



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ Μ/Υ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
ΣΧΟΛΗ ΝΑΥΤΙΛΙΑΣ ΚΑΙ ΒΙΟΜΗΧΑΝΙΑΣ
ΤΜΗΜΑΤΟΣ ΒΙΟΜΗΧΑΝΙΚΗΣ ΔΙΟΙΚΗΣΗΣ & ΤΕΧΝΟΛΟΓΙΑΣ
ΔΙΑΠΑΝΕΠΙΣΤΗΜΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
«ΤΕΧΝΟ-ΟΙΚΟΝΟΜΙΚΑ ΣΥΣΤΗΜΑΤΑ»



ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Ο ΡΟΛΟΣ ΤΗΣ ΕΠΙΣΤΗΜΗΣ ΔΕΔΟΜΕΝΩΝ ΣΤΗΝ ΤΕΤΑΡΤΗ ΒΙΟΜΗΧΑΝΙΚΗ ΕΠΑΝΑΣΤΑΣΗ

ΚΑΛΑΪΤΖΗΣ ΙΩΑΝΝΗΣ

ΕΠΙΒΛΕΠΩΝ

Βασίλειος Ασημακόπουλος, Καθηγητής Ε.Μ.Π.

Οκτώβριος 2022



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ Μ/Υ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

ΣΧΟΛΗ ΝΑΥΤΙΛΙΑΣ ΚΑΙ ΒΙΟΜΗΧΑΝΙΑΣ

ΤΜΗΜΑΤΟΣ ΒΙΟΜΗΧΑΝΙΚΗΣ ΔΙΟΙΚΗΣΗΣ & ΤΕΧΝΟΛΟΓΙΑΣ

ΔΙΑΠΑΝΕΠΙΣΤΗΜΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ

«ΤΕΧΝΟ-ΟΙΚΟΝΟΜΙΚΑ ΣΥΣΤΗΜΑΤΑ»



ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Ο ΡΟΛΟΣ ΤΗΣ ΕΠΙΣΤΗΜΗΣ ΔΕΔΟΜΕΝΩΝ ΣΤΗΝ ΤΕΤΑΡΤΗ ΒΙΟΜΗΧΑΝΙΚΗ ΕΠΑΝΑΣΤΑΣΗ

ΚΑΛΑΪΤΖΗΣ ΙΩΑΝΝΗΣ

ΕΠΙΒΛΕΠΩΝ

Βασίλειος Ασημακόπουλος, Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την Οκτωβρίου 2022.

(Υπογραφή)

.....

Βασίλειος Ασημακόπουλος
Καθηγητής Ε.Μ.Π.

(Υπογραφή)

.....

Ιωάννης Ψαρράς
Καθηγητής Ε.Μ.Π.

(Υπογραφή)

.....

Χρυσόστομος Δούκας
Αναπλ. Καθηγητής Ε.Μ.Π.

(Υπογραφή)

Οκτώβριος 2022

.....

Ιωάννης Καλαϊτζής

Διπλωματούχος Οικονομικών Επιστημών, Αριστοτελείου Πανεπιστημίου
Θεσσαλονίκης

Διπλωματούχος Διατμηματικού Προγράμματος Μεταπτυχιακών Σπουδών Τεχνο-
οικονομικών Συστημάτων

Copyright © Ιωάννης Καλαϊτζής, 2022.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Ευχαριστίες

Η παρούσα διπλωματική εργασία εκπονήθηκε κατά το ακαδημαϊκό έτος 2021-2022 στα πλαίσια της ολοκλήρωσης του Διατμηματικού Προγράμματος Μεταπτυχιακών Σπουδών (ΔΠΜΣ) "Τεχνο-Οικονομικά Συστήματα" που έχει οργανωθεί και υλοποιηθεί μέσα από τη συνεργασία της Σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών (ΗΜΜΥ) του Εθνικού Μετσόβιου Πολυτεχνείου (ΕΜΠ) και του Τμήματος Βιομηχανικής Διοίκησης και Τεχνολογίας του Πανεπιστημίου Πειραιώς.

Πριν την παρουσίαση της παρούσας διπλωματικής εργασίας, αισθάνομαι την ανάγκη να ευχαριστήσω ορισμένους ανθρώπους, χωρίς τη βοήθεια των οποίων δεν θα ήταν εφικτή η περάτωση της εν λόγω εργασίας.

Προσωπικά, θα ήθελα να εκφράσω τις ευχαριστίες μου στον επιβλέποντα Καθηγητή μου, κ. Βασίλειο Ασημακόπουλο, ο οποίος μου έδωσε την ευκαιρία να ασχοληθώ με τον πραγματικά ενδιαφέροντα τομέα των προβλέψεων στα πλαίσια της ερευνητικής δραστηριότητας της Μονάδας Προβλέψεων και Στρατηγικής.

Επίσης, θα ήθελα να ευχαριστήσω τον Καθηγητή κ. Ι. Ψαρρά και τον αναπληρωτή Καθηγητή, κ. Χ. Δούκα, για τη συμμετοχή τους στην επιτροπή εξέτασης της διπλωματικής εργασίας.

Θα ήθελα να ευχαριστήσω θερμά τον κ. Ευάγγελο Σπηλιώτη, ο οποίος ήταν παρών καθ' όλη τη διάρκεια εκπόνησης της διπλωματικής μου εργασίας, με πολύτιμες συμβουλές και παρατηρήσεις που έπαιξαν καθοριστικό ρόλο στην ολοκλήρωση της παρούσας εργασίας, καθώς και για την παροχή βοήθειας, γενικά γύρω από την ακαδημαϊκή εξέλιξή μου.

Κλείνοντας, θα ήθελα να ευχαριστήσω όλους εκείνους τους ανθρώπους από το κοντινό μου περιβάλλον που βρισκότουσαν κατά τη διάρκεια της φοίτησης μου στο πρόγραμμα αυτό δίπλα μου και με υποστήριζαν με τον δικό τους τρόπο, βοηθώντας με να πετύχω τους στόχους μου.

Περίληψη

Η παρούσα διπλωματική εργασία μελετά το ρόλο που έχει διαδραματίσει μέχρι σήμερα η επιστήμη των δεδομένων στην τέταρτη βιομηχανική επανάσταση καθώς και τις μελλοντικές της κατευθύνσεις. Παράλληλα, παραθέτει προτάσεις και παραδείγματα ψηφιακών λύσεων διαφορετικών τεχνολογιών που σχετίζονται με την εν λόγω επιστήμη.

Αρχικά περιγράφονται εισαγωγικές έννοιες της επιστήμης των δεδομένων και αναλύεται ο βαθμός εισχώρησής της σε κομβικούς τομείς της κοινωνίας σε παγκόσμιο αλλά και σε εθνικό επίπεδο. Ακόμη, περιγράφονται οι χρήσεις, οι μελλοντικές ευκαιρίες και τα οφέλη που μπορούν να υπάρξουν από αυτήν στην οικονομία, στην υγεία, στις μεταφορές και στο σχεδιασμό «έξυπνων» πόλεων. Στη συνέχεια προσδιορίζονται τα τρία κυριότερα επίπεδα μιας ομάδας ανάλυσης δεδομένων και οι ρόλοι που πηγάζουν από αυτά στα πλαίσια ενός σύγχρονου οργανισμού.

Αφού μελετηθεί η διαχείριση των δεδομένων από πλευράς οργάνωσης ανθρώπινου δυναμικού, εν συνεχεία παρουσιάζεται από λειτουργικής άποψης η διαδρομή που συνήθως ακολουθείται για την υλοποίησή της. Με βάση τα παραπάνω μελετάται η περίπτωση του συστήματος συστάσεων της εταιρείας Netflix. Η εν λόγω μελέτη οδηγεί στον εντοπισμό χρήσιμων παραδειγμάτων σωστής αξιοποίησης της επιστήμης των δεδομένων και στην εξαγωγή προτάσεων προς άλλες εταιρείες του κλάδου.

Τέλος, η εργασία παρουσιάζει σύγχρονα πληροφοριακά συστήματα, ψηφιακά εργαλεία και γλώσσες προγραμματισμού και συνοψίζει τα οφέλη που μπορούν να προκύψουν από αυτά με την ταυτόχρονη αξιοποίηση των δεδομένων, προσφέροντας παράλληλα προτάσεις εφαρμογής και ανάπτυξης ομάδων δεδομένων σε επιχειρησιακό πλαίσιο με στόχο την ομαλή και ασφαλή τους μετάβαση σε ένα σύγχρονο περιβάλλον δεδομένων.

Λέξεις Κλειδιά: Επιστήμη Δεδομένων, Τέταρτη Βιομηχανική Επανάσταση, Ανάλυση Δεδομένων, Επιχειρηματική Ευφυΐα, Ψηφιακά Εργαλεία

Abstract

The present thesis studies the role that data science has played so far in the fourth industrial revolution as well as its future directions. In addition, it lists proposals and examples of digital solutions of different technologies related to this science.

Initially, introductory concepts of data science are described and the degree of its penetration into important sectors of society at global and national level is analyzed. Furthermore, the uses, future opportunities and benefits that can be obtained from it in the economy, health, transportation, and the design of "smart" cities are described. Moreover, the three main levels of a data analysis team and the roles that flow from them within a modern organization are identified.

After studying the management of the data from the point of view of organization's human resources, the route that is usually followed for its implementation is then presented from an operational point of view. According to the above, the case of the recommendation system of the Netflix company is studied. This study leads to the identification of useful examples of the correct use of data science and the extraction of proposals to other companies in the sector.

Finally, the paper presents modern information systems, digital tools and programming languages and summarizes the benefits that can arise from these with the utilization of data, while offering proposals for the application and development of data groups in a business context with the aim of their smooth and safe transition to a modern data environment.

Key words: Data Science, The Fourth Industrial Revolution, Data Analysis, Business Intelligence, Digital Tools

Πίνακας περιεχομένων

Ευχαριστίες.....	1
Περίληψη.....	2
Λέξεις Κλειδιά	2
Πίνακας περιεχομένων.....	4
Ευρετήριο εικόνων	6
1. Η επιστήμη των δεδομένων στην 4η βιομηχανική επανάσταση: ευκαιρίες, προκλήσεις και προτάσεις	9
1.1. Η επιστήμη των δεδομένων και ο ρόλος της στην επιχειρηματική ευφυΐα	9
1.2. Βασικές έννοιες	12
1.2.1. Ανάλυση δεδομένων	12
1.2.2. Κατηγορίες δεδομένων	14
1.3. Επικρατούσα κατάσταση στον κόσμο και στην ελληνική πραγματικότητα	15
1.4. Ευκαιρίες, προκλήσεις και προτάσεις.....	21
1.4.1. Οικονομία-Επιχειρηματικότητα	21
1.4.2. Μεταφορές.....	27
1.4.3. Υγεία	33
1.4.4. Ενέργεια και έξυπνες πόλεις	40
2. Η αρχιτεκτονική μίας ομάδας διαχείρισης δεδομένων και ο ρόλος της στις επιχειρήσεις	49
2.1. Οι ρόλοι μίας ομάδας διαχείρισης.....	49
2.1.1. Τεχνικό επίπεδο	49
2.1.2. Λειτουργικό επίπεδο	53
2.1.3. Συνολικό επίπεδο και εξαγωγή αποτελεσμάτων	56
2.2. Κριτήρια σχεδιασμού ομάδων διαχείρισης δεδομένων.....	59
2.2.1. Τα μεγέθη και ο βαθμός αυτονομίας.....	61
2.2.2. Υποβολή αποτελεσμάτων	63
2.2.3. Πηγές δεδομένων.....	64
2.2.4. Το επίπεδο αξιοποίησης των δεδομένων από την επιχείρηση	66
3. Η διαδρομή των δεδομένων για τη δημιουργία προστιθέμενης αξίας.....	68
3.1 Τυπική διαδρομή.....	68
3.1.1. Συλλογή	68
3.1.2 Αποθήκευση	72

3.1.3. Επεξεργασία-Προετοιμασία.....	76
3.1.4. Ανάλυση	76
3.1.5. Αξιοποίηση	78
3.1.6. Σύνταξη αναφορών- Οπτικοποίηση.....	79
3.2. Μελέτη αρχιτεκτονικής συστήματος Netflix.....	81
3.2.1. Ιστορική αναδρομή	81
3.2.2. Σύστημα συστάσεων (recommendation system).....	84
4. Διαδικασίες και μέσα για την επίτευξη επιχειρησιακών στόχων σε ένα σύγχρονο περιβάλλον	89
4.1. Συστατικά στοιχεία επιχειρήσεων - Επιχειρησιακό πλάνο	89
4.2. Οικοσύστημα δεδομένων και διασύνδεση των εργαλείων.....	92
4.3. Συστήματα διαχείρισης πόρων (Enterprise Resource Planning)	95
4.4. Συστήματα διαχείρισης πελατειακών σχέσεων (Customer relationship management) ...	97
4.5. Γλώσσες προγραμματισμού στην επιστήμη δεδομένων	99
4.6. Εργαλεία διαχείρισης δεδομένων.....	101
5. Συμπεράσματα και προτάσεις	104
5.1. Συμπεράσματα	104
5.2. Προτάσεις	106
ΒΙΒΛΙΟΓΡΑΦΙΑ.....	114

Ευρετήριο εικόνων

Εικόνα 1. Ο ετήσιος όγκος δεδομένων παγκοσμίως και ο ρυθμός αύξησής του γεωγραφικά	10
Εικόνα 2. Η ετήσια αύξηση της υπολογιστικής δύναμης των υπερ-υπολογιστών παγκοσμίως [2]	11
Εικόνα 3. Η επαγγελματική προοπτική των Επιστημόνων Έρευνας Υπολογιστών και Πληροφοριών στις ΗΠΑ [11] Πηγή: Το επίσημο κυβερνητικό γραφείο στατιστικών για θέματα εργασίας και εργασιακών προγραμμάτων των ΗΠΑ	16
Εικόνα 4. Ο ρυθμός αύξησης της παγκόσμιας ροής δεδομένων μέσω διαδικτύου τα τελευταία 30 χρόνια [12]	17
Εικόνα 5.. Ο χάρτης των 174 χωρών που εξετάστηκαν στην Παγκόσμια Έκθεση Ανάπτυξης 2021 βάση των Στατιστικών Δεικτών Απόδοσης, οι οποίες χρωματίστηκαν βάση της βαθμολογίας τους. [12] Πηγή: Παγκόσμια Έκθεση Ανάπτυξης 2021, Παγκόσμια Τράπεζα.....	18
Εικόνα 6. Η αναλυτική βαθμολογία της Ελλάδας, στην Παγκόσμια Έκθεση Ανάπτυξης 2021 βάση των Στατιστικών Δεικτών Απόδοσης. [12].....	19
Εικόνα 7. το διάγραμμα του ρυθμού ανάπτυξης του κλάδου ποσοστιαία σε εθνικό επίπεδο και σε σύγκριση με εκείνο της Ε.Ε.[13].....	20
Εικόνα 8. Αποτυπώνεται μια οπτικοποίηση σε μορφή χάρτη, της αποτίμησης σε εκατομμύρια ευρώ, των βιομηχανιών επεξεργασίας και φιλοξενίας δεδομένων ανά χώρα της Ε.Ε. πλην Τσεχίας και Ολλανδίας. [13].....	20
Εικόνα 9. Έρευνα με δείγμα με επαγγελματίες της Επιχειρηματικής Ευφυίας και Επιστήμης Δεδομένων, για το τι εμποδίζει τις εταιρείες εξελιχθούν σε ένα επιχειρησιακό μοντέλο βασισμένο περισσότερο στα δεδομένα [15]	24
Εικόνα 10. Έρευνα με δείγμα επαγγελματίες της Επιχειρηματικής Ευφυίας και Επιστήμης Δεδομένων, για τα βήματα που πρέπει να ακολουθήσει η εταιρεία ώστε εξελιχθεί σε ένα επιχειρησιακό μοντέλο βασισμένο περισσότερο στα δεδομένα[15]	25
Εικόνα 11. Data-Driven Business Mode. [12] Source: World Development Report 2021, World Bank	26
Εικόνα 12. Επίπεδα αυτοματισμού οχημάτων σύμφωνα με την SAE International και η διάκριση των ρόλων μεταξύ οδηγού και αυτοκινήτου στα επίπεδα αυτά. [17]	30

Εικόνα 13. Data-Driven Μοντέλο Οδικής Κυκλοφορίας. Source: Mobility Patterns, Big Data and Transport Analytics, Tools and Applications for Modeling [16]	31
Εικόνα 14: Οι πηγές των δεδομένων της υγείας στα Health Informatics [20] Source: Health Informatics: Practical Guide.....	34
Εικόνα 15. Το προσδόκιμο ζωής του πλανήτη από την έναρξη 20ου αιώνα έως σήμερα μέσω οπτικοποίησης χάρτη (πάνω) [2].....	37
Εικόνα 16. Ο ρυθμός μετάδοσης και θανάτων ορισμένων ασθενειών με κύρια γεγονότα την εισαγωγή εμβολίων ως μέτρο πρόληψης [2].....	38
Εικόνα 17. Η αστικοποίηση με την πάροδο του χρόνου και προβλέψεις έως το 2050 παγκοσμίως[2]	41
Εικόνα 18. Θεωρητικό μοντέλο μία έξυπνης πόλης του μέλλοντος Source: Aliga IT Company	45
Εικόνα 19. Θεωρητικό μοντέλο ενός έξυπνου κτηρίου όπου με γαλάζιο χρώμα είναι χρωματισμένες οι λειτουργικές τεχνολογίες ενός "έξυπνου" κτιρίου και με κίτρινο οι εξωτερικές πηγές πληροφοριών και λειτουργιών. Source: Johnson Controls (Group of Companies).....	46
Εικόνα 20. Θεωρητικό μοντέλο αγωγού δεδομένων (data pipeline) Source: Pentaho Corporation	51
Εικόνα 21. Θεωρητική οπτική προσέγγιση της λειτουργίας μίας ομάδας δεδομένων εντός τους οργανισμού. Το Λειτουργικό Επίπεδο υποστηρίζεται από τα άλλα δύο μέρη και αποστέλλει κατευθύνσεις.....	59
Εικόνα 22. Οπτικοποίηση των ανωτέρω μεγεθών	62
Εικόνα 23. Θεωρητική προσέγγιση της ιεραρχίας μίας αυτόνομης ομάδας διαχείρισης δεδομένων	63
Εικόνα 24. Διαμόρφωση της ομάδας δεδομένων βάση του επιπέδου αξιοποίησης δεδομένων από της επιχείρηση	67
Εικόνα 25. Θεωρητική διαδρομή των δεδομένων	68
Εικόνα 26. Συλλογής Δεδομένων Source: QuestionPro	72
Εικόνα 27. Παραδείγματα των Διαγραμμάτων Source: Microsoft Excel.....	81

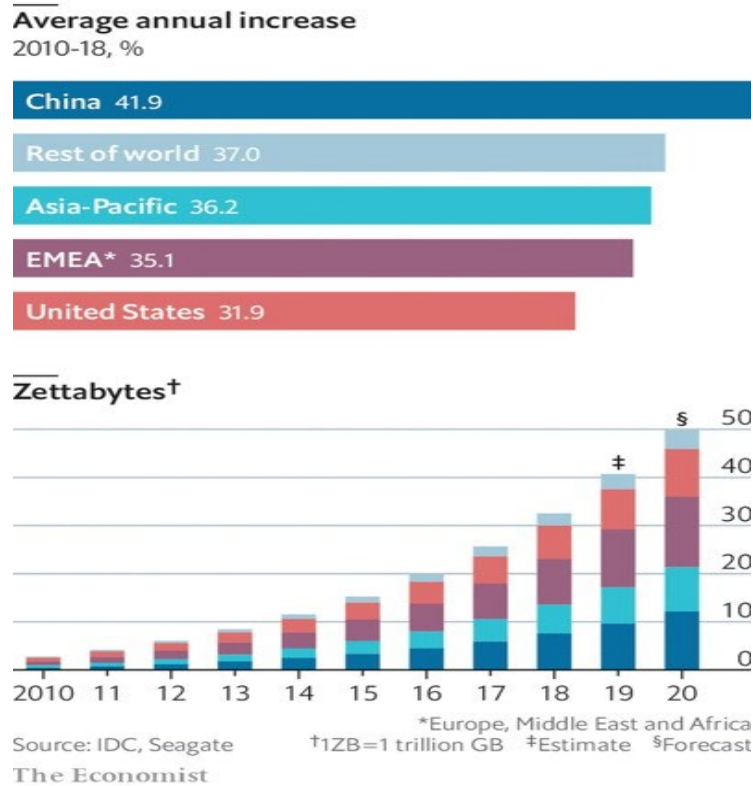
Εικόνα 28. Προσβασιμότητα στην πλατφόρμα Netflix σε όλο τον κόσμο. Με κόκκινο παρουσιάζονται οι χώρες που έχουν πρόσβαση ενώ με μαύρο αυτές που δεν έχουν. Source: Netflix’s Help Center για τον Μάρτιο 2022.[35]	84
Εικόνα 29. Το μοντέλο αρχιτεκτονικής του συστήματος σύστασης της Netflix[34]	88
Εικόνα 30. Απλουστευμένο μοντέλο επιχειρησιακής λειτουργίας	89
Εικόνα 31. Συστατικά στοιχεία μίας επιχείρησης.....	91
Εικόνα 32. Παραδοσιακό δίκτυο έναντι δικτύου υπολογιστικού νέφους.[37] Source: Alibaba Cloud	94
Εικόνα 33. Τα επίπεδα ανάθεσης υπηρεσιών υπολογιστικού νέφους σε άλλη εταιρεία Source: www.medium.com “Types of Cloud Services” by IDM	95
Εικόνα 34. Ροή δεδομένων με επίκεντρο ένα ERP σύστημα [40] Πηγή: Σύγχρονα Πληροφοριακά Συστήματα Επιχειρήσεων 2015, του Πάνου Φιτσιλή	97
Εικόνα 35. Πηγές δεδομένων μέσω CRM *(AM:Αυτοματοποιημένο Μάρκετινγκ) [40] Πηγή: Σύγχρονα Πληροφοριακά Συστήματα Επιχειρήσεων 2015, του Πάνου Φιτσιλή..	99
Εικόνα 36. Διάγραμμα με τις πιο δημοφιλείς γλώσσες προγραμματισμού στο πεδίο της επιστήμης των δεδομένων.	100
Εικόνα 37. Ο δρόμος για την μετάβαση στην Βιομηχανία 4.0.....	105
Εικόνα 38. Βασικές επιλογές επένδυσης μίας επιχείρησης.....	107
Εικόνα 39. Σταδιακή ανάπτυξης της ομάδας διαχείρισης δεδομένων εντός της επιχείρησης	110
Εικόνα 40. Θεωρητικό μοντέλο Πιλοτικής Εφαρμογής	111

1. Η επιστήμη των δεδομένων στην 4η βιομηχανική επανάσταση: ευκαιρίες, προκλήσεις και προτάσεις

1.1. Η επιστήμη των δεδομένων και ο ρόλος της στην επιχειρηματική ευφυΐα

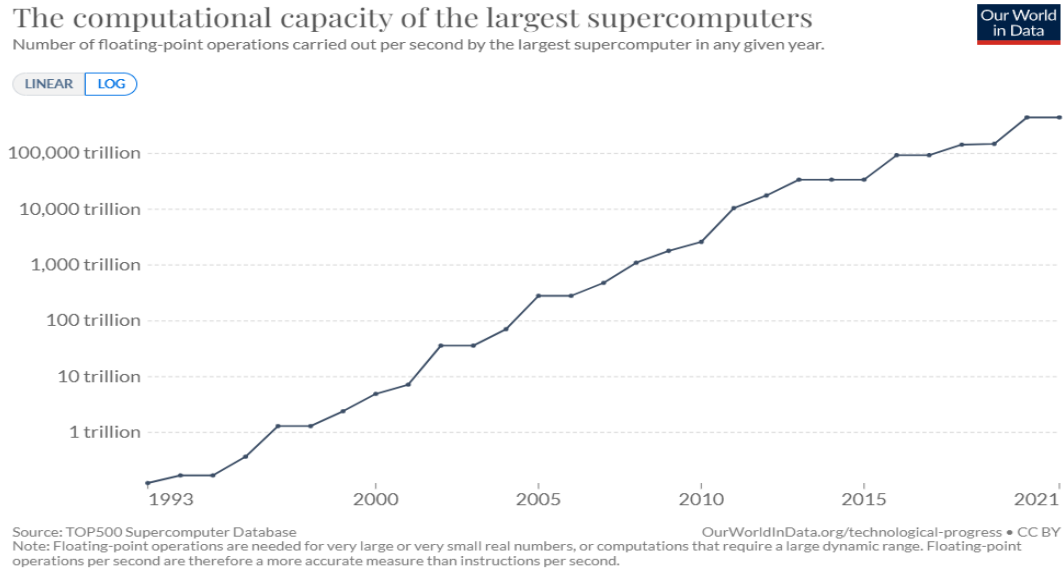
Ολοένα και περισσότερο ο όρος “Επιστήμη Δεδομένων” γίνεται γνωστός και διαδίδεται με αυξανόμενο ρυθμό σε πολλούς τομείς της κοινωνίας και κυρίως στο τομέα της οικονομίας και των επιχειρήσεων. Και αυτό επειδή αποτέλεσε λύση σε πολλές προκλήσεις της σύγχρονης εποχής και ανταγωνιστικό πλεονέκτημα όσων επιχειρήσεων την ενέταξαν στις διαδικασίες τους. Μέσω αυτής πολλοί οργανισμοί κατάφεραν να εκσυγχρονιστούν και να προσαρμοστούν ή να δημιουργηθούν πολλές νεοφυείς επιχειρήσεις (start-ups) οι οποίες προώθησαν καινοτόμα προϊόντα και υπηρεσίες με επιτυχία και μεγάλη οικονομική ανάπτυξη. Έτσι άρχισε να δημιουργείται σταδιακά μία νέα οικονομία όσο αφορά την εκμετάλλευση των δεδομένων, τα οποία μπορούν να έχουν χαρακτήρα ως χρήσιμα εργαλεία για την εκπόνηση των επιχειρησιακών εργασιών για την παραγωγή προϊόντων ή υπηρεσιών μέχρι και να αποτελέσουν το ίδιο το προϊόν. [5]

Αρχικά είναι ένα αρκετά διευρυμένο πεδίο έρευνας το οποίο αποτελείται από στοιχεία άλλων επιστημονικών πεδίων όπως της στατιστικής, της επιστήμης των υπολογιστών και των μαθηματικών και περιλαμβάνει κυρίως την στατιστική ανάλυση, την εξόρυξη, την επεξεργασία και την οπτικοποίηση δεδομένων, την μηχανική μάθηση και γενικότερα τον εντοπισμό μοτίβων και σχέσεων ανάμεσα στα ποικίλα και πολυάριθμα δεδομένα. Ο όγκος και ο ρυθμός αύξησης των δεδομένων την τελευταία δεκαετία είναι κατακόρυφοι καθώς μετά το 2018 ξεπέρασαν τα 30 Zettabytes ο όγκος των δεδομένων ενώ το 2020 προσέγγισε τα 50 Zettabytes σύμφωνα με άρθρο του The Economist (A deluge of data is giving rise to a new economy, Φεβρουάριος 2020) όπου μεγαλύτερος ρυθμός παραγωγής παρουσιάζεται από την Κίνα.[1]



Εικόνα 1. Ο ετήσιος όγκος δεδομένων παγκοσμίως και ο ρυθμός αύξησής του γεωγραφικά

Αυτή η εκτίναξη τα τελευταία χρόνια έλαβε χώρο κυρίως εξαιτίας της διάδοσης του διαδικτύου σε όλο σχεδόν τον πλανήτη αλλά και ταυτόχρονα της μεγάλης ανάπτυξης της υπολογιστικής δύναμης, δηλαδή ισχυρότερες αποθηκευτικές μνήμες υπολογιστών, καλύτεροι επεξεργαστές και πιο εξελιγμένο λογισμικό. [2] Σημαντικό, επίσης, ρόλο παίζει και ο ανθρώπινος παράγοντας όπου με την πάροδο του χρόνου, υπάρχουν ολοένα και περισσότερα άτομα με τις κατάλληλες ικανότητες και γνώσεις ώστε να συνδυάσουν όλα τα προηγούμενα, δημιουργώντας έτσι αξία και να υποστηρίξουν αυτήν την εκτίναξη. [1]



Εικόνα 2. Η ετήσια αύξηση της υπολογιστικής δύναμης των υπερ-υπολογιστών παγκοσμίως [2]

Για αυτό έχουν δημιουργηθεί αρκετοί διακριτοί ρόλοι στην οικονομία των δεδομένων όπου ο καθένας εξειδικεύεται σε ένα συγκεκριμένο κομμάτι της αλυσίδας της διαχείρισης των πληροφοριών και αρχίζει να παύει η θεώρηση, ένας επιστήμονας δεδομένων να είναι υπεύθυνος για το κομμάτι των πληροφοριών ενός οργανισμού με υποστήριξη μηχανικών υπολογιστών και αναλυτών. Όλοι αυτοί οι ρόλοι αποτελούν ένα πολύ μεγάλο κομμάτι της “Επιχειρηματικής Ευφυΐας”, αφού αυτή περιλαμβάνει τεχνολογικές πρακτικές για την ανάλυση, πρόβλεψη, αναφορών καθώς και κατάλληλη παρουσίαση αυτών, με σκοπό την αποτελεσματικότερη υποστήριξη λήψης αποφάσεων και σχεδιασμού στρατηγικών.[3] Για να υποστηριχθεί έχουν δημιουργηθεί διάφορα “εργαλεία”, όπως:

- πληροφοριακά συστήματα (π.χ. ERP, CRM)
- συστήματα διαχείρισης βάσεων δεδομένων (π.χ. MySQL)
- διαδικτυακές ψηφιακές πλατφόρμες με χρήση Cloud (π.χ. Looker)
- λογισμικά εξειδικευμένα για συγκεκριμένες διαδικασίες όπως η οπτικοποίηση των δεδομένων ή μηχανική μάθηση (π.χ. Power BI, TensorFlow).

Σε όλα αυτά τα εργαλεία έχουν επενδυθεί τεράστια κεφάλαια από κολοσσούς της τεχνολογίας όπως η IBM, η Google ή έχουν ξεπροβάλει από νεοφυείς εταιρείες με καινοτόμες ιδέες (εξασφαλίζοντας κεφάλαια συνήθως από επενδυτές ή τραπεζικά ιδρύματα), οι οποίες εταιρείες στην συνέχεια αφού τα προσαρμόσουν με τις ανάγκες της

οικονομίας, τα προωθούν στην αγορά τα οποία αγοράζουν οι οργανισμοί με την σειρά τους, δαπανώντας συνεχώς σε συστήματα Επιχειρηματικής Ευφυΐας μέσω των οποίων μπορούν να διαχειριστούν αποτελεσματικότερα τα δεδομένα τους και να δημιουργήσουν επιπρόσθετη αξία στα προϊόντα τους ή στις υπηρεσίες τους, αποφεύγοντας έτσι να θεωρηθούν παραχωρημένοι έναντι τυχόν ανταγωνιστών και της ίδιας της τεχνολογικής εξέλιξης. Με αυτόν τον τρόπο, τα προϊόντα και υπηρεσίες γίνονται ολοένα πιο προσωποποιημένα και ικανοποιούν σε μεγαλύτερο βαθμό τις ανάγκες του κόσμου, φτάνοντας σε σημείο ακόμα και να τις προλαμβάνουν και ταυτόχρονα να ελαττώνουν τα κόστη μέσω της βελτιστοποίησης των διαδικασιών, όπως για παράδειγμα την καλύτερη αναπλήρωση αποθεμάτων μειώνοντας έτσι τυχόν φύρα ή κόστη αποθήκευσης.[3]

1.2. Βασικές έννοιες

1.2.1. Ανάλυση δεδομένων

Αρχικά τα δεδομένα αποτελούν μη επεξεργασμένα στοιχεία τα οποία χαρακτηρίζονται από ποικιλομορφία όπως λέξεις, αριθμοί, βίντεο, εικόνες, ήχοι, και δεν έχει αξιολογηθεί η σημαντικότητά τους και αν μπορούν να αξιοποιηθούν. Αν μπορέσουν να προσφέρουν αξία μέσω της κατάλληλης επεξεργασίας και να μπορέσουν να απαντήσουν σε ερωτήματα τότε έχουν μετατραπεί σε πληροφορία. Η ανωτέρω επεξεργασία δεδομένων περιλαμβάνει την επιλογή των κατάλληλων δεδομένων από όλα αυτά που έχουμε στην διάθεσή μας και απαιτούνται για την λύση ενός προβλήματος, να τα “καθαρίσουμε” από τυχόν “θόρυβο” όπως ακραίες τιμές ή κάποιο ειδικό γεγονός και έτσι δίνοντας τους μία κατανοητή μορφή ώστε να μπορούν να υποστηρίξουν τον υπόλοιπο οργανισμό και την λήψη αποφάσεων. Αυτή η επεξεργασία είναι μέρος της ανάλυση δεδομένων. Στην ανάλυση δεδομένων συναντάμε τους παρακάτω τρεις κύριους τύπους [4]:

- Περιγραφική Αναλυτική (Descriptive Analytics)
- Διαγνωστική Αναλυτική (Diagnostic Analytics)
- Προγνωστική Αναλυτική (Predictive Analytics)
- Καθοδηγητική Αναλυτική (Prescriptive Analytics)

Η Περιγραφική Αναλυτική (Descriptive Analytics) απαντάει σε ερωτήματα τύπου “Τι συνέβη;” με χρήση ιστορικών στοιχείων. Καταλήγει σε απλά συμπεράσματα ωστόσο είναι η πιο αξιόπιστη καθώς οι περιγραφές του αναφέρονται στο παρελθόν και στο παρόν. Είναι το πρώτο στάδιο της ανάλυσης και αποφεύγει να ερμηνεύσει και να γενικεύσει καθώς αυτό θα απαιτούσε και άλλη προσπάθεια.

Η Διαγνωστική Αναλυτική (Diagnostic Analytics) ουσιαστικά είναι μέρος του προηγούμενου βήματος και επικεντρώνεται στους λόγους που συνέβη κάτι δηλαδή στο “γιατί.” Ο παρατηρητής εξερευνώντας τα δεδομένα, προσπαθεί να εντοπίσει τους λόγους που συνέβη κάτι να κατανοήσει αν υπάρχει αιτιότητα ή απλή συσχέτιση μεταξύ τους, καθώς ο διαχωρισμός αυτών των εννοιών και κατά επέκταση η αναγνώρισή τους στις σχέσεις των δεδομένων του προβλήματος αποτελεί μείζον ζήτημα. Γενικότερα, να καταλήξει σε κάποιο συμπέρασμα του προηγούμενου τύπου.

Η Προγνωστική Αναλυτική (Predictive Analytics) προσπαθεί να προβλέψει το τι θα συμβεί και η ακρίβεια της βασίζεται στην κατάλληλη επιλογή μεταβλητών και το κατά πόσο επηρεάζουν την τελική πρόβλεψη. Οι προγνώσεις (forecasting) αυτές βασίζονται σε προηγούμενα δεδομένα και στα διάφορα ποιοτικά χαρακτηριστικά της χρονοσειράς των δεδομένων που έχει διαμορφωθεί και όχι σε προβλέψεις (prediction) που βασίζονται στην εμπειρική ανάλυση από άτομα ακόμη και αυτά να είναι υψηλά καταρτισμένα άμεσα με το πεδίο προβλέψεων. Αυτή η μέθοδος επιτυγχάνεται κυρίως είτε μέσω της στατιστικής ανάλυσης και των κατάλληλων μοντέλων της είτε προσομοιώνοντας ένα σενάριο δίνοντάς του όσα περισσότερα δεδομένα ώστε να ανταποκρίνονται στην πραγματικότητα. Με την ανάπτυξη της αποθηκευτικής δυνατότητας και επεξεργασίας, οι προσομοιώσεις θα γίνονται ολοένα πιο διαδεδομένες και στο μέλλον θα αποτελούν ένα αναπόσπαστο και αξιόπιστο εργαλείο προβλέψεων.

Τέλος, η Καθοδηγητική Αναλυτική (Prescriptive Analytics) προσπαθεί να καθοδηγήσει τον παρατηρητή στο τι θα πρέπει να κάνει αν προκύψει κάτι. Είναι σημαντική η γνώση σχέσεων μεταξύ των διάφορων μεταβλητών και ο σωστός χαρακτηρισμός τους κατά το στάδιο της Διαγνωστικής Ανάλυσης καθώς και η ύπαρξη ενός κατάλληλου μηχανισμού μηχανικής μάθησης που θα έχει διαμορφωθεί από τα προηγούμενα βήματα(ιστορικά στοιχεία, η απόκλιση των προγνώσεων μας από τα πραγματικά) ώστε να

είναι δυνατή η επεξεργασία υποθετικών σεναρίων και η παροχή χρήσιμων και αξιόπιστων συμβουλών.

1.2.2. Κατηγορίες δεδομένων

Στη συνέχεια υπάρχουν πολλοί όροι που έχουν ως συνθετικό τη λέξη “δεδομένα” και έχουν γίνει γνωστοί στην κοινωνία μας τα τελευταία χρόνια χωρίς να είναι πλήρως κατανοητοί. Μερικούς από τους πιο διαδεδομένους όρους, είναι τα Μεγάλα Δεδομένα (Big Data), τα Ανοιχτά Δεδομένα (Open Data), τα Προσωπικά Δεδομένα (Personal Data), και Μετα-Δεδομένα (MetaData)

Τα Μεγάλα Δεδομένα (Big Data) είναι αυτά τα οποία οι συμβατικές βάσεις δεδομένων δεν μπορούν να επεξεργαστούν καθώς ο όγκος τους είναι τεράστιος και ανανεώνονται συνεχώς με πολύ γρήγορο ρυθμό από πολλές και διαφορετικές πηγές. Έχοντας την δυνατότητα να καταφέρουμε σε να επεξεργαστούμε τα δεδομένα σε πραγματικά χρόνο, προωθείται ένα δυναμικό μοντέλο και ξεπερνάει την στάσιμη στατιστική προσέγγιση. Χαρακτηρίζονται από τρία χαρακτηριστικά, τα 3Vs [6] ενώ σε κάποιες βιβλιογραφίες θα συναντήσουμε να αναλύονται σε πέντε, δηλαδή 5Vs [7]. Αυτά είναι η Ένταση(Volume), η Ταχύτητα(Velocity), η Ποικιλία(Variety) και επίσης η Αξιοπιστία(Veracity) και τέλος να μπορεί μέσω όλων αυτών να δημιουργηθεί Αξία(Value) που είναι και ο τελικός σκοπός της ενασχόλησης με τα Μεγάλα Δεδομένα.

Ανοιχτά Δεδομένα (Open Data) μπορούν να χαρακτηριστούν σύμφωνα με την παγκόσμια Μη Κερδοσκοπική Οργάνωση (Open Knowledge Foundation)[8], με υποκατάστημα και στην χώρα μας, τα στοιχεία αυτά που ακολουθούν αυτά τα παρακάτω χαρακτηριστικά:

- Διαθεσιμότητα και πρόσβαση για όλους
- Επαναχρησιμοποίηση και αναδιανομή
- Καθολική συμμετοχή χωρίς περιορισμούς

Ουσιαστικά έχει πρόσβαση σε αυτά ο καθένας ελεύθερα χωρίς να υπόκεινται σε πνευματικά δικαιώματα τα οποία μπορεί να είναι από λέξεις, αριθμοί, εικόνες μέχρι και έρευνες ή γλώσσα αναγνώσιμη από υπολογιστές, δηλαδή κάποιος κώδικας. Έχουν

δυναμικές χρήσεις και εφαρμογές σε διάφορους τομείς της κοινωνίας όπως στον Πολιτισμό και τις Τέχνες, στις Επιστήμες, στην Οικονομία, στις Στατιστικές Έρευνες, στην Μετεωρολογία ακόμη στη Δημόσια Διοίκηση όπου και άνοιξε τον δρόμο για την ανοικτή διακυβέρνηση, εκδημοκρατίζοντας περαιτέρω τις κοινωνίες μας.

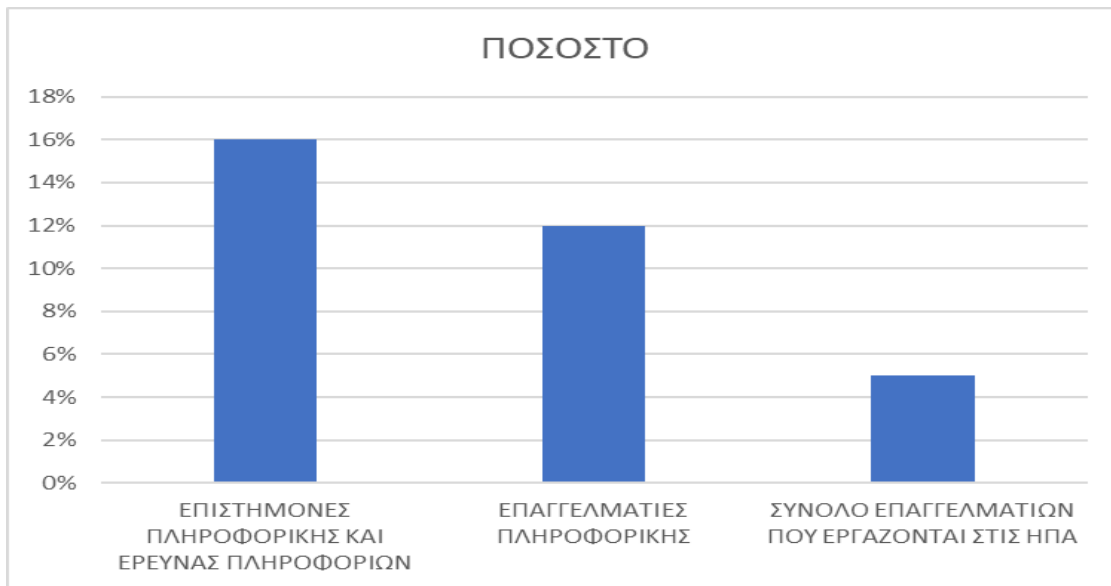
Τα Προσωπικά Δεδομένα ή Δεδομένα Προσωπικού Χαρακτήρα σύμφωνα με τον επίσημο Ιστότοπο της Ευρωπαϊκής Ένωσης [9] και το Ευρωπαϊκό Ενωσιακό νομοθετικό πλαίσιο θεωρούνται: το ονοματεπώνυμο, στοιχεία επικοινωνίας, τόπος διαμονής, τραπεζικά στοιχεία, ιατρικά στοιχεία κ.λπ., δηλαδή όλες αυτές οι πληροφορίες που συνθέτουν ένα μοναδικό άτομο και αποτελούν ένας είδος ταυτότητάς του. Σύμφωνα με τον Γενικό Κανονισμό Προστασίας Δεδομένων(ΓΚΠΔ) αυτά τα δεδομένα έχουν καταστεί ανώνυμα μέσω κρυπτογράφησης ή κάποια άλλης πρακτικής και πρέπει να παραμένουν έτσι. Αυτό ονομάζεται ανωνυμοποίηση και ουσιαστικά δεν είναι δυνατή η εύρεση του κατόχου των δεδομένων από βάσεις δεδομένων ή διάφορες εγγραφές σε καταλόγους. Στη χώρα μας την ευθύνη για την τήρηση της ανωνυμίας αυτής δηλαδή της προστασίας δεδομένων σε κοινωνικό επίπεδο είναι η ανεξάρτητη Αρχή Προστασίας Δεδομένων Προσωπικού Χαρακτήρα που συστάθηκε το 1997.

Τα Μετα-Δεδομένα (MetaData) είναι δεδομένα τα οποία χαρακτηρίζουν και προσδιορίζουν άλλα δεδομένα [10]. Είναι ιδιαίτερα σημαντικά έτσι ώστε να μπορεί ένας εξωτερικός παρατηρητής να μπορεί να αντιληφθεί και να αναγνωρίσει σε τι αναφέρονται τα στοιχεία που εξετάζει καθώς και να διευκολύνεται η ταξινόμησή τους. Μπορεί να περιγράφουν φυσικά χαρακτηριστικά ή να προσδιορίζουν την πηγή τους ή δομικά που ανήκουν. Επιπρόσθετα, συνήθως αποτελούν παράγωγο ανάλυσης των αρχικών δεδομένων, οπότε βοηθάνε στην καλύτερη κατανόηση των χαρακτηριστικών τους.

1.3. Επικρατούσα κατάσταση στον κόσμο και στην ελληνική πραγματικότητα

Σήμερα και ιδιαίτερα στις τεχνολογικά αναπτυγμένες χώρες οι περισσότεροι μεγάλοι οργανισμοί έχουν ήδη αντιληφθεί την σημασία της επιστήμης των δεδομένων και αφιερώνουν πολλούς πόρους σε έργα με αντικείμενο αυτήν, δημιουργώντας και αυτόνομα τμήμα, επανδρώνοντας τα με επαγγελματίες του χώρου. Η αυξανόμενη δυναμική αυτού

του εργασιακού κλάδου αποτυπώνεται και στην επίσημη ιστοσελίδα του Γραφείου Στατιστικών Εργασίας των ΗΠΑ [11], όπου αναμένεται οι θέσεις εργασίας που αφορούν το επάγγελμα διαχείρισης δεδομένων θα αυξηθούν κατά 14% στην αμερικάνικη αγορά την τωρινή δεκαετία, 2020-2030. Επίσης, ο ίδιος φορέας δημοσίευσε έρευνα όπου για την ίδια δεκαετία, οι Επιστήμονες Έρευνας Υπολογιστών και Πληροφοριών θα αυξηθούν ποσοστιαία κατά 22% όταν ο μέσος όρος ποσοστιαίας αύξησης όλων των επαγγελματιών πληροφορικής θα αυξηθεί κατά 13% ενώ γενικά όλες οι εργασιακές θέσεις θα αυξηθούν κατά 5%.



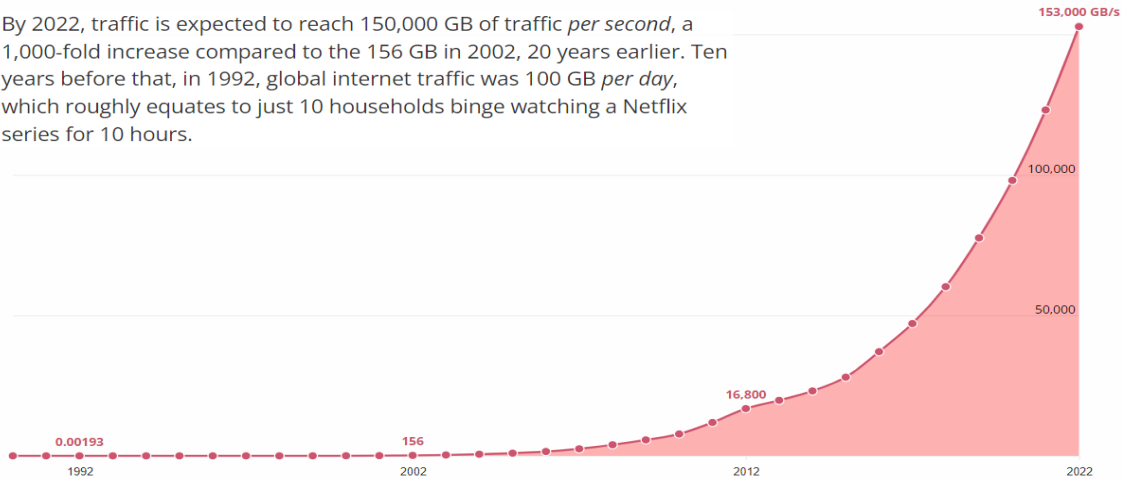
Εικόνα 3. Η επαγγελματική προοπτική των Επιστημόνων Έρευνας Υπολογιστών και Πληροφοριών στις ΗΠΑ [11] Πηγή: Το επίσημο κυβερνητικό γραφείο στατιστικών για θέματα εργασίας και εργασιακών προγραμμάτων των ΗΠΑ

Από τα ανωτέρα προκύπτει ότι η συγκεκριμένη επιστήμη συμβαδίζει με την ανάπτυξη της οικονομίας και αποτελεί σημαντικό στοιχείο του εκσυγχρονισμού της και της μετάβασης της στην επόμενη γενιά, της Τέταρτης Βιομηχανικής Επανάστασης (Industry 4.0). Για παράδειγμα, η Παγκόσμια Τράπεζα το 2021 στην έκθεσή της για την Παγκόσμια Ανάπτυξη, που εκδίδει κάθε χρόνο από το 1978 και εξετάζει καίριες πλευρές της παγκόσμιας οικονομίας, αποφάσισε να αναλύσει την προοπτική βελτίωσης του βιοτικού επιπέδου των ανθρώπων και ιδιαίτερα εκείνων που ζουν στα όρια της φτώχειας και τις μεγάλες αλλαγές στο σύγχρονο χρηματοοικονομικό, κοινωνικό και γεωπολιτικό πεδίο του κόσμου μέσω της αποτελεσματικής διαχείρισης των δεδομένων. Συγκεκριμένα, στην έκθεσή αυτή με όνομα “Δεδομένα για Καλύτερες Ζωές” [12] τονίζει με δεικτικό

τρόπο σε ένα σημείο της ότι η ροή δεδομένων στο διαδίκτυο καθημερινώς είναι τόσο μεγάλη σε σχέση με το παρελθόν, όπου η παγκόσμια ροή δεδομένων στο διαδίκτυο το 1992 ισοδυναμεί με δέκα συνεχόμενες ώρες θέασης τηλεοπτικής σειράς σε μια ψυχαγωγική πλατφόρμας ταινιών και σειρών [12].

Growth of global internet traffic in the past 30 years

By 2022, traffic is expected to reach 150,000 GB of traffic *per second*, a 1,000-fold increase compared to the 156 GB in 2002, 20 years earlier. Ten years before that, in 1992, global internet traffic was 100 GB *per day*, which roughly equates to just 10 households binge watching a Netflix series for 10 hours.



Source: WDR 2021 team calculations and Cisco Visual Networking Index: Forecast and Trends, 2017-2022.

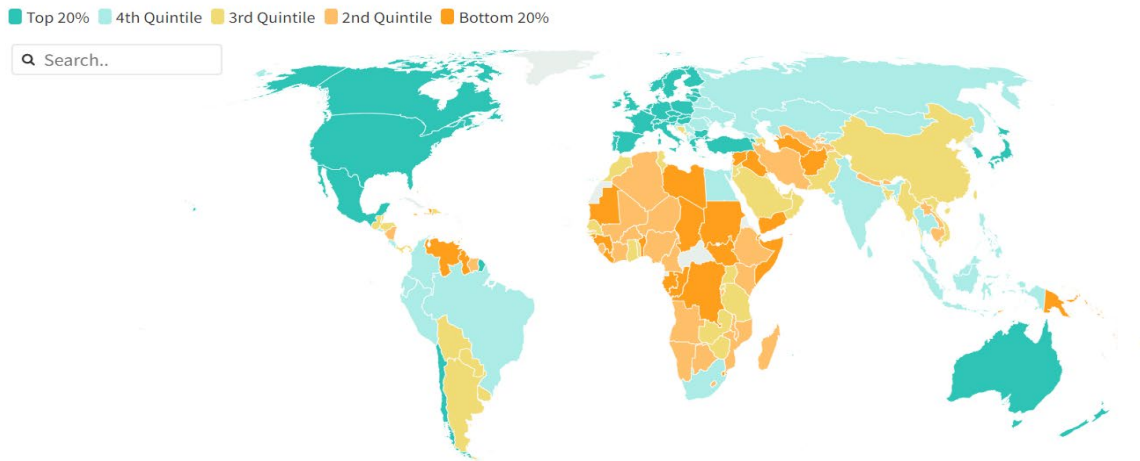
Εικόνα 4. Ο ρυθμός αύξησης της παγκόσμιας ροής δεδομένων μέσω διαδικτύου τα τελευταία 30 χρόνια [12]

Αυτή η μεγάλη διαφορά μεγθών καταδεικνύει την ανάγκη για την χρήση “έξυπνων εργαλείων” και είναι μονόδρομος η μετάβαση της συμβατικής ανάλυσης δεδομένων σε εκείνη που συνοδεύεται με την επιχειρηματική ευφυΐα. Επίσης, παρατηρώντας το ανωτέρω διάγραμμα αξίζει να σημειωθεί, ότι σε αυτήν την απότομη αλλαγή μεγθών σημαντικό ρόλο έπαιξαν τα Μέσα Κοινωνικής Δικτύωσης αφού η καμπύλη αρχίζει να γείρει απότομα προς τα πάνω από το 2006 και μετά, όπου προς στο τέλος αυτής της χρονιάς το Facebook έγινε προσβάσιμο σε όλο τον πλανήτη και επιπλέον δημιουργήθηκε και το Twitter. Στην συνέχεια, παρόμοιες τέτοιου είδους δωρεάν ψηφιακές υπηρεσίες δημοσιεύτηκαν, ενισχύοντας περαιτέρω τον ανωτέρω ρυθμό πληροφοριών και αποτελούν μία από τις σημαντικότερες πηγές παραγωγής δεδομένων, με αναγνωρισμένη και ισχυρή επιρροή.

Η ίδια έκθεση επίσης αναφέρει ότι η Ελλάδα βρίσκεται στο πρώτο γκρουπ δυναμικότητας από τα πέντε ανάμεσα σε άλλες 174 χώρες του κόσμου ως προς τους

στατιστικούς δείκτες απόδοσης, οι οποίοι αποτιμούν τα εθνικά στατιστικά συστήματα και στρατηγικούς σχεδιασμούς, βασιζόμενοι σε πέντε βασικούς πυλώνες [12]:

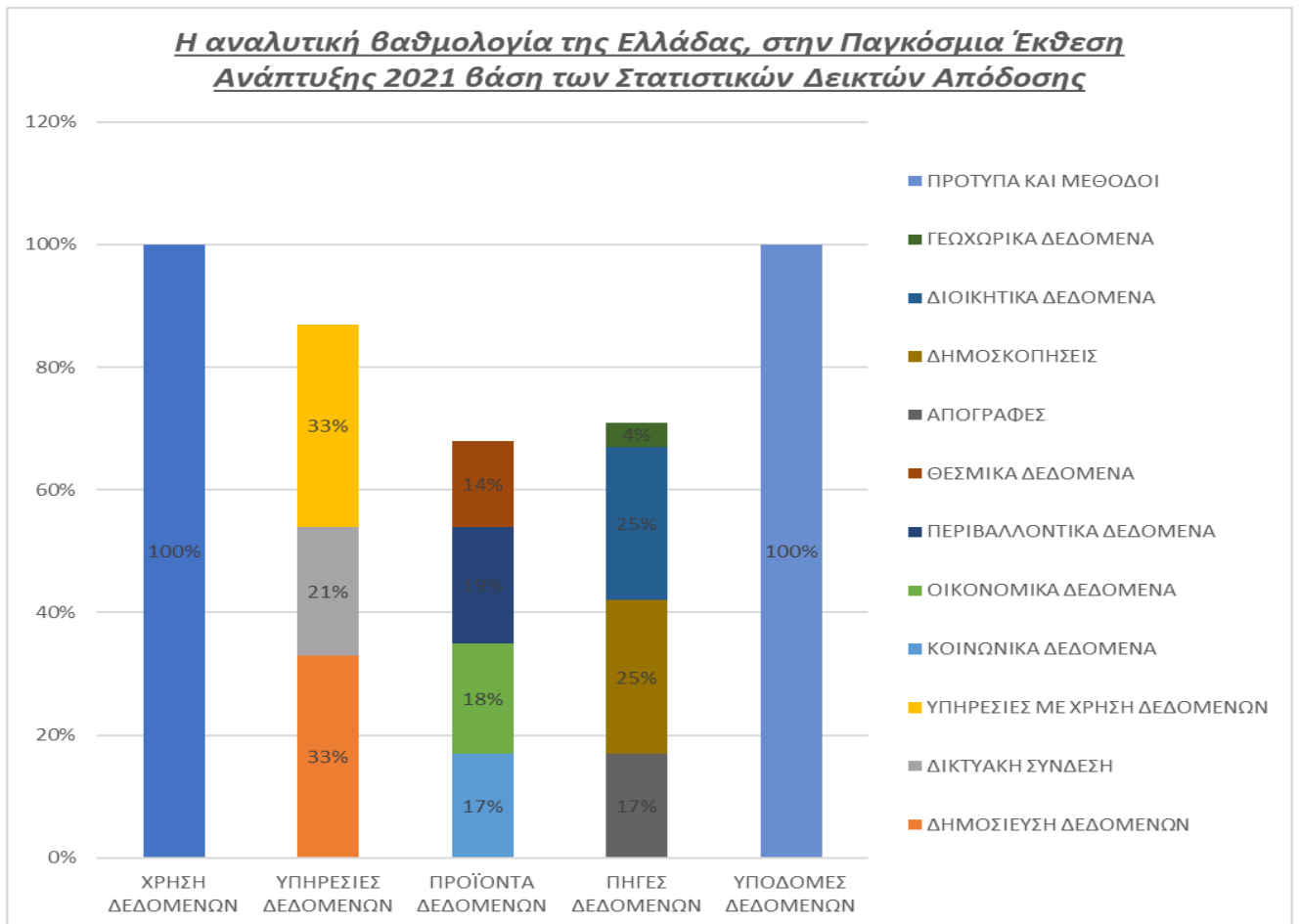
- Η χρήση των δεδομένων, δηλαδή να μπορούν να χρησιμοποιηθούν ευρέως τα δεδομένα.
- Οι υπηρεσίες δεδομένων, αξιόπιστες σχέσεις χρηστών και παραγωγών δεδομένων.
- Τα προϊόντα δεδομένων τα οποία αποτελούν κομμάτι της επιχειρηματικής ευφυίας και πάνω σε αυτά μπορούν να στηριχθούν οι επιχειρήσεις και η κοινωνία ώστε να αναπτύξουν την δραστηριότητά τους.
- Οι πηγές των δεδομένων οι οποίες χρειάζεται να είναι ποικίλες και να προέρχονται από όλους τους φορείς της κοινωνίας και συλλογή από αυτές να πραγματοποιείται όσο το δυνατόν με καινοτόμους τρόπους
- Και τέλος η υποδομή δεδομένων, η οποία προϋποθέτει μια εθνική στρατηγική με στέρεες βάσεις οι οποίες θα αποτελούνται από ένα κατάλληλο νομοθετικό πλαίσιο, ενισχυμένους οικονομικούς και τεχνολογικούς πόρους, εκπαιδευμένους πολίτες και την ύπαρξη ενισχυμένης διαλειτουργικότητας ανάμεσα στους φορείς της διοίκησης, των επιχειρήσεων και πολιτών μέσω των κατάλληλων προτύπων επικοινωνίας.



The boundaries, colors, denominations and any other information shown on this map do not imply, on the part of the World Bank Group, any judgment on the legal status of any territory, or any endorsement or acceptance of such boundaries.

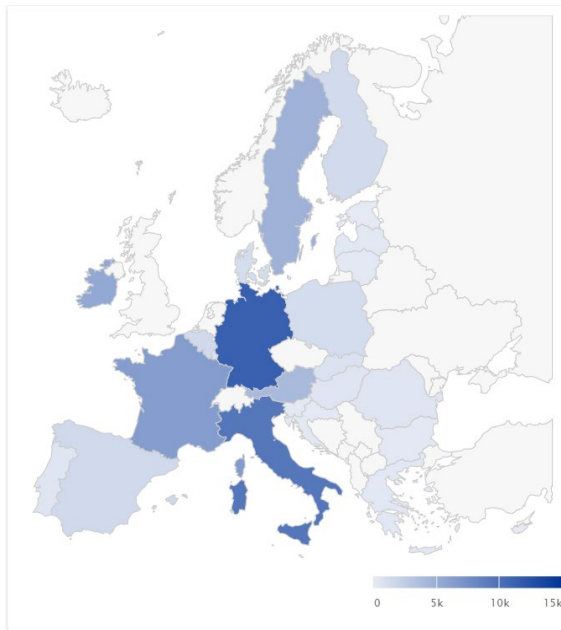
Εικόνα 5.. Ο χάρτης των 174 χωρών που εξετάστηκαν στην Παγκόσμια Έκθεση Ανάπτυξης 2021 βάση των Στατιστικών Δεικτών Απόδοσης, οι οποίες χρωματίστηκαν βάση της βαθμολογίας τους. [12] Πηγή: Παγκόσμια Έκθεση Ανάπτυξης 2021, Παγκόσμια Τράπεζα

Στην συνέχεια αυτής της έκθεσης η Ελλάδα έχει πετύχει την υψηλότερη βαθμολογία στα Βαλκάνια και βρίσκεται στη μέση περίπου των άλλων Ευρωπαϊκών χωρών . Εξετάζοντας αναλυτικά την βαθμολογίας της στους πέντε άξονες η χώρα μας υστερεί στις πηγές και στα προϊόντα δεδομένων, τα οποία χρήζουν περαιτέρω βελτίωσης, η οποία βάση του ρυθμού της Ελλάδας τα τελευταία χρόνια φαντάζει ρεαλιστική καθώς το έτος 2016 είχε γενική απόδοση 82,55 και αυξάνοντας κάθε χρόνο κατά μια περίπου μονάδα έφτασε το έτος 2019 στις 85,36 μονάδες.

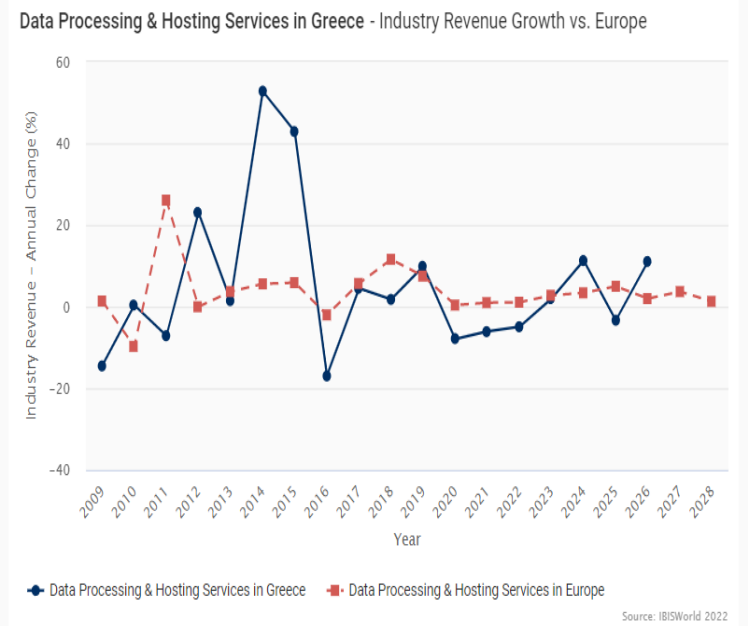


Εικόνα 6. Η αναλυτική βαθμολογία της Ελλάδας, στην Παγκόσμια Έκθεση Ανάπτυξης 2021 βάση των Στατιστικών Δεικτών Απόδοσης. [12]

Ωστόσο, στον τομέα των επιχειρήσεων, μεμονωμένα η βιομηχανία επεξεργασία και φιλοξενίας δεδομένων στην Ελλάδα είναι ακόμη περιορισμένη και δεν έχει μεγάλο αντίκτυπο στην οικονομία καθώς κατέχει μόλις την 85η θέση ανάμεσα στους 107 διακριτούς βιομηχανικούς κλάδους που όρισε η IbisWorld, (μία από τις μεγαλύτερες συμβουλευτικές εταιρείες ανάλυσης δεδομένων με παγκόσμια παρουσία) σε φετινή έκθεσή της, και επιπρόσθετα η Ελλάδα είναι 18η ανάμεσα σε είκοσι πέντε μέλη-κράτη (δεν αναφέρονται στοιχεία για την Ολλανδία και την Τσεχία στην έκθεση αυτή). [13]



Εικόνα 8. Αποτυπώνεται μια οπτικοποίηση σε μορφή χάρτη, της αποτίμησης σε εκατομμύρια ευρώ, των βιομηχανιών επεξεργασίας και φιλοξενίας δεδομένων ανά χώρα της Ε.Ε. πλην Τσεχίας και Ολλανδίας. [13]



Εικόνα 7. το διάγραμμα του ρυθμού ανάπτυξης του κλάδου ποσοστιαία σε εθνικό επίπεδο και σε σύγκριση με εκείνο της Ε.Ε.[13]

Στην ίδια έκθεση, ο κλάδος Υπηρεσιών Επεξεργασίας και Φιλοξενίας Δεδομένων στην Ελλάδα ή Βιομηχανία Στατιστικών Υπηρεσιών αποτιμάται στα διακόσια τριάντα επτά εκατομμύρια ευρώ για το έτος 2022 με περίπου χίλιες τετρακόσιες επιχειρήσεις, όταν το Ακαθάριστο Εγχώριο Προϊόν (Α.Ε.Π.) διαμορφώθηκε στα 181 δισεκατομμύρια την προηγούμενη χρονιά σύμφωνα με την Ελληνική Στατιστική Αρχή (ΕΛΣΤΑΤ) [14] δηλαδή θα αποτελούσε μόλις το 0.13% της οικονομίας αν παρέμενε το Α.Ε.Π. σταθερό και την νέα χρονιά. Επιπλέον, από το διάγραμμα της εικόνα 8, η προοπτική για ανάπτυξη του κλάδου προβλέπεται να είναι στάσιμη με αυξομειώσεις σε εθνικό επίπεδο μέχρι το 2026 ενώ στην αντίστοιχη περίοδο σε ευρωπαϊκό επίπεδο διατηρείται ένα σταθερό θετικό πρόσημο [13].

Η ανάπτυξη του κλάδου στην Ελλάδα μπορεί να κλίσει προς τα θετικά αν ενισχυθεί η εξωτερική ανάθεση (outsourcing) στις στατιστικές εργασίες από πλευράς ελληνικών εταιρειών όπως έχει διαδοθεί και σε άλλου είδους εργασίες, όπως λογιστικές, τεχνολογικές, νομικές. Και αυτό καθώς ο κλάδος μεμονωμένα δεν έχει μεγάλη ανάπτυξη, αφού πολλές επιχειρήσεις ενισχύονται συνεχώς με επαγγελματίες του κλάδου αυτού συμπληρωματικά ώστε να βελτιστοποιήσουν τις διαδικασίες τους.

1.4. Ευκαιρίες, προκλήσεις και προτάσεις

Η άνοδος της επιστήμης των δεδομένων σε συνδυασμό και με την επιχειρηματική ευφυΐα, έχει ως αποτέλεσμα να επαναπροσδιοριστεί ο τρόπος που αντιλαμβανόμαστε το “επιχειρείν”, τις απαιτήσεις που έχουν οι καταναλωτές από τους εμπόρους, οι πολίτες από το κράτος και γενικότερα οι δεσμοί και οι σχέσεις στην κοινωνία. Ωστόσο, βρισκόμαστε ακόμη στην αυγή της επόμενης τεχνολογικής εποχής κατά την οποία θα πραγματοποιηθεί η μετάβαση από τα “έξυπνα” συστήματα σε αυτά της Τεχνητής Νοημοσύνης [3].

1.4.1. Οικονομία-Επιχειρηματικότητα

Ήδη έχει εισέλθει σε μεγάλο βαθμό στην οικονομία η επιστήμη δεδομένων και γίνεται χρήση της με διάφορους τρόπους. Ο πιο γνωστός είναι η υποστήριξη αποφάσεων σε έναν οργανισμό με σκοπό την αποδοτικότερη λειτουργία του και την σχεδίαση στρατηγικών. Συνήθως, αυτό επιτυγχάνεται με την συρροή δεδομένων σε αυτόν είτε αυτά προέρχονται μέσω της τυπικής λειτουργίας της και το εσωτερικό της περιβάλλον, δηλαδή τιμολόγια, πωλήσεις, αγορές, υπάλληλοι κ.τ.λ. είτε καταβάλλει επιπλέον προσπάθεια και αναζητάει εξωτερικά πληροφόρηση η οποία μπορεί να “εφάπτεται” με την ίδια την εταιρεία όπως είναι η ικανοποίηση των πελατών της, οι ανταγωνιστές της, οι καιρικές ή κοινωνικές συνθήκες που επικρατούν στον τόπο δραστηριοτήτων της ή μπορεί να “μακριά” από αυτήν όπως στοιχεία διαφορετικών κλάδων, ξένων χωρών και τόπων.

Σε όλο αυτό το μείγμα δεδομένων, επιλέγονται εκείνα τα δεδομένα τα οποία θα υποστηρίξουν τις αποφάσεις του οργανισμού και αφού επεξεργαστούν, θα παραδοθούν με

την κατάλληλη οπτική απεικόνιση στην διοίκηση ή στο τμήμα που την χρειάζεται και θα αναλάβουν τις ανάλογες ενέργειες οι διευθύνοντες με βάση και της δικής τους εμπειρίας με υποβοήθηση, ίσως, διάφορων στατιστικών αλγορίθμων οι οποίοι προσπαθούν να εξάγουν ποσοστά πιθανοτήτων σε διάφορα σενάρια. Άρα, ακόμη ο ανθρώπινος παράγοντας παίζει σημαντικό ρόλο καθώς είναι αυτός που θα πάρει την τελική απόφαση με ότι συνεπάγεται με αυτό καθώς το κάθε άτομο ανεξάρτητα από την εργασιακή του εμπειρία και οξυδέρκεια μπορεί να υποπέσει σε παραπτώματα εξαιτίας προκαταλήψεων ή λόγω λανθασμένης αξιολόγησης της επίδρασης διάφορων μεταβλητών.

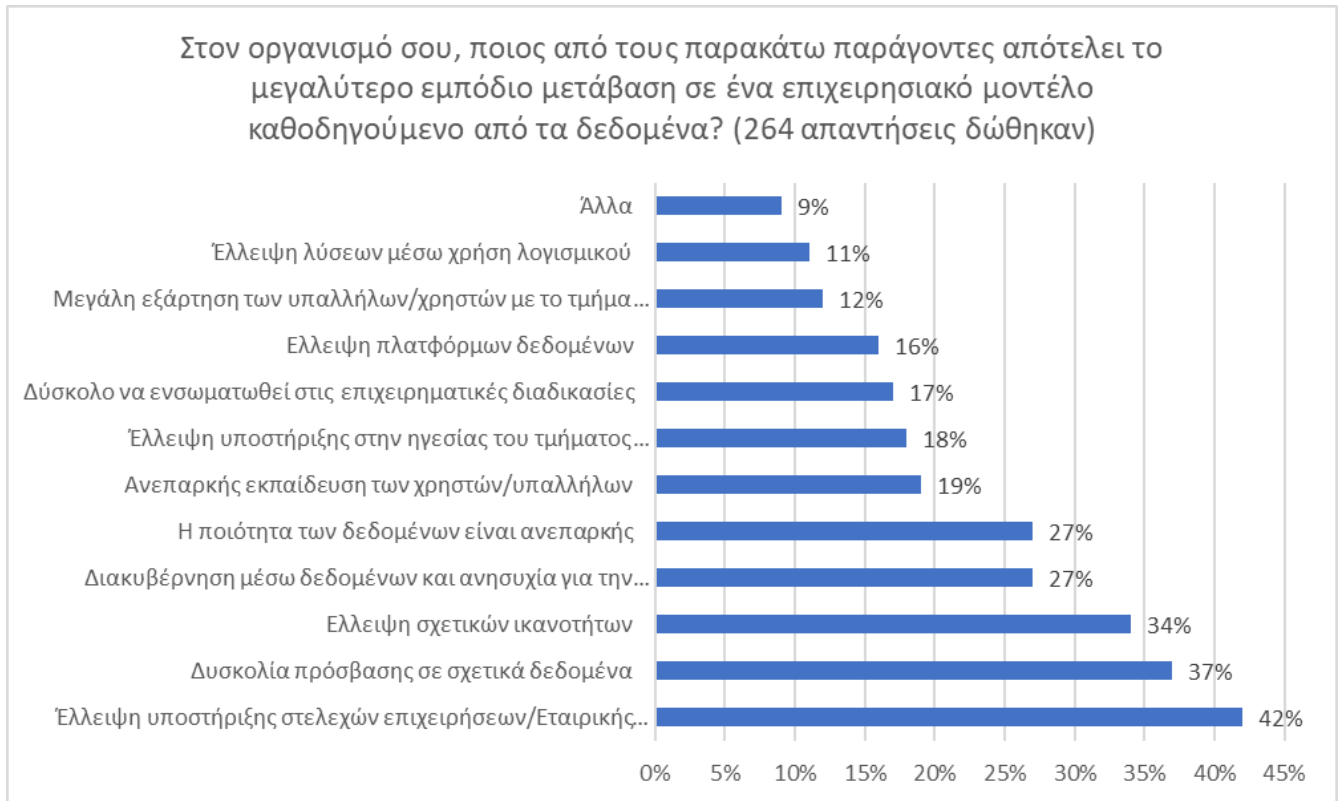
Το επόμενο βήμα αυτής της κατάστασης είναι οι Επιχειρήσεις που λειτουργούν εξολοκλήρου βασιζόμενες στα γεγονότα (Data-Driven Businesses). Φυσικά όλες οι επιχειρήσεις θα ήθελαν οι αποφάσεις τους να βασίζονται σε μία αυστηρά υπολογιστική ανάλυση ωστόσο προϋποθέτει την τροποποίηση της ίδιας της διαδικασίας λήψης αποφάσεων η οποία πρέπει να είναι αποκεντρωμένη και προσανατολισμένη μονάχα στο στόχο που προσπαθεί να επιλύσει κάθε φορά[15]. Οι στόχοι διαφέρουν ως προς την σημαντικότητά τους και την διάρκεια τους, δηλαδή αν είναι ημερήσιες και εβδομαδιαίες εργασίες ή σχεδίαση στρατηγικών.

Η πρώτη περίπτωση στόχων είναι οι πιο εύκολα επιτεύξιμοι και ήδη σε πολλούς οργανισμούς αντιμετωπίζονται με αυτό τον τρόπο, για παράδειγμα σε ένα χρηματοπιστωτικό ίδρυμα αν κάποιος πελάτης πληροί τα κριτήρια για ένα δάνειο ή ασφάλιση[4], αλλά η δεύτερη περίπτωση είναι πιο περίπλοκη και θα απαιτούσε ροή πληροφοριών από όλα τα τμήματα του οργανισμού και ύστερα μέσω αλγορίθμων θα παρήγαγε κατευθύνσεις, προειδοποιήσεις ή θα επισήμανε ευκαιρίες, αλλάζοντας ακόμη και την ίδια την στρατηγική της εταιρείας, όπως η αποεπένδυση από ρυπογόνα προϊόντα και παραγωγές λόγω της ολοένα και αυξανόμενης πράσινης πολιτικής που εφαρμόζεται από τα όλα τα κράτη ή η δημιουργία νέων δικτύων της εφοδιαστικής αλυσίδας και ανεύρεσης νέων συνεργατών λόγω μίας έκτακτης γεωπολιτική κρίσης η οποία βάση των δεδομένων θα δημιουργούσε ελλείψεις, των οποίων τα διαφυγόντα κέρδη θα αντισταθμίζονταν από το τις νέες επενδύσεις. Αν και περίπλοκη διαδικασία, θα μπορούσε να εντοπίσει ευκολότερα μοτίβα πίσω από μία τόση ογκώδη και ποικιλόμορφη βάση δεδομένων, προωθώντας σημαντικά την συνδυαστική στρατηγική και την διαλειτουργικότητα εντός του οργανισμού. Αυτό έρχεται σε σύγκρουση με την

παραδοσιακή κεντροποιημένη δομή των επιχειρήσεων, όπου οι αποφάσεις τελικά θα παρθούν μέσω της διοίκησης και χρησιμοποιούν τα στατιστικά εργαλεία ως υποστηρικτικά και όχι ως καταλυτικά στις αποφάσεις τους, ενώ σε μία μη-κεντροποιημένη επιχειρησιακή δομή, η ίδια η έννοια της διοίκησης θα αναθεωρούνταν και περιορίζονταν σε ένα ρόλο επίβλεψης της ομαλής λειτουργίας της επιχείρησης καθοδηγούμενη από τα δεδομένα και ότι τα δεδομένα αυτά τηρούν τις προϋποθέσεις που έχουν θέσει εξαρχής.

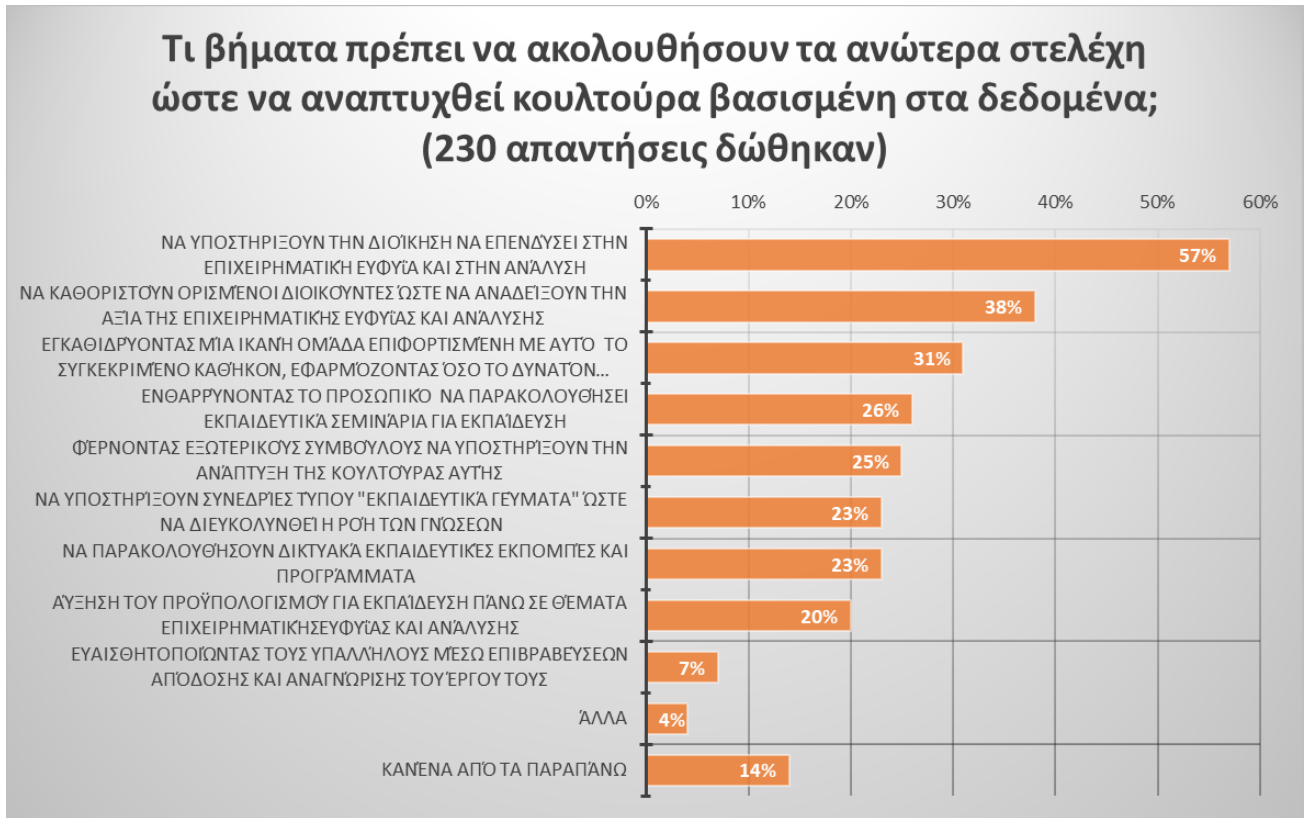
Από τα ανωτέρω θα μπορούσε να ειπωθεί ότι η επιτυχία της επιχείρησης θα βασίζεται κυρίως στην ποιότητα των δεδομένων της διαδικασίας ανατροφοδότησης του μηχανισμού λήψης αποφάσεων και όχι τόσο στον ανθρώπινο παράγοντα και αν έπρεπε να υπάρχει η έννοια του κέντρου αυτά θα ήταν τα δεδομένα. Οι λόγοι που οι οργανισμοί είναι διστακτικοί στην μετάβαση αυτή, είναι κυρίως επειδή απαιτείται ριζική αναθεώρηση των συστατικών στοιχείων της επιχείρησης, όπως είναι η κουλτούρα της όπου χρειάζεται να εγκαθιδρυθεί στους εργαζόμενους η νοοτροπία της αυτοματοποιημένης εκτέλεσης ενεργειών βασισμένες στα δεδομένα και στα εργαλεία τα οποία τους έχει προμηθεύσει η εταιρεία και ταυτόχρονα να αντιλαμβάνομαι τη σημασία των δεδομένων. Από την πλευρά της διοίκησης, να εμπιστευτεί τα χαμηλότερα στην ιεραρχία στελέχη αλλά και την ίδια την διαδικασία. Προφανώς απαιτείται εξάσκηση και εκπαίδευση σε αυτή την διαδικασία αλλά και στα εργαλεία επιχειρηματικής ευφυΐας. Άλλος ένας παράγοντας που καθυστερεί την εξέλιξη των εταιρειών είναι έλλειψη των κατάλληλων επαγγελματιών οι οποίοι θα ηγηθούν στο έργο της μετάβασης, θα βοηθήσουν τους άλλους υπαλλήλους στη χρήση των νέων ψηφιακών λύσεων και την “ανταλλαγή” πληροφοριών στα τμήματα του οργανισμού.

Σε έρευνα με μορφή ερωτοαπαντήσεων σε επαγγελματίες που εργάζονται, χωρίς κανέναν περιορισμό, της επιχειρηματικής ευφυΐας και της διαχείρισης των δεδομένων το έτος 2017 , που πραγματοποίησε μία διεθνής εταιρεία εξειδικευμένη σε παροχή εκπαιδεύσεων και υποστήριξης της επιστήμης των δεδομένων, η TDWI (TRANSFORMING DATA WITH INTELLIGENCE) [15], το δείγμα “έδειξε” ως κυριότερα εμπόδια προς αυτήν την κατεύθυνση θέματα που αφορούν στρατηγική της εταιρείας, έλλειψη αντίστοιχων δεξιοτήτων, κακή διαχείριση δεδομένων καθώς και ανησυχίες ασφαλείας.



Εικόνα 9. Έρευνα με δείγμα με επαγγελματίες της Επιχειρηματικής Ευφυΐας και Επιστήμης Δεδομένων, για το τι εμποδίζει τις εταιρείες εξελιχθούν σε ένα επιχειρησιακό μοντέλο βασισμένο περισσότερο στα δεδομένα [15]

Τα παραπάνω εμπόδια μπορούν να υπερκεραστούν με κατάλληλες μεθόδους και πρακτικές ανάλογα το είδος και μέγεθος της εταιρείας καθώς και των προϊόντων της. Σίγουρα την πρωτοβουλία πρέπει να πάρει το διοικητικό συμβούλιο και να δημιουργήσει μία ομάδα, η οποία θα αποτελείται από ένα μείγμα ανθρώπων από όλη την ιεραρχία της καθώς και με εξωτερικούς συνεργάτες, με κοινό χαρακτηριστικό όλων την γνώση της επιστήμης των δεδομένων και εξοικείωσης με την επιχειρηματική ευφυΐα, που θα έχει ως αυτοσκοπό της ομαλότερης και αποτελεσματικότερης μετάβαση της εταιρείας στην εποχή των Data-Driven Businesses. Στην ίδια έρευνα [15], στην ερώτηση ποια βήματα πρέπει να ακολουθήσει η διοίκηση προς αυτήν την κατεύθυνση, αναφέρεται σαν προαπαιτούμενη η επένδυση σε έξυπνες λύσεις, μετά η διοίκηση να εμπιστεύεται την επιχειρηματική ευφυΐα και επιστήμη δεδομένων καθώς και σύσταση μία ομάδας εκσυγχρονισμού και επίβλεψης της μεταβολή της γενικότερης αντίληψης των πραγμάτων.



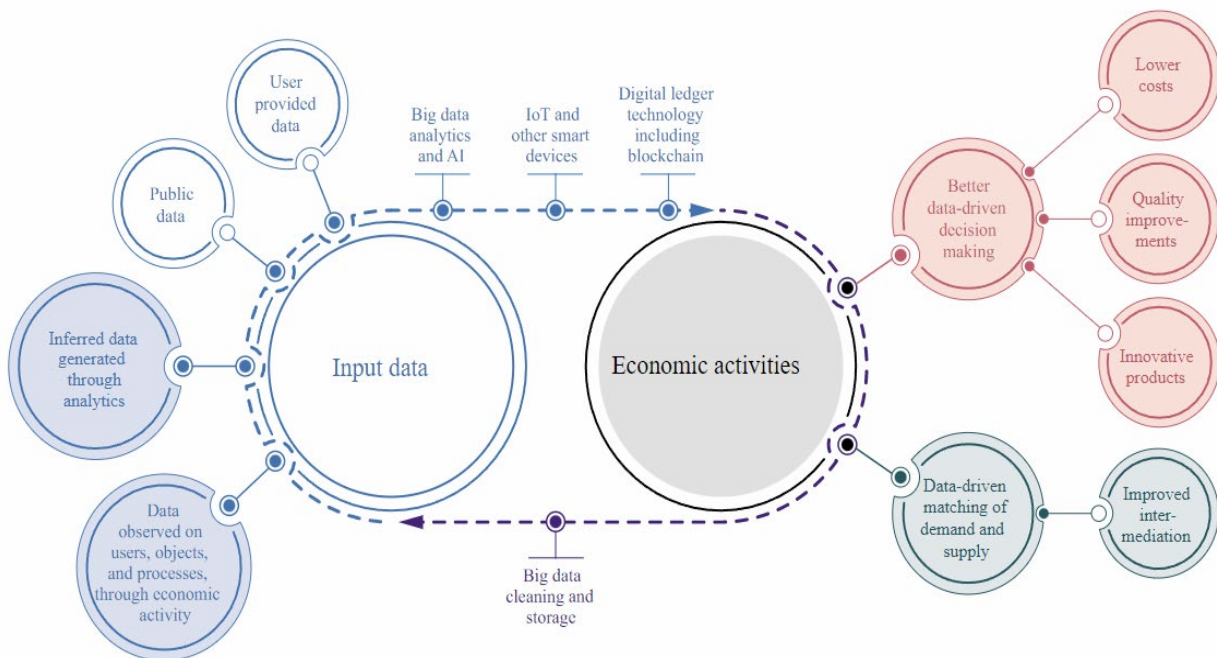
Εικόνα 10. Έρευνα με δείγμα επαγγελματιών της Επιχειρηματικής Ευφύιας και Επιστήμης Δεδομένων, για τα βήματα που πρέπει να ακολουθήσει η εταιρεία ώστε εξελιχθεί σε ένα επιχειρησιακό μοντέλο βασισμένο περισσότερο στα δεδομένα.[15]

Στην παραπάνω εικόνα, επίσης, αποτυπώνεται ένα ολοκληρωμένο μοντέλο Data-Driven επιχείρησης και στην δεξιά πλευρά αποτυπώνονται τα πλεονεκτήματα που προκύπτουν μέσω ενός τέτοιου μοντέλου επιχειρηματικότητας κατά το οποίο λαμβάνονται καλύτερες αποφάσεις και διαχείριση της προσφοράς και ζήτησης. Τα οφέλη αυτά είναι:

- Μειωμένα κόστη σε όλη την γραμμή παραγωγής και της εφοδιαστικής αλυσίδας μέσω μίας αποδοτικότερης χρήση των διαθέσιμων πόρων. Αναφορικά μερικά τέτοια κόστη μπορεί να αντιστοιχούν στην αποθήκευση, συντήρηση και μεταφορά.
- Η αύξηση της ποιότητας των παρεχόμενων προϊόντων και υπηρεσιών κατανοώντας καλύτερα τις ανάγκες του καταναλωτικού κοινού, επιδιώκοντας συνεχώς της αύξηση της ικανοποίηση του κοινού. Αυτό μπορεί να επιτευχθεί μέσω στοχευμένων διαφημιστικών εκστρατειών και πληροφόρησης, αναδεικνύοντας τα δυνατά σημεία και περιορίζοντας τα αρνητικά. Για παράδειγμα, μία αλυσίδα γρήγορου φαγητού αντιλαμβάνομενη ότι απευθύνεται σε μεγάλο βαθμό σε

οικογένειες, θα ενίσχυε τα καταστήματά της με παιδικούς καταλόγους και παιδότοπους, αυξάνοντας έτσι την προσφερόμενη αξία της επίσκεψης στα καταστήματα της προς τις οικογένειες όπου με την σειρά τους θα προτιμούσαν περισσότερο τα εστιατόρια αυτά.

- Καινοτόμα προϊόντα τα οποία είναι εξατομικευμένα για να ικανοποιούν αποτελεσματικότερα τις σύγχρονες ανάγκες των πελατών όσο γίνεται πιο καιρία από τα υπάρχοντα προϊόντα. Ένα σύγχρονο παράδειγμα στο εμπόριο τροφίμων είναι η ανάγκη των αγοραστών για τρόφιμα υψηλότερης διατροφικής αξίας για μυϊκή ενδυνάμωση, την οποία προσπαθούν να ικανοποιήσουν με πρωτεϊνούχα προϊόντα με χαμηλά λιπαρά, τα οποία έχουν παρόμοια γεύση με τα συμβατικά. Τέτοια τρόφιμα (High protein), δεν υπήρχαν σε απλά καταστήματα λιανικής και περιορίζονταν σε μέρη αθλητισμού και με πολύ υψηλότερες τιμές.
- Βελτιωμένη διαμεσολάβηση μεταξύ παραγωγού και εμπόρου με σκοπό την αποδοτικότερη αναπλήρωση αποθεμάτων από την χονδρική αγορά μέχρι το λιανικό εμπόριο. Έτσι, θα αποφεύγονται τυχόν ζημίες από φύρα ευαίσθητων προϊόντων ή διαφυγόντα κέρδη λόγω ελλείψεων και καθυστερήσεων στην εφοδιαστική αλυσίδα.



Εικόνα 11. Data-Driven Business Mode. [12] Source: World Development Report 2021, World Bank

1.4.2. Μεταφορές

Ίσως ένας από τους πιο πολλά υποσχόμενους τομείς κατά τον οποίο η επίδραση της επιστήμης των δεδομένων μπορεί να φέρει πραγματικά τεχνολογική επανάσταση και να αλλάξει την καθημερινότητα της ανθρωπότητας. Έχει γίνει πόλος συζητήσεων ως προς την οδική εμπειρία του μέλλοντος και τον καταλυτικό ρόλο που μπορεί να διαδραματίσει η Τεχνητή Νοημοσύνη η οποία θα μπορεί να μαθαίνει, να αυτο-βελτιώνεται και να αποφασίζει με τον δικό της ηθικό κώδικα χωρίς την ανθρώπινη παρέμβαση μέσω της μηχανικής μάθησης μίας μεθόδου που μπορεί να διαχειριστεί τα Big Data και βασίζεται στην στατιστική, στην βελτιστοποίηση και την επιστήμη των υπολογιστών, ανήκοντας στην επιστήμη δεδομένων.

Σήμερα, η βελτίωση στις μεταφορές εμφανίζεται έντονα στον πιο αποδοτικό προγραμματισμό δρομολογίων, χρησιμοποιώντας στατιστικά μοντέλα τα οποία στοχεύουν στη μεγιστοποίηση της αξίας ή την ελαχιστοποίηση στα κόστη, μία συνεχής διαδικασία κατά την οποία γίνεται προσπάθεια μείωσης των αποκλίσεων με την διαρκή ανανέωση των δεδομένων[16]. Επίσης, επιπρόσθετα εφαρμογές οι οποίες έχουν καταστεί απαραίτητα εργαλεία της καθημερινότητά μας βασίζονται στη διαχείριση δεδομένων και στη μηχανική μάθηση. Μία από αυτές είναι η Google Maps. Η εφαρμογή χαρτογράφησης της εταιρείας Google, ουσιαστικά είναι ένας αλγόριθμος που μαθαίνει συνεχώς από τα δεδομένα που συλλέγει σε πραγματικό χρόνο από τους ίδιους τους χρήστες της εφαρμογής, στοιχεία ως προς την κίνηση και την οδική τους συμπεριφορά και σε συνδυασμό με παρελθόντα μοτίβα και πληροφόρηση για κυκλοφορικές ρυθμίσεις από τις τοπικές αρχές, προτείνει την βέλτιστη διαδρομή στον χρήστη. Επιπλέον, έχει επεκταθεί ως βάση δεδομένων ως προς τις τοποθεσίες, με αξιολογήσεις χρηστών και έτσι ο καθένας εύκολα μπορεί να πλοηγηθεί και αναζητήσει στην εφαρμογή πληροφορίες για μία συγκεκριμένη τοποθεσία .

Ωστόσο, ο “ελέφαντας στο δωμάτιο” δεν είναι άλλος από την νέα γενιά αυτοκινήτων, τα οποία υπόσχονται αυτονομία σε τέτοιο βαθμό όπου δεν θα χρειάζονται πια οδηγοί σε ένα σύγχρονο κυκλοφοριακό μοντέλο. Τα αυτόνομα αυτοκίνητα είναι ένας συνδυασμός τεχνολογιών όπου βασικό ρόλο παίζει Τεχνητή Νοημοσύνη, της οποίας συστατικό στοιχείο είναι η επιστήμη των δεδομένων μέσω τεχνικών μηχανικής μάθησης όπως για παράδειγμα η βαθιά μάθηση (Deep Learning). Χονδρικά, ένα αυτόνομο όχημα

θα είναι απαραίτητο να διαθέτει αισθητήρες (sensors), δείκτες και κάμερες, τα οποία θα αλληλοεπιδρούν με το εξωτερικό περιβάλλον αλλά και με εξωτερικούς αισθητήρες και δείκτες που θα υπάρχουν στους δρόμους και στις τοπικές περιοχές, παράγοντας με αυτό τον τρόπο δεδομένα τα οποία θα εισάγονται στο λογισμικό τους και με την κατάλληλη επεξεργασία μέσω αλγορίθμων, θα μπορεί να εντοπίζει άλλα αυτοκίνητα, πεζούς καθώς και αντικείμενα, και έτσι θα μπορεί να πλοηγείτε αυτόνομα. Επιπρόσθετα, η μηχανική των οχημάτων θα έχει ως άξονα την λήψη αποφάσεων με γνώμονα τα δεδομένα, για παράδειγμα ποια διαδρομή να ακολουθήσει ή αν εντοπίσει αντικείμενο πως θα αντιδράσει. Ως προς τον τρόπο λήψης αποφάσεων και για το τι θα πρέπει να θεωρείται σωστό ή όχι, έχουν δημιουργηθεί ορισμένα ηθικά και φιλοσοφικά διλήμματα κυρίως στην περίπτωση ατυχήματος, δηλαδή ένα αυτοοδηγούμενο όχημα αν θα πρέπει να παίρνει αποφάσεις βάση του κοινωνικού συμφέροντος ή του ατομικού, δηλαδή του οδηγού[16].

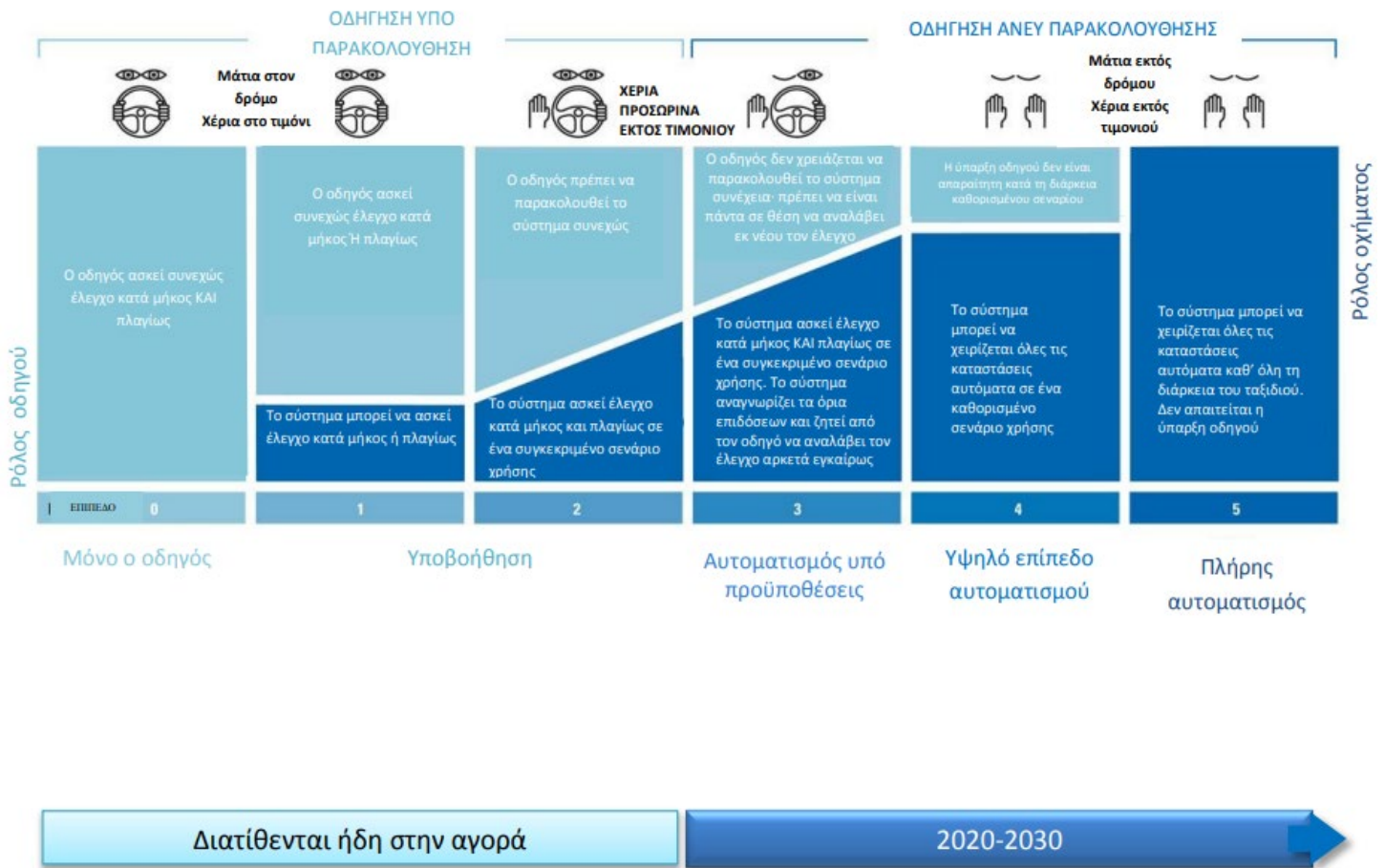
Για γίνει αντιληπτή η εξέλιξη της αυτονομίας στα οχήματα, η SAE International (Society of Automotive Engineers) [17], ένας διεθνής οργανισμός ανάπτυξης προτύπων για επαγγελματίες μηχανικούς σε διάφορους βιομηχανικούς κλάδους, πρότεινε έξι επίπεδα αυτονομίας στα οχήματα, την οποία κατηγοριοποίηση υιοθέτησε και ο οργανισμός NHTSA (National Highway Traffic Safety Administration) που ανήκει στο υπουργείο Μεταφορών των ΗΠΑ. Αναλυτικά τα επίπεδα αυτοματοποίησης των οχημάτων αποτυπώνεται παρακάτω, στα εξής αριθμημένα επίπεδα[18]:

1. Κανένας αυτοματισμός. Δεν υπάρχει καμία υποστήριξη, όπως ενδείξεις του
2. συστήματος ελέγχου πρόσφυσης, και η οδήγηση βασίζεται αποκλειστικά στις δυνατότητες του οδηγού. Αυτά τα αυτοκίνητα είναι παλαιού τύπου και δεν παράγονται πλέον σε μαζική παραγωγή.
3. Υποστήριξη του οδηγού. Υποστηρικτές λειτουργίες καθώς συγκεκριμένες παρεμβάσεις στο σύστημα οδήγησης και σε ορισμένες περιπτώσεις μόνο. Ο οδηγός και σε αυτό το επίπεδο έχει την απόλυτη ευθύνη του αυτοκινήτου. Περισσότερα αυτοκίνητα που κυκλοφορούν στους δρόμους ανήκουν σε αυτό το επίπεδο.
4. Μερικός Αυτοματισμός. Το όχημα μπορεί να εκτελέσει ορισμένες βασικές ενέργειες αυτόνομα, όπως να επιταχύνει ή να φρενάρει, μελετώντας το περιβάλλον του. Ελαφρύνεται το οδηγικό έργο του ατόμου ωστόσο ακόμη ο άνθρωπος παράγοντας έχει τον έλεγχο, κρίνει και παίρνει τις τελικές αποφάσεις. Παράγονται

- αρκετά ήδη αρκετά ωστόσο δεν είναι οικονομικά προσβάσιμα για όλες τις οικονομικές τάξεις, καθώς το κόστος του είναι αρκετά υψηλότερο από εκείνα της προηγούμενης κατηγορίας.
5. Αυτοματισμός Υπό Όρους. Σε συγκεκριμένες συνθήκες, το αυτοκίνητο έχει την δυνατότητα να αναλάβει τις σχεδόν όλες τις πτυχές της οδήγησης, ακόμη και της ανάλυσης του περιβάλλοντος του. Ο οδηγός μπορεί να αναλάβει τον έλεγχο όποτε θέλει και επιπρόσθετα το σύστημα πλοήγησης ειδοποιεί τον άνθρωπο να επέμβει όταν αντιμετωπίζει ένα σενάριο το δεν γνωρίζει. Η μετάβαση σε αυτό το επίπεδο είναι κομβικής σημασίας καθώς εδώ ο έλεγχος του αυτοκινήτου περνάει από τον άνθρωπο προς το σύστημα αυτόνομης πλοήγησης. Ο ανταγωνισμός μεταξύ των αυτοκινητοβιομηχανιών είναι τεράστιος για την επίτευξη μαζικής παραγωγής οχημάτων αυτής της κατηγορίας και προβλέπεται μέχρι τέλος αυτού του έτους να έχουν βγει προς πώληση τα πρώτα αυτοκίνητα.
 6. Υψηλό Επίπεδο Αυτοματισμού. Το αυτοκίνητο μπορεί να λειτουργεί και χωρίς ανθρώπινη συνδρομή ωστόσο μόνο υπό επιλεγμένες συνθήκες που διαμορφώνονται από παράγοντες που έχουν να κάνουν με το εξωτερικό περιβάλλον όπως ο τύπος του δρόμου. Τέτοια αυτοκίνητα έχουν δημιουργηθεί σε δοκιμαστικό επίπεδο από την εταιρεία Google, με όνομα μάρκας Waymo[19], ενώ πιθανολογείται στα αυτοκίνητα που θα κυκλοφορήσουν στο επόμενο διάστημα και θα ανήκουν στο προηγούμενο επίπεδο, να έχουν στο λογισμικό τους μία έκδοση “BETA” για ομαλότερη και γρηγορότερη μετάβαση, η οποία μετάβαση προβλέπεται να λάβει χώρα ευρέως έως το έτος 2025[18]. Αξίζει να σημειωθεί, ωστόσο, ότι ακόμη και να οι αυτοκινητοβιομηχανίες να ξεκινήσουν την παραγωγή τέτοιου τύπου αυτοκινήτων, θα έχουν λογισμικό που θα το εναλλάσσουν ανάλογα με την χώρα που βρίσκονται καθώς δεν είναι όλα τα εθνικά οδικά δίκτυα να διαχειριστούν αυτό το είδος τεχνολογίας. Μέχρι στιγμής, μόνο πολύ αναπτυγμένες χώρες όπως για παράδειγμα οι ΗΠΑ, η Κίνα και η Γερμανία, έχουν επιτρέψει την οδήγηση με χρήση αυτής της τεχνολογία και αυτό σε ορισμένες περιοχές και συνθήκες. Αυτό αναμένεται να αλλάξει, καθώς πολλές χώρες επενδύουν κεφάλαια στον εκσυγχρονισμό των οδικών δικτύων τους.

7. Απόλυτος Αυτοματισμός. Ο ανθρώπινος παράγοντας δεν είναι πια απαραίτητος καθώς το αυτοκίνητο μπορεί να εκτελέσει όλες τις ενέργειες ανεξαρτήτου συνθηκών. Όπως αναφέρθηκε προηγουμένως είναι διαφορετικό γεγονός η δυνατότητα δημιουργίας εξελιγμένων αυτοκινήτων από εκείνης της διαχείρισης αυτών στα εθνικά δίκτυα έτσι οι προβλέψεις για αυτοδηγούμενα οχήματα στις αστικές περιοχές δεν είναι ξεκάθαρες και βασίζονται και στην πολιτική των επί αυτού του θέματος. Πιθανή χρονική περίοδος θα είναι η επόμενη πενταετία της δεκαετίας μας, δηλαδή 2025-2030.[18]

Σχήμα: Διάφορα επίπεδα αυτοματισμού (πηγή: Society of Automotive Engineers-SAE¹¹)



Εικόνα 12. Επίπεδα αυτοματισμού οχημάτων σύμφωνα με την SAE International και η διάκριση των ρόλων μεταξύ οδηγού και αυτοκινήτου στα επίπεδα αυτά. [17]

Γίνεται επομένως αντιληπτό ότι είναι εξίσου σημαντική η εξέλιξη των μοντέλων οδικής κυκλοφορίας σε μοντέλα καθοδηγούμενα από τα δεδομένα, τα οποία θα είναι

σχεδιασμένα ώστε να έχουν την δυνατότητα να αλληλεπιδρούν με τα οχήματα υψηλής αυτοματοποίησης (επίπεδα 4 και 5). [16] Μια τέτοια σχεδίαση παρουσιάζεται στην επόμενη εικόνα:



Εικόνα 13. Data-Driven Μοντέλο Οδικής Κυκλοφορίας. Source: Mobility Patterns, Big Data and Transport Analytics, Tools, and Applications for Modeling [16]

Διακρίνουμε δύο μέρη του Data-Driven μοντέλου, αυτό της εκπαίδευσης (training) και εκείνο της εφαρμογής (application). Αρχικά υπάρχει μία βάση δεδομένων με ιστορικά στοιχεία, με έντονη ετερογένεια, για παράδειγμα αριθμοί όπως μετρήσεις ατυχημάτων, όρια ταχυτήτων, ποσοστό υγρασίας, θερμοκρασίες, ή εικόνες όπως πρόσωπα ανθρώπων ή ζώων, με προέλευση από πολλές πηγές όπως από την τροχαία, υπουργείο μεταφορών, μετεωρολογική υπηρεσία, τοπικούς φορείς. Η αφετηρία του μοντέλου είναι ο προσδιορισμός των συστάδων ανάμεσα στα υπάρχοντα δεδομένα. Οι συστάδες αυτές δεν θα είναι γνωστό, πριν την έναρξη του προβλήματος πόσες θα είναι ή με ποια χαρακτηριστικό θα ομαδοποιούνται αλλά το ίδιο το μοντέλο μηχανικής μάθησης θα επιλέγει αλλά και να ανακαλύπτει νέα μοτίβα που υποβόσκουν ανάμεσα στα δεδομένα. Αποθηκεύεται και η ομαδοποίηση των δεδομένων στην αρχική βάση δεδομένων και στη συνέχεια εφαρμόζεται η τεχνική παλινδρόμησης σε κάθε συστάδα ξεχωριστά προσπαθώντας, με ευέλικτο τρόπο, να σχηματίσει όσο το δυνατόν αντιπροσωπευτικά

μοντέλα προβλέψεων για κάθε μία ομάδα, λαμβάνοντας υπόψη την επίδραση των διάφορων μεταβλητών σε αυτήν την συγκεκριμένη συστάδα καθώς διαφέρει το βάρος των μεταβλητών ανάλογα την ομάδα δεδομένων. Εδώ ολοκληρώνεται το στάδιο της εκπαίδευσης. Προχωράμε στο στάδιο της εφαρμογής κατά το οποίο γίνονται νέες παρατηρήσεις, είτε συνεχής π.χ. από κάμερες κυκλοφορίας, drones ή αισθητήρες είτε διακεκομμένες όπως πληροφορίες που αναρτώνται στο μοντέλο κυκλοφορίας από το υπουργείο μεταφορών σε σταθερή ροή ή έκτακτα. Τα νέα στοιχεία βάση των χαρακτηριστικών τους ταξινομούνται στις ομάδες δεδομένων που έχει ορίσει το σύστημα από μόνο του μέσω των του αντιπροσωπευτικών μοντέλων των ομάδων, δηλαδή εντάσσονται στην κλάση που είναι πιο κοντά. Γίνεται μέτρηση του σφάλματος μεταξύ των προβλέψεων και των πραγματικών αξιών των νέων παρατηρήσεων και συνεκτιμάται για την νέα επανάληψη με σκοπό την συνεχής μείωση του σφάλματος και την γενική βελτίωση του μοντέλου [16].

Πιθανόν, τα πρώτα τέτοια μοντέλα οδικής κυκλοφορίας να αρχίζουν να εφαρμόζονται για Μέσα Μαζικής Μεταφοράς τα οποία κινούνται μεμονωμένα, όπως το υπόγειο μετρό και σταδιακά θα περάσει και στις ιδιωτικές μεταφορές, αφαιρώντας σταδιακά τον πιο αδύναμο κρίκο της αλυσίδας των μεταφορών από το ενεργό πεδίο, τον άνθρωπο. Είναι ιδιαίτερα σημαντικό ο περιορισμός του ανθρώπινου παράγοντα στις μεταφορές ειδικά στην οδήγηση οχημάτων σε αυτοκινητόδρομους, ο οποίος αποτελεί τον κύριο παράγοντα ατυχήματος. Άρα αύξηση του γενικού προσδόκιμου ζωής αλλά και ελάφρυνση του συστήματος υγείας από βαριά τραυματισμένους που έχουν εμπλακεί σε τέτοιου είδους καταστάσεις, όπου σε ορισμένες περιπτώσεις οι βλάβες των τραυματισμών είναι μόνιμες και η υποστήριξή τους απαιτεί αρκετούς πόρους από το κράτος.

Επιπρόσθετα, η μη ανάγκη χειρισμού οχήματος από άνθρωπο στα αυτοδηγούμενα οχήματα θα μειώσει το κόστος και το χρόνο μεταφορών ειδικά στις μεταφορές προϊόντων όπου απαιτούνται συνεχής στάσεις για ξεκούραση του οδηγού όταν τα δρομολόγια είναι πολύωρα. Στη συνέχεια, λόγω της αλληλοεπίδρασης των οχημάτων μεταξύ τους καθώς και με ένα κεντρικό σύστημα οδικής κυκλοφορίας της θα υπάρχει ομαλότερη ροή των αυτοκινήτων, αποσυμφορίζοντας επιβαρυμένα οδικά δίκτυα όπως αυτά των αστικών περιοχών σε ώρες αιχμής ή σε τελωνεία που διενεργούνται έλεγχοι και σχηματίζονται συχνά μεγάλες ουρές αναμονής. Τέλος, προς ίδια κατεύθυνση θα αποσυμπιέζονταν, επίσης,

και άλλοι κόμβοι μεταφορών επιβατών και υλικών όπως είναι τα αεροδρόμια και τα λιμάνια με αποτελεσματικότερη διαχείριση των δρομολογίων χρησιμοποιώντας τε δεδομένα τέτοιο τρόπο για την αποφυγή καθυστερήσεων και την συνεχή βελτίωση και εύρεση νέων τρόπων αύξησης της δυναμικής των κόμβων αυτών [18].

1.4.3. Υγεία

Η επιστήμη δεδομένων έχει την προοπτική να επιδράσει θετικά και στον τομέα υγείας σε δύο επίπεδα, αυτού της διάγνωσης και θεραπείας μεμονωμένα ενός ασθενή και του γενικότερου σε επίπεδο κοινωνίας και συστήματος υγείας. Στις ΗΠΑ ήδη από την δεκαετία του 1960, είχε γίνει αντιληπτό η σημασία των δεδομένων κατά την αποτελεσματική άσκηση της Ιατρικής και η διαχείρισής του με ψηφιακό τρόπο. Έτσι δημιουργήθηκε η έννοια των Health Informatics (ή medical informatics), ενός λογισμικού το οποίο θα συγκέντρωνε ιατρική γνώση, κλινικές υποθέσεις και στοιχεία ασθενών με σκοπό την υποστήριξη των ιατρών και ερευνητών στο έργο τους [20]. Από τότε μέχρι σήμερα έχει περάσει και σε πολλές ακόμη χώρες και στο μέλλον προμηνύεται να αποτελεί αναπόσπαστο κομμάτι της ιατρικής εξαιτίας των συνεχόμενων εξελίξεων στην Ιατρική επιστήμη αλλά και των κοντινών επιστημόνων, η αύξηση του ρυθμού δεδομένων και οι προοπτικές εκμετάλλευσής τους. Στην παρακάτω εικόνα απεικονίζεται από που μπορούμε να εκμαιεύσουμε δεδομένα και να αξιοποιηθούν στα Health Informatics.



Εικόνα 14: Οι πηγές των δεδομένων της υγείας στα Health Informatics [20] Source: Health Informatics: Practical Guide

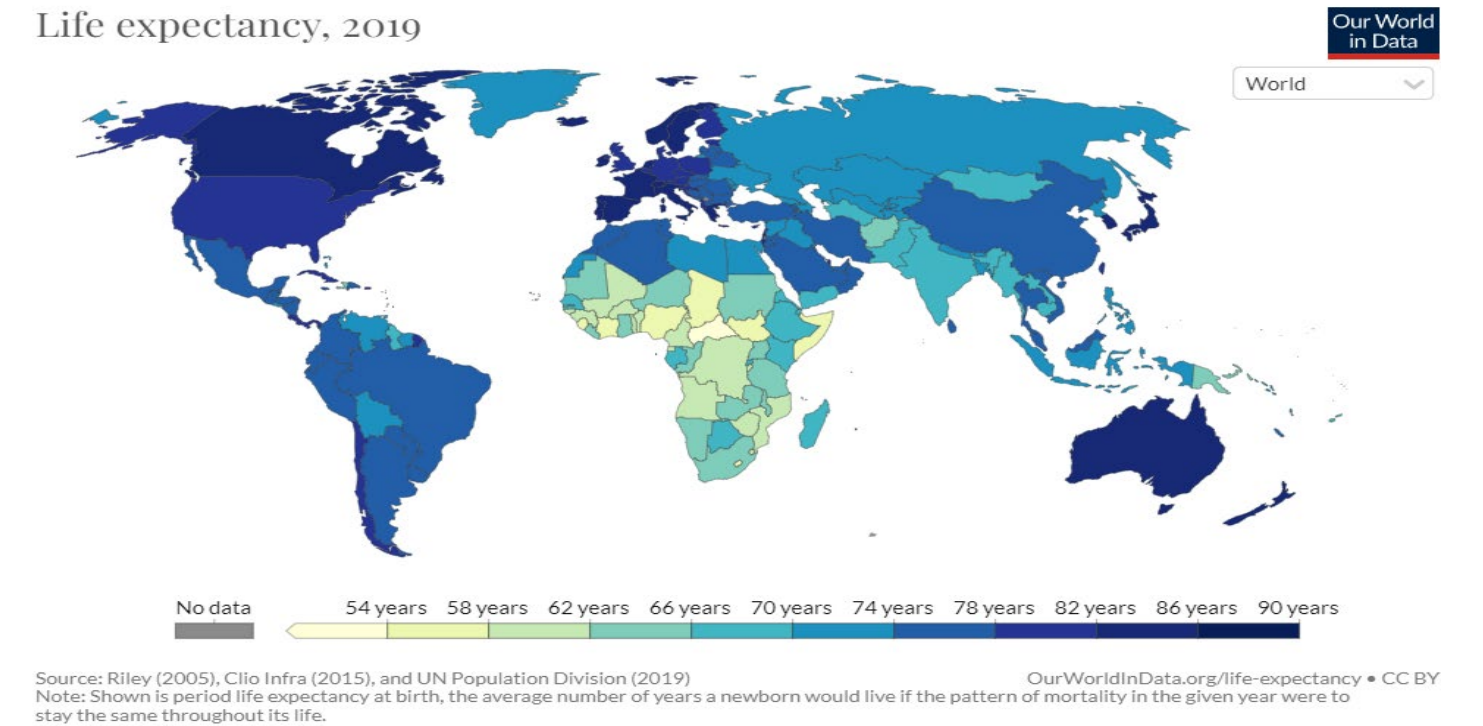
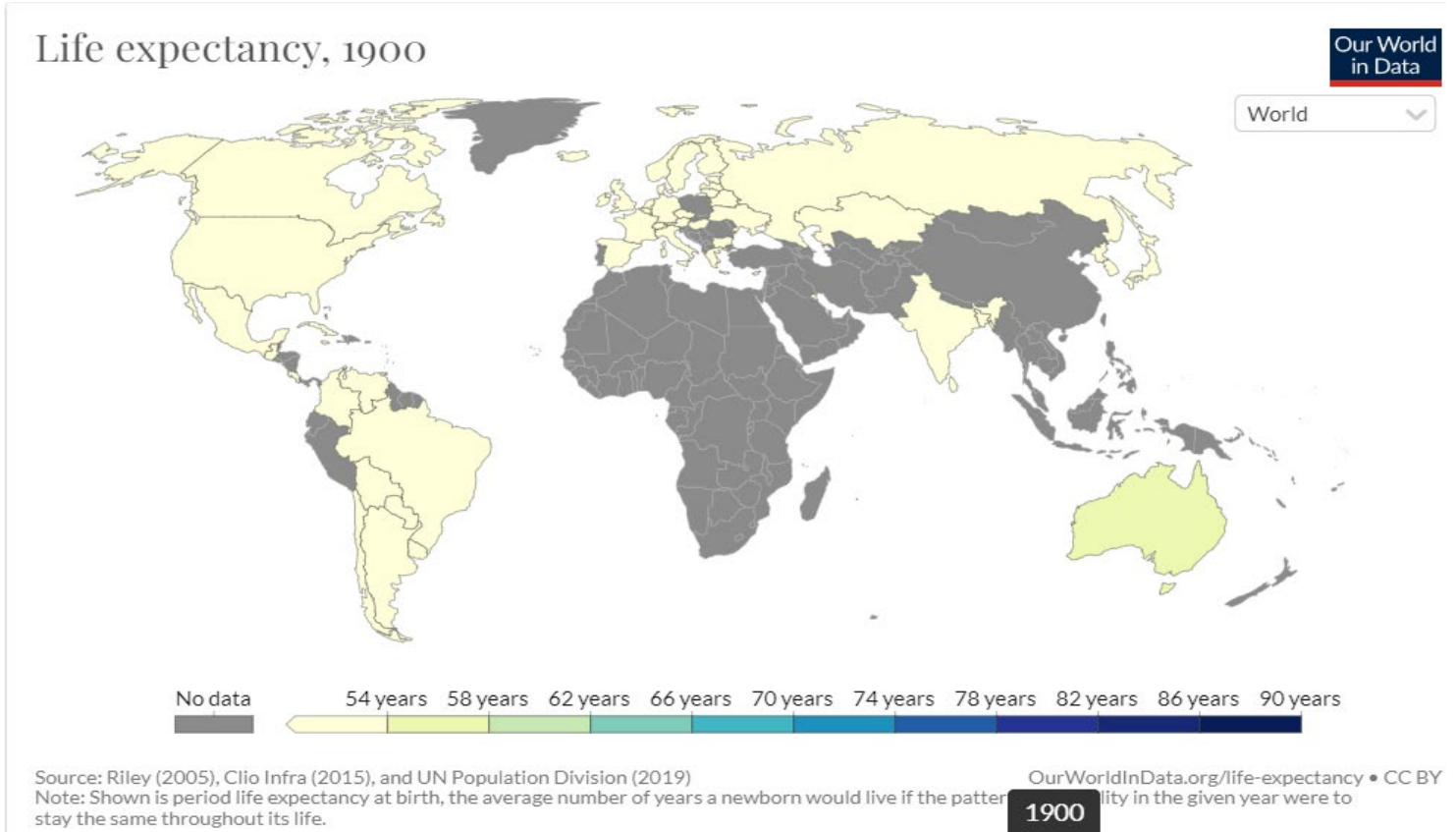
Η συγκέντρωση των ανωτέρω δεδομένων σε μία βάση όπου μέσω ενός λογισμικού θα μπορούσε εύκολα ένας επαγγελματίας της υγείας να ανατρέξει ώστε να βρει απαντήσεις, λιγότερο επίπονα, θα εξοικονομούσε σημαντικές ώρες εργασίας οι οποίες θα μπορούσαν να επενδυθούν πιο αποδοτικά. Πολλές οργανώσεις σε όλο τον κόσμο εργάζονται προς αυτήν την κατεύθυνση, όπως η AMIA (American Medical Informatics Association) μία μη κυβερνητική κοινότητα με έδρα τις ΗΠΑ, που απαρτίζεται από επαγγελματίες και επιστήμονες της επιστήμης δεδομένων και της υγείας, οι οποίοι ανταλλάζουν πληροφορίες και εργάζονται για την βελτίωση της υγειονομικής περίθαλψης. Στην Ευρώπη ο πιο γνωστός οργανισμός με παρουσία σε 32 ευρωπαϊκές χώρες, συμπεριλαμβανομένου και της Ελλάδας, υπάρχει ο μη κερδοσκοπικός οργανισμός EFMI (European Federation for Medical Informatics) που εργάζεται για τον ίδιο σκοπό και προσπαθεί να προωθήσει υψηλά πρότυπα βασισμένα στα δεδομένα, σε διάφορες εφαρμογές της υγείας.

Πολλές από αυτές τις εφαρμογές στηρίζονται στην επιστήμη των δεδομένων και την μηχανική μάθηση(machine learning), με κυριότερο παράδειγμα την βαθιά μάθηση (deep learning) με ελπιδοφόρα για την ίδια την ανθρωπότητα προοπτική εξέλιξης, εισάγοντας επιπλέον μία νέα έννοια, αυτή της Τεχνολογίας της Υγείας(HealthTech), ενισχύοντας τον ρόλο της τεχνολογίας στην ιατρική επιστήμη και σηματοδοτώντας μία

νέα αντίληψη όπου είναι αναγκαίο ένας υγειονομικός επαγγελματίας να είναι εξοικειωμένος με την χρήση πληροφοριακών και ψηφιακών εργαλείων. Μερικές από αυτές συναντάμε [20]:

- Στην ανακάλυψη νέων φαρμάκων και τοξικολογίας, με χαρακτηριστικό παράδειγμα AtomNet [23] το σύστημα Deep Learning το οποίο απαρτίζεται από ένα συνελκτικό νευρωνικό δίκτυο για την πρόβλεψη βιοδραστικότητας κατά την έρευνα νέων φαρμάκων ώστε να προβλεφθούν πιθανές επιπτώσεις ασθενείς πιο αποτελεσματικά, μειώνοντας και τα τωρινά χρονικά όρια που απαιτούνται στις διαδικασίες αυτές.
- Στην πρόληψη νοσημάτων μέσω εφαρμογών παρακολούθησης, ειδικά σε απομακρυσμένες περιοχές που εφαρμόζεται Τηλεϊατρική, ανιχνεύοντας έγκαιρα βλάβες, ασθένειες και ανωμαλίες του ανθρώπινου οργανισμού και γονιδιακή προδιάθεση[28].
- Στην καλύτερη διάγνωση, προσφέροντας βελτίωση επί της εικόνας του ανθρώπινου οργανισμού κατά το στάδιο των διαγνωστικών εξετάσεων όπως του κλάδου της ακτινολογίας
- Στην ιατροφαρμακευτική περίθαλψη ασθενών για πιο στοχευμένη αντιμετώπιση νοσημάτων περιορίζοντας τυχόν παρενέργειες για την θεραπεία ασθενειών, για παράδειγμα η ταξινόμηση καρκινικών κυττάρων ώστε να απομονωθούν οι κλάσεις κυττάρων τα οποία είναι υγιή.
- Στην οργάνωση και διοίκηση νοσοκομείο και κλινικών, μεταβαίνοντας σε ένα μοντέλο Data-Driven ιδρυμάτων για την καλύτερη αξιοποίηση πόρων και κυρίως του υγειονομικού προσωπικού.[15]
- Την χάραξη των εθνικών στρατηγικών του συστήματος υγείας αλλά και σε διεθνές επίπεδο όπως απαιτήθηκε και από την προσπάθεια αντιμετώπισης και περιορισμού της πανδημίας σε ολόκληρο τον κόσμο με την αξιοποίηση δεδομένων που αφορούν το βιοτικό επίπεδο των πληθυσμών: η διατροφή, πρόσβαση σε καθαρό νερό και οργανωμένο σύστημα αποχέτευσης, το προσδόκιμο ζωής, μετρήσεις κρουσμάτων ασθενειών και το ποσοστό θνησιμότητας τους, αναπαραγωγικός ρυθμός, επίπεδο συστήματος υγείας ανά περιοχές/πόλεις/χώρες. Μέσω αυτών των στοιχείων και με την κατάλληλη οπτικοποίηση αποτυπώνεται οι επιδράσεις στους πληθυσμούς των

διάφορων υγειονομικών δράσεων που εφαρμόζουν οι κυβερνήσεις και ύστερα αποφασίζουν για την αποτελεσματικότητά τους. Για παράδειγμα, η εμβολιαστική προσέγγιση στην υγειονομική πρόληψη των πολιτών φαίνεται να έχει αποδώσει και αποτελεί τα τελευταία χρόνια η κυρία μέθοδος ανοσοποίησης των ανθρώπων από τον προηγούμενο όπως φαίνονται στα παρακάτω ενδεικτικά διαγράμματα, με απλοϊκή γενίκευση. [2]



Εικόνα 15. Το προσδόκιμο ζωής του πλανήτη από την έναρξη 20ου αιώνα έως σήμερα μέσω οπτικοποίησης χάρτη (πάνω) [2]

Infectious diseases before and after a vaccine became available



Data sources: Smallpox: Our World in Data based on Guy (1862) and several publications of the Registrar General between 1896 and 1903
 Polio: Our World in Data based on US Public Health Service (1910-1951) and US Center for Disease Control (1960-2010)
 Measles: Our World in Data based on several publications from the Public Health Reports, the US Public Health Service's Morbidity and Mortality Weekly Report Annual Supplements, and the US Census's Annual Statistical Abstracts
 OurWorldinData.org – Research and data to make progress against the world's largest problems. Licensed under CC-BY by the author Max Roser.

Εικόνα 16. Ο ρυθμός μετάδοσης και θανάτων ορισμένων ασθενειών με κύρια γεγονότα την εισαγωγή εμβολίων ως μέτρο πρόληψης [2]

Από τα ανωτέρω, γίνεται εύκολα αντιληπτό το θετικό αντίκτυπο γενικότερα προς όφελος των ανθρώπων αλλά και της οικονομίας της υγείας, οι μελλοντικές εφαρμογές της εκμετάλλευσής των δεδομένων στον κλάδο αυτό. Συγκεκριμένα ορισμένα από τα θετικά αναλύονται ως εξής:

- Η έγκαιρη διάγνωση ασθενειών ακόμη και πριν εκδηλωθούν μέσω ανάλυση στοιχείων από νοσούντες και εντοπίζοντας μοτίβα , δίνοντας, έτσι, την ευκαιρία να δράσει ο ασθενής ώστε να προετοιμαστεί κατάλληλα ή και να προφυλαχθεί πλήρως, εξοικονομώντας ταυτόχρονα οικονομικούς πόρους του υγειονομικού συστήματος από μακροχρόνιες και ακριβές θεραπείες.
- Η μείωση του χρόνου ίασης των ασθενών μέσω καινοτόμων και εξατομικευμένων φαρμάκων πλήρως στο προφίλ του ασθενή, το οποίο προφίλ θα ανήκει σε συγκεκριμένες συστάδες ασθενών, η οποία κάθε συστάδα θα απαιτεί συγκεκριμένες ενέργειες για αντιμετώπισης του νοσήματος λαμβάνοντας υπόψη και τα λοιπά χαρακτηριστικά που θα παίζουν σημαντικό ρόλο όπως το φύλλο, ηλικία , ιατρικό παρελθόν κ.τ.λ.
- Η αποτελεσματικότερη χρήση των διαθέσιμων πόρων, θα εξασφάλιζε την απεμπλοκή του ανθρώπινου παράγοντα από απλές λειτουργίες και θα μείωνε το φόρτο εργασίας σε ιατρούς και νοσηλευτές, οι οποίοι είναι υποστελεχωμένοι συνήθως στα νοσηλευτικά ιδρύματα ιδιαίτερα στην χώρας μας, λόγω του υψηλού απαιτούμενου γνωστικού υπόβαθρου των ρόλων αυτών.
- Επιπλέον, θα δημιουργούνταν τα περιθώρια για περαιτέρω έρευνα σε κλινικό επίπεδο ακόμη επίπεδο μέσω της καλύτερης αξιοποίησης του ανθρώπινου παράγοντα εντός των νοσοκομείων ανατροφοδοτώντας το πληροφοριακό σύστημα της υγείας με ακόμη ποιοτικότερα δεδομένα, μειώνοντας τον χρόνο και το κόστος των ερευνών.

Πέρα την χρήση της επιστήμης των δεδομένων από τους επαγγελματίες και του συστήματος υγείας, θα επιφέρει επιπρόσθετες θετικές συνέπειες στην γενική υγεία των ανθρώπων λόγω της απόρριψης της εφαρμογής της επιστήμης δεδομένων σε άλλους τομείς όπως στις μεταφορές προκαλώντας μείωση στα ατυχήματα ή στην διαχείριση των πόλεων με την αντιμετώπιση προβλημάτων υγείας που έχουν δημιουργηθεί λόγω της επιβαρυνμένης διαβίωσης σε μητροπόλεις (μολυσμένος αέρας και νερό ή άγχος λόγω της πυκνοκατοικημένης και άναρχης οικοδόμησης) και δίνοντας ταυτόχρονα την δυνατότητα να απαντηθούν ερωτήματα και να αναγνωρίζονται οι λόγοι της αποτυχίας μίας υγειονομικής στρατηγικής, αντιλαμβάνοντας τυχόν κενά σε αυτήν. Για παράδειγμα, σύμφωνα με μετρήσεις του Παγκόσμιου Οργανισμού Υγείας το 2019 [24], στην χώρα των

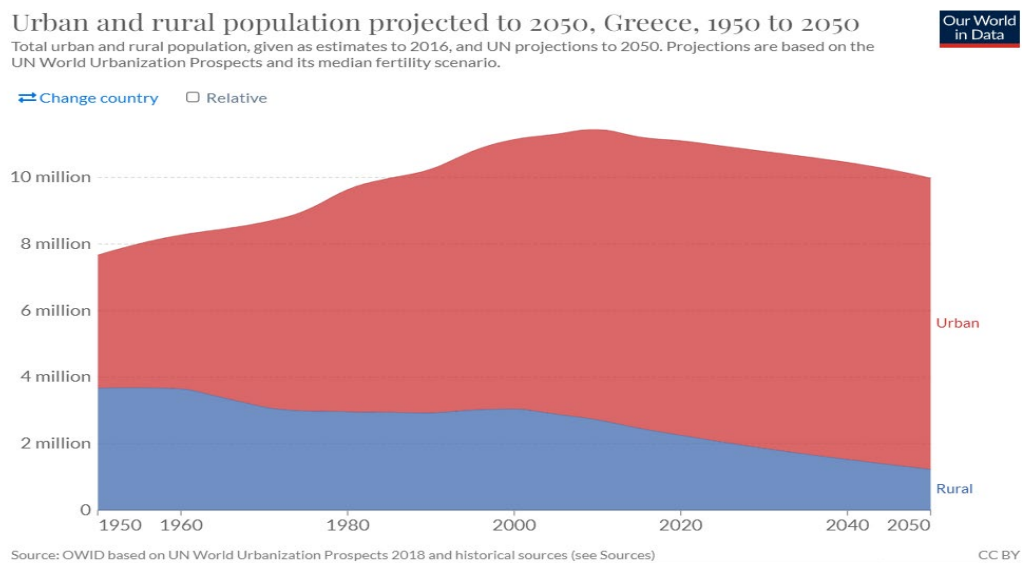
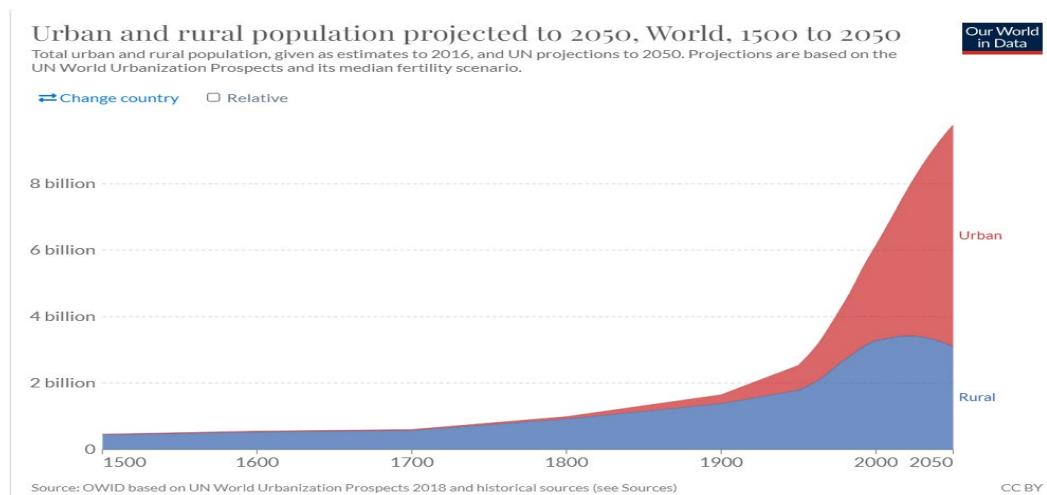
ΗΠΑ ενώ δαπανιούνται τα περισσότερα χρήματα στον κόσμο σε υγειονομικά έξοδα ανά κάτοικο ετησίως (10.921,01\$) τα οποία αποτελούν το 16.77% του Α.Ε.Π., καθιστώντας την πάλι στην πρώτη στην κατηγορία αυτή (εξαιρώντας την ακραία περίπτωση του πολύ μικρού νησιωτικού κράτους Τουβαλέ) βρίσκεται για το ίδιο έτος σύμφωνα με τον Οργανισμό Ηνωμένων Εθνών[25] στην 70η θέση ως προς το προσδόκιμο ζωής των κατοίκων της χώρας αυτής, γεγονός που δημιουργεί εύλογα ερωτήματα συνυπολογίζοντας το υψηλό τεχνολογικό και επιστημονικό επίπεδο της περιοχής αυτής. Αυτό οφείλεται για πολλούς λόγους όπως η υψηλή θνησιμότητα σε εφήβους σε συγκεκριμένες περιοχές της χώρας λόγω εγκληματικών δράσεων ή τα μεγάλα ποσοστά παχυσαρκίας από την παιδική ηλικία κιόλας ή της αδυναμία πρόσβασης στο σύστημα υγειονομικής ασφάλισης από το σύνολο του πληθυσμού(ίσως η πιο σημαντική αιτία) όπου σε όλες τις αναπτυγμένες χώρες του κόσμου και σε πολλές αναπτυσσόμενες σχεδόν το 100% του πληθυσμού του παρέχονται υπηρεσίες υγείας σε αντίθεση με το 85% των ΗΠΑ. Με αυτό το παράδειγμα καταδεικνύεται ότι η αλόγιστη διάθεση οικονομικών πόρων χωρίς να έχουν αναγνωριστεί τα πραγματικά προβλήματα δεν μπορεί να λυθούν σημαντικά προβλήματα της σύγχρονης εποχής όπως η επιμήκυνση της ανθρώπινης ζωής.

1.4.4. Ενέργεια και έξυπνες πόλεις

Στις προηγούμενες βιομηχανικές επαναστάσεις, η κυριότερη πτυχή που άλλαξε είναι ο τρόπος που ζούμε μέσα στις πόλεις και έτσι και στην 4η Βιομηχανική Επανάσταση (Βιομηχανία 4.0), προμηνύεται να γίνει κάτι αντίστοιχο. Συγκεκριμένα, οι πόλεις σταδιακά θα γίνονται “έξυπνότερες”, η οποία έξυπνάδα θα στηρίζεται στην τεχνολογία IoT (Internet of Things) για να συλλέγει τα δεδομένα τα οποία θα απαιτούνται για μίας καλύτερη οργάνωση, διοίκηση και διανομή αγαθών, υπηρεσιών και ενέργειας προς στους κατοίκους.

Συγκεκριμένα, η πιο αποδοτική δημιουργία, διανομή και κατανάλωση της ενέργειας θα αποτελεί κλειδί για την επιτυχία μίας σύγχρονης πόλης για την ευημερία και επιβίωσή της. Μέσω της ανάλυσης των δεδομένων και της λήψης αποφάσεων βάση των αποτελεσμάτων της ανάλυσης θα σχεδιάζονται όλα τα αστικά έργα όπως το σύστημα αποχέτευσης, καθαριότητας και συλλογής σκουπιδιών μέχρι και τον ίδιο τον αστικό

σχεδιασμό και την διοίκηση της πόλεως[27]. Αυτό το γεγονός είναι ιδιαίτερα κρίσιμο καθώς προβλέπεται, σύμφωνα με τον οργανισμό Ηνωμένων Εθνών, μέχρι το 2050 ο παγκόσμιος αστικός πληθυσμός θα αυξηθεί κατά 2,25 δισεκατομμύρια και θα ανέρχεται στα 6,7 δισεκατομμύρια το οποίο θα αντιστοιχεί στα δύο τρίτα περίπου του συνολικού παγκοσμίου πληθυσμού (9,8 δισεκατομμύρια) ενώ ο αγροτικός θα διαμορφώνεται στα 3,1 δισεκατομμύρια. Στην Ελλάδα ήδη η αστικοποίηση είναι έντονη καθώς σήμερα το 80% των Ελλήνων ζουν σε αστικές περιοχές και μέχρι το 2050 το ποσοστό θα έχει ανέβει στο 88% περίπου. [25]



Εικόνα 17. Η αστικοποίηση με την πάροδο του χρόνου και προβλέψεις έως το 2050 παγκοσμίως[2]

Καθοριστικό ρόλο, λοιπόν, στις ανωτέρω προβλέψεις για την αύξηση της αστικοποίησης διαδραματίζουν οι μελλοντικές έξυπνες πόλεις και οι τεχνολογίες που θα τις υποστηρίζουν. Οι τεχνολογίες αυτές θα αφορούν κυρίως τα πεδία [26]:

- Της Τεχνητής Νοημοσύνης (AI)
- IoT

Οι συνδυασμοί των ανωτέρω στα διάφορα αστικά έργα θα δημιουργήσουν υπεραξία, αξιοποιώντας πλήρως τις δυνατότητες που μπορεί να προσφέρει η επιστήμη δεδομένων και θα εξασφαλίσουν τις δύο βασικές προϋποθέσεις των σύγχρονων αστικών κέντρων, να είναι βιώσιμα και φιλικά προς τους κατοίκους.

Με την φράση «φιλικά προς τους κατοίκους» συνοψίζουμε όλα αυτά που είναι απαραίτητα ώστε να μπορούν οι άνθρωποι να έχουν μία ήρεμη, υγιή και ευτυχισμένη ζωή ζώντας αρμονικά σε μία πυκνοκατοικημένη περιοχή. Κάποια από τα βασικά συστατικά στοιχεία ενός τέτοιου εγχειρήματος θα ήταν[27]:

α) Η αστική συγκοινωνία προσαρμοσμένη με ένα κατάλληλο Data-Driven Traffic Model [16], όπως αναφέρθηκε και στην υποενότητα 1.4.2. (Μεταφορές), που θα μπορούσε να αυξήσει την χρησιμότητά τους. Ταυτόχρονα η ενσωμάτωση έξυπνων εφαρμογών όπως μίας τέτοιας εφαρμογής παρκαρίσματος η οποία θα μπορούσε να είναι ενσωματωμένη στα λογισμικά των έξυπνων αυτοκινήτων ή σε άλλες έξυπνες συσκευές των οδηγών και την ίδια ώρα θα αλληλοεπιδρούσε με αισθητήρες εγκατεστημένους στις θέσεις παρκαρίσματος, μειώνοντας με αυτό τον τρόπο τον χρόνο που απαιτείται για εύρεση κενής θέσης για να αφήσει το αυτοκίνητο του ο οδηγός[18].

β) Η χρήση ενός συστήματος ανάδρασης πληροφοριών και αξιολογήσεων από αυτούς που τις παρέχουν σε αυτούς που τις λαμβάνουν και την τοπική διοίκηση, με σκοπό την αναβάθμιση του επιπέδου των παρεχόμενων υπηρεσιών. Ταυτόχρονα, η διάθεση των πόρων από τους παρόχους υπηρεσιών είτε από τον ιδιωτικό είτε από τον δημόσιο φορέα θα μπορούσε να πραγματοποιηθεί με βέλτιστη στόχευση προς αυτούς που τις έχουν ανάγκη και τις χρησιμοποιούν και όχι προς το σύνολο των κατοίκων, οδεύοντας έτσι σε ολοένα πιο εξατομικευμένες υπηρεσίες αλλά και ικανοποιημένους πολίτες[28].

γ) Συστήματα δεδομένων ασφαλείας έναντι της εγκληματικότητας όπως χρήση καμερών για αναγνώριση προσώπων ή πινακίδων κυκλοφορίας οχημάτων και γενικότερα της δημοσίας ασφάλειας (π.χ. αναγνώριση επερχόμενων έκτακτων ακραίων φαινομένων

έχοντας καλύτερη εικόνα σε τοπικό επίπεδο, ειδοποιώντας αυτόματα και έγκαιρα το αντίστοιχο λογισμικό διαχείρισης δεδομένων της τοπικής διοίκησης για επικίνδυνα σενάρια στα οποία θα είχε εκπαιδευτεί μέσω της μηχανικής μάθησης).

δ) Ενός αποκεντρωμένου δικτύου λήψης και καταμέτρησης τιμών για ορισμένους δείκτες που αποσκοπούν σε ένα υψηλό βιοτικό επίπεδο όπως τα μικροσωματίδια του αέρα, η αλλαγή στα επίπεδα των χημικών στοιχείων του δικτύου υδροδότησης, η ηχορύπανση του κέντρου και στη συνέχεια να διενεργούνται ανάλογες και στοχευμένες ενέργειες [20].

ε) Η επίβλεψη των υπηρεσιών κοινής ωφέλειας (η παροχή ύδρευσης, αποβολή και ανακύκλωση απορριμμάτων, προστασία του περιβάλλοντος) μέσω μίας αποτελεσματικής διαχείριση των Big Data τα οποία σε ένα μεγάλο και πυκνοκατοικημένο αστικό κέντρο θα ήταν δύσκολη η επίβλεψή τους μέσω συμβατικών τρόπων. Μέσω της χρήσης δεδομένων και της ανατροφοδότησης των δεδομένων των πολιτών μπορούν να εξαλειφθούν φαινόμενα κακοδιαχείρισης ή διαφθοράς, όπως συνέβη σε μία επαρχία του Πακιστάν το 2008 σύμφωνα με την Παγκόσμια Έκθεση Ανάπτυξης 2021[12]. Στην επαρχία αυτή οι κρατικές υπηρεσίες δεν ήταν επαρκείς και υπήρχαν φαινόμενα δωροδοκίας και εκβιασμών από υπαλλήλους. Για την εξάλειψη τέτοιων φαινομένων οι κρατικοί φορείς υιοθέτησαν ένα απλό και εύχρηστο σύστημα ανατροφοδότησης των υπηρεσιών που λαμβάνανε οι πολίτες και όσο συνέχιζε και τροφοδοτούνταν με νέα δεδομένα, ήταν εφικτό να εντοπιστούν επαναλαμβανόμενα μοτίβα διαφθοράς και ανεπάρκειας και κατά επέκταση ο εντοπισμός υπεύθυνων και ελλείψεων με σκοπό να ενισχύσουν σε συγκεκριμένα σημεία αλλά με συγκεκριμένους πόρους και εργαλεία. Η επιτυχία του προγράμματος ήταν μεγάλη και επεκτάθηκε και σε άλλες περιοχές της χώρας αλλά και τομείς, διευρύνοντας έτσι την βάση δεδομένων ανακαλύπτοντας νέες σχέσεις μεταξύ διάφορων παραγόντων.

Στη συνέχεια για την εξασφάλιση της βιωσιμότητας των πόλεων, αποτελεί μείζον ζήτημα η ενέργεια και παροχή της στους πολίτες σε υψηλά επίπεδα απόδοσης και αξιοπιστίας.[27] Για να εξασφαλιστούν αυτά τα επίπεδα, η διαχείριση των Big Data από τις εταιρείες παραγωγής ενέργειας και των δικτύων του μπορεί να προσφέρει καινοτόμες ιδέες όπως:

α) Καλύτερο προγραμματισμό παραγωγής ειδικά όταν βασίζονται σε ανανεώσιμες πηγές ενέργειας οι οποίες βασίζονται στις καιρικές συνθήκες, οι οποίες δεν είναι σταθερές. Έγκυρες προβλέψεις βασιζόμενες σε μετεωρολογικά και περιβαλλοντολογικά δεδομένα

θα ενίσχυαν έναν αποδοτικό σχεδιασμό ενός μείγματος διάφορων πηγών ενέργειας ώστε να μπορεί τροφοδοτεί το δίκτυο ηλεκτροδότησης της πόλεως οικονομικά και αδιάκοπα.

β) Η αξιοπιστία των δικτύων ηλεκτροδότησης και η εξασφάλιση αδιάκοπης παροχής ακόμη και σε περίοδο υψηλής ζήτησης αποτελεί κρίσιμο στοιχείο για την βιωσιμότητα μίας πόλης. Αυτό θα μπορούσε να επιτευχθεί αποτελεσματικά με την παρακολούθηση της τάσης και κατανάλωσης σε πραγματικό χρόνο και σε συνδυασμό με ιστορικά στοιχεία να λαμβάνονται σωστές αποφάσεις προς αποφυγή τυχόν black-out καθώς και την συνεχή βελτίωση του δικτύου.

γ) Χρήση δεικτών ενεργειακής απόδοσης οι οποίοι θα έχουν διαμορφωθεί βάση συνδυαστικών ιστορικών στοιχείων με σκοπό τον ανίχνευση ευάλωτων σημείων στο ενεργειακό δίκτυο. Οι δείκτες θα τροποποιούνται συνεχώς μέσω της μηχανικής μάθησης προσπαθώντας να έχουν χαμηλό ποσοστό σφάλματος και θα ανατροφοδοτούνται συνεχώς με νέα δεδομένα.

δ) Την δημιουργία έξυπνων δικτύων ηλεκτρικής ενέργειας (smart grids)[28] τα οποία θα επέτρεπαν μία αμφίδρομη ροή ηλεκτρικής ενέργειας αλλά και δεδομένων με την χρήση ψηφιακών εργαλείων. Η παραγωγή ενέργειας θα ήταν αποκεντρωμένη με την χρήση μικρών και πολλών συλλεκτών ενέργειας όπως ηλιακά πάνελ ή μικρές ανεμογεννήτριες σε διάφορα σημεία (ταράτσες πολυκατοικιών, πάρκα), αντίθετα με το μοντέλο που εστιάζει σε λίγα και μεγάλα σημεία παραγωγής ενέργειας και ύστερα μεταφοράς μέσω καλωδίων στους καταναλωτές. Επίσης, η αμφίδρομη μεταφορά δεδομένων θα βελτίωνε την απόδοση των δεικτών ενέργειας και θα επέτρεπε σε προληπτικές δράσεις για αποφυγή προβλημάτων. Σημαντικό ακόμη στο ενεργειακό τομέα των πόλεων διαδραματίζει ο ανεφοδιασμός καυσίμων τα κέντρα διανομής, ο οποίος βασιζόμενος σε ένα σύστημα διαχείρισης δεδομένων, θα εξασφάλιζε μειωμένα κόστη χωρίς να τίθεται κίνδυνος επάρκειας αποθεμάτων. Τα χαμηλά κόστη θα προκύπταν λόγω μειωμένου κόστους αποθήκευσης αφού θα λαμβάνονταν υπόψη περίοδοι χαμηλής ή υψηλής κατανάλωσης και σε συνδυασμό με τις αυξομειώσεις των χρηματιστηριακών τιμών των καυσίμων θα μπορούσε να επιτευχθεί μέγιστη εξοικονόμηση. Για παράδειγμα σε περίοδο που οι τιμές των καυσίμων είναι χαμηλές πρόσκαιρα, ίσως εξαιτίας κάποιου ειδικού γεγονότος του οποίου δεν θα διαρκέσει για αρκετό καιρό, και αναμένοντας υψηλή ζήτηση στο επόμενο διάστημα, ακόμη και αν η παρούσα να είναι χαμηλής κατανάλωσης, θα μπορούσε να προμηθευτούν καύσιμα

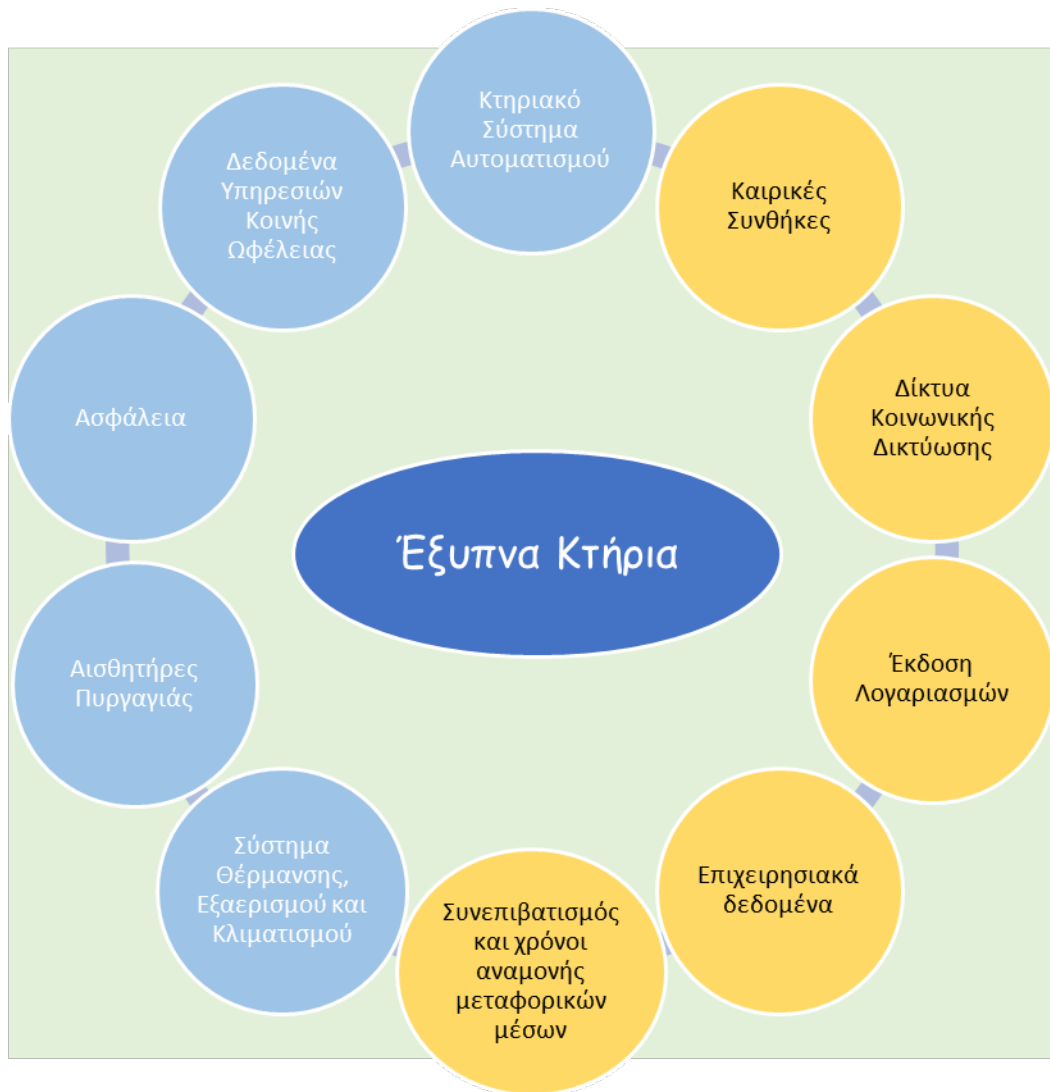
είτε τα κέντρα διανομής ή αγοραστές μεγάλων ποσοτήτων με δυνατότητα αποθήκευσης, όπως μεγάλα δημόσια κτηριακά συγκροτήματα ή τοπικές βιομηχανίες σε τέτοιο σημείο όπου το κόστος αποθήκευσης δεν θα υπερβαίνει την ωφέλεια από τις χαμηλές τιμές κτήσης καυσίμων.



Εικόνα 18. Θεωρητικό μοντέλο μία έξυπνης πόλης του μέλλοντος Source: Aliga IT Company

Στην ανωτέρω απεικόνιση, αποτυπώνονται οι τομείς που θα επηρεάζονταν θετικά κατά κύριο λόγο με την εισαγωγή των νέων τεχνολογιών στην εξέλιξη και σύνθεση των πόλεων. Προς τη ίδια κατεύθυνση θα είναι και η αρχιτεκτονική και δόμηση των κτηρίων τα οποία θα προσαρμοστούν σε ένα σύγχρονο περιβάλλον και θα αποτελούν κόμβοι πληροφοριών καθώς θα αποστέλλουν και θα λαμβάνουν δεδομένα, δηλαδή θα αλληλεπιδρούν άμεσα με το περιβάλλον τους ώστε να διεκπεραιώνονται καθημερινές εργασίες. Ουσιαστικά θα δημιουργούνται πολλές έξυπνες υπό-κοινότητες μέσα σε ένα πιο μεγάλο εύρος με τα ίδια χαρακτηριστικά. Σε συνδυασμό με την διαχείριση των δεδομένων και των τεχνολογιών AI και IoT θα εξασφαλίζεται ένα ασφαλές και φιλικό περιβάλλον για τους ενοίκους (καθαρό αέρα και νερό, συστήματα ασφαλείας και συναγερμών,

αισθητήρες πυρκαγιάς και διαρροών, αυτοματοποιημένες εφαρμογές κ.τ.λ.) με ορθολογική χρήση των διαθέσιμων πόρων όπως φαίνεται στο παράδειγμα της παρακάτω εικόνας. Επιπρόσθετα, θα αποτελούσαν κύρια συστατικά ενός έξυπνου δικτύου ηλεκτρικής ενέργειας (smart grids) [28] από τα οποία θα γινόταν ταυτόχρονα κατανάλωση και παραγωγή, δημιουργώντας ποιοτικά δεδομένα με σκοπό μία αποδοτικής και αποκεντρωμένης ενεργειακής κοινότητας.



Εικόνα 19. Θεωρητικό μοντέλο ενός έξυπνου κτηρίου όπου με γαλάζιο χρώμα είναι χρωματισμένες οι λειτουργικές τεχνολογίες ενός "έξυπνου" κτηρίου και με κίτρινο οι εξωτερικές πηγές πληροφοριών και λειτουργιών. Source: Johnson Controls (Group of Companies)

Η μετάβαση των παραδοσιακών πόλεων σε έξυπνες πόλεις θα επέφερε σημαντικά οφέλη ως προς την καθημερινότητα και την διαβίωση[27] σε αυτές:

- Θα μείωνε την ανάγκη για χρήση ιδιωτικών οχημάτων με ταυτόχρονη βελτίωση της οδηγικής εμπειρίας εντός των πόλεων.
- Κατά συνέπεια, ο μολυσμένος ατμοσφαιρικός αέρας, το οποίο είναι έντονο πρόβλημα των τωρινών πόλεων και πηγή πολλών προβλημάτων υγείας, θα ήταν λιγότερο επιβαρυνμένος. Σε αυτό το γεγονός, θα συνέβαλλαν και άλλες θετικές απόρροιες της εφαρμογής των έξυπνων πόλεων όπως μειωμένου χρόνου εύρεσης θέση στάθμευσης ή της συνεχούς παρακολούθησης της ποιότητας του ατμοσφαιρικού αέρα.
- Θα διευκολυνόταν η ανάπτυξη του μεγέθους της πόλης σε έκταση και πληθυσμό παραμένοντας βιώσιμη.
- Εξοικονόμηση ενέργειας μέσω της αποδοτικότερης χρήσης του μείγματος πηγών με περαιτέρω αξιοποίηση των ανανεώσιμων πηγών και επιπρόσθετα από την μείωση της απώλειας μεταφοράς της ενέργειας εξαιτίας ενός αποκεντρωμένου ενεργειακού ηλεκτρικού δικτύου όπου σε μεγάλο ποσοστό της διακινούμενης ενέργειας του δικτύου η κατανάλωση γίνεται στο ίδιο σημείο όπου γίνεται και η παραγωγή.
- Ορθότερη τιμολόγηση των καταναλωτών ενέργειας με χρήση των κατάλληλων ενεργειακών δεικτών και μείωση της αγοράς καυσίμων διατηρώντας επαρκή αποθέματα.
- Επαύξηση του αισθήματος εμπιστοσύνης προς τις τοπικές αρχές. Σε μια έρευνα που χρηματοδοτήθηκε από την Παγκόσμια Τράπεζα για το πρόγραμμα που αναφέρθηκε στην περιοχή του Πακιστάν[12], περίπου το 90% των ερωτηθέντων απάντησε ότι ενισχύθηκε η εμπιστοσύνη μεταξύ πολιτών και κρατικών φορέων ενώ άνω του 50% είπαν ότι όλες οι παρεχόμενες υπηρεσίες είχαν βελτιωθεί, όπως για παράδειγμα το ποσοστό της δωρεάν διανομή φαρμάκων από τα νοσηλευτικά ιδρύματα στην πρώτη επαρχία του Πακιστάν που εφαρμόστηκε το πρόγραμμα ανατροφοδότησης των πολιτών, ανήλθε στο 77% το 2019 έναντι του 46% του 2015 ύστερα από έρευνα σε άτομα που είχαν επισκεφτεί κάποιο νοσοκομείο και χρειάστηκε φαρμακευτική περίθαλψη. Αν αναλογιστεί κανείς ότι το παραπάνω παράδειγμα βασίστηκε σε μία απλή διαδικασία αποστολή μηνυμάτων κινητών(sms) και η επεξεργασία αυτών από υπαλλήλους, τι δυνατότητες θα είχε μία τέτοια

προσπάθεια με χρήση πιο σύγχρονων τεχνολογιών όπως αξιολόγηση μέσω διαδικτυακών εφαρμογών και διαχείριση των αξιολογήσεων μέσω ενός λογισμικού Deep Learning το οποίο θα αναγνώριζε τις θετικές και αρνητικές αξιολογήσεις και θα τις ταξινομούσε ανάλογα.

2. Η αρχιτεκτονική μίας ομάδας διαχείρισης δεδομένων και ο ρόλος της στις επιχειρήσεις

2.1. Οι ρόλοι μίας ομάδας διαχείρισης

Η δομή μίας ομάδας διαχείρισης δεδομένων είναι σημαντική, καθώς οι τεχνολογίες στο τομέα αυτό εξελίσσονται και αλλάζουν συνεχώς, καθώς και οι ανάγκες και οι πιθανές χρήσεις. Έτσι, μια επιτυχημένη αρχιτεκτονική μίας τέτοιας ομάδας πρέπει να είναι ευπροσάρμοστη στις αλλαγές, ακολουθώντας τα γεγονότα αλλά και ταυτόχρονα διατηρώντας τους διακριτούς ρόλους μέσα σε αυτή. Η αντίληψη ύπαρξης ατόμων με συσσωρευμένες ικανότητες σε όλο το εύρος της επιστήμης δεδομένων, από προγραμματιστικές γνώσεις έως και αυξημένη κριτική αντίληψη και ανάλυση, είναι παρωχημένη και οδηγεί σε αποτυχία[29]. Αντιθέτως, η διάκριση των εργασιών σε στάδια των οποίων την ευθύνη θα έχουν αντίστοιχες υπο-ομάδες, ακόμη και αν ο οργανισμός είναι μικρός, θα επέφερε εξειδίκευση σε κάθε στάδιο των διάφορων έργων και διεργασιών και παράλληλα θα μπορούσε η ομάδα να παράγει καλύτερα και αξιόπιστα προϊόντα δεδομένων, είτε αυτά είναι απλές προγνώσεις είτε ‘βαριά’ λογισμικά προσομοιώσεων ή μηχανικής μάθησης, προς υποστήριξη ολόκληρου του οργανισμού. Η διάκριση αυτή θα μπορούσε να συνοψιστεί σε τρία βασικά επίπεδα, α) Τεχνικό Επίπεδο, β) Λειτουργικό Επίπεδο γ) Συνολικό Επίπεδο και Εξαγωγή Αποτελεσμάτων, στα οποία υπάρχει ένας εσωτερικός κεντρικός έλεγχος. Το μέγεθος και ιδιότητες των μελών των υπο-ομάδων διαφέρει αναλόγως πολλών παραγόντων οι οποίοι όπως είναι φυσικό διαφέρουν σε κάθε περίπτωση.

2.1.1. Τεχνικό επίπεδο

Στόχος αυτής της υπο-ομάδας είναι να δώσει υλική και ψηφιακή υπόσταση στο έργο όλης της ομάδας, κατασκευάζοντας αρχικά έναν “αγωγό” δεδομένων (data pipeline) κατά τον οποίο θα εισέρχεται ακατέργαστη πληροφορία από πολλές προελεύσεις όπου με τα κατάλληλα εργαλεία και τεχνολογίες που επιλέχθηκαν από τους τεχνικούς, θα μετατρέπεται σε χρήσιμα προϊόντα δεδομένων τα οποία θα διαχέονται στο σύνολο του

οργανισμού.[3] Τα προϊόντα αυτά πρέπει να διατηρούνται, με εύκολη προσβασιμότητα προς τα τμήματα που τα έχουν ανάγκη και διαχειρίσιμα ακόμη και από άτομα χωρίς υπόβαθρο στην επιστήμη δεδομένων. Για αυτό το σκοπό απαιτείται η δημιουργία κοινών πρωτοκόλλων επικοινωνίας για επίτευξη διαλειτουργικότητας σε τεχνικό, οργανωτικό και σημασιολογικό επίπεδο ανάμεσα στα τμήματα για καλύτερη ροή και συνέπεια των δεδομένων από την πρωταρχική τους μορφή σε τελική χρήσιμη πληροφορία ώστε να μπορεί να επιτευχθεί μία αποδοτική κλιμάκωση. Η κλιμάκωση μπορεί να είναι είτε οριζόντια όπου γίνεται προσθήκη επιπλέον υπολογιστών και κόμβων πληροφορίας, είτε καθετοποιημένη όπου ενισχύονται συγκεκριμένοι υπολογιστές με μεγαλύτερης δυνατότητας επεξεργασίας και μνήμης. Συνήθως, σε μία επιχείρηση όπου υπάρχουν κατανεμημένα συστήματα με διαφορετική δυναμική κάθε κόμβου, η κλιμάκωση επιτυγχάνεται οριζόντια, αυξάνοντας ωστόσο την πολυπλοκότητα του δικτύου. Άρα η δημιουργία ενός διαλειτουργικού υπολογιστικού και ψηφιακού περιβάλλοντος κατά την ροή δεδομένων είναι απαραίτητη για αποφόρτιση της πολυπλοκότητας του έργου αυτού.

Επιπρόσθετα, η επιλογή κατάλληλων ψηφιακών εργαλείων εκτελείται από τα μέλη αυτού του επιπέδου, αφού λάβουν κατευθυντήριες οδηγίες του προϊσταμένου της ομάδας διαχείρισης δεδομένων για τις ανάγκες της ομάδας αυτής και ολόκληρου του οργανισμού. Η ευθύνη τους αφορά, λοιπόν, από την σχεδίαση βάσεων δεδομένων, δηλαδή αν είναι σχεσιακή(SQL) ή όχι, την συνδεσιμότητα αυτής με τα υπόλοιπα πληροφοριακά συστήματα της επιχείρησης (π.χ. ERP systems) ή με κάποιου υπολογιστικού νέφους (cloud) και επεκτείνονται ως την επιλογή κατάλληλων λογισμικών για επεξεργασία και οπτικοποίηση τα οποία συμβαδίζουν με τις ανάγκες και δυνατότητες του παρόντος οργανισμού.



Εικόνα 20. Θεωρητικό μοντέλο αγωγού δεδομένων (data pipeline) Source: Pentaho Corporation

Οι εργασίες σε Τεχνικό επίπεδο είναι πιο έντονες κατά την ίδρυση της ομάδας δεδομένων όπου μπαίνουν οι ψηφιακές βάσεις για την ομαλή λειτουργίας της. Συγκεκριμένα, το τμήμα αυτό είναι υπεύθυνο να σχεδιάσουν προγραμματιστικά οι βάσεις δεδομένων σύμφωνα με τις ανάγκες αλλά και τις οικονομικές δυνατότητες που υπάρχουν, υποστηρίζοντας αυτές με τον κατάλληλο υλικοτεχνικό εξοπλισμό, με ταυτόχρονη σύνδεση με τις πηγές δεδομένων[29]. Ανάλογα με το τρόπο που σχεδιάστηκαν και της προγραμματιστικής γλώσσας που χρησιμοποιήθηκε, επιλέγεται το κατάλληλο σύστημα διαχείρισής τους και εισάγονται μέσω αυτού τα υπάρχοντα δεδομένα στις βάσεις, έτοιμα για ανάλυση ανατροφοδοτώντας την συνεχώς με νέα. Επίσης, πραγματοποιείται διασύνδεση με υπόλοιπα λογισμικά που χρησιμοποιούνται κατά την λειτουργία της εταιρείας και αντιμετωπίζονται τα πιθανά προβλήματα που προκύπτουν στην αρχή όπως ασυνέπεια των εισαχθέντων δεδομένων (π.χ. να μην καταχωρούνται όλα τα δεδομένα σε μία κοινή μορφή ή τύπο, διπλοεγγραφές, λανθασμένες ή κενές εγγραφές κ.τ.λ.), μη συνδεσιμότητα διαφορετικών τεχνολογιών ή ακόμη και αδυναμία υποστήριξης των ψηφιακών διεργασιών με τις διαθέσιμες αποθηκευτικές και επεξεργαστικές δυνατότητες.

Το προσωπικό αυτό διαφέρει, διοικητικά και λειτουργικά από το εκάστοτε υποστηρικτικό τμήμα τεχνολογίας και πληροφορικής του κάθε οργανισμού ωστόσο μαζί συνεργάζονται στενά κατά την εγκατάσταση και σύνδεση του αγωγού δεδομένων στις υπόλοιπες διαδικασίες της επιχείρησης και αναλαμβάνουν από κοινού ενέργειες σε περίπτωση που δημιουργηθούν προβλήματα κακής συνεργασίας των τμημάτων εξαιτίας τεχνικών ασυνεπειών. Γενικότερα, λόγω των κατανεμημένων συστημάτων που απαιτεί μία επιχείρηση και των δυσκολιών που δημιουργούν, η προσέγγιση σε αυτό το επίπεδο είναι η εκτέλεση πολλών διασταυρωμένων εργασιών με χρήση διαφορετικών πρακτικών και η προσαρμογή του προγραμματιστικού κώδικα στις απαιτήσεις αυτές[29]. Στην συνέχεια, το τμήμα αυτό επεκτείνεται στο ρόλο της παρακολούθησης, συντήρησης και αναβάθμισης των συστημάτων, προτείνοντας παράλληλα βελτιώσεις και αλλαγές βάση των συνεχών τεχνολογικών εξελίξεων, και αν απαιτηθεί να τις φέρει εις πέρας. Υποστηρίζουν γνωστικά σε εξειδικευμένα αντικείμενα, τα υπόλοιπα μέλη της ομάδας διαχείρισης δεδομένων και προσπαθούν να προλαμβάνουν αδυναμίες διαχείρισης μεγάλου όγκου δεδομένων, να διευρύνουν τις δυνατότητες της εταιρείας βάση του διατιθέμενου προϋπολογισμού καθώς ορισμένες βελτιώσεις μπορεί να μην είναι οικονομικά ανεκτές και υιοθετείτε μία σχέση κόστους ωφέλειας η οποία μεταφέρεται στο επόμενο επίπεδο καθώς και στην διοίκηση[34]. Τέλος, να διορθώνουν και να εφαρμόζουν τυχόν παρατηρήσεις από τον προϊστάμενο της ομάδας αλλά και της κεντρικής διοίκησης.

Ορισμένοι ρόλων που δύναται να υπάρξουν στην υπο-ομάδα αυτή ανάλογα με τις δυνατότητες και απαιτήσεις της εταιρείας, διαφέρουν από άλλες ομάδες τεχνολογίας και είναι οι παρακάτω[29]:

- Μηχανικός Δεδομένων (Data Engineer), ο πιο κρίσιμος ρόλος και σε περίπτωση που οι τεχνικές εργασίες περιοριστούν σε λίγα ή ακόμη και σε ένα μόνο άτομο, κατά πάσα πιθανότητα το άτομο αυτό θα διαθέτει αυτόν τον εργασιακό τίτλο. Πρέπει να διαθέτει προγραμματιστικές γνώσεις υψηλού επιπέδου, με σπουδές στην επιστήμη των υπολογιστών ή παρεμφερές καθώς και να μπορεί να διαχειρίζεται συστήματα υψηλής κλίμακας επεξεργασίας. Γενικότερα, χρειάζεται να διαθέτει τεχνολογικές γνώσεις σε όλο το εύρος της επιστήμης δεδομένων[4].
- Μηχανικός Αποθηκών Δεδομένων (Data Warehouse Engineer), εξειδικεύεται στην κατασκευή και διαχείριση βάσεων δεδομένων, προσεγγίζοντας με ακρίβεια

ζητήματα αποθήκευσης, συντήρησης και κλιμάκωσης δεδομένων, ιδιαίτερα κρίσιμο όταν ένας φορέας διαχειρίζεται Μεγάλα Δεδομένα (Big Data).

- Αρχιτέκτονας Δεδομένων (Data Architect), ένας ρόλος οποίος συναντάται κυρίως σε ένα απαιτητικό περιβάλλον με έντονα κατανεμημένα συστήματα, όπου τα δεδομένα προέρχονται από πολλές πηγές και χρειάζεται μια αποτελεσματική σχεδίαση ενός δικτύου πληροφοριών με πολλές διασταυρώσεις. Συντονίζει πειθαρχημένα τον αγωγό δεδομένων και τα σημεία σύνδεσης του με τα υπόλοιπα πληροφοριακά συστήματα της επιχείρησης. Με τη μεγάλη εμπειρία που θα διαθέτει από προγενέστερες θέσεις πληροφοριακών τμημάτων και ομάδων λογισμικών, προσπαθεί να σχεδιάσει μια αποτελεσματική στρατηγική ανάπτυξη του δικτύου δεδομένων υιοθετώντας μία σφαιρική αντίληψη.
- Λοιποί προγραμματιστές, μηχανικοί και τεχνικοί υπολογιστών ή λογισμικού, μπορούν να υποστηρίξουν το έργο των ανωτέρω ρόλων, οι οποίοι μπορεί να μην ανήκουν στην υπο-ομάδα αυτή άμεσα αλλά στο τμήμα τεχνολογίας και πληροφορικής της εταιρείας ή να είναι εξωτερικοί συνεργάτες. Γενικά αναλόγως των συνθηκών μπορεί να απαιτηθεί συμπληρωματική υποστήριξη.

2.1.2. Λειτουργικό επίπεδο

Είναι ο κεντρικός πυλώνας της ομάδας διαχείρισης δεδομένων και από εδώ ασκείται ο έλεγχος της. Αποστολή σε αυτό το επίπεδο είναι η παραγωγή των προϊόντων δεδομένων και των στατιστικών διεργασιών, εκμεταλλευόμενη το τεχνολογικό έργο της προηγούμενης υπο-ομάδας. Ανάλογα του βαθμού διείσδυσης της μηχανικής μάθησης στις διαδικασίες και του αυτοματισμού τους, διαμορφώνεται και το απαιτούμενο λειτουργικό έργο που πρέπει να φέρει εις πέρα. Είναι το κεντρικό κομμάτι της ομάδας και σε περίπτωση αδυναμίας οικοδόμησης τριών υπο-ομάδων με διαφορετικούς εργαζόμενους, τότε όλο το έργο διαχείρισης δεδομένων θα συσπειρωθεί σε αυτή, διατηρώντας ωστόσο τα τρία διακριτά στάδια εργασιών καθώς είναι πολύ σημαντική η κατανόηση του κύκλου εργασιών. Εργάζεται με μεθοδικότητα ανάμεσα στα άλλα δύο τμήμα επικοινωνώντας μαζί τους ώστε να καταλάβουν τους στόχους όλης της ομάδας. Για παράδειγμα οι μηχανικοί δεδομένων

θα συνθέσουν ένα πιο στοχευμένο μείγμα λογισμικών και τεχνολογιών αν ξέρουν ακριβώς ποιοι και πως θα το χρησιμοποιούν, ενώ οι αναλυτές της επόμενης ομάδας θα μπορέσουν πιο εξειδικευμένα να αναλύσουν και να οπτικοποιήσουν τα προϊόντα δεδομένων γνωρίζοντας σε τι αποσκοπούν. Επιβλέπει ολόκληρο το έργο διαχείρισης δεδομένων, λαμβάνοντας υπόψη τις απόψεις και προτάσεις από όλα τα μέλη της ομάδας, επιστρέφοντας τους κατευθυντήριες οδηγίες για τις καθημερινές εργασίες καθώς και τα σχέδια μακροπρόθεσμης ανάπτυξης[29]. Συνομιλεί με την διοίκηση για τις ανάγκες πληροφόρησης ολόκληρου του οργανισμού, λαμβάνει οδηγίες για τα επόμενα επιχειρησιακά πλάνα και πως γίνεται αυτά να υποστηριχθούν με την κατάλληλη χρήση δεδομένων. Γενικότερα, επιδιώκει να βρει δυνατότητες περαιτέρω χρήσης των δεδομένων προς υποστήριξη του συνολικού επιχειρησιακού έργου, εξασφαλίζοντας υψηλότερη αποδοτικότητα όλων των πόρων και προσπαθεί να εκμεταλλευτεί τυχόν ευκαιρίες που θα παρουσιαστούν στο τομέα της επιστήμης δεδομένων.

Εξαρτάται σε μεγάλο βαθμό κατά πόσο στο προηγούμενο στάδιο επιτευχθεί αυτοματοποίηση στην παραγωγή αποτελεσμάτων δηλαδή αν δεν παράγονται αυτόματα τα προϊόντα δεδομένων, τότε με χρήση στατιστικής ανάλυσης, μαθηματικών και των διαθέσιμων ψηφιακών εργαλείων θα προσπαθήσει αυτή η υπο-ομάδα να δώσει την τελική μορφή στα δεδομένα, η οποία πρέπει να είναι εύκολα κατανοητή και αξιοποιήσιμη για τις λοιπές διεργασίες της εταιρείας, δημιουργώντας ένα μοντέλο μηχανικής μάθησης. Αντίθετα, αν ο αγωγός δεδομένων παράγει αυτόματα τα απαιτούμενα αποτελέσματα, π.χ. μέσω της τεχνολογίας της Τεχνητής Νοημοσύνης, τότε αξιολογείται συνεχώς το τεχνητό μοντέλο αυτό για μείωση αποκλίσεων με σκοπό την ενίσχυση της αξιοπιστίας των παραγόμενων. Παρακολουθεί την εισροή δεδομένων από όλες τις πηγές και ελέγχει για τυχόν περίεργες τιμές (π.χ. ακραίες ή ασυνήθιστες σταθερές), τις οποίες οι αλγόριθμοι δεν καταφέρουν να εντοπίσουν, αναζητεί νέες πηγές ώστε να ανατροφοδοτείται η βάση με περισσότερα δεδομένα, επεξεργάζεται τα στοιχεία που προκύπτουν, προσπαθώντας να τα ταξινομήσει κατάλληλα στις βάσεις δεδομένων αν δεν γίνεται αυτόματα, ενώ αν γίνεται να ελέγχει τις κλάσεις για τυχόν σφάλματα.

Συμμετέχει, ακόμη, στο διοικητικό έργο της επιχείρησης ώστε να γνωρίζει τα μελλοντικά πλάνα, συνομιλεί με ανώτερα στελέχη για να αντιληφθεί τυχόν προβλήματα και να ενημερωθεί για το δικό τους έργο, όπου συνδυάζοντας το δικό τους γνωστικό

αντικείμενο με την επιστήμη δεδομένων να δημιουργηθούν οι προϋποθέσεις για περαιτέρω αναβάθμισης της υποστήριξης των εταιρικών εργασιών, μειώνοντας έτσι την απαιτούμενη ανθρώπινη προσπάθεια και τα κόστη. Σε περίπτωση αδυναμίας ή υποστελέχωσης από τα άλλα δύο μέρη, ενισχύει το έργο τους και επεμβαίνει αν απαιτηθεί, και γι' αυτό είναι και απαραίτητη η γνωστική επαφή κατά μήκους της επιστήμης δεδομένων, από τα άτομα που εργάζονται σε αυτό για να έχουν μια συνδυαστική αντίληψη των ζητημάτων κατά την διαδικασία επεξεργασίας δεδομένων.

Στην υπο-ομάδα αυτή ανήκουν εξειδικευμένοι επαγγελματίες της επιστήμης δεδομένων έχοντας λάβει και την αντίστοιχη εκπαίδευση. Έτσι οι ρόλοι που απαρτίζουν το κομμάτι αυτό είναι συγκεκριμένοι αν και προσαρμόζονται στις απαιτήσεις της εταιρείας και στις συνεχείς εξελίξεις[29]:

- **Επιστήμονας Δεδομένων (Data Scientist)**, προσπαθεί να απαντήσει στα ερωτήματα που του θέτουν μέσω των διαθέσιμων δεδομένων, να δώσει μακροπρόθεσμες και βιώσιμες λύσεις σε προβλήματα μέσω μοντέλων μηχανικής μάθησης που έχει σχεδιάσει και να συντονίσει όλη την διαδικασία επεξεργασίας δεδομένων. Είναι ίσως ο πιο γνωστός ρόλος στις επιχειρήσεις σε μία ομάδα δεδομένων και από αυτόν ξεκινάει η σύσταση και η οργάνωσή της. Μπορεί να εκτελέσει και συντονιστικό ρόλο αν απαιτηθεί[4].
- **Προϊστάμενος Ομάδας Διαχείρισης Δεδομένων (Data Project Manager)**, συναντάται συνήθως σε μεγάλους οργανισμούς όπου η διαδικασία επεξεργασίας δεδομένων έχει ενσωματωθεί πλήρως στην ομαλή λειτουργία της επιχείρησης η οποία λειτουργεί με γνώμονα τα δεδομένα (Data-Driven Company). Ουσιαστικά, διαθέτει τις ικανότητες ενός Επιστήμονα Δεδομένων και είναι επιπλέον επιφορτισμένος με τον έλεγχο και διαχείριση των μελών της ομάδας δεδομένων.
- **Προϊστάμενος Προϊόντων Δεδομένων (Data Product Manager)**, και αυτός ο ρόλος συναντάται συνήθως σε μεγάλες επιχειρήσεις που λειτουργούν με βάση τα δεδομένα και είναι ιδιαίτερα κρίσιμος και αναγκαίος όταν οι παραγόμενες υπηρεσίες ή προϊόντα είναι προϊόντα δεδομένων. Για παράδειγμα, συμβουλευτικές επιχειρήσεις που υποστηρίζουν άλλους οργανισμούς κατά την λήψη αποφάσεων με μεγάλες στατιστικές αναλύσεις και μελέτες ή δημιουργοί εφαρμογών βασιζόμενες στη λήψη δεδομένων και η λειτουργία τους επιτυγχάνεται μέσω ενός

μοντέλου μηχανικής μάθησης. Ο κύριος σκοπός είναι η γεφύρωση των κατασκευαστών των προϊόντων αυτών με τις ανάγκες των στελεχών της εταιρείας ή των πελατών. Ιδιαίτερα για τους πελάτες που θα καταναλώσουν τα προϊόντα μία σημαντική ειδοποιός διαφορά είναι, όταν τα προϊόντα είναι για μαζική κατανάλωση, όπως μία εφαρμογή σε έξυπνο κινητό τηλέφωνο είτε είναι κατά παραγγελία ενός έργου, όπως μίας μεγάλης στατιστικής μελέτης, και η ανάθεση του γίνεται με πολύ συγκεκριμένες απαιτήσεις. Στην πρώτη περίπτωση, προσπαθεί επεξεργαστεί τις κριτικές από διάφορες πλατφόρμες ή δίκτυα κοινωνικής δικτύωσης, ως δεδομένα και να εφαρμόσει κάποιες πρακτικές όπως αυτής της Βαθιάς Μάθησης (Deep Learning) για να αναγνωρίσει εύκολα τις αρνητικές από τις θετικές αντιδράσεις ενώ στην άλλη περίπτωση η επικοινωνία μπορεί είναι άμεση με πολλοί εξειδικευμένους όρους.

2.1.3. Συνολικό επίπεδο και εξαγωγή αποτελεσμάτων

Το τελευταίο κομμάτι της αλυσίδας στο οποίο καταλήγει το έργο των προηγούμενων δύο τμημάτων και οι υπάλληλοι σε αυτό το στάδιο εργάζονται σύμφωνα με τις κατευθυντήριες οδηγίες που έχουν λάβει από τον προϊστάμενο της ομάδας και την κεντρική διοίκηση, χρησιμοποιώντας τα ψηφιακά εργαλεία που τους προσφέρει το τεχνικό τμήμα, εφαρμόζοντας παράλληλα πρακτικές Επιχειρηματικής Ευφυΐας. Έχουν πιο περιορισμένο έργο ως προς το δημιουργικό κομμάτι και επικεντρώνονται περισσότερο στο ερμηνευτικό και εκτελεστικό.[29] Υποστηρίζουν τα υπόλοιπα μέλη της ομάδας στην επεξεργασία δεδομένων και επικοινωνούν τα αποτελέσματα που προκύπτουν στον υπόλοιπο οργανισμό ώστε να είναι πλήρως κατανοητά από τα στελέχη της εταιρείας. Παρότι δεν είναι τόσο παραγωγικό κομμάτι όσο τα υπόλοιπα της ομάδας, αποτελεί κρίσιμο σημείο καθώς είναι αυτό που θα προσθέσει την επιπρόσθετη αξία στα προϊόντα δεδομένων, καθώς αν δεν μπορούν να αξιολογηθούν από τον υπόλοιπο οργανισμό η συνολική προσπάθεια θα είναι μάταια. Ακόμη, εκτελούν πρωτογενή εργασία, από την εισαγωγή των δεδομένων μέχρι το τέλος της διαδικασίας, έχοντας την δυνατότητα να εξηγήσουν σε έναν εξωτερικό παρατηρητή ολόκληρη την διαδικασία με συνδυαστικό και αναλυτικό τρόπο

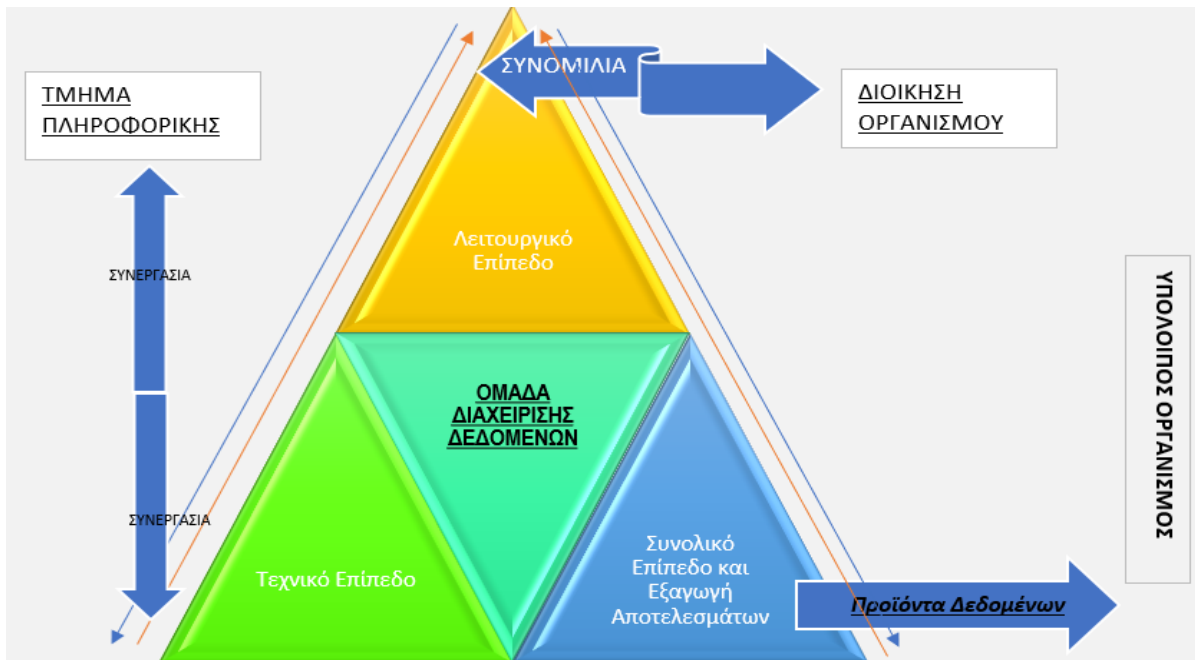
σκέψης καθώς μπορούν να εργαστούν συμπληρωματικά σε όλο το εύρος των εργασιών της ομάδας από την αρχή μέχρι το τέλος.

Για τον λόγο ότι είναι ένα υποστηρικτικό και εκτελεστικό τμήμα οι λειτουργίες του εξαρτώνται αρκετά από το έργο των υπόλοιπων υπο-ομάδων και από τις ανάγκες των άλλων τμημάτων της επιχείρησης, γι' αυτό εδώ οι εργασιακοί ρόλοι μπορεί να ποικίλουν. Ωστόσο, μερικές βασικές λειτουργίες είναι η ανάλυση και οπτικοποίηση δεδομένων με κατάλληλα ψηφιακά εργαλεία επιχειρηματικής ευφυΐας, τα οποία συνήθως είναι εξατομικευμένα για συγκεκριμένες εργασίες. Πιο συγκεκριμένα, οι υπάλληλοι χειρίζονται κατάλληλα το σύστημα διαχείρισης βάσης δεδομένων μπορούν να εξάγουν τα στοιχεία που τους είναι απαραίτητα ώστε να συντάξουν μια αναλυτική αναφορά σε φυσική γλώσσα ή να τα μεταφέρουν σε ευανάγνωστη μορφή σε υπολογιστικά φύλλα για πιο εύκολη επεξεργασία και οργάνωση από έτερα άτομα.

Επιπρόσθετα, με την χρήση κατάλληλων λογισμικών οπτικοποίησης μπορούν να τους δώσουν μία πιο διαδραστική απεικόνιση και να γίνονται εύκολα αντιληπτά τα ευρήματα των διάφορων αναλύσεων[4]. Οι αναλύσεις αυτές πρέπει να συνοδεύονται από επεξηγήσεις πως προέκυψαν και αν απαιτηθεί να τις παρουσιάσουν και σε άλλα άτομα, γι' αυτό είναι αναγκαίο να γνωρίζουν πως προέκυψαν τα δεδομένα, έστω και επιφανειακά, για να μπορούν να επικοινωνήσουν σωστά τα μηνύματα που επιθυμούν σε πραγματικό χρόνο. Επίσης, μπορεί να υποστηρίξουν και σε άλλα στάδια τον αγωγό δεδομένων, π.χ. εισάγοντας χειρωνακτικά τα δεδομένα μέσω ενός κατάλληλου περιβάλλοντος πλοήγησης, ειδικά στην έναρξη λειτουργίας της ομάδας δεδομένων, εισάγοντας ιστορικά στοιχεία αλλά και κατά την διάρκεια καθώς κάποιες πηγές δεδομένων δεν θα είναι διασυνδεδεμένες ψηφιακά με τις βάσεις. Τα δεδομένα πρέπει να έχουν υποστεί την κατάλληλη επεξεργασία από την Λειτουργική Ομάδα της επιστήμης δεδομένων, απομακρύνοντας πιθανούς “θορύβους” όπως ακραίες τιμές ή εποχικότητα σε μετρήσεις τιμών. Επιπρόσθετα, μπορεί να συνεργαστούν με ένα επιστήμονα δεδομένων από το προηγούμενο στάδιο ώστε να ερμηνεύσουν πιθανά επαναλαμβανόμενα μοτίβα ή σχέσεις μεταξύ των δεδομένων, χαρακτηρίζοντας αυτές κατάλληλα.

Αν και το είδος των εργαζομένων σε αυτό το επίπεδο μπορεί να ποικίλει αρκετά σε κάθε φορέα, κάποιοι ενδεικτικοί τίτλοι εργασίας είναι οι κάτωθι[29]:

- Αναλυτής Δεδομένων (Data Analyst), ικανός να αναλύει μεγάλες ποσότητες δεδομένων που εισέρχονται και εξέρχονται από εκεί που είναι αποθηκευμένα ώστε να εξάγει χρήσιμες πληροφορίες για την υποστήριξη του συνολικού εταιρικού έργου. Με συνδυαστική και στατιστική προσέγγιση καταλήγει σε συμπεράσματα τα οποία είναι σε θέση να εξηγήσει και να παρουσιάσει εύκολα σε άλλα άτομα μέσω διαγραμμάτων και πινάκων[4].
- Αναλυτής Επιχειρηματικής Ευφυΐας (BI Analyst), καλείται να χειρίζεται με αποδοτικό τρόπο τα διάφορα λογισμικά και ψηφιακές υπηρεσίες που έχει στην διάθεσή του ώστε να εκτελέσει το έργο του Αναλυτή Δεδομένων με ακόμη πιο “έξυπνο τρόπο”. Προτείνει νέα πιθανά τεχνολογικά εργαλεία τα οποία μπορούν να διαμορφώσουν πιο κατανοητά αποτελέσματα και να διακρίνουν πιθανά μοτίβα και συνδέσεις μεταξύ ορισμένων παραγόντων. Ειδικά όταν η εταιρεία πρέπει να διαχειριστεί Μεγάλα Δεδομένα απαιτείται η υιοθέτηση της προσέγγισης της επιχειρηματικής ευφυΐας.
- Υπάλληλος Καταχώρησης Δεδομένων (Data Entry Clerk), έντονα τα καθήκοντα του στην αρχή καθώς εισάγονται τα πρώτα παρελθοντικά στοιχεία και στην συνέχεια ανατροφοδοτώντας με νέα από τις πηγές όπου δεν γίνεται αυτόματα ή έστω με μία απλή φόρτωση δεδομένων από λογισμικό σε λογισμικό. Είναι σημαντικό να είναι εξοικειωμένος με το περιβάλλον πλοήγησης του συστήματος διαχείρισης βάσεων δεδομένων που έχει υιοθετήσει η εταιρεία καθώς είναι σημαντικό να εισάγονται με απόλυτη συνέπεια τα δεδομένα. Επομένως, η διαδικασία καταχώρησης πρέπει να αντιμετωπίζεται με μεθοδική και οργανωσιακή προσέγγιση για ταχύτερη και αξιόπιστη αποθήκευση δεδομένων. Μεμονωμένα, ίσως, κάθε τμήμα μπορεί να προσπαθήσει να εισάγει τα δεδομένα όταν αυτά είναι απλές τιμές στο πληροφοριακό σύστημα ωστόσο τίθεται κίνδυνος ασυνέπειας ως προς τον τρόπο γραφής ή διπλοεγγραφών, και η έλλειψη ελέγχου για τις αλλαγές που πραγματοποιούνται με πολύ αρνητικές συνέπειες. Γι’ αυτό πολλές εταιρείες αναθέτουν την καταχώρηση σε έμπειρα άτομα στην διαδικασία αυτή τα οποία δημιουργούν ταυτόχρονα και αντίγραφα της βάση για μεγαλύτερη ασφάλεια.



Εικόνα 21. Θεωρητική οπτική προσέγγιση της λειτουργίας μίας ομάδας δεδομένων εντός τους οργανισμού. Το Λειτουργικό Επίπεδο υποστηρίζεται από τα άλλα δύο μέρη και αποστέλλει κατευθύνσεις.

2.2. Κριτήρια σχεδιασμού ομάδων διαχείρισης δεδομένων

Έχοντας προσεγγίσει τη δομή μίας ομάδας διαχείρισης δεδομένων εντός μίας επιχείρησης, τις λειτουργίες της, τους διάφορους ρόλους των εργαζομένων εντός αυτής και την αλληλοεπίδραση με το υπόλοιπο εταιρικό περιβάλλον είναι σημαντικό να εξηγηθεί με ποια κριτήρια διαμορφώνονται όλα αυτά. Στον παρακάτω πίνακα αποτυπώνονται περιληπτικά ορισμένοι παράγοντες οι οποίοι επηρεάζουν το σχεδιασμό και την στρατηγική της ομάδας.

<u>ΠΑΡΑΓΟΝΤΕΣ</u>	<u>ΠΕΡΙΓΡΑΦΗ</u>
Τα μεγέθη και ο βαθμός αυτονομίας	Το μέγεθος της επιχείρησης, του κλάδου που ανήκει εντός της γενικής οικονομικής δραστηριότητας αλλά και της ίδιας της ομάδας εντός της επιχείρησης και κατά πόσο αυτόνομα λειτουργεί.
Υποβολή αποτελεσμάτων	Τύπος αποτελεσμάτων, πως υποβάλλονται, που

	απευθύνονται, ποιοι τα χρησιμοποιούν και σε τι αποσκοπούν.
Πηγές δεδομένων	Τα είδη των πηγών, η ποσότητα, η ποικιλομορφία, η ποιότητα και πως εισέρχονται στις βάσεις δεδομένων.
Το επίπεδο αξιοποίησης των δεδομένων από την επιχείρηση	Σε τι βαθμό συμμετέχουν στην παραγωγική διαδικασία της εταιρείας και τον βαθμό εκμετάλλευσής τους.
Εργαλεία διαχείρισης δεδομένων	Ποιες τεχνολογίες είναι διαθέσιμες, πως χρησιμοποιούνται και πως συνδέονται όλα αυτά.
Οικονομικές δυνατότητες	Ο διαθέσιμος προϋπολογισμός και οι δυνατότητες επέκτασης, οι διαθέσιμοι πόροι.
Το προσωπικό της	Οι ικανότητες του προσωπικού της ομάδας καθώς και αυτού που θα αλληλοεπιδρούν με αυτήν. Αν εργάζονται δια ζώσης ή εξ αποστάσεως,
Στρατηγική και όραμα	Η γενική επιχειρησιακή στρατηγική, οι πολιτικές και η νοοτροπία ανάμεσα στους εργαζόμενους καθώς και το όραμα για την λειτουργία της εταιρείας και της ίδιας της ομάδας.
Το εξωτερικό περιβάλλον της εταιρείας	Ο κλάδος που ανήκει, οι ανταγωνιστές της, οι πελάτες της και γενικά οι οργανισμοί που αλληλοεπιδρά
Χαρακτηριστικά της περιοχής	Τεχνολογικές δυνατότητες (δίκτυα τηλεπικοινωνιών, ταχύτητες διαδικτύου), το τεχνολογικό επίπεδο του πληθυσμού της περιοχής για εύρεση νέων εργαζόμενων ή εξωτερικών συνεργατών, πιθανές τοπικές συνεργασίες τμημάτων διαχείρισης δεδομένων

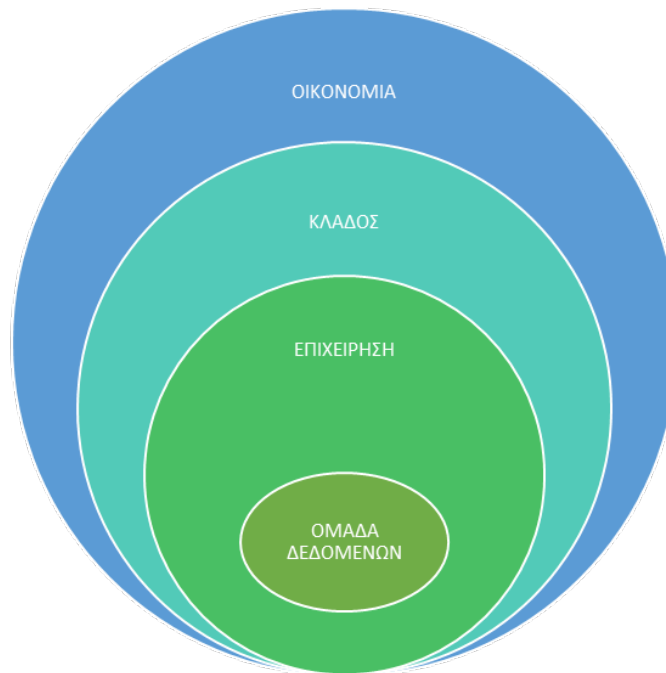
Από τα ανωτέρω κριτήρια θα αναλυθούν τα πρώτα τέσσερα τα οποία είναι περισσότερο εστιασμένα στο αντικείμενο της συγκεκριμένης ομάδας σε σχέση με οποιαδήποτε άλλη ομάδα εργασίας και διαδραματίζουν καθοριστικό ρόλο για την δόμηση και στελέχωση του τμήματος.

2.2.1. Τα μεγέθη και ο βαθμός αυτονομίας

Το μέγεθος του κλάδου που συμμετέχει μια επιχείρηση και η θέση της μέσα σε αυτόν θα οριοθετήσει την δυνατότητα επένδυσης σε καινοτόμες ιδέες ιδιαίτερα όταν ο ανταγωνισμός είναι μεγάλος και κάθε εταιρεία θέλει να αποκτήσει ανταγωνιστικό πλεονέκτημα στον κλάδο. Για παράδειγμα μία νέα υγιής αλυσίδα τουριστικών καταστημάτων (ξενοδοχεία, κέντρα εστίασης και διασκέδασης κ.τ.λ.) στην Ελλάδα, θα έχει τις δυνατότητες για μία γρήγορη και μεγάλη πρόοδο, παρότι του έντονου ανταγωνισμού που θα αντιμετωπίσει, καθώς ο κλάδος του Τουρισμού ήταν συνεχώς ανοδικός τα τελευταία χρόνια μέχρι την έναρξη της πανδημίας COVID-19, σύμφωνα με τα στοιχεία του Συνδέσμου Ελληνικών Τουριστικών Επιχειρήσεων[30] όπου το 2019 το ποσοστό συμμετοχής στο Α.Ε.Π. της Ελλάδας ήταν 20,8% με συνολικά έσοδα 17,7 δις εκατομμύρια ευρώ με συνεχή άνοδο σε σχέση με τα προηγούμενα χρόνια. Έτσι, η νέα εταιρεία για να καταφέρει να εισέλθει στο χώρο και να αποκτήσει σημαντικό μερίδιο αγοράς εις βάρος, ουσιαστικά, πιο έμπειρων επιχειρήσεων με δημοφιλή εμπορικά σήματα χρειάζεται να στραφεί στην καινοτομία και να εκμεταλλευτεί τα δεδομένα στο έπακρο ώστε να προσφέρει το ένα πακέτο με καλύτερη σχέση τιμής-ποιότητας.

Σε μία τέτοια περίπτωση, γνωρίζει η διοίκηση της εταιρείας ότι θα γίνει απόσβεση της επένδυσης επειδή ο κλάδος είναι μεγάλος και αναπτύσσεται ετησίως. Έτσι, θα υπάρχουν τα περιθώρια οικοδόμησης μίας μεγάλης ομάδας διαχείρισης με διαφορετικούς ρόλους εργαζομένων ώστε να επιτευχθεί μεγαλύτερη εκμετάλλευση των δεδομένων, δημιουργώντας οικονομίες κλίμακας, είτε καθετοποιημένες είτε οριζόντιες, για περαιτέρω αναπτυξιακές δυνατότητες της επιχείρησης. Στο συγκριμένο παράδειγμα της τουριστικής εταιρείας, οι εργασίες μίας τέτοιας θα μπορούσαν να αποσκοπούσαν στην αποτελεσματικότερη αναπλήρωση αποθεμάτων, αποδοτικότερη διαχείριση πόρων και

κυρίως των εργαζομένων, διαχείριση αξιολογήσεων πελατών και ταχύτερο εντοπισμό τυχόν δυσαρέσκειας από πελάτες, πιο έγκυρες προβλέψεις ζήτησης ανά κατάσταση, πιο εξατομικευμένες διαφημιστικές εκστρατείες και γενικότερα υποστήριξη αποφάσεων και σχεδίαση μελλοντικών σχεδίων. Αντίθετα, αν ο κλάδος ήταν μικρός σε αποτίμηση οικονομικών μεγεθών τότε η επένδυση σε τμήματα τα οποία δεν αφορούν άμεσα την κύρια παραγωγική διαδικασία, θα βασιζόνταν από το πρόσημο της ετήσιας ανάπτυξής του, δηλαδή αν το πρόσημο είναι θετικό και έχει προοπτικές ανάπτυξης είναι πιο εύκολο η διοίκηση να στηρίζει ένα τέτοιο υποστηρικτικό τμήμα σε μεγάλο βαθμό.



Εικόνα 22. Οπτικοποίηση των ανωτέρω μεγεθών

Η αυτονομία του τμήματος από την άλλη μπορεί να είναι ακόμη πιο σημαντική από το ίδιο το μέγεθος της ομάδας. Συγκεκριμένα, μια εταιρεία μπορεί να έχει αντιληφθεί την σημασία να έχει μία αξιόπιστη πληροφόρηση ωστόσο να μην δημιουργήσει ένα αυτόνομο τμήμα για το αντικείμενο αυτό και να λειτουργεί συμπληρωματικά στο έργο του κάθε άλλου τμήματος (παραγωγής, λογιστηρίου, logistics) προσλαμβάνοντας αρκετούς επαγγελματίες της επιστήμης δεδομένων. Αυτή η προσέγγιση διαθέτει ως θετικά την πιο άμεση και προσαρμοσμένη ένταξη των δεδομένων στην εκάστοτε επιχειρησιακή λειτουργία, δημιουργώντας ένα αποκεντρωμένο δίκτυο επεξεργασίας των δεδομένων υποστηριζόμενο από το τμήμα πληροφοριών και τεχνολογίας του οργανισμού[29], αλλά

ίσως να μην προσεγγίσει τόσο εξειδικευμένα το θέμα αυτό αφήνοντας αρκετά περιθώρια βελτίωσης.

Από την άλλη ακόμη και αν το τμήμα είναι ξεχωριστό από τα υπόλοιπα, δεν σημαίνει από μόνο του ότι διαθέτει μεγάλο βαθμό αυτονομίας. Ανάλογα και με την εσωτερική πολιτική της εταιρείας περί της αυτονομίας των τμημάτων, το τμήμα μπορεί να είναι ικανοποιητικά μεγάλο ωστόσο ο έλεγχος να ασκείται εκτός αυτού και να προέρχεται για παράδειγμα από την διοίκηση, αφαιρώντας έτσι ρόλους πιο ηγετικούς όπως του Προϊσταμένου της Ομάδας Διαχείρισης Δεδομένων. Γενικότερα για να χαρακτηριστεί μία ομάδα ότι λειτουργεί πλήρως αυτόνομα, είναι αναγκαίο η ύπαρξη ενός που θα επιβλέπει το συνολικό έργο, θα ασκεί έλεγχο και θα συνομιλεί απευθείας με την διοίκηση. Μία τέτοια προσέγγιση αποτυπώνεται στην παρακάτω εικόνα.



Εικόνα 23. Θεωρητική προσέγγιση της ιεραρχίας μίας αυτόνομης ομάδας διαχείρισης δεδομένων

2.2.2. Υποβολή αποτελεσμάτων

Αρχικά είναι σημαντικό να προσδιοριστεί η πιθανή μορφή των προϊόντων δεδομένων κατά το στάδιο εξαγωγής από το τελευταίο κομμάτι του αγωγού δεδομένων και σε τι αποσκοπούν. Συνοπτικά, μπορεί να είναι[4]:

- Πίνακες, διαγράμματα και γενικά στατιστικές αναλύσεις για λήψη αποφάσεων, ανάπτυξη σχεδίων και σχεδίαση μελλοντικών στρατηγικών.
- Προτάσεις και προγνώσεις για βελτιώσεις των τωρινών διαθέσιμων προϊόντων και την παραγωγή νέων καινοτόμων ή στρατηγικών marketing.
- Συμβουλευτικές υπηρεσίες και παρουσιάσεις προς στα ανώτερα στελέχη για ευκολότερη διευθέτηση των καθημερινών ζητημάτων, τον προγραμματισμό και οργάνωση της εταιρείας με συμμετοχή σε συναντήσεις και συμβούλια.
- Επεξεργασμένα Δεδομένα κατάλληλα προσαρμοσμένα τα οποία τροφοδοτούν αυτόματα, από βάσεις δεδομένων έως λογισμικά μηχανικής μάθησης ή πλατφόρμες ακόμη και εφαρμογές που λειτουργούν βάση των στοιχείων που λαμβάνουν.

Η κατανάλωση των ανωτέρω μπορεί να γίνεται είτε εσωτερικά, είτε εξωτερικά του οργανισμού με αποκλίσεις ως προς το χρονοδιάγραμμα. Συγκεκριμένα, εσωτερικά πιθανόν να απευθύνεται σε έτερα στελέχη ή προϊστάμενους τμημάτων για να εκτελέσουν τις καθημερινές τους εργασίες και στην ίδια την διοίκηση για να εκτελέσει τον σκοπό της ακόμη και στους ιδιοκτήτες και μετόχους για απλή ενημέρωση. Εξωτερικά μπορεί να απευθύνεται σε άλλους οργανισμούς-συνεργάτες, κρατικούς φορείς, συνεργαζόμενα ερευνητικά κέντρα είτε απευθείας στους πελάτες με κάποιες ίσως τροποποιήσεις και ανάλογα το είδος της εταιρείας. Η επικοινωνία των αποτελεσμάτων μπορεί να γίνεται σε ένα σταθερό χρονοδιάγραμμα ή κατά παραγγελία ή με αυτόματη τροφοδότηση όταν είναι έτοιμα στα αντίστοιχα πληροφοριακά συστήματα που απευθύνονται τα δεδομένα. Είναι σημαντικό τα παραγόμενα αποτελέσματα να είναι σωστά, αξιόπιστα, εύχρηστα και έτοιμα για κατανάλωση χωρίς να απαιτείται περαιτέρω ανάλυση μετά την ομάδα διαχείρισης δεδομένων, ώστε να μην σπαταλάτε επιπλέον χρόνος για περισσότερους ελέγχους και δημιουργηθούν αμφιβολίες ως προς την αξία της ομάδας, της ίδιας της διαδικασίας και των δεδομένων.

2.2.3. Πηγές δεδομένων

Η μεγαλύτερη διάκριση ανάμεσα στις πηγές δεδομένων είναι ως προς την μορφή των παραγόμενων δεδομένων, δηλαδή ψηφιακά ή μη ψηφιακά. Τα μη ψηφιακά δεδομένα

μπορούν να αποθηκευτούν και αυτά ψηφιακά, καταβάλλοντας ανάλογη ανθρώπινη προσπάθεια, καθιστώντας και αυτά αξιοποιήσιμα από τα σύγχρονα λογισμικά. Τα μη ψηφιακά δεδομένα μπορούν και αυτά να διακριθούν σε δύο κατηγορίες, σε αυτά που μπορούν να ψηφιοποιηθούν οι διαδικασίες που τα παράγουν και δεν έχουν αναληφθεί ακόμη ενέργειες όπως χειρόγραφα τιμολόγια και έγγραφα, γραπτές εξετάσεις μαθητών και η βαθμολόγησή τους ή οικονομικές συναλλαγές με μετρητά, και η δεύτερη κατηγορία είναι αυτές που δεν μπορούν να ψηφιοποιηθούν ακόμη με τις παρούσες τεχνολογίες όπως οι ανθρώπινες σκέψεις, τα συναισθήματα ή επιδόσεις αθλητών στο πραγματικό κόσμο.

Ωστόσο ολοένα και περισσότερο οι προσπάθειες επικεντρώνονται στην ψηφιοποίηση των διεργασιών και απόρροια αυτού, τα δεδομένα που παράγονται να είναι ψηφιακά αφαιρώντας τον κίνδυνο αλλοίωσης τους ή σφάλματος τους κατά την καταχώρησή τους, λόγω του ανθρώπινου παράγοντα. Γενικά εξαιτίας της ψηφιοποίησης αναδείχθηκε η γνωστή έκρηξη των δεδομένων και της δημιουργίας της έννοια των Μεγάλων Δεδομένων (Big Data). Ένας άλλος λόγος ύπαρξης των Μεγάλων Δεδομένων είναι και οι πολλαπλές πηγές από τις οποίες υπάρχει η δυνατότητα να εκμαιευτούν τιμές, στοιχεία, εικόνες, βίντεο ή μετρήσεις, τα οποία δημιουργούν ένα τεράστιο μείγμα με έντονη ποικιλομορφία κατά το οποίο πρέπει να απομονωθούν εκείνα τα σημεία που μπορούν να αξιοποιηθούν και να παράσχουν αξία [6].

Αξίζει να σημειωθεί, ακόμη μία διάκριση στις πηγές δεδομένων που επηρεάζει σημαντικά το είδος των εργαζομένων που απαιτείται, και αυτή είναι αν η εκάστοτε πηγή δεδομένων έχει ως αυτοσκοπό την παραγωγή δεδομένων ή όχι. Πιο συγκεκριμένα, σε ορισμένες περιπτώσεις επιδιώκεται η δημιουργία δεδομένων μέσω απογραφών, δειγματοληπτικής έρευνας, καταγραφή τιμών (π.χ. θερμοκρασίας), ή ανταλλαγή δεδομένων έναντι κάποιας υπηρεσίας ενώ σε άλλες περιπτώσεις τα δεδομένα παράγονται ως απόρροια άλλων ενεργειών των οποίων ο κύριος σκοπός είναι άλλος, όπως το ηλεκτρονικό εμπόριο, η χρήση των μέσων κοινωνικής δικτύωσης ή καθημερινή χρήση διάφορων υπηρεσιών ή υποδομών (π.χ. Μέσων Μαζικής Μεταφοράς). Τέλος, ο τρόπος μεταφοράς και εισαγωγής των δεδομένων στον υπόλοιπο αγωγό δεδομένων από τις πηγές, επιδρά σημαντικά ως προς την δόμηση και τρόπο λειτουργίας της ομάδας δεδομένων. Οι κυριότεροι τρόποι μπορεί να είναι:

- Απλή χειρωνακτική εισαγωγή στοιχείων ή ολόκληρων αρχείων.

- Καταχώρηση δεδομένων προγραμματιστικά.
- Απευθείας σύνδεση σε κάποια εφαρμογή, λογισμικό ή ιστοσελίδα.

2.2.4. Το επίπεδο αξιοποίησης των δεδομένων από την επιχείρηση

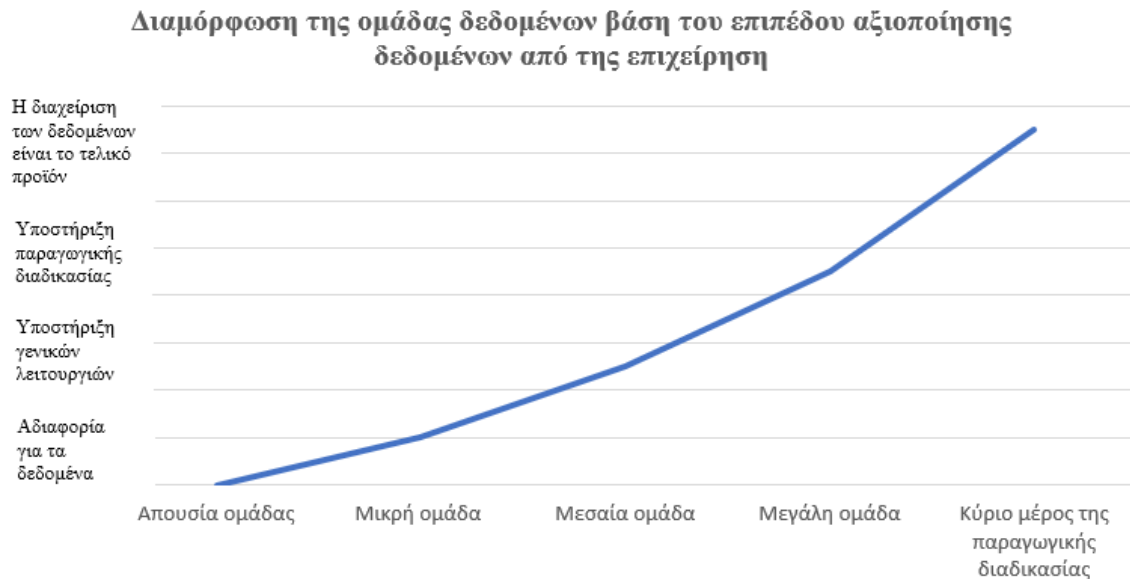
Στο παραπάνω παράδειγμα με το έξυπνο κλειστό σύστημα τηλεόρασης η ανάγκη για εξειδικευμένους αναλυτές δεδομένων θα ήταν μειωμένη σε περίπτωση που η διοίκηση δεν έδινε ιδιαίτερα βάση στα δεδομένα που παράγονται ως προς την κίνηση της πελατείας κατά το σχεδιασμό της στρατηγικής της και των διάφορων διαφημιστικών κινήσεων της. Ακολουθώντας παρόμοια προσέγγιση και άλλοι οργανισμοί, δεν αξιοποιούν πλήρως τα δεδομένα τους, ακόμη και αν αυτό δεν επιβάρυνε τον προϋπολογισμό τους αφού ήδη υπάρχει εγκαταστημένος εξοπλισμός καταγραφής και τα άτομα να υποστηρίζουν την επέκταση της χρήσης των δεδομένων. Έτσι η διαχείριση των δεδομένων δεν επηρεάζει σημαντικά το τελικό προϊόν ή υπηρεσία και είτε λόγω αδυναμίας της διοίκησης είτε λόγω αδιαφορίας και διαμορφώνεται ένα μειωμένο επίπεδο αξιοποίησης των δεδομένων, περιορίζοντας κατά επέκταση και την ομάδα που θα τα διαχειριζόταν.

Από την άλλη, ο ρόλος των δεδομένων είναι κρίσιμος για κάποιες εταιρείες, σε τέτοιο βαθμό που παύει να είναι μόνο υποστηρικτικός και επεκτείνεται και στην παραγωγική διαδικασία και αναλόγως το τελικό προϊόν μπορεί να είναι καίριο τμήμα, για να φτάσει στον τελικό καταναλωτή με συγκεκριμένες προδιαγραφές. Όταν αυτό που παράγει η εταιρεία, αποτελούν προϊόντα δεδομένων τότε η ομάδα δεδομένων μπορεί να ανήκει στην ομάδα παραγωγής και υποστηρίζοντας παράλληλα την υπόλοιπη επιχείρηση σε δεύτερο βαθμό για τις υπόλοιπες διεργασίες. Προφανώς σε τέτοιες περιπτώσεις απαιτούνται αρκετοί και άρτια καταρτισμένοι επαγγελματίες αυτού του πεδίου με μεγαλύτερη έμφαση στο διαχειριστικό κομμάτι των εργαζομένων καθώς η διαχείριση των δεδομένων δεν είναι πια συμπληρωματική αλλά νευραλγικό σημείο για την επιτυχία και επιβίωση της ίδιας της επιχείρησης.

Ένα πρότυπο τέτοιας επιχείρησης είναι η αμερικάνικη εταιρεία λογισμικού Palantir Technologies[31] με διεθνή παρουσία, η οποία εξειδικεύεται στην διαχείριση Μεγάλων Δεδομένων δημιουργώντας πληροφοριακά συστήματα για την προστασία της δημόσιας

ασφάλειας, διοίκησης και οργάνωσης επιχειρήσεων, υποστήριξη χρηματοπιστωτικών ιδρυμάτων, υπηρεσιών υπολογιστικού νέφους (Cloud) και γενικότερα έξυπνων ψηφιακών λύσεων ακόμη και της αντιμετώπισης της πανδημίας COVID-19. Συνδυάζει ένα μεγάλο εύρος τεχνολογιών και σύμφωνα με την Επιτροπή Κεφαλαιαγοράς των ΗΠΑ[10], η αξία της κεφαλαιοποίησης της για το 2021 ήταν περί των 2,291 δις εκατομμύρια δολάρια και ταυτόχρονα απασχολούσε, 2920 εργαζομένους εκ των οποίων οι περισσότεροι ήταν μηχανικοί υπολογιστών και αρκετοί επιστήμονες δεδομένων οι οποίοι δουλεύανε απευθείας στα προϊόντα και υπηρεσίες, αποτελώντας βασικό στοιχείο της παραγωγής.

Γενικά ο τρόπος που λαμβάνονται οι αποφάσεις ή που εκτελούνται οι καθημερινές εργασίες, το τελικό προϊόν, οι εσωτερικές διεργασίες είναι ορισμένοι παράγοντες που θα κρίνουν την εξάρτηση της επιχείρησης από τα δεδομένα ως προς την λειτουργία της και η ανάγκη να διαθέτει μία ομάδα δεδομένων ώστε να την υποστηρίξει να επιτύχει τους στόχους της.

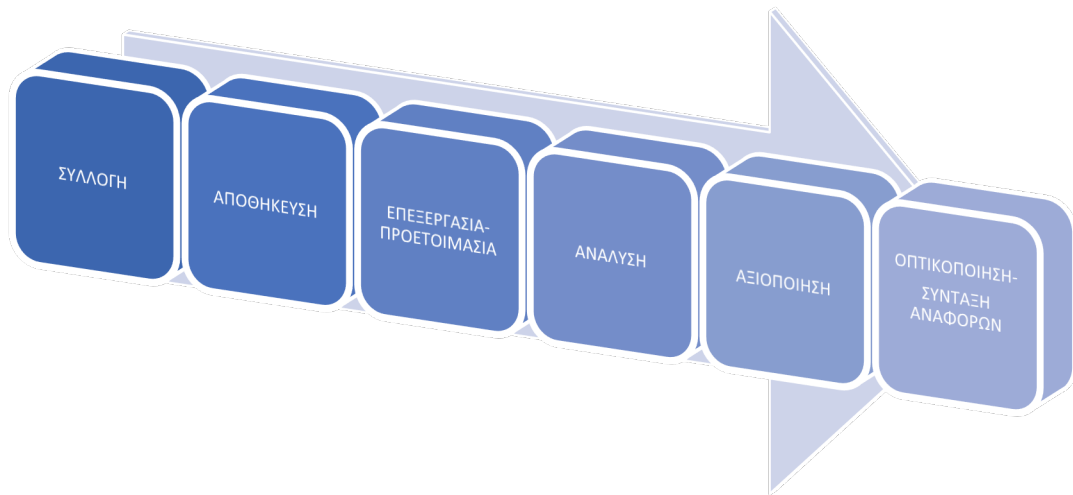


Εικόνα 24. Διαμόρφωση της ομάδας δεδομένων βάση του επιπέδου αξιοποίησης δεδομένων από της επιχείρηση

3. Η διαδρομή των δεδομένων για τη δημιουργία προστιθέμενης αξίας

3.1 Τυπική διαδρομή

Η επιστήμη των δεδομένων έχει ως σκοπό να δώσει λύσεις σε περίπλοκα προβλήματα τα οποία απαιτούν πολύπλευρη πληροφόρηση και να απαντήσει με ακρίβεια σε ερωτήματα βάση των στοιχείων που υπάρχουν. Τα ερωτήματα αυτά μπορεί να είναι απλά τύπου “ Ποιον μήνα έγιναν οι περισσότερες πωλήσεις ενός προϊόντος;” μέχρι και το “Πως θα αντιδράσει το καταναλωτικό κοινό σε ορισμένες τροποποιήσεις του προϊόντος και ποια από αυτές να υιοθετήσει το τμήμα παραγωγής;”. Έτσι έχοντας θέσει ως σημείο μηδέν την αναγνώριση των ερωτημάτων[6] και προβλημάτων που αντιμετωπίζει κάθε οργανισμός, διαμορφώνεται μία διαδρομή ώστε να φέρει εις πέρας την αποστολή αυτή, ακολουθώντας ωστόσο συγκεκριμένα βήματα τα οποία μεμονωμένα διαμορφώνονται ανάλογα με τις ανάγκες. Είναι φυσικό λόγω των αναγκών των εκάστοτε αναγκών και προκλήσεων ορισμένα βήματα να συγχωνεύονται ή να προσπερνιούνται ή να ανακατεύονται, πάντα διατηρώντας την ξεχωριστή νοητική οντότητα τους.



Εικόνα 25. Θεωρητική διαδρομή των δεδομένων

3.1.1. Συλλογή

Το πρώτο βήμα μετά την αφετηρία είναι η συλλογή όλων εκείνων των δεδομένων

από τις πηγές τους τα οποία μπορούν να αξιοποιηθούν και με την κατάλληλη επεξεργασία προσφέροντας μελλοντική αξία. Παλαιότερα λόγω των πιο περιορισμένων αποθηκευτικών και επεξεργαστικών μέσων ήταν σημαντικό να συλλέγονται μονάχα τα απαραίτητα δεδομένα με αποτέλεσμα όταν προέκυπτε ένα τελείως διαφορετικό περιστατικό, διαταράσσοντας την τυπική καθημερινότητα το οποίο θα απαιτούσε διαφορετικό τρόπο προσέγγισης, να μην έχουν διασωθεί τα παρελθοντικά στοιχεία εκείνα που θα μπορούσαν να υποστηρίξουν την νέα προσέγγιση. Όμως, λόγω της ανάπτυξης των αποθηκευτικών μέσων και την δυνατότητα μεγαλύτερης υπολογιστικής κλιμάκωσης, αυτού του είδους αποτυχίες παύουν να υφίσταται. Η συλλογή μπορεί να επιτευχθεί με διάφορες μεθόδους ωστόσο συνοψίζονται στους παρακάτω κύριους πέντε πυλώνες σύμφωνα με την διεθνή εταιρεία ερευνών μέσω λογισμικού QuestionPro (SaaS company)[32]:

- Παρατήρηση. Αποτελεί τον βασικότερο πυλώνα καθώς μέσω αυτού συλλέγουμε τα δεδομένα αβίαστα όπως αυτά δημιουργούνται στην πραγματική ζωή χωρίς την επέμβαση κάποιου επαγγελματία που προσπαθεί ουσιαστικά να τα εκμαιεύσει όπως γίνεται περίπου στις άλλες περιπτώσεις. Ουσιαστικά στην περίπτωση αυτή τα ίδια τα δεδομένα οδηγούν σε δημιουργία αξίας και όχι τα δείγματα. Η συνεχής παρατήρηση του αντικείμενου έρευνας ώστε να απομονωθούν οι αντιδράσεις του και να αντιστοιχηθούν με τις αιτίες που τις προκάλεσαν. Είτε το αντικείμενο είναι ένας πελάτης είτε ένας χρηματιστηριακός δείκτης, με την χρήση αυτής της μεθόδου πέρα από την τάση, την εποχικότητα και την κυκλικότητα, οι οποίες έτσι κι αλλιώς μπορούν να εντοπιστούν και στατιστικά, μπορεί να αποκρυπτογραφηθεί η επίδραση ενός μη συνηθισμένο ειδικού γεγονότος το οποίο θα διαταράξει την ομαλή λειτουργία του αντικείμενου στο οποίο γίνεται η παρατήρηση. Ο παρατηρητής προσπαθεί να αντιληφθεί όλες τις πτυχές που επιδέχονται επίδρασης, να την εντάξει σε μία μετρήσιμη κλίμακα και κατά πόσο μπορούν οι μετρήσεις αυτές να εφαρμοστούν και σε παρόμοια σενάρια. Η μέθοδος αυτή αποτελεί ένα υβρίδιο εμπειρικής και στατιστικής ανάλυσης και έτσι σημαντικό ρόλο παίζει ο ανθρώπινος παράγοντας ώστε να ερμηνευτούν πιο αποτελεσματικά μοναδικά γεγονότα και οι ομοιότητές τους με άλλα καθώς μεν συμβαίνουν σπάνια μεμονωμένα αλλά συχνά λαμβάνει χώρα ένα σπάνιο γεγονός και έτσι μπορεί να

- δημιουργηθεί ένας κοινός τρόπος αντίληψης για την αντιμετώπιση τέτοιων γεγονότων, αποφεύγοντας να τα χαρακτηρίσουμε Μαύρους Κύκλους λόγω της αδυναμίας για έγκαιρη και κατάλληλη αντιμετώπισή τους κατά το αρχικό στάδιο.
- Συνεντεύξεις. Παρουσιάζουν ομοιότητα με την δια ζώσης δειγματοληπτική έρευνα ωστόσο αναφέρεται σε πιο εξειδικευμένα αντικείμενα και το δείγμα είναι σαφώς πιο περιορισμένο και στοχευμένο, όπου μέσω αυτής της μεθόδου, ο αναλυτής εισέρχεται εις βάθος στο κάθε υποκείμενο της έρευνας. Πραγματοποιείται ακόμη και τηλεφωνικά ή μέσω κάποια εφαρμογής τηλεδιάσκεψης όταν δεν απαιτείται φυσική παρουσία και η καταγραφή των φυσικών αντιδράσεων. Δεν μπορεί να επιτευχθεί μεγάλη οριζόντια κλιμάκωση λόγω της μεγάλης απαίτησης ανθρώπινων πόρων και του χρόνου αλλά η καθετοποιημένη κλιμάκωση είναι εφικτή δηλαδή η διεξοδική συνέντευξη ενός ερωτηθέντα σκιαγραφώντας πλήρως το προφίλ του και τις αντιδράσεις του.
 - Δημοσκοπήσεις-Έρευνες. Η πιο διαδεδομένη είναι η δειγματοληπτική δημοσκόπηση κατά την οποία επιλέγεται ένα δείγμα του πληθυσμού ενδιαφέροντος, είτε τυχαία είτε επιλεγμένα, θέτοντας σε αυτό το δείγμα μία τυποποιημένη φόρμα ερωτήσεων η οποία συνήθως είναι προσπελάσιμη από αλγορίθμους ώστε στην συνέχεια εύκολα να αποδώσει αποτελέσματα. Οι ερωτήσεις μπορεί να είναι άμεσες ως προς το επιθυμητό τελικό αποτέλεσμα ή πλάγιες και μέσω των αλγορίθμων να εξαχθούν τα τελικά αποτελέσματα. Οι απαντήσεις στις άμεσες ερωτήσεις προφανώς είναι πιο ευδιάκριτες και απαλλαγμένες από αμφιβολίες ως προς την ερμηνεία τους ωστόσο οι συμμετέχοντες στην έρευνα, πιθανόν να μην είναι σε θέση να απαντήσουν καθώς η ερώτηση χρήζει μεγάλης ανθρώπινης σκέψης ή να αγγίζει ευαίσθητα θέματα ως προς τον ερωτηθέντα και έτσι οι απαντήσεις να είναι αναληθής. Αντίθετα, μέσω μίας σειράς πλάγιων ερωτήσεων με διαφορετικό βάρος η κάθε μία, μπορεί να εκμαιευτεί η απάντηση στο πραγματικό ερώτημα, καθοδηγώντας τον ερωτηθέντα να δώσει πιο αληθής απαντήσεις. Η δημοσκόπηση μπορεί να διενεργηθεί είτε τηλεφωνικά είτε διαδικτυακά είτε διαπροσωπικά. Ο κάθε τρόπος παρουσιάζει τα δικά του πλεονέκτημα και μειονεκτήματα ως προς την ταχύτητα, ανάγκη για προσωπικό ή αξιοπιστία και γι' αυτό ανάλογα την περίπτωση επιλέγεται ο αντίστοιχος τρόπος δημοσκόπησης ή ένα υβρίδιο όλων αυτών. Σε

περίπτωση που ελεγχθεί όλο το σύνολο, παύει να είναι δημοσκόπηση και είναι απογραφή.

- Στοχευμένες Ομάδες. Μια περιορισμένη μέθοδος έχοντας πειραματικό χαρακτήρα καθώς μελετώνται σε υποθετικά κοινά σενάρια και σε ένα συγκεκριμένο περιβάλλον διάφορα αντικείμενα διερεύνησης, τα οποία προέρχονται από διαφορετικές κλάσεις συνθέτοντας μια ποικιλόμορφη ομάδα. Υπάρχει μία συντονιστική αρχή είτε γνωστή είτε άγνωστη στο υπόλοιπο δείγμα ώστε να υπάρχει μία κατεύθυνση στο πειραματικό σενάριο ώστε να εξελιχθεί ομαλά. Γενικότερα μέσω αυτής προσέγγισης, οι αναλυτές προκαλούν τα δεδομένα να δημιουργηθούν μέσω μίας ανάδρασης. Τα περιβάλλοντα αυτά μπορεί να είναι απλά όπως ένας χώρος συνεδριάσεων μέχρι μία πολύπλοκη προσομοίωση και τα σενάρια να αφορούν σπάνια περιστατικά ή νέα προϊόντα και καινοτομίες. Είθισται αυτή η μέθοδος να προηγείται από κάθε προώθηση νέου προϊόντος ώστε να δοκιμαστεί από ένα ευρύ δείγμα καταναλωτών μία περιληπτική έκδοση του προϊόντος, όπως για παράδειγμα οι beta εκδόσεις των βιντεοπαιχνιδιών.
- Ιστορικά στοιχεία χρήσης. Η μέθοδος που καλύπτει το μεγαλύτερο κομμάτι αυτού του βήματος της διαδρομής δεδομένων είναι η εισροή των παρελθοντικών δεδομένων και αυτών που δημιουργούνται λόγω της εκτέλεσης εργασιών από τις πηγές τους στις αποθηκευτικές μονάδες. Μεγάλο ρόλο παίζει το είδος της πηγής και ο τύπος δεδομένων τα οποία μπορεί είτε να προκύπτουν από την λειτουργία της επιχείρησης είτε από διάφορα ψηφιακά εργαλεία όπως ηλεκτρονικές εφαρμογές ή τερματικές συσκευές πληρωμής (PoS) είτε εξωτερικά δεδομένα όπως ο πληθωρισμός ή οι τιμές συναλλάγματος. Η συλλογή αυτών μπορεί να γίνεται μέσω ανθρώπινης προσπάθειας είτε ψηφιακά μέσω μίας κατάλληλης λύσης. Μία αρκετά διαδεδομένη ψηφιακή τεχνική είναι τα δημοφιλή cookies τα οποία αποθηκεύονται στον φυλλομετρητή (browser) του κάθε χρήστη διαδικτύου όταν μεταβαίνει στην αντίστοιχη ιστοσελίδα που τα έχει δημιουργήσει και μέσω αυτών, ο ιστότοπος γνωρίζει το προφίλ και τον αριθμό των επισκεπτών του. Είναι, επίσης, σημαντικό η διαδικασία συλλογής να τηρεί ορισμένους κανόνες και να τους τηρεί συνεχώς ώστε τα δεδομένα που προκύπτουν να είναι συνεπή.



Εικόνα 26. Συλλογής Δεδομένων Source: QuestionPro

3.1.2 Αποθήκευση

Οι σύγχρονες υπολογιστικές δυνατότητες αποθήκευσης έχουν εξελιχθεί κατακόρυφα ως προς το μέγεθος των αρχείων και προγραμμάτων που μπορούν να υποστηρίξουν αλλά και την πολυπλοκότητά τους.[2] Έτσι, οι βάσεις δεδομένων έχουν αναπτυχθεί επιτρέποντας την διαχείριση των Μεγάλων Δεδομένων και δημιουργώντας ευκαιρίες για περαιτέρω καινοτομίες. Υπάρχουν σε αυτό το στάδιο δύο έννοιες κατά τις οποίες συχνά παρατηρείτε σύγχυση μεταξύ τους και μοιάζουν ταυτόσημες ενώ υπάρχει διακριτή διαφορά ανάμεσά τους και αυτές είναι η Βάση Δεδομένων(Database) και η Αποθήκη Δεδομένων (Data Warehouse).[4] Η βάση δεδομένων είναι οποιαδήποτε διακριτή συλλογή δεδομένων αποθηκευμένες σε κάποιο υπολογιστικό σύστημα με δυνατότητα πρόσβασης, αναζήτησης και ανάκτησης πληροφοριών μέσω ενός Συστήματος Διαχείρισης Βάσης Δεδομένων (DBMS). Ο πιο διαδεδομένος τύπος βάσης είναι οι

Σχεσιακές (Relational) ενώ τα τελευταία χρόνια άρχισαν να χρησιμοποιούνται πιο έντονα και άλλοι τύποι όπως οι Μη Σχεσιακές (NoSQL) και οι βάσεις στηριζόμενες σε υπολογιστικό νέφος (Cloud Databases). Υπάρχουν και άλλοι τύποι που χρησιμοποιούνται για συγκεκριμένους σκοπούς όπως τα Ιεραρχικά Μοντέλα Βάσεων Δεδομένων (Hierarchical database models) , Αντικειμενοστραφής Βάσεις Δεδομένων (Object - Oriented database) και Βάσεις Δεδομένων Κλειδιού-Τιμής (Key-Value database). Η άλλη έννοια, αυτής της Αποθήκης Δεδομένων ένα είδος βάσης δεδομένων ενσωματώνει δεδομένα από πολλές προελεύσεις με σκοπό την εξαγωγή αναφορών και αναλύσεων. Ιδιαίτερα διαδεδομένη σε επιχειρήσεις και οργανισμούς όπου υπάρχουν δεδομένα συναλλαγών, πελατών και εταιρικής λειτουργίας και με τους κατάλληλους συνδυασμούς μπορούν να εξαχθούν χρήσιμα αποτελέσματα.

Η διαδικασία της αποθήκευσης διαφέρει ανάλογα με τον τύπο της βάσης δεδομένων και των ίδιων των δεδομένων. Υπάρχει η χειρωνακτική μέθοδος κατά την οποία είτε προγραμματιστικά είτε μέσω του Συστήματος Διαχείρισης Βάσης Δεδομένων (ΣΔΒΔ) μπορεί να εισάγει ο χρήστης εγγραφές, λαμβάνοντας τα στοιχεία από τις πηγές.[3] Για παράδειγμα, σε μία σχεσιακή βάση μέσω ενός λογισμικού διαχείρισής της (π.χ. MySQL) με την χρήση της Δομημένης Γλώσσας Ερωτήσεων, SQL (Structured Query Language) μπορούμε να εισάγουμε νέα δεδομένα ως εγγραφές με τον παρακάτω τρόπο:

```
INSERT INTO table_name (column 1, column 2, column 3....)
VALUES (value 1, value 2, value 3....)
```

όπου:

- table_name, είναι το όνομα του πίνακα στο οποίο θα καταχωρηθούν οι νέες εγγραφές. (Μία σχεσιακή βάση δεδομένων αποτυπώνεται ως πίνακες και αυτοί με την σειρά τους με γραμμές και στήλες. Οι γραμμές κάθε πίνακα αποτελούν τις εγγραφές του οι οποίες απαρτίζονται από ορισμένα κελιά-πεδία.)
- column, τα ονόματα των στηλών, οι οποίες είναι τα χαρακτηριστικά τα οποία εκφράζουν κάθε εγγραφή.
- value, οι τιμές που θα εισαχθούν στα πεδία και θα απαρτίζουν την κάθε εγγραφή.
- INSERT INTO, VALUES εντολές αποδεκτές του ΣΔΒΔ.

Ωστόσο στην σύγχρονη με την έκρηξη των Μεγάλων Δεδομένων η ανωτέρω πρακτική είναι ανέφικτη να είναι η αποκλειστική μέθοδος εισαγωγής δεδομένων χειρωνακτικά λόγω ότι θα απαιτούσε πολλές εργατοώρες για την χειρωνακτική καταχώρηση των διάφορων παρατηρήσεων και στοιχείων, γι' αυτό πραγματοποιείται μέσα σε ένα κατανεμημένο υπολογιστικό πλαίσιο (Edge computing) η σύνδεση των επιχειρησιακών συστημάτων και εφαρμογών με εκείνα των διάφορων αποθηκευτικών μέσων και συσκευών, όπου τα δεδομένα θα μεταφέρονται σε αυτές είτε απευθείας είτε μέσω ενός δικτύου. Στην συνέχεια μέσω υπολογιστικών συστημάτων προωθούνται στην κατάλληλη μορφή ώστε να είναι προσπελάσιμα τα δεδομένα και να μπορούν να χρησιμοποιηθούν. Υπάρχουν τρεις βασικές μορφές αποθηκευτικών συσκευών στις οποίες τα δεδομένα μπορούν να αποθηκευτούν σε αυτές [33]:

- Αποθήκευση Αρχείων, ένα ιεραρχικό μοντέλο αποθήκευσης των αρχείων σε φακέλους και υποφακέλους όπου συνήθως είναι για περιπτώσεις όπως γεωγραφικά στοιχεία ή εργαζόμενοι τμημάτων. Ο πιο ενδεδειγμένος τύπος της βάσης δεδομένων, είναι οι σχεσιακές βάσεις δεδομένων (SQL) με αντιπροσωπευτικά παραδείγματα, το λογισμικό IMS (Information Management Software) της IBM για υποστήριξη επεξεργασίας συναλλαγών και το Windows Registry που υποστηρίζει τις ρυθμίσεις του λογισμικού Microsoft Office.
- Αποθήκευση σε επίπεδα (block), κατά την οποία τα δεδομένα αποθηκεύονται σε ξεχωριστά κομμάτια, όπου το κάθε κομμάτι διαθέτει ένα μοναδικό αναγνωριστικό. Αυτά τα επίπεδα μπορεί να ακολουθούν:
 - Μία ιεραρχική δομή με την λογική της προηγούμενης μορφής
 - Σχεσιακά block βάσεων δεδομένων με γλώσσα SQL, τα οποία είναι ένα μείγμα ψηφιακών περιβαλλόντων τα οποία αποτελούνται υπολογιστικά φύλλα, πίνακες εγγραφών και αυτά με την σειρά τους από στήλες και συναρτήσεις, όπως στην περίπτωση του DB2 συστήματος διαχείρισης δεδομένων της IBM.
 - Μη σχεσιακά block βάσεων δεδομένων τα οποία είναι επέκταση των ανωτέρων τα οποία μπορούν να κωδικοποιηθούν με διάφορους τρόπους, όπως XML, YAML, JSON , BSON ή ως απλό κείμενο και δημιουργήθηκαν με σκοπό να παρέχουν επεκτασιμότητα και ευελιξία για την κάλυψη των μεταβαλλόμενων επιχειρηματικών απαιτήσεων, σε πολύ μειωμένο χρόνο και δομή τα οποία

block μπορεί να είναι τύπου: είτε document based όπου εδώ τα επίπεδα-οντότητες αναφέρονται ως έγγραφα τα οποία είναι εύκολα προσπελάσιμα για γρήγορη αναζήτηση και φιλτράρισμα δεδομένων από διάφορες εφαρμογές όπως MongoDB ή Apache CouchDB είτε column based όπου εδώ οι οντότητες έχουν μορφή στηλών συγγενικών δεδομένων τα οποία μπορούν να γίνει η διαχείρισή τους πολύ γρήγορα ακόμα και αν είναι τεράστιο το πλήθος των εγγραφών τους (η διαφορά με τις σχεσιακές είναι στο ότι το μοναδικό κλειδί δεν είναι ανά εγγραφή αλλά ανά “οικογένεια” δεδομένων από προγράμματα όπως το Apache Cassandra ή το Bigtable της Google.

- Block βάσεων δεδομένων διαδικτύου όπου αυτά τα δεδομένα σε συνδυασμό άλλες τεχνολογίες μπορούν να υποστηρίξουν διάφορες εφαρμογές, με πιο διαδεδομένο και γνωστό παράδειγμα είναι τα κρυπτονομίσματα τα οποία ακολουθούν μια κρυπτοαλυσίδα (blockchain) η οποία μπορεί να προσδιοριστεί ως μία ακολουθία από blocks τα οποία διαθέτουν το καθένα ένα μοναδικό αναγνωριστικό και περιέχουν δεδομένα παλαιότερων συναλλαγών και στοιχείων τα οποία θα επιβεβαιώσουν την εγκυρότητα του νομίσματος.
- Αποθήκευση αντικειμένων, είναι η περίπτωση μη δομημένων στοιχείων τα οποία είναι δύσκολα να αποθηκευτούν και να αποτιμηθούν με τις κλασικές τεχνικές τα οποία δεδομένα προκύπτουν από ολόκληρες καρτέλες κειμένων ή από αρχεία ήχου, βίντεο, φωτογραφιών ακόμη και οι σχέσεις μεταξύ διάφορων οντοτήτων που έχουν παρατηρηθεί. Μη Σχεσιακές βάσεις δεδομένων τύπου Κλειδιού-Τιμής (Key-Value database) ή τύπου Γραφημάτων (Graph Database) μπορούν για την κάθε περίπτωση ανάλογα με το είδος του αντικειμένου και τις ανάγκες, να υποστηρίξουν την αποθήκευση αυτών των στοιχείων. Για παράδειγμα, αν σημασία έχει η απλότητα και η ταχύτητα τότε οι Βάσεις Δεδομένων Κλειδιού-Τιμής π.χ. (Amazon DynamoDB) μέσω μία απλής αντίστοιχης των αντικειμένων με ένα μοναδικό κλειδί είναι η λύση ενώ σκοπός είναι η διαχείριση των αντικειμένων καθώς και των σχέσεων τους θα χρησιμοποιηθούν Βάσεις Δεδομένων Γραφημάτων (Graph Database).

3.1.3. Επεξεργασία-Προετοιμασία

Ιδιαίτερα σημαντικό βήμα ως προς την αξιοπιστία ολόκληρης της διαδρομής καθώς εδώ επιλέγονται και εξάγονται τα δεδομένα από τις αποθηκευτικές μονάδες, τα οποία θα απαντήσουν συγκεκριμένα το κάθε ερώτημα ή πρόβλημα αφού πρώτα τους έχουν αφαιρεθεί οι διάφοροι “θόρυβοι” οι οποίοι θα δημιουργούσαν εσφαλμένα συμπεράσματα. Σύμφωνα με τη βιβλιογραφία,[4] στο βήμα αυτό πρέπει να εκτελεστούν τα εξής:

- Αφαίρεση διπλοεγγραφών.
- Συμπλήρωση κενών τιμών.
- Προσαρμογές και κανονικοποιήσεις.
- Στατιστική καταγραφή των δεδομένων και σύνοψη αυτών.
- Διαχείριση των ακραίων τιμών με τέτοιο τρόπο ώστε να μην διαταράσσει την συνολική ανάλυση αλλά και να λαμβάνουμε όποια χρησιμότητα μπορεί να δώσουν.
- Αντιπροσωπευτικά και αμερόληπτα δεδομένα.
- Ποικιλομορφία και πυκνότητα δεδομένων.
- Γενικότερα σε αυτό το επίπεδο, ενισχύεται η ποιότητα των δεδομένων και συνδέεται άμεσα με το επόμενο βήμα αυτό της ανάλυσης.

3.1.4. Ανάλυση

Γενικότερα ο όρος της ανάλυσης δεδομένων μπορεί να χρησιμοποιηθεί για να περιγράψει ολόκληρη την διαδρομή. Ωστόσο, περιορίζοντας τις ενέργειες διαχείρισης, του ήδη επεξεργασμένου δείγματος ή συνόλου δεδομένων με σκοπό την εξαγωγή συμπερασμάτων με κριτική προσέγγιση, μπορεί να οριοθετήσει το βήμα της ανάλυσης στην διαδρομή των δεδομένων. Έχουν ήδη περιγράψει στην παρούσα εργασία στο πρώτο κεφάλαιο οι κύριοι τύποι αναλύσεων, η Περιγραφική, η Διαγνωστική, η Προγνωστική και η Καθοδηγητική. Ανάλογα με τον αρχικό πρόβλημα που έχει τεθεί σαν αφετηρία επιλέγεται και ο αντίστοιχος τύπος, καθώς αυτά τα είδη των αναλύσεων διαμορφώθηκαν κατά τέτοιο τρόπο ώστε να αντιμετωπίζουν συγκεκριμένα προβλήματα. Γενικότερα, εδώ ο αναλυτής εξετάζει και εξερευνάει την προετοιμασία που έχει προηγηθεί και προσπαθεί με τα

κατάλληλα στοιχεία να σχηματίσει μοντέλα δεδομένων τα οποία θα μπορούν να κατανοήσει τα επεξεργασμένα δεδομένα και αν γίνεται να προχωρήσει και στην μοντελοποίηση τους προβλήματος.[29] Αυτό μπορεί να αναπτυχθεί με χρήση γλωσσών προγραμματισμού π.χ. R ή Python, κατάλληλων ψηφιακών εργαλείων όπως Microsoft Excel ή μέσω τυπικής στατιστικής ανάλυσης. Τα μοντέλα συνεχώς αξιολογούνται ως προς τα αποτελέσματα που προκύπτουν και διορθώνονται όταν καταλήγουν σε σφάλματα.

Τα συμπεράσματα που θα προκύψουν χωρίζονται σε δύο μεγάλες κατηγορίες σε αυτά που από την αρχή του ερωτήματος ήταν γνωστός ο τελικός προορισμός, Επιτηρούμενη Μάθηση και αυτά που δεν είναι, Μη Επιτηρούμενη Μάθηση[10].

Πιο συγκεκριμένα, στην πρώτη περίπτωση της Επιτηρούμενης Μάθησης ο αλγόριθμος αναπαράγει συναρτήσεις οι οποίες τροφοδοτούνται από τα νέα δεδομένα και από ήδη επισημασμένα δεδομένα. Τα επισημασμένα έχουν δημιουργηθεί από παλαιότερα παραδείγματα όπου έχει εκπαιδευτεί ο αλγόριθμος ή έχουν οριστεί εξαρχής από τον αναλυτή. Ουσιαστικά είναι ένα σύνολο παραδειγμάτων που προέκυψαν μέσω της εκπαίδευσης και παλαιότερης χρήσης του αλγόριθμου και έχουν μορφή πιθανόν λύσεων. Αυτού του είδους αλγορίθμου εφαρμόζεται για προβλήματα παλινδρόμησης ή ταξινόμησης όπου το λάθος της πρόβλεψης είναι εύκολα μετρήσιμο. Η μέτρηση γίνεται συγκρίνοντας τις προβλέψεις-λύσεις που πραγματοποίησε αρχικά ο αλγόριθμος με τις αντίστοιχες νέες παρατηρήσεις. Το ποσοστό απόκλισης αυτό καταγράφεται για βελτίωση του αλγορίθμου.

Στην δεύτερη περίπτωση, είναι ένας τύπος αλγορίθμου που μαθαίνει μοτίβα από τα δεδομένα που λαμβάνει χωρίς να υπάρχουν κάποια ήδη κάποιες γνωστές λύσεις από επισημασμένα δεδομένα όπως στην προηγούμενη περίπτωση και άρα δεν υπάρχει σφάλμα στις προβλέψεις. Οι εφαρμογές του αλγορίθμου είναι για προβλήματα συσταδοποίησης δεδομένων σε ομάδες(ή συστάδες) με κοινά χαρακτηριστικά και η επιτυχία του μπορεί να γίνει μέσω εμπειρικής ανάλυσης, χωρίς δηλαδή με στατιστικές μεθόδους ανάλυσης αλλά από τις εμπειρίες που διαθέτει ένας επαγγελματία του αντίστοιχου εργασιακού ή ακαδημαϊκού χώρου, από όπου και προέρχονται τα δεδομένα[16].

Ακόμη, υπάρχουν και άλλα είδη αναλύσεων πιο περιορισμένα αλλά και καλύτερα επικεντρωμένες σε συγκεκριμένους σκοπούς άρα και αποτελεσματικότερες σε αυτούς.

Μπορούν, ακόμη, να χαρακτηριστούν και ως υποκατηγορίες των ανωτέρω κύριων τύπων αναλύσεων και ορισμένες από αυτές είναι οι εξής[6]:

- Η Διερευνητική (EDA, Exploratory Data Analysis) μέσω της οποίας ο αναλυτής εξετάζει και εξερευνά τα δεδομένα ώστε να τα κατανοήσει εις βάθος και να αντιληφθεί τυχόν σχέσεις που υπάρχουν που δεν ήταν γνωστές. Πολύ σημαντικό σε αυτήν την προσέγγιση να γίνει κατανοητό ότι συσχέτιση δεν σημαίνει και αιτία καθώς το ότι μία μεταβλητή κινείται παράλληλα ή αντιθέτως παράλληλα με κάποια άλλη δεν σημαίνει ότι προκάλεσε κιάλας και την συμπεριφορά αυτή. Ωστόσο, από μόνη της δεν μπορεί να οδηγήσει σε μεγάλα συμπεράσματα και απλά βοηθάει την διαμόρφωση υποθέσεων για περαιτέρω μελέτη.
- Συμπερασματική Ανάλυση (Inferential Analysis) χρησιμοποιείτε ένα σχετικά μικρό δείγμα δεδομένων το οποίο να είναι ικανό ώστε να καταλήξουμε σε ένα συμπέρασμα για τον πληθυσμό γενικά όπως ένα στατιστικό μοντέλο δίνοντας παράλληλα μια εκτίμηση αβεβαιότητάς σας.
- Αιτιολογική Ανάλυση (Causal Analysis) χρησιμοποιείται για να δει τι συμβαίνει σε μια μεταβλητή όταν χειριζόμαστε μια άλλη μεταβλητή εξετάζοντας την αιτία και το αποτέλεσμα μιας σχέσης. Επίσης, καλύπτει το κενό της διερευνητικής ανάλυσης πάνω από την αιτιότητα και συνήθως λαμβάνει χώρα σε επιστημονικό περιβάλλον και την διενέργεια πειραματικών προγραμμάτων όπου ελεγχόμενα μπορούν να δοκιμάσουν να επηρεάσουν διάφορες μεταβλητές για να εκμαιεύσουν δεδομένα.

3.1.5. Αξιοποίηση

Σε αυτό το βήμα γίνονται αναφορές στις διεργασίες που εκτελούνται ύστερα από την ανάλυση των δεδομένων ώστε να επιτευχθεί μέγιστη αξία στο τελικό αποτέλεσμα ολόκληρου του οργανισμού. Επιπρόσθετα, στην αρχή του κεφαλαίου τέθηκε σαν αφετηρία του έργου η αναγνώριση ενός συγκεκριμένου προβλήματος και όλα τα βήματα κινούνται προς αυτήν την κατεύθυνση. Αυτού του είδους τα προβλήματα μίας επιχείρησης θα μπορούσε να προέρχονται από όλα τα τμήματα της όπως από το τμήμα πωλήσεων,

παραγωγής, λογιστηρίου και την ίδια την διοίκηση, ή να αποστέλλονται σε διάφορα πληροφοριακά συστήματα με αυτόματο ή και μη τρόπο, ώστε να υποστηριχθούν οι λειτουργίες τους. Τέτοια παραδείγματα συναντιόνται στις επιχειρήσεις τα οποία τα έχουν υιοθετήσει στις βασικές λειτουργίες τους όπως είναι η αναπλήρωση αποθεμάτων, ο προγραμματισμός παραγωγής, η σχέση προσφοράς και ζήτησης, η ικανοποίηση των πελατών, έγκριση δανείων και ασφαλίσεων, η σχεδίαση νέων ή εξατομικευμένων προϊόντων και η αξιόπιστη ενημέρωση σε πραγματικό χρόνο[3]. Γενικά η αξία που αποκτάται στηρίζεται στους τέσσερεις παρακάτω άξονες:

- Απόκτηση επιχειρησιακής αποτελεσματικότητας με αποδοτικότερη χρήση των διαθέσιμων πόρων.
- Προώθηση καινοτόμων ιδεών που προέκυψαν από την ορθή ερμηνεία των δεδομένων.
- Ταχύτητα και διαφάνεια στις επιχειρησιακές διεργασίες.
- Γρήγορη και έγκαιρη πληροφόρηση των απαραίτητων τμημάτων.

3.1.6. Σύνταξη αναφορών- Οπτικοποίηση

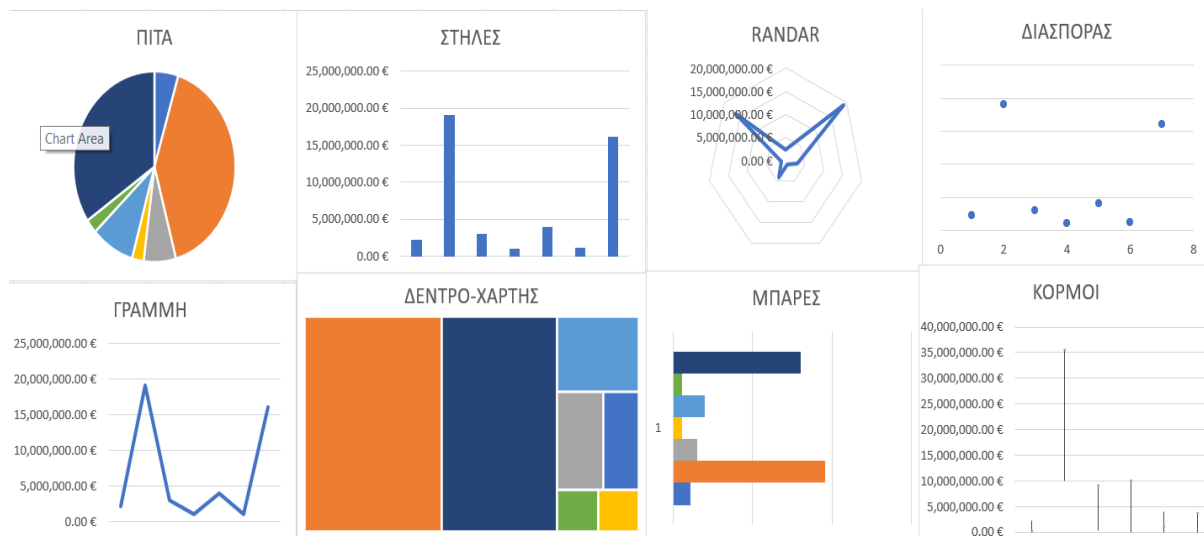
Το τελευταίο βήμα της διαδρομής αυτής το οποίο σε ορισμένες περιπτώσεις, αν δεν υπάρχει η απαίτηση μπορεί να παραληφθεί και να περιοριστεί ολόκληρη η διαχείριση των δεδομένων μέχρι το επίπεδο της ανάλυσής και της αξιοποίησής τους. Ωστόσο λόγω της μεγάλης προστιθέμενης αξίας που προσδίδεται σε όλη την διαχείριση των δεδομένων, έχει καταστεί σημαντικό βήμα και εντάσσεται στη συνήθη διαδρομή των δεδομένων. Η επικοινωνία της πληροφορίας που έχει προκύψει από τα προηγούμενα βήματα πραγματοποιείται ψηφιακά και ύστερα το αποτέλεσμα αποδίδεται σε συγκεκριμένα άτομα ή μπορεί και μαζικά μέσω διαδικτύου ή κάποιου άλλου μέσου ευρείας χρήσης. Στις επιχειρήσεις και οργανισμούς, συνήθως, αυτές οι αναφορές βασισμένες σε δεδομένα γίνονται γνωστές στα ενδιαφερόμενα μέρη είτε μέσω παρουσιάσεων (δια ζώσης ή εξ αποστάσεως) είτε απευθείας μέσω πλατφορμών δεδομένων, όπως το Tableau ή το PowerBI τα οποία δύο λογισμικά κυριαρχούν στις επιχειρήσεις και προσφέρουν ένα διαδραστικό περιβάλλον πλοήγησης των χρηστών κατά το οποίο μπορούν να ανατρέχουν σε αναφορές,

γραφήματα και γενικά οπτικές απεικονίσεις των δεδομένων. Σε κάθε περίπτωση, πρέπει να είναι κατανοητές και διαδραστικές, με τέτοιο τρόπο δομημένες ώστε να καλύπτει τις ανάγκες αυτών που τις βλέπουν.

Δεν υπάρχει νόημα η λεπτομερής προβολή δεδομένων, οι οποίες λεπτομέρειες δεν προσφέρουν τίποτα στους συγκεκριμένους σκοπούς του θεατή, ιδιαίτερα όταν αυτός δεν διαθέτει το αντίστοιχο γνωσιακό υπόβαθρο ώστε να αντιληφθεί τη σημασία τους, με την ταυτόχρονη ύπαρξη της πιθανότητας λάθους να υπερκαλύψουν το συνολικό νόημα. Αντίθετα, μία πλήρης εστιασμένη και ευκολονόητη παρουσίαση δεδομένων πάνω στο αντικείμενό της αλλά και τις ανάγκες αυτών που την παρακολουθούν, μπορεί να παίζει καθοριστικό ρόλο στην διαμόρφωση αποφάσεων και την σχεδίαση στρατηγικών με αυτοπεποίθηση από πλευράς των ατόμων που έχουν αναλάβει αυτές τις ευθύνες. Είναι σημαντικό να αναφερθεί ότι στην προσπάθεια για απλούστευση και μείωση των περιττών στοιχείων, εγκυμονείτε ο κίνδυνος να παραληφθούν ορισμένα δεδομένα και έτσι να διαστρεβλωθεί η τελική εικόνα. Συνοψίζοντας μία οπτικοποίηση δεδομένων πρέπει να διαθέτει τα παρακάτω χαρακτηριστικά[4]:

- Να είναι συνοπτική
- Να είναι εύκολα κατανοητή
- Να είναι πλήρης
- Να εξυπηρετεί τον αρχικό σκοπό της σε όλη την διάρκεια
- Να συνδυάζει τα δεδομένα, την στατιστική και την δύναμη της εικόνα με τέτοιο τρόπο ώστε είναι ενδιαφέρον στον θεατή και να δεχθεί το σωστό μήνυμα

Η μεγάλη πλειοψηφία των απεικονίσεων των δεδομένων σε επιχειρήσεις γίνεται με ορισμένα διαγράμματα, όπου τονίζουν με διαφορετικό τρόπο το ίδιο θέμα, επιδρώντας ταυτόχρονα στην ψυχολογία αυτού που θα καταναλώσει αυτήν την υπηρεσία.



Εικόνα 27. Παραδείγματα των Διαγραμμάτων Source: Microsoft Excel

Από τα ανωτέρω, τα πιο δημοφιλή είναι τα γραμμικά διαγράμματα στα οποία ενώνονται διάφορα σημεία κατά μήκος του κάθετου άξονα σε σχέση με τον οριζόντιο. Χρησιμοποιείται κυρίως για να φανεί η σχέση μεταξύ ορισμένων μεταβλητών ή πορεία ενός μεγέθους στην πορεία του άξονα X όπου συνήθως αποτυπώνεται μία χρονική περίοδος. Άλλη διαδεδομένη, είναι λεγόμενη “πίτα”, με την χρήση της οποίας αποτυπώνεται η κατανομή των επιμέρους ομάδων εντός ενός ευρύτερου συνόλου και η διαφορά των μεγεθών. Τέλος, το διάγραμμα στηλών όπου μπορεί εύκολα και κατανοητά να συγκριθούν διαφορετικά αντικείμενα ως προς μία μεταβλητή.

3.2. Μελέτη αρχιτεκτονικής συστήματος Netflix

3.2.1. Ιστορική αναδρομή

Ίσως η πιο γνωστή εταιρεία δικτυακών ενοικιάσεων και προβολών ταινιών στον κόσμο, η εταιρεία Netflix, κατάφερε να εξελιχθεί και να ξεχωρίσει έναντι των ανταγωνιστών της, εξαιτίας των καινοτόμων πρακτικών και τεχνολογιών της. Ιδρύθηκε το 1997 στην Καλιφόρνια των ΗΠΑ, εισέχοντας στον τομέα εκμετάλλευσης και διανομής ταινιών αλλά με πολύ πιο διαφορετική προσέγγιση από την σημερινή καθώς η διανομή

των ταινιών προς του πελάτες γινόταν μέσω αλληλογραφίας. Ήδη από την επόμενη χρονιά η εταιρεία αντιλήφθηκε την αξία της μετάβασης της αποθήκευσης ταινιών από βιντεοκασέτα σε DVD (ευκολότερη μεταφορά και αποθήκευση) και σε συνδυασμό με την ανερχόμενη επέκταση του διαδικτύου, λάνσαρε την ιστοσελίδα της, Netflix.com περιέχοντας έναν ψηφιακό κατάλογο ενοικίασης DVD.[10]

Ήδη η πολιτική χρέωσης ανά ενοικίαση θεωρήθηκε παρωχημένη από τους ιθύνοντες της εταιρείας και υιοθετήθηκε η συνδρομητική πολιτική χρέωσης ανά μήνα από τις αρχές του 2000, μία τακτική που την αγάλιασε το καταναλωτικό κοινό καθώς έπαψε να πληρώνει ογκώδης χρεώσεις αν αθετούσε τους περιορισμένους χρόνους ενοικίασης. Επίσης, εισήχθη η καινοτόμα ιδέα στην ιστοσελίδα να προσφέρει εξατομικευμένες προτάσεις στους πελάτες της μέσω ενός αλγορίθμου με το όνομα CineMatch, οποίος είχε ποσοστό επιτυχίας 75% επί τις άριστες βαθμολογίες των χρηστών οι οποίοι διαλέξανε ταινίες που τους είχαν προταθεί και στην συνέχεια βαθμολογήσανε άριστα την ταινία [4]. Ο αλγόριθμος λάμβανε υπόψη για προτείνει κάποια ταινία σε ένα θεατή τα εξής: την κατηγορία είδους που άνηκε η ταινία, το προφίλ του υποψήφιου ενοικιαστή όπως παλαιότερες ενοικιάσεις και βαθμολογίες του ή τις πρόσφατες κινήσεις του, και τέλος όλες οι βαθμολογίες που έχουν προσφέρει οι χρήστες. Μέσω αυτών των κινήσεων, της συνεχής προσπάθειας μείωσης των μηνιαίων τελών καθώς και άλλων τακτικών μέχρι το 2006 η εταιρεία είχε διαφοροποιηθεί από τους ανταγωνιστές της και πρωταγωνιστούσε στο χώρο εκμετάλλευσης τέτοιου περιεχομένου. Ωστόσο λόγω της επιθετικής δραστηριότητας της είχε αναγκαστεί να λάβει υψηλές δανειοδοτήσεις το οποίο είχε αυξήσει τον κίνδυνο και κατά επέκταση την πίεση στα στελέχη της.

Το 2006 ξεκίνησε ένα διαγωνισμό μηχανικής μάθησης και εξόρυξης δεδομένων με έπαθλο 1.000.000\$[4] για εκείνον που θα δημιουργήσει έναν αλγόριθμο με μικρότερη απόκλιση. Δόθηκε παράλληλα σε όποιον ενδιαφερόμενο έναν τεράστιο αριθμό ακατέργαστο δεδομένων για να εργαστεί πάνω σε αυτά, αμφιλεγόμενη τακτική για τα σημερινά δεδομένα καθώς θα είχε εγείρει αρκετά ερωτήματα και ως προς την διαφύλαξη προσωπικών δεδομένων ή ότι θα μπορούσε να αποτελέσει πεδίο εκμετάλλευσης από τους ανταγωνιστές. Ωστόσο, εκείνη την περίοδο αποτελούσε άγνωστη για το μεγαλύτερο μέρος του πληθυσμού η αξία της επιστήμης δεδομένων και γενικά των δεδομένων. Ο αλγόριθμος CineMatch παρήγαγε σφάλμα 0,9525[34] χρησιμοποιώντας ως τρόπο υπολογισμού την

Ρίζα Μέσου Τετραγωνικού Σφάλματος(RMSE)[34] η οποία υπολογίζεται ουσιαστικά από τον παρακάτω τύπο:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - F_i)^2}$$

όπου:

n= ο αριθμός των παρατηρήσεων που έχουμε διαθέσιμα

i= η χρονική περίοδος που για το οποίο αναφέρονται τα μεγέθη

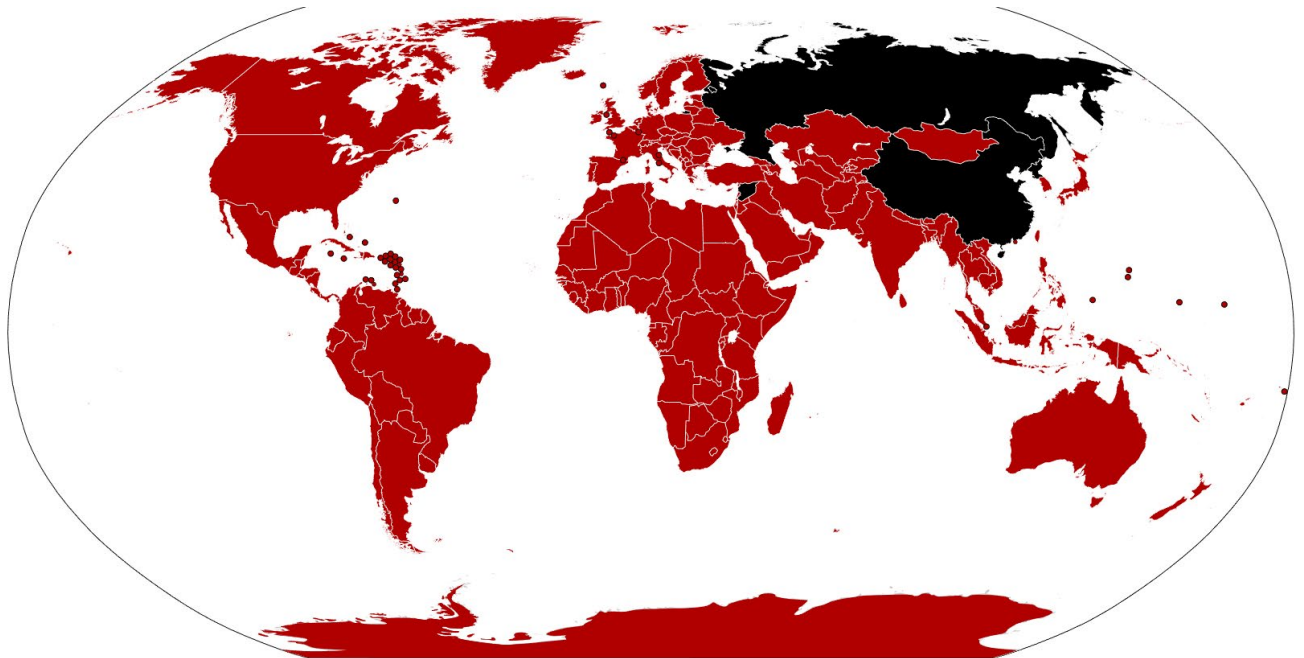
Y= η πραγματική παρατήρηση μίας τιμής

F= η πρόβλεψη της τιμής που είχε γίνει

Το 2009 απονεμηθεί το βραβείο σε μία ομάδα επαγγελματιών εφαρμόζοντας ένα μείγμα αλγορίθμων και τεχνικών να μειώσουν τουλάχιστον κατά 10% το σφάλμα (έφτασαν περίπου στο $RMSE=0,85$)[34] που ήταν η προϋπόθεση ώστε να κερδίσουν τον διαγωνισμό ωστόσο η εταιρεία ποτέ δεν υιοθέτησε αυτούς τους αλγόριθμους λόγω της υπολογιστικής δύναμης που θα απαιτούνταν για να κερδηθεί μία τέτοια οριακή αύξηση της ακρίβειας. Αξίζει, επιπρόσθετα, εδώ να δοθεί έμφαση στη σχέση μεταξύ του κόστους και της ωφέλειας που απορρέουν από κάθε αλλαγή ή επένδυση όπου στην προκειμένου περίπτωση θα ήταν αρνητική η σχέση μεταξύ των βελτιώσεων του μοντέλου συστάσεως και της υπολογιστικής και μηχανικής προσπάθειας και επένδυσης.[34] Το σύστημα προτάσεων που θα είχε η εταιρεία ήταν ιδιαίτερο σημαντικό αν αναλογιστεί κανείς ότι η θέαση γινόταν μέσω αποστολής DVD οπότε σε περίπτωση που δεν ήταν αρεστή μία ταινία στον πελάτη δεν είχε την δυνατότητα να μεταβεί σε μία επόμενη όπως συμβαίνει σήμερα με τις δικτυακές συνδρομητικές που προσφέρουν οι σύγχρονες πλατφόρμες, άρα ο πελάτης ή θα έβλεπε μία ταινία που δεν τον ενδιέφερε ή δεν έβλεπε μέχρι να του σταλθεί μέσω ταχυδρομείου μία άλλο. Όπως είναι φυσικό αυτό θα μείωνε την ικανοποίηση από το συνολικό προϊόν που λαμβάνει ο πελάτης εξαιτίας του συστήματος προτάσεων το οποίο λόγω της εποχής (μειωμένες πηγές ενημέρωσης σε σχέση με σήμερα) αποτελούσε βασικό κομμάτι επιλογής ταινιών.

Το 2007 έγινε η έναρξη της θέασης των ταινιών δικτυακά σε πραγματικό χρόνο μειώνοντας έτσι την αναμονή για την παρακολούθηση ταινιών σε μηδενικό χρόνο. Αρχικά ξεκίνησε για μία μικρή ποικιλία ταινιών και συνεχώς διευρυνόταν. Σταδιακά η εταιρεία

μεταμορφώθηκε στην μορφή που την γνωρίζει η πλειοψηφία του κόσμου παύοντας να διανέμει ταχυδρομικά σε υλική υπόσταση αλλά μέσω διαδικτύου σε πραγματικό χρόνο. Από το 2011 ήταν η κυρίαρχη εταιρεία στην διανομή on-line ταινιών και λόγω όλων των ανωτέρων ήταν εύκολη η σταδιακή επέκταση σε όλο τον κόσμο όπου σήμερα πέρα από ορισμένες περιπτώσεις (Κίνα, Ρωσία, Βόρεια Κορέα, Συρία) η πρόσβαση στην πλατφόρμα της εταιρείας είναι δυνατή παντού μέσω μηνιαίας συνδρομής προσφέροντας εξατομικευμένο περιεχόμενο σε κάθε χώρα ανάλογα με τα δεδομένα που έχει συλλέξει. [35]



Εικόνα 28. Προσβασιμότητα στην πλατφόρμα Netflix σε όλο τον κόσμο. Με κόκκινο παρουσιάζονται οι χώρες που έχουν πρόσβαση ενώ με μαύρο αυτές που δεν έχουν. Source: Netflix's Help Center για τον Μάρτιο 2022.[35]

3.2.2. Σύστημα συστάσεων (recommendation system)

Ένα από τα πιο κρίσιμα και γνωστά αντικείμενα για την επιτυχία της εταιρείας όπως αναφέρθηκε, είναι το σύστημα συστάσεων (recommendation system) το οποίο προτείνει στον χρήστη ταινίες και τηλεοπτικές σειρές ανάλογα με τα ενδιαφέροντα του μέσα από ένα τεράστιο εύρος επιλογών το οποίο θα ήταν χρονοβόρο κάθε φορά να εξετάζει ο ενδιαφερόμενος. Σήμερα, και άλλες γνωστές και μεγάλες εταιρείες στο χώρο των διαδικτυακών υπηρεσιών χρησιμοποιούν παρόμοιο σύστημα ώστε να προσφέρουν

εξατομικευμένες υπηρεσίες στους καταναλωτές τους, όπως η Spotify, η YouTube και Amazon. Αυτού του είδους συστήματα παρουσιάζουν δύο κύριες προσεγγίσεις, το Συνεργατικό Φιλτράρισμα (Collaborative Filtering) και το Φιλτράρισμα Περιεχομένου (Content-Filtering).[3] Μελετώντας την περίπτωση της εταιρείας Netflix, η πρώτη προσέγγιση θα πρότεινε ταινίες στον θεατή οι οποίες προτιμήθηκαν και αξιολογήθηκαν θετικά από άλλους θεατές με κοινά ενδιαφέροντα θέασης στο παρελθόν, δηλαδή υπάρχει η έννοια των οντοτήτων χωρίς να σημαίνει ότι ένα άτομο θα ανήκει μόνο σε μία οντότητα χρηστών. Η δεύτερη περίπτωση φιλτράρει τα χαρακτηριστικά του περιεχομένου που έχει παρακολουθήσει στο παρελθόν και προσπαθεί να βρει μοτίβα σε αυτά τα οποία προτιμά να διαθέτει μια ταινία, όπως σκηνοθέτης, ηθοποιοί ή είδος ταινίας. Και οι δύο περιπτώσεις διαθέτουν την έννοια της ταξινόμησης και εντάσσονται στα προβλήματα Μη Επιτηρούμενης Μάθησης καθώς οι πραγματικές λύσεις, στην προκειμένου περίπτωση οι προτιμήσεις των θεατών, είναι άγνωστες κατά την επίλυση του προβλήματος.

Το σύστημα σύστασης της Netflix έχει ως σκοπό ουσιαστικά να προτείνει εξατομικευμένα στον κάθε χρήστη ώστε να βελτιώσει την εμπειρία της περιήγησης στην πλατφόρμα της [36]. Σύμφωνα με την ίδια την εταιρεία οι κύριες πηγές όπου συλλέγει δεδομένα είναι οι παρακάτω:

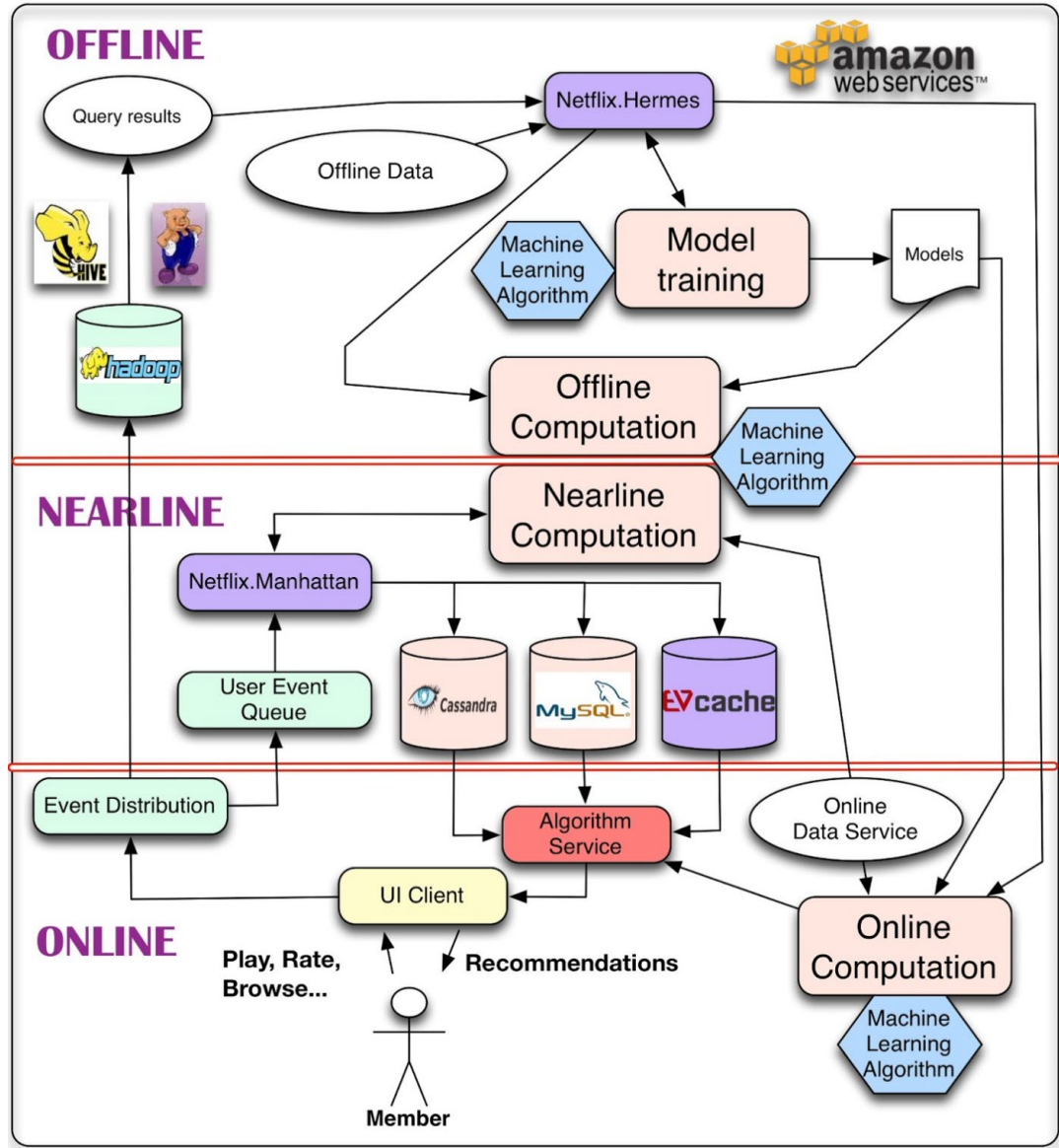
- Οι βαθμολογίες των χρηστών για κάθε αντικείμενο.
- Μετρητές δημοτικότητας ανάλογα το κάθε παράγοντα που λαμβάνεται υπόψη όπως γεωγραφικός, ωρολογιακός, ηλικιακός, ημερολογιακός, γλωσσικός, φύλλο.
- Οι αναζητήσεις μέσω της μηχανής αναζήτησης της πλατφόρμας.
- Τα χαρακτηριστικά του κάθε θεάματος και τα μεταδεδομένα που το ακολουθούν όπως η χρονιά κυκλοφορίας, η γλώσσα, η διάρκεια, οι καλλιτέχνες.
- Τύπος συσκευής παρακολούθησης, ο συνολικός χρόνος παρακολούθησής του κάθε αντικειμένου και όλων μαζί σε μια χρονική περίοδο.
- Πέρα από τα δεδομένα που εκμαιεύονται εσωτερικά, ενσωματώνονται από εκείνα που δημιουργούνται εξωτερικά όπως κριτικές ταινιών και πωλήσεις εισιτηρίων από τους κινηματογράφους.
- Από τα μέσα κοινωνικής δικτύωσης των χρηστών ως προς τους φίλους τους ή τα ενδιαφέροντά τους ή τις ομάδες που ανήκουν.

Είναι ευκόλως κατανοητό ότι τα δεδομένα είναι απεριόριστα και η διαχείρισή τους αποτελεί περίπτωση Μεγάλων Δεδομένων. Η συνολική αρχιτεκτονική της διαδρομής των δεδομένων ακολουθεί ένα υβριδικό τρόπο δομής καθώς διαθέτει επίπεδα που λειτουργούν εκτός σύνδεσης με το διαδίκτυο, εντός σύνδεσης και ενδιάμεσα. Κάθε επίπεδο προσφέρει με διαφορετικό τρόπο και γι' αυτό το λόγο σχεδιάστηκε έτσι. Η αποθήκευση βρίσκεται στο πρώτο καθώς οι ανάγκες αναφέρονται σε κατάλληλο υλικό εξοπλισμό συνδυασμένο με τεχνολογία που θα το ελέγχει και όχι σε διεργασίες που πραγματοποιούνται σε πραγματικό χρόνο.

Το 2016 ξεκίνησε συνεργασία με την Amazon Web Services ώστε να μπορέσει να ικανοποιήσει τις τεράστιες αποθηκευτικές και υπολογιστικές ανάγκες με χρήση χιλιάδων διακομιστών και υπολογιστών, και παράλληλα ασφαλή και δυναμική κλιμάκωση των δεδομένων της. Ταυτόχρονα με χρήση κατάλληλων ψηφιακών εργαλείων, όπως το Apache Hadoop επεξεργάζεται σε κατανεμημένες υπολογιστικές μονάδες, και με χρήση λογισμικών εξασφαλίζει βέλτιστη ροή και συνέπεια στα δεδομένα όπως το Hermes που η ίδια η εταιρεία σχεδίασε. Με την χρήση αλγορίθμων μηχανικής μάθησης εκπαιδεύεται το μοντέλο και προχωρά στο επόμενο και ενδιάμεσο επίπεδο, με την επίβλεψη του λογισμικού Manhattan, που και αυτό η ίδια εταιρεία σχεδίασε, για αποτελεσματική ροή, όπου τα δεδομένα επεξεργάζονται και αναλύονται με τα κατάλληλα Συστήματα Διαχείρισης Βάσεων Δεδομένων όπως το Cassandra, MySQL και EvCache με σκοπό τη σωστή οργάνωσή τους. Αξίζει να σημειωθεί ότι και τα τρία αυτά συστήματα διαφέρει η τεχνολογία τους και καταδεικνυε τις περιπλοκές και διαφορετικές ανάγκες στο επίπεδο αυτό.

Έτσι, συνεχώς βελτιώνεται αυτόματα το σύστημα και ανατροφοδοτείται με εκ νέου δεδομένα που αποστέλλονται από τελευταίο επίπεδο. Σε αυτό το επίπεδο είναι ο τελικός αλγόριθμος που θα αναλύσει τα υπάρχον στοιχεία και αποστέλλει προτάσεις, όπου ανάλογα με την επιτυχία ή όχι της πρότασης δημιουργούνται τα αντίστοιχα δεδομένα με βάση των διάφορων μετρικών που λαμβάνονται υπόψη. Στο επίπεδο αυτό απαιτείται να υπάρχει εύχρηστη και άμεση διεπαφή με τον χρήστη μέσω διαδικτύου ώστε η αλληλεπίδραση να είναι σε πραγματικό χρόνο και να προσμετράει και τα τελευταία δεδομένα και είναι πλήρως ενήμερο[34].

Χρησιμοποιώντας περαιτέρω τα δεδομένα η εταιρεία ξεκίνησε να δημιουργεί δικό της πρωτότυπο περιεχόμενο ανάλογα με τις απαιτήσεις και επιθυμίες του κοινού της. Έτσι, γνωρίζοντας την τάση αλλά και τα κενά της αγοράς εστίασε στην κατασκευή ταινιών και σειρών οι οποίες θα έχουν άμεσης αποδοχής από το κοινό τους την περίοδο εκείνη. Ανεξάρτητα από την ποιότητα ή την καλλιτεχνική τους αξία, τα θεάματα αυτά είχαν προβάδισμα σε σχέση με τα άλλα που δημιουργούνται εκτός της εταιρείας αυτής, καθώς η διάθεσή τους γινόταν με τέτοιο τρόπο όπου λαμβανόταν υπόψη οι συσχετίσεις μεταξύ παραγόντων όπως οι ώρες θέασης ανάλογα με την εποχή, οι αναζητήσεις στην μηχανή αναζήτησης ανάλογα με την κατηγορία της ταινίας/σειράς, οι πρόσφατες κινηματογραφικές επιτυχίες, κουλτούρα και έθιμα της περιοχής[36].

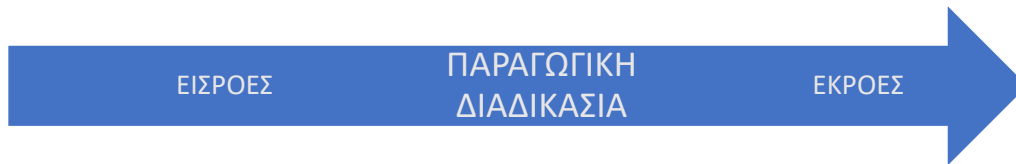


Εικόνα 29. Το μοντέλο αρχιτεκτονικής του συστήματος σύστασης της Netflix[34]

4. Διαδικασίες και μέσα για την επίτευξη επιχειρησιακών στόχων σε ένα σύγχρονο περιβάλλον

4.1. Συστατικά στοιχεία επιχειρήσεων - Επιχειρησιακό πλάνο

Κάθε εταιρεία κατά την ίδρυση της και όσο εξελίσσεται αποτελείται από ορισμένα στοιχεία τα οποία την ξεχωρίζουν και την καθιστούν μοναδική, συνθέτουν δηλαδή μία ταυτότητα. Πέρα από αυτό, η επιχείρηση αναπτύσσεται με έναν συγκεκριμένο τρόπο ακολουθώντας ορισμένες τακτικές ανάλογα με την ταυτότητα που έχει καθώς και το περιβάλλον της. Μια απλουστευμένη παραγωγική διαδικασία είναι αυτής των τριών βημάτων: εισροής, παραγωγική διαδικασία και τέλος εκροή[37], όπου οι εκροές πρέπει να είναι μεγαλύτερες από τις εισροές, εκφρασμένες πάντα σε χρηματικές μονάδες ώστε να είναι βιώσιμη η επιχείρηση. Όταν είναι ισοδύναμες είναι το νεκρό σημείο.



Εικόνα 30. Απλουστευμένο μοντέλο επιχειρησιακής λειτουργίας

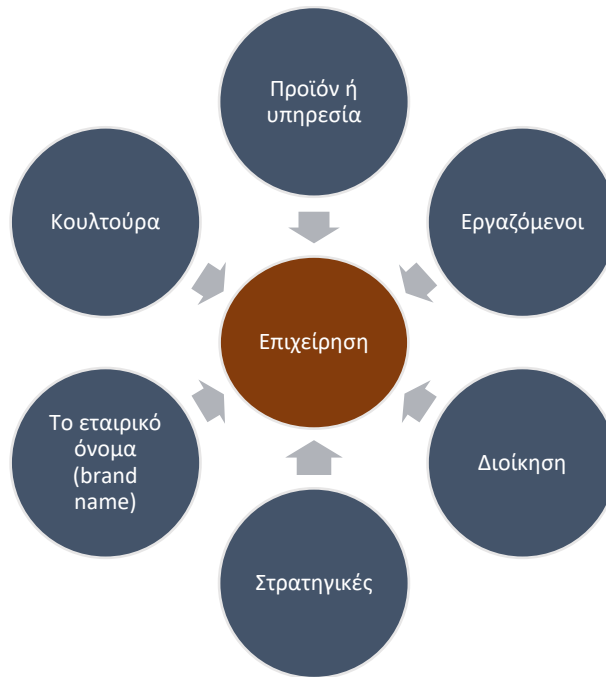
Αν και το ανώτερο είναι τελείως απλουστευμένο και δεν μπορεί να έχει καθαυτή ρεαλιστή εφαρμογή στην πραγματική ζωή, μπορεί να αποτελέσει η θεμελιώδης εισαγωγή κατάρτισης ενός επιχειρηματικού σχεδίου. Εδώ αξίζει να σημειωθεί ότι είναι διαφορετικές έννοιες τα συστατικά/θεμελιώδη στοιχεία μία εταιρείας και άλλη έννοια είναι το επιχειρηματικό πλάνο, καθώς αυτές οι δύο έννοιες είναι κοντινές και εύκολα συγχέονται.

Αρχικά, όταν αναφερόμαστε σε συστατικά στοιχεία μίας επιχείρησης είναι όλα αυτά που τι απαρτίζουν ανεξάρτητα την υλική υπόστασή τους. Σε πολλά άρθρα και βιβλιογραφίες αναφέρονται αρκετά σημεία ως θεμελιώδη και αξίζουν να μελετώνται ξεχωριστά για την αποτελεσματικότερη κατανόησή τους[37]. Σε μία προσπάθεια σύνοψης όλης της πληροφορίας που κυκλοφορεί, τα πιο σημαντικά στοιχεία τα οποία αποτελούν και ταυτότητα της εταιρείας είναι τα εξής:

- Το προϊόν ή υπηρεσία, δηλαδή ο λόγος ύπαρξης του οργανισμού και το τι προσφέρει στην κοινωνία.
- Οι εργαζόμενοι, οι οποίοι μέσα από τις δικές τους ικανότητες και πρωτοβουλίες το παραγόμενο προϊόν θα διαμορφωθεί και θα πωληθεί στους πελάτες.
- Η διοίκηση, είτε είναι διοικητικό συμβούλιο είτε είναι μέτοχοι είτε είναι απευθείας ιδιοκτήτες. Ελέγχει την επιχείρηση ώστε να λειτουργούν ομαλά και θα έχει τον τελευταίο λόγο ως προς την άσκηση των επιχειρηματικών πλάνων και την σχεδίαση στρατηγικών, ωστόσο λόγω της σημασίας αυτών κρίνεται σημαντικό να μελετηθούν ξεχωριστά.
- Το όνομα της επιχείρησης ή πιο ευρέως γνωστά το brand name, και γενικά ο τρόπος που θέλει και προσπαθεί η διοίκηση και οι εργαζόμενοι να προβάλλεται η ίδια και αυτά που παράγει, στο καταναλωτικό κοινό.
- Η κουλτούρα, που εισάγει η διοίκηση για διάφορα θέματα όπως ηθικής ή οι σχέσεις της επιχείρησης με το κράτος, το περιβάλλον και την κοινωνία αλλά και εσωτερικές σχέσεις μεταξύ των εργαζομένων καθώς και οι διάφορες πολιτικές περί επιβράβευσης, τιμωρίας και γενικά εργασιακών θεμάτων.
- Στρατηγικές σε διάφορα ζητήματα, ο τρόπος με τον οποίο θα πορευτεί η επιχείρηση όπως για παράδειγμα αν ασκήσει επιθετική πολιτική έναντι των ανταγωνιστών της, να επιτρέπει εύκολα τις καινοτόμες ιδέες, να εισέρχεται σε νέα αγορές. Σε αυτό το στοιχείο θα μπορούσε να ειπωθεί ότι καταρτίζεται το επιχειρηματικό πλάνο, το οποίο δύναται συνεχώς να ανανεώνεται.

Δεν είναι εύκολο να αναγνωριστεί ποιο είναι το πιο σημαντικό καθώς κάθε επιχείρηση είναι διαφορετική περίπτωση και μία γενίκευση δεν θα μπορούσε να συμπεριλάβει όλες τις περιπτώσεις. Για παράδειγμα, μία εταιρεία εμπορίας πολυτελών ρούχων είναι σημαντικό όλες οι διεργασίες της να επικεντρώνονται γύρω από την διαφήμιση του brand name της ώστε να συνεχιστεί η αντίληψη από πλευράς πελατών ότι χρησιμοποιώντας αυτά τα προϊόντα αποτελεί πολυτέλεια και ανωτερότητα, με ότι συνεπάγεται αυτό, ενώ παράλληλα μπορεί να υπάρχει αδιαφορία ως προς την κουλτούρα σε θέματα προστασίας περιβάλλοντος και ζώων λόγω χρήσης ζωικών προϊόντων (π.χ. δέρμα, γούνα) για

εξασφάλιση υψηλότερης ποιότητας προϊόντων. Από την άλλη μια ισχυρή επενδυτική επιχείρηση, δεν θα την ενδιαφέρει να αναλωθεί σε διαφημίστηκες καμπάνιες ώστε να χτίσει ένα δυνατό όνομα καθώς θα στοχεύει σε συγκεκριμένο επενδυτικό κοινό και θα προσπαθεί να διαθέτει ικανά ανώτερα στελέχη και εργαζομένους, τα οποία θα αποτελούν από μόνα τους διαφήμιση σε ειδικούς του επενδυτικού χώρου. Έτσι, θα προβάλλονται ως εγγύηση αυθεντίας και εμπιστοσύνης σε υποψήφιους επενδυτές καθώς θα έχουν την δυνατότητα αυτοί οι υπάλληλοι να πετυχαίνουν πιο εύκολα τα επιθυμητά αποτέλεσμα για τους πελάτες τους.



Εικόνα 31. Συστατικά στοιχεία μίας επιχείρησης.

Επιπρόσθετα, υπάρχει και η έννοια του επιχειρηματικού πλάνου με πρωταρχικό σκοπό η επιχείρηση να είναι βιώσιμη και να παρουσιάζει κερδοφορία ή έστω να μην παρουσιάζει ζημίες, δηλαδή εκροές \geq εισροές. Προς σε αυτήν την κατεύθυνση η διοίκηση καταρτίζει στρατηγικές ώστε να το εξασφαλίσει, οι οποίες στρατηγικές επειδή μπορεί να έχουν μακροπρόθεσμο χαρακτήρα να ξεπερνούν τα χρονικά όρια παραμονής των στελεχών της διοίκησης, αποκτούν σαν έννοια την δική τους υπόσταση. Το πιο σημαντικό μέσο εφαρμογής των στρατηγικών είναι αυτό της κατάρτισης ενός επιχειρηματικού πλάνου, το οποίο θα αποτελεί οργάνογραμμα και οδηγό ως προς τις τακτικές και ενέργειες που θα ακολουθηθούν και θα βοηθάει τα στελέχη και τους εργαζομένους να παραμένουν

συγκεντρωμένοι σε αυτό. Το επιχειρησιακό πλάνο μπορεί να αναφέρεται σε μεγάλης ευρείας προκλήσεις και να αναφέρεται ακόμη σε έργα και προγράμματα που αφορούν δεδομένα, όπως είναι το REACH2025: Σχέδιο Δράσης για την Ψηφιακή Οικονομία της Ιορδανίας, που παρουσιάστηκε στην τεχνολογική συνάντηση των κρατών της Μέσης Ανατολής και των Βορείων της Αφρικής το 2016 (απαρτίζεται κυρίως από Αραβικά Κράτη) [12]. Αυτό προϋποθέτει την επένδυση και ανάπτυξη υποδομών υποστήριξης της διαχείρισης δεδομένων και ανάδειξη των αντίστοιχων νομοθετικών ρυθμίσεων. Ακόμη, ένα αποτελεσματικό επιχειρηματικό πλάνο πρέπει να τηρεί ορισμένες προϋποθέσεις ώστε να είναι πετυχημένο και αποτελεσματικό. Συγκεκριμένα[39]:

- Ανάλυση εξωτερικού (ανταγωνιστές, κράτος, κοινωνία, φύση) και εσωτερικού (εργαζόμενοι, διαθέσιμοι πόροι, εργασιακά σωματεία) περιβάλλοντος. Η σημαντικότερη ειδοποιός διαφορά ότι το εσωτερικό περιβάλλον η εταιρεία έχει την δυνατότητα εύκολα να το επηρεάσει.
- Ανάλυση Δυνατών και Αδύνατων Σημείων καθώς και αναγνώριση ευκαιριών και Απειλών (SWOT Analysis).
- Κατανόηση της Αγοράς και των συστατικών της όπως, τα διαθέσιμα, ανταγωνιστές, πελάτες.
- Ενσωμάτωση της στρατηγικής στις επιχειρησιακές λειτουργίες και γενικά στην νοοτροπία των εργαζομένων.
- Εξασφάλιση των διαθέσιμων πόρων και ανάλυση της σχέσης κόστους και ωφέλειας
- Δυνατότητα ανανέωσης, τροποποίησης ακόμη και κατάργησης του.

4.2. Οικοσύστημα δεδομένων και διασύνδεση των εργαλείων

Πέρα από την κλασική επιχειρηματική προσέγγιση, σε ένα σύγχρονο οικοσύστημα δεδομένων υπάρχουν διάφορες οντότητες οι οποίες κινούνται σε ένα κοινό δίκτυο όπου συνεχώς εξελίσσονται και αλλάζουν. Ενδεικτικά των οντοτήτων, είναι τα δεδομένα σε διάφορες μορφές και οι πηγές τους, τα επιχειρησιακά πληροφοριακά συστήματα, οι επαγγελματίες των δεδομένων αλλά και οι χρήστες τους.[28] Για να μπορεί να καταστεί βιώσιμο και λειτουργικό έχει δημιουργηθεί ένας αγωγός δεδομένων στον οποίο έχουν

ενσωματωθεί τα κατάλληλα εργαλεία, με σκοπό την ομαλή ροή των δεδομένων τα οποία ακολουθούν ένα μονοπάτι, στο οποίο εισέρχονται ως ακατέργαστη πληροφορία και μέσα από μία επεξεργασία μετατρέπονται σε χρήσιμα και αξιόπιστα στοιχεία ώστε να τα εκμεταλλευτούν εκείνα τα άτομα με τέτοιο τρόπο ώστε να παραχθεί αξία.[4] Επομένως, η διασυνδεσιμότητα των επιμέρους μερών είναι κρίσιμη ώστε να το οικοσύστημα να είναι ενωμένο, συνεκτικό και παράλληλα ευέλικτο και αποδοτικό.[29] Οι πιο κύριες περιπτώσεις συνδεσιμότητας είναι οι παρακάτω:

α) Μέσω ενός συγκεκριμένου δικτύου υπολογιστών. Συνήθως, οι επιχειρήσεις εργάζονται σε ένα τοπικό δίκτυο(LAN) αλλά οι μεγαλύτεροι οργανισμοί μπορεί να έχουν δημιουργήσει δίκτυα πιο ευρείας χρήσης (WAN) ή ακόμη και μέσω διαδικτύου. Σε αυτό το δίκτυο οι διάφορες συσκευές, όπως οι υπολογιστές, τα τερματικά πληρωμών (PoS), διάφορες κινητές συσκευές παρουσιάζονται ως κόμβοι που επικοινωνούν μεταξύ τους. Αυτές οι συσκευές δύναται παράλληλα να διαθέτουν λογισμικά τα οποία να εκτελούν επιχειρησιακές λειτουργίες και παράλληλα να αποστέλλονται δεδομένα σε άλλους κόμβους μέσω συσκευών όπως είναι οι δρομολογητές (routers) και διακομιστές (servers), οι οποίοι κόμβοι καταναλώνουν τα δεδομένα για να εκτελέσουν ορισμένες διαδικασίες όπως προετοιμασία, επεξεργασία, ανάλυση και αποθήκευσή τους και μπορούν να χαρακτηριστούν ως σταθμούς εργασίας. Παράλληλα, αποθηκεύονται τα δεδομένα σε διάφορα αποθηκευτικά μέσα ανάλογα με τις ανάγκες και δυνατότητες της επιχείρησης. Συνήθως για να επιτευχθεί η δυναμική κλιμάκωση τα πρωταρχικά δεδομένα δεν αποθηκεύονται για καιρό ύστερα από την επεξεργασία τους. Επίσης, οι διαδικασίες που ακολουθούνται περιορίζονται στην εξαγωγή, επεξεργασία και αποθήκευση δηλαδή ένα πρότυπο ETL pipeline (Extract-Transform-Load), στο οποίο δεν ενεργοποιούνται άλλες διαδικασίες αυτόματα.

β) Μέσω ενός δικτύου υπολογιστικού νέφους. Η πιο σύγχρονη διασύνδεση ενός αγωγού δεδομένων η οποία προσφέρει στον οργανισμό απεριόριστες δυνατότητες επεξεργασίας και αποθήκευσης, αντιμετωπίζοντας αποτελεσματικά προβλήματα Μεγάλων Δεδομένων και την αξιοποίησή άλλων τεχνολογιών όπως του Διαδικτύου Πραγμάτων και της Μηχανικής Μάθησης. Βασίζεται στην ίδια φιλοσοφία λειτουργίας ωστόσο όλα τα επιμέρους μέρη είναι διασυνδεδεμένα σε ένα ιδιωτικό υπολογιστικό νέφος. Σε αυτό το δίκτυο επιτρέπεται η πρόσβαση από απομακρυσμένες συσκευές μέσω δικτυακών επαφών όπως

εφαρμογών ή ιστοσελίδων και το επίπεδο πρόσβασης αντιστοιχεί με το ρόλο που έχει ο κάθε χρήστης σε αυτό, εισάγοντας τα προσωπικά του διαπιστευτήρια. Η διανομή των ρόλων και τα όρια προσβασιμότητας γίνεται από έναν κεντρικό διαχειριστή που έχει την ευθύνη του δικτύου. Στο δίκτυο αυτό έχουν ενσωματωθεί λογισμικά για την διαχείριση δεδομένων και την λειτουργικότητα των συσκευών μεταξύ τους. Έτσι ο χρήστης που του επιτρέπεται να επεξεργαστεί ή να οπτικοποιήσει δεδομένα, μπορεί να εκτελέσει τις διαδικασίες αυτές και να αναρτήσει τα αποτελέσματα στο δίκτυο ώστε να τα δουν εκείνοι οι χρήστες που τους επιτρέπεται. Οι υποδομές, επιπρόσθετα, είναι απομακρυσμένες, όπως οι διάφοροι διακομιστές και αποθηκευτικά μέσα, και μέσω της τεχνολογίας του υπολογιστικού νέφους, οι δυνατότητες όλου του εξοπλισμού μπορεί να αξιοποιηθεί πλήρως ενιαία.

Η κύρια διαφορά μεταξύ του παραδοσιακού δικτύου και εκείνου του υπολογιστικού νέφους είναι ως προς την προσέγγιση του λογισμικού και συγκεκριμένα τα παραδοσιακά δίκτυα βασίζονται σε διακόπτες και δρομολογητές, ενώ το δίκτυο cloud βασίζεται σε εικονικό επίπεδο μέσω του επιπέδου ελέγχου.[37]

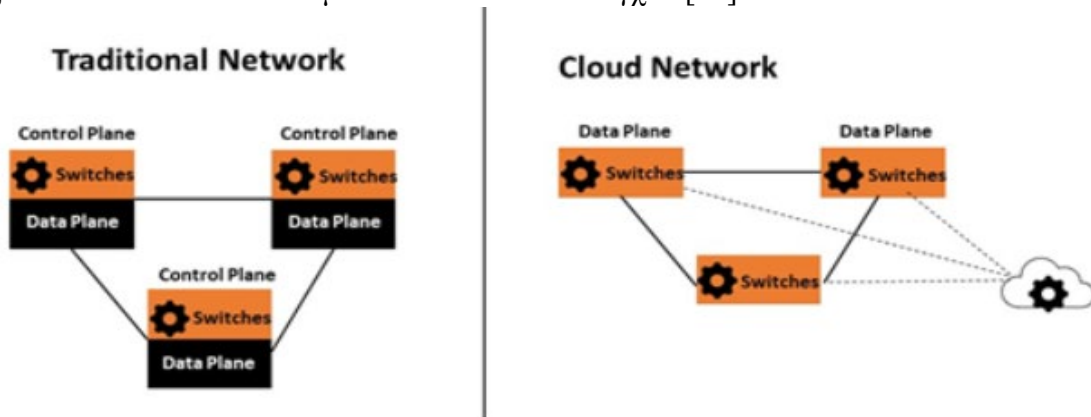
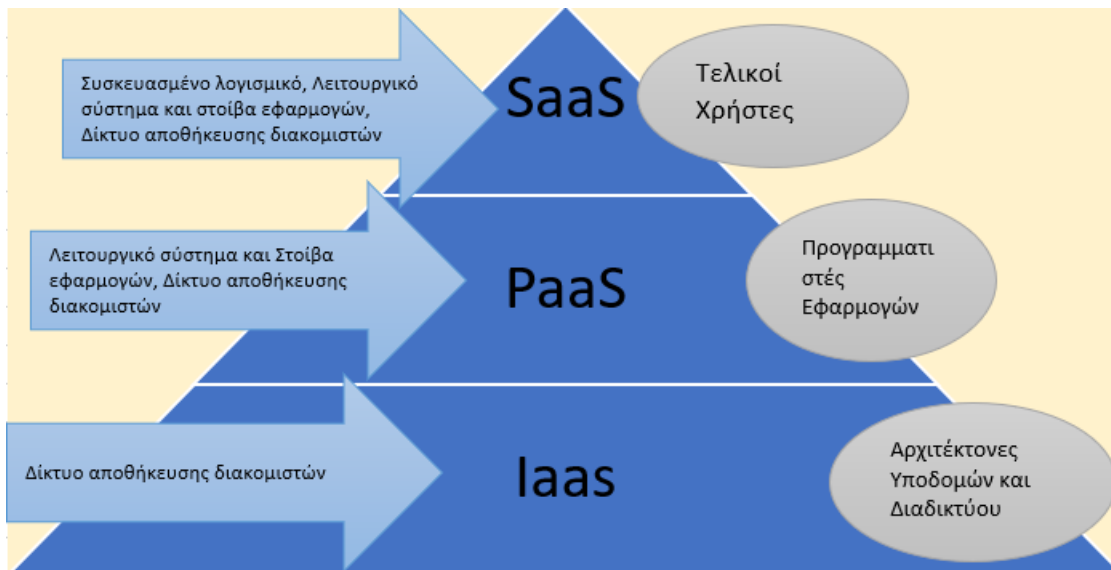


Fig: Traditional network and Cloud Network

Εικόνα 32. Παραδοσιακό δίκτυο έναντι δικτύου υπολογιστικού νέφους.[37] Source: Alibaba Cloud

γ) Ολοκληρωμένες πλατφόρμες διαχείρισης δεδομένων με χρήση υπολογιστικού νέφους, οι οποίες επιτρέπουν ανάλογα με τις ανάγκες του πελάτη να αναλύσει τα δεδομένα του και να τα μετατρέψει στην επιθυμητή μορφή, μέσα σε ένα διαδραστικό και εύκολο στην πλοήγηση ψηφιακό περιβάλλον. Σε αυτήν την πλατφόρμα μπορεί ο χρήστης να εισάγει

συνεχώς νέα στοιχεία είτε αυτόματα από τις πηγές δεδομένων που έχει ορίσει και συνδέσει είτε χειρωνακτικά. Ορισμένες τέτοιες πλατφόρμες έχουν κυκλοφορήσει από μεγάλες τεχνολογικές εταιρείες και έχουν γίνει απαραίτητα εργαλεία άλλων εταιρειών, οι οποίες χωρίς να επενδύσουν μεγάλα κεφάλαια στην διαχείριση των δεδομένων τους, έχουν αναθέσει, ουσιαστικά, αυτές τις εργασίες ουσιαστικά σαν μία μορφή outsourcing, ωστόσο την τελική ευθύνη την έχουν οι πελάτες-εταιρείες. Αυτή η προσέγγιση του υπολογιστικού νέφους θεωρείται παροχή πλατφόρμας ως υπηρεσία (PaaS), ενώ υπάρχουν άλλα δύο επίπεδα αυτού της παροχής λογισμικού ως υπηρεσίας απομακρύνοντας την διαχείριση των εφαρμογών από την πελάτη (SaaS), ενώ σε περίπτωση που ο πελάτης θέλει να έχει πλήρη έλεγχο των λειτουργιών αλλά έχει ανάγκη από αξιόπιστες υποδομές δεδομένων και τεχνολογίας καταφεύγει στην παροχή υποδομών ως υπηρεσία (IaaS) και είναι η περίπτωση που αναφέρθηκε στην προηγούμενη παράγραφο.



Εικόνα 33. Τα επίπεδα ανάθεσης υπηρεσιών υπολογιστικού νέφους σε άλλη εταιρεία. Source: www.medium.com "Types of Cloud Services" by IDM

4.3. Συστήματα διαχείρισης πόρων (Enterprise Resource Planning)

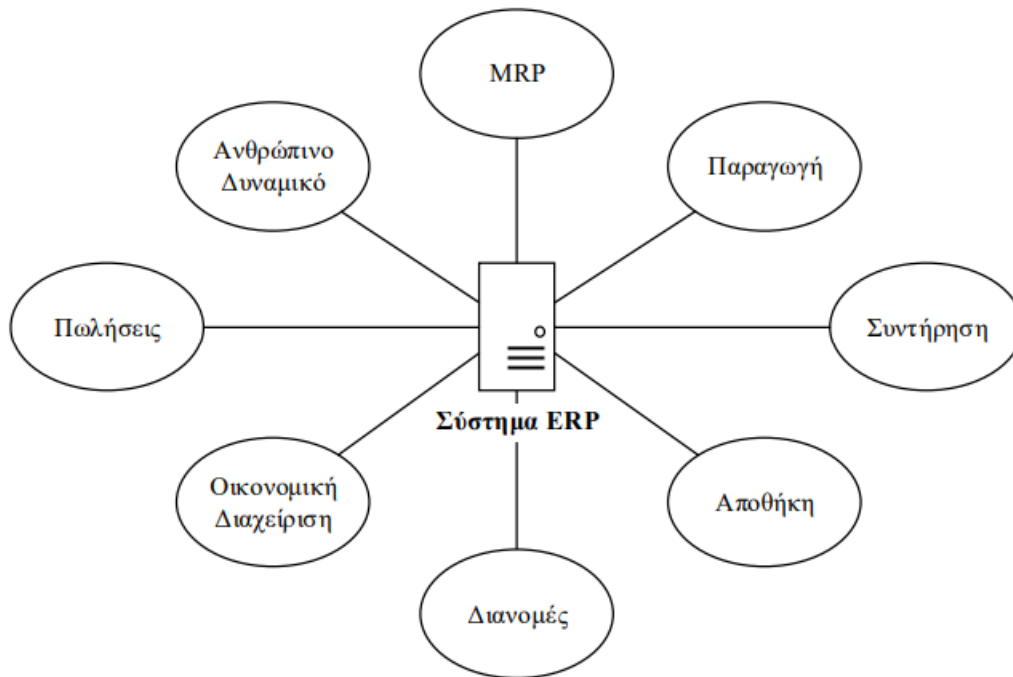
Τα Συστήματα Διαχείρισης Πόρων ή με το αγγλικό ακρωνύμιο τους, με οποίο είναι περισσότερο διαδεδομένα στις επιχειρήσεις ERP συστήματα, έχουν ως βασικό στόχο να διευκολύνει, να βελτιώσει έως και να αυτοματοποιήσει επιχειρησιακές διαδικασίες ως

προς τους τομείς της λογιστικής, προγραμματισμού παραγωγής και παραγγελιών και γενικότερα του χρηματοοικονομικού τομέα [40]. Ο συνδυασμός αυτών με τις τεχνικές της επιστήμης δεδομένων μπορεί να εξασφαλίσει την μέγιστη αξιοποίηση των συστημάτων και την απόσβεση των κεφαλαίων που δαπανήθηκαν σε αυτά γρήγορα. Αρχικά σημαντική είναι η εισαγωγή παλαιότερων δεδομένων πριν την εγκατάσταση ενός συστήματος ERP είτε αυτά ήταν ψηφιακά αποθηκευμένα είτε σε υλική μορφή. Σε αυτό το στάδιο μπορεί να απαιτηθεί μεγάλη ανθρώπινη προσπάθεια κυρίως ώστε να έρθουν όλα τα παρελθοντικά στοιχεία σε τέτοια μορφή η οποία θα είναι εύκολα αναγνώσιμη για το σύστημα όπως η εισαγωγή τους σε κοινά ψηφιακά υπολογιστικά φύλλα τα οποία τα περισσότερα τέτοιου είδους συστήματα μπορούν να τα δεχτούν χωρίς ιδιαίτερα προβλήματα.

Το ανώτερο έργο μπορούν να το αναλάβουν εργαζόμενοι από την ομάδα δεδομένων όπως υπάλληλοι Καταχώρησης Δεδομένων οι οποίοι έχουν την εμπειρία ώστε να επιτευχθεί σωστά και να αποφευχθούν ασυνέπειες και δυσλειτουργίες το οποίο δεν θα έθετε γερά θεμέλια στη χρησιμότητα του συστήματος[29]. Στην συνέχεια καθώς ενσωματώνεται στην λειτουργία της επιχείρησης, παράγονται πολλά δεδομένα που αφορούν πολλούς τομείς της εταιρείας, το οποίο ουσιαστικά για την επιστήμη δεδομένων σημαίνει ότι πολλές πηγές δεδομένων ενοποιήθηκαν σε μία μεγάλη κοινή πηγή, η οποία έχει την μορφή μίας σχεσιακής βάσης δεδομένων, στην οποία όλα τα δεδομένα μπορούν να συγκριθούν και έχουν συγκεκριμένους τύπους. Είναι απόλυτα κατανοητό, λοιπόν, την ευκολία που προσφέρει ένα τέτοιο γεγονός, δημιουργώντας την ευκαιρία για ταχύτερες και αποτελεσματικότερες διαδικασίες που προσφέρει η επιστήμη δεδομένων όπως, προγνώσεις για θέματα παραγωγής και αναπλήρωση αποθεμάτων, κατάρτιση χρονοδιαγραμμάτων εργασίας προσωπικού ή συντηρήσεις εξοπλισμού τα οποία λαμβάνουν υπόψη όλους τους υπόλοιπους παράγοντες και περιορισμούς, εισάγοντας και εξωτερικά στοιχεία στις δύσκολες εξισώσεις “αυτές” για μεγαλύτερη συσχέτιση με την αγορά και το κλίμα που επικρατεί [40].

Επίσης, είναι πιο εύκολα να εντοπιστούν μοτίβα και σχέσεις μεταξύ διάφορων μεταβλητών που επηρεάζονται από διαφορετικά τμήματα της εταιρείας τα οποία δεν έχουν καμία επαφή και η επιχείρηση να εκμεταλλευθεί ευκαιρίες ή να αντιμετωπίσει προβλήματα τα οποία δεν είχε αντιληφθεί[4]. Τα δεδομένα μπορεί να εξάγονται και να φορτώνονται για περαιτέρω ανάλυση και αξιοποίηση σε άλλα λογισμικά από το προσωπικό ή να υπάρχει

διασύνδεση με άλλα λογισμικά και να γίνεται αυτόματα, το οποίο αν και φαντάζει η πιο επιθυμητή λύση καθώς μπορεί να δημιουργηθούν προβλήματα συνδεσιμότητας μεταξύ των λογισμικών αλλά και τα ίδια τα εξαγόμενα δεδομένα των ERP συστημάτων να μην αναγνωρίζονται στο σύνολό τους από άλλα λογισμικά και να απαιτείται συνεχής έλεγχος και ανθρώπινη προσπάθεια για την εξασφάλιση της αξιοπιστίας και συνέπειας. Στην παρακάτω εικόνα φαίνεται η ροή των δεδομένων, κατά την οποία γίνεται εκκίνηση όταν ο πελάτης τοποθετεί μία εντολή/παραγγελία στο τμήμα πωλήσεων και τελειώνει με την διανομή της παραγγελίας [40].



Εικόνα 34. Ροή δεδομένων με επίκεντρο ένα ERP σύστημα [40] Πηγή: Σύγχρονα Πληροφοριακά Συστήματα Επιχειρήσεων 2015, του Πάνου Φιτσιλή

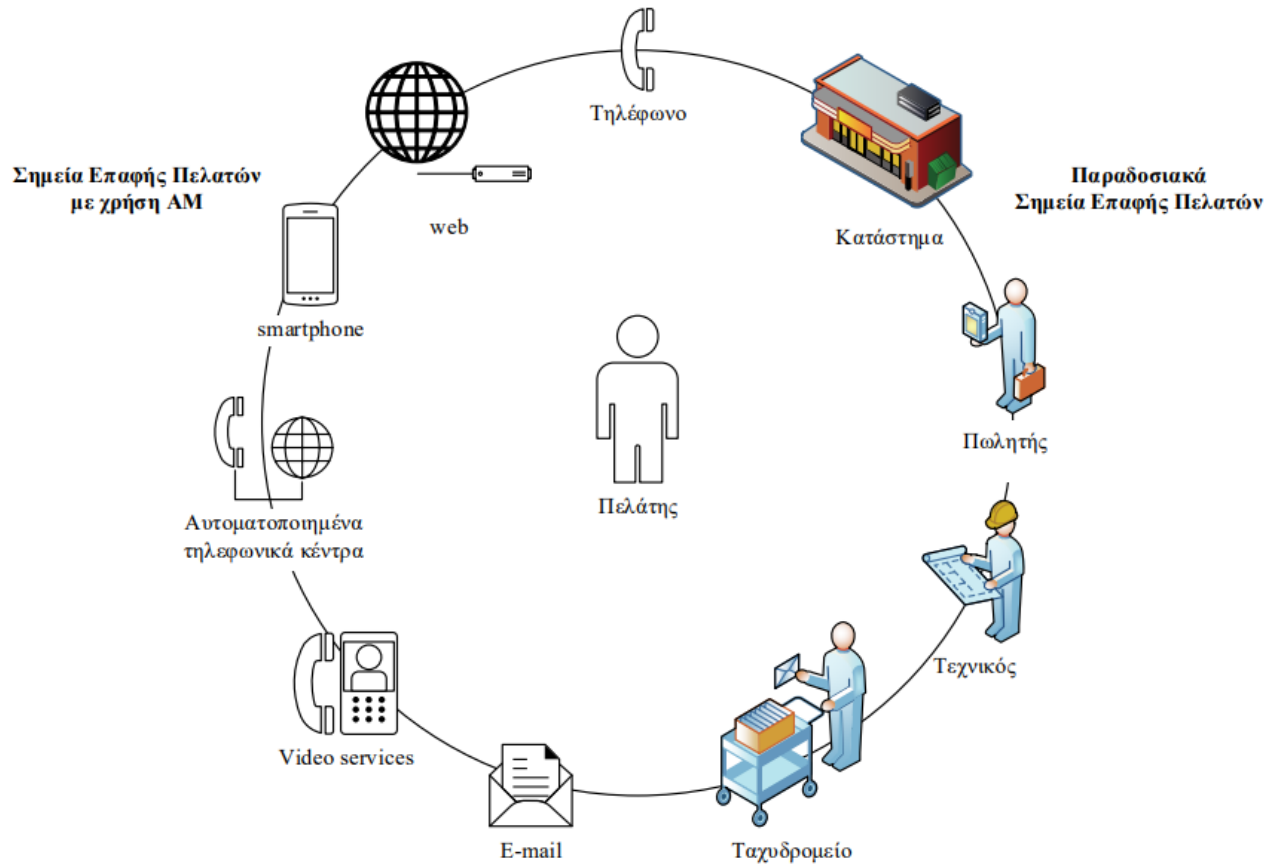
4.4. Συστήματα διαχείρισης πελατειακών σχέσεων (Customer relationship management)

Ένα άλλο πολύ διαδεδομένο σύστημα στην αγορά με το αγγλικό ακρωνύμιο CRM, το οποίο επικεντρώνεται στον πελάτη και χρησιμοποιείται κυρίως από τα τμήματα των πωλήσεων, marketing και εξυπηρέτησης πελατών. Βοηθάει στην αλληλοεπίδραση που έχει η εταιρεία με τους πελάτες της, ειδικά με τους μακροχρόνιους λόγω είτε προτίμησης στο προϊόν ή σύναψης σύμβασης συνεργασίας είτε υπογραφή συμβολαίου υπηρεσιών ή

συνδρομής. Στην βιβλιογραφία και σε διάφορα οικονομικά άρθρα προτείνονται ορισμένοι τύποι αυτών των συστημάτων και η κατηγοριοποίηση σχετίζονται με τον τρόπο που οι επιχειρήσεις τα αξιοποιούν. Συγκεκριμένα μία συνοπτική κατηγοριοποίηση που μπορεί να προταθεί είναι τρεις κατηγορίες για την χρήση των CRM συστημάτων[40]:

- Αναλυτική
- Λειτουργική
- Στρατηγική/Συνεργατική

Η πρώτη χρήση μπορεί να βρει έφορο πεδίο συνδυασμών με την επιστήμη δεδομένων. Και αυτό καθώς η προσέγγιση αυτή εστιάζει κατά κύριο λόγο την εύρεση μοτίβων πελατών και την κατανόηση της συμπεριφοράς των καταναλωτών σε σχέση με ορισμένες μεταβλητές και γεγονότα. Έτσι, με την ποσοτικοποίηση της επίδρασης ορισμένων γεγονότων, μπορεί εκ του αποτελέσματος να αποδειχθεί αν μία διαφημιστική ενέργεια κατάφερε να αποσβέσει το κόστος της αλλά και σε προγενέστερο στάδιο την ενίσχυση της εμπιστοσύνης των προγνώσεων για πιο επιτυχημένες δράσεις marketing. Επιπρόσθετα, ο εντοπισμός μοτίβων θα οδηγήσει σε συσταδοποίηση των πελατών και έτσι μπορεί να προβλεφθούν νωρίτερα τυχόν διαφυγή πελατών, όπως είχε συμβεί στο παρελθόν, λόγω ορισμένων έκτακτων γεγονότων και με εξατομικευμένες και βιώσιμες κινήσεις του τμήματος πωλήσεων, να διατηρήσει την συνεργασία. Από την άλλη πλευρά, μπορεί προσελκύσει πιο εύκολα μερίδιο της αγοράς από τους ανταγωνιστές, π.χ. προσφέροντας παρόμοιες παροχές σε πελάτες των οποίων το προφίλ ταιριάζει με ορισμένους ήδη υπάρχοντες πελάτες και απολαμβάνουν τις συγκεκριμένες παροχές.

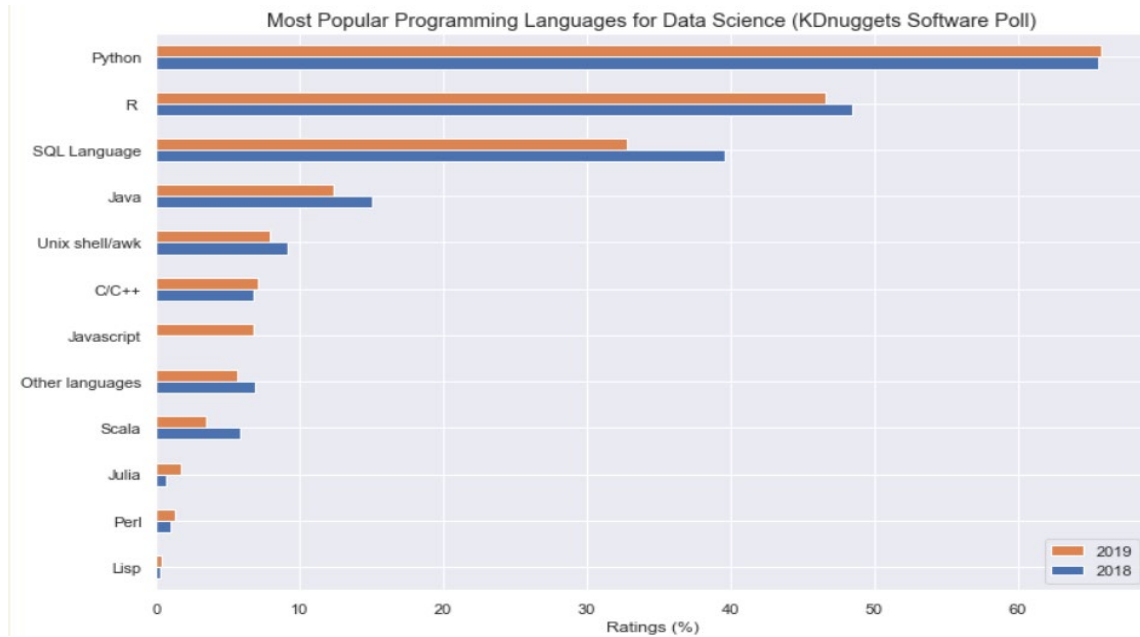


Εικόνα 35. Πηγές δεδομένων μέσω CRM *(AM: Αυτοματοποιημένο Μάρκετινγκ) [40] Πηγή: Σύγχρονα Πληροφοριακά Συστήματα Επιχειρήσεων 2015, του Πάνου Φιτσιλή

4.5. Γλώσσες προγραμματισμού στην επιστήμη δεδομένων

Βασικό κομμάτι της επιστήμης δεδομένων κατά την χρήση σε πραγματικά σενάρια είναι ο προγραμματισμός κυρίως για την δημιουργία αλγορίθμων οι οποίοι θα είναι αναγνώσιμοι από μηχανές και υπολογιστές. Επίσης, όλα τα ψηφιακά εργαλεία και πληροφοριακά συστήματα που έχουν αναφερθεί στην παρούσα εργασία έχουν κατασκευαστεί προγραμματιστικά με σκοπό την ανάλυση και αξιοποίηση των δεδομένων. Στην επιστήμη των δεδομένων, λόγω των δικών της ιδιαιτεροτήτων, συγκεκριμένες γλώσσες είναι πιο δημοφιλείς και προσεγγίζουν πιο αποτελεσματικά ίσως τα ζητήματα. Συγκεκριμένα, σε έρευνα που πραγματοποιήθηκε από μία από τις πιο γνωστές ιστοσελίδες για θέματα τεχνολογίας και επιστήμης δεδομένων την KDnuggets [41], ανάμεσα σε ενδιαφερόμενους

για το πεδίο αυτό αναδείχθηκαν ως πιο δημοφιλείς γλώσσες οι Python, R και SQL με μεγάλη διαφορά από τις υπόλοιπες όπως φαίνεται και στο παρακάτω διάγραμμα.



Εικόνα 36. Διάγραμμα με τις πιο δημοφιλείς γλώσσες προγραμματισμού στο πεδίο της επιστήμης των δεδομένων.

Πιο αναλυτικά για τις πιο δημοφιλείς ανωτέρω γλώσσες είναι τα εξής[10]:

- Python, η πιο διαδεδομένη γλώσσα προγραμματισμού στην επιστήμη αυτή αλλά και σε άλλους τομείς όπως την κατασκευή ιστοσελίδων ή βιντεοπαιχνιδιών. Μπορεί να εμπλουτίσει με πολλές δυνατότητες τα προγράμματα που δομούνται με αυτή λόγω της απλότητά της και την ευρεία χρησιμότητά της. Ακόμη δύναται να χρησιμοποιηθεί για όλες τα στάδια της διαχείρισης δεδομένων κατασκευάζοντας αντίστοιχα μοντέλα και δημιουργώντας αποτελεσματικά αλγορίθμους μηχανικής μάθησης.
- R, όχι τόσο δημοφιλής στο χώρο του προγραμματισμού όσο η Python, ωστόσο στο πεδίο των δεδομένων έχει πολύ σημαντική επιρροή και εφαρμογή, ειδικά σε έργα στατιστικής φύσης καθώς μέσω αυτής μπορεί να γίνει υψηλού επιπέδου χειρισμού των δεδομένων και να εξαχθούν οπτικοποιήσεις. Ταυτόχρονα μπορούν και μέσω αυτής να αναπτυχθούν αλγόριθμοι μηχανικής μάθησης με ευρεία εφαρμογή σε τομείς η οικονομία, στατιστική και για ακαδημαϊκούς σκοπούς.
- SQL, αποτελεί γλώσσα υπολογιστών ωστόσο δεν χρησιμοποιείται για προγραμματισμό καθώς μέσω αυτής δεν μπορούν να παραχθούν προγράμματα και

- λογισμικά. Ουσιαστικά είναι μια γλώσσα ερωτοαπαντήσεων, η πιο γνωστή για αυτό το σκοπό και αναπτύχθηκε ώστε να μπορεί γίνει εύκολη η διαχείριση των Σχεσιακών Συστημάτων Διαχείρισης Βάσεων Δεδομένων. Λόγω του γεγονότος ότι τα περισσότερα δεδομένα είναι αποθηκευμένα σε αποθηκευτικά μέσα σχεσιακού τύπου, η γλώσσα αυτή αποτελεί αναπόσπαστο κομμάτι της επιστήμης δεδομένων.
- Java, μία γλώσσα προγραμματισμού με ευρεία χρήση πέρα της επιστήμης δεδομένων. Λόγω της ταχύτητας που προσφέρει προγραμματιστικά σε σχέση με τις ανωτέρω γλώσσες για εισαγωγή και εξαγωγή δεδομένων, ιδιαίτερα χρήσιμη για διαχείριση Big Data και για πιο εξεζητημένες και δύσκολες υπολογιστικά εφαρμογές της Επιστήμης Δεδομένων όπως η Βαθιά Μάθηση ώστε να αποφεύγεται η επιβάρυνση όσο γίνεται υπολογιστικής δυναμικής.
 - Awk, είναι μια γλώσσα σεναρίων ενεργειών που χρησιμοποιείται για τον χειρισμό δεδομένων και τη δημιουργία αναφορών. Επιτρέπει στο χρήστη να χρησιμοποιεί μεταβλητές, αριθμητικές συναρτήσεις, συναρτήσεις συμβολοσειρών και λογικούς τελεστές. Αν και περίπλοκη μέσω αυτής μπορούν να γραφτούν μικροί κώδικες για προγράμματα με υψηλή απόδοση.

4.6. Εργαλεία διαχείρισης δεδομένων

Στην αγορά του κλάδου αυτού, κυκλοφορούν διάφορα λογισμικά τα οποία εξειδικεύονται σε διάφορες διεργασίες της διαχείρισης των δεδομένων, όπως την εξόρυξη, οπτικοποίηση, επεξεργασία ή οργάνωση, συνθέτοντας ένα μείγμα επιστήμης δεδομένων και επιχειρηματικής ευφυΐας το οποίο δημιουργεί αξία. Διάφορες τεχνολογίες και πρακτικές συνδυάζονται για την εκτέλεση των ανωτέρω διεργασιών, υποστηριζόμενες με τον κατάλληλο υλικοτεχνικό εξοπλισμό, σχηματίζοντας έτσι σε μεγάλο βαθμό τον αγωγό δεδομένων.[4] Για παράδειγμα, ένα έξυπνο κλειστό σύστημα τηλεόρασης (AI CCTV) σε ένα εμπορικό κέντρο, μπορεί να διαθέτει αισθητήρες και κάμερες, οι οποίες θα καταγράφουν συνεχώς τους χώρους και στην συνέχεια μέσω ενός αλγοριθμικού μοντέλου βαθιάς μάθησης θα αναγνωρίζει τους ανθρώπους, πιθανόν από τα χαρακτηριστικά του προσώπου τους, τους οποίους θα μπορεί να ταξινομήσει σε “υποκλάσεις” όπως γυναίκα,

άντρα ή παιδιά. Μέσω της τεχνολογίας του Διαδικτύου των Πραγμάτων και της Επιστήμης Δεδομένων, είναι δυνατόν να είναι γνωστό σε πραγματικό χρόνο ο αριθμός των επισκεπτών και σε περίπτωση που κάποιος από τους αισθητήρες, π.χ. θερμοκρασίας, καταγράψει ασυνήθιστες τιμές τότε μπορεί να προειδοποιήσει έγκαιρα το σύστημα το γραφείο ασφαλείας ακριβώς που υπάρχει το πρόβλημα και να απομακρυνθούν οι άνθρωποι που βρίσκονται εκεί από την ασφαλέστερη διαδρομή χωρίς μεγάλη ανθρώπινη προσπάθεια[27].

Παράλληλα, τα δεδομένα αποστέλλονται μέσω τοπικού δικτύου στο σύστημα ασφαλείας από το οποίο, αποθηκεύονται οι μετρήσεις των επισκεπτών ανά χρονική περίοδο, τα οποία θα μπορούν να χρησιμοποιηθούν για ενημέρωση της διοίκησης για την πελατεία του εμπορικού κέντρου ως προς τον αριθμό και τον τύπο της (π.χ. πότε επισκέπτονται τον χώρο περισσότερο τα παιδιά) αλλά και να συσχετιστούν και με άλλα μεγέθη δηλαδή τις πωλήσεις των καταστημάτων και να προκύψει πόσα χρήματα ξοδεύει κατά μ.ό.. ένας επισκέπτης του καταστήματος και πιο συγκεκριμένα για παράδειγμα ένας άντρας που φοράει γυαλιά.

Στο ανωτέρω παράδειγμα, αναγνωρίζεται η ανάγκη για ύπαρξη ενός μηχανικού δεδομένων με γνώση αποκεντρωμένων συστημάτων και λειτουργίας διαφορετικών συσκευών για βέλτιστη συνδεσιμότητα των διάφορων λογισμικών ή ενός αναλυτή επιχειρηματικής ευφυίας όπου θα διαθέτει τις κατάλληλες ικανότητες ώστε να εντοπίσει και να ερμηνεύσει κατάλληλα τα πιθανά μοτίβα με χρήση των κατάλληλων εργαλείων διαχείρισης και επεξεργασίας δεδομένων, συσχετίζοντας αυτά με πιθανούς παράγοντες και στην συνέχεια, να επικοινωνήσει, τα μηνύματα που χρειάζεται στην διοίκηση με τέτοιο τρόπο ώστε να είναι κατανοητά και αξιοποιήσιμα. Αυτό μπορεί να επιτευχθεί κατασκευάζοντας κατάλληλους πίνακες και διαγράμματα μέσω ενός κατάλληλου λογισμικού οπτικοποίησης δεδομένων. Προφανώς, οι επιχειρήσεις θα ήθελα να είχαν στην διάθεση τους, τέτοιου είδους εργαλεία ωστόσο λόγω περιορισμών όπως οικονομικών ή μη γνώσης για την ύπαρξή τους ή τον τρόπο λειτουργίας αυτών από τους διευθύνοντες ενός οργανισμού, δεν επενδύουν στις τεχνολογίες αυτές και κατά επέκταση δεν δημιουργείται και η ανάγκη για πρόσληψη ατόμων που θα τις γνωρίζουν και να τις εκμεταλλεύονται[15].

Μερικά παραδείγματα από τα πιο δημοφιλή λογισμικά για την εκτέλεση των παραπάνω διεργασιών τα οποία κυριαρχούν στην αγορά, είναι τα παρακάτω:

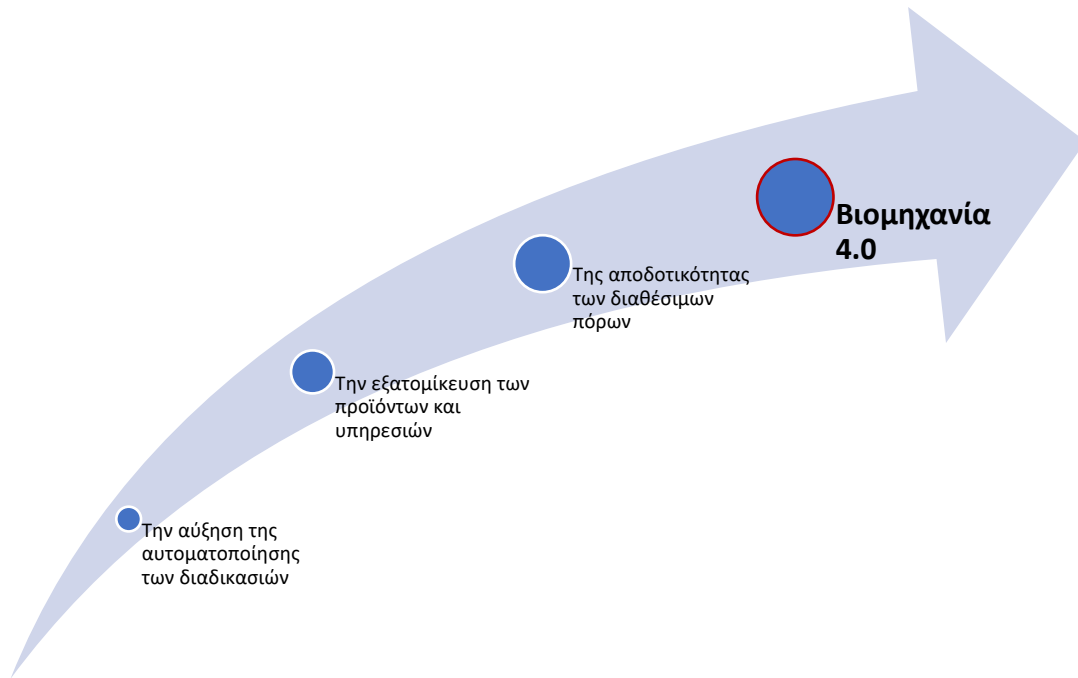
- Microsoft Power BI. Στοχεύει στην ανάλυση, στην οπτικοποίηση και στην εξόρυξη δεδομένων με σκοπό την υποβολή εκθέσεων και στην παροχή επιχειρηματικής πληροφόρησης. Λαμβάνει δεδομένα από διάφορες βάσεις και άλλες πηγές ώστε να κατασκευαστούν διαγράμματα, πίνακες και αναφορές.
- Tableau. Είναι εργαλεία οπτικοποίησης και ανάλυσης δεδομένων μέσω μίας πιο διαισθητικής προσέγγισης, όπως δημιουργία ερωτημάτων φυσικής γλώσσας και χρήση πολλαπλών φίλτρων. Είναι γενικότερα ευπροσάρμοστο στις διάφορες πηγές δεδομένων και μπορεί να υποστηρίξει κατανοητά την λήψη αποφάσεων.
- Looker (Google). Είναι μία επιχειρησιακή πλατφόρμα που στηρίζεται στο διαδίκτυο με σκοπό την διαχείριση και εξερεύνηση των δεδομένων. Μέσω αυτής μπορούν οι χρήστες μεταξύ τους να συνεργαστούν και να ανταλλάξουν δεδομένα και αναλύσεις, στοχεύοντας στην καλύτερη κατανόηση από πλευράς στελεχών στην οργάνωση και λειτουργίας της επιχείρησης.
- Oracle Analytics Cloud. Στηρίζεται στην Τεχνητή Νοημοσύνη το οποίο της εξασφαλίζει μεγάλες δυνατότητες ανάλυσης των δεδομένων, δημιουργώντας υψηλής ποιότητας αναφορές και αναλύσεις. Οι χρήστες μπορούν να λαμβάνουν αυτές την αναφορές στην επιθυμητή μορφή που θέλουν ώστε να προβούν σε ακριβείς προβλέψεις και να καταστρώσουν επιχειρηματικά σχέδια.
- MicroStrategy. Είναι μία πλατφόρμα μηχανικής μάθησης που χρησιμοποιούν διάφοροι οργανισμοί ώστε να διαχειρίζονται εύκολα τα δεδομένα τους και οι χρήστες της μπορούν να έχουν συνεχή πληροφόρηση με ταυτόχρονη συνεργασία όλων των συμμετεχόντων. Επίσης, οι χρήστες έχουν την δυνατότητα να συνδέονται από διάφορες απομακρυσμένες συσκευές όπως το κινητό τους, ενισχύοντας έτσι την διασύνδεση των δεδομένων και τα κατανεμημένα επιχειρησιακά συστήματα.

5. Συμπεράσματα και προτάσεις

5.1. Συμπεράσματα

Η διείσδυση της επιστήμης των δεδομένων σε όλο το εύρος της κοινωνίας είναι εμφανής και με το πέρασμα του χρόνου θα αλλάξει τον τρόπο που αντιλαμβάνεται ο άνθρωπος τον κόσμο. Στην παρούσα εργασία έγινε προσπάθεια ανάδειξης της μεγάλης επιρροής που θα διαδραματίσει στην σύγχρονη βιομηχανική επανάσταση, Βιομηχανία 4.0, την οποία θα υποστηρίξουν κατά κύριο λόγο οι τεχνολογίες της Τεχνητής Νοημοσύνης και του Διαδικτύου των Πραγμάτων οι οποίες για να επιτευχθούν απαιτείται αποτελεσματική εκμετάλλευση των δεδομένων μέσω των τεχνικών της επιστήμης αυτής. Ήδη η επιχειρηματικότητα έχει αλλάξει ο τρόπος που προσεγγίζεται καθώς ολοένα και περισσότερο κατευθύνεται προς τρεις κύριες κατευθύνσεις:

- α) Την αύξηση της αυτοματοποίησης των διαδικασιών, μειώνοντας έτσι τον χρόνο ανταπόκρισης των διαδεχόμενων ενεργειών και τον ανθρώπινο παράγοντα στη σχεδίαση και εκτέλεση με αποτέλεσμα την ελάττωση σφαλμάτων λόγω φυσιολογικής ανθρώπινης απροσεξίας.
- β) Την εξατομίκευση των προϊόντων και υπηρεσιών σύμφωνα με τις πραγματικές ανάγκες της κοινωνίας, αυξάνοντας την ικανοποίηση των καταναλωτών και βιοτικό τους επίπεδο χωρίς ταυτόχρονη οικονομική επιβάρυνση.
- γ) Της αποδοτικότητας των διαθέσιμων πόρων μέσω της βέλτιστης σχεδίασης παραγωγής και αναπλήρωσης αποθεμάτων, την κατανομή προσωπικού αποτελεσματικά και γενικότερα της αύξησης της παραγωγικότητας ανά διαθέσιμο επενδυτικό κεφάλαιο, μειώνοντας το τελικό κόστος.



Εικόνα 37. Ο δρόμος για την μετάβαση στην Βιομηχανία 4.0

Οι ανωτέρω αλλαγές θα ξεπεράσουν τα όρια των επιχειρήσεων και θα περάσουν σε όλο το εύρος της κοινωνίας, αλλάζοντας έτσι τον τρόπο που αντιλαμβανόμαστε την ανθρώπινη κοινωνία.

Σε ακαδημαϊκό επίπεδο εδώ και πολλά χρόνια η επιστήμη δεδομένων έχει αποκτήσει αυτόνομη πορεία από άλλες επιστήμες όπως εκείνων των μαθηματικών, της στατιστικής και των υπολογιστών. Αυτή η ανάγκη έχει αναδειχθεί και στην οικονομία για την δημιουργία αυτόνομου κλάδου διαχείρισης δεδομένων ο οποίος προσφέρει στην συνολική παραγωγική διαδικασία των χωρών. Η Ελλάδα δεν παρουσιάζει ακόμη την ίδια ανάπτυξη σε αυτόν τον κλάδο αν και αυτό ίσως να οφείλεται εξαιτίας της μακροχρόνιας οικονομικής και κοινωνικής κρίσης που επικρατούσε σε αυτήν, τα ίδια χρόνια όπου λάβανε χώρα η διερεύνηση του κλάδου πέρα από τις πολύ υψηλά προηγμένες τεχνολογικές περιοχές[13]. Ωστόσο πολλές χώρες, επιχειρήσεις και οργανισμοί επένδυσαν σε αυτόν τον κλάδο και γενικότερα στην ψηφιοποίηση για να καταφέρουν να ξεφύγουν από την οικονομική δυσπραγία και να ξεχωρίσουν από το υπόλοιπο σύνολο αποκτώντας ανταγωνιστικό πλεονέκτημα. Τέτοιο παράδειγμα αποτελεί η χώρα της Ιορδανίας[12] που προσπαθεί τα τελευταία χρόνια να μετασχηματιστεί ψηφιακά και να αναδειχθεί στο τομέα

της ψηφιοποίησης στην περιοχή και στις αραβικές χώρες πρωτοπόρος. Λόγω των περιορισμένων φυσικών πόρων που διαθέτει η χώρα και μην την πρόσβαση στη θάλασσα, η τεχνολογική ανάπτυξη φαντάζει μία ρεαλιστική διέξοδο της χώρας αυτής από τα χαμηλότερα επίπεδα οικονομικής ανάπτυξης τις περιοχής.

Έτσι έχει καθιερωθεί ήδη μία προτυποποίηση των διαδικασιών της διαχείρισης των δεδομένων σε επίπεδο κρατών ή μεγάλων οργανισμών όπως συμβαίνει στις περισσότερες εργασίες, η οποία περιλαμβάνει βήματα για ευκολότερη κατανόηση και εφαρμογή, τα οποία υποστηρίζουν νέοι εργασιακοί ρόλοι των οποίων τα καθήκοντα είναι προσανατολισμένα σε κάθε μεμονωμένο στάδιο. Προφανώς η επένδυση των επιχειρήσεων για την δημιουργία τμημάτων μέσα σε αυτές ενέχει κίνδυνο να μην είναι επιτυχής ωστόσο ολοένα και περισσότερο θα γίνεται ολοένα και περισσότερο αναγκαίο. Όποιοι οργανισμοί καταφέρουν να ευθυγραμμιστούν με τις νέες απαιτήσεις θα επιβιώσουν και θα αναπτυχθούν και οι υπόλοιποι θα αρχίσουν να χάνουν έδαφος συνεχώς. Ήδη το σύγχρονο τεχνολογικό περιβάλλον μπορεί εύκολα να υποστηρίξει ένα αποτελεσματικό οικοσύστημα δεδομένων χωρίς να απαιτούνται μεγάλα κεφάλαια για έρευνα και εύρεση καινοτομιών καθώς υπάρχουν πολλά παραδείγματα επιτυχίας και επιλογές στην αγορά για τις ανάγκες και δυνατότητες της κάθε εταιρείας.

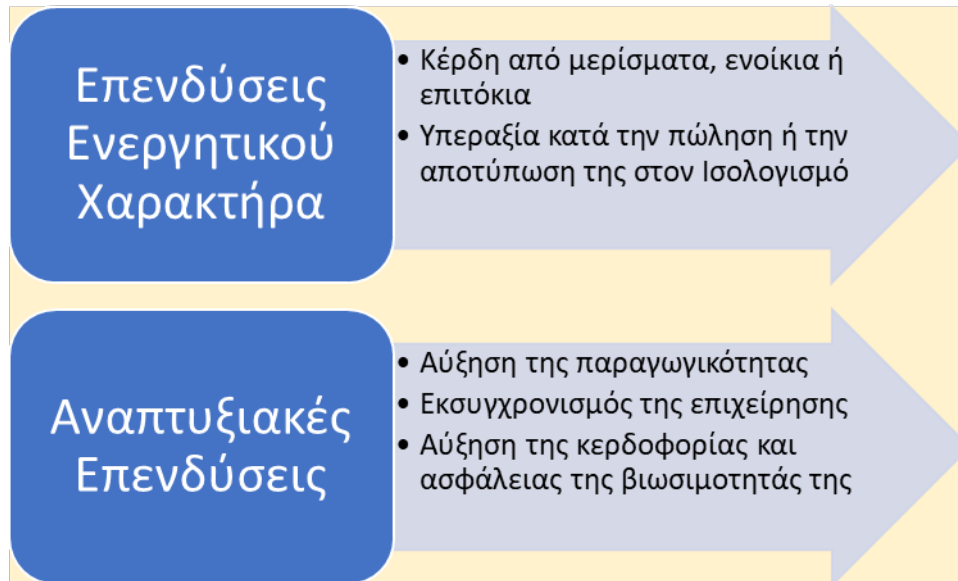
5.2. Προτάσεις

Όλες επενδύσεις που πραγματοποιούν οι επιχειρήσεις περιέχουν κάποιο κίνδυνο αποτυχίας με δυσάρεστες συνέπειες, όπως απώλεια του επενδυτικού κεφαλαίου ή να προκληθούν προβλήματα στις ήδη επιχειρησιακές λειτουργίες. Έτσι, οι επιχειρήσεις πριν προβούν στην κάθε επένδυση πρέπει να εκτιμήσουν σωστά τις πιθανότητες επιτυχίας ή αποτυχίας καθώς και να προετοιμαστούν κατάλληλα, εξασφαλίζοντας τα προαπαιτούμενα για την επιτυχία της επένδυσης[38]. Οι επενδύσεις μπορούν να διακριθούν σε αρκετές κατηγορίες οι οποίες θα μπορούσαν να τις συνοψιστούν σε δύο βασικές:

- Η απευθείας τοποθέτηση χρηματικών κεφαλαίων σε ορισμένες ευκαιρίες που υπόσχονται επιστροφή των επενδυόμενων κεφαλαίων με υπεραξία, για

παράδειγμα αγορά μετοχών, ομολόγων ή ακίνητης περιουσίας. Ουσιαστικά αλλάζει μορφή το Ενεργητικό το κομμάτι της περιουσίας της επιχείρησης.

- Οι δαπάνες που αφορούν στην ενίσχυση της καινοτομίας και βελτιώσεις των υπάρχουσών διαθέσιμων πόρων (ανανέωση εξοπλισμού, ψηφιοποίηση, εκπαίδευση προσωπικού, πρόσληψη νέων επαγγελματιών) με σκοπό την αύξηση της παραγωγικότητας και κατά επέκταση των κερδών της επιχείρησης.



Εικόνα 38. Βασικές επιλογές επένδυσης μίας επιχείρησης

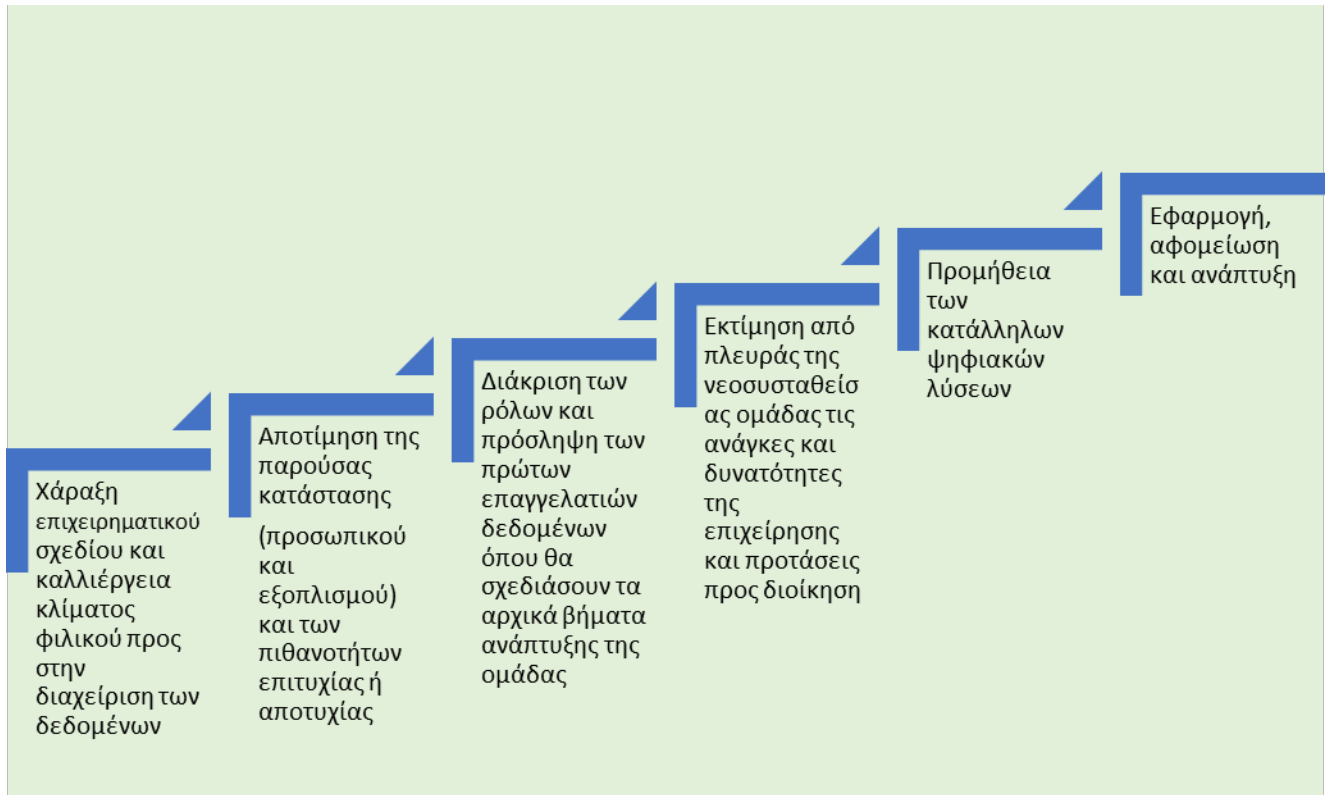
Όπως είναι εύκολα κατανοητό η δημιουργία μίας ομάδας διαχείρισης δεδομένων ή περαιτέρω εμβάθυνση της επιστήμης δεδομένων σε έναν οργανισμό θα άνηκε στην δεύτερη ομάδα επενδύσεων και είναι σημαντικό η διοίκηση να καλλιεργήσει από νωρίς ένα κλίμα ότι η επένδυση αυτή έχει ως σκοπό να βελτιώσει την επιχειρησιακή λειτουργία και να αυξήσει τελικά την κερδοφορία, με σκοπό όλοι οι εργαζόμενοι να υιοθετήσουν πιο πρόθυμα τις νέες πρακτικές[15]. Παράλληλα, η προσπάθεια αυτή θα απαιτήσει χρηματοδότηση για νέο εξοπλισμό, ψηφιακά εργαλεία, προσλήψεις νέων υπαλλήλων καθώς και την εκπαίδευση των ήδη εργαζομένων άρα είναι σημαντικό οι αλλαγές να αφομοιωθούν εύκολα από τον οργανισμό. Σε αυτό θα βοηθούσε η ένταξη των αλλαγών αυτών στο επιχειρησιακό πλάνο της εταιρείας ώστε να υπάρχει η απαραίτητη μεθοδικότητα κατά την σταδιακή εφαρμογή τους.

Στην συνέχεια πρέπει καθοριστούν οι ρόλοι των εργαζομένων ανάλογα με τους παράγοντες που έχουν αναφερθεί στο υποκεφάλαιο 2.2. αυτής της εργασίας[29]. Με δεδομένο ότι ο προϋπολογισμός της κάθε εταιρείας είναι συγκεκριμένος είναι πολύ σημαντική η αξιοποίηση του υπάρχοντος ανθρώπινου δυναμικού καθώς των υποδομών που υπάρχουν στην εταιρείας με ταυτόχρονη πρόσληψη μόνο των απαραίτητων επαγγελματιών, με πρώτο από όλους έναν επιστήμονα δεδομένων ο οποίος θα αρχίσει να διαμορφώνει την ομάδα ανάλογα με τις ανάγκες αλλά και τις δυνατότητες με πρώτες προσθήκες έναν μηχανικό και έναν αναλυτή δεδομένων ώστε να σχηματιστεί μία απλουστευμένη πολυεπίπεδη ομάδα που έχει αναφερθεί και προηγουμένως.

Εξασφαλίζοντας την πρωταρχική ομάδα η οποία θα χαράζει και τον αρχικό σχεδιασμό, θα γίνουν και οι ανάλογες προτάσεις προς την διοίκηση για αγορά νέων συστημάτων που θα υποστηρίζει το έργο τους και ταιριάζουν με το περιβάλλον της επιχείρησης, διαμορφώνοντας ένα νέο. Λόγω των διαφορετικών τμημάτων πρέπει να πρέπει να διαθέτει το νέο περιβάλλον ένα καλά κατανοημένο υπολογιστικό σύστημα (edge computing) ώστε η ροή των δεδομένων να πραγματοποιείται αβίαστα και να υπάρχει συνέπεια ως προς την αποθήκευση και την διαχείρισή τους. Επίσης, αυτό το περιβάλλον πρέπει να αναπτυχθεί με ευελιξία (agile methods) για κάθε κομμάτι του αγωγού δεδομένων που θα δημιουργηθεί, από το τεχνικό επίπεδο ακολουθώντας τις υποδείξεις των άλλων δύο υπομάδων. Τα κυριότερα βήματα που πρέπει να γίνουν περιληπτικά ώστε να εξασφαλιστεί τεχνολογική επάρκεια και να ακολουθείται μία σωστή διαδρομή είναι τα εξής:

- Η συγκέντρωση των δεδομένων από όλο το κατανοημένο υπολογιστικό σύστημα (Συλλογή) σε βάσεις δεδομένων προσπαθώντας όσο το δυνατόν αυτοματοποιημένα (Αποθήκευση). Υπάρχουν διάφορα εργαλεία για το σκοπό αυτό όπως το Apache Hadoop που αναφέρθηκε στην περίπτωση της εταιρείας Netflix[36], το οποίο επειδή είναι ανοικτού κώδικα, η κάθε ομάδα δεδομένων μπορεί να επιλέξει μία από τις διάφορες εκδόσεις που κυκλοφορούν και την προσαρμόσει στις ανάγκες της. Επιπρόσθετα, λόγω ότι είναι ανοικτού κώδικα πολλές εκδόσεις του Apache Hadoop που κυκλοφορούν είναι δωρεάν και αξιόπιστες, γεγονός πολύ σημαντικό στην αρχή που διαμορφώνεται το πρόγραμμα αυτό.

- Σημαντικό, εδώ, είναι η επιλογή του τύπου βάσεων δεδομένων, όπου ουσιαστικά οι επιλογές κυρίως είναι οι σχεσιακές (πιο εύκολες στην χρήση) και οι μη σχεσιακές (ανταποκρίνονται καλύτερα σε προβλήματα Μεγάλων Δεδομένων) καθώς των συστημάτων διαχείρισής τους.
- Στην συνέχεια πρέπει να αποφασιστεί πως θα επιτυγχάνεται η προετοιμασία των δεδομένων, δηλαδή αν τα μέλη της ομάδας προσπαθήσουν να φτιάξουν προγραμματιστικά από την αρχή ένα απλό μοντέλο επεξεργασίας δεδομένων ή να προμηθευτούν ένα ήδη εγγυημένο δημοφιλές λογισμικό, όπως το MicroStrategy, ανάλογα με τις ικανότητες και το διαθέσιμο προϋπολογισμό (Επεξεργασία-Ανάλυση).
- Η ανάλυση, ομοίως μπορεί να γίνει είτε από στατιστικά μοντέλα τα οποία έχουν γραφεί με κώδικες της ίδια της ομάδας είτε μέσω εργαλείων ανάλυσης π.χ. Oracle Analytics Cloud ή Anaconda(Ανάλυση). Στην αρχή όπου οι απαιτήσεις δεν θα είναι μεγάλες, για πιο εύκολες αλλά και αξιόπιστες λύσεις μπορούν να χρησιμοποιηθούν τα συστήματα διαχείρισης δεδομένων των βάσεων ή μέσω υπολογιστικών φύλλων (Microsoft Excel).
- Υπάρχουν επίσης διάφορα εργαλεία Επιχειρηματικής Ευφυΐας που μπορούν να ενταχθούν στην επιχείρηση, όπως Microsoft Power BI, Tableau ή Looker μπορούν να προσφέρουν διασύνδεση πολλών χρηστών και υποστηρίζοντας κατά την εκτέλεση των λειτουργιών τους διάφορα άλλα τμήματα, ενισχύοντας την συνολική αξία της διαχείρισης των δεδομένων (Αξιοποίηση). Συνήθως, όλα αυτά τα εργαλεία λειτουργούν ως επιχειρησιακές πλατφόρμες όπου μέσα σε αυτές γίνονται ανταλλαγή στοιχείων, αναφορών και παρουσιάσεων. Ακόμη μέσω αυτών μπορούν να αποτυπωθούν πιο παραστατικά σε εικόνες και διαγράμματα μεγάλες ποσότητες δεδομένων για καλύτερη κατανόηση από τρίτους (Οπτικοποίηση).
- Πρέπει να επισημανθεί ότι η επιλογή των ανωτέρω πέρα από το κόστος και τις προτιμήσεις των εργαζομένων, πρέπει να λαμβάνεται υπόψη και η συμβατότητα των διάφορων λογισμικών. Προτείνεται, δηλαδή, η προμήθεια ψηφιακών διάφορων εργαλείων για όλες τις φάσεις της διαχείρισης των δεδομένων, από ίδιο πάροδο ή έστω παρόμοιες τεχνολογίες.

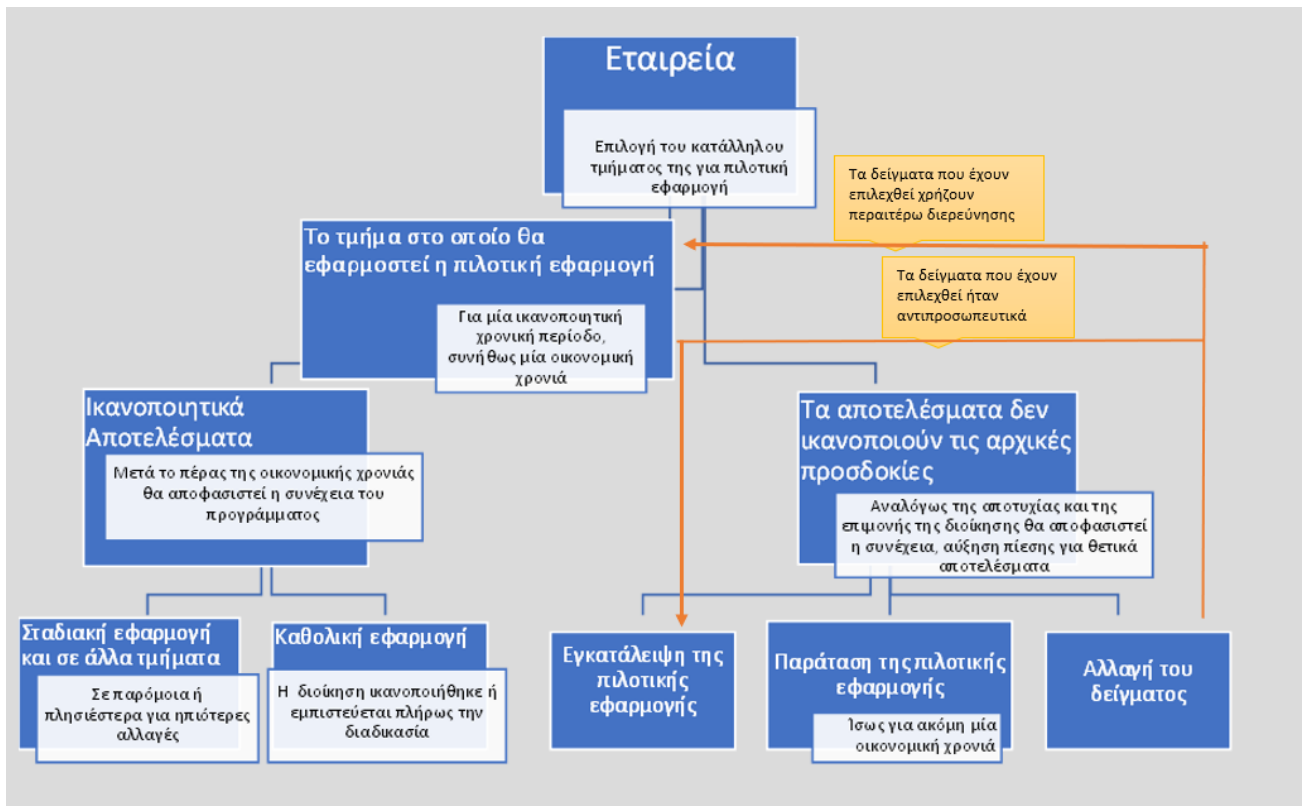


Εικόνα 39. Σταδιακή ανάπτυξης της ομάδας διαχείρισης δεδομένων εντός της επιχείρησης

Αφού πάρει μία πρωταρχική μορφή η ομάδα δεδομένων και γίνουν οι κατάλληλες προσθήκες, μπορεί σταδιακά να μεγαλώσει και να επεκταθεί εντός του οργανισμού. Είναι απαραίτητο η προσθήκη στο ανθρώπινο δυναμικό πιο εξειδικευμένων στελεχών όσο μεγαλώνει η ομάδα. Κάθε αλλαγή ή καινοτόμο πρόγραμμα πέρα από τις θετικές επιδράσεις, μπορεί να επιφέρει και αρνητικά αποτελέσματα. Αναφέρθηκε ότι πέρα από την απώλεια κεφαλαίων από μία άστοχη επένδυση, μπορούν να προκύψουν και προβλήματα στην κανονική λειτουργία της εταιρείας. Παρακάτω αναφέρονται ενδεικτικά τρεις διαφορετικοί τρόποι για πιο εύκολη μετάβαση, μετριάζοντας το ρίσκο της επένδυσης.

Αρχικά, μία πρακτική που μπορεί να προταθεί για την αφομοίωση των οργανισμών ενός νέου μοντέλου επιχειρηματικότητας που να έχει ως επίκεντρο τα δεδομένα, η οποία συνηθίζεται για ριζοσπαστικές αλλαγές στο χώρο των επιχειρήσεων, είναι η πιλοτική εφαρμογή εντός μίας ικανοποιητικής χρονικής περιόδου (συνήθως ένα οικονομικό έτος) σε ένα κομμάτι της εταιρείας στο οποίο μπορούν πιο εύκολα να απομονωθούν οι αρνητικές

επιπτώσεις και ζητημάτων των οποίων δεν είχαν υπολογιστεί πρότερα κατά την διάρκεια της αλλαγής διαδικασιών και ενσωμάτωσης των νέων ψηφιακών λύσεων. Έτσι, εκείνο το κομμάτι στο οποίο δοκιμάζεται, μπορεί να είναι ένα τμήμα ή ένα κατάστημα ή μία ολόκληρη θυγατρική, γίνεται πεδίο έρευνας και αναλύονται όσο τον δυνατόν διεξοδικά τα δεδομένα που θα παραχθούν. Αν τα αποτελέσματα που θα προκύψουν ικανοποιήσουν την διοίκηση και το τμήμα διαχείρισης των δεδομένων, δύναται η περαιτέρω σταδιακή εφαρμογή και σε άλλα τμήματα ή και ακόμη και την καθολική υιοθέτηση. Αν τα αποτελέσματα δεν ανταποκρίνονται στις αρχικές προσδοκίες και τα προβλήματα που προέκυψαν δεν έχουν ξεπεραστεί, αντί να εγκαταλειφθεί η πιλοτική εφαρμογή και κριθεί ως μη εφαρμόσιμη ολόκληρο το πρόγραμμα, μπορεί να αλλάξει το δείγμα, ώστε να επιλεγθούν την επόμενη φορά διαφορετικοί εργαζόμενοι, τμήμα, θυγατρική, τόπος ή να γίνει παράταση του πειράματος και ύστερα να αποφασιστούν οι μετέπειτα ενέργειες. Αυτό μπορεί να προκαλέσει διαφυγή κερδών και καθυστερήσεις, δημιουργώντας αμφιβολίες στους διοικούντες και μετόχους αλλά και πίεση στην ομάδα διαχείρισης δεδομένων για καλύτερη επίβλεψη και γρηγορότερα αποτελέσματα με κίνδυνο βεβιασμένων κινήσεων.



Εικόνα 40. Θεωρητικό μοντέλο Πιλοτικής Εφαρμογής

Μία άλλη τεχνική, είναι η “μίμηση” άλλων εταιρειών του κλάδου που έχουν επιτύχει το συγκεκριμένο σχέδιο, και αν δεν υπάρχουν στο κλάδο τότε η αναζήτηση θα γίνει σε εταιρείες πανομοιότυπες σε όσο το δυνατόν κοντινούς κλάδους. Αυτό έχει ως πλεονέκτημα την μείωση του ρίσκου αποτυχίας και την πιθανότητα απώλειας πόρων και αργοποριών. Από την άλλη, η μεθοδολογία αυτή προϋποθέτει ικανοποιητική πληροφόρηση των ξένων υποθέσεων το οποίο μπορεί να καταστεί δύσκολο και ιδιαίτερα μεταξύ ανταγωνιστών καθώς μπορεί κατηγορηθεί η εταιρεία που προσπαθεί να “μιμηθεί”, για βιομηχανική κατασκοπία και να θιχτεί η φήμη της (brand name) ως “δεύτερη” τάξεως στο κλάδο ενώ να αναδειχθεί η άλλη ως καινοτόμα και πρωτοποριακή. Ένα επιτυχημένο και διάσημο παράδειγμα “μίμησης” στο χώρο των δεδομένων, συνέβη όταν η ομάδα Oakland Athletics την χρονιά 2002 πρωτοπόρησε και εφάρμοσε Περιγραφική Ανάλυση(Descriptive Analysis) [4] κατά την επιλογή των παικτών της για να αντισταθμίσει τα χαμηλά χρηματικά της διαθέσιμα έναντι των ισχυρότερων οικονομικά αντιπάλων της και προσπαθώντας να “δει” στοιχεία στους παίκτες που οι άλλες ομάδες δεν τα έβλεπαν μέσω της εμπειρικής ανάλυσης που εφάρμοζαν όλες μέχρι τότε μέσω εμπειρών αθλητικών ανιχνευτών. Ο κίνδυνος αποτυχίας ήταν μεγάλος και αποτυπώνεται από τις κακές αρχικές εμφανίσεις η οποία προκάλεσε αμφιβολίες ως προς την επιτυχία του εγχειρήματος και με κίνδυνο την διακοπή του στη μέση της αγωνιστικής χρονιάς. Εν τέλη, η χρονιά ήταν από τις πιο επιτυχημένες της ομάδας αλλάζοντας τον τρόπο σκέψης των ομάδων και να την μιμηθούν ορισμένες άμεσα όπως έπραξε η Boston Red Sox, η οποία κατέκτησε το πρωτάθλημα δύο χρόνια μετά ύστερα από 86 χρόνια, εγκαινιάζοντας μια πολυετή επιτυχημένη πορεία με αρκετούς τίτλους.

Άλλη μία προσέγγιση αναφέρεται σε μεγάλο εύρος οργανισμών, η οποία θα ταίριαζε στην δική μας χώρα όντας μικρή πληθυσμιακά και χωρικά, άρα και εύκολα η διαχείριση της, θα ήταν η κατάρτιση ενός εθνικού σχεδίου δράσης ώστε να υποστηρίξει τις υπάρχουσες επιχειρήσεις για ψηφιοποίηση και μετάβαση σε ένα πιο σύγχρονο και εξελιγμένο μοντέλο επιχειρηματικότητας. Τέτοιο παράδειγμα είναι το REACH2025 που υιοθετήθηκε από την Ιορδανία[12] όπου εντός ενός συγκεκριμένου χρονοδιαγράμματος που έχει θέσει εξαρχής η ίδια η χώρα, θα πρέπει να πραγματοποιηθούν ορισμένες δράσεις σε τρεις άξονες, Υποδομές, Νόμοι και Διοίκηση με βάση σε ένα πολυετές πλάνο, το οποίο θα μεταμορφώσει την χώρα και θα μπορέσει να υποστηρίξει και την υπάρχουσα ιδιωτική

επιχειρηματικότητα καθώς και την ανάδειξη νέων καινοτόμων νεοφυούς (start-ups) επιχειρήσεων όπως για παράδειγμα στην Ιορδανία οι παρακάτω:

- MeteoWeather. Συλλέγει μετεωρολογικά δεδομένα μέσα από ένα σύγχρονο δίκτυο πηγών, τα οποία τα επεξεργάζεται και αναλύει μέσα από μοντέλα Τεχνητής Νοημοσύνης προσφέροντας προγνώσεις και υποστήριξη μέσω εφαρμογών στο κλάδο της μετεωρολογίας και το καιρό.
- Nestrom και Whyise, δύο διαφορετικές επιχειρήσεις ανάλυσης δεδομένων που υποστηρίζουν άλλες για λήψη αποφάσεων και διαχείριση των δεδομένων του κάνοντας χρήση τεχνολογιών όπως της Μηχανικής Μάθησης ή το Διαδίκτυο των Πραγμάτων.
- Η Electronic Health Solutions (EHS), μία μη κερδοσκοπική οργάνωση όπου εργάζεται για την ψηφιοποίηση του υγειονομικού συστήματος της Ιορδανίας, το 2019 εισήγαγε ένα πρόγραμμα ανάλυσης υγειονομικών στοιχείων μέσω της πρακτικής εξόρυξης δεδομένων το οποίο μπορεί να υποστηρίξει επαγγελματίες της υγείας και ερευνητές στις αποφάσεις τους.

ΒΙΒΛΙΟΓΡΑΦΙΑ

1. The Economist, Article «A deluge of data is giving rise to a new economy», Feb 2020
2. <https://ourworldindata.org/>
3. Διαχείριση Δεδομένων και Επιχειρηματική Ευφυΐα, Θεωρία και Εφαρμογές για Στελέχη Επιχειρήσεων, 2015 από τους Γεώργιο Σταλίδη και Δημητρίου Καρδάρα
4. Data Analytics Made Accessible: 2022 edition by Anil Maheshwari
5. The Economist, Article «Data is giving rise to a new economy», May 2017
6. The Data Scientist's Toolbox Course by Johns Hopkins University
7. <https://medium.com/> Article "The 5 Vs of Big Data" by Jose Antonio Ribeiro Neto
8. Open Knowledge Foundation <https://okfn.org/>
9. Ευρωπαϊκή Επιτροπή https://ec.europa.eu/info/law/law-topic/data-protection/reform/what-personal-data_el
10. Βικιπαίδεια <https://el.wikipedia.org/>
11. Γραφείο Στατιστικών Εργασίας ΗΠΑ <https://www.bls.gov>
12. World Bank, World Development Report 2021. Data for better lives. <https://wdr2021.worldbank.org/>
13. IBISWorld 2022, Data Processing & Hosting Services in Greece - Industry Statistics 2008 2026 <https://www.ibisworld.com/greece/industry-statistics/data-processing-hosting-services/3625/>
14. Ελληνική Στατιστική Αρχή <https://www.statistics.gr/>
15. TDWI (TRANSFORMING DATA WITH INTELLIGENCE) 2017, What It Takes to Be Data-Driven Technologies and Best Practices for Becoming a Smarter Organization by Fern Halper and David Stodder.
16. Mobility Patterns, Big Data and Transport Analytics: Tools and Applications for Modeling by Konstantinos Antoniou, Loukas Dimitriou, Francisco Pereira
17. SAE International (Society of Automotive Engineers) <https://www.sae.org/>

18. Autonomous Vehicles: Your Ultimate Guide to the Past, Present and Future of Autonomous Vehicles by C D Leonard
19. Waymo LLC Autonomous Car Company <https://waymo.com/>
20. Health Informatics: Practical Guide, Seventh Edition Health Informatics: Practical Guide, Seventh Edition by Robert E. Hoyt and William R. Hersh
21. American Medical Informatics Association <https://amia.org/>
22. Ευρωπαϊκή Ομοσπονδία Ιατρικής Πληροφορικής <https://efmi.org/>
23. Atomwise, Technology-enabled pharmaceutical company <https://www.atomwise.com/>
24. Παγκόσμιο Οργανισμό Υγείας, <https://www.who.int/>
25. Οργανισμό Ηνωμένων Εθνών, <https://www.un.org>
26. Toward Science Data, Article “Big Data & Smart Cities: How can we prepare for them?” by Alexandre Gonfalonieri <https://medium.com/dataseries/big-data-and-smart-cities-why-we-need-them-now-a194b2498fb1>
27. Smart Cities and Artificial Intelligence: Convergent Systems for Planning, Design, and Operations 1st Edition by Christopher Grant Kirwan
28. Κοινωνία της Πληροφορίας και Οικονομία της Γνώσης 2016, από τους Καλογήρου Γιάννη, Παναγιώτη Παναγιωτόπουλο, Τσακανίκας Άγγελος, Σιώκας Ευάγγελος, Καρούνος Θεόδωρος, Μαγκλάρης Βασίλειος, Τρούλος Κώστας, Καλογεράς Δημήτριος, Τσιάβος Πρόδρομος, Κανέλλος Νίκος, Μερεκουλιάς Βασίλης
29. Data Teams: A Unified Management Model for Successful Data-Focused Teams 1st ed. Edition, Kindle Edition by Jesse Anderson
30. Σύνδεσμο Ελληνικών Τουριστικών Επιχειρήσεων <https://sete.gr/el/>
31. Palantir Technologies Εταιρεία Λογισμικού Ανάλυσης Δεδομένων <https://www.palantir.com/>
32. QuestionPro Εταιρεία Λογισμικού και παροχή υπηρεσιών επιπέδου SaaS <https://www.questionpro.com/>

33. IBM Articles “File Storage” <https://www.ibm.com/cloud/learn/file-storage>, Rational Programming Patterns/9.5.0/ The Data Block entity <https://www.ibm.com/docs/en/rpp/9.5.0?topic=entities-database-block-entity>
34. Toward Science Data, Article “Deep Dive into Netflix’s Recommender System” by David Chong <https://towardsdatascience.com/deep-dive-into-netflixs-recommender-system-341806ae3b48>
35. Netflix’s Help Center <https://help.netflix.com/en>
36. Netflix’s Technology Blogspot Article Netflix Recommendations: Beyond the 5 stars (Part 1 and 2) by Xavier Amatriain and Justin Basilico. <https://netflixtechblog.com/netflix-recommendations-beyond-the-5-stars-part-1-55838468f429>, <https://netflixtechblog.com/netflix-recommendations-beyond-the-5-stars-part-2-d9b96aa399f5>
37. Alibaba Cloud Article Cloud Networking: A Global Cloud Network for Distributed Business System https://www.alibabacloud.com/blog/cloud-networking-a-global-cloud-network-for-distributed-business-system_596252
38. Η Επιχειρηματικότητα έκδοση 2008, του Πετράκη Ε. Παναγιώτη
39. Forbes Article “Seven Components Of Strong Business Plans” <https://www.forbes.com/sites/forbescoachescouncil/2020/11/25/seven-components-of-strong-business-plans/>
40. Σύγχρονα Πληροφοριακά Συστήματα Επιχειρήσεων 2015, του Πάνου Φιτσιλή
41. KDnuggets <https://www.kdnuggets.com/2019/05/poll-top-data-science-machine-learning-platforms.html>