



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ
ΥΠΟΛΟΓΙΣΤΩΝ

ΕΡΓΑΣΤΗΡΙΟ ΣΥΣΤΗΜΑΤΩΝ ΑΠΟΦΑΣΕΩΝ ΚΑΙ ΔΙΟΙΚΗΣΗΣ
ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΩΝ ΒΙΟΜΗΧΑΝΙΚΩΝ ΔΙΑΤΑΞΕΩΝ ΚΑΙ
ΣΥΣΤΗΜΑΤΩΝ ΑΠΟΦΑΣΕΩΝ

Big Data και εφαρμογές στην Ενέργεια

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

Ευστάθιου Σούφλα

Επιβλέπων : Ασκούνης Δημήτριος
Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούλιος 2023



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΩΝ ΒΙΟΜΗΧΑΝΙΚΩΝ
ΔΙΑΤΑΞΕΩΝ ΚΑΙ ΣΥΣΤΗΜΑΤΩΝ ΑΠΟΦΑΣΕΩΝ

Big Data και εφαρμογές στην Ενέργεια

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

Ευστάθιου Σούφλα

Επιβλέπων: Ασκούνης Δημήτριος

Καθηγητής Ε.Μ.Π

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 12^η Ιουλίου 2023.

.....
Δημήτριος Ασκούνης

Καθηγητής Ε.Μ.Π

.....
Ιωάννης Ψαρράς

Καθηγητής Ε.Μ.Π

.....
Χρυσόστομος Δούκας

Καθηγητής Ε.Μ.Π

Αθήνα, Ιούλιος 2023

.....

Ευστάθιος Σούφλας

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών
Ε.Μ.Π.

Copyright © ΣΟΥΦΛΑΣ ΕΥΣΤΑΘΙΟΣ, 2023

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.

Περίληψη

Η τεχνολογία των Big Data αποτελεί έναν τομέα που έχει κεντρίσει σε μεγάλο βαθμό το ενδιαφέρον της ακαδημαϊκής και ερευνητικής κοινότητας. Οι ποσότητες των δεδομένων που παράγονται και αποθηκεύονται στον σύγχρονο ψηφιακό κόσμο έχουν αυξηθεί ραγδαία, κάτι που έχει δημιουργήσει πολλές προκλήσεις στην ανάπτυξη εφαρμογών της τεχνολογίας των Big Data. Ωστόσο, οι παραδοσιακές τεχνικές αποθήκευσης και επεξεργασίας των δεδομένων κρίνονται λιγότερο αποτελεσματικές, καθώς παρουσιάζουν αργή απόκριση και έλλειψη επεκτασιμότητας, απόδοσης και ακρίβειας. Η αντιμετώπιση αυτών των προκλήσεων απαιτεί τη δημιουργία νέων τεχνικών και τεχνολογιών αποθήκευσης και επεξεργασίας των δεδομένων, ώστε να είναι εφικτή η υλοποίηση της τεχνολογίας των Big Data σε πλήθος εφαρμογών διαφόρων τομέων και κλάδων.

Σκοπός της διπλωματικής εργασίας, είναι η παρουσίαση μιας όσο το δυνατόν πιο ολοκληρωμένης βιβλιογραφικής ανασκόπησης της τεχνολογίας των Big Data, βάσει των πρόσφατων εξελίξεων σε διάφορες πτυχές του, όπως οι χρησιμοποιούμενες τεχνολογίες και οι προκλήσεις που αντιμετωπίζει, αλλά και η παρουσίαση της εφαρμογής της τεχνολογίας στον τομέα της ενέργειας. Αρχικά παρουσιάζεται το γενικό υπόβαθρο της τεχνολογίας, καθώς και μια παρουσίαση των πλέον αντιπροσωπευτικών εφαρμογών της, όπως οι υγειονομικές και ιατρικές εφαρμογές, τα μέσα κοινωνικής δικτύωσης, ο βιομηχανικός τομέας και άλλες. Στη συνέχεια, αναλύονται οι έννοιες της αξίας και της αλυσίδας αξίας των Big Data με τις τέσσερις φάσεις της. Για κάθε φάση, παρουσιάζεται το γενικό υπόβαθρο, οι τεχνικές προκλήσεις και οι τελευταίες εξελίξεις, με βάση τις μελέτες των τελευταίων χρόνων. Τέλος, παρουσιάζεται η εφαρμογή της τεχνολογίας των Big Data στον τομέα της ενέργειας, την αξιοποίηση των συλλεγόμενων δεδομένων για βελτίωση της λειτουργικής απόδοσης των συστημάτων αλλά και την ενίσχυση των προσπαθειών βέλτιστης διαχείρισης της ενέργειας. Η εργασία ολοκληρώνεται με τα συμπεράσματα που προκύπτουν, εστιάζοντας στα υφιστάμενα ζητήματα και τις μελλοντικές κατευθύνσεις.

Λέξεις κλειδιά

Ανάλυση, διαχείριση της ενέργειας, ενέργεια, εφαρμογές, μέσα κοινωνικής δικτύωσης, μηχανική μάθηση, Big Data, Big Data analytics, Cloud Computing, IoT, smart grid.

Abstract

Big Data is an area receiving considerable attention from academic and research communities. The amounts of data generated and stored in today's digital world have grown rapidly in recent years. This fast evolving rate of produced and stored data has created many challenges in developing Big Data applications in various fields and areas. The development of such applications has become increasingly important that several organizations from different sectors depend increasingly on knowledge extracted from huge volumes of generated and stored data. However, in Big Data context, traditional data processing and storing techniques are less efficient, since they show a slow responsiveness and lack of scalability, performance and accuracy. To face all these challenges, various types of storing and processing techniques and technologies must be developed so as to achieve the implementation of Big Data in various fields' applications.

The purpose of this master thesis is to present the most comprehensive literature review of Big Data based on the latest developments in various aspects, such as used technologies and challenges, and to present the application of Big Data in the energy sector. Firstly, the general background of Big Data is presented focusing on its definition and its characteristics, as well as an examination of the several representative applications of Big Data is done, including healthcare and medical applications, online social networks, industry applications, etc. Next, Big Data value and value chain, including its four phases, are presented, i.e., data generation, data acquisition, data storage, and data analysis. For each phase, the general background is presented, the technical challenges are discussed, and the latest advances are reviewed. Finally, a comprehensive study of big data driven smart energy management is presented so as to fulfill the potential of energy big data and to obtain insights in smart energy management. This study is concluded with a discussion of open problems and future directions.

Keywords

Analytics, Big Data, Big Data analytics, Cloud Computing, energy, energy management, Internet of things, machine learning, social media, smart grid.

Περιεχόμενα

ΚΕΦΑΛΑΙΟ 1: ΕΙΣΑΓΩΓΗ	9
1.1 Η τεχνολογία των Big Data	9
1.2 Εφαρμογή των Big Data στον τομέα της ενέργειας.....	13
1.3 Αντικείμενο διπλωματικής.....	15
1.4 Οργάνωση κειμένου	16
ΚΕΦΑΛΑΙΟ 2: BIG DATA.....	18
2.1 Έναρξη της εποχής των Big Data	18
2.2 Ορισμός των Big Data.....	20
2.3 Χαρακτηριστικά των Big Data.....	22
2.4 Εφαρμογές των Big Data	25
2.4.1 Υγειονομική περίθαλψη	26
2.4.2 Μεταφορές	28
2.4.3 Οικονομικός τομέας	29
2.4.4 Μέσα κοινωνικής δικτύωσης	31
2.4.5 Ηλεκτρονικό εμπόριο	31
2.4.6 Παρακολούθηση περιβάλλοντος και διαχείριση φυσικών καταστροφών	32
2.4.7 Ασφάλεια υποδομών ζωτικής σημασίας	33
2.4.8 Τηλεπικοινωνίες.....	34
2.4.9 Βιομηχανικές εφαρμογές.....	35
ΚΕΦΑΛΑΙΟ 3: ΑΛΥΣΙΔΑ ΑΞΙΑΣ BIG DATA.....	37
3.1 Η αξία των Big Data.....	37
3.2 Αλυσίδα αξίας δεδομένων.....	39
3.3 Αλυσίδα αξίας των Big Data.....	40
3.4 Παραγωγή δεδομένων	42
3.5 Ανάκτηση δεδομένων.....	44
3.5.1 Συλλογή δεδομένων	45
3.5.2 Μετάδοση Δεδομένων.....	46
3.5.3 Προεπεξεργασία δεδομένων	46
3.6 Αποθήκευση δεδομένων	48
3.6.1 Υποδομή αποθήκευσης	48
3.6.2 Πλαίσιο διαχείρισης δεδομένων.....	49
3.7 Ανάλυση Δεδομένων.....	52
3.7.1 Μηχανική μάθηση.....	52
3.7.2 Εξόρυξη δεδομένων	55
ΚΕΦΑΛΑΙΟ 4: BIG DATA ΣΤΟΝ ΤΟΜΕΑ ΤΗΣ ΕΝΕΡΓΕΙΑΣ.....	56

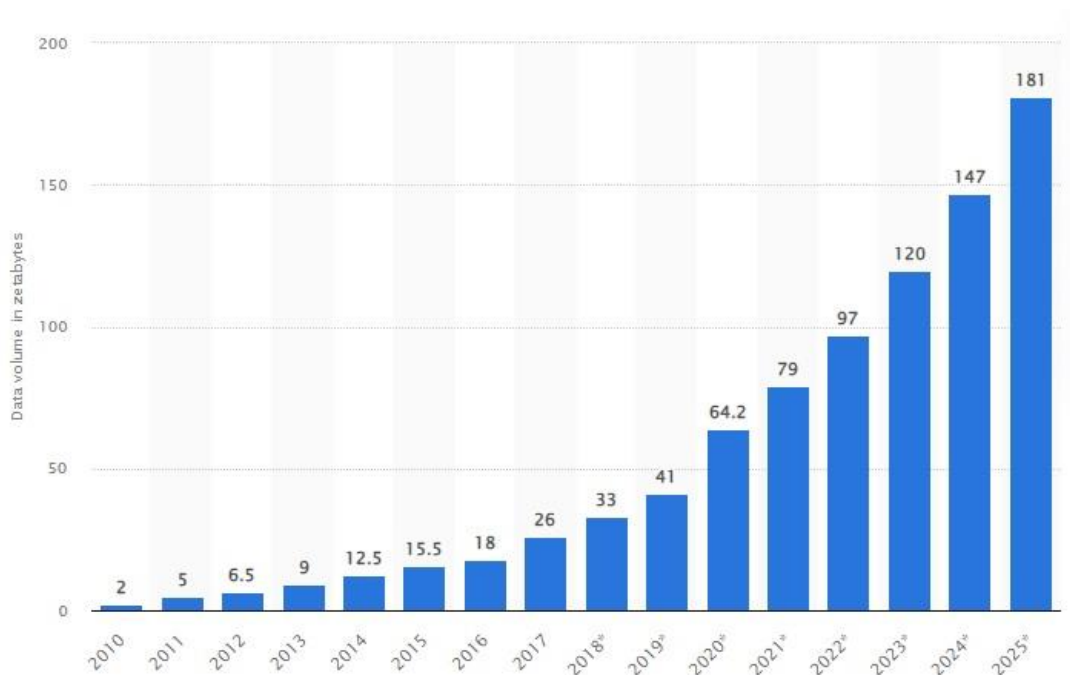
4.1	Πηγές παραγωγής δεδομένων ενέργειας.....	56
4.2	Χαρακτηριστικά των δεδομένων ενέργειας.....	60
4.3	Τεχνολογίες υποστήριξης.....	63
4.3.1	Smart Grid.....	64
4.3.2	IoT.....	68
4.3.3	Cloud computing.....	69
4.3.4	Edge computing.....	71
4.3.5	Fog Computing.....	72
ΚΕΦΑΛΑΙΟ 5: ΕΦΑΡΜΟΓΕΣ BIG DATA ΣΤΟΝ ΤΟΜΕΑ ΤΗΣ ΕΝΕΡΓΕΙΑΣ... 74		
5.1	Ανάγκες διαχείρισης ενέργειας στα smart grid.....	74
5.2	Έξυπνη διαχείριση ενέργειας με γνώμονα τα δεδομένα.....	77
5.3	Διαχείριση ενεργειακής παραγωγής.....	81
5.3.1	Σχεδιασμός ενεργειακής παραγωγής.....	81
5.3.2	Οικονομική κατανομή φορτίου.....	82
5.4	Διαχείριση μεταφοράς ενέργειας.....	83
5.4.1	Έλεγχος μεταβατικής ευστάθειας.....	84
5.5	Διαχείριση διανομής ενέργειας.....	85
5.5.1	Διαχείριση της ενέργειας μικροδικτύων και ΑΠΕ.....	87
5.5.2	Ανίχνευση και αναγνώριση βλαβών.....	90
5.5.3	Παρακολούθηση κατάστασης εξοπλισμού.....	92
5.5.4	Προληπτική συντήρηση.....	93
5.5.5	Ανίχνευση τοπολογίας δικτύου.....	95
5.6	Διαχείριση ενεργειακής ζήτησης.....	96
5.6.1	Παρακολούθηση ποιότητας παρεχόμενης ενέργειας.....	97
5.6.2	Πρόβλεψη φορτίων.....	98
5.6.3	Δημιουργία προφίλ φορτίων – Μοντέλο Demand/Response (DR).....	100
5.6.4	Διαχωρισμός φορτίων.....	103
5.6.5	Ανίχνευση ενεργειακής απάτης.....	104
ΚΕΦΑΛΑΙΟ 6: ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΜΕΛΛΟΝΤΙΚΕΣ ΚΑΤΕΥΘΥΝΣΕΙΣ .. 106		
ΒΙΒΛΙΟΓΡΑΦΙΚΕΣ ΑΝΑΦΟΡΕΣ.....		110

ΚΕΦΑΛΑΙΟ 1: ΕΙΣΑΓΩΓΗ

1.1 Η τεχνολογία των Big Data

Η σύγχρονη εποχή χαρακτηρίζεται από την καθημερινή παραγωγή και δημιουργία δεδομένων. Οι εξελίξεις στην τεχνολογία του Διαδικτύου των Πραγμάτων (Internet of Things – IoT), η εξάπλωση του υπολογιστικού νέφους (cloud computing) και η συνεχής αύξηση της χρήσης έξυπνων συσκευών, έχουν δημιουργήσει ένα ψηφιακό οικοσύστημα στο οποίο αυτή η δημιουργία δεδομένων γίνεται σε πρωτοφανές ποσοστό και προέρχεται από ετερογενείς πηγές και τομείς, όπως η υγεία, τα κοινωνικά δίκτυα, το μάρκετινγκ, η οικονομία, κλπ. [1]. Ένα τέτοιο ψηφιακό οικοσύστημα συμπληρώνεται από ισχυρά υπολογιστικά συστήματα και κατακεντρωμένες εφαρμογές που υποστηρίζουν διαδικτυακά συστήματα πολλαπλών συνδέσεων, όπως συστήματα έξυπνων δικτύων παροχής ηλεκτρικής ενέργειας, συστήματα υγειονομικής περίθαλψης, συστήματα λιανικής πώλησης, κυβερνητικά συστήματα, κλπ. [2].

Σύμφωνα με έρευνα της εταιρείας Statista του Ιουλίου του 2021, το 2020 ο συνολικός όγκος των δεδομένων που δημιουργήθηκαν, συλλέχθηκαν, αντιγράφηκαν και καταναλώθηκαν παγκοσμίως έφτασαν τα 64,2 zettabyte (10^{21} byte), δημιουργώντας νέο ρεκόρ. Η αύξηση αυτή ήταν υψηλότερη από ότι αναμενόταν, καθώς λόγω της πανδημίας COVID-19 οι περισσότεροι άνθρωποι εργάζονταν από το σπίτι και χρησιμοποιούσαν πιο συχνά επιλογές οικιακής ψυχαγωγίας [3]. Οι χρήστες του Διαδικτύου δημιουργούν περίπου 1,7 MB δεδομένων το δευτερόλεπτο. Μόνο στο Facebook παράγονται 4 petabyte (10^{15} byte) δεδομένων την ημέρα. Μέσα στο 2020, οι χρήστες παρακολούθησαν 10,5 εκατομμύρια βίντεο στο Facebook, πραγματοποίησαν σχεδόν 2 τρισεκατομμύρια αιτήματα αναζήτησης στην Google, ενώ οι νέοι χρήστες του Διαδικτύου έφτασαν τα 319 εκατομμύρια [4]. Όλα αυτά τα στοιχεία αποδεικνύουν περίτρανα ότι ο όγκος των δεδομένων αυξάνεται συνεχώς και αναμένεται, έως το 2025, η παγκόσμια δημιουργία δεδομένων να αυξηθεί περισσότερο και να ξεπεράσει τα 180 zettabyte (Εικ. 1.1) [3]. Κάτι τέτοιο δημιουργεί τεράστιες προκλήσεις όσον αφορά τη συλλογή, την αποθήκευση, τη διανομή και την ανάλυση των παραγόμενων δεδομένων [5].



Εικόνα 1.1: Ποσότητα δεδομένων που δημιουργήθηκαν, καταναλώθηκαν και αποθηκεύτηκαν (2010-2025) [3]

Τα παραδοσιακά τεχνολογικά συστήματα δεν είναι σε θέση να διαχειριστούν όλη αυτή την τεράστια ποσότητα δεδομένων που αποθηκεύονται στον ψηφιακό κόσμο. Ως εκ τούτου, απαιτήθηκαν νέες τεχνολογίες και μέθοδοι για τη συλλογή, την αποθήκευση και την ανάλυσή τους [6]. Ο όρος Big Data δημιουργήθηκε ακριβώς για να ορίσει τους τεράστιους όγκους δομημένων, ημιδομημένων και μη δομημένων συνόλων δεδομένων, τα οποία δεν μπορούσαν να επεξεργαστούν από τις παραδοσιακές τεχνικές διαχείρισης δεδομένων και αναφέρεται σε αναδυόμενες τεχνολογίες που έχουν σχεδιαστεί για την εξαγωγή πολύτιμων πληροφοριών από τα σύνολα αυτά. Παρά τα όσα έχουν αναφερθεί στη βιβλιογραφία σχετικά με το τι χαρακτηρίζει την τεχνολογία των Big Data, στις περισσότερες μελέτες τα Big Data χαρακτηρίζονται από πέντε V, δηλαδή, τον όγκο (Volume), την ποικιλία (Variety), την ταχύτητα (Velocity), την αξία (Value) και την ποιότητα (Veracity) [7].

Οι παραδοσιακές τεχνολογίες έχουν περιορισμένη χωρητικότητα αποθήκευσης και δεν παρουσιάζουν την επεκτασιμότητα, την ευελιξία και την απόδοση που απαιτούνται στο πλαίσιο των Big Data. Η διαχείριση των Big Data απαιτεί σημαντικούς πόρους, νέες μεθόδους και ισχυρές τεχνολογίες, καθώς τα Big Data απαιτούν την κατάλληλη επεξεργασία, ανάλυση και ασφάλεια, ώστε να μετατραπούν σε σύνολα δεδομένων τα οποία μπορούν να χρησιμοποιηθούν από τους οργανισμούς των διάφορων τομέων. Η

σημασία της σωστής ανάλυσης των δεδομένων για τους οργανισμούς των διάφορων τομέων έχει αναγνωριστεί και από τους ίδιους τους οργανισμούς, καθώς αυξάνει την ανταγωνιστικότητα και ενισχύει την ανεύρεση νέων ιδεών και την εξατομίκευση των παρεχόμενων υπηρεσιών. Αυτή η αναγνώριση εναρμονίζεται με τους Manyika και συν. (2011), οι οποίοι περιέγραψαν τα Big Data ως το επόμενο τεράστιο βήμα προς τον ανταγωνισμό, την παραγωγικότητα και την καινοτομία [8].

Λόγω της αξίας που παρουσιάζει η ανάλυση των Big Data, δημιουργήθηκαν διάφορα project, με σκοπό την δημιουργία των κατάλληλων συνθηκών επεξεργασίας τους. Τον Μάρτιο του 2012, η κυβέρνηση Ομπάμα ξεκίνησε το project “*Έρευνας και Ανάπτυξης των Big Data*”, με προϋπολογισμό 200 εκατομμυρίων δολαρίων. Στην Ιαπωνία, η ανάπτυξη των Big Data αποτέλεσε σημαντικό άξονα της εθνικής τεχνολογικής στρατηγικής. Τον Απρίλιο του 2012, η υπηρεσία της τεχνολογίας των πληροφοριών της Ν. Κορέας ξεκίνησε το project “*Υπηρεσίες Big Data*”. Τον Μάιο του 2012, τα Ηνωμένα Έθνη δημοσίευσαν μια έκθεση με τίτλο “*Big Data for Development: Opportunities and Challenges*”, στην οποία αναλύονται οι κύριες ανησυχίες σχετικά με τις προκλήσεις που παρουσιάζουν τα Big Data και προτείνεται η προώθηση του διαλόγου σχετικά με το πώς τα Big Data μπορούν να ενισχύσουν τη διεθνή ανάπτυξη. Τέλος, τον Δεκέμβριο του 2013, οι αρμόδιες κυβερνητικές υπηρεσίες της Ν. Κορέας ανακοίνωσαν από κοινού τη Στρατηγική Ανάπτυξης της Βιομηχανίας των Big Data και τον Ιούνιο του 2014 ξεκίνησαν τη χρήση των Big Data σε καινοτόμες εργασίες των τμημάτων του υπουργείου Εμπορίου, Βιομηχανίας και Ενέργειας [9].

Αποτέλεσμα της ανάπτυξης όλων αυτών των διαφορετικών project σε όλο τον κόσμο, ήταν η δημιουργία μοντέλων, πλαισίων και νέων προηγμένων τεχνικών εξόρυξης και αποθήκευσης των Big Data, με σκοπό την παροχή μεγαλύτερης χωρητικότητας αποθήκευσης, διαχείρισης, παράλληλης επεξεργασίας και ανάλυσης τους σε πραγματικό χρόνο από διαφορετικές ετερογενείς πηγές [8]. Επιπλέον, αναπτύχθηκαν νέες λύσεις για τη διασφάλιση του απορρήτου και της ασφάλειας των δεδομένων. Σε σύγκριση με τις παραδοσιακές τεχνολογίες, οι λύσεις αυτές παρέχουν μεγαλύτερη ευελιξία, επεκτασιμότητα, και απόδοση, προσπαθώντας παράλληλα να εξασφαλίσουν πιο ακριβή και αξιόπιστα αποτελέσματα για τις εφαρμογές των Big Data. Ωστόσο, η επιλογή της καταλληλότερης λύσης, ανάλογα με την εφαρμογή, είναι ιδιαίτερα δύσκολη και χρονοβόρα, καθώς θα πρέπει να ληφθούν υπόψη πολλές παράμετροι, όπως η τεχνολογική συμβατότητα, η πολυπλοκότητα ανάπτυξης, το κόστος, η

αποτελεσματικότητα, η απόδοση, η αξιοπιστία, η υποστήριξη και οι κίνδυνοι ασφάλειας [2]. Οι κύριες προκλήσεις για τους ερευνητές και τους επαγγελματίες προκύπτουν από τον εκθετικό ρυθμό αύξησης των δεδομένων, ο οποίος ξεπερνά την ικανότητα των ανθρώπων να σχεδιάζουν κατάλληλα συστήματα αποθήκευσης και ανάλυσης δεδομένων, για να διαχειρίζονται αποτελεσματικά τους μεγάλους όγκους πληροφοριών που περιλαμβάνονται στα Big Data [10].

Παρά τις όποιες προκλήσεις, η αξία των Big Data είναι αδιαμφισβήτητη. Η σημασία των Big Data δεν αφορά τον όγκο των δεδομένων που έχει ένας οργανισμός σε οποιοδήποτε τομέα, αλλά μάλλον τον τρόπο με τον οποίο αυτά χρησιμοποιούνται. Καθώς, στη σύγχρονη εποχή, δεδομένα μπορούν να βρεθούν παντού, η ανάλυσή τους παράγει πληροφορίες που αφορούν διάφορα θέματα των οργανισμών, όπως μείωση κόστους και χρόνου, ανάπτυξη νέων προϊόντων, βελτιστοποίηση παρεχόμενων υπηρεσιών και έξυπνη λήψη αποφάσεων. Ο συνδυασμός τους με σύγχρονες μεθόδους ανάλυσης, μπορεί να οδηγήσει στην επίτευξη των στόχων πολλών πολύτιμων επιχειρηματικών δραστηριοτήτων, όπως τον σχεδόν άμεσο εντοπισμό των αιτιών που οδηγούν σε σφάλματα, τη δημιουργία εκπαιδευτικών κουπονιών βάσει των αγοραστικών συνηθειών των πελατών, την ταχύτατη αναδιάρθρωση του κινδύνου χαρτοφυλακίου και την ανίχνευση δόλιων συμπεριφορών προτού επηρεάσουν την εταιρεία. Αυτά είναι μόνο μερικά παραδείγματα των δυνατοτήτων των Big Data analytics [11].

Για τους λόγους αυτούς, πολλοί οργανισμοί, ανεξαρτήτως τομέα, υλοποιούν εφαρμογές Big Data, αποσκοπώντας στη βελτίωση της αποδοτικότητας και της απόδοσης, όπως επίσης και στη μείωση του κόστους και της κατανάλωσης πόρων. Η αποτελεσματική ανάλυση και χρήση των Big Data αποτελεί βασικό παράγοντα επιτυχίας πολλών τομέων και κλάδων [12]. Οι εφαρμογές των Big Data σε διάφορους τομείς, είναι ένα θέμα που έχει απασχολήσει κατά κόρον την ακαδημαϊκή και ερευνητική κοινότητα. Αυτό αποδεικνύεται από το μεγάλο αριθμό σχετικών μελετών που έχουν εμφανιστεί μέχρι τώρα στη βιβλιογραφία. Οι μελέτες αυτές έχουν επικεντρωθεί στην εφαρμογή των Big Data σε συγκεκριμένους τομείς, όπως η διαχείριση των επιχειρήσεων, το IoT, τα διαδικτυακά κοινωνικά δίκτυα, οι υγειονομικές και ιατρικές εφαρμογές, οι έξυπνες πόλεις, κλπ., ή παρουσιάζουν μια γενική σύνοψη των εφαρμογών αυτών. Για παράδειγμα, οι Salomi & Balamurugan (2016) ανέλυσαν την ανάγκη εφαρμογής των Big Data στον τομέα της υγειονομικής περίθαλψης [13], οι Brinch, Stentoft & Jensen (2017) παρουσίασαν μια σύνοψη της

τεχνολογίας των Big Data και της εφαρμογής τους στη διαχείριση των εφοδιαστικών αλυσίδων [14], οι Mosavi, Lopez & Varkonyi-Koczy (2017) ανέλυσαν τις βιομηχανικές εφαρμογές των Big Data [15], οι Lin & Li (2018) ανέλυσαν την εφαρμογή των Big Data στη χαρτογράφηση και στην τεχνολογία των γεωγραφικών πληροφοριών [16], οι Sun & Scanlon (2019) ανέλυσαν τους τρόπους και τις εφαρμογές με τα οποία τα Big Data μπορούν να συνεισφέρουν στη διαχείριση της εξοικονόμησης νερού στον πλανήτη [17], οι Munawar και συν. (2020) παρουσίασαν μια συστηματική ανάλυση των εφαρμογών των Big Data στην έξυπνη ακίνητη περιουσία [18], οι Villarejo-Ramos και συν. (2021) παρουσίασαν ένα μοντέλο αξιολόγησης και πρόβλεψης της υιοθέτησης των Big Data από τις επιχειρήσεις [19] και οι Blank και συν. (2021) παρουσίασαν μια σύνοψη των τεχνικών Big Data που χρησιμοποιούνται για την εξόρυξη πληροφοριών από τα μέσα κοινωνικής δικτύωσης [20], ενώ οι S. K. Rubal (2016) [21], M. Goru (2017) [22], W. M. Marries (2017) [23], P. V. Desai (2018) [24] και Janev και συν. (2020) [25] παρουσίασαν μια σύνοψη των εφαρμογών των Big Data.

1.2 Εφαρμογή των Big Data στον τομέα της ενέργειας

Με την ταχεία ανάπτυξη των τεχνολογιών των αισθητήρων, της ασύρματης μετάδοσης, των δικτυακών επικοινωνιών, του τεχνολογία επικοινωνίας δικτύου, του Cloud Computing και των έξυπνων φορητών συσκευών, η καθημερινή συλλογή δεδομένων κάθε δομής και μορφής είναι τεράστια σε όλους τους τομείς. Στον τομέα της ενέργειας, καθημερινά συλλέγονται τεράστιες ποσότητες δεδομένων που αφορούν της παραγωγή και την κατανάλωση ενέργειας. Μια τέτοια κατάσταση οδήγησε στην ανάγκη ψηφιοποίησης των ενεργειακών συστημάτων, η οποία υποστηρίχθηκε από τη χρήση αναδυόμενων τεχνολογιών πληροφοριών [26].

Τα χαρακτηριστικά της τεχνολογίας των Big Data μπορούν να αλλάξουν το υφιστάμενο τοπίο του ενεργειακού τομέα, αντιμετωπίζοντας διάφορες προκλήσεις, όπως η λειτουργική αποτελεσματικότητα και ο έλεγχος του κόστους, η σταθερότητα και η αξιοπιστία των ενεργειακών συστημάτων, η ορθή διαχείριση των ανανεώσιμων πηγών ενέργειας (ΑΠΕ), η ενεργειακή απόδοση και τα περιβαλλοντικά ζητήματα, καθώς και η βελτίωση των παρεχόμενων υπηρεσιών προς τους καταναλωτές [27]. Η αντιμετώπιση αυτών των προκλήσεων μπορεί να επιτευχθεί με την όσο το δυνατόν καλύτερη ανάλυση των δεδομένων που συλλέγονται καθημερινά από όλα τα

ενεργειακά συστήματα, η οποία δημιουργεί νέες ευκαιρίες για την επίτευξη μιας έξυπνης διαχείρισης της παραγόμενης και καταναλωμένης ενέργειας. Πιο συγκεκριμένα, στη βιβλιογραφία έχουν παρουσιαστεί αρκετές μελέτες οι οποίες αναλύουν τη χρήση της τεχνολογίας των Big Data στο πλαίσιο των έξυπνων δικτύων παροχής ηλεκτρικής ενέργειας (smart grid) με σκοπό την επίτευξη των στόχων της καθαρής παραγωγής ενέργειας, της αποτελεσματικής μετάδοσης της ισχύος, τη δυναμική κατανομή της ισχύος και την ορθολογική κατανάλωση της ηλεκτρικής ενέργειας [28] – [34].

Τα έξυπνα δίκτυα παροχής ηλεκτρικής ενέργειας έχουν σχεδιαστεί με σκοπό την ενσωμάτωση των εννοιών της ροής πληροφοριών και της ενεργειακής ροής, επιτυγχάνοντας την ταυτόχρονη συλλογή δεδομένων και μεταφορά ενέργειας [28]. Η φιλοσοφία μιας τέτοιας λειτουργίας αποσκοπεί στη συλλογή μεγάλων ποσοτήτων δεδομένων διαφόρων τύπων, όπως δεδομένα κατάστασης συσκευής, δεδομένα κατανάλωσης ηλεκτρικής ενέργειας και δεδομένα αλληλεπίδρασης με τον χρήστη [29], και στην εν συνεχεία ανάλυσή τους μέσω τεχνικών ανάλυσης, όπως οι τεχνικές βελτιστοποίησης, πρόβλεψης, ταξινόμησης και ομαδοποίησης [30] – [32]. Μια τέτοια ανάλυση των Big Data στον τομέα της ενέργειας μπορεί να οδηγήσει στη βελτιστοποίηση της παραγωγής ενέργειας και της λειτουργίας των ενεργειακών συστημάτων σε πραγματικό χρόνο, στην πρόβλεψη με ακρίβεια της ζήτησης για ηλεκτρική ενέργεια, στην ανεύρεση με ακρίβεια των προτύπων κατανάλωσης ηλεκτρικής ενέργειας και στην αποτελεσματική ανάπτυξη δυναμικών μηχανισμών τιμολόγησης [33]. Με βάση την ανάλυση των Big Data, τα δίκτυα smart grid μπορούν επίσης να ανιχνεύσουν και να αποκαταστήσουν γρήγορα τις βλάβες που τους εμφανίζονται, να ανταποκριθούν άμεσα στις μεταβολές της ζήτησης για ηλεκτρική ενέργεια, να παρέχουν πιο αξιόπιστη και οικονομική ενέργεια, καθώς και να υποστηρίξουν τους καταναλωτές στο να έχουν μεγαλύτερο έλεγχο της ενέργειας που καταναλώνουν. Τέλος, η ανάλυση των Big Data μπορεί να παρέχει τις απαιτούμενες λύσεις για τη λήψη των εκάστοτε βέλτιστων αποφάσεων από όλους τους εμπλεκόμενους φορείς στο πλαίσιο των έξυπνων δικτύων παροχής ηλεκτρικής ενέργειας, όπως τους φορείς εκμετάλλευσης, τους καταναλωτές και τις ρυθμιστικές αρχές των δικτύων [34].

Η τεχνολογία των Big Data είναι σε θέση να μετασχηματίσει τους τρόπους παραγωγής και κατανάλωσης της ενέργειας. Παρόλα αυτά, η χρήση της τεχνολογίας στον τομέα

της ενέργειας καλείται να αντιμετωπίσει και κάποιες προκλήσεις, όπως [27]: (α) πώς να γίνει αποτελεσματικά η συλλογή, αποθήκευση και διαχείριση των δεδομένων, (β) πώς να γίνει αποτελεσματικά η ανάλυσή τους και η εξόρυξη των πολύτιμων πληροφοριών που εμπεριέχουν, (γ) πώς να γίνει αποτελεσματικά η χρήση τους για τη λήψη των εκάστοτε βέλτιστων αποφάσεων, (δ) πώς να γίνει η σωστή δημιουργία και αποτύπωση των αξιών που εμπεριέχουν και (ε) πώς να γίνει αποτελεσματικά η αποτροπή των κινδύνων παραβίασης του απορρήτου κατά τη χρήση τους.

1.3 Αντικείμενο διπλωματικής

Σκοπός της παρούσας διπλωματικής εργασίας, είναι η παρουσίαση μιας όσο το δυνατόν πιο ολοκληρωμένης βιβλιογραφικής ανασκόπησης της τεχνολογίας των Big Data με βάση τις πλέον πρόσφατες εξελίξεις σε διάφορες πτυχές της, όπως είναι οι χρησιμοποιούμενες τεχνολογίες και οι προκλήσεις που αντιμετωπίζει, αλλά και η παρουσίαση της εφαρμογής της τεχνολογίας στον τομέα της ενέργειας. Αρχικά παρουσιάζεται το γενικό υπόβαθρο της τεχνολογίας, στο οποίο περιλαμβάνονται ο ορισμός, τα χαρακτηριστικά, η αξία και η αλυσίδα αξίας των Big Data.. Στη συνέχεια, παρουσιάζονται οι τέσσερις φάσεις της αλυσίδας αξίας των Big Data, δηλαδή, η δημιουργία, η ανάκτηση, η αποθήκευση και η ανάλυση δεδομένων. Για κάθε φάση, παρουσιάζεται το γενικό υπόβαθρο, οι τεχνικές προκλήσεις και οι τελευταίες εξελίξεις, με βάση τις μελέτες που έχουν εμφανιστεί τα τελευταία χρόνια στη βιβλιογραφία. Κατόπιν, παρουσιάζονται οι πλέον αντιπροσωπευτικές εφαρμογές της τεχνολογίας των Big Data, όπως οι υγειονομικές και ιατρικές εφαρμογές, τα μέσα κοινωνικής δικτύωσης, ο βιομηχανικός τομέας και άλλες. Τέλος, παρουσιάζεται η εφαρμογή της τεχνολογίας των Big Data στον τομέα της ενέργειας, με έμφαση την παρουσίαση των τεχνολογιών που είναι σε θέση να υποστηρίξουν αυτήν την εφαρμογή, όπως το cloud computing, το IoT και το έξυπνο δίκτυο παροχής ηλεκτρικής ενέργειας (smart grid), την αξιοποίηση των συλλεγόμενων δεδομένων για βελτίωση της λειτουργικής απόδοσης των συστημάτων αλλά και την ενίσχυση των προσπαθειών βέλτιστης διαχείρισης της ενέργειας. Η εργασία ολοκληρώνεται με μια αναφορά στα συμπεράσματα που προκύπτουν από όλα τα παραπάνω, εστιάζοντας στα υφιστάμενα ζητήματα και τις μελλοντικές κατευθύνσεις.

Στόχος είναι η παροχή μιας ολοκληρωμένης επισκόπησης ενός τομέα που έχει κεντρίσει σε μεγάλο βαθμό το ενδιαφέρον της ακαδημαϊκής και ερευνητικής

κοινότητας, με την ελπίδα ότι θα μπορέσει να αποτελέσει μια χρήσιμη αναφορά για φοιτητές και επαγγελματίες σε συναφείς τομείς και γενικά σε αναγνώστες που θα επωφεληθούν από την κατανόηση του πεδίου των Big Data. Για να επιτευχθεί αυτός ο στόχος, θα πραγματοποιηθεί μια όσο το δυνατόν πιο διεξοδική και περιεκτική ανασκόπηση της διεθνούς βιβλιογραφίας. Για την αναζήτηση της βιβλιογραφίας θα χρησιμοποιηθεί η βάση δεδομένων Google Scholar, όσο και ανασκοπήσεις από διάφορες μελέτες, άρθρα και πηγές μέσω του Διαδικτύου.

1.4 Οργάνωση κειμένου

Στα πλαίσια της παρούσας πτυχιακής διπλωματικής εργασίας, για την παρουσίαση μιας όσο το δυνατόν πιο ολοκληρωμένης βιβλιογραφικής ανασκόπησης της τεχνολογίας των Big Data και των εφαρμογών τους, επιλέχθηκε η ακόλουθη δομή.

Στο κεφάλαιο 2 παρουσιάζεται ένα γενικό υπόβαθρο της τεχνολογίας, στο οποίο περιλαμβάνονται ο ορισμός και τα χαρακτηριστικά των Big Data, καθώς και μια παρουσίαση των πλέον αντιπροσωπευτικών εφαρμογών της τεχνολογίας των Big Data, όπως οι υγειονομικές και ιατρικές εφαρμογές, οι μεταφορές, οι εφαρμογές στον οικονομικό τομέα, τα διαδικτυακά κοινωνικά δίκτυα, το ηλεκτρονικό εμπόριο, η παρακολούθηση του περιβάλλοντος και η διαχείριση των φυσικών καταστροφών, η ασφάλεια υποδομών ζωτικής σημασίας, οι τηλεπικοινωνίες και οι βιομηχανικές εφαρμογές. Στόχος αυτής της παρουσίασης είναι η ανάδειξη της σημασίας της εφαρμογής των Big Data στους τομείς αυτούς.

Στο κεφάλαιο 3 αναλύεται η έννοια της αξίας και της αλυσίδας αξίας των δεδομένων και παρουσιάζονται οι τέσσερις φάσεις της αλυσίδας αξίας των Big Data, δηλαδή, η δημιουργία και η ανάκτηση, η αποθήκευση και η ανάλυση των δεδομένων. Για κάθε φάση, παρουσιάζεται το γενικό υπόβαθρο, οι τεχνικές προκλήσεις και οι τελευταίες εξελίξεις με βάση τις μελέτες που έχουν εμφανιστεί τα τελευταία χρόνια στη βιβλιογραφία.

Στο κεφάλαιο 4 αναφέρονται οι πηγές και τα χαρακτηριστικά των Big Data του τομέα της ενέργειας και γίνεται μια ανασκόπηση των τεχνολογιών που είναι σε θέση να υποστηρίξουν μια τέτοια εφαρμογή των Big Data, όπως το cloud computing, το edge computing, το Internet of Things (IoT) και το έξυπνο δίκτυο παροχής ηλεκτρικής ενέργειας (smart grid).

Στο κεφάλαιο 5 παρουσιάζεται ένα γενικό μοντέλο έξυπνης διαχείρισης της ενέργειας με γνώμονα τη χρήση των συλλεγόμενων δεδομένων από τους αισθητήρες και τις τεχνολογίες ασύρματης μετάδοσης, δικτυακής επικοινωνίας και υπολογιστικού νέφους που χρησιμοποιούνται στα ενεργειακά συστήματα. Επίσης, λαμβάνοντας ως ερευνητικό υπόβαθρο τα έξυπνα δίκτυα ηλεκτρικής ενέργειας, παρουσιάζονται μελέτες εφαρμογών της τεχνολογίας των Big Data σε περιπτώσεις, όπως η ενεργειακή παραγωγή και η διαχείρισή της, η διαχείριση των μικροδικτύων και των ανανεώσιμων πηγών ενέργειας και η διαχείριση της ενεργειακής ζήτησης.

Τέλος, η εργασία ολοκληρώνεται με τα συμπεράσματα που προκύπτουν από όλα τα παραπάνω, εστιάζοντας στα υφιστάμενα ζητήματα και τις μελλοντικές κατευθύνσεις.

ΚΕΦΑΛΑΙΟ 2: BIG DATA

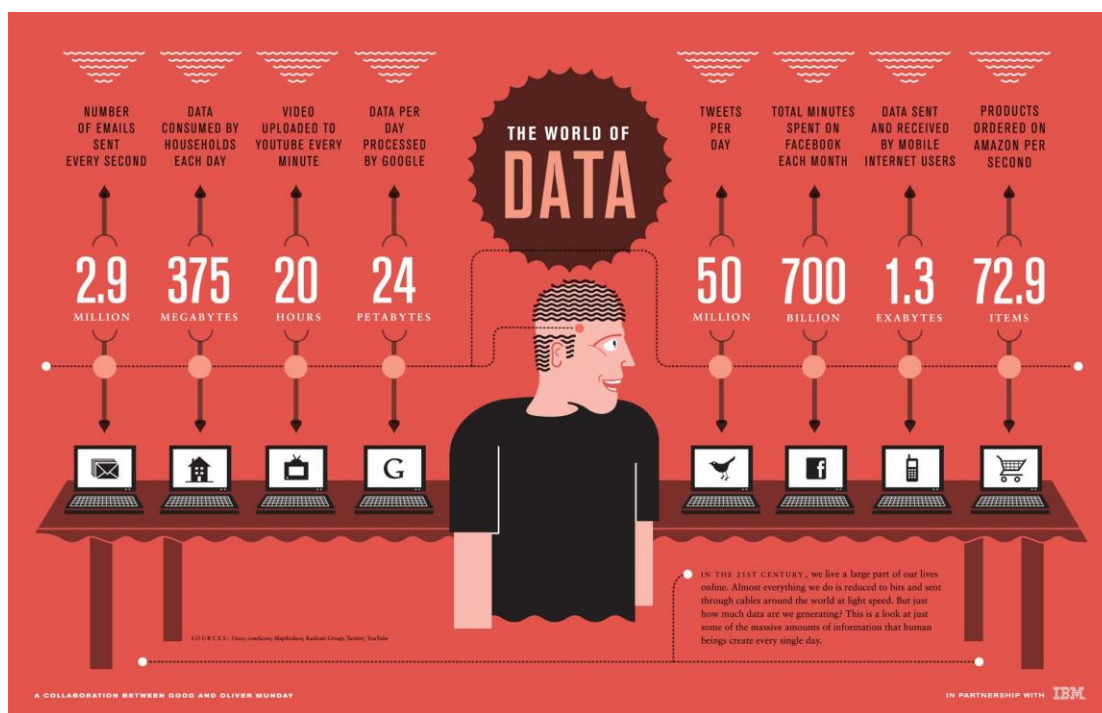
2.1 Έναρξη της εποχής των Big Data

Τα τελευταία 20 χρόνια, τα δεδομένα έχουν αυξηθεί σε μεγάλη κλίμακα και σε διάφορους τομείς. Σύμφωνα με έκθεση της εταιρείας ερευνών International Data Corporation (IDC) του 2021, μιας από τις μεγαλύτερες εταιρείες που ασχολείται με την τεχνολογία των Big Data και τα ερευνητικά τους πεδία, ο συνολικός όγκος δεδομένων που δημιουργήθηκαν και αντιγράφηκαν σε ολόκληρο τον κόσμο το 2020, ήταν 64,2 zettabyte, λόγω της δραματικής αύξησης των ατόμων που εργάστηκαν, εκπαιδεύτηκαν ή ψυχαγωγήθηκαν από το σπίτι τους χρησιμοποιώντας το Διαδίκτυο. Ο μεγάλος αυτός όγκος δεδομένων δημιουργήθηκε από την πίεση που προκάλεσε ο ιός COVID-19 σε όλους τους τομείς και τα αποτελέσματα αυτής της κατάστασης θα έχουν αντίκτυπο για πολλά χρόνια. Σύμφωνα με την ίδια έκθεση, αυτός ο αριθμός αναμένεται να αυξηθεί ακόμα περισσότερο τα επόμενα πέντε χρόνια και θα είναι πολύ μεγαλύτερος από το διπλάσιο του όγκου των δεδομένων που δημιουργήθηκε από την έναρξη της εποχής της ψηφιακής αποθήκευσης [35].

Ο όρος Big Data επινοήθηκε μετά την εκρηκτική αύξηση που παρουσίασε η παραγωγή ψηφιακών δεδομένων σε παγκόσμιο επίπεδο και χρησιμοποιήθηκε κυρίως για την περιγραφή της δημιουργίας των τεράστιων συνόλων δεδομένων [36]. Σε σύγκριση με τα παραδοσιακά σύνολα δεδομένων, τα Big Data γενικά περιλαμβάνουν μεγάλο πλήθος μη δομημένων δεδομένων που χρειάζονται περισσότερη ανάλυση σε πραγματικό χρόνο. Η ανάλυσή τους δημιουργεί ευκαιρίες εξόρυξης πολύτιμων πληροφοριών που μπορεί να ενισχύσουν την εις βάθος κατανόηση κρυμμένων αξιών και την αντιμετώπιση προκλήσεων, για παράδειγμα για τον αποτελεσματικό τρόπο οργάνωσης και διαχείρισης τέτοιων δεδομένων. Αυτοί άλλωστε είναι και οι λόγοι για τους οποίους η τεχνολογία των Big Data έχει κεντρίσει το ενδιαφέρον της ακαδημαϊκής και ερευνητικής κοινότητας, των βιομηχανικών και επιχειρηματικών τομέων, καθώς και τους κρατικούς φορείς [37]. Για παράδειγμα, ζητήματα της τεχνολογίας έχουν συχνά καλυφθεί από παγκοσμίου φήμης περιοδικά, όπως τα The Economist και New York Times, όπως επίσης και από κορυφαία επιστημονικά περιοδικά, όπως τα Nature και Science, στα οποία υπάρχουν ειδικές στήλες αναφοράς της σημασίας και των προκλήσεων που αντιμετωπίζουν τα Big Data. Επίσης, πολλοί βιομηχανικοί και

επιχειρηματικοί τομείς και κλάδοι έχουν εκφράσει το ενδιαφέρον τους για τις μεγάλες δυνατότητες της τεχνολογίας, καθώς και κυβερνητικές υπηρεσίες σε πολλές χώρες παγκοσμίως ανακοίνωσαν σχέδια για την επιτάχυνση των ερευνών σχετικά με τις εφαρμογές των Big Data [38].

Η ταχεία ανάπτυξη των Big Data προέρχεται κυρίως από την ψηφιοποίηση που χαρακτηρίζει πλέον την καθημερινότητα των ανθρώπων, ειδικά σε σχέση με την συνδεσιμότητά τους στο Διαδίκτυο. Για παράδειγμα, η Google επεξεργάζεται δεδομένα εκατοντάδων petabyte και στο Facebook παράγονται περίπου 4 petabyte δεδομένων την ημέρα. Η κινεζική εταιρεία Baidu επεξεργάζεται δεδομένα δεκάδων petabyte και η Taobao, θυγατρική της Alibaba, παράγει δεδομένα δεκάδων terabyte σε διαδικτυακές συναλλαγές την ημέρα [2].



Εικόνα 2.1: Η εκρηκτική αύξηση του όγκου των ψηφιακών δεδομένων θα απαιτήσει περισσότερες δυνατότητες από τα υφιστάμενα πληροφοριακά συστήματα [37]

Άλλος ένας παράγοντας που οδήγησε στην απότομη αύξηση της δημιουργίας δεδομένων ήταν οι εξελίξεις στις τεχνολογίες του υπολογιστικού νέφους (Cloud Computing – CC) και του Διαδικτύου των Πραγμάτων (Internet of Things - IoT) [39]. Το CC παρέχει προστασία, τοποθεσίες πρόσβασης και κανάλια μεταφοράς των δεδομένων. Στο πλαίσιο του IoT, αισθητήρες που υπάρχουν πλέον παντού έχουν τη δυνατότητα συλλογής και μετάδοσης δεδομένων τα οποία αποθηκεύονται και

υπόκεινται σε επεξεργασία σε περιβάλλον CC. Τα δεδομένα αυτά είναι μεγάλου αριθμού σε ποσότητα και παρουσιάζουν μια μεταξύ τους σχετικότητα που ξεπερνούν κατά πολύ τις δυνατότητες των αρχιτεκτονικών και των υποδομών των πληροφοριακών συστημάτων των υφιστάμενων επιχειρήσεων, ενώ η απαίτηση για την επεξεργασία τους σε πραγματικό χρόνο, αναδεικνύει το ποσό της διαθέσιμης υπολογιστικής ικανότητάς τους (Εικ. 2.1) [37].

Καθώς ο όγκος των μεγάλων συνόλων δεδομένων αυξάνεται δραματικά, δημιουργούνται διάφορα ζητήματα και προκλήσεις που απαιτούν άμεσες λύσεις. Οι συνεχείς εξελίξεις των τεχνολογιών της πληροφορίας καθιστούν πιο εύκολη τη δημιουργία δεδομένων. Για παράδειγμα, σύμφωνα με έρευνα της εταιρείας Statista του Σεπτεμβρίου του 2021, τον Φεβρουάριο του 2020, κάθε λεπτό μεταφορτώθηκαν κατά μέσο όρο πάνω από 500 ώρες βίντεο στο YouTube [40]. Αυτό το στατιστικό στοιχείο αποδεικνύει περίτρανα την πρώτη βασική πρόκληση που αντιμετωπίζεται στο πλαίσιο των Big Data ως προς τη συλλογή και ενσωμάτωση μεγάλου όγκου δεδομένων από ευρέως διανεμημένες πηγές δεδομένων [41]. Μια δεύτερη πρόκληση αφορά τον συνεχώς αυξανόμενο όγκο ήδη συλλεχθέντων δεδομένων, γεγονός που προκαλεί πρόβλημα αποθήκευσης και διαχείρισης μιας τόσο τεράστιας και ετερογενούς φύσης ποσότητας συνόλων δεδομένων μέσω των υφιστάμενων υποδομών hardware υλικού και λογισμικού [42]. Μια τρίτη πρόκληση αφορά την αποτελεσματική εξόρυξη των συνόλων δεδομένων σε διαφορετικά επίπεδα μέσω τεχνικών ανάλυσης, μοντελοποίησης, οπτικοποίησης, πρόβλεψης και βελτιστοποίησης, ώστε να εξαχθούν πολύτιμες πληροφορίες ικανές να βελτιώσουν τις λήψεις αποφάσεων, λαμβάνοντας υπόψη την ετερογένεια, την επεκτασιμότητα, τον πραγματικό χρόνο, την πολυπλοκότητα και τη διατήρηση του απορρήτου των Big Data [43].

2.2 Ορισμός των Big Data

Τα Big Data αποτελούν μια πολύ αφηρημένη έννοια. Αν και, σε γενικές γραμμές, η σημασία της τεχνολογίας έχει αναγνωριστεί, το γεγονός ότι ακόμα βρίσκεται σε πρώιμο στάδιο ανάπτυξης ίσως να αποτελεί μια βασική παράμετρο εμφάνισης διαφορετικών απόψεων σχετικά με τον ορισμό της και ανυπαρξίας ενός ορισμού κοινής αποδοχής. Γενικότερα, τα Big Data αναφέρονται σε σύνολα δεδομένων που δεν ήταν δυνατό να γίνουν αντιληπτά, να αποκτηθούν, να διαχειριστούν και να επεξεργαστούν με χρήση παραδοσιακών εργαλείων λογισμικού ή hardware υλικού, εντός ανεκτού χρόνου [44].

Παρόλα αυτά, αυτός ο ορισμός δεν καλύπτει όλες τις πτυχές που περικλείονται στον όρο των Big Data. Με βάση την εκάστοτε οπτική, επιστημονικοί και τεχνολογικοί οργανισμοί, μελετητές και αναλυτές έχουν δώσει κατά καιρούς διαφορετικούς ορισμούς της τεχνολογίας. Σύμφωνα με τους Wang και συν. (2016), οι ορισμοί αυτοί μπορούν να ταξινομηθούν σε τέσσερις μεγάλες κατηγορίες [45]: (α) της οπτικής που προσανατολίζεται στο προϊόν, (β) της οπτικής που προσανατολίζεται στη διαδικασία, (γ) της οπτικής που προσανατολίζεται στη γνώση και (δ) της οπτικής του κοινωνικού κινήματος.

Οι ορισμοί της οπτικής που προσανατολίζεται στο προϊόν (product oriented) δίνουν έμφαση στα χαρακτηριστικά των δεδομένων, όπως ο όγκος, ο ρυθμός παραγωγής ή δημιουργίας και η δομή τους. Κοινή βάση όλων αυτών των ορισμών αποτελεί η σύγκριση των ποσοτήτων των δεδομένων που παράγονται σε σύγκριση με το παρελθόν [45]. Υπό αυτή την έννοια, σε αυτήν την κατηγορία ορισμών περιλαμβάνονται ορισμοί όπως αυτός των Favaretto και συν. (2020), οι οποίοι αναφέρουν ότι [46]:

“Τα Big Data είναι μεγάλα, διαφορετικά, σύνθετα, διαμήκη και/ή κατανεμημένα σύνολα δεδομένων που δημιουργούνται από όργανα, αισθητήρες, συναλλαγές στο Διαδίκτυο, email, βίντεο, ροές κλικ και/ή όλες τις υπόλοιπες ψηφιακές πηγές που είναι διαθέσιμες σήμερα και θα υπάρξουν στο μέλλον”

Οι ορισμοί της οπτικής που προσανατολίζεται στη διαδικασία (process oriented) υπογραμμίζουν την καινοτομία των διαδικασιών που απαιτούνται και χρησιμοποιούνται για την αποθήκευση, διαχείριση, συγκέντρωση, αναζήτηση και ανάλυση των Big Data. Προκειμένου να υπογραμμιστούν οι προκλήσεις που αντιμετωπίζουν οι διαδικασίες επεξεργασίας των Big Data, στους ορισμούς αυτής της οπτικής συμπεριλαμβάνονται και όσοι αναφέρονται στην απαιτούμενη τεχνολογική υποδομή, δηλαδή στα τεχνικά εργαλεία, τις τεχνικές προγραμματισμού, αλλά και στις εξελίξεις της υπολογιστικής, της στατιστικής και των τεχνικών [45]. Έτσι, σε αυτήν την κατηγορία ορισμών περιλαμβάνονται ορισμοί όπως αυτός που αναφέρθηκε από τους Bhadani & Jothimani (2016) [47]:

“Η τεχνολογία των Big Data αποτελεί μια νέα γενιά τεχνολογιών και αρχιτεκτονικών, σχεδιασμένη ώστε να εξάγει οικονομική αξία από πολύ μεγάλους όγκους δεδομένων μεγάλης ποικιλίας, επιτρέποντας την υψηλής ταχύτητας λήψη, συλλογή και/ή ανάλυσή τους”

Οι ορισμοί της οπτικής που προσανατολίζεται στη γνώση (cognition oriented) εστιάζουν στις προκλήσεις που δημιουργούνται από τη χρήση της τεχνολογίας όσον αφορά τη γνώση και τις πληροφορίες που εμπεριέχει. Επομένως, οι ορισμοί που ανήκουν σε αυτήν την κατηγορία εξηγούν την έννοια των Big Data με βάση την υπέρβαση της ανθρώπινης ικανότητας ως προς την κατανόησή τους, υπογραμμίζοντας την αναγκαιότητα διαμεσολάβησης για την ενίσχυση της ερμηνείας τους [45]. Για παράδειγμα, οι Chen, Miao & Humar (2019) αναφέρουν ότι τα Big Data αφορούν [48]:

“τέτοια σύνολα δεδομένων που δεν ήταν δυνατό να αποκτηθούν, να αποθηκευτούν και να διαχειριστούν μέσω των λογισμικών των κλασικών βάσεων δεδομένων”

Τέλος, οι ορισμοί της οπτικής του κοινωνικού κινήματος (social movement) υπογραμμίζουν το χάσμα μεταξύ οράματος και πραγματικότητας, ειδικά τις κοινωνικοοικονομικές, πολιτιστικές και πολιτικές μεταβολές που αναμένεται να προκύψουν από τη χρήση των Big Data. Οι ορισμοί αυτής της κατηγορίας δίνουν επίσης έμφαση στο όραμα των Big Data για ενίσχυση των επιστημονικών ανακαλύψεων, της περιβαλλοντικής και βιοϊατρικής έρευνας, της εκπαίδευσης και της εθνικής ασφάλειας. Βάσει αυτής της οπτικής, οι Wang και συν. (2016) ορίζουν τα Big Data ως [45]:

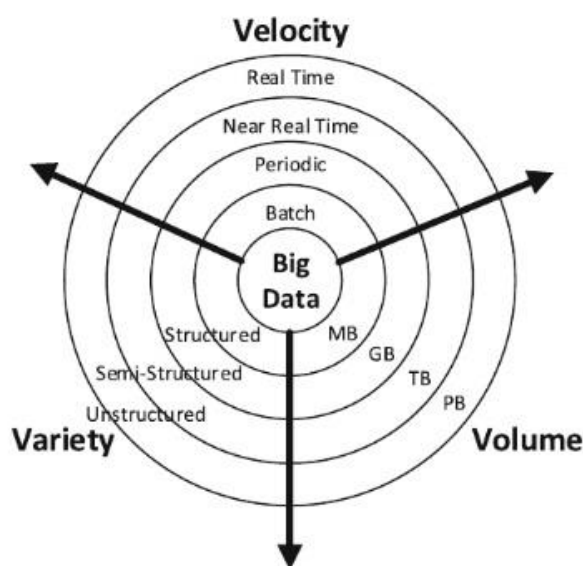
“μια τεχνολογία με δυνατότητες ποσοτικοποίησης και μεταβολής διαφόρων πτυχών της σύγχρονης ζωής, δημιουργίας επαναστατικών αλλαγών στον χώρο του management ή συμμετοχής στην μεγάλη μεταμόρφωση που απαιτεί εθνική προσπάθεια”

Η ανάλυση αυτών των τεσσάρων κατηγοριών των ορισμών των Big Data, μπορεί να παρέχει χρήσιμες πληροφορίες για την εννοιολογική ερμηνεία της τεχνολογίας, αλλά και για την άντληση των χαρακτηριστικών τους, κάποια από τα οποία θα αναφερθούν στην επόμενη ενότητα.

2.3 Χαρακτηριστικά των Big Data

Λαμβάνοντας υπόψη τον ορισμό των Chen, Miao & Humar (2019), προκύπτουν δύο συμπεράσματα. Πρώτον, ο όγκος των δεδομένων που μπορούν να χαρακτηριστούν ως Big Data αυξάνεται με την πάροδο του χρόνου ή με τις τεχνολογικές εξελίξεις. Δεύτερον, ο όγκος των δεδομένων που μπορούν να χαρακτηριστούν ως Big Data είναι διαφορετικός για τις εκάστοτε εφαρμογές διαφέρουν μεταξύ τους. Προς το παρόν, η

καθημερινή δημιουργία Big Data κυμαίνεται γενικά από αρκετά terabyte έως αρκετά petabyte [49]. Από τον ορισμό της εταιρείας McKinsey & Company, φαίνεται ότι ο όγκος δεν αποτελεί το μόνο χαρακτηριστικό των Big Data. Η συνεχώς αυξανόμενη κλίμακα παραγωγής τους και η διαχείρισή τους που δεν μπορεί να πραγματοποιηθεί μέσω των παραδοσιακών τεχνολογιών βάσεων δεδομένων, αποτελούν άλλα δύο βασικά χαρακτηριστικά της τεχνολογίας [48].

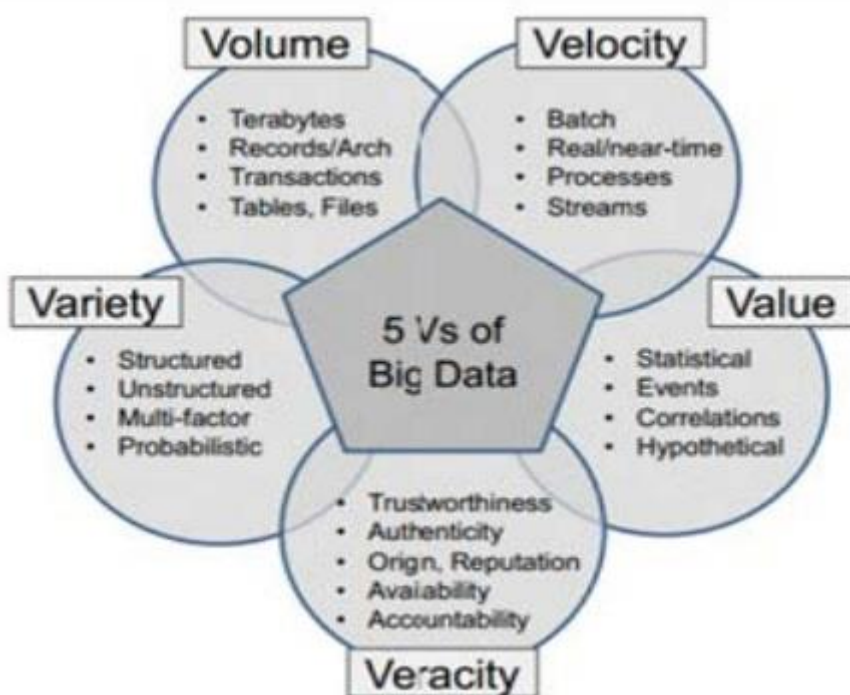


Εικόνα 2.2: Το μοντέλο 3V [50]

Αρχικά, και πιο συγκεκριμένα το 2001, οι προκλήσεις αλλά και οι ευκαιρίες που δημιουργήθηκαν από την τότε αύξηση των παραγόμενων ψηφιακών δεδομένων, καθορίστηκαν μέσω ενός μοντέλου που έγινε γνωστό ως το μοντέλο των τριών V, δηλαδή τον όγκο (Volume), την ταχύτητα (Velocity) και την ποικιλία (Variety). Αν και το συγκεκριμένο μοντέλο δεν χρησιμοποιήθηκε για τον ορισμό των Big Data, εντούτοις περιέγραφε την τεχνολογία για σχεδόν μια δεκαετία. Ο όγκος χαρακτηρίζει την μεγάλη ποσότητα δεδομένων που αποθηκεύονται σε μια υποδομή IT, και όσον αφορά τα Big Data αυτή η ποσότητα είναι τόσο πολύ μεγάλη που υπερβαίνει την χωρητικότητα της διαχειριστικής ικανότητας μιας παραδοσιακής βάσης δεδομένων. Η ταχύτητα στο πλαίσιο των Big Data μπορεί να σημαίνει τον συνεχώς αυξανόμενο ρυθμό παραγωγής δεδομένων σε όλο τον κόσμο, όπως επίσης και να αφορά το ρυθμό μετάδοσης των δεδομένων από το πέρασμα της μαζικής επεξεργασίας (batch processing) σε λειτουργία πραγματικού χρόνου. Η ποικιλία αφορά τη δυνατότητα εμφάνισης των Big Data σε διάφορες μορφές, όπως δομημένα, ημιδομημένα και μη δομημένης μορφής, όπως

αρχεία ήχου και εικόνας, ιστολογία, μηνύματα ηλεκτρονικού ταχυδρομείου και γραπτά μηνύματα (Εικ. 2.2) [50].

Με την πάροδο των ετών και την εξέλιξη της τεχνολογίας, η έννοια των Big Data άρχισε να παίρνει άλλες διαστάσεις. Με βάση τον ορισμό του R. Ramadan (2017), τα χαρακτηριστικά των Big Data γίνονται πέντε V, δηλαδή στα ήδη υπάρχοντα χαρακτηριστικά (όγκος, ποικιλία και ταχύτητα) προστίθενται η αξιοπιστία (Veracity), που χαρακτηρίζει το διαφορετικό βαθμό αξιοπιστίας των δεδομένων, ανάλογα με την προέλευση, τη διαχείριση και την επεξεργασία τους και η αξία (Value) των πληροφοριών που εμπεριέχονται στα δεδομένα (Εικ. 2.3) [51]. Ο ορισμός των Big Data μέσω του μοντέλου των πέντε V, αναγνωρίστηκε ευρέως αφού τονίζει το νόημα και την αναγκαιότητα της τεχνολογίας, δηλαδή την εξερεύνηση των τεράστιων αξιών που κρύβουν. Ταυτόχρονα όμως, ο συγκεκριμένος ορισμός εμφανίζει και το μεγαλύτερο ζήτημα που αντιμετωπίζει η τεχνολογία, που αφορά τον τρόπο εύρεσης αυτών των αξιών μέσα σύνολα δεδομένων τεράστιας κλίμακας, διαφόρων τύπων και ταχείας παραγωγής [52].



Εικόνα 2.3: Το μοντέλο 5V [51]

Εκτός από αυτά τα δύο βασικά μοντέλα ορισμού των Big Data, η συνεχιζόμενη μελέτη και έρευνα της τεχνολογίας έχουν οδηγήσει πολλούς μελετητές στο συμπέρασμα ότι

πάντα κάτι λείπει από αυτά. Μια έρευνα στο Διαδίκτυο μπορεί να ανασύρει μια σειρά από διαφορετικά μοντέλα ορισμών των Big Data, όπως τα 6V, τα 8V, τα 10V και 1C, τα 17V και 1C (volume, velocity, value, variety, veracity, validity, visualization, virality, viscosity, variability, volatility, venue, vocabulary, vagueness, verbosity, voluntariness, versatility and complexity), κλπ. [53]. Για παράδειγμα, αντί αναπαράστασης από πολλά V, οι Elankavi, Kalairasath & Udayakumar (2017) υπογράμμισαν τα χαρακτηριστικά των Big Data χρησιμοποιώντας το λεγόμενο θεώρημα HACE [54]:

“Τα Big Data δημιουργούνται από μεγάλο όγκο ετερογενών και αυτόνομων πηγών με κατανομημένο και αποκεντρωμένο έλεγχο, επιδιώκοντας να εξερευνήσουν περίπλοκες και εξελισσόμενες σχέσεις μεταξύ δεδομένων”

Αυτό όμως που έχει μεγαλύτερη σημασία, τελικά, είναι ο τρόπος διαχείρισης αυτών των δεδομένων και το πως θα γίνει η βέλτιστη εξαγωγή των αξιών που περιλαμβάνουν. Μόνο έτσι το συνονθύλευμα ενός “μάτσου δεδομένων” θα μετατραπεί σε πραγματικά Big Data [48].

2.4 Εφαρμογές των Big Data

Πρωταρχικός στόχος των εφαρμογών Big Data σε όλους τους τομείς και κλάδους είναι η ενίσχυση των επιχειρήσεων στο πλαίσιο της λήψης αποφάσεων μέσω ανάλυσης των τεράστιων όγκων δεδομένων που παράγονται καθημερινά σε όλες τις επιχειρηματικές διαδικασίες. Όπως έχει ήδη αναφερθεί, ο όγκος αυτός δεδομένων μπορεί να προέρχεται από αρχεία καταγραφής διακομιστών Ιστού, δεδομένα ροής στο Διαδίκτυο, περιεχόμενα κοινωνικών μέσων και αναφορές δραστηριότητας, κείμενα από μηνύματα email πελατών, λεπτομέρειες κλήσεων κινητών τηλεφώνων και δεδομένα μηχανημάτων που καταγράφονται από πολλούς αισθητήρες. Σε αυτό το πλαίσιο, οι επιχειρήσεις από διάφορους τομείς επενδύουν σε εφαρμογές Big Data, με σκοπό την ανάλυση των μεγάλων συνόλων δεδομένων και την ανακάλυψη όλων των κρυφών μοτίβων, των άγνωστων συσχετίσεων, τις τάσεις της αγοράς, τις προτιμήσεις των πελατών και όλες τις υπόλοιπες χρήσιμες επιχειρηματικές πληροφορίες [43].

Στο πλαίσιο της παρούσας εργασίας θα γίνει μια παρουσίαση των πλέον αντιπροσωπευτικών εφαρμογών της τεχνολογίας των Big Data, όπως οι υγειονομικές και ιατρικές εφαρμογές, οι μεταφορές, οι εφαρμογές στον οικονομικό τομέα, τα

διαδικτυακά κοινωνικά δίκτυα, το ηλεκτρονικό εμπόριο, η παρακολούθηση του περιβάλλοντος και η διαχείριση των φυσικών καταστροφών, η ασφάλεια υποδομών ζωτικής σημασίας, οι τηλεπικοινωνίες και οι βιομηχανικές εφαρμογές. Στόχος αυτής της παρουσίασης είναι η ανάδειξη της σημασίας της εφαρμογής των Big Data στους τομείς αυτούς.

2.4.1 Υγειονομική περίθαλψη

Οι εξελίξεις στον τομέα του IoT και των διασυνδεδεμένων συσκευών με αισθητήρες έχουν επιτρέψει την ολοκληρωμένη επεξεργασία των δεδομένων που μπορούν να συλλεχθούν από διάφορες υπηρεσίες παροχής υγειονομικής περίθαλψης σε πραγματικό χρόνο [55]. Εκτός από τις αμιγώς ιατρικές πηγές δεδομένων, όπως τα ηλεκτρονικά αρχεία υγείας (Electronic Health Records – EHR) και τις κλινικές αναφορές, οι πάροχοι υπηρεσιών υγειονομικής περίθαλψης μπορούν να χρησιμοποιήσουν νέες πηγές δεδομένων, όπως οι πλατφόρμες των μέσων κοινωνικής δικτύωσης, η τηλεματική και τα wearable, ώστε να είναι σε θέση να παρέχουν εξατομικευμένες υπηρεσίες ιατρικής φροντίδας και θεραπείας. Η υιοθέτηση της τεχνολογίας των Big Data σε έναν τέτοιο πλαίσιο μπορεί να οδηγήσει στη δημιουργία ενός οικοσυστήματος υγειονομικής περίθαλψης και φροντίδας που να έχει ως επίκεντρο τους ασθενείς. Η συνεργασία και η αποτελεσματική χρήση των δεδομένων που κατέχουν όλα τα ενδιαφερόμενα μέρη ενός τέτοιου οικοσυστήματος, όπως το υγειονομικό προσωπικό, το ιατρικό προσωπικό, οι ασφαλιστές και οι πάροχοι υγειονομικών υπηρεσιών, θα δώσει τη δυνατότητα παροχής υγειονομικών λύσεων με οικονομικά αποδοτικό τρόπο και βελτιωμένη εξατομικευμένη φροντίδα και περίθαλψη στους ασθενείς, δημιουργώντας τις κατάλληλες προϋποθέσεις ανάπτυξης εφαρμογών στο πλαίσιο της έξυπνης υγείας [56].

Ωστόσο, οι οργανισμοί υγειονομικής περίθαλψης αντιμετωπίζουν μοναδικές προκλήσεις όσον αφορά την υιοθέτηση των Big Data στο πλαίσιο της ανάπτυξης και της εφαρμογής της έννοιας της έξυπνης υγείας, ιδιαίτερα στην περίπτωση που αυτή βασίζεται στη χρήση απομακρυσμένων διακομιστών cloud με πολύ μεγάλες δυνατότητες υπολογιστικής [57]. Σε μια τέτοια περίπτωση, οι προκλήσεις αυτές αφορούν τα ιδιαίτερα χαρακτηριστικά των δεδομένων και των πληροφοριών υγείας, όπως ο όγκος, η ποικιλία, η αξιοπιστία και η ταχύτητα, που αναφέρθηκαν σε προηγούμενο κεφάλαιο, και που στον τομέα της υγειονομικής περίθαλψης σχετίζονται

με την επεκτασιμότητα, την αποτελεσματικότητα, τη διαφάνεια, τη διαθεσιμότητα, την ασφάλεια αλλά και την ποιότητα των παρεχόμενων εξατομικευμένων υπηρεσιών. Ως εκ τούτου, στη βιβλιογραφία, ως λύσεις αντιμετώπισης αυτών των προκλήσεων έχουν προταθεί οι περιπτώσεις χρήσης κόμβων τεχνολογίας fog με υπολογιστικές και αποθηκευτικές δυνατότητες ως συμπλήρωμα των κεντρικών διακομιστών cloud [58].

Η παροχή υπηρεσιών εξατομικευμένης ιατρικής αποτελεί μια προσέγγιση που χρησιμοποιεί πληροφορίες σχετικά με τη μοναδική γενετική σύνθεση και το περιβάλλον ενός ασθενούς, με σκοπό την προσαρμογή της ιατρικής του περίθαλψης ώστε να ανταποκρίνεται στις ιδιαίτερες ανάγκες του. Η υιοθέτηση της τεχνολογίας των Big Data στον τομέα της υγειονομικής περίθαλψης και φροντίδας έχει δώσει τη δυνατότητα εφαρμογής αλγορίθμων τεχνητής νοημοσύνης και μηχανικής μάθησης στο τεράστιο σύνολο δεδομένων που δημιουργείται συνεχώς από τα δεδομένα που συλλέγονται από τις ιατρικές έρευνες πάνω στο μοριακό προφίλ των ασθενών, από τα δεδομένα που συλλέγουν οι ίδιοι οι ασθενείς μέσω των wearable συσκευών και τις ιατρικές εφαρμογές των φορητών συσκευών, καθώς και από τις διάφορες νοσοκομειακές εξετάσεις. Μια τέτοια υιοθέτηση έχει δημιουργήσει νέες ευκαιρίες έρευνας σε προγνωστικά διαγνωστικά, ιατρική ακριβείας, εικονική διάγνωση, παρακολούθηση ασθενών και ανακάλυψη φαρμάκων για στοχευμένες θεραπείες [59].

Η χρήση της τεχνολογίας των Big Data στην ιατρική, σε συνδυασμό με άλλες αναδυόμενες τεχνολογίες, όπως η τεχνική νοημοσύνη και η μηχανική μάθηση, έχει αφυπνίσει το ενδιαφέρον της ακαδημαϊκής και ερευνητικής κοινότητας του κλάδου, αλλά και των ρυθμιστικών φορέων, για την παροχή νέων εργαλείων που μπορούν να χρησιμοποιηθούν πάνω στην ιατροφαρμακευτική έρευνα, όπως ο γονιδιωματικός υπολογισμός [60], σχετικά με την ανακάλυψη προτύπων ασθενειών και την ανίχνευση και πρόβλεψη πρώιμων επιδημιών και πανδημιών [25]. Για παράδειγμα, οι Das και συν. (2018) ανέλυσαν τις γενικές πιθανές χρήσεις των Big Data στην ιατρική, με σκοπό την πρόβλεψη των περιπτώσεων καρδιακής προσβολής, εγκεφαλικών νόσων, φυματίωσης και HIV/AIDS, τη διάγνωση των περιπτώσεων χρόνιων νεφρικών νόσων, καθώς και την ανάλυση των δεδομένων συγκεκριμένων νόσων με σκοπό την πρόβλεψη της έκβασης της εκάστοτε νόσου [61]. Επίσης, ο L. E. Bayne (2018) διερεύνησε τη χρήση των Big Data στην περίπτωση της ιατρικής φροντίδας των νεογνών [62].

Η χρήση της τεχνολογίας των Big Data στην ιατρική, όμως, έχει μελετηθεί ιδιαίτερα και για άλλους λόγους. Οι Lee και Yoon (2017) συζητούν ορισμένες τεχνικές πτυχές των εφαρμογών των Big Data στην ιατρική, εστιάζοντας στις προκλήσεις που αναμένεται να αντιμετωπίσουν [63]. Οι Ristevski και Chen (2018) αναφέρουν οι σημαντικότερες από αυτές τις προκλήσεις είναι η διατήρηση του απορρήτου και η ασφάλεια [64]. Τέλος, οι Istepanian & Al-Anzi (2018) αναδεικνύουν τη σημασία των Big Data για τον τομέα του m-health [65].

2.4.2 Μεταφορές

Σύμφωνα με τους Janev και συν. (2020), οι μεταφορές στο πλαίσιο των έξυπνων πόλεων (έξυπνες μεταφορές) αποτελούν ίσως το μεγαλύτερο θέμα μελέτης των εφαρμογών της τεχνολογίας των Big Data [25]. Για παράδειγμα, οι Grant-Muller και συν. (2017) περιέγραψαν τις επιπτώσεις που μπορεί να υπάρξουν από την εξόρυξη των Big Data που συλλέγονται από τον τομέα των μεταφορών για άλλους τομείς, όπως αυτών της ενέργειας, της υγείας και της ασφάλειας [66]. Οι Zhang και συν. (2018) μελέτησαν την εφαρμογή των Big Data για τη μείωση των ναύλων σε σύγχρονες πόλεις με κυκλοφοριακή συμφόρηση [67]. Οι Torre-Bastida και συν. (2018) εξέτασαν τις προκλήσεις των εφαρμογών των Big Data στον τομέα των μεταφορών [68]. Οι Priyan & Devi (2019) ανέλυσαν την ένταξη της τεχνολογίας των Big Data στο πλαίσιο του Διαδικτύου των Οχημάτων (IoV) [69]. Οι Mohandu & Kubendiran (2021) ανέλυσαν τη σημαντική προσφορά των Big Data στα σύγχρονα ευφυή συστήματα μεταφορών, καθώς και τις τεχνικές που μπορεί να χρησιμοποιηθούν, όπως η οπτικοποίηση σε εφαρμογές οδικής κυκλοφορίας [70]. Σχετική επίσης, είναι η έρευνα των Ghofrani και συν. (2018), οι οποίοι αναπτύσσουν τις εφαρμογές των Big Data στα συστήματα μεταφορών των σιδηροδρόμων [71].

Σύμφωνα με τους Zhu και συν. (2018), στα έξυπνα συστήματα μεταφορών, οι πηγές προέλευσης των Big Data μπορούν να ταξινομηθούν σε έξι κατηγορίες [72]: (α) αισθητήρες ροής κυκλοφορίας, (β) επεξεργαστές εικόνας-βίντεο, (γ) οχήματα και άτομα, (δ) τοποθεσίες διαδικτυακών κοινοτήτων, (ε) έξυπνες κάρτες μεταφοράς και (στ) γνώση περιοχών. Οι συγγραφείς αναφέρουν ότι τα δεδομένα που προέρχονται από τους επεξεργαστές εικόνας-βίντεο μπορούν να χρησιμοποιηθούν σε πολύ περισσότερες εφαρμογές μεταφορών σε σύγκριση με τους υπόλοιπους αισθητήρες ροής κυκλοφορίας, λόγω της ταχείας εξέλιξης στους τομείς της υπολογιστικής όρασης και

της επεξεργασίας εικόνας. Η συλλογή των δεδομένων επιτυγχάνεται με χρήση τεχνολογιών, όπως το παγκόσμιο σύστημα εντοπισμού θέσης (GPS), το bluetooth και τα ασύρματα κυβελοειδή δίκτυα. Μετά την ανάλυση των δεδομένων, οι εξαγόμενες πληροφορίες μπορούν να χρησιμοποιηθούν σε πλήθος εφαρμογών, όπως η ανίχνευση της ροής κυκλοφορίας, η ανίχνευση τυχόν προβλημάτων σε αυτή, ο έλεγχος της σηματοδότησης, ο προγραμματισμός των δημόσιων μεταφορών, κλπ. [72].

Οι Cui και συν. (2019) ανέφεραν ότι η χρήση του GPS, μπορεί να επιτρέψει στους χρήστες των smartphone να επωφεληθούν από συστήματα δρομολόγησης, λαμβάνοντας πληροφορίες σχετικά με τη συντομότερη ή ταχύτερη διαδρομή μεταξύ δύο επιθυμητών σημείων. Η χρησιμότητα ενός τέτοιου συστήματος οφείλεται στα αθροιστικά αποτελέσματα που τελικά παρουσιάζονται στους χρήστες, τα οποία προκύπτουν μετά από την ανάλυση και επεξεργασία των Big Data που συλλέγονται από πολλές διαφορετικές πηγές. Η όποια αβεβαιότητα των δεδομένων υπολογίζεται κατά μέσο όρο, με αποτέλεσμα η τελική πρόταση της συντομότερης διαδρομής να είναι σχεδόν ακριβής. Για την παροχή γρηγορότερης απόκρισης στις ανάγκες των χρηστών, τα σύγχρονα συστήματα υπολογισμού διαδρομής περιλαμβάνουν χωρικά διάσπαρτες συσκευές (διακομιστές) edge, που μπορούν να σχηματίσουν ένα δίκτυο edge cloud με σκοπό την παροχή υπολογιστικών πόρων, αλλά και πόρων αποθήκευσης και δικτύωσης, για τη διευκόλυνση της ανάλυσης των Big Data [73].

Τα αυτόνομα οχήματα αποτελούν μια ακόμη εφαρμογή των Big Data στον τομέα των έξυπνων μεταφορών. Η λειτουργία τους βασίζεται στις τεράστιες ποσότητες δεδομένων που παρέχονται συνεχώς από τους χρήστες τους και χρησιμοποιούνται για την εκπαίδευση των αλγορίθμων που διέπουν το όχημα στην λειτουργία αυτόματου πιλότου [74].

2.4.3 Οικονομικός τομέας

Εργαλεία επιχειρηματικής ευφυΐας (business intelligence) χρησιμοποιούνται ήδη από τις αρχές της δεκαετίας του 1990 με σκοπό την αύξηση της κερδοφορίας, τη μείωση του κινδύνου και τη δημιουργία ανταγωνιστικών πλεονεκτημάτων [10]. Στον οικονομικό τομέα, πολλές τράπεζες και ασφαλιστικές εταιρείες άρχισαν να χρησιμοποιούν τεχνικές μηχανικής μάθησης για την κατηγοριοποίηση και ιεράρχηση των πελατών, την αξιολόγηση του πιστωτικού κινδύνου μεμονωμένων πελατών ή

εταιρειών, κλπ. [25]. Γενικότερα, ο συγκεκριμένος τομέας είναι από τους πρώτους που υιοθετεί νέες τεχνολογίες.

Τη σύγχρονη εποχή και χάρη στην πρόοδο των τεχνολογιών της γνωστικής υπολογιστικής και της τεχνητής νοημοσύνης, οι οικονομικές εταιρείες μπορούν πλέον να χρησιμοποιούν εξελιγμένους αλγόριθμους για να αποκτήσουν γνώσεις σχετικά με τη συμπεριφορά των καταναλωτών [75]. Επίσης, η ανάλυση των δεδομένων που πηγάζουν από εσωτερικές και εξωτερικές πηγές αποτελεί σήμερα το σημαντικότερο όπλο για τον εντοπισμό των περιπτώσεων απάτης και τρωτών σημείων ασφάλειας των οικονομικών πληροφοριακών συστημάτων [76]. Επί πλέον, σε πολλές μελέτες αναφέρεται ότι η εφαρμοσμένη μηχανική μάθηση μπορεί να συμπληρωθεί με σημασιολογική γνώση, βελτιώνοντας έτσι τις ζητούμενες προβλέψεις και ταξινομήσεις, εμπλουτίζοντάς τες με συλλογιστικές εξηγήσεις που οι προσεγγίσεις της καθαρής μηχανικής μάθησης στερείται [77].

Η χρήση των Big Data μπορεί να εφαρμοστεί και σε άλλα χρηματοπιστωτικά ιδρύματα, όπως για παράδειγμα στα χρηματιστήρια, όπου ο τεράστιος όγκος και η συχνότητα των συναλλαγών καθιστούν τις παραδοσιακές λύσεις μεθόδων επεξεργασίας και υπολογισμού αργές και απαρχαιωμένες [25].

Η ανάλυση των μελετών που έχουν παρουσιαστεί μέχρι τώρα στη βιβλιογραφία και αφορούν τη χρήση των Big Data στον οικονομικό τομέα, αναδεικνύει ότι η εύρεση μοτίβων μέσω της ανάλυσης του τεράστιου όγκου οικονομικών δεδομένων αποτελεί το κλειδί για τη βελτιστοποίηση των οικονομικών διαδικασιών και την πρόληψη των περιπτώσεων απάτης [78]. Για παράδειγμα, οι Hasan, Kalipsiz & Akyokus (2017) παρουσιάζουν προσεγγίσεις πρόβλεψης των συνθηκών της αγοράς χρησιμοποιώντας μεθόδους βαθιάς μάθησης και εφαρμόζοντας τη θεωρία του προφίλ της αγοράς [79]. Οι Begenau, Farboodi & Veldkamp (2018) μελέτησαν την εξάρτηση της ανάπτυξης του οικονομικού τομέα από τη χρήση των Big Data [80]. Οι Óskarsdóttir και συν. (2019), μελετώντας το ίδιο ακριβώς θέμα, κατέληξαν στο συμπέρασμα ότι στην ανάπτυξη αυτή μεγάλο ρόλο παίζουν τα δεδομένα που συλλέγονται από τα μέσα κοινωνικής δικτύωσης και τα κινητά τηλέφωνα των χρηστών [81].

Οι Hasan, Popp & Oláh (2020), αναλύοντας την μέχρι τότε βιβλιογραφία σχετικά με την υιοθέτηση των Big Data στον οικονομικό τομέα, παρουσίασαν μια σύνοψη των εφαρμογών που μπορούν να υλοποιηθούν σε αυτό το πλαίσιο [82].

2.4.4 Μέσα κοινωνικής δικτύωσης

Η ανάπτυξη και εξέλιξη των μέσων κοινωνικής δικτύωσης, όπως τα YouTube, Facebook, Twitter, κλπ., είχε τεράστιο αντίκτυπο στα ψηφιακά μέσα και το ηλεκτρονικό εμπόριο [83]. Τα μέσα κοινωνικής δικτύωσης αποτελούν πηγή παραγωγής εξατομικευμένων Big Data κατάλληλη για εξόρυξη δεδομένων, με αρκετές εκατοντάδες χιλιάδες νέες αναρτήσεις να δημοσιεύονται κάθε λεπτό. Επίσης, αποτελούν πλατφόρμες ιδιαίτερης σημασίας για την εφαρμογή λύσεων Big Data, για διαφήμιση, προτάσεις αναζήτησης, ερωτήματα ανάρτησης ή προτάσεις σύνδεσης [84]. Η δομή των μέσων κοινωνικής δικτύωσης έχει επίσης παρακινήσει αρκετούς μελετητές να ερευνήσουν παρόμοιες αρχιτεκτονικές στον τομέα των Big Data [25]. Για παράδειγμα, οι Persico και συν. (2018) αξιολογούν την απόδοση των αρχιτεκτονικών Lambda και Kappa [85], ενώ οι Ghani και συν. (2019) παρουσιάζουν μια σύνοψη των προσεγγίσεων ανάλυσης των Big Data που πηγάζουν από τα μέσα κοινωνικής δικτύωσης [86].

2.4.5 Ηλεκτρονικό εμπόριο

Καθώς οι χρήστες του Διαδικτύου έχουν πλέον στη διάθεσή τους έναν τεράστιο αριθμό υπηρεσιών, η μεγάλη ποικιλία διαδικτυακών ιστότοπων αγορών προϊόντων αποτελεί επίσης μια συνεχή πηγή παραγωγής τεράστιου όγκου δεδομένων που μπορεί να αποθηκευτεί, να επεξεργαστεί, να αναλυθεί και να συναχθεί για τη δημιουργία μηχανών συστάσεων, μέσω χρήσης τεχνικών προγνωστικής ανάλυσης [87]. Στις περισσότερες περιπτώσεις εφαρμόζεται η έννοια του πολυ-καναλικού μάρκετινγκ (multi-channel marketing) για τη βελτιστοποίηση της παρουσίας των προϊόντων στα μέσα που τροφοδοτούνται στους χρήστες, ως μέσο για την αύξηση της αφοσίωσής τους [84]. Δεν είναι τυχαίο, άλλωστε, ότι μια συγκεκριμένη διαφήμιση αρχίζει να εμφανίζεται στους χρήστες αμέσως μετά την αναζήτησή τους για μια συγκεκριμένη κατηγορία προϊόντων. Η εξέταση των μοτίβων και των τάσεων συμπεριφοράς των χρηστών επιτρέπει την κατηγοριοποίηση των προσφορών σε αυτούς με τον καλύτερο δυνατό τρόπο, ώστε η σωστή προσφορά να παρουσιάζεται ακριβώς όταν χρειάζεται, μεγιστοποιώντας έτσι τις μετατροπές των προσφορών σε πωλήσεις [88]. Ωστόσο, το περιεχόμενο των συστάσεων, που συνάγονται από πηγές Big Data, δεν σχετίζεται μόνο με το μάρκετινγκ και τις πωλήσεις, αλλά χρησιμοποιείται και για τη σωστή εμφάνιση των πληροφοριών που σχετίζονται με τον χρήστη. Ορισμένες εταιρείες μηχανών

αναζήτησης έχουν δηλώσει δημόσια ότι η αρχιτεκτονική της υποδομής τους βασίζεται στην χρήση της τεχνολογίας των Big Data, κάτι που δεν προκαλεί έκπληξη, λαμβάνοντας υπόψη τον όγκο των δεδομένων που πρέπει να επεξεργαστούν [25].

2.4.6 Παρακολούθηση περιβάλλοντος και διαχείριση φυσικών καταστροφών

Η παρακολούθηση του περιβάλλοντος περιλαμβάνει τη συλλογή μετρήσεων με σκοπό την αξιολόγηση της κατάστασής του. Οι τεχνολογικές εξελίξεις στον τομέα της απομακρυσμένης ανίχνευσης με χρήση δορυφόρων και ραντάρ έχουν δημιουργήσει νέες δυνατότητες στους κλάδους της ωκεανογραφίας, της μετεωρολογίας, της δασοκομίας, της γεωργίας και των κατασκευών (πολεοδομίας) [89]. Σύμφωνα με τους Liang & Wang (2019), τα συστήματα απομακρυσμένης ανίχνευσης του περιβάλλοντος, με βάση την απόσταση των αισθητήρων και της περιοχής που παρακολουθείται, μπορεί να υποδιαιρεθούν σε τρεις μεγάλες κατηγορίες [90]: (α) τα δορυφορικά συστήματα μέτρησης, (β) τα εναέρια συστήματα παρακολούθησης και (γ) τα συστήματα παρακολούθησης εδάφους.

Η πρώτη κατηγορία, τα δορυφορικά συστήματα μέτρησης, χρησιμοποιείται κυρίως για τη μελέτη της γης και του μεταβαλλόμενου περιβάλλοντός της. Η πιο πολύτιμη πηγή δεδομένων αυτής της κατηγορίας είναι το Landsat, ένα κοινό δορυφορικό πρόγραμμα του USGS και της NASA, που παρατηρεί συνεχώς τη Γη από το 1972 μέχρι σήμερα [91]. Τα εναέρια συστήματα παρακολούθησης περιλαμβάνουν όργανα παρακολούθησης, όπως τα συστήματα ανίχνευσης και εμβέλειας φωτός (LIDAR), που επιτρέπουν την καλύτερη παρακολούθηση σημαντικών ατμοσφαιρικών ειδών όπως το όζον, το μονοξειδίο του άνθρακα, οι υδρατμοί, οι υδρογονάνθρακες και το υποξείδιο του αζώτου, καθώς και μετεωρολογικές παραμέτρους, όπως η ατμοσφαιρική πυκνότητα, η πίεση και η θερμοκρασία [92]. Στην τρίτη κατηγορία περιλαμβάνονται επίγεια όργανα παρακολούθησης και ασύρματα δίκτυα αισθητήρων (WSN), που χρησιμοποιούνται για την παρακολούθηση εξωτερικών χώρων, όπως αγροκτημάτων και τροπικών δασών, αλλά και για το συνεχή έλεγχο της ποιότητας του νερού και των φυσικών καταστροφών (ηφαιστειακών εκρήξεων, σεισμών, κλπ.) [93].

Στην περίπτωση της παρακολούθησης του περιβάλλοντος, η τεχνολογία των Big Data σε συνδυασμό με αυτή του IoT, δημιουργούν ένα οικοσύστημα εφαρμογών απομακρυσμένης ανίχνευσης, όπως της έξυπνης γεωργίας [94], της συνεχούς

παρακολούθησης των τροπικών δασών [95], του σχεδιασμού της βιοποικιλότητας biodiversity planning [96], κλπ.

Λόγω των συνεχώς μεταβαλλόμενων κλιματικών συνθηκών, οι φυσικές καταστροφές, όπως οι πλημμύρες, οι κατολισθήσεις, οι ξηρασίες, οι σεισμοί, οι πυρκαγιές, κλπ., είναι πλέον πολύ συχνά φαινόμενα. Η συνεχής παρακολούθηση αυτών των γεγονότων δημιουργεί έναν σημαντικό όγκο δεδομένων, τα οποία πρέπει να υποβάλλονται σε επεξεργασία σε πραγματικό χρόνο, για την αποφυγή των άμεσων ή έμμεσων κινδύνων που πηγάζουν από αυτά. Και στην περίπτωση της διαχείρισης των φυσικών καταστροφών, η τεχνολογία των Big Data σε συνδυασμό με άλλες αναδυόμενες τεχνολογίες, όπως το IoT, η μηχανική μάθηση, τα ασύρματα κυψελοειδή δίκτυα πέμπτης γενιάς (5G), κλπ., είναι σε θέση να δημιουργήσουν μια μεγάλη γκάμα εφαρμογών απομακρυσμένης ανίχνευσης, βελτιώνοντας την αποτελεσματικότητα των στρατηγικών διαχείρισης των καταστροφών και διευκολύνοντας την υλοποίηση των διαδικασιών εκκένωσης [97].

2.4.7 Ασφάλεια υποδομών ζωτικής σημασίας

Η εφαρμογή τεχνικών ανάλυσης των Big Data είναι ιδιαίτερα σημαντική για την ασφάλεια των υποδομών ζωτικής σημασίας, καθώς μπορεί να εξαγάγει την αξία που εμπεριέχεται σε δεδομένα που σχετίζονται με την ασφάλεια και προέρχονται από πηγές, όπως ενεργοποιητές έγκυρης προειδοποίησης, και να οδηγήσει σε νέες μορφές επιτήρησης και πρόληψης. Υποστήριξη αυτής της εφαρμογής, αναμένεται να υπάρξει από τον συνεχώς αυξανόμενο αριθμό των διασυνδεδεμένων συσκευών στο πλαίσιο της χρήσης της τεχνολογίας της τεχνικής νοημοσύνης στα δίκτυα 5G [98]. Ως υποδομές ζωτικής σημασίας θεωρούνται τα αεροδρόμια, οι σιδηροδρομικοί σταθμοί, οι σταθμοί του μετρό και τα δίκτυα παροχής ηλεκτρικής ενέργειας. Οι υποδομές τέτοιου μεγέθους είναι δύσκολο να παρακολουθούνται λόγω της πολύπλοκης διάταξής τους και της ποικιλίας των οντοτήτων που μπορεί να περιλαμβάνουν. Σε καταστάσεις έκτακτης ανάγκης, τα διάφορα συστήματα ελέγχου και παρακολούθησης που περιέχουν, όπως τα συστήματα πυροπροστασίας, θέρμανσης, εξαερισμού και κλιματισμού, εκκένωσης και ελέγχου πρόσβασης, απεικόνισης πληροφοριών, κλπ., μπορούν να στείλουν συνολικά χιλιάδες ειδοποιήσεις στην αίθουσα ελέγχου κάθε δευτερόλεπτο. Ο συνδυασμός όλων αυτών των ειδοποιήσεων με ουσιαστικό τρόπο, αυξάνει την επίγνωση της κατάστασης [99].

Η χρήση εργαλείων επεξεργασίας και ανάλυσης των Big Data μπορεί να καθορίσει και να ενισχύσει την απόδοση των συστημάτων επεξεργασίας των ειδοποιήσεων για την καλύτερη επίγνωση της κατάστασης και τη λήψη των απαραίτητων μέτρων, καθώς και να βελτιώσει τις διαδικασίες διαχείρισης των έκτακτων αναγκών [25]. Μια τέτοια χρήση παρόλα αυτά καλείται να αντιμετωπίσει, εκτός από τις προκλήσεις του όγκου, της ταχύτητας και της ποικιλίας, που προκύπτουν από την επεξεργασία του μεγάλου αριθμού ετερογενών δεδομένων, τα οποία εξάγονται από πολλαπλές πηγές [43], και προκλήσεις που αφορούν: (α) την οπτικοποίηση σε πραγματικό χρόνο και την επακόλουθη αλληλεπίδραση με υπολογιστικές μονάδες, προκειμένου να βελτιωθεί η επίγνωση της κατάστασης και να επιταχυνθεί η λήψη αποφάσεων [100] και (β) την ανάπτυξη προηγμένων τεχνικών σημασιολογικής ανάλυσης και μηχανικής μάθησης για αναγνώριση νέων προτύπων που θα βασίζεται σε προκαθορισμένα σενάρια έκτακτης ανάγκης και δημιουργία νέων διαδικασιών έγκαιρης προειδοποίησης ή αξιόπιστων σχεδίων δράσης [101].

2.4.8 Τηλεπικοινωνίες

Όπως έχει ήδη αναφερθεί, τα smartphone αποτελούν μία από τις πηγές παραγωγής δεδομένων. Αυτό συνεπάγεται την εφαρμογή των Big Data στον τομέα των τηλεπικοινωνιών. Ιδιαίτερα, η ανάπτυξη των ασύρματων κυψελοειδών δικτύων πέμπτης γενιάς (5G) αναμένεται να δημιουργήσει τις κατάλληλες συνθήκες συλλογής και ανάλυσης δεδομένων σε πραγματικό χρόνο και να υποστηρίξει νέες δυνατότητες των συστημάτων επιχειρηματικής νοημοσύνης και τεχνητής νοημοσύνης [102].

Ένα από τα βασικά μελήματα των παρόχων υπηρεσιών κινητής τηλεφωνίας, τηλεόρασης και Διαδικτύου, είναι η διατήρηση της βιωσιμότητας των επιχειρήσεών τους, κάτι που επιτυγχάνεται μέσω διατήρησης των πελατών τους. Επομένως, η ανάλυση των μοτίβων συμπεριφοράς των πελατών με σκοπό την πρόβλεψη της στάσης τους ως προς την αλλαγή παρόχου, αποτελεί μία εφαρμογή των Big Data που μπορεί να επιτρέψει στις συγκεκριμένες εταιρείες να ενεργήσουν έγκαιρα, προσφέροντας διάφορα κίνητρα διατήρησης των πελατών που ήδη έχουν [103]. Εκτός από αυτή την επιχειρηματική πτυχή, οι εταιρείες τηλεπικοινωνιών χρησιμοποιούν λύσεις ανάλυσης των Big Data που συλλέγονται από χρήστες κινητών τηλεφώνων, με σκοπό την άντληση πληροφοριών που μπορούν να χρησιμοποιήσουν, για την αξιολόγηση και την

αντιμετώπιση των προβλημάτων των δικτύων τους, βελτιώνοντας έτσι την ποιότητα των υπηρεσιών που παρέχουν [104].

Οι παραπάνω χρήσεις των Big Data στον τομέα των επικοινωνιών είναι εφικτές, καθώς όλη η επικοινωνία των χρηστών μεταφέρεται μέσω των δεδομένων των υπηρεσιών 4G και 5G που παρέχονται από τους φορείς εκμετάλλευσης των δικτύων αυτών. Τα υπάρχοντα εργαλεία ανάλυσης των Big Data που συλλέγονται σε αυτές τις υπηρεσίες δεν είναι σε θέση να πραγματοποιήσουν καθολική επεξεργασία αυτών των δεδομένων, αφήνοντας πολλά χρήσιμα στοιχεία των πληροφοριών που προκύπτουν ανεκμετάλλευτα. Στη βιβλιογραφία έχει παρουσιαστεί μεγάλος αριθμός μελετών που καλύπτουν αυτήν την πτυχή. Για παράδειγμα, οι Zahid και συν. (2019) παρουσίασαν μια βιβλιογραφική ανασκόπηση των αρχιτεκτονικών δομών των συγκεκριμένων εφαρμογών [105]. Οι Xu και συν. (2019) εξέτασαν τη χρήση των Big Data στις πλατφόρμες των ασύρματων κυψελοειδών δικτύων LTE/5G [106]. Τέλος, οι Amin και συν. (2019) μελέτησαν τρόπους ανάλυσης των Big Data για την πρόβλεψη της στάσης των πελατών των εταιρειών τηλεπικοινωνίας ως προς την αλλαγή παρόχου υπηρεσιών [107].

2.4.9 Βιομηχανικές εφαρμογές

Οι εξελίξεις στον τομέα της βιομηχανίας είναι συνεχείς, εξαρτώμενες από τις τεχνολογικές εξελίξεις. Η σύγχρονη βιομηχανική περίοδος χαρακτηρίζεται ως η τέταρτη περίοδος της βιομηχανικής επανάστασης (Industry 4.0), μια ονομασία που της δόθηκε για να περιγράψει την υφιστάμενη τάση της αυτοματοποίησης και της ανταλλαγής δεδομένων στις τεχνολογίες παραγωγής. Το Industry 4.0 αφορά την αυτοματοποίηση των διαδικασιών και τη βελτίωση της αποτελεσματικότητάς τους, που επιτυγχάνονται μέσω χρήσης αναδυόμενων τεχνολογιών, όπως το IoT και το edge computing [108]. Στην εποχή του Industry 4.0, οι έξυπνες υπηρεσίες παραγωγής πρέπει να λειτουργούν λαμβάνοντας υπόψη τις πολλαπλές ροές δεδομένων, οι οποίες συνήθως δημιουργούνται από καταναμημένους αισθητήρες, σχεδόν σε πραγματικό χρόνο [109]. Όπως συμβαίνει και σε άλλους κλάδους, ο πλήρης ψηφιακός μετασχηματισμός των βιομηχανικών μονάδων παραγωγής απαιτεί την ύπαρξη μιας αποτελεσματικής και ευέλικτης υποδομής που να είναι σε θέση να διαχειρίζεται τα δεδομένα αυτά και να συνδέεται με υπολογιστικά συστήματα και μηχανές γνωστικής συλλογιστικής [110].

Η τεχνολογία του edge computing παίζει σημαντικό ρόλο στις εφαρμογές του Industry 4.0. Στις εφαρμογές αυτές, που είναι γνωστές ως έξυπνες κατασκευές (smart manufacturing), θα πρέπει να υποστηρίζεται η αμφίδρομη επικοινωνία μεταξύ των μηχανημάτων, μεταξύ των μηχανημάτων και του cloud ή μεταξύ των μηχανημάτων και των πυλών επικοινωνίας. Με βάση μια τέτοια επικοινωνιακή δυνατότητα, τα δεδομένα πρέπει να μεταφέρονται από και προς τους χρήστες και τους παρόχους των υπηρεσιών, να αποθηκεύονται και να επεξεργάζονται, προκειμένου να είναι δυνατή η εκμετάλλευση της συναγόμενης γνώσης από τα δεδομένα που λαμβάνονται από τους αισθητήρες. Σε μια τέτοια πολυεπίπεδη αρχιτεκτονική των εφαρμογών των έξυπνων κατασκευών, οι παρεχόμενες υπηρεσίες παίζουν κεντρικό ρόλο και ο σχεδιασμός της (επιλογή αλγορίθμων/μοντέλων) εξαρτάται από την εκάστοτε εφαρμογή [25].

Οι Vater, Harscheidt & Knoll (2019), μελετώντας τον τρόπο με τον οποίο οι αναδυόμενες τεχνολογίες, όπως οι IoT, Big Data και cloud computing, διαδραματίζουν σημαντικό ρόλο στην μετάβαση στην επόμενη γενιά βιομηχανικών εφαρμογών, αναφέρουν ότι η επεξεργασία και ανάλυση των δεδομένων που χρησιμοποιούνται σε αυτές τις εφαρμογές μπορούν να παρέχουν πολύτιμες πληροφορίες που να βελτιώσουν το επίπεδο του εποπτικού ελέγχου των βιομηχανικών διεργασιών. Η ανάλυση των Big Data σε συνδυασμό με τεχνικές τεχνητής νοημοσύνης, ενισχύουν επίσης τις προσπάθειες βελτίωσης της ποιότητας των προϊόντων και των διαδικασιών [111].

ΚΕΦΑΛΑΙΟ 3: ΑΛΥΣΙΔΑ ΑΞΙΑΣ BIG DATA

3.1 Η αξία των Big Data

Το 2011, η εταιρεία McKinsey & Company αναγνώρισε τη δημιουργία αξιών από τη χρήση της τεχνολογίας των Big Data, πραγματοποιώντας μια εις βάθος έρευνα σε τομείς, όπως η υγειονομική περίθαλψη στις ΗΠΑ, η διοίκηση του δημόσιου τομέα στην ΕΕ, το λιανικό εμπόριο στις ΗΠΑ, η παγκόσμια παραγωγή και τα παγκόσμια προσωπικά δεδομένα τοποθεσίας. Το συμπέρασμα που προέκυψε από τη συγκεκριμένη μελέτη της εταιρείας ήταν ότι τα Big Data μπορούν να ενισχύσουν την παγκόσμια οικονομία, να βελτιώσουν την παραγωγικότητα και την ανταγωνιστικότητα των επιχειρήσεων και του δημόσιου τομέα, καθώς και να δημιουργήσουν τεράστια οφέλη για τους καταναλωτές [112].

Στον κόσμο της ψηφιοποίησης, οι εταιρείες και οι οργανισμοί, όλο και περισσότερο, αναζητούν τρόπους χρήσης των Big Data για τη δημιουργία και την αποτύπωση των αξιών που εμπεριέχουν [113]. Η χρήση των Big Data δίνει στους CEO των επιχειρήσεων τη δυνατότητα να γνωρίζουν ουσιαστικά περισσότερα για αυτές, με σκοπό το μετασχηματισμό αυτής της γνώσης σε καλύτερη λήψη αποφάσεων και βελτιωμένη απόδοση [114]. Για παράδειγμα, το Netflix μπορεί να συλλέγει δεδομένα παρακολουθώντας τη συμπεριφορά περιήγησης και αναζήτησης των πελατών του και προσδιορίζει τις προτιμήσεις τους με βάση τις βαθμολογίες που δίνουν στο περιεχόμενο. Αναλύοντας αυτά τα δεδομένα, το Netflix είναι σε θέση να προβλέψει τη δημοτικότητα του περιεχομένου του και να το σχεδιάσει με βάση τις πληροφορίες των δεδομένων που παράγονται από τους πελάτες του [112].

Κατά την ανάλυση των Big Data επιδιώκεται η εξαγωγή και συγκέντρωση πολύτιμων πληροφοριών και η μετατροπή τους σε επιχειρηματικά πλεονεκτήματα. Κάτι τέτοιο είναι εφικτό, καθώς, λόγω των βασικών χαρακτηριστικών τους (όγκος, ταχύτητα και ποικιλία), η ανάλυση των Big Data εξάγει πολύ περισσότερες πολύτιμες πληροφορίες από αυτές που προέκυπταν από την ανάλυση των παραδοσιακών βάσεων δεδομένων [115]. Επιπλέον, η σταθερή μείωση του κόστους των υπολογιστικών συστημάτων, λόγω της τεχνολογικής εξέλιξης στα διάφορα τμήματά τους, όπως η αποθήκευση, η μνήμη, η επεξεργασία και το εύρος ζώνης, συνεπάγεται ότι οι παλαιότερες δαπανηρές

προσεγγίσεις και τεχνικές ανάλυσης των δεδομένων καθίστανται πλέον πιο οικονομικές [116].

Η χρήση της τεχνολογίας των Big Data σε συνδυασμό με τις αυξημένες δυνατότητες υπολογιστικής ισχύος των σύγχρονων πληροφοριακών συστημάτων, επιτρέπουν την αύξηση της ευφυΐας και της καινοτομίας των εταιρειών και των οργανισμών, σε επίπεδο που δεν μπορούσε να επιτευχθεί παλιότερα [2]. Σύμφωνα με τους Lee & Holt (2021), όσο περισσότερο οι εταιρείες αναγνωρίζουν το γεγονός ότι η λειτουργία τους βασίζεται στα δεδομένα, τόσο καλύτερα θα επιδίδονται σε αντικειμενικές μετρήσεις και αναλύσεις τους, με τα αντίστοιχα οικονομικά και λειτουργικά αποτελέσματα. Οι συγγραφείς αναφέρουν ότι η αποδοτικότητα και ο ανταγωνισμός των εταιρειών συνδέεται σε μέγιστο βαθμό με την αξία της ανάλυσης των δεδομένων [117].

Παρόλο που οι οργανισμοί διερευνούν τρόπους ανάπτυξης και αξιοποίησης των Big Data για τη δημιουργία και την αποτύπωση των αξιών που εμπεριέχουν, οι περισσότεροι εξακολουθούν να έχουν πρόβλημα με το πώς να ανακτήσουν την αξία αυτή. Το ζήτημα αυτό προκύπτει από το γεγονός ότι οι εταιρείες είναι σε θέση να αναλύσουν μόνο ένα μικρό ποσοστό των δεδομένων που συλλέγουν και αποθηκεύουν, κάτι που οδηγεί εύκολα σε ένα υπερκορεσμό πληροφοριών. Για παράδειγμα, οι πάροχοι χρηματοπιστωτικών καρτών και υγειονομικής περίθαλψης απορρίπτουν περίπου το 80 με 90% των δεδομένων που δημιουργούν [118]. Πράγματι, η απλή κατοχή ενός πολύτιμου πόρου δεν εγγυάται την ικανότητα δημιουργίας αξίας [112]. Οι Davis, Pinto & Di Maddaloni (2020) αναφέρουν ότι η δημιουργία αξίας αποτελεί το σχετικό ποσό της αξίας που επιτυγχάνεται υποκειμενικά από τον στόχο που έχει θέσει κάποιος και το οποίο θα πρέπει τουλάχιστον να μεταφραστεί στην προθυμία του να ανταλλάξει την αξία που λαμβάνει με ένα χρηματικό ποσό [119]. Επομένως, για τη δημιουργία αξίας, οι επιχειρήσεις πρέπει να συσσωρεύουν, να συνδυάζουν και να εκμεταλλεύονται αποτελεσματικά τους πόρους που διαθέτουν, οι οποίοι στη συγκεκριμένη περίπτωση είναι τα δεδομένα [120]. Ενώ παλαιότερα, η κατανόηση του τρόπου διαχείρισης των πόρων μιας εταιρείας για τη δημιουργία αξίας αποτελούσε σοβαρό ζήτημα, τα τελευταία χρόνια, μελετητές σημείωσαν την αναγκαιότητα για εμπειρική διερεύνηση της διαδικασίας μετασχηματισμού των Big Data για την επίτευξη κάτι τέτοιου [121].

Επομένως, τα δεδομένα αποτελούν έναν σημαντικό παράγοντα παραγωγής που θα μπορούσε να συγκριθεί με τα υλικά περιουσιακά στοιχεία και το ανθρώπινο κεφάλαιο. Καθώς, οι τεχνολογικές εξελίξεις στα πολυμέσα, στα μέσα κοινωνικής δικτύωσης και στο IoT είναι ραγδαίες, οι επιχειρήσεις θα εξακολουθούν να συλλέγουν ολοένα και περισσότερες πληροφορίες, οδηγώντας σε μια εκθετική αύξηση του όγκου των δεδομένων. Τα Big Data τείνουν να έχουν τεράστιες και αυξανόμενες δυνατότητες δημιουργίας αξιών για τις επιχειρήσεις και τους καταναλωτές [28].

3.2 Αλυσίδα αξίας δεδομένων

Η δημιουργία αξίας αποτελεί στόχο κάθε οργανισμού. Ωστόσο, σε έναν κόσμο που αλλάζει όσον αφορά τη δημιουργία δεδομένων, η εύρεση στρατηγικών για τη βιώσιμη δημιουργία νέας αξίας και τη βελτίωση των λειτουργιών των οργανισμών, είναι μια αρκετά δύσκολη και επώδυνη διαδικασία. Αυτό σημαίνει ότι για τη δημιουργία αξίας μπορεί να απαιτηθεί επανεξέταση των ήδη χρησιμοποιούμενων διαδικασιών δημιουργίας της, καθώς και μια αποτελεσματική και ενημερωμένη διαχείριση τους, τόσο για τους ίδιους τους οργανισμούς όσο και για τους συνεργάτες τους. Η αλυσίδα αξίας αποτελεί μια από τις βασικές λύσεις σε τέτοια ζητήματα [122].

Η αλυσίδα αξίας αποτελείται από την ανάλυση όλων των δραστηριοτήτων καθώς και των αλληλεπιδράσεών τους, προσδιορίζει τις πηγές ενός δυνητικού ανταγωνιστικού πλεονεκτήματος και οργανώνει διαδικασίες που συμβάλλουν άμεσα ή έμμεσα στη δημιουργία προστιθέμενης αξίας για τον οποιοδήποτε οργανισμό. Αυτό σημαίνει ότι η ανάλυση της αλυσίδας αξίας περιλαμβάνει τον τεμαχισμό του συνόλου των δραστηριοτήτων μιας επιχείρησης σε στρατηγικά σημαντικές δραστηριότητες με σκοπό την κατανόηση του αντίκτυπου τους [123]. Ωστόσο, στην εποχή της ψηφιοποίησης των δραστηριοτήτων, η προσαρμογή των διαδικασιών έχει γίνει όλο και πιο περίπλοκη. Παρόλα αυτά, η ψηφιοποίηση των διαδικασιών των οργανισμών είναι πιο αποτελεσματική και αναδεικνύει μια συχνά παραμελημένη πηγή αξίας, που δεν είναι άλλη από τις πληροφορίες. Εξαιτίας αυτού, η παραδοσιακή έννοια της αλυσίδας αξίας έχει καταστεί ακατάλληλη, πράγμα που σημαίνει ότι έπρεπε να υιοθετηθεί μια προσέγγιση που να προσανατολίζεται στα δεδομένα παρά στο προϊόν ή στις διαδικασίες. Αυτός είναι ο λόγος για τον οποίο η αλυσίδα αξίας δεδομένων υιοθετείται όλο και περισσότερο από τους οργανισμούς που βασίζονται κυρίως στα δεδομένα για εξαγωγή χρήσιμων πληροφοριακών συσχετισμών [124].

Η αλυσίδα αξίας δεδομένων είναι ένας μηχανισμός που ορίζει ένα σύνολο επαναλαμβανόμενων διεργασιών με σκοπό την εξαγωγή της αξίας των δεδομένων. Ο μηχανισμός αυτός αποτελείται από τέσσερα βασικά διακριτά βήματα, τα οποία επαναλαμβάνονται σε ολόκληρο τον κύκλο ζωής των ακατέργαστων δεδομένων με σκοπό την εξαγωγή πραγματικών πληροφοριών. Τα βήματα αυτά αφορούν (Εικ. 3.1) [122]:

- **Παραγωγή δεδομένων (Data generation):** Λήψη και καταγραφή δεδομένων
- **Συλλογή δεδομένων (Data collection):** Συλλογή, επικύρωση και αποθήκευση δεδομένων
- **Ανάλυση δεδομένων (Data analysis):** Επεξεργασία και ανάλυση δεδομένων για τη δημιουργία νέων πιθανών γνώσεων
- **Ανταλλαγή δεδομένων (Data exchange):** Παροχή των δεδομένων για χρήση, εντός του οργανισμού ή από τους εξωτερικούς συνεργάτες του



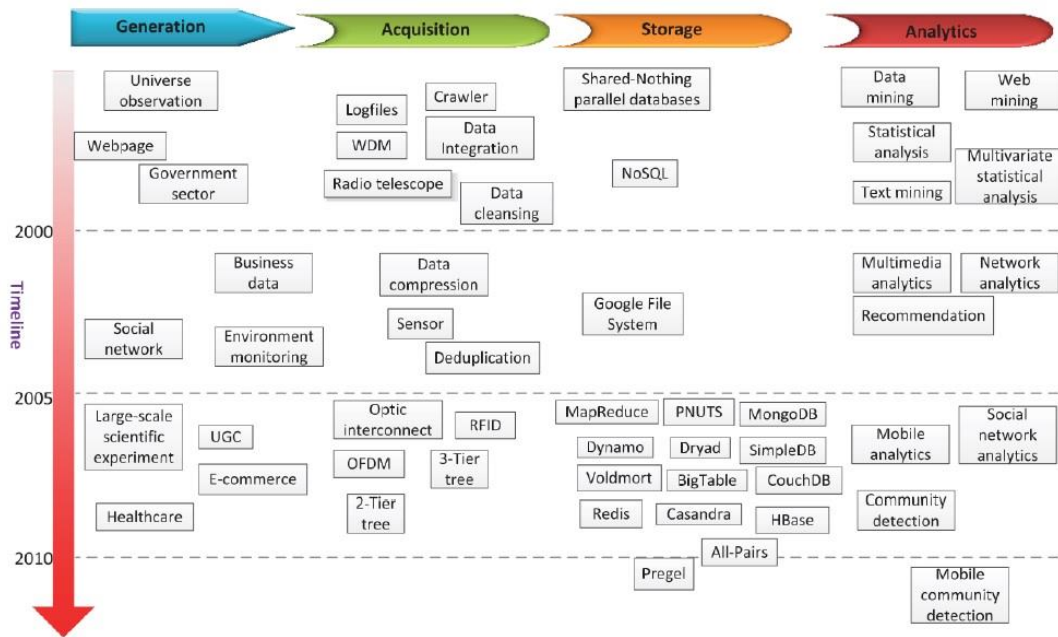
Εικόνα 3.1: Αλυσίδα αξίας δεδομένων [122]

3.3 Αλυσίδα αξίας των Big Data

Η διαχείριση των δεδομένων έχει αλλάξει πολύ με την πάροδο των ετών, λόγω της ανάδυσης νέων τεχνολογιών αποθήκευσης και επεξεργασίας, οι οποίες όχι μόνο μπορούν να διαχειριστούν καλύτερα την ποσότητα των δεδομένων που αποθηκεύεται, αλλά και τον τρόπο αποθήκευσης και διαχείρισης αυτών των δεδομένων. Αυτή η εξέλιξη κατέστη δυνατή με την εμφάνιση των Big Data [125]. Τα χαρακτηριστικά των Big Data, όπως αυτά αναφέρθηκαν σε προηγούμενη ενότητα, οδήγησαν στη

δημιουργία ενός νέου παραδείγματος, τέταρτου κατά σειρά, της επιστημονικής έρευνας, που είναι γνωστό ως υποστήριξη των ερευνητικών δραστηριοτήτων με μεγάλες απαιτήσεις σε διαχείριση δεδομένων (Data-Intensive Science) [126]. Το συγκεκριμένο παράδειγμα αποτελεί πραγματική επανάσταση στον τομέα της επιστημονικής έρευνας. Επιπλέον, η τεχνολογία των Big Data συγκεντρώνει ένα σύνολο προηγμένων εργαλείων που υποστηρίζουν κάθε χαρακτηριστικό της ακόμη και σε πραγματικό χρόνο. Αλλά, η πραγματική επανάσταση των Big Data δεν σταματά στις δυνατότητες ανάλυσης ή αποθήκευσης. Τα εργαλεία των Big Data είναι πλέον σε θέση να προβλέψουν συμπεριφορές ή γεγονότα και ακόμη και να αντιδράσουν ανάλογα [127]. Όλο αυτό το σκηνικό έχει ωθήσει τους οργανισμούς να ξανασκεφτούν τον τρόπο δημιουργίας αξίας υιοθετώντας ένα παράγωγο της αλυσίδας αξίας δεδομένων, που δεν είναι άλλο από την αλυσίδα αξίας των Big Data. Μια τέτοια αλυσίδα αξίας είναι καλά ευθυγραμμισμένη με το οικοσύστημα των Big Data και επιτρέπει την εξαγωγή κρυμμένων αξιόπιστων πληροφοριών, βασιζόμενη στα χαρακτηριστικά των Big Data [124].

Αν και ένα σύστημα Big Data είναι συχνά πολύπλοκο στο σχεδιασμό, παρέχει λειτουργίες και διεπαφές για την αντιμετώπιση των διαφορετικών φάσεων του κύκλου ζωής των ψηφιακών δεδομένων, από τη δημιουργία τους μέχρι την καταστροφή τους. Ως εκ τούτου, για την εξαγωγή αξίας, απαιτείται η υιοθέτηση μιας προσέγγισης που θα αναλύει ένα τυπικό σύστημα Big Data σε πολλές φάσεις. Κάτι τέτοιο δημιουργεί την αλυσίδα αξίας των Big Data [128]. Αν και στη βιβλιογραφία έχουν παρουσιαστεί πολλές μελέτες που ασχολούνται με την αλυσίδα αξίας των Big Data, στις οποίες η αλυσίδα αυτή περιλαμβάνει διαφορετικό αριθμό διακριτών φάσεων, στο πλαίσιο της παρούσας εργασίας θα παρουσιαστεί η πιο απλή μορφή της, η οποία αποτελείται από τέσσερις διακριτές φάσεις [129]: παραγωγή δεδομένων (Generation), ανάκτηση δεδομένων (Acquisition), αποθήκευση δεδομένων (Storage) και ανάλυση δεδομένων (Analytics). Η παραγωγή και η ανάκτηση των δεδομένων καλύπτουν τις φάσεις εκμετάλλευσης των ακατέργαστων δεδομένων, ενώ η φάση της αποθήκευσης χρησιμοποιείται για τη διατήρησή τους μέχρι να χρησιμοποιηθούν στη φάση ανάλυσης, όπου τα δεδομένα μετατρέπονται σε αξία. Στην εικόνα 3.2 παρουσιάζεται η πορεία της αλυσίδας αξίας των Big Data σε σχέση με το χρόνο, συμπεριλαμβανομένων των τεχνολογικών εξελίξεων που εμπλέκονται.



Εικόνα 3.3: Αλυσίδα αξίας των Big Data [129]

Η χρήση των αλυσίδων αξίας Big Data δημιουργεί περισσότερα οφέλη από ότι οι παραδοσιακές αλυσίδες αξίας, καθώς μετά τη δημιουργία τους, τα δεδομένα μπορούν να επεξεργαστούν πολλαπλές φορές μέσω τεχνικών εξόρυξης και για διαφορετικές ανάγκες. Τα ίδια δεδομένα μπορούν να χρησιμοποιηθούν και σε άλλες περιπτώσεις με διαφορετικό τρόπο ή να διαμορφωθούν κατάλληλα και να αξιοποιηθούν ακόμη περισσότερες φορές, δημιουργώντας ακόμη περισσότερα οφέλη [122].

Στις επόμενες ενότητες του κεφαλαίου θα παρουσιαστεί μια ανάλυση αυτών των τεσσάρων φάσεων της αλυσίδας αξίας των Big Data. Για κάθε φάση, θα παρουσιαστεί το γενικό υπόβαθρο, οι τεχνικές προκλήσεις και οι τελευταίες εξελίξεις με βάση τις μελέτες που έχουν εμφανιστεί τα τελευταία χρόνια στη βιβλιογραφία.

3.4 Παραγωγή δεδομένων

Το σκεπτικό ότι τα δεδομένα αποτελούν ακατέργαστο υλικό, οδήγησε στην επιτάχυνση των διαδικασιών παραγωγής τους, που με τη σειρά τους αποτέλεσαν τη δημιουργία της τεχνολογίας των Big Data. Σύμφωνα με τον O. Kiranlar (2017), ο διαχωρισμός της εξέλιξης των διαδικασιών παραγωγής δεδομένων σε τρία στάδια, μπορεί να βοηθήσει στην καλύτερη κατανόηση αυτής της πρώτης φάσης της αλυσίδας αξίας των Big Data. Το πρώτο στάδιο ξεκίνησε την δεκαετία του 1990, με την υιοθέτηση των ψηφιακών

τεχνολογιών και των βάσεων δεδομένων από τις επιχειρήσεις και τους οργανισμούς. Στο στάδιο αυτό, οι εταιρείες άρχισαν να αποθηκεύουν μεγάλο όγκο δεδομένων, όπως τραπεζικές και εμπορικές συναλλαγές, αρχεία δημόσιων υπηρεσιών, κλπ.. Καθώς αυτά τα σύνολα δεδομένων ήταν δομημένα και αρκετά μικρά, μπορούσαν εύκολα να διαχειριστούν και να αποθηκευτούν στα συστήματα αποθήκευσης των επιχειρήσεων. Το δεύτερο στάδιο ξεκίνησε με την εισαγωγή των συστημάτων Web 1.0, στα τέλη της δεκαετίας του 1990. Το Web 1.0 αναφέρεται σε μηχανές αναζήτησης και επιχειρήσεις ηλεκτρονικού εμπορίου που μπορούσαν να δημιουργήσουν μια αξιοσημείωτη ποσότητα ημιδομημένων και μη δομημένων δεδομένων (ιστοσελίδες, αρχεία καταγραφής συναλλαγών, κλπ.). Στις αρχές της δεκαετίας του 2000, η εισαγωγή της τεχνολογίας Web 2.0 επέτρεψε την παραγωγή δεδομένων και από τους χρήστες. Πηγές παραγωγής δεδομένων στο web 2.0 αποτέλεσαν τα φόρουμ, τα ιστολόγια, καθώς και οι πρώτες ιστοσελίδες των μέσων κοινωνικής δικτύωσης. Το υπάρχον, τρίτο στάδιο εξέλιξης των διαδικασιών παραγωγής δεδομένων, ξεκίνησε με την εισαγωγή συσκευών, όπως οι φορητές συσκευές, τα smartphone και οι συσκευές με δυνατότητα σύνδεσης στο Διαδίκτυο που περιλαμβάνουν μία σειρά από αισθητήρες. Η ένταξη των συσκευών αυτών στην κατηγορία των πηγών παραγωγής δεδομένων, οδήγησε στην ραγδαία καθημερινή παραγωγή τεραστίων ποσοτήτων δεδομένων [129].

Στην σύγχρονη ψηφιακή εποχή, κύριες πηγές παραγωγής δεδομένων αποτελούν οι δραστηριότητες των επιχειρήσεων και των οργανισμών (εταιρικά δεδομένα), το IoT (δεδομένα IoT), οι έξυπνες συσκευές και τα μέσα κοινωνικής δικτύωσης (δεδομένα Διαδικτύου), όπως επίσης και ο τομέας της επιστημονικής έρευνας (επιστημονικά δεδομένα) [7]. Οι δραστηριότητες των επιχειρήσεων και των οργανισμών παράγουν δεδομένα που αφορούν τις διαδικτυακές συναλλαγές, την παραγωγή, τα αποθέματα, τις πωλήσεις, τη χρηματοδότηση, κλπ. [130]. Σύμφωνα με στατιστικά στοιχεία του Ιουλίου του 2021, της ιρλανδικής εταιρείας Waterford Technologies, η οποία ειδικεύεται στη σωστή διαχείριση των δεδομένων που παράγονται από τις επιχειρήσεις, αυτή τη στιγμή το μεγαλύτερο μερίδιο των πηγών παραγωγής δεδομένων ανήκει στα εταιρικά δεδομένα, τα οποία εκτιμάται ότι διπλασιάζονται κάθε 1,2 χρόνια [131].

Τα δεδομένα IoT εκτιμάται ότι θα αποτελέσουν τη μεγαλύτερη πηγή παραγωγής Big Data στο μέλλον. Σύμφωνα με πρόσφατη έρευνα της Statista (Οκτώβριος 2021), ο αριθμός των διασυνδεδεμένων έξυπνων συσκευών μέχρι το τέλος του 2021, ανήλθε σε σχεδόν 10 δισεκατομμύρια. Σύμφωνα με την ίδια μελέτη, ο αριθμός αυτός αναμένεται

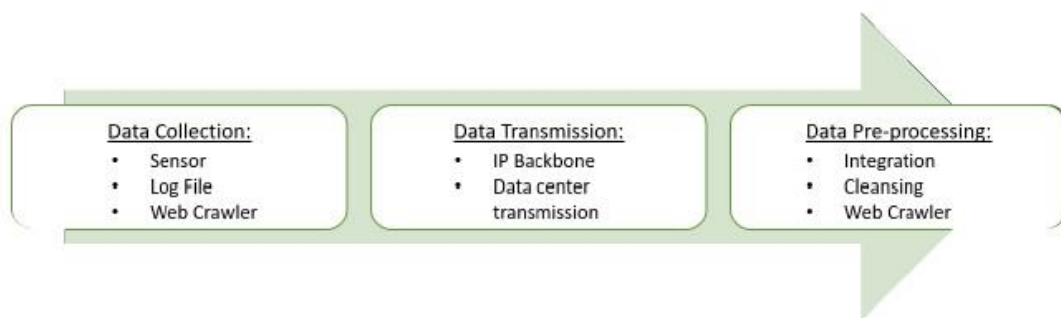
να φτάσει τις 25,45 δισεκατομμύρια συσκευές το 2030 [132]. Τα δεδομένα IoT είναι μεγάλης κλίμακας, συλλέγονται από καταναμημένες συσκευές IoT και παρουσιάζουν ετερογένεια, καθώς η μορφή τους διαφέρει ανάλογα με την εφαρμογή. Για παράδειγμα, οι πληροφορίες τοποθεσίας είναι δομημένες, όταν η ροή βίντεο δεν είναι δομημένη. Τα δεδομένα που συλλέγονται από τις καταναμημένες, σε συγκεκριμένες γεωγραφικές θέσεις, συσκευές IoT, περιλαμβάνουν πληροφορίες σχετικά με τον χρόνο και τον χώρο, κάτι που διευκολύνει την μετέπειτα ανάλυση τους. Παρόλα αυτά, μόνο ένα μικρό μέρος τους μπορεί να περιέχει αξία. Για παράδειγμα, μια κάμερα ασφαλείας που λειτουργεί 24 ώρες το 24ωρο και 7 ημέρες τη εβδομάδα καταγράφει πολύτιμες πληροφορίες μόνο σε περιπτώσεις έκτακτης ανάγκης [133].

Στην κατηγορία των δεδομένων του Διαδικτύου περιλαμβάνονται τα δεδομένα που παράγονται τόσο από το Διαδίκτυο όσο και από τα ασύρματα κυψελοειδή δίκτυα. Οι τεχνολογικές εξελίξεις στον τομέα του Διαδικτύου αλλά και των γενιών των ασύρματων κυψελοειδών δικτύων έχει οδηγήσει στην ταχύτερη παραγωγή τέτοιου είδους δεδομένων, η ανάλυση των οποίων απαιτεί προηγμένες τεχνικές επεξεργασίας [129].

Τέλος, τα επιστημονικά δεδομένα παράγονται σε προηγμένα επιστημονικά πειράματα ή παρατηρήσεις, όπως ο μεγάλος επιταχυντής αδρονίων στο CERN, η αστρονομική έρευνα Sloane Digital Sky Survey (SDSS) και τα διάφορα πειράματα που πραγματοποιούνται στον τομέα της υπολογιστικής βιολογίας [134].

3.5 Ανάκτηση δεδομένων

Το δεύτερο μέρος της αλυσίδας αξίας των Big Data αφορά την απόκτηση δεδομένων και περιλαμβάνει τρεις διαδικασίες (Εικ. 3.3) [129]: (α) τη συλλογή δεδομένων (data collection), (β) τη μετάδοση δεδομένων (data transmission) και (γ) την προεπεξεργασία δεδομένων (data pre-processing). Αρχικά, τα δεδομένα, οποιασδήποτε δομής, συλλέγονται από το περιβάλλον παραγωγής τους. Στη συνέχεια, τα ανακτημένα δεδομένα μεταφέρονται προς την υποδομή αποθήκευσής τους. Τέλος, η εφαρμογή τεχνικών προεπεξεργασίας αποσκοπεί στην απαλοιφή του θορύβου και των δεδομένων που πλεονάζουν.



Εικόνα 3.5: Διαδικασίες ανάκτησης δεδομένων [129]

3.5.1 Συλλογή δεδομένων

Οι μέθοδοι συλλογής δεδομένων ποικίλλουν ανάλογα με τα χαρακτηριστικά των πηγών παραγωγής τους και με τους εκάστοτε σκοπούς της μετέπειτα ανάλυσής τους. Στις εφαρμογές των Big Data χρησιμοποιούνται τρία βασικά είδη μέσων συλλογής δεδομένων [37]: τα αρχεία καταγραφής, οι αισθητήρες και οι ανιχνευτές Ιστού. Στον Πίνακα 3.1 παρουσιάζεται μια σύγκριση των χαρακτηριστικών αυτών των μέσων συλλογής δεδομένων.

Πίνακας 3.5.1: Σύγκριση των χαρακτηριστικών των βασικών μέσων συλλογής δεδομένων

Μέσο συλλογής	Δομή δεδομένων	Όγκος δεδομένων	Πολυπλοκότητα συλλογής	Εφαρμογές
Αρχεία καταγραφής	Δομημένα - Ημιδομημένα	Μικρός	Καμία	Αρχεία καταγραφής Ιστού Ροές κλικ
Αισθητήρες	Δομημένα - Ημιδομημένα	Μέτριος	Εξαιρετικά μεγάλη	Συστήματα παρακολούθησης Διαχείριση αποθεμάτων
Ανιχνευτές Ιστού	Διάφορα είδη	Μεγάλος	Μεσαία	Αναζητήσεις Ιστού Ανάλυση ιστότοπων κοινωνικής δικτύωσης

Τα αρχεία καταγραφής (log files) αποτελούν ένα από τα πιο κοινά μέσα συλλογής δεδομένων, καθώς δημιουργούνται αυτόματα και χρησιμοποιούνται σε όλες σχεδόν τις εφαρμογές που εκτελούνται σε συσκευές ψηφιακής τεχνολογίας [2]. Για παράδειγμα, οι δραστηριότητες των χρηστών, όπως τα κλικ και οι επισκέψεις σε διακομιστές ιστού,

καταγράφονται ως αρχεία καταγραφής [135]. Οι τρεις βασικότεροι τύποι αρχείων καταγραφής αφορούν τη δημόσια μορφή αρχείων καταγραφής (NCSA), τη διευρυμένη μορφή αρχείων καταγραφής (W3C) και τη μορφή αρχείων καταγραφής IIS (Microsoft) [37].

Οι αισθητήρες (sensors) είναι στοιχεία που μετρούν φυσικά μεγέθη και τα μετατρέπουν σε ψηφιακά σήματα. Τα είδη δεδομένων που συλλέγονται από τους αισθητήρες αφορούν μεγέθη όπως ήχο, δόνηση, ηλεκτρικό ρεύμα, πίεση, θερμότητα, κλπ. Η συλλογή των δεδομένων μέσω αισθητήρων μπορεί να εφαρμοστεί σε ενσύρματα ή ασύρματα συστήματα ανάλογα με την εφαρμογή [136].

Τέλος, οι ανιχνευτές Ιστού (Web crawlers) αποτελούν λογισμικά που χρησιμοποιούνται για τη λήψη και την αποθήκευση των ιστοσελίδων που επισκέπτονται οι χρήστες, από τις μηχανές αναζήτησης [137].

3.5.2 Μετάδοση Δεδομένων

Η μετάδοση δεδομένων αφορά τη διαδικασία που πραγματοποιείται για τη μεταφορά των ακατέργαστων δεδομένων προς την υποδομή αποθήκευσής τους, όπως συστήματα αποθήκευσης αντικειμένων, κέντρα δεδομένων και καταναμημένους χώρους αποθήκευσης στο cloud, για περαιτέρω ανάλυση. Η συγκεκριμένη διαδικασία χωρίζεται σε 2 φάσεις [138]: (α) τη μετάδοση μέσω IP backbone και (β) τη μετάδοση εντός του κέντρου δεδομένων. Στην πρώτη φάση μετάδοσης, τα δεδομένα μεταφέρονται από τις καταναμημένες πηγές προς το κέντρο δεδομένων μέσω της γραμμής κορμού του δικτύου. Ο ρυθμός μεταφοράς των δεδομένων σε αυτήν τη φάση εξαρτάται από τα φυσικά μέσα που χρησιμοποιούνται, αλλά και τη διαχείριση των μεθόδων σύνδεσης. Μετά τη μεταφορά τους στο κέντρο δεδομένων, τα δεδομένα μεταφέρονται εντός του κέντρου δεδομένων για περαιτέρω ανάλυση και επεξεργασία. Τα χαρακτηριστικά της συγκεκριμένης μετάδοσης εξαρτώνται από τον αρχιτεκτονικό σχεδιασμό και τα πρωτόκολλα του εκάστοτε κέντρου δεδομένων [138].

3.5.3 Προεπεξεργασία δεδομένων

Η διαδικασία της προεπεξεργασίας των δεδομένων είναι απαραίτητη, καθώς τα δεδομένα που έχουν συλλεχθεί εμπεριέχουν θόρυβο, μπορεί να περιλαμβάνουν διπλότυπα ή να χαρακτηρίζονται από ασυνέπεια, όπως αναφέρθηκε στο προηγούμενο κεφάλαιο. Η συγκεκριμένη διαδικασία είναι επίσης απαραίτητη καθώς μειώνει τον

χρησιμοποιούμενο χώρο αποθήκευσης, αλλά και αυξάνει την αποτελεσματικότητα της ανάλυσης των δεδομένων. Οι πιο κοινές τεχνικές που χρησιμοποιούνται στη διαδικασία της προεπεξεργασίας των δεδομένων αφορούν την ομαδοποίηση, την εκκαθάριση και την εξάλειψη των δεδομένων που πλεονάζουν [139].

Η τεχνική της ομαδοποίησης των δεδομένων (data integration) αφορά την ομαδοποίηση των συνόλων δεδομένων που έχουν συλλεχθεί από διαφορετικές πηγές, με σκοπό την παρουσίασή τους στους χρήστες με ενοποιημένη μορφή. Οι κυριότερες μέθοδοι ομαδοποίησης των δεδομένων αφορούν: (α) την ομαδοποίηση των δεδομένων που έχουν συλλεχθεί από διαφορετικές πηγές στην ίδια φυσική βάση δεδομένων, ανάλογα του περιεχομένου τους (data warehousing) και (β) την ενοποίηση των δεδομένων σε ομάδες (ομοσπονδίες) με βάση σχετικά ερωτήματα που αφορούν την προέλευση των δεδομένων (data federation) [140].

Η τεχνική της εκκαθάρισης των δεδομένων (data cleansing) αφορά τη διαδικασία που εφαρμόζεται για τον εντοπισμό δεδομένων που παρουσιάζουν ανακρίβειες, ελλείψεις ή που δεν βγάζουν νόημα. Μετά τον εντοπισμό τους, τα δεδομένα αυτά μπορούν να διαγραφούν. Σκοπός της χρήσης της συγκεκριμένης τεχνικής είναι η αύξηση της ποιότητας των δεδομένων που έχουν συλλεχθεί. Σε γενικές γραμμές, η εκκαθάριση των δεδομένων περιλαμβάνει τέσσερα στάδια [141]: (α) τον προσδιορισμό των τύπων των σφαλμάτων, (β) τη διόρθωση των σφαλμάτων, (γ) την τεκμηρίωση των τύπων των σφαλμάτων και (δ) την τροποποίηση των διαδικασιών εισαγωγής των δεδομένων για βελτίωση των μελλοντικών εφαρμογών. Παρόλο που η τεχνική της εκκαθάρισης των δεδομένων δημιουργεί τις ιδανικές συνθήκες καλύτερης μετέπειτα ανάλυσής τους, απαιτεί επιπλέον υπολογιστικούς πόρους, οι οποίοι μπορεί να χρειάζονται στη μετέπειτα φάση της ανάλυσης των δεδομένων. Ως εκ τούτου, θα πρέπει να δημιουργείται μια υπολογιστική ισορροπία μεταξύ αυτών των δύο σταδίων [142].

Η τεχνική της εξάλειψης των δεδομένων που πλεονάζουν (redundancy elimination) είναι ζωτικής σημασίας για τη μείωση της υπολογιστικής επιβάρυνσης των διαδικασιών μεταφοράς και αποθήκευσης. Στη βιβλιογραφία, και με βάση τις εξελίξεις στον τομέα της πληροφορικής, έχουν προταθεί κατά καιρούς διάφορες μέθοδοι εξάλειψης των δεδομένων που πλεονάζουν, όπως η ανίχνευση του πλεονασμού, το φιλτράρισμα των δεδομένων, η συμπίεση των δεδομένων, κλπ. [143]. Όπως συμβαίνει και στην περίπτωση της τεχνικής εκκαθάρισης των δεδομένων, κατά την

πραγματοποίηση της τεχνικής εξάλειψης των δεδομένων που πλεονάζουν, θα πρέπει να λαμβάνεται υπόψη η εξισορρόπηση της χρήσης των υπολογιστικών πόρων, καθώς αυτές οι μέθοδοι απαιτούν επιπλέον υπολογιστικούς πόρους [142].

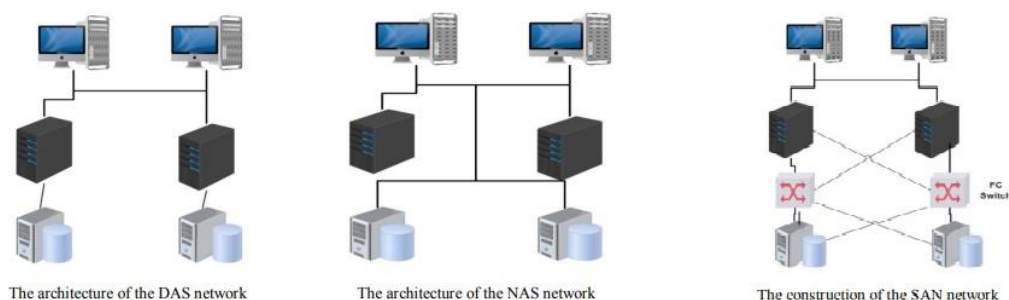
3.6 Αποθήκευση δεδομένων

Το τρίτο μέρος της αλυσίδας αξίας των Big Data αφορά την αποθήκευση των δεδομένων, που αποσκοπεί στην οργάνωσή τους σε μορφή, η οποία να είναι κατάλληλη για ανάλυση και εξαγωγή αξίας. Ένα σύστημα αποθήκευσης δεδομένων θα πρέπει να παρουσιάζει δύο βασικά χαρακτηριστικά. Πρώτον, η υποδομή του συστήματος θα πρέπει να είναι σε θέση να ενισχύει την αξιοπιστία των δεδομένων και δεύτερον, το σύστημα θα πρέπει να περιλαμβάνει μια επεκτάσιμη διεπαφή πρόσβασης που να υποστηρίζει τον καθορισμό ερωτημάτων και την ανάλυση των δεδομένων. Με βάση αυτά τα δύο χαρακτηριστικά, ένα σύστημα αποθήκευσης δεδομένων στο πλαίσιο των Big Data μπορεί να χωριστεί σε δύο μεγάλα μέρη, από τα οποία, το πρώτο ασχολείται με τη διαδικασία αποθήκευσης των δεδομένων, ενώ το δεύτερο με τη διαχείρισή τους [144].

3.6.1 Υποδομή αποθήκευσης

Η μελέτη της υποδομής των συστημάτων αποθήκευσης μπορεί να πραγματοποιηθεί με βάση τις τεχνολογίες αποθήκευσης ή τις αρχιτεκτονικές των τεχνολογιών δικτύωσης που χρησιμοποιούν [145]. Οι πιο κοινές τεχνολογίες αποθήκευσης δεδομένων περιλαμβάνουν τις μνήμες τυχαίας πρόσβασης (RAM), τους μαγνητικούς δίσκους (π.χ. τις μονάδες σκληρού δίσκου (HDD)) και τις μνήμες αποθήκευσης Storage Class Memories (π.χ. τις μονάδες αποθήκευσης στερεάς κατάστασης (SSD)) [146]. Οι μνήμες RAM αποτελούν έναν πτητικό μηχανισμό αποθήκευσης δεδομένων, στον οποίο τα δεδομένα χάνονται σε περίπτωση διακοπής της τροφοδοσίας. Μια μονάδα σκληρού δίσκου (HDD) αποτελεί μια παραδοσιακή συσκευή αποθήκευσης που χρησιμοποιεί έναν αριθμό μηχανικών πλατό και κινούμενες κεφαλές ανάγνωσης και εγγραφής για πρόσβαση σε δεδομένα. Τα εγγεγραμμένα δεδομένα σε μια μονάδα HDD παραμένουν ακόμα και μετά την απενεργοποίηση της τροφοδοσίας της. Ωστόσο, καθώς η τεχνολογία βασίζεται στις μηχανικές κινήσεις, οι ταχύτητες εγγραφής και ανάγνωσης των δεδομένων είναι σχετικά μικρές. Μια μονάδα δίσκου στερεάς κατάστασης (SSD) αποτελεί έναν νεότερο, ταχύτερο και πιο αθόρυβο τύπο συσκευής που αποθηκεύει

δεδομένα σε άμεσα προσβάσιμα chip μνήμης. Σε σύγκριση με τις μονάδες HDD, το μειονέκτημα των SSD είναι το υψηλότερο κόστος αγοράς τους, αν και τα τελευταία χρόνια, οι τιμές αγοράς τους έχουν αρχίσει να πέφτουν σημαντικά [146].



Εικόνα 3.6.1: Βασικές αρχιτεκτονικές δικτύωσης των υποδομών των συστημάτων αποθήκευσης δεδομένων [147]

Οι βασικότερες αρχιτεκτονικές των τεχνολογιών δικτύωσης που χρησιμοποιούνται στις υποδομές των συστημάτων αποθήκευσης δεδομένων αφορούν (Εικ. 3.4) [147]: (α) την άμεση συνημμένη αποθήκευση (Direct Attached Storage - DAS), (β) τη δικτυακά συνημμένη αποθήκευση (Network Attached Storage - NAS) και (γ) τα δίκτυα περιοχής αποθήκευσης (Storage Area Networks - SAN). Η αρχιτεκτονική δομή των συστημάτων αποθήκευσης DAS αποτελεί την απλούστερη δομή συστήματος αποθήκευσης, όπου πολλές συσκευές αποθήκευσης συνδέονται απευθείας σε έναν υπολογιστή. Καθώς η επεκτασιμότητα της αρχιτεκτονικής είναι περιορισμένη, χρησιμοποιείται συνήθως σε προσωπικούς υπολογιστές και σε μικρού μεγέθους διακομιστές. Τα συστήματα NAS χρησιμοποιούν ένα δίκτυο για να παρέχουν στους χρήστες πρόσβαση και κοινή χρήση δεδομένων με τη μορφή αρχείων. Το δίκτυο είναι διαχωρισμένο από τον κύριο διακομιστή και περιλαμβάνει το δικό του λογισμικό, κάτι που δίνει μεγαλύτερη επεκτασιμότητα στην αρχιτεκτονική. Τέλος, τα δίκτυα SAN αποτελούν εξειδικευμένα δίκτυα υψηλής ταχύτητας που παρέχεται από καλώδια οπτικών ινών. Ο σχεδιασμός τους είναι τέτοιος που να υποστηρίζει τη μέγιστη δυνατή επεκτασιμότητα, ένα από τα βασικότερα χαρακτηριστικά των εφαρμογών Big Data [147].

3.6.2 Πλαίσιο διαχείρισης δεδομένων

Το πλαίσιο διαχείρισης δεδομένων στοχεύει στην παροχή της πιο βολικής μεθόδου οργάνωσης των δεδομένων για επεξεργασία και ανάλυση. Στις περισσότερες εφαρμογές των Big Data χρησιμοποιούνται τρία βασικά πλαίσια διαχείρισης

δεδομένων, όπως τα συστήματα αρχείων, οι βάσεις δεδομένων και τα μοντέλα προγραμματισμού [2]. Τα συστήματα αρχείων που απαιτούνται στα συστήματα SAN και NAS αποτελούν τη βάση των τεχνολογιών αποθήκευσης δεδομένων. Τα συστήματα αυτά διαφέρουν μεταξύ των εταιρειών και σχεδιάζονται ανάλογα με τις εκάστοτε ανάγκες. Για παράδειγμα, η Google σχεδίασε το δικό της σύστημα αρχείων, το GFS, το οποίο χρησιμοποιεί φθηνούς διακομιστές βασικών προϊόντων για να παρέχει ανοχή σφαλμάτων και υψηλή απόδοση [142].

Οι βάσεις δεδομένων αποτελούν οργανωμένες συλλογές δεδομένων και βασικό μέρος των συστημάτων αποθήκευσης δεδομένων. Δεδομένου ότι οι εφαρμογές Big Data χρησιμοποιούν διάφορους τύπους δεδομένων που προέρχονται από πολλές και διαφορετικές πηγές, οι παραδοσιακές σχεσιακές βάσεις δεδομένων δεν μπορούν να χρησιμοποιηθούν, καθώς δεν είναι σε θέση να αντιμετωπίσουν προκλήσεις, όπως η πολυπλοκότητα των δεδομένων [47]. Λόγω των ιδιαίτερων χαρακτηριστικών τους, όπως η υποστήριξη εύκολης αντιγραφής (replication), η διάθεση απλών διεπαφών API, η συνέπεια και η δυνατότητα υποστήριξης τεράστιων όγκων δεδομένων, κατέστησαν τις βάσεις δεδομένων NoSQL ως την πιο βολική επιλογή για την αποθήκευση των Big Data σε τεράστια γκάμα εφαρμογών τους [148]. Οι κυριότεροι τύποι βάσεων δεδομένων NoSQL είναι τέσσερις [149]: (α) οι βάσεις δεδομένων κλειδιού-τιμής, (β) οι βάσεις δεδομένων γράφων, (γ) οι βάσεις δεδομένων στηλών και (δ) οι βάσεις δεδομένων εγγράφων. Οι βάσεις δεδομένων κλειδιού-τιμής (key-value databases) έχουν σχεδιαστεί για αποθήκευση, ανάκτηση και διαχείριση δεδομένων, συσχετίζοντας κάθε δεδομένο με διαφορετική και μοναδική τιμή, βάσει της οποίας πραγματοποιούν την αποθήκευσή τους. Οι βάσεις δεδομένων Dynamo της Amazon¹, Voldemort του LinkedIn² και Redis της Salvatore Sanfilippo³ αποτελούν παραδείγματα τέτοιου είδους. Οι βάσεις δεδομένων γράφων (graph databases) αποθηκεύουν τα δεδομένα σε δομή γραφήματος, στο οποίο τα δεδομένα αναπαρίστανται από ακμές και κόμβους, το καθένα με τα χαρακτηριστικά και τις ιδιότητές τους. Η δημοφιλέστερη βάση δεδομένων γράφων είναι η Neo4j⁴ της ομώνυμης εταιρείας. Οι βάσεις δεδομένων στηλών αποθηκεύουν τα δεδομένα σε πλαίσιο στηλών, σε αντίθεση με τις παραδοσιακές βάσεις δεδομένων όπου οι μεταβλητές τοποθετούνται σε σειρά. Η πρώτη

1 <https://www.amazon.com/>

2 <https://www.linkedin.com/>

3 <https://redis.com/>

4 <https://neo4j.com/>

βάση δεδομένων στηλών ήταν η BigTable της Google⁵, ενώ στη συνέχεια δημιουργήθηκαν και πολλές άλλες, όπως η Cassandra του Facebook⁶, η HBase της Apache⁷ και η Hypertable⁸. Τέλος, οι βάσεις δεδομένων εγγράφων, που είναι οι πλέον συχνά χρησιμοποιούμενες σε εφαρμογές Big Data, είναι σε θέση να διαχειρίζονται πιο σύνθετες δομές δεδομένων σε σύγκριση με τα υπόλοιπα είδη. Δημοφιλέστερες βάσεις δεδομένων εγγράφων είναι οι SimpleDB της Amazon, MongoDB της 10gen και CouchDB της Couchbase [149].

Το βασικό μειονέκτημα των βάσεων δεδομένων NoSQL είναι το γεγονός ότι δεν μπορούν να υποστηρίξουν πλήρως συνδυασμό λειτουργιών και ερωτημάτων [150]. Αυτό το μειονέκτημα αντιμετωπίζεται μέσω της χρήσης των μοντέλων προγραμματισμού (programming models) που υλοποιούν την εφαρμογή λογικής στις βάσεις δεδομένων NoSQL. Βασικότερους τύπους μοντέλων προγραμματισμού αποτελούν τα μοντέλα επεξεργασίας γενικής φύσης, γράφων και ροής [151]. Τα μοντέλα επεξεργασίας γενικής φύσης (generic processing models) διευκολύνουν την προσβασιμότητα των βάσεων δεδομένων NoSQL από το μοντέλο MapReduce, ένα μοντέλο προγραμματισμού που αυτοματοποιεί τον παραλληλισμό και τη διαχείριση των υπολογιστικών πόρων με βάση τη χρήση ενός μεγάλου αριθμού υπολογιστικών στοιχείων (commodity computers). Ο συγκεκριμένος αυτοματισμός επιτυγχάνεται μέσω δύο λειτουργιών, της χαρτογράφησης (map function), που επεξεργάζεται ένα ζεύγος κλειδιού-τιμής σύμφωνα με τα δεδομένα εισόδου και δημιουργεί ενδιάμεσα ζεύγη κλειδιών-τιμών, και της μείωσης (reduce function), που μειώνει το μέγεθος των δεδομένων συγχωνεύοντας όλα τα ενδιάμεσα ζεύγη που σχετίζονται με κάποιο αντίστοιχο κλειδί. Τα μοντέλα επεξεργασίας γράφων (graph processing model) εκφράζουν τις διαδικασίες υπολογιστικής ως γράφημα με κορυφές και ακμές. Οι κορυφές περιλαμβάνουν μια τροποποιήσιμη και καθορισμένη από το χρήστη τιμή. Τα προγράμματα εκτελούνται ως μια ακολουθία επαναλήψεων που ονομάζονται supersteps μέχρι να τερματιστεί ο αλγόριθμος. Τα βασικότερα εμπορικά μοντέλα είναι τα Pregel και Graphlab, με το τελευταίο να αποτελεί μοντέλο επεξεργασίας ροής δεδομένων [151].

5 <https://www.google.com/>

6 <https://www.facebook.com/>

7 <https://www.apache.org/>

8 <https://hypertable.org/>

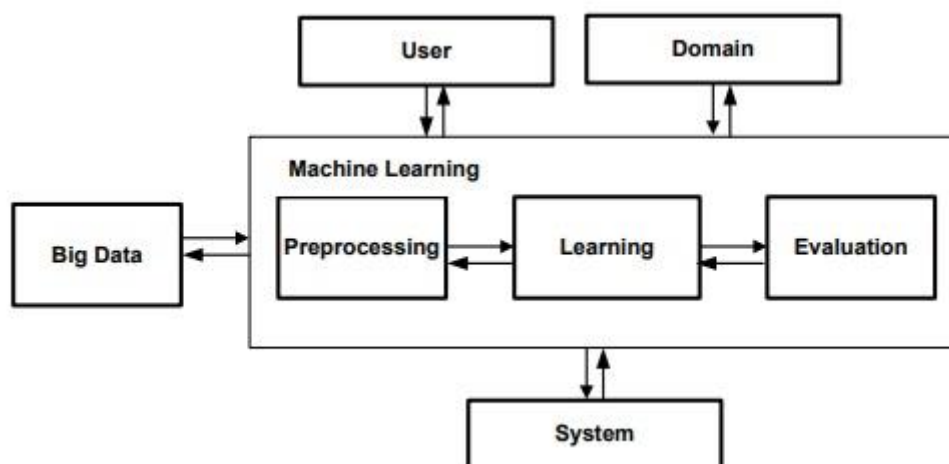
3.7 Ανάλυση Δεδομένων

Η ανάλυση των δεδομένων αποτελεί την τελευταία φάση της αλυσίδας αξίας των Big Data. Η φάση αυτή θεωρείται ως η σημαντικότερη της αλυσίδας αξίας, καθώς ένας τεράστιος αριθμός εφαρμογών των Big Data πηγάζει από την ανάλυση των δεδομένων. Σε γενικές γραμμές, σκοπός της ανάλυσης δεδομένων είναι ο έλεγχος της αξιοπιστίας τους, πολύτιμο στοιχείο για την ενίσχυση των διαδικασιών λήψης αποφάσεων και μελλοντικής πρόβλεψης καταστάσεων [152]. Με βάση τον αντικειμενικό σκοπό της ανάλυσης των δεδομένων, μπορούν να διακριθούν τρία βασικά είδη ανάλυσης [153]: (α) περιγραφική, (β) προγνωστική και (γ) καθοδηγητική ανάλυση. Η περιγραφική ανάλυση (perspective analysis) στοχεύει στον προσδιορισμό των αιτιών των περιστατικών αξιοποιώντας ιστορικά δεδομένα και χρησιμοποιείται συνήθως σε επιχειρηματικές εφαρμογές. Η προγνωστική ανάλυση (predictive analysis) αναφέρεται στην αποτύπωση μελλοντικών τάσεων και δυνατοτήτων μέσω προγνωστικής μοντελοποίησης. Τέλος, η καθοδηγητική ανάλυση (prescriptive analysis) εστιάζει στη λήψη αποφάσεων μέσω παρατήρησης ενός συστήματος [153].

Κατά τη φάση της ανάλυσης των δεδομένων χρησιμοποιούνται διάφορες τεχνικές ανάλυσης ανάλογα με τον απώτερο σκοπό της εκάστοτε εφαρμογής. Οι συνηθέστερες τεχνικές ανάλυσης των δεδομένων αφορούν την μηχανική μάθηση και την εξόρυξη δεδομένων [154].

3.7.1 Μηχανική μάθηση

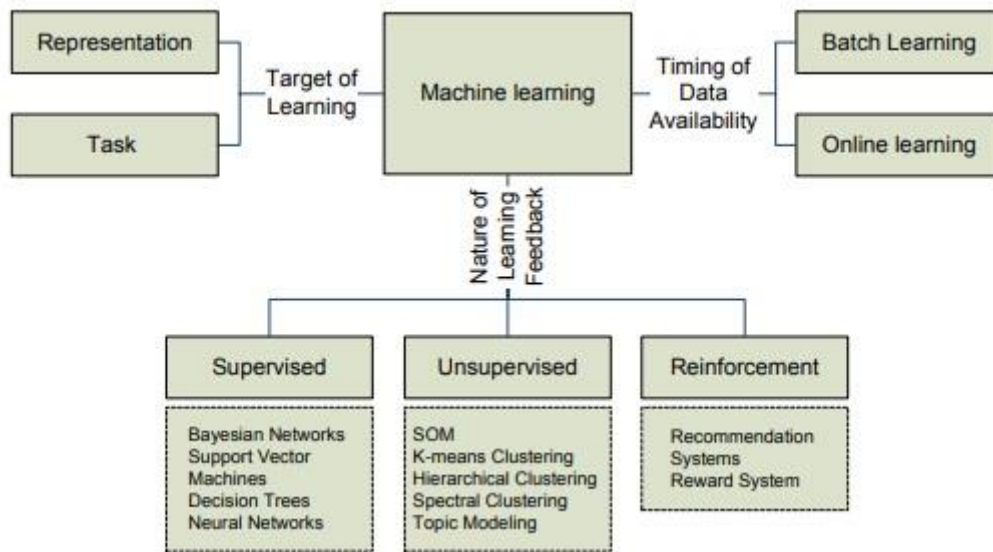
Η μηχανική μάθηση (machine learning) αποτελεί μια τεχνική που μπορεί να αποκαλύψει πολύτιμες γνώσεις μέσω της μάθησης της συμπεριφοράς από εμπειρικά δεδομένα. Στο πλαίσιο της αλυσίδας αξίας των Big Data, συνήθως ακολουθεί τις φάσεις προεπεξεργασίας, εκμάθησης και αξιολόγησης των δεδομένων (Εικ. 3.5) [55]. Η προεπεξεργασία των δεδομένων χρησιμοποιείται για την κατάλληλη διαμόρφωση των δεδομένων με σκοπό την εφαρμογή αλγορίθμων μηχανικής μάθησης, όπως η επιλογή, η μέτρηση της απόδοσης, η εκτίμηση σφαλμάτων και η στατιστική ανάλυση. Μετά την ολοκλήρωση του εκάστοτε αλγορίθμου, υπάρχει πιθανότητα εφαρμογής νέου ή η ρύθμιση των παραμέτρων του προηγούμενου, εάν η απόδοσή του δεν είναι ικανοποιητική [55].



Εικόνα 3.7.1: Χρήση της τεχνικής μηχανικής μάθησης στο πλαίσιο των Big Data [55]

Οι διάφοροι αλγόριθμοι/τεχνικές μηχανικής μάθησης μπορούν να ομαδοποιηθούν με βάση κάποιες παραμέτρους μάθησης, όπως είναι η φύση της μαθησιακής ανατροφοδότησης, ο στόχος των μαθησιακών εργασιών και ο χρονισμός της διαθεσιμότητας των δεδομένων (Εικ. 3.6) [55]. Με βάση τη φύση της μάθησης, οι τεχνικές μηχανικής μάθησης χωρίζονται σε τρεις τύπους [155]: (α) εποπτευόμενης μάθησης, (β) μη εποπτευόμενης μάθησης και (γ) ενισχυτικής μάθησης. Στην εποπτευόμενη μάθηση (supervised learning), ο αντίστοιχος αλγόριθμος εφαρμόζεται σε ζεύγη δεδομένων εισόδου και εξόδου μιας εφαρμογής, με σκοπό την αναγνώριση του μοτίβου της σχέσης μεταξύ των ζευγών αυτών. Τυπικοί αλγόριθμοι εποπτευόμενης μάθησης αποτελούν τα μπεϋζιανά δίκτυα (Bayesian networks), οι μηχανές υποστήριξης διανυσμάτων (support vector machines), τα δέντρα αποφάσεων (decision trees) και τα νευρωνικά δίκτυα (neural networks). Στην μη εποπτευόμενη μηχανική μάθηση, δεν λαμβάνεται υπόψη αυτή η σχέση μεταξύ των ζευγών εισόδου και εξόδου, αλλά η εύρεση του μοτίβου μεταξύ των δεδομένων εισόδου. Τυπικούς αλγόριθμους μη εποπτευόμενης μάθησης αποτελούν οι αυτό-οργανωμένοι χάρτες (self organising maps), οι αλγόριθμοι K-means, οι αλγόριθμοι ομαδοποίησης (ιεραρχικής ή φασματικής) (hierarchical – spectral clustering) και οι αλγόριθμοι μοντελοποίησης θεμάτων (topic modeling). Τέλος, η ενισχυτική μάθηση (reinforcement learning) αποτελεί ένα συνδυασμό εποπτευόμενης και μη εποπτευόμενης μάθησης, καθώς τροφοδοτείται από τις ανατροφοδοτήσεις των προηγούμενων εμπειριών (όπως στην εποπτευόμενη μάθηση), αλλά προσπαθεί να βρει το μοτίβο της σχέσης μεταξύ των δεδομένων εισόδου (όπως στην μη εποπτευόμενη μάθηση). Τυπικοί αλγόριθμοι

ενισχυτικής μάθησης θεωρούνται τα συστήματα συστάσεων (recommendation systems) και τα συστήματα ανταμοιβής (reward systems) [155].



Εικόνα 3.7.2: Κατηγορίες αλγορίθμων μηχανικής μάθησης [55]

Με βάση τον στόχο των μαθησιακών εργασιών, οι τεχνικές μηχανικής μάθησης μπορεί να χωριστούν σε τεχνικές αναπαράστασης ή εργασίας [55]. Ο στόχος των τεχνικών αναπαράστασης (representation learning) είναι η εκμάθηση νέων και πιο βολικών τρόπων αναπαράστασης των δεδομένων, που να διευκολύνουν την εξαγωγή πληροφοριών στα επόμενα βήματα της φάσης της ανάλυσης. Αντίθετα, οι τεχνικές εργασίας (task learning), όπως οι τεχνικές ομαδοποίησης και ταξινόμησης, προσπαθούν να φτάσουν στο επιθυμητό αποτέλεσμα του συστήματος μηχανικής μάθησης [55].

Με βάση το χρονισμό της διαθεσιμότητας των δεδομένων, οι τεχνικές μηχανικής μάθησης μπορούν να διακριθούν σε τεχνικές μάθησης δέσμης (batch learning) και σε τεχνικές απευθείας μάθησης (online learning) [156]. Οι τεχνικές μάθησης δέσμης χρησιμοποιούν ολόκληρο το σύνολο των δεδομένων για την αναγνώριση του μοτίβου, ενώ οι τεχνικές απευθείας μάθησης ενημερώνουν το μοντέλο σε κάθε νέα είσοδο δεδομένων. Οι τεχνικές απευθείας μάθησης χρησιμοποιούνται μόνο στις περιπτώσεις έλλειψης υπολογιστικών πόρων για την κάλυψη της ανάλυσης ολόκληρου του συνόλου δεδομένων ή όταν τα δεδομένα παράγονται συνεχώς [156].

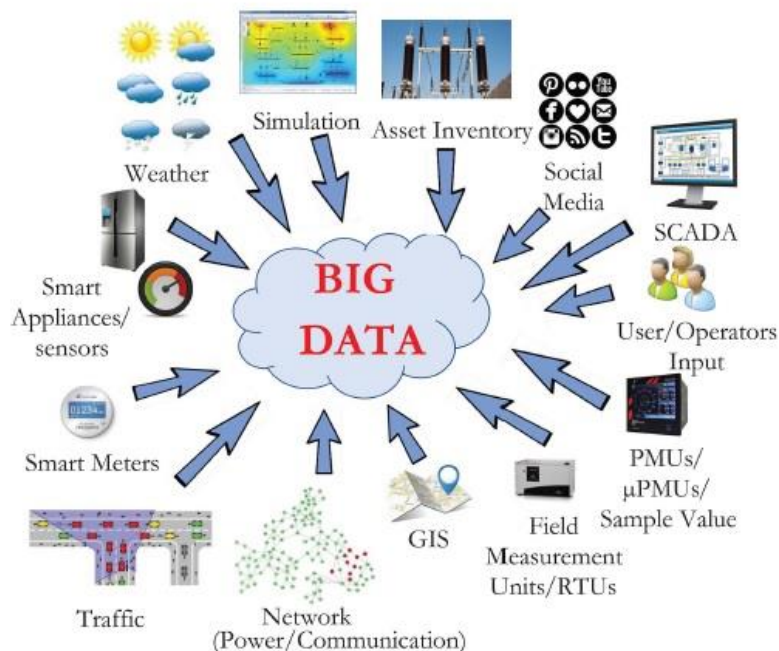
3.7.2 Εξόρυξη δεδομένων

Η εξόρυξη δεδομένων είναι μια υπολογιστική διαδικασία που χρησιμοποιείται για τον καθορισμό του μοτίβου σχέσης ενός συνόλου δεδομένων. Η διαδικασία ξεκινά με την κατανόηση του τομέα και συνεχίζει με την επιλογή, την προεπεξεργασία, την εκκαθάριση και την τροποποίησή τους, την εύρεση του μοτίβου σχέσης και την ερμηνεία του [157]. Η κατανόηση του τομέα εφαρμόζεται για να καθοριστεί ο σκοπός χρήσης των αποτελεσμάτων της διαδικασίας, κάτι που ενισχύει την καλύτερη επιλογή των σχετικών πληροφοριών από το σύνολο των δεδομένων που έχουν ήδη συλλεχθεί. Η επιλογή, η προεπεξεργασία και η εκκαθάριση των δεδομένων χρησιμοποιούνται για την κατάλληλη διαμόρφωσή τους με σκοπό την εφαρμογή των αλγορίθμων εξόρυξης. Το στάδιο της εύρεσης του μοτίβου της σχέσης μεταξύ των δεδομένων αποτελεί το βασικό μέρος της συνολικής διαδικασίας. Στο συγκεκριμένο μέρος χρησιμοποιούνται διάφοροι αλγόριθμοι μηχανικής μάθησης, στατιστικής ανάλυσης δεδομένων, νευρωνικών δικτύων, ασαφούς λογικής και αναγνώρισης μοτίβων, όπως αλγόριθμοι συσχετισμού, ταξινόμησης, ομαδοποίησης, ανάλυσης χρονοσειρών και οπτικοποίησης. Η διαδικασία της ερμηνείας χρησιμοποιείται για την επικύρωση των μοτίβων που έχουν βρεθεί στο προηγούμενο στάδιο. Η συνολική διαδικασία της εξόρυξης των δεδομένων ολοκληρώνεται με την αναφορά των αποτελεσμάτων της [157].

ΚΕΦΑΛΑΙΟ 4: BIG DATA ΣΤΟΝ ΤΟΜΕΑ ΤΗΣ ΕΝΕΡΓΕΙΑΣ

4.1 Πηγές παραγωγής δεδομένων ενέργειας

Η εφαρμογή της ψηφιακού μετασχηματισμού των συστημάτων που χρησιμοποιούνται σήμερα στον ενεργειακό τομέα έχει αυξήσει τον ρυθμό παραγωγής δεδομένων από μεγάλο πλήθος πηγών (Εικ. 4.1). Κάτι τέτοιο συνεπάγεται ότι τα Big Data αποτελούν πλέον βασική και αναπόσπαστη τεχνολογία του τομέα της ενέργειας [30].



Εικόνα 4.4: Πηγές δεδομένων ενέργειας [30]

Σύμφωνα με τους Kezunovic και συν. (2020), οι πηγές παραγωγής δεδομένων στον τομέα της ενέργειας, τα λεγόμενα δεδομένα ενέργειας (energy Big Data), μπορούν να ταξινομηθούν σε πέντε μεγάλες κατηγορίες [158]: (α) τα δεδομένα που προέρχονται από την υποδομή των δικτύων παροχής ηλεκτρικής ενέργειας, (β) τα δεδομένα που προέρχονται από τους σταθμούς ηλεκτροπαραγωγής από Ανανεώσιμες Πηγές Ενέργειας (ΑΠΕ), (γ) τα δεδομένα που προέρχονται από τους καταναλωτές, (δ) τα δεδομένα που προέρχονται από την ενεργειακή αγορά και (ε) τα δεδομένα που προέρχονται από το περιβάλλον. Η εφαρμογή λύσεων Big Data στον τομέα της ενέργειας επικεντρώνεται στην ανάλυση και επεξεργασία όλων αυτών των ετερογενών

πηγών δεδομένων, που παρουσιάζουν διαφορετικά στοιχεία ποιότητας, χωρικής ή/και χρονικής ανάλυσης και παρουσίας των πληροφοριών. Η αξιοποίηση όλων των δεδομένων που προέρχονται από αυτές τις πηγές ενέργειας, μπορεί να επιτευχθεί μέσω χρήσης προσεγγίσεων εξαγωγής γνώσης, όπως συνδυασμό των δεδομένων μέσω χρήσης της θεωρίας σύντηξης τους [159], εξαγωγή και συνδυασμό πληροφοριών διαφορετικών τύπων δεδομένων (π.χ. εικόνες, κείμενα, κλπ.) χρησιμοποιώντας τεχνικές πολυτροπικής μάθησης [160] ή κάποιο ετερογενές δίκτυο πληροφοριών [161] και συνδυασμό δεδομένων από γεωγραφικά καταναμημένες πηγές δεδομένων μέσω χρήσης κλασικών μεθόδων διανυσματικής αυτοπαλίνδρομης [162] ή βαθιάς μάθησης [163].

Οι σύγχρονες υποδομές των δικτύων παροχής ηλεκτρικής ενέργειας (grid infrastructure) περιλαμβάνουν μια σειρά από έξυπνες μονάδες και αισθητήρες που χρησιμοποιούνται για τη βελτίωση της παρακολούθησης της λειτουργίας τους. Τέτοιες μονάδες μπορεί να είναι μονάδες μέτρησης ανύσματος φάσης (Phasor Measurement Units - PMU), οι οποίες μπορούν να παρέχουν δεδομένα αναφοράς κατάστασης με πολύ υψηλό ρυθμό (π.χ. 30 μετρήσεις ανά δευτερόλεπτο μεγέθους τάσης/ρεύματος, φάσης και συχνότητας), καθώς και απομακρυσμένες τερματικές μονάδες (Remote Terminal Units - RTU), που τοποθετούνται σε υποσταθμούς, και έξυπνους μετρητές, που τοποθετούνται στους χώρους των καταναλωτών. Οι αισθητήρες και οι μονάδες RTU χρησιμοποιούνται επίσης για παρακολούθηση της κατάστασης των στοιχείων των υποσταθμών και τη βελτίωση της ποιότητας των παρεχόμενων υπηρεσιών ενέργειας [164].

Στους σταθμούς ηλεκτροπαραγωγής από ΑΠΕ εγκαθίστανται αισθητήρες παρακολούθησης σε επίπεδο ανεμογεννητριών και φωτοβολταϊκών πάνελ, που παράγουν μεγάλο όγκο δεδομένων (π.χ. μια ανεμογεννήτρια μπορεί να έχει περισσότερους από 100 αισθητήρες μέσα στον ρότορα, οι οποίοι συγκεντρώνουν περισσότερα από 10.000 δεδομένα το δευτερόλεπτο). Αυτού του είδους οι αισθητήρες μπορούν να χρησιμοποιηθούν για τοπική προγνωστική συντήρηση, η οποία μειώνει το κόστος λειτουργίας και συντήρησης των σταθμών. Η γεωγραφική κατανομή αισθητήρων παρακολούθησης, όπως ανεμογεννήτριες, εναέριες κάμερες παρακολούθησης και δορυφόροι, και ο συνδυασμός των δεδομένων που συλλέγουν, μπορούν να χρησιμοποιηθούν για την πρόγνωση των καιρικών συνθηκών, με απώτερο σκοπό την πρόβλεψη της παραγωγής του εκάστοτε είδους ΑΠΕ [165].

Η συνεχιζόμενη αύξηση του αριθμού των συσκευών IoT που χρησιμοποιούνται στη σύγχρονη εκδοχή των έξυπνων σπιτιών και κτιρίων δημιουργεί τις κατάλληλες συνθήκες παροχής υπηρεσιών ενέργειας που βασίζονται στα δεδομένα που συλλέγονται από τους ίδιους τους καταναλωτές. Προκλήσεις που αντιμετωπίζουν αυτού του είδους οι υπηρεσίες αποτελούν η προστασία του απορρήτου κατά τη συλλογή των δεδομένων, αλλά και η δέσμευση των καταναλωτών προς τους παρόχους των υπηρεσιών. Ιδιαίτερα η δεύτερη πρόκληση θα μπορούσε να αντιμετωπιστεί από τα ψηφιακά αποτυπώματα που δημιουργούν οι χρήστες στα μέσα κοινωνικής δικτύωσης, τα οποία είναι σε θέση να χρησιμοποιηθούν από τους παρόχους ηλεκτρικής ενέργειας για την καλύτερη κατανόηση των αναγκών των καταναλωτών, βάσει των οποίων θα μπορούν να προσελκύσουν πελάτες με τον βέλτιστο τρόπο [166].

Η συνεχιζόμενη αύξηση παραγωγής δεδομένων, οδήγησε στη δημιουργία της λεγόμενης αγοράς των Big Data, η οποία αναπτύχθηκε με σκοπό την αποτελεσματική εμπορία των δεδομένων που παράγονται. Σε μια τέτοια αγορά, τα δεδομένα θεωρούνται ως ένα είδος εμπορεύματος. Αντίστοιχα στην ενεργειακή αγορά, η κοινή χρήση των δεδομένων ενέργειας μεταξύ των καταναλωτών και των παρόχων αντίστοιχων υπηρεσιών μπορεί να οδηγήσει σε καλύτερη αξιοποίηση της αξίας τους [167]. Η κατάλληλη αξιοποίηση της αξίας των Big Data στον τομέα της ενέργειας μπορεί να οδηγήσει στην καλύτερη κατανόηση της ενεργειακής κατανάλωσης των χρηστών, με ανάλογη μεταβολή του κοστολογίου κατανάλωσης. Επίσης, μπορεί να αποδειχθεί πολύτιμη πληροφορία για την πρόβλεψη της χρήσης των ΑΠΕ στις μονάδες παραγωγής ηλεκτρικής ενέργειας [158].

Τα δεδομένα που προέρχονται από το περιβάλλον και τις καιρικές συνθήκες είναι υψίστης σημασίας για την πρόβλεψη των συνθηκών λειτουργίας και της πιθανότητας εμφάνισης βλαβών στα συστήματα παραγωγής ηλεκτρικής ενέργειας, ιδιαίτερα όταν αυτά περιλαμβάνουν μονάδες ΑΠΕ. Η ενσωμάτωση των ΑΠΕ στο ηλεκτρικό δίκτυο παρουσιάζει πολλές προκλήσεις όσον αφορά τη σταθερότητα, την αξιοπιστία και την αποδεκτή λειτουργία του δικτύου παραγωγής ηλεκτρικής ενέργειας. Οι ΑΠΕ, όπως η αιολική και η ηλιακή ενέργεια, παρουσιάζουν συχνά διακυμάνσεις και επομένως προβλήματα ποιότητας της παραγόμενης ηλεκτρικής ισχύος, λόγω της ποικίλης φύσης του ανέμου και του ήλιου (απότομες μεταβολές στην κατάσταση του ανέμου και της ηλιακής ακτινοβολίας). Επομένως, η δυνατότητα συλλογής, επεξεργασίας και ανάλυσης των δεδομένων που προέρχονται από πηγές, όπως επίγειους

μετεωρολογικούς σταθμούς, δορυφόρους και ραντάρ, κρίνεται ιδιαίτερα σημαντική για τον ενεργειακό τομέα [168].

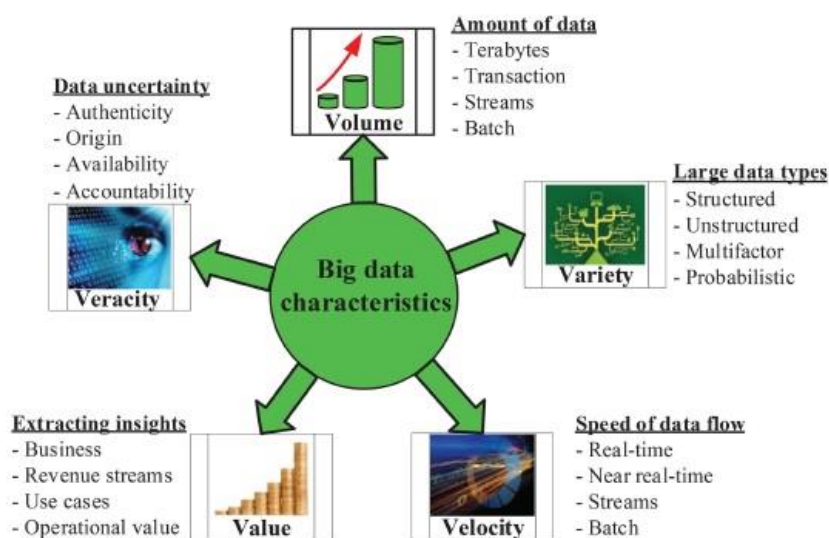
Μια τέτοια αγορά είναι τα data spaces. Η έννοια των data spaces εισήχθη στην επιστήμη των υπολογιστών πριν από περίπου 15 χρόνια ως ένας τρόπος για την ενσωμάτωση δεδομένων χωρίς τη φυσική συγκέντρωσή τους ή την απαίτηση ενός κοινού σχήματος βάσης δεδομένων. Αντίθετα, η ενοποίηση επιτυγχάνεται σε σημασιολογικό επίπεδο χρησιμοποιώντας κοινόχρηστα λεξιλόγια, επιτρέποντας πλεονασμούς δεδομένων και την συνύπαρξή τους. Ωστόσο, ο όρος έχει εξελιχθεί στην επιχειρηματική κοινότητα, όπου οι χώροι δεδομένων θεωρούνται πλέον ως μια μορφή συνεργασίας που καθοδηγείται από κοινούς στόχους και διαδικασίες λήψης αποφάσεων μεταξύ κοινοπραξιών συμμετεχόντων.

Τα data spaces βρίσκουν εφαρμογή και στον τομέα της ενέργειας. Το Energy Data Space μπορεί να βοηθήσει παρέχοντας μια αρχιτεκτονική για τη διάθεση δεδομένων. Η παροχή δεδομένων και η επικοινωνία της αγοράς στις λειτουργίες του δικτύου ηλεκτρικής ενέργειας αποτελούν βασική περίπτωση χρήσης. Άλλες πιθανές περιπτώσεις χρήσης περιλαμβάνουν την ενοποίηση δεδομένων από έξυπνους μετρητές, την προγνωστική συντήρηση και την παροχή ενέργειας για κτίρια. Το Energy Data Space μπορεί να επιτρέψει την αποτελεσματική ψηφιοποίηση των δεδομένων παρέχοντας μια αρχιτεκτονική για να καταστήσει τα δεδομένα διαθέσιμα προκειμένου να αυξηθεί η αποτελεσματικότητα στην διαχείριση κεφαλαίων και συστημάτων. Η δυναμική ανάπτυξη ως μέρος της ενεργειακής μετάβασης προς τις ανανεώσιμες πηγές δημιούργησε μεγάλη ανάγκη για δεδομένα και επικοινωνία μεταξύ παραγωγών και καταναλωτών ηλεκτρικής ενέργειας καθώς και διαχειριστών δικτύου, καθιστώντας τον ενεργειακό τομέα έναν πολλά υποσχόμενο τομέα για το Energy Data Space. Όπως και σε άλλους κλάδους της βιομηχανίας η προγνωστική συντήρηση είναι μια σημαντική τάση για αύξηση της αποτελεσματικότητας του υπάρχοντος εξοπλισμού με την ανίχνευση ανωμαλιών στη λειτουργική συμπεριφορά του και να αντιδρούν με ενέργειες συντήρησης πριν από μια αστοχία. Στον ενεργειακό τομέα, τα αολικά πάρκα αποτελούν μια καλή απεικόνιση των πλεονεκτημάτων της ανταλλαγής δεδομένων και του σχηματισμού οικοσυστημάτων δεδομένων. Το όραμα του Energy Data Space είναι ενός στοχευμένου, domain-specific και ευημερούντος οικοσυστήματος δεδομένων. Συνδυάζει τη βάση δεδομένων των επιχειρησιακών και εργοστασιακών δεδομένων με ψηφιακές πλατφόρμες και τεχνολογίες cloud. Όλα είναι δικτυωμένα και

διαλειτουργικά: υπηρεσίες, διαφορετικοί συμμετέχοντες και στοιχεία δεδομένων βρίσκονται και χρησιμοποιούνται δυναμικά.[361]

4.2 Χαρακτηριστικά των δεδομένων ενέργειας

Τα ιδιαίτερα χαρακτηριστικά των δεδομένων ενέργειας αποτελούν πηγές προκλήσεων εφαρμογής των Big Data στον τομέα της ενέργειας, όπως αυτές που παρουσιάστηκαν σε προηγούμενο κεφάλαιο. Σύμφωνα με τους Bhattarai και συν. (2019), τα αξιοσημείωτα χαρακτηριστικά των δεδομένων ενέργειας αφορούν (Εικ. 4.2) [30]: (α) τον όγκο, (β) την αξιοπιστία, (γ) την ταχύτητα, (δ) την ποικιλία και (ε) την αξία.



Εικόνα 4.5α: Χαρακτηριστικά δεδομένων ενέργειας [30]

Ο όγκος των δεδομένων που παράγονται από τις μονάδες των δικτύων παροχής ηλεκτρικής ενέργειας αυξάνεται με εκθετικό ρυθμό. Ως εκ τούτου, οι προκλήσεις που αντιμετωπίζει η χρήση των Big Data στον ενεργειακό τομέα και αφορούν την αποθήκευση, την εξόρυξη, την επεξεργασία και την ταξινόμηση των δεδομένων, αναμένεται να αυξάνονται αντίστοιχα. Ταυτόχρονα, ο αυξανόμενος ρυθμός χρήσης έξυπνων συσκευών στην πλευρά των καταναλωτών, αλλά και η ενεργός συμμετοχή τους στις διάφορες παρεχόμενες υπηρεσίες των έξυπνων δικτύων ηλεκτρικής ενέργειας, έχει οδηγήσει στην αναγκαστική επέκταση της διαχείρισης των δεδομένων ενέργειας σε επίπεδο καταναλωτή. Ακόμη όμως και σε αυτό το επίπεδο, ο όγκος των δεδομένων που παράγεται από διάφορες συσκευές, όπως οι έξυπνοι μετρητές, τα

ηλεκτρικά οχήματα, τα αυτόματα συστήματα φόρτισής τους, κλπ., είναι της τάξης των TB (Εικ. 4.3) [30]. Επομένως, η συνολική και αποτελεσματική διαχείριση του τεράστιου όγκου των δεδομένων ενέργειας αποτελεί ολοένα και μεγαλύτερο ζήτημα για τον τομέα της ενέργειας. Λύσεις στο συγκεκριμένο ζήτημα θα μπορούσαν να αποτελέσουν η χρήση καταναμημένης και επεκτάσιμης υπολογιστικής αρχιτεκτονικής των έξυπνων δικτύων παροχής ηλεκτρικής ενέργειας ή η χρήση τεχνικών μείωσης της αναπαράστασης του συνόλου των δεδομένων, με ταυτόχρονη διατήρηση της ακεραιότητας των αρχικών δεδομένων, που μπορεί να μειώσουν σημαντικά την πολυπλοκότητά τους [169].



Εικόνα 4.2β: Γράφημα παραγωγής όγκου δεδομένων ενέργειας από τα διάφορα συστήματα και μονάδες των έξυπνων δικτύων [30]

Η αξιοπιστία των δεδομένων είναι ένα από τα καθοριστικά χαρακτηριστικά των δεδομένων ενέργειας και πηγάζει βασικά από την ελλιπή κατανόηση του λειτουργικού πλαισίου των έξυπνων δικτύων παροχής ηλεκτρικής ενέργειας. Λαμβάνοντας υπόψη ότι η ποιότητα των δεδομένων, η οποία αποδίδεται από την ακρίβεια, την πληρότητα και την συνέπειά τους, αποτελεί μια από τις μεγαλύτερες ανησυχίες των έξυπνων δικτύων παροχής ηλεκτρικής ενέργειας, η ορθή λειτουργία των διαφόρων τμημάτων τους εξαρτάται εξ ολοκλήρου από την ποιότητα των δεδομένων ενέργειας. Ωστόσο, στις πραγματικές συνθήκες τα δεδομένα ενέργειας περιλαμβάνουν σφάλματα και επομένως δεν μπορούν να είναι 100% αξιόπιστα. Κύριες αιτίες αυτής της μη αξιοπιστίας τους, και επομένως της μικρότερης ποιότητάς τους, μπορούν να θεωρηθούν οι ανακριβείς μετρήσεις και ενδείξεις των αισθητήρων, οι καθυστερήσεις επικοινωνίας, οι περιπτώσεις κυβερνοεπιθέσεων, οι βλάβες του εξοπλισμού, η μη ύπαρξη συγχρονισμού, η ύπαρξη θορύβου, κλπ. Λόγοι ύπαρξης αυτών των αιτιών

μπορεί να αποτελέσουν η γήρανση των αισθητήρων, οι κακόβουλες επιθέσεις κατά τη διάρκεια των διαδικασιών ανάκτησης και ελέγχου των δεδομένων ενέργειας, κλπ. [170]. Όπως αναφέρθηκε σε προηγούμενο κεφάλαιο, τα ζητήματα της αξιοπιστίας των δεδομένων μπορούν να αντιμετωπιστούν με χρήση καινοτόμων τεχνικών εξόρυξης και ανάλυσης δεδομένων, όπως οι τεχνικές που χρησιμοποιούνται κατά τη φάση προεπεξεργασίας τους, όπως οι τεχνικές εκκαθάρισης των δεδομένων.

Με την αυξανόμενη ανάγκη για έλεγχο και επικοινωνία σε πραγματικό χρόνο στα έξυπνα δίκτυα παροχής ηλεκτρικής ενέργειας, η ύπαρξη συγχρονισμού αποτελεί ένα πολύ βασικό μέλημα. Επί του παρόντος, οι μονάδες PMU παρέχουν συγχρονισμένα δεδομένα, με τον συγχρονισμό να επιτυγχάνεται μέσω των δορυφορικών ρολογιών. Τα συγχρονισμένα δεδομένα επιτρέπουν στους αναλυτές να διακρίνουν συσχετισμούς μεταξύ συμβάντων και δράσεων αντιμετώπισής τους, τόσο κατά την ανάλυση σε πραγματικό χρόνο όσο και για τις αναλύσεις συμβάντων του παρελθόντος, να έχουν επίγνωση του οποιουδήποτε συμβάντος σχεδόν σε πραγματικό χρόνο και να λαμβάνουν τεκμηριωμένες προγνωστικές αποφάσεις [158]. Ο προσδιορισμός της αλληλουχίας προηγούμενων γεγονότων και η επίγνωση του οποιουδήποτε συμβάντος σε πραγματικό χρόνο, αποτελούν σημαντικά στοιχεία του ελέγχου της υπάρχουσας κατάστασης του έξυπνου δικτύου παροχής ηλεκτρικής ενέργειας, αλλά χρησιμοποιούνται και για την παροχή προληπτικών ή επανορθωτικών λύσεων. Ωστόσο, η επικοινωνία, η αποθήκευση και η ανάλυση των ροών δεδομένων ενέργειας από τις περισσότερες συσκευές του συστήματος διανομής και των πελατών, είναι διαδικασίες που επί του παρόντος δεν παρουσιάζουν συγχρονισμό [171]. Επομένως, η ταχύτητα συλλογής, επεξεργασίας και χρήσης των δεδομένων ενέργειας θα πρέπει να είναι τέτοια που να μην δημιουργεί κινδύνους λήψης εσφαλμένων αποφάσεων, οι οποίες μπορούν να αποδειχτούν καταστροφικές ακόμα και για τα ίδια τα έξυπνα δίκτυα [172].

Τα δεδομένα ενέργειας αποτελούν ένα συνονθύλευμα δομημένων, ημιδομημένων και μη δομημένων δεδομένων. Στα έξυπνα ενεργειακά συστήματα, τα ημιδομημένα δεδομένα αφορούν κυρίως δεδομένα καιρού και δεδομένα υπηρεσιών Ιστού, ενώ τα μη δομημένα δεδομένα, δεδομένα συμπεριφοράς πελατών και δεδομένα ήχου και βίντεο. Με την αυξανόμενη αξιοποίηση των διαλόγων μεταξύ των χρηστών των μέσων κοινωνικής δικτύωσης ή των υπαλλήλων των τηλεφωνικών κέντρων και των

καταναλωτών ενέργειας για την υποστήριξη των συστημάτων λήψης αποφάσεων, τα δεδομένα ενέργειας αναμένεται να παρουσιάσουν μεγαλύτερη ποικιλία [25].

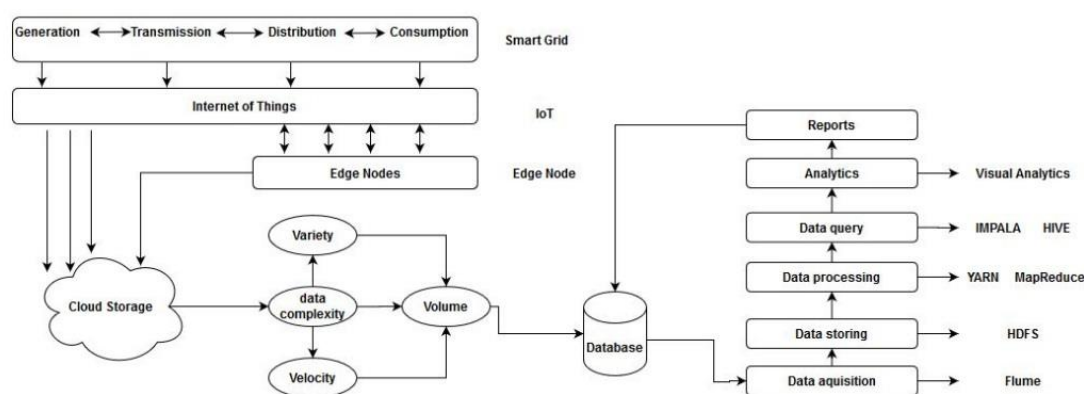
Από μόνα τους, τα δεδομένα ενέργειας δεν έχουν νόημα, παρά μόνο αν η σωστή ανάλυσή τους δημιουργήσει πολύτιμες γνώσεις που μπορούν να χρησιμοποιηθούν για την υποστήριξη της αποτελεσματικής και αποδοτικής λήψης αποφάσεων σε όλες τις διαδικασίες διαχείρισης της ενέργειας. Οι πληροφορίες που μπορούν να εμπεριέχονται στα δεδομένα ενέργειας μπορούν να χρησιμοποιηθούν επίσης για την προώθηση των δεσμεύσεων των καταναλωτών και τη βελτίωση της αποτελεσματικότητας των παρεχόμενων υπηρεσιών ενέργειας, για τη βελτίωση της αξιοπιστίας του συστήματος, για την κατανόηση των προτύπων κατανάλωσης ενέργειας και για την ανάπτυξη ανταγωνιστικών στρατηγικών μάρκετινγκ. Το μεγάλο ζήτημα της εύρεσης αξίας στα δεδομένα ενέργειας οφείλεται στο γεγονός ότι η γνώση που εξορύσσεται είναι περιορισμένη σε σύγκριση με τον όγκο των δεδομένων που αναλύεται. Επομένως, στην εποχή των Big Data, θα πρέπει να δίνεται μεγαλύτερη προσοχή στο σύνολο των δεδομένων και όχι στα μεμονωμένα δεδομένα που συλλέγονται [173].

Σύμφωνα με τους Akerkar & Hong (2021), τα δεδομένα ενέργειας, εκτός από τα παραπάνω χαρακτηριστικά (των 5V), παρουσιάζουν επίσης και τα χαρακτηριστικά των 3E [172]: (α) ενέργεια (Energy), (β) ανταλλαγή (Exchange) και (γ) ενσυναίσθηση (Empathy). Το χαρακτηριστικό της ενέργειας (δεδομένα ως ενέργεια) αφορά την εξοικονόμηση της ενέργειας που μπορεί να επιτευχθεί μέσω της ανάλυσης των Big Data. Το χαρακτηριστικό της ανταλλαγής (δεδομένα ως ανταλλαγή) αναφέρεται στο γεγονός ότι μεγαλύτερη αξιοποίηση της αξίας των δεδομένων ενέργειας μπορεί να επιτευχθεί μέσω της ανταλλαγής και της ενσωμάτωσης των Big Data των ενεργειακών συστημάτων με Big Data που προέρχονται από άλλες πηγές. Τέλος, τα χαρακτηριστικό της ενσυναίσθησης (δεδομένα ως ενσυναίσθηση) αφορά τη χρήση της ανάλυσης των δεδομένων ενέργειας για την παροχή καλύτερων υπηρεσιών ενέργειας και την μεγαλύτερη ικανοποίηση των αναγκών των καταναλωτών [172].

4.3 Τεχνολογίες υποστήριξης

Σύμφωνα με τον O. Kiranlar (2017), η εφαρμογή των Big Data στον τομέα της ενέργειας υποστηρίζεται από τρεις βασικές τεχνολογίες. Οι τεχνολογίες αυτές αφορούν τα έξυπνα δίκτυα παροχής ηλεκτρικής ενέργειας, το Διαδίκτυο των Πραγμάτων και το

υπολογιστικό νέφος [142]. Οι Ullah και συν. (2021) προσθέτουν σε αυτές και την τεχνολογία του edge computing [174]. Ουσιαστικά η συνεισφορά αυτών των τεχνολογιών αποσκοπεί στην υποστήριξη και την αξιοποίηση της αξίας των πληροφοριών που προκύπτουν από την ανάλυση των δεδομένων ενέργειας, τα οποία συλλέγονται στα διάφορα τμήματα των έξυπνων δικτύων παροχής ηλεκτρικής ενέργειας. Με τον τρόπο αυτό μπορεί να επιτευχθεί η εκπλήρωση των λειτουργικών και επιχειρησιακών απαιτήσεων του σύγχρονου ενεργειακού τομέα. Στην εικόνα 4.3 παρουσιάζεται ένα παράδειγμα συνδυασμού των παραπάνω τεχνολογιών με στόχο την εφαρμογή των Big Data στον τομέα της ενέργειας [174].

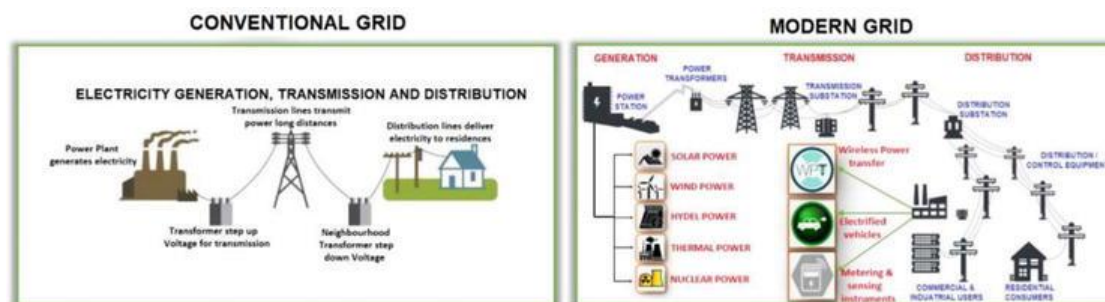


Εικόνα 4.3: Εφαρμογή των Big Data σε smart grid με χρήση των τεχνολογιών IoT, cloud computing και edge computing [174]

4.3.1 Smart Grid

Η ιδέα της τεχνολογίας των έξυπνων δικτύων παροχής ηλεκτρικής ενέργειας βασίστηκε στη βελτίωση της ευελιξίας και της αποτελεσματικότητας των παραδοσιακών ηλεκτρικών δικτύων, παρέχοντας ταυτόχρονα τη δυνατότητα ένταξης σε αυτά νέων μεθόδων παραγωγής ενέργειας, όπως η αιολική, η ηλιακή, κλπ., δηλαδή των ΑΠΕ (Εικ. 5.5). Ένα έξυπνο δίκτυο παροχής ηλεκτρικής ενέργειας είναι ουσιαστικά ένα ηλεκτρικό δίκτυο που αποτελείται από υποδομή ηλεκτρικών συστημάτων και πληροφοριακά συστήματα λογισμικού και hardware, που επιτρέπει την αμφίδρομη επικοινωνία μεταξύ όλων των τμημάτων του και των εμπλεκόμενων μερών σε αυτά, με απώτερο σκοπό την αποτελεσματική παραγωγή ενέργειας και την σωστή κατανομή και διανομή της στους καταναλωτές. Ως ένα αυτόνομο καταναμημένο σύστημα που μπορεί να παρέχει ενέργεια, η οποία μπορεί να προέρχεται ακόμα και από διαφορετικές πηγές αποθηκευμένης ενέργειας ΑΠΕ, το έξυπνο δίκτυο παροχής

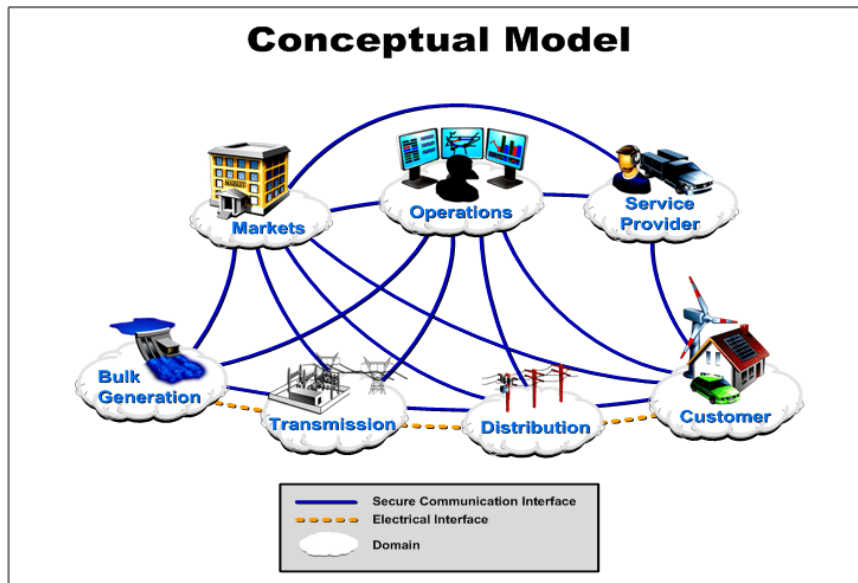
ηλεκτρικής ενέργειας επιτρέπει επίσης στους παρόχους αλλά και στους καταναλωτές να έχουν πρόσβαση σε διαδικασίες ελέγχου και διαχείρισης της ενέργειας [175].



Εικόνα 4.6α: Σύγκριση παραδοσιακού με έξυπνου δικτύου παροχής ηλεκτρικής ενέργειας [175]

Σε ένα έξυπνο δίκτυο παροχής ηλεκτρικής ενέργειας η αμφίδρομη επικοινωνία μεταξύ των πλευρών παραγωγής και κατανάλωσης πραγματοποιείται με τη βοήθεια συσκευών όπως οι έξυπνοι μετρητές, οι έξυπνες συσκευές, τα συστήματα μπαταριών αποθήκευσης ενέργειας, οι μετατροπείς ηλεκτρονικών ισχύος και άλλους αποδοτικά ενεργειακούς πόρους. Επίσης, τα συγκεκριμένα δίκτυα εφαρμόζουν τεχνολογίες που μπορούν να υποστηρίξουν τη βελτίωση του αυτοματισμού των συστημάτων τους, της ανταλλαγής πληροφοριών μεταξύ καταναλωτών, παρόχων υπηρεσιών ενέργειας, εγκαταστάσεων παραγωγής ενέργειας, κλπ., καθώς και της συνδεσιμότητας μεταξύ των τμημάτων τους [176].

Τα συμπεράσματα που προκύπτουν από την παρουσίαση των βασικών χαρακτηριστικών των έξυπνων δικτύων παροχής ηλεκτρικής ενέργειας είναι ότι η λειτουργία τους (οι διαδικασίες που πραγματοποιούνται από την παραγωγή της ηλεκτρικής ενέργειας μέχρι τη διανομή της στους καταναλωτές) είναι διαφορετική σε σύγκριση με τα παραδοσιακά ηλεκτρικά δίκτυα και ότι αυτή η λειτουργία παρουσιάζει μεγάλη πολυπλοκότητα. Οι βασικότερες λειτουργίες ενός έξυπνου δικτύου παροχής ηλεκτρικής ενέργειας αφορούν (Εικ. 4.3.1β) [177]: (α) την παραγωγή, (β) την μεταφορά, (γ) την διανομή και (δ) την κατανάλωση της ενέργειας. Στην εικόνα, εκτός από τη διαδρομή της ενέργειας παρουσιάζεται και η επικοινωνία μεταξύ όλων των τμημάτων και φορέων των έξυπνων δικτύων παροχής ηλεκτρικής ενέργειας.



Εικόνα 4.3.1β: Βασικότερες λειτουργίες ενός smart grid [178]

Η ενέργεια παράγεται από κατανεμημένες πηγές, όπως παραδοσιακούς σταθμούς ηλεκτροπαραγωγής και ΑΠΕ (π.χ. ηλιακή, αιολική, κλπ.). Στην λειτουργία της παραγωγής της ηλεκτρικής ενέργειας περιλαμβάνεται και η διαδικασία της αποθήκευσής της για σωστότερη διαχείριση της κατανάλωσής της. Η αποθήκευση αυτή αφορά την ενέργεια που παράγεται πλέον της ζήτησης και μπορεί να γίνεται στην πλευρά των σταθμών παραγωγής ή στην πλευρά του καταναλωτή. Η αποθήκευση της παραγόμενης ενέργειας αποτελεί σημαντική λειτουργία για τα έξυπνα δίκτυα παροχής ηλεκτρικής ενέργειας. Όταν γίνεται στους σταθμούς ηλεκτροπαραγωγής, η παροχή της προς τους καταναλωτές θα γίνει σε περιόδους υψηλής ζήτησης ή σε περιπτώσεις βλαβών των συστημάτων παραγωγής. Όταν γίνεται στην πλευρά των καταναλωτών, χρησιμοποιείται από αυτούς σε περιπτώσεις διακοπής λειτουργίας των τμημάτων παραγωγής, μεταφοράς ή διανομής ή όταν η τιμή του παρεχόμενου ρεύματος είναι υψηλή. Στην περίπτωση της οικιακής αποθήκευσης ενέργειας, τα τοπικά δίκτυα ενέργειας εξαρτώνται πλήρως από ΑΠΕ και παράγουν πλεόνασμα ενέργειας που μπορεί να αποθηκευτεί για τοπική κατανάλωση, όπως η φόρτιση των ηλεκτρικών οχημάτων, ή για μελλοντική χρήση [179].

Η παραγόμενη (ή η αποθηκευμένη) ενέργεια μεταφέρεται μέσω δικτύου υποσταθμών, γραμμών μεταφοράς και συστημάτων διανομής που καλύπτει μεγάλες περιοχές. Στα έξυπνα δίκτυα παροχής ηλεκτρικής ενέργειας, το τμήμα μετάδοσης της ενέργειας περιλαμβάνει τρία μέρη [176]: (α) τα έξυπνα κέντρα ελέγχου, (β) τα έξυπνα δίκτυα

μεταφοράς ηλεκτρικής ενέργειας και (γ) τους έξυπνους υποσταθμούς. Τα έξυπνα δίκτυα μεταφοράς βασίζονται στις παραδοσιακές υποδομές μεταφοράς ηλεκτρικής ενέργειας σε συνδυασμό με προηγμένες τεχνολογίες (αισθητήρες, υπολογιστές, επικοινωνία και επεξεργασία σήματος), παρέχοντας υπηρεσίες όπως η χρήση ενέργειας, η ποιότητα ισχύος, η ασφάλεια δικτύου και αξιοπιστία [176].

Η διανομή της ενέργειας γίνεται μέσω δικτύου υποσταθμών, γραμμών μεταφοράς και συστημάτων διανομής που καλύπτει μικρότερες περιοχές, έχοντας ως τελικούς αποδέκτες τους καταναλωτές. Καθώς τα έξυπνα δίκτυα παροχής ηλεκτρικής ενέργειας μεταφέρουν ενέργεια με κεντρικό αλλά και με κατανεμημένο τρόπο, τα δίκτυα διανομής παρουσιάζουν αμφίδρομη ροή ηλεκτρικής ενέργειας [177].

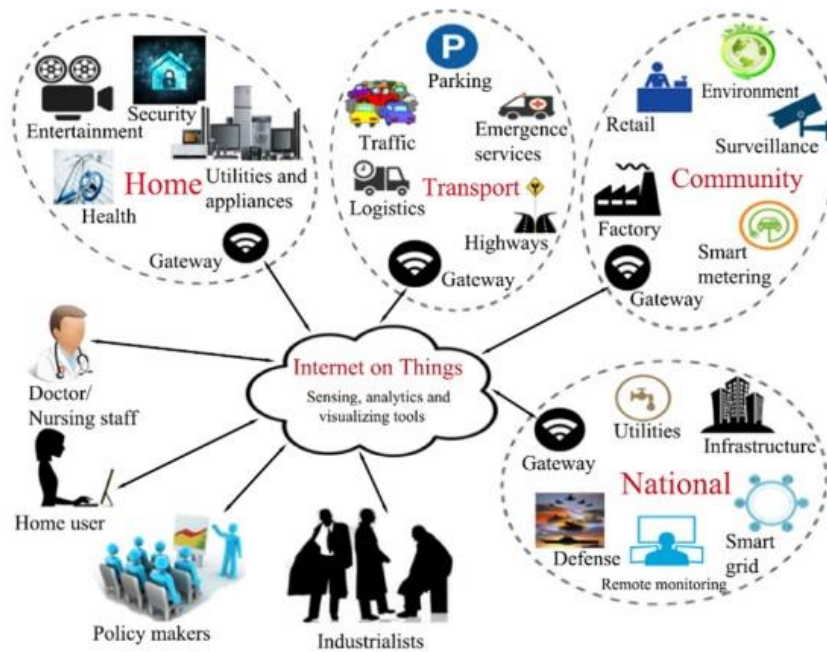
Στα έξυπνα δίκτυα παροχής ηλεκτρικής ενέργειας, η κατανάλωση ενέργειας, από τους καταναλωτές, γίνεται με ελεγχόμενο και διαχειρίσιμο τρόπο, χρησιμοποιώντας έξυπνους μετρητές, αισθητήρες σε συσκευές, έξυπνα βύσματα και πρίζες. Ο ίδιος ο καταναλωτής είναι σε θέση να παρακολουθεί, να ελέγχει και να διαχειρίζεται την ηλεκτρική ενέργεια που καταναλώνει, χρησιμοποιώντας ανάλογες εφαρμογές για κινητά τηλέφωνα ή εφαρμογές ιστότοπου. Η δυνατότητα ελέγχου και διαχείρισης της ενέργειας, που υπάρχει στα έξυπνα δίκτυα παροχής ηλεκτρικής ενέργειας, είναι σημαντική για την παρακολούθηση της χρήσης και της κατανάλωσης της, από τους καταναλωτές, αλλά και από όλους τους εμπλεκόμενους φορείς στον τομέα της ενέργειας. Τα δεδομένα που αφορούν την κατανάλωση της ενέργειας από τα έξυπνα σπίτια, τα έξυπνα κτίρια, τις έξυπνες πόλεις, κλπ., συλλέγονται, επεξεργάζονται και αναλύονται χρησιμοποιώντας εργαλεία ανάλυσης και οπτικοποίησης δεδομένων. Οι πληροφορίες που εξάγονται από αυτές τις διαδικασίες χρησιμοποιούνται από τις εταιρείες ενέργειας και τους πελάτες τους για τις τρέχουσες και τις μελλοντικές τους αποφάσεις. Για παράδειγμα, οι εταιρείες ενέργειας μπορούν να χρησιμοποιήσουν τις πληροφορίες για πρόβλεψη της συντήρησης των υποδομών τους, οι εταιρείες παροχής υπηρεσιών ενέργειας μπορούν να τις χρησιμοποιήσουν για την ανάπτυξη προγραμμάτων ζήτησης και απόκρισης, ενώ οι καταναλωτές μπορούν να τις χρησιμοποιήσουν για τη μείωση της κατανάλωσής τους σε ώρες αιχμής και γενικότερα για σωστότερη διαχείριση της τελικής τιμολόγησης της κατανάλωσής τους [178].

4.3.2 IoT

Το Διαδίκτυο των Πραγμάτων (IoT) είναι μία τεχνολογία που αποτελείται από μικρά και πολύπλοκα συστήματα και δίκτυα, τα οποία επιτρέπουν σε επιχειρήσεις, κυβερνήσεις και πολίτες να διασυνδέουν φυσικά με εικονικά αντικείμενα (πράγματα), μέσω των υπαρχουσών αλλά και των αναδύμενων διαλειτουργικών τεχνολογιών πληροφορίας και επικοινωνίας (ΤΠΕ) [174]. Η βασική φιλοσοφία ανάπτυξης της τεχνολογίας ήταν η δημιουργία ενός παγκόσμιου δικτύου μηχανών και συσκευών που να υποστηρίζει τη μεταξύ τους αλληλεπίδραση και τη μεταξύ τους ανταλλαγή πληροφοριών, με σκοπό την επίτευξη ενός κοινού στόχου. Ένα τέτοιο παγκόσμιο δίκτυο περιέχει επίσης πλατφόρμες συλλογής, ανάλυσης, αποθήκευσης και οπτικοποίησης των δεδομένων που παράγονται από τις συσκευές και τις μηχανές, χρησιμοποιώντας έξυπνες συσκευές, επικοινωνίες και εργαλεία ανάλυσης [179].

Οι εφαρμογές του IoT συνεπάγονται τη χρήση έξυπνων συσκευών (αισθητήρων και ενεργοποιητών) και συσκευών IoT, όπως φορητούς υπολογιστές, smartphone, έξυπνα gadget, έξυπνα οχήματα, κλπ, τα οποία έχουν τη δυνατότητα σύνδεσης με το Διαδίκτυο (Εικ. 4.7). Αυτό σημαίνει ότι το IoT μπορεί να θεωρηθεί και ως επέκταση του σύγχρονου Διαδικτύου και ως σημαντικό μέρος του μελλοντικού Διαδικτύου. Σε σύγκριση με το σύγχρονο Διαδίκτυο, το IoT μπορεί να υποστηρίξει [180]: (α) μια τεράστια γκάμα τερματικού εξοπλισμού, (β) αυτόματη λήψη δεδομένων και (γ) έξυπνα τερματικά.

Η εφαρμογή του IoT αποτελεί σήμερα βασικό στοιχείο για την αειφόρο χρήση των συστημάτων ΤΠΕ που πλέον χρησιμοποιούνται σε κάθε τομέα της καθημερινότητας. Καθώς, το IoT αποτελεί ένα σύνολο αλληλένδετων πόρων και διαδικασιών, ο συνδυασμός της τεχνολογίας με τα Big Data, δεν αποσκοπεί μόνο στη συλλογή, την επεξεργασία και την ανάλυση των δεδομένων, αλλά και στη δημιουργία αξίας [122]. Με βάση αυτά τα στοιχεία, στο πλαίσιο των έξυπνων δικτύων παροχής, το IoT χρησιμοποιείται ως μέσο συλλογής, σε πραγματικό χρόνο, όλων των δεδομένων ενέργειας που παράγονται από όλες τις λειτουργίες τους, δηλαδή της παραγωγής, της μεταφοράς, της διανομής και της κατανάλωσης ενέργειας, καθώς και ως τεχνολογία αξιοποίησης της αξίας που εμπεριέχεται στα δεδομένα αυτά [174].



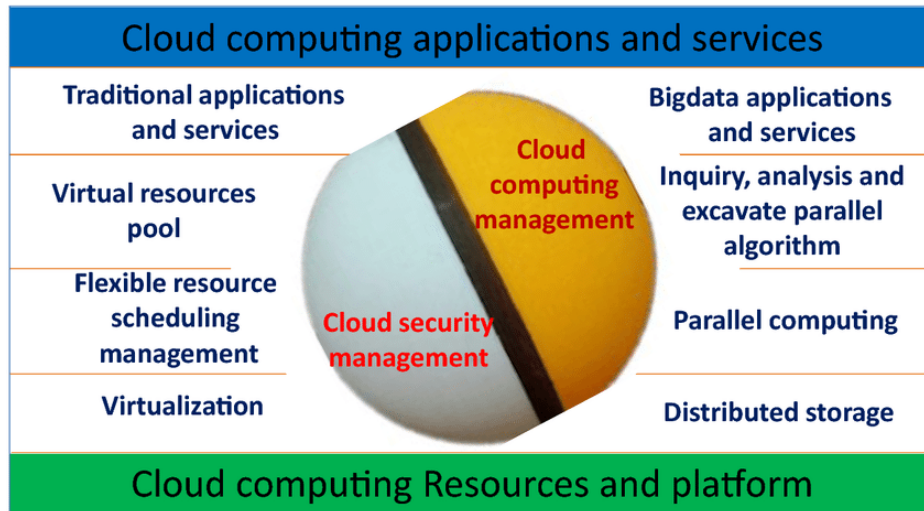
Εικόνα 4.3.3: Παραδείγματα εφαρμογών IoT [180]

4.3.3 Cloud computing

Το υπολογιστικό νέφος (cloud computing) είναι μια τεχνολογία που δημιουργήθηκε ως εξέλιξη των τεχνολογιών της κατακεντημένης υπολογιστικής, της παράλληλης υπολογιστικής και της υπολογιστικής πλέγματος, αν και από πολλούς θεωρείται και ως μια εμπορική υλοποίηση της έννοιας της επιστημονικής πληροφορικής. Κατά μία έννοια, η τεχνολογία αποτελεί μία μέθοδο μεταφοράς και αποθήκευσης δεδομένων ή παροχής υπηρεσιών, μέσω χρήσης των υποδομών των πληροφοριακών συστημάτων και της χρησιμοποίησης των απαραίτητων διαδικτυακών πόρων κατά παραγγελία ή με επεκτάσιμο τρόπο. Επίσης, μπορεί να θεωρηθεί ως μια τεχνολογία μέσω της οποίας οι χρήστες έχουν πρόσβαση σε έναν δικτυακό διακομιστή που βρίσκεται σε κάποια απομακρυσμένη τοποθεσία με σκοπό τη χρήση υπηρεσιών που παρέχονται από αυτόν [181].

Το cloud computing μπορεί να θεωρηθεί πλέον μια τεχνολογία που μπορεί να χρησιμοποιηθεί για τη διαχείριση της αποθήκευσης και της επεξεργασίας του μεγάλου όγκου δεδομένων που παράγονται από το σύνολο των πηγών που μπορούν να υπάρχουν σε ένα περιβάλλον IoT. Όπως φαίνεται στην εικόνα 4.3.3, όπου παρουσιάζονται τα βασικά στοιχεία του cloud computing, η τεχνολογία μπορεί να χρησιμοποιηθεί ως βασικό στοιχείο αποθήκευσης φυσικών αλλά και εικονικών υποδομών, στις οποίες

μπορεί να αποθηκευτεί και να αναλυθεί ο μεγάλος όγκος των δεδομένων που συλλέγονται από τις διάφορες σύγχρονες εφαρμογές [182].



Εικόνα 4.7: Βασικά στοιχεία του cloud computing [182]

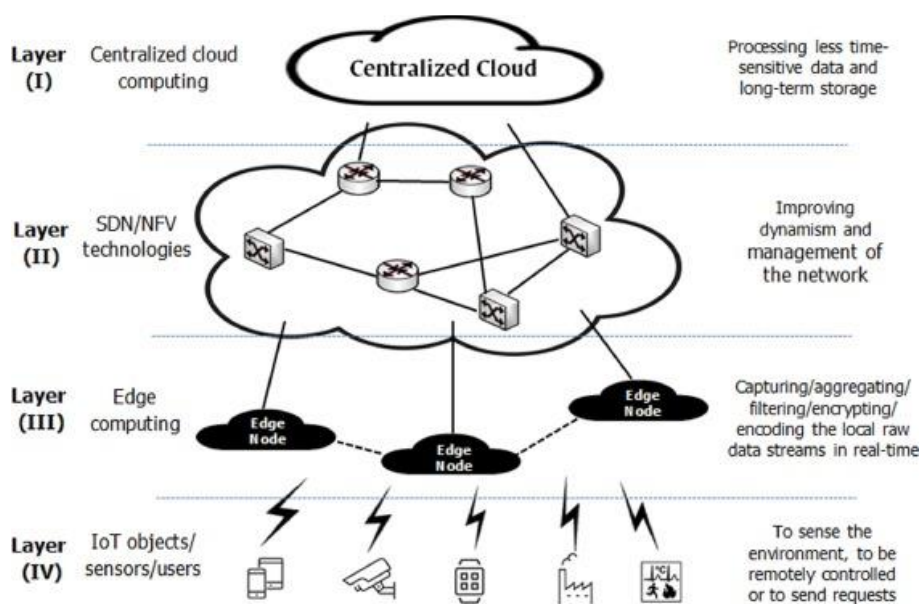
Οι υπηρεσίες που παρέχονται από το cloud computing μπορούν να περιγραφούν με βάση τα μοντέλα των υπηρεσιών και τα μοντέλα της ανάπτυξής τους. Τα μοντέλα υπηρεσιών ταξινομούνται με βάση τους τύπους των υπηρεσιών που παρέχονται από το cloud, ενώ τα μοντέλα ανάπτυξης ταξινομούνται με βάση το πώς και από ποιον χρησιμοποιούνται οι υπηρεσίες αυτές. Ως βασικότερα μοντέλα υπηρεσιών cloud θεωρούνται τα IaaS (υποδομή ως υπηρεσία), PaaS (πλατφόρμα ως υπηρεσία) και SaaS (λογισμικό ως υπηρεσία), ενώ άλλα πολύ χρησιμοποιούμενα μοντέλα υπηρεσιών cloud είναι τα DaaS (δεδομένα ως υπηρεσία), MaaS (παρακολούθηση ως υπηρεσία) κλπ. Ως βασικότερα μοντέλα ανάπτυξης υπηρεσιών cloud θεωρούνται τα δημόσια, ιδιωτικά, κοινοτικά και υβριδικά cloud, ενώ στην πράξη χρησιμοποιούνται και άλλα μοντέλα ανάπτυξης cloud, όπως τα κατανεμημένα, ενδοεταιρικά, κλπ. [183].

Ένας τέτοιος συνδυασμός μοντέλων cloud παρουσιάζει πολλά σημαντικά πλεονεκτήματα, όπως της αυτοεξυπηρέτησης με βάση τις εκάστοτε απαιτήσεις, της ευρείας πρόσβασης στο δίκτυο, της συγκέντρωσης πόρων, της ταχύτητας, της ελαστικότητας και της διαχείρισης των παρεχόμενων υπηρεσιών, καλύπτοντας έτσι τις απαιτήσεις πολλών εφαρμογών. Επομένως, η χρήση των υπηρεσιών cloud στις εφαρμογές και στα συστήματα ανάλυσης των Big Data, αποτελεί μεγάλο πλεονέκτημα, καθώς είναι σε θέση να υποστηρίξει ανάλυση δεδομένων σε πραγματικό χρόνο, καθώς και να αποτελέσει μια οικονομικά αποδοτική λύση αποθήκευσής τους [182]. Όσον

αφορά τον τομέα της ενέργειας, οι Zhou & Yang (2018) αναφέρουν ότι οι υπηρεσίες SaaS και PaaS χρησιμοποιούνται συνήθως στις εφαρμογές των Big Data στον τομέα αυτό [184].

4.3.4 Edge computing

Το edge computing αφορά μία τεχνολογία που υποστηρίζει την εκτέλεση διαδικασιών υπολογιστικής στις παρυφές ενός δικτύου. Λαμβάνοντας υπόψη ότι σε σχέση με την τεχνολογία του cloud, τα δεδομένα edge είναι δεδομένα downstream, ενώ σε σχέση με τις υπηρεσίες του IoT, είναι δεδομένα upstream, ως παρυφές ενός δικτύου μπορούν να οριστούν οποιεσδήποτε υπολογιστικοί και δικτυακοί πόροι που βρίσκονται μεταξύ των πηγών δεδομένων και του νέφους (Εικ. 4.9). Στόχος του edge computing είναι να εξασφαλίσει ότι οι διαδικασίες υπολογιστικής εκτελούνται κοντά στις πηγές δεδομένων [185].



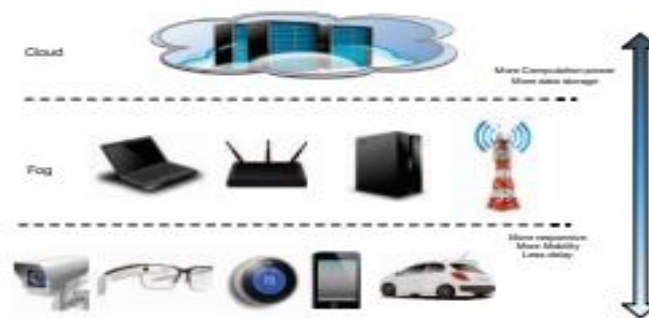
Εικόνα 4.8: Παράδειγμα πλαισίου edge computing [185]

Οι κόμβοι στις παρυφές του δικτύου εκτελούν πολλές διεργασίες όπως επεξεργασία των δεδομένων, προσωρινή αποθήκευση, διαχείριση συσκευών, κλπ., με σκοπό τη μείωση της κυκλοφορίας των δεδομένων από τις συσκευές προς το νέφος. Για να εκτελεστούν όλες αυτές οι διεργασίες, ο σχεδιασμός των παρυφών του δικτύου θα πρέπει να είναι τέτοιος που να μπορεί να αντιμετωπίσει αποτελεσματικά τις υπάρχουσες απαιτήσεις του δικτύου που αφορούν την ασφάλεια, την αξιοπιστία και τη διατήρηση του απορρήτου. Τα πλεονεκτήματα της τεχνολογίας του edge computing

αποδεικνύουν όλο και περισσότερο την αποτελεσματικότητα και την αξία της σε πολλές εφαρμογές και δείχνουν την ικανότητά της να εκπληρώσει τις απαιτήσεις αυτές, χρησιμοποιώντας τους διακομιστές που είναι φυσικά εγκατεστημένοι πιο κοντά στους τελικούς χρήστες. Αυτά τα πλεονεκτήματα καθιστούν την τεχνολογία edge ως μία από τις πλέον βέλτιστες επιλογές για ορισμένες κρίσιμες εφαρμογές που χρειάζονται τη δυνατότητα λήψης αποφάσεων για δράση σε πραγματικό χρόνο, όπως οι διεργασίες που εκτελούνται στις μονάδες των σταθμών παραγωγής ενέργειας [186].

4.3.5 Fog Computing

Το Fog Computing συνήθως συνεργάζεται με το υπολογιστικό νέφος (cloud computing) . Σαν αποτέλεσμα, οι τελικοί χρήστες μαζί με το νέφος και την ομίχλη σχηματίζουν ένα μοντέλο παροχής υπηρεσιών τριών στρωμάτων , όπως φαίνεται στην παρακάτω εικόνα [362].



Εικόνα 4.3.5: Fog computing [362]

Το Fog computing επίσης συνδέεται με το cloud computing σε όρους χαρακτηρισμών. Για παράδειγμα , ελαστικοί πόροι (υπολογισμός, αποθήκευση και δικτύωση) είναι τα δομικά στοιχεία και των δύο, υποδεικνύοντας ότι οι περισσότερες τεχνολογίες cloud computing μπορούν να εφαρμοστούν απευθείας στο fog computing. Όμως ,το fog computing έχει μερικές μοναδικές ιδιότητες που το ξεχωρίζουν από τις άλλες υπάρχουσες αρχιτεκτονικές υπολογισμού. Τα πλεονεκτήματα του Fog Computing είναι:

(α)Γρηγορότεροι χρόνοι απόκρισης και ειδοποιήσεις λόγω κατανεμημένων υπολογισμών που πραγματοποιείται πιο κοντά στις συσκευές edge αντί να εκτελούνται εξ ολοκλήρου στο cloud.

(β)Βελτιωμένη επεκτασιμότητα της λύσης καθώς ο υπολογισμός κατανέμεται σε τρία διαφορετικά επίπεδα - το Edge, το Fog και το Cloud. (γ)Μειωμένη κίνηση δικτύου και χαμηλότερη καθυστέρηση καθώς τα δεδομένα υποβάλλονται σε επεξεργασία και αναλύονται στο edge, μειώνοντας την ανάγκη μετάδοσης μεγάλου όγκου δεδομένων στο cloud. (δ)Βελτιωμένη αξιοπιστία και διαθεσιμότητα υπηρεσιών, ακόμη και όταν διακόπτεται η σύνδεση δικτύου, καθώς οι δυνατότητες επεξεργασίας και αποθήκευσης δεδομένων κατανέμονται σε πολλαπλά επίπεδα. (ε)Βελτιωμένο απόρρητο και ασφάλεια, καθώς τα ευαίσθητα δεδομένα μπορούν να υποβληθούν σε επεξεργασία και να αναλυθούν τοπικά στο edge, χωρίς να χρειάζεται να μεταδοθούν στο cloud.

Εν κατακλείδι , το fog computing μπορεί να οδηγήσει στην επεξεργασία και μείωση του όγκου δεδομένων σε πολύ πρώιμο στάδιο ,οδηγώντας σε μείωση καθυστερήσεων και εξοικονόμηση μεταφοράς δεδομένων. Το fog computing έχει ένα ευρύ φάσμα εφαρμογών,πχ σε δίκτυα έξυπνων δικτύων και ασύρματων αισθητήρων [365].

ΚΕΦΑΛΑΙΟ 5: ΕΦΑΡΜΟΓΕΣ BIG DATA ΣΤΟΝ ΤΟΜΕΑ ΤΗΣ ΕΝΕΡΓΕΙΑΣ

5.1 Ανάγκες διαχείρισης ενέργειας στα smart grid

Το smart grid θεωρείται η επόμενη γενιά ηλεκτρικών δικτύων. Χρησιμοποιεί αμφίδρομες ροές ηλεκτρισμού και πληροφοριών για την δημιουργία ενός ευρέως διανεμημένου και αυτοματοποιημένου δικτύου παροχής ενέργειας, σε αντίθεση με τα παραδοσιακά δίκτυα ηλεκτρικής ενέργειας, που χρησιμοποιούνται για την μεταφορά ηλεκτρισμού από λίγες κεντρικές γεννήτριες σε πολλούς τελικούς χρήστες/καταναλωτές. Ένα smart grid, που μπορεί να ονομάζεται και electrical/power grid, intelligrid, futuregrid, intergrid ή intragrid, είναι μια βελτίωση του ηλεκτρικού δικτύου του 20ου αιώνα. Στον κάτω πίνακα φαίνεται μια σύγκριση μεταξύ ενός υπάρχοντος δικτύου και ενός smart grid [363].

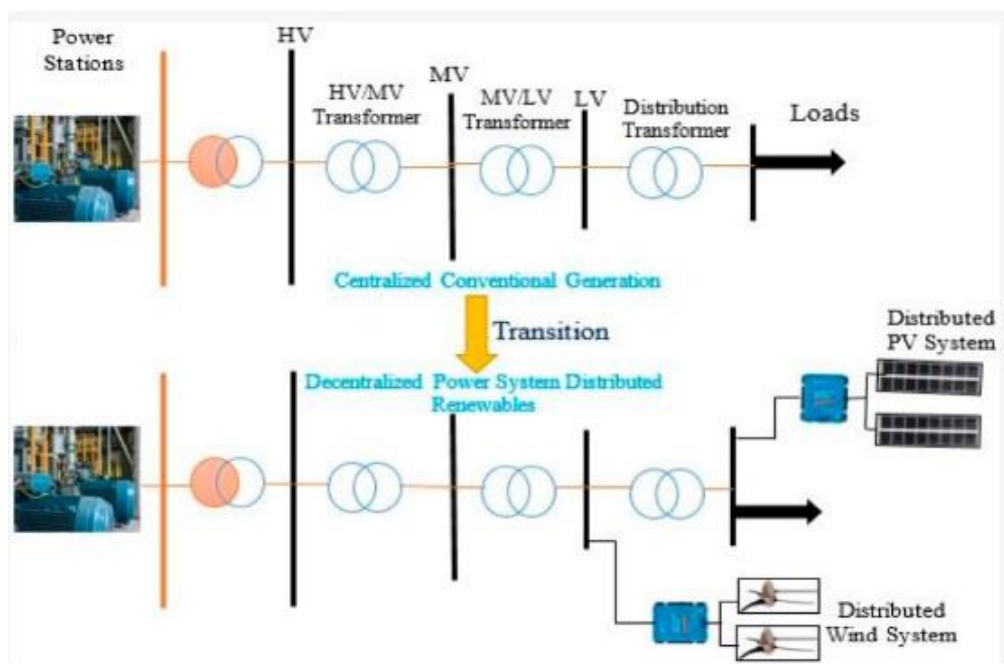
Existing Grid	Smart Grid
Electromechanical	Digital
One-way communication	Two-way communication
Centralized generation	Distributed generation
Few sensors	Sensors throughout
Manual monitoring	Self-monitoring
Manual restoration	Self-healing
Failures and blackouts	Adaptive and islanding
Limited control	Pervasive control
Few customer choices	Many customer choices

Εικόνα 5.1α: Σύγκριση υπάρχοντος δικτύου και smart grid [363]

Κάποιες από τις προκλήσεις που παρουσιάστηκαν στα smart grids είναι η μείωση κατανάλωσης ενέργειας και αύξηση παραγωγής ενέργειας από ΑΠΕ και η ενεργειακή απόδοση στον κτιριακό τομέα, στην αντιμετώπιση των οποίων συνέβαλαν έργα όπως το i-ENERGY [366], MATRYCS και BRICK [367,368,369].

Ο υφιστάμενος μετασχηματισμός του τομέα της ενέργειας αφορά αλλαγές στη σύνθεση και στη λειτουργία των ηλεκτρικών δικτύων. Μια από τις βασικότερες απαιτήσεις των σύγχρονων έξυπνων δικτύων παροχής ηλεκτρικής ενέργειας, είναι η ανά πάσα στιγμή επίτευξη μιας ισορροπίας μεταξύ της παραγωγής και της κατανάλωσης ενέργειας. Με άλλα λόγια, τα σύγχρονα έξυπνα δίκτυα παροχής ηλεκτρικής ενέργειας θα πρέπει να είναι σε θέση να διαχειρίζονται με τον βέλτιστο τρόπο την ενέργεια που παράγουν, μεταφέρουν και διανέμουν προς κατανάλωση. Η διαχείριση της ενέργειας στα παραδοσιακά ηλεκτρικά δίκτυα πραγματοποιείται μέσω προσαρμογής της παραγωγής

ενέργειας στις απαιτήσεις των καταναλωτών, οι οποίες διαπιστώνονταν μέσω της συνεχούς παρακολούθησης της καταναλισκόμενης ενέργειας. Ωστόσο, μια τέτοια μέθοδος δεν δείχνει να είναι αποτελεσματική, όσο η αρχιτεκτονική δομή των έξυπνων δικτύων παροχής ηλεκτρικής ενέργειας τείνει προς μία αποκέντρωση των λειτουργιών τους [187].



Εικόνα 5.1β: Σημεία ενσωμάτωσης των ΑΠΕ σε ένα smart grid [188]

Σύμφωνα με τον Ο. Kiranlar (2017), η ενσωμάτωση των ΑΠΕ στις κατακεκομημένες μονάδες παραγωγής ενέργειας, όπως οι σταθμοί θερμικής, ηλεκτρικής, ηλιακής, αιολικής και υδροηλεκτρικής ενέργειας (Εικ. 5.1), προκαλεί κάποιες διαταραχές στη λειτουργία του ηλεκτρικού δικτύου, και πιο συγκεκριμένα, στο δίκτυο διανομής. Οι διαταραχές αυτές αφορούν κυρίως διακυμάνσεις στην παρεχόμενη τάση, αντίστροφες ροές ισχύος και απώλειες ενέργειας [142].

Η επιλογή του κατάλληλου σημείου σύνδεσης των κατακεκομημένων μονάδων παραγωγής ενέργειας με το ηλεκτρικό δίκτυο αποτελεί σημαντικό στοιχείο της διαδικασίας ενσωμάτωσης των ΑΠΕ σε αυτό. Αν κάτι τέτοιο γίνει κοντά στα σημεία κατανάλωσης ή σε μετασχηματιστές διανομής, θα προκαλέσει διακυμάνσεις στην παρεχόμενη ισχύ, οι οποίες θα επηρεάσουν ολόκληρη την τάση του δικτύου. Οι διακυμάνσεις ισχύος σε ένα δίκτυο χαμηλής τάσης προκαλούν ανεπιθύμητη διακύμανση της τάσης, κάτι που επιδεινώνεται όταν το ποσοστό ενσωμάτωσης των ΑΠΕ είναι υψηλό, καθώς δημιουργείται απότομη άνοδος του επιπέδου τάσης, κυρίως

στις περιπτώσεις όπου η κατανάλωση είναι μικρή. Επομένως, το δίκτυο χαμηλής τάσης θα πρέπει να είναι εξοπλισμένο με αυτόματους ρυθμιστές τάσης προκειμένου να ελέγχεται το επίπεδο τάσης σε αυτό το τμήμα του δικτύου [188].

Κάθε φορά που μια κατανεμημένη μονάδα παραγωγής ενέργειας εγχέει ισχύ στο δίκτυο, παρατηρείται μεταβολή της στάθμης της τάσης, η οποία εξαρτάται σε μεγάλο βαθμό από τη ροή της ισχύος στο δίκτυο. Εάν η μονάδα βρίσκεται κοντά στα σημεία κατανάλωσης και ο συντελεστής ισχύος της είναι ανάλογος αυτού των φορτίων, τότε η μεταβολή της στάθμης της τάσης δεν θα οδηγήσει σε υπέρβαση των καθορισμένων ορίων της τιμής της [189]. Αντίθετα, αν η ηλεκτρική ενέργεια που παράγεται από την κατανεμημένη μονάδα υπερβαίνει το όριο ανοχής των φορτίων ή παρουσιάζει ακραίο συντελεστή ισχύος, θα υπάρξει αύξηση της τάσης. Αυτή η αύξηση της τάσης προκαλείται από μια αντίστροφη ροή της ισχύος και εξαρτάται από την αντίσταση, τη χωρητικότητα και τον συντελεστή ισχύος της μονάδας. Οι συγκεκριμένες διακυμάνσεις της τάσης μπορούν να οδηγήσουν στην εμφάνιση βλαβών στις ηλεκτρικές συσκευές των πελατών και στον εξοπλισμό του δικτύου, μειώνοντας την αξιοπιστία του ηλεκτρικού δικτύου [188].

Η ενσωμάτωση των κατανεμημένων μονάδων παραγωγής ενέργειας στο δίκτυο διανομής μπορεί να είναι επωφελής για τους διαχειριστές των συστημάτων ηλεκτρικής ενέργειας, τις εταιρείες παροχής υπηρεσιών ενέργειας και τους καταναλωτές. Ωστόσο, εάν κρίσιμες πτυχές όπως το σημείο σύνδεσης και η χωρητικότητά τους δεν καθοριστούν κατάλληλα, θα μπορούσε να προκληθεί υποβάθμιση της απόδοσης του δικτύου. Αυτή η κατάσταση θα μπορούσε επίσης να οδηγήσει σε αύξηση των απωλειών ενέργειας και της διακύμανσης της τάσης. Η επιλογή ενός βέλτιστου σημείου σύνδεσης των κατανεμημένων μονάδων παραγωγής ενέργειας βελτιώνει την αξιοπιστία και την ασφάλεια του συστήματος και μειώνει το ποσοστό των απωλειών ενέργειας [190].

Τέτοιου είδους διαταραχές στις λειτουργίες του ηλεκτρικού δικτύου, καθιστούν τη βέλτιστη διαχείριση της ενέργειας, κάτι παραπάνω από επιτακτική, ώστε να αυξηθεί η αξιοπιστία και η απόδοση των ηλεκτρικών δικτύων. Σύμφωνα με τους Das και συν. (2020), ο μετριασμός αυτών των διαταραχών μπορεί να επιτευχθεί με την παροχή σταθερής χωρητικότητας (firm capacity) και λειτουργικού αποθέματος (operating reserve). Με το όρο σταθερή χωρητικότητα, οι συγγραφείς εννοούν την σταθερότητα των διαθέσιμων φορτίων για ορισμένο χρόνο, ενώ το λειτουργικό απόθεμα αφορά τη

βραχυπρόθεσμη διαθέσιμη χωρητικότητα που γίνεται κρίσιμη λόγω των κυμαινόμενων αναγκών ενεργειακής παραγωγής [191].

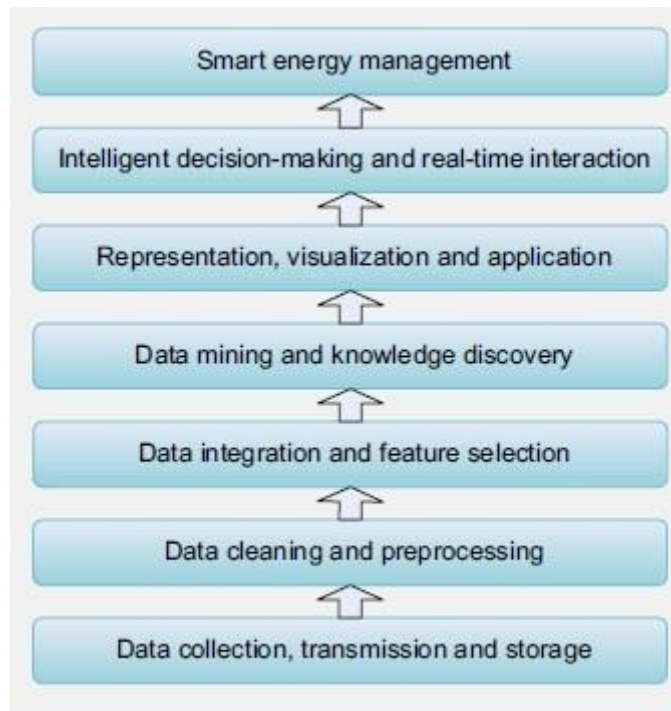
Όπως έχει ήδη αναφερθεί, τα έξυπνα δίκτυα παροχής ηλεκτρικής ενέργειας έχουν σχεδιαστεί με σκοπό την ενσωμάτωση των εννοιών της ροής πληροφοριών και της ενεργειακής ροής, επιτυγχάνοντας την ταυτόχρονη συλλογή δεδομένων και μεταφορά ενέργειας. Λόγω αυτής της φιλοσοφίας σχεδιασμού των έξυπνων δικτύων παροχής ηλεκτρικής ενέργειας, τα Big Data μπορούν να χρησιμοποιηθούν για την ενίσχυση της πρόβλεψης και του ελέγχου των επιπέδων της ζήτησης για ενεργειακή κατανάλωση, όπως επίσης και για την καλύτερη κατανόηση των δεδομένων που προέρχονται από τις καταναλωμένες μονάδες παραγωγής ενέργειας, αυξάνοντας με τον τρόπο αυτό την αποτελεσματικότητα των διαδικασιών διαχείρισης της ενέργειας μέσω υιοθέτησης πιο προηγμένων μεθόδων [142].

5.2 Έξυπνη διαχείριση ενέργειας με γνώμονα τα δεδομένα

Η επίτευξη των στόχων της έξυπνης διαχείρισης της ενέργειας με γνώμονα τα δεδομένα έχει αποτελέσει θέμα πολλών μελετών στη βιβλιογραφία. Για παράδειγμα, οι Wei και συν. (2017) παρουσίασαν έναν τρόπο βέλτιστης διαχείρισης και ελέγχου της ενέργειας σε έξυπνα οικιακά συστήματα μικροδικτύων (microgrid), με βάση τα δεδομένα που συλλέγονται για το μέγεθος των φορτίων και τον ρυθμό της κατανάλωσής τους [192]. Ο M. Rahmani-Andebili (2017) παρουσίασε μια προσαρμοστική και δυναμική προσέγγιση διαχείρισης της ενέργειας που αποθηκεύεται σε συστήματα αποθήκευσης ενέργειας και προέρχεται από καταναλωμένες μονάδες παραγωγής ΑΠΕ, οι οποίες είναι εγκατεστημένες σε κεντρικό ηλεκτρικό σύστημα διανομής. Ο προτεινόμενος τρόπος διαχείρισης βασίζεται στη χρήση ιστορικών δεδομένων σχετικά με τις καιρικές συνθήκες, αλλά και στα δεδομένα που συλλέγονται συνεχώς για τις συνθήκες αυτές, βάσει των οποίων προκύπτουν μοντέλα υπολογισμού της παραγόμενης ενέργειας από τις καταναλωμένες μονάδες ΑΠΕ [193]. Οι Mirakhorli & Dong (2018) παρουσίασαν ένα μοντέλο διαχείρισης της ενέργειας που καταναλώνεται σε μεγάλα κτιριακά συγκροτήματα, με βάση τα δεδομένα που λαμβάνονται για την καταναλωτική συμπεριφορά των ενοίκων. Τα δεδομένα αυτά αφορούσαν την λειτουργία μεγάλων καταναλωτικών μονάδων των συγκροτημάτων, όπως τα κλιματιστικά, οι θερμοσίφωνες και τα συστήματα φόρτισης των ηλεκτρικών οχημάτων [194]. Οι Wu & Wang (2018) αντιμετώπισαν τη διαχείριση της καταναλωμένης ενέργειας των

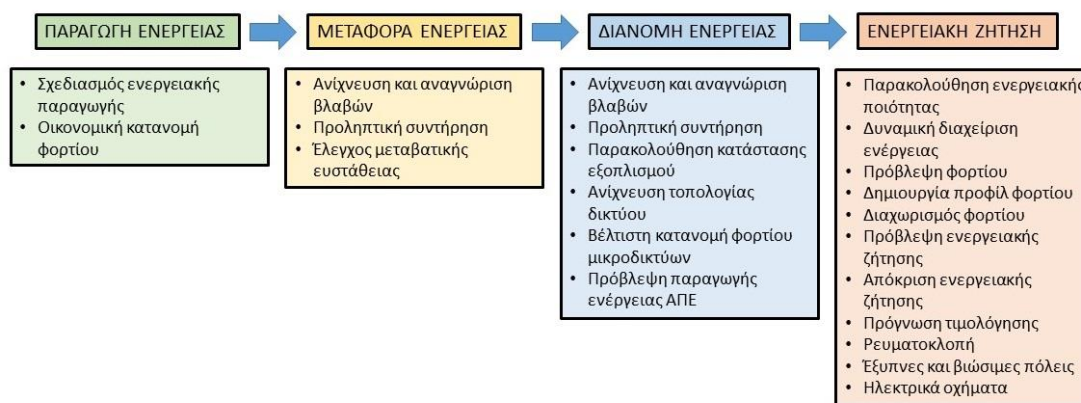
μικροδικτύων ως ζήτημα βελτιστοποίησης του ελέγχου και παρουσίασαν μία μέθοδο διαχείρισής της με χρήση προσαρμοστικού και δυναμικού αλγορίθμου βαθιάς μάθησης. Στο συγκεκριμένο μοντέλο, τα συλλεγόμενα δεδομένα αφορούσαν την ενέργεια που παράγεται από τις κατανεμημένες μονάδες των μικροδικτύων σε πραγματικό χρόνο. Τα δεδομένα χρησιμοποιούνται από τον αλγόριθμο βαθιάς μάθησης για την εκπαίδευση και τη βελτίωση της μονάδας ελέγχου των μικροδικτύων [195]. Τέλος, οι Bhattarai και συν. (2019) [30], καθώς και οι Ponnusamy και συν. (2021) [196], παρουσιάζουν δύο μελέτες στις οποίες αναφέρεται η χρήση των Big Data σε εφαρμογές διαχείρισης της ενέργειας στα έξυπνα δίκτυα παροχής ηλεκτρικής ενέργειας.

Σε γενικές γραμμές ένα μοντέλο έξυπνης διαχείρισης της ενέργειας με γνώμονα τα δεδομένα, έχει την ίδια φιλοσοφία με την ακολουθία των διαδικασιών της αλυσίδας αξίας των Big Data που παρουσιάστηκε σε προηγούμενο κεφάλαιο. Στο σημείο αυτό θα πρέπει να σημειωθεί ότι στο μοντέλο χρησιμοποιούνται Big Data ενέργειας. Έτσι, όπως φαίνεται στην εικόνα 5.2α, ένα τέτοιο μοντέλο αποτελείται από επτά βασικά στάδια. Από αυτά, τα τρία πρώτα αποτελούν σημαντικότερα στάδια προετοιμασίας των δεδομένων για τη διαδικασία εξόρυξης τους. Η εξόρυξη δεδομένων, σε συνδυασμό με τις πληροφορίες που προκύπτουν μέσα από την ανάλυση τους αποτελούν το πλέον σημαντικό στάδιο της συνολικής διαδικασίας της έξυπνης διαχείρισης της ενέργειας με γνώμονα τα δεδομένα. Στη συνέχεια, οι πληροφορίες που έχουν εξαχθεί από όλες τις προηγούμενες διεργασίες πάνω στα δεδομένα ενέργειας θα πρέπει να οπτικοποιηθούν και να παρουσιαστούν με τέτοιο τρόπο ώστε η εφαρμογή τους να μπορεί να ενισχύσει τη λειτουργία των συστημάτων λήψης αποφάσεων και ελέγχου ολόκληρου του έξυπνου δικτύου παροχής ηλεκτρικής ενέργειας. Το τελευταίο στάδιο της συνολικής διαδικασίας αφορά την επίτευξη των στόχων της έξυπνης διαχείρισης της ενέργειας, η οποία μπορεί να πραγματοποιηθεί μέσα από εφαρμογές όπως η διαχείριση της ενεργειακής απόδοσης, η δέσμευση των καταναλωτών, η παρακολούθηση σε πραγματικό χρόνο, η απόκριση στην ενεργειακή ζήτηση, ο έξυπνος έλεγχος κατανάλωσης, η δυναμική τιμολόγηση, κλπ.



Εικόνα 5.2α: Μοντέλο έξυπνης διαχείρισης ενέργειας με γνώμονα τα Big Data ενέργειας

Λαμβάνοντας υπόψη το παραπάνω μοντέλο έξυπνης διαχείρισης της ενέργειας με γνώμονα τα δεδομένα, προκύπτει το συμπέρασμα ότι η τεχνολογία των Big Data παίζει σημαντικό ρόλο στην όλη διαδικασία διαχείρισης ενός έξυπνου δικτύου παροχής ηλεκτρικής ενέργειας, αφού μπορεί να χρησιμοποιηθεί σε όλες τις εφαρμογές διαχείρισης της ενέργειας που αντιστοιχούν στις λειτουργίες ενός τέτοιου δικτύου, όπως η παραγωγή, η μεταφορά, η διανομή και η κατανάλωση της ενέργειας [31]. Η μελέτη της βιβλιογραφίας ανέδειξε ότι το μεγαλύτερο ποσοστό των ερευνών πάνω στη χρήση των Big Data στον τομέα της ενέργειας, και πιο συγκεκριμένα στα έξυπνα ηλεκτρικά δίκτυα, για την ενίσχυση της λήψης αποφάσεων σχετικά με την έξυπνη διαχείριση και έλεγχο της ενέργειας, εστιάζεται σε τέσσερα βασικά ζητήματα, όπως η διαχείριση της παραγωγής ενέργειας, η διαχείριση της ενέργειας των μικροδικτύων και των ΑΠΕ, η διαχείριση του εξοπλισμού και της λειτουργίας του και η διαχείριση της ενεργειακής ζήτησης [197]. Στην εικόνα 5.2β παρουσιάζονται κάποιες από αυτές τις εφαρμογές διαχείρισης της ενέργειας, οι οποίες αφορούν την παραγωγή, τη μεταφορά και τη διανομή ενέργειας, καθώς και την ενεργειακή ζήτηση. Οι εφαρμογές αυτές θα αναλυθούν στις επόμενες ενότητες του κεφαλαίου.



Εικόνα 5.2β: Εφαρμογές διαχείρισης ενέργειας smart grid με γνώμονα τα Big Data ενέργειας

Όπως γίνεται εύκολα κατανοητό από την εικόνα 5.2β, κάποιες από αυτές τις εφαρμογές μπορούν να συμπεριληφθούν σε περισσότερα από ένα τμήματα του συνολικού ηλεκτρικού δικτύου. Για παράδειγμα, η ανίχνευση και αναγνώριση βλαβών μπορεί να εφαρμοστεί στο τμήμα της μεταφοράς, αλλά και της διανομής της ενέργειας. Υπάρχουν και περιπτώσεις εφαρμογών, όμως, που εμπεριέχονται μεταξύ τους. Για παράδειγμα, η πρόβλεψη της ενεργειακής ζήτησης μπορεί να εμπεριέχεται στην πρόβλεψη των φορτίων, ενώ η απόκριση της ενεργειακής ζήτησης μπορεί να εμπεριέχεται στη δημιουργία των προφίλ των φορτίων. Επίσης, οι εφαρμογές στο πλαίσιο των έξυπνων πόλεων και των ηλεκτρικών οχημάτων εμπεριέχονται στο πλαίσιο του μοντέλου απόκρισης στην ενεργειακή ζήτηση. Το μόνο σίγουρο, που μπορεί να προκύψει ως συμπέρασμα είναι ότι όσο η ενέργεια προχωράει από το στάδιο της παραγωγής της προς την κατανάλωση, τόσο οι εφαρμογές των Big Data αυξάνονται.

Ιδιαίτερα για το δίκτυο διανομής, όπως θα αναφερθεί σε επόμενη ενότητα, πολλές μελέτες διακρίνουν τις εφαρμογές των Big Data σε βραχυπρόθεσμες και μακροπρόθεσμες. Στο πλαίσιο της παρούσας εργασίας, οι εφαρμογές των Big Data στο δίκτυο διανομής της ενέργειας θα χωριστούν με βάση το επίπεδο στο οποίο εφαρμόζονται, δηλαδή σε επίπεδο δικτύου (διαχείριση διανομής ενέργειας) και σε επίπεδο πελάτη/καταναλωτή (ενεργειακή ζήτηση). Αυτή η διαφοροποίηση επιλέχθηκε για να μπορέσει να γίνει και μια αναφορά στη διαχείριση της ενέργειας των μικροδικτύων και των ΑΠΕ, τα οποία όμως, όπως φαίνεται στην εικόνα 5.1α, δεν εντάσσονται και στα δύο επίπεδα, καθώς τα μικροδίκτυα εντάσσονται μόνο σε επίπεδο δικτύου διανομής, ενώ οι ΑΠΕ και στα δύο επίπεδα (δικτύου διανομής και πελάτη/καταναλωτή).

5.3 Διαχείριση ενεργειακής παραγωγής

Η βελτιστοποίηση της ενεργειακής παραγωγής και του προγραμματισμού της είναι δυνατή μέσω της χρήσης των Big Data και των τεχνικών ανάλυσής τους. Η εκμετάλλευση του μεγάλου όγκου των συλλεγόμενων δεδομένων ενέργειας και η χρήση μεθόδων και τεχνικών ανάλυσής του, δίνει τη δυνατότητα σημαντικής βελτίωσης της απόδοσης της ενεργειακής παραγωγής, καθώς επίσης και μείωσης του κόστους της. Έτσι, κατά τις διαδικασίες λήψης αποφάσεων σχετικά με το επίπεδο της παραγωγής της ηλεκτρικής ενέργειας, ο σχεδιασμός και προγραμματισμός της ενεργειακής παραγωγής, αλλά και η οικονομική κατανομή του φορτίου θεωρούνται από τα πλέον σημαντικά ζητήματα [184].

Στην περίπτωση της διαχείρισης της παραγωγής ενέργειας, τα δεδομένα που συλλέγονται αφορούν δεδομένα απόδοσης του έξυπνου δικτύου παροχής ηλεκτρικής ενέργειας, δεδομένα κόστους ενεργειακής παραγωγής, δεδομένα φορτίου ενεργειακής ζήτησης, δεδομένα ζωνών λειτουργίας, κλπ. Ο συνδυασμός των τεχνικών ανάλυσης των δεδομένων ενέργειας που συλλέγονται με την τεχνολογία του IoT, μπορεί να βελτιώσει σημαντικά την ενεργειακή απόδοση και να μειώσει το κόστος παραγωγής [198].

5.3.1 Σχεδιασμός ενεργειακής παραγωγής

Ο σχεδιασμός της ενεργειακής παραγωγής αποσκοπεί στη δημιουργία ενός βέλτιστου μείγματος παραγωγής ενέργειας με όσο το δυνατόν μεγαλύτερη αύξηση ενός δείκτη απόδοσης του έξυπνου ηλεκτρικού δικτύου, όπως για παράδειγμα η ελαχιστοποίηση του κόστους λειτουργίας του, και η ταυτόχρονη ικανοποίηση ενός συνόλου συνθηκών που σχετίζονται, με την ασφάλεια της ενεργειακής μεταφοράς και κατανομής, ο περιορισμός των ενεργειακών πόρων, η ενεργειακή ποικιλομορφία, οι περιβαλλοντικές επιπτώσεις καθώς και οι παράγοντες διείσδυσης των ΑΠΕ στο συνολικό ενεργειακό σύστημα. Όλα αυτά τα στοιχεία καθιστούν το σχεδιασμό της ενεργειακής παραγωγής μια ιδιαίτερα δύσκολη διαδικασία που απαιτεί την εξέταση πολλών, και συχνά αλληλένδετων, πτυχών [199].

Σύμφωνα με τους Jiang και συν. (2019), η χρήση των Big Data στις περιπτώσεις του σχεδιασμού και προγραμματισμού της ενεργειακής παραγωγής μπορεί να γίνει με τρεις τρόπους. Ο πρώτος τρόπος είναι η συλλογή δεδομένων των εκπομπών του διοξειδίου

του άνθρακα και των μονάδων παραγωγής και η χρήση τεχνικών προγραμματισμού περιορισμών (constraint programming) και ασαφούς πιθανολογικού μοντέλου (fuzzy possibilistic model). Ο δεύτερος τρόπος αφορά το συνδυασμό συλλογής δεδομένων απόδοσης του έξυπνου δικτύου παροχής ηλεκτρικής ενέργειας και χρήσης του προγραμματισμού περιορισμών. Ο τρίτος τρόπος αφορά τη συλλογή δεδομένων του κόστους ενεργειακής παραγωγής και τη χρήση τεχνικών ανάλυσης μέσης διακύμανσης (portfolio theory) και δυναμικού προγραμματισμού με βάση την ενεργειακή ζήτηση. Ανεξαρτήτως τρόπου χρήσης των Big Data, ο στόχος της σημαντικής βελτίωσης της ενεργειακής απόδοσης και της μείωσης του κόστους παραγωγής είναι εφικτός [198].

5.3.2 Οικονομική κατανομή φορτίου

Η οικονομική κατανομή φορτίου (Economic Load Dispatch – ELD) αποτελεί ένα σημαντικότερο ζήτημα των δικτύων ενεργειακής παραγωγής καθώς αφορά το λειτουργικό τους κόστος. Η επίλυση του συγκεκριμένου ζητήματος αποσκοπεί στην βέλτιστη κατανομή των φορτίων στα δίκτυα, έτσι ώστε να ελαχιστοποιηθεί το κόστος της ενεργειακής παραγωγής, με παράλληλη ικανοποίηση της συνολικής ενεργειακής ζήτησης, χωρίς να υπερβαίνονται τα λειτουργικά όρια των σταθμών παραγωγής ενέργειας [200].

Σύμφωνα με τους Jiang και συν. (2019), η χρήση των Big Data στις περιπτώσεις της οικονομικής κατανομής φορτίου μπορεί να γίνει με τέσσερις τρόπους. Ο πρώτος τρόπος αφορά τη συλλογή δεδομένων της ενεργειακής ζήτησης, των παραμέτρων λειτουργίας των σταθμών παραγωγής ενέργειας και των απωλειών μεταφοράς της ενέργειας και η χρήση του αλγόριθμου βελτιστοποίησης με βάση τη βιογεωγραφία (Biogeography-Based Optimization (BBO) algorithm). Στον δεύτερο τρόπο, τα δεδομένα που συλλέγονται αφορούν την ενεργειακή ζήτηση, τις ζώνες λειτουργίας και τις απώλειες μεταφοράς της ενέργειας, τα οποία χρησιμοποιούνται ως είσοδοι σε αλγόριθμο βελτιστοποίησης σμήνους σωματιδίων (Particle Swarm Optimization - PSO). Ο τρίτος τρόπος περιλαμβάνει τη συλλογή δεδομένων της ενεργειακής ζήτησης, των ζωνών λειτουργίας και των ορίων του ρυθμού μεταβολής της ενεργειακής παραγωγής (ramp-rate limits), καθώς και τη χρήση μεθόδων διαφορικής εξέλιξης (differential evolution based methods). Τέλος, στον τέταρτο τρόπο, τα δεδομένα που συλλέγονται και αφορούν την ενεργειακή ζήτηση και τις παραμέτρους λειτουργίας των

σταθμών παραγωγής ενέργειας, χρησιμοποιούνται ως είσοδοι του αλγόριθμου αναζήτησης αρμονίας (harmony search algorithm) [198].

5.4 Διαχείριση μεταφοράς ενέργειας

Η ενέργεια που παράγεται από τους σταθμούς παραγωγής, μεταφέρεται στο δίκτυο διανομής μέσω γραμμών μεταφοράς υψηλής τάσης. Επειδή οι σταθμοί παραγωγής ηλεκτρικής ενέργειας βρίσκονται συνήθως σε ιδιαίτερα απομακρυσμένα σημεία από τα δίκτυα διανομής, η χρήση υψηλής τάσης για τη μεταφορά της παραγόμενης ενέργειας προτιμάται, για την ελάττωση των ενεργειακών απωλειών, που μπορούν να υπάρξουν σε μια τέτοια μεταφορά [201]. Όπως αναφέρθηκε σε προηγούμενο κεφάλαιο, ένα δίκτυο μεταφοράς ενέργειας είναι ένα αυτόματο σύστημα, το οποίο αποτελείται από διάφορα μέρη, όπως υποσταθμούς, γραμμές μεταφοράς και απομακρυσμένες τερματικές μονάδες (RTU), που τοποθετούνται στους υποσταθμούς. Οι μονάδες RTU χρησιμοποιούνται στο δίκτυο μεταφοράς για τον έλεγχο και την παρακολούθηση της κατάστασης των στοιχείων των υποσταθμών μέσω της εφαρμογής σύγχρονων τεχνολογιών [164].

Τα δεδομένα που παράγονται από τα διαφορετικά μέρη και τον εξοπλισμό του δικτύου μεταφοράς είναι ζωτικής σημασίας για την εύρυθμη λειτουργία του συστήματος. Τα δεδομένα που συλλέγονται από τις μονάδες RTU είναι μεγάλου όγκου. παράγονται γρήγορα, προέρχονται από διάφορες πηγές-υποσταθμούς, είναι συνήθως διαφορετικής μορφής και η διαχείρισή τους παρουσιάζει πολλές προκλήσεις. Μια τέτοια διαχείριση, επεξεργασία και ανάλυση των δεδομένων που συλλέγονται από τις μονάδες RTU είναι πιο εύκολο να πραγματοποιηθεί χρησιμοποιώντας την τεχνολογία των Big Data. Η εκμετάλλευση του μεγάλου όγκου των συλλεγόμενων δεδομένων από τις μονάδες RTU και η χρήση μεθόδων και τεχνικών ανάλυσής του, δίνει τη δυνατότητα καλύτερης αντιμετώπισης ζητημάτων, όπως η προσφορά και ζήτηση ενέργειας, η εύρυθμη λειτουργία του δικτύου μεταφοράς, η βελτιστοποίηση της απόδοσής του, κλπ. [202]. Για την διευκόλυνση της αντιμετώπισης αυτών των ζητημάτων, η χρήση της τεχνολογίας των Big Data είναι σε θέση να υποστηρίξει εφαρμογές, όπως η ανίχνευση και ο εντοπισμός βλαβών, η προληπτική συντήρηση του εξοπλισμού του δικτύου μεταφοράς και ο έλεγχος της μεταβατικής ευστάθειας [184]. Καθώς οι εφαρμογές της ανίχνευσης και του εντοπισμού βλαβών, αλλά και της προληπτικής συντήρησης του εξοπλισμού, υλοποιούνται και στο δίκτυο διανομής της ενέργειας, η ανάλυσή τους θα

γίνει σε αυτό το δίκτυο, αφού συμπεριλαμβάνει και τις βλάβες και τη συντήρηση του εξοπλισμού των μικροδικτύων και των συστημάτων ΑΠΕ, που ενσωματώνονται σε αυτό.

5.4.1 Έλεγχος μεταβατικής ευστάθειας

Με τον όρο μεταβατική ευστάθεια (transient stability) εννοείται η ικανότητα του συστήματος να διατηρήσει συγχρονισμό μετά την πρόκληση κάποιας σοβαρής μεταβατικής διαταραχής, η οποία συνήθως εξαρτάται από την κατάσταση του συστήματος πριν την εμφάνιση κάποιας βλάβης, τη σοβαρότητα της βλάβης και τον τρόπο αποκατάστασής της. Αν και η μεταβατική αστάθεια σε ένα σύστημα μεταφοράς είναι σπάνια, εντούτοις η εμφάνισή της είναι αρκετά σοβαρή, καθώς μπορεί να προκαλέσει διαδοχικές διακοπές στην ηλεκτροδότηση και, στη χειρότερη περίπτωση, εκτεταμένες διακοπές ρεύματος [203]. Για το λόγο αυτό, ο έλεγχος της μεταβατικής ευστάθειας είναι πολύ σημαντικός ως προς την ασφαλή λειτουργία ολόκληρου του ηλεκτρικού δικτύου. Ωστόσο, η σύγχρονη ενσωμάτωση των ΑΠΕ στα δίκτυα διανομής και η μεταβαλλόμενη ενεργειακή ζήτηση αποτελούν σημαντικές προκλήσεις ενός τέτοιου ελέγχου, καθώς αναγκάζουν το ηλεκτρικό δίκτυο να λειτουργεί κοντά στα όρια ασφαλούς λειτουργίας του [29].

Η αντιμετώπιση αυτών των προκλήσεων στο πλαίσιο του ελέγχου της μεταβατικής ευστάθειας των δικτύων μεταφοράς της ενέργειας έχει αποτελέσει φλέγον ζήτημα προς μελέτη και έρευνα. Στις μελέτες αυτές, η εκάστοτε προτεινόμενη τεχνική ελέγχου της μεταβατικής ευστάθειας του συστήματος θα πρέπει να λαμβάνει υπόψη τη συνολική προστασία του ηλεκτρικού δικτύου και την εμφάνιση κάποιας μεταβατικής αστάθειας μετά την πρόκληση κάποιας σοβαρής βλάβης που είχε σαν συνέπεια τη διακοπή της ηλεκτροδότησης. Η χρήση παραδοσιακά χρησιμοποιούμενων τεχνικών που βασίζονται στη χρήση χρονικής προσομοίωσης, όμως, δεν είναι σε θέση να παρέχει καθολικά αποτελέσματα, λόγω των πολλών μεταβλητών που παρουσιάζουν τα σύγχρονα συστήματα ενέργειας [204] και λόγω της δέσμευσης τεράστιων υπολογιστικών πόρων για την υλοποίησή τους [31]. Για το λόγο αυτό, σε άλλες μελέτες, χρησιμοποιούνται αποτελεσματικές προσεγγίσεις που εξάγουν χρήσιμες, για το σκοπό αυτό, πληροφορίες και μοτίβα, μέσα από τον τεράστιο όγκο των δεδομένων που συλλέγονται για τον έλεγχο της μεταβατικής ευστάθειας. Σύμφωνα με τους Zhang, Zhu & Li (2021), οι

προσεγγίσεις που προτείνονται μπορούν να ταξινομηθούν σε δύο μεγάλες κατηγορίες [205]: (α) τις προσεγγίσεις της αυτόματης μάθησης και (β) τις άμεσες τεχνικές.

Στις άμεσες τεχνικές ελέγχου της μεταβατικής ευστάθειας συνήθως χρησιμοποιούνται ενεργειακές συναρτήσεις που δημιουργούνται με εφαρμογή της θεωρίας Lyapunov [206]. Σε μελέτες που έχουν παρουσιαστεί στη βιβλιογραφία για το σκοπό αυτό, η εκτίμηση της μεταβατικής ευστάθειας ενός δικτύου μεταφοράς μπορεί να επιτευχθεί μέσω χρήσης μοντέλου Koopman (το οποίο δεν απαιτεί σύνθετη χρονική προσομοίωση του συστήματος μετά την εμφάνιση κάποιας βλάβης) [207], μέσω χρήσης απλών ισοδύναμων των γεννητριών των υποσταθμών [208] και μέσω χρήσης του γενικευμένου κριτηρίου ίσων εμβαδών (Extended Equal Area Criterion - EEAC) [209]. Επιπλέον, σε αρκετές από αυτές τις μελέτες, έχουν χρησιμοποιηθεί μονάδες μέτρησης φάσης (PMU) [210] και δυναμικής εκτίμησης της κατάστασης (DSE) [211] για τη συλλογή πληροφοριών σε πραγματικό χρόνο, βάσει των οποίων θα γίνει εφαρμογή κάποιας από τις παραπάνω μεθόδους για τον έλεγχο της μεταβατικής ευστάθειας.

Σε αντίθεση με τις άμεσες τεχνικές, οι προσεγγίσεις αυτόματης μάθησης αντιμετωπίζουν τον έλεγχο της μεταβατικής ευστάθειας ως πρόβλημα αναγνώρισης μοτίβων. Βασική φιλοσοφία των συγκεκριμένων προσεγγίσεων είναι η αξιοποίηση του μεγάλου όγκου δεδομένων που συλλέγονται μέσω δημιουργίας ενός μοντέλου αξιολόγησης των παραμέτρων λειτουργίας του δικτύου μεταφοράς και της κατάστασης της μεταβατικής ευστάθειάς του. Αυτή η φιλοσοφία παρουσιάζει τα πλεονεκτήματα της ισχυρής ικανότητας μάθησης και της γρήγορης ταχύτητας αξιολόγησης της μεταβατικής ευστάθειας των συστημάτων ισχύος [205]. Στη βιβλιογραφία και στις μελέτες που έχουν παρουσιαστεί, για τη δημιουργία αυτού του μοντέλου αξιολόγησης, χρησιμοποιούνται τεχνικές τεχνητών νευρωνικών δικτύων (Artificial Neural Networks – ANN) [212-215], μηχανών διανυσμάτων υποστήριξης (Support Vector Machines – SVM) [216,217], συνολικής μάθησης (Ensemble Learning –EL) [218,219] και βαθιάς μάθησης (Deep Learning – DL) [220-222].

5.5 Διαχείριση διανομής ενέργειας

Τα παραδοσιακά συστήματα διανομής της ηλεκτρικής ενέργειας παρουσίαζαν πάντα μια εξάρτηση από τις τεχνολογίες των πληροφοριών και των επικοινωνιών (ΤΠΕ), προκειμένου να επιτευχθεί η παρακολούθηση και ο αποτελεσματικότερος έλεγχος του

συνολικού δικτύου. Από τα μέσα της δεκαετίας του 1980 και μετά, οι διαδικασίες αυτές γίνονταν με χρήση των συστημάτων SCADA (Supervisory Control And Data Acquisition). Με την αύξηση όμως της πολυπλοκότητας των ηλεκτρικών δικτύων, από την εισαγωγή των κατανεμημένων μονάδων παραγωγής ενέργειας (που περιλαμβάνουν και τις ΑΠΕ) και των μικροδικτύων στο δίκτυο διανομής της ενέργειας, δημιουργήθηκε η ανάγκη για ενίσχυση της απόδοσής του όσον αφορά την αποτελεσματικότητα, την αξιοπιστία και την ευελιξία [223]. Παρόλο που τα σύγχρονα ηλεκτρικά δίκτυα είναι σε θέση να υποστηρίξουν μια πληθώρα νέων βελτιωμένων εφαρμογών, όπως η δυναμική βελτιστοποίηση των λειτουργιών τους, η ενσωμάτωση της απόκρισης στην ενεργειακή ζήτηση και η ενσωμάτωση των κατανεμημένων ενεργειακών πόρων (συμπεριλαμβανομένων των ΑΠΕ και των μικροδικτύων), η αλληλεξάρτηση του δικτύου ηλεκτρικής ενέργειας από τις ΤΠΕ, όχι μόνο δεν μειώθηκε, αλλά γνώρισε πρωτοφανή αύξηση [224]. Κάτι τέτοιο οδήγησε στη χρήση ΤΠΕ, που να μπορούν να υποστηρίξουν την υψηλή ενσωμάτωση των ΑΠΕ και των μικροδικτύων στα δίκτυα διανομής των σύγχρονων ηλεκτρικών δικτύων, αλλά και να μπορούν να ανταπεξέλθουν στις προκλήσεις που προκύπτουν από τον τεράστιο όγκο δεδομένων και πληροφοριών, ο οποίος ανταλλάσσεται πλέον μεταξύ του εξοπλισμού των συστημάτων διανομής και των συστημάτων που τον διαχειρίζονται [225].

Όπως αναφέρθηκε σε προηγούμενο κεφάλαιο, οι σύγχρονες υποδομές των δικτύων διανομής ηλεκτρικής ενέργειας, περιλαμβάνουν μια σειρά από έξυπνες μονάδες, μετρητές και αισθητήρες που χρησιμοποιούνται για τη βελτίωση της παρακολούθησης της λειτουργίας τους [164]. Πριν από την ευρεία χρήση των έξυπνων μετρητών, οι εταιρείες ηλεκτρισμού δεν μπορούσαν να έχουν πλήρη εικόνα της ενεργειακής ζήτησης, πλην των υποσταθμών των δικτύων διανομής. Επίσης, τα μόνα δεδομένα που είχαν στη διάθεσή τους ήταν αυτά της μηνιαίας κατανάλωσης των πελατών τους. Με τη χρήση των έξυπνων μετρητών, η εικόνα αυτή άλλαξε. Οι έξυπνοι μετρητές, εξοπλισμένοι με εξελιγμένους αισθητήρες, είναι σε θέση να συλλέγουν πληροφορίες που αφορούν, εκτός από την μηνιαία κατανάλωση, τα πρότυπα κατανάλωσης, την ποιότητα της παρεχόμενης ενέργειας, κλπ. Όλα αυτά τα δεδομένα μπορούν να χρησιμοποιηθούν για τη λήψη αποφάσεων που σχετίζονται με τη λειτουργία του δικτύου διανομής, το σχεδιασμό των ηλεκτρικών του συστημάτων, κλπ. Το πρόβλημα που τίθεται, όμως από μια τέτοια χρήση των έξυπνων μετρητών είναι ο τεράστιος όγκος των δεδομένων που συλλέγεται, καθώς η συλλογή δεδομένων από το σύνολο των

μετρητών σε ένα δίκτυο διανομής γίνεται κάθε 15 ή 30 λεπτά [226]. Με τον τρόπο αυτό, μελέτες έχουν δείξει ότι μόνο στο δίκτυο διανομής συλλέγονται τουλάχιστον 50Gb δεδομένων την ημέρα [227]. Αυτός ο τεράστιος όγκος των δεδομένων που συλλέγεται με αυτό τον ρυθμό, και με σκοπό την επίτευξη των στόχων της συλλογής του, απαιτεί αξιόπιστη επεξεργασία και ανάλυση, τις περισσότερες φορές σε πραγματικό χρόνο, κάτι που μπορεί να επιτευχθεί μόνο με τη χρήση της τεχνολογίας των Big Data [225].

Πολλές μελέτες που ασχολούνται με το ζήτημα των εφαρμογών της τεχνολογίας των Big Data στο τμήμα της διανομής της ηλεκτρικής ενέργειας, τις ταξινομούν σε δύο μεγάλες κατηγορίες. Για παράδειγμα, οι εφαρμογές της ανίχνευσης ενεργειακής απάτης (ρευματοκλοπής), της ανίχνευσης σύνδεσης των οχημάτων EV στο ηλεκτρικό δίκτυο, η πρόβλεψη του ενεργειακού φορτίου μιας περιοχής, η πρόβλεψη της παραγωγής ενέργειας από ΑΠΕ, η παρακολούθηση της κατάστασης του εξοπλισμού και η ανίχνευση και αναγνώριση βλαβών σε πραγματικό χρόνο θεωρούνται από κάποιους μελετητές ως βραχυπρόθεσμες εφαρμογές. Αντίθετα, εφαρμογές όπως η δημιουργία καταναλωτικών προφίλ στο πλαίσιο της προγνωστικής τιμολόγησης, η ανίχνευση της τοπολογίας του δικτύου και η προληπτική συντήρηση θεωρούνται ως μακροπρόθεσμες εφαρμογές [225-227].

5.5.1 Διαχείριση της ενέργειας μικροδικτύων και ΑΠΕ

Όπως έχει ήδη αναφερθεί, οι ΑΠΕ αποτελούν σημαντικό μέρος των σύγχρονων ενεργειακών συστημάτων και η ενέργεια που παράγουν ενσωματώνεται στα μικροδίκτυα, ένα μοντέλο κατανεμημένης παραγωγής ενέργειας. Αυτή η παραγόμενη ενέργεια όμως είναι ευμετάβλητη, καθώς επηρεάζεται σημαντικά από τις καιρικές συνθήκες. Λόγω αυτών των ζητημάτων, η ανάλυση των δεδομένων ενέργειας διαδραματίζει σημαντικό ρόλο στη διαχείριση της ενέργειας που προέρχεται από τις ΑΠΕ και χρησιμοποιείται στα μικροδίκτυα. Για παράδειγμα, η πρόβλεψη της ενεργειακής παραγωγής των ΑΠΕ θα είναι πιο ακριβής και αποτελεσματική αν πραγματοποιηθεί μέσω της ανάλυσης των τεράστιου όγκου δεδομένων που μπορούν να συλλεχθούν και αφορούν τις καιρικές συνθήκες. Ο συνδυασμός των δεδομένων παραγωγής και κατανάλωσης ενέργειας, με τα δεδομένα που προέρχονται από συστήματα GIS και τα δεδομένα της θερμοκρασίας, της ατμοσφαιρικής πίεσης, της υγρασίας, της ταχύτητας και της διεύθυνσης του ανέμου, κλπ., μπορεί να υποστηρίξει

την επιλογή τοποθεσιών εγκατάστασης μονάδων παραγωγής ενέργειας από ΑΠΕ, ώστε να βελτιωθεί η ενεργειακή παραγωγή και η απόδοση των συστημάτων [31]. Δύο από τις σημαντικότερες εφαρμογές των Big Data για τη διαχείριση της ενέργειας των μικροδικτύων και των ΑΠΕ, αφορούν τη βέλτιστη κατανομή των φορτίων στα μικροδίκτυα και την πρόβλεψη της παραγωγής ενέργειας από τις ΑΠΕ.

1 Βέλτιστη κατανομή φορτίων στα μικροδίκτυα

Η χρήση των μικροδικτύων στα έξυπνα δίκτυα παροχής ηλεκτρικής ενέργειας αποσκοπεί στον συντονισμό των κατανεμημένων μονάδων παραγωγής ενέργειας με το δίκτυο ηλεκτρικής ενέργειας. Παρά τα σημαντικά οφέλη μιας τέτοιας χρήσης, η εφαρμογή των μικροδικτύων στα έξυπνα δίκτυα παροχής ηλεκτρικής ενέργειας αντιμετωπίζει ορισμένες προκλήσεις όσον αφορά τη διαμόρφωση του συστήματος, την επαρκή απαίτηση χωρητικότητας αποθήκευσης ενέργειας, τη διαχείριση της ενέργειας, την κατανομή της εφεδρικής ισχύος και τον έλεγχο. Ένα από τα κρίσιμα ζητήματα είναι ο βέλτιστος συντονισμός των υβριδικών πηγών ενέργειας στο μικροδίκτυο με το κύριο δίκτυο. Τρόπο επίλυσης του συγκεκριμένου ζητήματος μπορεί να αποτελέσει η βέλτιστη κατανομή των φορτίων των μικροδικτύων, η οποία είναι σε θέση να μειώσει το κόστος λειτουργίας των μικροδικτύων και τις εκπομπές των ρύπων, να αυξήσει την αξιοπιστία του μικροδικτύου αλλά και του συνολικού δικτύου ηλεκτρικής ενέργειας, κλπ., ικανοποιώντας παράλληλα διάφορους περιορισμούς, όπως τους λειτουργικούς περιορισμούς των κατανεμημένων μονάδων παραγωγής ενέργειας, τους περιορισμούς φόρτισης, εκφόρτισης και χωρητικότητας των μπαταριών αποθήκευσης ενέργειας και τους περιορισμούς της ενεργειακής ισορροπίας [228].

Κατά τη διαδικασία βέλτιστης κατανομής των φορτίων των μικροδικτύων συλλέγονται δεδομένα ενέργειας που αφορούν τις παραμέτρους των κατανεμημένων μονάδων παραγωγής ενέργειας, τη διαθεσιμότητα των μονάδων αποθήκευσης ενέργειας, την πρόβλεψη για ενεργειακή ζήτηση, την παραγωγή ενέργειας από τις ΑΠΕ και την τιμή κοστολόγησης της ενεργειακής μονάδας. Τα δεδομένα αυτά αποστέλλονται ως παράμετροι εισόδου σε αλγόριθμο βελτιστοποίησης, οι έξοδοι του οποίου παρουσιάζουν τον καλύτερο προγραμματισμό παραγωγής ενέργειας για όλες τις ώρες της ημέρας [228]. Λόγω της δυνατότητας επίτευξης όλων των παραπάνω στόχων οι χρησιμοποιούμενοι αλγόριθμοι είναι συνήθως γνωστοί με τον όρο αλγόριθμοι βελτιστοποίησης πολλαπλών στόχων (multi-objective optimization algorithms) και

μπορούν να περιλαμβάνουν πολλά είδη τεχνικών, όπως γραμμικού προγραμματισμού [229], μη γραμμικού προγραμματισμού [230], στοχαστικού προγραμματισμού [231], ευρετικής προσέγγισης [232], εξελικτικής προσέγγισης [233], ελέγχου μοντέλου πρόβλεψης [234] και στιβαρής βελτιστοποίησης [235].

2 Πρόβλεψη παραγωγής ενέργειας από ΑΠΕ

Παρόλα τα πλεονεκτήματα που παρουσιάζουν οι ΑΠΕ, η αβέβαιη και διακοπτόμενη συμπεριφορά τους ως προς την παραγωγή ενέργειας μπορεί να δημιουργήσει διάφορα προβλήματα στη σωστή διαχείριση της ενέργειας του δικτύου διανομής. Η πρόβλεψη της παραγωγής ενέργειας από τις ΑΠΕ, παίζει σημαντικό ρόλο στην εύρυθμη λειτουργία ολόκληρου του ηλεκτρικού δικτύου, αντιμετωπίζοντας ζητήματα που θα μπορούσαν να προκύψουν στον προγραμματισμό της παραγωγής ενέργειας. Όσον αφορά το δίκτυο διανομής, η πρόβλεψη της παραγωγής ενέργειας από τις ΑΠΕ υποβοηθά εφαρμογές όπως η διαχείριση του φορτίου, η αντιμετώπιση περιπτώσεων αυξημένης ζήτησης, η διαχείριση της παρεχόμενης ενέργειας, η προληπτική συντήρηση του εξοπλισμού, κλπ. [29, 31, 236]. Η σημαντικότητα της πρόβλεψης της παραγωγής ενέργειας από τις ΑΠΕ, έχει οδηγήσει την ακαδημαϊκή και ερευνητική κοινότητα στην διεξαγωγή μεγάλου αριθμού μελετών, στις οποίες έχουν παρουσιαστεί πολλές προσεγγίσεις, όπως φυσικά μοντέλα, στατιστικά μοντέλα, μοντέλα μηχανικής μάθησης, υβριδικά μοντέλα, κλπ. [31, 193, 237].

Τα φυσικά μοντέλα περιλαμβάνουν προσεγγίσεις προσομοίωσης των γεωγραφικών χαρακτηριστικών μιας περιοχής, στις οποίες τα δεδομένα που χρησιμοποιούνται αφορούν την πρόβλεψη των καιρικών συνθηκών σε τοπικό επίπεδο. Τα ζητήματα που παρουσιάζουν τα φυσικά μοντέλα είναι η ανάγκη χρήσης τεράστιας ποσότητας υπολογιστικών πόρων, η μικρότερη ακρίβεια πρόβλεψης, καθώς και η ακαταλληλότητα για περιπτώσεις βραχυπρόθεσμης πρόβλεψης [31]. Τα στατιστικά μοντέλα εφαρμόζουν μαθηματική μοντελοποίηση, με σκοπό την αναγνώριση των μοτίβων σε δεδομένα χρονοσειρών των ΑΠΕ. Σε μελέτες έχουν παρουσιαστεί στατιστικά μοντέλα που χρησιμοποιούν μοντέλα Markov [238], φίλτρα Kalman [239], μοντέλα αυτόματου επιταχυνόμενου-κινούμενου μέσου όρου [240], κλπ.

Με την ευρεία αύξηση της δημοτικότητας των μοντέλων μηχανικής μάθησης, η εφαρμογή τους στον τομέα της πρόβλεψης της παραγωγής ενέργειας από τις ΑΠΕ αύξησε την αξιοπιστία των προσεγγίσεων πρόβλεψης. Σε μελέτες έχουν παρουσιαστεί

αλγόριθμοι μηχανικής μάθησης που χρησιμοποιούν μοντέλα γραμμικής παλινδρόμησης [241], πολυστρωματικών νευρωνικών δικτύων [242], μηχανών υποστήριξης διανυσμάτων [243], παλινδρόμησης δέντρων αποφάσεων [244], κλπ. [245]. Επίσης, λόγω της εγγενούς διαλείπουσας και μη γραμμικής φύσης της παραγωγής ενέργειας από τις ΑΠΕ, η χρήση μοντέλων βαθιάς μάθησης έχει αποδειχθεί εξαιρετικά αποδοτική και αποτελεσματική, για την πρόβλεψή της [31, 231]. Σε αντίστοιχες μελέτες έχουν χρησιμοποιηθεί μοντέλα βαθιάς μάθησης, όπως δίκτυα βαθιάς πίστης, αυτοκωδικοποιητές, συνελκτικά νευρωνικά δίκτυα, μακροπρόθεσμης και βραχυπρόθεσμης μνήμης, μοντέλα συνόλων βαθιάς μάθησης, κλπ. [246-248].

Η ανάλυση της ακρίβειας και της αξιοπιστίας όλων αυτών των μοντέλων, υποδηλώνει ότι με την πάροδο των χρόνων, για την πρόβλεψη της ενέργειας από τις ΑΠΕ χρησιμοποιούνται μοντέλα με όλο και μεγαλύτερη ακρίβεια πρόβλεψης, κάτι που αναμένεται να βοηθήσει στη βελτιστοποίηση του προγραμματισμού παρεχόμενης από το ηλεκτρικό δίκτυο ενέργειας και την επίτευξη μιας πιο αξιόπιστης ενσωμάτωσης των ΑΠΕ στα δίκτυα διανομής μέσω των μικροδικτύων [31].

5.5.2 Ανίχνευση και αναγνώριση βλαβών

Η ιδέα της δημιουργίας των έξυπνων δικτύων παροχής ηλεκτρικής ενέργειας βασίστηκε στην προσπάθεια μείωσης των εκπομπών άνθρακα και στη επίτευξη όσο το δυνατόν καλύτερων συνθηκών βιωσιμότητας του περιβάλλοντος, καθώς επίσης και στην προσπάθεια εκπλήρωσης των αναγκών για παραγωγή καθαρών και εναλλακτικών μορφών ενέργειας. Ιδιαίτερα η χρήση καταναλωμένων μονάδων στα δίκτυα διανομής ηλεκτρικής ενέργειας δημιούργησε τις κατάλληλες συνθήκες ενσωμάτωσης των ΑΠΕ στα έξυπνα δίκτυα παροχής ηλεκτρικής ενέργειας. Όπως αναφέρθηκε σε προηγούμενη ενότητα, μια τέτοια υλοποίηση επέτρεψε τη δημιουργία μικροδικτύων, τα οποία, υποστηρίζοντας την παραγωγή ενέργειας πιο κοντά στα φορτία των καταναλωτών, ενισχύουν την αξιοπιστία του ηλεκτρικού δικτύου, βοηθούν στη βελτίωση της παροχής ισχύος και στη μείωση των απωλειών μεταφοράς. Επιπλέον, η χρήση των μικροδικτύων σε λειτουργία νησίδας, επιτρέπει την προστασία των φορτίων από βλάβες που μπορούν να προκύψουν λόγω διακυμάνσεων της τάσης, της συχνότητας, κλπ. [29], [31].

Παρόλα αυτά, η ευμετάβλητη φύση των περισσότερων ΑΠΕ δημιουργεί διακυμάνσεις στην παραγόμενη ενέργεια από αυτές, με αποτέλεσμα η αποτελεσματικότητα, η

σταθερότητα και η αξιοπιστία του ηλεκτρικού δικτύου να αποτελούν πραγματικές προκλήσεις [249]. Η χρήση των δικτύων IIDG (Inverted Integrated Distribution Grids) αποτελεί μία λύση βελτίωσης της ποιότητας ισχύος των μικροδικτύων, που όμως, λόγω της χαμηλής αδράνειας που παρουσιάζουν, στην περίπτωση πρόκλησης βλαβών στα μικροδίκτυα, ο μη έγκαιρος εντοπισμός και αποκατάστασή τους, μπορούν να αποτελέσουν σημαντικά προβλήματα για την λειτουργία τους. Οι κλασικές προσεγγίσεις εντοπισμού και αποκατάστασης αυτών των βλαβών βασίζονται στη μέτρηση των υπερεντάσεων και των ρευμάτων αρνητικής ακολουθίας, οι οποίες όμως δεν είναι κατάλληλες για τα μικροδίκτυα [250].

Για τα μικροδίκτυα που συνδέονται στο έξυπνο δίκτυο παροχής ηλεκτρικής ενέργειας, οι έντονες καιρικές συνθήκες ή οι διακοπές ρεύματος στο ηλεκτρικό δίκτυο ενδέχεται να προκαλέσουν βλάβες στις νησίδες, οι οποίες μπορούν να αποτελέσουν απειλή για την ασφαλή λειτουργία και να δημιουργήσουν άλλα τεχνικά ζητήματα σε ολόκληρο το ηλεκτρικό δίκτυο. Η χρήση μεθόδων και τεχνικών ανάλυσης των δεδομένων ενέργειας, που αναφέρθηκαν σε προηγούμενο κεφάλαιο, μπορεί να αποτελέσει λύση των συγκεκριμένων ζητημάτων. Οι Hashemi, Mohammadi & Kargarian (2017) παρουσίασαν έναν τρόπο εντοπισμού βλαβών στις νησίδες των μικροδικτύων, χρησιμοποιώντας τεχνητά νευρωνικά δίκτυα (Artificial Neural Networks - ANN), με σκοπό την εκπαίδευσή τους με χαρακτηριστικά που εξάγονται από την παρακολούθηση του σήματος του ρυθμού μεταβολής της συχνότητας (ROCOF) [251]. Οι Alam, Muttaqi & Bouzardoum (2017) παρουσίασαν μια προσέγγιση ανίχνευσης βλαβών στις νησίδες των μικροδικτύων, στην οποία χρησιμοποιείται μηχανή υποστήριξης διανυσμάτων (Support Vector Machine - SVM) που δέχεται ως δεδομένα εισόδου χαρακτηριστικά που εξάγονται από διάφορες μεταβλητές του συστήματος. Η διαδικασία εξαγωγής των χαρακτηριστικών υλοποιείται με τη χρήση ενός μεταβλητού παραθύρου, η βελτιστοποίηση του πλάτους του οποίου επιτυγχάνει υψηλότερο ρυθμό ανίχνευσης βλαβών [252]. Τέλος, οι James και συν. (2017) χρησιμοποίησαν μοντέλο βαθιάς μάθησης, το οποίο εκπαιδεύεται από στατιστικά χαρακτηριστικά διαφόρων παραμέτρων ενός μικροδικτύου, με σκοπό τον εντοπισμό και την ταξινόμηση των βλαβών σε αυτά [253].

Σύμφωνα με τους Bauman, Tuzhilin & Zaczynski (2017), τα Big Data που παράγονται από διάφορα μέσα κοινωνικής δικτύωσης, όπως το Twitter ή το Facebook, θα μπορούσαν να περιέχουν χρήσιμες πληροφορίες που να υποδεικνύουν την εμφάνιση

και τη θέση των διακοπών ρεύματος του έξυπνου ηλεκτρικού δικτύου σε πραγματικό χρόνο, αλλά τον έγκαιρο εντοπισμό των πιθανών σημείων επικείμενων διακοπών ρεύματος για την προληπτική αποκατάσταση της βλάβης [254]. Για παράδειγμα, οι Khan & Wei (2018) ανέπτυξαν έναν μηχανισμό εντοπισμού των διακοπών ρεύματος σε πραγματικό χρόνο χρησιμοποιώντας νευρωνικά δίκτυα MLP (Multi-Layer Perception), τα οποία εκπαιδεύονται από πληροφορίες που εξάγονται από την ανάλυση Big Data που παράγονται από το Twitter [161].

5.5.3 Παρακολούθηση κατάστασης εξοπλισμού

Η εμφάνιση βλάβης στον κρίσιμο εξοπλισμό του δικτύου διανομής ηλεκτρικής ενέργειας, όπως για παράδειγμα στους μετασχηματιστές ισχύος, είναι σημαντικό ζήτημα καθώς μπορεί να προκαλέσει διαδοχικές διακοπές στην ηλεκτροδότηση και, στη χειρότερη περίπτωση, εκτεταμένες διακοπές ρεύματος. Επομένως, η παρακολούθηση της κατάστασης του εξοπλισμού είναι ζωτικής σημασίας [203]. Η παρακολούθηση αυτή, όταν πραγματοποιείται στα επιμέρους στοιχεία του εξοπλισμού και όχι στο σύνολό του, διασφαλίζει σε μεγάλο βαθμό τη μείωση της κατανάλωσης ενέργειας από το συνολικό δίκτυο, βοηθώντας, τελικά, στην εξοικονόμηση ενέργειας και στην αποτελεσματικότερη διαχείρισή της [31].

Στα παραδοσιακά δίκτυα διανομής ενέργειας, ένα σύστημα παρακολούθησης της κατάστασης του εξοπλισμού τους βασίζεται στη χρήση μηχανισμών και μεθόδων ελέγχου του κατωφλίου λειτουργίας του εξοπλισμού, παρακολουθώντας διαφορετικές παραμέτρους και λαμβάνοντας τις αντίστοιχες μετρήσεις [31]. Στα σύγχρονα ηλεκτρικά δίκτυα όμως, η ενσωμάτωση των ΑΠΕ, σε συνδυασμό με τα χαρακτηριστικά της διαλείπουσας φύσης που τις συνοδεύουν, δημιουργεί ζητήματα ως προς την ορθή εκτίμηση του εναπομείναντος χρόνου λειτουργίας των βασικών στοιχείων του εξοπλισμού των δικτύων διανομής, όπως είναι οι μετασχηματιστές ισχύος, κλπ. [255]. Επομένως, η διαχείριση του κύκλου ζωής των μετασχηματιστών ισχύος, για την επίτευξη ενός πιο σταθερού και αξιόπιστου συστήματος διανομής της ηλεκτρικής ενέργειας, έχει προσελκύσει το ενδιαφέρον πολλών ερευνητών [29].

Στη βιβλιογραφία έχουν παρουσιαστεί διάφορες προσεγγίσεις παρακολούθησης της κατάστασης του εξοπλισμού των δικτύων διανομής ενέργειας, οι οποίες βασίζονται στην εφαρμογή τεχνικών τεχνητής νοημοσύνης και στη χρήση Big Data ενέργειας. Τέτοιες τεχνικές αφορούν τη χρήση γενετικών αλγορίθμων [256], ειδικών συστημάτων

τεχνητής νοημοσύνης (artificial intelligence expert systems) [257], τεχνητών νευρωνικών δικτύων [258], ασαφούς λογικής [259], μοντέλων βαθιάς μάθησης [260], κλπ.

Εκτός από αυτές τις προσεγγίσεις, στη βιβλιογραφία έχουν παρουσιαστεί και κάποια υβριδικά μοντέλα παρακολούθησης της κατάστασης του εξοπλισμού των δικτύων διανομής, που έχουν ως σκοπό την συμπλήρωση των όποιων ελλείψεων μπορούν να περιλαμβάνουν τα μοντέλα στα οποία χρησιμοποιείται μία μόνο τεχνική. Παράδειγμα υβριδικής προσέγγισης αποτελεί το μοντέλο παρακολούθησης των Sheng και συν. (2018), στο οποίο χρησιμοποιούνται τρεις αλγόριθμοι (Apriori, AprioriTid και AprioriHybrid) σε συνδυασμό με τεχνικές εξόρυξης δεδομένων, με σκοπό την λήψη πληροφοριών που σχετίζονται με τη λειτουργία του συστήματος διανομής και τις επικρατούσες μετεωρολογικές συνθήκες. Στη συνέχεια, οι πληροφορίες που προκύπτουν συνδυάζονται με ένα πιθανοτικό γραφικό μοντέλο για την πρόβλεψη πιθανής εμφάνισης δυσλειτουργίας στον εξοπλισμό του συστήματος διανομής [261]. Οι Aizuruwa και συν. (2019) παρουσίασαν μία αντίστοιχη υβριδική προσέγγιση, στην οποία η παρακολούθηση της κατάστασης του εξοπλισμού των δικτύων διανομής πραγματοποιείται με χρήση πιθανολογικής πρόβλεψης σε συνδυασμό με τεχνικές φιλτραρίσματος Kalman και Monte Carlo. Στη συγκεκριμένη προσέγγιση, η εκτίμηση της διάρκειας ζωής των μετασχηματιστών ισχύος βασίζεται στην εξαγωγή πληροφοριών βάσει των οποίων δημιουργείται ένα πιθανοτικό θερμικό μοντέλο και ένα πιθανοτικό μοντέλο διάρκειας της ζωής τους [262].

5.5.4 Προληπτική συντήρηση

Λόγω της μεγάλης πιθανότητας εμφάνισης κάποιου είδους δυσλειτουργίας σε διάφορα σημεία του εξοπλισμού του δικτύου διανομής, η ύπαρξη προληπτικής συντήρησής τους μπορεί να παίξει καθοριστικό ρόλο στη μείωση της πιθανότητας εμφάνισης διακοπών στην ηλεκτροδότηση [263]. Όπως έχει αναφερθεί, αυτή η πιθανότητα εμφάνισης κάποιου είδους δυσλειτουργίας του εξοπλισμού ενισχύεται από την ένταξη των ΑΠΕ στο δίκτυο διανομής. Η αβεβαιότητα παραγωγής ενέργειας από τις ΑΠΕ συνοδεύεται και από διακυμάνσεις της συχνότητας και της τάσης της παρεχόμενης ενέργειας, με τελικό αποτέλεσμα να αυξάνονται οι πιθανότητες εμφάνισης κάποιας δυσλειτουργίας του εξοπλισμού των δικτύων διανομής [264]. Η προληπτική συντήρηση του εξοπλισμού δίνει τη δυνατότητα ανίχνευσης των όποιων πιθανών βλαβών μπορεί να

εμφανιστούν σε αυτόν και επομένως επιτρέπει στο συνολικό σύστημα να λειτουργεί ομαλά και χωρίς καθυστερήσεις παροχής ενέργειας, οι οποίες θα ήταν αναπόφευκτες για την αποκατάσταση της όποιας βλάβης θα μπορούσε να προκύψει [29,31].

Στον τομέα της ενέργειας η προληπτική συντήρηση μπορεί να κατηγοριοποιηθεί σε δύο βασικούς τύπους [265]: (α) βάσει χρόνου και (β) βάσει συνθηκών. Στη συντήρηση βάσει χρόνου, ο εξοπλισμός υπόκειται σε συντήρηση σε τακτά χρονικά διαστήματα ανεξαρτήτως από την κατάστασή τους. Το πρόβλημα που παρουσιάζει η προληπτική συντήρηση βάσει χρόνου είναι ότι δεν αξιοποιεί αποτελεσματικά το χρόνο λειτουργίας του εξοπλισμού. Αντίθετα, στη συντήρηση βάσει συνθηκών, ο προγραμματισμός της συντήρησης του εξοπλισμού προκύπτει από την παρακολούθηση της κατάστασης του εξοπλισμού και του συσχετισμού της με τα πιθανά μελλοντικά σφάλματα του εξοπλισμού [216,266].

Σε παλαιότερες προσεγγίσεις προληπτικής συντήρησης, ο προγραμματισμός της προέκυπτε μετά την ανάλυση των δεδομένων που συλλέγονταν από τους διακόπτες αυτόματης επαναφοράς, οι οποίοι τοποθετούνται στις συνδέσεις των καλωδίων διανομής ενέργειας [267]. Χάρη, όμως, στην ανάπτυξη της τεχνολογίας των ΤΠΕ στα συστήματα ισχύος, ένας τεράστιος όγκος δεδομένων μπορεί να συλλεχθεί μέσω των μετρητών, των αισθητήρων, καθώς και των διάφορων άλλων μονάδων που υπάρχουν στο δίκτυο διανομής, αλλά και μέσω των υποδομών επικοινωνίας που το υποστηρίζει. Τα δεδομένα αυτά, που αφορούν τη λειτουργία του συστήματος ισχύος, τις καιρικές συνθήκες, τις διαγνώσεις των συστημάτων των ρελέ προστασίας των κυκλωμάτων ισχύος, τις διακυμάνσεις της ενέργειας που μεταφέρεται μέσω των γραμμών μεταφοράς (λόγω μεταβολών της τάσης και του ρεύματος παροχής), κλπ., υποβάλλονται σε επεξεργασία και χρησιμοποιούνται ως είσοδος σε συστήματα ταξινόμησης δεδομένων, δημιουργώντας μοντέλα σχεδιασμού των διαδικασιών προληπτικής συντήρησης του εξοπλισμού [29,31]. Στη βιβλιογραφία έχουν παρουσιαστεί διάφορες προσεγγίσεις δημιουργίας τέτοιων μοντέλων προληπτικής συντήρησης, οι οποίες βασίζονται στη χρήση μηχανών SVM [268], μακροπρόθεσμων μνημών (Long Short Term Memory – LSTM) [269], μηχανών ακραίας μάθησης (Extreme Learning Machines – EML) [270], δικτύων ANN [271], μοντέλων υβριδικών συνόλων [272], κλπ. Σε όλες αυτές τις προσεγγίσεις, βασικό μέλημα αποτελεί η συσχέτιση μεταξύ του ιστορικού των βλαβών που έχουν παρουσιαστεί στον εξοπλισμό του δικτύου διανομής στο παρελθόν και στα

χαρακτηριστικά του δικτύου που εξήχθησαν από την ανάλυση των συλλεγόμενων δεδομένων.

Η χρήση της τεχνολογίας των Big Data στην περίπτωση της προληπτικής συντήρησης θεωρείται από πολλούς μελετητές ως ιδανική, καθώς οι πληροφορίες που περιλαμβάνουν μπορούν να αξιοποιηθούν από τα παραπάνω μοντέλα, δημιουργώντας τις κατάλληλες συνθήκες μιας πιο στοχευμένης και με μεγαλύτερη ακρίβεια συντήρησης του εξοπλισμού [273].

5.5.5 Ανίχνευση τοπολογίας δικτύου

Η αξιοποίηση των πληροφοριών που μπορούν να συλλεχθούν στο πλαίσιο των σύγχρονων συστημάτων παροχής ηλεκτρικής ενέργειας, αποτελεί ένα αποτελεσματικό μέσο αντιμετώπισης των προκλήσεων που δημιουργούνται από την ενσωμάτωση των ΑΠΕ στο δίκτυο διανομής. Η συνεχόμενη μεταβολή των φορτίων σε ένα σύγχρονο δίκτυο παροχής ηλεκτρικής ενέργειας αυξάνει την ανάγκη για λήψη ολοκληρωμένων αποφάσεων με βάση τις πληροφορίες που εξάγονται από την επεξεργασία του τεράστιου όγκου δεδομένων που συλλέγονται. Η πραγματοποίηση μετρήσεων, διαδικασιών παρακολούθησης, επικοινωνίας και ελέγχου των ενεργειακών δικτύων από προηγμένες μονάδες, μετρητές και αισθητήρες, καθιστούν το σύνθετο πλέον δίκτυο πιο αντιληπτό [29].

Αυτή η συνεχόμενη μεταβολή των φορτίων που χαρακτηρίζει τα σύγχρονα συστήματα παροχής ηλεκτρικής ενέργειας δημιουργεί ένα επιπλέον ζήτημα, αυτό της ανίχνευσης της τοπολογίας του δικτύου, το οποίο αφορά τον προσδιορισμό της δομής του δικτύου διανομής, την ανίχνευση της συνδεσιμότητας φάσης των καταναλωτών και το συσχετισμό τους με τους μετασχηματιστές διανομής ισχύος [274]. Ο προσδιορισμός της συνδεσιμότητας φάσης είναι ζωτικής σημασίας για την ανάλυση του συστήματος διανομής, η οποία περιλαμβάνει την εκτίμηση του καταναμημένου δικτύου, την ανάλυση και τη βελτιστοποίηση της ενεργειακής ροής, την αναδιαμόρφωση και την αποκατάσταση του δικτύου διανομής, την εξισορρόπηση των φορτίων, κλπ. [275-279].

Σε κάποιες μελέτες, το ζήτημα της ανίχνευσης της τοπολογίας του δικτύου διανομής αντιμετωπίστηκε με προσεγγίσεις, οι οποίες βασίστηκαν στη χρήση εξειδικευμένων αισθητήρων, όπως συγχρονισμένοι φασιθέτες (synchrophasors), αισθητήρες ρεύματος γραμμής, μετρητές φάσης, κλπ. [280]. Η χρήση όμως εξειδικευμένων αισθητήρων για

κάθε καταναλωτή είναι μη πρακτική και δαπανηρή [281]. Σε άλλες προσεγγίσεις, η ανίχνευση της τοπολογίας του δικτύου πραγματοποιείται χρησιμοποιώντας τα δεδομένα που συλλέγονται από την υπάρχουσα υποδομή του ηλεκτρικού δικτύου, δηλαδή με χρήση της προηγμένης υποδομής μέτρησης (Advanced Metering Infrastructure - AMI), των συστημάτων SCADA, των συστημάτων GIS, των συστημάτων διαχείρισης των διακοπών της ηλεκτροδότησης (Outage Management Systems - OMS), κλπ. [31]. Τέλος, το ζήτημα της ανίχνευσης της τοπολογίας του δικτύου διανομής έχει αντιμετωπιστεί σε νεότερες προσεγγίσεις με χρήση τεχνικών μηχανικής μάθησης, στις οποίες δεδομένα κατάρτισης χρησιμοποιούνται για την ανεύρεση περιπτώσεων επικυρωμένης συνδεσιμότητας φάσης [282].

5.6 Διαχείριση ενεργειακής ζήτησης

Η διαχείριση της ενεργειακής ζήτησης αφορά την αποτελεσματική διαχείριση των φορτίων του καταναλωτή, μέσω της χρήσης καινοτόμων προσεγγίσεων. Τις περισσότερες φορές η διαχείριση της ενεργειακής ζήτησης επιτυγχάνεται μέσω της υλοποίησης συστημάτων πραγματικού χρόνου, η λειτουργία των οποίων βασίζεται στην επεξεργασία και ανάλυση των δεδομένων που αφορούν την παρακολούθηση της ενεργειακής κατανάλωσης και στο σχεδιασμό της βελτιστοποίησης της ενεργειακής ροής, βάσει των πληροφοριών που μπορούν να εξαχθούν από αυτήν την ανάλυση [283].

Καθώς τα σύγχρονα συστήματα παροχής ηλεκτρικής ενέργειας έχουν ως φορτία έξυπνες συσκευές και καταναλωτές ενεργειακούς πόρους, ο εξοπλισμός τους με προηγμένα συστήματα ελέγχου με επικοινωνιακές δυνατότητες, δίνει τη δυνατότητα δυναμικής διαχείρισης της ενεργειακής ζήτησης. Με τον τρόπο αυτό, τα σύγχρονα συστήματα παροχής ηλεκτρικής ενέργειας ενσωματώνουν τις συμβατικές αρχές διαχείρισης της ενέργειας, που επιτυγχάνονται με τον σωστό σχεδιασμό προγραμμάτων απόκρισης σε αυτή την ενεργειακή ζήτηση, με βάση δεδομένα που προκύπτουν από μεθόδους πρόβλεψής της. Έτσι, τα σύγχρονα συστήματα παροχής ηλεκτρικής ενέργειας είναι σε θέση να αντιμετωπίσουν ταυτόχρονα διάφορα ζητήματα, όπως η μόνιμη εξοικονόμηση ενέργειας, οι μόνιμες αυξήσεις της ενεργειακής ζήτησης, κλπ. Επομένως, στο πλαίσιο των σύγχρονων συστημάτων παροχής ηλεκτρικής ενέργειας, η δυνατότητα δυναμικής διαχείρισης της ενεργειακής ζήτησης εξαλείφει τα όποια ζητήματα αναποτελεσματικότητας προκύπτουν από τη μέθοδο της συμβατικής

διαχείρισης της ενεργειακής χρήσης, που αποτελούσε το βασικό σκεπτικό των παραδοσιακών ηλεκτρικών δικτύων [201,225].

Η διαχείριση της ενεργειακής ζήτησης μπορεί να αποτελέσει μία ακόμη βραχυπρόθεσμη εφαρμογή της τεχνολογίας των Big Data στον τομέα της ενέργειας. Κάτι τέτοιο μπορεί να επιτευχθεί μέσω επιμέρους εφαρμογών, όπως είναι η πρόβλεψη των φορτίων και της ενεργειακής ζήτησης, η απόκριση σε αυτή τη ζήτηση, η κοστολόγηση και τιμολόγηση της χρησιμοποιούμενης ενέργειας, η παρακολούθηση της ποιότητας της παρεχόμενης ενέργειας, ο διαχωρισμός των φορτίων, κλπ. [283]. Σε όλες αυτές τις εφαρμογές χρησιμοποιούνται διάφορες προηγμένες μέθοδοι και τεχνικές επεξεργασίας δεδομένων, λαμβάνοντας υπόψη την ενσωμάτωση στο δίκτυο διανομής των ΑΠΕ και των μικροδικτύων, καθώς και των ηλεκτρικών οχημάτων, αλλά και με το σκεπτικό ότι τα σύγχρονα συστήματα ηλεκτρικής ενέργειας θα πρέπει να ανταπεξέλθουν στις ενεργειακές απαιτήσεις των έξυπνων και βιώσιμων πόλεων, κτιρίων, οικιών, κλπ. [225].

5.6.1 Παρακολούθηση ποιότητας παρεχόμενης ενέργειας

Η υψηλή ποιότητα της παρεχόμενης ενέργειας ορίζεται από την εντός των προβλεπόμενων τιμών της συχνότητας και του πλάτους του ρεύματος και της τάσης, αλλά και από τη σταθερότητα των κυματομορφών τους [284]. Η ποιότητα της παρεχόμενης ενέργειας αποτελεί ένα χαρακτηριστικό που καθορίζει την απόδοση και την καλή κατάσταση των στοιχείων ενός ηλεκτρικού δικτύου, αλλά και την ακρίβεια των μετρήσεων που μπορούν να πραγματοποιηθούν στο πλαίσιο του. Με την ενσωμάτωση, όμως, στα σύγχρονα ηλεκτρικά δίκτυα, των μη γραμμικών πηγών ενέργειας και τη χρήση ενός τεράστιου αριθμού ηλεκτρονικών συσκευών ισχύος ως φορτία, δημιουργήθηκε μια αύξηση των αρμονικών που εμφανίζονται στις κυματομορφές της τάσης ή του ρεύματος της παρεχόμενης ενέργειας [285]. Η αύξηση των αρμονικών είναι σε θέση να δημιουργήσει σοβαρές διαταραχές, αλλά και καταστάσεις αστάθειας στην παρεχόμενη ενέργεια, κάτι που αποτελεί σοβαρή απειλή για την ασφαλή λειτουργία τόσο του ίδιου του ηλεκτρικού δικτύου, αλλά και των συσκευών που είναι συνδεδεμένες σε αυτό [29, 31].

Για την αντιμετώπιση αυτής της κατάστασης και των ζητημάτων της μειωμένης ποιότητας της παρεχόμενης ενέργειας από τα σύγχρονα ηλεκτρικά δίκτυα, επί του παρόντος γίνεται χρήση διάφορων διατάξεων, όπως δυναμικοί ρυθμιστές τάσης,

μετατροπείς (inverters), στατικοί συγχρονισμένοι αντισταθμιστές (Static Synchronous Compensators – STATCOM), σύνθετοι ρυθμιστές ποιότητας ισχύος (Unified Power Quality Conditioners – UPQC), τεχνικές συνεχούς παρακολούθησης της ποιότητας της παρεχόμενης ενέργειας, κλπ. [286,287]. Οι τεχνικές της συνεχούς παρακολούθησης της ποιότητας της παρεχόμενης ενέργειας περιλαμβάνουν τη χρήση συστημάτων SCADA, μονάδων μέτρησης, κλπ. ή τη χρήση προσεγγίσεων που βασίζονται σε τεχνικές τεχνητής νοημοσύνης [288].

Στη βιβλιογραφία έχουν παρουσιαστεί διάφορες προσεγγίσεις τεχνικών μηχανικής μάθησης, οι οποίες αποσκοπούν στην παρακολούθηση της εμφάνισης διαταραχών στην ποιότητα της παρεχόμενης ενέργειας. Τέτοιες τεχνικές αφορούν τη χρήση δέντρων αποφάσεων [289], δικτύων Bayes [290], k-πλησιέστερων γειτόνων [291], συνελκτικών δικτύων βαθιάς μάθησης (Convolutional Neural Networks - CNN) [292], δικτύων βαθιάς πεποίθησης (Deep Belief Networks – DBN) [293], μηχανών SVM [294], κλπ. Σε όλες αυτές τις προσεγγίσεις, η χρήση των Big Data αποσκοπεί κυρίως στην καλύτερη παρουσίαση και οπτικοποίηση των αποτελεσμάτων της παρακολούθησης των διαταραχών στην ποιότητα της παρεχόμενης ενέργειας [288].

5.6.2 Πρόβλεψη φορτίων

Όπως συμβαίνει και στην εφαρμογή της πρόβλεψης της παραγωγής ενέργειας από τις ΑΠΕ, μια ακριβής βραχυπρόθεσμη πρόβλεψη των φορτίων σε ένα ηλεκτρικό δίκτυο, αποτελεί τη βάση της ορθότερης διαχείρισης της ενέργειας και της λειτουργίας του συνολικού συστήματος. Σύμφωνα με τους Jordan & Messner (2020), η αύξηση της ακρίβειας των προβλέψεων μπορεί να φέρει πολλά οφέλη και εξοικονόμηση των επενδύσεων [295]. Με τον αναδυόμενο ενεργό ρόλο των καταναλωτών στο πλαίσιο των smart grid, η υψηλής απόδοσης δυναμική αγορά ηλεκτρικής ενέργειας βασίζεται σε μέγιστο βαθμό σε μια επίσης αποδοτική πρόβλεψη της συνολικής κατανάλωσης ηλεκτρικής ενέργειας [296].

Λόγω της σημαντικότητας που παρουσιάζει η πρόβλεψη των φορτίων ενός ηλεκτρικού δικτύου, αλλά και της αυξημένης ακρίβειας που πρέπει να παρουσιάζει, το συγκεκριμένο θέμα έχει αποτελέσει αντικείμενο μελέτης, ιδιαίτερα τα τελευταία χρόνια. Με δεδομένο ότι η κατανάλωση ηλεκτρικής ενέργειας επηρεάζεται σε κάποιο βαθμό από τις εκάστοτε καιρικές συνθήκες, οι Liu και συν. (2018) παρουσίασαν μια προσέγγιση πρόβλεψης των κατανεμημένων φορτίων ενός ηλεκτρικού δικτύου, στην

οποία το προγραμματιστικό παράδειγμα Map/Reduce χρησιμοποιείται για την κατάτμηση της περιοχής του δικτύου, σύμφωνα με τις πληροφορίες που αφορούν την πρόγνωση των καιρικών συνθηκών της συγκεκριμένης περιοχής [297].

Ένας συνδυασμός τεχνικών EML και EL, που λαμβάνουν ως είσοδο τις κυματομορφές της παρεχόμενης ενέργειας, χρησιμοποιείται για την πρόβλεψη της κατανάλωσης ηλεκτρικής ενέργειας με βάση την επιλογή συγκεκριμένων χαρακτηριστικών των κυματομορφών, παρουσιάστηκε από τους Ahmad και συν. (2017) [298], μια προσέγγιση που χρησιμοποιήθηκε επίσης από τους Lai και συν. (2020) [299].

Για να ξεπεραστεί η αστάθεια και η αβεβαιότητα των προφίλ των φορτίων, οι Shi, Xu & Li (2018) χρησιμοποίησαν ένα επαναλαμβανόμενο νευρικό δίκτυο (Recurrent Neural Network – RNN), στο οποίο πρόσθεσαν ένα επιπλέον στρώμα για την αύξηση της απόδοσης πρόβλεψης των φορτίων των καταναλωτών [300]. Η χρήση δικτύου RNN για την πρόβλεψη των φορτίων των καταναλωτών, παρουσιάστηκε και από τους Fekri και συν. (2021), οι οποίοι όμως κατάφεραν να αυξήσουν ακόμα περισσότερο την απόδοση πρόβλεψης, μέσω αυτόματης ενημέρωσης των δεδομένων της εισόδου του δικτύου RNN [301].

Κάποιοι μελετητές, θεωρώντας ότι η επίτευξη μεγάλης ακρίβειας στην πρόβλεψη του συνόλου των φορτίων μιας περιοχής είναι αρκετή δύσκολη, επέλεξαν τη μελέτη της πρόβλεψης της κατανάλωσης ενέργειας των μεμονωμένων κατοικιών. Λόγω της αυξημένης δημοτικότητας που παρουσιάζουν οι τεχνικές βαθιάς μάθησης, οι περισσότερες από αυτές τις μελέτες έχουν επικεντρωθεί στη χρήση δικτύου RNN μνήμης LSTM. Το αποτέλεσμα των μελετών έδειξε ότι η χρήση της μνήμης LSTM αυξάνει σε μέγιστο βαθμό την ακρίβεια της πρόβλεψης της κατανάλωσης ενέργειας των μεμονωμένων κατοικιών, το ποσοστό της οποίας, όμως, εξαρτάται σε μεγάλο βαθμό από τη συμβατότητα που παρουσιάζουν τα είδη των φορτίων των κατοικιών [302-305]. Όσον αφορά την οικιακή κατανάλωση ενέργειας, άλλες μελέτες ασχολήθηκαν με την πρόβλεψη της συμπεριφοράς των καταναλωτών σε πραγματικό χρόνο. Για το λόγο αυτό χρησιμοποιήθηκαν προσεγγίσεις διαδικασιών αποφάσεων με μοντέλα Markov [306].

Τέλος, κάποιες μελέτες έχουν ασχοληθεί με τη δυνατότητα αξιοποίησης των αναδυόμενων τάσεων της σύγχρονης εποχής, που αφορούν τη χρήση των μέσων κοινωνικής δικτύωσης μέσω εφαρμογών για κινητά (mobile apps), για την πρόβλεψη

των συνολικών φορτίων μιας περιοχής [307]. Για παράδειγμα, οι Cao και συν. (2020) εξέτασαν τη σχέση που μπορούν να έχουν οι δραστηριότητες των καταναλωτών στα μέσα κοινωνικής δικτύωσης στη χρήση του ηλεκτρικού δικτύου και μελέτησαν το αντίκτυπο που θα μπορούσαν να έχουν αυτές στο συνολικό φορτίο μιας περιοχής [308].

5.6.3 Δημιουργία προφίλ φορτίων – Μοντέλο Demand/Response (DR)

Η δημιουργία προφίλ των φορτίων αποτελεί έναν τρόπο περιγραφής της τυπικής συμπεριφοράς της ηλεκτρικής κατανάλωσης, η οποία συνήθως αναπαρίσταται στον τομέα του χρόνου για την πρόβλεψη των φορτίων και τη διαχείριση της ενεργειακής ζήτησης [309]. Σε παλαιότερες προσεγγίσεις, η τιμολόγηση των καταναλωτών με βάση το είδος των δραστηριοτήτων θεωρούταν ως αποτελεσματική μέθοδος διαχείρισης της παρεχόμενης ενέργειας και της ενεργειακής ζήτησης, κάτι που όμως δεν είναι σε θέση να υποδείξει την καταναλωτική τους συμπεριφορά με ολοκληρωμένο τρόπο [310]. Για το λόγο αυτό, οι ερευνητές προσπάθησαν τα τελευταία χρόνια να αναδείξουν νέους τρόπους δημιουργίας ενός προφίλ ηλεκτρικής κατανάλωσης, το οποίο θα ενισχύσει ακόμα περισσότερο τις προσπάθειες ορθότερης διαχείρισης της ενεργειακής ζήτησης [29, 31].

Ένας τρόπος δημιουργίας τέτοιων προφίλ είναι η ταξινόμηση των καταναλωτών με βάση τις καμπύλες των ηλεκτρικών φορτίων τους, χρησιμοποιώντας αλγόριθμους ομαδοποίησης, όπως μηχανών SVM [311], τεχνικών χαρτών με αυτόματη οργάνωση (Self-Organizing Maps – SOM) [312], K-means [313], κλπ. Όλες οι χρησιμοποιούμενες τεχνικές στις παραπάνω μελέτες αφορούν ανάλυση των χαρακτηριστικών των ηλεκτρικών φορτίων των καταναλωτών στο πεδίο του χρόνου [314]. Ανάλογες τεχνικές όμως έχουν χρησιμοποιηθεί και σε μελέτες, όπου η ανάλυση των χαρακτηριστικών των ηλεκτρικών φορτίων των καταναλωτών δεν γίνεται στο πεδίο του χρόνου. Για την καλύτερη κατανόηση των δεδομένων που αφορούν την ασυμβατότητα των διαφόρων φορτίων των κατοικιών, οι Haben, Singleton & Grindrod (2017) παρουσίασαν μια προσέγγιση ανάλυσης των δεδομένων αυτών, η οποία βασίζεται στη χρήση μοντέλων ομαδοποίησης πεπερασμένου μείγματος δεδομένων [315]. Σύμφωνα με τους Rajabi και συν. (2020), χάρτες SOM, ως τύπος δικτύων ANN, έχουν χρησιμοποιηθεί για τη μείωση της διάστασης των συλλεγόμενων πρωτογενών δεδομένων, που μπορούν να αξιοποιηθούν για την εξαγωγή πληροφοριών δημιουργίας

προφίλ ηλεκτρικών φορτίων. Οι συγγραφείς αναφέρουν ότι η χρήση των χαρτών SOM για ανάλυση δεδομένων στο πεδίο της συχνότητας υπερτερεί της αντίστοιχης χρήσης στο πεδίο του χρόνου, παρουσιάζοντας μεγαλύτερη ακρίβεια στην ταξινόμηση των χαρακτηριστικών των φορτίων των καταναλωτών [316]. Τέλος, οι R. Li, F. Li & Smith (2017) χρησιμοποίησαν μεθόδους ανάλυσης, όπως ο διακριτός μετασχηματισμός κυματομορφών (Discrete Wavelet Transformation – DWT) και ο διακριτός μετασχηματισμός Fourier (Discrete Fourier Transformation – DFT), στο πεδίο του φάσματος, για την αποσύνθεση των δεδομένων που λαμβάνονται από τους μετρητές των ηλεκτρικών δικτύων και την εξαγωγή των απαιτούμενων χαρακτηριστικών των φορτίων κατανάλωσης. Σύμφωνα με τους συγγραφείς ο μετασχηματισμός DWT αποδίδει καλύτερα από τον DFT σε επίπεδο ενός νοικοκυριού, ενώ ο DFT είναι πιο κατάλληλο για χρήση σε περιπτώσεις ανάλυσης φορτίων μιας ολόκληρης περιοχής [317].

Ένα από τα μεγαλύτερα πλεονεκτήματα της δημιουργίας προφίλ για τα οικιακά ηλεκτρικά φορτία και την αντίστοιχη κατανάλωσή τους, είναι η καλύτερη κατανόησή τους με τέτοιο τρόπο που να οδηγήσει σε μια δυναμική ευελιξία των καταναλώσεων. Μελέτες έχουν δείξει ότι η ομαδοποίηση των καταναλωτών βάσει παρόμοιων καταναλωτικών προφίλ, δίνει τη δυνατότητα καλύτερου σχεδιασμού στοχευμένης χρονικής μετατόπισης των φορτίων στις ώρες αιχμής και, επομένως, μείωσης της συνολικής ενεργειακής ζήτησης [318]. Σύμφωνα με τους Panda και συν. (2022), τα οικιακά ηλεκτρικά φορτία συνήθως μπορεί να χωριστούν σε τρεις βασικές κατηγορίες [319]: (α) τα σταθερά φορτία, (β) τα χρονικά ελαστικά φορτία και (γ) τα χρονικά ανελαστικά φορτία. Η διάκριση αυτή στις κατηγορίες των φορτίων μπορεί να αποτελέσει τη θεωρητική βάση για τη βέλτιστη ενεργειακή διαχείριση του μοντέλου ζήτησης/απόκρισης (Demand/Response - DR) [319]. Το μοντέλο DR χρησιμοποιείται ως τρόπο αντιμετώπισης της ενεργειακής ζήτησης και αποσκοπεί στο μετασχηματισμό του προφίλ ηλεκτρικής κατανάλωσης των καταναλωτών, χρησιμοποιώντας ως κίνητρο τις πληροφορίες που προκύπτουν μετά την ανάλυση των δεδομένων κόστους και οικολογικών επιπτώσεων που μπορεί να έχει αυτό το προφίλ [320]. Το μοντέλο DR μπορεί επίσης να χρησιμοποιηθεί για το έντονο ζήτημα που αντιμετωπίζουν τα σύγχρονα συστήματα διανομής, όσον αφορά την επίτευξη σταθεροποίησης της παρεχόμενης τάσης [321].

Το ζήτημα της χρήσης του μοντέλου DR στα σύγχρονα ηλεκτρικά δίκτυα έχει απασχολήσει εκτεταμένα την ερευνητική κοινότητα. Αρκετές μελέτες έχουν ασχοληθεί με την ανάπτυξη μηχανισμών που συνδυάζουν τεχνικές μηχανικής μάθησης και σχεδιασμού δομών δεδομένων, για τη δημιουργία συστημάτων διαχείρισης της οικιακής ενέργειας που μπορούν να καλύψουν τις ενεργειακές ανάγκες ενός σπιτιού ή ενός κτιρίου με βάση το μοντέλο DR. Στις περισσότερες από αυτές έχει δοθεί ιδιαίτερη έμφαση σε χρονικά ελαστικά φορτία, όπως είναι τα συστήματα HVAC (θέρμανσης, εξαερισμού και κλιματισμού) των σπιτιών/κτιρίων. Με την κατανόηση των καταναλωτικών προφίλ αυτών των συστημάτων, μπορεί να σχεδιαστεί ένα κατάλληλο μοντέλο DR ρύθμισής τους που να βασίζεται στην πρόγνωση των καιρικών συνθηκών [322-326].

Με την ολοένα αυξανόμενη υιοθέτηση των ηλεκτρικών οχημάτων, η εκμάθηση των μοτίβων φόρτισής τους μπορεί να αποτελέσει επίσης ένα αποφασιστικό βήμα προς την επίτευξη σταθερότητας της παρεχόμενης τάσης από τα δίκτυα διανομής ηλεκτρικής ενέργειας. Δεδομένης της αστάθειας της κατανεμημένης παραγωγής και της ισχύος φόρτισης των ηλεκτρικών οχημάτων στο επίπεδο χαμηλής τάσης, απαιτούνται νέες υπηρεσίες στην πλευρά του δικτύου διανομής για τη διατήρηση της τάσης, εντός αποδεκτών ορίων, ελαχιστοποιώντας παράλληλα το φορτίο και τη μείωση της παραγωγής. Στην περίπτωση αυτή, η χρήση του μοντέλου DR μπορεί να αποδειχθεί ιδιαίτερα χρήσιμη στην πράξη, επιτρέποντας την ενσωμάτωση περισσότερων κατανεμημένων πηγών ενέργειας και ηλεκτρικών οχημάτων στο σύστημα διανομής, χωρίς ιδιαίτερη οικονομική επιβάρυνση [327]. Για το λόγο αυτό, μεγάλο μέρος της ερευνητικής κοινότητας επικεντρώθηκε στην εύρεση αλγορίθμων DR για τον συντονισμό της φόρτισης των ηλεκτρικών οχημάτων. Ο σχεδιασμός τέτοιων αποτελεσματικών και πρακτικών αλγορίθμων DR απαιτεί την ακριβή κατανόηση της μεταβλητότητας των φορτίων των ηλεκτρικών οχημάτων και του πώς αυτή μπορεί να επηρεάσει τα οφέλη του μοντέλου DR [328]. Για παράδειγμα, αλγόριθμοι ομαδοποίησης μη εποπτευόμενης μάθησης έχουν χρησιμοποιηθεί για την εξαγωγή του μοτίβου των φορτίων φόρτισης των ηλεκτρικών οχημάτων, μέσω της ανάλυσης των δεδομένων που συλλέγονται από τις πραγματικές μετρήσεις της ενεργειακής κατανάλωσης [329-333]. Επιπλέον, στη βιβλιογραφία έχουν παρουσιαστεί μοντέλα Bayes για την εκτίμηση της μέγιστης πιθανοφάνειας (Bayesian Maximum Likelihood)

εφαρμογή του μοντέλου DR στις περιπτώσεις συγχρονισμού της φόρτισης των ηλεκτρικών οχημάτων [334-337].

5.6.4 Διαχωρισμός φορτίων

Η διαδικασία του διαχωρισμού των φορτίων (load disaggregation), γνωστή και ως μη παρεμβατική παρακολούθηση φορτίων (Non-Intrusive Load Monitoring - NILM), έχει ως στόχο τον διαχωρισμό των προφίλ του συνόλου των φορτίων, σε επίπεδο νοικοκυριού, σε μεμονωμένη κατανάλωση ενέργειας, σε επίπεδο συσκευής [338]. Σε αντίθεση με τη διαδικασία της απευθείας παρακολούθησης των συσκευών για άμεση καταγραφή της κατανάλωσής τους, η παρακολούθηση NILM από έναν μόνο έξυπνο μετρητή, που είναι εγκατεστημένος στο χώρο της οικείας, είναι ευκολότερο να γίνει αποδεκτή από τους καταναλωτές [339].

Καθώς ένα νοικοκυριό αποτελείται από διαφορετικούς τύπους ηλεκτρικών συσκευών, καθένας από τους οποίους παρουσιάζει διαφορετικές δυνατότητες συμμετοχής στο μοντέλο DR, η διαδικασία διαχωρισμού των φορτίων και η δημιουργία καταναλωτικών προφίλ σε επίπεδο συσκευής, δίνουν τη δυνατότητα καλύτερης κατανόησης της ενεργειακής συμπεριφοράς των καταναλωτών, κάτι που ενισχύει την ανάπτυξη ενεργειακά αποδοτικότερων στρατηγικών [340-342]. Στη βιβλιογραφία, η διαδικασία του διαχωρισμού των φορτίων έχει αντιμετωπιστεί ως πρόβλημα αναγνώρισης μοτίβων και οι διάφορες προσεγγίσεις που έχουν παρουσιαστεί βασίζονται στην ανίχνευση συμβάντων, δηλαδή στον εντοπισμό οποιασδήποτε μεταβολής της κατάστασης των συσκευών ως προς την κατανάλωση ενέργειας [343]. Οι αλγόριθμοι που παρουσιάζονται σε αυτές τις προσεγγίσεις, έχουν ως σκοπό, μετά τον εντοπισμό αυτής της κατάστασης, την ταξινόμηση των συσκευών με βάση τα πιο αντιπροσωπευτικά καταναλωτικά χαρακτηριστικά τους, ώστε να διαφοροποιούνται και να αναγνωρίζονται μεταξύ τους [340]. Οι προσεγγίσεις διαχωρισμού των φορτίων που έχουν παρουσιαστεί στη βιβλιογραφία μπορούν να ταξινομηθούν σε τρεις βασικές κατηγορίες [341]: (α) ειδικής ευρετικής αξιολόγησης, (β) πιθανοτικών μοντέλων και (γ) προσαρμοσμένων φίλτρων.

Οι αλγόριθμοι που βασίζονται σε τεχνικές ειδικής ευρετικής αξιολόγησης (expert heuristic evaluations) προσπαθούν να διαφοροποιήσουν τις συσκευές μέσω της χρήσης ενός συνόλου κανόνων, οι οποίοι λαμβάνουν υπόψη τους μεταβλητές, όπως η διακύμανση ή η κατανάλωση της παρεχόμενης ενέργειας [344]. Οι προσεγγίσεις στις

οποίες χρησιμοποιούνται πιθανοτικά μοντέλα (probabilistic models) έχουν ως σκοπό την απομόνωση της χρονικής σύμπτωσης των συμβάντων. Οι συγκεκριμένες προσεγγίσεις απαιτούν μοντέλα εκπαίδευσης για την προσαρμογή των μεταβλητών και τη δημιουργία στατιστικών μοντέλων, κάτι παρόμοιο με αυτό που πραγματοποιείται στην περίπτωση της μεθόδου GLR (Generalized Likelihood Ratio) [345]. Οι Puente και συν. (2020) αναφέρουν ότι η σύγκριση αυτών των δύο κατηγοριών δείχνει ότι τα αποτελέσματα των προσεγγίσεων με τα πιθανοτικά μοντέλα παρουσιάζουν μεγαλύτερη ακρίβεια και περιέχουν λιγότερες ψευδώς θετικές καταχωρήσεις κατανάλωσης [341].

Στην τρίτη κατηγορία προσεγγίσεων, των προσαρμοσμένων φίλτρων (matched filters), ο εντοπισμός του εκάστοτε τύπου συσκευής πραγματοποιείται μέσω χρήσης μοτίβων που συσχετίζονται με τις κυματομορφές του σήματος ισχύος. Σε αυτή την περίπτωση, η χρήση της τεχνολογίας των Big Data είναι ιδιαίτερης σημασίας, αφού το σύνολο των δεδομένων που πρέπει να επεξεργασθεί για την εξαγωγή χρήσιμων συμπερασμάτων απαιτείται να είναι πολύ μεγάλο [341]. Λόγω αυτής της ιδιαιτερότητας, στη βιβλιογραφία έχουν παρουσιαστεί διάφορες μελέτες που συνδυάζουν προηγμένες τεχνικές μηχανικής μάθησης και αλγόριθμους τεχνητής νοημοσύνης, όπως κρυφά μοντέλα Markov [346,347] και εξελιγμένα νευρωνικά δίκτυα, όπως BP-ANN (Back-Propagation-Artificial Neural Networks) [348] ή δίκτυα CNN [349–351].

5.6.5 Ανίχνευση ενεργειακής απάτης

Με τον όρο ενεργειακή απάτη εννοείται η περίπτωση της ρευματοκλοπής ή η περίπτωση της ύπαρξης λογιστικού σφάλματος, και αποτελεί μια από τις εξέχουσες ανησυχίες που έχουν ταλαιπωρήσει τις εταιρείες παροχής ηλεκτρικής ενέργειας για μεγάλο χρονικό διάστημα [29]. Από τις περιπτώσεις αυτές, η ρευματοκλοπή, που ορίζεται ως η πράξη δόλιας μετατροπής της μέτρησης της ηλεκτρικής κατανάλωσης, προκειμένου να μειωθεί ο λογαριασμός πληρωμής της, μέσω φυσικής παρέμβασης, όπως η παράκαμψη ή η παραβίαση του μετρητή κατανάλωσης, αποτελεί μείζον ζήτημα, το οποίο έχει απασχολήσει ιδιαίτερα την ερευνητική κοινότητα [31].

Στη βιβλιογραφία έχουν παρουσιαστεί μέχρι τώρα πολλές προσεγγίσεις που βασίζονται στη συλλογή δεδομένων, από την επεξεργασία των οποίων μπορεί να επιτευχθεί προσδιορισμός περιπτώσεων ρευματοκλοπής. Οι προσεγγίσεις αυτές μπορούν να ταξινομηθούν σε διαφορετικές κατηγορίες ανάλογα με το είδος των μετρητών που χρησιμοποιούνται και το σύνολο των δεδομένων που συλλέγουν [352]. Στις

περιπτώσεις που χρησιμοποιούνται παραδοσιακοί μετρητές ελέγχου της ηλεκτρικής κατανάλωσης, το ζήτημα της ρευματοκλοπής μπορεί να αντιμετωπιστεί μέσω μοντέλων μηχανικής μάθησης, όπως ομαδοποίηση μέσω ασαφούς λογικής [353], μηχανών SVM [354], κλπ. Σε αυτές τις μελέτες, ο εντοπισμός της ρευματοκλοπής πραγματοποιείται μέσω σύγκρισης της ετήσιας κατανάλωσης ενέργειας και της κατανάλωσης σε διάφορες χρονικές και εποχικές περιόδους [355]. Στις περιπτώσεις που χρησιμοποιούνται έξυπνοι μετρητές, το ζήτημα της ρευματοκλοπής μπορεί να αντιμετωπιστεί μέσω μοντέλων εποπτευόμενης μηχανικής μάθησης, όπως νευρωνικά δίκτυα [356], μοντέλα βαθιάς μάθησης [357], κλπ. Σε αυτές τις μελέτες, ο εντοπισμός της ρευματοκλοπής μπορεί να πραγματοποιηθεί χρησιμοποιώντας τα δεδομένα που συλλέγονται από τους έξυπνους μετρητές σε συνδυασμό με πληροφορίες που αφορούν την τοπολογία του ηλεκτρικού δικτύου, και εφαρμόζοντας προσεγγίσεις εκτίμησης της κατάστασης [358], αλλά και άλλες τεχνικές ανίχνευσης ανωμαλιών [359,360].

ΚΕΦΑΛΑΙΟ 6: ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΜΕΛΛΟΝΤΙΚΕΣ ΚΑΤΕΥΘΥΝΣΕΙΣ

Σκοπός της παρούσας διπλωματικής εργασίας, ήταν η παρουσίαση μιας όσο το δυνατόν πιο ολοκληρωμένης βιβλιογραφικής ανασκόπησης της τεχνολογίας των Big Data, με βάση τις πλέον πρόσφατες εξελίξεις σε διάφορες πτυχές του, όπως είναι οι χρησιμοποιούμενες τεχνολογίες και οι προκλήσεις που αντιμετωπίζει, αλλά και η παρουσίαση της εφαρμογής της ειδικά στον τομέα της ενέργειας. Για το λόγο αυτό, ολόκληρη η αλυσίδα αξίας του τομέα της ηλεκτρικής ενέργειας χωρίστηκε σε τέσσερα μέρη (την παραγωγή, την μεταφορά, την διανομή και την κατανάλωση της ενέργειας), με σκοπό τον καθορισμό των εφαρμογών της τεχνολογίας των Big Data σε καθένα από αυτά. Η παρουσίαση των εφαρμογών αυτών έγινε με γνώμονα τις μελέτες του ακαδημαϊκού και ερευνητικού χώρου που έχουν κάνει την εμφάνισή τους στην αντίστοιχη βιβλιογραφία της τελευταίας πενταετίας. Κατά περίπτωση, παρουσιάστηκαν οι διάφορες προσεγγίσεις και τα μοντέλα των εφαρμογών που βασίζονται στη χρήση της τεχνολογίας των Big Data.

Από την παρουσίαση των εφαρμογών της τεχνολογίας των Big Data στον τομέα της ενέργειας προκύπτει ένα βασικό συμπέρασμα, το οποίο αφορά τη δυναμικότητα που παρουσιάζει το συγκεκριμένο θέμα. Ειδικά στην περίπτωση των smart grid, η υιοθέτηση της τεχνολογίας των Big Data και η χρήση των διάφορων εφαρμογών της αποσκοπεί στη βελτίωση των λειτουργιών του, αλλά και στη δημιουργία νέων εφαρμογών μεγαλύτερης ενεργειακής εξοικονόμησης. Καθώς, η χρήση της τεχνολογίας των Big Data στον τομέα της ενέργειας δείχνει να αυξάνεται όλο και περισσότερο, αναμένεται η εμφάνιση ακόμη νεότερων ερευνών και μελετών για καλύτερη αξιοποίησή της, κάτι που φαίνεται να ενισχύει περαιτέρω τη δυναμικότητα του θέματος.

Όσον αφορά το κομμάτι της ενεργειακής παραγωγής, οι εφαρμογές των Big Data ενισχύουν την υιοθέτηση και την απόδοση των ΑΠΕ, καθώς επικεντρώνονται στη μείωση του κόστους παραγωγής, παρέχοντας καλύτερη πρόβλεψη σχετικά με την ενεργειακή ζήτηση, αλλά και καλύτερη διαχείριση της παραγόμενης ενέργειας. Επίσης, η ανάλυση των πληροφοριών που αφορούν την κατάσταση λειτουργίας των συστημάτων ΑΠΕ, όπως είναι οι ανεμογεννήτριες και τα φωτοβολταϊκά πάνελ,

ενισχύει την αποτελεσματικότητα παραγωγής ενέργειας των συγκεκριμένων συστημάτων και αυξάνουν την προώθηση της υιοθέτησής τους, μέσω μείωση του κόστους χρηματοδότησης και εγκατάστασής τους (ιδιαίτερα στην περίπτωση των φωτοβολταϊκών πάνελ που θα πρέπει να χρησιμοποιηθούν από τους ίδιους τους καταναλωτές). Όλες αυτές οι εφαρμογές δημιουργούν καλύτερες συνθήκες βελτιστοποίησης της διαχείρισης της παραγόμενης ενέργειας.

Όσον αφορά τις εφαρμογές της τεχνολογίας των Big Data στα δίκτυα μεταφοράς και διανομής της ενέργειας, οι διάφορες προσεγγίσεις που έχουν παρουσιαστεί στη βιβλιογραφία στοχεύουν στη μείωση του λειτουργικού τους κόστους. Εφαρμογές, όπως η παρακολούθηση της κατάστασης λειτουργίας του εξοπλισμού, η προληπτική συντήρησή του και η ανίχνευση και ανεύρεση βλαβών, μπορούν να χρησιμοποιηθούν για την ενδυνάμωση της λειτουργικής αποτελεσματικότητας του συγκεκριμένου εξοπλισμού, απλοποιώντας την πολυπλοκότητα και την διαφορετικότητα που παρουσιάζουν τα συλλεγόμενα σύνολα δεδομένων και δημιουργώντας εύκολα κατανοητές αντίστοιχες εργασίες του προσωπικού. Η καλύτερη παρακολούθηση και ο έλεγχος των δικτύων αυτών, επιτρέπουν επίσης την έγκαιρη πρόβλεψη πιθανών διαδοχικών διακοπών στην ηλεκτροδότηση και, στη χειρότερη περίπτωση, εκτεταμένων διακοπών ρεύματος, που σχετίζονται με έντονες καιρικές συνθήκες, μέσω βελτίωσης του προγραμματισμού προληπτικής διαχείρισης της παρεχόμενης ενέργειας.

Η σωστή διαχείριση της παρεχόμενης ενέργειας αποτελεί μια από τις σημαντικότερες προκλήσεις που θα πρέπει να αντιμετωπίσουν τα σύγχρονα δίκτυα διανομής. Η χρήση εφαρμογών της τεχνολογίας των Big Data μπορεί να μετατρέψει αυτήν την πρόκληση σε επιχειρηματικές ευκαιρίες, μέσω της υλοποίησης των διάφορων προτεινόμενων προσεγγίσεων διαχείρισης της ενέργειας των μικροδικτύων και των ΑΠΕ.

Σε επίπεδο καταναλωτών, η χρήση των Big Data, εκτός από τη βελτιστοποίηση της διαχείρισης της ενεργειακής ζήτησης, δίνει και τη δυνατότητα ανάπτυξης καλύτερων πελατειακών σχέσεων μεταξύ εταιρειών παροχής ενέργειας και καταναλωτών, δύο στοιχείων αρκετά αλληλένδετων. Αυτή η αλληλεξάρτηση μεταξύ διαχείρισης της ενεργειακής ζήτησης και πελατειακών σχέσεων, οφείλεται κυρίως στην αλλαγή των προσδοκιών των καταναλωτών από τις εταιρείες παροχής ηλεκτρικής ενέργειας και στην απόκριση που θα πρέπει να έχουν αυτές σε αυτήν την αλλαγή. Η ενεργειακή απόκριση των εταιρειών στις εκάστοτε ανάγκες των καταναλωτών, μπορεί να

επιτευχθεί μέσω χρήσης των δεδομένων που συλλέγονται από τους μετρητές κατανάλωσης, τα οποία επιτρέπουν τη δημιουργία καταναλωτικών προφίλ. Η δημιουργία τέτοιων προφίλ, δίνουν τη δυνατότητα στους καταναλωτές να γνωρίζουν την ενέργεια που καταναλώνουν και να τη διαχειρίζονται ανάλογα, και στις εταιρείες παροχής για την αποστολή σχετικών συστάσεων. Τέτοιου είδους εφαρμογών μπορούν να κλιμακωθούν και σε περιπτώσεις διαχείρισης της ενέργειας που καταναλώνεται ακόμα και σε επίπεδο κτιρίων, παρέχοντας τη δυνατότητα στους ενοίκους να αποφασίσουν για την ενεργειακή αναβάθμιση του κτιρίου που διαμένουν (π.χ. υιοθέτηση φωτοβολταϊκών πάνελ), σύμφωνα με τις εκάστοτε ανάγκες. Εκτός από την αντιμετώπιση των ζητημάτων της ορθότερης διαχείρισης της ενεργειακής ζήτησης, οι εφαρμογές της τεχνολογίας των Big Data στον τομέα της ενέργειας και σε επίπεδο καταναλωτών, δίνουν τη δυνατότητα για μια πιο ακριβή ανίχνευση και εντοπισμό των περιπτώσεων ενεργειακής απάτης και ρευματοκλοπής, ζητήματα που είναι ιδιαίτερα σημαντικά και που έχουν ταλαιπωρήσει τις εταιρείες παροχής ηλεκτρικής ενέργειας για μεγάλο χρονικό διάστημα.

Όλα τα παραπάνω, αλλά και όλα όσα αναφέρονται στις σελίδες της παρούσας πτυχιακής εργασίας, δείχνουν τη σημασία που έχει η τεχνολογία των Big Data για τον τομέα της ενέργειας, όπως επίσης και την αξιοπιστία που παρουσιάζει η χρήση των πληροφοριών, που προκύπτουν μετά την ανάλυσή τους, στη διαχείριση της ηλεκτρικής ενέργειας. Παρόλα αυτά, λόγω της δυναμικής που παρουσιάζει ο τομέας της ενέργειας, αυτή η χρήση των Big Data μπορεί να θεωρηθεί ότι ακόμα βρίσκεται σε πρώιμο στάδιο, καθώς σύμφωνα με τα συμπεράσματα των μελετών της τελευταίας πενταετίας, υπάρχει δυνατότητα ανάπτυξης ακόμα περισσότερων εφαρμογών.

Μερικές κατευθύνσεις και τάσεις για μελλοντική έρευνα είναι η επανεξέταση παραδοχών αναλυτικής μοντελοποίησης, η εκμάθηση μηχανών σε συστήματα παραγωγής ,η προγνωστική συντήρηση [364], βελτιστοποιήσεις και πλήρεις εφαρμογές σε πλατφόρμες fog [362], η εξέλιξη της υποδομής smart grid ενδέχεται να οδηγήσει σε πιο προχωρημένες τεχνολογίες επικοινωνίας και πληροφοριών, μείωση των εκπομπών διοξειδίου του άνθρακα μέσω και της παρακίνησης των καταναλωτών για αναβάθμιση σε φιλικότερες προς το περιβάλλον συσκευές [363] ,η αντιμετώπιση ζητημάτων απόρρητου [365] και τέλος η ανάπτυξη του κώδικα δεοντολογίας πρακτικής, του ελέγχου δεοντολογίας και άλλων πλαισίων διακυβέρνησης που σχετίζονται με τη διαχείριση μεγάλων δεδομένων [370].

ΒΙΒΛΙΟΓΡΑΦΙΚΕΣ ΑΝΑΦΟΡΕΣ

- [1] Botta, A., De Donato, W., Persico, V., & Pescapé, A. (2016). Integration of cloud computing and internet of things: a survey. *Future generation computer systems*, 56, 684-700.
- [2] Oussous, A., Benjelloun, F. Z., Lahcen, A. A., & Belfkih, S. (2018). Big Data technologies: A survey. *Journal of King Saud University-Computer and Information Sciences*, 30(4), 431-448.
- [3] Holst, A. (2021, July). Amount of data created, consumed, and stored 2010-2025. Statista. <https://www.statista.com/statistics/871513/worldwide-data-created/> (Πρόσβαση την 28 ΔΕΚ 2021).
- [4] Bulao, J. (2021, December). How Much Data Is Created Every Day in 2021? Techjury. <https://techjury.net/blog/how-much-data-is-created-every-day/#gref> (Πρόσβαση την 28 ΔΕΚ 2021).
- [5] Lytras, M. D., Raghavan, V., & Damiani, E. (2017). Big data and data analytics research: From metaphors to value space for collective wisdom in human decision making and smart machines. *International Journal on Semantic Web and Information Systems (IJSWIS)*, 13(1), 1-10.
- [6] Pusala, M. K., Salehi, M. A., Katukuri, J. R., Xie, Y., & Raghavan, V. (2016). Massive Data Analysis: Tasks, Tools, Applications, and Challenges. *Big Data Analytics*. Chapter 2, pp: 11-40. Springer, New Delhi.
- [7] Arfat, Y., Usman, S., Mehmood, R., & Katib, I. (2020). Big data tools, technologies, and applications: A survey. *Smart Infrastructure and Applications*. Chapter 19, pp: 453-490. Springer, Cham.
- [8] Yaqoob, I., Hashem, I. A. T., Gani, A., Mokhtar, S., Ahmed, E., Anuar, N. B., & Vasilakos, A. V. (2016). Big data: From beginning to future. *International Journal of Information Management*, 36(6), 1231-1247.
- [9] Kim, S. S., Lim, S., & Kim, S. (2021). Real-time analysis and predictability of the health functional food market using big data. *Food Science and Biotechnology*, 1-8.
- [10] Vassakis, K., Petrakis, E., & Kopanakis, I. (2018). Big data analytics: applications, prospects and challenges. In *Mobile big data* (pp. 3-20). Springer, Cham.
- [11] Lehtisalo, O. (2018). The application of digital technologies in supply chain management. Master's thesis, Degree in Business Administration, School of Business and Management, Lappeenranta University of Technology.
- [12] Qazi, R. U. R., & Sher, A. (2016). Big data applications in businesses: An overview. *The International Technology Management Review*, 6(2), 50-63.
- [13] Salomi, M., & Balamurugan, S. A. A. (2016). Need, application and characteristics of big data analytics in healthcare—a survey. *Indian Journal of Science and Technology*, 9(16), 1-5.

- [14] Brinch, M., Stentoft, J., & Jensen, J. K. (2017, January). Big data and its applications in supply chain management: Findings from a Delphi study. In Proceedings of the 50th Hawaii International Conference on System Sciences.
- [15] Mosavi, A., Lopez, A., & Varkonyi-Koczy, A. R. (2017, September). Industrial applications of big data: State of the art survey. In International Conference on Global Research and Education (pp. 225-232). Springer, Cham.
- [16] Lin, D., & Li, B. (2018, October). The Application of Big Data in Mapping and Geographic Information Technology. In 8th International Conference on Management and Computer Science (ICMCS 2018) (pp. 26-30). Atlantis Press.
- [17] Sun, A. Y., & Scanlon, B. R. (2019). How can Big Data and machine learning benefit environment and water management: a survey of methods, applications, and future directions. *Environmental Research Letters*, 14(7), 073001.
- [18] Munawar, H. S., Qayyum, S., Ullah, F., & Sepasgozar, S. (2020). Big data and its applications in smart real estate and the disaster management life cycle: A systematic analysis. *Big Data and Cognitive Computing*, 4(2), 4.
- [19] Villarejo-Ramos, Á. F., Cabrera-Sánchez, J. P., Lara-Rubio, J., & Liébana-Cabanillas, F. (2021). Predicting Big Data Adoption in Companies With an Explanatory and Predictive Model. *Frontiers in Psychology*, 12, 978.
- [20] Blank, C., McBurney, M., Morgan, M., & Seetan, R. (2021). A Survey of Big Data Techniques for Extracting Information from Social Media Data. *Adv. Sci. Technol. Eng. Syst. J. Volume 6, Issue 3*, pp: 189-204.
- [21] Rubal, S. K. (2016). Real-Time Applications of Big Data- A Survey. *International Journal of Engineering Research & Technology (IJERT)*. Volume 05, Issue 03.
- [22] Gopu, M. (2017). Big Data and Its Applications: A Survey. *Research Journal of Pharmaceutical, Biological and Chemical Sciences*. 8.
- [23] Marries, W.M. (2017). Big Data and its Applications - A Survey. *Computer Science*.
- [24] Desai, P. V. (2018). A survey on big data applications and challenges. *Second International Conference on Inventive Communication and Computational Technologies (ICICCT)*. pp: 737-740.
- [25] Janev, V., Pujić, D., Jelić, M., & Vidal, M. E. (2020). Survey on Big Data Applications. In *Knowledge Graphs and Big Data Processing* (pp. 149-164). Springer, Cham.
- [26] Brown, M., Woodhouse, S., & Sioshansi, F. (2019). Digitalization of energy. Consumer, Prosumer, Prosumer: How Service Innovations will Disrupt the Utility Business Model, 3-25.
- [27] Zhou, K., Fu, C., & Yang, S. (2016). Big data driven smart energy management: From big data to big insights. *Renewable and Sustainable Energy Reviews*, 56, 215-225.
- [28] Daki, H., El Hannani, A., Aqqal, A., Haidine, A., & Dahbi, A. (2017). Big Data management in smart grid: concepts, requirements and implementation. *Journal of Big Data*, 4(1), 1-19.

- [29] Zhang, Y., Huang, T., & Bompard, E. F. (2018). Big data analytics in smart grids: a review. *Energy informatics*, 1(1), 1-24.
- [30] Bhattarai, B. P., Paudyal, S., Luo, Y., Mohanpurkar, M., Cheung, K., Tonkoski, R., ... & Zhang, X. (2019). Big data analytics in smart grids: state-of-the-art, challenges, opportunities, and future directions. *IET Smart Grid*, 2(2), 141-154.
- [31] Syed, D., Zainab, A., Ghrayeb, A., Refaat, S. S., Abu-Rub, H., & Bouhali, O. (2020). Smart grid big data analytics: Survey of technologies, techniques, and applications. *IEEE Access*, 9, 59564-59585.
- [32] Wang, S., Sun, X., Geng, J., Han, Y., Zhang, C., & Zhang, W. (2020, October). Application and Analysis of Big Data Technology in Smart Grid. In *Journal of Physics: Conference Series* (Vol. 1639, No. 1, p. 012043). IOP Publishing.
- [33] Pham, Q. V., Liyanage, M., Deepa, N., VVSS, M., Reddy, S., Maddikunta, P. K. R., ... & Hwang, W. J. (2021). Deep Learning for Intelligent Demand Response and Smart Grids: A Comprehensive Survey. arXiv preprint arXiv:2101.08013.
- [34] Herodotou, H. (2021). Introduction to the Special Issue on Data-Intensive Computing in Smart Microgrids. *Energies*, 14(9), 2704.
- [35] IDC Corporate. (2021, March). Data Creation and Replication Will Grow at a Faster Rate than Installed Storage Capacity, According to the IDC Global DataSphere and StorageSphere Forecasts. International Data Corporation, Needham, MA. <https://www.idc.com/getdoc.jsp?containerId=prUS47560321> (Πρόσβαση την 11 ΙΑΝ 2022).
- [36] Sah, H. R., & Negi, Y. (2021). Big Data for Smart Energy. *Big Data Analytics and Intelligent Techniques for Smart Cities*. Chapter 8, pp: 149-167. CRC Press.
- [37] Hwang, K., & Chen, M. (2017). *Big Data Science and Machine Intelligence*. Big-data analytics for cloud, IoT and cognitive computing. Chapter 1, pp: 3-44. John Wiley & Sons.
- [38] Mishra, S., Tripathy, H. K., Mishra, B. K., & Sahoo, S. (2022). Usage and analysis of big data in E-health domain. *Research Anthology on Big Data Analytics, Architectures, and Applications*. Chapter 13, pp: 417-430. IGI Global.
- [39] Sadeeq, M. M., Abdulkareem, N. M., Zeebaree, S. R., Ahmed, D. M., Sami, A. S., & Zebari, R. R. (2021). IoT and Cloud computing issues, challenges and opportunities: A review. *Qubahan Academic Journal*, 1(2), 1-7.
- [40] Ceci, L. (2021, September). Hours of video uploaded to YouTube every minute 2007-2020. Statista. <https://www.statista.com/statistics/259477/hours-of-video-uploaded-to-youtube-every-minute/> (Πρόσβαση την 15 ΙΑΝ 2022).
- [41] Murazzo, M., Gómez, P., Rodríguez, N., & Medel, D. (2019, June). Database NewSQL performance evaluation for big data in the public cloud. *Conference on Cloud Computing and Big Data*. pp: 110-121. Springer, Cham.

- [42] Liu, Y., Qiu, M., Liu, C., & Guo, Z. (2016, July). Big data in ocean observation: Opportunities and challenges. *International Conference on Big Data Computing and Communications*. pp: 212-222. Springer, Cham.
- [43] Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*, 70, 263-286.
- [44] Bharti, A. K., & NehaVerma, D. K. V. (2019). A Review on Big Data Analytics Tools in Context with Scalability. *International Journal of Computer Sciences and Engineering*, 7(2), 273-277.
- [45] Wang, H., Xu, Z., Fujita, H., & Liu, S. (2016). Towards felicitous decision making: An overview on challenges and trends of Big Data. *Information Sciences*, 367, 747-765.
- [46] Favaretto, M., De Clercq, E., Schneble, C. O., & Elger, B. S. (2020). What is your definition of Big Data? Researchers' understanding of the phenomenon of the decade. *PloS one*, 15(2), e0228987.
- [47] Bhadani, A. K., & Jothimani, D. (2016). Big data: challenges, opportunities, and realities. *Effective Big Data management and opportunities for implementation*. Chapter 1, pp: 1-24. IGI Global.
- [48] Chen, M., Miao, Y., & Humar, I. (2019). *OPNET IoT Simulation*. Chapter 1, pp: 1-76. Springer Nature.
- [49] Lei, J., & Kong, L. (2020). *Fundamentals of big data in radio astronomy*. *Big Data in Astronomy*. Chapter 2, pp: 29-58. Elsevier.
- [50] Yang, L., Li, J., Elisa, N., Prickett, T., & Chao, F. (2019). Towards big data governance in cybersecurity. *Data-Enabled Discovery and Applications*, 3(1), 1-12.
- [51] Rao, A. V. S. R., & Lakshmi, R. D. (2017). A survey on challenges in integrating big data. *Proceedings of 2nd International Conference on Intelligent Computing and Applications* (pp. 571-581). Springer, Singapore.
- [52] Furht B., Villanustre F. (2016). *Introduction to Big Data*. *Big Data Technologies and Applications*. Chapter 1, pp: 3-11. Springer, Cham.
- [53] Rivera, E. (2020, May). The V's of Big Data. *Marbella International University Centre*. <https://miuc.org/vs-big-data/> (Πρόσβαση την 18 ΙΑΝ 2022).
- [54] Elankavi, R., Kalaiprasath, R., & Udayakumar, R. (2017). Data mining with big data revolution hybrid. *International Journal on Smart Sensing and Intelligent Systems*, 10(5).
- [55] Zhou, L., Pan, S., Wang, J., & Vasilakos, A. V. (2017). Machine learning on big data: Opportunities and challenges. *Neurocomputing*, 237, 350-361.
- [56] Palanisamy, V., & Thirunavukarasu, R. (2019). Implications of big data analytics in developing healthcare frameworks—A review. *Journal of King Saud University-Computer and Information Sciences*, 31(4), 415-425.
- [57] Alharthi, H. (2018). Healthcare predictive analytics: An overview with a focus on Saudi Arabia. *Journal of infection and public health*, 11(6), 749-756.

- [58] Tang, W., Zhang, K., Zhang, D., Ren, J., Zhang, Y., & Shen, X. (2019). Fog-enabled smart health: toward cooperative and secure healthcare service provision. *IEEE Communications Magazine*, 57(5), 42-48.
- [59] Seyhan, A. A., & Carini, C. (2019). Are innovation and new technologies in precision medicine paving a new era in patients centric care?. *Journal of translational medicine*, 17(1), 1-28.
- [60] Reali, G., Femminella, M., Nunzi, E., & Valocchi, D. (2018). Genomics as a service: A joint computing and networking perspective. *Computer Networks*, 145, 27-51.
- [61] Das, N., Das, L., Rautaray, S. S., & Pandey, M. (2018). Big data analytics for medical applications. *International Journal of Modern Education and Computer Science*, 11(2), 35.
- [62] Bayne, L. E. (2018). Big data in neonatal health care: big reach, big reward?. *Critical Care Nursing Clinics*, 30(4), 481-497.
- [63] Lee, C. H., & Yoon, H. J. (2017). Medical big data: promise and challenges. *Kidney research and clinical practice*, 36(1), 3.
- [64] Ristevski, B., & Chen, M. (2018). Big data analytics in medicine and healthcare. *Journal of integrative bioinformatics*, 15(3).
- [65] Istepanian, R. S., & Al-Anzi, T. (2018). m-Health 2.0: new perspectives on mobile health, machine learning and big data analytics. *Methods*, 151, 34-40.
- [66] Grant-Muller, S., Hodgson, F., Malleson, N., & Snowball, R. (2017, June). Enhancing energy, health and security policy by extracting, enriching and interfacing next generation data in the transport domain (a study on the use of big data in cross-sectoral policy development). In *2017 IEEE International Congress on Big Data (BigData Congress)* (pp. 515-518). IEEE.
- [67] Zhang, S., Jia, S., Ma, C., & Wang, Y. (2018, April). Impacts of public transportation fare reduction policy on urban public transport sharing rate based on big data analysis. In *2018 IEEE 3rd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA)* (pp. 280-284). IEEE.
- [68] Torre-Bastida, A. I., Del Ser, J., Laña, I., Ilardia, M., Bilbao, M. N., & Campos-Cordobés, S. (2018). Big Data for transportation and mobility: recent advances, trends and challenges. *IET Intelligent Transport Systems*, 12(8), 742-755.
- [69] Priyan, M. K., & Devi, G. U. (2019). A survey on internet of vehicles: applications, technologies, challenges and opportunities. *International Journal of Advanced Intelligence Paradigms*, 12(1-2), 98-119.
- [70] Mohandu, A., & Kubendiran, M. (2021). Survey on big data techniques in intelligent transportation system (its). *Materials Today: Proceedings*, 47, 8-17.
- [71] Ghofrani, F., He, Q., Goverde, R. M., & Liu, X. (2018). Recent applications of big data analytics in railway transportation systems: A survey. *Transportation Research Part C: Emerging Technologies*, 90, 226-246.
- [72] Zhu, L., Yu, F. R., Wang, Y., Ning, B., & Tang, T. (2018). Big data analytics in intelligent transportation systems: A survey. *IEEE Transactions on Intelligent Transportation Systems*, 20(1), 383-398.

- [73] Cui, Q., Gong, Z., Ni, W., Hou, Y., Chen, X., Tao, X., & Zhang, P. (2019). Stochastic online learning for mobile edge computing: Learning from changes. *IEEE Communications Magazine*, 57(3), 63-69.
- [74] Fathi, F., Abghour, N., & Ouzzif, M. (2018, August). From big data to better behavior in self-driving cars. In *Proceedings of the 2018 2nd International Conference on Cloud and Big Data Computing* (pp. 42-46).
- [75] Zohuri, B., & Moghaddam, M. (2020). From Business Intelligence to Artificial Intelligence. *Journal of Material Sciences & Manufacturing Research*. SRC/JMSMR/102 Page, 3.
- [76] Moşteanu, N. R., Faccia, A., & Cavaliere, L. P. L. (2020, August). Disaster Management, Digitalization and Financial Resources: Key factors to keep the organization ongoing. In *Proceedings of the 2020 4th International Conference on Cloud and Big Data Computing* (pp. 118-122).
- [77] Bellomarini, L., Fakhoury, D., Gottlob, G., & Sallinger, E. (2019, April). Knowledge graphs and enterprise AI: the promise of an enabling technology. In *2019 IEEE 35th International Conference on Data Engineering (ICDE)* (pp. 26-37). IEEE.
- [78] Yadav, A. K. S., & Sora, M. (2021). Fraud detection in financial statements using text mining methods: A review. In *IOP Conference Series: Materials Science and Engineering* (Vol. 1020, No. 1, p. 012012). IOP Publishing.
- [79] Hasan, A., Kalıpsız, O., & Akyokuş, S. (2017, October). Predicting financial market in big data: deep learning. In *2017 International Conference on Computer Science and Engineering (UBMK)* (pp. 510-515). IEEE.
- [80] Begenau, J., Farboodi, M., & Veldkamp, L. (2018). Big data in finance and the growth of large firms. *Journal of Monetary Economics*, 97, 71-87.
- [81] Óskarsdóttir, M., Bravo, C., Sarraute, C., Vanthienen, J., & Baesens, B. (2019). The value of big data for credit scoring: Enhancing financial inclusion using mobile phone data and social network analytics. *Applied Soft Computing*, 74, 26-39.
- [82] Hasan, M. M., Popp, J., & Oláh, J. (2020). Current landscape and influence of big data on finance. *Journal of Big Data*, 7(1), 1-17.
- [83] Singh, M., & Singh, G. (2018). Impact of social media on e-commerce. *International Journal of Engineering & Technology*, 7(2.30), 21-26.
- [84] Dwivedi, Y. K., Ismagilova, E., Hughes, D. L., Carlson, J., Filieri, R., Jacobson, J., ... & Wang, Y. (2021). Setting the future of digital and social media marketing research: Perspectives and research propositions. *International Journal of Information Management*, 59, 102168.
- [85] Persico, V., Pescapé, A., Picariello, A., & Sperlí, G. (2018). Benchmarking big data architectures for social networks data processing using public cloud platforms. *Future Generation Computer Systems*, 89, 98-109.
- [86] Ghani, N. A., Hamid, S., Hashem, I. A. T., & Ahmed, E. (2019). Social media big data analytics: A survey. *Computers in Human Behavior*, 101, 417-428.

- [87] Shchutskaya, V., & Spirina, K. (2019, October). Big Data Behind Recommender Systems. InData Labs. <https://indatalabs.com/blog/big-data-behind-recommender-systems> (Πρόσβαση την 06 FEB 2022).
- [88] Bucko, J., Kakalejčík, L., & Ferencová, M. (2018). Online shopping: Factors that affect consumer purchasing behaviour. *Cogent Business & Management*, 5(1), 1535751.
- [89] Roy, P. S., Behera, M. D., & Srivastav, S. K. (2017). Satellite remote sensing: sensors, applications and techniques. *Proceedings of the National Academy of Sciences, India Section A: Physical Sciences*, 87(4), 465-472.
- [90] Liang, S., & Wang, J. (Eds.). (2019). A systematic view of remote sensing. *Advanced remote sensing: terrestrial information extraction and applications*. Chapter1, pp: 1-57. Academic Press.
- [91] Ustin, S. L., & Middleton, E. M. (2021). Current and near-term advances in Earth observation for ecological applications. *Ecological processes*, 10(1), 1-57.
- [92] Fahey, T., Islam, M., Gardi, A., & Sabatini, R. (2021). Laser Beam Atmospheric Propagation Modelling for Aerospace LIDAR Applications. *Atmosphere*, 12(7), 918.
- [93] Ullo, S., Gallo, M., Palmieri, G., Amenta, P., Russo, M., Romano, G., ... & De Angelis, M. (2018, March). Application of wireless sensor networks to environmental monitoring for sustainable mobility. In 2018 IEEE International Conference on Environmental Engineering (EE) (pp. 1-7). IEEE.
- [94] Iaksch, J., Fernandes, E., & Borsato, M. (2021). Digitalization and Big data in smart farming—a review. *Journal of Management Analytics*, 8(2), 333-349.
- [95] Akter, S., & Wamba, S. F. (2019). Big data and disaster management: a systematic review and agenda for future research. *Annals of Operations Research*, 283(1), 939-959.
- [96] Kamilaris, A., Kartakoullis, A., & Prenafeta-Boldú, F. X. (2017). A review on the practice of big data analysis in agriculture. *Computers and Electronics in Agriculture*, 143, 23-37.
- [97] van Oorschot, N., & van Leeuwen, B. (2017). Intelligent fire risk monitor based on Linked Open Data. In ISCRAM.
- [98] Yao, M., Sohul, M., Marojevic, V., & Reed, J. H. (2019). Artificial intelligence defined 5G radio access networks. *IEEE Communications Magazine*, 57(3), 14-20.
- [99] Mijović, V., Tomašević, N., Janev, V., Stanojević, M., & Vraneš, S. (2019). Emergency management in critical infrastructures: a complex-event-processing paradigm. *Journal of Systems Science and Systems Engineering*, 28(1), 37-62.
- [100] Eberhard, K. (2021). The effects of visualization on judgment and decision-making: a systematic literature review. *Management Review Quarterly*, 1-48.
- [101] Qadir, J., Ali, A., ur Rasool, R., Zwitter, A., Sathiaseelan, A., & Crowcroft, J. (2016). Crisis analytics: big data-driven crisis response. *Journal of International Humanitarian Action*, 1(1), 1-21.

- [102] Haidine, A., Salmam, F. Z., Aqqal, A., & Dahbi, A. (2021). Artificial intelligence and machine learning in 5G and beyond: a survey and perspectives. *Moving Broadband Mobile Communications Forward: Intelligent Technologies for 5G and Beyond*, Chapter 3, pp: 47-68. BoD – Books on Demand.
- [103] Zhao, M., Zeng, Q., Chang, M., Tong, Q., & Su, J. (2021). A Prediction Model of Customer Churn considering Customer Value: An Empirical Research of Telecom Industry in China. *Discrete Dynamics in Nature and Society*, 2021.
- [104] Kastouni, M. Z., & Lahcen, A. A. (2020). Big data analytics in telecommunications: Governance, architecture and use cases. *Journal of King Saud University-Computer and Information Sciences*.
- [105] Zahid, H., Mahmood, T., Morshed, A., & Sellis, T. (2019). Big data analytics in telecommunications: literature review and architecture recommendations. *IEEE/CAA Journal of Automatica Sinica*, 7(1), 18-38.
- [106] Xu, L., Shao, G., Cao, Y., Yang, H., Sun, C., Zhang, T., ... & He, X. (2019, October). Research on telecom big data platform of LTE/5G mobile networks. In *2019 IEEE International Conferences on Ubiquitous Computing & Communications (IUCC) and Data Science and Computational Intelligence (DSCI) and Smart Computing, Networking and Services (SmartCNS)* (pp. 756-761). IEEE.
- [107] Amin, A., Al-Obeidat, F., Shah, B., Adnan, A., Loo, J., & Anwar, S. (2019). Customer churn prediction in telecommunication industry using data certainty. *Journal of Business Research*, 94, 290-301.
- [108] Rocha-Jácome, C., Carvajal, R. G., Chavero, F. M., Guevara-Cabezas, E., & Hidalgo Fort, E. (2022). Industry 4.0: A Proposal of Paradigm Organization Schemes from a Systematic Literature Review. *Sensors*, 22(1), 66.
- [109] Sun, S., Zheng, X., Villalba-Díez, J., & Ordieres-Meré, J. (2020). Data handling in industry 4.0: Interoperability based on distributed ledger technology. *Sensors*, 20(11), 3046.
- [110] Gaurav, J., & Kongar, E. (2021). Value creation via accelerated digital transformation. *IEEE Engineering Management Review*, 49(2), 63-72.
- [111] Vater, J., Harscheidt, L., & Knoll, A. (2019, March). Smart manufacturing with prescriptive analytics. In *2019 8th International Conference on Industrial Technology and Management (ICITM)* (pp. 224-228). IEEE.
- [112] Zeng, J., & Glaister, K. W. (2018). Value creation from big data: Looking inside the black box. *Strategic Organization*, 16(2), 105-140.
- [113] Urbinati, A., Bogers, M., Chiesa, V., & Frattini, F. (2019). Creating and capturing value from Big Data: A multiple-case study analysis of provider companies. *Technovation*, 84, 21-36.
- [114] Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big data analytics and firm performance: Findings from a mixed-method approach. *Journal of Business Research*, 98, 261-276.

- [115] Gad-Elrab, A. A. (2021). Modern Business Intelligence: Big Data Analytics and Artificial Intelligence for Creating the Data-Driven Value. *E-Business: Higher Education and Intelligence Applications*, 135.
- [116] Ajah, I. A., & Nweke, H. F. (2019). Big data and business analytics: Trends, platforms, success factors and applications. *Big Data and Cognitive Computing*, 3(2), 32.
- [117] Lee, J. R., & Holt, T. J. (2021). Big Data, Cyber Security and Liberty. *The Pre-Crime Society: Crime, Culture and Control in the Ultramodern Age*. Chapter 18, pp: 409-432. Policy Press.
- [118] Pyne, S., Rao, B. P., & Rao, S. B. (Eds.). (2016). *Big data analytics: Methods and applications*. Springer.
- [119] Davis, K., Pinto, J., & Di Maddaloni, F. (2020). Significance: the need for better benefits realisation in megaprojects. *Routledge Handbook of Planning and Management of Global Strategic Infrastructure Projects*. Chapter 2, pp: 19-46. Routledge.
- [120] Salazar, L. A. L. (2017). The resource-based view and the concept of value: The role of emergence in value creation. *Mercados y Negocios (2594-0163 en línea; 1665-7039 en impreso)*, (35), 27-46.
- [121] Raguseo, E. (2018). Big data technologies: An empirical investigation on their adoption, benefits and risks for companies. *International Journal of Information Management*, 38(1), 187-195.
- [122] Faroukhi, A. Z., El Alaoui, I., Gahi, Y., & Amine, A. (2020). Big data monetization throughout Big Data Value Chain: a comprehensive review. *Journal of Big Data*, 7(1), 1-22.
- [123] Jurevicius, O. (2021, November). Value Chain Analysis. *Strategic Management Insight*. <https://strategicmanagementinsight.com/tools/value-chain-analysis/> (Πρόσβαση την 24 ΙΑΝ 2022).
- [124] Faroukhi, A. Z., El Alaoui, I., Gahi, Y., & Amine, A. (2020). An Adaptable Big Data Value Chain Framework for End-to-End Big Data Monetization. *Big Data and Cognitive Computing*, 4(4), 34.
- [125] Saleh, S. H., Ismail, R., Ibrahim, Z., & Hussin, N. (2018). Issues, challenges and solutions of big data in information management: an overview. *Int. J. Acad. Res. Bus. Soc. Sci*, 8, 12.
- [126] Draxl, C., & Scheffler, M. (2019). Big-Data-Driven Materials Science and its FAIR Data Infrastructure. *arXiv preprint arXiv:1904.05859*.
- [127] Brady, H. E. (2019). The challenge of big data and data science. *Annual Review of Political Science*, 22, 297-323.
- [128] El Arass, M., & Souissi, N. (2018, October). Data lifecycle: From big data to smartdata. In *2018 IEEE 5th international congress on information science and technology (CiSt)* (pp. 80-87). IEEE.
- [129] Gaurav, D., Yadav, J. K. P. S., Kaliyar, R. K., & Goyal, A. (2018, October). An outline on big data and big data analytics. In *2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)* (pp. 74-79). IEEE..

- [130] Darvazeh, S. S., Vanani, I. R., & Musolu, F. M. (2020). Big data analytics and its applications in supply chain management. *New Trends in the Use of Artificial Intelligence for the Industry 4.0*. Chapter 10, pp: 175-193. IntechOpen.
- [131] Waterford Technologies. (2021, July). Big Data Statistics & Facts- Are you in Control? <https://waterfordtechnologies.com/big-data-interesting-facts/> (Πρόσβαση την 01 FEB 2022).
- [132] von See, A. (2021, October). Number of IoT connected devices worldwide 2019-2030. Statista. <https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide/> (Πρόσβαση την 01 FEB 2022).
- [133] Gil, D., Johnsson, M., Mora, H., & Szymański, J. (2019). Review of the complexity of managing big data of the Internet of Things. *Complexity*, 2019.
- [134] Rao, A., Dallmeier-Tiessen, S., Lassila-Perini, K., McCauley, T., & Šimko, T. (2019). Early Experience with Open Data from CERN's Large Hadron Collider. In *Digital Innovation: Harnessing the Value of Open Data* (pp. 227-245).
- [135] Sharma, N., & Shamkuwar, M. (2019). Big Data Analysis in Cloud and Machine Learning. *Big Data Processing Using Spark in Cloud*. Chapter 3, pp: 51-85. Springer, Singapore.
- [136] Javaid, M., Haleem, A., Rab, S., Singh, R. P., & Suman, R. (2021). Sensors for daily life: a review. *Sensors International*, 2, 100121.
- [137] Genovese, L. M., & Geraci, F. (2018). Web Crawling and Processing with Limited Resources for Business Intelligence and Analytics Applications. *J. Softw.*, 13(5), 300-316.
- [138] Manikandan, G., & Abirami, S. (2017). Big Data Layers and Analytics: A Survey. *Computer Communication, Networking and Internet Security. Lecture Notes in Networks and Systems*, vol 5. pp: 383-393. Springer, Singapore.
- [139] Jamshed, H., Khan, M. S. A., Khurram, M., Inayatullah, S., & Athar, S. (2019). Data Preprocessing: A preliminary step for web data mining. *3C Tecnología. Glosas de innovación aplicadas a la pyme. Edición Especial*, Mayo 2019, pp. 206–221.
- [140] Karcher, S., Willighagen, E. L., Rumble, J., Ehrhart, F., Evelo, C. T., Fritts, M., ... & Hendren, C. O. (2018). Integration among databases and data sets to support productive nanotechnology: Challenges and recommendations. *NanoImpact*, 9, 85-101.
- [141] Ilyas, I. F., & Chu, X. (2019). Data cleaning. Morgan & Claypool. Chapter 1, pp: 1-10. Morgan & Claypool.
- [142] Kiranlar, O. (2017, December). Big Data Applications in the Energy Sector - A Review of the Current Status. Master of Science Degree in Energy Engineering and Management. Tecnico Lisboa.
- [143] Potdar, V., Chandan, A., Batool, S., & Patel, N. (2018). Big energy data management for smart grids—Issues, challenges and recent developments. *Smart Cities: Development and Governance Frameworks*. Chapter 8, pp: 177-205. Springer, Cham.

- [144] Khan, S., Liu, X., Ali, S. A., & Alam, M. (2019). Storage solutions for big data systems: A qualitative study and comparison. arXiv preprint arXiv:1904.11498.
- [145] Roozbeh, A., Soares, J., Maguire, G. Q., Wuhib, F., Padala, C., Mahloo, M., ... & Kostić, D. (2018). Software-defined “hardware” infrastructures: A survey on enabling technologies and open research directions. *IEEE Communications Surveys & Tutorials*, 20(3), 2454-2485.
- [146] Coughlin, T., Hoyt, R., & Handy, J. (2017). Digital storage and memory technology (part 1). *IEEE discussing developments in digital storage technology*, November 2017.
- [147] Ibrahim, S., K., & Abdulhussien, S., A. (2020). Improved storage area network method for backup approach. *Indonesian Journal of Electrical Engineering and Computer Science*, Vol.17, No.3, pp: 1493-1498.
- [148] Raj, P. (2018). A detailed analysis of nosql and newsql databases for bigdata analytics and distributed computing. *Advances in Computers*. Vol. 109, pp: 1-48. Elsevier.
- [149] Rasheed, Y., Qutqut, M., & Almasalha, F. (2019). Overview of the current status of NoSQL database. *Int. J. Comput. Sci. Netw. Secur*, 19(4), 47-53.
- [150] Gueidi, A., Gharsellaoui, H., & Ahmed, S. B. (2021). Towards Unified Modeling for NoSQL Solution Based on Mapping Approach. *Procedia Computer Science*, 192, 3637-3646.
- [151] Presser, D., Siqueira, F., Rodrigues, L., & Romano, P. (2020, July). Edgescaler: effective elastic scaling for graph stream processing systems. In *Proceedings of the 14th ACM International Conference on Distributed and Event-based Systems* (pp. 39-50).
- [152] Stobierski, T. (2019, August). The Advantages of Data-Driven Decision-Making. *Harvard Business School Online*. <https://online.hbs.edu/blog/post/data-driven-decision-making> (Πρόσβαση την 03 FEB 2022).
- [153] Sedkaoui, S., & Khelfaoui, M. (2019). Building an Analytics Culture to Boost a Data-Driven Entrepreneur's Business Model. *Managerial Perspectives on Intelligent Big Data Analytics*. Chapter 14, pp: 260-291. IGI Global.
- [154] Jeyaraj, R., Pugalendhi, G., & Paul, A. (2020). *Big Data with Hadoop MapReduce: A Classroom Approach*. Chapter 7. Apple Academic Press.
- [155] Painuli, D., Mishra, D., Bhardwaj, S., & Aggarwal, M. (2021). Forecast and prediction of COVID-19 using machine learning. *Data Science for COVID-19*. Chapter 20, pp: 381-397. Academic Press.
- [156] Hoi, S. C., Sahoo, D., Lu, J., & Zhao, P. (2021). Online learning: A comprehensive survey. *Neurocomputing*, 459, 249-289.
- [157] Jassim, M. A., & Abdulwahid, S. N. (2021, March). Data Mining preparation: Process, Techniques and Major Issues in Data Analysis. In *IOP Conference Series: Materials Science and Engineering* (Vol. 1090, No. 1, p. 012053). IOP Publishing.

- [158] Kezunovic, M., Pinson, P., Obradovic, Z., Grijalva, S., Hong, T., & Bessa, R. (2020). Big data analytics for future electricity grids. *Electric Power Systems Research*, 189, 106788.
- [159] Wu, T., Chung, C. Y., & Kamwa, I. (2017). A fast state estimator for systems including limited number of PMUs. *IEEE Transactions on Power Systems*, 32(6), 4329-4339.
- [160] Srivastava, N. (2017). *Deep Learning Models for Unsupervised and Transfer Learning*. Doctoral dissertation. University of Toronto, Canada.
- [161] Khan, S. S., & Wei, J. (2018, November). Real-time power outage detection system using social sensing and neural networks. In *2018 IEEE global conference on signal and information processing (GlobalSIP)* (pp. 927-931). IEEE.
- [162] Cavalcante, L., Bessa, R. J., Reis, M., & Browell, J. (2017). LASSO vector autoregression structures for very short-term wind power forecasting. *Wind Energy*, 20(4), 657-675.
- [163] Zhu, Q., Chen, J., Shi, D., Zhu, L., Bai, X., Duan, X., & Liu, Y. (2019). Learning temporal and spatial correlations jointly: A unified framework for wind speed prediction. *IEEE Transactions on Sustainable Energy*, 11(1), 509-523.
- [164] Kodaira, D., Park, J., Kim, S. Y., Han, S., & Han, S. (2019). Impedance estimation with an enhanced particle swarm optimization for low-voltage distribution networks. *Energies*, 12(6), 1167.
- [165] Sweeney, C., Bessa, R. J., Browell, J., & Pinson, P. (2020). The future of forecasting for renewable energy. *Wiley Interdisciplinary Reviews: Energy and Environment*, 9(2), e365.
- [166] Sakthivel, T., & Saravanan, V. (2019). A survey on internet of things based water and energy management open research issues. *Int. J. Sci. Res. Eng. Trends*, 5(6), 2139-2142.
- [167] Junaidi, N., & Shaaban, M. (2018, August). Big Data Applications in Electric Energy Systems. In *2018 international conference on computational approach in smart systems design and applications (ICASSDA)* (pp. 1-5). IEEE.
- [168] Ourahou, M., Ayrir, W., Hassouni, B. E., & Haddi, A. (2020). Review on smart grid control and reliability in presence of renewable energies: Challenges and prospects. *Mathematics and computers in simulation*, 167, 19-31.
- [169] Tu, C., He, X., Shuai, Z., & Jiang, F. (2017). Big data issues in smart grid—A review. *Renewable and Sustainable Energy Reviews*, 79, 1099-1107.
- [170] Zacheilas, N. (2018). *Deterministic, elastic and real-time processing in the big data era*. Doctoral dissertation. Department of Informatics. Athens University of Economics and Business. Athens, Greece.
- [171] Hong, T. (2018). Big data analytics: making the smart grid smarter [Guest Editorial]. *IEEE Power and Energy Magazine*, 16(3), 12-16.
- [172] Akerkar, R., & Hong, M. (2021). *Big Data in Electric Power Industry. Opportunities and Challenges for Sogn og Fjordane region*. Vestlandsforskning. Electric hydropower collage, Sogndal.

- [173] Liu, Z., & Zhang, A. (2020). A survey on sampling and profiling over big data (technical report). arXiv preprint arXiv:2005.05079.
- [174] Ullah, M., Narayanan, A., Wolff, A., & Nardelli, P. (2021, November). Smart Grid Information Processes Using IoT and Big Data with Cloud and Edge Computing. In 2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO) (pp. 956-961). IEEE.
- [175] Dhara, S., Shrivastav, A. K., & Sadhu, P. K. (2021). Smart Grid Modernization: Opportunities and Challenges. *Electric Grid Modernization*. IntechOpen.
- [176] Reka, S. S., & Dragicevic, T. (2018). Future effectual role of energy delivery: A comprehensive review of Internet of Things and smart grid. *Renewable and Sustainable Energy Reviews*, 91, 90-108.
- [177] Al-Badi, A. H., Ahshan, R., Hosseinzadeh, N., Ghorbani, R., & Hossain, E. (2020). Survey of smart grid concepts and technological demonstrations worldwide emphasizing on the Oman perspective. *Applied system innovation*, 3(1), 5.
- [178] Caputo, F., Buhnova, B., & Wallezky, L. (2018). Investigating the role of smartness for sustainability: insights from the Smart Grid domain. *Sustainability Science*, 13(5), 1299-1309.
- [179] Ullah, M., & Smolander, K. (2019, May). Highlighting the key factors of an IoT platform. In 2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO) (pp. 901-906). IEEE.
- [180] Marcu, I., Suci, G., Bălăceanu, C., Vulpe, A., & Drăgulinescu, A. M. (2020). Arrowhead technology for digitalization and automation solution: smart cities and smart agriculture. *Sensors*, 20(5), 1464.
- [181] Rastogi, S. (2021). *Cloud Computing Simplified: Explore Application of Cloud, Cloud Deployment Models, Service Models and Mobile Cloud Computing*. Chapter 1. BPB Publications.
- [182] Srinivas, J., Das, A. K., & Rodrigues, J. J. (2020). 2PBDC: privacy-preserving bigdata collection in cloud environment. *The Journal of Supercomputing*, 76(7), 4772-4801.
- [183] Mukundha, C., & Vidyamadhuri, K. (2017). Cloud computing models: A survey. *Adv. Comput. Sci. Technol.*, 10(5), 747-761.
- [184] Zhou, K., & Yang, S. (2018). Smart Energy Management. *Comprehensive Energy Systems*. Volume 5, Chapter 11, pp: 423-456. Elsevier.
- [185] Taherizadeh, S., Jones, A. C., Taylor, I., Zhao, Z., & Stankovski, V. (2018). Monitoring self-adaptive applications within edge computing frameworks: A state-of-the-art review. *Journal of Systems and Software*, 136, 19-38.
- [186] Sulieman, N. A., Ricciardi Celsi, L., Li, W., Zomaya, A., & Villari, M. (2022). Edge-Oriented Computing: A Survey on Research and Use Cases. *Energies*, 15(2), 452.

- [187] Akrami, A., Doostizadeh, M., & Aminifar, F. (2019). Power system flexibility: an overview of emergence to evolution. *Journal of Modern Power Systems and Clean Energy*, 7(5), 987-1007.
- [188] Iweh, C. D., Gyamfi, S., Tanyi, E., & Effah-Donyina, E. (2021). Distributed Generation and Renewable Energy Integration into the Grid: Prerequisites, Push Factors, Practical Options, Issues and Merits. *Energies*, 14(17), 5375.
- [189] Khetrupal, P. (2020). Distributed Generation: A Critical Review of Technologies, Grid Integration Issues, Growth Drivers and Potential Benefits. *International Journal of Renewable Energy Development*, 9(2).
- [190] Norshahrani, M., Mokhlis, H., Bakar, A., Halim, A., Jamian, J. J., & Sukumar, S. (2017). Progress on protection strategies to mitigate the impact of renewable distributed generation on distribution systems. *Energies*, 10(11), 1864.
- [191] Das, P., Mathuria, P., Bhakar, R., Mathur, J., Kanudia, A., & Singh, A. (2020). Flexibility requirement for large-scale renewable energy integration in Indian power system: Technology, policy and modeling options. *Energy Strategy Reviews*, 29, 100482.
- [192] Wei, Q., Liu, D., Lewis, F. L., Liu, Y., & Zhang, J. (2017). Mixed iterative adaptive dynamic programming for optimal battery energy control in smart residential microgrids. *IEEE Transactions on Industrial Electronics*, 64(5), 4110-4120.
- [193] Rahmani-Andebili, M. (2017). Stochastic, adaptive, and dynamic control of energy storage systems integrated with renewable energy sources for power loss minimization. *Renewable Energy*, 113, 1462-1471.
- [194] Mirakhorli, A., & Dong, B. (2018). Model predictive control for building loads connected with a residential distribution grid. *Applied energy*, 230, 627-642.
- [195] Wu, N., & Wang, H. (2018). Deep learning adaptive dynamic programming for real time energy management and control strategy of micro-grid. *Journal of cleaner production*, 204, 1169-1177.
- [196] Ponnusamy, V. K., Kasinathan, P., Madurai Elavarasan, R., Ramanathan, V., Anandan, R. K., Subramaniam, U., ... & Hossain, E. (2021). A Comprehensive Review on Sustainable Aspects of Big Data Analytics for the Smart Grid. *Sustainability*, 13(23), 13322.
- [197] Amaro, N., & Pina, J. M. (2017). Big Data in Power Systems. *IEEE, International Conference on Engineering, Technology and Innovation (ICE/ITMC)*, Funchal, Portugal.
- [198] Jiang, A., Yuan, H., Li, D., & Tian, J. (2019). Key technologies of ubiquitous power Internet of Things-aided smart grid. *Journal of Renewable and Sustainable Energy*, 11(6), 062702.
- [199] Ioannou, A., Fuzuli, G., Brennan, F., Yudha, S. W., & Angus, A. (2019). Multi-stage stochastic optimization framework for power generation system planning integrating hybrid uncertainty modelling. *Energy Economics*, 80, 760-776.
- [200] Bhattacharjee, K., & Patel, N. (2020). A comparative study of Economic Load Dispatch with complex non-linear constraints using Salp Swarm Algorithm. *Scientia Iranica*.

- [201] Şen, O. (2020). A Comprehensive Review of Big Data Applications for Power Systems Doctoral dissertation, Ankara Yıldırım Beyazıt Üniversitesi Fen Bilimleri Enstitüsü.
- [202] Perçuku, A., Minkovska, D., & Stoyanova, L. (2017). Modeling and processing big data of power transmission grid substation using neo4j. *Procedia computer science*, 113, 9-16.
- [203] Sobbouhi, A. R., & Vahedi, A. (2021). Transient stability prediction of power system; a review on methods, classification and considerations. *Electric Power Systems Research*, 190, 106853.
- [204] Suvorov, A. A., Diab, A. A. Z., Gusev, A. S., Andreev, M. V., Ruban, N. Y., Askarov, A. B., ... & Al-Sumaiti, A. S. (2020). Comprehensive validation of transient stability calculations in electric power systems and hardware-software tool for its implementation. *IEEE Access*, 8, 136071-136091.
- [205] Zhang, S., Zhu, Z., & Li, Y. (2021). A critical review of data-driven transient stability assessment of power systems: principles, prospects and challenges. *Energies*, 14(21), 7238.
- [206] Jozs, C., Molzahn, D. K., Tacchi, M., & Sojoudi, S. (2019, February). Transient stability analysis of power systems via occupation measures. In 2019 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT) (pp. 1-5). IEEE.
- [207] Jafarzadeh, S., Genc, I., & Nehorai, A. (2021). Real-time transient stability prediction and coherency identification in power systems using Koopman mode analysis. *Electric Power Systems Research*, 201, 107565.
- [208] Cepeda, J., Salazar, P., Echeverría, D., & Arcos, H. (2018). Implementation of the Single Machine Equivalent (SIME) Method for Transient Stability Assessment in DIGSILENT PowerFactory. In *Advanced Smart Grid Functionalities Based on PowerFactory* (pp. 319-353). Springer, Cham.
- [209] Huang, T., & Wang, J. (2019). A practical method of transient stability analysis of stochastic power systems based on EEAC. *International Journal of Electrical Power & Energy Systems*, 107, 167-176.
- [210] Shereen, E. (2021). Security of Time Synchronization for PMU-based Power System State Estimation: Vulnerabilities and Countermeasures. Doctoral dissertation, KTH Royal Institute of Technology.
- [211] Rivera, I. L. O., Vittal, V., Heydt, G. T., Fuerte-Esquivel, C. R., & Angeles-Camacho, C. (2018). A dynamic state estimator based control for power system damping. *IEEE Transactions on Power Systems*, 33(6), 6839-6848.
- [212] James, J. Q., Hill, D. J., Lam, A. Y., Gu, J., & Li, V. O. (2017). Intelligent time-adaptive transient stability assessment system. *IEEE Transactions on Power Systems*, 33(1), 1049-1058.
- [213] Shi, Z., Yao, W., Zeng, L., Wen, J., Fang, J., Ai, X., & Wen, J. (2020). Convolutional neural network-based power system transient stability assessment and instability mode prediction. *Applied Energy*, 263, 114586.

- [214] Huang, J., Guan, L., Su, Y., Yao, H., Guo, M., & Zhong, Z. (2020). Recurrent graph convolutional network-based multi-task transient stability assessment framework in power system. *IEEE Access*, 8, 93283-93296.
- [215] Luo, Y., Lu, C., Zhu, L., & Song, J. (2021). Data-driven short-term voltage stability assessment based on spatial-temporal graph convolutional network. *International Journal of Electrical Power & Energy Systems*, 130, 106753.
- [216] Hu, W., Lu, Z., Wu, S., Zhang, W., Dong, Y., Yu, R., & Liu, B. (2019). Real-time transient stability assessment in power system based on improved SVM. *Journal of Modern Power Systems and Clean Energy*, 7(1), 26-37.
- [217] Wang, Q., Pang, C., & Alnami, H. (2021, February). Transient stability assessment of a power system using multi-layer SVM method. In *2021 IEEE Texas Power and Energy Conference (TPEC)* (pp. 1-5). IEEE.
- [218] Liu, X., Zhang, X., Chen, L., Xu, F., & Feng, C. (2020). Data-driven transient stability assessment model considering network topology changes via mahalanobis kernel regression and ensemble learning. *Journal of Modern Power Systems and Clean Energy*, 8(6), 1080-1091.
- [219] Sarajcev, P., Kunac, A., Petrovic, G., & Despalatovic, M. (2021). Power system transient stability assessment using stacked autoencoder and voting ensemble. *Energies*, 14(11), 3148.
- [220] Zhu, Q., Dang, J., Chen, J., Xu, Y. P., Li, Y. H., & Duan, X. Z. (2018). A method for power system transient stability assessment based on deep belief networks. *Proc. CSEE*, 38(3), 735-743.
- [221] Yin, X., & Liu, Y. (2018, October). Deep learning based feature reduction for power system transient stability assessment. In *TENCON 2018-2018 IEEE Region 10 Conference* (pp. 2308-2312). IEEE.
- [222] Azman, S. K., Isbeih, Y. J., El Moursi, M. S., & Elbassioni, K. (2020). A unified online deep learning prediction model for small signal and transient stability. *IEEE Transactions on Power Systems*, 35(6), 4585-4598.
- [223] Siqueira de Carvalho, R. (2019). Integrating big data analytics and cybersecurity for power distribution networks with distributed energy resources. Doctoral dissertation, Colorado School of Mines.
- [224] Moussa, B., Akaber, P., Debbabi, M., & Assi, C. (2017). Critical links identification for selective outages in interdependent power-communication networks. *IEEE Transactions on Industrial Informatics*, 14(2), 472-483.
- [225] Şen, O., & Bayındır, K. Ç. (2019). Big data usage in electrical distribution systems: A review. *International Journal of Applied Business and Management Studies*, 4(2), 2548-0448.
- [226] Joseph, S., & Jasmin, E. A. (2017). Big Data Analytics for Distribution System Monitoring in Smart Grid. *International Journal of Smart Home*, 11(5), 21-32.
- [227] Koshy, S., Rahul, S., Sunitha, R., & Cheriyan, E. P. (2021). Smart grid-based big data analytics using machine learning and artificial intelligence: a survey. *Artificial Intelligence and Internet of Things for Renewable Energy Systems*. Chapter 8, pp: 241-277. The Series: Frontiers In Computational Intelligence.

- [228] Murty, V. V. S. N., & Kumar, A. (2020). Multi-objective energy management in microgrids with hybrid energy sources and battery energy storage systems. *Protection and Control of Modern Power Systems*, 5(1), 1-20.
- [229] Tabar, V. S., Jirdehi, M. A., & Hemmati, R. (2017). Energy management in microgrid based on the multi objective stochastic programming incorporating portable renewable energy resource as demand response option. *Energy*, 118, 827–839.
- [230] Hamidi, A., Nazarpour, D., & Golshannavaz, S. (2018). Multiobjective scheduling of microgrids to harvest higher photovoltaic energy. *IEEE Transactions on Industrial Informatics*, 14(1), 47–57.
- [231] Kou, P., Liang, D., & Gao, L. (2018). Stochastic energy scheduling in microgrids considering the uncertainties in both supply and demand. *IEEE Systems Journal*, 12(3), 2589–2600.
- [232] Li, X., Deb, K., & Fang, Y. (2017). A derived heuristics based multi-objective optimization procedure for micro-grid scheduling. *Engineering Optimization*, 49(6), 1078–1096.
- [233] Liu, J., Chen, H., Zhang, W., et al. (2017). Energy management problems under uncertainties for grid-connected microgrids: A chance constrained programming approach. *IEEE Transactions on Smart Grid*, 8(6), 2585–2596.
- [234] Cominesi, S. R., Farina, M., Giulioni, L., et al. (2018). A two-layer stochastic model predictive control scheme for microgrids. *IEEE Transactions on Control Systems Technology*, 26(1), 1–13.
- [235] Guo, Y., & Zhao, C. (2018). Islanding-aware robust energy management for microgrids. *IEEE Transactions on Smart Grid*, 9(2), 1301–1309.
- [236] Hasanuzzaman, M., & Kumar, L. (2020). Energy supply. *Energy for Sustainable Development*. Chapter 4, pp: 89-104. Academic Press.
- [237] Rahman, M. M., Shakeri, M., Tiong, S. K., Khatun, F., Amin, N., Pasupuleti, J., & Hasan, M. K. (2021). Prospective methodologies in hybrid renewable energy systems for energy prediction using artificial neural networks. *Sustainability*, 13(4), 2393.
- [238] Hocaoglu, F. O., & Serttas, F. (2017). A novel hybrid (Mycielski-Markov) model for hourly solar radiation forecasting. *Renewable Energy*, 108, 635-643.
- [239] Samalot, A., Astitha, M., Yang, J., & Galanis, G. (2019). Combined kalman filter and universal kriging to improve storm wind speed predictions for the northeastern United States. *Weather and Forecasting*, 34(3), 587-601.
- [240] Singh, S. N., & Mohapatra, A. (2019). Repeated wavelet transform based ARIMA model for very short-term wind speed forecasting. *Renewable energy*, 136, 758-768.
- [241] Sangrody, H., Sarailoo, M., Zhou, N., Tran, N., Motalleb, M., & Foruzan, E. (2017). Weather forecasting error in solar energy forecasting. *IET Renewable Power Generation*, 11(10), 1274-1280.
- [242] Wu, W., & Peng, M. (2017). A data mining approach combining k-means clustering with bagging neural network for short-term wind power forecasting. *IEEE Internet of Things Journal*, 4(4), 979-986.

- [243] Zendehboudi, A., Baseer, M. A., & Saidur, R. (2018). Application of support vector machine models for forecasting solar and wind energy resources: A review. *Journal of cleaner production*, 199, 272-285.
- [244] Cai, L., Gu, J., Ma, J., & Jin, Z. (2019). Probabilistic wind power forecasting approach via instance-based transfer learning embedded gradient boosting decision trees. *Energies*, 12(1), 159.
- [245] Lai, J. P., Chang, Y. M., Chen, C. H., & Pai, P. F. (2020). A survey of machine learning models in renewable energy predictions. *Applied Sciences*, 10(17), 5975.
- [246] Wang, H., Lei, Z., Zhang, X., Zhou, B., & Peng, J. (2019). A review of deep learning for renewable energy forecasting. *Energy Conversion and Management*, 198, 111799.
- [247] Abdel-Nasser, M., & Mahmoud, K. (2019). Accurate photovoltaic power forecasting models using deep LSTM-RNN. *Neural Computing and Applications*, 31(7), 2727-2740.
- [248] Syed, D., Refaat, S. S., Abu-Rub, H., Bouhali, O., Zainab, A., & Xie, L. (2019, December). Averaging ensembles model for forecasting of short-term load in smart grids. In *2019 IEEE International Conference on Big Data (Big Data)* (pp. 2931-2938). IEEE.
- [249] Bajaj, M., & Singh, A. K. (2020). Grid integrated renewable DG systems: A review of power quality challenges and state-of-the-art mitigation techniques. *International Journal of Energy Research*, 44(1), 26-69.
- [250] Zainab, A., Refaat, S. S., Syed, D., Ghayeb, A., & Abu-Rub, H. (2019, December). Faulted line identification and localization in power system using machine learning techniques. In *2019 IEEE International Conference on Big Data (Big Data)* (pp. 2975-2981). IEEE.
- [251] Hashemi, F., Mohammadi, M., & Kargarian, A. (2017). Islanding detection method for microgrid based on extracted features from differential transient rate of change of frequency. *IET Generation, Transmission & Distribution*, 11(4), 891-904.
- [252] Alam, M. R., Muttaqi, K. M., & Bouzerdoum, A. (2017). Evaluating the effectiveness of a machine learning approach based on response time and reliability for islanding detection of distributed generation. *IET renewable power generation*, 11(11), 1392-1400.
- [253] James, J. Q., Hou, Y., Lam, A. Y., & Li, V. O. (2017). Intelligent fault detection scheme for microgrids with wavelet-based deep neural networks. *IEEE Transactions on Smart Grid*, 10(2), 1694-1703.
- [254] Bauman, K., Tuzhilin, A., & Zaczynski, R. (2017). Using social sensors for detecting emergency events: a case of power outages in the electrical utility industry. *ACM Transactions on Management Information Systems (TMIS)*, 8(2-3), 1-20.
- [255] García Vera, Y. E., Dufo-López, R., & Bernal-Agustín, J. L. (2019). Energy management in microgrids with renewable energy sources: A literature review. *Applied Sciences*, 9(18), 3854.

- [256] Gui, G., Pan, H., Lin, Z., Li, Y., & Yuan, Z. (2017). Data-driven support vector machine with optimization techniques for structural health monitoring and damage detection. *KSCE Journal of Civil Engineering*, 21(2), 523-534.
- [257] Žarković, M., & Stojković, Z. (2017). Analysis of artificial intelligence expert systems for power transformer condition monitoring and diagnostics. *Electric Power Systems Research*, 149, 125-136.
- [258] Woon, W. L., Aung, Z., & El-Hag, A. (2018, September). Intelligent monitoring of transformer insulation using convolutional neural networks. In *International Workshop on Data Analytics for Renewable Energy Integration* (pp. 127-136). Springer, Cham.
- [259] Chothani, N. G., Raichura, M. B., Patel, D. D., & Mistry, K. D. (2018, October). Real-time monitoring & protection of power transformer to enhance smart grid reliability. In *2018 IEEE electrical power and energy conference (EPEC)* (pp. 1-6). IEEE.
- [260] Tan, H. F. C., Woo, W. L., Sharma, A., Logenthiran, T., & Kumar, D. S. (2019, May). Study of smart condition monitoring using deep neural networks with dropouts and cross-validation. In *2019 IEEE Innovative Smart Grid Technologies-Asia (ISGT Asia)* (pp. 3965-3969). IEEE.
- [261] Sheng, G., Hou, H., Jiang, X., & Chen, Y. (2018). A novel association rule mining method of big data for power transformers state parameters based on probabilistic graph model. *IEEE Transactions on Smart Grid*, 9(2), 695-702.
- [262] Aizpurua, J. I., Garro, U., Muxika, E., Mendicute, M., Gilbert, I. P., Stewart, B. G., ... & Lambert, B. (2019, October). Probabilistic power transformer condition monitoring in smart grids. In *2019 6th International Advanced Research Workshop on Transformers (ARWtr)* (pp. 42-47). IEEE.
- [263] Haes Alhelou, H., Hamedani-Golshan, M. E., Njenda, T. C., & Siano, P. (2019). A survey on power system blackout and cascading events: Research motivations and challenges. *Energies*, 12(4), 682.
- [264] Jayachandran, M., Reddy, C., Padmanaban, S., & Milyani, A. H. (2021). Operational planning steps in smart electric power delivery system. *Scientific Reports*, 11(1), 1-21.
- [265] Ali, A., & Abdelhadi, A. (2022). Condition-Based Monitoring and Maintenance: State of the Art Review. *Applied Sciences*, 12(2), 688.
- [266] Liu, B., Wu, S., Xie, M., & Kuo, W. (2017). A condition-based maintenance policy for degrading systems with age-and state-dependent operating cost. *European Journal of Operational Research*, 263(3), 879-887.
- [267] Wang, X., McArthur, S. D., Strachan, S. M., Kirkwood, J. D., & Paisley, B. (2017). A data analytic approach to automatic fault diagnosis and prognosis for distribution automation. *IEEE Transactions on Smart Grid*, 9(6), 6265-6273.
- [268] Peng, G., Tang, S., Lin, Z., & Zhang, Y. (2017, March). Applications of fuzzy multilayer support vector machines in fault diagnosis and forecast of electric power equipment. In *2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)* (pp. 457-461). IEEE.

- [269] Zhang, S., Wang, Y., Liu, M., & Bao, Z. (2017). Data-based line trip fault prediction in power systems using LSTM networks and SVM. *Ieee Access*, 6, 7675-7686.
- [270] Rafiei, M., Niknam, T., Aghaei, J., Shafie-Khah, M., & Catalão, J. P. (2018). Probabilistic load forecasting using an improved wavelet neural network trained by generalized extreme learning machine. *IEEE Transactions on Smart Grid*, 9(6), 6961-6971.
- [271] Rocchetta, R., Bellani, L., Compare, M., Zio, E., & Patelli, E. (2019). A reinforcement learning framework for optimal operation and maintenance of power grids. *Applied energy*, 241, 291-301.
- [272] Hu, C. (2019). Ensemble feature learning-based event classification for cyber-physical security of the smart grid. Doctoral dissertation, Concordia University.
- [273] Vallim Filho, A. R. D. A., Farina Moraes, D., Bhering de Aguiar Vallim, M. V., Santos da Silva, L., & da Silva, L. A. (2022). A Machine Learning Modeling Framework for Predictive Maintenance Based on Equipment Load Cycle: An Application in a Real World Case. *Energies*, 15(10), 3724.
- [274] Qu, M., Wang, T., Li, F., Liu, L., Li, Q., Long, H., ... & Xu, J. (2021). Distribution Network Electrical Topology Identification Algorithm Based on Deep Learning. In *IOP Conference Series: Earth and Environmental Science* (Vol. 645, No. 1, p. 012076). IOP Publishing.
- [275] Abu-Elanien, A. E., Salama, M. M. A., & Shaban, K. B. (2018). Modern network reconfiguration techniques for service restoration in distribution systems: A step to a smarter grid. *Alexandria engineering journal*, 57(4), 3959-3967.
- [276] Liu, G., Jiang, T., Ollis, T. B., Li, X., Li, F., & Tomsovic, K. (2020). Resilient distribution system leveraging distributed generation and microgrids: A review. *IET Energy Systems Integration*, 2(4), 289-304.
- [277] Mahdavi, M., Alhelou, H. H., Hatziargyriou, N. D., & Jurado, F. (2021). Reconfiguration of Electric Power Distribution Systems: Comprehensive Review and Classification. *IEEE Access*.
- [278] Manousakis, N. M., & Korres, G. N. (2021). Application of State Estimation in Distribution Systems with Embedded Microgrids. *Energies*, 14(23), 7933.
- [279] Skolfield, J. K., & Escobedo, A. R. (2021). Operations research in optimal power flow: A guide to recent and emerging methodologies and applications. *European Journal of Operational Research*.
- [280] Mohsenian-Rad, H. (2021). *Smart Grid Sensors: Principles and Applications*. Cambridge University Press.
- [281] Pengwah, A. B., Fang, L., Razzaghi, R., & Andrew, L. L. (2021). Topology Identification of Radial Distribution Networks Using Smart Meter Data. *IEEE Systems Journal*.
- [282] Razmi, P., Ghaemi Asl, M., Canarella, G., & Emami, A. S. (2021). Topology identification in distribution system via machine learning algorithms. *Plos one*, 16(6), e0252436.

- [283] Lydia, E. L., Kumar, B. P., & Ramya, D. (2018). Generation of dynamic energy management using data mining techniques basing on big data analytics issues in smart grids. *International Journal of Engineering & Technology*, 7(2.26), 85-89.
- [284] Busacca, A., Di Tommaso, A. O., Miceli, R., Nevoloso, C., Schettino, G., Scaglione, G., ... & Colak, I. (2022). Switching Frequency Effects on the Efficiency and Harmonic Distortion in a Three-Phase Five-Level CHBMI Prototype with Multicarrier PWM Schemes: Experimental Analysis. *Energies*, 15(2), 586.
- [285] Kabir, K. H., Aurko, S. Y., & Rahman, M. (2021). Smart Power Management in OIC Countries: A Critical Overview Using SWOT-AHP and Hybrid MCDM Analysis. *Energies*, 14(20), 6480.
- [286] Ray, P. K., Dash, S. K., Subudhi, B., & Korkua, S. K. (2020). Mitigation of power quality issues using UPQC. *International Journal of Emerging Electric Power Systems*, 21(5).
- [287] Mansor, M. A., Othman, M. M., Musirin, I., & Mohammad Noor, S. Z. (2021). Mitigation of power quality issues using UPQC with STF-SRF technique under adverse voltage source condition. *Journal of Electrical and Electronic Systems Research (JEESR)*, 19, 105-116.
- [288] Wang, H., Huang, Z., Zhang, X., Huang, X., wei Zhang, X., & Liu, B. (2022). Intelligent power grid monitoring and management strategy using 3D model visual computation with deep learning. *Energy Reports*, 8, 3636-3648.
- [289] Mahela, O. P., & Shaik, A. G. (2017). Recognition of power quality disturbances using S-transform based ruled decision tree and fuzzy C-means clustering classifiers. *Applied Soft Computing*, 59, 243-257.
- [290] Luo, Y., Li, K., Li, Y., Cai, D., Zhao, C., & Meng, Q. (2017). Three-layer Bayesian network for classification of complex power quality disturbances. *IEEE Transactions on Industrial Informatics*, 14(9), 3997-4006.
- [291] Zhu, R., Gong, X., Hu, S., & Wang, Y. (2019). Power quality disturbances classification via fully-convolutional siamese network and k-Nearest neighbor. *Energies*, 12(24), 4732.
- [292] Wang, S., & Chen, H. (2019). A novel deep learning method for the classification of power quality disturbances using deep convolutional neural network. *Applied energy*, 235, 1126-1140.
- [293] Chen, Z., Li, M., Ji, T., & Wu, Q. (2020). Real-time recognition of power quality disturbance-based deep belief network using embedded parallel computing platform. *IEEE Transactions on Electrical and Electronic Engineering*, 15(4), 519-526.
- [294] Choudhary, B. (2021). An advanced genetic algorithm with improved support vector machine for multi-class classification of real power quality events. *Electric Power Systems Research*, 191, 106879.
- [295] Jordan, S., & Messner, M. (2020). The use of forecast accuracy indicators to improve planning quality: Insights from a case study. *European Accounting Review*, 29(2), 337-359.

- [296] Guerrero-Prado, J. S., Alfonso-Morales, W., Caicedo-Bravo, E., Zayas-Pérez, B., & Espinosa-Reza, A. (2020). The power of big data and data analytics for AMI data: A case study. *Sensors*, 20(11), 3289.
- [297] Liu, D., Zeng, L., Li, C., Ma, K., Chen, Y., & Cao, Y. (2018). A Distributed Short-Term Load Forecasting Method Based on Local Weather Information. *IEEE Systems Journal*, 12(1), 208-215.
- [298] Ahmad, A., Javaid, N., Guizani, M., Alrajeh, N., & Khan, Z. A. (2017). An Accurate and Fast Converging Short-Term Load Forecasting Model for Industrial Applications in a Smart Grid. *IEEE Transactions on Industrial Informatics*, 13(5), 2587.
- [299] Lai, C. S., Yang, Y., Pan, K., Zhang, J., Yuan, H., Ng, W. W., ... & Lai, L. L. (2020). Multi-view neural network ensemble for short and mid-term load forecasting. *IEEE Transactions on Power Systems*, 36(4), 2992-3003.
- [300] Shi, H., Xu, M., & Li, R. (2018). Deep Learning for Household Load Forecasting—A Novel Pooling Deep RNN. *IEEE Transactions on Smart Grids*, 9(5), 5271-5280.
- [301] Fekri, M. N., Patel, H., Grolinger, K., & Sharma, V. (2021). Deep learning for load forecasting with smart meter data: Online Adaptive Recurrent Neural Network. *Applied Energy*, 282, 116177.
- [302] Kong, W., Dong, Z. Y., Jia, Y., Hill, D. J., Xu, Y., & Zhang, Y. (2017). Short-term residential load forecasting based on LSTM recurrent neural network. *IEEE Transactions on Smart Grid*, 10(1), 841-851.
- [303] Jiao, R., Zhang, T., Jiang, Y., & He, H. (2018). Short-term non-residential load forecasting based on multiple sequences LSTM recurrent neural network. *IEEE Access*, 6, 59438-59448.
- [304] Sun, G., Jiang, C., Wang, X., & Yang, X. (2020). Short-term building load forecast based on a data-mining feature selection and LSTM-RNN method. *IEEE Transactions on Electrical and Electronic Engineering*, 15(7), 1002-1010.
- [305] Wang, Y., Zhang, N., & Chen, X. (2021). A Short-Term Residential Load Forecasting Model Based on LSTM Recurrent Neural Network Considering Weather Features. *Energies*, 14(10), 2737.
- [306] Hansen, T. M., Chong, E. K., Suryanarayanan, S., Maciejewski, A. A., & Siegel, H. J. (2018). A Partially Observable Markov Decision Process Approach to Residential Home Energy Management. *IEEE Transactions on Smart Grid*, 9(2), 1271.
- [307] Zafar, R., Mahmood, A., Razzaq, S., Ali, W., Naeem, U., & Shehzad, K. (2018). Prosumer based energy management and sharing in smart grid. *Renewable and Sustainable Energy Reviews*, 82, 1675-1684.
- [308] Cao, Y., Li, Y., Liu, X., & Rehtanz, C. (2020). Self-sustainable community of electricity prosumers in distribution system. In *Cyber-Physical Energy and Power Systems* (pp. 119-138). Springer, Singapore.
- [309] Singh, S., & Yassine, A. (2019). Mining Energy Consumption Behavior Patterns for Households in Smart Grid. *IEEE Transactions on Emerging Topics in Computing*, 7(03), 404-419.

- [310] Ahmed, N., Levorato, M., & Li, G. P. (2017). Residential consumer-centric demand side management. *IEEE Transactions on Smart Grid*, 9(5), 4513-4524.
- [311] Peng, B., Wan, C., Dong, S., Lin, J., Song, Y., Zhang, Y., & Xiong, J. (2016, November). A two-stage pattern recognition method for electric customer classification in smart grid. In *2016 IEEE International Conference on Smart Grid Communications (SmartGridComm)* (pp. 758-763). IEEE.
- [312] Biscarri, F., Monedero, I., Garcia, A., Guerrero, J. I., & Leon, C. (2017). Electricity clustering framework for automatic classification of customer loads. *Expert Systems with applications*, 86, 54-63.
- [313] Jeong, H. C., Jang, M., Kim, T., & Joo, S. K. (2021). Clustering of load profiles of residential customers using extreme points and demographic characteristics. *Electronics*, 10(3), 290.
- [314] Ramos, S., Soares, J., Cembranel, S. S., Tavares, I., Foroozandeh, Z., Vale, Z., & Fernandes, R. (2021). Data mining techniques for electricity customer characterization. *Procedia Computer Science*, 186, 475-488.
- [315] Haben, S., Singleton, C., & Grindrod, P. (2017). Analysis and Clustering of Residential Customers Energy Behavioral Demand Using Smart Meter Data. *IEEE Transactions on Smart Grid*, 1(7), 136-144.
- [316] Rajabi, A., Eskandari, M., Ghadi, M. J., Li, L., Zhang, J., & Siano, P. (2020). A comparative study of clustering techniques for electrical load pattern segmentation. *Renewable and Sustainable Energy Reviews*, 120, 109628.
- [317] Li, R., Li, F., & Smith, N. D. (2017). Load Characterization and Low-order Approximation for Smart Metering Data in the Spectral Domain. *IEEE Transactions on Industrial Informatics*, 13(3), 976-984.
- [318] Trotta, G. (2020). An empirical analysis of domestic electricity load profiles: Who consumes how much and when?. *Applied Energy*, 275, 115399.
- [319] Panda, S., Mohanty, S., Rout, P. K., Sahu, B. K., Bajaj, M., Zawbaa, H. M., & Kamel, S. (2022). Residential Demand Side Management model, optimization and future perspective: A review. *Energy Reports*, 8, 3727-3766.
- [320] Lopes, F., & Algarvio, H. (2018). Demand response in electricity markets: an overview and a study of the price-effect on the Iberian daily market. *Electricity markets with increasing levels of renewable generation: Structure, operation, agent-based simulation, and emerging designs*, 265-303.
- [321] Xie, Q., Hui, H., Ding, Y., Ye, C., Lin, Z., Wang, P., ... & Chen, R. (2020). Use of demand response for voltage regulation in power distribution systems with flexible resources. *IET Gener. Transm. Distrib*, 14(5), 883-892.
- [322] Keshtkar, A., & Arzanpour, S. (2017). An adaptive fuzzy logic system for residential energy management in smart grid environments. *Applied Energy*, 186, 68-81.
- [323] Shafie-Khah, M., & Siano, P. (2017). A stochastic home energy management system considering satisfaction cost and response fatigue. *IEEE Transactions on Industrial Informatics*, 14(2), 629-638.

- [324] Homod, R. Z. (2018). Analysis and optimization of HVAC control systems based on energy and performance considerations for smart buildings. *Renewable Energy*, 126, 49-64.
- [325] Al Essa, M. J. M. (2019). Home energy management of thermostatically controlled loads and photovoltaic-battery systems. *Energy*, 176, 742-752.
- [326] Cardoso, E. S. (2020). Advanced Energy Management Strategies for HVAC Systems in Smart Buildings. Doctoral dissertation, Universitat Politècnica de Catalunya.
- [327] Antonopoulos, I., Robu, V., Couraud, B., Kirli, D., Norbu, S., Kiprakis, A., ... & Wattam, S. (2020). Artificial intelligence and machine learning approaches to energy demand-side response: A systematic review. *Renewable and Sustainable Energy Reviews*, 130, 109899.
- [328] Sadeghianpourhamami, N., Refa, N., Strobbe, M., & Develder, C. (2018). Quantitative analysis of electric vehicle flexibility: A data-driven approach. *International Journal of Electrical Power & Energy Systems*, 95, 451-462.
- [329] Munshi, A. A., & Mohamed, Y. A. R. I. (2017). Extracting and defining flexibility of residential electrical vehicle charging loads. *IEEE Transactions on Industrial Informatics*, 14(2), 448-461.
- [330] Munshi, A. A., & Mohamed, Y. A. R. I. (2018). Unsupervised nonintrusive extraction of electrical vehicle charging load patterns. *IEEE Transactions on Industrial Informatics*, 15(1), 266-279.
- [331] Zhao, H., Yan, X., & Ma, L. (2019). Training-free non-intrusive load extracting of residential electric vehicle charging loads. *IEEE Access*, 7, 117044-117053.
- [332] Xiang, Y., Wang, Y., Xia, S., & Teng, F. (2021). Charging Load Pattern Extraction for Residential Electric Vehicles: A Training-Free Nonintrusive Method. *IEEE Transactions on Industrial Informatics*, 17(10), 7028-7039.
- [333] Rauma, K., Funke, A., Simolin, T., Järventausta, P., & Rehtanz, C. (2021). Electric Vehicles as a Flexibility Provider: Optimal Charging Schedules to Improve the Quality of Charging Service. *Electricity*, 2(3), 225-243.
- [334] Sadeghianpourhamami, N. (2018). Data-driven statistical modeling and reinforcement learning based control of flexibility in domestic and electric vehicle power consumption. Doctoral dissertation, Ghent University.
- [335] Dong, G., Ma, J., Wei, R., & Haycox, J. (2019). Electric vehicle charging point placement optimisation by exploiting spatial statistics and maximal coverage location models. *Transportation Research Part D: Transport and Environment*, 67, 77-88.
- [336] Munshi, A. (2019). Big Data Framework for Analytics in Smart Grids and Applications on Electric Vehicle Loads. Doctoral dissertation, University of Alberta.
- [337] Khazaei, H., & Tareq, M. A. (2021). Moderating effects of personal innovativeness and driving experience on factors influencing adoption of BEVs in Malaysia: An integrated SEM–BSEM approach. *Heliyon*, 7(9), e08072.

- [338] Bucci, G., Ciancetta, F., Fiorucci, E., Mari, S., & Fioravanti, A. (2021). State of art overview of Non-Intrusive Load Monitoring applications in smart grids. *Measurement: Sensors*, 18, 100145.
- [339] Zhou, Y., Li, F., Liu, L., Wang, T., Cheng, Z., Li, R., & Gao, J. (2022). Non-intrusive load monitoring method based on the time-segmented state probability. *Energy Reports*, 8, 1418-1423.
- [340] Ruano, A., Hernandez, A., Ureña, J., Ruano, M., & Garcia, J. (2019). NILM techniques for intelligent home energy management and ambient assisted living: A review. *Energies*, 12(11), 2203.
- [341] Puente, C., Palacios, R., González-Arechavala, Y., & Sánchez-Úbeda, E. F. (2020). Non-intrusive load monitoring (NILM) for energy disaggregation using soft computing techniques. *Energies*, 13(12), 3117.
- [342] Chen, Y. Y., Chen, M. H., Chang, C. M., Chang, F. S., & Lin, Y. H. (2021). A smart home energy management system using two-stage non-intrusive appliance load monitoring over fog-cloud analytics based on Tridium's Niagara framework for residential demand-side management. *Sensors*, 21(8), 2883.
- [343] Hosseini, S. S., Agbossou, K., Kelouwani, S., & Cardenas, A. (2017). Non-intrusive load monitoring through home energy management systems: A comprehensive review. *Renewable and Sustainable Energy Reviews*, 79, 1266-1274.
- [344] Himeur, Y., Ghanem, K., Alsalemi, A., Bensaali, F., & Amira, A. (2021). Artificial intelligence based anomaly detection of energy consumption in buildings: A review, current trends and new perspectives. *Applied Energy*, 287, 116601.
- [345] Klemenjak, C., Kovatsch, C., Herold, M., & Elmenreich, W. (2020). A synthetic energy dataset for non-intrusive load monitoring in households. *Scientific Data*, 7(1), 1-17.
- [346] Lin, Y. H. (2018). Design and implementation of an IoT-oriented energy management system based on non-intrusive and self-organizing neuro-fuzzy classification as an electrical energy audit in smart homes. *applied sciences*, 8(12), 2337.
- [347] Wang, H., & Yang, W. (2018). An iterative load disaggregation approach based on appliance consumption pattern. *Applied Sciences*, 8(4), 542.
- [348] Andread, V., Zhao, X. H., Teshome, D. F., Huang, T. D., & Lian, K. L. (2018). A hybrid method of cascade-filtering and committee decision mechanism for non-intrusive load monitoring. *IEEE Access*, 6, 41212-41223.
- [349] Harell, A., Makonin, S., & Bajić, I. V. (2019, May). Wavenilm: A causal neural network for power disaggregation from the complex power signal. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 8335-8339). IEEE.
- [350] Wu, Q., & Wang, F. (2019). Concatenate convolutional neural networks for non-intrusive load monitoring across complex background. *Energies*, 12(8), 1572.

- [351] Massidda, L., Marrocu, M., & Manca, S. (2020). Non-intrusive load disaggregation by convolutional neural network and multilabel classification. *Applied Sciences*, 10(4), 1454.
- [352] Shehzad, F., Asif, M., Aslam, Z., Anwar, S., Rashid, H., Ilyas, M., & Javaid, N. (2021, July). Comparative Study of Data Driven Approaches Towards Efficient Electricity Theft Detection in Micro Grids. In *International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing* (pp. 120-131). Springer, Cham.
- [353] Iniyathan, R., Balaramakrishnan, B., & Vanila, S. (2021). Energy Theft Detection Issues For Advanced Metering Infrastructure Using Iot In Smart Grid. *Technology (IJEET)*, 12(3), 62-67.
- [354] Ghasemi, A. A., & Gitizadeh, M. (2018). Detection of illegal consumers using pattern classification approach combined with Levenberg-Marquardt method in smart grid. *International Journal of Electrical Power & Energy Systems*, 99, 363-375.
- [355] Messinis, G. M., & Hatziargyriou, N. D. (2018). Review of non-technical loss detection methods. *Electric Power Systems Research*, 158, 250-266.
- [356] Huang, H., Liu, S., & Davis, K. (2018, October). Energy theft detection via artificial neural networks. In *2018 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)* (pp. 1-6). IEEE.
- [357] Yao, D., Wen, M., Liang, X., Fu, Z., Zhang, K., & Yang, B. (2019). Energy theft detection with energy privacy preservation in the smart grid. *IEEE Internet of Things Journal*, 6(5), 7659-7669.
- [358] Wen, M., Yao, D., Li, B., & Lu, R. (2018, May). State estimation based energy theft detection scheme with privacy preservation in smart grid. In *2018 IEEE International Conference on Communications (ICC)* (pp. 1-6). IEEE.
- [359] Yip, S. C., Tan, W. N., Tan, C., Gan, M. T., & Wong, K. (2018). An anomaly detection framework for identifying energy theft and defective meters in smart grids. *International Journal of Electrical Power & Energy Systems*, 101, 189-203.
- [360] Gao, Y., Foggo, B., & Yu, N. (2019). A physically inspired data-driven model for electricity theft detection with smart meter data. *IEEE Transactions on Industrial Informatics*, 15(9), 5076-5088.
- [361] Boris Otto , Michael Ten Hompel, Stefan Wrobel,(2022). *Designing Data Spaces*, 20, 329-332
- [362] Shanhe Yi, Zijiang Hao, Zhengrui Qin, and Qun Li (2015), 2015 Third IEEE Workshop on Hot Topics in Web Systems and Technologies .Fog Computing: Platform and Applications, 73
- [363] Xi Fang ,Satyajayant Misra ,Guoliang Xue ,Dejun Yang (2012).*IEEE Communications Surveys and tutorials*, Vol. 14, No. 4, fourth quarter 2012. Smart Grid – The New and Improved Power Grid: A Survey , 944-947
- [364] Yong-Hong Kuo, Andrew Kusiak (2019), *International Journal of Production Research*, 2019 Vol. 57, Nos. 15–16, From data to big data in production research: the past and future trends, 4841-4845

- [365] Evangelos Karakolis, Konstantinos Alexakis Panagiotis Kapsalis, Spiros Mouzakitis ,John Psarras (2022), An end-to-end approach for scalable real time Anomaly detection in smart buildings
- [366] Evangelos Karakolis, Sotiris Pelekis, Spiros Mouzakitis, Ourania Markaki, Katerina Papapostolou, Giorgos Korbakis and John Psarras (2022), ARTIFICIAL INTELLIGENCE FOR NEXT GENERATION ENERGY SERVICES ACROSS EUROPE – THE I-ENERGY PROJECT
- [367] Panagiotis Kapsalis ,Giorgos Kormpakis ,Konstantinos Alexakis ,Evangelos Karakolis ,Spiros Mouzakitis ,Dimitris Askounis (2022), A Reasoning Engine Architecture for Building Energy Metadata Management
- [368] Konstantinos Alexakis ,Panagiotis Kapsalis, Zoi Mylona, Georgios Kormpakis ,Evangelos Karakolis, Christos Ntanos, Dimitris Askounis (2022), Intelligent Querying for Implementing Building Aggregation Pipelines
- [369] Georgios Kormpakis, Panagiotis Kapsalis, Konstantinos Alexakis,Sotiris Pelekis, Evangelos Karakolis, Haris Doukas (2022), An Advanced Visualisation Engine with Role-Based Access Control for Building Energy Visual Analytics
- [370] S. Bag, M.S. Rahman, G. Srivastava, A. Shore,P. Ram (2022), Examining the role of virtue ethics and big data in enhancing viable, sustainable, and digital supply chain performance