



Πανεπιστήμιο Μακεδονίας

Τμήμα Οργάνωσης και Διοίκησης Επιχειρήσεων
Πρόγραμμα Μεταπτυχιακών Σπουδών στο Τουριστικό Μάνατζμεντ Επιχειρήσεων και
Οργανισμών

Διπλωματική Εργασία

Θέμα:

**«Η διαχείριση μεγάλου όγκου δεδομένων στον Τουρισμό- Μια προσπάθεια
βιβλιογραφικής και εμπειρικής προσέγγισης του θεματος»**

Του Γεώργιου Μιχαηλίδη ΜΤΧ17005

Επιβλέπων Καθηγητής: Χρήστος Βασιλειάδης

Υποβλήθηκε ως προαπαιτούμενο για την απόκτηση μεταπτυχιακού διπλώματος

Απρίλιος 2019

Περίληψη

Σκοπός της παρούσας εργασίας είναι η διερεύνηση του ρόλου και των πλεονεκτημάτων της ανάλυσης των μεγάλων δεδομένων στον τουρισμό, τους τομείς στους οποίους μπορεί να ενσωματωθεί ώστε να προσφέρει λύσεις και βελτιώσεις σε προβλήματα που απασχολούν τον κλάδο. Η ευρεία χρήση του διαδικτύου και η ραγδαία ανάπτυξη της τεχνολογίας των έξυπνων συσκευών αναπόφευκτα επηρέασε και τον τουριστικό τομέα όπου η εφαρμογή των εργαλείων που προσφέρει η ανάλυση των μεγάλων δεδομένων μπορεί να ενσωματωθεί σε πάρα πολλές λειτουργίες των επιχειρήσεων. Συνεπώς, οι τομείς του τουρισμού που επιλέχθηκαν στην παρούσα εργασία αφορούν στο διαδίκτυο, τα μέσα κοινωνικής δικτύωσης και τις ηλεκτρονικές κριτικές. Επίσης, εξετάζεται η αποτελεσματικότητα των εργαλείων ανάλυσης και επεξεργασίας των μεγάλων δεδομένων όπως η ανάλυση συναισθημάτων μέσω προσεγγίσεων μηχανικής μάθησης. Τα αποτελέσματα της μελέτης αυτής δείχνουν ότι αν και βρισκόμαστε ακόμα σε αρχικό στάδιο η ως τώρα εφαρμογή των μεγάλων δεδομένων έχει στεφθεί με μεγάλη επιτυχία και οι προοπτικές ανάπτυξης του τουρισμού καθώς οι τεχνολογίες εξελίσσονται και γίνονται πιο προσιτές είναι τεράστιες.

Περιεχόμενα

Τι είναι η Ανάλυση;	7
Περιγραφική Ανάλυση	8
Διαγνωστική Ανάλυση	8
Προγνωστική Ανάλυση	10
Συμβουλευτική Ανάλυση	10
Τι είναι τα Μεγάλα Δεδομένα;	11
Τα Χαρακτηριστικά των Μεγάλων Δεδομένων	13
1 Όγκος	14
2 Ταχύτητα	14
3 Ποικιλία	15
4 Εγκυρότητα	15
5 Αξία	15
Εφαρμογή Μεγάλων Δεδομένων στον Τουρισμό	15
Η Έννοια των Μεγάλων Δεδομένων στον Τουρισμό	17
Τα Πλεονεκτήματα χρήσης Μεγάλων Δεδομένων στον Τουρισμό	19
1. Αξιοπιστία	20
2. Νέες ροές πληροφοριών	20
3. Δεδομένα σε πραγματικό χρόνο και η άμεση πρόβλεψη (nowcasting)	20
Τα Οφέλη της Χρήσης Μεγάλων Δεδομένων στις Τουριστικές Επιχειρήσεις	21
Η Συμπεριφορά των Καταναλωτών	22
Μηχανισμοί Ανατροφοδότησης	24
Συλλογή Μεγάλων Δεδομένων για την Τουριστική Πρόγνωση	26
1. Ο Σκοπός	26
2. Η Απεικόνιση των Μεγάλων Δεδομένων	26
3. Η Δόμηση των Μεγάλων Δεδομένων	27
Η Επιλογή και η Συρρίκνωση των Μεγάλων Δεδομένων	27
1. Το Παραγοντικό Μοντέλο (factor model)	27
2. Η Μέθοδος LASSO	28
Πλαίσιο για την Τουριστική Πρόβλεψη Ζήτησης με την Χρήση Μεγάλων Δεδομένων	28
1. Μοντέλο Μικτής Συχνότητας στα Μεγάλα Δεδομένα	29
2. Παραγοντικό μοντέλο και συνδυασμοί πρόβλεψης	30
Ο Ψηφιακός Τουρισμός	30

Εισαγωγή	30
Η Ανάδυση του Ψηφιακού Ταξιδιώτη και οι Φορητές Συσκευές	32
Ο Ψηφιακός Ταξιδιώτης και η Συναίσθηση του Περιβάλλοντος.....	35
Ο Ψηφιακός Ταξιδιώτης στην Καθημερινότητα	36
Ο Ψηφιακός Ταξιδιώτης και η Ανάπτυξη του Έξυπνου Τουρισμού	37
Η Αξιοποίηση των Κριτικών της Τουριστικής Βιομηχανίας.....	40
Εισαγωγή	40
Η Εμφάνιση των Ηλεκτρονικών Κριτικών στον Ξενοδοχειακό Κλάδο	41
Τα Είδη και τα Χαρακτηριστικά των Ηλεκτρονικών Κριτικών	43
Η Επιρροή των Ηλεκτρονικών Κριτικών στον Ξενοδοχειακό Κλάδο.....	45
Η Επιρροή των Ηλεκτρονικών Κριτικών στην Καταναλωτική Συμπεριφορά	47
Εξαγωγή Γνώσεων για τους Πελάτες μέσω Ηλεκτρονικών Κριτικών	49
Αποκρίσεις στις Ηλεκτρονικές Κριτικές.....	51
Ιδέες και Μέθοδοι Ανάλυσης του Διαδικτύου και των Μέσω Κοινωνικής Δικτύωσης	54
Εισαγωγή	54
1 Η Συναισθηματική Ανάλυση στον Τουριστικό Τομέα	56
1α Η Επιβλεπόμενη Μηχανική Μάθηση.....	57
1β Μάθηση Βάσει Λεξιλογίου.....	59
1γ Η Μη Επιβλεπόμενη Μηχανική Μάθηση	60
1δ Η Σημασιολογική Προσέγγιση	60
1ε Η Υβριδική Προσέγγιση	61
2 Η Ανίχνευση Θέματος.....	63
2α Επιβλεπόμενη Ανίχνευση Θέματος	64
2β Μη Επιβλεπόμενη Ανίχνευση Θέματος.....	65
3 Ανίχνευση Συναισθημάτων.....	69
3α Ανίχνευση Υποκειμενικότητας	69
3β Ανίχνευση Συναισθήματος.....	70
Ερευνητικά ερωτήματα.....	72
Συνεντεύξεις.....	74
Αρθούρος Τσιλίδης (Mouzenidis Travel)	74
Σωκράτης Παπουνίδης (Ellinair).....	75
Συμπεράσματα ποιοτικής έρευνας.....	78
Συμπεράσματα εμπειρικής έρευνας	80

Η διαχείριση μεγάλου όγκου δεδομένων στον Τουρισμό- Μια προσπάθεια βιβλιογραφικής και εμπειρικής προσέγγισης του θεματος

Προτάσεις για περαιτέρω έρευνα	82
Περιορισμοί της έρευνας	83
Προτάσεις για διοικητική εφαρμογή στην πράξη	84
Βιβλιογραφία	87
Παράρτημα 1	93
Παράρτημα 2	96

Εισαγωγή

Ο τουρισμός αποτελεί σημαντικό συστατικό πολλών εθνικών και τοπικών οικονομιών. Ενώ η επιτυχία της διαχείρισης του τουρισμού βασίζεται σε τομείς πολιτικής και διαχείρισης, αυξανόμενη είναι η εξάρτηση της για την κατανόηση εις βάθος της συνεχώς μεταβαλλόμενης καταναλωτικής συμπεριφοράς προκειμένου να κινητοποιηθούν οι αναγκαίοι πόροι για την κάλυψη των αναγκών και επιθυμιών των καταναλωτών. Ευρέως τεκμηριωμένη είναι η άποψη ότι η πληροφορική σήμερα αφενός μετέβαλε θεμελιωδώς τον τρόπο με τον οποίο οι ταξιδιώτες καταναλώνουν και λαμβάνουν πρόσβαση σε προϊόντα και υπηρεσίες τουρισμού, αφετέρου έχει δημιουργήσει νέες ανάγκες και ευκαιρίες για την πρόσβαση στα δεδομένα και την καλύτερη κατανόηση της συμπεριφοράς των τουριστών. Από τη σκοπιά αυτή ο τουρισμός είναι ένας πλούσιος και ιδανικός χώρος για την ανάλυση μεγάλων δεδομένων και αυτό γιατί οι δυνατότητες οποιασδήποτε επιχείρησης ή προορισμού για την καταγραφή, την παρακολούθηση, την ανάλυση και την ερμηνεία των συμπεριφορών των ταξιδιωτών είναι ουσιώδης.

Η τεχνολογία έχει μεταμορφώσει την τουριστική εμπειρία και αυτό αποδεικνύεται από το γεγονός ότι η τεχνολογία συνδέει την καθημερινή μας ζωή με τον τρόπο που βιώνουμε το ταξίδι. Το Facebook, το Twitter, το Youtube και το Pinterest είναι βασικές εφαρμογές web 2.0 δεδομένου ότι αποτελούν καινοτόμους τρόπους για τη διευκόλυνση της ανταλλαγής πληροφοριών και της κοινωνικής δικτύωσης. Στα δίκτυα που υποστηρίζονται από την τεχνολογία ανήκουν επίσης και αυτά της κινητής τηλεφωνίας με τα smartphones και τα tablets τα οποία διευκολύνουν τις συναλλαγές και ενισχύσουν τους κοινωνικούς δεσμούς των ταξιδιωτών. Για πολλούς το κινητό τηλέφωνο είναι κάτι περισσότερο από ένα εργαλείο επικοινωνίας ή ένα αξεσουάρ της καθημερινότητας, στην πραγματικότητα τα smartphones αποτελούν αναπόσπαστο κομμάτι της ζωής μας. Η κινητή τεχνολογία οδηγεί αναμφισβήτητα σε μια πιο δημιουργική χρήση της. Πράγματι, έχει υποστηριχθεί ότι η ανάπτυξη στις υπηρεσίες βάσει τοποθεσίας (LBS) καθιστούν τους προορισμούς πιο εντυπωσιακούς και ελκυστικούς για τους ταξιδιώτες (Hannam, Butler, & Paris, 2014). Αυτό σημαίνει ότι αυτή η αλλαγή της τουριστικής εμπειρίας που ήταν αποτέλεσμα της αλληλεπίδρασής της με την τεχνολογία πληροφοριών είναι πολύπλευρη και λαμβάνει χώρα σε διάφορους τεχνολογικούς και κοινωνικούς τομείς.

Η διαχείριση μεγάλου όγκου δεδομένων στον Τουρισμό- Μια προσπάθεια βιβλιογραφικής και εμπειρικής προσέγγισης του θεματος

Επίσης η αλλαγή στο τεχνολογικό περιβάλλον δημιουργεί νέες ευκαιρίες αλλά και ανάγκες για την κατανόηση και την περιγραφή των συνθηκών που επηρεάζουν τα ταξίδια και την εμπειρία του τουρισμού.

Τι είναι η Ανάλυση;

Η ανάλυση ως όρος έχει ευρεία έννοια και περιλαμβάνει τις διαδικασίες, τα πλαίσια και τους αλγορίθμους για την εξαγωγή σημαντικών πληροφοριών από δεδομένα. Τα ακατέργαστα δεδομένα καθαυτά είναι απροσδιόριστα μέχρι να κατηγοριοποιηθούν και να μετατραπούν σε χρήσιμες πληροφορίες. Ανάλυση είναι η διαδικασία της εξαγωγής και δημιουργίας πληροφορίας από ακατέργαστα δεδομένα μέσω σύμπτυξης, πλαισιοποίησης, φιλτραρίσματος και κατηγοριοποίησης των δεδομένων. Κατόπιν οι πληροφορίες που συλλέγονται οργανώνονται και δομούνται ώστε να εξάγουν πληροφορίες για το σύστημα και τους χρήστες, τις λειτουργίες και την πρόοδο αναφορικά με τους στόχους καθιστώντας το έτσι πιο έξυπνο και αποτελεσματικό.

Η επιλογή των τεχνολογιών, των αλγορίθμων και των πλαισίων για την ανάλυση βασίζεται στους στόχους της εφαρμογής της ανάλυσης. Για παράδειγμα οι στόχοι της ανάλυσης μπορεί να είναι 1) η πρόβλεψη (όπως η πρόβλεψη τουριστικών αφίξεων την επόμενη σεζόν), 2) η εύρεση μοτίβων στα δεδομένα (για παράδειγμα η εύρεση σελίδων σε έναν ιστότοπο που οι χρήστες δείχνουν μεγαλύτερη προτίμηση ή η εύρεση του πιο δημοφιλούς καταλύματος σε μια συγκεκριμένη περιοχή), 3) η εύρεση συσχετισμών στα δεδομένα (για παράδειγμα η εύρεση τιμών για σχετικές υπηρεσίες από ανταγωνιστικές εταιρίες ή η εύρεση συσχετισμού μεταξύ ειδήσεων και τιμών μετοχών).

Το Εθνικό Συμβούλιο Έρευνας (National Research Council, 2013) [1] έχει κατηγοριοποιήσει όλες τις υπολογιστικές εργασίες που εμπλέκουν μαζική ανάλυση δεδομένων και τις ονομάζει ως "επτά γίγαντες". Αυτές οι διαδικασίες περιλαμβάνουν, 1) τη Βασική Στατιστική (Basis Statistics), 2) τα Γενικευμένα Προβλήματα N Σωμάτων (Generalized N-Body Problems), 3) τους Γραμμικούς Αλγεβρικούς Υπολογισμούς (Linear Algebraic Computations), 4) τους Υπολογισμούς Θεωρητικών Γραφημάτων (Graph-Theoretic Computations), 5) την Βελτιστοποίηση (Optimization), 6) τα Ολοκληρώματα (Integration) και 7) τα Προβλήματα Στοίχισης Ακολουθιών (Alignment Problems). Αυτή η κατηγοριοποίηση των υπολογιστικών διαδικασιών έχει ως σκοπό την ταξινόμηση των εργασιών που αποδείχθηκαν χρήσιμες στην

Η διαχείριση μεγάλου όγκου δεδομένων στον Τουρισμό- Μια προσπάθεια βιβλιογραφικής και εμπειρικής προσέγγισης του θεματος

ανάλυση δεδομένων και την προσεγγιστική ομαδοποίηση τους με μαθηματική δομή και υπολογιστική στρατηγική.

Περιγραφική Ανάλυση

Η περιγραφική ανάλυση είναι η ανάλυση παλαιότερων δεδομένων για την παρουσίασή τους σε συνοπτική μορφή προκειμένου να ερμηνευτούν ευκολότερα. Σκοπός της περιγραφικής ανάλυσης είναι η απάντηση στο ερώτημα “τι συνέβη;”. Ένα μεγάλο μέρος των αναλύσεων που έγιναν σήμερα είναι περιγραφικής φύσεως με την χρήση στατιστικών συναρτήσεων (maximum, minimum, counts, top-N, percentage, mean κ.α.). Αυτά τα στατιστικά εργαλεία βοηθούν στην περιγραφή των δειγμάτων και στην συνοπτική παρουσίασή τους. Για παράδειγμα, ο υπολογισμός του συνολικού αριθμού θετικών σχολίων για το προσωπικό ξενοδοχειακής μονάδας στο TripAdvisor ή ο υπολογισμός του μέσου αριθμού επισκεπτών ανά μήνα σε έναν προορισμό.

Μεταξύ των επτά υπολογιστικών εργασιών (Σχήμα 1), εργασίες όπως η Βασική Στατιστική και οι Γραμμικοί Αλγεβρικοί Υπολογισμοί μπορούν να χρησιμοποιηθούν στις περιγραφικές αναλύσεις.

Διαγνωστική Ανάλυση

Η διαγνωστική ανάλυση είναι η ανάλυση παλαιών δεδομένων για την διάγνωση των λόγων για τους οποίους συνέβησαν ορισμένα γεγονότα. Σκοπός της διαγνωστικής ανάλυσης είναι η απάντηση στο ερώτημα “γιατί συνέβη;”. Η διαφορά της διαγνωστικής ανάλυσης με την περιγραφική έγκειται στο ότι ενώ η περιγραφική ανάλυση είναι χρήσιμη στην περιληπτική παρουσίαση μεγάλου όγκου δεδομένων με την χρήση στατιστικών εργαλείων η διαγνωστική ανάλυση μπορεί να παράξει πληροφορίες για το λόγο τον οποίο ορισμένα σφάλματα μπορεί να συνέβησαν βάσει αναζήτησης δειγμάτων με παρόμοια σφάλματα σε παλαιότερα δεδομένα. Για τον λόγο αυτό η διαγνωστική ανάλυση βρίσκει καλή εφαρμογή στον τομέα της υγείας.

Μεταξύ των επτά υπολογιστικών εργασιών, εργασίες όπως οι Γραμμικοί Αλγεβρικοί Υπολογισμοί, τα Γενικευμένα Προβλήματα N Σωμάτων και οι Υπολογισμοί Θεωρητικών Γραφημάτων μπορούν να χρησιμοποιηθούν για διαγνωστικές αναλύσεις.

Σχήμα 1
Τύποι Αναλύσεων



Υπολογιστικοί Γίγαντες Μεγάλων Δεδομένων

Προγνωστική Ανάλυση

Η προγνωστική ανάλυση είναι η ανάλυση κατά την οποία γίνεται είτε η πρόβλεψη της εμφάνισης ενός γεγονότος, είτε η πρόβλεψη του πιθανού αποτελέσματος ενός γεγονότος είτε η πρόβλεψη αξιών με τη χρήση εργαλείων πρόγνωσης. Σκοπός της προγνωστικής ανάλυσης είναι η απάντηση στο ερώτημα “τι είναι πιθανόν να συμβεί;”. Τα πεδία εφαρμογής της συγκεκριμένης ανάλυσης είναι ευρεία καθώς βρίσκει χρησιμότητα σε όλους τους κλάδους θετικών επιστημών. Η προγνωστική ανάλυση γίνεται με τη χρήση εργαλείων πρόβλεψης που λαμβάνουν πληροφορίες από υπάρχοντα δεδομένα. Αυτά τα εργαλεία μαθαίνουν τα μοτίβα και τις τάσεις των δεδομένων κι έτσι προβλέπουν την εμφάνιση ενός συμβάντος ή την πιθανή έκβαση ενός γεγονότος (μοντέλα ταξινόμησης) ή την πρόβλεψη αριθμών (μοντέλα παλινδρόμησης). Η ακρίβεια των εργαλείων πρόβλεψης εξαρτάται από την ποιότητα και τον όγκο των διαθέσιμων δεδομένων για την εκμάθηση τους. Προτού χρησιμοποιηθεί ένα μοντέλο πρόβλεψης πρέπει να επικυρωθεί με τα υπάρχοντα δεδομένα. Η συνηθέστερη προσέγγιση που υιοθετήθηκε κατά την ανάπτυξη εργαλείων πρόβλεψης είναι η διάσπαση των υφιστάμενων δεδομένων σε εκπαιδευτικά δεδομένα και δοκιμαστικά δεδομένα, Συγκεκριμένα, το 75% των δεδομένων χρησιμοποιείται για την εκμάθηση και το 25% χρησιμοποιείται για τη δοκιμή του μοντέλου πρόβλεψης.

Μεταξύ των επτά υπολογιστικών εργασιών, εργασίες όπως οι Γραμμικοί Αλγεβρικοί Υπολογισμοί, τα Γενικευμένα Προβλήματα N Σωμάτων, οι Υπολογισμοί Θεωρητικών Γραφημάτων, τα Ολοκληρώματα και τα Προβλήματα Στοίχισης Ακολουθιών μπορούν να χρησιμοποιηθούν για προγνωστικές αναλύσεις.

Συμβουλευτική Ανάλυση

Ενώ η προγνωστική ανάλυση χρησιμοποιεί εργαλεία και μοντέλα για την πρόβλεψη πιθανού αποτελέσματος ενός γεγονότος, η συμβουλευτική ανάλυση χρησιμοποιεί πληθώρα εργαλείων για την πρόβλεψη διαφορετικών αποτελεσμάτων και προτείνει το βέλτιστο σχέδιο δράσης για το κάθε αποτέλεσμα. Σκοπός της συμβουλευτικής ανάλυσης είναι η απάντηση στο ερώτημα “ποια τακτική θα φέρει το επιθυμητό αποτέλεσμα;”.

Η συμβουλευτική ανάλυση μπορεί να προβλέψει τα πιθανά αποτελέσματα που μπορεί να φέρει η συγκεκριμένη πορεία δράσης. Μπορούμε να θεωρήσουμε την

Η διαχείριση μεγάλου όγκου δεδομένων στον Τουρισμό- Μια προσπάθεια βιβλιογραφικής και εμπειρικής προσέγγισης του θεματος

συμβουλευτική ανάλυση ως την ανάλυση που χρησιμοποιεί διαφορετικά μοντέλα πρόβλεψης για διαφορετικά δεδομένα. Ένα παράδειγμα συμβουλευτικής ανάλυσης είναι οι ταξιδιωτικοί προορισμοί που θα μπορούσε να προτείνει η ιστοσελίδα ενός ταξιδιωτικού πρακτορείου βάσει το ιστορικό περιήγησης του επισκέπτη.

Μεταξύ των επτά υπολογιστικών εργασιών, εργασίες όπως τα Γενικευμένα Προβλήματα Ν Σωμάτων, οι Υπολογισμοί Θεωρητικών Γραφημάτων, τα Ολοκληρώματα και τα Προβλήματα Στοιχίσης Ακολουθιών μπορούν να χρησιμοποιηθούν για συμβουλευτικές αναλύσεις.

Τι είναι τα Μεγάλα Δεδομένα;

Τα μεγάλα δεδομένα ορίζονται ως συλλογές συνόλων δεδομένων των οποίων ο όγκος, η ταχύτητα και η ποικιλία είναι τόσο μεγάλα ώστε είναι δύσκολο να αποθηκευτούν, να διαχειριστούν, να επεξεργαστούν και να αναλυθούν χρησιμοποιώντας παραδοσιακές βάσεις δεδομένων και εργαλεία επεξεργασίας δεδομένων. Τα τελευταία χρόνια υπήρξε μια αλματώδης αύξηση στον όγκο τόσο των δομημένων όσο και των μη δομημένων δεδομένων που παρήχθησαν από τους τομείς της πληροφορικής, της βιομηχανίας, της υγείας, του Διαδικτύου Πραγμάτων (Internet of Things) και άλλων συστημάτων.

Η IBM εκτιμά ότι δημιουργούνται 2,5 πεντάκις εκατομμύρια μπάιτς δεδομένων κάθε μέρα [R. Taft, M. Vartak, N.R. Satish, N. Sundaram, S. Madden, M. Stonebraker, 2014]. Παρακάτω παρατίθενται ορισμένα βασικά στοιχεία από την αναφορά του DOMO αναφορικά με τον όγκο των δεδομένων που παράγονται κάθε λεπτό σε δημοφιλείς διαδικτυακές πλατφόρμες [OpenRefine, 2016]:

- Οι χρήστες του Facebook μοιράζονται περί τις 4,16 εκατομμύρια πληροφορίες.
- Οι χρήστες του Twitter στέλνουν κοντά στα 300.000 τουίτς (tweets).
- Οι χρήστες του Instagram κάνουν λάικ περίπου σε 1,73 εκατομμύρια φωτογραφίες.
- Οι χρήστες του Youtube ανεβάζουν βίντεο διάρκειας 300 ωρών.
- Οι χρήστες της Apple κατεβάζουν περίπου 51.000 εφαρμογές.
- Οι χρήστες του Skype κάνουν περίπου 110.000 κλήσεις.
- Η Amazon λαμβάνει 4.300 νέους επισκέπτες

Η διαχείριση μεγάλου όγκου δεδομένων στον Τουρισμό- Μια προσπάθεια βιβλιογραφικής και εμπειρικής προσέγγισης του θεματος

- Οι επιβάτες της Uber κάνουν 694 διαδρομές.
- Οι συνδρομητές του Netflix βλέπουν σε ζωντανή ροή (stream) βίντεο διάρκειας 77.000 ωρών.

Τα μεγάλα δεδομένα έχουν τη δυνατότητα να τροφοδοτήσουν την επόμενη γενιά εφαρμογών που θα αξιοποιήσουν τη δύναμη των δεδομένων για να κάνουν τις εφαρμογές έξυπνες. Οι εφαρμογές μεγάλων δεδομένων καλύπτουν ένα ευρύ φάσμα κλάδων όπως αυτό του διαδικτύου, του λιανικού εμπορίου και του μάρκετινγκ, του τραπεζικού και χρηματοπιστωτικού, του βιομηχανικού, του υγειονομικού, το περιβαλλοντικού, του Διαδικτύου των Πραγμάτων και τα κυβερνο-φυσικά συστήματα (cyber-physical systems).

Τα μεγάλα δεδομένα ασχολούνται με τη συλλογή, αποθήκευση, επεξεργασία και ανάλυση αυτών των δεδομένων μαζικής κλίμακας. Ειδικά εργαλεία και πλαίσια απαιτούνται για την ανάλυση μεγάλων δεδομένων όταν:

1. Ο όγκος των δεδομένων είναι τόσο μεγάλος που είναι δύσκολο να αποθηκευτούν, να επεξεργαστούν και να αναλυθούν με έναν μηχανήμα μόνο.
2. Η ταχύτητα των δεδομένων είναι πολύ υψηλή και τα δεδομένα πρέπει να αναλύονται σε πραγματικό χρόνο.
3. Υπάρχουν ποικίλα δεδομένα που μπορούν να δομηθούν, να αποδομηθούν ή να ημι-δομηθούν και συλλέγονται από διαφορετικές πηγές δεδομένων.
4. Διάφορα είδη αναλύσεων πρέπει να εκτελεστούν προκειμένου να εξαχθεί αξία από τα δεδομένα, όπως η περιγραφική ανάλυση, η διαγνωστική ανάλυση, η προγνωστική ανάλυση και η συμβουλευτική ανάλυση.

Οι δυνατότητες και η αρχιτεκτονική παράλληλης επεξεργασίας των εργαλείων και πλαισίων των μεγάλων δεδομένων μπορούν να συγκριθούν με αυτές ενός μεγάλου συμπλέγματος υπολογιστικών συστημάτων.

Η ανάλυση μεγάλων δεδομένων περιλαμβάνει ορισμένα βήματα που ξεκινούν από τον καθαρισμό δεδομένων, τον μετασχηματισμό δεδομένων (data munging), την επεξεργασία και την απεικόνιση δεδομένων. Ο κύκλος ζωής της ανάλυσης μεγάλων δεδομένων ξεκινά από το συλλογή δεδομένων από πολλαπλές πηγές. Εξειδικευμένα εργαλεία και πλαίσια απαιτούνται για τη λήψη δεδομένων από διαφορετικές πηγές στη

Η διαχείριση μεγάλου όγκου δεδομένων στον Τουρισμό- Μια προσπάθεια βιβλιογραφικής και εμπειρικής προσέγγισης του θεματος

βάση δεδομένων. Τα δεδομένα αποθηκεύονται σε εξειδικευμένους υπηρεσίες αποθήκευσης δεδομένων (όπως καταναμημένα συστήματα αρχείων και μη σχεσιακές βάσεις δεδομένων) τα οποία έχουν σχεδιαστεί για να μπορούν να κλιμακοθετούν (scale).

Με βάση τις απαιτήσεις της ανάλυσης (σε παρτίδα ή σε πραγματικό χρόνο) και τον τύπο (περιγραφική, διαγνωστική, προγνωστική ή συμβουλευτική) χρησιμοποιούνται τα αντίστοιχα εξειδικευμένα πλαίσια. Οι αναλύσεις μεγάλων δεδομένων υποβοηθούνται από διάφορες τεχνολογίες, όπως για παράδειγμα το υπολογιστικό νέφος (cloud computing), τα καταναμημένα και παράλληλα πλαίσια επεξεργασίας, οι μη σχεσιακές βάσεις δεδομένων και η in-memory computing.

Παρακάτω παρατίθενται ορισμένα παραδείγματα μεγάλων δεδομένων:

- Τα δεδομένα που παράγονται από κοινωνικά δίκτυα, όπως κείμενα, εικόνες, ηχητικά και οπτικοακουστικά δεδομένα.
- Δεδομένα χρήσης (clickstream data) που παράγονται από διαδικτυακές εφαρμογές όπως το e-Commerce, για την ανάλυση της συμπεριφοράς των χρηστών.
- Δεδομένα που συλλέγονται από αισθητήρες σε ενσωματωμένες μηχανές βιομηχανικών και ενεργειακών συστημάτων για την παρακολούθηση της κατάστασής τους και την ανίχνευση σφαλμάτων.
- Δεδομένα υγειονομικής περίθαλψης που συλλέγονται στα συστήματα ηλεκτρονικής καταγραφής της υγείας (EHR).
- Αρχεία καταγραφής που δημιουργούνται από διαδικτυακές εφαρμογές.
- Στοιχεία χρηματιστηριακών αγορών.
- Δεδομένα συναλλαγών που παράγονται από τραπεζικές και χρηματοοικονομικές εφαρμογές.

Τα Χαρακτηριστικά των Μεγάλων Δεδομένων

Τα βασικά χαρακτηριστικά των μεγάλων δεδομένων είναι τα εξής:

1 Όγκος

Τα μεγάλα δεδομένα είναι είδος δεδομένων των οποίων ο όγκος είναι τόσο μεγάλος ώστε να μην χωράει σε ένα μόνο υπολογιστή, επομένως απαιτούνται εξειδικευμένα εργαλεία και πλαίσια για την διεκπεραίωση των διαδικασιών και αναλύσεων τέτοιων δεδομένων. Για παράδειγμα οι εφαρμογές κοινωνικής δικτύωσης επεξεργάζονται δισεκατομμύρια μηνύματα καθημερινά, βιομηχανικά και ενεργειακά συστήματα παράγουν καθημερινά τεραμπάιτς δεδομένων κλπ. Οι όγκοι των δεδομένων που παράγονται από τους τομείς τις πληροφορικής, της βιομηχανίας, της υγείας, του Διαδικτύου των Πραγμάτων και άλλων συστημάτων, αυξάνεται εκθετικά λόγω της μείωσης του κόστους των τεχνολογιών αποθήκευσης και επεξεργασίας δεδομένων και της ανάγκης εξαγωγής πολύτιμων πληροφοριών για τη βελτίωση των επιχειρηματικών διαδικασιών, την αποτελεσματικότητα και την εξυπηρέτηση των καταναλωτών. Παρόλο που δεν υπάρχει κάποιο όριο ως προς τον όγκο των δεδομένων ώστε να θεωρούνται μεγάλα δεδομένα, ωστόσο ο όρος μεγάλα δεδομένα χρησιμοποιείται για δεδομένα μεγάλης κλίμακας, τέτοιας που είναι δύσκολο να αποθηκευτούν, να διαχειριστούν και να επεξεργαστούν χρησιμοποιώντας παραδοσιακές βάσεις δεδομένων και μεθόδους επεξεργασίας δεδομένων.

2 Ταχύτητα

Η ταχύτητα των δεδομένων αναφέρεται στο πόσο γρήγορα παράγονται τα δεδομένα. Τα δεδομένα που από ορισμένες πηγές μπορεί να λαμβάνονται σε πολύ υψηλές ταχύτητες. Για παράδειγμα δεδομένα μέσω κοινωνικής δικτύωσης ή δεδομένα από αισθητήρες μηχανημάτων. Η ταχύτητα είναι ένα άλλο σημαντικό χαρακτηριστικό των μεγάλων δεδομένων και ο πρωταρχικός λόγος για την ραγδαία ανάπτυξή τους. Η υψηλή ταχύτητα των δεδομένων έχει ως αποτέλεσμα ο όγκος των δεδομένων να συλλέγεται γρήγορα και να καθίσταται πολύ μεγάλος σε σύντομο χρονικό διάστημα. Ορισμένες εφαρμογές έχουν αυστηρές προθεσμίες για την ανάλυση των δεδομένων (όπως αυτές στον κλάδο των πωλήσεων ή αυτές που ανήκουν στην κατηγορία της ανίχνευσης ηλεκτρονικής απάτης) όπου τα δεδομένα πρέπει να αναλύονται σε πραγματικό χρόνο.

Η διαχείριση μεγάλου όγκου δεδομένων στον Τουρισμό- Μια προσπάθεια βιβλιογραφικής και εμπειρικής προσέγγισης του θεματος

Εξειδικευμένα εργαλεία απαιτούνται για την λήψη δεδομένων τόσο υψηλής ταχύτητας στο σύστημα ώστε η ανάλυση των δεδομένων να πραγματοποιείται σε πραγματικό χρόνο.

3 Ποικιλία

Η ποικιλία αναφέρεται στα είδη των δεδομένων. Τα μεγάλα δεδομένα έχουν διάφορες μορφές, όπως είναι τα δομημένα, τα μη δομημένα ή τα ημι-δομημένα, συμπεριλαμβανομένων και άλλων ειδών δεδομένων όπως αυτές του κειμένου, εικόνας, ήχου, βίντεο και αισθητήρων. Τα συστήματα μεγάλων δεδομένων πρέπει να είναι αρκετά ευέλικτα ώστε να διαχειρίζονται μεγάλο εύρος δεδομένων.

4 Εγκυρότητα

Η εγκυρότητα αναφέρεται στο πόσο ακριβή είναι τα δεδομένα. Προκειμένου να εξαχθεί αξία από τα δεδομένα, αυτά θα πρέπει να φιλτραριστούν ώστε να αφαιρεθούν οι περιττές πληροφορίες. Οι εφαρμογές που βασίζονται σε δεδομένα μπορούν να επωφεληθούν από τα μεγάλα δεδομένα μόνο όταν τα δεδομένα είναι ουσιαστικά και ακριβή. Ως εκ τούτου, ο καθαρισμός των δεδομένων είναι σημαντικός ώστε τυχόν λανθασμένα και ελαττωματικά δεδομένα να φιλτράρονται.

5 Αξία

Η αξία των δεδομένων έχει να κάνει με την χρησιμότητα των δεδομένων για τον επιδιωκόμενο σκοπό. Ο απώτερος σκοπός οποιουδήποτε συστήματος ανάλυσης μεγάλων δεδομένων είναι να εξάγει αξία από τα δεδομένα. Η αξία των δεδομένων σχετίζεται επίσης με την εγκυρότητα ή την ακρίβεια των δεδομένων. Για ορισμένες εφαρμογές επίσης η αξία εξαρτάται από την ταχύτητα επεξεργασίας των δεδομένων.

Εφαρμογή Μεγάλων Δεδομένων στον Τουρισμό

Η τεχνολογία των πληροφοριών οδηγεί στην δημιουργία νέων μοτίβων στη συμπεριφορά των ταξιδιωτών κατά την λήψη αποφάσεων. Γενικά νοείται ότι η διαδικασία ταξιδιού περιλαμβάνει τρία στάδια, στο προ-ταξιδιωτικό, στο καθ' οδόν και επί τόπου, και στο μετά-ταξιδιωτικό, κατά τα οποία ο ταξιδιώτης αναλαμβάνει

διαφορετικές δραστηριότητες χρήσης πληροφοριών και αλληλεπίδρασης με το περιβάλλον. Πρόσφατες μελέτες έδειξαν ότι η χρήση των smartphones κατά τη διάρκεια του ταξιδιού είναι δυνατό να αλλάξουν την ταξιδιωτική εμπειρία "ξεκλειδώνοντας" την διαδικασία των τριών σταδίων (Wang, D., Xiang, Z., & Fesenmaier, D. R. 2016). Οι δραστηριότητες στις οποίες επιβάλλονταν οι τουρίστες στο προ-ταξιδιωτικό και στο μετά-ταξιδιωτικό στάδιο, αρχίζουν ολοένα και περισσότερο να εκπληρώνονται στο καθ' οδόν και επί τόπου στάδιο λόγω της γενικευμένης σύνδεσης στο διαδίκτυο οδηγώντας σε σημαντικές αλλαγές συμπεριφοράς. Το επίπεδο της ευελιξίας λήψης αποφάσεων κατά τη διάρκεια του ταξιδιού είναι πιθανό να αυξηθεί μιας και ο ταξιδιώτης θα μπορεί εύκολα να αλλάξει το αρχικό του σχέδιο λόγω της διαθεσιμότητας νέων πληροφοριών. Τελικά οι ταξιδιωτικές δραστηριότητες μπορεί να γίνουν πιο αυθόρμητες, με συνέπεια να πραγματοποιούνται περισσότερα απρογραμματίιστα ταξίδια ή δραστηριότητες. Η διαδικασία λήψης αποφάσεων στο καθ' οδόν στάδιο είναι πιο δυναμική δεδομένου ότι περιλαμβάνει μια σειρά από αλληλεξαρτώμενες αποφάσεις εντός των οποίων τα πλαίσια μεταγενέστερων αποφάσεων εξαρτώνται από προηγούμενες αποφάσεις. Συνεπώς η χρήση έξυπνων συσκευών όπως smartphones αλλάζει τα δεδομένα σε επίπεδο λήψης αποφάσεων για τα καθ' οδόν και επί τόπου στάδιο ειδικά λαμβάνοντας υπόψη την ευκολία πρόσβασης σε μηχανές αναζήτησης και μέσω κοινωνικής δικτύωσης οπουδήποτε και οποτεδήποτε. Αυτό προκαλεί τις ως τώρα γνώσεις μας αναφορικά με τη διαδικασία λήψης αποφάσεων ταξιδιών και δημιουργεί την ανάγκη για μια νέα σειρά εργαλείων ανάλυσης που μπορούν να αποθηκεύουν και να μετρούν τη διαδικασία και τη δομή της ταξιδιωτικής συμπεριφοράς.

Νέες τεχνολογικές καινοτομίες συνεχίζουν να αναδύονται και απαιτούν νέα οράματα για την ανάπτυξη του τουρισμού (Gretzel, U., Sigala, M., Xiang, Z. και Koo, C., 2015). Ένα τέτοιο παράδειγμα αποτελεί το Διαδίκτυο των Πραγμάτων το οποίο δηλώνει την παρουσία του γύρω μας μέσω ποικιλίας αντικειμένων όπως είναι οι αισθητήρες, οι ενεργοποιητές, οι κινητές συσκευές κλπ. Τα αντικείμενα αυτά είναι σε θέση να αλληλεπιδρούν μεταξύ τους και να συνεργάζονται με κοντινά τους αντικείμενα για την διεκπεραίωση κοινού έργου (Atzori, Iera, και Morabito, 2010). Επίσης μέσω της δυνατότητας σύνδεσης στο διαδίκτυο γεφυρώνεται το χάσμα μεταξύ του πραγματικού κόσμου και του ψηφιακού χώρου. Επιπλέον, η ανάπτυξη φορητών υπολογιστών

Η διαχείριση μεγάλου όγκου δεδομένων στον Τουρισμό- Μια προσπάθεια βιβλιογραφικής και εμπειρικής προσέγγισης του θεματος

υποστηρίζει μια πληθώρα εφαρμογών συμπεριλαμβανομένης και της οπτικής σήμανσης των αντικειμένων και των συσκευών Near Field Communication (NFC) που συμβάλλουν στην ανάπτυξη του Διαδικτύου Πραγμάτων (Borrego - Jaraba, Ruiz, και Gomez - Nieto, 2011). Το σημαντικό είναι ότι η εμφάνιση του Διαδικτύου Πραγμάτων μετατόπισε την παροχή υπηρεσιών από το σημερινό όραμα των υπηρεσιών συνεχούς σύνδεσης στις πάντα διαθέσιμες υπηρεσίες, οι οποίες δημιουργούνται και συγκροτούνται κατά την εκτέλεση ώστε να ανταποκρίνονται στις αυξημένες ανάγκες των χρηστών. Έτσι, προβλέπεται ότι μέσα στην επόμενη δεκαετία το διαδίκτυο θα μπορεί να πραγματοποιήσει το όραμα πολλών ετών με το να επιτύχει την σύνδεση αντικειμένων που μπορούν να αναγνωριστούν, να εντοπιστούν, να παρακολουθηθούν και να διαχειριστούν οποιαδήποτε στιγμή και οπουδήποτε.

Αυτή η νέα τεχνολογική δομή δημιουργεί νέες δυνατότητες σύνδεσης και τρόπους αλληλεπίδρασης εντός και εκτός ταξιδιού καθιστώντας δυνατό να επηρεαστεί ο τρόπος με τον οποίο αντιλαμβανόμαστε τη διαδικασία ταξιδιού. Ως εκ τούτου, είναι σαφές ότι οι τεχνολογικές εξελίξεις στον τομέα των κινητών, κοινωνικών, επικοινωνιακών και τεχνολογιών εντοπισμού θέσης έχουν επηρεάσει τις αισθήσεις και τις εμπειρίες των τουριστών μέσω της συναισθηματικής, αισθητικής, ενημερωτικής, παιχνιδιάρικης και κοινωνικής τους εμπλοκής, προκαλώντας τους να είναι πιο δημιουργικοί και αυθόρμητοι (Richards, 2011). Αυτές οι πρόσφατες εξελίξεις απαιτούν τη δημιουργία νέων μοντέλων συμπεριφοράς ταξιδιών, νέων μοντέλων σχεδιασμού προϊόντων και νέων μοντέλων έρευνας και αξιολόγησης, που με τη σειρά τους θα θεσπίσουν ένα νέο πρότυπο τουριστικής διοίκησης.

Η Έννοια των Μεγάλων Δεδομένων στον Τουρισμό

Η πληροφορία ευδοκίμει στην τουριστική βιομηχανία. Οι τεράστιες νέες αποθήκες δεδομένων (πολύ μεγαλύτερες από αυτές που συλλέγονται σε παραδοσιακές βάσεις δεδομένων) μπορούν να προσφέρουν στους ερευνητές, τους διευθυντές και τους υπεύθυνους χάραξης πολιτικής, τα στοιχεία που απαιτούνται για τη λήψη αποφάσεων βάσει αριθμών και αναλύσεων αντί διαισθήσεων ή προηγούμενων εμπειριών (Frederiksen, 2012). Η αφθονία των Μεγάλων Δεδομένων στον Τουρισμό έχει τη δυνατότητα να παράσχει νέα - άκρως ενημερωμένα συμπεράσματα σχετικά με την

ανθρώπινη δραστηριότητα και συμπεριφορά που θα προσδώσει μεγάλη ώθηση στην τουριστική βιομηχανία και θα ωφελήσει όχι μόνο τους καταναλωτές αλλά και επαγγελματίες που δραστηριοποιούνται στην τουριστική βιομηχανία (Fuchs, Horken και Lexhagen, 2014).

Οι ταξιδιώτες αφήνουν διαφορετικά ψηφιακά ίχνη στο διαδίκτυο καθώς χρησιμοποιούν κινητές τεχνολογίες. Ο κάθε ταξιδιώτης μπορεί να διαθέσει μεγάλες ποσότητες δεδομένων για οποιοδήποτε από τα τρία ταξιδιωτικά στάδια: προ-ταξιδιωτικό, καθ' οδόν και επί τόπου και μετά-ταξιδιωτικό (Hendrik & Perdana, 2014). Τα περισσότερα από αυτά τα δεδομένα είναι εξωτερικής φύσης όπως είναι τα tweets ή τροφοδοσίες (feeds) άλλων μέσων κοινωνικής δικτύωσης.

Λόγω του μεγάλου όγκου αποθηκευμένων ακατέργαστων δεδομένων στο νέφος (cloud) και προκειμένου να εξαχθεί αξία και χρήσιμη πληροφορία από αυτά, απαραίτητη είναι η επεξεργασία και ανάλυσή τους. Πολλοί από τους δυνητικούς τουρίστες που προγραμματίζουν ένα ταξίδι δεν περιορίζονται σε μια απλή έρευνα στο διαδίκτυο κατά την αναζήτηση προορισμού, εισιτηρίων και διαμονής. Οι καταναλωτές των τουριστικών υπηρεσιών στρέφονται ολοένα και περισσότερο στα Μεγάλα Δεδομένα για την βελτίωση της λήψης αποφάσεων, την ανεύρεση ευκαιριών και την βελτίωση των εν γένει επιδόσεων τους (Irudeen και Samaraweera, 2013). Για παράδειγμα, τα μεγάλα δεδομένα μπορούν να χρησιμοποιηθούν για τη διασύνδεση των διάχυτων πληροφοριών από διαφορετικά συστήματα ώστε να βελτιωθεί η δυνατότητα λήψης αποφάσεων.

Τα μεγάλα δεδομένα παρέχουν πρωτοφανείς πληροφορίες για τον τρόπο λήψης αποφάσεων των πελατών, επιτρέποντας στις εταιρείες να εντοπίζουν και να αναλύουν τα καταναλωτικά πρότυπα, τις συστάσεις, την αγοραστική συμπεριφορά και άλλα κίνητρα που επηρεάζουν τις πωλήσεις. Οι εταιρείες που εμπλέκονται με τον τουρισμό μπορούν να βρουν καινοτόμους τρόπους χρήσης των ποικίλων πηγών δεδομένων ώστε να συνδεθούν με τους πιθανούς επισκέπτες σε κάθε στάδιο του ταξιδιού και να χρησιμοποιήσουν τις αποκτηθείσες πληροφορίες για την καλύτερη και έγκαιρη κατανόηση των δημογραφικών στοιχείων των επισκεπτών. Επίσης μπορούν να στοχεύσουν διαφημίσεις σε δυνητικούς αγοραστές που έχουν αναζητήσει κάποιον συγκεκριμένο προορισμό σε ιστοσελίδα ταξιδιωτικού πρακτορείου (Sust, V. O., Illera, E. G., Berengue, A. S., Garcia, R. G., Alonso, M. V. P. και Torres, M. J. T., 2014). Με την

Η διαχείριση μεγάλου όγκου δεδομένων στον Τουρισμό- Μια προσπάθεια βιβλιογραφικής και εμπειρικής προσέγγισης του θεματος

χρήση μεγάλων δεδομένων οι βιομηχανίες γίνονται πιο αποτελεσματικές. Όλο και περισσότερες εταιρείες αρχίζουν να εξειδικεύονται στην αποθήκευση και αξιολόγηση μεγάλου όγκου δεδομένων που σχετίζονται με διαμονές τουριστών σε ξενοδοχεία, τις συναλλαγές αγοράς και τις πληροφορίες πελατών προκειμένου να παρέχουν αποτελεσματικότερες και υψηλότερης ποιότητας υπηρεσίες.

Τα Πλεονεκτήματα χρήσης Μεγάλων Δεδομένων στον Τουρισμό

Το εξατομικευμένο μάρκετινγκ και η στοχευμένη σχεδίαση προϊόντος παρουσιάζουν εξαιρετικά ισχυρές ευκαιρίες τόσο για τους προμηθευτές τουριστικών προϊόντων όσο και για τους καταναλωτές. Είναι ξεκάθαρο ότι τα μεγάλα δεδομένα δίνουν τη δυνατότητα στις επιχειρήσεις να παρέχουν καλύτερες, στοχοποιημένες και πιο κερδοφόρες υπηρεσίες και προϊόντα στους καταναλωτές (Pries και Dunnigan, 2015). Για παράδειγμα οι αναλυτές Μεγάλων Δεδομένων μπορούν να ανιχνεύσουν πληροφορίες σχετικά με τα ενδιαφέροντα κάποιου καταναλωτή από φωτογραφίες που έχει αναρτήσει στο Facebook ή σε άλλα μέσα κοινωνικής δικτύωσης (πχ. ένας πάροχος τουριστικών υπηρεσιών που βλέποντας αναρτήσεις σχετικές με ορεινή ποδηλασία θα μπορούσε να προωθήσει τοπικούς προορισμούς ποδηλασίας ή λέσχες ποδηλασίας).

Προηγούμενες μελέτες με θέμα τον τουρισμό βασίστηκαν σε έρευνες ή απόψεις εμπειρογνομόνων πράγμα που σημαίνει ότι τα δείγματα λήφθηκαν από το σύνολο του πληθυσμού και όχι αμιγώς από το τουριστικό πλήθος. Αντίθετα, σε άλλη μελέτη με θέμα τα μεγάλα δεδομένα στον τουρισμό έγινε προσπάθεια παρουσίασης δεδομένων που βασίζονταν σε πραγματικές δραστηριότητες από όλους τους χρήστες τουριστικών προϊόντων αντί να αντληθούν πληροφορίες από δείγματα ερευνών (Irudeen και Samaraweera, 2013). Στο παρόν κεφάλαιο θα προσπαθήσουμε να παρουσιάσουμε ένα πλαίσιο που ενσωματώνει μεγάλα δεδομένα που παράγονται από τους ίδιους τους τουρίστες (μέσω κινητών τηλεφώνων που συνδέονται στο δίκτυο τηλεπικοινωνιών ή μέσω τραπεζικών καρτών που συνδέονται με τα τερματικά POS) και το οποίο αυξάνει τη γνώση της αγοράς-στόχου για την πρόβλεψη τουριστικής ζήτησης.

Όπως παρουσιάζεται παρακάτω τα μεγάλα δεδομένα στον τουρισμό με την χρήση καινοτόμων μεθόδων έχουν πλεονεκτήματα έναντι των παραδοσιακών μεθόδων:

1. Αξιοπιστία

Τα μεγάλα δεδομένα βασίζονται σε πραγματικά στοιχεία που προκύπτουν από τις ενέργειες των χρηστών και όχι σε έρευνες. Δηλαδή, έχουν αναλυθεί πραγματικές ενέργειες και όχι δηλωμένες προθέσεις ή απαντήσεις σε ερωτήσεις. Λαμβάνοντας υπόψη όλες τις πηγές πληροφόρησης μπορεί να πιστοποιηθεί ότι τα μεγάλα δεδομένα αυξάνουν το πλήθος δειγμάτων στο οποίο η συμβατική έρευνα τείνει να βασίζεται σε αρκετές τάξεις μεγέθους (Meeker και Hong, 2014). Η αξιοπιστία της ανάλυσης μεγάλων δεδομένων μας επιτρέπει να εξετάσουμε όλες τις πτυχές των πληροφοριών προκειμένου να παρέχουμε ολοκληρωμένα αποτελέσματα αντί για μεροληπτικά συμπεράσματα λόγω απώλειας πληροφοριών στο πλήθος δειγμάτων.

2. Νέες ροές πληροφοριών

Τα τουριστικά μεγάλα δεδομένα αποτελούν ένα είδος πληροφοριών που παράγονται από τους ίδιους τους ταξιδιώτες, εμπλουτίζει τη γνώση της τουριστικής αγοράς στόχου και είναι πολύ χρήσιμη για την ανάλυση της καταναλωτικής ζήτησης για διάφορα τουριστικά προϊόντα και υπηρεσίες (Hendrik και Perdana, 2014). Δεδομένου ότι τα τουριστικά Μεγάλα Δεδομένα είναι δομημένα και επανατοποθετημένα είναι δυνατόν να τα παραπέμψουμε σε άλλες πηγές όπως αυτές των μέσων κοινωνικής δικτύωσης και των ανοικτών δεδομένων, είτε πρόκειται για πηγές που δομούνται επί του παρόντος, είτε για πηγές πληροφοριών που ενδέχεται να δημιουργηθούν ή να δημοσιοποιηθούν στο μέλλον. Η ανάλυση των τουριστικών μεγάλων δεδομένων μπορεί να συγκριθεί με τα εσωτερικά δεδομένα της κάθε τουριστικής επιχείρησης προκειμένου να καθοριστεί εάν η προσφορά τουριστικών προϊόντων / υπηρεσιών στην περιοχή μιας πόλης είναι σε αρμονία με την τουριστική ζήτηση.

3. Δεδομένα σε πραγματικό χρόνο και η άμεση πρόβλεψη (nowcasting)

Μία από τις καινοτόμες χρήσεις των μεγάλων δεδομένων είναι η άμεση πρόβλεψη (nowcasting), το οποίο ουσιαστικά είναι η χρήση των δεδομένων σε πραγματικού χρόνου για να περιγραφούν οι πρόσφατες δραστηριότητες προτού δημοσιοποιηθούν οι επίσημες πηγές δεδομένων (Bollier και Firestone, 2010). Σύμφωνα με τον Varian (2014) ένα καλό παράδειγμα άμεσης πρόβλεψης καταναλωτικής συμπεριφοράς αποτελούν οι αναζητήσεις της Google σε πραγματικό χρόνο, καθώς η

Η διαχείριση μεγάλου όγκου δεδομένων στον Τουρισμό- Μια προσπάθεια βιβλιογραφικής και εμπειρικής προσέγγισης του θεματος

ανάλυση ταυτόχρονου συσχετισμού που προσφέρεται από την Google Correlate εξακολουθεί να είναι κατά έξι εβδομάδες μπροστά από τις αναφερόμενες τιμές (reported values). Αξιοσημείωτο παράδειγμα της χρήσης των αναζητήσεων της Google για άμεση πρόβλεψη αποτελεί η Google Flu Trends η οποία εντοπίζει πιθανές εκδηλώσεις γρίπης 1-2 εβδομάδες πριν από τις επίσημες υγειονομικές εκθέσεις από τις λέξεις κλειδιά σχετικές με την γρίπη που χρησιμοποιήθηκαν στη μηχανή αναζήτησης της.

Πολλές μελέτες χρησιμοποίησαν δομημένα δεδομένα μηχανών αναζήτησης για την άμεση πρόβλεψη (nowcasting) και την πρόβλεψη στον τουρισμό (Artola, Pinto και Pedraza, 2015, Bangwayo - Skeete και Skeete, 2015, Yang, Pan, Evans και Lv, 2015). Εκτός από τα δεδομένα μηχανών αναζήτησης υπάρχουν και άλλοι τύποι πραγματικών δεδομένων που μπορούν να συγκεντρωθούν και να αναλυθούν, όπως για παράδειγμα είναι τα δεδομένα που αντλούνται από τις αγορές πιστωτικών καρτών, η αποστολή και ο εντοπισμός δεμάτων και η χρήση κινητών τηλεφώνων. Πολλά από αυτά τα δεδομένα γίνονται διαθέσιμα σχεδόν σε πραγματικό χρόνο και τα οποία μπορούν να χρησιμοποιηθούν για την πρόβλεψη των μακροδεδομένων που θα συγκεντρωθούν κάποια στιγμή στο μέλλον (Jeng & Fesenmaier, 2002).

Ο απώτερος στόχος της χρήσης μεγάλων δεδομένων σε πραγματικό χρόνο είναι η ανάπτυξη εφαρμογών που θα είναι σε θέση να παράσχουν προτάσεις μόλις ληφθεί ο οικονομικός παλμός της επιχείρησης, αυτό βέβαια πρέπει να γίνεται υπό ελεγχόμενες συνθήκες και να υπάρχει η δυνατότητα διακοπής τους ανά πάσα στιγμή.

Τα Οφέλη της Χρήσης Μεγάλων Δεδομένων στις Τουριστικές Επιχειρήσεις

Η δυναμική της ανάλυσης Μεγάλων Δεδομένων αλλάζει όλες τις βιομηχανίες, όχι μόνο τις συνηθισμένες της λιανικής, της λογιστικής και της υψηλής τεχνολογίας. Αλλάζει επίσης και τον κόσμο της φιλοξενίας και ταξιδιού καθώς οι τουριστικές επιχειρήσεις έχουν να κάνουν με πληθώρα δεδομένων που καλύπτουν όλες τις πτυχές των πληροφοριών όπως είναι οι επιβεβαιώσεις πτήσεων ή οι προτιμήσεις δωματίου πελατών, δημιουργώντας έτσι δυνατότητες συσχετισμού δεδομένων για την εύρεση πληροφοριών που άλλοτε ήταν άγνωστες (Turner, 2014). Επιπλέον παρατηρούνται σημαντικές αλλαγές και στον τομέα των μεγάλων δεδομένων λόγω του φθίνοντος κόστους των συστημάτων ανάλυσης και της εξοικείωσης των υπαλλήλων με τις

Η διαχείριση μεγάλου όγκου δεδομένων στον Τουρισμό- Μια προσπάθεια βιβλιογραφικής και εμπειρικής προσέγγισης του θεματος

δυνατότητες των συστημάτων αυτών. Ουσιαστικά τα μεγάλα δεδομένα μπορούν να χρησιμοποιηθούν για την προσαρμογή εκστρατειών μάρκετινγκ στην πραγματικότητα της εκάστοτε αγοράς και στην εύρεση ανεπαρκειών στα επιχειρηματικά μοντέλα. Η ανάλυση μεγάλων δεδομένων μπορεί να προσφέρει κομβικής σημασίας επιχειρηματικές πληροφορίες και να αποτελέσει την πηγή καινοτομίας για τους τουριστικούς οργανισμούς και τη βιομηχανία εν γένει. Οι δυνατότητες των μεγάλων δεδομένων στον τουρισμό είναι τεράστιες και οι τουριστικοί οργανισμών δεν πρέπει να υποτιμούν τη σημασία τους (Pries και Dunnigan, 2015).

Με τη σωστή προσέγγιση ο τουριστικός κλάδος μπορεί να μάθει πολλά για τις προτιμήσεις των καταναλωτών και να χρησιμοποιεί αυτές τις πληροφορίες για την οικοδόμηση σχέσεων με τους μεμονωμένους τουρίστες. Η ικανότητά να προσφέρεται η σωστή υπηρεσία ή προϊόν τη σωστή στιγμή είναι καίριας σημασίας. Δίχως τις σωστές πληροφορίες και μια καλή στρατηγική στοχοποίησης η διαφήμιση δεν θα αποδίδει και δεν θα μπορεί να παράξει αξία. Ο τουρισμός είναι μια βιομηχανία που κινείται σε υψηλούς ρυθμούς και αυτό γεννά την ανάγκη ταχείας ανάλυσης δεδομένων και γρήγορων αποφάσεων. Στον τουρισμό η οποιαδήποτε ζήτηση πρέπει να ικανοποιείται άμεσα προκειμένου ο πάροχος του προϊόντος / υπηρεσίας να παραμείνει προσίτος στους πελάτες, πράγμα που καθιστά τα μεγάλα δεδομένα εξαιρετικά σημαντικά. Με την ραγδαία αύξηση του αριθμού των εφαρμογών των μεγάλων δεδομένων ανακύπτει ο κίνδυνος της αλληλεπίδρασης των παραδοσιακών τουριστικών δεδομένων με τα νέα. Για παράδειγμα τα τηλεφωνικά κέντρα θα αλληλεπιδρούν με τις ηλεκτρονικές κριτικές των πελατών, τα προγράμματα αφοσίωσης-ανταμοιβής θα συνδέεται με το ιστορικό κρατήσεων και οι προτιμήσεις διαμονής θα συνδυάζονται με τα τσατ των μέσων κοινωνικής δικτύωσης.

Η Συμπεριφορά των Καταναλωτών

Βρισκόμαστε σε μια εποχή πρωτοφανούς ρευστότητας καταναλωτικής συμπεριφοράς, προσδοκιών πελατών και επιχειρηματικών μοντέλων που δημιουργήθηκαν από τις ίδιες τεχνολογίες που διαταράσσουν την λειτουργία καταξιωμένων επιχειρήσεων και ταυτόχρονα δημιουργούν νέες (Marko, 2015, McAfee, Brynjolfsson, Davenport, Patil και Barton, 2012). Ωστόσο τα στοιχεία των τουριστικών

μεγάλων δεδομένων καταδεικνύουν σημαντικές αλλαγές στη σχέση μεταξύ των επιχειρήσεων και των πελατών τους. Έτσι τα μεγάλα δεδομένα μπορούν να χρησιμοποιηθούν για την παροχή ανώτερων υπηρεσιών αγοράς και υποστήριξης με σκοπό την ενίσχυση της επιλογής και των προσδοκιών των πελατών. Καταλυτικό ρόλο για την επανα-αναγνώριση της καταναλωτικής συμπεριφοράς με την χρήση μεγάλων δεδομένων, παίζει η εκτεταμένη χρήση κινητών συσκευών, εφαρμογών και άλλων μέσων κοινωνικής δικτύωσης, τα οποία διαδραματίζουν ολοένα και μεγαλύτερο ρόλο στη συλλογή ακατέργαστων δεδομένων.

Τα μεγάλα δεδομένα μπορούν να παρέχουν πολλές πληροφορίες σχετικά με τη καταναλωτική συμπεριφορά, ορισμένες από τις οποίες ήδη αποδίδουν ενώ άλλες δεν έχουν ανακαλυφθεί ακόμη. Οι δυνατότητες που δημιουργούνται από τη χρήση μεγάλων δεδομένων είναι μεγάλες ιδιαίτερα στο λιανικό εμπόριο καθώς οι βιομηχανίες και οι επιχειρησιακές διεργασίες μπορούν να εκμεταλλευτούν επιτυχώς τα νέα κανάλια επικοινωνίας και τις επιλογές παροχής υπηρεσιών (Marko, 2015). Η συλλογή, η συσχέτιση και η ανάλυση όλων των καναλιών των μεγάλων δεδομένων του τουρισμού από την αλληλεπίδραση με τους πελάτες είναι το κλειδί για τη μετατροπή της εμπειρίας των πελατών από εφιάλτη σε νιρβάνα (Chase, 2013). Ο σύνδεσμος μεταξύ των μεγάλων δεδομένων και της μηχανικής μάθησης σε όλες τις εκφάνσεις, συμπεριλαμβανομένης της προγνωστικής ανάλυσης όπως και των νευρωνικών δικτύων βαθιάς μάθησης, αποτελεί το θεμέλιο μιας καλά ενημερωμένης, άκρως αποτελεσματικής και ιδιαίτερα ικανοποιητικής αλληλεπίδρασης που ωφελεί τόσο τους πελάτες όσο και τις επιχειρήσεις.

Ο στόχος της χρήσης των τουριστικών μεγάλων δεδομένων είναι η δημιουργία μιας αυθεντικής συναισθηματικής σχέσης μεταξύ των πελατών και των συμμετεχόντων στην τουριστική βιομηχανία προκειμένου να βελτιωθεί σημαντικά ο τομέας της εξυπηρέτησης και υποστήριξης πελατών. Η εξερεύνηση των μεγάλων δεδομένων στον τουρισμό έχει τεράστιες επιπτώσεις και παρέχει ευκαιρίες για την απρόσκοπτη δικτύωση των καταναλωτικών εμπειριών κατά τη χρήση κινητών συσκευών, ιστοσελίδων και επικοινωνίας μέσω διάφορων καναλιών (τηλέφωνο, μηνύματα, ηλεκτρονικό ταχυδρομείο, τσατ και τα μέσα κοινωνικής δικτύωσης). Το κλειδί για την χρήση των τουριστικών μεγάλων δεδομένων δεν είναι μόνο η παροχή ολοκληρωμένων

Η διαχείριση μεγάλου όγκου δεδομένων στον Τουρισμό- Μια προσπάθεια βιβλιογραφικής και εμπειρικής προσέγγισης του θεματος

υπηρεσιών αλλά η προορατικότητα. Πρέπει να μπορούμε να προβλέψουμε τις ανάγκες των πελατών και να προλαμβάνουμε τα προβλήματα. Με άλλα λόγια, μπορούμε να προβλέψουμε προβλήματα και απορίες με τη χρήση στατιστικών μοντέλων πρόβλεψης πριν μας γίνει η ίδια ερώτηση ή μας ζητηθεί να εξηγήσουμε επανειλημμένα την ίδια κατάσταση.

Μηχανισμοί Ανατροφοδότησης

Η ανατροφοδότηση ή η κριτική στην τουριστική βιομηχανία είναι σημαντική για τον εντοπισμό των προτιμήσεων των πελατών και για την μετέπειτα παροχή θετικών εμπειριών. Η επεξεργασία κριτικών είναι ένα από τα πιο σημαντικά στοιχεία για την επίτευξη ραγδαίας ανάπτυξης της επιχείρησης και για την οικοδόμηση μιας στρατηγικής που θα καλύπτει καλύτερα της ανάγκες των πελατών. Ανατροφοδότηση βάσει τουριστικών μεγάλων δεδομένων από πελάτες, υπαλλήλους, συνεργάτες, προμηθευτές και κοινότητες έχει επίσης βελτιώσει τις δυνατότητες της ανάλυσης μεγάλων δεδομένων. Οι επιχειρηματικές και καταναλωτικές εφαρμογές που βασίζονται στα δεδομένα αποτελούν τις πιο διαδεδομένες μεθόδους συλλογής κριτικών οποτεδήποτε και οπουδήποτε. Ολοένα και περισσότερες υπηρεσίες νεφών (cloud) μας δίνουν την άμεση και καθολική ικανότητα αλληλεπίδρασης μέσω έξυπνων τηλεφώνων, τάμπλετς ή ακόμα και ρολογιών.

Η αυξημένη χρήση σύγχρονων τεχνικών για την συλλογή κριτικών έχει οδηγήσει το παραδοσιακό μάρκετινγκ κριτικών να αντικατασταθεί σταδιακά από εμπορικά μηνύματα που είναι άμεσα, κοφτά, μοναδικά και προσαρμοσμένα στον κάθε πελάτη. Μία από τις εφαρμογές των μηχανισμών ανατροφοδότησης που εφαρμόζουν οι τουριστικοί οργανισμοί είναι αυτή της προσαρμογής τιμών, κατά την οποία μια μεταβολή στην ζήτηση ταξιδιών που προκύπτει από προγνωστική ανάλυση μεγάλων δεδομένων μπορεί να παρέχει χρήσιμες πληροφορίες για την γρήγορη και αποτελεσματική προσαρμογή των τιμών.

Η μηχανική μάθηση είναι μία από τις βασικές μεθόδους που χρησιμοποιείται στον τουριστικό κλάδο για την κατασκευή του μηχανισμού ανατροφοδότησης μεταξύ πελατών και τουριστικών πρακτορείων (Bajari, Nekipelov, Ryan, και Yang, 2015). Παραδείγματος χάρη, μέσω συνέργειας μεταξύ τουριστικών οργανισμών, πιστωτικών

ιδρυμάτων και οργανισμών τηλεπικοινωνιών, η μηχανική μάθηση μπορεί να προσδιορίσει εάν ένα άτομο έχει αλλάξει την διεύθυνση κατοικίας του ή ταξιδεύει διεθνώς, μέσω ελέγχου για ασυνήθιστες χρεώσεις. Η χρήση της μηχανικής μάθησης μεγάλων δεδομένων σε εμπειρίες πελατών δίνει την δυνατότητα στις τουριστικές επιχειρήσεις και τους ταξιδιωτικούς πράκτορες να προωθούν προδραστικά προσφορές μέσω μηνυμάτων ή τηλεφωνικών κλήσεων. Πιο συγκεκριμένα, η μηχανική μάθηση θα μπορούσε να τροποποιήσει το σύστημα ανάδρασης προσδιορίζοντας παλαιότερες κινήσεις του χρήστη και υπολογίζοντας το ποσοστό επιτυχίας τους. Με την χρήση αυτών των πληροφοριών οι τουριστικοί οργανισμοί θα μπορούσαν να παρέχουν λύσεις για την αναποτελεσματικότητα κάποιων διεργασιών, και την ενόχληση των πελατών.

Οι προγνωστικές αναλύσεις παρουσιάζονται συχνά ως πανάκεια για τις επιχειρήσεις και μπορούν όντως να φανούν εξαιρετικά αποτελεσματικές. Η προγνωστική ανάλυση στα τουριστικά μεγάλα δεδομένα που χρησιμοποιείται στους σύγχρονους μηχανισμούς ανατροφοδότησης είναι σημαντική βελτίωση από τους αρχαιότερους τρόπους συλλογής πληροφοριών. Οι προγνωστική ανάλυση μπορεί να δώσει στα διευθυντικά στελέχη του μάρκετινγκ περισσότερες πληροφορίες για τις προτιμήσεις των πελατών, κάτι που μπορεί να τους βοηθήσει να κατανοήσουν καλύτερα τους πελάτες και να οδηγήσει στην αύξηση των πωλήσεων (LaValle, Lesser, Shockley, Hopkins και Kruschwitz, 2011). Ωστόσο εδώ αξίζει να σημειωθεί ότι η επιτυχία των προγνωστικών αναλύσεων εξαρτάται τόσο από την ποιότητα των μεγάλων δεδομένων όσο και από τους μηχανισμούς ανατροφοδότησης των πελατών.

Οι μηχανισμοί ανατροφοδότησης των πελατών πρέπει να είναι καλά σχεδιασμένοι και κατανοητοί ώστε αποδίδουν έγκαιρα χρήσιμα δεδομένα. Τα έγκαιρα και αξιόπιστα τουριστικά μεγάλα δεδομένα μπορεί να παρέχουν πλήρη εικόνα των πελατών και των δυνητικών πελατών που ακολούθως μεταφράζονται σε υψηλής ακρίβειας επενδύσεις μάρκετινγκ σε πιο παραγωγικά κανάλια.

Η χρήση των Μεγάλων Δεδομένων στην Πρόγνωση του Τουρισμού

Στα παρακάτω κεφάλαια θα αναλύσουμε τα κύρια βήματα για την πρόγνωση του τουρισμού με τη χρήση μεγάλων δεδομένων γνωρίζοντας ότι προσφέρουν πολλαπλά οφέλη στην τουριστική βιομηχανία.

Συλλογή Μεγάλων Δεδομένων για την Τουριστική Πρόγνωση

Οι εταιρίες που καταφέρνουν να σχεδιάζουν και να υλοποιούν αποτελεσματικά στρατηγικές μεγάλων δεδομένων αποκτούν ανταγωνιστικό πλεονέκτημα καθώς η τεχνολογία που απαιτείται για την διαχείριση τους αποτελεί εμπόδιο για πολλές επιχειρήσεις λόγω της πολυπλοκότητας και του κόστους. Προτού φτάσουμε στην χρήση των μεγάλων δεδομένων, θα αναλύσουμε τα τρία βήματα της διαδικασίας συλλογής τους.

1. Ο Σκοπός

Το πρώτο βήμα της συλλογής μεγάλων δεδομένων είναι ο σκοπός της χρήσης τους, ο οποίος πρέπει να πηγάζει από τα οφέλη και τα κέρδη της επιχείρησης (Pellegrini, 2013). Όταν συλλέγουμε μεγάλα δεδομένα είναι σημαντικό να έχουμε πρόσβαση σε αυτά και να βρισκόμαστε στη θέση να γνωρίζουμε τι είναι διαθέσιμο ώστε να μπορούμε να καθορίσουμε πού την αξία για την επιχείρηση. Με άλλα λόγια, θα πρέπει να γνωρίζουμε τις δυνατότητες των μεγάλων δεδομένων και ακριβώς αυτό που ψάχνουμε ώστε να αναγνωρίζουμε τις αξίες τους. Είναι σημαντικό να θέτουμε συγκεκριμένους στόχους πριν ξεκινήσουμε τη διαδικασία συλλογής και επεξεργασίας μεγάλων δεδομένων.

2. Η Απεικόνιση των Μεγάλων Δεδομένων

Το δεύτερο βήμα ορίζεται ως η απεικόνιση της εξόρυξης μεγάλων δεδομένων, το οποίο στην πράξη σημαίνει ότι τα δεδομένα θα πρέπει να τεθούν ορατά στους χρήστες της εταιρίας / οργανισμού. Αυτό θα επιτρέψει στα τμήματα πρόβλεψης ζήτησης να προσδιορίσουν τη βέλτιστη ποσότητα προϊόντος που απαιτείται ώστε να προσαρμόσουν τις διαδικασίες υλικοτεχνικής υποστήριξης για την μεγιστοποίηση της αποτελεσματικότητας (Weiler και Black, 2014). Ο σκοπός της απεικόνισης των δεδομένων είναι η εύρεση τρόπων με τους οποίους τα δεδομένα θα μπορούσαν να συλλεχθούν αποτελεσματικά από διάφορες πηγές και να παρουσιαστούν με κατανοητό τρόπο στους χρήστες. Επίσης αυτό θα βοηθήσει τα τμήματα πρόβλεψης να χρησιμοποιήσουν αποτελεσματικότερα τα μεγάλα δεδομένα για την εκπλήρωση των προγνώσεων τους.

3. Η Δόμηση των Μεγάλων Δεδομένων

Τρίτο βήμα αποτελεί η δόμηση των μη δομημένων δεδομένων. Πρακτικά αυτό σημαίνει ότι τα μεγάλα δεδομένα πρέπει να οργανωθούν σύμφωνα με συγκεκριμένα πρότυπα μεγέθους και μορφής ώστε να μπορούν να εισαχθούν σε υπολογιστικά φύλλα. Τα δομημένα δεδομένα συνήθως βρίσκονται σε σχεσιακές βάσεις δεδομένων με αποτέλεσμα να ονομάζονται σχεσιακά δεδομένα (Akerkar, 2013). Τα μη δομημένα δεδομένα μπορούν να χαρτογραφηθούν εύκολα σε προκαθορισμένα πεδία. Για παράδειγμα, τα δομημένα δεδομένα ενός τηλεφωνικού κέντρου αποτελούνται από αριθμούς, ημερομηνίες και ομάδες λέξεων και αριθμών που ονομάζονται ακολουθίες. Γενικά είναι αποδεκτό αυτού του είδους τα δεδομένα να αντιπροσωπεύουν περίπου το 20% του συνολικού όγκου των μεγάλων δεδομένων. Τα μη δομημένα δεδομένα είναι πολύ δύσκολο να αναλυθούν δεδομένου ότι αποτελούν το μεγαλύτερο μέρος των μεγάλων δεδομένων και περιέχουν πληθώρα πολύτιμων πληροφοριών που δεν είναι συμβατά με τα προκαθορισμένα συστήματα δεδομένων. Για το λόγο αυτό έχουν δημιουργηθεί διαφορετικά λογισμικά τα οποία αναζητούν μη δομημένα δεδομένα και εξαγάγουν από αυτά σημαντικές πληροφορίες.

Η Επιλογή και η Συρρίκνωση των Μεγάλων Δεδομένων

Τα μεγάλα δεδομένα περιέχουν πολλές πληροφορίες γεγονός που δημιουργεί όχι μόνο προβλήματα αποθήκευσης αλλά και τεράστιο πρόβλημα ανάλυσης. Η εύρεση χρησιμότητας για αυτές τις ογκώδεις βάσεις δεδομένων αποτελεί το μεγαλύτερο πρόβλημα στην πρόβλεψη του τουρισμού χρησιμοποιώντας δομημένα μεγάλα δεδομένα. Οι δύο πιο δημοφιλείς μέθοδοι που χρησιμοποιούνται για την επιλογή και συρρίκνωση μεγάλων όγκων δομημένων δεδομένων είναι το παραγοντικό μοντέλο (factor model) και η μέθοδος LASSO (least absolute shrinkage and selection operator).

1. Το Παραγοντικό Μοντέλο (factor model)

Το παραγοντικό μοντέλο είναι η πιο συνηθέστερα χρησιμοποιούμενη μέθοδος για την επιλογή και συρρίκνωση δομημένων μεγάλων δεδομένων. Όπως επισημαίνουν ορισμένες πρόσφατες μελέτες στον τομέα των οικονομικών τα παραγοντικά μοντέλα

Η διαχείριση μεγάλου όγκου δεδομένων στον Τουρισμό- Μια προσπάθεια βιβλιογραφικής και εμπειρικής προσέγγισης του θεματος

αποτελούν χρήσιμα εργαλεία στην πρόγνωση μεγάλων συνόλων δεδομένων. Επιπροσθέτως, η χρήση της ανάλυσης ανεξάρτητων συνιστωσών (ICA) και της αραιής ανάλυσης κύριων συνιστωσών (SPCA) σε συνδυασμό με πληθώρα άλλων μεθόδων εκτίμησης και συρρίκνωσης δεδομένων μεταξύ άλλων και των τεχνικών bagging, boosting και elastic net προσφέρει σημαντικά προγνωστικά οφέλη στις επιχειρήσεις που επιλέγουν να τις υιοθετήσουν.

2. Η Μέθοδος LASSO

Η μέθοδος LASSO είναι μια μέθοδος επιλογής συμμεταβλητών στο πλαίσιο γραμμικής παλινδρόμησης (Tibshirani, 1996). Ο τρόπος λειτουργίας της τεχνικής αυτής είναι μέσω της επιβολής ποινής στο πρόβλημα βελτιστοποίησης που σχετίζεται με την παλινδρόμηση υπό τον όρο της L1-νόρμας των συντελεστών. Ανήκει στην οικογένεια τεχνικών συρρίκνωσης μοντέλων μέσω ποινών κατά τα οποία εκτελείται η μέθοδος ελαχίστων τετραγώνων με ορισμένους περιορισμούς στους συντελεστές και συγκεκριμένα την L1-νόρμα στην περίπτωση της LASSO. Σύμφωνα με την βιβλιογραφία η μέθοδος LASSO τείνει να έχει χαμηλότερο ποσοστό κίβδηλων αποτελεσμάτων σε σχέση με τις συνήθεις τεχνικές πρόβλεψης (Ng, 2012). Η μέθοδος LARS μπορεί να συνδυαστεί με το παραγοντικό μοντέλο για τη συρρίκνωση μεγάλων όγκων δεδομένων και για την πρόβλεψη οικονομικών σειρών.

Πλαίσιο για την Τουριστική Πρόβλεψη Ζήτησης με την Χρήση Μεγάλων Δεδομένων

Μια ευρέως διαδεδομένη αντίληψη που επικρατεί είναι ότι τα μεγάλα δεδομένα μπορούν να βοηθήσουν στη βελτίωση των προβλέψεων υπό την προϋπόθεση ότι τα κρυμμένα μοτίβα και οι τάσεις μέσα στον όγκο πληροφοριών που έχουν συλλεχθεί μπορούν να ανακαλυφθούν και να αναλυθούν και ότι οι προβλέψεις μπορούν να βελτιωθούν μέσω της λήψης αποφάσεων με την χρήση δεδομένων. Ορισμένοι ερευνητές υποστηρίζουν ότι οι μέθοδοι εξόρυξης δεδομένων μπορούν να παρέχουν βοήθεια στις διαδικασίες πρόγνωσης των μεγάλων δεδομένων. Ωστόσο οι τεχνικές εξόρυξης δεδομένων χρησιμοποιούν πάντοτε στατικά δεδομένα αντί χρονοσειρών και σπάνια χρησιμοποιούνται στην πρόβλεψη ζήτησης τουρισμού. Όταν στρεφόμεστε στις

παραδοσιακές μεθόδους πρόβλεψης της τουριστικής ζήτησης μεγάλων δεδομένων, το μεγαλύτερο πρόβλημα που παρουσιάζεται είναι ότι τα παραδοσιακά εργαλεία πρόβλεψης δεν έχουν τη δυνατότητα να επεξεργαστούν το μέγεθος, την ταχύτητα και την πολυπλοκότητα των μεγάλων δεδομένων ακόμα κι αν αυτά είναι στη δομημένη τους μορφή.

Προκειμένου να εφαρμόσουμε μια παραδοσιακή μέθοδο πρόβλεψης στα μεγάλα δεδομένα πρέπει πρώτα να απλοποιήσουμε τα δομημένα δεδομένα. Μία λύση είναι η συρρίκνωση των μεγάλων δεδομένων και στη συνέχεια η λήψη των σημαντικότερων πληροφοριών σε μορφή που να μπορούν να εφαρμοστούν εύκολα στο παραδοσιακό μοντέλο. Τα παραγοντικά μοντέλα είναι από τις πιο διαδεδομένες και δημοφιλείς τεχνικές στατιστικής και εξόρυξης δεδομένων για προβλέψεις μεγάλων δεδομένων.

1. Μοντέλο Μικτής Συχνότητας στα Μεγάλα Δεδομένα

Η έρευνα για την πρόβλεψη μεγάλων δεδομένων στον τουρισμό έχει αποδώσει καρπούς, συγκεκριμένα δε το εργαλείο της Google Trends έχει αποδειχθεί εξαιρετικά αποτελεσματικό στην πρόβλεψη αφίξεων σε συγκεκριμένους προορισμούς. Για παράδειγμα, η πρόβλεψη επισκεψιμότητας ενός αξιοθέατου μέσω παρακολούθησης της χρήσης συγκεκριμένων λέξεων κλειδιών σε μηχανές αναζήτησης δείχνει σημαντική μείωση σφαλμάτων πρόβλεψης και χαμηλό ποσοστό αποκλίσεων. Ωστόσο αυτές οι μελέτες συγκεντρώνουν ή αγνοούν τις εβδομαδιαίες παρατηρήσεις ώστε να καταστήσουν τα δεδομένα συμβατά με τα παραδοσιακά εργαλεία πρόβλεψης. Για παράδειγμα, προκειμένου να προβλεφθεί ο συνολικός αριθμός επισκεπτών κάποιου αξιοθέατου σε έναν μήνα μπορεί να χρησιμοποιηθούν δεδομένα από τις δύο πρώτες εβδομάδες του μήνα, απορρίπτοντας πληροφορίες για τις τελευταίες δύο εβδομάδες. Πράγματι, αυτές οι έρευνες θα μπορούσαν να βελτιωθούν με την χρήση καινοτόμου μεθόδου πρόβλεψης όπως η mixed-data sampling (MIDAS) και κατ' αυτόν τον τρόπο να αξιοποιήσουν πλήρως τις δυνατότητες των δεδομένων που παράγουν οι μηχανές αναζήτησης υψηλής συχνότητας. Άλλο μοντέλο μικτής συχνότητας που μπορεί να εκπληρώσει την ίδια εργασία είναι το μοντέλο μικτής συχνότητας VAR, η οποία χειρίζεται δεδομένα χαμηλής συχνότητας ως υψηλής συχνότητας και στη συνέχεια τα επεξεργάζεται με την χρήση του μοντέλου state-space.

2. Παραγοντικό μοντέλο και συνδυασμοί πρόβλεψης

Είναι γεγονός ότι ο καλύτερος τρόπος για την πρόβλεψη σειρών χαμηλών συχνοτήτων (όπως η τουριστική ζήτηση) χρησιμοποιώντας δεδομένα υψηλών συχνοτήτων είναι με τον συνδυασμό μοντέλων συρρίκνωσης με τα μοντέλα μικτών συχνοτήτων. Ορισμένες μελέτες πέτυχαν καλύτερα αποτελέσματα με την χρήση μοντέλου μικτών συχνοτήτων με συντελεστές τα δεδομένα υψηλών συχνοτήτων για την πρόβλεψη των μακροοικονομικών δεικτών. Σύμφωνα με την υπάρχουσα βιβλιογραφία ο συνδυασμός μεθόδων πρόβλεψης μπορεί να βελτιώσει την ακρίβεια των αποτελεσμάτων.

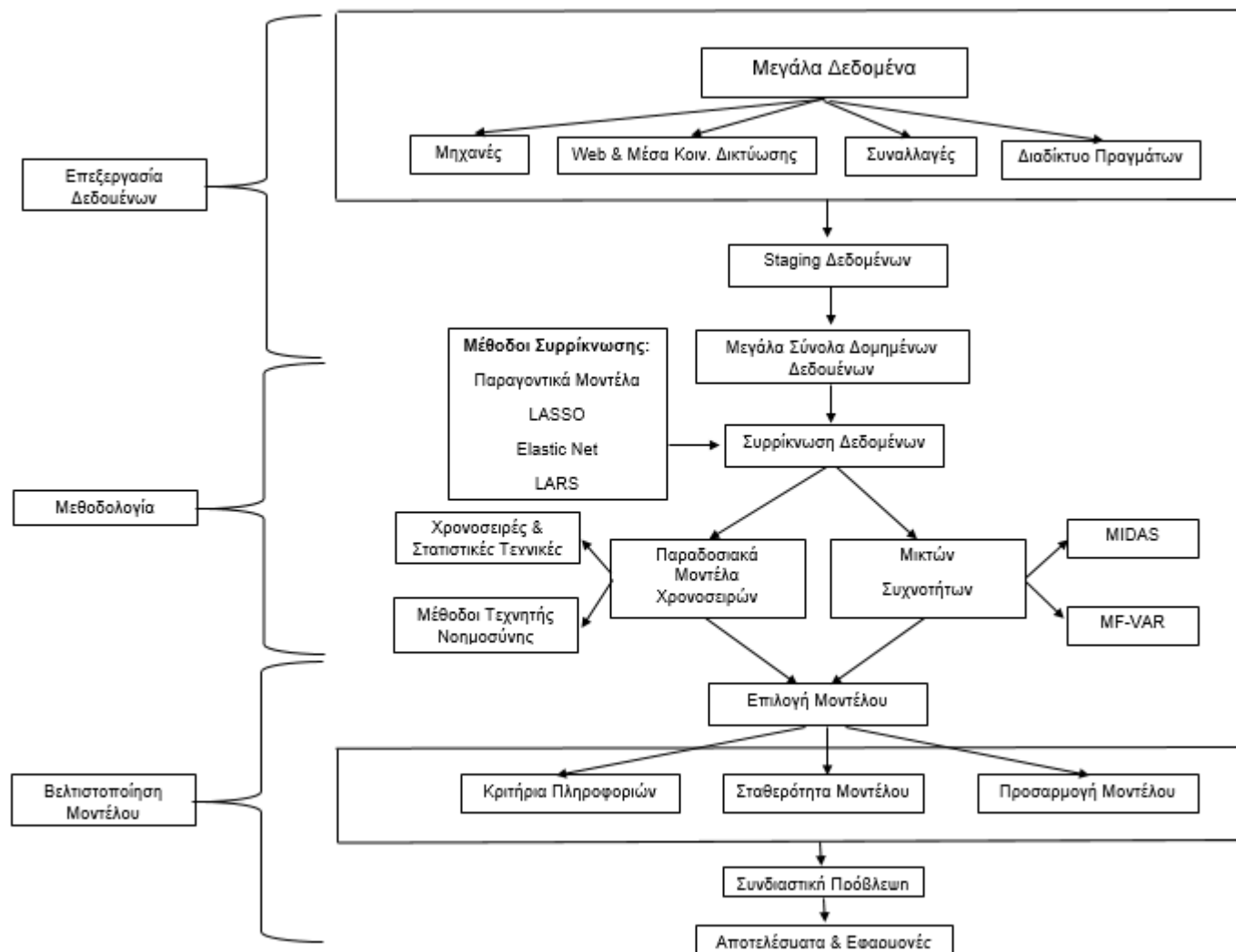
Το παρακάτω σχήμα (Σχήμα 2) δείχνει το πλαίσιο της πρόβλεψης στον τουρισμό με τη χρήση μεγάλων δεδομένων. Όπως φαίνεται, η διαδικασία αυτή περιλαμβάνει τρία βασικά βήματα: 1) την εξερεύνηση δεδομένων που είναι η επεξεργασία και εξαγωγή δεδομένων σε κατάλληλη μορφή για την εισαγωγή στο μοντέλο, 2) η χρήση τεχνικών μοντελοποίησης για την πρόβλεψη της καταναλωτικής συμπεριφοράς βάσει προηγούμενων συναλλαγών και προτιμήσεων, 3) η βελτιστοποίηση των αποτελεσμάτων πρόβλεψης και η μείωση του ρίσκου σφάλματος μέσω της επιλογής κατάλληλου μοντέλου και συνδυασμού μεθόδων προβλέψεως.

Ο Ψηφιακός Τουρισμός

Εισαγωγή

Η συνεχής εξέλιξη της τεχνολογίας της πληροφορίας και της επικοινωνίας (ΤΠΕ) έχει μεταμορφώσει τον τουρισμό με πολλούς τρόπους. Συγκεκριμένα τα smartphones και οι σχετικές με αυτά εφαρμογές επέκτειναν το πεδίο της τουριστικής εμπειρίας επιτρέποντας στους ταξιδιώτες να επικοινωνούν και να μοιράζονται τις εμπειρίες τους με τις οικογένειες και τους φίλους τους όπου και όποτε θέλουν. Παράλληλα με αυτές τις

Σχήμα 2. Το πλαίσιο της τουριστικής πρόβλεψης με τη βοήθεια μεγάλων δεδομένων



εξελίξεις, φορητές συσκευές (Google Glass, Apple iWatch, βραχιόλια άθλησης κλπ.) χρησιμοποιούνται ευρέως λόγω της φορητότητάς τους και της δυνητικής χρήσης τους στα ταξίδια. Αυτές οι τεχνολογίες μας κατευθύνουν προς μια ψηφιοποιημένη κοινωνία, όπου ο κάθε χρήστης ανεβάζει μεγάλο όγκο δεδομένων στη διάρκεια της ημέρας, δημιουργώντας ευκαιρίες για επιχειρηματική δράση.

Μια χρήση του όρου ψηφιακός-εαυτός χρησιμοποιήθηκε για να περιγράψει τη βελτίωση της ποιότητας ζωής μέσα από την αυτογνωσία και την ανακάλυψη, όπου οι φορητές συσκευές παρακολουθούν την καθημερινότητά των χρηστών και αποθηκεύουν τα συγκεντρωμένα δεδομένα ώστε ενδεχομένως να μοιραστούν τις πληροφορίες αυτές με άλλους άλλους χρήστες με κοινά ενδιαφέροντα. Με την ευρεία χρήση των φορητών συσκευών στον τουρισμό, δεν θα ήταν καθόλου απίθανο το ενδεχόμενο μιας “επανατοποθέτησης” και “επέκτασης” της χρήσης των πληροφοριών με τέτοιο τρόπο ώστε να παρέχουν περισσότερες λεπτομέρειες για τον ταξιδιώτη αναφορικά με την κατάσταση υγείας του, την τήρηση της διατροφής του, πιθανές αλλαγές σχεδίων, την ποιότητα ύπνου, της συναισθηματικής του κατάστασης, την επικοινωνία με φίλους και συγγενείς και ακόμα την αίσθηση ότι βρίσκεται στο σπίτι του. Βάσει των παραπάνω, στο παρόν κεφάλαιο θα συζητήσουμε την έννοια του ψηφιακού-εαυτού ως προς τον τρόπο που οι σύγχρονες τεχνολογίες συνδέουν την καθημερινότητα με την τουριστική εμπειρία και προσφέρουν τρόπους για την κατανόηση αυτής της τεχνολογίας στο πλαίσιο του ταξιδιού.

Η Ανάδυση του Ψηφιακού Ταξιδιώτη και οι Φορητές Συσκευές

Μια νέα αναδυόμενη τάση είναι αυτή της χρήσης συσκευών αυτο-εντοπισμού, καταγραφής βιωμάτων, προσωπικής ανάλυσης και πληροφόρησης. Η έννοια της αυτο-παρακολούθησης και του αυτο-εντοπισμού έχει μια αρκετά μακρά ιστορία που μπορεί να εντοπιστεί στη δεκαετία του 1970. Από τότε, η θεώρηση της αυτο-παρακολούθησης έχει αποδειχθεί αποτελεσματική στη μεταβολή του χαρακτήρα και της συμπεριφοράς των ανθρώπων, που επίσης αποτελεί στόχο των λειτουργιών των τεχνολογιών ανίχνευσης. Στόχος αυτού του κινήματος είναι η αυτογνωσία μέσω της παρακολούθησης των βιωμάτων του χρήστη για την αυτοβελτίωση του. Οι χρήστες

αυτών των συσκευών έχουν εντοπίσει πολλά οφέλη, μεταξύ των οποίων η απόκτηση χρήσιμων πληροφοριών, η παρακολούθηση και η δοκιμασία του εαυτού τους σε νέες προκλήσεις και τέλος η λήψη ανάδρασης μέσω της σύγκρισης των αποτελεσμάτων τους με τα αποτελέσματα άλλων χρηστών που μοιράζονται κοινά ενδιαφέροντα.

Κατηγορίες	Παράδειγμα Πιθανού Μέτρου
Φυσική κατάσταση & δραστηριότητες	Κίνηση, θερμοκρασία, θερμίδες
Ψυχολογικές & πνευματικές καταστάσεις & χαρακτηριστικά	Διάθεση, συναισθήματα, αυτοεκτίμηση, εστίαση, προσοχή, μνήμη, άγχος, ένταση
Περιβαλλοντικές & περιστασιακές μεταβλητές	Τοποθεσία, καιρός, θόρυβος, ρύπανση, περιβάλλον, ώρα της ημέρας, ταξίδια, χρονικά διαστήματα
Κοινωνικές μεταβλητές	Επιρροή, εμπιστοσύνη, αλληλεπιδράσεις, παρέα, αντιλαμβανόμενη ασφάλεια

Πίνακας 1 Κατηγορίες ψηφιακού-εαυτού και μετρήσεις

Το σημαντικότερο είναι ότι η ανάπτυξη φορητών συσκευών όπως τα βραχιόλια άθλησης και τα έξυπνα ρολόγια, τα οποία έχουν πρόσβαση στο διαδίκτυο και στο υπολογιστικό νέφος (cloud computing), έχουν αλλάξει εντελώς τον τρόπο με τον οποίο οι άνθρωποι παρακολουθούν τη ζωή τους. Η τάση της ψηφιοποίησης έχει επεκταθεί σε διάφορους τομείς (υγεία, αθλητισμός κλπ) και παρέχει διάφορων ειδών πληροφορίες για τις ζωές των χρηστών. Όπως φαίνεται στον Πίνακα 1, σε ορισμένες περιπτώσεις είναι απαραίτητος ο κόπος και ο τρόπος από πλευράς χρηστών για την καταγραφή των δραστηριοτήτων και των συναισθημάτων τους (απόσταση που καλύφθηκε, ευεξία, ευτυχία, πρόσληψη θερμίδων και κάψιμο θερμίδων). Εντούτοις, υπάρχουν διαφορετικοί τρόποι που προσφέρουν τη δυνατότητα καταμέτρησης / παρακολούθησής των χρηστών σχεδόν άκοπα. Αυτές οι έξυπνες συσκευές έχουν τη δυνατότητα να καταγράφουν το περιβάλλον απαραίτητα, να παρακολουθούν τις συμπεριφορές των χρηστών σε

πραγματικό χρόνο και αλληλεπιδρώντας μεταξύ τους να κατανοούν την εκάστοτε κατάσταση του χρήστη, πχ. οι οδηγοί θα μπορούσαν να προειδοποιούνται εάν εμφανίζουν συμπτώματα υπνηλίας. Ενώ οι πρακτικές αυτο-παρακολούθησης συνεπάγονται τη συνεχή καταγραφή των καθημερινών δραστηριοτήτων, η μορφή των αποθηκευμένων δεδομένων δεν αποτελεί περιοριστικό παράγοντα καθώς μπορεί να είναι σε μορφή εικόνας, βίντεο, και ήχου. Κατά συνέπεια, ανάλογα με τη συσκευή και τη μέθοδο συλλογής πληροφοριών, τα δεδομένα που συλλέγονται μπορεί να είναι δομημένα ή μη δομημένα.

Γενικά, οι φορητές συσκευές δίνουν τη δυνατότητα σύνδεσης στο διαδίκτυο, με άλλες συσκευές και με το εξωτερικό περιβάλλον μέσω ψηφιακών αισθητήρων. Ορισμένες από αυτές τις συσκευές μπορούν να ανταλλάξουν τις αποθηκευμένες πληροφορίες ασύρματα, μέσω επικοινωνίας κοντινού πεδίου και iBeacon, ώστε να διευκολύνουν τους χρήστες, να αποκτήσουν καλύτερη αντίληψη του περιβάλλοντος χώρου και προκειμένου να τους ενθαρρύνουν να αλλάξουν ορισμένες συμπεριφορές. Παραδείγματος χάρη, τα έξυπνα παπούτσια μπορούν να δονούνται προκειμένου να υποδείξουν στον χρήστη τη σωστή κατεύθυνση ώστε να ακολουθήσει διαδρομή με ευχάριστο τοπίο. τα έξυπνα θερμομέτρα που είναι ενσωματωμένα στο ρούχα μπορούν να ανταλλάσσουν πληροφορίες με άλλες συσκευές για να ρυθμίζουν τη θερμοκρασία του δωματίου, αντίστοιχοι αισθητήρες που είναι ενσωματωμένοι στην κουβέρτα παρακολουθώντας τον ύπνο θα μπορούν να εκτιμήσουν το χρονικό διάστημα και το ρυθμό του βαθύ ύπνου.

Πολλοί άνθρωποι χρησιμοποιούν συστηματικά “αισθητήρες” για να συλλέγουν δεδομένα σε πραγματικό χρόνο για το περιβάλλον τους καθώς και για τη φυσική και συναισθηματική τους κατάσταση, παράγοντας τεράστιους όγκους δεδομένων. Για παράδειγμα, πολλοί λάτρεις των υπαίθριων δραστηριοτήτων συλλέγουν και μοιράζονται πληροφορίες σχετικά με τα πτηνά, συλλέγουν ανελλιπώς μετεωρολογικές πληροφορίες για τη δημοσίευση τους σε τοπικό επίπεδο ή παρακολουθούν τον ουρανό για παρουσία τυχόν νέων φαινομένων. Στο πλαίσιο του τουρισμού, οι υπεύθυνοι ενός θεματικού πάρκου μπορούν πλέον εύκολα να παρακολουθούν τη ροή των επισκεπτών σε

συγκεκριμένες ώρες κατά τη διάρκεια της ημέρας μέσω των δεδομένων τοποθεσίας των χρηστών από τα κινητά τους ή μέσω των εισιτηρίων με ενσωματωμένα RFID. Το σημαντικότερο είναι ότι οι τεχνολογίες αυτές προσφέρουν μεγάλο όγκο ψηφιακών “αποτυπωμάτων” βοηθώντας τους υπεύθυνους του τουριστικού προορισμού να παρακολουθούν αυτές τις πληροφορίες προκειμένου να αποκτούν πιο ολοκληρωμένη εικόνα για τον κάθε επισκέπτη καθώς δραστηριοποιείται (και κάνει επιλογές) στον χώρο. Ως εκ τούτου οι νέες αυτές τεχνολογίες ενισχύουν και τους τουρίστες και τους τουριστικούς οργανισμούς με το να συνδέουν τον πραγματικό με τον ψηφιακό κόσμο.

Ο Ψηφιακός Ταξιδιώτης και η Συναίσθηση του Περιβάλλοντος

Καθώς οι ταξιδιώτες κινούνται από το ένα μέρος (ή δραστηριότητα) στο άλλο κατά τη διάρκεια του ταξιδιού τους, οι μεταβαλλόμενες συνθήκες και το περιβάλλον τους μπορούν να προκαλέσουν αλλαγές στη λήψη αποφάσεων και συμπεριφορών τους. Για παράδειγμα, οι ταξιδιώτες συχνά επαναδιαπραγματεύονται τις λεπτομέρειες ενός ταξιδιού όταν η πτήση τους έχει καθυστέρηση για πολλές ώρες, λόγω σωματικής κόπωσης μπορεί να αναβάλλουν το δείπνο, τον περίπατο σε ένα πάρκο ή την επίσκεψή τους σε ένα μουσείο. Οι μεταβολές στο περιβάλλον και στη μεταγενέστερη συμπεριφορά μπορούν να αλλάξουν τον τρόπο με τον οποίο οι ταξιδιώτες αλληλεπιδρούν και βιώνουν την εμπειρία του προορισμού τους. Ως εκ τούτου, οι φορητές συσκευές μας επιτρέπουν να παρακολουθούμε όχι μόνο τις συμπεριφορές από τις πληροφορίες που μας παρέχουν, αλλά επίσης μας επιτρέπουν να προβλέψουμε με ακρίβεια τι σκέφτονται και πώς νιώθουν την κάθε στιγμή. Συναφώς το περιβάλλον είναι μια θεμελιώδης πτυχή της τουριστικής εμπειρίας και η γνώση του περιβάλλοντος των ταξιδιωτών αποτελεί βάση για το σχεδιασμό και την ανάπτυξη του τουρισμού. Από την οπτική γωνία του τουριστικού προορισμού η κατανόηση του περιβάλλοντος δίνει τη δυνατότητα να επηρεάζει τις αποφάσεις των ταξιδιωτών σε πραγματικό χρόνο. Επομένως, παρουσιάζονται πολλές ευκαιρίες για τους τουριστικούς προορισμούς να συλλέξουν, να κατανοήσουν και να ερμηνεύσουν τις πληροφορίες του περιβάλλοντος που παράγονται από τις φορητές συσκευές των συνδεδεμένων στο διαδίκτυο χρηστών.

Ο Ψηφιακός Ταξιδιώτης στην Καθημερινότητα

Τα δεδομένα που δημιουργούνται καθημερινά επηρεάζουν σε μεγάλο βαθμό την ταξιδιωτική συμπεριφορά και κατά συνέπεια ενισχύουν την ταξιδιωτική εμπειρία. Τα τελευταία χρόνια αρκετές έρευνες της τουριστικής βιβλιογραφίας έχουν δείξει ότι ο τεχνολογικοί παράγοντες (smartphones, φορητές συσκευές) αλλάζουν δραστικά τον τρόπο με τον οποίο ταξιδεύει ο κόσμος σε όλα του τα στάδια, συνδέοντας το ταξίδι με την καθημερινότητα. Επιπλέον, υποστηρίζεται έντονα η αντίληψη ότι η τουριστική εμπειρία δεν διαχωρίζεται πλέον από την καθημερινή ζωή. Αν και το βασικό κίνητρο των ταξιδιών είναι να ξεφύγει κανείς από την καθημερινότητα και να αναζητήσει νέες εμπειρίες, πολλοί ταξιδιώτες προτιμούν να ασχολούνται με τις ίδιες δραστηριότητες που ασχολούνται στην καθημερινότητά τους. Παραδείγματος χάρη, αν οι άνθρωποι ακολουθούν συγκεκριμένη διατροφή, μέσα σε συγκεκριμένα πλαίσια θα προσπαθούν να την ακολουθούν και κατά τη διάρκεια του ταξιδιού τους ή αν ο ταξιδιώτης αθλείται καθημερινά, θα προτιμήσει να τρέξει σε ένα δρόμο, στην παραλία ή να προπονηθεί στο γυμναστήριο του ξενοδοχείου. Υπό αυτό το πρίσμα, η σχέση μεταξύ της καθημερινής ζωής και της τουριστικής εμπειρίας συμβάλλει στην μεγιστοποίηση της ικανοποίησης της τουριστικής εμπειρίας και την καθιστά ακόμη πιο αξέχαστη.

Αν και η έννοια του ψηφιακού ταξιδιώτη δίνει έμφαση στο άτομο, μπορεί εύκολα να επεκταθεί στις κοινωνικές ομάδες. Αυτό οφείλεται στο γεγονός ότι οι άνθρωποι συχνά μοιράζονται πληροφορίες σχετικές με τη ζωή τους (επίπεδο ευτυχίας, απόσταση περπατήματος ανά ημέρα) με άλλους με σκοπό την συλλογική ανάπτυξη του γνωστικού επιπέδου, τη συγκριτική αξιολόγηση των επιδόσεων και τη συμμετοχή σε διάφορες κοινότητες. Συνεπώς, είναι πιθανό άλλες οντότητες πέρα του προσωπικού και ιδιωτικού τομέα, όπως τα ταξιδιωτικά πρακτορεία και οι τουριστικοί οργανισμοί να έχουν πρόσβαση στις συγκεκριμένες πληροφορίες μέσω των κοινοτήτων αυτών ή των υπηρεσιών υπολογιστικού νέφους και με τη σειρά τους να παρέχουν ανατροφοδοτήσεις (λύσεις, εκπαιδευτικά κουπόνια κλπ) σε πραγματικό χρόνο (Lupton, 2014). Αυτό μπορεί να έχει τεράστια σημασία και να αποτελεί κοινή πρακτική στο κοντινό μέλλον καθώς οι τεχνολογικές εξελίξεις μας επιτρέπουν να μαθαίνουμε εύκολα από τους άλλους μοιράζοντας και συγκρίνοντας τον τρόπο με τον οποίο ο καθένας κάνει πράγματα και

έτσι λαμβάνοντας σημαντικές πληροφορίες και ιδέες από τις συλλογικές δραστηριότητες. Με αυτά τα πλεονεκτήματα οι επιχειρήσεις μπορούν να επαναπροσανατολίσουν αυτά τα δεδομένα και να δημιουργήσουν εμπορική αξία, αν και τα βασικά δεδομένα που δημιουργούνται βασίζονται αποκλειστικά σε προσωπικές δραστηριότητες με εθελοντική δέσμευση. Καθώς αυτές οι τεχνολογίες ενσωματώνονται ολοένα και περισσότερο στην καθημερινότητα μέσω των έξυπνων συσκευών, τα ταξίδια και ο τουρισμός μπορούν να θεωρηθούν ως πεδίο λογικής επέκτασης της έννοιας της ψηφιοποίησης, ιδιαίτερα λόγω των πιθανών εφαρμογών στην ανάπτυξη του έξυπνου τουρισμού.

Ο Ψηφιακός Ταξιδιώτης και η Ανάπτυξη του Έξυπνου Τουρισμού

Ο έξυπνος τουρισμός αναφέρεται στη σύγκλιση των τεχνολογιών πληροφορικής, των επιχειρηματικών δραστηριοτήτων και των τουριστικών εμπειριών. Οι Gretzel, Sigala, Xiang και Koo (2015) υποστηρίζουν ότι στο επίκεντρο της τεχνολογίας του έξυπνου τουρισμού βρίσκονται οι αισθητήρες και οι φορητές συσκευές που επιτρέπουν στους προορισμούς να δημιουργούν το διάχυτο τεχνολογικό περιβάλλον το οποίο μπορούν να χρησιμοποιήσουν οι τουριστικοί οργανισμοί για να προβλέψουν τις ανάγκες των ταξιδιωτών σε πραγματικό χρόνο προκειμένου να βελτιώσουν τις εμπειρίες τους και να ενθαρρύνουν την ανταλλαγή των τουριστικών εμπειριών. Κατά συνέπεια, υποστηρίζεται ότι προκειμένου να αναπτυχθεί ο έξυπνος τουρισμός απαιτείται από τους προορισμούς και τις επιχειρήσεις να ενσωματώσουν στις βασικές τους δραστηριότητες, την εξατομίκευση, την περιβαλλοντική συναίσθηση, την παρακολούθηση σε πραγματικό χρόνο μέσω της συλλογής πληροφοριών, την ευρεία συνδεσιμότητα και τη δυνατότητα συγχρονισμού σε πραγματικό χρόνο. (Gretzel, Sigala, Xiang και Koo, 2015). Στο πλαίσιο αυτό, υποστηρίζεται επιπλέον ότι η έννοια του ψηφιακού ταξιδιώτη κατέχει το κλειδί για την κατανόηση του τρόπου με τον οποίο οι σημερινές συσκευές και φορητές τεχνολογίες συμβάλλουν στην τουριστική εμπειρία και πώς μπορούν να χρησιμοποιηθούν για την βελτίωση της έξυπνης τουριστικής ανάπτυξης. Συγκεκριμένα, από άποψη σχεδιασμού υπηρεσιών και συστημικής ανάπτυξης, ο ψηφιακός ταξιδιώτης 1) παρέχει πληροφορίες για την περιβαλλοντική συναίσθηση, 2) μέσω των

δραστηριοτήτων του συνδέει την τουριστική εμπειρία με την καθημερινή του ζωή και συνεπώς 3) μας επιτρέπει να κατανοήσουμε τις αλληλεπιδράσεις του ταξιδιώτη με το περιβάλλον.

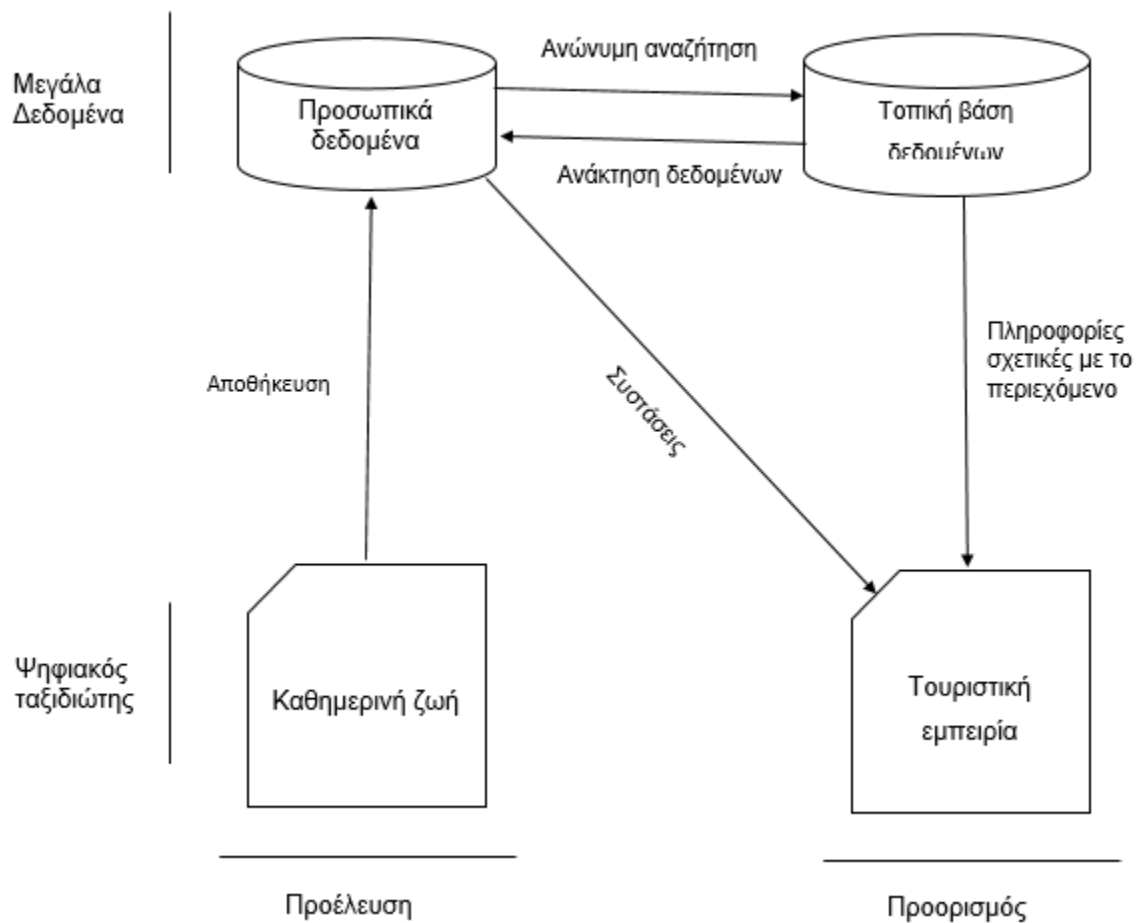
Άλλο ένα κλειδί για την ανάπτυξη του έξυπνου τουρισμού είναι η κατανόηση του τρόπου με τον οποίο ο ταξιδιώτης αλληλεπιδρά με το φυσικό και το κοινωνικό του περιβάλλον. Η τεχνολογία (φορητές συσκευές, αισθητήρες και Διαδικτύου Πραγμάτων) παίζει σημαντικό αλλά σιωπηρό ρόλο στη διευκόλυνση της αλληλεπίδρασης μεταξύ των ταξιδιωτών και του περιβάλλοντος τους. Πράγματι, η τεχνολογία θεωρείται αποτελεσματικό μέσο στη δημιουργία, την υποστήριξη και την ενίσχυση των τουριστικών εμπειριών μέσω της παροχής πληροφοριών, τη διεύρυνση των επιλογών της τουριστικής συμπεριφοράς και την δυνατότητα να μοιράζονται τις εμπειρίες τους με την οικογένεια και τους φίλους τους ακόμη και κατά τη διάρκεια των διακοπών τους. Μεταξύ των λειτουργιών αυτών, η δυνατότητα αναζήτησης πληροφοριών και ανάκτησης συμπεριφορών αποτελούν τις πιο ζωτικές λειτουργίες που επηρεάζουν τις συμπεριφορές και τις εμπειρίες των ταξιδιωτών. Η χρήση αυτών των τεχνολογιών επιτρέπει στους ταξιδιώτες να δημιουργούν και να διαχειρίζονται τις τουριστικές τους εμπειρίες όχι μόνο παθητικά, λαμβάνοντας πληροφορίες από τους τουριστικούς οργανισμούς αλλά και ενεργητικά συμμετέχοντας σε δραστηριότητες εντός του προορισμού.

Ο ψηφιακός ταξιδιώτης δεν παρέχει μόνο πληροφορίες για το περιβάλλον του κατά τη διάρκεια του ταξιδιού αλλά και προσωπικά δεδομένα κατά τη διάρκεια της καθημερινότητάς του. Τα αποτελέσματα σύνδεσης των προσωπικών δεδομένων με την τουριστική εμπειρία μπορούν να χρησιμοποιηθούν για την ανάπτυξη του έξυπνου τουρισμού. Ένα παράδειγμα τέτοιου συστήματος απεικονίζεται στο Σχήμα 3. Με άλλα λόγια, μέσω της αυξημένης χρήσης των φορητών συσκευών και αισθητήρων, η φυσική κατάσταση (ιστορικό αγοράς, κινήσεις και ιστορικό αναζήτησης) και η συναισθηματική κατάσταση (διάθεση και συναίσθημα) μπορούν να παρακολουθούνται και στη συνέχεια να αποθηκεύονται σε πραγματικό χρόνο. Επιπλέον, τα δεδομένα αυτά θα διευρυνθούν

Η διαχείριση μεγάλου όγκου δεδομένων στον Τουρισμό- Μια προσπάθεια βιβλιογραφικής και εμπειρικής προσέγγισης του θεματος

καθώς τα υπάρχοντα δεδομένα ενσωματώνονται στο σύστημα (Andrejevic & Burdon, 2014).

Σχήμα 3 Ένα βασικό σύστημα που εντάσσει τον ψηφιακό ταξιδιώτη και την τουριστική εμπειρία για τον έξυπνο τουριστικό προορισμό



Η Αξιοποίηση των Κριτικών της Τουριστικής Βιομηχανίας

Εισαγωγή

Η εμφάνιση του web 2.0 άλλαξε τον τρόπο με τον οποίο οι ξενοδοχειακοί πελάτες αναζητούν πληροφορίες. Πλέον δεν βασίζονται μόνο σε μηνύματα από τα ξενοδοχεία, αλλά διαβάζουν και μοιράζονται εμπειρίες σε διαφορετικές ιστοσελίδες κοινωνικών δικτύων. Προηγούμενες μελέτες έδειξαν ότι ο μέσος ταξιδιώτης περιηγεί περίπου 40 ιστοσελίδες πριν προχωρήσει στην αγορά κάποιου τουριστικού πακέτου από ταξιδιωτικά γραφεία (Schaal, 2013). Οι ιστοσελίδες των ξενοδοχείων αντιστοιχούν σε 4% των πηγών που χρησιμοποιούνται στο διαδίκτυο, ενώ οι ιστοσελίδες των κριτικών σε 7% αυτού του συνόλου (Schaal, 2013). Περίπου το 62% των Ευρωπαίων συμμετέχουν ενεργά σε δραστηριότητες κοινωνικής δικτύωσης όπως η ανάγνωση και η σύνταξη ηλεκτρονικών κριτικών (Forrester Research, 2007). Το TripAdvisor, ως μία από τις μεγαλύτερες ταξιδιωτικές ιστοσελίδες, έχει λάβει πάνω από 200 εκατομμύρια κριτικές παγκοσμίως (TripAdvisor, 2014). Περίπου το 68% των χρηστών θεωρούν τις ηλεκτρονικές κριτικές ως μία από τις πιο αξιόπιστες πηγές πληροφοριών (NielsenWire, 2009). Συνεπώς, η παρουσία σε μια τέτοια ιστοσελίδα μπορεί να αυξήσει την πιθανότητα ενός ξενοδοχείου να αποτελέσει την τελική επιλογή του ταξιδιωτικού προορισμού.

Είναι προφανές ότι ως αποτέλεσμα της ταχείας ανάπτυξης του διαδικτύου και των έξυπνων συσκευών, οι καταναλωτές γίνονται πιο ενημερωμένοι, εκλεπτυσμένοι και απαιτητικοί. Μια κριτική πελάτη είναι “ένα μείγμα γεγονότων και απόψεων, εντυπώσεων και συναισθημάτων, εμπειριών, ακόμη και φημών”. Ως εκ τούτου, οι πληροφορίες αυτές ονομάζονται και μεγάλα δεδομένα. Συνεπώς, τα μοτίβα που δημιουργούνται στις προτιμήσεις των ξενοδοχείων πρέπει να προσδιορίζονται από μεγάλης κλίμακας περιεχόμενα που παράγονται από τους χρήστες (ΠΠΧ). Η καταναλωτική δόραση που προσφέρεται από τις ηλεκτρονικές κριτικές αναμένεται να δημιουργήσει αξία και κίνητρα στους διευθυντές των ξενοδοχείων ώστε να ενισχύσουν την ανταγωνιστικότητα και τη φήμη τους στην αγορά.

Στο παρόν κεφάλαιο θα συζητηθεί, η εμφάνιση του διαδικτυακού κινήματος κριτικών στον ξενοδοχειακό κλάδο, η περιγραφή της φύσης και των χαρακτηριστικών των ηλεκτρονικών κριτικών, η επεξήγηση της επιρροής των ηλεκτρονικών κριτικών στον ξενοδοχειακό κλάδο, η συμπεριφορά των καταναλωτών και πιο συγκεκριμένα η παρουσίαση των τρόπων που οι ακαδημαϊκοί και οι ξενοδόχοι λαμβάνουν και αξιοποιούν τις πληροφορίες από τις ηλεκτρονικές κριτικές προκειμένου να παρουσιάσουν τόσο οι επιτυχημένες όσο και οι αποτυχημένες απαντήσεις των ξενοδοχείων για να επισημανθούν οι βέλτιστες πρακτικές στη διαχείριση της φήμης. Μια ανάλυση του θεωρητικού πλαισίου των ακαδημαϊκών μελετών και των πρακτικών που επιστρατεύονται από τα ξενοδοχεία μας δίνει τη δυνατότητα να συγκεντρώσουμε και να διευκρινίσουμε τα ζητήματα που σχετίζονται με τις ηλεκτρονικές κριτικές καθώς και την επιρροή τους στην απόδοση των ξενοδοχείων.

Η Εμφάνιση των Ηλεκτρονικών Κριτικών στον Ξενοδοχειακό Κλάδο

Η ταχεία ανάπτυξη του διαδικτύου και του web 2.0 επέτρεψε στους χρήστες να διαδίδουν, να συνεργάζονται και να ανταλλάσσουν ταξιδιωτικές εμπειρίες διαδικτυακά. Οι πελάτες δεν μπορούν να αξιολογήσουν τα προϊόντα και τις υπηρεσίες ενός ξενοδοχείου προτού τα καταναλώσουν. Τέτοιοι πελάτες μπορούν να ερευνήσουν διεξοδικά την επιλογή τους χρησιμοποιώντας διάφορες πηγές πληροφοριών για τον περιορισμό του ρίσκου. Οι ηλεκτρονικές κριτικές συγκαταλέγονται στις αναδυόμενες μορφές των ΠΠΧ και του ηλεκτρονικού word of mouth και οι κριτικές αυτές παίζουν σημαντικό ρόλο στη λήψη αποφάσεων των καταναλωτών. Σύμφωνα με έρευνες που διεξήχθησαν το 2012, περισσότερο από το 50% των ταξιδιωτών δεν θα έκλεινε ξενοδοχείο που δεν έχει κριτικές. Συνεπώς, οι ηλεκτρονικές κριτικές επηρεάζουν σε μεγάλο βαθμό τον ξενοδοχειακό κλάδο.

Στο διαδίκτυο υπάρχει μεγάλη ποικιλία ιστοσελίδων τουριστικών κριτικών. Ορισμένες τέτοιες ιστοσελίδες όπως το TripAdvisor και το Yelp φιλοξενούν κριτικές που είναι γραμμένες από τους ίδιους τους πελάτες των ξενοδοχείου χωρίς τη δυνατότητα κρατήσεων. Ως εκ τούτου, οι κριτικές που δημοσιεύονται σε αυτούς τους ιστότοπους θεωρούνται αξιόπιστες πηγές πληροφοριών (Travel Media Group, 2015). Η Google+, η

Yahoo!, η Travel και άλλες μηχανές αναζήτησης επιτρέπουν στους χρήστες τους να αναρτούν κριτικές ξενοδοχείων υπό τον όρο της εγγραφής τους. Η Expedia, η Hotels, η Orbitz, η Travelocity, η Booking και άλλα ηλεκτρονικά ταξιδιωτικά πρακτορεία (OTA) έχουν ως πολιτική να δημοσιεύουν επαληθευμένες κριτικές. Συγκεκριμένα, μόνο οι χρήστες που έχουν κάνει κράτηση από τις αντίστοιχες ιστοσελίδες προσκαλούνται μέσω e-mail να δημοσιεύσουν σχόλια.

Η περίοδος έντασης με τις αγωγές των ξενοδοχείων κατά των ιστοσελίδων ηλεκτρονικών κριτικών σχετικά με τη δημοσίευση ψεύτικων και παραπλανητικών πληροφοριών ανήκει στο παρελθόν και από το 2012 η σχέση των δυο μερών έχει βελτιωθεί. Αρκετές ξενοδοχειακές αλυσίδες έχουν συνεργαστεί με ιστοσελίδες αναθεώρησης για να αξιοποιήσουν αυτήν την τάση. Ο όμιλος Four Seasons δίνει τη δυνατότητα στους πιθανούς πελάτες τους να ελέγξουν τις κριτικές και τα σχόλια άλλων χρηστών στην ιστοσελίδα τους. Κάθε μια από ιστοσελίδες των ξενοδοχείων της Four Seasons περιλαμβάνει κριτικές που προέρχονται από το TripAdvisor, το Twitter και το Facebook (Four Seasons, 2012). Η Wyndham και η Accor ενθαρρύνουν τους επισκέπτες τους, μετά τη διαμονή τους να γράψουν κριτικές στο TripAdvisor. Ο συνολικός αριθμός των αξιολογήσεων, η μέση βαθμολογία και οι πρόσφατες κριτικές στη συνέχεια εμφανίζονται στην ιστοσελίδα της κάθε ιδιοκτησίας. Ο Eric Danziger, Διευθύνων Σύμβουλος της Wyndham Hotel Group, εξηγεί ότι "οι τουρίστες έχουν μια ισχυρή τάση να μελετούν και να αναρτούν κριτικές, οι κριτικές με τη σειρά τους επιτρέπουν τα ξενοδοχεία να διορθώσουν τα όποια προβλήματα, τέλος η εύκολη πρόσβαση των χρηστών σε κριτικές έχει ήδη μεταφραστεί σε αύξηση των κρατήσεων". Πιο συγκεκριμένα, οι κρατήσεις των ξενοδοχείων της Wyndham αυξήθηκαν κατά 30% με την προσθήκη του λογότυπου του TripAdvisor, της εμφάνισης των βαθμολογιών και της σύνδεσης με το TripAdvisor για δημοσίευση των πρόσφατων εμπειριών (USA Today, 2012).

Με την αυξανόμενη δημοτικότητα των ιστότοπων ηλεκτρονικών κριτικών, αρκετές ξενοδοχειακές επιχειρήσεις εξέτασαν το ενδεχόμενο της ενσωμάτωσης των δυνατοτήτων δημοσίευσης κριτικών στις ιστοσελίδες τους. Η Starwood είναι η πρώτη

ξενοδοχειακή αλυσίδα παγκοσμίως που καλεί τα μέλη της Starwood Preferred Guest να μοιραστούν τις σκέψεις και τις απόψεις τους στις ιστοσελίδες των ξενοδοχείων που διέμεναν. Πρόκειται για επισκέπτες που διαμένουν σε οποιοδήποτε από τα ξενοδοχεία της αλυσίδας για περισσότερες από 50 διανυκτερεύσεις ετησίως. Το πρόγραμμα αυτό τέθηκε σε εφαρμογή στα τέλη του 2011 και έχει αποδειχθεί επιτυχημένο από τις 529.000 νέες δημοσιεύσεις κριτικών στις ιστοσελίδες όλων των ξενοδοχείων της αλυσίδας από τον Ιανουάριο έως τα μέσα Σεπτεμβρίου του 2013. Επιπλέον, η ανταποκρίσεις στις κριτικές από τα ξενοδοχεία ήταν πολύ υψηλότερες (46%) από το μέσο όρο του κλάδου (23%) κατά το ίδιο έτος. Η Marriott ακολουθεί παρόμοια πρακτική, ενσωματώνοντας την παρουσίαση των ηλεκτρονικών κριτικών στην ιστοσελίδα της Marriott Rewards Insiders. Ωστόσο, μόνο τα ενημερωμένα μέλη της Marriott Club, οι οποίοι είναι και τακτικοί πελάτες έχουν τη δυνατότητα δημοσίευσης σχολίων που είναι διαθέσιμα στο ευρύ κοινό. Οι ιστοσελίδες αξιολόγησης ταξιδιών όπως το TripAdvisor, δεν μπορούν να επαληθεύσουν εάν ο χρήστης έχει μείνει όντως στο ξενοδοχείο που αποτελεί αντικείμενο κριτικής, Έτσι, οι ξενοδόχοι με δική τους πρωτοβουλία, δίνουν τη δυνατότητα στους επισκεπτών τους, να αξιολογήσουν και σχολιάσουν, ενισχύοντας έτσι την γνησιότητα των κριτικών. Το σημαντικότερο στοιχείο είναι ότι δεν παρέχεται κανένα κίνητρο στους σχολιαστές και οι κριτικές που αναρτώνται μένουν αυτούσιες. Η τακτική αυτή επιτρέπει στις ξενοδοχειακές επιχειρήσεις να αποφεύγουν σημαντικά έξοδα από τις προμήθειες που καταβάλλουν στα ΟΤΑ, όπως επίσης να ενθαρρύνουν τους πιθανούς πελάτες να περιηγούνται και να κάνουν κρατήσεις μέσω της ιστοσελίδας τους.

Τα Είδη και τα Χαρακτηριστικά των Ηλεκτρονικών Κριτικών

Οι ηλεκτρονικές κριτικές, ως νέα μεγάλη τάση που επηρεάζει την τουριστική βιομηχανία, αποτελούν ένα σημαντικό αντικείμενο που παρακινεί την ακαδημαϊκή κοινότητα στη δημιουργία γνώσης και ενός θεωρητικού πλαισίου για τη δημιουργία ενός νέου κλάδου (Leung, Law , van Hoof, & Buhalis, 2013). Προηγούμενες μελέτες που εξέτασαν τη φύση και τα χαρακτηριστικά των ηλεκτρονικών κριτικών σε ξενοδοχεία εντόπισαν δύο ενδιαφέροντα χαρακτηριστικά. Το πρώτο είναι ότι οι κριτικές των ξενοδοχείων παρουσιάζουν έντονη κλίση προς τις θετικές αξιολογήσεις. Πιο

συγκεκριμένα, πάνω από το 70% των 1,3 εκατομμυρίων κριτικών ξενοδοχείων στο TripAdvisor ήταν θετικές (4 ή άνω σε κλίμακα βαθμολόγησης 5 σημείων), ενώ μόλις το 15% ήταν αρνητικές (κάτω από 3). Το δεύτερο χαρακτηριστικό είναι ότι η αντικειμενικότητα της αξιολόγησης ενός ξενοδοχείου γίνεται πιο αντικειμενική καθώς αυξάνεται ο όγκος των κριτικών, περιορίζοντας έτσι τις επιπτώσεις των αρνητικών αξιολογήσεων. Έτσι, η αρνητική επιρροή των κακών κριτικών μπορεί να περιοριστεί εάν οι ξενοδόχοι ενθαρρύνουν τους επισκέπτες τους να αναρτούν θετικές κριτικές στις αντίστοιχες ιστοσελίδες.

Σε μια διαφορετική κατεύθυνση έρευνας, μελετάται το κίνητρο για τη δημοσίευση και την ανάγνωση των ηλεκτρονικών σχολίων. Οι Yoo και Gretzel (2008) κατέληξαν στο συμπέρασμα ότι τα κίνητρα για τη δημοσίευση σχολίων είναι μεταξύ άλλων η ηθική αυτοϊκανοποίηση που λαμβάνουν οι χρήστες, η διάθεση να βοηθήσουν τον παροχέα των ταξιδιωτικών υπηρεσιών που τους φιλοξένησε και προκειμένου να βοηθήσουν άλλους καταναλωτές. Η πράξη της εκτόνωσης αρνητικών συναισθημάτων μέσω των κριτικών δεν είναι τόσο συνηθισμένη στο ευρύ κοινό όσο αναμενόταν. Οι Oguta και Cezara (2012) διερεύνησαν περαιτέρω το θέμα των κινήτρων πίσω από τις κριτικές των ξενοδοχείου και τα αποτελέσματα της μελέτης αυτής έδειξαν ότι οι πελάτες που είναι ευχαριστημένοι με την τιμή και την εξυπηρέτηση του ξενοδοχείου είναι πιο πρόθυμοι να δημοσιεύσουν κριτικές σε σύγκριση με τους δυσαρεστημένους πελάτες. Μολονότι η βαθμολογίες των ξενοδοχείων στις εκάστοτε ιστοσελίδες κριτικών δεν φαίνεται να επηρεάζουν σημαντικά την πρόθεση αξιολόγησης των χρηστών, ούτε η υπερβάλλουσα ικανοποίηση ή δυσαρέσκεια φαίνεται να επηρεάζει τα κίνητρα δημοσίευσης κριτικών.

Από την άλλη πλευρά, αρκετές έρευνες διευκρινίζουν το λόγο που οι καταναλωτές διαβάζουν ηλεκτρονικές κριτικές. Οι τρεις βασικοί κινητήριοι παράγοντες είναι, ευκολία και ποιότητα (γρήγορη και αποτελεσματική μέθοδο απόκτησης πληροφοριών για την καλύτερη αξία και το χαμηλό κόστους), μείωση κινδύνου (σωστή λήψη αποφάσεων αγοράς) και κοινωνική επιβεβαίωση (η διαπίστωση ότι και άλλος κόσμος μοιράζονται τις ίδιες απόψεις, η συμμετοχή σε μια κοινότητα, η σύγκριση αξιολογήσεων με άλλους χρήστες). Τα αποτελέσματα της έρευνας που αναφέρθηκαν

παραπάνω κατέδειξαν ότι γενικά οι γυναίκες βασίζονται σε κριτικές ξενοδοχείων για την ελαχιστοποίηση των κινδύνων και την αναζήτηση ευχερειών και ποιότητας. Αντίθετα, οι άνδρες όταν διαβάζουν κριτικές ξενοδοχείων βασίζονται κυρίως στις προσωπικές τους γνώσεις στις ηλεκτρονικές κρατήσεις. Μελετώντας τις κριτικές ξενοδοχείων τα πέντε σημαντικότερα χαρακτηριστικά αξιολόγησης είναι, η καθαριότητα των δωματίων, η τοποθεσία και η προσβασιμότητα σε άλλα σημεία ενδιαφέροντος, η σχέση ποιότητας-τιμής, η εξυπηρέτηση και η ασφάλεια (Ong, 2012). Η μελέτη του Ong αποκάλυψε ότι οι αναγνώστες κριτικών δίνουν σημασία σε δύο βασικά χαρακτηριστικά, εκ των οποίων το ένα είναι το προφίλ των χρηστών που αναρτούν κριτικές και το άλλο η ημερομηνία της διαμονής στο ξενοδοχείο ή η ημερομηνία δημοσίευσης της κριτικής. Επίσης, άλλο λόγο ανάγνωσης κριτικών αποτελούν οι ασαφείς πληροφορίες που περιλαμβάνονται στις διαφημίσεις και τους ταξιδιωτικούς οδηγούς των ξενοδοχείων.

Η κατανόηση του λόγου που οι άνθρωποι διαβάζουν ή γράφουν κριτικές στο διαδίκτυο είναι πολύ σημαντικό για τους ξενοδόχους καθώς πολλά ηλεκτρονικά συστήματα κατάταξης βασίζονται στη δραστηριότητα των κριτικών. Για παράδειγμα, ο Δείκτης Δημοτικότητας του TripAdvisor καθορίζεται από την ποσότητα, την ποιότητα και την επικαιρότητα των κριτικών. Το περιεχόμενο και η δραστηριότητα μιας επιχείρησης στη Google+ επηρεάζουν επίσης τη σειρά με την οποία αυτή εμφανίζεται στις αναζητήσεις της Google. Περισσότερες, καλύτερες και πιο πρόσφατες κριτικές ενδέχεται να βελτιώσουν τη θέση ενός ξενοδοχείου στο Δείκτη Δημοτικότητας του TripAdvisor ή στη μηχανή αναζήτησης Google (Travel Media Group, 2015). Συνεπώς, μια τέτοια στρατηγική μπορεί να ενισχύσει την ηλεκτρονική φήμη και να αυξήσει τις κρατήσεις ενός ξενοδοχείου.

Η Επιρροή των Ηλεκτρονικών Κριτικών στον Ξενοδοχειακό Κλάδο

Μια συνηθισμένη κριτική ξενοδοχείου παρέχει τόσο στατιστικά στοιχεία (το συνολικό αριθμό των σχολίων, βαθμολογίες ταξιδιωτών και βαθμολογίες για διαφορετικά χαρακτηριστικά του ξενοδοχείου, όπως η τοποθεσία, τα δωμάτια, η εξυπηρέτηση, η αξία και της χρησιμότητας της κριτικής) όσο και περιγραφικά στοιχεία (η γνώμη του σχολιαστή για το ξενοδοχείο, “Εγώ και η οικογένειά μου απολαύσαμε τις

διακοπές μας σε αυτό το υπέροχο ξενοδοχείο", "Οι εγκαταστάσεις είναι σύγχρονες και η τοποθεσία πολύ βολική", "τα δωμάτια με θέα θάλασσα αξίζουν τα λεφτά τους"), καθώς και συμβουλές ("Προσπαθήστε να αποφύγετε το οικογενειακό δωμάτιο", "Δοκιμάστε οπωσδήποτε το φαγητό στο εστιατόριο του ξενοδοχείου"). Επίσης, οι σχολιαστές και οι ξενοδόχοι μπορούν ακόμα και να δημοσιεύουν φωτογραφίες στην εκάστοτε ιστοσελίδα κριτικών. Πολλοί ιστότοποι παρέχουν τη δυνατότητα επικοινωνίας με το ξενοδοχείο μέσω ερωτο-απαντήσεων για τη διευκόλυνση των ηλεκτρονικών συζητήσεων. Οι άφθονες αριθμητικές, λεκτικές και οπτικές πληροφορίες που μπορούν να ληφθούν από τις ιστοσελίδες κριτικών έχει προσελκύσει το ενδιαφέρον των ακαδημαϊκών μελετητών για τον καθορισμό του τρόπου με τον οποίο οι κριτικές επηρεάζουν τον ξενοδοχειακό κλάδο.

Αρκετές μελέτες έχουν ερευνήσει πιο διεξοδικά την επιρροή των κριτικών στην απόδοση των ξενοδοχείων. Η πιο πρωτοποριακή μελέτη έγινε από τους Ye, Q., Law, R., Gu, B., και Chen, W.. (2011) όπου χρησιμοποιήθηκε το λογαριθμικό μοντέλο γραμμικής παλινδρόμησης για την πρόβλεψη των ηλεκτρονικών κρατήσεων ξενοδοχειακών δωματίων στην Κίνα. Τα αποτελέσματα έδειξαν ότι η διακύμανση στις βαθμολογίες των αξιολογήσεων δεν επηρέασε σημαντικά τις πωλήσεις. Εντούτοις, μια αύξηση βαθμολογιών κατά 10% οδηγεί σε αύξηση των ηλεκτρονικών κρατήσεων κατά 5%. Παρόλο που οι υψηλές τιμές δωματίων μπορεί να μειώσουν τον αριθμό των κρατήσεων, η βολική τοποθεσία του ξενοδοχείου, ειδικά σε μεγάλες πόλεις συμβάλλει στην ενίσχυση των κρατήσεων.

Οι επαγγελματίες του κλάδου αξιολογούν την απόδοση ενός ξενοδοχείου βάσει της πραγματικής πληρότητας του ξενοδοχείου, τη μέση ημερήσια τιμή (ADR) και τα έσοδα ανά διαθέσιμο δωμάτιο (RevPAR). Υπό αυτό το πρίσμα, οι ακαδημαϊκοί επιστήμονες προσπάθησαν να συνεργαστούν με κοινοπρακτικές ερευνητικές εταιρίες ή με ηλεκτρονικά ταξιδιωτικά πρακτορεία (ReviewPro, Smith Travel Research, comScore, Travelocity και TripAdvisor) προκειμένου να αποκτήσουν πραγματικά δεδομένα για εμπειρική μοντελοποίηση. Ο Anderson (2012) μελέτησε τις επιπτώσεις των μέσων κοινωνικής δικτύωσης στην απόδοση επενδύσεων ενός ξενοδοχείου (ROI) μέσω

λογιστικής παλινδρόμησης. Σύμφωνα με τη μελέτη, η τιμή ενός ξενοδοχείου (όπως υπολογίζεται με βάση το ADR) αυξάνεται κατά 12% με την αύξηση ενός βαθμού στη βαθμολογία αξιολογήσεων (από 7 σε 8 σε κλίμακα 10 σημείων), όταν η πληρότητα ή το μερίδιο αγοράς ενός ξενοδοχείου παραμένουν σταθερά. Την ίδια στιγμή, η αύξηση της ηλεκτρονικής φήμης ενός ξενοδοχείου κατά 1% (όπως μετράται από το Δείκτη Global Review της ReviewPro) έχει ως αποτέλεσμα την αύξηση κατά 2% του RevPAR. Το αποτέλεσμα αυτό είναι πολύ σημαντικότερο σε ένα ξενοδοχείο μεσαίας κατηγορίας απ ότι σε ένα πολυτελές ξενοδοχείο. Η χρήση του μοντέλου γενικευμένης μεθόδου ροπών (GMM) επέτρεψε στον Duverger (2013) να προσδιορίσει ότι οι σύντομες θετικές κριτικές (ως προς τον αριθμό λέξεων) μειώνουν την αρνητική επίδραση στο μερίδιο αγοράς ενός ξενοδοχείου. Από την άλλη, όσο μεγαλύτερη είναι μια κριτική, ανεξαρτήτως αν είναι θετική ή αρνητική, τόσο αρνητικό αντίκτυπο έχει στο μερίδιο αγοράς του ξενοδοχείου. Οι Blal και Sturman (2014) προσδιόρισαν τις διαφορετικές επιδράσεις που έχει ο αριθμός και η βαθμολογία των κριτικών στο RevPAR ενός ξενοδοχείου και έδειξαν ότι η απόδοση των ξενοδοχείων χαμηλής κατηγορίας αυξάνεται όταν αυξηθεί ο αριθμός των κριτικών. Αντίθετα, οι πολυτελείς ξενοδοχειακές μονάδες θα πρέπει να επικεντρώνονται στην αύξηση των βαθμολογιών των αξιολογήσεων αντί να στοχεύουν στην αύξηση του αριθμού των κριτικών. Η υπέρβαση των προσδοκιών των πελατών μέσω της παροχής άριστων υπηρεσιών είναι πολύ σημαντικό για την αύξηση των πωλήσεων και των εσόδων σε πολυτελή ξενοδοχεία. Με την χρήση της νέας μεθοδολογικής προσέγγισης των τεχνητών νευρωνικών δικτύων (ANN), οι Phillips, Zigan, Silva και Schegg (2015) κατάφεραν να προσδιορίσουν ότι ο συντελεστής της βαθμολογίας δωματίου επηρέαζε θετικά το RevPAR ενός ελβετικού ξενοδοχείου, ενώ οι συντελεστές ποιότητας δωματίου και φήμης επηρεάζουν αρνητικά την απόδοση του ξενοδοχείου.

Η Επιρροή των Ηλεκτρονικών Κριτικών στην Καταναλωτική Συμπεριφορά

Οι ηλεκτρονικοί κριτικές μπορούν να μελετηθούν από διάφορες οπτικές όπως ο τύπος γραφής, οι συναισθηματικές εκφράσεις, η εξυπηρετικότητα, το προφίλ των χρηστών, η αξιοπιστία, η εμπιστοσύνη, η βαθμολογία, το μήκος και ο όγκος της

κριτικής. Συνεπώς, οι προηγούμενες μελέτες ανέλυαν υπό διαφορετικές οπτικές τις κριτικές και τον τρόπο με τον οποίο αυτές επηρεάζουν την καταναλωτική συμπεριφορά.

Οι Vermeulen και Seegers (2009) κατέληξαν στο συμπέρασμα ότι οι θετικές κριτικές βελτιώνουν τη στάση των καταναλωτών απέναντι σε ένα ξενοδοχείο. Αυτό το φαινόμενο είναι ιδιαίτερα αισθητό για ξενοδοχεία με χαμηλή αναγνωρισιμότητα επωνυμίας. Οι Sparks and Browning (2011) διερεύνησαν το ρόλο που παίζουν οι αξιολογήσεις των ξενοδοχείων στην έμπνευση εμπιστοσύνης και την πρόθεση πραγματοποίησης κρατήσεων. Αρχικά επιλέχθηκαν τέσσερα βασικά χαρακτηριστικά κριτικών ως ανεξάρτητες μεταβλητές και συγκεκριμένα, ο στόχος της κριτικής (εάν το περιεχόμενο σχετίζεται με τα βασικά χαρακτηριστικά ενός ξενοδοχείου, όπως η καθαριότητα, οι εγκαταστάσεις, η εξυπηρέτηση και το προσωπικό), το γενικό περιεχόμενο της κριτικής (αν τα σχόλια που αναρτήθηκαν από τους επισκέπτες ήταν θετικά ή αρνητικά), η αξιολόγηση της κριτικής (εάν η αξιολόγηση είναι θετική ή αρνητική), καθώς και η ύπαρξη ή η απουσία αριθμητικής βαθμολόγησης. Σύμφωνα με τα αποτελέσματα της μελέτης, οι καταναλωτές υιοθετούν την απλή προσέγγιση αξιολόγησης ενός ξενοδοχείου. Η επιρροή μιας κριτικής πολλαπλασιάζεται όταν το περιεχόμενο της και η αξιολόγηση της είναι αρνητικά πλαισιωμένα. Είναι αξιοσημείωτο ότι μεμονωμένη η βαθμολογία της κριτικής δεν αυξάνει τις κρατήσεις αλλά ούτε ενισχύει την εμπιστοσύνη των χρηστών. Όταν ένα σύνολο κριτικών έχουν θετικό περιεχόμενο βάσει των πρότυπων εξυπηρέτησης και υποστηρίζονται από καλές βαθμολογίες, οι καταναλωτές συνήθως δείχνουν εμπιστοσύνη στο ξενοδοχείο και κατά συνέπεια προτίθενται να κάνουν κράτηση. Κρίνοντας τη μελέτη ενός οικολογικού ξενοδοχειακού θέρετρου σε πειραματικό πλαίσιο, οι Sparks, B. A., Perkins, H. E., και Buckley, R., (To 2013) εξέτασαν την επίδραση της πηγής πληροφοριών (αν η κριτική δημοσιεύτηκε από πελάτη ή από το ίδιο το ξενοδοχείο), του περιεχομένου (εάν οι πληροφορίες που σχετίζονται με τις οικολογικές εγκαταστάσεις και οι διαδικασίες του θέρετρου είναι αόριστες ή συγκεκριμένες), τα επουσιώδη στοιχεία εγκυρότητας (παρουσία λογοτύπων οικολογικής πιστοποίησης), την πρόθεση κρατήσεων και τις πεποιθήσεις των επισκεπτών του ξενοδοχείου στην ωφελιμότητα των αξιολογήσεων, της αξιοπιστίας και

της εταιρικής κοινωνικής ευθύνης του εν λόγω ξενοδοχείου. Τα αποτελέσματα της συγκεκριμένης μελέτης κατέδειξαν ότι η συμπερίληψη τόσο των ΠΠΧ όσο και των παραγόμενων πληροφοριών από τις επιχειρήσεις με έμφαση στις ηλεκτρονικές προωθήσεις διευκολύνει την χρησιμότητα και την αποτελεσματικότητα της επικοινωνίας μηνυμάτων. Στοιχεία όπως βραβεία, λογότυπα και διαπιστευτήρια, αυξάνουν περαιτέρω την εμπιστοσύνη και την εταιρική κοινωνική ευθύνη του ξενοδοχείου. Τέλος, η εμπιστοσύνη αποτελεί ιδιαίτερα σημαντικό στοιχείο για την οικοδόμηση του προφίλ ενός ξενοδοχείου καθώς και την πρόθεση αγοράς.

Εξαγωγή Γνώσεων για τους Πελάτες μέσω Ηλεκτρονικών Κριτικών

Οι ξενοδόχοι έχουν αναγνωρίσει από καιρό τη σημασία των απόψεων και των σχολίων των πελατών στις αγοραστικές αποφάσεις. Ένα θέμα που απασχολεί τους ξενοδόχους είναι ο τύπος της γνώσης των πελατών και η διαδικασία εξαγωγής της (δημογραφικά χαρακτηριστικά, ψυχολογικά προφίλ, κύρια ενδιαφέροντα, επιθυμητές ξενοδοχειακές υπηρεσίες και προτιμώμενες παροχές) από διαφορετικές πηγές δεδομένων για την αύξηση της ικανοποίησης των πελατών και της απόδοσης του ξενοδοχείου. Οι κύριες παραδοσιακές μέθοδοι εξαγωγής πληροφοριών που χρησιμοποιούνται από τους ακαδημαϊκούς και τους επαγγελματίες είναι οι έρευνες, οι δημοσκοπήσεις, οι ομάδες έρευνας αγοράς και η μέθοδος “mystery guest” (Withiam, 1995). Ωστόσο το μεγαλύτερο μειονέκτημα αυτών των μεθόδων είναι η χρήση ενός δείγματος του πληθυσμού αντί δεδομένων απ όλο τον πληθυσμό κατά την χρονική περίοδο της έρευνας. Η χρήση των ταξιδιωτικών κριτικών ως πηγής δεδομένων για ανάλυση μπορεί να ελαχιστοποιήσει τον περιορισμό του μικρού μεγέθους δειγμάτων καθώς μεγάλοι όγκοι δεδομένων είναι άμεσα διαθέσιμοι στις ιστοσελίδες ηλεκτρονικών κριτικών. Οι πληροφορίες σε αυτούς τους ιστότοπους βασίζονται σε πραγματικές εμπειρίες τουριστικών συναλλαγών. Οι ηλεκτρονικές κριτικές έχουν αναδειχθεί ως πηγή δεδομένων για τη εξαγωγή γνώσεων για τους πελάτες τα τελευταία έτη, χάρη στα πλεονεκτήματα της απλότητας, του χαμηλού κόστους, της ταχύτητας και της μη παρεμβατικότητας,.

Οι προηγούμενες μελέτες υιοθέτησαν διαφορετικές προσεγγίσεις για την λεξική ανάλυση των κριτικών. Συγκεκριμένα, την ανάλυση περιεχομένου (Au, Buhalis, & Law, 2014, Barreda & Bilgihan, 2013, O'Connor 2010) και την εξόρυξη κειμένου (Berezina, Bilgihan, Cobanoglu, & Okumus, 2015).

Με την εφαρμογή της ανάλυσης περιεχομένου, ο O'Connor (2010) κατάφερε να εντοπίσει τα κοινά χαρακτηριστικά της ικανοποίησης και δυσαρέσκειας των πελατών του Λονδίνου όπως αυτή αντικατοπτρίζεται στο TripAdvisor. Η τοποθεσία του ξενοδοχείου, το μέγεθος του δωματίου, η καλή εξυπηρέτηση του προσωπικού, η καθαριότητα, η ποικιλία στις επιλογές του πρωινού μπουφέ, οι εγκαταστάσεις του δωματίου, η άνεση, η θερμοκρασία, η συντήρηση και ο θόρυβος αποτελούν τις 10 πρώτες ανησυχίες των ερωτηθέντων. Οι Barreda και Bilgihan (2013) προσδιόρισαν τα στοιχεία εκείνα των εμπειριών κατά τη διαμονή σε ξενοδοχείο (είτε θετικά, είτε αρνητικά) που είναι ικανά να παρακινήσουν τους πελάτες ώστε να αξιολογήσουν ένα ξενοδοχείο και να δημοσιεύσουν μια κριτική στο TripAdvisor. Σύμφωνα με τη μελέτη τα κύρια θέματα που απασχολούν τα σχόλια των ηλεκτρονικών κριτικών και κατά συνέπεια επηρεάζουν την εμπειρία των επισκεπτών κατά τη διαμονή τους σε ένα ξενοδοχείο είναι, οι εμπειρίες υπηρεσιών, οι εσωτερικοί χώροι δωματίων, η τοποθεσία και η καθαριότητα, η ποιότητα του ύπνου, η τιμή, η ατμόσφαιρα του ξενοδοχείου, καθώς και οι ανέσεις και οι συμπληρωματικές υπηρεσίες. Η βολική τοποθεσία είναι το κοινό χαρακτηριστικό που εκτιμάται από τις περισσότερες θετικές κριτικές. Από την άλλη, στις αρνητικές κριτικές ο βασικός παράγοντας που φαίνεται να επηρεάζει τις εμπειρίες των πελατών είναι η καθαριότητα. Προκειμένου να βελτιωθεί η εικόνα της εμπορικής επωνυμίας τους, οι ξενοδόχοι πρέπει να επικεντρωθούν στην παροχή ποιοτικών υπηρεσιών, μίας και αποτελεί κύριο παράγοντα στην παρακίνηση ανάρτησης θετικών κριτικών. Όπως αναφέρθηκε προηγουμένως, η εξόρυξη κειμένου χρησιμοποιείται όλο και περισσότερο για την ανάλυση των ηλεκτρονικών κριτικών. Η χρήση των μηχανών αναζήτησης στις εφαρμογές εξόρυξης κειμένου (Intelligent Miner for Text της IBM και το Mine for Text της SAS) ελαχιστοποιεί την ανθρώπινη παρέμβαση στην εξερεύνηση των

μοτίβων και κανόνων του ογκωδών λεκτικών δεδομένων. Ως αποτέλεσμα ο αριθμός των ανθρώπινων λαθών που καταγράφηκαν ήταν περιορισμένος.

Οι Berezina, K., Bilgihan, A., Cobanoglu, C. και Okumus, F., (2015) μέσω ανάλυσης κειμένου προσδιόρισαν ότι οι ξενοδοχειακές υπηρεσίες είναι αυτές που αφήνουν ευχαριστημένους τους πελάτες, ενώ η επίπλωση (κρεβάτια, χαλιά και πετσέτες) και τα οικονομικά ζητήματα (χρήματα, χρεώσεις και μετρητά) παρακινούν τους δυσαρεστημένους πελάτες να δημοσιεύσουν αρνητικές κριτικές.

Η ανάλυση περιεχομένου και η εξόρυξη κειμένου στοχεύουν στον εντοπισμό μεμονωμένων ζητημάτων ενός ξενοδοχείου, με αποτέλεσμα να μην μπορούν να ανιχνεύσουν τις αλλαγές στα υπό ερώτηση ζητήματα με την πάροδο του χρόνου. Σε πρόσφατη μελέτη, οι Li, G., Law, R., Vu, H. Q., Rong, J., και Zhao, X. R., (2015) πρότειναν μια νέα μέθοδο έρευνας που ονομάζεται τεχνική εξόρυξης προτύπων (EPM) για τον εντοπισμό των χαρακτηριστικών των ξενοδοχείων που θα αυξήσουν τη ζήτηση τους. Στην έρευνα αυτή συλλέχθηκε μεγάλος όγκος ηλεκτρονικών κριτικών που δημοσιεύθηκε από το 2009 έως το 2013 (συνολικός όγκος 120.000 κριτικών σε περίοδο 5 ετών) και εντοπίστηκαν τα χαρακτηριστικά των ξενοδοχείων που ενδιαφέρουν τους τουρίστες. Τα δωμάτια, ο αριθμός του προσωπικού, η τοποθεσία, η ποικιλία του πρωινού μπουφέ, η εξυπηρέτηση, η καθαριότητα, η διατροφή, η πισίνα, τα δάπεδα και η θέα ήταν αυτά που καθόρισαν τα 10 δημοφιλέστερα χαρακτηριστικά ξενοδοχείων βάσει των ηλεκτρονικών κριτικών. Μεταξύ αυτών των χαρακτηριστικών, οι ξενοδοχειακές εγκαταστάσεις, οι υπηρεσίες και η διατροφή υπήρξαν και εξακολουθούν να αποτελούν σημαντικά χαρακτηριστικά των ξενοδοχείων, όπως αντικατοπτρίζονται στις ηλεκτρονικές κριτικές. Με την υιοθέτηση της μεθόδου EPM, οι ξενοδόχοι μπορούν να προσαρμόσουν τις προσφορές τους στην αγορά και να μεγιστοποιήσουν την απόδοση των επενδύσεων τους βελτιώνοντας τους πόρους τους ώστε να ικανοποιήσουν τις μελλοντικές ανάγκες και απαιτήσεις των πελατών.

Αποκρίσεις στις Ηλεκτρονικές Κριτικές

Οι θετικές κριτικές προσδίδουν αξία στον ξενοδοχειακό κλάδο, διότι η πρακτική της κριτικής προσφέρει στην ουσία επιπλέον δημοσιότητα. Από την άλλη, οι αρνητικές

κριτικές είναι ζημιογόνες για την εταιρική φήμη αν τα εν λόγω ξενοδοχεία αδυνατούν να τις χειριστούν σωστά. Δυστυχώς δεν γνωρίζουν όλοι οι διευθυντές ξενοδοχείων το βαθμό της επιρροής που ασκεί το αρνητικό word of mouth στους καταναλωτές, ορισμένοι μάλιστα υποτιμούν ακόμη και τις επιπτώσεις που προκαλεί στην εταιρική φήμη. Παρόλο που αρκετοί μάνατζερς αντιλαμβάνονται την αυξανόμενη σημασία της χρήσης των ηλεκτρονικών κριτικών ώστε να ασχοληθούν με τις έγνοιες των πελατών, δεν απαντούν όλα τα ξενοδοχεία σε τέτοιου είδους κριτικές. Λίγα είναι τα ξενοδοχεία που διαχειρίζονται ενεργά τις απαιτήσεις των πελατών τους σε ιστοσελίδες ηλεκτρονικών κριτικών. Συγκεκριμένα, λιγότερο από 0,5% των κριτικών έλαβαν απαντήσεις από τις διευθύνσεις των ξενοδοχείων παρόλο που ιστοσελίδες όπως το TripAdvisor, παρέχουν τη δυνατότητα άμεσης ανταπόκρισης σε αρνητικά σχόλια. Σε συνέντευξη με διευθυντές ξενοδοχείων που ανταποκρίνονται συχνά σε ηλεκτρονικές κριτικές, οι Park και Allen (2013) διαπίστωσαν ότι οι διευθυντές αυτοί αντιλαμβάνονται τις κριτικές των πελατών ως ειλικρινείς εκφράσεις των συναισθημάτων τους. Εν αντιθέσει, ο λόγος που άλλοι διευθυντές δεν ανταποκρίνονται σε (αρνητικές) κριτικές είναι ότι τις αντιμετωπίζουν ως την “εξαίρεση του κανόνα”. Αρκετοί ξενοδόχοι μάλιστα αισθάνονται ότι απειλούνται από πελάτες που αναρτούν κακές κριτικές με σκοπό να αποσπάσουν δωρεάν υπηρεσίες (BBC, 2014). Ως συνέπεια, οι διευθυντές διστάζουν να αναγνωρίσουν τα παράπονα των πελατών.

Οι αποκρίσεις των ξενοδοχείων σε ηλεκτρονικές κριτικές παίζουν σημαντικό ρόλο στη διαχείριση των πελατειακών σχέσεων και της εταιρικής φήμης. Οι Zhang και Vasquez (2014) ανέλυσαν 80 απαντήσεις ξενοδοχείων σε αρνητικές κριτικές και κατέληξαν ότι το ένα τρίτο των απαντήσεων δεν είναι σχετικού περιεχομένου, συγκεκριμένα οι διευθυντές της ίδιας αλυσίδα ξενοδοχείων έδωσαν την ίδια απάντηση σε διαφορετικές κριτικές.

Σε ορισμένες ακραίες περιπτώσεις οι αποκρίσεις σε δυο διαφορετικά είδη παραπόνων ήταν ίδιες (η μια κριτική αφορούσε στα κρεβάτια, ενώ η άλλη αφορούσε στην εξυπηρέτηση). Μόνο το 25% των απαντήσεων των ξενοδοχείων παρείχε λεπτομερείς εξηγήσεις ή συγκεκριμένα βήματα για την επίλυση των προβλημάτων. Αυτό το πόρισμα

είναι σύμφωνο με πρόσφατη μελέτη που πραγματοποίησαν οι Sparks και Bradley (2014) κατά την οποία λίγες αποκρίσεις ξενοδοχείων σε αρνητικές κριτικές περιλαμβάνουν εξηγήσεις. Στην πραγματικότητα, το 31% των απαντήσεων δεν κάνουν καν λόγο για αποτυχίες στην εξυπηρέτηση πελατών.

Ξενοδοχεία τα οποία αποκρίνονται συγκεκριμένα και επί του θέματος σε αρνητικές κριτικές εμπνέουν πολύ υψηλότερο επίπεδο εμπιστοσύνης και ποιότητας επικοινωνίας από εκείνα που αποκρίνονται μέσω τυποποιημένων μηνυμάτων. Οι απαντήσεις των ξενοδοχείων που περιλαμβάνουν δηλώσεις κατανόησης προς τα παράπονα των πελατών και παραφράσεις των αναφερθέντων προβλημάτων αυξάνουν την πιθανότητα αξιολόγησης τέτοιων αποκρίσεων ευνοϊκά. Επίσης ο χρόνος απόκρισης δεν φαίνεται να επηρεάζει σημαντικά την ικανοποίηση του πελάτη. Αυτό το γεγονός έρχεται σε αντίθεση με την κοινή πεποίθηση ότι η ταχύτητα απόκρισης είναι καθοριστικής σημασίας στη διαδικασία επανόρθωσης υπηρεσιών. Μια τυπική απάντηση από ξενοδοχείο μειώνει την αξιοπιστία του μηνύματός του επειδή οι χρήστες αντιλαμβάνονται ότι έγινε για διαφημιστικούς σκοπούς. Οι απαντήσεις των ξενοδοχείων επηρεάζουν τις επιδόσεις και τις αξιολογήσεις τους για διάφορα χαρακτηριστικά υπηρεσιών σε σχέση με τον αριθμό και την παραλλαγή των απαντηθέντων κριτικών.

Οι σαφείς απαντήσεις ενός ξενοδοχείου μπορεί επηρεάσουν θετικά τις κρατήσεις και τη βαθμολογία του. Αντίθετα, οι πομπώδεις απαντήσεις είναι πιθανό να έχουν αρνητικό αντίκτυπο στις κρατήσεις και τη βαθμολογία του αντίστοιχα. Ως εκ τούτου, οι διευθυντές των ξενοδοχείων θα πρέπει να απευθύνονται στα παράπονα των πελατών διαφορετικά κατά περίπτωση, κάνοντας μνεία στις υπηρεσίες υπό ερώτηση για να μη ληφθούν οι απαντήσεις ως πολύ γενικές ή απολογητικές.

Τέλος, οι ξενοδόχοι θα πρέπει να καθορίσουν προσεκτικά τη βάση κατά το σχεδιασμό των στρατηγικών απόκρισης, δηλαδή εάν η απάντηση θα στοχεύει απλώς στην βραχυπρόθεσμη επίλυση του προβλήματος ή στη μακροπρόθεσμη βελτίωση του ξενοδοχείου. Οι κριτικές των πελατών μπορούν να θεωρηθούν ως μέρος των στοιχείων που αποτελούν το στρατηγικό σχεδιασμό. Ενέργειες που μπορούν δυνητικά να ενισχύσουν την επιχειρησιακή αποδοτικότητα και την αποτελεσματικότητα ενός

ξενοδοχείου είναι η ενασχόληση με τους πελάτες και η ανάπτυξη καινοτόμων υπηρεσιών καθώς και οι δράσεις που μπορούν να αυξήσουν τη διαθεσιμότητα ειδικών πόρων για τη δημιουργία ενός ιδιαίτερα συνεργατικού εσωτερικού περιβάλλοντος. Παρόλο που είναι αδύνατη η εξάλειψη όλων των αρνητικών κριτικών, ο καλός ποιοτικός έλεγχος των υπηρεσιών, σε συνδυασμό με την επίλυση των παραπόνων και τη διαχείριση σχέσεων με τους πελάτες μετά την παροχή των υπηρεσιών είναι καθοριστικής σημασίας για την άμβλυνση των επιζήμιων επιπτώσεων των αρνητικών κριτικών.

Ιδέες και Μέθοδοι Ανάλυσης του Διαδικτύου και των Μέσω Κοινωνικής Δικτύωσης

Εισαγωγή

Τα περιεχόμενα που παράγονται από τους χρήστες (ΠΠΧ) με τη μορφή ανατροφοδοτήσεων, μπλογκς ή μηνυμάτων στα μέσα κοινωνικής δικτύωσης αποτελούν ολοένα και μεγαλύτερο μέρος του συνολικού διαδικτυακού περιεχομένου. Στον τουριστικό τομέα τα δεδομένα αυτά λαμβάνουν τη μορφή των μπλογκς (όπως είναι τα ταξιδιωτικά ημερολόγια ή αναφορές), μηνύματα σε μέσα κοινωνικής δικτύωσης όπως είναι το Facebook ή το Twitter ή οι κριτικές πελατών σε πλατφόρμες όπως το TripAdvisor ή το Booking.com. Ειδικά στον τουρισμό τα μέσα κοινωνικής δικτύωσης κερδίζουν μεγαλύτερη προσοχή και διαδραματίζουν σημαντικό ρόλο στη διαδικασία λήψης αποφάσεων των πελατών. Από τη μία πλευρά οι τουρίστες χρησιμοποιούν τα μέσα κοινωνικής δικτύωσης προκειμένου να εκφράσουν τη γνώμη τους για τα προϊόντα και τις υπηρεσίες που χρησιμοποίησαν και από την άλλη πλευρά για να ενημερωθούν για την ποιότητα και την καταλληλότητα του προϊόντος πριν από την κατανάλωση.

Το TripAdvisor είναι μία από τις σημαντικότερες πλατφόρμες αξιολόγησης ταξιδιών, στο οποίο έχουν αναρτηθεί πάνω από 250 εκατομμύρια κριτικές για πάνω από 5,2 εκατομμύρια τουριστικές επιχειρήσεις. Ενώ τα εργαλεία των μέσων κοινωνικής δικτύωσης όπως είναι το Facebook χρησιμοποιούνται από τους τουρίστες κυρίως για την ενημέρωση των φίλων και γνωστών για τις ταξιδιωτικές τους εμπειρίες, πλατφόρμες

ανασκόπησης όπως το TripAdvisor χρησιμοποιούνται συγκεκριμένα για την δημιουργία κριτικών της ποιότητας των προσφερόμενων προϊόντων και υπηρεσιών του τουρισμού. Τα ΠΠΧ και ειδικότερα οι κριτικές προϊόντων ασκούν μεγάλη επιρροή στις αποφάσεις των τουριστών. Το 90% των ταξιδιωτών κατά την διαδικασία προγραμματισμού του ταξιδιού τους μελετούν τις κριτικές άλλων ταξιδιωτών, το 87% των ταξιδιωτών δήλωσαν ότι οι κριτικές επηρέασαν την επιλογή τους και το 70% των ταξιδιωτών εμπιστεύονται τις διαδικτυακές συστάσεις. Τα ΠΠΧ θεωρούνται ως πιο ενημερωμένα, ευχάριστα και αξιόπιστα σε σχέση με τις πληροφορίες που προσφέρουν οι πάροχοι ταξιδιωτικών υπηρεσιών.

Ωστόσο, τα ΠΠΧ δεν είναι πολύτιμη πηγή γνώσης μόνο για τους πελάτες αλλά και για τους παρόχους τουριστικών υπηρεσιών. Οι κριτικές αποτελούν σαφείς και αμερόληπτες γνώμες πελατών σχετικά με την ποιότητα και την καταλληλότητα του προϊόντος και συνεπώς συμβάλουν στη βελτιστοποίηση του και στη διαχείριση των πελατειακών σχέσεων. Γενικά η ανάλυση των ΠΠΧ σημαίνει συγκέντρωση πληροφοριών σχετικά με την γνώμη που εξέφρασε ο κάθε χρήστης (θετική ή αρνητική) για συγκεκριμένο τουριστικό προϊόν ή υπηρεσία (η αίθουσα του ξενοδοχείου, το προσωπικό, η διατροφή). Ενώ τα κίνητρα ταξιδιού και τα δημογραφικά στοιχεία των χρηστών που αναρτούν κριτικές σε πλατφόρμες ανατροφοδότησης μπορούν να αποτυπωθούν εύκολα, η άποψη και τα θέματα πρέπει να εξαχθούν από το ελεύθερο κείμενο της κριτικής. Καθώς ο αριθμός των κριτικών έχει αυξηθεί δραματικά η χειροκίνητη ανάλυση τους δεν είναι πλέον εφικτή. Συνεπώς αυτόματες μέθοδοι εξαγωγής πληροφοριών από ελεύθερα κείμενα (που ονομάζονται ανάλυση συναισθημάτων ή εξόρυξη γνώμης) είναι απαραίτητες και αποτελούν σημαντικό αντικείμενο έρευνας τα τελευταία χρόνια. Οι συγκεκριμένοι μέθοδοι προσπαθούν να εξαγάγουν τις γνώμες των πελατών από κείμενα με την χρήση στατιστικών μεθόδων ή μεθόδων μηχανικής μάθησης. Επιπλέον αυτές οι μέθοδοι προσπαθούν να προσδιορίσουν το θέμα στο οποίο αναφέρεται η κριτική (ανίχνευση θεμάτων) και το συναίσθημα ή την πόλωση του χρήστη (ανίχνευση συναισθήματος).

Το παρόν κεφάλαιο παρέχει μια γενική εικόνα των διαφορετικών προσεγγίσεων της ανάλυσης συναισθημάτων στον τομέα του τουρισμού. Το κεφάλαιο είναι δομημένο ως εξής: Το πρώτο μέρος παρουσιάζει το γενικό θέμα της ανάλυσης των συναισθημάτων και συνοπτικά τις υπάρχουσες τεχνικές προσεγγίσεις. Το δεύτερο μέρος εξετάζει διαφορετικές προσεγγίσεις, που αποτελούν υποκατηγορία της ανάλυσης συναισθημάτων, για την ανίχνευση του θέματος της κριτικής. Το τρίτο μέρος ασχολείται με την ανίχνευση της υποκειμενικότητας και συναισθημάτων με στόχο τον προσδιορισμό του συναισθήματος ή της πόλωσης της κριτικής. Η παρακάτω ενότητα παρουσιάζει την εφαρμογή των προσεγγίσεων που συζητήθηκαν προηγουμένως και καταδεικνύει την εφαρμοσιμότητά τους και τα επιχειρηματικά οφέλη για τα πληροφοριακά συστήματα διαχείρισης προορισμών (ΠΣΔΠ).

1 Η Συναισθηματική Ανάλυση στον Τουριστικό Τομέα

Ο όγκος των κριτικών στο διαδίκτυο με τη μορφή κειμένου αυξάνεται ραγδαία. Συνεπώς οι μέθοδοι ανάλυσης του συγκεκριμένου είδους δεδομένων έλαβαν μεγαλύτερο ενδιαφέρον από τους ερευνητές και επαγγελματίες τα τελευταία χρόνια. Τα αντικείμενα της ανάλυσης συναισθημάτων μπορούν να χωριστούν σε δύο κατηγορίες, την ανίχνευση του θέματος και την ανίχνευση των συναισθημάτων. Η ανίχνευση θεμάτων έχει ως στόχο την αναγνώριση του θέματος της κριτικής. Για παράδειγμα, στην περίπτωση κριτικής του δωματίου ενός ξενοδοχείου το θέμα συνήθως είναι το ίδιο το δωμάτιο (το μέγεθός του, η θέα ή οι παροχές του). Έπειτα η ανίχνευση συναισθημάτων προσπαθεί να προσδιορίσει το συναίσθημα ή της πόλωση της κριτικής δηλαδή αν είναι θετική ή αρνητική ή απλώς αποτελεί μια ουδέτερη παρατήρηση. Η ανίχνευση συναισθήματος μπορεί να γίνει σε όλο το κείμενο της κριτικής, σε μία πρόταση ή ακόμα και σε συγκεκριμένο θέμα για το σύνολο των κριτικών.

Κάθε μία από τις παραπάνω εργασίες μπορούν να προσεγγιστούν με διαφορετική μέθοδο. Μια απλή αλλά δημοφιλής προσέγγιση είναι αυτή του λεξιλογίου. Βάσει ενός καταλόγου λέξεων μπορεί να προσδιοριστεί εάν ένα κείμενο έχει θετική ή αρνητική έννοια, με απλή μέτρηση του αριθμού θετικών ή αρνητικών λέξεων. Παρά τη λιτότητα της συγκεκριμένης προσέγγισης η εφαρμογή της δείχνει καλά αποτελέσματα

υπό ορισμένες συνθήκες καθώς έχει ως βάση ένα σαφώς καθορισμένο λεξικό. Επιπλέον, δημοφιλείς μέθοδοι μηχανικής μάθησης χρησιμοποιούνται ως βοηθήματα τόσο για την επιβλεπόμενη μάθηση όσο και τη μη επιβλεπόμενη μάθηση. Η επιβλεπόμενη μάθηση είναι η εκμάθηση του λεξιλογίου από μια προ-ταξινομημένη βάση εκπαιδευτικών δεδομένων, στην περίπτωση της ανίχνευσης θέματος η ανίχνευση επιτυγχάνεται μέσω της ταξινόμησης του θέματος της κριτικής σε μια από τις προκαθορισμένες κατηγορίες της βάσης δεδομένων η οποία αντλήθηκε από προ-ταξινομημένα εκπαιδευτικά δεδομένα. Από την άλλη, η μη επιβλεπόμενη μάθηση κάνοντας χρήση μοντέλων όπως η παραγοντική ανάλυση ή η ανάλυση κατά συστάδες μαθαίνει κατά την διαδικασία της επίλυσης του έργου. Παραδείγματος χάρη, η ανίχνευση του θέματος ενός συνόλου κριτικών επιτυγχάνεται με την αναγνώριση λέξεων που έχουν ξαναπαρουσιαστεί (χρήση παραγοντικής ανάλυσης - ανάλυσης κατά συστάδες) και την ταξιθέτηση τους και σε ένα πιθανό θέμα. Ενώ οι παραπάνω προσεγγίσεις αντιμετωπίζουν τις κριτικές ως σύνολα λέξεων, οι σημασιολογικές προσεγγίσεις προσπαθούν να καθορίσουν την έννοια της κάθε λέξης μέσω επεξεργασίας της δομής ολόκληρης της πρότασης. Προφανώς προκειμένου να αξιοποιηθούν τα αντίστοιχα πλεονεκτήματα της κάθε προσέγγισης, συγκεκριμένες εφαρμογές της συναισθηματικής ανάλυσης συχνά συνδυάζουν τις παραπάνω προσεγγίσεις με αποτέλεσμα τη δημιουργία των λεγόμενων υβριδικών προσεγγίσεων.

Βάσει των παραπάνω, θα προσαρμόζουμε την ταξινόμηση που πρότειναν οι Tsytsarau και Palpanas (2011) και θα ομαδοποιήσουμε την συναισθηματική ανάλυση σε πέντε μεγάλες κατηγορίες: 1) την επιβλεπόμενη μηχανική μάθηση, 2) τη βάση λεξιλογίου, 3) τη μη επιβλεπόμενη μηχανική μάθηση, 4) τη σημασιολογική και 5) την υβριδική προσέγγιση.

1α Η Επιβλεπόμενη Μηχανική Μάθηση

Η μελέτη των Ye, Zhang και Law (2009) κάνει χρήση των εργαλείων συναισθηματικής κατηγοριοποίησης στον τομέα της συλλογής κριτικών από ταξιδιωτικά μπλογκς. Η μελέτη συγκρίνει τρεις τεχνικές επιβλεπόμενης μηχανικής μάθησης, τη Naive Bayes, τις μηχανές διανυσμάτων υποστήριξης (ΜΔΥ) και τη γράφο N-

γραμμάτων (ΓΝΓ) για την κατηγοριοποίηση των συναισθημάτων των κριτικών από ταξιδιωτικά μπλογκς για δημοφιλείς ταξιδιωτικούς προορισμούς των ΗΠΑ και Ευρώπης. Τα αποτελέσματα δείχνουν ότι προσεγγίσεις ΜΔΥ και ΓΝΓ σημείωσαν καλύτερες επιδόσεις από τη Naeve Bayes. Ενδιαφέρον παρουσιάζει το γεγονός ότι αν τα εκπαιδευτικά δεδομένα είχαν σχετικά μεγάλο όγκο κριτικών και οι τρεις προσεγγίσεις θα έφθαναν επίπεδα ακρίβειας τουλάχιστον 80% (Ye, Zhang και Law, 2009).

Η μελέτη των Lin και Chao (2010) επικεντρώνεται στη συλλογή απόψεων με την χρήση δεδομένων από σχόλια σε τουριστικά μπλογκς. Συγκεκριμένα, οι σχολιαστές κλήθηκαν να σημειώσουν την πόλωση και το αντικείμενο της άποψης της κάθε πρότασης ξεχωριστά. Ακολούθως, εφαρμόστηκαν τεχνικές μηχανικής μάθησης για την εκπαίδευση των ταξινομητών. Τα αποτελέσματα ακρίβειας και ανάκλησης της ανίχνευσης συναισθήματος στον τουρισμό αντιστοιχούσαν σε 56% και 60% αντίστοιχα. Αντίθετα τα αποτελέσματα αναγνώρισης αντικειμένου (ανίχνευση θέματος) σε προτάσεις με ξεκάθαρα διατυπωμένες απόψεις κυμαίνονται στο 91% και 90% αντίστοιχα.

Η μελέτη των Kasper και Vela (2011) χρησιμοποιεί ένα υπάρχον λεξικό για την συνδρομή της συναισθηματικής ανάλυσης σε σύνολο κριτικών ξενοδοχείου (Waltinger, 2010). Προκειμένου να επιτευχθεί ο στόχος της ανίχνευσης συναισθημάτων, τα αποθηκευμένα 7200 λεξικά τμήματα του λεξικού χρησιμοποιούνται για την εκπαίδευση των ταξινομητών μηχανικής μάθησης με δύο κατηγορίες πόλωσης (θετικές - αρνητικές). Μια τελική επαλήθευση αποδεικνύει την ικανοποιητική απόδοση της ακρίβειας του μοντέλου.

Ομοίως με τη μελέτη των Ye, Zhang και Law (2009), οι Alves, Baptista, Firmino, de Oliveira και de Paiva (2014) συγκρίνουν τις ΜΔΥ με τους ταξινομητές Naeve Bayes για την ανάλυση των συναισθημάτων στα tweets του Κύπελλου Συνομοσπονδιών της FIFA το 2013. Τα αποτελέσματα έδειξαν επανειλημμένα ότι οι ΜΔΥ αποδίδουν καλύτερα από τη μέθοδο Naeve Bayes (Alves, Baptista, Firmino, de Oliveira και de Paiva, 2014). Οι Markopoulos, Mikros, Iliadi και Lontos (2015) δημιουργούν έναν ταξινομητή για την ανίχνευση συναισθήματος εφαρμόζοντας τη μέθοδο μηχανικής

μάθησης των ΜΔΥ σε κριτικές ξενοδοχείων γραμμένες στην ελληνική γλώσσα. Τα αποτελέσματα ήταν ικανοποιητικά μετά τη χρήση μονογραμματικού μοντέλου γλώσσας. Τέλος, η μελέτη των Pablos, Cuadros και Linaza (2015) παρουσιάζει το Ευρωπαϊκό έργο OreNER, ένα σύνολο έτοιμων εργαλείων ανοικτής πρόσβασης ανάλυσης κειμένου (ΜΔΥ) για την εκτέλεση εργασιών επεξεργασίας κειμένου, όπως η ανίχνευση άποψης. Επιπλέον η μελέτη προσφέρει μια ενδιαφέρουσα εφαρμογή του OreNER ως προς τη γεωγραφική θέση των κριτικών των ξενοδοχείων.

1β Μάθηση Βάσει Λεξιλογίου

Οι Garcia, Gaines και Linaza (2012) παρουσιάζουν μια προσέγγιση ανάλυσης συναισθημάτων βάσει λεξιλογίου στον τομέα του τουρισμού. Η μελέτη παρουσιάζει τη χρήση μιας λεξικής βάσης δεδομένων για τη συναισθηματική ανάλυση των κριτικών του TripAdvisor στους τομείς της εστίασης, των τροφίμων και ποτών. Κάνοντας χρήση μιας λεξικής βάσης δεδομένων που απαρτίζεται από άνω των 6000 λέξεων μπορεί να υπολογίσει το επίπεδο συναισθήματος μίας κριτικής βάσει των αρνητικών και θετικών λέξεων που εμφανίζονται σε αυτήν. Χρησιμοποιώντας το λεξικό οι προτάσεις που περιλαμβάνονται σε μια κριτική σημειώνονται για το είδος της πόλωσης τους. Τέλος, η εν λόγω προσέγγιση περιλαμβάνει μια ταξινόμηση για την κατηγοριοποίηση των φράσεων ανά θέμα χρησιμοποιώντας μια λίστα λημματοποιημένων και κανονικοποιημένων λέξεων, κάθε μία από τις οποίες ανήκει σε διαφορετική θεματική κατηγορία.

Ομοίως, μια μελέτη από τους Grabner, Zanker, Fliedl και Fuchs (2012) προτείνει ένα σύστημα που ταξινομεί τις ξενοδοχειακές κριτικές. Κατά τη διαδικασία της κατάρτιση συντάσσεται ένα λεξικό ειδικού-πεδίου, σημασιολογικά συναφών λέξεων που απορρέουν από το σύνολο των κριτικών. Το λεξικό που προκύπτει βοηθάει τη συναισθηματική ανάλυση παράγοντας μια ταξινόμηση κριτικών. Η αξιολόγηση της ταξινόμησης αυτής σε εκπαιδευτικά δεδομένα δείχνει ότι η προτεινόμενη προσέγγιση αποδίδει καλύτερα εν συγκρίσει με προκαθορισμένες αναφορές: εάν η κριτική ταξινομείται χειροκίνητα ως καλή ή κακή, η ταξινόμηση θα είναι σωστή κατά 90%.

1γ Η Μη Επιβλεπόμενη Μηχανική Μάθηση

Στη μελέτη των Xiang, Schwartz και Uysal (2015) διερευνάται η χρησιμότητα των διαστάσεων των εμπειριών βάσει πραγματικών κριτικών πελατών προκειμένου να προσδιοριστεί ποία ξενοδοχεία αφήνουν τους πελάτες τους ικανοποιημένους και ποια όχι. Χρησιμοποιώντας ανάλυση κατά συστάδες (μη επιβλεπόμενη μηχανική μάθηση) τα ξενοδοχεία ομαδοποιούνται βάσει βαθμολογιών και ικανοποίησης πελατών. Έπειτα, κάνοντας χρήση ανάλυσης αντιστοιχιών τα ομαδοποιημένα ξενοδοχεία εξετάζονται σε σχέση με τις θεματικές λέξεις που εμπεριέχονται στις κριτικές πελατών. Τα αποτελέσματα καταδεικνύουν την ύπαρξη διαφορετικών τύπων ξενοδοχείων με μοναδικά χαρακτηριστικά που ικανοποιούν τους πελάτες τους, από την άλλη τα ξενοδοχεία που δεν κατάφεραν να ικανοποιήσουν τους πελάτες δείχνουν να έχουν θέματα με την καθαριότητα και τη συντήρηση. Η μελέτη αυτή υποδεικνύει μια πολλά υποσχόμενη κατεύθυνση προς την χρήση πραγματικών δεδομένων καταναλωτικών εμπειριών για τη στήριξη της αντιληπτικής χαρτογράφησης και της κατάτμησης της αγοράς στον κλάδο της φιλοξενίας.

Οι Rossetti, Stella, Cao και Zanker (2015) μελετούν εναλλακτικούς τρόπους ανάλυσης τουριστικών κριτικών μέσω θεματικών μοντέλων. Η συγκεκριμένη μέθοδος χρησιμοποιεί στατιστική προσέγγιση και τη δυνατότητα επεξεργασίας κριτικών σε μορφή κειμένου. Εκτός από τη συμβολή ενός νέου μοντέλου στις υπάρχουσες μεθόδους θεματικών μοντέλων, αυτή η μελέτη περιλαμβάνει και εμπειρικά στοιχεία από μελέτες σε κριτικές χρηστών από βάσεις δεδομένων του YELP καθώς και του TripAdvisor.

1δ Η Σημαιολογική Προσέγγιση

Οι Kasper και Vela (2012) παρουσιάζουν ένα σύστημα που παρακολουθεί αυτόματα τις κριτικές και τα σχόλια χρηστών σε ξενοδοχεία από διάφορες πηγές μέσω κοινωνικής δικτύωσης χρησιμοποιώντας σημασιολογικές τεχνικές. Ως σημαντική βάση πληροφοριών του ποιοτικού ελέγχου ενός ξενοδοχείου, το σύστημα παρέχει ταξινομημένες περιλήψεις θετικών και αρνητικών χαρακτηριστικών του ξενοδοχείου. Η μελέτη των Xiang, Schwartz, Gerdes και Uysal (2015) διερευνά την χρησιμότητα της ανάλυσης μεγάλων δεδομένων για την καλύτερη κατανόηση της σχέσης μεταξύ της

εμπειρίας και της ικανοποίησης που λαμβάνει ο πελάτης ενός ξενοδοχείου. Η μελέτη τους περιλαμβάνει την ανάλυση μεγάλου όγκου κριτικών της Expedia για τη λεπτομερή επεξεργασία της εμπειρίας διαμονής σε ξενοδοχείο σε σχέση με την αξιολόγηση της ικανοποίησης. Τα αποτελέσματα αποκαλύπτουν διάφορες διαστάσεις της εμπειρίας των πελατών που έφεραν διαφορετική βαρύτητα και συνεπώς έλαβαν νέες σημαντικές σημασιολογικές συνθέσεις. Η σημασιολογική συσχέτιση ανάμεσα στην εμπειρία του πελάτη και την ικανοποίηση φαίνεται ισχυρή υποδηλώνοντας ότι οι δυο αυτοί τομείς της καταναλωτικής συμπεριφοράς είναι εγγενώς συνδεδεμένοι.

1ε Η Υβριδική Προσέγγιση

Τελευταία από τις κατηγορίες συναισθηματικής ανάλυσης που θα καλύψουμε στην παρούσα μελέτη είναι η υβριδική προσέγγιση. Η συγκεκριμένη προσέγγιση συνδυάζει την επιβλεπόμενη μηχανική μάθηση, τη μάθηση βάσει λεξιλογίου, τη μη επιβλεπόμενη μηχανική μάθηση και τη σημασιολογική προσέγγιση. Μια μελέτη που έγινε από τους Waldhor και Rind (2008) έκανε χρήση ενός συνδυασμού προσεγγίσεων όπως της μάθησης βάσει λεξιλογίου για την ανίχνευση του θέματος, της μηχανικής μάθησης για τη συλλογή απόψεων και ορισμένες πτυχές σημασιολογικής προσέγγισης. Το προτεινόμενο ημιαυτόματο λογισμικό που θα αναλύει τουριστικά μπλογκς, περιλαμβάνει διαδικασίες κατηγοριοποίησης κειμένου, ανίχνευσης και εξαγωγής συναισθημάτων. Πιο συγκεκριμένα, συνδυάζει τη μέθοδο γλωσσικής ανάλυσης με τις μεθόδους εξαγωγής πληροφοριών και ορολογίας προκειμένου να προσδιορίσει την πόλωση και την ένταση των εκφράσεων που χρησιμοποιούνται. Κατ' αυτόν τον τρόπο, η συγκεκριμένη προσέγγιση αποδεικνύεται ιδιαίτερα χρήσιμη στην εξέταση των σημασιολογικών πτυχών όπως είναι η άρνηση (όπως "όχι") ή λέξεις που προσδίδουν έμφαση σε μια φράση (όπως "πολύ").

Σε άρθρο των Weichselbraun, Gindl και Scharl (2013), συστήθηκε μια υβριδική προσέγγιση που συνδυάζει την ανάλυση βάσει λεξιλογίου με την ευελιξία της μηχανικής μάθησης για την επίλυση ζητημάτων ασάφειας και την εξέταση του θεματικού πλαισίου των συναισθηματικών όρων. Η προτεινόμενη μέθοδος προσδιορίζει διαφορετικούς όρους που ποικίλλουν ως προς την πόλωση ανάλογα με το νόημα της πρότασης και

στη συνέχεια τις αποθηκεύει και τις ταξινομεί (ανάλογα με το συναίσθημα που εκφράζουν) σε λεξικές βάσεις δεδομένων. Σε συνδυασμό με τις σημασιολογικές βάσεις δεδομένων, αυτά τα λεξικά συμβάλλουν στη συσχέτιση των αμφιλεγόμενων συναισθηματικών όρων με έννοιες που αντιστοιχούν σε ανάλογη πόλωση.

Σε πρόσφατη μελέτη των Schmunk, Hocken, Fuchs και Lexhagen (2014) παρουσιάζεται μια υβριδική προσέγγιση για την συλλογή πληροφοριών που συνδράμουν στη λήψη αποφάσεων μέσω ΠΠΧ και συγκρίνουν διαφορετικές τεχνικές εξόρυξης δεδομένων (ΕΔ) ανάλογα με την ακρίβειά προσδιορισμού της πόλωσης των απόψεων των χρηστών και της απόδοσης απόψεων σε θέματα κριτικών. Πιο συγκεκριμένα, η μελέτη στοχεύει στην κατανόηση της συνολικής διαδικασίας εξόρυξης πληροφοριών από τις κριτικές πελατών που δημοσιεύτηκαν σε τουριστικές πλατφόρμες ανασκόπησης όπως το TripAdvisor και το Booking και στη σύγκριση διαφόρων τεχνικών ΕΔ (μάθηση βάσει λεξιλογίου, μηχανικής μάθησης όπως η Naëve Bayes και ΜΔΥ) για τον προσδιορισμό τόσο του θέματος όσο και του συναισθήματος της άποψης. Οι προτεινόμενες τεχνικές αξιολογούνται με γνώμονα την ποιότητα των συγκεντρωμένων πληροφοριών, την ακρίβεια τους και την χρησιμότητά τους σε ένα πληροφοριακό σύστημα διαχείρισης προορισμών.

Σε άλλη πρόσφατη μελέτη, οι Chiu, Chiu, Sunga, and Hsieh (2015) παρουσίασαν μια μέθοδο εξόρυξης απόψεων με βάση την ταξινόμηση των συναισθημάτων για την συλλογή του e-word-of-mouth από μπλογκς. Για την εξόρυξη απόψεων εφαρμόστηκε ο αλγόριθμος σημειακής αμοιβαίας πληροφορίας (ΣΑΠ) ώστε να αναγνωριστούν λέξεις που σχετίζονται με θετικές ή αρνητικές έννοιες. Επιπροσθέτως, χρησιμοποιήθηκε ένας ευρετικός κανόνας N-φράσεων για τον εντοπισμό των απόψεων των πελατών σχετικά με τα χαρακτηριστικά του ξενοδοχείου, όπως η εικόνα του ξενοδοχείου, οι υπηρεσίες, η σχέση ποιότητας - τιμής, τα τρόφιμα και τα ποτά, τα δωμάτια, οι παροχές και η τοποθεσία. Τα αποτελέσματα φάνέρωσαν ότι η προτεινόμενη υβριδική προσέγγιση παρουσιάζει αποτελεσματική επίδοση στις εργασίες ταξινόμησης και πρόβλεψης. Τέλος, ένας αντιληπτικός χάρτης που βασίζεται στην ανάλυση αντιστοιχιών απεικονίζει

συγκρίσεις απόψεων και μας δίνει μια εικόνα για τον ανταγωνισμό μεταξύ των ξενοδοχείων.

2 Η Ανίχνευση Θέματος

Η ανίχνευση θέματος μπορεί να πραγματοποιηθεί είτε υπό επίβλεψη (επιβλεπόμενη) είτε δίχως επίβλεψη (μη επιβλεπόμενη). Στην πρώτη περίπτωση, τα θέματα είναι προκαθορισμένα και ως εκ τούτου ο αριθμός των θεμάτων είναι περιορισμένος. Παραδείγματα τέτοιων προκαθορισμένων θεμάτων στην περίπτωση των ξενοδοχειακών κριτικών είναι τα δωμάτια, τα τρόφιμα και ποτά, η υπηρεσία και το προσωπικό, η τοποθεσία κλπ. Η επιβλεπόμενη ανίχνευση θέματος συνήθως πραγματοποιείται στο επίπεδο μιας δήλωσης (σε πρόταση) που περιλαμβάνει μια κριτική, καθώς τα πλήρη κείμενα των κριτικών τείνουν να έχουν περισσότερα από ένα θέματα. Πιθανές μέθοδοι που θα μπορούσαν να αναλάβουν την επιβλεπόμενη ανίχνευση θέματος είναι η βάση λεξιλογίου ή η επιβλεπόμενης μηχανικής μάθησης (τεχνικές ταξινόμησης όπως Naive Bayes ή ΜΔΥ). Ένα σαφές όφελος της επιβλεπόμενης ανίχνευσης θέματος είναι ότι τα θέματα είναι προκαθορισμένα και επομένως συγκρίσιμα με όλες τις κριτικές καθώς και το γεγονός ότι μπορούν να αποτελέσουν σημείο αναφοράς για μελλοντικές αναλύσεις.

Στην περίπτωση της ανίχνευσης μη επιβλεπόμενου θέματος, τα θέματα δεν είναι προκαθορισμένα ωστόσο τυχόν νέα θέματα που μπορεί να προκύπτουν από τα σχόλια και τις κριτικές μπορούν να προσδιοριστούν. Κατά συνέπεια ο αριθμός των θεμάτων στην προσέγγιση αυτή είναι απεριόριστος και ξεπερνάει κατά πολύ τον αριθμό προκαθορισμένων θεμάτων. Η μέθοδος μη επιβλεπόμενης ανίχνευσης θέματος συνήθως αναλύει την κάθε λέξη ξεχωριστά και επομένως μπορεί να εντοπίσει πολλαπλά θέματα σε μία πρόταση (η ανάλυση επίσης μπορεί να πραγματοποιηθεί και σε επίπεδο προτάσεων το οποίο έχει νόημα ιδιαίτερα αν η ανίχνευση θέματος θα συνδυαστεί με την ανίχνευση συναισθήματος). Πιθανές προσεγγίσεις που θα μπορούσαν να υιοθετηθούν στην περίπτωση της ανίχνευσης μη επιβλεπόμενου θέματος είναι η μη επιβλεπόμενη μηχανική μάθηση, οι στατιστικές προσεγγίσεις όπως η συσταδοποίηση και παραγοντική ανάλυση ή η επιβλεπόμενη μηχανική μάθηση. Ένα

σαφές πλεονέκτημα της μη επιβλεπόμενης ανίχνευσης θέματος είναι ότι τα πιο σημαντικά θέματα που απασχολούν τους χρήστες μπορούν να προσδιοριστούν αυτομάτως χωρίς να χρειάζεται να προκαθοριστούν εκ των προτέρων. Ακολούθως, νέα θέματα, αλλαγές θεμάτων η τάσεις θεματολογίας μπορούν να αναγνωριστούν και να προσδιοριστούν.

2α Επιβλεπόμενη Ανίχνευση Θέματος

Η επιβλεπόμενη ανίχνευση θέματος εκτελείται συνήθως είτε με τη μέθοδο βάσει λεξιλογίου, είτε με τη μέθοδο ταξινόμησης (επιβλεπόμενη μηχανική μάθησης). Η ανίχνευση θέματος βάσει λεξιλογίου σημαίνει ότι για την κάθε κατηγορία θέματος ορίζεται ένα λεξικό ή ένας κατάλογος λέξεων, που περιλαμβάνει ένα σύνολο αντιπροσωπευτικών λέξεων ή τίτλων. Στη συνέχεια προσδιορίζεται κατά πλειοψηφία το θέμα μιας πρότασης με την καταμέτρηση του αριθμού των λέξεων από τον κάθε λεξικό κατάλογο. Στα πλαίσια μιας πρωτότυπης εφαρμογής ανάλυσης συναισθημάτων για το σουηδικό ορεινό προορισμό Are, εξετάστηκε η προσέγγιση βάσει λεξιλογίου σε 210 κριτικές καταλυμάτων της TripAdvisor και Booking.com, αποτελούμενες από 1520 μεμονωμένες προτάσεις. Οι κατάλογοι λέξεων που ορίστηκαν αφορούσαν στη διατροφή / πρωινό, ξενοδοχείο, δωμάτιο, υπηρεσία / προσωπικό, τοποθεσία και ευεξία (που περιλαμβάνουν μεταξύ τριών και επτά λέξεων). Τα αποτελέσματα συγκρίθηκαν με αντίστοιχες ταξινομήσεις που πραγματοποιήθηκαν χειροκίνητα και η προσέγγιση βάσει λεξιλογίου έφθασε σε ακρίβεια το 72%.

Από την άλλη πλευρά, οι επιβλεπόμενες προσεγγίσεις μηχανικής μάθησης ακολουθούν τη μέθοδο του προσδιορισμού του θέματος μιας διατύπωσης (δήλωση μιας κριτικής) από τα χαρακτηριστικά της (οι λέξεις του κειμένου της κριτικής) βάσει προ-ταξινομημένων εκπαιδευτικών δεδομένων. Το κείμενο της κριτικής αποτελεί απλώς ένα σύνολο - διάνυσμα λέξεων που βασίζεται στην ύπαρξη λέξεων ή πιο συγκεκριμένα στις τιμές TF-IDF / ΣΕ-ΑΣΕ (συχνότητα εμφάνισης-αντίστροφη συχνότητα εγγράφου), οι οποίες αντικατοπτρίζουν το βάρος (σημαντικότητα) μιας συγκεκριμένη λέξης - όρου σε ένα κείμενο. Στη συνέχεια, ο ταξινομητής μάθησης αποφασίζει ποιο θέμα αντιστοιχεί καλύτερα στην κάθε πρόταση του κειμένου της κριτικής με βάση τη συχνότητα

εμφάνισής συγκεκριμένων λέξεων. Ως επιβλεπόμενες προσεγγίσεις μηχανικής μάθησης, συγκρίθηκαν η μηχανή διανυσμάτων υποστήριξης (ΜΔΥ), η Naeve Bayes (και οι δύο κατάλληλες για τη διαδικασία κατηγοριοποίησης κειμένων) και ο κατηγοριοποιητής k-εγγύτατου γείτονα (k-NN). Το αποτέλεσμα της σύγκρισης αυτής έδειξε ότι η ΜΔΥ ξεπερνάει με 73% τόσο τον k-NN (57%) όσο και τη Naeve Bayes (50%). Σε όλες τις περιπτώσεις, η σημείωση μερών του λόγου ΕΜΛ (POS) χρησιμοποιήθηκε για την εξαγωγή των ουσιαστικών του κειμένου της κριτικής, το οποίο βελτίωσε την ακρίβεια της προσέγγισης ΜΔΥ.

Συνοψίζοντας, η ΜΔΥ ως προσέγγιση επιβλεπόμενης μηχανικής μάθησης πέτυχε τα καλύτερα αποτελέσματα, ειδικά σε σύγκριση με τις άλλες προσεγγίσεις μηχανικής μάθησης (Πίνακας 2). Με γνώμονα την απλότητα, παρόλο που η προσέγγιση βάσει λεξιλογίου είναι ελαφρώς κατώτερη σε σύγκριση με τη ΜΔΥ, εξακολουθεί να αποτελεί τη βέλτιστη επιλογή σε πρακτικές περιπτώσεις.

Προσέγγιση Επιβλεπόμενης Ανίχνευσης Θέματος	Ορθότητα (%)
Βάσει λεξιλογίου	72
k-NN	57
ΜΔΥ	73
Naeve Bayes	50

Πίνακας 2 Αξιολόγηση των προσεγγίσεων επιβλεπόμενης ανίχνευσης θέματος

2β Μη Επιβλεπόμενη Ανίχνευση Θέματος

Σκοπός της ανίχνευσης μη επιβλεπόμενου θέματος είναι ο προσδιορισμός των θεμάτων που απασχολούν οι κριτικές χωρίς την ανάγκη προκαθορισμού θεμάτων. Η ανίχνευση μη επιβλεπόμενου θέματος συνήθως εκτελείται με τη βοήθεια στατιστικών μεθόδων όπως η παραγοντική ανάλυση ή οι τεχνικές μηχανικής μάθησης, όπως η συσταδοποίηση και η σειριακή εξόρυξη προτύπων. Η τεχνική των συχνών λέξεων είναι μια απλή προσέγγιση εάν το πρόβλημα της ανίχνευσης θέματος είναι η αναγνώριση

λέξεων που προσδιορίζουν το θέμα. Η βασική υπόθεση είναι απλή, αν ορισμένες αναφορές κριτικών σχετίζονται με το ίδιο προϊόν (πχ ξενοδοχείο), τότε είναι πολύ πιθανό να ξανά-αναφερθούν στο ίδιο θέμα. Συνεπώς λέξεις σχετιζόμενες με το θέμα θα εμφανίζονται συχνότερα. Αντιθέτως λέξεις μη-σχετιζόμενες με το θέμα θα εμφανίσουν μεγαλύτερη ετερογένεια, με αποτέλεσμα κάθε μία από αυτές να εμφανίζεται σπανιότερα από τις σχετιζόμενες λέξεις. Αντιστρόφως μπορούμε να συμπεράνουμε ότι οι λέξεις που χρησιμοποιούνται συχνότερα (τα ουσιαστικά) σε πολλές περιπτώσεις προδίδουν το θέμα της κριτικής. Όταν η προσέγγιση αυτή εξετάστηκε στις κριτικές των ξενοδοχείων του Are, απέδωσε με ορθότητα (accuracy) 83%. Η ορθότητα υπολογίζεται συγκρίνοντας τα αποτελέσματα με χειροκίνητα σημειωμένα εκπαιδευτικά δεδομένα, δηλαδή τις κριτικές στις οποίες οι σωστές ετικέτες λέξεων (μέρη του λόγου) έχουν σημειωθεί από ανθρώπους. Ωστόσο τα αποτελέσματα των μετρικών της ακρίβειας (precision) στο 54% και της ανάκλησης (recall) στο 95% αποκαλύπτουν ότι υπερβολικά πολλές λέξεις αναγνωρίζονται ως θέμα κάτι που αποτελεί περιορισμό της προσέγγισης.

Η ανάλυση συστάδων χρησιμοποιείται όταν δύο προτάσεις της ίδιας κριτικής αναφέρονται στο ίδιο θέμα υπό την υπόθεση ότι αποτελούνται από τις ίδιες σχετιζόμενες με το θέμα λέξεις. Εάν συνδυάζουμε (συσταδοποιούμε) τις προτάσεις με βάση τις λέξεις από τις οποίες αποτελούνται (κάθε συστάδα αντιπροσωπεύει έναν συνδυασμό λέξεων που εμφανίζεται συχνά) τότε οι συστάδες μπορούν να θεωρηθούν ως δυνητικά θέματα και κάθε θέμα να περιγράφεται από τις λέξεις που εμφανίζονται στις αντίστοιχες συστάδες (Kiran, Shankar και Pudi, 2010). Αντίστοιχα με την παραπάνω προσέγγιση θα είχε νόημα να περιορίσουμε την ανάλυση σε ουσιαστικά και πιο συγκεκριμένα σε αντιπροσωπευτικά ουσιαστικά, τα αποκαλούμενα ως λέξεις-κλειδιά. Τα ουσιαστικά φιλτράρονται με τη μέθοδο EMΛ και τα σημαντικά ουσιαστικά με τη μέθοδο ΣΕ-ΑΣΕ. Οι λέξεις-κλειδιά, που ανήκουν στο προσδιορισμένο θέμα μπορούν στη συνέχεια να επισημανθούν ως λέξεις που σχετίζονται με τα θέματα των κριτικών. Η αξιολόγηση της προσέγγισης γίνεται ξανά με τη σύγκριση των αποτελεσμάτων με χειροκίνητα σημειωμένες προτάσεις κριτικών. Μια συσσώρευση k-means 80 συστάδων (και η συνημιτονοειδής ομοιότητα ως μετρητής απόστασης) αποδίδει ορθότητα 89%.

Η λανθάνουσα σημασιολογική ανάλυση (ΛΣΑ) είναι μια άλλη πολύ γνωστή μέθοδος ανίχνευσης θεμάτων. Αντίστοιχα με την ανάλυση συστάδων που περιγράψαμε παραπάνω, υποθέτουμε ότι οι κριτικές που αναφέρονται στο ίδιο θέμα θα αποτελούνται από τις ίδιες σχετιζόμενες με το θέμα λέξεις. Ωστόσο, η προσέγγιση ΛΣΑ βασίζεται στην αρχή της ελαχιστοποίησης των διαστάσεων. Διαστάσεις είναι οι μεμονωμένες λέξεις που εμφανίζονται στις προτάσεις κριτικών και οι λέξεις που συχνά συνυπάρχουν στην ίδια πρόταση, έπειτα αυτές ομαδοποιούνται σε παράγοντες μέσω παραγοντικής ανάλυσης ή πιο συγκεκριμένα μέσω διάσπασης σε ιδιάζουσες τιμές (ΔΙΤ - SVD). Οι παράγοντες που προκύπτουν έπειτα θα συμβολίζουν τα λανθάνοντα θέματα. Η εκτίμηση της παραπάνω προσέγγισης για τις κριτικές ξενοδοχείων του Are, με την ΛΣΑ να περιλαμβάνει 40 παράγοντες έφτασε το ποσοστό ορθότητας 88%.

Οι παραπάνω προσεγγίσεις ανίχνευσης θεμάτων χειρίζονται τα κείμενα ανασκόπησης ως σύνολα λέξεων, δηλαδή δεν λαμβάνεται υπόψη η σειρά των λέξεων και η θέση τους μέσα σε στην πρόταση. Η βασική ιδέα της σειριακής εξόρυξης προτύπων για την ανίχνευση θεμάτων είναι να λαμβάνεται υπόψη το εννοιολογικό πλαίσιο της κάθε λέξης, δηλαδή τις λέξεις που προηγούνται και ακολουθούν. Συνεπώς οι προτάσεις που συγκροτούν τις κριτικές παύουν να προσδιορίζονται ως διανύσματα λέξεων (σύνολα λέξεων) αλλά η κάθε λέξη αποθηκεύεται με το εννοιολογικό της πλαίσιο, δηλαδή αποθηκεύεται και ένας αριθμός λέξεων που πλαισιώνουν την αποθηκευμένη λέξη (προηγούμενες και επόμενες) καθώς και τα χαρακτηριστικά τους όπως η θέση της μέσα στην πρόταση, το μήκος τους κλπ. Προκειμένου η τεχνική σειριακής εξόρυξης προτύπων να προσδιορίσει τις λέξεις που προδίδουν το θέμα μιας πρότασης ανάλογα με το εννοιολογικό τους πλαίσιο και τα χαρακτηριστικά τους, απαιτούνται εκπαιδευτικά δεδομένα με ήδη σημειωμένες λέξεις που σχετίζονται με το θέμα, αντίστοιχες με τα εκπαιδευτικά δεδομένα που χρησιμοποιήθηκαν στις προηγούμενες μεθόδους. Τα εκπαιδευτικά σειριακά πρότυπα που καθιστούν δυνατή την αναγνώριση των λέξεων που σχετίζονται με το θέμα, μπορούν να κάνουν χρήση οποιασδήποτε τεχνικής ταξινόμησης. Παρόλο που τέτοιες τεχνικές ταξινόμησης εμπίπτουν στην κατηγορία της επιβλεπόμενης μάθησης, η γενική προσέγγιση

εξακολουθεί να αποτελεί περίπτωση μη επιβλεπόμενης ανίχνευσης θεμάτων καθώς δεν υπάρχουν προκαθορισμένα θέματα και οι λέξεις δεν είναι ταξινομημένες σε συγκεκριμένα θέματα (όπως δωμάτιο, παροχές, τοποθεσία, εξυπηρέτηση κλπ.). Η συγκεκριμένη προσέγγιση εφαρμόστηκε στις κριτικές των ξενοδοχείων του Age, χρησιμοποιώντας τη Naive Bayes ως τεχνική ταξινόμησης και ένα πλαίσιο λέξεων δύο προηγούμενων και επόμενων λέξεων με αποτέλεσμα την επίτευξη ποσοστού ορθότητας της τάξεως των 93%. Ωστόσο το αρκετά υψηλό ποσοστό ορθότητας οφείλεται στο γεγονός ότι ορισμένες από τις λέξεις που ήταν σημειωμένες ως θέματα από τα εκπαιδευτικά δεδομένα ήταν κοινές με τις λέξεις των κριτικών που αναλύθηκαν με αποτέλεσμα να ξανά-αναγνωριστούν από το σύστημα. Λαμβάνοντας υπόψη τη συγκεκριμένη παράμετρο η ορθότητα της προσέγγισης εξετάστηκε εκ νέου χωρίς την ύπαρξη οποιωνδήποτε προ-σημειωμένων λέξεων-κλειδιά με αποτέλεσμα το ποσοστό η ορθότητα να φτάνει το 84%.

Συνοψίζοντας, αν συγκρίνουμε τις τέσσερις προσεγγίσεις μη επιβλεπόμενης ανίχνευσης θεμάτων που παρουσιάστηκαν παραπάνω, μπορούμε να καταλήξουμε ότι η ανάλυση συστάδων και η λανθάνουσα σημασιολογική ανάλυση παρέχουν τα καλύτερα αποτελέσματα (Πίνακας 3). Εκτός των χαμηλότερων ποσοστών ορθότητας, η σειριακή εξόρυξη προτύπων φέρει το μειονέκτημα της εξάρτησης από τα προ-σημειωμένα εκπαιδευτικά δεδομένα.

Προσέγγιση Ανίχνευσης Θέματος	Ορθότητα (%)
Τεχνική συχνών λέξεων	83
Ανάλυση συστάδων	89
Λανθάνουσα σημασιολογική ανάλυση	88
Σειριακή εξόρυξη προτύπων	84 (93)

Πίνακας 3 Εκτίμηση των μη επιβλεπόμενων προσεγγίσεων ανίχνευσης θέματος

3 Ανίχνευση Συναισθημάτων

Η ανίχνευση συναισθημάτων έχει να κάνει με τον προσδιορισμό του συναισθήματος ή της πόλωσης μιας κριτικής ή της πρότασης μιας κριτικής. Αυτό το έργο μπορεί να υποστηριχθεί από ανίχνευση υποκειμενικότητας ως προκαταρκτικό βήμα. Σε αυτή την περίπτωση η ανίχνευση υποκειμενικότητας προσδιορίζει εάν η πρόταση μιας κριτικής διέπεται από υποκειμενικότητα ή αντικειμενικότητα και συνεπώς κατά πόσο η πρόταση αυτή εκφράζει μια γνώμη (πχ η τοποθεσία του ξενοδοχείου είναι καλή ή κακή) ή απλώς είναι μια ουδέτερη παρατήρηση (πχ το εστιατόριο είναι στο ισόγειο). Έπειτα η ανάλυση συναισθημάτων πρέπει επεξεργάζεται τις προτάσεις που διέπονται από υποκειμενικότητα και ανάλογα με την εφαρμοζόμενη τεχνική επιτυγχάνει καλύτερα αποτελέσματα από την αντίστοιχη ανάλυση συναισθημάτων που επεξεργάζεται υποκειμενικές και αντικειμενικές προτάσεις ταυτόχρονα.

3α Ανίχνευση Υποκειμενικότητας

Η διαδικασία ανίχνευσης υποκειμενικότητας αποτελείται από το αρχικό βήμα της ταξινόμησης της πρότασης μιας κριτικής ως προς την υποκειμενικότητα ή αντικειμενικότητα της και στη συνέχεια την ολοκλήρωση της διαδικασίας με μια από τις τέσσερις τεχνικές που αναφέρθηκαν παραπάνω (βάσει λεξιλογίου, επιβλεπόμενες προσεγγίσεις μηχανικής μάθησης: k-NN, ΜΔΥ και Naive Bayes). Για την προσέγγιση βάσει λεξιλογίου χρησιμοποιείται ένας κατάλογος 6.800 λέξεων που περιέχει θετικές και αρνητικής λέξεις. Εάν μια πρόταση περιέχει οποιεσδήποτε από τις λέξεις του καταλόγου η πρόταση ταξινομείται αυτομάτως ως υποκειμενική. Όλες οι υπόλοιπες περιπτώσεις ταξινομούνται ως αντικειμενικές. Η δοκιμασία της αποτελεσματικότητας της προσέγγισης βάσει λεξιλογίου στο περιβάλλον των κριτικών των ξενοδοχείων του Are, έφερε ποσοστά ορθότητας της τάξεως των 83%. Η αναζήτηση της ρίζας μίας λέξης (stemming) ή η λημματοποίηση (lemmatization) δεν αύξησε την ορθότητα των αποτελεσμάτων.

Το έργο της ανίχνευσης υποκειμενικότητας αρχικά εκπληρωνόταν από τρεις προσεγγίσεις μηχανικής μάθησης, την k-NN, την ΜΔΥ και την Naive Bayes. Όπως και με τις παραπάνω προσεγγίσεις η συγκεκριμένη μέθοδος εκπαιδεύεται με τη βοήθεια

χειροκίνητα σημειωμένων προτάσεων κριτικών ως προς την υποκειμενικότητα και αντικειμενικότητα τους. Πραγματοποιείται δεκαπλή διασταύρωση για την εκτίμηση όλων μοντέλων μηχανικής μάθησης και για τον υπολογισμό της ορθότητάς τους. Την υψηλότερη ορθότητα έδειξε η προσέγγιση k-NN (72% με το k=120), επόμενη είναι η ΜΔΥ (70%) και τελευταία η Naive Bayes (62%). Σε αντίθεση με τις άλλες δύο προσεγγίσεις, η k-NN επωφελείται σημαντικά από τη χρήση bigramms (δηλαδή προσθήκης ομάδων λέξεων δύο γραμμάτων) και το φιλτράρισμα ουσιαστικών και επιθέτων με τη μέθοδο EML.

Συνοψίζοντας, η βέλτιστη ορθότητα ανίχνευσης υποκειμενικότητας επιτεύχθηκε με την προσέγγιση βάσει λεξιλογίου (81%), η οποία ξεπερνά κατά πολύ τη μέθοδο k-NN (72%) που αποτελεί την καλύτερη προσέγγιση μηχανικής μάθησης (Πίνακας 4). Θα μπορούσε ευλόγως να υποτεθεί ότι οι υψηλές επιδώσεις της προσέγγισης βάσει λεξιλογίου επιτυγχάνονται με τη βοήθεια των σχετικά μεγάλων καταλόγων λέξεων που αποτελούνται από άνω των 6.800 λέξεων, σε σχέση με το περιορισμένο μέγεθος των εκπαιδευτικών δεδομένων των 1.516 προ-ταξινομημένων προτάσεων.

Προσέγγιση Ανίχνευσης Υποκειμενικότητας	Ορθότητα (%)
Βάσει λεξιλογίου	81
k-NN	72
ΜΔΥ	70
Naive Bayes	62

Πίνακας 4 Εκτίμηση των προσεγγίσεων ανίχνευσης υποκειμενικότητας

3β Ανίχνευση Συναισθήματος

Η ανίχνευση συναισθήματος βασίζεται στην ανίχνευση της υποκειμενικότητας και ταξινομεί τις υποκειμενικές δηλώσεις σε θετικές και αρνητικές. Όπως στην ανίχνευση υποκειμενικότητας έτσι και στην ανίχνευση συναισθήματος, η εργασία διεκπεραιώνεται με τη βοήθεια μεθόδων που βασίζονται σε λεξιλόγια, καθώς και σε προσεγγίσεις

επιβλεπόμενης μηχανικής μάθησης, όπως η k-NN, η ΜΔΥ και η Naive Bayes. Η ανίχνευση συναισθημάτων με βάση λεξιλογίου κάνει χρήση καταλόγου λέξεων που περιέχουν περί τις 2000 θετικές και 4800 αρνητικές λέξεις. Σε περίπτωση απουσίας πλειοψηφίας θετικών ή αρνητικών λέξεων, το συναίσθημα προσδιορίζεται ως ουδέτερο. Ο έλεγχος απόδοσης της ανίχνευσης συναισθήματος βάσει λεξιλογίου σημείωσε ποσοστό ορθότητας της τάξης των 72%. Λαμβάνοντας υπόψη αρνητικές λέξεις (όπως “όχι”, “δεν” κλπ) οι οποίες αλλάζουν την σημασία μίας πρότασης δεν επηρέασε το αποτέλεσμα.

Όπως στην ανίχνευση υποκειμενικότητας, η ανίχνευση συναισθήματος εκτελέστηκε από τεχνικές ταξινόμησης k-NN, ΜΔΥ και Naive Bayes. Τα εκπαιδευτικά δεδομένα των προτάσεων που συγκροτούν τις κριτικές, ταξινομήθηκαν χειροκίνητα σε θετικά και αρνητικά. Τέλος, δεκαπλή διασταύρωση πραγματοποιήθηκε για την εκτίμηση της ορθότητας. Την υψηλότερη ορθότητα έφερε η προσέγγιση ΜΔΥ (77%, κάνοντας χρήση bigrams), επόμενη ήταν η Naive Bayes (71%, κάνοντας χρήση trigrams) και τελευταία ήταν η k-NN (70%, με k=10).

Συνοψίζοντας, τις καλύτερες επιδόσεις για την ανίχνευση συναισθημάτων πετυχαίνει η προσέγγιση ΜΔΥ (77%) (Πίνακας 5). Πιθανός λόγος της χαμηλότερης επίδοσης της προσέγγισης βάσει λεξιλογίου (72%) μπορεί να αποδοθεί στο γεγονός ότι η πρόσθετη κατηγορία (ουδέτερη) περιπλέκει τη διαδικασία στις περιπτώσεις που οι θετικές και αρνητικές λέξεις έχουν ίδια συχνότητα.

Προσέγγιση Ανίχνευσης Συναισθήματος	Ορθότητα (%)
Βάσει λεξιλογίου	72
k-NN	70
ΜΔΥ	77
Naive Bayes	71

Πίνακας 5 Εκτίμηση των προσεγγίσεων ανίχνευσης συναισθήματος

Ερευνητικά ερωτήματα

Στα προηγούμενα κεφάλαια καλύψαμε το ρόλο, τη χρησιμότητα και την αξιοποίηση των μεγάλων δεδομένων στο διεθνές τουριστικό περιβάλλον, προκειμένου να λάβουμε μια εικόνα της ελληνικής πραγματικότητας επιλέξαμε να διενεργήσουμε συνεντεύξεις με ειδικούς στον τομέα της πληροφορικής και συγκεκριμένα τους προϊσταμένους IT της Mouzenidis Travel και της Ellinair. Ο λόγος που επιλέχθηκαν οι δυο συγκεκριμένες εταιρείες είναι ότι ανήκουν σε έναν όμιλο που κατέχει ηγετική θέση στον τομέα, έχει ευρύ δίκτυο συνεργατών, εικοσαετή πετυχημένη δράση στην αγορά, προσφέρει μεγάλη ποικιλία και ποσότητα προϊόντων, εξυπηρετεί μεγάλο αριθμό τουριστών και συνεπώς διαχειρίζεται μεγάλους όγκους πληροφοριών. Μέσω των συνεντεύξεων σκοπεύουμε να απαντήσουμε στα παρακάτω ερευνητικά ερωτήματα:

E1= Ποιοι μέθοδοι επεξεργασίας και ανάλυσης μεγάλων δεδομένων εφαρμόζονται στην εταιρεία;

Έχοντας καλύψει σε μεγάλο βαθμό τα προσφερόμενα εργαλεία ανάλυσης μεγάλων δεδομένων καθώς και τα εργαλεία που υιοθετούνται από τουριστικές επιχειρήσεις σε διεθνές επίπεδο έχει ενδιαφέρον να μάθουμε ποια εργαλεία ανάλυσης μεγάλου όγκου δεδομένων χρησιμοποιούνται από τις δύο εταιρείες προκειμένου να προσδιορίσουμε το επίπεδο της τεχνολογίας που χρησιμοποιείται σε σχέση με το διεθνές επίπεδο.

E2= Πως αξιοποιούνται τα αποτελέσματα των μεγάλων δεδομένων για την βελτίωση της απόδοσης της εταιρείας;

Μέσα από την βιβλιογραφική έρευνα που έγινε στην εργασία αυτή επισημάνθηκε η αποτελεσματικότητα των διάφορων προσεγγίσεων και εργαλείων ανάλυσης μεγάλων δεδομένων για τις επιχειρήσεις που επέλεξαν να τις εφαρμόσουν. Μέσω του εμπειρικού μέρους της έρευνας θα επιχειρήσουμε να επαληθεύσουμε τη θεωρία αυτή στα πλαίσια εγχώριας τουριστικής αγοράς, δηλαδή πως γίνεται η διαχείριση και η αξιοποίηση των πληροφοριών για την αντιμετώπιση προβλημάτων και την επίτευξη των εταιρικών στόχων.

E3 = Πως έχουν επηρεάσει οι κανονισμοί Προστασίας Προσωπικών Δεδομένων (GDPR) τις λειτουργίες της εταιρείας;

Ένας σημαντικός παράγοντας της ανάλυσης πληροφοριών που σχετίζονται με τους καταναλωτές είναι τα προσωπικά δεδομένα. Από το Μάιο του 2018 η Ευρωπαϊκή Ένωση επέβαλε το Γενικό Κανονισμό Προστασίας Προσωπικών Δεδομένων (GDPR) ο οποίος ισχύει καθολικά σε όλους του τομείς και σε όλους τους επιχειρηματικούς κλάδους και εξασφαλίζει την ασφάλεια των προσωπικών δεδομένων των καταναλωτών μέσω κρυπτογράφησης και ανωνυμίας. Συνεπώς, το εμπειρικό σκέλος της μελέτης έχει σκοπό να αποσαφηνίσει πως και κατά πόσο επηρέασε ο κανονισμός αυτός τις λειτουργίες των επιχειρήσεων.

E4 = Σε τι βαθμό έχουν ενσωματωθεί στις επιχειρησιακές διαδικασίες οι τεχνολογίες μεγάλων δεδομένων;

Προκειμένου να προσδιορίσουμε το επίπεδο της ελληνικής τουριστικής πραγματικότητας ως προς την εφαρμογή των τεχνολογιών διαχείρισης μεγάλου όγκου δεδομένων είναι σημαντικό να εντοπίσουμε το βαθμό ενσωμάτωσης των εργαλείων ανάλυσης μεγάλων δεδομένων στις επιχειρησιακές διαδικασίες και να τις συγκρίνουμε με τον αντίστοιχο βαθμό ενσωμάτωσης της διεθνούς αγοράς.

Συνηεντεύξεις

Οι συνηεντεύξεις έγιναν μέσω email, με δομημένα ερωτηματολόγια και ερωτήσεις ανοικτού τύπου ώστε να δοθεί η δυνατότητα στους ερωτώμενους να εκφράσουν ελεύθερα τις απόψεις τους και να αναπτύξουν τις αναλύσεις τους επί των θεμάτων. Παρακάτω παρουσιάζονται τα στοιχεία που συλλέχθηκαν από τις απαντήσεις των ερωτηματολογίων του Αρθούρου Τσιλίδη - προϊστάμενου IT στη Mouzenidis Travel και του Σωκράτη Παπουνίδη - προϊστάμενου IT στην Ellinair.

Αρθούρος Τσιλίδης (Mouzenidis Travel)

Ο Αρθούρος Τσιλίδης έχει δεκαετή εμπειρία στο τμήμα IT της εταιρείας, όπου τα τελευταία τέσσερα χρόνια κατέχει τη θέση του προϊσταμένου, ηγείται μιας ομάδας 13 ατόμων, όπου οι 10 εξ αυτών είναι προγραμματιστές και οι 3 ασχολούνται με hardware και support. Σύμφωνα με τον κ. Τσιλίδη, τα μεγάλα δεδομένα αποτελούν ένα ισχυρό εργαλείο για την ανίχνευση των τουριστικών τάσεων κάτι που παλαιότερα ήταν ακατόρθωτο δίχως την ύπαρξη εκτεταμένου δικτύου συνεργατών. Ένας τρόπος ανίχνευσης τέτοιων τάσεων είναι μέσω της ανάλυσης μεγάλου όγκου δεδομένων που έχουν συλλεχθεί από την ιστοσελίδα της εταιρείας και των συνεργατών της και την αξιοποίηση τους για την διαπίστωση των καταναλωτικών προτιμήσεων για συγκεκριμένες χρονικές περιόδους.

Κυριότερη πηγή δεδομένων της εταιρείας αποτελούν τα portal sites της εταιρείας και των παραρτημάτων της και ο καθημερινός όγκος δεδομένων που καλούνται να διαχειριστούν κυμαίνεται από μερικές εκατοντάδες megabytes μέχρι μερικά gigabytes, που αντιστοιχούν σε εκατοντάδες επισκέψεις στις αντίστοιχα portal sites της εταιρείας. Η αποθήκευση των δεδομένων γίνεται αυτόματα, σε σχεσιακή βάση δεδομένων κατά την επίσκεψη των χρηστών στις ιστοσελίδες που διαχειρίζεται η εταιρεία. Γίνεται καταγραφή όλων των κινήσεων και των στοιχείων των επισκεπτών όπως τα δημογραφικά στοιχεία, οι προτιμήσεις, οι λέξεις κλειδιά που χρησιμοποίησαν κατά την αναζήτηση των τουριστικών πακέτων τους, ο χρόνος παραμονής τους στην ιστοσελίδα κλπ. Η ταξινόμηση και επεξεργασία των δεδομένων γίνεται με την χρήση database schemas (προδιαγραφές αποθήκευσης δεδομένων) και στατιστικών κατανομών βάσει

προτιμήσεων των καταναλωτών, την ηλικιακή τους ομάδα και την τουριστική περίοδο. Η πρόβλεψη της τουριστικής ζήτησης γίνεται μέσω πιο παραδοσιακών τεχνικών, όπως η χρήση δεδομένων προηγούμενων ετών, ενώ ο προσδιορισμός ενδοεταιρικών προβλημάτων δεν ανήκει στις εργασίες του τμήματος. Ωστόσο τα συμπεράσματα που προκύπτουν από την ανάλυση δεδομένων όπως είναι οι τάσεις της αγοράς, αξιοποιούνται για την χάραξη της μεσοπρόθεσμης και μακροπρόθεσμης στρατηγικής της εταιρείας καθώς και τις ευκαιρίες που παρουσιάζονται στις υπάρχουσες και νέες αγορές. Μέσω της παρακολούθησης των μεταβολών στις προτιμήσεις των target groups προσδιορίζονται τα τουριστικά προϊόντα που έχουν μεγαλύτερη ζήτηση σε συγκεκριμένη εποχή. Η αξιολόγηση της επιτυχίας των νέων προϊόντων γίνεται βάσει των αφίξεων στους προορισμούς, της ζήτησης των τουριστικών πακέτων και των shares που έχουν γίνει στα μέσα κοινωνικής δικτύωσης. Τέλος, λόγω των αυξημένων απαιτήσεων υπολογιστικής ισχύος και network bandwidth των νέων τεχνικών συλλογής, αποθήκευσης και επεξεργασίας δεδομένων, στους άμεσους στόχους του τμήματος βρίσκεται η αναβάθμιση του υλικού και λογισμικού των πληροφοριακών συστημάτων καθώς και η στενή παρακολούθηση και εφαρμογή των τεχνικών της ανάλυσης μεγάλων δεδομένων.

Αναφορικά με τους κανονισμούς των προσωπικών δεδομένων (GDPR) και κατά πόσο έχουν επηρεάσει τις διαδικασίες συλλογής και αποθήκευσης δεδομένων σχολιάστηκε ότι τηρούνται όλοι οι κανονισμοί ασφαλείας και ότι σε τεχνικό επίπεδο οι αλλαγές που έγιναν ήταν λίγες και δίχως επιβαρυντικά επιπτώσεις στις διαδικασίες συλλογής και αποθήκευσης δεδομένων, απεναντίας παρατηρήθηκε αύξηση των feedbacks στις ιστοσελίδες που αποδίδεται στα αυξημένα μέτρα ασφαλείας.

Σωκράτης Παπουνίδης (Ellinair)

Ο Σωκράτης Παπουνίδης ξεκίνησε την επαγγελματική του σταδιοδρομία στην Mouzenidis Travel όπου εργάστηκε για οκτώ έτη και τα τελευταία δυο χρόνια υπήρξε προϊστάμενος του τμήματος IT, με την ίδρυση της Ellinair μεταφέρθηκε στη νέα εταιρεία του ομίλου ως ο νέος προϊστάμενος του τμήματος IT. Το τμήμα αποτελείται από πέντε άτομα όπου οι τρεις είναι προγραμματιστές και οι δύο hardware και support. Σύμφωνα

με τον κ Παπουνίδη τα δεδομένα, είτε είναι μεγάλα, είτε μικρά, αποτελούν τη βασική πηγή πληροφόρησης για τη λήψη αποφάσεων σε κάθε σύγχρονη επιχείρηση. Ως παράδειγμα αναφέρεται η αξιοποίηση των αποθηκευμένων δεδομένων για την μελέτη των τιμών των εισιτηρίων, την πρόβλεψη της ζήτησης, την ανάλυση στατιστικών και την απλοποίηση και βελτιστοποίηση των διαδικασιών της εταιρείας. Κύριες πηγές δεδομένων αποτελούν το ERP της εταιρείας, τα στατιστικά επισκεψιμότητας των ιστοσελίδων καθώς και τα logs & metrics και για την εξαγωγή των δεδομένων χρησιμοποιούνται αυτοματοποιημένες διαδικασίες όπως η ETL (Extract, Transform, Load) που εξάγει, τροποποιεί, συνδυάζει και αποθηκεύει δεδομένα για την εύκολη ανάλυση στην τοπική βάση δεδομένων. Τα δεδομένα αυτά σε πολλές περιπτώσεις αποτελούν το input για διαφορά reports. Οι παραπάνω πηγές δεδομένων τροφοδοτούν το τμήμα Revenue Analysis το οποίο χρησιμοποιώντας το σύστημα Revenue Management πρόβλεπει τη μελλοντική ζήτηση. Έκτος από τα δεδομένα πωλήσεων και επισκεψιμότητας των ιστοσελίδων, συλλέγονται και δεδομένα webserver logs τα οποία χρησιμεύουν στην επίλυση προβλημάτων που σχετίζονται με τις λειτουργίες της ιστοσελίδας. Επίσης για την επίλυση των ενδοεταιρικών προβλημάτων χρησιμοποιούνται πιο παραδοσιακοί μέθοδοι όπως οι επισκοπήσεις και τα ερωτηματολόγια για τη συλλογή ποιοτικών δεδομένων. Η αξιοποίηση των δεδομένων για την διευκόλυνση και την υποστήριξη της διεύθυνσης γίνεται με την αποστολή αναφορών από τις αναλύσεις των συντρεχθέντων δεδομένων, είτε πιο άμεσα μέσω του intranet portal. Για τον προσδιορισμό νέων ευκαιριών, των επιχειρηματικών ιδεών και των νέων αγορών η ανάμειξη του τμήματος περιορίζεται στη συλλογή και επεξεργασία δεδομένων όπως είναι οι αναφορές των αεροδρομίων, της πολιτικής αεροπορίας κλπ. Η αξιολόγηση της επίδοσης της εταιρείας πραγματοποιείται επίσης με τον παραδοσιακό τρόπο της σύγκρισης δεδομένων προηγούμενων περιόδων κάτι που επιτρέπει την εταιρεία να μην ξεφεύγει από τους αρχικούς στόχους και τον προϋπολογισμό. Στους μελλοντικούς στόχους του τμήματος βρίσκεται η αυτοματοποίηση των διαδικασιών συλλογής, αποθήκευσης και ανάλυσης δεδομένων, η εξαγωγή των αποτελεσμάτων

Η διαχείριση μεγάλου όγκου δεδομένων στον Τουρισμό- Μια προσπάθεια βιβλιογραφικής και εμπειρικής προσέγγισης του θεματος

αναλύσεων των δεδομένων σε μορφή αναφορών ή OLAP και η χρήση των εργαλείων μηχανικής μάθησης για τη βελτίωση της εξόρυξης και της ανάλυσης των δεδομένων.

Όσον αφορά την εφαρμογή των κανονισμών των προσωπικών δεδομένων (GDPR) φαίνεται ότι όχι μόνο δεν δυσκολεύσαν τη λειτουργία του τμήματος αλλά αντιθέτως βελτίωσαν τις διαδικασίες ταξινόμησης και ανάλυσης δεδομένων.

Συμπεράσματα ποιοτικής έρευνας

Η διάδοση και η ραγδαία αύξηση της εφαρμογής μεγάλων δεδομένων τα έχει καταστήσει εργαλεία ευρείας χρήσης στις εταιρικές αποφάσεις με αποτέλεσμα τη δημιουργία νέων κοινωνικών κίνδυνων. Ταυτοχρόνως, καθώς διαδίδονται οι τεχνολογίες ανάλυσης δεδομένων εγείρονται ερωτήματα προστασίας προσωπικών δεδομένων και το ενδεχόμενο κατάχρησής τους. Η χρήση των μεγάλων δεδομένων για πρόβλεψη ενέχει ορισμένες προκλήσεις. Η πρώτη πρόκληση αφορά στη δυσκολία εντοπισμού των σωστών δεδομένων και του βέλτιστου τρόπου αξιοποίησής τους. Η δεύτερη πρόκληση είναι η εύρεση του κατάλληλου εξειδικευμένου προσωπικού, ικανού τόσο να εργαστεί με τις νέες τεχνολογίες όσο και να ερμηνεύει τα δεδομένα προκειμένου να εξάγει χρήσιμες πληροφορίες. Η τρίτη πρόκληση είναι να ξεπεραστεί το εμπόδιο της προσβασιμότητας και της συνδεσιμότητας δεδομένων, το οποίο απαιτεί τα κατάλληλα συστήματα που θα συγκεντρώνουν και θα διαχειρίζονται μεγάλα δεδομένα. Η τέταρτη πρόκληση είναι η εύρεση νέων τρόπων αξιοποίησης των μεγάλων δεδομένων. Τέλος, η πέμπτη πρόκληση αφορά στην ασφάλεια των μεγάλων δεδομένων και του τρόπου διατήρησής του πλεονεκτήματος της χρήσης τους.

Υπάρχουν πολλές πιθανές λύσεις για να αντιμετωπιστούν αυτές οι προκλήσεις. Αρχικά, τα αποτελέσματα των προβλέψεων των μεγάλων δεδομένων πρέπει να σχετίζονται άμεσα με τις επιχειρηματικές αποφάσεις. Σκοπός της τουριστικής πρόβλεψης είναι η ταχεία εύρεση και ανάλυση σημαντικών πληροφοριών με ακρίβεια. Η οπτικοποίηση είναι ένας καλός τρόπος για την παρουσίαση αποτελεσμάτων και για τη διευκόλυνση λήψης γρήγορων αποφάσεων. Με τη βελτίωση του υλισμικού και των μοντέλων πρόβλεψης δίνεται η δυνατότητα επεξεργασίας τεράστιων όγκων δεδομένων και εξαγωγής πληροφοριών σε πραγματικό χρόνο. Δευτερευόντως, μια εναλλακτική λύση είναι η απόκτηση μιας γενικής γνώσης των μεγάλων δεδομένων η οποία είναι ζωτικής σημασίας για την οπτικοποίηση και την ερμηνεία των δεδομένων. Συγκεκριμένα, πρέπει να υπάρχει καλή γνώση των πηγών των δεδομένων, του κοινού που θα καταναλώνει τα δεδομένα και του τρόπου που θα αντιλαμβάνεται το κοινό τις πληροφορίες αυτές. Ακολούθως, τρίτη λύση είναι η προοληπτική χρήση των μεγάλων δεδομένων καθώς οι περισσότερες από τις πληροφορίες συλλέγονται σε πραγματικό

χρόνο και έχουν κολοσσιαίο όγκο. Συνεπώς, η έγκαιρη χρήση των μεγάλων δεδομένων για την πρόβλεψη και την χρήση των κατάλληλων προσεγγίσεων και μεθόδων είναι ο καλύτερος τρόπος αξιοποίησης των ωφελημάτων των μεγάλων δεδομένων.

Δεδομένης της ευρείας χρήσης των “έξυπνων” συσκευών, όπως τα smartphones και τα tablets, οι καταναλωτές μπορούν εύκολα να αναζητούν και να διαδίδουν πληροφορίες word of mouth στο διαδίκτυο από τα μέσα κοινωνικής δικτύωσης. Οι ηλεκτρονικές κριτικές ασκούν σημαντική επιρροή στις επιλογές των καταναλωτών και στην αγοραστική τους συμπεριφορά. Οι ξενοδόχοι επίσης, εξετάζουν όλους τους πιθανούς τρόπους αξιοποίησης των ηλεκτρονικών κριτικών ώστε να κατανοήσουν, να επικοινωνήσουν, να ευχαριστήσουν και να προσελκύσουν πελάτες. Μέσω συνεργασιών με ιστοσελίδες ηλεκτρονικών κριτικών ή μέσω της ενσωμάτωσης δυνατοτήτων δημοσίευσης κριτικών στις ιστοσελίδες τους, οι ξενοδοχειακές επιχειρήσεις επιδιώκουν να οδηγήσουν τους πελάτες στις ιστοσελίδες τους προκειμένου να αυξήσουν τις κρατήσεις και να διαχειρίζονται οι ίδιοι την ηλεκτρονική τους φήμη. Οι ηλεκτρονικές κριτικές, ως συνιστώσα των μεγάλων δεδομένων παρέχουν πολύτιμες πληροφορίες στους ξενοδόχους για τους πελάτες τους και τη βελτίωση των παρεχόμενων υπηρεσιών. Οι διευθυντές ξενοδοχείων πρέπει να ακολουθούν το ρυθμό των αλλαγών των πελατειακών προτιμήσεων και να έχουν καλή αντίληψη των τάσεων στα μέσα κοινωνικής δικτύωσης ώστε να παρακολουθούν τις ηλεκτρονικές κριτικές. Ως εκ τούτου, η πολιτική ανταπόκρισης ενός ευρέως καταξιωμένου ξενοδοχείου αναμένεται να ενισχύσει την εταιρική του εικόνα και να διατηρήσει την αφοσίωση των πελατών.

Σήμερα ο τρόπος που ταξιδεύουμε σε έναν προορισμό και οι εμπειρίες που αποκομίζουμε από αυτόν εξελίσσονται συνεχώς λόγω της τεχνολογίας. Η εμφάνιση των συστημάτων που πλαισιώνουν τον ψηφιακό ταξιδιώτη παίζουν το ρόλο της νέας γενιάς εργαλείων που αλλάζουν τον τρόπο που ταξιδεύουν οι άνθρωποι. Σε αυτό τον νέο κόσμο του ψηφιακού ταξιδιώτη, οι φορητές συσκευές απαθανατίζουν ολόκληρη την εμπειρία του ταξιδιού, όπου όλες οι πτυχές του είναι διαμορφωμένες απόλυτα στα μέτρα του ταξιδιώτη. Επιπλέον, υποστηρίζεται ότι οι νέες αυτές τεχνολογίες θα επιφέρουν αλλαγές στη διαδικασία δημιουργίας αξίας, όπου οι ταξιδιώτες θα γίνονται πιο

δημιουργικοί κατά το σχεδιασμό του ταξιδιού τους με τρόπο που να ταιριάζει απόλυτα στις ξεχωριστές ταξιδιωτικές τους ανάγκες, αξίες και προτιμήσεις. Ασφαλώς, οι νέες συσκευές και το νέο πληροφοριακό περιβάλλον απειλούν τα παραδοσιακά κανάλια πληροφοριών που παρέχουν συστάσεις και βασικές πληροφορίες σχετικά με τον προορισμό. Ως εκ τούτου, η εμφάνιση του ψηφιακού ταξιδιώτη δημιουργεί την ανάγκη ανάπτυξης πιο δυναμικών στρατηγικών από τους προορισμούς ώστε να δώσουν τη δυνατότητα στον κάθε επισκέπτη να επιλέξει τον τρόπο που θα βιώσει την ταξιδιωτική του εμπειρία ανάλογα με τις ανάγκες του. Προκειμένου να επωφεληθούν από τις νέες αυτές τεχνολογίες, οι φορείς και παράγοντες των εκάστοτε τουριστικών προορισμών πρέπει να σχεδιάσουν τα προϊόντα και τις υπηρεσίες που προσφέρουν έτσι ώστε να αλληλεπιδρούν άμεσα με τους επισκέπτες. Θα πρέπει να δοθεί έμφαση σε αρκετούς παράγοντες των καινοτομιών αυτών ώστε να ληφθούν υπόψη από τους υπεύθυνους των προορισμών. Τεχνολογίες όπως η παιχνιδοποίηση, η ειδοποίηση περιβάλλοντος και η αφηγηματική παραγωγή ερεθισμάτων πρέπει να ενσωματώνονται στις λειτουργίες των επιχειρήσεων ώστε ο κόσμος να προσαρμόζεται σε αυτές. Μολαταύτα, τα μεγάλα δεδομένα που παράγονται κατά τις καθημερινές δραστηριότητες των ταξιδιωτών, κατά τη διάρκεια των ταξιδιών καθώς και οι πληροφορίες που συλλέγονται από τις επιχειρήσεις μπορούν να χρησιμεύσουν ως δομικά συστατικά της ανάπτυξης των έξυπνων τουρισμών προορισμών. Συνοψίζοντας, η τουριστική βιομηχανία βρίσκεται στα πρόθυρα μιας νέας επανάστασης που θα αλλάξει όχι μόνο τα εργαλεία που χρησιμοποιούνται για το σχεδιασμό των ταξιδιών και τον τρόπο που δημιουργούνται οι ταξιδιωτικές εμπειρίες, αλλά και τη φύση της ίδιας της τουριστικής βιομηχανίας.

Συμπεράσματα εμπειρικής έρευνας

Η χρήση των μεγάλων δεδομένων στον κλάδο του τουρισμού και της φιλοξενίας διεθνώς βρίσκεται ακόμα σε πρώιμο στάδιο, ωστόσο η δυνητική ανάπτυξη των εφαρμογών της είναι τεράστια. Η ίδια εικόνα επιβεβαιώνεται και από την εμπειρική μελέτη καθώς οι πληροφορίες που αποκομίσαμε από τις συνεντεύξεις με ανώτερα στελέχη της Mouzenidis Travel και της Ellinair επιβεβαιώνουν το σχετικά χαμηλό

ποσοστό ενσωμάτωσης των τεχνολογιών μεγάλων δεδομένων στις επιχειρησιακές διαδικασίες. Οι διαδικασίες που εφαρμόζονται τεχνολογίες διαχείρισης μεγάλου όγκου δεδομένων είναι αυτές της αποθήκευσης δεδομένων (σχεσιακές βάσεις δεδομένων), της ταξινόμησης δεδομένων (database schemas - προδιαγραφές αποθήκευσης δεδομένων) και της εξόρυξης δεδομένων (ETL - Extract, Transform, Load). Οι χρήση των εργαλείων αυτών που περιλαμβάνουν πολλούς αυτοματισμούς έχει ως αποτέλεσμα την αύξηση της αποδοτικότητας των τμημάτων IT των δύο εταιρειών που μεταφράζεται σε δεκάδες εργατοώρες ανά βδομάδα. Επιπλέον, αναφορικά με την εφαρμογή του Κανονισμού Προστασίας Προσωπικών Δεδομένων (GDPR), παρά τις αρχικές επιφυλάξεις των επιχειρήσεων περί επιβάρυνσης και περιπλοκής των λειτουργιών, η εμπειρική ανάλυση έδειξε ότι εκτός από ορισμένες τροποποιήσεις σε τεχνικό επίπεδο (αλλαγές στο front-end των ιστοσελίδων και καταγραφή της αποδοχής ή απόρριψης των όρων GDPR) όχι μόνο δεν τις επιβάρυνε αλλά συνέβαλε σε επίπεδο βελτίωσης των διαδικασιών ταξινόμησης και ανάλυσης δεδομένων καθώς επίσης και σε επίπεδο συμπεριφοράς των καταναλωτών στο διαδίκτυο, οι οποίοι από τη στιγμή της εφαρμογής του νέου κανονισμού παρατηρήθηκε ότι αναρτούν πιο ενεργά feedbacks.

Η γνώση των δυνατοτήτων των μεγάλων δεδομένων τα καθιστά ως πρώτη προτεραιότητα στα μελλοντικά σχέδια των δυο εταιρειών προκειμένου να αντικατασταθούν οι παραδοσιακές μέθοδοι που χρησιμοποιούνται σε αρκετές από τις διαδικασίες. Πριν την αντικατάσταση των παλαιών τεχνολογιών από τις νέες όμως υπάρχουν πολλές παρασκηνακές διεργασίες των μεγάλων δεδομένων που πρέπει να ολοκληρωθούν πριν φθάσουν τα αποτελέσματά στους καταναλωτές. Ευτυχώς, οι λύσεις για τη χρήση των μεγάλων δεδομένων είναι πλέον ευρέως διαθέσιμες και το κόστος είναι πλέον πολύ χαμηλό. Η χρήση των μεγάλων δεδομένων από βιομηχανίες που δραστηριοποιούνται στον τουριστικό κλάδο δίνει τη δυνατότητα εξαγωγής πληροφοριών για τις προτιμήσεις των πελατών σε μακρο-επίπεδο και μεγάλων προοπτικών κερδοφορίας με αντάλλαγμα τις σχετικά μικρού μεγέθους επενδύσεις στο εγγύς μέλλον.

Προτάσεις για περαιτέρω έρευνα

Η ανάλυση μεγάλων δεδομένων στον τουρισμό προσφέρει νέα διάσταση στην έρευνα της φύσης και της διαδικασίας της τουριστικής εμπειρίας. Πολλές εφαρμογές αναλύσεων που είναι διαθέσιμες έχουν τη δυνατότητα να δώσουν άλλο νόημα στην έννοια του τουρισμού. Για παράδειγμα οι τεχνολογίες έξυπνων συσκευών μελλοντικά θα μπορούν να βρίσκονται σε θέση να γνωρίζουν και να ικανοποιούν τις ανάγκες των ταξιδιωτών. Με τη βοήθεια μικροκυκλωμάτων ενσωματωμένων σε τουριστικά αξιοθέατα οι πάροχοι τουριστικών υπηρεσιών θα μπορούν να εντοπίζουν βάση γεωγραφικής τοποθεσίας τους τουρίστες και να προσδιορίζουν τη συμπεριφορά τους ώστε να προσφέρουν υπηρεσίες βάσει τοποθεσίας. Οι τουρίστες θα μπορούν να χρησιμοποιούν τα smartphones τους για να ερευνήσουν τα τουριστικά αξιοθέατα και τα σημεία ενδιαφέροντος χρησιμοποιώντας επιτόπια συλλογή και αναφορά δεδομένων. Επίσης, οι διαδικτυακές δραστηριότητες που παράγουν ψηφιακά ίχνη αφήνουν πίσω τους ποικιλόμορφα δεδομένα που επιτρέπουν στους τουριστικούς οργανισμούς να αναπτύξουν νέα επιχειρηματικά μοντέλα βάσει των τουριστικών εμπειριών. Μέσα σε ένα κοινωνικό περιβάλλον θα δίνεται η δυνατότητα να εντοπιστούν και να συλλεχθούν πληροφορίες σχετικά με εκδηλώσεις και τα μέρη που συναθροίζεται ο κόσμος ώστε να παρέχουν πολύτιμες πληροφορίες. Αυτό συνεπάγεται ότι το ίδιο ταξίδι δεν θα αποτελεί πια προσωπική εμπειρία ενός ατόμου αλλά πολλαπλές εμπειρίες ενός κοινωνικού συνόλου όπου ο χρόνος, ο χώρος και η αλληλεπίδραση με το φυσικό περιβάλλον είναι άμεσα κατανεμημένη (με συγγενείς, φίλους και γνωστούς) μεταξύ πολλών ηλεκτρονικά συνδεδεμένων κοινωνικών δικτύων. Περισσότερες τέτοιες ευκαιρίες θα ανακαλύπτονται καθώς συνεχίζονται οι έρευνες προκειμένου να κατανοήσουμε τις συνθήκες της αγοράς καθώς και τις πραγματικές σχέσεις μεταξύ της προσφοράς και της ζήτησης του τουρισμού.

Το πεδίο έρευνας των τεχνολογιών διαχείρισης μεγάλου όγκου δεδομένων στον τουρισμό είναι ευρύ καθώς η ανάπτυξη στον τομέα των επιστήμων πληροφορικής και του διαδικτύου είναι ραγδαία ενώ η εφαρμογή τους στην τουριστική βιομηχανία βρίσκεται σε πρώιμο στάδιο. Μελλοντικές μελέτες θα μπορούσαν να επικεντρωθούν στην έρευνα τεχνικών μέτρησης ανθρωπίνων αισθήσεων και τουριστικών εμπειριών. Η

μελέτες αυτές θα δώσουν τη δυνατότητα να μελετηθεί ο ρόλος των ανθρώπινων συναισθημάτων κατά τη διάρκεια του ταξιδιού. Μια προτεινόμενη προσέγγιση θα ήταν αυτή της ανθρωποκεντρικής προσέγγισης εξαγωγής συναισθηματικών πληροφοριών με τη χρήση φορητών αισθητήρων ώστε να κατανοηθεί πως ακριβώς δημιουργούνται οι τουριστικές εμπειρίες. Άλλη κατεύθυνση που θα μπορούσαν να λάβουν οι μελλοντικές έρευνες είναι αυτή της μελέτης Συστημάτων Γεωγραφικών Πληροφοριών (GIS) για την παρακολούθηση των ταξιδιωτικών προορισμών βάσει μεγάλων δεδομένων. Οι μελέτες αυτές θα είχαν στόχο τη διερεύνηση των μοτίβων των ταξιδιωτικών ροών κατά τη διάρκεια εθνικών ή θρησκευτικών εορτών. Τέλος, σημαντικές γνώσεις θα μπορούσαν να προσφέρουν οι μελέτες περιπτώσεων με τη χρήση διαδικτυακών αναλύσεων και αναλύσεων των μέσων κοινωνικής δικτύωσης. Συγκεκριμένα, οι έρευνες αυτές θα μπορούσαν να δείξουν πώς οι μεγάλες διοργανώσεις όπως το Mundial ή οι Ολυμπιακοί Αγώνες απεικονίζονται στα μέσα κοινωνικής δικτύωσης από τους οικοδεσπότες και τους φιλοξενούμενους κάνοντας έμφαση στη γεωγραφική αναπαράσταση και τα καίρια θέματα που τους απασχολούν πριν, κατά τη διάρκεια και μετά την εκδήλωση.

Περιορισμοί της έρευνας

Στην παρούσα εργασία υποστηρίζεται ότι οι ανάλυση μεγάλων δεδομένων συνδέεται εγγενώς με την πρόσφατη ανάδειξη του τουριστικού σχεδιασμού και της ανάπτυξης του έξυπνου τουρισμού, γεγονός που είναι λογικό επακόλουθο της εξέλιξης της πληροφορικής και της ευρείας εφαρμογής της τόσο στην καταναλωτική όσο και στη βιομηχανική αγορά τα τελευταία είκοσι χρόνια. Η χρήση δεδομένων βρίσκεται στο επίκεντρο όλων των έξυπνων τουριστικών δραστηριοτήτων. Ως εκ τούτου, η αξιοποίηση και η εκμετάλλευση των μεγάλων δεδομένων θα οδηγήσει στη δημιουργία νέων επιχειρηματικών μοντέλων και σε καινοτομίες στον τομέα του τουρισμού. Ωστόσο, υπάρχουν πολλά ζητήματα και προκλήσεις στη χρήση μεγάλων δεδομένων. Για παράδειγμα, η προστασία της ιδιωτικότητας αποτελεί προφανή ανησυχία στα πλαίσια της εφαρμογής του έξυπνου τουρισμού, ιδιαίτερα δε των υπηρεσιών που βασίζονται στην τοποθεσία που είναι εξαιρετικά χρήσιμη για τους τουρίστες καθιστά όμως

ευάλωτους τους καταναλωτές. Συνεπώς, ένας από τους περιορισμούς της συγκεκριμένης έρευνας έχει να κάνει με την απουσία πληροφοριών για την προστασία των προσωπικών δεδομένων. Δυστυχώς είναι πολύ νωρίς για να μπορούμε να συλλέξουμε πληροφορίες σχετικά με την τήρηση και την αποτελεσματικότητα του κανονισμού της Ευρωπαϊκής Ένωσης καθώς δεν έχει περάσει αρκετός καιρός ώστε να μπορέσουν να αξιολογηθούν τα αποτελέσματα. Ωστόσο, τα στοιχεία που συλλέχθηκαν από το εμπειρικό μέρος της εργασίας δείχνουν ότι ενώ ο κανονισμός εφαρμόστηκε πλήρως δεν είχε επιβαρυντικά αποτελέσματα στις επιχειρηματικές λειτουργίες των τουριστικών οργανισμών.

Έναν ακόμα περιορισμό αποτελεί το ελλιπές δείγμα που λήφθηκε για την πραγματοποίηση της εμπειρικής έρευνας. Σκοπός της έρευνας ήταν ο προσδιορισμός του βαθμού χρήσης των τεχνολογιών μεγάλων δεδομένων στην εγχώρια τουριστική αγορά. Ωστόσο, οι πληροφορίες που λάβαμε ήταν από δυο εταιρείες που αν και ανήκουν σε έναν από τους μεγαλύτερους παράγοντες του εγχώριου τουρισμού δεν ήταν αρκετές για να μας δώσουν την αντικειμενική εικόνα της αγοράς. Σε ιδανική περίπτωση η λήψη πληροφοριών από φορείς και παράγοντες τουρισμού όπως του Υπουργείου Τουρισμού, του ΕΟΤ, της ΕΛΣΤΑΤ, του Ξενοδοχειακού Επιμελητηρίου και από άλλους ιδιωτικούς οργανισμούς θα μας έδινε μια πολύ πιο ολοκληρωμένη εικόνα της ελληνικής πραγματικότητας.

Τέλος, άλλο περιοριστικό στοιχείο σχετικό με την εμπειρική έρευνα έχει να κάνει με την ποιότητα των απαντήσεων καθώς σε πολλές από τις ερωτήσεις μπορεί να μην δόθηκαν ολοκληρωμένες ή εμπεριστατωμένες απαντήσεις για λόγους προστασίας ευαίσθητων εταιρικών δεδομένων. Αυτό σημαίνει ότι η ποιότητα των απαντήσεων μπορεί να επηρέασε το αποτέλεσμα της παρούσας έρευνας.

Προτάσεις για διοικητική εφαρμογή στην πράξη

Όπως έδειξαν τα αποτελέσματα της παρούσας εργασίας η ανάλυση μεγάλων δεδομένων μπορεί να προσφέρει στην τουριστική βιομηχανία με πολλούς τρόπους. Παρέχοντας πληροφορίες που θα ενισχύσουν τις διοικητικές ικανότητες των

τουριστικών οργανισμών, τις διαδικασίες λήψης αποφάσεων, τον εντοπισμό ομάδων στόχου και την ενίσχυση της αποτελεσματικότητας και της ποιότητας των προσφερόμενων υπηρεσιών. Μια τέτοια εφαρμογή θα μπορούσε να λάβει χώρα στους αερολιμένες τουριστικών προορισμών και όχι μόνο. Τα δεδομένα που παρέχουν οι ταξιδιώτες κατά τη διέλευσή τους από τα αεροδρόμια, μπορούν να βοηθήσουν τις αεροπορικές εταιρείες να σχεδιάσουν αποτελεσματικά τις υπηρεσίες και τους πόρους που έχουν στη διάθεσή τους. Στη συνέχεια κάνοντας χρήση της τεχνολογίας “αναζήτησης” (seeking) συλλέγονται ανώνυμες πληροφορίες για τις κινήσεις των ταξιδιωτών στο αεροδρόμιο. Η εφαρμογή τεχνολογιών μεγάλων δεδομένων στον ξενοδοχειακό κλάδο, μπορεί να βελτιώσει τις προσφερόμενες υπηρεσίες και να ενισχύσει την αποτελεσματικότητα των λειτουργιών μέσω της προσφοράς προσωποποιημένων υπηρεσιών. Συλλέγοντας δεδομένα από τα μέσα κοινωνικής δικτύωσης των πελατών εξάγονται πληροφορίες για τους επισκέπτες, όπως οι γαστρονομικές προτιμήσεις, τα χόμπι, τα χαρακτηριστικά της προσωπικότητάς τους κλπ τα οποία αποτελούν σημαντικούς παράγοντες για την προσφορά προσωποποιημένων υπηρεσιών. Η λειτουργία άμεσης χρέωσης στα μίνι-μπαρ κάνει χρήση ενσωματωμένου σαρωτή των διαθέσιμων προϊόντων τα οποία χρεώνονται μέσω αισθητήρων όποτε αυτά καταναλώνονται. Αυτές οι εφαρμογές των τεχνολογιών μεγάλων δεδομένων πέρα από τις θετικές εμπειρίες που δημιουργούν στους ταξιδιώτες δίνουν τη δυνατότητα στους ξενοδόχους να διατηρούν πιο προσωπικές σχέσεις με τους επισκέπτες τους αυξάνοντας έτσι τις πιθανότητες να επιστρέψουν. Άλλη εφαρμογή των μεγάλων δεδομένων είναι αυτή της χρήση έξυπνων βραχιολιών. Οι έξυπνες αυτές συσκευές δίνουν τη δυνατότητα στους επισκέπτες ενός καταλύματος μέσω του ενσωματωμένου συστήματος ραδιοσυχνότητας που έχουν, να ανοίξουν το δωμάτιο τους δίχως τη χρήση κλειδιού. Επίσης, έχουν τη δυνατότητα αποθήκευσης πληροφοριών των πιστωτικών ή χρεωστικών καρτών για ευκολότερες αγορές και μπορούν να χρησιμοποιηθούν σε πολυσύχναστα μουσεία και αξιοθέατα όπου θα επιτρέπουν την αποφυγή της συμφόρησης μέσω δυνατότητας γρήγορης εισόδου. Επιπλέον, οι κρυμμένοι αισθητήρες προσφέρουν πληροφορίες σχετικά με τον εκάστοτε χρήστη (όπως όνομα, γενέθλια κλπ)

Η διαχείριση μεγάλου όγκου δεδομένων στον Τουρισμό- Μια προσπάθεια βιβλιογραφικής και εμπειρικής προσέγγισης του θεματος

που αποκαλύπτονται με τη σάρωση τους οι οποίες μπορούν να συμβάλουν στην παροχή πιο προσωποποιημένων υπηρεσιών.

Βιβλιογραφία

- Alves, A., Baptista, C., Firmino, A., de Oliveira, M., & de Paiva, A. (2014). Comparison of SVM versus Naive-Bayes techniques for sentiment analysis in tweets: A case study with the 2013 FIFA Confederations Cup. Proceedings of the 20th Brazilian Symposium on Multimedia and the Web, Web-Media 2014 pp. 119–129.
- Anderson, C. (2012). The impact of social media on lodging performance. Cornell Hospitality Reports, pp 2–15.
- Andrea Ceron, Luigi Curini and Stefano M. Iacus, (2017). Politics and Big Data Nowcasting and Forecasting Elections with Social Media, pp. 1-4.
- Andrejevic, M., & Burdon, M. (2014). Defining the sensor society. Television & New Media.
- Arvind Sathi, (2014). Engaging Customers Using Big Data Analytics, How Marketing Analytics are Transforming Business, pp. 76-85.
- Barreda, A., & Bilgihan, A. (2013). An analysis of user-generated content for hotel experiences. Journal of Hospitality and Tourism, pp 259–279.
- BBC. (2014, November 19). Trip Advisor bad review ‘fine’ to be refunded by Blackpool hotel.
- Berezina, K., Bilgihan, A., Cobanoglu, C., & Okumus, F. (2015). Understanding satisfied and dissatisfied hotel customers: Texting mining of online hotel reviews. Journal of Hospitality Marketing & Management, pp 2–25.
- Blal, I., & Sturman, M. C. (2014). The differential effects of the quality and quantity of online reviews on hotel room sales. Cornell Hospitality Quarterly, pp 362–380.
- Chase, C. W. (2013). Using big data to enhance demand-driven forecasting and planning. Journal of Business Forecasting, pp. 26–33.

- Chiu, C., Chiu, N.-H., Sunga, R.-J., & Hsieh, P.-Y. (2015). Opinion mining of hotel customer-generated contents in Chinese weblogs. *Current Issues in Tourism*, pp 480–493.
- Duverger, P. (2013). Curvilinear effects of user-generated content of hotels' market share: A dynamic panel-data analysis. *Journal of Travel Research*, pp 462–480.
- Forrester Research. (2007). 60 percent of Europeans have adopted social computing. <https://www.forrester.com/Europeans+Have+Adopted+Social+Computing+Differently/fulltext/-/E-res42156>
- Four Seasons. (2012). The luxury consumer in the new digital world: Then & now. Four Seasons Group. http://www.fourseasons.com/content/dam/fourseasons/web/pdfs/landing_page_pdfs/2012_TRD_Report_final.pdf
- Garcí'a, A., Gaines, S., & Linaza, M. (2012). A lexicon-based sentiment analysis retrieval system for tourism domain. *e-Review of Tourism Research*, pp 31–40.
- Grabner, D., Zanker, M., Fliedl, G., & Fuchs, M. (2012). Classification of customer reviews based on sentiment analysis. In M. Fuchs, F. Ricci, & L. Cantoni (Eds.), *Information and communication technologies in tourism* pp. 459–469.
- Gretzel, U., Sigala, M., Xiang, Z., & Koo, C. (2015). Smart tourism: Foundations and developments. *Electronic Markets*
- Kasper, W., & Vela, M. (2011). Sentiment analysis for hotel reviews, *Proceedings of the Computational Linguistics-Applications Conference*, pp. 41–55.
- Kiran, G., Shankar, R., & Pudi, V. (2010). Frequent itemset based hierarchical document clustering using Wikipedia as external knowledge. In R. Setchi, I. Jordanov, R. Howlett, & L. Jain (Eds.), *Knowledge-based and intelligent information and engineering systems—lecture notes in computer science* (Vol. 6277, pp. 9–19).
- Leung, D., Law, R., van Hoof, H., & Buhalis, D. (2013). Social media in tourism and hospitality: A literature review. *Journal of Travel & Tourism Marketing*, pp 1–25.
- Lexhagen, M., Kuttainen, C., Fuchs, M., & Höpken, W. (2012). Destination talk in social media: a content analysis for innovation. In E. Christou, D. Chionis, D. Gursory,

& M. Sigala (Eds.), *Advances in hospitality and tourism marketing & management*. Corfu.

- Li, G., Law, R., Vu, H. Q., Rong, J., & Zhao, X. R. (2015). Identifying emerging hotel preferences using emerging pattern mining technique. *Tourism Management*, pp 309–319.
- Lupton, D. (2014). Self-tracking modes: Reflexive self-monitoring and data practices. In the “Imminent Citizenships: Personhood and Identity Politics in the Informatic Age” workshop.
- Markopoulos, G., Mikros, G., Iliadi, A., & Lontos, M. (2015). Sentiment analysis of hotel reviews in Greek: A comparison of unigram features of cultural tourism in a digital era. In *Springer proceedings in business and economics 2015* pp. 369–379.
- Miner, G., Delen, D., Elder, J., Fast, A., Hill, T., & Nisbet, R. (2012). *Practical text mining and statistical analysis for non-structured text data applications*. Waltham: Elsevier.
- National Academies Press, (2013). *Frontiers in Massive Data Analysis*, Chapter: 10 The Seven Computational Giants of Massive Data Analysis. <https://www.nap.edu/read/18374/chapter/12>
- Ng, E. C. (2012). Forecasting US recessions with various risk factors and dynamic probit models. *Journal of Macroeconomics*, pp 110–127.
- O’Connor, P. (2010). Managing a hotel’s image on TripAdvisor. *Journal of Hospitality Marketing & Management*, pp 724–730.
- Oguta, H., & Cezara, A. (2012). The factors affecting writing reviews in hotel websites. *Procedia- Social and Behavioral Sciences*, pp 978–987.
- Ong, B. S. (2012). The perceived influence of user reviews in the hospitality industry. *Journal of Hospitality Marketing & Management*, pp 460–487.
- Pablos, A., Cuadros, M., & Linaza, M. (2015). OpeNER: Open tools to perform natural language processing on accommodation. In I. Tussyadiah & A. Inversini (Eds.), *Information and communication technologies in tourism* pp. 115–140.

- Park, S. Y., & Allen, J. P. (2013). Responding to online reviews: Problem solving and engagement in hotels. *Cornell Hospitality Quarterly*, pp 62–75.
- Peter Buneman, Susan Davidson, Mary Fernandez και Dan Suciu, (1996). Adding Structure to Unstructured Data, pp. 1-3
- Phillips, P., Zigan, K., Silva, M. M. S., & Schegg, R. (2015). The interactive effects of online reviews on the determinants of Swiss hotel performance: A neural network analysis. *Tourism Management*, 129–139.
- Ralph Jacobson, (2013). 2.5 quintillion bytes of data created every day. How does CPG & Retail manage it? <https://www.ibm.com/blogs/insights-on-business/consumer-products/2-5-quintillion-bytes-of-data-created-every-day-how-does-cpg-retail-manage-it/>
- Schaal, D. (2013, August 26). Travelers visit 38 sites before booking a vacation, study says. <http://skift.com/2013/08/26/travelers-visit-38-sitesbefore-booking-a-vacation-study-says/>
- Schmunk, S., Hopken, W., Fuchs, M., & Lexhagen, M. (2014). Sentiment analysis—extracting decision-relevant knowledge from UGC. In Z. Xiang & I. Tussyadiah (Eds.), *Information and communication technologies in tourism*, pp. 249–270.
- Sparks, B. A., & Bradley, G. L. (2014). A “Triple A” typology of responding to negative consumer-generated online reviews. *Journal of Hospitality & Tourism Research*.
- Sparks, B. A., Perkins, H. E., & Buckley, R. (2013). Online travel reviews as persuasive communication: The effects of content type, source, and certification logos on consumer behaviour. *Tourism Management*, pp 2–11.
- Sparks, B., & Browning, V. (2011). The impact of online reviews on hotel booking intentions and perception of trust. *Tourism Management*, pp 1300–1319.
- Steve LaValle, Eric Lesser, Rebecca Shockley, Michael S. Hopkins και Nina Kruschwitz, (2011). Big Data Analytics and the Path From Insights to Value. *MIT Sloan Management Review*.

- Tansel Ozyer και Reda Alhaji, (2018). Machine Learning Techniques for Online Social Networks, pp. 26-57
- Thi Mai Le και Shu-Yi Liaw, (2017). Effects of Pros and Cons of Applying Big Data Analytics to Consumers' Responses in an E-Commerce Context. Department of Tropical Agriculture and International Cooperation, National Pingtung University of Science and Technology, Management College, Computer Centre, National Pingtung University of Science and Technology
- Tibshirani, R. (1996). Regression shrinkage and selection via the LASSO. Journal of the Royal Statistical Society Series B, pp. 270-290.
- Travel Media Group. (2015, March). Hoteliers' guide to responding online reviews. <http://travelmediagroup.com/wp-content/uploads/2015/03/TMG-How-To-Respond-To-Online-Reviews-Whitepaper.pdf>
- TripAdvisor. (2014). Fact Sheet. TripAdvisor. http://www.tripadvisor.com.tw/PressCenter-c4-Fact_Sheet.html
- Tsytarau, M., & Palpanas, T. (2011). Survey on mining subjective data on the web. Trento: Springer.
- USA Today. (2012, October 9). Wyndham to invite guests to review its hotels on TripAdvisor. <http://www.usatoday.com/story/hotelcheckin/2012/10/09/wyndham-tripadvisor-consumer-reviews/1622111/>
- Vermeulen, I. E., & Seegers, D. (2009). Tried and tested: The impact of online hotel reviews on consumer consideration. Tourism Management, pp 119–130.
- Waldhor, K., & Rind, A. (2008). E-BlogAnalysis—Mining virtual communities using statistical and linguistic methods for quality control in tourism. In P. O'Connor, W. Hopken, & U. Gretzel (Eds.), Information and communication technologies in tourism pp. 449–459.
- Wanchun Zhu, (2016). Analysis of the Application of Big Data in Intelligent Tourism Mode. Guizhou Minzu University, Guiyang.
- Weichselbraun, A., Gindl, S., & Scharl, A. (2013). Extracting and grounding context-aware. Sentiment Lexicons, pp 40–50.

- Withiam, G. (1995). Measuring guest perceptions: Combine measurement tools. *Cornell Hotel and Restaurant Administration Quarterly*, pp 13-17.
- Xiang, Z., Schwartz, Z., & Uysal, M. (2015). What types of hotels make their guests (un-) happy? Text analytics of customer experiences in online reviews. In I. Tussyadiah & A. Inversini (Eds.), *Information and communication technologies in tourism* pp. 29–44.
- Ye, Q., Law, R., Gu, B., & Chen, W. (2011). The influence of user-generated content on traveler behavior: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings. *Computers in Human Behavior*, pp 631–641.
- Ye, Q., Zhang, Z., & Law, R. (2009). Sentiment classification of online reviews to travel destinations by supervised machine learning approaches. *Expert Systems with Applications*, pp 657–670.
- Zhang, Y., & Va'squez, C. (2014). Hotel's responses to online reviews: Managing consumer dissatisfaction. *Discourse, Context and Media*, pp 51–67.

Παράρτημα 1

ΕΡΩΤΗΜΑΤΟΛΟΓΙΟ

1. Αρχικά λίγα λόγια για σας και για τα καθήκοντά σας στην εταιρεία;

Είμαι ο προϊστάμενος του τμήματος πληροφοριακών υποδομών του ομίλου Mouzenidis Travel.

2. Από πόσα άτομα αποτελείται το τμήμα σας και ποιες είναι οι κύριες αρμοδιότητές σας;

Αυτή τη στιγμή το τμήμα μας, απασχολεί 34 επαγγελματίες με εξειδίκευση στον τομέα της πληροφορικής.

3. Κατά τη γνώμη σας ποιος είναι ο ρόλος της ανάλυσης μεγάλων δεδομένων στον τουριστικό τομέα και τι προσφέρει;

Η ανάλυση μεγάλων δεδομένων (big data analytics) έχει επηρεάσει πολλούς τομείς της κοινωνίας όπως επίσης και τον τουριστικό τομέα. Είναι ένα επιπλέον ισχυρό εργαλείο για την διαπίστωση τάσεων τουριστικών πακέτων. Γειά παράδειγμα, με την ανάλυση μεγάλου όγκου δεδομένων που έχουν συλλεχθεί από site ταξιδιωτικών πακέτων, μπορεί να διαπιστώσει κανείς προτιμήσεις καταναλωτών για συγκεκριμένες χρονικές περιόδους. Παλαιότερα, κανείς χρειαζόταν να έχει εκτεταμένο δίκτυο συνεργατών ανα τον κόσμο, για να μπορέσει προσεγγίσει παρόμοια πηγή πληροφοριών.

4. Ποιες είναι οι κύριες πηγές δεδομένων που λαμβάνετε και ποιος είναι ο όγκος τους σε καθημερινή βάση;

Οι κυριότερες πηγές προέρχονται απ'τα portal site τουριστικών πακέτων. Από εκεί μπορεί κανείς να “αλιεύσει” πληροφορίες σχετικά με ποιές κοινωνικές ομάδες τι προτιμήσεις έχουν κατα τη διάρκεια της χρονιάς. Ο όγκος δεδομένων που συλλέγουμε κυμαίνεται από μερικές εκατοντάδες Megabytes έως μερικά Gigabytes καθημερινά. Αυτό αντιστοιχεί σε εκατοντάδες επισκέψεις σε portal sites, συμπεριλαμβανομένων αυτών των παραρτημάτων μας.

5. Ποιες είναι οι διαδικασίες εξαγωγής, αποθήκευσης και επεξεργασίας τους;

Αρικά, όλα τα ακατέργαστα δεδομένα αποθηκεύονται σε μια σχεσιακή βάση δεδομένων. Αυτό επιτυγχάνεται κατα την επίσκεψη των site που έχουν επιλεχθεί ως πηγές περισυλλογής δεδομένων. Όλες οι κινήσεις των επισκεπτών καταγράφονται, όπως ποιές ταξιδιωτικές επιλογές επιλέχθηκαν, η ηλικία τους (η οποία καταγράφεται κατα την εθελοντική συμπλήρωση σχετικού πεδίου απ'τον χρήστη, πόσο χρόνο δαπάνησαν για την ανάγνωση σχετικού προϊόντος, καθώς και λέξεις κλειδιά που χρησιμοποίησαν για την αναζήτηση ταξιδιωτικών προορισμών. Η κατηγοριοποίηση και επεξεργασία γίνεται με βάση database schemas (προδιαγραφές αποθήκευσης δεδομένων) που έχουν αναπτυχθεί σε συνεργασία με ειδικούς αναλυτές του τουριστικού τομέα οι οποίοι αξιολογούν τις πληροφορίες αυτές για την ανάπτυξη νέων προϊόντων.

6. Ποιες τεχνικές ανάλυσης δεδομένων χρησιμοποιείτε για τον προσδιορισμό δημογραφικών στοιχείων τουριστών, αγαπημένων προορισμών, αγαπημένων ξενοδοχείων, αγαπημένων τουριστικών πακέτων, περίοδοι υψηλής – χαμηλής ζήτησης;

Καταγράφουμε τις στατιστικές κατανομές των δεδομένων που έχουμε περισυλλέξει με βάση κάποια κριτήρια ενδιαφέροντος, συμπεριλαμβανομένων αρκετών που αναφέρονται στην παραπάνω ερώτηση. Μετά από στατιστική ανάλυση, όπως των υπολογισμό του συσχετισμού (correlation) των προτιμήσεων των χρηστών ανάλογα με την ηλικιακή τους ομάδα, και την τουριστική περίοδο.

7. Ο όμιλος Mouzenidis Travel διαχειρίζεται πολλές ιστοσελίδες που δίνουν τη δυνατότητα στους πελάτες να εξερευνήσουν τα προσφερόμενα προϊόντα, να κάνουν κρατήσεις καθώς και να αφήνουν feedback. Πως αξιοποιούνται αυτές οι πληροφορίες;

Όπως ανέφερα παραπάνω, οι ιστοσελίδες αυτές είναι οι κύριες πηγές συλλογής δεδομένων για την καταγραφή των προτιμήσεων των πελατών μας. Διαπιστώνουμε ότι το εθελοντικό feedback των χρηστών, αν και περιορισμένο σε αριθμό, είναι ιδιαίτερα χρήσιμο καθώς μπορεί να χρησιμοποιηθεί για να επαληθεύσει ή να διαψεύσει τα οποία συμπεράσματα έχουμε βγάλει απ'τη στατιστική ανάλυση που έχουμε κάνει στα raw δεδομένα.

8. Πως αξιοποιούνται τα δεδομένα για την πρόβλεψη της τουριστικής ζήτησης;

Αυτή τη στιγμή η ανάλυση των δεδομένων είναι σε πρωταρχική μορφή. Στηρίζομαστε στην αξιοποίηση στοιχείων παλαιότερων ετών για την πρόβλεψη της τουριστικής ζήτησης. Στο μέλλον σκοπεύουμε να βελτιώσουμε τις μεθόδους ανάλυσης δεδομένων ώστε να αναλύσουμε επιπλέον παραμέτρους όπως την χρονική διάρκεια που δαπανησε ο χρήστης για την ανάγνωση κάποιου προϊόντος.

9. Πως αξιοποιούνται τα δεδομένα για τον προσδιορισμό των αναγκών και των ενδοεταιρικών προβλημάτων;

Αυτή τη στιγμή δεν αξιοποιούμε τεχνικές big data analytics, για τον προσδιορισμό ενδοεταιρικών θεμάτων.

10. Πως αξιοποιούνται τα δεδομένα για την διευκόλυνση και υποστήριξη των αποφάσεων της διεύθυνσης;

Τα συμπεράσματα που έχουν προκύψει απ'την ανάλυση δεδομένων αξιοποιούνται για την μεσοπρόθεσμη και μακροπρόθεσμη στρατηγική της επιχείρησης. Η ταση της αγοράς, όπως προκύπτει απ'την ανάλυση, επηρεάζει το κέντρο λήψης αποφάσεων για την επέκταση ή ενίσχυση νέων αγορών.

11. Πως αξιοποιούνται τα δεδομένα για τον προσδιορισμό νέων ευκαιριών, επιχειρηματικών ιδεών και νέων αγορών;

Έχουμε διαπιστώσει ταχύτερες μεταβολές στις προτιμήσεις στο target group. Τεχνικές βασισμένες σε big data analytics μας επιτρέπουν να αντιδράσουμε σε αυτές μεταβολές. Για παράδειγμα αυτές μπορούν να αξιοποιηθούν για την ανάπτυξη προϊόντων με αντικείμενο τον θρησκευτικό τουρισμό. Απ'την ανάλυση δεδομένων μπορεί να προκύψει ποιές τουριστικές δραστηριότητες είναι δημοφιλείς τη δεδομένη εποχή.

12. Πως αξιοποιούνται τα δεδομένα για την αξιολόγηση της επίδοσης του ομίλου;

Αξιολογούμε τον βαθμό επιτυχίας των νέων προϊόντων που έχουν αναπτυχθεί απ'τον όμιλο, με βάση την επισκεψιμότητα, αγορά και τον αριθμό των shares που έχουν γίνει στα κοινωνικά δίκτυα. Επίσης δίνουμε μεγάλη βαρύτητα στο feedback των πελατών μας ούτως ώστε συνέχεια να βελτιώνουμε τις υπηρεσίες και τα προϊόντα μας.

13. Πως έχουν επηρεάσει οι κανονισμοί Προστασίας Δεδομένων (GDPR) την χρήση και αξιοποίηση των δεδομένων;

Είναι πλέον απαραίτητο να ζητάμε την συναίνεση των χρηστών για την νόμιμη συλλογή δεδομένων των κινήσεών τους. Σε τεχνικό επίπεδο, οι αλλαγές ήταν μικρότερες απ' ότι περιμέναμε. Μια αλλαγή στο front-end των ιστοσελίδων μας, και την καταγραφή της αποδοχής ή απόρριψης των όρων καταγραφής με βάση των προδιαγραφών GDPR ήταν αρκετό καθώς οι καταγραφές που κάνουμε είναι ανώνυμη. Σε επίπεδο συμπεριφοράς των χρηστών μας έχει γενικότερα θετική επίδραση καθώς τα feedback αυξήθηκαν πιθανώς επειδή έχει ενισχύσει την εμπιστοσύνη των χρηστών στο θέμα των προσωπικών δεδομένων μας.

14. Τέλος, ποιοι είναι οι μελλοντικοί στόχοι του τμήματος σας και ποιες προκλήσεις έχετε να αντιμετωπίσετε;

Οι προκλήσεις που αντιμετωπίζουμε έχουν να κάνουν με τις αναβαθμίσεις του υλικού και λογισμικού των πληροφοριακών συστημάτων μας επειδή οι νέες τεχνικές συλλογής, αποθήκευσης και επεξεργασίας δεδομένων, απαιτεί μεγάλη υπολογιστική ισχύει και network bandwidth. Μια επιπλέον πρόκληση έχει να κάνει με την παρακολούθηση και αξιοποίηση νέων τεχνολογιών του ταχύτατα εξελισσόμενο τομέα των big data analytics.

Παράρτημα 2

ΕΡΩΤΗΜΑΤΟΛΟΓΙΟ

1. Αρχικά λίγα λόγια για σας και για τα καθήκοντά σας στην εταιρεία;

Ονομάζομαι Σωκράτης Παπουνίδης και είμαι επικεφαλής του τμήματος IT της αεροπορικής εταιρίας Ellinair.

2. Από πόσα άτομα αποτελείται το τμήμα σας και ποιες είναι οι κύριες αρμοδιότητές σας;

Το τμήμα αποτελείται από 5 άτομα: 2 – προγραμματιστές, 2 – hardware & support και εγώ.

3. Κατά τη γνώμη σας ποιος είναι ο ρόλος της ανάλυσης μεγάλων δεδομένων στον τομέα των αερομεταφορών και ευρύτερα στον τουριστικό τομέα και τι προσφέρει;

Τα δεδομένα, είτε είναι μεγάλα, είτε μικρά, αποτελούν την βασική πηγή πληροφόρησης για την λήψη αποφάσεων σε κάθε σύγχρονη επιχείρηση. Η αναγκαιότητα εξόρυξης γνώσης από την ωμή πληροφορία και χρήση της οδηγεί σε ανάπτυξη και υιοθέτηση ειδικευμένων συστημάτων και τεχνικών για άμεση ανάλυση και προώθηση αποτελεσμάτων της. Στην αεροπορική εταιρία μας τα δεδομένα που συλλέγουμε χρησιμοποιούνται για την μελέτη και πρόβλεψη τρέχουσας τιμής του εισιτηρίου, στατιστικά, απλοποίηση και βελτιστοποίηση διαδικασιών κλπ.

4. Ποιες είναι οι κύριες πηγές δεδομένων που λαμβάνετε και ποιος είναι ο όγκος τους σε καθημερινή βάση;

Η κύρια πηγή είναι το ERP της επιχείρησης, στατιστικά επισκεψιμότητας των ιστοσελίδων, logs & metrics

5. Ποιες είναι οι διαδικασίες εξαγωγής, αποθήκευσης και επεξεργασίας τους;

Έχουμε διάφορους τρόπους εξαγωγής δεδομένων, όπως οι αυτοματοποιημένη διαδικασία ETL (Extract, Transform, Load), που κατεβάζει, τροποποιεί, συνδυάζει και αποθηκεύει δεδομένα για την εύκολη ανάλυση στην τοπική βάση. Τα δεδομένα αυτά σε πολλές περιπτώσεις αποτελούν το input για διάφορα reports.

6. Η ιστοσελίδες της Ellinair και της Mouzenidis Travel θα αποτελούν σημαντικές πηγές πληροφοριών για την εταιρεία. Μπορείτε να μας περιγράψετε συνοπτικά τα είδη των δεδομένων, τις μεθόδους επεξεργασίας καθώς και τις διαδικασίες ανάλυσης και αξιοποίησης των δεδομένων που συλλέγονται;

Εκτός από τα δεδομένα πωλήσεων συλλέγουμε και δεδομένα επισκεψιμότητας. Τα στατιστικά αυτά είναι προσβάσιμα από την ίδια την πλατφόρμα GA, αλλά ταυτόχρονα χρησιμοποιούνται και στα daily reports για τα KPI της εταιρίας.

Επίσης συλλέγουμε και τα webserver logs, συνήθως για την επίλυση των προβλημάτων στα websites μας.

7. Πως αξιοποιούνται τα δεδομένα για την πρόβλεψη της ζήτησης;

Για αυτό το σκοπό έχουμε ειδικό τμήμα Revenue Analysis Department, το οποίο χρησιμοποιεί σύστημα Revenue Management για την πρόβλεψη ζήτησης. Όλα τα ιστορικά δεδομένα που αναφέραμε πριν τροφοδοτούν το σύστημα αυτό.

8. Πως αξιοποιούνται τα δεδομένα για τον προσδιορισμό των αναγκών και των ενδοεταιρικών προβλημάτων;

Εκτός από τα ποσοτικά δεδομένα, οι αεροπορικές εταιρίες πολύ συχνά συλλέγουν και τα ποιοτικά δεδομένα σε μορφή surveys and questionnaires.

9. Πως αξιοποιούνται τα δεδομένα για την διευκόλυνση και υποστήριξη των αποφάσεων της διεύθυνσης;

Όπως ανέφερα και πριν, τα δεδομένα συλλέγονται από τις διάφορες πηγές και μετατρέπονται σε μορφή reports σε ευανάγνωστη μορφή. Οι αναφορές αυτές καθημερινά αποστέλλονται στα emails της διοίκησης ή είναι προσβάσιμα μέσω intranet portal.

10. Πως αξιοποιούνται τα δεδομένα για τον προσδιορισμό νέων ευκαιριών, επιχειρηματικών ιδεών και νέων αγορών;

Για της νέες αγορές χρησιμοποιούνται τα δεδομένα από τις τρίτες εταιρίες και αρχές, όπως οι αναφορές των αεροδρομίων, πολιτικής αεροπορίας κλπ.

11. Πως αξιοποιούνται τα δεδομένα για την αξιολόγηση της επίδοσης του ομίλου;

Για την αξιολόγηση της επίδοσης κατά κύριο λόγο γίνονται οι συγκρίσεις με τα δεδομένα των προηγούμενων περιόδων. Αυτό μας επιτρέπει να καταλάβουμε την κατεύθυνση, προς την οποία κινείται η επιχείρησή μας και αν ακολουθεί το αρχικό πλάνο των budgets and forecasts.

12. Πως έχουν επηρεάσει οι κανονισμοί Προστασίας Δεδομένων (GDPR) την χρήση και αξιοποίηση των δεδομένων;

Παρά την διαδεδομένη γνώμη ότι οι κανονισμοί αυτοί δυσκολεύουν και περιπλέκουν τις ήδη υπάρχουσες διαδικασίες, κατά την γνώμη μας αντιθέτως βοηθάνε στην κατηγοριοποίηση και ανάλυση των δεδομένων αυτόν με την μεγαλύτερη προσοχή.

Με την πάροδο χρονικών περιθωρίων τα δεδομένα αποπροσωποποιούνται και χρησιμοποιούνται μόνο για στατιστικούς λόγους.

13. Τέλος, ποιοι είναι οι μελλοντικοί στόχοι του τμήματος σας και ποιες προκλήσεις έχετε να αντιμετωπίσετε;

Οι στόχοι μας είναι να αυτοματοποιηθούν οι διαδικασίες συλλογής, αποθήκευσης και ανάλυσης των δεδομένων. Να απλοποιηθεί η πρόσβαση στα δεδομένα αυτά και στα αποτελέσματα των αναλύσεων σε μορφή αναφορών ή OLAP. Να χρησιμοποιηθούν οι τεχνολογίες machine learning για την ανάλυση, εξόρυξη και λήψη αποφάσεων.

Ευχαριστώ για τις απαντήσεις και τον χρόνο σας!