



ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΣΧΟΛΗ ΟΙΚΟΝΟΜΙΚΩΝ ΚΑΙ ΠΟΛΙΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΤΜΗΜΑ ΕΠΙΚΟΙΝΩΝΙΑΣ ΚΑΙ ΜΕΣΩΝ ΜΑΖΙΚΗΣ ΕΝΗΜΕΡΩΣΗΣ

ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ

**ΚΑΤΕΥΘΥΝΣΗ ΨΗΦΙΑΚΑ ΜΕΣΑ ΕΠΙΚΟΙΝΩΝΙΑΣ ΚΑΙ ΠΕΡΙΒΑΛΛΟΝΤΑ
ΑΛΛΗΛΕΠΙΔΡΑΣΗΣ**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**«Ανάλυση συναισθήματος σε ελληνικά και γερμανικά δεδομένα του
Twitter γύρω από την εκπαίδευση των προσφύγων με τη χρήση
lexicon- based και machine learning μεθόδων και εργαλείων της
γλώσσας R»**

Έλλη Βλάχου

ΑΘΗΝΑ

ΑΥΓΟΥΣΤΟΣ 2017

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

«Ανάλυση συναισθήματος σε ελληνικά και γερμανικά δεδομένα του Twitter γύρω από την εκπαίδευση των προσφύγων με τη χρήση lexicon- based και machine learning μεθόδων και εργαλείων της γλώσσας R»

Έλλη Βλάχου

A.M. 35057

ΕΠΙΒΛΕΠΩΝ:

Κωνσταντίνος Μουρλάς

ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ:

Στους γονείς μου, Ανέστη και Μαρία

ΠΕΡΙΛΗΨΗ

Στόχος της παρούσας διπλωματικής είναι η ταξινόμηση σύντομων μηνυμάτων από το κοινωνικό δίκτυο Twitter με γνώμονα το συναίσθημα που εκφράζεται μέσα από αυτά, χρησιμοποιώντας τεχνικές εξόρυξης δεδομένων. Τα μηνύματα στο Twitter, τα λεγόμενα tweets, περιορίζονται στους 140 χαρακτήρες. Οι χρήστες δυσκολεύονται να εκφράσουν κάποιο συναίσθημα μέσα σε ένα τόσο σύντομο μήνυμα, γεγονός που καθιστά την κατηγοριοποίηση του συναισθήματος ιδιαίτερος απαιτητική.

Αρχικά, θα γίνει προσπάθεια να κατηγοριοποιηθούν χειροκίνητα ως προς την πολικότητά τους μηνύματα του Twitter από Ελλάδα και Γερμανία που έχουν να κάνουν με την προσφυγική κρίση και συγκεκριμένα με την εκπαίδευση των προσφύγων και στη συνέχεια θα ανιχνευτεί ο στόχος του συναισθήματος, σε ποιον δηλαδή απευθύνεται το συναίσθημα που εκφράζεται σε καθένα από τα tweets. Στο πλαίσιο αυτό θα συζητηθούν όψεις του κοινωνικού αυτού θέματος, όπως το συναίσθημα που εκδηλώνουν τα Μ.Μ.Ε και οι πολίτες της κάθε χώρας προς τον εκάστοτε στόχο του συναισθήματος και το ενδεχόμενο τα μηνύματα αυτά να εκφράζουν ξενοφοβία ή ρατσισμό.

Κατόπιν, θα γίνει προσπάθεια πρόβλεψης του συναισθήματος με δύο διαφορετικές μεθόδους, αυτή του λεξικού και αυτή της μηχανικής μάθησης με τη χρήση του αλγορίθμου κατηγοριοποίησης Naive Bayes. Μετά την αξιολόγηση της αποτελεσματικότητας των δύο μεθόδων, θα επιλεγεί η πιο επιτυχημένη από τις δύο και θα γίνει πρόβλεψη του συναισθήματος που εκφράζεται για τον καθένα από τους στόχους του συναισθήματος ξεχωριστά.

ΤΙΤΛΟΣ: «Ανάλυση συναισθήματος σε ελληνικά και γερμανικά δεδομένα του Twitter γύρω από την εκπαίδευση των προσφύγων με τη χρήση lexicon- based και machine learning μεθόδων και εργαλείων της γλώσσας R»

ΘΕΜΑΤΙΚΗ ΠΕΡΙΟΧΗ : Κοινωνικές στάσεις σε πλαίσιο αυτόματης επεξεργασίας κειμενικών δεδομένων/επεξεργασίας φυσικού λόγου/ εξόρυξης κειμενικών δεδομένων

ΛΕΞΕΙΣ – ΚΛΕΙΔΙΑ : προσφυγικό, εκπαίδευση, ανάλυση συναισθήματος, κοινωνικά δίκτυα, (μη) επιβλεπόμενη μάθηση, λεξικά, μικρο- ιστολογία, ταξινομητής

ABSTRACT

The main aim of this thesis is to classify small messages from Twitter (tweets), according to their sentiment, using data mining techniques. 'Tweets' are limited to a maximum of 140 characters of text, which makes it even harder for people to express their sentiments. Thus, the classification process becomes even more difficult.

In the beginning, the researcher will attempt to detect the opinion targets of the tweets that have to do with the refugee crisis and particularly with the refugees' education and to categorize them according to their polarity. Thus, some aspects of the social issue of refugees are going to be discussed, for example the sentiment that the media and the users of each country express towards the opinion targets in Twitter and their possibility of expressing xenophobia or racism.

After that, the sentiment of the above-mentioned tweets is going to be predicted automatically by using two different methods, a lexicon- based, as well as a machine learning method based on the Naive Bayes algorithm. After evaluating both methods, the most successful is going to be chosen in order to predict the sentiment that is expressed for each one of the opinion targets.

TITLE: 'Sentiment analysis of Greek and German Twitter data regarding the education of refugees using lexicon-based and machine learning methods and tools of the R language'

RESEARCH FIELD: Social attitudes as part of automatic processing of textual data/natural language processing/ textual data mining

KEYWORDS: refugee crisis, education, sentiment analysis, social media, (un)supervised learning, lexicon, micro-blogging, classifier.

ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να ευχαριστήσω, καταρχάς, όλους τους καθηγητές του Π.Μ.Σ. «Ψηφιακά Μέσα Επικοινωνίας & Περιβάλλοντα Αλληλεπίδρασης» για τις γνώσεις που μου προσέφεραν κατά τη διάρκεια φοίτησης μου στο εν λόγω μεταπτυχιακό πρόγραμμα.

Ιδιαίτερος ευχαριστώ τον επιβλέποντά μου, Αναπληρωτή Καθηγητή κ. Κωνσταντίνο Μουρλά, για τη βοήθεια και τη στήριξη του κατά τη διάρκεια εκπόνησης της παρούσας μελέτης, αλλά και για την ευκαιρία που μου έδωσε να ασχοληθώ με ένα άκρως ενδιαφέρον και σύγχρονο αντικείμενο.

Ακόμη, θα ήθελα να ευχαριστήσω τους συμφοιτητές μου για όλες τις στιγμές που περάσαμε μαζί τα τελευταία δύο χρόνια.

Τέλος, ένα μεγάλο ευχαριστώ οφείλω στους γονείς μου που στέκονται πάντα δίπλα μου και στηρίζουν όλες τις επιλογές μου. Σε αυτούς οφείλω όλη τη διαδρομή των σπουδών μου μέχρι σήμερα.

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ	VIII
ΕΥΡΕΤΗΡΙΟ ΠΙΝΑΚΩΝ	XI
ΕΥΡΕΤΗΡΙΟ ΓΡΑΦΗΜΑΤΩΝ.....	XII
ΕΥΡΕΤΗΡΙΟ ΕΙΚΟΝΩΝ.....	XIV
ΠΙΝΑΚΑΣ ΣΥΝΤΜΗΣΕΩΝ.....	XV
ΕΙΣΑΓΩΓΗ	16
ΘΕΩΡΗΤΙΚΟ ΠΛΑΙΣΙΟ	18
1. Η ΠΡΟΣΦΥΓΙΚΗ ΚΡΙΣΗ	18
1.1 Ο ΠΡΟΣΦΥΓΑΣ ΚΑΙ Ο ΜΕΤΑΝΑΣΤΗΣ	18
1.2 ΟΙ ΠΡΟΣΦΥΓΕΣ ΣΤΗΝ ΕΥΡΩΠΗ	21
1.3 ΤΟ ΠΡΟΦΙΛ ΤΩΝ ΠΡΟΣΦΥΓΩΝ	28
1.4 ΑΝΗΛΙΚΟΙ ΠΡΟΣΦΥΓΕΣ.....	29
1.5 ΑΣΥΝΟΔΕΥΤΟΙ ΑΝΗΛΙΚΟΙ ΠΡΟΣΦΥΓΕΣ.....	30
1.6 ΤΑ ΔΙΚΑΙΩΜΑΤΑ ΤΩΝ ΠΡΟΣΦΥΓΩΝ.....	33
1.7 Η ΕΝΣΩΜΑΤΩΣΗ ΤΩΝ ΠΡΟΣΦΥΓΩΝ.....	37
1.8 ΤΟ ΔΙΚΑΙΩΜΑ ΤΩΝ ΠΡΟΣΦΥΓΩΝ ΣΤΗΝ ΕΚΠΑΙΔΕΥΣΗ	40
1.8.1 ΤΟ ΝΟΜΙΚΟ ΠΛΑΙΣΙΟ ΣΤΗΝ ΕΛΛΑΔΑ.....	42
1.8.2 ΤΟ ΝΟΜΙΚΟ ΠΛΑΙΣΙΟ ΣΤΗ ΓΕΡΜΑΝΙΑ.....	45
1.8.3 ΟΙ ΔΙΕΘΝΕΙΣ ΚΑΙ ΕΥΡΩΠΑΪΚΕΣ ΣΥΝΘΗΚΕΣ	46
1.9 Η ΠΡΑΓΜΑΤΙΚΟΤΗΤΑ ΓΙΑ ΤΗΝ ΕΚΠΑΙΔΕΥΣΗ ΤΩΝ ΠΡΟΣΦΥΓΩΝ	51
1.9.1 Η ΕΚΠΑΙΔΕΥΣΗ ΤΩΝ ΠΡΟΣΦΥΓΩΝ ΣΤΗΝ ΕΛΛΑΔΑ.....	53
1.9.2 Η ΕΚΠΑΙΔΕΥΣΗ ΤΩΝ ΠΡΟΣΦΥΓΩΝ ΣΤΗ ΓΕΡΜΑΝΙΑ	54
1.10 ΡΑΤΣΙΣΜΟΣ	61
1.11 ΞΕΝΟΦΟΒΙΑ.....	64
1.12 ΑΚΡΟΔΕΞΙΑ.....	71
1.12.1 Η ΑΚΡΟΔΕΞΙΑ ΣΤΗΝ ΕΛΛΑΔΑ	72
1.12.2 Η ΑΚΡΟΔΕΞΙΑ ΣΤΗ ΓΕΡΜΑΝΙΑ	73
1.13 Η ΞΕΝΟΦΟΒΙΑ ΚΑΙ Ο ΡΑΤΣΙΣΜΟΣ ΣΤΗΝ ΕΛΛΑΔΑ ΚΑΙ ΣΤΗ ΓΕΡΜΑΝΙΑ.....	75
1.13.1 ΤΙ ΛΕΝΕ ΟΙ ΕΡΕΥΝΕΣ.....	77

2.	ΤΑ ΚΟΙΝΩΝΙΚΑ ΔΙΚΤΥΑ (SOCIAL MEDIA)	87
2.1	ΚΑΤΗΓΟΡΙΟΠΟΙΗΣΗ ΤΩΝ ΜΕΣΩΝ ΚΟΙΝΩΝΙΚΗΣ ΔΙΚΥΤΩΣΗΣ	88
2.2	ΤΟ TWITTER	90
2.3	ΤΟ TWITTER ΓΙΑ ΤΟ ΠΡΟΣΦΥΓΙΚΟ ΖΗΤΗΜΑ	94
3.	ΕΞΟΡΥΞΗ ΔΕΔΟΜΕΝΩΝ ΚΑΙ ΑΝΑΛΥΣΗ ΣΥΝΑΙΣΘΗΜΑΤΟΣ	96
3.1	ΕΞΟΡΥΞΗ ΔΕΔΟΜΕΝΩΝ	96
3.1.1	Η ΔΙΑΔΙΚΑΣΙΑ ΕΞΟΡΥΞΗΣ ΔΕΔΟΜΕΝΩΝ	97
3.1.2	ΔΙΕΡΓΑΣΙΕΣ ΕΞΟΡΥΞΗΣ ΔΕΔΟΜΕΝΩΝ	100
3.2	ΕΞΟΡΥΞΗ ΓΝΩΜΗΣ ΚΑΙ ΑΝΑΛΥΣΗ ΣΥΝΑΙΣΘΗΜΑΤΟΣ	102
3.2.1	ΕΦΑΡΜΟΓΕΣ	105
3.2.2	ΟΡΙΣΜΟΣ ΤΗΣ ΑΠΟΨΗΣ	106
3.2.3	Η ΔΙΑΔΙΚΑΣΙΑ ΕΞΟΡΥΞΗΣ ΚΑΙ ΑΝΑΛΥΣΗΣ ΑΠΟΨΕΩΝ	108
3.2.4	ΔΙΑΦΟΡΕΤΙΚΑ ΕΠΙΠΕΔΑ ΑΝΑΛΥΣΗΣ	109
3.2.5	ΤΕΧΝΙΚΕΣ ΚΑΤΗΓΟΡΙΟΠΟΙΗΣΗΣ ΣΥΝΑΙΣΘΗΜΑΤΟΣ	112
3.2.5.1	ΤΕΧΝΙΚΕΣ ΒΑΣΙΣΜΕΝΕΣ ΣΕ ΛΕΞΙΚΟ (LEXICON BASED TECHNIQUES)	113
3.2.5.2	ΤΕΧΝΙΚΕΣ ΒΑΣΙΣΜΕΝΕΣ ΣΕ ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ (MACHINE LEARNING TECHNIQUES)	115
3.2.5.2.1	Η ΜΠΕΪΖΙΑΝΗ ΣΤΑΤΙΣΤΙΚΗ ΚΑΙ Ο ΑΛΓΟΡΙΘΜΟΣ ΝΑΪΒΕ ΒΑΥΕΣ	116
3.2.6	ΔΥΣΚΟΛΙΕΣ ΚΑΙ ΠΡΟΚΛΗΣΕΙΣ	121
3.2.7	ΑΝΑΛΥΣΗ ΣΥΝΑΙΣΘΗΜΑΤΟΣ ΣΤΟ TWITTER	125
	ΕΡΕΥΝΗΤΙΚΕΣ ΥΠΟΘΕΣΕΙΣ – ΕΡΕΥΝΗΤΙΚΑ ΕΡΩΤΗΜΑΤΑ	130
	ΜΕΘΟΔΟΛΟΓΙΑ ΕΡΕΥΝΑΣ	133
	ΥΛΟΠΟΙΗΣΗ ΕΡΕΥΝΑΣ	139
1.	ΣΥΛΛΟΓΗ ΔΕΔΟΜΕΝΩΝ	139
1.1	ΔΗΜΙΟΥΡΓΙΑ ΜΙΑΣ TWITTER- ΕΦΑΡΜΟΓΗΣ	139
1.2	ΠΙΣΤΟΠΟΙΗΣΗ ΜΕ ΤΗ ΧΡΗΣΗ ΤΟΥ TWITTER ROAUTH	139
1.3	ΔΙΕΠΑΦΕΣ ΤΟΥ TWITTER- REST vs STREAMING API	141
1.4	ΛΗΨΗ ΔΕΔΟΜΕΝΩΝ	142
2.	ΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ	144
2.1	ΚΑΤΗΓΟΡΙΟΠΟΙΗΣΗ ΣΥΝΑΙΣΘΗΜΑΤΟΣ ΚΑΙ ΕΞΑΓΩΓΗ ΤΩΝ OPINION TARGETS	144
2.2	ΔΙΑΧΩΡΙΣΜΟΣ ΤΩΝ TWEETS ΣΕ MEDIA ΚΑΙ USERS ΣΤΟ RSTUDIO	145
2.3	ΔΙΑΧΩΡΙΣΜΟΣ ΤΩΝ TWEETS ΜΕ ΒΑΣΗ ΤΟΥΣ ΣΤΟΧΟΥΣ ΣΥΝΑΙΣΘΗΜΑΤΟΣ ΣΤΟ RSTUDIO	146
3.	ΑΝΑΛΥΣΗ ΔΕΔΟΜΕΝΩΝ	147
3.1	ΥΠΟΛΟΓΙΣΜΟΣ ΤΟΥ ΣΥΝΑΙΣΘΗΜΑΤΟΣ ΚΑΙ ΕΞΑΓΩΓΗ ΤΩΝ ΣΤΟΧΩΝ ΤΟΥ ΜΕ ΒΑΣΗ ΤΗ ΧΕΙΡΩΝΑΚΤΙΚΗ ΕΠΙΣΗΜΕΙΩΣΗ	147
3.2	ΜΕΘΟΔΟΣ ΚΑΘΑΡΙΣΜΟΥ ΔΕΔΟΜΕΝΩΝ	147
3.3	ΑΝΑΛΥΣΗ ΣΥΝΑΙΣΘΗΜΑΤΟΣ ΜΕ ΤΗ ΜΕΘΟΔΟ ΤΟΥ ΛΕΞΙΚΟΥ	152

3.3.1	ΔΗΜΙΟΥΡΓΙΑ ΛΕΞΙΚΩΝ	152
3.3.2	ΕΚΤΕΛΕΣΗ ΜΕΘΟΔΟΥ.....	156
3.4	ΑΝΑΛΥΣΗ ΣΥΝΑΙΣΘΗΜΑΤΟΣ ΜΕ ΤΟΝ ΑΛΓΟΡΙΘΜΟ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ ΝΑΙΒΕ ΒΑΥΕΣ.....	157
3.4.1	ΤΡΟΠΟΠΟΙΗΣΕΙΣ ΑΛΓΟΡΙΘΜΟΥ ΓΙΑ ΒΕΛΤΙΩΣΗ ΤΗΣ ΑΠΟΔΟΣΗΣ.....	160
3.5	ΑΝΑΛΥΣΗ ΣΥΝΑΙΣΘΗΜΑΤΟΣ ΜΕ ΤΗ ΜΕΘΟΔΟ ΤΟΥ ΛΕΞΙΚΟΥ ΑΝΑ ΚΑΤΗΓΟΡΙΑ ΤΩΝ ΣΤΟΧΩΝ ΤΟΥ ΣΥΝΑΙΣΘΗΜΑΤΟΣ.....	168
	ΑΠΟΤΕΛΕΣΜΑΤΑ ΕΡΕΥΝΑΣ & ΣΥΜΠΕΡΑΣΜΑΤΑ.....	169
1.	ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΝΑΛΥΣΗΣ ΠΕΡΙΕΧΟΜΕΝΟΥ	169
1.1	ΣΤΟΧΟΙ ΣΥΝΑΙΣΘΗΜΑΤΟΣ	169
1.2	ΣΥΝΑΙΣΘΗΜΑ	172
2.	ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΝΑΛΥΣΗΣ ΣΥΝΑΙΣΘΗΜΑΤΟΣ	177
2.1	ΑΠΟΔΟΣΗ ΤΩΝ ΔΥΟ ΜΕΘΟΔΩΝ.....	177
2.1.1	ΑΠΟΔΟΣΗ ΤΟΥ ΑΛΓΟΡΙΘΜΟΥ ΝΑΙΒΕ ΒΑΥΕΣ	178
2.1.2	ΑΠΟΔΟΣΗ ΤΗΣ LEXICON- BASED ΜΕΘΟΔΟΥ.....	191
2.2	ΣΥΓΚΡΙΣΗ ΜΕΘΟΔΩΝ.....	177
3.	ΣΥΜΠΕΡΑΣΜΑΤΑ	201
4.	ΠΕΡΙΟΡΙΣΜΟΙ ΕΡΕΥΝΑΣ & ΠΡΟΤΑΣΕΙΣ ΓΙΑ ΜΕΛΛΟΝΤΙΚΕΣ ΕΡΕΥΝΕΣ.....	204
	ΒΙΒΛΙΟΓΡΑΦΙΚΕΣ ΑΝΑΦΟΡΕΣ	206
	ΠΑΡΑΡΤΗΜΑ Α: ΠΙΝΑΚΕΣ ΑΠΟΤΕΛΕΣΜΑΤΩΝ.....	230
	ΠΑΡΑΡΤΗΜΑ Β: ΚΩΔΙΚΑΣ	238

ΕΥΡΕΤΗΡΙΟ ΠΙΝΑΚΩΝ

Πίνακας 1. Δείκτες μέτρησης ενσωμάτωσης προσφύγων	38
Πίνακας 2. Αλγόριθμοι μηχανικής μάθησης	120
Πίνακας 3. Ελληνικά και γερμανικά keywords	143
Πίνακας 4. Διαδικασία καθαρισμού δεδομένων	151
Πίνακας 5. Θετικές και αρνητικές λέξεις που προστέθηκαν στο ελληνικό λεξικό	153
Πίνακας 6. Θετικές και αρνητικές λέξεις που προστέθηκαν στο γερμανικό λεξικό	156
Πίνακας 7. Οι στόχοι συναισθήματος στα ελληνικά tweets.....	172
Πίνακας 8. Οι στόχοι συναισθήματος στα γερμανικά tweets.....	172
Πίνακας 9. Σύγκριση μεθόδων ανάλυσης συναισθήματος.....	178

ΕΥΡΕΤΗΡΙΟ ΓΡΑΦΗΜΑΤΩΝ

Γράφημα 1. Στόχοι συναισθήματος.....	170
Γράφημα 2. Συναισθημα- Greek media.....	173
Γράφημα 3. Συναισθημα- Greek users.....	173
Γράφημα 4. Συναισθημα- German media	173
Γράφημα 5. Συναισθημα- German users	173
Γράφημα 6. Συναισθημα για ακροδεξιά- Greek media	174
Γράφημα 7. Συναισθημα για ακροδεξιά- Greek users.....	174
Γράφημα 8. Συναισθημα για ακροδεξιά- German media.....	174
Γράφημα 9. Συναισθημα για ακροδεξιά- German users.	174
Γράφημα 10. Συναισθημα για πολιτική- Greek media.....	175
Γράφημα 11. Συναισθημα για πολιτική- Greek users.	175
Γράφημα 12. Συναισθημα για πολιτική- German media..	175
Γράφημα 13. Συναισθημα για πολιτική- German users.....	175
Γράφημα 14. Συναισθημα για κοινωνία- Greek media.....	176
Γράφημα 15. Συναισθημα για κοινωνία- Greek users.	176
Γράφημα 16. Συναισθημα για κοινωνία- German media.....	176
Γράφημα 17. Συναισθημα για κοινωνία- German users	176
Γράφημα 18. Συναισθημα για πρόσφυγες- Greek media.....	177
Γράφημα 19. Συναισθημα για πρόσφυγες- Greek users.....	177
Γράφημα 20. Συναισθημα για πρόσφυγες- German media.....	177
Γράφημα 21. Συναισθημα για πρόσφυγες- German users.	177
Γράφημα 22. Σύγκριση μεθόδων λεξικού και αλγορίθμου μηχανικής μάθησης	178
Γράφημα 23. Εξάρτηση υπολογιστικού χρόνου σε σχέση με το πλήθος των πρώτων γραμμάτων των αρχικών συνολικών λέξεων στην περίπτωση των MME.....	179
Γράφημα 24. Πλήθος κοινών ριζών σε σχέση με το πλήθος των πρώτων γραμμάτων των αρχικών συνολικών λέξεων στην περίπτωση των MME.....	180
Γράφημα 25. Εξάρτηση υπολογιστικού χρόνου σε σχέση με το πλήθος των πρώτων γραμμάτων των αρχικών συνολικών λέξεων στην περίπτωση των ιδιωτών χρηστών..	181
Γράφημα 26. Πλήθος κοινών ριζών σε σχέση με το πλήθος των πρώτων γραμμάτων των αρχικών συνολικών λέξεων στην περίπτωση των ιδιωτών χρηστών.....	182
Γράφημα 27. Απόδοση αλγορίθμου σε σχέση με το πλήθος των πρώτων γραμμάτων των αρχικών συνολικών λέξεων στην περίπτωση των ιδιωτών χρηστών.....	183

Γράφημα 28. Απόδοση αλγόριθμου σε σχέση με το πλήθος των πρώτων γραμμάτων των αρχικών συνολικών λέξεων στην περίπτωση των ΜΜΕ.....	184
Γράφημα 29. Εξάρτηση του υπολογιστικού χρόνου σε σχέση με το πλήθος των συνθετικών tweets.....	186
Γράφημα 30. Απόδοση του αλγορίθμου σε σχέση με το πλήθος των συνθετικών tweets.....	188
Γράφημα 31. Απόδοση του αλγορίθμου σε σχέση με τον κανόνα-κριτήριο χαρακτηρισμού συναισθήματος των συνθετικών tweets.....	190
Γράφημα 32. Απόδοση του αλγορίθμου σε σχέση με το μέγεθος του δείγματος ελέγχου ως ποσοστό του αρχικού δείγματος των συνθετικών tweets .	191
Γράφημα 33. Αποτελέσματα μεθόδου λεξικού- Greek media.	192
Γράφημα 34. Αποτελέσματα μεθόδου λεξικού- Greek users.....	192
Γράφημα 35. Αποτελέσματα μεθόδου λεξικού- German media.	192
Γράφημα 36. Αποτελέσματα μεθόδου λεξικού- German users.....	192
Γράφημα 37. Αποτελέσματα μεθόδου λεξικού για ακροδεξιά- Greek media.....	193
Γράφημα 38. Αποτελέσματα μεθόδου λεξικού για ακροδεξιά - Greek users.....	193
Γράφημα 39. Αποτελέσματα μεθόδου λεξικού για ακροδεξιά - German media.....	193
Γράφημα 40. Αποτελέσματα μεθόδου λεξικού για ακροδεξιά - German users.....	193
Γράφημα 41. Αποτελέσματα μεθόδου λεξικού για πολιτική- Greek media.....	194
Γράφημα 42. Αποτελέσματα μεθόδου λεξικού για πολιτική - Greek users.....	194
Γράφημα 43. Αποτελέσματα μεθόδου λεξικού για πολιτική - German media.....	194
Γράφημα 44. Αποτελέσματα μεθόδου λεξικού για πολιτική - German users.....	194
Γράφημα 45. Αποτελέσματα μεθόδου λεξικού για κοινωνία- Greek media.....	195
Γράφημα 46. Αποτελέσματα μεθόδου λεξικού για κοινωνία - Greek users.....	195
Γράφημα 47. Αποτελέσματα μεθόδου λεξικού για κοινωνία - German media.....	195
Γράφημα 48. Αποτελέσματα μεθόδου λεξικού για κοινωνία - German users.....	195
Γράφημα 49. Αποτελέσματα μεθόδου λεξικού για πρόσφυγες- Greek media.....	196
Γράφημα 50. Αποτελέσματα μεθόδου λεξικού για πρόσφυγες - Greek users.....	196
Γράφημα 51. Αποτελέσματα μεθόδου λεξικού για πρόσφυγες - German media.....	196
Γράφημα 52. Αποτελέσματα μεθόδου λεξικού για πρόσφυγες - German users.....	196

ΕΥΡΕΤΗΡΙΟ ΕΙΚΟΝΩΝ

Εικόνα 1. Οι διαδρομές των προσφύγων.	24
Εικόνα 2. Χώρες προορισμού των προσφύγων.....	26
Εικόνα 3. Άποψη προσφύγων για τις δυνατότητες επαγγελματικής αποκατάστασης στην Ελλάδα.	27
Εικόνα 4. Άποψη προσφύγων για χώρες και οργανισμούς.....	28
Εικόνα 5. Η γνώση των δικαιωμάτων των προσφύγων.....	34
Εικόνα 6. Αιτήματα ασύλου (συνολικός αριθμός) στη Γερμανία από τον Ιανουάριο έως τον Απρίλιο 2016.....	57
Εικόνα 7. Η σημασία του να μιλάει κανείς τη γλώσσα της χώρας.....	66
Εικόνα 8. Η σημασία του να γνωρίζει κάποιος τα ήθη και τα έθιμα της χώρας.....	67
Εικόνα 9. Η σημασία της θρησκείας	68
Εικόνα 10. Η σημασία του τύπου γέννησης	69
Εικόνα 11. Κλίμακα εθνικής ταυτότητας	70
Εικόνα 12. Τα συναισθήματα που προκαλεί η προσφυγική κρίση στους Έλληνες.....	78
Εικόνα 13. Οι τομείς στους οποίους θα συμβάλει η υποδοχή των προσφύγων	79
Εικόνα 14. Οι ανησυχίες των Γερμανών σε σχέση με την προσφυγική κρίση	80
Εικόνα 15. Η συμπάθεια των Ελλήνων απέναντι στους πρόσφυγες.....	81
Εικόνα 16. Άποψη Ελλήνων για την προοπτική μόνιμης εγκατάστασης των προσφύγων στην Ελλάδα	82
Εικόνα 17. Άποψη Γερμανών για το δικαίωμα παραμονής των προσφύγων στη Γερμανία.	83
Εικόνα 18. Άποψη Γερμανών για την προοπτική μόνιμης εγκατάστασης των προσφύγων στη Γερμανία	83
Εικόνα 19. Άποψη πολιτών για την υποδοχή των προσφύγων στη χώρα τους.....	84
Εικόνα 20. Άποψη πολιτών για την υποδοχή των προσφύγων στη χώρα/πόλη/γειτονιά/σπίτι τους.	85
Εικόνα 21. Δείκτης υποδοχής προσφύγων.....	87
Εικόνα 22. . Οι χρήστες του Twitter σε παγκόσμιο επίπεδο.	92
Εικόνα 23. . Διαδικασία εξόρυξης δεδομένων.....	99
Εικόνα 24. Τεχνικές κατηγοριοποίησης συναισθήματος.	113
Εικόνα 25. Το μοντέλο του Naive Bayes.	117
Εικόνα 26. Το μοντέλο της αλυσίδας Markov.....	119
Εικόνα 27. Το μοντέλο Support Vector Machines.....	120
Εικόνα 28. Το μοντέλο Maximum Entropy.....	120
Εικόνα 29. Τεκμηρίωση του πρωτοκόλλου OAuth στο Twitter.	140
Εικόνα 30. Κώδικας για τη λήψη tweets μέσω του REST API στο περιβάλλον του RStudio.....	143

ΠΙΝΑΚΑΣ ΣΥΝΤΜΗΣΕΩΝ

ΣΥΝΤΟΜΟΓΡΑΦΙΑ	ΠΛΗΡΕΣ ΟΝΟΜΑ
Βλ.	Βλέπε
Ε.Υ.	Ερευνητική Υπόθεση
Ερ.Ε.	Ερευνητικό Ερώτημα
κ.α.	Και άλλα
Δ.χ	Λόγου χάρη
ΜΜΕ	Μέσα Μαζικής Επικοινωνίας
Ό.π.	Όπως παραπάνω
Π.χ.	Παραδείγματος χάρη

ΕΙΣΑΓΩΓΗ

Μία από τις βασικότερες ανάγκες του ανθρώπου είναι αυτή της αλληλεπίδρασης και της συναναστροφής με άλλους ανθρώπους. Η ραγδαία ανάπτυξη της τεχνολογίας των επικοινωνιών έχει αλλάξει ριζικά τον τρόπο με τον οποίο οι άνθρωποι κοινωνικοποιούνται και συναναστρέφονται. Ο όγκος των δεδομένων που παράγεται τα τελευταία χρόνια στο διαδίκτυο έχει αυξηθεί δραματικά και προκειμένου να αξιοποιηθούν τα εν λόγω δεδομένα και να εξαχθούν χρήσιμα συμπεράσματα από αυτά είναι απαραίτητη η γρήγορη και η αποδοτική τους επεξεργασία. Επομένως, οι τεχνικές ανάλυσης μεγάλου όγκου δεδομένων (Big Data) έχουν γνωρίσει ιδιαίτερη άνθηση. Στόχος της εξόρυξης γνώσης είναι η πληροφορία που θα εξαχθεί και τα πρότυπα που θα προκύψουν να έχουν δομή κατανοητή προς τον άνθρωπο, ώστε να τον βοηθήσουν να πάρει τις κατάλληλες αποφάσεις.

Πρόσφατες έρευνες δείχνουν ότι το μεγαλύτερο μέρος του χρόνου που καταναλώνεται στο διαδίκτυο είναι στην κοινωνική δικτύωση, με πολλές εκατοντάδες ώρες μηνιαίως να δαπανώνται σε ιστοτόπους, όπως το Twitter. Έτσι οι πλατφόρμες κοινωνικής δικτύωσης αποτελούν σημαντικές πηγές πληροφοριών, όσον αφορά στην άποψη και στα συναισθήματα των ανθρώπων, τα οποία καταγράφονται πλέον απευθείας σε ηλεκτρονική μορφή. Η ανάγκη ανάλυσης και αξιοποίησης αυτού του τεράστιου όγκου δεδομένων που παράγεται από τους χρήστες (user-generated content) με αυτοματοποιημένο τρόπο οδήγησε στην εμφάνιση του επιστημονικού πεδίου της Ανάλυσης Συναισθήματος (Sentiment Analysis). Η ανάλυση συναισθήματος σε δεδομένα κοινωνικών δικτύων αποκτά ολοένα και περισσότερο έδαφος στον ακαδημαϊκό χώρο, αλλά και στον επιχειρηματικό, χάρη στις πολλά υποσχόμενες προοπτικές της. Οι κύριες μέθοδοι για την εξαγωγή του συναισθήματος από τα δεδομένα είναι η χρήση λεξικών και η χρήση αλγορίθμων ταξινόμησης, οι οποίοι κατατάσσουν αυτόματα τα κείμενα σε κατηγορίες.

Τα συναισθήματα που κρύβουν τα δεδομένα του διαδικτύου έχουν μεγάλη αξία και γι' αυτό οι εφαρμογές της εξόρυξης και ανάλυσης απόψεων έχουν εξαπλωθεί σε κάθε πιθανό θεματικό τομέα, από τα προϊόντα, τις υπηρεσίες και την οικονομία στις κοινωνικές εκδηλώσεις, τις πολιτικές εκλογές και τα κοινωνικά θέματα, με ένα από τα οποία καταπιάνεται η παρούσα εργασία και συγκεκριμένα με την εκπαίδευση των προσφύγων.

Η ανάλυση συναισθήματος σε θέματα που έχουν να κάνουν με την κοινωνία είναι ίσως ο ερευνητικός τομέας του συγκεκριμένου πεδίου που βρίσκεται σε μεγαλύτερη ανάπτυξη αυτή τη στιγμή. Όσο σημαντικό είναι για μία εταιρεία ή έναν κατασκευαστή να γνωρίζει τη γνώμη και τα σχόλια των καταναλωτών για ένα προϊόν, τόσο και ακόμα πιο σημαντικό είναι για μία κυβέρνηση ή έναν οργανισμό να γνωρίζει την άποψη των πολιτών για κάποιο φλέγον κοινωνικό θέμα, προκειμένου να πάρουν τις καταλληλότερες αποφάσεις στη χάραξη πολιτικής και να αντιμετωπίσουν πιο αποτελεσματικά το εκάστοτε ζήτημα.

Η προσφυγική κρίση των τελευταίων ετών έχει συγκλονίσει ολόκληρη την υφήλιο. Μία από τις πολλές πτυχές της είναι το κομμάτι της εκπαίδευσης των παιδιών προσφύγων, τα οποία απαρτίζουν ένα πολύ μεγάλο μέρος του προσφυγικού πληθυσμού. Το κάθε κράτος αντιμετωπίζει το εν λόγω θέμα σε διαφορετικό βαθμό και με διαφορετικό τρόπο. Άλλες χώρες έχουν μεγάλη εμπειρία στον τομέα της πολιτικής για την ένταξη αυτών των παιδιών στο εκπαιδευτικό σύστημα, ενώ άλλες, που δεν έχουν βιώσει ξανά ένα τέτοιο φαινόμενο, αναζητούν τώρα τις απαραίτητες πολιτικές και εκπαιδευτικές μεταρρυθμίσεις. Στην παρούσα μελέτη θα γίνει προσπάθεια ανίχνευσης του συναισθήματος σε σχέση με την εκπαίδευση των προσφύγων με βάση τα δεδομένα από δύο χώρες, της Ελλάδας και της Γερμανίας που, υπό διαφορετικές συνθήκες και κυρίως με διαφορετικό οικονομικό και κοινωνικό υπόβαθρο, έρχονται αντιμέτωπες με το εν λόγω ζήτημα.

ΘΕΩΡΗΤΙΚΟ ΠΛΑΙΣΙΟ

1. Η ΠΡΟΣΦΥΓΙΚΗ ΚΡΙΣΗ

Ο εμφύλιος πόλεμος στη Συρία, σε συνδυασμό με τις συγκρούσεις σε χώρες της Μέσης Ανατολής και της Αφρικής έχουν αναγκάσει τους πολίτες των χωρών αυτών να ακολουθήσουν το δρόμο της προσφυγιάς. Η προσφυγική-μεταναστευτική κρίση των τελευταίων ετών έχει χαρακτηριστεί από πολλούς ως η χειρότερη στην ιστορία μετά τον Β' παγκόσμιο πόλεμο. Καθημερινά, αναγκάζονται να μετακινηθούν εκατομμύρια άνθρωποι. Μόνο το 2014, ο αριθμός των προσφύγων και των εσωτερικά εκτοπισμένων έφτασε σε παγκόσμια κλίμακα σχεδόν τους 60.000.000 ανθρώπους, από τους οποίους το 51% ήταν παιδιά κάτω των 18 ετών. Όπως χαρακτηριστικά αναφέρεται στην έκθεση της Ύπατης Αρμοστείας του ΟΗΕ για τους Πρόσφυγες (UNHCR), ο διωγμός, η σύγκρουση, η γενικευμένη βία και οι παραβιάσεις των ανθρωπίνων δικαιωμάτων δημιούργησαν το «έθνος των εκτοπισμένων», που αν ήταν χώρα, θα ήταν η 24^η μεγαλύτερη στον κόσμο.¹ Μέχρι και σήμερα οι προσφυγικές ροές συνεχίζονται αμείωτες, καθώς μόλις τους τέσσερις πρώτους μήνες του 2017 έφτασαν στην Ευρώπη μέσω θαλάσσης περίπου στους 45.000 πρόσφυγες.²

1.1 Ο ΠΡΟΣΦΥΓΑΣ ΚΑΙ Ο ΜΕΤΑΝΑΣΤΗΣ

Όπως ορίζει η συνθήκη της Γενεύης, πρόσφυγας είναι *«κάθε πρόσωπο το οποίο, επειδή έχει δικαιολογημένο φόβο διωγμού λόγω φυλής, θρησκείας, εθνικότητας, πολιτικών πεποιθήσεων ή συμμετοχής σε ορισμένη κοινωνική ομάδα βρίσκεται έξω από τη χώρα*

¹ UNHCR, Global Trends 2014, 2015, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.unhcr.org/statistics/country/556725e69/unhcr-global-trends-2014.html>, [προσπελάστηκε στις 24 Φεβρουαρίου 2017], σελ. 2, 3 και 5.

² UNHCR, Refugees/Migrants Emergency Response-Mediterranean, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://data.unhcr.org/mediterranean/regional.php>, [προσπελάστηκε στις 8 Μαρτίου 2017].

της υπηκοότητάς του, και δεν μπορεί, ή εξ' αιτίας αυτού του φόβου, δεν θέλει να προσφύγει στην προστασία της χώρας αυτής».³

Η ιδιότητα του πρόσφυγα αποτελεί μια διαφορετική κατηγορία από αυτή του μετανάστη. Συγκεκριμένα, δύο είναι τα βασικά χαρακτηριστικά του πρόσφυγα που τον διαχωρίζουν από τον μετανάστη:

1) Η μετακίνηση του πρόσφυγα χαρακτηρίζεται ως άμεσο αποτέλεσμα του κινδύνου που διατρέχει η ζωή του, σε περίπτωση παραμονής του στην χώρα προέλευσης.

2) Η παρουσία του πρόσφυγα έξω από τα σύνορα της χώρας του, λόγω παραβίασης των ανθρωπίνων δικαιωμάτων του αν παραμείνει μέσα σε αυτήν, τον θέτει στη νομική προστασία του διεθνούς νομικού συστήματος, σύμφωνα με την Συνθήκη της Γενεύης του 1951.⁴

Όπως χαρακτηριστικά αναφέρει η Ύπατη Αρμοστεία για τους Πρόσφυγες: «Οι μετανάστες, ιδιαίτερα οι οικονομικοί μετανάστες, επιλέγουν να μετακινηθούν προκειμένου να βελτιώσουν τη ζωή τους. Οι πρόσφυγες αναγκάζονται να μετακινηθούν προκειμένου να σώσουν τη ζωή τους ή να διαφυλάξουν την ελευθερία τους».⁵

Η διάκριση μεταξύ προσφύγων και μεταναστών είναι πολύ σημαντική, καθώς οι εξελίξεις στο διεθνές σύστημα τείνουν να περιορίσουν τη σημασία ανάγκης της προστασίας των προσφύγων, καθώς, κατηγοριοποιώντας τους ως μετανάστες, εντάσσουν την αντιμετώπιση των ζητημάτων τους κάτω από το πρίσμα πολιτικών που αφορούν τη μετανάστευση.⁶

³ Ελληνικό Συμβούλιο για τους Πρόσφυγες, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.gcr.gr/index.php/el/about-gcr/what-we-do>, [προσπελάστηκε στις 8 Μαρτίου 2017].

⁴ Hathaway, J.C., *Why Refugee Law Still Matters*, *Melbourne Journal of International Law*, 2007, 89 (8), σελ. 351.

⁵ UNHCR, Προστατεύοντας τους πρόσφυγες, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://www.unhcr.gr/fileadmin/Greece/General/PDF-LOW.pdf>, [προσπελάστηκε στις 9 Μαρτίου 2017], σελ. 7.

⁶ Παρδαλίδης, Θ., Στρατηγικές επιβίωσης και κοινωνικής προσαρμογής των αιτούντων άσυλο και των αναγνωρισμένων προσφύγων στην Ελλάδα. Μία εθνογραφική προσέγγιση, Διδακτορική Διατριβή, Πανεπιστήμιο Μακεδονίας, 2013, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: https://dspace.lib.uom.gr/bitstream/2159/15789/1/PardalidisTheodosios_PhD2013.pdf, [προσπελάστηκε στις 8 Μαρτίου 2017], σελ. 2.

«Συχνά οι όροι που παρουσιάζονται παραπάνω διαπλέκονται μεταξύ τους λόγω των πολυποίκιλων παραγόντων που τους ορίζουν με αποτέλεσμα να δυσχεραίνεται η διαδικασία κατηγοριοποίησής τους, ή να μην είναι επαρκής. Η σημερινή έκφραση της προσφυγικής κρίσης συμπλέκεται και περιπλέκεται με το φαινόμενο της μαζικής οικονομικής μετανάστευσης και γίνεται δυσδιάκριτη από αυτήν.»⁷ Για παράδειγμα, με την κατάρρευση της Συρίας, εκατομμύρια άνθρωποι έφευγαν από τη χώρα σε μεγάλα προσφυγικά και μεταναστευτικά ρεύματα προς την Ευρώπη εδώ και περίπου τρεις δεκαετίες. Οι εν λόγω άνθρωποι φεύγουν από περιοχές πολιτικής και οικονομικής αποσταθεροποίησης όλα αυτά τα χρόνια και γι' αυτό θα μπορούσε να θεωρηθεί ότι το σύνολο των αιτιών που προκαλούν προσφυγικές κρίσεις συμπεριλαμβάνεται, λιγότερο επιτακτικά, στα αίτια της μεταναστευτικής κρίσης. Επομένως, πολλές φορές είναι δύσκολο να πραγματοποιηθεί η διάκριση ανάμεσα σε πρόσφυγα και οικονομικό μετανάστη, καθώς οι αιτίες μετανάστευσής τους πιθανώς να σχετίζονται με την πολιτική ή πολεμική αποσταθεροποίηση και την οικονομική διάλυση συγχρόνως.⁸

Έτσι, ενώ οι πρόσφυγες αντιμετωπίζονται στο διεθνές δίκαιο ως ξεχωριστή κατηγορία από τους μετανάστες, συχνά ταξιδεύουν μαζί χρησιμοποιώντας τους ίδιους μεταναστευτικούς δρόμους, κινδυνεύουν στις ίδιες θάλασσες στην προσπάθειά τους να φτάσουν στην Ευρώπη ή πέφτουν θύματα των ίδιων κυκλωμάτων εμπορίας ανθρώπων, με αποτέλεσμα πολλές φορές να βιώνουν τις ίδιες παραβιάσεις ανθρωπίνων δικαιωμάτων στις χώρες μετάβασης ή στις χώρες προορισμού.

Στα ΜΜΕ αποτελεί πλέον σύνηθες φαινόμενο να αναφέρονται μαζικά στους ανθρώπους που βρίσκονται σε κίνηση, χωρίς να ξεχωρίζουν τις διαφορετικές περιπτώσεις, δημιουργώντας εν προκειμένω μεγαλύτερη σύγχυση στην κοινή γνώμη. Σε αυτό συμβάλλει και η έλλειψη ενημέρωσης του κοινού καθώς και των σχετικών φορέων.⁹ Επίσης, στα μέσα κοινωνικής δικτύωσης, οι δύο αυτοί όροι, «πρόσφυγας» και «μετανάστης», συγχέονται και χρησιμοποιούνται εναλλακτικά. Αυτό συμβαίνει, είτε

⁷ Παναγιωτάκης, Μ., 10 σημεία για τη γεωπολιτική συγκυρία της προσφυγικής κρίσης, Το πιο κρύο καλοκαίρι, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://histologion-gr.blogspot.gr/2016/03/10.html>, [προσπελάστηκε στις 28 Φεβρουαρίου 2017].

⁸ Ο.π.

⁹ Παπαπάνου, Π., Το Προσφυγικό Ζήτημα, Οι ελληνοτουρκικές σχέσεις και ο ρόλος της Ευρωπαϊκής Ένωσης, Μεταπτυχιακή Διπλωματική Εργασία, Πανεπιστήμιο Μακεδονίας, 2017, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://dspace.lib.uom.gr/bitstream/2159/20086/4/PapapanouPanagiotaAikateriniMsc2017.pdf>, [προσπελάστηκε στις 8 Ιανουαρίου 2017], σελ. 14.

επειδή ο συντάκτης αγνοεί τη διαφοροποίηση των δύο όρων, είτε επειδή αμφισβητεί την προσφυγική ιδιότητα των ανθρώπων αυτών, θεωρώντας τους οικονομικούς μετανάστες που θέλουν να αναγνωριστούν ως πρόσφυγες, προκειμένου να επωφεληθούν από τα δικαιώματα που τυγχάνουν οι πρόσφυγες. Γι' αυτό το λόγο συχνά αναφέρονται στους πρόσφυγες και με τον όρο «λαθρομετανάστες».

1.2 ΟΙ ΠΡΟΣΦΥΓΕΣ ΣΤΗΝ ΕΥΡΩΠΗ

Η Ευρώπη, σε παγκόσμια κλίμακα, φιλοξενεί έναν σχετικά μικρό αριθμό προσφύγων, καθώς το μεγαλύτερο μέρος φιλοξενείται σε αναπτυσσόμενες χώρες, κοντά στις πληγείσες περιοχές, όπως είναι η Τουρκία, το Πακιστάν και ο Λίβανος.¹⁰

Παρόλαυτά, οι εισροές προς την Ευρώπη μέσω της Μεσογείου, οι οποίες ξεκίνησαν ουσιαστικά το 2010, αυξάνονται σε μεγάλο ποσοστό χρόνο με το χρόνο. Ενδεικτικά, από το 2014 έφτασαν στην Ευρώπη κατά 83% περισσότεροι άνθρωποι από ότι το 2013. Το 2014 διέσχισαν τη Μεσόγειο θάλασσα 219.000 πρόσφυγες αναζητώντας επιβίωση και ασφάλεια, ενώ μόλις το πρώτο εξάμηνο του 2015 ο αριθμός αυτός άγγιξε τις 137.000.¹¹

Η πλειοψηφία των προσφύγων, με βάση τις καταγραφές που πραγματοποιούνται κατά την άφιξή τους στην Ευρώπη, προέρχεται από χώρες στις οποίες η πολιτική κατάσταση και οι διαρκείς παραβιάσεις των ανθρωπίνων δικαιωμάτων έχουν αναγκάσει έναν πολύ μεγάλο αριθμό πολιτών να εκτοπιστεί. Συγκεκριμένα, το 2016, από τους 362.376 πρόσφυγες που έφτασαν στην Ευρώπη μέσω θαλάσσης, οι περισσότεροι προέρχονται από τη Συρία (21,2%), το Αφγανιστάν (10,6%) και το Ιράκ (10,3%) . Αντίστοιχα, οι περισσότερες αιτήσεις χορήγησης ασύλου στην Ευρώπη το 2015 υποβλήθηκαν από πολίτες της Συρίας, του Αφγανιστάν και του Ιράν.¹² Τόσο η κλιμάκωση της βίας στις χώρες αυτές, όσο η πείνα, η άρνηση των κοινωνικών δικαιωμάτων των πολιτών, αλλά

¹⁰ UNHCR, Global Trends 2014, ό.π. παραπομπή 1, σελ. 3 & 5.

¹¹ UNHCR, The sea route to Europe: The Mediterranean passage in the age of refugees, 2015, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.unhcr.org/5592bd059.html>, [προσπελάστηκε στις 3 Μαρτίου 2017], σελ. 2, 5 & 6.

¹² Eurostat, Asylum Statistics, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: http://ec.europa.eu/eurostat/statistics-explained/index.php/Asylum_statistics, [προσπελάστηκε στις 6 Μαρτίου 2017].

και η πολιτική διαφθορά έχουν γίνει η αιτία για την κλιμάκωση των προσφυγικών ροών προς την Ευρώπη.¹³

Συγκεκριμένα, στη Συρία, με τον εμφύλιο πόλεμο που ξέσπασε έξι χρόνια πριν, οι παραβιάσεις των ανθρωπίνων δικαιωμάτων είναι διαρκείς και κατάφωρες. Τα εγκλήματα πολέμου, οι επιθέσεις κατά των αμάχων, οι εκτελέσεις, οι απαγωγές, τα βασανιστήρια και οι βιασμοί κατά των γυναικών, έχουν οδηγήσει πολλά εκατομμύρια Σύριους να μετακινηθούν τόσο σε άλλες χώρες, όσο και μέσα στο ίδιο το κράτος.¹⁴

Στο Αφγανιστάν επικρατεί έντονη πολιτική αστάθεια εξαιτίας των συγκρούσεων των κυβερνητικών δυνάμεων με αντικυβερνητικές ένοπλες δυνάμεις. Πόλεις που έχουν καταληφθεί από τους Ταλιμπάν βιώνουν καθημερινά εκτελέσεις, επιθέσεις κατά των πολιτών, απαγωγές και βασανιστήρια. Τα βασικότερα ανθρώπινα δικαιώματα των πολιτών καταπατώνται διαρκώς, ιδίως αυτά των γυναικών, σε βάρος των οποίων ασκείται κάθε μορφή βίας.¹⁵

Επιπλέον, στο Ιράκ και στη Συρία, που βρίσκονται σε εμφύλιο πόλεμο από το 2014, η εξτρεμιστική οργάνωση ISIS, σύμφωνα με τα Ηνωμένα Έθνη, παραβιάζει το διεθνές δίκαιο και τα ανθρώπινα δικαιώματα με συστηματικές επιθέσεις εναντίον άμαχου πληθυσμού, με απαγωγές, βασανιστήρια και εκτελέσεις ατόμων που ανήκουν κυρίως σε θρησκευτικές μειονότητες, ενώ ιδιαιτέρως σκληρής μεταχείρισης τυγχάνουν οι γυναίκες και τα παιδιά.¹⁶

¹³ UNHCR, The sea route to Europe: The Mediterranean passage in the age of refugees, ό.π. παραπομπή 11.

¹⁴ OHCHR, Report of the Independent International Commission of Inquiry on the Syrian Arab Republic. A/HRC/30/48, 2015, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: http://www.ohchr.org/EN/HRBodies/HRC/RegularSessions/Session30/Documents/A.HRC.30.48_AEV.pdf, [προσπελάστηκε στις 9 Μαρτίου 2017].

¹⁵ UNAMA/OHCHR, Afghanistan, Human rights and protection of civilians in armed conflict, Special report on Kunduz province, 2015, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: https://unama.unmissions.org/sites/default/files/special_report_on_kunduz_province_12_december_2015.pdf, [προσπελάστηκε στις 23 Μαρτίου 2017].

¹⁶ UN/ Meetings coverage and press releases, Security Council 'Unequivocally' Condemns ISIL Terrorist Attacks, Unanimously Adopting Text that Determines Extremist Group Poses 'Unprecedented' Threat, 2015, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.un.org/press/en/2015/sc12132.doc.htm>, [προσπελάστηκε στις 25 Φεβρουαρίου 2017].

Τη Συρία, το Αφγανιστάν και το Ιράκ ακολουθούν χώρες της Ασίας και της Αφρικής, όπως η Νιγηρία, η Ερυθραία, το Πακιστάν, η Ακτή του Ελεφαντοστού, η Γκάμπια, Η Γουινέα και το Σουδάν.¹⁷

Η πορεία των προσφύγων προς την Ευρώπη γίνεται είτε μέσω ξηράς είτε μέσω της Μεσογείου θάλασσας, με τη δεύτερη να είναι δημοφιλέστερη και να περιλαμβάνει τρεις βασικούς άξονες. Συγκεκριμένα, η πρώτη διαδρομή πραγματοποιείται μέσω της κεντρικής Μεσογείου με κατεύθυνση προς τη Μάλτα και την Ιταλία. Αυτή ήταν και η προτιμώμενη διαδρομή μέχρι το 2015.¹⁸ Η δεύτερη διαδρομή γίνεται μέσω της δυτικής Μεσογείου και κατευθύνεται προς την Ισπανία. Η εν λόγω διαδρομή ήταν και παραμένει η λιγότερο δημοφιλής από τις τρεις. Τέλος, υπάρχει η διαδρομή μέσω της ανατολικής Μεσογείου με κατεύθυνση τη Βουλγαρία, την Κύπρο και την Ελλάδα, η οποία, έχοντας ως προορισμό τα νησιά του Αιγαίου, αποτελεί πλέον σχεδόν τη μόνη χρησιμοποιούμενη από τις παραπάνω διαδρομές.¹⁹

Μέχρι το 2015, οι περισσότερες αφίξεις στη Μεσόγειο θάλασσα καταγράφονταν στην Ιταλία. Το 2014 έφτασαν στην Ιταλία 170.000 πρόσφυγες, ενώ στην Ελλάδα μόλις 43.500, λιγότερο δηλαδή από το 1/5 του συνολικού πληθυσμού που έφτασε στην Ευρώπη μέσω της Μεσογείου. Το 2015 όμως η εικόνα αυτή άλλαξε, καθώς το πρώτο εξάμηνο του έτους έφτασαν στην Ιταλία 67.500 άνθρωποι και στην Ελλάδα αντίστοιχα 68.000, περισσότεροι από ό,τι όλη την προηγούμενη χρονιά.²⁰

¹⁷ UNHCR, Refugees/Migrants Emergency Response-Mediterranean, ό.π. παραπομπή 2.

¹⁸ UNHCR, The sea route to Europe: The Mediterranean passage in the age of refugees, ό.π. παραπομπή 11, σελ. 11.

¹⁹ Frontex, Migratory routes map, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://frontex.europa.eu/trends-and-routes/migratory-routes-map/>, [προσπελάστηκε στις 3 Ιουλίου 2017].

²⁰ UNHCR, The sea route to Europe: The Mediterranean passage in the age of refugees, ό.π. παραπομπή 11, σελ. 11.



Εικόνα 1. Οι διαδρομές των προσφύγων²¹

Το ταξίδι των προσφύγων όμως δεν τελειώνει με την άφιξή τους στην Ελλάδα, στην Ιταλία ή στην Ισπανία, καθώς ο τελικός τους προορισμός βρίσκεται στα βορειοδυτικά τμήματα της Ευρώπης. Η πλειονότητα των προσφύγων δεν επιθυμεί να διαμείνει στη χώρα άφιξης. Σε έρευνα που διενεργήθηκε με τους Σύριους πρόσφυγες στην Ελλάδα, το 90% ήθελε να αιτηθεί άσυλο σε κάποια άλλη χώρα της Ευρωπαϊκής Ένωσης και μάλιστα κατά κύριο λόγο στη Γερμανία και στη Σουηδία, προκειμένου να έχουν καλύτερη προστασία και περισσότερες ευκαιρίες εύρεσης εργασίας.²²

Από έρευνα που πραγματοποίησε η Κάπα Research για την Ένωση Περιφερειών Ελλάδος (ΕΝΠΕ) σε πρόσφυγες σε κέντρα υποδοχής/παραμονής μεταναστών στο Λεκανοπέδιο Αττικής τον Μάρτιο του 2016²³ προκύπτει ότι η συντριπτική πλειοψηφία των προσφύγων δεν επιθυμεί να παραμείνει στην Ελλάδα, αλλά να συνεχίσει το ταξίδι

²¹ Businessinsider.com, This map shows the routes of Europe's refugee nightmare — and how it's getting worse, 2015, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.businessinsider.com/map-of-europe-refugee-crisis-2015-9>, [προσπελάστηκε στις 15 Ιουλίου 2017].

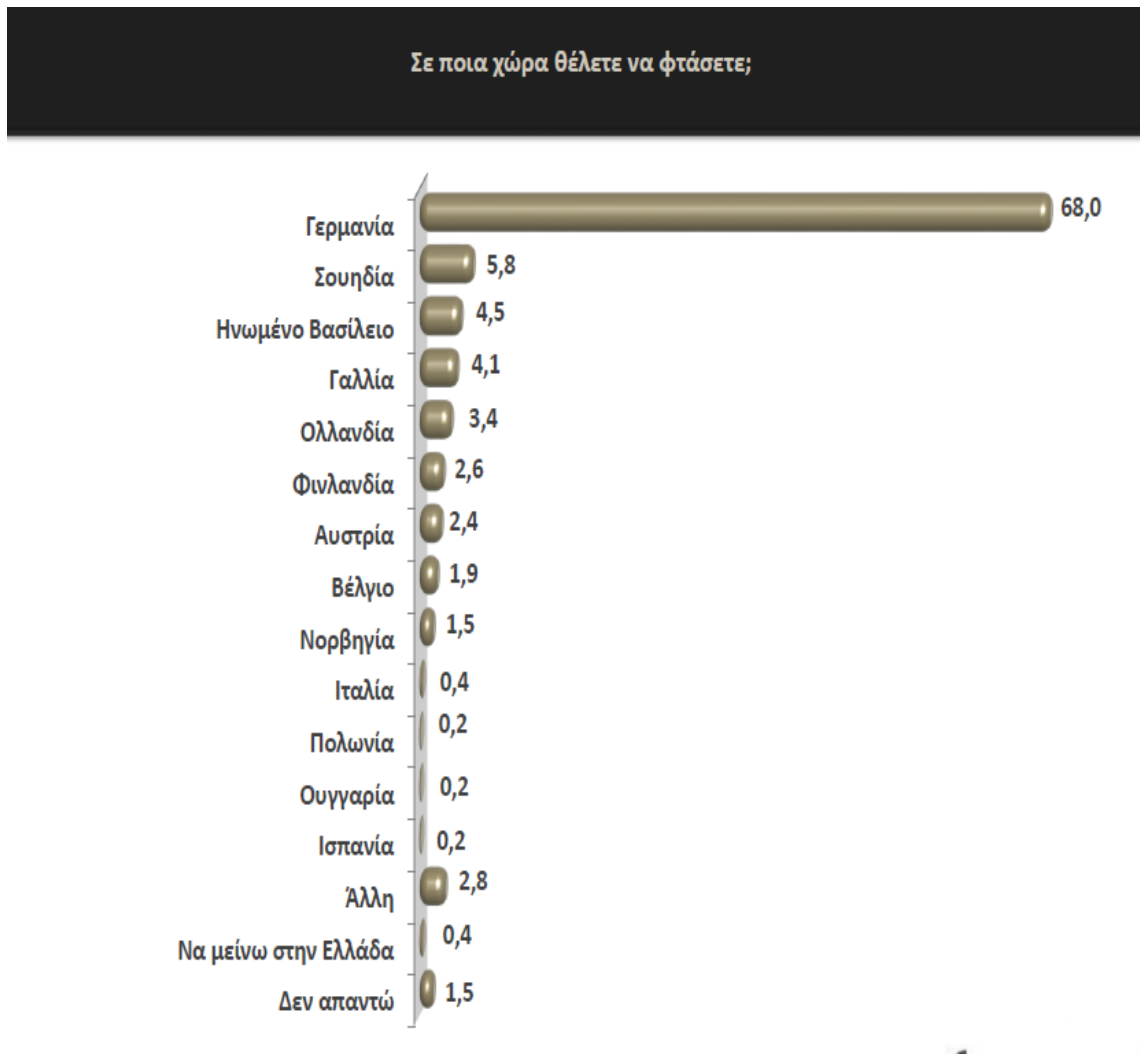
²² UNHCR, The sea route to Europe: The Mediterranean passage in the age of refugees, ό.π. παραπομπή 11, σελ. 11.

²³ Κάπα research, Έρευνα σε πρόσφυγες, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: http://www.enpe.gr/media/368840/2016_03_16%20CE%88%CF%81%CE%B5%CF%85%CE%BD%CE%B1%20%CF%83%CE%B5%20%CF%80%CF%81%CF%8C%CF%83%CF%86%CF%85%CE%B3%CE%B5%CF%82%20%CE%95%CE%9D%CE%A0%CE%95.pdf, [προσπελάστηκε στις 14 Φεβρουαρίου 2017].

της στην Ευρώπη και να εγκατασταθεί σε κάποια άλλη χώρα (εικόνα 2). Συγκεκριμένα, μόλις το 0,4% των ερωτωμένων θέλει να μείνει στην Ελλάδα, ενώ η Γερμανία αποτελεί την πρώτη προτίμηση των προσφύγων με ποσοστό 68%, γεγονός που αποδεικνύει ότι η Ελλάδα δεν αποτελεί χώρα προορισμού. Πράγματι, το 2014, η Γερμανία και η Σουηδία έλαβαν σχεδόν τις μισές από τις 625.000 αιτήσεις για άσυλο που υποβλήθηκαν συνολικά στην Ευρωπαϊκή Ένωση.²⁴

Παρά το γεγονός όμως ότι στο σύνολό τους οι πρόσφυγες δεν επιθυμούν να εγκατασταθούν στην Ελλάδα, υπό ορισμένες συνθήκες, σχεδόν οι 3 στους 10 θα επέλεγαν να μείνουν: το 27,9% θα έμενε αν έβρισκε δουλειά και το 30,1% θα κάνει αίτηση ασύλου στη χώρα μας αν δεν καταφέρει να περάσει τα σύνορα. Το 31,8% θα έμενε στην Ελλάδα γιατί θα είναι πιο εύκολο από άλλες χώρες, το 25,8% γιατί υπάρχει ήδη κοινότητα από τη χώρα του εδώ ή έχει συγγενείς και το 14% επειδή έχει καλό σύστημα κοινωνικής προστασίας. Το 24,1% - εάν έμενε τελικά στην Ελλάδα - θα επεδίωκε την επανένωση της οικογένειάς του προσκαλώντας την εδώ.

²⁴ FRA, Fundamental rights: Challenges and achievements in 2014, Annual Report 2014, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: http://fra.europa.eu/sites/default/files/fra-annual-report-2014_en.pdf, σελ. 85.



Εικόνα 2. Χώρες προορισμού των προσφύγων²⁵

Επιπλέον το 68% φαίνεται πως γνωρίζοντας την οικονομική κρίση που βιώνει η χώρα, δηλώνει πως δεν πιστεύει ότι μπορεί να βρει εργασία εδώ, ώστε να παραμείνει (εικόνα 3).

²⁵ Ο.π. παραπομπή 23, σελ. 34.

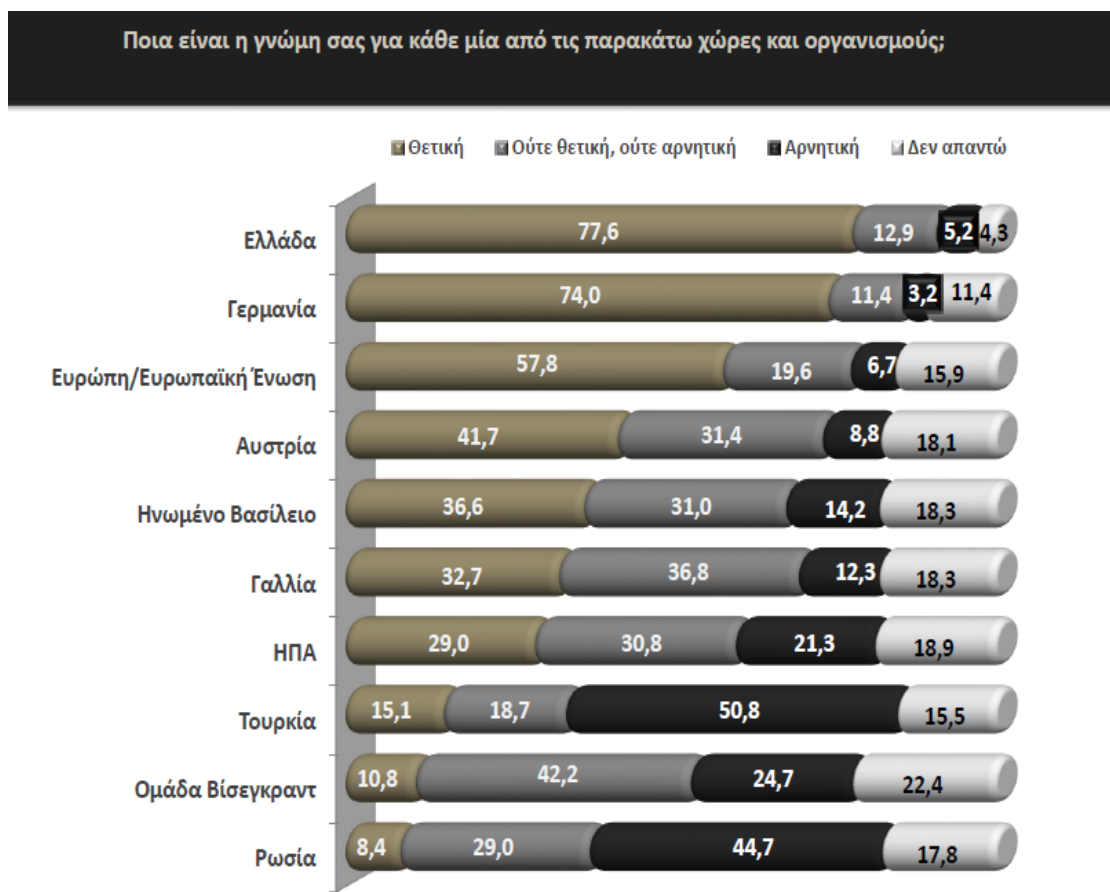
Νιώθετε ότι μπορείτε να βρείτε κάποια δουλειά στην Ελλάδα ώστε να έχετε τα βασικά για να ζήσετε;



Εικόνα 3. Άποψη προσφύγων για τις δυνατότητες επαγγελματικής αποκατάστασης στην Ελλάδα²⁶

Επίσης, από τη συγκεκριμένη έρευνα προκύπτει ότι η Ελλάδα (77,6%) και η Γερμανία (74%) συγκεντρώνουν τις περισσότερες θετικές γνώμες από την πλευρά των προσφύγων ανάμεσα στις χώρες που διαδραματίζουν ρόλο στο προσφυγικό ζήτημα, σε αντίθεση με την Τουρκία και τη Ρωσία που αντιμετωπίζονται αρνητικά με τις αρνητικές γνώμες να αγγίζουν το 50,8% και 44,7% και τις θετικές μόλις το 15,1% και 8,4% αντίστοιχα (εικόνα 4).

²⁶ Ο.π. παραπομπή 23, σελ. 54.

Εικόνα 4. Αποψη προσφύγων για χώρες και οργανισμούς²⁷

1.3 ΤΟ ΠΡΟΦΙΛ ΤΩΝ ΠΡΟΣΦΥΓΩΝ

Η πλειοψηφία των ανθρώπων αυτών είναι άντρες, οι γυναίκες αποτελούν το 18%, ενώ ιδιαιτέρως υψηλό είναι το ποσοστό των παιδιών, καθώς αγγίζει το 31%.²⁸ Συγκεκριμένα, το 2015 έφτασαν στην Ευρώπη 314.873 παιδιά²⁹, εκ των οποίων τα 26.000 ήταν ασυνόδευτα.³⁰ Το 2016, 398.065 παιδιά αιτήθηκαν άσυλο σε χώρες της

²⁷ Ό.π. παραπομπή 23, σελ. 37.

²⁸ UNHCR, Refugees/Migrants Emergency Response-Mediterranean, ό.π. παραπομπή 2.

²⁹ European Commission, Regional refugee and migrant response plan for Europe Eastern Mediterranean and Western Balkans route, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: http://ec.europa.eu/justice/fundamental-rights/files/rights_child/data_children_in_migration.pdf, [προσπελάστηκε στις 16 Φεβρουαρίου 2017], σελ. 21.

³⁰ Save the Children.org, Response to the IDC Report on Syria, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.savethechildren.org.uk/2016-01/response-idc-report-syria>, [προσπελάστηκε στις 25 Φεβρουαρίου 2017].

Ευρωπαϊκής Ένωσης, εκ των οποίων τα 62.930 ήταν ασυνόδευτα.³¹ Στις 24.135 ανέρχονται και τα ασυνόδευτα παιδιά που έφτασαν σε Ελλάδα και Ιταλία τους πρώτους δέκα μήνες του 2016, ποσοστό που ανέρχεται στο 28% των παιδιών που καταγράφηκαν στις δύο χώρες (82.741 παιδιά).³²

Από τις αρχές του 2017 και μέχρι τις 12 Απριλίου του ίδιου έτους, 32.750 πρόσφυγες έφτασαν στην Ευρώπη μέσω θαλάσσης, εκ των οποίων το 26% ήταν παιδιά, το 17% γυναίκες και το 58% άντρες. Συγκεκριμένα, στην Ελλάδα, το ποσοστό των παιδιών που κατέφθασε το ίδιο χρονικό διάστημα στη χώρα ανέρχεται στο 30%, ενώ στην Ιταλία ήταν μόλις το 13%.

1.4 ΑΝΗΛΙΚΟΙ ΠΡΟΣΦΥΓΕΣ

Οι ανήλικοι αντιπροσωπεύουν ένα αυξανόμενο ποσοστό των μεταναστών και των προσφύγων, όπως αναφέρει η UNICEF σε έκθεσή της³³, καθώς αναγκάζονται να εγκαταλείψουν την πατρίδα τους για να αποφύγουν και επιπλέον κινδύνους, όπως για παράδειγμα αυτόν της σεξουαλικής εκμετάλλευσης.³⁴

Περίπου 4.700 παιδιά αγνοούνται στη Γερμανία, 4.000 παιδιά στην Ιταλία, 1.000 στη Σουηδία, ενώ στην Ελλάδα αγνοούνται 1.096, από τα οποία τα 16 είναι κορίτσια. Η πλειοψηφία των ανηλίκων που έχουν χαθεί τα ίχνη τους είναι κυρίως από 15 έως 18 ετών.³⁵

Από τους περίπου 206.200 ανθρώπους που έφθασαν το 2016 στην Ευρώπη διά θαλάσσης έως τις 4 Ιουνίου, ένας στους τρεις ήταν παιδί, αναφέρει η UNICEF,

³¹ European Commission, Regional refugee and migrant response plan for Europe Eastern Mediterranean and Western Balkans route, ό.π. παραπομπή 29, σελ. 21.

³² European Commission, Regional refugee and migrant response plan for Europe Eastern Mediterranean and Western Balkans route, ό.π. παραπομπή 29, σελ. 50.

³³ UNICEF, Danger Every Step of the Way, 2016, [διαδίκτυο], Διαθέσιμο στη διεύθυνση: https://www.unicef.org/emergencies/childrenonthemove/files/Child_Alert_Final_PDF.pdf, [προσπελάστηκε στις 25 Ιανουαρίου 2017], σελ. 2.

³⁴ IOM, Children on the Move, 2013, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: http://publications.iom.int/system/files/pdf/children_on_the_move_15may.pdf, [προσπελάστηκε στις 6 Μαρτίου 2017], σελ.32.

³⁵ Στοιχεία σοκ: Έντεκα χιλιάδες ασυνόδευτα προσφυγόπουλα έχουν εξαφανιστεί, *Αυγή*, 2016, 6 Ιουνίου, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.avgi.gr/article/6594752/stoixeia-sok-enteka-xiliades-asunodeuta-prosfugoroula-exoun-exafanistei>, [προσπελάστηκε στις 6 Φεβρουαρίου 2017].

παραθέτοντας στοιχεία της υπηρεσίας του ΟΗΕ για τους πρόσφυγες. «Κάθε βήμα του ταξιδιού είναι γεμάτο κινδύνους, πόσο μάλλον για σχεδόν ένα στα τέσσερα παιδιά που ταξιδεύουν χωρίς έναν γονέα ή κηδεμόνα». Το ποσοστό είναι πολύ υψηλότερο στα σκάφη που ξεκινούν από τη Λιβύη, όπου περισσότερα από εννέα στα δέκα παιδιά είναι ασυνόδευτα. Η UNICEF ανέφερε ότι υπάρχουν ισχυρές αποδείξεις ότι εγκληματικά δίκτυα διακίνησης ανθρώπων θέτουν στο στόχαστρο τους πιο ευάλωτους, ιδιαίτερα γυναίκες και παιδιά³⁶. Συχνά, επομένως, γίνονται θύματα παράνομης διακίνησης ανθρώπων (trafficking), αλλά αρκετές φορές απελευθώνονται χωρίς να χαρακτηριστούν ως θύματα.³⁷

Σε αναφορά της η Unicef τον Μάρτιο του 2017 εκτιμά τον αριθμό των παιδιών προσφύγων στην Ελλάδα, σε 20.300³⁸. Στη Γερμανία, μόνο το 2016, σύμφωνα πάλι με τη Unicef, κατατέθηκαν αιτήσεις ασύλου για 261.383 ανηλίκους.³⁹

1.5 ΑΣΥΝΟΔΕΥΤΟΙ ΑΝΗΛΙΚΟΙ ΠΡΟΣΦΥΓΕΣ

Σύμφωνα με την Ύπατη Αρμοστεία ασυνόδευτα είναι τα παιδιά που έχουν αποχωριστεί τους γονείς τους και δεν ανατρέφονται από κάποιον ενήλικα που αναλαμβάνει την ευθύνη τους⁴⁰. Οι όροι «ασυνόδευτοι ανήλικοι» ή «ασυνόδευτα παιδιά» προτιμώνται

³⁶ Παιδιά που πραγματοποιούν το επικίνδυνο ταξίδι των μεταναστών και των προσφύγων προς την Ευρώπη για να ξεφύγουν από τον πόλεμο και τη φτώχεια αντιμετωπίζουν το ενδεχόμενο ξυλοδαρμού, βιασμών και αναγκαστικής εργασίας εκτός από τον κίνδυνο να πνιγούν στη Μεσόγειο, όπως ανακοίνωσε το Ταμείο των Ηνωμένων Εθνών για τα Παιδιά (UNICEF, ό.π. παραπομπή 33).

³⁷ Τσιγκρής, Α., Άλλο παράνομοι μετανάστες και... άλλο πρόσφυγες και μάλιστα ανήλικοι που ζητούν άσυλο, *Το Βήμα*, 2015, 21 Απριλίου, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.tovima.gr/opinions/article/?aid=696478>, [προσπελάστηκε στις 6 Μαρτίου 2017].

³⁸ Unicef, Refugee and Migrant Children in Greece - by Region, 2017, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: https://www.unicef.org/eca/Map_RefugeeMigrantChildren_25MAR2017.pdf, [προσπελάστηκε στις 4 Αυγούστου 2017].

³⁹ Unicef, Kindheit im Wartezustand, 2017, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://www.unicef.de/blob/137704/053ab16048c3f443736c4047694cc5d1/studie--kindheit-im-wartezustand-data.pdf>, [προσπελάστηκε στις 4 Αυγούστου 2017].

⁴⁰ UNHCR, Κατευθυντήριες οδηγίες για τα ασυνόδευτα παιδιά που ζητούν άσυλο, 2005, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: https://www.unhcr.gr/no_cache/prostasia/paidia.html?cid=801&did=490&sechash=1fbe2e71, [προσπελάστηκε στις 6 Φεβρουαρίου 2017].

αντί του «ορφανά», καθώς ένα παιδί θα πρέπει να χαρακτηρίζεται ως ορφανό εάν έχουν αποβιώσει και οι δύο γονείς του.⁴¹

Τα ασυνόδευτα ανήλικα άτομα αποτελούν ιδιαίτερα ευάλωτη κατηγορία μεταναστών. Είναι υπήκοοι τρίτων χωρών ή ανιθαγενείς, οι οποίοι εισέρχονται ή βρίσκονται στην χώρα υποδοχής, χωρίς να συνοδεύονται από τους κατά νόμο ή έθιμο υπευθύνους για την επιμέλειά τους. Πολλές φορές προέρχονται από συνθήκες πολέμου και εμφύλιες αναταραχές, έχουν χάσει τις οικογένειές τους και τα σπίτια τους και έχουν υποστεί βασανιστήρια και κακομεταχείριση. Φθάνουν μόνα τους στις χώρες υποδοχής⁴², όπου πολλές φορές είναι δύσκολη η διασφάλιση των δικαιωμάτων τους, τα οποία οφείλουν να τους αναγνωρίσουν οι αρχές της εκάστοτε χώρας.⁴³

Γενικότερα, τα ασυνόδευτα ανήλικα άτομα χρήζουν ιδιαίτερης προστασίας και αρωγής, είτε πρόκειται για παιδιά-πρόσφυγες είτε όχι, ενώ στην Ελλάδα έχουν το δικαίωμα να υποβάλουν και από μόνα τους αίτηση για άσυλο, αν είναι άνω των 14 ετών. Σε διαφορετική περίπτωση η αίτηση γίνεται υποχρεωτικά δια του εκπροσώπου τους. Στη Γερμανία η αίτηση μπορεί να γίνει μόνο δια εκπροσώπου αν είναι κάτω των 18 ετών.⁴⁴ Επιπλέον, η απομάκρυνση του ανηλίκου από τη χώρα είναι νόμιμη μόνο εφόσον εξυπηρετεί το βέλτιστο συμφέρον του παιδιού και εφόσον έχει διακριβωθεί ότι ο ανήλικος θα επιστραφεί στην οικογένειά του ή σε κατάλληλες εγκαταστάσεις υποδοχής στο κράτος επιστροφής.

⁴¹ Σπανού Φ. & Αλεξάκης Β., Ασυνόδευτοι ανήλικοι πρόσφυγες στην Ελλάδα, Πτυχιακή εργασία, Τ.Ε.Ι. Ηρακλείου, 2007, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://nefeli.lib.teicrete.gr/browse/seyp/ker/2007/SpanouFoteini,AlexakisVasilis/attached-document/spanou.pdf>, [προσπελάστηκε στις 16 Ιανουαρίου 2017], σελ. 49.

⁴² Τσιγκρής, Α., Τα προβλήματα στην παροχή ασύλου σε ασυνόδευτα ανήλικα άτομα και η εκμετάλλευση παιδιών, *Το Βήμα*, 2013, 24 Οκτωβρίου, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.tovima.gr/opinions/article/?aid=536277>, [προσπελάστηκε στις 16 Μαρτίου 2017].

⁴³ UNICEF, Convention on the Rights of the Child, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.unicef.org/crc/>, [προσπελάστηκε στις 24 Ιανουαρίου 2017].

⁴⁴ Bundesamt für Migration und Flüchtlinge, Unbegleitete Minderjährige, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.bamf.de/DE/Fluechtlingsschutz/UnbegleiteteMinderjaehrige/unbegleitete-minderjaehrige-node.html>, [προσπελάστηκε στις 14 Φεβρουαρίου 2017].

Συνήθως, το προφίλ των ασυνόδευτων ανηλίκων μπορεί να διακριθεί στις εξής κατηγορίες:⁴⁵

- α) ασυνόδευτοι ανήλικοι που ο γονέας τους βρίσκεται σε κάποια χώρα εντός ΕΕ,
- β) ασυνόδευτοι ανήλικοι που ο γονέας τους βρίσκεται σε κάποιο χώρο παραμονής προσφύγων εκτός ΕΕ,
- γ) ασυνόδευτοι ανήλικοι που ο γονέας τους βρίσκεται σε εμπόλεμη ζώνη στη Συρία και αλλού,
- δ) ασυνόδευτοι ανήλικοι που οι γονείς τους δεν ζουν πια.

Το 2014 και το 2015, συνολικά 4.638 ασυνόδευτοι ανήλικοι, 2.390 και 2.248 αντίστοιχα, τοποθετήθηκαν σε δομές φιλοξενίας στην Ελλάδα, οι περισσότεροι εκ των οποίων αποχώρησαν μετακινούμενοι προς άλλες ευρωπαϊκές χώρες, κάτι πλέον έχει καταστεί ιδιαίτερος δύσκολο. Η έλευση των παιδιών αυτών συνεχίζεται με αμείωτο ρυθμό, καθώς μόνο στους τρεις πρώτους μήνες του 2016 καταγράφηκαν 1.150 ασυνόδευτα παιδιά.⁴⁶ Στη Γερμανία, το 2015, έφτασαν στη χώρα 11.936 ασυνόδευτοι ανήλικοι πρόσφυγες, ενώ μέχρι τον Αύγουστο του 2016 καταγράφηκαν άλλοι 12.981.⁴⁷

Μετά τη Συμφωνία Ευρωπαϊκής Ένωσης και Τουρκίας πραγματοποιήθηκε παρέμβαση⁴⁸ του Συνηγόρου του Πολίτη για τα ασυνόδευτα παιδιά, ζητώντας να διασφαλιστούν τα κατοχυρωμένα δικαιώματα των ασυνόδευτων ανηλίκων, μεταναστών και προσφύγων, να εξασφαλιστεί η αδιάλειπτη λειτουργία και ενίσχυση των εξειδικευμένων δομών φιλοξενίας τους και να προβλεφθεί η άμεση

⁴⁵ Νοτοπούλου, Κ., Μεταναστευτική Κρίση στην Ε.Ε., Μεταπτυχιακή Διπλωματική εργασία, Πανεπιστήμιο Μακεδονίας, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://dspace.lib.uom.gr/bitstream/2159/19621/6/NotopoulouAikateriniMsc2016.pdf>, [προσπελάστηκε στις 24 Ιανουαρίου 2017], σελ. 46.

⁴⁶ Συνήγορος του Πολίτη, Παρέμβαση Συνηγόρου του Πολίτη για τα ασυνόδευτα παιδιά, πρόσφυγες και μετανάστες, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://www.synigoros.gr/resources/dt-asynodeytoi-anilikoi-3032016.pdf>, [προσπελάστηκε στις 14 Φεβρουαρίου 2017].

⁴⁷ Diakonie Deutschland, Unbegleitete minderjährige Flüchtlinge, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://info.diakonie.de/infothek/wissen-kompakt/detail/unbegleitete-minderjaehrige-fluechtlinge/>, [προσπελάστηκε στις 4 Μαρτίου 2017].

⁴⁸ Επιστολή Συνηγόρου του Πολίτη στους αρμόδιους υπουργούς δηλαδή τους Αναπληρωτές Υπουργούς Μεταναστευτικής Πολιτικής, Προστασίας του Πολίτη και Κοινωνικής Αλληλεγγύης, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://www.synigoros.gr/resources/160330-epistoli-asynodeytoi-anilikoi.pdf>, [προσπελάστηκε στις 4 Ιουνίου 2017].

μετεγκατάσταση σε χώρες της Ευρωπαϊκής Ένωσης μεγάλου αριθμού ασυνόδευτων ανηλίκων.

Ιδιαίτερως σημαντική είναι η επίτευξη της ομαλής ενσωμάτωσης των παιδιών αυτών στη χώρα ασύλου, καθώς θα πρέπει να έχουν πρόσβαση σε ό,τι έχουν και τα παιδιά της εκάστοτε χώρας. Θα πρέπει να δοθεί ιδιαίτερη βαρύτητα στο να εξασφαλιστεί για τα παιδιά αυτά η πρόσβαση σε κατάλληλη ισορροπημένη διατροφή και σε υπηρεσίες υγείας, στην εκπαίδευση, συμπεριλαμβανομένης της εκπαίδευσης στην γλώσσα και στον πολιτισμό της χώρας ενσωμάτωσης, σε εκπαιδευτικά επιδόματα ή υποτροφίες, καθώς και στην επαγγελματική εκπαίδευση σε εφήβους.⁴⁹

Σημαντικό ρόλο για τα ασυνόδευτα παιδιά και την ενσωμάτωσή τους παίζει και ο κηδεμόνας. Στην Ελλάδα η διαδικασία καθορισμού του κηδεμόνα καθυστερεί τη μεταφορά του παιδιού στις κατάλληλες εγκαταστάσεις για την προστασία του, ενώ στη Γερμανία, όπου ο κάθε κηδεμόνας μπορεί να είναι υπεύθυνος για μέχρι και 150 παιδιά, σημειώνονται αρκετά μεγάλες καθυστερήσεις στο διορισμό τους που φτάνουν και τους οκτώ μήνες, γεγονός που προκαλεί προβλήματα στην ένταξη του παιδιού στο σχολικό περιβάλλον.⁵⁰

1.6 ΤΑ ΔΙΚΑΙΩΜΑΤΑ ΤΩΝ ΠΡΟΣΦΥΓΩΝ

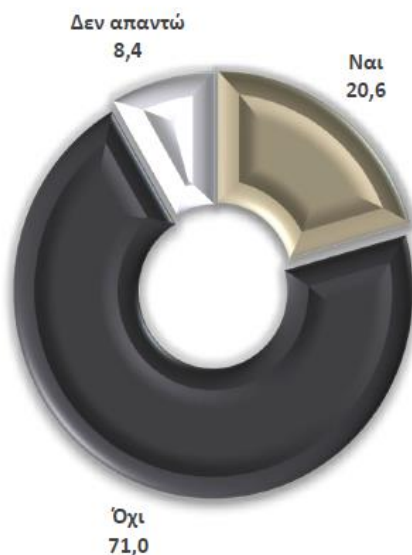
Από την έρευνα της ΚΑΠΑ RESEARCH⁵¹ προκύπτει ότι η συντριπτική πλειοψηφία των προσφύγων και συγκεκριμένα το 71%, δεν γνωρίζει τα δικαιώματά του ως πρόσφυγας, ενώ μόλις το 20,6% φαίνεται να έχει γνώση των δικαιωμάτων του (εικόνα 5).

⁴⁹ Σπανού Φ. & Αλεξάκης Β., Ασυνόδευτοι ανήλικοι πρόσφυγες στην Ελλάδα, ό.π. παραπομπή 41, σελ. 56.

⁵⁰ European Commission, Regional refugee and migrant response plan for Europe Eastern Mediterranean and Western Balkans route, ό.π. παραπομπή 29, σελ 46.

⁵¹ Κάπα research, Έρευνα σε πρόσφυγες, ό.π. παραπομπή 23.

Γνωρίζετε τα δικαιώματά σας ως πρόσφυγας;



Εικόνα 5. Η γνώση των δικαιωμάτων των προσφύγων⁵²

Προκειμένου ένας άνθρωπος να αποκτήσει τα δικαιώματα του πρόσφυγα, θα πρέπει πρώτα να του απονεμηθεί η εν λόγω ιδιότητα. Η απονομή της ιδιότητας του πρόσφυγα στην Ελλάδα γίνεται από της αρμόδιες αρχές της χώρας, δηλαδή την Υποδιεύθυνση Αλλοδαπών της Διεύθυνσης Εθνικής Ασφαλείας του Υπουργείου Δημοσίας Τάξης και όπως ορίζεται στη διυπουργική απόφαση του 1977.

Συγκεκριμένα, μία αρμόδια επιτροπή εξετάζει, τόσο την αίτηση, όσο και τον ίδιο τον αιτούντα άσυλο. Αφού εξεταστούν τα όποια έγγραφα ή αποδεικτικά στοιχεία έχει προσκομίσει ο αιτών, εάν κριθεί αναγκαίο και συνήθως στην περίπτωση που τα αποδεικτικά στοιχεία απουσιάζουν, η επιτροπή προσπαθεί με επανειλημμένες συζητήσεις να διαπιστώσει την πραγματική κατάσταση του αιτούντος, τους λόγους, έτσι ώστε να διαπιστωθεί ο «δικαιολογημένος φόβος δίωξης του.»⁵³

⁵² Κάπα research, Έρευνα σε πρόσφυγες, ό.π. παραπομπή 23, σελ. 12.

⁵³ Νάσκου-Περράκη Π., *Το νομικό καθεστώς των προσφύγων στη διεθνή και ελληνική έννομη τάξη. Η Σύμβαση της Γενεύης του 1951 για το καθεστώς των προσφύγων*, Αντ. Ν. Σάκκουλας, Αθήνα-Κομοτηνή, 1991, σελ. 279.

Επιπλέον, τον Ιανουάριο του 2011, υιοθετήθηκε ο νέος νόμος (3907/2011), για την αναμόρφωση του συστήματος ασύλου στην Ελλάδα. Ο νόμος αυτός προβλέπει τη δημιουργία αυτοτελούς Υπηρεσίας Ασύλου και ενός συστήματος υποδοχής και διαπίστωσης του «προφίλ» και των αναγκών, όσων εισέρχονται στην Ελλάδα χωρίς τις νόμιμες διατυπώσεις, συμπεριλαμβανομένων και ατόμων που επιθυμούν να ζητήσουν άσυλο.

Παρόμοια διαδικασία ακολουθείται και στη Γερμανία. Η Ομοσπονδιακή Υπηρεσία Μετανάστευσης και Προσφύγων διεξάγει μία σειρά συνεντεύξεων με τους αιτούντες άσυλο, προκειμένου να διερευνηθούν οι λόγοι της φυγής από το κράτος προέλευσης. Στη συνέχεια, ο αιτών λαμβάνει μία προσωρινή άδεια παραμονής μέχρι να βγει η τελική απόφαση για την αίτησή του, που στην περίπτωση που είναι θετική, τότε παραπέμπεται σε ένα κρατίδιο ή μία περιφέρεια, η οποία είναι από εκεί και πέρα υπεύθυνη για τη στέγαση και την περαιτέρω μέριμνα του πρόσφυγα.⁵⁴

Όσοι λαμβάνουν καθεστώς διεθνούς προστασίας έχουν πρόσβαση στην παιδεία, τις υπηρεσίες υγείας, την αγορά εργασίας και την κοινωνική ασφάλιση, ενώ οι αρμόδιες αρχές οφείλουν να εκδώσουν στους αναγνωρισμένους πρόσφυγες και ταξιδιωτικά έγγραφα.⁵⁵

Τα δικαιώματα που αναγνωρίζονται σε έναν άνθρωπο από τη στιγμή που θα γίνει δεκτή η αίτηση ασύλου που έχει υποβάλει και θεωρηθεί πρόσφυγας διακρίνονται σε γενικά και ειδικά:

Στα γενικά δικαιώματα των προσφύγων περιλαμβάνονται η αρχή της ίσης μεταχείρισης, στην οποία συμπεριλαμβάνεται η αρχή της μη διάκρισης ως προς την φυλή, τη θρησκεία και τη χώρα καταγωγής, η προστασία δικαιωμάτων και

⁵⁴ Wer bekommt Asyl in Deutschland?, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://www.caritas.de/hilfeundberatung/ratgeber/migration/asylundflucht/wer-bekommt-asyl-in-deutschland>, [προσπελάστηκε στις 12 Φεβρουαρίου 2017].

⁵⁵ Υπουργείο Προστασίας του Πολίτη, 2014, Ελληνική Δημοκρατία, Συχνές ερωτήσεις, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: http://www.minocp.gov.gr/index.php?option=ozo_content&lang=EN&perform=view&id=5055&Itemid=466&lang=GR, [προσπελάστηκε στις 6 Φεβρουαρίου 2017].

ευεργετημάτων και η εξαίρεση των προσφύγων από έκτακτα μέτρα που τυχόν λαμβάνονται σε βάρος του ατόμου.⁵⁶

Στα ειδικά δικαιώματα⁵⁷ περιλαμβάνονται η μεταχείριση το ίδιο ευνοϊκή με αυτή των ημεδαπών, η μεταχείριση περισσότερο ευνοϊκή από υπηκόους αλλοδαπής πολιτείας και η μεταχείριση κατά το δυνατόν ευνοϊκή, σε καμία περίπτωση λιγότερο ευνοϊκή από αυτήν που παρέχεται στους αλλοδαπούς γενικά κάτω από τις ίδιες περιστάσεις. Ως προς την ίση μεταχείριση με τους ημεδαπούς, οι πρόσφυγες απολαμβάνουν όπως και οι ημεδαποί τη θρησκευτική τους ελευθερία, το δικαίωμα να προσφεύγουν στην δικαιοσύνη, να σπουδάζουν τα παιδιά τους στα δημόσια σχολεία, να προστατεύονται τα δικαιώματα της πνευματικής τους ιδιοκτησίας, υπό μία σειρά προϋποθέσεων να μην εφαρμόζονται τα περιοριστικά μέτρα που εφαρμόζονται στους αλλοδαπούς σε θέματα εργασίας, να έχουν την ίδια μεταχείριση σε περίπτωση που ανεφοδιασμός του πληθυσμού γίνεται με το σύστημα δελτίου, ενώ απαλλάσσονται και από την εγγυοδοσία του αλλοδαπού.

Επίσης, το 2014 τέθηκε σε ισχύ από την Ευρωπαϊκή Ένωση η νέα οδηγία για την αναγνώριση των προσφύγων. Οι διατάξεις της προβλέπουν μια σειρά δικαιωμάτων σχετικά με την προστασία από την επαναπροώθηση, τις άδειες διαμονής, τα ταξιδιωτικά έγγραφα, την πρόσβαση στην απασχόληση, την πρόσβαση στην εκπαίδευση, την κοινωνική πρόνοια, την υγειονομική περίθαλψη, την πρόσβαση σε κατάλυμα, την πρόσβαση σε υπηρεσίες κοινωνικής ένταξης, καθώς και ειδικές διατάξεις για τα παιδιά και τα ευάλωτα άτομα.

Τα ελάχιστα πρότυπα στην προηγούμενη οδηγία ήταν έως ένα βαθμό ασαφή, γεγονός που συνέβαλε στη διατήρηση των αποκλίσεων στις εθνικές νομοθεσίες και πρακτικές περί ασύλου. Οι πιθανότητες εξασφάλισης διεθνούς προστασίας από ένα πρόσωπο μπορούσαν να παρουσιάζουν μεγάλες διακυμάνσεις ανάλογα με το ποιο κράτος μέλος εξέταζε την αίτηση ασύλου. Η νέα οδηγία για την αναγνώριση ουσιαστικά συμβάλει στη βελτίωση της ποιότητας της λήψης αποφάσεων και στη διασφάλιση της δίκαιης

⁵⁶ Νάσκου-Περράκη, Μ., Δικαιώματα και υποχρεώσεις στα πλαίσια της Σύμβασης της Γενεύης του 1951 για το καθεστώς των προσφύγων: Γενική θεώρηση, *Κοινωνική Εργασία*, 1992, 26, σελ. 93.

⁵⁷ Νάσκου-Περράκη, Μ., Δικαιώματα και υποχρεώσεις στα πλαίσια της Σύμβασης της Γενεύης του 1951 για το καθεστώς των προσφύγων, ό.π. παραπομπή 56, σελ. 93.

μεταχείρισης, με ομοιόμορφο τρόπο, των προσώπων που προσπαθούν να αποφύγουν τη δίωξη, τον πόλεμο και τα βασανιστήρια.

Σημαντικό είναι ότι η νέα οδηγία διασφαλίζει να λαμβάνονται καλύτερα υπόψη το υπέρτατο συμφέρον του παιδιού και τα ζητήματα που σχετίζονται με το φύλο κατά την αξιολόγηση των αιτήσεων ασύλου, καθώς και στην εφαρμογή των κανόνων για το περιεχόμενο της διεθνούς προστασίας.⁵⁸ Τόσο η Γερμανία, όσο και η Ελλάδα, έχουν ενσωματώσει την παραπάνω οδηγία⁵⁹ της Ευρωπαϊκής Ένωσης στη νομοθεσία τους.

1.7 Η ΕΝΣΩΜΑΤΩΣΗ ΤΩΝ ΠΡΟΣΦΥΓΩΝ

Η αναγνώριση των δικαιωμάτων του πρόσφυγα κρίνεται απαραίτητη για την ομαλή ένταξη του στη χώρα υποδοχής. Η ενσωμάτωση των προσφύγων θα πρέπει να επιτυγχάνεται στο συντομότερο δυνατό χρονικό διάστημα, καθώς συμβάλλει στην αποκατάσταση της ανθρώπινης αξιοπρέπειας. Τα συστήματα κοινωνικής ένταξης αποτελούν μια δυναμική, αμφίδρομη καθώς και δισδιάστατη διαδικασία μεταξύ των προσφύγων και των κοινωνιών ένταξης και περιλαμβάνουν δικαιώματα και υποχρεώσεις και για τις δύο πλευρές, ενώ η ίδια η διαδικασία ένταξης προσαρμόζεται ανάλογα με τις συνθήκες που επικρατούν σε κάθε περιοχή.⁶⁰

Συγκεκριμένα, η κοινωνική ενσωμάτωση των προσφύγων αφορά, τόσο στην ικανότητα του πρόσφυγα να προσαρμοστεί στην κοινωνία στην οποία αναζητά άσυλο, όσο και στην ίδια την κοινωνία να ανταποκριθεί στις υποχρεώσεις της για την προστασία των αιτούντων άσυλο, μέσω της εφαρμογής συγκεκριμένων πολιτικών που τους αφορούν (policy making). Επίσης, η ενσωμάτωση είναι μια διαδικασία, η οποία

⁵⁸ Ευρωπαϊκή Επιτροπή, Ετήσια έκθεση δραστηριοτήτων 2013, 2014, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://www.easo.europa.eu/sites/default/files/public/BZAD14001ELN.pdf>, [προσπελάστηκε στις 7 Φεβρουαρίου 2017], σελ. 7.

⁵⁹ Επίσημη Εφημερίδα της Ευρωπαϊκής Ένωσης, Οδηγία 2013/33/ΕΕ του Ευρωπαϊκού Κοινοβουλίου και του συμβουλίου της 26ης Ιουνίου 2013 σχετικά με τις απαιτήσεις για την υποδοχή των αιτούντων διεθνή προστασία (αναδιτύπωση), 2013, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:180:0096:0116:EL:PDF>, [προσπελάστηκε στις 15 Μαρτίου 2017].

⁶⁰ Κέντρο Ευρωπαϊκής Πληροφόρησης Europe Direct, Η κοινωνική ένταξη και η ενσωμάτωση των προσφύγων στην αγορά εργασίας, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.acci.gr/acci/articles/article.jsp?context=103&categoryid=8289&articleid=1038>, [προσπελάστηκε στις 11 Μαρτίου 2017].

ξεκινάει από την στιγμή της άφιξης του πρόσφυγα και ολοκληρώνεται με την εξασφάλιση συνθηκών διαβίωσης εφάμιλλων με αυτών που απολαμβάνει ο εγχώριος πληθυσμός.⁶¹

Οι Ager και Strang⁶² κατασκεύασαν μια σειρά από δείκτες που επιχειρούν να μετρήσουν ποσοτικά και ποιοτικά τον βαθμό με τον οποία επιτυγχάνεται η διαδικασία της ενσωμάτωσης και χρησιμοποιήθηκαν και από τη βρετανική κυβέρνηση, προκειμένου να εξαχθούν συμπεράσματα σχετικά με την επιτυχία των επίσημων πολιτικών για τους πρόσφυγες. Το μοντέλο αυτό, με τους συγκεκριμένους δείκτες μέτρησης έχει ως εξής:

Μέσα συμμετοχής	Εργασία και Απασχόληση	Στέγαση	Εκπαίδευση	Υγεία
Παράγοντες Διευκόλυνσης	Γλωσσική Επικοινωνία και διαπολιτισμική επαφή		Ασφάλεια και Σταθερότητα	
Θεμελίωση της κοινωνικής συμμετοχής	Κοινωνικά και Πολιτικά Δικαιώματα			

Πίνακας 1. Δείκτες μέτρησης ενσωμάτωσης των προσφύγων⁶³

Ο παραπάνω πίνακας μας δείχνει τον τρόπο με τον οποίο μπορούμε να διερευνήσουμε το βαθμό της κοινωνικής συμμετοχής των προσφύγων στην κοινωνία υποδοχής, έχοντας ως αφετηρία τις βασικές δραστηριότητες που σχετίζονται με την επιβίωση των προσφύγων, εν συνεχεία τους παράγοντες που συντελούν στην ομαλότερη ενοποίηση με το εγχώριο κοινωνικό σώμα, και τέλος τα δικαιώματα που απολαμβάνουν οι πρόσφυγες σε σχέση με τον εγχώριο πληθυσμό.

Η πρώτη κατηγορία, αυτή των μέσων συμμετοχής, όπως η εργασία, η στέγαση, η πρόσβαση στις υπηρεσίες εκπαίδευσης και υγείας, αφορά στους τομείς εκείνους που

⁶¹ Phillimore J. & Goodson L., Making a place in the Global City: The relevance of Indicators of Integration, *Journal of Refugee Studies*, 2008, 21 (3), Oxford Uni.Press, σελ. 305-325.

⁶² Ager, A. & Strang, A., Indicators of Integration : Final Report, *Home Office Development and Practice Report 28*, 2004, Home Office, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://webarchive.nationalarchives.gov.uk/20110218141321/http://rds.homeoffice.gov.uk/rds/pdfs04/dpr28.pdf>, [προσπελάστηκε στις 23 Μαρτίου 2017], σελ. 13- 23.

⁶³ Ο.π.

σχετίζονται με την καθημερινή επιβίωση των προσφύγων και την εξασφάλιση ενός είδους κανονικότητας (normality) στη ζωή τους στη χώρα υποδοχής. Επίσης, αποτελούν τους τομείς εκείνους των κοινωνικών δραστηριοτήτων και της πρόσβασης σε βασικές κοινωνικές παροχές, οι οποίοι μπορούν να εξασφαλίσουν στους πρόσφυγες καλύτερη ενσωμάτωση στην κοινωνία υποδοχής.

Παρομοίως, οι Mestheneos και Ioannidi⁶⁴ θεωρούν ότι ο βαθμός ενσωμάτωσης των προσφύγων και μεταναστών στις ευρωπαϊκές κοινωνίες δύναται να μετρηθεί ποσοτικά με την εισαγωγή μίας σειράς από δείκτες που αφορούν πρακτικά ζητήματα της ζωής αυτών στις κοινωνίες υποδοχής και που προκύπτουν από την παρουσία τους, σε βραχυχρόνιο και μακροχρόνιο επίπεδο. Τέτοιοι δείκτες είναι: η (νόμιμη) απασχόληση και εργασία, η δυνατότητα συμμετοχής στο εκπαιδευτικό σύστημα, η στέγαση, καθώς και η δυνατότητα για πολιτική συμμετοχή και εκπροσώπηση στα κοινά.

Επομένως, τόσο από τους δείκτες των Anger και Strang και συγκεκριμένα από την τρίτη υποκατηγορία των μέσων συμμετοχής, όσο και από τους δείκτες των Mestheneos και Ioannidi παρατηρούμε το ρόλο που παίζει στην ολοκλήρωση της ενσωμάτωσης η δυνατότητα συμμετοχής των προσφύγων στο εκπαιδευτικό σύστημα της χώρας υποδοχής.

Στην πράξη, για την επίτευξη της ενσωμάτωσης των προσφύγων, ψηφίστηκε στη Γερμανία στις 31 Ιουλίου του 2016 και τέθηκε σε εφαρμογή από τις 31 Αυγούστου του ίδιου έτους ο «Νόμος Ενσωμάτωσης Προσφύγων»⁶⁵. Ο εν λόγω νόμος ρυθμίζει έξι βασικά θέματα δίνοντας ιδιαίτερη βάση στην εκπαίδευση και συγκεκριμένα προβλέπει μαθήματα ενσωμάτωσης στη γερμανική κουλτούρα, την κοινωνία και τις γερμανικές αξίες, μαθήματα γλώσσας και προγράμματα εργασίας, ενώ ρυθμίζει την εργατική νομοθεσία, την πρόληψη της δημιουργίας «γκέτο» και τα ζητήματα της μόνιμης κατοικίας.

Σύμφωνα με την Καγκελάριο Μέρκελ, ο νόμος αυτός αποτελεί «μία προσφορά, αλλά παράλληλα περιέχει υποχρεώσεις και καθήκοντα για όλους», ενώ έχει χαρακτηριστεί ως

⁶⁴ Mestheneos, E. & Ioannidi, E., Obstacles to Refugee Integration in the European Union Member States, *Journal of Refugee Studies*, 2002, 15, Oxford Uni. Press, σελ. 304-320.

⁶⁵ Στελλάκη, Φ., Ο Γερμανικός «Νόμος Ενσωμάτωσης» Προσφύγων, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://powerpolitics.eu/o-germanikos-nomos-ensomatwshs-prosphy/>, [προσπελάστηκε στις 8 Φεβρουαρίου 2017].

ένα ιστορικό βήμα για την καλύτερη συνειδητοποίηση του εκσυγχρονισμού και του «ανοίγματος» της γερμανικής κοινωνίας.

Ανάμεσα στις αντιδράσεις που προκλήθηκαν ήταν και αυτή του κρατιδίου της Βαυαρίας, όπου ψηφίστηκε νόμος, ο οποίος υποχρεώνει τους πρόσφυγες να σέβονται την επικρατούσα τοπική κουλτούρα, προκειμένου να προστατευτεί η κουλτούρα της Βαυαρίας⁶⁶. Σύμφωνα με την υπουργό Κοινωνικής Ασφάλειας της Βαυαρίας Emilia Müller, «ο νέος νόμος εμποδίζει την εμφάνιση παράλληλων κοινωνιών που διέπονται από την Σαρία, θρησκευτικούς κώδικες ή φυλετικά έθιμα, αντί των νόμων μας και των αξιών μας». Σύμφωνα με το Γερμανικό Ομοσπονδιακό σύστημα, ο νέος νόμος θα εφαρμοστεί μόνο στη Βαυαρία και δεν επηρεάζει τα υπόλοιπα 15 κρατίδια.

1.8 ΤΟ ΔΙΚΑΙΩΜΑ ΤΩΝ ΠΡΟΣΦΥΓΩΝ ΣΤΗΝ ΕΚΠΑΙΔΕΥΣΗ

Το ποσοστό των παιδιών προσφύγων που έχει πρόσβαση στην εκπαίδευση είναι αρκετά χαμηλό, ενώ τις περισσότερες φορές η εκπαίδευση που παρέχεται δεν είναι ούτε η στοιχειώδης. Αυτό είναι κάτι που αφενός παραβιάζει τα δικαιώματα του παιδιού και αφετέρου επηρεάζει την υπόλοιπη ζωή του.

Πολλά είναι τα κράτη που δεν έχουν τη δυνατότητα να παρέχουν πλήρως τις αναγκαίες υπηρεσίες για την εκπαίδευση των παιδιών αυτών. Τα πιο συχνά προβλήματα που προκύπτουν είναι οι ελλείψεις σε δομές και σε ανθρώπινο δυναμικό. Επιπλέον, το γεγονός ότι πολλές φορές δεν υπάρχει η δυνατότητα για εκπαίδευση στη μητρική γλώσσα των προσφύγων, αλλά και το ότι οι δυνατότητες για περαιτέρω πρόσβαση σε όλες τις βαθμίδες της εκπαίδευσης είναι περιορισμένη, επηρεάζει αρνητικά την ευημερία και την οικονομική αυτάρκεια των προσφύγων.⁶⁷

Προτεραιότητα της Ύπατης Αρμοστείας αποτελεί η ένταξη των παιδιών των προσφύγων στο εθνικό εκπαιδευτικό σύστημα. Ήδη από το 2009, η Ύπατη Αρμοστεία

⁶⁶ Huggler, J., Bavaria passes new law to make migrants respect 'dominant' local culture, *The Telegraph*, 2016, 9 Δεκεμβρίου, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.telegraph.co.uk/news/2016/12/09/bavaria-passes-new-law-make-migrants-respect-local-culture/>, [προσπελάστηκε στις 16 Μαρτίου 2017].

⁶⁷ Σπανού Φ. & Αλεξάκης Β., Ασυνόδετοι ανήλικοι πρόσφυγες στην Ελλάδα, ό.π. παραπομπή 41, σελ. 47.

έχει ενισχύσει τις προσπάθειές της, ώστε τα παιδιά των προσφύγων να έχουν πρόσβαση στα τοπικά εκπαιδευτικά ιδρύματα και έχει αυξήσει τις θέσεις στα σχολεία, όπου αυτό ήταν δυνατόν, προκειμένου να αποκτήσει πρόσβαση στη στοιχειώδη εκπαίδευση το μεγαλύτερο ποσοστό των παιδιών.⁶⁸

Βασικός στόχος της στοιχειώδους εκπαίδευσης είναι να διασφαλιστεί η πρόσβαση και να αμβλυνθούν οι διαφορές μεταξύ των φύλων, ενώ θα πρέπει να περιλαμβάνει τουλάχιστον γνώσεις γραφής, ανάγνωσης και αριθμητικής. Επιπλέον, για τη δευτεροβάθμια εκπαίδευση αναγκαία κρίνεται η πληροφόρηση σχετικά με τις δυνατότητες που υπάρχουν σε αυτή και για τις ευκαιρίες που τους δίνονται. Τέλος, θα πρέπει να γίνεται προσπάθεια, έτσι ώστε η παρεχόμενη εκπαίδευση να διατίθεται με την έναρξη της προσφυγικής κατάστασης, κάτι που σημαίνει ότι θα πρέπει να υφίσταται και στα κέντρα υποδοχής προσφύγων, αλλά και καθ' όλη την διάρκεια εξέτασης του αιτήματος ασύλου.

Τα περισσότερα προβλήματα εντοπίζονται συνήθως όταν οι πρόσφυγες έχουν προσφάτως μεταβεί στη χώρα προορισμού τους. Ιδιαίτερως στις χώρες που δέχτηκαν μεγάλο αριθμό προσφύγων το τελευταίο διάστημα, δεν μπορεί να εγγυηθεί κανείς στα παιδιά αυτά ότι θα έχουν άμεση πρόσβαση στην εκπαίδευση. Για παράδειγμα, σε χώρες, όπως είναι το Βέλγιο και η Ολλανδία, εντοπίζεται μεγάλη καθυστέρηση στην ένταξη των παιδιών προσφύγων στην εκπαίδευση. Σε άλλες χώρες πάλι, όπως είναι η Εσθονία και η Πολωνία, η κατάσταση είναι ακόμα πιο κρίσιμη, με τα παιδιά των προσφύγων να μην έχουν καμία απολύτως πρόσβαση στην εκπαίδευση. Στην Ελλάδα και στην Ιταλία, που δεν αποτελούν χώρες προορισμού, αλλά χώρες υποδοχής, ενώ προβλέπεται η ένταξη των παιδιών αυτών στο εκπαιδευτικό σύστημα, η διαδικασία παρουσιάζει πολλά προβλήματα, με αποτέλεσμα τα παιδιά, που ενδεχομένως στερήθηκαν για ένα διάστημα, μικρότερο ή μεγαλύτερο, το δικαίωμα να πηγαίνουν

⁶⁸ UNHCR, *Οι πρόσφυγες του κόσμου: Σε αναζήτηση αλληλεγγύης (σύννοψη)*, Ύπατη Αρμοστεία του ΟΗΕ για τους Πρόσφυγες, 2012, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: https://www.unhcr.gr/fileadmin/Greece/General/publications/OI_PROSFYGES_TOY_KOSMOY.pdf, [προσπελάστηκε στις 23 Μαρτίου 2017],σελ. 24.

σχολείο, εξαιτίας της εμπόλεμης κατάστασης στην πατρίδα τους, να μένουν ίσως και για χρόνια εκτός σχολείων.⁶⁹

Επομένως, πολλά είναι τα ερωτήματα που προκύπτουν σχετικά με την εκπαίδευση των προσφύγων, όπως για παράδειγμα ποια είναι ακριβώς τα δικαιώματα των παιδιών αυτών, εάν η χώρα υποδοχής έχει υποχρέωση να εντάξει τα παιδιά στο εκπαιδευτικό της σύστημα, εάν είναι το κράτος που έχει την ευθύνη για την εκπαίδευσή τους ή εάν μπορεί να αρνηθεί να παράσχει τέτοιες υπηρεσίες.⁷⁰

1.8.1 ΤΟ ΝΟΜΙΚΟ ΠΛΑΙΣΙΟ ΣΤΗΝ ΕΛΛΑΔΑ

Στην Ελλάδα, το δικαίωμα στην εκπαίδευση κατοχυρώνεται στις διατάξεις του άρθρου 16 του Συντάγματος, το οποίο θεσπίστηκε το 1975 και έχει χαρακτηριστεί ως το πιο πολυσύνθετο από τα δικαιώματα που κατοχυρώνονται στο Ελληνικό Σύνταγμα.⁷¹

Συγκεκριμένα στο άρθρο 16, παράγραφο 4⁷² διατυπώνεται το δικαίωμα στην εκπαίδευση ως εξής: *«Όλοι οι Έλληνες έχουν δικαίωμα δωρεάν παιδείας, σε όλες τις βαθμίδες της, στα κρατικά εκπαιδευτήρια. Το Κράτος ενισχύει τους σπουδαστές που διακρίνονται, καθώς και αυτούς που έχουν ανάγκη από βοήθεια ή ειδική προστασία, ανάλογα με τις ικανότητές τους.»*

Το πρώτο εδάφιο της εν λόγω παραγράφου έχει μία ασυνήθιστα ακριβή και αναλυτική διατύπωση σε σχέση με αυτές που καθορίζουν τα υπόλοιπα κοινωνικά δικαιώματα και είναι πιο γενικόλογες. Αν σταθεί κανείς στον όρο *«όλοι οι Έλληνες»*, διαπιστώνει ότι χρήζει περαιτέρω ερμηνείας. Το Σύνταγμα φαίνεται να κατοχυρώνει το δικαίωμα της

⁶⁹ Fagerholm, K., & Verheul, R., Safety and Fundamental Rights at Stake for Children on the Move, ENOC, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://www.synigoros.gr/resources/160211-ekthesi.pdf> , [προσπελάστηκε στις 12 Μαρτίου 2017].

⁷⁰ Τσιτσελίκης Κ., Ανήλικοι πρόσφυγες και υπηρεσίες εκπαίδευσης. Δικαίωμα στην εκπαίδευση και υποχρεώσεις του κράτους, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: https://www.esos.gr/sites/default/files/articles-legacy/keimeno_tekmiriosis_dik_ekpaideusi2_1.pdf , [προσπελάστηκε στις 12 Φεβρουαρίου 2017].

⁷¹ Καϊδατζής, Α., Το δικαίωμα δωρεάν παιδείας μεταξύ (συνταγματικού) δικαίου και (νομοθετικής) πολιτικής. Με αφορμή την απόφαση για τα δίδακτρα στις μεταπτυχιακές σπουδές, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.constitutionalism.gr/2558-to-dikaiwma-dwrean-paideias-metaxy-syntagmatikoy-d/> , [προσπελάστηκε στις 12 Φεβρουαρίου 2017].

⁷² Σύνταγμα της Ελλάδας, Όπως αναθεωρήθηκε με το ψήφισμα της 27^{ης} Μαΐου 2008 της Η' Αναθεωρητικής Βουλής των Ελλήνων, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.hellenicparliament.gr/UserFiles/8c3e9046-78fb-48f4-bd82-bbba28ca1ef5/SYNTAGMA.pdf> , [προσπελάστηκε στις 23 Φεβρουαρίου 2017], σελ. 30.

δωρεάν παιδείας ως δικαίωμα του Έλληνα πολίτη και μόνο. Όμως, όπως έχει επισημανθεί, το βάρος στη συγκεκριμένη διατύπωση δεν πέφτει τόσο στο «Έλληνες», όσο στο «όλοι», το οποίο έρχεται να υποδηλώσει την καθολικότητα της δωρεάν παιδείας.⁷³ Μπορεί δηλαδή το Σύνταγμα να αναφέρεται αποκλειστικά στους Έλληνες ως φορείς του δικαιώματος στην παιδεία, ωστόσο αυτό δε συνεπάγεται τον αποκλεισμό των αλλοδαπών από την εκπαίδευση.⁷⁴ Άλλωστε, ο χαρακτήρας του Συντάγματος είναι τέτοιος που αρκείται σε μία γενική στοχοθεσία και αφήνει τον καθορισμό των μέσων στον κοινό νομοθέτη.⁷⁵

Επίσης, στην παράγραφο 3 του άρθρου 16, όπου αναφέρεται πως «τα έτη υποχρεωτικής φοίτησης δεν μπορεί να είναι λιγότερα από εννέα», δεν γίνεται διάκριση μεταξύ ημεδαπών και αλλοδαπών και κατ' επέκταση ούτε περαιτέρω διάκριση μεταξύ των τελευταίων, σε ό,τι αφορά στη νομιμότητα ή μη της διαμονής τους στη χώρα. Για τον λόγο αυτό δεν θα πρέπει να ζητώνται δικαιολογητικά παραμονής στη χώρα ως προϋπόθεση για την εγγραφή των αλλοδαπών στα σχολεία, καθώς τόσο δικαίωμα, όσο και υποχρέωση φοίτησης, φαίνεται να έχουν ακόμα και τα παιδιά των παράνομων μεταναστών.⁷⁶

Επιπλέον, οι υποχρεώσεις της χώρας για την εφαρμογή του δικαιώματος στην εκπαίδευση των προσφύγων καθορίζονται από μία σειρά νόμων του ελληνικού κράτους. Ενδεικτικά, σύμφωνα με τον νόμο 2910/2001 και το άρθρο 40, όλοι οι αλλοδαποί έχουν το δικαίωμα να συμμετέχουν στο εκπαιδευτικό σύστημα της χώρας ισότιμα με τους Έλληνες. Ακριβώς το ίδιο προβλέπει και το άρθρο 72 του νόμου 3386/2005 που τον διαδέχτηκε.

⁷³ Καϊδατζής, Α., Το δικαίωμα δωρεάν παιδείας μεταξύ (συνταγματικού) δικαίου και (νομοθετικής) πολιτικής, ό.π. παραπομπή 71.

⁷⁴ Δημητρόπουλος, Α., *Τα Συνταγματικά δικαιώματα των αλλοδαπών στην Ελλάδα: Constitutional rights of foreign citizens in Greece*, Αντ. Ν. Σάκκουλα, Αθήνα, 2004, σελ. 7

⁷⁵ Χρυσόγονος, Κ., *Ατομικά και Κοινωνικά Δικαιώματα*, Νομική βιβλιοθήκη, 2006, σελ. 334

⁷⁶ Ο. π., σελ. 337

Πιο αναλυτικά, το άρθρο 40⁷⁷ αναφέρει, μεταξύ άλλων, ότι τα παιδιά των προσφύγων έχουν αφενός την υποχρέωση να παρακολουθήσουν τις εννέα πρώτες τάξεις της πρωτοβάθμιας και δευτεροβάθμιας εκπαίδευσης και αφετέρου έχουν στη συνέχεια το δικαίωμα να συνεχίσουν τη φοίτησή τους. Το κράτος οφείλει να παρέχει όχι μόνο την ελάχιστη σχολική φοίτηση στους αλλοδαπούς ανήλικους, αλλά να διασφαλίζει ότι οι ανήλικοι θα λαμβάνουν τουλάχιστον εννεαετή εκπαίδευση και κατόπιν να τους παρέχει ελεύθερη πρόσβαση σε όλες τις εκπαιδευτικές βαθμίδες. Παρόμοιες διατάξεις περιλαμβάνει και ο νόμος νόμος 4251/2014⁷⁸ άρθρο 21 (ΦΕΚ Α' 80).

Μεταγενέστερα, με τον νόμο 4415/2016 (ΦΕΚ Α 159) και το άρθρο 38, οι υπουργοί Παιδείας και Οικονομικών εξουσιοδοτούνται, ώστε να ιδρύουν δομές υποδοχής για την εκπαίδευση των προσφύγων, καθώς και να καθορίζουν την οργάνωση, τη λειτουργία και το πρόγραμμα της εκπαίδευσης αυτής. Έτσι, σε εκτέλεση της εν λόγω διάταξης, εκδόθηκε Κοινή Υπουργική Απόφαση (Κ.Υ.Α.) με αριθμό 152360/ΓΔ4 (ΦΕΚ 3049 Β 2006), η οποία προβλέπει την ίδρυση δομών υποδοχής για την εκπαίδευση των προσφύγων (ΔΥΕΠ), οι οποίες λειτουργούν από το σχολικό έτος 2016-2017, είτε εντός των σχολικών μονάδων, είτε στα κέντρα φιλοξενίας των προσφύγων. Οι παραπάνω δομές αποτελούν παραρτήματα των ήδη υφιστάμενων παρακείμενων σχολικών μονάδων σε επίπεδο νηπιαγωγείου, πρωτοβάθμιας και δευτεροβάθμιας εκπαίδευσης.

Επιπλέον, με απόφαση του υπουργού Παιδείας (131024/Δ1 στο ΦΕΚ Β 2687/2016) ορίζονται Ζώνες Εκπαιδευτικής Προτεραιότητας (ΖΕΠ) για τις ανάγκες των παιδιών

⁷⁷ Άρθρο 40: 1. Ανήλικοι αλλοδαποί, που διαμένουν στην ελληνική επικράτεια, υπάγονται στην υποχρέωση της ελάχιστης σχολικής φοίτησης, όπως και οι ημεδαποί. 2. Οι ανήλικοι αλλοδαποί που φοιτούν σε όλες τις βαθμίδες της εκπαίδευσης έχουν χωρίς περιορισμούς πρόσβαση στις δραστηριότητες της σχολικής ή εκπαιδευτικής κοινότητας. 3. Για την εγγραφή αλλοδαπών ανηλίκων στα δημόσια σχολεία απαιτούνται τα αντίστοιχα με τα προβλεπόμενα για τους ημεδαπούς δικαιολογητικά. Κατ' εξαίρεση, με ελλιπή δικαιολογητικά μπορεί να εγγράφονται στα δημόσια σχολεία τέκνα: α. όσων προστατεύονται από το ελληνικό κράτος ως πρόσφυγες και όσων τελούν υπό την προστασία της Υπατης Αρμοστείας των Ηνωμένων Εθνών, β. όσων προέρχονται από περιοχές στις οποίες επικρατεί έκρυθμη κατάσταση, γ. όσων έχουν υποβάλει αίτηση για τη χορήγηση ασύλου και δ. αλλοδαπών που διαμένουν στην Ελλάδα, ακόμη και αν δεν έχει ρυθμιστεί η νόμιμη παραμονή τους σε αυτήν.

⁷⁸ Άρθρο 21: 7. Ανήλικοι πολίτες τρίτων χωρών, που διαμένουν στην ελληνική επικράτεια, υπάγονται στην υποχρεωτική σχολική φοίτηση, όπως και οι ημεδαποί. Οι ανήλικοι πολίτες τρίτων χωρών, που φοιτούν σε όλες τις βαθμίδες της εκπαίδευσης, έχουν, χωρίς περιορισμούς, πρόσβαση στις δραστηριότητες της σχολικής ή εκπαιδευτικής κοινότητας. 8. Για την εγγραφή ανηλίκων πολιτών τρίτων χωρών στα ελληνικά σχολεία, όλων των βαθμίδων, απαιτούνται τα αντίστοιχα με τα προβλεπόμενα για τους ημεδαπούς δικαιολογητικά. Κατ' εξαίρεση, με ελλιπή δικαιολογητικά μπορεί να εγγράφονται στα δημόσια σχολεία και τέκνα πολιτών τρίτων χωρών, εφόσον: α. Προστατεύονται από το ελληνικό κράτος ως δικαιούχοι διεθνούς προστασίας και όσων τελούν υπό την προστασία της Υπατης Αρμοστείας των Ηνωμένων Εθνών. β. Προέρχονται από περιοχές, στις οποίες επικρατεί έκρυθμη κατάσταση. γ. Έχουν υποβάλει αίτηση διεθνούς προστασίας. δ. Είναι πολίτες τρίτων χωρών που διαμένουν στην Ελλάδα, ακόμη και αν δεν έχει ρυθμισθεί η νόμιμη διαμονή τους σε αυτή.

προσφύγων, όπου ιδρύονται Τάξεις Υποδοχής (ΤΥ), Ενισχυτικά Φροντιστηριακά Τμήματα (ΦΤ) και Δομές Υποδοχής για την Εκπαίδευση των Προσφύγων (Δ.Υ.Ε.Π.). Το νέο αυτό πλαίσιο στοχεύει στην ομαλή και ισότιμη ένταξη των προσφύγων στο ελληνικό εκπαιδευτικό σύστημα, καθώς από τη στιγμή που τα παιδιά αυτά δεν μπορούν να ανταποκριθούν στις γλωσσικές απαιτήσεις της τάξης που αντιστοιχεί στην ηλικία τους, κρίνεται αναγκαίο να ενταχθούν σε ένα εναλλακτικό σχήμα, όπως είναι τα παραπάνω. Προκειμένου όμως να μην οδηγηθούμε στον κοινωνικό αποκλεισμό και την περαιτέρω περιθωριοποίηση των προσφύγων, οι παράλληλες αυτές τάξεις, μόνο προσωρινή λύση μπορούν να αποτελέσουν.

Σε καμία περίπτωση όμως το κράτος δεν μπορεί να αρνηθεί το δικαίωμα των παιδιών αυτών στην εκπαίδευση και να μην υλοποιήσει την εν λόγω παροχή.

1.8.2 ΤΟ ΝΟΜΙΚΟ ΠΛΑΙΣΙΟ ΣΤΗ ΓΕΡΜΑΝΙΑ

Το Σύνταγμα της Ομοσπονδιακής Δημοκρατίας της Γερμανίας δεν αναφέρεται ρητώς στο δικαίωμα στην εκπαίδευση. Παρόλα αυτά το δικαίωμα στην εκπαίδευση προκύπτει από άλλα δικαιώματα που θεμελιώνονται στο γερμανικό Σύνταγμα. Συγκεκριμένα, στο άρθρο 1 παράγραφος 1, το απαραβίαστο της ανθρώπινης αξιοπρέπειας απαγορεύει την αυθαίρετη στέρηση των ανθρώπων από τις ευκαιρίες στην εκπαίδευση. Επίσης, η συνταγματική αρχή της ισότητας, στο άρθρο 3, απαγορεύει το να γίνονται διακρίσεις στους ανθρώπους εξαιτίας του φύλου τους, της καταγωγής τους, της φυλής τους, της γλώσσας τους, της πατρίδας τους, της πίστης τους, των θρησκευτικών ή πολιτικών τους απόψεων ή κάποιας αναπηρίας, όσον αφορά στην εξασφάλιση της εκπαίδευσής τους.⁷⁹

Το θέμα της υποχρεωτικής εκπαίδευσης των μεταναστών στη Γερμανία προκαλεί συχνά τόσο νομικές αντιπαραθέσεις, όσο και πολιτικές συζητήσεις γύρω από το δικαίωμα στην εκπαίδευση, με αφορμή την ενσωμάτωση των προσφύγων στο κράτος υποδοχής. Για παράδειγμα, η υποχρεωτική εκπαίδευση των παιδιών προσφύγων αποτέλεσε κεντρικό θέμα αντιπαραθέσεως στη συζήτηση του «Νόμου ενσωμάτωσης»

⁷⁹ Grundgesetz für die Bundesrepublik Deutschland, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://www.bundestag.de/gg>, [προσπελάστηκε στις 23 Φεβρουαρίου 2017].

στη Βαυαρία. Στον νέο νόμο θίγονται ζητήματα, όπως η κατάργηση της μαντίλας από τα σχολεία και τα μικτά τμήματα στο μάθημα της γυμναστικής.

1.8.3 ΟΙ ΔΙΕΘΝΕΙΣ ΚΑΙ ΕΥΡΩΠΑΪΚΕΣ ΣΥΝΘΗΚΕΣ

Επιτακτικοί κανόνες τόσο του διεθνούς, όσο και του ευρωπαϊκού δικαίου επεκτείνουν το δικαίωμα στην εκπαίδευση σε όλα τα παιδιά ανεξαρτήτως εθνικότητας, θρησκείας ή οποιασδήποτε άλλης διάκρισης, τουλάχιστον όσον αφορά στην υποχρεωτική εκπαίδευση.

Τα παιδιά- πρόσφυγες, καθώς είναι ταυτόχρονα και παιδιά και πρόσφυγες, χρήζουν διπλής προστασίας, κάτι το οποίο αναγνωρίζεται από την Ύπατη Αρμοστεία των Ηνωμένων Εθνών για τους Πρόσφυγες, η οποία υιοθέτησε τη λεγόμενη «Πολιτική για τα Παιδιά Πρόσφυγες», η οποία και επικυρώθηκε από την Εκτελεστική Επιτροπή του Οργανισμού τον Οκτώβριο του 1993.⁸⁰ Αξίζει να αναφερθεί ότι οι ανήλικοι πρόσφυγες, πέραν του γεγονότος ότι είναι και ανήλικοι και πρόσφυγες, γεγονός που τους καθιστά ιδιαιτέρως ευάλωτους, ενδέχεται να επιβαρύνονται και από επιπλέον χαρακτηριστικά: Οι χρόνια ασθενείς, οι ανάπηροι, καθώς και τα ασυνόδευτα παιδιά αποτελούν υποκατηγορίες ανήλικων προσφύγων με ειδικές ανάγκες που δημιουργούν επιπλέον υποχρεώσεις για το κράτος.

Οι βασικότερες συμβάσεις που εξασφαλίζουν σε διεθνές επίπεδο το δικαίωμα των παιδιών στην εκπαίδευση είναι οι εξής:

- **Διεθνής Σύμβαση για τα Δικαιώματα του Παιδιού**

Η Διεθνής Σύμβαση για τα Δικαιώματα του παιδιού, που είναι και το σημαντικότερο νομικό κείμενο το οποίο κατοχυρώνει το δικαίωμα στην εκπαίδευση των παιδιών, υπογράφηκε στη Νέα Υόρκη στις 26 Ιανουαρίου του 1990 και κατόπιν κυρώθηκε από την Ελλάδα με τον νόμο 2101/92, ΦΕΚ 192 τ. Α'. Στην εν λόγω σύμβαση, η οποία

⁸⁰ Τζωρτζοπούλου, Μ., & Κοτζαμάνη, Α., Η εκπαίδευση των αλλοδαπών μαθητών. Διερεύνηση των προβλημάτων και της προοπτικής επίλυσής τους, Εθνικό Κέντρο Κοινωνικών Ερευνών, 2008, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.ekke.gr/publications/wp/wp19.pdf> , [προσπελάστηκε στις 23 Φεβρουαρίου 2017], σελ. 24.

εγκρίθηκε ομόφωνα από όλα τα μέλη του ΟΗΕ στις 20/11/1989, καθορίζεται αρχικά ως παιδί κάθε άνθρωπος μικρότερος των 18 ετών.

Στο άρθρο 2 ορίζεται ότι: *«Τα Συμβαλλόμενα Κράτη υποχρεούνται να σέβονται τα δικαιώματα που αναφέρονται στην παρούσα Σύμβαση και να τα εγγυώνται σε κάθε παιδί που υπάγεται στη δικαιοδοσία τους, χωρίς καμιά διάκριση φυλής, χρώματος, φύλου, γλώσσας, θρησκείας, πολιτικών ή άλλων πεποιθήσεων του παιδιού ή των γονέων του ή των νόμιμων εκπροσώπων του ή της εθνικής, εθνικιστικής ή κοινωνικής καταγωγής τους, της περιουσιακής τους κατάστασης, της ανικανότητάς τους, της γέννησής τους ή οποιασδήποτε άλλης κατάστασης».*

Επιπλέον, Στο άρθρο 2 παράγραφος 2 ορίζεται ρητά ότι κάθε παιδί θα πρέπει να προστατεύεται *«έναντι κάθε μορφής διάκρισης ή κύρωσης, βασισμένης στη νομική κατάσταση, στις δραστηριότητες, στις εκφρασμένες απόψεις ή στις πεποιθήσεις των γονέων του, των νόμιμων εκπροσώπων του ή των μελών της οικογένειάς του.»* Αναφορικά με την εκπαίδευση, ορίζεται ότι κάθε παιδί έχει δικαίωμα στην εκπαίδευση, η οποία αναγνωρίζεται ως ένα διεθνές ανθρώπινο δικαίωμα ζωτικής σημασίας για την ανάπτυξη των παιδιών.⁸¹

- **Σύμβαση για το καθεστώς των προσφύγων/ Σύμβαση της Γενεύης του 1951 άρθρο 22 (ν. 3989/1959) ΦΕΚ Α 201⁸²**

Στις 28 Ιουλίου του 1951, η Γενική Συνέλευση των Ηνωμένων Εθνών ενέκρινε τη σύμβαση για το «Καθεστώς των Προσφύγων», γνωστή και ως «Σύμβαση της Γενεύης του 1951», σύμφωνα με το άρθρο 22 της οποίας, κάθε χώρα ασύλου υποχρεώνεται να παρέχει εκπαίδευση στους πρόσφυγες, τουλάχιστον τη στοιχειώδη, ενώ θα πρέπει να τους μεταχειρίζεται ευνοϊκά, όσον αφορά, τόσο στην πρόσβασή τους στα εκπαιδευτικά ιδρύματα, όσο και στην αναγνώριση αλλοδαπών πιστοποιητικών, διπλωμάτων και πανεπιστημιακών τίτλων, στην καταβολή μειωμένων διδάκτρων και στη χορήγηση υποτροφιών.

Επιπλέον, τα κάτωθι νομικά κείμενα κατοχυρώνουν νομικά το δικαίωμα των προσφύγων στην εκπαίδευση:

⁸¹ Βλ. όλο το κείμενο της σύμβασης στο <https://www.unicef.gr/to-πλήρες-κείμενο-της-σύμβασης-για-τα-δικαιώματα-του-παιδιού/a4-366-22>

⁸² Βλ. όλο το κείμενο της σύμβασης στο <https://www.unric.org/html/greek/pdf/refugeeconvention-GRE.pdf>

- Σύμβαση κατά των διακρίσεων στην εκπαίδευση (1960)⁸³
- Διακήρυξη για τα δικαιώματα προσώπων που ανήκουν σε εθνικές ή εθνοτικές, θρησκευτικές και γλωσσικές μειονότητες, άρθρο 4 (1992)⁸⁴
- Διεθνής Σύμβαση για την Εξάλειψη κάθε μορφής Φυλετικών Διακρίσεων (1965)⁸⁵
- Διεθνές σύμφωνο για τα οικονομικά, κοινωνικά και πολιτιστικά δικαιώματα, άρθρο 13 (1966)⁸⁶
- Οικουμενική διακήρυξη των δικαιωμάτων του ανθρώπου, άρθρο 26 (1948)⁸⁷

Όσον αφορά στην ευρωπαϊκή νομοθεσία, οι βασικότερες συμβάσεις που εξασφαλίζουν σε ευρωπαϊκό επίπεδο το δικαίωμα των παιδιών στην εκπαίδευση είναι οι εξής:

- Χάρτης θεμελιωδών δικαιωμάτων της Ευρωπαϊκής Ένωσης, άρθρο 14 παράγραφος 1,2 ⁸⁸
- Πρώτο πρόσθετο πρωτόκολλο της Ευρωπαϊκής σύμβασης των δικαιωμάτων του ανθρώπου, άρθρο 2 παρ. 1 (1952)⁸⁹

Σε γενικές γραμμές, η εκπαιδευτική πολιτική της Ευρωπαϊκής Ένωσης κινείται γύρω από το πλαίσιο του δικαιώματος της διαφορετικότητας, της ίσης μεταχείρισης και των ίσων δικαιωμάτων. Παρόλο που όλα τα κράτη- μέλη της Ε.Ε. είναι υποχρεωμένα να

⁸³ Βλ. όλο το κείμενο της σύμβασης στο http://6dim-diap-elfth.thess.sch.gr/Greek/Diapolitismiki_Ekpaidefsi/DiethneisSymvaseis_Diakirixeis/Unesco1960.htm

⁸⁴ Βλ. όλο το κείμενο της διακήρυξης στο http://6dim-diap-elfth.thess.sch.gr/Greek/Diapolitismiki_Ekpaidefsi/DiethneisSymvaseis_Diakirixeis/OHE_18_12_1992.htm

⁸⁵ Βλ. όλο το κείμενο της σύμβασης στο <http://www.refworld.org/cgi-bin/texis/vtx/rwmain/opendocpdf.pdf?reldoc=y&docid=4c2aff2b2>

⁸⁶ Βλ. όλο το κείμενο του νόμου στο <https://eclass.duth.gr/modules/document/file.php/KOM01151/ΔΙΕΘΝΕΣ%20ΣΥΜΦΩΝΟ%20ΓΙΑ%20ΤΑ%20ΟΙΚΟΝΟΜΙΚΑ.pdf>

⁸⁷ Βλ. όλο το κείμενο της διακήρυξης στο http://www.ohchr.org/EN/UDHR/Documents/UDHR_Translations/grk.pdf

⁸⁸ Άρθρο 14: Δικαίωμα εκπαίδευσης 1. Κάθε πρόσωπο έχει δικαίωμα στην εκπαίδευση και στην πρόσβαση στην επαγγελματική και συνεχή κατάρτιση. 2. Το δικαίωμα αυτό περιλαμβάνει την ευχέρεια δωρεάν παρακολούθησης της υποχρεωτικής εκπαίδευσης. 3. Η ελευθερία ίδρυσης εκπαιδευτικών ιδρυμάτων με σεβασμό των δημοκρατικών αρχών καθώς και το δικαίωμα των γονέων να εξασφαλίζουν την εκπαίδευση και τη μόρφωση των τέκνων τους σύμφωνα με τις θρησκευτικές, φιλοσοφικές και παιδαγωγικές πεποιθήσεις τους, γίνονται σεβαστά σύμφωνα με τις εθνικές νομοθεσίες που διέπουν την άσκηση τους.

⁸⁹ Βλ. όλο το κείμενο της σύμβασης στο http://www.echr.coe.int/Documents/Convention_ELL.pdf

κινούνται γύρω από το εν λόγω πλαίσιο, ως προς την ένταξη των αλλοδαπών παιδιών στην εκπαίδευση, διακρίνονται σημαντικά μεταξύ τους, γεγονός που οφείλεται σε πολλούς παράγοντες. Καταρχάς, το κάθε κράτος στην παρούσα φάση αντιμετωπίζει την προσφυγική κρίση σε διαφορετική έκταση. Επίσης, άλλες χώρες έχουν μεγάλη εμπειρία στον τομέα της πολιτικής για την ένταξη αυτών των παιδιών, ενώ άλλες που δεν έχουν βιώσει ξανά ένα τέτοιο φαινόμενο- τουλάχιστον όχι σε τέτοιο βαθμό- αναζητούν τώρα τις απαραίτητες πολιτικές και εκπαιδευτικές μεταρρυθμίσεις. Σημαντικός παράγοντας είναι και το γεγονός ότι η κάθε μία χώρα, με διαφορετική ιστορία και πολιτισμό, ασκεί πολιτική υπό διαφορετικές κοινωνικές, αλλά και οικονομικές συνθήκες. Για τους λόγους αυτούς, η Ευρώπη καλείται να διαχειριστεί και να αντιμετωπίσει τα προβλήματα που προκύπτουν εξαιτίας της διαφορετικότητας της κάθε μία χώρας.⁹⁰

Τα άρθρα των διεθνών και ευρωπαϊκών συμβάσεων αποτελούν δεσμευτική νομοθεσία τόσο για την Ελλάδα, όσο και για τη Γερμανία. Συγκεκριμένα, σύμφωνα με το άρθρο 59 παράγραφο 2 του Γερμανικού Συντάγματος⁹¹ αποτελούν ένα δεσμευτικό τμήμα της Γερμανικής νομοθεσίας με τη μορφή ομοσπονδιακού νόμου. Επιπλέον, σύμφωνα με το άρθρο 20 παράγραφος 3 του Γερμανικού Συντάγματος θα πρέπει οι κανόνες των συμβάσεων που αφορούν στα ανθρώπινα δικαιώματα να τηρούνται ως ομοσπονδιακοί νόμοι από όλα τα κρατικά όργανα της ομοσπονδίας, ενώ σύμφωνα με το άρθρο 31, σε περίπτωση σύγκρουσης υπερισχύουν ως ομοσπονδιακοί νόμοι από τους νόμους των κρατιδίων. Στην Γερμανία υπεύθυνα για την εκπαίδευση είναι σχεδόν κατά αποκλειστικότητα τα ομόσπονδα κρατίδια. Έτσι τα τοπικά κοινοβούλια έχουν την υποχρέωση να παίρνουν τις απαραίτητες αποφάσεις για την εφαρμογή των διεθνών υποχρεώσεων.⁹²

Αντίστοιχα στην Ελλάδα, το άρθρο 28 παρ. 1 του Συντάγματος⁹³ εκθέτει τα εξής: «Οι γενικά παραδεγμένοι κανόνες του διεθνούς δικαίου και οι διεθνείς συμβάσεις, από την επικύρωσή τους με νόμο και τη θέση τους σε ισχύ σύμφωνα με τους όρους καθεμιάς, αποτελούν αναπόσπαστο μέρος του εσωτερικού ελληνικού δικαίου και υπερισχύουν από

⁹⁰ Τζωρτζοπούλου, Μ., & Κοτζαμάνη, Α., Η εκπαίδευση των αλλοδαπών μαθητών, ό.π. παραπομπή 80, σελ. 26.

⁹¹ Grundgesetz für die Bundesrepublik Deutschland, ό.π. παραπομπή 79.

⁹² Niendorf, M. & Reitz, S., Das Menschenrecht auf Bildung im deutschen Schulsystem

Was zum Abbau von Diskriminierung notwendig ist, Deutsches Institut für Menschenrechte, 2016, σελ. 17.

⁹³ Σύνταγμα της Ελλάδας, ό.π. παραπομπή 72.

κάθε άλλη αντίθετη διάταξη νόμου. Η εφαρμογή των κανόνων του διεθνούς δικαίου και των διεθνών συμβάσεων στους αλλοδαπούς τελεί πάντοτε υπό τον όρο της αμοιβαιότητας».

Συνοψίζοντας, σύμφωνα με το ελληνικό Σύνταγμα (άρθρο 16), η ανάπτυξη της διδασκαλίας αποτελεί υποχρέωση του Κράτους (παρ. 1), η παιδεία αποτελεί βασική αποστολή του Κράτους (παρ. 2) και όλοι οι έλληνες έχουν δικαίωμα δωρεάν παιδείας (παρ. 4). Αν και το Σύνταγμα, είναι δομημένο και προσανατολισμένο σε δικαιώματα και υποχρεώσεις Ελλήνων πολιτών, εντούτοις γίνεται δεκτό ότι, μέσω των διεθνών συμβάσεων και λοιπών διεθνών κειμένων, αλλά και των αντίστοιχων ευρωπαϊκών που η Ελλάδα έχει αναγνωρίσει και κυρώσει με νόμο, δεσμεύεται από αυτά και, κατά συνέπεια, το δικαίωμα στην παιδεία επεκτείνεται και καλύπτει και όλους όσους πολίτες της αλλοδαπής ζουν και δραστηριοποιούνται εντός της ελληνικής επικράτειας. Με όμοιο τρόπο δεσμεύεται και η Γερμανία απέναντι στα εν λόγω νομικά κείμενα, τα οποία είναι δεσμευτικά για τη χώρα και γι'αυτό εναρμονίζει τη νομοθεσία της με αυτά.⁹⁴ Παραδείγματος χάρη, αναφορικά με την παράγραφο 1 του άρθρου 28 της Συνθήκης για τα Δικαιώματα του παιδιού, που προβλέπει, μεταξύ άλλων, τη θέσπιση της δωρεάν εκπαίδευσης και της προσφοράς χρηματικής βοήθειας σε περίπτωση ανάγκης, το γερμανικό εκπαιδευτικό σύστημα δίνει τη δυνατότητα στους μαθητές με το πέρας της υποχρεωτικής εκπαίδευσης να έχουν αξίωση σε κρατική οικονομική υποστήριξη με τη μορφή επιδόματος.

Σύμφωνα με τον Pierre Bourdieu, *«για τα άτομα που προέρχονται από τα λιγότερο προνομιούχα στρώματα, το σχολείο παραμένει η μία και μοναδική οδός πρόσβασης στην κουλτούρα [...]»*.⁹⁵ Ο αποκλεισμός των παιδιών των προσφύγων από την εκπαίδευση αποτελεί κατάφωρη παραβίαση των ανθρωπίνων δικαιωμάτων. Η σημασία που έχει η εκπαίδευση για την ανάπτυξη των παιδιών είναι εξαιρετικά μεγάλη και αυτό ακριβώς την έχει καταστήσει ένα διεθνώς αναγνωρισμένο δικαίωμα. Τα κράτη δεσμεύονται από συνθήκες και συμβάσεις που έχουν υπογράψει, ώστε να εκπληρώσουν τις υποχρεώσεις τους απέναντι στα παιδιά και να τους παρέχουν πρόσβαση την εκπαίδευση. Οι εν λόγω

⁹⁴ Bax. M., *Wie wird das Recht auf Bildung gesetzlich geregelt?*, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.bildungsexperten.net/wissen/welche-gesetze-gibt-es-zur-bildung/> , [προσπελάστηκε στις 11 Μαρτίου 2017].

⁹⁵ Bourdieu, P., *Οι Κληρονόμοι. Οι φοιτητές και η κουλτούρα*, μετάφραση: Ν. Παναγιωτόπουλος - Μ. Βιδάλη, Ινστιτούτο του βιβλίου – Καρδαμίτσας, Αθήνα, 1966, σελ.73.

συμβάσεις δεν εξαρτώνται από την εκάστοτε μεταναστευτική πολιτική του κάθε κράτους και έχουν υπέρτατη τυπική νομοθετική ισχύ. Ούτε η προσφυγιά, ούτε κανένας άλλος λόγος δεν θα πρέπει να αποτελεί αιτία για τη μη εκπλήρωση αυτών των υποχρεώσεων. Και ενώ οι ρυθμίσεις υπάρχουν και μάλιστα σε διεθνές επίπεδο, μόνο ένα μικρό ποσοστό των παιδιών αυτών φαίνεται να έχει πρόσβαση στην εκπαίδευση.⁹⁶

Επομένως, όλα τα παιδιά που ζουν, τόσο στην Ελλάδα, όσο και στη Γερμανία έχουν δικαίωμα στην εκπαίδευση, ανεξαρτήτως από το αν οι γονείς τους είναι πρόσφυγες ή μετανάστες, αλλά και από το καθεστώς παραμονής τους, καθώς το ατομικό αυτό δικαίωμα είναι αναπόσπαστο. Το κάθε σχολείο οφείλει να είναι ελεύθερο, δημοκρατικό και ανοιχτό στη διαφορετικότητα. Μακριά από προκαταλήψεις και στερεότυπα θα πρέπει να προτάσσει το δικαίωμα του κάθε παιδιού να μορφωθεί ισότιμα και χωρίς να αφομοιωθεί πολιτισμικά, ώστε να διαμορφώσει ελεύθερους και δημοκρατικούς πολίτες.

1.9 Η ΠΡΑΓΜΑΤΙΚΟΤΗΤΑ ΓΙΑ ΤΗΝ ΕΚΠΑΙΔΕΥΣΗ ΤΩΝ ΠΡΟΣΦΥΓΩΝ

Μόλις τα μισά από τα παιδιά των προσφύγων έχουν πρόσβαση στην πρωτοβάθμια εκπαίδευση και το ένα τέταρτο στη δευτεροβάθμια.⁹⁷ Παγκοσμίως, από τα 6 εκατομμύρια παιδιών προσφύγων που βρίσκονται υπό την εντολή της Ύπατης Αρμοστείας των Ηνωμένων Εθνών για τους Πρόσφυγες και σε ηλικία πρωτοβάθμιας και δευτεροβάθμιας εκπαίδευσης, τα 3,7 εκατομμύρια δεν πηγαίνουν σχολείο.⁹⁸ Σύμφωνα με έρευνες, τα προσφυγόπουλα έχουν πέντε φορές περισσότερες πιθανότητες να μην πάνε στο σχολείο από ότι τα υπόλοιπα παιδιά. Το 2015, δεν είχαν πρόσβαση στην πρωτοβάθμια εκπαίδευση 1,75 εκατομμύρια παιδιά και στη δευτεροβάθμια 1,95. Στην τριτοβάθμια εκπαίδευση είχε πρόσβαση μόλις το 1% των προσφύγων, ενώ για τους μη εκτοπισμένους παγκοσμίως το ποσοστό αυτό ανέρχεται στο 34%. Η UNHCR εκτιμά ότι οι πρόσφυγες χάνουν κατά μέσο όρο τρία με τέσσερα

⁹⁶ Τζωρτζοπούλου, Μ., & Κοτζαμάνη, Α., Η εκπαίδευση των αλλοδαπών μαθητών, ό.π. παραπομπή 80, σελ.25.

⁹⁷ European Commission. Regional refugee and migrant response plan for Europe Eastern Mediterranean and Western Balkans route, ό.π. παραπομπή 29, σελ. 59.

⁹⁸ European Commission. Regional refugee and migrant response plan for Europe Eastern Mediterranean and Western Balkans route, ό.π. παραπομπή 29, σελ. 56.

χρόνια από τη σχολική τους εκπαίδευση εξαιτίας του αναγκαίου εκτοπισμού τους. Σε παγκόσμια κλίμακα, για κάθε δέκα αγόρια πρόσφυγες στην πρωτοβάθμια εκπαίδευση υπάρχουν λιγότερα από οκτώ κορίτσια. Στη δευτεροβάθμια εκπαίδευση η εικόνα είναι ακόμα χειρότερη, με λιγότερα από επτά κορίτσια για κάθε δέκα αγόρια προσφυγόπουλα. Ιδιαίτέρως σε πληθυσμούς όπου υπάρχουν έντονοι πολιτισμικοί φραγμοί στην εκπαίδευση των κοριτσιών, η κατάσταση είναι ακόμα χειρότερη.⁹⁹

Η ποιοτική εκπαίδευση παρέχει στα παιδιά έναν ασφαλή τόπο και μπορεί να μειώσει τον γάμο σε παιδική ηλικία, την παιδική εργασία, την εγκυμοσύνη στους εφήβους και την παροχή εργασίας με καταχρηστικούς όρους και σε επικίνδυνες συνθήκες. Τους δίνει επιπλέον τη δυνατότητα να κάνουν φίλους, να βρουν καθοδηγητές και τους εφοδιάζει με δεξιότητες, προκειμένου να έχουν εμπιστοσύνη στον εαυτό τους, κριτική σκέψη, να μπορούν να επιλύσουν προβλήματα και να δουλέψουν σε ομάδες. Επίσης, βελτιώνει τις εργασιακές προοπτικές τους και τονώνει την αυτοπεποίθηση και την αυτοεκτίμησή τους. Χωρίς την προστασία που μπορεί να τους παρέχει η εκπαίδευση, οι έφηβοι πρόσφυγες μπορούν να γίνουν εξαιρετικά ευάλωτοι. Ακόμα και αν δεν οδηγηθούν στην παιδική εργασία, μπορεί να νιώσουν πλήξη, ότι είναι αβοήθητοι, έρμαιοι και αποτυχημένοι. Αυτό μπορεί να τους καταστήσει εύκολη λεία για στρατολόγηση από ένοπλες δυνάμεις, ενώ για τα κορίτσια εγκυμονούν επιπλέον οι κίνδυνοι του γάμου σε παιδική ηλικία, της εφηβικής εγκυμοσύνης, της κράτησης για οικιακή εργασία ως υπηρέτριες ή της σεξουαλικής εκμετάλλευσης.

Οι πρόσφυγες θα πρέπει να εντάσσονται στο εκάστοτε εθνικό εκπαιδευτικό σύστημα και να ακολουθούν εθνικά προγράμματα σπουδών και όχι να ακολουθούν απλώς κάποιους κύκλους μαθημάτων, τους οποίους το κράτος υποδοχής δεν μπορεί ούτε να τους ελέγξει, ούτε να εγγυηθεί γι'αυτούς. Επειδή όμως πολλά παιδιά και έφηβοι χάνουν την εκπαίδευσή τους εξαιτίας της φτώχειας, της περιθωριοποίησης, των συρράξεων και της κρίσης, κρίνεται απαραίτητη στην περίπτωση των προσφύγων η παροχή πιο ευέλικτων μορφών εκπαίδευσης. Η ταχύρρυθμη εκπαίδευση θα πρέπει να περιλαμβάνει ευέλικτα και κατάλληλα για την εκάστοτε ηλικία προγράμματα που να στοχεύουν σε μειονεκτούσες ομάδες και σε και παιδιά και εφήβους που σταμάτησαν το σχολείο και διέκοψαν την εκπαίδευσή τους.

⁹⁹ Ο.π. , παραπομπή 98.

1.9.1 Η ΕΚΠΑΙΔΕΥΣΗ ΤΩΝ ΠΡΟΣΦΥΓΩΝ ΣΤΗΝ ΕΛΛΑΔΑ

Στην Ελλάδα αρχικά, η εκπαίδευση δεν αποτέλεσε προτεραιότητα για την ελληνική κυβέρνηση, η οποία προσπάθησε να παράσχει τα βασικά στους 60.000 αιτούντες άσυλο που ήταν παγιδευμένοι στη χώρα. Έτσι, μέχρι το Μάιο του 2016 υπολογίζεται ότι τα παιδιά των προσφύγων έμειναν κατά μέσο όρο εκτός σχολείου για 1,5 χρόνο, σύμφωνα με στοιχεία της Μ.Κ.Ο. «Save the Children». ¹⁰⁰

Μπορεί η Ελλάδα να αποτελεί χώρα υποδοχής και όχι χώρα προορισμού για τους πρόσφυγες και η παραμονή τους εδώ να είναι σε γενικές γραμμές προσωρινή, όμως, με τα προγράμματα μετεγκατάστασης να προχωρούν με πολύ αργό ρυθμό, επικρατεί πλέον η αίσθηση ότι δεκάδες χιλιάδες πρόσφυγες, που θέλουν να συνεχίσουν το ταξίδι τους στην Ευρώπη, θα παραμείνουν στην Ελλάδα για αρκετό καιρό, ίσως και για κάποια χρόνια.

Από τα 27.000 παιδιά που δέχτηκε η Ελλάδα, τουλάχιστον τα 18.000 υπολογίζεται ότι είναι παιδιά σχολικής ηλικίας. Μέχρι και τον Σεπτέμβριο του 2016, ένας μικρός αριθμός παιδιών προσφύγων είχε ξεκινήσει το σχολείο στην Ελλάδα, ενώ κρίθηκαν αναγκαίες οι προσπάθειες, προκειμένου παιδιά ακόμα και των πιο απομακρυσμένων περιοχών, να έχουν πρόσβαση στην εκπαίδευση. ¹⁰¹ Έτσι, ένα μέρος από τα παιδιά αυτά εντάχθηκαν από τον Σεπτέμβριο του 2016 στις λεγόμενες «τάξεις υποδοχής», όπου ξεχωριστά από τους Έλληνες μαθητές διδάσκονται τέσσερις ώρες τη μέρα κάθε απόγευμα Ελληνικά, Αγγλικά, Μαθηματικά και Πληροφορική. Στόχος της ελληνικής κυβέρνησης είναι να προετοιμάσει τα παιδιά κατάλληλα, ώστε από την επόμενη σχολική χρονιά να μπορέσουν να ενσωματωθούν πλήρως στο ελληνικό εκπαιδευτικό σύστημα και να ενταχθούν στις κανονικές τάξεις.

Σύμφωνα με το υπουργείο Παιδείας της Ελλάδας, 2.500 προσφυγόπουλα ξεκίνησαν από την αρχή της σχολικής χρονιάς 2016-2017 να παρακολουθούν μαθήματα σε

¹⁰⁰ <http://www.savethechildren.org.uk/2016-05/child-refugees-greece-out-school-15-years>

¹⁰¹ European Commission. Regional refugee and migrant response plan for Europe Eastern Mediterranean and Western Balkans route, ό.π. παραπομπή 29, σελ. 54.

δημόσια σχολεία της ηπειρωτικής χώρας, ενώ στη Λέρο διαμορφώθηκε ένας ανεπίσημος εκπαιδευτικός χώρος, με χρηματοδότηση της Ευρωπαϊκής Επιτροπής, ο οποίος δίνει τη δυνατότητα σε 150 περίπου παιδιά ηλικίας από έξι έως δεκαοχτώ ετών να λάβουν εκπαίδευση μέχρι την κανονική ένταξή τους στο ελληνικό εκπαιδευτικό σύστημα.¹⁰²

Σύμφωνα με στοιχεία του υπουργείου Παιδείας, μέχρι το Φεβρουάριο του 2017, υπολογίζεται ότι ο αριθμός των παιδιών προσφύγων ηλικίας έως 15 ετών που ξεκίνησαν το σχολείο στην Ελλάδα έφτασε τις 10.000,¹⁰³ εκ των οποίων τα 7.000 είναι από κέντρα φιλοξενίας και τα υπόλοιπα 3.000 μένουν μαζί με τις οικογένειές τους σε ξενοδοχεία και διαμερίσματα νοικιασμένα από την Ύπατη Αρμοστεία του ΟΗΕ.

1.9.2 Η ΕΚΠΑΙΔΕΥΣΗ ΤΩΝ ΠΡΟΣΦΥΓΩΝ ΣΤΗ ΓΕΡΜΑΝΙΑ

Ο Ειδικός Εισηγητής του ΟΗΕ για το δικαίωμα στην εκπαίδευση Vernor Muñoz επισκέφθηκε το 2006 τη Γερμανία για να εξετάσει το γερμανικό εκπαιδευτικό σύστημα και να ενημερωθεί για το επίπεδο εφαρμογής του δικαιώματος στην εκπαίδευση. Ένα χρόνο αργότερα δημοσίευσε την αναφορά του, στην οποία χαρακτήριζε το γερμανικό εκπαιδευτικό σύστημα ως «επιλεκτικό», καθώς αδικούσε συγκεκριμένες ομάδες και οδηγούσε στην de facto διάκρισή τους. Σύμφωνα με τον Muñoz για να εφαρμοστεί στη Γερμανία το δικαίωμα στην εκπαίδευση εξίσου για όλα τα παιδιά, θα πρέπει να παραμεριστούν οι κοινωνικές και οικονομικές διακρίσεις.¹⁰⁴

«Τα συμφέροντα και οι ανάγκες των παιδιών που προέρχονται από εθνοτικές μειονότητες υπολογίζονται σε πολύ μικρό βαθμό και αυτό γιατί η Γερμανία αρνείται να συνειδητοποιήσει ότι εδώ και πολλές δεκαετίες αποτελεί μία μεταναστευτική κοινωνία.»

¹⁰² European Commission. Regional refugee and migrant response plan for Europe Eastern Mediterranean and Western Balkans route, ό.π. παραπομπή 29, σελ. 36

¹⁰³ Λάκασας Α., Επαγγελματική εκπαίδευση νέων προσφύγων, *Καθημερινή*, 2017, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: https://www.minedu.gov.gr/publications/docs2017/site_swsto_pemptpi.pdf, [προσπελάστηκε στις 28 Φεβρουαρίου 2017], σελ. 24.

¹⁰⁴ Muñoz, V., Report of the Special Rapporteur on the right to education, 2007, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: http://www.netzwerk-bildungsfreiheit.de/pdf/Munoz_Mission_on_Germany.pdf, [προσπελάστηκε στις 21 Φεβρουαρίου 2017], σελ. 22.

αναφέρει χαρακτηριστικά ο Muñoz, καθώς επίσης ότι «σε καμία άλλη βιομηχανική χώρα της Δύσης δεν εξαρτάται σε τέτοιο μεγάλο βαθμό η επιτυχία στην εκπαίδευση ενός παιδιού από την κοινωνική θέση της οικογένειά του, όσο στη Γερμανία. Αυτή η κοινωνική επιλεκτικότητα επηρεάζει ιδιαίτερος τα παιδιά των μεταναστών και των προσφύγων, καθώς κατά μέσο όρο μεγαλώνουν σε κοινωνικά μειονεκτούσες οικογένειες.»

Για τα παιδιά των προσφύγων ιδιαίτερος στη Γερμανία δεν εξασφαλίζεται σε μεγάλο βαθμό το δικαίωμα στην εκπαίδευση. Συγκεκριμένα, τα παιδιά που έχουν προσωρινή ή καμία άδεια παραμονής, έχουν περιορισμένη ή και καθόλου πρόσβαση στην εκπαίδευση, καθώς δεν συμπεριλαμβάνονται στην υποχρεωτική εκπαίδευση. Γι' αυτήν της την αντιμετώπιση προς τα παιδιά των προσφύγων έχει δεχτεί μαζική κριτική η γερμανική κυβέρνηση, όπως για παράδειγμα από την Επιτροπή του ΟΗΕ για τα Δικαιώματα των Παιδιών.¹⁰⁵ Μέχρι τους πρώτους μήνες του 2016, τα παιδιά, που παρέμεναν στα κέντρα φιλοξενίας των προσφύγων στη Γερμανία, δεν είχαν πρόσβαση στο τακτικό εκπαιδευτικό σύστημα της χώρας.¹⁰⁶

Παρόλα αυτά, η ενσωμάτωση των παιδιών στα σχολεία έχει βελτιωθεί κατά πολύ από τον Σεπτέμβριο του 2016. Η Γερμανία, όπως και η Αυστρία και η Σουηδία που αποτελούν τις σημαντικότερες χώρες προορισμού των προσφύγων, προσφέρουν προπαρασκευαστική εκπαίδευση πριν την ενσωμάτωση των παιδιών στις κανονικές τάξεις. Ιδιαίτερος η Σουηδία επιτυγχάνει την ένταξη των παιδιών στο σχολικό σύστημα μέσα σε ένα μήνα.¹⁰⁷

Στη Γερμανία, το εάν και το πότε το κάθε παιδί θα αποκτήσει πρόσβαση στην εκπαίδευση, εξαρτάται από το κρατίδιο στο οποίο έχει εγκατασταθεί. Σε μερικά κρατίδια τα παιδιά φαίνεται να ξεκινούν σε σύντομο χρονικό διάστημα το σχολείο, ενώ σε άλλα θα πρέπει πρώτα η οικογένειά τους να φύγει από το κέντρο υποδοχής προσφύγων και να εγκατασταθεί κάπου μόνιμα, διαφορετικά είναι αδύνατον για το παιδί να ξεκινήσει να παρακολουθεί μαθήματα σε ένα σχολείο.

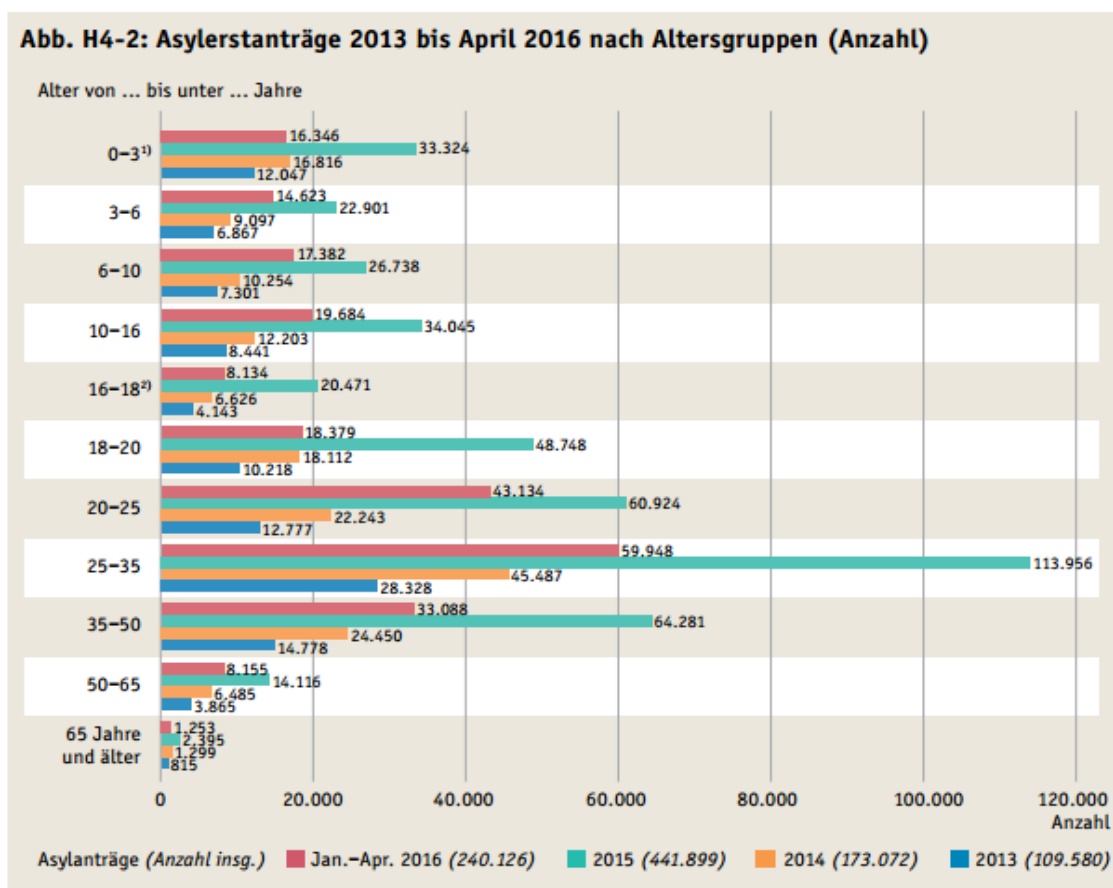
¹⁰⁵ Das Menschenrecht auf Bildung, 2013, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.bpb.de/gesellschaft/kultur/zukunft-bildung/156819/menschenrecht?p=all>, [προσπελάστηκε στις 17 Φεβρουαρίου 2017].

¹⁰⁶ European Commission. Regional refugee and migrant response plan for Europe Eastern Mediterranean and Western Balkans route, ό.π. παραπομπή 29, σελ. 47.

¹⁰⁷ European Commission. Regional refugee and migrant response plan for Europe Eastern Mediterranean and Western Balkans route, ό.π. παραπομπή 29, σελ. 40.

Σε μερικά δημοτικά σχολεία τα παιδιά των προσφύγων εντάσσονται απευθείας στις κανονικές τάξεις μαζί με τα υπόλοιπα παιδιά, ενώ στις μεγαλύτερες ηλικίες κρίνεται απαραίτητη η ένταξη σε μία τάξη προετοιμασίας για ένα σχολικό έτος, προκειμένου να μπορέσουν τα παιδιά να μάθουν όσο το δυνατόν καλύτερα τη γλώσσα και στη συνέχεια να μπορούν να παρακολουθήσουν το μάθημα στις κανονικές τάξεις.¹⁰⁸ Δεν υπάρχουν επίσημα νούμερα για το πόσα παιδιά προσφύγων έχουν πάει σε ποια σχολεία και οι περισσότεροι αριθμοί βασίζονται σε εκτιμήσεις. Περίπου 300.000 παιδιά φαίνεται πως εντάχθηκαν στα σχολεία της Γερμανίας, χωρίς να έχει υποβληθεί για όλα αίτηση χορήγησης ασύλου, γεγονός που φαίνεται από τον κατά πολύ μικρότερο αριθμό των παιδιών που έχουν υποβάλει αίτηση για άσυλο και είναι υποχρεωμένα να φοιτήσουν σε σχολείο, τον οποίο γνωστοποίησε η Ομοσπονδιακή Υπηρεσία Μετανάστευσης και Προσφύγων (εικόνα 6).

¹⁰⁸ Fakten zur Flüchtlingskrise, 2017, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.spiegel.de/politik/deutschland/fluechtlinge-und-einwanderer-die-wichtigsten-fakten-a-1030320.html#sponfakt=10>, [προσπελάστηκε στις 22 Φεβρουαρίου 2017].



Εικόνα 6. Αιτήσεις ασύλου (συνολικός αριθμός) στη Γερμανία από τον Ιανουάριο έως τον Απρίλιο 2016¹⁰⁹

Από έρευνα που πραγματοποίησε το Γερμανικό Ινστιτούτο για τα Ανθρώπινα Δικαιώματα τον Ιανουάριο του 2016, προέκυψε ότι μόνο σε πέντε από τα δεκαέξι γερμανικά κρατίδια ήταν υποχρεωμένα τα παιδιά των προσφύγων έως τότε να πηγαίνουν άμεσα στο σχολείο με την άφιξή τους,¹¹⁰ καθώς ακόμα και οι ειδικές διατάξεις για την υποχρεωτική φοίτηση διαφέρουν στα επιμέρους ομόσπονδα κρατίδια.¹¹¹

Συγκεκριμένα, στα κρατίδια Βερολίνο (Berlin), Αμβούργο (Hamburg), Βρέμη (Bremen), Σλέσβιχ-Χόλσταϊν (Schleswig Holstein) και Σάαρλαντ (Saarland) ισχύει για όλα τα

¹⁰⁹ Βλ. <http://www.bildungsbericht.de/de/bildungsberichte-seit-2006/bildungsbericht-2016/pdf-bildungsbericht-2016/bildungsbericht-2016>

¹¹⁰ Welchen Zugang haben geflüchtete Kinder zu Bildung?, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://landkarte-kinderrechte.de/>, [προσπελάστηκε στις 17 Φεβρουαρίου 2017].

¹¹¹ Καλώς ήλθατε στη Γερμανία, Πληροφορίες για μετανάστες, Bundesamt für Migration und Flüchtlinge, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: https://www.bmi.bund.de/SharedDocs/Downloads/DE/Broschueren/Sprachvarianten/Willkommen_in_Deutschland_gr.pdf?__blob=publicationFile, [προσπελάστηκε στις 9 Φεβρουαρίου 2017], σελ. 73.

παιδιά η υποχρεωτική εκπαίδευση ανεξαρτήτως από το καθεστώς παραμονής τους, ενώ η ένταξή τους στο εκπαιδευτικό σύστημα επιτυγχάνεται άμεσα. Μόνο στην περίπτωση της Βρέμης εξαιρούνται τα ασυνόδευτα παιδιά, για τα οποία συχνά χρειάζεται να περάσουν κάποιοι μήνες μέχρι να πάνε στο σχολείο.

Αντιθέτως, στα υπόλοιπα έντεκα κρατίδια, Θουριγγία (Thüringen), Σαξονία-Άνχαλτ (Sachsen-Anhalt), Μεκλεμβούργο-Δυτική Πομερανία (Mecklenburg-Vorpommern), Κάτω Σαξονία (Niedersachsen), Βρανδεμβούργο (Brandenburg), Βόρεια Ρηνανία-Βεστφαλία (Nordrhein-Westfalen), Σαξονία (Sachsen), Έσση (Hessen), Ρηνανία-Παλατινάτο (Rheinland-Pfalz), Βαυαρία (Bayern), Σαξονία-Άνχαλτ (Sachsen-Anhalt) και Βάδη-Βυρτεμβέργη (Baden-Württemberg) απαραίτητη προϋπόθεση για ένταξη στο εκπαιδευτικό σύστημα είναι το να έχουν φύγει τα παιδιά από το κέντρο υποδοχής και να έχουν ενταχθεί σε κάποιο δήμο ή κοινότητα. Σε πολλές περιπτώσεις είναι απαραίτητο τα παιδιά να έχουν συμπληρώσει ένα συγκεκριμένο χρονικό διάστημα παραμονής στο κρατίδιο που συνήθως κυμαίνεται γύρω στους έξι μήνες.

Από τη στιγμή που τα παιδιά εγγραφούν σε ένα σχολείο, εξαρτάται και πάλι από το κρατίδιο στο οποίο ανήκει το σχολείο αυτό για το σε ποια μορφή θα είναι η εκπαίδευση που θα παρακολουθήσουν και ποια θα είναι η προετοιμασία που θα κάνουν για να ενταχθούν στο νέο σχολικό σύστημα. Το ποσοστό των προσφύγων από τάξη σε τάξη διαφέρει σημαντικά. Επίσης δεν υπάρχει καθορισμένη μεταβατική περίοδος από τις τάξεις υποδοχής στις κανονικές τάξεις. Η αλλαγή αυτή πραγματοποιείται βήμα-βήμα και εξαρτάται από τις ικανότητες και την πρόοδο του κάθε παιδιού.

Υπάρχουν περιπτώσεις που τα παιδιά έχουν ενταχθεί σε κανονικές τάξεις σε μαθήματα όπως η Γυμναστική και η Μουσική, αλλά η μετάβαση για μαθήματα όπως είναι τα Μαθηματικά και η Φυσική γίνεται πολύ αργότερα. Ανάλογα σε ποιο Ομόσπονδο Κρατίδιο της χώρας αναφερόμαστε, γίνεται λόγος για «τάξεις προετοιμασίας», «μεταβατικές τάξεις» ή «τάξεις υποδοχής». Σε όλες τις περιπτώσεις όμως ισχύει το ότι προκειμένου να ενταχθούν τα παιδιά των προσφύγων στις κανονικές τάξεις, θα πρέπει πρώτα να μάθουν τη γλώσσα. Το πώς επιτυγχάνεται αυτό, εξαρτάται τόσο από κρατίδιο σε κρατίδιο, όσο και από σχολείο σε σχολείο. Παντού ισχύει όμως ο γενικός κανόνας ότι οι πρόσφυγες θα πρέπει να είναι έτοιμοι για την ένταξή τους στις

κανονικές τάξεις μέσα σε περίπου ένα χρόνο.¹¹² Μπορεί οι Γερμανοί ειδικοί να συμφωνούν στο ότι η ενσωμάτωση στην κοινωνία περνάει μέσα από την εκπαίδευση, οι απόψεις τους όμως σχετικά με την επίτευξη της ενσωμάτωσης δίστανται. Η μία πλευρά υποστηρίζει ότι οι ειδικές τάξεις υποδοχής δεν αποτελούν καλή λύση, διότι τα παιδιά είναι συνεχώς μεταξύ τους και δεν μαθαίνουν γρήγορα τη γερμανική γλώσσα, ούτε ενσωματώνονται στο σχολείο, οπότε και θα πρέπει να εντάσσονται αμέσως σε κανονικές τάξεις, ενώ η άλλη πλευρά λαμβάνει υπόψη την πολιτιστική, θρησκευτική και γεωγραφική ετερογένεια των παιδιών αυτών και υποστηρίζει ότι είναι αναγκαία η παραμονή τους για ένα διάστημα, από ένα έως δύο χρόνια σε ειδικές τάξεις προετοιμασίας.¹¹³ Οι δάσκαλοι δείχνουν να τάσσονται υπέρ της ένταξης των παιδιών στις κανονικές τάξεις, καθώς πιστεύουν ότι μόνο έτσι επιτυγχάνεται ουσιαστικά η ενσωμάτωση των προσφύγων και μαθαίνεται γρηγορότερα η γλώσσα.¹¹⁴

Παρόλο που στο γερμανικό κράτος κατέστη δυνατό να διοριστούν επιπλέον καθηγητές για να καλυφθούν οι νέες ανάγκες σε σύντομο χρονικό διάστημα, η έλλειψη καταρτισμένων εκπαιδευτικών αποτελεί ένα από τα μεγαλύτερα προβλήματα, σύμφωνα με τα ίδια τα σχολεία, τα οποία έχουν ως στόχο τους το να μάθουν τα παιδιά όσο το δυνατόν καλύτερα Γερμανικά για να μπορούν να ανταπεξέλθουν στη ζωή στη Γερμανία.¹¹⁵

Στην πράξη, όπως αναφέρθηκε παραπάνω, τα σχολεία ακολουθούν διαφορετικό σύστημα, ανάλογα με το κρατίδιο στο οποίο υπάγονται. Για παράδειγμα, στο Αμβούργο, τα προσφυγόπουλα επισκέπτονται αρχικά την λεγόμενη βασική τάξη ή μια διεθνή τάξη προετοιμασίας για την εκμάθηση της γλώσσας. Μαθήματα μουσικής, τεχνικών και αθλητισμός γίνονται από κοινού με τους άλλους μαθητές, ώστε να προετοιμάζεται η ένταξή τους. Όταν οι γνώσεις γερμανικών είναι επαρκείς

¹¹² Agarwala, A., Schenk, A., & Spiwak, M., Flüchtlinge im Schulalltag, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.zeit.de/2016/29/integration-fluechtlinge-schule-kinder-jugendliche-deutschunterricht-sprachbarriere-bildungspolitik>, [προσπελάστηκε στις 17 Φεβρουαρίου 2017].

¹¹³ Ο.π.

¹¹⁴ Blaß, K., & Himmelrath, A., Wenn Grundschüler plötzlich von "Kanaken" sprechen, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.spiegel.de/lebenundlernen/schule/fluechtlingkinder-in-der-schule-tipps-fuer-lehrer-a-1100905.html>, [προσπελάστηκε στις 17 Φεβρουαρίου 2017].

¹¹⁵ Agarwala, A., & Hamann, G., Schaffen die das?, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.zeit.de/gesellschaft/2016-07/fluechtlingkinder-schule-fluechtlinge-deutschland>, [προσπελάστηκε στις 17 Φεβρουαρίου 2017].

εντάσσονται σε κανονικές τάξεις.¹¹⁶ Τα παιδιά όμως που δεν μπορούν να γράψουν ή να διαβάσουν στη μητρική τους γλώσσα εντάσσονται στις λεγόμενες «βασικές τάξεις» (Basisklassen).¹¹⁷

Στη Βαυαρία υπάρχουν μεταβατικές τάξεις προετοιμασίας ήδη από το δημοτικό, όπου γίνεται η εκμάθηση της γλώσσας, των αξιών της δυτικής κοινωνίας και διαφόρων άλλων μαθημάτων. Ακολούθως τα παιδιά εντάσσονται σε κανονικές τάξεις, όπου συνεχίζουν να παρακολουθούν ενισχυτικά μαθήματα γερμανικών, τα οποία προορίζονται όμως και για όλους τους υπόλοιπους μαθητές.

Στο Βερολίνο, όπου το ένα στα τρία σχολεία έχει πρόσφυγες, υπάρχουν τάξεις γλωσσικής προετοιμασίας διάρκειας ενός ή δύο ετών. Στο διάστημα αυτό συμμετέχουν σε κανονικές τάξεις στα μαθήματα της μουσικής και του αθλητισμού. Εν συνεχεία εντάσσονται πλήρως σε κανονικές τάξεις. Από φέτος καθιερώνονται σε ορισμένα σχολεία και οι λεγόμενες «τάξεις-γέφυρες», ώστε να γίνεται κατά περίπτωση και σε διαφορετικό χρόνο η ένταξη στις κανονικές τάξεις. Συγκεκριμένα, εντάχθηκαν το 2011 1.378 παιδιά προσφύγων σε 112 τάξεις εκμάθησης Γερμανικών χωρίς να έχουν καθόλου γνώση της γλώσσας. Το σχολικό έτος 2015-2016 ο αριθμός αυτός τετραπλασιάστηκε και έτσι εντάχθηκαν σχεδόν 5.000 νέοι μαθητές σε 431 τάξεις.¹¹⁸

Στη Βάδη-Βυρτεμβέργη, τα παιδιά των προσφύγων, αφού επισκεφθούν τις τάξεις γλωσσικής προετοιμασίας, εντάσσονται σε κανονικές τάξεις. Από τη σχολική χρονιά 2016-2017 δοκιμάζεται ένα πρωτοποριακό σύστημα, το οποίο προβλέπει τεστ δεξιοτήτων και ικανοτήτων πριν ενταχθούν οι μαθητές στο ανάλογο είδος σχολείου.

¹¹⁶ Bildung in Deutschland 2016, Bundesministerium für Bildung und Forschung, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.bildungsbericht.de/de/bildungsberichte-seit-2006/bildungsbericht-2016/pdf-bildungsbericht-2016/bildungsbericht-2016>, [προσπελάστηκε στις 9 Φεβρουαρίου 2017], σελ. 189.

¹¹⁷ Hollenstein, O., Will gelernt sein, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.zeit.de/2016/18/fluechtlingskinder-bildung-schule-integration>, [προσπελάστηκε στις 17 Φεβρουαρίου 2017].

¹¹⁸ Schuler, K., & und Sadigh, P., Ab in die Schule, Alima, Bassam und Chipo!, 2015, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.zeit.de/gesellschaft/schule/2015-09/schule-fluechtlingskinder-jugendliche>, [προσπελάστηκε στις 17 Φεβρουαρίου 2017].

Στα κρατίδια Ζάαρλαντ και Μεκλεμβούργο-Πομερανία τα παιδιά εντάσσονται από την πρώτη μέρα σε κανονικές τάξεις και παράλληλα κάνουν εντατικά μαθήματα Γερμανικών.

Ένα καίριο πρόβλημα που αντιμετωπίζουν, σύμφωνα με δημοσιεύματα του τύπου, οι δάσκαλοι στη Γερμανία κατά την εκπαιδευτική πράξη, είναι το ότι έρχονται πολλές φορές αντιμέτωποι με ξενοφοβικά επεισόδια, τα οποία δεν ξέρουν πώς να αντιμετωπίσουν. Έτσι από τη μία πλευρά βρίσκονται ορισμένα παιδιά που εμφανίζουν ρατσιστική συμπεριφορά και από την άλλη τα παιδιά- πρόσφυγες, πολλά από τα οποία φέρουν έντονα ψυχικά τραύματα, τόσο από τις εμπειρίες τους στη χώρα προέλευσης, όσο και από το δύσκολο ταξίδι που έκαναν. Γύρω στο 20- 30% των παιδιών- προσφύγων πρέπει να υποφέρουν από μετατραυματικό στρες σύμφωνα με τον καθηγητή Ψυχολογίας Frank Neuer του πανεπιστημίου του Bielefeld. Το εν λόγω στρες μπορεί να οδηγήσει σε διαταραχή συγκέντρωσης, επιθετική συμπεριφορά και απόγνωση.¹¹⁹

Γι' αυτό κρίνεται απαραίτητο να εξεταστεί το πώς δέχονται οι Έλληνες και οι Γερμανοί τους πρόσφυγες στην εκπαίδευση και γενικότερα, αλλά και το κατά πόσο τα περιστατικά ρατσισμού και ξενοφοβίας δείχνουν να είναι μεμονωμένα ή γενικευμένα.

1.10 ΡΑΤΣΙΣΜΟΣ

Στην Ελλάδα, η οικονομική κρίση που επηρέασε βαθιά όλη τη χώρα οδήγησε να επιδεινωθεί σε μεγάλο βαθμό η κοινωνική ασφάλεια των προσφύγων που βρίσκονται στο ελληνικό κράτος. Όπως χαρακτηριστικά αναφέρουν δημοσιεύματα του διεθνούς τύπου, ήδη από το 2008 έως το 2012 αυξήθηκαν κατακόρυφα τα εγκλήματα ρατσιστικής βίας που είχαν ως στόχο μετανάστες και πρόσφυγες.¹²⁰ Μόλις τους πρώτους εννέα μήνες του 2012, το Δίκτυο Καταγραφής Ρατσιστικής Βίας στην Ελλάδα

¹¹⁹ Blaß, K., & Himmelrath, A., Wenn Grundschüler plötzlich von "Kanaken" sprechen, ό.π. παραπομπή 103.

¹²⁰ Racist Attacks in Greece Hit 'Alarming' Levels: UNHCR, 2012, , [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.reuters.com/article/2012/10/23/us-greece-violenceidUSBRE89M12R20121023>, [προσπελάστηκε στις 9 Φεβρουαρίου 2017].

κατέγραψε 87 περιστατικά ρατσιστικής βίας, που αφορούσαν κυρίως σε σωματικές επιθέσεις κατά προσφύγων και μεταναστών, οι οποίες ως επί το πλείστον έγιναν σε δημόσιους χώρους. Ο Γερμανός βουλευτής Andrej Hunko, σε δημοσίευσή του στον τύπο τον Νοέμβριο του 2012, δήλωσε χαρακτηριστικά ότι *«η Ευρώπη είναι συνένοχη στο ρατσισμό και στο φασισμό στην Ελλάδα»*.¹²¹

Παρομοίως στη Γερμανία, η έντονη προσφυγή ροή των τελευταίων ετών φαίνεται πως αύξησε τα επεισόδια ρατσισμού και ξενοφοβίας. Παραδείγματος χάρη, αυξήθηκε ο αριθμός των επιθέσεων εναντίον των τζαμιών από το 2011 και μετά. Ενώ από το 2001 μέχρι το 2011 καταγράφονταν κατά μέσο όρο 22 επιθέσεις το χρόνο, ο αριθμός αυτός αυξήθηκε σε 35 και 36 το 2012 και το 2013 αντίστοιχα.¹²²

Η Ευρωπαϊκή Επιτροπή σε αναφορά της, τονίζει την ανάγκη εφαρμογής του πλαισίου σχετικά με τον ρατσισμό και την ξενοφοβία από όλα τα κράτη- μέλη, καθώς και την αντιμετώπιση, τόσο του λεκτικού, όσο και του εγκλήματος μίσους μέσω του ποινικού δικαίου. Παρόλο που η πλειονότητα των κρατών- μελών τιμωρεί την πρόκληση ρατσιστικής και ξενοφοβικής βίας και μίσους, οι νομικές τους διατάξεις δεν ανταποκρίνονται πάντα πλήρως σε αυτές του ευρωπαϊκού πλαισίου. Συγκεκριμένα, στην αναφορά γίνεται λόγος για *«μερικά κενά [...] σε σχέση με τα ρατσιστικά και ξενοφοβικά κίνητρα των εγκλημάτων, την ευθύνη και τη δικαιοδοσία νομικών προσώπων, οι οποίοι αποτελούν το αντικείμενο της παρούσας απόφασης-πλαίσιο»*.¹²³

Έτσι, μερικά κράτη- μέλη, μεταξύ των οποίων η Ελλάδα, η Γερμανία και η Λετονία, θέσπισαν νέους νόμους και προχώρησαν σε τροποποιήσεις του ποινικού κώδικα. Η ελληνική κυβέρνηση συγκεκριμένα, τροποποίησε τον προηγούμενο αντιρατσιστικό νόμο¹²⁴ και αύξησε τις ελάχιστες ποινές για τα πλημμελήματα και τα κακουργήματα που προκαλούνται από μίσος ή προκαταλήψεις, προκειμένου να τιμωρήσει όποιον *«με πρόθεση, δημόσια, προφορικά ή δια του τύπου, μέσω του διαδικτύου ή με οποιοδήποτε άλλο μέσο ή τρόπο, υποκινεί, προκαλεί, διεγείρει ή προτρέπει σε πράξεις ή ενέργειες που*

¹²¹ Hunko, A., The EU is becoming complicit in racism and fascism in Greece, 2012, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.statewatch.org/news/2012/nov/eu-greece-racism-prel-hunko.pdf>, [προσπελάστηκε στις 16 Φεβρουαρίου 2017].

¹²² FRA, Fundamental rights: Challenges and achievements in 2014, Annual Report 2014, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: http://fra.europa.eu/sites/default/files/fra-annual-report-2014_en.pdf, [προσπελάστηκε στις 16 Φεβρουαρίου 2017], σελ. 85.

¹²³ Ό.π.

¹²⁴ ΦΕΚ Α' 191/2014

μπορούν να προκαλέσουν διακρίσεις, μίσος ή βία κατά προσώπου ή ομάδας προσώπων, που προσδιορίζονται με βάση τη φυλή, το χρώμα, τη θρησκεία, τις γενεαλογικές καταβολές, την εθνική ή εθνοτική καταγωγή, το σεξουαλικό προσανατολισμό, την ταυτότητα φύλου ή την αναπηρία, κατά τρόπο που εκθέτει σε κίνδυνο τη δημόσια τάξη ή ενέχει απειλή για τη ζωή, την ελευθερία ή τη σωματική ακεραιότητα των ως άνω προσώπων [...]»

Ο όρος ρατσισμός συνδέεται με το μίσος και τον φόβο για ανθρώπους ανάλογα με την διαφορετικότητά τους στο χρώμα και την φυλή, την καταγωγή, το φύλο τους και το επάγγελμά τους.¹²⁵ Άτομα που βιώνουν τον ρατσισμό αντιμετωπίζουν άنيση μεταχείριση, μεροληπτική και άδικη αντιμετώπιση, περιθωριοποίηση και κοινωνικό αποκλεισμό από τους υπόλοιπους. Επίσης στερούνται την ελευθερία τους και, ενίοτε, την ισοτιμία. Οι ρατσιστές υποστηρίζουν την διαφορετικότητα των φυλών έχοντας την αντίληψη ότι η δική τους φυλή είναι ανώτερη από τις υπόλοιπες που έχουν καταδικαστεί από την φύση τους σε κληρονομική κατωτερότητα.

Ο ρατσισμός διακρίνεται σε τρεις βασικές μορφές, τον φυλετικό-εθνικό, τον θρησκευτικό και τον κοινωνικό. Στον φυλετικό-εθνικό ρατσισμό επικρατεί η αντίληψη ότι λόγω των βιολογικών διαφορών μεταξύ των φυλών, μπορούμε να θεωρήσουμε άλλες ανώτερες και άλλες κατώτερες. Έτσι οι άνθρωποι εξυμνούν φανατικά την εθνική ζωή και κρατούν δυσμενή στάση σε κάθε ξένο στοιχείο. Ο θρησκευτικός είναι ο ρατσισμός με βάση το θρήσκευμα. Χαρακτηριστικό του αποτελεί η αντίληψη ότι ένα δόγμα είναι το μόνο αληθινό και πρέπει να επικρατήσει έναντι των υπολοίπων. Συχνά παρατηρείται η προσπάθεια επιβολής της εκάστοτε θρησκείας με χρήση βίας. Τέλος, ο κοινωνικός ρατσισμός χαρακτηρίζεται από την εχθρική συμπεριφορά απέναντι σε κοινωνικές ομάδες όπως είναι οι τοξικομανείς, οι άνεργοι, οι άποροι, οι άστεγοι, άτομα με ειδικές ανάγκες, σεξισμός με θύματα γυναίκες καθώς και περιθωριακών ομάδων (π.χ. οι φορείς του AIDS, οι ομοφυλόφιλοι, οι φυλακισμένοι κ.λπ.), και τέλος από τη δυσανεξία απέναντι σε μετανάστες και πρόσφυγες.¹²⁶

¹²⁵ Μπαμπινιώτης, Γ., Λεξικό της νέας ελληνικής γλώσσας (Γ' έκδοση), Κέντρο Λεξικολογίας, Αθήνα, 2008, σελ. 1548.

¹²⁶ Ο.π.

Οι πρόσφυγες είναι πιθανόν να έρθουν αντιμέτωποι και με τις τρεις μορφές ρατσισμού, τόσο με τον φυλετικό-εθνικό λόγω της καταγωγής τους και με τον θρησκευτικό λόγω του διαφορετικού δόγματος στο οποίο πιστεύουν, όσο και με τον κοινωνικό.

Στη διαμόρφωση των αιτιών ρατσισμού που εμφανίζεται κυρίως τη σημερινή εποχή παίζει σημαντικό ρόλο η οικονομική και κατ' επέκταση η αξιακή κρίση που πλήττει πολλά κράτη. Σε τέτοιες περιόδους ο ρατσισμός βρίσκει γόνιμο έδαφος, καθώς ο άνθρωπος αδυνατεί να κατανοήσει τα πραγματικά αίτια της πτώσης του βιοτικού επιπέδου του και οδηγείται σε λανθασμένα συμπεράσματα, και φανταστικούς υπεύθυνους που τον οδηγούν σε εκρήξεις ρατσιστικής βίας. Παραδείγματος χάρη, συχνά αποδίδονται τα αίτια της κρίσης και των μεγάλων ποσοστών ανεργίας στους μετανάστες και η απομάκρυνσή τους από την εκάστοτε χώρα θεωρείται από πολλούς η λύση για τέτοιου είδους προβλήματα. Επιπλέον, ο ρατσισμός συχνά εξυπηρετεί κάποια συμφέροντα. Αναφερόμενος στον αντισημιτισμό, μία από τις πλέον ειδεχθείς μορφές σύγχρονου ρατσισμού, ο Sartre υποστηρίζει ότι αποτελεί δικλείδα «ασφαλείας για τις τάξεις των κατεχόντων, οι οποίες τον ενισχύουν και αντικαθιστούν έτσι ένα επικίνδυνο μίσος ενάντια στο κατεστημένο με ένα ανώδυνο μίσος, εναντίον συγκεκριμένων προσώπων»¹²⁷. Τέλος, αιτίες ρατσισμού μπορούν να αποτελέσουν η ελλιπής παιδεία και μόρφωση, η μη αποδοχή της διαφορετικότητας, η πνευματική ένδεια, ο φόβος για το άγνωστο, η έλλειψη αυτοσεβασμού και το σύμπλεγμα κατωτερότητας ή ανωτερότητας.

1.11 ΞΕΝΟΦΟΒΙΑ

Οι περισσότεροι ορισμοί της «ξеноφοβίας» στηρίζονται στη σημασιολογία της ίδιας της λέξης, η οποία αναλύεται στα δύο συστατικά της: ξένος και φόβος. Έτσι η ξеноφοβία συχνά ταυτίζεται με τη *«δυσπιστία, τον παράλογο φόβο ή το μίσος [...] απέναντι σε οτιδήποτε θεωρείται ξένο»*.¹²⁸

¹²⁷ Sartre, J.- P., Στοχασμοί για το εβραϊκό ζήτημα, μετάφραση: Αθανάσιος Σαμαρτζής, Βιβλιοπωλείον της Εστίας, Αθήνα, 2006, σελ. 48.

¹²⁸ Oksana, Y., Xenophobia: Understanding the roots and consequences of negative attitudes towards immigrants, *Counseling Psychologist*, 2009, 37 (1), σελ. 56.

Ο ρατσισμός και η ξενοφοβία είναι συγγενείς έννοιες, σύμφωνα με τον Küchler. Η ξενοφοβία εκδηλώνεται ως αποκλεισμός ή διαχωρισμός μιας ομάδας, ως ένας μηχανισμός απόκρουσης του ξένου, που όμως διακρίνεται από τον ρατσισμό, στον οποίο κριτήριο αποκλεισμού είναι εκείνο της φυλετικής ή εθνικής υπεροχής, καθώς έχει ως σκοπό τη διατήρηση των αξιών, κανόνων και πρακτικών της δικής μας εθνικής ομάδας.¹²⁹

Η άνοδος της ξενοφοβίας συνδέεται και αυτή, όπως ο ρατσισμός, τόσο με την παγκοσμιοποίηση, όσο και με τη μετανάστευση, ενώ σημαντικό ρόλο φαίνεται να παίζει και η οικονομική κρίση στις διαφορετικές της εκφάνσεις, καθώς και η εν εξελίξει προσφυγική κρίση.¹³⁰ Επιπλέον, σημαντικό παράγοντα αποτελεί και η άνοδος του δεξιού λαϊκισμού και εξτρεμισμού, που στην περίπτωση της Ελλάδας αποτυπώνεται στο κόμμα της Χρυσής Αυγής, ενώ σε ευρύτερο ευρωπαϊκό πλαίσιο συνδέεται με την ενίσχυση των εθνικολαϊκιστικών και ευρωσκεπτικιστικών κομμάτων, αλλά και την παρουσία μεταδημοκρατικών σχηματισμών¹³¹ στον πολιτικό χώρο.

Σημαντικό ρόλο στην ανάπτυξη της ξενοφοβίας παίζουν και τα ΜΜΕ, καθώς συχνά κατασκευάζουν και αναπαράγουν στερεότυπα, όπως αυτό του «επικίνδυνου Αλβανού», μία έκφραση που συναντάται συχνά, ακόμα και αυτούσια, σε δημοσιεύματα του ελληνικού Τύπου.¹³² Η εγκληματικότητα των μεταναστών, που στις διαστάσεις που της έχουν δοθεί, ενδεχομένως να αποτελεί έναν κατασκευασμένο μύθο, ενισχύεται σε μεγάλο βαθμό από την έκταση που της δίνουν τα ΜΜΕ. Κάθε έγκλημα αλλοδαπού συνοδεύεται συχνά από βαρείς και αιματοβαμμένους τίτλους, καθώς και από κηρύγματα ρατσιστικού περιεχομένου¹³³.

¹²⁹ Küchler, M., Xenophobie im internationalen Vergleich. In: Falter J.W., Jaschke HG., Winkler J.R. (eds) Rechtsextremismus. Politische Vierteljahresschrift, *Verlag für Sozialwissenschaften*, 1996, 27, σελ. 248-262.

¹³⁰ Λιαλιούτη, Ζ., Γεωργιάδου, Β., Γαλαριώτης, Γ., & Καφέ, Α., Η Ξενοφοβία στην Ελλάδα ως ερευνητικό πρόβλημα: Εννοιολογικά και μεθοδολογικά ζητήματα, στο: Επιστήμη και κοινωνία: Επιθεώρηση πολιτικής και ηθικής θεωρίας, *Gutenberg (περιοδικό)*, 2017, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://ejournals.e-publishing.ekt.gr/index.php/sas/article/view/11494/11559> , [προσπελάστηκε στις 20 Μαρτίου 2017], σελ. 162.

¹³¹ Μαυροζαχαράκης, Μ., Colin Crouch, Σειρά μεγάλοι στοχαστές, 24 γράμματα, 2014, , [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: http://www.24grammata.com/wp-content/uploads/2014/01/COLIN-CROUCH-manrozarachakis-24grammata.com_.pdf , [προσπελάστηκε στις 18 Μαρτίου 2017], σελ. 29.

¹³² Λιαλιούτη et al., ό.π. παραπομπή 130, σελ. 171.

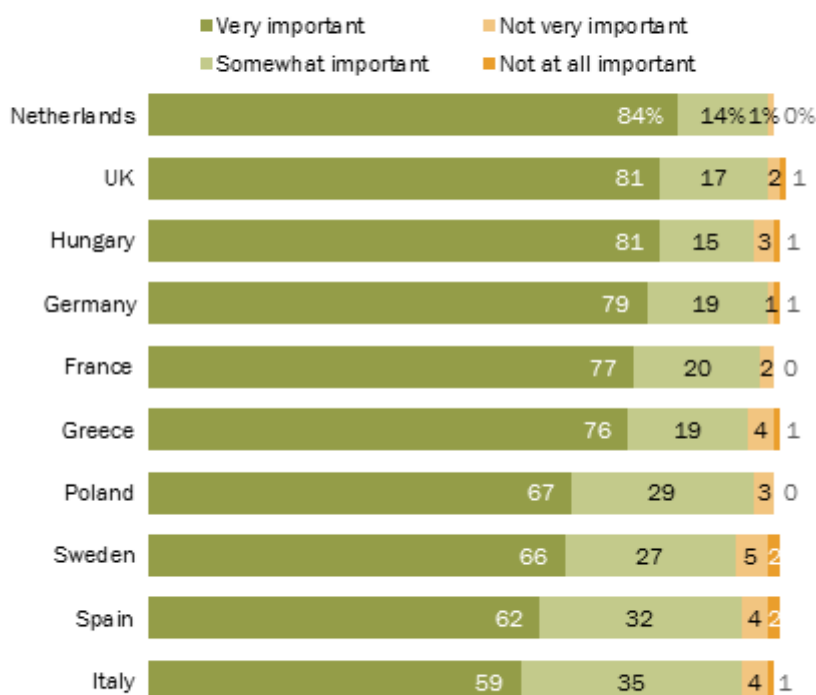
¹³³ Έμκε - Πουλοπούλου, Ή., Μετανάστες και πρόσφυγες στην Ελλάδα 1970-1990, *Εκλογή Θεμάτων Κοινωνικής Πρόνοιας*, 1990, 85/86, σελ. 46.

Ερευνητικά έχει τεκμηριωθεί και η ανάπτυξη της ξενοφοβίας σε σχέση με την εθνική ταυτότητα. Έρευνα του PewResearch Center επεξεργάστηκε, μεταξύ άλλων, έναν «δείκτη εθνικής ταυτότητας» με βάση τέσσερις θεματικές που συγκροτούν την εθνική ταυτότητα. Στόχος της έρευνας ήταν να μετρηθεί ο «αποκλειστικός» (exclusionary) χαρακτήρας της εθνικής ταυτότητας στις χώρες-μέλη της Ε.Ε..

Πιο αναλυτικά, εστιάζοντας στις δύο χώρες της παρούσας μελέτης, ως προς το εάν είναι σημαντικό να μπορεί κάποιος να μιλήσει τη γλώσσα της χώρας, το 98% των Γερμανών και το 95% των Ελλήνων το θεωρούν κάτι το σημαντικό (εικόνα 7).

Ability to speak the national language viewed as very important across Europe

Being able to speak our national language is ___ for being truly (survey country nationality)



Source: Spring 2016 Global Attitudes Survey. Q85b.

PEW RESEARCH CENTER

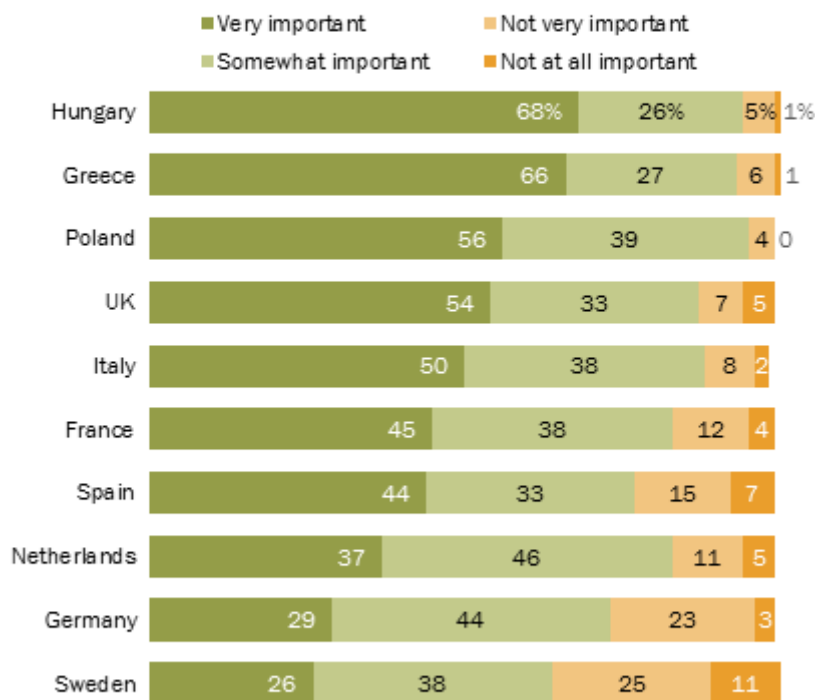
Εικόνα 7. Η σημασία του να μιλάει κανείς τη γλώσσα της χώρας¹³⁴

¹³⁴ Βλ. <http://www.pewglobal.org/2017/02/01/language-the-cornerstone-of-national-identity/>

Επιπλέον, το να γνωρίζει κάποιος τα ήθη και τα έθιμα της χώρας είναι σημαντικό για το 93% των Ελλήνων και για το 73% των Γερμανών (εικόνα 8).

Europeans differ on importance of customs and traditions to national identity

Sharing our national customs and traditions is ___ for being truly (survey country nationality)



Source: Spring 2016 Global Attitudes Survey. Q85d.

PEW RESEARCH CENTER

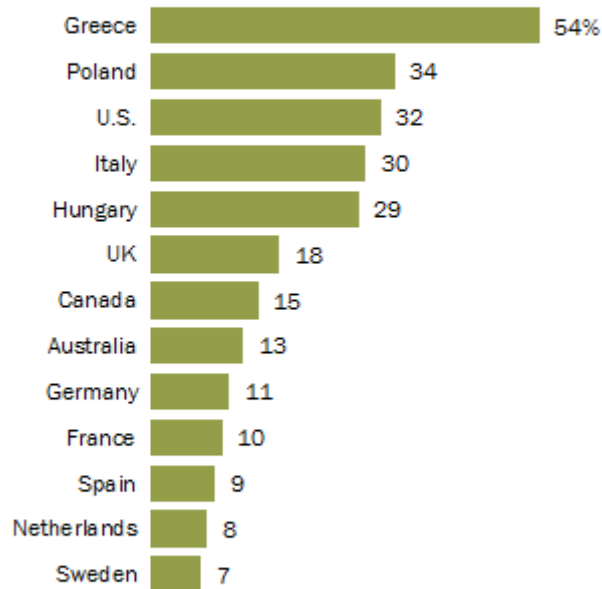
Εικόνα 8. Η σημασία του να γνωρίζει κάποιος τα ήθη και τα έθιμα της χώρας ¹³⁵

Η θρησκεία παίζει επίσης σημαντικό ρόλο στη διαμόρφωση της εθνικής ταυτότητας για το 54% των Ελλήνων, αλλά μόλις για το 11% των Γερμανών (εικόνα 9).

¹³⁵ Βλ. <http://www.pewglobal.org/2017/02/01/wide-disparity-on-the-importance-of-national-customs-and-traditions/>

Relatively few say religion essential to national identity

Being a Christian is very important for being truly (survey country nationality)*



* In Italy, Poland and Spain, asked "Catholic." Question not asked in Japan.

Source: Spring 2016 Global Attitudes Survey. Q85c.

PEW RESEARCH CENTER

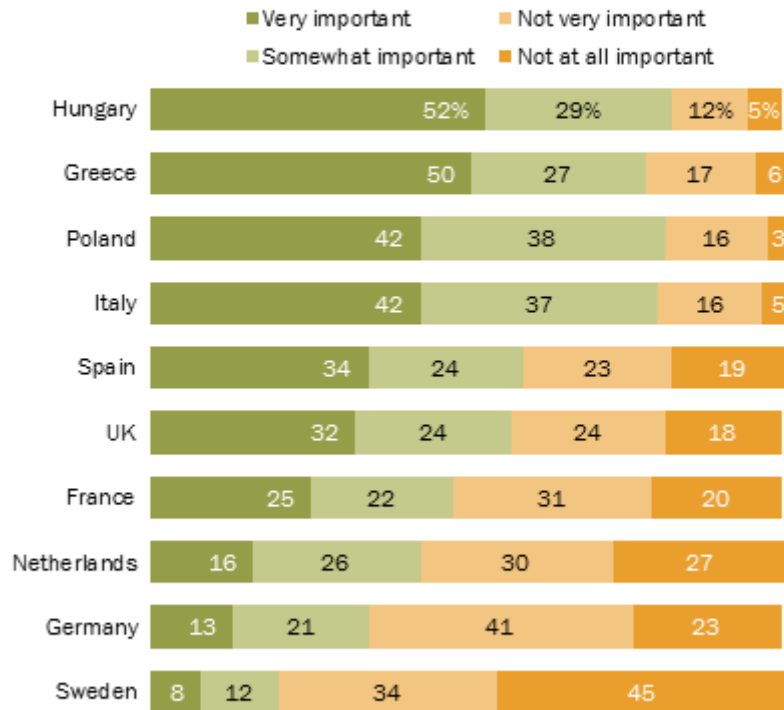
Εικόνα 9. Η σημασία της θρησκείας ¹³⁶

Τέλος, όπως φαίνεται στην παρακάτω εικόνα, το να έχει γεννηθεί κάποιος στην εκάστοτε χώρα, θεωρείται σημαντικό για το 77% των Ελλήνων και για το 34% των Γερμανών (εικόνα 10).

¹³⁶ Βλ. <http://www.pewglobal.org/2017/02/01/faith-few-strong-links-to-national-identity/>

Hungarians and Greeks most likely to say birthplace is very important to national identity

Having been born in our country is ___ for being truly (survey country nationality)



Source: Spring 2016 Global Attitudes Survey. Q85a.

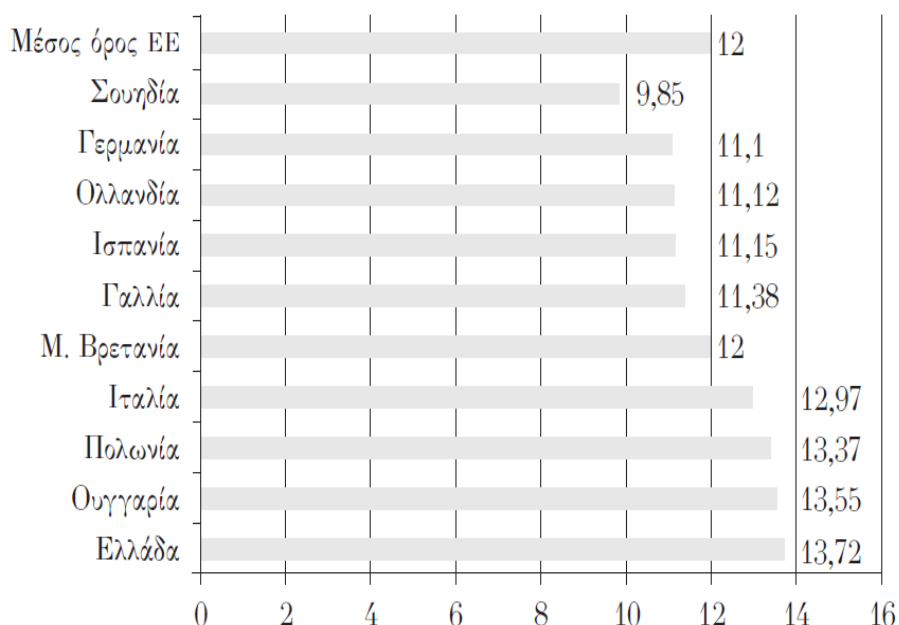
PEW RESEARCH CENTER

Εικόνα 10. Η σημασία του τόπου γέννησης ¹³⁷

Παρατηρώντας τα συγκεντρωτικά αποτελέσματα που προκύπτουν από τις τέσσερις θεματικές που προαναφέρθηκαν- γλώσσα, έθιμα, θρησκεία και τόπος γέννησης- μπορεί κανείς να παρατηρήσει ότι η Ελλάδα βρίσκεται στην κορυφή της κλίμακας εθνικής ταυτότητας, ενώ η Γερμανία στην προτελευταία θέση της σχετικής κλίμακας και κάτω από το μέσο όρο των κρατών της Ευρωπαϊκής Ένωσης, όπως φαίνεται στο παρακάτω γράφημα:

¹³⁷ Βλ. <http://www.pewglobal.org/2017/02/01/birthright-nationality/>

Γράφημα 1: Κλίμακα εθνικής ταυτότητας



Πηγή: PewResearchCenter 2016

Εικόνα 11. Κλίμακα εθνικής ταυτότητας ¹³⁸

Επομένως, οι παραπάνω παράγοντες, της γλώσσας, της θρησκείας, των πολιτισμικών διαφορών και της καταγωγής, οδηγούν πολύ συχνά στην περιθωριοποίηση των προσφύγων και κατ' επέκταση στο να δυσχεραίνεται η ενσωμάτωσή τους στη χώρα υποδοχής.¹³⁹ Χαρακτηριστικό είναι επίσης ότι, τουλάχιστον μεταπολεμικά, οι ξενοφοβικοί και οι ρατσιστές δεν επικαλούνται πλέον τόσο την ανωτερότητα κάποιων φυλών απέναντι σε κάποιες άλλες, αλλά τονίζουν τη ζημιά που θα προκαλέσουν οι πολιτισμικές ασυμβατότητες.^{140, 141}

¹³⁸ Λιαλιούτη, Ζ., Γεωργιάδου, Β., Γαλαριώτης, Γ., & Καφέ, Α., Η Ξενοφοβία στην Ελλάδα ως ερευνητικό πρόβλημα, ό.π. παραπομπή 130, σελ. 178.

¹³⁹ Έμκε - Πουλοπούλου, Ή., Μετανάστες και πρόσφυγες στην Ελλάδα 1970-1990, ό.π. παραπομπή 133, σελ. 44.

¹⁴⁰ Ντε Μπενουά, Α., Το θινκ τανκ του σύγχρονου ρατσισμού, *Ελευθεροτυπία*, 2000, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.iospress.gr/ios2000/ios20000305a.htm>, [προσπελάστηκε στις 12 Ιουνίου 2017].

¹⁴¹ Τρουμπέτα, Σ., Μερικές σκέψεις για την παράσταση του «άλλου», *Επιθεώρηση Κοινωνικών Ερευνών*, 2000, 101-102, σελ. 139.

1.12 ΑΚΡΟΔΕΞΙΑ

Την τελευταία δεκαετία η Ευρώπη αντιμετωπίζει μία αξιοσημείωτη άνοδο της ακροδεξιάς που συμπεριλαμβάνει τόσο εθνικιστικά κόμματα, ορισμένα εκ των οποίων έχουν νεοναζιστικές καταβολές, όσο και κάποιες αντιδραστικές πολιτικές ομάδες. Παρόμοια άνοδος της ακροδεξιάς είχε παρατηρηθεί και τη δεκαετία του 1930, όπου η μεγάλη ύφεση είχε ενισχύσει τέτοιου είδους κινήματα και τα είχε κάνει ιδιαίτερος δημοφιλή στην Ευρώπη¹⁴². Στις μέρες μας, η μεγάλη οικονομική και κατ' επέκταση κοινωνική κρίση που βιώνουμε έχει προκαλέσει αντίστοιχα κοινωνικά αντανakλαστικά και έτσι η αναγέννηση της ακροδεξιάς αποτελεί για ακόμα μία φορά το προϊόν αποτυχημένων οικονομικών και κοινωνικών πολιτικών, καθώς *«μπορούμε να πούμε ότι υπάρχει σύνδεση μεταξύ της εμφάνισης ενός ακροδεξιού κόμματος και των οικονομικών διακυμάνσεων»*.¹⁴³

Τα κόμματα της ακροδεξιάς, που χαρακτηρίζονται ως «κόμματα με εχθρότητα προς τον ξένο» ή αλλιώς «κόμματα ξενοφοβίας», έχουν την τάση να ανεβάζουν τα ποσοστά τους υπό τέτοιες συνθήκες, όμοιες με αυτές που βιώνει η Ευρώπη σήμερα, καθώς επωφελούνται από την οργή της κατώτερης και νεόπτωχης μεσαίας τάξης, η οποία ψάχνει τρόπους να εκφράσει την αγανάκτησή της για τα πολιτικές λιτότητας του εκάστοτε κράτους. Η στροφή μίας μερίδας του κόσμου στα ακροδεξιά κόμματα οφείλεται ουσιαστικά στο ότι *«προσφέρουν απλοϊκές απαντήσεις σε υπερβολικά σύνθετα προβλήματα και έχουν αναπτύξει αποτελεσματικές ρητορικές στρατηγικές για να κινητοποιούν πολίτες να ψηφίζουν υπέρ αυτών των κομμάτων και συνεπώς ενάντια στα πραγματικά τους συμφέροντα. Η ακροδεξιά επικαλείται το κοινό αίσθημα των ανθρώπων περί «προδοσίας και τιμωρίας» ή περί «προτίμησης και παροχής βοήθειας στους ξένους έναντι όσων διασφαλίζουν την εθνική καθαρότητα»*.¹⁴⁴

Επιπλέον, τα ακροδεξιά κόμματα εκμεταλλεύονται σε μεγάλο βαθμό την προσφυγική κρίση. Με αφορμή τις επιθέσεις των τζιχαντιστών, εξομοιώνουν όλους τους πρόσφυγες

¹⁴² Γκολντσάιν, Ρ., Η άνοδος της ακροδεξιάς: μια δυσσιώνη ηχώ, Asia Times Online, μετάφραση: Κουκουμά, Α., [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://rnbnet.gr/details.php?id=821?id=821>, [προσπελάστηκε στις 12 Μαρτίου 2017].

¹⁴³ Νοτοπούλου, Κ., Μεταναστευτική Κρίση στην Ε.Ε., ό.π. παραπομπή 45, σελ. 66.

¹⁴⁴ Η άνοδος της ακροδεξιάς: μια δυσσιώνη ηχώ, <http://rnbnet.gr/details.php?id=821?id=821>

με τρομοκράτες και τη φιλομεταναστευτική πολιτική με απειλή για την ασφάλεια και την οικονομία της εκάστοτε χώρας. Παράλληλα, παρουσιάζουν τους πρόσφυγες ως ανθρώπους χωρίς δικαιώματα, ενώ τους κατηγορούν ως εν δυνάμει τρομοκράτες ή υπαίτιους σεξουαλικών εγκλημάτων.¹⁴⁵

Η ανάπτυξη της ακροδεξιάς ιδεολογίας στα κράτη της Ευρωπαϊκής Ένωσης αποτελεί πρόκληση τόσο για την ίδια την Ένωση, όσο και για την ηγεσία του κάθε κράτους. Με τους πυρήνες ρατσισμού και ξеноφοβίας να αποκτούν κοινοβουλευτική δύναμη και εκπροσώπηση, κρίνεται αναγκαία η αναζήτηση των αιτιών αυτής της τάσης σε πολλαπλές αρνητικές συγκυρίες, όπως *«τα αδιέξοδα σε κοινωνικό και οικονομικό επίπεδο, τις ανεπαρκείς εθνικές ηγεσίες αλλά και τη δραματική εκτροπή της Ένωσης από αξίες όπως η δημοκρατία, η αλληλεγγύη, η συνοχή και η στήριξη του κράτους δικαίου»*.¹⁴⁶

1.12.1 Η ΑΚΡΟΔΕΞΙΑ ΣΤΗΝ ΕΛΛΑΔΑ

Στην Ελλάδα η ακροδεξιά εκπροσωπείται κοινοβουλευτικά από τη Χρυσή Αυγή, η οποία παραμένει σε γενικές γραμμές περιθωριοποιημένη, παρόλο που τα τελευταία χρόνια κατάφερε να αυξήσει τη δημοτικότητά της.

Η Χρυσή Αυγή είναι ένα αντιμεταναστευτικό κόμμα. Σύμφωνα με τον ξένο τύπο, η αύξηση των ρατσιστικών επιθέσεων στη χώρα φαίνεται να συνδέεται και με την εκλογική άνοδο της Χρυσής Αυγής, που στις εκλογές του Ιουνίου του 2012 έλαβε 21 έδρες στη Βουλή των Ελλήνων.¹⁴⁷ Παρά την αύξηση όμως των προσφυγικών ροών, τα ποσοστά του κόμματος στις βουλευτικές εκλογές παρέμειναν γύρω στο 7% από το Μάιο του 2012 μέχρι και τις τελευταίες εκλογές τον Σεπτέμβριο του 2015.¹⁴⁸ Τα αιτήματα της Χρυσής Αυγής σε σχέση με το προσφυγικό περιλαμβάνουν το κλείσιμο των συνόρων και τη διασφάλισή τους από τις Ειδικές Δυνάμεις του ελληνικού στρατού.

¹⁴⁵ Συνεχής άνοδος των ακροδεξιών κομμάτων στην Ευρώπη -Σε ποιες χώρες ανεβαίνουν επικίνδυνα τα ποσοστά τους, [iefimerida.gr](http://www.iefimerida.gr), 2015, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.iefimerida.gr/news/236718/synehis-anodos-ton-akrodexion-kommaton-stin-eyropi-se-poies-hores-anevainoun-epikindyna#ixzz4DGyrZdU6>, [προσπελάστηκε στις 11 Μαρτίου 2017].

¹⁴⁶ Βοσκόπουλος, Γ., Η άνοδος της ακροδεξιάς στην Ευρώπη, Ελεύθερη Ζώνη, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.elzoni.gr/html/ent/942/ent.15942.asp>, [προσπελάστηκε στις 11 Μαρτίου 2017].

¹⁴⁷ Βλ. [http://ekloges-prev.singularlogic.eu/v2012a/public/index.html#{"cls":"main","params":{}}](http://ekloges-prev.singularlogic.eu/v2012a/public/index.html#{)

¹⁴⁸ Βλ. [http://ekloges.yves.gr/current/v/public/index.html#{"cls":"main","params":{}}](http://ekloges.yves.gr/current/v/public/index.html#{)

Πιο συγκεκριμένα, η Χρυσή Αυγή προτείνει τη διασπορά των ναρκών στα σύνορα,¹⁴⁹ ενώ ζητά να φυλάσσονται από Έλληνες εθνικιστές, οι οποίοι κάνοντας χρήση όπλων δε θα αφήνουν κανέναν να μπαίνει παράνομα στη χώρα.¹⁵⁰ Οι μετανάστες θα συλλαμβάνονται άμεσα και οποιοδήποτε έγκλημα τελεί κάποιος αλλοδαπός θα είναι ιδιώνυμο, θα δικάζεται δηλαδή με βάση ιδιαίτερο νόμο, διαχωριζόμενο από τη γενικότερη κατηγορία αδικημάτων και θα τιμωρείται με ξεχωριστές ποινές από τις συνηθισμένες. Σύμφωνα με τη Χρυσή Αυγή, απαιτείται «*κάθαρση του συστήματος από τους υπηρέτες ξένων συμφερόντων*»¹⁵¹ και γι' αυτό όσοι μετανάστες έρχονται παράνομα στην Ελλάδα, θα πρέπει να τιμωρούνται με απέλαση και η ποινή τους θα εκτελείται στον τόπο από τον οποίο προέρχονται. Στόχος της Χρυσής Αυγής, σύμφωνα με τη διακήρυξη του κόμματος, είναι η μείωση της εγκληματικότητας που προκαλείται από τους αλλοδαπούς μετανάστες. Συχνά οι οπαδοί του κόμματος στρέφονται εναντίον των αλλοδαπών με παράνομα και βίαια μέσα, όπως με άγριους ξυλοδαρμούς, με ανθρωποκυνηγητά και μαχαιρώματα, που ενίοτε καταλήγουν και σε δολοφονίες.

Πολλές είναι οι επιθέσεις που έχει οργανώσει η Χρυσή Αυγή σε γειτονιές αλλά και σχολεία, όπου φοιτούν ή επρόκειτο να φοιτήσουν παιδιά προσφύγων. Μέλη της Χρυσής Αυγής μαζί με βουλευτές του κόμματος προπηλάκιζαν και απειλούσαν μαθητές, γονείς, εκπαιδευτικούς και δημάρχους, με χαρακτηριστικό παράδειγμα την επίθεση στο Πέραμα Αττικής τον Ιανουάριο του 2017.¹⁵²

1.12.2 Η ΑΚΡΟΔΕΞΙΑ ΣΤΗ ΓΕΡΜΑΝΙΑ

Στη Γερμανία, το ακροδεξιό κόμμα Alternative für Deutschland (AfD) εκφράζοντας την άποψη ότι το Ισλάμ δεν είναι συμβατό με το γερμανικό σύνταγμα, έφτασε ήδη από το 2015 να καταλαμβάνει στις δημοσκοπήσεις την τρίτη θέση ανάμεσα στα γερμανικά

¹⁴⁹ Βλ. <http://www.xryshaygh.com/index.php/kinima/thesis>, 6/1/2014.

<http://www.antinews.gr/2012/04/23/157504/>, 10/1/2014.

¹⁵⁰ Βλ. <http://www.antinews.gr/2012/04/23/157504/>, 10/1/2014.

¹⁵¹ Βλ. <http://www.xryshaygh.com/index.php/kinima/thesis>, 6/1/2014.

¹⁵² Η εισβολή της Χρυσής Αυγής στο Δημοτικό σχολείο στο Πέραμα, *alfavita.gr*, 2017, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.alfavita.gr/arthron/koinonia/vinteo-i-eisvoli-tis-hrysis-aygis-sto-dimotiko-sholeio-sto-perama>, [προσπελάστηκε στις 11 Μαρτίου 2017].

κόμματα με ποσοστό 10,5%¹⁵³, ενώ τον Δεκέμβριο του 2016, μετά την τρομοκρατική επίθεση σε χριστουγεννιάτικο παζάρι του Βερολίνου, άγγιξε το 15,5%¹⁵⁴. Σε επίπεδο κρατιδίων, το AfD κατάφερε στις εκλογές του 2016 να εκλέξει εκπροσώπους στα τοπικά κοινοβούλια με ποσοστά που κυμάνθηκαν μεταξύ 12 και 24%.¹⁵⁵

Παράλληλα, το αντι- μεταναστευτικό και αντι- ισλαμικό κίνημα PEGIDA, ή αλλιώς «Ευρωπαίοι Πατριώτες εναντίον της Ισλαμοποίησης της Δύσης», αποτελεί την πιο ακραία έκφραση της ακροδεξιάς στη χώρα, τις αξίες του οποίου συμερίζεται και το AfD. Η Γερμανίδα καγκελάρια, Άνγκελα Μέρκελ, καταδίκασε το κίνημα PEGIDA, το οποίο ιδρύθηκε στη Δρέσδη και ξεκίνησε από μία κλειστή ομάδα στο κοινωνικό δίκτυο Facebook και σύντομα έλαβε μεγάλες διαστάσεις, συγκεντρώνοντας ένα οργισμένο πλήθος πολιτών, οι οποίοι συζητούν για θέματα εθνικής ταυτότητας, μετανάστευσης και ενσωμάτωσης.¹⁵⁶

Σε γενικές γραμμές οι Γερμανοί δείχνουν να εκφράζουν όλο και πιο έντονα τη διαφωνία τους, αλλά κυρίως το φόβο για την άφιξη προσφύγων στη χώρα τους. Με τις πράξεις βίας κατά των προσφυγικών καταλυμάτων να έχουν υπερδιπλασιαστεί στη Γερμανία τα τελευταία χρόνια¹⁵⁷, τόσο το κόμμα AfD, όσο και το κίνημα PEGIDA, οργανώνουν συχνά διαδηλώσεις κατά των προσφύγων¹⁵⁸, αλλά και βίαιες επιθέσεις σε κέντρα φιλοξενίας και αλλού.¹⁵⁹

¹⁵³ Συνεχής άνοδος των ακροδεξιών κομμάτων στην Ευρώπη -Σε ποιες χώρες ανεβαίνουν επικίνδυνα τα ποσοστά τους, iefimerida.gr, ό.π. παραπομπή 144.

¹⁵⁴ AfD erreicht mit 15,5 Prozent erneut Höchstwert, *welt.de*, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://www.welt.de/politik/deutschland/article160569813/AfD-erreicht-mit-15-5-Prozent-erneut-Hoehchstwert.html>, [προσπελάστηκε στις 11 Μαρτίου 2017].

¹⁵⁵ Ergebnisse der Abgeordnetenhauswahlen in Berlin (bis 1989 West-Berlin), *wahlrecht.de*, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.wahlrecht.de/ergebnisse/berlin.htm>, [προσπελάστηκε στις 11 Φεβρουαρίου 2017].

¹⁵⁶ . (FRA, 'Fundamental rights: Challenges and achievements in 2014', Annual Report 2014, [Διαδίκτυο]. Διαθέσιμη στη διεύθυνση: http://fra.europa.eu/sites/default/files/fra-annual-report-2014_en.pdf, p. 53.)

¹⁵⁷ zeit.de, Gewalt gegen Flüchtlinge, <http://www.zeit.de/gesellschaft/2015-08/rechte-gewalt-fluechtling-sachsen-pegida>

¹⁵⁸ skai.gr, Χιλιάδες υποστηρικτές του Pegida διαδήλωσαν κατά των προσφύγων στη Γερμανία, <http://www.skai.gr/news/world/article/299198/hiliades-upostiriktes-tou-pegida-diadilosan-kata-ton-prosfugon-sti-germania/>

¹⁵⁹ web.de, «Pegida, AfD, NPD gegen Flüchtlinge: So gefährlich ist rechtsextreme Gewalt für Deutschland» <https://web.de/magazine/politik/pegida-afd-npd-fluechtlinge-gefaehrlich-rechtsextreme-gewalt-deutschland-31014520>

1.13 Η ΞΕΝΟΦΟΒΙΑ ΚΑΙ Ο ΡΑΤΣΙΣΜΟΣ ΣΤΗΝ ΕΛΛΑΔΑ ΚΑΙ ΣΤΗ ΓΕΡΜΑΝΙΑ

Καταγγελίες που γίνονται σε εθνικούς οργανισμούς ισότητας σε πολλά κράτη- μέλη της Ευρωπαϊκής Ένωσης, μεταξύ των οποίων η Ελλάδα και η Γερμανία, δείχνουν πως η φυλή, η εθνικότητα και το χρώμα του δέρματος παραμένουν τα πιο συχνά αίτια για τις διακρίσεις που γνωστοποιούνται. Αυτά τα δεδομένα των καταγγελιών όμως αποτελούν μόνο την κορυφή του παγόβουνου.¹⁶⁰

Στην Ελλάδα, το 2013, πριν ακόμα πυκνώσουν οι προσφυγικές ροές, το Δίκτυο Καταγραφής Περιστατικών Ρατσιστικής Βίας κατέγραψε 166 ρατσιστικά εγκλήματα μέσα σε ένα χρόνο, εκ των οποίων τα 143 στρέφονταν ενάντια σε μετανάστες ή πρόσφυγες. Στην πλειοψηφία αυτών των εγκλημάτων, το θύμα υπέφερε από «σοβαρή σωματική βλάβη», η οποία προκλήθηκε από διάφορα όπλα. Οι πρώτες έντονες αντιδράσεις στην Ελλάδα ξεκίνησαν με την έναρξη της δημόσιας συζήτησης για τη δημιουργία δομών προσωρινής φιλοξενίας για τους πρόσφυγες. Ακραία περιστατικά έλαβαν χώρα, όπως, ενδεικτικά, αυτό με τις γουρουνοκεφαλές στο στρατόπεδο στη Σκύδρα, οι οποίες τοποθετήθηκαν εκεί για να «μολυνθεί» ο χώρος για τους μουσουλμάνους, ο εμπρησμός αποθήκης με δωρεές και ανθρωπιστικό υλικό για τους πρόσφυγες στο Καστελόριζο, η γουρουνοκεφαλή στον προαύλιο χώρο του υπό ανέγερση Μουσείου Προσφυγικού Ελληνισμού στην Καβάλα και η έφοδος σε πρώην στρατόπεδο στην Λάρισα, προκειμένου να ξηλωθούν οι σκηνές που προορίζονταν για τους πρόσφυγες.¹⁶¹

Μεγάλη ένταση παρατηρήθηκε και στα νησιά, όταν ο τοπικός πληθυσμός άρχισε να αντιλαμβάνεται τις ελλείψεις στη διαχείριση της τεράστιας προσφυγικής ροής και τον αντίκτυπο που αυτή είχε, τόσο στην καθημερινότητά των κατοίκων, όσο και στον τουρισμό. Τα περιστατικά ρατσιστικής βίας στα νησιά άρχισαν να κάνουν την εμφάνισή τους από νωρίς, ενώ με την έλευση του φθινοπώρου του 2016 και το τέλος της τουριστικής περιόδου, αν και η μεγάλη πλειοψηφία των κατοίκων δεν φάνηκε να εγκρίνει τις ακρότητες, άρχισαν να πληθαίνουν οι συγκεντρώσεις διαμαρτυρίας των κατοίκων κατά της διαμονής προσφύγων στη Σάμο και στη Χίο.

¹⁶⁰ FRA, Fundamental rights: Challenges and achievements in 2014, ό.π. παραπομπή 22, σελ. 50.

¹⁶¹ Δίκτυο Καταγραφής Περιστατικών Ρατσιστικής Βίας, Ετήσια Έκθεση 2016, , [διαδίκτυο]. Διαθέσιμο στη διεύθυνση:http://rvrn.org/wp-content/uploads/2017/04/Report_2016gr.pdf, [προσπελάστηκε στις 11 Μαρτίου 2017], σελ. 1.

Στη συνέχεια, με την έναρξη της νέας σχολικής χρονιάς και την ένταξη των παιδιών-προσφύγων στο εκπαιδευτικό σύστημα της χώρας, υπήρξαν μεμονωμένες, αλλά ιδιαίτερα ακραίες αντιδράσεις κάποιων συλλόγων γονέων και κηδεμόνων, αλλά και εκπαιδευτικών, αν και οι τελευταίες σε μικρότερο βαθμό. Τα πιο έντονα περιστατικά παρατηρήθηκαν στο Ωραιόκαστρο της Θεσσαλονίκης, το οποίο βρέθηκε στο επίκεντρο και παρέμεινε στην επικαιρότητα για αρκετές εβδομάδες. Το βασικότερο επιχείρημα των ανθρώπων που τάχθηκαν κατά της ένταξης των παιδιών- προσφύγων στα ελληνικά σχολεία ήταν αυτό του εμβολιασμού, το κατά πόσο δηλαδή τα παιδιά αυτά είναι εμβολιασμένα ή υπάρχει ο κίνδυνος μετάδοσης ασθενειών. Έτσι τέθηκε το ζήτημα για το δικαίωμα των παιδιών αυτών στην εκπαίδευση.¹⁶²

Παράλληλα, ένα τμήμα της ελληνικής κοινωνίας βρέθηκε στο πλευρό των προσφύγων και ιδίως όσον αφορά στο ζήτημα της εκπαίδευσης των παιδιών, το οποίο, ένωσε και ενώνει πολλούς ανθρώπους, ενώ αναδεικνύεται ως σύμβολο μηδενικής ανοχής του ρατσισμού κατά των παιδιών. Από την άλλη όμως, σε ένα τμήμα του ελληνικού πληθυσμού φαίνεται πως υπάρχει ρατσισμός, ο οποίος μάλιστα είναι βαθιά ριζωμένος.

Παρόλο που τα παιδιά των προσφύγων δεν επρόκειτο να ενταχθούν στις ίδιες τάξεις με τα παιδιά των Ελλήνων, δεν έλειψαν οι αντιδράσεις από την πλευρά των γονέων σε πολλά μέρη της Ελλάδας. Για παράδειγμα, στην πόλη της Φιλιππιάδας οι γονείς τάχθηκαν κατά της εκπαίδευσης των προσφύγων στο σχολείο των παιδιών τους, επικαλούμενοι ιατρικούς, εκπαιδευτικούς, θρησκευτικούς και πολιτισμικούς λόγους, εκφράζοντας παράλληλα τους φόβους τους για τη μετάδοση ασθενειών και για ότι τα παιδιά τους δεν θα μπορέσουν «να συνυπάρξουν με παιδιά μεταναστών». Όπως χαρακτηριστικά αναφέρει η ανακοίνωση του συλλόγου γονέων «τα παιδιά των μεταναστών έχουν διαφορετικές αντιλήψεις για το ρόλο της οικογένειας, τη θέση της γυναίκας, τη θρησκεία, με ελληνόπουλα που οι γονείς τους στερήθηκαν και στερούνται για να προσφέρουν στα παιδιά τους την καλύτερη μόρφωση, αυτή που αναγνωρίζουν και

¹⁶² Δίκτυο Καταγραφής Περιστατικών Ρατσιστικής Βίας, ό.π. παραπομπή 161, σελ. 2.

συγχαίρουν τα σχολεία της Ευρώπης.»¹⁶³ Παρόμοιες αντιδράσεις υπήρξαν και από σχολεία άλλων ελληνικών πόλεων, όπως είναι η Ορεστιάδα.¹⁶⁴

Στη Γερμανία η κατάσταση φαίνεται να κινείται σε παρόμοιο πλαίσιο. Τόσο το 2014, όσο και το 2015, λάμβαναν χώρα στη Γερμανία κάθε εβδομάδα κατά μέσο όρο σχεδόν έξι διαδηλώσεις εναντίον των προσφύγων. Οι διαδηλώσεις αυτές οργανώθηκαν από διαφορετικές ομάδες, μεταξύ άλλων από ακροδεξιές οργανώσεις ή από πολιτικά κόμματα, όπως είναι το Εθνικοδημοκρατικό Κόμμα Γερμανίας (NPD), οι Δεξιοί (Die Rechte) και ο Τρίτος Δρόμος (Der Dritte Weg), καθώς και από πιο νέες ομάδες ή κόμματα, όπως οι Πατριώτες Ευρωπαίοι κατά του εξισλαμισμού της Δύσης (PEGIDA) και η Εναλλακτική για τη Γερμανία (AfD), για τις οποίες έγινε λόγος παραπάνω. Σύμφωνα με στοιχεία της Διεθνούς Αμνηστίας, το 2015 καταγράφηκαν στις αρχές 1.031 επιθέσεις εναντίον προσφυγικών καταλυμάτων, πέντε φορές περισσότερες από το 2014, που είχαν καταγραφεί 199 επιθέσεις.¹⁶⁵ Έπειτα, μόνο τους πρώτους τρεις μήνες του 2016, καταγράφηκαν 347 επιθέσεις, εκ των οποίων οι 319 είχαν οργανωθεί ή υποκινηθεί από ακροδεξιές ομάδες ή οργανώσεις.¹⁶⁶

1.13.1 ΤΙ ΛΕΝΕ ΟΙ ΕΡΕΥΝΕΣ

Σε δημοσκόπηση που πραγματοποιήθηκε τον Μάρτιο του 2016 από την Κάπα Research για την Περιφερειακή Ένωση Δήμων Αττικής¹⁶⁷, το 67,4% των Ελλήνων απάντησε ότι η προσφυγική κρίση, έτσι όπως εξελίσσεται στην Ελλάδα, τους προκαλεί ανησυχία, ενώ

¹⁶³ Ανεπιθύμητα τα προσφυγόπουλα στο 2ο Δημοτικό Σχολείο Φιλιπιάδας, kathimerini.gr, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.kathimerini.gr/875595/article/epikairothta/ellada/anepi8ymhta-ta-prosfgoporoyla-sto-2o-dhmotiko-sxoleio-filippiadas>, [προσπελάστηκε στις 7 Φεβρουαρίου 2017], σελ. 73.

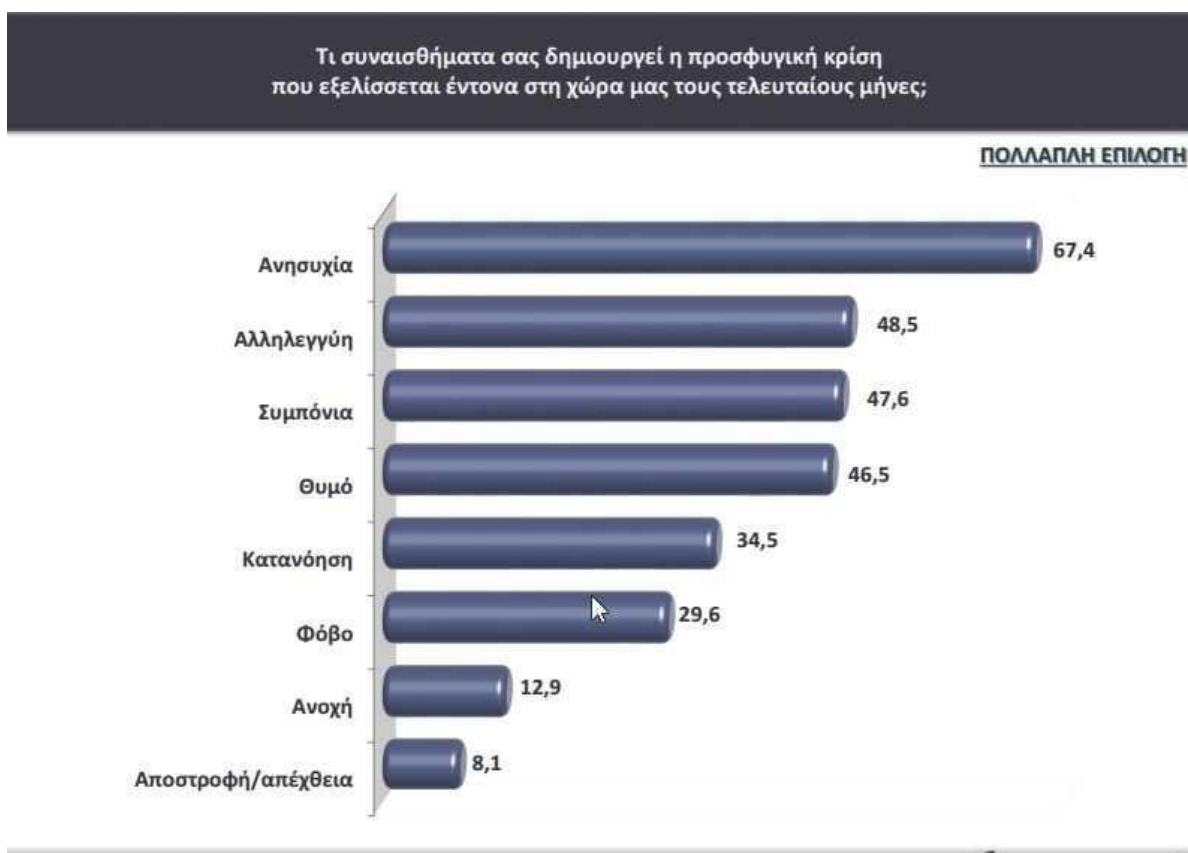
¹⁶⁴ Weise, Z., Refugee children get little education in Europe, politico.eu, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.politico.eu/article/refugee-children-education-europe-volunteers-greece-migrant-camp/>, [προσπελάστηκε στις 7 Φεβρουαρίου 2017].

¹⁶⁵ Kampf, L., Viele Straftaten gegen Asylunterkünfte, tagesschau.de, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://www.tagesschau.de/inland/bka-asylunterkuenfte-101.html>, [προσπελάστηκε στις 9 Φεβρουαρίου 2017].

¹⁶⁶ Amnesty International, Leben in Unsicherheit- Wie Deutschland die Opfer rassistischer Gewalt im Sicht lässt, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://www.amnesty.de/files/Amnesty-Bericht-Rassistische-Gewalt-in-Deutschland-Juni2016.pdf>, [προσπελάστηκε στις 9 Φεβρουαρίου 2017], σελ. 50.

¹⁶⁷ Κάπα research, Έρευνα στην Αττική για την προσφυγική κρίση, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: http://i2.aftocdn.gr/mediafiles/2016/05/erevna_peda_aftodioikisi.pdf, [προσπελάστηκε στις 18 Φεβρουαρίου 2017].

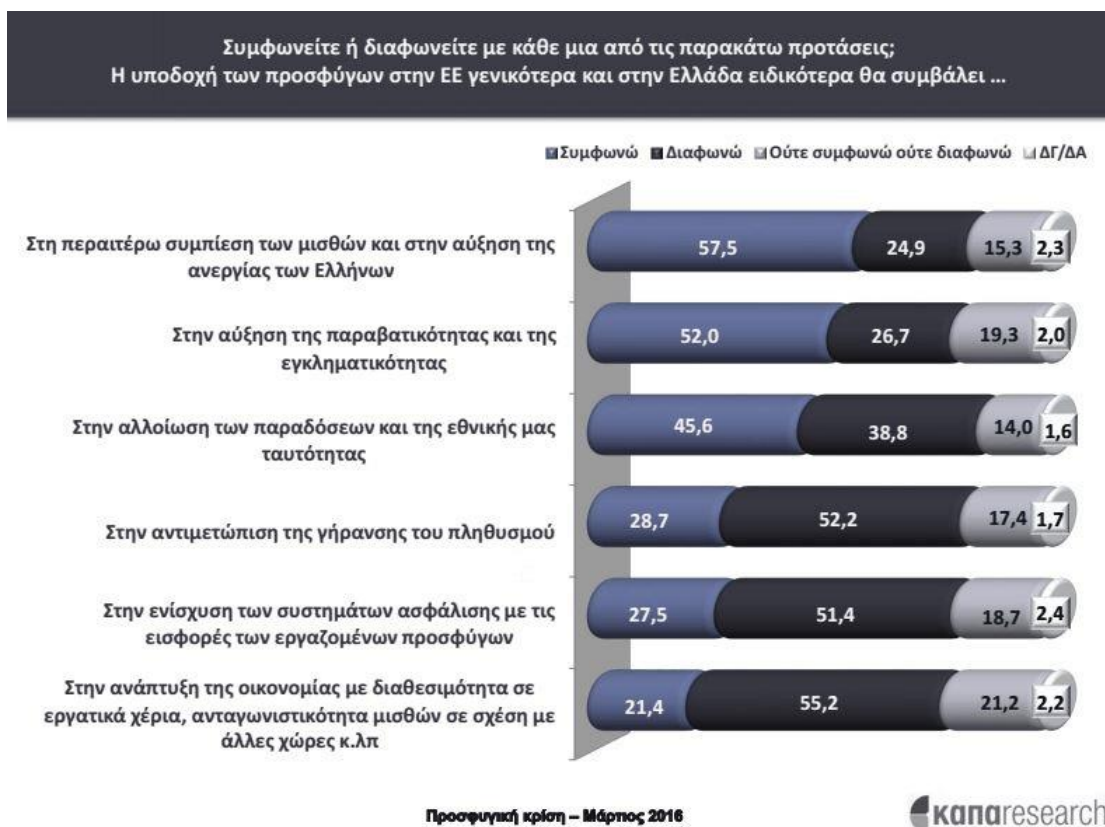
στο 48,5% προκαλεί αισθήματα αλληλεγγύης, στο 47,6% συμπόνιας και στο 46,5% θυμού (εικόνα 12).



Εικόνα 12. Τα συναισθήματα που προκαλεί η προσφυγική κρίση στους Έλληνες ¹⁶⁸

Συγκεκριμένα, οι Έλληνες ανησυχούν περισσότερο για θέματα τρομοκρατίας, ανεργίας και εθνικής ταυτότητας. Το 57,5% θεωρεί ότι η υποδοχή των προσφύγων θα συμβάλει στην περαιτέρω συμπίεση των μισθών και την αύξηση της ανεργίας, ενώ το 52% ότι θα συμβάλει στην αύξηση της παραβατικότητας και της εγκληματικότητας. Επίσης, το 45,6% ανησυχεί πως οι πρόσφυγες θα αλλοιώσουν τις παραδόσεις και την εθνική μας ταυτότητα, ενώ για το ενδεχόμενο μίας τρομοκρατικής επίθεσης στην Ελλάδα, παρόμοιας με αυτής στο αεροδρόμιο των Βρυξελλών, ανησυχεί το 54,5% των ερωτηθέντων (εικόνα 13).

¹⁶⁸ Ό.π. παραπομπή 167, σελ. 4.



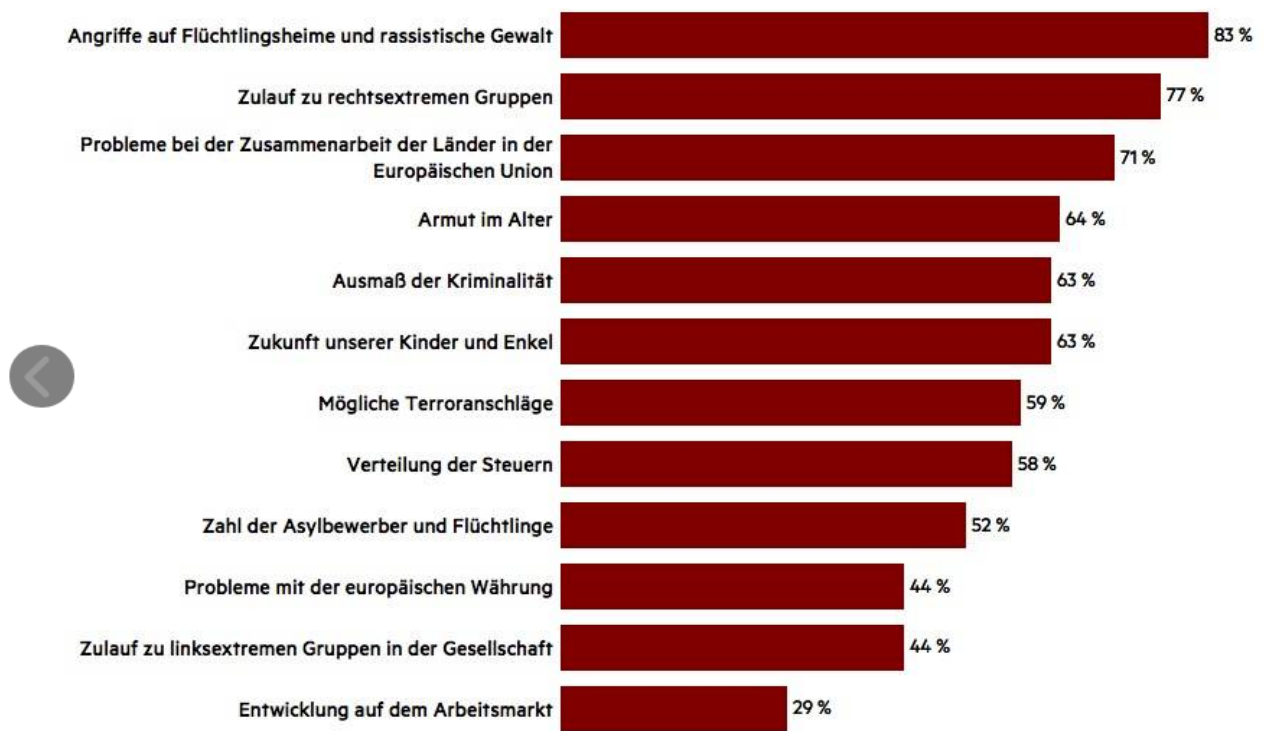
Εικόνα 13. Οι τομείς στους οποίους θα συμβάλει η υποδοχή των προσφύγων ¹⁶⁹

Σε αντίστοιχη δημοσκόπηση που πραγματοποιήθηκε επίσης τον Μάρτιο του 2016 στη Γερμανία από τη Forsa για το περιοδικό Stern¹⁷⁰, οι δύο μεγαλύτερες ανησυχίες των Γερμανών σχετικά με το θέμα των προσφύγων είναι οι επιθέσεις στα προσφυγικά καταλύματα και η ρατσιστική βία, καθώς είναι κάτι που ανησυχεί το 83% των ερωτηθέντων και ακολουθεί η εισροή του κόσμου στις ακροδεξιές οργανώσεις, καθώς προκαλεί ανησυχία στο 77% των ερωτηθέντων. Το μέλλον των παιδιών και των εγγονών τους απασχολεί το 63% του πληθυσμού, ενώ το ενδεχόμενο φτώχειας στο γηραιότερο πληθυσμό και η αύξηση της εγκληματικότητας απασχολεί το 64% και το 63% του πληθυσμού αντίστοιχα. Η εξέλιξη στην αγορά εργασίας βρίσκεται στην τελευταία θέση και ανησυχεί μόλις το 29% των ερωτηθέντων (εικόνα 14).

¹⁶⁹ Ο.π. παραπομπή 167, σελ. 16.

¹⁷⁰ Das denken die Deutschen wirklich über Flüchtlinge, Merkel - und die Medien, *stern.de*, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.stern.de/politik/deutschland/fluechtlinge--das-denken-die-deutschen-wirklich-6737204.html>, [προσπελάστηκε στις 11 Φεβρουαρίου 2017].

Deutschland 2016: Was die Menschen bewegt und ängstigt



Εικόνα 14. Οι ανησυχίες των Γερμανών σε σχέση με την προσφυγική κρίση ¹⁷¹

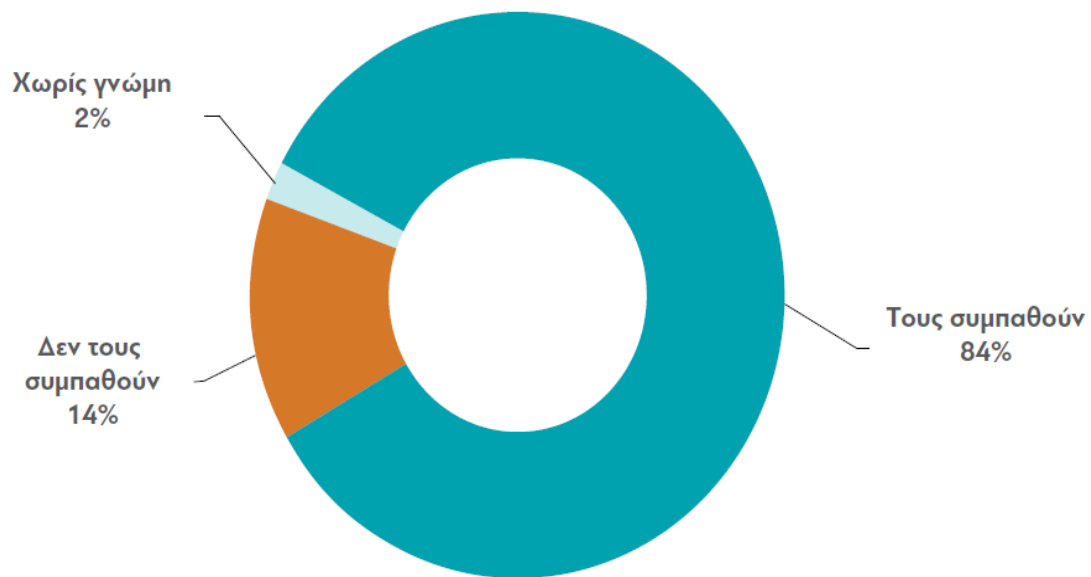
Στο σημείο αυτό να αναφέρουμε ότι στην πρώτη έρευνα δεν ζητήθηκε από τους Έλληνες η άποψή τους σε σχέση με την άνοδο της ακροδεξιάς ή τις ρατσιστικές επιθέσεις εναντίον των προσφύγων.

Κοινό χαρακτηριστικό Ελλήνων και Γερμανών φαίνεται να αποτελεί το γεγονός ότι, ενώ και οι δύο δείχνουν πως βλέπουν με συμπόνια τους πρόσφυγες, ωστόσο δεν επιθυμούν τη μόνιμη εγκατάστασή τους σε Ελλάδα και Γερμανία αντίστοιχα.

Η συμπάθεια των Ελλήνων απέναντι στους πρόσφυγες είναι σχεδόν καθολική και κυμαίνεται σε ποσοστό 84% σύμφωνα με έρευνα της Public Issue για τον οργανισμό έρευνας και ανάλυσης διαΝΕΟσις που διενεργήθηκε το Φεβρουάριο του 2016 (εικόνα 15).¹⁷²

¹⁷¹ Ο.π. παραπομπή 168, σελ. 16.

¹⁷² Διανέοσις, Το προσφυγικό πρόβλημα και οι Έλληνες, 2016, [διαδίκτυο], Διαθέσιμο στη διεύθυνση: http://www.dianeosis.org/wp-content/uploads/2016/02/immigration_04.pdf.

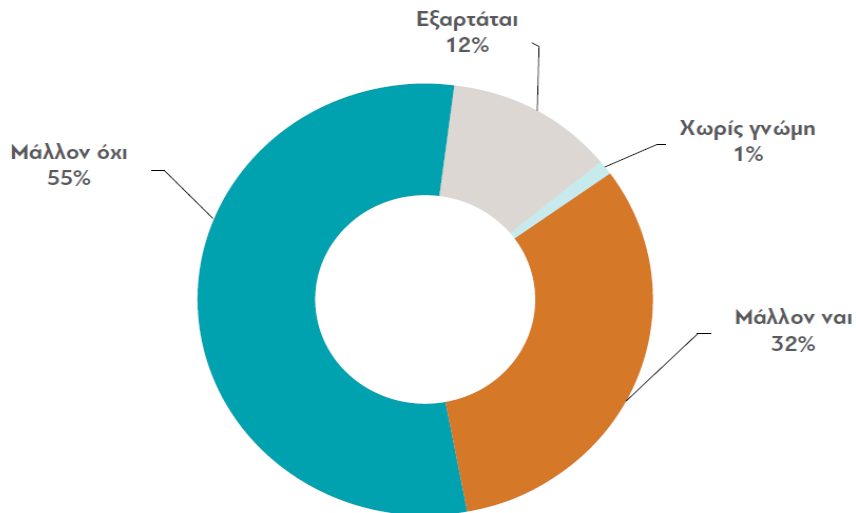


Εικόνα 15. Η συμπάθεια των Ελλήνων απέναντι στους πρόσφυγες ¹⁷³

Το ζήτημα της μόνιμης εγκατάστασης προσφύγων στην Ελλάδα όμως δε γίνεται αποδεκτό από την πλειοψηφία της κοινής γνώμης (55%). Υπέρ της μόνιμης εγκατάστασης τάσσεται μόνο το 32% των ερωτηθέντων, ενώ το 12% δηλώνει ότι η στάση του εξαρτάται από διάφορους παράγοντες και προϋποθέσεις (εικόνα 16).

¹⁷³ Ό.π. παραπομπή 172, σελ. 53.

Θ3. ΔΙΑΓΡΑΜΜΑ 43: Εσείς προσωπικά, θα θέλατε να εγκατασταθούν πρόσφυγες μόνιμα στην Ελλάδα; Μάλλον ναι ή μάλλον όχι;



Εικόνα 16. Άποψη Ελλήνων για την προοπτική μόνιμης εγκατάστασης των προσφύγων στην Ελλάδα¹⁷⁴

Παρόμοια είναι και τα αποτελέσματα στη Γερμανία. Ενώ το 81% των ερωτηθέντων σε έρευνα που πραγματοποίησε το Deutsches Institut für Wirtschaftsforschung (DIW) τον Μάρτιο του 2016¹⁷⁵ τάσσονται υπέρ των προσφύγων και του να γίνουν δεκτοί στη χώρα (εικόνα 17), το 55% του πληθυσμού απάντησε ότι θα ήθελε οι πρόσφυγες να γυρίσουν πίσω μετά από μερικά χρόνια στην πατρίδα τους και μόλις το 28% πιστεύει ότι θα πρέπει οι πρόσφυγες να παραμείνουν στη Γερμανία, εφόσον το θέλουν, ακόμα κι όταν καλυτερεύσει η κατάσταση στη χώρα τους (εικόνα 18).

¹⁷⁴ Ό.π. παραπομπή 172, σελ. 53.

¹⁷⁵ Eisnecker, P., & Schupp, J., Stimmungsbarometer zu Geflüchteten in Deutschland, Deutsches Institut für Wirtschaftsforschung, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: https://www.diw.de/documents/publikationen/73/diw_01.c.534430.de/16-21.pdf, [προσπελάστηκε στις 11 Μαρτίου 2017].

Einstellungen der BürgerInnen zum Aufenthaltsrecht unterschiedlicher Gruppen von Flüchtlingen

In Prozent

Fluchtgrund	Ablehnung	Ambivalenz	Zustimmung	Mittelwert ¹
Subsidiär Schutzberechtigte (EU-Recht)	10	8	81	8,9
Verfolgte wegen... (Genfer Konvention)	20	16	63	7,4
Engagements für Menschenrechte	14	12	74	8,3
Engagements für Gewerkschaften	31	20	49	6,5
Religion (Christen)	14	14	72	8,2
Religion (Muslime)	31	18	51	6,7
Zugehörigkeit zu ethnischer Minderheit	21	15	64	7,6
Homosexualität	27	16	57	7,1
Gesamteinschätzung aller Fluchtgründe	19	13	69	7,4

Εικόνα 17. Άποψη Γερμανών για το δικαίωμα παραμονής των προσφύγων στη Γερμανία¹⁷⁶

Einstellungen der Bürger zur Dauer des Aufenthaltsrechts

In Prozent

	Einstellung zum Aufenthaltsrecht für Flüchtlinge ¹			Gesamt
	Ablehnung	Ambivalenz	Zustimmung	
Anerkannte Flüchtlinge, die seit einigen Jahren in Deutschland leben und deren Situation im Herkunftsland sich gebessert hat, sollten ...				
... zurückgeschickt werden (Werte 1 bis 5)	82	65	45	55
neutral (6)	10	18	19	17
... hier bleiben dürfen (Werte 7 bis 11)	9	17	35	28

Εικόνα 18. Άποψη Γερμανών για την προοπτική μόνιμης εγκατάστασης των προσφύγων στη Γερμανία¹⁷⁷

Τέλος, η Διεθνής Αμνηστία, στο πλαίσιο μίας έρευνας που πραγματοποίησε¹⁷⁸, έθεσε μία σειρά κλιμακωτών ερωτήσεων σε 27.000 άτομα από 27 χώρες με στόχο να μάθει τι

¹⁷⁶ Ό.π. παραπομπή 175, σελ. 468.

¹⁷⁷ Ό.π. παραπομπή 175, σελ. 470.

¹⁷⁸ Amnesty International, Refugees Welcome Survey, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: https://www.amnesty.at/de/view/files/download/showDownload/?tool=12&feld=download&sprach_connect=424, [προσπελάστηκε στις 11 Μαρτίου 2017].

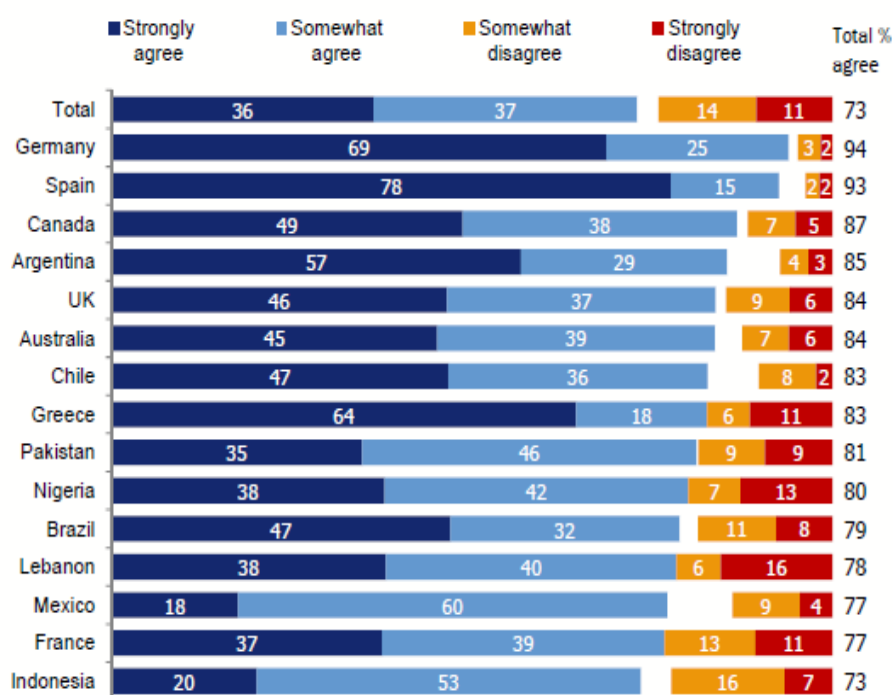
πραγματικά σκέφτονται για τους πρόσφυγες και το βαθμό στον οποίο οι ίδιοι θα αποδέχονταν πρόσφυγες στο σπίτι τους, στη γειτονιά τους, την πόλη / κωμόπολη / το χωριό όπου μένουν ή στη χώρα τους – ή αν θα τους αρνούσαν εντελώς την είσοδο στη χώρα.

Η συντριπτική πλειοψηφία των ερωτηθέντων, και συγκεκριμένα το 80%, δήλωσε ότι τάσσεται υπέρ της υποδοχής των προσφύγων. Θετικά στην παραπάνω ερώτηση απάντησε το 94% των Γερμανών καταλαμβάνοντας τη 2η θέση και το 82% των Ελλήνων καταλαμβάνοντας την 7η θέση (εικόνα 19).

1. People should be able to take refuge in other countries to escape from war or persecution



By Country, 2016



Εικόνα 19. Άποψη πολιτών για την υποδοχή των προσφύγων στη χώρα τους¹⁷⁹

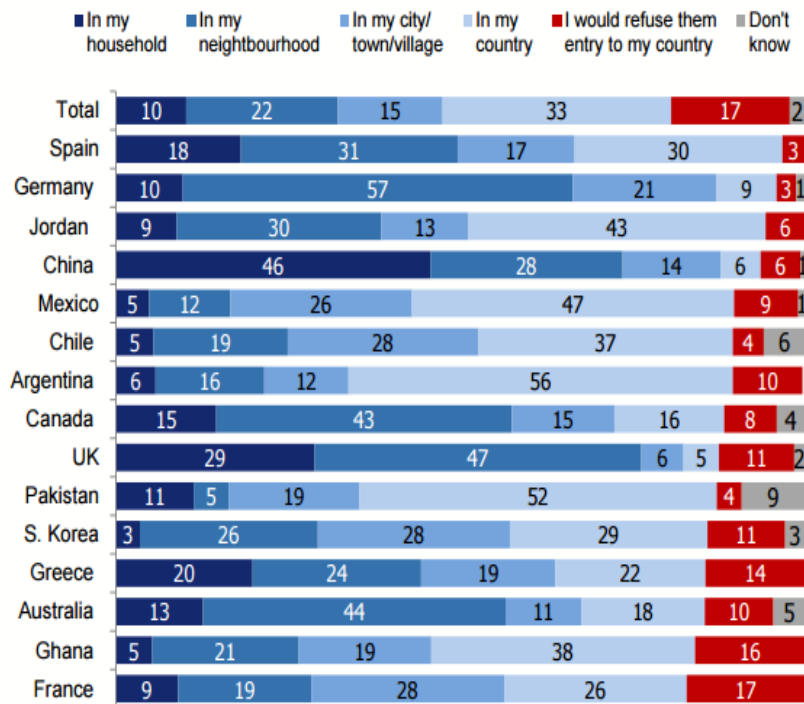
¹⁷⁹ Ό.π. παραπομπή 178, σελ. 7.

Παρόλα αυτά, μόλις το 10% των Γερμανών θα δεχόταν πρόσφυγες μέσα στο ίδιο του το σπίτι, ενώ οι Έλληνες απάντησαν θετικά σε ποσοστό 20% (εικόνα 20).

3. How closely would you personally accept people fleeing war or persecution?



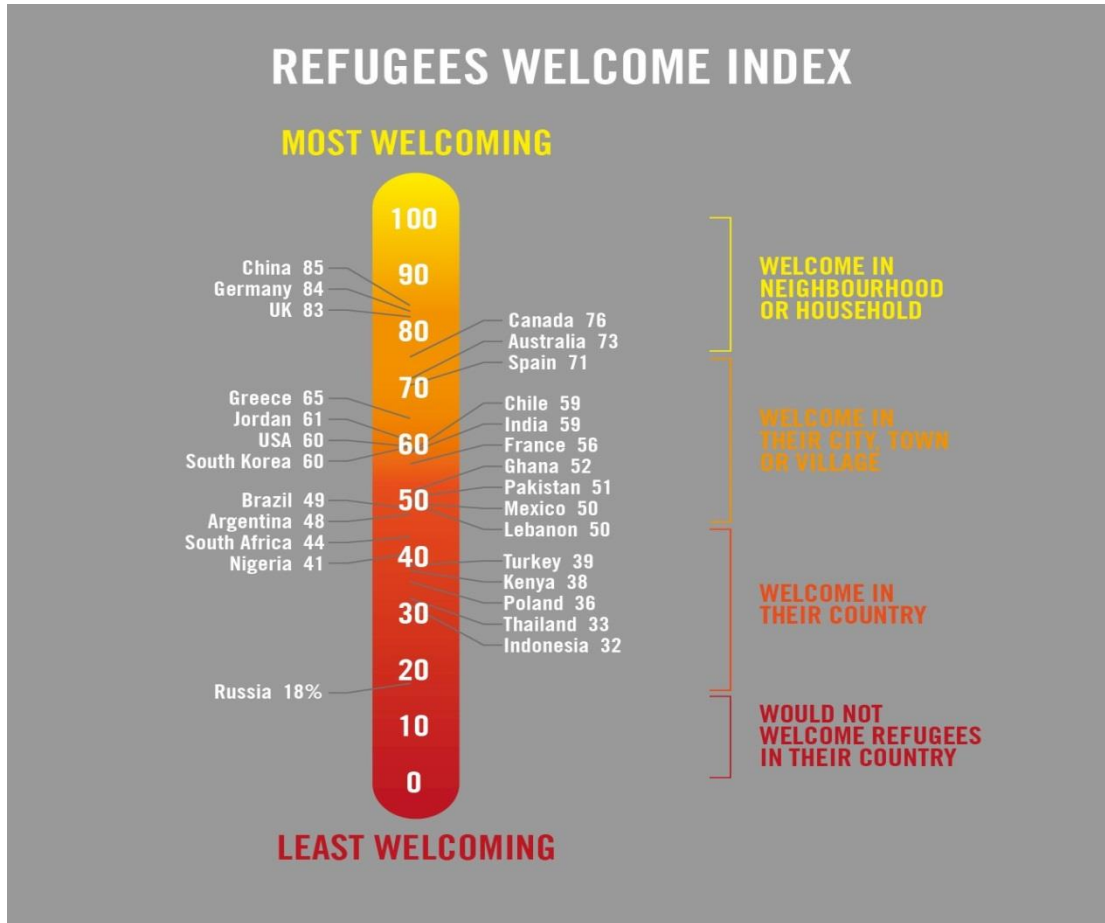
By Country, 2016



Εικόνα 20. Αποψη πολιτών για την υποδοχή προσφύγων στη χώρα/πόλη/γειτονιά/σπίτι τους¹⁸⁰

Με βάση όλες τις απαντήσεις που έδωσαν οι ερωτηθέντες, η Γερμανία κατέλαβε τη 2η θέση στο «Δείκτη Υποδοχής Προσφύγων» της Διεθνούς Αμνηστίας με ποσοστό 84%, ενώ η Ελλάδα την 7η με ποσοστό 65% (εικόνα 21).

¹⁸⁰ Ό.π. παραπομπή 177, σελ. 11.



Εικόνα 21. Δείκτης υποδοχής προσφύγων¹⁸¹

¹⁸¹ Ό.π. παραπομπή 178, σελ. 14.

2. ΤΑ ΚΟΙΝΩΝΙΚΑ ΔΙΚΤΥΑ (SOCIAL MEDIA)

Μία ομάδα ανθρώπων, η οποία συνδέεται από έναν κοινό δεσμό, μπορεί να χαρακτηριστεί ως ένα παραδοσιακό κοινωνικό δίκτυο.¹⁸² Το τι είναι όμως τα νέα μέσα κοινωνικής δικτύωσης είναι δύσκολο να το ορίσει κανείς, καθώς εξελίσσονται διαρκώς με ταχύτατους ρυθμούς και επηρεάζουν παγκοσμίως την καθημερινότητα εκατομμυρίων ανθρώπων.

Αρχικά, θα πρέπει να εστιάσει κανείς στο Web 2.0, το οποίο έπαιξε καθοριστικό ρόλο για τη δημιουργία των μέσων κοινωνικής δικτύωσης. Ο εμπνευστής του Web 2.0, Tim O'Reilly, το ορίζει ως μια δικτυωμένη πλατφόρμα που προσφέρει στο χρήστη τη δυνατότητα να δημιουργεί, να σχεδιάζει, να μοιράζεται περιεχόμενο και υπηρεσίες. Με αυτόν τον τρόπο ο χρήστης, εκτός από καταναλωτής, γίνεται και παραγωγός περιεχομένου. Συγκεκριμένα, ο O'Reilly αναφέρει πως «*το Web 2.0 είναι η επιχειρηματική επανάσταση στη βιομηχανία των υπολογιστών που συντελείται από τη μετακίνηση στο Internet σαν πλατφόρμα, και μια προσπάθεια να κατανοήσουμε τους όρους της επιτυχίας αυτής της νέας πλατφόρμας. Ο βασικότερος ανάμεσα τους όρους είναι ο εξής: Δημιουργήστε εφαρμογές που θα μεγαλώσουν την απήχηση του δικτύου κάνοντας περισσότερο κόσμο να το χρησιμοποιεί*». Επιπλέον, το Web 2.0 ξεπερνάει τις δυνατότητες του Web 1.0, δημιουργώντας δικτυακές επιδράσεις μέσα από μια «*αρχιτεκτονική της συμμετοχής*» και προσφέροντας στο χρήστη πλούσιες εμπειρίες.¹⁸³ Ένα από τα βασικά χαρακτηριστικά του Web 2.0 είναι η συμμετοχική ιδιοφυΐα που εμπεριέχει, καθώς «*ό,τι ανήκει στον χρήστη, προέρχεται από τον χρήστη και κυρίως προορίζεται για τον χρήστη*».¹⁸⁴

Επιστρέφοντας στα μέσα κοινωνικής δικτύωσης, θα μπορούσε να τα χαρακτηρίσει κανείς ως μια κατηγορία μέσων, με τα οποία οι άνθρωποι μιλούν, συμμετέχουν, μοιράζονται, δικτυώνονται, επισημαίνουν ιστοσελίδες, ενώ βρίσκονται στο

¹⁸² Turkat, D., Social networks: Theory and practice, *Journal of Community Psychology*, 1980, 3, σελ. 99-109.

¹⁸³ O'Reilly, T., What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software, *International Journal of Digital Economics*, 2007, 65, σελ. 17-37.

¹⁸⁴ Xu, C., Ouyang, F., & Chu, H., The academic library meets web 2.0: Applications and implications. *The Journal of Academic Librarianship*, 2009, 35 (4), σελ. 324-331.

διαδίκτυο.¹⁸⁵ Βασικό χαρακτηριστικό των μέσων κοινωνικής δικτύωσης είναι ότι ενθαρρύνουν και ενισχύουν την ανατροφοδότηση, τη συμμετοχή σε συζητήσεις και ψηφοφορίες, το σχολιασμό και τη διάχυση της πληροφορίας σε όλα τα ενδιαφερόμενα μέλη και γι' αυτό το λόγο έχει παρατηρηθεί ότι συχνά δίνεται μεγαλύτερη έμφαση στην αλληλεπίδραση, παρά στο ίδιο το περιεχόμενο.¹⁸⁶

Συμπερασματικά, από τους ορισμούς των παραπάνω όρων διαπιστώνεται ότι η μεγάλη σημασία των μέσων κοινωνικής δικτύωσης αφορά στην προσωπική έκφραση, την επικοινωνία, την πρόσβαση, την αλληλεπίδραση με την αξιοποίηση της τεχνολογίας.¹⁸⁷

2.1 ΚΑΤΗΓΟΡΙΟΠΟΙΗΣΗ ΤΩΝ ΜΕΣΩΝ ΚΟΙΝΩΝΙΚΗΣ ΔΙΚΤΥΩΣΗΣ

Τα μέσα κοινωνικής δικτύωσης πολλαπλασιάζονται με γρήγορο ρυθμό και εξαπλώνονται ραγδαία σε πολλούς χώρους, ενώ ταυτόχρονα η χρήση τους εξελίσσεται και αλλάζει. Για τους παραπάνω λόγους υπάρχουν ποικίλες κατηγοριοποιήσεις των εν λόγω μέσων. Μία συνοπτική κατηγοριοποίηση είναι αυτή του Zhang και έχει ως εξής¹⁸⁸:

- **Ιστολόγια (blogging / microblogging):** Ιδιαίτερως δημοφιλή είναι τα Twitter, Blogger και WordPress, στα οποία μπορεί κανείς να ενημερωθεί για ζητήματα με ποικίλο περιεχόμενο, όπως για παράδειγμα απόψεις-σχόλια, αλλά και προσωπικές καταχωρήσεις για κάθε είδους θέμα π.χ. τεχνολογία, μόδα, καλλιτέχνες, εθελοντικές οργανώσεις κ.ά.

- **Κοινωνικά δίκτυα (social networking),** με πιο γνωστά τα Facebook, LinkedIn, Myspace και Ning, τα οποία επιτρέπουν στους χρήστες να αναζητήσουν και να

¹⁸⁵ Jones, R., Social Media Marketing 101, Part 1, 2009, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://searchenginewatch.com/sew/opinion/2064413/social-media-marketing-101-part> , [προσπελάστηκε στις 2 Φεβρουαρίου 2017].

¹⁸⁶ Zhang, J., Social media and distance education, 2010, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://www.scribd.com/document/160525480/Social-Media-and-Distance-Education-De-Oracle>, [προσπελάστηκε στις 5 Φεβρουαρίου 2017].

¹⁸⁷ Μανούσου, Ε., Χαρτοφύλακα, Τ., Κοινωνικά δίκτυα και μέσα κοινωνικής δικτύωσης στην εξ αποστάσεως τριτοβάθμια εκπαίδευση, 2^ο Πανελλήνιο Συνέδριο, Πάτρα, 28-30/4/2011, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.etpe.gr/custom/pdf/etpe1730.pdf>, [προσπελάστηκε στις 28 Φεβρουαρίου 2017], σελ. 499.

¹⁸⁸ Zhang, J., Social media and distance education, ό.π. παραπομπή 185.

συνδεθούν με άλλους χρήστες με παρόμοια ενδιαφέροντα και χαρακτηριστικά, π.χ. φίλους, συνεργάτες κ.ά. και να επικοινωνήσουν άμεσα μεταξύ τους, να ενημερωθούν, να ανταλλάξουν πληροφορίες, αρχεία, σκέψεις κ.ά.

- **Κοινωνική σελιδοσήμανση (social bookmarking)**, όπως τα: Delicious, Digg, Faves, StumbleUpon, BlogMarks, τα οποία επιτρέπουν στον χρήστη να επισημαίνει τις ιστοσελίδες που τον ενδιαφέρουν και στη συνέχεια να τις μοιράζεται με τους ανθρώπους που επιθυμεί.

- **Συνεργατική συγγραφή (collaborative authoring)**, με πιο γνωστά τα Wikipedia, Google docs και Zoho office suite.

- **Διαμοιρασμός πολυμέσων (multimedia sharing)**, όπως είναι τα: YouTube, Flickr, Qik, Snapfish, Jumpcut, Vimeo, Decianart, με τα οποία οι χρήστες δημιουργούν και διαμοιράζονται αρχεία ήχου, εικόνας και video. Να σημειωθεί ότι σε αυτή την κατηγορία εξαιρετικά δυναμική παρουσία έχουν το Instagram και το Snapchat από το 2010 και το 2011 αντίστοιχα.

- **Διαδικτυακές τηλεδιασκέψεις (web conferencing)**, όπως τα: WebEx, GoToMeeting, DimDim.

- **Μέσα κοινωνικών ειδήσεων (newssite ή social news)** με δικτυακούς τόπους, όπως τα Digg, Sphinn, Newsvine, και BallHype που δίνουν τη δυνατότητα στους χρήστες να ενημερωθούν για ζητήματα της επικαιρότητας και στη συνέχεια τους δίνεται η δυνατότητα να τα ψηφίσουν ή να τα σχολιάσουν. Η συμμετοχή των χρηστών στα άρθρα είναι αυτή που αλλάζει και τη θέση τους στο site, καθώς προωθούνται τα πιο δημοφιλή από αυτά.

2.2 TO TWITTER

Το Twitter, το οποίο αποτελεί αντικείμενο μελέτης της παρούσας εργασίας, αποτελεί ένα μέσο κοινωνικής δικτύωσης, το οποίο φαίνεται να κερδίζει έδαφος τα τελευταία χρόνια. Το Twitter ιδρύθηκε το Μάρτιο του 2006 από τον Jack Dorsey και αποτελεί ένα μέσο online κοινωνικής δικτύωσης και microblogging, το οποίο δίνει τη δυνατότητα στους χρήστες του να βρίσκονται σε επαφή και να ενημερώνονται για τις καθημερινές δραστηριότητες των ανθρώπων που τους ενδιαφέρουν, αλλά και να ενημερώνουν οι ίδιοι το δικό τους κοινό. Η επικοινωνία πραγματοποιείται μέσω μικρών κειμένων, έως 140 χαρακτήρων, τα οποία αναρτούν οι χρήστες στο προσωπικό τους προφίλ, με στόχο να μοιραστούν τους προβληματισμούς, τις ιδέες, τις δραστηριότητες, τις απορίες τους ή απλώς να εκφράσουν αυτό που σκέφτονται με όσους έχουν επιλέξει να τους ακολουθούν μέσω της υπηρεσίας.

Ένα χαρακτηριστικό του Twitter που το κάνει ιδιαιτέρως δημοφιλές είναι το γεγονός ότι, λόγω του περιορισμού της έκτασης των δημοσιεύσεων με το όριο 140 χαρακτήρων ανά δημοσίευση, καθίσταται σε μεγάλο βαθμό άμεσο και κατάλληλο για σχολιασμό γεγονότων, ακόμα και σε πραγματικό χρόνο. Καθώς αποτελεί μια ζωντανή μηχανή αναζήτησης, η οποία ανανεώνεται διαρκώς με αναρτήσεις συνεχούς ροής από τους ίδιους τους χρήστες, υπάρχει η δυνατότητα για τον καθένα να ενημερώνεται ανά πάσα στιγμή ή και να αναμεταδίδει ό,τι τον ενδιαφέρει.¹⁸⁹

Σύμφωνα με το trending.gr, που αναλύει καθημερινά τα tweets των Ελλήνων, υπάρχουν αυτή τη στιγμή πάνω από 600.000 Έλληνες χρήστες στο Twitter.¹⁹⁰ Επιπλέον, ο ένας στους τέσσερις χρήστες του internet δηλώνουν ότι παρακολουθούν τις εξελίξεις της επικαιρότητας μέσω του συγκεκριμένου μέσου. Σύμφωνα με το 45% των χρηστών του Twitter στην Ελλάδα, ένα γεγονός, το οποίο ενίσχυσε τη χρήση του εν λόγω μέσου και τους ώθησε να εξοικειωθούν με αυτό, ήταν η εμφάνιση των capital controls το καλοκαίρι του 2015. Κατά τη διάρκεια της συγκεκριμένης περιόδου, οι χρήστες επιθυμούσαν να ενημερώνονται άμεσα για τις εξελίξεις, από διαφορετικές

¹⁸⁹ Σιδέρη, Μ., *Το βιβλίο του twitter - Ένας οδηγός για "αθώους" χρήστες*, Κλειδάριθμος, Αθήνα, 2011, σελ. 24.

¹⁹⁰ Βλ. στη διεύθυνση <http://trending.gr>, στοιχεία από τις 29/5/2017.

πηγές, προκειμένου να έχουν τη δυνατότητα να σχηματίσουν την προσωπική τους άποψη. Γρήγορα αντιλήφθηκαν, ότι το Twitter είναι το μέσο που μπορούσε να ικανοποιήσει τη συγκεκριμένη ανάγκη τους.¹⁹¹

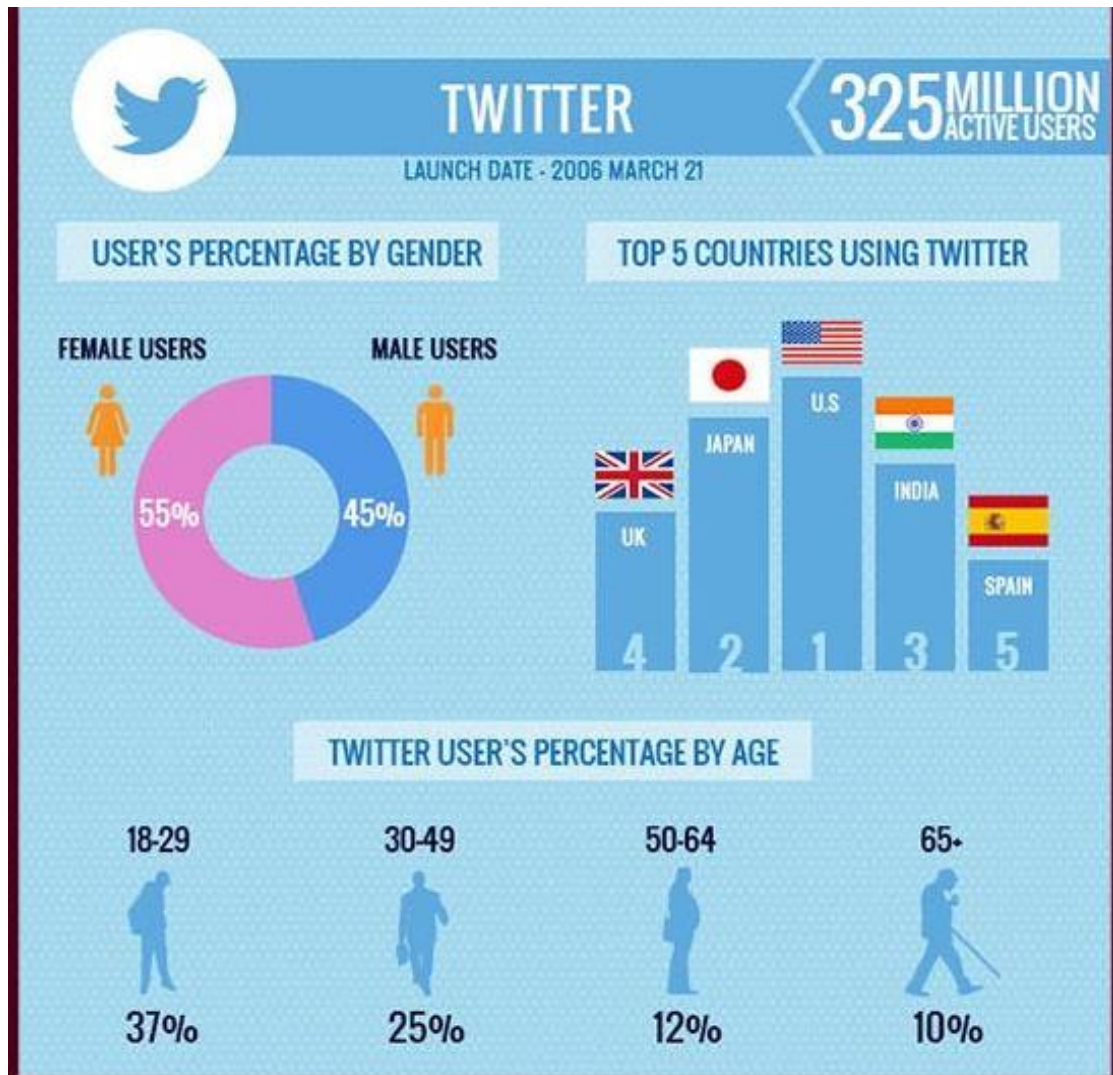
Στη Γερμανία, το ίδιο το Twitter ανακοίνωσε στο τέλος του 2016, ότι 12 εκατομμύρια Γερμανοί χρησιμοποιούν την εν λόγω υπηρεσία. Στο συγκεκριμένο αριθμό όμως συμπεριλαμβάνονται τόσο οι εγγεγραμμένοι χρήστες του Twitter στη Γερμανία, αλλά και αυτοί που μπαίνουν μόνο για να διαβάσουν tweets, δε διαθέτουν προσωπικό προφίλ και επομένως δεν κάνουν αναρτήσεις. Με βάση μία έρευνα, οι εγγεγραμμένοι χρήστες του συγκεκριμένου μέσου στη Γερμανία υπολογίζονται στα 3 εκατομμύρια, εκ των οποίων το 1 εκατομμύριο είναι ενεργό.¹⁹²

Σε παγκόσμιο επίπεδο, το Twitter, σύμφωνα με στοιχεία του Μαρτίου του 2016 (εικόνα 22), φαίνεται να έχει 325 εκατομμύρια ενεργούς χρήστες, εκ των οποίων το 55% είναι γυναίκες και το 45% άντρες. Όπως παρουσιάζεται στην παρακάτω εικόνα, η πλειονότητα των χρηστών και συγκεκριμένα το 37% βρίσκεται στην ηλικιακή ομάδα των 18-29, ενώ ακολουθούν οι χρήστες ηλικίας 30-49 (25%), 50-64 (12%) και τέλος οι άνω των 65 (10%). Παρόμοια εικόνα, όσον αφορά στις ηλικίες, παρουσιάζουν τα περισσότερα μέσα κοινωνικής δικτύωσης.¹⁹³ Επομένως, επειδή το Twitter είναι σε αρκετά μεγάλο βαθμό δημοφιλές, μπορεί να φανεί σαν μια καλή αντανάκλαση του τι συμβαίνει σε όλο τον κόσμο.

¹⁹¹ ELTRUN, Πιο ώριμοι και συνειδητοποιημένοι δηλώνουν οι Internet users στην Ελλάδα, όσον αφορά την χρήση των Social Media και των σχετικών εργαλείων digital marketing από τις επιχειρήσεις, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: www.greekecommerce.gr/gr/file-download/eltrun-digital-marketing-2016, [προσπελάστηκε στις 5 Φεβρουαρίου 2017], σελ. 3.

¹⁹² Buggisch, C., Social Media und Messenger – Nutzerzahlen in Deutschland 2017, 2017, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://buggisch.wordpress.com/2017/01/02/social-media-und-messenger-nutzerzahlen-in-deutschland-2017/>, [προσπελάστηκε στις 5 Μαρτίου 2017].

¹⁹³ Kroker, M., Social Media 2016: Die Zahlen & Fakten zu Facebook, Twitter & Co. im Überblick, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://blog.wiwo.de/look-at-it/2016/06/30/social-media-2016-die-zahlen-fakten-zu-facebook-twitter-co-im-ueberblick/>, [προσπελάστηκε στις 5 Φεβρουαρίου 2017].



Εικόνα 22. Οι χρήστες του Twitter σε παγκόσμιο επίπεδο¹⁹⁴

Παρακάτω παρουσιάζονται συνοπτικά ορισμένα από τα βασικότερα συστατικά του Twitter:

- **Tweet:** Το tweet αποτελεί τον πυρήνα αυτής της πλατφόρμας, καθώς είναι ο μόνος τρόπος δημοσίευσης και μεταφοράς πληροφορίας ανάμεσα στους χρήστες, με εξαίρεση τα προσωπικά μηνύματα. Ουσιαστικά αποτελεί την ανάρτηση του κάθε χρήστη, που μπορεί να έχει τη μορφή σχολίου, άποψης ή ακόμα και φωτογραφίας ή βίντεο. Βασικά του χαρακτηριστικά είναι ο περιορισμός στο μέγεθός του και η ιδιαίτερη δομή του.

¹⁹⁴ Βλ. <http://blog.wiwo.de/look-at-it/2016/06/30/social-media-2016-die-zahlen-fakten-zu-facebook-twitter-co-im-ueberblick/>

- **Follow:** Ο χρήστης μπορεί να «ακολουθήσει» τον λογαριασμό όσων χρηστών επιθυμεί, με αποτέλεσμα να ενημερώνεται άμεσα κάθε φορά που οι χρήστες αυτοί αναρτούν ένα tweet.
- **Retweet:** Ένα άλλο διαδεδομένο σύμβολο που εμφανίζεται συχνά στην αρχή των tweets, είναι το «RT», το οποίο σημαίνει «retweet». Ο χρήστης κάνοντας «retweet» αναδημοσιεύει στο προσωπικό του προφίλ μία ανάρτηση άλλου χρήστη που του φάνηκε ενδιαφέρουσα και τον εκφράζει. Με αυτό τον τρόπο τη γνωστοποιεί και στους δικούς του ακολούθους.
- **Favorite:** Με το «favorite» ή αλλιώς «like» ο χρήστης εκφράζει την επιδοκιμασία του σε ένα tweet με το οποίο συμφωνεί.
- **Reply:** Με αυτή την ενέργεια ο χρήστης απαντάει σε ένα tweet σχολιάζοντάς το.
- **Mention:** Με το σύμβολο «@» ακολουθούμενο από το όνομα ενός χρήστη, συμπεριλαμβάνεται το εν λόγω άτομο σε ένα tweet, καθώς απευθύνεται σε αυτό ή το αφορά.
- **Hashtag:** Το «hashtag» που απεικονίζεται με το σύμβολο της δίσσης (#) δημιουργεί ένα θέμα συζήτησης μεταξύ των χρηστών και με τη χρήση του γίνεται μια αναφορά σε κάποιο θέμα (topic). Τα hashtags είναι εμπνευσμένα από τους ίδιους τους χρήστες, προκειμένου να κατηγοριοποιούν τα tweets και έτσι να γίνονται άμεσα ορατά σε όλους τους χρήστες που ενδιαφέρονται για το εκάστοτε θέμα, χωρίς απαραίτητα να ακολουθούν ο ένας τον άλλον¹⁹⁵. Συνεπώς με αυτόν τον τρόπο προσφέρεται η δυνατότητα ομαδοποίησης πολλών δημοσίων tweets γύρω από ένα συγκεκριμένο θέμα και καθίσταται εύκολη η αναζήτηση πληροφοριών σχετικά με το θέμα αυτό.

¹⁹⁵ Wenger, E., White, N., & Smith, J.D., *Digital Habitats: Stewarding Technologies for Communities*, CPsquare, Portland, 2009, σελ. 17.

- **Search:** Το Twitter περιλαμβάνει μία μηχανή αναζήτησης, στην οποία ο χρήστης εισάγοντας λέξεις-κλειδιά μπορεί να βρει θέματα και συζητήσεις που τον ενδιαφέρουν.
- **URLs:** Αρκετά συχνή είναι και η χρήση urls στα tweets, καθώς εκτός από την επιπρόσθετη αναφορά σε κάτι, δίνει και μια λύση στην περιορισμένη δυνατότητα ανάλυσης ενός θέματος εξαιτίας του περιορισμού στο πλήθος των χαρακτήρων.

2.3 ΤΟ TWITTER ΓΙΑ ΤΟ ΠΡΟΣΦΥΓΙΚΟ ΖΗΤΗΜΑ

Στο Twitter τα τελευταία χρόνια γίνεται σε μεγάλο βαθμό λόγος για την προσφυγική κρίση, καθώς αποτελεί μία από τις μεγαλύτερες και σκληρότερες ανθρωπιστικές κρίσεις που έχει συγκλονίσει όλη την υφήλιο. Σύμφωνα με έρευνα¹⁹⁶, από την αρχή του εμφυλίου πολέμου στη Συρία το 2011 και μέχρι το τέλος του 2015 είχαν αναρτηθεί στο Twitter πάνω από 20 εκατομμύρια δημοσιεύσεις στα Αγγλικά γύρω από το προσφυγικό. Ο μεγαλύτερος όγκος των εν λόγω δημοσιεύσεων προέρχεται από το Ηνωμένο Βασίλειο, την Τουρκία, τη Γαλλία, τη Γερμανία και από τα Αραβικά κράτη, όπως τον Λίβανο, την Αίγυπτο, τα Ηνωμένα Αραβικά Εμιράτα, την Ιορδανία και τη Σαουδική Αραβία.

Το Twitter επιλέχθηκε από το BeyondCSRNet¹⁹⁷ ως μέσο για την πρώτη εκστρατεία Δημόσιας Διπλωματίας στα social media για την προσφυγική κρίση. Η καμπάνια με θέμα «*Πώς να βρούμε λύση για την προσφυγική κρίση*» έλαβε χώρα από τις 14 Μαρτίου έως τις 8 Απριλίου του 2016. Κατά τη διάρκεια της υλοποίησης της συγκεκριμένης ιδέας, οι συμμετέχοντες κλήθηκαν να μοιραστούν τις σκέψεις τους, είτε γραπτώς, είτε δημιουργώντας κάποιο βίντεο ή φωτογραφία. Το υλικό που συλλέχθηκε, αξιοποιήθηκε και με βάση αυτό συντάχθηκε μία αναφορά, η οποία ήταν η πρώτη που βασίστηκε σε

¹⁹⁶ Crimson Hexagon, The Syrian Refugee Crisis on Social Media, 2015, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://partners.twitter.com/content/dam/partners-twitter/success-stories/pdf/Syrian%20Refugee%20Crisis%20Report%20-%20Partnerships.pdf>, [προσπελάστηκε στις 2 Φεβρουαρίου 2017], σελ. 3.

¹⁹⁷ Βλ. <http://beyondcsr.org/>,

δημόσιο διάλογο και πρότεινε λύσεις για την αντιμετώπιση του προσφυγικού. Με τη χρήση hashtags και συγκεκριμένα των #refolution, #refducatation και #globalsouth, πολίτες, διεθνείς οργανισμοί, δημόσιοι και ιδιωτικοί φορείς, Μ.Κ.Ο., εταιρίες, η πανεπιστημιακή κοινότητα και τα Μ.Μ.Ε. ενεπλάκησαν σε έναν διάλογο γύρω από τα προβλήματα και τις πιθανές λύσεις για την προσφυγική κρίση¹⁹⁸. Κατά τη διάρκεια της δράσης, το βασικό hashtag της εκστρατείας #refolution έγινε ορατό σε ένα κοινό μεγέθους 11,8 εκατομμυρίων, ενώ βρέθηκε στην πεντάδα των δημοφιλέστερων hashtags σχετικά με την προσφυγική κρίση.¹⁹⁹

Το Twitter αποτελεί ένα ψηφιακό εργαλείο επικοινωνίας, με το οποίο οι χρήστες ασκούν κριτική, συζητούν, καταγγέλλουν, αλληλεπιδρούν και μοιράζονται με άλλους χρήστες απόψεις σχετικά με συγκεκριμένα θέματα και γεγονότα. Επιπλέον, ξεχωρίζει χάρη στη δυνατότητα που έχει να μεταφέρει μηνύματα γρήγορα και άμεσα, να «γεννάει» δημόσιες συζητήσεις, να κινητοποιεί και να συνδέει χρήστες μέσω των hashtags ομαδοποιώντας τα κοινά θέματα που συζητούν.²⁰⁰ Επομένως, το Twitter προσφέρει έναν νέο τύπο δεδομένων, τα οποία παράγονται σε πολύ μεγάλο όγκο και κυρίως κατευθύνονται από τα πραγματικά γεγονότα και όχι από τις αποφάσεις του εκάστοτε επιστήμονα. Η έρευνα στο Twitter δεν έχει μεγάλη ιστορία, ούτε συλλογική μνήμη για το τι λειτουργεί και τι όχι. Χρησιμοποιεί όμως νέες τεχνολογίες σε νέους δρόμους, οι οποίοι μπορεί να είναι ασυνήθεις στις κοινωνικές επιστήμες, αλλά πολλές φορές παρέχουν νέες και σημαντικές εφαρμογές για έρευνα²⁰¹. Όλα τα παραπάνω, σε συνδυασμό με την στροφή των ανθρώπων από τα παραδοσιακά μέσα στο διαδικτυακό περιεχόμενο και στα μέσα κοινωνικής δικτύωσης, τα οποία παρέχουν έναν νέο, δυναμικό και λιγότερο ιεραρχικό χώρο, κατάφεραν να ανοίξουν την πόρτα για την παραγωγή συζητήσεων γύρω από θέματα, όπως αυτό της προσφυγικής κρίσης.²⁰²

¹⁹⁸ Βλ. <http://beyondcsr.org/refugees-call/>

¹⁹⁹ BeyondCSRNet, Twitter resolution call for refugees- How to solve the refugee crisis, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://beyondcsr.org/wp-content/uploads/2017/02/BEYONDRESOLUTION.pdf>, [προσπελάστηκε στις 2 Ιανουαρίου 2017], σελ. 14.

²⁰⁰ Gualda E. et al., The Refugee Crisis on Twitter: A Diversity of Discourses at a European Crossroads, *Journal of Spatial and Organizational Dynamics*, 2016, 4 (3), σελ. 200.

²⁰¹ Bartlett, J., & Norrie, R., Immigration on twitter- understanding public attitudes online, 2015, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: https://www.demos.co.uk/files/immigration_on_twitter.pdf?1428506056, [προσπελάστηκε στις 2 Μαρτίου 2017], σελ. 30.

²⁰² Mohamud, A., 2013 UK Digital Future in Focus, 2013, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: http://www.comscore.com/Insights/Presentations-and-Whitepapers/2013/2013-UK-Digital-Future-in-Focus?cs_edgescape_cc=GR, [προσπελάστηκε στις 21 Μαρτίου 2017].

3. ΕΞΟΡΥΞΗ ΔΕΔΟΜΕΝΩΝ ΚΑΙ ΑΝΑΛΥΣΗ ΣΥΝΑΙΣΘΗΜΑΤΟΣ

Το πλήθος των δεδομένων που είναι αυτή τη στιγμή διαθέσιμα για έρευνα είναι τεράστιο και αυξάνεται εκθετικά κάθε μέρα. Παράλληλα, μειώνεται διαρκώς η δυσκολία στη συλλογή και αποθήκευση των δεδομένων, αλλά και το κόστος συλλογής τους, με αποτέλεσμα να αναπτύσσεται ραγδαία το πεδίο της εξόρυξης δεδομένων. Παρόλα αυτά, ο τεράστιος όγκος των δεδομένων που συσσωρεύεται και συλλέγεται, δεν μπορεί να αξιοποιηθεί έτσι όπως είναι. Αρχικά πρέπει να γίνουν κάποιες ενέργειες για να δομηθούν κατάλληλα τα δεδομένα, ώστε στη συνέχεια να μπορούν να αξιοποιηθούν. Όπως λέγεται χαρακτηριστικά, μπορεί να υπάρχει πληθώρα δεδομένων, αλλά απουσιάζει η πληροφορία, η οποία θα πρέπει να είναι κατανοητή και αξιοποιήσιμη για τον άνθρωπο.²⁰³

3.1 ΕΞΟΡΥΞΗ ΔΕΔΟΜΕΝΩΝ

Αρχικά, ο όρος «εξόρυξη δεδομένων» ή «αποκάλυψη δεδομένων» ήταν ένας μειωτικός όρος που αναφερόταν στις προσπάθειες να εξαχθεί πληροφορία που δεν υποστηριζόταν από τα δεδομένα.²⁰⁴ Η εξόρυξη γνώσης από δεδομένα (data mining) αποτελεί πλέον την εξεύρεση πληροφορίας ή προτύπων, από μεγάλες βάσεις δεδομένων, με τη χρήση αλγορίθμων, των αρχών της στατιστικής, της τεχνητής νοημοσύνης, της μηχανικής μάθησης (machine learning) και των συστημάτων βάσεων δεδομένων.²⁰⁵ Ο πλέον αποδεκτός ορισμός της «εξόρυξης δεδομένων» είναι η ανακάλυψη των «μοντέλων» των δεδομένων.²⁰⁶

²⁰³ Κύρκος, Ε., *Επιχειρηματική ευφυΐα και εξόρυξη δεδομένων*, Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών, 2015, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://hdl.handle.net/11419/1226>, [προσπελάστηκε στις 8 Ιουνίου 2017], σελ. 127.

²⁰⁴ Rajaraman, A., Ullman, J., *Εξόρυξη από μεγάλα σύνολα δεδομένων*, μετάφραση: Αναστάσιος Αθ. Γούναρης, Ιωάννης Μανωλόπουλος, Απόστολος Ν. Παπαδόπουλος, Κωνσταντίνος Τσίχλας, Εκδόσεις Νέων Τεχνολογιών, 2014, σελ. 2.

²⁰⁵ Hand, D. J., Mannila, H., & Smyth, P., *Principles of Data Mining*, The MIT Press, London, 2001, σελ. 27.

²⁰⁶ Rajaraman, A., Ullman, J., *Εξόρυξη από μεγάλα σύνολα δεδομένων*, ό.π. παραπομπή 204, σελ. 1.

Το πεδίο της εξόρυξης γνώσης από δεδομένα γνωρίζει μεγάλη άνθιση στη σύγχρονη εποχή, τόσο σε ερευνητικό επίπεδο, όσο και σε επίπεδο πρακτικής εφαρμογής. Με το data mining γίνεται προσπάθεια, ώστε η πληροφορία που θα εξαχθεί και τα πρότυπα που θα προκύψουν να έχουν δομή κατανοητή από τον άνθρωπο, έτσι ώστε να τον βοηθήσουν να πάρει τις κατάλληλες αποφάσεις.²⁰⁷ Συγκεκριμένα, στόχος της εξόρυξης δεδομένων είναι η αυτόματη ή ημιαυτόματη ανάλυση μεγάλων δεδομένων (big data) για την εξαγωγή κάποιου ενδιαφέροντος προτύπου που ήταν άγνωστο μέχρι εκείνη τη στιγμή.²⁰⁸ Με την εξόρυξη δεδομένων θα μπορούσαν να προσδιοριστούν πολλαπλά σύνολα στα δεδομένα, τα οποία θα μπορούσαν να χρησιμοποιηθούν για να εξασφαλίσουν περισσότερο ακριβή αποτελέσματα από ένα σύστημα υποστήριξης αποφάσεων.²⁰⁹

3.1.1 Η ΔΙΑΔΙΚΑΣΙΑ ΕΞΟΡΥΞΗΣ ΔΕΔΟΜΕΝΩΝ

Η εξόρυξη δεδομένων συχνά συναντάται ως συνώνυμο με έναν άλλο όρο, αυτόν της ανακάλυψης γνώσης από δεδομένα (knowledge discovery from data). Η εξόρυξη δεδομένων θεωρείται ως ένα σημαντικό βήμα στη διαδικασία της ανακάλυψης γνώσης, η οποία αποτελείται από την επαναληπτική ακολουθία των παρακάτω βημάτων (εικόνα 23):

1. Καθαρισμός δεδομένων (data cleaning), κατά τον οποίο απομακρύνεται ο θόρυβος, τα εσφαλμένα δεδομένα και οι περιττές πληροφορίες.

2. Ενσωμάτωση δεδομένων (data integration), όπου μπορούν να συνδυαστούν πολλαπλές πηγές δεδομένων.

²⁰⁷ Απειρανθίτης, Α., Τσαχανσάχης, Β., Ανάλυση Δεδομένων από Κοινωνικά Δίκτυα για τη Μελέτη των Συναισθημάτων των Χρηστών, Διπλωματική εργασία, Εθνικό Μετσόβιο Πολυτεχνείο, Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://dspace.lib.ntua.gr/bitstream/handle/123456789/43734/Diploma%20Thesis.pdf?sequence=1>, [προσπελάστηκε στις 24 Ιανουαρίου 2017], σελ . 20.

²⁰⁸ Ό.π., σελ . 64.

²⁰⁹ Ό.π., παραπομπή 206, σελ . 65.

3. Συλλογή δεδομένων (data selection), όπου τα δεδομένα που σχετίζονται με την διαδικασία της ανάλυσης ανακτώνται από τη βάση δεδομένων.

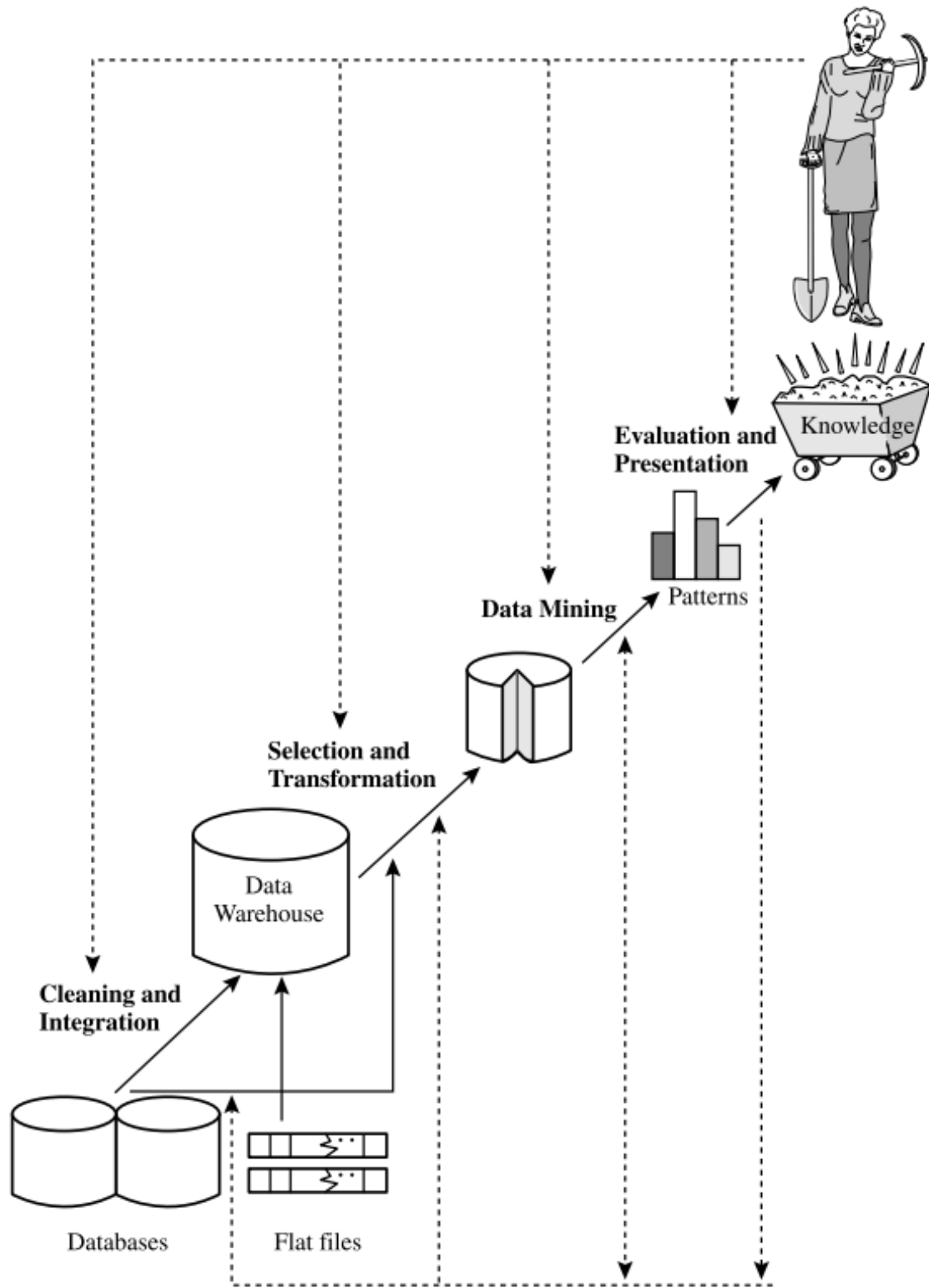
4. Μετασχηματισμός δεδομένων (data transformation), όπου τα δεδομένα μετασχηματίζονται ή ενοποιούνται σε μορφές κατάλληλες για εξόρυξη.

5. Εξόρυξη δεδομένων (data mining), μια ουσιαστική διαδικασία, όπου ευφυείς μέθοδοι εφαρμόζονται προκειμένου να εξαχθούν μοτίβα δεδομένων.

6. Αξιολόγηση μεθόδου (pattern evaluation), για τον εντοπισμό των πραγματικά ενδιαφερόντων μεθόδων που αναπαριστούν τη γνώση με βάση κάποια αξιολογικά μέτρα.

7. Παρουσίαση της γνώσης (knowledge presentation), όπου οι τεχνικές απεικόνισης και αναπαράστασης της γνώσης χρησιμοποιούνται για να παρουσιάσουν την εξορυσσόμενη γνώση στο χρήστη.²¹⁰

²¹⁰ Han, J., & Kamber, M., *Data Mining: Concepts and Techniques, 2nd Ed.*, Morgan Kaufmann Publishers, U.S.A., 2006, σελ. 7.



Εικόνα 23. Διαδικασία εξόρυξης δεδομένων²¹¹

²¹¹ Ό.π. παραπομπή 210, σελ. 6.

3.1.2 ΔΙΕΡΓΑΣΙΕΣ ΕΞΟΡΥΞΗΣ ΔΕΔΟΜΕΝΩΝ

Στο στάδιο της εξόρυξης δεδομένων ο αναλυτής εξάγει πρότυπα από το σύνολο δεδομένων που έχει δημιουργήσει. Ανάλογα με το είδος του προβλήματος και τον τύπο της ανάλυσης που θα διεξάγει, μπορεί να αναζητήσει πρότυπα διαφόρων τύπων. Οι εργασίες εξόρυξης δεδομένων που υπάρχουν είναι ποικίλες, με βασικότερες τις ακόλουθες:²¹²

- **Προβλεπτικού τύπου ή μοντέλου**
 - **Κατηγοριοποίηση (Classification)**
Η κατηγοριοποίηση αποτελεί εργασία επιβλεπόμενης μάθησης, είναι μία από τις συνηθέστερες εργασίες εξόρυξης δεδομένων και θα εκτελεστεί και στην παρούσα μελέτη. Σε προβλήματα κατηγοριοποίησης είναι γνωστό εκ των προτέρων ότι τα δεδομένα υπάγονται σε κατηγορίες και σε ένα από τα γνωρίσματα καταγράφεται η κατηγορία των αντικειμένων.
 - **Παλινδρόμηση (Regression)**
Η παλινδρόμηση είναι μια εργασία επιβλεπόμενης μάθησης που μοιάζει πολύ με την κατηγοριοποίηση. Υπάρχει πάλι ένα γνώρισμα-στόχος, οι τιμές του οποίου υπολογίζονται από τα υπόλοιπα γνωρίσματα και οι αλγόριθμοι παλινδρόμησης εξετάζουν τις σχέσεις μεταξύ του γνωρίσματος-στόχου και των υπόλοιπων γνωρισμάτων κατασκευάζοντας έναν μηχανισμό υπολογισμού. Η διαφορά με την κατηγοριοποίηση είναι ότι στην περίπτωση της παλινδρόμησης υπολογίζονται αριθμητικές τιμές.
 - **Ανάλυση χρονοσειρών (Time series analysis)**
Υπάρχουν μεγέθη τα οποία παρουσιάζουν μια χρονική εξέλιξη. Η εξέλιξη αυτή αναπαρίσταται με τη βοήθεια χρονοσειρών, δηλαδή ακολουθιών σημείων που αποτελούν μετρήσεις του μεγέθους στη διάρκεια του

²¹² Κύρκος, Ε., *Επιχειρηματική ευφυΐα και εξόρυξη δεδομένων*, ό.π. παραπομπή 203, σελ. 132.

χρόνου. Οι μέθοδοι ανάλυσης χρονοσειρών αναλύουν τα δεδομένα διαφορετικών χρονικών περιόδων και εξάγουν χρήσιμα συμπεράσματα για το φαινόμενο.

- **Περιγραφικού τύπου ή μοντέλου**

- **Συσταδοποίηση (Clustering)**

Η ανάλυση συστάδων είναι και αυτή μια πολύ συνηθισμένη εργασία εξόρυξης δεδομένων και ανήκει στην κατηγορία εργασιών μη επιβλεπόμενης μάθησης. Στόχος της ανάλυσης συστάδων είναι ο επιμερισμός ενός συνόλου αντικειμένων σε ομάδες. Η διαφορά με την κατηγοριοποίηση έγκειται στο γεγονός ότι δεν υπάρχουν κατηγορίες, οι οποίες είναι εκ των προτέρων γνωστές, δεν υπάρχει δηλαδή ένα γνώρισμα στο οποίο καταγράφεται η κατηγορία των αντικειμένων. Οι αλγόριθμοι της ανάλυσης συστάδων προσπαθούν να δημιουργήσουν ομάδες αναζητώντας ομοειδή αντικείμενα. Στόχος τους είναι να μεγιστοποιήσουν την ομοιότητα εντός των ομάδων και την ανομοιότητα μεταξύ των διαφορετικών ομάδων. Αυτό σημαίνει ότι τα αντικείμενα της μιας ομάδας θα πρέπει να μοιάζουν μεταξύ τους και να μη μοιάζουν με τα αντικείμενα των άλλων ομάδων. Αφού σχηματιστούν οι ομάδες, μπορούν να θεωρηθούν ως κατηγορίες και να δημιουργηθούν κανόνες που να τις περιγράφουν.

- **Εύρεση Κανόνων Συσχέτισης (Association Rules)**

Η ανάλυση κανόνων συσχέτισης θεωρείται το πιο γνήσιο τέκνο της εξόρυξης δεδομένων, καθώς οι άλλες μέθοδοι προέρχονται από τη μηχανική μάθηση, τη στατιστική κλπ. Στόχος των κανόνων συσχέτισης είναι η ανακάλυψη σχέσεων μεταξύ τιμών των γνωρισμάτων, οι οποίες εμφανίζονται συχνά μαζί.

3.2 ΕΞΟΡΥΞΗ ΓΝΩΜΗΣ ΚΑΙ ΑΝΑΛΥΣΗ ΣΥΝΑΙΣΘΗΜΑΤΟΣ

Η «ανάλυση συναισθήματος» (sentiment analysis) και η «εξόρυξη γνώμης» (opinion mining) είναι ερευνητικοί τομείς στους οποίους με τη χρήση τεχνικών, όπως είναι αυτές της επεξεργασίας φυσικής γλώσσας (natural language processing), της ανάλυσης κειμένου (text analysis) και της υπολογιστικής γλωσσολογίας (computational linguistics), αναγνωρίζονται και εξάγονται υποκειμενικές πληροφορίες που έχουν παραχθεί από τον άνθρωπο. Ουσιαστικά, πρόκειται για υπολογιστικές μελέτες για τις ανθρώπινες απόψεις, τα συναισθήματα, τις αξιολογήσεις και τις στάσεις προς οντότητες (entities), όπως προϊόντα, υπηρεσίες, οργανισμούς, πρόσωπα, εκδηλώσεις, θεματικές ενότητες, αλλά και προς τις ιδιότητες των οντοτήτων αυτών. Σκοπός της εξαγωγής και της ανάλυσης απόψεων είναι το να προσδιοριστεί η στάση ενός ατόμου σε σχέση με κάποιο θέμα. Η εν λόγω στάση μπορεί να αποτελεί την κρίση ή την αξιολόγηση του ομιλητή ή συγγραφέα, τη συναισθηματική του κατάσταση την εκάστοτε στιγμή, καθώς και του συναισθηματικού αντικτύπου που επιθυμεί να έχει στον ακροατή ή αναγνώστη.²¹³

Κάνοντας λόγο για εξόρυξη άποψης συνεπάγεται μία διαδικασία εξόρυξης κειμένου (text mining). Το text mining ή αλλιώς text data mining είναι ο ερευνητικός τομέας που αναφέρεται στη διαδικασία άντλησης πληροφοριών υψηλής ποιότητας από κειμενικές πηγές. Στη διαδικασία του text mining χρησιμοποιούνται τεχνικές ανάλυσης κειμένου, όπως είναι οι τεχνικές ανάκτησης πληροφορίας (information retrieval), λεξικής ανάλυσης (lexical analysis), αναγνώρισης προτύπων (pattern recognition), και εξόρυξης δεδομένων (data mining), καθώς επίσης και τεχνικές επεξεργασίας φυσικής γλώσσας (natural language processing).²¹⁴

Πολλά είναι τα ονόματα και οι διαδικασίες που σχετίζονται με το opinion mining και το sentiment analysis, καθώς πρόκειται για έναν ευρύ τομέα. Ενδεικτικά, συχνά συναντάει κανείς τους όρους και τις αντίστοιχες διαδικασίες που αυτοί συνεπάγονται, του opinion

²¹³ Παναγόπουλος, Π., 2015, Opinion Mining and Sentiment Analysis, Πανεπιστήμιο Πατρών, Τμήμα Μηχανικών Η/Υ και Πληροφορικής. Διπλωματική εργασία, http://83.212.99.169/wp/wp-content/uploads/2015/12/opinion_mining_and_sentiment_analysis.pdf, [προσπελάστηκε στις 24 Ιανουαρίου 2017], σελ. 1-2

²¹⁴ Ο.π.

extraction, sentiment mining, subjectivity analysis, affect analysis, emotion analysis και του review mining. Όλες αυτές οι διαδικασίες όμως βρίσκονται κάτω από την ομπρέλα του opinion mining και του sentiment analysis. Από τη μία, στο χώρο των εταιρειών, είναι συχνότερος ο όρος «sentiment analysis», ενώ από την άλλη στην ακαδημαϊκή κοινότητα και οι δύο όροι, «opinion mining» και «sentiment analysis», χρησιμοποιούνται το ίδιο συχνά.²¹⁵ Οι δύο αυτές έννοιες, ενώ μπορεί φαινομενικά να ταυτίζονται, παρουσιάζουν στην πραγματικότητα διαφορές.²¹⁶ Από τη μία πλευρά, το opinion mining αναλύει τη γνώμη των ανθρώπων για μία οντότητα, ενώ το sentiment analysis αναγνωρίζει το συναίσθημα που εκφράζεται σε μία γνώμη και το αναλύει. Στόχος της ανάλυσης συναισθήματος είναι η αναγνώριση του συναισθήματος σε ένα κείμενο που εκφράζει γνώμη και η κατηγοριοποίησή του ανάλογα με την πολικότητα του, για παράδειγμα σε θετικό ή αρνητικό.²¹⁷ Επομένως, η ειδοποιός διαφορά βρίσκεται στο ότι η ανάλυση συναισθήματος ασχολείται με τη διάκριση και την κατηγοριοποίηση του συναισθήματος για μία οντότητα, ενώ η εξόρυξη γνώμης εστιάζει στην εις βάθος περιγραφή της ίδιας της οντότητας και των χαρακτηριστικών της.²¹⁸

Ο επιστημονικός τομέας της εξόρυξης γνώμης και της ανάλυσης συναισθήματος σχετίζεται και με τον τομέα του Natural Language Processing (NLP), καθώς ο δεύτερος ασχολείται με την επεξεργασία της ανθρώπινης γλώσσας και την εξαγωγή νοήματος από αυτή. Παρόλα αυτά, αν και η γλωσσολογία και η επεξεργασία φυσικής γλώσσας έχουν μία μακρά ερευνητική ιστορία, λίγη είναι η έρευνα που έχει πραγματοποιηθεί πριν το 2000 για τις τεχνικές εξαγωγής και ανάλυσης απόψεων από κειμενικές πηγές. Από το 2000 όμως και μετά, ο τομέας του opinion mining έχει γίνει μία πολύ ενεργή ερευνητική περιοχή, γεγονός που οφείλεται σε πολλούς λόγους. Αρχικά, πρόκειται για έναν τομέα που έχει ένα ευρύ φάσμα εφαρμογών, εκ των οποίων οι εμπορικές ήταν οι

²¹⁵ Ό.π. παραπομπή 213, σελ. 3.

²¹⁶ Τζίμας, Δ., Επεξεργασία Δεδομένων του Twitter Σχετικά με Παρενέργειες Φαρμάκων στο Κατανεμημένο Σύστημα Storm, Διπλωματική Εργασία, Πολυτεχνείο Κρήτης Σχολή Ηλεκτρονικών Μηχανικών & Μηχανικών Υπολογιστών, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://purl.tuc.gr/dl/dias/24B7029F-4418-435F-BEBA-7B37DBE46E43>, [προσπελάστηκε στις 24 Ιανουαρίου 2017], σελ. 33.

²¹⁷ Medhat, W., Hassan, A., Korashy, H., Sentiment analysis algorithms and applications: A survey, Ain Shams Engineering Journal, 2014, 5 (4), σελ. 1093.

²¹⁸ Tsytarau, M., & Palpanas, T., Survey on mining subjective data on the web. Data Mining and Knowledge Discovery, 2012, 24(3), σελ. 478-514.

σημαντικότερες κατά τα προηγούμενα χρόνια και αυτές που συνετέλεσαν στην ανάπτυξη του εν λόγω τομέα. Δεύτερον, πρόκειται για έναν ερευνητικό πεδίο που παρέχει πολλά ενδιαφέροντα ερευνητικά προβλήματα, τα οποία δεν έχουν μελετηθεί στο παρελθόν. Επιπλέον, για πρώτη φορά στην ανθρώπινη ιστορία, με τη ραγδαία ανάπτυξη των τεχνολογιών του διαδικτύου, υφίσταται ένας τεράστιος όγκος από κείμενα και κατ' επέκταση από απόψεις για γεγονότα και άτομα στο διαδίκτυο και κατά κύριο λόγο στα κοινωνικά δίκτυα, χωρίς τα οποία, μεγάλο μέρος της έρευνας δεν θα μπορούσε να πραγματοποιηθεί. Καθόλου τυχαία δεν είναι η ταυτόχρονη ταχεία ανάπτυξη των κοινωνικών δικτύων και της εξόρυξης και ανάλυσης απόψεων.²¹⁹ Προκειμένου να προσπελάσει κανείς αυτό το πλήθος των απόψεων, απαιτείται η κατανάλωση σημαντικού χρόνου και ενέργειας, έτσι ώστε να γίνει κατανοητή η συνολική εικόνα του εύρους των απόψεων. Για τον λόγο αυτό, η ανάπτυξη υπολογιστικών τεχνικών που βοηθούν τον χρήστη στην εξόρυξη και στην αποδοτική χρήση όλων των απόψεων αποτελεί μία ενδιαφέρουσα ερευνητική πρόκληση.²²⁰

Αυτό που κάνει τις απόψεις ένα ιδιαιτέρως σημαντικό κομμάτι της ανθρώπινης δραστηριότητας, είναι το γεγονός ότι επηρεάζουν σε μεγάλο βαθμό τη συμπεριφορά του κάθε ανθρώπου. Κάθε φορά που κάποιος χρειάζεται να πάρει μία απόφαση, θέλει πολλές φορές να γνωρίζει τις απόψεις του κοινωνικού συνόλου. Για παράδειγμα, οι καταναλωτές, θέλουν να γνωρίζουν τις απόψεις των άλλων χρηστών για ένα προϊόν ή μία υπηρεσία, πριν προχωρήσουν σε κάποια αγορά. Αντίστοιχα οι εταιρείες και οι οργανισμοί θέλουν να έχουν εικόνα των απόψεων των καταναλωτών ή της κοινής γνώμης για τα προϊόντα ή τις υπηρεσίες τους. Ακόμα η γνώση των απόψεων της κοινής γνώμης παίζει σημαντικό ρόλο στη διαδικασία μιας πολιτικής ψηφοφορίας, επηρεάζοντας ακόμα και το τελικό αποτέλεσμα.²²¹

Για τους παραπάνω λόγους ο τομέας αυτός βρίσκεται στο επίκεντρο της ερευνητικής δραστηριότητας που έχει να κάνει με τα κοινωνικά δίκτυα. Άλλωστε, η έρευνα πάνω στις απόψεις χρηστών δεν έχει μόνο σημαντικό αντίκτυπο για τον τομέα της επεξεργασίας της φυσικής γλώσσας, αλλά μπορεί επίσης να έχει μια βαθιά επίδραση

²¹⁹ Ό.π., παραπομπή 218, σελ. 3.

²²⁰ Ό.π., σελ. 1-2.

²²¹ Παναγόπουλος, Π., 2015, Opinion Mining and Sentiment Analysis, ό.π. παραπομπή 213, σελ. 1-2.

στις επιστήμες διαχείρισης, τις πολιτικές επιστήμες, τα οικονομικά και τις κοινωνικές επιστήμες, καθώς όλες αυτές επηρεάζονται από τις απόψεις των ανθρώπων.²²²

3.2.1 ΕΦΑΡΜΟΓΕΣ

Η απόκτηση της εικόνας των απόψεων της κοινής γνώμης και των καταναλωτών, αποτελούσε πάντα μία τεράστια επιχείρηση για το marketing, τις δημόσιες σχέσεις και τις εκστρατείες πολιτικών κινήσεων, με τη διαφορά ότι πριν από περίπου μία δεκαετία, όταν κάποιος χρειαζόταν απόψεις σχετικά με κάποιο θέμα, τις αναζητούσε στον κοινωνικό κύκλο των φίλων ή της οικογένειάς του. Όταν μία επιχείρηση χρειαζόταν τις απόψεις των καταναλωτών, διεξήγαγε έρευνες κυρίως με την μορφή των δημοσκοπήσεων.²²³ Επίσης, όσο σημαντική είναι η φωνή του καταναλωτή σχετικά με ένα προϊόν για μία εταιρεία, τόσο σημαντική είναι και η φωνή των πολιτών σχετικά με κοινωνικά θέματα για την κυβέρνηση και τους οργανισμούς. Λόγω των παραπάνω, οι εφαρμογές της εξόρυξης και ανάλυσης απόψεων έχουν εξαπλωθεί σε κάθε πιθανό θεματικό τομέα, από τα προϊόντα, τις υπηρεσίες και την οικονομία στις κοινωνικές εκδηλώσεις, τις πολιτικές εκλογές και τα κοινωνικά θέματα.²²⁴

Συνοπτικά, η ανάλυση συναισθήματος εφαρμόζεται κατά κύριο λόγο στους παρακάτω τομείς:²²⁵

- Βιομηχανία:
 - Φωνή καταναλωτή
 - Φήμη μάρκας
 - Διαδικτυακή διαφήμιση
 - Διαδικτυακό εμπόριο

²²² Ο. π. παραπομπή 213, σελ. 3.

²²³ Ο.π., σελ. 4.

²²⁴ Ο.π., σελ. 5.

²²⁵ Alessia, D., Ferri, F., Grifoni, P., & Guzzo, T., Approaches, Tools and Applications for Sentiment Analysis Implementation, *International Journal of Computer Applications*, 2015, 125(3), σελ. 31.

- Πολιτική:
 - Συμβουλευτική δημοσκοπήσεων
 - Ταξινόμηση πολιτικών τάσεων

- Κοινωνία:
 - Παρακολούθηση διοργανώσεων
 - Δημόσια διαβούλευση
 - Συμβουλευτική επενδύσεων
 - Ευφυή συστήματα μεταφορών
 - Κοινωνικά θέματα

3.2.2 ΟΡΙΣΜΟΣ ΤΗΣ ΑΠΟΨΗΣ

Μία άποψη αποτελείται από ένα στόχο και ένα συναίσθημα. Παίρνοντας ως παράδειγμα την πρόταση «Το παγωτό είναι καλό», διαπιστώνει κανείς ότι αποτελεί μία πρόταση που εκφράζει μία άποψη, στόχος της οποίας είναι το «παγωτό» και το συναίσθημα που εκφράζει είναι το «καλό». Σε αυτή την πρόταση, επομένως, εκφράζεται μία άποψη θετικής πολικότητας. Έστω (g,s) ο ορισμός της άποψης, όπου g είναι ο στόχος της άποψης και s είναι το συναίσθημα που εκφράζεται για το στόχο g. Ο στόχος g μπορεί να είναι μία οντότητα, ή ένα μέρος ή μία ιδιότητα της οντότητας. Το s είναι ένα θετικό, αρνητικό ή ουδέτερο συναίσθημα, ή ακόμα και μία αριθμητική βαθμολογία, η οποία εκφράζει την ένταση του συναισθήματος (αξιολόγηση 1-10 ή 0-5 αστέρια).²²⁶

Πιο αναλυτικά, όπως αναφέρθηκε παραπάνω, ο στόχος της άποψης μπορεί να αποτελείται από μία οντότητα στο σύνολό της, αλλά και από ένα μέρος της οντότητας αυτής ή ακόμα και από μία ιδιότητά της. Για παράδειγμα στην πρόταση «η γεύση του παγωτού είναι καλή», ο στόχος της άποψης είναι μία ιδιότητα της οντότητας «παγωτό» και συγκεκριμένα «η γεύση του παγωτού». Μία οντότητα ορίζεται ως $e(T,W)$, όπου T

²²⁶ Liu, B. , *Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data*, Springer, 2006, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: http://sirius.cs.put.poznan.pl/~inf89721/Seminarium/Web_Data_Mining_2nd_Edition_Exploring_Hyperlinks_Contents_and_Usage_Data.pdf, [προσπελάστηκε στις 12 Φεβρουαρίου 2017], σελ. 461.

είναι η ιεραρχία με τα επιμέρους τμήματα (sub-parts) που αποτελούν την οντότητα, και W είναι το σύνολο των ιδιοτήτων της οντότητας. Η αναγνώριση των μερών και των ιδιοτήτων μιας οντότητας είναι ένα δύσκολο NLP πρόβλημα και συνήθως χρησιμοποιείται μια ιεραρχία 2 επιπέδων: Την ίδια την οντότητα σε πρώτο επίπεδο και τα χαρακτηριστικά (aspects) της οντότητας σε δεύτερο επίπεδο, όπου αποτελούν το σύνολο των επιμέρους τμημάτων και ιδιοτήτων της εκάστοτε οντότητας.

Ο αναλυτικός ορισμός της άποψης είναι ο εξής²²⁷:

Μια άποψη είναι μία πεντάδα ($ei, aij, sijkl, hk, tl$), όπου ei είναι μία οντότητα, aij είναι ένα χαρακτηριστικό (aspect/feature) της οντότητας ei , $sijkl$ είναι το συναίσθημα το οποίο εκφράζεται για το aspect aij της οντότητας ei , hk είναι αυτός που εκφράζει την άποψη (opinion holder), και tl είναι η χρονική στιγμή την οποία εκφράζεται η άποψη από τον hk .

Το συναίσθημα $sijkl$ είναι θετικό, αρνητικό ή ουδέτερο, ή εκφράζεται με διαφορετικά επίπεδα έντασης, για παράδειγμα 1–5 αστέρια, όπως χρησιμοποιείται συνήθως σε ιστοσελίδες με reviews στο διαδίκτυο. Σε αυτό τον ορισμό τα ei και aij μαζί αντιπροσωπεύουν τον στόχο της άποψης (opinion target). Όταν μία άποψη έχει ως στόχο την οντότητα ως ολότητα και όχι κάποιο χαρακτηριστικό της οντότητας αυτής, τότε το aij μπορεί να πάρει την τιμή «general».

Ο ορισμός αυτός καλύπτει τις περισσότερες, όχι όμως όλες τις πιθανές οπτικές της σημασιολογικής έννοιας μιας άποψης, η οποία μπορεί να είναι περισσότερο σύνθετη. Για παράδειγμα, δεν καλύπτει την εξής περίπτωση «Το σκόπευτρο και ο φακός είναι πολύ κοντά», η οποία εκφράζει άποψη για την απόσταση μεταξύ των δύο μερών. Επίσης αυτό το μοντέλο της άποψης είναι ένας από τους τύπους απόψεων που χρησιμοποιούνται αρκετά συχνά και ονομάζεται γενική άποψη. Παρά τους περιορισμούς που εισάγει στην ανάλυση, μπορούμε ακόμα να το χρησιμοποιούμε μιας και αποδεικνύεται επαρκές για τις περισσότερες εφαρμογές. Άλλωστε ένα πολύ

²²⁷ Ο. π. παραπομπή 226, σελ. 463.

σύνθετο μοντέλο μπορεί να κάνει το πρόβλημα εξαιρετικά δύσκολο, ώστε να βρεθεί κάποια λύση.²²⁸

3.2.3 Η ΔΙΑΔΙΚΑΣΙΑ ΕΞΟΡΥΞΗΣ ΚΑΙ ΑΝΑΛΥΣΗΣ ΑΠΟΨΕΩΝ

Έχοντας δώσει τον ορισμό της άποψης και τους συνήθεις τρόπους με τους οποίους οι απόψεις τείνουν να εμφανίζονται, έπεται ο ορισμός της διαδικασίας της εξόρυξης και ανάλυσης απόψεων, ο οποίος έχει ως εξής:

Δοθέντος ενός κειμένου d (ή μιας συλλογής κειμένων D), μπορεί κανείς να ανακαλύψει όλες τις πλειάδες απόψεων ($ei, aij, sijkl, hk, tl$) που υπάρχουν μέσα στο d . Η διαδικασία περιέχει έξι διεργασίες (tasks)²²⁹:

- Task 1: Εξαγωγή όλων των οντοτήτων (entities) από το D και κατηγοριοποίηση των συνώνυμων σε ομάδες (entity clusters). Σε κάθε cluster δίνεται το αναγνωριστικό ei .
- Task 2: Εξαγωγή όλων των χαρακτηριστικών (aspects) από το D και κατηγοριοποίηση των συνώνυμων σε ομάδες (aspects clusters). Σε κάθε cluster δίνεται το αναγνωριστικό aij του ei .
- Task 3: Εξαγωγή των opinion holders και κατηγοριοποίηση σε ομάδες με αναγνωριστικό hk .
- Task 4: Εξαγωγή του χρόνου και κανονικοποίηση σε ένα πρότυπο.
- Task 5: Προσδιορισμός της πολικότητας της άποψης σε θετική, αρνητική, ουδέτερη, ή με βαθμολόγηση σε ποσοστό επί τις % ή με 1-5 αστέρια.
- Task 6: Παραγωγή όλων των πλειάδων απόψεων ($ei, aij, sijkl, hk, tl$).

²²⁸ Παναγόπουλος, Π., 2015, *Opinion Mining and Sentiment Analysis*, ό.π. παραπομπή 213, σελ. 5-7.

²²⁹ Liu, B. , *Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data*, ό.π. παραπομπή 226, σελ. 465.

Ο παραπάνω ορισμός περιγράφει την πλήρη διαδικασία εξόρυξης και ανάλυσης απόψεων. Πρέπει να τονιστεί πως δεν απαιτείται από όλες τις εφαρμογές η πλήρης διαδικασία. Συνήθως ζητούνται πλειάδες του τύπου (ei, aij, sijkl), οπότε οι διεργασίες τρία και τέσσερα παραλείπονται.

Ακόμα και αναφορικά με την εξαγωγή οντοτήτων υπάρχουν περιπτώσεις, στις οποίες μπορεί να αποφευχθεί. Για παράδειγμα, σε ένα σύστημα αναζήτησης απόψεων για οντότητες, ο χρήστης εισάγει το όνομα ή το είδος της οντότητας για την οποία ενδιαφέρεται, οπότε το σύστημα μπορεί να επεξεργάζεται μόνο τα κείμενα ή τις προτάσεις οι οποίες αναφέρουν αυτή την οντότητα. Θα πρέπει όμως να τονιστεί πως οι χρήστες συνήθως χρησιμοποιούν διαφορετικά ονόματα για να αναφερθούν στην ίδια οντότητα, με αποτέλεσμα να είναι επιθυμητό από το σύστημα να έχει μία λίστα με τις διαφορετικές ονομασίες μιας οντότητας.²³⁰ Η κατηγοριοποίηση των συνωνύμων σε ομάδες που πραγματοποιείται στα παραπάνω tasks είναι απαραίτητη διότι, όπως συμβαίνει σε όλα τα κείμενα φυσικής γλώσσας, χρησιμοποιούνται διαφορετικές, συνώνυμες, λέξεις ή φράσεις για να γίνει αναφορά στις ίδιες οντότητες ή τα ίδια χαρακτηριστικά.²³¹

3.2.4 ΔΙΑΦΟΡΕΤΙΚΑ ΕΠΙΠΕΔΑ ΑΝΑΛΥΣΗΣ

Η εξόρυξη και ανάλυση απόψεων έχει ερευνηθεί κυρίως σε τρία επίπεδα, στο επίπεδο κειμένου (document level), στο επίπεδο πρότασης (sentence level) και στο επίπεδο οντότητας ή χαρακτηριστικού (entity or aspect level).²³²

Στο επίπεδο κειμένου, ένα ολόκληρο κείμενο υπολογίζεται ως η βασική πληροφοριακή μονάδα και επιχειρείται να κατηγοριοποιηθεί ως κείμενο με θετική ή αρνητική γνώμη ή συναίσθημα. Σε αυτό το επίπεδο θεωρείται πως το έγγραφο αναφέρεται σε μία και μόνο οντότητα. Έτσι η διαδικασία είναι ευκολότερη, διότι δεν χρειάζεται να

²³⁰ Παναγόπουλος, Π., 2015, *Opinion Mining and Sentiment Analysis*, ό.π. παραπομπή 213, σελ. 44.

²³¹ Ό.π., σελ. 9-10.

²³² Liu, B., *Sentiment Analysis and Opinion Mining*, Morgan & Claypool Publishers, 2012, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://www.cs.uic.edu/~liub/FBS/SentimentAnalysis-and-OpinionMining.pdf>, [προσπελάστηκε στις 12 Φεβρουαρίου 2017], σελ. 10-12.

πραγματοποιηθεί εξαγωγή οντοτήτων και είναι ασφαλές να γίνεται η υπόθεση πως όποια aspects υπάρχουν στο κείμενο, αναφέρονται στη μοναδική οντότητα του κειμένου. Σε αυτήν την περίπτωση ο ορισμός της διαδικασίας της εξόρυξης και ανάλυσης απόψεων διαμορφώνεται ως εξής:

«Δοθέντος ενός κειμένου απόψεων d , να προσδιοριστεί η συνολική sentiment πολικότητα s που εκφράζει ο συγγραφέας (opinion holder) για την οντότητα, ως σύνολο (general). Σε αυτό το επίπεδο ανάλυσης η ζητούμενη πεντάδα είναι η $(-, general, s, -, -)$, δηλαδή δεν ενδιαφερόμαστε για την εξαγωγή της οντότητας e , του opinion holder h , και του χρόνου t που διατυπώθηκε η άποψη.»

Στο επίπεδο ανάλυσης κειμένου το πρόβλημα της εξαγωγής και ανάλυσης απόψεων παρουσιάζει δύο μορφές, ανάλογα με τη είδος του πεδίου τιμών της επιθυμητής εξόδου. Αν το πεδίο τιμών αποτελείται από διακριτές, κατηγορηματικές, τιμές, όπως θετικό ή αρνητικό, τότε είναι ένα πρόβλημα κατηγοριοποίησης. Αν το πεδίο τιμών αποτελείται από αριθμητικές τιμές ή βαθμολογίες κατάταξης σε ένα προκαθορισμένο εύρος, όπως για παράδειγμα 1-5 αστέρια, τότε απαιτείται μία στατιστική διαδικασία για την εκτίμηση των σχέσεων μεταξύ μεταβλητών και περιλαμβάνει τεχνικές για την μοντελοποίηση και την ανάλυση των μεταβλητών και των σχέσεών τους, διαδικασία που δεν αποτελεί αντικείμενο της παρούσας έρευνας και γι'αυτό δεν θα αναλυθεί περαιτέρω.

Στο επίπεδο πρότασης, αναλύεται η κάθε πρόταση με σκοπό να αναγνωριστεί τι συναίσθημα εκφράζει. Στόχος είναι η κατηγοριοποίηση των προτάσεων σε προτάσεις που εκφράζουν θετικό, αρνητικό ή ουδέτερο συναίσθημα. Το ουδέτερο συναίσθημα συχνά συνεπάγεται ότι δεν εκφράζεται κάποια άποψη. Συχνά προηγείται κατηγοριοποίηση υποκειμενικότητας, δηλαδή κατηγοριοποίηση της πρότασης σε ουδέτερη ή πρόταση με πολικότητα, ανάλογα με το αν εκφράζει ή όχι κάποιο συναίσθημα, προκειμένου να μελετηθούν μόνο οι προτάσεις που έχουν θετικό ή αρνητικό συναίσθημα. Παρόλα αυτά θα πρέπει να σημειωθεί ότι δεν πρέπει να ταυτίζεται η υποκειμενικότητα με το συναίσθημα, καθώς μία αντικειμενική πρόταση δεν συνεπάγεται έλλειψη συναισθήματος, καθώς πολλές είναι αυτές που μπορούν να εκφέρουν μία άποψη, όπως για παράδειγμα η πρόταση «Αγοράσαμε το X αυτοκίνητο

τον προηγούμενο μήνα και οι υαλοκαθαριστήρες του παρμπρίζ έπεσαν».²³³ Ο ορισμός της διαδικασίας της εξόρυξης και ανάλυσης απόψεων κατά την ανάλυση σε επίπεδο πρότασης μετατρέπεται ως εξής:

«Δοθέντος μίας πρότασης x, το ζητούμενο είναι να προσδιοριστεί αν η πρόταση εκφράζει θετική, αρνητική, ή ουδέτερη άποψη.»

Αν και με βάση την παρατήρηση ότι οι προτάσεις είναι κείμενα μικρού μήκους, δεν φαίνεται να υπάρχει σημαντική διαφορά μεταξύ της ανάλυσης σε επίπεδο κειμένου και πρότασης, η υπόθεση που κάνουν συχνά οι ερευνητές, και είναι αυτή που διαφοροποιεί τα δύο επίπεδα, είναι πως μία πρόταση συχνά φέρει μία άποψη. Αυτή η υπόθεση απλοποιεί την ανάλυση, αλλά όπως είναι φανερό, δεν ισχύει σε πολλές περιπτώσεις.

Σε έρευνα με αντικείμενο την κατηγοριοποίηση tweets διαπιστώθηκε ότι κάθε tweet, εξαιτίας του περιορισμού των 140 χαρακτήρων που έχει, μπορεί να θεωρηθεί ως μία πρόταση και αυτή η υπόθεση θα υιοθετηθεί και στην παρούσα μελέτη.²³⁴

Στο επίπεδο ανάλυσης χαρακτηριστικού, στόχος είναι ο προσδιορισμός των χαρακτηριστικών (aspects) της οντότητας, δηλαδή των επιμέρους τμημάτων και των ιδιοτήτων της, καθώς και της πολικότητας της άποψης που εκφράζεται για αυτά τα χαρακτηριστικά, προκειμένου να κατηγοριοποιηθεί το συναίσθημα σε σχέση με τις συγκεκριμένες πτυχές των διαφορετικών οντοτήτων που υπάρχουν μέσα στην πρόταση. Μία πρόταση συνήθως περιέχει περισσότερα από ένα χαρακτηριστικά. Παίρνοντας ως παράδειγμα την πρόταση *«Αν και ο καφές δεν είναι ο καλύτερος, μου αρέσει αυτή η καφετέρια»*, παρατηρείται ότι εκφράζεται διαφορετική άποψη για διαφορετικά χαρακτηριστικά της ίδιας οντότητας. Το πρώτο βήμα σε αυτό το επίπεδο είναι να εντοπιστούν οι διάφορες οντότητες μέσα στην πρόταση, καθώς και οι πτυχές τους.²³⁵ Σε αυτό το επίπεδο γίνεται ανάλυση κατευθείαν στην ίδια την άποψη, ενώ για

²³³ Ό.π. παραπομπή 232, σελ. 11.

²³⁴ Dmitry, D., Tsur, O., & Rappoport, A., Enhanced sentiment learning using twitter hashtags and smileys., in Proceedings of Coling, 2010, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.aclweb.org/anthology/C10-2028>, [προσπελάστηκε στις 28 Φεβρουαρίου 2017], [προσπελάστηκε στις 28 Φεβρουαρίου 2017], σελ. 244.

²³⁵ Medhat, W., Hassan, A., Korashy, H., Sentiment analysis algorithms and applications: A survey, ό.π. παραπομπή 217, σελ. 1094.

την εξαγωγή αυτών των λεπτομερειών είναι απαραίτητο να πραγματοποιηθεί ανάλυση απόψεων σε επίπεδο aspect.

Αυτό σημαίνει πως κρίνεται αναγκαίο να χρησιμοποιηθεί ο πλήρης ορισμός της εξόρυξης και ανάλυσης απόψεων με τις έξι διεργασίες, όπως αυτός αναφέρεται παραπάνω. Σε αυτό το επίπεδο η διαδικασία του opinion mining γίνεται πιο δύσκολη, διότι η αναγνώριση των οντοτήτων και των συναισθημάτων που εκφράζονται για αυτές, απαιτεί βαθύτερη NLP ανάλυση.²³⁶

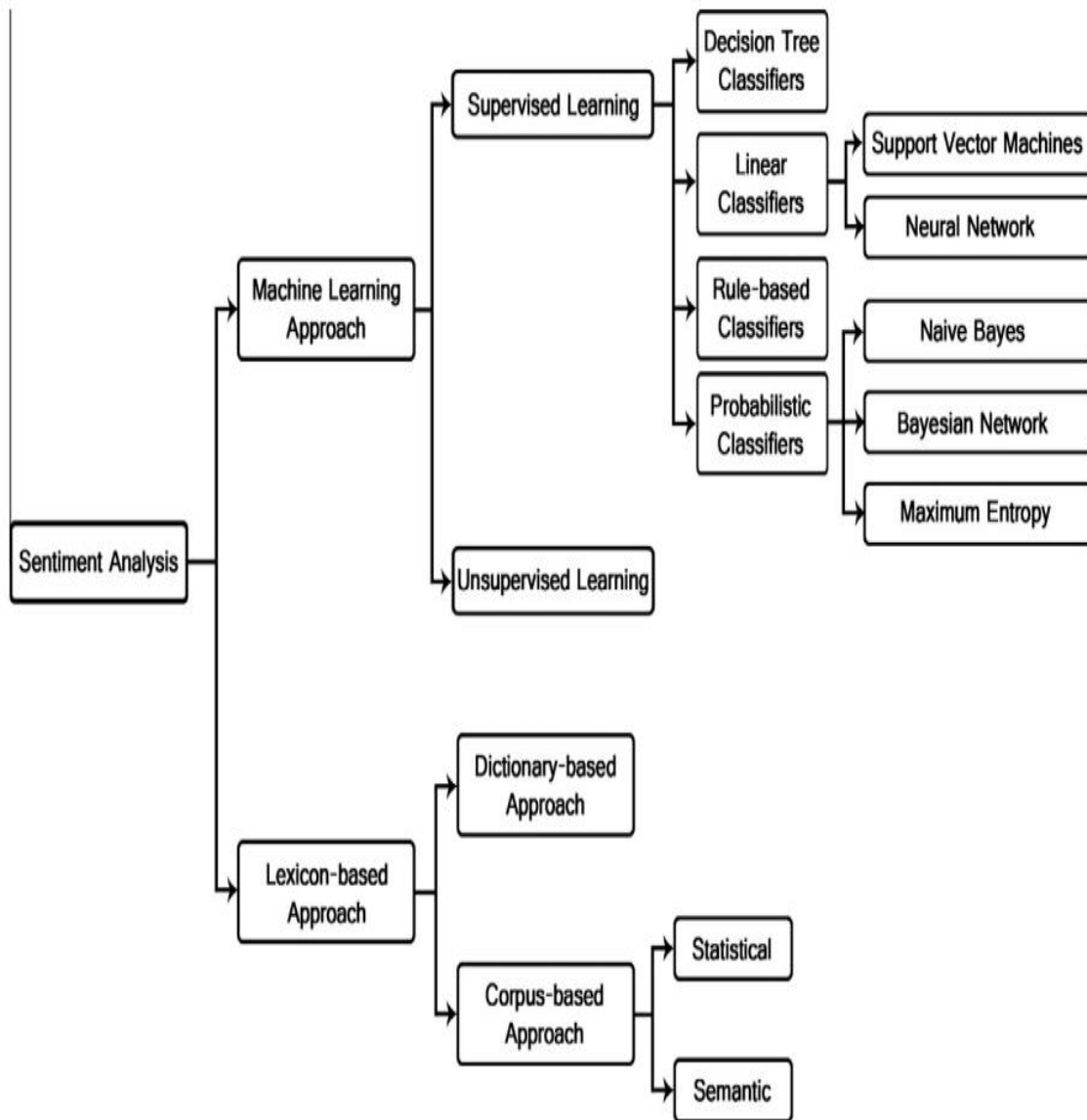
3.2.5 ΤΕΧΝΙΚΕΣ ΚΑΤΗΓΟΡΙΟΠΟΙΗΣΗΣ ΣΥΝΑΙΣΘΗΜΑΤΟΣ

Στο πεδίο του sentiment analysis χρησιμοποιούνται διάφορες τεχνικές ανίχνευσης συναισθήματος, όπου κάθε μία χρησιμοποιεί και διαφορετικό είδος αλγορίθμων ή και συνδυασμό αυτών.²³⁷ Οι δύο μεγάλες κατηγορίες στις οποίες μπορούμε να χωρίσουμε τις εν λόγω τεχνικές είναι οι εξής (εικόνα 24):

- Τεχνικές βασισμένες σε λεξικό (lexicon based techniques)
- Τεχνικές βασισμένες σε μηχανική μάθηση (machine learning based techniques)

²³⁶ Παναγόπουλος Π., 2015, Opinion Mining and Sentiment Analysis, ό.π. παραπομπή 213, σελ. 9-10.

²³⁷ Vohra, S., Teraiya, J. B., A Comparative Study of Sentiment Analysis Techniques., *Journal of Information Knowledge and Research in Computer Engineering*, 2 (2), 2012, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.ejournal.aessangli.in/ASEEJournals/CE63.pdf> [προσπελάστηκε στις 18 Μαρτίου 2017], σελ. 314.



Εικόνα 24. Τεχνικές κατηγοριοποίησης συναισθήματος²³⁸

3.2.5.1 ΤΕΧΝΙΚΕΣ ΒΑΣΙΣΜΕΝΕΣ ΣΕ ΛΕΞΙΚΟ (LEXICON BASED TECHNIQUES)

Στις τεχνικές που χρησιμοποιούν λεξικό, η κατηγοριοποίηση γίνεται συγκρίνοντας τα χαρακτηριστικά από μία κειμενική μονάδα με λέξεις που υπάρχουν στο λεξικό και στις οποίες έχει αντιστοιχηθεί κάποιο συναίσθημα από πριν. Να σημειωθεί εδώ ότι ως λεξικό (lexicon) νοούνται λίστες με λέξεις ή φράσεις που χρησιμοποιούνται από τους

²³⁸ Medhat, W., Hassan, A., Korashy, H., Sentiment analysis algorithms and applications: A survey, ό.π. παραπομπή 217, σελ. 1095.

ανθρώπους για να εκφράσουν συναίσθημα ή γνώμη (opinion words) και όχι το λεξικό με την τυπική του έννοια, αυτή του dictionary, το οποίο απαρτίζεται από μία λίστα λέξεων μαζί με τις σημασίες τους.

Η διαδικασία που ακολουθείται είναι αρχικά ο ορισμός των λεξικών για θετικές και αρνητικές λέξεις, ακολουθεί η ανάλυση του κειμένου και στο τέλος εάν οι θετικές λέξεις είναι περισσότερες από τις αρνητικές, το συναίσθημα χαρακτηρίζεται ως θετικό, αλλιώς ως αρνητικό.

Σε πολλές τεχνικές ανάλυσης συναισθήματος χρησιμοποιούνται λίστες με λέξεις ή φράσεις, οι οποίες συνήθως εκφράζουν θετική ή αρνητική πολικότητα στο λόγο. Αυτές οι λίστες ονομάζονται sentiment λεξικά ή opinion λεξικά. Οι λέξεις θετικής πολικότητας εκφράζουν επιθυμητές καταστάσεις ή ιδιότητες, ενώ οι αρνητικές μη επιθυμητές καταστάσεις ή ιδιότητες. Επίσης είναι δυνατόν να χρησιμοποιηθούν και sentiment εκφράσεις. Ένα sentiment λεξικό μπορεί να περιέχει μόνο λέξεις, ή μόνο εκφράσεις, ή λέξεις και εκφράσεις μαζί, από μία ή και από τις δύο πολικότητες.

Οι κυριότερες μέθοδοι για την σύνταξη των λεξικών είναι οι εξής:

- **Dictionary based techniques:** Οι dictionary-based προσεγγίσεις χρησιμοποιούν ένα λεξικό, όπως για παράδειγμα το σημασιολογικό δίκτυο WordNet, για τη δημιουργία ενός sentiment λεξικού. Τα λεξικά αυτά, τύπου WordNet, διαθέτουν λίστες με τα συνώνυμα και αντώνυμα των λέξεων. Έτσι, μία απλή τεχνική για τη δημιουργία ενός sentiment λεξικού, είναι η επιλογή ενός μικρού συνόλου λέξεων από κάποιο λεξικό και η χρήση των λιστών με τα συνώνυμα και τα αντώνυμα αυτών των λέξεων για τον εμπλουτισμό του sentiment λεξικού. Το πλεονέκτημα αυτής της προσέγγισης είναι ο εύκολος και γρήγορος εντοπισμός μεγάλου αριθμού sentiment λέξεων. Πρέπει να τονιστεί πως οι sentiment λέξεις, οι οποίες συγκεντρώνονται με αυτό τον τρόπο είναι domain-independent και context-independent, αδυνατεί δηλαδή η εν λόγω μέθοδος να βρει opinion words για ειδικότερα θέματα, γεγονός που αποτελεί και το βασικό μειονέκτημα αυτής της μεθόδου.

- **Corpus based techniques:** Αυτές οι τεχνικές βασίζονται στα συντακτικά μοτίβα μεγάλων corpora (σωμάτων κειμένων). Μπορούν να παράγουν opinion words με σχετικά μεγάλη ακρίβεια, χρειάζονται όμως ένα πολύ μεγάλο όγκο δεδομένων για εκπαίδευση. Το πλεονέκτημα αυτών των τεχνικών είναι ότι μπορούν να παράξουν opinion words για ειδικότερα θέματα, ανάλογα και με τα δεδομένα που παρέχονται για εκπαίδευση.²³⁹

Επιπλέον, υπάρχει και η χειρωνακτική προσέγγιση (manual), η οποία απαιτεί ανθρώπινη, χειρωνακτική, εργασία και όπως είναι φανερό είναι χρονοβόρα, αλλά με τον τρόπο αυτό εξασφαλίζεται η απόδοση της ανθρώπινης ακρίβειας. Συνήθως όμως αυτή η τεχνική χρησιμοποιείται συνδυαστικά με τις αυτόματες προσεγγίσεις, ως το βήμα του τελικού ελέγχου των αποτελεσμάτων, διότι συχνά στις αυτόματες προσεγγίσεις παρατηρούνται λάθη.

3.2.5.2 ΤΕΧΝΙΚΕΣ ΒΑΣΙΣΜΕΝΕΣ ΣΕ ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ (MACHINE LEARNING TECHNIQUES)

Η εξόρυξη δεδομένων χρησιμοποιεί με κατάλληλο τρόπο αλγορίθμους της μηχανικής μάθησης. Ένας μεγάλος αριθμός από τεχνικές machine learning έχει χρησιμοποιηθεί για να την κατηγοριοποίηση γνώμων. Οι προσεγγίσεις μηχανικής μάθησης, όπου είναι εφαρμόσιμες στο sentiment analysis ανήκουν στην κατηγορία της εποπτευόμενης μάθησης (supervised learning).

Συγκεκριμένα, οι επαγγελματίες της μηχανικής μάθησης χρησιμοποιούν τα δεδομένα σαν ένα σύνολο εκπαίδευσης για να εκπαιδεύσουν έναν αλγόριθμο από τους πολλούς της μηχανικής μάθησης, όπως είναι τα δίκτυα Bayes (Bayes nets), οι μηχανές διανυσμάτων υποστήριξης (support-vector machines), τα δένδρα αποφάσεων (decision trees), τα κρυμμένα μοντέλα Markov (hidden Markov models) και πολλοί άλλοι.²⁴⁰

²³⁹ Παναγόπουλος Π., 2015, Opinion Mining and Sentiment Analysis, ό.π. παραπομπή 213, σελ. 55-56

²⁴⁰ Rajaraman, A., Ullman, J., *Εξόρυξη από μεγάλα σύνολα δεδομένων*, ό.π. παραπομπή 204, σελ. 2.

Για να εφαρμοστούν οι τεχνικές μηχανικής μάθησης χρειάζονται δύο διαφορετικές κατηγορίες κειμένων, τα κείμενα για εκπαίδευση και τα κείμενα όπου θα γίνουν οι εφαρμογές (test). Ένας αυτόματος ταξινομητής (classifier) χρησιμοποιεί τα κείμενα εκπαίδευσης για να μάθει τα διαφοροποιητικά χαρακτηριστικά των κειμένων και στη συνέχεια να δοκιμάσει την ακρίβεια και την αποδοτικότητά του στα επόμενα κείμενα. Πολύ σημαντικό ρόλο σε αυτό παίζει η ακρίβεια με την οποία θα εξαχθεί το συναίσθημα, ενώ ένας ακόμα σημαντικός παράγοντας είναι η χρήση μεγάλου όγκου δεδομένων, τόσο κατά την εκπαίδευση, όσο και κατά τη χρήση του ταξινομητή στην κατάταξη νέων κειμένων σε κατηγορίες.

3.2.5.2.1 Η ΜΠΕΥΖΙΑΝΗ ΣΤΑΤΙΣΤΙΚΗ ΚΑΙ Ο ΑΛΓΟΡΙΘΜΟΣ NAIVE BAYES

Η Μπεϋζιανή στατιστική (Bayesian statistics), αν και έχει τις ρίζες της στα μέσα του 18^{ου} αιώνα, μόλις τις δύο τελευταίες δεκαετίες έχει ευρέως καθιερωθεί στο χώρο της επιστημονικής έρευνας, καθώς η μεθοδολογία που ακολουθείται, η οποία σχετίζεται με την πρότερη γνώση βασικών παραμέτρων του εκάστοτε προβλήματος, θα μπορούσε να οδηγήσει τη στατιστική μελέτη σε υποκειμενικά συμπεράσματα. Εάν μάλιστα κάποιος υποθέσει ότι η πρότερη αυτή γνώση πηγάζει από την προσωπική πεποίθηση για κάποιο γεγονός, τότε τα συμπεράσματα θα μπορούσαν να είναι απόλυτα εσφαλμένα. Στην πραγματικότητα όμως, εάν η πρότερη διαθέσιμη πληροφορία αξιοποιηθεί δόκιμα, τα αποτελέσματα είναι πολύτιμα.²⁴¹

Το βασικό χαρακτηριστικό της Μπεϋζιανής μεθοδολογίας είναι ότι περιλαμβάνει τη χρήση μέτρων πιθανότητας ως μέσο ποσοτικοποίησης της αβεβαιότητας των εμπλεκόμενων παραμέτρων ενός προβλήματος.²⁴² Οι παράμετροι του προβλήματος που εξετάζεται θεωρούνται πλέον τυχαίες μεταβλητές και συνεπώς δεν έχουν αναγκαστικά σταθερή τιμή. Με τον τρόπο αυτό, εκτιμάται το διάστημα μέσα στο οποίο

²⁴¹ Ο' Hagan, A., *The Bayesian Approach to Statistics*, Chapter 6, in: *Handbook of Probability: Theory and Applications*, Corwin, 2008, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: https://www.corwin.com/sites/default/files/upm-binaries/18550_Chapter6.pdf, [προσπελάστηκε στις 12 Ιουνίου 2017], σελ. 94.

²⁴² Congdon, P., *Bayesian Statistical Modelling*, Second Edition, Wiley Series in Probability and Statistics, 2006, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://bacbuc.hd.free.fr/WebDAV/data/Bouquins/Congdon%20-%20Bayesian%20Statistical%20Modelling.pdf>, [προσπελάστηκε στις 22 Ιουνίου 2017], σελ. 2.

αναμένεται να βρίσκεται μια ποσότητα με βάση το κατά πόσο μεγάλη είναι η πιθανότητα που αντιστοιχεί σε αυτή, σε αντίθεση με τον αυστηρό χαρακτήρα της εκτίμησης ενός διαστήματος εμπιστοσύνης που προκύπτει στηριζόμενο στην επαναληπτικότητα του ίδιου πειράματος τύχης.

Στην ουσία, η Μπεϋζιανή στατιστική χρησιμοποιεί την έννοια της πιθανότητας με μία πιο γενικευμένη έννοια από ότι η κλασική στατιστική (frequentist statistics), η οποία και τη χρησιμοποιεί κάτω από την αυστηρή ερμηνεία της σχετικής συχνότητας εμφάνισης ενός γεγονότος, όταν επαναληφθεί πολλές φορές. Αντιθέτως, η Μπεϋζιανή στατιστική ερμηνεύει την έννοια της πιθανότητας ως την ευλογοφάνεια (likelihood) της δυνατότητας πραγματοποίησης ενός ενδεχομένου, χωρίς δηλαδή να έχει υποχρεωτικά από πριν προηγηθεί για αυτό πείραμα τύχης. Αυτό είναι και το χαρακτηριστικό που προσδίδει στη συνεπαγόμενη συμπερασματολογία μεγάλη ευελιξία, καθώς, έτσι καθίσταται εφικτή μέσω της ύστερης και μόνο κατανομής.

Το μοντέλο κατηγοριοποίησης Naive Bayes, που θα χρησιμοποιηθεί και στην παρούσα έρευνα, υπολογίζει την a-posteriori πιθανότητα μίας κλάσης, βασισμένο στην κατανομή των λέξεων της κλάσης αυτής στο κείμενο. Ο Naive Bayes χρησιμοποιεί για την εξαγωγή χαρακτηριστικών (feature extraction) το μοντέλο σάκου λέξεων (bag of words), μη λαμβάνοντας υπ' όψιν τη θέση της λέξης στο κείμενο. Στη συνέχεια, υπολογίζεται η πιθανότητα ένα δεδομένο σύνολο από χαρακτηριστικά (feature set) να ανήκει σε μία συγκεκριμένη ετικέτα (label), όπως για παράδειγμα, θετικό, αρνητικό ή ουδέτερο συναίσθημα. Θεωρώντας ότι όλα τα χαρακτηριστικά είναι ανεξάρτητα μεταξύ τους, προκύπτει ότι:

$$P(\text{label}|\text{features}) = \frac{P(\text{label}) * P(\text{features}|\text{label})}{P(\text{features})}$$

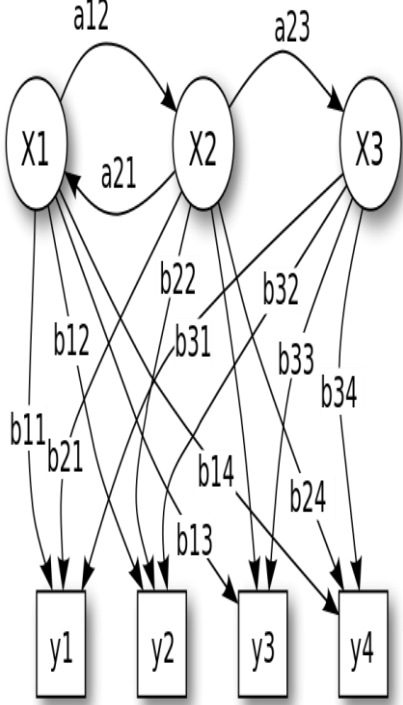
Εικόνα 25. Το μοντέλο του Naive Bayes²⁴³

²⁴³ Medhat, W., Hassan, A., Korashy, H., Sentiment analysis algorithms and applications: A survey, ό.π. παραπομπή 217, σελ. 1099.

Όπου $P(\text{label})$ είναι η a- priori πιθανότητα για ένα τυχαίο χαρακτηριστικό (feature) να ορίσει μία ετικέτα (label), $P(f|\text{label})$ είναι η a- priori πιθανότητα για κάθε σύνολο από χαρακτηριστικά να κατηγοριοποιηθεί ως μία ετικέτα και $P(\text{features})$ είναι η a-priori πιθανότητα ότι ένα δεδομένο σύνολο από χαρακτηριστικά εμφανίζεται.²⁴⁴

Πέρα από τον Naive Bayes, οι αλγόριθμοι που συναντάει κανείς περισσότερο στην βιβλιογραφία και χρησιμοποιούνται ευρέως για κατηγοριοποίηση γνώμης, είναι οι παρακάτω:

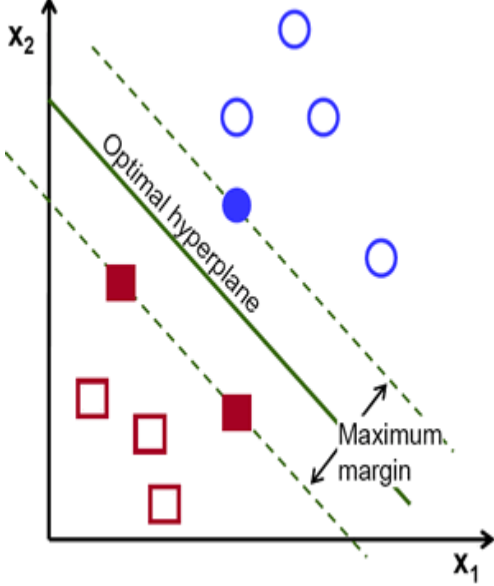
²⁴⁴ Pang, B., Lee, L., & Vaithyanathan, S., Thumbs up?: sentiment classification using machine learning techniques, in Proceedings of the ACL-02 conference on Empirical methods in natural language processing, 10, 2002, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://www.cs.cornell.edu/home/llee/papers/sentiment.pdf>, [προσπελάστηκε στις 28 Φεβρουαρίου 2017], σελ. 81-82.

<p>Hidden Markov Model</p>	<p>Η αλυσίδα Markov (Markov chain) είναι ένα μοντέλο πιθανότητας, το οποίο μεταβαίνει από μία κατάσταση σε επόμενη κατάσταση μέσα από ένα πεπερασμένο σύνολο πιθανών καταστάσεων.</p> <p>Βασικό χαρακτηριστικό της αλυσίδας Markov είναι το σε ποια κατάσταση θα μεταβεί το σύστημα εξαρτάται από την τωρινή κατάσταση και όχι από την προηγούμενη αλληλουχία καταστάσεων που αυτό διένυσε.</p> <p>Η ιδιότητα αυτή καλείται έλλειψη μνήμης. Έτσι, η μετάβαση από τη μία κατάσταση στην επόμενη γίνεται βάσει των πιθανοτήτων μετάβασης.</p> <p>Επιπλέον, το σύστημα παράγει μία έξοδο, ανάμεσα στις μεταβάσεις του, από ένα σύνολο συμβόλων εξόδου, η κατανομή πιθανότητας των οποίων είναι διαφορετική για κάθε κατάσταση.²⁴⁵</p>	 <p>The diagram illustrates a Hidden Markov Model (HMM) with three hidden states, X_1, X_2, and X_3, and four observed states, y_1, y_2, y_3, and y_4. The hidden states are represented by ovals, and the observed states are represented by rectangles. Transitions between hidden states are labeled with a_{ij} (e.g., a_{12}, a_{21}, a_{23}). Transitions from hidden states to observed states are labeled with b_{ij} (e.g., b_{11}, b_{12}, b_{13}, b_{14}, b_{21}, b_{22}, b_{23}, b_{24}, b_{31}, b_{32}, b_{33}, b_{34}).</p>
-----------------------------------	---	--

Εικόνα 26. Το μοντέλο της αλυσίδας Markov²⁴⁶

²⁴⁵ Pang, B., Lee, L., & Vaithyanathan, S., Thumbs up?: sentiment classification using machine learning techniques, ό.π. παραπομπή 244, σελ. 79–86.

²⁴⁶ Βλ. http://homepages.inf.ed.ac.uk/group/sli_archive/slip0809_c/s0562005/theory.html

<p>Support Vector Machines</p>	<p>Η βασική αρχή της μεθόδου αυτής είναι ο εντοπισμός στο χώρο αναζήτησης των γραμμικών διαχωριστών (linear separators), οι οποίοι μπορούν καλύτερα να χωρίσουν τις διαφορετικές κλάσεις, έτσι ώστε όλα τα στοιχεία της κάθε κλάσης να έχουν το μεγαλύτερο διάστημα από τα στοιχεία της άλλης κλάσης²⁴⁷.</p>	 <p>Εικόνα 27. Το μοντέλο Support Vector Machines²⁴⁸</p>
<p>Maximum Entropy</p>	<p>Η μέθοδος αυτή κωδικοποιεί σύνολα από χαρακτηριστικά (feature sets), τα οποία έχουν ήδη κάποια ετικέτα (label), σε διανύσματα. Αυτό το κωδικοποιημένο διάνυσμα χρησιμοποιείται για να υπολογιστεί ένα βάρος για κάθε χαρακτηριστικό, τα οποία στη συνέχεια συνδυάζονται για να καθοριστεί το η πιο πιθανή ετικέτα για το σύνολο χαρακτηριστικών.²⁴⁹</p>	$P(c d, \lambda) \stackrel{def}{=} \frac{\exp \sum_i \lambda_i f_i(c, d)}{\sum_{c' \in C} \exp \sum_i \lambda_i f_i(c', d)}$ <p>Εικόνα 28. Το μοντέλο Maximum Entropy²⁵⁰</p>

Πίνακας 2. Αλγόριθμοι μηχανικής μάθησης

²⁴⁷ Pang, B., Lee, L., & Vaithyanathan, S., Thumbs up?: sentiment classification using machine learning techniques, ό.π. παραπομπή 244, σελ. 82.

²⁴⁸ Βλ. http://docs.opencv.org/2.4/doc/tutorials/ml/introduction_to_svm/introduction_to_svm.html

²⁴⁹ Pang, B., Lee, L., & Vaithyanathan, S., Thumbs up?: sentiment classification using machine learning techniques, ό.π. παραπομπή 244, σελ. 82.

²⁵⁰ Βλ. <http://sentiment.christopherpotts.net/classifiers.html#maxent>

Συμπερασματικά, θα μπορούσε να υποθέσει κανείς ότι οι τεχνικές που χρησιμοποιούν μεθόδους μηχανικής μάθησης μπορούν να παρέχουν καλύτερα αποτελέσματα από τις τεχνικές με λεξικό. Οι τεχνικές με λεξικό φαίνεται να έχουν σαφείς περιορισμούς, καθώς λαμβάνουν κυρίως υπ' όψιν κάθε λέξη ξεχωριστά, πράγμα που οδηγεί πολύ εύκολα σε αστοχίες. Οι τεχνικές machine learning, μπορεί να απαιτούν μεγάλο όγκο δεδομένων για να εκπαιδευτούν, όμως εστιάζουν πέρα από κάθε λέξη και στη σύνδεση με τις υπόλοιπες λέξεις για να κάνουν την κατηγοριοποίηση, γεγονός που προσεγγίζει περισσότερο την ανθρώπινη κρίση.²⁵¹ Επιπλέον, στην παρούσα έρευνα, όπου τα δεδομένα που εξετάζονται είναι από το Twitter, στο οποίο η γλώσσα και ο τρόπος που τη χειριζόμαστε είναι πολύ διαφορετικός, μία τεχνική με λεξικό είναι πιθανό να μην έχει τα επιθυμητά αποτελέσματα. Προκειμένου να αποδειχτεί αυτό, θα εξεταστούν για δεδομένα που ελήφθησαν από το Twitter API και οι δύο μέθοδοι συγκριτικά, μία μέθοδος βασισμένη σε λεξικό και μία μηχανικής μάθησης και με τη χρήση του ταξινομητή Naive Bayes.

3.2.6 ΔΥΣΚΟΛΙΕΣ ΚΑΙ ΠΡΟΚΛΗΣΕΙΣ

Αρκετοί είναι οι παράγοντες που καθιστούν δύσκολη την εξόρυξη και ανάλυση απόψεων. Αρχικά, ο τομέας του opinion mining είναι ένας σχετικά νέος ερευνητικός τομέας, και η διαδικασία της εξόρυξης και ανάλυσης απόψεων θεωρείται ως μία δύσκολη διαδικασία για την οποία δεν υπάρχει κάποια συγκεκριμένη προσέγγιση που να προσφέρει ικανοποιητικά αποτελέσματα. Οι δυσκολίες απορρέουν από το γεγονός ότι απαιτείται από τους αλγορίθμους να είναι σε θέση να κατανοούν πλήρως τη σημασιολογική πληροφορία των κειμένων, καθώς επίσης και από τη φύση των απόψεων, οι οποίες συχνά εκφράζονται με σύνθετο και πολύπλοκο τρόπο. Μέχρι στιγμής, δεν απαιτείται από τις μηχανές και τους αλγορίθμους να αντιλαμβάνονται πλήρως τη σημασιολογική των κειμένων, αλλά να εντοπίζουν τη στάση του συγγραφέα και να την ταξινομούν ως προς την πολικότητά της. Αυτή η διαδικασία βέβαια πολλές φορές παρουσιάζει δυσκολίες ακόμα και για τον άνθρωπο.

²⁵¹ Pang, B., Lee, L., & Vaithyanathan, S., Thumbs up?: sentiment classification using machine learning techniques, ό.π. παραπομπή 244, σελ. 79.

Ακόμα και αν ο υπολογιστής ήταν σε θέση να διεκπεραιώσει κάτι τέτοιο, τα αποτελέσματα πάλι δεν θα ήταν 100% ικανοποιητικά, διότι και οι ίδιοι οι άνθρωποι αντιμετωπίζουν προβλήματα και συχνά παρατηρούνται διαφορετικές προσεγγίσεις των κειμένων. Συγκεκριμένα, οι ανθρώπινοι βαθμολογητές δείχνουν να συμφωνούν σε ποσοστό της τάξης του 80%.²⁵²

Για την κατηγοριοποίηση της πολικότητας των απόψεων, οι οποίες εκφράζονται για οντότητες και τα χαρακτηριστικά τους, σε απόψεις με θετικό, αρνητικό και ουδέτερο συναίσθημα, σημαντικά σημεία αποτελούν οι sentiment λέξεις. Η ύπαρξη sentiment λέξεων σε μία πρόταση συνήθως υποδηλώνει την ύπαρξη κάποιας άποψης, γεγονός όμως που δεν συμβαίνει πάντα. Ορισμένες απόψεις εκφράζονται χωρίς τη χρήση sentiment λέξεων ή αντίστοιχα μία πρόταση μπορεί να περιέχει sentiment λέξεις αλλά να μην εκφράζει άποψη. Πιο εξειδικευμένες τεχνικές για την αναγνώριση και κατηγοριοποίηση των απόψεων βασίζονται σε συντακτικά πρότυπα, τα οποία χρησιμοποιούνται συχνά για την απόδοση των απόψεων στον ανθρώπινο λόγο.

Επίσης, υπάρχουν αρκετά είδη απόψεων όπως γενικές, συγκριτικές, άμεσες και έμμεσες. Το πλήθος των διαφορετικών μορφών με τους οποίους μπορεί να εκφραστεί μία άποψη καθιστά δύσκολη την κατηγοριοποίησή τους σε θετικές, αρνητικές και ουδέτερες απόψεις. Επίσης, η ύπαρξη sentiment λέξεων θετικής ή αρνητικής πολικότητας, σε μια πρόταση η οποία εκφράζει άποψη, δεν συνεπάγεται απαραίτητα πως η πολικότητα της άποψης θα είναι θετική ή αρνητική αντίστοιχα. Ο προσδιορισμός της πολικότητας των απόψεων, πρέπει να πραγματοποιηθεί με προσοχή, διότι πρέπει να ληφθούν υπόψιν σχήματα άρνησης, σαρκασμού, σχήματα σύγκρισης και άλλα. Η ποικιλομορφία της ανθρώπινης έκφρασης είναι αυτή που αποτελεί έναν από τους βασικότερους παράγοντες οι οποίοι καθιστούν δύσκολη μία διαδικασία ανάλυσης απόψεων.

Επιπλέον, πολλές φορές το συναίσθημα ενός κειμένου εξαρτάται από το σημασιολογικό πλαίσιο στο οποίο τοποθετείται. Για παράδειγμα, η ίδια πρόταση μπορεί να έχει

²⁵² Παναγόπουλος, Π., 2015, *Opinion Mining and Sentiment Analysis*, ό.π. παραπομπή 213, σελ. 13.

διαφορετική πολικότητα σε διαφορετικό περιεχόμενο.²⁵³ Εκτός αυτού, σε μία διαδικασία ανάλυσης συναισθήματος, θα πρέπει να προσπελαστούν οι απόψεις πολλών χρηστών διότι, λόγω της υποκειμενικής φύσης των απόψεων, αναλύοντας μόνο τις απόψεις ενός χρήστη ή ενός μικρού συνόλου χρηστών, δεν αναγνωρίζεται η συνολική εικόνα της κατανομής των απόψεων.²⁵⁴

Άλλος ένας παράγοντας που μπορεί να επηρεάσει το συναίσθημα είναι η σειρά των λέξεων μέσα στο κείμενο, καθώς οι ίδιες λέξεις αλλά τοποθετημένες σε διαφορετική σειρά μπορούν να οδηγήσουν στην αντίθετη πολικότητα. Τέλος, οι προκλήσεις της Επεξεργασίας Φυσικής Γλώσσας (NLP) αποτελούν δυσκολίες που συναντάει η ανάλυση συναισθήματος, όπως λ.χ. η αμφισημία, ο χειρισμός της άρνησης, η ειρωνεία και ο σαρκασμός.

Όσον αφορά στην ανάλυση συναισθήματος σε επίπεδο πρότασης, με την οποία καταπιάνεται η παρούσα έρευνα, εντοπίζονται τα εξής προβλήματα:

- Πολύπλοκες προτάσεις περιέχουν διαφορετικές απόψεις για διαφορετικούς στόχους (opinion targets).
- Ακόμα και αν μία πρόταση έχει γενικά θετική ή αρνητική πολικότητα, κάποια μέρη της ίσως εκφράζουν αντίθετες απόψεις, όπως φαίνεται για παράδειγμα στην πρόταση *«Αν και η ανεργία βρίσκεται σε υψηλά επίπεδα, η οικονομία είναι σε καλή κατάσταση»*.
- Η ανάλυση σε επίπεδο πρότασης δεν μπορεί να αντιμετωπίσει την περίπτωση των συγκριτικών προτάσεων (comparative sentences). Για παράδειγμα: *«Η γεύση του καφέ είναι καλύτερη από αυτή του χυμού βύσσινο»*. Αν και η πρόταση ξεκάθαρα εκφράζει άποψη, δεν μπορούμε να την κατηγοριοποιήσουμε σε θετική, αρνητική ή ουδέτερη.²⁵⁵

²⁵³ Pang B., & Lee, L., Opinion mining and sentiment analysis, *Foundations and Trends in Information Retrieval*, 2 (1-2), 2008, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.cs.cornell.edu/home/llee/omsa/omsa.pdf>, [προσπελάστηκε στις 5 Αυγούστου 2017], σελ. 11.

²⁵⁴ Παναγόπουλος, Π., 2015, *Opinion Mining and Sentiment Analysis*, ό.π. παραπομπή 213, σελ. 15-16.

²⁵⁵ Ό.π., σελ. 36.

- Η υπόθεση που γίνεται στις περισσότερες έρευνες για την sentiment κατηγοριοποίηση μιας πρότασης, πως μία πρόταση δηλαδή περιέχει μόνο μία ή καμία άποψη, οδηγεί στην προσέγγιση πως οι sentiment λέξεις της πρότασης, εφόσον υπάρχουν, θα αποκαλύπτουν την πολικότητα αυτής της άποψης ή την ουδετερότητα της αντίστοιχα. Αυτή η προσέγγιση, γενικά, δεν μπορεί να εφαρμοστεί σε μία υπό συνθήκη πρόταση όπως «*Αν η x βιβλιοθήκη δεν σε καλύπτει, προσπάθησε να δουλέψεις με την y*». Σε αυτή την πρόταση το πρώτο μέρος δεν εκφράζει κάποια άποψη, αλλά το δεύτερο εκφράζει θετική άποψη. Η παραπάνω πρόταση δεν περιέχει καμία sentiment λέξη, αλλά περιέχει άποψη. Για τη διαχείριση των υπό συνθήκη προτάσεων συνήθως χρησιμοποιούνται εποπτευόμενες τεχνικές, οι οποίες χρησιμοποιούν σύνολα γλωσσολογικών χαρακτηριστικών (linguistic features), όπως η θέση συγκεκριμένων λέξεων στις προτάσεις, POS tags, πρότυπα, και συντακτικοί σύνδεσμοι.
- Ένα άλλο είδος δύσκολων προτάσεων, ως προς την διαχείριση, είναι οι προτάσεις ερωτήσεων, όπως για παράδειγμα «*Μπορεί κάποιος να μου προτείνει ένα μοντέλο smartphone με καλή οθόνη;*», η οποία δεν περιέχει κάποια άποψη, αν και περιέχει μία sentiment λέξη, η οποία στοχεύει ένα χαρακτηριστικό μιας οντότητας, στη συγκεκριμένη περίπτωση την καλή οθόνη του smartphone.
- Επίσης υπάρχουν και οι προτάσεις οι οποίες περιέχουν σαρκασμό. Αυτό το είδος προτάσεων είναι αρκετά συχνό στις συζητήσεις στο διαδίκτυο, όπως και στα tweets. Οι προτάσεις σαρκασμού είναι μία ειδική μορφή του λόγου, με την οποία ο ομιλητής ή συγγραφέας λέει ή γράφει το αντίθετο από αυτό που πραγματικά εννοεί. Ο σαρκασμός έχει μελετηθεί στη γλωσσολογία, την ψυχολογία και τη γνωστική επιστήμη. Όσον αφορά στη διαδικασία της ανάλυσης συναισθήματος, όταν εντοπίζεται ένα τέτοιο είδος πρότασης, η πολικότητά της πρέπει να αντιστρέφεται. Αν και αυτό το είδος προτάσεων, όπως είναι φανερό, είναι δύσκολο να εντοπιστεί και να διαχειριστεί από έναν αλγόριθμο, υπάρχουν έρευνες οι οποίες προτείνουν τρόπους για την διαχείρισή του.

Στην έρευνα González-Ibáñez et al. μελετήθηκε το πρόβλημα της αναγνώρισης των σαρκαστικών προτάσεων σε δεδομένα από το Twitter. Συγκεκριμένα,

προτάθηκε μία τεχνική εποπτευόμενης μάθησης, η οποία χρησιμοποιεί Support Vector Machines (SVM) και παλινδρόμηση (logistic regression). Ως χαρακτηριστικά χρησιμοποιήθηκαν unigrams και κάποιες πληροφορίες βασισμένες σε λεξικό, όπως κατηγορίες λέξεων, επιφωνήματα, σημεία στίξης, καθώς επίσης και emoticons και ToUser (το οποίο δείχνει πως ένα tweet είναι απάντηση σε κάποιο tweet που έθεσε ο χρήστης <@user>). Τα πειράματα έδειξαν πως η κατηγοριοποίηση των προτάσεων στις τρεις κλάσεις (σαρκαστικές, θετικές, αρνητικές), μπορεί να επιτύχει ακρίβεια μόνο 57%.²⁵⁶

Συμπερασματικά, οι περισσότερες από τις έρευνες πάνω στην κατηγοριοποίηση συναισθήματος στο επίπεδο πρότασης επικεντρώνονται στη λύση του γενικού προβλήματος, χωρίς να λαμβάνουν υπ' όψιν τους το γεγονός πως υπάρχουν διαφορετικά είδη προτάσεων τα οποία επιβάλουν ειδική μεταχείριση.

3.2.7 ΑΝΑΛΥΣΗ ΣΥΝΑΙΣΘΗΜΑΤΟΣ ΣΤΟ TWITTER

Σε έρευνα πάνω στην ανάλυση συναισθήματος στο Twitter με δεδομένα γύρω από την προσφυγική κρίση²⁵⁷, έγινε προσπάθεια να εντοπιστεί ο τρόπος που τα γεγονότα στον «πραγματικό» κόσμο επηρεάζουν το συναίσθημα των χρηστών του εν λόγω μέσου, όπως αυτό εκφράζεται μέσα από τα tweets που αναρτούν. Στο πλαίσιο αυτό, συλλέχθηκαν και αναλύθηκαν tweets στην αγγλική και στη γερμανική γλώσσα και με τη χρήση του εργαλείου Linguistic Inquiry and Word Count (LIWC)²⁵⁸ έγινε προσπάθεια να εντοπιστεί το θετικό και το αρνητικό συναίσθημα των αναρτήσεων στο Twitter, καθώς και το συναίσθημα του άγχους, του θυμού και της λύπης. Τα αποτελέσματα έδειξαν ότι το συναίσθημα στο Twitter επηρεάζεται από τα γεγονότα

²⁵⁶ González-Ibáñez, R., Muresan, S., & Wacholder, N., Identifying sarcasm in Twitter: a closer look., in Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: shortpapers (ACL-2011), 2011, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://pdfs.semanticscholar.org/55e3/6d6b45c91a0daa49234bd47b856470d6825c.pdf>, [προσπελάστηκε στις 28 Φεβρουαρίου 2017], σελ. 583.

²⁵⁷ Pope, D., & Griffith, J., An Analysis of Online Twitter Sentiment Surrounding the European Refugee Crisis., in Proceedings of the 8th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: https://nuigulresearchday2016.files.wordpress.com/2016/04/nuigul-2016_paper_38.pdf, [προσπελάστηκε στις 22 Φεβρουαρίου 2017].

²⁵⁸ Βλ. <https://liwc.wpengine.com/>

αυτά, όπως για παράδειγμα από την τρομοκρατική επίθεση στο Παρίσι τον Νοέμβριο του 2015. Αυτό που δεν εντοπίστηκε από την εν λόγω έρευνα είναι αν το άγχος και το αρνητικό συναίσθημα των χρηστών έχει ως στόχο την προσφυγική κρίση ευρύτερα ή τους ίδιους τους πρόσφυγες.

Στην παρούσα έρευνα θα γίνει προσπάθεια να εντοπιστεί σε ποιον απευθύνεται κάθε φορά το συναίσθημα, ποιος είναι δηλαδή ο στόχος της άποψης (opinion target).

Σε έρευνα τον Ιούνιο του 2011 από το πανεπιστήμιο του Ιλινόις²⁵⁹ συνδυάστηκε η μέθοδος βασισμένη σε λεξικό με τη μηχανική μάθηση και συγκεκριμένα τον αλγόριθμο Support Vector Machines (SVM), προκειμένου να επιτευχθεί μεγαλύτερο ποσοστό επιτυχίας στην ανάλυση συναισθήματος στο Twitter. Στο πλαίσιο αυτό, έγινε ανάλυση σε επίπεδο οντοτήτων και όχι σε επίπεδο πρότασης, όπως συνηθίζεται να γίνεται με τα tweets και η μέθοδος του λεξικού φάνηκε να έχει μεγάλη ακρίβεια, αλλά σχετικά μικρό βαθμό ανάκλησης, καθώς η εν λόγω μέθοδος, προκειμένου να προσδιορίσει την κατεύθυνση του συναισθήματος, βασίζεται αποκλειστικά στην παρουσία λέξεων που εκφράζουν άποψη. Προκειμένου να αντιμετωπιστεί το εν λόγω πρόβλημα, χρησιμοποιήθηκε ο αλγόριθμος SVM, αλλά αντί να πραγματοποιηθεί χειρωνακτική επισημείωση στα tweets για την εκπαίδευση του συστήματος, χρησιμοποιήθηκαν τα αποτελέσματα από τη μέθοδο με το λεξικό, γεγονός που επέφερε τα επιθυμητά αποτελέσματα, έναν μεγαλύτερο δηλαδή βαθμό ανάκλησης.

Σε έρευνα του Οικονομικού Πανεπιστημίου Αθηνών²⁶⁰ πραγματοποιήθηκε ανάλυση συναισθήματος σε tweets στην ελληνική και στην αγγλική γλώσσα. Για το σκοπό αυτό συλλέχθηκαν και επισημειώθηκαν χειροκίνητα tweets στα Ελληνικά, ενώ για τα αντίστοιχα κείμενα στα Αγγλικά χρησιμοποιήθηκαν έτοιμα επισημειωμένα tweets. Συγκεκριμένα για τα ελληνικά tweets πραγματοποιήθηκε χειρωνακτική επισημείωση 8.000 tweets για την εκπαίδευση του συστήματος και επιπλέον 888 tweets για να πραγματοποιηθεί το τεστ. Για τα αγγλικά tweets χρησιμοποιήθηκαν έτοιμα

²⁵⁹ Lei, Z., Hsu, M., Liu, B., Ghosh, R., & Dekhil, M., Combining Lexicon-based and Learning-based Methods for Twitter Sentiment Analysis., *HP Laboratories*, 2011, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.hpl.hp.com/techreports/2011/HPL-2011-89.pdf>, [προσπελάστηκε στις 18 Μαρτίου 2017].

²⁶⁰ Makrynioti, N., & Vassalos, V., Sentiment Extraction from Tweets: Multilingual Challenges, *Big Data Analytics and Knowledge Discovery Lecture Notes in Computer Science*, 2015, σελ. 136-148.

επισημειωμένα tweets από τους οργανωτές του SemEval 2013.²⁶¹ Κατόπιν πραγματοποιήθηκαν πειράματα και αξιολογήθηκαν τρεις μέθοδοι ανάλυσης συναισθήματος βασισμένες στη μηχανική μάθηση, με τα δεδομένα των δύο γλωσσών να μελετώνται παράλληλα, εξετάζοντας την αποδοτικότητα των τριών μεθόδων στην κάθε γλώσσα και τη δυνατότητα γενίκευσης των υπό μελέτη τεχνικών, προκειμένου να αντιμετωπιστούν οι πολυγλωσσικές προκλήσεις.

Στην παρούσα έρευνα πραγματοποιείται ανάλυση συναισθήματος σε δεδομένα στην ελληνική και στη γερμανική γλώσσα ταυτόχρονα και αξιολογείται η αποδοτικότητα των εργαλείων στην κάθε μία από αυτές.

Σε ανάλογη έρευνα ανάλυσης συναισθήματος πολύγλωσσων tweets²⁶² πραγματοποιήθηκε μηχανική μετάφραση αγγλικών tweets σε άλλες τέσσερις γλώσσες, Ιταλικά, Γαλλικά, Γερμανικά και Ισπανικά. Η ίδια μέθοδος ανάλυσης συναισθήματος εφαρμόστηκε σε κάθε μία γλώσσα ξεχωριστά, σε ζευγάρια γλωσσών συνδυαστικά, αλλά και στις πέντε διαφορετικές γλώσσες ταυτόχρονα, όπου και παρατηρήθηκε ότι ήταν πιο αποδοτική. Τα περίπου 1000 tweets που χρησιμοποιήθηκαν ήταν έτοιμα επισημειωμένα tweets από τους οργανωτές του SemEval 2013.²⁶³ Ο συγκεκριμένος αριθμός των tweets επιλέχθηκε, καθώς ήταν αρκετά μεγάλος, ώστε να εξαχθούν συμπεράσματα, και αρκετά μικρός, ώστε να καταστεί δυνατή η διόρθωση των λαθών της μηχανικής μετάφρασης, που χωρίς αυτή θα αλλοιώνονταν τα αποτελέσματα.

Η μέθοδος που χρησιμοποιήθηκε ήταν μία υβριδική μέθοδος που συνδύασε την εποπτευόμενη τεχνική μηχανικής μάθησης και συγκεκριμένα τον αλγόριθμο Support Vector Machines με το πρόγραμμα Weka²⁶⁴, το οποίο έχει σχεδιαστεί για να προβλέπει το συναίσθημα σε νέα tweets με βάση τις πληροφορίες που έχει από άλλα, ήδη επισημειωμένα, tweets. Η έρευνα κατέληξε ότι τα δεδομένα από διαφορετικές γλώσσες, ειδικά από αυτές που έχουν παρόμοια δομή, αν συνδυαστούν, μπορούν να πετύχουν στατιστικά καλύτερα αποτελέσματα από ότι τα δεδομένα από μία μόνο γλώσσα.

²⁶¹ Nakov, P., Rosenthal, S., Kozareva, Z., Stoyanov, V., Ritter, A., & Wilson, T., Semeval-2013 task 2: Sentiment analysis in twitter, in Proceedings of the Seventh International Workshop on Semantic Evaluation, 2013, σελ. 312- 320.

²⁶² Balahur, A., & Turchi, M., Improving Sentiment Analysis in Twitter Using Multilingual Machine Translated Data., in Proceedings of Recent Advances in Natural Language Processing, 2013, σελ. 49-55.

²⁶³ Ό.π.

²⁶⁴ Βλ. <http://www.cs.waikato.ac.nz/ml/weka/>

Σε γενικές γραμμές μπορεί να παρατηρήσει κανείς ότι η ανάλυση συναισθήματος στο Twitter, σε σύγκριση με αυτήν σε άλλα κοινωνικά δίκτυα, συνεπάγεται τόσο ορισμένες διευκολύνσεις, όσο και ορισμένες δυσκολίες, οι οποίες οφείλονται στην ιδιαίτερη φύση του εν λόγω μέσου.

Ανάμεσα στα χαρακτηριστικά που κάνουν το Twitter να υπερτερεί, όσον αφορά στην ανάλυση συναισθήματος, βρίσκεται ο μοναδικός τρόπος αλληλεπίδρασης. Συγκεκριμένα, το συναίσθημα «εγκλωβίζεται» στα σύντομα μηνύματα των έως 140 χαρακτήρων που μπορούν να δημοσιεύσουν οι χρήστες. Επίσης, το μεγαλύτερο μέρος των δεδομένων στο Twitter είναι ελεύθερο στο κοινό, μέσω του Twitter API, και αυτό καθιστά δυνατή τη συλλογή ενός μεγάλου αριθμού δεδομένων. Εκτός αυτού, τα tweets περιέχουν χρονοσφραγίδα (timestamp), δείχνοντας έτσι την αλληλουχία των γεγονότων.

Η ιδιαίτερη φύση των κειμένων στο Twitter μπορεί όμως να προκαλέσει και δυσκολίες στην εφαρμογή της ανάλυσης συναισθήματος. Το περιορισμένο μήκος των κειμένων συνεπάγεται απουσία του ευρύτερου εννοιολογικού πλαισίου, με αποτέλεσμα να μην είναι συχνά σαφής η πολικότητα του κειμένου.²⁶⁵ Όσον αφορά στο λεξιλόγιο, τα περισσότερα κείμενα αυτού του είδους διατυπώνονται σε ανεπίσημη, καθομιλουμένη γλώσσα και εμφανίζουν ποικιλομορφία σε πολύ μεγαλύτερο βαθμό σε σχέση με άλλα είδη κειμένου. Συγκεκριμένα, στα tweets περιλαμβάνονται νεολογισμοί, συντομογραφίες, αργκό, αλλά και εσκεμμένες παραλλαγές των λέξεων που εξυπηρετούν συγκεκριμένους επικοινωνιακούς σκοπούς.

Όλα τα παραπάνω καθιστούν δύσκολη την εφαρμογή λεξικών που στηρίζονται στην πιο επίσημη μορφή της γλώσσας. Επιπλέον, το γεγονός ότι οι πλατφόρμες επιτρέπουν την αυθόρμητη επικοινωνία σε πραγματικό χρόνο, οδηγεί τους χρήστες στο να δημοσιεύουν μηνύματα, χωρίς να ελέγχουν για γραμματικά, συντακτικά ή ορθογραφικά λάθη, τα οποία συνιστούν ουσιαστικά θόρυβο για την ανάλυση συναισθήματος. Τέλος, το περιεχόμενο στα μέσα κοινωνικής δικτύωσης είναι

²⁶⁵ Bermingham, A. & Smeaton, A., Classifying sentiment in 93 microblogs: is brevity an advantage?, in Proceedings of the 19th ACM international conference on Information and knowledge management, ACM, 2010, σελ. 1833–1836.

πολυγλωσσικό, καθώς οι χρήστες προέρχονται από όλο τον κόσμο, με αποτέλεσμα να γράφουν σε διαφορετικές γλώσσες ακόμα και σε επίπεδο πρότασης ή μηνύματος, γεγονός που πολλές φορές καθιστά μη εφαρμόσιμες ακόμα και τεχνικές στοχευμένες σε συγκεκριμένες γλώσσες (language-specific).

ΕΡΕΥΝΗΤΙΚΕΣ ΥΠΟΘΕΣΕΙΣ – ΕΡΕΥΝΗΤΙΚΑ ΕΡΩΤΗΜΑΤΑ

Παρακάτω παρουσιάζονται οι κύριες ερευνητικές υποθέσεις και τα ερευνητικά ερωτήματα της παρούσας έρευνας:

E.Y. 1: *Οι Έλληνες και οι Γερμανοί πολίτες φαίνεται να τάσσονται κατά της εγκατάστασης των προσφύγων στη χώρα τους και κατά της ένταξης των παιδιών προσφύγων στο εκπαιδευτικό σύστημα.*

Το ζήτημα της μόνιμης εγκατάστασης των προσφύγων, όπως αναφέρεται παραπάνω, δε γίνεται αποδεκτό, σύμφωνα με έρευνες, από την πλειοψηφία της κοινής γνώμης τόσο της Ελλάδας, όσο και της Γερμανίας. Ενώ η πλειοψηφία των ερωτηθέντων από τη μία πλευρά τάσσεται υπέρ της προσωρινής διαμονής των προσφύγων στην εκάστοτε χώρα, από την άλλη, δηλώνει ότι θα ήθελε οι πρόσφυγες να γυρίσουν μετά από κάποιο διάστημα πίσω στην πατρίδα τους. Με την έναρξη της σχολικής χρονιάς και στις δύο χώρες, ιδιαιτέρως στην Ελλάδα, υπήρξαν κάποιες ακραίες αντιδράσεις, κατά κύριο λόγο από γονείς, οι οποίοι τάχθηκαν κατά της ένταξης των παιδιών προσφύγων στα σχολεία.

E.Y.2: *Οι Γερμανοί δείχνουν να δέχονται σε μεγαλύτερο βαθμό τους πρόσφυγες από ό,τι οι Έλληνες.*

Σύμφωνα με έρευνες, οι Γερμανοί πολίτες εμφανίζονται πιο ανοιχτοί απέναντι στους πρόσφυγες από ό,τι οι Έλληνες. Για παράδειγμα, όπως αναφέρθηκε παραπάνω, σε έρευνα της Διεθνούς Αμνηστίας η Γερμανία κατέλαβε τη 2η θέση στο «Δείκτη Υποδοχής Προσφύγων» με ποσοστό 84%, ενώ η Ελλάδα την 7η με ποσοστό 65%.

E.Y.3: *Τα Μ.Μ.Ε εκφράζουν ουδέτερο συναίσθημα.*

Η κάλυψη της ειδησεογραφίας οφείλει να διακρίνεται στις μέρες μας από τη λεγόμενη «αντικειμενικότητα» και «αξιολογική ουδετερότητα».²⁶⁶ Μέσα από τα μέσα μαζικής επικοινωνίας θα πρέπει να διασφαλίζεται η αντικειμενικότητα της

²⁶⁶ Λεμπέσης, Ε., Κοινωνιολογία του Τύπου, Αρσενίδης, 1992, σελ. 61.

πληροφόρησης. Επομένως, το συναίσθημα που εκφράζεται μέσα από τα tweets των media αναμένεται να είναι κατά κύριο λόγο ουδέτερο.

Ερ.Ε.1: *Σε ποιους απευθύνεται το συναίσθημα που εκφράζουν οι Έλληνες και Γερμανοί στο Twitter αναφορικά με την εκπαίδευση των προσφύγων;*

Στόχος της εξαγωγής των opinion targets είναι το να γίνει αντιληπτό σε ποιες πτυχές του θέματος της εκπαίδευσης των προσφύγων αναφέρονται με τις αναρτήσεις τους οι χρήστες του Twitter. Ο εντοπισμός και η ανάλυση των opinion targets πραγματοποιείται, καθώς κρίνεται σημαντικό να γνωρίζει κανείς, όχι μόνο ποιο είναι το συναίσθημα και ποιος το εκφράζει, αλλά και σε ποιον απευθύνεται το εκάστοτε συναίσθημα.

Ερ.Ε. 2: *Συμβαδίζει το συναίσθημα που εκφράζουν τα Μ.Μ.Ε της κάθε χώρας με αυτό των πολιτών της;*

Τα Μ.Μ.Ε αποτελούν τους κυριότερους διαμορφωτές της κοινής γνώμης, καθώς έχουν τη δύναμη να επηρεάσουν και να προδιαθέσουν ευμενώς ή δυσμενώς το κοινό για πρόσωπα, γεγονότα ή καταστάσεις. Επομένως, θα διερευνηθεί κατά πόσο το συναίσθημα που εκφράζουν μέσω του Twitter τα μέσα μαζικής επικοινωνίας για την εκπαίδευση των προσφύγων σχετίζεται με αυτό που εκφράζουν οι πολίτες.

Ερ.Ε.3: *Πώς κρίνουν οι πολίτες, Έλληνες και Γερμανοί, τη στάση της ακροδεξιάς απέναντι στο θέμα της εκπαίδευσης των προσφύγων;*

Τα ακροδεξιά κόμματα, όπως αναφέρθηκε παραπάνω, εκμεταλλεύονται σε μεγάλο βαθμό την προσφυγική κρίση, προκειμένου να αυξήσουν τη δημοτικότητά τους. Στο πλαίσιο αυτό τείνουν να παρουσιάζουν τους πρόσφυγες ως ανθρώπους χωρίς δικαιώματα, ενώ τους κατηγορούν ως δυνάμει τρομοκράτες ή υπαίτιους τέλεσης σεξουαλικών εγκλημάτων. Αναφορικά με το ζήτημα της εκπαίδευσης των προσφύγων, τα μέλη των εν λόγω κομμάτων αντιδρούν έντονα, ενώ προκαλούν και επεισόδια σε σχολεία, για να αποτρέψουν την ένταξη των παιδιών-προσφύγων σε αυτά. Στην παρούσα μελέτη θα μετρηθεί το συναίσθημα που

εκφράζουν οι πολίτες μέσω του Twitter για τη στάση της ακροδεξιάς σε σχέση με την εκπαίδευση των προσφύγων.

Ερ.Ε. 4: *Ποια μέθοδος ανάλυσης συναισθήματος εμφανίζεται να είναι πιο αποτελεσματική;*

Θα εφαρμοστούν δύο τεχνικές ανίχνευσης συναισθήματος, μία βασισμένη σε λεξικό και μία στη μηχανική μάθηση, έτσι ώστε να εντοπιστεί η πιο αποτελεσματική από τις δύο για την παρούσα έρευνα.

Ερ.Ε. 5: *Ποια κατηγορία συναισθήματος (θετικό/αρνητικό/ουδέτερο) μπορεί να κατηγοριοποιηθεί σωστά με μεγαλύτερη επιτυχία;*

Αφού εντοπιστεί η πιο αποτελεσματική μέθοδος από τις δύο και χρησιμοποιηθεί για να γίνει πρόβλεψη του συναισθήματος για κάθε μία από τις ομάδες στόχων του συναισθήματος, θα υπολογιστεί ποια κατηγορία συναισθήματος, το θετικό, το αρνητικό ή το ουδέτερο συναίσθημα, προβλέπεται με μεγαλύτερη επιτυχία από τη μέθοδο που θα εφαρμοστεί.

ΜΕΘΟΔΟΛΟΓΙΑ ΕΡΕΥΝΑΣ

ΕΠΙΛΟΓΗ ΠΛΗΘΥΣΜΟΥ ΤΗΣ ΜΕΛΕΤΗΣ

Προς ανάλυση επιλέχθηκαν tweets τα οποία έχουν αναρτηθεί από λογαριασμούς στο Twitter που ανήκουν σε μέσα μαζικής επικοινωνίας, αλλά και σε απλούς χρήστες. Όλα τα tweets είναι γραμμένα στα Ελληνικά και στα Γερμανικά και η ανάρτησή τους έχει γίνει αποκλειστικά από την Ελλάδα και τη Γερμανία αντίστοιχα.

Ο εν λόγω πληθυσμός επιλέχτηκε, καθώς προέρχεται από δύο χώρες που έρχονται αντιμέτωπες με την προσφυγική κρίση σε πολύ μεγάλο βαθμό, αλλά με διαφορετικό τρόπο και υπό διαφορετικές συνθήκες, η μεν Ελλάδα ως χώρα υποδοχής και η δε Γερμανία ως χώρα τελικού προορισμού. Επιπλέον, η Ελλάδα βιώνει την ίδια περίοδο τη σφοδρή οικονομική κρίση των τελευταίων ετών, ενώ η Γερμανία κάνει άνοιγμα στο να δεχτεί και να ενσωματώσει μεγάλο μέρος των προσφύγων.

ΠΕΡΙΟΔΟΣ ΑΝΑΛΥΣΗΣ

Τα δεδομένα του Twitter που αναλύθηκαν προέρχονται από την περίοδο 1/7/2016 έως 31/3/2017. Η συγκεκριμένη περίοδος επιλέχθηκε σκοπίμως, καθώς περιλαμβάνει τους καλοκαιρινούς μήνες που είναι έντονη η προσφυγική μετακίνηση και κυρίως τον Σεπτέμβριο, τον μήνα που ξεκινούν τα σχολεία, αλλά και τους αμέσως επόμενους μήνες που γίνεται προσπάθεια ενσωμάτωσης των παιδιών- προσφύγων στο εκπαιδευτικό σύστημα της κάθε χώρας.

ΑΝΑΛΥΣΗ ΠΕΡΙΕΧΟΜΕΝΟΥ (CONTENT ANALYSIS)

Αρχικά, στην παρούσα μελέτη εφαρμόζεται η μέθοδος της ανάλυσης περιεχομένου. Συγκεκριμένα, εντοπίζονται οι στόχοι του συναισθήματος (opinion targets) των δεδομένων από το Twitter, αλλά και το ίδιο το συναίσθημα, κατόπιν της κατηγοριοποίησής του με χειροκίνητο τρόπο.

Στόχος του εντοπισμού των opinion targets είναι αρχικά το να γίνει αντιληπτό σε ποιες πτυχές του θέματος της εκπαίδευσης των προσφύγων αναφέρονται με τις αναρτήσεις

τους οι χρήστες του Twitter, πράγμα που φανερώνει τις όψεις του εν λόγω θέματος. Αφετέρου, με τον εντοπισμό και την ανάλυση των opinion targets, γίνεται πιο ουσιαστική η μετέπειτα ανάλυση του συναισθήματος, καθώς είναι σημαντικό να γνωρίζει κανείς όχι μόνο ποιο είναι το συναίσθημα και ποιος το εκφράζει, αλλά και σε ποιον απευθύνεται.

Μετέπειτα αναλύεται το συναίσθημα που εκφράζουν τα ελληνικά και γερμανικά Μ.Μ.Ε, αλλά και οι Έλληνες και Γερμανοί χρήστες, τόσο ευρύτερα για το θέμα της εκπαίδευσης των προσφύγων, αλλά και για κάθε έναν opinion target ξεχωριστά. Με αυτό τον τρόπο φαίνεται αν υπάρχει συνάφεια μεταξύ του συναισθήματος των media με αυτό των χρηστών σε κάθε χώρα, αλλά και οι ομοιότητες και οι διαφορές που υπάρχουν στο συναίσθημα που εκφράζεται από τους χρήστες του Twitter στη μία χώρα και στην άλλη.

ΑΝΑΛΥΣΗ ΣΥΝΑΙΣΘΗΜΑΤΟΣ (SENTIMENT ANALYSIS)

Στο επόμενο τμήμα της παρούσας έρευνας πραγματοποιείται ανάλυση συναισθήματος με δύο διαφορετικές μεθόδους, μία μη-εποπτευόμενης μάθησης, βασισμένη σε λεξικό και μία εποπτευόμενης μάθησης με τη χρήση του αλγορίθμου Naive Bayes.

Σκοπός της παραπάνω ανάλυσης είναι να διαπιστωθεί αν είναι δυνατόν να γίνει η κατηγοριοποίηση του συναισθήματος με αυτοματοποιημένο τρόπο χωρίς να απαιτείται ανθρώπινος κόπος για μία χειρωνακτική ταξινόμηση, η οποία είναι μία άκρως χρονοβόρα διαδικασία. Για να διαπιστωθεί αυτό, θα γίνει αξιολόγηση της αποδοτικότητας των δύο μεθόδων, θα διατυπωθούν υποθέσεις για τους λόγους που οι παραπάνω μέθοδοι επέτυχαν ή απέτυχαν να κατηγοριοποιήσουν σωστά το συναίσθημα και κατόπιν θα χρησιμοποιηθεί η πιο πετυχημένη στη συγκεκριμένη περίπτωση μέθοδος, προκειμένου να γίνει πρόβλεψη του συναισθήματος πιο αναλυτικά, για κάθε μία δηλαδή από τις ομάδες στόχων του συναισθήματος.

ΕΡΓΑΛΕΙΑ

• Η ΓΛΩΣΣΑ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΥ R

Η R είναι μία γλώσσα προγραμματισμού, η οποία συνοδεύεται από ένα περιβάλλον για επεξεργασία δεδομένων, υπολογισμών και γραφημάτων. Στην ουσία πρόκειται για ένα υπολογιστικό πακέτο που προσφέρει δυνατότητες διαχείρισης και στατιστικής ανάλυσης δεδομένων καθώς και δυνατότητες κατασκευής γραφημάτων. Βασίζεται στην γλώσσα προγραμματισμού S και αποτελεί ένα λογισμικό ανοικτού κώδικα (open source) που διατίθεται ελεύθερα.²⁶⁷ Καθώς η R δεν διαθέτει γραφικό περιβάλλον εργασίας, απαιτεί αρκετή ενασχόληση για να γίνει οικεία στο χρήστη, παρόλο που θεωρείται εύκολη στην εκμάθησή της. Το λογισμικό R γράφτηκε αρχικά από τους Ross Ihaka και Robert Gentleman στα μέσα της δεκαετίας του '90 (σε αυτούς οφείλει το όνομά του). Από το 1997 αναπτύσσεται από το R Development Core Team.²⁶⁸

Η R διαθέτει έναν μεγάλο αριθμό έτοιμων πακέτων με καλογραμμένα εγχειρίδια χρήσης, ενώ, όποιος επιθυμεί μπορεί να συνεισφέρει στα ήδη υπάρχοντα πακέτα και να προσφέρει νέες δυνατότητες στο σύστημα.

Η εγκατάσταση των πακέτων γίνεται με τη συνάρτηση `install.packages`, ενώ το φόρτωμά τους στο περιβάλλον πραγματοποιείται με τη συνάρτηση `library`.

Στην παρούσα έρευνα χρησιμοποιήθηκαν τα παρακάτω πακέτα, τα οποία συνέβαλλαν στην συλλογή, επεξεργασία και ανάλυση των δεδομένων:

- ROAuth (για πιστοποίηση της εφαρμογής και σύνδεση με το Twitter)
- twitter (για σύνδεση με το API του Twitter)
- methods (για την εκτέλεση βασικών εντολών με την R)
- tm (για την επεξεργασία των δεδομένων)
- stringr (για ανάλυση συναισθήματος με τη μέθοδο του λεξικού)

²⁶⁷ Venables, W. N., Smith, D. M. & the R Core Team, An Introduction to R, 2017, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://cran.r-project.org/doc/manuals/r-release/R-intro.pdf>, [προσπελάστηκε στις 5 Αυγούστου 2017], σελ. 2.

²⁶⁸ Αντζουλάκος, Δ., Ανάλυση Δεδομένων με τη Χρήση Στατιστικών Πακέτων- Εισαγωγή στο R, 2013, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: [http://www.unipi.gr/faculty/dantz/Introduction to R.pdf](http://www.unipi.gr/faculty/dantz/Introduction%20to%20R.pdf), [προσπελάστηκε στις 3 Αυγούστου 2017], σελ. 1.

- e1071 (για ανάλυση συναισθήματος με τον αλγόριθμο μηχανικής μάθησης Naive Bayes)

Στις περισσότερες γλώσσες προγραμματισμού υπάρχουν μεταβλητές και τύποι μεταβλητών. Ωστόσο η R βλέπει τα πάντα ως αντικείμενα (objects) που ανήκουν σε μία κλάση (class). Στην ουσία τα αντικείμενα αποτελούν τις μεταβλητές και η κλάση είναι ο τύπος των μεταβλητών αυτών. Στην R δεν απαιτείται η ρητή δήλωση της κλάσης στην οποία ανήκουν τα αντικείμενα, καθώς αυτή καθορίζεται αυτόματα από την τιμή που θα ανατεθεί στο αντικείμενο. Η R έχει πέντε βασικές κλάσεις αντικειμένων:

- χαρακτήρας (character)
- αριθμητικός- πραγματικοί αριθμοί (numeric)
- ακέραιος (integer)
- μιγαδικός (complex)
- λογικός (logical- True/False)

Οι βασικές δομές των αντικειμένων που χρησιμοποιούνται ευρέως και χρησιμοποιήθηκαν και στην παρούσα έρευνα είναι οι εξής:

- **vector**: Το διάνυσμα (vector) αποτελεί τη βασικότερη δομή. Ένα διάνυσμα μπορεί να περιέχει αντικείμενα του ίδιου μόνο τύπου. Η δημιουργία ενός διανύσματος είναι εφικτή είτε με τη συνάρτηση `c`, είτε με τη συνάρτηση `vector`.
- **list**: Μία λίστα, όπως και το διάνυσμα, είναι ένα σύνολο από αντικείμενα, που όμως μπορούν να ανήκουν σε διαφορετική κλάση. Για τη δημιουργία της λίστας χρησιμοποιείται η συνάρτηση `list`.
- **matrix**: Ένας μητρώο (matrix) είναι ουσιαστικά πολλά διανύσματα ενωμένα. Αποτελεί δηλαδή μία ειδική δομή, η οποία έχει ως επιπλέον ιδιότητα τη διάσταση (dimension). Ένα μητρώο αποτελεί στην ουσία ένα δισδιάστατο διάνυσμα, το οποίο είναι εξ' ορισμού ατομικό, τόσο δηλαδή οι γραμμές όσο και οι στήλες του θα πρέπει να περιέχουν αντικείμενα της ίδιας κλάσης. Ένας από τους τρόπους δημιουργίας ενός μητρώου είναι η δημιουργία ενός διανύσματος και στη συνέχεια ο καθορισμός των διαστάσεών του με τη συνάρτηση `dim`.

- **dataframe:** Το πλαίσιο δεδομένων (dataframe) χρησιμοποιείται για την αποθήκευση δεδομένων σε μορφή πίνακα. Έχει παρόμοια δομή με το μητρώο, καθώς είναι δισδιάστατο, αλλά μπορεί, όπως και η λίστα, να περιέχει διαφορετικού τύπου δεδομένα σε κάθε στήλη του. Για τη δημιουργία ενός πλαισίου δεδομένων χρησιμοποιείται η συνάρτηση `data.frame`.

Επιπλέον στην παρούσα έρευνα γίνεται χρήση των παρακάτω συναρτήσεων και εντολών:

- **as.vector:** Με το `as.vector` πραγματοποιείται η αναγωγή ενός αντικειμένου σε διάνυσμα.
- **cat και print:** Οι δύο αυτές συναρτήσεις εκτυπώνουν στην οθόνη την τιμή του ορίσματος που τους έχουμε δώσει.
- **data.frame:** Η εντολή `data.frame` ορίζει ένα πλαίσιο δεδομένων παίρνοντας αριθμητικά διανύσματα.
- **dim:** Με το `dim` αποδίδουμε διάσταση σε ένα διάνυσμα, προκειμένου να μετατραπεί σε πίνακα.
- **grep:** το `grep` αναζητά γραμμές κειμένου που ταιριάζουν σε μία ή περισσότερες κανονικές εκφράσεις και δίνει σαν έξοδο μόνο τις γραμμές που ταιριάζουν.
- **lapply:** Η `lapply` υπολογίζει το αποτέλεσμα μίας συνάρτησης πάνω στο κάθε στοιχείο μίας λίστας.
- **matrix:** Με τη συνάρτηση `matrix` μετατρέπεται ένα διάνυσμα σε πίνακα με άμεσο τρόπο.
- **predict:** Για να πάρουμε διαστήματα εμπιστοσύνης μέσης και ατομικής πρόβλεψης χρησιμοποιούμε τη συνάρτηση `predict`.
- **read.csv:** Το `read.csv` αποτελεί συνάρτηση για εισαγωγή δεδομένων στην R και συγκεκριμένα για αρχεία μορφής `.csv`, στα οποία τα δεδομένα διαχωρίζονται με κόμμα.
- **sapply:** Η συνάρτηση `sapply` λειτουργεί όπως η `lapply`, αλλά επιχειρεί να απλοποιήσει το αποτέλεσμα που επιστρέφεται, με αποτέλεσμα η μόνη ουσιαστική διαφορά τους να βρίσκεται στην επιστρεφόμενη τιμή.
- **str.match:** Το `str.match` αντιστοιχεί μία συμβολοσειρά (string) με μία κανονική έκφραση (regular expression) και επιστρέφει μία σειρά που περιλαμβάνει τα αποτελέσματα της αναζήτησης.

- **sub:** Το sub επιλέγει τμήμα ενός vector, matrix, ή dataframe.
 - **sum:** Το sum υπολογίζει το άθροισμα των τιμών.
 - **tm_map:** Το tm_map είναι μία διεπαφή, προκειμένου να πραγματοποιηθούν αλλαγές στα σώματα κειμένων.
 - **unlist:** Η συνάρτηση unlist μετατρέπει τη λίστα σε διάνυσμα.
 - **VCorpus:** Με το VCorpus δημιουργείται ένα σώμα κειμένων το οποίο είναι ασταθές (volatile), δηλαδή κρατάται πλήρως στη μνήμη και όλες οι αλλαγές επηρεάζουν μόνο το εν λόγω αντικείμενο στην R. Συνήθως χρησιμοποιείται, προκειμένου να γίνει εφικτός ο καθαρισμός των δεδομένων στο πλαίσιο του text mining.
 - **write.table:** Το write.table αποτελεί συνάρτηση για αποθήκευση δεδομένων στην R και συγκεκριμένα πινάκων δύο διαστάσεων.
- **EXCEL**

Στο Excel πραγματοποιήθηκε χειρωνακτικά η κατηγοριοποίηση των tweets ως προς την πολικότητά τους, αλλά και η επισημείωση των opinion targets. Επιπλέον, στο Excel δημιουργήθηκαν και τα train και test- sets για την εφαρμογή του αλγορίθμου Naive Bayes, τα οποία στη συνέχεια εισήχθησαν στο RStudio.

- **ORIGIN**

Το πρόγραμμα Origin και συγκεκριμένα η έκδοση OriginPro 7.5 χρησιμοποιήθηκε για την οπτικοποίηση των αποτελεσμάτων σε γραφήματα. Το Origin αποτελεί ένα πακέτο γραφικής αναπαράστασης και ανάλυσης δεδομένων, το οποίο παρέχει σε γραφικό περιβάλλον, δυνατότητες απευθείας εισαγωγής δεδομένων ή εισαγωγή από εξωτερικά αρχεία δεδομένων, δημιουργίας επαγγελματικών γραφικών και εξαγωγής τους σε πλήθος διαφορετικών μορφοποιήσεων.

ΥΛΟΠΟΙΗΣΗ ΕΡΕΥΝΑΣ

1. ΣΥΛΛΟΓΗ ΔΕΔΟΜΕΝΩΝ

1.1 ΔΗΜΙΟΥΡΓΙΑ ΜΙΑΣ TWITTER- ΕΦΑΡΜΟΓΗΣ

Το Twitter, προκειμένου να καταστήσει πιο εύκολη την πρόσβαση και την εκμετάλλευση του τεράστιου όγκου πληροφορίας που διαθέτει, προσφέρει τη δυνατότητα δημιουργίας εφαρμογών τρίτων- άλλων χρηστών (third-party applications). Συγκεκριμένα, μέσα από την πλατφόρμα του Twitter για προγραμματιστές, μπορεί κάποιος, μέσω μίας ασφαλούς και σύντομης διαδικασίας, να πιστοποιήσει μια τέτοια εφαρμογή, δίνοντας δικαιώματα πρόσβασης σε Διεπαφές Προγραμματισμού Εφαρμογών (APIs)²⁶⁹ από τον υπάρχοντα λογαριασμό του στο Twitter.

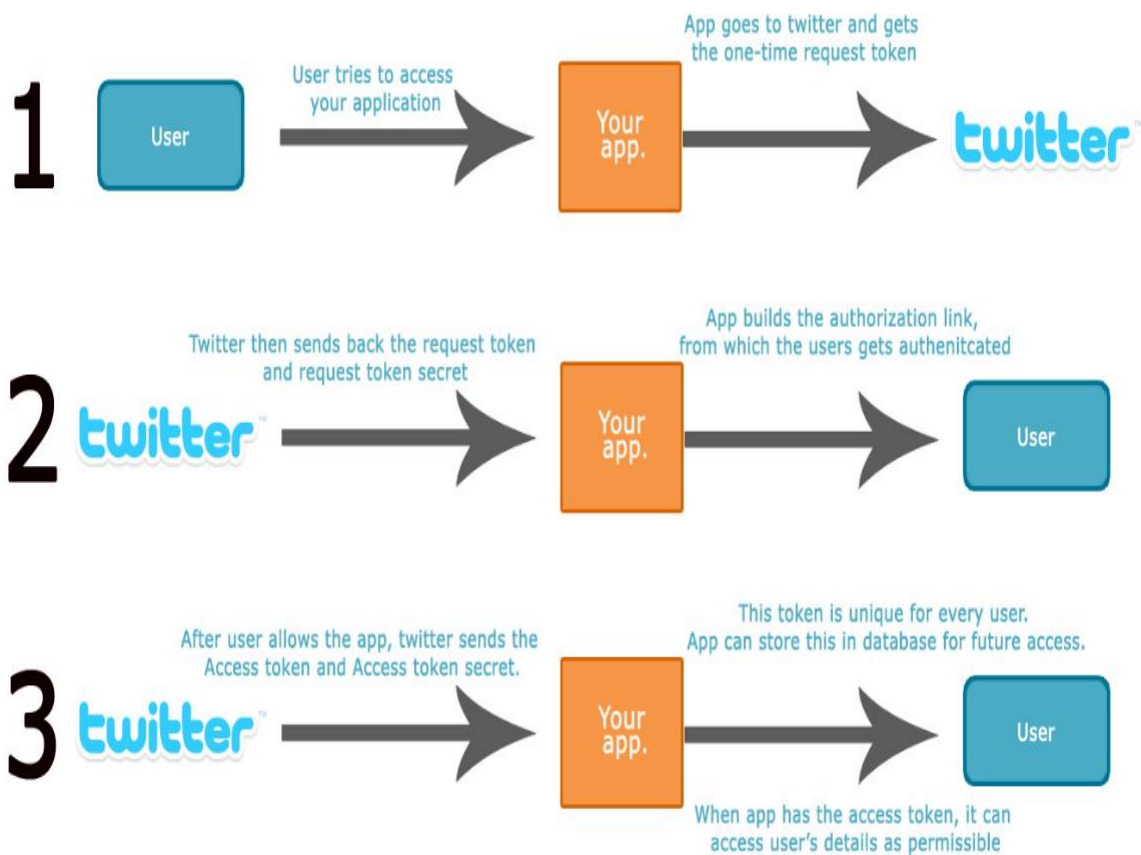
Στο πλαίσιο αυτής της μελέτης δημιουργήθηκε μία τέτοια εφαρμογή, προκειμένου να μπορεί να υπάρξει πρόσβαση στα επιθυμητά δεδομένα, μέσω κλήσεων σε συγκεκριμένα web-services του Twitter.

1.2 ΠΙΣΤΟΠΟΙΗΣΗ ΜΕ ΤΗ ΧΡΗΣΗ ΤΟΥ TWITTER OAuth

Στη συνέχεια είναι απαραίτητη η πιστοποίηση της εφαρμογής μέσω των αντίστοιχων consumer-key και consumer-secret που χορηγήθηκαν κατά την δημιουργία της εφαρμογής και στη συνέχεια με τη χρήση μιας έτοιμης βιβλιοθήκης (OAuth) να πραγματοποιηθεί μια χειραψία (handshake) με το Twitter, ώστε να γίνει εφικτή η πρόσβαση στα ελεύθερα δεδομένα. Συγκεκριμένα, για την πιστοποίηση των

²⁶⁹ Η Διεπαφή Προγραμματισμού Εφαρμογών είναι η διεπαφή των προγραμματιστικών διαδικασιών που παρέχει ένα λειτουργικό σύστημα, μία βιβλιοθήκη ή μία εφαρμογή, προκειμένου να μπορούν να γίνονται προς αυτά αιτήσεις από άλλα προγράμματα και ανταλλαγή δεδομένων.

εφαρμογών, για να μπορέσει δηλαδή κανείς να συνδεθεί, το Twitter χρησιμοποιεί την έτοιμη βιβλιοθήκη OAuth, που παρέχει ένα τύπο πιστοποίησης εφαρμογής πραγματοποιώντας αυτόνομα api calls. Αρχικά, η εφαρμογή κωδικοποιεί και στέλνει στο Twitter το consumer key & consumer secret που έχει παραχθεί κατά τη δημιουργία της ίδιας της εφαρμογής και στη συνέχεια στέλνει αυτή την πληροφορία στο OAuth με σκοπό να πάρει τον κωδικό access token,²⁷⁰ ώστε να μπορεί να αποκτήσει κανείς πρόσβαση στα ελεύθερα δεδομένα του Twitter.



Εικόνα 29. Τεκμηρίωση του πρωτοκόλλου OAuth στο Twitter²⁷¹

²⁷⁰ Βλ. <https://dev.twitter.com/oauth>

²⁷¹ Βλ. <http://www.phpbuilder.com/columns/oauth-figure.jpg>

1.3 ΔΙΕΠΑΦΕΣ ΤΟΥ TWITTER- REST vs Streaming API

Προκειμένου να αναζητήσει κανείς τα επιθυμητά tweets και να δημιουργήσει τα αναγκαία σύνολα δεδομένων προς ανάλυση, παρέχονται από το Twitter δύο ξεχωριστές διεπαφές (APIs), η REST (ή αλλιώς search) API και η Streaming API, τις οποίες μπορεί να καλέσει δημιουργώντας έναν απλό web service client και καθορίζοντας συγκεκριμένα χαρακτηριστικά στο αίτημα που θα στείλει.

Οι δύο αυτές διεπαφές έχουν συγκεκριμένες ιδιότητες και αποσκοπούν στην εξυπηρέτηση συγκεκριμένων αναγκών ανά περίπτωση. Τα κοινά τους χαρακτηριστικά είναι ότι και οι δύο απαιτούν πιστοποίηση μέσω του OAuth, παρέχουν τη δυνατότητα κάποιος να εισάγει κάποιες λέξεις- κλειδιά, ώστε να περιορίσει την αναζήτησή του, και ακόμα να πραγματοποιήσει αναζήτηση για συγκεκριμένο χρήστη ή λογαριασμό, χρονική περίοδο, ακόμα και για κάποια συγκεκριμένη γεωγραφική περιοχή.

Τα δεδομένα που επιστρέφουν περιέχουν όλη την πληροφορία που σχετίζεται με τα συγκεκριμένα tweets, όπως ποιος τα δημοσίευσε, από ποιο μέρος, πότε, αν έγιναν retweeted κ.α. Στην παρούσα έρευνα, από όλα τα παραπάνω στοιχεία, θα μελετηθούν το στοιχείο «text», που αποτελεί στην ουσία το κείμενο του εκάστοτε tweet και περιέχει τις λέξεις- κλειδιά με βάση τα οποία έγινε η αναζήτηση, καθώς και ο χρήστης που ανέβασε το κάθε tweet.

Η βασικότερη διαφορά μεταξύ των δύο διεπαφών έγκειται στο ότι η REST δίνει τη δυνατότητα πρόσβασης σε μερικά από τα αρχέτυπα του πυρήνα του Twitter, όπως είναι οι πληροφορίες για έναν χρήστη, τα tweets που δημοσιεύει κλπ. Όλα αυτά μπορούν να υλοποιηθούν με λίγες ξεκούραστες κλήσεις στο συγκεκριμένο API. Αυτός ήταν και ο λόγος που επιλέχθηκε για την παρούσα μελέτη. Αντιθέτως, η Streaming API, δημιουργεί μια σταθερή σύνδεση με το twitter και ανοίγει μια ροή που φέρνει διαρκώς δεδομένα για όσο μένει ανοιχτή ή μέχρις ότου ικανοποιηθεί ένα επιθυμητό μέγιστο που

μπορεί να οριστεί για την αναζήτηση των tweets.²⁷² Για το λόγο αυτό θεωρείται κατάλληλη για προγραμματιστές που επιθυμούν τη δημιουργία μίας μακρόβιας σύνδεσης και τη διατήρησή της.

Σημαντικό είναι να αναφερθεί ότι είναι δυνατή η πρόσβαση σε δεδομένα της τελευταίας εβδομάδας και όχι σε ιστορικά δεδομένα του Twitter, τα οποία είναι μόνο διαθέσιμα από εταιρείες επί πληρωμή. Στην παρούσα μελέτη χρησιμοποιήθηκε η διεπαφή REST API, μέσω της οποίας πραγματοποιήθηκαν κλήσεις, μία φορά κάθε εβδομάδα, το διάστημα 1/7/2016 έως 31/3/2017.

1.4 ΛΗΨΗ ΔΕΔΟΜΕΝΩΝ

Με τη γλώσσα προγραμματισμού R ζητήθηκε μέσω του REST API από το Twitter πρόσβαση σε αναρτήσεις χρηστών στα Ελληνικά και στα Γερμανικά που να προέρχονται από Ελλάδα και από Γερμανία αντίστοιχα και να σχετίζονται με την εκπαίδευση των προσφύγων. Για να γίνει αυτό εφικτό, πραγματοποιήθηκε αναζήτηση με βάση τις παρακάτω λέξεις- κλειδιά, οι οποίες είναι οι ίδιες και για τις δύο γλώσσες:

Ελληνικά keywords	Γερμανικά keywords
πρόσφυγες+σχολείο	Flüchtlinge+Schule
πρόσφυγες+σχολεία	Flüchtlinge+Schulen
προσφυγόπουλα+σχολείο	Flüchtlingskinder+Schule
προσφυγόπουλα+σχολεία	Flüchtlingskinder+Schulen
μετανάστες+σχολείο	Immigranten+Schule
μετανάστες+σχολεία	Immigranten+Schulen

²⁷² Βλ. <https://dev.twitter.com/streaming/overview>

Σύριοι+παιδιά+σχολείο	Syrer+Kinder+Schule
Σύριοι+παιδιά+σχολεία	Syrer+Kinder+Schulen
εκπαίδευση+παιδιά+μεταναστών	Bildung+Kinder+Immigranten
εκπαίδευση+παιδιά+προσφύγων	Bildung+Kinder+Flüchtlinge
άσυλο+παιδιά+σχολεία	Asyl+Kinder+Schule

Πίνακας 3. Ελληνικά και γερμανικά keywords

Όπως αναφέρθηκε παραπάνω, πραγματοποιήθηκαν κλήσεις, μία φορά κάθε εβδομάδα, το διάστημα 1/7/2016 έως 31/3/2017 και τα δεδομένα που επιστρέφονταν, αποθηκεύονταν στο RStudio.

```

23 search_all_keyw <- function(keywords)
24 {
25
26
27 lapply(keywords[sapply(keywords, length) > 0], function(kw){
28
29   rfg <- try( searchTwitter(kw, n=1500, locale='el_GR.UTF-8',
30                           lang='el'))
31   if (length(rfg) != 0) {
32     rfgdf = twListToDF(rfg)
33     fname <- paste0(paste0('twit_search_', kw, '_'), strftime(Sys.time(), format
34                                     = "%j_%H%M"), '.csv')
35     write.table(rfgdf, file = fname, row.names = F, sep=',', fileEncoding='UTF-8')
36   } else {'No tweets found!'}
37
38 })
39 }
40
41
42 test1 <- search_all_keyw(keywords = keywords)

```

Εικόνα 30. Κώδικας για τη λήψη tweets μέσω του REST API στο περιβάλλον του RStudio

2. ΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ

2.1 ΚΑΤΗΓΟΡΙΟΠΟΙΗΣΗ ΣΥΝΑΙΣΘΗΜΑΤΟΣ ΚΑΙ ΕΞΑΓΩΓΗ ΤΩΝ OPINION TARGETS

Τα αρχεία που συλλέχθηκαν με την παραπάνω μέθοδο εξήχθησαν από το RStudio και συγχωνεύθηκαν σε δύο υπολογιστικά φύλλα του Excel, εκ των οποίων το ένα περιλαμβάνει τα ελληνικά και το άλλο τα γερμανικά tweets μαζί με την ημερομηνία ανάρτησης και το όνομα του λογαριασμού που ανήρτησε το καθένα από αυτά. Συνολικά συγκεντρώθηκαν 22.454 tweets, εκ των οποίων τα 12.212 προέρχονται από την Ελλάδα, ενώ τα 10.242 από τη Γερμανία. Ανάμεσα στα παραπάνω tweets συμπεριλαμβάνονται και retweets, αναδημοσιεύσεις δηλαδή των πρωτότυπων tweets, τα οποία δεν αφαιρέθηκαν, καθώς, όταν ένας χρήστης αναπαράγει ένα tweet, θεωρείται ότι και ο ίδιος αποδέχεται το περιεχόμενο του tweet και θέλει να εκφράσει το ίδιο συναίσθημα²⁷³.

Στη συνέχεια τα tweets επισημειώθηκαν χειρωνακτικά και κατηγοριοποιήθηκαν ως προς την πολικότητά τους, ως προς το αν εκφράζουν δηλαδή θετικό, αρνητικό ή ουδέτερο συναίσθημα. Παράλληλα έγινε επισημείωση των opinion targets σε κάθε tweet και συγκεκριμένα των λέξεων- κλειδιών με τις οποίες αναφέρεται στα tweets ο εκάστοτε στόχος του συναισθήματος, προκειμένου να μπορεί κατόπιν να μετρηθεί το συναίσθημα για τον κάθε ένα στόχο του συναισθήματος ξεχωριστά. Για παράδειγμα, αναφορά στην οντότητα της κυβέρνησης γίνεται με τις λέξεις «κυβέρνηση», «Συριζαίοι», «κυβερνώντες» κλπ. Στη συνέχεια τα keywords ομαδοποιήθηκαν και αρχικά προέκυψαν οι εξής ομάδες, τόσο για τα ελληνικά όσο και για τα γερμανικά tweets:

- Πρόσφυγες
- Γονείς μαθητών
- Κυβέρνηση

²⁷³ Bosch, R., Sentiment Analysis: Incremental learning to build domain models, Master thesis, Universitat Pompeu Fabra, 2013, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: [https://www.upf.edu/iis/pdf/Theses/Incremental learning to build domain models-Master Thesis UPF-Raimon Bosch.pdf](https://www.upf.edu/iis/pdf/Theses/Incremental%20learning%20to%20build%20domain%20models-Master%20Thesis%20UPF-Raimon%20Bosch.pdf), [προσπελάστηκε στις 12 Ιουλίου 2017], σελ. 10.

- Ευρώπη- Ο.Η.Ε
- Μ.Μ.Ε
- Ακροδεξιά
- Κοινωνία
- Σχολείο
- ΜΚΟ
- Τοπική Αυτοδιοίκηση
- Κόμματα- Πολιτικοί

Στη συνέχεια, για λόγους οικονομίας, οι παραπάνω κατηγορίες ομαδοποιήθηκαν περαιτέρω και έτσι προέκυψαν τέσσερις κατηγορίες. Πρώτη κατηγορία είναι αυτή που το συναίσθημα των tweets αφορά στην ακροδεξιά. Η δεύτερη κατηγορία έχει να κάνει με την πολιτική και περιλαμβάνει τις κατηγορίες της κυβέρνησης, των κομμάτων-πολιτικών, της τοπικής αυτοδιοίκησης, της Ευρώπης- Ηνωμένων Εθνών, τα Μ.Μ.Ε και τις ΜΚΟ. Στην τρίτη κατηγορία στόχοι του συναισθήματος είναι τα μέλη της κοινωνίας και γι'αυτό περιλαμβάνονται οι κατηγορίες των γονέων των μαθητών, του σχολείου και της ευρύτερης κοινωνίας. Τέλος, η τέταρτη κατηγορία αναφέρεται αποκλειστικά στους πρόσφυγες και στα παιδιά τους.

2.2 ΔΙΑΧΩΡΙΣΜΟΣ ΤΩΝ TWEETS ΣΕ MEDIA ΚΑΙ USERS ΣΤΟ RSTUDIO

Στη συνέχεια έγινε η εισαγωγή των δύο αρχείων (του ενός με τα ελληνικά και του άλλου με τα γερμανικά tweets) στο περιβάλλον του RStudio σε μορφή .csv. Το κάθε αρχείο περιελάμβανε τέσσερις στήλες: την ημερομηνία ανάρτησης του tweet (date), το κείμενο του tweet (text), τον χρήστη που το ανήρτησε (user) και το συναίσθημα που εκφράζει το κάθε tweet, όπως αυτό προέκυψε από τη διαδικασία της χειρωνακτικής κατηγοριοποίησης. Με τις κατάλληλες εντολές της γλώσσας R έγινε ο διαχωρισμός των tweets σε αυτά που έχουν αναρτήσει απλοί χρήστες και σε αυτά που έχουν αναρτηθεί από λογαριασμούς των μέσων μαζικής επικοινωνίας, με την εξής διαδικασία: Εισήχθησαν στο RStudio δύο αρχεία τύπου .txt, ένα που περιελάμβανε ονόματα Μ.Μ.Ε

της Ελλάδας και ένα δεύτερο με ονόματα γερμανικών Μ.Μ.Ε, όπως αυτά εμφανίζονται στο Twitter. Οι δύο αυτές λίστες αποτέλεσαν το φίλτρο, έτσι ώστε να διαχωριστούν τα δύο μεγάλα αρχεία και να προκύψουν τέσσερα μικρότερα, με τα tweets από τους Έλληνες χρήστες, από τα ελληνικά media, από τους Γερμανούς χρήστες και από τα γερμανικά media, το καθένα από τα οποία είχε την εξής έκταση:

- **Greek media:** 2.448 tweets
- **Greek users:** 9.764 tweets
- **German media:** 1.743 tweets
- **German users:** 8.499 tweets

Στόχος αυτού του διαχωρισμού είναι να μελετηθεί ξεχωριστά το συναίσθημα που εκφράζεται από τους πολίτες και ξεχωριστά αυτό των Μ.Μ.Ε της κάθε χώρας.

2.3 ΔΙΑΧΩΡΙΣΜΟΣ ΤΩΝ TWEETS ΜΕ ΒΑΣΗ ΤΟΥΣ ΣΤΟΧΟΥΣ ΣΥΝΑΙΣΘΗΜΑΤΟΣ ΣΤΟ RSTUDIO

Κατόπιν έγινε εισαγωγή στο RStudio οκτώ αρχείων μορφής .txt, τεσσάρων για κάθε γλώσσα, όπου το καθένα περιλαμβάνει μία λίστα με τις λέξεις- κλειδιά, οι οποίες αναφέρονται στις τέσσερις κατηγορίες των στόχων του συναισθήματος, όπως αυτές καθορίστηκαν με τη χειρωνακτική επισημείωση. Με τις κατάλληλες εντολές της γλώσσας R έγινε περαιτέρω διαχωρισμός των tweets, με αποτέλεσμα να γίνει γνωστό ποια tweets των ελληνικών και γερμανικών media, αλλά και των users, Ελλήνων και Γερμανών, αναφέρονται σε κάθε μία κατηγορία των στόχων του συναισθήματος, δηλαδή στην ακροδεξιά, στην πολιτική, στην κοινωνία και στους πρόσφυγες. Στόχος αυτού του διαχωρισμού είναι να μελετηθεί ξεχωριστά το συναίσθημα που εκφράζουν οι Έλληνες και Γερμανοί χρήστες, καθώς και τα ελληνικά και γερμανικά μέσα μαζικής επικοινωνίας για την ακροδεξιά, την πολιτική, την κοινωνία και τους ίδιους τους πρόσφυγες σε σχέση πάντα με το θέμα της εκπαίδευσης των προσφύγων, αλλά και το σε ποιο βαθμό αναφέρονται στην κάθε μία από τις τέσσερις κατηγορίες.

3. ΑΝΑΛΥΣΗ ΔΕΔΟΜΕΝΩΝ

3.1 ΥΠΟΛΟΓΙΣΜΟΣ ΤΟΥ ΣΥΝΑΙΣΘΗΜΑΤΟΣ ΚΑΙ ΕΞΑΓΩΓΗ ΤΩΝ ΣΤΟΧΩΝ ΤΟΥ ΜΕ ΒΑΣΗ ΤΗ ΧΕΙΡΩΝΑΚΤΙΚΗ ΕΠΙΣΗΜΕΙΩΣΗ

Στη συνέχεια έγιναν οι απαραίτητες μετρήσεις, προκειμένου να γίνει φανερό για τον κάθε opinion holder (greek media, greek users, german media και german users) το πλήθος των tweets που έχει αναρτήσει για την κάθε μία από τις τέσσερις κατηγορίες των opinion targets (ακροδεξιά, πολιτική, κοινωνία και πρόσφυγες), διαπιστώνοντας έτσι σε ποια έκταση αναφέρονται οι χρήστες του Twitter στις τέσσερις κατηγορίες των στόχων του συναισθήματος. Μετέπειτα, ανιχνεύτηκε και το συναίσθημα που εκφράζουν οι opinion holders και για όλους τους στόχους του συναισθήματος μαζί, αλλά και για κάθε έναν ξεχωριστά, με βάση τα δεδομένα που προέκυψαν από τη χειρωνακτική κατηγοριοποίηση. Επειδή όμως στόχος της έρευνας είναι να διαπιστωθεί το κατά πόσο είναι δυνατόν να γίνει η εξαγωγή του συναισθήματος με αυτοματοποιημένες μεθόδους, χωρίς να χρειάζεται η ανθρώπινη κατηγοριοποίηση, έγινε προσπάθεια ανάλυσης του συναισθήματος με τη μέθοδο του λεξικού, αλλά και του αλγορίθμου της μηχανικής μάθησης Naive Bayes, προκειμένου να ελεγχθεί η αποδοτικότητα των δύο μεθόδων, το κατά πόσο δηλαδή τα αποτελέσματά τους συμφωνούν με αυτά της ανθρώπινης κατηγοριοποίησης.

3.2 ΜΕΘΟΔΟΣ ΚΑΘΑΡΙΣΜΟΥ ΔΕΔΟΜΕΝΩΝ

Προκειμένου να εφαρμοστεί κανείς τεχνικές ανάλυσης συναισθήματος, θα πρέπει πρώτα να απομακρύνει από το κείμενο προς ανάλυση τον λεγόμενο «θόρυβο», έτσι ώστε το κείμενο να είναι καθαρό και να περιλαμβάνει μόνο τις λέξεις που χρειάζονται για να γίνει η ανάλυση, χωρίς αριθμούς, σημεία στίξης, ειδικούς χαρακτήρες, κτλ. Η εν λόγω διαδικασία είναι απαραίτητη τόσο για τα δεδομένα που χρησιμοποιούνται για την εκπαίδευση του αλγορίθμου Naive Bayes που κάνει την κατηγοριοποίηση, όσο και για δεδομένα που θα χρησιμοποιηθούν για την πρόβλεψη του συναισθήματος με βάση το λεξικό.

Η διαδικασία καθαρισμού των δεδομένων που ακολουθήθηκε στην παρούσα έρευνα και υλοποιήθηκε με τη χρήση του πακέτου της R για το text mining (tm) είναι η εξής:

1. **Μετατροπή του συνόλου του κειμένου σε πεζά γράμματα (lower text)**, ώστε να υπάρχει μια ενιαία και κοινή δομή. Η χρήση κεφαλαίων γραμμάτων, ειδικά για λόγους έμφασης, είναι πολύ συνηθισμένο φαινόμενο στο Twitter. Δεδομένου όμως ότι η έμφαση δεν επηρεάζει την πολικότητα των λέξεων, παρά μόνο ενισχύει το συναίσθημα, και ότι η αντιστοίχιση των λέξεων των tweets με αυτές του λεξικού απαιτεί ομοιογένεια, δημιουργείται η ανάγκη μετατροπής όλων των κεφαλαίων χαρακτήρων σε πεζούς.
2. **Αφαίρεση των αριθμητικών συμβόλων (numbers).**
3. **Αφαίρεση των υπερσυνδέσμων (links / urls)** που μπορεί να περιέχονται στα tweets. Οι υπερσύνδεσμοι σε ένα tweet συνήθως οδηγούν σε ένα νέο ιστότοπο που περιέχει ακατέργαστο κείμενο και ενδεχομένως και άλλου είδους πολυμεσικό περιεχόμενο όπως εικόνες και βίντεο. Εφόσον δεν μελετάται στην παρούσα έρευνα ένα σύστημα που ανιχνεύει το ευρύτερο σημασιολογικό πλαίσιο των tweets ερμηνεύοντας πολυμεσικό περιεχόμενο, οι υπερσύνδεσμοι δεν προσφέρουν καμία αξιοποιήσιμη πληροφορία και συνεπώς κρίνεται σκόπιμη η αφαίρεση κάθε συμβολοσειράς που ξεκινάει με “www.”, “http://” ή https:// κλπ.
4. **Αφαίρεση των mentions που υπάρχουν μέσα στο κείμενο (@username)**, καθώς στη συγκεκριμένη ανάλυση το όνομα του χρήστη που αναφέρεται μέσα σε ένα tweet έχει μηδενική επιρροή στο συναίσθημα του.
5. **Αφαίρεση της λέξης «RT»** που συναντάται μπροστά από τα retweets.
6. **Αφαίρεση όλων των σημείων στίξης και όλων των μορφών punctuation** που μπορεί να υπάρχουν στο κείμενο. Τα σημεία στίξης μπορούν να συμβάλλουν στην ενίσχυση του συναισθήματος, αλλά δεν επιδρούν, συνήθως, στην πολικότητα της λέξης, επομένως μπορούν να αφαιρεθούν. Σε ορισμένες περιπτώσεις επιδρούν και στο συναίσθημα όλης της πρότασης (π.χ. «Μα είσαι σοβαρός;;;;;;»), αλλά ο χειρισμός τέτοιων περιπτώσεων απαιτεί ενδελεχή γλωσσολογική ανάλυση. Έτσι, θα αφαιρεθούν και όλα τα emoticons που υπάρχουν στα tweets, τα οποία μπορεί να

εκφράζουν συναίσθημα, ωστόσο στόχος της παρούσας έρευνας είναι η πρόβλεψη που βασίζεται καθαρά στις λέξεις του κειμένου.

7. Αφαίρεση των stopwords, λέξεων δηλαδή, οι οποίες είναι οι πιο συνηθισμένες-κοινές λέξεις μιας γλώσσας, δεν έχουν κανένα συναισθηματικό υπόβαθρο και για αυτό το λόγο αφαιρούνται. Τέτοιες λέξεις είναι συνήθως τα άρθρα, οι προθέσεις, οι σύνδεσμοι και οι αντωνυμίες, οι οποίες είναι πολύ συνηθισμένες και δεν αποτελούν σε καμία περίπτωση φορείς συναισθήματος. Προκειμένου να γίνει η αφαίρεση των stopwords, εισήχθησαν στην R δύο αρχεία τύπου .txt, εκ των οποίων το ένα περιλαμβάνει ελληνικά και το άλλο γερμανικά stopwords²⁷⁴. Με τις κατάλληλες εντολές απομακρύνθηκαν από το σύνολο των tweets όσες λέξεις βρίσκονται στις δύο λίστες.

8. Αφαίρεση περιττών κενών (spaces), ώστε τα διανύσματα κειμένου που προκύπτουν από τα επεξεργασμένα tweets να είναι απολύτως ωμά (raw text).

Η παραπάνω διαδικασία εφαρμόστηκε, τόσο στα ελληνικά, όσο και στα γερμανικά tweets. Όσον αφορά στα Ελληνικά, εφαρμόστηκαν επιπλέον δύο εντολές, αυτή της αφαίρεσης των τόνων των λέξεων και αυτή της αφαίρεσης όλων των λατινικών χαρακτήρων, κάτι που ευλόγως δεν έγινε για τα γερμανικά tweets. Το φαινόμενο της εμφάνισης ξενόγλωσσων λέξεων, χαρακτήρων ή συντομογραφιών, αλλά και η χρήση greeklish, είναι εξαιρετικά συνηθισμένο στα ελληνικά tweets.

²⁷⁴ Τα ελληνικά stopwords διαθέσιμα στη διεύθυνση: <https://github.com/xtsimpouris/gr-nlp-law/blob/master/Greek%20Stopwords/stopwords.txt>, [προσπελάστηκε στις 10 Αυγούστου 2017] και τα γερμανικά stopwords διαθέσιμα στη διεύθυνση: <https://github.com/stopwords-iso/stopwords-de/blob/master/stopwords-de.txt>, [προσπελάστηκε στις 10 Αυγούστου 2017].

Ακολουθεί ένα παράδειγμα στο οποίο φαίνονται αναλυτικά όλα τα στάδια του καθαρισμού των tweets:

Βήμα	Tweet
1. Αρχικό tweet	RT @VMourtz85: Μη σας παραξενεύει που οι γονείς από το Ωραιόκαστρο ζήτησαν να μην πάνε τα προσφυγόπουλα σχολείο. Ούτε αυτοί έχουν πάει. https://vice.com/gr/read/katali\u0085pic.twitter.com/tlRUfZALza
2. Μετατροπή κεφαλαίων σε πεζά	rt @vmourtz85: μη σας παραξενεύει που οι γονείς από το Ωραιόκαστρο ζήτησαν να μην πάνε τα προσφυγόπουλα σχολείο. ούτε αυτοί έχουν πάει. https://vice.com/gr/read/katali\u0085pic.twitter.com/tlRUfZALza
3. Αφαίρεση αριθμών	rt @vmourtz: μη σας παραξενεύει που οι γονείς από το Ωραιόκαστρο ζήτησαν να μην πάνε τα προσφυγόπουλα σχολείο. ούτε αυτοί έχουν πάει. https://vice.com/gr/read/katali\u0085pic.twitter.com/tlRUfZALza
4. Αφαίρεση υπερσυνδέσμων	rt @vmourtz: μη σας παραξενεύει που οι γονείς από το Ωραιόκαστρο ζήτησαν να μην πάνε τα προσφυγόπουλα σχολείο. ούτε αυτοί έχουν πάει. pic.twitter.com/tlRUfZALza

5. Αφαίρεση αναφορών σε χρήστες	rt : μη σας παραξενεύει που οι γονείς από το Ωραιόκαστρο ζήτησαν να μην πάνε τα προσφυγόπουλα σχολείο. ούτε αυτοί έχουν πάει. pic.twitter.com/tlRUfZALza
6. Αφαίρεση του «RT»	: μη σας παραξενεύει που οι γονείς από το Ωραιόκαστρο ζήτησαν να μην πάνε τα προσφυγόπουλα σχολείο. ούτε αυτοί έχουν πάει. pic.twitter.com/tlRUfZALza
7. Αφαίρεση σημείων στίξης	μη σας παραξενεύει που οι γονείς από το Ωραιόκαστρο ζήτησαν να μην πάνε τα προσφυγόπουλα σχολείο ούτε αυτοί έχουν πάει pictwittercomtlRUfZALza
8. Αφαίρεση των stopwords	παραξενεύει γονείς Ωραιόκαστρο ζήτησαν πάνε προσφυγόπουλα σχολείο έχουν πάει pictwittercomtlRUfZALza
9. Αφαίρεση περιττών κενών	παραξενεύει γονείς Ωραιόκαστρο ζήτησαν πάνε προσφυγόπουλα σχολείο έχουν πάει pictwittercomtlRUfZALza
10. Αφαίρεση λατινικών χαρακτήρων	παραξενεύει γονείς Ωραιόκαστρο ζήτησαν πάνε προσφυγόπουλα σχολείο έχουν πάει
11. Αφαίρεση τονισμού	παραξενευει γονεις Ωραιόκαστρο ζητησαν πανε προσφυγοπουλα σχολειο εχουν παει

Πίνακας 4. Διαδικασία καθαρισμού δεδομένων

3.3 ΑΝΑΛΥΣΗ ΣΥΝΑΙΣΘΗΜΑΤΟΣ ΜΕ ΤΗ ΜΕΘΟΔΟ ΤΟΥ ΛΕΞΙΚΟΥ

3.3.1 ΔΗΜΙΟΥΡΓΙΑ ΛΕΞΙΚΩΝ

Ως λεξικό, όπως αναφέρθηκε και παραπάνω, ορίζουμε στο πλαίσιο της ανάλυσης συναισθήματος, τις λίστες λέξεων οι οποίες περιλαμβάνουν λέξεις που εκφράζουν θετικό ή αρνητικό συναίσθημα και μπορούν να χρησιμοποιηθούν για να προσδώσουν μία βαθμολογία σε μία ολόκληρη πρόταση. Έπειτα, αυτή η βαθμολογία μπορεί να χαρακτηρίσει τον συναισθηματικό τόνο της πρότασης, αφού για παράδειγμα η χρήση αρνητικά βαθμολογημένων λέξεων γίνεται στην περίπτωση που η πρόταση στο σύνολό της εννοεί κάτι το αρνητικό. Αντίστοιχα, η χρήση θετικά βαθμολογημένων λέξεων γίνεται στην περίπτωση που η πρόταση στο σύνολό της εννοεί κάτι το θετικό.

- **Ελληνικό λεξικό**

Στην παρούσα μελέτη χρησιμοποιήθηκε για τα Ελληνικά ένα λεξικό συναισθήματος που δημιουργήθηκε από τον Αδάμ Τσακαλίδη (Ινστιτούτο Τεχνολογιών Πληροφορικής και Επικοινωνιών) σε συνεργασία με τον Συμεών Παπαδόπουλο (Ινστιτούτο Τεχνολογιών Πληροφορικής και Επικοινωνιών) και τη συνεισφορά των Ουρανία Βοσκάκη (Κέντρο Ελληνικής Γλώσσας), Κυριακή Ιωαννίδου (Κέντρο Ελληνικής Γλώσσας) και Χριστίνα Βοϊδίδου (Κέντρο Ελληνικής Γλώσσας).²⁷⁵ Το λεξικό αυτό δημιουργήθηκε με την υποστήριξη του προγράμματος SocialSensor 16 και βρίσκεται διαθέσιμο στο διαδίκτυο για κοινή χρήση στο κοινωνικό εναποθετήριο εργασιών ανοιχτού κώδικα GitHub²⁷⁶. Πέραν όμως από τις θετικές και αρνητικές λέξεις που περιλαμβάνει το παραπάνω λεξικό, προστέθηκαν και μερικές ακόμα που έχουν να κάνουν με το θέμα των προσφύγων.

²⁷⁵ Tsakalidis, A., Papadopoulos, S., & Kompatsiaris, I., An Ensemble Model for Cross-Domain Polarity Classification on Twitter, *Web Information Systems Engineering– WISE*, Springer International Publishing, 2014, σελ. 168-177.

²⁷⁶ Βλ. <https://github.com/MKLab-ITI/greek-sentiment-lexicon>

Συγκεκριμένα, προστέθηκαν οι παρακάτω λέξεις:

Θετικές λέξεις	Αρνητικές λέξεις
συνδρομή	ζωσμένος
αγκαλιά	ναζί
	εκρηκτικά
	φασίστας
	μαντρώστε
	φασιστικός
	bullying
	ταγαρόβλαχος
	μπολιτιζμός
	σκατάνθρωποι
	αντίχριστος
	κακομαθημένος
	ευαισθητούληδες
	lidl

Πίνακας 5. Θετικές και αρνητικές λέξεις που προστέθηκαν στο ελληνικό λεξικό

Στη συνέχεια αποκόπηκαν οι καταλήξεις των λέξεων του λεξικού (stemming) και διατηρήθηκε μόνο το θέμα της κάθε λέξης. Με το stemming των λέξεων του λεξικού, κάθε θέμα ταυτίζεται στη συνέχεια με πολλαπλές μορφές μιας λέξης, ενώ διαφορετικές λέξεις με ίδιο θέμα και άρα σχεδόν ταυτόσημη έννοια, λαμβάνονται υπόψιν ως μία, βελτιώνοντας έτσι σημαντικά την απόδοση της ανάλυσης. Το stemming έγινε με τη χρήση της stemmer για την ελληνική γλώσσα, που είναι διαθέσιμος στο διαδίκτυο.²⁷⁷ Στην τελική του μορφή το ελληνικό λεξικό που χρησιμοποιήθηκε στην παρούσα έρευνα αποτελείται από 1.119 θετικές και 2.198 αρνητικές λέξεις.

²⁷⁷ Βλ. <http://deixto.com/greek-stemmer/>

- **Γερμανικό λεξικό**

Όσον αφορά στο γερμανικό λεξικό χρησιμοποιήθηκε το SentiWS του πανεπιστημίου της Λειψίας.²⁷⁸ Στο συγκεκριμένο λεξικό δεν χρειάστηκε να γίνει η διαδικασία του stemming, καθώς χρησιμοποιήθηκε η έκδοσή του που περιλαμβάνει τις λέξεις σε όλες τους τις μορφές (π.χ. τα ουσιαστικά σε όλες τις πτώσεις, τα επίθετα με τις δυνατές καταλήξεις κ.α.). Παράλληλα, προστέθηκαν και εδώ κάποιες λέξεις, θετικές και αρνητικές, που έχουν να κάνουν με το θέμα της εκπαίδευσης των προσφύγων:

Θετικές λέξεις	Αρνητικές λέξεις
Kindheit	Talibanköpfe
pauken	prügeln
lieb	protzen
motiviert	Tuberkulose
engagieren	hassen
willkommen	Schlägereien
beherrschen	Ghettoisierung
cute	gemoppt
Engagement	geschlagen
Bleibrecht	skandalös
Ehrgeiz	kriminell
Glücksfall	Ünerforderung
büffeln	Kriegsgebiet
Kindheit	Propagandalüge
pauken	nutzlose
lieb	Apathie
motiviert	Privatvormundschaft
engagieren	vergammeln
willkommen	Schulmiserie

²⁷⁸ Βλ. <http://www.wortschatz.uni-leipzig.de/de/download>

beherrschen	Spaltung
cute	vergewaltigen
	Deserteur
	Dabakel
	Belagerung
	Kleinmut
	Feigheit
	verbarrikadieren
	Kanake
	#abmerkeln
	vollverschleiert
	Skepsis
	kotzen
	verreck
	Trottel
	Intoleranz
	scmuggeln
	geschmuggelt
	Depp
	Deppen
	Talibanköpfe

Πίνακας 6. Θετικές και αρνητικές λέξεις που προστέθηκαν στο γερμανικό λεξικό

Επίσης αφαιρέθηκαν από τις αρνητικές λέξεις όλες όσες είχαν να κάνουν με τη λέξη «πρόσφυγας», γιατί προφανώς κάτι τέτοιο θα αλλοίωνε τα αποτελέσματα. Συγκεκριμένα, ανάμεσα στις αρνητικές λέξεις βρέθηκαν και αφαιρέθηκαν οι εξής: Flucht , Fluchten, Flüchtiger, Flüchtigkeit, Flüchtigkeiten, Flüchtling, Flüchtlings, Flüchtlinge, Flüchtlinges, Flüchtlingen. Στην τελική του μορφή το γερμανικό λεξικό που χρησιμοποιήθηκε στην παρούσα έρευνα αποτελείται από 1.671 θετικές και 1.855

αρνητικές λέξεις, οι οποίες, προσθέτοντας όλες τις πιθανές τους μορφές, φτάνουν τις 15.669 θετικές και 15.661 αρνητικές.

3.3.2 ΕΚΤΕΛΕΣΗ ΜΕΘΟΔΟΥ

Βήμα 1^ο

Αρχικά εγκαταστάθηκε και φορτώθηκε στο περιβάλλον του RStudio το πακέτο «stringr», το οποίο δίνει τη δυνατότητα να εντοπίζει συγκεκριμένα στοιχεία σε μία συμβολοσειρά, στη συγκεκριμένη περίπτωση σε ένα tweet.

Βήμα 2^ο

Στη συνέχεια κλήθηκαν τα αρχεία με τις λίστες των θετικών και των αρνητικών λέξεων, τα tweets χωρίστηκαν στις επί μέρους λέξεις τους και ζητήθηκε για κάθε λέξη να γίνει έλεγχος για το αν υπάρχει στο λεξικό.

Βήμα 3^ο

Με τη συνάρτηση «str.match» μετρήθηκαν οι θετικές και οι αρνητικές λέξεις σε κάθε tweet και ανάλογα με το ποιες ήταν κάθε φορά οι περισσότερες, το tweet χαρακτηριζόταν ως θετικό ή αρνητικό. Σε περίπτωση μη εντοπισμού κάποιας θετικής ή αρνητικής λέξης ή στη σπάνια περίπτωση εντοπισμού ίδιου αριθμού θετικών και αρνητικών λέξεων σε ένα tweet, τότε αυτό χαρακτηριζόταν ως ουδέτερο.

Επομένως, αυτό που κάνει ουσιαστικά ο κατηγοριοποιητής του λεξικού είναι να υπολογίζει πόσο θετικό, πόσο αρνητικό και πόσο ουδέτερο είναι ένα tweet, ανάλογα με το ποιες λέξεις του λεξικού περιέχει. Το συναίσθημα που εκφράζεται σε κάθε tweet αντιστοιχεί στον μετρητή που συγκέντρωσε τη μεγαλύτερη τιμή, ενώ, όπως αναφέρθηκε παραπάνω, στην σπάνια περίπτωση που οι μετρητές ισοβαθούν, τότε υποθετικά, το συναίσθημα που εκφράζεται είναι ουδέτερο.

3.4 ΑΝΑΛΥΣΗ ΣΥΝΑΙΣΘΗΜΑΤΟΣ ΜΕ ΤΟΝ ΑΛΓΟΡΙΘΜΟ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ NAIVE BAYES

Βήμα 1^ο

Αρχικά εγκαταστάθηκε και φορτώθηκε στο περιβάλλον του RStudio το πακέτο «xlsx», το οποίο δίνει τη δυνατότητα διαχείρισης φύλλων του MS Excel στο προγραμματιστικό περιβάλλον της R, καθώς και το πακέτο «e1071», το οποίο υλοποιεί τη μέθοδο Naive Bayes.

Βήμα 2^ο

Στη συνέχεια, έγινε εισαγωγή των δεδομένων που περιέχει το φύλλο Excel και συγκεκριμένα των γερμανικών tweets που αντιστοιχούν σε συντάκτες μέσω μαζικής επικοινωνίας (media). Το συγκεκριμένο αρχείο, όπως προαναφέρθηκε, περιλαμβάνει τέσσερις στήλες: Η πρώτη αφορά στον χρόνο δημοσίευσής του, ενώ η δεύτερη στο κείμενο του tweet, το οποίο έχει ήδη καθαριστεί από ανεπιθύμητους χαρακτήρες που δυσχεραίνουν τη διαδικασία εφαρμογής της μεθόδου (π.χ. χαρακτήρες emoticon, τονισμός, κεφαλαία γράμματα κλπ.), σύμφωνα με την προεργασία που έχει περιγραφεί σε προηγούμενη ενότητα. Η τρίτη στήλη εκφράζει τον συναισθηματικό χαρακτηρισμό που εκφράζει το εκάστοτε tweet (sentiment) και αποτελεί και τη μεταβλητή στόχο της εν λόγω προσέγγισης. Τέλος, η τέταρτη στήλη αφορά στο όνομα του συντάκτη. Λόγω της γραφής των tweet στην ελληνική και γερμανική γλώσσα, κλήθηκε το σχήμα κωδικοποίησης χαρακτήρων UTF-8 ώστε να αναγνωρίζονται στο περιβάλλον της R.

Βήμα 3^ο

Επόμενο βήμα είναι η καταμέτρηση των λέξεων που περιλαμβάνει κάθε tweet ξεχωριστά. Η απαρίθμηση στηρίζεται πρώτα στη διακριτοποίηση του κάθε tweet με βάση την επιλογή ενός κατάλληλου οριοθέτη, που στη συγκεκριμένη περίπτωση έχει τεθεί το κενό διάστημα (" ") μέσω της εντολής «strsplit». Δημιουργείται λοιπόν ένας πίνακας-στήλη (num_words), του οποίου κάθε στοιχείο αποτελεί το πλήθος των λέξεων που καταμετρήθηκαν στο κάθε tweet μέσω της εντολής «length». Είναι προφανές ότι

το πλήθος γραμμών του πίνακα αυτού θα είναι ίσος με το πλήθος των tweets που έχουν εισαχθεί προς επεξεργασία (`num_tweets`).

Βήμα 4^ο

Το συνολικό πλήθος των λέξεων «`tot_num_words`» που βρέθηκαν σε όλα τα tweets μπορεί να υπολογιστεί ως το άθροισμα των στοιχείων του πίνακα «`num_words`» μέσω της εντολής «`sum`». Πολλές είναι οι λέξεις που επαναλαμβάνονται, καθώς όσες λέξεις εντοπίζονται, καταγράφονται για κάθε tweet ξεχωριστά, χωρίς να υπάρχει γνώση, εάν οι συγκεκριμένες έχουν ήδη βρεθεί σε κάποιο άλλο tweet.

Βήμα 5^ο

Αφού πλέον το πλήθος των καταμετρημένων λέξεων είναι γνωστό, κατασκευάζεται ο πίνακας-στήλη «`words`», του οποίου το κάθε στοιχείο αντιστοιχεί σε καθεμία λέξη, η οποία εντοπίστηκε από την προηγούμενη διαδικασία. Οι πρώτες θέσεις του πίνακα καταλαμβάνονται από τις λέξεις του πρώτου tweet και η ίδια πορεία ακολουθείται διαδοχικά μέχρι να πληρωθούν και οι «`tot_num_words`» συνολικά θέσεις. Η πορεία κατάληψης των θέσεων ελέγχεται μέσω του δείκτη «*i*» για τα tweets και για κάθε *i*-οστό tweet μετακινούνται οι περιεχόμενες σε αυτό λέξεις μέσω του δείκτη «*j*». Δηλαδή το «*j*» μετράει μεταξύ των `num_words[i]` του *i*-οστού tweet για να προχωρήσει στη συνέχεια στο αμέσως επόμενο *i*+1.

Βήμα 6^ο

Όπως αναφέρθηκε, ο πίνακας `words` περιλαμβάνει όλες πλέον τις λέξεις που βρέθηκαν από όλα τα tweets ανεξαιρέτως. Η R δίνει τη δυνατότητα να απομακρύνει κανείς τις επαναλαμβανόμενες λέξεις μέσω της εντολής «`unique`» και να σχηματιστεί έτσι ένας νέος πίνακας-στήλη, ο «`common_words`», ο οποίος θα περιλαμβάνει τις λέξεις που εντοπίστηκαν από μία μόνο φορά, χωρίς δηλαδή επαναλήψεις. Το πλήθος των λέξεων πλέον υποβιβάζεται από «`tot_num_words`» σε ένα σημαντικά μικρότερο «`num_com_words`» που μετράται και πάλι με την εντολή «`length`». Ο πίνακας αυτός αποτελεί και το σημείο αναφοράς της διαδικασίας, καθώς βάσει των στοιχείων που περιλαμβάνει θα εκτελούνται πλέον όλοι οι έλεγχοι για την κατασκευή του πίνακα εισόδου στην τεχνική του Naive Bayes (bag of words).

Βήμα 7^ο

Ο πίνακας εισόδου, «tweet_words», της τεχνικής Naive Bayes περιλαμβάνει πληροφορίες σχετικά με τις καταμετρημένες λέξεις από όλα τα tweets. Πρόκειται για έναν πίνακα διαστάσεων «num_tweets» x «num_com_words», δηλαδή πλήθους γραμμών ίσου με το πλήθος των διαθέσιμων tweets και πλήθους στηλών ίσου με το πλήθος των κοινών λέξεων (text-document matrix, tdm). Τα στοιχεία του πίνακα αυτού αφορούν στο πλήθος της κάθε λέξης που εντοπίστηκε, έστω και μια φορά σε ένα μόνο tweet, σε κάθε tweet ανεξάρτητα. Δηλαδή η τυχαία θέση (i,j) εκφράζει το πλήθος της λέξης της στήλης «j» στο tweet της γραμμής «i». Είναι προφανές ότι ένας σημαντικά μεγάλος αριθμός στοιχείων θα καταλαμβάνεται από τον αριθμό 0, καθώς υπάρχουν πολλές λέξεις, οι οποίες εντοπίζονται σε ένα και μόνο tweet, αλλά δεν ξαναβρίσκονται ποτέ ξανά, σε κανένα άλλο δηλαδή tweet. Η συγκεκριμένη περίπτωση λέξης της στήλης έστω m θα έχει ένα στοιχείο 1 (στον δείκτη του tweet που εντοπίστηκε έστω n) και όλες οι υπόλοιπες θα είναι 0, δηλαδή θα ισχύει $\text{tweet_words}[m,n] = 1$ και $\text{tweet_words}[i,j] = 0$ με $(i,j) \neq (m,n)$. Γενικά, η συχνότητα εμφάνισης της ίδιας λέξης σε ένα και μόνο tweet είναι αρκετά σπάνια και συνεπώς ο μεγαλύτερος αριθμός που εμφανίζουν τα στοιχεία του «tdm» δεν ξεπερνάει το 3.

Βήμα 8^ο

Έχοντας κατασκευάσει το πίνακα εισόδου «tweet_words» για την τεχνική Naive Bayes, καλέστηκε η σχετική ρουτίνα naiveBayes που παρέχεται από το πακέτο «e1071», προκειμένου να εκπαιδευτεί το δείγμα (train set). Ως δείγμα εκπαίδευσης επιλέγεται το 1/3 του δείγματος ελέγχου (test set), δηλαδή ένας πίνακας ο οποίος έχει πλήθος γραμμών «num_tweets/3» και πλήθος στηλών όσο και το αρχικό πλήθος των κοινών λέξεων «num_com_words». Ως μεταβλητή-στόχος στο δείγμα επιλέγεται προφανώς ο χαρακτηρισμός συναισθήματος (sentiment) που αφορά στο συγκεκριμένο τμήμα του δείγματος ελέγχου. Η πρόβλεψη για το σύνολο του δείγματος ελέγχου δίνεται στη συνέχεια μέσω της εντολής «predict», ένας πίνακας-στήλη, ο οποίος περιλαμβάνει τον εκτιμώμενο χαρακτηρισμό συναισθήματος για το σύνολο των tweets που εξετάστηκαν, «num_tweets».

Βήμα 9^ο

Τέλος, η απόδοση του αλγόριθμου υπολογίζεται με βάση το πλήθος των επιτυχιών που είχε σε σχέση με τους παρατηρούμενους, δηλαδή επισημειωμένους, χαρακτηρισμούς συναισθήματος για το σύνολο του δείγματος ελέγχου. Το ποσοστό επιτυχίας υπολογίζεται ως το πλήθος των επιτυχιών σε σχέση με το πλήθος των εξεταζόμενων tweets, όσες φορές δηλαδή ο χαρακτηρισμός εκτιμήθηκε ίδιος με τον επισημειωμένο. Για πρακτικούς λόγους, κατασκευάστηκε ένας πίνακας-στήλη, ο «success», στον οποίο κατά τον έλεγχο αποδόθηκε η τιμή 1 στις περιπτώσεις που υπήρξε επιτυχής ταύτιση πρόβλεψης-επισημείωσης και η τιμή 0 σε περίπτωση αποτυχίας. Το πλήθος των επιτυχιών προσδιορίζεται εύκολα από το άθροισμα όλων των μονάδων, δηλαδή των επιτυχιών, των στοιχείων του πίνακα success, ενώ το ποσοστό επιτυχίας, «accuracy», μέσω του λόγου των επιτυχιών αυτών, «sum(success)», προς το πλήθος των tweets, «num_tweets».

Η ίδια διαδικασία επαναλήφθηκε και για τα υπόλοιπα τρία αρχεία, αυτά των ελληνικών media, καθώς και των Ελλήνων και Γερμανών χρηστών.

Οι επιδόσεις του αλγόριθμου, όμως, ειδικά στην περίπτωση των MME, δεν αποδείχθηκαν ιδιαίτερα ικανοποιητικές, καθώς δεν μπόρεσαν να ξεπεράσουν το 20%, γεγονός που κατέστησε αναγκαία την αναζήτηση κατάλληλων τροποποιήσεων που θα μπορούσαν να συνεισφέρουν στην περαιτέρω βελτίωση των επιδόσεων.

3.4.1 ΤΡΟΠΟΠΟΙΗΣΕΙΣ ΑΛΓΟΡΙΘΜΟΥ ΓΙΑ ΒΕΛΤΙΩΣΗ ΤΗΣ ΑΠΟΔΟΣΗΣ

- **1η Τροποποίηση: Διαδικασία εκτίμησης απόδοσης αλγόριθμου**

Ο βασικός αλγόριθμος, αρχικά, σχεδιάστηκε να εκτιμά την επίδοσή του θέτοντας ως δείγμα εκπαίδευσης (train set) το 33% του συνολικού δείγματος και ως δείγμα ελέγχου (test set) το ίδιο το αρχικό δείγμα. Η προσέγγιση αυτή δεν ήταν απόλυτα σωστή, γιατί αφενός το δείγμα εκπαίδευσης ήταν πολύ μικρό σε σχέση με το συνολικό δείγμα, αφετέρου το δείγμα ελέγχου ήταν δυσανάλογα μεγάλο. Επομένως, κρίθηκε αναγκαία η αντικατάσταση του τμήματος αυτού και η εκπαίδευση του 70% του συνολικού δείγματος, αφήνοντας το υπόλοιπο 30% προς έλεγχο της απόδοσης. Για λόγους

ευκολίας στη διαδικασία σύνταξης του κώδικα, το δείγμα εκπαίδευσης αφορά χωροταξικά στο τελευταίο 70% του συνολικού δείγματος, ενώ το δείγμα ελέγχου το αρχικό 30%. Ως παράμετρος, η ποσόστωση αυτή έχει εισαχθεί με τέτοιο τρόπο ώστε να είναι ελεύθερη και εύκολα ρυθμήσιμη από τον χρήστη.

Η εν λόγω τροποποίηση οδήγησε σε μικρή αύξηση της απόδοσης, μέχρι 5%, καθώς η προηγούμενη προσέγγιση υποεκτιμούσε σε κάποιο βαθμό τις επιδόσεις. Παρόλα αυτά το ποσοστό επιτυχίας παρέμεινε μικρότερο του 50% και για το λόγο αυτό αναζητήθηκαν νέες λύσεις προς την κατεύθυνση των βελτιώσεων.

- **2^η Τροποποίηση: Περιορισμός λέξεων με το ίδιο νόημα αλλά με διαφορετική δομή**

Είναι προφανές ότι το μεγάλο πλήθος των λέξεων που εισάγονται στη διαδικασία εκπαίδευσης ως επεξηγηματικές μεταβλητές εισόδου, δυσχεραίνει την απόδοση του αλγορίθμου, καθώς οι περισσότερες από αυτές δεν εμφανίζονται παρά μία μόνο φορά σε ένα και μοναδικό tweet. Το αποτέλεσμα είναι ο πίνακας εισόδου λέξεων-κειμένου (text-document matrix, tdm) να είναι πολύ αραιός, δηλαδή επιφορτισμένος με πολλά μηδενικά. Οι λέξεις αυτές δεν έχουν σημαντική επίδραση στη μεταβλητή-στόχο, τον χαρακτηρισμό δηλαδή του συναισθήματος, ενώ επιβαρύνουν χρονικά και αποπροσανατολίζουν την υπολογιστική διαδικασία από τον επιτυχή χαρακτηρισμό.

Η τροποποίηση στηρίχθηκε στην απλή σκέψη της παρουσίας πολυάριθμων λέξεων με κοινή ρίζα αλλά διαφορετική δομή, όπως π.χ. γένος, κλίση κλπ. Στην πράξη δηλαδή, στόχος είναι ο περιορισμός των λέξεων, κρατώντας μόνο ορισμένα από τα αρχικά τους γράμματα, με την υπόθεση ότι πολλές από αυτές έχουν την ίδια ερμηνεία και την ίδιο προσανατολισμό ταξινόμησης. Επομένως, μετά την κατασκευή του πίνακα-στήλη «words», ο οποίος περιλαμβάνει όλες τις λέξεις που εντοπίστηκαν σε όλα τα tweets ανεξαιρέτως (ανεξάρτητα αν ήταν κοινές μεταξύ τους ορισμένες από αυτές), κατασκευάστηκε ένας ακόμα πίνακας-στήλη, `trunc_words`, με `tot_num_words` συνολικά θέσεις, τον μέγιστο δηλαδή δυνατό αριθμό, ώστε να εξασφαλιστεί ότι δεν θα υπάρχουν ελλείψεις, ανεξάρτητα από τον τρόπο επιλογής που θα ακολουθηθεί. Οι κοινές ρίζες

παράγονται μέσω της εξαγωγής των πρώτων `num_letters` από τις αρχικές λέξεις και της περικοπής της υπόλοιπης κατάληξης. Η λειτουργία αυτή υποστηρίζεται από την εντολή `"substr"` της R η οποία δίνει τη δυνατότητα απομόνωσης οποιουδήποτε τμήματος λέξεων είναι επιθυμητό. Στην περίπτωση π.χ. της λέξης «πολυκατοικία», `num_letters = 5` σημαίνει ότι θα παραχθεί η ρίζα «πολυκ», ενώ για `num_letters = 2` η ρίζα «πο». Τον ρόλο πλέον των αρχικών λέξεων έχουν αντικαταστήσει οι λέξεις-ρίζες.

Στη συνέχεια, απομακρύνθηκαν οι επαναλαμβανόμενες λέξεις-ρίζες μέσω της εντολής `"unique"` και σχηματίστηκε ο πίνακας-στήλη, `common_words`, ο οποίος περιλαμβάνει τις λέξεις-ρίζες που εντοπίστηκαν από μία μόνο φορά, χωρίς δηλαδή επαναλήψεις. Μάλιστα, εξαιτίας της συνήθους παρουσίας ενός «κενού», το οποίο εκλαμβάνεται ως λέξη, απομακρύνθηκε και αυτό από τη διαδικασία εκπαίδευσης μέσω της εντολής `"!="`. Το πλήθος των λέξεων-ριζών πλέον υποβιβάστηκε σε ένα σημαντικά μικρότερο αριθμό `num_com_words` που μετρήθηκε και πάλι με την εντολή `"length"`, ακολουθώντας πλέον την ίδια πορεία με εκείνη του βασικού αλγόριθμου.

Η τροποποίηση αυτή είχε σημαντικές βελτιώσεις στην απόδοση του αλγόριθμου της τάξης του 10%, όταν η θέση περικοπής αφορούσε τα δύο ή τρία πρώτα γράμματα των λέξεων (παράγοντας δηλαδή ρίζες δύο ή τριών γραμμάτων), καθώς πολλές ελληνικές λέξεις είχαν την ίδια ερμηνεία και προσανατολισμό συναισθήματος με βάση την ίδια ρίζα και συνεπώς ο πίνακας λέξεων-κειμένου (`tdm`) περιορίστηκε σημαντικά. Αντίθετα, στην περίπτωση των γερμανικών λέξεων δεν παρατηρήθηκε βελτίωση, η απόδοση απλά διατηρήθηκε στα επίπεδα της απόδοσης του αρχικού αλγόριθμου. Πιθανή εξήγηση της διαπίστωσης αυτής θα μπορούσε να αφορά στο μικρό πλήθος ριζών στη γερμανική γλώσσα που είναι ικανές να παράγουν λέξεις με το ίδιο νόημα. Ενδεχόμενη δηλαδή παρατήρηση κοινών πρώτων γραμμάτων σε ορισμένες γερμανικές λέξεις δεν οδηγεί αναγκαστικά σε κοινή νοηματική ερμηνεία.

- **Αξιολόγηση Αλγόριθμου με τη βοήθεια Συνθετικών Δειγμάτων**

Είναι φανερό ότι η απόδοση του αλγορίθμου Naive Bayes δεν είναι τόσο αποτελεσματική συγκριτικά με εκείνη της χρήσης λεξικού, παρά τις τροποποιήσεις που

πραγματοποιήθηκαν. Διερευνώντας πιθανά αίτια στη διαδικασία εκπαίδευσης, σχεδιάστηκε ένας κώδικας αξιολόγησης του αλγόριθμου κάτω από ελεγχόμενες συνθήκες, υποβάλλοντας δηλαδή τη μεθοδολογία σε δοκιμασίες που πρόκειται να αναδείξουν το βαθμό επιρροής συγκεκριμένων παραγόντων, οι οποίοι είχαν αμεληθεί ή τουλάχιστον υποεκτιμηθεί. Η αξιολόγηση έγινε με τη βοήθεια της τεχνικής Monte Carlo, με σκοπό την παραγωγή μεγάλου πλήθους συνθετικών δειγμάτων και ιδιότητες της επιλογής του ερευνητή. Στη συνέχεια, τα συνθετικά tweets υποβλήθηκαν στη διαδικασία εκπαίδευσης ακολουθώντας την ίδια ακριβώς μεθοδολογία του αρχικού αλγόριθμου, με σκοπό τη συγκριτική παρουσίαση των αποτελεσμάτων. Αξιολογώντας τον απαιτούμενο υπολογιστικό χρόνο, οι δοκιμασίες χρονομετρήθηκαν με τη βοήθεια των εντολών “`rtm<-proc.time()`” και “`proc.time()-rtm`” κατά την έναρξη και την ολοκλήρωση της διαδικασίας αντίστοιχα.

Βήμα 1^ο

Κορμός της προσέγγισης που ακολουθήθηκε, αποτέλεσε η τεχνική αναδειγματοληψίας με επανάθεση (bootstrap), με σκοπό την παραγωγή δειγμάτων (tweets) του ίδιου μεγέθους με ένα προεπιλεγμένο δείγμα αναφοράς (συνθετικό bag of words) από το οποίο προέρχονται όλες οι δυνατές λέξεις. Η επιλογή κάθε λέξης κατά τη δειγματοληψία είναι ισοπίθανη με όλες τις υπόλοιπες, εκείνο δηλαδή το οποίο διαφοροποιείται σε κάθε συνθετικό tweet είναι η συχνότητα εμφάνισης των λέξεων. Είναι προφανές ότι, εφόσον τα συνθετικά tweets είναι ισομεγέθη, κάποιες λέξεις δεν θα επιλεγθούν καθόλου, ενώ κάποιες άλλες θα συνοδεύονται από πολλαπλές εμφανίσεις. Εκτός από το πακέτο «e1071», το οποίο υλοποιεί τη μέθοδο Naive Bayes, κλήθηκε και το πακέτο «bootstrap» το οποίο υλοποιεί την τεχνική αναδειγματοληψίας επανάθεσης.

Βήμα 2^ο

Αποσκοπώντας στην προσομοίωση του προβλήματος που μελετάται, κατασκευάστηκαν δύο επιμέρους υποπίνακες, `pos.words` και `neg.words`, οι οποίοι περιλαμβάνουν από 15 λέξεις έκαστος που υποδηλώνουν θετικό ή αρνητικό χαρακτήρα αντίστοιχα. Στη συνέχεια οι δύο υποπίνακες συγχωνεύθηκαν σε έναν νέο πίνακα «`com.words`», ο οποίος αποτελεί πλέον τον συνθετικό πίνακα αναφοράς, τον γεννήτορα δυνατών λέξεων (συνθετικό bag of words). Έπειτα, με την εντολή «`length`» είναι δυνατόν να υπολογιστεί το πλήθος των λέξεων που περιλαμβάνουν, δηλαδή των

num_pos.words, num_neg.words, και num_com.words (15, 15, 30 αντίστοιχα) στην περίπτωση που αργότερα είναι επιθυμητή κάποια μεταβολή.

Οι πίνακες που δημιουργήθηκαν είναι οι εξής:

pos.words	neg.words
good	bad
nice	sad
gentle	mean
happy	ugly
smiling	selfish
beautiful	cruel
kind	awful
warm	horrible
handsome	unhappy
sweet	miserable
friendly	dreadful
lovely	deficient
sympathetic	evil
welcome	inhuman
helpful	brutal

Βήμα 2^ο

Με τη βοήθεια της συνάρτησης `boot.words(x,data)`, η οποία κατασκευάστηκε ώστε να υλοποιήσει την αναδειγματοληψία κάθε στοιχείου-λέξης x από τον πίνακα αναφοράς `com.words (data)`, εφαρμόστηκε η τεχνική μέσω της εντολής «bootstrap» με σκοπό την παραγωγή `num_tweets` συνολικά συνθετικών tweets, δηλαδή `num_tweets` συνθετικών πινάκων, οι οποίοι περιέχουν `num_com.words = 30` πάντοτε λέξεις ισοπίθανα επιλεγμένες από τον γεννήτορα πίνακα `com.words`. Οι διαστάσεις του πίνακα tweets είναι `num_com.words × num_tweets`, δηλαδή για κάθε στήλη $j = 1, \dots, \text{num_tweets}$ περιλαμβάνονται οι `num_com.words` λέξεις που συνθέτουν το tweet (στις $i = 1, \dots, \text{num_com.words}$ του γραμμής).

Βήμα 3^ο

Ομοίως με την εφαρμογή του αλγόριθμου Naive Bayes στα πραγματικά tweets, κατασκευάστηκε ο πίνακας εισόδου `tweet_words (text-document matrix, tdm)`, ο οποίος διαθέτει διαστάσεις `num_tweets × num_com_words`, δηλαδή πλήθους γραμμών

ίσου με το πλήθος των συνθετικών tweets και πλήθους στηλών ίσου με το πλήθος των λέξεων του πίνακα γεννήτορα λέξεων. Τα στοιχεία του πίνακα αυτού αφορούν στο πλήθος της κάθε λέξης που εντοπίστηκε, έστω και μια φορά σε ένα μόνο tweet, σε κάθε tweet ανεξάρτητα. Δηλαδή η τυχαία θέση (i,j) εκφράζει το πλήθος της λέξης της στήλης j στο συνθετικό tweet της γραμμής i .

Μια σημαντική διαφορά μεταξύ των δυνατοτήτων που προσφέρει η διερεύνηση των επιδόσεων του αλγόριθμου Naive Bayes σε συνθετικά δείγματα έναντι των πραγματικών αφορά στην αντικειμενική επισύναψη του κάθε tweet με έναν σαφή, αλλά το κυριότερο, όμοιο τρόπο για όλα τα tweets ανεξαιρέτως. Ο χαρακτηρισμός συναισθήματος στα πραγματικά tweets με βάση την επισημείωση στηρίζεται κυρίως σε υποκειμενικά κριτήρια τα οποία εξαρτώνται από τις προσωπικές πεποιθήσεις, την τρέχουσα συναισθηματική διάθεση, το κοινωνικό και μορφωτικό υπόβαθρο του κριτή. Τα κριτήρια αυτά μάλιστα δεν είναι βέβαιο ότι εφαρμόζονται ισότιμα σε όλα τα tweets, καθώς ο κριτής αδυνατεί κάθε φορά να υλοποιεί την ίδια στρατηγική με τον ίδιο αυστηρό τρόπο.

Στη συγκεκριμένη περίπτωση, στρατηγική αποτέλεσε η διαμόρφωση ενός απλού κανόνα με βάση την υπεροχή ή όχι του πλήθους λέξεων με θετικό χαρακτήρα έναντι εκείνων με αρνητικό χαρακτήρα. Επομένως, απαραίτητη προϋπόθεση είναι η καταμέτρηση λέξεων που ανήκουν στους υποπίνακες pos.words και neg.words για κάθε συνθετικό tweet, ώστε στη συνέχεια οι δύο αυτές τιμές να συγκριθούν. Κατασκευάστηκαν αρχικά οι πίνακες tweet_pos.words και tweet_neg.words, οι οποίοι έχουν διαστάσεις $\text{num_tweets} \times \text{num_pos.words}$ και $\text{num_tweets} \times \text{num_neg.words}$, αντίστοιχα, με την γραμμή i να εκφράζει το i -οστό από τα num_tweets συνολικά tweets και τη στήλη j να εκφράζει την j -οστή λέξη από τις num_pos.words συνολικά θετικού χαρακτήρα και τις num_neg.words αρνητικού. Η τυχαία θέση (i,j) του πίνακα tweet_pos.words καταλαμβάνεται, μετά από καταμέτρηση μέσω της εντολής «length», από τη συχνότητα εμφάνισης της j -οστής λέξης θετικού χαρακτήρα στο i -οστό tweet και με τον ίδιο τρόπο καταλαμβάνεται η τυχαία θέση του πίνακα tweet_neg.words για τις λέξεις αρνητικού χαρακτήρα. Στη συνέχεια, μέσω της εντολής “sum”, αθροίζονται οι συνολικές λέξεις θετικού ή αρνητικού χαρακτήρα που περικλείει το i -οστό tweet και το αποτέλεσμα τοποθετείται στην αντίστοιχη θέση των πινάκων num_tweet_pos.words και num_tweet_neg.words, αντίστοιχα.

Βήμα 4^ο

Το τελευταίο βήμα πριν την εκτέλεση του αλγόριθμου Naive Bayes είναι πλέον η θέσπιση του κανόνα ταξινόμησης των συνθετικών tweets. Αρχικά κατασκευάστηκε ένας πίνακας-στήλη, `sentim`, ο οποίος περιλαμβάνει τον χαρακτηρισμό συναισθήματος για τα `num_tweets` συνολικά tweets. Στην περίπτωση μας ως κριτήριο «θετικού» χαρακτηρισμού τέθηκε το κατώφλι του 60% των λέξεων με θετικό χαρακτήρα (δηλαδή λέξεων που αποτελούν στοιχεία του υποπίνακα `pos.words`) σε σχέση με το σύνολο λέξεων που περιέχει το κάθε tweet ξεχωριστά. Εάν το κριτήριο αυτό δεν ικανοποιείται, επαναλαμβάνεται ο ίδιος αριθμητικός έλεγχος για την περίπτωση των λέξεων με αρνητικό χαρακτήρα, δηλαδή λέξεων που αποτελούν στοιχεία του υποπίνακα `neg.words`) και εφόσον επαληθεύεται, το tweet χαρακτηρίζεται ως «αρνητικό». Αν κανένα από τα ανωτέρω κριτήρια δεν ικανοποιείται, το tweet χαρακτηρίζεται ως «ουδέτερο». Η επιβολή του 60% δεν είναι αυστηρή. Αφήνεται ελεύθερη από τον ερευνητή μέσω της παραμέτρου `sent.thres` και μπορεί να λάβει τιμές στο διάστημα (50%,100%).

Βήμα 5^ο

Έχοντας κατασκευάσει τον πίνακα εισόδου `tweet_words` για την τεχνική Naive Bayes, όπως έγινε και στον βασικό αλγόριθμο μελέτης των πραγματικών tweets, κλήθηκε η σχετική ρουτίνα `naiveBayes` που παρέχεται από το πακέτο «`e1071`» και εκπαιδεύτηκε το δείγμα (`train set`). Ως δείγμα εκπαίδευσης επιλέχθηκε το 70% του συνολικού δείγματος, αφήνοντας το υπόλοιπο 30% προς έλεγχο της απόδοσης (`test set`). Όπως ήδη ειπώθηκε κατά την πρώτη τροποποίηση της προσέγγισης, για λόγους ευκολίας στη διαδικασία σύνταξης του κώδικα, το δείγμα εκπαίδευσης τέθηκε χωροταξικά το τελευταίο 70% του συνολικού δείγματος, ενώ το δείγμα ελέγχου το αρχικό 30%. Η ποσόστωση αυτή έχει εισαχθεί με τέτοιο τρόπο ώστε να είναι ελεύθερη και εύκολα ρυθμίσιμη από τον χρήστη μέσω της παραμέτρου `test.thres`. Ως μεταβλητή-στόχος επιλέγεται προφανώς ο χαρακτηρισμός συναισθήματος (`sentiment`) που αφορά στο δείγμα εκπαίδευσης και παρέχεται μέσω του πίνακα-στήλη `sentim` μεταξύ των θέσεων `test.thres` και `num_tweets`. Η πρόβλεψη δίνεται στη συνέχεια μέσω της εντολής “`predict`”, οδηγώντας σε έναν πίνακα-στήλη, `results`, ο οποίος περιλαμβάνει τον

εκτιμώμενο χαρακτηρισμό συναισθήματος για το σύνολο των tweets στο δείγμα ελέγχου, δηλαδή μεταξύ των θέσεων 1 και `test.thres` του αρχικού δείγματος.

Η εκτίμηση του βαθμού απόδοσης του αλγόριθμου στηρίζεται στη σύγκριση των χαρακτηρισμών συναισθήματος που προέκυψε μέσω της πρόβλεψης σε σχέση με τον υπάρχοντα χαρακτηρισμό από το δείγμα ελέγχου (επισημείωση). Κατασκευάστηκε ένας πίνακας-στήλη, `success`, τα στοιχεία του οποίου περιλαμβάνουν είτε την τιμή 1 είτε την τιμή 0 ανάλογα με το αν παρατηρηθεί ταύτιση των χαρακτηρισμών (επιτυχία) ή όχι (αποτυχία) κατά τη σύγκριση επισημείωσης (του δείγματος ελέγχου) και του εκτιμώμενου χαρακτηρισμού συναισθήματος (του προβλεπτικού πίνακα). Το ποσοστό απόδοσης, `accuracy`, υπολογίζεται τελικά από το πλήθος των επιτυχιών (άθροισμα μονάδων) σε σχέση με το πλήθος των tweets που εξετάστηκαν, δηλαδή `test.thres` στο σύνολο.

Οι δοκιμασίες στις οποίες υποβλήθηκε ο αλγόριθμος αφορούσαν στη διερεύνηση της απόδοσης και του απαιτούμενου υπολογιστικού χρόνου κάτω από την επιρροή τριών παραγόντων: (i) του πλήθους συνθετικών tweets, (ii) του βαθμού αυστηρότητας του κανόνα ταξινόμησης. δηλαδή το κατώφλι μετάβασης από θετικό σε αρνητικό χαρακτηρισμό συναισθήματος και (iii) του μεγέθους του δείγματος εκπαίδευσης. Είναι προφανές ότι κατά την αξιολόγηση καθενός από τους πιο πάνω παράγοντες, οι υπόλοιποι δύο ήταν σταθεροποιημένοι. Να σημειωθεί εδώ οι υπολογισμοί έγιναν σε ηλεκτρονικό υπολογιστή με επεξεργαστή τύπου `i3-3217U` δύο πυρήνων συχνότητας 1.80 GHz και 4.0 GB μνήμης RAM.

3.5 ΑΝΑΛΥΣΗ ΣΥΝΑΙΣΘΗΜΑΤΟΣ ΜΕ ΤΗ ΜΕΘΟΔΟ ΤΟΥ ΛΕΞΙΚΟΥ ΑΝΑ ΚΑΤΗΓΟΡΙΑ ΤΩΝ ΣΤΟΧΩΝ ΤΟΥ ΣΥΝΑΙΣΘΗΜΑΤΟΣ

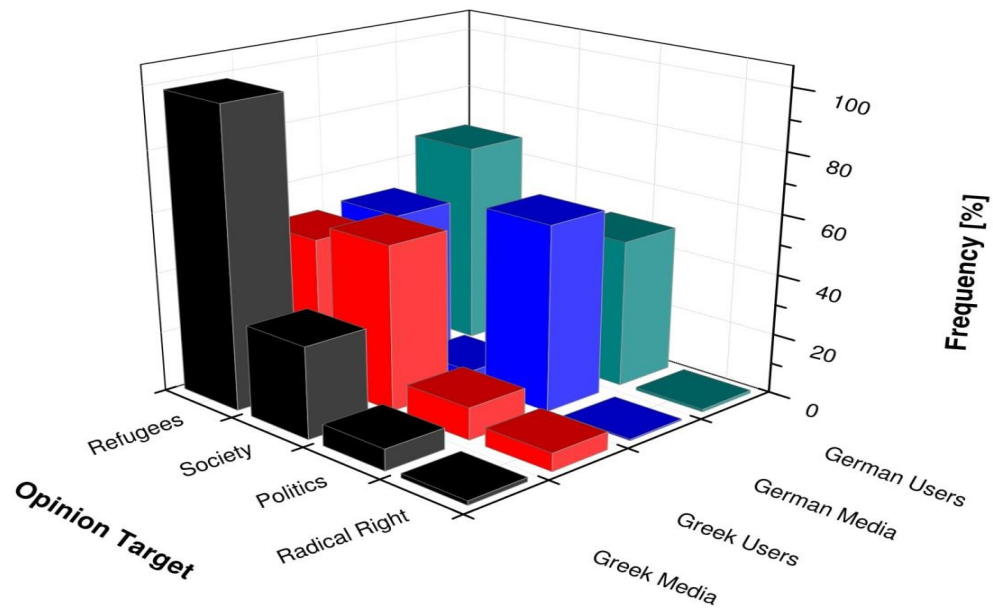
Λαμβάνοντας υπόψιν την αποδοτικότητα των δύο μεθόδων στα πειράματα που διεξήχθησαν, πραγματοποιήθηκε στη συνέχεια ανάλυση συναισθήματος με τη μέθοδο του λεξικού, αυτή τη φορά ξεχωριστά για κάθε μία από τις τέσσερις κατηγορίες των στόχων του συναισθήματος και για τις δύο γλώσσες, προκειμένου να μετρηθεί πιο αναλυτικά η αποδοτικότητα της συγκεκριμένης μεθόδου.

Η διαδικασία που ακολουθήθηκε είναι η ίδια με αυτήν που εφαρμόστηκε για ολόκληρο το σώμα των tweets σε προηγούμενο στάδιο της παρούσας μελέτης.

ΑΠΟΤΕΛΕΣΜΑΤΑ ΕΡΕΥΝΑΣ & ΣΥΜΠΕΡΑΣΜΑΤΑ

1. ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΝΑΛΥΣΗΣ ΠΕΡΙΕΧΟΜΕΝΟΥ

1.1 ΣΤΟΧΟΙ ΣΥΝΑΙΣΘΗΜΑΤΟΣ



Γράφημα 1

Στο παραπάνω γράφημα φαίνεται ο βαθμός στον οποίο τα Μ.Μ.Ε στην Ελλάδα και στη Γερμανία, αλλά και οι Έλληνες και οι Γερμανοί χρήστες εκφράζουν με τις αναρτήσεις τους στο Twitter συναίσθημα για κάθε μία από τέσσερις ομάδες των στόχων του συναισθήματος, όπως αυτές προέκυψαν μέσω της R και έτσι δίνεται απάντηση στο 1^ο ερευνητικό ερώτημα. Αρχικά μπορεί να παρατηρήσει κανείς ότι το συναίσθημα στο μεγαλύτερο μέρος του εκφράζεται αποκλειστικά για τους πρόσφυγες, ιδιαίτερα από τα ελληνικά Μ.Μ.Ε, γεγονός που φαίνεται λογικό, καθώς όλα τα tweets έχουν ως θέμα την εκπαίδευση των προσφύγων. Αντιθέτως, σε ελάχιστο βαθμό φαίνεται το συναίσθημα να έχει ως στόχο την ακροδεξιά.

- **Greek media και greek users**

Όσον αφορά στην Ελλάδα, οι Έλληνες χρήστες είναι αυτοί που εκφράζονται περισσότερο για την ακροδεξιά σε σχέση με τα Μ.Μ.Ε, τα οποία εκφράζουν συναίσθημα προς το συγκεκριμένο στόχο σε πολύ μικρό ποσοστό. Το ίδιο ισχύει και για τα θέματα γύρω από την πολιτική. Οι Έλληνες χρήστες εκφράζουν σε μεγαλύτερο βαθμό άποψη για τους Έλληνες πολιτικούς και την κυβέρνηση από ότι τα ελληνικά media. Ακόμα πιο έντονη είναι η διαφορά αναφορικά με την κοινωνία, με την οποία οι Έλληνες χρήστες καταπιάνονται στο μεγαλύτερο ποσοστό. Αντιθέτως, τα Μ.Μ.Ε της Ελλάδας εκφράζουν αισθητά σε μεγαλύτερο βαθμό άποψη αποκλειστικά για τους πρόσφυγες από ότι οι Έλληνες χρήστες, τους οποίους δείχνει να τους ενδιαφρουν περισσότερο οι υπόλοιποι παράγοντες που εμπλέκονται στο θέμα της εκπαίδευσης των προσφύγων.

- **German media και german users**

Αρκετά διαφοροποιημένη φαίνεται να είναι η κατάσταση στη Γερμανία. Τα γερμανικά Μ.Μ.Ε φαίνεται πως εκφράζουν κυρίως συναίσθημα προς τις οντότητες γύρω από την πολιτική και δευτερευόντως αποκλειστικά για τους πρόσφυγες, ενώ σε αισθητά μικρό βαθμό αναφέρονται στην ακροδεξιά και στην υπόλοιπη κοινωνία. Από την άλλη πλευρά, οι Γερμανοί χρήστες εκφράζουν άποψη σε μεγαλύτερο βαθμό αποκλειστικά για τους πρόσφυγες και την εκπαίδευσή τους, ενώ στη δεύτερη θέση έρχονται οι οντότητες της πολιτικής. Ελάχιστο είναι το ποσοστό των tweets των Γερμανών χρηστών που κατευθύνεται στην ακροδεξιά και στην ευρύτερη κοινωνία.

- **Greek και german media**

Εστιάζοντας στα media των δύο χωρών μπορεί κανείς να παρατηρήσει μία αρκετά διαφορετική εικόνα, με τα γερμανικά να δίνουν βάρος στους πολιτικούς και την κυβέρνηση, ενώ τα ελληνικά στην ελληνική κοινωνία.

- **Greek και german users**

Διαφορετικές ανησυχίες δείχνουν να έχουν οι Έλληνες και οι Γερμανοί χρήστες. Οι μεν Έλληνες ασκούν έντονη κριτική στην ελληνική κοινωνία, οι δε Γερμανοί στους πολιτικούς της χώρας τους.

Στον παρακάτω πίνακα φαίνεται αναλυτικά τη συχνότητα αναφοράς των tweets στον κάθε έναν opinion target. Να σημειωθεί ότι μερικά tweets εκφράζουν στόχο προς δύο ή και παραπάνω οντότητες και γι' αυτό το λόγο υπάρχουν επικαλύψεις των κατηγοριών.

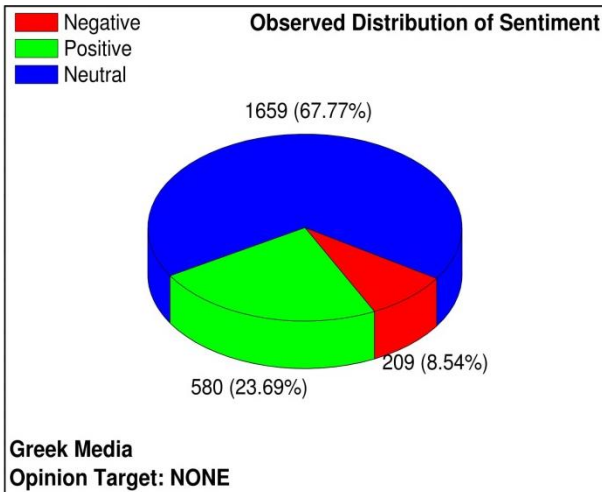
Target Group	Greek Media: 2448		Greek Users: 9764	
Opinion Target	Observed	Frequency [%]	Observed	Frequency [%]
Ακροδεξιά	34	1.39	585	5.99
Πολιτική	173	7.07	1046	10.71
Κοινωνία	744	30.39	5358	54.88
Πρόσφυγες	2414	98.61	4818	49.34

Πίνακας 7. Οι στόχοι συναισθήματος στα ελληνικά tweets

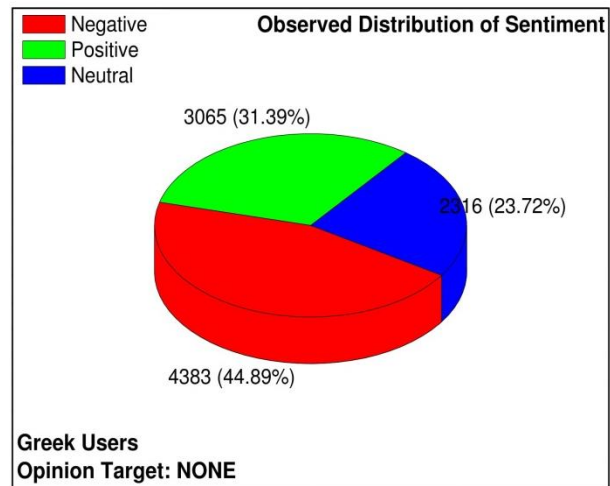
Target Group	German Media: 1743		German Users: 8499	
Opinion Target	Observed	Frequency [%]	Observed	Frequency [%]
Ακροδεξιά	8	0.46	77	0.91
Πολιτική	1072	61.50	4155	48.89
Κοινωνία	79	4.53	351	4.13
Πρόσφυγες	881	50.55	5723	67.34

Πίνακας 8. Οι στόχοι συναισθήματος στα γερμανικά tweets

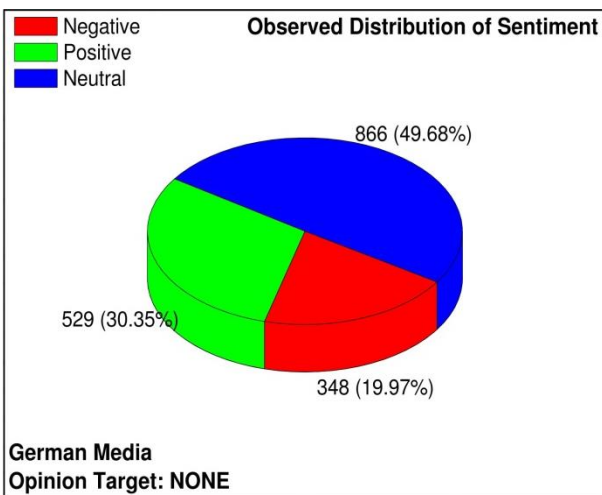
1.2 ΣΥΝΑΙΣΘΗΜΑ



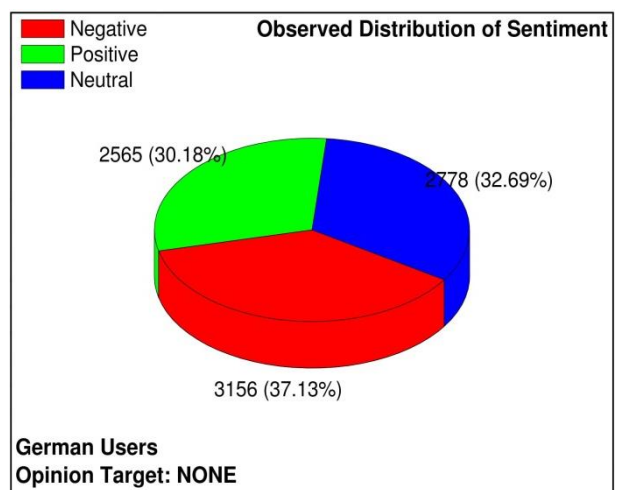
Γράφημα 2



Γράφημα 3



Γράφημα 4

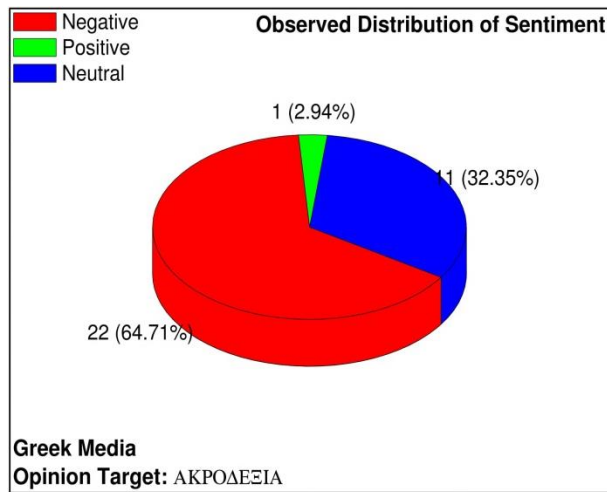


Γράφημα 5

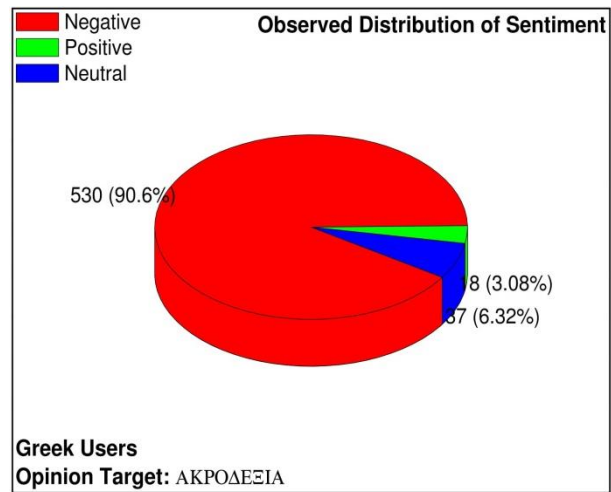
Παρατηρώντας το συναίσθημα, είναι εμφανές ότι τόσο τα ελληνικά όσο και τα γερμανικά M.M.E εκφράζουν σε μεγαλύτερο βαθμό ουδέτερο συναίσθημα επιβεβαιώνοντας έτσι την ερευνητική υπόθεση 3, που θέλει τα media να αναρτούν κατά κύριο λόγο ειδήσεις με ουδέτερο sentiment. Τα ελληνικά M.M.E εμφανίζονται πιο μετριοπαθή από τα γερμανικά, τα οποία αναρτούν ειδήσεις σε ελαφρώς μεγαλύτερο ποσοστό tweets με θετικό ή αρνητικό συναίσθημα, με το θετικό να υπερτερεί του αρνητικού. Τους Έλληνες και Γερμανούς χρήστες η εικόνα είναι διαφορετική, καθώς και οι δύο ομάδες εκφράζουν κυρίως αρνητικό συναίσθημα, με τους Έλληνες να το εκφράζουν σε μεγαλύτερο βαθμό. Οι Γερμανοί εμφανίζονται πιο μετριοπαθείς, καθώς τα ποσοστά του αρνητικού, θετικού και ουδέτερου συναισθήματος είναι αρκετά κοντινά, με το ποσοστό του ουδέτερου συναισθήματος να εμφανίζεται αρκετά υψηλό. Οι Έλληνες αναρτούν tweets με αρνητικό συναίσθημα σε πολύ υψηλό ποσοστό, με το θετικό να έπεται, ενώ μικρά σε αριθμό είναι τα tweets με ουδέτερο συναίσθημα. Έτσι απαντάται το 2^ο ερευνητικό ερώτημα, καθώς το συναίσθημα των media τους κάθε χώρας δεν

φαίνεται να συμβαδίζει με αυτό των πολιτών. Αντιθέτως είναι διαφοροποιημένο σε πολύ μεγάλο βαθμό.

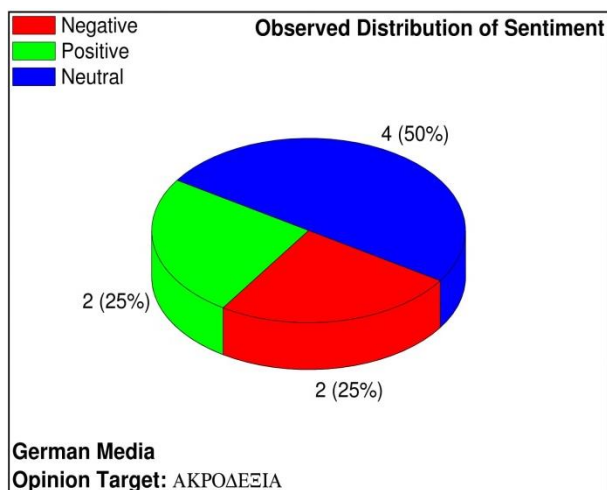
ΑΚΡΟΔΕΞΙΑ



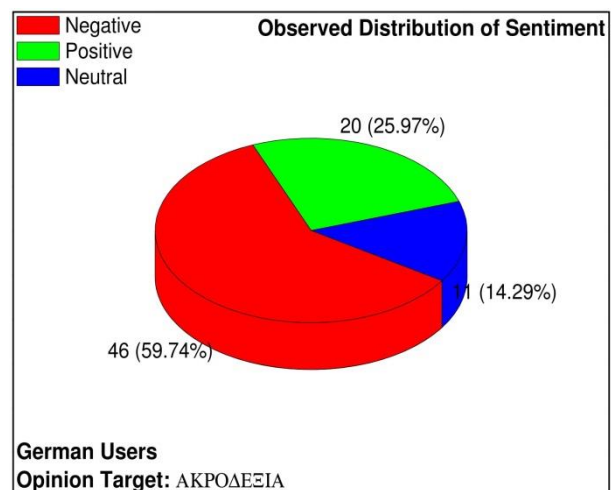
Γράφημα 6



Γράφημα 7

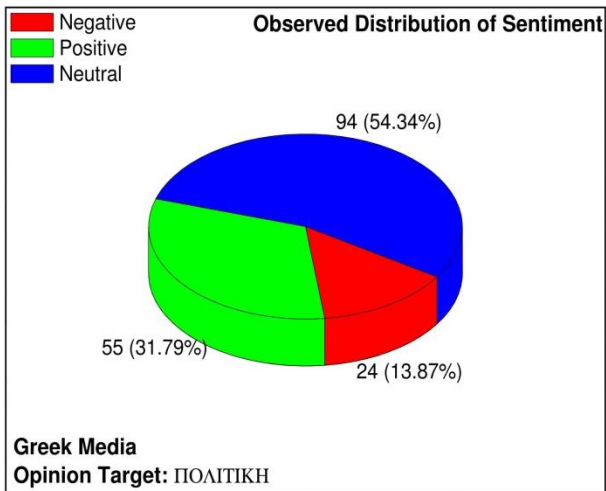


Γράφημα 8

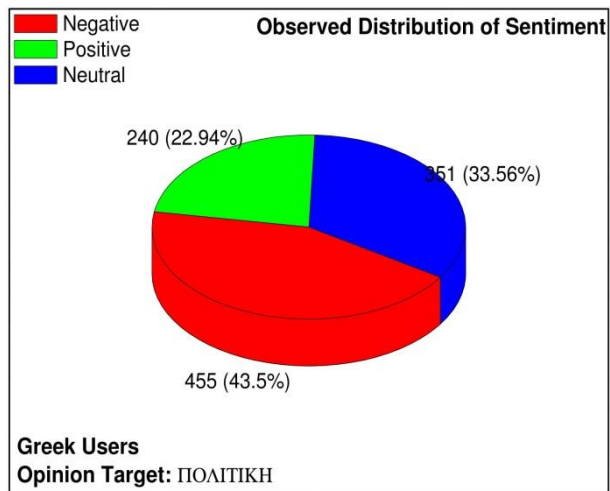


Γράφημα 9

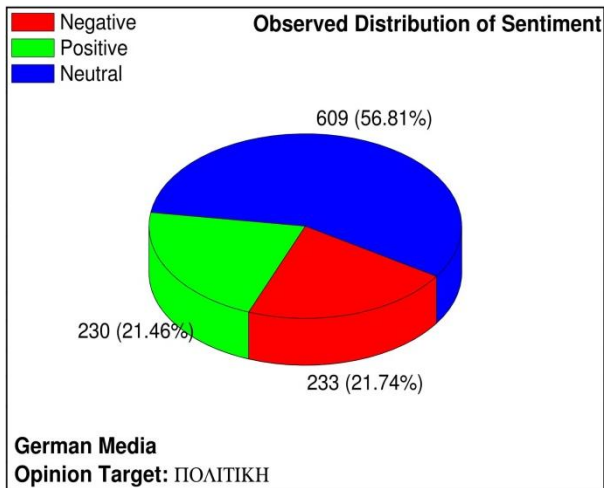
Στα παραπάνω γραφήματα φαίνεται το συναίσθημα που εκφράζουν οι Έλληνες και Γερμανοί χρήστες, καθώς και τα ελληνικά και γερμανικά media για την ακροδεξιά και έτσι δίνεται απάντηση στο 3ο ερευνητικό ερώτημα, δηλαδή στο πώς κρίνουν οι πολίτες, Έλληνες και Γερμανοί, τη στάση τους ακροδεξιάς απέναντι στο θέμα της εκπαίδευσης των προσφύγων. Συγκεκριμένα, οι Έλληνες πολίτες τάσσονται στη συντριπτική τους πλειοψηφία εναντίον της ακροδεξιάς, γεγονός που πιθανώς σχετίζεται και με τις εφόδους που έκαναν μέλη της Χρυσής Αυγής σε σχολεία, για να μην επιτρέψουν την εγγραφή παιδιών προσφύγων. Το αρνητικό συναίσθημα κυριαρχεί και στους Γερμανούς χρήστες, αν και σε μικρότερο ποσοστό. Τα γερμανικά M.M.E εμφανίζουν μία μετριοπαθή στάση, με το ουδέτερο συναίσθημα να επικρατεί, αν και το δείγμα είναι πολύ μικρό για να μπορούμε να βγάλουμε κάποιο συμπέρασμα, όπως και αυτό για τα ελληνικά M.M.E, που όμως διακρίνεται μία τάση προς το αρνητικό συναίσθημα.

ΠΟΛΙΤΙΚΗ

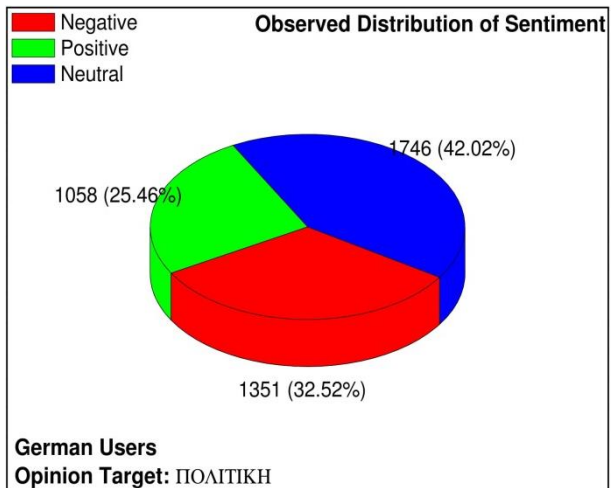
Γράφημα 10



Γράφημα 11

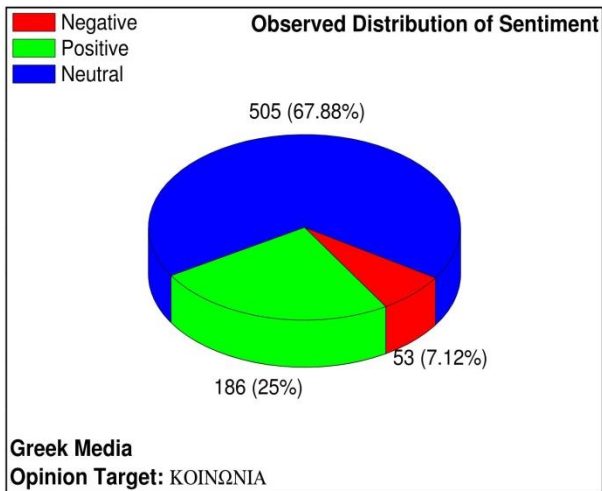


Γράφημα 12

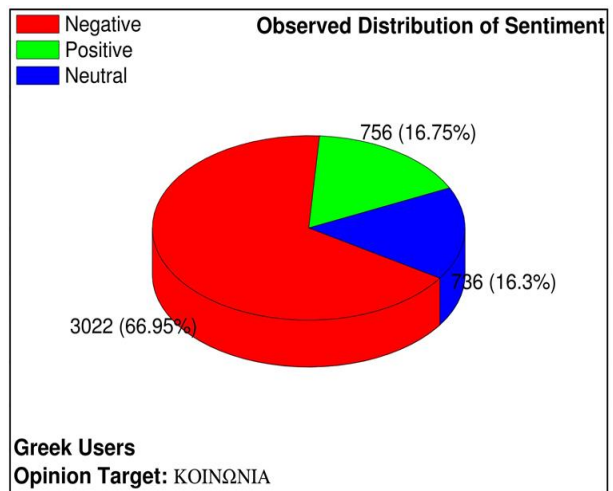


Γράφημα 13

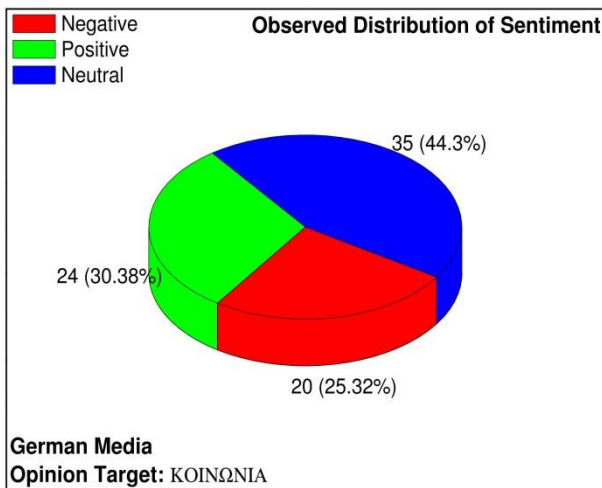
Όσον αφορά στις οντότητες που έχουν να κάνουν με το χώρο της πολιτικής, είναι εμφανής η δυσαρέσκεια των Ελλήνων χρηστών απέναντί τους, ενώ το ίδιο ισχύει και για τους Γερμανούς χρήστες, αν και σε μικρότερο βαθμό. Τα Μ.Μ.Ε και των δύο χωρών εκφράζονται με ουδετερότητα, ενώ εντύπωση προκαλεί ότι αμέσως μετά το ουδέτερο συναίσθημα, όσον αφορά στα ελληνικά Μ.Μ.Ε έρχεται το θετικό, ενώ όσον αφορά στα γερμανικά Μ.Μ.Ε το αρνητικό.

ΚΟΙΝΩΝΙΑ

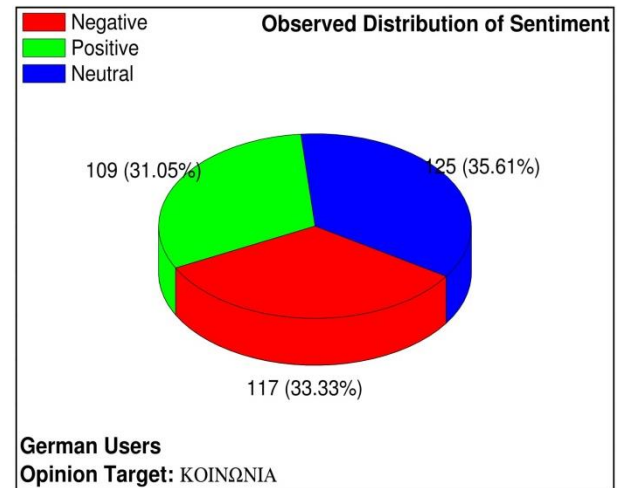
Γράφημα 14



Γράφημα 15

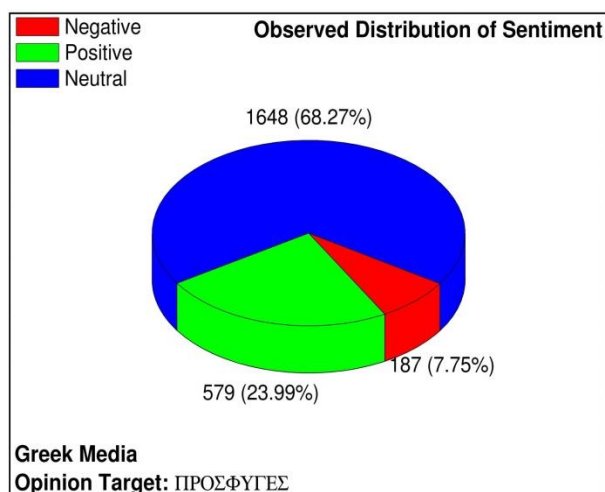


Γράφημα 16

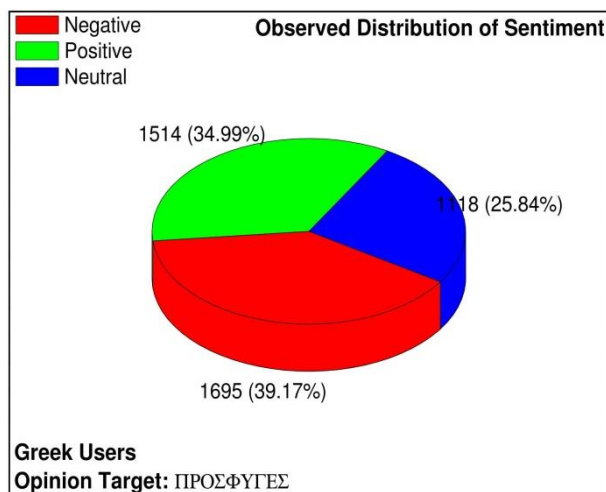


Γράφημα 17

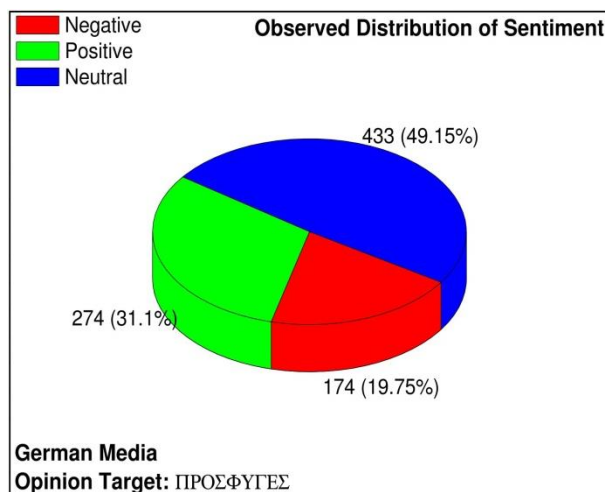
Έντονο αρνητικό συναίσθημα εκφράζουν οι Έλληνες πολίτες για τα μέλη της κοινωνίας, πράγμα που συνδέεται με τα ξενοφοβικά γεγονότα στο Ωραιόκαστρο και σε άλλα μέλη της Ελλάδας, όπου γονείς παιδιών έβαζαν λουκέτο στα σχολεία για να μην έχουν πρόσβαση τα παιδιά των προσφύγων. Αντιθέτως στη Γερμανία, το συναίσθημα για τα μέλη της κοινωνίας φαίνεται να είναι μοιρασμένο εξίσου στις τρεις κατηγορίες συναισθήματος.

ΠΡΟΣΦΥΓΕΣ

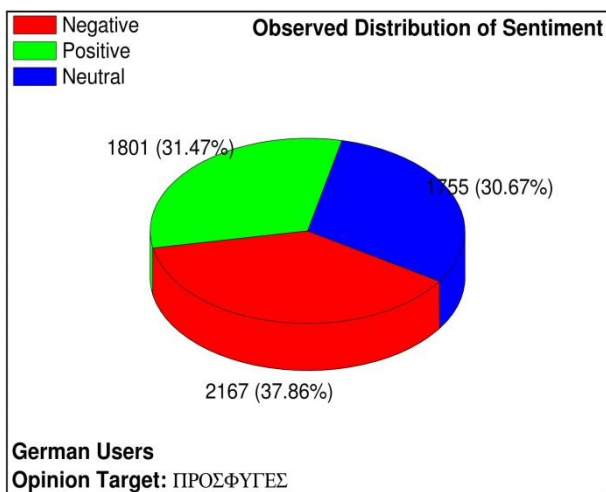
Γράφημα 18



Γράφημα 19



Γράφημα 20

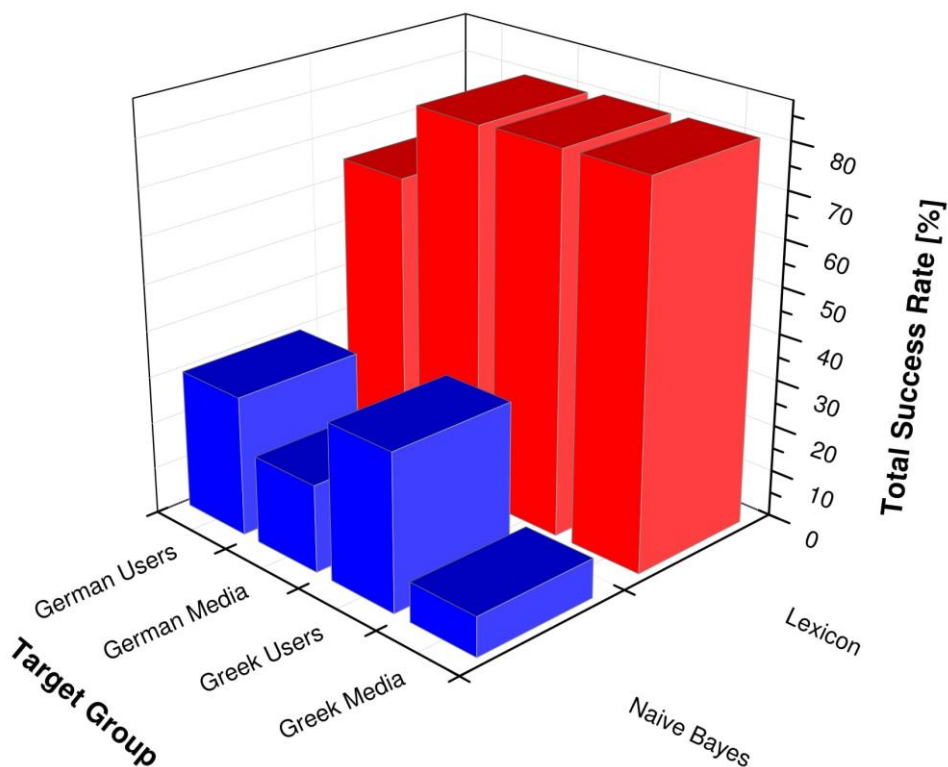


Γράφημα 21

Τέλος, τόσο οι Έλληνες, όσο και οι Γερμανοί χρήστες δείχνουν να εκφράζονται αρνητικά για τους ίδιους τους πρόσφυγες και την εκπαίδευσή τους, αλλά, όπως μπορεί να παρατηρήσει κανείς, η διαφορά των αρνητικών και των θετικών αναρτήσεων είναι σχετικά μικρή, της τάξης περίπου του 5-6%. Επομένως, αναφορικά με την 1^η ερευνητική υπόθεση, δεν θα μπορούσε κανείς να ισχυριστεί ότι οι αναρτήσεις στο Twitter εκφράζουν ξενοφοβία απέναντι στα παιδιά των προσφύγων και στο ενδεχόμενο ενσωμάτωσής τους στο εκπαιδευτικό σύστημα της εκάστοτε χώρας. Όσον αφορά στη 2^η ερευνητική υπόθεση, οι Έλληνες δεν φαίνεται να είναι πιο ξενοφοβικοί από τους Γερμανούς, καθώς μπορεί να εκφράζουν κατά μόλις 2% σε μεγαλύτερο βαθμό αρνητικό συναίσθημα, αλλά από την άλλη πλευρά εκφράζουν κατά περίπου 3% περισσότερο θετικό συναίσθημα, με τους Γερμανούς να εμφανίζονται ελαφρώς πιο ουδέτεροι.

2. ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΝΑΛΥΣΗΣ ΣΥΝΑΙΣΘΗΜΑΤΟΣ

2.1 ΑΠΟΔΟΣΗ ΤΩΝ ΔΥΟ ΜΕΘΟΔΩΝ



Γράφημα 22

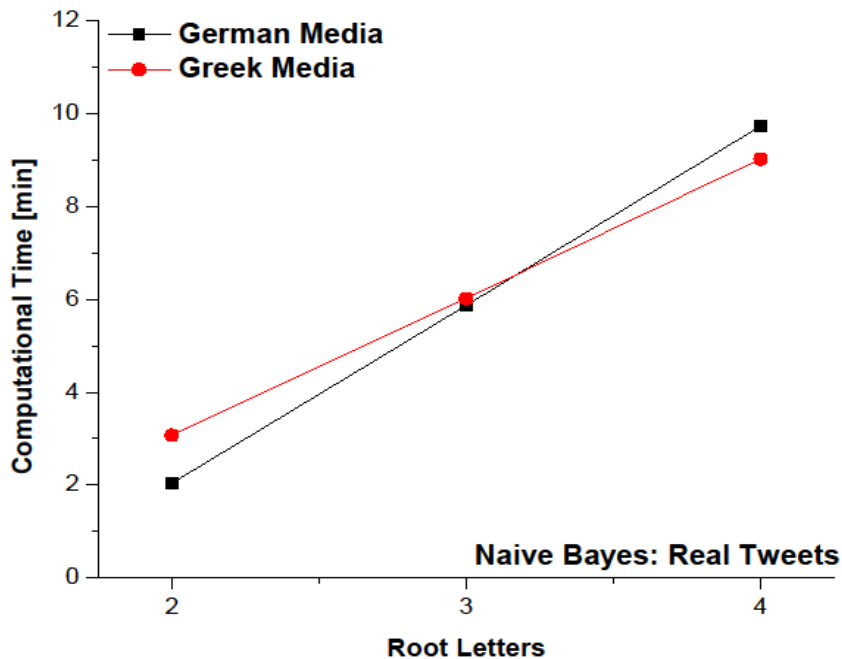
Method	Naive Bayes	Lexicon
Target Group	Total Success Rate [%]	
Greek Media	9.81	81.53
Greek Users	18.95	81.38
German Media	21.03	80.93
German Users	44.17	64.53

Πίνακας 9. Σύγκριση μεθόδων ανάλυσης συναισθήματος

Εφαρμόζοντας και τις δύο μεθόδους ανάλυσης συναισθήματος μπορεί να παρατηρήσει κανείς ότι η μέθοδος που βασίστηκε στο λεξικό παρουσιάζει σε μεγάλο βαθμό επιτυχία, σε αντίθεση με τον αλγόριθμο Naive Bayes, ο οποίος παρουσιάζει μικρά ποσοστά. Απαντώντας στο 4^ο ερευνητικό ερώτημα, η μέθοδος του λεξικού προέβη σε μεγάλο βαθμό σε σωστή κατηγοριοποίηση του συναισθήματος, με βάση πάντα την χειρωνακτική επισημείωση που έγινε και γι' αυτό στη συνέχεια χρησιμοποιήθηκε και για να προβλέψει το συναίσθημα για κάθε μία κατηγορία των opinion targets ξεχωριστά. Οι λόγοι της χαμηλής επιτυχίας του αλγορίθμου Naive Bayes θα αναλυθούν παρακάτω.

2.1.1 ΑΠΟΔΟΣΗ ΤΟΥ ΑΛΓΟΡΙΘΜΟΥ NAIVE BAYES

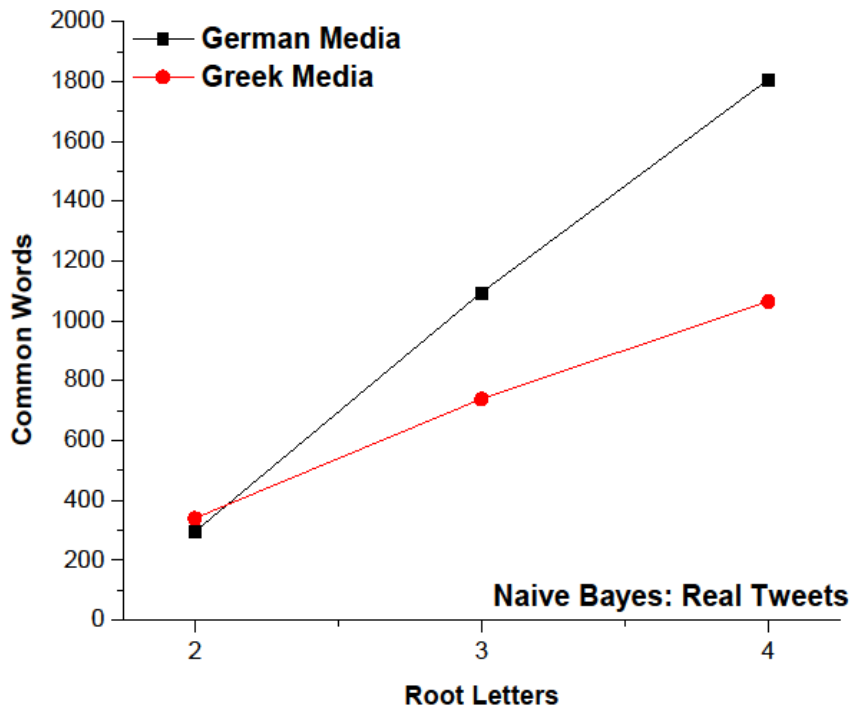
- Πραγματικά tweets



Γράφημα 23

Το παραπάνω συγκριτικό διάγραμμα απεικονίζει τον τρόπο εξάρτησης του υπολογιστικού χρόνου σε σχέση με το πλήθος των πρώτων γραμμάτων (που συνθέτουν τις κοινές ρίζες) των αρχικών συνολικών λέξεων στην περίπτωση των ΜΜΕ. Στην περίπτωση των ελληνικών ΜΜΕ ο χρόνος φαίνεται να εξαρτάται γραμμικά από το μέγεθος της ρίζας, δηλαδή αύξηση των γραμμάτων της ρίζας οδηγεί σε αύξηση του απαιτούμενου χρόνου με την ίδια αναλογία. Στην περίπτωση των γερμανικών ΜΜΕ ο χρόνος φαίνεται να αυξάνει με γρηγορότερο ρυθμό όταν τα γράμματα της ρίζας ξεπεράσουν τα δύο. Είναι προφανές ότι ο συγκεκριμένος τρόπος εξάρτησης είναι επακόλουθος του πλήθους των δυνατών κοινών ριζών που παράγονται με την αύξηση του μεγέθους της ρίζας, δηλαδή η

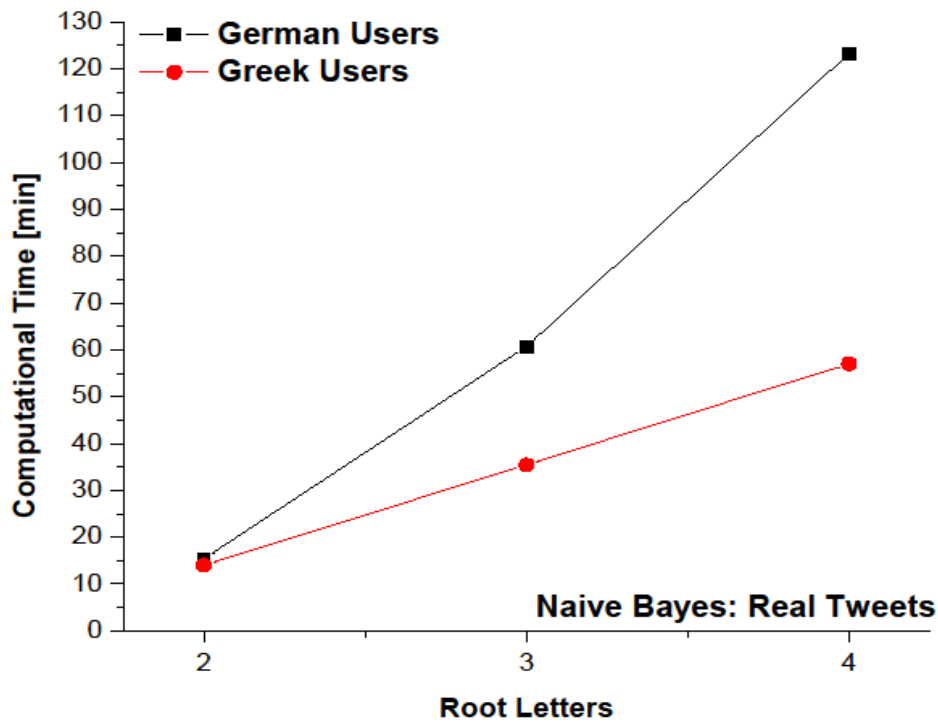
αύξηση του πλήθους των λέξεων είναι εκείνη που οδηγεί στην αύξηση του υπολογιστικού χρόνου.



Γράφημα 24

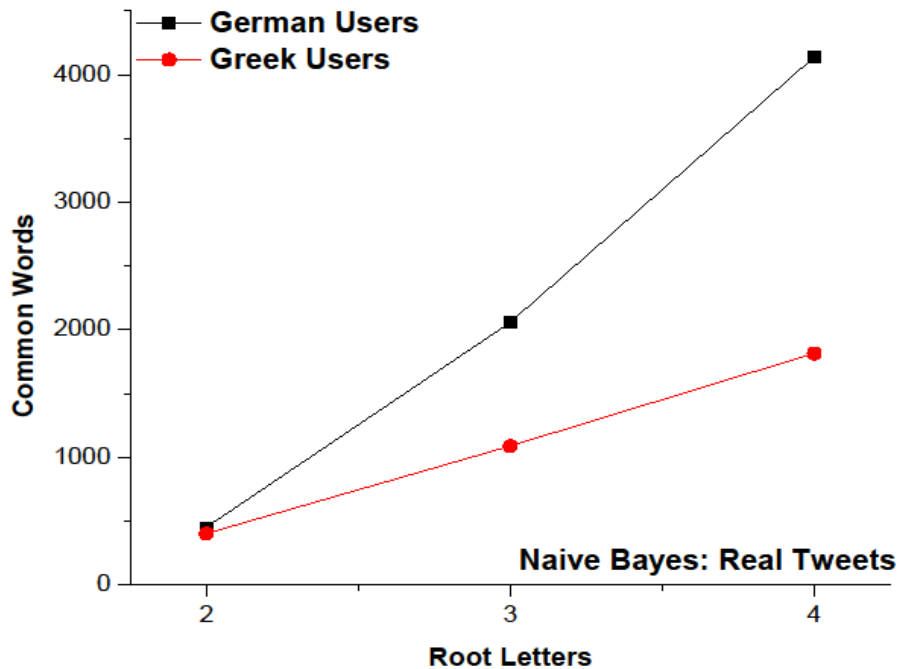
Το εν λόγω συγκριτικό διάγραμμα απεικονίζει το πλήθος των κοινών ριζών που παράγονται σε σχέση με το πλήθος των πρώτων γραμμάτων (που συνθέτουν τις κοινές ρίζες) των αρχικών συνολικών λέξεων στην περίπτωση των ΜΜΕ. Στην περίπτωση των ελληνικών ΜΜΕ, ο ρυθμός αύξησης του πλήθους λέξεων φαίνεται να είναι ανάλογος με την αύξηση των γραμμάτων, ενώ στην περίπτωση των γερμανικών ΜΜΕ, ο ρυθμός αύξησης φαίνεται να είναι δυσανάλογα μεγαλύτερος μόλις τα γράμματα της ρίζας ξεπεράσουν τα δύο. Αίτιο της δυσαναλογίας αυτής είναι ο σαφώς μεγαλύτερος αριθμός αρχικών λέξεων που ανιχνεύθηκαν στα tweets των γερμανικών μέσων (3404) σε σχέση με τα ελληνικά (1862). Το γεγονός αυτό μπορεί εν μέρει να αποδοθεί στην αυξημένη αναπαραγωγή tweets (retweets) που ενδεχομένως παρατηρείται στον ελληνικό ηλεκτρονικό τύπο και οδηγεί σε μικρότερο κοινό αριθμό λέξεων. Στο σημείο αυτό να αναφερθεί ότι το πλήθος λέξεων για κάθε αριθμό γραμμάτων είναι

σωρευτικός, δηλαδή προσαυξάνεται κάθε φορά στο πλήθος λέξεων των προηγούμενων (μικρότερων) κοινών ριζών (με λίγα λόγια το πλήθος ριζών για τα τέσσερα πρώτα γράμματα περιλαμβάνει και τις ρίζες ενός, δύο και τριών γραμμάτων).



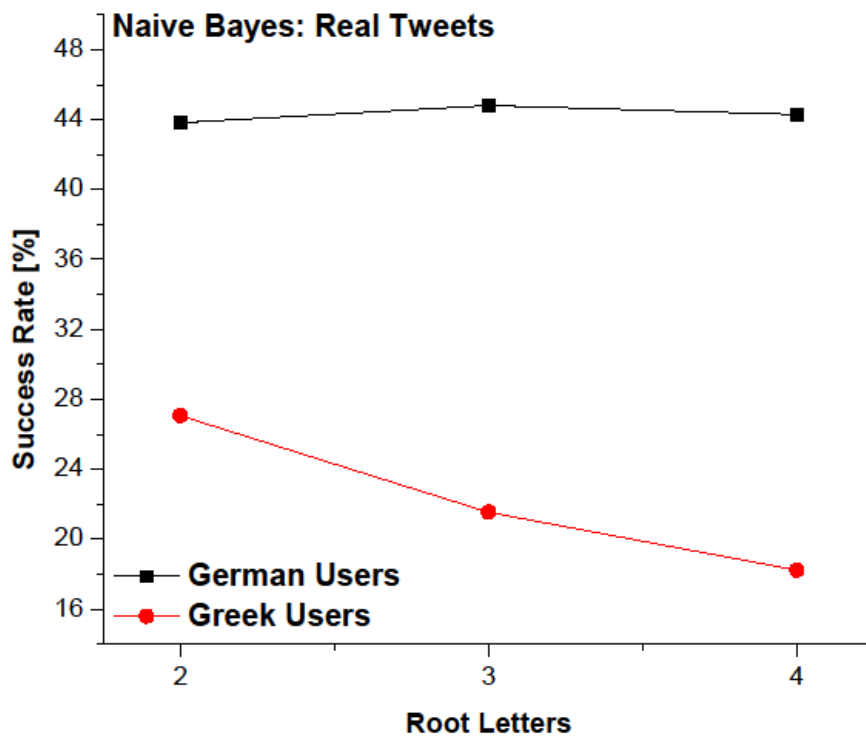
Γράφημα 25

Το συγκριτικό διάγραμμα στην προκειμένη περίπτωση απεικονίζει τον τρόπο εξάρτησης του υπολογιστικού χρόνου σε σχέση με το πλήθος των πρώτων γραμμάτων (που συνθέτουν τις κοινές ρίζες) των αρχικών συνολικών λέξεων στην περίπτωση των ιδιωτών χρηστών. Η περιγραφή και η ερμηνεία είναι όμοια με εκείνη των ΜΜΕ (γράφημα 23) με την πρόσθετη παρατήρηση της ραγδαίας πλέον αύξησης του απαιτούμενου πλέον υπολογιστικού χρόνου μετά τα δύο πρώτα γράμματα. Είναι προφανές ότι η διαπίστωση αυτή συνδέεται με το αδιαμφισβήτητο μεγάλο πλήθος γερμανικών λέξεων (10555) που προέκυψαν εξαιτίας του μεγαλύτερου διαθέσιμου δείγματος σε σχέση με τις ελληνικές (3718) και συνεπώς του αναλόγως μεγαλύτερου πλήθους κοινών ριζών που χρησιμοποιήθηκαν ως μεταβλητές εισόδου.



Γράφημα 26

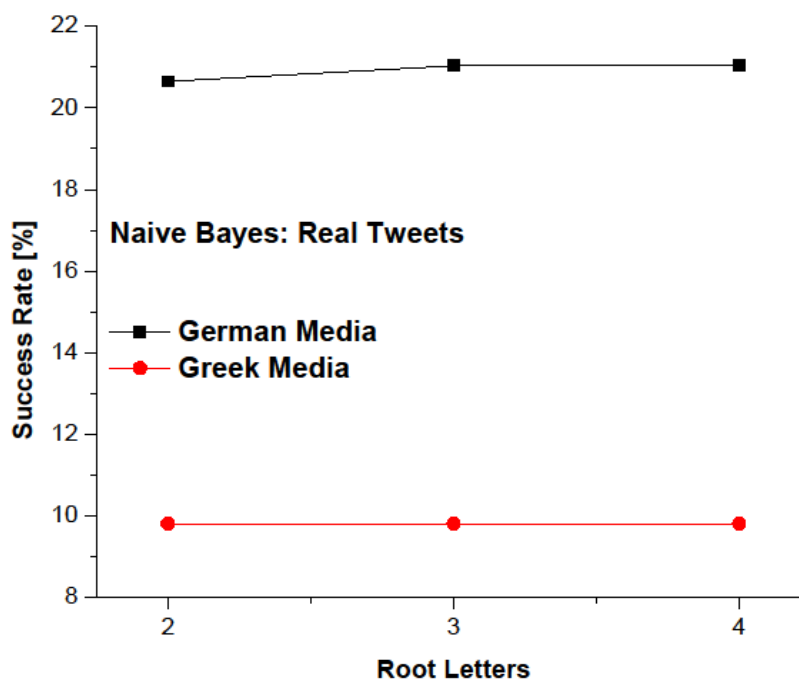
Το παραπάνω συγκριτικό διάγραμμα απεικονίζει το πλήθος των κοινών ριζών που παράγονται σε σχέση με το πλήθος των πρώτων γραμμάτων (που συνθέτουν τις κοινές ρίζες) των αρχικών συνολικών λέξεων στην περίπτωση των ιδιωτών χρηστών. Η περιγραφή και η ερμηνεία είναι όμοια με εκείνη των ΜΜΕ (γράφημα 24) με την πρόσθετη παρατήρηση της ραγδαίας πλέον αύξησης των κοινών λέξεων μετά τα δύο πρώτα γράμματα για τους γερμανούς χρήστες. Όπως ήδη εξηγήθηκε κατά την ερμηνεία του γραφήματος 24, η διαπίστωση αυτή συνδέεται με το μεγάλο πλήθος γερμανικών λέξεων (10555) που προέκυψαν εξαιτίας του μεγαλύτερου διαθέσιμου δείγματος σε σχέση με τις ελληνικές (3718).



Γράφημα 27

Το συγκεκριμένο συγκριτικό διάγραμμα απεικονίζει την απόδοση του αλγόριθμου σε σχέση με το πλήθος των πρώτων γραμμάτων (που συνθέτουν τις κοινές ρίζες) των αρχικών συνολικών λέξεων στην περίπτωση των ιδιωτών χρηστών. Η αύξηση της απόδοσης του αλγόριθμου είναι φανερή αγγίζοντας το 45% στην περίπτωση των γερμανών χρηστών, διαπίστωση πλήρως συμβατή με την αξιολόγησή του κάτω από ελεγχόμενες συνθήκες, η οποία έδειξε ότι μεγιστοποίηση της απόδοσης επιτυγχάνεται μετά τα 10000 περίπου tweets (εδώ είχαμε 8499 γερμανικά και 9764 ελληνικά). Όπως αναμενόταν, στα ελληνικά tweets η απόδοση αυξάνεται με τη μείωση των πρώτων γραμμάτων και συνεπώς τη μείωση των κοινών ριζών, καθώς πολλές λέξεις αποδεικνύεται ότι είχαν ακριβώς το ίδιο νόημα κάτω από την ίδια ρίζα αλλά διαφορετική δομή (π.χ. γένος, κλίση κ.α.). Η διαπίστωση αυτή δε φαίνεται να ισχύει στην περίπτωση των γερμανικών tweets, των οποίων οι ανιχνευόμενες λέξεις φαίνεται να είναι ανεξάρτητες και όχι με κοινό νόημα, δηλαδή ενδεχόμενη κοινή

ρίζα δεν εξασφαλίζει την ίδια ερμηνεία των λέξεων. Επίσης, ειδικά η διερεύνηση των λέξεων-μεταβλητών στα ελληνικά tweets αποκαλύπτει ότι μέρος της ασθενούς απόδοσης οφείλεται στον ελλιπή καθαρισμό από ανεπιθύμητους χαρακτήρες. Πέραν των 24 ελληνικών γραμμάτων, βρέθηκαν πλήθος αγγλικών αλλά και 17 ακόμα εκ των οποίων σημαντικό μέρος αφορούσε διαφόρων ειδών ομοιωματικά, σημεία στίξης και παρενθέσεις (π.χ. «, », ’, ’, -,] κ.α.), ενώ οι υπόλοιποι ήταν μη αναγνωρίσιμοι κατά την ανάγνωση των δεδομένων στο πρόγραμμα που εκτελέστηκε.



Γράφημα 28

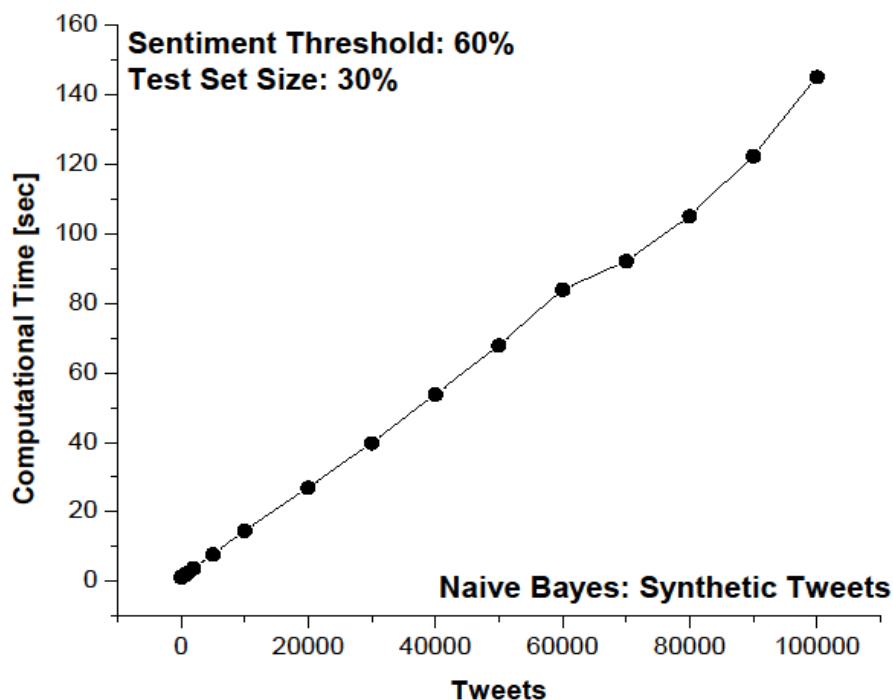
Το παραπάνω συγκριτικό διάγραμμα απεικονίζει την απόδοση του αλγόριθμου σε σχέση με το πλήθος των πρώτων γραμμάτων (που συνθέτουν τις κοινές ρίζες) των αρχικών συνολικών λέξεων στην περίπτωση των MME. Τόσο στην περίπτωση των γερμανικών όσο και των ελληνικών MME η απόδοση εμφανίζεται χαμηλή αλλά και σταθερή, ανεξάρτητη δηλαδή του μεγέθους της κοινής ρίζας. Εν μέρει η αδυναμία αυτή μπορεί να αποδοθεί στο μικρό μέγεθος του εξεταζόμενου δείγματος (1743 γερμανικά tweets και 2448 ελληνικά). Οι επιδόσεις της τεχνικής έχουν άμεση εξάρτηση από το μέγεθος του δείγματος εκπαίδευσης, αλλά και της πολυπλοκότητας των αλληλεπιδράσεων μεταξύ των

μεταβλητών εισόδου (εδώ είναι οι λέξεις-ρίζες) και της μεταβλητής εξόδου-στόχου (εδώ η κλάση συναισθήματος). Η αξιολόγηση του αλγορίθμου με τη βοήθεια της τεχνικής Monte Carlo κάτω από ελεγχόμενες συνθήκες έδειξε ότι με 30 μόνο λέξεις και απλούς κανόνες χαρακτηρισμού συναισθήματος, η επίδοση μεγιστοποιείται μετά τα 10000 περίπου συνθετικά tweets.

- **Συνθετικά tweets**

- **Επίδραση μεγέθους δείγματος στον υπολογιστικό χρόνο**

Κατά την εκτέλεση των δοκιμών στις οποίες υποβλήθηκε ο αλγόριθμος κατασκευάστηκαν από 50 μέχρι 100000 συνθετικά tweets. Με τη βοήθεια χρονομέτρησης την οποία παρέχει η R, **αναδείχθηκε η γραμμική μεταβολή του απαιτούμενου υπολογιστικού χρόνου**, τουλάχιστον για το εύρος συνθετικών tweets που αναφέρθηκε, εφόσον πάντοτε όλοι οι πιθανοί παράγοντες που θα μπορούσαν να τον επηρεάσουν ήταν σταθεροποιημένοι. Μάλιστα προέκυψε η σχέση $\text{CPU Time [sec]} = 0.00137 \times \text{Tweets Number}$ με τη βοήθεια της οποίας είναι δυνατή η εκτίμηση του χρόνου για έναν συγκεκριμένο αριθμό tweets κάτω από τις ίδιες συνθήκες και το ίδιο υπολογιστικό σύστημα.



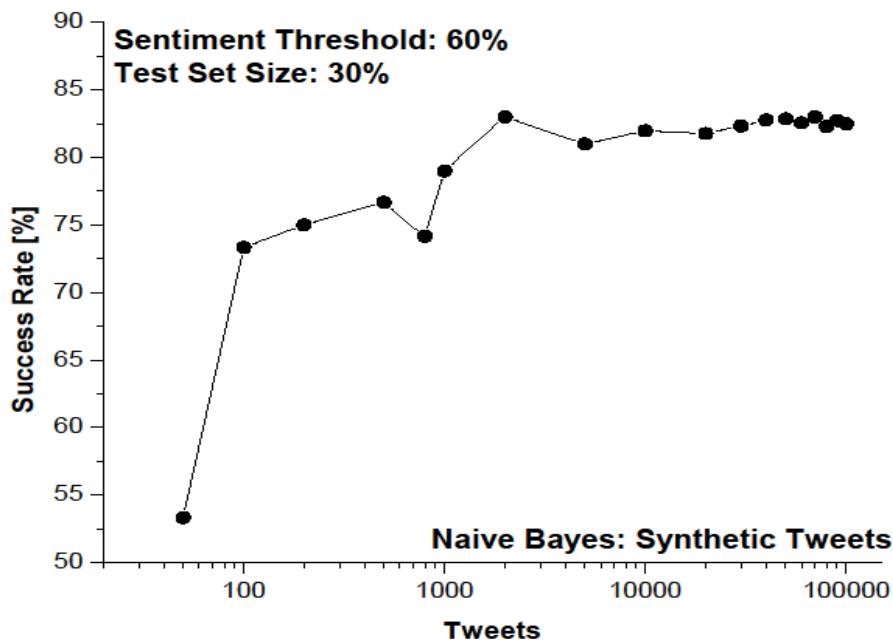
Γράφημα 29

Ας σημειωθεί ότι η συγκεκριμένη δοκιμασία πραγματοποιήθηκε κάτω από ελεγχόμενες συνθήκες, δηλαδή συγκεκριμένο κανόνα χαρακτηρισμού

συναισθήματος (εδώ «θετικός» χαρακτηρισμός αποδόθηκε όταν το 60% των λέξεων του εκάστοτε tweet αφορούσε λέξεις με θετικό νόημα, «αρνητικός» χαρακτηρισμός με το ακριβώς αντίστροφο κριτήριο αρνητικών λέξεων και «ουδέτερος» όταν κανένα από τα προηγούμενα δύο κριτήρια δεν επαληθευόταν), συγκεκριμένο μέγεθος δείγματος εκπαίδευσης (εδώ το 70% του συνολικού, δηλαδή τέθηκε το 30% ως δείγμα ελέγχου).

➤ **Επίδραση μεγέθους δείγματος στο συντελεστή απόδοσης**

Διερευνήθηκε η απόδοση του αλγόριθμου με βάση έναν απλό κανόνα επισημείωσης και την ίδια σταθερή αναλογία μεγέθους δείγματος εκπαίδευσης – ελέγχου. Ο κανόνας απέδιδε «θετικό» χαρακτηρισμό συναισθήματος, όταν το 60% των λέξεων του εκάστοτε tweet αποτελούνταν από λέξεις με θετικό νόημα, «αρνητικό» χαρακτηρισμό με το ακριβώς αντίστροφο κριτήριο αρνητικών λέξεων και «ουδέτερο» όταν κανένα από τα προηγούμενα δύο κριτήρια δεν επαληθευόταν. Κατά την εκτέλεση των δοκιμών μεταξύ 50 και 100000 συνθετικών tweets η απόδοση ήταν αρκετά υψηλή, αγγίζοντας ένα συντελεστή μεταξύ 75% στα πρώτα 1000 tweets. Αποκαλύφθηκε όμως ότι μεγιστοποίηση των επιδόσεων επιτεύχθηκε μετά τα 10000 tweets, με το συντελεστή απόδοσης να πλησιάζει το 85% στα 100000 tweets. Το εύρημα αυτό είναι πολύ σημαντικό, καθώς **με χρήση απλών ελεγχόμενων συνθηκών, ο αλγόριθμος συνεχίζει να απαιτεί ένα πολύ μεγάλο δείγμα ώστε να εξασφαλιστεί βελτιστοποίηση των επιδόσεών του.**



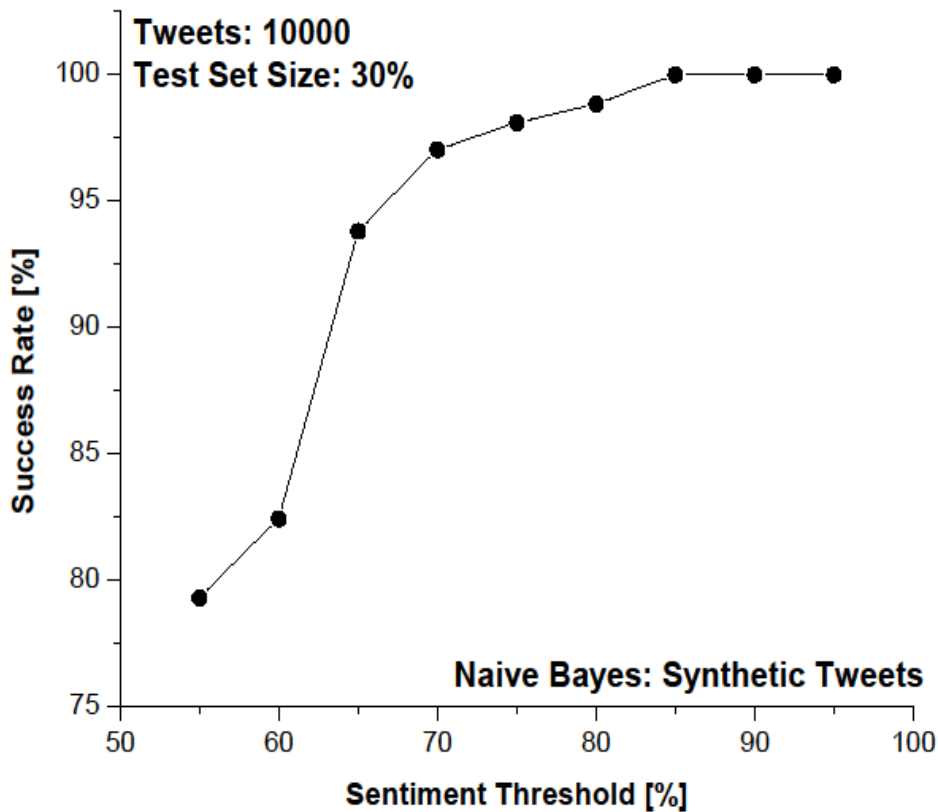
Γράφημα 30

Στο παραπάνω γράφημα, λόγω του πολύ μεγάλου εύρους συνθετικών δειγμάτων (50 - 100000), έχει επιλεγθεί λογαριθμική κλίμακα στον (οριζόντιο) άξονα που εκφράζει το μέγεθος του δείγματος. Είναι εμφανής η συνεχής αύξηση της απόδοσης με την αύξηση του δείγματος με σταθεροποίηση κοντά στο 85% μετά τα 10000 συνθετικά tweets.

➤ Επίδραση βαθμού πολυπλοκότητας ταξινόμησης στο συντελεστή απόδοσης

Διερευνήθηκε η απόδοση του αλγόριθμου διαφοροποιώντας κάθε φορά τον κανόνα επισημείωσης, αλλά διατηρώντας πάντοτε σταθερή την αναλογία μεγέθους δείγματος εκπαίδευσης - ελέγχου και το πλήθος των συνθετικών tweets (ρυθμίστηκε στα 10000, σύμφωνα με τα ευρήματα του προηγούμενου ελέγχου). Η δομή του κανόνα επισημείωσης παρέμεινε ίδια, αφέθηκε όμως ελεύθερο το κατώφλι (threshold) μετάβασης μεταξύ θετικού, αρνητικού και ουδέτερου χαρακτηρισμού. Συγκεκριμένα, αξιολογήθηκε η επίδραση του ποσοστιαίου κριτηρίου μεταξύ αριθμού λέξεων με θετικό ή αρνητικό νόημα έναντι του συνολικού πλήθους λέξεων. **Αποκαλύφθηκε έτσι η κρισιμότητα**

του βαθμού αυστηρότητας που διακατέχει τον κριτή κατά την επισημείωση των tweets με έναν «ελαστικό» κριτή να περιορίζει τις προγνωστικές επιδόσεις του αλγόριθμου και έναν «απαιτητικό» κριτή να ευνοεί μια σαφώς πιο αποτελεσματική λειτουργία. Ένα χαλαρό κριτήριο με κατώφλι μεταξύ 55-60%, το οποίο δηλαδή αποδίδει χαρακτηρισμό με οριακή υπεροχή των λέξεων του αντίστοιχου προσανατολισμού, οδήγησε σε συντελεστή απόδοσης μικρότερο του 85%. Αντίθετα, ένα αυστηρό κριτήριο με κατώφλι μεταξύ μεγαλύτερο του 80%, το οποίο δηλαδή αποδίδει χαρακτηρισμό με συντριπτική υπεροχή των λέξεων του αντίστοιχου προσανατολισμού, οδήγησε σε συντελεστή απόδοσης μεγαλύτερο του 95%. Το υψηλό κατώφλι οδηγεί πολύ εύκολα στον «ουδέτερο» χαρακτηρισμό και στην απουσία τόσο του «θετικού» όσο και του «αρνητικού». Ο αλγόριθμος συνεπώς έχει πολύ εύκολο προγνωστικό έργο καθώς τα ενδεχόμενα είναι πλέον πολύ περιορισμένα.



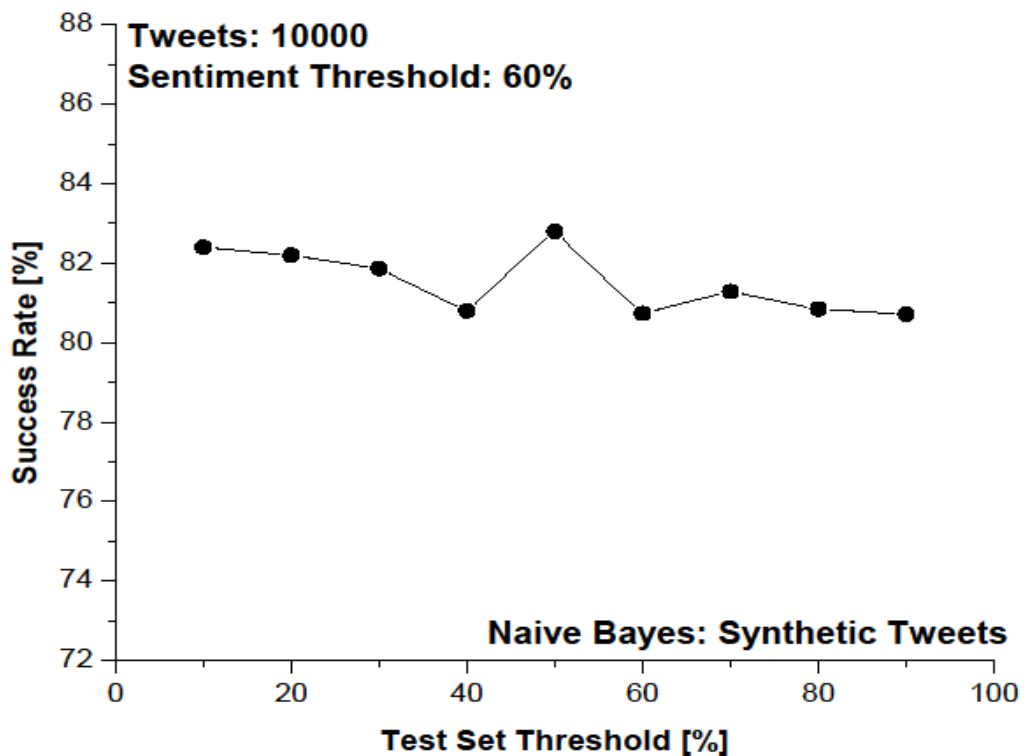
Γράφημα 31

Υπενθυμίζεται ότι τα συνθετικά tweets κατασκευάστηκαν από αναδειγματοληψία ισάριθμων προεπιλεγμένων λέξεων με «θετικό» ή «αρνητικό» χαρακτήρα (15+15) και είχαν όλα τους κοινό μέγεθος (30 λέξεων), χωρίς την εμπλοκή καμίας άλλης λέξης πέραν των 30 αρχικών.

➤ **Επίδραση Αναλογίας Μεγέθους Δείγματος Εκπαίδευσης-Ελέγχου στο Συντελεστή Απόδοσης**

Διερευνήθηκε η απόδοση του αλγόριθμου διαφοροποιώντας κάθε φορά το μέγεθος του δείγματος εκπαίδευσης σε σχέση με εκείνο του δείγματος ελέγχου διατηρώντας σταθερό το μέγεθος του δείγματος και τον κανόνα επισημείωσης (τα συνθετικά tweets παρέμειναν στα 10000 και το ποσοστιαίο κατώφλι ρυθμίστηκε στο 60%). **Αποδείχθηκε η ασθενής επίδραση του συγκεκριμένου παράγοντα στην απόδοση του αλγορίθμου, καθώς ένα αρκετά μεγάλο**

δείγμα (όπως π.χ. εκείνο των 10000 tweets) εξασφαλίζει επαρκή εκμάθηση ανεξάρτητα από την αναλογία δειγμάτων εκπαίδευσης-ελέγχου. Μια αναλογία 10-90% υπέρ του δείγματος εκπαίδευσης οδηγεί σε επιδόσεις της τάξης του 80-85%, ενώ μια αναλογία 90-10% υπέρ του δείγματος ελέγχου συνεχίζει να διατηρεί το συντελεστή απόδοσης εξίσου σε υψηλά επίπεδα, μεγαλύτερα του 80%.

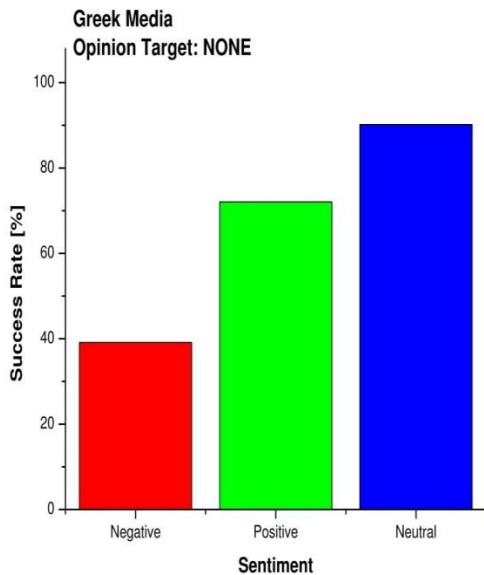


Γράφημα 32

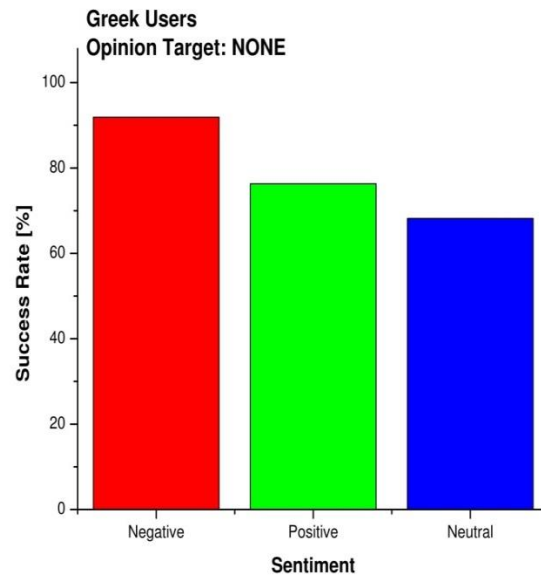
Από το παραπάνω γράφημα είναι εμφανές ότι ο παράγοντας αυτός δεν είναι κρίσιμος, καθώς η απόδοση διατηρείται σταθερή ανεξάρτητα από το συγκεκριμένο μέγεθος, σημειώνοντας μια ανεπαίσθητη πτώση, καθώς το δείγμα εκπαίδευσης περιορίζεται όπως άλλωστε είναι αναμενόμενο.

2.1.2 ΑΠΟΔΟΣΗ ΤΗΣ LEXICON- BASED ΜΕΘΟΔΟΥ

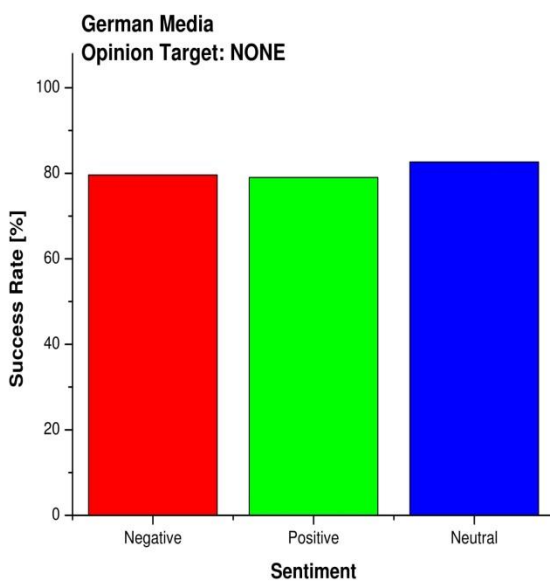
Στη συνέχεια παρατίθενται οι επιδόσεις της μεθόδου βασισμένης σε λεξικό και συγκεκριμένα το κατά πόσο κατηγοριοποιήθηκε σωστά το συναίσθημα ανά πολικότητα (θετικό/αρνητικό/ουδέτερο), τόσο στο σύνολο των tweets, όσο και ανά στόχο συναισθήματος.



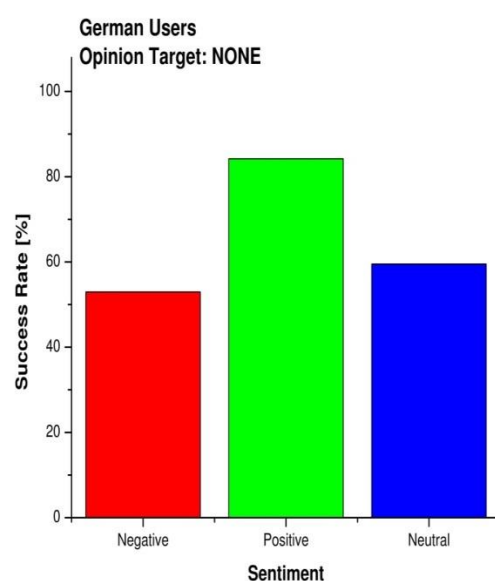
Γράφημα 33



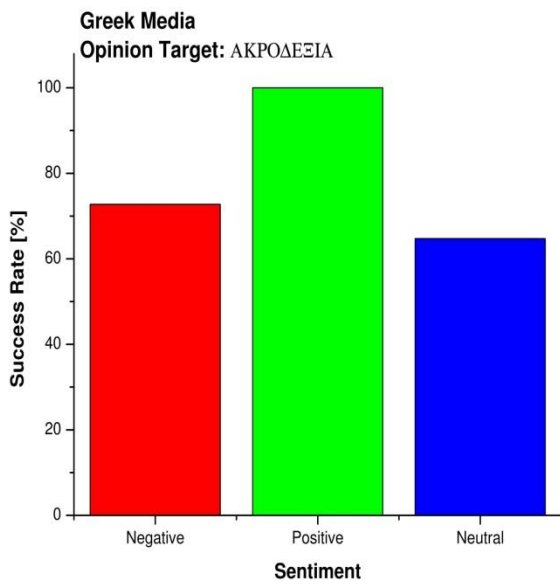
Γράφημα 34



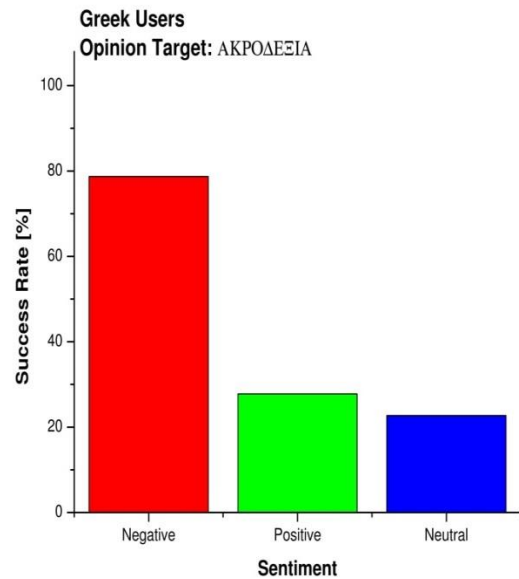
Γράφημα 35



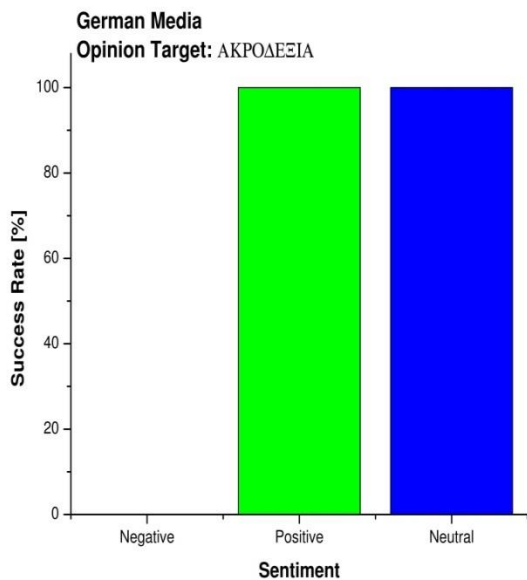
Γράφημα 36

ΑΚΡΟΔΕΞΙΑ

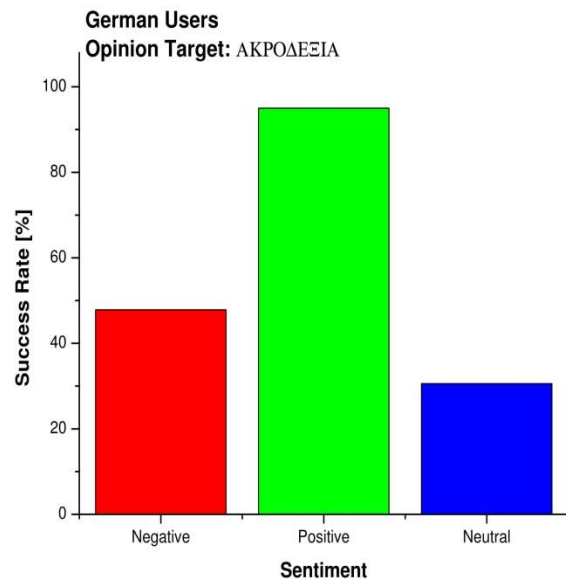
Γράφημα 37



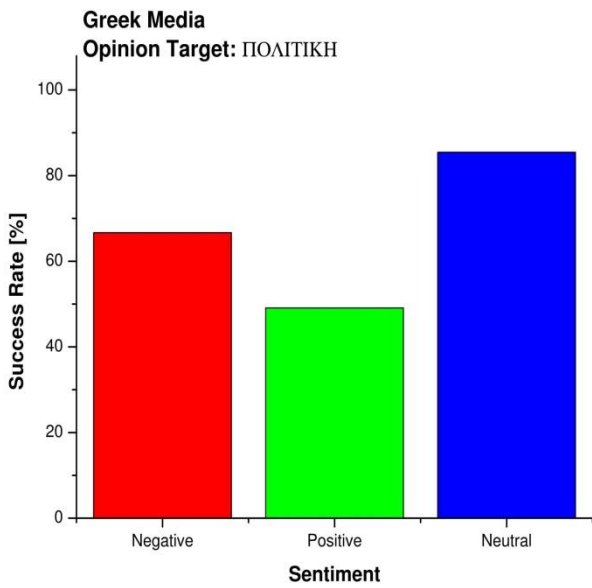
Γράφημα 38



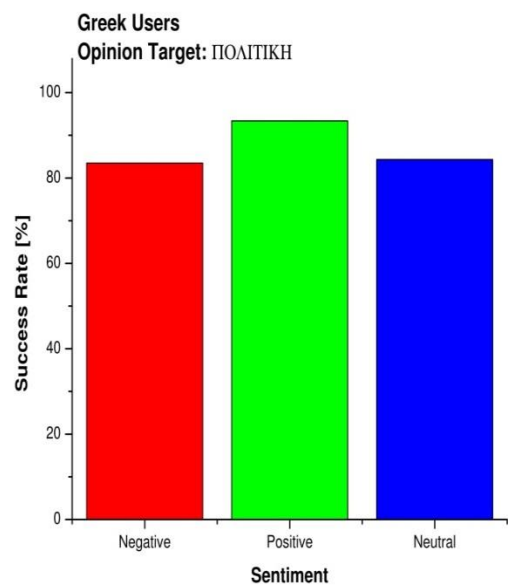
Γράφημα 39



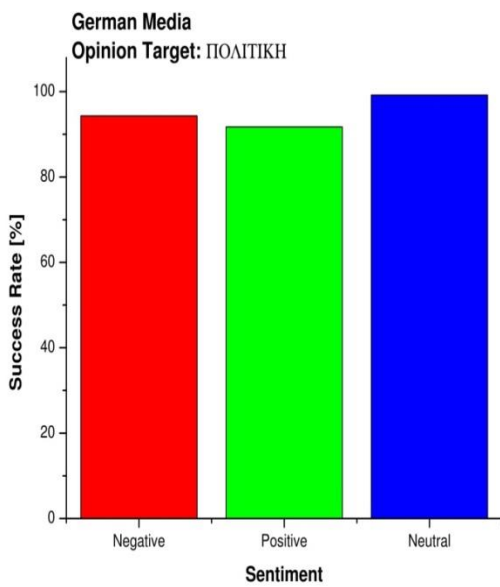
Γράφημα 40

ΠΟΛΙΤΙΚΗ

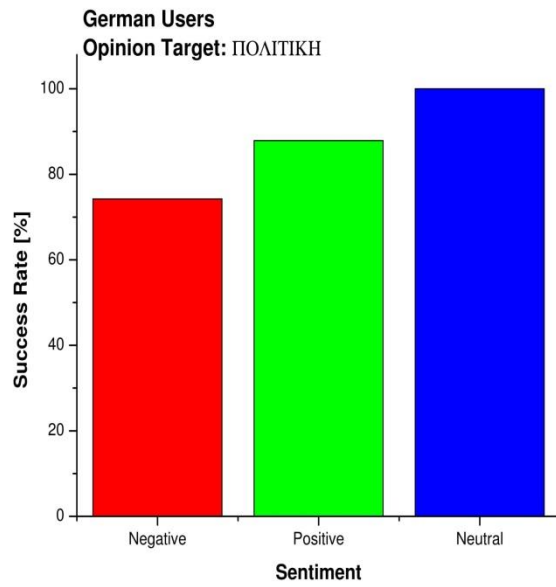
Γράφημα 41



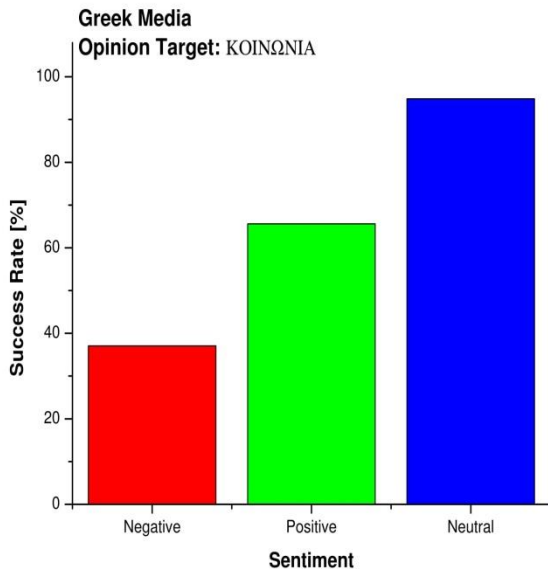
Γράφημα 42



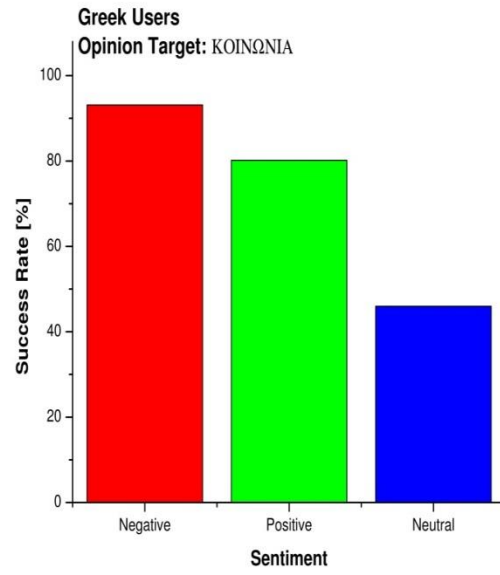
Γράφημα 43



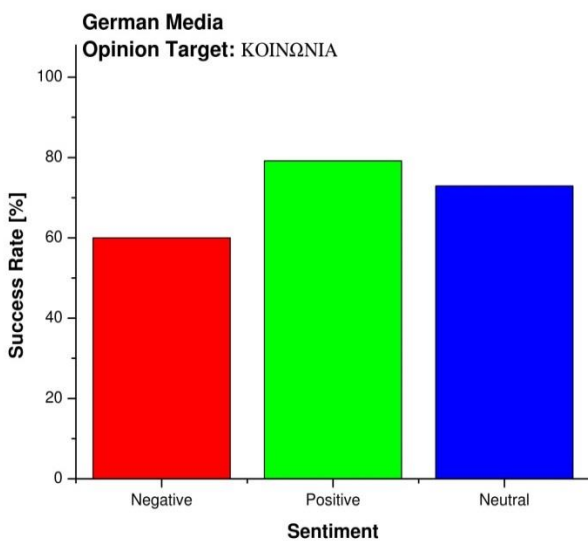
Γράφημα 44

KOINΩNIA

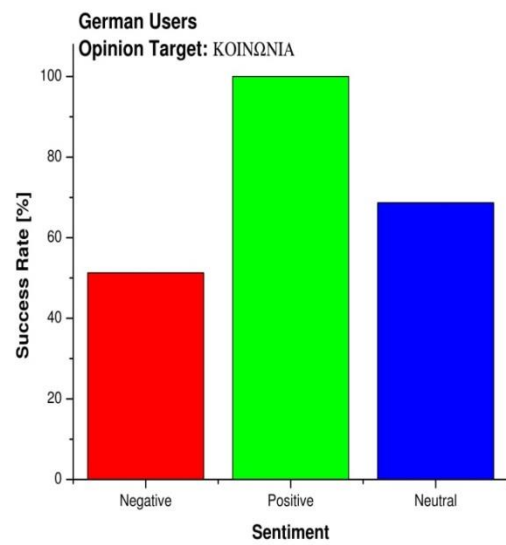
Γράφημα 45



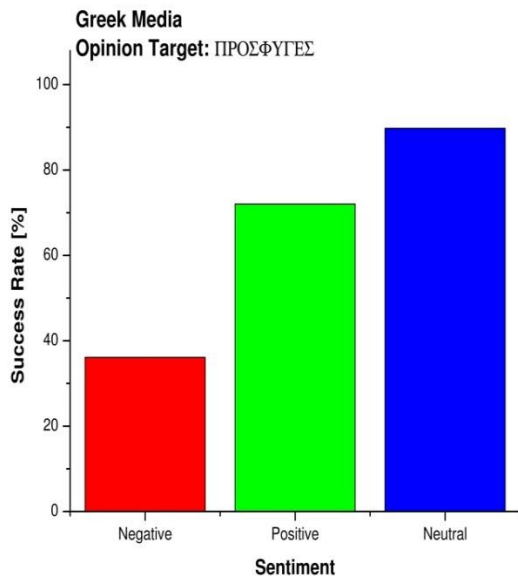
Γράφημα 46



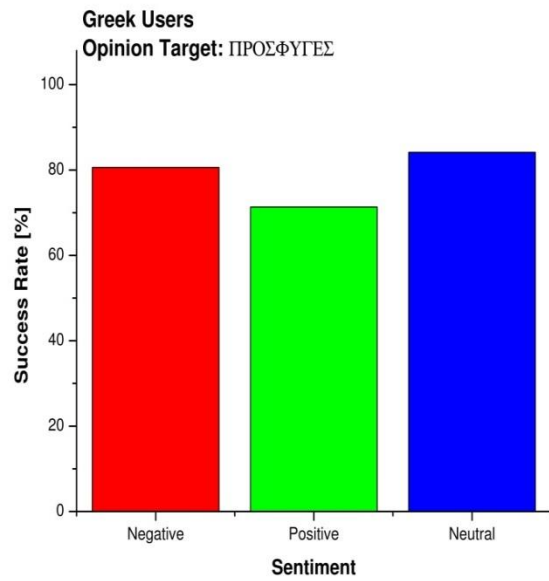
Γράφημα 47



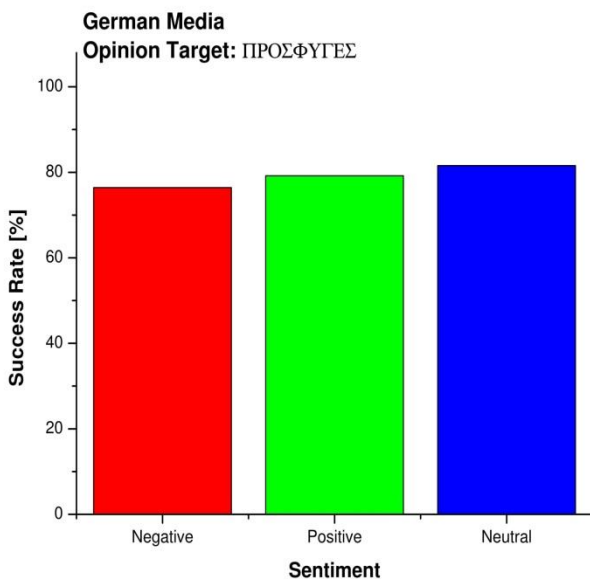
Γράφημα 48

ΠΡΟΣΦΥΓΕΣ

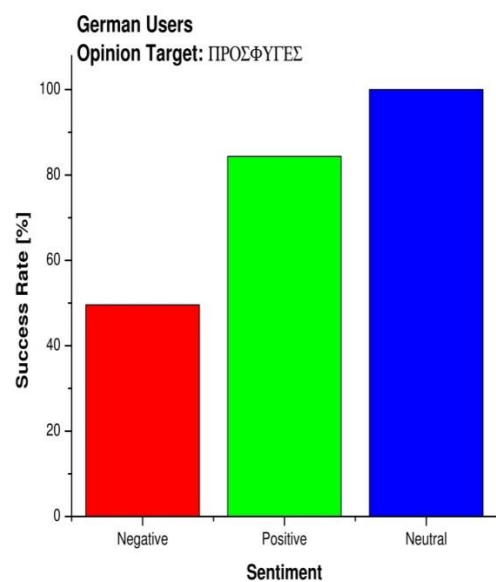
Γράφημα 49



Γράφημα 50



Γράφημα 51



Γράφημα 52

Με βάση τα παραπάνω γραφήματα, καθώς και τους αναλυτικούς πίνακες που βρίσκονται στο παράρτημα της παρούσας εργασίας, μπορεί να απαντήσει κανείς στο 5^ο ερευνητικό ερώτημα. Η μέθοδος του λεξικού φαίνεται πως προβλέπει με μεγαλύτερη ακρίβεια το θετικό συναίσθημα, κατόπιν το ουδέτερο και πιο δύσκολα το αρνητικό. Αυτό ενδεχομένως να οφείλεται στο ότι δεν εφαρμόστηκε κάποια μέθοδος που να λαμβάνει υπόψη την άρνηση σε μία πρόταση, όπως για παράδειγμα την παρουσία μορίων τύπου «δεν» και «μη». Αυτό θα μπορούσε να αντιμετωπιστεί με το να γίνεται αντιστροφή της πολικότητας μία πρότασης, όταν εντοπίζονται τα αρνητικά μόρια. Παρόλα αυτά η μέθοδος επέδειξε μεγάλο ποσοστό επιτυχίας, γεγονός που έχει να κάνει με την επέκταση των λεξικών με τις θετικές και αρνητικές λέξεις. Εάν το λεξικό που χρησιμοποιείται σε μία τέτοια έρευνα δεν περιλαμβάνει λέξεις που να σχετίζονται με το εκάστοτε θέμα, παρά χρησιμοποιεί ένα γενικό sentiment λεξικό, τα ποσοστά επιτυχίας της μεθόδου θα είναι κατά πολύ μειωμένα.

2.2 ΣΥΓΚΡΙΣΗ ΜΕΘΟΔΩΝ

Χωρίς καμία αμφιβολία, η μέθοδος του «λεξικού» παρουσιάζεται περισσότερο αποτελεσματική από εκείνη της «Naive Bayes», όσον αφορά στα πραγματικά tweets που μελετήθηκαν από τα ελληνικά και γερμανικά κοινωνικά μέσα δικτύωσης. Η πρώτη έφτανε εύκολα σε επιδόσεις της τάξης του 85%, όταν η δεύτερη δεν μπόρεσε να ξεπεράσει το 50% μετά από πλήθος τροποποιήσεων. Στηριζόμενοι στα ευρήματα της αξιολόγησης της δεύτερης μεθοδολογίας μέσω συνθετικών tweets και η οποία κατέδειξε τις υψηλές της επιδόσεις σε ελεγχόμενο περιβάλλον, οι οποίες ήταν μάλιστα ανάλογες ή υψηλότερες της πρώτης μεθόδου, κρίνεται απαραίτητο να διερευνηθούν στη συνέχεια τα πιθανά αίτια που οδήγησαν στην περιορισμένη αποτελεσματικότητά της σε πραγματικές συνθήκες.

▪ Αίτια υπερεκτίμησης μεθόδου «λεξικού»

- i. Στη μέθοδο του λεξικού, η καταμέτρηση επιτυχιών, δηλαδή των γεγονότων ταύτισης επισημείωσης και εκτιμώμενου χαρακτηρισμού συναισθήματος, απαντά μόνο στα tweets εκείνα που οι λέξεις- κλειδιά, αυτές δηλαδή που περιλαμβάνονται στα λεξικά, εντοπίζονται. Πρόκειται για ένα γνώρισμα της μεθόδου, το οποίο υπερεκτιμά εκ των προτέρων τις επιδόσεις, καθώς η μεθοδολογία αξιολογείται στην πράξη μόνο σε tweets που ευνοούν ένα θετικό αποτέλεσμα και όχι σε όλο το δείγμα καθολικά. Αντίθετα, η μέθοδος Naive Bayes, αξιολογείται αξιοκρατικά επί του συνολικού δείγματος με όλα τα μειονεκτήματα που συνεπάγεται. Ας σημειωθεί ότι η αξιολόγηση της μεθόδου του λεξικού έγινε και σε επιμέρους θεματικές ενότητες (πρόσφυγες, ακροδεξιά, κοινωνία, πολιτική) και χρησιμοποιήθηκαν στατιστικά βάρη ανάλογα του πλήθους των tweets που αντιστοιχούσαν στις υποκατηγορίες αυτές με σκοπό τον σταθμισμένο προσδιορισμό ενός ολικού συντελεστή απόδοσης.
- ii. Ο τρόπος χαρακτηρισμού συναισθήματος είναι ιδιαίτερα απλός και ευνοεί ισχυρά την επιτυχή πρόβλεψη της αντίστοιχης επισημείωσης, καθώς στηρίζεται σε ένα αριθμητικό κριτήριο υπεροχής λέξεων με θετικό ή αρνητικό

χαρακτήρα, προεπιλεγμένες όμως μέσα από το λεξικό και αυστηρά μόνο σε όσα tweets εντοπίστηκαν οι συγκεκριμένες λέξεις. Η προσέγγιση αυτή προσομοιάζει σε σημαντικό βαθμό τον τρόπο αξιολόγησης του κριτή και συνεπώς αναμένεται να έχει θετική έκβαση τις περισσότερες φορές. Με λίγα λόγια, ο μηχανικός τρόπος ταξινόμησης της μεθόδου του λεξικού βρίσκεται πολύ κοντά στον ανθρώπινο, τον φυσικό τρόπο ταξινόμησης. Στη μέθοδο Naive Bayes, οι κανόνες δεν είναι γνωστοί, ο αλγόριθμος είναι υπεύθυνος να εντοπίσει και να αξιοποιήσει όχι μόνο τις επεξηγηματικές μεταβλητές (λέξεις) με υψηλό αντίκτυπο στη μεταβλητή στόχο (χαρακτηρισμό συναισθήματος), αλλά και τους τρόπους και κανόνες συσχέτισης μεταξύ τους, γεγονός το οποίο τον καθιστά λιγότερο ανταγωνιστικό σε σχέση με την προσέγγιση του λεξικού.

- iii. Ενδεχόμενο ανεπαρκές φιλτράρισμα από ανεπιθύμητους χαρακτήρες κατά τη διαδικασία προ-επεξεργασίας των tweets δεν επηρεάζει την απόδοσή του, σε πλήρη αντίθεση με τον Naive Bayes ο οποίος είναι εξαιρετικά ευαίσθητος από την παράμετρο αυτή. Η εκτίμηση των επιδόσεων στη μέθοδο του λεξικού πραγματοποιείται μόνο στα tweets με λέξεις που αναγνωρίζονται αυτούσιες από το πρόγραμμα.
- iv. Τέλος, το γεγονός ότι πραγματοποιήθηκε επέκταση του λεξικού, το οποίο εμπλουτίστηκε με λέξεις που σχετίζονταν με το προσφυγικό, έπαιξε σημαντικό ρόλο στα υψηλά ποσοστά επιτυχίας που σημείωσε η εν λόγω μέθοδος. Αν είχε γίνει αυτούσια χρήση των λεξικών, δεν θα μπορούσε να εντοπιστεί μεγάλος αριθμός θετικών και αρνητικών λέξεων, από τη στιγμή που αυτές δεν θα εμπεριέχονταν στα λεξικά και έτσι τα ποσοστά επιτυχίας της μεθόδου θα ήταν σαφώς χαμηλότερα.

▪ Αίτια υποεκτίμησης μεθόδου «Naive Bayes»

- i. Η υποκειμενικότητα και η ασαφής τακτική επισημείωσης. Ο χαρακτηρισμός συναισθήματος στα πραγματικά tweets (επισημείωση) στηρίζεται κυρίως σε κριτήρια τα οποία εξαρτώνται από τις προσωπικές πεποιθήσεις, το προσωπικό ύφος, το κοινωνικό και μορφωτικό υπόβαθρο του κριτή. Τα κριτήρια αυτά μάλιστα δεν είναι βέβαιο ότι εφαρμόζονται ισότιμα σε όλα τα

- tweets, καθώς ο κριτής αδυνατεί κάθε φορά να υλοποιεί την ίδια στρατηγική με τον ίδιο αυστηρό τρόπο. Η προσέγγιση αυτή φαίνεται να θυμίζει τον «χαλαρό» κριτή που περιγράφηκε κατά την αξιολόγηση του αλγορίθμου και η οποία αποδείχθηκε ότι δυσχεραίνει τη διαδικασία αποτελεσματικής πρόγνωσης και ταξινόμησης.
- ii. Το μικρό μέγεθος του δείγματος το οποίο εξετάστηκε. Η αξιολόγηση του αλγορίθμου με τη βοήθεια της τεχνικής Monte Carlo έδειξε ότι ο παράγοντας αυτός είναι από τους πλέον κρίσιμους, καθώς και ότι κάτω από απλές ελεγχόμενες συνθήκες και απλούς κανόνες επισημείωσης, 10000 tweets είναι ο ελάχιστος απαιτούμενος αριθμός ώστε ο αλγόριθμος να μεγιστοποιήσει την απόδοσή του. Θυμίζουμε ότι στην περίπτωση των MME, αναλύθηκαν 1743 γερμανικά και 2448 ελληνικά tweets αντίστοιχα, ενώ στην περίπτωση των ιδιωτών χρηστών 8499 γερμανικά και 9764 ελληνικά. Δεδομένης και της πολυπλοκότητας των κανόνων που συνδέουν τον χαρακτηρισμό συναισθήματος με τις λέξεις-μεταβλητές, είναι εμφανές ότι το δείγμα δεν είναι επαρκές.
 - iii. Το μεγάλο πλήθος λέξεων το οποίο, ενώ έχει κοινό προσανατολισμό, διαφέρει στη δομή του, όπως π.χ. λέξεις σε διαφορετική κλίση, γένος κλπ. Πρόκειται στην ουσία για ομόρριζες λέξεις οι οποίες όμως κατά τη διαδικασία εκπαίδευσης αντιμετωπίζονται ως πλήρως αυτόνομες επεξηγηματικές μεταβλητές εισόδου. Οι τροποποιήσεις που πραγματοποιήθηκαν προς την κατεύθυνση αυτή επιβεβαίωσαν την υπόθεση αυτή αλλά μόνο για τα ελληνικά tweets και με βελτίωση η οποία δεν ξεπέρασε το 10% στο συντελεστή απόδοσης. Ενδεχομένως, πολλές γερμανικές λέξεις, λόγω της φύσης της συγκεκριμένης γλώσσας, να έχουν κοινή ρίζα (κοινά αρχικά δηλαδή γράμματα) αλλά διαφορετική ερμηνεία. Ακόμα όμως και στην περίπτωση των ελληνικών tweets, η βελτίωση δεν ήταν ικανοποιητική και φαίνεται ότι υπάρχουν σημαντικότεροι παράγοντες που διαμορφώνουν τη μειωμένη αποτελεσματικότητα του αλγορίθμου.
 - iv. Το μεγάλο πλήθος λέξεων οι οποίες δεν έχουν σημαντικό αντίκτυπο στη μεταβλητή-στόχο, δηλαδή τον χαρακτηρισμό συναισθήματος. Οι λέξεις αυτές έχουν ουδέτερο χαρακτήρα, όπως π.χ. τα βοηθητικά ρήματα «είμαι», «έχω»

- κ.α., αλλά διογκώνουν τον πίνακα λέξεων-κειμένου (tdm) και αποπροσανατολίζουν τη διαδικασία εκπαίδευσης.
- v. Το καθάρισμα των κειμένων από ανεπιθύμητους χαρακτήρες είναι σχετικά ελλιπές. Μετά από σχετική διερεύνηση αποδείχθηκε ότι πέραν των 24 ελληνικών γραμμάτων, εντοπίστηκε πλήθος αγγλικών, αλλά και 17 ακόμα εκ των οποίων σημαντικό μέρος αφορούσε διαφόρων ειδών ομοιωματικά, σημεία στίξης και παρενθέσεις (π.χ. «, », ’, ’, -,] κ.α.), ενώ οι υπόλοιποι ήταν μη αναγνωρίσιμοι κατά την ανάγνωση των δεδομένων στο πρόγραμμα που εκτελέστηκε.
- vi. Η αδυναμία αναγνώρισης του προσανατολισμού συναισθήματος με βάση τον τρόπο διατύπωσης παρά μόνο με τη συχνότητα εμφάνισης συγκεκριμένων λέξεων. Επομένως, όταν λέξεις, οι οποίες φέρουν θετικό χαρακτήρα συνοδεύονται από ορισμένες άλλες βοηθητικές που αντιστρέφουν όμως το νόημα (π.χ. «μη», «δεν» κ.α.), συνεχίζουν να αποδίδουν θετικό προσανατολισμό στον χαρακτηρισμό (και το αντίστροφο). Στην ίδια κατηγορία ανήκει και η «ειρωνική» διατύπωση, της οποίας ο αρνητικός προσανατολισμός δεν μπορεί να αποκωδικοποιηθεί.

3. ΣΥΜΠΕΡΑΣΜΑΤΑ

Στην παρούσα διπλωματική, συλλέχθηκε ένα σύνολο από 22.454 tweets, εκ των οποίων τα 12.212 ήταν στα Ελληνικά και τα 10.242 στα Γερμανικά, αντλημένα από επίσημους λογαριασμούς ΜΜΕ της Ελλάδας και της Γερμανίας, αλλά και από προσωπικούς λογαριασμούς χρηστών. Το σύνολο των tweets χωρίστηκε στα τέσσερα επιμέρους σύνολα, δηλαδή στα tweets των ελληνικών και των γερμανικών media, των Ελλήνων και των Γερμανών χρηστών. Στόχος ήταν η αυτόματη εξαγωγή, κατόπιν χειρωνακτικής επισημείωσης, των στόχων του συναισθήματος (opinion targets) που εκφράζεται στα tweets και η κατηγοριοποίηση του συναισθήματος, προκειμένου να μελετηθούν και στη συνέχεια να εφαρμοστούν δύο αυτόματες μέθοδοι ανάλυσης συναισθήματος, αυτή του λεξικού και αυτή της μηχανικής μάθησης με τη χρήση του αλγορίθμου Naive Bayes.

Συγκεκριμένα, από την ανάλυση περιεχομένου, προέκυψε ότι τα ΜΜΕ τείνουν να δημοσιεύουν tweets με ουδέτερο συναίσθημα, καθώς προφανώς πρόκειται για τίτλους ειδήσεων που εμπεριέχουν απλή αναφορά γεγονότων στο μεγαλύτερο μέρος τους. Αντιθέτως, Έλληνες και Γερμανοί χρήστες εκφράζουν κυρίως αρνητικό συναίσθημα, με τους Έλληνες να το εκφράζουν σε μεγαλύτερο βαθμό. Επίσης, το συναίσθημα που εκφράζουν οι χρήστες των δύο χωρών δεν φαίνεται να έχει καμία συνάφεια με το συναίσθημα που εκφράζεται από τα media κάθε χώρας, αντιθέτως, είναι σε μεγάλο βαθμό διαφοροποιημένα.

Επίσης, οι Έλληνες πολίτες τάσσονται εναντίον της ακροδεξιάς και το ίδιο αρνητικό συναίσθημα κυριαρχεί και στους Γερμανούς χρήστες, αν και σε μικρότερο ποσοστό. Τόσο οι Έλληνες, όσο και οι Γερμανοί χρήστες, μπορεί να εκφράζονται στο μεγαλύτερο ποσοστό αρνητικά για τους πρόσφυγες και την εκπαίδευσή τους, αλλά, οι αναρτήσεις με αρνητικό συναίσθημα είναι κατά τι μόνο περισσότερες από αυτές με θετικό, με αποτέλεσμα να μην μπορεί να ισχυριστεί κανείς ότι διακρίνεται ξενοφοβία απέναντι στα παιδιά των προσφύγων και στο ενδεχόμενο ενσωμάτωσής τους στο εκπαιδευτικό σύστημα

της εκάστοτε χώρας. Επιπλέον δεν φαίνεται οι πολίτες της μίας χώρας να είναι πιο ξενοφοβικοί από αυτούς της άλλης.

Στη συνέχεια, πραγματοποιήθηκε μια συγκριτική μελέτη μεταξύ δυο διαφορετικών μεθόδων ανάλυσης, τη μέθοδο βασισμένη στο λεξικό και του Naive Bayes. Η μέθοδος που βασίστηκε στο λεξικό παρουσίασε σε μεγάλο βαθμό επιτυχία, σε αντίθεση με τον αλγόριθμο Naive Bayes, ο οποίος εμφάνισε ιδιαίτερος μικρά ποσοστά επιτυχίας.

Οι λόγοι που το λεξικό σημείωσε υψηλά ποσοστά επιτυχίας είναι εύκολο να εντοπιστούν. Βασικό ρόλο παίζει ο τρόπος χαρακτηρισμού συναισθήματος που εφαρμόζει η εν λόγω μέθοδος, ο οποίος είναι ιδιαίτερα απλός και ευνοεί ισχυρά την επιτυχή πρόβλεψη της αντίστοιχης επισημείωσης, καθώς στηρίζεται σε ένα αριθμητικό κριτήριο υπεροχής λέξεων με θετικό ή αρνητικό χαρακτήρα, προεπιλεγμένες όμως μέσα από το λεξικό και αυστηρά μόνο σε όσα tweets εντοπίστηκαν οι συγκεκριμένες λέξεις. Η προσέγγιση αυτή προσομοιάζει σε σημαντικό βαθμό τον τρόπο αξιολόγησης του κριτή και συνεπώς αναμένεται να έχει θετική έκβαση τις περισσότερες φορές. Επίσης, το ενδεχόμενο ανεπαρκές φιλτράρισμα από ανεπιθύμητους χαρακτήρες κατά τη διαδικασία προεπεξεργασίας των tweets δεν επηρεάζει την απόδοσή της συγκεκριμένης μεθόδου, σε πλήρη αντίθεση με τον Naive Bayes, ο οποίος είναι εξαιρετικά ευαίσθητος από την παράμετρο αυτή. Τέλος, το γεγονός ότι πραγματοποιήθηκε επέκταση του λεξικού, το οποίο εμπλουτίστηκε με λέξεις που σχετίζονταν με το προσφυγικό, έπαιξε σημαντικό ρόλο στα υψηλά ποσοστά επιτυχίας που σημείωσε η εν λόγω μέθοδος.

Ο αλγόριθμος Naive Bayes, από την άλλη πλευρά, δεν κατάφερε να πετύχει τις υψηλές επιδόσεις του λεξικού. Ένας βασικός παράγοντας είναι ο χαρακτηρισμός συναισθήματος στα πραγματικά tweets (επισημείωση), ο οποίος στηρίζεται κυρίως σε κριτήρια τα οποία εξαρτώνται από τις προσωπικές πεποιθήσεις, το προσωπικό ύφος, το κοινωνικό και μορφωτικό υπόβαθρο του κριτή. Τα κριτήρια αυτά μάλιστα δεν είναι βέβαιο ότι εφαρμόζονται ισότιμα σε όλα τα tweets, καθώς ο κριτής αδυνατεί κάθε φορά να υλοποιεί την ίδια στρατηγική με τον ίδιο αυστηρό τρόπο.

Επιπλέον, το δείγμα το οποίο εξετάστηκε, αποδείχτηκε περιορισμένο για τη σωστή εκπαίδευση του αλγορίθμου. Αυτό αποδείχτηκε από την αξιολόγησή του με τη βοήθεια της τεχνικής Monte Carlo και την κατασκευή συνθετικών tweets, η οποία έδειξε ότι ο παράγοντας αυτός είναι από τους πλέον κρίσιμους, καθώς και ότι κάτω από απλές ελεγχόμενες συνθήκες και απλούς κανόνες επισημείωσης, 10000 tweets είναι ο ελάχιστος απαιτούμενος αριθμός, ώστε ο αλγόριθμος να μεγιστοποιήσει την απόδοσή του. Εκτός αυτού, το μεγάλο πλήθος λέξεων το οποίο, ενώ έχει κοινό προσανατολισμό, διαφέρει στη δομή του, όπως π.χ. λέξεις σε διαφορετική κλίση, γένος κλπ. επηρέασε την απόδοση της μεθόδου. Πρόκειται στην ουσία για ομόρριζες λέξεις οι οποίες όμως κατά τη διαδικασία εκπαίδευσης αντιμετωπίζονται ως πλήρως αυτόνομες επεξηγηματικές μεταβλητές εισόδου.

Τέλος, τη διαδικασία δυσχέρανε και το μεγάλο πλήθος λέξεων, οι οποίες δεν έχουν σημαντικό αντίκτυπο στη μεταβλητή-στόχο, δηλαδή τον χαρακτηρισμό συναισθήματος. Οι λέξεις αυτές έχουν ουδέτερο χαρακτήρα, όπως π.χ. τα βοηθητικά ρήματα «είμαι», «έχω» κ.α., αλλά διογκώνουν τον πίνακα λέξεων-κειμένου (tdm) και αποπροσανατολίζουν τη διαδικασία εκπαίδευσης.

Από περαιτέρω ανάλυση, φάνηκε πως η μέθοδος του λεξικού προέβλεψε με μεγαλύτερη ακρίβεια το θετικό συναίσθημα, κατόπιν το ουδέτερο και πιο δύσκολα το αρνητικό. Αυτό ενδεχομένως να οφείλεται στο ότι δεν έγινε εφαρμογή κάποιας μεθόδου που να λαμβάνει υπόψιν της την άρνηση σε μία πρόταση, όπως για παράδειγμα την παρουσία αρνητικών μορίων. Αυτό θα μπορούσε να αντιμετωπιστεί με το να γίνεται αντιστροφή της πολικότητας μία πρότασης, κάθε φορά που εντοπίζονται τα αρνητικά μόρια.

4. ΠΕΡΙΟΡΙΣΜΟΙ ΕΡΕΥΝΑΣ & ΠΡΟΤΑΣΕΙΣ ΓΙΑ ΜΕΛΛΟΝΤΙΚΕΣ ΕΡΕΥΝΕΣ

Κατά τη διεξαγωγή της έρευνας προέκυψαν οι παρακάτω περιορισμοί:

- Οι περισσότερες τεχνικές που έχουν αναπτυχθεί για την ανάλυση συναισθήματος αφορούν στην αγγλική γλώσσα. Για τις υπόλοιπες γλώσσες και ιδιαίτερα για τα Ελληνικά δεν είναι διαθέσιμα πολλά από τα εργαλεία που υπάρχουν για τα Αγγλικά. Παραδείγματος χάρη, δεν υπάρχει η δυνατότητα για αυτόματο stemming στο ελληνικό κείμενο, δεν υπάρχουν έτοιμα stopwords στις βιβλιοθήκες της R, ενώ τα διαθέσιμα λεξικά είναι ελάχιστα.
- Ο αριθμός των tweets ήταν αρκετά μεγάλος για τα δεδομένα ενός τυπικού υπολογιστικού συστήματος, με αποτέλεσμα για την εκτέλεση του αλγορίθμου να απαιτούνται πολλές ώρες ανάλυσης.

Οι επεκτάσεις αυτής της έρευνας μπορεί να είναι πολλές ανάλογα το προσωπικό ενδιαφέρον και τις επιδιώξεις του εκάστοτε ερευνητή. Αρχικά θα μπορούσε πέρα από την πολικότητα των tweets, να γίνει προσπάθεια εντοπισμού μέσα από ένα μεγαλύτερο σύνολο πιθανών συναισθημάτων, είτε λαμβάνοντας υπόψη τις έξι συναισθηματικές καταστάσεις, τα λεγόμενα POMS (Tension, Depression, Anger, Vigour, Fatigue, Confusion), είτε τα οχτώ κύρια συναισθήματα (acceptance, fear, anger, joy, anticipation, sadness, disgust, surprise).

Πολλοί χρήστες στο Twitter τείνουν να γράφουν ειρωνικά ή αστεία μηνύματα, καθώς και μηνύματα με διφορούμενη σημασία. Θα ήταν ενδιαφέρον να φανεί το κατά πόσο επηρεάζουν την απόδοση των τεχνικών της ανάλυσης συναισθήματος τα εν λόγω μηνύματα, εάν αντιμετωπιστούν διαφορετικά. Επίσης, θα μπορούσε να δημιουργηθεί μία εφαρμογή που θα συλλέγει tweets σε πραγματικό χρόνο και αυτόματα θα τα κατηγοριοποιεί βάσει του συναισθήματος που εκφράζουν. Τέλος, θα μπορούσε να χρησιμοποιηθεί για την ανάλυση συναισθήματος μία υβριδική μέθοδος που να συνδυάζει το λεξικό με τη

μηχανική μάθηση, ώστε να αξιοποιηθούν τα πλεονεκτήματα και των δύο μεθόδων.

ΒΙΒΛΙΟΓΡΑΦΙΚΕΣ ΑΝΑΦΟΡΕΣ

ΒΙΒΛΙΑ

Δημητρόπουλος, Α., *Τα Συνταγματικά δικαιώματα των αλλοδαπών στην Ελλάδα: Constitutional rights of foreign citizens in Greece*, Αντ. Ν. Σάκκουλα, Αθήνα, 2004, σελ. 7

Μπαμπινιώτης, Γ., *Λεξικό της νέας ελληνικής γλώσσας (Γ' έκδοση)*, Κέντρο Λεξικολογίας, Αθήνα, 2008, σελ. 1548.

Νάσκου-Περράκη Π., *Το νομικό καθεστώς των προσφύγων στη διεθνή και ελληνική έννομη τάξη. Η Σύμβαση της Γενεύης του 1951 για το καθεστώς των προσφύγων*, Αντ. Ν. Σάκκουλας, Αθήνα-Κομοτηνή, 1991, σελ. 279.

Σιδέρη, Μ., *Το βιβλίο του twitter - Ένας οδηγός για "αθώους" χρήστες*, Κλειδάριθμος, Αθήνα, 2011, σελ. 24.

Χρυσόγονος, Κ., *Ατομικά και Κοινωνικά Δικαιώματα*, Νομική βιβλιοθήκη, 2006, σελ. 334

Bourdieu, P., *Οι Κληρονόμοι. Οι φοιτητές και η κουλτούρα*, μετάφραση: Ν. Παναγιωτόπουλος - Μ. Βιδάλη, Ινστιτούτο του βιβλίου - Καρδαμίτσας, Αθήνα, 1966, σελ.73

Han, J., & Kamber, M., *Data Mining: Concepts and Techniques, 2nd Ed.*, Morgan Kaufmann Publishers, U.S.A., 2006, σελ. 7.

Hand, D. J., Mannila, H., & Smyth, P., *Principles of Data Mining*, The MIT Press, London, 2001, σελ. 27.

Rajaraman, A., Ullman, J., *Εξόρυξη από μεγάλα σύνολα δεδομένων*, μετάφραση: Αναστάσιος Αθ. Γούναρης, Ιωάννης Μανωλόπουλος, Απόστολος Ν. Παπαδόπουλος, Κωνσταντίνος Τσίχλας, Εκδόσεις Νέων Τεχνολογιών, 2014, σελ. 1.

Sartre, J.- P., *Στοχασμοί για το εβραϊκό ζήτημα*, μετάφραση: Αθανάσιος Σαμαρτζής, Βιβλιοπωλείον της Εστίας, Αθήνα, 2006, σελ. 48.

Wenger, E., White, N., & Smith, J.D., *Digital Habitats: Stewarding Technologies for Communities*, CPsquare, Portland, 2009, σελ. 17.

ΗΛΕΚΤΡΟΝΙΚΑ ΒΙΒΛΙΑ

Κύρκος, Ε., *Επιχειρηματική ευφυΐα και εξόρυξη δεδομένων*, Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών, 2015, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://hdl.handle.net/11419/1226>, [προσπελάστηκε στις 8 Ιουνίου 2017], σελ. 127.

Congdon, P., *Bayesian Statistical Modelling*, Second Edition, Wiley Series in Probability and Statistics, 2006, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://bacuc.hd.free.fr/WebDAV/data/Bouquins/Congdon%20-%20Bayesian%20Statistical%20Modelling.pdf> , [προσπελάστηκε στις 22 Ιουνίου 2017], σελ. 2.

Liu, B. , *Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data*, Springer, 2006, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://sirius.cs.put.poznan.pl/~inf89721/Seminarium/Web Data Mining 2nd Edition Exploring Hyperlinks Contents and Usage Data.pdf>, [προσπελάστηκε στις 12 Φεβρουαρίου 2017], σελ. 461.

Liu, B., *Sentiment Analysis and Opinion Mining*, Morgan & Claypool Publishers, 2012, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://www.cs.uic.edu/~liub/FBS/SentimentAnalysis-and-OpinionMining.pdf>, [προσπελάστηκε στις 12 Φεβρουαρίου 2017], σελ. 10-12.

O' Hagan, A., *The Bayesian Approach to Statistics*, Chapter 6, in: *Handbook of Probability: Theory and Applications*, Corwin, 2008, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: https://www.corwin.com/sites/default/files/upm-binaries/18550_Chapter6.pdf, [προσπελάστηκε στις 12 Ιουνίου 2017], σελ. 94.

UNHCR, *Οι πρόσφυγες του κόσμου: Σε αναζήτηση αλληλεγγύης (σύνοψη)*, Υπατη Αρμοστεία του ΟΗΕ για τους Πρόσφυγες, 2012, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: https://www.unhcr.gr/fileadmin/Greece/General/publications/OI_PROSFYGEES_TOY_KOSMOY.pdf, [προσπελάστηκε στις 23 Μαρτίου 2017], σελ. 24.

ΑΡΘΡΑ ΕΝΤΥΠΩΝ ΠΕΡΙΟΔΙΚΩΝ

Έμκε - Πουλοπούλου, Η., Μετανάστες και πρόσφυγες στην Ελλάδα 1970-1990, *Εκλογή Θεμάτων Κοινωνικής Πρόνοιας*, 1990, 85/86, σελ. 46.

Νάσκου-Περράκη, Μ., Δικαιώματα και υποχρεώσεις στα πλαίσια της Σύμβασης της Γενεύης του 1951 για το καθεστώς των προσφύγων: Γενική θεώρηση, *Κοινωνική Εργασία*, 1992, 26, σελ. 93.

Τρουμπέτα, Σ., Μερικές σκέψεις για την παράσταση του «άλλου», *Επιθεώρηση Κοινωνικών Ερευνών*, 2000, 101-102, σελ. 139.

Alessia, D., Ferri, F., Grifoni, P., & Guzzo, T., Approaches, Tools and Applications for Sentiment Analysis Implementation, *International Journal of Computer Applications*, 2015, 125(3), σελ. 31.

Gualda, E. & Rebollo, C., The Refugee Crisis on Twitter: A Diversity of Discourses at a European Crossroads, *Journal of Spatial and Organizational Dynamics*, 2016, 4 (3), σελ. 199-212.

Hathaway, J.C., Why Refugee Law Still Matters, *Melbourne Journal of International Law*, 2007, 89 (8), σελ. 351.

Küchler, M., Xenophobie im internationalen Vergleich. In: Falter J.W., Jaschke HG., Winkler J.R. (eds) *Rechtsextremismus. Politische Vierteljahresschrift, Verlag für Sozialwissenschaften*, 1996, 27, σελ. 248-262.

Medhat, W., Hassan, A., Korashy, H., Sentiment analysis algorithms and applications: A survey, *Ain Shams Engineering Journal*, 2014, 5 (4), σελ. 1093.

Makrynioti, N., & Vassalos, V., Sentiment Extraction from Tweets: Multilingual Challenges, *Big Data Analytics and Knowledge Discovery Lecture Notes in Computer Science*, 2015, σελ. 136-148.

Mestheneos, E. & Ioannidi, E., Obstacles to Refugee Integration in the European Union Member States, *Journal of Refugee Studies*, 2002, 15, Oxford Uni. Press, σελ. 304-320.

Niendorf, M. & Reitz, S., Das Menschenrecht auf Bildung im deutschen Schulsystem
Was zum Abbau von Diskriminierung notwendig ist, *Deutsches Institut für Menschenrechte*, 2016, σελ. 17.

Oksana, Y., Xenophobia: Understanding the roots and consequences of negative attitudes towards immigrants, *Counseling Psychologist*, 2009, 37 (1), σελ. 56.

O'Reilly, T., What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software, *International Journal of Digital Economics*, 2007, 65, σελ. 17-37.

Phillimore J. & Goodson L., Making a place in the Global City: The relevance of Indicators of Integration, *Journal of Refugee Studies*, 2008, 21 (3),

Oxford Uni.Press, σελ. 305-325.

Tsakalidis, A., Papadopoulos, S., & Kompatsiaris, I., An Ensemble Model for Cross-Domain Polarity Classification on Twitter, *Web Information Systems Engineering- WISE*, Springer International Publishing, 2014 , σελ.. 168-177.

Tsytsarau, M., & Palpanas, T., Survey on mining subjective data on the web. *Data Mining and Knowledge Discovery*, 2012, 24(3), σελ. 478-514.

Turkat, D., Social networks: Theory and practice, *Journal of Community Psychology*, 1980, 3, σελ. 99-109.

Xu, C., Ouyang, F., & Chu, H., The academic library meets web 2.0: Applications and implications. *The Journal of Academic Librarianship*, 2009, 35 (4), σελ. 324-331.

ΑΡΘΡΑ ΗΛΕΚΤΡΟΝΙΚΩΝ ΠΕΡΙΟΔΙΚΩΝ

Ager, A. & Strang, A., Indicators of Integration : Final Report, *Home Office Development and Practice Report 28*, 2004, Home Office, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://webarchive.nationalarchives.gov.uk/20110218141321/http://rds.homeoffice.gov.uk/rds/pdfs04/dpr28.pdf> , [προσπελάστηκε στις 23 Μαρτίου 2017], σελ. 13- 23.

Lei, Z., Hsu, M., Liu, B., Ghosh, R., & Dekhil, M., Combining Lexicon-based and Learning-based Methods for Twitter Sentiment Analysis., *HP Laboratories*, 2011, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.hpl.hp.com/techreports/2011/HPL-2011-89.pdf>, [προσπελάστηκε στις 18 Μαρτίου 2017].

Pang B., & Lee, L., Opinion mining and sentiment analysis, *Foundations and Trends in Information Retrieval*, 2 (1-2), 2008, [διαδίκτυο]. Διαθέσιμο στη

διεύθυνση:<http://www.cs.cornell.edu/home/llee/omsa/omsa.pdf>, [προσπελάστηκε στις 5 Αυγούστου 2017], σελ. 11.

Vohra, S., Teraiya, J. B., A Comparative Study of Sentiment Analysis Techniques., *Journal of Information Knowledge and Research in Computer Engineering*, 2 (2), 2012, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.ejournal.aessangli.in/ASEEJournals/CE63.pdf>, [προσπελάστηκε στις 18 Μαρτίου 2017], σελ. 314.

Λιαλιούτη, Ζ., Γεωργιάδου, Β., Γαλαριώτης, Γ., & Καφέ, Α., Η Ξενοφοβία στην Ελλάδα ως ερευνητικό πρόβλημα: Εννοιολογικά και μεθοδολογικά ζητήματα, στο: *Επιστήμη και κοινωνία: Επιθεώρηση πολιτικής και ηθικής θεωρίας, Gutenberg (περιοδικό)*, 2017, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://ejournals.epublishing.ekt.gr/index.php/sas/article/view/11494/11559>, [προσπελάστηκε στις 20 Μαρτίου 2017], σελ. 162.

Μαυροζαχαράκης, Μ., Colin Crouch, Σειρά μεγάλοι στοχαστές, *24 γράμματα*, 2014, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.24grammata.com/wp-content/uploads/2014/01/COLIN-CROUCH-manrozaharakis-24grammata.com.pdf>, [προσπελάστηκε στις 18 Μαρτίου 2017], σελ. 29.

ΔΙΑΔΙΚΤΥΑΚΟΙ ΤΟΠΟΙ

Αντζουλάκος, Δ., Ανάλυση Δεδομένων με τη Χρήση Στατιστικών Πακέτων-Εισαγωγή στο R, 2013, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: [http://www.unipi.gr/faculty/dantz/Introduction to R.pdf](http://www.unipi.gr/faculty/dantz/Introduction%20to%20R.pdf), [προσπελάστηκε στις 3 Αυγούστου 2017], σελ. 1.

Βοσκόπουλος, Γ., Η άνοδος της ακροδεξιάς στην Ευρώπη, *Ελεύθερη Ζώνη*, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.elzoni.gr/html/ent/942/ent.15942.asp>, [προσπελάστηκε στις 11 Μαρτίου 2017].

Δίκτυο Καταγραφής Περιστατικών Ρατσιστικής Βίας, Ετήσια Έκθεση 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: http://rvrn.org/wp-content/uploads/2017/04/Report_2016gr.pdf, [προσπελάστηκε στις 11 Μαρτίου 2017], σελ. 1.

Επίσημη Εφημερίδα της Ευρωπαϊκής Ένωσης, Οδηγία 2013/33/ΕΕ του Ευρωπαϊκού Κοινοβουλίου και του συμβουλίου της 26ης Ιουνίου 2013 σχετικά με τις απαιτήσεις για την υποδοχή των αιτούντων διεθνή προστασία (αναδιατύπωση), 2013, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:180:0096:0116:EL:PDF>, [προσπελάστηκε στις 151 Μαρτίου 2017].

Ευρωπαϊκή Επιτροπή, Ετήσια έκθεση δραστηριοτήτων 2013, 2014, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://www.easo.europa.eu/sites/default/files/public/BZAD14001ELN.pdf>, [προσπελάστηκε στις 7 Φεβρουαρίου 2017], σελ. 7.

Η εισβολή της Χρυσής Αυγής στο Δημοτικό σχολείο στο Πέραμα, *alfavita.gr*, 2017, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.alfavita.gr/arhron/koinonia/vinteo-i-eisvoli-tis-hrysis-aygis-sto-dimotiko-sholeio-sto-perama>, [προσπελάστηκε στις 11 Μαρτίου 2017].

Καλώς ήλθατε στη Γερμανία, Πληροφορίες για μετανάστες, Bundesamt für Migration und Flüchtlinge, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: https://www.bmi.bund.de/SharedDocs/Downloads/DE/Broschueren/Sprachvarianten/Willkommen_in_Deutschland_gr.pdf?blob=publicationFile, [προσπελάστηκε στις 9 Φεβρουαρίου 2017], σελ. 73.

Κάπα research, Έρευνα σε πρόσφυγες, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: http://www.enpe.gr/media/368840/2016_03_16_%20Έρευνα%20σε%20πρόσφυγες%20ENΠΕ.pdf, [προσπελάστηκε στις 14 Φεβρουαρίου 2017].

Κάπα research, Έρευνα στην Αττική για την προσφυγική κρίση, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: http://i2.aftocdn.gr/mediafiles/2016/05/erevna_peda_aftodioikisi.pdf, [προσπελάστηκε στις 18 Φεβρουαρίου 2017].

Κέντρο Ευρωπαϊκής Πληροφόρησης Europe Direct, Η κοινωνική ένταξη και η ενσωμάτωση των προσφύγων στην αγορά εργασίας, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.acci.gr/acci/articles/article.jsp?context=103&categoryid=8289&articleid=1038>, [προσπελάστηκε στις 11 Μαρτίου 2017].

Ντε Μπενουά, Α., Το θινκ τανκ του σύγχρονου ρατσισμού, *Ελευθεροτυπία*, 2000, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.iospress.gr/ios2000/ios20000305a.htm>, [προσπελάστηκε στις 12 Ιουνίου 2017].

Συνεχής άνοδος των ακροδεξιών κομμάτων στην Ευρώπη -Σε ποιες χώρες ανεβαίνουν επικίνδυνα τα ποσοστά τους, *iefimerida.gr*, 2015, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.iefimerida.gr/news/236718/synehis-anodos-ton-akrodexion-kommaton-stin-eyropi-se-poiies-hores-anevainoyn-epikindyna#ixzz4DGyrZdU6>, [προσπελάστηκε στις 11 Μαρτίου 2017].

AfD erreicht mit 15,5 Prozent erneut Höchstwert, *welt.de*, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://www.welt.de/politik/deutschland/article160569813/AfD-erreicht-mit-15-5-Prozent-erneut-Hoehstwert.html>, [προσπελάστηκε στις 11 Μαρτίου 2017].

Amnesty International, Leben in Unsicherheit- Wie Deutschland die Opfer rassistischer Gewalt im Sicht lässt, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://www.amnesty.de/files/Amnesty-Bericht-Rassistische->

[Gewalt-in-Deutschland-Juni2016.pdf](#), [προσπελάστηκε στις 9 Φεβρουαρίου 2017], σελ. 50.

Amnesty International, Refugees Welcome Survey, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: https://www.amnesty.at/de/view/files/download/showDownload/?tool=12&feld=download&sprach_connect=424, [προσπελάστηκε στις 11 Μαρτίου 2017].

Bartlett, J., & Norrie, R., Immigration on twitter- understanding public attitudes online, 2015, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: https://www.demos.co.uk/files/immigration_on_twitter.pdf?1428506056, [προσπελάστηκε στις 2 Μαρτίου 2017].

Bax. M., Wie wird das Recht auf Bildung gesetzlich geregelt?, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.bildungsexperten.net/wissen/welche-gesetze-gibt-es-zur-bildung/>, [προσπελάστηκε στις 11 Μαρτίου 2017].

BeyondCSRNet, Twitter resolution call for refugees- How to solve the refugee crisis, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://beyonddcsr.org/wp-content/uploads/2017/02/BEYONDRESOLUTION.pdf>, [προσπελάστηκε στις 2 Ιανουαρίου 2017], σελ. 14.

Businessinsider.com, This map shows the routes of Europe's refugee nightmare — and how it's getting worse, 2015, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.businessinsider.com/map-of-europe-refugee-crisis-2015-9>, [προσπελάστηκε στις 15 Ιουλίου 2017].

Buggisch, C., Social Media und Messenger – Nutzerzahlen in Deutschland 2017, 2017, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://buggisch.wordpress.com/2017/01/02/social-media-und->

[messenger-nutzerzahlen-in-deutschland-2017/](#), [προσπελάστηκε στις 5 Μαρτίου 2017].

Crimson Hexagon, The Syrian Refugee Crisis on Social Media, 2015, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://partners.twitter.com/content/dam/partners-twitter/success-stories/pdf/Syrian%20Refugee%20Crisis%20Report%20-%20Partnerships.pdf>, [προσπελάστηκε στις 2 Φεβρουαρίου 2017], σελ. 3.

Das denken die Deutschen wirklich über Flüchtlinge, Merkel - und die Medien, *stern.de*, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.stern.de/politik/deutschland/fluechtlinge--das-denken-die-deutschen-wirklich-6737204.html>, [προσπελάστηκε στις 11 Φεβρουαρίου 2017].

Eisnecker, P., & Schupp, J., Stimmungsbarometer zu Geflüchteten in Deutschland, Deutsches Institut für Wirtschaftsforschung, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: https://www.diw.de/documents/publikationen/73/diw_01.c.534430.de/16-21.pdf, [προσπελάστηκε στις 11 Μαρτίου 2017].

ELTRUN, Πιο ώριμοι και συνειδητοποιημένοι δηλώνουν οι Internet users στην Ελλάδα, όσον αφορά την χρήση των Social Media και των σχετικών εργαλείων digital marketing από τις επιχειρήσεις, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: www.greekecommerce.gr/gr/file-download/eltrun-digital-marketing-2016, [προσπελάστηκε στις 5 Φεβρουαρίου 2017], σελ. 3.

Ergebnisse der Abgeordnetenhauswahlen in Berlin (bis 1989 West-Berlin), *wahlrecht.de*, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.wahlrecht.de/ergebnisse/berlin.htm>, [προσπελάστηκε στις 11 Φεβρουαρίου 2017].

Jones, R., Social Media Marketing 101, Part 1, 2009, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://searchenginewatch.com/sew/opinion/2064413/social-media-marketing-101-part>, [προσπελάστηκε στις 2 Φεβρουαρίου 2017].

Kroker, M., Social Media 2016: Die Zahlen & Fakten zu Facebook, Twitter & Co. im Überblick, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://blog.wiwo.de/look-at-it/2016/06/30/social-media-2016-die-zahlen-fakten-zu-facebook-twitter-co-im-ueberblick/>, [προσπελάστηκε στις 5 Φεβρουαρίου 2017].

Mohamud, A., 2013 UK Digital Future in Focus, 2013, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: http://www.comscore.com/Insights/Presentations-and-Whitepapers/2013/2013-UK-Digital-Future-in-Focus?cs_edgescape_cc=GR, [προσπελάστηκε στις 21 Μαρτίου 2017].

Muñoz, V., Report of the Special Rapporteur on the right to education, 2007, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: http://www.netzwerkbildungsfreiheit.de/pdf/Munoz_Mission_on_Germany.pdf, [προσπελάστηκε στις 21 Φεβρουαρίου 2017].

Unicef, Refugee and Migrant Children in Greece - by Region, 2017, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: https://www.unicef.org/eca/Map_RefugeeMigrantChildren_25MAR2017.pdf, [προσπελάστηκε στις 4 Αυγούστου 2017].

Unicef, Kindheit im Wartezustand, 2017, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://www.unicef.de/blob/137704/053ab16048c3f443736c4047694cc5d1/studie--kindheit-im-wartezustand-data.pdf>, [προσπελάστηκε στις 4 Αυγούστου 2017].

Venables, W. N., Smith, D. M. & the R Core Team, An Introduction to R, 2017, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://cran.r-project.org/doc/manuals/r-release/R-intro.pdf>, [προσπελάστηκε στις 5 Αυγούστου 2017], σελ. 2.

Zhang, J., Social media and distance education, 2010, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://www.scribd.com/document/160525480/Social-Media-and-Distance-Education-De-Oracle>, [προσπελάστηκε στις 5 Φεβρουαρίου 2017].

Ανεπιθύμητα τα προσφυγόπουλα στο 2ο Δημοτικό Σχολείο Φιλιπιάδας, kathimerini.gr, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.kathimerini.gr/875595/article/epikairothta/ellada/anepi8ymhta-ta-ta-prosfygo-poula-sto-2o-dhmotiko-sxoleio-filippiadas>, [προσπελάστηκε στις 7 Φεβρουαρίου 2017], σελ. 73.

Διανέοσις, Το προσφυγικό πρόβλημα και οι Έλληνες, 2016, [διαδίκτυο], Διαθέσιμο στη διεύθυνση: http://www.dianeosis.org/wp-content/uploads/2016/02/immigration_04.pdf, σελ. 53.

Ελληνικό Συμβούλιο για τους Πρόσφυγες, [διαδίκτυο]. Διαθέσιμο στο: <http://www.gcr.gr/index.php/el/about-gcr/what-we-do>, [προσπελάστηκε στις 8 Μαρτίου 2017].

Επιστολή Συνήγορου του Πολίτη στους αρμόδιους υπουργούς δηλαδή τους Αναπληρωτές Υπουργούς Μεταναστευτικής Πολιτικής, Προστασίας του Πολίτη και Κοινωνικής Αλληλεγγύης, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://www.synigoros.gr/resources/160330-epistoliasynodeytoi-anilikoi.pdf>, [προσπελάστηκε στις 4 Ιουνίου 2017].

Καϊδατζής, Α., Το δικαίωμα δωρεάν παιδείας μεταξύ (συνταγματικού) δικαίου και (νομοθετικής) πολιτικής. Με αφορμή την απόφαση για τα δίδακτρα στις μεταπτυχιακές σπουδές, [διαδίκτυο]. Διαθέσιμο στη

διεύθυνση: <http://www.constitutionalism.gr/2558-to-dikaiwma-dwrean-paideias-metaxy-syntagmatikoy-d/> , [προσπελάστηκε στις 12 Φεβρουαρίου 2017].

Συνήγορος του Πολίτη, Παρέμβαση Συνηγόρου του Πολίτη για τα ασυνόδευτα παιδιά, πρόσφυγες και μετανάστες, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://www.synigoros.gr/resources/dt-asynodeytoi-anilikoi-3032016.pdf> , [προσπελάστηκε στις 14 Φεβρουαρίου 2017].

Σύνταγμα της Ελλάδας, Όπως αναθεωρήθηκε με το ψήφισμα της 27^{ης} Μαΐου 2008 της Η' Αναθεωρητικής Βουλής των Ελλήνων, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.hellenicparliament.gr/UserFiles/8c3e9046-78fb-48f4-bd82-bbba28ca1ef5/SYNTAGMA.pdf> , [προσπελάστηκε στις 23 Φεβρουαρίου 2017], σελ. 30.

Τζωρτζοπούλου, Μ., & Κοτζαμάνη, Α., Η εκπαίδευση των αλλοδαπών μαθητών. Διερεύνηση των προβλημάτων και της προοπτικής επίλυσής τους, Εθνικό Κέντρο Κοινωνικών Ερευνών, 2008, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.ekke.gr/publications/wp/wp19.pdf> , [προσπελάστηκε στις 23 Φεβρουαρίου 2017], σελ. 24.

Τσιτσελίκης Κ., Ανήλικοι πρόσφυγες και υπηρεσίες εκπαίδευσης. Δικαίωμα στην εκπαίδευση και υποχρεώσεις του κράτους, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: https://www.esos.gr/sites/default/files/articles-legacy/keimeno_tekmiriosis_dik_ekpaideusi2_1.pdf , [προσπελάστηκε στις 12 Φεβρουαρίου 2017].

Υπουργείο Προστασίας του Πολίτη, 2014, Ελληνική Δημοκρατία, Συχνές ερωτήσεις, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: http://www.minocp.gov.gr/index.php?option=ozo_content&lang=EN&perform=view&id=5055&Itemid=466&lang=GR, [προσπελάστηκε στις 6 Φεβρουαρίου 2017].

Agarwala, A., & Hamann, G., Schaffen die das?, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.zeit.de/gesellschaft/2016-07/fluechtlingskinder-schule-fluechtlinge-deutschland>, [προσπελάστηκε στις 17 Φεβρουαρίου 2017].

Agarwala, A., Schenk, A., & Spiewak, M., Flüchtlinge im Schulalltag, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.zeit.de/2016/29/integration-fluechtlinge-schule-kinder-jugendliche-deutschunterricht-sprachbarriere-bildungspolitik>, [προσπελάστηκε στις 17 Φεβρουαρίου 2017].

Bildung in Deutschland 2016, Bundesministerium für Bildung und Forschung, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.bildungsbericht.de/de/bildungsberichte-seit-2006/bildungsbericht-2016/pdf-bildungsbericht-2016/bildungsbericht-2016>, [προσπελάστηκε στις 9 Φεβρουαρίου 2017], σελ. 189.

Blaß, K., & Himmelrath, A., Wenn Grundschüler plötzlich von "Kanaken" sprechen, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.spiegel.de/lebenundlernen/schule/fluechtlingskinder-in-der-schule-tipps-fuer-lehrer-a-1100905.html>, [προσπελάστηκε στις 17 Φεβρουαρίου 2017].

Bundesamt fuer Migration und Fluechtlinge, Unbegleitete Minderjaehrige, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.bamf.de/DE/Fluechtlingsschutz/UnbegleiteteMinderjaehrige/unbegleitete-minderjaehrige-node.html>, [προσπελάστηκε στις 14 Φεβρουαρίου 2017].

Das Menschenrecht auf Bildung, 2013, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.bpb.de/gesellschaft/kultur/zukunft->

bildung/156819/menschenrecht?p=all, [προσπελάστηκε στις 17 Φεβρουαρίου 2017].

Diakonie Deutschland, Unbegleitete minderjaehrige Fluechtlinge, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://info.diakonie.de/infothek/wissen-kompakt/detail/unbegleitete-minderjaehrige-fluechtlinge/>, [προσπελάστηκε στις 4 Μαρτίου 2017].

European Commission, Regional refugee and migrant response plan for Europe Eastern Mediterranean and Western Balkans route, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: http://ec.europa.eu/justice/fundamental-rights/files/rights_child/data_children_in_migration.pdf), [προσπελάστηκε στις 16 Φεβρουαρίου 2017], σελ. 21.

Eurostat, Asylum Statistics, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: http://ec.europa.eu/eurostat/statistics-explained/index.php/Asylum_statistics, [προσπελάστηκε στις 6 Μαρτίου 2017].

Fakten zur Flüchtlingskrise, 2017, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.spiegel.de/politik/deutschland/fluechtlinge-und-einwanderer-die-wichtigsten-fakten-a-1030320.html#sponfakt=10>, [προσπελάστηκε στις 22 Φεβρουαρίου 2017].

Fagerholm, K., & Verheul, R., Safety and Fundamental Rights at Stake for Children on the Move, ENOC, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://www.synigoros.gr/resources/160211-ekthesi.pdf>, [προσπελάστηκε στις 12 Μαρτίου 2017].

FRA, Fundamental rights: Challenges and achievements in 2014, Annual Report 2014, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση:

http://fra.europa.eu/sites/default/files/fra-annual-report-2014_en.pdf ,
[προσπελάστηκε στις 16 Φεβρουαρίου 2017], σελ. 85.

FRA, Fundamental rights: Challenges and achievements in 2014, Annual Report 2014, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: http://fra.europa.eu/sites/default/files/fra-annual-report-2014_en.pdf, σελ. 85.

Frontex, Migratory routes map, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://frontex.europa.eu/trends-and-routes/migratory-routes-map/>,
[προσπελάστηκε στις 3 Ιουλίου 2017].

Grundgesetz für die Bundesrepublik Deutschland, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://www.bundestag.de/gg>. [προσπελάστηκε στις 23 Φεβρουαρίου 2017].

Hollenstein, O., Will gelernt sein, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.zeit.de/2016/18/fluechtlingskinder-bildung-schule-integration>, [προσπελάστηκε στις 17 Φεβρουαρίου 2017].

Hunko, A., The EU is becoming complicit in racism and fascism in Greece, 2012, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.statewatch.org/news/2012/nov/eu-greece-racism-prel-hunko.pdf>, [προσπελάστηκε στις 16 Φεβρουαρίου 2017].

IOM, Children on the Move, 2013, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: http://publications.iom.int/system/files/pdf/children_on_the_move_15may.pdf, [προσπελάστηκε στις 6 Μαρτίου 2017], σελ.32.

Kampf, L., Viele Straftaten gegen Asylunterkünfte, *tagesschau.de*, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://www.tagesschau.de/inland/bka-asylunterkuenfte-101.html> ,
[προσπελάστηκε στις 9 Φεβρουαρίου 2017].

OHCHR, Report of the Independent International Commission of Inquiry on the Syrian Arab Republic. A/HRC/30/48, 2015, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση:

<http://www.ohchr.org/EN/HRBodies/HRC/RegularSessions/Session30/Documents/A.HRC.30.48.AEV.pdf>, [προσπελάστηκε στις 9 Μαρτίου 2017].

Racist Attacks in Greece Hit 'Alarming' Levels: UNHCR, 2012, , [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.reuters.com/article/2012/10/23/us-greece-violenceidUSBRE89M12R20121023>, [προσπελάστηκε στις 9 Φεβρουαρίου 2017].

Schuler, K., & und Sadigh, P., Ab in die Schule, Alima, Bassam und Chipol, 2015, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.zeit.de/gesellschaft/schule/2015-09/schule-fluechtlingskinder-jugendliche>, [προσπελάστηκε στις 17 Φεβρουαρίου 2017].

Weise, Z., Refugee children get little education in Europe, *politico.eu*, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.politico.eu/article/refugee-children-education-europe-volunteers-greece-migrant-camp/>, [προσπελάστηκε στις 7 Φεβρουαρίου 2017].

Welchen Zugang haben geflüchtete Kinder zu Bildung?, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://landkarte-kinderrechte.de/>, [προσπελάστηκε στις 17 Φεβρουαρίου 2017].

Wer bekommt Asyl in Deutschland?, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://www.caritas.de/hilfeundberatung/ratgeber/migration/asylundflucht/wer-bekommt-asyl-in-deutschland>, [προσπελάστηκε στις 12 Φεβρουαρίου 2017].

Save the Children.org, Response to the IDC Report on Syria, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.savethechildren.org.uk/2016-01/response-idc-report-syria>, [προσπελάστηκε στις 25 Φεβρουαρίου 2017].

UN/ Meetings coverage and press releases, Security Council 'Unequivocally' Condemns ISIL Terrorist Attacks, Unanimously Adopting Text that Determines Extremist Group Poses 'Unprecedented' Threat, 2015, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.un.org/press/en/2015/sc12132.doc.htm>, [προσπελάστηκε στις 25 Φεβρουαρίου 2017].

UNAMA/OHCHR, Afghanistan, Human rights and protection of civilians in armed conflict, Special report on Kunduz province, 2015, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: https://unama.unmissions.org/sites/default/files/special_report_on_kunduz_province_12_december_2015.pdf, [προσπελάστηκε στις 23 Μαρτίου 2017].

UNHCR, Global Trends 2014, 2015, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.unhcr.org/statistics/country/556725e69/unhcr-global-trends-2014.html>, [προσπελάστηκε στις 24 Φεβρουαρίου 2017], σελ. 2, 3 και 5.

UNHCR, Refugees/Migrants Emergency Response-Mediterranean, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://data.unhcr.org/mediterranean/regional.php>, [προσπελάστηκε στις 8 Μαρτίου 2017].

UNHCR, The sea route to Europe: The Mediterranean passage in the age of refugees, 2015, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.unhcr.org/5592bd059.html>, [προσπελάστηκε 3 Μαρτίου 2017], σελ. 2, 5 & 6.

UNHCR, Προστατεύοντας τους πρόσφυγες, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://www.unhcr.gr/fileadmin/Greece/General/PDF-LOW.pdf>, [προσπελάστηκε στις 9 Μαρτίου 2017], σελ. 7.

UNHCR, Κατευθυντήριες οδηγίες για τα ασυνόδευτα παιδιά που ζητούν άσυλο, 2005, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: https://www.unhcr.gr/no_cache/prostasia/paidia.html?cid=801&did=490&sechash=1fbe2e71, [προσπελάστηκε στις 6 Φεβρουαρίου 2017].

UNICEF, Danger Every Step of the Way, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: https://www.unicef.org/emergencies/childrenonthemove/files/Child_Alert_Final_PDF.pdf, [προσπελάστηκε στις 25 Ιανουαρίου 2017], σελ. 2.

UNICEF, Convention on the Rights of the Child, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.unicef.org/crc/>, [προσπελάστηκε στις 24 Ιανουαρίου 2017].

ΔΙΠΛΩΜΑΤΙΚΕΣ ΕΡΓΑΣΙΕΣ- ΔΙΔΑΚΤΟΡΙΚΕΣ ΔΙΑΤΡΙΒΕΣ

Bosch, R., Sentiment Analysis: Incremental learning to build domain models, Master thesis, Universitat Pompeu Fabra, 2013, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: https://www.upf.edu/iis/pdf/Theses/Incremental_learning_to_build_domain_models-Master_Thesis_UPF-Raimon_Bosch.pdf, [προσπελάστηκε στις 12 Ιουλίου 2017], σελ. 10.

Απειρανθίτης, Α., Τσαχανσάχης, Β., Ανάλυση Δεδομένων από Κοινωνικά Δίκτυα για τη Μελέτη των Συναισθημάτων των Χρηστών, Διπλωματική εργασία, Εθνικό Μετσόβιο Πολυτεχνείο, Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://dspace.lib.ntua.gr/bitstream/handle/123456789/43734/Diploma%20Thesis.pdf?sequence=1>, [προσπελάστηκε στις 24 Ιανουαρίου 2017], σελ. 20.

Νοτοπούλου, Κ., Μεταναστευτική Κρίση στην Ε.Ε., Μεταπτυχιακή Διπλωματική εργασία, Πανεπιστήμιο Μακεδονίας, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://dspace.lib.uom.gr/bitstream/2159/19621/6/NotopoulouAikateriniMsc2016.pdf>, [προσπελάστηκε στις 24 Ιανουαρίου 2017], σελ. 46.

Παναγόπουλος, Π., Opinion Mining and Sentiment Analysis, Διπλωματική εργασία, Πανεπιστήμιο Πατρών, Τμήμα Μηχανικών Η/Υ και Πληροφορικής, 2015, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: http://83.212.99.169/wp/wp-content/uploads/2015/12/opinion_mining_and_sentiment_analysis.pdf, [προσπελάστηκε στις 24 Ιανουαρίου 2017], σελ. 1-2.

Παπαπάνου, Π., Το Προσφυγικό Ζήτημα, Οι ελληνοτουρκικές σχέσεις και ο ρόλος της Ευρωπαϊκής Ένωσης, Μεταπτυχιακή Διπλωματική Εργασία, Πανεπιστήμιο Μακεδονίας, 2017, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://dspace.lib.uom.gr/bitstream/2159/20086/4/PapapanouPanagiotaAikateriniMsc2017.pdf>, [προσπελάστηκε στις 8 Ιανουαρίου 2017], σελ. 14.

Παρδαλίδης, Θ., Στρατηγικές επιβίωσης και κοινωνικής προσαρμογής των αιτούντων άσυλο και των αναγνωρισμένων προσφύγων στην Ελλάδα. Μία εθνογραφική προσέγγιση, Διδακτορική Διατριβή, Πανεπιστήμιο Μακεδονίας, 2013, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://dspace.lib.uom.gr/bitstream/2159/15789/1/PardalidisTheodosiosPhd2013.pdf>, [προσπελάστηκε στις 8 Μαρτίου 2017], σελ. 2.

Σπανού Φ. & Αλεξάκης Β., Ασυνόδευτοι ανήλικοι πρόσφυγες στην Ελλάδα, Πτυχιακή εργασία, Τ.Ε.Ι. Ηρακλείου, 2007, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://nefeli.lib.teicrete.gr/browse/seyp/ker/2007/SpanouFoteini,AlexakisVasilis/attached-document/spanou.pdf>, [προσπελάστηκε στις 16 Ιανουαρίου 2017], σελ. 49.

Τζίμας, Δ., Επεξεργασία Δεδομένων του Twitter Σχετικά με Παρενέργειες Φαρμάκων στο Κατανεμημένο Σύστημα Storm, Διπλωματική Εργασία, Πολυτεχνείο Κρήτης, Σχολή Ηλεκτρονικών Μηχανικών & Μηχανικών Υπολογιστών, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://purl.tuc.gr/dl/dias/24B7029F-4418-435F-BEBA-7B37DBE46E43>, [προσπελάστηκε στις 24 Ιανουαρίου 2017], σελ. 33.

ΑΡΘΡΑ ΕΦΗΜΕΡΙΔΩΝ

Γκολντστάιν, Ρ., Η άνοδος της ακροδεξιάς: μια δυσσίωνη ηχώ, *Asia Times Online*, μετάφραση: Κουκουμά, Α., [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://rmbnet.gr/details.php?id=821?id=821>, [προσπελάστηκε στις 12 Μαρτίου 2017].

Λάκασας Α., Επαγγελματική εκπαίδευση νέων προσφύγων, *Καθημερινή*, 2017, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: https://www.minedu.gov.gr/publications/docs2017/site_swsto_pemptpi.pdf, [προσπελάστηκε στις 28 Φεβρουαρίου 2017], σελ. 24

Στοιχεία σοκ: Έντεκα χιλιάδες ασυνόδευτα προσφυγόπουλα έχουν εξαφανιστεί, *Αυγή*, 2016, 6 Ιουνίου, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.avgi.gr/article/6594752/stoixeia-sok-enteka-xiliades-asunodeuta-prosfugopoula-exoun-exafanistei>, [προσπελάστηκε στις 6 Φεβρουαρίου 2017].

Τσιγκρής, Α., Άλλο παράνομοι μετανάστες και... άλλο πρόσφυγες και μάλιστα ανήλικοι που ζητούν άσυλο, *Το Βήμα*, 2015, 21 Απριλίου, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.tovima.gr/opinions/article/?aid=696478>, [προσπελάστηκε στις 6 Μαρτίου 2017].

Τσιγκρής, Α., Τα προβλήματα στην παροχή ασύλου σε ασυνόδευτα ανήλικα άτομα και η εκμετάλλευση παιδιών, *Το Βήμα*, 2013, 24 Οκτωβρίου,

[διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.tovima.gr/opinions/article/?aid=536277>, [προσπελάστηκε στις 16 Μαρτίου 2017].

Παναγιωτάκης, Μ., 10 σημεία για τη γεωπολιτική συγκυρία της προσφυγικής κρίσης, Το πιο κρύο καλοκαίρι, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://histologion-gr.blogspot.gr/2016/03/10.html>, [προσπελάστηκε στις 28 Φεβρουαρίου 2017].

Στελλάκη, Φ., Ο Γερμανικός «Νόμος Ενσωμάτωσης» Προσφύγων, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://powerpolitics.eu/o-germanikos-nomos-ensomatwshs-prosfo/>, [προσπελάστηκε στις 8 Φεβρουαρίου 2017].

Huggler, J., Bavaria passes new law to make migrants respect 'dominant' local culture, *The Telegraph*, 2016, 9 Δεκεμβρίου, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.telegraph.co.uk/news/2016/12/09/bavaria-passes-new-law-make-migrants-respect-local-culture/>, [προσπελάστηκε στις 16 Μαρτίου 2017].

ΠΡΑΚΤΙΚΑ ΣΥΝΕΔΡΙΩΝ

Μανούσου, Ε., Χαρτοφύλακα, Τ., Κοινωνικά δίκτυα και μέσα κοινωνικής δικτύωσης στην εξ αποστάσεως τριτοβάθμια εκπαίδευση, 2^ο Πανελλήνιο Συνέδριο, Πάτρα, 28-30/4/2011, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.etpe.gr/custom/pdf/etpe1730.pdf>, [προσπελάστηκε στις 28 Φεβρουαρίου 2017], σελ. 499.

Balahur, A., & Turchi, M., Improving Sentiment Analysis in Twitter Using Multilingual Machine Translated Data., in Proceedings of Recent Advances in Natural Language Processing, 2013, σελ. 49-55.

Birmingham, A. & Smeaton, A., Classifying sentiment in 93 microblogs: is brevity an advantage?, in Proceedings of the 19th ACM international

conference on Information and knowledge management, ACM, 2010, σελ. 1833–1836.

Dmitry, D., Tsur, O., & Rappoport, A., Enhanced sentiment learning using twitter hashtags and smileys., in Proceedings of Coling, 2010, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <http://www.aclweb.org/anthology/C10-2028>, [προσπελάστηκε στις 28 Φεβρουαρίου 2017], [προσπελάστηκε στις 28 Φεβρουαρίου 2017], σελ. 244.

González-Ibáñez, R., Muresan, S., & Wacholder, N., Identifying sarcasm in Twitter: a closer look., in Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: shortpapers (ACL-2011), 2011, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://pdfs.semanticscholar.org/55e3/6d6b45c91a0daa49234bd47b856470d6825c.pdf>, [προσπελάστηκε στις 28 Φεβρουαρίου 2017], σελ. 583.

Nakov, P., Rosenthal, S., Kozareva, Z., Stoyanov, V., Ritter, A., & Wilson, T., Semeval-2013 task 2: Sentiment analysis in twitter, in Proceedings of the Seventh International Workshop on Semantic Evaluation, 2013, σελ. 312-320.

Pang, B., Lee, L., & Vaithyanathan, S., Thumbs up?: sentiment classification using machine learning techniques, in Proceedings of the ACL-02 conference on Empirical methods in natural language processing, 10, 2002, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση: <https://www.cs.cornell.edu/home/llee/papers/sentiment.pdf>, [προσπελάστηκε στις 28 Φεβρουαρίου 2017], σελ. 81-82.

Pope, D., & Griffith, J., An Analysis of Online Twitter Sentiment Surrounding the European Refugee Crisis., in Proceedings of the 8th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, 2016, [διαδίκτυο]. Διαθέσιμο στη διεύθυνση:

https://nuigulresearchday2016.files.wordpress.com/2016/04/nuigul-2016_paper_38.pdf, [προσπελάστηκε στις 22 Φεβρουαρίου 2017].

ΠΑΡΑΡΤΗΜΑ Α

ΠΙΝΑΚΕΣ ΑΠΟΤΕΛΕΣΜΑΤΩΝ

ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΗΣ ΜΕΘΟΔΟΥ ΤΟΥ ΛΕΞΙΚΟΥ

Target Group		Greek Media	
Opinion Target		NONE	
Sentiment	Observed	Predicted	Success Rate [%]
Negative	209	534	39
Positive	580	418	72
Neutral	1659	1496	90

Target Group		Greek Media	
Opinion Target		ΑΚΡΟΔΕΞΙΑ	
Sentiment	Observed	Predicted	Success Rate [%]
Negative	22	16	73
Positive	1	1	100
Neutral	11	17	65

Target Group		Greek Media	
Opinion Target		ΠΟΛΙΤΙΚΗ	
Sentiment	Observed	Predicted	Success Rate [%]
Negative	24	36	67
Positive	55	27	49
Neutral	94	110	85

Target Group		Greek Media	
Opinion Target		ΚΟΙΝΩΝΙΑ	
Sentiment	Observed	Predicted	Success Rate [%]
Negative	53	143	37
Positive	186	122	66
Neutral	505	479	95

Target Group		Greek Media	
Opinion Target		ΠΡΟΣΦΥΓΕΣ	
Sentiment	Observed	Predicted	Success Rate [%]
Negative	187	518	36
Positive	579	417	72
Neutral	1648	1479	90

Target Group		Greek Users	
Opinion Target		NONE	
Sentiment	Observed	Predicted	Success Rate [%]
Negative	4383	4028	92
Positive	3065	2339	76
Neutral	2316	3397	68

Target Group		Greek Users	
Opinion Target		ΑΚΡΟΔΕΞΙΑ	
Sentiment	Observed	Predicted	Success Rate [%]
Negative	530	417	79
Positive	18	5	28
Neutral	37	163	23

Target Group		Greek Users	
Opinion Target		ΠΟΛΙΤΙΚΗ	
Sentiment	Observed	Predicted	Success Rate [%]
Negative	455	545	83
Positive	240	257	93
Neutral	351	296	84

Target Group		Greek Users	
Opinion Target		ΚΟΙΝΩΝΙΑ	
Sentiment	Observed	Predicted	Success Rate [%]
Negative	3022	2814	93
Positive	756	943	80
Neutral	736	1601	46

Target Group		Greek Users	
Opinion Target		ΠΡΟΣΦΥΓΕΣ	
Sentiment	Observed	Predicted	Success Rate [%]
Negative	1695	1366	81
Positive	1514	2123	71
Neutral	1118	1329	84

Target Group		German Media	
Opinion Target		NONE	
Sentiment	Observed	Predicted	Success Rate [%]
Negative	348	277	80
Positive	529	418	79
Neutral	866	1048	83

Target Group		German Media	
Opinion Target		ΑΚΡΟΔΕΞΙΑ	
Sentiment	Observed	Predicted	Success Rate [%]
Negative	2	0	0
Positive	2	2	100
Neutral	4	4	100

Target Group		German Media	
Opinion Target		ΠΟΛΙΤΙΚΗ	
Sentiment	Observed	Predicted	Success Rate [%]
Negative	233	247	94
Positive	230	211	92
Neutral	609	614	99

Target Group		German Media	
Opinion Target		ΚΟΙΝΩΝΙΑ	
Sentiment	Observed	Predicted	Success Rate [%]
Negative	20	12	60
Positive	24	19	79
Neutral	35	48	73

Target Group		German Media	
Opinion Target		ΠΡΟΣΦΥΓΕΣ	
Sentiment	Observed	Predicted	Success Rate [%]
Negative	174	133	76
Positive	274	217	79
Neutral	433	531	82

Target Group		German Users	
Opinion Target		NONE	
Sentiment	Observed	Predicted	Success Rate [%]
Negative	3156	1671	53
Positive	2565	2160	84
Neutral	2778	4668	60

Target Group		German Users	
Opinion Target		ΑΚΡΟΔΕΞΙΑ	
Sentiment	Observed	Predicted	Success Rate [%]
Negative	46	22	48
Positive	20	19	95
Neutral	11	36	31

Target Group		German Users	
Opinion Target		ΠΟΛΙΤΙΚΗ	
Sentiment	Observed	Predicted	Success Rate [%]
Negative	1351	1003	74
Positive	1058	930	88
Neutral	1746	2222	79

Target Group		German Users	
Opinion Target		ΚΟΙΝΩΝΙΑ	
Sentiment	Observed	Predicted	Success Rate [%]
Negative	117	60	51
Positive	109	109	100
Neutral	125	182	69

Target Group		German Users	
Opinion Target		ΠΡΟΣΦΥΓΕΣ	
Sentiment	Observed	Predicted	Success Rate [%]
Negative	2167	1074	50
Positive	1801	1519	84
Neutral	1755	3130	56

Target Group	Greek Media: 2448		Greek Users: 9764	
Opinion Target	Observed	Frequency [%]	Observed	Frequency [%]
Ακροδεξιά	34	1.39	585	5.99
Πολιτική	173	7.07	1046	10.71
Κοινωνία	744	30.39	5358	54.88
Πρόσφυγες	2414	98.61	4818	49.34

Target Group	German Media: 1743		German Users: 8499	
Opinion Target	Observed	Frequency [%]	Observed	Frequency [%]
Ακροδεξιά	8	0.46	77	0.91
Πολιτική	1072	61.50	4155	48.89
Κοινωνία	79	4.53	351	4.13
Πρόσφυγες	881	50.55	5723	67.34

ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΗΣ ΜΕΘΟΔΟΥ ΤΟΥ NAIVE BAYES

- Πραγματικά Tweets

Target Group		German Users	
Total Tweets		8499	
Total Words		147746	
Root Letters	Common Words	Computational Time [min]	Success Rate [%]
2	449	15.26	43.82
3	2055	60.61	44.80
4	4134	123.12	44.29
Full Length	10555	294.82	44.17

Target Group		German Media	
Total Tweets		1743	
Total Words		27572	
Root Letters	Common Words	Computational Time [min]	Success Rate [%]
2	294	2.03	20.65
3	1094	5.86	21.03
4	1806	9.73	21.03
Full Length	3404	17.23	21.03

Target Group		Greek Users	
Total Tweets		9764	
Total Words		95622	
Root Letters	Common Words	Computational Time [min]	Success Rate [%]
2	401	13.95	27.07
3	1087	35.40	21.54
4	1813	56.95	18.23
Full Length	3718	127.36	18.95

Target Group		Greek Media	
Total Tweets		2448	
Total Words		17910	
Root Letters	Common Words	Computational Time [min]	Success Rate [%]
2	340	3.07	9.81
3	739	6.02	9.81
4	1065	9.02	9.81
Full Length	1862	14.61	9.81

- Συνθετικά Tweets

Tested Factor		Sample Size (number of tweets)	
Sentiment Threshold (pos/neg words out of the total words)		60%	
Test Set Size (out of the total sample size)		30%	
Tweets	Computational Time [sec]	Success Rate [%]	
50	1.09	53.33	
100	1.16	73.33	
200	1.36	75.00	
500	1.66	76.67	
800	2.05	74.17	
1000	2.34	79.00	
2000	3.74	83.00	
5000	7.64	81.00	
10000	14.48	81.97	
20000	26.90	81.77	
30000	39.76	82.33	
40000	53.73	82.77	
50000	67.86	82.85	
60000	83.87	82.56	
70000	92.11	82.99	
80000	105.03	82.29	
90000	122.35	82.74	
100000	145.09	82.47	

Tested Factor	Sentiment Threshold
Sample Size (number of tweets)	10000
Test Set Size (out of the total sample size)	30%
Sentiment Threshold [%]	Success Rate [%]
55	79.30
60	82.43
65	93.80
70	97.03
75	98.10
80	98.83
85	100
90	100
95	100

Tested Factor	Test Set Size
Sample Size (number of tweets)	10000
Sentiment Threshold (pos/neg words out of the total words)	60%
Test Set Size [%]	Success Rate [%]
10	82.40
20	82.20
30	81.87
40	80.80
50	82.80
60	80.73
70	81.29
80	80.85
90	80.71

ΣΥΓΚΡΙΤΙΚΟΣ ΠΙΝΑΚΑΣ ΕΠΙΔΟΣΗΣ ΜΕΘΟΔΩΝ ΛΕΞΙΚΟΥ ΚΑΙ NAIVE BAYES

Αρχική μέτρηση

Method	Naive Bayes	Lexicon
Target Group	Total Success Rate [%]	
Greek Media	8.54	81.53
Greek Users	33.68	81.38
German Media	18.77	80.93
German Users	29.99	64.53

Μέτρηση μετά τις τροποποιήσεις

Method	Naive Bayes	Lexicon
Target Group	Total Success Rate [%]	
Greek Media	9.81	81.53
Greek Users	18.95	81.38
German Media	21.03	80.93
German Users	44.17	64.53

ΠΑΡΑΡΤΗΜΑ Β

ΚΩΔΙΚΑΣ

Κώδικας για τα ελληνικά tweets

- **Σύνδεση με το Twitter API**

Setting up Twitter Authentication

```
library(ROAuth)
```

```
library(twitterR)
```

```
library(methods)
```

```
api_key <- '...'
```

```
api_secret <- '...'
```

```
access_token <- '...'
```

```
access_token_secret <- '...'
```

```
setup_twitter_oauth(api_key,
```

```
  api_secret,
```

```
  access_token,
```

```
  access_token_secret)
```

- **Λέξεις-κλειδιά για το κατέβασμα των tweets**

Read in Keywords

```
keywords <- readLines('mykeys.txt', encoding='UTF-8')
```

- **Κατέβασμα των tweets**

Function Definition

```
search_all_keyw <- function(keywords)
```

```
{
```

```
  lapply(keywords[sapply(keywords, length) > 0], function(kw){
```

```
    rfg <- try( searchTwitter(kw, n=1500, locale='el_GR.UTF-8',
```

```
      lang='el'))
```

```
    if (length(rfg) != 0) {
```

```
      rfgdf = twListToDF(rfg)
```

```

fname <- paste0(paste0('twit_search_', kw, '_'), strftime(Sys.time(), format
                    = "%j_%H%M"), '.csv')

write.table(rfgdf, file = fname, row.names = F, sep=',', fileEncoding='UTF-8')

} else {'No tweets found!'}

})

}

test1 <- search_all_keyw(keywords = keywords)

```

- **Εύρεση αρχείου**

```
arxeia<-dir(pattern='^tw.*csv$')
```

- **Διάβασμα αρχείου**

```
tweets_gr <- read.csv(arxeia[1], header=T, stringsAsFactors = F, sep=';',fileEncoding="ISO-8859-7")
```

- **Λίστα με τα ονόματα των media**

```
media_gr <- readLines('media_gr', encoding='UTF-8' )
```

- **Διαχωρισμός των tweets σε media και users**

```
tweets_gr_media <- tweets_gr[tweets_gr$username %in% media_gr, ]
```

```
tweets_gr_users <- tweets_gr[!(tweets_gr$username %in% media_gr), ]
```

- **Διαχωρισμός των tweets με βάση τους opinion targets**

#Split tweets

#Radical right

```
radicalright_gr <- readLines('radicalright_gr', encoding='UTF-8')
```

```
rows_found_radicalright <- lapply(radicalright_gr, grep, tweets_gr_users$text)
```

```
rows_found_list_radicalright <- unlist(rows_found_radicalright)
```

```
tweets_gr_users_radicalright <- tweets_gr_users[rows_found_list_radicalright, ]
```

```
tweets_gr_users_radicalright[, 'text']
```

#Politics

```
politics_gr <- readLines('politics_gr', encoding='UTF-8')
```

```
rows_found_politics <- lapply(politics_gr, grep, tweets_gr_users$text)
```

```
rows_found_list_politics <- unlist(rows_found_politics)
```

```
tweets_gr_users_politics <- tweets_gr_users[rows_found_list_politics, ]
tweets_gr_users_politics_final <- tweets_gr_users_politics[-rows_found_list_radicalright, ]
tweets_gr_users_politics_final[, 'text']
```

#Society

```
society_gr <- readLines('society_gr', encoding='UTF-8')
rows_found_society <- lapply(society_gr, grep, tweets_gr_users$text)
rows_found_list_society <- unlist(rows_found_society)
tweets_gr_users_society <- tweets_gr_users[rows_found_list_society, ]
tweets_gr_users_society_semifinal <- tweets_gr_users_society[-rows_found_list_radicalright, ]
tweets_gr_users_society_final <- tweets_gr_users_society_semifinal[-rows_found_list_politics, ]
tweets_gr_users_society_final[, 'text']
```

#Refugees

```
tweets_gr_users_refugees <- tweets_gr_users[-rows_found_list_radicalright, ]
tweets_gr_users_refugees_semifinal <- tweets_gr_users_refugees[-rows_found_list_politics, ]
tweets_gr_users_refugees_final <- tweets_gr_users_refugees_semifinal[-rows_found_list_society, ]
tweets_gr_users_refugees_final[, 'text']
```

- **Καθαρισμός media tweets**

```
library(tm)
tweetsDataCorpus_gr_media <- VCorpus(VectorSource(tweets_gr_media$text))
```

#Transform to lower

```
tweetsDataCorpus_gr_media_clean <- tm_map(tweetsDataCorpus_gr_media,
content_transformer(tolower))
```

#Replace «σ» with an «ς»

```
tweetsDataCorpus_gr_media_clean <- tm_map(tweetsDataCorpus_gr_media_clean,
content_transformer(gsub), pattern = "σ ", replacement = "ς ")
```


#Remove numbers

```
tweetsDataCorpus_gr_media_clean <- tm_map(tweetsDataCorpus_gr_media_clean,
removeNumbers)
```

- **Διαγραφή stopwords**

```
grsw <- read.csv("gr_stop_words_UC", header=F, stringsAsFactors = F, sep=',')
```

```
grswCorpus <- VCorpus(VectorSource(grsw))
```

```
grswCorpus_clean <- tm_map(grswCorpus, content_transformer(tolower))
```

```
text = sapply(grswCorpus_clean, as.character)
```

```
vtext = as.vector(text)
```

```
vtext <- sub('σ$', "ς", vtext)
```

```
vtext <- sub('Δ', "δ", vtext)
```

```
vtext <- sub('^ ', "", vtext)
```

- **Διαγραφή υπερσυνδέσμων**

```
tweetsDataCorpus_gr_media_clean <- tm_map(tweetsDataCorpus_gr_media_clean,
content_transformer(gsub), pattern = "https:.*", replacement = "")
```

- **Διαγραφή mentions**

```
tweetsDataCorpus_gr_media_clean <- tm_map(tweetsDataCorpus_gr_media_clean,
content_transformer(gsub), pattern = "@[a-z,A-Z,0-9,:]*", replacement = "")
```

- **Διαγραφή των «rt»**

```
tweetsDataCorpus_gr_media_clean <- tm_map(tweetsDataCorpus_gr_media_clean,
removeWords, c("rt"))
```

- **Διαγραφή των σημείων στίξης**

```
#tweetsDataCorpus_gr_media_clean <- tm_map(tweetsDataCorpus_gr_media_clean,
removePunctuation)
```

```
tweetsDataCorpus_gr_media_clean <- tm_map(tweetsDataCorpus_gr_clean,
content_transformer(gsub), pattern = "[!\"#$%&'()*+,-./:;<=>?[\\^_`{|}~.]", replacement = "
")
```

- **Απομάκρυνση τονισμού**

```
tweetsDataCorpus_gr_media_cleanNoAcc <- tm_map(tweetsDataCorpus_gr_media_clean,
content_transformer( function(x) gsub("ά", "α", x)))
```

```
tweetsDataCorpus_gr_media_cleanNoAcc <-
tm_map(tweetsDataCorpus_gr_media_cleanNoAcc, content_transformer( function(x)
gsub("έ", "ε", x)))
```

```
tweetsDataCorpus_gr_media_cleanNoAcc <-
tm_map(tweetsDataCorpus_gr_media_cleanNoAcc, content_transformer( function(x)
gsub("ή", "η", x)))
```

```
tweetsDataCorpus_gr_media_cleanNoAcc <-
tm_map(tweetsDataCorpus_gr_media_cleanNoAcc, content_transformer( function(x)
gsub("ί", "ι", x)))
```

```
tweetsDataCorpus_gr_media_cleanNoAcc <-
tm_map(tweetsDataCorpus_gr_media_cleanNoAcc, content_transformer( function(x)
gsub("ό", "ο", x)))
```

```
tweetsDataCorpus_gr_mediacleannoAcc <-
tm_map(tweetsDataCorpus_gr_media_cleanNoAcc, content_transformer( function(x)
gsub("ύ", "υ", x)))
```

```
tweetsDataCorpus_gr_mediacleannoAcc <-
tm_map(tweetsDataCorpus_gr_media_cleanNoAcc, content_transformer( function(x)
gsub("ώ", "ω", x)))
```

- **Καθαρισμός users tweets**

```
library(tm)
```

```
tweetsDataCorpus_gr_users <- VCorpus(VectorSource(tweets_gr_users$text))
```

#Transform to lower

```
tweetsDataCorpus_gr_users_clean <- tm_map(tweetsDataCorpus_gr_users,
content_transformer(tolower))
```

#Replace «σ» with an «ς»

```
tweetsDataCorpus_gr_users_clean <- tm_map(tweetsDataCorpus_gr_users_clean,
content_transformer(gsub), pattern = "σ ", replacement = "ς ")
```

#Remove numbers

```
tweetsDataCorpus_gr_users_clean <- tm_map(tweetsDataCorpus_gr_users_clean,
removeNumbers)
```

- **Διαγραφή stopwords**

```
grsw <- read.csv("gr_stop_words_UC", header=F, stringsAsFactors = F, sep=',')
```

```
grswCorpus <- VCorpus(VectorSource(grsw))
```

```
grswCorpus_clean <- tm_map(grswCorpus, content_transformer(tolower))
```

```
text = sapply(grswCorpus_clean, as.character)
```

```
vtext = as.vector(text)
```

```
vtext <- sub('σ$', "ς", vtext)
```

```
vtext <- sub('Δ', "δ", vtext)
```

```
vtext <- sub('^ ', "", vtext)
```

- **Διαγραφή υπερσυνδέσμων**

```
tweetsDataCorpus_gr_users_clean <- tm_map(tweetsDataCorpus_gr_users_clean,
content_transformer(gsub), pattern = "https:.*", replacement = "")
```

- **Διαγραφή mentions**

```
tweetsDataCorpus_gr_users_clean <- tm_map(tweetsDataCorpus_gr_users_clean,
content_transformer(gsub), pattern = "@[a-z,A-Z,0-9,:]*", replacement = "")
```

- **Διαγραφή των «rt»**

```
tweetsDataCorpus_gr_users_clean <- tm_map(tweetsDataCorpus_gr_users_clean,
removeWords, c("rt"))
```

- **Διαγραφή των σημείων στίξης**

```
#tweetsDataCorpus_gr_users_clean <- tm_map(tweetsDataCorpus_gr_users_clean,
removePunctuation)
```

```
tweetsDataCorpus_gr_users_clean <- tm_map(tweetsDataCorpus_gr_clean,
content_transformer(gsub), pattern = "[!\"#$%&'()*+,-./:;<=>?[\\"^_`{|}~.]", replacement = ""
")
```

- **Απομάκρυνση τονισμού**

```
tweetsDataCorpus_gr_users_cleanNoAcc <- tm_map(tweetsDataCorpus_gr_users_clean,
content_transformer( function(x) gsub("ά", "α", x)))
```

```
tweetsDataCorpus_gr_users_cleanNoAcc <-
tm_map(tweetsDataCorpus_gr_users_cleanNoAcc, content_transformer( function(x)
gsub("έ", "ε", x)))
```

```
tweetsDataCorpus_gr_users_cleanNoAcc <-
tm_map(tweetsDataCorpus_gr_users_cleanNoAcc, content_transformer( function(x)
gsub("ή", "η", x)))
```

```
tweetsDataCorpus_gr_users_cleanNoAcc <-
tm_map(tweetsDataCorpus_gr_users_cleanNoAcc, content_transformer( function(x)
gsub("ί", "ι", x)))
```

```
tweetsDataCorpus_gr_users_cleanNoAcc <-
tm_map(tweetsDataCorpus_gr_users_cleanNoAcc, content_transformer( function(x)
gsub("ό", "ο", x)))
```

```
tweetsDataCorpus_gr_userscleanNoAcc <-
tm_map(tweetsDataCorpus_gr_users_cleanNoAcc, content_transformer( function(x)
gsub("ύ", "υ", x)))
```

```
tweetsDataCorpus_gr_userscleanNoAcc <-
tm_map(tweetsDataCorpus_gr_users_cleanNoAcc, content_transformer( function(x)
gsub("ώ", "ω", x)))
```

- **Ανάλυση συναισθήματος με τη μέθοδο του λεξικού**

#Lexicon based sentiment analysis

#Media tweets sentiment

```
library(stringr)

tweets_dtm <- DocumentTermMatrix(tweetsDataCorpus_gr_media_cleanNoAcc)

as.matrix(tweets_dtm[76,])

gr.pos <- readLines('greek_positive_words', encoding='UTF-8' )
gr.pos <- gsub('-', '\\\\w+', gr.pos)
gr.pos <- gsub('\\^', '\\\\s+', gr.pos)

gr.neg <- readLines('greek_negative_words', encoding='UTF-8' )
gr.neg <- gsub('-', '\\\\w+', gr.neg)
gr.neg <- gsub('\\^', '\\\\s+', gr.neg)

str.match <- function(pattern, x)
{ no_matches = 0
for (i in 1:length(pattern)) {
  matches <- str_extract_all(x, pattern[i])
  no_matches <- no_matches + length(matches[[1]])
  if (length(matches[[1]]) > 0)
    print(matches[[1]])
}
return(no_matches)
}

global_media_pos <- 0
global_media_neg <- 0
global_media_neu <- 0
```

```

for (i in 1:length(tweetsDataCorpus_gr_media_cleanNoAcc)) {

  no_pos_matches <- str.match(gr.pos,
tweetsDataCorpus_gr_media_cleanNoAcc[[i]]$content)

  no_neg_matches <- str.match(gr.neg, paste0(" ",
tweetsDataCorpus_gr_media_cleanNoAcc[[i]]$content))

  if (no_pos_matches > no_neg_matches)

    {global_media_pos <- global_media_pos + 1}

  else if (no_pos_matches < no_neg_matches)

    {global_media_neg <- global_media_neg + 1}

  else

    {global_media_neu <- global_media_neu + 1}

  cat(i, ":", no_pos_matches, ":", no_neg_matches, "\n")

  print("-----")

}

#Users tweets sentiment

library(stringr)

tweets_dtm <- DocumentTermMatrix(tweetsDataCorpus_gr_users_cleanNoAcc)

as.matrix(tweets_dtm[1,])

gr.pos <- readLines('greek_positive_words', encoding='UTF-8' )

gr.pos <- gsub('-', '\\\\w+', gr.pos)

gr.pos <- gsub('\\^', '\\\\s+', gr.pos)

gr.neg <- readLines('greek_negative_words', encoding='UTF-8' )

gr.neg <- gsub('-', '\\\\w+', gr.neg)

gr.neg <- gsub('\\^', '\\\\s+', gr.neg)

str.match <- function(pattern, x)

{ no_matches = 0

for (i in 1:length(pattern)) {

  matches <- str_extract_all(x, pattern[i])

```

```

no_matches <- no_matches + length(matches[[1]])
if (length(matches[[1]]) > 0)
  print(matches[[1]])
}
return(no_matches)
}

global_media_pos <- 0
global_media_neg <- 0
global_media_neu <- 0

for (i in 1:length(tweetsDataCorpus_gr_users_cleanNoAcc)) {
  no_pos_matches <- str.match(gr.pos,
tweetsDataCorpus_gr_users_cleanNoAcc[[i]]$content)

  no_neg_matches <- str.match(gr.neg, paste0(" ",
tweetsDataCorpus_gr_users_cleanNoAcc[[i]]$content))

  if (no_pos_matches > no_neg_matches)
    {global_media_pos <- global_media_pos + 1}
  else if (no_pos_matches < no_neg_matches)
    {global_media_neg <- global_media_neg + 1}
  else
    {global_media_neu <- global_media_neu + 1}

  cat(i, ":", no_pos_matches, ":", no_neg_matches, "\n")
  print("-----")
}

```

Naive Bayes για τα ελληνικά media

Loading automatically the required package

Implementing the Naive bayes approach

```
library(e1071)
```

```
file<-tweets_gr_media
```

```
file[,2]<-dataframe
```

```
file[,3]<-tweets_gr_media$sentiment
```

Finding the number of tweets

```
num_tweets<-dim(file)[1]
```

```
num_tweets
```

Determining the number of words for each tweet

```
num_words<-matrix(nrow=num_tweets,ncol=1)
```

```
for(i in 1:num_tweets)
```

```
{
```

```
  num_words[i]<-length(strsplit(as.character(file[i,2])," ")[[1]])
```

```
}
```

```
num_words
```

Finding the total number of words (regarding all available tweets)

```
tot_num_words<-sum(num_words)
```

```
tot_num_words
```

Building a matrix comprising the total number of words

```
words<-matrix(nrow=tot_num_words,ncol=1)
```

```
k<-1
```

```
for(i in 1:num_tweets)
```

```
{
```

```
  for(j in 1:num_words[i])
```

```
  {
```

```

words[k]<-strsplit(as.character(file[i,2])," ")[[1]][j]

k<-k+1

}

}

# Building a matrix comprising the words without repetitions

common_words<-unique(words)

common_words

# Finding the number of common words

num_com_words<-length(common_words)

# Building the text-word matrix

tweet_words=matrix(nrow=num_tweets,ncol=num_com_words)

for(i in 1:num_tweets)

{

  for(j in 1:num_com_words)

  {

    tweet_words[i,j]<-length(which(strsplit(as.character(file[i,2]),"
")[[1]]==common_words[j]))

  }

}

# Performing the Naive bayes technique

naiveBayes.train<-
naiveBayes(tweet_words[1:(num_tweets/3),1:(num_tweets/3)],file[1:(num_tweets/3),3],lap
lace=1)

results<-predict(naiveBayes.train,tweet_words)

# Estimating the success rate

success<-matrix(nrow=num_tweets,ncol=1)

for (i in 1:num_tweets)

{

  if (results[i]==file[i,3])

```



```

{ success[i]<-1 }
else { success[i]<-0 }
}
accuracy<-sum(success)/num_tweets
accuracy
write(dataframe3[,1],file="tweets_de_media.csv",ncolumns=1)
tweets_gr_media_unclean<-matrix(nrow=dim(tweets_gr_media)[1],ncol=4)
for (i in 1:dim(tweets_de_media)[1])
{
tweets_gr_media_unclean[i,1]<-tweets_gr_media[i,1]
tweets_gr_media_unclean[i,2]<-tweets_gr_media[i,2]
tweets_gr_media_unclean[i,3]<-tweets_gr_media[i,3]
tweets_gr_media_unclean[i,4]<-tweets_gr_media[i,4]
}
write.csv(tweets_gr_media_unclean,file="tweets_gr_media_unclean.csv")

```

Naive Bayes για τους Έλληνες χρήστες

Loading automatically the required package

```
# implementing the Naive bayes approach
```

```
library(e1071)
```

```
file<-tweets_gr_users
```

```
file[,2]<-dataframe
```

```
file[,3]<-tweets_gr_users$sentiment
```

Finding the number of tweets

```
num_tweets<-dim(file)[1]
```

```
num_tweets
```

Determining the number of words for each tweet

```
num_words<-matrix(nrow=num_tweets,ncol=1)
```

```

for(i in 1:num_tweets)
{
  num_words[i]<-length(strsplit(as.character(file[i,2])," ")[[1]])
}
num_words
# Finding the total number of words (regarding all available tweets)
tot_num_words<-sum(num_words)
tot_num_words
# Building a matrix comprising the total number of words
words<-matrix(nrow=tot_num_words,ncol=1)
k<-1
for(i in 1:num_tweets)
{
  for(j in 1:num_words[i])
  {
    words[k]<-strsplit(as.character(file[i,2])," ")[[1]][j]
    k<-k+1
  }
}
# Building a matrix comprising the words without repetitions
common_words<-unique(words)
common_words
# Finding the number of common words
num_com_words<-length(common_words)
# Building the text-word matrix
tweet_words=matrix(nrow=num_tweets,ncol=num_com_words)
for(i in 1:num_tweets)

```

```

{
  for(j in 1:num_com_words)
  {
    tweet_words[i,j]<-length(which(strsplit(as.character(file[i,2]),"
    ")[[1]]==common_words[j]))
  }
}

```

Performing the Naive bayes technique

```

naiveBayes.train<-
naiveBayes(tweet_words[1:(num_tweets/3),1:(num_tweets/3)],file[1:(num_tweets/3),3],lap
lace=1)

```

```

results<-predict(naiveBayes.train,tweet_words)

```

Estimating the success rate

```

success<-matrix(nrow=num_tweets,ncol=1)

```

```

for (i in 1:num_tweets)

```

```

{

```

```

  if (results[i]==file[i,3])

```

```

  { success[i]<-1 }

```

```

  else { success[i]<-0 }

```

```

}

```

```

accuracy<-sum(success)/num_tweets

```

```

accuracy

```

```

write(dataframe3[,1],file="tweets_de_users.csv",ncolumns=1)

```

```

tweets_gr_users_unclean<-matrix(nrow=dim(tweets_gr_users)[1],ncol=4)

```

```

for (i in 1:dim(tweets_de_users)[1])

```

```

{

```

```

  tweets_gr_users_unclean[i,1]<-tweets_gr_users[i,1]

```

```

  tweets_gr_users_unclean[i,2]<-tweets_gr_users[i,2]

```

```

  tweets_gr_users_unclean[i,3]<-tweets_gr_users[i,3]

```

```
tweets_gr_users_unclean[i,4]<-tweets_gr_users[i,4]
}
write.csv(tweets_gr_users_unclean,file="tweets_gr_users_unclean.csv")
```

Κώδικας για τα γερμανικά tweets

- **Σύνδεση με το Twitter API**

Setting up Twitter Authentication

```
library(ROAuth)
library(twitterR)
library(methods)
api_key <- '...'
api_secret <- '...'
access_token <- '...'
access_token_secret <- '...'
setup_twitter_oauth(api_key,
                    api_secret,
                    access_token,
                    access_token_secret)
```

- **Λέξεις κλειδιά για το κατέβασμα των tweets**

Read in Keywords

```
keywords <- readLines('mykeys.txt', encoding='UTF-8')
```

- **Κατέβασμα των tweets**

Function Definition

```
search_all_keyw <- function(keywords)
{
  lapply(keywords[sapply(keywords, length) > 0], function(kw){
    rfg <- try( searchTwitter(kw, n=1500, locale='de_DE.UTF-8',
                             lang='el'))
    if (length(rfg) != 0) {
      rfgdf = twListToDF(rfg)
      fname <- paste0(paste0('twit_search_', kw, '_'), strftime(Sys.time(), format
```

```

= "%j_%H%M"), '.csv')

write.table(rfgdf, file = fname, row.names = F, sep=',', fileEncoding='UTF-8')

} else {'No tweets found!'}

})

}

test1 <- search_all_keyw(keywords = keywords)

```

- **Εύρεση αρχείου**

```
arxeia<-dir(pattern='^tw.*csv$')
```

- **Διάβασμα αρχείου**

```
tweets_de <- read.csv("tweets_de.csv", header=T, stringsAsFactors = F,
sep=';',fileEncoding="ISO-8859-7")
```

- **Λίστα με τα ονόματα των media**

```
media_de <- readLines('media_de', encoding='UTF-8' )
```

- **Διαχωρισμός των tweets σε media και users**

```
tweets_de_media <- tweets_de[tweets_de$username %in% media_de, ]
```

```
tweets_de_users <- tweets_de[!(tweets_de$username %in% media_de), ]
```

- **Διαχωρισμός των tweets με βάση τους opinion targets**

```
#Split tweets
```

```
#Radical right
```

```
radicalright_de <- readLines('radicalright_de', encoding='UTF-8')
```

```
rows_found_radicalright <- lapply(radicalright_de, deep, tweets_de_users$text)
```

```
rows_found_list_radicalright <- unlist(rows_found_radicalright)
```

```
tweets_de_users_radicalright <- tweets_de_users[rows_found_list_radicalright, ]
```

```
tweets_de_users_radicalright[, 'text']
```

```
#Politics
```

```
politics_de <- readLines('politics_de', encoding='UTF-8')
```

```
rows_found_politics <- lapply(politics_de, deep, tweets_de_users$text)
```

```
rows_found_list_politics <- unlist(rows_found_politics)
```

```
tweets_de_users_politics <- tweets_de_users[rows_found_list_politics, ]
```

```
tweets_de_users_politics_final <- tweets_de_users_politics[-rows_found_list_radicalright, ]
```

```
tweets_de_users_politics_final[, 'text']
```

#Society

```
society_de <- readLines('society_de', encoding='UTF-8')
```

```
rows_found_society <- lapply(society_de, deep, tweets_de_users$text)
```

```
rows_found_list_society <- unlist(rows_found_society)
```

```
tweets_de_users_society <- tweets_de_users[rows_found_list_society, ]
```

```
tweets_de_users_society_semifinal <- tweets_de_users_society[-  
rows_found_list_radicalright, ]
```

```
tweets_de_users_society_final <- tweets_de_users_society_semifinal[-  
rows_found_list_politics, ]
```

```
tweets_de_users_society_final[, 'text']
```

#Refugees

```
tweets_de_users_refugees <- tweets_de_users[-rows_found_list_radicalright, ]
```

```
tweets_de_users_refugees_semifinal <- tweets_de_users_refugees[-  
rows_found_list_politics, ]
```

```
tweets_de_users_refugees_final <- tweets_de_users_refugees_semifinal[-  
rows_found_list_society, ]
```

```
tweets_de_users_refugees_final[, 'text']
```

- **Καθαρισμός media tweets**

```
tweetsDataCorpus_de_media <- VCorpus(VectorSource(tweets_de_media$text))
```

#Transform to lower

```
tweetsDataCorpus_de_media_clean <- tm_map(tweetsDataCorpus_de_media,  
content_transformer(tolower))
```

```
tweetsDataCorpus_de_media_clean <- tm_map(tweetsDataCorpus_de_media_clean,  
removeNumbers)
```

#Remove links

```
tweetsDataCorpus_de_media_clean <- tm_map(tweetsDataCorpus_de_media_clean,  
content_transformer(gsub), pattern = "https:.*", replacement = "")
```

```
tweetsDataCorpus_de_media_clean <- tm_map(tweetsDataCorpus_de_media_clean,  
content_transformer(gsub), pattern = "http:.*", replacement = "")
```

- **Διαγραφή stopwords**

```
desw <- readLines('de_stop_words_UC', encoding='ISO-8859-1')

deswCorpus <- VCorpus(VectorSource(desw))

deswCorpus_clean <- tm_map(deswCorpus, content_transformer(tolower))

text = sapply(deswCorpus_clean, as.character)

vtext = as.vector(text)

tweetsDataCorpus_de_media_cleanNoAcc <- tm_map(tweetsDataCorpus_de_media_clean,
                                               removeWords,vtext)
```

#Remove numbers

```
tweetsDataCorpus_de_media_clean <- tm_map(tweetsDataCorpus_de_media_clean,
removeNumbers)
```

- **Διαγραφή υπερσυνδέσμων**

```
tweetsDataCorpus_de_media_clean <- tm_map(tweetsDataCorpus_de_media_clean,
content_transformer(gsub), pattern = "https:.*", replacement = "")
```

- **Διαγραφή των σημείων στίξης**

```
#tweetsDataCorpus_de_media_clean <- tm_map(tweetsDataCorpus_de_media_clean,
removePunctuation)
```

```
tweetsDataCorpus_de_media_clean <- tm_map(tweetsDataCorpus_de_clean,
content_transformer(gsub), pattern = "[!\"#$%&'()*+,-./:;<=>?[\\^_`{|}~.]", replacement = ""
")
```

- **Διαγραφή mentions**

```
tweetsDataCorpus_de_media_clean <- tm_map(tweetsDataCorpus_de_media_clean,
content_transformer(gsub), pattern = "@[a-z,A-Z,0-9,:]*", replacement = "")
```

- **Διαγραφή των «rt»**

```
tweetsDataCorpus_de_media_clean <- tm_map(tweetsDataCorpus_de_media_clean,
removeWords, c("rt"))
```

```
#tweetsDataCorpus_de_media_clean <- tm_map(tweetsDataCorpus_de_media_clean,
removePunctuation)
```

```
tweetsDataCorpus_de_media_clean <- tm_map(tweetsDataCorpus_de_clean,
content_transformer(gsub), pattern = "[!\"#$%&'()*+,-./:;<=>?[\\^_`{|}~.]", replacement = ""
")
```


- **Ανάλυση συναισθήματος με τη μέθοδο του λεξικού**

#Lexicon based sentiment analysis

#Media tweets sentiment

```

library(stringr)

tweets_dtm <- DocumentTermMatrix(tweetsDataCorpus_de_media_cleanNoAcc)

as.matrix(tweets_dtm[76,])

de.pos <- readLines('deeeek_positive_words', encoding='UTF-8' )
de.pos <- gsub('-', '\\\\w+', de.pos)
de.pos <- gsub("\\^", '\\\\s+', de.pos)

de.neg <- readLines('deeeek_negative_words', encoding='UTF-8' )
de.neg <- gsub('-', '\\\\w+', de.neg)
de.neg <- gsub("\\^", '\\\\s+', de.neg)

str.match <- function(pattern, x)
{ no_matches = 0
  for (i in 1:length(pattern)) {
    matches <- str_extract_all(x, pattern[i])
    no_matches <- no_matches + length(matches[[1]])
    if (length(matches[[1]]) > 0)
      print(matches[[1]])
  }
  return(no_matches)
}

global_media_pos <- 0
global_media_neg <- 0
global_media_neu <- 0

for (i in 1:length(tweetsDataCorpus_de_media_cleanNoAcc)) {
  no_pos_matches <- str.match(de.pos,
tweetsDataCorpus_de_media_cleanNoAcc[[i]]$content)

```

```

no_neg_matches <- str.match(de.neg, paste0(" ",
tweetsDataCorpus_de_media_cleanNoAcc[[i]]$content))

if (no_pos_matches > no_neg_matches)

{global_media_pos <- global_media_pos + 1}

else if (no_pos_matches < no_neg_matches)

{global_media_neg <- global_media_neg + 1}

else

{global_media_neu <- global_media_neu + 1}

cat(i, ":", no_pos_matches, ":", no_neg_matches, "\n")

print("-----")
}

```

#Users tweets sentiment

```

library(stringr)

tweets_dtm <- DocumentTermMatrix(tweetsDataCorpus_de_users_cleanNoAcc)

as.matrix(tweets_dtm[1,])

de.pos <- readLines('deek_positive_words', encoding='UTF-8' )

de.pos <- gsub('-', '\\\\w+', de.pos)

de.pos <- gsub("\\^", '\\\\s+', de.pos)

de.neg <- readLines('deek_negative_words', encoding='UTF-8' )

de.neg <- gsub('-', '\\\\w+', de.neg)

de.neg <- gsub("\\^", '\\\\s+', de.neg)

str.match <- function(pattern, x)

{ no_matches = 0

for (i in 1:length(pattern)) {

matches <- str_extract_all(x, pattern[i])

no_matches <- no_matches + length(matches[[1]])

if (length(matches[[1]]) > 0)

print(matches[[1]])
}
}

```

```

}

return(no_matches)

}

global_media_pos <- 0

global_media_neg <- 0

global_media_neu <- 0

for (i in 1:length(tweetsDataCorpus_de_users_cleanNoAcc)) {

  no_pos_matches <- str.match(de.pos,
tweetsDataCorpus_de_users_cleanNoAcc[[i]]$content)

  no_neg_matches <- str.match(de.neg, paste0(" ",
tweetsDataCorpus_de_users_cleanNoAcc[[i]]$content))

  if (no_pos_matches > no_neg_matches)

    {global_media_pos <- global_media_pos + 1}

  else if (no_pos_matches < no_neg_matches)

    {global_media_neg <- global_media_neg + 1}

  else

    {global_media_neu <- global_media_neu + 1}

  cat(i, ":", no_pos_matches, ":", no_neg_matches, "\n")

  print("-----")

}

```

- **Καθαρισμός users tweets**

```
tweetsDataCorpus_de_users <- VCorpus(VectorSource(tweets_de_users$text))
```

```
tweetsDataCorpus_de_users_clean <- tm_map(tweetsDataCorpus_de_users,
content_transformer(tolower))
```

```
tweetsDataCorpus_de_users_clean <- tm_map(tweetsDataCorpus_de_users_clean,
removeNumbers)
```

#Remove links

```
tweetsDataCorpus_de_users_clean <- tm_map(tweetsDataCorpus_de_users_clean,
content_transformer(gsub), pattern = "https:.*", replacement = "")
```

```
tweetsDataCorpus_de_users_clean <- tm_map(tweetsDataCorpus_de_users_clean,
content_transformer(gsub), pattern = "http:.*", replacement = "")
```

#Remove numbers

```
tweetsDataCorpus_de_users_clean <- tm_map(tweetsDataCorpus_de_users_clean,
removeNumbers)
```

- **Διαγραφή υπερσυνδέσμων**

```
tweetsDataCorpus_de_users_clean <- tm_map(tweetsDataCorpus_de_users_clean,
content_transformer(gsub), pattern = "https:.*", replacement = "")
```

- **Διαγραφή των σημείων στίξης**

```
#tweetsDataCorpus_de_users_clean <- tm_map(tweetsDataCorpus_de_users_clean,
removePunctuation)
```

```
tweetsDataCorpus_de_users_clean <- tm_map(tweetsDataCorpus_de_clean,
content_transformer(gsub), pattern = "[!\"#$%&'()*+,-./:;<=>?[\\^_`{|}~.]", replacement = "
")
```

- **Διαγραφή stopwords**

```
desw <- readLines('de_stop_words_UC', encoding='ISO-8859-1')
```

```
deswCorpus <- VCorpus(VectorSource(desw))
```

```
deswCorpus_clean <- tm_map(deswCorpus, content_transformer(tolower))
```

```
text = sapply(deswCorpus_clean, as.character)
```

```
vtext = as.vector(text)
```

```
tweetsDataCorpus_de_users_cleanNoAcc <- tm_map(tweetsDataCorpus_de_users_clean,
removeWords,vtext)
```

- **Διαγραφή mentions**

```
tweetsDataCorpus_de_users_clean <- tm_map(tweetsDataCorpus_de_users_clean,
content_transformer(gsub), pattern = "@[a-z,A-Z,0-9;:]*", replacement = "")
```

- **Διαγραφή των «rt»**

```
tweetsDataCorpus_de_users_clean <- tm_map(tweetsDataCorpus_de_users_clean,
removeWords, c("rt"))
```

```
#tweetsDataCorpus_de_users_clean <- tm_map(tweetsDataCorpus_de_users_clean,
removePunctuation)
```

```
tweetsDataCorpus_de_users_clean <- tm_map(tweetsDataCorpus_de_clean,
content_transformer(gsub), pattern = "[!\"#$%&'()*+,-./:;<=>?[\^_`{|}~.]", replacement = "
")
```

- **Ανάλυση συναισθήματος με τη μέθοδο του λεξικού**

#Lexicon based sentiment analysis

#Users tweets sentiment

```
library(stringr)
```

```
tweets_dtm <- DocumentTermMatrix(tweetsDataCorpus_de_users_cleanNoAcc)
```

```
as.matrix(tweets_dtm[76,])
```

```
de.pos <- readLines('deeeek_positive_words', encoding='UTF-8' )
```

```
de.pos <- gsub('-', "\\\w+", de.pos)
```

```
de.pos <- gsub('\^', "\\\s+", de.pos)
```

```
de.neg <- readLines('deeeek_negative_words', encoding='UTF-8' )
```

```
de.neg <- gsub('-', "\\\w+", de.neg)
```

```
de.neg <- gsub('\^', "\\\s+", de.neg)
```

```
str.match <- function(pattern, x)
```

```
{ no_matches = 0
```

```
for (i in 1:length(pattern)) {
```

```
  matches <- str_extract_all(x, pattern[i])
```

```
  no_matches <- no_matches + length(matches[[1]])
```

```
  if (length(matches[[1]]) > 0)
```

```
    print(matches[[1]])
```

```
}
```

```
return(no_matches)
```

```
}
```

```
global_users_pos <- 0
```

```
global_users_neg <- 0
```

```
global_users_neu <- 0
```

```
for (i in 1:length(tweetsDataCorpus_de_users_cleanNoAcc)) {
```

```

no_pos_matches <- str.match(de.pos,
tweetsDataCorpus_de_users_cleanNoAcc[[i]]$content)

no_neg_matches <- str.match(de.neg, paste0(" ",
tweetsDataCorpus_de_users_cleanNoAcc[[i]]$content))

if (no_pos_matches > no_neg_matches)

{global_users_pos <- global_users_pos + 1}

else if (no_pos_matches < no_neg_matches)

{global_users_neg <- global_users_neg + 1}

else

{global_users_neu <- global_users_neu + 1}

cat(i, ":", no_pos_matches, ":", no_neg_matches, "\n")

print("-----")

}

```

#Users tweets sentiment

```

library(stringr)

tweets_dtm <- DocumentTermMatrix(tweetsDataCorpus_de_users_cleanNoAcc)

as.matrix(tweets_dtm[1,])

de.pos <- readLines('deek_positive_words', encoding='UTF-8' )

de.pos <- gsub('-', '\\\\w+', de.pos)

de.pos <- gsub("\\^", '\\\\s+', de.pos)

de.neg <- readLines('deek_negative_words', encoding='UTF-8' )

de.neg <- gsub('-', '\\\\w+', de.neg)

de.neg <- gsub("\\^", '\\\\s+', de.neg)

str.match <- function(pattern, x)

{ no_matches = 0

for (i in 1:length(pattern)) {

matches <- str_extract_all(x, pattern[i])

no_matches <- no_matches + length(matches[[1]])

```

```

if (length(matches[[1]]) > 0)
  print(matches[[1]])
}
return(no_matches)
}

global_users_pos <- 0
global_users_neg <- 0
global_users_neu <- 0

for (i in 1:length(tweetsDataCorpus_de_users_cleanNoAcc)) {
  no_pos_matches <- str.match(de.pos,
tweetsDataCorpus_de_users_cleanNoAcc[[i]]$content)

  no_neg_matches <- str.match(de.neg, paste0(" ",
tweetsDataCorpus_de_users_cleanNoAcc[[i]]$content))

  if (no_pos_matches > no_neg_matches)
    {global_users_pos <- global_users_pos + 1}
  else if (no_pos_matches < no_neg_matches)
    {global_users_neg <- global_users_neg + 1}
  else
    {global_users_neu <- global_users_neu + 1}
  cat(i, ":", no_pos_matches, ":", no_neg_matches, "\n")
  print("-----")
}

```

Naive Bayes για τα tweets των γερμανικών media

Loading automatically the required package

implementing the Naive bayes approach

```
library(e1071)
```

Reading data from .xlsx file

```
file<-tweets_de_media
```

```
file[,2]<-dataframe
```

```
file[,3]<-tweets_de_media$sentiment
```

Finding the number of tweets

```
num_tweets<-dim(file)[1]
```

```
num_tweets
```

Determining the number of words for each tweet

```
num_words<-matrix(nrow=num_tweets,ncol=1)
```

```
for(i in 1:num_tweets)
```

```
{
```

```
  num_words[i]<-length(strsplit(as.character(file[i,2])," ")[[1]])
```

```
}
```

```
num_words
```

Finding the total number of words (regarding all available tweets)

```
tot_num_words<-sum(num_words)
```

```
tot_num_words
```

Building a matrix comprising the total number of words

```
words<-matrix(nrow=tot_num_words,ncol=1)
```

```
k<-1
```

```
for(i in 1:num_tweets)
```

```
{
```



```

for(j in 1:num_words[i])
{
  words[k]<-strsplit(as.character(file[i,2])," ")[[1]][j]
  k<-k+1
}
}

# Building a matrix comprising the words without repetitions

common_words<-unique(words)

common_words

# Finding the number of common words

num_com_words<-length(common_words)

# Building the text-word matrix

tweet_words=matrix(nrow=num_tweets,ncol=num_com_words)

for(i in 1:num_tweets)
{
  for(j in 1:num_com_words)
  {
    tweet_words[i,j]<-length(which(strsplit(as.character(file[i,2]),"
    ")[1]==common_words[j]))
  }
}

# Performing the Naive bayes technique

naiveBayes.train<-
naiveBayes(tweet_words[1:(num_tweets/3),1:(num_tweets/3)],file[1:(num_tweets/3),3],lap
lace=1)

results<-predict(naiveBayes.train,tweet_words)

# Estimating the success rate

success<-matrix(nrow=num_tweets,ncol=1)

for (i in 1:num_tweets)

```

```

{
  if (results[i]==file[i,3])
  { success[i]<-1 }
  else { success[i]<-0 }
}

accuracy<-sum(success)/num_tweets

accuracy

write(dataframe3[,1],file="tweets_de_media.csv",ncolumns=1)

tweets_de_media_unclean<-matrix(nrow=dim(tweets_de_media)[1],ncol=4)

for (i in 1:dim(tweets_de_media)[1])
{
  tweets_de_media_unclean[i,1]<-tweets_de_media[i,1]
  tweets_de_media_unclean[i,2]<-tweets_de_media[i,2]
  tweets_de_media_unclean[i,3]<-tweets_de_media[i,3]
  tweets_de_media_unclean[i,4]<-tweets_de_media[i,4]
}

write.csv(tweets_de_media_unclean,file="tweets_de_media_unclean.csv")

```

Naive Bayes για τα tweets των Γερμανών χρηστών

Loading automatically the required package

implementing the Naive bayes approach

```
library(e1071)
```

```
file<-tweets_de_users
```

```
file[,2]<-dataframe
```

```
file[,3]<-tweets_de_users$sentiment
```

Finding the number of tweets

```
num_tweets<-dim(file)[1]
```

```
num_tweets
```

Determining the number of words for each tweet

```

num_words<-matrix(nrow=num_tweets,ncol=1)
for(i in 1:num_tweets)
{
  num_words[i]<-length(strsplit(as.character(file[i,2])," ")[[1]])
}
num_words

```

Finding the total number of words (regarding all available tweets)

```

tot_num_words<-sum(num_words)
tot_num_words

```

Building a matrix comprising the total number of words

```

words<-matrix(nrow=tot_num_words,ncol=1)
k<-1
for(i in 1:num_tweets)
{
  for(j in 1:num_words[i])
  {
    words[k]<-strsplit(as.character(file[i,2])," ")[[1]][j]
    k<-k+1
  }
}

```

Building a matrix comprising the words without repetitions

```

common_words<-unique(words)
common_words

```

Finding the number of common words

```

num_com_words<-length(common_words)

```

Building the text-word matrix

```

tweet_words=matrix(nrow=num_tweets,ncol=num_com_words)
for(i in 1:num_tweets)
{
  for(j in 1:num_com_words)
  {
    tweet_words[i,j]<-length(which(strsplit(as.character(file[i,2]),"
")[[1]]==common_words[j]))
  }
}

```

Performing the Naive bayes technique

```

naiveBayes.train<-
naiveBayes(tweet_words[1:(num_tweets/3),1:(num_tweets/3)],file[1:(num_tweets/3),3],lap
lace=1)

results<-predict(naiveBayes.train,tweet_words)

```

Estimating the success rate

```

success<-matrix(nrow=num_tweets,ncol=1)
for (i in 1:num_tweets)
{
  if (results[i]==file[i,3])
  { success[i]<-1 }
  else { success[i]<-0 }
}

accuracy<-sum(success)/num_tweets
accuracy

tweets_de_users_unclean<-matrix(nrow=dim(tweets_de_users)[1],ncol=4)
for (i in 1:dim(tweets_de_users)[1])
{
  tweets_de_users_unclean[i,1]<-tweets_de_users[i,1]
  tweets_de_users_unclean[i,2]<-tweets_de_users[i,2]

```

```

tweets_de_users_unclean[i,3]<-tweets_de_users[i,3]
tweets_de_users_unclean[i,4]<-tweets_de_users[i,4]
}
write.csv(tweets_de_users_unclean,file="tweets_de_users_unclean.csv")

```

Τροποποίηση αλγορίθμου Naive Bayes

- **1^η τροποποίηση: Διαδικασία εκτίμησης απόδοσης αλγόριθμου**

```
# Performing the Naive Bayes technique
```

```
test.thres<-round(0.3*num_tweets,digits=0)
```

```
naiveBayes.train<-
```

```
naiveBayes(tweet_words[test.thres:num_tweets,],file[test.thres:num_tweets,3],laplace=1)
```

```
results<-predict(naiveBayes.train,tweet_words[1:test.thres,])
```

```
# Estimating the success rate
```

```
success<-matrix(nrow=test.thres,ncol=1)
```

```
for (i in 1:test.thres)
```

```
{
```

```
if (results[i]==file[i,3])
```

```
{ success[i]<-1 }
```

```
else { success[i]<-0 }
```

```
}
```

```
accuracy<-sum(success)/test.thres
```

```
accuracy
```

- **2^η τροποποίηση: Περιορισμός Λέξεων με το ίδιο νόημα αλλά διαφορετική δομή**

```
# Building a matrix comprising the words truncated to a specified letter position
```

```
trunc_words<-matrix(nrow=tot_num_words,ncol=1)
num_letters<-3

for(i in 1:tot_num_words)
{
trunc_words[i]<-substr(words[i],1,num_letters)
}

# Building a matrix comprising the words without repetitions

common_words<-unique(trunc_words)
common_words<-common_words[common_words!=""]
```

- **Αξιολόγηση αλγόριθμου με τη βοήθεια συνθετικών δειγμάτων**

```
# Loading automatically the required packages

library(e1071) # implementing the Naive Bayes approach
library(bootstrap) # implementing the bootstrap technique

# Specifying the words of positive and negative meaning

pos.words<-
c("good","nice","gentle","happy","smiling","beautiful","kind","warm","handsome","sweet","friendly","lovely","sympathetic","welcome","helpful")

neg.words<-
c("bad","sad","mean","ugly","selfish","cruel","awful","horrible","unhappy","miserable","dreadful","deficient","evil","inhuman","brutal")

com.words<-c(pos.words,neg.words) # building the synthetic bag of words

# Finding the number of words of both positive and negative meaning

num_pos.words<-length(pos.words)
num_neg.words<-length(neg.words)
num_com.words<-length(com.words)

# Generating synthetic tweets

num_tweets<-10000
```

```

boot.words<-function(x,data)
{
boot.words<-com.words[x]
return(boot.words)
}

tweets<-bootstrap(1:num_com.words,num_tweets,boot.words,com.words)$thetastar
# Building the text-word matrix

tweet_words=matrix(nrow=num_tweets,ncol=num_com.words)

for(i in 1:num_tweets)
{
for(j in 1:num_com.words)
{
tweet_words[i,j]<-length(which(tweets[,i]==com.words[j]))
}
}
# Determining the number of positive words

tweet_pos.words=matrix(nrow=num_tweets,ncol=num_pos.words)
num_tweet_pos.words=matrix(nrow=num_tweets,ncol=1)

for(i in 1:num_tweets)
{
for(j in 1:num_pos.words)
{
tweet_pos.words[i,j]=length(which(tweets[,i]==pos.words[j]))
}
}
num_tweet_pos.words[i]=sum(tweet_pos.words[i,])
}

# Determining the number of negative words

tweet_neg.words=matrix(nrow=num_tweets,ncol=num_neg.words)
num_tweet_neg.words=matrix(nrow=num_tweets,ncol=1)

for(i in 1:num_tweets)
{
for(j in 1:num_neg.words)
{
tweet_neg.words[i,j]=length(which(tweets[,i]==neg.words[j]))
}
}
num_tweet_neg.words[i]=sum(tweet_neg.words[i,])
}
# Setting the rule for labeling a tweet

sentim<-matrix(nrow=num_tweets,ncol=1)

```

```

sent.thres<-0.6

for (i in 1:num_tweets)
{
if (num_tweet_pos.words[i]>=round((sent.thres*num_com.words),digits=0))
{ sentim[i]<-"positive" }
else if (num_tweet_neg.words[i]>=round((sent.thres*num_com.words),digits=0))
{ sentim[i]<-"negative" }
else { sentim[i]<-"neutral" }
}

sentim<-as.factor(sentim) # converting to a factor-type vector (numeric codes with character-
valued levels)
levels(sentim)<-c("positive","negative","neutral") # assigning the full set of levels

# Performing the naiveBayes technique

test.thres<-round(0.3*num_tweets,digits=0)

naiveBayes.train<-
naiveBayes(tweet_words[test.thres:num_tweets,],sentim[test.thres:num_tweets,],laplace=1)
results<-predict(naiveBayes.train,tweet_words[1:test.thres,])

# Estimating the success rate

success<-matrix(nrow=test.thres,ncol=1)

for (i in 1:test.thres)
{
if (results[i]==sentim[i])
{ success[i]<-1 }
else { success[i]<-0 }
}

accuracy<-sum(success)/test.thres

```