

Εθνικό και Καποδιστριακό
Πανεπιστήμιο Αθηνών



Τμήμα Επικοινωνίας και
Μέσων Μαζικής Ενημέρωσης

Πρόγραμμα Μεταπτυχιακών Σπουδών στην
Επικοινωνία και τα Μέσα Μαζικής Ενημέρωσης

Κατεύθυνση: Ψηφιακά Μέσα Επικοινωνίας
και Περιβάλλοντα Αλληλεπίδρασης

Οπτικοποίηση (μεγάλων) δεδομένων για την ανακάλυψη γνώσης

Χρυσούλα Νομικού (Α.Μ. 32026)

Διπλωματική εργασία που κατατίθεται
ως μέρος των απαιτήσεων του
Προγράμματος Μεταπτυχιακών Σπουδών στην
Επικοινωνία και τα Μέσα Μαζικής Ενημέρωσης

Επιβλέπων:
Επίκουρος Καθηγητής Δ. Χαρίτος

Αθήνα, Σεπτέμβριος 2014

Πίνακας περιεχομένων

Κατάλογος εικόνων	v
Κατάλογος πινάκων	ix
Περίληψη	x
Abstract	xi
Εισαγωγή	1
Οργάνωση κεφαλαίων	3
2 Μεγάλα Δεδομένα	5
2.1 Ορισμοί	5
2.2 Προέλευση ΜΔ	12
2.2.1 Κοινωνικά μέσα	14
2.2.2 Φορητές Συσκευές	18
2.2.3 Ποσοτικοποιημένος Εαυτός	19
2.2.4 Διαδίκτυο των Πραγμάτων	22
2.3 Διαχείριση ΜΔ	24
2.4 Η σκοτεινή πλευρά των ΜΔ	27
2.5 Ανάλυση ΜΔ	30
2.6 Νέα πεδία επαγγελματικής εξειδίκευσης	36
2.6.1 Επιστήμη δεδομένων	36
2.6.2 Δημοσιογραφία δεδομένων	38
3 Οπτικοποίηση δεδομένων: γενικό πλαίσιο	42
3.1 Εισαγωγικές επισημάνσεις	42
3.2 Ιστορική αναδρομή	43
3.2.1 Προ ηλεκτρονικών υπολογιστών	45
3.2.2 Μετά την εμφάνιση των ηλεκτρονικών υπολογιστών	56

3.3 Ορισμοί	58
3.4 Επιστήμη, τέχνη ή κάτι άλλο;	66
3.5 Οπτική Ανάλυση Δεδομένων	68
4 Η οπτικοποίηση δεδομένων ως δημιούργημα: σχεδιασμός από ανθρώπους για ανθρώπους	73
4.1 Εισαγωγικές επισημάνσεις	73
4.2 Οπτικός σχεδιασμός	75
4.2.1 Συστατικά οπτικοποίησης	77
4.2.2 Μέθοδοι οπτικοποίησης	84
4.2.3 Εστιάζοντας στο χρώμα	89
4.3 Εμπλέκοντας τον χρήστη	93
4.3.1 Αλληλεπίδραση	96
4.3.2 Αφηγηματικές τεχνικές	99
5 Μελέτες περίπτωσης	108
5.1 Επιλογή και εξέταση περιπτώσεων	109
5.1.1 How riot rumours spread on Twitter	111
5.1.2 Twitter Reverb - Mentions on Twitter of #Ferguson	117
5.1.3 Feltron 2013 Annual Report	124
5.1.4 Fitbit Dashboard	132
5.2 Συζήτηση	138
6 Τελικά συμπεράσματα	143
Βιβλιογραφικές παραπομπές	154
Απόδοση ξενόγλωσσων όρων	175
Παράρτημα	178

Κατάλογος εικόνων

Εικόνα 1: Οι φάσεις εξέλιξης του Διαδικτύου, όπως αποτυπώνονται σχηματικά από τους Perera et al. (2014: 416)	14
Εικόνα 2: Εκτύπωση οθόνης σελίδας προφίλ χρήστη Facebook το 2006 (Πηγή: Sun and Iyer, 2013)	17
Εικόνα 3: Η διαστρωματωμένη δομή ΜΔ (Πηγή: Davenport and Dyché, 2013: 10)	25
Εικόνα 4: Άντληση σοφίας από τα δεδομένα (Πηγή: Evans, 2011: 6)	32
Εικόνα 5: Ιεραρχική δομή νοητικής διεργασίας για παραγωγή σοφίας από τα δεδομένα, σύμφωνα με Ackoff (1989: 3 σε Rowley, 2007: 166)	33
Εικόνα 6: Ανακάλυψη Γνώσης από Βάσεις Δεδομένων (Πηγή: Tan et al., 2014: 3)	35
Εικόνα 7: Σχηματική απεικόνιση φάσεων εξέλιξης της οπτικοποίησης δεδομένων που εντοπίζει ο Friendly (2008a)	44
Εικόνα 8: Γράφημα ράβδων του Playfair που απεικονίζει τις εξαγωγές και τις εισαγωγές της Σκωτίας (Πηγή: Rogers, 2013a: 62)	46
Εικόνα 9: Ο πρώτος χωροπληθής χάρτης από τον Dupin (Πηγή: Meirelles, 2013: 117)	47
Εικόνα 10: Χάρτης του Guerry που απεικονίζει την εγκληματική δραστηριότητα (Πηγή: Meirelles, 2013: 143)	48
Εικόνα 11: Λεπτομέρεια από τον χάρτη του Snow που απεικονίζει τους θανάτους από χολέρα (Πηγή: Tufte, 1997: 31)	50
Εικόνα 12: Διαγράμματα της Nightingale που απεικονίζουν τις αιτίες θανάτου Βρετανών στρατιωτών (Πηγή: Rogers, 2013a: 67)	51
Εικόνα 13: Διάγραμμα του Minard που παρουσιάζει τις απώλειες του γαλλικού στρατού κατά την εκστρατεία στη Ρωσία (Πηγή: Tufte, 2001: 176)	53
Εικόνα 14: Διάγραμμα πεταλούδας του Maunder (Πηγή: Tufte, 1990: 22)	55

Εικόνα 15: Απόσπασμα από κανόνες που περιλαμβάνονται στην πρώτη αναφορά της JCSGP (Πηγή: Joint Committee on Standards for Graphic Presentation, 1915: 795)	55
Εικόνα 16: Διάγραμμα αψίδας που οπτικοποιεί τη μουσική σύνθεση <i>Für Elise</i> του Beethoven, από Wattenberg (Πηγή: Wattenberg, 2002: 4)	58
Εικόνα 17: Imageplot από το εγχείρημα " <i>Selfiecity</i> ", που συντονίζει ο Manovich (Πηγή: Selfiecity, 2014)	65
Εικόνα 18: Παράδειγμα καλλιτεχνικής οπτικοποίησης (Πηγή: Viégas and Wattenberg, 2007: 188)	68
Εικόνα 19: Η Αξία ως το 4ο "V" των ΜΔ για τους οργανισμούς και τους εργαζόμενους σε αυτούς, όπως αποτυπώνεται σχηματικά από Intel (Intel IT Center, 2013: 6)	70
Εικόνα 20: Απεικόνιση κατάταξης στοιχειωδών αντιληπτικών εργασιών βάσει ακριβείας, κατά την αρχική εκτίμηση των Cleveland και McGill (1984: 536)	80
Εικόνα 21: Οπτικές μεταβλητές (Πηγή: Bertin, 2011: 43)	82
Εικόνα 22: Στοιχειώδεις αντιληπτικές εργασίες (Πηγή: Cleveland and McGill, 1984: 532)	82
Εικόνα 23: Αντιληπτικές εργασίες ανά είδος μεταβλητής (Πηγή: Mackinlay, 1984: 125)	82
Εικόνα 24: Παράδειγμα Dashboard (Πηγή: Few, 2013: 7)	87
Εικόνα 25: Παράδειγμα πολλαπλών μικρογραφημάτων (Πηγή: Tufte, 2001: 169)	88
Εικόνα 26: Προσθετικό χρωματικό μοντέλο (Πηγή: Stone, 2003: 139)	91
Εικόνα 27: Σύγκριση χρωματικής γκάμας RGB και CMYK (Πηγή: Kadavy, 2011: 229)	92
Εικόνα 28: Επιφάνεια εργασίας <i>Tableau Desktop</i> - Λειτουργικότητα <i>Story Points</i>	102
Εικόνα 29: Τα είδη αφηγηματικής οπτικοποίησης, σύμφωνα με Segel και Heer (Πηγή: Segel and Heer, 2010: 1145)	103

Εικόνα 30: Αντιπαραβολή αφηγηματικών τεχνικών ιστοριών μυστηρίου με επιλεγμένες οπτικοποιήσεις δεδομένων (Πηγή: TapestryConference, 2014)	106
Εικόνα 31: Εκτύπωση οθόνης <i>Fizz</i> (Πηγή: Carden, 2011)	113
Εικόνα 32: Αρχική οθόνη " <i>How riot rumours spread on Twitter</i> " (Πηγή: Guardian Interactive Team et al., 2011)	114
Εικόνα 33: Σελίδα οπτικοποίησης " <i>Rioters attack London zoo and release animals</i> " (Πηγή: Guardian Interactive Team et al., 2011)	115
Εικόνα 34: Εμφάνιση υποκείμενων δεδομένων (Πηγή: Guardian Interactive Team et al., 2011)	116
Εικόνα 35: Πλήρης ανάπτυξη σελίδας " <i>Tweeting 'happy new year' around the world</i> " (Πηγή: Twitter, 2014)	118
Εικόνα 36: Πλήρης ανάπτυξη σελίδας άρθρου (Πηγή: Zak, 2014a)	120
Εικόνα 37: Οπτικοποίηση " <i>Mentions on Twitter of #Ferguson</i> " (Πηγή: Zak, 2014b)	121
Εικόνα 38: Εμφάνιση σημείωσης (Πηγή: Zak, 2014b)	122
Εικόνα 39: Επιλογή τιτιβίσματος (Πηγή: Zak, 2014b)	123
Εικόνα 40: Οπτικοποίηση <i>#IceBucketChallenge</i> μέσω <i>Twitter Reverb</i> (Πηγή: Twitter Reverb, 2014a)	123
Εικόνα 41: Οπτικοποίηση <i>#HurricaneArthur</i> κ.λπ. μέσω <i>Twitter Reverb</i> (Πηγή: Twitter Reverb, 2014b)	124
Εικόνα 42: Καταγραφή συζητήσεων από Felton σε 400 σελίδες (Πηγή: Feltron, 2014c)	125
Εικόνα 43: Ετήσιες αναφορές <i>Feltron</i> - Ευρετήριο (Πηγή: Feltron, 2014d)	126
Εικόνα 44: Ετήσια αναφορά <i>Feltron 2013</i> - Αρχική οθόνη (πλήρης ανάπτυξη) (Πηγή: Feltron, 2014b)	127
Εικόνα 45: Έναρξη πλοήγησης στην ετήσια αναφορά <i>Feltron 2013</i> (Πηγή: Feltron, 2014b)	128
Εικόνα 46: Σύνοψη καταγραφών του συνόλου των επικοινωνιών (Πηγή: Feltron, 2014b: 3)	129

Εικόνα 47: Καταμέτρηση εμφάνισης λέξεων, που έλαβε ως δέκτης ο Felton κατά τη διάρκεια των επικοινωνιών του (Πηγή: Feltron, 2014b: 8)	131
Εικόνα 48: Εξατομικευμένο <i>Fitbit Dashboard</i> - Πλήρης ανάπτυξη σελίδας (Πηγή: Fitbit, 2014b)	133
Εικόνα 49: Εξέλιξη προόδου (Πηγή: Fitbit, 2014b)	135
Εικόνα 50: Διαφορετικές εκδοχές μετρητή παρακολούθησης ισοζυγίου πρόσληψης-καύσης θερμίδων (Πηγή: Fitbit, 2014b)	136
Εικόνα 51: Διαφορετικές εκδοχές επίτευξης στόχου καύσης θερμίδων (Πηγή: Fitbit, 2014b)	137
Εικόνα 52: Αρχική οθόνη <i>Fitbit Dashboard</i> για <i>iPhone</i>	178
Εικόνα 53: Οθόνη παρακολούθησης εξέλιξης βάρους	178
Εικόνα 54: Οθόνη παρακολούθησης ύπνου	178

Κατάλογος πινάκων

Πίνακας 1: Παραδείγματα δομημένων και μη δομημένων δεδομένων, σύμφωνα με Hurwitz et al. (2013: 26-30)	7
Πίνακας 2: Σύγκριση κόστους αποθήκευσης, σύμφωνα με Davenport και Dyché (2013: 3)	26
Πίνακας 3: Ορισμοί και χαρακτηριστικά, σύμφωνα με Ackoff (1989 σε Rowley, 2007: 166)	34
Πίνακας 4: Πίνακας ορισμών από Card et al. (Πηγή: Card et al., 1999: 7)	59
Πίνακας 5: Σύγκριση ειδών οπτικοποίησης από Gershon και Eick (Πηγή: Gershon and Eick, 1997: 29)	63
Πίνακας 6: Καταγραφή μεθόδων οπτικοποίησης κατά Kirk (2012: 120-158)	86
Πίνακας 7: Επιλεγμένες μελέτες περίπτωσης	110

Περίληψη

Το πολυδιάστατο φαινόμενο "*Μεγάλα Δεδομένα*" (ΜΔ - Big Data) διαμορφώνεται από τις τεχνολογικές δυνατότητες που επιτρέπουν τη συλλογή, την αποθήκευση, την επεξεργασία και την ανάλυση δεδομένων που χαρακτηρίζονται από πρωτοφανή όγκο, υψηλή ταχύτητα παραγωγής και ανανέωσης ή/και ποικιλομορφία. Ωστόσο, η αξία που εμπεριέχεται στα ΜΔ εξαρτάται εν πολλοίς από τις δυνατότητες ανάλυσης και άντλησης πληροφορίας από αυτά. Η αναγνώριση της οπτικοποίησης ως μέσου άντλησης αξίας από τα ΜΔ, συνδέεται με τον σημαντικό ρόλο που διαδραματίζει στη διαδικασία μετατροπής των ανεπεξέργαστων δεδομένων σε χρήσιμη πληροφορία.

Στο πλαίσιο της παρούσας εργασίας, η οπτικοποίηση εξετάζεται ως μέσο ανακάλυψης πληροφοριών, νοημάτων και γνώσης από τα (μεγάλα) δεδομένα, συνεκτιμώντας παράλληλα την επίδραση των νέων τεχνολογιών στην αύξηση των διαθέσιμων δεδομένων και στην ενίσχυση των δυνατοτήτων παραγωγής οπτικοποιήσεων. Στόχος είναι η εποπτική θεώρηση της διαδικασίας σχεδίασης μίας οπτικοποίησης (μεγάλων) δεδομένων, για την κατανόηση των απαραίτητων σχεδιαστικών αποφάσεων που λαμβάνονται στο πλαίσιο αυτής.

Μέσω της βιβλιογραφικής έρευνας και της εξέτασης των μελετών περίπτωσης επιχειρείται η κατανόηση βασικών πτυχών του φαινομένου ΜΔ, η αναγνώριση της οπτικοποίησης ως μέσο άντλησης γνώσης από τα δεδομένα και η μελέτη βασικών αρχών και πρακτικών στην οπτικοποίηση δεδομένων. Έμφαση δίδεται στη συνολική σχεδιαστική διαδικασία, για τη διαμόρφωση ενός γενικού θεωρητικού υποβάθρου ως αφετηρία για περαιτέρω καλλιέργεια των απαραίτητων σχετικών δεξιοτήτων.

Λέξεις-κλειδιά: Οπτικοποίηση δεδομένων, Μεγάλα Δεδομένα, Big Data, Οπτικός Σχεδιασμός, Επικοινωνία Ανθρώπου - Υπολογιστή

Abstract

Big Data is a multifaceted phenomenon, defined by the technological advances that enable collecting, storing, processing and analyzing data, which are characterized by unprecedented volume, velocity and/or variety. However, the value that lies within Big Data largely depends on analysis and information extraction capabilities. Visualization is widely recognized as a means of extracting value from Big Data, since it plays a key role in the process of transforming raw data into useful information.

In the scope of the current thesis, visualization is examined as a means of discovering information, meaning and knowledge from (big) data, while taking into account the increase in available data and the proliferation of visualization methods and tools. The aim is to provide a high-level view of the process of designing a (big) data visualization, in order to grasp fundamental design decisions that are made throughout the creative process.

Literature review and examination of a selection of case studies are employed with the intention of understanding essential aspects of the Big Data phenomenon, acknowledging visualization as means of extracting knowledge from data, as well as studying basic data visualization principles and practices. Focus is on the overall design process, in order to provide some theoretical background as a starting point for further developing the necessary skills.

Keywords: Data visualization, Big Data, Visual Design, Human - Computer Interaction

Εισαγωγή

Τα τελευταία χρόνια παρατηρούνται ραγδαίες τεχνολογικές εξελίξεις σε επίπεδο υποδομών, υλισμικού και λογισμικού που έχουν καταστήσει εφικτή την καταγραφή, τη συλλογή, την αποθήκευση και τη διακίνηση μεγάλου όγκου δεδομένων. Μέσω του Διαδικτύου, των κινητών συσκευών, των κοινωνικών μέσων (social media) και του Διαδικτύου των Πραγμάτων (Internet of Things) συγκεντρώνονται δεδομένα, κατά κύριο λόγο μη δομημένα, τα οποία συχνά μπορούν να αξιοποιηθούν πέρα από τους σκοπούς για τους οποίους παράγονται και να αποκαλύψουν -είτε μεμονωμένα, είτε σε συνδυασμό με άλλα- σχέσεις ή/και γνώση πέρα από τα προφανή. Τα δεδομένα αυτά, τα οποία έχει καθιερωθεί να αποκαλούνται "*Big Data*"¹, προέρχονται από πηγές που είτε δεν υπήρχαν παλαιότερα, είτε η συλλογή τους ήταν τεχνικώς δυσχερής, αδύνατη ή εξαιρετικά κοστοβόρα. Χαρακτηρίζονται δε από πρωτοφανή όγκο, υψηλή ταχύτητα παραγωγής και ανανέωσης και ποικιλομορφία.

Η μετατροπή ανεπεξέργαστων δεδομένων σε χρήσιμη πληροφορία περιλαμβάνει στάδια όπως η συλλογή των δεδομένων, η αρχική επεξεργασία τους, η εξόρυξη δεδομένων (data mining) και η τελική επεξεργασία τους. Κεντρικό ρόλο στις δύο τελευταίες διεργασίες διαδραματίζει η οπτικοποίηση δεδομένων: η οπτική αναπαράσταση δεδομένων αξιοποιεί τις δυνατότητες του ανθρώπινου οπτικού συστήματος, για να διευκολύνει τον άμεσο εντοπισμό σχέσεων και μοτίβων σε αφηρημένα δεδομένα. Ως αποτέλεσμα, είναι δυνατή η γρήγορη ερμηνεία των δεδομένων που αναπαρίστανται και η άντληση γνώσης από αυτά.

Βάσει των προαναφερθέντων, καθίσταται αντιληπτή η σύνδεση της οπτικοποίησης με τα μεγάλα σύνολα δεδομένων. Το θετικό και αισιόδοξο κλίμα διεθνώς για τα ΜΔ πηγάζει από τον δυνητικό πλούτο πληροφορίας που

¹ Στη συνέχεια θα αναφέρονται ως "*Μεγάλα Δεδομένα*" (για λόγους συντομίας ΜΔ).

εμπεριέχουν. Ωστόσο, η αξία τους εξαρτάται εν πολλοίς από τις δυνατότητες ανάλυσής τους και εξαγωγής συμπερασμάτων από αυτά. Η αναγνώριση της οπτικοποίησης ως μέσου άντλησης αξίας από τα μεγάλα σύνολα δεδομένων, συνδέεται με τον σημαντικό ρόλο που διαδραματίζει στη διαδικασία μετατροπής των ανεπεξέργαστων δεδομένων σε χρήσιμη πληροφορία. Ένδειξη της αυξημένης σημασίας που αποδίδεται στην οπτικοποίηση για την αξιοποίηση της δυναμικής των (μεγάλων) δεδομένων αποτελεί το γεγονός ότι κορυφαίες εταιρείες που αναπτύσσουν εφαρμογές επιχειρηματικής ευφυΐας, προσφέρουν λειτουργικότητες οπτικοποίησης για την ανάλυση δεδομένων και την εξαγωγή συμπερασμάτων από αυτά.

Παράλληλα, τα τελευταία χρόνια παρατηρείται έντονο ερευνητικό ενδιαφέρον για την οπτικοποίηση δεδομένων, ωστόσο δεν πρόκειται για πρωτοεμφανιζόμενο κλάδο. Η σύγχρονη περίοδος της οπτικοποίησης καθορίζεται από τις εξελίξεις στους ηλεκτρονικούς υπολογιστές και τα γραφικά υπολογιστών. Εντούτοις, αποτελεί την πλέον πρόσφατη φάση εξέλιξης μίας μακροχρόνιας παράδοσης στην κατασκευή οπτικών αναπαραστάσεων δεδομένων.

Με βάση τα προαναφερθέντα, στο πλαίσιο της παρούσας εργασίας, η οπτικοποίηση εξετάζεται ως μέσο ανακάλυψης πληροφοριών, νοημάτων και γνώσης από τα (μεγάλα) δεδομένα, συνεκτιμώντας παράλληλα την επίδραση των νέων τεχνολογιών στην αύξηση των διαθέσιμων δεδομένων και στην ενίσχυση των δυνατοτήτων παραγωγής οπτικοποιήσεων. Στόχος είναι η εποπτική θεώρηση της διαδικασίας σχεδίασης μίας οπτικοποίησης (μεγάλων) δεδομένων, για την κατανόηση των απαραίτητων σχεδιαστικών αποφάσεων που λαμβάνονται στο πλαίσιο αυτής.

Κατ' αρχάς, καταγράφονται οι τεχνολογικές εξελίξεις που καθιστούν εφικτή τη συλλογή και ανάλυση τεράστιου όγκου δεδομένων. Στη συνέχεια, μελετώνται βασικές αρχές και πρακτικές στην οπτικοποίηση δεδομένων, με έμφαση στις σχεδιαστικές αποφάσεις που λαμβάνονται στο πλαίσιο δημιουργίας

οπτικοποιήσεων από ανθρώπους για ανθρώπους, με σκοπό την ανακάλυψη γνώσης. Τέλος, εξετάζονται επιλεγμένες μελέτες περίπτωσης παρουσίασης δεδομένων που έχουν συλλεχθεί από πηγές προέλευσης ΜΔ, που έχουν οπτικοποιηθεί σε πραγματικό και σε ύστερο χρόνο, για τον εντοπισμό των βασικών σχεδιαστικών επιλογών και των επιπλοκών τους.

Μέσω της βιβλιογραφικής έρευνας και της εξέτασης των μελετών περίπτωσης επιχειρείται η κατανόηση βασικών πτυχών του φαινομένου ΜΔ, η σύνδεση της οπτικοποίησης δεδομένων ως μέσο άντλησης γνώσης από αυτά και η μελέτη βασικών αρχών και πρακτικών στην οπτικοποίηση δεδομένων. Δεδομένου του ενδιαφέροντος της γράφουσας για μελλοντική ενασχόλησή της με την κατασκευή οπτικοποιήσεων δεδομένων, έμφαση δίδεται στη συνολική σχεδιαστική διαδικασία, για τη διαμόρφωση ενός γενικού θεωρητικού υποβάθρου ως αφετηρία για περαιτέρω καλλιέργεια των απαραίτητων σχετικών δεξιοτήτων.

Οργάνωση κεφαλαίων

Στο **Κεφάλαιο 2** καταγράφονται ορισμοί και περιγράφονται οι κυριότερες πηγές προέλευσης ΜΔ που σχετίζονται με το Διαδίκτυο. Επιπρόσθετα, τίγονται ζητήματα αναφορικά με τη διαχείριση και την ανάλυση Μεγάλων Δεδομένων και επισημαίνονται προβληματισμοί που έχουν διατυπωθεί στο πλαίσιο της κριτικής επί αυτών. Τέλος, αναφέρονται τα νέα πεδία επαγγελματικής εξειδίκευσης που έχουν αναδειχθεί και απαιτούν δεξιότητες ανάλυσης ΜΔ και παρουσίασης της πληροφορίας που μπορεί να αντληθεί από αυτά.

Στο **Κεφάλαιο 3** περιλαμβάνεται ιστορική αναδρομή για τον εντοπισμό σημαντικών γεγονότων που σχετίζονται με την οπτική αναπαράσταση δεδομένων. Επίσης, καταγράφονται βασικοί ορισμοί και επιλεγμένες οπτικές που αναδεικνύουν ενδιαφέρουσες πτυχές της οπτικοποίησης, βάσει της σχετικής βιβλιογραφίας. Τέλος, μελετάται η Οπτική Ανάλυση Δεδομένων, ως λύση που αξιοποιεί την οπτικοποίηση δεδομένων για τη διερεύνηση και την άντληση

γνώσης από ΜΔ.

Στο **Κεφάλαιο 4** παρουσιάζονται ζητήματα που αφορούν στις σχεδιαστικές αποφάσεις που λαμβάνονται στο πλαίσιο του οπτικού σχεδιασμού μίας οπτικοποίησης δεδομένων. Επιπλέον, εντοπίζονται τρόποι με τους οποίους μπορούν να εμπλουτιστούν οπτικοποιήσεις που έχουν κατασκευαστεί ως διεπαφές παρουσίασης δεδομένων με σκοπό την ανακάλυψη γνώσης, προκειμένου να παρέχουν το κατάλληλο ερμηνευτικό πλαίσιο στον χρήστη και να ενθαρρύνουν τον διάλογό του με τα υποκείμενα δεδομένα.

Στο **Κεφάλαιο 5** εξετάζονται επιλεγμένες οπτικοποιήσεις δεδομένων, που προέρχονται από πηγές παραγωγής ΜΔ, ως προς το τελικό αποτέλεσμα που προετοιμάζει ένας δημιουργός-άνθρωπος για να προσφέρει τη δυνατότητα ανακάλυψης γνώσης σε έναν χρήστη-άνθρωπο. Στόχος της μελέτης είναι η εποπτική θεώρηση της διαδικασίας σχεδίασης μίας οπτικοποίησης δεδομένων που έχουν συλλεχθεί από πηγές προέλευσης ΜΔ, για τον εντοπισμό των βασικών σχεδιαστικών επιλογών και των επιπλοκών τους.

Στο **Κεφάλαιο 6** συζητούνται τα συμπεράσματα από τη βιβλιογραφική έρευνα (Κεφάλαια 2-4) και την εξέταση των επιλεγμένων μελετών περίπτωσης (Κεφάλαιο 5).

2 Μεγάλα Δεδομένα

Στο παρόν κεφάλαιο εξετάζονται βασικές πτυχές του φαινομένου "Μεγάλα Δεδομένα" (Big Data)².

Κατ' αρχάς, καταγράφονται ορισμοί και περιγράφονται οι κυριότερες πηγές προέλευσης Μεγάλων Δεδομένων (εφεξής ΜΔ) που σχετίζονται με το Διαδίκτυο. Στη συνέχεια, τίγονται ζητήματα αναφορικά με τη διαχείριση και την ανάλυση ΜΔ και επισημαίνονται προβληματισμοί που έχουν διατυπωθεί στο πλαίσιο της κριτικής επί αυτών. Τέλος, αναφέρονται τα νέα πεδία επαγγελματικής εξειδίκευσης που έχουν αναδειχθεί και απαιτούν δεξιότητες ανάλυσης ΜΔ και παρουσίασης της πληροφορίας που μπορεί να αντληθεί από αυτά.

2.1 Ορισμοί

Η δυνατότητα συλλογής τεράστιου όγκου ετερογενών και συχνά ανανεωνόμενων δεδομένων, σε συνδυασμό με το πολύ χαμηλό κόστος

² Εν συντομία και βάσει όσων εξετάζονται στο παρόν κεφάλαιο, ο γενικός όρος "Μεγάλα Δεδομένα" περιγράφει ένα πολυδιάστατο φαινόμενο που διαμορφώνεται από τις τεχνολογικές δυνατότητες που επιτρέπουν τη συλλογή, την αποθήκευση, την επεξεργασία και την ανάλυση δεδομένων, τα οποία χαρακτηρίζονται από πρωτοφανή όγκο, υψηλή ταχύτητα παραγωγής και ανανέωσης ή/και ποικιλομορφία.

Το 2013 και ως αποτέλεσμα της αυξανόμενης χρήσης του, ο όρος εισήχθη για πρώτη φορά στο Λεξικό Αγγλικής Γλώσσας της Οξφόρδης (Oxford Dictionaries, 2013). Σύμφωνα με τον σχετικό ορισμό, ως Μεγάλα Δεδομένα ορίζονται γενικώς τα "δεδομένα [σύνολα δεδομένων] μεγάλου μεγέθους, κατά κανόνα σε τέτοιο βαθμό που ο χειρισμός και η διαχείρισή τους θέτουν σημαντικές υλικοτεχνικές προκλήσεις· επίσης, ο κλάδος της υπολογιστικής που ασχολείται με αυτού του είδους δεδομένα" (OED Online, 2014 και Press, 2013). Σήμερα, η αναθεωρημένη εκδοχή του εν λόγω ορισμού περιγράφει τα Μεγάλα Δεδομένα ως "σύνολα δεδομένων εξαιρετικά μεγάλου μεγέθους που μπορούν να αναλυθούν υπολογιστικά για να αποκαλύψουν μοτίβα, τάσεις και συσχετίσεων, ειδικά όσων σχετίζονται με ανθρώπινες συμπεριφορές και αλληλεπιδράσεις" (Oxford Dictionaries, 2014).

αποθήκευσής τους κατέστησαν εφικτή τη συγκέντρωση ΜΔ. Ο Laney³ (2001) είχε εντοπίσει την τάση αύξησης των δεδομένων ειδικά στον τομέα του ηλεκτρονικού εμπορίου, που έθετε *"νέες προκλήσεις στη διαχείριση των δεδομένων σε τρεις διαστάσεις"*. Οι τρεις αυτές διαστάσεις, οι οποίες έχει καθιερωθεί να αναφέρονται ως τα *"3Vs"* που χαρακτηρίζουν τα ΜΔ, είναι ο όγκος (volume), η ταχύτητα (velocity) και η ποικιλία (variety).

Κατ' αρχάς, ο όγκος περιγράφει την αύξηση των διαθέσιμων συλλεγόμενων δεδομένων που προέρχονται από παραδοσιακές και από νέες πηγές προέλευσης, και θέτει προκλήσεις τόσο στην αποθήκευση, όσο και στην ανάλυση αυτών (Gartner, 2011). Η ταχύτητα αναφέρεται τόσο στην ταχύτητα παραγωγής δεδομένων, όσο και στην αναγκαιότητα γρήγορης επεξεργασίας των δεδομένων (Gartner, 2011), ακόμη και σε πραγματικό χρόνο. Τέλος, η ποικιλία συνίσταται στους διαφορετικούς τύπους των δεδομένων προς ανάλυση και στις πηγές προέλευσής τους: αυτά είτε προέρχονται από την αλληλεπίδραση ανθρώπων με μηχανές, είτε παράγονται απευθείας από μηχανές χωρίς να παρεμβάλλεται ανθρώπινη παρέμβαση (Hurwitz et al., 2013: 26, 29).

Οι δύο κύριοι τύποι δεδομένων είναι τα δομημένα και τα μη δομημένα δεδομένα, τα οποία διαφοροποιούνται βάσει της ύπαρξης ή μη συγκεκριμένης μορφής (format). Τα δομημένα δεδομένα (structured data) έχουν συγκεκριμένο μήκος και μορφή -παραδείγματα περιλαμβάνουν αριθμούς, ημερομηνίες και ακολουθίες χαρακτήρων (Hurwitz et al., 2013: 26). Αντιθέτως, τα μη δομημένα δεδομένα (unstructured data), που αποτελούν τον κύριο όγκο πληροφορίας που διαθέτει ένας οργανισμός, δεν χαρακτηρίζονται από συγκεκριμένη μορφή και μέχρι πρότινος οι διαθέσιμες τεχνολογικές λύσεις περιορίζονταν στην αποθήκευση και στη χειροκίνητη ανάλυσή τους (Hurwitz et al., 2013: 29). Αξίζει μάλιστα να σημειωθεί ότι, σύμφωνα με τη SAS (2013: 1), εκτιμάται ότι τα δομημένα (αριθμητικά) δεδομένα συνιστούν μόλις το 10% της διαθέσιμης

³ Αναλυτής της META Group, η οποία σήμερα αποτελεί μέρος της συμβουλευτικής εταιρείας Gartner.

πληροφορίας που διαθέτει έναν οργανισμός, ενώ το υπόλοιπο 90% αφορά σε μη δομημένα δεδομένα, όπως ελεύθερο κείμενο και αρχεία πολυμέσων (φωτογραφίες, βίντεο ή ήχος).

Στον ακόλουθο πίνακα (Πίνακας 1) καταγράφονται παραδείγματα ανά τύπο δεδομένων και φορέα παραγωγής:

Πίνακας 1: Παραδείγματα δομημένων και μη δομημένων δεδομένων, σύμφωνα με Hurwitz et al. (2013: 26-30)

	Δομημένα δεδομένα	Μη δομημένα δεδομένα
Παραγόμενα από μηχανή	• καταγραφές από αισθητήρες • δεδομένα καταγραφών Ιστού⁴ • συναλλαγές στα σημεία πώλησης • χρηματιστηριακά δεδομένα	• δορυφορικές εικόνες • επιστημονικά δεδομένα • φωτογραφίες & βίντεο • δεδομένα από ραντάρ
Παραγόμενα από άνθρωπο	• δεδομένα εισόδου • click-stream data • gaming-related data	• κείμενα • δεδομένα από κοινωνικά μέσα • δεδομένα από κινητές συσκευές • περιεχόμενο ιστοσελίδων

Στο σημείο αυτό θα πρέπει να σημειωθεί μία επιπλέον διάσταση, η απουσία της οποίας υπονομεύει την αξία που μπορεί να προκύψει μετά την επεξεργασία των δεδομένων: πρόκειται για την αξιοπιστία (veracity) των δεδομένων που συνδέεται με την καθαρότητα της εξαγόμενης πληροφορίας (Ohlhorst, 2013: 3).

Οι πρώτοι ορισμοί των ΜΔ έδιναν έμφαση στον πρωτοφανή όγκο των παραγόμενων δεδομένων (Mayer-Schönberger and Cukier, 2013: 6). Ωστόσο, θα μπορούσε να υποστηριχθεί ότι το όριο πέραν του οποίου τα συλλεγόμενα δεδομένα χαρακτηρίζονται ως μεγάλα δεν είναι αντικειμενικό, καθώς εξαρτάται

⁴ Απόδοση του αγγλικού όρου "web log".

και συνεχώς επαναπροσδιορίζεται τόσο από τις τεχνολογικές δυνατότητες συλλογής και αποθήκευσης κάθε εποχής, όσο και από τις εγγενείς δυνατότητες ενός οργανισμού να παράγει και να διαχειρίζεται δεδομένα (Manyika et al., 2011). Άλλωστε, η επικέντρωση της προσοχής στη συγκέντρωση τεράστιου όγκου δεδομένων υποβαθμίζει τα υπόλοιπα χαρακτηριστικά που διαφοροποιούν τα συλλεγόμενα δεδομένα σε σχέση με αυτά που ήταν παλαιότερα διαθέσιμα: την ποικιλομορφία και την υψηλή ταχύτητα ανανέωσης.

Όπως παραστατικά περιγράφει ο Davenport (2014: 1), τα ΜΔ διαφέρουν από ό,τι ήταν σύνηθες μέχρι σήμερα καθώς *"είναι τόσο μεγάλα για να χωρέσουν σε έναν μόνο εξυπηρετητή, τόσο αδόμητα για να χωρέσουν σε μία παραδοσιακή βάση δεδομένων ή με τόσο συνεχή ροή για να χωρέσουν σε μία στατική αποθήκη δεδομένων"*. Επισημαίνεται η διαζευκτική διατύπωση που χρησιμοποιεί, η οποία αιτιολογείται από τις επιφυλάξεις που εκφράζει στη συνέχεια για την ακρίβεια του όρου *"big data"*: υπογραμμίζει ότι είναι μάλλον αδόκιμος αφού ο όγκος δεν είναι το διαφοροποιό ούτε το σημαντικότερο στοιχείο, ενώ η προϋπόθεση συνύπαρξης των 3Vs παραβλέπει περιπτώσεις όπου ενδέχεται να υπάρχουν μόνο δύο από τα τρία (Davenport, 2014: 6-7). Σημειώνεται ότι η πολυπλοκότητα (complexity) και η μεταβλητότητα (variability) έχουν προταθεί ως επιπλέον ιδιότητες που χαρακτηρίζουν τα ΜΔ (SAS, 2013: 3), ωστόσο κρίνεται ότι δεν συνιστούν χαρακτηριστικά αλλά μάλλον επιπτώσεις της ποικιλίας και της ταχύτητας αντιστοίχως, που μεγεθύνονται εξαιτίας του μεγάλου όγκου των συλλεγόμενων δεδομένων.

Όσο ο ενθουσιασμός για τον όγκο των συλλεγόμενων δεδομένων καταλαγιάζει, τα δεδομένα συνοδεύονται από νέους επιθετικούς προσδιορισμούς που εστιάζουν σε διαφορετικές όψεις των ΜΔ. Ο συνδυασμός ΜΔ και των δυνατοτήτων επεξεργασίας και ανάλυσής τους σε πραγματικό χρόνο συνοψίζεται στον όρο *"fresh data"*, τα οποία διευκολύνουν τους οργανισμούς στην εξαγωγή άμεσων συμπερασμάτων και στην αναβάθμιση της εμπειρίας του χρήστη, μέσω της παροχής υπηρεσιών προσαρμοσμένων στην τρέχουσα θέση και δραστηριότητά του (Ericsson, 2014).

Τα ίχνη που αφήνουν μεμονωμένοι καταναλωτές κατά την καθημερινή χρήση υπηρεσιών, όπως ηλεκτρονικές συναλλαγές ή αλληλεπιδράσεις στα κοινωνικά μέσα, συνιστούν τα *"small data"*, η επεξεργασία των οποίων είναι δυνατόν να αποφέρει οφέλη, εκτός από τους οργανισμούς που τα συγκεντρώνουν, και στους ίδιους τους παραγωγούς τους (Estrin, 2013 σε Lundquist, 2013)⁵. Πρόκειται για πιο διαχειρίσιμα σύνολα δεδομένων που διακινούνται σε *"ένα κατανεμημένο οικοσύστημα πληροφορίας"*, το οποίο παρέχει δυνατότητα πρόσβασης σε πολλαπλές πηγές για την ανάλυση του - κατάλληλου κατά περίπτωση- συνδυασμού επιμέρους συνόλων δεδομένων (Pollock, 2013). Ακόμη και έτσι όμως, θα μπορούσε να υποστηριχθεί ότι τα *small data* συνδέονται με τα ΜΔ, καθώς προϋπόθεση για την αξιοποίησή τους στο πλαίσιο εξαγωγής συμπερασμάτων, αποτελεί ο συνδυασμός επιμέρους συνόλων δεδομένων που αθροίζουν σε ένα νέο, δυνητικά πολύ μεγαλύτερο και ποικιλόμορφο σύνολο δεδομένων.

Ο Arbesman (2012) επισημαίνει ότι *"τα ΜΔ έχουν βάθος, αλλά δεν έχουν εύρος"*, καθώς παρότι σήμερα είναι δυνατή η συγκέντρωση τεράστιου όγκου δεδομένων, εντούτοις αυτός αφορά συγκεκριμένες εκφάνσεις της ανθρώπινης δραστηριότητας και επομένως η εξαγωγή συμπερασμάτων που βασίζονται σε αυτόν εμπεριέχει a priori μία δόση μεροληψίας. Η αντιπρότασή του για την αντιμετώπιση της μονομέρειας των ΜΔ είναι η συλλογή διαχρονικών μετρήσεων, ή αλλιώς *"long data"*, που καθιστούν εφικτή την παρακολούθηση της διαχρονικής εξέλιξης της ανθρώπινης δραστηριότητας και την παροχή πλαισίου εντός του οποίου ερμηνεύεται η παρατήρηση αυτής (Arbesman, 2013).

Η σημασία του πλαισίου τονίζεται από τους Boyd και Crawford (2012), στην απόπειρά τους να προσδιορίσουν τα κρίσιμα ζητήματα που αφορούν στα ΜΔ. Όπως επισημαίνουν, χωρίς πλαίσιο *"τα δεδομένα χάνουν το νόημα και την αξία τους"*, και επομένως η ύπαρξή του αποτελεί ζητούμενο κατά την εξέταση ενός

⁵ Η Estrin (2013 σε Lundquist, 2013) υποστηρίζει ότι οι χρήστες υπηρεσιών που καταγράφουν τα προσωπικά ίχνη της δραστηριότητάς τους, επωφελούνται όταν αυτά τα δεδομένα που παράγουν αξιοποιούνται για την παροχή εξατομικευμένων υπηρεσιών.

συνόλου ΜΔ (Boyd and Crawford, 2012: 670-671). Επομένως, προκύπτει ότι η τοποθέτηση των δεδομένων εντός πλαισίου είναι απαραίτητη για την ανάλυση και την εξαγωγή συμπερασμάτων για τα ΜΔ.

Η αξία που δυνητικά εμπεριέχεται στα ΜΔ αναγνωρίζεται σε διεθνές επίπεδο. Για παράδειγμα, το World Economic Forum αναγνωρίζει την αξία των προσωπικών δεδομένων (personal data) που *"παράγονται από και για τους ανθρώπους"* (World Economic Forum, 2011: 5) και των δεδομένων που αντλούνται από τις φορητές συσκευές. Θεωρεί δε ότι τα ΜΔ θα επιδράσουν θετικά σε τομείς όπως οι οικονομικές υπηρεσίες, η εκπαίδευση και η υγεία (World Economic Forum, 2012).

Παράλληλα, το Παγκόσμιο Ινστιτούτο της συμβουλευτικής εταιρείας McKinsey (Manyika et al., 2011: 4-6) υποστηρίζει ότι τα ΜΔ έχουν μετασχηματιστική επίδραση σε οργανισμούς και επιχειρήσεις και διατυπώνει τις ακόλουθες κατευθύνσεις για την αποκόμιση αξίας από αυτά:

- παροχή στα ενδιαφερόμενα μέρη άμεσης και εύκολης πρόσβασης στα δεδομένα, για την ταχεία ενσωμάτωσή τους στις εσωτερικές διεργασίες του οργανισμού ή της επιχείρησης
- διενέργεια πειραματισμών στην εξέταση και ανάλυση του αυξημένου όγκου των διαθέσιμων δεδομένων, με σκοπό την ανακάλυψη αιτιωδών σχέσεων και γενικότερα την καλύτερη κατανόηση των εξεταζόμενων περιπτώσεων
- διαίρεση της πελατειακής βάσης σε τμήματα για την παροχή εξατομικευμένων προσφορών προϊόντων και υπηρεσιών
- υποστήριξη της διαδικασίας λήψης αποφάσεων μέσω εκτέλεσης αυτοματοποιημένων αλγορίθμων
- χρήση των συλλεγόμενων δεδομένων για την παροχή καινοτόμων προϊόντων και υπηρεσιών.

Αυτή η εκτιμώμενη αξία έχει οδηγήσει στην ανάπτυξη ενός οικοσυστήματος εταιρειών που εξειδικεύονται στα ΜΔ και προσφέρουν λύσεις σε διαφορετικά

ζητήματα. Οι Mayer-Schönberger και Cukier (2013: 124-125) εντοπίζουν τρία είδη: (1) εταιρείες που κατέχουν δεδομένα ή έχουν πρόσβαση σε αυτά, χωρίς όμως να διαθέτουν την απαραίτητη τεχνογνωσία για να τα επεξεργαστούν και για να αντλήσουν αξία από αυτά, (2) συμβουλευτικές εταιρείες, πάροχοι τεχνολογίας ή λογισμικού ανάλυσης, που δεν κατέχουν δεδομένα αλλά διαθέτουν εξειδίκευση και αναλαμβάνουν την επεξεργασία δεδομένων, και (3) οργανισμοί που έχουν ενστερνιστεί πλήρως τα ΜΔ, διαθέτουν δηλαδή αυτό που οι συγγραφείς αποκαλούν *"big-data mindset"*: εκτός από την κατοχή δεδομένων και τεχνογνωσίας επεξεργασίας τους, απασχολούν ανθρώπινο δυναμικό με καινοτόμες ιδέες για την εξαγωγή αξίας από τα δεδομένα.

Συνοψίζοντας, καθίσταται αντιληπτό ότι, αν και ετυμολογικά ο όρος *"Μεγάλα Δεδομένα"* αυθόρμητα παραπέμπει στον όγκο, δεν θα πρέπει να παραβλέπονται οι υπόλοιπες ιδιότητες (ποικιλομορφία και ταχύτητα ανανέωσης) που διαφοροποιούν τα συλλεγόμενα δεδομένα από το παρελθόν. Επιπρόσθετα, όπως συζητείται και στη συνέχεια, τα ΜΔ δεν προκύπτουν απλά από τη συνύπαρξη των 3Vs, αλλά από τον συνδυασμό back-end τεχνολογιών⁶ που καθιστούν εφικτή τη συλλογή, την αποθήκευση και επεξεργασία τους με front-end λύσεις⁷ που επιτρέπουν την ανάλυση δεδομένων που χαρακτηρίζονται από όγκο, ποικιλομορφία και ταχύτητα, και τελικά διευκολύνουν την διερεύνηση και την εξαγωγή συμπερασμάτων από αυτά (Intel IT Center, 2013: 5). Όπως χαρακτηριστικά επισημαίνουν οι Hurwitz et al. (2013: 15), τα ΜΔ *"δεν είναι απλώς ένα εργαλείο ή μία τεχνολογία"*, αλλά *"ο τρόπος που όλες αυτές οι τεχνολογίες συνδυάζονται για να δώσουν τη σωστή εικόνα, τη σωστή στιγμή με βάση τα κατάλληλα δεδομένα, είτε αυτά παράγονται από ανθρώπους, μηχανές ή τον Ιστό"*.

⁶ Το back-end τμήμα ενός υπολογιστικού συστήματος περιλαμβάνει τεχνολογίες, π.χ. δρομολογητές (routers) και εξυπηρετητές (servers), που υποστηρίζουν τις ενέργειες ή τα αιτήματα του τελικού χρήστη (Techopedia, 2010).

⁷ Το front-end τμήμα ενός υπολογιστικού συστήματος αποτελεί τη διεπαφή μέσω της οποίας ο τελικός χρήστης αλληλεπιδρά με το back-end τμήμα (Techopedia, 2010).

Με άλλα λόγια, θα μπορούσε να ειπωθεί ότι ο όρος *"Μεγάλα Δεδομένα"* δεν θα πρέπει να χρησιμοποιείται μονοδιάστατα για να περιγράψει τα χαρακτηριστικά των συλλεγόμενων δεδομένων, καθώς αυτά δεν αρκούν από μόνα τους για να προκύψει η πολυθρύλητη αξία. Όπως άλλωστε επισημαίνει η Gartner (2014), τα ΜΔ αποτελούν *"πληροφοριακά περιουσιακά στοιχεία μεγάλου μεγέθους, υψηλής ταχύτητας και έντονης ποικιλίας που απαιτούν αποδοτικές και καινοτόμες μορφές επεξεργασίας πληροφορίας για την ενίσχυση της επίγνωσης και της λήψης αποφάσεων"*. Επομένως, θα μπορούσε να θεωρηθεί ότι τα ΜΔ συνιστούν ένα πολυδιάστατο φαινόμενο που διαμορφώνεται από τις τεχνολογικές δυνατότητες που επιτρέπουν τη συλλογή, την αποθήκευση, την επεξεργασία και την ανάλυσή τους.

2.2 Προέλευση ΜΔ

Στην παρούσα ενότητα περιγράφονται οι βασικότερες πηγές προέλευσης ΜΔ που σχετίζονται με το Διαδίκτυο: τα κοινωνικά μέσα (social media), οι κινητές συσκευές, το Διαδίκτυο των Πραγμάτων (Internet of Things) και οι εφαρμογές Ποσοτικοποιημένου Εαυτού (Quantified Self). Τα δεδομένα που συλλέγονται από αυτές τις σχετικά νέες πηγές, μόνα τους ή σε συνδυασμό με δεδομένα που παραδοσιακά καταγράφονται από τους οργανισμούς για διάφορους σκοπούς⁸, συνιστούν κατά κύριο λόγο μη δομημένα δεδομένα που προκύπτουν από τη δραστηριότητα εκατοντάδων χιλιάδων χρηστών και ανανεώνονται με μεγάλη συχνότητα. Επομένως, ο τεράστιος όγκος δεδομένων που τελικά συγκεντρώνεται βάσει των προαναφερθέντων πληροί τα κριτήρια για να χαρακτηριστεί ως ΜΔ.

Πριν την περιγραφή των πηγών αυτών, κρίνεται σκόπιμη η εξέταση δύο όρων: του Διαδικτύου (Internet) και του Ιστού (Web). Το Διαδίκτυο, ή αλλιώς *"το Δίκτυο των Δικτύων, που καθορίζεται από τα πρότυπα TCP/IP"* (W3C, 2009), και ο

⁸ Ως ενδεικτικό παράδειγμα αναφέρεται η συλλογή στοιχείων συναλλαγών για την παρακολούθηση πωλήσεων.

Ιστός ως "ένας πληροφοριακός χώρος εντός του οποίου τα αντικείμενα ενδιαφέροντος (πόροι) ταυτοποιούνται με παγκόσμια αναγνωριστικά" (W3C, 2004) έχουν συμβάλει σημαντικά στη διαμόρφωση των ΜΔ.

Από τους πιο πάνω ορισμούς, καθίσταται αντιληπτό ότι το Διαδίκτυο αποτελεί το σύνολο των υποδομών και των τεχνολογιών που καθιστούν εφικτή τη διασύνδεση μεταξύ υπολογιστών, ενώ ο Ιστός είναι η επινόηση που υποστηρίζει τη διασύνδεση μηχανών, πληροφοριών και ανθρώπων σε παγκόσμιο επίπεδο.

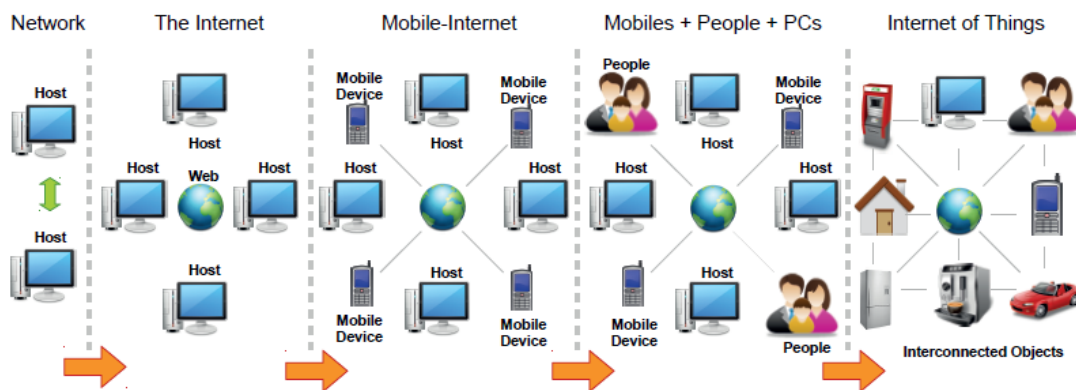
Αν και συχνά οι δύο όροι χρησιμοποιούνται για να περιγράψουν αυτό που εν τέλει συνιστά ο Ιστός (Pew Research Center, 2014a), ο Evans (2011: 5) επιμένει να τους διαχωρίζει, τονίζοντας τις λειτουργικές διαφορές τους και επισημαίνοντας ότι σε αντίθεση με το Διαδίκτυο⁹, εντοπίζονται τέσσερα σημαντικά εξελικτικά στάδια στην ανάπτυξη του Ιστού:

- **στάδιο 1:** αρχικά περιορισμένη χρήση, κυρίως από τα μέλη της ακαδημαϊκής κοινότητας για ερευνητικούς σκοπούς
- **στάδιο 2:** ευρεία διάδοση και αξιοποίησή του ως μέσο προβολής και ενημέρωσης για προϊόντα και υπηρεσίες
- **στάδιο 3:** άνθιση ηλεκτρονικού εμπορίου και συναλλαγών
- **στάδιο 4:** δραστηριοποίηση εταιρειών διαδικτυακών προϊόντων και υπηρεσιών που παράγουν κέρδος, παρέχοντας στους χρήστες τη δυνατότητα να επικοινωνούν και να διαμοιράζονται περιεχόμενο με άλλους χρήστες, οι οποίοι συνιστούν τον κοινωνικό τους περίγυρο.

Σε μία διαφορετική προσέγγιση, όπου Διαδίκτυο και Ιστός δεν διαχωρίζονται και το δεύτερο παρουσιάζεται ως εξελικτικό στάδιο του πρώτου, οι Perera et al. (2014: 416) εντοπίζουν πέντε φάσεις εξέλιξης του Διαδικτύου, που σηματοδοτούνται από τα ακόλουθα ορόσημα:

⁹ Επισημαίνει ότι, παρά τη σταθερή ανάπτυξη και βελτίωση του Διαδικτύου, η πρώτη ουσιαστική εξέλιξη από καταβολής του είναι το Διαδίκτυο των Πραγμάτων (Evans, 2011: 5).

- τη δυνατότητα επικοινωνίας μεταξύ δύο υπολογιστών μέσω δικτύου υπολογιστών στα τέλη της δεκαετίας του 1960
- την εισαγωγή πρωτοκόλλων TCP/IP τη δεκαετία του 1980 και τη διάθεση του Παγκόσμιου Ιστού (World Wide Web) από το 1991
- τη σύνδεση των κινητών συσκευών στο Διαδίκτυο και τη διαμόρφωση του κινητού Διαδικτύου (mobile-Internet)
- τη διασύνδεση μεταξύ χρηστών, μέσω των διαδικτυακών υπηρεσιών κοινωνικής δικτύωσης
- τη διασύνδεση αντικειμένων, τα οποία συνδέονται και επικοινωνούν με την ελάχιστη ανθρώπινη παρέμβαση μέσω του Διαδικτύου (Διαδίκτυο των Πραγμάτων).



Εικόνα 1: Οι φάσεις εξέλιξης του Διαδικτύου, όπως αποτυπώνονται σχηματικά από τους Perera et al. (2014: 416)

Για τους σκοπούς της παρούσας εργασίας, υιοθετείται η την προσέγγιση των Perera et al. Κατ' επέκταση, το Διαδίκτυο και ο Ιστός αντιμετωπίζονται ως ενιαίο σύνολο, το οποίο θα αναφέρεται ως *Διαδίκτυο* και θεωρείται η κυριότερη υποδομή προέλευσης ΜΔ.

2.2.1 Κοινωνικά μέσα

Οι ορισμοί για τα κοινωνικά μέσα περιλαμβάνουν την κοινωνική

αλληλεπίδραση που λαμβάνει χώρα μέσω αυτών και αναγνωρίζουν -σε διαφορετικές διαβαθμίσεις- τη σημασία των τεχνολογιών που έχουν καταστήσει εφικτή τη λειτουργία τους.

Ο Blossom (2009: 29) εστιάζει στη δυνατότητα που παρέχεται στους χρήστες να δημοσιεύουν και να διαμοιράζονται περιεχόμενο, χωρίς όμως να κατονομάζει συγκεκριμένη τεχνολογία, και ορίζει τα κοινωνικά μέσα ως *"οποιαδήποτε ευέλικτη και προσβάσιμη επικοινωνιακή τεχνολογία ή τεχνική που παρέχει σε οποιοδήποτε μεμονωμένο άτομο τη δυνατότητα να επηρεάσει ομάδες άλλων ατόμων εύκολα"*. Ωστόσο, οι Kaplan και Haenlein (2010: 61) προσδιορίζουν τον Ιστό 2.0 (Web 2.0) ως την τεχνολογική υποδομή που επέτρεψε τη μετάβαση από στατικές ιστοσελίδες, το περιεχόμενο των οποίων παραγόταν από μεμονωμένους χρήστες, σε μία δυναμική πλατφόρμα με περιεχόμενο που δημιουργείται και αναθεωρείται συνεχώς χάρη στη συμμετοχή και αλληλεπίδραση όλων των χρηστών. Υπό αυτό το πρίσμα, ορίζουν τα κοινωνικά μέσα ως *"μία ομάδα εφαρμογών Διαδικτύου που βασίζονται στα ιδεολογικά και τεχνολογικά θεμέλια του Ιστού 2.0 και που επιτρέπουν τη δημιουργία και ανταλλαγή Περιεχομένου Παραγόμενου από τον Χρήστη (User Generated Content)"* (Kaplan and Haenlein, 2010: 61).

Και στις δύο περιπτώσεις, κεντρικό ρόλο διαδραματίζει ο χρήστης που δημιουργεί περιεχόμενο και αλληλεπιδρά με άλλους χρήστες: αυτός είναι ο *"παραγωγός-καταναλωτής"* (prosumer), ο οποίος μπορεί να παράγει και να συνεισφέρει το δικό του περιεχόμενο και δεν αρκείται πλέον στο να καταναλώνει παθητικά την πληροφορία που του μεταδίδεται μονόδρομα από παραδοσιακά μέσα (Tapscott and Williams, 2008: 143-146). Σε τι όμως συνίσταται το Περιεχόμενο Παραγόμενο από τον Χρήστη; Ο Οργανισμός για την Οικονομική Συνεργασία και Ανάπτυξη (ΟΟΣΑ/OECD, 2007: 8-9) θέτει τρία κριτήρια που διαμορφώνουν το πλαίσιο εντός του οποίου ένα περιεχόμενο εμπίπτει σε αυτή την κατηγορία: (1) το περιεχόμενο είναι ευρέως διαθέσιμο μέσω Διαδικτύου ή ιστοσελίδων κοινωνικής δικτύωσης, (2) ο χρήστης έχει καταβάλει δημιουργική προσπάθεια για την δημιουργία του, είτε πρωτογενώς ή μετασκευάζοντας ένα

υφιστάμενο, και (3) το περιεχόμενο έχει παραχθεί εκτός των συνήθων επαγγελματικών πρακτικών και από μη εμπορικές οντότητες. Έχοντας προσδιορίσει το πλαίσιο, ο ΟΟΣΑ καταγράφει τους τύπους Περιεχομένου Παραγόμενου από τον Χρήστη: κείμενο, φωτογραφίες και εικόνες, μουσική και αρχεία ήχου, βίντεο και ταινίες, καθώς και συμβουλευτικό περιεχόμενο για αγορές ή ενδιαφέροντα (OECD, 2007: 16-18).

Ωστόσο, η φύση του περιεχομένου που σήμερα παράγεται από τους χρήστες και διακινείται μέσω των κοινωνικών μέσων έχει αλλάξει. Τα συλλεγόμενα δεδομένα έχουν αυξηθεί και περιλαμβάνουν νέες μορφές, όπως οι αξιολογήσεις στα κοινωνικά μέσα και τα μετα-δεδομένα¹⁰, που παράγονται από τη δραστηριότητα των χρηστών εν αγνοία τους ή/και δεν απαιτούν την καταβολή δημιουργικής προσπάθειας από πλευράς τους. Τα δεδομένα αυτά αποτελούν πλούσιες και πολύτιμες πηγές πληροφοριών και η επεξεργασία τους είναι δυνατό να οδηγήσει σε συμπεράσματα για τον χρήστη και να αξιοποιηθεί ποικιλοτρόπως.

Προκειμένου να εξεταστεί η μεταβολή των δεδομένων που προέρχονται από τη δραστηριότητα των χρηστών κοινωνικών μέσων, περιγράφεται το παράδειγμα του *Facebook*¹¹. Μέχρι το 2010, όλες σχεδόν οι πληροφορίες για τα προσωπικά στοιχεία και τις προτιμήσεις των χρηστών του *Facebook* καταγράφονταν σε πεδία απλού κειμένου (Sun and Iyer, 2013), επομένως τα δεδομένα που συλλέγονταν ήταν κυρίως μη δομημένα.

Τον Απρίλιο του 2010 επιχειρήθηκε η μαζική μετατροπή αυτών των πεδίων απλού κειμένου σε δομημένες συνδέσεις, αρχικά με αυτόματο τρόπο βάσει παραδοχών και αλγορίθμων, και στη συνέχεια προσκαλώντας τους χρήστες να παράσχουν ανατροφοδότηση και να αναφέρουν σφάλματα (Sun and Iyer, 2013

¹⁰ Ενδεικτικά αναφέρεται το παράδειγμα της καταγραφής της γεωγραφικής θέσης του χρήστη τη στιγμή μίας ανάρτησης.

¹¹ Πρόκειται για το δημοφιλέστερο κοινωνικό μέσο παγκοσμίως (Lunden, 2014) και αριθμεί καθημερινά πάνω από 800 εκατομμύρια ενεργούς χρήστες, με στοιχεία Μαρτίου 2014 (Facebook Newsroom, 2014).

και Hernandez, 2013). Οι αλλαγές αυτές παρουσιάστηκαν στους χρήστες ως βελτίωση της λειτουργικότητας των σελίδων προφίλ, οι οποίες πλέον περιλάμβαναν "συνδέσεις"¹² σε αντικατάσταση "των βαρετών κειμένων" και μετατράπηκαν σε "ζωντανό χάρτη όλων των συνδέσεων" σε μέρη, αντικείμενα και εμπειρίες που ενδιαφέρουν τον χρήστη (Li, 2010).



Εικόνα 2: Εκτύπωση οθόνης σελίδας προφίλ χρήστη Facebook το 2006
(Πηγή: Sun and Iyer, 2013)

Έκτοτε, το Facebook συγκεντρώνει με ευρηματικές μεθόδους δεδομένα από την αλληλεπίδραση των χρηστών: (1) ενθαρρύνοντάς τους να μοιράζονται πληροφορίες, από τα δημογραφικά τους στοιχεία έως και σημαντικές προσωπικές τους στιγμές, (2) χαρακτηρίζοντας κάθε φωτογραφία και μήνυμα που διαμοιράζεται με την ακριβή γεωγραφική θέση του χρήστη, (3) αντλώντας δεδομένα για τη δραστηριότητα των χρηστών από εφαρμογές που δεν σχετίζονται άμεσα με το εν λόγω κοινωνικό μέσο¹³ (Simonite, 2012).

¹² Απόδοση του αγγλικού όρου "connections".

¹³ Το Facebook ενθαρρύνει τους κατασκευαστές εφαρμογών να παρέχουν στους χρήστες λειτουργικότητα διαμοιρασμού της δραστηριότητας τους σε αυτές, μέσω του εν λόγω κοινωνικού μέσου (Fedorov, 2012).

2.2.2 Φορητές Συσκευές

Εκτός όμως από τα κοινωνικά μέσα, που αποτελούν φορείς σημαντικών πληροφοριών για τη δραστηριότητα, τα ενδιαφέροντα και τις προτιμήσεις των χρηστών τους, οι φορητές συσκευές συνιστούν μία ακόμη πηγή δεδομένων που έχει γνωρίσει άνθιση την τελευταία δεκαετία. Η ευρεία διάδοση των έξυπνων τηλεφώνων (smartphones) επιτρέπει τη χρήση των συσκευών αυτών όχι μόνο για τη διεκπεραίωση τηλεφωνικών κλήσεων, αλλά και για την πρόσβαση στο Διαδίκτυο.

Σε έκθεση της Ευρωπαϊκής Επιτροπής για την παρακολούθηση των στόχων που έχουν τεθεί στο πλαίσιο του Ψηφιακού Θεματολογίου (Digital Agenda), επισημαίνεται ότι η χρήση κινητού τηλεφώνου για πρόσβαση στο Διαδίκτυο στην Ευρώπη σχεδόν διπλασιάστηκε (27% του πληθυσμού το 2012 από 14% το προηγούμενο έτος), με τον όγκο των δεδομένων που διακινούνται με αυτόν τον τρόπο να αναμένεται να δεκαπλασιαστεί μέσα στην πενταετία 2010-2015 (European Commission, 2013: 26). Σε παγκόσμιο επίπεδο παρατηρείται αντίστοιχη τάση: σύμφωνα με το Ericsson Mobility Report (Ericsson, 2013: 11), η αύξηση στα δεδομένα που διακινούνται μέσω κινητών συσκευών, με αιχμή τα έξυπνα τηλέφωνα, συνεχίστηκε και το 2013, χωρίς μάλιστα να διαφαίνονται σημεία επιβράδυνσης.

Τα στοιχεία αυτά υποδηλώνουν τεράστια ζήτηση από πλευράς καταναλωτών, η οποία συνεπάγεται αύξηση του όγκου και της ποικιλίας των δεδομένων που μπορούν να συλλεχθούν και να ταυτοποιηθούν με συγκεκριμένους χρήστες. Δεδομένου ότι κάθε συσκευή είναι προσωπική και φορητή, μέσω αυτής είναι δυνατή η καταγραφή μιας συνεχούς ροής ατομικών δεδομένων εν κινήσει. Πρόκειται λοιπόν για φορέα επιπλέον πληροφορίας για την ανθρώπινη δραστηριότητα (Davenport, 2014: 13), κρυμμένης σε μετα-δεδομένα, όπως η τοποθεσία ή η διάρκεια κλήσεων, οι αλληλεπιδράσεις στα κοινωνικά μέσα, τα ίχνη πλοήγησης στο Διαδίκτυο και οι ηλεκτρονικές συναλλαγές.

Και ενώ κάποια χρόνια νωρίτερα, η φράση "φορητές συσκευές" παρέπεμπε σε κινητά τηλέφωνα ή φορητούς ηλεκτρονικούς υπολογιστές, πλέον εμπερικλείει πλήθος συσκευών, από ταμπλέτες μέχρι και φορέσιμους υπολογιστές (wearable computers), που λειτουργούν ως εναλλακτικό σημείο πρόσβασης στο Διαδίκτυο ή/και ως νέες πηγές εισόδου δεδομένων. Σύμφωνα με έρευνα του Ericsson ConsumerLab (2013: 4), το 40% των χρηστών έξυπνων κινητών τηλεφώνων θα ήθελε η συσκευή να καταγράφει δεδομένα σχετικά με άσκηση και ύπνο, ενώ το 60% θα επιθυμούσε η καταγραφή αυτή να γίνεται μέσω συσκευής καρπού.

Σύμφωνα με τα προαναφερθέντα, γίνεται κατανοητό ότι και στην περίπτωση της ευρείας κατηγορίας των φορητών συσκευών, είναι δυνατόν να συλλεχθούν δεδομένα με ποικιλία στη μορφή και τη δομή, με μεγάλη ταχύτητα ανανέωσης που προκαλεί συνακόλουθα τη συγκέντρωση ενός τεράστιου όγκου ετερογενών δεδομένων.

2.2.3 Ποσοτικοποιημένος Εαυτός

Μέχρι πρότινος, τα μέσα για την καταγραφή προσωπικών δεδομένων, όπως ζωτικές λειτουργίες, περιορίζονταν στο χαρτί. Σύμφωνα με τον Wolf (2010), η σμίκρυνση των ηλεκτρονικών αισθητήρων, η εξάπλωση ισχυρών υπολογιστικών συσκευών όπως τα έξυπνα κινητά τηλέφωνα, η συνήθεια του διαμοιρασμού ατομικών πληροφοριών μέσω των κοινωνικών μέσων και η νεφοϋπολογιστική (cloud computing)¹⁴, κατέστησαν εφικτή την ακριβέστερη και αντικειμενική

¹⁴ Σύμφωνα με τον αναλυτικό ορισμό των Mell και Grance (2011: 2), η νεφοϋπολογιστική είναι "ένα μοντέλο που καθιστά εφικτή την από οποιοδήποτε σημείο, εύκολη και κατόπιν αιτήματος πρόσβαση σε μία κοινή δεξαμενή παραμετροποιήσιμων υπολογιστικών πόρων (π.χ. δίκτυα, εξυπηρετητές, υποδομές αποθήκευσης, εφαρμογές και υπηρεσίες) που μπορούν να τροφοδοτηθούν γρήγορα και να αποδεσμευθούν με την ελάχιστη δυνατή διαχειριστική προσπάθεια ή αλληλεπίδραση με τον πάροχο".

Κεντρικό ρόλο κατέχει η έννοια του νέφους (cloud), το οποίο αποτελεί την υποδομή, σε επίπεδο υλισμικού και λογισμικού, που υποστηρίζει την παροχή υπηρεσιών νεφοϋπολογιστικής. Ανάλογα με το είδος πρόσβασης που ο πάροχος χορηγεί στον χρήστη, διακρίνονται τρία μοντέλα υπηρεσιών (Mell and Grance, 2011: 2-3): (1) οι υπηρεσίες SaaS (Software as a Service) παρέχουν πρόσβαση σε λογισμικό που τρέχει σε υποδομή νέφους, (2) οι υπηρεσίες PaaS (Platform as a Service) παρέχουν

καταγραφή αυτών των δεδομένων.

Παρατηρώντας ότι κοινό στοιχείο μεταξύ των τεχνολογιών καταγραφής βιομετρικών δεδομένων, γεωγραφικής θέσης και καθημερινών δραστηριοτήτων αποτελεί η *"προσθήκη μίας υπολογιστικής διάστασης στη συνηθισμένη ύπαρξη"* (Wolf, 2011), οι Wolf και Kelly καθιέρωσαν τον όρο *"Quantified Self"*¹⁵. Ο όρος αυτός περιγράφει την πεποίθηση ότι μέσω της συστηματικής καταγραφής και ανάλυσης ποσοτικών δεδομένων σε ατομικό επίπεδο είναι δυνατή η επίτευξη αυτογνωσίας.¹⁶

Οι εφαρμογές του *Ποσοτικοποιημένου Εαυτού* δεν στοχεύουν μόνο στην αυτοβελτίωση του ατόμου. Σε ακραίες εκφάνσεις του, το σώμα νοείται ως συνδεδεμένη *"έξυπνη μηχανή"* που διασυνδέεται με άλλες έξυπνες μηχανές, αποτελώντας έναν ακόμη κόμβο στο Διαδίκτυο των Πραγμάτων (Lupton, 2013: 27). Επιπρόσθετα, υπάρχουν ήδη επιχειρήσεις που συλλέγουν και αναλύουν δεδομένα σχετικά με την ποιότητα του ύπνου, τη φυσική δραστηριότητα, την πρόσληψη θερμίδων καθώς και τη διάθεση των εργαζομένων τους, προκειμένου

τη δυνατότητα ανάπτυξης εφαρμογών με τη χρήση εργαλείων και γλωσσών προγραμματισμού που διαθέτει ο πάροχος, και (3) οι υπηρεσίες IaaS (Infrastructure as a Service) παρέχουν την πρόσβαση σε επεξεργαστική ισχύ, αποθηκευτικά μέσα και άλλους υπολογιστικούς πόρους, στους οποίους ο χρήστης μπορεί να αναπτύξει ή να εγκαταστήσει λογισμικό.

Αυτή η δυνατότητα πρόσβασης σε κατανεμημένους υπολογιστικούς πόρους αίρει τους οικονομικούς και τεχνολογικούς περιορισμούς του παρελθόντος, καθώς επιτρέπει τη συλλογή, την αποθήκευση και την επεξεργασία δεδομένων με σημαντικά χαμηλότερο κόστος για τους οργανισμούς. Παράλληλα, οι ανάγκες για αυξημένη επεξεργαστική ισχύ και αποθηκευτικά μέσα διαρκώς ενισχύονται από τη μαζική χρήση προϊόντων και υπηρεσιών που αξιοποιούν τις δυνατότητες της νεφοϋπολογιστικής και από την ενσωμάτωσή τους στην καθημερινότητα των χρηστών.

¹⁵ Αρχικά, δημιούργησαν μία ιστοσελίδα καταγραφής των νέων συστημάτων αυτό-καταγραφής, ονομάζοντας το εγχείρημα *"Quantified Self"* (Wolf, 2009). Στη συνέχεια, αυτό εξελίχθηκε σε μία διεθνή κοινότητα χρηστών και κατασκευαστών εργαλείων αυτο-καταγραφής (Quantified Self, 2012).

¹⁶ Χαρακτηριστικά, η φράση που συνοδεύει το λογότυπο *"Quantified Self"* στη σχετική ιστοσελίδα www.quantifiedself.com είναι *"αυτογνωσία μέσω αριθμών"* (*"self-knowledge through numbers"*).

να εντοπίσουν τους παράγοντες που μπορούν να αυξήσουν την παραγωγικότητά τους¹⁷.

Ενώ λοιπόν στο παρελθόν οι μέθοδοι αυτο-καταγραφής (self-tracking) απαιτούσαν προσπάθεια από πλευράς χρήστη και δεν ήταν απολύτως ακριβείς, η αυτοματοποίηση της διαδικασίας συγκέντρωσης δεδομένων και η δυνατότητα ανάλυσής τους μέσω εύχρηστων διαδικτυακών εφαρμογών καθιστούν εφικτή την επίγνωση του εαυτού με διαφορετικό τρόπο (Antepphase, 2014). Οι υποστηρικτές των εφαρμογών του *Ποσοτικοποιημένου Εαυτού* θεωρούν ότι τα αριθμητικά δεδομένα που προκύπτουν από την αυτόματη καταγραφή είναι απαλλαγμένα από υποκειμενικότητα, διότι συλλέγονται με περιορισμένη ή και μηδαμινή καταβολή συνειδητής προσπάθειας για την ενδεδειγμένη καταγραφή των δραστηριοτήτων, και επομένως δεν αλλοιώνονται από τον ανθρώπινο παράγοντα.

Για παράδειγμα, οι συσκευές της εταιρείας *Fitbit* παρέχουν στους χρήστες τους τη δυνατότητα καταγραφής δεδομένων σχετικά με την καθημερινή φυσική δραστηριότητα, την κατανάλωση τροφής, τη μέτρηση βάρους και την ποιότητα ύπνου. Εκτός από το υλισμικό, που συνίσταται σε ένα οικοσύστημα φορητών κυρίως συσκευών με αισθητήρες ανίχνευσης κίνησης, ιδιαίτερη έμφαση δίδεται στον συνδυασμό του με το ειδικά ανεπτυγμένο λογισμικό εμπλουτισμού και ανάλυσης των συλλεγόμενων δεδομένων: οι ειδικά σχεδιασμένες διεπαφές αναπαριστούν οπτικά τα συλλεχθέντα δεδομένα και βοηθούν τον χρήστη να θέσει στόχους βελτίωσης της τρέχουσας φυσικής του κατάστασης (π.χ. μείωση βάρους) και να "*παρακολουθεί την πρόοδό του*" για να "*αποκτήσει βαθύτερη γνώση*" σχετικά με την κατάσταση της υγείας του (Fitbit, 2014a).

Η πρόσβαση του χρήστη στα δεδομένα που παράγει είναι κατ' αρχάς εφικτή μέσω πιστοποιημένης εισόδου στη σχετική ιστοσελίδα ή/και ειδικής εφαρμογής

¹⁷ Ως παράδειγμα αναφέρεται η εταιρεία ανάλυσης δεδομένων *The Outside View* που υλοποιεί το πρόγραμμα Υγείας, Ευεξίας και Ευτυχίας, στο οποίο συμμετέχει υποχρεωτικά το σύνολο του προσωπικού της (Dato, 2014).

για έξυπνα κινητά τηλέφωνα. Μέσω των ειδικά σχεδιασμένων διεπαφών, ο χρήστης μπορεί παράλληλα να εμπλουτίσει τις αυτόματες καταγραφές με χειροκίνητες καταχωρίσεις, για την πληρέστερη καταγραφή των δραστηριοτήτων του. Η *Fitbit* ισχυρίζεται ότι μέσω της συστηματικής καταγραφής και της ανάλυσης των δεδομένων, ο χρήστης μπορεί να εξάγει συμπεράσματα για τη φυσική του κατάσταση, παρακολουθώντας τη διαχρονική εξέλιξη και εντοπίζοντας τάσεις στα δεδομένα, και να λάβει αποφάσεις για τη βελτίωσή της.

Επισημαίνεται όμως ότι, παρά τη δυνατότητα παρακολούθησης δραστηριότητας μέσω αισθητήρων που μειώνει την απαιτούμενη προσπάθεια καταγραφής από τον χρήστη και την υποκειμενικότητά του, η ακρίβεια των δεδομένων που συλλέγονται εξακολουθεί να αποτελεί ζητούμενο. Για παράδειγμα, οι συσκευές που καταγράφουν δεδομένα σχετικά με την ποιότητα του ύπνου, βασίζονται στις κινήσεις του μέλους στο οποίο φοριούνται, και όχι στα εγκεφαλικά κύματα που παραδοσιακά αποτελούν ακριβές μέτρο παρακολούθησης του ύπνου, με αποτέλεσμα να μην είναι πάντα δυνατή η ακριβής καταγραφή των μοτίβων ύπνου και επομένως να εξάγονται συμπεράσματα σε λάθος βάση (Cooper, 2014).

2.2.4 Διαδίκτυο των Πραγμάτων

Το Διαδίκτυο των Πραγμάτων συνίσταται στην *"ενσωμάτωση πομπών σε συσκευές και καθημερινά αντικείμενα, που καθιστά εφικτές νέες μορφές επικοινωνίας μεταξύ ανθρώπων και πραγμάτων, αλλά και μεταξύ πραγμάτων καθ'αυτών"* (International Telecommunications Union, 2005: 2). Ο ορισμός αυτός διατυπώθηκε σε μία εποχή που ακόμη ο αριθμός των διασυνδεδεμένων συσκευών υπολειπόταν του συνολικού ανθρώπινου πληθυσμού (Evans, 2011: 3) και το Διαδίκτυο των Πραγμάτων φάνταζε ως προοπτική. Σήμερα πια όμως σχηματίζεται *"το δίκτυο φυσικών αντικειμένων, προσβάσιμων μέσω του Διαδικτύου, τα οποία είναι τεχνολογικώς εφοδιασμένα για να αλληλεπιδρούν με εσωτερικές καταστάσεις ή με το εξωτερικό περιβάλλον"* (Cisco Systems, 2014).

Για να συμβεί αυτή η αλληλεπίδραση, τα αντικείμενα θα πρέπει να ταυτοποιούνται μονοσήμαντα, να φέρουν αισθητήρες μέσω των οποίων θα λαμβάνουν πληροφορίες από το εξωτερικό περιβάλλον, καθώς και να χρησιμοποιούν ενεργοποιητές (actuators) προκειμένου να αντιδρούν στα εξωτερικά ερεθίσματα (Commission of the European Communities, 2009: 2).

Οι βελτιώσεις των αισθητήρων, τόσο σε επίπεδο υλισμικού (μέγεθος, ισχύς) όσο και σε κόστος, έχουν συντελέσει στην ευρεία χρήση τους, που δεν περιορίζεται σε κινητές συσκευές. Αυτό όμως είναι το αρχικό στάδιο, καθώς σύμφωνα με πρόσφατες εκτιμήσεις (Gartner, 2013), το 2020 οι συσκευές που φέρουν αισθητήρες, εξαιρουμένων των προσωπικών υπολογιστών, των ταμπλετών και των έξυπνων τηλεφώνων, θα ανέρχονται σε 26 δισεκατομμύρια μονάδες, τριάντα φορές περισσότερες από το 2009.

Σε έρευνα του Pew Research Center (2014b: 2) για τις προοπτικές του Διαδικτύου των Πραγμάτων, ειδικοί απάντησαν ότι οι πρακτικές εφαρμογές του αναμένεται να είναι εμφανείς:

- στο ανθρώπινο σώμα, καθώς συσκευές συνδεδεμένες στο Διαδίκτυο θα καταγράφουν δεδομένα για να παράσχουν ανατροφοδότηση στον χρήστη σχετικά με την υγεία και τη φυσική του κατάσταση
- στις κατοικίες, καθώς θα είναι εφικτός ο έλεγχος και ο χειρισμός συσκευών από απόσταση
- στις κοινότητες, καθώς θα επιτυγχάνεται συνετή διαχείριση του ηλεκτρισμού και των υδάτινων αποθεμάτων
- σε προϊόντα και υπηρεσίες, καθώς θα παρακολουθείται και θα διασφαλίζεται η ομαλότητα σε ό,τι αφορά στην παραγωγή και στην διακίνηση των αγαθών
- στο περιβάλλον, καθώς θα καταγράφονται σε πραγματικό χρόνο δεδομένα π.χ. για τα επίπεδα ρύπανσης ή εξόρυξης πόρων που θα επιτρέπουν τη στενή παρακολούθηση τυχόν προβλημάτων και κατ' επέκταση την έγκαιρη αντιμετώπισή τους.

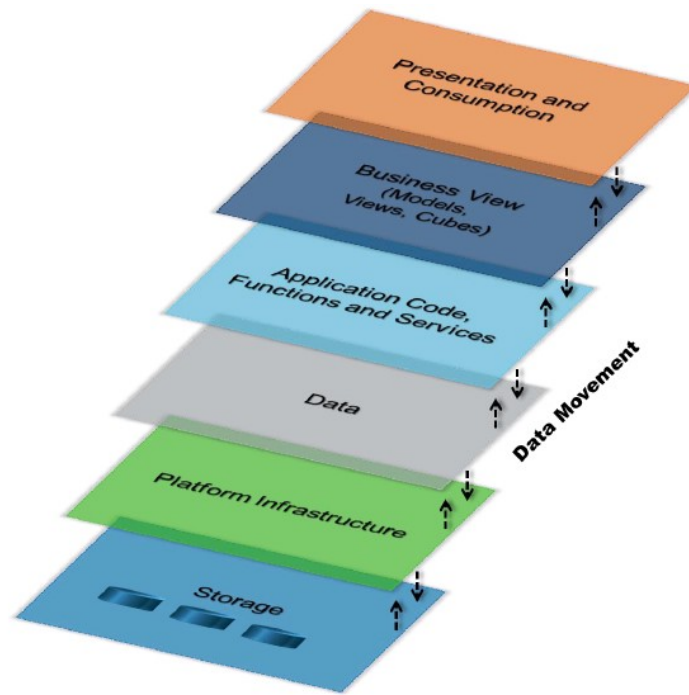
Από τα παραπάνω καθίσταται αντιληπτό ότι αυτή η διασύνδεση των πάντων με τα πάντα συνεπάγεται την παραγωγή και συλλογή τεράστιου όγκου δεδομένων που χαρακτηρίζονται από μεγάλη ποικιλία και ετερογένεια. Η διαχείριση αυτού του όγκου προκαλεί προβληματισμό για επιπλέον απαιτήσεις σε χρόνο επεξεργασίας και αποθήκευση (Perera et al., 2014: 433) και θέτει σοβαρό ζήτημα επεκτασιμότητας (Brech et al., 2013).

2.3 Διαχείριση ΜΔ

Για την αποτελεσματική διαχείριση των ΜΔ είναι απαραίτητος ο συνδυασμός επιμέρους τεχνολογικών λύσεων. Μία τέτοια ολιστική προσέγγιση αποτελεί η *διαστρωματωμένη δομή ΜΔ*¹⁸ (Davenport and Dyché, 2013: 9-12) που περιλαμβάνει τις ακόλουθες διαστρωματώσεις:

- την τεχνολογική υποδομή αποθήκευσης δεδομένων, είτε σε τοπικούς δίσκους είτε σε απομακρυσμένους εξυπηρετητές
- την πλατφόρμα διαχείρισης και εφαρμογής υπολογιστικών μεθόδων επεξεργασίας δεδομένων
- τα δεδομένα καθαυτά, που απαιτούν ιδιαίτερες πολιτικές διαχείρισης και διακυβέρνησης
- τους κώδικες προγραμματισμού για τον χειρισμό και την επεξεργασία των δεδομένων, οι οποίοι προσαρμόζονται στα εξεταζόμενα δεδομένα
- την προβολή για επιχειρηματικούς σκοπούς, που αφορά την παροχή ενός επεξεργασμένου ενδιαμέσου συνόλου δεδομένων για περαιτέρω ανάλυση με τα συνήθη εργαλεία επεξεργασίας και ανάλυσης
- την παρουσίαση και κατανάλωση, όπου κυριαρχούν τα λογισμικά οπτικοποίησης δεδομένων για την παρουσίαση της πληροφορίας με εύληπτο τρόπο.

¹⁸ Απόδοση του αγγλικού όρου "*big data stack*".



Εικόνα 3: Η διαστρωματωμένη δομή ΜΔ
(Πηγή: Davenport and Dyché, 2013: 10)

Παρατηρείται ότι στη βάση της προσέγγισης αυτής βρίσκεται η αποθήκευση των παραγόμενων δεδομένων, που αφορά στην καταγραφή και στη δυνατότητα ανάκτησής τους για σκοπούς ανάλυσης και περαιτέρω αξιοποίησής τους. Άλλωστε η πρόοδος που παρατηρείται τα τελευταία χρόνια στις τεχνολογίες που υποστηρίζουν τη συλλογή του συνεχώς αυξανόμενου όγκου δεδομένων (Piramanayagam, 2012: 2) και η δραστική μείωση του κόστους αποθήκευσης σε βάσεις δεδομένων (Davenport, 2014: 120), έχουν συμβάλει στη συγκέντρωση των ΜΔ. Ένα χαρακτηριστικό παράδειγμα μείωσης κόστους αναφέρεται από τους Davenport και Dyché (2013: 3): σύμφωνα με μαρτυρία που επικαλούνται, το ετήσιο κόστος αποθήκευσης ενός terabyte¹⁹ ισούται (1) με \$37,000 σε μία παραδοσιακή σχεσιακή βάση δεδομένων (traditional relational database), (2) με \$5,000 σε ένα ολοκληρωμένο σύστημα βάσεων δεδομένων με εξοπλισμό και λογισμικό (database appliance) και (3) με μόλις \$2,000 όταν χρησιμοποιείται το

¹⁹ 1 TB ισούται με 1.000 GB.

ανοικτό λογισμικό Hadoop σε μία ομάδα υπολογιστών (cluster).

Στον ακόλουθο πίνακα (Πίνακας 2), υπολογίζεται, βάσει του προαναφερθέντος παραδείγματος, η ποσοστιαία μείωση του κόστους αποθήκευσης και διαπιστώνεται περιορισμός του έως και κατά 95%:

Πίνακας 2: Σύγκριση κόστους αποθήκευσης, σύμφωνα με Davenport και Dyché (2013: 3)

αποθήκευση	ετήσιο κόστος ανά TB	διαφορά σε σύγκριση με (1)
(1) παραδοσιακή σχεσιακή βάση δεδομένων	\$37,000	
(2) ολοκληρωμένο σύστημα βάσεων δεδομένων	\$5,000	-86%
(3) χρήση Hadoop σε μία ομάδα υπολογιστών	\$2,000	-95%

Σε ό,τι αφορά στη συνολική διαχείριση δεδομένων, οι Hurwitz et al. (2013: 11-15) διακρίνουν τρία κύματα εξελίξεων μέχρι σήμερα:

- Το *πρώτο κύμα* σηματοδοτείται από το μοντέλο σχεσιακών δεδομένων (relational data) και τα εργαλεία λογισμικού, μέσω των οποίων και στο πλαίσιο της διαδικασίας λήψης αποφάσεων καθίσταται δυνατή η οργάνωση των δεδομένων για τη διενέργεια συγκρίσεων και μελετών. Το μοντέλο αυτό, επινοήθηκε στα τέλη της δεκαετίας του 1970, και χρησιμοποιείται ακόμη και σήμερα για τη διαχείριση δομημένων δεδομένων που συλλέγονται από συναλλαγές. Προς το τέλος αυτής της φάσης, αξιοποιούνται αρχικά οι αποθήκες δεδομένων (data warehouses) και στη συνέχεια οι αγορές δεδομένων (data marts) που αποτελούν πιο εξειδικευμένο υποσύνολο των πρώτων, ωστόσο σύντομα προκύπτουν ζητήματα που αφορούν στις απαιτήσεις για εξέταση των δεδομένων σε

πραγματικό χρόνο και στον ολοένα αυξανόμενο όγκο των μη δομημένων δεδομένων.

- Το *δεύτερο κύμα* πυροδοτείται από έναν συνδυασμό εξελίξεων όπως η εμπορική εξάπλωση του Διαδικτύου τη δεκαετία του 1990, η εκτεταμένη χρήση των συστημάτων διαχείρισης περιεχομένου (content management systems) και η νεφοϋπολογιστική. Οι εξελίξεις αυτές αύξησαν τις πηγές προέλευσης των δεδομένων και τις μορφές που αυτά μπορούσαν να λάβουν, καθιστώντας επιτακτική την αναζήτηση αποτελεσματικών και γρήγορων τρόπων επεξεργασίας τους.
- Το *τρίτο κύμα* αφορά τα ΜΔ και βασίζεται στα δύο προηγούμενα, με τους οργανισμούς να μπορούν πλέον να συλλέγουν όλες τις πληροφορίες που είναι απαραίτητες για μια διεξοδική ανάλυση και με σημαντικά μειωμένο κόστος, όπως έχει ήδη αναφερθεί πιο πάνω.

Βάσει των προαναφερθέντων, γίνεται κατανοητό ότι η άρση των οικονομικών και τεχνολογικών περιορισμών του παρελθόντος επιτρέπει τη συλλογή, την αποθήκευση και την επεξεργασία ΜΔ από πολλαπλές πηγές και με σημαντικά χαμηλότερο κόστος για τους οργανισμούς. Ωστόσο, η συνεχώς αυξανόμενη παραγωγή δεδομένων από τη μαζική χρήση προϊόντων και υπηρεσιών όπως τα κοινωνικά μέσα ή οι εφαρμογές *Ποσοτικοποιημένου Εαυτού*, σε συνδυασμό με το ότι η αξία των ΜΔ εξαρτάται από τις δυνατότητες ανάλυσής τους και εξαγωγής συμπερασμάτων από αυτά, τροφοδοτούν και ενισχύουν τις ανάγκες για αυξημένη επεξεργαστική ισχύ και αποθηκευτικά μέσα.

2.4 Η σκοτεινή πλευρά των ΜΔ

Εκτός από τη διάχυτη αισιοδοξία για την αξία που εμπεριέχεται στα ΜΔ, έχουν διατυπωθεί και εύλογοι προβληματισμοί. Σε ειδικό αφιέρωμα για τα ΜΔ το 2010, ο Economist (2010) επισημαίνει ότι οι κυβερνήσεις θα πρέπει να θεσπίσουν ένα κανονιστικό πλαίσιο που θα προσδιορίζει τις βέλτιστες πρακτικές στη συλλογή και επεξεργασία δεδομένων, με έμφαση στους

ακόλουθους τομείς:

- στη διευθέτηση της σύγκρουσης μεταξύ της προστασίας των χρηστών υπηρεσιών που συλλέγουν δεδομένα και του συμφέροντος των εν λόγω υπηρεσιών να εκμεταλλεύονται αυτές τις προσωπικές πληροφορίες
- στην προστασία των υπολογιστικών συστημάτων και δικτύων από απειλές που θέτουν σε κίνδυνο την ασφάλεια των συλλεγόμενων δεδομένων
- στον καθορισμό περιορισμένου χρονικού διαστήματος διατήρησης των αποθηκευμένων δεδομένων, μετά το πέρας του οποίου αυτά θα διαγράφονται, για την αποφυγή κακόβουλης χρήσης ή δημοσιοποίησης
- στην πρόληψη διακρίσεων από πιθανή κατάχρηση των συλλεγόμενων δεδομένων
- στην αντιμετώπιση των προσωπικών πληροφοριών ως πνευματικό δικαίωμα
- στη διασφάλιση της ακεραιότητας της πληροφορίας και στην αποφυγή λογοκρισίας ή νόθευσης αυτής.

Οι προβληματισμοί αυτοί, οι οποίοι εστιάζονται κυρίως στις απειλές για την ιδιωτικότητα (privacy), παραμένουν επίκαιροι ακόμη και σήμερα. Ο Morozov (2013: 230-235), στο πλαίσιο της κριτικής του για την κουλτούρα του *Ποσοτικοποιημένου Εαυτού* και της αμφισβήτησης της επίτευξης της αυτογνωσίας μέσω αριθμών που αυτή πρεσβεύει, αφενός αμφισβητεί την εξιδανίκευση της ποσοτικοποίησης και την πεποίθηση ότι η ανάλυση αυτών των δεδομένων θα βελτιώσει την ποιότητα των λαμβανόμενων αποφάσεων, αφετέρου θέτει ζήτημα επαναδιαπραγμάτευσης της έννοιας της ιδιωτικότητας, διαβλέποντας ακόμη και την εκχώρησή της έναντι αντιτίμου.

Αν οι θέσεις του είναι αναμενόμενες, δεδομένης της κριτικής του στάσης για τη δυνατότητα των νέων τεχνολογιών να επιφέρουν θετικές αλλαγές, αξίζει να καταγραφούν οι προβληματισμοί που διατυπώνουν οι θετικά προσκείμενοι στα ΜΔ, Mayer-Schönberger και Cukier. Υποστηρίζουν ότι μέσα στο γενικότερο

κλίμα ενθουσιασμού²⁰, οι άνθρωποι είναι πιο επιρρεπείς στη "δικτατορία των δεδομένων", εξαιτίας της οποίας εμπιστεύονται τα αποτελέσματα των αναλύσεων ακόμη και όταν υπάρχουν ενδείξεις για το αντίθετο, συλλέγουν μανιωδώς δεδομένα χωρίς σκοπό, ή συχνά παραβλέπουν ότι αυτό που αποσκοπούν να ποσοτικοποιήσουν δεν αποτυπώνεται τελικά στα συλλεχθέντα δεδομένα (Mayer-Schönberger and Cukier, 2013: 166).

Παράλληλα, αναγνωρίζουν ότι οι χρήστες θα έχουν περιορισμένες δυνατότητες ελέγχου της χρήσης των δεδομένων τους και ότι η ιδιωτικότητα θα βρίσκεται υπό μεγάλο κίνδυνο -μεγαλύτερο από το παρελθόν-, διότι (1) η αξία της πληροφορίας των ΜΔ έγκειται στις δευτερογενείς χρήσεις τους, οι οποίες συχνά δεν είναι προσδιορισμένες κατά τον χρόνο αρχικής συλλογής των δεδομένων, και (2) ακόμη και σε περίπτωση αφαίρεσης αναγνωριστικών (identifiers) από μεμονωμένα σύνολα δεδομένων για την ανάλυσή τους προκειμένου να μην θιγεί η ιδιωτικότητα των χρηστών, είναι δυνατή η επαναταυτοποίηση μέσω του συγκερασμού επιμέρους διαθέσιμων συνόλων δεδομένων (Mayer-Schönberger and Cukier, 2013: 153-154).

Προβλέπουν δε την ανάδυση μίας αγοράς προσωπικών δεδομένων, στην οποία θα δραστηριοποιηθούν μεσάζοντες που θα παρέχουν στους καταναλωτές τη δυνατότητα να κατοχυρώνουν τα δεδομένα που παράγουν, αυτοματοποιώντας τις διαδικασίες συναλλαγής-πώλησής τους σε οργανισμούς που ενδιαφέρονται να τα αποκτήσουν (Mayer-Schönberger and Cukier, 2013: 147-148). Παρατηρείται δηλαδή αντιστοιχία με το δεύτερο σκέλος της κριτικής του Morozon, που αναφέρθηκε προηγουμένως, σχετικά με τον έλεγχο των ατόμων επί των δεδομένων που παράγουν.

Το σκάνδαλο των μαζικών παρακολουθήσεων από την κυβέρνηση των ΗΠΑ που αποκαλύφθηκε από τον Edward Snowden, επιβεβαίωσε τους φόβους και

²⁰ Οι Boyd and Crawford (2012: 663) διακρίνουν μία μυθολογική διάσταση του φαινομένου ΜΔ, η οποία συνίσταται στην "ευρέως διαδεδομένη πεποίθηση ότι μεγάλα σύνολα δεδομένων παρέχουν μια ανώτερη μορφή ευφυΐας και γνώσης που μπορεί να παράγει συμπεράσματα (...), με την αύρα της αλήθειας, της αντικειμενικότητας και της ακρίβειας".

προκάλεσε δημόσιο διάλογο διεθνώς (Guardian, 2013). Προσχηματικές ή μη, καταγράφονται πρωτοβουλίες σε θεσμικό επίπεδο, τόσο στην Ευρώπη όσο και στις ΗΠΑ. Για την αποκατάσταση του πλήγματος στην εμπιστοσύνη των πολιτών απέναντι στις κυβερνήσεις, η Ευρωπαϊκή Ένωση επεξεργάζεται ένα σύμφωνο για την προστασία των δεδομένων στην Ευρώπη, το οποίο μεταξύ άλλων θα προβλέπει ότι *"η συλλογή δεδομένων θα πρέπει να έχει συγκεκριμένους στόχους και να περιορίζεται σε ό,τι αφορά τους στόχους αυτούς"*, καθώς και ότι *"οι κανόνες προστασίας των δεδομένων θα πρέπει να εφαρμόζονται ασχέτως ιθαγένειας"* (Ευρωπαϊκή Επιτροπή, 2014).

Παράλληλα, στις ΗΠΑ, στην τελική έκθεση της ειδικής ομάδας εργασίας για την εξέταση ζητημάτων σχετικά με τα ΜΔ²¹, επισημαίνεται ότι επειδή τα ΜΔ ενδέχεται να χρησιμοποιηθούν εκούσια ή ακούσια με τρόπο που να πλήττουν το κοινωνικό σύνολο, *"η κοινωνία πρέπει να λάβει μέτρα για την προστασία από αυτούς τους πιθανούς κινδύνους, διασφαλίζοντας ότι η ισχύς είναι κατάλληλα εξισοροποιημένη μεταξύ ατόμων και θεσμών· πολιτών και κυβερνήσεων, καταναλωτών και επιχειρήσεων, ή εργαζομένων και εργοδοτών"* (Executive Office of the President, 2014: 59). Συμπεριλαμβάνει δε παραίνεση για την προστασία των προσωπικών δεδομένων, τόσο εντός των ΗΠΑ, όσο και μέσω της θέσπισης διακρατικών κανόνων προστασίας της ιδιωτικότητας (Executive Office of the President, 2014: 59, 63).

2.5 Ανάλυση ΜΔ

Σε ό,τι αφορά τις σύγχρονες εξειδικευμένες λύσεις επεξεργασίας και ανάλυσης που λαμβάνουν υπ' όψιν τα ιδιαίτερα χαρακτηριστικά των ΜΔ, ο Davenport (2014: 114) θεωρεί ως σημαντικότερες τις ακόλουθες:

- το λογισμικό ανοικτού κώδικα *Hadoop*, που επιτρέπει την αποθήκευση και επεξεργασία δεδομένων χρησιμοποιώντας κατανεμημένους

²¹ Συστάθηκε το 2014 από τον Πρόεδρο Ομπάμα για την εξέταση ζητημάτων σχετικά με τα ΜΔ και την ιδιωτικότητα (The White House, 2014).

υπολογιστικούς πόρους

- το πλαίσιο *MapReduce*, στο οποίο βασίζεται το *Hadoop*
- γλώσσες προγραμματισμού (Scripting languages), όπως η *Python*, για τον χειρισμό και την ανάλυση των δεδομένων
- τη Μηχανική Μάθηση (Machine Learning), που βασίζεται στην Τεχνητή Νοημοσύνη²²
- την Οπτική Ανάλυση Δεδομένων (Visual Analytics), για την οπτική απεικόνιση του αποτελέσματος της ανάλυσης
- την Επεξεργασία Φυσικής Γλώσσας (Natural Language Processing), που επιτρέπει την ανάλυση κειμένων
- την ανάλυση στη μνήμη του υπολογιστικού συστήματος (In-memory analytics), που επιταχύνει τη διαδικασία επεξεργασίας.

Για τους σκοπούς της παρούσας εργασίας, δεν κρίνεται σκόπιμη η περαιτέρω επέκταση στα ιδιαίτερα χαρακτηριστικά κάθε μιας από τις προαναφερθείσες λύσεις, πλην της Οπτικής Ανάλυσης Δεδομένων που εξετάζεται στο Κεφάλαιο 3 (Ενότητα 3.5) πιο κάτω.²³

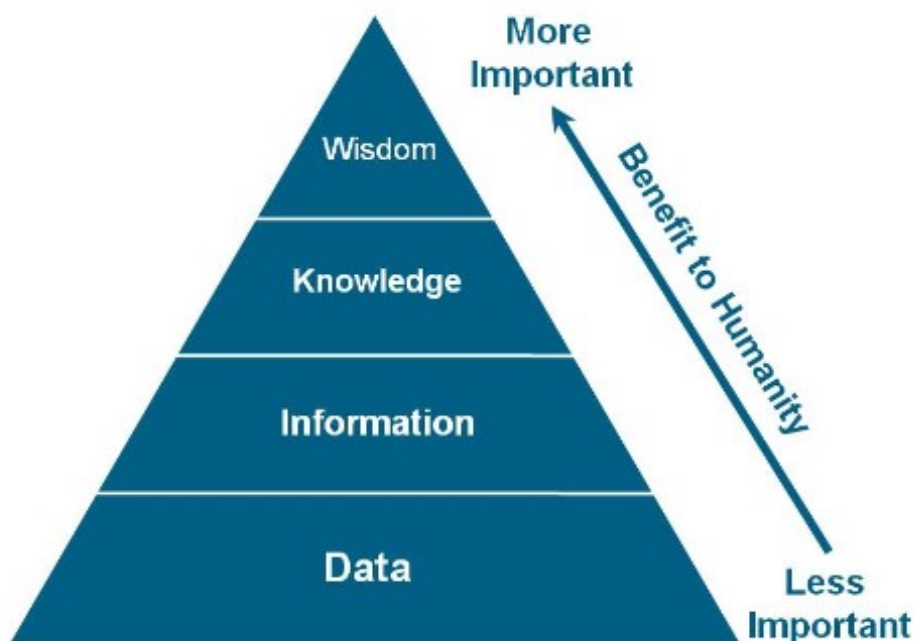
Η βαθιά εδραιωμένη πεποίθηση ότι η ανάλυση δεδομένων συνδέεται με τη λήψη κατάλληλων αποφάσεων (Davenport, 2014: 10), μπορεί να αιτιολογήσει τον ενθουσιασμό για την υπερσυγκέντρωση δεδομένων από πολλαπλές πηγές. Ο Evans (2011: 6)²⁴ αποτυπώνει σχηματικά τη διαδικασία επεξεργασίας δεδομένων που οδηγεί στην απόκτηση γνώσης ως πυραμίδα, με τα δεδομένα να βρίσκονται

²² Σύμφωνα με τον Hack (2013 σε Simon, 2013: 89), η Μηχανική Μάθηση διευκολύνει την εξέταση ενός συνόλου δεδομένων ακόμη και όταν δεν υπάρχουν συγκεκριμένα ερωτήματα προς διερεύνηση.

²³ Ωστόσο, για λόγους πληρότητας αξίζει να επισημανθεί ότι (1) οι τεχνολογίες που διευκολύνουν την ανάλυση ΜΔ προκρίνουν την χρήση κατανεμημένων δικτύων υπολογιστικών πόρων (Intel IT Center, 2012: 2), καθώς και ότι (2) η στατιστική, η Μηχανική Μάθηση και η Επεξεργασία Φυσικής Γλώσσας μπορούν να συνδυαστούν για την αυτοματοποίηση του εντοπισμού τάσεων και μοτίβων από μεγάλα σύνολα δομημένων και μη δομημένων δεδομένων (SAS, 2013: 10).

²⁴ Συντάκτης αναφοράς της Cisco, σχετικά με τις αλλαγές που εκτιμάται ότι θα επιφέρει το Διαδίκτυο των Πραγμάτων.

στη βάση της, όπως φαίνεται στην ακόλουθη εικόνα (Εικόνα 4):



Εικόνα 4: Άντληση σοφίας από τα δεδομένα
(Πηγή: Evans, 2011: 6)

Αν και ο ισχυρισμός του ότι όσο περισσότερα δεδομένα παράγονται, τόσο μεγαλύτερη γνώση και σοφία προκύπτει είναι μάλλον απλοϊκός, καθώς δεν λαμβάνει υπ' όψιν του την ποιότητα των δεδομένων, εντούτοις μπορούμε να δεχθούμε ότι η ροή που περιγράφει από τη συγκέντρωση των δεδομένων, στην επεξεργασία τους που παράγει πληροφορία, η οποία με τη σειρά της συνεισφέρει στην απόκτηση γνώσης, σε γενικές γραμμές αποτυπώνει τον τρόπο άντλησης γνώσης από τα ΜΔ.

Επί της ουσίας, πρόκειται για την ιεραρχική δομή *DIKW* (Data-Information-Knowledge-Wisdom), που αποτυπώνει την αλληλεξάρτηση μεταξύ των δεδομένων, της πληροφορίας, της γνώσης και της σοφίας, καθώς και τη συνολική διαδικασία μετατροπής των δεδομένων σε σοφία. Η ιεραρχία αυτή βασίζεται στην "υπονοούμενη παραδοχή (...) ότι τα δεδομένα μπορούν να χρησιμοποιηθούν για να παραχθεί πληροφορία, η πληροφορία μπορεί να χρησιμοποιηθεί για να παραχθεί γνώση, και η γνώση μπορεί να χρησιμοποιηθεί για να παραχθεί σοφία"

(Rowley, 2007: 164).

Όπως επισημαίνει η Rowley (2007: 166), η σύγχρονη βιβλιογραφία, στο πλαίσιο της εξέτασης της εν λόγω δομής υπό το πρίσμα των πληροφοριακών συστημάτων, επικαλείται τους ορισμούς που έχουν διατυπωθεί από τον Ackoff για *"την αρχική περιγραφή των τύπων στην ιεραρχία ως περιεχόμενο του ανθρώπινου νου"*. Σύμφωνα με αυτόν, η σοφία βρίσκεται στην κορυφή της ιεραρχίας, με τις υπόλοιπες έννοιες (κατανόηση, γνώση, πληροφορία και δεδομένα) να προηγούνται ως στάδια· το χαμηλότερο στάδιο αποτελεί την προϋπόθεση για μετάβαση στο επόμενο (Ackoff, 1989: 3 σε Rowley, 2007: 166).

Επομένως, η ιεραρχική δομή θα μπορούσε να απεικονιστεί ως ακολούθως (Εικόνα 5):



Εικόνα 5: Ιεραρχική δομή νοητικής διεργασίας για παραγωγή σοφίας από τα δεδομένα, σύμφωνα με Ackoff (1989: 3 σε Rowley, 2007: 166)

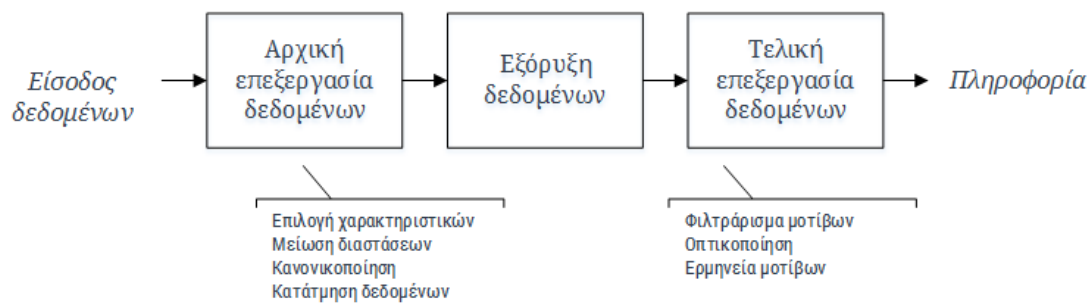
Δεδομένου ότι η κατανόηση ως στάδιο τείνει να εκλείψει από μεταγενέστερες αναλύσεις της ιεραρχικής δομής (Rowley, 2007: 166-167), στον ακόλουθο πίνακα (Πίνακας 3) καταγράφονται τα υπόλοιπα επιμέρους τμήματα και τα βασικά χαρακτηριστικά τους, σύμφωνα με τον Ackoff (1989 σε Rowley, 2007: 166):

Πίνακας 3: Ορισμοί και χαρακτηριστικά, σύμφωνα με Ackoff (1989 σε Rowley, 2007: 166)

Όρος (από το χαμηλότερο στο υψηλότερο επίπεδο)	Ορισμός και βασικά χαρακτηριστικά
Δεδομένα (Data)	• σύμβολα που αναπαριστούν ιδιότητες αντικειμένων, γεγονότων και του περιβάλλοντος αυτών • προϊόντα παρατήρησης • μη αξιοποιήσιμη μορφή
Πληροφορία (Information)	• απαντήσεις σε ερωτήματα του τύπου “ποιος;”, “τι;”, “πότε;”, “πόσο;” • τα πληροφοριακά συστήματα παράγουν, αποθηκεύουν, αντλούν και επεξεργάζονται δεδομένα • η πληροφορία συνάγεται από τα δεδομένα
Γνώση (Knowledge)	• εξειδικευμένες γνώσεις και μέθοδοι που μπορούν να αξιοποιηθούν σε έναν τομέα²⁵, καθώς και ό,τι καθιστά εφικτή τη μετατροπή της πληροφορίας σε οδηγίες • αποκτάται από τη μετάδοσή της από κάποιο άτομο που ήδη την κατέχει, μέσω κατάρτισης ή μέσω εμπειρίας
Σοφία (Wisdom)	• η ικανότητα αύξησης της αποτελεσματικότητας • προσθέτει αξία, η οποία προϋποθέτει τη νοητική λειτουργία που αποκαλείται κρίση

Η Ανακάλυψη Γνώσης από Βάσεις Δεδομένων (Knowledge Discovery in Databases) συνίσταται στη συνολική διαδικασία μετατροπής ανεπεξέργαστων δεδομένων σε χρήσιμη πληροφορία, και περιλαμβάνει στάδια όπως η συλλογή των δεδομένων, η αρχική επεξεργασία τους, η εξόρυξη δεδομένων και η τελική επεξεργασία τους, όπως φαίνεται και στην ακόλουθη εικόνα (Εικόνα 6):

²⁵ Περιφραστική απόδοση του αγγλικού όρου “know-how”, χρησιμοποιώντας τον ορισμό του όρου “τεχνογνωσία” (Μπαμπινιώτης, 2011).



Εικόνα 6: Ανακάλυψη Γνώσης από Βάσεις Δεδομένων
(Πηγή: Tan et al., 2014: 3)

Κεντρικό ρόλο στις δύο τελευταίες διεργασίες διαδραματίζει η οπτικοποίηση δεδομένων (Tan et al., 2014: 4, 105), η οποία, ως μέθοδος δημιουργίας αναπαραστάσεων σε αισθητηριακά αντιληπτή μορφή, επιτρέπει τη διαμόρφωση μίας γρήγορης άποψης σχετικά με τα δεδομένα που αναλύονται ή παρουσιάζονται αντιστοίχως, ενώ καθιστά ορατά σχέσεις και μοτίβα που δεν είναι άμεσα αντιληπτά.

Σύμφωνα με τους Hand et al. (2001: 1), η εξόρυξη δεδομένων (data mining) συνίσταται στην *"ανάλυση (συχνά μεγάλων) συνόλων δεδομένων παρατήρησης για τον εντοπισμό μη αναμενόμενων σχέσεων και για την σύνοψη των δεδομένων με νέους τρόπους, που κατανοούνται και χρησιμεύουν σε αυτόν που τα κατέχει [τα δεδομένα]"*. Τα αποτελέσματά της αναφέρονται ως *"μοντέλα"* ή *"μοτίβα"* (Hand et al., 2001: 1). Πρόκειται για ένα διεπιστημονικό πεδίο που αντλεί κυρίως από τη στατιστική, την Τεχνητή Νοημοσύνη, τη Μηχανική Μάθηση και την αναγνώριση μοτίβων (Tan et al. 2014: 6).

Οι βασικές εργασίες που διενεργούνται στο πλαίσιο της εξόρυξης δεδομένων έχουν είτε προγνωστικό είτε περιγραφικό χαρακτήρα (Fayyad et al., 1996: 44 και Tan et al., 2014: 7), με τις βασικότερες να είναι οι ακόλουθες (Tan et al. 2014, 8-11):

- Τα προγνωστικά μοντέλα (predictive modelling) στοχεύουν στην πρόβλεψη της τιμής μίας εξαρτημένης μεταβλητής βάσει των τιμών ανεξάρτητων μεταβλητών. Η κατηγοριοποίηση (classification) και η παλινδρόμηση (regression), για διακριτές και για συνεχείς εξαρτημένες

μεταβλητές αντιστοίχως, είναι δύο τύποι προγνωστικών μοντέλων που στοχεύουν στην ελαχιστοποίηση του σφάλματος στην πρόβλεψη της μεταβλητής και των πραγματικών τιμών.

- Η ανάλυση σχέσεων (association analysis) αποκαλύπτει μοτίβα που περιγράφουν τους ισχυρούς δεσμούς μεταξύ των χαρακτηριστικών των δεδομένων.
- Η ανάλυση συστάδων (cluster analysis) εντοπίζει ομάδες παρατηρήσεων, με τις μεμονωμένες παρατηρήσεις που απαρτίζουν μία συστάδα να έχουν περισσότερες ομοιότητες μεταξύ τους από ό,τι θα είχαν με παρατηρήσεις που απαρτίζουν άλλες συστάδες.
- Μέσω της ανίχνευσης ανωμαλιών (anomaly detection) εντοπίζονται παρατηρήσεις που διαφέρουν σημαντικά σε σχέση με τις υπόλοιπες, προκειμένου να διαπιστωθεί αν πρόκειται όντως για έγκυρες ή εσφαλμένες τιμές.

Η τάση στα σύγχρονα συστήματα επιχειρηματικής ευφυΐας (business intelligence) που διευκολύνουν την ανακάλυψη γνώσης από τα δεδομένα είναι η παροχή δυνατότητας οπτικοποίησης ως βασική λειτουργικότητα, που επιτρέπει στους χρήστες *"να διεξάγουν αναλύσεις πρόβλεψης σε πραγματικό χρόνο και να παρουσιάζουν τα αποτελέσματα σε εντυπωσιακές, αλληλεπιδραστικές και εύκολα κατανοητές οπτικές απεικονίσεις"* (Intel IT Center, 2013: 3).

2.6 Νέα πεδία επαγγελματικής εξειδίκευσης

Η έκρηξη του όγκου δεδομένων και η ευρεία διαθεσιμότητά τους έχουν αναδείξει νέα πεδία επαγγελματικής εξειδίκευσης, όπως οι επιστημονικοί υπεύθυνοι δεδομένων (data scientists) και η δημοσιογραφία δεδομένων (data journalism).

2.6.1 Επιστήμη δεδομένων

Ο επιστημονικός υπεύθυνος δεδομένων, όρος που διατυπώθηκε για πρώτη

φορά το 2008 από τους Patil και Hammerbacher (Davenport and Patil, 2012), εξειδικεύεται στην ερμηνεία ΜΔ και χαρακτηρίζεται από υψηλό βαθμό διεπιστημονικότητας, καθώς συνδυάζει "γνώσεις επιστήμης υπολογιστών, εξόρυξης δεδομένων, αλγορίθμων, στατιστικών, μαθηματικών, οπτικοποίησης δεδομένων και αντίληψης του οργανισμού εντός του οποίου εργάζεται και των στόχων αυτού, προκειμένου να παρέχει την αντίστοιχη γνώση στους κατάλληλους αποδέκτες" (Big Data Insight Group, 2012).

Στο δημοφιλές κείμενο "*Data Scientist: The Sexiest Job of the 21st Century*", οι Davenport και Patil (2012) επισημαίνουν ότι ο επιστημονικός υπεύθυνος δεδομένων, εκτός από την ικανότητα να προσδίδει δομή σε μη δομημένα δεδομένα, να τα συνδυάζει με άλλα και γενικότερα να εμπλέκεται σε έναν συνεχή διάλογο με τα δεδομένα, χαρακτηρίζεται επιπλέον από ιδιαίτερες δεξιότητες αφήγησης με τα δεδομένα, λεκτικής ή/και οπτικής, αλλά και από συνδυαστική σκέψη.

Ωστόσο επισημαίνουν ότι το μεγαλύτερο εμπόδιο στην αξιοποίηση ΜΔ από τις επιχειρήσεις αποτελεί η απουσία εξειδικευμένων ακαδημαϊκών προγραμμάτων στην επιστήμη δεδομένων και η επακόλουθη έλλειψη καταρτισμένων επιστημονικών υπευθύνων δεδομένων. Σύμφωνα μάλιστα με τον Chan (2012 σε Big Data Insight Group, 2012)²⁶, η έλλειψη αυτή συνδέεται και με τον τρόπο διδασκαλίας των μαθηματικών και των επιστημονικών μαθημάτων κατά τις τελευταίες δύο δεκαετίες.

Κατά την άποψη του Ohlhorst (2013: 30), "*το σημαντικότερο προσόν*" που θα πρέπει να διαθέτει ένας επιστημονικός υπεύθυνος δεδομένων είναι "*η ικανότητα να μεταφράζει τη σημασία των δεδομένων με τρόπο εύκολα κατανοητό από άλλους*", χωρίς όμως να δίνει περισσότερες λεπτομέρειες. Σημειώνει ωστόσο ότι η δυνητική δεξαμενή υποψηφίων για τη στελέχωση θέσεων που αφορούν στην ανάλυση ΜΔ απαιτεί ένα "*υβριδικό ταλέντο*", το οποίο δεν συναντάται απαραίτητα μόνο σε άτομα με εξειδίκευση στην Πληροφορική: άτομα "*με*

²⁶ Διευθυντής του Κέντρου για την Ηγεσία Πληροφορίας του Πανεπιστημίου City στο Λονδίνο

κατανόηση της αξίας των δεδομένων και της ιδεολογίας του πώς να τα ερμηνεύουν" μπορούν να ενταχθούν σε ομάδα ανάλυσης ΜΔ εφόσον διαθέτουν θεμελιώδεις γνώσεις (1) εξόρυξης, (2) οπτικοποίησης, (3) ανάλυσης, (4) χειρισμού δεδομένων, καθώς και (5) ανακάλυψης από τα δεδομένα (Ohlhorst, 2013: 33).

Σύμφωνα δε με τους Rios και Lin (2012: 25), οι οποίοι υποστηρίζουν ότι όσοι ασχολούνται με οπτικοποίηση δεδομένων πρέπει να διαθέτουν δεξιότητες χειρισμού των μη επεξεργασμένων δεδομένων, ο ιδανικός επιστημονικός υπεύθυνος δεδομένων πρέπει να λειτουργεί *"ως μηχανικός και ως καλλιτέχνης"*. Επικαλούμενοι την εμπειρία τους ως αναλυτές δεδομένων στο *Twitter*, αναγνωρίζουν ότι άτομα με αυτόν τον συνδυασμό ικανοτήτων είναι δυσεύρετα, αλλά αποτελούν σημαντικό πλεονέκτημα για τους οργανισμούς που έρχονται αντιμέτωποι με τις προκλήσεις που θέτουν τα ΜΔ (Rios and Lin, 2012: 25).

2.6.2 Δημοσιογραφία δεδομένων

Η διάθεση και η διαρροή δεδομένων που αφορούν σε δραστηριότητες κρατικών θεσμών οδήγησε στην άνθιση της δημοσιογραφίας δεδομένων (Rogers, 2013a: 30), η οποία συνίσταται στην αξιοποίηση δεδομένων και στην ενσωμάτωσή τους στην παραγωγή ειδησεογραφικού περιεχομένου. Θα μπορούσε να θεωρηθεί ότι αποτελεί προέκταση της παραδοσιακής δημοσιογραφίας, η οποία αντιμετωπίζει νέες προκλήσεις λόγω του όγκου και της ποικιλομορφίας της πληροφορίας που είναι διαθέσιμη. Όπως επισημαίνει ο Bradshaw (2012 σε Gray et al., 2012: 3), *"τα δεδομένα μπορούν είτε να είναι η πηγή της δημοσιογραφίας δεδομένων είτε το εργαλείο μέσω του οποίου λέγεται μια ιστορία -ή συνδυασμός και των δύο"*.

Ο Sir Tim Berners-Lee (2012 σε Gray et al., 2012: 6) υποστηρίζει ότι *"η δημοσιογραφία που βασίζεται στα δεδομένα είναι το μέλλον"* και για τον λόγο αυτόν οι δημοσιογράφοι θα πρέπει να διαθέτουν τα κατάλληλα εφόδια ανάλυσης των δεδομένων και επισήμανσης των σημαντικών ζητημάτων, εντάσσοντας τα σε ένα ευρύτερο πλαίσιο. Επομένως, θα μπορούσε να υποστηριχθεί ότι οι

δημοσιογράφοι δεδομένων θα πρέπει να αναπτύξουν και ικανότητες αντίστοιχες με αυτές των επιστημονικών υπεύθυνων δεδομένων. Οι δεξιότητες αυτές είναι απαραίτητες προκειμένου να είναι σε θέση *"να παρέχουν το νοηματικό πλαίσιο, τη σαφήνεια και ίσως την αλήθεια στο συνεχώς αυξανόμενο ψηφιακό περιεχόμενο"*, με στόχο *"την εξαγωγή νοήματος κυρίως από τα μη δομημένα δεδομένα"* (Howard, 2012 σε Gray et al., 2012: 8-9).

Ο Lorenz (2012: 4) υποστηρίζει ότι *"οι δημοσιογράφοι πρέπει να αντιμετωπίσουν τα δεδομένα ως ευκαιρία"*. Επισημαίνει μάλιστα ότι οι δημοσιογράφοι δεδομένων, όπως και οι επιστημονικοί υπεύθυνοι δεδομένων, είναι περιζήτητοι ως επαγγελματίες που γνωρίζουν πώς να επεξεργάζονται δεδομένα και *"να τα μετατρέπουν σε κάτι απτό"* (Lorenz, 2012: 4). Σημειώνει όμως ότι, προκειμένου να είναι δυνατή η μετατροπή των δεδομένων σε χρήσιμη πληροφορία, οι δημοσιογράφοι πρέπει να αναζητήσουν ευκαιρίες για να αποκτήσουν δεξιότητες και να εξοικειωθούν με όλα τα στάδια της διαχείρισης των δεδομένων.

Ποιες όμως είναι οι απαραίτητες επιπλέον γνώσεις που πρέπει να διαθέτουν οι δημοσιογράφοι δεδομένων; Ο Burn-Murdoch (2013: 154-155) θεωρεί απαραίτητη τη βασική γνώση αρχών στατιστικής, ωστόσο επισημαίνει ότι η άποψη αυτή δεν είναι καθολικώς αποδεκτή εντός των δημοσιογραφικών οργανισμών, καθώς *"οι στατιστικές μέθοδοι θεωρούνται σκοτεινές τέχνες και η επάρκεια στη χρήση υπολογιστικών φύλλων και στα εργαλεία οπτικοποίησης (...) αντιμετωπίζεται με καχυποψία"*.

Τα έργα δημοσιογραφίας δεδομένων συχνά παρουσιάζονται με αλληλεπιδραστικές διεπαφές οπτικοποίησης, η υλοποίηση των οποίων απαιτεί τη συνδρομή προγραμματιστών. Η Hannaford (2013: 180) σημειώνει ότι στις ΗΠΑ, οι δημοσιογραφικοί οργανισμοί αναζητούν προσωπικό με *υβριδική* επαγγελματική υπόσταση, η οποία συνίσταται στον συνδυασμό

δημοσιογραφικής επάρκειας και προγραμματιστικών δεξιοτήτων.²⁷ Οι δημοσιογραφικοί οργανισμοί φαίνεται ότι καλύπτουν την ανάγκη αυτού του συνδυασμού δεξιοτήτων μέσω της δημιουργίας ομάδων που αποτελούνται από εξειδικευμένα άτομα σε διαφορετικά αντικείμενα (Hannaford, 2013: 180).

Επομένως, γίνεται αντιληπτό ότι βασικές γνώσεις προγραμματισμού είναι απαραίτητες για τους δημοσιογράφους, προκειμένου να μπορούν να επικοινωνήσουν με τους προγραμματιστές συναδέλφους τους. Όπως άλλωστε επισημαίνει η Hannaford (2013: 183), στο πλαίσιο σχεδιασμού και υλοποίησης ενός εγχειρήματος δημοσιογραφίας δεδομένων, ο ρόλος του δημοσιογράφου είναι να εποπτεύει τη συνολική διαδικασία. Αναλαμβάνοντας τον ρόλο αυτό, ο δημοσιογράφος πρέπει να είναι σε θέση να κατανοήσει τα προβλήματα που αντιμετωπίζει ο προγραμματιστής, να προσαρμόζεται στους τεχνικούς και χρονικούς περιορισμούς, και να διασφαλίζει την παροχή της κατάλληλης εμπειρίας χρήστη (Hannaford, 2013: 183).

Οι νέες δεξιότητες που πρέπει να καλλιεργήσουν οι δημοσιογράφοι εμπλουτίζουν τις παραδοσιακά απαιτούμενες. Ο δημοσιογράφος χρειάζεται αφενός να εφαρμόζει μεθόδους ανάλυσης για να *"εξάγει νόημα και δομή από την ασταμάτητη ροή δεδομένων"*, αφετέρου να μεριμνά για την παρουσίαση των δεδομένων για να αναδείξει τι είναι σημαντικό στο κοινό στο οποίο απευθύνεται (Meyer, 2012 σε Gray et al., 2012: 6). Παράλληλα, οι δυνατότητες αφήγησης εμπλουτίζονται, δεδομένου ότι η δημοσιογραφία δεδομένων *"εμπερικλείει ένα συνεχώς αυξανόμενο πλήθος εργαλείων, τεχνικών και προσεγγίσεων στην υπηρεσία της αφήγησης, [και] περιλαμβάνει οτιδήποτε από το παραδοσιακό ρεπορτάζ με υποβοήθηση ηλεκτρονικών υπολογιστών, όπου τα δεδομένα είναι η πηγή, μέχρι τις πιο σύγχρονες εφαρμογές οπτικοποίησης δεδομένων και ειδήσεων"* (Pilhofer, 2012 σε Gray et al., 2012).

²⁷ Όπως σημειώνει η Hannaford (2013: 175), δεν υπάρχει ακόμη ορολογία για αυτή την επαγγελματική ιδιότητα· ενδεικτικά αναφέρει τους όρους *"journo-coder"*, *"programmer-journalist"*, *"hacker-journalist"* και *"journo-programmer"* που χρησιμοποιούνται για να την περιγράψουν.

Στο παρόν κεφάλαιο, τα ΜΔ ορίστηκαν ως ένα πολυδιάστατο φαινόμενο που διαμορφώνεται από τις τεχνολογικές δυνατότητες που επιτρέπουν τη συλλογή, την αποθήκευση, την επεξεργασία και την ανάλυση δεδομένων, τα οποία χαρακτηρίζονται από πρωτοφανή όγκο, υψηλή ταχύτητα παραγωγής και ανανέωσης ή/και ποικιλομορφία.

Συγκεκριμένα, η δυνατότητα πρόσβασης σε κατανεμημένους υπολογιστικούς πόρους αίρει τους οικονομικούς και τεχνολογικούς περιορισμούς του παρελθόντος, και επιτρέπει τη συγκέντρωση ΜΔ από πολλαπλές πηγές και με σημαντικά χαμηλότερο κόστος για τους οργανισμούς. Ωστόσο, η αξία των ΜΔ εξαρτάται από τις δυνατότητες ανάλυσής τους και εξαγωγής συμπερασμάτων από αυτά. Επιπρόσθετα, η διάχυτη αισιοδοξία για τις ευκαιρίες που δημιουργούνται από τη συγκέντρωση ΜΔ, οδήγησε στην εμφάνιση νέων πεδίων επαγγελματικής εξειδίκευσης που απαιτούν δεξιότητες ανάλυσης και παρουσίασης της πληροφορίας που μπορεί να αντληθεί από αυτά.

Η διαδικασία μετατροπής ανεπεξέργαστων δεδομένων σε χρήσιμη πληροφορία περιλαμβάνει στάδια όπως η συλλογή των δεδομένων, η αρχική επεξεργασία τους, η εξόρυξη δεδομένων και η τελική επεξεργασία τους. Σημαντικό ρόλο στα δύο τελευταία στάδια διαδραματίζει η οπτικοποίηση δεδομένων, που παρουσιάζεται στα επόμενα κεφάλαια.

3 Οπτικοποίηση δεδομένων: γενικό πλαίσιο

Στο παρόν κεφάλαιο, παρουσιάζεται ένα γενικό πλαίσιο για την κατανόηση των παραγόντων που καθιστούν την οπτικοποίηση δεδομένων (data visualization) πεδίο έντονου ερευνητικού ενδιαφέροντος.

Μετά τις απαραίτητες εισαγωγικές επισημάνσεις, ακολουθεί η ιστορική αναδρομή για τον εντοπισμό σημαντικών γεγονότων που σχετίζονται με την οπτική αναπαράσταση δεδομένων. Στη συνέχεια, καταγράφονται βασικοί ορισμοί και επιλεγμένες οπτικές που αναδεικνύουν ενδιαφέρουσες πτυχές της οπτικοποίησης, βάσει της σχετικής βιβλιογραφίας. Τέλος, μελετάται η Οπτική Ανάλυση Δεδομένων, ως λύση ανάλυσης που αξιοποιεί την οπτικοποίηση δεδομένων για τη διερεύνηση και την άντληση γνώσης από ΜΔ.

3.1 Εισαγωγικές επισημάνσεις

Μέσω της όρασης λαμβάνονται περισσότερες πληροφορίες απ' ό,τι μέσω όλων των υπόλοιπων αισθήσεων· αυτές οι πληροφορίες αναλύονται από τους 20 δισεκατομμύρια εγκεφαλικούς νευρώνες που είναι αφιερωμένοι σε αυτό τον σκοπό και παρέχουν τον μηχανισμό εντοπισμού μοτίβων, στον οποίον βασίζεται το μεγαλύτερο μέρος της ανθρώπινης γνωστικής δραστηριότητας (Ware, 2013: 2). Επομένως, θα μπορούσε να θεωρηθεί ότι η οπτική αναπαράσταση δεδομένων, ως πρακτική εφαρμογή των ανωτέρω, αξιοποιεί τις δυνατότητες του ανθρώπινου οπτικού συστήματος για να διευκολύνει τον άμεσο εντοπισμό σχέσεων και μοτίβων σε αφηρημένα δεδομένα. Ως αποτέλεσμα, είναι δυνατή η γρήγορη ερμηνεία των δεδομένων που αναπαρίστανται και η άντληση γνώσης από αυτά.

Βάσει των προαναφερθέντων, καθίσταται αντιληπτή η σύνδεση της οπτικοποίησης με το φαινόμενο ΜΔ. Συγκεκριμένα, η αξία που εμπεριέχεται στα

ΜΔ εξαρτάται εν πολλοίς από τις δυνατότητες ανάλυσης και άντλησης πληροφορίας από αυτά. Η αναγνώριση της οπτικοποίησης ως μέσου άντλησης αξίας από τα ΜΔ²⁸, συνδέεται με τον σημαντικό ρόλο που διαδραματίζει στη διαδικασία μετατροπής των ανεπεξέργαστων δεδομένων σε χρήσιμη πληροφορία²⁹.

Παράλληλα, η αύξηση των δημοσιεύσεων και των συνεδρίων σχετικά με την οπτικοποίηση δεδομένων που παρατηρείται τα τελευταία χρόνια (Lindquist, 2011: 4-5), συνιστά ισχυρή ένδειξη έντονης ερευνητικής δραστηριότητας στο εν λόγω αντικείμενο. Χαρακτηρίζεται από υψηλό βαθμό διεπιστημονικότητας, καθώς με αυτό ασχολούνται ερευνητές με υπόβαθρο στις θετικές, τεχνολογικές και ανθρωπιστικές επιστήμες, καθώς και καλλιτέχνες. Αυτή η ποικιλία επιστημονικών υποβάθρων οδηγεί σε διαφορετικές αντιμετώπισεις και ορισμούς της οπτικοποίησης δεδομένων (Yau, 2013: 89), όπως επιβεβαιώνεται στη συνέχεια (Ενότητες 3.3 και 3.4).

Ωστόσο, η οπτικοποίηση δεδομένων δεν συνιστά πρωτοεμφανιζόμενο κλάδο, όπως προκύπτει από την ιστορική αναδρομή που ακολουθεί.

3.2 Ιστορική αναδρομή

Ο Friendly (2008a), στο κείμενο *"A Brief History of Data Visualization"*, καταγράφει τα ορόσημα στην εξέλιξη της οπτικοποίησης δεδομένων και επισημαίνει ότι *"οι εξελίξεις στις τεχνολογίες (εκτύπωση, αναπαραγωγή), στη μαθηματική θεωρία και πράξη, καθώς και στην εμπειρική παρατήρηση και καταγραφή επέτρεψαν την ευρύτερη χρήση γραφημάτων και την πρόοδο στις μορφές και στο περιεχόμενο"* (Friendly, 2008a: 16).

Οι φάσεις εξέλιξης που εντοπίζει, από πρώιμες χαρτογραφικές αναπαραστάσεις έως τα περίπλοκα γραφήματα της σύγχρονης περιόδου που

²⁸ Βλ. Intel IT Center, 2013 και Trendviz, 2014.

²⁹ Όπως έχει αναφερθεί στην Ενότητα 2.5 πιο πάνω.

χαρακτηρίζονται από υψηλό βαθμό αλληλεπίδρασης με τα δεδομένα, απεικονίζονται σχηματικά στην ακόλουθη εικόνα (Εικόνα 7):



Εικόνα 7: Σχηματική απεικόνιση φάσεων εξέλιξης της οπτικοποίησης δεδομένων που εντοπίζει ο Friendly (2008a)

Πράγματι, από την ιστορική του καταγραφή προκύπτει ότι η οπτικοποίηση δεδομένων εξελίχθηκε παράλληλα με την αύξηση των διαθέσιμων δεδομένων και κατ' επέκταση με την ανάγκη επεξεργασίας και ερμηνείας τους για την επίλυση προβλημάτων και την αιτιολόγηση φαινομένων. Όπως άλλωστε

επισημαίνει και ο Spence (2007: 6-7), η σημαντική πρόοδος που συντελέστηκε κατά την τελευταία εικοσαετία στο πεδίο της οπτικοποίησης της πληροφορίας, οφείλεται κατά βάση στις τεχνολογικές εξελίξεις, με σημαντικότερες: (1) τη συνεχή μείωση του κόστους αποθήκευσης δεδομένων και τη συνακόλουθη ραγδαία αύξηση του όγκου τους, (2) την αύξηση της υπολογιστικής ισχύος, που επιτρέπει την γρήγορη επιλογή και διερεύνηση δεδομένων, και (3) τη διαθεσιμότητα απεικονιστικών συσκευών υψηλής ανάλυσης.

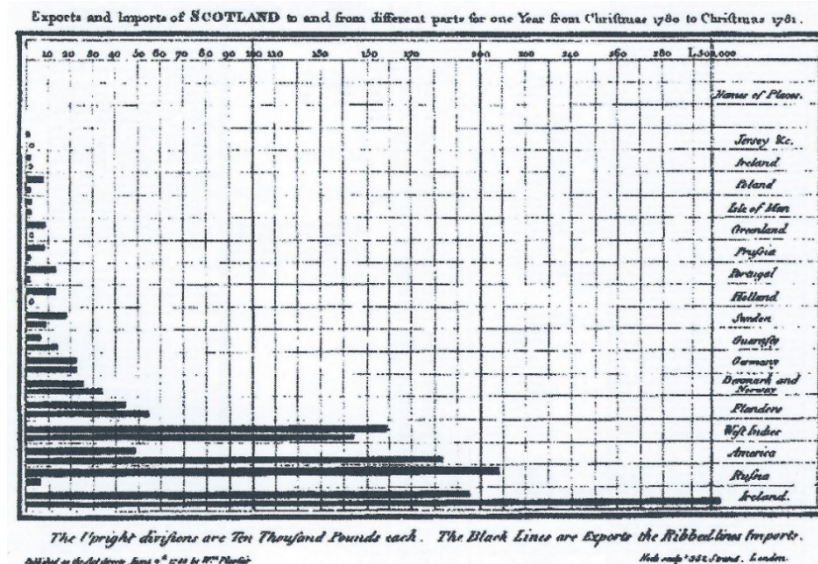
Στη συνέχεια, αναπαράγονται συνοπτικά τα ορόσημα που εντοπίζονται από τον Friendly (2008a). Όπου κρίνεται σκόπιμο, η καταγραφή του εμπλουτίζεται με επιπλέον πηγές.

3.2.1 Προ ηλεκτρονικών υπολογιστών

Η καταγραφή του Friendly ξεκινά από τον 17^ο αιώνα. Μέχρι τότε, οι οπτικές αναπαραστάσεις περιορίζονταν σε πίνακες, στους οποίους καταγράφονταν οι θέσεις των ουράνιων σωμάτων, και σε χάρτες, που ήταν απαραίτητοι για τις εξερευνήσεις (Friendly, 2008a: 18). Τον 17^ο αιώνα, οι συνθήκες ήταν πια ώριμες για την ανάπτυξη μεθόδων γραφικής αναπαράστασης: η ανάπτυξη εργαλείων παρατήρησης και μέτρησης καθιστά εφικτή την καταγραφή ποσοτικών δεδομένων που προσέλκυαν ενδιαφέρον, ενώ παράλληλα διατυπώνονται θεωρίες για την ανάλυσή τους και ιδέες για την οπτική τους αναπαράσταση (Friendly, 2008a: 19-22).

Κατά τη διάρκεια του 18^{ου} αιώνα, εντοπίζονται πειραματισμοί σχετικά με τις μορφές γραφικής αναπαράστασης: για παράδειγμα, παρατηρούνται οι πρώτες απόπειρες εμπλουτισμού γεωγραφικών χαρτών με γεωλογικά, οικονομικά και ιατρικά δεδομένα (Friendly, 2008a: 22). Παράλληλα, αναπτύσσονται τεχνολογικές καινοτομίες που διευκολύνουν την αναπαραγωγή των γραφικών παραστάσεων, όπως η τρίχρωμη εκτύπωση και η λιθογραφία (Friendly, 2008a: 22). Επίσης, πιο αφηρημένες γραφικές αναπαραστάσεις εμφανίζονται -ο William Playfair συνιστά εμβληματική μορφή, καθώς επινοεί μία σειρά γραφικών

αναπαραστάσεων που χρησιμοποιούνται ακόμη και σήμερα (Friendly, 2008a: 23-24)³⁰. Στην ακόλουθη εικόνα (Εικόνα 8), παρατίθεται γράφημα από το βιβλίο του Playfair με τίτλο *"The Commercial and Political Atlas"*³¹, το οποίο εκτιμάται ότι περιλαμβάνει το πρώτο γράφημα ράβδων (Rogers, 2013a: 62):



Εικόνα 8: Γράφημα ράβδων του Playfair που απεικονίζει τις εξαγωγές και τις εισαγωγές της Σκωτίας (Πηγή: Rogers, 2013a: 62)

Η ευρεία διάδοση των γραφικών αναπαραστάσεων δεδομένων καθυστερεί για λίγες ακόμη δεκαετίες, διότι η ποικιλία και η περιπλοκότητα των μέχρι τότε διαθέσιμων ποσοτικών δεδομένων δεν επιτάσσουν την ανάπτυξη νέων μεθόδων οπτικοποίησης (Friendly, 2008a: 25). Στο σημείο αυτό, αξίζει να σημειωθεί ότι ο Manovich (2010: 5-8) εντοπίζει τις απαρχές της οπτικοποίησης πληροφορίας στο δεύτερο μισό του 18^{ου} αιώνα, υποστηρίζοντας ότι βασίζεται σε δύο αρχές που την χαρακτηρίζουν έως και σήμερα: τη μείωση (reduction) και τις χωρικές μεταβλητές (spatial variables).

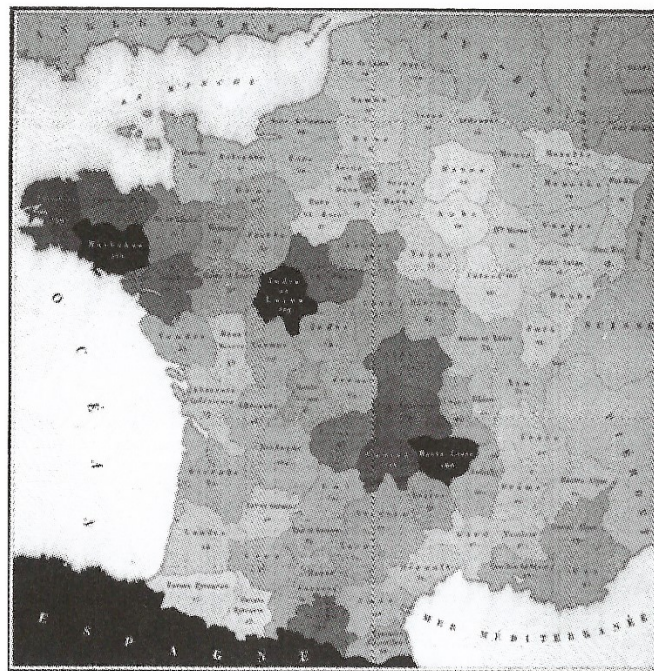
Κατά το πρώτο μισό του 19^{ου} αιώνα, καταγράφονται σημαντικές εξελίξεις (Friendly, 2008a: 25), καθώς:

³⁰ Σύμφωνα με τον Friendly (2008a: 23-24), στον Playfair αποδίδεται η επινόηση, μεταξύ άλλων, των γραφημάτων ράβδων (bar charts) και πιτών (pie charts).

³¹ Όπως σημειώνει ο Rogers (2013a: 62), το εν λόγω βιβλίο κυκλοφόρησε το 1786 και θεωρείται πρωτοπόρο στη χρήση γραφημάτων για την επεξήγηση οικονομικών δεδομένων.

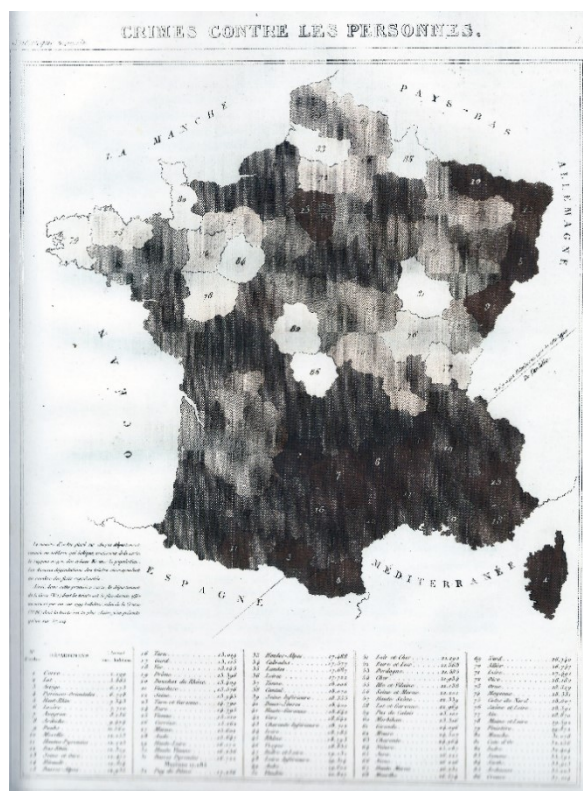
- εφευρίσκονται οι μορφές αναπαράστασης δεδομένων που χρησιμοποιούνται μέχρι και σήμερα, όπως γραφήματα γραμμών (line charts), χρονολογικές σειρές (time-series plots), γραφήματα διασποράς (scatter plots) κ.λπ.
- απεικονίσεις φυσικών φαινομένων συμπεριλαμβάνονται σε επιστημονικές εκδόσεις, σε τακτική πλέον βάση
- κυκλοφορούν αναλυτικοί άτλαντες που απεικονίζουν δεδομένα που αφορούν σε οικονομικά, κοινωνικά και άλλα ζητήματα, εισάγοντας έτσι νέες μορφές συμβολισμού.

Αξίζει να σημειωθούν δύο σημαντικές καινοτομίες που εισάγονται στη Γαλλία. Το 1826, ο Charles Dupin χρησιμοποιεί συνεχείς τονικές διαβαθμίσεις, από το λευκό ως το μαύρο, για να συμβολίσει την κατανομή και τον βαθμό αναλφαβητισμού στη Γαλλία και δημιουργεί τον πρώτο χωροπληθή χάρτη (Friendly, 2008a: 26· Meirelles, 2013: 117). Στην ακόλουθη εικόνα (Εικόνα 9), παρατίθεται ο εν λόγω χάρτης, που παρουσιάζει την ένταση του αναλφαβητισμού ανά διοικητικό διαμέρισμα:



Εικόνα 9: Ο πρώτος χωροπληθής χάρτης από τον Dupin
(Πηγή: Meirelles, 2013: 117)

Λίγα χρόνια αργότερα, το 1833, ο André-Michel Guerry συνδυάζει δεδομένα που αφορούν σε καταγραφές εγκληματικής δραστηριότητας με δεδομένα που αφορούν κοινωνικά ζητήματα, προκειμένου να εξετάσει τη σχέση μεταξύ εγκληματικότητας και αλφαριθμητισμού· απεικονίζοντας σε χάρτες τη σύγκριση μεταξύ μεταβλητών, κλονίζει τελικά την πεποίθηση ότι η μείωση της εγκληματικότητας μπορεί να επιτευχθεί μέσω της ενίσχυσης της εκπαίδευσης (Friendly, 2008a: 26· Meirelles, 2013:143). Στην ακόλουθη εικόνα (Εικόνα 10), παρατίθεται χάρτης που απεικονίζει την εγκληματική δραστηριότητα ανά διοικητικό διαμέρισμα, με τη χρήση τονικών διαβαθμίσεων³²:



Εικόνα 10: Χάρτης του Guerry που απεικονίζει την εγκληματική δραστηριότητα (Πηγή: Meirelles, 2013: 143)

Ο χάρτης αυτός συνοδεύεται από υπομνηματικό πίνακα, στον οποίο παραπέμπουν οι αριθμοί που αναγράφονται στο κέντρο κάθε διοικητικού διαμερίσματος. Κάθε αριθμός αντιστοιχίζεται με το πλήθος εγκλημάτων που

³² Ο πιο σκούρος τόνος συμβολίζει υψηλή εγκληματική δραστηριότητα, ενώ αντίθετα ο πιο ανοιχτός υποδηλώνει λιγότερα εγκλήματα.

έχουν διαπραχθεί στην περιοχή.

Το δεύτερο μισό του 19^{ου} αιώνα χαρακτηρίζεται από τον Friendly (2008a και 2008b) ως η *"Χρυσή Εποχή των στατιστικών γραφημάτων"*. Κατ' αρχάς, καταγράφεται μία σειρά από εξελίξεις στον τομέα της συλλογής δεδομένων και στη βελτίωση των διαθέσιμων τεχνολογικών μέσων. Ήδη, από τις αρχές του αιώνα, στην Ευρώπη συγκεντρώνονται δεδομένα σχετικά με τον πληθυσμό και τις εμπορικές δραστηριότητες -η αύξηση των διαθέσιμων δεδομένων σχετικά με κοινωνικά φαινόμενα τις δεκαετίες του 1820 και του 1830 εμπνέει την κατασκευή γραφικών παραστάσεων για την διερεύνηση κοινωνικών ζητημάτων (Friendly, 2008b: 505). Παράλληλα, έχουν ήδη ιδρυθεί στην Ευρώπη επίσημες κρατικές στατιστικές υπηρεσίες, γεγονός που καταδεικνύει την έμπρακτη αναγνώριση της σημασίας των ποσοτικών δεδομένων σε τομείς όπως ο κοινωνικός σχεδιασμός και το εμπόριο (Friendly, 2008a: 29). Επιπρόσθετα, έχει ενισχυθεί το θεωρητικό υπόβαθρο, καθώς η στατιστική θεωρία καθιστά εφικτή την εξαγωγή συμπερασμάτων από μεγάλα σύνολα δεδομένων (Friendly, 2008a: 29). Τέλος, η εξέλιξη της τεχνολογίας επιτρέπει την παραγωγή και την αναπαραγωγή των γραφικών παραστάσεων³³ και διευκολύνει την καταγραφή δεδομένων (Friendly, 2008b: 507).

Οι κυρίαρχες τάσεις της περιόδου στην κατασκευή γραφικών παραστάσεων και θεματικών χαρτών συνοψίζονται:

- στην ανάπτυξη τεχνικών για την τρισδιάστατη απεικόνιση δεδομένων σε χαρτί (Friendly, 2008a: 30)
- σε καινοτομίες στη γραφική απεικόνιση για την ανάδειξη πολυδιάστατων δεδομένων, με κυρίαρχες φυσιογνωμίες τον Charles Joseph Minard και την Florence Nightingale (Friendly, 2008a: 30-31)
- στη συνεισφορά του Francis Galton στην οπτικοποίηση μετεωρολογικών δεδομένων (Friendly, 2008a: 32)

³³ Ως ιδιαίτερα θετική αξιολογείται η συμβολή της λιθογραφίας: επέτρεψε την εκτύπωση έγχρωμων εικόνων και εξηγεί τη διαδεδομένη χρήση χρώματος στις οπτικές αναπαραστάσεις δεδομένων που κατασκευάζονται αυτή την περίοδο (Friendly, 2008b: 507).

- στη διάδοση των στατιστικών ατλάντων (Friendly, 2008a: 32-36).

Οι καινοτόμες γραφικές αναπαραστάσεις της εν λόγω περιόδου έχουν ιδιαίτερο ενδιαφέρον και προκαλούν τον θαυμασμό ακόμη και σήμερα. Το 1854, ο John Snow συγκεντρώνει επίσημα στοιχεία σχετικά με τα περιστατικά θανάτων από χολέρα στο Νότιο Λονδίνο και τα συνδυάζει με εμπειρικές παρατηρήσεις και τοπικές πληροφορίες (Tufte, 1997: 28· Meirelles, 2013: 135). Επιλέγει να απεικονίσει σε χάρτη το πλήθος των θανάτων, συμβολίζοντάς τους ως γραμμές, και αποκαλύπτει ότι αυτοί εντοπίζονται κυρίως γύρω από μία συγκεκριμένη αντλία νερού³⁴ (Tufte, 1997: 28-29· Friendly, 2008a: 27). Μέσω αυτής της απεικόνισης (Εικόνα 11), επιχειρεί να υποστηρίξει την υπόθεσή του ότι η συγκεκριμένη ασθένεια δεν μεταδίδεται μέσω του αέρα, σύμφωνα με τις τότε εκτιμήσεις, αλλά μέσω του νερού.

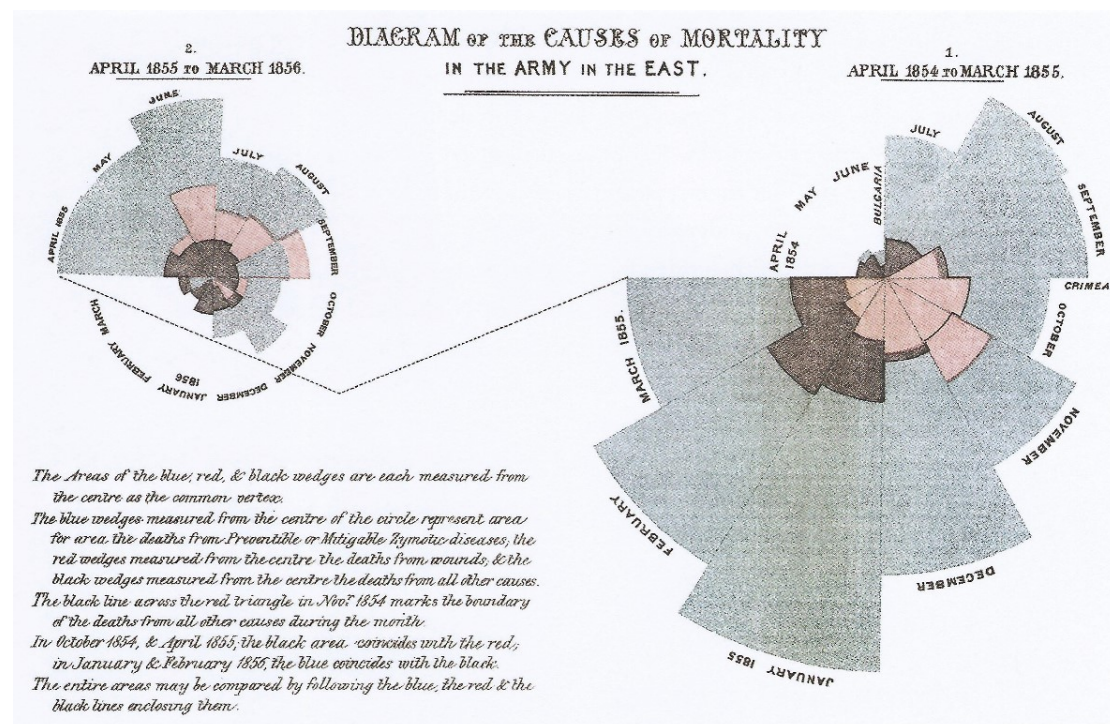


Εικόνα 11: Λεπτομέρεια από τον χάρτη του Snow που απεικονίζει τους θανάτους από χολέρα (Πηγή: Tufte, 1997: 31)

³⁴ Συμβολίζει τις αντλίες νερού ως κύκλους.

Αξίζει να σημειωθεί ότι η πρακτική απεικόνισης δεδομένων σε χάρτες για την κατανόηση των περιβαλλοντικών παραγόντων που συνδέονται με την εκδήλωση επιδημιών εφαρμόζεται από τις αρχές του 19^{ου} αιώνα στην Ευρώπη (Meirelles, 2013: 135). Ωστόσο, η πρωτοτυπία του συγκεκριμένου έγκειται στον συνδυασμό της επιλογής απεικόνισης των δεδομένων στον χάρτη με την αντίληψη του πλαισίου, που επέτρεψε στον Snow να συνδέσει τα περιστατικά θανάτων με τις αντλίες νερού. Όπως επισημαίνει ο Johnson (2007 σε Meirelles, 2013: 135), "η πραγματική καινοτομία βρίσκεται στα δεδομένα από τα οποία παρήχθη αυτό το διάγραμμα, καθώς και στην έρευνα που κατέληξε στην κατ' αρχάς συγκέντρωση αυτών των δεδομένων".

Το 1858, η Nightingale κυκλοφορεί μία ιδιαίτερη γραφική αναπαράσταση που ενισχύει την επιχειρηματολογία της σχετικά με τα αίτια θανάτου των Βρετανών στρατιωτών στον Κριμαϊκό Πόλεμο (Εικόνα 12):



Εικόνα 12: Διαγράμματα της Nightingale που απεικονίζουν τις αιτίες θανάτου Βρετανών στρατιωτών (Πηγή: Rogers, 2013a: 67)

Πρόκειται για μία σύνθεση που περιλαμβάνει δύο κυκλικά διαγράμματα. Το δεξί διάγραμμα απεικονίζει τους θανάτους μέχρι τον Μάρτιο του 1855· τότε απεστάλη στα πεδία των μαχών ειδική υγειονομική επιτροπή για να βελτιώσει τις συνθήκες υγιεινής στις κλινικές και στα στρατόπεδα (Friendly, 2008b: 509). Το δε αριστερό διάγραμμα απεικονίζει τους θανάτους μετά την ανάληψη δράσης από την εν λόγω επιτροπή.

Κάθε κυκλικό διάγραμμα περιλαμβάνει 12 κυκλικούς τομείς ίσης επίκεντρης γωνίας, καθένας εκ των οποίων αντιστοιχεί σε ένα μήνα του έτους. Κάθε τομέας περιλαμβάνει αλληλεπικαλυπτόμενες σφήνες (wedges), η ακτίνα των οποίων υποδηλώνει το μέγεθος των απωλειών ανά αιτία. Οι σφήνες μπλε χρώματος συμβολίζουν τους θανάτους από προβλέψιμες ή ιάσιμες ασθένειες, οι σφήνες κόκκινου χρώματος συμβολίζουν τους θανάτους από τραυματισμούς στη μάχη, ενώ οι μαύρες σφήνες υποδηλώνουν θανάτους από άλλα αίτια.

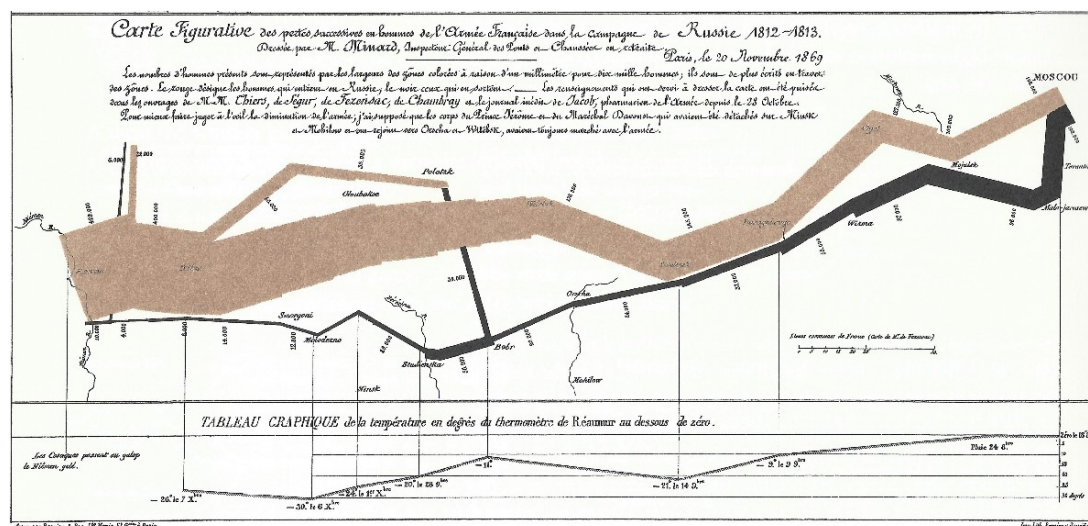
Αν και αυτός ο τύπος διαγράμματος έχει επινοηθεί παλαιότερα από τον Guerry³⁵, η Nightingale είναι η πρώτη που χρησιμοποιεί αυτό και άλλα γραφήματα ως μέσο άσκησης πολιτικής πίεσης και πειθούς (Friendly, 2008b: 509). Όπως και στην περίπτωση του Snow, η καινοτομία δεν έγκειται στην επινόηση ενός νέου τύπου οπτικής αναπαράστασης δεδομένων. Στη συγκεκριμένη περίπτωση, η Nightingale αξιοποιεί αυτόν τον τύπο γραφήματος για να αναδείξει τη βασιμότητα της επιχειρηματολογίας της και την άποψή της ότι η συντριπτική πλειοψηφία των ανθρώπινων απωλειών δεν προκλήθηκε από τραυματισμούς αλλά από προβλέψιμες ή ιάσιμες ασθένειες.

Το 1869, ο Minard δημιουργεί μία γραφική παράσταση που παρουσιάζει τις απώλειες του γαλλικού στρατού κατά την εκστρατεία στη Ρωσία, που ξεκίνησε το 1812 και έληξε το 1813. Σύμφωνα με τον Tufte (2001: 40), συνδυάζει αποτελεσματικά τη χρήση χρονολογικών σειρών με την προσθήκη χωρικών διαστάσεων για να δείξει τη μεταβολή των δεδομένων στον χώρο και στον χρόνο

³⁵ Ο Friendly (2008b: 509) σημειώνει ότι συχνά αποδίδεται στην Nightingale η επινόηση αυτού του τύπου διαγράμματος.

-καταλήγει δε να υποστηρίξει πως πρόκειται ίσως "για το καλύτερο στατιστικό γράφημα που σχεδιάστηκε ποτέ".

Στο εν λόγω γράφημα (Εικόνα 13), ο Minard περιορίζει στο ελάχιστο τις γεωγραφικές πληροφορίες, καθώς εμφανίζονται μόνο τα ονόματα πόλεων και ποταμών (τοποθετημένα σύμφωνα με τις γεωγραφικές συντεταγμένες).



Εικόνα 13: Διάγραμμα του Minard που παρουσιάζει τις απώλειες του γαλλικού στρατού κατά την εκστρατεία στη Ρωσία (Πηγή: Tufte, 2001: 176)

Η καφέ³⁶ γραμμή, από τα αριστερά προς τα δεξιά, δείχνει τους σταθμούς της εκστρατείας μέχρι την κατάληψη της Μόσχας (δεξί άκρο). Η δε μαύρη γραμμή, από τα δεξιά προς τα αριστερά, δείχνει τις περιοχές από τις οποίες πέρασε ο γαλλικός στρατός κατά την υποχώρησή του. Παράλληλα, στο κάτω μέρος έχει τοποθετηθεί γράφημα χρονολογικής σειράς που παρέχει πληροφόρηση για τις θερμοκρασίες κατά την υποχώρηση των γαλλικών στρατευμάτων.

Το πάχος της καφέ και της μαύρης γραμμής συμβολίζει το πλήθος των στρατιωτών· ένα χιλιοστό πάχους γραμμής αντιστοιχεί σε 10.000 στρατιώτες. Επομένως, οι αυξομειώσεις στο πάχος υποδηλώνουν αντίστοιχη μεταβολή στην στρατιωτική δύναμη. Συνδυάζοντας με την πληροφόρηση σχετικά με τις

³⁶ Στο επεξηγηματικό κείμενο του Minard, το χρώμα αναφέρεται ως κόκκινο.

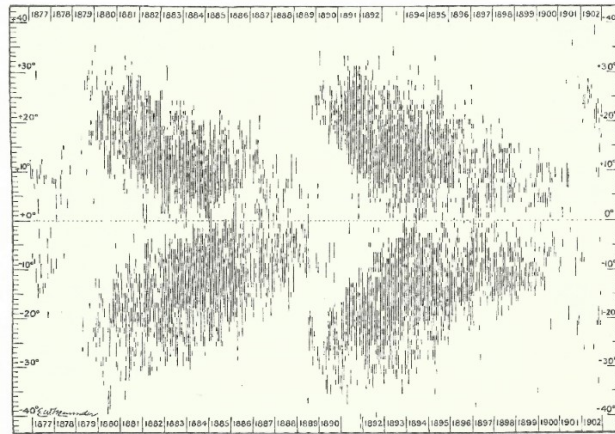
καιρικές συνθήκες, παρέχεται η δυνατότητα εξαγωγής περαιτέρω συμπερασμάτων: για παράδειγμα, όπως σημειώνει η Meirelles (2013: 163), μπορεί να υποτεθεί ότι υπάρχει σχέση μεταξύ των αυξημένων θυμάτων και της επιδείνωσης των καιρικών συνθηκών κατά τη διάσχιση του ποταμού Berezina.

Οι τρεις αυτές περιπτώσεις οπτικών αναπαραστάσεων δεδομένων που περιγράφηκαν πιο πάνω, καθώς και οι υπόλοιπες προσπάθειες που καταγράφονται κατά τη διάρκεια αυτής της περιόδου³⁷, εντυπωσιάζουν ακόμη και σήμερα, ιδιαίτερα λαμβάνοντας υπ' όψιν τα τότε διαθέσιμα τεχνικά μέσα κατασκευής τους. Όπως σημειώνει ο Friendly (2008b: 532), αν και έχουν σχεδιαστεί με το χέρι, αυτά τα διαγράμματα και οι χάρτες "καταφέρνουν να προκαλούν ισχυρό οπτικό αντίκτυπο".

Στις αρχές του 20^{ου} αιώνα, η ποσοτικοποίηση και η εφαρμογή στατιστικών μοντέλων στις κοινωνικές επιστήμες κυριαρχούν· οι γραφικές παραστάσεις θεωρούνται "απλώς εικόνες" που δεν είναι εξίσου ακριβείς με τα αριθμητικά αποτελέσματα των αναλυτικών τεχνικών (Friendly, 2008b: 530). Ωστόσο, είναι μια εποχή διάδοσης των μεθόδων που αναπτύχθηκαν τις προηγούμενες περιόδους: καθιερώνεται η χρήση στατιστικών γραφημάτων για κυβερνητικούς, εμπορικούς και επιστημονικούς σκοπούς, ενώ παράλληλα κυκλοφορούν εγχειρίδια με οδηγίες σχεδιασμού γραφικών παραστάσεων (Friendly, 2008a: 37).

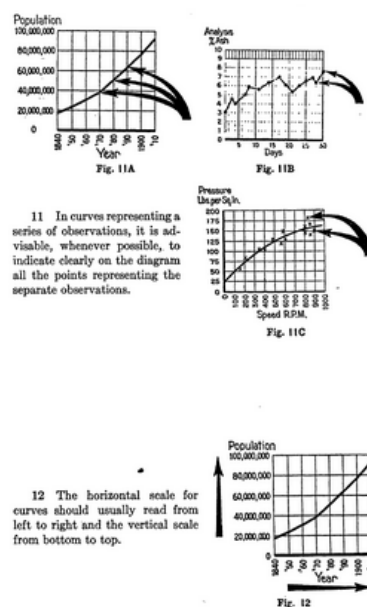
Η αστρονομία, η φυσική, η βιολογία και άλλες επιστήμες επωφελούνται από τις γραφικές αναπαραστάσεις, καθώς χάρη σε αυτές λαμβάνουν χώρα ανακαλύψεις και διατυπώνονται θεωρίες (Friendly, 2008a: 38). Μία από αυτές είναι ανακάλυψη της μείωσης της συχνότητας εμφάνισης των ηλιακών κηλίδων μεταξύ των ετών 1645-1715, χάρη στο διάγραμμα πεταλούδας (butterfly diagram) που δημιούργησε E.W. Maunder το 1904 (Friendly, 2008a: 38), το οποίο παρατίθεται στην ακόλουθη εικόνα (Εικόνα 14):

³⁷ Στο κείμενο *"The Golden Age of Statistical Graphics"*, ο Friendly (2008b) συγκεντρώνει πλούσιο οπτικό υλικό σχετικά με στατιστικά γραφήματα και θεματικούς χάρτες που κατασκευάζονται την περίοδο μεταξύ 1850 και 1900.



Εικόνα 14: Διάγραμμα πεταλούδας του Maunder
(Πηγή: Tufte, 1990: 22)

Παράλληλα, στις ΗΠΑ το 1914, συστήνεται η *Κοινή Επιτροπή για τα Πρότυπα Γραφικής Παρουσίασης* (Joint Committee on Standards for Graphic Presentation - JCSGP), για τη θέσπιση προτύπων και κανόνων σχετικά με τη γραφική παράσταση δεδομένων (Friendly, 2008a: 39). Η πρώτη προκαταρκτική αναφορά της εκδίδεται έναν χρόνο αργότερα και περιλαμβάνει 17 βασικούς κανόνες που συνοδεύονται από αντίστοιχα εικονογραφημένα παραδείγματα (Schmid, 1978: 71). Στην ακόλουθη εικόνα (Εικόνα 15), παρατίθεται μία από τις εικονογραφημένες σελίδες της εν λόγω αναφοράς:



Εικόνα 15: Απόσπασμα από κανόνες που περιλαμβάνονται στην πρώτη αναφορά της JCSGP
(Πηγή: Joint Committee on Standards for Graphic Presentation, 1915: 795)

3.2.2 Μετά την εμφάνιση των ηλεκτρονικών υπολογιστών

Οι Post et al. (2003: ix) εντοπίζουν τις ρίζες του πεδίου της οπτικοποίησης δεδομένων στις αρχές της δεκαετίας του 1950, περίοδο κατά την οποία παρήχθησαν οι πρώτες εικόνες και γραφήματα μέσω ηλεκτρονικού υπολογιστή.

Σύμφωνα με τον Friendly (2008a: 39), από τη δεκαετία του 1960 και μετά, νέες εξελίξεις ανανεώνουν την οπτικοποίηση δεδομένων. Στις Ηνωμένες Πολιτείες, ο John W. Tukey υποστηρίζει τη διάκριση της Ανάλυσης Δεδομένων από τη Μαθηματική Στατιστική και επινοεί νέες γραφικές αναπαραστάσεις υπό την ονομασία *"Διερευνητική Ανάλυση Δεδομένων"* (Exploratory Data Analysis). Η περίοδος αυτή χαρακτηρίζεται συνολικά από την εκρηκτική ανάπτυξη μεθόδων και τεχνικών οπτικοποίησης, η οποία μπορεί να αποδοθεί στην αυξημένη συνεργασία μεταξύ επιστημονικών πεδίων και στην ανάπτυξη νέων λογισμικών που επιτρέπουν την έκφραση στατιστικών ιδεών και την εφαρμογή γραφικών απεικονίσεων δεδομένων (Friendly, 2008a: 39-40).

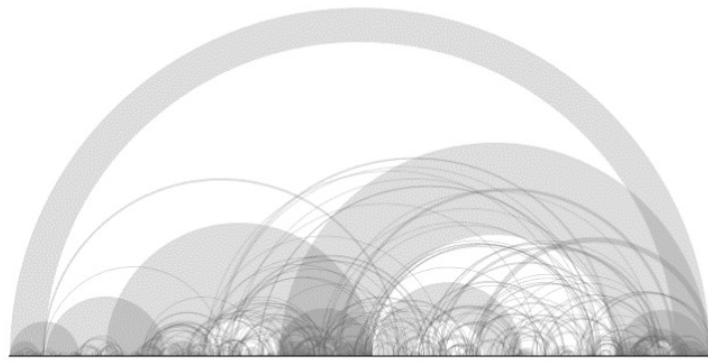
Στο σημείο αυτό, αξίζει να τονιστεί η σημασία που απέδιδε ο Tukey στις δυνατότητες ανάλυσης μέσω της οπτικής αναπαράστασης δεδομένων: κατά τον προσδιορισμό των απαραίτητων εργαλείων για τη διεξαγωγή ανάλυσης δεδομένων, είχε προβλέψει ότι *"τα γραφήματα θα σχεδιάζονται ολοένα και περισσότερο από τους ηλεκτρονικούς υπολογιστές"* χωρίς ανθρώπινη παρέμβαση, ενώ η αυτοματοποίηση της διαδικασίας διάγνωσης και ανίχνευσης ενδείξεων, που είναι κεντρικής σημασίας για την ανάλυση δεδομένων, θα καταστήσει τα γραφήματα και τις υπόλοιπες μορφές παρουσίασης *"το όχημα μέσω του οποίου ο άνθρωπος θα επιβλέπει και θα αποδέχεται ή θα απορρίπτει τις ενδείξεις"* που θα έχει εντοπίσει η μηχανή (Tukey, 1962: 60-61). Η πρόβλεψή του αυτή φαίνεται να επιβεβαιώνεται από την αυξημένη προσφορά εργαλείων οπτικής ανάλυσης που παρέχονται από εταιρείες κατασκευής λογισμικού επιχειρηματικής ευφυΐας για την εξαγωγή αξίας από σύνολα ΜΔ.

Το τελευταίο τέταρτο του εικοστού αιώνα, η οπτικοποίηση δεδομένων μετασχηματίζεται σε διεπιστημονικό ερευνητικό πεδίο, ως υποπεδίο των

γραφικών υπολογιστών (Post et al., 2003: ix· Wright, 2007: 1-2· National Science Foundation, 2000: 92). Το 1987, κυκλοφορεί από το National Science Foundation (NSF) η αναφορά *"Visualization in Scientific Computing"*, στην οποία αναγνωρίζεται η διττή χρησιμότητα της οπτικοποίησης, ως μέσου κατανόησης οπτικών κωδικοποιήσεων αφηρημένων δεδομένων και ως μέσο δημιουργίας τέτοιων αναπαραστάσεων από σύνθετα και πολυδιάστατα σύνολα δεδομένων (NSF, 1987: 3). Ταυτόχρονα εκφράζεται έντονα η ανάγκη για την ανάπτυξη νέων τεχνικών οπτικοποίησης μέσω ηλεκτρονικών υπολογιστών (National Science Foundation, 1987: 10). Δεδομένου ότι τα διαθέσιμα δεδομένα προς αξιοποίηση που αναφέρονται στο εν λόγω κείμενο αφορούν κυρίως μετρήσεις και αυτοί που ασχολούνται με το πεδίο είναι ως επί το πλείστον επιστήμονες με θετικό υπόβαθρο, θα μπορούσε να υποστηριχθεί ότι οι ιδιότητες που προσδίδονται στην οπτικοποίηση, παραπέμπουν στους στενούς ορισμούς της επιστημονικής οπτικοποίησης, που παρατίθενται στη συνέχεια (Ενότητα 3.3).

Η συνεχώς αυξανόμενη συγκέντρωση δεδομένων από ποικίλες πηγές προέλευσης, που πλέον δεν περιορίζονται στις καταγραφές από το φυσικό περιβάλλον, αλλά αφορούν σε συναλλαγές και αλληλεπιδράσεις μεταξύ ατόμων έχει πλέον μεταβάλει τη φύση των υπό εξέταση δεδομένων. Από τις αρχές της δεκαετίας του 1990, χρησιμοποιείται ο όρος *οπτικοποίηση πληροφορίας*, για να περιγράψει την οπτική αναπαράσταση δεδομένων από πολυμετάβλητες βάσεις δεδομένων ή από συλλογές εγγράφων (Card et al., 1999: 8).

Παράλληλα με την αύξηση των συλλεγόμενων δεδομένων, οι τεχνολογικές εξελίξεις συμβάλλουν στην υιοθέτηση της οπτικοποίησης δεδομένων ως μέσο ανάλυσης και παρουσίασης δεδομένων. Αξίζει να σημειωθεί ότι κάποιες από τις μεθόδους οπτικοποίησης που επινοήθηκαν στις αρχές του 21^{ου} αιώνα, όπως τα διαγράμματα αψίδας (arc diagrams) (Wattenberg, 2002), αξιοποιούν τις τεχνολογικές δυνατότητες για καινοτόμες μορφές αναπαράστασης δεδομένων.



Εικόνα 16: Διάγραμμα αφίδας που οπτικοποιεί τη μουσική σύνθεση *Für Elise* του Beethoven, από Wattenberg
(Πηγή: Wattenberg, 2002: 4)

Επιπρόσθετα, αναπτύσσονται νέες γλώσσες προγραμματισμού και λογισμικά που απευθύνονται σε όσους έχουν διάθεση πειραματισμού και ανάλυσης των δεδομένων, χωρίς να διαθέτουν υπόβαθρο προγραμματισμού. Το 2001, οι Ben Fry και Casey Reas ξεκινούν την ανάπτυξη της προγραμματιστικής γλώσσας *Processing*, που φιλοδοξούν "να καταστρέψει τις καριέρες ταλαντούχων σχεδιαστών, δελεάζοντάς τους να παρατήσουν τα συνήθη εργαλεία τους και να στραφούν στον κόσμο του προγραμματισμού και της υπολογιστικής" (*Processing*, 2014). Επίσης, λογισμικά εξειδικευμένα στην οπτικοποίηση δεδομένων, όπως το *Many Eyes* της IBM και το *Tableau Public* της Tableau, αποκτούν δημοφιλία. Η χρήση αυτών των τεχνολογικών λύσεων από δημοσιογραφικά μέσα για την πλαισίωση ιστοριών που βασίζονται στην ανάλυση μεγάλων συνόλων δεδομένων (Rogers, 2012: 206-212), συνδέεται με την άνθιση της δημοσιογραφίας δεδομένων.

3.3 Ορισμοί

Σύμφωνα με τα προαναφερθέντα, καθίσταται αντιληπτό ότι η σύγχρονη περίοδος της οπτικοποίησης καθορίζεται από τις εξελίξεις στους ηλεκτρονικούς υπολογιστές και τα γραφικά υπολογιστών (*computer graphics*) ειδικότερα. Πώς όμως ορίζεται η οπτικοποίηση; Αφετηρία για την εξερεύνηση των προτεινόμενων ορισμών αποτελεί ένα κείμενο της σύγχρονης περιόδου που

συγκαταλέγεται στα πιο συχνά επικαλούμενα στη βιβλιογραφία.³⁸ Σε αυτό, οι Card et al. (1999: 7) διατυπώνουν ορισμούς εργασίας, τους οποίους συσχετίζουν ιεραρχικά, προκειμένου να ξεκαθαρίσουν τις σχέσεις μεταξύ ορισμών και εννοιών που συνδέονται με την οπτικοποίηση. Η καταγραφή αυτή φαίνεται στον ακόλουθο πίνακα (Πίνακας 4):

Πίνακας 4: Πίνακας ορισμών από Card et al.
(Πηγή: Card et al., 1999: 7)

Όρος	Ορισμός
Εξωτερική νόηση (<i>External cognition</i>)	χρήση του εξωτερικού κόσμου για την επίτευξη νόησης
Σχεδιασμός Πληροφορίας (<i>Information Design</i>)	σχεδιασμός εξωτερικών αναπαραστάσεων, προκειμένου να διευρυνθεί η νόηση
Γραφικά δεδομένων (<i>Data graphics</i>)	χρήση αφηρημένων, μη αναπαραστατικών οπτικών αναπαραστάσεων δεδομένων, προκειμένου να διευρυνθεί η νόηση
Οπτικοποίηση (<i>Visualization</i>)	χρήση υποστηριζόμενων από ηλεκτρονικούς υπολογιστές, αλληλεπιδραστικών οπτικών αναπαραστάσεων δεδομένων, προκειμένου να διευρυνθεί η νόηση
Επιστημονική οπτικοποίηση (<i>Scientific visualization</i>)	χρήση αλληλεπιδραστικών οπτικών αναπαραστάσεων επιστημονικών δεδομένων, συνήθως προερχόμενων από τον φυσικό κόσμο, προκειμένου να διευρυνθεί η νόηση
Οπτικοποίηση πληροφορίας (<i>Information visualization</i>)	χρήση αλληλεπιδραστικών οπτικών αναπαραστάσεων αφηρημένων δεδομένων, μη προερχόμενων από τον φυσικό κόσμο, προκειμένου να διευρυνθεί η νόηση

³⁸ Το κείμενο αυτό περιλαμβάνεται στην πρώτη συλλογή επιστημονικών δημοσιεύσεων με στόχο τη συστηματική καταγραφή των πτυχών του πεδίου της οπτικοποίησης πληροφορίας. Ο τίτλος αυτής, *"Information Visualization: Using Vision to Think"* (Card et al., 1999), προϋδεάζει για τις προθέσεις των επιμελητών της έκδοσης.

Αυτές αναφέρονται επιγραμματικά στο οπισθόφυλλο και συνίστανται στην "παρουσίαση δημοσιεύσεων που επικεντρώνονται στην χρήση της οπτικοποίησης για την ανακάλυψη σχέσεων, μέσω της χρήσης αλληλεπιδραστικών γραφικών για τη διεύρυνση της σκέψης" (Card et al., 1999).

Επισημαίνεται ότι ο όρος "*cognition*" αποδίδεται ως *νόηση*, λαμβάνοντας υπόψη ότι οι Card et al. (1999: 6) προσδίδουν σε αυτήν τη σημασία "*της απόκτησης ή της χρήσης της γνώσης*" και επιδιώκουν οι ορισμοί τους να περιλαμβάνουν τόσο τον σκοπό της οπτικοποίησης³⁹, όσο και τα μέσα επίτευξής του.

Συγκρίνοντας τους ορισμούς που διατυπώνουν οι Card et al., παρατηρείται ότι η διεύρυνση της νόησης αποτελεί κοινό ζητούμενο σε κάθε περίπτωση οπτικής αναπαράστασης. Με βάση την ιεραρχική δομή που προτείνουν, προκύπτει ότι ο όρος "*οπτικοποίηση*" αποτελεί υποκατηγορία των γραφικών δεδομένων, και χρησιμοποιείται για να εξειδικεύσει εκείνες τις οπτικές αναπαραστάσεις που δημιουργούνται μέσω ηλεκτρονικού υπολογιστή και παρέχουν από δυνατότητα αλληλεπίδρασης. Δεδομένου ότι εξειδίκευση αυτή δεν προκύπτει ετυμολογικά, διακρίνεται ένας βαθμός αυθαιρεσίας στη χρήση της λέξης "*οπτικοποίηση*" για την υποδήλωση της χρήσης ηλεκτρονικών υπολογιστών και των δυνατοτήτων αλληλεπίδρασης που παρέχονται στον χρήστη. Στο πλαίσιο του συγκεκριμένου κειμένου δεν προκύπτει ζήτημα παρερμηνείας καθώς προσδιορίζεται ρητά, ωστόσο επισημαίνεται αυτή τη λεπτομέρεια, διότι φαίνεται να έχει εκλείψει σε μεταγενέστερους ορισμούς.

Για παράδειγμα, ο Spence (2007: 5-6) διατυπώνει την άποψη ότι η οπτικοποίηση είναι μία "*αποκλειστικά ανθρώπινη γνωστική δραστηριότητα που δεν έχει σχέση με τους ηλεκτρονικούς υπολογιστές*", και ορίζει συγκεκριμένα την οπτικοποίηση πληροφορίας ως τη διαδικασία κατά την οποία δεδομένα μετατρέπονται σε εικόνες, οι οποίες με τη σειρά τους ερμηνεύονται από τον άνθρωπο, με απώτερο στόχο την εξαγωγή της πληροφορίας από αυτά. Επισημαίνει δε ότι η δραστηριότητα αυτή επηρεάζεται και από άλλα αισθητηριακά ερεθίσματα πέραν των οπτικών, όπως ο ήχος, προκειμένου να

³⁹ Ο οποίος συνίσταται στην επίτευξη κατανόησης ("*insight*") με στόχο "*την ανακάλυψη, τη λήψη αποφάσεων και την επεξήγηση*" (Card et al. 1999: 6).

σχηματιστεί ένα νοητικό μοντέλο των δεδομένων που αναπαρίστανται.

Αυτός ο συλλογισμός του είναι εν μέρει σύμφωνος με την αναγνώριση από τους Card et al. (1999: 7) ότι η οπτικοποίηση αποτελεί υποκατηγορία της αισθητοποίησης (perceptualization). Παρότι αναγνωρίζουν ότι τα δεδομένα μπορούν να αναπαρασταθούν είτε ακουστικά είτε απτικά, δηλώνουν ότι επιλέγουν να εστιάσουν στην οπτικοποίηση λόγω των δυνατοτήτων του ανθρώπινου οπτικού συστήματος (Card et al., 1999: 7).

Στον αντίποδα των συγκεκριμένων ορισμών που παρέχουν οι Card et al., άλλοι ερευνητές προτιμούν πιο γενικές περιγραφές. Σύμφωνα με τους Ward et al. (2010: 1) η οπτικοποίηση είναι *"η επικοινωνία πληροφοριών με τη χρήση γραφικών αναπαραστάσεων"*, ενώ ο Kirk (2012: 16-17) ορίζει την οπτικοποίηση δεδομένων ως την *"αναπαράσταση και παρουσίαση δεδομένων που εκμεταλλεύεται τις ικανότητες οπτικής αντίληψης με σκοπό τη διεύρυνση της νόησης"*. Στον τελευταίο μάλιστα διακρίνεται έμμεση αναφορά στον ορισμό των Card et al., λόγω της διατύπωσης για διεύρυνση της νόησης.

Ο Cairo (2013: 18-19) επισημαίνει ότι οι γραφικές αναπαραστάσεις είναι είτε μεταφορικές είτε κυριολεκτικές, και για τον λόγο αυτόν, υιοθετεί τον ορισμό του Costa ότι η οπτικοποίηση

"συνίσταται στο να κάνει κάποια φαινόμενα και τμήματα της πραγματικότητας ορατά και κατανοητά -αρκετά από αυτά τα φαινόμενα δεν είναι προσβάσιμα με φυσικό τρόπο στο γυμνό μάτι και αρκετά από αυτά δεν είναι καν οπτικής φύσης".

Προς επίρρωση αυτού, αναφέρει ως παράδειγμα κυριολεκτικής αναπαράστασης τους χάρτες ως πορτραίτο σε κλίμακα μίας γεωγραφικής περιοχής, και σημειώνει ότι η αναπαράσταση αφηρημένων φαινομένων είναι μεταφορική διότι η σχέση μεταξύ της αναπαράστασης και του αναπαριστώμενου είναι αποτέλεσμα σύμβασης και δεν εκπορεύεται από τη φύση.

Αυτή η αντιδιαστολή μεταξύ κυριολεκτικών και μεταφορικών γραφικών

αναπαραστάσεων, επιβάλλει την επιστροφή στο σχήμα που προτείνουν οι Card et al. (Πίνακας 4 πιο πάνω). Βάσει αυτού, η οπτικοποίηση διαχωρίζεται σε δύο κατηγορίες, ανάλογα με το είδος δεδομένων που αναπαρίστανται: εάν τα δεδομένα προέρχονται από οντότητες με φυσική υπόσταση⁴⁰ πρόκειται για *επιστημονική οπτικοποίηση*, ενώ εάν πρόκειται για αφηρημένα δεδομένα⁴¹ πρόκειται για *οπτικοποίηση πληροφορίας* (Card et al., 1999: 6-7).

Ο διαχωρισμός αυτός έχει ιστορικές καταβολές. Η οπτικοποίηση μέσω ηλεκτρονικών υπολογιστών, μέχρι τα τέλη της δεκαετίας του 1990, αφορούσε κυρίως δεδομένα που προέρχονταν από επιστημονικές καταγραφές και προσομοιώσεις (Gershon and Eick, 1997: 29).⁴² Ωστόσο, από τις αυξανόμενες ανάγκες για αντιμετώπιση των προκλήσεων ανάλυσης νέων τύπων δεδομένων, που προκύπτουν από τη διάδοση του Διαδικτύου και από την ανάπτυξη αποθηκών δεδομένων εντός των οργανισμών, προέκυψε το ερευνητικό πεδίο της *"οπτικοποίησης πληροφορίας"*: αυτό συνδυάζει *"πτυχές της επιστημονικής οπτικοποίησης, των διεπαφών ανθρώπου-υπολογιστή, της εξόρυξης δεδομένων, της απεικόνισης και των γραφικών"* (Gershon and Eick, 1997: 29).

Στον ακόλουθο πίνακα (Πίνακας 5), οι Gershon και Eick (1997: 29), συγκρίνουν τα δύο είδη οπτικοποίησης:

⁴⁰ Αναφέρουν ως ενδεικτικά παραδείγματα τη γη, το ανθρώπινο σώμα και τα μόρια (Card et al., 1999: 6).

⁴¹ Αναφέρουν ως ενδεικτικά παραδείγματα οικονομικά δεδομένα, επιχειρηματικές πληροφορίες, συλλογές αρχείων και αφηρημένες έννοιες (Card et al., 1999: 6-7).

⁴² Όπως σημειώνουν οι Gershon και Eick (1997: 29), μέχρι τότε, μεγάλο μέρος της έρευνας σχετικά με την οπτικοποίηση περιοριζόταν στην κάλυψη αυτών των αναγκών. Οι Bonneau et al. (2006: Preface), από την σκοπιά των θετικών επιστημών, σημειώνουν ότι η επιστημονική οπτικοποίηση *"ασχολείται με τεχνικές που επιτρέπουν στους επιστήμονες και στους μηχανικούς να εξάγουν γνώση από τα αποτελέσματα προσομοιώσεων και υπολογισμών"*. Σε παρόμοιο μοτίβο, για τους Post et al. (2003: ix) η επιστημονική οπτικοποίηση συνιστά τμήμα της επιστημονικής υπολογιστικής, δηλ. της *"χρήσης μοντέλων και προσομοιώσεων μέσω ηλεκτρονικών υπολογιστών στην επιστημονική και μηχανολογική πρακτική"*.

Πίνακας 5: Σύγκριση ειδών οπτικοποίησης από Gershon και Eick (Πηγή: Gershon and Eick, 1997: 29)

	<i>Κοινό</i>	<i>Εργασία</i>	<i>Είδος δεδομένων</i>
Επιστημονική οπτικοποίηση	<i>Εξειδικευμένο, με ιδιαίτερες τεχνικές γνώσεις</i>	<i>Βαθιά κατανόηση επιστημονικών φαινομένων</i>	<i>Φυσικά δεδομένα, μετρήσεις, αποτελέσματα προσομοιώσεων</i>
Οπτικοποίηση πληροφορίας	<i>Ποικιλόμορφο και ευρύ, χωρίς ιδιαίτερες τεχνικές γνώσεις</i>	<i>Αναζήτηση, ανακάλυψη σχέσεων</i>	<i>Σχέσεις, μη φυσικά δεδομένα, πληροφορίες</i>

Οι Ward et al. (2010: 23-24), θεωρούν ότι σήμερα δεν υφίσταται λόγος διαφοροποίησης μεταξύ της επιστημονικής οπτικοποίησης και της οπτικοποίησης πληροφορίας, καθώς και στις δύο περιπτώσεις αναπαρίστανται δεδομένα, ωστόσο αναγνωρίζουν ότι η αντιμετώπιση ενός συγκεκριμένου συνόλου δεδομένων ενδέχεται να διαφοροποιείται ανάλογα με τον σκοπό της οπτικοποίησης. Παράλληλα, ο Manovich (2010: 3-4) εκτιμά πως το γεγονός ότι ορισμένοι ερευνητές εξακολουθούν να αποδέχονται αυτόν τον διαχωρισμό, οφείλεται κυρίως στο ότι "η επιστημονική οπτικοποίηση και η οπτικοποίηση της πληροφορίας προέρχονται από διαφορετικές κουλτούρες, και η ανάπτυξή τους αντιστοιχεί σε διαφορετικές [ερευνητικές] περιοχές των τεχνολογιών των γραφικών υπολογιστών".

Ο Friendly (2009: 2), στο πλαίσιο της προσπάθειας να οριοθετήσει "τις ποικιλίες οπτικοποίησης", παρέχει μία σειρά ορισμών:

- η επιστημονική οπτικοποίηση είναι "η οπτικοποίηση τρισδιάστατων φαινομένων (αρχιτεκτονικών, μετεωρολογικών, ιατρικών, βιολογικών, κ.λπ.), όπου η έμφαση αποδίδεται στη ρεαλιστική απεικόνιση των όγκων, των πηγών φωτισμού, ενδεχομένως με μία δυναμική (χρονική) συνιστώσα"

- η *οπτικοποίηση πληροφορίας* αποτελεί ένα γενικό πλαίσιο που καλύπτει τις οπτικές αναπαραστάσεις *"μεγάλων συλλογών μη αριθμητικών πληροφοριών, όπως αρχεία, κώδικα λογισμικού, βιβλιογραφικές βάσεις δεδομένων, δίκτυα σχέσεων στο Διαδίκτυο"*
- η *οπτικοποίηση δεδομένων* είναι η *"επιστήμη της οπτικής αναπαράστασης των 'δεδομένων', που ορίζονται ως πληροφορία στην οποία, μέσω της αφαίρεσης, έχει αποδοθεί κάποια σχηματική μορφή, που συμπεριλαμβάνει χαρακτηριστικά ή μεταβλητές για τις μονάδες της πληροφορίας"*.

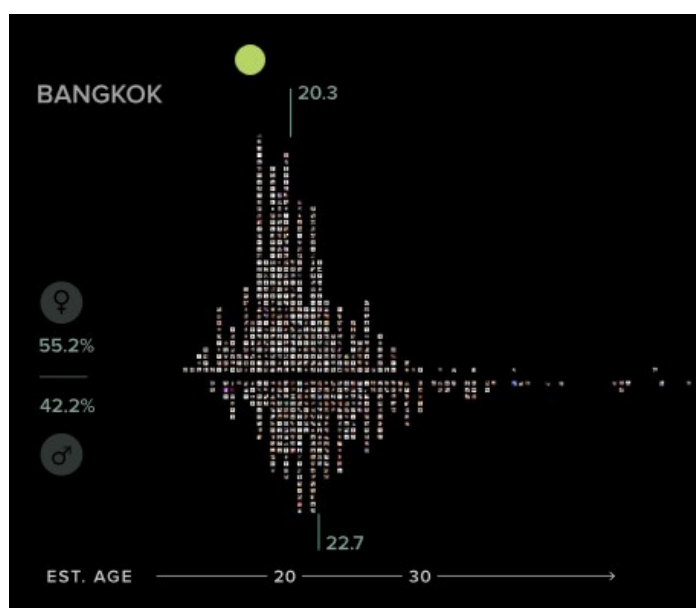
Θα μπορούσε να υποστηριχθεί ότι ο διαχωρισμός που επιχειρεί μεταξύ οπτικοποίησης δεδομένων και οπτικοποίησης πληροφορίας αντανακλά τη διάκριση μεταξύ δομημένων και μη δομημένων δεδομένων αντιστοίχως. Η σταδιακή και αυξανόμενη συγκέντρωση δεδομένων και από άλλες πηγές, όπως επιχειρήσεις ή ψηφιακά μέσα, έχει δημιουργήσει νέες ανάγκες. Επομένως, μία *οπτικοποίηση ΜΔ* θα απαιτούσε χωριστές υποκατηγορίες, ανάλογα με τα αναπαριστώμενα δεδομένα.

Συμπερασματικά, ο διαχωρισμός μεταξύ επιστημονικής οπτικοποίησης και οπτικοποίησης πληροφορίας έχει ιστορικές καταβολές και βασίζεται στην προέλευση των δεδομένων ή/και στον σκοπό αναπαράστασης αυτών. Μέχρι πρότινος, δεν υπήρχε κάποιο σημείο τομής μεταξύ των δύο ειδών οπτικοποίησης. Ωστόσο, σήμερα με τη συγκέντρωση ΜΔ, όπου ακόμη και μη δομημένα δεδομένα μπορούν να καταστούν επεξεργάσιμα και να υποστούν ανάλυση, τα όρια συγχέονται.

Δεν αποκλείεται βέβαια μελλοντικά να διατυπωθούν ορισμοί για καινοφανείς μορφές οπτικοποίησης που θα εξυπηρετούν τις τότε ανάγκες ανάλυσης και παρουσίασης των συλλεγόμενων δεδομένων. Όπως εύστοχα επισημαίνει ο Migurski (2014 σε Hoffer, 2014), *"η οπτικοποίηση δεδομένων είναι σχετικός όρος ... πάντα αναφέρεται σε κάτι καινούργιο που διακρίνεται στον ορίζοντα"*.

Για παράδειγμα, ο Manovich (2010: 23) έχει ήδη προτείνει ένα νέο είδος οπτικοποίησης, την οπτικοποίηση μέσων (media visualization) που συνίσταται στη *"δημιουργία νέων οπτικών αναπαραστάσεων από υπάρχοντα πολυμεσικά αντικείμενα ή μέρη αυτών"*, όπου *"οι εικόνες παραμένουν εικόνες, το κείμενο παραμένει κείμενο"*.

Κάποια χρόνια πριν θα ήταν αδιανόητη η ανάλυση των εικόνων με αυτόν τον τρόπο. Ωστόσο, ο συνδυασμός των τεχνολογικών εξελίξεων, οι οποίες έχουν καταστήσει εφικτή την ψηφιοποίηση και την επεξεργασία τους, με το ανήσυχο πνεύμα και τη δημιουργικότητα ανθρώπων όπως ο Manovich, καταλήγει σε μία οπτικοποίηση που μπορεί να περιγραφεί ακριβέστερα με τη διατύπωση ενός νέου όρου: ο όρος αυτός εκφράζει ταυτόχρονα το είδος των δεδομένων που οπτικοποιούνται και το αποτέλεσμα της οπτικοποίησης.



Εικόνα 17: Imageplot από το εγχείρημα *"Selfiecity"*, που συντονίζει ο Manovich (Πηγή: Selfiecity, 2014)

Καταλήγοντας και λαμβάνοντας υπ' όψιν τα προαναφερθέντα, στο πλαίσιο της παρούσας εργασίας, ο όρος *"οπτικοποίηση δεδομένων"*⁴³ περιγράφει τις μεθόδους και το τελικό αποτέλεσμα της αναπαράστασης δεδομένων, δομημένων

⁴³ Για την αποφυγή επαναλήψεων, εναλλάσσεται με τη διατύπωση *"οπτική αναπαράσταση δεδομένων"*.

ή μη, σε οπτικά αντιληπτή μορφή, που επιτρέπουν την ταχεία διαμόρφωση άποψης και την εξαγωγή νέας ή την ενίσχυση υφιστάμενης γνώσης από τα αναπαριστώμενα δεδομένα.

3.4 Επιστήμη, τέχνη ή κάτι άλλο;

Στην παρούσα ενότητα, αναφέρονται επιλεγμένες επιστημολογικές οπτικές που αναδεικνύουν ενδιαφέρουσες πτυχές της οπτικοποίησης, όπως προκύπτουν από τη σχετική βιβλιογραφία.

Ο Kosara (2006, σε Jankun-Kelly et al., 2006: 68-69) διακρίνει την αναγκαιότητα δημιουργίας μίας επιστήμης της οπτικοποίησης, η οποία θα εξειδικεύεται σε επιμέρους κλάδους, όπως *"σημειωτική της οπτικοποίησης, αισθητική της οπτικοποίησης, φιλοσοφία της οπτικοποίησης κ.ο.κ."*. Εκφράζει την άποψη ότι οι ερευνητές και οι επαγγελματίες του πεδίου θα πρέπει να συνεργαστούν με μέλη άλλων επιστημονικών κοινοτήτων που συνδέονται κατά κάποιον τρόπο με την οπτικοποίηση, προκειμένου να ανταλλάξουν ιδέες και ερωτήματα προς όφελος αμφοτέρων (Kosara, 2006 σε Jankun-Kelly et al., 2006: 69).

Ο North (2006, σε Jankun-Kelly et al., 2006: 69) επισημαίνει ότι τα βασικά χαρακτηριστικά της επιστήμης είναι οι μετρήσεις και η προτυποποίηση, επομένως *"μία επιστήμη της οπτικοποίησης θα πρέπει να θεσπίσει μεθόδους μετρήσεων και πρότυπα που να καλύπτουν τους θεμελιώδεις σκοπούς και στόχους της οπτικοποίησης"*. Παράλληλα, σημειώνει ότι αυτές οι μέθοδοι θα πρέπει να μην περιορίζονται στην καταγραφή της ανθρώπινης απόδοσης βάσει χρόνου και ακρίβειας εκτέλεσης συγκεκριμένων εργασιών, αλλά να θέτουν και κριτήρια (1) για τον προσδιορισμό των χαρακτηριστικών της αλληλεπίδρασης που λαμβάνει χώρα στο πλαίσιο της οπτικοποίησης και (2) για τη μέτρηση των συνεπειών της σε ανθρώπινους χρήστες (North, 2006 σε Jankun-Kelly et al., 2006: 69-70). Ωστόσο, αναγνωρίζει ότι η *επίτευξη κατανόησης*, που αποτελεί βασικό στόχο της οπτικοποίησης, και τα βήματα που οδηγούν σε αυτή δύσκολα μπορούν

να αξιολογηθούν (North, 2006 σε Jankun-Kelly et al., 2006: 70).

Οι Fekete et al. (2008: 3-4), οι οποίοι επίσης αναγνωρίζουν ότι *"από τη φύση της και τον σκοπό της, η οπτικοποίηση πληροφορίας παρουσιάζει θεμελιώδεις προκλήσεις για την αναγνώριση και τη μέτρηση της αξίας της"*, καταλήγουν ότι αποτελεί μάλλον μέθοδο παρά επιστήμη, που αξιοποιεί το ανθρώπινο οπτικό σύστημα για τον εντοπισμό μοτίβων μέσω της γρήγορης επεξεργασίας αφηρημένων οπτικών αναπαραστάσεων, εκμεταλλευόμενη τις ιδιότητες της ανθρώπινης αντίληψης.

Ο Ware (2006 σε Jankun-Kelly et al., 2006: 70) εκφράζει την άποψη ότι η οπτικοποίηση αποτελεί έναν επιστημονικό κλάδο των γνωσιακών επιστημών⁴⁴ και τονίζει ότι η διατύπωση μίας θεωρίας της οπτικοποίησης θα πρέπει να συνδέεται με τη μελέτη της ανθρώπινης αντίληψης και με την ακόμη πιο εξειδικευμένη έρευνα του ανθρώπινου οπτικού συστήματος, στοχεύοντας στην εξήγηση και στην ανάπτυξη των γνωσιακών διαδικασιών.⁴⁵

Οι Viégas και Wattenberg (2007: 183-184) υποστηρίζουν ότι καλλιτεχνικά έργα βασισμένα σε δεδομένα συνιστούν *"καλλιτεχνικές οπτικοποιήσεις"*, τις οποίες ορίζουν ως *"οπτικοποιήσεις δεδομένων που δημιουργούνται από καλλιτέχνες με την πρόθεση να δημιουργήσουν τέχνη"*. Εντοπίζουν δύο βασικούς παράγοντες που έχουν συμβάλει στην εμφάνιση αυτού του νέου είδους οπτικοποίησης (ή ακόμη και τέχνης): (1) σήμερα υπάρχουν τα κατάλληλα εργαλεία λογισμικού για την καλλιτεχνική παραγωγή οπτικοποιήσεων, και (2) τα δεδομένα που είναι ευρέως διαθέσιμα, εμπνέουν τους καλλιτέχνες να δημιουργήσουν έργα που θίγουν ζητήματα σχετικά με *"τον έλεγχο της ισχύος των δεδομένων"* (Viégas and Wattenberg, 2007: 184).

⁴⁴ Σε αυτές συγκαταλέγει την ψυχολογία, τη γνωσιακή νευροεπιστήμη, την τεχνητή νοημοσύνη, την επαυξημένη νοημοσύνη και την αλληλεπίδραση ανθρώπου-υπολογιστή (Ware, 2006 σε Jankun-Kelly et al., 2006: 70).

⁴⁵ Στο βιβλίο του *"Information Visualization: Perception for Design"* επιχειρεί να θέσει τα θεμέλια για μια *"εφαρμοσμένη επιστήμη Οπτικοποίησης Δεδομένων"* και εξετάζει τον τρόπο που η οπτικοποίηση *"μπορεί να λειτουργήσει ως γνωσιακό εργαλείο"* (Ware, 2013: 2).

Επισημαίνουν δε ότι διαφοροποιείται σε σχέση με τις υπόλοιπες οπτικοποιήσεις αφενός ως προς την πρόθεση του δημιουργού, καθώς επιδιώκεται η έκφραση συγκεκριμένης άποψης ή προβληματισμού για τα δεδομένα που αναπαρίστανται, αφετέρου ως προς το ιδιαίτερο στίγμα τους, που προκύπτει "*διαπράττοντας αμαρτήματα*", παραβιάζοντας δηλαδή βασικές αρχές που πρέπει να τηρούνται στο πλαίσιο της Οπτικής Ανάλυσης για να μην προκύπτει ένα υποκειμενικό αποτέλεσμα (Viégas and Wattenberg, 2007: 189-190).



Εικόνα 18: Παράδειγμα καλλιτεχνικής οπτικοποίησης
(Πηγή: Viégas and Wattenberg, 2007: 188)

Τέλος, καταλήγουν ότι οι οπτικοποιήσεις εμπεριέχουν το υποκειμενικό στοιχείο του δημιουργού, το οποίο είναι πιο έντονο στις καλλιτεχνικές οπτικοποιήσεις, και προτρέπουν τους ερευνητές που ασχολούνται με την οπτικοποίηση να αναρωτηθούν για το τι μπορούν να αποκομίσουν από αυτές τις καλλιτεχνικές εξερευνήσεις (Viégas and Wattenberg, 2007: 190-191).

Σκοπός της αναφοράς σε αυτές τις διαφορετικές απόψεις δεν αποτελεί η διατύπωση συγκεκριμένης άποψης, αλλά η ανάδειξη της ευρύτερης προβληματικής για τη θέση της οπτικοποίησης. Ωστόσο, θα μπορούσε να θεωρηθεί ότι ο χαρακτηρισμός της οπτικοποίησης ως "*γνωσιακή τέχνη*" (Tuft, 1990: 9) συνδυάζει τις φαινομενικά ασύμβατες προσεγγίσεις και περιγράφει εύστοχα την πολύπλευρη υπόστασή της.

3.5 Οπτική Ανάλυση Δεδομένων

Η Οπτική Ανάλυση Δεδομένων συνίσταται στον "*σχηματισμό αφηρημένων*

οπτικών μεταφορών σε συνδυασμό με ανθρώπινο πληροφοριακό διάλογο (αλληλεπίδραση) που καθιστά εφικτό τον εντοπισμό των αναμενόμενων και την ανακάλυψη των μη αναμενόμενων εντός τεράστιων, δυναμικά μεταβαλλόμενων χώρων πληροφοριών" (Wong and Thomas, 2004: 20).

Οι Thomas και Cook (2005: 4) παρέχουν ένα πλήρες πλαίσιο, ορίζοντας την Οπτική Ανάλυση Δεδομένων ως *"την επιστήμη της αναλυτικής συλλογιστικής που διευκολύνεται από αλληλεπιδραστικές οπτικές διεπιφάνειες"*, με περιοχές ενδιαφέροντος:

- τις τεχνικές αναλυτικής συλλογιστικής για την υποστήριξη σχεδιασμού και λήψης αποφάσεων
- τις οπτικές αναπαραστάσεις και τις τεχνικές αλληλεπίδρασης που επιτρέπουν την ταχεία επισκόπηση, εξερεύνηση και κατανόηση πολλών πληροφοριών ταυτόχρονα, αξιοποιώντας τις αντιληπτικές ικανότητες του ανθρώπινου οπτικού συστήματος
- τις αναπαραστάσεις και τους μετασχηματισμούς δεδομένων, για την υποστήριξη των σκοπών οπτικοποίησης και ανάλυσης
- τις τεχνικές που επιτρέπουν την επικοινωνία των αποτελεσμάτων της αναλυτικής διαδικασίας, παρέχοντας το κατάλληλο πλαίσιο κατανόησης των παρουσιαζόμενων πληροφοριών.

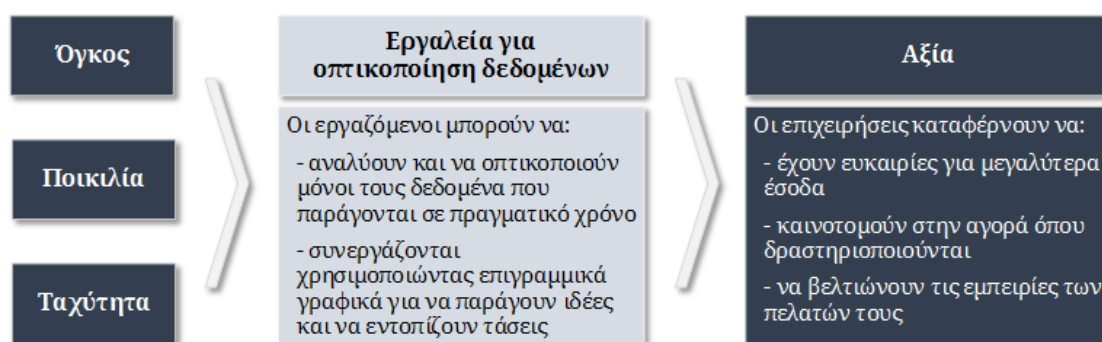
Η Οπτική Ανάλυση Δεδομένων αποτελεί μία από τις εξειδικευμένες λύσεις για την ανάλυση ΜΔ, όπως αναφέρθηκε στο Κεφάλαιο 2. Τα σύγχρονα συστήματα επιχειρηματικής ευφυΐας παρέχουν λειτουργικότητες οπτικοποίησης δεδομένων για να διευκολύνουν τους χρήστες στη διερεύνηση των δεδομένων και εν τέλει στην ανακάλυψη γνώσης.

Για την Tableau⁴⁶, η Οπτική Ανάλυση Δεδομένων είναι *"η εξερεύνηση των δεδομένων με οπτικό τρόπο"*, που επιτρέπει στον χρήστη (1) να μεταβάλλει την επιλογή των δεδομένων προς ανάλυση, προκειμένου να μπορεί να διερευνήσει

⁴⁶ Εταιρεία που έχει αναπτύξει λογισμικό οπτικοποίησης, το οποίο συγκαταλέγεται στις κορυφαίες πλατφόρμες επιχειρηματικής ευφυΐας και ανάλυσης (Sallam et al., 2014).

διαφορετικές πτυχές του εξεταζόμενου ζητήματος, και (2) να εναλλάσσει μεταξύ διαφορετικών οπτικοποιήσεων δεδομένων, επιλέγοντας αυτή που είναι κατάλληλη για την απάντηση του κατά περίπτωση ερωτήματος (Hanrahan et al., 2007: 3). Ο συνδυασμός αυτών των δυο δυνατοτήτων δημιουργεί τις προϋποθέσεις ώστε ακόμη και μη εξοικειωμένοι χρήστες να μπορούν να εντοπίσουν γρήγορα απαντήσεις στα ερωτήματα που θέτουν σχετικά με τα εξεταζόμενα δεδομένα (Hanrahan et al., 2007: 3). Συμβουλεύει δε τους οργανισμούς που επιδιώκουν να αποκομίσουν οφέλη από τα ΜΔ να υιοθετήσουν αυτό που αποκαλεί ως *"Κύκλο της Οπτικής Ανάλυσης Δεδομένων"* (Visual Analysis Cycle): πρόκειται για μία επαναληπτική διαδικασία που ξεκινά με την πρόσβαση στα δεδομένα, συνεχίζει με την οπτική τους αναπαράσταση, από την οποία προκύπτουν απαντήσεις ή/και νέα ερωτήματα που τροφοδοτούν έναν νέο κύκλο ανάλυσης (Jewett, 2014: 6).

Η Intel (Intel IT Center, 2013: 6) υποστηρίζει ότι η χρήση εργαλείων για την ανακάλυψη γνώσης που βασίζονται στην παροχή λειτουργικότητας οπτικοποίησης δεδομένων, μπορεί να βοηθήσει τους οργανισμούς να επωφεληθούν στο έπακρο από τα 3Vs⁴⁷. Αποτυπώνει μάλιστα σχηματικά τη συμβολή τους στην παραγωγή αξίας για τους οργανισμούς που τα εντάσσουν, όπως φαίνεται στην ακόλουθη εικόνα (Εικόνα 19):



Εικόνα 19: Η Αξία ως το 4ο "V" των ΜΔ για τους οργανισμούς και τους εργαζόμενους σε αυτούς, όπως αποτυπώνεται σχηματικά από Intel (Intel IT Center, 2013: 6)

⁴⁷ Δηλ. τις βασικές ιδιότητες των ΜΔ (βλ. Κεφάλαιο 2, Ενότητα 2.1).

Αν και είναι κατανοητό ότι η Οπτική Ανάλυση Δεδομένων αποτελεί σημαντικό εργαλείο, ο τρόπος που παρουσιάζεται από εταιρείες όπως οι δύο που προαναφέρθηκαν ενδέχεται να εμπεριέχει μεροληψία καθώς και οι δύο δραστηριοποιούνται στον τομέα της παροχής λύσεων Οπτικής Ανάλυσης Δεδομένων.

Μία από τις επιφυλακτικές απόψεις σχετικά με τη χρήση εργαλείων Οπτικής Ανάλυσης σε σύνολα ΜΔ εκφράζεται από τον Davenport (2014: 94), ο οποίος θεωρεί ότι η Οπτική Ανάλυση Δεδομένων πρέπει να περιορίζεται στη γενική παρουσίαση για τον σχηματισμό μίας πρώτης άποψης και να μην χρησιμοποιείται για τη διενέργεια προβλέψεων ή τη λήψη αποφάσεων, διότι δεν δύναται να εκφράσει σύνθετες πολυμεταβλητές σχέσεις ή στατιστικά μοντέλα. Εντούτοις, αξίζει να επισημανθεί ότι αναγνωρίζει ότι η Οπτική Ανάλυση Δεδομένων έλκει την προσοχή και είναι ερμηνεύσιμη ακόμη και από άτομα που δεν διαθέτουν γνώσεις ποσοτικών μεθόδων (Davenport, 2014: 94). Τελικώς, αποδέχεται ότι οι επιστημονικοί υπεύθυνοι δεδομένων, οι οποίοι είναι οι καθ' ύλην αρμόδιοι για την άντληση γνώσης από ΜΔ, θα πρέπει να διαθέτουν βασικές έστω γνώσεις για την οπτική απεικόνιση δεδομένων και αναλύσεων, προφανώς για να μπορούν να επικοινωνούν τα ευρήματά της εργασίας τους σε κοινό εκτός των ιδίων (Davenport, 2014: 96).

Εκτιμάται ότι πιο μετριοπαθείς προσεγγίσεις, όπως αυτές των Fekete et al. (2008: 14) που αναδεικνύουν την αναγκαιότητα κριτικής επιλογής των μεθόδων ανάλυσης βάσει των χαρακτηριστικών του υπό εξέταση συνόλου δεδομένων, είναι πιο ρεαλιστικές. Θα μπορούσε μάλιστα να υποστηριχθεί η άποψη ότι τα εργαλεία Οπτικής Ανάλυσης Δεδομένων, με τις αυξημένες δυνατότητες αλληλεπίδρασης που παρέχουν, θα πρέπει να χρησιμοποιούνται συνδυαστικά με τις υπόλοιπες τεχνικές υπολογιστικής ανάλυσης, συμφωνώντας με τους Kerren και Schreiber (2012: 1, 11) ότι *"θα πρέπει να υποστηρίζουν τους αναλυτές παρέχοντας έναν έξυπνο σχεδιασμό αλληλεπιδραστικών οπτικοποιήσεων και μεθόδων αυτόματης ανάλυσης"*, συμβάλλοντας έτσι στη διαχείριση της υπερπληθώρας διαθέσιμης πληροφορίας.

Η σύγχρονη περίοδος της οπτικοποίησης καθορίζεται από τις εξελίξεις στους ηλεκτρονικούς υπολογιστές και τα γραφικά υπολογιστών. Εντούτοις, όπως προκύπτει από την ιστορική αναδρομή, αποτελεί την πλέον πρόσφατη φάση εξέλιξης μίας μακροχρόνιας παράδοσης στην κατασκευή οπτικών αναπαραστάσεων δεδομένων.

Κατά την εξέταση ορισμών που έχουν διατυπωθεί υπό το πρίσμα των πρόσφατων τεχνολογικών εξελίξεων, εντοπίστηκε ο διαχωρισμός μεταξύ επιστημονικής οπτικοποίησης και οπτικοποίησης πληροφορίας. Αυτός ο διαχωρισμός έχει ιστορικές καταβολές και αντανακλά τα διαφορετικά υπόβαθρα αυτών που τους διατυπώνουν. Βασίζεται δε στην προέλευση των δεδομένων ή/και στον σκοπό αναπαράστασης αυτών. Ο διαχωρισμός αυτός τείνει να εκλείψει, καθώς η συγκέντρωση ΜΔ και οι τεχνικές δυνατότητες επεξεργασίας δομημένων και μη δεδομένων, συγχέουν τα όρια.

Επιπρόσθετα, ένα σύνολο δεδομένων μπορεί να επιδέχεται ρεαλιστικής και αφαιρετικής οπτικής αναπαράστασης. Η επιλογή του κατάλληλου τύπου οπτικής αναπαράστασης που εξυπηρετεί τον σκοπό της οπτικοποίησης αποτελεί σημαντική σχεδιαστική απόφαση, για τη λήψη της οποίας συνεκτιμώνται διάφοροι παράγοντες. Βασικά ζητήματα της σχεδιαστικής διαδικασίας μετατροπής των άμορφων, αφηρημένων δεδομένων σε οπτικά αντιληπτή μορφή παρουσιάζονται στο επόμενο κεφάλαιο.

4 Η οπτικοποίηση δεδομένων ως δημιούργημα: σχεδιασμός από ανθρώπους για ανθρώπους

Στο παρόν κεφάλαιο, εξετάζονται ζητήματα που σχετίζονται με την οπτικοποίηση δεδομένων, με εστίαση στην οπτική του σχεδιαστή.

Μετά τις απαραίτητες εισαγωγικές επισημάνσεις, παρουσιάζονται ζητήματα που αφορούν στις σχεδιαστικές αποφάσεις που λαμβάνονται στο πλαίσιο του οπτικού σχεδιασμού μίας οπτικοποίησης δεδομένων. Εστιάζοντας στις οπτικοποιήσεις που έχουν κατασκευαστεί ως διεπαφές παρουσίασης των δεδομένων με σκοπό την ανακάλυψη γνώσης, εντοπίζονται τρόποι με τους οποίους αυτές μπορούν να εμπλουτιστούν, προκειμένου να παρέχουν το κατάλληλο ερμηνευτικό πλαίσιο στον χρήστη και να ενθαρρύνουν τον διάλογό του με τα υποκείμενα δεδομένα.

4.1 Εισαγωγικές επισημάνσεις

Η οπτικοποίηση δεδομένων χρησιμοποιείται είτε για σκοπούς ανάλυσης, όταν ο χρήστης επιθυμεί να μελετήσει τα δεδομένα, είτε για σκοπούς παρουσίασης προκειμένου να τα επικοινωνήσει σε άλλους. Στην πρώτη περίπτωση επιδιώκεται η κατανόηση των πληροφοριών που εμπεριέχεται στα υποκείμενα δεδομένα, ενώ στη δεύτερη η μετάδοση αυτής της πληροφορίας με ευνόητο τρόπο (Cleveland, 1994: 1· Bertin, 2011: 419).

Οι σημερινές τεχνολογικές δυνατότητες καθιστούν εφικτή τη δημιουργία οπτικοποιήσεων που επιτρέπουν την εκτεταμένη αλληλεπίδραση του χρήστη με τα υποκείμενα δεδομένα, σε πρωτογενή ή επεξεργασμένη μορφή. Κατ' επέκταση, διακρίνεται μία τρίτη, υβριδική χρήση της οπτικοποίησης, που προκύπτει από τον συνδυασμό των δύο παραπάνω και συνίσταται στον εμπλουτισμό της οπτικής παρουσίασης του αποτελέσματος ανάλυσης με δυνατότητες

αλληλεπίδρασης. Με αυτόν τον τρόπο, ο χρήστης, εφόσον είναι διαφορετικό άτομο από τον αναλυτή, εξερευνά και αναλύει περαιτέρω τα υποκείμενα δεδομένα, με δυνητικό αποτέλεσμα την ανακάλυψη γνώσης.

Τα ΜΔ ενδείκνυνται για τέτοιου είδους οπτικοποιήσεις. Ο Ohlhorst (2013: 121-122) υποστηρίζει ότι η παρουσίαση της ανάλυσης ενός συνόλου ΜΔ, η οποία διευκολύνει την κατανόηση και την ερμηνεία των αποτελεσμάτων, πρέπει να περιλαμβάνει οπτικοποιήσεις *"βασισμένες σε αλληλεπιδραστικές πηγές που θα επιτρέπουν στους χρήστες να επιλέγουν και να επαναπροσδιορίζουν τα παρουσιαζόμενα στοιχεία"*. Για τον σκοπό αυτό, η διεπιφάνεια θα πρέπει να παρέχει δυνατότητες πειραματισμών που δοκιμάζουν την ευστάθεια θεωριών, επανεμφάνισης στοιχείων που έχουν αποκρυφθεί, διενέργειας υποθετικών σεναρίων αλλά και πρόσβασης στην πηγή των δεδομένων, ώστε *"οι χρήστες να έχουν τη δυνατότητα όχι μόνο να βλέπουν τα αποτελέσματα [της ανάλυσης] αλλά και να κατανοούν γιατί βλέπουν αυτά τα αποτελέσματα"* (Ohlhorst, 2013: 122).

Παλαιότερα, η οπτικοποίηση θεωρείτο θέμα χαμηλής σπουδαιότητας, ιδιαίτερα από άτομα με εξειδίκευση στις ποσοτικές μεθόδους ανάλυσης, που δεν θεωρούσαν κρίσιμο το να αφιερώσουν χρόνο και προσπάθεια για την επικοινωνία τους διότι πίστευαν πως *"τα αποτελέσματα μιλούν από μόνα τους"* (Davenport and Kim, 2013: 95-98). Πλέον όμως η οπτικοποίηση, είτε ως μέσο ανάλυσης, είτε -ακόμη περισσότερο- ως μέσο επικοινωνίας των ευρημάτων ανάλυσης ΜΔ, θεωρείται σημαντικό μέρος της διαδικασίας που αποσκοπεί στην άντληση αξίας από τα ΜΔ.

Αναγνωρίζοντας τη σημασία της οπτικοποίησης δεδομένων στην εξαγωγή αξίας από τα ΜΔ και προκειμένου να εντοπιστούν βέλτιστες και τρέχουσες πρακτικές, επιλέγεται εφεξής η εξέταση σχεδιαστικών αποφάσεων που λαμβάνονται στο πλαίσιο δημιουργίας σύγχρονων αλληλεπιδραστικών οπτικοποιήσεων από ανθρώπους για ανθρώπους, με σκοπό την ανακάλυψη γνώσης που δεν διέθεταν πριν εκτεθούν στην οπτικοποίηση.

4.2 Οπτικός σχεδιασμός

Οι οπτικές αναπαραστάσεις δεδομένων, δεδομένου ότι αποσκοπούν στη διεύρυνση της νόησης, δεν περιορίζονται στην παροχή μίας ευχάριστης οπτικής εμπειρίας. Η σχεδιαστική διαδικασία μετατροπής των άμορφων, αφηρημένων δεδομένων σε αισθητηριακά αντιληπτή μορφή περιλαμβάνει διάφορα στάδια, από την απόκτηση εξειδικευμένων γνώσεων σχετικά με το προς οπτικοποίηση σύνολο δεδομένων μέχρι και τη διατύπωση του σκοπού της οπτικοποίησης. Τα στάδια αυτά προηγούνται και αποτελούν τη βάση για την επιλογή εκείνων των οπτικών κωδικοποιήσεων που μπορούν να εξυπηρετήσουν τον σκοπό της οπτικοποίησης.

Όπως άλλωστε σημειώνουν οι Heer, Bostock και Ogievetsky (2010: 1), η σχεδιαστική διαδικασία εμπερικλείει μία σειρά αποφάσεων, όπως

"ο ορισμός των ερωτημάτων, ο προσδιορισμός των κατάλληλων δεδομένων και η επιλογή αποτελεσματικών οπτικών κωδικοποιήσεων για τη σύνδεση των τιμών των δεδομένων με γραφικά χαρακτηριστικά, όπως η θέση, το μέγεθος, το σχήμα και το χρώμα".

Γίνεται επομένως κατανοητό ότι η διατύπωση των ερωτημάτων και η επιλογή των συνόλων δεδομένων που μπορούν να τα απαντήσουν, προϋποθέτουν γνώση, τόσο του γενικότερου περιβάλλοντος από το οποίο έχουν συλλεχθεί τα δεδομένα προς οπτικοποίηση, όσο και των ειδικότερων λεπτομερειών για τις ιδιότητες τους. Βάσει των προαναφερθέντων και προτού εξεταστούν βασικά ζητήματα οπτικού σχεδιασμού για την κατασκευή οπτικοποιήσεων δεδομένων, κρίνεται σκόπιμη η αναφορά σε βασικές έννοιες που σχετίζονται με τα δεδομένα και τα είδη αυτών.⁴⁸

Κατ' αρχάς, ένα σύνολο δεδομένων (data set) είναι ένα σύνολο μετρήσεων από ένα περιβάλλον ή μία διαδικασία (Hand et al., 2001: 4) που αποτελείται από

⁴⁸ Στο πλαίσιο των στόχων της παρούσας εργασίας, περιγράφονται επιγραμματικά εκείνοι οι όροι που κρίνονται απολύτως απαραίτητοι για την κατανόηση του μηχανισμού επιλογής συγκεκριμένων μεθόδων οπτικοποίησης.

μεταβλητές, κάθε μια εκ των οποίων αφορά σε ένα εξεταζόμενο κριτήριο. Οι μεταβλητές διακρίνονται σε ποιοτικές (qualitative), οι οποίες αναφέρονται σε ποιοτικά χαρακτηριστικά που δεν αντιστοιχούν σε αριθμούς, και σε ποσοτικές (quantitative), οι οποίες μπορούν να απαριθμηθούν ή να μετρηθούν.

Οι μετρήσεις, δηλαδή οι τιμές που λαμβάνουν οι μεταβλητές, ταξινομούνται σε τέσσερις κλίμακες, ως ακολούθως (Αρμενάκης, 2011: 1-4):

- στις *ονομαστικές ή κατηγορικές* (nominal) κλίμακες, οι αριθμοί ή οι λέξεις περιγράφουν ή ταξινομούν ομάδες ανθρώπων, αντικειμένων κ.λπ., βάσει ποιοτικών χαρακτηριστικών, χωρίς όμως να ισχύει κάποιας μορφής διάταξη
- οι *τακτικές ή διατεταγμένες* (ordinal) κλίμακες είναι επίσης ποιοτικές, ωστόσο οι τιμές που λαμβάνουν εμπεριέχουν την έννοια της διάταξης, είτε ως διαδοχή είτε ως κατάταξη, χωρίς όμως οι διαφορές ή οι αποστάσεις μεταξύ των κατηγοριών να μπορούν να μετρηθούν
- οι *διαστημικές* (interval) κλίμακες είναι ποσοτικές, με τις διαφορές και τις αποστάσεις μεταξύ των κατηγοριών να είναι γνωστές και μετρήσιμες, ωστόσο η μέτρηση μηδέν ορίζεται αυθαίρετα και δεν συνεπάγεται έλλειψη χαρακτηριστικού
- οι *αναλογικές* (ratio) κλίμακες είναι επίσης ποσοτικές, με τη μέτρηση μηδέν να είναι απόλυτα ορισμένη και να σημαίνει έλλειψη χαρακτηριστικών.

Επισημαίνεται ότι οι ποιοτικές μεταβλητές μπορούν να λάβουν αριθμητικές τιμές κατά την καταχώριση σε λογισμικά επεξεργασίας δεδομένων. Για παράδειγμα, οι τιμές "άνδρας" ή "γυναίκα" στη μεταβλητή "φύλο" μπορεί να καταχωριστούν ως αριθμοί 1 ή 2 αντιστοίχως, χωρίς αυτό να σημαίνει ότι κατά την επεξεργασία και την ανάλυση αντιμετωπίζονται ως ποσοτικές.

Παράλληλα, οι μεταβλητές χαρακτηρίζονται ως διακριτές (discrete) όταν λαμβάνουν συγκεκριμένες, μεμονωμένες τιμές ή ως συνεχείς (continuous) όταν οι δυνατές τιμές είναι πραγματικοί αριθμοί.

Ανάλογα με τον τύπο των υπό εξέταση μεταβλητών, επιλέγεται η κατάλληλη μέθοδος ανάλυσης, από την οποία προκύπτουν απαντήσεις στα ερωτήματα που θέτει ο αναλυτής, καθώς και ευρήματα που κρίνονται άξια παρουσίασης σε άλλους πέραν αυτού. Αντίστοιχα, ο τύπος των υπό εξέταση μεταβλητών επηρεάζει την επιλογή του τρόπου οπτικοποίησης ενός συνόλου δεδομένων.

4.2.1 Συστατικά οπτικοποίησης

Μέσω της οπτικοποίησης δεδομένων, ο δημιουργός προσδίδει μορφή στα δεδομένα, αντιστοιχίζοντάς τα με οπτικά αντιληπτές μορφές. Από αυτές τις αντιστοιχίσεις, ο χρήστης μπορεί να αντλήσει πληροφορίες μέσω των νοητικών διεργασιών για να επιβεβαιώσει υπάρχουσα ή και για να αποκτήσει νέα γνώση σχετικά με τα αναπαριστώμενα υποκείμενα δεδομένα.

Ο Bertin, το έργο του οποίου αποτελεί σημείο αναφοράς για τη μελέτη των βασικών αρχών που διέπουν τις γραφικές αναπαραστάσεις, ορίζει αυτές ως *"τη μεταγραφή στο σύστημα γραφικών σημείων της 'πληροφορίας' που είναι γνωστή μέσω οποιουδήποτε συστήματος σημείων"* (Bertin, 2011: 4). Επισημαίνει δε ότι κατά την εξέταση των παραμέτρων που συνιστούν ένα γραφικό σύστημα, είτε στο πλαίσιο μελέτης είτε στο πλαίσιο σχεδιασμού, είναι κρίσιμο να εξεταστεί χωριστά το περιεχόμενο⁴⁹ από τις ιδιότητές του.

Επιπρόσθετα, εντοπίζει συγκεκριμένα στοιχεία που μπορεί να χαρακτηρίζουν ένα σημείο που αναπαριστά δεδομένα επί των δύο διαστάσεων ενός λευκού χαρτιού, όταν αυτό λαμβάνει τη μορφή μίας τελείας, μίας γραμμής ή ακόμη και μίας έκτασης (Bertin, 2011: 44). Συγκεκριμένα, υποστηρίζει ότι ο σχεδιαστής είναι σε θέση να δημιουργεί γραφικές αναπαραστάσεις συνδυάζοντας συγκεκριμένες *"οπτικές μεταβλητές"* με διαφορετικές ιδιότητες και δυνατότητες κωδικοποίησης της πληροφορίας: πρόκειται για τη θέση του σημείου στις δύο διαστάσεις του επιπέδου, καθώς και για διαφοροποιήσεις στο

⁴⁹ Που συνίσταται στη μεταδιδόμενη πληροφορία.

μέγεθος, στην τονικότητα, στην υφή, στο χρώμα, στον προσανατολισμό και στο σχήμα του εν λόγω στοιχείου (Bertin, 2011: 42).

Είναι μάλιστα πιθανό ένα στοιχείο (μία μεμονωμένη τιμή ή ένα σύνολο μετρήσεων μίας ποιοτικής ή ποσοτικής μεταβλητής) να μπορεί να αναπαρασταθεί με περισσότερες από μία οπτικές μεταβλητές. Γενικά, ο Bertin επισημαίνει ότι η επιλογή της κατάλληλης κατά περίπτωση οπτικής μεταβλητής προσδιορίζεται από το επίπεδο οργάνωσης κάθε στοιχείου προς οπτικοποίηση: (1) το ποιοτικό επίπεδο, το οποίο σχετίζεται με τις έννοιες της ομοιότητας ή της διαφοράς για την κατάταξη σε κατηγορίες, (2) το διατεταγμένο επίπεδο, το οποίο επιτρέπει την κατάταξη των στοιχείων σε μία συγκεκριμένη σειρά προκειμένου να είναι εφικτή η σύγκριση (μεγαλύτερο-μικρότερο) μεταξύ τους, και (3) το ποσοτικό επίπεδο, όπου χρησιμοποιούνται συγκεκριμένες μονάδες μέτρησης (Bertin, 2011: 36-38).

Εκτός από τις δύο διαστάσεις του επιπέδου που μπορούν να χρησιμοποιηθούν για την οπτική αναπαράσταση ενός στοιχείου ανεξαρτήτως επιπέδου οργάνωσής του, η επιλογή μίας ή κάποιων⁵⁰ εκ των υπόλοιπων οπτικών μεταβλητών συναρτάται με το επίπεδο οργάνωσης των στοιχείων (Bertin, 2011: 96). Για παράδειγμα, στην περίπτωση των ποσοτικών μεταβλητών, οι διαφοροποιήσεις στο μέγεθος του σημείου είναι κατάλληλες για να κωδικοποιήσουν τις ποσοτικές διακυμάνσεις. Αντίθετα, η χρήση του χρώματος δεν μπορεί να μεταδώσει εξίσου αποτελεσματικά ποσοτικές αυξομειώσεις· εντούτοις ενδείκνυται στην περίπτωση ποιοτικών χαρακτηριστικών, όπου το χρώμα μπορεί να χρησιμοποιηθεί, ώστε ο χρήστης να αντιληφθεί κατηγορίες σημείων με όμοια ή διαφορετικά χαρακτηριστικά.

Βάσει των προαναφερθέντων, συμπεραίνεται ότι οι τύποι δεδομένων περιορίζουν τις διαθέσιμες επιλογές οπτικής κωδικοποίησής τους. Όταν όμως για ένα επίπεδο οργάνωσης, είναι διαθέσιμες περισσότερες της μίας επιλογές,

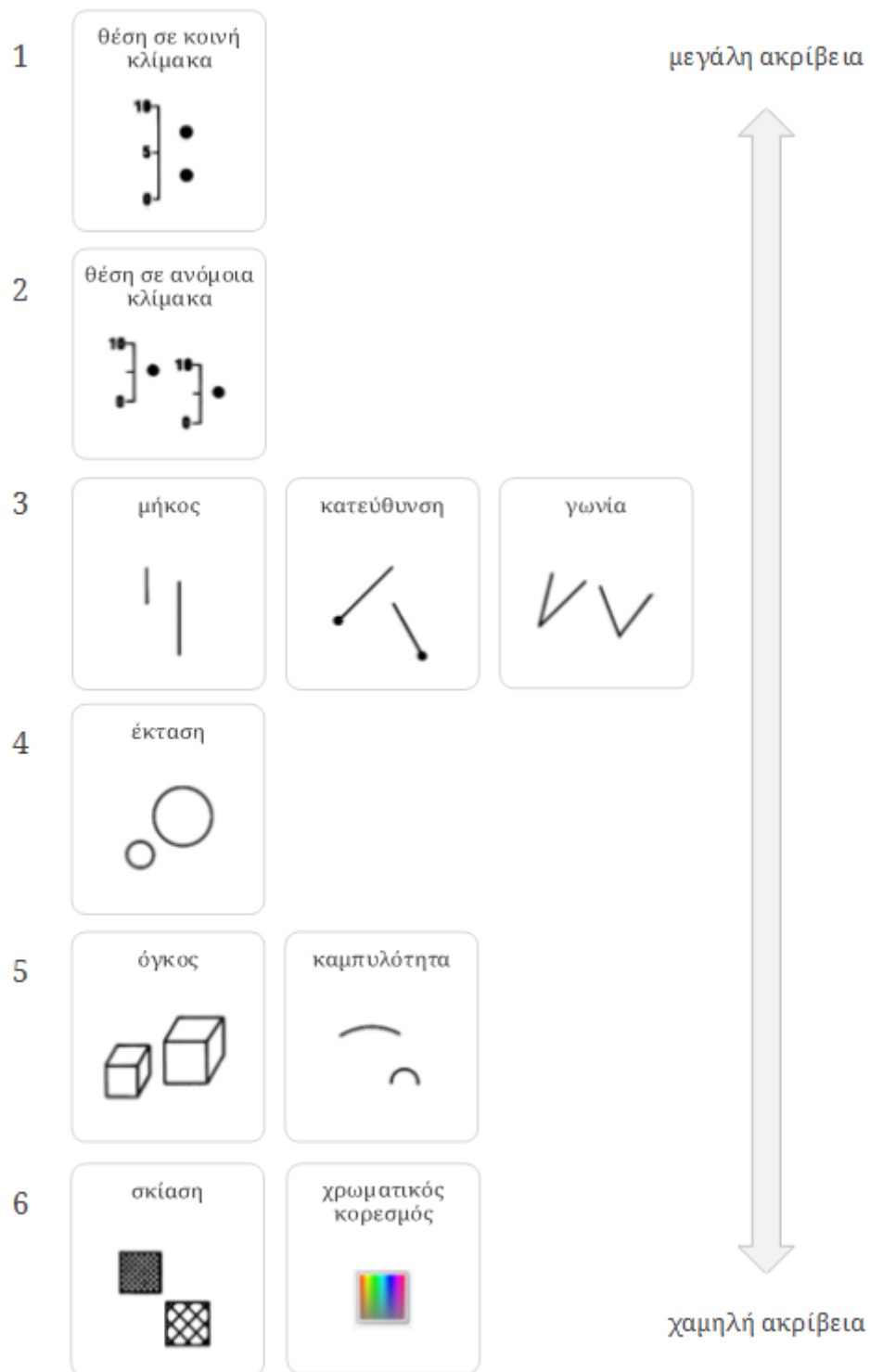
⁵⁰ Ανάλογα με το πλήθος των στοιχείων προς αναπαράσταση.

πώς επιλέγεται η κατάλληλη; Και, κατ' επέκταση, ποιος είναι ο κατάλληλος συνδυασμός οπτικών μεταβλητών όταν πρόκειται να οπτικοποιηθούν πολυδιάστατα δεδομένα;

Εδώ, ο Bertin προτείνει την έννοια της αποδοτικότητας, η οποία σχετίζεται με τον βαθμό ευκολίας ανάγνωσης μίας γραφικής αναπαράστασης, προφανώς βασιζόμενος στις δυνατότητες της ανθρώπινης αντίληψης. Μία σύνθεση των οπτικών μεταβλητών χαρακτηρίζεται ως πιο αποδοτική σε σχέση με μία άλλη όταν *"προκειμένου να ληφθεί μία σωστή και πλήρης απάντηση σε μία δεδομένη ερώτηση, (...), μία κατασκευή απαιτεί λιγότερο χρόνο παρατήρησης από μία άλλη κατασκευή"* (Bertin, 2011: 139).

Αναγνωρίζοντας επίσης τις δυνατότητες και τους περιορισμούς της ανθρώπινης αντίληψης και αναζητώντας παράλληλα ένα επιστημονικό υπόβαθρο για τις μεθόδους γραφικής αναπαράστασης στο πλαίσιο ανάλυσης και παρουσίασης δεδομένων, οι Cleveland και McGill προχωρούν σε μία αρχική καταγραφή *στοιχειωδών αντιληπτικών διεργασιών* (elementary perceptual tasks), οι οποίες -μεμονωμένα ή σε συνδυασμό- επιτελούνται στο μυαλό του χρήστη για την άντληση ποσοτικών πληροφοριών από γραφικές αναπαραστάσεις: αυτές αφορούν στην αντίληψη της θέσης, του μήκους, της κατεύθυνσης, της γωνίας, της έκτασης, του όγκου, της καμπυλότητας, της σκίασης και του χρωματικού κορεσμού των στοιχείων ενός γραφήματος για την αποκωδικοποίηση της αντίστοιχης κωδικοποιημένης ποσοτικής πληροφορίας (Cleveland and McGill, 1984: 531-532).

Επιχειρούν μάλιστα να κατατάξουν τις διεργασίες ως προς την ακρίβεια αποκωδικοποίησης της πληροφορίας από τους χρήστες, όπως φαίνεται στην Εικόνα 20, βασιζόμενοι σε προσωπικούς τους πειραματισμούς με γραφήματα, και στην ψυχοφυσική θεωρία και σε σχετικά πειραματικά δεδομένα (Cleveland and McGill, 1984: 536-537).



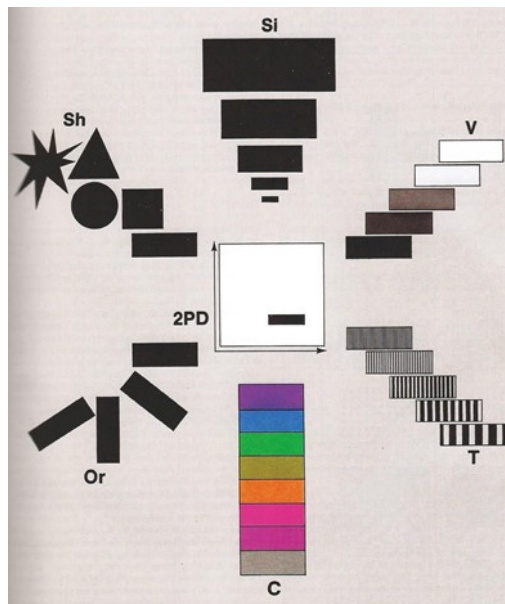
Εικόνα 20: Απεικόνιση κατάταξης στοιχειωδών αντιληπτικών εργασιών βάσει ακριβείας, κατά την αρχική εκτίμηση των Cleveland και McGill (1984: 536)⁵¹

⁵¹ Εμπνευσμένη από αντίστοιχη απεικόνιση του Mackinlay (1986: 125), χρησιμοποιώντας τα πρωτότυπα ασπρόμαυρα εικονίδια των Cleveland και McGill (1984: 532)

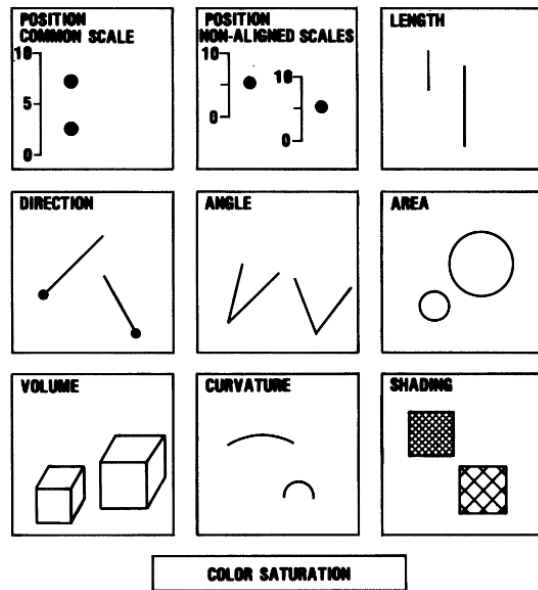
Λίγα χρόνια αργότερα, ο Mackinlay (1986: 110) επιχειρεί να προδιαγράψει ένα σύστημα που θα βασίζεται σε τεχνολογία Τεχνητής Νοημοσύνης και βάσει αυτής θα παράγει αυτόματα γραφικές αναπαραστάσεις κατάλληλες για τα δεδομένα προς παρουσίαση. Υπό αυτό το πρίσμα, επιχειρεί να επεκτείνει την κατάταξη των γνωστικών εργασιών των Cleveland και McGill, προσθέτοντας επιπλέον κωδικοποιήσεις και διακρίνοντας διαφοροποιήσεις ανάλογα με τον τύπο των δεδομένων προς αναπαράσταση (Mackinlay, 1986: 124-125 και Εικόνα 23).

Η προβληματική για την αξιολόγηση μίας οπτικής αναπαράστασης φαίνεται να απασχολεί και τον Mackinlay. Συγκεκριμένα, διατυπώνει την άποψη ότι ο προσδιορισμός της αποτελεσματικότητας (effectiveness) μίας κωδικοποίησης έναντι μίας άλλης είναι εφικτός μέσω μιας προσέγγισης που συνδυάζει την προσωπική αίσθηση του ερευνητή με την εκάστοτε διαθέσιμη τεκμηριωμένη γνώση για τις δυνατότητες της ανθρώπινης αντίληψης (Mackinlay, 1986: 124). Διαχωρίζει δε σαφώς την εκφραστικότητα (expressiveness) μίας οπτικής αναπαράστασης, η οποία αντανακλά το σχεδιαστικό όραμα του δημιουργού και βασίζεται στις συμβάσεις της οπτικής επικοινωνίας, από την αποτελεσματικότητά της, η οποία εξαρτάται επιπλέον από τις αντιληπτικές ικανότητες του χρήστη (Mackinlay, 1986: 124).

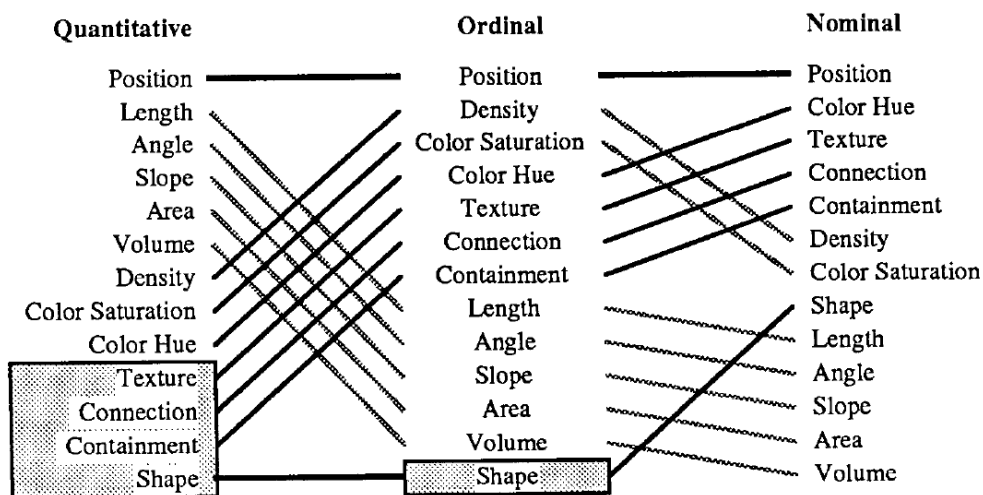
Στις επόμενες εικόνες παρατίθενται οι πρωτότυπες εικονογραφήσεις των δομικών στοιχείων με τις οποίες πλαισιώνουν τα κείμενά τους οι Bertin, Cleveland και McGill, και Mackinlay αντιστοίχως. Αξίζει να επισημανθεί μία λεπτομέρεια που είναι μάλλον ενδεικτική της πίστης των εν λόγω συγγραφέων στη χρήση της οπτικής αναπαράστασης πληροφοριών για τη διευκόλυνση των νοητικών διεργασιών του χρήστη: και στις τρεις περιπτώσεις επιλέγουν να απεικονίσουν σχηματικά τις ιδέες τους, είτε για να συνοψίσουν την αναλυτικά διατυπωθείσα θεωρία τους (Εικόνα 21), είτε για να αποφύγουν να αναλωθούν σε περίπλοκες λεκτικές διατυπώσεις για ζητήματα που γίνονται άμεσα αντιληπτά όταν παρουσιάζονται με γραφικό τρόπο (Εικόνα 22 και Εικόνα 23).



Εικόνα 21: Οπτικές μεταβλητές
(Πηγή: Bertin, 2011: 43)



Εικόνα 22: Στοιχειώδεις αντιληπτικές εργασίες
(Πηγή: Cleveland and McGill, 1984: 532)



Εικόνα 23: Αντιληπτικές εργασίες ανά είδος μεταβλητής
(Πηγή: Mackinlay, 1984: 125)

Οι προαναφερθείσες κατηγοριοποιήσεις διατυπώθηκαν σε μία εποχή που οι δυνατότητες των ηλεκτρονικών υπολογιστών να παράγουν γραφικές αναπαραστάσεις ήταν ακόμη σχετικά περιορισμένες, καίτοι ελπιδοφόρες, κρίνοντας από την επιδίωξη του Mackinlay να παραμετροποιήσει, στις αρχές της δεκαετίας του 1980, τη διαδικασία δημιουργίας οπτικών αναπαραστάσεων για την αυτοματοποίηση της παραγωγής βέλτιστων οπτικοποιήσεων. Κατά την ίδια

περίοδο, γράφει ο Bertin στην αγγλική έκδοση του *"Sémiologie Graphique"*⁵²:

"τα σύγχρονα γραφικά περιλαμβάνουν τη μεταμορφώσιμη και επαναδιατάξιμη εικόνα. Αυτό είναι το αυστηρό ερευνητικό εργαλείο που επιτρέπει σε αυτόν που λαμβάνει αποφάσεις να ανακαλύπτει τι πρέπει να ειπωθεί και τι πρέπει να γίνει. Ο υπολογιστής (...) θα βρει το πιο πλήρες και ισχυρό μέσο έκφρασης στα σύγχρονα γραφικά. Επιπλέον, αυτή θα λάβει τη μορφή της απλούστερης και επομένως πιο επικοινωνίσιμης εικόνας" (Bertin, 2011: xv).

Θα μπορούσε άραγε να υποτεθεί ότι οι αρχές αυτές παραμένουν επίκαιρες; Σε ένα από τα πλέον πρόσφατα εκδοθέντα βιβλία σχετικά με την κατασκευή μεστών οπτικοποιήσεων, ο Yau αναγνωρίζει ότι μεγάλο μέρος της σύγχρονης έρευνας για την οπτικοποίηση εξακολουθεί να παραπέμπει στο κείμενο των Cleveland και McGill (Yau, 2013: 104, 93).

Παράλληλα, παρέχει τη δική του εκδοχή για τα συστατικά μίας οπτικής αναπαράστασης δεδομένων, βασιζόμενος ρητά στους Bertin και Wilkinson (Yau, 2013: 104, 93).⁵³ Σύμφωνα λοιπόν με τη δική του προσέγγιση (Yau, 2013: 93-113), οι οπτικοποιήσεις αναλύονται σε τέσσερις βασικές κατηγορίες συστατικών μερών: (1) στις οπτικές ενδείξεις (visual cues), (2) στο σύστημα συντεταγμένων, (3) στην κλίμακα, και (4) στο γενικό ερμηνευτικό πλαίσιο. Οι οπτικές ενδείξεις συνίστανται (ibid): (1) στη θέση, (2) στο μήκος, (3) στη γωνία, (4) στην κατεύθυνση, (5) στην έκταση, (6) στον όγκο, (7) στη χρωματική απόχρωση, και (8) στον χρωματικό κορεσμό.

Αυτές οι οπτικές ενδείξεις τοποθετούνται σε ένα σύστημα συντεταγμένων (καρτεσιανό, πολικό ή γεωγραφικό), το οποίο προσδιορίζει τις διαστάσεις της

⁵² Κυκλοφόρησε στις αρχές της δεκαετίας του 1980.

⁵³ Ο Yau (2013: 93) αναφέρει ότι ο Wilkinson, στο βιβλίο *The Grammar of Graphics*, παρείχε μία παραλλαγή των οπτικών μεταβλητών, αντλώντας στοιχεία από το έργο του Bertin. Δεν κρίνεται σκόπιμη η περιγραφή της εν λόγω προσέγγισης, ωστόσο εκτιμάται ότι πιθανότατα αναφέρεται στον πίνακα καταγραφής αισθητικών ιδιοτήτων ανά γεωμετρία (Wilkinson, 2005: 317), ο οποίος θα μπορούσε να αντιδιασταλεί με τον συγκεντρωτικό πίνακα ιδιοτήτων των οπτικών μεταβλητών ανά επίπεδο οργάνωσης του Bertin (2011: 96).

οπτικοποίησης. Βάσει των επιλεχθέντων κλιμάκων (γραμμικής, λογαριθμικής, κατηγορικής, τακτικής, ποσοστιαίας ή χρονικής) και συστήματος συντεταγμένων, προκύπτει η τοποθέτηση κάθε οπτικής ένδειξης.

Τέλος, ιδιαίτερη σημασία προσδίδει στη δημιουργία από τον σχεδιαστή ενός γενικού ερμηνευτικού πλαισίου, που παρέχει στον χρήστη πληροφορίες σχετικά με τα δεδομένα που αναπαρίστανται -στην πιο απλή του μορφή, αυτό συνίσταται στην ονομασία των αξόνων και στον προσδιορισμό των μονάδων μέτρησης.

Όπως προκύπτει από τα παραπάνω, οι *οπτικές ενδείξεις* του Yau έχουν κοινά στοιχεία με τις *οπτικές μεταβλητές* του Bertin και με τις *στοιχειώδεις αντιληπτικές εργασίες* των Cleveland και McGill. Βάσει των προαναφερθέντων και δεδομένου ότι κάθε οπτικό μήνυμα προκύπτει από το συνδυασμό της τονικότητας, της γραμμής, του σχήματος, της κατεύθυνσης, της υφής, της κλίμακας, της διάστασης και της κίνησης, στοιχείων που είναι οικεία και αναγνωρίσιμα μέσω των ανθρώπινων αντιληπτικών ικανοτήτων (Dondis, 1973: 20-21), εξάγεται το συμπέρασμα ότι η ανάλυση μίας οπτικής αναπαράστασης δεδομένων σε συστατικά στοιχεία, ανεξαρτήτως της ονομασίας που τους αποδίδεται, παραπέμπει στα βασικά στοιχεία οπτικής επικοινωνίας.

Μετά τον εντοπισμό των επιμέρους οπτικών στοιχείων που συνιστούν μία οπτική αναπαράσταση δεδομένων σε δομικό επίπεδο, στη συνέχεια περιγράφεται πώς ο σχεδιαστής μπορεί να χρησιμοποιήσει συνδυασμούς αυτών για να οπτικοποιήσει τα δεδομένα με τρόπο κατανοητό από τον χρήστη.

4.2.2 Μέθοδοι οπτικοποίησης

Ανάλογα με τον σκοπό και το διαθέσιμο σύνολο δεδομένων, επιλέγεται η κατάλληλη μέθοδος οπτικοποίησης. Κάθε μέθοδος προκύπτει από τον συνδυασμό μεμονωμένων οπτικών κωδικοποιήσεων, όπως περιγράφηκαν πιο πάνω.

Η συνεχώς αυξανόμενη συγκέντρωση ΜΔ και η ανάγκη εξαγωγής γνώσης από αυτά, αναζωογόνησαν το πεδίο της οπτικοποίησης δεδομένων, με τα

σύγχρονα τεχνολογικά μέσα (υλισμικό και γλώσσες προγραμματισμού) να παρέχουν τη δυνατότητα δημιουργίας νέων ειδών οπτικοποίησης ικανών να αναπαραστήσουν πιο σύνθετα σύνολα δεδομένων. Σε αυτό το πλαίσιο, οι Heer et al. (2010) καταγράφουν μεθόδους οπτικοποίησης δεδομένων, εστιάζοντας σε αυτές που απαντούν στις σύγχρονες ανάγκες διερεύνησης και ανάλυσης περίπλοκων συνόλων δεδομένων απευθυνόμενες σε διαφορετικά και ετερογενή κοινά. Επισημαίνουν ωστόσο ότι, παρά την ποικιλία τους, όλες μοιράζονται ένα *"κοινό DNA"* που συνίσταται σε *"ένα σύνολο αντιστοιχίσεων μεταξύ των ιδιοτήτων των δεδομένων και οπτικών χαρακτηριστικών"*, (2010: 1-2).

Λίγα χρόνια αργότερα, ο Kirk (2012) κατηγοριοποιεί τις μεθόδους οπτικοποίησης,⁵⁴ ανάλογα με τον σκοπό που μπορεί να εκπληρωθεί από τη χρήση μίας οπτικής αναπαράστασης δεδομένων. Συγκεκριμένα, διακρίνει τις ακόλουθες κατηγορίες σκοπών (Kirk, 2012: 120):

1. τη διευκόλυνση της σύγκρισης κατηγορικών μεταβλητών
2. την εξέταση ιεραρχικών σχέσεων ή σχέσεων μέρους προς το σύνολο
3. τον εντοπισμό μεταβολών σε τάσεις και μοτίβα, εντός συγκεκριμένου χρονικού πλαισίου
4. την εξέταση των σχέσεων, των κατανομών και των μοτίβων που υπάρχουν σε πολυμετάβλητα σύνολα δεδομένων
5. την παρουσίαση συνόλων δεδομένων με γεωχωρικές ιδιότητες, χρησιμοποιώντας μία από τις διαθέσιμες μεθόδους χαρτογράφησης.

Η προσέγγισή του αυτή παρουσιάζει ιδιαίτερο ενδιαφέρον, καθώς εκτός από την ταξινόμηση των μεθόδων ανά σκοπό, αναφέρει τους τύπους και το πλήθος μεταβλητών που μπορούν να οπτικοποιηθούν ανά μέθοδο, καθώς και τις οπτικές μεταβλητές που συνθέτουν κάθε μέθοδο οπτικοποίησης. Στον ακόλουθο πίνακα (Πίνακας 6) καταγράφονται οι 45 μέθοδοι οπτικοποίησης που εντοπίζει, καθώς και οι οπτικές μεταβλητές που αποδίδει σε κάθε μια από αυτές.

⁵⁴ Καταγράφει, μέσω παρατήρησης, τις συχνότερα χρησιμοποιούμενες κατασκευές.

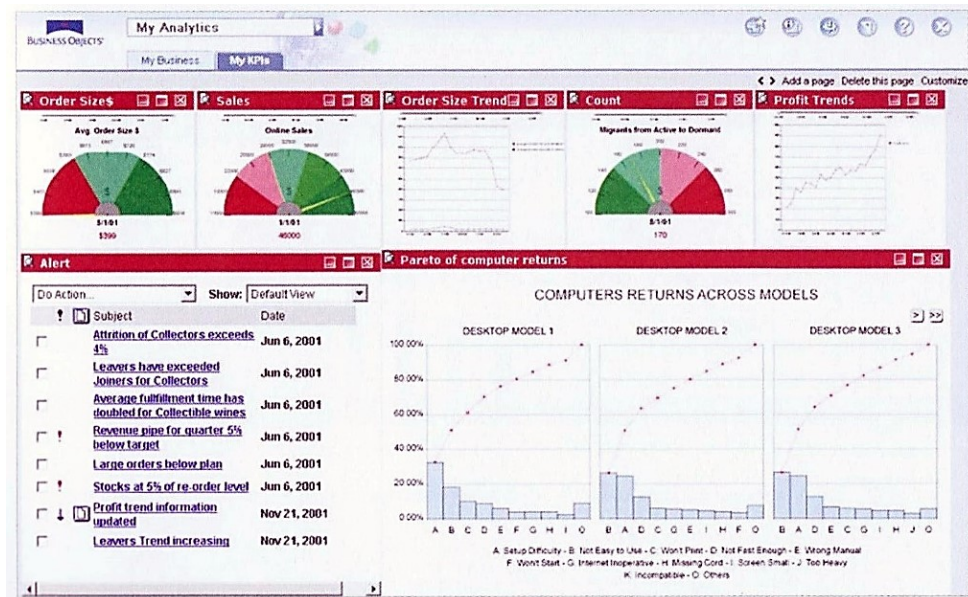
Πίνακας 6: Καταγραφή μεθόδων οπτικοποίησης κατά Kirk (2012: 120-158)

ταξινόμηση μεθόδου	μέθοδος οπτικοποίησης	οπτικές μεταβλητές												
		θέση	έκταση	γωνία	κλίση	χρωματική απόχρωση	χρωματικός κορεσμός	χρωματική φωτεινότητα	χρωματική σκουρότητα	υφή	σύμβολο	μήκος	ύψος	πλάτος
σύγκριση κατηγοριών (1)	Dot plot	■				■					■			
	Bar chart					■						■	■	
	Floating bar	■										■		
	Pixelated bar chart					■					■			
	Histogram											■	■	
	Slopegraph	■				■								■
	Radial chart	■				■	■	■		■				
	Glyph chart	■				■								
	Sankey diagram	■				■						■	■	
	Area size chart		■			■								■
	Small multiples	■	(οποιαδήποτε οπτική μεταβλητή)											
	Word cloud													■
ιεραρχίες & σχέσεις μέρους προς σύνολο (2)	Pie chart		■	■		■								
	Stacked bar chart	■				■	■	■				■		
	Square pie	■				■				■				
	Tree map	■	■			■	■	■						
	Circle packing diagram	■	■			■								
	Bubble hierarchy	■	■			■								
διαχρονικές αλλαγές (3)	Tree hierarchy	■	■	■		■								
	Line chart	■			■	■								
	Sparklines	■			■	■								
	Area chart		■		■	■						■	■	
	Horizon chart		■		■	■	■	■				■	■	
	Stacked area chart		■		■	■						■	■	
	Stream graph		■		■	■						■	■	
	Chandlestick chart	■			■	■						■	■	
συνδέσεις & σχέσεις (4)	Barcode chart	■			■	■				■				■
	Flow map	■			■	■						■	■	
	Scatter plot	■			■	■								
	Bubble-plot	■	■			■								
	Scatter plot matrix	■			■	■								
	Heatmap	■			■	■	■							
	Parallel sets	■			■	■						■	■	
	Radial network	■			■	■		■		■		■	■	
απεικόνιση γεωχωρικών δεδομένων (5)	Network diagram	■	■		■	■						■	■	
	Choropleth map	■				■	■							
	Dot plot map	■				■								
	Bubble plot map	■	■			■								
	Isarithmic map	■			■	■		■						
	Particle flow map	■				■							■	■
	Cartogram	■				■								■
	Dorling cartogram	■				■								■
	Network connection map	■				■							■	

Από αυτή την καταγραφή προκύπτει ότι κάθε μέθοδος αποτελεί συνδυασμό επιμέρους οπτικών κωδικοποιήσεων, κάθε μια από τις οποίες είναι κατάλληλη για συγκεκριμένους τύπους μεταβλητών. Ωστόσο, οι μέθοδοι αυτές μπορούν να χρησιμοποιούνται και συνδυαστικά: μία οπτικοποίηση δεδομένων μπορεί να αποτελεί σύνθεση περισσότερων της μίας μεθόδων οπτικοποίησης, προκειμένου να αναδείξει διαφορετικές πτυχές του υπό οπτικοποίηση συνόλου δεδομένων.

Μία εφαρμογή συνδυασμού διαφορετικών μεθόδων οπτικοποίησης, που συναντάται κυρίως σε λογισμικά επιχειρηματικής ευφυΐας, αποτελούν τα *dashboards*. Πρόκειται για τη συμπυκνωμένη προβολή οπτικών αναπαραστάσεων μεγάλου όγκου δεδομένων σε μία μόνο οθόνη, με σκοπό την άμεση κατανόηση των υποκείμενων δεδομένων (Few, 2013: 36). Αυτές οι συνθέσεις κατασκευάζονται προκειμένου να είναι δυνατή η παρακολούθηση βασικών μεγεθών και δεικτών απόδοσης, με απώτερο στόχο τη λήψη ενημερωμένων αποφάσεων.

Ο αγγλικός όρος "*dashboard*", που μεταφράζεται αυτολεξεί ως *πίνακας οργάνων μηχανήματος*, ανταποκρίνεται σε ένα βασικό χαρακτηριστικό αυτών των συνθέσεων οπτικοποίησης κυρίως κατά τα πρώτα χρόνια εμφάνισής τους. Συγκεκριμένα, εκτός από τις συνήθεις μεθόδους γραφικής αναπαράστασης, τα *dashboards* περιλάμβαναν οπτικές αναπαραστάσεις δεδομένων χρησιμοποιώντας σχετικές μεταφορές, όπως φαίνεται στην ακόλουθη εικόνα (Εικόνα 24):

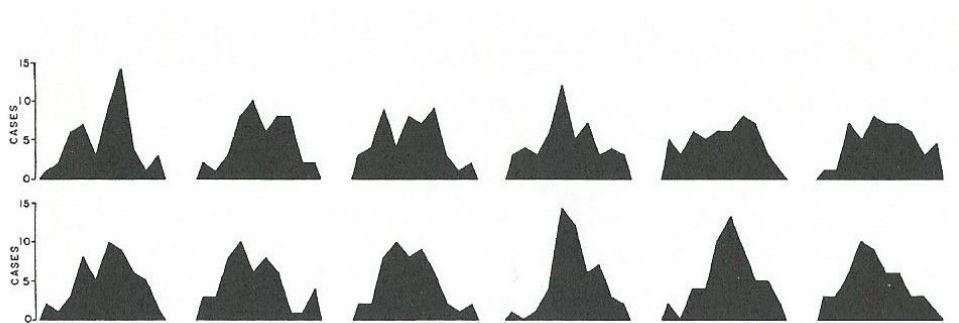


Εικόνα 24: Παράδειγμα Dashboard
(Πηγή: Few, 2013: 7)

Τα οφέλη της σύνθεσης επιμέρους οπτικοποιήσεων, ειδικά για τη διευκόλυνση διεξαγωγής συγκρίσεων, έχουν τονιστεί επανειλημμένα από τον

Tufte (ενδεικτικά 1990 και 1997), ο οποίος ονομάζει αυτή τη δομή "πολλαπλά μικρογραφήματα" (small multiples).

Βασίζόμενος στο ότι η εξέταση των ποσοτικών δεδομένων εστιάζει κατ' αρχήν στη σύγκριση, υποστηρίζει ότι τα *πολλαπλά μικρογραφήματα* αποτελούν τη βέλτιστη σχεδιαστική λύση όταν αναπαρίστανται δεδομένα (Tufte, 1990: 67). Επισημαίνει δε ότι η επανάληψη εικόνων *"αποκαλύπτει την επανάληψη και την αλλαγή, τα μοτίβα και τις εκπλήξεις"*, με αποτέλεσμα η σύνθεση αυτή να επιτρέπει την εξαγωγή ουσιωδών συμπερασμάτων από διαφορετικά μη προσβάσιμα δεδομένα (Tufte, 1997: 105).



Εικόνα 25: Παράδειγμα πολλαπλών μικρογραφημάτων
(Πηγή: Tufte, 2001: 169)

Αν και ο Kirk (2012: 130) αναφέρει τα πολλαπλά μικρογραφήματα ως μέθοδο οπτικοποίησης που διευκολύνει την διεξαγωγή συγκρίσεων μεταξύ κατηγοριών, αναγνωρίζει ότι μάλλον αποτελούν εργαλείο σύνθεσης που επιτρέπει τη διεξαγωγή συγκρίσεων. Δομικά στοιχεία της αποτελούν επιμέρους οπτικές αναπαραστάσεις ως αποτέλεσμα χρήσης διαφόρων μεθόδων, ανάλογα με την περίπτωση (ιστογράμματα, διαγράμματα σκέδασης κ.λπ.).

Καταλήγοντας, είναι σαφές ότι η επιλογή του τύπου οπτικής αναπαράστασης εξαρτάται από τον σκοπό της οπτικοποίησης. Εκτός από τους τύπους οπτικοποιήσεων που καταγράφονται σε κατηγοριοποιήσεις όπως αυτή του Kirk, ο σχεδιαστής μπορεί να επιλέξει να προχωρήσει στη δημιουργία μίας σύνθετης οπτικοποίησης, που περιλαμβάνει επιμέρους τμήματα, στα οποία τα δεδομένα αναπαρίστανται με διαφορετικές μεθόδους ή η ίδια μέθοδος

χρησιμοποιείται επαναλαμβανόμενα για να αναδείξει διαφορετικές διαστάσεις.

Ο σχεδιαστής πρέπει να κατέχει τις μεθόδους οπτικοποίησης και τις συμβάσεις για να μπορεί να χρησιμοποιήσει τυποποιημένα και ευρέως κατανοητά από τους χρήστες πρότυπα οπτικής αναπαράστασης. Η επίγνωση του τι μπορεί και τι δεν μπορεί να αναπαρασταθεί μέσω αυτών των τυποποιημένων μεθόδων σε συνδυασμό με τη γνώση των οπτικών κωδικοποιήσεων σε δομικό επίπεδο για την κατάλληλη αντιστοίχισή τους με τις ιδιότητες των δεδομένων, του επιτρέπουν να επινοήσει νέες μεθόδους που εξυπηρετούν καλύτερα τον σκοπό της οπτικοποίησης, όπως για παράδειγμα η *οπτικοποίηση μέσω* που αναφέρθηκε στο Κεφάλαιο 3.

4.2.3 Εστιάζοντας στο χρώμα⁵⁵

Ένα ενδιαφέρον εύρημα από την καταγραφή των οπτικών στοιχείων που συνθέτουν κάθε τύπο οπτικοποίησης σύμφωνα με τον Kirk (Πίνακας 6 πιο πάνω), είναι ότι φαίνεται πως η χρωματική απόχρωση χρησιμοποιείται εξίσου με τη θέση.

Στη συνέχεια, εξετάζεται η χρήση του χρώματος στις οπτικές αναπαραστάσεις δεδομένων. Χωρίς περαιτέρω επέκταση σε θεωρητικά ζητήματα που αφορούν το χρώμα εν γένει, κρίνεται σκόπιμη η αναφορά στα τρία συγκεκριμένα και μετρήσιμα χαρακτηριστικά του χρώματος: την απόχρωση, τον κορεσμό και τη φωτεινότητα (Dondis, 1973: 39). Η απόχρωση (hue) συχνά συγχέεται με το χρώμα ως συνολική εντύπωση του συνδυασμού των επιμέρους διαστάσεών του. Ο κορεσμός (saturation) αναφέρεται στην ένταση της απόχρωσης ενός χρώματος. Ένα κορεσμένο χρώμα προκύπτει από ένα ή δύο βασικά χρώματα, αναμεμειγμένα σε οποιαδήποτε αναλογία και δεν περιλαμβάνουν μαύρο, άσπρο ή γκρι (Holtzschue, 2011: 73). Τα λιγότερο

⁵⁵ Σε αυτή την ενότητα αξιοποιείται υλικό από την εργασία με θέμα *"Οπτικός Σχεδιασμός: Χρώμα & Διεπιφάνειες Ανθρώπου - Υπολογιστή"* που είχε εκπονήσει η γράφουσα στο πλαίσιο του μαθήματος *"Επικοινωνιακές Διαστάσεις του Νέου Τεχνολογικού Περιβάλλοντος"* (Οκτώβριος 2013).

κορεσμένα χρώματα τείνουν προς την ουδετερότητα (Dondis, 1973: 51). Τέλος, η φωτεινότητα (brightness) συνδέεται με τις τονικές διαβαθμίσεις του χρώματος, από τις πιο φωτεινές έως τις πιο σκούρες.

Όπως διακρίνει ο Tufte (1990: 81), στο πλαίσιο της αντιστοίχισης του χρώματος με την αναπαριστώμενη πληροφορία, το χρώμα αξιοποιείται από τον σχεδιαστή για να κατηγοριοποιήσει ή να ονοματίσει, να εκφράσει ποσότητες, να αναπαραστήσει την πραγματικότητα ή να λειτουργήσει διακοσμητικά. Εντοπίζεται ότι οι δύο πρώτες από αυτές τις χρήσεις του χρώματος συνδέονται με το είδος των δεδομένων που αναπαρίστανται: η μεν πρώτη αναφέρεται στις ποιοτικές μεταβλητές, η δε δεύτερη στις ποσοτικές.

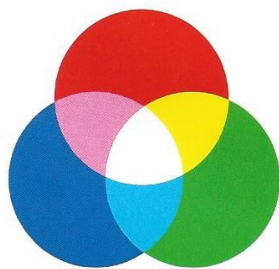
Ανάλογα με το είδος των αναπαριστώμενων μεταβλητών, ο σχεδιαστής πρέπει να επιλέξει την παράμετρο του χρώματος που μπορεί να αναδείξει με ακριβή και αποτελεσματικό τρόπο τις ιδιότητες των δεδομένων που αναπαρίστανται: η φωτεινότητα του χρώματος είναι προτιμότερη για την κωδικοποίηση ποσοτικών μεταβλητών, διότι μπορεί να υποδηλώσει μία έννοια σειράς, αντίθετα από τη χρωματική απόχρωση, που ενδείκνυται για τις κατηγορικές μεταβλητές (Driscoll, 2010: 63).

Ο σχεδιαστής παράλληλα όμως πρέπει να επιδιώκει τον περιορισμό της *χρωματικής ζημιάς* (color damage), τηρώντας δύο βασικές αρχές: τη χρήση έντονων χρωμάτων με φειδώ και την αποφυγή τοποθέτησης ανοιχτών, έντονων χρωμάτων μαζί με λευκό (Tufte, 1990: 81-82). Δεδομένου μάλιστα ότι η συχνότερα χρησιμοποιούμενη σήμερα διεπαφή προβολής μίας οπτικοποίησης είναι η οθόνη, είναι χρήσιμο για τον σχεδιαστή να κατανοεί τον τρόπο που αυτή λειτουργεί, ώστε να γνωρίζει τις δυνατότητες και τους περιορισμούς που επιβάλλει στην προβολή χρωμάτων.

Οι έγχρωμες οθόνες λειτουργούν με βάση το προσθετικό χρωματικό μοντέλο RGB⁵⁶ (Εικόνα 26) και περιλαμβάνουν μεγάλο αριθμό εικονοστοιχείων

⁵⁶ Αρχικά Red, Green, Blue.

(pixels), σε ορθογώνια διάταξη με τη μορφή πλέγματος, τα οποία έχουν τη δυνατότητα προβολής φωτός.



Εικόνα 26: Προσθετικό χρωματικό μοντέλο
(Πηγή: Stone, 2003: 139)

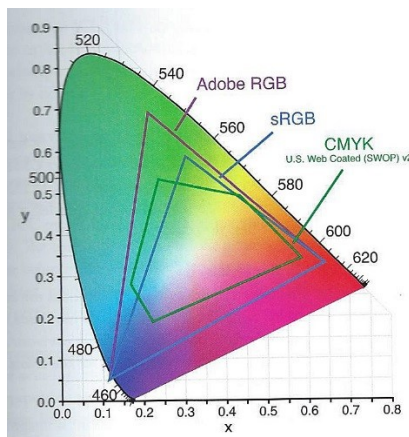
Κάθε εικονοστοιχείο ελέγχεται από την κάρτα γραφικών του υπολογιστή και οι δυνατότητες απεικόνισης χρώματος εξαρτώνται από το πλήθος των bits ανά εικονοστοιχείο. Κάθε bit διαχειρίζεται την πληροφορία on/off προκειμένου να ανάψει ή να σβήσει το εικονοστοιχείο (Dix et al., 2004: 79-80).

Στις οθόνες υγρών κρυστάλλων (Liquid Crystal Display - LCD) που είναι ευρέως διαδεδομένες σήμερα, κάθε εικονοστοιχείο αποτελείται από τρία υπο-εικονοστοιχεία (sub-pixels), καθένα εκ των οποίων αντιστοιχεί σε ένα από τα χρώματα RGB (Kadavy, 2011: 218-219). Το κόκκινο, το πράσινο και το μπλε μπορούν να λάβουν τιμές από το 0 έως και το 255, επομένως ο συνδυασμός των επιμέρους τιμών ανά χρώμα δημιουργεί τις διαφορετικές αποχρώσεις. Με άλλα λόγια, κάθε απόχρωση ορίζεται από τρεις αριθμούς, καθένας εκ των οποίων αντιστοιχεί σε ένα από τα χρώματα RGB. Σημειώνεται ότι ο συνδυασμός των τριών χρωμάτων σε πλήρη ένταση (όλα με τιμή 255) παράγει το λευκό, ενώ όταν και τα τρία έχουν τιμή μηδέν παράγεται μαύρο.

Σε σύγκριση με την εκτύπωση χρωμάτων σε χαρτί που βασίζεται στο αφαιρετικό χρωματικό μοντέλο CMYK⁵⁷, η προβολή μίας οπτικοποίησης δεδομένων στην οθόνη έχει το πλεονέκτημα προβολής μεγαλύτερης γκάμας χρωμάτων, διότι βασίζεται στο προσθετικό χρωματικό μοντέλο RGB. Τα δύο

⁵⁷ Αρχικά Cyan, Magenta, Yellow, Key.

αυτά χρωματικά μοντέλα έχουν αρκετές κοινές δυνατότητες παραγωγής χρωμάτων, ωστόσο παρουσιάζουν διαφορετικούς, εγγενείς περιορισμούς στην αναπαραγωγή αποχρώσεων, όπως φαίνεται στην ακόλουθη εικόνα (Εικόνα 27):



Εικόνα 27: Σύγκριση χρωματικής γκάμας RGB και CMYK
(Πηγή: Kadavy, 2011: 229)

Εφόσον ο σχεδιασμός οπτικοποιήσεων διέπεται από βασικές αρχές του γραφικού σχεδιασμού, σε ό,τι αφορά τις αισθητικές επιλογές και τον συνδυασμό της απόχρωσης, της φωτεινότητας και του κορεσμού σε επίπεδο σύνθεσης, μπορούν να εφαρμοστούν οι αρχές βαθιά ριζωμένες στις εικαστικές τέχνες (αρμονία, αντίθεση). Υπάρχουν μάλιστα εργαλεία στη διάθεση των σχεδιαστών, είτε ως μέρος πακέτων λογισμικών, είτε ως αυτόνομες υπηρεσίες που διευκολύνουν τις δοκιμές διάφορων συνδυασμών χρωμάτων και τελικά βοηθούν στη διαμόρφωση της χρωματικής παλέτας.

Επιπρόσθετα, το χρώμα θα πρέπει να χρησιμοποιείται στο πλαίσιο συμβάσεων, εμπεδωμένων από τους χρήστες και σε συμφωνία με τις προσδοκίες τους (Dix et al., 2004: 136), αφού ως φορέας μηνυμάτων, μπορεί να προκαλέσει συναισθηματικές αντιδράσεις ή/και να λειτουργήσει συμβολικά (Dondis, 1973: 50 και Holtzschue, 2011: 57-60). Υπάρχουν μάλιστα πειραματικά στοιχεία ότι η επιλογή χρωμάτων με σημειολογικά κριτήρια συνδέεται με ταχύτερη κατανόηση του περιεχομένου μίας γραφικής αναπαράστασης δεδομένων (Lin et al., 2013: 8).

Επιπλέον, πρέπει να συνεκτιμώνται οι ιδιαιτερότητες των χρηστών. Περίπου 10% του παγκόσμιου πληθυσμού πάσχει από κάποια μορφή

αχρωματοψίας (Sherin, 2012: 44), επομένως δυσκολεύεται ή, σε σπάνιες περιπτώσεις, αδυνατεί να διακρίνει διαφορές μεταξύ αποχρώσεων. Επομένως, ο σχεδιαστής θα πρέπει να φροντίζει να συνδυάζει το χρώμα με άλλα οπτικά στοιχεία, προκειμένου οι κρίσιμες πληροφορίες να είναι αντιληπτές και από χρήστες που αντιμετωπίζουν δυσκολίες στην αναγνώριση χρωμάτων. Εκτός από τη γενική οδηγία να χρησιμοποιούνται χρώματα με μεγάλη αντίθεση (Morioka and Stone, 2006: 64-69), ειδικά στην περίπτωση της οπτικοποίησης δεδομένων συστήνεται επιπλέον η χρήση μαύρου χρώματος γραμματοσειράς σε ανοιχτόχρωμο φόντο και η τοποθέτηση των ετικετών κοντά στα στοιχεία του γραφήματος (Wong, 2010: 45).

4.3 Εμπλέκοντας τον χρήστη

Η οπτικοποίηση φέρει το μήνυμα του σχεδιαστή, και κατ' επέκταση εμπεριέχει επικοινωνιακή διάσταση, όπως άλλωστε έχει επισημανθεί δεκαετίες νωρίτερα, με τον Bertin (2011) να συγκαταλέγει την επικοινωνία της πληροφορίας στις βασικές λειτουργίες της γραφικής αναπαράστασης. Σημειώνει χαρακτηριστικά ότι στο πλαίσió της *"το γραφικό είναι ένα 'μήνυμα'"* (Bertin, 2011: 162). Η άποψη αυτή συνεχίζεται να διατυπώνεται και στη σύγχρονη βιβλιογραφία (Heer et al., 2010· Kirk, 2012· Yau, 2013).

Αξίζει να δοθεί έμφαση στον ορισμό του Kirk (2012: 17), ο οποίος θεωρεί την οπτικοποίηση δεδομένων ως την *"αναπαράσταση και παρουσίαση των δεδομένων που αξιοποιεί τις ικανότητες της οπτικής μας αντίληψης με σκοπό τη διεύρυνση της νόησης"*. Ενώ η διατύπωσή του προβλέπει τη διεύρυνση της νόησης αξιοποιώντας τις δυνατότητες του ανθρώπινου οπτικού συστήματος και συμφωνεί με τους ορισμούς που αναφέρθηκαν στο προηγούμενο κεφάλαιο, εντοπίζεται ένας ενδιαφέρων διαχωρισμός: χρησιμοποιεί τις λέξεις *"αναπαράσταση"* και *"παρουσίαση"*, για να περιγράψει δύο διαφορετικές λειτουργίες της οπτικοποίησης. Αιτιολογεί αυτή τη διάκριση, εξηγώντας ότι η αναπαράσταση των δεδομένων αφορά στην επιλογή της μορφής απεικόνισης,

ενώ η παρουσίασή τους συνίσταται στην ενσωμάτωση της αναπαράστασης σε ένα ευρύτερο επικοινωνιακό πλαίσιο (Kirk, 2012: 17).

Με άλλα λόγια, η αναπαράσταση ενός συνόλου δεδομένων περιορίζεται στην επιλογή της αφαιρετικής απεικόνισης (π.χ. ένα διάγραμμα στηλών) ενώ η παρουσίασή του συνίσταται στον συνδυασμό των σχεδιαστικών επιλογών που διαμορφώνουν το συνολικό τελικό αποτέλεσμα που βλέπει ο χρήστης. Αναγνωρίζει λοιπόν (Kirk, 2012: 16) ότι η διαδικασία της οπτικοποίησης δεδομένων πληροί τα κριτήρια να θεωρηθεί ως επικοινωνιακή διαδικασία, καθώς περιλαμβάνει έναν πομπό (δημιουργός-σχεδιαστής), ένα μήνυμα (οπτικοποίηση) και έναν δέκτη (τελικός χρήστης).

Σύμφωνα με τον Ware (2013: 2), ενώ παλαιότερα⁵⁸ ο όρος *"οπτικοποίηση"* αφορούσε στην κατασκευή μιας οπτικής εικόνας που υπήρχε μόνο στον νου, σήμερα συνίσταται στη γραφική αναπαράσταση δεδομένων ή εννοιών και έχει καταστεί *"ένα εξωτερικό τεχνούργημα (artifact) που υποστηρίζει τη λήψη αποφάσεων"*. Σε αυτό το πλαίσιο, θεωρεί ότι ο χρήστης που χρησιμοποιεί ένα εργαλείο οπτικής ανάλυσης συνιστά ένα γνωσιακό σύστημα (cognitive system), το οποίο περιλαμβάνει αφενός το ανθρώπινο οπτικό σύστημα, αφετέρου τις πληροφορίες και το υπολογιστικό σύστημα που τις φέρει· η αλληλεπιδραστική οπτικοποίηση αποτελεί τη διεπαφή μεταξύ των δύο πλευρών, η βελτίωση της οποίας *"μπορεί να βελτιώσει σημαντικά την απόδοση ολόκληρου του συστήματος"* (Ware, 2013: 2).

Βάσει αυτού, εντοπίζει ότι βασικό ζητούμενο αποτελεί *"ο μετασχηματισμός των δεδομένων σε κάτι που οι άνθρωποι μπορούν να κατανοήσουν προκειμένου να βελτιώσουν τη λήψη αποφάσεων"* και προσδιορίζει τέσσερα βασικά στάδια για τη δημιουργία μιας οπτικοποίησης: (1) τη συλλογή και αποθήκευση των πρωτογενών δεδομένων, (2) τον μετασχηματισμό τους για τη διευκόλυνση του χειρισμού τους μέσω υπολογιστή, (3) την αντιστοίχιση των δεδομένων με την κατάλληλη οπτική αναπαράσταση, και (4) την αξιοποίηση των γνωστικών και

⁵⁸ Επικαλείται ορισμό που είχε διατυπωθεί το 1972.

αντιληπτικών ικανοτήτων του χρήστη (Ware, 2013: 4-5).

Σύμφωνα με το απλοποιημένο μοντέλο επεξεργασίας πληροφοριών από τον χρήστη μέσω της οπτικής αντίληψης που περιγράφει, αρχικά γίνεται μια γρήγορη σάρωση των βασικών στοιχείων, στη συνέχεια εντοπίζονται μοτίβα που σχετίζονται με την αναγνώριση, ενώ στο τέλος ενεργοποιείται η οπτική μνήμη εργασίας (Ware, 2013: 21-22). Θα μπορούσε κατ' επέκταση να θεωρηθεί ότι αυτό το μοντέλο δεν περιορίζεται στα εργαλεία οπτικής ανάλυσης και ισχύει γενικά για τις οπτικοποιήσεις που παρέχουν δυνατότητες εξερεύνησης των υποκείμενων δεδομένων για την ανακάλυψη γνώσης από τον χρήστη, καθώς τον τοποθετούν στη θέση του αναλυτή και του παρέχουν δυνατότητα διαλόγου με αυτά.

Διευρύνοντας την προβληματική, θα μπορούσε να θεωρηθεί ότι οι σύγχρονες οπτικοποιήσεις που σχεδιάζονται και υλοποιούνται κυρίως για χρήση μέσω ηλεκτρονικού υπολογιστή συνιστούν διεπαφές μεταξύ των υποκείμενων δεδομένων και του χρήστη που τις χρησιμοποιεί για να εξερευνήσει και για να αποκτήσει πρόσβαση στα δεδομένα. Η θεώρηση της οπτικοποίησης ως διεπαφή δεν είναι καινούργια (π.χ. Tidwell, 2011) και οδηγεί στον συλλογισμό ότι, στο πλαίσιο κατασκευής οπτικοποίησης δεδομένων που απευθύνεται σε κοινό πέραν του εαυτού του, ο σχεδιαστής πρέπει να υιοθετεί μία χρηστοκεντρική προσέγγιση για να βελτιώσει την επικοινωνία ανθρώπου-υπολογιστή.

Αναφερόμενος στις βέλτιστες πρακτικές κατά τον σχεδιασμό πληροφοριακών συστημάτων, ο Norman (1988: 189-190), αρκετές δεκαετίες νωρίτερα, τονίζει ότι το νοητικό μοντέλο που αναπτύσσει ο σχεδιαστής θα πρέπει να ταυτίζεται με αυτό του χρήστη, δεδομένου ότι ο σχεδιαστής και ο χρήστης ουσιαστικά επικοινωνούν μέσω του συστήματος. Προφανώς σε αυτή τη βάση, ο Ware (2013: 23) σημειώνει ότι η οπτικοποίηση μπορεί να ιδωθεί από δύο οπτικές, τόσο από αυτή του δημιουργού, όσο και από αυτή του χρήστη.

Συνοψίζοντας και συνδέοντας με όσα έχουν ήδη αναφερθεί σχετικά με την επιλογή μεθόδου οπτικής αναπαράστασης δεδομένων σύμφωνα με τον σκοπό

της οπτικοποίησης, ο σχεδιαστής θα πρέπει να συνεκτιμά την επικοινωνιακή της διάσταση, όταν η οπτικοποίηση δεν αποτελεί μία κατασκευή που προορίζεται για ιδία χρήση. Εκτός από τον οπτικό σχεδιασμό που αναδεικνύει τις ιδιότητες των δεδομένων, ο σχεδιαστής μπορεί να εμπλουτίσει αυτή την κατασκευή, χρησιμοποιώντας στοιχεία αλληλεπίδρασης ή/και επιλέγοντας αφηγηματικές τεχνικές, προκειμένου (1) να ενισχύσει το πλαίσιο ερμηνείας και κατανόησης των αναπαριστώμενων δεδομένων και (2) να ενθαρρύνει τον διάλογο του χρήστη με τα υποκείμενα δεδομένα. Με τον τρόπο αυτό, παρέχεται στον χρήστη ένα περιβάλλον ευνοϊκό για εξερεύνηση των δεδομένων και κατ' επέκταση για ανακάλυψη γνώσης.

4.3.1 Αλληλεπίδραση

Ένα βασικό χαρακτηριστικό των σημερινών οπτικών αναπαραστάσεων είναι η παροχή δυνατότητας αλληλεπίδρασης του χρήστη με τα δεδομένα. Στο κλασικό κείμενο *"The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations"*, ο Schneiderman (1996)⁵⁹ υποστηρίζει ότι για τη διευκόλυνση του χρήστη κατά την αναζήτηση πληροφοριών, η οπτικοποίηση δεδομένων θα πρέπει να παρέχει σε αυτόν κάποιες βασικές δυνατότητες εξερεύνησης: (1) να βλέπει μία πρώτη γενική εικόνα, (2) να εστιάζει και να περιορίζει την προβολή ώστε να επικεντρώνεται σε εκείνα τα δεδομένα που τον ενδιαφέρουν, περιορίζοντας τα μη απαραίτητα, και (3) να επιλέγει ένα σημείο για να αποκτήσει πρόσβαση σε περισσότερες λεπτομέρειες εφόσον κρίνει ότι του είναι απαραίτητες.

Αυτές οι δυνατότητες συνοψίζονται στο περίφημο μάντρα *"overview first, zoom and filter, then details-on-demand"* (Schneiderman, 1996: 336-337). Μαζί με τις επιπλέον δυνατότητες (1) προβολής σχέσεων μεταξύ αντικειμένων, (2) διατήρησης της ιστορικότητας των εξερευνητικών δραστηριοτήτων για την

⁵⁹ μάλλον εμπνεόμενος από την -τότε σχετικά πρόσφατη- επικράτηση των διεπιφανειών άμεσου χειρισμού που ανέδειξε την υπεροχή της χρήσης οπτικών αντί κειμενικών στοιχείων για την επικοινωνία ανθρώπου-υπολογιστή.

ανάκτηση ενεργειών, και (3) εξαγωγής μέρους του συνόλου δεδομένων που ικανοποιεί συγκεκριμένα κριτήρια εξερεύνησης, συνιστούν τις εργασίες που κατά την άποψή του πρέπει να υποστηρίζονται από τα πακέτα λογισμικού σύνταξης επερωτήσεων (query formulation) και οπτικοποίησης δεδομένων (Schneiderman, 1996: 340-341).

Το μάντρα του Schneiderman αναφέρεται συχνά στη βιβλιογραφία σχετικά με τις τρέχουσες πρακτικές οπτικοποίησης δεδομένων (π.χ. Lima, 2011: 92· Cairo, 2013: 195· Börner and Polley, 2014: 217). Πλέον είναι τεχνολογικώς εφικτή και ευρέως διαδεδομένη η παροχή λειτουργικότητας αλληλεπίδρασης με τα δεδομένα, τόσο στα εργαλεία οπτικής ανάλυσης (π.χ. Tableau Desktop, SAP Lumira), όσο και στις κατασκευές οπτικοποιήσεων με σκοπό την ανάλυση δεδομένων.

Σε τι όμως συνίσταται η παροχή αυτής της δυνατότητας; Η Tidwell (2011: 287-288), ασχολούμενη με τον σχεδιασμό οπτικών αναπαραστάσεων δεδομένων ως διεπιφανειών, εντοπίζει κάποιες συχνά χρησιμοποιούμενες τεχνικές πλοήγησης που εξυπηρετούν την εξερεύνηση των δεδομένων, οι οποίες συνίστανται:

- στην οριζόντια ή κάθετη κύλιση προκειμένου να καταστεί ορατό μέρος της οπτικοποίησης που δεν είναι άμεσα προσβάσιμο
- στη μεγέθυνση για την εστίαση σε κάποιο τμήμα της οπτικοποίησης
- στην ανάπτυξη ή στην σύμπτυξη των αντικειμένων που ανήκουν σε ανώτερο ιεραρχικά επίπεδο, η οποία επιτρέπει τον εντοπισμό ιεραρχικών σχέσεων
- στην αναλυτική κάθοδο (drill down) στα υποκείμενα δεδομένα.

Επιπρόσθετα, σημειώνει ότι μέσω της παροχής δυνατότητας ταξινόμησης και αναδιάταξης ο χρήστης μπορεί να διενεργήσει πολλαπλές συγκρίσεις και να ανακαλύψει αναπάντεχες σχέσεις. Αυτή η δυνατότητα συνίσταται στην κατάταξη με αλφαβητική ή αριθμητική σειρά, βάσει χρονικών ή γεωγραφικών κριτηρίων, σύμφωνα με συγκεκριμένες κατηγορίες ή λέξεις-κλειδιά, ή ακόμη και

με τυχαίο τρόπο (Tidwell, 2011: 289-291). Τέλος, αποδίδει σημασία στην παροχή δυνατότητας αναζητήσεων και φιλτραρίσματος των δεδομένων, καθώς και στην παροχή επεξηγήσεων, κατόπιν επιθυμίας του χρήστη ώστε αυτός να αποφασίζει, με την κίνηση της συσκευής εισόδου ή με την επιλογή ενός αντικειμένου, να καταστήσει ορατά τα δεδομένα ή τις λεπτομέρειες που τον ενδιαφέρουν (Tidwell, 2011: 291-294).

Μία κατηγοριοποίηση των τεχνικών αλληλεπίδρασης που μπορούν να χρησιμοποιηθούν για τη διευκόλυνση της οπτικής ανάλυσης διατυπώνεται από τους Heer και Schneiderman (2012), οι οποίοι οργανώνουν τις επιμέρους τεχνικές σε τρεις ευρείες κατηγορίες: (1) στην παροχή δυνατότητας προσδιορισμού των δεδομένων και προβολών με ενδιαφέρον για τον χρήστη, (2) στην παροχή ευχέρειας διαμόρφωσης της προβολής για σκοπούς διερεύνησης των δεδομένων, μέσω της επιλογής αντικειμένων, της περιήγησης από το υψηλότερο στο χαμηλότερο επίπεδο πληροφόρησης και του συνδυασμού εναλλακτικών προβολών για πληρέστερη διερεύνηση, και (3) στην καταγραφή της αναλυτικής διαδικασίας, προκειμένου να είναι δυνατή η εκ των υστέρων πρόσβαση σε καταγραφές των επιμέρους βημάτων, είτε από τον χρήστη είτε από άλλα ενδιαφερόμενα μέρη (π.χ. συνεργάτες).

Καταλήγοντας, εντοπίζονται αρκετά κοινά σημεία στις προσεγγίσεις της Tidwell και των Heer και Schneiderman. Κρίνεται σκόπιμη η εστίαση στη σταδιακή πρόσβαση, από τη γενική εικόνα στα δεδομένα που αυτή αναπαριστά, σύμφωνα με τις ανάγκες του χρήστη: φαίνεται να έχει πεδίο εφαρμογής στις περιπτώσεις οπτικοποιήσεων μεγάλων συνόλων δεδομένων, όπως για παράδειγμα τα διαγράμματα δικτύου που αποσκοπούν στην απεικόνιση σχέσεων (Lima, 2011: 195).

Βάσει των προαναφερθέντων, θα μπορούσε να υποστηριχθεί ότι τα γενικότερα οφέλη από την αξιοποίηση των δυνατοτήτων αλληλεπίδρασης είναι αμφίπλευρα. Ο σχεδιαστής έχει τη δυνατότητα να διαμορφώσει την οπτικοποίηση σύμφωνα με όσα κρίνει απαραίτητα βάσει του σκοπού της, χωρίς

να επιβαρύνει τον χρήστη προβάλλοντας μονομιάς όλη τη δυνατή πληροφόρηση. Παράλληλα, ο χρήστης έχει ενεργή συμμετοχή στη διαμόρφωση της προβαλλόμενης οπτικοποίησης, μέσω της επιλογής προβολής εναλλακτικών οπτικών, σύμφωνα με τις επιθυμίες του.

4.3.2 Αφηγηματικές τεχνικές

Η παροχή δυνατότητας αλληλεπίδρασης έχει εμπλουτίσει τις οπτικές αναπαραστάσεις δεδομένων, μετατρέποντάς τις από στατικές εικόνες σε δυναμικές προβολές διαφορετικών πλευρών των αναπαριστώμενων δεδομένων. Η μεταμόρφωση των ηλεκτρονικών υπολογιστών σε *"ένα νέο δυναμικό οπτικό μέσο"* που μπορεί να προβάλλει τεράστιους όγκους δεδομένων με αποτελεσματικό τρόπο σε μεγάλο κοινό, χάρη στις εξελίξεις σε επίπεδο λογισμικού αλλά και υλισμικού, ενέπνευσε προσδοκίες για την ανάδειξη νέων ειδών οπτικοποίησης που θα παρουσιάζουν πληροφορίες με τη χρήση αφηγηματικών μέσων (Gershon and Page, 2001: 36-37).

Τα τελευταία χρόνια μάλιστα παρατηρείται έντονο ενδιαφέρον για τον ρόλο που μπορεί να διαδραματίσει η αφήγηση στην οπτικοποίηση δεδομένων. Ίσως η αυξανόμενη χρήση πληροφοριογραφημάτων (infographics) για την πλαισίωση δημοσιογραφικών άρθρων και η άνθιση της δημοσιογραφίας δεδομένων έχει αναδείξει τις δυνατότητες εμπλουτισμού της οπτικοποίησης με αφηγηματικά μέσα για την παρουσίαση των δεδομένων.

Βάσει των προαναφερθέντων, η χρησιμότητα της οπτικοποίησης δεν περιορίζεται στο πλαίσιο διερεύνησης των δεδομένων, αλλά επεκτείνεται και στην παρουσίασή τους, διότι τα αποτελέσματα της ανάλυσης πρέπει να επικοινωνηθούν σε άλλους ανθρώπους από αυτούς που κατέληξαν σε αυτά. Όμως, όπως επισημαίνουν οι Kosara και Mackinlay (2013: 44-45), τα εργαλεία οπτικοποίησης δεδομένων που χρησιμοποιούνται για αναλυτικούς σκοπούς δεν είναι απαραίτητα κατάλληλα για την παρουσίαση των αποτελεσμάτων σε κοινό, διότι η παρουσίαση των αποτελεσμάτων αποσκοπεί *"στην υποστήριξη μίας θέσης*

ή στην επεξήγηση των ευρημάτων". Επομένως, προκύπτει ότι η χρήση εργαλείων Οπτικής Ανάλυσης, η ουδετερότητα των οποίων θεωρητικά διασφαλίζει την εξαγωγή αμερόληπτων συμπερασμάτων, δεν ενδείκνυται στην περίπτωση κατασκευής οπτικοποίησης που στοχεύει στη διατύπωση μίας επιχειρηματολογίας ή στην πρόκληση συναισθηματικών αντιδράσεων.

Οι Kosara και Mackinlay (2013: 44, 50) υποστηρίζουν ότι η χρήση αφηγηματικών τεχνικών για την παρουσίαση δεδομένων συνιστά το επόμενο βασικό αντικείμενο έρευνας στον πεδίο της οπτικοποίησης, λόγω (1) της ωρίμασης στην έρευνα σχετικά με τη χρήση της οπτικοποίησης για σκοπούς διερεύνησης και ανάλυσης, και (2) της αυξανόμενης χρήσης της οπτικοποίησης στο πλαίσιο της διαδικασίας λήψης αποφάσεων. Υπό αυτό το πρίσμα, διακρίνουν τρία πιθανά σενάρια παρουσίασης δεδομένων με τη χρήση αφηγηματικών τεχνικών (Kosara and Mackinlay, 2013: 45-46):

- Η *αυτοπροβαλλόμενη παρουσίαση σε μεγάλο κοινό* (self-running presentation to a large audience) χαρακτηρίζεται από γραμμικότητα και στοχεύει κυρίως στην ενημέρωση ή στην ευαισθητοποίηση του χρήστη σχετικά με ένα ζήτημα. Δημιουργείται για κατανάλωση από κάθε μεμονωμένο χρήστη χωριστά, χωρίς δυνατότητα άμεσης επαφής με τον δημιουργό. Η διεπαφή της οπτικοποίησης παρέχει περιορισμένες δυνατότητες αλληλεπίδρασης, ωστόσο είναι πιθανό να επιτρέπει την πρόσβαση του χρήστη στα υποκείμενα δεδομένα.
- Κατά τη *ζωντανή παρουσίαση από έναν ομιλητή σε κοινό* (live presentation by a speaker in front of an audience), ο ομιλητής οπτικοποιεί τα δεδομένα σε πραγματικό χρόνο, δέχεται ερωτήσεις ή προσαρμόζει τη ροή και το περιεχόμενο στις ανάγκες του κοινού, αλληλεπιδρώντας με αυτό και προσφέροντας απαντήσεις.
- Η *στοχευμένη παρουσίαση των αποτελεσμάτων* (individual or small-group presentation of results) προϋποθέτει πιο ευέλικτα εργαλεία παρουσιάσεων, μέσω των οποίων ο ομιλητής έχει δυνατότητα

πρόσβασης στα πρωτογενή δεδομένα για την απάντηση τυχόν ερωτημάτων που θέτει το κοινό κατά τη διάρκεια της παρουσίασης.

Αποδίδουν δε ιδιαίτερη σημασία στην παροχή αλληλεπίδρασης, η οποία, όταν παρέχεται στο τέλος μίας αφηγηματικής παρουσίασης δεδομένων, συνιστά *"μία κατάλληλη αφετηρία για διερεύνηση"* (Kosara and Mackinlay, 2013: 49). Επισημαίνεται ότι η άποψή τους αυτή είναι συμβατή με την υβριδική χρήση της οπτικοποίησης που συνίσταται στην οπτική παρουσίαση του αποτελέσματος ανάλυσης με δυνατότητες αλληλεπίδρασης, όπως έχει προταθεί στις εισαγωγικές επισημάνσεις του παρόντος κεφαλαίου.

Πρακτική εφαρμογή του ερευνητικού του ενδιαφέροντος του Kosara σχετικά με τη χρήση αφηγηματικών τεχνικών στην παρουσίαση δεδομένων συνιστά η λειτουργικότητα *"Story Points"* που περιλαμβάνεται στη νέα έκδοση των λογισμικών της *Tableau* (Tableau, 2014a)⁶⁰. Σύμφωνα με το σχετικό δελτίο τύπου, αυτή η νέα λειτουργικότητα (1) παρέχει τη δυνατότητα *δημιουργίας "οπτικά συναρπαστικών, αλληλεπιδραστικών αφηγήσεων"* και (2) διευκολύνει *"την ενίσχυση του πλαισίου, την ομαλή ροή [της παρουσίασης] και την παροχή καθοδήγησης για την αφήγηση πιο δομημένων ιστοριών βασιζόμενων σε δεδομένα"* για την παραγωγή γνώσης (Tableau, 2014a).

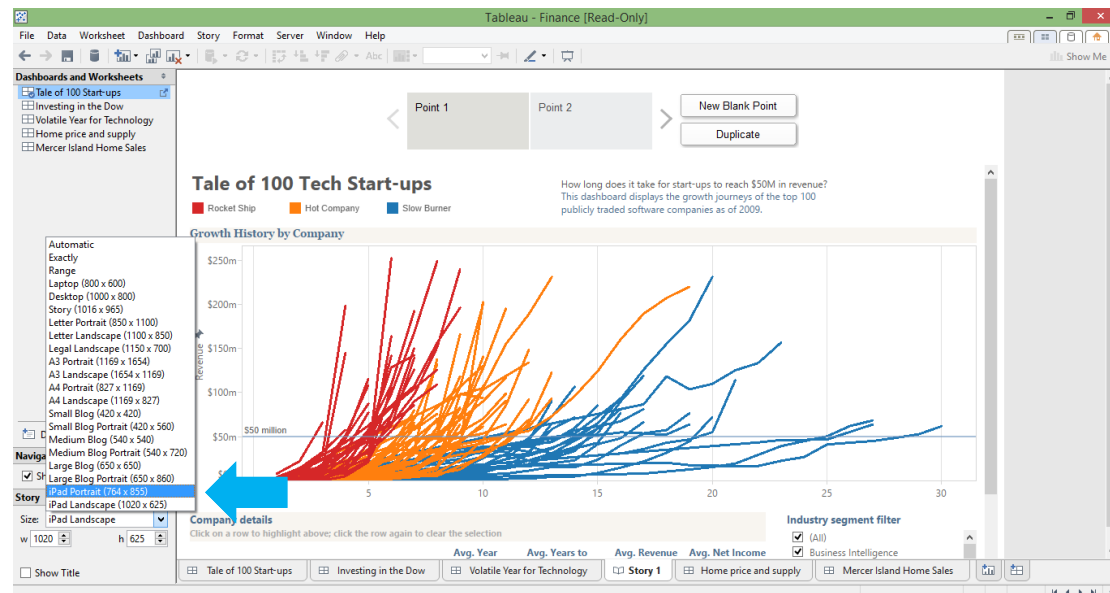
Οι ιστορίες που δημιουργούνται μέσω της εν λόγω λειτουργικότητας αποτελούνται από *"συγκεκριμένες προβολές ή dashboards, που παρουσιάζονται διαδοχικά"*⁶¹, οι οποίες όμως μπορούν και να εναλλάσσονται σύμφωνα με τις επιλογές του χρήστη (Tableau, 2014b: 2, 4)⁶². Ενδιαφέρον παρουσιάζει μία λεπτομέρεια τεχνικής φύσεως: στη διεπαφή της λειτουργικότητας, παρέχεται η δυνατότητα προσδιορισμού των διαστάσεων της συσκευής στην οποία θα

⁶⁰ Ο Kosara, ως επιστημονικός συνεργάτης της Tableau, συνέβαλε καθοριστικά στη σύλληψη και στην ανάπτυξη της λειτουργικότητας *"Story Points"* (Fields, 2014).

⁶¹ Αυτές οι επιμέρους προβολές συνιστούν τα *"σημεία"* (*"points"*), τα οποία συνολικά διαμορφώνουν την *"ιστορία"* (*"story"*). Κάθε σημείο μπορεί να περιλαμβάνει συνδυασμό μίας ή περισσότερων οπτικοποιήσεων, συνοδευτικά κείμενα κ.λπ.

⁶² Ως παράδειγμα αναφέρεται η παρουσίαση που εξηγεί σε κοινό τα βήματα της αναλυτικής διαδικασίας που τελικά οδήγησαν σε κάποιου είδους ανακάλυψη (Tableau, 2014b: 2).

προβληθεί η ιστορία (Εικόνα 28). Χάρη σε αυτή τη δυνατότητα, ο σχεδιαστής συνεκτιμά εξ αρχής τους περιορισμούς στον διαθέσιμο χώρο και επιλέγει τις κατάλληλες σχεδιαστικές λύσεις.



Εικόνα 28: Επιφάνεια εργασίας *Tableau Desktop* - Λειτουργικότητα *Story Points*

Ιδιαίτερη έμφαση δίδεται στα οφέλη της αξιοποίησης της εν λόγω λειτουργικότητας για τη δημιουργία ιστοριών που βασίζονται σε δεδομένα. Ο δημιουργός μπορεί να επιλέξει να αφηγηθεί μία ιστορία που βασίζεται σε δεδομένα, προκειμένου να παρουσιάσει τις πληροφορίες μαζί με το γενικότερο πλαίσιο, ώστε να διευκολύνει τον χρήστη να εντοπίσει τα βασικά σημεία στα οποία βασίζονται τα παρουσιαζόμενα συμπεράσματα. Σε αυτό το πλαίσιο και εμπλουτίζοντας την αφήγηση με δυνατότητες αλληλεπίδρασης,

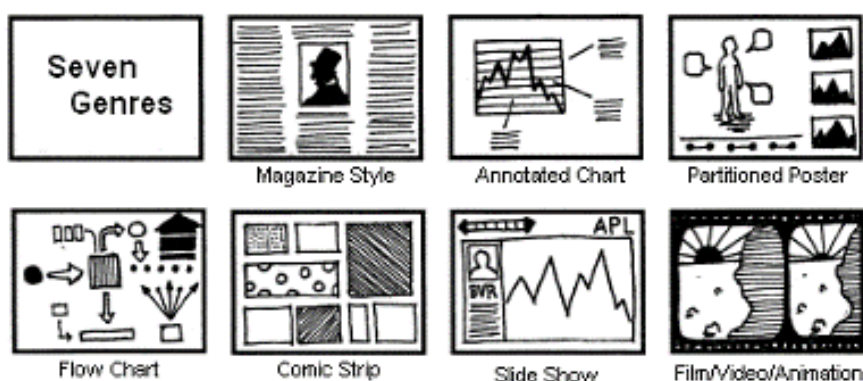
"(...) ο δημιουργός μπορεί να επισημάνει συγκεκριμένα συμπεράσματα και να εμπλουτίσει το γενικό πλαίσιο (...). Επίσης, η αφηγηματική δομή παρέχει στον δημιουργό την ευκαιρία να σπάσει την ιστορία σε επιμέρους τμήματα, καθένα εκ των οποίων προσθέτει κάτι επιπλέον στα προηγούμενα, ώστε ο αποδέκτης να μπορεί να τα αφομοιώσει πιο εύκολα" (Fields, 2014).

Μεταξύ των κατευθύνσεων για μελλοντικές έρευνες σχετικά με την εφαρμογή αφηγηματικών τεχνικών στην παρουσίαση δεδομένων, που

προτείνουν οι Kosara και Mackinlay (2013), συγκαταλέγεται η προτροπή για άντληση ιδεών από τομείς που έχουν μακροχρόνια παράδοση στην κατασκευή ιστοριών, υποδεικνύοντας εμμέσως τη δημοσιογραφία δεδομένων. Λίγα χρόνια νωρίτερα και στο πλαίσιο μελέτης περιπτώσεων που αφορούν κυρίως παραδείγματα δημοσιογραφίας δεδομένων, οι Segel και Heer (2010) εντοπίζουν τρεις κατηγορίες χαρακτηριστικών της αφηγηματικής οπτικοποίησης: (1) το είδος (genre), που διακρίνεται ανάλογα με το πλήθος των πλάνων και την ταξινόμηση των επιμέρους οπτικών στοιχείων, (2) τις τακτικές οπτικής αφήγησης (visual narrative tactics), και (3) τις τακτικές αφηγηματικής δομής (narrative structure tactics).

Τα επτά είδη αφηγηματικών οπτικοποιήσεων που εντοπίζουν είναι τα εξής:

- Διάταξη περιοδικού (Magazine Style)
- Υπομνηματισμένο Γράφημα (Annotated Chart)
- Κατατμημένη Αφίσα (Partitioned Poster)
- Διάγραμμα Ροής (Flow Chart)
- Ταινίες Κόμικ (Comic Strip)
- Προβολή Διαφανειών (Slide Show)
- Ταινία/Βίντεο/Κινούμενες εικόνες (Film/Video/Animation)



Εικόνα 29: Τα είδη αφηγηματικής οπτικοποίησης, σύμφωνα με Segel και Heer
(Πηγή: Segel and Heer, 2010: 1145)

Η πολυπλοκότητα των δεδομένων ή/και της ιστορίας, το κοινό αλλά και το μέσο αποτελούν παράγοντες που πρέπει να λαμβάνονται υπ' όψιν κατά την

επιλογή του αφηγηματικού είδους, αφού δεν ενδείκνυνται όλα για όλες τις περιπτώσεις (Segel and Heer, 2010: 1146). Επισημαίνουν δε ότι είναι πιθανό τα αφηγηματικά είδη να χρησιμοποιούνται μεμονωμένα ή σε συνδυασμό, προκειμένου να δημιουργήσουν πιο σύνθετες δομές (Segel and Heer, 2010: 1145).

Παράλληλα, υποστηρίζουν ότι οι ιστορίες χαρακτηρίζονται από το ποιος καθοδηγεί την εξέλιξή τους και περιγράφουν ένα συνεχές από την αμιγώς καθοριζόμενη από τον δημιουργό (author-driven) έως την αμιγώς καθοριζόμενη από τον χρήστη-αναγνώστη (reader-driven) προσέγγιση, με τα παραδείγματα από τη δημοσιογραφία δεδομένων να βρίσκονται μεταξύ αυτών των άκρων.

Ως αποτέλεσμα, διακρίνουν τρία είδη αφηγηματικών δομών, με διαφορετικό βαθμό εμπλοκής του δημιουργού ή/και του χρήστη:

- Η *δομή ποτηριού μαρτίνι* (Martini Glass structure) βρίσκεται πιο κοντά στην καθοριζόμενη από τον δημιουργό προσέγγιση, καθώς η οπτικοποίηση ξεκινά από σχόλια και εισαγωγικές πληροφορίες που παρέχονται από τον δημιουργό.
- Η *αλληλεπιδραστική προβολή διαφανειών* (Interactive Slideshow) παρέχει δυνατότητες αλληλεπίδρασης στον χρήστη, με τις μεμονωμένες διαφάνειες να εμφανίζουν στοιχεία της προηγούμενης δομής.
- Η *ιστορία αναλυτικής καθόδου* (Drill-Down Story) η εξέλιξη της οποίας εξαρτάται από τον χρήστη, που έχει τον έλεγχο -ωστόσο προϋποθέτει σημαντική προεργασία κατά τον σχεδιασμό από τον δημιουργό.

Πιο πρόσφατα, οι Viégas και Wattenberg (TapestryConference, 2014) διακρίνουν επτά είδη αφήγησης στην οπτικοποίηση δεδομένων, τα οποία παρομοιάζουν με λογοτεχνικά είδη:

- Οι ιστορίες *μυστηρίου* (mystery stories) μοιάζουν με τις οπτικοποιήσεις που παρέχουν στον χρήστη ίχνη, αποκαλύψεις και εξερευνήσεις.
- Η επιστημονική οπτικοποίηση παρομοιάζεται με τις *ιστορίες επιστημονικής φαντασίας* (science fiction), δεδομένου ότι βασίζεται σε

καθιερωμένες συμβάσεις που βοηθούν στην αποκωδικοποίηση του μηνύματος που αφορά σε επιστημονικά δεδομένα.

- Οι *ιστορίες εκλαϊκευμένης επιστήμης* (pop science) παρομοιάζονται με τα πληροφοριογραφήματα που βασίζονται σε δημοφιλή δεδομένα.
- Οι *επιφανειακές αντιγραφές* (skin-deep copies) αφορούν σε απομιμήσεις επιτυχημένων παραδειγμάτων, που επαναλαμβάνονται άκριτα και χάριν της φόρμας.
- Οι *χάρτες* παρομοιάζονται με *ιστορίες αγάπης* (love stories), διότι συνήθως χρησιμοποιούνται για να παρουσιάσουν δεδομένα που αφορούν στη γεωγραφική κατανομή του πλούτου και των κοινωνικών τάξεων.
- Οι *συλλογές* (collections) παρομοιάζονται με τις οπτικοποιήσεις ομοειδών αντικειμένων που αναδεικνύουν αναπάντεχες σχέσεις.
- Τα *διαγράμματα δικτύων* παρομοιάζονται με τις *ιστορίες φαντασίας* (fantasy), διότι ο δημιουργός τους μπορεί να παρασυρθεί από τον ενθουσιασμό του να επιμεληθεί ιδιαιτέρως την παρουσίαση και να χάσει την ουσία της ιστορίας που θέλει να διηγηθεί.

Υποστηρίζουν ότι η αύξηση του πλήθους των παραγόμενων οπτικοποιήσεων και η επαναλαμβανόμενη χρήση συγκεκριμένων τεχνικών έχουν οδηγήσει ήδη στην εμφάνιση συμβάσεων. Το μήνυμά τους προς τους δημιουργούς είναι σαφές: *"ελέγξτε το είδος σας, αλλιώς το είδος σας θα σας ελέγξει"*. Εφιστούν δε την προσοχή στην εμπέδωση των συμβάσεων που χαρακτηρίζουν κάθε είδος, ώστε οι σχεδιαστικές αποφάσεις που λαμβάνονται να είναι συνειδητές και να εξυπηρετούν τον σκοπό της οπτικοποίησης.

Παράλληλα, επισημαίνουν ότι η επίγνωση του δημιουργού για το είδος της οπτικοποίησης που κατασκευάζει του επιτρέπει να παράγει αποτελεσματικές οπτικοποιήσεις και να διευρύνει το εκάστοτε είδος προσθέτοντας νέα στοιχεία. Σε αντίθετη περίπτωση, το αποτέλεσμα είναι κενό περιεχομένου και δεν εκπληρώνει τον απώτερο στόχο της οπτικοποίησης που είναι να καταλήξει στην εξαγωγή γνώσης και νοήματος από τον τελικό χρήστη. Τέλος, σημειώνουν ότι η

επιτυχία μίας οπτικοποίησης δεν κρίνεται από τη μεταδοτικότητα (virality) της αλλά από το μήνυμα που εκφράζει, και καλούν τους δημιουργούς κατά την επιλογή του είδους να έχουν κατά νου το κοινό στο οποίο απευθύνονται: τι γνωρίζει και τι χρόνο μπορεί να διαθέσει.



Εικόνα 30: Αντιπαραβολή αφηγηματικών τεχνικών ιστοριών μυστηρίου με επιλεγμένες οπτικοποιήσεις δεδομένων (Πηγή: TapestryConference, 2014)

Συνοψίζοντας, η προσέγγιση των Segel και Heer (2010) είναι πιο τεχνική καθώς εστιάζει στον τρόπο παρουσίασης των δεδομένων, ενώ οι Viégas και Wattenberg (TapestryConference, 2014) φαίνεται να αποδίδουν μεγαλύτερη βαρύτητα στο περιεχόμενο και στο μήνυμα που εκπέμπει ο δημιουργός της οπτικοποίησης. Ο τομέας που βρίθει παραδειγμάτων που συνδυάζουν τις δύο αυτές προσεγγίσεις είναι αυτός της δημοσιογραφίας δεδομένων, από τον οποίο μπορούν να αντληθούν διδάγματα, καθώς όπως σημειώνουν οι Kosara και Mackinlay (2013: 54), η δημοσιογραφία μπορεί να συνεισφέρει λόγω "της αυξανόμενης εστίασης σε ιστορίες που περιέχουν κείμενο, εικόνες, ήχο και βίντεο, καθώς και όλο και περισσότερους αριθμούς".

Στο παρόν κεφάλαιο, εξετάστηκαν σχεδιαστικές αποφάσεις που λαμβάνονται στο πλαίσιο δημιουργίας σύγχρονων αλληλεπιδραστικών οπτικοποιήσεων από ανθρώπους για ανθρώπους, με σκοπό την ανακάλυψη γνώσης. Η ανάλυση μίας οπτικής αναπαράστασης δεδομένων σε συστατικά στοιχεία, παραπέμπει στα βασικά στοιχεία οπτικής επικοινωνίας, τα οποία είναι οικεία στον χρήστη λόγω της συνεχούς έκθεσής του σε αυτά και αναγνωρίσιμα μέσω των ανθρώπινων αντιληπτικών ικανοτήτων.

Επίσης, η επίγνωση του τι μπορεί και τι δεν μπορεί να αναπαρασταθεί μέσω των τυποποιημένων μεθόδων οπτικοποίησης σε συνδυασμό με τη γνώση των οπτικών κωδικοποιήσεων σε δομικό επίπεδο, επιτρέπουν την επινόηση μεθόδων προσαρμοσμένων στον εκάστοτε σκοπό της οπτικοποίησης. Όμως, εκτός από τον οπτικό σχεδιασμό που αναδεικνύει τις ιδιότητες των δεδομένων, ο σχεδιαστής μπορεί να εμπλουτίσει αυτή την κατασκευή, με στοιχεία αλληλεπίδρασης ή/και αφηγηματικές τεχνικές, για την παροχή στον χρήστη ενός περιβάλλοντος ευνοϊκού για την εξερεύνηση των δεδομένων και κατ' επέκταση για την ανακάλυψη γνώσης από αυτά.

5 Μελέτες περίπτωσης

Συνδέοντας με τα προηγούμενα, τα ΜΔ ορίστηκαν ως ένα πολυδιάστατο φαινόμενο που διαμορφώνεται από τις τεχνολογικές δυνατότητες που επιτρέπουν τη συλλογή, την αποθήκευση, την επεξεργασία και την ανάλυση δεδομένων που χαρακτηρίζονται από πρωτοφανή όγκο, υψηλή ταχύτητα παραγωγής και ανανέωσης ή/και ποικιλομορφία.

Ωστόσο, η αξία τους εξαρτάται από τις δυνατότητες ανάλυσής τους και εξαγωγής συμπερασμάτων από αυτά. Η οπτικοποίηση δεδομένων διαδραματίζει σημαντικό ρόλο στη διαδικασία μετατροπής ανεπεξέργαστων δεδομένων σε χρήσιμη πληροφορία, και μπορεί να αξιοποιηθεί είτε για σκοπούς ανάλυσης, όταν ο χρήστης επιθυμεί να τα μελετήσει, είτε για σκοπούς παρουσίασης προκειμένου να τα επικοινωνήσει σε άλλους.

Οι σημερινές τεχνολογικές δυνατότητες επιτρέπουν τη δημιουργία οπτικών αναπαραστάσεων που επιτρέπουν την αλληλεπίδραση του χρήστη με τα υποκείμενα δεδομένα, σε πρωτογενή ή επεξεργασμένη μορφή. Κατ' επέκταση, διακρίνεται μία υβριδική χρήση της οπτικοποίησης, που προκύπτει από τον συνδυασμό των δύο παραπάνω και συνίσταται στην οπτική παρουσίαση του αποτελέσματος ανάλυσης με δυνατότητες αλληλεπίδρασης. Με αυτόν τον τρόπο, ο χρήστης, εφόσον είναι διαφορετικό άτομο από τον αναλυτή, σχηματίζει μία αρχική άποψη βάσει της ερμηνείας της παρουσίασης και στη συνέχεια εξερευνά και αναλύει περαιτέρω τα υποκείμενα δεδομένα. Επομένως, οι σύγχρονες οπτικοποιήσεις που σχεδιάζονται και υλοποιούνται κυρίως για χρήση μέσω ηλεκτρονικού υπολογιστή μπορούν να θεωρηθούν ως διεπαφές μεταξύ των υποκείμενων δεδομένων και του χρήστη που τις χρησιμοποιεί για να εξερευνήσει και για να αποκτήσει πρόσβαση στα δεδομένα.

Γενικά, η σχεδιαστική διαδικασία για την κατασκευή μίας οπτικής αναπαράστασης δεδομένων περιλαμβάνει μία σειρά αποφάσεων, όπως η

διατύπωση των ερωτημάτων και η επιλογή των συνόλων δεδομένων που μπορούν να τα απαντήσουν, οι οποίες προηγούνται της επιλογής της αντιστοίχισης των δεδομένων με οπτικές κωδικοποιήσεις. Η λήψη των κατάλληλων αποφάσεων προϋποθέτει γνώση, τόσο του γενικότερου περιβάλλοντος από το οποίο έχουν συλλεχθεί τα δεδομένα προς οπτικοποίηση, όσο και των ειδικότερων λεπτομερειών για τις ιδιότητες τους.

Επιπρόσθετα, ο σχεδιαστής πρέπει να συνεκτιμά την επικοινωνιακή διάσταση της οπτικοποίησης, εφόσον πρόκειται για δημιούργημα που απευθύνεται σε κοινό πέραν του κατασκευαστή του. Εκτός από τον οπτικό σχεδιασμό που αναδεικνύει τις ιδιότητες των δεδομένων, ο σχεδιαστής μπορεί να εμπλουτίσει την κατασκευή, χρησιμοποιώντας στοιχεία αλληλεπίδρασης ή/και επιλέγοντας αφηγηματικές τεχνικές, προκειμένου (1) να ενισχύσει το πλαίσιο ερμηνείας και κατανόησης των αναπαριστώμενων δεδομένων και (2) να ενθαρρύνει τον διάλογο του χρήστη με τα υποκείμενα δεδομένα. Με τον τρόπο αυτό, παρέχει στον χρήστη ένα περιβάλλον ευνοϊκό για εξερεύνηση των δεδομένων και κατ' επέκταση για ανακάλυψη γνώσης.

Στο παρόν κεφάλαιο εξετάζονται επιλεγμένες οπτικοποιήσεις δεδομένων σε πραγματικό ή σε ύστερο χρόνο, ως προς το τελικό αποτέλεσμα που προετοιμάζει ένας δημιουργός-άνθρωπος για να προσφέρει τη δυνατότητα ανακάλυψης γνώσης σε έναν χρήστη-άνθρωπο. Στόχος της μελέτης είναι η εποπτική θεώρηση της διαδικασίας σχεδίασης μίας οπτικοποίησης δεδομένων που έχουν συλλεχθεί από πηγές προέλευσης ΜΔ, για τον εντοπισμό των βασικών σχεδιαστικών επιλογών και των επιπλοκών τους.

5.1 Επιλογή και εξέταση περιπτώσεων

Αντικείμενο μελέτης αποτελούν οι σχεδιαστικές επιλογές για την κατασκευή μίας οπτικοποίησης ΜΔ που διαμορφώνουν το τελικό αποτέλεσμα που παρέχεται στον χρήστη. Οι τέσσερις περιπτώσεις οπτικοποίησης που έχουν επιλεγεί βασίζονται σε διαφορετικές πηγές προέλευσης ΜΔ, προκειμένου μέσα

από την ανάλυση και τη σύνθεση των χαρακτηριστικών τους να προκύψουν ερεθίσματα για περαιτέρω διερεύνηση.

Ως προς το περιεχόμενο, επιλέγονται οπτικοποιήσεις που βασίζονται σε δεδομένα που προέρχονται από (1) το κοινωνικό μέσο *Twitter* (κειμενικές αναρτήσεις) και (2) εφαρμογές *Ποσοτικοποιημένου Εαυτού* (αυτο-καταγραφή). Επιπρόσθετα, για τον εντοπισμό των προκλήσεων που θέτει η ταχύτητα παραγωγής και ανανέωσης ΜΔ, επιλέγονται οπτικοποιήσεις ΜΔ που κατασκευάζονται (1) κατόπιν (σε ύστερο χρόνο) ή (2) κατά τη διάρκεια (σε πραγματικό χρόνο) συγκέντρωσης των δεδομένων.

Στον ακόλουθο πίνακα (Πίνακας 7), τοποθετούνται ανάλογα οι επιλεγμένες περιπτώσεις:

Πίνακας 7: Επιλεγμένες μελέτες περίπτωσης

είδος περιεχομένου	οπτικοποίηση σε ύστερο χρόνο	οπτικοποίηση σε πραγματικό ⁶³ χρόνο
κειμενικές αναρτήσεις	• How riot rumours spread on Twitter (Guardian)	• Mentions on Twitter of #Ferguson μέσω Twitter Reverb
αυτο-καταγραφή	• Feltron Annual Report 2013	• Fitbit Dashboard

Πληροφορίες αντλούνται κατ' αρχήν μέσω παρατήρησης και ελεύθερης πλοήγησης στις αντίστοιχες διεπαφές οπτικοποίησης. Παράλληλα, αναζητείται υλικό τεκμηρίωσης σχετικά με τις αποφάσεις που ελήφθησαν στο πλαίσιο της συνολικής σχεδιαστικής διαδικασίας. Όπου εντοπίζεται, αξιοποιείται υλικό που

⁶³ Η σε σχεδόν πραγματικό χρόνο. Η διευκρίνιση αφορά στην περίπτωση του *Twitter Reverb*, καθώς δεν κατέστη εφικτή η λήψη διευκρίνισης σχετικά με το αν υποστηρίζει οπτικοποίηση δεδομένων σε πραγματικό χρόνο. Η έρευνα σε οπτικοποιήσεις που παρήχθησαν μέσω αυτού του εργαλείου δείχνει ότι τουλάχιστον παρέχει τη δυνατότητα οπτικοποίησης δεδομένων σε πολύ σύντομο χρονικό διάστημα από τη συλλογή τους.

αφορά στα διαφορετικά στάδια κατασκευής, από την επιλογή των δεδομένων μέχρι και τον σχεδιασμό της οπτικοποίησης.

Στη συνέχεια, εξετάζονται μεμονωμένα οι επιλεγμένες περιπτώσεις οπτικοποίησης ως προς τις ακόλουθες παραμέτρους που, βάσει των προηγούμενων κεφαλαίων, διαμορφώνουν το τελικό αποτέλεσμα που παρέχεται στον χρήστη:

- συγκέντρωση και επεξεργασία δεδομένων
- τοποθέτηση δεδομένων σε πλαίσιο
- επιλογή κατάλληλων οπτικών κωδικοποιήσεων για την παρουσίαση των δεδομένων στον χρήστη
- εμπλουτισμός με (1) δυνατότητες αλληλεπίδρασης και (2) υποστηρικτικό υλικό που παρέχει ερμηνευτικό πλαίσιο στον χρήστη.

5.1.1 How riot rumours spread on Twitter

Ο *Guardian* αποτελεί διεθνώς πρωτοπόρο μέσο στον τομέα της δημοσιογραφίας δεδομένων, με το ιστολόγιο *Datablog* να αποτελεί σημείο αναφοράς για το είδος. Οπτικοποιήσεις δεδομένων συχνά συνοδεύουν ή συνοδεύονται από δημοσιογραφικά κείμενα παραδοσιακού τύπου. Ένα τέτοιο δείγμα γραφής αποτελεί το εγχείρημα "*Reading the Riots*"⁶⁴, το οποίο πλαισίωσε τη δημοσιογραφική κάλυψη των αναταραχών του 2011 στην Αγγλία (Lewis and Newburn, 2011).

Εν προκειμένω, αντικείμενο εξέτασης αποτελεί η σειρά αλληλεπιδραστικών οπτικοποιήσεων "*How riot rumours spread on Twitter*", που δημιουργήθηκε στο πλαίσιο του εν λόγω εγχειρήματος. Οι εν λόγω οπτικοποιήσεις βασίζονται σε δεδομένα που προέρχονται από τις αναρτήσεις των χρηστών του κοινωνικού μέσου *Twitter*, και αποσκοπούν στην παρουσίαση της εξέλιξης της δημιουργίας

⁶⁴ Πρόκειται για συνεργασία του *Guardian* και του *London School of Economics*, "με στόχο την παραγωγή κοινωνικής έρευνας που βασίζεται σε αποδείξεις για την ερμηνεία των αιτίων εξάπλωσης των αναταραχών στην Αγγλία" (Lewis and Newburn, 2011).

και της αποκλιμάκωσης των φημών κατά τη διάρκεια των αναταραχών. Η επεξεργασία των δεδομένων που συλλέχθηκαν από τα τιτιβίσματα (tweets) κατάφερε να ανατρέψει τη μέχρι τότε επικρατούσα άποψη ότι το *Twitter* χρησιμοποιήθηκε κυρίως για την οργάνωση εγκληματικών δραστηριοτήτων (Dant and Richards, 2011)⁶⁵.

Το σύνολο δεδομένων που αναπαρίσταται προέκυψε από τη συλλογή των τιτιβισμάτων που σχετίζονταν με τις αναταραχές· σχετικές ετικέτες (hashtags) περιλαμβάνονταν σε 2,6 εκατομμύρια τιτιβίσματα (Ball and Lewis, 2011). Όπως αναφέρεται σε άρθρο που παρουσιάζει εκτενώς τη μεθοδολογία που εφαρμόστηκε, για την επεξεργασία και την ανάλυση των τιτιβισμάτων, η ομάδα εργασίας του *Guardian* συνεργάστηκε με διεπιστημονική ομάδα ερευνητών από αγγλικά Πανεπιστήμια (Dant and Richards, 2011).

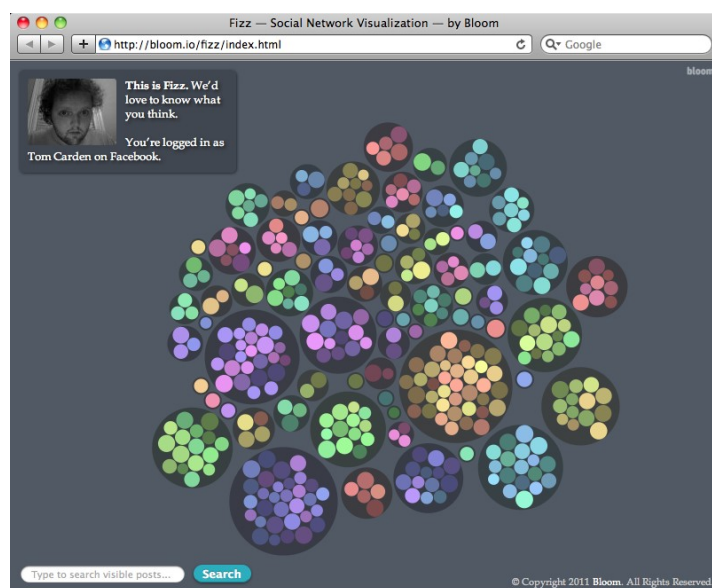
Προκειμένου να περιορίσουν το εύρος της έρευνας, οι δημοσιογράφοι κατέληξαν σε επτά βασικές φήμες, οι οποίες σχετίζονταν με συνολικά 10.000 τιτιβίσματα (Richards and Lewis, 2011). Για την ένταξη των τιτιβισμάτων σε συστάδες, οι ακαδημαϊκοί ερευνητές χρησιμοποίησαν αλγόριθμο απόστασης προκειμένου να εντοπίσουν αυτά με την κοντινότερη απόσταση, δηλ. τη μεγαλύτερη συνάφεια (Dant and Richards, 2011). Με αυτόν τον τρόπο, κάθε συστάδα περιλάμβανε τις αναπαραγωγές (retweets) του αρχικού τιτιβίσματος. Ως αποτέλεσμα, τα τιτιβίσματα που περιλαμβάνονταν σε καθένα από τα επτά μικρά σύνολα δεδομένων οργανώθηκαν σε συστάδες.

Επιπρόσθετα, επιχειρήθηκε η εξέταση του τόνου κάθε τιτιβίσματος σε σχέση με την αναπαραγόμενη φήμη. Αρχικά, τα τιτιβίσματα αναλύθηκαν με τη χρήση αλγορίθμων ανάλυσης συναισθήματος (Procter, 2011). Στη συνέχεια και για την επιβεβαίωση της ορθότητας των αποτελεσμάτων, εθελοντές ανέλαβαν να κωδικοποιήσουν τα τιτιβίσματα σύμφωνα με τη στάση (θετική, αρνητική ή ουδέτερη) που εξέφραζε ο εκάστοτε συντάκτης σχετικά με την αναπαραγόμενη

⁶⁵ Βάσει αυτής, προκύπτει ότι το 8% των τιτιβισμάτων αφορούσαν στην αυτο-οργάνωση πολιτών για την αποκατάσταση των ζημιών από τις αναταραχές (Dant and Richards, 2011).

φήμη (Procter, 2011). Τέλος, οι χαρακτηρισμοί από αυτές τις δύο πηγές συγκεράστηκαν για τη διαμόρφωση των τελικών αρχείων (Procter, 2011· Dant and Richards, 2011).

Σε ό,τι αφορά τον οπτικό σχεδιασμό της βασικής οπτικοποίησης που αναπαριστά την πορεία διάδοσης κάθε φήμης, πηγή έμπνευσης αποτέλεσε το εγχείρημα *Fizz* της νεοφυούς επιχείρησης *Bloom* (Dant and Richards, 2011)⁶⁶.



Εικόνα 31: Εκτύπωση οθόνης *Fizz*
(Πηγή: Carden, 2011)

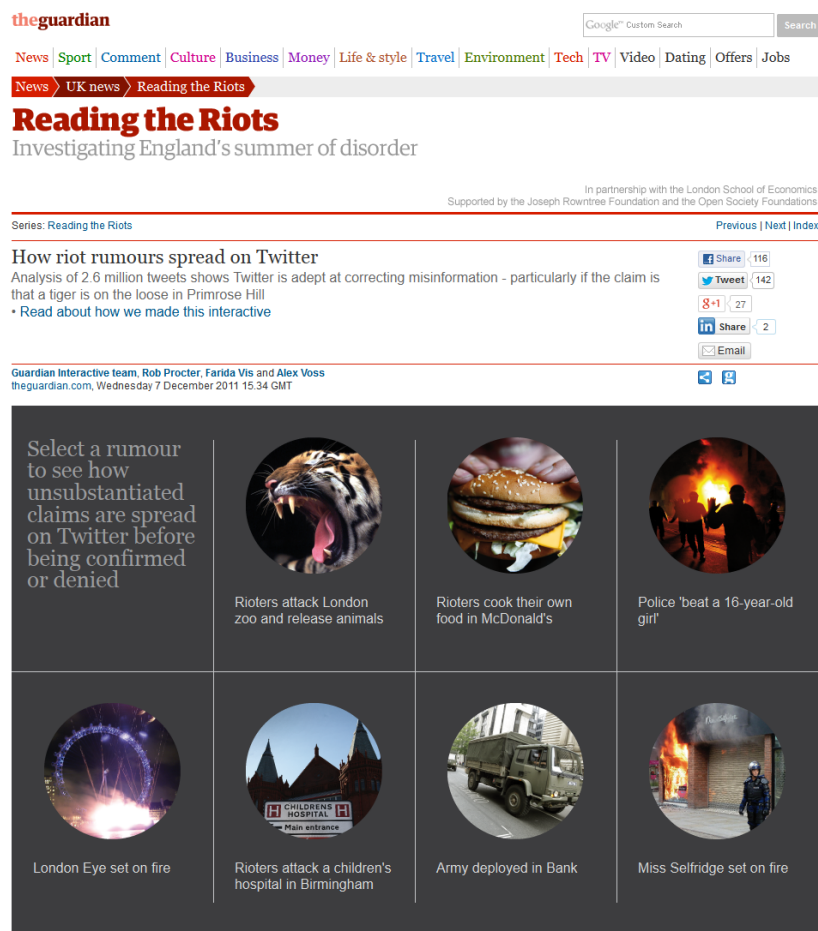
Στο πλαίσιο αυτού και για τον σχεδιασμό της βασικής οπτικοποίησης "*How riot rumours spread on Twitter*", κάθε συστάδα αναπαρίσταται ως κύκλος που περιλαμβάνει μικρότερους, οι οποίοι αντιστοιχούν στις αναπαραγωγές του δεδομένου τιτιβίσματος. Παράλληλα, το μέγεθος κάθε κύκλου κωδικοποιεί την επιρροή του συντάκτη κάθε τιτιβίσματος: όσο μεγαλύτερος ο κύκλος, τόσο μεγαλύτερος ο αριθμός ακολούθων (followers) του αντίστοιχου συντάκτη.

Επιπρόσθετα, η πληροφορία για τον τόνο κάθε τιτιβίσματος, κωδικοποιείται χρωματικά. Συγκεκριμένα, επιλέχθηκε μία περιορισμένη και έντονη χρωματικά παλέτα με σημασιολογική φόρτιση: κόκκινο χρώμα για την εναντίωση με την αναπαραγόμενη φήμη και πράσινο για την υποστήριξη. Με κίτρινο και γκρι

⁶⁶ Η εν λόγω επιχείρηση διέκοψε τη λειτουργία της τον Νοέμβριο του 2012 (Ben, 2012).

χρωματίστηκαν τα τιτιβίσματα που εξέφραζαν απορία ή σχόλιο, αντιστοίχως.

Οι προγραμματιστές που υλοποίησαν την οπτικοποίηση κατασκεύασαν μία αλληλεπιδραστική χρονογραμμή (interactive timeline), που απεικονίζει σε βίντεο την εξέλιξη της συζήτησης για κάθε φήμη (Dant and Richards, 2011). Διαδοχικές στατικές εικόνες της βασικής οπτικοποίησης συνιστούν τα καρέ του βίντεο. Καθεμιά από αυτές αντιστοιχεί σε συγκεκριμένη χρονική στιγμή στην εξέλιξη της συζήτησης, όπως προκύπτει από τον κινούμενο δείκτη επί της χρονογραμμής. Η οπτικοποίηση πλαισιώνεται από σύντομα δημοσιογραφικά σχόλια για τα βασικότερα τιτιβίσματα, προκειμένου να διευκολύνει τον χρήστη να εντάξει την οπτική αναπαράσταση δεδομένων σε ένα ευρύτερο πλαίσιο και να καταλήξει στα δικά του συμπεράσματα.

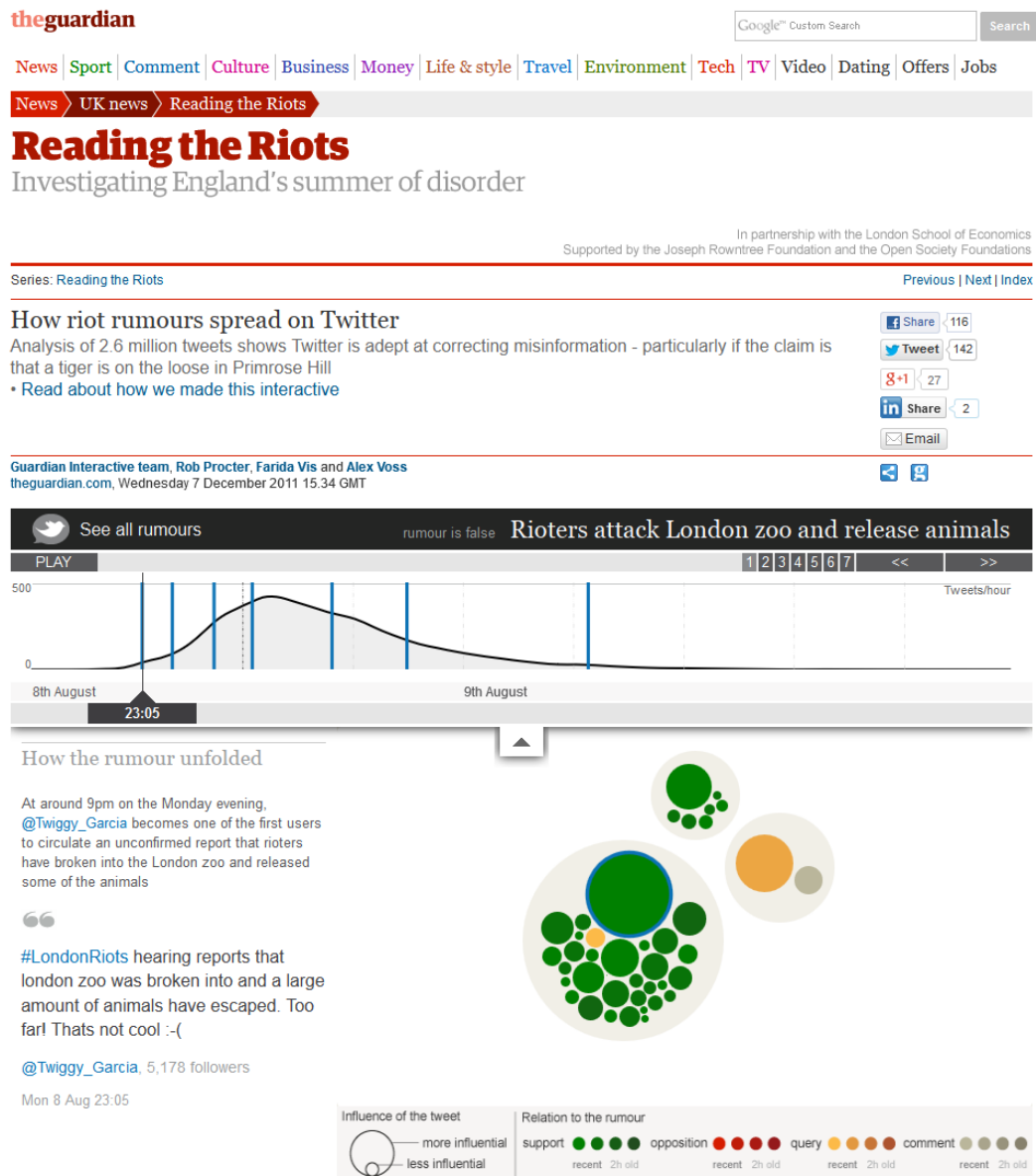


Εικόνα 32: Αρχική οθόνη "How riot rumours spread on Twitter"
(Πηγή: Guardian Interactive Team et al., 2011)

Η αρχική οθόνη της οπτικοποίησης που τελικά δημοσιεύθηκε στην

ιστοσελίδα του Guardian (Εικόνα 32) περιλαμβάνει επτά εικόνες, αντιπροσωπευτικές για κάθε μία από τις υπό εξέταση φήμες, και προτρέπει τον χρήστη να επιλέξει μία για να δει "πώς ανεπιβεβαιώτες φήμες διαδίδονται μέσω Twitter προτού αποδειχθεί αν ισχύουν ή είναι ανυπόστατες".

Επιλέγοντας μία φήμη, ο χρήστης μεταβαίνει στη σελίδα της αντίστοιχης οπτικοποίησης (Εικόνα 33).



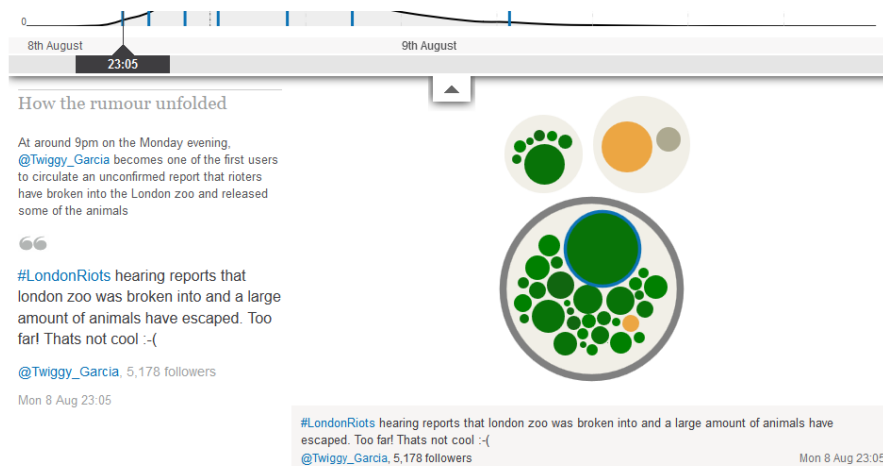
Εικόνα 33: Σελίδα οπτικοποίησης "Rioters attack London zoo and release animals"
(Πηγή: Guardian Interactive Team et al., 2011)

Στο πάνω μέρος, έχουν τοποθετηθεί τα κουμπιά ελέγχου της οπτικοποίησης:

η επιλογή έναρξης ή παύσης της οπτικοποίησης και η δυνατότητα μετάβασης σε άλλη φήμη ή επιστροφή στην αρχική οθόνη. Αμέσως πιο κάτω έχει τοποθετηθεί η χρονογραμμή που απεικονίζει τις αυξομειώσεις των τιτιβισμάτων σε βάθος χρόνου, με δυνατότητα απόκρυψης, ώστε να απελευθερώνεται χώρος. Οι κάθετες μπλε γραμμές υποδηλώνουν τη στιγμή που αναρτάται κάποιο τιτίβισμα που επηρεάζει την εξέλιξη της συζήτησης.

Στο κέντρο ο χώρος χωρίζεται σε δύο μέρη: η δεξιά και πιο ευρεία στήλη αποτελεί την κύρια περιοχή οπτικοποίησης ενώ η αριστερή στήλη αναφέρει το χαρακτηριστικότερο τιτίβισμα τη δεδομένη στιγμή της συζήτησης (αυτό που αντιστοιχεί στην προγενέστερη κάθετη μπλε γραμμή επί της χρονογραμμής). Στο κάτω μέρος, εμφανίζεται υπόμνημα με πληροφορίες για τις συμβάσεις της συγκεκριμένης οπτικοποίησης: το μέγεθος των κύκλων και τη χρωματική κωδικοποίηση.

Με την κίνηση του ποντικιού πάνω από έναν κύκλο, εμφανίζεται στην περιοχή του υπομνήματος το περιεχόμενο του τιτιβίσματος που αναπαρίσταται (Εικόνα 34). Με αυτόν τον τρόπο, ο χρήστης ελέγχει τα υποκείμενα δεδομένα και εντοπίζει τυχόν αστοχίες.



Εικόνα 34: Εμφάνιση υποκείμενων δεδομένων
(Πηγή: Guardian Interactive Team et al., 2011)

Η εν λόγω οπτικοποίηση, ενώ επιτρέπει την αλληλεπίδραση του χρήστη με τα δεδομένα, συνίσταται στον εκ των προτέρων καθορισμό της χρονικής σειράς

παρουσίασης από τους δημιουργούς, ο οποίος επιβάλλεται από τη φύση του παρουσιαζόμενου θέματος. Επιπρόσθετα, δεν εντοπίζεται σύνδεσμος στο πρωτογενές σύνολο δεδομένων, που θα επέτρεπε την περαιτέρω ανάλυσή του από τον χρήστη, εφόσον αυτός επιθυμεί να εξετάσει και άλλες πτυχές του ζητήματος.

5.1.2 Twitter Reverb - Mentions on Twitter of #Ferguson

Το *Twitter Reverb* είναι μία πλατφόρμα που αναπτύχθηκε από το κοινωνικό μέσο *Twitter* σε συνεργασία με την εταιρεία *Periscope* που ειδικεύεται στην οπτικοποίηση δεδομένων. Μέσω αυτής, παρέχεται η δυνατότητα οπτικοποίησης των τιτιβισμάτων που συντάσσονται σχετικά με ένα συγκεκριμένο θέμα ή γεγονός, χωρίς να απαιτούνται προγραμματιστικές γνώσεις (Rogers, 2014).

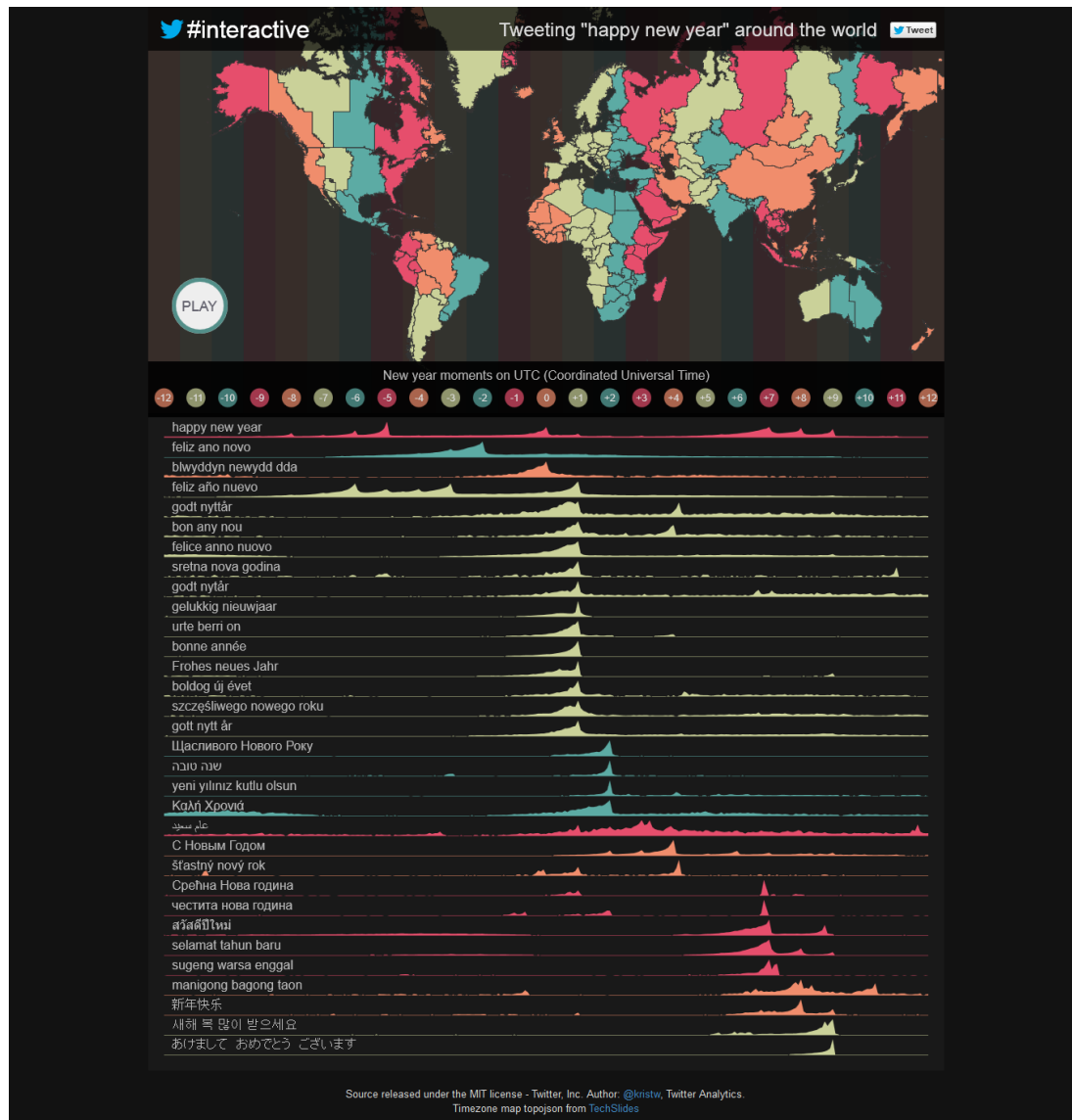
Γενικά, το *Twitter* έχει παράδοση στη δημιουργία αλληλεπιδραστικών οπτικοποιήσεων που βασίζονται σε τιτιβίσματα των χρηστών του. Για τις ανάγκες αυτής της δραστηριότητας, έχει δημιουργηθεί μία in-house ομάδα που περιλαμβάνει προγραμματιστές, επιστημονικούς υπεύθυνους δεδομένων και σχεδιαστές, η οποία αναλαμβάνει να κατασκευάσει αλληλεπιδραστικές οπτικοποιήσεις που πραγματεύονται διαφορετικά ζητήματα της επικαιρότητας (Rogers, 2014). Τον Μάιο του 2013, σε αυτή την ομάδα εντάχθηκε ο Simon Rogers⁶⁷ ως ο πρώτος "Επιμελητής Δεδομένων"⁶⁸ με βασική αρμοδιότητα "την αφήγηση ιστοριών που βασίζονται σε κάποια από τα συναρπαστικότερα δεδομένα που κυκλοφορούν" (Rogers, 2013b). Η προσθήκη του Rogers στο δυναμικό της ειδικής αυτής ομάδας, λαμβάνοντας υπ' όψιν τα δείγματα γραφής που είχε δώσει κατά τη θητεία του στον *Guardian*, εκτιμάται ότι υποδηλώνει το έντονο ενδιαφέρον του *Twitter* για την αξιοποίηση των δεδομένων που συλλέγει από τη δραστηριότητα των χρηστών του.

Οι οπτικοποιήσεις που κατασκευάζονται από αυτή την εσωτερική ειδική

⁶⁷ Δημιούργησε το *DataBlog* του *Guardian*, το 2009.

⁶⁸ Απόδοση του αγγλικού όρου "*Data Editor*".

ομάδα, συνήθως δημιουργούνται ειδικά για συγκεκριμένες περιστάσεις και αφορούν σύνολα δεδομένων που δεν μεταβάλλονται σε πραγματικό χρόνο. Κάθε οπτικοποίηση είναι μοναδική, ως αποτέλεσμα διαφορετικών σχεδιαστικών αποφάσεων που εξυπηρετούν διαφορετικούς σκοπούς (π.χ. Ρίος, 2013· Twitter, 2014). Μία από αυτές τις οπτικοποιήσεις παρατίθεται στην ακόλουθη εικόνα (Εικόνα 35):



Εικόνα 35: Πλήρης ανάπτυξη σελίδας "Tweeting 'happy new year' around the world" (Πηγή: Twitter, 2014)

Ωστόσο, η επεξεργασία των δεδομένων των τιτιβισμάτων πριν την οπτικοποίηση, είναι μία περίπλοκη διαδικασία με αρκετές δυσκολίες, που

αναγνωρίζονται ακόμη και από τους έχοντες προγραμματιστικές γνώσεις (Rios and Lin, 2012: 25). Το *Twitter Reverb* καλύπτει την ανάγκη άμεσης και εύκολης κατασκευής οπτικοποίησης, σχεδόν παράλληλα με την εξέλιξη του γεγονότος, τόσο για σκοπούς εσωτερικής παρακολούθησης από το *Twitter*, όσο και για τη διευκόλυνση δημοσιογράφων στην κάλυψη ενός θέματος επικαιρότητας (Rogers, 2014). Όπως επισημαίνει ο Buckhouse⁶⁹ (2014 σε Rogers, 2014), "*έμφαση δόθηκε στην ιστορία που θα ήθελε να πει ένας δημοσιογράφος, και όχι απλώς η επινόηση ενός νέου τρόπου παρουσίασης δεδομένων*".

Κατά τον προσδιορισμό των προδιαγραφών υλοποίησής του, τέθηκαν οι ακόλουθες απαιτήσεις που έπρεπε να ικανοποιεί το νέο εργαλείο (Rogers, 2014):

- να παρέχει δυνατότητα καταχώρισης σημειώσεων (annotations), ώστε ο δημιουργός της οπτικοποίησης να μπορεί να επισημάνει επιπλέον πληροφορίες σχετικά με τα αναπαριστώμενα δεδομένα
- να λειτουργεί σε κινητές συσκευές
- να προβάλλεται στην τηλεόραση
- να ενσωματώνεται σε ιστοσελίδες
- να απεικονίζει την εξέλιξη του διαλόγου σχετικά με ένα ζήτημα
- να συνιστά μία κινηματογραφική και συναρπαστική εμπειρία.

Το *Twitter Reverb* ξεκίνησε ως πειραματικό εγχείρημα, το καλοκαίρι του 2014. Στην παρούσα φάση, πρόσβαση έχουν επιλεγμένοι χρήστες που καλύπτουν γεγονότα ειδησεογραφικού ενδιαφέροντος. Στη συνέχεια, εξετάζεται μία από τις οπτικοποιήσεις που έχουν κατασκευαστεί στο σύντομο χρονικό διάστημα λειτουργίας του· αυτή πλαισιώνει δημοσιογραφικό άρθρο, το οποίο αναφέρεται στις αναταραχές που εκδηλώθηκαν στις ΗΠΑ, με αφορμή την δολοφονία ενός αфроαμερικανού πολίτη από αστυνομικό, τον Αύγουστο του 2014 (Zak, 2014a).

Η εν λόγω οπτικοποίηση απεικονίζει την εξέλιξη της συζήτησης το χρονικό διάστημα από τις 9 έως και τις 18 Αυγούστου, και δημοσιεύθηκε στην ηλεκτρονική έκδοση της *Wall Street Journal* στις 18 Αυγούστου (Εικόνα 36).

⁶⁹ Σχεδιαστής στο Twitter.

Πρόκειται δηλαδή για μία σχεδόν σε πραγματικό χρόνο⁷⁰ παρακολούθηση της εξέλιξης των σχετικών τιτιβισμάτων.



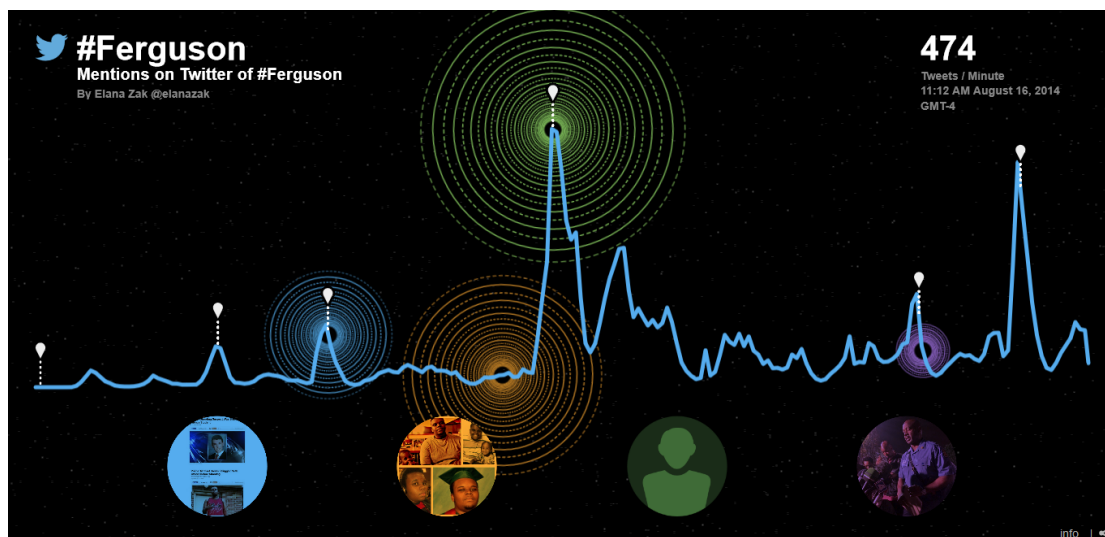
Εικόνα 36: Πλήρης ανάπτυξη σελίδας άρθρου (Πηγή: Zak, 2014a)

Για τη διαμόρφωση του συνόλου δεδομένων προς οπτικοποίηση, συλλέχθηκαν τα τιτιβίσματα που περιλάμβαναν την ετικέτα "*#ferguson*"⁷¹. Η δημιουργός της οπτικοποίησης εκτιμάται ότι, μέσω ειδικής επιλογής που παρέχει το σχεδιαστικό περιβάλλον του *Twitter Reverb*, έθεσε το εν λόγω κριτήριο επιλογής των δεδομένων προς οπτικοποίηση.

⁷⁰ Η εκτίμηση βασίζεται στη χρονική στιγμή κυκλοφορίας του άρθρου (18 Αυγούστου, 16:20 ET), που είναι πολύ κοντά στη λήξη της περιόδου που αφορούν τα αναπαριστώμενα δεδομένα (18 Αυγούστου, 14:13 GMT-4).

⁷¹ Η πόλη που διαπράχθηκε η δολοφονία.

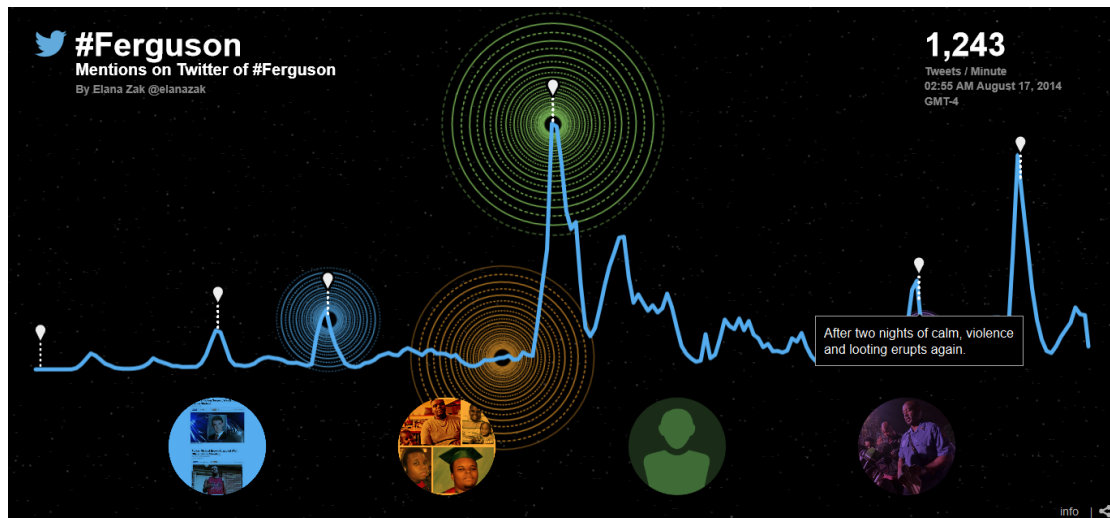
Στην Εικόνα 37, φαίνεται η αρχική οθόνη της οπτικοποίησης. Στο πάνω αριστερό μέρος, φαίνεται ο τίτλος της οπτικοποίησης και τα στοιχεία της δημιουργού.



Εικόνα 37: Οπτικοποίηση "Mentions on Twitter of #Ferguson"
(Πηγή: Zak, 2014b)

Στο κέντρο της οθόνης βρίσκεται ένα γράφημα γραμμής, μέσω του οποίου είναι ορατές οι αυξομειώσεις των σχετικών τιτιβισμάτων. Λόγω της εκτεταμένης περιόδου που καλύπτει η εν λόγω οπτικοποίηση, κάθε σημείο της γραμμής φαίνεται πως αφορά χρονική περίοδο δέκα περίπου λεπτών. Η κίνηση του δείκτη στον χώρο περιβάλλει τη γραμμή εμφανίζει, στην πάνω δεξιά γωνία της οθόνης, τον μέσο όρο τιτιβισμάτων ανά λεπτό και τον χρόνο καταγραφής τους.

Τα έξι λευκά σημεία επισήμανσης που προεξέχουν, αντιστοιχούν σε καταχωρίσεις σημειώσεων του δημιουργού που βοηθούν τον χρήστη να εντάξει την πληροφορία που λαμβάνει σε ένα γενικότερο πλαίσιο. Και σε αυτή την περίπτωση η κίνηση του δείκτη πάνω από ένα από τα λευκά σημεία, εμφανίζει την επεξήγηση του δημιουργού. Για παράδειγμα, η προτελευταία σημείωση εξηγεί για ποιον λόγο αυξάνονται σημαντικά τα σχετικά τιτιβίσματα ενώ διαφαινόταν μείωση του ενδιαφέροντος (Εικόνα 38).



Εικόνα 38: Εμφάνιση σημείωσης
(Πηγή: Zak, 2014b)

Στο κάτω μέρος της οθόνης, εμφανίζονται τέσσερις κύκλοι: καθένας εξ αυτών αντιστοιχεί σε ένα τιτίβισμα που κατά την κρίση του συντάκτη είναι σημαντικό. Εφόσον το εν λόγω τιτίβισμα περιέχει φωτογραφία, αυτή γεμίζει τον κύκλο. Παράλληλα, σε κάποια σημεία της γραμμής, ακουμπούν χρωματιστοί κύκλοι. Καθένας από αυτούς τους κύκλους έχει ως κέντρο ένα σημείο της γραμμής και αντιστοιχεί σε ένα από τα τιτιβίσματα που έχει ξεχωρίσει ο δημιουργός στο κάτω μέρος της οθόνης: ο χρήστης συνδέει νοερά κάθε τιτίβισμα με τον αντίστοιχο κύκλο επί της γραμμής, βάσει χρώματος⁷². Οι κύκλοι επί της γραμμής περιβάλλονται από ομόκεντρους ιδίου χρώματος. Ο χρήστης, παρατηρώντας τη συνολική επιφάνεια βάσει της ακτίνας του μεγαλύτερου εξ αυτών, μπορεί να σχηματίσει μία γενική άποψη για ποιο από τα τέσσερα τιτιβίσματα είχε μεγαλύτερη απήχηση, δηλαδή περισσότερες αναπαραγωγές.

Τέλος, πατώντας έναν από τους κύκλους στο κάτω μέρος της οθόνης, ο χρήστης αποκτά πρόσβαση στο περιεχόμενο του τιτιβίσματος, δηλαδή στο κείμενο ή/και στη φωτογραφία που περιέχονται στην ανάρτηση (Εικόνα 39). Αρχικά, το περιεχόμενο καλύπτεται από διαφανή χρωματική επίστρωση και

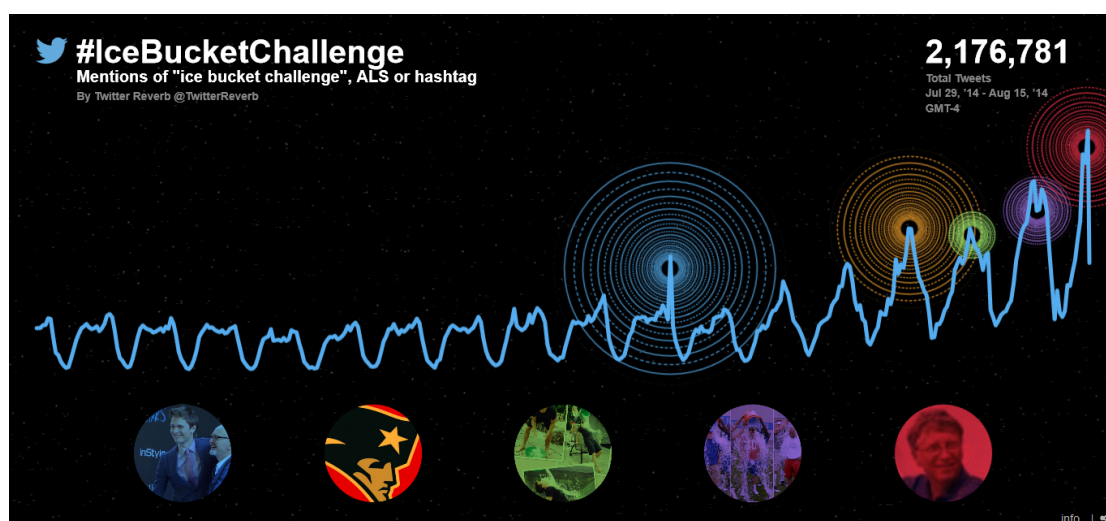
⁷² Συνδέει δηλαδή τη χρωματική επίστρωση στην εικόνα του τιτιβίσματος με το χρώμα του κύκλου επί της γραμμής.

εμφανίζονται στοιχεία σχετικά με το πλήθος χρηστών που το αναπαρήγαγαν ή το χαρακτήρισαν ως αγαπημένο, καθώς και οι τρεις από αυτούς τους χρήστες με τη μεγαλύτερη επιρροή. Ο χρήστης έχει τη δυνατότητα να αφαιρέσει αυτή την επιστροφή για να δει ευκρινώς το περιεχόμενο.

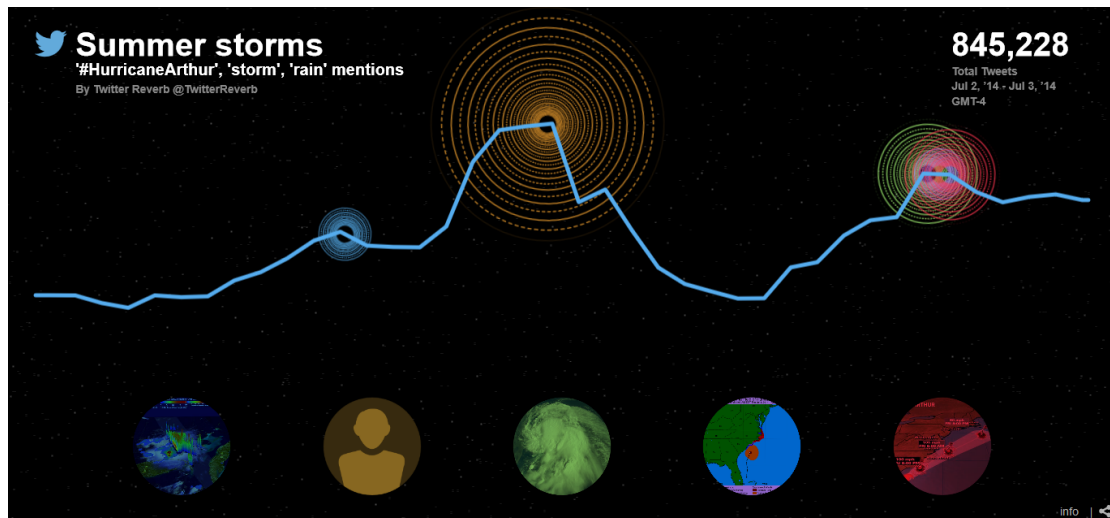


Εικόνα 39: Επιλογή τιτιβίσματος
(Πηγή: Zak, 2014b)

Επισημαίνεται ότι το τελικό αποτέλεσμα που προκύπτει από τη χρήση του *Twitter Reverb* είναι τυποποιημένο: όλες οι οπτικοποιήσεις που παράγονται μέσω αυτού έχουν την ίδια γενική εμφάνιση, αν και διαφέρουν σε περιεχόμενο. Ωστόσο, ο σχεδιαστής έχει τη δυνατότητα προσαρμογής της οπτικοποίησης με την καταχώριση σημειώσεων και την επιλογή φωτογραφιών. Στις ακόλουθες εικόνες, παρατίθενται σχετικά παραδείγματα:



Εικόνα 40: Οπτικοποίηση #IceBucketChallenge μέσω *Twitter Reverb*
(Πηγή: Twitter Reverb, 2014a)



Εικόνα 41: Οπτικοποίηση #HurricaneArthur κ.λπ. μέσω *Twitter Reverb*
(Πηγή: Twitter Reverb, 2014b)

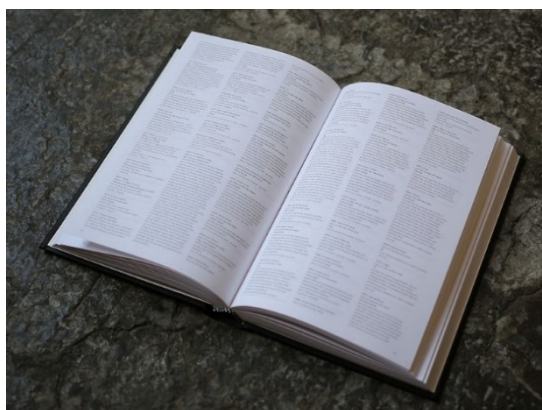
Η οπτικοποίηση που εξετάστηκε, όπως και άλλες που παράγονται μέσω του εργαλείου *Twitter Reverb*, χαρακτηρίζεται από απλότητα. Οι παρεχόμενες δυνατότητες αλληλεπίδρασης αφορούν (1) στην εμφάνιση αριθμητικών ή λεκτικών επεξηγήσεων σχετικά με τις τιμές που απεικονίζονται στο γράφημα γραμμής, και (2) στην πρόσβαση σε περιορισμένο αριθμό τιτιβισμάτων που κρίνονται ως σημαντικά από τον δημιουργό της οπτικοποίησης, χωρίς παροχή συνδέσμου στο πρωτογενές σύνολο δεδομένων. Επισημαίνεται ότι δεν παρέχεται βοηθητικό υλικό επεξήγησης των συμβάσεων της οπτικοποίησης, εκτός από μία γενική περιγραφή στο κείμενο του άρθρου που την περιλαμβάνει. Επομένως, ο χρήστης πλοηγείται διαισθητικά και αποκωδικοποιεί σταδιακά τα αναπαριστώμενα δεδομένα.

5.1.3 Feltron 2013 Annual Report

Το 2005, ο Nicholas Felton δημιούργησε -και στη συνέχεια κυκλοφόρησε- την πρώτη ετήσια αναφορά που περιελάμβανε οπτικές αναπαραστάσεις δεδομένων που προέρχονταν από προσωπικές αναμνήσεις, φωτογραφίες και δεδομένα από την υπηρεσία *last.fm*. Η απήχηση ήταν τέτοια που έκτοτε σχεδιάζει και κυκλοφορεί ετήσιους απολογισμούς της ζωής του, οπτικοποιώντας

δεδομένα που αφορούν σε διάφορα είδη δραστηριοτήτων του (Feltron, 2014a).

Η αναφορά για το έτος 2013 περιλαμβάνει οπτικοποιήσεις που προέρχονται από δεδομένα και μετα-δεδομένα σχετικά με τις επικοινωνίες του: συζητήσεις, μηνύματα σύντομου κειμένου (SMS), τηλεφωνικές κλήσεις, ηλεκτρονική και φυσική αλληλογραφία, καθώς και ανταλλαγές μηνυμάτων στο Facebook (Feltron, 2014b). Πρόκειται για δεδομένα που αφορούν σε δραστηριότητες του Felton στον φυσικό κόσμο, συνδυασμένα με δεδομένα σχετικά με τις αλληλεπιδράσεις του στον ψηφιακό κόσμο. Ως αποτέλεσμα εκτεταμένης αυτό-καταγραφής της επικοινωνίας του με ανθρώπους, άλλοτε με αυτόματο τρόπο και άλλοτε με την καταβολή σημαντικής προσπάθειας εκ μέρους του να καταγράψει οποιαδήποτε συνομιλία για να την ποσοτικοποιήσει στη συνέχεια (Evans, 2014), το σύνολο δεδομένων που προκύπτει και η δεδηλωμένη πρόθεσή του να εξετάσει και να αποκαλύψει μοτίβα από την οπτικοποίηση αυτού συνιστούν έκφραση των αξιών που συμμερίζονται οι υποστηρικτές του *Ποσοτικοποιημένου Εαυτού*.



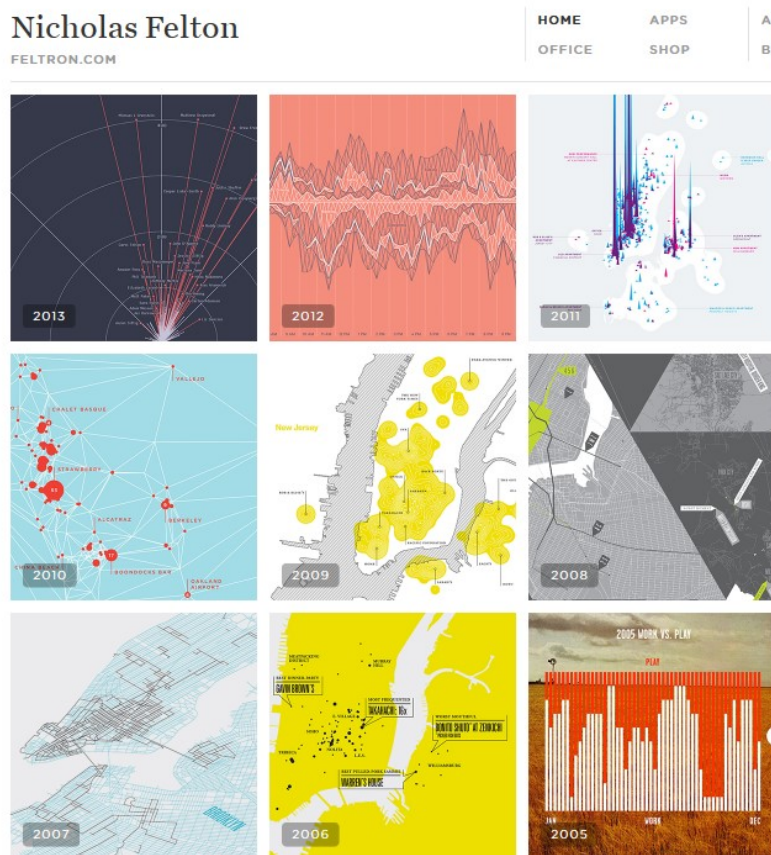
Εικόνα 42: Καταγραφή συζητήσεων από Felton σε 400 σελίδες⁷³
(Πηγή: Feltron, 2014c)

Η ετήσια αναφορά διατίθεται σε έντυπη μορφή σε αριθμημένα κομμάτια για συλλέκτες, και σε ηλεκτρονική μορφή στον προσωπικό του ιστοχώρο. Η διάθεση της συγκεκριμένης σειράς οπτικοποιήσεων δεδομένων σε έντυπη μορφή συνιστά ένα ενδιαφέρον γεγονός, λόγω της υβριδικής φύσης των δεδομένων που αναπαρίστανται. Στη συνέχεια, εξετάζεται η οπτικοποίηση όπως διατίθεται στο

⁷³ Σύμφωνα με δήλωσή του (Evans, 2014).

Διαδίκτυο, ως οπτική αναπαράσταση δεδομένων με περιορισμένες δυνατότητες αλληλεπίδρασης που αφορούν στην πλοήγηση μεταξύ σελίδων.

Στην αρχική σελίδα του ιστοχώρου εμφανίζονται εικόνες, καθεμιά εκ των οποίων αντιστοιχεί σε μία ετήσια αναφορά και φέρει ένδειξη για το έτος συλλογής των αναπαριστώμενων δεδομένων. Αυτή η χρονολογική διάταξη, από την πιο πρόσφατη ως την παλαιότερη, λειτουργεί ως παρουσίαση του συνολικού έργου του και ως ευρετήριο για πρόσβαση στην ετήσια αναφορά που ενδιαφέρει τον χρήστη.



Εικόνα 43: Ετήσιες αναφορές Feltron - Ευρετήριο
(Πηγή: Feltron, 2014d)

Πατώντας πάνω σε μία εικόνα, εν προκειμένω στην πρώτη που αφορά στην πλέον πρόσφατη, ο χρήστης μεταβαίνει στην περιοχή του ιστοχώρου που αντιστοιχεί στο συγκεκριμένο έτος (Εικόνα 44).

2013 Annual Report

FELTRON.COM

HOME

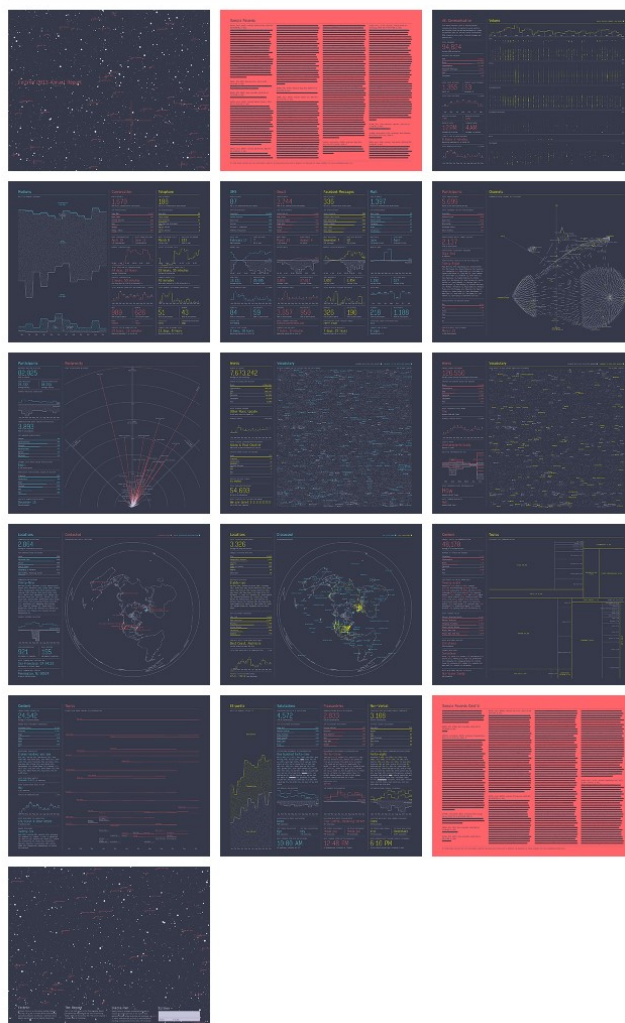
OFFICE

APPS

SHOP

ABOUT

BLOG



This report examines a year of Nicholas Felton's communication data. It aspires to uncover patterns and insights within the data and metadata of a large and personal data set. Sources include conversations, SMS, telephone calls, email, Facebook messages and physical mail.

Pre-Order

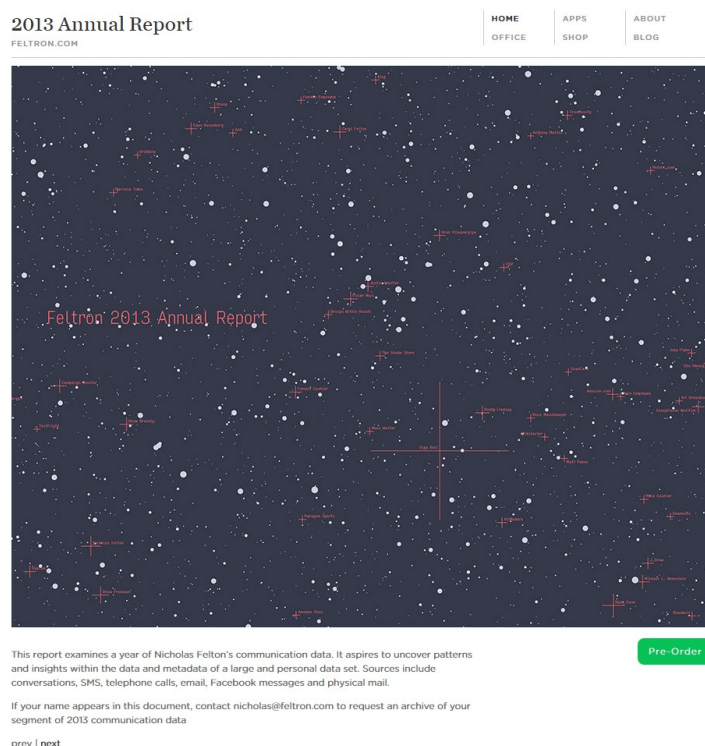
If your name appears in this document, contact nicholas@feltron.com to request an archive of your segment of 2013 communication data

Εικόνα 44: Ετήσια αναφορά *Feltron 2013* - Αρχική οθόνη (πλήρης ανάπτυξη)
(Πηγή: Feltron, 2014b)

Εκεί, παρατίθενται μικρογραφίες των σελίδων και ο χρήστης μπορεί να επιλέξει είτε την πρώτη από αυτές για να πλοηγηθεί σειριακά στην αναφορά ή μία από τις υπόλοιπες για να ορίσει ο ίδιος διαφορετικό σημείο εκκίνησης. Παρότι δεν είναι εύκολα αναγνώσιμα τα λεκτικά των επικεφαλίδων, η σύνθεση των μικρογραφιών λειτουργεί και ως εικονικός πίνακας περιεχομένων. Συνοδεύεται δε από μία σύντομη παράγραφο που αναφέρεται στο είδος των δεδομένων που αναπαρίστανται. Η παράγραφος αυτή παρέχει στον χρήστη

βασική γνώση του πλαισίου συλλογής τους και διευκολύνει την κατανόηση των όσων ακολουθούν⁷⁴.

Ξεκινώντας από την πρώτη μικρογραφία, προβάλλεται στην οθόνη το εξώφυλλο (Εικόνα 45).



Εικόνα 45: Έναρξη πλοήγησης στην ετήσια αναφορά *Feltron 2013*
(Πηγή: Feltron, 2014b)

Κάτω από αυτό, παραμένει η σύντομη εισαγωγική παράγραφος και εμφανίζονται οι επιλογές για μετάβαση στην επόμενη ή στην προηγούμενη σελίδα⁷⁵. Από το σημείο αυτό, η πλοήγηση στις σελίδες της αναφοράς είναι σειριακή: δεν είναι εφικτή η επιλογή μετάβασης σε οποιαδήποτε σελίδα, εκτός εάν ο χρήστης επιστρέψει στο ευρετήριο αναφορών για να επιλέξει εκ νέου την αναφορά που τον ενδιαφέρει.

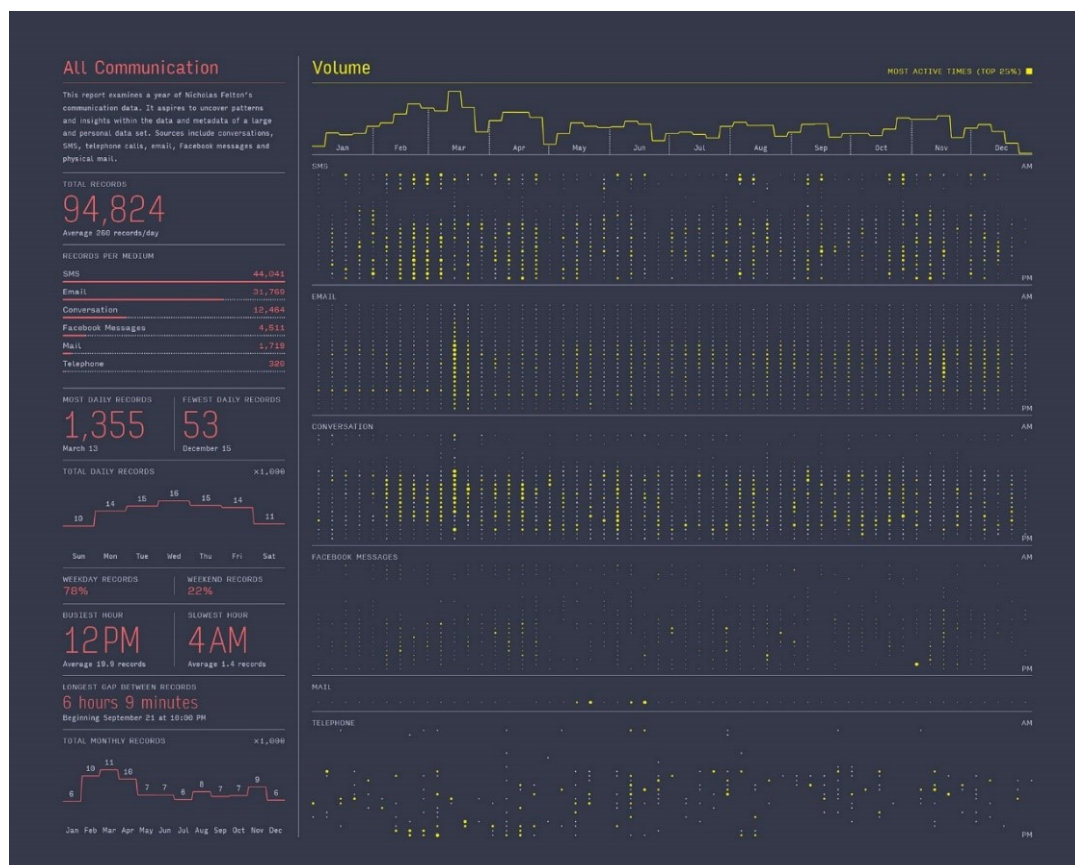
Η αναφορά περιλαμβάνει 12 σελίδες, καθεμιά εκ των οποίων συνιστά

⁷⁴ Σημειώνεται, για λόγους πληρότητας της παρουσίασης της συγκεκριμένης περίπτωσης, ότι, δίπλα από την εισαγωγική παράγραφο, έχει τοποθετηθεί ειδικό κουμπί έντονου πράσινου χρώματος που έλκει την προσοχή του χρήστη για προ-παραγγελία της αναφοράς σε έντυπη μορφή.

⁷⁵ Η δεύτερη επιλογή είναι αρχικά απενεργοποιημένη για ευνόητους λόγους.

σύνθεση διαφορετικών τύπων οπτικοποιήσεων. Ήδη από την αρχική οθόνη που περιλαμβάνει τις μικρογραφίες τους, σχηματίζεται μία πρώτη άποψη για τη δομή και τον τρόπο οργάνωσης των σελίδων. Κάθε σελίδα βασίζεται σε πλέγμα τεσσάρων στηλών ίδιου πλάτους. Στις περισσότερες περιπτώσεις, δύο ή τρεις από αυτές τις στήλες συγχωνεύονται για να περιλάβουν μία οπτική αναπαράσταση που απαιτεί αρκετό χώρο για να είναι αναγνώσιμη. Αναφορικά με το χρώμα, αυτό χρησιμοποιείται με σύνεση για να διαχωρίσει τις ενότητες περιεχομένου. Σε κάθε ενότητα περιεχομένου, ο Felton χρησιμοποιεί την αντίστοιχη απόχρωση για να τονίσει τα σημεία που επιθυμεί να στρέψει το βλέμμα του ο χρήστης.

Η πρώτη σελίδα της αναφοράς (Εικόνα 46) συνοψίζει τα δεδομένα που αφορούν στο σύνολο των επικοινωνιών του που έλαβαν χώρα εντός του 2013, και αποτελεί αντιπροσωπευτική εικόνα της δομής και του περιεχομένου των υπολοίπων.



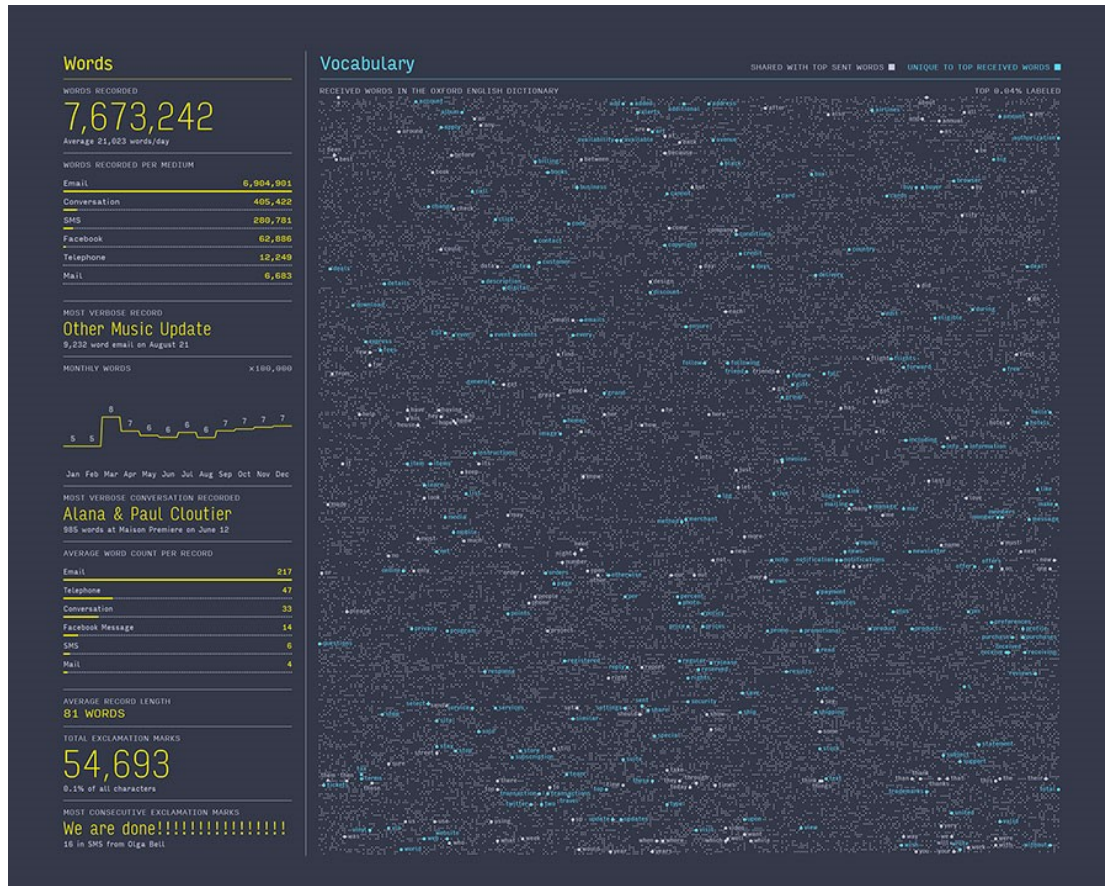
Εικόνα 46: Σύνοψη καταγραφών του συνόλου των επικοινωνιών
(Πηγή: Feltron, 2014b: 3)

Η σελίδα αυτή είναι χωρισμένη σε δύο τμήματα. Στην αριστερή στήλη, οι αριθμοί και τα βασικά οπτικά στοιχεία των γραφημάτων επισημαίνονται με έντονο κόκκινο χρώμα. Επιπρόσθετα, οι χρωματισμένοι αριθμοί διαφοροποιούνται από το υπόλοιπο κείμενο λόγω αυξημένου μεγέθους γραμματοσειράς. Επομένως, γίνεται αντιληπτό ότι εκτός από το χρώμα, ο Felton χρησιμοποιεί και την τυπογραφία για να κατευθύνει το βλέμμα του χρήστη στα σημεία που επιθυμεί. Το υπόλοιπο κείμενο, με λευκή και σημαντικά μικρότερη γραμματοσειρά, χρησιμοποιείται για να προσδιορίσει τις θεματικές υποενότητες, να υπομνηματίσει τους άξονες και να διευκρινίσει τις αναπαριστώμενες αξίες. Γενικά, σε αυτό το τμήμα της σελίδας, χρησιμοποιούνται συνήθεις μέθοδοι οπτικοποίησης με αφαιρετική διάθεση.

Η δεξιά περιοχή, με πλάτος που αντιστοιχεί σε τρεις στήλες, περιλαμβάνει γραφήματα που διατάσσονται κάθετα. Το πρώτο γράφημα γραμμής παρουσιάζει συγκεντρωτικά, για όλα τα κανάλια επικοινωνίας, τις αυξομειώσεις του συνολικού όγκου των επικοινωνιών μέσα στο έτος. Κάτω από αυτό, τοποθετούνται διαδοχικά γραφήματα που αναλύουν ανά κανάλι επικοινωνίας τον όγκο των επικοινωνιών: κάθε στήλη αντιστοιχεί στην εβδομάδα που αναπαρίσταται στο συγκεντρωτικό γράφημα, ενώ κάθε γραμμή αντιστοιχεί σε συγκεκριμένη ώρα της ημέρας. Η ένταση του χρώματος και το μέγεθος του εκάστοτε σημείου δεδομένων φαίνεται να υποδηλώνει την τιμή της μέτρησης του όγκου επικοινωνίας για το συγκεκριμένο μέσο κατά τη συγκεκριμένη χρονική περίοδο.

Η διάταξη των επιμέρους γραφημάτων οδηγεί στο συμπέρασμα ότι το άθροισμα των μετρήσεων που αντιστοιχούν σε όλα τα σημεία που βρίσκονται στην ίδια κάθετη ευθεία γραμμή αποτελεί την αξία του σημείου του συγκεντρωτικού γραφήματος που βρίσκεται στην προέκταση αυτής. Όπως και για τα υπόλοιπα γραφήματα που αναπτύσσονται σε δύο ή περισσότερες στήλες, φαίνεται ότι έχει καταβληθεί προσπάθεια να συμπεριληφθεί η μέγιστη δυνατή

πληροφόρηση με ισχυρά αφαιρετικό τρόπο, που όμως δυσχεραίνει την κατανόηση. Παρατηρείται ότι δεν υπάρχει αναφορά σε συγκεκριμένες τιμές, εκτός από το κλειδί στα δεξιά του τίτλου της περιοχής που περιλαμβάνει ένα μικρό κίτρινο τετράγωνο και το λεκτικό "πιο δραστήριες στιγμές (κορυφαίο 25%)".



Εικόνα 47: Καταμέτρηση εμφάνισης λέξεων, που έλαβε ως δέκτης ο Felton κατά τη διάρκεια των επικοινωνιών του (Πηγή: Feltron, 2014b: 8)

Η οπτικοποίηση που εξετάστηκε, εξ ορισμού δεν παρέχει δυνατότητες αλληλεπίδρασης του χρήστη με τα δεδομένα, καθώς έχει κατασκευαστεί για να διατεθεί κατ' αρχήν σε έντυπη μορφή. Στην ψηφιακή έκδοσή της, ο δημιουργός έχει ορίσει εκ των προτέρων τη σειρά εμφάνισης των σελίδων, χωρίς να αφήνει περιθώρια στον χρήστη να πλοηγηθεί ελεύθερα (όπως αν ξεφύλλιζε την έντυπη μορφή). Το σύνολο της αναφοράς αποπνέει μία σχεδιαστική κομψότητα και τη

διάθεση του δημιουργού για ένα άρτιο αισθητικά αποτέλεσμα⁷⁶. Αυτή η ιδιαίτερη φροντίδα της συνολικής αισθητικής εξηγείται από το υπόβαθρο του Felton, ο οποίος είναι απόφοιτος του *Rhode Island School of Design* και για χρόνια ασχολήθηκε με τον γραφικό σχεδιασμό (Essmaker and Essmaker, 2014).

5.1.4 Fitbit Dashboard

Οι συσκευές της εταιρείας *Fitbit* παρέχουν στους χρήστες τους τη δυνατότητα καταγραφής δεδομένων σχετικά με την καθημερινή φυσική δραστηριότητα, την κατανάλωση τροφής, τη μέτρηση βάρους και την ποιότητα ύπνου. Όπως έχει αναφερθεί στο Κεφάλαιο 2 πιο πάνω, το οικοσύστημα συσκευών αυτο-καταγραφής συνδυάζεται με ειδικά σχεδιασμένες διεπαφές που αναπαριστούν οπτικά τα συλλεχθέντα δεδομένα, με στόχο τη διευκόλυνση του χρήστη στον προσδιορισμό στόχων βελτίωσης της τρέχουσας φυσικής του κατάστασης και την παρακολούθηση της προόδου του. Η πρόσβαση στα παραγόμενα δεδομένα διενεργείται μέσω πιστοποιημένης εισόδου στην ιστοσελίδα ή/και της ειδικής εφαρμογής για έξυπνα κινητά τηλέφωνα.

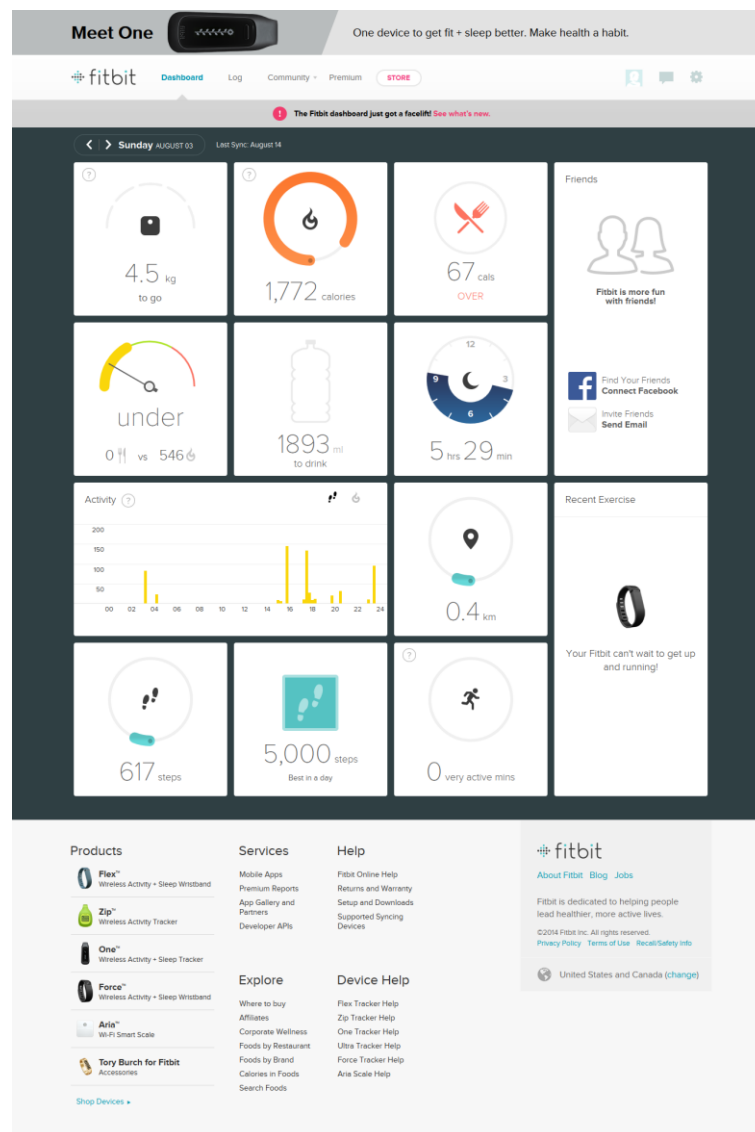
Για τη μελέτη της συγκεκριμένης περίπτωσης στο πλαίσιο της παρούσας εργασίας, χρησιμοποιήθηκαν οι συσκευές *Fitbit Flex* και *Fitbit Aria* για να δημιουργηθεί ένα ενδεικτικό σύνολο δεδομένων. Με τον τρόπο αυτόν καθίσταται δυνατή η -σε βασικό επίπεδο- εξέταση των εργαλείων οπτικής ανάλυσης που παρέχονται μέσω της σχετικής ιστοσελίδας⁷⁷. Παράλληλα, καταχωρίστηκαν χειροκίνητα μετρήσεις στην ειδική ενότητα καταγραφών, για τον εμπλουτισμό των δεδομένων που συλλέγονταν από τις εν λόγω συσκευές.

Μετά την πιστοποιημένη είσοδο στην ιστοσελίδα *fitbit.com*, ο χρήστης αποκτά πρόσβαση στο εξατομικευμένο *Dashboard* που περιλαμβάνει τη σύνοψη

⁷⁶ Παρατίθεται μία επιπλέον σελίδα της αναφοράς (Εικόνα 47), η οποία περιλαμβάνει σύνθεση διαφορετικών οπτικοποιήσεων, που παρουσιάζουν διαφορετικές πτυχές του μεγάλου συνόλου δεδομένων.

⁷⁷ Η ιστοσελίδα ανασχεδιάστηκε πρόσφατα και τέθηκε στη διάθεση του κοινού τον Ιούλιο του 2014 (Chris, 2014).

της καταγεγραμμένης δραστηριότητάς του. Στην Εικόνα 48, φαίνονται οι διαφορετικές μέθοδοι οπτικοποίησης που χρησιμοποιούνται, ανάλογα με το είδος δραστηριότητας που αφορούν τα αντίστοιχα δεδομένα. Σημειώνεται ότι κατά την είσοδο στην εφαρμογή για έξυπνα κινητά τηλέφωνα, η εικόνα του *Fitbit Dashboard* είναι διαφορετική, καθώς προσαρμόζεται στους περιορισμούς της μικρότερης οθόνης· στο Παράρτημα της παρούσας εργασίας, παρατίθενται ενδεικτικές οθόνες της εφαρμογής μέσω *iPhone*.



Εικόνα 48: Εξατομικευμένο *Fitbit Dashboard* - Πλήρης ανάπτυξη σελίδας
(Πηγή: Fitbit, 2014b)

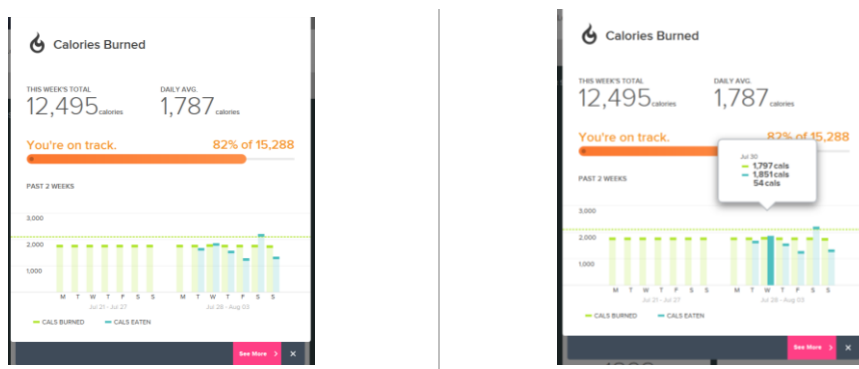
Σχετικά με τη δομή της βασικής περιοχής, αυτή είναι διαιρεμένη σε παραλληλόγραμμα, καθένα εκ των οποίων αντιστοιχεί σε μία κατηγορία

παρακολουθούμενης δραστηριότητας. Σε αρκετές περιπτώσεις, η κίνηση του κέρσορα πάνω από ένα παραλληλόγραμμα μεταβάλλει το περιεχόμενό του: προκαλεί εναλλαγή μεταξύ της προβολής του στόχου που έχει τεθεί και των δεδομένων που έχουν καταγραφεί, χωρίς όμως η προεπιλεγμένη παρουσίαση να αντιστοιχεί πάντα στη μία ή στην άλλη προβολή. Επιπρόσθετα, ο χρήστης μπορεί να αναδιατάξει τη σειρά εμφάνισης των παραλληλογράμμων, σύμφωνα με τις προτιμήσεις του, επιλέγοντας ποια παραλληλόγραμμα τον ενδιαφέρει περισσότερο να είναι άμεσα ορατά κατά την επίσκεψή του στην ιστοσελίδα. Η δυνατότητα παροχής βοήθειας παρέχεται σε λίγα μόνο παραλληλόγραμμα και είναι αντιληπτή από τη συμβολική ένδειξη στο πάνω αριστερό μέρος τους (κυκλωμένο ερωτηματικό). Σε όσα εξ αυτών παρέχεται, περιλαμβάνει μία σύντομη επεξήγηση και σύνδεσμο που παραπέμπει σε αναλυτική τεκμηρίωση.

Το *Fitbit Dashboard* διαμορφώνεται από τη σύνθεση εικόνων που μοιάζουν με ενδείξεις σε όργανα μέτρησης, συνήθων μεθόδων οπτικοποίησης και αφαιρετικών παραλλαγών τους. Η εμφάνιση των συστατικών αυτών μεταβάλλεται δυναμικά και σε πραγματικό χρόνο, όταν ανανεώνονται τα αντίστοιχα δεδομένα. Επισημαίνεται η σχεδόν απόλυτη απουσία τίτλων· ο χρήστης αντιλαμβάνεται το είδος δραστηριότητας που αντιστοιχεί σε κάθε παραλληλόγραμμα από τις οπτικές ενδείξεις, που συνίστανται στη χρήση εικονιδίων (π.χ. φλόγα για τις καύσεις, μπουκάλι νερό για την κατανάλωση νερού). Οι αριθμοί αναγράφονται με μεγάλη γραμματοσειρά ενώ το μέγεθος γραμματοσειράς που χρησιμοποιείται για λέξεις είναι σημαντικά μικρότερο, όταν αυτές δεν αφορούν σε ένδειξη επίτευξης του στόχου.

Από προεπιλογή, εμφανίζονται οι αναπαραστάσεις των δεδομένων που αφορούν στην τρέχουσα ημέρα. Χρησιμοποιώντας την επιλογή μετάβασης σε προηγούμενες ημέρες, ο χρήστης μπορεί να έχει μία γενική εικόνα της προόδου του, μέσω της παρατήρησης της αλλαγής εμφάνισης των παραλληλογράμμων. Επιλέγοντας ένα συγκεκριμένο παραλληλόγραμμα, αποκτά γρήγορη πρόσβαση σε ένα συγκριτικό γράφημα που παρουσιάζει την πρόοδό του εντός συγκεκριμένου χρονικού διαστήματος. Στην Εικόνα 49, φαίνεται η

οπτικοποιημένη έκθεση προόδου σχετικά με την καύση θερμίδων σε χρονικό διάστημα δύο εβδομάδων -εκτός από τη γενική εικόνα, ο χρήστης μπορεί να μετακινήσει τον κέρσορα πάνω από μία στήλη για να δει τις αντίστοιχες τιμές κατανάλωσης και καύσης θερμίδων.



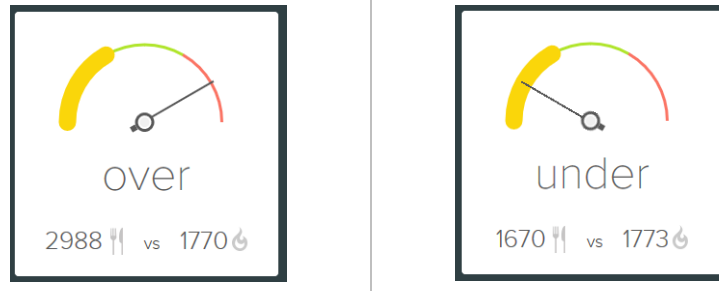
Εικόνα 49: Εξέλιξη προόδου
(Πηγή: Fitbit, 2014b)

Και σε αυτή την προβολή παρατηρείται το μεγάλο μέγεθος γραμματοσειράς που χρησιμοποιείται για τους αριθμούς και τη χρήση του έντονου πορτοκαλί στο γράφημα που παρουσιάζει την επίτευξη του στόχου. Μπορεί να υποθεθεί ότι αυτή η σχεδιαστική επιλογή απόδοσης έμφασης στην εμφάνιση των αριθμητικών τιμών, που είναι εμφανής καθ' όλη τη διάρκεια της πλοήγησης στο *Fitbit Dashboard*, συνδέεται με το γεγονός ότι πρόκειται για εφαρμογή *Ποσοτικοποιημένου Εαυτού* και ως τέτοια απευθύνεται σε άτομα που επιδιώκουν "αυτογνωσία μέσω αριθμών".

Γενικά, τα χρώματα που χρησιμοποιούνται είναι φωτεινά και κορεσμένα, και παραπέμπουν σε διάθεση χαράς και ενεργητικότητας. Εξαίρεση αποτελεί το σκούρο μπλε χρώμα που χρησιμοποιείται για την αναπαράσταση της διάρκειας του ύπνου και συμβολίζει τον νυχτερινό ουρανό. Ωστόσο παρατηρείται ότι η χρήση των χρωμάτων δεν είναι απαραίτητα συνδεδεμένη σημασιολογικά με τις τιμές που αναπαρίστανται.

Στην Εικόνα 50, φαίνονται διαφορετικές εκδοχές του μετρητή με τη μορφή κοντέρ, μέσω του οποίου ο χρήστης παρακολουθεί το ισοζύγιο πρόσληψης-καύσης θερμίδων σε σχέση με τον ημερήσιο στόχο που έχει θέσει. Η καμπύλη

γραμμή που περιγράφει ημικύκλιο, διαιρείται σε τρία τμήματα χρωματικά διαχωρισμένα: το τμήμα που αντιστοιχεί σε έλλειμα θερμίδων χρωματίζεται με κίτρινο⁷⁸, το τμήμα που αντιστοιχεί σε απόκλιση ± 50 θερμίδων από τον ημερήσιο στόχο χρωματίζεται με λαχανί, ενώ με κόκκινο-πορτοκαλί χρωματίζεται το τμήμα που αντιστοιχεί σε πλεόνασμα θερμίδων.



Εικόνα 50: Διαφορετικές εκδοχές μετρητή παρακολούθησης ισοζυγίου πρόσληψης-καύσης θερμίδων (Πηγή: Fitbit, 2014b)

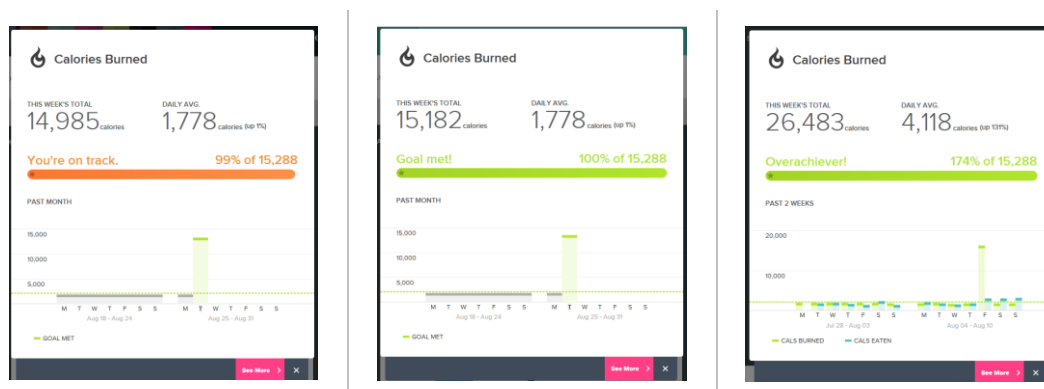
Παρατηρείται ότι οι σχεδιαστές επιλέγουν να διαιρέσουν χρωματικά το ημικύκλιο χρησιμοποιώντας διαφορετικές αποχρώσεις, αντιστοιχίζοντας τα θερμά χρώματα (κίτρινο και κόκκινο-πορτοκαλί) με τις μη αποδεκτές τιμές και το λαχανί με την επίτευξη του στόχου. Η εν λόγω αντιστοίχιση χαρακτηρίζεται από υψηλό βαθμό αυθαιρεσίας, δεδομένου ότι η χρήση τονικής διαβάθμισης και όχι χρωματικής απόχρωσης ενδείκνυται για την οπτικοποίηση συνεχών τιμών. Ακόμη και αν θεωρηθεί ότι ο στόχος αυτής της σχεδιαστικής επιλογής είναι η άμεση αντίληψη του χρήστη για απόκλιση από την επιθυμητή απόδοση σε μικρό ή μεγάλο βαθμό, η μη συνεπής εφαρμογή της σε άλλα σημεία του *Dashboard* δεν συνεπάγεται άμεση αντίληψη βάσει των οπτικών ενδείξεων και είναι απαραίτητη η ανάγνωση του συνοδευτικού κειμένου.

Για παράδειγμα, στην Εικόνα 49, το πορτοκαλί χρώμα χρησιμοποιείται για το γέμισμα της παχιάς γραμμής, το μήκος της οποίας υποδεικνύει το ποσοστό επίτευξης του εβδομαδιαίου στόχου καύσης θερμίδων, και βάσει των προαναφερθέντων έρχεται σε αντίφαση με το λεκτικό "*Είσαι σε καλό δρόμο*" ("*You're on track.*"). Για τον έλεγχο της υπόθεσης ότι οι ενδείξεις αλλάζουν όταν

⁷⁸ Παράλληλα η γραμμή έχει αυξημένο πάχος.

επιτυγχάνεται ο στόχος καύσης θερμίδων, καταχωρίζονται χειροκίνητα τιμές που οδηγούν σε επίτευξη άνω του 99% του εβδομαδιαίου στόχου.

Στην Εικόνα 51, εμφανίζονται τα αποτελέσματα αυτών των δοκιμών. Παρατηρείται ότι, ακόμη και για ποσοστό επίτευξης 99%, η ένδειξη και η χρωματική κωδικοποίηση δεν μεταβάλλονται. Επομένως συμπεραίνεται ότι το λαχανί χρώμα χρησιμοποιείται μόνο όταν η επίδοση του χρήστη ισοφαρίζει ή ξεπερνά το 100% του στόχου.



Εικόνα 51: Διαφορετικές εκδοχές επίτευξης στόχου καύσης θερμίδων
(Πηγή: Fitbit, 2014b)

Η χρήση του χρώματος εκτός των εμπεδωμένων συμβάσεων σε συνδυασμό με τη μη συνεπή εφαρμογή των σχετικών σχεδιαστικών επιλογών που θα συνέβαλλε στην εκπαίδευση των χρηστών, εκτιμάται ότι θέτει εμπόδια στην κατανόηση των οπτικών αναπαραστάσεων των συλλεχθέντων δεδομένων. Ήδη έχει καταγραφεί έντονη δυσaréσκεια από μερίδα χρηστών, όπως φαίνεται στα σχόλια που συνοδεύουν την επίσημη ανακοίνωση ανασχεδιασμού του *Fitbit Dashboard* στο ιστολόγιο της εταιρείας (Chris, 2014).

Η δομή του *Fitbit Dashboard* παραπέμπει σε κατατμημένη αφίσσα, τα τμήματα της οποίας (παραλληλόγραμμα) επιλέγονται και αναδιατάσσονται από τον χρήστη σύμφωνα με τις προτιμήσεις του. Με τον τρόπο αυτό, οι δημιουργοί παρέχουν στους χρήστες την ελευθερία διαμόρφωσης της προβολής, αφού όμως έχουν ορίσει σε προγενέστερο στάδιο τον τύπο οπτικοποίησης που περιλαμβάνεται σε κάθε παραλληλόγραμμα και κατά την κρίση τους αναπαριστά αποτελεσματικά τα υποκείμενα δεδομένα. Οι σχεδιαστές βασίζονται πολύ στην

παροχή δυνατότητας εμφάνισης λεπτομερειών με την κίνηση του κέρσορα πάνω από τα παραλληλόγραμμα αλλά και πάνω από επιμέρους στοιχεία (όπως π.χ. μία ράβδος σε γράφημα ράβδων) προκειμένου να περιορίσουν στο ελάχιστο τα συνοδευτικά λεκτικά. Η συνολική εικόνα που αποπνέει το *Fitbit Dashboard* είναι απαλλαγμένη από υπερβολική πληροφόρηση, εντούτοις λόγω της προβληματικής χρήσης του χρώματος ο χρήστης μπορεί να δυσκολευτεί στην κατανόηση των αναπαριστώμενων δεδομένων.

Κλείνοντας, σημειώνεται ότι το *Fitbit Dashboard* δεν αποτελεί τον μοναδικό τρόπο πρόσβασης του χρήστη στα δεδομένα που παράγει: αυτός μπορεί να κατεβάσει στον υπολογιστή του τα δεδομένα που έχουν καταγραφεί μέσω των συσκευών ή έχουν καταχωριστεί χειροκίνητα, σε αρχείο της μορφής .csv ή .xls για περαιτέρω επεξεργασία και ανάλυση.

5.2 Συζήτηση

Στην παρούσα ενότητα καταγράφονται τα γενικά ευρήματα από την εξέταση των περιπτώσεων που παρουσιάστηκαν.

Στις περιπτώσεις οπτικοποίησης δεδομένων που προέρχονται από τιτιβίσματα, ο δημιουργός της οπτικοποίησης έπρεπε να θέσει κριτήρια επιλογής των τιτιβισμάτων, βάσει της ετικέτας που αυτά περιείχαν. Τα τιτιβίσματα είναι αναρτήσεις που περιλαμβάνουν κείμενο ή/και φωτογραφίες, συνιστούν επομένως μη δομημένα δεδομένα. Ωστόσο, θέτοντας ως κριτήριο ένταξης την ύπαρξη συγκεκριμένης ετικέτας στην ανάρτηση, η οποία αξιοποιείται ως μετα-δεδομένο, κατέστη ευχερής ο χειρισμός τους. Κατά τη διαμόρφωση του συνόλου δεδομένων από το οποίο προέκυψε η οπτικοποίηση "*How riot rumours spread on Twitter*", οι ετικέτες αξιοποιήθηκαν για τον αρχικό εντοπισμό των τιτιβισμάτων που σχετίζονταν με τις αναταραχές. Για την κατασκευή της οπτικοποίησης που δημιουργήθηκε μέσω *Twitter Reverb*, η ετικέτα *#ferguson* ορίστηκε ως κλειδί για τον εντοπισμό των τιτιβισμάτων που κατά την εκτίμησή της δημιουργού αφορούσαν στο γεγονός που βρισκόταν σε εξέλιξη.

Και στις δύο περιπτώσεις, η επίγνωση του γενικού πλαισίου και η γνώση των συμβάσεων λειτουργίας του εν λόγω κοινωνικού μέσου, επέτρεψαν στον δημιουργό να επιλέξει το κατάλληλο κλειδί (εν προκειμένω η ετικέτα) για τον προσδιορισμό του υπό οπτικοποίηση συνόλου δεδομένων. Εκφράζονται ωστόσο επιφυλάξεις κατά πόσον αυτή η πρακτική είναι ασφαλής, στην περίπτωση που πρέπει να οπτικοποιηθούν γεγονότα που λαμβάνουν χώρα σε πραγματικό χρόνο, καθώς ο δημιουργός της οπτικοποίησης θα πρέπει να είναι σε εγρήγορση ώστε να τροποποιήσει την/ις ετικέτα/ες που παρακολουθεί, εφόσον εμφανιστούν νέες στο πλαίσιο του ίδιου διαλόγου.

Η μελέτη των οπτικοποιήσεων σε ύστερο χρόνο κατέστησε εμφανείς τις δυσκολίες στη διαδικασία διαμόρφωσης ενός διαχειρίσιμου συνόλου δεδομένων, όταν περιλαμβάνει κειμενικό περιεχόμενο και είναι επιθυμητή η απεικόνιση του σημασιολογικού φορτίου του. Εκτός από την εφαρμογή αλγορίθμων ή άλλων αυτοματοποιημένων ροών εργασίας, φαίνεται να είναι απαραίτητη η εμπλοκή του ανθρώπινου παράγοντα: στην περίπτωση του *Guardian*, μετά τη χρήση των αλγορίθμων για την ανάλυση του περιεχομένου των τιτιβισμάτων, μικρή ομάδα ατόμων ανέλαβε τον έλεγχο του περιεχομένου για να διασφαλίσει την αξιοπιστία του συνόλου δεδομένων. Ο Felton, συνδυάζοντας τις αυτόματες καταγραφές, τα μετα-δεδομένα των ηλεκτρονικών επικοινωνιών του και την 400 σελίδων χειρόγραφη καταγραφή των συνομιλιών του, κατέληξε σε ένα σύνολο δεδομένων, η επεξεργασία του οποίου ήταν μία επαναληπτική και χρονοβόρα διαδικασία.

Αντίθετα, από τη μελέτη των οπτικοποιήσεων σε πραγματικό ή σχεδόν πραγματικό χρόνο, φαίνεται ότι όταν τα δεδομένα που συλλέγονται είναι δομημένα ή διαχειρίσιμα μέσω αξιοποίησης των μετα-δεδομένων για την κατηγοριοποίησή τους, είναι εφικτή η διαμόρφωση ενός άμεσα οπτικοποιήσιμου συνόλου δεδομένων. Η εμφάνιση του *Fitbit Dashboard* ανανεώνεται αμέσως μόλις ενημερωθεί η αντίστοιχη βάση δεδομένων με τις μετρήσεις της συσκευής. Παρόμοια, φαίνεται ότι ο προσδιορισμός της ετικέτας που σχετίζεται με ένα θέμα επικαιρότητας αρκεί για να παραχθεί μέσω *Twitter Reverb* η οπτικοποίηση των

τιτιβισμάτων που την περιέχουν και αναρτήθηκαν στο άμεσο παρελθόν. Ωστόσο, δεδομένου ότι και στις δύο περιπτώσεις οι οντότητες που ανέπτυξαν τις υποδομές οπτικοποίησης είναι αυτές που διαχειρίζονται τον φορέα παραγωγής των συλλεγόμενων δεδομένων, κρίνεται ότι διαθέτουν την ευχέρεια και την τεχνογνωσία να υλοποιήσουν εφαρμογές που ανταποκρίνονται στους συγκεκριμένους τύπους δεδομένων.

Επιπρόσθετα, το μέσο στο οποίο θα προβληθεί η οπτικοποίηση φαίνεται πως έχει ιδιαίτερη σημασία σε ό,τι αφορά την κατασκευή οπτικοποιήσεων που βασίζονται σε δεδομένα που παράγονται και ανανεώνονται σε πραγματικό ή σχεδόν πραγματικό χρόνο. Οι προδιαγραφές υλοποίησης του εργαλείου *Twitter Reverb* ανέφεραν ρητά ότι οι οπτικοποιήσεις που παράγονται μέσω αυτού θα πρέπει να είναι λειτουργικές και σε κινητές συσκευές, δεδομένου ότι η πλειοψηφία των χρηστών τους αναμένεται να τις χρησιμοποιεί εν κινήσει. Η απαλλαγμένη από κείμενο τελική οθόνη και η λιτή σχεδιάσή της καθιστά αναγνώσιμη και εύχρηστη την οπτικοποίηση, ακόμη και όταν προβάλλεται στη μικρή οθόνη μιας κινητής συσκευής. Η περίπτωση του *Fitbit Dashboard* διαφέρει: η αντίστοιχη εφαρμογή για έξυπνα κινητά τηλέφωνα έχει πολύ διαφορετική εμφάνιση από την εκδοχή για πρόσβαση μέσω ηλεκτρονικού υπολογιστή. Η αισθητική παραπέμπει σε εφαρμογές κινητών τηλεφώνων, ενώ εντοπίζονται διαφοροποιήσεις στον τρόπο οπτικοποίησης κάποιων δεδομένων σε σύγκριση με τη διαδικτυακή εκδοχή.

Σε ό,τι αφορά τον οπτικό σχεδιασμό, οι οπτικοποιήσεις δεδομένων που παράγονται ή/και ανανεώνονται σε πραγματικό ή σχεδόν πραγματικό χρόνο έχουν μία σχετικά μονότονη εμφάνιση, καθώς ανεξαρτήτως περιεχομένου, η δομή δεν μεταβάλλεται. Το *Fitbit Dashboard* έχει μία προκαθορισμένη εμφάνιση για όλους τους χρήστες, οι οποίοι έχουν περιορισμένες δυνατότητες αναδιάταξης της προβολής (απόκρυψη ή μετακίνηση παραλληλογράμμων). Εντούτοις, χάρη στη δυνατότητα εξαγωγής αρχείου με τα υποκείμενα δεδομένα, ο χρήστης μπορεί να επεξεργαστεί, να εμπλουτίσει και να οπτικοποιήσει με τον δικό του τρόπο το σύνολο δεδομένων.

Η οπτικοποίηση *#ferguson* δεν διαφέρει από άλλες που δημιουργήθηκαν μέσω του *Twitter Reverb*: όλες έχουν στο κέντρο ένα διάγραμμα γραμμής που αναπαριστά το πλήθος τιτιβισμάτων ανά λεπτό για μια δεδομένη ετικέτα σε μία δεδομένη χρονική περίοδο και μία σειρά κύκλων στο κάτω μέρος της οθόνης που αντιστοιχεί σε σημαντικά -κατά την κρίση του δημιουργού- τιτιβίσματα. Επισημαίνεται ότι η απουσία υπομνημάτων και βοηθητικού υλικού δυσχεραίνει την κατανόηση της οπτικοποίησης, καθώς μόνο το διάγραμμα γραμμής συνιστά συνήθη τύπο οπτικής αναπαράστασης. Ωστόσο, εκτιμάται ότι η ομοιομορφία μεταξύ των παραγόμενων μέσω *Twitter Reverb* οπτικοποιήσεων και η πιθανή μαζική υιοθέτηση του εργαλείου για την παραγωγή οπτικοποιήσεων από δημοσιογραφικούς οργανισμούς, μπορούν να συντελέσουν βοηθητικά στην εκπαίδευση των χρηστών στην αποκωδικοποίηση της παρουσίασης των υποκείμενων δεδομένων.

Οι οπτικοποιήσεις που βασίζονται σε σύνολα δεδομένων που έχουν αντληθεί σε προγενέστερο χρόνο, έχουν πιο ιδιαίτερη μορφή, που προσαρμόζεται στον σκοπό και στα σημεία που επιθυμεί να τονίσει ο δημιουργός. Παράλληλα, ο χρόνος είναι προς όφελος των δημιουργών, διότι μπορούν να επεξεργαστούν διάφορες εναλλακτικές προτού καταλήξουν στο τελικό αποτέλεσμα. Η οπτικοποίηση του *Guardian* αποτελεί έναν συνδυασμό επιμέρους οπτικοποιήσεων, που συνοδεύονται από επαρκή υπομνηματισμό και σύντομο κείμενο που δίνει στον χρήστη συμπληρωματικές πληροφορίες για την ερμηνεία και την κατανόηση της εξέλιξης της συζήτησης που παρακολουθεί. Οι οπτικοποιήσεις του *Felton* είναι αισθητικά άρτιες αλλά δυσπρόσιτες για τον μη εξοικειωμένο χρήστη· το υψηλό επίπεδο αφαίρεσης δυσχεραίνει την κατανόηση.

Στις περισσότερες από τις επιλεγμένες μελέτες περίπτωσης, η χρήση του χρώματος συνδυάζεται με άλλα οπτικά στοιχεία (π.χ. μέγεθος ή θέση) για την οπτική κωδικοποίηση της πληροφορίας που φέρουν τα αναπαριστώμενα δεδομένα. Ωστόσο, στην περίπτωση του *Guardian*, ο διαχωρισμός της εναντίωσης από την υποστήριξη μίας φήμης που εκφράζεται ανά τιτίβισμα προκύπτει από τον διαφορετικό χρωματισμό των κύκλων που συμβολίζουν κάθε

τιτίβισμα. Τα χρώματα που χρησιμοποιούνται (κόκκινο και πράσινο, αντιστοίχως) για να μεταδώσουν τη συγκεκριμένη πληροφορία ενδέχεται να καθιστούν αυτό το τμήμα της οπτικοποίησης μη προσβάσιμο από άτομα που πάσχουν από αχρωματοψία. Εκτιμάται ότι ο σχεδιαστής στάθμισε το κόστος και το όφελος από την χρήση μίας πιο ασφαλούς παλέτας που θα αποδυνάμωνε το αποτέλεσμα για το μεγαλύτερο τμήμα του κοινού στο οποίο απευθύνεται. Η επιλογή αυτών των σημασιολογικά φορτισμένων χρωμάτων, παρά τις αδυναμίες που επισημάνθηκαν, φαίνεται ότι προτιμήθηκε για την ενίσχυση του μηνύματος της οπτικοποίησης.

Σε ό,τι αφορά τις δυνατότητες αλληλεπίδρασης που παρέχονται στους χρήστες, παρατηρείται συχνή χρήση της παροχής επεξηγήσεων με την κίνηση του δείκτη πάνω από μία περιοχή ή με την επιλογή ενός αντικειμένου της οπτικοποίησης, όπως έχει περιγραφεί πιο πάνω κατά περίπτωση. Η χρήση αυτής της δυνατότητας για τον εμπλουτισμό της οπτικοποίησης μέσω *Twitter Reverb* εμπλουτίζει το ερμηνευτικό πλαίσιο, καθώς συνιστά σχολιασμό του δημιουργού σχεδόν κατά τη διάρκεια εξέλιξης του γεγονότος και διευκολύνει τον χρήστη στην αφομοίωση της παρουσιαζόμενης πληροφορίας και στην εξαγωγή συμπερασμάτων. Επισημαίνεται ότι μόλις σε μία περίπτωση (*Fitbit Dashboard*) παρέχεται η δυνατότητα εξαγωγής αρχείου με τα δεδομένα που οπτικοποιούνται για περαιτέρω επεξεργασία ή ανάλυση από τον χρήστη.

Κλείνοντας, εντοπίζονται περιπτώσεις όπου το αποτέλεσμα της οπτικοποίησης επηρεάζεται σημαντικά από τον δημιουργό. Ο Felton φαίνεται ότι αντιμετωπίζει με καλλιτεχνική διάθεση την οπτικοποίηση, αφού παράγει μινιμαλιστικές συνθέσεις απαλλαγμένες από υπομνήματα. Το *Twitter Reverb*, το οποίο συνιστά εργαλείο μέσω του οποίου παράγονται τυποποιημένες αλλά μερικώς προσαρμόσιμες οπτικοποιήσεις, μάλλον εκφράζει το όραμα του Rogers για την υποστήριξη των δημοσιογράφων που επιθυμούν να αφηγηθούν ιστορίες με δεδομένα.

6 Τελικά συμπεράσματα

Στο πλαίσιο της παρούσας εργασίας, εξετάστηκαν ζητήματα σχετικά με το πολυδιάστατο φαινόμενο "*Μεγάλα Δεδομένα*", το οποίο διαμορφώνεται εν πολλοίς από τις τεχνολογικές δυνατότητες που επιτρέπουν τη συλλογή, την αποθήκευση, την επεξεργασία και την ανάλυση δεδομένων που χαρακτηρίζονται από πρωτοφανή όγκο, υψηλή ταχύτητα παραγωγής και ανανέωσης ή/και ποικιλομορφία.

Αν και ετυμολογικά ο όρος "*Μεγάλα Δεδομένα*" αυθόρμητα παραπέμπει στον όγκο, δεν θα πρέπει να παραβλέπονται οι υπόλοιπες ιδιότητες (ποικιλομορφία και ταχύτητα ανανέωσης) που διαφοροποιούν τα συλλεγόμενα δεδομένα από το παρελθόν. Ωστόσο, ο όρος αυτός δεν θα πρέπει να χρησιμοποιείται μονοδιάστατα για να περιγράψει τα εν λόγω χαρακτηριστικά των συλλεγόμενων δεδομένων (3Vs), καθώς αυτά δεν αρκούν από μόνα τους για να προκύψει η πολυθρύλητη αξία.

Για τη συγκέντρωση και την αποτελεσματική διαχείριση των ΜΔ είναι απαραίτητος ο συνδυασμός επιμέρους τεχνολογικών λύσεων. Η δυνατότητα πρόσβασης σε κατανεμημένους υπολογιστικούς πόρους αίρει τους οικονομικούς και τεχνολογικούς περιορισμούς του παρελθόντος, καθώς επιτρέπει τη συλλογή, την αποθήκευση και την επεξεργασία δεδομένων με σημαντικά χαμηλότερο κόστος για τους οργανισμούς. Παράλληλα, οι ανάγκες για αυξημένη επεξεργαστική ισχύ και αποθηκευτικά μέσα διαρκώς ενισχύονται από τη μαζική χρήση προϊόντων και υπηρεσιών που αξιοποιούν τις δυνατότητες της νεφοϋπολογιστικής και από την ενσωμάτωσή τους στην καθημερινότητα των χρηστών.

Τα σύνολα δεδομένων που προκύπτουν από σχετικά νέες πηγές όπως τα κοινωνικά μέσα, οι κινητές συσκευές, το Διαδίκτυο των Πραγμάτων και οι συσκευές *Ποσοτικοποιημένου Εαυτού*, συνιστούν κυρίως μη δομημένα δεδομένα που προκύπτουν από τη δραστηριότητα εκατοντάδων χιλιάδων

χρηστών και ανανεώνονται με μεγάλη συχνότητα. Ως αποτέλεσμα, τελικά προκαλείται η συγκέντρωση ενός τεράστιου όγκου δεδομένων, ο οποίος πληροί τα κριτήρια για να χαρακτηριστεί ως ΜΔ.

Τα δεδομένα που συλλέγονται από τα κοινωνικά μέσα έχουν αυξηθεί και περιλαμβάνουν νέες μορφές, όπως οι αξιολογήσεις και τα μετα-δεδομένα, που παράγονται από τη δραστηριότητα των χρηστών εν αγνοία τους ή/και δεν απαιτούν την καταβολή δημιουργικής προσπάθειας από πλευράς τους. Τα δεδομένα αυτά αποτελούν πλούσιες και πολύτιμες πηγές πληροφοριών και η επεξεργασία τους είναι δυνατό να οδηγήσει σε συμπεράσματα για τον χρήστη και να αξιοποιηθεί ποικιλοτρόπως. Εκτιμάται μάλιστα ότι η διαφαινόμενη τάση για τη διάθεση έξυπνων φορέσιμων συσκευών με δυνατότητες καταγραφής δεδομένων φυσικής δραστηριότητας και διαμοιρασμού τους μέσω κοινωνικών μέσων, θα αυξήσει τις απαιτήσεις για την καταγραφή συνεχούς ροής ατομικών δεδομένων εν κινήσει. Συνακόλουθα, εκτιμάται ότι θα αυξηθεί η ζήτηση για εφαρμογές *Ποσοτικοποιημένου Εαυτού*.

Η πρόοδος στις τεχνολογίες που υποστηρίζουν τη συλλογή συνεχώς αυξανόμενου όγκου δεδομένων, σε συνδυασμό με τη δραστική μείωση του κόστους αποθήκευσης σε βάσεις δεδομένων και τις διαθέσιμες μεθόδους εξαγωγής γνώσης από βάσεις δεδομένων, εμπνέουν διάχυτη αισιοδοξία για την αξία που μπορεί να αντληθεί από τα ΜΔ. Ήδη έχει αναπτυχθεί ένα οικοσύστημα εταιρειών με αντικείμενο τα ΜΔ, ενώ έχουν εμφανιστεί νέα πεδία επαγγελματικής εξειδίκευσης που απαιτούν δεξιότητες ανάλυσης ΜΔ και παρουσίασης της πληροφορίας που εμπεριέχεται σε αυτά: η επιστήμη δεδομένων και η δημοσιογραφία δεδομένων. Σημειώνεται ότι και στις δύο αυτές εξειδικεύσεις, η τοποθέτηση των δεδομένων εντός πλαισίου και η ικανότητα οπτικής ή/και λεκτικής αφήγησης είναι απαραίτητη για την ανάλυση και την παρουσίαση ΜΔ.

Η τάση στα σύγχρονα συστήματα επιχειρηματικής ευφυΐας που διευκολύνουν την ανακάλυψη γνώσης από ΜΔ είναι η παροχή δυνατότητας

οπτικοποίησης ως βασική λειτουργικότητα για τη διεξαγωγή αναλύσεων σε πραγματικό χρόνο και για την παρουσίασή τους με αλληλεπιδραστικό και κατανοητό τρόπο. Η προτίμηση στην παρουσίαση πληροφορίας με οπτικό τρόπο οφείλεται στο ότι μέσω της όρασης λαμβάνονται περισσότερες πληροφορίες απ' ό,τι μέσω όλων των υπόλοιπων αισθήσεων, ενώ αυτές οι πληροφορίες αναλύονται από τον ανθρώπινο εγκέφαλο, ο οποίος διαθέτει έναν πανίσχυρο μηχανισμό εντοπισμού μοτίβων.

Επομένως, η αναγνώριση της οπτικοποίησης ως μέσου ανάλυσης και άντλησης αξίας από τα ΜΔ συνδέεται με τον σημαντικό ρόλο που διαδραματίζει στη διαδικασία μετατροπής των ανεπεξέργαστων δεδομένων σε χρήσιμη πληροφορία. Η οπτική αναπαράσταση δεδομένων αξιοποιεί τις δυνατότητες του ανθρώπινου οπτικού συστήματος για να διευκολύνει τον άμεσο εντοπισμό σχέσεων και μοτίβων σε αφηρημένα δεδομένα. Ως αποτέλεσμα, είναι δυνατή η γρήγορη ερμηνεία των δεδομένων που αναπαρίστανται και η άντληση της γνώσης από αυτά.

Τα τελευταία χρόνια παρατηρείται έντονη ερευνητική δραστηριότητα στο εν λόγω αντικείμενο. Η ποικιλία επιστημονικών υποβάθρων των ερευνητών που ασχολούνται με αυτό εξηγεί τις διαφορετικές αντιμετωπίσεις και ορισμούς. Ωστόσο, όπως προκύπτει από την ιστορική αναδρομή για την ένταξή της σε ιστορικό πλαίσιο, η οπτικοποίηση δεν συνιστά πρωτοεμφανιζόμενο κλάδο αλλά αποτελεί την πλέον πρόσφατη φάση εξέλιξης μίας μακροχρόνιας παράδοσης στην κατασκευή οπτικών αναπαραστάσεων δεδομένων. Διαχρονικά, φαίνεται ότι η οπτικοποίηση εξελίχθηκε παράλληλα με την αύξηση δεδομένων και συνακόλουθα με την ανάγκη επεξεργασίας και ερμηνείας τους για την επίλυση προβλημάτων και την αιτιολόγηση φαινομένων.

Η σύγχρονη περίοδος της οπτικοποίησης έχει καθοριστεί από τις εξελίξεις στους ηλεκτρονικούς υπολογιστές και τα γραφικά υπολογιστών, που έχουν διευκολύνει τη διαδικασία μετατροπής των άμορφων, αφηρημένων δεδομένων σε οπτικά αντιληπτή μορφή. Κατά την εξέταση ορισμών που έχουν διατυπωθεί

υπό το πρίσμα των πρόσφατων τεχνολογικών εξελίξεων, εντοπίστηκε ότι μέχρι πρότινος, υπήρχε σαφής διαχωρισμός μεταξύ επιστημονικής οπτικοποίησης και οπτικοποίησης πληροφορίας. Ο διαχωρισμός αυτός έχει ιστορικές καταβολές και αντανακλά τα διαφορετικά υπόβαθρα αυτών που τους διατυπώνουν -βασίζεται δε στην προέλευση των δεδομένων ή/και στον σκοπό αναπαράστασης αυτών.

Ωστόσο, πλέον τείνει να εκλείψει, καθώς η συγκέντρωση ΜΔ και οι τεχνικές δυνατότητες επεξεργασίας δομημένων και μη δεδομένων, συγχέουν τα όρια. Στο πλαίσιο της παρούσας εργασίας, ο όρος *"οπτικοποίηση δεδομένων"* περιγράφει τις μεθόδους και το τελικό αποτέλεσμα της αναπαράστασης δεδομένων, δομημένων ή μη, σε οπτικά αντιληπτή μορφή, που επιτρέπουν την ταχεία διαμόρφωση άποψης, και την εξαγωγή νέας ή την ενίσχυση υφιστάμενης γνώσης από τα αναπαριστώμενα δεδομένα.

Έχουν διατυπωθεί διάφορες απόψεις σχετικά με τη φύση της οπτικοποίησης, αν δηλαδή αποτελεί επιστήμη, τέχνη ή κάτι άλλο. Στην παρούσα φάση, φαίνεται ότι πρόκειται μάλλον για μέθοδο, παρά για επιστήμη, που αξιοποιεί τις ιδιότητες της ανθρώπινες αντίληψης. Ο χαρακτηρισμός *"γνωσιακή τέχνη"* που της έχει αποδώσει ο Tufte (1990) συνδυάζει τις φαινομενικά ασύμβατες προσεγγίσεις και αναδεικνύει την πολύπλευρη υπόστασή της.

Γενικά, η οπτικοποίηση δεδομένων χρησιμοποιείται είτε για σκοπούς ανάλυσης, όταν ο χρήστης επιθυμεί να μελετήσει τα δεδομένα, είτε για σκοπούς παρουσίασης προκειμένου να τα επικοινωνήσει σε άλλους. Ωστόσο, οι σημερινές τεχνολογικές δυνατότητες καθιστούν εφικτή τη δημιουργία οπτικοποιήσεων που επιτρέπουν την εκτεταμένη αλληλεπίδραση του χρήστη με τα υποκείμενα δεδομένα, σε πρωτογενή ή επεξεργασμένη μορφή. Κατ' επέκταση, διακρίνεται μία τρίτη, υβριδική χρήση της οπτικοποίησης, που προκύπτει από τον συνδυασμό των δύο παραπάνω και συνίσταται στον εμπλουτισμό της οπτικής παρουσίασης του αποτελέσματος ανάλυσης με δυνατότητες αλληλεπίδρασης.

Με αυτόν τον τρόπο, ο χρήστης, εφόσον είναι διαφορετικό άτομο από τον αναλυτή, εξερευνά και αναλύει περαιτέρω τα υποκείμενα δεδομένα, με δυνητικό

αποτέλεσμα την ανακάλυψη γνώσης. Εστιάζοντας σε αυτή την υβριδική χρήση, εξετάστηκαν οι σχεδιαστικές αποφάσεις που λαμβάνονται στο πλαίσιο δημιουργίας σύγχρονων αλληλεπιδραστικών οπτικοποιήσεων από ανθρώπους για ανθρώπους, με σκοπό την ανακάλυψη γνώσης που δεν διέθεταν πριν εκτεθούν στην οπτικοποίηση.

Η σχεδιαστική διαδικασία μετατροπής των άμορφων, αφηρημένων δεδομένων σε αισθητηριακά αντιληπτή μορφή περιλαμβάνει διάφορα στάδια, από την απόκτηση εξειδικευμένων γνώσεων σχετικά με το προς οπτικοποίηση σύνολο δεδομένων μέχρι και τη διατύπωση του σκοπού της οπτικοποίησης. Τα στάδια αυτά προηγούνται και αποτελούν τη βάση για την επιλογή εκείνων των οπτικών κωδικοποιήσεων που μπορούν να εξυπηρετήσουν τον σκοπό της οπτικοποίησης.

Βάσει των προαναφερθέντων, γίνεται κατανοητό ότι η διατύπωση των ερωτημάτων και η επιλογή των συνόλων δεδομένων που μπορούν να τα απαντήσουν, προϋποθέτουν γνώση τόσο του γενικότερου περιβάλλοντος από το οποίο έχουν συλλεχθεί τα δεδομένα, όσο και των ειδικότερων λεπτομερειών για τις ιδιότητές τους. Πρόκειται δηλαδή για την αναγκαιότητα τοποθέτησης των δεδομένων εντός πλαισίου, η οποία έχει εντοπιστεί ως κρίσιμο ζήτημα για την άντληση αξίας κατά την εξέταση ενός συνόλου ΜΔ. Χωρίς πλαίσιο, ο σχεδιαστής δεν μπορεί να αναλύσει τα δεδομένα και να καταλήξει σε συμπέρασμα από αυτά.

Ανάλογα με τον τύπο των υπό εξέταση μεταβλητών, επιλέγεται η κατάλληλη μέθοδος ανάλυσης, από την οποία προκύπτουν απαντήσεις στα ερωτήματα που έχουν τεθεί, καθώς και ευρήματα που κρίνονται άξια παρουσίασης στο κοινό πέραν αυτού. Αντίστοιχα, η επιλογή του τρόπου οπτικοποίησης ενός συνόλου δεδομένων εξαρτάται από τον τύπο και τις ιδιότητες των μεταβλητών προς αναπαράσταση.

Μέσω της οπτικοποίησης, ο σχεδιαστής προσδίδει μορφή στα δεδομένα, αντιστοιχίζοντάς τα με οπτικά αντιληπτές μορφές. Από αυτές τις αντιστοιχίσεις, ο χρήστης μπορεί να αντλήσει πληροφορίες μέσω των νοητικών διεργασιών,

προκειμένου να επιβεβαιώσει υπάρχουσα ή και για να αποκτήσει νέα γνώση σχετικά με τα αναπαριστώμενα δεδομένα. Μέσω της ανάλυσης μίας οπτικής αναπαράστασης δεδομένων σε συστατικά στοιχεία, ανεξαρτήτως της ονομασίας που τους αποδίδεται, επιβεβαιώθηκε ότι διαχρονικά αυτά παραπέμπουν στα βασικά στοιχεία οπτικής επικοινωνίας: κάθε οπτικό μήνυμα προκύπτει από τον συνδυασμό αυτών των στοιχείων που είναι οικεία και αναγνωρίσιμα στον χρήστη μέσω των ανθρώπινων αντιληπτικών ικανοτήτων.

Ανάλογα με τον σκοπό και το διαθέσιμο σύνολο δεδομένων, ο σχεδιαστής επιλέγει την κατάλληλη μέθοδο οπτικοποίησης. Κάθε μέθοδος οπτικοποίησης αποτελεί συνδυασμό επιμέρους οπτικών κωδικοποιήσεων. Η χρήση τυποποιημένων μεθόδων οπτικοποίησης και συμβάσεων προϋποθέτει γνώση των αρχών που τις διέπουν, ώστε ο σχεδιαστής να εκμεταλλευθεί την εξοικείωση των χρηστών με τυποποιημένα και ευρέως κατανοητά πρότυπα οπτικής αναπαράστασης. Η επίγνωση του τι μπορεί και τι δεν μπορεί να αναπαρασταθεί μέσω αυτών των τυποποιημένων μεθόδων σε συνδυασμό με τη γνώση των οπτικών κωδικοποιήσεων σε δομικό επίπεδο, δίνουν στον σχεδιαστή την ευελιξία να δημιουργεί οπτικοποιήσεις που εξυπηρετούν καλύτερα τον σκοπό οπτικής αναπαράστασης σύνθετων συνόλων ΜΔ.

Το χρώμα αποτελεί μία από τις συχνότερα χρησιμοποιούμενες οπτικές κωδικοποιήσεις για την αναπαράσταση δεδομένων. Ο σχεδιαστής μπορεί να χρησιμοποιήσει το χρώμα για να κατηγοριοποιήσει ή να ονοματίσει, να εκφράσει ποσότητες, να αναπαραστήσει την πραγματικότητα ή να διακοσμήσει. Διαφορετικές ιδιότητες του χρώματος είναι κατάλληλες για την αναπαράσταση ποιοτικών και ποσοτικών μεταβλητών· για παράδειγμα, η φωτεινότητα του χρώματος είναι προτιμότερη για την κωδικοποίηση ποσοτικών μεταβλητών, διότι μπορεί να υποδηλώσει μία έννοια σειράς, αντίθετα από τη χρωματική απόχρωση, που ενδείκνυται για τις κατηγορικές μεταβλητές.

Επιπρόσθετα, το χρώμα ως φορέας μηνυμάτων, μπορεί να προκαλέσει συναισθηματικές αντιδράσεις ή/και να λειτουργήσει συμβολικά. Για τον λόγο

αυτόν, θα πρέπει να χρησιμοποιείται στο πλαίσιο συμβάσεων, εμπεδωμένων από τους χρήστες και να συμφωνεί με τις προσδοκίες τους. Υπάρχουν μάλιστα πειραματικές ενδείξεις ότι το περιεχόμενο μίας γραφικής αναπαράστασης μπορεί να κατανοηθεί ταχύτερα, όταν τα χρώματα που χρησιμοποιούνται έχουν επιλεγεί με σημειολογικά κριτήρια.

Εκτός όμως από τις αισθητικές και συμβολικές επιλογές, ο σχεδιαστής θα πρέπει να συνεκτιμά τις τυχόν ιδιαιτερότητες του κοινού στο οποίο απευθύνεται η οπτικοποίηση, ώστε να μην αποκλείονται χρήστες που δυσκολεύονται ή αδυνατούν να διακρίνουν διαφορές μεταξύ αποχρώσεων. Επομένως, ο σχεδιαστής θα πρέπει να φροντίζει να συνδυάζει το χρώμα με άλλα οπτικά στοιχεία, προκειμένου οι κρίσιμες πληροφορίες να είναι αντιληπτές και από χρήστες που αντιμετωπίζουν δυσκολίες στην αναγνώριση χρωμάτων. Γενικές οδηγίες για την καθολική σχεδίαση μίας οπτικοποίησης δεδομένων είναι η χρήση χρωμάτων με μεγάλη αντίθεση, η χρήση μαύρου χρώματος γραμματοσειράς σε ανοιχτόχρωμο φόντο και η τοποθέτηση των ετικετών κοντά στα στοιχεία του γραφήματος.

Δεδομένου ότι ο σχεδιαστής επιλέγει ποιες από τις διαστάσεις του υπό εξέταση συνόλου δεδομένων θα οπτικοποιήσει και με ποιον σκοπό, το αποτέλεσμα της οπτικοποίησης συνιστά μήνυμα του σχεδιαστή προς τον χρήστη αυτής. Επομένως, η οπτικοποίηση δεδομένων ως αναπαράσταση και παρουσίαση των υπό εξέταση δεδομένων εμπεριέχει επικοινωνιακή διάσταση. Επιπρόσθετα, δεδομένου ότι οι σύγχρονες οπτικοποιήσεις που σχεδιάζονται και υλοποιούνται κυρίως για χρήση μέσω ηλεκτρονικού υπολογιστή, συνιστούν διεπαφές μεταξύ των υποκείμενων δεδομένων και του χρήστη που τις χρησιμοποιεί για να εξερευνήσει και για να αποκτήσει πρόσβαση σε αυτά τα δεδομένα. Υπό το πρίσμα αυτό, υποστηρίχθηκε ότι, στο πλαίσιο κατασκευής οπτικοποίησης δεδομένων που απευθύνεται σε κοινό πέραν του εαυτού του, ο σχεδιαστής πρέπει να υιοθετεί μία χρηστοκεντρική προσέγγιση για να βελτιώσει την επικοινωνία ανθρώπου-υπολογιστή.

Εκτός από τον οπτικό σχεδιασμό που αναδεικνύει τις ιδιότητες των δεδομένων, ο σχεδιαστής μπορεί να εμπλουτίσει αυτή την κατασκευή, συμπεριλαμβάνοντας στοιχεία αλληλεπίδρασης ή/και χρησιμοποιώντας αφηγηματικές τεχνικές, προκειμένου (1) να ενισχύσει το πλαίσιο ερμηνείας και κατανόησης των αναπαριστώμενων δεδομένων και (2) να ενθαρρύνει τον διάλογο του χρήστη με τα υποκείμενα δεδομένα. Με τον τρόπο αυτό, παρέχει στον χρήστη ένα περιβάλλον ευνοϊκό για εξερεύνηση των δεδομένων και κατ' επέκταση για ανακάλυψη γνώσης.

Ο εμπλουτισμός των οπτικών αναπαραστάσεων δεδομένων με δυνατότητες αλληλεπίδρασης, τις καθιστά δυναμικές προβολές διαφορετικών πλευρών των αναπαριστώμενων δεδομένων. Η παροχή δυνατοτήτων αλληλεπίδρασης συνίσταται κυρίως στη σταδιακή πρόσβαση, σύμφωνα με τις ανάγκες του χρήστη, από τη γενική εικόνα στα δεδομένα που αυτή αναπαριστά. Τα γενικότερα οφέλη από την αξιοποίηση των δυνατοτήτων αλληλεπίδρασης είναι αμφίπλευρα. Ο σχεδιαστής έχει τη δυνατότητα να διαμορφώσει την οπτικοποίηση σύμφωνα με όσα κρίνει απαραίτητα βάσει του σκοπού της, χωρίς να επιβαρύνει τον χρήστη προβάλλοντας μονομιάς όλη τη δυνατή πληροφόρηση. Παράλληλα, ο χρήστης έχει ενεργή συμμετοχή στη διαμόρφωση της προβαλλόμενης οπτικοποίησης, μέσω της επιλογής προβολής εναλλακτικών οπτικών, σύμφωνα με τις επιθυμίες του.

Ο έλεγχος της έκτασης της αλληλεπίδρασης αποτελεί αφηγηματικό εργαλείο για τον σχεδιαστή. Η εφαρμογή αφηγηματικών τεχνικών παρέχει τη δυνατότητα παρουσίασης των πληροφοριών μαζί με το γενικότερο πλαίσιο. Η δε αφηγηματική δομή μίας οπτικοποίησης χαρακτηρίζεται από το ποιος καθοδηγεί την εξέλιξή της. Οι δυνατότητες ελέγχου της έκτασης της αλληλεπίδρασης εκτείνονται σε ένα συνεχές από την αμιγώς καθοριζόμενη από τον δημιουργό έως την αμιγώς καθοριζόμενη από τον χρήστη-αναγνώστη προσέγγιση. Ο σχεδιαστής επιλέγει τον βαθμό ελευθερίας που θα παρέχει στον χρήστη και το είδος αφήγησης που θα χρησιμοποιήσει, λαμβάνοντας υπ' όψιν τις ιδιαιτερότητες του κοινού στο οποίο απευθύνεται: για παράδειγμα τις γνώσεις

σχετικά με το αντικείμενο ή τον χρόνο που μπορεί να διαθέσει.

Για την εξέταση των προαναφερθέντων σε πρακτικό επίπεδο, εξετάστηκαν επιλεγμένες οπτικοποιήσεις δεδομένων σε πραγματικό ή σε ύστερο χρόνο, ως προς το τελικό αποτέλεσμα που προετοιμάζει ένας δημιουργός-άνθρωπος για να προσφέρει τη δυνατότητα ανακάλυψης γνώσης σε έναν χρήστη-άνθρωπο. Στόχος της μελέτης ήταν η εποπτική θεώρηση της διαδικασίας σχεδίασης μίας οπτικοποίησης δεδομένων που έχουν συλλεχθεί από πηγές προέλευσης ΜΔ, για τον εντοπισμό των βασικών σχεδιαστικών επιλογών και των επιπλοκών τους.

Ως προς το περιεχόμενο, επιλέχθηκαν οπτικοποιήσεις που βασίζονται σε δεδομένα που προέρχονται από (1) το κοινωνικό μέσο *Twitter* (κειμενικές αναρτήσεις) και (2) εφαρμογές *Ποσοτικοποιημένου Εαυτού* (αυτο-καταγραφή). Επιπρόσθετα, για τον εντοπισμό των προκλήσεων που θέτει η ταχύτητα παραγωγής και ανανέωσης ΜΔ, επιλέχθηκαν οπτικοποιήσεις ΜΔ που κατασκευάζονται (1) κατόπιν (σε ύστερο χρόνο) ή (2) κατά τη διάρκεια (σε πραγματικό χρόνο) συγκέντρωσης των δεδομένων.

Στις περιπτώσεις οπτικοποίησης των κειμενικών αναρτήσεων στο *Twitter* (μη δομημένα δεδομένα), ο δημιουργός χρησιμοποιεί μετα-δεδομένα για να καταλήξει σε ένα οπτικοποιήσιμο σύνολο δεδομένων. Η επίγνωση του γενικού πλαισίου και η γνώση των συμβάσεων λειτουργίας του εν λόγω κοινωνικού μέσου, επιτρέπουν στον δημιουργό να επιλέξει το κατάλληλο κλειδί (εν προκειμένω την ετικέτα) για τον προσδιορισμό του υπό οπτικοποίηση συνόλου δεδομένων.

Η μελέτη των οπτικοποιήσεων σε ύστερο χρόνο κατέστησε εμφανείς τις δυσκολίες στη διαδικασία διαμόρφωσης ενός διαχειρίσιμου συνόλου δεδομένων, όταν περιλαμβάνει κειμενικό περιεχόμενο και είναι επιθυμητή η απεικόνιση του σημασιολογικού φορτίου του: εκτός από την εφαρμογή αλγορίθμων ή άλλων αυτοματοποιημένων ροών εργασίας, φαίνεται να είναι απαραίτητη η εμπλοκή του ανθρώπινου παράγοντα. Αντίθετα, από τη μελέτη των οπτικοποιήσεων σε πραγματικό ή σχεδόν πραγματικό χρόνο, φαίνεται ότι

όταν τα δεδομένα που συλλέγονται είναι δομημένα ή διαχειρίσιμα μέσω αξιοποίησης των μετα-δεδομένων, είναι εφικτή η διαμόρφωση ενός άμεσα οπτικοποιήσιμου συνόλου δεδομένων.

Σε ό,τι αφορά τον οπτικό σχεδιασμό, οι οπτικοποιήσεις δεδομένων που παράγονται ή/και ανανεώνονται σε πραγματικό ή σχεδόν πραγματικό χρόνο χαρακτηρίζονται από ομοιομορφία, καθώς ανεξαρτήτως περιεχομένου, η δομή δεν μεταβάλλεται. Αντίθετα, οι οπτικοποιήσεις που βασίζονται σε σύνολα δεδομένων που έχουν αντληθεί σε προγενέστερο χρόνο, έχουν πιο ιδιαίτερη μορφή, που προσαρμόζεται στον σκοπό και στα σημεία που επιθυμεί να τονίσει ο δημιουργός. Επιπλέον, ο χρόνος είναι προς όφελος των δημιουργών, διότι μπορούν να επεξεργαστούν διάφορες εναλλακτικές προτού καταλήξουν στο τελικό αποτέλεσμα.

Στις περισσότερες από τις επιλεγμένες μελέτες περίπτωσης, το χρώμα συνδυάζεται με άλλα οπτικά στοιχεία (π.χ. μέγεθος ή θέση) για την οπτική κωδικοποίηση της πληροφορίας που φέρουν τα αναπαριστώμενα δεδομένα. Ωστόσο, σε μία από αυτές, το χρώμα χρησιμοποιείται για τη μετάδοση συγκεκριμένης πληροφορίας -οι αποχρώσεις που χρησιμοποιούνται (κόκκινο και πράσινο) καθιστούν το συγκεκριμένο τμήμα της οπτικοποίησης μη προσβάσιμο από άτομα με αχρωματοψία. Προφανώς η στάθμιση του κόστους και του οφέλους από την χρήση μίας πιο ασφαλούς παλέτας που αποδυνάμωνει το αποτέλεσμα για το μεγαλύτερο τμήμα του κοινού-στόχος, συνεκτιμάται κατά την επιλογή των χρωμάτων.

Σχετικά με τις δυνατότητες αλληλεπίδρασης που παρέχονται στους χρήστες, παρατηρείται συχνή χρήση της παροχής επεξηγήσεων με την κίνηση του δείκτη πάνω από μία περιοχή ή με την επιλογή ενός αντικειμένου της οπτικοποίησης. Η χρήση αυτής της δυνατότητας για τον εμπλουτισμό της οπτικοποίησης κειμενικών αναρτήσεων σε σχεδόν πραγματικό χρόνο εμπλουτίζει το ερμηνευτικό πλαίσιο, καθώς συνιστά σχολιασμό του δημιουργού σχεδόν κατά τη διάρκεια εξέλιξης του γεγονότος και διευκολύνει τον χρήστη στην αφομοίωση

της παρουσιαζόμενης πληροφορίας και στην εξαγωγή συμπερασμάτων.

Τέλος, το μέσο στο οποίο θα προβληθεί η οπτικοποίηση πρέπει να λαμβάνεται υπ' όψιν. Με δεδομένη την αύξηση των χρηστών κινητών συσκευών που επιλέγουν αυτές για πρόσβαση στο Διαδίκτυο, ο σχεδιαστής θα πρέπει να συνεκτιμά τους περιορισμούς που θέτουν οι μικρότερες οθόνες των εν λόγω συσκευών, στοχεύοντας στη διασφάλιση της αναγνωσιμότητας και της λειτουργικότητας της οπτικοποίησης.

Στο πλαίσιο της παρούσας εργασίας, εξετάστηκε περιορισμένο πλήθος περιπτώσεων οπτικοποιήσεων δεδομένων για την εξέταση των προαναφερθέντων σε πρακτικό επίπεδο και για τον εντοπισμό ερεθισμάτων για περαιτέρω διερεύνηση. Ιδιαίτερο ενδιαφέρον παρουσιάζουν οι οπτικοποιήσεις δεδομένων που ανανεώνονται σε πραγματικό χρόνο. Σημειώνεται ότι οι οντότητες που ανέπτυξαν τις εφαρμογές από τις οποίες προέρχονται οι οπτικοποιήσεις των σχετικών μελετών περίπτωσης είναι αυτές που διαχειρίζονται τον φορέα παραγωγής των συλλεγόμενων δεδομένων. Επομένως, θα είχε ενδιαφέρον η περαιτέρω εξέταση τυχόν οπτικοποιήσεων αντίστοιχων ειδών περιεχομένου από τρίτα μέρη, για τη σύγκριση του τελικού αποτελέσματος που παρέχεται στον χρήστη. Παράλληλα, εκτιμάται ότι η διάδοση των φορέσιμων συσκευών με δυνατότητες αυτο-καταγραφής θα συνοδευθεί από αυξημένη ζήτηση για ειδικές διεπαφές οπτικοποιήσεων, μέσω των οποίων ο χρήστης θα έχει πρόσβαση στα δεδομένα που συλλέγονται μέσω εφαρμογών *Ποσοτικοποιημένου Εαυτού*.

Βιβλιογραφικές παραπομπές

- Antephase, 2014. *Quantified Self*. Διαθέσιμο στο: <http://antephase.com/quantifiedself> [Πρόσβαση 17 Αυγούστου 2014].
- Arbesman, S., 2012. Big data: Mind the gaps. *The Boston Globe*, 30 Σεπτεμβρίου 2012. Διαθέσιμο στο: <http://www.bostonglobe.com/ideas/2012/09/29/big-data-mind-gaps/QClupxdwdPWhtRrZ002590/story.html> [Πρόσβαση 17 Αυγούστου 2014].
- Arbesman, S., 2013. Stop Hying Big Data and Start Paying Attention to ‘Long Data’. *Wired*, 29 Ιανουαρίου 2013. Διαθέσιμο στο: <http://www.wired.com/2013/01/forget-big-data-think-long-data/> [Πρόσβαση 17 Αυγούστου 2014].
- Αρμενάκης, Α., 2011. *Σημειώσεις Στατιστικής*. Εθνικό & Καποδιστριακό Πανεπιστήμιο Αθηνών, Τμήμα Επικοινωνίας και ΜΜΕ.
- Ball, J. and Lewis, P., 2011. Twitter and the riots: how the news spread. *The Guardian*, 07 Δεκεμβρίου 2011. Διαθέσιμο στο: <http://www.theguardian.com/uk/2011/dec/07/twitter-riots-how-news-spread> [Πρόσβαση 28 Αυγούστου 2014].
- Ben, 2012. This Bloom is off the Vine. *Bloom Blog*, 01 Νοεμβρίου 2012. Διαθέσιμο στο: <http://blog.bloom.io/2012/11/01/this-bloom-is-off-the-vine/> [Πρόσβαση 17 Αυγούστου 2014].
- Bertin, J., 2011. *Semiology of Graphics*. Αγγλική μετάφραση από τον W. J. Berg, της αναθεωρημένης γαλλικής έκδοσης του 1973. Redlands: Esri Press.
- Big Data Insight Group, 2012. Big Data Talk 007: The Data Scientist. *Big Data Insight Group*, 12 Ιουλίου 2012. Διαθέσιμο στο: <http://www.thebigdatainsightgroup.com/site/article/big-data-talk-007-data-scientist> [Πρόσβαση 17 Αυγούστου 2014].

- Blossom, J., 2009. *Content Nation: Surviving and Thriving as Social Media Changes Our Work, Our Lives and Our Future*. Indianapolis: Wiley.
- Bonneau, G.-P., Ertl, T. and Nielson, G., eds., 2006. *Scientific Visualization: The Visual Extraction of Knowledge from Data*. Berlin: Springer.
- Börner, K., and Polley, D. E., 2014. *Visual Insights: A Practical Guide to Making Sense of Data*. Cyberinfrastructure for Network Science Center - Massachusetts Institute of Technology.
- Boyd, D. and Crawford, K., 2012. Critical Questions for Big Data. *Information, Communication & Society*, 15 (5), 662-679. Διαθέσιμο στο: <http://www.tandfonline.com/doi/pdf/10.1080/1369118X.2012.678878> [Πρόσβαση 17 Αυγούστου 2014].
- Brech, B., Jamison, J., Shao, L. and Wightwick, G., 2013. *The Interconnecting of Everything*. IBM Academy of Technology. Διαθέσιμο στο: <http://www.redbooks.ibm.com/redpapers/pdfs/redp4975.pdf> [Πρόσβαση 17 Αυγούστου 2014].
- Burn-Murdoch, J., 2013. The advent of the statistician journalist. In: J. Mair, R. L. Keeble, P. Bradshaw and T. Beleaga, eds. *Data Journalism: Mapping the future*. Suffolk: Abramis academic publishing. 149-155.
- Card, S. K., Mackinlay, J. D. and Schneiderman, B., eds., 1999. *Readings in Information Visualization: Using Vision to Think*. San Francisco: Morgan Kaufmann.
- Cairo, A., 2013. *The functional art: an introduction to information graphics and visualization*. Berkeley: New Riders.
- Carden, T., 2011. Introducing Fizz. *Bloom Blog*, 25 Φεβρουαρίου 2011. Διαθέσιμο στο: <http://blog.bloom.io/2011/02/25/introducing-fizz/> [Πρόσβαση 28 Αυγούστου 2014].

- Chris, 2014. Fitbit Dashboard Updated with Weekly Activity and More. *Fitbit Blog*, 30 Ιουλίου 2014. Διαθέσιμο στο: <http://blog.fitbit.com/fitbit-dashboard-updated-with-weekly-activity-and-more/> [Πρόσβαση 28 Αυγούστου 2014].
- Cisco Systems, 2014. *Internet of Things*. Διαθέσιμο στο: <http://www.cisco.com/web/solutions/trends/iot/overview.html> [Πρόσβαση 17 Αυγούστου 2014].
- Cleveland, W. S., 1994. *The Elements of Graphing Data*. Summit: Hobart Press.
- Cleveland, W. S. and McGill, R., 1984. Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods. *Journal of the American Statistical Association*, 79 (387), 531-534. Διαθέσιμο στο: <http://www.cs.ubc.ca/~tmm/courses/cpsc533c-04-spr/readings/cleveland.pdf> [Πρόσβαση 17 Αυγούστου 2014].
- Commission of the European Communities, 2009. *Internet of Things - An action plan for Europe*. Commission of the European Communities. Διαθέσιμο στο: <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52009DC0278&from=EN> [Πρόσβαση 17 Αυγούστου 2014].
- Cooper, B. B., 2014. How does your fitness tracker know when you're asleep?. *Exist*, 16 Φεβρουαρίου 2014. Διαθέσιμο στο: <https://exist.io/blog/post/fitness-tracker-sleep/> [Πρόσβαση 17 Αυγούστου 2014].
- Dant, A. and Richards, J., 2011. Behind the rumours: how we built our Twitter riots interactive. *The Guardian DataBlog*, 08 Δεκεμβρίου 2011. Διαθέσιμο στο: <http://www.theguardian.com/news/datablog/2011/dec/08/twitter-riots-interactive> [Πρόσβαση 17 Αυγούστου 2014].
- Datoo, S., 2014. These companies are tracking the fitness of their employees. *The Guardian*, 17 Μαρτίου 2014. Διαθέσιμο στο: <http://www.theguardian.com/technology/2014/mar/17/why-companies-are-tracking-the-fitness-of-their-employees> [Πρόσβαση 17 Αυγούστου 2014].

- Davenport, T. H., 2014. *Big Data at Work: Dispelling the Myths, Uncovering the Opportunities*. Boston: Harvard Business Review Press.
- Davenport, T. H. and Dyché, J., 2013. *Big Data in Big Companies*. SAS. Διαθέσιμο στο: <http://www.sas.com/resources/asset/Big-Data-in-Big-Companies.pdf> [Πρόσβαση 17 Αυγούστου 2014].
- Davenport, T. H. and Kim, J., 2013. *Keeping Up with the Quants*. Boston: Harvard Business Review Press.
- Davenport, T. H. and Patil, D.J., 2012. Data Scientist: The Sexiest Job of the 21st Century. *Harvard Business Review*, October 2012. Διαθέσιμο στο: <http://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century/ar/pr> [Πρόσβαση 17 Αυγούστου 2014].
- Dix, A., Finlay, J., Abowd, G. D. and Beale, R., 2004. *Επικοινωνία Ανθρώπου - Υπολογιστή*. Αθήνα: Γκιούρδας.
- Dondis, D. A., 1973. *A Primer of Visual Literacy*. Cambridge: MIT Press.
- Driscoll, M., 2010. Color: The Cinderella of Data Visualization. In: J. Steele and N. Illinsky, eds. *Beautiful Visualization: Looking at Data Through the Eyes of Experts*. Sebastopol: O'Reilly, 59-67.
- Economist, 2010. New rules for big data. *The Economist*, 25 Φεβρουαρίου. Διαθέσιμο στο: <http://www.economist.com/node/15557487> [Πρόσβαση 17 Αυγούστου 2014].
- Ericsson, 2013. *Ericsson Mobility Report, November 2013*. Ericsson. Διαθέσιμο στο: <http://www.ericsson.com/res/docs/2013/ericsson-mobility-report-november-2013.pdf> [Πρόσβαση 17 Αυγούστου 2014].
- Ericsson, 2014. *OSS/BSS Agility - Fresh Data: Big data, real-time action*. Διαθέσιμο στο: <http://www.ericsson.com/oss-bss/agility/fresh-data/> [Πρόσβαση 17 Αυγούστου 2014].

- Ericsson ConsumerLab, 2013. *10 Hot Consumer Trends 2014*. Ericsson.
Διαθέσιμο στο: <http://www.ericsson.com/res/docs/2013/consumerlab/10-hot-consumer-trends-report-2014.pdf> [Πρόσβαση 17 Αυγούστου 2014].
- Essmaker, R. and Essmaker, T., 2014. Nicholas Felton on The Great Discontent. *The Great Discontent*, 26 Φεβρουαρίου 2014. Διαθέσιμο στο: <https://thegreatdiscontent.com/interview/nicholas-felton> [Πρόσβαση 28 Αυγούστου 2014].
- European Commission, 2013. *Digital Agenda Scoreboard 2013*. European Commission. Διαθέσιμο στο: <https://ec.europa.eu/digital-agenda/sites/digital-agenda/files/DAE%20SCOREBOARD%202013%20-%20SWD%202013%20217%20FINAL.pdf> [Πρόσβαση 17 Αυγούστου 2014].
- Evans, C. L., 2014. Nicholas Felton Quantified Literally Every Conversation He Had in 2013. *Motherboard*, 19 Αυγούστου 2014. Διαθέσιμο στο: <http://motherboard.vice.com/read/nicholas-felton-quantified-literally-everything-he-said-in-2013> [Πρόσβαση 28 Αυγούστου 2014].
- Evans, D., 2011. *The Internet of Things: How the Next Evolution of the Internet Is Changing Everything*. Cisco Internet Business Solutions Group. Διαθέσιμο στο: http://www.cisco.com/web/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf [Πρόσβαση 17 Αυγούστου 2014].
- Ευρωπαϊκή Επιτροπή, 2014. *Ημέρα Προστασίας Δεδομένων 2014: η αντιπρόεδρος Ρέντινγκ ζητά νέο σύμφωνο για την προστασία των δεδομένων στην Ευρώπη*. [δελτίο τύπου] 28 Ιανουαρίου 2014. Διαθέσιμο στο: http://europa.eu/rapid/press-release_IP-14-70_el.htm [Πρόσβαση 17 Αυγούστου 2014].
- Executive Office of the President, 2014. *Big Data: Seizing Opportunities, Preserving Values*. Washington: The White House. Διαθέσιμο στο: http://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_5.1.14_final_print.pdf [Πρόσβαση 17 Αυγούστου 2014].

- Facebook Newsroom, 2014. *Company Info*. Διαθέσιμο στο:
<http://newsroom.fb.com/company-info/> [Πρόσβαση 11 Ιουνίου 2014].
- Fayyad, U., Piatetsky-Shapiro, G. and Smyth, P., 1996. From Data Mining to Knowledge Discovery in Databases. *AI Magazine*, 17 (3), 37-54. Διαθέσιμο στο: <http://www.aaai.org/ojs/index.php/aimagazine/article/view/1230/1131> [Πρόσβαση 17 Αυγούστου 2014].
- Fedorov, V., 2012. Early Success Stories: Timeline Apps and Open Graph. *Facebook Developers Blog*, 15 Φεβρουαρίου 2012. Διαθέσιμο στο:
<https://developers.facebook.com/blog/post/2012/02/15/early-success-stories--timeline-apps-and-open-graph/> [Πρόσβαση 17 Αυγούστου 2014].
- Fekete, J.-D., van Wijk, J. J., Stasko, J. T. and North, C., 2008. The Value of Information Visualization. In: A. Kerren, J. T. Stasko, J.-D. Fekete and C. North, eds. *Information Visualization*. Berlin: Springer, 1-18.
- Feltron, 2014a. *About Nicholas Felton*. Διαθέσιμο στο: <http://feltron.com/about.html> [Πρόσβαση 17 Αυγούστου 2014].
- Feltron, 2014b. *2013 Annual Report*. Διαθέσιμο στο: <http://feltron.com/FAR13.html> [Πρόσβαση 17 Αυγούστου 2014].
- Feltron, 2014c. *To visualize a year of communication required...* Διαθέσιμο στο:
<http://feltron.tumblr.com/post/95211354121/to-visualize-a-year-of-communication-required> [Πρόσβαση 17 Αυγούστου 2014].
- Feltron, 2014d. *Home*. Διαθέσιμο στο: <http://feltron.com/> [Πρόσβαση 17 Αυγούστου 2014].
- Few, S., 2013. *Information Dashboard Design*. 2^η έκδοση. Burlingame: Analytics Press.
- Fields, E., 2014. *8.2 Preview: Story Points*. Διαθέσιμο στο:
<http://www.tableausoftware.com/about/blog/2014/5/82-preview-tell-story-your-data-story-points-30761> [Πρόσβαση 17 Αυγούστου 2014].

Fitbit, 2014a. *The Fitbit Story*. Διαθέσιμο στο: <https://www.fitbit.com/story> [Πρόσβαση 17 Αυγούστου 2014].

Fitbit, 2014b. *Fitbit Dashboard*. Διαθέσιμο στο: <https://www.fitbit.com/login> [Πρόσβαση 28 Αυγούστου 2014].⁷⁹

Friendly, M., 2008a. A Brief History of Data Visualization. In: C. Chen, W. K. Hardle and A. Unwin, eds. *Handbook of Data Visualization*. Berlin: Springer. 15-56.

Friendly, M., 2008b. The Golden Age of Statistical Graphics. *Statistical Science*, 23 (4), 502-535. Διαθέσιμο στο: http://projecteuclid.org/download/pdfview_1/euclid.ss/1242049392 [Πρόσβαση 28 Αυγούστου 2014].

Friendly, M., 2009. *Milestones in the history of thematic cartography, statistical graphics, and data visualization*. Διαθέσιμο στο: <http://www.math.yorku.ca/SCS/Gallery/milestone/milestone.pdf> [Πρόσβαση 17 Αυγούστου 2014].

Gartner, 2011. *Gartner Says Solving 'Big Data' Challenge Involves More Than Just Managing Volumes of Data*. [δελτίο τύπου] 27 Ιουνίου 2011. Διαθέσιμο στο: <http://www.gartner.com/newsroom/id/1731916> [Πρόσβαση 17 Αυγούστου 2014].

Gartner, 2013. *Gartner Says the Internet of Things Installed Base Will Grow to 26 Billion Units By 2020*. [δελτίο τύπου] 12 Δεκεμβρίου 2013. Διαθέσιμο στο: <http://www.gartner.com/newsroom/id/2636073> [Πρόσβαση 17 Αυγούστου 2014].

Gartner, 2014. *Big Data - BI - Free Gartner Research*. Διαθέσιμο στο: <http://www.gartner.com/it-glossary/big-data/> [Πρόσβαση 17 Αυγούστου 2014].

Gershon, N. and Eick, S. G., 1997. Information Visualization. *IEEE Computer Graphics and Applications*, 17 (4), 29-31. Διαθέσιμο στο:

⁷⁹ Πρόκειται για περιβάλλον που απαιτεί πιστοποιημένη είσοδο χρήστη. Το περιεχόμενο του ιστοχώρου είναι μοναδικό για κάθε χρήστη, καθώς περιλαμβάνει ατομικά δεδομένα.

<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=595265&isnumber=13035>

[Πρόσβαση 17 Αυγούστου 2014].

Gershon, N. and Page, W., 2001. What Storytelling Can Do for Information Visualization. *Communications of the ACM*, 44 (8), 31-37. Διαθέσιμο στο:

<http://delivery.acm.org/10.1145/390000/381653/p31-gershon.pdf> [Πρόσβαση 17

Αυγούστου 2014].

Gray, J., Bounegru, L. and Chambers, L., eds., 2012. The Data Journalism Handbook. Sebastopol: O'Reilly.

Guardian, 2013. NSA Files: Decoded - What the revelations mean for you. *The Guardian*, 01 Νοεμβρίου 2013. Διαθέσιμο στο:

<http://www.theguardian.com/world/interactive/2013/nov/01/snowden-nsa-files-surveillance-revelations-decoded> [Πρόσβαση 17 Αυγούστου 2014].

Guardian Interactive Team, Procter, R., Vis, F. and Voss, A., 2011. How riot rumours spread on Twitter. *The Guardian*, 07 Δεκεμβρίου 2011. Διαθέσιμο στο: <http://www.theguardian.com/uk/interactive/2011/dec/07/london-riots-twitter> [Πρόσβαση 17 Αυγούστου 2014].

Hand, D., Mannila, H. and Smyth, P., 2001. *Principles of Data Mining*. Cambridge: MIT Press.

Hannaford, L., 2013. Recalculating the newsroom: The rise the journo-coder? In: J. Mair, R. L. Keeble, P. Bradshaw and T. Beleaga, eds. *Data Journalism: Mapping the future*. Suffolk: Abramis academic publishing. 175-187.

Hanrahan, P., Stolte, C. and Mackinlay, J., 2007. *Visual analysis for everyone*. Tableau. Διαθέσιμο στο:

<http://www.tableausoftware.com/sites/default/files/whitepapers/visual-analysis-for-everyone.pdf> [Πρόσβαση 17 Αυγούστου 2014].

Heer, J., Bostock, M. and Ogievetsky, V., 2010. A Tour through the Visualization Zoo: A survey of powerful visualization techniques, from the obvious to the obscure. *Queue*, 8 (5), 1-22. Διαθέσιμο στο:

<http://delivery.acm.org/10.1145/1810000/1805128/p20-heer.pdf> [Πρόσβαση 17

Αυγούστου 2014].

Heer, J. and Schneiderman, B., 2012. Interactive Dynamics for Visual Analysis:

A taxonomy of tools that support the fluent and flexible use of visualizations. *Queue*, 10 (2), 1-26. Διαθέσιμο στο:

<http://delivery.acm.org/10.1145/2150000/2146416/p30-heer.pdf> [Πρόσβαση 17

Αυγούστου 2014].

Hernandez, D., 2013. How Facebook Builds a Digital Signature for You (And Your World). *Wired*, 18 Ιουλίου 2013. Διαθέσιμο στο:

<http://www.wired.com/2013/07/entities-graph/> [Πρόσβαση 17 Αυγούστου 2014].

Holtzschue, L., 2011. *Understanding Color: An Introduction for Designers*. New Jersey: John Wiley & Sons.

Hoffer, D., 2014. What Does Big Data Look Like? Visualization Is Key for

Humans. *Wired*, 10 Ιανουαρίου 2014. Διαθέσιμο στο:

<http://www.wired.com/2014/01/big-data-look-like-visualization-key-humans/> [Πρόσβαση

17 Αυγούστου 2014].

Hurwitz, J., Nugent, A., Halper, F. and Kaufman, M., 2013. *Big Data for Dummies*. Hoboken: Wiley.

Intel IT Center, 2012. *Big Data 101: Unstructured Data Analytics*. Intel Corporation. Διαθέσιμο στο:

<http://www.intel.com/content/dam/www/public/us/en/documents/solution-briefs/big-data-101-brief.pdf> [Πρόσβαση 17 Αυγούστου 2014].

Intel IT Center, 2013. *Big Data Visualization: Turning Big Data Into Big Insights*. Intel Corporation. Διαθέσιμο στο:

<http://www.intel.com/content/dam/www/public/us/en/documents/white-papers/big-data-visualization-turning-big-data-into-big-insights.pdf> [Πρόσβαση 17 Αυγούστου 2014].

International Telecommunications Union, 2005. *ITU Internet Reports 2005: The Internet of Things*. Geneva: International Telecommunication Union (ITU).

Διαθέσιμο στο: http://www.itu.int/dms_pub/itu-s/opb/pol/S-POL-IR.IT-2005-SUM-PDF-E.pdf
[Πρόσβαση 17 Αυγούστου 2014].

Jankun-Kelly, T. J., Kosara, R., Kindlmann, G., North, C., Ware, C. and Wes Bethel, E., 2006. Is There Science in Visualization?. *IEEE Visualization Conference Compendium*, 68-71. Διαθέσιμο στο:
http://kosara.net/papers/2006/Jankun-Kelly_VisPanel_2006.pdf [Πρόσβαση 17 Αυγούστου 2014].

Jewett, D., 2014. *7 Tips to Succeed with Big Data in 2014*. Tableau. Διαθέσιμο στο: http://www.tableausoftware.com/sites/default/files/media/7-tips-to-succeed-with-big-data-2014_0.pdf [Πρόσβαση 17 Αυγούστου 2014].

Joint Committee on Standards for Graphic Presentation, 1915. Preliminary Report Published for the Purpose of Inviting Suggestions for the Benefit of the Committee. *Publications of the American Statistical Association*, 14 (112), 790-797. Διαθέσιμο στο: <http://www.jstor.org/stable/2965153> [Πρόσβαση 17 Αυγούστου 2014].

Kadavy, D., 2011. *Design for Hackers*. West Sussex: Wiley.

Kaplan, A. M. and Haenlein, M., 2010. Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*, 53 (1), 59-68.
Διαθέσιμο στο: <http://www.sciencedirect.com/science/article/pii/S0007681309001232#>
[Πρόσβαση 17 Αυγούστου 2014].

Kerren, A. and Schreiber, F., 2012. Toward the role of interaction in Visual Analytics. In: C. Laroque, J. Himmelspace, R. Pasupathy, O. Rose, and A. M. Uhrmacher, eds. *Proceedings of the 2012 Winter Simulation Conference (WSC)*, 9-12 Δεκεμβρίου 2012, Berlin. IEEE, 1-13. Διαθέσιμο στο:
<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6465208> [Πρόσβαση 28 Αυγούστου 2014].

Kirk, A., 2012. *Data Visualization: a successful design process*. Birmingham: Packt.

- Kosara, R. and Mackinlay, J., 2013. Storytelling: The next step for Visualization. *Computer*, 46 (5), 44-50. Διαθέσιμο στο: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&number=6412677> [Πρόσβαση 28 Αυγούστου 2014].
- Laney, D., 2001. *3D Data Management: Controlling Data Volume, Velocity, and Variety*. META Group. Διαθέσιμο στο: <http://blogs.gartner.com/douglaney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf> [Πρόσβαση 17 Αυγούστου 2014].
- Lewis, P. and Newburn, T., 2011. The Reading the Riots project: our methodology explained. *The Guardian*, 05 Δεκεμβρίου 2011. Διαθέσιμο στο: <http://www.theguardian.com/uk/2011/dec/05/reading-the-riots-methodology-explained> [Πρόσβαση 17 Αυγούστου 2014].
- Li, A., 2010. Connecting to Everything You Care About. *Facebook*, [facebook note] 19 Απριλίου 2010. Διαθέσιμο στο: <https://www.facebook.com/notes/facebook/connecting-to-everything-you-care-about/382978412130> [Πρόσβαση 17 Αυγούστου 2014].
- Lima, M., 2011. *Visual Complexity: Mapping Patterns of Information*. New York: Princeton Architectural Press.
- Lin, S., Fortuna, J., Kulkarni, C., Stone, M. and Heer, J., 2013. Selecting Semantically-Resonant Colors for Data Visualization. *Computer Graphics Forum*, 32 (3), 401-410. Διαθέσιμο στο: <http://vis.stanford.edu/files/2013-SemanticColor-EuroVis.pdf> [Πρόσβαση 17 Αυγούστου 2014].
- Lindquist, E., 2011. Surveying the world of visualization. *Australian National University*. HC Coombs Policy Forum - Background paper. Διαθέσιμο στο: https://crawford.anu.edu.au/public_policy_community/content/doc/Visualisation_roundtable_2_Background_Paper.pdf [Πρόσβαση 17 Αυγούστου 2014].

- Lorenz, M., 2012. Why Journalists Should Use Data. In: J. Gray, L. Bounegru, and L. Chambers, eds. *The Data Journalism Handbook*. Sebastopol: O'Reilly. 3-5.
- Lundquist, E., 2013. 'Small Data' Analysis the Next Big Thing, Advocates Assert. *eWeek*, 10 Σεπτεμβρίου 2013. Διαθέσιμο στο: <http://www.eweek.com/enterprise-apps/small-data-analysis-the-next-big-thing-advocates-assert.html/> [Πρόσβαση 17 Αυγούστου 2014].
- Lunden, I., 2014. Instagram Is The Fastest-Growing Social Site Globally, Mobile Devices Rule Over PCs For Access. *TechCrunch*, 21 Ιανουαρίου 2014. Διαθέσιμο στο: <http://techcrunch.com/2014/01/21/instagram-is-the-fastest-growing-social-site-globally-mobile-devices-rule-over-pcs-for-social-access/> [Πρόσβαση 17 Αυγούστου 2014].
- Lupton, D., 2013. Understanding the Human Machine. *IEEE Technology and Society Magazine*, Winter 2013, 25-30. Διαθέσιμο στο: <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=06679313> [Πρόσβαση 17 Αυγούστου 2014].
- Mackinlay, J., 1986. Automating the Design of Graphical Presentations of Relational Information. *ACM Transactions on Graphics*, 5 (2). Διαθέσιμο στο: <http://delivery.acm.org/10.1145/30000/22950/p110-mackinlay.pdf> [Πρόσβαση 28 Αυγούστου 2014].
- Manovich, L., 2010. *What is Visualization?*. Διαθέσιμο στο: http://manovich.net/content/04-projects/062-what-is-visualization/61_article_2010.pdf [Πρόσβαση 28 Αυγούστου 2014].
- Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., Hung Byers, A. and McKinsey Global Institute, 2011. *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute. Διαθέσιμο στο: www.mckinsey.com/~media/McKinsey/dotcom/Insights%20and%20pubs/MGI/Research/Tec

hnology%20and%20Innovation/Big%20Data/MGI_big_data_full_report.ashx [Πρόσβαση 17 Αυγούστου 2014].

Mayer-Schönberger, V. and Cukier, K., 2013. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. London: John Murray.

Mell, P. and Grance, T., 2011. *The NIST Definition of Cloud Computing: Recommendations of the National Institute of Standards and Technology*. National Institute of Standards and Technology. Διαθέσιμο στο: <http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf> [Πρόσβαση 28 Αυγούστου 2014].

Meirelles, I., 2013. *Design for Information*. Beverly: Rockport.

Morioka, A. and Stone, T., 2006. *Color Design Workbook: A Real-World Guide to Using Color in Graphic Design*. Beverly: Rockport.

Morozov, E., 2013. *To Save Everything Click Here: Technology Solutionism and the urge to fix problems that don't exist*. Allen Lane.

Μπαμπινιώτης, Γ., 2011. Μικρό Λεξικό της Νέας Ελληνικής Γλώσσας. 2^η έκδοση. Lexicology Centre [iPhone application]. Διαθέσιμο στο: <https://itunes.apple.com/gr/app/g.-mpampiniote-mikro-lexiko/id421613903?mt=8> [Πρόσβαση 28 Αυγούστου 2014].

National Science Foundation, 2000. *America's Investment in the Future*. Arlington: National Science Foundation. Διαθέσιμο στο: <http://www.nsf.gov/about/history/nsf0050/pdf/aif.pdf> [Πρόσβαση 28 Αυγούστου 2014].

Norman, D. A., 1988. *The Psychology of Everyday Things*. Basic Books

OED Online, 2014. *big, adj. and adv.* Oxford University Press. Διαθέσιμο στο: <http://www.oed.com/view/Entry/18833#eid301162178> [Πρόσβαση 28 Αυγούστου 2014].

- Ohlhorst, F. J., 2013. *Big Data Analytics: Turning Big Data into Big Money*. Hoboken: Wiley.
- Organization for Economic Co-operation and Development (OECD), 2007. *Participative Web: User-Created Content*. OECD. Διαθέσιμο στο: <http://www.oecd.org/internet/ieconomy/38393115.pdf> [Πρόσβαση 17 Αυγούστου 2014].
- Oxford Dictionaries, 2013. *Tweet geekery and epic crowdsourcing: an Oxford English Dictionary update*. OxfordWords blog, 13 Ιουνίου 2013. Διαθέσιμο στο: <http://blog.oxforddictionaries.com/2013/06/oed-june-2013-update/> [Πρόσβαση 28 Αυγούστου 2014].
- Oxford Dictionaries, 2014. *big data*. Διαθέσιμο στο: <http://www.oxforddictionaries.com/definition/english/big-data> [Πρόσβαση 28 Αυγούστου 2014].
- Perera, C., Zaslavsky, A., Christen, P., and Georgakopoulos, D., 2014. Context Aware Computing for the Internet of Things: A Survey. *IEEE Communications Surveys & Tutorials*, 16 (1), 414-454. Διαθέσιμο στο: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6512846&isnumber=6734841> [Πρόσβαση 17 Αυγούστου 2014].
- Pew Research Center, 2014a. *The Web at 25 in the U.S.* Pew Research Center. Διαθέσιμο στο: http://www.pewinternet.org/files/2014/02/PIP_25th-anniversary-of-the-Web_0227141.pdf [Πρόσβαση 17 Αυγούστου 2014].
- Pew Research Center, 2014b. *The Internet of Things Will Thrive by 2025*. Pew Research Center. Διαθέσιμο στο: http://www.pewinternet.org/files/2014/05/PIP_Internet-of-things_0514142.pdf [Πρόσβαση 17 Αυγούστου 2014].
- Piramanayagam, S. N., 2012. Introduction. In: Piramanayagam, S. N. and Chong T. C. eds. *Developments in Data Storage: Materials Perspective*. New Jersey: Wiley, 1-9.

- Pollock, R., 2013. Forget big data, small data is the real revolution. *The Guardian DataBlog*, 25 Απριλίου 2013. Διαθέσιμο στο: <http://www.theguardian.com/news/datablog/2013/apr/25/forget-big-data-small-data-revolution> [Πρόσβαση 17 Αυγούστου 2014].
- Post, F. H., Nielson, G. M. and Bonneau, G.-P., eds., 2003. *Data Visualization: The State of the Art*. New York: Springer Science+Business Media.
- Press, G., 2013. Big Data News: A Revolution Indeed. *Forbes*, 18 Ιουνίου 2013. Διαθέσιμο στο: <http://www.forbes.com/sites/gilpress/2013/06/18/big-data-news-a-revolution-indeed/> [Πρόσβαση 28 Αυγούστου 2014].
- Processing, 2014. *Overview. A short introduction to the Processing software and projects from the community*. Διαθέσιμο στο: <http://www.processing.org/overview/> [Πρόσβαση 28 Αυγούστου 2014].
- Procter, R., 2011. How 2.6m tweets were analysed to understand reaction to the riots. *The Guardian*, 07 Δεκεμβρίου 2011. Διαθέσιμο στο: <http://www.theguardian.com/uk/2011/dec/07/how-tweets-analysed-understand-riots> [Πρόσβαση 28 Αυγούστου 2014].
- Quantified Self, 2012?. *About us*. Διαθέσιμο στο: <http://quantifiedself.com/about/> [Πρόσβαση 28 Αυγούστου 2014].
- Richards, J. and Lewis, P., 2011. How Twitter was used to spread – and knock down – rumours during the riots. *The Guardian*, 07 Δεκεμβρίου 2011. Διαθέσιμο στο: <http://www.theguardian.com/uk/2011/dec/07/how-twitter-spread-rumours-riots> [Πρόσβαση 17 Αυγούστου 2014].
- Ríos, M., 2013. The geography of Tweets. Twitter blog, 31 Μαΐου 2013. Διαθέσιμο στο: <https://blog.twitter.com/2013/the-geography-of-tweets> [Πρόσβαση 17 Αυγούστου 2014].
- Rios, M. and Lin, J., 2012. Distilling Massive Amounts of Data into Simple Visualizations: Twitter Case Studies. *Papers from the 2012 ICWSM Workshop*, 22-25. Διαθέσιμο στο:

http://www.umiacs.umd.edu/~jimmylin/publications/Rios_Lin_2012.pdf [Πρόσβαση 28 Αυγούστου 2014].

Rogers, S., 2012. Data Visualization DIY: Our Top Tools. *In: J. Gray, L. Bounegru, and L. Chambers, eds. The Data Journalism Handbook*. Sebastopol: O'Reilly. 206-212.

Rogers, S., 2013a. *Facts are sacred*. London: Faber and Faber.

Rogers, S., 2013b. Farewell Guardian, hello Twitter. *Simon Rogers*, 18 Απριλίου 2013. Διαθέσιμο στο: <http://simonrogers.net/2013/04/18/farewell-guardian-hello-twitter/> [Πρόσβαση 28 Αυγούστου 2014].

Rogers, S., 2014. How (and Why) We Made Twitter Reverb: Twitter and Periscope talk about their process and goals. *Source*, 13 Αυγούστου 2014. Διαθέσιμο στο: <https://source.opennews.org/en-US/articles/how-and-why-we-made-twitter-reverb/> [Πρόσβαση 28 Αυγούστου 2014].

Rowley, J., 2007. The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of Information Science*, 33 (2), 163-180. Διαθέσιμο στο: <http://jis.sagepub.com/content/33/2/163.full.pdf+html> [Πρόσβαση 28 Αυγούστου 2014].

Sallam, R. L., Tapadinhas, J., Parenteau, J., Yuen, D. and Hostmann, B., 2014. *Magic Quadrant for Business Intelligence and Analytics Platforms*. Gartner. Διαθέσιμο στο: <http://www.gartner.com/technology/reprints.do?id=1-1QLGACN&ct=140210> [Πρόσβαση 17 Αυγούστου 2014].

SAS, 2013. *From Big Data to Meaningful Information*. SAS. Διαθέσιμο στο: http://www.sas.com/content/dam/SAS/en_us/doc/conclusionpaper1/from-big-data-to-meaningful-information-106328.pdf [Πρόσβαση 17 Αυγούστου 2014].

Schmid, C. F., 1978. The role of Standards in Graphic Presentation. *In: Graphic Presentation of Statistical Information: Papers Presented at the 136th Annual Meeting of the American Statistical Association*. 23-26 Αυγούστου 1976, Boston. Bureau of the Census, 69-78. Διαθέσιμο στο:

<http://books.google.gr/books?id=FRQBYsncAJkC&hl=el&pg=PA69> [Πρόσβαση 28 Αυγούστου 2014].

Schneiderman, B., 1996. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. *In: Proceedings of IEEE Symposium on Visual Languages*. 03-06 Σεπτεμβρίου 1996, Boulder. IEEE, 336 - 343. Διαθέσιμο στο: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=545307> [Πρόσβαση 28 Αυγούστου 2014].

Segel, E. and Heer, J., 2010. Narrative Visualization: Telling Stories with Data. *IEEE Transactions on Visualization and Computer Graphics*, 16 (6), 1139-1148. Διαθέσιμο στο: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5613452&isnumber=5613414> [Πρόσβαση 28 Αυγούστου 2014].

Selfiecity, 2014. *Selfiecity: Investigating the style of self-portraits (selfies) in five cities across the world*. Διαθέσιμο στο: <http://selfiecity.net/> [Πρόσβαση 17 Αυγούστου 2014].

Sherin, A., 2012. *Design Elements: Color Fundamentals*. Beverly: Rockport.

Simon, P., 2013. *Too big to ignore: the business case for big data*. Hoboken: Wiley

Simonite, T., 2012. What Facebook Knows. *MIT Technology Review*, 13 Ιουνίου 2012. Διαθέσιμο στο: <http://www.technologyreview.com/featuredstory/428150/what-facebook-knows/> [Πρόσβαση 17 Αυγούστου 2014].

Spence, R., 2007. *Information Visualization: Design for Interaction*. 2^η έκδοση. Essex: Pearson Education.

Stone, M. C., 2003. *A Field Guide to Digital Color*. Natick: A K Peters.

Sun, E. and Iyer, V., 2013. Under the Hood: The Entities Graph. *Facebook Engineering*, [facebook] 6 Ιουνίου 2013. Διαθέσιμο στο: <https://www.facebook.com/notes/facebook-engineering/under-the-hood-the-entities-graph/10151490531588920> [Πρόσβαση 17 Αυγούστου 2014].

- Tableau, 2014a. *Tableau 8.2 Brings Data to the Mac and Data Storytelling to Everyone*. [δελτίο τύπου] 19 Ιουνίου 2014. Διαθέσιμο στο:
<http://www.tableausoftware.com/about/press-releases/2014/tableau-82-brings-data-mac-and-data-storytelling-everyone> [Πρόσβαση 17 Αυγούστου 2014].
- Tableau, 2014b. *On-Demand Training: Story Points Transcript*. Tableau.
Διαθέσιμο στο:
http://www.tableausoftware.com/sites/default/files/training/story_points_transcript.zip
[Πρόσβαση 17 Αυγούστου 2014].
- Tan, P.-N., Steinbach, M. and Kumar, V., 2014. *Introduction to Data Mining*.
Essex: Pearson Education.
- TapestryConference, 2014. *Fernanda Viegas and Martin Wattenberg Tapestry 2014*. Διαθέσιμο στο: <https://www.youtube.com/watch?v=VRgJnX1lr04> [Πρόσβαση 17 Αυγούστου 2014].
- Tapscott, D. and Williams, A. 2008. *Wikinomics: How Mass Collaboration Changes Everything*. 2η έκδοση. London: Atlantic Books.
- Techopedia, 2010?. *Front and Back Ends*. Διαθέσιμο στο:
<http://www.techopedia.com/definition/24794/front-and-back-ends> [Πρόσβαση 17 Αυγούστου 2014].
- The White House, 2014. *The Big Data and Privacy Review*. Διαθέσιμο στο:
<http://www.whitehouse.gov/issues/technology/big-data-review> [Πρόσβαση 17 Αυγούστου 2014].
- Thomas, J. J. and Cook, K. A., 2005 (eds.). *Illuminating the Path: The Research & Development Agenda for Visual Analytics*. National Visualization and Analytics Center. Διαθέσιμο στο:
http://vis.pnnl.gov/pdf/RD_Agenda_VisualAnalytics.pdf [Πρόσβαση 17 Αυγούστου 2014].
- Tidwell, 2011. *Designing Interfaces*. 2^η έκδοση. Sebastopol: O'Reilly.

- Trendviz, 2014. *Trendviz: Your organization's first small step in Big Data*. CLEVER°FRANKE and Ai Applied. Διαθέσιμο στο: http://www.cleverfranke.com/TrendViz_Whitepaper_v1.1.pdf [Πρόσβαση 17 Αυγούστου 2014].
- Tufte, E. R., 1990. *Envisioning Information*. Connecticut: Graphics Press.
- Tufte, E. R., 1997. *Visual Explanations: Images and Quantities, Evidence and Narrative*. Connecticut: Graphics Press.
- Tufte, E. R., 2001. *The Visual Display of Quantitative Information*. 2^η έκδοση. Connecticut: Graphics Press.
- Tukey, J. W., 1962. The future of data analysis. *Annals of Mathematical Statistics*, 33 (1), 1-67. Διαθέσιμο στο: http://web.stanford.edu/~gavish/documents/Tukey_the_future_of_data_analysis.pdf [Πρόσβαση 17 Αυγούστου 2014].
- Twitter, 2014. *Tweeting "happy new year" around the world*. Διαθέσιμο στο: <http://twitter.github.io/interactive/newyear2014/> [Πρόσβαση 17 Αυγούστου 2014].
- Twitter Reverb, 2014a. #IceBucketChallenge - Mentions of "ice bucket challenge", ALS or hashtag. *Twitter Reverb*. Διαθέσιμο στο: <http://reverb.guru/view/704481784367058763> [Πρόσβαση 28 Αυγούστου 2014].
- Twitter Reverb, 2014b. Summer storms - '#HurricaneArthur', 'storm', 'rain' mentions. *Twitter Reverb*. Διαθέσιμο στο: <http://reverb.guru/view/230172050563440044> [Πρόσβαση 28 Αυγούστου 2014].
- Viégas, F. B. and Wattenberg, M., 2007. Artistic Data Visualization: Beyond Visual Analytics. In: D. Schuler, ed. *Proceedings of the 2nd International Conference on Online Communities and Social Computing*. 22-27 Ιουλίου 2007, Beijing. Berlin: Springer, 182-191. Διαθέσιμο στο: http://link.springer.com/chapter/10.1007%2F978-3-540-73257-0_21 [Πρόσβαση 28 Αυγούστου 2014].

- W3C, 2004. *Architecture of the World Wide Web, Volume One*. Διαθέσιμο στο: <http://www.w3.org/TR/webarch/#index> [Πρόσβαση 17 Αυγούστου 2014].
- W3C, 2009. *Help and FAQ*. Διαθέσιμο στο: <http://www.w3.org/Help/#webinternet> [Πρόσβαση 17 Αυγούστου 2014].
- Ward, M., Grinstein, G. and Keim, D., 2010. *Interactive Data Visualization: Foundations, Techniques and Applications*. Natick: AK Peters.
- Ware, C., 2013. *Information Visualization: Perception for Design*. 3^η έκδοση. Waltham: Elsevier.
- Wattenberg, M. 2002. Arc Diagrams: Visualizing Structure in Strings. In: *Proceedings of the IEEE Symposium on Information Visualization (InfoVis'02)*. IEEE, 110-116. Διαθέσιμο στο: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1173155> [Πρόσβαση 17 Αυγούστου 2014].
- Wilkinson, L., 2005. *The Grammar of Graphics*. 2^η έκδοση. New York: Springer.
- Wolf, G., 2009. Know Thyself: Tracking Every Facet of Life, from Sleep to Mood to Pain, 24/7/365. *Wired*, 22 Ιουνίου 2009. Διαθέσιμο στο: http://archive.wired.com/medtech/health/magazine/17-07/lbnp_knowthyself?currentPage=all [Πρόσβαση 17 Αυγούστου 2014].
- Wolf, G., 2010. The Data-Driven Life. *The New York Times*, 01 Ιουνίου 2010. Διαθέσιμο στο: <http://www.nytimes.com/2010/05/02/magazine/02self-measurement-t.html?pagewanted=all&r=0> [Πρόσβαση 17 Αυγούστου 2014].
- Wolf, G. 2011. What is The Quantified Self?. *Quantified Self*, 03 Μαρτίου 2011. Διαθέσιμο στο: <http://quantifiedself.com/2011/03/what-is-the-quantified-self/> [Πρόσβαση 17 Αυγούστου 2014].
- Wong, D. M., 2010. *The Wall Street Journal Guide to Information Graphics: The Dos & Don'ts of Presenting Data, Facts, and Figures*. New York: W. W. Norton & Company.

Wong, P. C. and Thomas, J., 2004. Visual Analytics. *IEEE Computer Graphics and Applications*, 24 (5), 20-21. Διαθέσιμο στο:

<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1333623> [Πρόσβαση 28 Αυγούστου 2014].

World Economic Forum, 2011. *Personal Data: The Emergence of a New Asset Class*. World Economic Forum. Διαθέσιμο στο:

http://www3.weforum.org/docs/WEF_ITTC_PersonalDataNewAsset_Report_2011.pdf [Πρόσβαση 17 Αυγούστου 2014].

World Economic Forum, 2012. *Big Data, Big Impact: New Possibilities for International Development*. World Economic Forum. Διαθέσιμο στο:

http://www3.weforum.org/docs/WEF_TC_MFS_BigDataBigImpact_Briefing_2012.pdf [Πρόσβαση 17 Αυγούστου 2014].

Wright, H., 2007. *Introduction to Scientific Visualization*. London: Springer.

Yau, N., 2013. *Data Points: Visualization That Means Something*. Indianapolis: Wiley.

Zak, E., 2014a. How #Ferguson Has Unfolded on Twitter. *The Wall Street Journal*, 18 Αυγούστου 2014. Διαθέσιμο στο:

<http://blogs.wsj.com/dispatch/2014/08/18/how-ferguson-has-unfolded-on-twitter/> [Πρόσβαση 28 Αυγούστου 2014].

Zak, E., 2014b. #Ferguson - Mentions of Twitter of #Ferguson. *Twitter Reverb*.

Διαθέσιμο στο: <http://reverb.guru/view/126740849281262175> [Πρόσβαση 28 Αυγούστου 2014].

Απόδοση ξενόγλωσσων όρων

αγγλικός όρος	ελληνική απόδοση
actuators	ενεργοποιητές
anomaly detection	ανίχνευση ανωμαλιών
artifact	τεχνούργημα
association analysis	ανάλυση σχέσεων
bar charts	γραφήματα ράβδων
Big Data	Μεγάλα Δεδομένα
brightness	φωτεινότητα
business intelligence	επιχειρηματική ευφυΐα
butterfly diagram	διάγραμμα πεταλούδας
classification	κατηγοριοποίηση
cloud computing	νεφοϋπολογιστική
cluster analysis	ανάλυση συστάδων
color damage	χρωματικής ζημιάς
complexity	πολυπλοκότητα
computer graphics	γραφικά υπολογιστών
content management systems	συστήματα διαχείρισης περιεχομένου
continuous variables	συνεχείς μεταβλητές
data journalism	δημοσιογραφία δεδομένων
data marts	αγορές δεδομένων
data mining	εξόρυξη δεδομένων
data scientists	επιστημονικοί υπεύθυνοι δεδομένων
data set	σύνολο δεδομένων
data warehouses	αποθήκες δεδομένων
Digital Agenda	Ψηφιακό Θεματολόγιο
discrete variables	διακριτές μεταβλητές
drill down	αναλυτική καθόδος
effectiveness	αποτελεσματικότητα
elementary perceptual tasks	στοιχειώδεις αντιληπτικές διεργασίες
Exploratory Data Analysis	Διερευνητική Ανάλυση Δεδομένων
expressiveness	εκφραστικότητα
format	μορφή
genre	είδος
hue	απόχρωση
identifiers	αναγνωριστικά
In-memory analytics	ανάλυση στη μνήμη του υπολογιστικού συστήματος

αγγλικός όρος	ελληνική απόδοση
Internet	Διαδίκτυο
Internet of Things	Διαδίκτυο των Πραγμάτων
interval scales	διαστημικές κλίμακες
Knowledge Discovery in Databases	Ανακάλυψη Γνώσης από Βάσεις Δεδομένων
line charts	γραφήματα γραμμών
Liquid Crystal Display	οθόνες υγρών κρυστάλλων
Machine Learning	Μηχανική Μάθηση
media visualization	οπτικοποίηση μέσω
mobile-Internet	κινητό Διαδίκτυο
narrative structure tactics	τακτικές αφηγηματικής δομής
Natural Language Processing	Επεξεργασία Φυσικής Γλώσσας
nominal scales	ονομαστικές/κατηγορικές κλίμακες
ordinal scales	τακτικές ή διατεταγμένες κλίμακες
personal data	προσωπικά δεδομένα
pie charts	γραφήματα πιτών
pixels	εικονοστοιχεία
predictive modelling	προγνωστικά μοντέλα
privacy	ιδιωτικότητα
prosumer	παραγωγός-καταναλωτής
qualitative variables	ποιοτικές μεταβλητές
Quantified Self	Ποσοτικοποιημένος Εαυτός
quantitative variables	ποσοτικές μεταβλητές
ratio scales	αναλογικές κλίμακες
reduction	μείωση
regression	παλινδρόμηση
relational data	σχεσιακά δεδομένα
routers	δρομολογητές
saturation	κορεσμός
scatter plots	γραφήματα διασποράς
Scripting languages	γλώσσες προγραμματισμού
self-tracking	αυτο-καταγραφή
servers	εξυπηρετητές
small multiples	πολλαπλά μικρογραφήματα
smartphones	έξυπνα τηλέφωνα
social media	κοινωνικά μέσα
spatial variables	χωρικές μεταβλητές
structured data	δομημένα δεδομένα
sub-pixels	υπο-εικονοστοιχεία

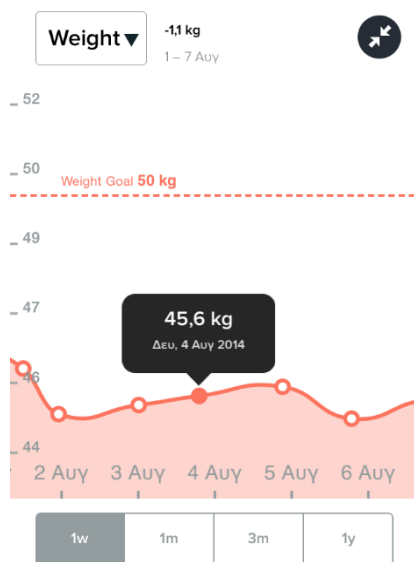
αγγλικός όρος	ελληνική απόδοση
tag	ετικέτα
time-series plots	χρονολογικές σειρές
tweet	τιτίβισμα
unclassified choropleth map	χωροπληθής χάρτης
unstructured data	μη δομημένα δεδομένα
User Generated Content	Περιεχόμενο Παραγόμενο από τον Χρήστη
variability	μεταβλητότητα
variety	ποικιλία
velocity	ταχύτητα
veracity	αξιοπιστία
Visual Analytics	Οπτική Ανάλυση Δεδομένων
visual cues	οπτικές ενδείξεις
visual narrative tactics	τακτικές οπτικής αφήγησης
volume	όγκος
wearable computers	φορέσιμοι υπολογιστές
Web	Ιστός
Web 2.0	Ιστός 2.0
web log	δεδομένα καταγραφών Ιστού
World Wide Web	Παγκόσμιος Ιστός

Παράρτημα

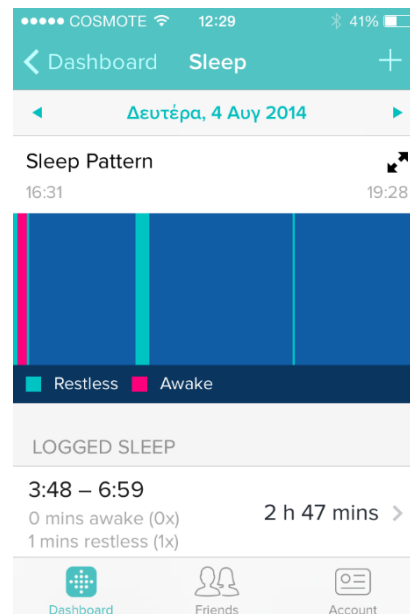
Ενδεικτικές οθόνες εφαρμογής *Fitbit Dashboard* για *iPhone*



Εικόνα 52: Αρχική οθόνη *Fitbit Dashboard* για *iPhone*



Εικόνα 53: Οθόνη παρακολούθησης εξέλιξης βάρους



Εικόνα 54: Οθόνη παρακολούθησης ύπνου

