

ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ (ΤΕΙ) ΑΘΗΝΑΣ

ΣΧΟΛΗ ΔΙΟΙΚΗΣΗΣ ΚΑΙ ΟΙΚΟΝΟΜΙΑΣ

ΤΜΗΜΑ ΒΙΒΛΙΟΘΗΚΟΝΟΜΙΑΣ ΚΑΙ ΣΥΣΤΗΜΑΤΩΝ

ΠΛΗΡΟΦΟΡΗΣΗΣ

**Big Data: Νέοι ρόλοι και ευκαιρίες
για τους επαγγελματίες
πληροφόρησης**

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

Θάνος Ευάγγελος

Επιβλέπων καθηγητής:

Αλέξανδρος Κουλούρης

Αθήνα, 2017

Ευχαριστίες

Θα ήθελα να εκφράσω τις πιο θερμές ευχαριστίες μου στους καθηγητές μου και ειδικά στον επιβλέπων καθηγητή μου, στην οικογένεια και στους φίλους μου που αποτέλεσαν συμπαραστάτες σε όλη την διάρκεια της εκπόνησης της πτυχιακής εργασίας μου.

Περίληψη

Σκοπός: Διερεύνηση του πεδίου των Big Data, τους νέους ρόλους και τις αναδυόμενες ευκαιρίες για τον επαγγελματία της Πληροφόρησης αλλά και τις προκλήσεις για τις Ελληνικές Βιβλιοθήκες.

Μεθοδολογία: Στην έρευνα συμμετείχαν 49 επαγγελματίες της Πληροφόρησης. Η έρευνα διεξήχθη με την χρήση ενός ερωτηματολογίου 22 ερωτήσεων για την εκτίμηση των απόψεων των επαγγελματιών της Πληροφόρησης σχετικά με τα Big Data. Πιο ειδικά για τεχνικά θέματα, ρόλους και ευκαιρίες για τον επαγγελματία της Πληροφόρησης και τις προκλήσεις για τις Ελληνικές Βιβλιοθήκες.

Αποτελέσματα: Η ανάλυση των αποτελεσμάτων κατέδειξε σημαντικά ευρήματα αναφορικά με τις απόψεις των επαγγελματιών της Πληροφόρησης σχετικά με τα Big Data. Προέκυψε ότι το 32,6% των φορέων στους οποίους απασχολούνται οι επαγγελματίες του δείγματος χρησιμοποιεί τα Big Data, με χαμηλό επίπεδο υιοθέτησης μόλις κατά 15,7%. Σχετικά με τεχνικά θέματα για τα Big Data αναδείχθηκε κύρια μέθοδος συλλογής τα Αρχεία και Δεδομένα Δικτύου (Searching, Indexing) (34,4%), τεχνολογία επεξεργασίας η NoSQL (16,1%) και μέθοδος επεξεργασίας το Indexing (22,6%). Τα κυριότερα προβλήματα των Big Data θεωρήθηκαν τα προβλήματα υποδομών αποθήκευσης πληροφοριών (44,9%) και η έλλειψη ταχύτητας στην άντληση πληροφοριών (31,1%). Οι σπουδαιότεροι ρόλοι που αναδείχθηκαν για τους επαγγελματίες της Πληροφόρησης είναι η προσφορά υπηρεσιών εντοπισμού, επεξεργασίας και αναμόρφωσης πληροφοριών (83,3%), η παροχή εξειδικευμένης πρόσβασης πελατών σε διευρυμένες πηγές πληροφόρησης (78,7%), η βοήθεια σε χρήστες για απόκτηση ικανοτήτων πληροφοριακής παιδείας για εκτίμηση των πληροφοριών (73,2%), η καθοδήγηση χρηστών στη διαδικασία αναζήτησης, ανάκτησης, ερμηνείας και χρήσης πληροφοριών (72,9%). Οι σημαντικότερες ευκαιρίες για τους επαγγελματίες της Πληροφόρησης είναι η απόκτηση εμπειρίας (82,9%), εξειδίκευσης (78,7%) και τεχνογνωσίας σε Big Data (78,7%). Ο βαθμός υιοθέτησης των Big Data από τις Ελληνικές Βιβλιοθήκες είναι χαμηλός μόλις στο 27,1% αλλά παρατηρήθηκε μια θετική τάση (31,9%) για την προοπτική υιοθέτησής τους. Επίσης, αναδύονται προκλήσεις για τις Ελληνικές Βιβλιοθήκες από την υιοθέτηση των Big Data όπως η πληροφορία (επικαιροποίηση, οργάνωση, άντληση, ανάλυση, σύνθεση, διαχείριση) (80,9%), η συνεργασία των επαγγελματιών (78,3%), η δημιουργία υποδομών για μετασχηματισμό δεδομένων σε πληροφορίες και των πληροφοριών σε γνώση (76,6%), οι επαγγελματικές δεξιότητες βιβλιοθηκονόμων (74,5%) και το ανταγωνιστικό πλεονέκτημα στην Οικονομία της Γνώσης (72,3%). Αν και η πρόκληση των Big Data φαίνεται να διεισδύει μελλοντικά σε κρίσιμους τομείς όπως η Τεχνολογική Έρευνα (55,1%), τα Κοινωνικά Δίκτυα (49%), οι Επιχειρήσεις (46,9%), η Οικονομία (26,5%) και η Υγεία (26,5%).

Συμπεράσματα: Τα Big Data μπορούν να προσδώσουν ένα ανταγωνιστικό πλεονέκτημα σε φορείς όπου εφαρμόζονται. Αναδύονται νέοι ρόλοι για τον επαγγελματία της Πληροφόρησης όπως του διαχειριστή της πληροφορίας, του διαμεσολαβητή, του διευκολυντή και του εκπαιδευτή αλλά και ευκαιρίες όπως η απόκτηση εμπειρίας, εξειδίκευσης και τεχνογνωσίας σε Big Data. Προκλήσεις των Big Data για τις Ελληνικές Βιβλιοθήκες αποτελούν η πληροφορία, η συνεργασία των επαγγελματιών, οι υποδομές και οι επαγγελματικές δεξιότητες. Μελλοντικά

διαφαίνεται να διεισδύει σε κρίσιμους τομείς όπως η τεχνολογική έρευνα, τα κοινωνικά δίκτυα, οι επιχειρήσεις, η οικονομία και η υγεία.

Λέξεις Κλειδιά: Big Data, Επαγγελματίας Πληροφόρησης, Βιβλιοθήκες, Ευκαιρία, Πρόκληση.

Summary

Purpose: Exploring Big Data, the new roles and emerging opportunities for professional in Information well as the challenges for the Greek Libraries.

Methodology: The survey involved 49 professionals of Information. The survey was conducted by using a questionnaire with 22 questions for assessing the opinions of professionals in Information about Big Data. Especially for technical issues, roles and opportunities for professional in Information and challenges for the Greek Libraries, for the time period January - February 2017.

Results: The statistical analysis of the results showed significant findings about Big Data. Only 32.6% of organizations use the Big Data with low level of adoption just 15.7%. On technical issues on Big Data main collection method was the Searching - Indexing (34,4%), processing technology NoSQL (16.1%) and the processing method Indexing (22.6%). The main problems of Big Data considered as storage process infrastructures (44.9%) and the lack of speed in the extraction of information (31.1%). The most important roles for professionals are to service in positioning, processing and information reformation (83.3%), to provide specialized customer access to expanded information sources (78.7%), to help users to acquire information literacy skills for assessment of the information (73.2%), to guide user in the search procedure, recovery, interpretation and the use of information (72.9%). The most important opportunities for information professionals is to gain experience (82.9%), specialization (78.7%) and expertise in Big Data (78,7%). The level of adoption of Big Data from Greek Libraries are low at just 27.1%, but a positive trend was observed (31.9%) for their perspective adoption. Also, challenges emerge for the Greek Libraries from their adoption of Big Data such as the information (updating, organization, raise, analysis, management) (80.9%), the cooperation of the professionals (78.3%), the creation of infrastructure for transforming data into information and information into knowledge (76.6%), professional librarians skills (74.5%) and competitive advantage in Economy of Knowledge (72.3%). Although in future the challenge of Big Data seems to penetrate in crucial sectors such as technological research (55.1%), social networks (49%), corporations (46.9%), the economy (26.5%) and health (26.5%).

Conclusions: Big Data can provide a competitive advantage to operators who apply. Are emerging new roles for information professionals such as information manager, mediator, facilitator and trainer as well as opportunities such as the acquisition of experience, specialization and expertise in Big Data. Challenges of Big Data for Greek Libraries are management of information, co-operation of professionals, infrastructure and professional skills. In future seems to penetrate in critical sectors such as technological research, social networks, enterprises, the economy and health.

Keywords: Big Data, Professional in Information, Role, Opportunity, Challenge.

Πίνακας Περιεχομένων

Ευχαριστίες	
Περίληψη	
Summary	
Εισαγωγή	
1. Έννοια Big Data	
1.1. Ορισμός των Big Data	
1.2. Ιστορική εξέλιξη των Big Data	
1.3. Χαρακτηριστικά των Big Data	
1.4. Η δημιουργία των Big Data	
2. Τεχνολογίες Big Data	
2.1. Η τεχνολογία Cloud και τα Big Data	
2.2. Η σχέση Internet of Things (IoT) και Big Data	
2.3. Data Center και Big Data	
2.4. Τεχνολογίες επεξεργασίας Big Data	
2.5. Μέθοδοι επεξεργασίας Big Data	
3. Υλοποίηση Big Data.....	
3.1. Η δημιουργία των δεδομένων (data generation)	
3.2. Η απόκτηση δεδομένων (data acquisition)	
3.2.1. Συλλογή δεδομένων (data collection)	
3.2.2. Μεταφορά δεδομένων (data transportation)	
3.2.3. Προεπεξεργασία των δεδομένων (data pre - processing)	
3.3. Συστήματα αποθήκευσης των Big Data	
3.4. Μηχανισμός αποθήκευσης Big Data	
3.4.1. Τεχνολογίες αποθήκευσης δεδομένων.....	
3.5. Η ανάλυση των Big Data	
3.5.1. Εργαλεία άντλησης και ανάλυσης δεδομένων.....	
3.6. Προβλήματα των Big Data	
4. Υλοποίηση Big Data Σε Βιβλιοθήκες	
4.1. Η εφαρμογή των Big Data από τη Βρετανική Βιβλιοθήκη (British Library)	
4.2. Η υλοποίηση στην Εθνική Γερμανική Βιβλιοθήκη (German National Library)	
4.3. Υλοποίηση Big Data στη Βιβλιοθήκη του Κογκρέσου (Library of Congress)	

5. Big Data & Επαγγελματίες Πληροφόρησης: Νέοι ρόλοι και ευκαιρίες	
5.1. Γνώσεις επαγγελματιών πληροφόρησης και επεξεργασία πληροφοριών	
5.2. Ο ρόλος των επαγγελματιών πληροφόρησης στη δημιουργία και παρουσίαση Big Data	
5.3. Νέοι ρόλοι και ευκαιρίες επαγγελματιών πληροφόρησης στα Big Data	
6. Προκλήσεις στα Big Data	
6.1. Οι προκλήσεις για τα Big Data	
6.2. Η ποιότητα των Big Data ως σημασιολογική πρόκληση	
6.3. Προκλήσεις βιβλιοθηκών και επαγγελματιών πληροφόρησης στα Big Data	
6.3.1. Προκλήσεις οργάνωσης ετερογένειας και ατελειών πληροφοριών	
6.3.2. Επικαιροποίηση πληροφοριών και ταχύτητα άντλησης	
6.3.3. Διασφάλιση ιδιωτικότητας των δεδομένων	
6.3.4. Συνεργασία επαγγελματιών πληροφόρησης	
6.4. Η εφαρμογή των Big Data στις ελληνικές βιβλιοθήκες	
7. Case Study Big Data	
7.1. Εφαρμογές Big Data	
7.1.1. Η εφαρμογή των Big Data στις επιχειρήσεις	
7.1.2. Online κοινωνικά δίκτυα προσαρμοσμένα στα Big Data	
7.1.3. Υγειονομικές και ιατρικές εφαρμογές Big Data	
8. Έρευνα	
9. Αποτελέσματα Έρευνας	
9.1. Στατιστική Ανάλυση Αποτελεσμάτων Έρευνας	
9.2. Συμπεράσματα Έρευνας	
Συμπεράσματα	
Βιβλιογραφικές Αναφορές	
Παράρτημα	
Φόρμα Ερωτηματολογίου	

Εισαγωγή

Η ψηφιοποίηση και οι τεχνολογικές εξελίξεις καθιστούν σημαντική την διαχείριση των Big Data. Εξέχουσα τεχνολογική και επιχειρηματική πρόκληση της σύγχρονης εποχής αποτελεί η διαχείριση του γιγάντιου όγκου δεδομένων «Big Data». Μάλιστα μια πρόσφατη έρευνα της Microsoft αξιολόγησε την αγορά που διαμορφώνεται γύρω από τα Big Data στα 1,6 τρισ. δολάρια στην επόμενη πενταετία. Η σημαντικότητα και η επικαιρότητα των Big Data που χαρακτηρίζονται ως το νέο σημαντικό τεχνολογικό «trend» στον τομέα της Πληροφόρησης έδωσε το έναυσμα για την ενασχόληση με το συγκεκριμένο θέμα.

Η παρούσα εργασία σκοπεύει να αναδείξει τα Big Data, τους νέους ρόλους και τις αναδυόμενες ευκαιρίες για τον επαγγελματία της Πληροφόρησης αλλά και τις προκλήσεις για τις Ελληνικές Βιβλιοθήκες. Για την διερεύνηση των παραπάνω στοιχείων η εργασία δομείται σε δύο μέρη. Στο πρώτο μέρος αναλύονται θεωρητικά μοντέλα αναφορικά με το υπό εξεταζόμενο θέμα μέσα από επτά (7) ενότητες σχετικά με τα Big Data, έννοιες όπως, τεχνολογίες, υλοποίηση, νέοι ρόλοι και ευκαιρίες για επαγγελματίες της πληροφόρησης, προκλήσεις για βιβλιοθήκες και επαγγελματίες, case study. Το δεύτερο μέρος της εργασίας περιλαμβάνει δύο (2) ενότητες, την μεθοδολογία έρευνας και τα αποτελέσματα όπου και παρατίθεται η στατιστική ανάλυση και τα συμπεράσματα της έρευνας που πραγματοποιήθηκε σε 49 επαγγελματίες της Πληροφόρησης για την καταγραφή των στάσεών τους αναφορικά με τα Big Data και ειδικότερα σε τεχνικά θέματα, ρόλους, ευκαιρίες και προκλήσεις.

1. Έννοια Big Data

1.1. Ορισμός των Big Data



Τα Big Data (Μεγάλα Δεδομένα) είναι μια αφηρημένη έννοια. Στην έννοιά τους πέραν του ποσοτικού χαρακτήρα των δεδομένων, υπάρχουν και κάποια άλλα χαρακτηριστικά, τα οποία καθορίζουν τη διαφορά μεταξύ της έννοιας των «μαζικών δεδομένων» και των «πολύ μεγάλων δεδομένων». Ο ορισμός των Big Data στη βάση των χαρακτηριστικών και της σημασίας τους τυγχάνει διαφορετικών προσεγγίσεων. Σε γενικές γραμμές, ως Big Data (Μεγάλα Δεδομένα) νοούνται *«σύνολα δεδομένων τα οποία δεν μπορεί να αντιληφθεί και να διαχειριστεί η παραδοσιακή πληροφορική και τα εργαλεία λογισμικού μέσα σε ένα ανεκτό επίπεδο χρόνου»* (Chen et al., 2014:172).

Οι Manyika et al. (2011) αναφέρουν ως Big Data *«ένα σύνολο στοιχείων και δεδομένων τα οποία δεν μπορούν να συγκεντρωθούν, να αποθηκευτούν και να επεξεργαστούν τα παραδοσιακά λογισμικά βάσεων δεδομένων»*. Στο συγκεκριμένο ορισμό αρχικά σημειώνεται η ευμεταβλητότητα του όγκου των δεδομένων και δευτερευόντως η αδυναμία επεξεργασίας τους με παραδοσιακές μεθόδους. Ομοίως, οι Davis & Patterson (2012:4) και οι Emani et al. (2015:71) αναφέρουν ότι *«τα Big Data είναι δεδομένα πολύ μεγάλου όγκου ώστε να τα διαχειριστούν και να τα αναλύσουν παραδοσιακά πρωτόκολλα βάσεων δεδομένων όπως η SQL»*.

Οι Mayer - Schonberger & Cukier (2013) από την πλευρά τους ορίζουν τα Big Data ως *«τεράστια σύνολα δεδομένων και πληροφοριών απ' όπου μπορούν να εξαχθούν χρήσιμα συμπεράσματα στη βάση χαρακτηριστικών όπως ο μεγάλος όγκος, η ποικιλία, η ταχύτητα και η αξία τους»*. Στην ίδια άποψη αναφορικά με τον ορισμό των Big Data κατευθύνονται και οι Krishnan (2013) και Reeve (2013) οι οποίοι σημειώνουν ότι η επεξεργασία τους καθίσταται αδύνατη από τα παραδοσιακά λογισμικά.

Προσεγγίζοντας εννοιολογικά τα Big Data ο Kraska (2013:84-85) αναφέρει ότι *«τα Big Data είναι η λύση όταν η κανονική εφαρμογή της τρέχουσας τεχνολογίας δεν επιτρέπει στους χρήστες να έχουν έγκαιρες, αποδοτικές και ποιοτικές απαντήσεις σε ερωτήσεις που βασίζονται σε δεδομένα»*. Ομοίως, οι Jacobs (2009:36-37) και Wang et

al. (2016:749) σημειώνουν ότι «αποτελούν δεδομένα τεράστιου όγκου που επιτρέπουν την αξιοποίηση πληροφοριών πέραν των παραδοσιακών τεχνολογιών σε θέματα δεδομένων».

Σε μια προσπάθεια να δοθεί ένας γενικός ορισμός των Big Data αναφέρεται ότι αποτελούν οποιαδήποτε συλλογή συνόλων δεδομένων, τα οποία έχουν τέτοιο όγκο και πολυπλοκότητα ώστε η διαχείριση και επεξεργασία τους καθίσταται αδύνατη αξιοποιώντας τις δυνατότητες των παραδοσιακών τεχνολογικών εργαλείων και βάσεων δεδομένων (Vitoło, 2015:186; Koman & Kundrikova, 2016:606).

1.2. Ιστορική εξέλιξη των Big Data

Η ιστορική εξέλιξη των Big Data βασίστηκε στη διαχρονική αύξηση της αποτελεσματικής αποθήκευσης και διαχείρισης μεγαλύτερων συνόλων δεδομένων. Η ιστορία των Big Data ακολούθησε τρία βασικά στάδια: από τα megabyte στα gigabyte, από τα gigabyte στα terabyte και από τα terabyte στα petabyte. Στη δεκαετία 1970 - 1980, εισήχθη η αρχική πρόκληση με τη μετάβαση από τα megabyte σε gigabyte για τις επιχειρήσεις. Η ανάλυση στοιχείων και εκθέσεων ουσιαστικά οδήγησε στη δημιουργία των βάσεων δεδομένων που χαρακτηρίστηκε ως ιδανική για την εποχή επιλογή στην επίλυση προβλημάτων. Ωστόσο, οι τεχνολογικές εξελίξεις ήταν τέτοιες με τις βάσεις δεδομένων να μην συμβαδίζουν (Sharma & Gupta, 2015:28).

Στα τέλη της δεκαετίας του 1980, η εκλαΐκευση της ψηφιακής τεχνολογίας δημιούργησε όγκους δεδομένων που εκτιμούνταν σε αρκετά gigabytes ή ακόμα και terabyte, τα οποία ξεπερνούσαν την ικανότητα αποθήκευσης και επεξεργασίας ενός συστήματος υπολογιστών. Το συγκεκριμένο πρόβλημα δημιούργησε τις παράλληλες βάσεις δεδομένων και επιμερισμού των πληροφοριών. Χαρακτηριστικά είναι η περίοδος όπου ξεχωρίζει η αρχιτεκτονική shared – nothing που βασίζεται σε ένα σύμπλεγμα συστημάτων, το καθένα με το δικό του επεξεργαστή, μνήμη και δίσκο (Dewitt & Gray, 1992:86-88).

Στα τέλη της δεκαετίας του 1990, όταν εξαπλώθηκαν τα δεδομένα σε μεγαλύτερο βαθμό με την ανάπτυξη του διαδικτύου, η αποθήκευση δεδομένων έφθασε να απαιτεί όγκους πολλών terabytes ή petabytes. Η ανάγκη διαχείρισης και επεξεργασίας δεδομένων αυξήθηκε ακόμη περισσότερο με τις παράλληλες βάσεις δεδομένων να

μην είναι σε θέση να ανταποκριθούν. Χαρακτηριστικά, είναι η εποχή όπου η Google δημιουργεί το Google File (GFS) και το Map Reduce ώστε να αξιοποιηθούν δεδομένα ευρείας κλίμακας (Ghemawat et al., 2003:29-31). Στις 2 Ιουνίου 1986, γεγονός ορόσημο στην ανάπτυξη των Big Data αποτέλεσε η πρώτη παράλληλη βάση δεδομένων (Teradata) με τη χωρητικότητα αποθήκευσης 1TB (Ghemawat & Leung, 2003:31).

Στα μέσα της δεκαετίας του 2000, επέρχεται η αλματώδης εξέλιξη στις βάσεις δεδομένων όπως η NoSQL δημιουργώντας ένα νέο περιβάλλον απαιτήσεων (Dean & Ghemawat, 2008:108). Σύμφωνα με τις τρέχουσες τάσεις ανάπτυξης, τα δεδομένα όπως αποθηκεύονται και αναλύονται αναμφίβολα θα αυξήσουν περαιτέρω τις τεχνολογικές απαιτήσεις. Τα τελευταία χρόνια, σχεδόν όλες οι μεγάλες εταιρείες (EMC, Oracle, IBM, Microsoft, Google, Amazon, Facebook κλπ) έχουν ξεκινήσει μεγάλα έργα για τα Big Data. Από το 2011 και μετά όλο και περισσότερο απασχολεί η οργάνωση και διαχείριση των Big Data ταυτόχρονα με τις τεχνολογικές εξελίξεις (Gantz & Reinsel, 2011:2-4). Ο όγκος δεδομένων παγκοσμίως αναμένεται να αυξηθεί κατά 40% ανά έτος ως το 2020 ενώ έχει παρατηρηθεί ότι το 90% των σημερινών δεδομένων δημιουργήθηκε από το 2014 και μετά (Yaqoob et al., 2016:1233). Λόγω της ταχείας ανάπτυξης στην παραγωγή δεδομένων το 2020 υπολογίζεται ότι ο όγκος τους θα είναι 44 φορές μεγαλύτερος από ότι το 2009 (Khan et al., 2014:1).

1.3. Χαρακτηριστικά των Big Data

Στα χαρακτηριστικά των Big Data οι περισσότεροι συγγραφείς αναφέρουν τα 4V (Volume, Variety, Velocity, Veracity). Αν και τα τρία (3) εξ αυτών συναντώνται συχνότερα στη βιβλιογραφία (volume, variety, velocity), η ειλικρίνεια (veracity) και η αξία (value) αποτελούν επιπλέον χαρακτηριστικά των Big Data. Ως προς τα χαρακτηριστικά των Big Data σημειώνονται τα εξής:

1. Όγκος (Volume). Ο όγκος αναφέρεται στο πλήθος των δεδομένων που δημιουργούνται. Ο τεράστιος όγκος είναι βασική ιδιότητα των Big Data. Γενικά ακολουθούν εκθετική αύξηση ενώ δημιουργούνται πολύ γρήγορα. Τα Big Data σε επίπεδο όγκου προσφέρουν τη δυνατότητα επεξεργασία πολλών δεδομένων. Το συγκεκριμένο χαρακτηριστικό είναι ιδιαίτερα χρήσιμο ιδιαίτερα τόσο στην αποθήκευση όσο και στην επεξεργασία για είδη

δεδομένων από κοινωνικά δίκτυα, υγειονομική περίθαλψη, χρηματοοικονομικά δεδομένα, βιοχημεία, γενετικά δεδομένα κ.λπ. (Emani et al., 2015:71; Vitolo, 2015:186; Singh & Singla, 2015:6).

2. Ποικιλία (Variety). Τα δεδομένα των Big Data δεν έχουν μια σταθερή δομή και σπάνια παρουσιάζονται σε μια απόλυτα επεξεργάσιμη μορφή. Τα δεδομένα αυτά μπορεί να είναι ιδιαίτερα δομημένα (δεδομένα από σχεσιακές βάσεις δεδομένων), ημιδομημένα (αρχεία καταγραφής ιστού, social media feeds κλπ) ή μη δομημένα (βίντεο, φωτογραφίες). Η ποικιλία είναι μία από τις πιο σημαντικές ιδιότητες των Big Data. Η απουσία συγκεκριμένης δομής δημιουργεί από μόνη της προκλήσεις για τις παραδοσιακές τεχνολογίες και βάσεις δεδομένων. Το εύρος και η ποικιλία των δεδομένων είναι τέτοια ώστε να επηρεάζουν τη σημασιολογία, ή τη μεταβλητότητα του νοήματος (Emani et al., 2015:72; Hoy, 2014:321; Wang et al., 2016:750; Singh & Singla, 2015:6).
3. Ταχύτητα (Velocity). Η ταχύτητα είναι ένα χαρακτηριστικό των Big Data που αφορά το χρόνο δημιουργίας και αποθήκευσης των δεδομένων, τη διαθεσιμότητα και την πρόσβαση. Το ζήτημα της ταχύτητας δεν εστιάζει τόσο στην αποθήκευση όσο στη ταχύτητα με την οποία δημιουργούνται τα δεδομένα. Η ταχύτητα είναι μια άλλη βασική ιδιότητα των μεγάλων δεδομένων που παράγονται πολύ γρήγορα καθημερινά στο ψηφιακό κόσμο (Krishnan, 2013; Hoy, 2014:321; Wang et al., 2016:750; Singh & Singla, 2015:6).
4. Ειλικρίνεια (Veracity). Η ειλικρίνεια των πληροφοριών βασίζεται στο βαθμό συμφωνίας με την αλήθεια ή την πραγματικότητα. Τυχόν αβεβαιότητες στα δεδομένα μπορούν να προκληθούν από λανθασμένες εφαρμογές μοντέλων, παραπλάνηση, επικάλυψη πληροφοριών (Emani et al., 2015:72; Wang et al., 2016:750; Yaqoob et al., 2016:1232).

1.4. Η δημιουργία των Big Data

Η δημιουργία των Big Data εξαρτήθηκε σε μεγάλο βαθμό από τις εξελίξεις που σημειώθηκαν σε επίπεδο ηλεκτρονικών υπολογιστών, εφαρμογών, λογισμικού, διαδικτύου και μεγάλων εφαρμογών δεδομένων (Abolfazli et al., 2014:342). Η ανάγκη δημιουργίας των Big Data προέκυψε στη βάση αξιοποίησης όσων το δυνατό περισσότερων πληροφοριών και δεδομένων στη λήψη μιας απόφασης. Τα Big Data ακολουθώντας κατά περίπτωση διαφορετικές τεχνικές και μεθόδους αποθήκευσης στο υπολογιστικό περιβάλλον, προσφέρουν δυνατότητες επεξεργασίας των δεδομένων που συντίθενται. Επίσης, δημιουργούν πρότυπα και επιμέρους αλληλεπιδράσεις μεταξύ των δεδομένων δημιουργώντας τις κατά το δυνατόν ιδανικότερες συνθήκες στη λήψη αποφάσεων.

Τα Big Data βασίζονται στην ύπαρξη διασυνδεδεμένων υπολογιστών μέσω δικτύου οι οποίοι λειτουργούν ως ενιαίο σύστημα. Αυτοί οι υπολογιστές μπορούν να είναι απομακρυσμένοι ο ένας από τον άλλο, συνδεδεμένοι σε ένα δίκτυο ευρείας περιοχής ή σε τοπικό δίκτυο. Στη δημιουργία των Big Data ξεχωρίζει η αποθήκευση των δεδομένων καθώς και το λογισμικό διαχείρισης των συστημάτων αποθήκευσης (Cooper, 2012:8). Η μελέτη της γένεσης των Big Data είναι σημαντική ώστε να γίνουν αντιληπτά το εννοιολογικό θεμέλιο, το όραμα και η τάση τους. Η εξέλιξη των Big Data βασίζεται στην πληροφορία η οποία αξιοποιείται εντασσόμενη σε ένα συνολικότερο σύστημα που περιλαμβάνει κατάλληλη αποθήκευση, λογισμικό επεξεργασίας και άντλησης (Abolfazli et al., 2014:351).

Οι ανάγκες για βελτίωση της ικανότητας αποθήκευσης δεδομένων έχουν αυξηθεί ραγδαία ενώ και οι απαιτήσεις των χρηστών συνεχίζουν να μεταβάλλονται με την πάροδο του χρόνου. Πλέον οι διαθέσιμες μορφές επεξεργασίας των δεδομένων σε ορισμένες περιπτώσεις θεωρούνται απαρχαιωμένες μη μπορώντας να ανταποκριθούν στις ανάγκες των χρηστών. Ως εκ τούτου, οι διάφορες επιλογές παρακολούθησης και αξιοποίησης των δεδομένων δείχνουν να εστιάζουν κυρίως στα Big Data. Η χρήση δεδομένων στη λήψη αποφάσεων αποτελεί πλέον βασική επιλογή με την ανάπτυξη των νέων τεχνολογιών και την εξάπλωση των κοινωνικών δικτύων (Kacprzyk & Zadrozny, 2001:85).

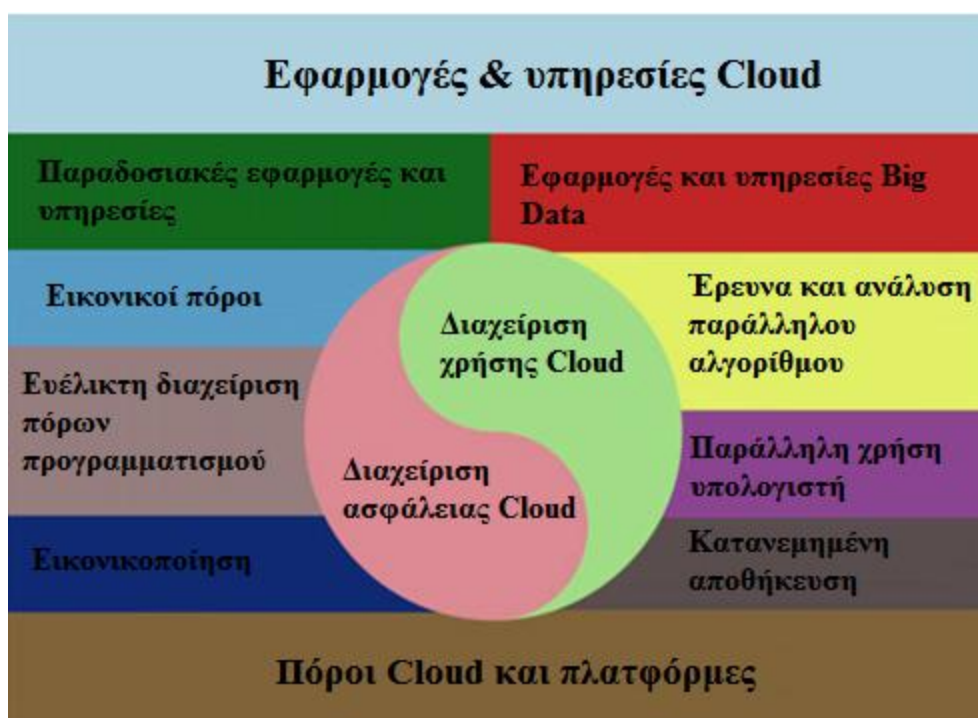
Τα Big Data δημιουργούνται ουσιαστικά σε κάθε χρονική στιγμή ενώ οποιαδήποτε δραστηριότητα (αλληλεπίδραση στα μέσα κοινωνικής δικτύωσης, browsing, online

αγορές, Facebook κ.ά.) παράγει δεδομένα από πολλαπλές πηγές. Η αξιοποίηση ωστόσο των Big Data απαιτεί κατάλληλη επεξεργαστική ισχύ, εργαλεία ανάλυσης, χώρους αποθήκευσης. Σε κάθε περίπτωση, η δημιουργία των Big Data δεν μπορεί να περιοριστεί, μπορεί όμως με κατάλληλη αξιοποίηση να προσφέρει σημαντική πληροφόρηση (Yaqoob et al., 2016:1233).

2. Τεχνολογίες Big Data

2.1. Η τεχνολογία Cloud και τα Big Data

Τα Big Data στη λειτουργία τους αξιοποιούν τις δυνατότητες αποθήκευσης που προσφέρει το Cloud. Ο κύριος στόχος του Cloud είναι να προσφέρει τεράστιες αποθηκευτικές δυνατότητες έτσι ώστε η αξιοποίηση των πληροφοριών από τα Big Data να είναι άμεση και ταχεία. Η ανάπτυξη του Cloud αφενός παρέχει λύσεις για την αποθήκευση και επεξεργασία των Big Data και αφετέρου εκείνα με τη σειρά τους επιταχύνουν την ανάπτυξη του Cloud. Η τεχνολογία της κατανεμημένης αποθήκευσης που βασίζεται σε Cloud μπορεί να διαχειριστεί αποτελεσματικά μεγάλα δεδομένα. Η παράλληλη υπολογιστική ικανότητα του Cloud μπορεί να βελτιώσει την αποτελεσματικότητα της απόκτησης και της ανάλυσης των Big Data. Το Cloud είναι στενά συνδεδεμένο με τα Big Data. Τα στοιχεία που συνθέτουν το Cloud παρουσιάζονται στο σχήμα 2.1 που ακολουθεί (Chen et al., 2014:176).



Σχήμα 2.1 : Στοιχεία σύνθεσης Cloud
Πηγή:Chen et al. (2014)

Ωστόσο υπάρχουν πολλές επικαλυπτόμενες τεχνολογίες και διαφορές ανάμεσα στο Cloud και τα Big Data. Αρχικά, ως έννοιες παρουσιάζουν διαφοροποιήσεις στο ότι το Cloud μετασχηματίζει την αρχιτεκτονική της πληροφορικής ενώ τα Big Data ασκούν μεγάλες επιρροές στη λήψη επιχειρηματικών αποφάσεων. Ωστόσο, τα Big Data εξαρτώνται από το Cloud για την ομαλή λειτουργία τους. Επίσης, παρουσιάζουν διαφοροποιήσεις ως προς το επίπεδο των χρηστών. Το Cloud είναι μία τεχνολογία που χρησιμοποιείται ευρέως από ανώτερα στελέχη (CIO) ως λύση προηγμένης τεχνολογίας. Τα Big Data είναι ένα προϊόν που αξιοποιούν ανώτατα στελέχη των επιχειρήσεων (CEO) εστιάζοντας σε επιχειρηματικές δραστηριότητες. Δεδομένου ότι οι φορείς λήψης αποφάσεων των επιχειρήσεων μπορεί να αισθανθούν άμεσα την πίεση από τον ανταγωνισμό της αγοράς, θα πρέπει να επικρατήσουν στην αγορά αξιοποιώντας όλους τους διαθέσιμους τρόπους (Simmhan et al., 2013:46; Chen et al., 2014:176; Kobiellus et al., 2014:7-8).

Με την πρόοδο των Big Data και του Cloud, αυτές οι δύο τεχνολογίες όλο και περισσότερο συνεργάζονται. Το Cloud λειτουργώντας με τρόπο που ομοιάζει με τα λειτουργικά συστήματα, παρέχει σημαντικούς πόρους σε επίπεδο συστήματος. Τα Big Data λειτουργούν σε ανώτερο επίπεδο υποστηριζόμενα από το Cloud και παρέχουν λειτουργίες παρόμοιες με εκείνες της βάσης δεδομένων όπως η αποτελεσματική επεξεργασία των δεδομένων. Η εξέλιξη των Big Data επηρεάζεται από την ταχεία ανάπτυξη των εφαρμογών και του Cloud. Ως εκ τούτου, το Cloud όχι μόνο παρέχει δυνατότητες υπολογισμού και επεξεργασίας στα Big Data αλλά και σημαντικά στοιχεία λειτουργίας (Chen et al., 2014:176)

2.2. Η σχέση Internet of Things (IoT) και Big Data

Στην περίπτωση του Internet of Things (Διαδίκτυο των Πραγμάτων), πολλά στοιχεία που συγκεντρώνονται προέρχονται από δεδομένα του πραγματικού κόσμου από κατάλληλες συσκευές και αισθητήρες που τα καταγράφουν. Οι αισθητήρες ενσωματώνονται σε διάφορες συσκευές και μηχανήματα στον πραγματικό κόσμο συλλέγοντας διάφορα είδη δεδομένων, όπως περιβαλλοντικά στοιχεία, γεωγραφικά δεδομένα, αστρονομικά δεδομένα, διοικητικά στοιχεία κλπ. Για την καταγραφή δεδομένων στο IoT μπορεί να αξιοποιηθεί και κινητός εξοπλισμός, εγκαταστάσεις

μεταφορών, δημόσιες εγκαταστάσεις και οικιακές συσκευές (Gubbi et al., 2013:1646).

Τα Big Data που παράγονται από IoT έχουν διαφορετικά χαρακτηριστικά σε σύγκριση με τα γενικά Big Data λόγω του διαφορετικού τύπου των δεδομένων που συλλέγονται με χαρακτηριστικά στοιχεία την ετερογένεια, την ποικιλία, τον αδόμητο χαρακτήρα. Παρά το γεγονός ότι τα τρέχοντα δεδομένα από το IoT δεν αποτελούν το κυρίαρχο μέρος των Big Data, από το 2030 όπου ο αριθμός των αισθητήρων θα φτάσει το 1 τρις. τα δεδομένα του IoT θα αποτελούν το βασικότερο τμήμα των Big Data. Μια έκθεση από την Intel επισήμανε ότι τα δεδομένα στο IoT έχουν τρία χαρακτηριστικά που συμμορφώνονται με το πρότυπο των Big Data: α) μεγάλος αριθμός τερματικών που δημιουργούν πληθώρα δεδομένων, β) τα δεδομένα που προέρχονται από το IoT είναι συνήθως ημιδομημένα ή αδόμητα και γ) τα δεδομένα του IoT είναι χρήσιμα μόνο όταν αναλύονται (Chen et al., 2014:177).

Προς το παρόν, η ικανότητα επεξεργασίας δεδομένων του IoT βρίσκεται ακόμα σε χαμηλότερα επίπεδα έναντι των Big Data καθιστώντας αναγκαία την εισαγωγή των νέων τεχνολογιών δεδομένων για την προώθηση της ανάπτυξης του IoT. Πολλοί φορείς εκμετάλλευσης του IoT έχουν κατανοήσει τη σημασία των Big Data για την τελική του επιτυχία μέσω αποτελεσματικής ενσωμάτωσής τους και αξιοποίησης του Cloud. Υπάρχει επιτακτική ανάγκη να υιοθετηθούν νέες τεχνολογίες για το IoT ενώ και η ανάπτυξη των Big Data μέσω IoT βρίσκεται σε πρώιμο στάδιο. Έχει αναγνωριστεί ότι οι δύο τεχνολογίες είναι αλληλένδετες και θα πρέπει να αναπτυχθούν από κοινού: από τη μία πλευρά, η ευρεία διάδοση του IoT οδηγεί σε υψηλή αύξηση των Big Data, τόσο σε ποσότητα και σε είδη παρέχοντας έτσι νέες ευκαιρίες στην εφαρμογή και ανάπτυξή τους ενώ από την άλλη πλευρά η απαίτηση νέων δεδομένων στο IoT επιταχύνει την έρευνα και την πρόοδο στα επιχειρηματικά μοντέλα (Zaslavsky et al., 2012:4-5).

2.3. Data Center και Big Data

Στα Big Data το Κέντρο Δεδομένων (Data Center) δεν αποτελεί μόνο μια πλατφόρμα για συγκέντρωση και αποθήκευση των δεδομένων αλλά αναλαμβάνει επίσης περισσότερες ευθύνες, όπως η απόκτησή τους, η κατάλληλη διαχείριση, η οργάνωση των δεδομένων και η αξιοποίησή τους. Τα Data Center οργανώνουν και διαχειρίζονται τα δεδομένα με κεντρικό στόχο την ανάκτησή τους σε κατάλληλο χρόνο παρέχοντας ουσιαστική πληροφόρηση. Η εμφάνιση των Big Data προσφέρει σημαντικές ευκαιρίες ανάπτυξης και έχει δημιουργήσει νέες προκλήσεις για τα Data Center. Τα Big Data αποτελούν μια χαρακτηριστική περίπτωση που μπορεί να προωθήσει την εκρηκτική ανάπτυξη των υποδομών και του λογισμικού που σχετίζονται με τα κέντρα δεδομένων (Sun et al., 2013:15).

Τα Big Data απαιτούν από τα κέντρα δεδομένων να παρέχουν ισχυρή παρασκηνιακή υποστήριξη. Τα Big Data έχουν αυστηρότερες απαιτήσεις αναφορικά με την αποθήκευση, την επεξεργασία και την διάθεση των πληροφοριών στο δίκτυο. Οι επιχειρήσεις πρέπει να αναλάβουν την ανάπτυξη κέντρων δεδομένων λαμβάνοντας υπόψη τη βελτίωση της ικανότητας ταχείας και αποτελεσματικής επεξεργασίας των Big Data και στο πλαίσιο βελτίωσης της σχέσης τιμής – απόδοσης. Το κέντρο δεδομένων παρέχει τη στήριξη με ένα μεγάλο αριθμό κόμβων, την κατασκευή ενός εσωτερικού δικτύου υψηλής ταχύτητας και αντίγραφα ασφάλειας δεδομένων.

Η ανάπτυξη των Big Data επιταχύνει την επανάσταση και την καινοτομία των κέντρων δεδομένων. Πολλές εφαρμογές Big Data έχουν αναπτύξει μοναδικές αρχιτεκτονικές και προωθούν άμεσα την ανάπτυξη της αποθήκευσης, δικτύων και υπολογιστικών τεχνολογιών που σχετίζονται με το κέντρο δεδομένων. Με τη συνεχή αύξηση του όγκου των δομημένων και αδόμητων δεδομένων, καθώς και η ποικιλία των πηγών των αναλυτικών δεδομένων, οι δυνατότητες επεξεργασίας δεδομένων και πληροφορικής του κέντρου δεδομένων θα πρέπει να ενισχυθεί σε μεγάλο βαθμό. Τέλος, τα Big Data προσδίδουν περισσότερες λειτουργίες στο κέντρο δεδομένων. Τα κέντρα δεδομένων θα πρέπει να διαθέτουν την απαραίτητη τεχνολογία ώστε να εξασφαλίζουν τη συγκέντρωση, την οργάνωση, την ανάλυση και την εφαρμογή των Big Data. Τα κέντρα δεδομένων μπορούν να βοηθήσουν το προσωπικό των επιχειρήσεων να αναλύσουν τα υπάρχοντα δεδομένα, να ανακαλύψουν προβλήματα

στη λειτουργία των επιχειρήσεων και να αναπτύξουν λύσεις από τα Big Data (Chui & Manyika, 2010:6-10; Chen et al., 2014:178-179).

2.4. Τεχνολογίες επεξεργασίας Big Data

Υπάρχουν διάφορες τεχνολογίες που μπορούν να χρησιμοποιηθούν στην επεξεργασία των Big Data. Οι περισσότερες τεχνολογίες είναι προεκτάσεις των έργων Apache και είναι κατασκευασμένα με Hadoop. Το Hadoop παρέχει την ικανότητα επεξεργασίας μεγάλου όγκου δεδομένων ανεξάρτητα από τη δομή τους ενώ αποτελείται από δύο βασικά έργα το Hadoop Distributed File System και το MapReduce. Όπως σημειώνουν οι Shvachko et al. (2010:2-6) το HDFS είναι σχεδιασμένο κατά τέτοιο τρόπο ώστε να αποθηκεύει και να επεξεργάζεται μεγάλο όγκο δεδομένων σε μεγάλο εύρος ζώνης για εφαρμογές χρηστών. Στις εφαρμογές HDFS, τα αρχεία που έχουν συνταχθεί μία φορά και προσπελαστεί πολλές φορές παρουσιάζουν αυξημένη συνέπεια και συνοχή ενώ είναι εύκολα προσβάσιμα. Εκτός από την επεξεργασία των δεδομένων, το HDFS είναι σε θέση να εντοπίζει και να διαχειρίζεται αποτυχίες κατά την εφαρμογή. Ο στόχος αυτός επιτυγχάνεται μέσω ενός οργανωμένου μηχανισμού αντιγραφής αρχείων (Turkington, 2013:29-30).

Οι διαθέσιμες τεχνολογίες επεξεργασίας των Big Data είναι πολυάριθμες καλύπτοντας διαφορετικές ανάγκες. Ωστόσο, είναι χρήσιμο πέραν των παραπάνω να αναφερθούν ορισμένες που ξεχωρίζουν. Αρχικά, αναφέρεται το MapReduce που είχε τεθεί σε εφαρμογή από την Google για την επίλυση του προβλήματος δημιουργίας ευρετηρίου αναζήτησης. Πλέον αποτελεί βασικό μοντέλο προγραμματισμού και υλοποίησης που συνδέεται με την επεξεργασία και τη δημιουργία μεγάλων συνόλων δεδομένων. Η μορφή των δεδομένων που εισέρχονται καθορίζονται από το χρήστη και είναι δομημένα ή αδόμητα (Dean & Ghemawat, 2008:109; Ranger et al., 2007:15-17). Πέραν του Hadoop και του MapReduce, υπάρχουν μια σειρά τεχνολογιών που βασίζονται σ' αυτά και τα οποία ταξινομούνται ανάλογα με τις δυνατότητές τους. Ειδικότερα, υπάρχουν τεχνολογίες που εστιάζουν στην αποθήκευση και διαχείριση των δεδομένων (π.χ. Cloudera Manager, RCFile), στις δυνατότητες της βάσης δεδομένων (π.χ. Oracle NoSQL, Apache HBase, Apache Cassandra, Apache ZooKeeper), στις δυνατότητες επεξεργασίας (Pig, Chukwa) και σε δυνατότητες

ενοποίησης των δεδομένων (Apache Sqoop, Flume) (Emani et al., 2015:73; Yaqoob et al., 2016:1235-1236; Singh & Singla, 2015:3-4; De Mauro, 2015:99).

Επίσης, η NoSQL είναι μια άλλη τεχνολογία που χρησιμοποιείται ευρέως για να χειριστεί τα Big Data ως επί το πλείστον αδόμητα δεδομένα. Παρέχει σημαντική ευελιξία και επεκτασιμότητα ενώ ο όγκος των δεδομένων είναι πολύ εύκολο να είναι διαχειρίσιμος από βάσεις δεδομένων NoSQL. Κύριες βάσεις δεδομένων με NoSQL τεχνολογία είναι οι οριοθετημένες με τη χρήση εγγράφων, βάσεις δεδομένων γραφημάτων και οι εκτεταμένες σχεσιακές βάσεις δεδομένων (Singh & Singla, 2015:9-10).

Ο Skytree Server χρησιμοποιείται για την επεξεργασία μεγάλων ποσοτήτων δεδομένων σε υψηλές ταχύτητες. Είναι φιλικό προς το χρήστη και παρέχει μια διεπαφή γραμμής εντολών, όπου οι χρήστες μπορούν να εισάγουν εντολές. Το επίκεντρο του Skytree Server είναι η ανάλυση των δεδομένων σε πραγματικό χρόνο με τη χρήση των κατάλληλων αλγορίθμων. Η συγκεκριμένη τεχνολογία μπορεί να χειριστεί σχεσιακές βάσεις δεδομένων, επίπεδα αρχεία, δομημένα και αδόμητα δεδομένα. Παρά τα πολλά πλεονεκτήματα στην επεξεργασία των Big Data όπως η υψηλή απόδοση, η αυξημένη πολυπλοκότητα αποτελεί έναν βασικό περιορισμό (Han et al., 2011:364).

Τέλος, το Jaspersoft χρησιμοποιείται για να παράγει εκθέσεις από βάση δεδομένων σε στήλες. Αποτελεί μια επεκτάσιμη πλατφόρμα, παρέχει γρήγορη οπτικοποίηση των δεδομένων ενώ μια από τις εξαιρετικές δυνατότητες της συγκεκριμένης τεχνολογίας είναι να επεξεργάζεται δεδομένα σε πολύ μικρό χρόνο χωρίς να καθίσταται αναγκαία η διαδικασία ETL. Η συγκεκριμένη τεχνολογία επεξεργάζεται μεγάλο όγκο δεδομένων μέσω της HTML απεικόνισης ενώ μπορεί να ενσωματωθεί και σε άλλες εφαρμογές των χρηστών. Παρά τα πολλά πλεονεκτήματα του Jaspersoft, όπως η χαμηλή τιμή, η εύκολη εγκατάσταση και η μεγάλη λειτουργικότητα, υπάρχουν ορισμένα μειονεκτήματα όπως αστοχίες τεκμηρίωσης που μπορεί να μην επιφέρουν επιθυμητά αποτελέσματα (Wayner, 2012).

2.5. Μέθοδοι επεξεργασίας Big Data

Ιδιαίτερο ενδιαφέρον έχει ανακύψει αναφορικά με τον τρόπο όπου εξάγονται πληροφορίες από τα Big Data. Οι μέθοδοι επεξεργασίας που χρησιμοποιούνται για τα Big Data είναι οι ακόλουθες:

1. Κατακερματισμός (Hashing). Σε ένα μεγάλο όγκο δεδομένων όπως τα Big Data, η ανάκτηση των πληροφοριών μέσω ευρετηρίου δεν έχει πάντοτε τα επιθυμητά αποτελέσματα. Ο κατακερματισμός των δεδομένων μπορεί να αποτελέσει μια ιδιαίτερα αποτελεσματική μέθοδο ανάκτησης των δεδομένων χωρίς τη χρήση ευρετηρίου. Η συγκεκριμένη τεχνική βασίζεται στη δημιουργία μιας αξίας σε αντίστοιχο κλειδί. Ένα πλεονέκτημα του κατακερματισμού είναι η ταχεία ανάγνωση των δεδομένων ενώ ως μέθοδος είναι ακατάλληλη όταν τα δεδομένα είναι οργανωμένα σε σειρά. Ο κατακερματισμός αποδίδει καλύτερα όταν τα δεδομένα είναι διακριτά και τυχαία. Παρά τα πολλά πλεονεκτήματα του κατακερματισμού, όπως η γρήγορη ανάγνωση των δεδομένων, υπάρχουν μειονεκτήματα όπως η υψηλή πολυπλοκότητα και η γραμμική ανίχνευση (Yaqoob et al., 2016:1238).
2. Ευρετηρίαση. Ο ταχύς εντοπισμός δεδομένων πολλές φορές επιτυγχάνεται μέσω της ευρετηρίασης. Στα Big Data και αξιοποιώντας το ευρετήριο οι πληροφορίες μπορούν να ομαδοποιούνται ή να δημιουργούνται συνολικά αρχεία πληροφοριών. Το μειονέκτημα της συγκεκριμένης μεθόδου είναι το αυξημένο κόστος ενώ σε στάδιο ανάπτυξης είναι σημαντικά συστήματα ευρετηρίασης (π.χ. VegaIndexer, sksOpen, Cintia) (Gani et al., 2016:243; Funaki et al., 2015:150).
3. Φίλτρο. Η χρήση του κατάλληλου φίλτρου αναζητήσεων επιτρέπει την εξοικονόμηση χώρου αποθήκευσης δεδομένων και αστοχίας εύρεσης λανθασμένης πληροφορίας. Η συγκεκριμένη μέθοδος παρουσιάζει σημαντικά πλεονεκτήματα όπως η ταχεία εύρεση πληροφορίας και η υψηλή απόδοση στο διαθέσιμο αποθηκευτικό χώρο, ωστόσο περιορίζει σημαντικά τις δυνατότητες διαγραφών (De Mauro et al., 2015:4-5; Yaqoob et al., 2016:1239).
4. Παράλληλη υπολογιστική υποδομή. Η χρήση παράλληλων υπολογιστικών δομών βελτιώνει τις δυνατότητες επεξεργασίας των δεδομένων. Στη συγκεκριμένη μέθοδο, ένα σύνθετο πρόβλημα χωρίζεται σε πολλαπλά μέρη τα

οποία επεξεργάζονται ταυτόχρονα ενώ με τη βοήθεια των παραλληλισμών προκύπτει η διαθέσιμη πληροφορία (De Mauro et al., 2015:4-5; Yaqoob et al., 2016:1239).

3. Υλοποίηση Big Data

3.1. Η δημιουργία των δεδομένων (data generation)

Η δημιουργία των δεδομένων (data generation) είναι το πρώτο βήμα για τα Big Data. Λαμβάνοντας υπόψη τα δεδομένα του διαδικτύου ως παράδειγμα, προκύπτει ότι δημιουργείται μια τεράστια ποσότητα δεδομένων από αναζητήσεις, καταχωρήσεις, δημοσιεύσεις, αρχεία και μηνύματα σε ιστοσελίδες και blog. Τα στοιχεία αυτά είναι στενά συνδεδεμένα με την καθημερινή ζωή των ανθρώπων και έχουν χαρακτηριστικά υψηλής και χαμηλής αξίας. Τα εν λόγω δεδομένα στο διαδίκτυο μπορεί να είναι μεμονωμένα και άνευ αξίας αλλά, μέσω της αξιοποίησης με τα Big Data μπορεί να προσφέρουν χρήσιμες πληροφορίες σχετικά με συνήθειες, χαρακτηριστικά, χόμπι κλπ. των χρηστών ενώ υπό προϋποθέσεις μπορούν να προβλέψουν ακόμα και τη συναισθηματική τους διάθεση (Howe1 et al., 2008:48-49).

Επιπλέον, δημιουργούνται μέσω διαμηκών ή και καταναμεμημένων πηγών δεδομένων, σύνολα μεγαλύτερης κλίμακας, ποικιλίας και πολυπλοκότητας. Τέτοιες πηγές δεδομένων περιλαμβάνουν βίντεο, κλικ χρηστών, μηνύματα κλπ. Προς το παρόν, οι κύριες πηγές των Big Data για τη δημιουργία των δεδομένων (data generation) είναι οι επιχειρήσεις, το διαδίκτυο, τα δεδομένα της επιστημονικής έρευνας, οι αναζητήσεις των χρηστών του διαδικτύου και τα διάφορα αρχεία (ήχου, εικόνας, κειμένου κλπ) που παράγονται online και offline (Agrawal et al., 2011:531). Πλέον, οι εξελίξεις σε τεχνολογικό επίπεδο και η πληροφορική έχουν καταστήσει περισσότερο αναγκαία από ποτέ τη δημιουργία Big Data στη βάση ότι εξελίσσονται συνεχώς και οι πηγές άντλησης νέων δεδομένων. Η δημιουργία των δεδομένων (data generation) αποτελεί τη βάση σύνθεσης του οικοδομήματος των Big Data απαιτώντας αυξημένη προσοχή στη συγκέντρωσή τους. Ειδικότερα, για τη δημιουργία των δεδομένων (data generation) αξιοποιούνται τα εξής:

1. **Δεδομένα επιχειρήσεων.** Τα εσωτερικά δεδομένα των επιχειρήσεων είναι η κύρια πηγή δημιουργίας Big Data. Τα δεδομένα των επιχειρήσεων είναι αποτέλεσμα των εμπορικών της συναλλαγών, της παραγωγής, στοιχείων αποθεμάτων, δεδομένα πωλήσεων, οικονομικά στοιχεία και δεδομένα με γνώμονα τις δραστηριότητες. Εκτιμάται ότι ο όγκος των επιχειρηματικών δεδομένων παγκοσμίως διπλασιάζεται κάθε 1 - 2 χρόνια, καθιστώντας

περισσότερο αναγκαία την αποτελεσματική σε πραγματικό χρόνο ανάλυση των δεδομένων (Manyika et al., 2011:64-66; Gantz & Reinsel, 2010:6).

2. **Internet of Things (IoT).** Όπως προαναφέρθηκε το IoT είναι μια σημαντική πηγή των Big Data. Αξιοποιώντας τις διεργασίες απόκτησης και μετάδοσης των δεδομένων στο IoT, η αρχιτεκτονική του δικτύου μπορεί να διαστρωματωθεί ως εξής: ανίχνευση, συγκέντρωση και αξιοποίηση των δεδομένων. Τα δεδομένα που δημιουργούνται από το IoT έχουν μεγάλη κλίμακα, ετερογένεια, ισχυρή συσχέτιση χώρου και χρόνου.
3. **Βιοϊατρικά δεδομένα.** Στοιχεία βιοϊατρικών καταγραφών, μετρήσεις, ερευνητικά στοιχεία στον τομέα της βιοϊατρικής έχουν εισέλθει στην εποχή των Big Data.
4. **Διάφορες πηγές.** Η δημιουργία των Big Data βασίζεται σε στοιχεία επιστημονικών εφαρμογών, στοιχεία αστρονομίας, φυσικών καταγραφών, εμπορικών συναλλαγών, διαδίκτυο κλπ. Τα στοιχεία από διαφορετικές πηγές παρουσιάζουν αυξημένη ετερογένεια ενώ είναι πολλές φορές αρκετά πολύπλοκα (Bhatotia et al., 2011:7).

3.2. Η απόκτηση δεδομένων (data acquisition)

Το επόμενο βήμα στη δημιουργία των Big Data είναι η απόκτηση των δεδομένων (data acquisition) που περιλαμβάνει τη συλλογή, τη μετάδοση και την προεπεξεργασία των δεδομένων. Κατά τη διάρκεια της συλλογής των δεδομένων, θα πρέπει ταυτόχρονα να εφαρμοστεί και κατάλληλος μηχανισμός αποτελεσματικής μετάδοσης για να μεταφερθούν τα δεδομένα σε ένα κατάλληλο σύστημα διαχείρισης και αποθήκευσης που υποστηρίζει τις διαφορετικές αναλυτικές εφαρμογές. Η συγκέντρωση δεδομένων μπορεί να περιλαμβάνει μερικές φορές πολύ περιττά ή άχρηστα δεδομένα, γεγονός που αυξάνει άσκοπα το χώρο αποθήκευσης και επηρεάζει τη μετέπειτα ανάλυση των δεδομένων. Ως εκ τούτου, καθίστανται αναγκαίες ενέργειες επεξεργασίας των δεδομένων για να εξασφαλιστεί η αποτελεσματική αποθήκευση και αξιοποίησή τους.

3.2.1. Συλλογή δεδομένων (data collection)

Για τη συλλογή των αρχικών δεδομένων χρησιμοποιούνται ειδικές τεχνικές που κυρίως εστιάζουν στις ακόλουθες μεθόδους:

1. **Σύνδεση αρχείων.** Ως μία ευρέως χρησιμοποιούμενη μέθοδος συλλογής δεδομένων είναι η σύνδεσή τους σε επιμέρους αρχεία καταγραφής. Χρησιμοποιούνται σχεδόν σε όλες τις ψηφιακές συσκευές ενώ παρέχουν πληροφόρηση αναφορικά με τη δραστηριότητα του χρήστη (αριθμός κλικ, επισκέψεις ιστοσελίδων κ.ά.) (Nanopoulos et al., 2002:65).
2. **Ανίχνευση.** Στην καθημερινότητα χρησιμοποιούνται αισθητήρες για τη μέτρηση φυσικών δεδομένων και τη μετατροπή τους σε αναγνώσιμα ψηφιακά σήματα για περαιτέρω επεξεργασία και αποθήκευση. Δεδομένα των αισθήσεων όπως το ηχητικό κύμα, η φωνή, το ρεύμα αέρα, η θερμοκρασία, κλπ. μεταφέρονται σε ένα σημείο συλλογής δεδομένων μέσω ενσύρματων ή ασύρματων δικτύων. Τα τελευταία χρόνια, τα ασύρματα δίκτυα αισθητήρων έχουν λάβει μεγάλο ενδιαφέρον και έχουν εφαρμοστεί σε πολλές εφαρμογές, όπως η περιβαλλοντική έρευνα, η παρακολούθηση της ποιότητας του νερού, η κατασκευή ενός έργου, η παρακολούθηση της πανίδας κλπ. (Selavo et al., 2007:104; Barrenetxea et al., 2008:335).
3. **Μέθοδοι απόκτησης δεδομένων δικτύου.** Προς το παρόν, το δίκτυο στην απόκτηση δεδομένων αξιοποιείται μέσω του web crawler. Το web crawler προβαίνει σε καταγραφές δεδομένων του χρήστη, σε διευθύνσεις που επισκέπτεται ενώ η διαδικασία επαναλαμβάνεται μέχρι να διακοπεί η χρήση του δικτύου. Η απόκτηση δεδομένων μέσω ενός web crawler εφαρμόζεται ευρέως σε εφαρμογές που βασίζονται σε ιστοσελίδες, όπως οι μηχανές αναζήτησης ή web caching (Dincturk et al., 2014:3-4).

3.2.2. Μεταφορά δεδομένων (data transportation)

Με την ολοκλήρωση της συλλογής των ανεπεξέργαστων δεδομένων, μεταφέρονται στην κατάλληλη υποδομή αποθήκευσης δεδομένων για την μετέπειτα επεξεργασία και ανάλυση. Τα Big Data αποθηκεύονται κυρίως στα κέντρα δεδομένων ενώ η διάταξή τους θα πρέπει να είναι τέτοια ώστε να βελτιώνεται η υπολογιστική απόδοση και να διευκολύνεται η συντήρηση του υλικού. Με άλλα

λόγια, μπορεί να ακολουθηθεί εσωτερική μεταφορά δεδομένων στο κέντρο δεδομένων. Η μετάδοση δεδομένων αποτελείται από δύο φάσεις: τις Inter - DCN και τις Intra - DCN μεταφορές. Οι Inter - DCN μεταφορές βασίζονται στην ύπαρξη κέντρου δεδομένων από το οποίο γίνονται μέσω φυσικού δικτύου οι όποιες μεταφορές. Λόγω της ραγδαίας ανάπτυξης της ζήτησης της μεταφοράς δεδομένων, η φυσική υποδομή δικτύου στις περισσότερες περιοχές σε όλο τον κόσμο αποτελούνται από υψηλής ποιότητας συνδέσεις. Η χρήση συνδέσεων οπτικών ινών αποτελεί μια σύγχρονη επιλογή μεταφοράς δεδομένων με ταχύτητες που πρακτικά μπορεί να αγγίζουν τα 40GB/s με στόχο τα 100Gb/s στο εγγύς μέλλον (Jinno et al., 2009:1-3).

Οι Intra - DCN μεταφορές είναι οι ροές επικοινωνίας δεδομένων εντός των κέντρων δεδομένων. Οι Intra - DCN μεταφορές εξαρτώνται από τον μηχανισμό επικοινωνίας μέσα στο κέντρο δεδομένων, δηλαδή από τη φυσική σύνδεση της μνήμης, της αρχιτεκτονικής δικτύου, των πρωτοκόλλων κλπ. Η εσωτερική σύνδεση στα περισσότερα κέντρα δεδομένων είναι, δύο στρωμάτων ή δομής τριών επιπέδων (Barroso & Holzle, 2009:65)

3.2.3. Προεπεξεργασία των δεδομένων (data pre - processing)

Πριν τη δημιουργία των Big Data είναι ιδιαίτερα σημαντικό να διενεργείται προεπεξεργασία των δεδομένων (data pre - processing). Εξαιτίας της ευρείας ποικιλίας πηγών δεδομένων, τα συλλεγόμενα δεδομένα παρουσιάζουν σημαντικές διαφοροποιήσεις καθιστώντας πολλές φορές την αποθήκευσή τους μη ουσιαστική. Επιπλέον, σε ορισμένες περιπτώσεις η ποιότητά τους είναι χαμηλή καθιστώντας μη αποτελεσματική την ανάλυση των δεδομένων. Η προεπεξεργασία των δεδομένων συμβάλλει όχι μόνο στην ενσωμάτωση δεδομένων από διαφορετικές πηγές αλλά επίσης βελτιώνει την ακρίβεια της ανάλυσης.

Οι τεχνικές προεπεξεργασίας των δεδομένων έχουν μεγάλο εύρος και επιλέγονται στη βάση δημιουργίας των Big Data. Χαρακτηριστικά, η μέθοδος ενσωμάτωσης (integration) περιλαμβάνει το συνδυασμό των δεδομένων από διαφορετικές πηγές παρέχοντας ομοιόμορφη προβολή των δεδομένων (Lenzerini, 2002:241). Μέθοδος κατάλληλης προεπεξεργασίας των δεδομένων είναι και εκείνη του καθαρισμού (cleaning) όπου εντοπίζονται ελλιπή ή παράλογα δεδομένα και στη συνέχεια τροποποιούνται ή διαγράφονται για τη βελτίωση της ποιότητας των δεδομένων. Ο καθαρισμός των δεδομένων είναι ζωτικής σημασίας για να διατηρηθεί η συνοχή των

δεδομένων, η οποία εφαρμόζεται ευρέως σε πολλούς τομείς, όπως οι τραπεζικές και ασφαλιστικές υπηρεσίες, ο κλάδος του λιανικού εμπορίου, οι τηλεπικοινωνίες, ο έλεγχος της κυκλοφορίας κλπ. (Maletic & Marcus, 2000:203).

3.3. Συστήματα αποθήκευσης των Big Data

Η εκρηκτική ανάπτυξη των δεδομένων έχει οδηγήσει σε πιο αυστηρές απαιτήσεις σε θέματα αποθήκευσης και διαχείρισης των δεδομένων. Η αποθήκευση των Big Data επικεντρώνεται και στη διαχείριση των συνόλων δεδομένων μεγάλης κλίμακας, επιτυγχάνοντας παράλληλα την αξιοπιστία και τη διαθεσιμότητα της πρόσβασης. Από τη μία πλευρά, η υποδομή αποθήκευσης πρέπει να παρέχει πληροφορίες αποθήκευσης και από την άλλη να εξασφαλίζει την άμεση και σωστή πρόσβαση. Παραδοσιακά, για την αποθήκευση των Big Data χρησιμοποιείται βοηθητικός εξοπλισμός αποθήκευσης δεδομένων. Η αύξηση των δεδομένων έχει ανάλογα αυξήσει και τις ανάγκες για συσκευές αποθήκευσης δεδομένων επιδιώκοντας τη δημιουργία όλο και μεγαλύτερων μονάδων αποθήκευσης. Για την αποθήκευση δεδομένων χρησιμοποιούνται συστήματα όπως:

1. **Σύστημα αποθήκευσης ογκωδών δεδομένων. Διάφορα συστήματα αποθήκευσης αναδύονται για να ανταποκριθούν στις απαιτήσεις των μαζικών δεδομένων.** Οι υφιστάμενες μαζικές τεχνολογίες αποθήκευσης μπορεί να είναι άμεσης αποθήκευσης (Direct Attached Storage, DAS) και αποθήκευσης στο δίκτυο (Network Attached Storage, NAS ή Storage Area Network, SAN). Το σύστημα αποθήκευσης στα Big Data επιλέγεται στη βάση του σκοπού και των εκάστοτε αναγκών. Ειδικότερα, τα συστήματα DAS είναι κατάλληλα για διασύνδεση με servers σε μικρή κλίμακα, καθώς λόγω της χαμηλής επεκτασιμότητας του συστήματος υπάρχει ενδεχόμενο η αναβάθμιση του χώρου αποθήκευσης να είναι δύσκολη. Έτσι, το σύστημα DAS χρησιμοποιείται κυρίως σε προσωπικούς υπολογιστές και μικρού μεγέθους servers. Αντίθετα, το σύστημα NAS χαρακτηρίζεται από υψηλή επεκτασιμότητα καθώς στην πραγματικότητα λειτουργεί ως βοηθητικός εξοπλισμός αποθήκευσης ενός δικτύου. Είναι άμεσα συνδεδεμένος με το δίκτυο μέσω ενός hub ή μέσω εναλλαγής πρωτοκόλλων TCP/IP. Στο σύστημα NAS, τα δεδομένα μεταφέρονται με τη μορφή αρχείων. Στο σύστημα SAN, η

διαχείριση αποθήκευσης δεδομένων είναι σχετικά ανεξάρτητη εντός του τοπικού δικτύου όπου αξιοποιούνται πολλαπλές διαδρομές στη μεταφορά δεδομένων (Rimal & Lumb, 2009:43-44).

2. **Κατανεμημένα συστήματα αποθήκευσης.** Στη χρήση κατανεμημένου συστήματος αποθήκευσης δεδομένων θα πρέπει να λαμβάνεται υπόψη η σύνδεση των server για αποφυγή αποτυχιών επικοινωνίας και ασυνέπειας μεταξύ διαφορετικών αντιγράφων των ίδιων δεδομένων, να εξασφαλίζεται η διαθεσιμότητα και να διαθέτει κατάλληλο επίπεδο ανοχής για τα προβλήματα που προκαλούνται από βλάβες του δικτύου. Τα συγκεκριμένα συστήματα θεωρούνται γενικά ιδιαίτερα ασφαλή και χρήσιμα στην αποθήκευση δεδομένων των Big Data ενώ επιτυγχάνουν υψηλά επίπεδα συνοχής (Chang et al., 2008:1-5).

3.4. Μηχανισμός αποθήκευσης Big Data

Η αποθήκευση των Big Data έχει οδηγήσει στην ανάπτυξη κατάλληλων μηχανισμών αποθήκευσης για μεγάλο όγκο δεδομένων. Οι υφιστάμενοι μηχανισμοί αποθήκευσης των Big Data μπορούν να ταξινομηθούν σε τρία διαφορετικά επίπεδα από κάτω προς τα πάνω: α) τα συστήματα αρχείων, β) βάσεις δεδομένων και γ) μοντέλα προγραμματισμού. Τα συστήματα αρχείων είναι η βάση για τα ανώτερα επίπεδα εξασφαλίζοντας τη λειτουργία του μηχανισμού αποθήκευσης. Το Google's GFS είναι ένα χαρακτηριστικό παράδειγμα μηχανισμού αποθήκευσης αρχείων το οποίο μπορεί να υποστηρίξει εφαρμογές υψηλών απαιτήσεων. Οι μηχανισμοί αποθήκευσης σε πολλές περιπτώσεις χαρακτηρίζονται από υψηλές επιδόσεις υποστηρίζοντας εφαρμογές μεγάλης κλίμακας. Ωστόσο, μπορεί να προκύψουν αδυναμίες και περιορισμοί όπως η δυσκολία παρουσίασης δεδομένων μικρής κλίμακας. Οι συγκεκριμένοι περιορισμοί σε μεγάλο βαθμό έχουν αντιμετωπιστεί με την ανάπτυξη μηχανισμών όπως ο Colossus διαδόχου του GFS (McKusick & Quinlan, 2009:11).

Επιπλέον, άλλες εταιρείες και οι ερευνητές έχουν επίσης αναφέρει συστήματα τα οποία μπορούν να ανταποκριθούν στις διαφορετικές απαιτήσεις για την αποθήκευση των Big Data. Για παράδειγμα, το HDFS και το Kosmosfs είναι αποτέλεσμα του ανοικτού κώδικα του GFS. Η Microsoft ανέπτυξε το Cosmos για την υποστήριξη της

αναζήτησης και της διαφήμισης των επιχειρήσεων ενώ το Facebook χρησιμοποιεί το Haystack για να αποθηκεύει το μεγάλο όγκο των φωτογραφιών μικρού μεγέθους. Γενικά, τα κατανεμημένα συστήματα αρχείων έχουν αναπτυχθεί ιδιαίτερα τα τελευταία χρόνια ενώ και οι σχετικές τεχνολογίες τυγχάνουν ανάπτυξης (Cattell, 2011:14-15; Chaiken et al., 2008:1266; Beaver et al., 2010:2-3).

3.4.1. Τεχνολογίες αποθήκευσης δεδομένων

Η τεχνολογία αποθήκευσης δεδομένων τις τελευταίες δεκαετίες έχει μεταβληθεί σημαντικά. Πλέον, τα διάφορα συστήματα βάσεων δεδομένων που έχουν αναπτυχθεί είναι σε θέση να χειρίζονται σύνολα δεδομένων σε διαφορετικές κλίμακες υποστηρίζοντας ταυτόχρονα διάφορες εφαρμογές. Οι παραδοσιακές σχεσιακές βάσεις δεδομένων δεν μπορούν πλέον να ανταποκριθούν στις προκλήσεις σχετικά με τις κατηγορίες και τις κλίμακες που επέφεραν τα Big Data (μεγάλα δεδομένα). Οι βάσεις δεδομένων NoSQL (μη παραδοσιακές σχεσιακές βάσεις δεδομένων) γίνονται όλο και πιο δημοφιλής για τα Big Data. Οι NoSQL διαθέτουν ευέλικτα χαρακτηριστικά και στοιχεία πλήρους υποστήριξης των Big Data μετατρέποντάς τις σε βασική τεχνολογία για τα Big Data. Οι σημαντικότερες βάσεις δεδομένων που χρησιμοποιούνται στην τεχνολογία NoSQL είναι οι εξής:

1. **Βάσεις δεδομένων κλειδιού - τιμής (key - value databases).** Οι βάσεις δεδομένων κλειδιού - τιμής αποτελούν ένα απλό μοντέλο δεδομένων με τα δεδομένα που αποθηκεύονται να αντιστοιχούν σε βασικές τιμές. Οι συγκεκριμένες βάσεις δεδομένων διαθέτουν απλή δομή και χαρακτηρίζονται από υψηλή επεκτασιμότητα και μικρότερο χρόνο απόκρισης ερωτήματος σε σύγκριση με τις σχεσιακές βάσεις δεδομένων (DeCandia et al., 2007:207-208).
2. **Βάση προσανατολισμένη σε στήλες.** Η βάση δεδομένων σε στήλες αποθηκεύει και επεξεργάζεται τα δεδομένα σύμφωνα με τις στήλες εκτός από σειρές. Οι συγκεκριμένες βάσεις δεδομένων είναι εμπνευσμένες κυρίως από το Google Big Table. Τα δεδομένα είναι κατανεμημένα και δομημένα σε ένα σύστημα αποθήκευσης δεδομένων, το οποίο έχει σχεδιαστεί για την επεξεργασία δεδομένων μεγάλης κλίμακας (Chang et al., 2008:4-6).
3. **Βάση δεδομένων εγγράφων.** Η συγκεκριμένη τεχνολογία μπορεί να υποστηρίξει πιο σύνθετες μορφές δεδομένων. Οι βασικότεροι εκπρόσωποι

των συστημάτων αποθήκευσης του εγγράφου είναι το MongoDB, το SimpleDB και το CouchDB (Chen et al., 2014:187-189).

3.5. Η ανάλυση των Big Data

Η ανάλυση των Big Data αφορά κυρίως σε μεθόδους ανάλυσης για τα παραδοσιακά και τα μεγάλα δεδομένα, σε αναλυτική αρχιτεκτονική για τα μεγάλα δεδομένα και σε αξιοποίηση του κατάλληλου λογισμικού. Η ανάλυση των δεδομένων είναι το τελευταίο και το πιο σημαντικό στάδιο στην αλυσίδα των Big Data. Η ανάλυση των Big Data καλύπτει μια ευρεία περιοχή η οποία συνεχώς μεταβάλλεται και είναι εξαιρετικά πολύπλοκη. Για την ανάλυση των Big Data αξιοποιούνται μέθοδοι, αρχιτεκτονικές και εργαλεία παράγοντας διαφορετικές πληροφορίες. Στην παραδοσιακή ανάλυση των δεδομένων απαιτείται να εφαρμόζονται οι κατάλληλες στατιστικές μέθοδοι ώστε να μεγιστοποιείται η αξία των πληροφοριών. Η ανάλυση των δεδομένων διαδραματίζει σημαντικό ρόλο στη λήψη αποφάσεων, στην κατανόηση απαιτήσεων και στον καθορισμό προβλέψεων. Για την ανάλυση των Big Data μπορούν να χρησιμοποιηθούν παραδοσιακές μέθοδοι με τις στατιστικές να ξεχωρίζουν (πχ. cluster analysis, factor analysis, correlation analysis, regression analysis, data mining algorithms) (Emani et al., 2015:74; Chen et al., 2014:190-191).

Πέραν της παραδοσιακής ανάλυσης, ανάλογα με τα δεδομένα μπορεί να ακολουθηθούν διαφορετικές προσεγγίσεις ανάλυσης. Ειδικότερα, ανάλογα με τις απαιτήσεις της επικαιρότητας, η ανάλυση των Big Data μπορεί να γίνει σε πραγματικό χρόνο (on line analysis) και σε μη πραγματικό (offline analysis). Η ανάλυση των Big Data μπορεί να ταξινομηθεί ανάλογα με το επίπεδο μνήμης, την επιχειρηματική δραστηριοποίηση αλλά και σε μαζικό επίπεδο. Τέλος, η ανάλυση των Big Data μπορεί να γίνει στη βάση διαφορετικής πολυπλοκότητας των αλγορίθμων που χρησιμοποιούνται (Emani et al., 2015:74; Chen et al., 2014:190-191).

3.5.1. Εργαλεία άντλησης και ανάλυσης δεδομένων

Η άντληση και η ανάλυση των δεδομένων Big Data γίνεται με την αξιοποίηση εργαλείων και λογισμικών, τόσο επαγγελματικών όσο και ερασιτεχνικών, ανοικτού και κλειστού κώδικα. Μεταξύ των σημαντικότερων εργαλείων άντλησης και ανάλυσης δεδομένων αναφέρονται τα ακόλουθα:

1. **R.** Το R είναι μια γλώσσα προγραμματισμού ανοικτού κώδικα με περιβάλλον σχεδιασμένο για άντληση και οπτικοποίηση δεδομένων. Οι εργασίες εκτελούνται με κωδικοποίηση με C, C⁺⁺ και Fortran. Επιπλέον, ειδικευμένοι χρήστες μπορούν να ανακαλούν R δεδομένα σε C. Στην πραγματικότητα, το R είναι μια υλοποίηση της γλώσσας S που όμως είναι ανοικτού κώδικα. Η επιλογή του R έφθασε να ξεπερνά και την SQL μέχρι τα προηγούμενα έτη. Λόγω της δημοτικότητας του R, κατασκευαστές βάσεων δεδομένων όπως η Teradata και η Oracle, έχουν κυκλοφορήσει προϊόντα που υποστηρίζονται από την R (Wang et al., 2016:754-756; Berthold et al., 2008:321-322).
2. **Excel.** Το Excel αποτελεί βασικό εργαλείο του Microsoft Office που παρέχει ισχυρή υποστήριξη στην επεξεργασία δεδομένων και στατιστικών με δυνατότητες ανάλυσης.
3. **Rapid - I Rapidminer.** Το Rapidminer είναι ένα ανοικτό λογισμικό που χρησιμοποιείται για την άντληση δεδομένων, μηχανική μάθηση και προγνωστική ανάλυση. Η άντληση δεδομένων από το Rapidminer περιλαμβάνει την εξαγωγή, την οπτικοποίηση, την αξιολόγηση και την ανάπτυξη. Η ροή άντλησης δεδομένων περιγράφεται σε XML και εμφανίζεται σ' ένα γραφικό περιβάλλον χρήστη (GUI) (Wang et al., 2016:754-756; Berthold et al., 2008:321-322).
4. **KNMINE.** Το KNIME είναι ένα φιλικό και ανοικτό προς το χρήστη εργαλείο ενοποίησης, άντλησης, επεξεργασίας και ανάλυσης δεδομένων. Επιτρέπει στους χρήστες να δημιουργήσουν ροές δεδομένων ή σχετικά κανάλια με έναν ορατό τρόπο. Το KNIME είναι γραμμένο σε Java και παρέχει λειτουργίες, διαμέσου των οποίων οι χρήστες μπορούν αν εισάγουν αρχεία δεδομένων.
5. **Weka/Pentaho.** Το Weka είναι ένα λογισμικό ανοικτού κώδικα για την άντληση δεδομένων γραμμένο σε Java. Το Weka παρέχει λειτουργίες όπως η επεξεργασία δεδομένων, η ταξινόμηση, η ομαδοποίηση, η απεικόνιση, κ.λπ. (Wang et al., 2016:754-756; Berthold et al., 2008:321-322).

3.6. Προβλήματα των Big Data

Πέραν των όποιων ωφελειών προσφέρουν τα Big Data, υπάρχουν μια σειρά προβλημάτων που πρέπει να αντιμετωπιστούν. Το σημαντικότερο πρόβλημα είναι το θέμα της προστασίας της ιδιωτικής ζωής και της ιδιοκτησίας των δεδομένων. Αν και η εποχή των Big Data βρίσκεται ακόμα στην απαρχή της, ωστόσο έχουν σημειωθεί περιπτώσεις παραβίασης της ιδιωτικής ζωής. Το συγκεκριμένο πρόβλημα ανακύπτει από ορισμένες δυσκολίες των υπολογιστικών συστημάτων να ορίσουν τις πληροφορίες που θεωρούνται προσωπικά δεδομένα. Επίσης, υφίσταται δυσκολία στη τήρηση διακριτικότητας σε ορισμένα στοιχεία όπως το φύλο, η ηλικία, οι καταναλωτικές προτιμήσεις κλπ, δημιουργώντας πολλές φορές δυνατότητες επηρεασμού της καταναλωτικής συμπεριφοράς (Hill, 2014).

Τα συστήματα που σχετίζονται με την καταναλωτική συμπεριφορά δεν είναι τα μοναδικά που δημιουργούν μεγάλη ανησυχία για το απόρρητο των δεδομένων. Η υγειονομική περίθαλψη, η κοινωνική δικτύωση και τα κυβερνητικά συστήματα περιέχουν επίσης μεγάλες ποσότητες ευαίσθητων πληροφοριών. Κάθε άτομο στον ανεπτυγμένο κόσμο μπορεί να συνδέεται με τουλάχιστον ένα γεγονός σε μια ηλεκτρονική βάση δεδομένων που ο οποιοσδήποτε θα μπορούσε να χρησιμοποιήσει για εκβιασμό, διακρίσεις, παρενόχληση, οικονομικές βλάβες (Lazer et al., 2014:1203-1204).

Ένα άλλο πρόβλημα με τα Big Data είναι ότι μπορεί να είναι παραπλανητικές οι πληροφορίες που παρουσιάζουν εξαιτίας των αρνητικών συσχετίσεων τους. Ουσιαστικά, παρουσιάζονται πληροφορίες που πολλές φορές δεν είναι αληθείς αλλά παρουσιάζουν μια αληθοφάνεια ή δεν έχουν κανένα απολύτως νόημα που να σχετίζεται με τις αναζητούμενες πληροφορίες. Στα Big Data ορισμένες φορές παρουσιάζονται στρεβλώσεις νοημάτων οι οποίες όχι μόνο δεν συνεισφέρουν στην πληροφόρηση αλλά αποτελούν βασικό εμπόδιο. Τέλος, η λειτουργία των Big Data απαιτεί κατάλληλες υποδομές και υπολογιστικά συστήματα τα οποία καταλαμβάνουν μεγάλο χώρο και απαιτούν αντίστοιχα μεγάλη ισχύ και μεγάλη δικτύωση. Τα συστήματα πληροφορικής αντιπροσωπεύουν σήμερα περίπου το 10% της παγκόσμιας χρήσης ηλεκτρικής ενέργειας. Ο αριθμός αυτός είναι βέβαιο ότι θα αυξηθεί ραγδαία, καθώς όλο και περισσότερο χρησιμοποιούνται αποθηκευτικοί δίσκοι και επεξεργαστές για τα δεδομένα που αποθηκεύονται. Η εκθετική αύξηση των

δεδομένων έχει οδηγήσει σε αύξηση των απαιτήσεων ενέργειας καθιστώντας τα Big Data ως ιδιαίτερα ενεργοβόρα (Hoy, 2014:5-6).

4. Υλοποίηση Big Data Σε Βιβλιοθήκες

4.1. Η εφαρμογή των Big Data από τη Βρετανική Βιβλιοθήκη (British Library)

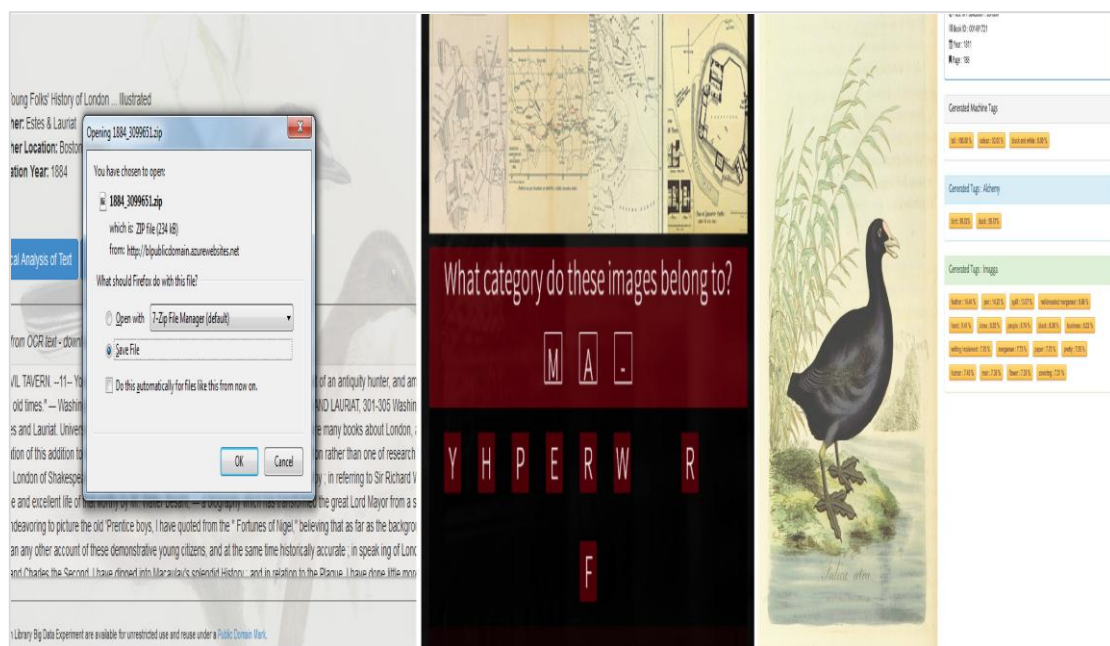
Η συλλογή της Βρετανικής Βιβλιοθήκης διαθέτει περισσότερα από 150 εκατομμύρια βιβλία, εφημερίδες, χάρτες και χειρόγραφα, τα οποία αποτελούν μεγάλο όγκο δεδομένων με βάση οποιοδήποτε κριτήριο. Επιπλέον, η βιβλιοθήκη βρίσκεται στη διαδικασία της ψηφιοποίησης τμήματος αυτής της συλλογής καθιστώντας τα δεδομένα χρήσιμα στην καινοτόμο έρευνα. Ωστόσο, παράλληλα με τις πολλές δυνατότητες του υλικού, τα Big Data παρουσιάζουν επίσης σημαντικές προκλήσεις για τους ερευνητές των κοινωνικών επιστημών (COWLS, 2015).

Τα Big Data στη Βρετανική Βιβλιοθήκη αποσκοπούν στην ανάπτυξη πλατφόρμας πρόσβασης και ανάκτησης ψηφιακών συλλογών με τη χρήση της υποδομής Cloud της Microsoft Azure. Η δυνατότητα οργάνωσης τεράστιων ποσοτήτων ψηφιοποιημένων δεδομένων με χρήση ευφυών συστημάτων είναι προς όφελος όχι μόνο των ακαδημαϊκών αλλά και εμπλουτίζουν τις γνώσεις της ευρύτερης ακαδημαϊκής κοινότητας (Baker, 2015^a).

Η αξιοποίηση των δυνατοτήτων της τεχνολογίας στο χώρο των βιβλιοθηκών έχει ένθερμα υποστηρικτεί. Η Βρετανική Βιβλιοθήκη (British Library) τα τελευταία χρόνια έχει δημιουργήσει ένα όραμα για μετασχηματισμό της πρόσβασης και της έρευνας σε ψηφιακές συλλογές του. Η προσπάθεια της Βρετανικής Βιβλιοθήκης στα Big Data αποτελεί σημαντικό συμπλήρωμα των εν εξελίξει δραστηριοτήτων των υποδομών της επιτρέποντας την ανάπτυξη των υπηρεσιών που προσφέρει με συμμετοχή και των ψηφιακών συλλογών. Το πείραμα της Βρετανικής Βιβλιοθήκης στα Big Data πέραν της οργάνωσης των δεδομένων περιλαμβάνει μια σειρά από δραστηριότητες με σκοπό και τον έλεγχο των υποδομών και συστημάτων (Baker et al., 2015^b:1-3).

Η Βρετανική Βιβλιοθήκη μέσω των Big Data επιτρέπει την πρόσβαση σε στοιχεία και δεδομένα τα οποία υποστηρίζονται από τους επαγγελματίες πληροφόρησης. Τον Ιούνιο του 2014, η πρώτη πειραματική ομάδα της Βρετανικής Βιβλιοθήκης στα Big Data συγκλήθηκε με σκοπό την ψηφιοποίηση 68.000 δεδομένων σε κατάλληλο σύστημα και τη παροχή τους μέσω κατάλληλου δικτύου.

Για τη λειτουργική οργάνωση των δεδομένων χρησιμοποιήθηκε το Microsoft Azure 5 APIs με τους φοιτητές να συνεργάζονται με ερευνητές ανθρωπιστικών επιστημών οι οποίοι εξέφρασαν την επιθυμία χρήσης των δυνατοτήτων ενός τέτοιου συστήματος. Η δημόσια πρόσβαση μέσω κατάλληλης πύλης (<http://b1publicdomain.azurewebsites.net/>) αντιπροσωπεύει μια προσπάθεια να εντοπίσει τις σύνθετες και πολύπλευρες ανάγκες των ανθρωπιστικών ερευνητών ενώ προσφέρει αντισυμβατικές υπηρεσίες, όπως η μαζική λήψη του κειμένου με βάση τα ερωτήματα, πίνακες συχνότητας λέξεων και OCR προεπισκόπηση κειμένου. Μετά από αυτήν την επιτυχημένη πιλοτική προσπάθεια, το πείραμα της Βρετανικής Βιβλιοθήκη στα Big Data συμπεριέλαβε προσπάθειες μηχανικής μάθησης και κινητών εφαρμογών στην προσπάθεια αντιμετώπισης πρόσθετων μαθησιακών αναγκών των φοιτητών (Baker et al., 2015^b:1-3; Farquhar & Baker, 2014:1).



Εικόνα 4.1 : Αξιοποίηση Big Data μέσω εφαρμογών στη Βρετανική Βιβλιοθήκη (British Library)
Πηγή: Baker, 2015^b

4.2. Η υλοποίηση στην Εθνική Γερμανική Βιβλιοθήκη (German National Library)

Η Εθνική Γερμανική Βιβλιοθήκη (German National Library) αποτελεί ένα σημαντικό παράδειγμα αξιοποίησης των δυνατοτήτων που προσφέρουν τα Big Data. Μέσω των Big Data ο χρήστης είναι σε θέση να εντοπίσει μεμονωμένα δεδομένα ενώ τα ψηφιακά έγγραφα και αρχεία ταξινομούνται αυτόματα αξιοποιώντας σχετικές λέξεις κλειδιά. Η Γερμανική Εθνική Βιβλιοθήκη (German National Library) διαθέτει μια συλλογή αρχειακού υλικού, βιβλιογραφιών, εντύπων, χειρογράφων και ηλεκτρονικών εγγράφων που χρονολογούνται από το 19^ο αιώνα. Συνολικά, το αρχειακό υλικό της Γερμανικής Εθνικής Βιβλιοθήκης αποτελείται από 26 εκατ. έγγραφα. Η αφομοίωση των ηλεκτρονικών εκδόσεων στις συλλογές της βιβλιοθήκης έχει οδηγήσει σε μεγάλη αύξηση τα ηλεκτρονικά έγγραφα τα οποία μεμονωμένα δεν θα μπορούσαν να οργανωθούν ή να προστατευθούν σε επίπεδο πνευματικής ιδιοκτησίας. Μέσω των Big Data επετεύχθη όχι μόνο η οργάνωση και η αυτοματοποίηση των διαδικασιών ευρετηρίασης με τη χρήση λέξεων – κλειδιών αλλά και οι εξής στόχοι (Deutsche National Bibliothek, 2015):

1. Μείωση χρόνου ευρετηρίασης
2. Ταχύτερη ανάκτηση δεδομένων
3. Αντιμετώπιση «κενών» ευρετηρίασης
4. Δημιουργία δεδομένων υψηλής ποιότητας
5. Καλύτερη ανάκτηση και επεξεργασία πληροφοριών
6. Οργάνωση δεδομένων και ευκολότερη ανανέωσή τους

Με τη χρήση των Big Data η Εθνική Γερμανική Βιβλιοθήκη πέτυχε αρτιότερη οργάνωση του αρχειακού υλικού, την ευκολότερη ανάκτηση, τη διασφάλιση της προστασίας των προσωπικών δεδομένων και της πνευματικής ιδιοκτησίας ενώ και η πρόσβαση είναι ευκολότερη για τους χρήστες. Τα ηλεκτρονικά έγγραφα και αρχεία ταξινομούνται αυτόματα ενώ η βιβλιοθήκη μπορεί να υποστηρίξει ψηφιακά δεδομένα με μεγάλη ετερογένεια (γλώσσα, είδος αρχείου, όγκος κ.ά.). Τα Big Data προσέφεραν στην Εθνική Γερμανική Βιβλιοθήκη δυνατότητες όχι μόνο οργάνωσης, διαχείρισης και επεξεργασίας μεγάλου όγκου δεδομένων αλλά και δυνατοτήτων πρόσβασης σε ένα ευρύτερο φάσμα χρηστών σε παγκόσμιο επίπεδο. Ο μακροπρόθεσμος στόχος της

βιβλιοθήκης είναι να προσελκύει συνεχώς νέες ομάδες χρηστών υποστηρίζοντας ταυτόχρονα τις απαιτήσεις τους (Deutsche National Bibliothek, 2015)

Η εφαρμογή των Big Data στην Εθνική Γερμανική Βιβλιοθήκη συνοδεύτηκε από την ενεργό συμμετοχή και υποστήριξη των επαγγελματιών πληροφόρησης, μηχανικών ηλεκτρονικών υπολογιστών, προγραμματιστών με στόχο να δημιουργηθεί μια παγκόσμια υποδομή πληροφοριών θέτοντας τις βάσεις για μια σύγχρονη διαχείριση της πληροφορίας και των δεδομένων. Παρά το γεγονός ότι ακόμη και σήμερα δεν έχουν προσδιοριστεί με ακρίβεια τα οφέλη των Big Data στη βιβλιοθήκη, εντούτοις είναι δεδομένο ότι έχει εισέλθει σε μια νέα εποχή διαχείρισης των δεδομένων (Averbis, 2015).

4.3. Υλοποίηση Big Data στη Βιβλιοθήκη του Κογκρέσου (Library of Congress)

Η Βιβλιοθήκη του Κογκρέσου (Library of Congress) είναι το αρχαιότερο ομοσπονδιακό πολιτιστικό ίδρυμα των Η.Π.Α. και εμμέσως, η εθνική βιβλιοθήκη της χώρας. Είναι επίσης η μεγαλύτερη βιβλιοθήκη στον κόσμο, με περισσότερα από 36 εκατομμύρια βιβλία και έντυπο υλικό, καθώς και περισσότερα από 121 εκατομμύρια χάρτες, χειρόγραφα, φωτογραφίες, ταινίες, ήχου και βίντεο, χαρακτηριστικά και σχέδια, καθώς και άλλες ειδικές συλλογές. Μεταξύ των ποικίλων πρωτοβουλιών της βιβλιοθήκης είναι η δημιουργία υποδομής ψηφιακής πληροφορίας, μια εθνική στρατηγική για τη συλλογή, διατήρηση και διάθεση σημαντικού όγκου ψηφιακού περιεχομένου ιδίως της πληροφορίας που δημιουργείται σε ψηφιακή μορφή όχι μόνο για την τρέχουσα αλλά και τις μελλοντικές γενιές (Weller et al., 2013:57-59).

Από το Μάρτιο του 2014, η Βιβλιοθήκη του Κογκρέσου έχει συλλέξει περίπου 525 terabytes δεδομένων από τον ιστό με αύξηση περίπου 5 terabytes ανά μήνα. Η συγκεκριμένη εξέλιξη οδήγησε στην ανάγκη ψηφιακής διαφύλαξης και αρχειοθέτησης δεδομένων. Ξεχωριστή περίπτωση αποτέλεσε η οργάνωση, αποθήκευση και διαχείριση των δισεκατομμυρίων tweets στα πλαίσια πολιτικών, πολιτιστικών και κοινωνικών γεγονότων (Library of Congress, 2014).

Το 2010 η Βιβλιοθήκη του Κογκρέσου και το Twitter ανακοίνωσαν μια συμφωνία διαχείρισης των δεδομένων (Raymond, 2010). Η συμφωνία προέβλεπε τόσο το είδος όσο και τον τρόπο προστασίας στη διάθεση των δεδομένων με δικαίωμα πρόσβασης

σε εκείνους τους χρήστες όπου η βιβλιοθήκη θα όριζε ως κατάλληλους και αξιόπιστους. Τον Ιανουάριο του 2013, η βιβλιοθήκη ανακοίνωσε ότι έχει οργανώσει και διαχειρίζεται ένα αρχείο περίπου 170 δισεκατομμύρια tweets, συνολικού όγκου 133,2 terabytes για το διάστημα 2006 – 2010. Η βιβλιοθήκη με τη χρήση της τεχνολογίας Big Data κατόρθωσε να επεξεργάζεται και να αποθηκεύει μεγάλο όγκο δεδομένων ενώ συνεχιζόταν η ροή δεδομένων (Allen, 2013).

Για τη Βιβλιοθήκη του Κογκρέσου η διαχείριση του όγκου των δεδομένων είναι μια συνεχής πρόκληση. Αν και διαχειρίζεται μεγάλο όγκο δεδομένων για πολλά χρόνια το αρχείο των tweets δημιούργησε μια μοναδική πρόκληση, αντιπροσωπεύει ένα νέο τύπο συλλογής δεδομένων. Οι προκλήσεις εστιάζουν στην επεξεργασία και οργάνωση του υλικού, την αποθήκευση, την πρόσβαση και την ανάκτηση. Αν και η Βιβλιοθήκη του Κογκρέσου είναι αρκετά έμπειρη σχετικά με τη διατήρηση μεγάλων ποσοτήτων ψηφιακών πληροφοριών και έχει συγκεντρώσει περισσότερα από 525 terabytes δεδομένων, τα tweets έχουν θέσει μοναδικές τεχνικές προκλήσεις λόγω του μεγέθους και της πολυπλοκότητάς τους. Ένας παράγοντας που συμβάλλει στο μεγάλο όγκο είναι το ποσό των μετα-δεδομένων που συνοδεύουν κάθε tweet (Dwoskin, 2014).

Προσθέτοντας στις προκλήσεις την παροχή σταθερής και βιώσιμης ψηφιακής αποθήκευσης για ένα τόσο μεγάλο όγκο, η βιβλιοθήκη εστίασε στην εύρεση κατάλληλων πόρων και ανθρώπινου δυναμικού (Nave, 2012). Επιλέγοντας την κατάλληλη υλικοτεχνική υποδομή, τους επαγγελματίες πληροφόρησης και εξασφαλίζοντας την προστασία των δεδομένων, η Βιβλιοθήκη του Κογκρέσου δημιούργησε μια βάση ψηφιακής αρχειοθέτησης και διαχείρισης μεγάλων βάσεων δεδομένων. Τα Big Data στη Βιβλιοθήκη του Κογκρέσου έδωσαν λύση όχι μόνο στη διαχείριση των δεδομένων αλλά και στην πρόσβαση και επεξεργασία ερωτημάτων των χρηστών (Gaffney & Puschmann, 2013:60-62).

5. Big Data & Επαγγελματίες Πληροφόρησης: Νέοι ρόλοι και ευκαιρίες

5.1. Γνώσεις επαγγελματιών πληροφόρησης και επεξεργασία πληροφοριών

Η ανάπτυξη των Big Data στο χώρο των βιβλιοθηκών είναι δύσκολο να επιτευχθεί χωρίς τον ενεργό ρόλο και την ανάπτυξη των γνώσεων των επαγγελματιών πληροφόρησης γύρω από το συγκεκριμένο ζήτημα. Τα Big Data έχουν δημιουργήσει νέες συνθήκες, ευκαιρίες και προκλήσεις για το περιβάλλον των βιβλιοθηκών και την ακαδημαϊκή κοινότητα. Η αύξηση των διαθέσιμων δεδομένων για τις βιβλιοθήκες έχει καταστήσει αναγκαία την επάρκεια των γνώσεων των επαγγελματιών πληροφόρησης αναφορικά με τους τρόπους που οι συγκεκριμένες πληροφορίες μπορούν να καταστούν αξιοποιήσιμες (Mavrinac, 2010:2-3).

Η εξέλιξη των Big Data και η υιοθέτησή τους στο χώρο της βιβλιοθήκης έχει υπερβεί από τα στενά πλαίσια της αποθήκευσης και οργάνωσης των δεδομένων. Οι επαγγελματίες πληροφόρησης πλέον καλούνται να διαθέτουν επαρκείς γνώσεις αναφορικά και με την επεξεργασία των δεδομένων που διαθέτουν στους χρήστες. Η σωστή ταξινόμηση, η οργάνωση, η διαχείριση του υλικού και κυρίως η κατάλληλη επεξεργασία ουσιαστικά καθιστούν τα Big Data ιδιαίτερα χρήσιμα για τις βιβλιοθήκες. Οι επαγγελματίες πληροφόρησης καλούνται να επεξεργάζονται τα δεδομένα κατά τέτοιο τρόπο ώστε να αποθηκεύονται και να διατηρούνται διαθέσιμα για μακροπρόθεσμη πρόσβαση, χρήση και επαναχρησιμοποίηση προσιτή σε ένα ευρύ φάσμα χρηστών (Lougee et al., 2007:9-11).

Η γνώση των επαγγελματιών πληροφόρησης στην επεξεργασία των πληροφοριών είναι ιδιαίτερα σημαντική. Ο βασικός λόγος είναι η κατάλληλη διαχείριση του συνόλου των δεδομένων των Big Data ώστε να εξασφαλίζεται εύκολη αξιοποίηση, ταχύτητα πληροφόρησης, ικανοποίηση των χρηστών των βιβλιοθηκών ακόμα και περαιτέρω δημιουργία πληροφοριών. Οι επαγγελματίες της πληροφόρησης ακολουθώντας τις εξελίξεις αναφορικά με τα Big Data καλούνται να ανταποκριθούν στις νέες ανάγκες και γνώσεις επεξεργασίας των πληροφοριών. Πλέον, έχει ξεφύγει από τα στενά πλαίσιά τους η οργάνωση και η διαχείριση απαιτώντας και επεξεργασίας τους. Η εκθετική αύξηση των πληροφοριών για τις βιβλιοθήκες τα

επόμενα χρόνια, καθιστά αναγκαία την εκπαίδευση και εκμάθηση των επαγγελματιών πληροφόρησης όχι μόνο στη διαχείριση, οργάνωση και πρόσβαση αλλά και στην επεξεργασία των πληροφοριών. Οι βιβλιοθηκονόμοι μπορούν δυνητικά να διαδραματίσουν σημαντικό ρόλο στην εν λόγω προσπάθεια, φέρνοντας τεχνογνωσία σε πληροφορίες και διαχείριση της γνώσης. Σε μια εποχή που πολλές βιβλιοθήκες αντιμετωπίζουν νέες προκλήσεις στην επεξεργασία της πληροφορίας, οι βιβλιοθηκονόμοι μπορούν να αναδείξουν τον ενεργό τους ρόλο σ' αυτήν την κατεύθυνση ενισχύοντας την υπάρχουσα γνώση τους (Federer, 2016:41-42).

5.2. Ο ρόλος των επαγγελματιών πληροφόρησης στη δημιουργία και παρουσίαση Big Data

Οι σύγχρονες βιβλιοθήκες έχουν συσσωρεύσει ένα τεράστιο όγκο δεδομένων ο οποίος είναι διαθέσιμος τόσο σε φυσική όσο και σε ηλεκτρονική μορφή. Ωστόσο, πολλές φορές αυτά τα δεδομένα παρουσιάζονται κατά τέτοιο τρόπο ώστε είναι κατανοητά μόνο από μέλη της ακαδημαϊκής κοινότητας περιορίζοντας το βαθμό των χρηστών. Το συγκεκριμένο πρόβλημα έχει δημιουργήσει νέες παραμέτρους στο ρόλο των επαγγελματιών πληροφόρησης αναφορικά με τη δημιουργία και την παρουσίαση δεδομένων και ιδιαίτερα εκείνων που οργανώνονται μέσω Big Data. Για αρκετούς αιώνες, οι βιβλιοθήκες ακολουθούσαν συγκεκριμένα πρότυπα αποθήκευσης και οργάνωσης των πληροφοριών καθιστώντας τις προσβάσιμες στους χρήστες. Η τεχνολογική ωστόσο εξέλιξη και η εκθετική αύξηση τόσο των πληροφοριών όσο και των αναγκών των χρηστών, οδήγησε σε αύξηση της ανάγκης δημιουργίας και οργάνωσης τους (Teets & Goldner, 2013:429; Hoy, 2014:7).

Ο επαγγελματίας πληροφόρησης καλείται πλέον να επιτελέσει σημαντικότερο ρόλο στον τρόπο με τον οποίο τα δεδομένα δημιουργούνται και παρουσιάζονται ιδιαίτερα στην περίπτωση των Big Data. Η δημιουργία και παρουσίαση των Big Data από τους επαγγελματίες πληροφόρησης απαιτεί αξιοποίηση όλων των διαθέσιμων δεδομένων τα οποία παραμένουν ανεκμετάλλευτα και ανοργάνωτα. Ο επαγγελματίας πληροφόρησης καλείται στο σύγχρονο περιβάλλον των βιβλιοθηκών να εντοπίζει, να οργανώνει και να επεξεργάζεται τα δεδομένα κατά τέτοιο τρόπο ώστε να είναι προσβάσιμα στο σύνολο των χρηστών (Teets & Goldner, 2013:430; Hoy, 2014:8).

Η οργάνωση των πληροφοριών στις βιβλιοθήκες έχει πραγματοποιηθεί εδώ και αρκετούς αιώνες. Ωστόσο, στην πορεία εξέλιξης της διαχείρισης και οργάνωσης της πληροφορίας σε συνδυασμό με την εμφάνιση των Big Data, προέκυψαν σημαντικά ερωτήματα αναφορικά με το ρόλο που επιτελούν οι επαγγελματίες πληροφόρησης στον τομέα της δημιουργίας και παρουσίασης. Ο μεγάλος όγκος δεδομένων προς επεξεργασία κατέστησε τους επαγγελματίες πληροφόρησης ως βασικό κρίκο στην αλυσίδα αποδοτικής επεξεργασίας των πληροφοριών. Πλέον, η τεχνολογική εξέλιξη και η διάδοση του διαδικτύου είναι τέτοια ώστε ο επαγγελματίας πληροφόρησης τόσο στη δημιουργία όσο και στην παρουσίαση των Big Data να πρέπει να διαθέτει τη σχετική γνώση (Teets & Goldner, 2013:430-431).

Ο χώρος των βιβλιοθηκών και ιδιαίτερα οι επαγγελματίες πληροφόρησης είναι οι καταλληλότεροι για τη δημιουργία και την επεξεργασία των Big Data. Οι βιβλιοθήκες έχουν μακρά παράδοση στο να υιοθετούν νωρίς την τεχνολογία και τα Big Data δεν αποτελούν εξαίρεση (Huwe, 2014:17-18). Υπάρχουν αρκετοί τρόποι που οι επαγγελματίες πληροφόρησης μπορούν να συμμετάσχουν. Ειδικότερα, μέσω της συλλογής, ανάπτυξης και διατήρησης του συνόλου των δεδομένων μπορούν να συνεισφέρουν σημαντικά καθώς όλο και περισσότεροι χρήστες ενδιαφέρονται να αξιοποιήσουν τις δυνατότητες των Big Data. Ο ρόλος των επαγγελματιών πληροφόρησης στη δημιουργία και παρουσίαση Big Data είναι καθοριστικός αν αναλογιστεί κανείς ότι επιτελούν και το ρόλο ενημέρωσης των χρηστών (Hoy, 2014:8).

Ένας άλλος τρόπος όπου οι επαγγελματίες πληροφόρησης και ειδικότερα οι βιβλιοθηκονόμοι μπορούν να ασχοληθούν με τα Big Data είναι βοηθώντας στη διαχείριση και επεξεργασία των ερευνητικών δεδομένων. Οι ερευνητές πολλές φορές χρειάζονται βοήθεια αναφορικά με τη διαχείριση και την επεξεργασία των Big Data έχοντας συγκεκριμένες κατευθυντήριες γραμμές που θα πρέπει να ακολουθηθούν. Οι βιβλιοθηκονόμοι μπορούν να παρέχουν βοήθεια στους ερευνητές, ακόμη και σε φάσεις του σχεδιασμού των έργων τους, αξιολογώντας τις αρχειακές αξίες διατήρησης και κοινής χρήση των δεδομένων (Creamer et al., 2014:255)

5.3. Νέοι ρόλοι και ευκαιρίες επαγγελματιών πληροφόρησης στα Big Data

Οι εξελίξεις στο χώρο των Big Data έχουν δημιουργήσει νέους ρόλους και ευκαιρίες για τα ενεργότερα μέλη στο χώρο των βιβλιοθηκών, τους βιβλιοθηκονόμους. Πλέον, οι βιβλιοθήκες έχουν ξεφύγει από τα στενά πλαίσια παροχής πληροφόρησης ευρετηριακού χαρακτήρα στους χρήστες δημιουργώντας ένα νέο περιβάλλον για τους επαγγελματίες πληροφόρησης. Τα προηγούμενα χρόνια και πριν την εντυπωσιακή ανάπτυξη των Big Data ο ρόλος για τους επαγγελματίες πληροφόρησης ήταν σαφώς πιο περιορισμένος. Οι τεχνολογικές εξελίξεις, η αύξηση των αναγκών των χρηστών, οι βελτιώσεις και η ανάπτυξη στο χώρο των βιβλιοθηκών έχουν δημιουργήσει νέους ρόλους και ευκαιρίες που μπορούν να αξιοποιήσουν οι επαγγελματίες πληροφόρησης αναφορικά με τα Big Data. Νέοι ρόλοι και ευκαιρίες επαγγελματιών πληροφόρησης στα Big Data αποτελούν η ενεργή συμμετοχή τους στο σχεδιασμό του τρόπου λειτουργίας τους, η συλλογή των πληροφοριών, η σωστή συντήρηση των δεδομένων και η σχετική επιμέλεια, η ενημέρωση των χρηστών αναφορικά με τις δυνατότητες των Big Data, η βοήθεια επί ερευνητικών προσπαθειών, η αξιοποίηση των διαθέσιμων πόρων και η απόκτηση επιπλέον εμπειρίας και ενημέρωσης σχετικά με τα Big Data (Federer, 2016:38).

Ο σχεδιασμός των Big Data είναι ένας τομέας όπου αποτελεί πρόκληση και σημαντική ευκαιρία αξιοποίησης των δυνατοτήτων των επαγγελματιών πληροφόρησης. Τα επίπεδα γνώσης των επαγγελματιών πληροφόρησης πολλές φορές κινούνται σε τέτοια επίπεδα ώστε η παροχή της να είναι πολύτιμη για τον κατά το δυνατό καλύτερο αρχικό σχεδιασμό των Big Data. Παρατηρείται όλο και περισσότερο οι βιβλιοθήκες να παρέχουν στήριξη στις προσπάθειες επαγγελματιών πληροφόρησης να ενημερώνουν σχετικά με τον τρόπο οργάνωσης και σχεδιασμού των Big Data που προορίζονται για βιβλιοθήκες. Το συγκεκριμένο σημείο είναι καθοριστικό για την εξέλιξη του ρόλου των επαγγελματιών πληροφόρησης οι οποίοι μπορούν να αποτελέσουν πολύτιμο κρίκο στην αλυσίδα περαιτέρω ανάπτυξης τόσο του ρόλου τους στη διαχείριση και οργάνωση των Big Data όσο και στην επεξεργασία των δεδομένων (Federer, 2016:39).

Τα Big Data έχουν δημιουργήσει ένα νέο περιβάλλον αναφορικά με το διαμοιρασμό των δεδομένων. Οι επαγγελματίες πληροφόρησης καλούνται πλέον να

κατευθύνουν αρτιότερα τους χρήστες των βιβλιοθηκών αλλά και ερευνητές ώστε να επιτύχουν ευκολότερα το στόχο της εύρεσης των κατάλληλων πληροφοριών. Ο τεράστιος όγκος δεδομένων πολλές φορές προκαλεί σύγχυση και παραπληροφόρηση, στοιχεία τα οποία θα πρέπει να αντιμετωπιστούν από τους επαγγελματίες πληροφόρησης. Ιδιαίτερα στην περίπτωση ερευνητών, οι επαγγελματίες πληροφόρησης καλούνται να παρέχουν πλήρη και σαφή ενημέρωση αναφορικά με τα διαθέσιμα δεδομένα ώστε να αποκλειστούν φαινόμενα παρερμηνειών ή παραπληροφόρησης (Pinfield et al., 2014:1-3; Federer, 2016:41).

Η εξέλιξη των Big Data στην περίπτωση των βιβλιοθηκών προσδίδει έναν νέο ρόλο στους επαγγελματίες πληροφόρησης, εκείνον του «πληροφοριοδότη». Ειδικότερα, ο επαγγελματίας πληροφόρησης καλούμενος να οργανώσει, να αναλύσει και να επεξεργαστεί μεγάλο όγκο δεδομένων είναι σε θέση να διαθέτει πολύτιμες γνώσεις και εμπειρίες αναφορικά με ατέλειες ή δυνατά σημεία. Ο σχεδιασμός ή η βελτίωση των Big Data μπορεί να επιτευχθεί με ενεργή συμμετοχή των επαγγελματιών πληροφόρησης αξιοποιώντας την εμπειρία τους από τη χρήση τους, την κατάρτιση, τις γνώσεις και το ταλέντο τους στην επίλυση προβλημάτων. Οι βιβλιοθηκονόμοι πλέον καλούνται να διαθέτουν εμπειρία και άριστη κατάρτιση στην αναζήτηση δεδομένων και σχετική εξοικείωση με τους τύπους των δεδομένων που μπορεί να είναι χρήσιμα. Η επίτευξη των συγκεκριμένων στόχων βασίζεται σε σημαντικό βαθμό στην αναγκαία ενημέρωση, εκπαίδευση και πληροφόρηση των επαγγελματιών (Pinfield et al., 2014:1-3; Federer, 2016:41).

Σημείο πρόκλησης και ευκαιριών για τους επαγγελματίες πληροφόρησης αποτελεί η οπτικοποίηση των δεδομένων στα Big Data. Ο τρόπος με τον οποίο τα δεδομένα οργανώνονται και απεικονίζονται είναι μια νέα πρόκληση για τους επαγγελματίες πληροφόρησης. Η οπτικοποίηση δεδομένων αρχικά αποτελούσε μια παράθεση βασικών πληροφοριών ενώ πλέον έχει μετατραπεί σε σημείο αναφοράς για τη συνολική λειτουργία των Big Data. Οι δυνατότητες της τεχνολογίας (3D αναζητήσεις, εκτυπώσεις και παρουσίασης των δεδομένων) έχουν διαφοροποιήσει σημαντικά την οπτικοποίηση των δεδομένων. Πλέον, ο επαγγελματίας πληροφόρησης καλείται να καταρτιστεί στη χρήση σύγχρονων τεχνολογιών ώστε η πληροφόρηση μέσω των Big Data να μεγιστοποιείται σε επίπεδο απόδοσης (Federer, 2016:41).

Ο επαγγελματίας της πληροφόρησης στις τελευταίες εξελίξεις στο χώρο της πληροφορίας δεν είναι απλά διαχειριστής ή διανεμητής της διαθέσιμης γνώσης. Ο στενός ρόλος των επαγγελματιών πληροφόρησης έχει μεταβληθεί ξεφεύγοντας από

τα στενά όρια εκείνου του προσώπου που αναλαμβάνει την αναζήτησης πληροφοριών. Πρόκληση για τον επαγγελματία πληροφόρησης αποτελεί η παροχή σύγχρονων υπηρεσιών στους χρήστες, οργάνωση των πηγών πληροφοριών κατά τρόπο χρήσιμο, διευκόλυνση στην αξιοποίηση των δυνατοτήτων των Big Data. Οι επαγγελματίες πληροφόρησης αναλαμβάνουν πλέον έναν εκπαιδευτικό ρόλο σε ένα χώρο όπως τα Big Data όπου η αναζήτηση, η ανάκτηση και η ερμηνεία των πληροφοριών πολλές φορές μπορεί να αποδειχθεί ιδιαίτερα δύσκολη (Erasmus, 2001:17).

Οι επαγγελματίες πληροφόρησης καλούνται πλέον να προσφέρουν πολύτιμες και καθοριστικές υπηρεσίες προστιθέμενης αξίας στις πληροφορίες των Big Data μέσω της κατάλληλης οργάνωσης, αποθήκευσης, διαχείρισης και επεξεργασίας. Ειδικότερα, στην εποχή της Πληροφορίας και της Γνώσης, οι ειδικοί στη διαχείριση πληροφοριών είναι σε θέση να προσφέρουν υπηρεσίες προστιθέμενης αξίας μέσω του εντοπισμού, επεξεργασίας και αναμόρφωσης των πληροφοριών με τρόπο που να ανταποκρίνεται στις εξειδικευμένες και εξατομικευμένες απαιτήσεις των ενδιαφερομένων (Chernatory & Harris, 2000:42).

Ο επαγγελματίας πληροφόρησης πλέον αποτελεί σημαντικό κομμάτι της λειτουργίας των Big data στις βιβλιοθήκες προσφέροντας κατά περίπτωση ακόμα και ανταγωνιστικά πλεονεκτήματα. Η εμπλοκή τους στη διαχείριση των πληροφοριών είναι ζωτικής σημασίας για τη συνολικότερη εξέλιξη των Big Data. Ειδικότερα, εξαιτίας των γνώσεων που αποκομίζουν από τα Big Data αποτελούν πλέον πολύτιμο κεφάλαιο για οποιαδήποτε βιβλιοθήκη, το οποίο αν αξιοποιηθεί κατάλληλα μπορεί να λειτουργήσει καταλυτικά σε επίπεδο βελτιώσεων ή αναδιαρθρώσεων των Big Data. Η ευκαιρία οργάνωσης των πληροφοριών στη βάση των αναγκών των χρηστών μπορεί να αξιοποιηθεί μέσω των επαγγελματιών πληροφόρησης. Αποτελούν την ιδανικότερη περίπτωση εμπλεκομένων με τα Big Data για τη δημιουργία κατάλληλων συστημάτων και περιβάλλοντος πληροφόρησης (Cline, 2000:25). Οι νέοι ρόλοι και ευκαιρίες των επαγγελματιών πληροφόρησης στα Big Data έχουν διαμορφώσει ταυτόχρονα ένα νέο πρότυπο αναγκών εκπαίδευσης και κατάρτισής τους. Οι εξελίξεις σε επίπεδο τεχνολογίας και πληροφορικής απαιτούν συνεχή εκπαίδευση, ενημέρωση και κατάρτιση σε ζητήματα που σχετίζονται με την οργάνωση, την αποθήκευση, τη διαχείριση και την επεξεργασία των Big Data αποτελώντας πλέον πρόκληση για τους επαγγελματίες πληροφόρησης (Hallam & Partridge, 2005:6-8).

6. Προκλήσεις στα Big Data

6.1. Οι προκλήσεις για τα Big Data

Η ραγδαία αύξηση των δεδομένων στην εποχή των Big Data επιφέρει τεράστιες προκλήσεις για την απόκτηση δεδομένων, την αποθήκευση, τη διαχείριση και την ανάλυση. Είναι προφανές πλέον ότι οι παραδοσιακές σχεσιακές βάσεις δεδομένων δεν μπορούν να χειριστούν τον τεράστιο όγκο και την ετερογένεια των Big Data. Η προσπάθεια διαχείρισης και οργάνωσης των Big Data πλέον βασίζεται σε τεχνολογίες cloud, σε βάσεις δεδομένων NoSQL και σε εφαρμογές οι οποίες είναι σε θέση να υποστηρίξουν τον όγκο των δεδομένων. Πλέον, τα Big Data βρίσκονται μπροστά σε μια σειρά προκλήσεων. Οι βασικές προκλήσεις εστιάζουν στα ακόλουθα σημεία:

1. **Κατάλληλη αναπαράσταση δεδομένων.** Πολλά σύνολα δεδομένων έχουν ορισμένα επίπεδα ετερογένειας ως προς το είδος, τη δομή, τη σημασιολογία, την οργάνωση, τη διακριτικότητα και την προσβασιμότητα. Η αναπαράσταση δεδομένων έχει ως στόχο να καταστήσει τα δεδομένα πιο ουσιαστικά σε ανάλυση και ερμηνεία από το χρήστη. Η αποτελεσματική αναπαράσταση των δεδομένων είναι πρόκληση ώστε να αντικατοπτρίζεται σαφώς η δομή των δεδομένων, ο τύπος τους και η σαφής έννοιά τους (Labrinidis & Jagadish, 2012:2032-2033; Chaudhuri et al., 2011:89-92; Agrawal, 2012:8).
2. **Μείωση όγκου αποθήκευσης και συμπίεση δεδομένων.** Γενικά, υπάρχει υψηλό επίπεδο πλεονασμού σε σύνολα δεδομένων. Η μείωση του όγκου αποθήκευσης και η συμπίεση δεδομένων μπορεί να είναι αποτελεσματικές στη μείωση του έμμεσου κόστους ενώ τυχόν πλεονάζοντα δεδομένα μπορούν να αποφεύγονται.
3. **Διαχείριση κύκλου ζωής των δεδομένων.** Σημαντική πρόκληση των Big Data είναι η διαχείριση του κύκλου ζωής των δεδομένων καθώς είναι πρακτικά αδύνατο τα εκάστοτε συστήματα αποθήκευσης να διαχειρίζονται πλήρως τα διαθέσιμα δεδομένα (Labrinidis & Jagadish, 2012:2032-2033; Chaudhuri et al., 2011:89-92; Agrawal, 2012:8). Ως εκ τούτου, είναι σημαντικό να δημιουργηθούν προϋποθέσεις αξιολόγησης της σπουδαιότητας των εκάστοτε δεδομένων για να αποφασίζεται ποια δεδομένα θα αποθηκεύονται και ποια πρέπει να απορρίπτονται.

4. **Εμπιστευτικότητα δεδομένων.** Το σύνολο των δεδομένων θα πρέπει πλέον να εξασφαλίζουν υψηλό επίπεδο ασφάλειας, τήρησης απορρήτου και διασφάλισης των προσωπικών δεδομένων (Labrinidis & Jagadish, 2012:2032-2033; Chaudhuri et al., 2011:89-92; Agrawal, 2012:8).
5. **Διαχείριση ενέργειας.** Η αύξηση του όγκου των δεδομένων και των απαιτήσεων σε επεξεργασία, αποθήκευση και μετάδοση των Big Data απαιτούν αναπόφευκτα υψηλότερα επίπεδα ενέργειας. Πλέον, απαιτείται η αξιοποίηση των κατάλληλότερων συστημάτων ώστε αφενός να διασφαλίζεται η άντληση της πληροφορίας και αφετέρου να μειώνονται κατά το δυνατόν τα επίπεδα ενέργειας.
6. **Επεκτασιμότητα και συνεργασία δεδομένων.** Τα Big Data καλούνται μελλοντικά να είναι σε θέση να επεξεργάζονται όλο και πιο πολύπλοκα σύνολα δεδομένων με ταυτόχρονο συνδυασμό των περισσότερων δυνατών πηγών (Labrinidis & Jagadish, 2012:2032-2033; Chaudhuri et al., 2011:89-92; Agrawal, 2012:8).

6.2. Η ποιότητα των Big Data ως σημασιολογική πρόκληση

Ένα βασικό σημείο για τα Big Data είναι σε ποιό βαθμό η ποιότητά τους μπορεί να αποκτήσει σημασιολογική πρόκληση. Οι επαγγελματίες της γνώσης καλούνται να δώσουν απάντηση στο κατά πόσο τα Big data μπορούν να αξιοποιηθούν σημασιολογικά. Σε κάθε περίπτωση, η απάντηση μόνο καταφατική μπορεί να είναι καθώς τα παραγόμενα δεδομένα μπορούν μέσω της κατάλληλης σύνδεσης να προσφέρουν χρήσιμη πληροφόρηση (Cai & Zhu, 2015:2).

Η ποιότητα των Big Data μπορεί να είναι υψηλού επιπέδου εφόσον έχει προηγηθεί κατάλληλη επεξεργασία ή αντιπροσωπεύουν σαφείς πληροφορίες. Στην πραγματικότητα, είναι αρκετά δύσκολο να επιτευχθεί άριστη ποιότητα για τα δεδομένα ενώ αποκτούν σημαντική σημασία στη περίπτωση που κινούνται σε υψηλά επίπεδα ποιότητας. Για τη επίτευξη ποιότητας στα Big Data, θα πρέπει να επιτευχθούν μια σειρά προκλήσεων οι οποίες εστιάζουν στα εξής (Bizer et al., 2012:57):

1. **Ενοποίηση δεδομένων με κατάλληλη επεξεργασία ως προς τον τύπο, τη δομή, το περιεχόμενο**

2. Πλήρης περιγραφή και οργάνωση δεδομένων μεγάλης κλίμακας

3. Αποτελεσματική διασύνδεση δεδομένων

Οι παραπάνω προκλήσεις για την επίτευξη ποιότητας (ενσωμάτωση δεδομένων, επεκτάσιμη συλλογιστική, κτλ.), θα μπορούσαν να θεωρηθούν η βάση αυξημένης σημασιολογίας των δεδομένων. Επιπλέον, η σημασιολογία των δεδομένων μπορεί να να χρησιμοποιηθεί σε ένα σύστημα το οποίο καθιστά δυνατή την ανίχνευση ασυνεπών δεδομένων αλλά και στη δημιουργία νέων γνώσεων. Η σχέση μεταξύ της ποιότητας των Big Data και της σημασιολογίας είναι αμφίδρομη. Τα Big Data δεν αποκτούν κάποιο νόημα εάν δεν συνοδεύονται αντίστοιχα από ακρίβεια πληροφοριών. Η ποιότητα των Big Data αποκτά σημασιολογική πρόκληση καθώς τυχόν μη ποιοτικά δεδομένα είναι βέβαιο ότι θα οδηγήσουν σε ανακριβή ανάλυση ή λήψη μιας απόφασης.

6.3. Προκλήσεις βιβλιοθηκών και επαγγελματιών πληροφόρησης στα Big Data

6.3.1. Προκλήσεις οργάνωσης ετερογένειας και ατελειών πληροφοριών

Σε καθημερινό επίπεδο η διαχείριση και οργάνωση ετερογενών πληροφοριών αποτελεί μια απλή διαδικασία. Στην περίπτωση ωστόσο διαχείρισης και οργάνωσης μεγάλου όγκου δεδομένων όπως τα Big Data, η ετερογένεια ή οι ατέλειες επί των πληροφοριών δημιουργεί μια πρόκληση τόσο για τις βιβλιοθήκες όσο και για τους επαγγελματίες πληροφόρησης. Κατά συνέπεια, τα δεδομένα πρέπει να είναι προσεκτικά δομημένα εξ αρχής για την ανάλυση των δεδομένων (Agrawal et al, 2011:12).

Χαρακτηριστικό παράδειγμα πρόκλησης της οργάνωσης της ετερογένειας των πληροφοριών θα μπορούσαν να αποτελούν οι πληροφορίες για έναν ασθενή που εισέρχεται σε μια νοσοκομειακή μονάδα. Η δημιουργία πληροφοριών ανά ιατρική διαδικασία ή εργαστηριακή εξέταση, πληροφοριακά στοιχεία, πορεία νόσου κλπ. ως προς τη συνολική εικόνα του ασθενούς θα μπορούσαν να δημιουργήσουν μια ετερογένεια. Για την οποιαδήποτε ιατρική απόφαση ο συνδυασμός των πληροφοριών μπορεί να οδηγήσει στη λήψη σωστών αποφάσεων ωστόσο η επεξεργασία σε μια και μόνο βάση δεδομένων δεν είναι βέβαιο ότι θα οδηγούσε προς αυτήν την κατεύθυνση.

Εναλλακτική επιλογή θα μπορούσε να αποτελεί η δημιουργία μικρότερων βάσεων ανάλυσης δεδομένων ώστε ανά περίπτωση να προσφέρεται και η καταλληλότερη πληροφορία. Ωστόσο, τα συστήματα ηλεκτρονικών υπολογιστών λειτουργούν πιο αποτελεσματικά εάν αποθηκεύουν πολλά στοιχεία που είναι όλα πανομοιότυπα σε μέγεθος και δομή. Αποτελεσματική πρόσβαση και ανάλυση πληροφοριών μπορεί να προκύψει μόνο μέσω περαιτέρω ενεργειών από την πλευρά των επαγγελματιών πληροφόρηση (Agrawal et al., 2011:12).

Η συγκέντρωση σε ένα ηλεκτρονικό αρχείο σε κατάλληλη βάση δεδομένων στοιχείων όλων των αναγκαίων δεδομένων όπως η ημερομηνία γέννησης, το επάγγελμα, η ομάδα αίματος κλπ. του παραπάνω παραδείγματος, αποτελεί κρίσιμο σημείο ώστε να εξασφαλιστεί η ομοιογένεια και η αρτιότητα των πληροφοριών. Η καταγραφή με ελλιπή στοιχεία προφανώς και δημιουργεί ετερογένεια στα δεδομένα με συνακόλουθα ελλειμματική ανάλυση (Agrawal et al., 2011:12).

Οι επαγγελματίες πληροφόρησης και οι βιβλιοθήκες με τις εξελίξεις που σημειώνονται σε τεχνολογικό επίπεδο και η ταυτόχρονη αύξηση του όγκου των δεδομένων προς διαχείριση αναμφίβολα συνοδεύονται από ετερογένεια και ατέλειες στις πληροφορίες. Η πρόκληση εστιάζει στον τρόπο με τον οποίο οι επαγγελματίες πληροφόρησης θα αξιοποιήσουν τα δεδομένα, θα τα επεξεργαστούν και θα τα διαμορφώσουν σε μια βάση όπου οι πληροφορίες θα χαρακτηρίζονται από ομοιογένεια. Οι όσο το δυνατόν αρτιότερες πληροφορίες είναι βέβαιο ότι θα συμβάλλουν όχι μόνο στην καλύτερη αξιοποίηση του υλικού αλλά και στη διευκόλυνση των χρηστών, στην επαγγελματική ανάπτυξη των επαγγελματιών και συνολικά στην εξέλιξη των Big Data (Agrawal et al., 2011:12).

6.3.2. Επικαιροποίηση πληροφοριών και ταχύτητα άντλησης

Μια σημαντική πρόκληση που καλούνται να αντιμετωπίσουν οι βιβλιοθήκες και οι επαγγελματίες πληροφόρησης στα Big Data είναι η επικαιροποίηση των πληροφοριών και η ταχύτητα άντλησης τους. Είναι παραδεκτό ότι όσο μεγαλύτερο είναι το σύνολο δεδομένων που πρέπει να υποβληθούν σε επεξεργασία, τόσο περισσότερος χρόνος απαιτείται και για την ανάλυσή τους. Ο σχεδιασμός ενός συστήματος Big Data που ασχολείται αποτελεσματικά με το μέγεθος και την επικαιροποίηση των πληροφοριών είναι πιθανό να οδηγήσει σε ένα σύστημα που μπορεί να επεξεργάζεται ένα μεγάλο όγκο δεδομένων πιο γρήγορα. Ωστόσο, δεν είναι

μόνο η ταχύτητα άντλησης που προβληματίζει τους επαγγελματίες πληροφόρησης και τις βιβλιοθήκες στα πλαίσια των Big Data αλλά κυρίως η επικαιροποίηση των πληροφοριών. Σε πολλές περιπτώσεις παρουσιάζονται αυξημένες απαιτήσεις ανάλυσης των δεδομένων στο οποίο πολλές φορές επικεντρώνονται και οι περισσότερες προσπάθειες, ωστόσο ίσως η συνεχής επικαιροποίηση αποτελεί τη σημαντικότερη πρόκληση (Agrawal et al., 2011:12)

Οι επαγγελματίες πληροφόρησης και οι βιβλιοθήκες καλούνται να στρέψουν το βλέμμα τους στη διαμόρφωση εκείνων των προϋποθέσεων ώστε τα δεδομένα συνεχώς να ανανεώνονται. Ο έλεγχος της καταλληλότητας, της σχετικότητας και της αρτιότητας των πληροφοριών θα ήταν πρακτικά αδύνατο να επιτευχθεί σε απόλυτο βαθμό. Ωστόσο, υπάρχουν δυνατότητες που αν αξιοποιηθούν μπορούν να συνεισφέρουν σημαντικά στην επικαιροποίηση των πληροφοριών και την ταχύτητα άντλησης. Οι τεχνολογίες των Big Data επιτρέπουν στους επαγγελματίες πληροφόρησης να δημιουργήσουν κατάλληλα ευρετήρια, βάσεις δεδομένων, επαρκείς χώρους αποθήκευσης και να εντοπίσουν ευκολότερα όλες εκείνες τις πληροφορίες που χαρακτηρίζονται ως ανεπαρκείς ή ακατάλληλες. Η επικαιροποίηση των πληροφοριών είναι μια από τις σημαντικότερες προκλήσεις που δεν αφορούν αποκλειστικά μόνο τους επαγγελματίες πληροφόρησης αλλά και τις ίδιες τις βιβλιοθήκες. Από την πλευρά των βιβλιοθηκών θα πρέπει συνεχώς να αναβαθμίζονται οι τεχνολογίες που υποστηρίζουν τα Big Data, να παρέχεται κατάλληλη εκπαίδευση και ενημέρωση στους επαγγελματίες πληροφόρησης και γενικά να κινούνται με γνώμονα την αναβάθμιση των Big Data. Οι επαγγελματίες πληροφόρησης από την πλευρά τους καλούνται να συμμετέχουν ενεργά σε όλες τις προσπάθειες επικαιροποίησης των δεδομένων τόσο μέσω των γνώσεών τους όσο και σε πρακτικό επίπεδο (Agrawal et al., 2011:12).

6.3.3. Διασφάλιση ιδιωτικότητας των δεδομένων

Η ιδιωτικότητα των δεδομένων είναι μια άλλη τεράστια πρόκληση για τους επαγγελματίες πληροφόρησης και τις βιβλιοθήκες που αυξάνει στο πλαίσιο των Big Data. Για τα ηλεκτρονικά δεδομένα έχει διαμορφωθεί ένα ιδιαίτερα αυστηρό νομικό πλαίσιο προστασίας αναφορικά με την ιδιωτικότητα των δεδομένων. Ωστόσο, η εξέλιξη των Big Data έχει δημιουργήσει ανησυχίες αναφορικά με την ιδιωτικότητα και την ακατάλληλη χρήση των προσωπικών δεδομένων, ιδίως μέσω της σύνδεσης

των δεδομένων από πολλαπλές πηγές. Η διαχείριση των δεδομένων της ιδιωτικής ζωής είναι τόσο τεχνικό όσο και κοινωνιολογικό πρόβλημα, το οποίο πρέπει να αντιμετωπιστεί από τους επαγγελματίες πληροφόρησης στη βάση των προοπτικών που δημιουργούν τα Big Data (Agrawal et al., 2011:14).

Η πρόσβαση στα Big Data δημιουργεί για το χρήστη ηλεκτρονικό αποτύπωμα από το οποίο μπορούν να αντληθούν στοιχεία αναφορικά με την τοποθεσία πρόσβασης, στοιχεία ταυτότητας, καταναλωτικές προτιμήσεις κλπ. Ο επαγγελματίας πληροφόρησης καλείται να κατευθύνει και να συμβουλεύει τους χρήστες στην αξιοποίηση των πληροφοριών των Big Data με τέτοιο τρόπο ώστε να διασφαλιστεί η ιδιωτικότητα. Το υπάρχον πρότυπο της προστασίας της ιδιωτικής ζωής αν και κινείται σε πολύ θετική κατεύθυνση εντούτοις οι συνεχείς μεταβολές στα δεδομένα θέτουν στον επαγγελματία πληροφόρησης συνεχώς νέες προκλήσεις. Ακόμη, είναι πολύ σημαντικό ο επαγγελματίας πληροφόρησης να επανεξετάζει το βαθμό ασφάλειας σε περιπτώσεις κοινής χρήσης δεδομένων Big Data καθώς πολλές online αναζητήσεις απαιτούν το διαμοιρασμό προσωπικών πληροφοριών. Σε κάθε περίπτωση, η διασφάλιση της ιδιωτικότητας των δεδομένων είναι μια συνεχή πρόκληση για τους επαγγελματίες πληροφόρησης στα πλαίσια λειτουργίας και ανάπτυξης των Big Data (Agrawal et al., 2011:14).

6.3.4. Συνεργασία επαγγελματιών πληροφόρησης

Παρά τις τεράστιες προόδους στην υπολογιστική ανάλυση των Big Data, βασικό σημείο αναφοράς και πρόκλησης αποτελεί η επίτευξη συνεργασιών μεταξύ των επαγγελματιών πληροφόρησης. Η παρουσία επαρκούς τεχνολογίας, η κατάλληλη οργάνωση των δεδομένων και η προστασία της ιδιωτικότητας αναμφίβολα συμπεριλαμβάνονται σε ένα πλέγμα προκλήσεων που καλούνται να διαχειριστούν οι επαγγελματίες πληροφόρησης. Υπάρχει ωστόσο ένα σημείο όπου το σύνολο των προσπαθειών μπορεί να πέσουν στον κενό: η κατάλληλη συνεργασία τόσο ανάμεσα στους επαγγελματίες πληροφόρησης όσο και στη σχέση τους με τις βιβλιοθήκες. Στο σημερινό πολύπλοκο κόσμο των δεδομένων, η συνεργασία των επαγγελματιών πληροφόρησης είναι όχι μόνο σημαντική αλλά πολλές φορές και η βάση επιτυχίας στη λειτουργία και την αξιοποίηση των Big Data στο χώρο των βιβλιοθηκών. Ένα ολοκληρωμένο σύστημα Big Data θα πρέπει να υποστηρίζεται από ένα οργανωμένο πλαίσιο καταρτισμένων και συνεργαζόμενων επαγγελματιών πληροφόρησης. Είναι

αποδεδειγμένο ότι η ομαλή λειτουργία και συνεισφορά των Big Data επιτυγχάνεται μέσω των συνεργασιών των επαγγελματιών πληροφόρησης για τον εντοπισμό αδυναμιών των δεδομένων, κενών ασφάλειας, τεχνολογικών αναγκών, αιτημάτων χρηστών (Agrawal et al., 2011:14).

6.4 Η εφαρμογή των Big Data στις ελληνικές βιβλιοθήκες

Η χρήση των Big Data στις ακαδημαϊκές βιβλιοθήκες αρχικά περιελάμβανε ερωτήματα που σχετίζονταν με τον τρόπο και το χρόνο εφαρμογής τους αν και πλέον τα συγκεκριμένα ερωτήματα έχουν σαφώς απαντηθεί. Τα Big Data έχουν εισέλθει δυναμικά στο χώρο των βιβλιοθηκών δημιουργώντας πλέον νέες απαιτήσεις εκπαίδευσης και κατάρτισης των εμπλεκομένων και ιδιαίτερα των επαγγελματιών πληροφόρησης (Cox & Jantti, 2012).

Οι ελληνικές ακαδημαϊκές βιβλιοθήκες, συντασσόμενες στις παγκόσμιες εξελίξεις στο χώρο των Big Data δεν θα μπορούσαν να παραμένουν αμέτοχες. Σε όλο και μεγαλύτερο φάσμα βιβλιοθηκών στην Ελλάδα καταβάλλονται προσπάθειες καθολικής υιοθέτησης των Big Data ενώ και η εκπαίδευση και κατάρτιση των επαγγελματιών πληροφόρησης στρέφεται σε ένα νέο πρότυπο. Οι ανάγκες των χρηστών ελληνικών και όχι μόνο βιβλιοθηκών δικαιολογούν σε μεγάλο βαθμό τη στροφή που παρατηρείται στο συγκεκριμένο χώρο (Gerolimos, 2011:109-111).

Ωστόσο, η εφαρμογή των Big Data στις ελληνικές βιβλιοθήκες πολλές φορές προσκρούει σε ηθικά ζητήματα και αξίες που έχουν σημασία για τις ακαδημαϊκές βιβλιοθήκες, τη διατήρηση της συμμόρφωσης στις νομικές απαιτήσεις βάσει των οποίων λειτουργούν και σε ζητήματα επαρκούς χρηματοδότησης, επιλογής κατάλληλων υπολογιστικών και λειτουργικών εφαρμογών. Η συγκεκριμένη προσέγγιση προσδιορίζει και εξετάζει μερικά από τα μεγαλύτερα προβλήματα αναφορικά με τη χρήση των Big Data που αφορούν τους ακαδημαϊκούς βιβλιοθηκονόμους. Οι επαγγελματίες της πληροφόρησης καλούνται να διαχειριστούν δεδομένα και να εφαρμόσουν πρακτικές προστασίας της ιδιωτικής ζωής και της ασφάλειας των χρηστών συναντώντας ωστόσο πρακτικές δυσκολίες (έλλειψη κατάρτισης, τεχνολογικών υποδομών κ.ά.) (Gerolimos, 2011:111-112).

Η τεχνολογία των Big Data τυγχάνει όλο και μεγαλύτερης αναγνώρισης από τις ελληνικές ακαδημαϊκές βιβλιοθήκες οι οποίες έχουν στραφεί πλέον σε μια κατεύθυνση που καλύπτει μεγαλύτερες δυνατότητες από απλή διασύνδεση των δεδομένων. Πλέον, τα Big Data έχουν κριθεί ως ιδιαίτερα επωφελή και για τις ελληνικές ακαδημαϊκές βιβλιοθήκες. Τα οφέλη αφορούν όχι μόνο την ίδια τη λειτουργία των βιβλιοθηκών αλλά και σε μεμονωμένα άτομα όπως οι βιβλιοθηκονόμοι, οι επαγγελματίες πληροφόρησης, οι ερευνητές και οι φοιτητές (Σπανός, 2015:17-19; Baker et al., 2011).

Συγκεκριμένα, η αποτύπωση των δεδομένων των βιβλιοθηκών υπό μορφή Big Data καθιστά ευκολότερη την εύρεση των σχετικών τεκμηρίων και πληροφοριών, την ανάκτηση ή ακόμα και την επεξεργασία τους. Επίσης, η ύπαρξη των Big Data στις ελληνικές βιβλιοθήκες επιτρέπει τη σύνδεση τεκμηρίων, ηλεκτρονικών αρχείων κλπ. Παράλληλα, προσφέρεται στις βιβλιοθήκες η δυνατότητα να βελτιώσουν σημαντικά την ποιότητα των δεδομένων τους. Το σύνολο των δεδομένων συνδυάζεται μεταξύ τους με τέτοιο τρόπο ώστε οι χρήστες, οι επαγγελματίες πληροφόρησης αλλά και η ίδια η λειτουργία της βιβλιοθήκης να χαρακτηρίζεται από αυξημένη ποιότητα και ταχύτητα. Με την εφαρμογή των Big Data στις ελληνικές βιβλιοθήκες παρέχεται η ευκαιρία αξιοποίησης και μετασχηματισμού των πληροφοριών με έναν τρόπο που διαφοροποιείται του παραδοσιακού, της αυτόνομης περιγραφικής παρουσίασης των πληροφοριών. Επίσης, τα Big Data δημιουργούν νέες προοπτικές στα σύνολα δεδομένων που μέχρι παλαιότερα ήταν διαθέσιμα μόνο μέσω των ολοκληρωμένων βιβλιοθηκονομικών συστημάτων στα οποία ήταν αποθηκευμένα. Τα Big Data όπως διαμορφώνονται και εφαρμόζονται σταδιακά από τις ελληνικές βιβλιοθήκες δημιουργούν νέες συνθήκες στη διαχείριση των δεδομένων. Στοιχεία όπως ο δανεισμός, ο εντοπισμός του κατάλληλου τεκμηρίου, η διαθεσιμότητα είναι όχι μόνο ευκολότερα αλλά και αμεσότερα ενώ μειώνεται και η εξάρτηση των βιβλιοθηκών από εξειδικευμένα λογισμικά (Σπανός, 2015:17-19).

Οι ελληνικές ακαδημαϊκές βιβλιοθήκες αποτελούν σημαντικά κέντρα συλλογής, επεξεργασίας και διάχυσης των πληροφοριών με βασικό σκοπό την υποστήριξη του εκπαιδευτικού, επιστημονικού και ερευνητικού έργου των πανεπιστημίων. Η συνεχής ανάπτυξη της τεχνολογίας και ιδιαίτερα των Big Data παρέχει τη δυνατότητα βελτίωσης των υπηρεσιών της βιβλιοθήκης, με κύριο στόχο την ικανοποίηση των πληροφοριακών αναγκών των χρηστών. Στις ελληνικές ψηφιακές βιβλιοθήκες αξιοποιώντας τις δυνατότητες των Big Data περιλαμβάνονται βάσεις δεδομένων όπου

οι χρήστες μπορούν να χρησιμοποιούν για εκπαιδευτικούς και ερευνητικούς λόγους. Ο Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών (Heal Link) στον οποίο εντάσσονται οι περισσότερες ακαδημαϊκές βιβλιοθήκες, παρέχει πρόσβαση σε πλήρες κείμενο ηλεκτρονικών πηγών περιοδικών και βιβλίων και σε βιβλιογραφικές βάσεις δεδομένων διαθέτοντας σχετικές πύλες (ΖΕΦΥΡΟΣ, ARTEMIS) (ΥΠ.Π.Ε.Θ., 2016).

7. Case Study Big Data

7.1. Εφαρμογές Big Data

Τα τελευταία χρόνια, οι εφαρμογές Big Data έχουν υιοθετηθεί από μια σειρά κλάδων και επιχειρήσεων παρέχοντας λύση σε σειρά ζητημάτων όπως η διαχείριση του χώρου, η λήψη αποφάσεων. Ήδη από τη δεκαετία του '90 αρκετές επιχειρηματικές εφαρμογές βασίζονταν στις μηχανές αναζήτησης του διαδικτύου ενώ μαζική ανάκτηση και διαχείριση δεδομένων εμφανίστηκε στις αρχές του 21^{ου} αιώνα. Μεταξύ των σημαντικότερων εφαρμογών των Big Data παρουσιάζονται στις ενότητες που θα ακολουθήσουν καταδεικνύοντας τον καταλυτικό ρόλο που μπορούν να επιτελέσουν στην τεχνολογική και όχι μόνο εξέλιξη. Ειδικότερα:

1. **Ιατρικές εφαρμογές.** Τα Big Data αξιοποιούνται στη διαχείριση μεγάλου όγκου ιατρικών δεδομένων των ασθενών, προγραμμάτων αποθεραπείας, χορήγησης φαρμάκων κλπ. Η επιλογή της διάγνωσης και της περίθαλψης βασίζεται στη δυνατότητα των Big Data να συνδυάζουν μεγάλο όγκο δεδομένων, δομημένων και αδόμητων (Huser & Cimino, 2015:891-893).
2. **Υπηρεσία Google Flu Trends.** Η συγκεκριμένη υπηρεσία της Google επιτρέπει την αξιολόγηση μεγάλου όγκου δεδομένων που σχετίζονται με την εκδήλωση νόσων για την έγκαιρη ανίχνευση της εξέλιξής τους. Απώτερος στόχος είναι η πρόληψη ή ο περιορισμός των κρουσμάτων (Ginsberg et al., 2009:102-1014).
3. **Λιανικές πωλήσεις.** Τα Big Data δίνουν στις επιχειρήσεις τη δυνατότητα να αξιολογούν τις καταναλωτικές συνήθειες και να προσαρμόζονται σ' αυτές. Ειδικότερα, μέσω της ανάλυσης των Big Data μπορούν να διαμορφώσουν και ένα πλαίσιο τάσεων ως προς συγκεκριμένα προϊόντα και υπηρεσίες (Buyya et al., 2013:78).
4. **Τραπεζικές εφαρμογές.** Η καταγραφή στοιχείων, η συναλλακτική συμπεριφορά, η προστασία της ιδιωτικότητας και η εξέλιξη σε επίπεδο παροχής χρηματοοικονομικών υπηρεσιών μέσω των Big Data έχουν κινήσει το ενδιαφέρον των τραπεζικών ιδρυμάτων. Η συλλογή, αποθήκευση, οργάνωση και επεξεργασία δεδομένων αποσκοπεί στο σαφή προσδιορισμό της οικονομικής συμπεριφοράς των πελατών (Dzik et al., 2015).

5. **Εκπαίδευση.** Η καταγραφή στοιχείων που αφορούν τους μαθητές (επιδόσεις, βαθμολογίες, συμπεριφορές, μάθηση κλπ) μπορούν να βοηθήσουν σημαντικά στη βελτίωση του προγράμματος μαθημάτων, στην οργάνωση του σχολείου, στη βελτίωση των εκπαιδευτικών κλπ. (Supovitz & Weathers, 2004:34)

7.1.1. Η εφαρμογή των Big Data στις επιχειρήσεις

Τα Big Data τυγχάνουν όλο και περισσότερο της εφαρμογής στις επιχειρήσεις. Μέχρι και πριν από μερικά χρόνια, η βασική επιλογή των επιχειρήσεων στη διαχείριση και επεξεργασία των πληροφοριών ήταν οι σχεσιακές βάσεις δεδομένων. Οι τεχνικές και μέθοδοι που χρησιμοποιούνταν στα συγκεκριμένα συστήματα βασίζονταν σε απάντηση απλών ερωτημάτων, αναζητήσεων, online επεξεργασία συναλλαγών, διαδραστικές οπτικοποιήσεις. Από τις αρχές του 21^{ου} αιώνα, τα παγκόσμια δίκτυα και το World Wide Web (WWW) παρείχε μια μοναδική ευκαιρία για τις επιχειρήσεις να διαθέτουν ηλεκτρονικές εφαρμογές και να αλληλεπιδρούν άμεσα με τους πελάτες (Sallam et al., 2011).

Η εξέλιξη των Big Data ώθησε τις επιχειρήσεις στην εφαρμογή τους για αξιοποίηση του συνόλου των δεδομένων. Η εφαρμογή των Big Data στις επιχειρήσεις μπορεί να ενισχύσει την αποτελεσματικότητα της παραγωγής τους και την ανταγωνιστικότητα σε πολλές πτυχές. Ειδικότερα, για την εμπορία, με τη ανάλυση συσχέτισης των Big Data, οι επιχειρήσεις μπορούν προβλέψουν με μεγαλύτερη ακρίβεια τη συμπεριφορά των καταναλωτών και να δημιουργήσουν νέες λειτουργίες και προϊόντα. Σχετικά με τον προγραμματισμό των πωλήσεων, μετά από σύγκριση των μαζικών δεδομένων, οι επιχειρήσεις μπορούν να βελτιστοποιήσουν βασικά προϊόντα τους σε επίπεδο τιμών, κόστους και ποιότητας. Αναφορικά με τη λειτουργία τους βελτιώνεται η αποτελεσματικότητα, η ικανοποίηση, η βελτιστοποίηση του εργατικού δυναμικού, η ακρίβεια πρόβλεψης απαιτήσεων στην κατανομή του προσωπικού, αποφεύγεται η πλεονάζουσα παραγωγική ικανότητα, μειώνεται το κόστος εργασίας. Η εφαρμογή των Big Data στις επιχειρήσεις πλέον επιτρέπει την ευκολότερη απογραφή, την υλικοτεχνική βελτιστοποίηση και τη συνεργασία με τους προμηθευτές για τον περιορισμό του χάσματος μεταξύ προσφοράς και ζήτησης ενώ και η χρηματοδότηση επιχειρήσεων που στρέφονται στα Big Data έχει αποκτήσει άλλη υπόσταση (Chen et al., 2014:194).

7.1.2. On line κοινωνικά δίκτυα προσαρμοσμένα στα Big Data

Αν υπάρχει ένα στοιχείο όπου οι εταιρείες social media και τα online κοινωνικά δίκτυα εξειδικεύονται είναι τα δεδομένα. Ο μεγάλος όγκος δεδομένων στα social media αντανακλά το πώς οι άνθρωποι αλληλεπιδρούν μεταξύ τους και στο επίκεντρο αυτών των αλληλεπιδράσεων βρίσκονται πολύτιμες πληροφορίες. Αυτός ο όγκος των δεδομένων, σε συνδυασμό με το γρήγορο ρυθμό της ροής δεδομένων για την οποία τα μέσα κοινωνικής δικτύωσης φημίζονται ουσιαστικά εκφράζουν τη φύση των Big Data (Huston, 2015).

Η εφαρμογή των Big Data στα κοινωνικά δίκτυα ξεπερνά τα στενά πλαίσια της αλληλεπίδρασης καθώς μπορεί να επηρεάσει συμπεριφορές (καταναλωτικές, κοινωνικές κλπ). Τα Big Data στα online κοινωνικά δίκτυα είναι η κινητήρια δύναμη πίσω από κάθε απόφαση marketing. Ακριβώς όπως συμβαίνει στα παραδοσιακά δίκτυα επικοινωνίας, στα social media εντοπίζονται στοιχεία χρήσιμα για την ανάλυση των αγορών και την πρόβλεψη της συμπεριφοράς των καταναλωτών. Το Facebook χειρίζεται πάνω από 500 terabytes δεδομένων κάθε μέρα, 300 εκατ. φωτογραφίες και 2,6 δισ. likes. Όλα αυτά τα δεδομένα υποβάλλονται σε επεξεργασία σε λίγα λεπτά δίνοντας στο Facebook πληροφορίες για τις αντιδράσεις των χρηστών και τη δυνατότητα να αναπτύξουν ή να τροποποιήσουν τη συμπεριφορά τους (Soundararajan, 2016; Addo - Tenkorang, 2016).

Το Facebook χρησιμοποιεί τα Big Data μέσω διαφόρων επιλογών. Χαρακτηριστικά, μέσω της αναδρομής προσφέρει στους χρήστες του τη δυνατότητα να βλέπουν και να μοιράζονται στοιχεία που σκιαγραφούν την πορεία της δραστηριότητάς τους στο συγκεκριμένο κοινωνικό δίκτυο από την ημερομηνία εγγραφής μέχρι σήμερα. Επίσης, μέσω της καταγραφής ψήφου των χρηστών επί διαφόρων θεμάτων όπως π.χ. οι εκλογές του 2010 μπορεί να δημιουργεί μια εικόνα των ψηφοφόρων ή ακόμα και πολιτικές τάσεις.



Εικόνα 7.1 : Εφαρμογή Big Data στο Facebook

Πηγή: Monnappa (2015)

Πολλές εταιρείες εκτιμούν την ισχυρή φύση των online κοινωνικών δικτύων για αλληλεπίδραση σε προσωπικό επίπεδο με τους πελάτες τους. Το Facebook, το Twitter, το Instagram και το Pinterest προσαρμόζονται στις δυνατότητες που μπορούν να προσφέρουν τα Big Data εξατομικεύοντας κόμη περισσότερο τις δυνατότητές τους προς τους χρήστες. Σε μελέτες πλέον έχει αποδειχθεί ότι το Facebook λόγω του μεγάλου όγκου των δεδομένων που διαχειρίζεται είναι ως ένα βαθμό ισχυρότερο εργαλείο αποτύπωσης της συμπεριφοράς και της προσωπικότητας ενός χρήστη ακόμα και από ανθρώπους του στενού περιβάλλοντος (Soundararajan, 2016; Addo - Tenkorang 2016).

Η απευθείας σύνδεση των online κοινωνικών δικτύων δημιουργεί τεράστιο όγκο δεδομένων και πληροφοριών τα οποία μέσω της κατάλληλης ανάλυσης, οργάνωσης αποθήκευσης και δυνατοτήτων ανάκτησης μπορεί να προσφέρει πολύ χρήσιμες πληροφορίες. Η πιο σημαντική θεμελιώδης αντίληψη της προσαρμογής των online κοινωνικών δικτύων στα Big Data είναι το γεγονός ότι τα social media είναι μέρος των μεγάλων δεδομένων. Δηλαδή, τα social media είναι μία από τις πιο σημαντικές πηγές Big Data. Το βασικό σημείο αναφοράς είναι ότι τα κοινωνικά δίκτυα (Facebook, Twitter, Instagram κλπ) δεν αποτελούν απλά πεδίο εφαρμογής των Big Data αλλά αποτελούν αναπόσπαστη πηγή δεδομένων και πληροφοριών

δημιουργώντας μια σχέση αλληλεξάρτησης (Soundararajan, 2016; Addo - Tenkorang, 2016).

7.1.3. Υγειονομικές και ιατρικές εφαρμογές Big Data

Ο τομέας της υγειονομικής περίθαλψης και των ιατρικών δεδομένων είναι συνεχώς και ταχέως εξελισσόμενος. Η αξιοποίηση των Big Data στο χώρο της υγείας προσφέρει πολύτιμη πληροφόρηση καθώς παρουσιάζουν απεριόριστες δυνατότητες για αποτελεσματική αποθήκευση, επεξεργασία, επερωτήσεις και ανάλυση ιατρικών δεδομένων. Η εφαρμογή των ιατρικών εφαρμογών Big Data στο μέλλον αναμένεται να επηρεάσει σημαντικά την οργάνωση και παροχή της υγειονομικής περίθαλψης. Για παράδειγμα, η Aetna Ασφαλιστική Εταιρεία Ζωής επέλεξε 102 ασθενείς μεταξύ 1.000 για να μελετήσει τη συνεισφορά των Big Data αναφορικά με την αντιμετώπιση των επιπτώσεων του μεταβολικού συνδρόμου. Ειδικότερα, αξιολογήθηκαν 600.000 αποτελέσματα εξετάσεων και 180.000 ελέγχων για το μεταβολικό σύνδρομο των ασθενών σε τρία συνεχόμενα έτη. Από την αξιολόγηση των στοιχείων προέκυψε το εξατομικευμένο σχέδιο θεραπείας για την αξιολόγηση των επικίνδυνων παραγόντων και τα κύρια σχέδια θεραπείας των ασθενών. Επίσης, από την αξιοποίηση των δεδομένων, το ιατρικό προσωπικό μπορεί να κατευθύνει σε μια μείωση της νοσηρότητας κατά 50% τα επόμενα 10 χρόνια με συνταγογράφηση στατίνων και βοηθώντας τους ασθενείς να χάσουν βάρος ή προτείνοντας τη μείωση των επιπέδων τριγλυκεριδίων (Chen et al., 2014:200).

Το Ιατρικό Κέντρο Mount Sinai στις Η.Π.Α. χρησιμοποιεί τεχνολογίες της Ayasdi, μιας εταιρείας Big Data, για να μελετά τις γενετικές αλληλουχίες του *Escherichia Coli*, συμπεριλαμβανομένων πάνω από 1 εκ. παραλλαγών του DNA και για να διερευνά «γιατί τα βακτηριακά στελέχη ανθίστανται των αντιβιοτικών». Επίσης, η HealthVault της Microsoft, που ξεκίνησε το 2007, είναι μια εξαιρετική εφαρμογή των ιατρικών Big Data. Στόχος της είναι να διαχειρίζεται ατομικές πληροφορίες αναφορικά με την υγεία των ασθενών. Η εφαρμογή υποδέχεται τις πληροφορίες για την υγεία μέσω κατάλληλων συσκευών είτε μέσω ανοικτής διασύνδεσης με τη χρήση κατάλληλου λογισμικού (Chen et al., 2014:201).

8. Ερευνητικό Μέρος

Η έρευνα διάρκειας δύο μηνών, πραγματοποιείται μέσα από τη δημιουργία ενός εργαλείου έρευνας του ερωτηματολογίου, ώστε από την στατιστική επεξεργασία των απαντήσεων να προκύψουν συμπεράσματα .

Το ερωτηματολόγιο δημιουργήθηκε με το Google Drive. Το συγκεκριμένο εργαλείο επιλέχθηκε καθώς είναι εύκολο στο χειρισμό. Είναι ανοιχτό και κυρίως πρόκειται για ένα εύκολο πρόγραμμα τόσο για τον ερευνητή όσο και για αυτόν που απαντάει στο ερωτηματολόγιο. Επιπλέον δίνει πολλές δυνατότητες στον δημιουργό και κατά τη διάρκεια της έρευνας αλλά και στο τέλος για την επεξεργασία των αποτελεσμάτων.

Η έρευνα πραγματοποιήθηκε με το διαμοιρασμό των ερωτηματολογίων σε δύο φάσεις. Η πρώτη φάση αποστολής-συλλογής απαντήσεων διήρκησε από τις 29/01/2017 που έγινε η διασπορά του ερωτηματολογίου μέχρι τις 10/02/2017. Ορισμένες δυσκολίες όπως η «απαίτηση να συμπληρώνει ο ερωτώμενος το email του ώστε να θεωρείται έγκυρη η απάντηση» απέτρεψε αρκετούς από την συμπλήρωση του ερωτηματολογίου. Η λεπτομέρεια αυτή εντοπίστηκε και έγιναν οι κατάλληλες ενέργειες για να μπορέσει η διαδικασία συμπλήρωσης του ερωτηματολογίου να είναι πιο εύκολη και πιο απλή.

Αφού προχωρήσαμε στις απαραίτητες διορθώσεις σε λάθη που εντοπίστηκαν σε ευρύτερο πλαίσιο, αλλά σε μερικές περιπτώσεις και μετά από υπόδειξη των συμμετεχόντων προχωρήσαμε στη δεύτερη φάση διαμοιρασμού.

Η δεύτερη αυτή φάση, η οποία ήταν αποτελεσματικότερη στη συλλογή απαντήσεων από την πρώτη, κράτησε από τις 13/02/2017 μέχρι τις 27/02/2017. Αυτή ήταν και η καταληκτική ημερομηνία συλλογής απαντήσεων.

Ο διαμοιρασμός των ερωτηματολογίων έγινε με ηλεκτρονικό τρόπο και με αποστολή στους επαγγελματίες πληροφόρησης, είτε σε επαγγελματικά ηλεκτρονικά ταχυδρομεία είτε σε προσωπικούς λογαριασμούς.

Οι ερωτήσεις που βρίσκονταν στο ερωτηματολόγιο ήταν κλειστού τύπου με προεπιλεγμένες τιμές και επιλογές.

Σκοπός της έρευνας αποτελεί η καταγραφή των απόψεων των επαγγελματιών της Πληροφόρησης σχετικά με τα Big Data στις ελληνικές βιβλιοθήκες. Συγκεκριμένα τέθηκαν ως ερευνητικοί στόχοι προς διερεύνηση, οι εξής:

1. Διερεύνηση χρήσης και υιοθέτησης των Big Data.
2. Διερεύνηση τεχνικών θεμάτων σχετικά με τα Big Data αναφορικά με: συστατικά στοιχεία, μεθόδους συλλογής, τεχνολογίες επεξεργασίας και μεθόδους επεξεργασίας.
3. Διερεύνηση προβλημάτων των Big Data.
4. Διερεύνηση νέων ρόλων επαγγελματιών της Πληροφόρησης από την εμπλοκή τους με τα Big Data.
5. Διερεύνηση αναδυόμενων ευκαιριών για τους επαγγελματίες της Πληροφόρησης από την εμπλοκή τους με τα Big Data.
6. Διερεύνηση προκλήσεων για τις Ελληνικές Βιβλιοθήκες από την εμπλοκή τους με τα Big Data.
7. Διερεύνηση μελλοντικής υιοθέτησης των Big Data από τις Ελληνικές Βιβλιοθήκες.
8. Διερεύνηση μελλοντικής διείσδυσης των Big Data σε κρίσιμους τομείς όπως η Τεχνολογική Έρευνα, τα Κοινωνικά Δίκτυα, οι Επιχειρήσεις, η Οικονομία και η Υγεία.

Ο σχεδιασμός του ερωτηματολογίου αποτέλεσε σημαντικό στοιχείο της έρευνας καθώς μέσω αυτού συλλέχθηκαν και επεξεργάστηκαν τα δεδομένα της έρευνας που θα αποδώσουν συμπεράσματα. Το ερωτηματολόγιο της παρούσας έρευνας περιλαμβάνει 22 ερωτήσεις κλειστού τύπου και προέκυψε μέσα από την βιβλιογραφία.

Το δείγμα της έρευνας αποτελούνταν από 49 Επαγγελματίες της Πληροφόρησης.

9. Αποτελέσματα Έρευνας

9.1. Ανάλυση Αποτελεσμάτων Έρευνας

Στην παρούσα ενότητα θα αποτυπωθούν τα αποτελέσματα των αναλύσεων της έρευνας που πραγματοποιήθηκε σε 49 επαγγελματίες της πληροφόρησης για την διερεύνηση των απόψεών τους σχετικά με τα Big Data αλλά και τους νέους ρόλους και τις ευκαιρίες που αναδύουν.

Για την ανάλυση της πρώτης ενότητας του ερωτηματολογίου αναφορικά με τα δημογραφικά στοιχεία του δείγματος μπορούμε να διακρίνουμε αρχικά τα ατομικά στοιχεία του δείγματος που συμμετείχε στην έρευνα. Πιο αναλυτικά:

1. **Φύλο:** Γυναίκες 32,7% και άνδρες 67,3%
2. **Ηλικία:** Μέσος όρος ηλικίας του δείγματος τοποθετείται στα 42,7 έτη με την μικρότερη ηλικία τα 30 έτη και την μεγαλύτερη ηλικία τα 57 έτη.
3. **Επίπεδο Μόρφωσης:** Το επίπεδο μόρφωσης του δείγματος είναι υψηλό με το 12,2% πτυχιούχων της Τριτοβάθμιας Εκπαίδευσης, 55,1% κάτοχοι Μεταπτυχιακού Διπλώματος και 32,7% κάτοχοι Διδακτορικού Διπλώματος .

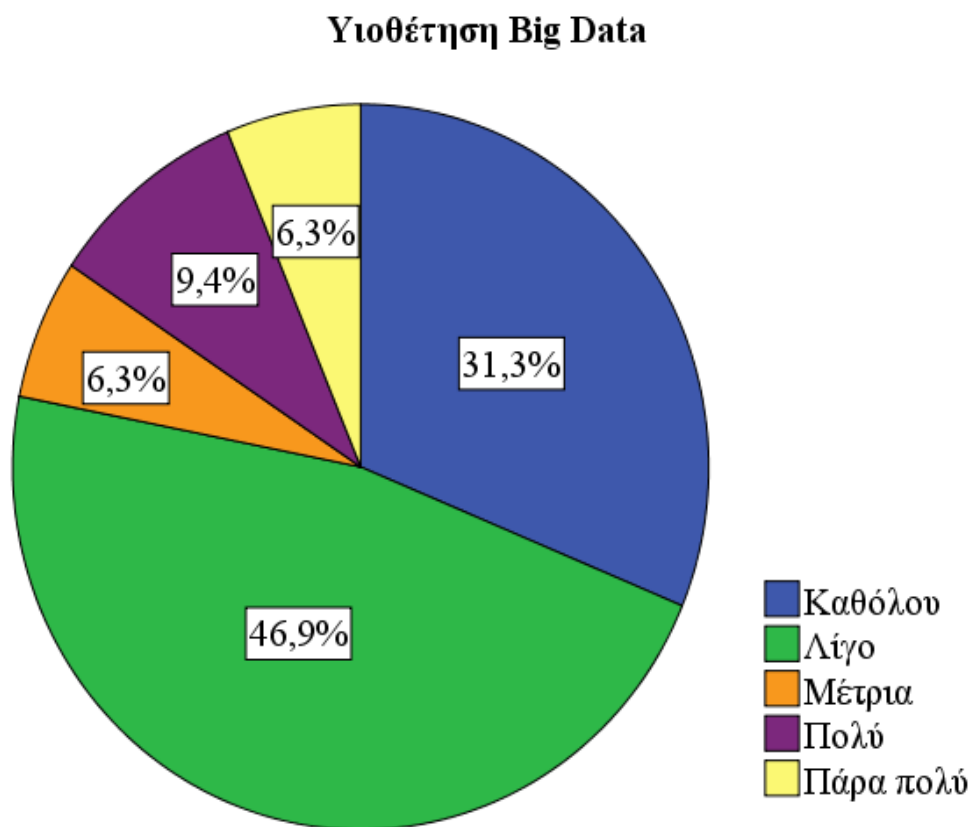
Εκτός των ατομικών στοιχείων πραγματοποιείται και ανάλυση των επαγγελματικών στοιχείων του δείγματος που συμμετείχε στην έρευνα. Πιο αναλυτικά:

- ✓ **Κλάδος Πληροφόρησης:** Το 67,3% του δείγματος απασχολείται στον κλάδο της Βιβλιοθηκονομίας .
- ✓ **Φορέας Απασχόλησης:** Το 79,6% του δείγματος απασχολείται σε Δημόσιο φορέα .
- ✓ **Τομέας Πληροφόρησης:** Το 69,4% του δείγματος απασχολείται σε κάποιο Ίδρυμα .

Για την ανάλυση της δεύτερης ενότητας του ερωτηματολογίου αναφορικά με τα Big Data αναδεικνύονται οι απόψεις του δείγματος που συμμετείχε στην έρευνα. Πιο αναλυτικά:

1. **Χρήση Big Data από Φορέα Απασχόλησης:** Το 32,6% των Φορέων στον οποίο απασχολείται το δείγμα χρησιμοποιεί τα Big Data.
2. **Σημασία Big Data:** Αναφορικά με την σημασία που αποδίδουν οι επαγγελματίες της πληροφόρησης στα Big Data σημειώνεται με 65,6% ο υψηλός όγκος data (δεδομένων) και η πολυσήμαντη διάσταση της πληροφόρησης με 18,8% .
3. **Υιοθέτηση Big Data:** Τα Big Data υιοθετούνται σε μικρό βαθμό μάλιστα σημειώθηκε 31,3% Καθόλου και 46,9% Λίγο (Διάγραμμα 1).

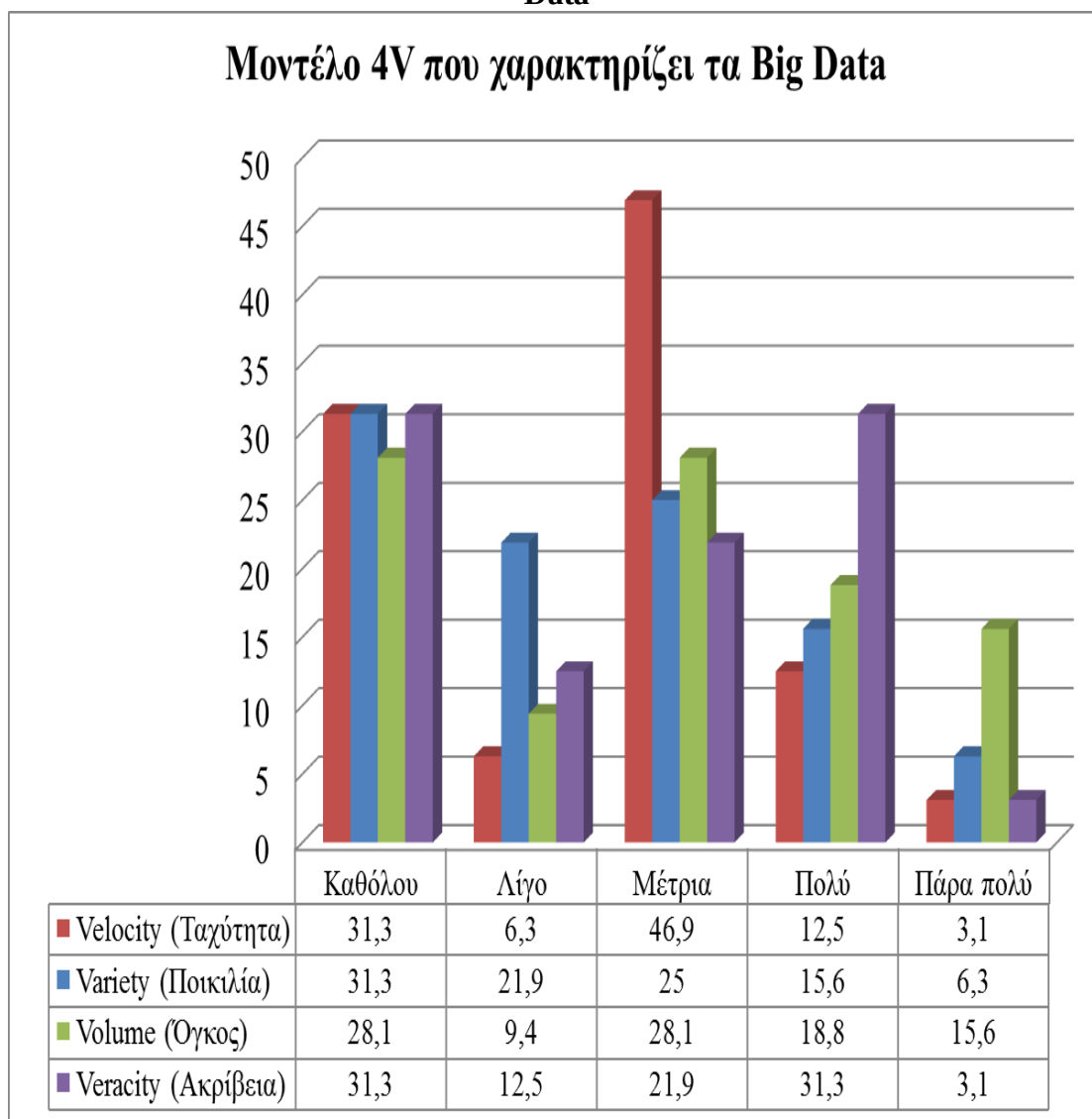
Διάγραμμα1 . Υιοθέτηση Big Data



4. **Χρήση Big Data:** Τα Big Data δεν χρησιμοποιούνται ιδιαίτερα, μάλιστα σημειώθηκε 29% Καθόλου και 48,4% Λίγο.

Η καταγραφή των απόψεων των επαγγελματιών της Πληροφόρησης σχετικά με τα συστατικά στοιχεία του Μοντέλου 4V: Volume, Velocity, Variety και Veracity που χαρακτηρίζουν τα Big Data της/του επιχείρησης/φορέα/οργανισμού όπου απασχολούνται είναι η εξής: το 46,9% των επαγγελματιών σημειώνει πως η Velocity (Ταχύτητα) χαρακτηρίζει τα Big Data σε μέτριο βαθμό, το 31,3% των επαγγελματιών σημειώνει ότι η Variety (Ποικιλία) δεν χαρακτηρίζει καθόλου τα Big Data και 21,9% τα χαρακτηρίζει σε μικρό βαθμό. Σχετικά με το Volume (Όγκο) το δείγμα διχάστηκε με 28,1% να σημειώνει τον ανύπαρκτο και αντίστοιχα με το ίδιο ποσοστό τον μέτριο επηρεασμό του στα Big Data. Τέλος, σχετικά με την Veracity (Ακρίβεια) οι επαγγελματίες διχάστηκαν και πάλι με το 31,3% να θεωρεί τον χαρακτηρισμό ανύπαρκτο, ενώ το ίδιο ποσοστό 31,3% σε μεγάλο βαθμό (Διάγραμμα 2).

Διάγραμμα 2. Μοντέλο 4V που χαρακτηρίζει τα Big Data



Στη συνέχεια ζητήθηκε η άποψη των επαγγελματιών της Πληροφόρησης σχετικά με τεχνικά θέματα αναφορικά με τα Big Data που χρησιμοποιεί η/ο επιχείρηση / φορέας / οργανισμός όπου και απασχολούνται. Σχετικά:

- ✓ **Μέθοδοι Συλλογής Big Data:** Το 46,9% του δείγματος σημείωσε πως δεν χρησιμοποιούνται τέτοιες μέθοδοι συλλογής Big Data στην εργασία τους, ενώ το 34,4% καταδεικνύει τα Αρχεία & Δεδομένα Δικτύου (Searching, Indexing).
- ✓ **Τεχνολογίες Επεξεργασίας Big Data:** Το 74,2% του δείγματος σημείωσε πως δεν χρησιμοποιούνται τέτοιες τεχνολογίες Big Data στην εργασία τους, ενώ το 12,9% καταδεικνύει την NoSQL.
- ✓ **Μέθοδοι Επεξεργασίας Big Data:** Το 64,5% του δείγματος σημείωσε πως δεν χρησιμοποιούνται τέτοιες τεχνολογίες επεξεργασίας στην εργασία τους, ενώ το 22,6% καταδεικνύει το Indexing .

Στη συνέχεια ζητήθηκε η άποψη των επαγγελματιών της Πληροφόρησης σχετικά με τα προβλήματα των Big Data που ταλανίζουν την/τον επιχείρηση / φορέα / οργανισμό όπου και απασχολούνται. Για την ανάλυση της παρούσας ερώτησης τα ποσοστά των απαντήσεων αναδεικνύουν την Διαφωνία (Καθόλου + Λίγο) των επαγγελματιών, την Ουδετερότητα (Μέτρια) και την Συμφωνία (Πολύ + Πάρα Πολύ). Σχετικά:

- ✓ **Έλλειψη ασφάλειας στοιχείων και προσωπικών δεδομένων:** Το 44,8% του δείγματος διαφωνεί, 27,6% είναι ουδέτερο και 27,6% συμφωνεί.
- ✓ **Έλλειψη ταχύτητας στην άντληση πληροφοριών:** Το 44,8% του δείγματος διαφωνεί, 24,1% είναι ουδέτερο και 31,1% συμφωνεί.
- ✓ **Λανθασμένες – συγκεχυμένες πληροφορίες:** Το 41,4% του δείγματος διαφωνεί, 37,9% είναι ουδέτερο και 20,7% συμφωνεί.
- ✓ **Μη τήρηση απορρήτου πληροφοριών:** Το 55,2% του δείγματος διαφωνεί, 27,6% είναι ουδέτερο και 17,2% συμφωνεί.
- ✓ **Προβλήματα υποδομών αποθήκευσης πληροφοριών:** Το 44,8% του δείγματος διαφωνεί, 10,3% είναι ουδέτερο και 44,9% συμφωνεί .

Για την ανάλυση της **τρίτης** ενότητας του ερωτηματολογίου αναφορικά με τους ρόλους και τις ευκαιρίες των επαγγελματιών της Πληροφόρησης παρατίθενται παρακάτω τα αποτελέσματα των απαντήσεών τους. Πιο ειδικά, οι ρόλοι των επαγγελματιών της Πληροφόρησης είναι δυνατό να ταξινομηθούν σε τέσσερις κύριες

κατηγορίες ανάλογα με το έργο που επιτελείται κάθε φορά:

1. Ο επαγγελματίας της Πληροφόρησης ως **Διαχειριστής Πληροφορίας και Γνώσης** είναι δυνατό να:
 - ✓ **Δημιουργεί οργανωτικά συστήματα διαχείρισης πληροφοριών:** Το 20,8% του δείγματος διαφωνεί, 14,6% είναι ουδέτερο και 64,6% συμφωνεί.
 - ✓ **Αναπτύσσει αρχιτεκτονικές γνωστικών και πληροφοριακών συστημάτων:** Το 22,9% του δείγματος διαφωνεί, 14,6% είναι ουδέτερο και 62,5% συμφωνεί .
 - ✓ **Ασχολείται με ψηφιοποίηση υλικού:** Το 10,4% του δείγματος διαφωνεί, 31,3% είναι ουδέτερο και 58,3% συμφωνεί .
 - ✓ **Προσφέρει υπηρεσίες εντοπισμού, επεξεργασίας και αναμόρφωσης πληροφοριών:** Το 6,3% του δείγματος διαφωνεί, 10,4% είναι ουδέτερο και 83,3% συμφωνεί .

2. Ο επαγγελματίας της Πληροφόρησης ως **Διαμεσολαβητής** είναι δυνατό να:
 - ✓ **Αντιμετωπίζει Δυσκολίες Χρηστών σε Αναζήτηση Πληροφοριών:** Το 8,3% του δείγματος διαφωνεί, 14,6% είναι ουδέτερο και 77,1% συμφωνεί .
 - ✓ **Παρέχει εξειδικευμένη πρόσβαση πελατών σε διευρυμένες πηγές πληροφόρησης:** Το 4,3% του δείγματος διαφωνεί, 17% είναι ουδέτερο και 78,7% συμφωνεί .

3. Ο επαγγελματίας της Πληροφόρησης ως **Διευκολυντής** είναι δυνατό να:
 - ✓ **Διευκολύνει τη σύνδεση χρηστών με έντυπα και ηλεκτρονικά προϊόντα:** Το 10,4% του δείγματος διαφωνεί, 29,2% είναι ουδέτερο και 60,4% συμφωνεί .
 - ✓ **Διευκολύνει τους χρήστες στη χρήση νέων πηγών πληροφόρησης & γνώσης:** Το 10,4% του δείγματος διαφωνεί, 18,8% είναι ουδέτερο και 70,8% συμφωνεί .
 - ✓ **Οργανώνει τις πηγές πληροφόρησης και γνώσης:** Το 14,6% του δείγματος διαφωνεί, 20,8% είναι ουδέτερο και 64,6% συμφωνεί .
 - ✓ **Παρέχει βοήθεια σε χρήστες για απόκτηση ικανοτήτων πληροφοριακής παιδείας για εκτίμηση των πληροφοριών:** Το 14,6% του δείγματος διαφωνεί, 12,2% είναι ουδέτερο και 73,2% συμφωνεί .

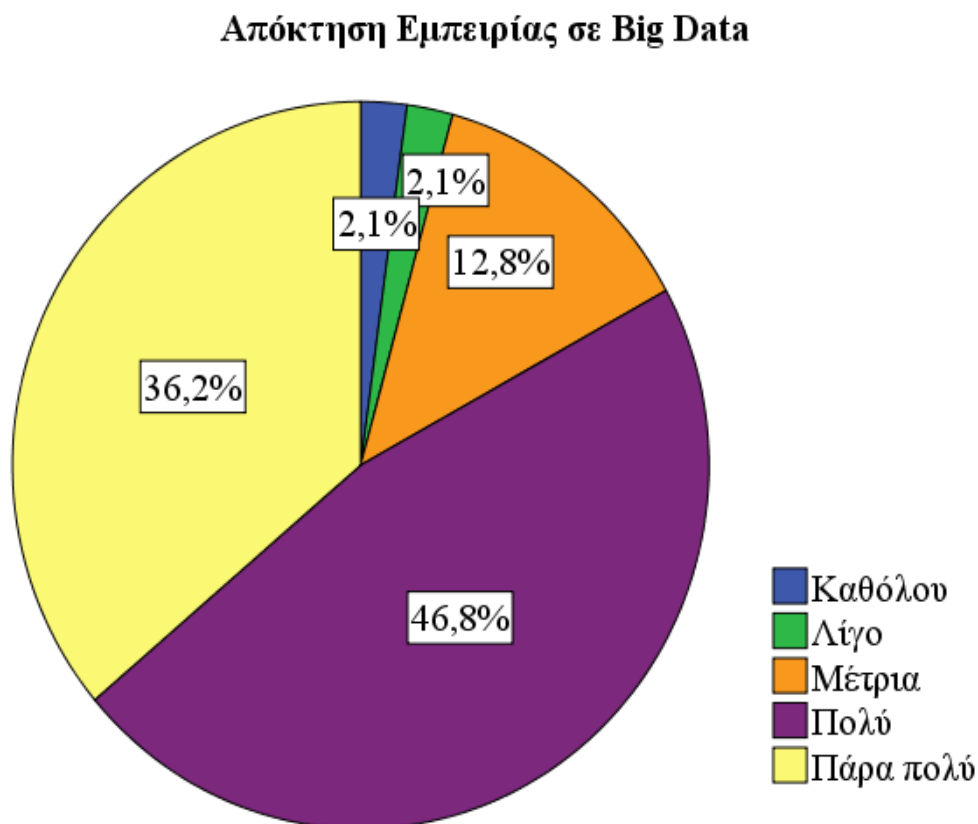
- ✓ **Παρέχει πληροφόρηση σε χρήστες:** Το 14,6% του δείγματος διαφωνεί, 12,5% είναι ουδέτερο και 72,9% συμφωνεί .
4. Ο επαγγελματίας της Πληροφόρησης ως **Εκπαιδευτής** είναι δυνατό να:
- ✓ **Καθοδηγεί χρήστες στη διαδικασία αναζήτησης, ανάκτησης, ερμηνείας και χρήσης πληροφοριών:** Το 14,6% του δείγματος διαφωνεί, 12,5% είναι ουδέτερο και 72,9% συμφωνεί .
 - ✓ **Παρέχει εκπαιδευτικά προγράμματα για απομακρυσμένους χρήστες:** Το 21,3% του δείγματος διαφωνεί, 21,3% είναι ουδέτερο και 57,4% συμφωνεί .

Σε επόμενη ερώτηση ζητήθηκε από τους επαγγελματίες της Πληροφόρησης να απαντήσουν αν κατά την εμπλοκή τους με τα Big Data είναι δυνατό να επωφεληθούν από ορισμένες αναδυόμενες ευκαιρίες.

Η άποψη των επαγγελματιών της Πληροφόρησης σχετικά με τεχνικά θέματα αναφορικά με τις αναδυόμενες ευκαιρίες των Big Data είναι οι εξής:

1. **Απόκτηση εμπειρίας σε Big Data:** Το 4,3% του δείγματος διαφωνεί, 12,8% είναι ουδέτερο και 82,9% συμφωνεί (Διάγραμμα 3) .

Διάγραμμα 3. Απόκτηση Εμπειρίας σε Big Data



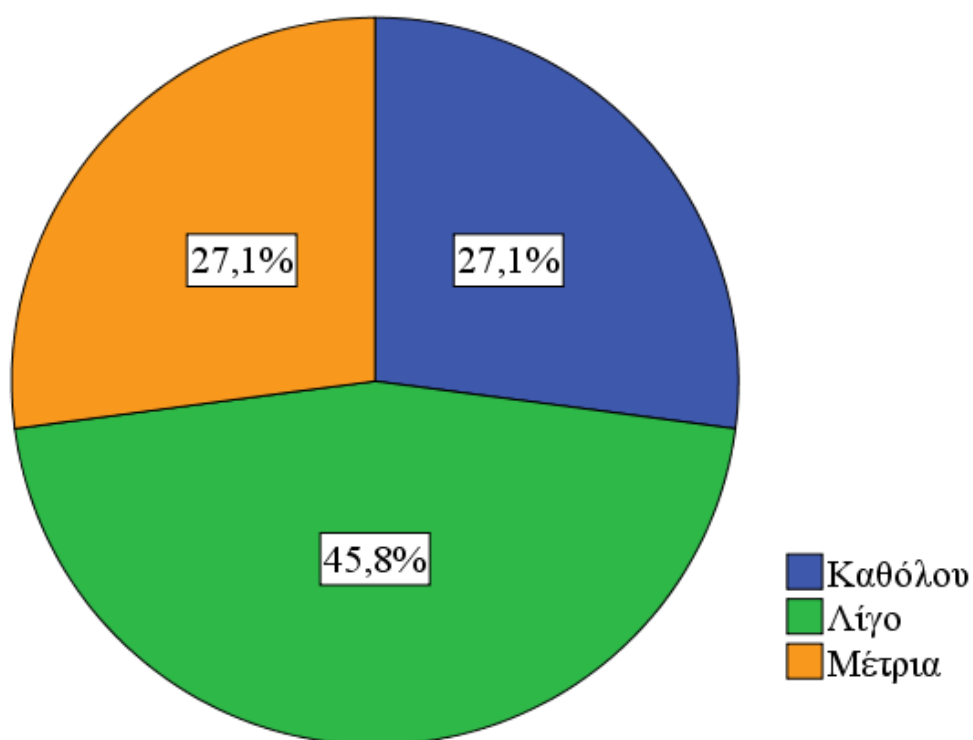
2. **Απόκτηση εξειδίκευσης σε Big Data:** Το 6,4% του δείγματος διαφωνεί, 14,9% είναι ουδέτερο και 78,7% συμφωνεί .
3. **Απόκτηση τεχνογνωσίας σε Big Data:** Το 8,5% του δείγματος διαφωνεί, 12,8% είναι ουδέτερο και 78,7% συμφωνεί .
4. **Διευκόλυνση εργασίας: Επάρκεια λογισμικού, βάσης δεδομένων & πληροφοριών:** Το 6,4% του δείγματος διαφωνεί, 21,3% είναι ουδέτερο και 72,3% συμφωνεί .
5. **Εξασφάλιση ανταγωνιστικού πλεονεκτήματος:** Το 10,6% του δείγματος διαφωνεί, 21,3% είναι ουδέτερο και 68,1% συμφωνεί .
6. **Εξέλιξη (πχ. προαγωγή, αύξηση οικονομικών αποδοχών):** Το 21,3% του δείγματος διαφωνεί, 34% είναι ουδέτερο και 44,7% συμφωνεί .
7. **Νέες θέσεις εργασίας σε βιβλιοθήκες:** Το 25,5% του δείγματος διαφωνεί, 23,4% είναι ουδέτερο και 51,1% συμφωνεί .
8. **Νέες θέσεις εργασίας σε νέα πεδία απασχόλησης (πχ. διαχείριση γνώσης, ανάπτυξη δικτυακών πόρων):** Το 12,8% του δείγματος διαφωνεί, 25,5% είναι ουδέτερο και 61,7% συμφωνεί .

Για την ανάλυση της **τέταρτης** ενότητας του ερωτηματολογίου αναφορικά με τις προκλήσεις Big Data για τις Ελληνικές Βιβλιοθήκες. Ειδικά σχετικά με το αν οι Ελληνικές Βιβλιοθήκες διαθέτουν θετική άποψη για τη νέα πρόκληση των Big Data. Οι επαγγελματίες της Πληροφόρησης συμφώνησαν κατά 12,5%, κατά 37,5% διαφώνησαν και 50% παρέμεναν ουδέτεροι.

Αναφορικά με το βαθμό υιοθέτησης στην πράξη των Big Data από τις Ελληνικές Βιβλιοθήκες, οι επαγγελματίες της Πληροφόρησης συμφώνησαν ότι έχουν υιοθετηθεί κατά 27,1%, κατά 45,8% διαφώνησαν και 27,1% παρέμεναν ουδέτεροι χωρίς να πάρουν θέση αν οι ελληνικές βιβλιοθήκες έχουν υιοθετήσει τα Big Data (Διάγραμμα 4)

Διάγραμμα 4. Ελληνικές Βιβλιοθήκες: Πρακτική Υιοθέτηση Big Data

Ελληνικές Βιβλιοθήκες: Πρακτική Υιοθέτηση Big Data



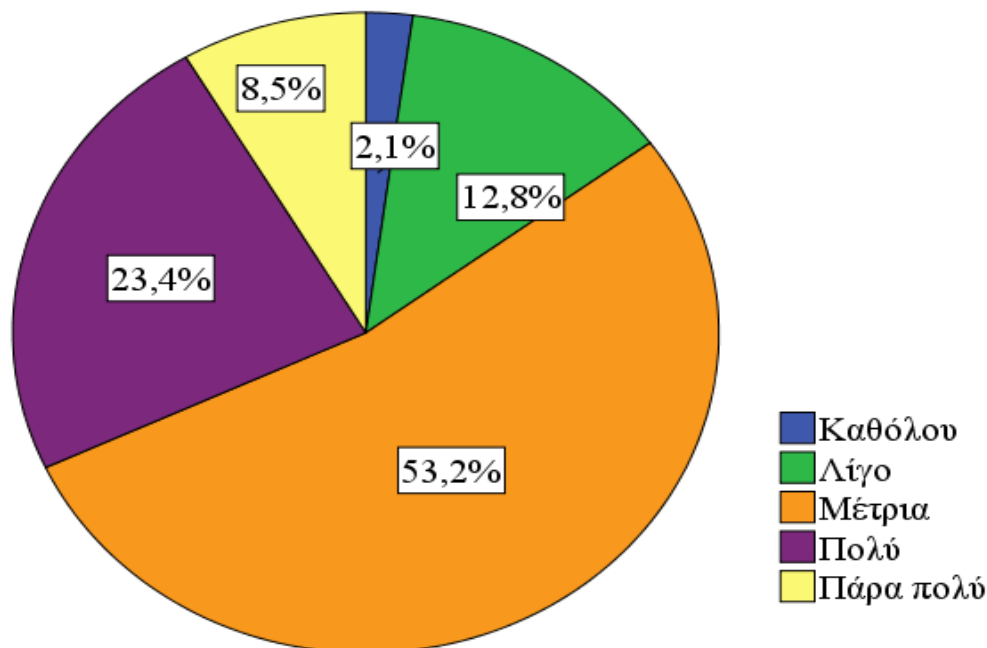
Η άποψη των επαγγελματιών της Πληροφόρησης σχετικά με τις προκλήσεις των Big Data για τις Ελληνικές Βιβλιοθήκες είναι οι εξής:

1. **Διασφάλιση ιδιωτικότητας των δεδομένων:** Το 6,4% του δείγματος διαφωνεί, 34% είναι ουδέτερο και 59,6% συμφωνεί .
2. **Πληροφορία: Επικαιροποίηση, οργάνωση ετερογένειας και ατελειών, υψηλή ταχύτητα άντλησης, ανάλυσης, σύνθεσης, διαχείρισης:** Το 2,1% του δείγματος διαφωνεί, 17% είναι ουδέτερο και 80,9% συμφωνεί .
3. **Συνεργασία επαγγελματιών της πληροφόρησης:** Το 4,3% του δείγματος διαφωνεί, 17,4% είναι ουδέτερο και 78,3% συμφωνεί .
4. **Ανταγωνιστικό πλεονέκτημα στην Οικονομία της Γνώσης:** Το 6,4% του δείγματος διαφωνεί, 21,3% είναι ουδέτερο και 72,3% συμφωνεί .
5. **Δημιουργία υποδομών για μετασχηματισμό δεδομένων σε πληροφορίες και των πληροφοριών σε γνώση:** Το 8,5% του δείγματος διαφωνεί, 14,9% είναι ουδέτερο και 76,6% συμφωνεί .
6. **Ειδικές επαγγελματικές δεξιότητες βιβλιοθηκονόμων:** Το 6,4% του δείγματος διαφωνεί, 19,1% είναι ουδέτερο και 74,5% συμφωνεί .

Αναφορικά με το βαθμό της υιοθέτησης των Big Data στο μέλλον από τις Ελληνικές Βιβλιοθήκες, οι επαγγελματίες της Πληροφόρησης συμφώνησαν κατά 31,9%, κατά 14,9% διαφώνησαν και 53,2% παρέμεναν ουδέτεροι (Διάγραμμα 5) .

Διάγραμμα 5. Ελληνικές Βιβλιοθήκες: Μελλοντική Υιοθέτηση Big Data

Ελληνικές Βιβλιοθήκες: Μελλοντική Υιοθέτηση Big Data



Αναφορικά με τους κρίσιμους τομείς που θα διεισδύσουν μελλοντικά τα Big Data, οι επαγγελματίες της Πληροφόρησης θεωρούν τα εξής: Οι κρίσιμοι τομείς είναι η Τεχνολογική Έρευνα (55,1%), τα Κοινωνικά Δίκτυα (49%), οι Επιχειρήσεις (46,9%), η Οικονομία (26,5%) και η Υγεία (26,5%), (Πίνακας 1).

Πίνακας 1. Κρίσιμοι Τομείς Μελλοντικής Διείσδυσης Big Data

Τομείς Μελλοντικής Διείσδυσης Big Data	Πλήθος	Ποσοστό (%)
Επιχειρήσεις	23	46,90%
Κοινωνικά Δίκτυα	24	49,00%
Οικονομία	13	26,50%
Τεχνολογική Έρευνα	27	55,10%
Υγεία	13	26,50%
Άλλο	2	4,10%

9.2. Συνδυαστικά Αποτελέσματα Έρευνας

Προέκυψαν σημαντικά αποτελέσματα από τον συνδυασμό ορισμένων ερωτήσεων που ουσιαστικά απαντούν σε ερωτήματα σχετικά με την έρευνα που έχουν τεθεί στην προσπάθεια του ερευνητή να διερευνήσει το πεδίο των Big Data στις ελληνικές βιβλιοθήκες και να αναδείξει τις απόψεις των επαγγελματιών της πληροφόρησης σχετικά. Ειδικότερα:

1. Οι γυναίκες *δεν πραγματοποιούν* χρήση των Big Data (83,4%) & δεν φαίνεται να τα υιοθετούν (83,3%) σε αντίθεση με τους άνδρες που τα χρησιμοποιούν κατά 15,8% και τα έχουν υιοθετήσει κατά 15% περισσότερο.
2. Ο Δημόσιος τομέας αναδεικνύει τα εξής προβλήματα των Big Data σε υψηλό βαθμό:
 - ✓ Προβλήματα υποδομών αποθήκευσης πληροφοριών (27,6%)
 - ✓ Λανθασμένες – συγκεχυμένες πληροφορίες (20,7%)
 - ✓ Έλλειψη ασφάλειας στοιχείων & προσωπικών δεδομένων (17,2%)
 - ✓ Έλλειψη ταχύτητας άντλησης πληροφοριών (10,3%)
 - ✓ Μη τήρηση απορρήτου πληροφοριών (10,3%)
3. Ο Ιδιωτικός τομέας αναδεικνύει τα εξής προβλήματα των Big Data σε υψηλό βαθμό:
 - ✓ Προβλήματα υποδομών αποθήκευσης πληροφοριών (17,2%)
 - ✓ Λανθασμένες – συγκεχυμένες πληροφορίες (10,3%)
 - ✓ Έλλειψη ασφάλειας στοιχείων & προσωπικών δεδομένων (10,3%)
 - ✓ Έλλειψη ταχύτητας άντλησης πληροφοριών (10,3%)
 - ✓ Μη τήρηση απορρήτου πληροφοριών (6,8%)
4. Ο Δημόσιος τομέας αναδεικνύει ως σημαντικότερους νέους ρόλους για τους επαγγελματίες της πληροφόρησης και τις ελληνικές βιβλιοθήκες τους εξής:
 - ✓ **Διαχειριστής Πληροφορίας Και Γνώσης:** Δημιουργεί οργανωτικά συστήματα διαχείρισης πληροφοριών (52,1%), Αναπτύσσει αρχιτεκτονικές γνωστικών & πληροφοριακών συστημάτων (50%), Ασχολείται με ψηφιοποίηση υλικού (50,1%), Προσφέρει υπηρεσίες εντοπισμού, επεξεργασίας & αναμόρφωσης πληροφοριών (66,7%),

- ✓ **Διαμεσολαβητής:** Αντιμετωπίζει τις δυσκολίες χρηστών κατά την αναζήτηση πληροφοριών (64,6%), Παρέχει εξειδικευμένη πρόσβαση πελατών σε διευρυμένες πηγές πληροφόρησης (53,2%),
 - ✓ **Διευκολυντής:** Διευκολύνει τη σύνδεση χρηστών με έντυπα και ηλεκτρονικά προϊόντα (47,9%), Διευκολύνει τους χρήστες στη χρήση νέων πηγών πληροφόρησης & γνώσης (56,3%), Οργανώνει τις πηγές πληροφόρησης & γνώσης (52,3%), Παρέχει βοήθεια σε χρήστες για απόκτηση ικανοτήτων πληροφοριακής παιδείας για εκτίμηση των πληροφοριών (50%), Παρέχει πληροφόρηση σε χρήστες (60,4%).
 - ✓ **Εκπαιδευτής:** Καθοδηγεί χρήστες στη διαδικασία αναζήτησης, ανάκτησης, ερμηνείας και χρήσης πληροφοριών (58,4%), παροχή εκπαιδευτικών προγραμμάτων για απομακρυσμένους χρήστες (42,6%).
5. Ο Ιδιωτικός τομέας είναι ιδιαίτερα επιφυλακτικός αναφορικά με νέους ρόλους για τους επαγγελματίες της πληροφόρησης και τις ελληνικές βιβλιοθήκες γεγονός που αναδεικνύεται από τα παρακάτω ποσοστά θετικών απαντήσεων (πολύ και πάρα πολύ):
- ✓ **Διαχειριστής Πληροφορίας Και Γνώσης:** Δημιουργεί οργανωτικά συστήματα διαχείρισης πληροφοριών (12,5%), αναπτύσσει αρχιτεκτονικές γνωστικών & πληροφοριακών συστημάτων (12,5%), ασχολείται με ψηφιοποίηση υλικού (8,3%), προσφέρει υπηρεσίες εντοπισμού, επεξεργασίας & αναμόρφωσης πληροφοριών (16,7%),
 - ✓ **Διαμεσολαβητής:** Αντιμετωπίζει τις δυσκολίες χρηστών κατά την αναζήτηση πληροφοριών (10,5%), παρέχει εξειδικευμένη πρόσβαση πελατών σε διευρυμένες πηγές πληροφόρησης (14,9%),
 - ✓ **Διευκολυντής:** Διευκολύνει τη σύνδεση χρηστών με έντυπα και ηλεκτρονικά προϊόντα (12,5%), διευκολύνει τους χρήστες στη χρήση νέων πηγών πληροφόρησης & γνώσης (14,6%), οργανώνει τις πηγές πληροφόρησης & γνώσης (12,3%), παρέχει βοήθεια σε χρήστες για απόκτηση ικανοτήτων πληροφοριακής παιδείας για εκτίμηση των πληροφοριών (17,6%), παρέχει πληροφόρηση σε χρήστες (12,5%).
 - ✓ **Εκπαιδευτής:** Καθοδηγεί χρήστες στη διαδικασία αναζήτησης, ανάκτησης, ερμηνείας και χρήσης πληροφοριών (14,6%), παροχή εκπαιδευτικών προγραμμάτων για απομακρυσμένους χρήστες (14,9%).

6. Οι **νέοι ρόλοι** για τους επαγγελματίες της πληροφόρησης λαμβάνουν διαφορετικής αντιμετώπισης ανάλογα με το φύλο των ερωτώμενων. Για τις **γυναίκες** σημαντικότεροι ρόλοι θεωρήθηκαν η αντιμετώπιση των δυσκολιών των χρηστών κατά την αναζήτηση πληροφοριών (22,9%) και η παροχή εξειδικευμένης πρόσβασης πελατών σε διευρυμένες πηγές πληροφόρησης (21,3%). Οι **άνδρες** θεωρούν σημαντικότερους τους εξής ρόλους: Διευκολύνει τη σύνδεση χρηστών με έντυπα και ηλεκτρονικά προϊόντα (48%), παρέχει εξειδικευμένη πρόσβαση πελατών σε διευρυμένες πηγές πληροφόρησης (48,9%), αντιμετωπίζει τις δυσκολίες χρηστών κατά την αναζήτηση πληροφοριών (52,1%), καθοδηγεί χρήστες στη διαδικασία αναζήτησης, ανάκτησης, ερμηνείας και χρήσης πληροφοριών (52,1%), διευκολύνει τους χρήστες στη χρήση νέων πηγών πληροφόρησης & γνώσης (54,2%) και παρέχει πληροφόρηση σε χρήστες (54,2%).
7. Οι **νέες ευκαιρίες** για τους επαγγελματίες της πληροφόρησης λαμβάνουν διαφορετικής αντιμετώπισης ανάλογα με το φύλο των ερωτώμενων. Για τις **γυναίκες** σημαντικότερες ευκαιρίες αποτελούν οι νέες θέσεις εργασίας σε νέα πεδία απασχόλησης (π.χ. διαχείριση γνώσης, ανάπτυξη δικτυακών πόρων) (19,1%), απόκτηση εξειδίκευσης (21,3%), απόκτηση εμπειρίας (23,4%) και τεχνογνωσίας σε Big Data (23,4%). Οι **άνδρες** θεωρούν σημαντικότερες τις εξής ευκαιρίες: την απόκτηση εμπειρίας (59,6%), εξειδίκευσης (57,5%), τεχνογνωσίας (55,4%) και διευκόλυνση εργασίας (επάρκεια λογισμικού, βάσης δεδομένων & πληροφοριών (51%).
8. Οι **νέες ευκαιρίες** για τους επαγγελματίες της πληροφόρησης λαμβάνουν διαφορετικής αντιμετώπισης ανάλογα με το φορέα δημόσιο ή ιδιωτικό. Στο δημόσιο τομέα αναδεικνύονται σημαντικότερες ευκαιρίες οι: απόκτηση εμπειρίας (66%), εξειδίκευσης (66%), τεχνογνωσίας (66%), διευκόλυνση εργασίας (επάρκεια λογισμικού, βάσης δεδομένων & πληροφοριών (57,5%) και νέες θέσεις εργασίας σε νέα πεδία απασχόλησης (π.χ. διαχείριση γνώσης, ανάπτυξη δικτυακών πόρων) (52,2%). Ενώ στον ιδιωτικό τομέα αναδεικνύονται απόκτηση εμπειρίας (16,8%) και με 12,8% η απόκτηση εξειδίκευσης, τεχνογνωσίας, η διευκόλυνση εργασίας, η εξασφάλιση ανταγωνιστικού πλεονεκτήματος και οι νέες θέσεις εργασίας σε βιβλιοθήκες.
9. Οι **προκλήσεις** των Big Data για τις Ελληνικές Βιβλιοθήκες λαμβάνουν διαφορετικής αντιμετώπισης ανάλογα με το φύλο των ερωτώμενων. Για τις

γυναίκες σημαντικότερες αποτελούν η πληροφορία (επικαιροποίηση, οργάνωση ετερογένειας και ατελειών, υψηλή ταχύτητα άντλησης, ανάλυσης, σύνθεσης, διαχείρισης) (23,4%) και το ανταγωνιστικό πλεονέκτημα στην Οικονομία της Γνώσης (25,5%). Οι **άνδρες** θεωρούν σημαντικότερες τις εξής προκλήσεις: πληροφορία (επικαιροποίηση, οργάνωση ετερογένειας και ατελειών, υψηλή ταχύτητα άντλησης, ανάλυσης, σύνθεσης, διαχείρισης) (57,5%), συνεργασία επαγγελματιών της πληροφόρησης (60,9%), ειδικές επαγγελματικές δεξιότητες βιβλιοθηκονόμων (53,2%).

10. Οι **προκλήσεις** των Big Data για τις Ελληνικές Βιβλιοθήκες λαμβάνουν διαφορετικής αντιμετώπισης ανάλογα με το φορέα δημόσιο ή ιδιωτικό. Στο **δημόσιο** τομέα αναδεικνύονται ως σημαντικότερες προκλήσεις οι: πληροφορία (επικαιροποίηση, οργάνωση ετερογένειας και ατελειών, υψηλή ταχύτητα άντλησης, ανάλυσης, σύνθεσης, διαχείρισης) (57,5%), συνεργασία επαγγελματιών της πληροφόρησης (60,9%), ανταγωνιστικό πλεονέκτημα στην Οικονομία της Γνώσης (59,6%) και οι ειδικές επαγγελματικές δεξιότητες βιβλιοθηκονόμων (61,7%). Ενώ στον **ιδιωτικό** τομέα αναδεικνύονται η πληροφορία (επικαιροποίηση, οργάνωση ετερογένειας και ατελειών, υψηλή ταχύτητα άντλησης, ανάλυσης, σύνθεσης, διαχείρισης) (17,1%) και η συνεργασία επαγγελματιών της πληροφόρησης (17,4%).

9.3. Συμπεράσματα Έρευνας

Προέκυψαν σημαντικά συμπεράσματα που ουσιαστικά απαντούν στα ερωτήματα της έρευνας που έχουν τεθεί στην προσπάθεια του ερευνητή να διερευνήσει το πεδίο των Big Data στις ελληνικές βιβλιοθήκες και να αναδείξει τις απόψεις των επαγγελματιών της πληροφόρησης σχετικά.

1. **Κοινωνικο/Δημογραφικά Στοιχεία Δείγματος:** Αναφορικά με τα ατομικά στοιχεία του δείγματος που συμμετείχε στην έρευνα παρατηρήθηκε ότι ήταν κατά κύριο λόγο ανδροκρατούμενο (67,3%), με το μέσο όρο ηλικίας τα 42,7 έτη, υψηλής μόρφωσης με το 55,1% των επαγγελματιών να είναι κάτοχοι Μεταπτυχιακού Διπλώματος και 32,7% κάτοχοι Διδακτορικού Διπλώματος. Σχετικά με τα επαγγελματικά χαρακτηριστικά του δείγματος και με την εξειδίκευση τους, οι επαγγελματίες απασχολούνται στον Κλάδο της Βιβλιοθηκονομίας κατά 67,3%, σε δημόσιο φορέα (79,6%) και ειδικότερα σε κάποιο Ίδρυμα (69,4%).
2. **Big Data:** Οι επαγγελματίες της Πληροφόρησης καταθέτουν χρήσιμες απόψεις αναφορικά με τα Big Data στην/στον επιχείρηση/φορέα/οργανισμό όπου και απασχολούνται. Το 32,6% των φορέων απασχόλησης χρησιμοποιεί τα Big Data, αποδίδεται σημασία στον υψηλό όγκο δεδομένων κατά 65,6% και στην διάσταση της πληροφόρησης κατά 18,8%. Μόλις το 15,7% των φορέων έχει υιοθετήσει τα Big Data σε υψηλό βαθμό, 6,3% σε μέτριο βαθμό, 46,9% σε μικρό βαθμό και 31,3% καθόλου. Ακόμη, τα συστατικά στοιχεία του Μοντέλου 4V χαρακτηρίζουν τα Big Data του φορέα σε υψηλό βαθμό: κατά 21,9% η Variety (Ποικιλία), 15,6% η Velocity (Ταχύτητα), 34,4% ο Volume (Όγκος) και 34,4% η Veracity (Ακρίβεια). Σχετικά με τεχνικά θέματα για τα Big Data αναδείχθηκε ως κύρια μέθοδος συλλογής Big Data κατά 34,4% τα Αρχεία και Δεδομένα Δικτύου (Searching, Indexing) ενώ κατά 46,9% δεν χρησιμοποιείται κάποια μέθοδο συλλογής. Ως κύρια τεχνολογία επεξεργασίας Big Data αναδείχθηκε με 12,9% η NoSQL ενώ 74,2% των φορέων δεν χρησιμοποιεί κάποια τεχνολογία. Αναφορικά με τη μέθοδο επεξεργασίας Big Data κατά 22,6% καταδεικνύει το Indexing ενώ 64,5% των φορέων δεν χρησιμοποιεί κάποια μέθοδο επεξεργασίας. Επίσης αναδεικνύονται προβλήματα των Big Data τα οποία ταξινομούνται ανάλογα με τον βαθμό συμφωνίας των

επαγγελματιών σε προβλήματα υποδομών αποθήκευσης πληροφοριών (44,9%), έλλειψη ταχύτητας στην άντληση πληροφοριών (31,1%), έλλειψη ασφάλειας στοιχείων και προσωπικών δεδομένων (27,6%), λανθασμένες – συγκεχυμένες πληροφορίες (20,7%) και μη τήρηση απορρήτου πληροφοριών (17,2%).

3. **Big Data: Ρόλοι του επαγγελματία της Πληροφόρησης.** Οι ρόλοι των επαγγελματιών της Πληροφόρησης ταξινομούνται ως προς την σημαντικότητά τους ανάλογα με τον βαθμό συμφωνίας:

- ✓ Ο επαγγελματίας της Πληροφόρησης ως **Διαχειριστής** της πληροφορίας και της γνώσης είναι δυνατό να προσφέρει υπηρεσίες εντοπισμού, επεξεργασίας και αναμόρφωσης πληροφοριών (83,3%), να δημιουργεί οργανωτικά συστήματα διαχείρισης πληροφοριών (64,6%), να ασχολείται με ψηφιοποίηση υλικού (58,3%) και να αναπτύσσει αρχιτεκτονικές γνωστικών και πληροφοριακών συστημάτων (62,5%).
- ✓ Ο επαγγελματίας της Πληροφόρησης ως **Διαμεσολαβητής** είναι δυνατό να παρέχει εξειδικευμένη πρόσβαση πελατών σε διευρυμένες πηγές πληροφόρησης (78,7%) και να δημιουργεί οργανωτικά συστήματα διαχείρισης πληροφοριών (77,1%).
- ✓ Ο επαγγελματίας της Πληροφόρησης ως **Διευκολυντής** είναι δυνατό να παρέχει βοήθεια σε χρήστες για απόκτηση ικανοτήτων πληροφοριακής παιδείας για εκτίμηση των πληροφοριών (73,2%), να παρέχει πληροφόρηση σε χρήστες (72,9%), να διευκολύνει τους χρήστες στη χρήση νέων πηγών πληροφόρησης & γνώσης (70,8%), να οργανώνει τις πηγές πληροφόρησης και γνώσης (64,6%) και να διευκολύνει τη σύνδεση χρηστών με έντυπα και ηλεκτρονικά προϊόντα (60,4%).
- ✓ Ο επαγγελματίας της Πληροφόρησης ως **Εκπαιδευτής** είναι δυνατό να καθοδηγεί χρήστες στη διαδικασία αναζήτησης, ανάκτησης, ερμηνείας και χρήσης πληροφοριών (72,9%) και να παρέχει εκπαιδευτικά προγράμματα για απομακρυσμένους χρήστες (57,4%).

4. **Big Data: Ευκαιρίες του επαγγελματία της Πληροφόρησης.** Ως αναδυόμενες ευκαιρίες των επαγγελματιών της Πληροφόρησης θεωρήθηκαν οι εξής: απόκτηση εμπειρίας σε big data (82,9%), απόκτηση εξειδίκευσης σε big data (78,7%), απόκτηση τεχνογνωσίας σε big data (78,7%), διευκόλυνση εργασίας: επάρκεια λογισμικού, βάσης δεδομένων & πληροφοριών (72,3%), εξασφάλιση ανταγωνιστικού πλεονεκτήματος (68,1%), νέες θέσεις εργασίας σε

νέα πεδία απασχόλησης (πχ. διαχείριση γνώσης, ανάπτυξη δικτυακών πόρων) (61,7%), νέες θέσεις εργασίας σε βιβλιοθήκες (51,1%) και εξέλιξη (πχ. προαγωγή, αύξηση οικονομικών αποδοχών) (44,7%).

5. **Big Data: Προκλήσεις για τις Ελληνικές Βιβλιοθήκες.** Οι επαγγελματίες της Πληροφόρησης παρουσιάστηκαν μετριοπαθείς (50%) για το αν οι Ελληνικές Βιβλιοθήκες διαθέτουν θετική άποψη για τη νέα πρόκληση των Big Data. Μάλιστα ο βαθμός υιοθέτησής τους από τις Ελληνικές Βιβλιοθήκες είναι χαμηλός μόλις στο 27,1%. Αν και οι αναδυόμενες προκλήσεις κερδίζουν έδαφος όπως η πληροφορία (επικαιροποίηση, οργάνωση ετερογένειας και ατελειών, υψηλή ταχύτητα άντλησης, ανάλυσης, σύνθεσης, διαχείρισης) (80,9%), η συνεργασία επαγγελματιών της πληροφόρησης (78,3%). Για το μέλλον διαφαίνεται μια θετική τάση (31,9%) για την υιοθέτηση των Big Data από τις Ελληνικές Βιβλιοθήκες. Τέλος, η πρόκληση των Big Data φαίνεται να διεισδύει μελλοντικά σε κρίσιμους τομείς, όπως η Τεχνολογική Έρευνα (55,1%), τα Κοινωνικά Δίκτυα (49%), οι Επιχειρήσεις (46,9%), η Οικονομία (26,5%) και η Υγεία (26,5%).

6. **Συνδυαστικά Αποτελέσματα Έρευνας:** Οι γυναίκες δεν πραγματοποιούν χρήση των Big Data (83,4%), δεν φαίνεται να δείχνουν διάθεση να τα υιοθετήσουν (83,3%) σε αντίθεση με τους άνδρες που τα χρησιμοποιούν κατά 15,8% και τα έχουν υιοθετήσει κατά 15% περισσότερο. Επίσης, προβλήματα που ταλανίζουν τον Δημόσιο τομέα αποτελούν οι υποδομές αποθήκευσης πληροφοριών (27,6%) και οι λανθασμένες – συγκεχυμένες πληροφορίες (20,7%). Ενώ ο ιδιωτικός τομέας ανησυχεί για τις υποδομές αποθήκευσης πληροφοριών (17,2%). Αναφορικά με τους νέους ρόλους οι γυναίκες: αναδεικνύουν την αντιμετώπιση των δυσκολιών των χρηστών κατά την αναζήτηση πληροφοριών (22,9%) και οι άνδρες την διευκόλυνση των χρηστών στη χρήση νέων πηγών πληροφόρησης & γνώσης και στην παροχή πληροφόρησης σ' αυτούς (54,2%). Ενώ ο δημόσιος και ο ιδιωτικός τομέας αναδεικνύουν τις υπηρεσίες εντοπισμού, επεξεργασίας - αναμόρφωσης πληροφοριών με 66,7% και 16,7% αντίστοιχα. Σχετικά με τις νέες ευκαιρίες οι γυναίκες θεωρούν σημαντική την απόκτηση εμπειρίας και τεχνογνωσίας σε Big Data κατά 23,4% και οι άνδρες την απόκτηση εμπειρίας (59,6%) εξειδίκευσης (57,5%) και τεχνογνωσίας σε Big Data (55,4%). Ενώ ο δημόσιος και ο ιδιωτικός τομέας αναδεικνύουν την απόκτηση εμπειρίας σε Big Data με 66%

και 16,8% αντίστοιχα. Αναφορικά με τις προκλήσεις, οι γυναίκες αναδεικνύουν το ανταγωνιστικό πλεονέκτημα (25,5%) και οι άνδρες την συνεργασία των επαγγελματιών της πληροφόρησης (60,9%). Ενώ στον δημόσιο τομέα σημασία αποκτούν οι ειδικές επαγγελματικές δεξιότητες (61,7%) και η συνεργασία μεταξύ των επαγγελματιών της πληροφόρησης (17,4%).

Συμπεράσματα

Η εξέλιξη στη διαχείριση μεγάλου όγκου δεδομένων κατέστησαν τα Big τεχνολογική πρόκληση της εποχής μας. Σκοπός της παρούσας εργασίας αποτέλεσε η διερεύνηση των Big Data, οι ρόλοι και οι ευκαιρίες για τον επαγγελματία της Πληροφόρησης και οι προκλήσεις για τις Ελληνικές Βιβλιοθήκες.

Η έρευνα που διεξήχθη κατέδειξε σημαντικά ευρήματα αναφορικά με τα Big Data. Το επίπεδο υιοθέτησης και χρήσης των Big Data από τις Ελληνικές Βιβλιοθήκες είναι χαμηλό μόλις στο 27,1% αν και διαφαίνεται μια αισιόδοξη στάση (31,9%) για την προοπτική υιοθέτησής τους. Προβλήματα προς διερεύνηση αναδείχθηκαν τα προβλήματα υποδομών αποθήκευσης πληροφοριών και η έλλειψη ταχύτητας στην άντληση πληροφοριών. Οι νέοι ρόλοι που έρχονται στο προσκήνιο για τους επαγγελματίες της Πληροφόρησης είναι σημαντικοί όπως η προσφορά υπηρεσιών εντοπισμού, επεξεργασίας και αναμόρφωσης πληροφοριών, η παροχή εξειδικευμένης πρόσβασης πελατών σε διευρυμένες πηγές πληροφόρησης, η βοήθεια σε χρήστες για απόκτηση ικανοτήτων πληροφοριακής παιδείας για εκτίμηση των πληροφοριών και η καθοδήγηση χρηστών στη διαδικασία αναζήτησης, ανάκτησης, ερμηνείας και χρήσης πληροφοριών. Οι αναδυόμενες ευκαιρίες για τους επαγγελματίες της Πληροφόρησης επίσης είναι πολλές όπως η απόκτηση εμπειρίας, εξειδίκευσης και τεχνογνωσίας σε θέματα Big Data. Αναδύονται προκλήσεις για τις Ελληνικές Βιβλιοθήκες από την υιοθέτηση των Big Data όπως η οργάνωση – άντληση – ανάλυση – σύνθεση – διαχείριση της πληροφορίας, η συνεργασία των επαγγελματιών, η δημιουργία υποδομών για μεταβολή των δεδομένων σε πληροφορίες και των πληροφοριών σε γνώση, οι επαγγελματικές δεξιότητες και το ανταγωνιστικό πλεονέκτημα. Η πρόκληση των Big Data φαίνεται να διεισδύει μελλοντικά σε κρίσιμους τομείς όπως η τεχνολογική έρευνα, τα κοινωνικά δίκτυα, οι επιχειρήσεις, η οικονομία και η υγεία.

Τα Big Data απασχολούν ήδη τον «κόσμο» της Πληροφορικής. Αν και η προσαρμογή σε νέες τεχνολογικές εξελίξεις προϋποθέτει ανθρώπινο δυναμικό με δεξιότητες και ένα χρονικό διάστημα προσαρμογής στο νέο περιβάλλον. Τα Big Data μέσα από την έγκαιρη αξιοποίηση της πληροφορίας θα προσδώσουν μια υψηλότερη «αξία» στον φορέα που θα τα υιοθετήσει με αποτέλεσμα να επιτύχει ένα

ανταγωνιστικό πλεονέκτημα. Τα Big Data αποτελούν τη νέα αξία για τον ψηφιακό κόσμο.

Βιβλιογραφικές Αναφορές

1. Abolfazli, S., Sanaei, Z., Ahmed, E., Gani, A., Buyya, R. (2014). "Cloud – based augmentation for mobile devices: motivation, taxonomies, and open challenges". *IEEE Communications Surveys & Tutorials*, 16(1), pp. 337–368.
2. Addo - Tenkorang , R., Helo, P. (2016). "Big data applications in operations/supply-chain management: A literature review", *Computers & Industrial Engineering*, pp. 1-16.
3. Agrawal, D., Bernstein, P., Bertino, E., Davidson, S., Dayal, U., Franklin, M., Gehrke, J., Haas, L., Halevy, A., Han J. (2012). "Challenges and opportunities with big data. A community white paper developed by leading researches across the United States", *Computing Community Consortium committee of the Computing Research Association*, pp. 1-17.
4. Agrawal, D., Bernstein, P., Bertino, E. (2011). "Challenges and Opportunities with Big Data 2011-1", *Cyber Center Technical Reports. Paper 1-19*, pp. 2032-2033.
5. Agrawal, D., Das, S., El Abbad, A. (2011). "Big data and cloud computing: current state and future opportunities". *Proceedings of the 14th International Conference on Extending Database Technology, Sweden: Uppsala*, pp. 530-533.
6. Allen, E. (2013). "Update on the Twitter Archive at the Library of Congress", [τεκμήριο *www*, URL: <http://blogs.loc.gov/loc/2013/01/update-on-the-twitter-archive-at-the-library-of-congress/>, ημερομηνία πρόσβασης: 19.11.2016].
7. Averbis, (2015). "Big Data Analytics in the German National Library", [τεκμήριο *www*, URL: <https://averbis.com/wp-content/uploads/2015/01/Averbis-CaseStudy-DNB-en.pdf>, ημερομηνία πρόσβασης: 02.11.2016].
8. Baker, J. (2015a). "The British Library Big Data Experiment", [τεκμήριο *www*, URL: <http://blogs.bl.uk/digital-scholarship/2014/06/the-british-library-big-data-experiment.html>, ημερομηνία πρόσβασης: 09.11.2016].
9. Baker, J., Terras, M., Mohamedally, D., Weyrich, T. (2015b). "The British Library Big Data Experiment: experimental interfaces, experimental teaching", *Sydney: Digital Humanities*.
10. Barrenetxea, G., Ingelrest, F., Schaefer, G., Vetterli, M., Couach, O., Parlange, M. (2008). "Sensorscope: out – of – the - box environmental monitoring". *Information processing in sensor networks, International Conference on IPSN'08. IEEE*, pp 332– 343.
11. Barroso, L., Holzle, U. (2009). "The datacenter as a computer: an introduction to the design of warehouse - scale machines". *Synt. Lect. Comput. Archit.*, 4(1), pp. 1-108.
12. Batini, C., Rula, A. (2015), "From data quality to Big Data quality", *Journal of Database Management*, 26(1), pp. 60-82.
13. Beaver, D., Kumar, S., Li, H., Sobel, J., Vajgel, P. (2010). "Finding a needle in haystack: facebook's photo storage". *OSDI*, 10(1), pp 1–8.
14. Berthold, M., Cebron, N., Dill, F., Gabriel, T., Kotter, T., Meinl, T., Ohl, P., Sieb, C., Thiel, K., Wiswedel, B. (2008). "Data Analysis, Machine Learning and Applications, KNIME: the Konstanz information miner. Springer, Data Analysis, Machine Learning and Applications, pp. 319-326.
15. Bhatotia, P., Wieder, A., Rodrigues, R., Acar, U., Pasquin, R. (2011). "Incoop: mapreduce for incremental computations". *Proceedings of the 2nd ACM Symposium on Cloud Computing. ACM*, pp. 7.

16. Bizer, C., Boncz, P., Brodie, M., Erling, O. (2012), "The meaningful use of big data: four perspectives - four challenges", SIGMOD Rec., 40(4), pp. 56-60.
17. Bughin, J., Chui, M., Manyika, J. (2010). "Clouds, big data, and smart assets: Ten tech-enabled business trends to watch". [τεκμήριο www, URL: <http://www.mckinsey.com/industries/high-tech/our-insights/clouds-big-data-and-smart-assets-ten-tech-enabled-business-trends-to-watch>, ημερομηνία πρόσβασης: 21.11.2016].
18. Buyya, R., Vecchiola, C., Selvi, T. (2013). "Chapter 3 – Virtualization", Mastering Cloud Computing, Technologies and Applications Programming, pp. 71-109.
19. Cai, L., Zhu, Y., (2015). "The Challenges of data quality and data quality assessment in the Big Data Era". Data Science Journal, 14(1), p.2.
20. Cattell, R. (2011). "Scalable sql and nosql data stores". ACM SIGMOD Record 39(4), pp. 12-27.
21. Chaiken, R., Jenkins, B., Larson, P., Ramsey, B., Shakib, D., Weaver, S., Zhou, J. (2008). "Scope: easy and efficient parallel processing of massive data sets". Proc. VLDB Endowment, 1(2), pp. 1265-1276.
22. Chang, F., Dean, J., Ghemawat, S. (2008). "Bigtable: A distributed storage for structured data". ACM Transactions on Computer Systems (TOCS), 26(2), pp. 1-14.
23. Chaudhuri, S., Dayal, U., Narasayya, V. (2011). "An overview of business intelligence technology". Commun ACM, 54(8), pp. 88-98.
24. Chen, M., Mao, S., Liu, Y. (2014). "Big Data: A survey". Mobile Netw. Appl., 19(1), pp. 171-209.
25. Chernatony, L., Harris, F. (2000). "Added value: its nature, roles and sustainability". European Journal of Marketing, 34(1/2), pp. 39-56.
26. Cline, N. (2000). "Virtual continuity: the challenge for research libraries today". Educause Review, 35(3), pp. 22-28.
27. Cooper, A. (2012). "What is analytics? Definition and essential characteristics". CETIS Analytics Series, 1(5), pp. 1-10.
28. Cows, J. (2015). "The Challenge of Big Data: A Day at the British Library", [τεκμήριο www, URL: <https://joshcows.com/2013/10/28/the-challenge-of-big-data-a-day-at-the-british-library/>, ημερομηνία πρόσβασης: 29.10.2016].
29. Cox, B., Jantti, M. (2012). "Discovering the impact of library use and student performance." EDUCAUSE Review Online [τεκμήριο www, URL: <http://www.educause.edu/ero/article/discovering-impact-library-use-and-student-performance>, ημερομηνία πρόσβασης: 20.11.2016].
30. Creamer, A., Martin, E., Kafel, D. (2014). "Research data management and the health sciences librarian". University of Massachusetts Medical School, Library Publications and Presentations, Paper 147, pp. 252-274.
31. Davis, K., Patterson, D. (2012). "Ethics of Big Data: Balancing risk and innovation". San Francisco: O'Reilly Media.
32. De Mauro, A., Greco, M., Grimaldi, M. (2015). "What is big data? A consensual definition and a review of key research topics". AIP Conference Proceedings, 1644, 97(1), pp. 7-35.
33. Dean, J., Ghemawat, S. (2008). "Mapreduce: Simplified data processing on large clusters". Commun. ACM, 51(1), pp. 107-113.
34. DeCandia, G., Hastorun, D., Jampani, M., et al. (2007). "Dynamo: amazon's highly available key - value store". SOSOP, 7(1), pp. 205-220.
35. Deutsche Nationalbibliothek (2015). "Linked data service of the german National Library", [τεκμήριο [www](http://www.dnb.de/en/linked-data), URL: <http://www.dnb.de/en/linked-data>]

- http://www.dnb.de/EN/Service/DigitaleDienste/LinkedData/linkedata_node.html, ημερομηνία πρόσβασης: 08.11.2016].
36. Dewitt, D., Gray, J. (1992). "Parallel database s: The future of high performance database s". *Commun. ACM*, 35(6), pp. 85–98.
 37. Dincturk, M., Jourdan, G., Bochmann, G. (2014). "A model - based approach for crawling rich internet applications". *ACM Transactions on the Web*, 8(3), pp. 1-39.
 38. Dwoskin, E. (2014), "In a single tweet, as many pieces of metadata as there are characters", [τεκμήριο www, URL: <http://blogs.wsj.com/digits/2014/06/06/in-a-single-tweet-as-many-pieces-of-metadata-as-there-are-characters/>, ημερομηνία πρόσβασης: 02.11.2016].
 39. Dzik, C. et al. (2015). "Big Data in banking". [τεκμήριο www, URL: <http://libeldoc.bsuir.by/bitstream/123456789/4938/1/BIGDATAinbanking.PDF>, ημερομηνία πρόσβασης: 07.11.2016].
 40. Emani, C., Cullot, N., Nicolle, C. (2015). "Understandable Big Data: A survey". *Computer Science Review*, 17(1), pp. 70-81.
 41. Erasmus, S. (2001). "Information literacy and distance education: the challenges of addressing the lack of (basic) information skills in a lifelong environment: a case study". *Mousaion*, 9(2), pp. 15-22.
 42. Farquhar, A., Baker, J. (2014). "Interoperable infrastructures for digital research: A proposed pathway for enabling transformation", Lausanne: Digital Humanities.
 43. Federer, L. (2016). "Research data management in the age of big data: Roles and opportunities for librarians". *Information Services & Use*, 36(1-2), pp. 35-43.
 44. Funaki, K., Hochin, T., Nomiya, H., Nakanishi, H. (2015). "Evaluation of ParallelIndexing Scheme for Big Data". *Applied Computing and InformationTechnology/2nd International Conference on Computational Science andIntelligence (ACIT-CSI)*, 3rd International Conference, IEEE, pp. 148-153.
 45. Gaffney, D., Puschmann, C. (2013). "Data collection on Twitter," In: Weller, K., Bruns, A., Burgess, J., Mahrt, M., Puschmann, C. (2013). "Twitter and society". New York: Peter Lang, pp. 55-68.
 46. Gani, A., Siddiq, A., Shamshirband, S., Hanum, F. (2016). "A survey on indexingtechniques for big data: taxonomy and performance evaluation". *Knowledge and Information s*, 46(2), pp. 241-284.
 47. Gantz, J., Reinsel, D. (2010). "The digital universe decade - are you ready". External publication of IDC (Analyse the Future) Information and Data, pp 1–16.
 48. Gantz, J., Reinsel, D. (2011). "Extracting value from chaos". *Proc. IDC iView*, pp. 1–12.
 49. Gerolimos, M., (2011). "Library education in Greece: history, current status and future prospects", *Library Review*, 60(2), pp.108 – 124.
 50. Ghemawat, S., Gobioff, H., Leung, S. (2003). "The google file ". *ACM roceedings of the nineteenth ACM symposium on Operating s principles*, 37(5), pp 29–43.
 51. Ginsberg, J. et al. (2009). "Detecting influenza epidemics using search engine query data". *Nature*, 457(7232), pp.1012-1014.
 52. Gubbi, J., Buyya, R. (2013). "Internet of Things (IoT): A vision, architectural elements, and future directions". *Future Generation Computer s*, 29(7), pp. 1645–1660.
 53. Hallam, G., Partridge, H. (2005). "Great expectations? Developing a profile of the 21st century library and information student: a Queensland University of Technology case study". Paper presented at IFLA Conference, pp. 1-12.

54. Han, J., Haihong, E., Le, G., Du, J. (2011). "Survey on NoSQL database. In Pervasive computing and applications (ICPCA), 6th International Conference, pp.363-366.
55. Hill, K. (2014). "You can hide your pregnancy online, but you'll feel like a
56. Howe, D., Costanzo, M., Fey, P. et al. (2008). "Big data: The future of biocuration". *Nature*, 455(7209), pp. 47-50.
57. Hoy, M. (2014). "Big Data: An introduction for librarians". *Medical Reference Services Quarterly*, 33(3), pp. 320-326.
58. Huser, V., Cimino, J. (2016). "Impending challenges for the use of Big Data". *International Journal of Radiation Oncology*, 95(3), pp. 890-894.
59. Huston, M. (2015). "Social Media and Big Data, Part 1: What is it?", *Social Data*, [τεκμήριο www, URL: <https://blogs.oracle.com/socialspotlight/social-media-and-big-data,-part-1:-what-is-it>, ημερομηνία πρόσβασης: 04.11.2016].
60. Huwe, T. (2014). "Big Data and the library: A natural fit". *Computers in Libraries*, 34(2), pp. 17-18.
61. Jacobs, A. (2009). "The pathologies of big data". *Commun. ACM*, 7(6), pp. 36-44 .
62. Javeau, C. (2010) "Η έρευνα με ερωτηματολόγιο", Εκδόσεις Τυπωθήτω. Αθήνα, σελ. 55.
63. Jinno, M., Takara, H., Kozicki, B. (2009). "Dynamic optical mesh networks: drivers, challenges and solutions for the future". *Optical Communication*, 35th European conference on ECOC'09. IEEE, pp. 1-4.
64. Kacprzyk, J., Zadrozny, S. (2001). "Computing with words in intelligent databasequerying: Standalone and Internet-based applications". *Information Sciences*, 134(1), pp. 71-109.
65. Khan, S., Shiraz, M., Abdul Wahab, A., Gani, A., Han, Q., BinAbdul Rahman, Z. (2014). "A comprehensive review on adaptabilityof network forensics frameworks for mobile cloud computing". *The Scientific World Journal*, pp. 1-24.
66. Kobielus, J., Marcus, B., Kobielus, J., et al. (2014). "Deploying Big Data analytics applications to the cloud: Roadmap for success", *Cloud Standards Customer Council*, pp. 1-20.
67. Koman, G., Kundrikova, J. (2016). "Application of Big Data technology in knowledge transfer process between business and academia". *Procedia Economics and Finance*, 39(1), pp. 605-611.
68. Kraska, T. (2013). "Finding the needle in the big data s haystack". *IEEE Intern. Comput.*, 17(1), pp. 84-86 .
69. Krishnan, K. (2013). "Data warehousing in the age of big data: A volume in MK series on business intelligence". Elsevier, pp. 3-27.
70. Labrinidis, A., Jagadish, H. (2012). "Challenges and opportunities with big data". *Proc VLDB Endowment*, 5(12), pp. 2032-2033.
71. Lazer, D., Kennedy, R., King, G., Vespignani, A. (2014). "Big Data. The Parable of Google Flu: Traps in Big Data Analysis". *Science*, 343(6176), pp. 1203-1205.
72. Lenzerini, M. (2002). "Data integration: a theoretical perspective". *Proceedings of the 31st ACM SIGMOD - SIGACT - SIGART symposium on principles of database s*. ACM, pp. 233-246.
73. Library of Congress, (2014). "Web archiving FAQs," [τεκμήριο www, URL: <http://www.loc.gov/webarchiving/faq.html>, ημερομηνία πρόσβασης: 15.11.2016].

74. Lougee, W., Choudhury, S., Gold, A. et al. (2007). "Association of research libraries, joint task force on library support for e-science, agenda for developing e-science in research libraries". Chicago, pp.130.
75. Maletic, J., Marcus, A. (2000). "Data cleansing: beyond integrity analysis". IQ Citeseer, pp. 200-209.
76. Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., Byers, A. (2011). "Big data: the next frontier for innovation, competition, and productivity". McKinsey Global Institute, pp. 1-143.
77. Mavrinac, M. (2010). "Transforming libraries: Big Data". Library Hazell McCallion Academic Learning Centre, 9(1), pp. 1-3.
78. Mayer – Schonberger, V., Cukier, K. (2013). "Big data: a revolution that will transform how we live, work, and think". New York: Eamon Dolan/Houghton Mifflin Harcourt.
79. McKusick, M., Quinlan, S. (2009). "Gfs: evolution on fastforward". ACM Queue 7(7), pp. 10.
80. Miller, E., Ogbuji, U., Mueller, V., MacDougall, K. (2012). "Bibliographic framework as a web of data: linked data model and supporting services". Washington: Library of Congress.
81. Monnappa, A. (2015). "How Facebook is using Big Data - The good, the bad, and the ugly [τεκμήριο www, URL: <https://www.simplilearn.com/how-facebook-is-using-big-data-article>, ημερομηνία πρόσβασης: 11.11.2016].
82. Nanopoulos, A., Manolopoulos, Y., Zakrzewicz, M., Morzy, T. (2002). "Indexing web access - logs for pattern queries". Proceedings of the 4th International workshop on web information and data management. ACM, pp 63–68.
83. Pinfield, S., Cox, A., Smith, J. (2014). "Research data management and libraries: Relationships, activities, drivers and influences", PLoS ONE, 9(12), pp. 1-28.
84. Ranger, C., Raghuraman, R., Penmetsa, A., Bradski, G., Kozyrakis, C. (2007). "Evaluating mapreduce for multi - core and multiprocessor s". Proceedings of the 2007 IEEE 13th International Symposium on High Performance Computer Architecture, HPCA '07, IEEE Computer Society, Washington, DC, USA, pp. 13-24.
85. Raymond, M. (2010a). "How tweet it is! Library acquires entire Twitter archive", [τεκμήριο www, URL: <http://blogs.loc.gov/loc/2010/04/how-tweet-it-is-library-acquires-entire-twitter-archive>, ημερομηνία πρόσβασης: 19.11.2016].
86. Reeve, A. (2013). "Managing data in motion: Data Integration best practice techniques and technologies". Elsevier: Morgan Kaufmann. pp.86.
87. Rimal, B., Choi, E., Lumb, I. (2009). "A taxonomy and survey of cloud computing s". 5th International Joint Conference on INC, IMS and IDC, pp. 44-51.
88. Sallam, R., Richardson, J., Hagerty, J., Hostmann, B. (2011). "Magic quadrant for business intelligence platforms". Stamford : CT, Gartner Group.
89. Selavo, L, Wood, A., Cao, Q., Sookoor, T., et al. (2007). "Luster: wireless sensor network for environmental research". Proceedings of the 5th International conference on Embedded networked sensor s. ACM, pp 103–116.
90. Sharma, I., Gupta, I., (2015). "Managing Big - Data over social media: A survey", International Journal of Advanced Research in Computer Science & Technology, 3(2), pp. 28-30.
91. Shvachko, K., Kuang, H., Radia, S., Chansler, R. (2010). "The Hadoop distributed file ". Proceedings of the 2010 IEEE 26th Symposium on Mass

- Storage s and Technologies (MSST), MSST '10, IEEE Computer Society, Washington, DC,USA, pp. 1–10.
92. Simmhan, Y., Aman, S., Kumbhare, A., et al. (2013). “Cloud - based software platform for Big Data analytics in smart grids”. *Comput. Sci. Eng.*, 15(4), pp. 38–47.
 93. Singh, J., Singla, V. (2015). “Big Data: Tools and technologies in Big Data”. *International Journal of Computer Applications*, 112(15), pp. 6–10.
 94. Soundararajan, S. (2016). “How social media companies use Big Data”, [τεκμήριο www, URL: <https://datafloq.com/read/how-social-media-companies-use-big-data/1957>, ημερομηνία πρόσβασης: 01.11.2016].
 95. Sun, Y., Chen, M., Liu, B., Mao, S. (2013). “Far: A fault - avoidant routing method for data center networks with regular topology”. *Proceedings of ACM/IEEE symposium on architectures for networking and communications s (ANCS'13)*. ACM.
 96. Supovitz, J., Weathers, J. (2004). “Dashboard lights: Monitoring implementation of district instructional reform strategies. Consortium for Policy Research in Education, USA: University of Pennsylvania, p.34.
 97. Teets, M., Goldner, M. (2013). “Libraries’ role in curating and exposing Big Data”. *Future Internet*, 5(1), pp. 429–438.
 98. Turkington, G. (2013). “Hadoop beginners guide”. Packt Publishing, Limited.
 99. Vitolo, C., Elkhatib, Y., Reusser, D., Macleod, C., Buytaert, W. (2015). “Web technologies for environmental Big Data”. *Environmental Modelling & Software*, 63(1), pp. 185–198.
 100. Wang, H., Xu, Z., Fujita, H., Liu, S. (2016). “Towards felicitous decision making: An overview on challenges and trends of Big Data”. *Information Sciences*, 367–368(1), pp. 747–765.
 101. Wayner, P. (2012). “7 top tools for taming big data”. [τεκμήριο www, URL: <http://www.networkworld.com/reviews/2012/041812-7-top-tools-for-taming-258398.html>, ημερομηνία πρόσβασης: 20.11.2016].
 102. Weller, K., Bruns, A., Burgess, J., Mahrt, M., Puschmann, C. (2013). “Twitter and society”. New York: Peter Lang, pp. 55–68.
 103. Yaqoob, I., Hashem, I., Gani, A., et al. (2016). “Big data: From beginning to future”. *International Journal of Information Management*, 36(6), pp. 1231–1247.
 104. Zaslavsky, A., Perera, C., Georgakopoulos, D. (2012). “Sensing as a service and Big Data”. *Proceedings of the International Conference on Advances in Cloud Computing*, Bangalore, India, pp. 1–8.
 105. Γναρδέλλης, Χ. (2006), “Ανάλυση δεδομένων με το SPSS 14.0 For Windows”, Εκδόσεις Παπαζήσης, Αθήνα. 2006; σελ. 378–380.
 106. Δικτυακή Εκπαιδευτική Πύλη ΥΠ.Π.Ε.Θ., (2016), “Συνδεδεμένα δεδομένα: Μια ευκαιρία για τις ελληνικές βιβλιοθήκες”, [τεκμήριο www, URL: http://www.e-yliko.gr/index.php?option=com_sppagebuilder&view=page&id=33&Itemid=129, ημερομηνία πρόσβασης: 14.11.2016].
 107. Κατσίλλης, Ι. (2006), “Επαγωγική στατιστική: εφαρμοσμένη στις κοινωνικές επιστήμες και την εκπαίδευση με έμφαση στην ανάλυση με υπολογιστές”, Εκδόσεις Gutenberg, Αθήνα, σελ. 95.
 108. Σπανός, Δ. (2015), “Συνδεδεμένα Δεδομένα: Μια ευκαιρία για τις ελληνικές Βιβλιοθήκες”, *Hellenic Academic Libraries Journal*, 1(1), σ. 14–22.

Παράρτημα

Φόρμα Ερωτηματολογίου

ΤΕΙ ΑΘΗΝΑΣ
Σχολή Διοίκησης & Οικονομίας
Τμήμα Βιβλιοθηκονομίας & Συστημάτων Πληροφόρησης



ΕΡΩΤΗΜΑΤΟΛΟΓΙΟ ΕΡΕΥΝΑΣ
ΓΙΑ ΕΚΠΟΝΗΣΗ ΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ

ΜΕ ΘΕΜΑ

«Big Data: Νέοι Ρόλοι Και Ευκαιρίες Για Τους Επαγγελματίες Πληροφόρησης»

Ο φοιτητής
Βαγγέλης Θάνος
email: noe.two.vthan@gmail.com
τ: 6949518930

Οι επιστημονικοί υπεύθυνοι
Κουλούρης Αλέξανδρος
email: akoul@teiath.gr

Κωνσταντίνος Κυπριανός
email: k.kyprianos@gmail.com

Αθήνα, 2016

Ερωτηματολόγιο Για Τους Επαγγελματίες Πληροφόρησης

Αγαπητοί Επαγγελματίες Της Πληροφόρησης,

Η έρευνα διεξάγεται στα πλαίσια της εκπόνησης της πτυχιακής μου εργασίας για το Τμήμα Βιβλιοθηκονομίας & Συστημάτων Πληροφόρησης της Σχολής Διοίκησης & Οικονομίας του ΤΕΙ Αθήνας στο οποίο φοιτώ.

Η εργασία μου τιτλοφορείται: «Big Data: Νέοι Ρόλοι Και Ευκαιρίες Για Τους Επαγγελματίες Πληροφόρησης».

Σκοπός της πτυχιακής μου εργασίας είναι η διερεύνηση του πεδίου των Big Data στις ελληνικές βιβλιοθήκες και η ανάδειξη των στάσεων των επαγγελματιών της πληροφόρησης σχετικά.

Το παρόν ερωτηματολόγιο θα παραμείνει αυστηρά ανώνυμο. Η συμμετοχή σας – μη υποχρεωτική – στην έρευνα θα είναι απόλυτα ανώνυμη ώστε να διασφαλιστεί το απόρρητο των προσωπικών σας δεδομένων.

*Σας ευχαριστώ πολύ εκ των προτέρων,
Θάνος Ευάγγελος*

ΕΝΟΤΗΤΑ 1. ΚΟΙΝΩΝΙΚΟ/ΔΗΜΟΓΡΑΦΙΚΑ ΣΤΟΙΧΕΙΑ

1. Φύλο

Γυναίκα	Άνδρας
---------	--------

2. Ηλικία

Σημειώστε:

3. Επίπεδο Μόρφωσης

Δημοτικό	Γυμνάσιο	Λύκειο	
Μεταλυκειακή Εκπαίδευση (IEK)	Τριτοβάθμια Εκπαίδευση	Μεταπτυχιακό Δίπλωμα	Διδακτορικό Δίπλωμα

4. Κλάδος της Πληροφόρησης που απασχολείστε:

Βιβλιοθηκονομία	Αρχειονομία	Μουσειολογία	Πληροφορική	Άλλο
-----------------	-------------	--------------	-------------	------

5. Φορέας που απασχολείστε:

Δημόσιος	Ιδιωτικός
----------	-----------

6. Τομέας Πληροφόρησης που απασχολείστε:

Επιχείρηση	Φορέας	Οργανισμός	Ίδρυμα	Άλλο
------------	--------	------------	--------	------

ΕΝΟΤΗΤΑ 2. BIG DATA

7. Ο φορέας σας χρησιμοποιεί Big Data;

Ναι

☐

Όχι

☐

8. Τι σημαίνουν για εσάς, ως επαγγελματίας της πληροφόρησης, τα Big Data;

Καινούρια Data (δεδομένα)	Πληροφόρηση σε πραγματικό χρόνο	Υψηλός όγκος data (δεδομένων)
Πολυσήμαντη διάσταση της πληροφόρησης	Είσοδος Data (δεδομένων) από την εκμετάλλευση νέων τεχνολογιών	Data (δεδομένα) που προκύπτουν από τα Social Media

9. Σε ποίο βαθμό θεωρείτε ότι έχουν υιοθετηθεί τα Big Data από την/τον επιχείρηση/φορέα/οργανισμό όπου και απασχολείστε;

Καθόλου

Λίγο

Μέτρια

Πολύ

Πάρα πολύ

☐
☐
☐
☐
☐

10. Σε ποίο βαθμό θεωρείτε ότι έχουν χρησιμοποιηθεί τα Big Data από την/τον επιχείρηση/φορέα/οργανισμό όπου και απασχολείστε;

Καθόλου

Λίγο

Μέτρια

Πολύ

Πάρα πολύ

☐
☐
☐
☐
☐

11. Σε ποίο βαθμό θεωρείτε ότι τα παρακάτω συστατικά στοιχεία του Μοντέλου 4V: Volume, Velocity, Variety και Veracity¹ χαρακτηρίζουν τα Big Data της/του επιχείρησης/φορέα/οργανισμού όπου και απασχολείστε;

Καθόλου

Λίγο

Μέτρια

Πολύ

Πάρα πολύ

1. Variety (Ποικιλία)

☐
☐
☐
☐
☐

2. Velocity (Ταχύτητα)

☐
☐
☐
☐
☐

3. Volume (Όγκος)

☐
☐
☐
☐
☐

4. Veracity (Ακρίβεια)

☐
☐
☐
☐
☐

12. Ποιες από τις παρακάτω μεθόδους συλλογής Big Data χρησιμοποιεί η/ο επιχείρηση/φορέας/οργανισμός όπου απασχολείστε;

Αρχεία καταγραφής (log files)	Δεδομένα αισθήσεων (sensory data)
Αρχεία και δεδομένα δικτύου (searching, indexing)	Δεν χρησιμοποιούνται τέτοιες τεχνολογίες

13. Ποιες από τις παρακάτω τεχνολογίες επεξεργασίας Big Data χρησιμοποιεί η/ο επιχείρηση/φορέας/οργανισμός όπου απασχολείστε;

¹ Variety (Ποικιλία): Υψηλό εύρος τύπων datas (δεδομένων) προς διαχείριση

Velocity (Ταχύτητα): Ρυθμός παραγωγής και επεξεργασίας datas (δεδομένων)

Volume (Όγκος): Ποσότητα – μέγεθος διαθέσιμων datas (δεδομένων)

Veracity (Ακρίβεια): Αξιοπιστία (Ειλικρίνεια) datas (δεδομένων)

NoSQL	Apache Hadoop
Skytree Server	Talend Open Studio
Pentaho	Splunk
MapReduce	Apache Spark
Δεν χρησιμοποιούνται τέτοιες τεχνολογίες	

14. Ποιες από τις παρακάτω μεθόδους επεξεργασίας Big Data χρησιμοποιεί η/ο επιχείρηση/φορέας/οργανισμός όπου απασχολείστε;

Hashing	Bloom Filter
Indexing	Parallel Computing
Δεν χρησιμοποιούνται τέτοιες τεχνολογίες	

15. Σε ποίο βαθμό θεωρείτε ότι τα παρακάτω προβλήματα ταλανίζουν τα Big Data της/του επιχείρησης/φορέα/οργανισμού όπου και απασχολείστε;

Καθόλου Λίγο Μέτρια Πολύ Πάρα πολύ

1. Έλλειψη ασφάλειας στοιχείων και προσωπικών δεδομένων	☐	☐	☐	☐	☐
2. Έλλειψη ταχύτητας στην άντληση πληροφοριών	☐	☐	☐	☐	☐
3. Λανθασμένες – συγκεχυμένες πληροφορίες	☐	☐	☐	☐	☐
4. Μη τήρηση απορρήτου πληροφοριών	☐	☐	☐	☐	☐
5. Προβλήματα υποδομών αποθήκευσης πληροφοριών	☐	☐	☐	☐	☐

ΕΝΟΤΗΤΑ 3. BIG DATA: ΡΟΛΟΙ – ΕΥΚΑΙΡΙΕΣ ΤΟΥ ΕΠΑΓΓΕΛΜΑΤΙΑ ΤΗΣ ΠΛΗΡΟΦΟΡΗΣΗΣ

16. Σε ποίο βαθμό θεωρείτε ότι ο επαγγελματίας της πληροφόρησης κατά την εμπλοκή του με τα Μεγάλα Δεδομένα (Big Data) διακρίνεται από τους παρακάτω νέους ρόλους;

Καθόλου Λίγο Μέτρια Πολύ Πάρα πολύ

Διαχειριστής Πληροφορίας Και Γνώσης

1. Δημιουργεί οργανωτικά συστήματα διαχείρισης πληροφοριών	☐	☐	☐	☐	☐
2. Αναπτύσσει αρχιτεκτονικές γνωστικών &	☐	☐	☐	☐	☐

πληροφοριακών συστημάτων					
3. Ασχολείται με ψηφιοποίηση υλικού	☐	☐	☐	☐	☐
4. Προσφέρει υπηρεσίες εντοπισμού, επεξεργασίας & αναμόρφωσης πληροφοριών	☐	☐	☐	☐	☐
Διαμεσολαβητής					
1. Αντιμετωπίζει τις δυσκολίες χρηστών κατά την αναζήτηση πληροφοριών	☐	☐	☐	☐	☐
2. Παρέχει εξειδικευμένη πρόσβαση πελατών σε διευρυμένες πηγές πληροφόρησης	☐	☐	☐	☐	☐
Διευκολυντής					
1. Διευκολύνει τη σύνδεση χρηστών με έντυπα και ηλεκτρονικά προϊόντα	☐	☐	☐	☐	☐
2. Διευκολύνει τους χρήστες στη χρήση νέων πηγών πληροφόρησης & γνώσης	☐	☐	☐	☐	☐
3. Οργανώνει τις πηγές πληροφόρησης & γνώσης	☐	☐	☐	☐	☐
	Καθόλου	Λίγο	Μέτρια	Πολύ	Πάρα πολύ
4. Παρέχει βοήθεια σε χρήστες για απόκτηση ικανοτήτων πληροφοριακής παιδείας για εκτίμηση των πληροφοριών	☐	☐	☐	☐	☐
5. Παρέχει πληροφόρηση σε χρήστες	☐	☐	☐	☐	☐
Εκπαιδευτής					
1. Καθοδηγεί χρήστες στη διαδικασία αναζήτησης, ανάκτησης, ερμηνείας & χρήσης πληροφοριών	☐	☐	☐	☐	☐
2. Παροχή εκπαιδευτικών προγραμμάτων για απομακρυσμένους χρήστες	☐	☐	☐	☐	☐
17. Σε ποιο βαθμό θεωρείτε ότι ο επαγγελματίας της πληροφόρησης κατά την εμπλοκή του με τα Big Data είναι δυνατό να επωφεληθεί από τις παρακάτω αναδυόμενες ευκαιρίες;					
	Καθόλου	Λίγο	Μέτρια	Πολύ	Πάρα πολύ
1. Απόκτηση εμπειρίας σε Big Data	☐	☐	☐	☐	☐
2. Απόκτηση εξειδίκευσης σε Big Data	☐	☐	☐	☐	☐
3. Απόκτηση τεχνογνωσίας σε Big Data	☐	☐	☐	☐	☐
4. Διευκόλυνση εργασίας: Επάρκεια λογισμικού, βάσης δεδομένων & πληροφοριών	☐	☐	☐	☐	☐
5. Εξασφάλιση ανταγωνιστικού πλεονεκτήματος	☐	☐	☐	☐	☐
6. Εξέλιξη (πχ. προαγωγή, αύξηση	☐	☐	☐	☐	☐

- οικονομικών αποδοχών)
- | | | | | | |
|--|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 7. Νέες θέσεις εργασίας σε βιβλιοθήκες | ☐ | ☐ | ☐ | ☐ | ☐ |
| 8. Νέες θέσεις εργασίας σε νέα πεδία απασχόλησης (πχ. διαχείριση γνώσης, ανάπτυξη δικτυακών πόρων) | ☐ | ☐ | ☐ | ☐ | ☐ |

ΕΝΟΤΗΤΑ 4. BIG DATA: ΠΡΟΚΛΗΣΕΙΣ ΓΙΑ ΤΙΣ ΕΛΛΗΝΙΚΕΣ ΒΙΒΛΙΟΘΗΚΕΣ

18. Σε ποió βαθμό θεωρείτε ότι οι Ελληνικές Βιβλιοθήκες διαθέτουν θετική άποψη για τη νέα πρόκληση των Big Data;

Καθόλου	Λίγο	Μέτρια	Πολύ	Πάρα πολύ
☐	☐	☐	☐	☐

19. Σε ποió βαθμό θεωρείτε ότι έχουν υιοθετηθεί στην πράξη τα Big Data από τις Ελληνικές Βιβλιοθήκες;

Καθόλου	Λίγο	Μέτρια	Πολύ	Πάρα πολύ
☐	☐	☐	☐	☐

20. Σε ποió βαθμό θεωρείτε ότι τα παρακάτω αποτελούν προκλήσεις των Big Data για τις Ελληνικές Βιβλιοθήκες;

	Καθόλου	Λίγο	Μέτρια	Πολύ	Πάρα πολύ
1. Διασφάλιση ιδιωτικότητας των δεδομένων	☐	☐	☐	☐	☐
2. Πληροφορία: <i>Επικαιροποίηση, οργάνωση ετερογένειας και ατελειών, υψηλή ταχύτητα άντλησης, ανάλυσης, σύνθεσης, διαχείρισης</i>	☐	☐	☐	☐	☐
3. Συνεργασία επαγγελματιών της πληροφόρησης	☐	☐	☐	☐	☐
4. Ανταγωνιστικό πλεονέκτημα στην Οικονομία της Γνώσης	☐	☐	☐	☐	☐
5. Δημιουργία υποδομών για μετασχηματισμό δεδομένων σε πληροφορίες και των πληροφοριών σε γνώση	☐	☐	☐	☐	☐
6. Ειδικές επαγγελματικές δεξιότητες βιβλιοθηκονόμων	☐	☐	☐	☐	☐

21. Σε ποιο βαθμό θεωρείτε ότι θα υιοθετηθούν μελλοντικά τα Big Data από τις Ελληνικές Βιβλιοθήκες;

Καθόλου

☐

Λίγο

☐

Μέτρια

☐

Πολύ

☐

Πάρα πολύ

☐

22. Κατά τη προσωπική σας άποψη σε ποιόν από τους παρακάτω κρίσιμους τομείς θα διεισδύσει μελλοντικά η νέα πρόκληση των Μεγάλων Δεδομένων (Big Data);

Επιχειρήσεις	Κοινωνικά Δίκτυα
Οικονομία	Τεχνολογική Έρευνα
Υγεία	Άλλο

ΕΥΧΑΡΙΣΤΩ ΠΟΛΥ ΓΙΑ ΤΗ ΣΥΜΜΕΤΟΧΗ ΣΑΣ