



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

**Συλλογή, Ανάλυση και Αξιολόγηση απόψεων Ελλήνων  
Πολιτών σε σχέση με παρεχόμενες Υπηρεσίες Ηλεκτρονικής  
Διακυβέρνησης, όπως αυτές εκφράζονται σε πηγές Web 2.0**

**ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ**

**Ευαγγελία Δ. Καλύβα**

**Επιβλέπων : Ασκούνης Δημήτριος**

Αναπληρωτής Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούλιος 2011





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

**Συλλογή, Ανάλυση και Αξιολόγηση απόψεων Ελλήνων  
Πολιτών σε σχέση με παρεχόμενες Υπηρεσίες Ηλεκτρονικής  
Διακυβέρνησης, όπως αυτές εκφράζονται σε πηγές Web 2.0**

## ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Ευαγγελία Δ. Καλύβα

**Επιβλέπων : Ασκούνης Δημήτριος**

Αναπληρωτής Καθηγητής Ε.Μ.Π

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 20ή Ιουλίου 2011.

.....  
Δημήτριος Ασκούνης  
Αναπ. Καθηγητής Ε.Μ.Π

.....  
Ιωάννης Ψαρράς  
Καθηγητής Ε.Μ.Π

.....  
Βασίλειος Ασημακόπουλος  
Καθηγητής Ε.Μ.Π

Αθήνα, Ιούλιος 2011

Ευαγγελία Δ. Καλύβα

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

Copyright © Ευαγγελία Δ. Καλύβα, 2011

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

*Αφιερώνεται στην οικογένειά μου.*

## Περίληψη

Η αλματώδης ανάπτυξη της τεχνολογίας επιφέρει ριζικές αλλαγές στην καθημερινότητα των ανθρώπων, δημιουργώντας νέες καινοτόμες υπηρεσίες και εκμεταλλεύοντας στο έπακρον όποια δυνατότητα παρουσιαστεί. Οι κυβερνήσεις σε παγκόσμια κλίμακα δεν θα μπορούσαν να απέχουν από την παραπάνω πρόκληση, εισάγοντας την έννοια της ηλεκτρονικής διακυβέρνησης στα πλαίσια της μετάβασης στην ευρύτερη ηλεκτρονική δημοκρατία. Η παροχή δημόσιων υπηρεσιών προς πολίτες και επιχειρήσεις ηλεκτρονικά αποτέλεσε αναμφίβολα καίριο βήμα προς την βελτίωση της ποιότητας αυτών, την ταχύτερη και αμεσότερη εξυπηρέτηση των ενδιαφερομένων και την τελική ανάπτυξη υγιέστερης σχέσης μεταξύ πολιτών και δημόσιου τομέα.

Ενώ αρχικά, η ηλεκτρονική διακυβέρνηση ήταν άμεσα συνυφασμένη με το τεχνολογικό υπόβαθρο των υπηρεσιών και την ανάπτυξη σύγχρονων λογισμικών και εξελιγμένων συστημάτων δικτύων, στην πορεία έγινε σαφές ότι για να καταστεί το εν λόγω εγχείρημα επιτυχές χρειάζεται κάτι πολύ περισσότερο, με πρωταγωνιστή την δημιουργία ψηφιακής και τεχνολογικής κουλτούρας στον πολίτη. Η ανάγκη να γνωρίζουμε την υποκειμενική άποψη των άλλων σχετικά με κάθε είδους ζήτημα υπήρξε πάντοτε υψίστης σημασίας για τον άνθρωπο και αναπόφευκτα αποτελεί αναπόσπαστο κομμάτι προς την επιτυχή μετάβαση στην παροχή υπηρεσιών ηλεκτρονικής διακυβέρνησης. Καθοριστική κρίνεται και η ανάπτυξη του Ιστού 2.0 που συνέβαλε στην άμεση αλληλεπίδραση των χρηστών του διαδικτύου και άλλαξε ριζικά τον τρόπο επικοινωνίας αλλά και μεταφοράς πληροφοριών σε παγκόσμια κλίμακα. Οι χρήστες πλέον έχουν την δυνατότητα να διατυπώνουν ελεύθερα κάθε είδους γνώμη και άποψη και να την μοιράζονται μέσω του Παγκόσμιου Ιστού, επηρεάζοντας κατ' αυτό τον τρόπο την υποκειμενική στάση των υπόλοιπων χρηστών.

Η υπερχειλίση πληροφοριών στο διαδίκτυο αποτελεί θησαυρό και ταυτόχρονα πρόκληση για τους ερευνητές για την όσο το δυνατόν καλύτερη αξιοποίησή τους. Φυσικά, η αυτόματη κατηγοριοποίηση και ταξινόμηση άποψης πάνω σε ένα οποιοδήποτε θέμα μόνο εύκολη δεν μπορεί να χαρακτηριστεί. Παρ' όλα αυτά έχουν αναπτυχθεί διάφοροι αλγόριθμοι κατηγοριοποίησης αλλά και έτοιμα λογισμικά που είναι σε θέση να επιτύχουν τον παραπάνω στόχο. Το πεδίο της αυτόματης εξόρυξης συναισθήματος έχει κεντρίσει το ενδιαφέρον των ερευνητών την τελευταία δεκαετία και αποτελεί ταχέως αναπτυσσόμενο κλάδο στα πλαίσια της ευρύτερης ανάπτυξης του πεδίου της Τεχνητής Νοημοσύνης.

Ο εν λόγω κλάδος έχει απασχολήσει τις εταιρίες που επιθυμούν να γνωρίζουν την άποψη των καταναλωτών τους σχετικά με τις παρεχόμενες υπηρεσίες και τα προϊόντα τους, ώστε να αποκτήσουν πλεονέκτημα έναντι του ανταγωνισμού, αλλά φυσικά και τις κυβερνήσεις για τις οποίες η γνώση της άποψης των πολιτών

σχετικά με τις κινήσεις τους κρίνεται μεγίστης σημασίας. Προς αυτή την κατεύθυνση, στόχος των κυβερνήσεων αποτελεί πλέον η εξόρυξη γνώμης των πολιτών σχετικά με τις παρεχόμενες υπηρεσίες ηλεκτρονικής διακυβέρνησης με στόχο την βελτιστοποίηση αυτών, στα πλαίσια πάντα της γενικότερης ανάγκης για συνεργασία και διαδραστικότητα μεταξύ των δύο πλευρών.

**Λέξεις Κλειδιά:** Ηλεκτρονική Διακυβέρνηση, Ιστός 2.0, Κοινωνικά Μέσα, Αυτόματη Εξόρυξη Γνώμης, Αυτόματη Εξόρυξη Συναισθήματος.

## Abstract

The exponential technological development changes radically human daily routine by creating new innovative services and exploiting every potential occurrence. The governments worldwide could not abstain from the above mentioned challenge, so they have imported the electronic government (e-government) in the context of transition to electronic democracy. The electronic delivery of public services to citizens and businesses constituted undoubtedly a vital action for the improvement of their quality, the creation of more rapid and more direct services and the final development of a healthier relationship between citizens and public sector.

Although, in the beginning, the electronic government was directly interwoven with the technological infrastructure and the development of contemporary software and advanced computer networks, later it was clear that in order to be this project successful there needed to be something more, mainly the introduction of digital and technological culture to the citizens. The knowledge of the subjective opinion of the others was always of paramount importance to the people and inevitably constitutes an integral factor to the success of electronic public services' delivery. In addition, the growth of Web 2.0 is of utmost importance as it contributed to the direct interaction between the web users and changed radically the usual way of communication as also as information delivery all around the world. Web users are now able to share their thought and opinion without limits and make the aforementioned available through the World Wide Web, affecting by this way the subjective thought of other users.

The information overflow through the World Wide Web consists a treasure and simultaneously a challenge for the researchers in order to make the most out of it. Without doubt, automatic classification and categorization of opinion, no matter the subject, is far from being described as an easy task. However, many classification algorithms as well as a lot of programs have been created which can achieve this goal. The field of automatic opinion extraction has gained the researchers' interest through the last decade and has been developed rapidly in the terms of the general development of the wider field of Artificial Intelligence.

The aforementioned field of opinion extraction has been significant not only for the businesses, as they need to get knowledge of how their customers feel about their services or products in order to take advantage in the competition between other companies, but also for the governments as the knowledge of their citizens' opinion about their campaigns and strategies is of paramount importance. To this direction, governments' goal is now to extract the opinion of their citizens about the electronic delivery of public services in order to improve the quality of them because of the general necessity of collaboration and interactivity between the two sides.



**Keywords:** Electronic Government, e- government, Web 2.0, Social Media, Opinion Extraction, Opinion Mining, Sentiment Analysis.

## Πρόλογος

Η παρούσα διπλωματική εργασία εκπονήθηκε στον τομέα Ηλεκτρικών Βιομηχανικών Διατάξεων και Συστημάτων Απόφασης της Σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του ΕΜΠ, στα πλαίσια των ερευνητικών δραστηριοτήτων του Εργαστηρίου Συστημάτων Αποφάσεων και Διοίκησης. Σημαντικό ρόλο για την επιλογή του θέματος αυτού αποτέλεσε η συμμετοχή του εργαστηρίου στο ερευνητικό έργο Cockpit (Citizens Collaboration & Co-Creation In Public Service Delivery) το οποίο εστιάζει στην δημιουργία ενός νέου μοντέλου ηλεκτρονικής διακυβέρνησης με βασική προϋπόθεση τον ενεργό ρόλο και συμμετοχή των πολιτών σε όλα τα στάδια οργάνωσης της νεοσύστατης ηλεκτρονικής υπηρεσίας.

Αντικείμενο της διπλωματικής εργασίας αποτελεί η δημιουργία μιας πρότυπης εφαρμογής αυτόματης εξόρυξης συναισθήματος (sentiment analysis), όπως αυτό αποτυπώνεται σε διαδικτυακές πηγές, με έμφαση στα κοινωνικά μέσα. Σκοπός της εφαρμογής αυτής είναι η αυτοματοποιημένη ταξινόμηση της άποψης των Ελλήνων πολιτών σχετικά με την ηλεκτρονική διακυβέρνηση, όπως αυτή εκφράζεται σε blogs, forums και λουπά κοινωνικά δίκτυα (social networks).

Υπεύθυνος κατά την εκπόνηση της διπλωματικής εργασίας ήταν ο Αναπληρωτής Καθηγητής κ. Δ. Ασκούνης, στον οποίο οφείλω τις θερμές μου ευχαριστίες για την ανάθεση αυτής και την δυνατότητα ενασχόλησης με ένα τόσο σημαντικό και ταυτόχρονα ενδιαφέρον θέμα.

Τέλος, θα ήθελα να ευχαριστήσω ιδιαίτερα τους συνεργάτες του Εργαστηρίου Συστημάτων Αποφάσεων και Διοίκησης κ.κ. Δημήτρη Πανόπουλο και Παναγιώτη Κοκκινάκο, τόσο για την καθοδήγηση όσο και για την υποστήριξη που μου παρείχαν καθ' όλη την διάρκεια εκπόνησης της διπλωματικής μου εργασίας, στα πλαίσια της γενικότερης άριστης συνεργασίας που αναπτύχθηκε.

Ευαγγελία Καλύβα  
Ιούλιος 2011

## Πίνακας Περιεχομένων

ΚΕΦΑΛΑΙΟ 1: Εισαγωγή.....	13
1.1. Αντικείμενο – Σκοπός .....	<b>Error! Bookmark not defined.</b> 4
1.2. Φάσεις υλοποίησης .....	15
1.3. Οργάνωση τόμου.....	16
ΚΕΦΑΛΑΙΟ 2: Εξόρυξη γνώμης/ συναισθήματος....	<b>Error! Bookmark not defined.</b> 8
2.1. Data Mining.....	19
2.2. Text Mining.....	<b>Error! Bookmark not defined.</b> 19
2.3. Opinion Mining/ Sentiment Analysis.....	<b>Error! Bookmark not defined.</b> 20
2.4. Επιμέρους στάδια Εξόρυξης Γνώμης/ Συναισθήματος .....	21
2.5. Πλεονεκτήματα .....	23
2.5. Προκλήσεις/ Δυσκολίες .....	25
ΚΕΦΑΛΑΙΟ 3: Ηλεκτρονική Διακυβέρνηση και Opinion Mining.....	28
3.1. Πολιτική 2.0 .....	29
3.2. Ηλεκτρονική Διακυβέρνηση και Εξόρυξη Δεδομένων.....	30
3.2.1. Στάδια Ηλεκτρονικής Διακυβέρνησης.....	30
3.2.2. Συνεισφορά Εξόρυξης Γνώμης .....	32
3.3. Ιστότοποι Ηλεκτρονικής Διακυβέρνησης .....	<b>Error! Bookmark not defined.</b>
3.4. Παραδείγματα .....	<b>Error! Bookmark not defined.</b> 3
ΚΕΦΑΛΑΙΟ 4: Εφαρμογή Opinion Mining .....	41
4.1. Εισαγωγή στο RapidMiner.....	42
4.1.1. Σχεδίαση 1 <sup>ης</sup> διεργασίας.....	44
4.1.2. Σχεδίαση 2 <sup>ης</sup> διεργασίας.....	51
4.2. Εκτέλεση .....	55
4.3. Κώδικας.....	55
ΚΕΦΑΛΑΙΟ 5: Συγκριτικά στοιχεία με χρήση αγγλικής γλώσσας .....	59
5.1. Ηλεκτρονική Διακυβέρνηση και Web 2.0 .....	60
5.2. Στατιστικά στοιχεία για ΕΕ και Ελλάδα .....	61
5.2.1. Μέση διαθεσιμότητα διαδικτυακών υπηρεσιών .....	61
5.2.2. Ποσοστό χρησιμοποίησης και εμπειρία χρήστη .....	63

5.3. Άποψη Ελλήνων πολιτών για ηλεκτρονική διακυβέρνηση .....	64
5.3.1. Σημαντικότερες ηλεκτρονικές υπηρεσίες στην Ελλάδα .....	65
5.3.2. Σημαντικότεροι ιστότοποι άντλησης γνώμης .....	66
5.3.3. Κύριες καταγεγραμμένες απόψεις.....	66
5.4. Σύγκριση αποτελεσμάτων με χρήση εφαρμογής στην αγγλική γλώσσα .....	68
5.4.1. Τροποποίηση εφαρμογής RapidMiner .....	69
5.4.2. Σύγκριση ακρίβειας αποτελεσμάτων .....	71
5.5. Παράθεση εφαρμογών αυτόματης εξόρυξης γνώμης	<b>Error! Bookmark not defined.</b> 74
ΚΕΦΑΛΑΙΟ 6: Συμπερασματα και Προοπτικές .....	82
6.1. Συμπεράσματα.....	83
6.2. Προοπτικές .....	84
Βιβλιογραφία: .....	86

## ***Κεφάλαιο 1. Εισαγωγή***

---

## 1.1 Αντικείμενο – Σκοπός

Η τεχνολογία αναφέρεται στο αποτέλεσμα της εφαρμογής της (θεωρητικής) επιστημονικής γνώσης με στόχο την δημιουργία ενός αντικειμένου με πρακτικό όφελος. Τα τελευταία χρόνια παρατηρείται η τάση η έννοια να αναφέρεται μόνο στην υψηλή τεχνολογία ή/και στην τεχνολογία υπολογιστών μόνο, αν και κατά βάση δεν περιορίζεται μόνο σε αυτούς τους τομείς. Το σίγουρο είναι ότι ο κοινός ανθρώπινος νους, χρόνια πριν, δεν μπορούσε να συλλάβει και πόσο μάλλον να πιστέψει την δημιουργία τεχνολογικών επιτευγμάτων που πλέον έχουν εισβάλει στην καθημερινότητά μας, προσδίδοντας ευκολία, άνεση, ευχαρίστηση και ικανοποίηση.

Ποιος θα μπορούσε δεκαετίες πριν να φανταστεί ότι με ένα «κλικ» θα μπορούσε να εξερευνήσει έναν ατελείωτο πλούτο πληροφορίας, να επικοινωνήσει με άλλα άτομα καταργώντας κάθε έννοια γεωγραφικών και πολιτιστικών συνόρων και, για να προσεγγίσουμε περισσότερο το αντικείμενο της παρούσας εργασίας, να αποκτήσει πρόσβαση σε δημόσιες υπηρεσίες όχι μόνο από άποψη πλήρους ενημέρωσης σχετικά με την διεκπεραίωση των εκκρεμοτήτων του απέναντι σε κρατικούς φορείς αλλά πολύ περισσότερο από άποψη άμεσης εκτέλεσης αυτών ανεξαρτήτου ωραρίου. Αυτή είναι η εποχή της ηλεκτρονικής διακυβέρνησης, που με τη χρήση πληροφοριακών συστημάτων στη δημόσια διοίκηση αποσκοπεί στη ριζική αλλαγή της λειτουργίας όλων των σύγχρονων κρατών.

Θα μπορούσε να θεωρηθεί μάλλον άσκοπο κι ίσως ανούσιο να εμμείνει κάποιος αποκλειστικά στην τεχνολογική υποδομή των υπηρεσιών της ηλεκτρονικής διακυβέρνησης. Είναι πλέον περισσότερο σαφές από ποτέ ότι το νέο αυτό πολύτιμο εργαλείο δεν είναι κάτι άψυχο αλλά βασίζεται πάνω από όλα στις επιθυμίες και τις ανάγκες πολιτών και επιχειρήσεων που κάνουν χρήση των εν λόγω υπηρεσιών. Ακόμα και η πιο τέλεια τεχνολογικά πλατφόρμα ηλεκτρονικής διακυβέρνησης θα ήταν καταδικασμένη να αποτύχει εάν οι χρήστες δεν αποκτήσουν εμπιστοσύνη στο νέο αυτό μοντέλο και δεν συνειδητοποιήσουν τα οφέλη που θα μπορέσουν να αποκομίσουν μόνο εντός των πλαισίων συνεργασίας και διαδραστικότητας με το κράτος.

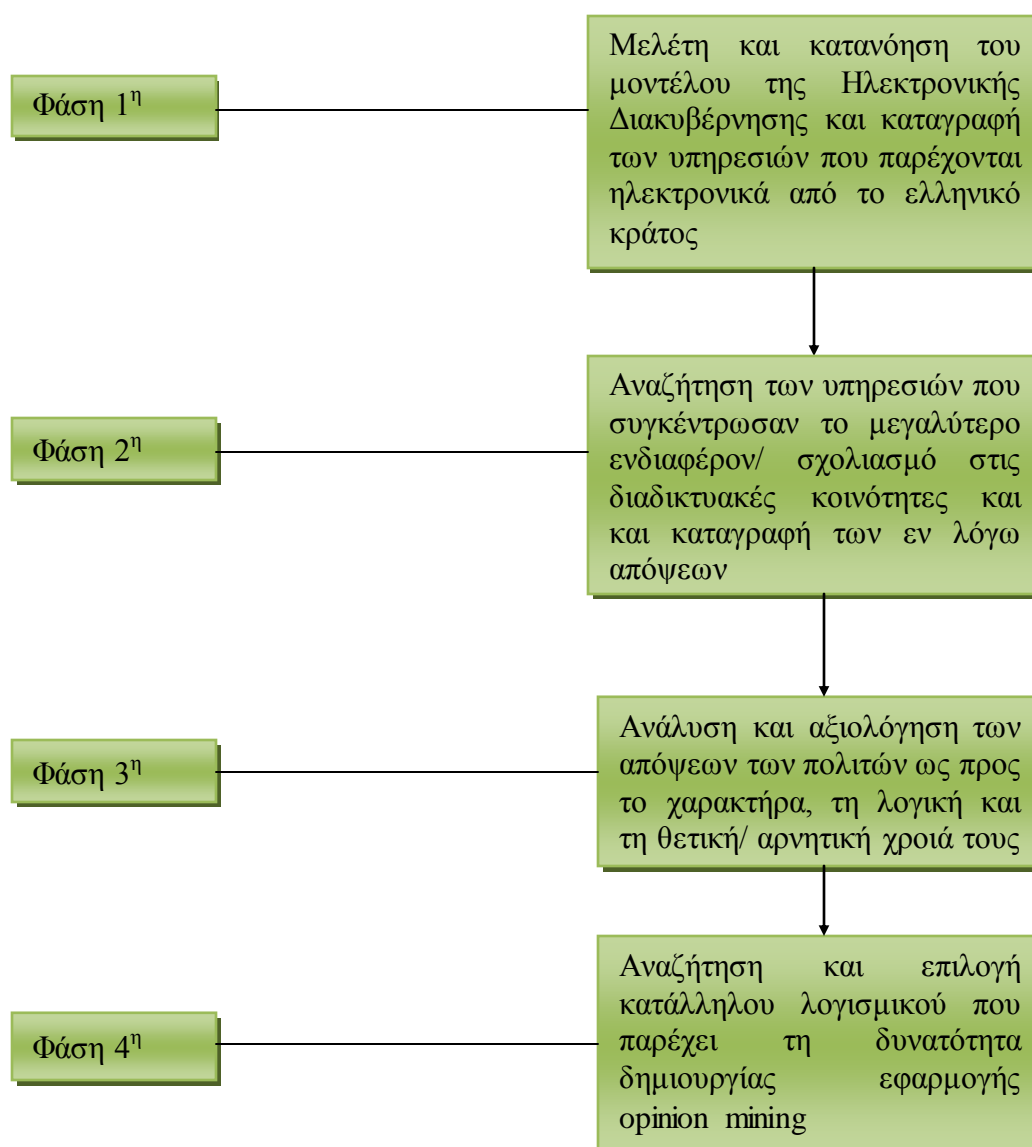
Είναι υψίστης σημασίας για τις κυβερνήσεις να «ακούσουν» την γνώμη των πολιτών τους και να αφουγκραστούν την γενικότερη στάση τους απέναντι στο εν λόγω εργαλείο, κάτι που μπορεί να γίνει άμεσα λόγω της πληθώρας διατυπωθέντων απόψεων στο διαδίκτυο με έμφαση στα κοινωνικά μέσα. Η πρόκληση και ο προβληματισμός πηγάζει από το μέγεθος της διαθέσιμης αυτής πληροφορίας που αναπόφευκτα οδηγεί στην ανάγκη για δημιουργία σύγχρονων λογισμικών που θα έχουν την δυνατότητα να ταξινομούν το εν λόγω υλικό αυτόματα, καθώς ο χρόνος

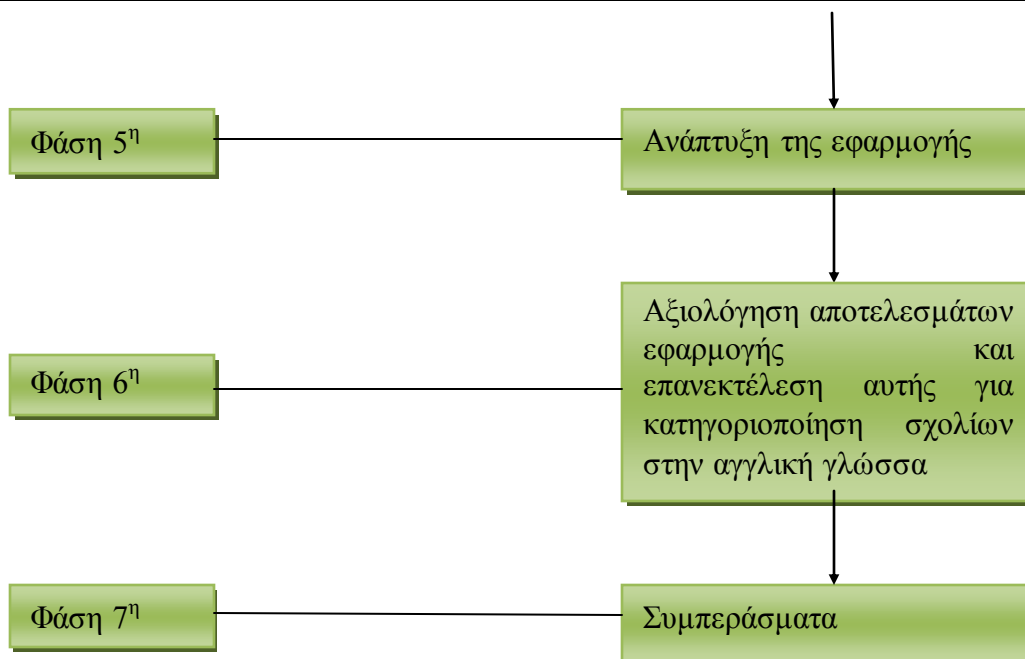
επεξεργασίας κρίνεται πολύτιμος για να σπαταληθεί σε μία εποχή που χαρακτηρίζεται από ανάγκες που αλλάζουν καθημερινά και ανεξέλεγκτα.

Μέσα σε αυτή την φιλοσοφία βρίσκει πρόσφορο έδαφος ο σκοπός της παρούσας διπλωματικής εργασίας καθώς αποσκοπήθηκε η συγκέντρωση απόψεων των Ελλήνων πολιτών σχετικά με το εγχείρημα της ηλεκτρονικής διακυβέρνησης, εκμεταλλεύοντας ουσιαστικά τα πλεονεκτήματα που παρέχει ο Ιστός 2.0, και μετέπειτα η δημιουργία μιας εφαρμογής που θα είναι σε θέση θεωρητικά να εντοπίζει το συναίσθημα και την υποκειμενική στάση του γράφοντος.

## 1.2 Φάσεις υλοποίησης

Η εκπόνηση της διπλωματικής εργασίας πραγματοποιήθηκε μεταξύ Νοεμβρίου του 2010 και Ιουλίου του 2011 και η πορεία αυτής ακολούθησε τις εξής φάσεις, που παρουσιάζονται παρακάτω στο σχήμα 1.1.





Σχήμα 1.1: Φάσεις υλοποίησης διπλωματικής εργασίας

### 1.3 Οργάνωση τόμου

Η διπλωματική εργασία «Συλλογή, Ανάλυση και Αξιολόγηση απόψεων Ελλήνων Πολιτών σε σχέση με παρεχόμενες Υπηρεσίες Ηλεκτρονικής Διακυβέρνησης, όπως αυτές εκφράζονται σε πηγές Web 2.0» περιλαμβάνει τα ακόλουθα πέντε κεφάλαια:

- Εξόρυξη γνώμης/ συναισθήματος (Opinion Mining/ Sentiment Analysis)
- Ηλεκτρονική διακυβέρνηση και Opinion Mining
- Εφαρμογή Opinion Mining
- Συγκριτικά στοιχεία με χρήση αγγλικής γλώσσας
- Συμπεράσματα και Προοπτικές

Πιο αναλυτικά η δομή της παρούσας διπλωματικής έχει ως εξής:

Στο κεφάλαιο 2 «Εξόρυξη γνώμης/ συναισθήματος (Opinion Mining/ Sentiment Analysis)», ξεκινώντας από το πεδίο της εξόρυξης γνώσης από δεδομένα (data mining), προχωρούμε στην ανάλυση της τεχνικής εξόρυξης κειμένου (text mining). Έναν πιο εξειδικευμένο κλάδο αποτελεί η τεχνική της εξόρυξης γνώμης (opinion mining) ή εξόρυξης συναισθήματος (sentiment analysis), ο οποίος και παρουσιάζεται αναλυτικά. Ακολουθεί η παράθεση των στόχων της εν λόγω τεχνικής, των εν δυνάμει πλεονεκτημάτων αυτής μέσω χαρακτηριστικών παραδειγμάτων και τέλος των προκλήσεων και δυσκολιών που αντιμετωπίζει ο παραπάνω κλάδος.



Στο κεφάλαιο 3 «Ηλεκτρονική διακυβέρνηση και Opinion Mining», αναλύεται αρχικά η συνεισφορά της ανάπτυξης του Ιστού 2.0 στην βελτιστοποίηση λειτουργίας επιχειρήσεων και κυβερνήσεων και η σημασία της αυτόματης εξόρυξης γνώμης μέσω διαδικτυακών πηγών για την επιτυχία της ηλεκτρονικής διακυβέρνησης. Ακολουθεί η ανάλυση των επιμέρους σταδίων της ηλεκτρονικής διακυβέρνησης, όπως αυτά έχουν διατυπωθεί από ερευνητές, με έμφαση στο στάδιο της ηλεκτρονικής συμμετοχής και τέλος παρατίθενται κάποια χαρακτηριστικά παραδείγματα κυβερνητικών πλατφόρμων που αποσκοπούν στην γνώση της άποψης των πολιτών σχετικά με το προαναφερθέν εγχείρημα.

Το κεφάλαιο 4 «Εφαρμογή Opinion Mining», αποτελεί το ερευνητικό μέρος της παρούσας εργασίας καθώς παρουσιάζεται αναλυτικά η ανάπτυξη πρότυπης εφαρμογής αυτόματης εξόρυξης συναισθήματος. Συγκεκριμένα, για το σκοπό αυτό δημιουργήθηκαν δύο διεργασίες μέσω του ελεύθερου λογισμικού προγράμματος RapidMiner και η διαδικασία που ακολουθήθηκε καθώς και τα επιμέρους βήματα περιγράφονται αναλυτικά.

Στο κεφάλαιο 5 «Συγκριτικά στοιχεία με χρήση αγγλικής γλώσσας», αρχικά παραθέεται ο ορισμός της ηλεκτρονικής διακυβέρνησης και τα αντίστοιχα στατιστικά στοιχεία για τη διείσδυση αυτής στις χώρες της ΕΕ. Έμφαση δίνεται στην ανάλυση της κατάστασης που επικρατεί στην χώρα μας, σύμφωνα πάντα με τους αντίστοιχους επίσημους δείκτες. Ακολουθεί η παράθεση των σημαντικότερων διαδικτυακών πηγών από όπου συγκεντρώθηκαν τα σχόλια των Ελλήνων πολιτών σχετικά με την ηλεκτρονική διακυβέρνηση και μία ανάλυση του κύριου περιεχομένου αυτών. Κλείνοντας το εν λόγω κεφάλαιο, τροποποιούμε την εφαρμογή που δημιουργήσαμε στο κεφάλαιο 4 και παραθέτουμε συγκριτικά στοιχεία για την ακρίβεια κατηγοριοποίησης των δεδομένων για την ελληνική και αγγλική γλώσσα αντίστοιχα. Επισυνάπτονται επίσης κάποιες εφαρμογές αυτόματης εξόρυξης συναισθήματος και αναλυτικός κατάλογος αυτών με τα κύρια χαρακτηριστικά τους.

Στο κεφάλαιο 6 «Συμπεράσματα και Προοπτικές», παρουσιάζονται τα συμπεράσματα από την προσέγγιση του εν λόγω ζητήματος κατά την εκπόνηση της παρούσας εργασίας αλλά και οι προτάσεις/ προοπτικές που διαφαίνονται πάνω στο συγκεκριμένο εγχείρημα.

## ***Κεφάλαιο 2. Εξόρυξη γνώμης / συναισθήματος***

---

## 2.1 Data Mining

*«Πνιγόμαστε στις πληροφορίες, αλλά διψάμε για γνώση»*

John Naisbitt

Η παραπάνω γνωστή ρήση είναι πιο επίκαιρη από ποτέ καθώς ζούμε στην «Εποχή της Πληροφορίας». Η συμβολή της ανάπτυξης του διαδικτύου κρίνεται αναμφίβολα καθοριστική καθώς άλλαξε όχι μόνο τον τρόπο αλλά και την νοοτροπία δισεκατομμυρίων χρηστών ανά τον κόσμο ως προς την πρόσβαση σε πηγές πληροφοριών όσον αφορά την ενημέρωσή τους. Ακόμα περισσότερο προς αυτή την κατεύθυνση συνέβαλε ο Ιστός 2.0, που πρόσθεσε τη δυνατότητα άμεσης αλληλεπίδρασης και ανταλλαγής πληροφοριών, καταργώντας κάθε έννοια γεωγραφικών αλλά και πολιτιστικών συνόρων.

Η αλματώδης αύξηση εγγράφων κειμένου που δημοσιεύονται καθημερινά στο διαδίκτυο και η αναζήτηση της επιθυμητής πληροφορίας μέσω αυτών αποτέλεσε ιδανικό κίνητρο προς την έρευνα και ανάπτυξη τεχνικών εξόρυξης γνώσης από δεδομένα. Πιο συγκεκριμένα, η εξόρυξη γνώσης από δεδομένα (data mining) αναφέρεται στην εφαρμογή τεχνικών και αλγορίθμων μηχανικής μάθησης (machine learning και machine learning algorithms). Οι τεχνικές και οι αλγόριθμοι αυτοί αποτελούν έναν από τους επιστημονικούς κλάδους του ευρύτερου πεδίου της Τεχνητής Νοημοσύνης και σημείο συνάντησης των ερευνητικών περιοχών όπως οι βάσεις δεδομένων, η επεξεργασία φυσικής γλώσσας, η ανάκτηση πληροφορίας, η εξαγωγή πληροφορίας και η μοντελοποίηση χρηστών.

## 2.2 Text Mining

Η εξόρυξη κειμένου (text mining) βασίζεται στις παραπάνω τεχνικές και προσπαθεί να επιλύσει το πρόβλημα της υπερφόρτωσης πληροφοριών, βοηθώντας τους χρήστες να εξάγουν αυτόματα από μεγάλους κειμενικούς πόρους την επιθυμητή γι' αυτούς πληροφορία κυρίως μέσω της κατηγοριοποίησης και ομαδοποίησης εγγράφων. Εδώ πρέπει να τονίσουμε ότι η εξόρυξη κειμένου δεν πρέπει να συγχέεται με την απλή αναζήτηση πληροφορίας μέσω διαδικτύου (web search). Όσον αφορά την απλή αναζήτηση πληροφορίας, με την οποία είμαστε όλοι λιγότερο ή περισσότερο εξοικειωμένοι, ο χρήστης αναζητά κάτι που είναι ήδη γνωστό και έχει γραφτεί προηγουμένως από κάποιον άλλο. Εδώ το μοναδικό πρόβλημα του χρήστη είναι να ξεχωρίσει το υλικό που είναι σχετικό και προσαρμοσμένο στις εκάστοτε ανάγκες του. Αντίθετα, στην εξόρυξη κειμένου στόχος είναι η ανακάλυψη καινούριας πληροφορίας, άγνωστης μέχρι στιγμής και συνεπώς μη ακόμα καταγεγραμμένης από κάποιον άλλο χρήστη στο παρελθόν.

Η εξόρυξη κειμένου αποτελεί ένα κλάδο του ευρύτερου πεδίου εξόρυξης γνώσης από δεδομένα και αποτελούν έννοιες που συχνά συγχέονται. Η διαφορά τους εδώ έγκειται στο γεγονός ότι στην εξόρυξη κειμένου τα υποδείγματα (γνώση) εξάγονται από κείμενο φυσικής γλώσσας και όχι από δομημένες βάσεις των γεγονότων. Οι

βάσεις δεδομένων άλλωστε σχεδιάζονται ώστε να μπορούν να προσπελαστούν αυτόματα από διάφορα προγράμματα σε αντίθεση με το κείμενο γραμμένο σε φυσική γλώσσα που διαβάζεται από ανθρώπους. Πολλοί ερευνητές πιστεύουν ότι χρειάζεται πλήρης προσομοίωση του τρόπου με τον οποίο το ανθρώπινο μυαλό λειτουργεί ώστε να υπάρξουν προγράμματα που θα μπορούν να «διαβάζουν» όπως οι άνθρωποι. Ωστόσο, η υπολογιστική γλωσσολογία (γνωστή και ως Επεξεργασία Φυσικής Γλώσσας) έχει σημειώσει αξιόλογη πρόοδο στην διεκπεραίωση μικρών δευτερευουσών εργασιών σχετικά με την ανάλυση κειμένου, όπως για παράδειγμα η αυτόματη σύνοψη περιεχομένου και εξαγωγή περίληψης.

## 2.3 Opinion Mining / Sentiment Analysis

Έναν πιο εξειδικευμένο κλάδο της επεξεργασίας φυσικής γλώσσας και της εξόρυξης κειμένου αποτελεί η εξόρυξη γνώμης/ συναισθήματος (opinion mining ή sentiment analysis). Η εξόρυξη γνώμης/ συναισθήματος στοχεύει στο να προσδιορίσει την υποκειμενική στάση του ομιλούντος ή του γράφοντα σχετικά με ένα ζήτημα ή την συνολική άποψη που επικρατεί σε ένα έγγραφο. Η άποψη ή η στάση αυτή μπορεί να συνίσταται στην κρίση ή αξιολόγηση, την συναισθηματική κατάσταση του ομιλούντα ή γράφοντα ή την προσδοκώμενη από τον τελευταίο συναισθηματική επικοινωνία και αντίκτυπο που επιθυμεί για όσα λέει ή γράφει. Άλλωστε ανέκαθεν, σημαντικό κομμάτι της συλλογής πληροφοριών υπήρξε η ανάγκη για γνώση σχετικά με το τι σκέφτονται οι υπόλοιποι για ένα συγκεκριμένο θέμα. Με την ραγδαία εξάπλωση πηγών άντλησης γνώμης όπως τα προσωπικά ιστολόγια (blogs), οι ηλεκτρονικοί τόποι δημόσιας συζήτησης (forum) και τα διάφορα sites που περιέχουν κριτικές, εμφανίστηκαν αναμφίβολα νέες ευκαιρίες και προκλήσεις προς την κατεύθυνση εξόρυξης συναισθήματος και γενικότερης κατανόησης απόψεων. Πλέον, η αυτόματη εξόρυξη γνώμης, κυρίως μέσω κειμένου, έχει αναδειχτεί σε ένα από τα αντικείμενα που προσελκύουν ολοένα και αυξανόμενο ενδιαφέρον από την ακαδημαϊκή κοινότητα αλλά και τους εταιρικούς οργανισμούς.

Ο όρος opinion mining εμφανίζεται για πρώτη φορά σε δημοσίευση των Kushal Dave, Steve Lawrence, David M. Pennock στα πλαίσια του συνεδρίου WWW (WWW conference) κατά το έτος 2003. Η δημοσίευση στο συγκεκριμένο συνέδριο μπορεί εν μέρει να εξηγήσει την δημοτικότητα του όρου opinion mining μεταξύ των κοινοτήτων που είναι προσανατολισμένες στην κατεύθυνση αναζήτησης στο διαδίκτυο (Web search) ή ανάκτησης πληροφορίας (information retrieval). Σύμφωνα με την εν λόγω δημοσίευση, η ιδανική εφαρμογή για εξόρυξη συναισθήματος «θα μπορούσε να επεξεργαστεί ένα σύνολο από δεδομένα αναζήτησης, δημιουργώντας μία λίστα των κύριων χαρακτηριστικών αυτών και συνοψίζοντας τις απόψεις που επικρατούν για κάθε ένα από αυτά τα χαρακτηριστικά σε θετικές, ουδέτερες και αρνητικές». Οι περισσότερες έρευνες που έχουν διεξαχθεί στο πεδίο του opinion mining έρχονται σε συμφωνία με τον παραπάνω ορισμό, με έμφαση κυρίως στο κομμάτι της εξόρυξης και ανάλυσης κρίσεων πάνω σε διάφορες πλευρές των εκάστοτε δοθέντων αντικειμένων. Ωστόσο, ο εν λόγω όρος πρόσφατα έχει διερμηνευτεί καλύπτοντας μεγαλύτερο εύρος και περιέχοντας πολύ περισσότερους και ταυτόχρονα διαφορετικούς τύπους ανάλυσης των κειμένων προς αξιολόγηση.

Η ιστορία του όρου sentiment analysis συμπίπτει με αυτή του opinion mining από αρκετές πλευρές. Ο όρος «sentiment» που χρησιμοποιείται αναφορικά με την αυτόματη ανάλυση του κειμένου προς αξιολόγηση και την πρόβλεψη άποψης μέσω αυτού εμφανίζεται για πρώτη φορά εν έτη 2001 στις δημοσιεύσεις των Sanjiv Das, Mike Chen και Richard M. Tong αντίστοιχα, λόγω του ενδιαφέροντος των τελευταίων για την ανάλυση αίσθησης στην αγορά. Τον επόμενο χρόνο ακολουθούν οι δημοσιεύσεις των Peter Turney και Bo Pang, Lillian Lee, Shivakumar Vaithyanathan στην ετήσια συνάντηση της Εταιρίας Υπολογιστικής Γλωσσολογίας (Association for Computational Linguistics (ACL)) καθώς και στο ετήσιο συνέδριο Εμπειρικών Μεθόδων για την Επεξεργασία Φυσικής Γλώσσας (Empirical Methods in Natural Language Processing (EMNLP)). Τα παραπάνω γεγονότα δύναται να εξηγήσουν την δημοφιλία του όρου sentiment analysis μεταξύ των κοινοτήτων με έμφαση στην Επεξεργασία Φυσικής Γλώσσας. Ένας αξιόλογος αριθμός δημοσιεύσεων που αναφέρεται στον εν λόγω όρο εμβαθύνει κυρίως σε εφαρμογές κατηγοριοποίησης κριτικών σχετικά με την πόλωση συναισθήματος που περιέχεται σε αυτές, γεγονός που φαίνεται να έχει οδηγήσει κάποιους ερευνητές στο να συνδέσουν τον όρο sentiment analysis συγκεκριμένα με αυτή την μικρή σε εύρος λειτουργία. Παρ' όλα αυτά, σήμερα, πολλοί χρησιμοποιούν τον όρο με την ευρύτερη έννοια για να αναφερθούν στην υπολογιστική αντιμετώπιση της γνώμης, του συναισθήματος και της υποκειμενικότητας σε ένα κείμενο. Με αυτό τον τρόπο, οι όροι opinion mining και sentiment analysis χρησιμοποιούνται πλέον ταυτόσημα και παρουσιάζονται στο ίδιο πεδίο μελέτης, το οποίο με την σειρά του μπορεί να θεωρηθεί μια υποπεριοχή της υποκειμενικής ανάλυσης. Σημειώνεται εδώ ότι στην παρούσα διπλωματική εργασία δεν υπήρξε διαχωρισμός των παραπάνω εννοιών και στο εξής οι όροι opinion mining και sentiment analysis θα αναφέρονται εξίσου στην εξόρυξη συναισθήματος.

Ανακεφαλαιώνοντας έως εδώ την ιστορική αναδρομή των παραπάνω όρων, αναφέρουμε ότι το 2004 η Αμερικανική Εταιρεία Τεχνητής Νοημοσύνης (American Association for Artificial Intelligence) οργάνωσε συμπόσιο επί της συγκεκριμένης θεματικής περιοχής που έφερε τον τίτλο «Εξερευνώντας Επιρροή και Στάση σε Κείμενο» («Exploring Affect and Attitude in Text»). Έκτοτε, η ανάλυση γνώμης έχει αποτελέσει αντικείμενο συνεχούς ενδιαφέροντος σε πολλά συνέδρια στις Ηνωμένες Πολιτείες, σε αντίθεση με την Ευρώπη όπου το σχετικό ενδιαφέρον σημείωσε μια συγκριτική υστέρηση. Τέλος αξιοσημείωτο είναι και το γεγονός ότι λίγες δημοσιεύσεις απαντώνται σχετικά με την αυτόματη εξόρυξη συναισθήματος σε γλώσσες εκτός της αγγλικής.

## **2.4 Επιμέρους στάδια εξόρυξης γνώμης / συναισθήματος**

Στο σημείο αυτό θα αναλυθούν οι τρεις κύριοι στόχοι της τεχνικής εξόρυξης γνώμης/ συναισθήματος. Εξ' αιτίας της πολυπλοκότητας της εν λόγω τεχνικής, η εξόρυξη συναισθήματος περιλαμβάνει αρκετές επιμέρους ξεχωριστές εργασίες, οι οποίες συνήθως συνδυάζονται για την τελική εξαγωγή γνώσης μέσω των απόψεων που εκφράζονται σε ένα κείμενο.

Η πρώτη, λοιπόν, εργασία είναι ο εντοπισμός γνώμης ή συναισθήματος, την οποία μπορούμε να περιγράψουμε ως μία ευρύτερη κατηγοριοποίηση του κειμένου σε αντικειμενικό ή υποκειμενικό. Η κατηγοριοποίηση αυτή συνήθως είναι βασισμένη στην εξέταση των υπαρχόντων επιθέτων στις προτάσεις. Χάριν παραδείγματος, η πόλωση της πρότασης: «η εικόνα είναι όμορφη» μπορεί εύκολα να προσδιοριστεί, εξετάζοντας το επίθετο που περιέχεται σε αυτή. Παλαιότερη μελέτη<sup>1</sup> περιγράφει αναλυτικά την επίδραση των επιθέτων στην υποκειμενικότητα μίας πρότασης, ενώ πιο πρόσφατες μελέτες<sup>2</sup> έχουν δείξει ότι τα επιρρήματα αποτελούν εξίσου σημαντικούς παράγοντες. Για τον αναγνώστη που επιθυμεί περισσότερες πληροφορίες σχετικά με τις τεχνικές αναγνώρισης υποκειμενικότητας μπορεί να ανατρέξει επίσης στην παρακάτω έρευνα<sup>3</sup>.

Η δεύτερη εργασία έχει να κάνει με την καθ' αυτή κατηγοριοποίηση της πόλωσης. Δοσμένου ενός μέρος του κειμένου, ο στόχος είναι να κατηγοριοποιηθεί η άποψη που εκφράζεται μεταξύ των δύο αντίθετων κατηγοριών συναισθήματος ή σε περίπτωση σύγχυσης στην πλησιέστερη εκ των δύο αυτών. Βλέποντάς το ως δυαδικό χαρακτηριστικό, η κατηγοριοποίηση πόλωσης είναι η δυαδική κατηγοριοποίηση συναισθήματος εντός του κειμένου είτε σε συνολικά θετική ή σε συνολικά αρνητική άποψη. Οι περισσότερες έρευνες έχουν διεξαχθεί επάνω σε κριτικές προϊόντων όπου οι ορισμοί «θετικό» και «αρνητικό» είναι ξεκάθαροι. Σε άλλες εφαρμογές όπως για παράδειγμα η κατηγοριοποίηση ειδήσεων σε «καλή» και «κακή» παρουσιάζεται μεγαλύτερος βαθμός δυσκολίας καθώς ένα ειδησεογραφικό άρθρο μπορεί να περιέχει «κακά» νέα χωρίς ουσιαστικά να χρησιμοποιεί κανένα υποκειμενικό όρο. Επιπλέον, μεγάλο πλήθος κειμένων περιέχει τόσο θετικές όσο αρνητικές απόψεις ώστε τελικά ο στόχος θα πρέπει να είναι ο προσδιορισμός του κυρίαρχου συναισθήματος εντός αυτού. Με αυτόν τον τρόπο, η εξόρυξη συναισθήματος γίνεται τελικά ένα πολυσύνθετο πρόβλημα κατηγοριοποίησης κειμένου. Σε αντίθεση όμως με την κατηγοριοποίηση κειμένου ανάλογα με το θέμα που περιέχεται, όπου και ο διαχωρισμός του λεξιλογίου είναι ξεκάθαρος, στην κατηγοριοποίηση με βάση το συναίσθημα το αντίστοιχο λεξιλόγιο παρουσιάζει ομοιότητα και διαφέρει μόνο σε μερικές κρίσιμες λέξεις, οι οποίες δυνητικά είναι οι μοναδικές που κρίνουν την κατηγορία στην οποία εμπίπτει. Όταν πρόκειται για αρχεία που περιέχουν «μιξαρισμένη» γνώμη, η κατηγορία τελικά επιλέγεται να είναι ένας συνδυασμός θετικού και αρνητικού.

Οι παραπάνω δύο εργασίες μπορούν να γίνουν σε πολλά επίπεδα: σε επίπεδο όρου, φράσης, πρότασης, ή αρχείου. Είναι επίσης σύνηθες να χρησιμοποιείται η έξοδος

---

<sup>1</sup> Vasileios Hatzivassiloglou, Janyce Wiebe: Effects of Adjective Orientation and Gradability on Sentence Subjectivity.

<sup>2</sup> Farah Benamara, Carmine Cesarano, Antonio Picariello, Diego Reforgiato: Sentiment Analysis: Adjectives and Adverbs are better than Adjectives Alone.

<sup>3</sup> Learning Subjective Language, Janyce Wiebe, Theresa Wilson, Rebecca Bruce, Matthew Bell, Melanie Martin

ενός επιπέδου ως είσοδος για τα υψηλότερα επίπεδα (Turney and Littman, 2003; Dave et al., 2003; Kanayama et al., 2004). Παραδείγματος χάριν, μπορούμε να εφαρμόσουμε μία διαδικασία εξόρυξης συναισθήματος σε μία φράση και έπειτα να κάνουμε χρήση της εν λόγω πληροφορίας ώστε να αποτιμήσουμε προτάσεις, έπειτα παραγράφους και ακολούθως. Φυσικά, έχουν δημιουργηθεί διαφορετικές τεχνικές που εφαρμόζονται ορθότερα ανάλογα με το επίπεδο κατηγοριοποίησης.

Η τρίτη εργασία, συμπληρωματική της κατηγοριοποίησης της πόλωσης, είναι η ανακάλυψη του αντικειμένου- στόχου στο οποίο αναφέρεται η γνώμη που εντοπίστηκε προηγουμένως. Η δυσκολία της παρούσας εργασίας εξαρτάται κατά πολύ από το τομέα της ανάλυσης. Όπως αναφέρθηκε και προηγουμένως, είναι σχετικά ασφαλές να υποθέσουμε ότι μία κριτική προϊόντος αναφέρεται στο συγκεκριμένο και μόνο προϊόν. Από την άλλη πλευρά και θέλοντας να προσεγγίσουμε μία γενικότερη ανάλυση, τα περιεχόμενα των ιστοσελίδων και τα άρθρα των ιστολογίων δεν ανήκουν πάντα σε μία προκαθορισμένη θεματική περιοχή και συχνά υπάρχει αναφορά πολλών και διαφορετικών αντικειμένων σε αυτά. Έτσι λοιπόν, μία σημαντική περιοχή έρευνας έχει αποτελέσει ο προσδιορισμός αντικειμένων, δοθέντος δηλαδή ενός κειμένου να μπορούν να εξαχθούν τα επιμέρους συστατικά του, ως αντικείμενα αυτού. Προσδιορίζοντας τα εν λόγω συστατικά οδηγούμαστε σε μία πιο ακριβή ανάλυση των συναισθημάτων που περιέχονται σε ένα κείμενο και μία πιο λεπτομερή ανακεφαλαίωση των αποτελεσμάτων. Άλλωστε πολλές φορές σε μία πρόταση που περιέχει συναισθημα αναφέρονται περισσότερα του ενός αντικειμένου, όπως συμβαίνει στην περίπτωση των συγκριτικών προτάσεων. Αυτές οι προτάσεις μπορούν να προσδιοριστούν από την αντίστοιχη ύπαρξη συγκριτικών επιθέτων και επιρρημάτων, υπερθετικών επιθέτων και κάποιων άλλων χαρακτηριστικών λέξεων όπως ίδιος, διαφέρω, προτιμώ και άλλων συναφών. Αφού γίνει ο παραπάνω προσδιορισμός, τα εν λόγω αντικείμενα μπορούν να παρατεθούν σε μία περισσότερο αντιπροσωπευτική ως προς τα προτερήματά τους σειρά, όπως αυτή προκύπτει μέσα από το κείμενο.

## 2.5 Πλεονεκτήματα

Έπεται η παράθεση των εν δυνάμει πλεονεκτημάτων που μπορεί να επιφέρει η αυτόματη εξόρυξη συναισθήματος και κάποιων εκ των πιο χαρακτηριστικών παραδειγμάτων τομέων που μπορούν να επωφεληθούν από την ανάπτυξη της τεχνικής αυτής. Μέχρι τώρα, ο παραδοσιακός τρόπος λήψης της γνώμης χρηστών, καταναλωτών ή πολιτών υπήρξαν οι έρευνες και δημοσκοπήσεις επί κάθε ζητήματος ενδιαφέροντος. Σαφώς, τα προβλήματα που ανακύπτουν από μία τέτοιου είδους προσέγγιση σχετίζονται κυρίως με τη δαπάνη σημαντικού χρόνου καθώς και χρηματικού ποσού για την κατάστρωση των εν λόγω ερευνών/ δημοσκοπήσεων, τη διανομή τους και την τελική αξιολόγηση αυτών. Επιπλέον, οι παραπάνω μέθοδοι στηρίζονται κατά κόρον στην «καλή θέληση» των συμμετεχόντων, γεγονός που πολλές φορές καθιστά την αξιοπιστία των αποτελεσμάτων τους αμφίβολη. Ακόμη, θα μπορούσαμε να ισχυριστούμε ότι με την έλευση της άντλησης των ζητούμενων πληροφοριών με αυτοματοποιημένο τρόπο από διαδικτυακές πηγές η διεξαγωγή κάθε είδους ερευνών μέσω ερωτηματολογίων

κρίνεται δυνητικά παρωχημένη. Σημαντικό παράδειγμα που τεκμηριώνει τον παραπάνω ισχυρισμό είναι η ευρεία ανάπτυξη και διάδοση των προσωπικών ιστολογιών, με κύριο σκοπό την παράθεση σκέψεων, παρουσίαση ιδεών και ανταλλαγή γνώσης, γεγονός που τα καθιστά μια από τις πιο πολύτιμες πηγές εξόρυξης γνώμης και γενικότερης ανάλυσης τάσεων. Έτσι λοιπόν, τα διάφορα σχόλια στα άρθρα των ιστολογιών αποτελούν μία πλούσια πηγή πληροφοριών που αντικατοπτρίζει άμεσα τις απόψεις των χρηστών σε μία ευρεία γκάμα θεματικών περιοχών και η αυτοματοποιημένη εξόρυξη και μετέπειτα ταξινόμηση αυτών θα μπορούσε αναμφίβολα να αντικαταστήσει τις πατροπαράδοτες έρευνες και δημοσκοπήσεις, ξεπερνώντας μάλιστα αυτές σε αξιοπιστία.

Εδώ κρίνεται σκόπιμο να εξεταστεί αναλυτικότερα το παράδειγμα των ιστολογιών, παραθέτοντας κάποια στατιστικά στοιχεία και επιχειρώντας μία προσέγγιση του πως η αυτόματη εξόρυξη συναισθήματος έχει συνδεθεί πλέον με την λεγόμενη «φωνή του καταναλωτή». Ο προσδιορισμός των απόψεων, όπως αυτές διατυπώνονται μέσω των σχολίων στα διάφορα ιστολόγια, μπορεί να χρησιμοποιηθεί για την εξαγωγή γνώσης σχετικά με το πώς βλέπουν οι καταναλωτές ένα προϊόν ή μία υπηρεσία. Σχετικές μελέτες έχουν δείξει ότι:

- το 60% των κατοίκων των Ηνωμένων Πολιτειών έχουν κάνει διαδικτυακή έρευνα αγοράς τουλάχιστον μία φορά, ενώ το 15% αυτών κάνει κάτι τέτοιο καθημερινά.
- ένα ποσοστό που κυμαίνεται μεταξύ 73%-87% ισχυρίζεται ότι οι διαδικτυακές κριτικές που διάβασε είχαν σημαντική επιρροή στην αγορά του.
- το 30% των κατοίκων των Ηνωμένων Πολιτειών έχει γράψει μία διαδικτυακή κριτική ή έχει παραθέσει ένα αντίστοιχο σχόλιο.
- όμως, το 58% των χρηστών του Internet στις Ηνωμένες Πολιτείες δηλώνει ότι οι διαθέσιμες πληροφορίες ήταν ελλιπείς, μπερδεμένες και πολυάριθμες.

Ήταν πλέον εμφανές ότι η δημιουργία τεχνολογιών που θα μπορούσαν να εξαγάγουν και να αναλύουν την άποψη των χρηστών θα έλυne ένα τεράστιο πρόβλημα της ανάγκης για πληροφορία. Προς αυτή την κατεύθυνση, η ανάλυση κειμενικών πόρων απέκτησε τεράστια σημασία κυρίως λόγω της συνειδητοποίησης ότι η πληθώρα αναξιόποιτης έως τότε πληροφορίας θα μπορούσε να χρησιμοποιηθεί προς όφελος εταιριών και οργανισμών. Οι αυτόματες τεχνικές εξόρυξης κειμένου όπως η κατηγοριοποίηση, η εξόρυξη συναισθήματος και η περίληψη επιτρέπουν ικανοποιητική και επαρκή επεξεργασία των σχολίων των καταναλωτών και αναδεικνύουν τα σημαντικότερα θέματα καθώς και τα κυρίαρχα συναισθήματα. Ταυτοποιώντας τα θέματα τα οποία σχολιάζονται και κατηγοριοποιώντας το συναίσθημα που περιέχεται μπορούμε τελικά να χαρακτηρίσουμε την συνολική άποψη που εκφράζεται. Αξιοποιώντας την εν λόγω γνώση, οι εταιρίες και οι οργανισμοί μπορούν να βελτιώσουν σημαντικά την αξία των παρεχόμενων προϊόντων και υπηρεσιών, καθιερώνοντας νέα επιχειρησιακά μοντέλα προσαρμοσμένα κατάλληλα στις ανάγκες των πελατών. Κάτι τέτοιο συμβάλλει στην



καινοτομία και θέτει πλεονέκτημα έναντι του ανταγωνισμού καθώς λαμβάνει υπόψη τις συνεχώς μεταβαλλόμενες ανάγκες που προκύπτουν στην αγορά. Από τη μία πλευρά, η «φωνή των καταναλωτών» όπως παρουσιάζεται μέσω διαδικτυακών πηγών βοηθάει στο να προσδιοριστούν νέες ευκαιρίες για εφεύρεση και βελτίωση των προϊόντων ενώ από την άλλη συνεισφέρει στον ευρύτερο προσδιορισμό των τάσεων της αγοράς και των προτιμήσεων των καταναλωτών που οδηγεί σε καινούριες επιχειρησιακές ευκαιρίες.

Συνοψίζοντας έως εδώ, αναφέρουμε ότι υπάρχει πληθώρα πρόσθετων πλεονεκτημάτων στην αυτοματοποιημένη εξόρυξη γνώμης, σε όποιο πεδίο και αν αυτή εφαρμόζεται. Πρωτίστως, τα άτομα που μοιράζονται τις απόψεις τους συνήθως τείνουν να έχουν περισσότερο παγιωμένες γνώμες από το μέσο όρο, κάτι που επιπρόσθετα επηρεάζει τους αναγνώστες τους στα πλαίσια μιας διάδοσης «από στόμα σε στόμα». Η εξαγωγή τέτοιου είδους απόψεων κρίνεται επομένως πολύτιμη. Δευτερευόντως, οι απόψεις εξάγονται σε πραγματικό χρόνο και άμεσα, επιτρέποντας καλύτερους χρόνους αντίδρασης σε αλλαγές, βασισμένων στην παράμετρο των τάσεων συναρτήσεως του χρόνου. Άλλωστε, η εξόρυξη συναισθήματος είναι άρρηκτα συνδεδεμένη με την ίδια την γνώμη του συγγραφέα σχετικά με ένα θέμα, παρά με απλώς το ίδιο το θέμα αυτό καθ' αυτό. Ένας αναλυτής, ιδανικά, θα επιθυμούσε να έχει στη διάθεσή του μία ανακεφαλαίωση των αποτυπωμένων απόψεων παρά να διαβάζει πολυάριθμα σχόλια και συζητήσεις χρηστών. Οι οργανισμοί, επίσης, επωφελούνται από την αυτοματοποιημένη εξόρυξη γνώμης διατηρώντας μία συνεχή εικόνα του πώς οι χρήστες, καταναλωτές ή πολίτες αισθάνονται απέναντι στις υπηρεσίες τους, αποσπώντας συγκεκριμένες πληροφορίες που βρίσκονται «κρυμμένες» μέσα σε ένα κείμενο όπως παράπονα, προβλήματα και γενικότερες παρατηρήσεις.

## 2.6 Προκλήσεις / Δυσκολίες

Για μία πιο ολοκληρωμένη άποψη των όσων αναφέρθηκαν μέχρι στιγμής, θα πρέπει να καταστεί σαφές ότι η κατηγοριοποίηση της άποψης σε θετική και αρνητική είναι κάτι περίπλοκο, ακόμα και για τους ίδιους τους ανθρώπους. Για παράδειγμα, ενώ έχουν υπάρξει προσπάθειες για κατηγοριοποίηση περιεχομένου βασισμένη αποκλειστικά στο ίδιο το κείμενο, τα σχόλια πάνω σε αυτό και οι παραθέσεις από άλλες πηγές θεωρούνται επίσης μέρος του κειμένου αυτού. Όταν, λοιπόν, ένα άρθρο ή μία κριτική αντίστοιχα γίνονται αντικείμενο συζήτησης η κατηγοριοποίηση άποψης συνολικά μπορεί να γίνει πολύ δύσκολη. Αυτό ισχύει κυρίως όταν πρόκειται για πολιτικές συζητήσεις, όπου οι χρήστες-σχολιαστές προβαίνουν συχνά σε συγκρίσεις μεταξύ προσώπων, πολιτικών τακτικών ή γεγονότων. Ακόμα περισσότερο, η παράθεση αστείων, ιδιωματισμών, ειρωνείας και πολιτιστικών αναφορών δυσχεραίνει την αυτόματη εξόρυξη συναισθήματος καθώς απαιτεί ευρύτερη γνώση και εξοικείωση με τα σχετικά γεγονότα και τις αναφορές. Ο παράγοντας αυτός καθιστά την κατηγοριοποίηση κειμένου ιδιαίτερα πολύπλοκη για τους περισσότερους αλγόριθμους. Σε σχετικές μελέτες έχει παρατηρηθεί ότι οι αλγόριθμοι κατηγοριοποίησης τείνουν να έχουν μεγαλύτερη ακρίβεια όταν πρόκειται για περισσότερο τεχνικά θέματα σε σύγκριση με κοινωνικά. Πιο

συγκεκριμένα, η ακρίβεια κατηγοριοποίησης κριτικών για αυτοκίνητα και τράπεζες αγγίζει το 84% και 80% αντίστοιχα, ενώ οι ίδιες μέθοδοι παρέχουν ακρίβεια 66% για ταινίες και 71% για ταξιδιωτικούς προορισμούς. Όσον αφορά πολιτικά ζητήματα η ακρίβεια τείνει να κυμαίνεται σε ακόμα χαμηλότερα ποσοστά.

Επιπλέον, η εξαγωγή γνώμης θέτει ένα πλήθος προκλήσεων τόσο από πλευράς έρευνας όσο και από μηχανικής μάθησης. Συγκεκριμένα παρατίθενται:

- **Εισαγωγή κειμένων:** Είναι πολύ δύσκολη η ανάπτυξη ενός γενικού αλγορίθμου που λειτουργεί με συνέπεια σε όλα τα είδη σωμάτων κειμένου διότι κάθε είδος έχει τις δικές του ιδιαιτερότητες από πλευράς διάταξης (layout), έκτασης και λεξιλογίου. Χάρην παραδείγματος, τα άρθρα των ιστολογιών είναι πολύ πιο «θορυβώδη» και συνήθως περιέχουν περισσότερες διαφημίσεις και παραπλανητικές πληροφορίες (spam) σε σχέση με τα αντίστοιχα των επίσημων ιστοτόπων ενημέρωσης. Επιπλέον, στα πρώτα απαντώνται συχνότερα φαινόμενα ανορθογραφίας και χρήσης αργκούς, κάτι που δυσχεραίνει περαιτέρω το όλο εγχείρημα.
- **Μονάδα μέτρησης κατηγοριοποίησης:** Η άποψη που εκφράζεται σε ένα κείμενο μπορεί να διατρέχει όλο το κείμενο ή να αναφέρεται σε ένα μικρότερο μέρος, μία παράγραφο, μία πρόταση ή ακόμη στο μέρος μίας πρότασης. Η επιλογή της κατάλληλης μονάδας συνδέεται αναπόφευκτα με την ποσότητα της απαιτούμενης γλωσσικής επεξεργασίας και κατ' αυτόν τον τρόπο επηρεάζει την αποδοτικότητα του εγχειρήματος.
- **Αίσθημα/ γνώμη σε σχέση με μία οντότητα:** Ενώ τα περισσότερα πειράματα που περιγράφονται στην ακαδημαϊκή βιβλιογραφία γενικά κατηγοριοποιούν μία πρόταση ή ένα κείμενο με βάση το μεταδιδόμενο αίσθημα, είναι πιο ενδιαφέρον να αναζητηθεί το τελευταίο επί μίας συγκεκριμένης οντότητας (π.χ. μίας υπηρεσίας). Σε αυτήν την περίπτωση απαιτείται ένα επιπλέον βήμα, αυτό της «εξαγωγής» μονάδων σχετικών με την οντότητα- είσοδο, γεγονός που αποτελεί καθ' αυτό μία πρόκληση.
- **Πολυγλωσσία:** Η ανάπτυξη αλγορίθμων σε λέξεις εκτός της αγγλικής περιλαμβάνει τη δόμηση λεξικών και σωμάτων εκπαίδευσης (training corpora) σε αυτές τις γλώσσες, μία διαδικασία ταυτόχρονα δημιουργική και δαπανηρή.
- **Ποιότητα:** Η ποιότητα της κατηγοριοποίησης συναισθήματος μετριέται με όρους βαθμού ορθότητας (το ποσοστό των παραδειγμάτων που κατηγοριοποιούνται σωστά από τον αλγόριθμο)

ή με όρους ακρίβειας και ανάκλησης (το πλήθος των σωστών παραδειγμάτων αποτελεί την ακρίβεια ενώ το πλήθος των σωστών παραδειγμάτων που ανιχνεύτηκαν από τον αλγόριθμο επί του συνόλου αυτών ονομάζεται ανάκληση). Η επίτευξη βαθμού ποιότητας αρκούντως υψηλού για χρήση προσανατολισμένη στα αποτελέσματα παραμένει ένα δύσκολο εγχείρημα και μία συνεχής πρόκληση για τους ερευνητές.

## ***Κεφάλαιο 3. Ηλεκτρονική Διακυβέρνηση και Opinion Mining***

---

### 3.1 Πολιτική 2.0

Ανακεφαλαιώνοντας με την εξόρυξη συναισθήματος, θα αναλύσουμε την συνεισφορά που μπορεί να έχει στον τομέα της ηλεκτρονικής διακυβέρνησης, αξιοποιώντας πάντα τα πλεονεκτήματα του Ιστού 2.0 (Web 2.0). Με την άφιξή του, ο Ιστός 2.0, αναμφίβολα, δημιούργησε ενθουσιασμό και συνέβαλε στην δημιουργία μιας πληθώρας νέων ευκαιριών προς την βελτιστοποίηση επιχειρήσεων και κυβερνήσεων. Οι τελευταίες προεδρικές εκλογές στις ΗΠΑ, υπήρξαν μία πρώτη ένδειξη της επιτυχίας που μπορούν να έχουν οι διαδικτυακές καμπάνιες και η πολιτική συμμετοχή μέσω του Ιστού 2.0. Σε αυτό που αναφέρεται ως «Πολιτική 2.0», οι πολιτικοί χρησιμοποιούν σύγχρονες διαδικτυακές πλατφόρμες για επιτυχή συζήτηση, διαφήμιση της εκστρατείας τους, ανακοίνωση γεγονότων, κινητοποίηση ψηφοφόρων ακόμα και για δωρεές. Καινοτομίες προς αυτή την κατεύθυνση, είναι η δημιουργία ιστοσελίδων και προσωπικών ιστολογίων πολιτικών προσώπων, wikis για πολιτικά θέματα και συζητήσεις, αμεσότητα και διαφάνεια στο γενικότερο σχεδιασμό και τις δαπάνες, κοινωνίες ειδικών σχολιαστών και πολλά ακόμα. Καθώς οι πολιτικές διαδικασίες γίνονται πιο συμμετοχικές με την ύπαρξη διαδικτυακών πηγών, γεννιέται για σημαντική ευκαιρία για ανάδειξη και καθιέρωση ευφών συστημάτων έρευνας σχετικά με την ηλεκτρονική διακυβέρνηση και τις εφαρμογές του ευρύτερου πλαισίου της «Πολιτικής 2.0». Έτσι λοιπόν, επιλεγμένες εφαρμογές των πεδίων της εξόρυξης δεδομένων, κειμένου αλλά και συναισθήματος μπορούν να υποστηρίξουν την διαδικτυακή πολιτική συμμετοχικότητα, την ηλεκτρονική δημοκρατία, τα πολιτικά blogs και forums, τη διάθεση υπηρεσιών ηλεκτρονικής διακυβέρνησης, την υπευθυνότητα και τη διαφάνεια.

Σε αντίθεση με το μοντέλο βασικής πρόσβασης, δημοσίευσης και μεταφοράς περιεχομένου του Ιστού 1.0, ο Ιστός 2.0 πρόσθεσε διαλειτουργικότητα, συμμετοχή, περιεχόμενο πλούσιο και προσανατολισμένο στις υπηρεσίες. Ο όρος Ιστός 2.0, που επινοήθηκε από τον Tim O'Reilly, «είναι η τάση για χρησιμοποίηση τεχνολογιών του παγκόσμιου ιστού και σχεδιασμού του διαδικτύου που αποσκοπεί στην εξυπηρέτηση της δημιουργικότητας, της ανταλλαγής πληροφορίας και της συνεργασίας μεταξύ των χρηστών. Ο Ιστός 2.0 είναι μία επιχειρηματική επανάσταση στην βιομηχανία των υπολογιστών, που προκλήθηκε από την εξέλιξη του διαδικτύου ως πλατφόρμας, και μία προσπάθεια να κατανοηθούν οι κανόνες της επιτυχίας της πλατφόρμας αυτής».

Προς το παρόν, συνεχής πρόκληση για τις κυβερνήσεις αποτελεί η μετάβαση από την απλή διανομή υπηρεσιών στην παροχή προσαρμοσμένων εφαρμογών αναλόγων των εκάστοτε αναγκών. Με άλλα λόγια, θα μπορούσαμε να ισχυριστούμε ότι η επιτυχία των κυβερνήσεων είναι πλέον συνυφασμένη με την επιτυχή μεταφορά των μηνυμάτων τους στους πολίτες και το χτίσιμο δυνατών συμμαχιών μαζί τους, ενισχύοντας την συμμετοχή των τελευταίων στη διαδικασία λήψης αποφάσεων. Το διαδίκτυο έχει τη δυνατότητα να αλλάξει ριζικά το προφίλ των κυβερνήσεων, βελτιώνοντας την επικοινωνία μεταξύ πολιτών και δημόσιων υπηρεσιών. Πλέον, είναι γνωστό ότι οι κυβερνήσεις σε παγκόσμια κλίμακα έχουν λανσάρει φιλόδοξα σχέδια για τη δημιουργία ηλεκτρονικών εφαρμογών και

υπηρεσιών προς την ευρύτερη κατεύθυνση της ηλεκτρονικής διακυβέρνησης. Ο στόχος της καθιέρωσης των ηλεκτρονικά παρεχόμενων υπηρεσιών είναι πολυδιάστατος: να διευκολύνει τον δημόσιο τομέα στην επαφή του με τους πολίτες μέσω ενός απλού, άμεσου και οικονομικού τρόπου, να αυξήσει την υπευθυνότητα των κυβερνήσεων έναντι των πολιτών, να μειώσει την γραφειοκρατία και την διαφθορά μέσω της διαδραστικής σχέσης μεταξύ πολιτών, δημόσιου και ιδιωτικού τομέα και της κυβέρνησης. Παρά τους πόρους που χρειάζεται να κατανεμηθούν για την επιτυχή καθιέρωση της ηλεκτρονικής διακυβέρνησης, οι περισσότεροι τείνουν να αντιλαμβάνονται αυτή ως ένα σύστημα υπηρεσιών που υποστηρίζει τις ενέργειες των κυβερνήσεων ενώ παράλληλα θέτει ζητήματα αντιμετώπισης των επιρροών των εν λόγω ενεργειών τόσο σε κοινωνικό όσο σε πολιτικό επίπεδο.

### **3.2 Ηλεκτρονική Διακυβέρνηση και Εξόρυξη Δεδομένων**

Η εκμετάλλευση τεχνικών εξόρυξης δεδομένων πρωτίστως για τον εντοπισμό της κοινής γνώμης (όπως αυτή εκφράζεται διαδικτυακά) για τις κυβερνητικές αποφάσεις και δευτερευόντως για την ανάλυση της πόλωσης αυτών αποτελεί σημαντικό παράγοντα για την επιτυχία του εγχειρήματος της ηλεκτρονικής διακυβέρνησης. Στον τομέα αυτό έχουν διεξαχθεί αρκετές έρευνες και έχουν προταθεί αντίστοιχες μεθοδολογίες από ερευνητές ώστε με την αξιολόγηση των σχολίων από διάφορες διαδικτυακές πηγές, με έμφαση στα blogs και forums, να υπάρχει μια συνεχής αποτίμηση της άποψης των πολιτών σχετικά με τις κυβερνητικές αποφάσεις. Η γνώση αυτή είναι πολύτιμη καθώς μπορεί να αποτελέσει το υπόβαθρο που θα οδηγεί τις κυβερνήσεις στις κατάλληλες ενέργειες και αποφάσεις για την ταυτόχρονη ικανοποίηση και των δύο πλευρών. Η αναβάθμιση των κυβερνητικών ιστοτόπων κρίνεται απαραίτητη, ενσωματώνοντας παράλληλα εφαρμογές που θα παρέχουν την δυνατότητα να ακουστεί η φωνή των πολιτών, κάτι που μέχρι τώρα δεν ήταν εφικτό για το ευρύ κοινό. Με αυτό τον τρόπο η ηλεκτρονική διακυβέρνηση μετατρέπεται από μία τεχνολογική υποδομή σε ένα διαδραστικό εγχείρημα που ενισχύει την ηλεκτρονική συμπερίληψη και συμμετοχή. Μία πλήρης συμμετοχική κοινωνία πληροφορίας θα εκμεταλλευτεί τις εξελίξεις στον τομέα «Τεχνολογίας Πληροφορικής και Επικοινωνιών (ΤΠΕ)» ώστε να καθιερωθεί η ηλεκτρονική διακυβέρνηση με όλες τις διαθέσιμες δυνατότητες της. Αναμένεται άλλωστε η ηλεκτρονική διακυβέρνηση να διευρυνθεί ώστε να αντιμετωπίσει μία σειρά προκλήσεων όπως η κλιματική αλλαγή ή η τρομοκρατία, να εξασφαλίσει καλύτερες υπηρεσίες, μεγαλύτερη ασφάλεια και καλύτερες δημοκρατικές διαδικασίες, να παρέχει δημόσιες υπηρεσίες ανεξαρτήτως συνόρων και να αυξήσει τις ευκαιρίες των πολιτών προς την κατεύθυνση της επιχειρηματικότητας.

#### **3.2.1 Στάδια Ηλεκτρονικής Διακυβέρνησης**

Λίγα χρόνια πριν, η ηλεκτρονική διακυβέρνηση περιείχε δύο στάδια (Reddick C, 2004): το πρώτο αναφερόταν στην φάση διάδοσης πληροφορίας ώστε να είναι διαθέσιμη για δημόσια χρήση και το δεύτερο περιείχε την διάθεση ηλεκτρονικά

διαθέσιμων υπηρεσιών όπως την πληρωμή φόρων μέσω διαδικτύου. Επομένως, θεωρούνταν περισσότερο σαν ένας τρόπος για τις κυβερνήσεις να κάνουν χρήση των πιο καινοτόμων Τεχνολογιών Πληροφορίας και Επικοινωνιών ώστε να προσφέρουν στους πολίτες, στις επιχειρήσεις και στους υπαλλήλους επαρκή πρόσβαση σε πληροφορίες και υπηρεσίες (Hayat 2009). Η ηλεκτρονική διακυβέρνηση συμβάλλει στην ανάπτυξη προσωπικά προσαρμοσμένων υπηρεσιών μεταξύ πολιτών και της κυβέρνησής τους ενώ η μείωση του κόστους λειτουργίας και διαχείρισης υπήρξε κινητήρια δύναμη για την εφαρμογή υποδομών ηλεκτρονικής διακυβέρνησης. Εδώ και περισσότερα από δέκα χρόνια οι κυβερνήσεις σε όλο τον κόσμο οργανώνουν και παρουσιάζουν τα δικά τους σχέδια προς αυτή την κατεύθυνση, ενώ αξίζει να σημειωθεί ότι ανάμεσα στους πρωτοπόρους ήταν η κυβέρνηση της Αυστραλίας που ανέπτυξε τις δικές της στρατηγικές από τις αρχές της δεκαετίας του 1990.

Αρκετοί ερευνητές έχουν προσπαθήσει να προσδιορίσουν από ποιες δημόσιες υπηρεσίες που θα διατίθενται ηλεκτρονικά θα επωφεληθούν περισσότερο οι επιχειρήσεις και οι πολίτες. Έτσι δημιουργήθηκε μία λίστα με 20 βασικές υπηρεσίες, των οποίων η ανάπτυξη ελέγχεται από την Ευρωπαϊκή Επιτροπή σε όλες τις φάσεις αυτών. Οι 12 από αυτές είναι υπηρεσίες προς τους πολίτες και είναι οι ακόλουθες: φόρος εισοδήματος, υπηρεσίες αναζήτησης εργασίας, εισφορές κοινωνικής ασφάλισης, προσωπικά έγγραφα (διαβατήριο και άδεια οδήγησης), καταχώρηση οχήματος, έκδοση οικοδομικής άδειας, δήλωση προς την αστυνομία (π.χ. σε περίπτωση κλοπής), δημόσιες βιβλιοθήκες, πιστοποιητικά (γέννησης και γάμου), εισαγωγή στην ανώτατη εκπαίδευση, ανακοίνωση μετακόμισης, υπηρεσίες υγείας. Οι υπόλοιπες 8 ανήκουν στην κατηγορία υπηρεσιών προς επιχειρήσεις: εισφορές κοινωνικής ασφάλισης προς τους εργαζομένους, φόρος επιχειρήσεων, ΦΠΑ (δήλωση και ειδοποίηση εκκαθάρισης), έναρξη επιχείρησης, υποβολή στοιχείων σε στατιστικές υπηρεσίες, τελωνειακές διασαφήσεις, περιβαλλοντικές άδειες, δημόσιες προμήθειες. Ωστόσο, ακόμα και αν το διαδίκτυο φαίνεται να είναι ο μοχλός των κυβερνήσεων και των πολιτών στα πλαίσια διαδραστικότητας και συμμετοχικής δημοκρατίας, η ηλεκτρονική διακυβέρνηση είναι κάτι πολύ περισσότερο από την παροχή υπηρεσιών και πληροφοριών online: είναι η εφαρμογή νέων τεχνολογιών για την παράδοση κυβερνητικών υπηρεσιών με ένα πιο πειστικό για τους πολίτες τρόπο. Με στόχο τη δημιουργία υπηρεσιών υψηλής ποιότητας με επίκεντρο τον πολίτη και παροχή αξιόπιστων πληροφοριών σε σύγκριση με τον μέχρι τώρα παραδοσιακό τρόπο, οι κυβερνητικοί φορείς πρέπει πρώτα να κατανοήσουν τους παράγοντες που επηρεάζουν τους πολίτες προς την υιοθέτηση των εν λόγω καινοτομιών.

Στην προσέγγιση των δύο σταδίων προστέθηκε μετέπειτα το στάδιο που αποτελείται από την ανατροφοδότηση και την ηλεκτρονική συμμετοχή (Lenova 2009, Osimo 2008), το οποίο θεωρείται ως το κύριο βήμα προς την εξέλιξη της ηλεκτρονικής διακυβέρνησης γενικότερα. Σύμφωνα με αυτό, η πιο παραδοσιακή προσέγγιση της ηλεκτρονικής διακυβέρνησης, που εστιάζει αποκλειστικά στην παροχή υπηρεσιών, θα έπρεπε να αντικατασταθεί με ένα μοντέλο γενικότερης ηλεκτρονικής συμμετοχικότητας και οι διάφορες αξιολογήσεις να έχουν ως επίκεντρο τον ίδιο τον χρήστη. Είναι φανερό άλλωστε ότι το εγχείρημα της

ηλεκτρονικής διακυβέρνησης πρέπει να προσαρμόζεται συνεχώς στις νέες προκλήσεις και ανάγκες. Έτσι λοιπόν, το πρώτο βήμα προς την ηλεκτρονική συμμετοχικότητα είναι να «ακούσεις» την γνώμη του κοινού, κάτι που μπορεί να καλυφθεί από την ανάπτυξη αποτελεσματικών και αξιόπιστων μέσων που θα αφουγκράζονται τις απόψεις όπως αυτές διατυπώνονται μέσω του παγκόσμιου ιστού.

### 3.2.2 Συνεισφορά Εξόρυξης Γνώμης

Η εξόρυξη γνώμης αποτελεί συνώνυμο του να ανακαλύψεις τι σκέφτονται οι άλλοι. Η διαδικασία αυτή, στις περισσότερες περιπτώσεις, αποτελεί πολύτιμη πληροφορία για την λήψη αποφάσεων κατά κοινή ομολογία. Πολύ πριν τη διάδοση του Παγκόσμιου Ιστού, οι άνθρωποι πάντα ρωτούσαν τους φίλους τους ποιον σκοπεύουν να ψηφίσουν στις επερχόμενες εκλογές ή συμβουλευόνταν κριτικές καταναλωτών ώστε να αποφασίσουν ποιο προϊόν θα αγοράσουν. Το Internet έδωσε τη δυνατότητα να μαθαίνουμε τις απόψεις και τις εμπειρίες ανθρώπων που δεν αποτελούν ούτε προσωπικές μας γνωριμίες ούτε επαγγελματίες κριτικούς, στην πραγματικότητα αποτελούν ανθρώπους που δεν γνωρίζουμε τίποτα για αυτούς. Από την άλλη οι εταιρίες εδώ και δεκαετίες προσπαθούν να συλλάβουν τις κοινές τάσεις και την κοινή γνώμη για συγκεκριμένα προϊόντα και υπηρεσίες μέσω ερωτηματολογίων και ερευνών. Φυσικά το θέμα αυτό έχει απασχολήσει κατά πολύ και τον τομέα της πολιτικής καθώς τα πολιτικά πρόσωπα σχεδόν πάντα προσπαθούν να κατανοήσουν τι σκέφτονται οι ψηφοφόροι σχετικά με τις επικείμενες προτάσεις και αποφάσεις τους. Είναι επομένως εμφανές ότι οι αρνητικές και θετικές απόψεις μπορούν να χρησιμοποιηθούν σαν οδηγός για τις εταιρίες προκειμένου να αλλάξουν την στρατηγική τους για διάφορα προϊόντα, για τους καταναλωτές ώστε να αποφασίσουν για τις αγορές τους και πρόσφατα για τις κυβερνήσεις ώστε να βελτιώσουν τις υπηρεσίες τους και να εγκαινιάσουν τις εκστρατείες τους. Εδώ πρέπει να σημειώσουμε, ότι όλο και περισσότεροι άνθρωποι καταθέτουν την άποψή τους σε αγνώστους μέσω του διαδικτύου (Pang and Lee, 2008). Με την έκρηξη των πλατφόρμων του Ιστού 2.0 όπως τα ιστολόγια, οι ηλεκτρονικοί τόποι δημόσιας συζήτησης και διάφορα άλλα είδη κοινωνικών δικτύων, οι χρήστες έχουν πλέον στη διάθεσή τους μία δυνατότητα άνευ προηγουμένου να μοιράζονται τις εμπειρίες και τις απόψεις τους, αρνητικές και θετικές, σχετικά με κάθε είδους θέμα ή αντικείμενο (Zabin and Jefferies, 2008).

### 3.3 Ιστότοποι Ηλεκτρονικής Διακυβέρνησης

Οι ιστότοποι ηλεκτρονικής διακυβέρνησης βρίσκονται ανάμεσα στους μεγαλύτερους υπάρχοντες ιστότοπους σε μέγεθος, αριθμό χρηστών και διάθεσης πληροφοριών. Επομένως, η δημιουργία μιας σημασιολογικής δομής διαδικτύου που θα οργανώσει ουσιαστικά τους εν λόγω ιστότοπους είναι καθ' όλα επιθυμητή. Οι ευρωπαϊκές χώρες έχουν υπογράψει κοινή δέσμευση, να αναπτύξουν ως το 2015 ευφυέστερες ηλεκτρονικές υπηρεσίες τόσο προς τους πολίτες όσο προς τις επιχειρήσεις. Όπως αναφέρθηκε και προηγουμένως, για την επίτευξη του ανωτέρου



στόχου δεν αρκεί απλά ο σχεδιασμός, η υλοποίηση και η παρακολούθηση της απλής μεταφοράς των υφιστάμενων υπηρεσιών σε ηλεκτρονικό περιβάλλον, αλλά απαιτείται καινοτομία που θα μεταβάλλει ουσιαστικά τη σχέση κυβέρνησης και πολίτη. Κυριότερο «τεχνολογικό καταλύτη» προς αυτή την κατεύθυνση αποτελούν οι δυνατότητες που προσφέρει η δεύτερη γενιά διαδικτύου, που χαρακτηρίζονται από την ανάπτυξη εργαλείων ευρύτερης επικοινωνίας και ηλεκτρονικής συνεργασίας. Ο τρόπος αλληλεπίδρασης των χρηστών αναπτύσσεται με ραγδαίο ρυθμό, αντικαθιστώντας τον όρο «ενημερώνομαι» με όρους όπως «μοιράζομαι», «εκφράζομαι», «επηρεάζω», «δικτυώνομαι». Οι κυβερνήσεις σε παγκόσμια κλίμακα δεν θα μπορούσαν παρά να συμβαδίσουν με την νέα αυτή τάση, δημιουργώντας ιστότοπους που δεν αρκούνται στην απλή μεταφορά πληροφοριών και διεκπεραίωση των αντίστοιχων δημόσιων υπηρεσιών διαδικτυακά αλλά στοχεύουν στην επικοινωνία με τον πολίτη, τον αφουγκρασμό της κοινής γνώμης και την συμμετοχικότητα που θα τις οδηγήσει σε ορθότερες και αποτελεσματικότερες δράσεις και αποφάσεις μελλοντικά. Ακολουθεί η παράθεση μερικών εκ των πιο χαρακτηριστικών και πιο ενδιαφερόντων ηλεκτρονικών τόπων που σχετίζονται με την ηλεκτρονική διακυβέρνηση και την ταυτόχρονη ηλεκτρονική συνεργασία κυβέρνησης και πολιτών.

### 3.4 Παραδείγματα

Αρχικά παραθέτουμε την κυβερνητική πλατφόρμα: <http://change.gov> που δημιουργήθηκε από τον πρόεδρο των Ηνωμένων Πολιτειών και το επιτελείο αυτού. Μετά την επιτυχημένη προεκλογική εκστρατεία, που στηρίχθηκε για πρώτη φορά στα χρονικά στον Παγκόσμιο Ιστό με έμφαση στα κοινωνικά δίκτυα, η δημιουργία του συγκεκριμένου διαδικτυακού τύπου ήρθε ως φυσιολογική εξέλιξη. Οι πολίτες μπορούν να βρουν όλες τις πληροφορίες γύρω από την ατζέντα του νυν προέδρου και να ενημερώνονται άμεσα, όμως το σημαντικότερο γεγονός αποτελεί η παρότρυνση της συμμετοχής των τελευταίων στα κοινά. Οι δημιουργοί της σελίδας μοιράζονται τα νέα σχέδια και δράσεις του προεδρικού επιτελείου με τους πολίτες, αποσκοπώντας στην ενεργή τους συμμετοχή και την παράθεση απόψεων και ιδεών. Επιπλέον, σημαντική πρωτοβουλία θεωρείται η υποβολή ερωτήσεων από μεριάς πολιτών σχετικά με κάθε είδους ζητήματα που τους απασχολούν και οι αντίστοιχες απαντήσεις εκ μέρους του προεδρικού επιτελείου. Χαρακτηριστικές είναι και οι επιλογές: «Share your vision» και «Share your story» που παροτρύνουν ακόμα περισσότερο τους χρήστες να καταθέσουν τις προσωπικές τους εμπειρίες αλλά και ιδέες για την βελτίωση της καθημερινότητάς τους αλλά και γενικότερα της λειτουργίας του δημοκρατικού πολιτεύματος. Αναμφίβολα, το εν λόγω site αποτελεί ένδειξη της μετάβασης προς την ηλεκτρονική δημοκρατία (e-democracy), την ανάγκη ηλεκτρονικής διαδραστικότητας κυβέρνησης και πολιτών και όχι απλά την ηλεκτρονική διάθεση υπηρεσιών στα πλαίσια της ηλεκτρονικής διακυβέρνησης καθώς και την ανάγκη ανάπτυξης σύγχρονων μεθόδων εξόρυξης κειμένου και συναισθήματος για την ταχύτερη αξιολόγηση του παρεχόμενου υλικού. Δίνονται τα αντίστοιχα screenshots ώστε ο αναγνώστης να αποκτήσει μία καλύτερη εικόνα των όσων αναλύθηκαν σχετικά με την λειτουργία της συγκεκριμένης διαδικτυακής πλατφόρμας:

change.gov/page/content/americanmoment

HOME | NEWSROOM | BLOG | LEARN | AGENDA | AMERICAN MOMENT | AMERICA SERVES

## An American Moment

[Share Your Story](#)  
[Share Your Vision](#)

The story of the campaign and this historic moment has been your story. It is about the great things we can do when we come together around a common purpose. The story of bringing this country together as a healed and united nation will be led by President-Elect Obama, but written by you. The millions of you who built this campaign from the ground up, and echoed your call for the change you wanted to see implemented by the Obama Administration - this process of setting up that new government is about you.

This transition is about selecting a new staff and agenda that will help reclaim the American dream and bring about positive lasting change to this country. In order to do that, we want to hear from *you*.

**Tell us your story** and the issues that matter most to you. **Share with us your concerns and hopes** – the policies you want to see carried out in the next four years.

"I ask you to believe - not just in my ability to bring about change, but in yours. I know this change is possible...because in this campaign, I have had the privilege to witness what is best in America. I've seen it in lines of voters that stretched around schools and churches; in the young people who cast their ballot for the first time, and those not so young folks who got involved again ;after a very long time. I've seen it in the workers who would rather cut back their hours than see their friends lose their jobs; in the neighbors who take a stranger in when the floodwaters rise; in the soldiers who re-enlist after losing a limb. I've seen it in the faces of the men and women I've met at countless rallies and town halls across the country, men and women who speak of their struggles but also of their hopes and dreams."

OPEN GOVERNMENT

  
**CITIZEN'S BRIEFING BOOK**  
Give us your ideas and we'll give them to the President.

  
**JOIN THE DISCUSSION**  
Our policy teams will be sharing new developments with you, the American people, and asking for feedback.

  
**YOUR SEAT *at* THE TABLE**  
Track the Transition's meetings with outside groups, and leave comments and ideas for the team.

  
**OPEN FOR QUESTIONS  
RESPONSE**  
Incoming White House Press Secretary Robert Gibbs responds to your questions.

QUESTION FROM THE TRANSITION TEAM  
What social causes and service organizations are you a part of that make a difference in your community?  
Comments have closed. See the follow-up video here.

Comments (4,208)

Sort by: Date Rating Last Activity

Commenting Disabled  
Further commenting on this page has been disabled by the blog admin.

Josh · 120 weeks ago +34

Dear Obama Inauguration Team,

Not only has Pastor Rick Warren compared gay relationships to pedophilia and given vocal support to Uganda's criminalization of homosexuality, but he has also had abusive "ex-gay" programs at his church.

As I'm sure you know, the American Psychological Association and the American Psychiatric Association do NOT approve or advocate these "ex-gay" therapies as being gay/homosexual is not a disorder and is an innate orientation.

Therefore, "ex-gay" therapies can definitely be viewed as psychological abuse. There are many ex-"ex-gays" who return to identifying as gay because while they temporarily changed their behavior they never changed their innate feelings and situations.

They usually report that the "ex-gay" experience caused them severe emotional trauma and some even consider suicide over it. Some spend years and a lot of money in these abusive programs.

Despite what he sometimes says, he is against our very existence. He pretty much said as much when he was in Uganda.

Sincerely,  
-Josh

Report

Jannic\_68 · 121 weeks ago -17

Everyone in this household understands exactly why President-elect Obama chose Pastor Warren. It's time to stop slinging the venom and begin to reconcile. I'm ashamed of some of the comments coming from the far left. When did we all become so intolerant? When did we become so childish as a nation that we cannot abide being around people who disagree with our viewpoints. Sad...sad, indeed.

Report

Clive\_Thomas · 120 weeks ago +16

Judging from the comments already posted here on his own web blog, our President Elect's sincere desire to bring us together has started off with a colossal blunder. The Rev Warren choice has brought all anger, frustration, fear, and prejudice in our country right to the surface. That would be great if his objective was polarization, but it's no way to encourage flexibility. What scares me is if Mr. Obama (for whom I worked & voted) so seriously misjudged the effects on our country of this decision to honor Rev Warren, what other misjudgments is he capable of I mean, this really does scare me!!!

Report

Tom · 121 weeks ago +11

Rick Warren?...I hope this is not the 'Change We can Believe In' that Obama talked about?

Could he have found someone more divisive to give the invocation if he had tried??

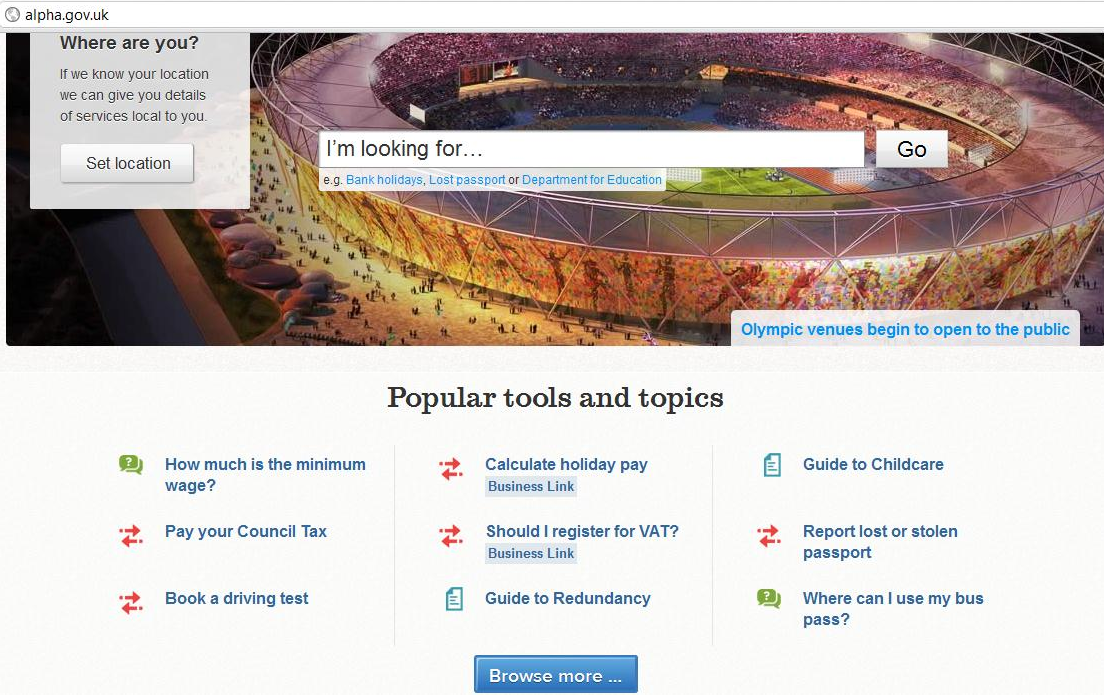
I voted for Obama but I'll be sitting out this inauguration. His choice of Rick Warren shows all too clearly that only certain types of people are welcome...and that is not me.

Report

Mark Heman · 120 weeks ago +14

Ακολουθεί η παράθεση του ηλεκτρονικού ιστότοπου: <http://www.alpha.gov.uk> που βρίσκεται σε πειραματική λειτουργία. Προοπτική του εν λόγω site είναι να φέρει «επανάσταση» στις παρεχόμενες υπηρεσίες ηλεκτρονικής διακυβέρνησης του Ηνωμένου Βασιλείου. Βασισμένο στη φιλοσοφία ότι: «Η κυβέρνηση πρέπει να εκμεταλλευτεί το πλεονέκτημα των πιο ανοιχτών, ευφυών και φθηνών ψηφιακών τεχνολογιών ώστε να παρέχει πιο απλές και αποτελεσματικές ψηφιακές υπηρεσίες στους χρήστες», ο συγκεκριμένος ιστότοπος πρωταρχικά αποσκοπεί στο να δοκιμάσει την λειτουργία μίας και μοναδικής κυβερνητικής πλατφόρμας που θα παρέχει συγκεντρωτικά τις διαθέσιμες ηλεκτρονικές υπηρεσίες στους χρήστες. Ο εν λόγω ιστότοπος δεν αποσκοπεί στο να αντικαταστήσει άμεσα τα υπάρχοντα κυβερνητικά sites, αλλά να συλλέξει τις απόψεις των χρηστών σχετικά με το συγκεκριμένο εγχείρημα. Με την δοκιμαστική αυτή λειτουργία, η νέα δομή και οι

υπηρεσίες που θα παρέχονται ηλεκτρονικά θα σχεδιαστούν εξ' ολοκλήρου πάνω στις ανάγκες των χρηστών για την μέγιστη ικανοποίησή τους. Το alpha.gov.uk δοκιμάζει ένα εύρος εργαλείων που στοχεύουν στην διεκπεραίωση των πιο κοινών απαιτήσεων των χρηστών και προσφέρει μία εικόνα του πώς θα πρέπει να παρουσιάζονται οι δραστηριότητες σε μία κυβερνητική πλατφόρμα ώστε να επιτευχθεί το βέλτιστο από πλευράς κόστους και πολυπλοκότητας αποτέλεσμα. Άλλωστε, η γνώση των πραγματικών αναγκών των χρηστών είναι υψίστης σημασίας για την επιτυχή μετάβαση σε έναν συγκεντρωτικό ιστότοπο παροχής ηλεκτρονικών υπηρεσιών και την αποφυγή άσκοπης σπατάλης χρήματος και χρόνου. Ακολουθούν κάποια χαρακτηριστικά στιγμιότυπα από το εν λόγω site:



alpha.gov.uk

Where are you?  
If we know your location we can give you details of services local to you.  
Set location

I'm looking for...  
Go  
e.g. Bank holidays, Lost passport or Department for Education

Olympic venues begin to open to the public

### Popular tools and topics

- How much is the minimum wage?
- Calculate holiday pay  
Business Link
- Guide to Childcare
- Pay your Council Tax
- Should I register for VAT?  
Business Link
- Report lost or stolen passport
- Book a driving test
- Guide to Redundancy
- Where can I use my bus pass?

Browse more ...

The screenshot displays the Alpha.gov.uk website interface, which is an experimental prototype for UK Government services. The main content area is divided into several sections:

- Ask a question:** A form with options to 'Ask a question', 'Share an idea', 'Report a problem', and 'Give praise'. A 'Continue' button is visible.
- FREQUENTLY ASKED QUESTIONS:** A list of questions with their respective authors, reply counts, and last reply dates. Examples include 'What backend technologies does alphagov use?' (12 replies) and 'Text too big?' (8 replies).
- IDEAS UNDER CONSIDERATION:** A list of ideas submitted by users, such as 'Spell checker on search box' (10 votes) and 'Make a Mobile Version' (7 votes).
- COMMON PROBLEMS:** A list of reported issues, including 'Set Location links do not work' (14 votes) and 'Avoid pointless junk logins and sign-ups' (10 votes).
- RECENT PRAISE:** A section for positive feedback, featuring a comment: 'Great project!' by Dickon, 3 hours ago.

On the right side of the page, there is a 'Message from alphagov prototype' explaining the site's purpose as an experimental prototype. Below this, there are 'Additional support links' for the website and blog, and a 'Join the community' section with checkboxes for various preferences. A 'Community members' section lists recent activity, such as 'ahmednuaman shared their first idea: Carousel for hero background'. At the bottom of the right sidebar, there is a 'Welcome to our new members' section with profile pictures and names of users who have recently joined, including Dickon, Jason Edwards, Andrew Cameron, and Desmondll.

Συνοψίζοντας, κρίνεται σκόπιμο να αναφερθούμε και στην αντίστοιχη χαρακτηριστική ελληνική πρωτοβουλία, το portal: <http://www.opengov.gr>. Η εν λόγω κυβερνητική πλατφόρμα αποτελεί σημαντικό και καιρίο βήμα προς την επιτυχία της ηλεκτρονικής διακυβέρνησης μέσω της ανάπτυξης άμεσης επικοινωνίας μεταξύ κυβέρνησης και πολιτών και την ενίσχυση της εμπιστοσύνης των τελευταίων λόγω της εδραίωσης περισσότερο διαφανών μεθόδων. Η ηλεκτρονική συμμετοχικότητα, περισσότερο απαραίτητη από ποτέ για την υγιή λειτουργία του ηλεκτρονικού μοντέλου διακυβέρνησης, αντικατοπτρίζεται άμεσα μέσω του συγκεκριμένου site, καθώς τίθενται σε διαβούλευση τα επικείμενα νομοσχέδια όλων των υπουργείων ώστε οι πολίτες να έχουν την δυνατότητα όχι μόνο να ενημερωθούν αλλά και να συμβάλλουν καθοριστικά στην λήψη των τελικών αποφάσεων. Η κάθε διαβούλευση είναι ανοιχτή και διαθέσιμη στο κοινό για ένα συγκεκριμένο χρονικό διάστημα και οι χρήστες είναι σε θέση να



καταθέσουν ελεύθερα την άποψή τους πάνω στο αντίστοιχο ζήτημα. Τα παρακάτω στατιστικά στοιχεία είναι ενδεικτικά του ενδιαφέροντος που έχουν επιδείξει έως τώρα οι πολίτες απέναντι στο εν λόγω εγχείρημα:

Τελευταία ενημέρωση, Τρίτη 04 Μαΐου 2011.

Απο τις 6/10/2009 έως την 23/2/2011 οι δικτυακοί τόποι του [OpenGov.gr](http://www.opengov.gr) έχουν δεχτεί συνολικά 4.511.097 επισκέψεις.

Υπουργείο	Αριθμός				
	Διαβουλεύσεων	Σχολίων	Προσκλήσεων	Θέσεων	Αιτήσεων
Διοικητικής Μεταρρύθμισης & Ηλεκτρονικής Διακυβέρνησης	0	0	0	0	0
Εσωτερικών	10	14.069	6	45	1.758
Οικονομικών	9	18.713	4	63	2.645
Εξωτερικών	0	0	0	0	0
Εθνικής Άμυνας	0	0	1	2	167
Ανάπτυξης, Ανταγωνιστικότητας & Ναυτιλίας *	24	4.138	18	218	3.607
Περιβάλλοντος, Ενέργειας και Κλιματικής Αλλαγής	29	7.760	11	160	4.517
Παιδείας, Δια Βίου Μάθησης και Θρησκευμάτων	19	9.577	9	538	9.313
Υποδομών, Μεταφορών και Δικτύων	14	2.154	7	65	1.540
Εργασίας και Κοινωνικής Ασφάλισης	7	1.790	6	70	1.843
Υγείας και Κοινωνικής Αλληλεγγύης	1	1.146	1	535	4.446
Αγροτικής Ανάπτυξης και Τροφίμων	8	3.053	5	20	1.508
Δικαιοσύνης	17	1.956	2	2	189
Προστασίας του Πολίτη	5	1.910	0	0	0
Πολιτισμού και Τουρισμού	2	580	23	88	3.047
Υπουργός Επικρατείας & Κυβερνητικός Εκπρόσωπος	2	182	1	1	36
Γενική Γραμματεία της Κυβέρνησης	1	535	1	1	40
OpenGov	5	366	1	1	213
<b>Σύνολα</b>	<b>153</b>	<b>67.929</b>	<b>96</b>	<b>1.809</b>	<b>34.889</b>

Τέλος, παραθέτουμε το παράδειγμα του ηλεκτρονικού ιστοτόπου: <http://www.labs.opengov.gr>, που αποτελεί εξίσου σημαντική πρωτοβουλία στα πλαίσια της ανοιχτής διακυβέρνησης. Άμεσος στόχος του παραπάνω διαδικτυακού τόπου είναι να φέρει στο προσκήνιο νέες και δημιουργικές ιδέες και τρόπους που θα επιφέρουν ουσιαστική καινοτομία στις σχέσεις του πολίτη με τον δημόσιο τομέα. Η ύπαρξη καινοτόμων ιδεών εκ μέρους των πολιτών θα πρέπει να συνοδεύεται από τον κατάλληλο τόπο, στον οποίο μπορούν να εκφραστούν και να ληφθούν περαιτέρω υπόψη. Το site εν προκειμένω αποτελεί κόμβο επικοινωνίας και βήμα διαλόγου ώστε να αποκτήσουν νέα διάσταση, συμμετοχική και αποκεντρωμένη, ο σχεδιασμός και η υλοποίηση των δημόσιων ηλεκτρονικών

υπηρεσιών. Σημαντικό δε είναι το γεγονός ότι μπορούν να υποβληθούν προτάσεις από πολίτες της χώρας που αποτελούν είτε ανεξάρτητα πρόσωπα είτε μέλη φορέων του ιδιωτικού και δημόσιου τομέα. Μερικά χαρακτηριστικά αποσπάσματα είναι τα παρακάτω:

#### Φάσεις labs.opengov

##### Προτάσεις Γ' Φάσης

- Ιδέες Άμεσης Εφαρμογής
- Τεχνικές Προτάσεις

##### Προτάσεις Β' Φάσης

##### Προτάσεις Α' Φάσης

#### Διάγραμμα Ροής / labs.opengov

#### #σχόλια στις προτάσεις (1,830)

» Εγώ πάλι πιστεύω πως το πρόβλημα είναι η υπερ-ρύθμιση που γίνεται για τα πάντα....

» αν και το λογισμικό έχει λάβει τον 1ο βραβείο καινοτομίας της MS και είναι πραγματικά απίστευτο σε αυτά που κάνει , όπως και η γενικότερη...

» πραγματικά μπράβο για την πρωτοβουλία!...

» χρειαζόμαστε πολλά wikis για να πάει μπροστά αυτός ο τόπος! wiki θα πρέπει να έχει κάθε υπουργείο και κάθε υπηρεσία καθώς πέραν του χρόν...

» Πολύ καλή ιδέα! εύκολη υλοποίηση θα σώσει πολλά λίτρα βενζίνης και πολλά λεπτά ομιλίας στα τηλέφωνα! Απορώ γιατί δν έχει γίνει ακόμη!...

» [http://www.elsekek.org/index.php?option=com\\_fireboard&Itemid=73&lang=el](http://www.elsekek.org/index.php?option=com_fireboard&Itemid=73&lang=el)  
Προσκαλώ τα μέλη της κοινότητας.opengov σε αυτή την ισ...

» Μέχρι σήμερα 5-8-2011 ισχύουν τα ίδια που ίσχυαν. Αυτά όλα τελικά μένουν στα λόγια, το τέρας καλά κρατεί!

---

## Η 4η Φάση των labs.OpenGov ολοκληρώθηκε!

### 1. Πρόγραμμα Ημερίδας & Παρουσιάσεις

### 2. Προτάσεις που προκρίθηκαν

Κάλεσμα για την 4η Φάση labs.OpenGov - Ηλεκτρονικές Υπηρεσίες

Καθημερινά διαπιστώνουμε ότι πολλές από τις ήδη υπάρχουσες ηλεκτρονικές υπηρεσίες του δημοσίου προς τους πολίτες, χρειάζονται σημαντικές βελτιώσεις ώστε να είναι χρήσιμες και φιλικές. Άλλες λειτουργούν μερικά και απαιτείται να αναβαθμιστούν ώστε να καλύπτουν τις σύγχρονες ανάγκες των πολιτών και των επιχειρήσεων.



Η βελτίωση αυτών των υπηρεσιών μπορεί να έχει άμεσα θετικά αποτελέσματα στην καλύτερη εξυπηρέτηση του πολίτη, στην εξοικονόμηση χρόνου και στην καταπολέμηση της γραφειοκρατίας. Για να επιτευχθεί αυτό είναι απαραίτητη η ενεργή συμμετοχή όλων των εμπλεκόμενων.

Η 4η φάση του labs.opengov, είναι αφιερωμένη στην υποβολή προτάσεων για την βελτίωση των ήδη παρεχόμενων υπηρεσιών με έμφαση στην Υγεία, στην Ασφάλιση και στις Οικονομικές συναλλαγές.

Καλούμε όλους τους ενδιαφερόμενους

α. **Φορείς**, να παρουσιάσουν τον οδικό χάρτη για την βελτίωση των υφιστάμενων υπηρεσιών που παρέχουν σε πολίτες και επιχειρήσεις

β. **Πολίτες** να υποβάλουν τις προτάσεις τους με στόχο την βελτίωση και αναβάθμιση ηλεκτρονικών υπηρεσιών που χρησιμοποιούν

γ. **Στελέχη & Επαγγελματίες** στον κλάδο των ΤΠΕ να υποβάλουν προτάσεις για την πιο αποτελεσματική αξιοποίηση υπαρχόντων συστημάτων λαμβάνοντας υπόψη και αντίστοιχες διευθετήσεις πρακτικές.

Οι προτάσεις μπορούν να είναι ομαδικές αλλά και ατομικές.

Η υποβολή προτάσεων θα ολοκληρωθεί στις **3 / 4 / 2011**. Θα επιλεγούν από την επιστημονική επιτροπή οι καλύτερες, λαμβάνοντας υπόψη και την αξιολόγηση των εγγεγραμμένων χρηστών των labs και θα παρουσιαστούν σε ανοικτή εκδήλωση. Στην επιλογή των προτάσεων η επιστημονική επιτροπή θα συνεκτιμήσει στοιχεία όπως η ευστοχία στην επιλογή μιας υφιστάμενης υπηρεσίας, η ωριμότητα της πρότασης από πλευράς υλοποίησης και η καλύτερη εξυπηρέτηση του πολίτη.

Στις προτάσεις σας μπορούν να αξιοποιηθούν οι παρακάτω υπηρεσίες:

<https://www.gsis.gr/>

<http://www.ika.gr/>

<http://www.oaed.gr/>

<http://www.lidika.gr/>

<http://www.noleadonia.gov.gr/>



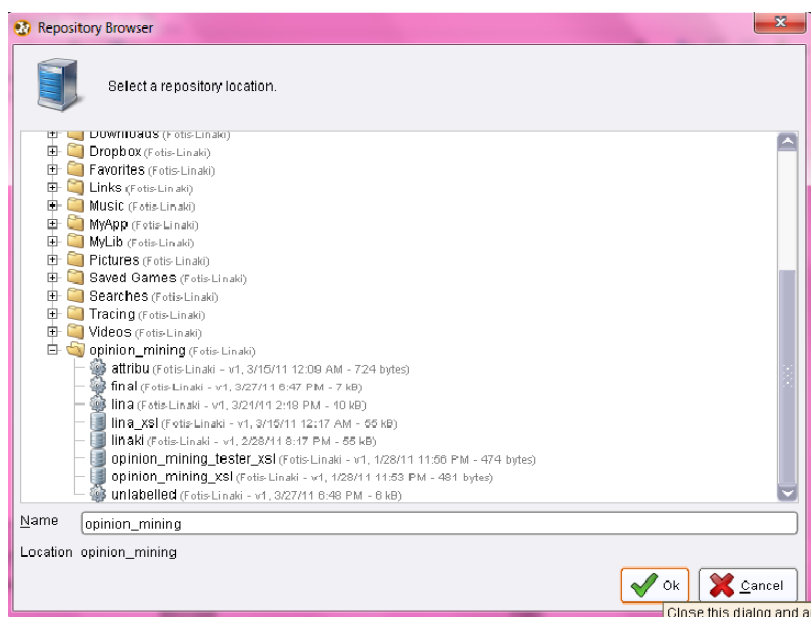
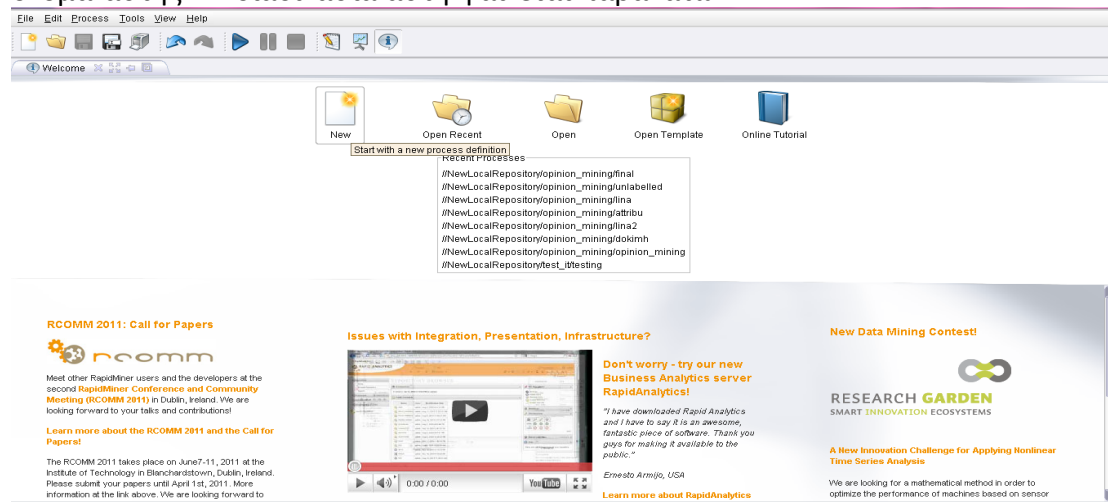
## ***Κεφάλαιο 4. Εφαρμογή Opinion Mining***

---

## 4.1 Εισαγωγή στο RapidMiner

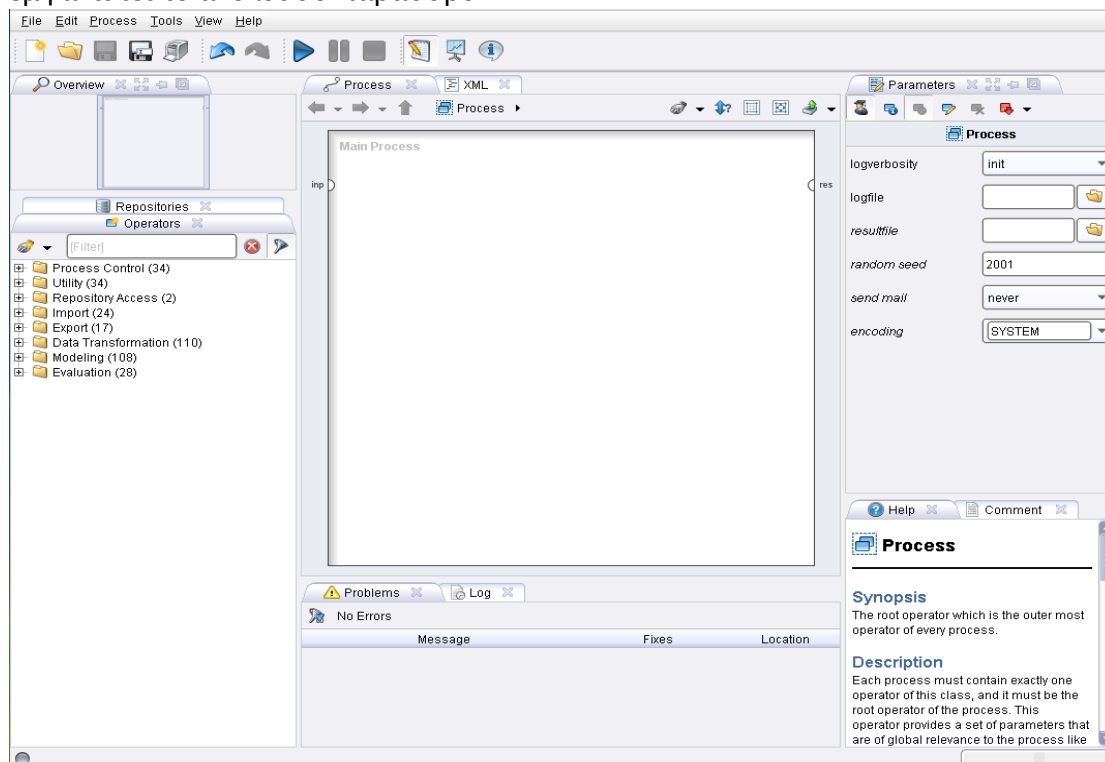
Για την αξιολόγηση των σχολίων που συγκεντρώθηκαν από διάφορες διαδικτυακές πηγές, με έμφαση κυρίως σε απόψεις που συγκεντρώθηκαν μέσω blogs και forums, και την κατηγοριοποίησή τους σε θετικές και αρνητικές χρησιμοποιήθηκε το λογισμικό RapidMiner. Οι κύριοι λόγοι για την επιλογή του συγκεκριμένου λογισμικού υπήρξαν η ελεύθερη χρήση του, καθώς διατίθεται μέσω της επίσημης ιστοσελίδας του δωρεάν, σε συνδυασμό με το γεγονός ότι είναι ένα λογισμικό αναγνωρισμένο, που χρησιμοποιείται ευρέως σε παγκόσμια κλίμακα και χαίρει υψηλής αξιολόγησης από τους χρήστες του στο πεδίο του data mining.

Αφού κατεβάσουμε την κατάλληλη για τον υπολογιστή μας έκδοση από την επίσημη ιστοσελίδα κι εγκαταστήσουμε το RapidMiner σύμφωνα με τις οδηγίες, δημιουργούμε μία νέα διεργασία ανάλυσης (analysis process) και την αποθηκεύουμε για μελλοντική χρήση. Ανοίγοντας το λογισμικό και πατώντας το εικονίδιο «New» εμφανίζεται το παράθυρο «RepositoryBrowser» από όπου επιλέγουμε την τοποθεσία που θα αποθηκεύσουμε τη διεργασία μας καθώς και το όνομα αυτής. Η διαδικασία αυτή φαίνεται παρακάτω :



Στο μέλλον μπορούμε να ανατρέξουμε στη διεργασία αυτή για οποιαδήποτε επεξεργασία μέσω του εικονιδίου «Open» και αναζητώντας την με το όνομά της στην τοποθεσία που επιλέξαμε να την αποθηκεύσουμε. Προχωρούμε στην περαιτέρω σχεδίαση της διεργασίας μας.

Ανοίγοντας την κύρια περιοχή σχεδίασης («Design Perspective») του RapidMiner θα εμφανιστεί το ακόλουθο παράθυρο:



Στα αριστερά της οθόνης μπορούμε να διακρίνουμε τους διαθέσιμους Operators του λογισμικού, οι οποίοι παρουσιάζονται σε ομάδες, και είναι αυτοί που μπορούν να συμπεριληφθούν στην εκάστοτε διεργασία μας. Ουσιαστικά, η κάθε διεργασία στο RapidMiner είναι μία κατάλληλη αλληλουχία των διαθέσιμων Operators, όπως θα φανεί και στη συνέχεια. Το λογισμικό μας διαθέτει και ένα πλήθος άλλων Operators τους οποίους μπορούμε να εγκαταστήσουμε από το menu «Help» επιλέγοντας «UpdateRapidMiner». Για εφαρμογές κατηγοριοποίησης κειμένου (Document classification/categorization) απαραίτητο είναι να εγκαταστήσουμε τη λίστα: «TextProcessing» που περιέχει ένα σύνολο Operators κατάλληλο για επεξεργασία των επιμέρους στοιχείων ενός κειμένου :



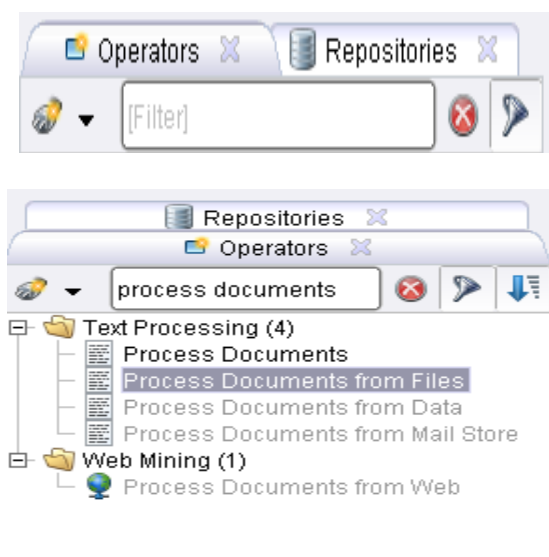
#### 4.1.1 Σχεδίαση 1<sup>ης</sup> διεργασίας

Είμαστε πλέον έτοιμοι να επιλέξουμε, να συνδέσουμε με την ορθή σειρά και να εισάγουμε τα αντίστοιχα δεδομένα στο σύνολο των Operators που κρίνονται απαραίτητοι για την υλοποίηση της εφαρμογής κατηγοριοποίησης κειμένου που χρησιμοποιήθηκε στην παρούσα εργασία. Σημειώνεται εδώ ότι συνολικά χρειάστηκε η δημιουργία δύο ξεχωριστών διεργασιών. Στην πρώτη δημιουργούμε, εκπαιδεύουμε και αποθηκεύουμε το μοντέλο αυτόματης κατηγοριοποίησης και στη δεύτερη εφαρμόζουμε το μοντέλο αυτό σε νέα δεδομένα ώστε να ταξινομηθούν αυτόματα σε θετικά και αρνητικά.

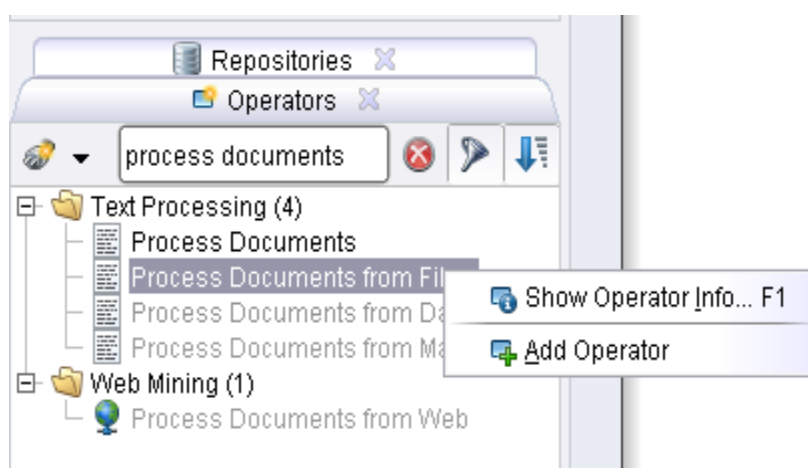
Ξεκινάμε με την υλοποίηση της πρώτης μας διεργασίας. Ο αρχικός Operator που διαλέγεται ονομάζεται «Process Documents From Files», και ανήκει στην ομάδα των «TextProcessing Operators».



Η διαδικασία εύρεσης του κατάλληλου Operator από τη λίστα απλοποιείται μέσω της επιλογής αναζήτησης «Filter» που είναι διαθέσιμη, ώστε πληκτρολογώντας τον τίτλο του Operator που αναζητάμε ή μέρος αυτού εμφανίζονται τα αντίστοιχα αποτελέσματα όπως φαίνεται παρακάτω:

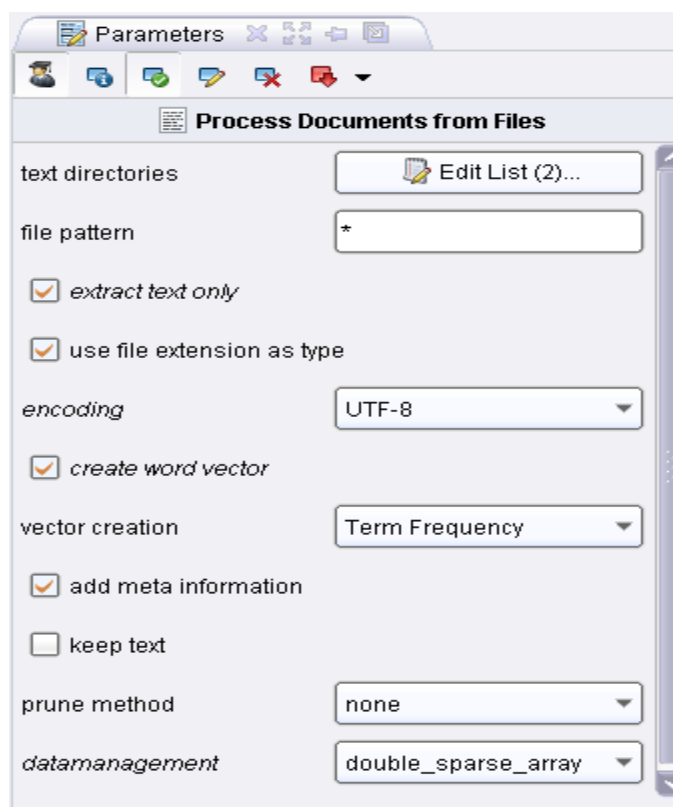


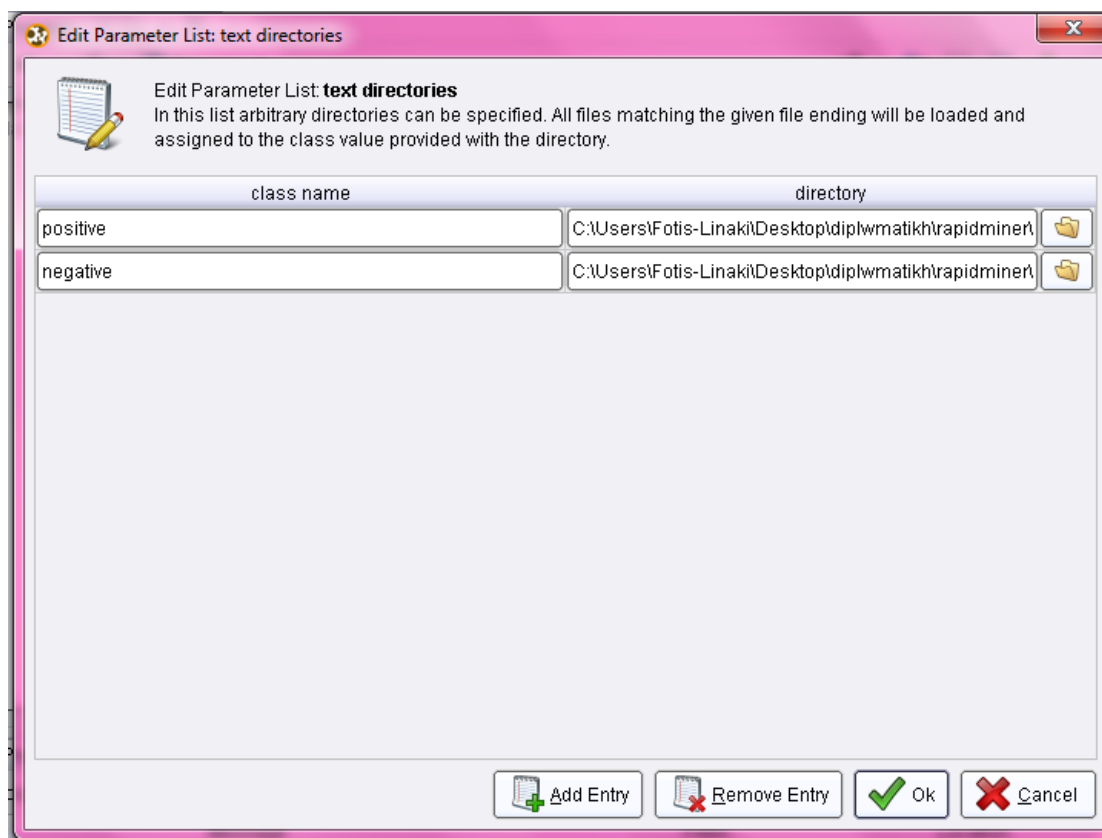
Με δεξί κλικ στον Operator μπορούμε να τον προσθέσουμε στη διεργασία μας ή να δούμε περισσότερες λεπτομέρειες για την χρήση του ώστε να είμαστε σε θέση να συμπληρώσουμε όσο το δυνατόν ορθότερα τα αντίστοιχα πεδία.



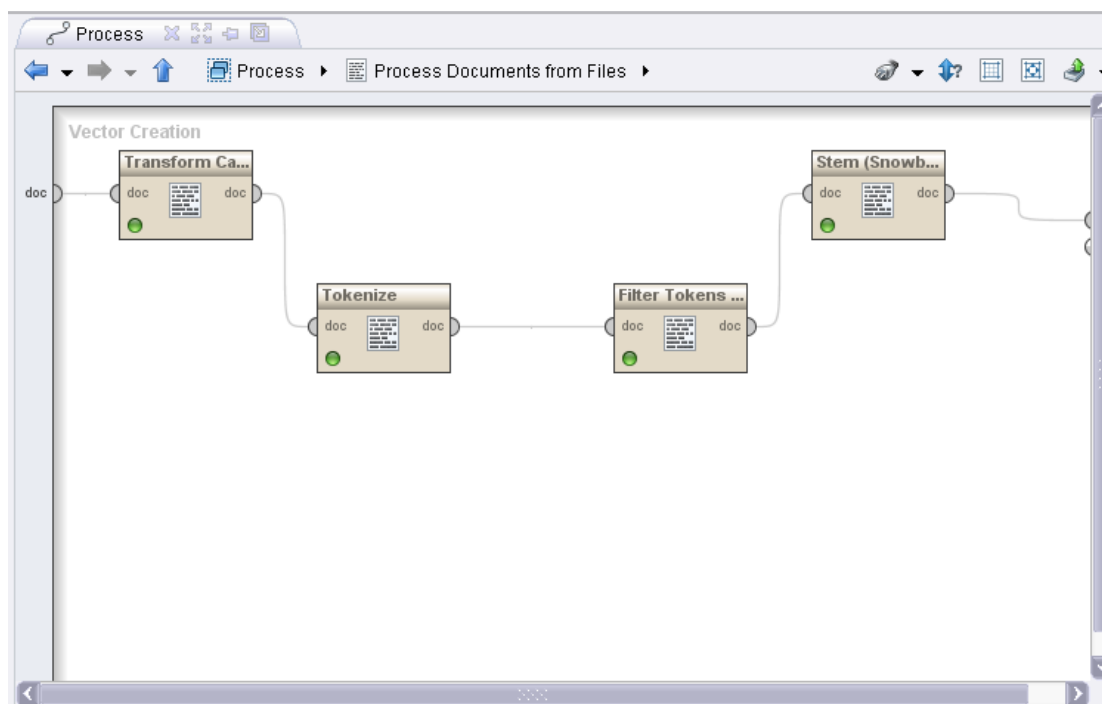
Αφού λοιπόν προσθέσουμε τον συγκεκριμένο Operator στη διεργασία μας, στο πεδίο «text directories» που βρίσκεται στα δεξιά του περιβάλλοντος εργασίας τοποθετούμε τα σχόλια που έχουμε συλλέξει από τις διάφορες διαδικτυακές πηγές. Τα σχόλια τα έχουμε κατηγοριοποιήσει προηγουμένως και χειροκίνητα σε θετικά και αρνητικά και θα αποτελέσουν την βάση για την εκπαίδευση του μοντέλου που θα χρησιμοποιηθεί στη συνέχεια για αυτόματη κατηγοριοποίηση. Για το λόγο αυτό, τα σχόλια χωρίζονται σε δύο διαφορετικά αρχεία προσθέτοντας το κατάλληλο label (ετικέτα, που υποδηλώνει την αντίστοιχη κατηγορία) σε καθένα από αυτά, δηλαδή positive και negative αντίστοιχα. Αξίζει να αναφερθεί πως για ορθότερα και πληρέστερα αποτελέσματα εκτός των σχολίων καθ' αυτών προστέθηκαν και δύο επιπλέον αρχεία. Τα αρχεία αυτά περιέχουν το καθένα το πλέον χαρακτηριστικό λεξιλόγιο σχολιασμού, θετικού κι αρνητικού αντίστοιχα, όπως αυτό καταγράφηκε μέσω της συλλογής των δεδομένων μας. Όπως είναι φυσικό, επειδή πρόκειται για σχολιασμό ενός συγκεκριμένου ζητήματος, αυτού της ηλεκτρονικής διακυβέρνησης, το λεξιλόγιο επιλέχθηκε λαμβάνοντας υπ' όψιν μόνο λέξεις που μπορούν να

χρησιμοποιηθούν για τον χαρακτηρισμό ενός τέτοιου θέματος και όχι λέξεις που εκφράζουν γενικά θετικό ή αρνητικό συναίσθημα αλλά δε θα ήταν πιθανόν να χρησιμοποιηθούν γι' αυτό το σκοπό. Από τις λέξεις που εντοπίσαμε συχνότερα και με τη βοήθεια λεξικού, καταγράψαμε όλες τις συνώνυμες που θα μπορούσαν να χρησιμοποιηθούν για τον χαρακτηρισμό της ηλεκτρονικής διακυβέρνησης ως θετικής ή αρνητικής. Ο φάκελος Positive λοιπόν αποτελείται από δύο αρχεία, ένα με τα σχόλια καθ' αυτά κι ένα που περιέχει το κατάλληλο λεξιλόγιο που παραπέμπει σε θετικό σχολιασμό. Τα ίδια ισχύουν και για το φάκελο Negative. Οι παράμετροι προς συμπλήρωση του εν λόγω Operator καθώς και η προσθήκη των θετικών και αρνητικών σχολίων στο πεδίο text directories με τα αντίστοιχα labels φαίνονται παρακάτω:





Σημασία δίνεται ώστε να επιλεγεί η συμβατή με τα αρχεία μας κωδικοποίηση από τη λίστα encoding όπως επίσης και στην επιλογή του πεδίου «vector creation», όπου εδώ επιλέγουμε από τη διαθέσιμη λίστα την επιλογή «term frequency», ώστε να μετρηθεί η συχνότητα εμφάνισης της κάθε λεκτικής οντότητας. Συνεχίζοντας, με διπλό κλικ στον Operator που έχουμε ήδη προσθέσει δημιουργούμε μία υποδιεργασία (subprocess) εντός αυτού. Εδώ θα επιλεγούν και θα συνδεθούν σε σειρά οι κατάλληλοι Operators από την ομάδα «TextProcessing», οι οποίοι και θα βοηθήσουν στην περαιτέρω επεξεργασία των δεδομένων που έχουμε εισάγει. Δυστυχώς, κάποιοι χρήσιμοι Operators δεν είναι προσαρμοσμένοι ώστε να επεξεργάζονται κατάλληλα την ελληνική γλώσσα παρά μόνο την αγγλική κι ένα μικρό σύνολο κάποιων άλλων γλωσσών ευρέως χρησιμοποιούμενων, γεγονός που δυσκόλεψε την εργασία μας. Η υποδιεργασία που αναπτύχθηκε φαίνεται παρακάτω ενώ στην συνέχεια εξηγείται η χρησιμότητα κάθε Operator ξεχωριστά:



«Transform Cases»: Μετατρέπει όλους τους χαρακτήρες ενός κειμένου ή σε κεφαλαία είτε σε μικρά γράμματα αντίστοιχα.

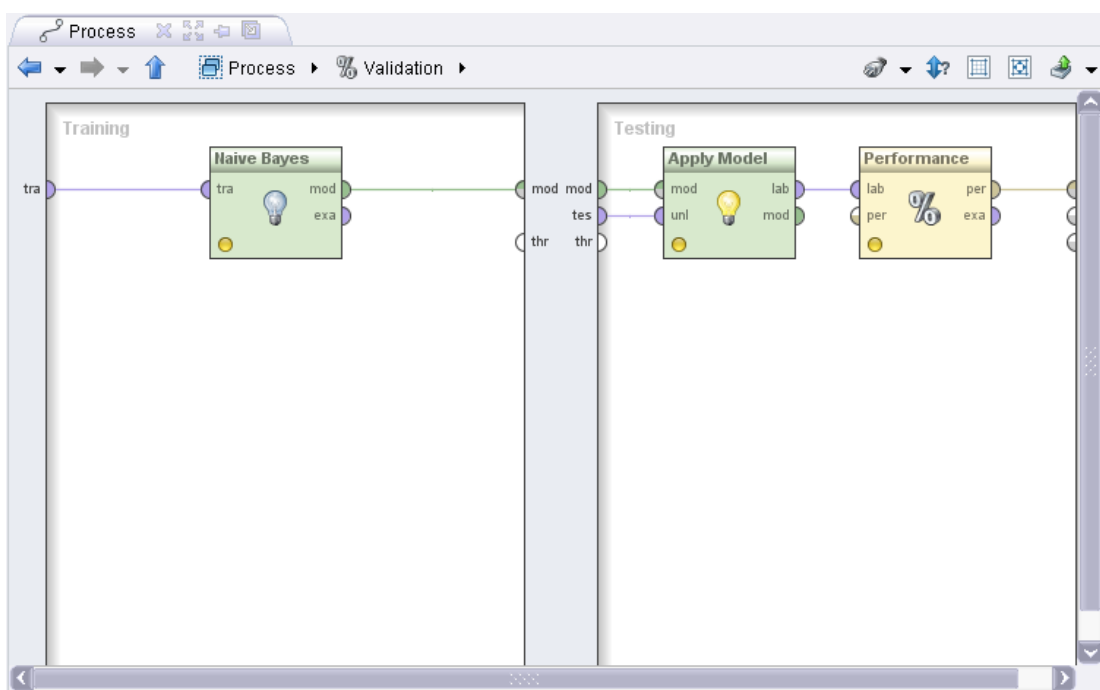
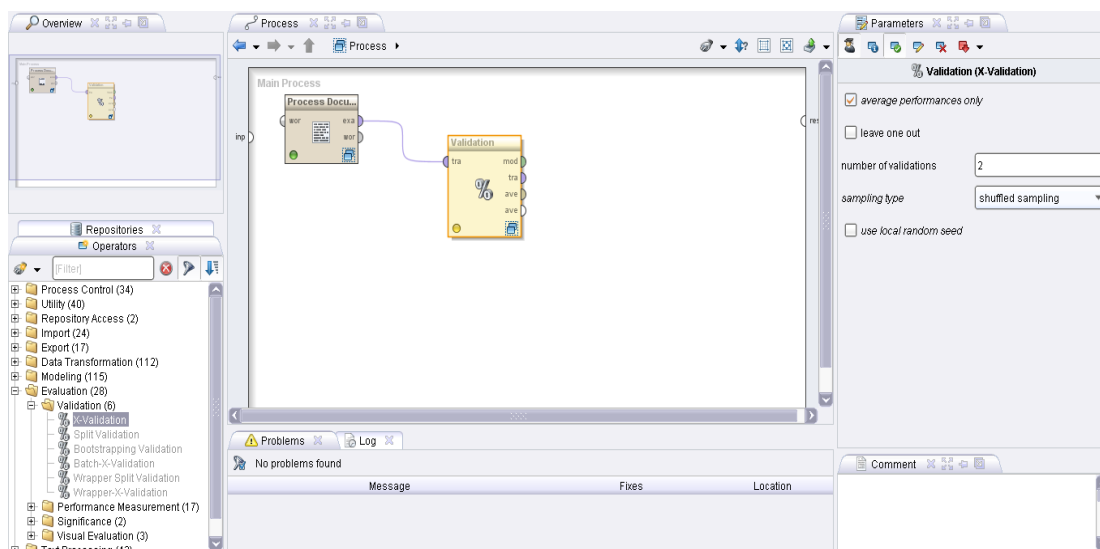
«Tokenize»: Πατώντας στο συγκεκριμένο Operator κι επιλέγοντας «non letters» χωρίζει το κείμενο προς επεξεργασία σε λεκτικές μονάδες.

«Filter Tokenize (by length)»: Αφαιρεί τις λέξεις που περιέχουν λιγότερους ή περισσότερους χαρακτήρες από ένα συγκεκριμένο όριο. Κατώτερο όριο χαρακτήρων επιλέχθηκε το τρία, δηλαδή αφαιρούνται οι λέξεις που περιέχουν ένα ή δύο χαρακτήρες, και ανώτερο όριο το 25.

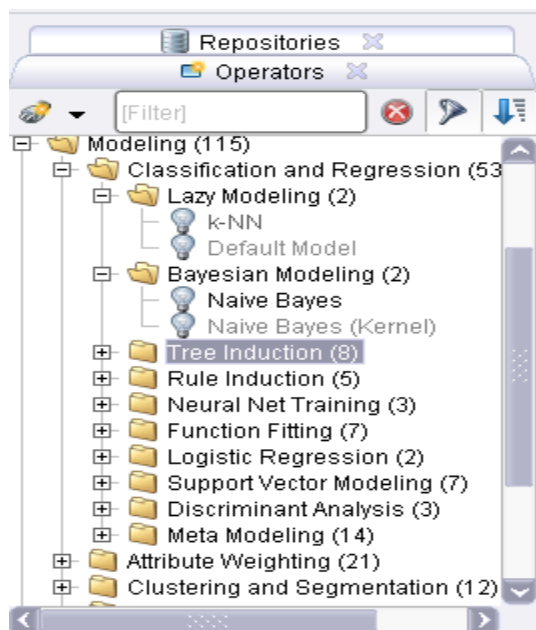
«Stem (Snowball)»: Ο συγκεκριμένος Operator είναι ιδιαίτερα σημαντικός και χρήσιμος καθώς εντοπίζει και αφαιρεί από κάθε λέξη την κατάληξή της (επίθεμα) αφήνοντας μόνο το κύριο στέλεχος της. Το RapidMiner δε διαθέτει ελληνικό στελεχωτή κι έτσι επιλέχθηκε ο συγκεκριμένος που είναι πλησιέστερος στο ζητούμενο, αλλά κυρίως η προσθήκη του στη διεργασία γίνεται εδώ για λόγους πληρότητας.

Γυρνάμε στην κύρια διεργασία μας και προσθέτουμε τον επόμενο Operator, ο οποίος ανήκει στην ομάδα «Evaluation» και ονομάζεται «X-Validation». Ο συγκεκριμένος Operator εκτιμά την απόδοση του μοντέλου που θα εκπαιδεύσουμε και θα εφαρμόσουμε στη συνέχεια στα μη κατηγοριοποιημένα δεδομένα μας. Ο «X-Validation» περιέχει εσωτερικά μία υποδιεργασία, η οποία και επιστρέφει το ζητούμενο αυτό μοντέλο το οποίο εκπαιδεύεται μέσω των δεδομένων που εισαγάγαμε αρχικά στη διεργασία μας, εν προκειμένω τα αρχεία που περιέχουν τις θετικές και αρνητικές απόψεις όπως αυτές έχουν συγκεντρωθεί. Σημειώνεται ότι τα αποτελέσματα που επιστρέφονται αποτελούν απλά μία εκτίμηση της απόδοσης του μοντέλου και δεν πρέπει να ταυτίζονται με τα ακριβή και πραγματικά. Η κύρια διεργασία έως εδώ καθώς και η υποδιεργασία από την οποία αποτελείται ο «X-Validation» φαίνονται στο παρακάτω σχήμα:



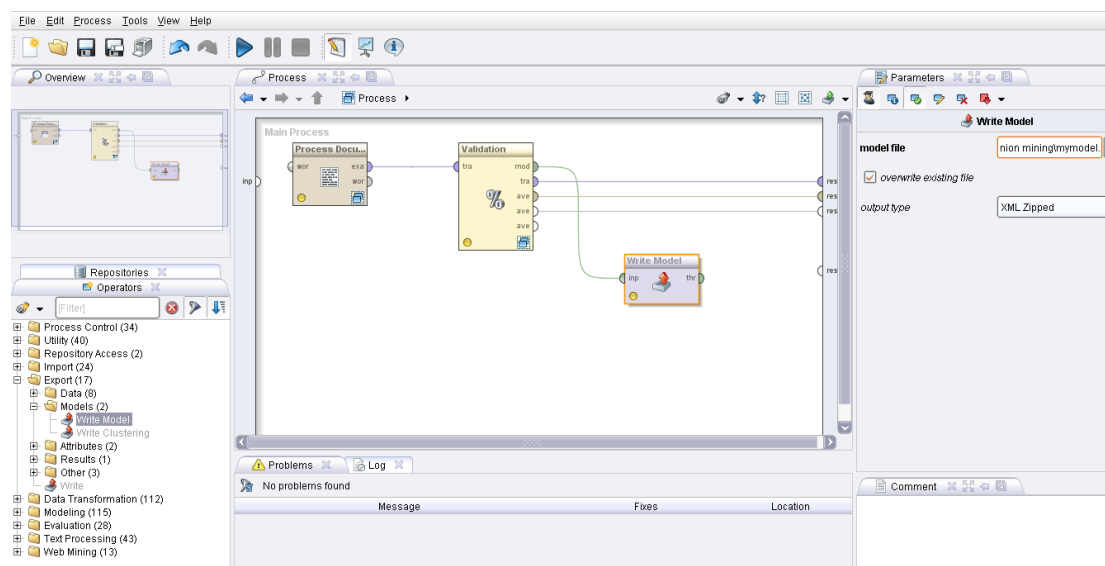


Σχετικά με την υποδιεργασία, όπως φαίνεται και από το παραπάνω σχήμα, αποτελείται από δύο φάσεις, τη φάση της εκπαίδευσης του μοντέλου και τη φάση της εφαρμογής αυτού στο σύνολο των δεδομένων. Αναλυτικότερα, το RapidMiner διαθέτει ένα πλήθος έτοιμων Operators, εξειδικευμένων στην κατηγοριοποίηση δεδομένων, όπως φαίνεται κάτωθι:



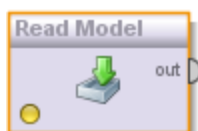
Από τους διαθέσιμους Operators, έπειτα από δοκιμές, εκλέχθηκε τελικά ο αλγόριθμος «Naïve Bayes», ο οποίος είχε την μεγαλύτερη απόδοση, δηλαδή σωστότερη πρόβλεψη σε σύγκριση με τους υπολοίπους. Ο «Naïve Bayes» αποτελεί ένα μοντέλο που βασίζεται στην αποτίμηση πιθανότητας ώστε να παράγει την ορθότερη πρόβλεψη και συνιστά μία απλή και ιδιαίτερως δημοφιλή μέθοδο μηχανικής μάθησης που δύναται να εφαρμοστεί και στο πεδίο κατηγοριοποίησης φυσικής γλώσσας. Με τον επόμενο Operator, τον «Apply Model» από το group «Modeling» - «Model Application», ουσιαστικά εφαρμόζουμε το μοντέλο που δημιουργήθηκε μέσω του «Naïve Bayes» στο σύνολο των δεδομένων του παραδείγματός μας. Ο Naïve Bayes Operator είναι αυτός που περιέχει τις απαραίτητες πληροφορίες και τα δεδομένα πάνω και στα οποία εκπαιδεύτηκε και θα χρησιμοποιηθεί αργότερα για την πρόβλεψη κατηγοριοποίησης της νέας πληροφορίας που θα εισαχθεί σε θετική ή αρνητική. Ο τελευταίος Operator αυτής της υποδιεργασίας είναι ο Operator «Performance» που βρίσκεται αντίστοιχα στην ομάδα: «Evaluation» - «Performance Measurement» και υπολογίζει μία πρώτη εκτίμηση της ακρίβειας της πρόβλεψής μας. Σημειώνεται εδώ ότι η εκτίμηση είναι καθαρά θεωρητική και η πραγματική απόδοση του μοντέλου δύναται να απέχει από αυτή.

Γυρνώντας πίσω στην κύρια διεργασία μας, προσθέτουμε τον τελευταίο Operator, ο οποίος ονομάζεται «Write Model» και ανήκει στο group «Export» - «Models». Με τον συγκεκριμένο Operator αποθηκεύουμε σε φάκελο της επιλογής μας το μοντέλο που δημιουργήσαμε ώστε μελλοντικά και εδώ συγκεκριμένα στην επόμενη διεργασία μας να είμαστε σε θέση να το εφαρμόσουμε σε νέα δεδομένα. Η πρώτη αυτή διεργασία μας λοιπόν έφτασε στο τέλος με τη δημιουργία, εκπαίδευση και τελικά την αποθήκευση του κατηγοριοποιητή μας. Η ολοκληρωμένη αυτή διεργασία φαίνεται στο σχήμα που ακολουθεί:



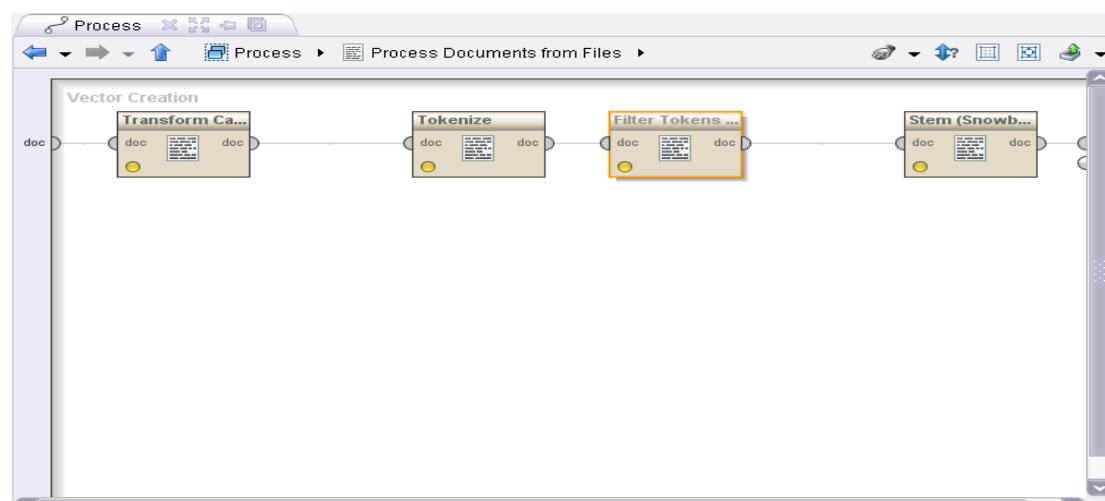
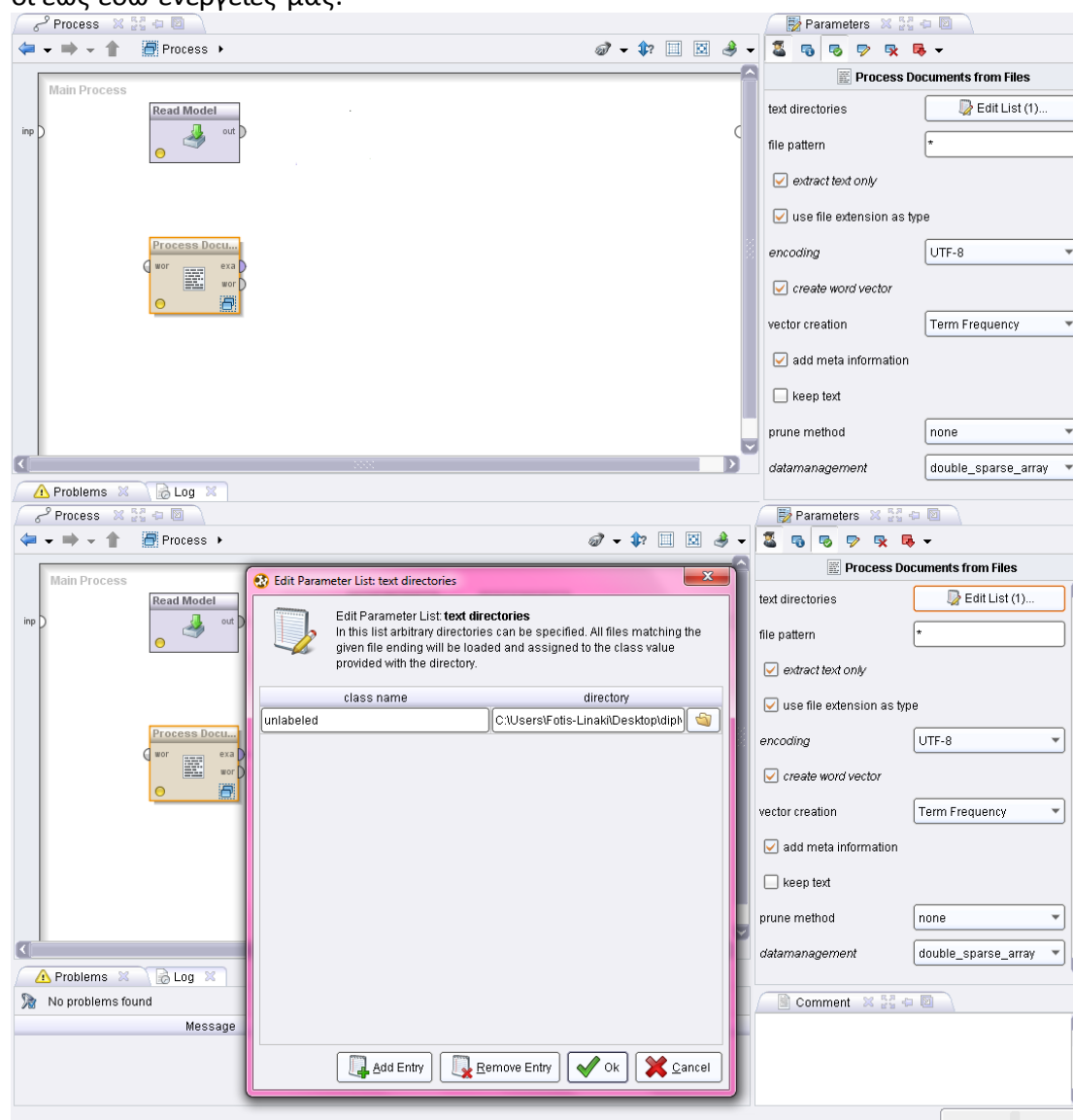
## 4.1.2 Σχεδίαση 2<sup>ης</sup> διεργασίας

Ακολουθεί η δημιουργία μέσω του ίδιου λογισμικού της δεύτερης διεργασίας. Στη διεργασία αυτή θα εφαρμόσουμε το μοντέλο - κατηγοριοποιητή που δημιουργήθηκε και αποθηκεύτηκε προηγουμένως σε νέα δεδομένα, ώστε αυτά να κατηγοριοποιηθούν αυτόματα σε θετικά και αρνητικά. Ξεκινάμε να εισάγουμε τους Operators της διεργασίας αυτής και να εξηγήσουμε τη λειτουργία και τη χρησιμότητά τους αναλυτικά. Στην κύρια διεργασία, δηλαδή στο χώρο main process του λογισμικού μας, εισάγουμε τον Operator: «Read Model» που ανήκει στην κατηγορία «Import» - «Models».

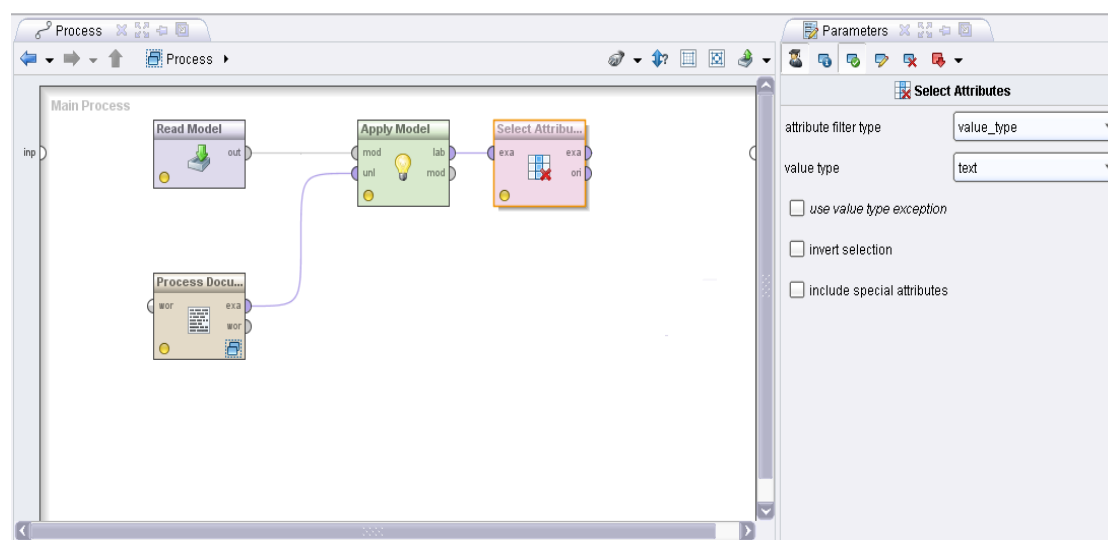


Ο εν λόγω Operator δέχεται σαν παράμετρο το μοντέλο το οποίο θα εφαρμοστεί στα δεδομένα που θα εισαχθούν μετέπειτα και αρκεί να ορίσουμε στο πεδίο «Model file» την τοποθεσία στην οποία αυτό έχει αποθηκευτεί από την προηγούμενη διεργασία μας. Συνεχίζοντας εισάγουμε τον Operator «Process documents from files», ο οποίος είχε χρησιμοποιηθεί και στην προηγούμενη διεργασία μας. Η διαφορά έγκειται στο γεγονός ότι στο πεδίο: «Text directories», που εμφανίζεται στα δεξιά του παραθύρου εργασίας, εισάγουμε τα σχόλια προς κατηγοριοποίηση, ήτοι τα σχόλια στα οποία θέλουμε να εφαρμοστεί το μοντέλο-κατηγοριοποιητής και να γίνει η αντίστοιχη πρόβλεψη, και όχι τα σχόλια κατηγοριοποιημένα όπως έγινε στη φάση της εκπαίδευσης του μοντέλου. Τα υπόλοιπα πεδία του Operator μένουν ως είχαν και στην προηγούμενη διεργασία που αυτός χρησιμοποιήθηκε. Απαραίτητο ώστε να μπορεί ο κατηγοριοποιητής να εφαρμοστεί σωστά στα καινούρια δεδομένα είναι η δημιουργία της ίδιας ακριβώς υποδιεργασίας εσωτερικά του Operator: «Process documents from files» με αυτή που δημιουργήθηκε στη φάση εκπαίδευσης του μοντέλου αρχικά. Παρακάτω

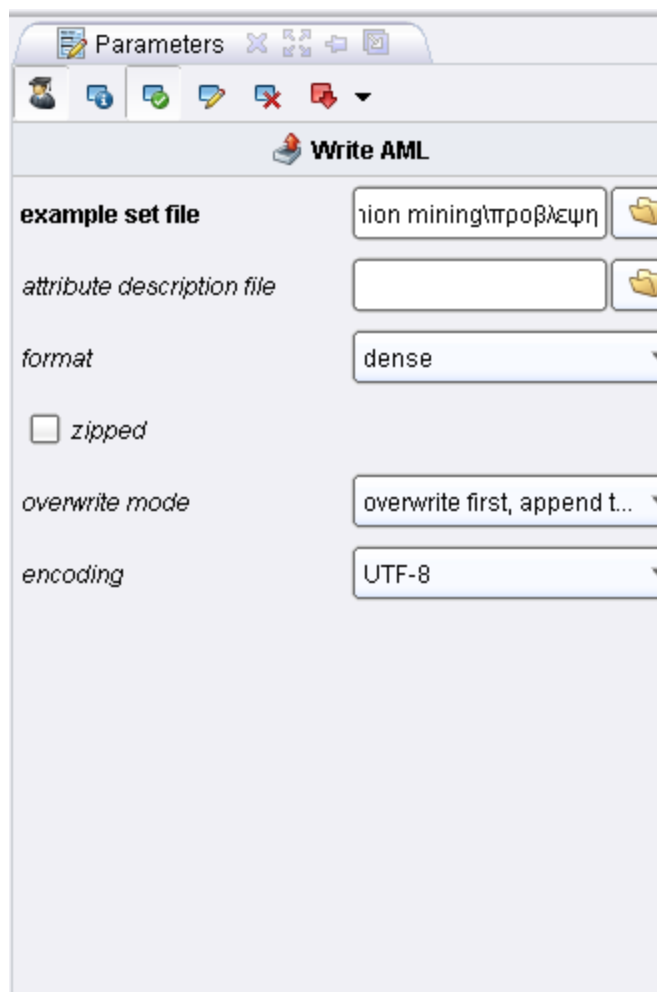
δίνονται κάποια χαρακτηριστικά screenshots ώστε να γίνουν σαφέστερα αντιληπτές οι έως εδώ ενέργειες μας:



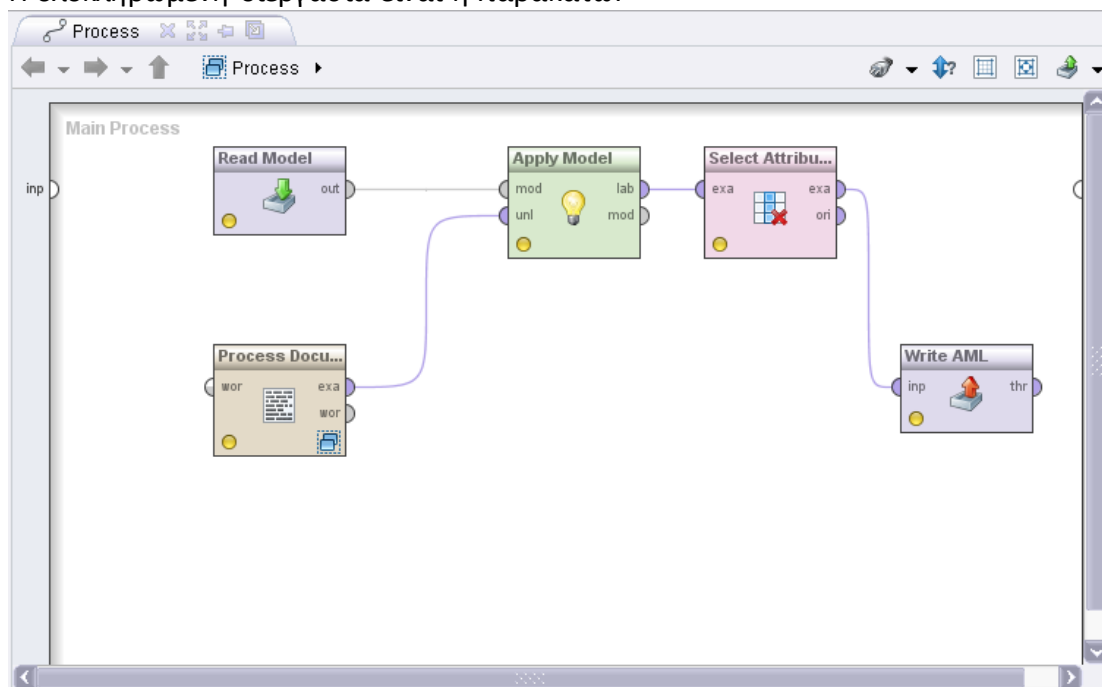
Για τη δημιουργία της συγκεκριμένης υποδιεργασίας και τη λειτουργία του καθενός από τους παραπάνω Operators μιλήσαμε αναλυτικά την προηγούμενη φορά που χρησιμοποιήθηκαν γι' αυτό κι εδώ παραλείπεται περαιτέρω αναφορά σε αυτούς. Ο αναγνώστης εάν επιθυμεί μπορεί να ανατρέξει και να θυμηθεί κάποιες λεπτομέρειες για τη χρησιμότητα των Operators εν λόγω. Επιστρέφουμε στην κύρια διεργασία μας για να προσθέσουμε και να συνδέσουμε κατάλληλα τους τελευταίους Operators που θα χρειαστούν για την ολοκλήρωσή της. Ο επόμενος Operator ονομάζεται «Apply Model», ανήκει στην κατηγορία: «Modeling» - «Model Application» και θα συνδεθεί με τους δύο προηγούμενους Operators της κύριας διεργασίας όπως θα φανεί σε παρακάτω σχήμα. Αναφέρουμε ότι για το συγκεκριμένο Operator δεν χρειάζεται να προσθέσουμε κάτι παραπάνω ή να τροποποιήσουμε κάποιο από τα πεδία του, καθώς απλά εφαρμόζει το μοντέλο που επιλέξαμε στα προς κατηγοριοποίηση δεδομένα. Στη συνέχεια προστίθεται από την ομάδα «Data Transformation» - «Attribute Set Reduction and Transformation» - «Selection» ο Operator: «Select Attributes». Ο εν λόγω Operator θα επιλέξει τα attributes που θα πρέπει να διατηρηθούν και αυτά που θα πρέπει να αφαιρεθούν αντίστοιχα από το παράδειγμα προς κατηγοριοποίηση που εισάγαμε ώστε να γίνει η ζητούμενη πρόβλεψη. Πιο συγκεκριμένα, εδώ ως «attribute filter type» επιλέγουμε «value type», το οποίο στη συνέχεια ορίζουμε ως «text» καθώς έχουμε να κάνουμε με αρχεία κειμένου. Στο σημείο αυτό παραθέτουμε το αντίστοιχο screenshot ώστε να υπάρξει μια πιο σαφής εικόνα όσων περιγράφηκαν μέχρι στιγμής:



Φτάσαμε πλέον στο σημείο όπου προσθέτουμε τον τελευταίο Operator της δεύτερης διεργασίας, εν προκειμένω από το group «Export» - «Data», επιλέγουμε τον Operator «Write AML». Στο πεδίο «example set file», επιλέγουμε την τοποθεσία στον υπολογιστή μας όπου επιθυμούμε να αποθηκευτούν τα αποτελέσματα της διεργασίας μας, δηλαδή η πρόβλεψη κατηγοριοποίησης, και δίνουμε ένα όνομα της επιλογής μας στο αρχείο που θα περιέχει την πρόβλεψη αυτή. Συμπληρώνουμε τα υπόλοιπα πεδία του Operator όπως απεικονίζεται στο παρακάτω σχήμα:



Η ολοκληρωμένη διεργασία είναι η παρακάτω:



## 4.2 Εκτέλεση

Πλέον έχουμε ολοκληρώσει την κατασκευή των δύο διεργασιών μέσω του RapidMiner ώστε να προβλέψουμε το συναίσθημα θετικό ή αρνητικό που εκφράζεται σε σχόλια που επιθυμούμε να ταξινομήσουμε αυτόματα. Αποθηκεύουμε τις διεργασίες αυτές ώστε να μπορούν να χρησιμοποιηθούν μελλοντικά οποιαδήποτε στιγμή το επιθυμούμε. Η διαδικασία που χρειάζεται να ακολουθήσουμε κάθε φορά είναι η εξής: ανοίγουμε τη δεύτερη διεργασία, εισάγουμε μέσω του Operator: «Process documents from files» ένα κείμενο της επιλογής μας στο πεδίο «Text Directories», όπως περιγράφηκε κι αναλυτικότερα προηγουμένως, και εκτελούμε τη διεργασία πατώντας στο κουμπί «Play»



Το αρχείο-πρόβλεψη του συναισθήματος είναι πλέον διαθέσιμο και βρίσκεται στην τοποθεσία όπου επιλέξαμε να αποθηκευτεί μέσω του Operator: «Write AML» και έχει το όνομα που ορίσαμε. Το αρχείο αυτό περιέχει τον χαρακτηρισμό positive ή negative με το αντίστοιχο ποσοστό ώστε να έχουμε και ένα μέτρο της πόλωσης της κάθε άποψης και να μπορούμε να αποφανθούμε και για σχόλια που τείνουν προς το ουδέτερο.

## 4.3 Κώδικας

Οι διεργασίες του RapidMiner σε μορφή κώδικα, οι οποίοι περιέχουν και καταδεικνύουν αναλυτικά τις λεπτομέρειες του κάθε Operator ξεχωριστά, επισυνάπτονται παρακάτω:

Κώδικας πρώτης διεργασίας:

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<process version="5.1.001">
  <context>
    <input/>
    <output/>
    <macros/>
  </context>
  <operator activated="true" class="process" compatibility="5.1.001" expanded="true"
name="Process">
    <process expanded="true" height="396" width="687">
      <operator activated="true" class="text:process_document_from_file" compatibility="5.1.001"
expanded="true" height="76" name="Process Documents from Files" width="90" x="45" y="30">
        <list key="text_directories">
          <parameter key="positive" value="C:\Users\Fotis-
Linaki\Desktop\diplwmatikh\rapidminer\opinion mining\pos"/>
          <parameter key="negative" value="C:\Users\Fotis-
Linaki\Desktop\diplwmatikh\rapidminer\opinion mining\neg"/>
        </list>
        <parameter key="encoding" value="UTF-8"/>
        <parameter key="vector_creation" value="Term Occurrences"/>
      </process expanded="true" height="396" width="705">
    </operator>
  </operator>
</process>
```

```
<operator activated="true" class="text:transform_cases" compatibility="5.1.001"
expanded="true" height="60" name="Transform Cases" width="90" x="45" y="30"/>
<operator activated="true" class="text:tokenize" compatibility="5.1.001" expanded="true"
height="60" name="Tokenize" width="90" x="179" y="120"/>
<operator activated="true" class="text:filter_by_length" compatibility="5.1.001"
expanded="true" height="60" name="Filter Tokens (by Length)" width="90" x="380" y="120">
  <parameter key="min_chars" value="3"/>
</operator>
<operator activated="true" class="text:stem_snowball" compatibility="5.1.001" expanded="true"
height="60" name="Stem (Snowball)" width="90" x="514" y="30">
  <parameter key="language" value="Spanish"/>
</operator>
<connect from_port="document" to_op="Transform Cases" to_port="document"/>
<connect from_op="Transform Cases" from_port="document" to_op="Tokenize"
to_port="document"/>
<connect from_op="Tokenize" from_port="document" to_op="Filter Tokens (by Length)"
to_port="document"/>
<connect from_op="Filter Tokens (by Length)" from_port="document" to_op="Stem
(Snowball)" to_port="document"/>
<connect from_op="Stem (Snowball)" from_port="document" to_port="document 1"/>
<portSpacing port="source_document" spacing="0"/>
<portSpacing port="sink_document 1" spacing="18"/>
<portSpacing port="sink_document 2" spacing="0"/>
</process>
</operator>
<operator activated="true" class="x_validation" compatibility="5.1.001" expanded="true"
height="130" name="Validation" width="90" x="246" y="75">
  <parameter key="number_of_validations" value="7"/>
  <parameter key="sampling_type" value="shuffled sampling"/>
  <process expanded="true" height="396" width="327">
    <operator activated="true" class="naive_bayes" compatibility="5.1.001" expanded="true"
height="76" name="Naive Bayes" width="90" x="112" y="30"/>
    <connect from_port="training" to_op="Naive Bayes" to_port="training set"/>
    <connect from_op="Naive Bayes" from_port="model" to_port="model"/>
    <portSpacing port="source_training" spacing="0"/>
    <portSpacing port="sink_model" spacing="0"/>
    <portSpacing port="sink_through 1" spacing="36"/>
    </process>
    <process expanded="true" height="396" width="327">
      <operator activated="true" class="apply_model" compatibility="5.1.001" expanded="true"
height="76" name="Apply Model" width="90" x="45" y="30">
        <list key="application_parameters"/>
      </operator>
      <operator activated="true" class="performance" compatibility="5.1.001" expanded="true"
height="76" name="Performance" width="90" x="179" y="120"/>
      <connect from_port="model" to_op="Apply Model" to_port="model"/>
      <connect from_port="test set" to_op="Apply Model" to_port="unlabelled data"/>
      <connect from_op="Apply Model" from_port="labelled data" to_op="Performance"
to_port="labelled data"/>
      <connect from_op="Performance" from_port="performance" to_port="averagable 1"/>
      <portSpacing port="source_model" spacing="0"/>
      <portSpacing port="source_test set" spacing="36"/>
      <portSpacing port="source_through 1" spacing="0"/>
      <portSpacing port="sink_averagable 1" spacing="0"/>
      <portSpacing port="sink_averagable 2" spacing="0"/>
      <portSpacing port="sink_averagable 3" spacing="0"/>
      </process>
    </operator>
    <operator activated="true" class="write_model" compatibility="5.1.001" expanded="true"
height="60" name="Write Model" width="90" x="405" y="182">
```



```
<parameter key="model_file" value="C:\Users\Fotis-  
Linaki\Desktop\diplwmatikh\rapidminer\opinion mining\ok.mod"/>  
</operator>  
<connect from_op="Process Documents from Files" from_port="example set" to_op="Validation"  
to_port="training"/>  
<connect from_op="Validation" from_port="model" to_op="Write Model" to_port="input"/>  
<connect from_op="Validation" from_port="training" to_port="result 1"/>  
<connect from_op="Validation" from_port="averagable 1" to_port="result 2"/>  
<connect from_op="Validation" from_port="averagable 2" to_port="result 3"/>  
<portSpacing port="source_input 1" spacing="18"/>  
<portSpacing port="sink_result 1" spacing="0"/>  
<portSpacing port="sink_result 2" spacing="72"/>  
<portSpacing port="sink_result 3" spacing="0"/>  
<portSpacing port="sink_result 4" spacing="0"/>  
</process>  
</operator>  
</process>
```

### Κώδικας δεύτερης διεργασίας:

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>  
<process version="5.1.001">  
<context>  
<input/>  
<output/>  
<macros/>  
</context>  
<operator activated="true" class="process" compatibility="5.1.001" expanded="true"  
name="Process">  
<process expanded="true" height="396" width="705">  
<operator activated="true" class="read_model" compatibility="5.1.001" expanded="true"  
height="60" name="Read Model" width="90" x="112" y="30">  
<parameter key="model_file" value="C:\Users\Fotis-  
Linaki\Desktop\diplwmatikh\rapidminer\opinion mining\ok.mod"/>  
</operator>  
<operator activated="true" class="text:process_document_from_file" compatibility="5.1.001"  
expanded="true" height="76" name="Process Documents from Files" width="90" x="112" y="165">  
<list key="text_directories">  
<parameter key="unlabeled" value="C:\Users\Fotis-  
Linaki\Desktop\diplwmatikh\rapidminer\opinion mining\pos"/>  
</list>  
<parameter key="encoding" value="UTF-8"/>  
<parameter key="vector_creation" value="Term Occurrences"/>  
<process expanded="true" height="396" width="705">  
<operator activated="true" class="text:transform_cases" compatibility="5.1.001"  
expanded="true" height="60" name="Transform Cases" width="90" x="45" y="30"/>  
<operator activated="true" class="text:tokenize" compatibility="5.1.001" expanded="true"  
height="60" name="Tokenize" width="90" x="246" y="30"/>  
<operator activated="true" class="text:filter_by_length" compatibility="5.1.001"  
expanded="true" height="60" name="Filter Tokens (by Length)" width="90" x="380" y="30">  
<parameter key="min_chars" value="3"/>  
</operator>  
<operator activated="true" class="text:stem_snowball" compatibility="5.1.001" expanded="true"  
height="60" name="Stem (Snowball)" width="90" x="581" y="30"/>  
<connect from_port="document" to_op="Transform Cases" to_port="document"/>  
<connect from_op="Transform Cases" from_port="document" to_op="Tokenize"  
to_port="document"/>
```

```
<connect from_op="Tokenize" from_port="document" to_op="Filter Tokens (by Length)"
to_port="document"/>
<connect from_op="Filter Tokens (by Length)" from_port="document" to_op="Stem
(Snowball)" to_port="document"/>
<connect from_op="Stem (Snowball)" from_port="document" to_port="document 1"/>
<portSpacing port="source_document" spacing="0"/>
<portSpacing port="sink_document 1" spacing="0"/>
<portSpacing port="sink_document 2" spacing="0"/>
</process>
</operator>
<operator activated="true" class="apply_model" compatibility="5.1.001" expanded="true"
height="76" name="Apply Model" width="90" x="313" y="30">
<list key="application_parameters"/>
</operator>
<operator activated="true" class="select_attributes" compatibility="5.1.001" expanded="true"
height="76" name="Select Attributes" width="90" x="447" y="30">
<parameter key="attribute_filter_type" value="value_type"/>
<parameter key="value_type" value="text"/>
</operator>
<operator activated="true" class="write_aml" compatibility="5.1.001" expanded="true"
height="60" name="Write AML" width="90" x="581" y="165">
<parameter key="example_set_file" value="C:\Users\Fotis-
Linaki\Desktop\diplwmatikh\rapidminer\opinion mining\προβλεψη 1λκ283"/>
<parameter key="encoding" value="UTF-8"/>
</operator>
<connect from_op="Read Model" from_port="output" to_op="Apply Model" to_port="model"/>
<connect from_op="Process Documents from Files" from_port="example set" to_op="Apply
Model" to_port="unlabelled data"/>
<connect from_op="Apply Model" from_port="labelled data" to_op="Select Attributes"
to_port="example set input"/>
<connect from_op="Select Attributes" from_port="example set output" to_op="Write AML"
to_port="input"/>
<portSpacing port="source_input 1" spacing="0"/>
<portSpacing port="sink_result 1" spacing="0"/>
</process>
</operator>
</process>
```

## ***Κεφάλαιο 5. Συγκριτικά στοιχεία με χρήση αγγλικής γλώσσας***

---

## 5.1 Ηλεκτρονική Διακυβέρνηση και Ιστός 2.0

Ο όρος «Ηλεκτρονική Διακυβέρνηση» αναφέρεται στη χρησιμοποίηση Τεχνολογιών Πληροφορικής και Επικοινωνιών (ΤΠΕ) με σκοπό να βελτιώσει την ποιότητα των υπηρεσιών και πληροφοριών που προσφέρονται στους πολίτες, να ενδυναμώσει το θεσμό της δημοκρατίας και να υποστηρίξει τις δημόσιες πολιτικές προς την κατεύθυνση καλύτερης και ταχύτερης εξυπηρέτησης των πολιτών, αλλά και πάταξης φαινομένων γραφειοκρατίας και διαφθοράς. Παρ' όλο που η «Ηλεκτρονική Διακυβέρνηση» (e- government ή αλλιώς e- governance στα αγγλικά ή απλά e-gov που είναι η συντομογραφία με την οποία συναντάμε συχνότερα τον όρο) είναι άρρηκτα συνδεδεμένη με σύγχρονες τεχνολογίες λογισμικών και εξελιγμένα συστήματα δικτύων, στην πράξη πρωταγωνιστής δεν είναι το τεχνολογικό αυτό κομμάτι καθώς για να πετύχει η ηλεκτρονική διακυβέρνηση απαιτείται αλλαγή στις οργανωτικές δομές του δημόσιου τομέα, με ταυτόχρονη εκπαίδευση των δημόσιων υπαλλήλων, δημιουργία ψηφιακής και τεχνολογικής κουλτούρας στον πολίτη, και, πάνω από όλα, ενημέρωση προς όλες τις πλευρές. Έτσι λοιπόν, στις μέρες μας θέματα σχετικά με κυβερνητικές αποφάσεις που αφορούν την ηλεκτρονική διακυβέρνηση βρίσκονται ανάμεσα στα πιο ευρέως συζητούμενα στις ψηφιακές κοινωνίες σε παγκόσμιο επίπεδο. Σε αυτό αναμφίβολα συνέβαλε καθοριστικά η χρησιμοποίηση της νέας γενιάς του Παγκόσμιου Ιστού, του Web 2.0 (Ιστός 2.0), που δίνει τη δυνατότητα στους χρήστες του διαδικτύου να επικοινωνούν και να αλληλεπιδρούν μεταξύ τους, σχηματίζοντας μια νέα μορφή «κοινωνίας», την λεγόμενη ψηφιακή, ανταλλάζοντας γνώμες και απόψεις πάνω στα θέματα της επικαιρότητας. Προς αυτή την κατεύθυνση και κάνοντας χρήση των πλεονεκτημάτων της νέας γενιάς του Διαδικτύου, οι κυβερνήσεις σε παγκόσμιο επίπεδο έχουν ξεκινήσει να μοιράζονται πλέον τον κυριότερο όγκο των αποφάσεων και μεταρρυθμίσεων τους διαδικτυακά, δίνοντας στους πολίτες την ευκαιρία όχι μόνο να ενημερωθούν πληρέστερα και ταχύτερα για αυτές, αλλά κυρίως να συμμετέχουν ενεργά στη διαμόρφωσή τους παραθέτοντας τις προσωπικές τους κρίσεις, την αξιολόγησή τους ακόμα και τις προτάσεις τους πάνω σε αυτές. Μπορεί εύκολα, λοιπόν, ο καθένας να κατανοήσει ότι η δυνατότητα αξιολόγησης των απόψεων των πολιτών όπως αυτές εκφράζονται στα διάφορα κοινωνικά μέσα (social media) είναι υψίστης σημασίας για την επιτυχή μετάβαση προς την ψηφιακή εποχή διακυβέρνησης.

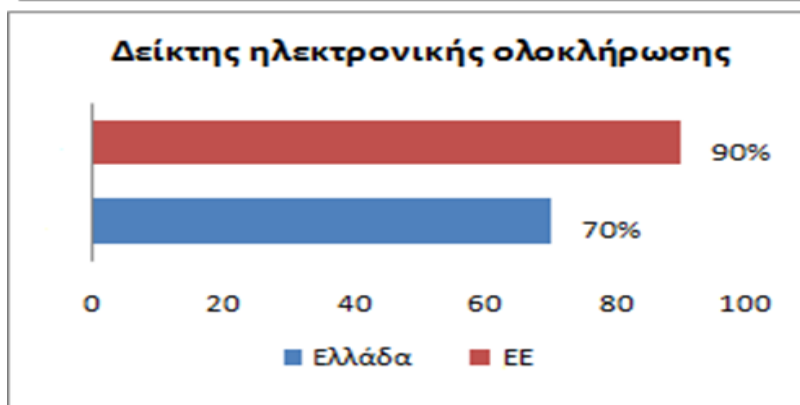
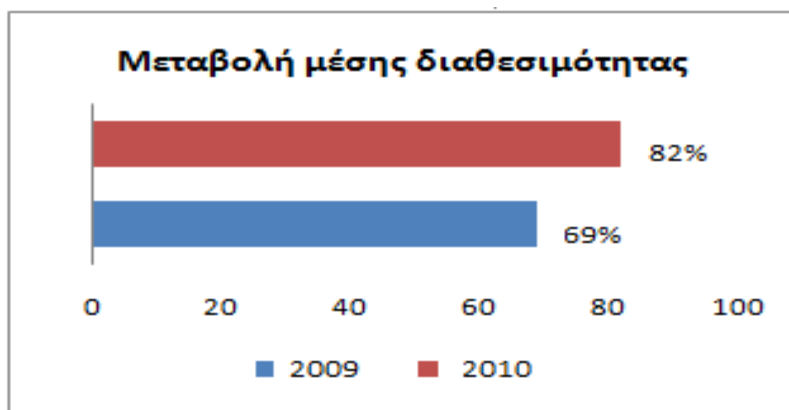
## 5.2 Στατιστικά στοιχεία για ΕΕ και Ελλάδα

Εδώ κρίνεται σκόπιμο να παραθέσουμε κάποια στατιστικά στοιχεία για τη διείσδυση και την εξέλιξη της «Ηλεκτρονικής Διακυβέρνησης» στη χώρα μας σε σύγκριση και με τις υπόλοιπες Ευρωπαϊκές χώρες, όπως αυτά δημοσιεύτηκαν

σύμφωνα με την «9η έκθεση συγκριτικής αξιολόγησης της ηλεκτρονικής διακυβέρνησης στην Ευρώπη» για το έτος 2010 και δημοσιεύονται στην ιστοσελίδα του Εθνικού Κέντρου Τεκμηρίωσης: [www.ekt.gr](http://www.ekt.gr). Η συγκεκριμένη έκθεση, που πραγματοποιείται από το 2001, έχει ως αντικείμενο τη σύγκριση της προόδου στις διάφορες ευρωπαϊκές χώρες και η ανάλυση καλύπτει περισσότερους από 10.000 ιστότοπους. Για τη συγκριτική αξιολόγηση λαμβάνονται υπόψη 12 βασικές διαδικτυακές υπηρεσίες προς τους πολίτες και 8 ηλεκτρονικές δημόσιες υπηρεσίες προς επιχειρήσεις.

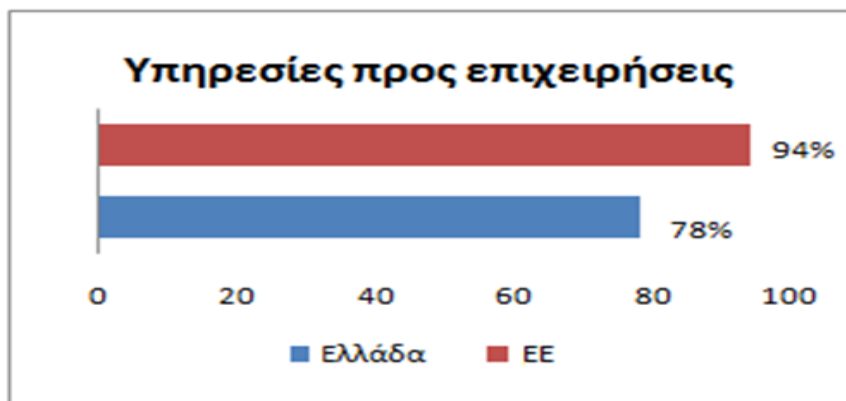
### **5.2.1 Μέση διαθεσιμότητα ηλεκτρονικών υπηρεσιών**

Από την τελευταία έκθεση προκύπτει ότι η μέση διαθεσιμότητα των διαδικτυακών δημοσίων υπηρεσιών στην Ευρωπαϊκή Ένωση αυξήθηκε από 69% το 2009 σε 82% το 2010, προσεγγίζοντας τον στόχο που έχει τεθεί ώστε, ως το 2015, ένα στους δύο πολίτες και τέσσερις στις πέντε επιχειρήσεις να χρησιμοποιούν υπηρεσίες ηλεκτρονικής διακυβέρνησης. Η Ελλάδα, με 10 από τις 20 υπηρεσίες πλήρως διαθέσιμες ηλεκτρονικά (5 για πολίτες και 5 για επιχειρήσεις), βρίσκεται στην τελευταία θέση της σχετικής κατάταξης. Πιο συγκεκριμένα, ο αντίστοιχος δείκτης ηλεκτρονικής ολοκλήρωσης (online sophistication) για τις εν λόγω 20 βασικές υπηρεσίες για τη χώρα μας είναι 70% ενώ ο μέσος όρος για τις υπόλοιπες χώρες αγγίζει το 90%. Για πληρέστερη εικόνα, μετά την παράθεση ποσοστών θα ακολουθούν κάθε φορά και τα αντίστοιχα διαγράμματα:



Αναλυτικότερα, ο δείκτης για τις παρεχόμενες ηλεκτρονικά υπηρεσίες προς τους πολίτες είναι 65% και για τις υπηρεσίες προς επιχειρήσεις είναι 78%, τη στιγμή που οι αντίστοιχοι δείκτες για την ΕΕ διαμορφώνονται στο 87% και 94% αντίστοιχα.





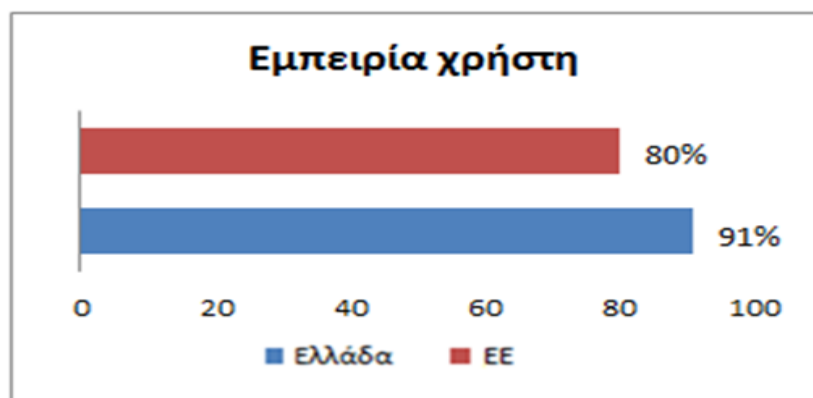
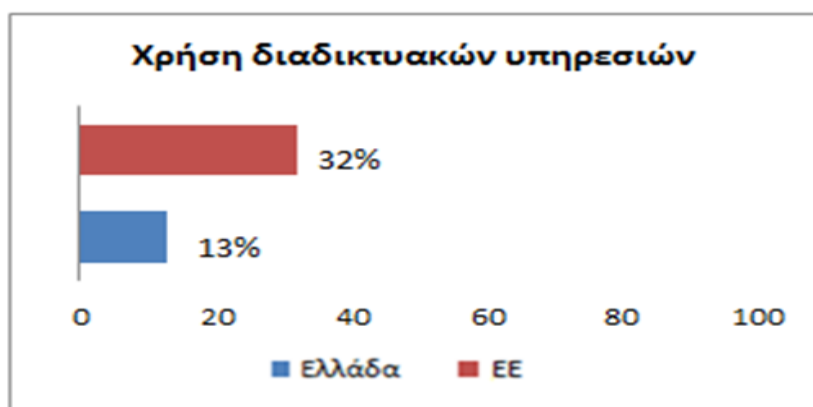
Για να γίνει καλύτερα αντιληπτό, σημειώνουμε εδώ ότι οι δείκτες αφορούν την ηλεκτρονική ολοκλήρωση (online sophistication) που αποδίδει στην κάθε υπηρεσία μια επίδοση (%) ανάλογη του επιπέδου ηλεκτρονικής ωριμότητας σε πληροφόρηση, ηλεκτρονική λήψη εντύπων, ηλεκτρονική αποστολή συμπληρωμένων εντύπων, ολοκλήρωση συναλλαγής και εξατομικευμένη παροχή υπηρεσίας. Από τον μέσο όρο των μετρήσεων για τις 20 υπηρεσίες προκύπτει τελικώς ο συνολικός δείκτης ηλεκτρονικής ολοκλήρωσης της χώρας ενώ αντίστοιχα υπολογίζεται ο δείκτης πλήρους ηλεκτρονικής διαθεσιμότητας (full online availability).

Όσον αφορά τη χώρα μας, εκτός από τους δείκτες που ήδη παρατέθηκαν προηγουμένως, σκόπιμο επίσης κρίνεται να αναφερθεί πως το portal opengov.gr θεωρείται η σημαντικότερη πρωτοβουλία για την ηλεκτρονική διακυβέρνηση ενώ η χειρότερη επίδοση καταγράφεται στον τομέα των προμηθειών λόγω της έλλειψης σχετικής πλατφόρμας για την ηλεκτρονική διεκπεραίωση αυτών. Για λόγους πληρότητας σημειώνουμε εδώ πως οι καλύτερες επιδόσεις σημειώθηκαν στην Αυστρία, την Ιταλία, την Μάλτα, την Πορτογαλία και την Σουηδία, όπου το σύνολο του καταλόγου των βασικών υπηρεσιών παρέχεται ηλεκτρονικά, ενώ σημαντική βελτίωση στη διαδικτυακή παροχή υπηρεσιών σημειώθηκε το τελευταίο έτος στην Βουλγαρία, την Ιταλία, την Κροατία, την Λετονία και την Σουηδία. Τέλος, όπως προκύπτει από την έκθεση, οι δημόσιες υπηρεσίες που αφορούν τις επιχειρήσεις έχουν σημειώσει μεγαλύτερη πρόοδο από εκείνες που αφορούν τους πολίτες.

### 5.2.2 Ποσοστό χρησιμοποίησης και εμπειρία χρήστη

Βέβαια, και για την παρούσα εργασία, σημαντικότερο στοιχείο αποτελεί το ποσοστό των ατόμων που χρησιμοποιούν το διαδίκτυο για συναλλαγές με το δημόσιο το οποίο, σύμφωνα πάντα με την εν λόγω έκθεση, ανέρχεται στο 13% για την Ελλάδα, ενώ για την υπόλοιπη Ευρώπη διαμορφώνεται στο 32%. Από όλα τα κράτη καταβάλλονται σημαντικές προσπάθειες ώστε να βελτιωθεί η εμπειρία του χρήστη (user experience) κατά τη χρήση βασικών υπηρεσιών. Η εμπειρία του χρήστη βασίζεται σε κριτήρια που έχουν να κάνουν με: την ευκολία χρήσης υπηρεσίας,

διαφάνειας, προστασίας των προσωπικών δεδομένων, παροχής της υπηρεσίας μέσω πολλαπλών διαύλων και δυνατότητας καταγραφής της άποψης του χρήστη για την ποιότητα της υπηρεσίας. Αυτά αποτελούν και τα βασικότερα στοιχεία πάνω στα οποία καταγράφηκαν οι περισσότερες γνώμες χρηστών του διαδικτύου σχετικά με την ηλεκτρονική διακυβέρνηση και αποτελούν τον κύριο όγκο συζητήσεων στις «ψηφιακές κοινωνίες». Στατιστικά, η επίδοση της Ελλάδας για το συγκεκριμένο δείκτη διαμορφώνεται στο 91%, έναντι 80% του μέσου όρου των 32 χωρών της έρευνας, γεγονός που φέρνει τη χώρα μας στην έκτη θέση και αποτελεί σαφώς βελτιωμένη επίδοση.



### 5.3 Άποψη Ελλήνων πολιτών για Ηλεκτρονική Διακυβέρνηση

Λαμβάνοντας υπόψη τα παραπάνω στατιστικά στοιχεία και κυρίως βασιζόμενοι στο γεγονός ότι περίπου μόλις ένας στους δέκα Έλληνες χρησιμοποιεί το διαδίκτυο για τις συναλλαγές του με το δημόσιο, είναι εύκολο να αναλογιστεί κανείς πως η χώρα μας βρίσκεται ακόμα σε πρώιμο στάδιο όσον αφορά την διείσδυση της ηλεκτρονικής διακυβέρνησης. Μάλιστα, σε ευρωπαϊκό επίπεδο, το αντίστοιχο ποσοστό είναι διπλάσιο, και εδώ ακριβώς έγκειται η πρώτη και ίσως μεγαλύτερη δυσκολία που αντιμετωπίστηκε κατά την εκπόνηση της παρούσας διπλωματικής. Η



μη εξοικείωση και πολλές φορές η άγνοια των Ελλήνων χρηστών του διαδικτύου σχετικά με υπηρεσίες ηλεκτρονικής διακυβέρνησης σίγουρα δυσχεραίνει το έργο καταγραφής, ανάλυσης και αξιολόγησης των απόψεων τους πάνω στο συγκεκριμένο θέμα, καθώς μεγαλύτερος όγκος πληροφορίας θα οδηγούσε σε ορθότερα και σαφέστερα συμπεράσματα.

Ακόμη, η ύπαρξη περισσότερων και πολυπληθέστερων «ψηφιακών κοινωνιών» ενεργών στο θέμα της ηλεκτρονικής διακυβέρνησης θα βοηθούσε στη δημιουργία ενός πληρέστερου λεξιλογίου χρήσης που απαντάται συχνότερα για τον σχολιασμό αυτό. Εδώ σημειώνουμε ότι ο διαχωρισμός λέξεων που εκφράζουν θετική ή αρνητική άποψη δεν είναι αρκετός καθώς συγκεκριμένες θετικές ή αρνητικές λέξεις θα μπορούσαν να χρησιμοποιηθούν ώστε να χαρακτηρίσουν ένα κυβερνητικό ζήτημα όπως αυτό της ηλεκτρονικής διακυβέρνησης. Για παράδειγμα, αναμφίβολα, οι λέξεις όμορφος ή άσχημος περιέχουν θετικό και αρνητικό συναίσθημα αντίστοιχα, αλλά μάλλον είναι σχεδόν απίθανο να εντοπίσουμε σχόλια του τύπου : «Η ηλεκτρονική διακυβέρνηση είναι όμορφη/ άσχημη». Έτσι λοιπόν, περισσότερος σχολιασμός θα οδηγούσε στη δημιουργία μεγαλύτερου όγκου εξειδικευμένου λεξιλογίου απαραίτητου ως βάσης για εφαρμογές στο πεδίο του opinion mining. Άλλωστε, για να γίνει και μία πρώτη σύγκριση με τα αποτελέσματα που θα είχαμε εάν τυχόν επιχειρούσαμε αντίστοιχη ανάλυση σε μια γλώσσα όπως την αγγλική, εκτός του γεγονότος ότι πρόκειται για μία γλώσσα ευρέως διαδεδομένη σε παγκόσμια κλίμακα, επιπρόσθετα πρόκειται και για μία γλώσσα που χρησιμοποιείται σε κράτη όπου οι πολίτες έχουν μεγαλύτερη εξοικείωση και είναι πιο ενεργοί «ψηφιακά» στο θέμα της ηλεκτρονικής διακυβέρνησης, σύμφωνα και με τα στοιχεία που προκύπτουν από τις αντίστοιχες μελέτες.

### **5.3.1 Σημαντικότερες ηλεκτρονικές υπηρεσίες στην Ελλάδα**

Για την Ελλάδα, οι υπηρεσίες που βρίσκονται στο επίκεντρο χρήσης διαδικτυακά και συνεπώς είναι και αυτές που σχολιάζονται περισσότερο είναι η εθνική πύλη δημόσιας διοίκησης Ermis: [www.ermis.gov.gr](http://www.ermis.gov.gr), που αποτελεί την ηλεκτρονική πύλη των ΚΕΠ και οι υπηρεσίες του ΙΚΑ που προσφέρονται διαδικτυακά στην αντίστοιχη ιστοσελίδα: [www.ika.gr](http://www.ika.gr), με κυρίαρχες την ηλεκτρονική υποβολή Αναλυτικής Περιοδικής Δήλωσης (ΑΠΔ) και ηλεκτρονική έκδοση Ασφαλιστικής Ενημερότητας. Βέβαια, η υπηρεσία ηλεκτρονικής διακυβέρνησης που κατέχει τα πρωτεία σε χρήση και σχολιασμό αντίστοιχα είναι το ονομαζόμενο TAXISnet, που αποτελεί την ηλεκτρονική πύλη του δημοσίου που παρέχει τη δυνατότητα σε πολίτες και επιχειρήσεις να διεκπεραιώνουν τις φορολογικές τους υποχρεώσεις διαδικτυακά.

### 5.3.2 Σημαντικότεροι ιστότοποι άντλησης γνώμης

Σχετικά με τον σχολιασμό των υπηρεσιών, όπως ήταν αναμενόμενο άλλωστε, πιο ενεργές υπήρξαν οι διαδικτυακές κοινότητες επαγγελματιών που αντικείμενο της δουλειάς τους είναι η διεκπεραίωση φορολογικών υποθέσεων πολιτών και επιχειρήσεων απέναντι στο κράτος, όπως για παράδειγμα η «ψηφιακή» κοινωνία του portal: [www.taxheaven.gr](http://www.taxheaven.gr), όπου στο αντίστοιχο forum υπάρχουν πολλές ενότητες συζητήσεων και θεμάτων σχετικά με τις παρεχόμενες υπηρεσίες ηλεκτρονικής διακυβέρνησης και την ποιότητα αυτών. Ουσιαστική είναι και η ύπαρξη αντίστοιχου forum στην ηλεκτρονική πύλη ermis, όπου οι πολίτες μπορούν να υποβάλουν τις ερωτήσεις, σχόλια και προτάσεις τους σχετικά με τις υπηρεσίες που παρέχει το εν λόγω site.

Φυσικά, όπως αναφέρθηκε και παραπάνω, την ουσιαστικότερη πρωτοβουλία αποτελεί ο διαδικτυακός ιστότοπος: [www.opengov.gr](http://www.opengov.gr), που αντικατοπτρίζει άμεσα τις αντιδράσεις των Ελλήνων πολιτών σχετικά με την εξέλιξη της ηλεκτρονικής διακυβέρνησης στη χώρα μας και αποτελεί πόλο έλξης σχολιασμού. Στο εν λόγω site τίθενται σε διαβούλευση τα νομοσχέδια όλων των υπουργείων και οι πολίτες μπορούν ελεύθερα να υποβάλουν τις σκέψεις, αντιδράσεις, προβληματισμούς και προτάσεις πάνω σε αυτές. Το [opengov.gr](http://opengov.gr) αποτέλεσε μία από τις κύριες πηγές από όπου συλλέχθηκαν απόψεις χρηστών του διαδικτύου σχετικά με την ηλεκτρονική διακυβέρνηση, απαραίτητες για τη διεξαγωγή της παρούσας εργασίας.

Τέλος, σχολιασμός περί του θέματος υπήρξε και σε διάφορα άλλα portals ποικίλου περιεχομένου, χωρίς όμως οι εκάστοτε «ψηφιακές» κοινωνίες να παραμένουν ενεργές για αξιοσημείωτο χρονικό διάστημα στο συγκεκριμένο ζήτημα καθώς δίνεται περισσότερο βάση σε θέματα επικαιρότητας και ο σχολιασμός για τις διάφορες υπηρεσίες αυξάνεται και μειώνεται ανάλογα. Για παράδειγμα, σε περιόδους που ανακοινώνεται ένα νέο νομοσχέδιο σχετικά με την ηλεκτρονική διακυβέρνηση είτε παρουσιάζεται μία νέα υπηρεσία που μπορεί πλέον να εκτελεστεί και διαδικτυακά ή αντίστοιχα υπάρχουν αλλαγές και βελτιώσεις σε μία ήδη υπάρχουσα υπηρεσία, ο σχολιασμός αυξάνεται, αναλόγως και του ενδιαφέροντος των πολιτών στο αντίστοιχο θέμα.

### 5.3.3 Κύριες καταγεγραμμένες απόψεις

Στο σημείο αυτό κρίνουμε σκόπιμο να επιχειρήσουμε μια μικρή ανάλυση των σχολίων και γνωμών που συγκεντρώθηκαν από τις διάφορες διαδικτυακές πηγές και να διεξάγουμε κάποιες παρατηρήσεις αλλά και συμπεράσματα. Αρχικά, σημαντικό είναι το γεγονός ότι επικρατεί μία σχετική ισορροπία μεταξύ θετικών και αρνητικών απόψεων, ενώ υπάρχει απουσία έντονων θετικών ή αρνητικών σχολίων αντίστοιχα.

Πολλά είναι τα σχόλια που κυμαίνονται σε πιο ουδέτερους τόνους, είτε υποδηλώνοντας κάποια επιφύλαξη στο να εκφέρουν ακόμα ολοκληρωμένη άποψη επί της ηλεκτρονικής διακυβέρνησης και των αντίστοιχων παρεχόμενων υπηρεσιών ή μέσα στο ίδιο κείμενο που προέρχεται από τον ίδιο συγγραφέα εκφράζονται τόσο θετικές όσο και αρνητικές απόψεις, οπότε και το συνολικό κείμενο κρίνεται ως ουδέτερο. Φυσικά, η ουδετερότητα ενός σχολίου δεν βοηθά σε εφαρμογές κατηγοριοποίησης κειμένου καθώς ένας αλγόριθμος θα πρέπει να έχει εκπαιδευτεί με χιλιάδες αντίστοιχων σχολίων ώστε να είναι σε θέση να παρέχει μία ικανοποιητική εκτίμηση αυτών. Ακόμα, η έντονη παρουσία ουδέτερων σχολίων αποτέλεσε εμπόδιο στη δημιουργία πληρέστερης βάσης λεξιλογίου θετικού και αρνητικού σχολιασμού λόγω της απουσίας αντίστοιχων χαρακτηριστικών εκφράσεων.

Βέβαια, παρά την σχετική ισορροπία των απόψεων παρατηρείται, όπως ήταν φυσικό, μια υπεροχή του πλήθους των αρνητικών σχολίων. Άλλωστε σε γενικές γραμμές, όπως συμβαίνει και όταν πρόκειται για το σχολιασμό παρόμοιων καταστάσεων, είναι πιο πιθανό ένας χρήστης του διαδικτύου να παραθέσει την άποψή του προκειμένου να επισημάνει μια προβληματική, κατά την γνώμη του, κατάσταση παρά για να επιδοκιμάσει κάτι θετικό. Για να γίνει πιο κατανοητό, η δημιουργία ενός θέματος συζήτησης που αφορά τυχόν προβλήματα, καθυστερήσεις και παραλείψεις στο σχεδιασμό και τη χρήση μιας διαδικτυακής υπηρεσίας ηλεκτρονικής διακυβέρνησης είναι σαφώς πιο συχνή καθώς ο χρήστης με τον τρόπο αυτό επιθυμεί να αναδείξει τα κακώς κείμενα και να συμβάλει στην επιτυχή βελτίωσή τους από την οποία και θα επωφεληθεί μελλοντικά. Αντίθετα και συγκριτικά πάντα, είναι λιγότερο πιθανό να ανοιχθεί ένα νέο θέμα συζήτησης ώστε απλά να επαινεθεί η ποιότητα και τα πλεονεκτήματα χρήσης μιας υπηρεσίας. Το παραπάνω συμπέρασμα δεν προκύπτει βέβαια θεωρητικά, αλλά επιβεβαιώθηκε κρίνοντας από την αναλογία περιεχομένου και πλήθους διαδικτυακών συζητήσεων στο θέμα της ηλεκτρονικής διακυβέρνησης.

Όσον αφορά τις απόψεις που κυριαρχούν στις διαδικτυακές κοινότητες, πέρα από την επισήμανση τυχόν προβλημάτων που παρατηρούνται κατά τη χρήση των υπηρεσιών, προκύπτει ότι ένα μεγάλο ποσοστό είναι θετικά προσκείμενο στην ιδέα της ηλεκτρονικής διακυβέρνησης και αναγνωρίζει τα πλεονεκτήματα που μπορεί να προσφέρει σε θεωρητικό επίπεδο, προβληματίζεται όμως για το κατά πόσο μπορεί να εφαρμοστεί ικανοποιητικά και χωρίς παρενέργειες κάτι τέτοιο στην πράξη. Υπάρχει, λοιπόν, μια γενική έντονη δυσαρέσκεια ως προς το επίπεδο των δημόσιων υπηρεσιών που παρέχονται από το κράτος, κυρίως λόγω της γραφειοκρατίας που αντιμετωπίζεται, κάτι που οδηγεί τους πολίτες να είναι επιφυλακτικοί (εάν όχι απαισιόδοξοι) ως προς οποιαδήποτε βελτίωση που μπορεί να επέλθει με την εισαγωγή των αντίστοιχων ηλεκτρονικά διαθέσιμων υπηρεσιών. Επαναλαμβάνεται εδώ ότι η συγκεκριμένη προκατάληψη και επομένως αρνητική στάση δεν πηγάζει

από την έως τώρα πορεία της ηλεκτρονικής διακυβέρνησης στη χώρα μας και δεν αφορά αυτή καθ' αυτή, αλλά προκύπτει από διάφορους παράγοντες συνυφασμένους με τις δημόσιες υπηρεσίες.

Από το σύνολο των επαγγελματιών που έχουν πιο συχνές συναλλαγές με το δημόσιο σε σχέση με τον μέσο πολίτη, διάχυτη είναι η άποψη ότι η ηλεκτρονική διακυβέρνηση είναι κάτι που θα έπρεπε να έχει εφαρμοστεί πολύ νωρίτερα, παρ' όλο που συχνά και αυτοί εκφράζουν τη δυσαρέσκειά τους σχετικά με κενά και παραλείψεις που παρατηρούν. Από την άλλη πλευρά, συχνή είναι η επισήμανση της μη εξοικείωσης μεγάλου μέρους του πληθυσμού με το διαδίκτυο και τις σύγχρονες τεχνολογίες, κυρίως όσον αφορά άτομα μεγαλύτερης ηλικίας, το λεγόμενο σύνολο των «τεχνολογικά και διαδικτυακά αναλφάβητων». Το συγκεκριμένο κομμάτι της κοινωνίας, πέρα του γεγονότος ότι μπορεί να μην διαθέτει καν πρόσβαση στο διαδίκτυο, χαρακτηρίζεται και από έντονο φόβο χρήσης των συγκεκριμένων υπηρεσιών κυρίως ως προς την επιτυχή διεκπεραίωση των υποθέσεων του με το δημόσιο, αλλά και την ασφάλεια των προσωπικών του δεδομένων. Επιπλέον, προβληματισμός πάνω στο συγκεκριμένο ζήτημα δεν υπάρχει μόνο για τους πολίτες που θα χρησιμοποιήσουν τις εν λόγω υπηρεσίες, αλλά και ως προς τους αρμόδιους υπαλλήλους που ίσως να μην είναι εκπαιδευμένοι κατάλληλα για να χειριστούν τις αλλαγές που επιφέρει η τεχνολογία στην παραδοσιακή χρήση των υπηρεσιών.

## **5.4 Σύγκριση αποτελεσμάτων με χρήση εφαρμογής στην αγγλική γλώσσα**

Ανακεφαλαιώνοντας ως εδώ, το μικρό συγκριτικά ποσοστό Ελλήνων που είναι ενημερωμένο σε θέματα ηλεκτρονικής διακυβέρνησης και το ακόμα μικρότερο που κάνει χρήση των διαθέσιμων υπηρεσιών, 13% συγκεκριμένα όπως αναφέρθηκε και παραπάνω, δυσκόλεψαν την συλλογή πληροφοριών για τη δημιουργία κατάλληλου λεξιλογίου χαρακτηρισμού απόψεων. Σε συνδυασμό με κάποιους άλλους δυσμενείς παράγοντες που αναλύθηκαν προηγουμένως και φυσικά λαμβάνοντας υπ' όψιν την πολυπλοκότητα της ελληνικής γλώσσας και τη μικρή της διάδοση σε παγκόσμιο επίπεδο, αποτέλεσαν τελικά τα κύρια εμπόδια στην εφαρμογή αυτόματης ταξινόμησης συναισθήματος που πραγματοποιήθηκε στην παρούσα διπλωματική εργασία. Στη συνέχεια παρατίθεται σύγκριση της ακρίβειας των αποτελεσμάτων που θα είχαμε εάν εκτελούσαμε την εφαρμογή που δημιουργήσαμε μέσω του RapidMiner στην αγγλική γλώσσα.

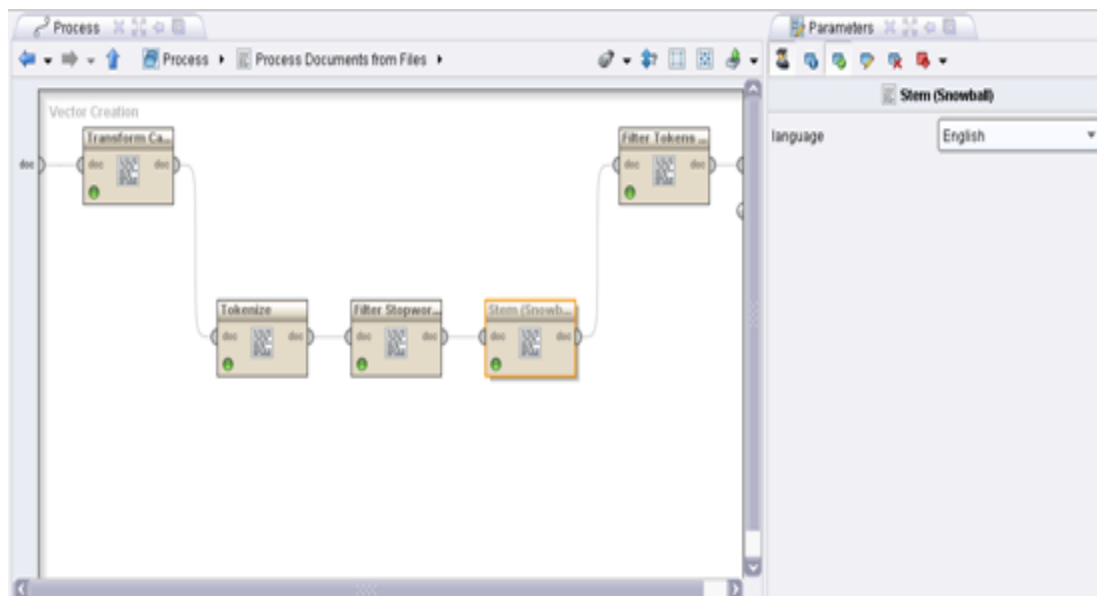
### 5.4.1 Τροποποίηση εφαρμογής RapidMiner

Αρχικά, και όσον αφορά το τεχνικό κομμάτι της εφαρμογής μας, η προσθήκη κάποιων επιπλέον συναρτήσεων που είναι έτοιμες και διαθέσιμες από το πρόγραμμα για την αγγλική γλώσσα, καθώς επίσης και για ένα πλήθος άλλων γλωσσών ευρέως διαδεδομένων, θα οδηγήσει σε πιο ακριβή αποτελέσματα και συμπεράσματα. Το γεγονός αυτό είναι και η σημαντικότερη παράμετρος που διαφοροποιεί, όσον αφορά πάντα την παρούσα εργασία, τη χρήση αγγλικής και ελληνικής γλώσσας σε εφαρμογές opinion mining. Οι προσθήκες, λοιπόν, θα γίνουν στον Operator: «Process Documents from Files» και πιο συγκεκριμένα στην υποδιεργασία που δημιουργήθηκε. Η υποδιεργασία αυτή έχει ιδιαίτερη βαρύτητα όσον αφορά το τελικό αποτέλεσμα, καθώς επεξεργάζεται ουσιαστικά τη βάση των ταξινομημένων σχολίων, θετικών και αρνητικών, που έχουμε εισάγει προηγουμένως.

Όπως θα φανεί και παρακάτω όπου θα εξηγηθούν αναλυτικότερα οι προσθήκες, είναι πολύ σημαντικό, για παράδειγμα, από τα δεδομένα που εισάγουμε να μπορεί το πρόγραμμα αυτόματα να εντοπίζει το κύριο θέμα της κάθε λέξης ώστε όταν θα βρει την ίδια λέξη με διαφορετική κατάληξη να την αναγνωρίζει. Γίνεται κατανοητό ότι σε μία γλώσσα όπως η ελληνική που η ίδια λέξη μπορεί να βρεθεί με διάφορες καταλήξεις λόγω αλλαγής γένους, πτώσης ή αριθμού (για άρθρα, ουσιαστικά, επίθετα και αντωνυμίες) αλλά και προσώπου, χρόνου ή φωνής (για τα ρήματα), κάτι τέτοιο κρίνεται υψίστης σημασίας. Συνολικά, τα επιθέματα για ουσιαστικά, επίθετα και ρήματα υπολογίζονται στα 166 για την ελληνική γλώσσα. Την επεξεργασία αυτή που περιγράφεται κάνει ένας στελεχωτής (stemmer). Δυστυχώς, για τα ελληνικά δεν υπάρχει έτοιμος στελεχωτής στο RapidMiner, γεγονός που αναμφίβολα δυσχεραίνει την όποια επεξεργασία των δεδομένων, πόσο μάλλον αν αναλογιστεί κανείς το πλήθος των επιθεμάτων της γλώσσας.

Αντίθετα, για δεδομένα στην αγγλική γλώσσα υπάρχει διαθέσιμος ο Operator: «Stem (Porter)», βασιζόμενος στον ομώνυμο αλγόριθμο Porter, που χρησιμοποιείται για την αφαίρεση επιθεμάτων των αγγλικών λέξεων. Ακόμα, μπορούμε να διαλέξουμε αντί αυτού τον Operator: «Stem (Snowball)», επιλέγοντας από τη διαθέσιμη λίστα γλωσσών την αγγλική. Επιπλέον, ένας άλλος σημαντικός Operator που μπορεί να προστεθεί στην εν λόγω υποδιεργασία είναι ο «Filter Stopwords (English)», που αφαιρεί από το αρχείο που έχουμε εισάγει λέξεις που ταυτίζονται με μία λίστα που διαθέτει. Η συγκεκριμένη επεξεργασία, που χρησιμοποιείται κατά κόρον από τις μηχανές αναζήτησης, βοηθάει στο να αφαιρεθούν οι πολύ κοινές λέξεις που μπορεί να υπάρχουν σε ένα κείμενο και δεν προσφέρουν κάτι στην επεξεργασία αυτού ώστε να επιταχυνθούν τα αποτελέσματα της διαδικασίας. Παρακάτω παρουσιάζεται η ολοκληρωμένη υποδιεργασία με την οποία αντικαθίσταται η ήδη υπάρχουσα στο RapidMiner. Προσθέτουμε εδώ ότι

απαραίτητα είναι η αντικατάσταση της ίδιας υποδιεργασίας και στη δεύτερη διεργασία καθώς οι δύο αυτές υποδιεργασίες πρέπει να συμπίπτουν αφού ανήκουν στον ίδιο Operator : «Process Documents from Files».



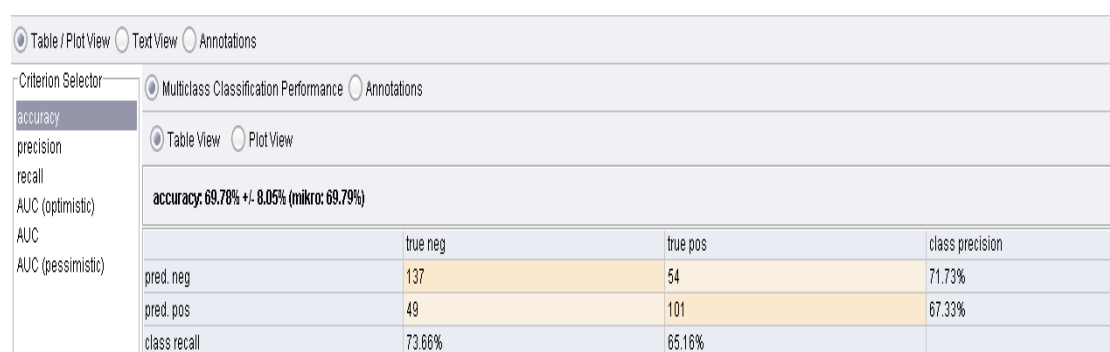
Οι παραπάνω αλλαγές αφορούν το αμιγώς τεχνικό κομμάτι των διαδικασιών που δημιουργήθηκαν μέσω του RapidMiner. Φυσικά, απαραίτητη είναι και πάλι η ύπαρξη των κατάλληλων και ταξινομημένων σχολίων για την εκπαίδευση του μοντέλου που θα εφαρμοστεί για την πρόβλεψη συναισθήματος σε νέα σχόλια. Για εφαρμογές opinion mining στην αγγλική γλώσσα, η βάση σχολίων που χρησιμοποιείται κατά κόρον για την εκπαίδευση τέτοιου είδους αλγορίθμων είναι μία βάση σχολίων που βασίζεται σε κριτικές ταινιών. Η εν λόγω βάση περιέχει ταξινομημένες 1000 θετικές κριτικές και αντίστοιχα 1000 αρνητικές, ένα αν μη τι άλλο ικανοποιητικό αριθμό που θα οδηγήσει σε ασφαλή συμπεράσματα. Μάλιστα, λόγω της ευρείας χρήσης της συγκεκριμένης βάσης σε εφαρμογές opinion mining, οι ερευνητές έχουν τελειοποιήσει την παρουσίαση αυτών των σχολίων ώστε να καλύπτει όσο το δυνατόν μεγαλύτερο και πληρέστερο εύρος της αγγλικής γλώσσας στο πεδίο του σχολιασμού. Για τη σύγκριση των αποτελεσμάτων που θα προκύψουν αν χρησιμοποιήσουμε την αγγλική γλώσσα στην εφαρμογή που ήδη δημιουργήσαμε μέσω RapidMiner, η βάση αυτή κρίθηκε ως η πλέον κατάλληλη καθώς άλλωστε δεν μας ενδιέφερε η συλλογή πληροφοριών σχετικά με τις απόψεις των πολιτών άλλων χωρών στο ζήτημα της ηλεκτρονικής διακυβέρνησης, κάτι που ξεφεύγει από τα πλαίσια της εκπόνησης της παρούσας διπλωματικής εργασίας.

Τονίζουμε εδώ, ότι οι δύο διεργασίες που δημιουργήθηκαν μέσω RapidMiner είναι κατάλληλες για εφαρμογές opinion mining σε οποιαδήποτε γλώσσα, αρκεί να αλλάξουμε τη βάση δεδομένων που εισάγουμε για την εκπαίδευση του μοντέλου

και φυσικά να τροποποιήσουμε κατάλληλα τους Operators που χρησιμοποιούνται για την επεξεργασία κειμένου ώστε να ταιριάζουν καλύτερα στην κάθε γλώσσα, όπως εξηγήθηκε και παραπάνω με την τροποποίηση της διεργασίας για την αγγλική γλώσσα. Σημαντικό κρίνεται να διευκρινιστεί εδώ ο λόγος που δεν χρησιμοποιήθηκε μία βάση ταξινομημένων σχολίων πάνω σε ένα γενικότερο θέμα κατά την εκτέλεση της εφαρμογής στα ελληνικά. Αφενός σκοπός κατά την εκπόνηση της παρούσας εργασίας ήταν η συλλογή πληροφοριών σχετικά με τις απόψεις των Ελλήνων πολιτών για την ηλεκτρονική διακυβέρνηση, οπότε κρίθηκε και σκόπιμο να εκπαιδύσουμε το μοντέλο πάνω σε αυτά και μόνο τα δεδομένα, τα οποία με την πάροδο του χρόνου θα αυξάνονται και θα μπορούν να οδηγήσουν σε μία συνεχώς διευρυνόμενη βάση, και αφετέρου η μη ύπαρξη αντίστοιχης βάσης σχολιασμού που εξακριβωμένα οδηγεί σε ασφαλή συμπεράσματα και χρησιμοποιείται κατά κόρον σε αντίστοιχες εφαρμογές, όπως συνέβη με τη βάση που εντοπίσαμε για την αγγλική γλώσσα. Το μοναδικό αρνητικό σημείο κατά την επιλογή αυτή είναι το γεγονός ότι το πλήθος των σχολίων που συγκεντρώθηκαν υστερεί για την εκπαίδευση του μοντέλου και οδηγεί σε μειωμένη ακρίβεια.

## 5.4.2 Σύγκριση ακρίβειας αποτελεσμάτων

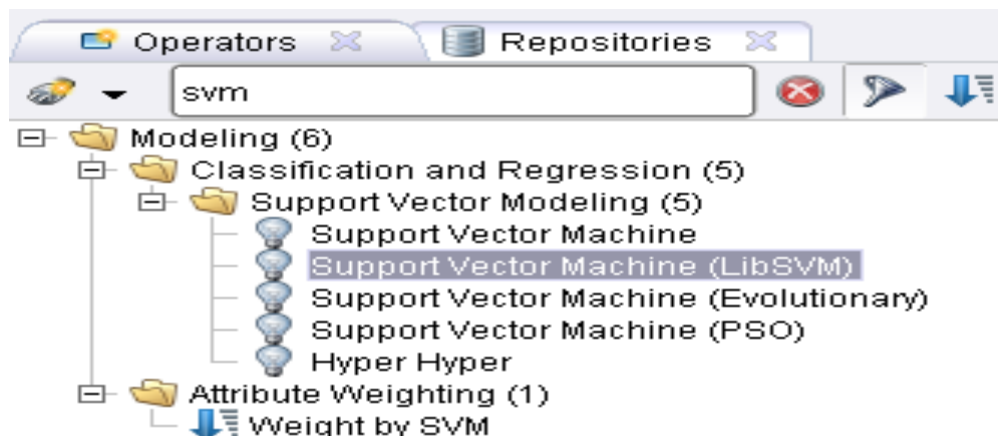
Επανερχόμαστε στην πρώτη διεργασία που δημιουργήσαμε και εκτελώντας την, το πρόγραμμα ως αποτέλεσμα επιστρέφει μία πρόβλεψη της ακρίβειας των αποτελεσμάτων που θα έχουμε κατά την εφαρμογή του μοντέλου σε μη ταξινομημένα σχόλια. Έτσι, κατά την εκτέλεση της εν λόγω διεργασίας για την ελληνική γλώσσα και με βάση εκπαίδευσης τα σχόλια που συγκεντρώθηκαν περί της ηλεκτρονικής διακυβέρνησης, τα αποτελέσματα είναι τα ακόλουθα:



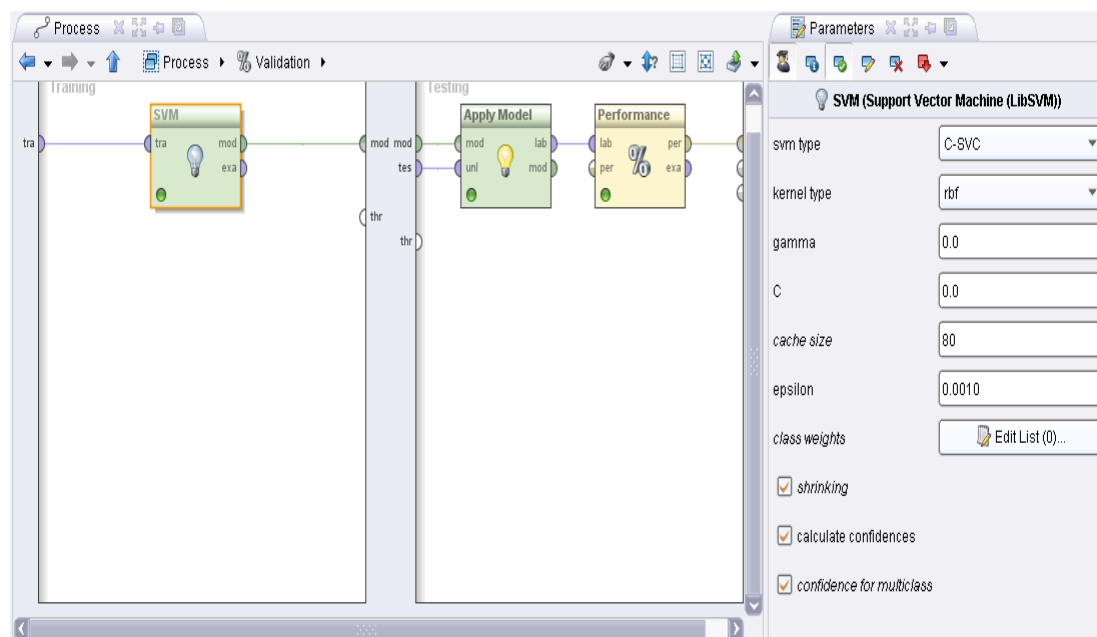
	true neg	true pos	class precision
pred. neg	137	54	71.73%
pred. pos	49	101	67.33%
class recall	73.66%	65.16%	

Παρατηρούμε ότι η ακρίβεια των αποτελεσμάτων ισούται με 69,78%, με τα θετικά να αγγίζουν το ποσοστό 67,33% και τα αρνητικά 71,73%. Η καλύτερη κατηγοριοποίηση όσον αφορά τα αρνητικά σχόλια οφείλεται στο γεγονός ότι ήταν περισσότερα σε αριθμό κατά την εισαγωγή τους ως βάσης εκπαίδευσης του μοντέλου που θα εφαρμοστεί.

Ακολουθεί η ίδια πρόβλεψη για την εφαρμογή που δημιουργήθηκε στην αγγλική γλώσσα. Προσθέτουμε ότι μετά από δοκιμές για την επίτευξη όσο το δυνατόν μεγαλύτερης ακρίβειας αποτελεσμάτων, για την συγκεκριμένη εφαρμογή αντικαταστήσαμε τον αλγόριθμο κατηγοριοποίησης: «Naïve Bayes» με τον αλγόριθμο: «Support Vector Machine(LibSVM)». Ο συγκεκριμένος αλγόριθμος κατηγοριοποίησης αποτελεί μία από τις πολλές SVM υλοποιήσεις που είναι διαθέσιμες. Οι Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines, SVM), που αναπτύχθηκαν από τον Vladimir Vapnik, αποτελούν σήμερα μία εκ των οικογενειών αλγορίθμων επιβλεπόμενης μάθησης (supervised learning) που χρησιμοποιούνται ευρέως σε μία μεγάλη γκάμα προβλημάτων κατάταξης. Η χρήση τους είναι ιδανική για την κατηγοριοποίηση δεδομένων που ανήκουν σε δύο κατηγορίες, για παράδειγμα μεταξύ θετικών και αρνητικών όπως συμβαίνει εν προκειμένω, και έχουν το πλεονέκτημα να λειτουργούν καλύτερα όταν εισάγουμε μεγαλύτερο πλήθος πληροφοριών. Ο εν λόγω Operator βρίσκεται στην ομάδα: «Modeling» - «Classification and Regression» -«Support Vector Modeling» και κατά την προσθήκη του ο χρήστης καλείται να διαλέξει ανάμεσα σε 4 βασικούς πυρήνες: γραμμικός, πολυωνυμικός, RBF, σιγμοειδής. Κρατάμε την επιλογή RBF, που είναι και αυτή που συνιστάται από τους κατασκευαστές του αλγορίθμου, καθώς και τις επόμενες παραμέτρους τις αφήνουμε ως έχουν στις προκαθορισμένες τους τιμές από το πρόγραμμα. Για να γίνει καλύτερα κατανοητό παραθέτουμε και τα αντίστοιχα screenshots:







Στη συνέχεια παρατίθεται η πρόβλεψη ακρίβειας του RapidMiner σχετικά με την κατηγοριοποίηση των δεδομένων, τόσο συνολικά όσο και ξεχωριστά για την ομάδα θετικών και αρνητικών σχολίων:

Table / Plot View | Text View | Annotations

Criterion Selector

- accuracy
- precision
- recall
- AUC (optimistic)
- AUC
- AUC (pessimistic)

Multiclass Classification Performance | Annotations

Table View | Plot View

**accuracy: 82.30% +/- 2.48% (mikro: 82.30%)**

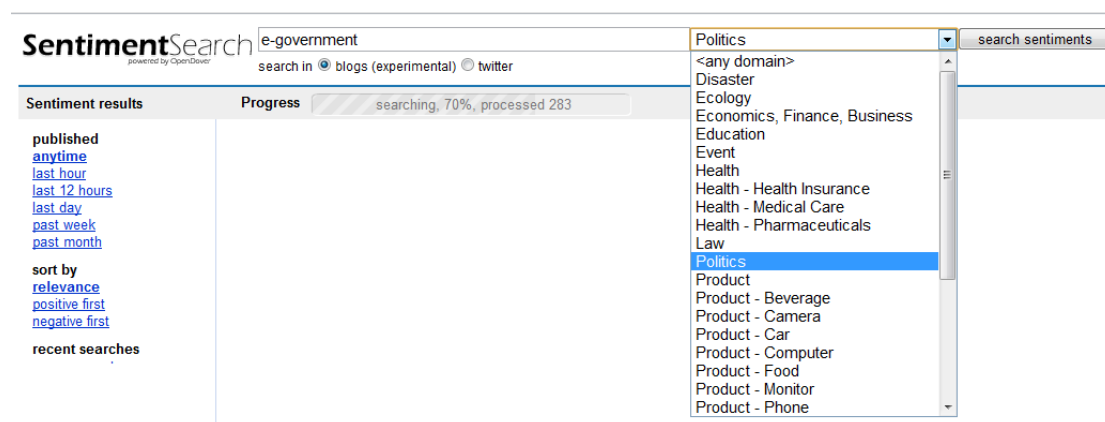
	true neg	true pos	class precision
pred. neg	822	176	82.36%
pred. pos	178	824	82.24%
class recall	82.20%	82.40%	

Όπως φαίνεται, η συνολική ακρίβεια της πρόβλεψης αγγίζει το 82,3%, ποσοστό σίγουρα ικανοποιητικό αν κρίνει κανείς από τις μέγιστες επιδόσεις αντίστοιχων μοντέλων τόσο θεωρητικά όσο και στην πράξη. Τα αντίστοιχα ποσοστά για κατηγοριοποίηση σε θετικό και αρνητικό ανέρχονται σε 82,24% και 82,36%. Η ελάχιστη αυτή απόκλιση οφείλεται στο γεγονός ότι το μοντέλο έχει εκπαιδευτεί πάνω σε ισάριθμα θετικά και αρνητικά σχόλια. Ανακεφαλαιώνοντας ως εδώ, τα παραπάνω στοιχεία επιβεβαιώνουν την ορθότερη και αποτελεσματικότερη χρήση της εφαρμογής μας όταν η γλώσσα που θα χρησιμοποιηθεί είναι η αγγλική έναντι της ελληνικής. Άλλωστε, σε αυτό συνηγορεί και η εισαγωγή αριθμητικώς περισσότερων δεδομένων, που θα αποτελέσουν τη βάση εκπαίδευσης ώστε αργότερα να επέλθει η σωστή κατηγοριοποίηση σε νέα δεδομένα.

## 5.5 Παράθεση εφαρμογών αυτόματης εξόρυξης γνώμης

Συνοψίζοντας το παρόν κεφάλαιο, παραθέτουμε κάποιες από τις πλέον ενδιαφέρουσες εφαρμογές στο πεδίο αυτόματης εξόρυξης συναισθήματος από κοινωνικά μέσα, οι οποίες είναι διαθέσιμες δωρεάν διαδικτυακά. Στα πλεονεκτήματα των εφαρμογών αυτών συγκαταλέγεται το γεγονός ότι είναι εύχρηστες καθώς δεν απαιτούν κανενός είδους επεξεργασία από τη χρήστη και καμία εξειδικευμένη γνώση κατά τη χρήση τους. Τα αποτελέσματα επιστρέφονται άμεσα και έτσι μπορούν να χρησιμοποιηθούν κυρίως για μία αρχική εξόρυξη της αναλογίας θετικών και αρνητικών απόψεων που κυριαρχούν σε ένα θέμα. Φυσικά, δεν μπορούν να επεξεργαστούν μεγάλο σε όγκο πλήθος πληροφορίας και για κάτι τέτοιο συνίσταται η χρήση μιας πιο εξειδικευμένης εφαρμογής, όπως αυτής που δημιουργήθηκε μέσω RapidMiner.

Ξεκινάμε με την online εφαρμογή που είναι διαθέσιμη μέσω της εξής ιστοσελίδας: <http://www.sentimentsearch.nl>. Πληκτρολογούμε στο πεδίο αναζήτησης το θέμα που μας ενδιαφέρει και επιλέγουμε εάν θέλουμε η αναζήτηση να γίνει πάνω σε απόψεις όπως έχουν διατυπωθεί μέσω ιστολογίων (blogs) είτε σε απόψεις που έχουν διατυπωθεί στην δημοφιλή πλατφόρμα κοινωνικής δικτύωσης, το twitter. Επιπλέον, μπορούμε να επιλέξουμε την κύρια κατηγορία θέματος με την οποία διαπραγματεύονται τα μέσα στα οποία θα γίνει η αναζήτηση, ώστε να έχουμε όσο το δυνατόν πιο συναφή με το αντικείμενό μας αποτελέσματα. Η διαφορά της εν λόγω μηχανής αναζήτησης έναντι των συνηθισμένων είναι ότι επιστρέφει τα σχετικά με τα κριτήριά μας αποτελέσματα, έχοντας κατηγοριοποιήσει αυτόματα το περιεχόμενό τους σε θετικό και αρνητικό. Άλλωστε, όπως προϋποθέτει και ο τίτλος της πρόκειται για μηχανή αναζήτησης συναισθήματος. Μάλιστα, δίνεται και σχετική βαθμολόγηση ανάλογα με το πόσο έντονα θετική ή αρνητική είναι η άποψη που διατυπώνεται. Κάποιες παραπάνω επιλογές που είναι διαθέσιμες κατά την εν λόγω αναζήτηση είναι η ημερομηνία δημοσίευσης του άρθρου ή του σχολίου αντίστοιχα και η ταξινόμηση ανάλογα είτε με τη σχετικότητα του θέματος είτε ανάλογα με το συναίσθημα του περιεχομένου. Παρακάτω δίνεται το αντίστοιχο screenshot κατά την εισαγωγή των παραμέτρων αναζήτησης:



The screenshot shows the SentimentSearch application interface. The search bar contains the text "e-government". Below the search bar, there are options for "search in" with "blogs (experimental)" selected and "twitter" as an alternative. A "search sentiments" button is visible on the right. The search progress is indicated as "searching, 70%, processed 283". The search results are categorized by domain, with "Politics" selected. The search results are sorted by "relevance".

Βέβαια, η συγκεκριμένη εφαρμογή μπορεί να χρησιμοποιηθεί και για την αναζήτηση ενός ελληνικού όρου και κατά συνέπεια να επιστραφούν αναρτήσεις που είναι γραμμένες στην ελληνική γλώσσα. Διαπιστώνεται όμως, ότι η ακρίβεια κατηγοριοποίησης δεν κυμαίνεται στα επιθυμητά επίπεδα όπως συμβαίνει κατά τη χρήση αγγλικών όρων. Ακολουθεί ένα απόσπασμα των αποτελεσμάτων που επιστράφηκαν κατά την αναζήτηση του όρου: e-government σε ιστολόγια πολιτικού περιεχομένου:

← → ↻ 🏠 [www.sentimentsearch.nl/search?query=e-government](http://www.sentimentsearch.nl/search?query=e-government)

**SentimentSearch** powered by OpenDover e-government Politics search sentiments

search in  blogs (experimental)  twitter results  5  10  20

**Sentiment results** Progress done, found 20 of 20

**published**  
[anytime](#)  
[last hour](#)  
[last 12 hours](#)  
[last day](#)  
[past week](#)  
[past month](#)

**sort by**  
[relevance](#)  
[positive first](#)  
[negative first](#)

**recent searches**  
[e-government](#)

**Moroccan government delegation familiarized with Dubai's leading ...**  
**Politics** 19 Jan 2011 by unknown, ExcitingAds! AME Info Latest News  
+2 A Moroccan government delegation has recently visited Dubai eGovernment Department to be familiarized with Dubai Government's leading eTransformation experience through the applications of the...  
<http://www.ameinfo.com/254224.html>

**Tunisia: Blog Declines Award Sponsored by Bahrain Government ...**  
**Politics** 09 May 2011 by admin, Current Affairs  
0 Since the eContent Award is officially sponsored by the Bahraini eGovernment Authority and since Nawaat's position has always been to avoid any form of support from government organisations, we reject this award in protest against the ...  
[http://pakistanvoices.com/Current\\_affairs/tunisia-blog-declines-award-sponsored-by-bahrain-government/](http://pakistanvoices.com/Current_affairs/tunisia-blog-declines-award-sponsored-by-bahrain-government/)

**Russian government must go digital by 2015 — RT**  
**Politics** 23 Dec 2009 by unknown, RT  
+2 And by the way, Bogdanov, I do believe American taxpayers contributed quite few dollars to Russia since 1991, too, but Russia has not used them to set up e-government. In fact, no one seems to know what Russia used them for at all – the ...  
<http://rt.com/news/sci-tech/russian-government-electronic-digital/>

**CEDEM11 Day 1 e Digital Government & Society**  
**Politics** 05 May 2011 by jsduk, Digital Government & Society  
-2 Another factor is the financial crisis (Duisborg for instance is totally skint) that raises the question whether eGovernment can be a problem solver for all this. Twitter Revolution in Duisburg? :). Using a case study analyzing the ...  
<http://digitalgovernment.wordpress.com/2011/05/05/cedem11-day-1/>

**E - Government - Singapore experience.**  
**Politics** 23 Mar 2011 by singapore, world story  
+2 The catechism is: What is E-government? The United Nations defines e-Government as "Permanent allegation by government to advance the accord amid the clandestine aborigine and the attainable area through enhanced, cost-effective and ...  
<http://uwin.com/government-singapore-experience/>

Συνεχίζουμε με την παράθεση της εφαρμογής που είναι διαθέσιμη διαδικτυακά μέσω του ιστοτόπου: <http://www.opendover.nl/>. Η εν λόγω εφαρμογή επεξεργάζεται ένα κείμενο που εισάγουμε και εντοπίζει τις λέξεις που εκφράζουν συναίσθημα καθώς και την ένταση του συναισθήματος αυτού. Ο χρήστης μπορεί να επιλέξει την οντότητα πάνω στην οποία θα αναζητηθεί κάθε είδους χαρακτηρισμός εντός του κειμένου ή απλά να παραθέσει το κείμενο και η εφαρμογή θα επιστρέψει ως αποτέλεσμα τις λέξεις- κλειδιά που σχετίζονται με κάποιου είδους κατηγορία, όπως για παράδειγμα οικονομία, πολιτική, υγεία, εκπαίδευση, σε συνδυασμό με λέξεις που εκφράζουν κάποιου είδους επιδοκιμασία ή δυσαρέσκεια. Παρακάτω δίνεται το αντίστοιχο screenshot ώστε να φανεί καλύτερα το αποτέλεσμα της συγκεκριμένης εφαρμογής. Στο παράδειγμά μας, απλά εισάγαμε μία κριτική, χωρίς να ορίσουμε ποια είναι η οντότητα πάνω στην οποία μας ενδιαφέρει να επιστραφούν αποτελέσματα:

**Input Parameters**

- Take a piece of text, for example, a review of a hotel. Paste this piece of text in the text area.
- Define an object to search in the text (optional). An object is a word describing the entity on which you would like to receive sentiments. It should exist in the text itself. For a hotel review, a hotel name would be a good choice for the object.
- Press the button "Process" and see the sentiments, the severity of the emotion, domain words OpenDover has found.
- An example of the RDF XML output is also given.

**Story**

Automating public services with technology can save countries a significant amount of money, as well as make citizens feel connected to their government.

**Results**

Economics, Finance, Business ☆☆☆☆ Politics ☆☆☆☆

Automating public services with technology can save countries a significant amount of money, as well as make citizens feel connected to their government.

Sentiment: significant  
Appreciation, +2

Παρατηρούμε, λοιπόν, ότι η εφαρμογή έχει επισημάνει αυτόματα τις λεκτικές οντότητες που παραπέμπουν συνήθως σε μία δεδομένη κατηγορία, καθώς επίσης και τη λέξη που περιέχεται στο κείμενο και σχετίζεται με την έκφραση άποψης θετικής ή αρνητικής. Η λέξη χρωματίζεται αντίστοιχα είτε με πράσινο ή με κόκκινο χρώμα, όπως άλλωστε συνηθίζεται, και δίνεται σε αυτή η αντίστοιχη βαθμολογία ανάλογα με το πόσο έντονο είναι το συναίσθημα που εκφράζεται μέσω αυτής. Σημειώνουμε ότι η εν λόγω εφαρμογή, προς το παρόν τουλάχιστον, δεν είναι διαθέσιμη στην ελληνική γλώσσα καθώς αντίστοιχες ελληνικές λέξεις δεν αναγνωρίζονται και δεν βαθμολογούνται από το σύστημα. Το γεγονός αυτό συνηγορεί στην άποψη ότι όσον αφορά εφαρμογές opinion mining ακόμα δεν υπάρχει διαδικτυακά έτοιμο υλικό προς χρήση όταν πρόκειται για την ελληνική γλώσσα.

Κλείνοντας την ενότητα αυτή, θα θέλαμε να αναφερθούμε στις πολυάριθμες εφαρμογές sentiment analysis που είναι διαθέσιμες διαδικτυακά και είναι εστιασμένες στα σχόλια των χρηστών μέσω της πλατφόρμας κοινωνικής δικτύωσης: <http://www.twitter.com>. Η εν λόγω πλατφόρμα κερδίζει όλο και περισσότερους χρήστες και η λειτουργία της είναι συνυφασμένη με τη δημοσίευση από μέρους των χρηστών μικρών και άμεσων κειμένων, ενώ τα θέματα σχολιασμού ποικίλουν και καλύπτουν ουσιαστικά κάθε είδους κατηγορία. Η δυνατότητα αλληλεπίδρασης μεταξύ των χρηστών και άμεσης ανταλλαγής απόψεων κυρίως πάνω στα θέματα της επικαιρότητας έχουν καταστήσει την εν λόγω πλατφόρμα μια ευρέως αναπτυσσόμενη πηγή άντλησης της κοινής γνώμης των πολιτών σε διάφορα ζητήματα. Βέβαια, η ηλεκτρονική διακυβέρνηση δε θα μπορούσε να λείπει από τις κατηγορίες θεμάτων που έχουν δημιουργηθεί και συζητούνται μέσω της δημοφιλούς αυτής «ψηφιακής κοινότητας». Παρατηρείται για άλλη μία φορά ότι χρήστες από την Ελλάδα δεν είναι τόσο ενεργοί στο συγκεκριμένο ζήτημα όπως συμβαίνει με χρήστες που προέρχονται από χώρες όπου η ηλεκτρονική διακυβέρνηση έχει διεισδύσει περισσότερο.

Η παρακάτω εφαρμογή λειτουργεί ως μηχανή αναζήτησης και επιστρέφει αποτελέσματα που αντλούνται αποκλειστικά μέσω των σχολίων των χρηστών του twitter, τα οποία και ταξινομεί σε θετικά και αρνητικά. Ο ιστότοπος της εν λόγω εφαρμογής είναι: <http://twittersentiment.appspot.com> και επιστρέφονται τα πιο πρόσφατα αποτελέσματα που εντοπίζονται. Ένα χαρακτηριστικό screenshot είναι το

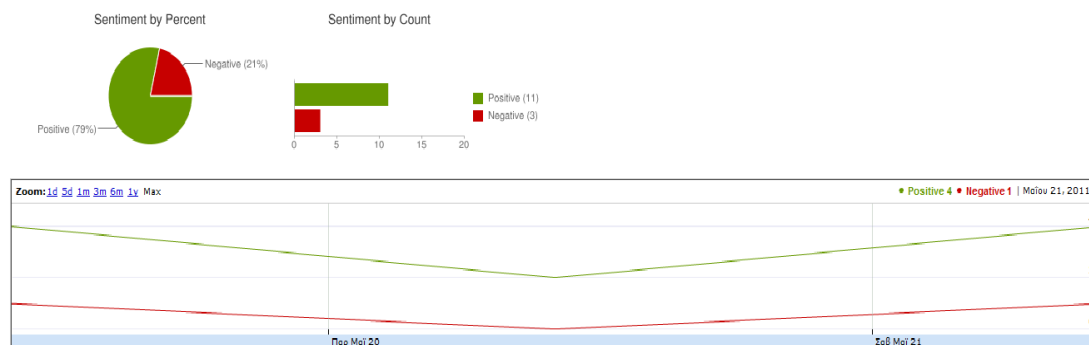
παρακάτω, στο οποίο επισημαίνεται η αναλογία των κυριότερων θετικών και αρνητικών σχολίων κατά τη διάρκεια των δύο τελευταίων ημερών:

## Twitter Sentiment

Type in a word and we'll highlight the good and the bad

egov  [Save this search](#)

### Sentiment analysis for egov



Για τον αναγνώστη που αναζητά περισσότερες ιστοσελίδες με online εφαρμογές sentiment analysis ή λογισμικά που διατίθενται είτε δωρεάν ή όχι και προορίζονται για την συγκεκριμένη χρήση, παρατίθεται ο επόμενος πίνακας που περιέχει μία λίστα αυτών, μία σύντομη περιγραφή και κάποια από τα βασικά τους χαρακτηριστικά:

Όνομα	Περιγραφή	URL	Sentiment Analysis που εστιάζεται το Twitter	Διαθέσιμο online δωρεάν
TweetFeel	Η κατηγοριοποίηση γίνεται μέσω λέξεων-κλειδιών στην ελεύθερη έκδοση της εφαρμογής ενώ στην επαγγελματική χρησιμοποιούνται αλγόριθμοι μηχανικής μάθησης.	<a href="http://www.tweetfeel.com/">http://www.tweetfeel.com/</a>	NAI	NAI
Twendz	Ελεύθερη και επαγγελματική έκδοση	<a href="http://twendz.waggeneredstrom.com/">http://twendz.waggeneredstrom.com/</a>	NAI	NAI
PeopleBrowsr	Sentiment analysis εστιασμένη σε κοινωνικά μέσα με πολλαπλές επιλογές αναζήτησης	<a href="http://brand.s.peoplebrowsr.com/">http://brand.s.peoplebrowsr.com/</a>	NAI	NAI

Groubal	Κριτικές καταναλωτών πάνω σε προϊόντα που κατηγοριοποιούνται μέσω του αλγορίθμου Naïve Bayes	<a href="http://www.groubalcsi.com/">http://www.groubalcsi.com/</a>	NAI	NAI
Pulse of the Nation	Αυτόματη ανάλυση της διάθεσης που επικρατεί σε κάθε πολιτεία των Η.Π.Α όπως αυτή καταγράφεται μέσω Twitter.	<a href="http://www.ccs.neu.edu/home/amislove/twittemod/">http://www.ccs.neu.edu/home/amislove/twittemod/</a>	NAI	NAI
Twitrratr	Εισάγοντας μία λέξη-κλειδί επιστρέφονται τα σχόλια ταξινομημένα βάσει συναισθήματος σε θετικά, αρνητικά και ουδέτερα	<a href="http://twitrratr.com/">http://twitrratr.com/</a>	NAI	NAI
Brand Bowl	Sentiment analysis σε διαφημίσεις εμπορικών προϊόντων	<a href="http://brandbowl2011.com/">http://brandbowl2011.com/</a>	NAI	NAI
EmoLib	Προσδιορίζει το εύρος συναισθήματος που εκφράζεται σε ένα κείμενο	<a href="http://dtmin.redis.housing.salle.url.edu:8080/EmoLib/">http://dtmin.redis.housing.salle.url.edu:8080/EmoLib/</a>	OXI	NAI
OpinionCrawl	Online sentiment analysis για επιλεγμένα γεγονότα επικαιρότητας, εταιρείες, προϊόντα, προσωπικότητες	<a href="http://opinioncrawl.com/">http://opinioncrawl.com/</a>	OXI	NAI
Python NLTK Demos by Streamhacker	Online sentiment analysis και κάποιες επιπλέον σχετικές εφαρμογές	<a href="http://text-processing.com/demo/">http://text-processing.com/demo/</a>	OXI	NAI
RankSpeed	Αναζήτηση με βάση το συναίσθημα που περιέχεται	<a href="http://rankspeed.com">http://rankspeed.com</a>	OXI	NAI
OPFINE	Sentiment analysis για εμπορικά προϊόντα	<a href="http://www.jane16.com/">http://www.jane16.com/</a>	OXI	NAI
Moodbase	Το περιεχόμενο χαρακτηρίζεται κάθε φορά με μία μοναδική λέξη που εκφράζει συναίσθημα	<a href="http://moodbase.com/">http://moodbase.com/</a>	OXI	NAI

General Sentiment	Λογισμικό για Sentiment analysis πάνω σε διάφορα θέματα. Η εφαρμογή προορίζεται κυρίως για επαγγελματικούς σκοπούς καθώς δεν είναι διαθέσιμη δωρεάν, παρά μόνο ένα demo.	<a href="http://www.generalsentiment.com/">http://www.generalsentiment.com/</a>	OXI	OXI
Affect Labs	Λογισμικό για Sentiment analysis πάνω σε διάφορα θέματα. Η εφαρμογή προορίζεται κυρίως για επαγγελματικούς σκοπούς καθώς δεν είναι διαθέσιμη δωρεάν, παρά μόνο ένα demo.	<a href="http://www.affectlabs.com/">http://www.affectlabs.com/</a>	OXI	OXI
Sysomos	Λογισμικό για Sentiment analysis πάνω σε διάφορα θέματα. Η εφαρμογή προορίζεται κυρίως για επαγγελματικούς σκοπούς καθώς δεν είναι διαθέσιμη δωρεάν, παρά μόνο ένα demo.	<a href="http://www.sysomos.com/">http://www.sysomos.com/</a>	OXI	OXI
Sentimatrix	Λογισμικό για Sentiment analysis πάνω σε διάφορα θέματα. Η εφαρμογή προορίζεται κυρίως για επαγγελματικούς σκοπούς καθώς δεν είναι διαθέσιμη δωρεάν, παρά μόνο ένα demo.	<a href="http://www.sentimatrix.com/">http://www.sentimatrix.com/</a>	OXI	OXI
Lexalytics	Λογισμικό για Sentiment analysis πάνω σε διάφορα θέματα. Η εφαρμογή προορίζεται κυρίως για επαγγελματικούς σκοπούς καθώς δεν είναι διαθέσιμη δωρεάν, παρά μόνο ένα demo.	<a href="http://www.lexalytics.com/">http://www.lexalytics.com/</a>	OXI	OXI

Radian6	Λογισμικό για Sentiment analysis πάνω σε διάφορα θέματα. Η εφαρμογή προορίζεται κυρίως για επαγγελματικούς σκοπούς καθώς δεν είναι διαθέσιμη δωρεάν, παρά μόνο ένα demo.	<a href="http://www.radian6.com/">http://www.radian6.com/</a>	OXI	OXI
Sentiment Metrics	Λογισμικό για Sentiment analysis πάνω σε διάφορα θέματα. Η εφαρμογή προορίζεται κυρίως για επαγγελματικούς σκοπούς καθώς δεν είναι διαθέσιμη δωρεάν, παρά μόνο ένα demo.	<a href="http://www.sentimentmetrics.com/">http://www.sentimentmetrics.com/</a>	OXI	OXI
I-Sieve Technologies Ltd.	Λογισμικό για Sentiment analysis πάνω σε διάφορα θέματα. Η εφαρμογή προορίζεται κυρίως για επαγγελματικούς σκοπούς καθώς δεν είναι διαθέσιμη δωρεάν, παρά μόνο ένα demo.	<a href="http://www.i-sieve.com/">http://www.i-sieve.com/</a>	OXI	OXI
SentiStrength	Λογισμικό για Sentiment analysis πάνω σε διάφορα θέματα. Η εφαρμογή προορίζεται κυρίως για επαγγελματικούς σκοπούς καθώς δεν είναι διαθέσιμη δωρεάν, παρά μόνο ένα demo.	<a href="http://sentistrength.wlv.ac.uk/">http://sentistrength.wlv.ac.uk/</a>	OXI	OXI
Saplo	Λογισμικό για Sentiment analysis πάνω σε διάφορα θέματα. Η εφαρμογή προορίζεται κυρίως για επαγγελματικούς σκοπούς καθώς δεν είναι διαθέσιμη δωρεάν, παρά μόνο ένα demo.	<a href="http://www.saplo.com/">http://www.saplo.com/</a>	OXI	OXI



Social Mention	Online εφαρμογή για sentiment analysis σε μέσα κοινωνικής δικτύωσης	<a href="http://www.socialmention.com">http://www.socialmention.com</a>	NAI	NAI
Viral Heat	Λογισμικό για Sentiment analysis πάνω σε διάφορα θέματα. Η εφαρμογή προορίζεται κυρίως για επαγγελματικούς σκοπούς καθώς δεν είναι διαθέσιμη δωρεάν, παρά μόνο ένα demo.	<a href="http://www.viralheat.com/">http://www.viralheat.com/</a>	OXI	OXI
Autonomy Idol	Λογισμικό για Sentiment analysis πάνω σε διάφορα θέματα. Η εφαρμογή προορίζεται κυρίως για επαγγελματικούς σκοπούς καθώς δεν είναι διαθέσιμη δωρεάν, παρά μόνο ένα demo.	<a href="http://www.autonomy.com/content/functionality/idol-functionality-sentiment/index.en.html">http://www.autonomy.com/content/functionality/idol-functionality-sentiment/index.en.html</a>	OXI	OXI
Neputation	Λογισμικό για Sentiment analysis πάνω σε διάφορα θέματα. Η εφαρμογή προορίζεται κυρίως για επαγγελματικούς σκοπούς καθώς δεν είναι διαθέσιμη δωρεάν, παρά μόνο ένα demo.	<a href="http://www.neputation.com">http://www.neputation.com</a>	OXI	OXI

## ***Κεφάλαιο 6. Συμπεράσματα και Προοπτικές***

---

## 6.1 Συμπεράσματα

Όπως προκύπτει μέσα από τις σελίδες της παρούσας διπλωματικής εργασίας, η μετάβαση στην εποχή της ηλεκτρονικής διακυβέρνησης αποτελεί ένα από τα πιο σύγχρονα και σπουδαία εγχειρήματα της εποχής μας. Η συνειδητοποίηση τόσο από πλευράς κυβερνήσεων όσο από πλευράς πολιτών της αναγκαιότητας της εν λόγω μετάβασης είναι καθοριστική καθώς μόνο οφέλη μπορούν να αποκομισθούν. Η προσπάθεια για την βελτιστοποίηση και την μέγιστη δυνατή επιτυχία της ηλεκτρονικής διακυβέρνησης πρέπει να καταβληθεί αμφίδρομα και από τις δύο πλευρές, στα πλαίσια μιας υγιούς σχέσης συνεργασίας, συμμετοχής και εν τέλει διαδραστικότητας.

Και σε αυτή την ηλεκτρονική συμμετοχή και διαδραστικότητα βρίσκει πάτημα η τεχνική της αυτόματης εξόρυξης γνώμης/ συναισθήματος μέσω διαδικτυακών πηγών. Με έμφαση πάντα στα κοινωνικά μέσα και εν έτη 2011, ο Παγκόσμιος Ιστός αποτελεί την αμεσότερη πύλη επαφής με την κοινή γνώμη και τον παλμό της επικαιρότητας. Οι χρήστες πλέον νιώθουν έντονη την ανάγκη να μοιραστούν σκέψεις, απόψεις, προτάσεις και προβληματισμούς, κάτι το οποίο κατά κόρον συμβαίνει αυθόρμητα και όχι εσκεμμένα. Αυτός ο «πλούτος» που είναι διαθέσιμος ελεύθερα στο διαδίκτυο πρέπει να αξιοποιηθεί ώστε να οδηγηθούμε αρχικά σε υπηρεσίες ηλεκτρονικής διακυβέρνησης κατάλληλα προσαρμοσμένες στις ανάγκες και τις απαιτήσεις του χρήστη και εν τέλει στη βέλτιστη λειτουργία του κρατικού μηχανισμού αλλά και την εξοικονόμηση πόρων.

Εδώ εμπίπτει η μεγαλύτερη πρόκληση αλλά και ο προβληματισμός καθώς οι πληροφορίες δεν είναι πάντα διαθέσιμες στην κατάλληλη μορφή για επεξεργασία, αλλά κυρίως το πρόβλημα δημιουργείται από τον τεράστιο όγκο των δεδομένων αυτών. Εύκολα μπορεί ο καθένας να κατανοήσει πόσο χρονοβόρα και ταυτόχρονα επίπονη είναι η διαδικασία εύρεσης αρχικά των διατυπωθέντων απόψεων πάνω σε ένα συγκεκριμένο ζήτημα μέσω των αναρίθμητων και ραγδαίως αυξανόμενων πηγών στο διαδίκτυο και έπειτα ταξινόμησης αυτών σε θετικές και αρνητικές ώστε να κατανοηθεί η κοινή γνώμη επί του ζητήματος. Οι παραπάνω λόγοι οδήγησαν τους ερευνητές κατά την τελευταία δεκαετία στην δημιουργία αλγορίθμων και τεχνικών που θα έχουν θεωρητικά την δυνατότητα να κατηγοριοποιούν αυτόματα το κείμενο βάση του συναισθήματος που περιέχεται και γενικότερα βάση της άποψης, θετικής ή αρνητικής.

Ενώ οι αλγόριθμοι και τα διάφορα λογισμικά προγράμματα ξεπερνούν τον ανθρώπινο νου σε ταχύτητα και αποτέλεσμα όταν πρόκειται για ποσοτικές μεταβλητές και υπολογιστικές διαδικασίες, στην περίπτωση του opinion mining, που στηρίζεται κυρίως σε ποιοτικές μεταβλητές, η ακρίβεια της διαδικασίας κυμαίνεται αναμφίβολα σε χαμηλότερα επίπεδα. Οι παράγοντες που διαφοροποιούν την ακρίβεια της διαδικασίας αναλύθηκαν στην παρούσα εργασία και γίνονται εύκολα αντιληπτοί αν διανοηθεί κανείς ότι πρόκειται για μία διαδικασία πολλές φορές δύσκολη ακόμα και για τον κοινό ανθρώπινο νου.

Με τη δημιουργία εφαρμογής κατηγοριοποίησης γνώμης, στα πλαίσια της εργασίας, προσπαθήσαμε να προσεγγίσουμε όσο το δυνατόν ακριβέστερα και πληρέστερα τον ζητούμενο στόχο. Θεωρητικά, πρόκειται για μία διαδικασία που ακολουθεί τα απαραίτητα βήματα και περιέχει τους κατάλληλους αλγορίθμους για

τον σκοπό αυτό, στην πράξη όμως η ακρίβεια απέχει από τα επιθυμητά επίπεδα για λόγους που αναλύθηκαν διεξοδικά.

## 6.2 Προοπτικές

Ολοκληρώνοντας την παρούσα διπλωματική εργασία, θα θέλαμε να παρουσιάσουμε κάποιες προτάσεις που θα μπορούσαν να συνεισφέρουν στην βελτίωση της λειτουργίας του μοντέλου της ηλεκτρονικής διακυβέρνησης.

- Απαραίτητη κρίνεται η επένδυση σε αξιόπιστα και σύγχρονα πληροφοριακά συστήματα μέσω των οποίων η κυβέρνηση αρχικά θα επεξεργαστεί μεγάλο όγκο πληροφορίας σχετικά με την άποψη των Ελλήνων πολιτών περί της ηλεκτρονικής διακυβέρνησης γενικότερα, αλλά κυρίως ειδικότερα σχετικά με το σχεδιασμό συγκεκριμένων υπηρεσιών. Με αυτό τον τρόπο θα εξοικονομηθούν πολύτιμοι πόροι, καθώς μελλοντικά η σχεδίαση οποιασδήποτε ηλεκτρονικά παρεχόμενης υπηρεσίας θα είναι προσαρμοσμένη όσο το δυνατόν περισσότερο ικανοποιητικά στις ανάγκες και τις απαιτήσεις των χρηστών. Τα εν λόγω συστήματα αργότερα θα αξιοποιηθούν για την πλήρη ηλεκτρονική διεκπεραίωση των εν λόγω υπηρεσιών. Ο σχεδιασμός μίας ηλεκτρονικής υπηρεσίας άνευ γνώσης και αξιολόγησης της άποψης των πολιτών μόνο δαπανηρός μπορεί να αποβεί.
- Η μέση διαθεσιμότητα των διαδικτυακών δημόσιων υπηρεσιών στη χώρα μας κρίνεται απογοητευτική, συγκρινόμενη πάντα με τους αντίστοιχους δείκτες στην ΕΕ, καταλαμβάνοντας συγκεκριμένα την τελευταία θέση στην σχετική κατάταξη. Η πλήρης ηλεκτρονική διεκπεραίωση τουλάχιστον των 20 βασικών δημόσιων υπηρεσιών προς πολίτες και επιχειρήσεις πρέπει να είναι άμεση όπως επίσης και η διαμόρφωση συγκεκριμένων διαδικτυακών κυβερνητικών πλατφόρμων με απώτερο στόχο την ευχρηστία και την καλύτερη αναζήτηση πληροφοριών. Αναμφίβολα, θα αποτελέσει καίριο βήμα για την αύξηση του ποσοστού των Ελλήνων πολιτών που κάνουν χρήση των διαδικτυακών υπηρεσιών, το οποίο προς το παρόν κυμαίνεται σε εξίσου χαμηλά επίπεδα, και θα συνεισφέρει στην επιτυχία του εγχειρήματος της ηλεκτρονικής διακυβέρνησης.
- Η συνεργασία των ευρωπαϊκών χωρών μέσω ενός κοινού πλαισίου για την συλλογή, ανάλυση και αξιολόγηση των απόψεων των πολιτών σχετικά με την ηλεκτρονική διακυβέρνηση θα μπορούσε να συμβάλλει τα μέγιστα για την βελτίωση αυτής. Η ύπαρξη μεγαλύτερου όγκου πληροφορίας θα μπορούσε να οδηγήσει μελλοντικά στην δημιουργία νέων και καινοτόμων υπηρεσιών.
- Τέλος και αναφορικά με τις τεχνικές αυτόματης εξόρυξης γνώμης μέσω διαδικτυακών πηγών, το πεδίο αυτό θα πρέπει να απασχολήσει περισσότερο τους Έλληνες ερευνητές μελλοντικά, κυρίως λόγω της ιδιαιτερότητας της ελληνικής γλώσσας. Η ανάπτυξη μίας βάσης θετικού και αρνητικού λεξιλογίου καθώς και αντίστοιχων εκφράσεων κρίνεται επιτακτική και ταυτόχρονα υλοποιήσιμη. Η πρόκληση έγκειται στην δημιουργία ειδικού λεξιλογίου σχολιασμού, προσαρμοσμένου κατάλληλα στα αντίστοιχα θέματα που επρόκειτο να γίνει η κατηγοριοποίηση άποψης, καθώς είναι σαφές πως

το χρησιμοποιούμενο λεξιλόγιο σχολιασμού παρουσιάζει αρκετές ιδιομορφίες από θέμα σε θέμα.

## ΒΙΒΛΙΟΓΡΑΦΙΑ

1. [http://nemertes.lis.upatras.gr/dspace/bitstream/123456789/518/1/Nimertis\\_Nasikas.pdf](http://nemertes.lis.upatras.gr/dspace/bitstream/123456789/518/1/Nimertis_Nasikas.pdf)
2. [http://el.wikipedia.org/wiki/%CE%95%CF%80%CE%BF%CF%87%CE%AE\\_%CF%84%CE%B7%CF%82\\_%CF%80%CE%BB%CE%B7%CF%81%CE%BF%CF%86%CE%BF%CF%81%CE%AF%CE%B1%CF%82](http://el.wikipedia.org/wiki/%CE%95%CF%80%CE%BF%CF%87%CE%AE_%CF%84%CE%B7%CF%82_%CF%80%CE%BB%CE%B7%CF%81%CE%BF%CF%86%CE%BF%CF%81%CE%AF%CE%B1%CF%82)
3. <http://cgi.di.uoa.gr/~sigletos/thesis/Final.pdf>
4. [http://en.wikipedia.org/wiki/Data\\_mining](http://en.wikipedia.org/wiki/Data_mining)
5. [http://el.wikipedia.org/wiki/%CE%9C%CE%B7%CF%87%CE%B1%CE%BD%CE%B9%CE%BA%CE%AE\\_%CE%BC%CE%AC%CE%B8%CE%B7%CF%83%CE%B7](http://el.wikipedia.org/wiki/%CE%9C%CE%B7%CF%87%CE%B1%CE%BD%CE%B9%CE%BA%CE%AE_%CE%BC%CE%AC%CE%B8%CE%B7%CF%83%CE%B7)
6. [http://en.wikipedia.org/wiki/Machine\\_learning](http://en.wikipedia.org/wiki/Machine_learning)
7. [http://en.wikipedia.org/wiki/List\\_of\\_machine\\_learning\\_algorithms](http://en.wikipedia.org/wiki/List_of_machine_learning_algorithms)
8. [http://en.wikipedia.org/wiki/Artificial\\_intelligence](http://en.wikipedia.org/wiki/Artificial_intelligence)
9. <http://en.wikipedia.org/wiki/Database>
10. [http://en.wikipedia.org/wiki/Natural\\_language\\_processing](http://en.wikipedia.org/wiki/Natural_language_processing)
11. [http://en.wikipedia.org/wiki/Information\\_retrieval](http://en.wikipedia.org/wiki/Information_retrieval)
12. [http://en.wikipedia.org/wiki/Information\\_extraction](http://en.wikipedia.org/wiki/Information_extraction)
13. [http://en.wikipedia.org/wiki/User\\_modeling](http://en.wikipedia.org/wiki/User_modeling)
14. <http://people.ischool.berkeley.edu/~hearsst/text-mining.html>
15. [http://en.wikipedia.org/wiki/Sentiment\\_analysis](http://en.wikipedia.org/wiki/Sentiment_analysis)
16. Kushal Dave, Steve Lawrence, and David M. Pennock. Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. In Proceedings of WWW
17. Bing Liu. *Web data mining; Exploring hyperlinks, contents, and usage data*, chapter 11: Opinion Mining. Springer, 2006.
18. Sanjiv Das and Mike Chen. Yahoo! for Amazon: Extracting market sentiment from stock message boards. In *Proceedings of the Asia Pacific Finance Association Annual Conference (APFA)*, 2001.
19. Richard M. Tong. An operational system for detecting and tracking opinions in on-line discussion. In *Proceedings of the Workshop on Operational Text Classification (OTC)*, 2001.

20. Peter Turney. Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In Proceedings of the Association for Computational Linguistics (ACL)
21. Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. Thumbs up? Sentiment classification using machine learning techniques. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*
22. <http://www.aaai.org/Library/Symposia/Spring/ss04-07.php>
23. <http://naaclhlt2010.isi.edu/>
24. [http://en.wikipedia.org/wiki/List\\_of\\_artificial\\_intelligence\\_conferences](http://en.wikipedia.org/wiki/List_of_artificial_intelligence_conferences)
25. [http://en.wikipedia.org/wiki/Annual\\_Meeting\\_of\\_the\\_Association\\_for\\_Computational\\_Linguistics](http://en.wikipedia.org/wiki/Annual_Meeting_of_the_Association_for_Computational_Linguistics)
26. R. Mihalcea, C. Banea and Janyce Wiebe (2007). Learning Multilingual Subjective Language via Cross-Lingual Projections. In Proceedings of the 45th Annual Meeting of ACL.
27. Y. A. Mathieu (2004). A Computational Semantic Lexicon of French Verbs of Emotion. In Computing Attitude and Affect in Text: Theory and Applications. Ed. James G. Shanahan, Yan Qu and Janyce Wiebe.
28. Using SentiWordNet for Multilingual Sentiment Analysis Kerstin Denecke [http://www.l3s.de/web/upload/documents/1/ICDE\\_workshop\\_9002.pdf](http://www.l3s.de/web/upload/documents/1/ICDE_workshop_9002.pdf)
29. Multilingual Corpus Development for Opinion Mining by Julia Maria Schulz , Christa Womser-hacker , Thomas M <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.180.8321>
30. Sentiment Analysis: An Overview Yelena Mejova, Computer Science Department University of Iowa, November 2009. [http://uiowa.academia.edu/YelenaMejova/Papers/241860/Sentiment\\_Analysis\\_An\\_Overview](http://uiowa.academia.edu/YelenaMejova/Papers/241860/Sentiment_Analysis_An_Overview)
31. Vasileios Hatzivassiloglou, Janyce Wiebe: Effects of Adjective Orientation and Gradability on Sentence Subjectivity. COLING 2000
32. Sentiment Analysis: Adjectives and Adverbs are better than Adjectives Alone <http://www.icwsm.org/papers/3--Benamara-Cesarano-Picariello-Reforgiato-Subrahmanian.pdf>
33. Janyce Wiebe, Theresa Wilson, Rebecca Bruce, Matthew Bell, Melanie Martin: Learning Subjective Language <http://portal.acm.org/citation.cfm?id=1105598>
34. Turney, P. D. and Littman, M. L. (2003). Measuring praise and criticism: Inference of semantic orientation from association. ACM Transactions on Information Systems

35. Dave, K., Lawrence, S., and Pennock, D. M. (2003). Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. Proceedings of the World Wide Web Conference.
36. Kanayama, H., Nisukawa, T., and Watanabe, H. (2004). Deeper sentiment analysis using machine translation technology. Proceedings of the International Conference on Computational Linguistics Web Data Mining, Bing Liu Dec 2006
37. Extracting Opinions from Blog Comments: Analysis, Design and Applications Preethi Raghavan, B. Tech Graduate Program in Computer Science and Engineering The Ohio State University 2009 <http://etd.ohiolink.edu/send-pdf.cgi/Raghavan%20Preethi.pdf?osu1259756172>
38. Lillian Lee, AAAI 2008, Invited talk, [www.cs.cornell.edu/home/llee/talks/lleeaaai08.pdf](http://www.cs.cornell.edu/home/llee/talks/lleeaaai08.pdf)
39. [http://en.wikipedia.org/wiki/Voice\\_of\\_the\\_customer](http://en.wikipedia.org/wiki/Voice_of_the_customer)
40. G. Adomavicius and A. Tuzhilin. "Using Data Mining Methods to Build Customer Profiles." IEEE Computer, vol. 34, no. 2, February 2001. <http://www.cs.pitt.edu/~mrotaru/comp/rs/Adomavicius%20IEEE%202001.pdf>
41. Sentiment Analysis of Blogs by Combining Lexical Knowledge with Text Classification <http://www.prem-melville.com/pooling-multinomials-kdd09.pdf>
42. Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews <http://acl.ldc.upenn.edu/P/P02/P02-1053.pdf>
43. AI, E-government, and Politics 2.0 Hsinchun Chen, University of Arizona
44. T. O'Reilly, "What Is Web 2.0? Design Patterns and Business Models for the Next Generation of Software," presented at O'Reilly Media Web 2.0 Conf., 2004; <http://oreilly.com/web2/archive/what-is-web-20.html>.
45. Reddick C, 2004, "Citizen interaction with eGovernment: form the streets to servers?".
46. Haya. Alimagwashi and Steve McIntosh. 2009. "Understanding the Government to eGovernment Transition using a Soft Systems Approach: What is eGovernment Supposed to do?". Proceedings of the 9th European Conference on eGovernment
47. "i2010 2005– "i 2010 A European Information Society for growth and employment", Communication from the Commission to the European Parliament, the Council, the EESC and the Committee of the Regions, Brussels, COM(2005)



48. Osmo. D. 2008. "Benchmarking Government in the web 2.0 era: what to measure, and how". European journal of e-Practice, V 4, August 2008. <http://www.epracticejournal.eu>. ISSN:1988-625X.
49. Lenova. M 2009 "New Index for Measuring Feedback and e-Participation effectiveness of eGovernment in Russia". Proceedings of the 9th European Conference on eGovernment
50. Pang Bo and Lee Lillian, 2008. "Opinion Mining and Sentiment Analysis." Foundations and Trends in information Retrieval.
51. Zabin J and Jefferies A, (2008), "Social media monitoring and analysis: Generating consumer insights from online conversation", Aberdeen group Benchmark Report.
52. [http://www.infosoc.gr/NR/ronlyres/5CDB2236-DD00-410C-B5AB-D4CF80E5AF93/6870/EGOV\\_%CE%97%CE%BB%CE%94%CE%B9%CE%B1%CE%BA%CF%85%CE%B2%CE%AD%CF%81%CE%BD%CE%B7%CF%83%CE%B7Web2.pdf](http://www.infosoc.gr/NR/ronlyres/5CDB2236-DD00-410C-B5AB-D4CF80E5AF93/6870/EGOV_%CE%97%CE%BB%CE%94%CE%B9%CE%B1%CE%BA%CF%85%CE%B2%CE%AD%CF%81%CE%BD%CE%B7%CF%83%CE%B7Web2.pdf)
53. <http://en.wikipedia.org/wiki/E-democracy>
54. <http://www.dataminingtools.net/viewtutorials.php?id=3667258>
55. [http://en.wikipedia.org/wiki/Operator\\_\(programming\)](http://en.wikipedia.org/wiki/Operator_(programming))
56. [http://en.wikipedia.org/wiki/Document\\_classification](http://en.wikipedia.org/wiki/Document_classification)
57. [http://rapid-i.com/wiki/index.php?title=Text:Process\\_Documents\\_from\\_Files](http://rapid-i.com/wiki/index.php?title=Text:Process_Documents_from_Files)
58. <http://dictionary.sensagent.com/%CE%BB%CE%B5%CE%BE%CE%B9%CE%BA%CF%8C+%CF%83%CF%85%CE%BD%CF%89%CE%BD%CF%8D%CE%BC%CF%89%CE%BD/el/el/>
59. [http://www.neurolingo.gr/el/online\\_tools/lexiscope.htm](http://www.neurolingo.gr/el/online_tools/lexiscope.htm)
60. <http://office.microsoft.com/el-gr/word-help/HP010151361.aspx>
61. <http://www.alfavita.gr/lexikosinonim.htm>
62. [http://en.wikipedia.org/wiki/Character\\_encoding](http://en.wikipedia.org/wiki/Character_encoding)
63. <http://www.slideshare.net/rapidminercontent/rapidminer-word-vector-tool-and-rapid-miner-3667261>
64. <http://www.searchenginedictionary.com/terms-term-frequency.shtml>
65. <http://www.dataminingtools.net/viewtutorials.php?id=3627412>
66. <http://google.csd.uoc.gr/apache2-default/index.php/Stemmer>

67. <http://rapid-i.com/api/rapidminer-5.1/com/rapidminer/operator/validation/XValidation.html>
68. <http://www.dataminingtools.net/viewtutorials.php?id=3667253>
69. <http://www.slideshare.net/rapidminercontent/rapidminer-advanced-processes-and-operators-3667255>
70. <http://www.statsoft.com/textbook/naive-bayes-classifier/>
71. <http://www.cs.cmu.edu/~tom/mlbook/NBayesLogReg.pdf>
72. [http://en.wikipedia.org/wiki/Naive\\_Bayes\\_classifier](http://en.wikipedia.org/wiki/Naive_Bayes_classifier)
73. [http://rapid-i.com/wiki/index.php?title=Apply\\_Model&oldid=2772](http://rapid-i.com/wiki/index.php?title=Apply_Model&oldid=2772)
74. [http://rapid-i.com/wiki/index.php?title=Performance\\_\(Regression\)](http://rapid-i.com/wiki/index.php?title=Performance_(Regression))
75. [http://rapid-i.com/wiki/index.php?title=Write\\_Model&oldid=2646](http://rapid-i.com/wiki/index.php?title=Write_Model&oldid=2646)
76. [http://rapid-i.com/wiki/index.php?title=Select\\_Attributes](http://rapid-i.com/wiki/index.php?title=Select_Attributes)
77. [http://el.wikipedia.org/wiki/%CE%97%CE%BB%CE%B5%CE%BA%CF%84%CF%81%CE%BF%CE%BD%CE%B9%CE%BA%CE%AE\\_%CE%B4%CE%B9%CE%B1%CE%BA%CF%85%CE%B2%CE%AD%CF%81%CE%BD%CE%B7%CF%83%CE%B7](http://el.wikipedia.org/wiki/%CE%97%CE%BB%CE%B5%CE%BA%CF%84%CF%81%CE%BF%CE%BD%CE%B9%CE%BA%CE%AE_%CE%B4%CE%B9%CE%B1%CE%BA%CF%85%CE%B2%CE%AD%CF%81%CE%BD%CE%B7%CF%83%CE%B7)
78. <http://www.ejeg.com/search/index.html?name=keywords&value=opinion%20mining>
79. [http://www.go-online.gr/ebusiness/specials/article.html?article\\_id=1518](http://www.go-online.gr/ebusiness/specials/article.html?article_id=1518)
80. [http://ec.europa.eu/information\\_society/activities/egovernment/index\\_en.htm](http://ec.europa.eu/information_society/activities/egovernment/index_en.htm)
81. [http://en.wikipedia.org/wiki/Information\\_and\\_communication\\_technologies](http://en.wikipedia.org/wiki/Information_and_communication_technologies)
82. [http://www.go-online.gr/ebusiness/specials/article.html?article\\_id=1516](http://www.go-online.gr/ebusiness/specials/article.html?article_id=1516)
83. <http://www.w3.org/2008/09/msnws/papers/egov-social-ws.html>
84. [http://en.wikipedia.org/wiki/Web\\_2.0](http://en.wikipedia.org/wiki/Web_2.0)
85. <http://www-05.ibm.com/dk/files/EGovernment08.pdf>
86. <http://www.slideshare.net/osimod/web2-0-for-egovernment-why-and-how>
87. [http://www.observatory.gr/files/meletes/EGOV\\_A070110TX\\_%CE%97%CE%BB%20%CE%94%CE%B9%CE%B1%CE%BA%CF%85%CE%B2%CE%AD%CF%81%CE%BD%CE%B7%CF%83%CE%B7%20Web2.pdf](http://www.observatory.gr/files/meletes/EGOV_A070110TX_%CE%97%CE%BB%20%CE%94%CE%B9%CE%B1%CE%BA%CF%85%CE%B2%CE%AD%CF%81%CE%BD%CE%B7%CF%83%CE%B7%20Web2.pdf)
88. [http://ec.europa.eu/information\\_society/activities/egovernment/docs/reports/eu\\_eparticipation\\_summary\\_nov\\_09.pdf](http://ec.europa.eu/information_society/activities/egovernment/docs/reports/eu_eparticipation_summary_nov_09.pdf)
89. [http://en.wikipedia.org/wiki/Social\\_media](http://en.wikipedia.org/wiki/Social_media)

90. <http://www.scribd.com/doc/50863975/eGovernment-in-GR-april-2010-Local-Edition-2-0>
91. <http://europa.eu/rapid/pressReleasesAction.do?reference=IP/11/206&format=HTML&aged=1&language=EN&guiLanguage=en>
92. [http://ec.europa.eu/information\\_society/newsroom/cf/item-detail-dae.cfm?item\\_id=6537](http://ec.europa.eu/information_society/newsroom/cf/item-detail-dae.cfm?item_id=6537)
93. [http://en.wikipedia.org/wiki/Open\\_Method\\_of\\_Coordination](http://en.wikipedia.org/wiki/Open_Method_of_Coordination)
94. <http://www.ekt.gr/content/display?prnbr=82722>
95. [http://ec.europa.eu/information\\_society/eeurope/2002/action\\_plan/pdf/egovindicators.pdf](http://ec.europa.eu/information_society/eeurope/2002/action_plan/pdf/egovindicators.pdf)
96. <http://www.observatory.gr/files/meletes/eGov20basicServices.pdf>
97. <http://www.epractice.eu/en/document/288256>
98. <http://www.epractice.eu/en/document/288257>
99. [http://ec.europa.eu/information\\_society/activities/egovernment/action\\_plan\\_2011\\_2015/index\\_en.htm](http://ec.europa.eu/information_society/activities/egovernment/action_plan_2011_2015/index_en.htm)
100. <http://techblog.gr/internet/e-government-europe-2002/>
101. [http://www.opengov.gr/home/?page\\_id=136](http://www.opengov.gr/home/?page_id=136)
102. [http://www.observatory.gr/files/meletes/%CE%97%CE%BB\\_%CE%A0%CF%81%CE%BF%CE%BC%CE%AE%CE%B8%CE%B5%CE%B9%CE%B5%CF%82-17-12-2007.pdf](http://www.observatory.gr/files/meletes/%CE%97%CE%BB_%CE%A0%CF%81%CE%BF%CE%BC%CE%AE%CE%B8%CE%B5%CE%B9%CE%B5%CF%82-17-12-2007.pdf)
103. <http://www.capgemini.com/news-and-events/news/continued-improvement-in-european-egovernment-services/>
104. <http://www.capgemini.com/insights-and-resources/by-publication/2010-egovernment-benchmark/>
105. [http://en.wikipedia.org/wiki/User\\_experience](http://en.wikipedia.org/wiki/User_experience)
106. [http://en.wikipedia.org/wiki/User\\_experience\\_evaluation](http://en.wikipedia.org/wiki/User_experience_evaluation)
107. [http://static.googleusercontent.com/external\\_content/untrusted\\_dlcp/research.google.com/el//pubs/archive/36299.pdf](http://static.googleusercontent.com/external_content/untrusted_dlcp/research.google.com/el//pubs/archive/36299.pdf)
108. <http://www.globaluserresearch.com/blog/2010/03/user-experience-egovernment>
109. <http://www.epractice.eu/en/news/290427>
110. [http://www.rantos.com/Papers/The\\_Realization\\_of\\_the\\_Greek\\_E-Gif.pdf](http://www.rantos.com/Papers/The_Realization_of_the_Greek_E-Gif.pdf)

111. [http://www.items.fr/IMG/pdf/Constantine\\_Steriadis\\_GF\\_2008.pdf](http://www.items.fr/IMG/pdf/Constantine_Steriadis_GF_2008.pdf)