισαγωγή 2ο Slide.txt

Στο αρχείο αρφ, η ηλικία και το εισόδημα είναι αριθμητικές τιμές , νουμέρικ. Οι υπόλοιπες τιμές είναι κανονικοποιημένες

## 3 Παρουσίαση μη κανονικοποιημένων δεδομένων.txt

Αρχικά ανοίγουμε από το περιβάλλον του Γουέκα το αρχείο αρφ.

Υπάρχουν εξακόσιες εγγραφές με έντεκα πεδία.

Η ηλικία είναι επιλεγμένη. Η ελάχιστη τιμή είναι δεκαοκτώ, η μέγιστη εξήντα επτά και ο μέσος όρος είναι σαράντα δύο.

Επιλέγοντας κλάς καθορίζουμε το χρωματισμό. Στο παράδειγμα αυτό θα δούμε με μπλε χρώμα τις γυναίκες και με κόκκινο τους άντρες.

Αυτομάτως τα δεδομένα οπτικοποιούνται. Επιλέγοντας κάθε ξεχωριστό χαρακτηριστικό μπορούμε να δούμε γραφήματα δεδομένων. Χωρίς να εκτελέσουμε καμία εντολή ουσιαστικά αξιοποιούμε τεχνικές εξόρυξης δεδομένων.

Παρατηρήστε τα δεδομένα σχετικά με την οικογενειακή κατάσταση..... τον αριθμό των παιδιών..... Ο αριθμός των παιδιών μπορείτε να δείτε ότι είναι μια τιμή από το μηδέν έως το 3.... επίσης παρατηρήστε αν διαθέτουν αυτοκίνητο..... αν οι πελάτες έχουν λογαριασμό αποταμίευσης στη συγκεκριμένη τράπεζα..... αν διαθέτουν λογαριασμό όψεως..... αν έχουν κάποιο δάνειο..... ή αν έχουν κάποιο προσωπικό πλάνο διαχείρισης κεφαλαίων.

Στη συνέχεια θα παρουσιαστεί ένα σενάριο ταξινόμησης, κλάσιφαί και ένα σενάριο ομαδοποίησης, κλάστερινγκ



Για το σενάριο ταξινόμησης εκτελέστηκε ο αλγόριθμος ζήροου άρ. Επιλέγουμε στον άξονα χι το ετήσιο εισόδημα, στον άξονα ψι την ηλικία και είναι εύκολο να παρατηρήσουμε ότι όσο αυξάνεται η ηλικία τόσο αυξάνεται και το ετήσιο εισόδημα. Κάθε ξεχωριστή εγγραφή των δεδομένων από τις συνολικά εξακόσιες εμφανίζεται ως ένα σημείο στο γράφημα. Όταν περισσότερες από μία εγγραφές συμπίπτουν στο ίδιο σημείο τότε το Χι στο γράφημα έχει μεγαλύτερο μέγεθος. Σχετικά με το εισόδημα μπορούμε να κατηγοριοποιήσουμε το σύνολο των δεδομένων σε τρεις κατηγορίες. Σε αυτούς που έχουν εισόδημα από πέντε χιλιάδες μέχρι εικοσιτέσσερις περίπου χιλιάδες, από εικοσιτέσσερις χιλιάδες μέχρι περίπου σαραντατρις χιλιάδες και σε αυτούς που έχουν εισόδημα παραπάνω από σαραντατρις χιλιάδες.

Το σενάριο ομαδοποίησης απαιτεί αρκετό χρόνο για να εκτελεστεί καθώς δοκιμάζονται χιλιάδες πιθανοί συνδυασμοί. Ένα μοντέλο δημιουργείται μέσω αλγορίθμου.

Όταν αυτό ολοκληρωθεί μπορούμε να πατήσουμε δεξί κλικ πάνω στο αποτέλεσμα και κατόπιν να επιλέξουμε την οπτικοποίηση των αποτελεσμάτων ομαδοποίησης.

Ο αλγόριθμος δημιούργησε 3 ομάδες. Για να κατανοήσουμε καλύτερα επιλέγουμε στον άξονα Χι την ηλικία, στον άξονα Ψι το ετήσιο εισόδημα και το χρώμα να αφορά αν ο κάθε πελάτης διαθέτει λογαριασμό αποταμίευσης ή όχι. Με μπλε χρώμα εμφανίζονται όσοι ΔΕΝ διαθέτουν τέτοιο λογαριασμό ενώ με κόκκινο όσοι διαθέτουν. Μπορεί κάποιος να παρατηρήσει ότι όσο μεγαλύτερος σε ηλικία είναι ο πελάτης τόσο πιο πιθανό είναι να έχει λογαριασμό αποταμίευσης. Επίσης όσο μεγαλύτερο είναι το εισόδημα, τόσο πιο πιθανό να υπάρχει λογαριασμός αποταμίευσης. Με ασφάλεια εξάγεται το συμπέρασμα ότι όσοι έχουν εισόδημα μεγαλύτερο από τριάντα έξι χιλιάδες ευρώ, έχουν όλοι λογαριασμό αποταμίευσης.

Το σενάριο σχέσεων , ασοσιέσιον δεν μπορεί να εκτελεστεί σε μη κανονικοποιημένα δεδομένα!

#### 4 Παρουσίαση 3ου Slide.txt

Έχοντας εξάγει κάποια συμπεράσματα από το προηγούμενο παράδειγμα μπορούμε να προχωρήσουμε στην κανονικοποίηση των δεδομένων. Το δεύτερο αρχείο που θα

κατεβάσετε περιέχει τις ίδιες εξακόσιες καταχωρήσεις σχετικά με ενήλικες πελάτες μιας τράπεζας. Σε αυτή την περίπτωση όλα τα δεδομένα είναι κανονικοποιημένα. Πως επιτυγχάνεται αυτό ;

Αντικαθιστώντας την πραγματική ηλικία με μια διακριτή τιμή. Για παράδειγμα από μηδέν μέχρι τριάντα τέσσερα, από τριάντα πέντε μέχρι πενήντα ένα και από πενήντα δύο και πάνω.

Ομοίως αντικαθιστούμε το πραγματικό εισόδημα με μια διακριτή τιμή. Από μηδέν μέχρι εικοσιτέσσερις χιλιάδες, από εικοσιτέσσερις μέχρι περίπου σαράντα τρεις και από σαράντα τέσσερις και πάνω.

#### 5 Παρουσίαση 4ου Slide.txt

Εδώ παρατηρείτε το νέο αρχείο μετά από την απαραίτητη επεξεργασία. Η επεξεργασία και μετατροπή των μη κανονικοποιημένων δεδομένων σε κανονικοποιημένα μπορεί να γίνει με διάφορους τρόπους ακόμα και χρησιμοποιώντας το εξέλ. Εσείς θα χρησιμοποιήσετε το έτοιμο αρχείο που σας προσφέρει η σελίδα του μαθήματος.

#### 6 Παρουσίαση κανονικοποιημένων δεδομένων 1ο μέρος.txt

Ομοίως με το προηγούμενο παράδειγμα τα δεδομένα εμφανίζονται άμεσα σε γραφήματα. Αυτή τη φορά όμως όλα τα δεδομένα έχουν διακριτές τιμές. Μπορείτε τώρα μέσα από το γραφικό περιβάλλον του Γουέκα να παρατηρήσετε τις νέες τιμές της ηλικίας και κατόπιν του εισοδήματος.

Θα προχωρήσουμε εκτελώντας ένα σενάριο ομαδοποίησης.

Και αυτή τη φορά θα χρειαστεί πολύς χρόνος για να ολοκληρωθεί η διαδικασία αυτή. Σκεφτείτε ότι τα δεδομένα μας αποτελούν μονάχα εξακόσιες εγγραφές, κάθε μία από τις οποίες έχει μόλις έντεκα στοιχεία. Φανταστείτε λοιπόν πόσο χρόνο και τι υπολογιστικά συστήματα απαιτεί η εξόρυξη δεδομένων από πραγματικές αποθήκες δεδομένων ενός νοσοκομείου ή μιας για παράδειγμα ασφαλιστικής εταιρίας!

Εδώ πρέπει να πούμε ότι το Γουέκα όπως και οι αντίστοιχες εφαρμογές μίζινες ιντέλιτζενς διαθέτουν δεκάδες υλοποιημένους διαφορετικούς αλγόριθμους, κάθε ένας από τους οποίους έχει διαφορετική αποτελεσματικότητα και ταχύτητα εκτέλεσης. Το επιστημονικό πεδίο υλοποίησης και βελτιστοποίησης των αλγορίθμων αυτών είναι πολύ ενεργό και δεκάδες διδακτορικές έρευνες απασχολούνται με τα θέματα αυτά.

Μπορεί κάποιος εύκολα να καταλάβει βλέποντας πόσο χρόνο απαιτεί η κάθε ενέργεια ότι τα αποτελέσματα για να είναι αξιοποιήσιμα πρέπει να εξάγονται το συντομότερο δυνατόν. Πολλές τέτοιες διεργασίες μπορεί να απαιτήσουν ώρες ή ακόμα και ημέρες για την εξαγωγή συμπερασμάτων.

Αφού ολοκληρωθεί η διαδικασία της ομαδοποίησης πατάμε δεξί κλικ και επιλέγουμε βιζουαλαίζ κλάστερ ασάινμεντς.

Στον άξονα Χι επιλέγουμε το εισόδημα και στον άξονα Ψι την ηλικία. Σε αυτή την περίπτωση θα επιλέξουμε την κατοχή αυτοκινήτου ως την τρίτη διάσταση των δεδομένων. Με κόκκινο χρώμα είναι αυτοί που διαθέτουν αυτοκίνητο.

Ανεβάζοντας την τιμή του τζίτερ, οι τιμές στο γράφημα απομακρύνονται και μπορούμε οπτικά να κατανοήσουμε τον ακριβή αριθμό των εγγραφών. Παρατηρούμε ότι τέσσερις εγγραφές αφορούν άτομα υψηλού εισοδήματος και ηλικίας από τριανταπέντε έως πενήντα ένα. Μπορούμε στη συνέχεια να επιλέξουμε διαφορετική τρίτη διάσταση των δεδομένων. Να δούμε δηλαδή ότι οι δύο εξ αυτών είναι παντρεμένοι, ότι και οι τέσσερις είναι άνδρες, ότι μένουν στα προάστια, ότι οι δύο δεν έχουν παιδιά, ενώ οι άλλοι δύο έχουν από δύο παιδιά, ότι όλοι έχουν λογαριασμό αποταμίευσης, ότι κανείς δεν έχει δάνειο. Μια πραγματική τράπεζα με πραγματικά δεδομένα θα μπορούσε να δημιουργήσει προφίλ πελατών και αναλόγως του εισοδήματος, της ηλικίας και άλλων στοιχείων να προβλέψει τη συμπεριφορά του κάθε πελάτη με βάση τις υπάρχουσες στατιστικές. Αυτή είναι η σπουδαιότητα της εξόρυξης δεδομένων.

## 7 Παρουσίαση κανονικοποιημένων δεδομένων 2ο μέρος.txt

Ανοίγουμε λοιπόν το αρχείο με τα δεδομένα και εκτελούμε τον αλγόριθμο α πραϊόρι. Εντοπίζονται δέκα συσχετίσεις. Η κάθε συσχέτιση ακολουθείται από μια τιμή που

ονομάζεται κόνφιντενς.

Αυτοί που έχουν μεγάλο εισόδημα είναι σγδόντα εγγραφές και ΟΛΕΣ έχουν λογαριασμό αποταμίευσης. Η σιγουριά ή αλλιώς κόνφιντενς είναι λοιπόν ένα.

Όσοι έχουν μεγάλη ηλικία ΚΑΙ μεγάλο εισόδημα είναι εβδομήντα έξι εγγραφές και ΟΛΕΣ έχουν λογαριασμό αποταμίευσης. Αυτή η συσχέτιση είναι εμφανής από τον προηγούμενο κανόνα, οπότε και την αγνοούμε, όπως και τις αμέσως παρακάτω.

Στην πέμπτη συσχέτιση παρατηρούμε ότι όσοι δεν έχουν παιδιά, έχουν λογαριασμό αποταμίευσης, δεν έχουν δάνειο και προσωπικό πλάνο επενδύσεων, είναι κατά ενενήντα εννιά τις εκατό παντρεμένοι.

Στην έκτη συσχέτιση παρατηρούμε ότι οι γυναίκες χωρίς παιδιά και χωρίς δάνειο είναι κατά κανόνα παντρεμένες.

Κατόπιν παρατηρούμε και τους υπόλοιπους κανόνες συσχετίσεων μαζί με την βεβαιότητά τους. Εδώ πρέπει να τονίσουμε ότι ζητήθηκε από το Γουέκα να δημιουργήσει δέκα συσχετίσεις. Επειδή δεν είναι πάντα όλες οι συσχετίσεις χρήσιμες, θα μπορούσαν να ζητηθούν περισσότεροι κανόνες.

## 8 Επίλογος.txt

Τώρα προχωρήστε στην πρακτική εξάσκηση και χρησιμοποιήστε τα αρχεία που σας παρέχονται στο Γουέκα με διαφορετικούς αλγορίθμους. Προσπαθήστε να εξάγετε τα πιο ασφαλή και χρήσιμα συμπεράσματα χρησιμοποιώντας τεχνικές εξόρυξης δεδομένων.

## 5. Βιβλιογραφία

Blain, J. (2010). *Cegos International Partners Network. Current learning trends in Europe and the United States*. Retrieved 4 2, 2010, from [http://www.trainingindustry.com/media/2499749/cegos%20white%20paper\\_europe%20and%20us%20learning%20trends\\_10\\_09.pdf](http://www.trainingindustry.com/media/2499749/cegos%20white%20paper_europe%20and%20us%20learning%20trends_10_09.pdf)

Clark, R. C. (2008). *Developing technical training: a structured approach for developing Classroom and Computer-Based Instructional Materials*. Third edition. Pfeiffer.

Clark, R. C., & Mayer, R. E. (2008). *E-Learning and the Science of Instruction*. 2nd Edition. Pfeiffer.

Hammergren, T. C., & Simon, A. R. (2009). *Data Warehousing For Dummies 2nd.Edition*.

Βανδουλάκης, Ι. (n.d.). *Οπτικό, Ακουστικό και Κινησθητικό Στιλ*. Retrieved from [http://plnct.ning.com/group/web20/forum/attachment/download?id=3136368%3AUploadedFile%3A880&ei=DEDVS6v0F8ry-Qbh99i\\_DA&usg=AFQjCNGTi6B\\_A3tVzofXsKpBD-fd5HqhAA&sig2=VpmuzDcPKMf-Ml\\_MtQzIfQ](http://plnct.ning.com/group/web20/forum/attachment/download?id=3136368%3AUploadedFile%3A880&ei=DEDVS6v0F8ry-Qbh99i_DA&usg=AFQjCNGTi6B_A3tVzofXsKpBD-fd5HqhAA&sig2=VpmuzDcPKMf-Ml_MtQzIfQ)

Βελίκης, Ι. (n.d.). *Οι αντιληπτικοί τύποι*. Retrieved from <http://sites.google.com/site/yannisvelikis/essays/oi-antileptikoi-typoi>

Mayer , R.E. , Heiser , J. , & Lonn , S. ( 2001 ). Cognitive constraints on multimedia learning: When presenting more material results in less understanding . *Journal of Educational Psychology* , 93 , 187 – 198 .

Moreno , R. , & Mayer , R.E. ( 2000b ). Engaging students in active learning: The case for personalized multimedia messages. *Journal of Educational Psychology* , 93 , 724 – 733 .

Mayer , R.E. , & Anderson , R.B. ( 1991 ). Animations need narrations: An experimental test of a dual-coding hypothesis . *Journal of Educational Psychology* , 83 , 484 – 490 .

Δημήτρης Ζαχαρόπουλος (2005) , Διεθνή πρότυπα στην ηλεκτρονική μάθηση, Διπλωματική εργασία σε Μεταπτυχιακό «Τεχνολογίες της Πληροφορικής και των επικοινωνιών στην εκπαίδευση». Διαθέσιμο από:

<http://nemertes.lis.upatras.gr/dspace/bitstream/123456789/173/1/248.pdf>

Galvin T. Industry 2001 report. Training Magazine. 2001;38(10):40–75.

Ξυδίας Παναγιώτης, «E-Learning μια νέα διάσταση στην εκπαίδευση». Διπλωματική εργασία για το ΜΔΕ στα Υπολογιστικά Μαθηματικά και την Πληροφορική.

Διαθέσιμο από:

[http://nemertes.lis.upatras.gr/dspace/bitstream/123456789/513/1/Nimertis\\_Xydias.pdf](http://nemertes.lis.upatras.gr/dspace/bitstream/123456789/513/1/Nimertis_Xydias.pdf)

Ανεζίνης Βασίλης (2009), «Υλοποίηση οπτικού εργαλείου κατασκευής περιεχομένου μεταδεδομένων για εξ' αποστάσεως εκπαίδευση με την χρήση προτύπων.»

Διπλωματική εργασία σε Μεταπτυχιακό «Τεχνολογίες της Πληροφορικής και των επικοινωνιών στην εκπαίδευση». Διαθέσιμο από:

[http://nemertes.lis.upatras.gr/dspace/bitstream/123456789/1605/1/Anezinis\\_diplomai\\_ki.pdf](http://nemertes.lis.upatras.gr/dspace/bitstream/123456789/1605/1/Anezinis_diplomai_ki.pdf)

Φιλίππιδη Ανδρομάχη (2008). Σχεδίαση, ανάπτυξη, εφαρμογή και αξιολόγηση ενός συμβατικού πανεπιστημιακού μαθήματος με τη χρήση ενός υπολογιστικού περιβάλλοντος σύγχρονης και ασύγχρονης εκπαίδευσης: αντιλήψεις στάσεις και πρακτικές των φοιτητών. Διπλωματική εργασία σε Μεταπτυχιακό «Διδακτική των Θετικών Επιστημών: Εκπαιδευτικά Προγράμματα, Αξιολόγηση και Τεχνολογίες της Πληροφορίας και της Επικοινωνίας στην Εκπαίδευση». Διαθέσιμο από:

[http://nemertes.lis.upatras.gr/dspace/bitstream/123456789/1991/1/Nimertis\\_Filippidi.pdf](http://nemertes.lis.upatras.gr/dspace/bitstream/123456789/1991/1/Nimertis_Filippidi.pdf)

Κουτσοουρίδης Ιωάννης (2008). «Συστήματα διαχείρισης μάθησης (LMS).

Παρουσίαση και αξιολόγηση των moodle, blackboard και e-class με κριτήριο τις θεωρίες μάθησης στις οποίες στηρίζονται». Διπλωματική εργασία σε μεταπτυχιακό στις επιστήμες της γλώσσας και επικοινωνίας. Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης. Διαθέσιμο από:

<http://invenio.lib.auth.gr/record/114027/files/KOUTSOURIDIS.pdf?version=1>

## APPENDIX

**Questionnaire module για ερωτηματολόγια**

Αρχικά εγκαθιστούμε το module Questionnaire:

**Modules and plugins**

Σελίδα: (Προηγούμενο) 1 ... 623 624 625 626 627 628 629 630 631 632 633 634 635 636 (Επόμενο)

**Activity Module: Questionnaire**

Create your own surveys using this tool.

<http://moodle.org/mod/data/view.php?d=13&rid=84>

Για τη δημιουργία των ερωτημάτων πηγαίνουμε στο Questionnaire, επιλέγουμε Rate (scale 1..5) και πατάμε Add selected question type.

**Questions ?**

Rate (scale 1.5) Add selected question type

Στη συνέχεια συμπληρώνουμε τον τίτλο, το αν είναι υποχρεωτικό να απαντηθεί το ερώτημα αυτό, 5 ως τον αριθμό των πιθανών απαντήσεων και μια σύντομη περιγραφή.





6. Πόσο ικανοποιητικό βρίσκετε το κύριο βιβλίο(α) ή τις σημειώσεις;
7. Πόσο εύκολα διαθέσιμη είναι η βιβλιογραφία στην Πανεπιστημιακή Βιβλιοθήκη;
8. Πόσο απαραίτητα κρίνετε τα προαπαιτούμενα του μαθήματος;
9. Χρήση γνώσεων από / σύνδεση με άλλα μαθήματα.
10. Πώς κρίνετε το επίπεδο δυσκολίας του μαθήματος για το έτος του;
11. Πώς κρίνετε τον αριθμό Διδακτικών Μονάδων σε σχέση με το φόρτο εργασίας;
12. Διαφάνεια των κριτηρίων βαθμολόγησης.

Για την κατηγορία «**2. Στις περιπτώσεις όπου υπήρχαν γραπτές ή/και προφορικές εργασίες:**»

- 1=Καθόλου
- 2=Λίγο
- 3=Μέτρια
- 4=Πολύ
- 5=Πάρα πολύ
13. Το θέμα δόθηκε εγκαίρως;
14. Η καταληκτική ημερομηνία για υποβολή ή παρουσίαση των εργασιών ήταν λογική;
15. Υπήρχε σχετικό ερευνητικό υλικό στη βιβλιοθήκη;
16. Υπήρχε καθοδήγηση από το διδάσκοντα;
17. Τα σχόλια του διδάσκοντος ήταν εποικοδομητικά και αναλυτικά;
18. Δόθηκε η δυνατότητα βελτίωσης της εργασίας;
19. Η συγκεκριμένη εργασία σας βοήθησε να κατανοήσετε το συγκεκριμένο θέμα;

Για την κατηγορία «**3. Ο/Η διδάσκων/ουσα:**».

- 1=Καθόλου
- 2=Λίγο
- 3=Μέτρια
- 4=Πολύ
- 5=Πάρα πολύ
20. Οργανώνει καλά την παρουσίαση της ύλης στα μαθήματα;
21. Επιτυγχάνει να διεγείρει το ενδιαφέρον για το αντικείμενο του μαθήματος;
22. Αναλύει και παρουσιάζει τις έννοιες με τρόπο απλό χρησιμοποιώντας παραδείγματα;
23. Ενθαρρύνει τους φοιτητές να διατυπώνουν απορίες και ερωτήσεις για να αναπτύξουν την κρίση τους;

24. Ήταν συνεπής στις υποχρεώσεις του/της (παρουσία στα μαθήματα, έγκαιρη διόρθωση εργασιών ή εργαστηριακών αναφορών, ώρες συνεργασίας με τους φοιτητές);
25. Είναι γενικά προσιτός στους φοιτητές;

Για την κατηγορία «4. Εγώ ο/η φοιτητής/τρια:»


- 1=Καθόλου  
2=Λίγο  
3=Μέτρια  
4=Πολύ  
5=Πάρα πολύ
26. Παρακολουθώ τακτικά τις διαλέξεις.  
27. Μελετώ συστηματικά την ύλη.


Για την κατηγορία «5. Αξιολογήστε με άριστα το 10 πώς σας φάνηκε το:»


- 1ο μάθημα  
2ο μάθημα  
3ο μάθημα  
4ο μάθημα  
5ο μάθημα  
6ο μάθημα  
7ο μάθημα  
8ο μάθημα  
9ο μάθημα



Να σημειωθεί ότι για την κατηγορία 5 έχουν ορισθεί ως 10 οι πιθανές βαθμολογήσεις σε κάθε ερώτημα.


#### Editing Rate (scale 1..5) question

Question Name  Πώς σας φάνηκε το κάθε μάθι

Response Required   Yes  No

Nb of scale items  10

Type of rate scale  Normal 

Question Text\* 

Τέλος για ελεύθερη έκθεση παρατηρήσεων και σχολίων τοποθετούμε ένα Essay Box.

### Editing Essay Box question

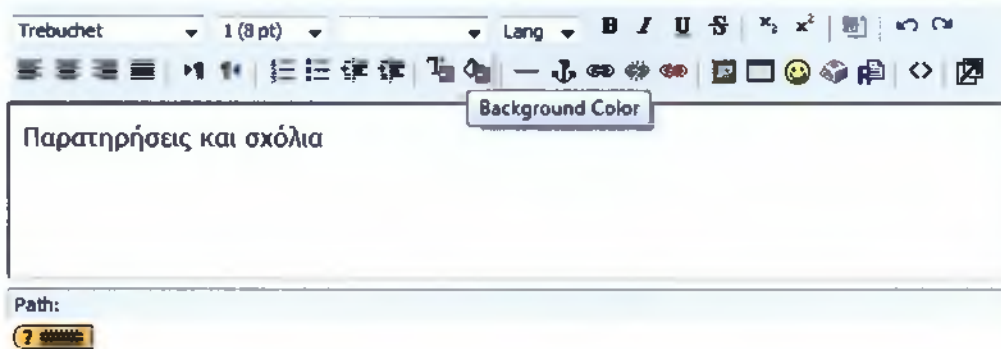
Question Name  Παρατηρήσεις και σχόλια




Response Required   Yes  No


Textarea columns  80

Textarea rows  8


Question Text\* 



Trebuchet 1 (8 pt) Lang **B** *I* U ~~S~~ [x](#) [x<sup>2</sup>](#)   

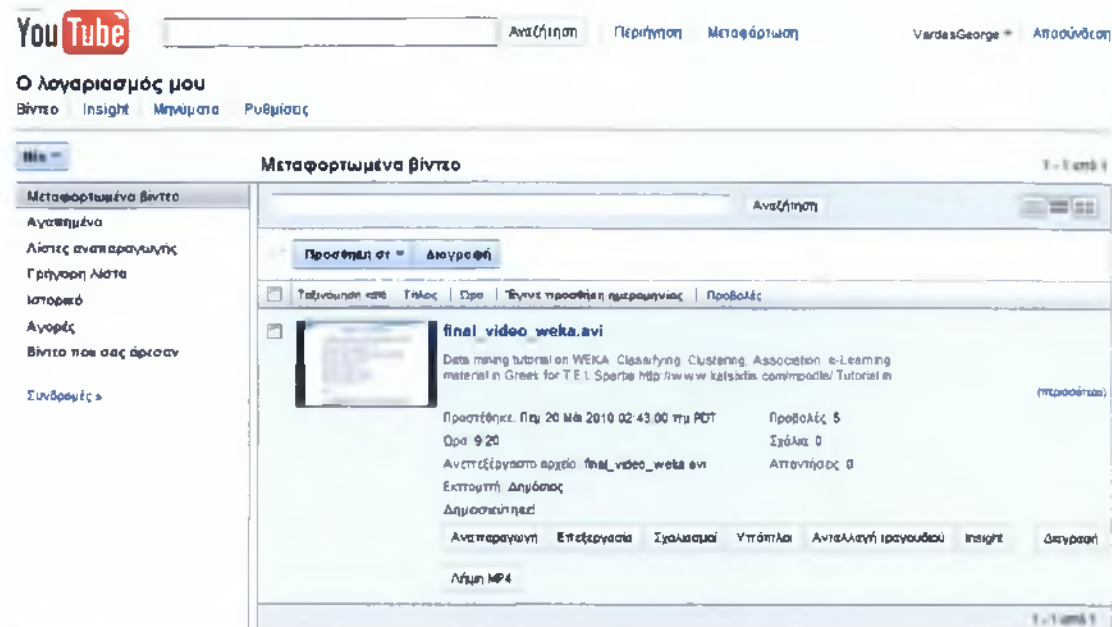
 Background Color

Παρατηρήσεις και σχόλια

Path: 

## Τοποθέτηση Video στο YouTube

Αρχικά πρέπει να δημιουργήσουμε έναν λογαριασμό στο YouTube και να ανεβάσουμε το αρχείο Video εκεί.



Κατόπιν και αφού έχουμε ορίσει ότι το video είναι public (δηλαδή μπορεί να το δει οποιοσδήποτε) και έχουμε φροντίσει να συμπληρώσουμε τα απαραίτητα μεταδεδομένα, χρησιμοποιούμε τον παρακάτω HTML κώδικα για να κάνουμε το Video embed σε μια σελίδα του Moodle:

```
<div style="text-align: center;">
  <object width="680" height="535"><param name="movie"
value="http://www.youtube.com/v/2UhlwOmz9Hk&hl=el&fs=1"
/><param name="allowFullScreen" value="true" /><param
name="allowscriptaccess" value="always" /><embed width="680"
height="535"
src="http://www.youtube.com/v/2UhlwOmz9Hk&hl=el&fs=1"
type="application/x-shockwave-flash" allowscriptaccess="always"
allowfullscreen="true" />
  </object></div>
```