



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ Μ/Υ  
ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ  
ΣΧΟΛΗ ΝΑΥΤΙΛΙΑΣ ΚΑΙ ΒΙΟΜΗΧΑΝΙΑΣ  
ΤΜΗΜΑΤΟΣ ΒΙΟΜΗΧΑΝΙΚΗΣ ΔΙΟΙΚΗΣΗΣ &  
ΤΕΧΝΟΛΟΓΙΑΣ  
ΔΙΑΠΑΝΕΠΙΣΤΗΜΙΑΚΟ ΠΡΟΓΡΑΜΜΑ  
ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ  
«ΤΕΧΝΟ-ΟΙΚΟΝΟΜΙΚΑ ΣΥΣΤΗΜΑΤΑ»



## ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

“ΕΜΠΛΟΥΤΙΣΜΟΣ ΤΗΣ ΑΝΑΛΥΣΗΣ ΚΟΙΝΩΝΙΚΩΝ ΔΙΚΤΥΩΝ ΜΕ ΠΕΡΙΕΧΟΜΕΝΟ  
ΑΠΟ ΕΙΚΟΝΕΣ ΤΟΥ TWITTER”

ΘΕΟΦΙΛΟΥ ΓΙΩΡΓΟΣ

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ  
ΑΝΑΓΝΩΣΤΟΠΟΥΛΟΣ ΙΩΑΝΝΗΣ

Ιούνιος 2020

## Ευχαριστίες

Με την ολοκλήρωση της μεταπτυχιακής διπλωματικής μου εργασίας, θα ήθελα να εκφράσω τις θερμές ευχαριστίες μου σε όλους όσους συνέβαλαν στην εκπόνηση της.

Ευχαριστώ θερμά τον καθηγητή κ. Αναγνωστόπουλο Ιωάννη για την εμπιστοσύνη του, αναθέτοντας μου το συγκεκριμένο θέμα.

Ευχαριστώ θερμά τον κ. Ραζή Γεράσιμο για την καθοδήγησή του, την πολύτιμη βοήθειά του και την άριστη συνεργασία μας.

Επίσης ευχαριστώ και τους δύο για την δυνατότητα που μου έδωσαν να συμμετέχω στη συγγραφή μιας δημοσίευσης ως αποτέλεσμα αυτής της εργασίας.

Τέλος, ευχαριστώ τους γονείς μου Γιώργο και Ζαχαρούλα, για την στήριξή τους και την φροντίδα τους καθ' όλη την διάρκεια των σπουδών μου.

## Abstract

In this thesis, we propose a framework that uses latent information from Twitter images by employing the Google Cloud Vision API platform aiming at enriching social analytics with semantics and textual information. Our study reveals that user-generated content, linked data as well as hidden concepts and textual information from social images can be highly considered for enriching social analytics. Finally, we publish our annotated dataset for further use and evaluation from our research community.

**Keywords:** Social labeling, Twitter, Images, Google Cloud Vision API, OCR

## Περίληψη

Σε αυτή τη διπλωματική εργασία περιγράφεται και υλοποιείται μία μεθοδολογία για τον εμπλουτισμό των χαρακτηριστικών των χρηστών Κοινωνικών Δικτύων, με βάση τις διαμοιραζόμενες από αυτούς εικόνες. Αυτές οι πληροφορίες μπορούν να συνδυαστούν με το παραδοσιακό περιεχόμενο κειμένου των δημοσιεύσεων για την πιο ολοκληρωμένη και σφαιρική ανάλυση του Twitter. Για την ανάλυση των εικόνων χρησιμοποιήθηκε η πλατφόρμα Google Cloud Vision API, η οποία δίνει τη δυνατότητα για επισήμανση εικόνων με ετικέτες (labels), αλλά και εξαγωγή κειμένου με τεχνικές ανάγνωσης κειμένου (Optical Character Recognition - OCR).

Η αξιολόγηση των πειραμάτων έδειξε ότι το περιεχόμενο που διαμοιράζονται οι χρήστες, τα διασυνδεδεμένα δεδομένα (linked data) καθώς και οι οντότητες συμπληρωματικά με το κείμενο που εξάγονται από τις εικόνες Κοινωνικών Δικτύων είναι εξαιρετικές πηγές για τον εμπλουτισμό της κοινωνικής ανάλυσης. Εκτός αυτού, η διερεύνηση της εξαχθείσας πληροφορίας προσφέρει πολύτιμα στοιχεία για τα μοτίβα της λανθάνουσας πληροφορίας που εμπεριέχεται στις εικόνες που δημοσιεύουν οι χρήστες, οι οποίοι ανήκουν στην ίδια ή σε επικαλυπτόμενες κοινότητες. Συνεπώς, η αξιοποίησή τους μπορεί να προσφέρει πολύτιμες πληροφορίες, οι οποίες μέχρι τώρα αγνοούνταν.

Τέλος, δημοσιεύουμε το σύνολο των δεδομένων το οποίο παράχθηκε κατά την εκπόνηση αυτής της διπλωματικής εργασίας για περαιτέρω χρήση και αξιολόγηση από την επιστημονική κοινότητα.

# Περιεχόμενα

[Ευχαριστίες](#)

[Περίληψη](#)

[Περιεχόμενα](#)

[Πίνακας περιεχομένων εικόνων](#)

[Πίνακας περιεχομένων πινάκων](#)

## [1. Εισαγωγή](#)

[1.1 Περιγραφή Προβλήματος και Σκοπός της Διπλωματικής](#)

[1.1.1 Στόχος της Εργασίας](#)

[1.1.2 Ανάλυση του προβλήματος](#)

[1.1.3 Βασικά Συμπεράσματα και Συνεισφορά](#)

[1.1.4 Σύνοψη επομένων κεφαλαίων](#)

[1.2 Σχετική Βιβλιογραφία](#)

## [2. Η προτεινόμενη εφαρμογή](#)

[2.1 Επισκόπηση](#)

[2.2 Τεχνική Ανάλυση](#)

[2.2.1 Frontend](#)

[2.2.1.1 Γλώσσα Προγραμματισμού & Frameworks](#)

[2.2.1.2 AJAX](#)

[2.2.1.3 Functionality](#)

[2.2.2 Backend](#)

[Controller](#)

[Service](#)

[DAO \(Data Access Layer\)](#)

[2.2.2.1 Γλώσσα Προγραμματισμού & Frameworks](#)

[Εικ.4 Αρχιτεκτονική τριών επιπέδων](#)

[2.2.2.2 Βάση Δεδομένων](#)

[USER](#)

[FILE](#)

[GOOGLE\\_RESPONSE](#)

[FILE\\_USER\\_ASSOC](#)

[NLP](#)

[STOP\\_WORDS](#)

[TAG, TAG\\_USER\\_ASSOC](#)

- [2.2.2.3 REST API](#)
- [2.2.3 Google Vision API](#)
- [2.2.4 Stanford CoreNLP](#)
- [2.3 Δεδομένα Twitter και Συλλογή Twitter Images](#)
  - [2.3.1 Προσδιορισμός των λογαριασμών Twitter](#)
  - [2.3.2 Συλλογή Δεδομένων](#)
- [2.4 Ανάλυση χρήσης εφαρμογής](#)
  - [2.4.1 Upload Tab](#)
    - [Εισαγωγή - Επιλογή Χρήστη](#)
    - [Μεταφόρτωση Αρχείου](#)
    - [Λίστα Αρχείων](#)
  - [2.4.2 Statistics Tab](#)
    - [Count](#)
    - [Average](#)
    - [TF](#)
    - [TFIDF](#)
    - [Dominant Words](#)

[3. Αποτελέσματα και Αξιολόγηση](#)

- [3.1 Πειράματα - Μεθοδολογία](#)
- [3.2 Αξιολόγηση Αποτελεσμάτων](#)
- [3.3 Συμπεράσματα και Μελλοντικές Επεκτάσεις Εφαρμογής](#)

[Βιβλιογραφία](#)

## Πίνακας περιεχομένων εικόνων

Εικ.1 Η Αρχιτεκτονική της προτεινόμενης εφαρμογής	12
Εικ.2 Τεχνική ανάλυση αρχιτεκτονικής	12
Εικ.3 Αρχιτεκτονική Single-page application	13
Εικ.4 Αρχιτεκτονική τριών επιπέδων	15
Εικ.5 ER-Diagram	18
Εικ.6 Εικόνα από τον λογαριασμό @CNN, όπως αναλύθηκε από το GC Vision API. Με πράσινο πλαίσιο είναι η επισήμανση του αναγνωρισμένου κειμένου.	19
Εικ.7 Δείγμα από response του GC Vision API για labels	21
Εικ.8 Δείγμα από response του GC Vision API για labels	21
Εικ.9 Διάγραμμα λειτουργίας Stanford CoreNLP διάγραμμα	22
Εικ.10 Διεπαφή επιλογής χρήστη	24
Εικ.11 Διεπαφή μεταφόρτωσης αρχείων	25
Εικ.12 Διεπαφή μαζικής υποβολής	26
Εικ.13 Διεπαφή λίστας αρχείων	26
Εικ.14 Διεπαφή επισκόπησης δεδομένων	27
Εικ.15 Διεπαφή επιλογής γραφημάτων	28
Εικ.16 Tag Cloud των labels του ccnbrk με μετρική count	30
Εικ.17 Histogram των labels του ccnbrk με μετρική count	30
Εικ.18 Tag Cloud των labels όλων των πολιτικών με μετρική average	31
Εικ.19 Histogram των labels όλων των πολιτικών με μετρική average	31
Εικ.20 Tag Cloud των OCR όλων των πολιτικών με μετρική Dominant Words	32
Εικ.21 Histogram των OCR όλων των πολιτικών με μετρική Dominant Words	32
Εικ.22 Τα πρώτα 15 labels ανά κατηγορία	35
Εικ.23 Οι 20 πιο συχνόι όροι με βάση τη μετρική Dominance Metric ανά κατηγορία	37

## Πίνακας περιεχομένων πινάκων

Πιν.1 Κατηγορίες των λογαριασμών	33
Πιν.2 Επισκόπηση δεδομένων	34

# 1. Εισαγωγή

## 1.1 Περιγραφή Προβλήματος και Σκοπός της Διπλωματικής

### 1.1.1 Στόχος της Εργασίας

Στόχος της εργασίας είναι η περιγραφή και υλοποίηση μίας προσαρμόσιμης μεθοδολογίας για τη λήψη, ανάλυση, επεξεργασία και τελικώς τον εμπλουτισμό των χαρακτηριστικών των χρηστών Κοινωνικών Δικτύων (ΚΔ), με βάση τις διαμοιραζόμενες από αυτούς εικόνες. Αυτές οι πληροφορίες μπορούν να συνδυαστούν με το παραδοσιακό περιεχόμενο κειμένου των δημοσιεύσεων για την πιο ολοκληρωμένη και σφαιρική ανάλυση του Twitter. Για την υλοποίηση τμημάτων των προαναφερθέντων διαδικασιών χρησιμοποιήθηκε η πλατφόρμα Google Cloud (GC) Vision API.

Το GC Vision API δίνει τη δυνατότητα για επισήμανση εικόνων με ετικέτες (labels), μέσω προ-εκπαιδευμένων μοντέλων μηχανικής μάθησης (Machine Learning), αλλά και εξαγωγή κειμένου με τεχνικές ανάγνωσης κειμένου (Optical Character Recognition - OCR). Για τους σκοπούς της διπλωματικής εργασίας αναλύθηκαν 39.499 εικόνες, οι οποίες επισημάνθηκαν με 4.059 μοναδικά labels και εξήχθησαν 26.953 τμήματα κειμένου. Οι εικόνες προήλθαν από 130 χρήστες του Twitter που ανήκουν στον τομέα της πολιτικής και είναι διαθέσιμες στον δικτυακό τόπο InfluenceTracker<sup>1</sup>.

Τέλος, δημοσιεύουμε το προαναφερθέν σύνολο δεδομένων για περαιτέρω χρήση και αξιολόγηση από την επιστημονική κοινότητα (<https://www.doi.org/10.34740/KAGGLE/DS/732777>).

### 1.1.2 Ανάλυση του προβλήματος

Στις μέρες μας τα μέσα κοινωνικής δικτύωσης αποτελούν ολοένα και πιο διαδεδομένο τρόπο για τη διάδοση περιεχομένου και πληροφοριών με χρήση εικόνων και βίντεο. Στην τάση αυτή συμβάλλουν οι βελτιωμένες κάμερες που φέρουν οι συσκευές κινητής τηλεφωνίας, αλλά και οι ευρυζωνικές συνδέσεις στο διαδίκτυο.

---

<sup>1</sup> <http://InfluenceTracker.com/>

Όπως αναφέρεται στο [1], πολυμεσικό περιεχόμενο εντοπίζεται στο μεγαλύτερο ποσοστό των tweets, τα οποία και προσελκύουν περισσότερο το ενδιαφέρον των χρηστών από άποψη likes, retweets και replies, σε σχέση με εκείνα που φέρουν μόνο κείμενο. Ως εκ τούτου οι εικόνες έχουν εξελιχθεί σε σημαντικό στοιχείο των μηνυμάτων που διαμοιράζονται στα ΚΔ, δρώντας συμπληρωματικά των λεκτικών πληροφοριών. Ωστόσο, λίγη προσοχή έχει δοθεί σε αυτό το περιεχόμενο και τους τρόπους που μπορεί να αξιοποιηθεί για τον εμπλουτισμό των κοινωνικών προφίλ των χρηστών.

Μέχρι πρόσφατα, η ανάλυση των ΚΔ επικεντρωνόταν αποκλειστικά στο κείμενο, αγνοώντας εντελώς το πολυμεσικό περιεχόμενο που τα συνοδεύει. Ο όγκος του πολυμεσικού περιεχομένου που κοινοποιείται στα ΚΔ αυξάνεται συνεχώς και μπορεί να εμπεριέχει πληροφορίες που δεν περιλαμβάνονται στο κείμενο. Ως εκ τούτου μπορεί να αποτελέσει πολύτιμη πηγή δεδομένων για τον εμπλουτισμό [2] των πληροφοριών που διαδίδονται.

### 1.1.3 Βασικά Συμπεράσματα και Συνεισφορά

Η συνεισφορά της εργασίας είναι διττή. Πρώτον, προτείνουμε μια προσαρμόσιμη μεθοδολογία που χρησιμοποιεί λανθάνουσες πληροφορίες οι οποίες έχουν εξαχθεί από εικόνες του Twitter, επιστρατεύοντας την πλατφόρμα Google Cloud Vision API, με στόχο τον εμπλουτισμό της ανάλυσης Κοινωνικών Δικτύων με σημασιολογικά δεδομένα και πληροφορίες κειμένου. Επιπλέον, χρησιμοποιώντας τεχνικές αναγνώρισης κειμένου OCR εξάγουμε τις λεκτικές πληροφορίες που απεικονίζονται στις εικόνες. Η μελέτη μας έδειξε ότι το περιεχόμενο που διαμοιράζονται οι χρήστες, τα διασυνδεδεμένα δεδομένα (linked data) καθώς και οι οντότητες συμπληρωματικά με το κείμενο που εξάγονται από εικόνες ΚΔ είναι εξαιρετικές πηγές για τον εμπλουτισμό της κοινωνικής ανάλυσης. Εκτός αυτού, η διερεύνηση της εξαχθείσας πληροφορίας προσφέρει πολύτιμα στοιχεία για τα μοτίβα της λανθάνουσας πληροφορίας που εμπεριέχεται στις εικόνες που δημοσιεύουν οι λογαριασμοί των ΚΔ, οι οποίοι ανήκουν στην ίδια ή σε επικαλυπτόμενες κοινότητες. Τέλος, η αξιοποίησή τους θα μπορούσε να προσφέρει πολύτιμες πληροφορίες, οι οποίες μέχρι τώρα αγνοούνταν.

Δεύτερον, δημοσιεύουμε ένα σύνολο δεδομένων που αποτελείται από 39.499 φωτογραφίες, από τις οποίες εξήχθησαν 4.059 labels και δημοσιεύτηκαν από 130 λογαριασμούς χρηστών στο Twitter που ανήκουν στον πολιτικό τομέα. Επιπλέον, η συντριπτική πλειονότητα των εικόνων (68,2%) περιείχε κείμενο που επιβεβαιώνει την έρευνά μας. Τέλος δημοσιεύουμε το σύνολο των δεδομένων που δημιουργήθηκε κατά την εκπόνηση της διπλωματικής εργασίας για περαιτέρω χρήση και αξιολόγηση από την ερευνητική κοινότητα. Επιπρόσθετα, παρέχεται ο πηγαίος κώδικας της εφαρμογής στο δημόσιο αποθετήριο <https://github.com/gtheofilou/gvapi>.

Αποτέλεσμα αυτής της εργασίας είναι η συγγραφή μιας δημοσίευσης η οποία έχει γίνει αποδεκτή για παρουσίαση στο διεθνές συνέδριο SMAP 2020 (<https://smap2020.eu/>).



### 1.1.4 Σύνοψη επομένων κεφαλαίων

Στην επόμενη ενότητα γίνεται αναφορά στη σχετική βιβλιογραφία και στις εργασίες πάνω στην ανάλυση του πολυμεσικού περιεχομένου στα Κοινωνικά Δίκτυα. Στο κεφάλαιο 2 παρουσιάζεται η προτεινόμενη μεθοδολογία, η τεχνική ανάλυση της εφαρμογής, καθώς και ένα εγχειρίδιο χρήσης της. Τέλος, στο κεφάλαιο 3 παρουσιάζονται αναλυτικά τα αποτελέσματα της μεθοδολογίας καθώς και η αξιολόγησή τους.

## 1.2 Σχετική Βιβλιογραφία

Ο εντοπισμός της συσχέτισης μεταξύ κειμένου και πολυμεσικού περιεχομένου δεν είναι καινούριο θέμα. Ωστόσο, κερδίζει όλο και περισσότερη προσοχή και χρησιμοποιείται σε πλήθος νέων εφαρμογών. Επιπροσθέτως, χρήσιμες πληροφορίες για τους χρήστες των ΚΔ και τα χαρακτηριστικά τους μπορούν επίσης να εξαχθούν, τα οποία μπορούν να χρησιμοποιηθούν με τη σειρά τους για τη βελτίωση συστημάτων σχετικά με τη δημιουργία προφίλ χρηστών, τις προτάσεις για σχετικές οντότητες (recommendation systems), και τον εμπλουτισμό της πληροφορίας και την κατηγοριοποίηση.

Οι συγγραφείς στο [2] ερευνούν την ύπαρξη συσχέτισης μεταξύ κειμένου και εικόνας σε επίπεδο tweet προτείνοντας μια μεθοδολογία δύο σταδίων. Κατά την πρώτη φάση εξετάζεται η σημασιολογική συσχέτιση μεταξύ κειμένου και εικόνας και κατά τη δεύτερη εάν η εικόνα παρέχει επιπρόσθετες πληροφορίες. Ο βαθμός της συσχέτισης αξιολογείται με ανθρώπινη παρέμβαση.

Ομοίως, η μελέτη στο [4] αναλύει ένα σύνολο εικόνων ΚΔ για την αξιολόγηση της συσχέτισης μεταξύ της δημοσιευμένης εικόνας και του συνοδευτικού κειμένου. Η μελέτη καταλήγει στο συμπέρασμα ότι υπάρχουν δημοσιεύσεις με υψηλή σημασιολογική συσχέτιση μεταξύ κειμένου και εικόνας, δρώντας συμπληρωματικά, όπως επίσης και ότι υπάρχουν δημοσιεύσεις όπου οι εικόνες χρησιμοποιούνταν για την αύξηση του βαθμού αλληλεπίδρασης με άλλους χρήστες.

Στη [5] προτείνεται μια χωροχρονική ανάλυση κάνοντας χρήση προεκπαιδευμένων μοντέλων Νευρωνικών Δικτύων. Αυτή η μελέτη, επέκταση της προηγούμενης των συγγραφέων, ταξινομεί εικόνες από tweet με γεωγραφικές ετικέτες ως σχετικές σε μία κατάσταση έκτακτης ανάγκης τόσο γεωγραφικά όσο και χρονικά.

Μια μεθοδολογία 2 βημάτων με βάση το φύλο χρηστών ΚΔ παρουσιάζεται στο [6], αναλύοντας τις δημοσιεύσεις τους τόσο από πλευράς κειμένου όσο και πολυμεσικού περιεχομένου. Κάνοντας χρήση ενός SVM μοντέλου, με την ανάλυση του κειμένου υπολογίζεται η πιθανότητα για το φύλο, ενώ με την ανάλυση των εικόνων και την παραπάνω πιθανότητα γίνεται ταξινόμηση σε ένα σύνολο προκαθορισμένων κατηγοριών.

Η μελέτη στο [7] διερευνά την πιθανή σημασία των χαρακτηριστικών των εικόνων του Instagram στις προσωπικότητες των χρηστών. Δύο τύποι χαρακτηριστικών εξάγονται, οπτικά, όπως η απόχρωση και κορεσμός, και περιεχόμενο, δηλαδή οντότητες που απεικονίζονται στις εικόνες. Η μελέτη καταλήγει στο συμπέρασμα ότι και τα δύο οπτικά χαρακτηριστικά περιέχουν πολύτιμες πληροφορίες για την πρόβλεψη της προσωπικότητας του χρήστη.

Πρόσφατα, προϊόντα επί πληρωμή για ανάλυση και ταξινόμηση εικόνων έγιναν διαθέσιμα (π.χ. Microsoft Azure Computer Vision<sup>2</sup>, Google Cloud Vision API<sup>3</sup>), με πολλές μελέτες να τα χρησιμοποιούν όλο και περισσότερο. Συγκεκριμένα, στο [8] προτείνεται μια μεθοδολογία για γεωχωρική ανάλυση συναισθημάτων σε εικόνες και μηνύματα σχετιζόμενα με καταστροφές σε διάφορες πληγείσες περιοχές. Τα μοντέλα που χρησιμοποιήθηκαν για την ανάλυση συναισθημάτων, στηρίζονται στην ανάλυση προσώπων, απέτυχαν λόγω του ότι οι εικόνες δεν περιείχαν πρόσωπα. Για την αντιμετώπιση αυτού του περιορισμού, χρησιμοποιήθηκε το Google Cloud Vision API για την εξαγωγή του περιεχομένου από τις εικόνες με τη μορφή λέξεων-κλειδιών και ετικετών. Έπειτα η ανάλυση εφαρμόζεται επί των εξαχθέντων όρων.

Οι συγγραφείς του [9] κάνουν επίσης χρήση του Google Cloud Vision API για ανάλυση εικόνων που δημοσιεύονται στο Twitter από κρατικούς λογαριασμούς. Στόχος είναι να αναλυθεί η τάση διάδοσης «ψεύτικων ειδήσεων» και κατά πόσο οι εικόνες βοηθούν στην αξιοπιστία τέτοιων μηνυμάτων.

Μια μεθοδολογία για τη δημιουργία μιας γνωσιακής βάσης αποτελούμενη από οντότητες του Twitter, καθώς και από λεκτικό και πολυμεσικό περιεχόμενο, παρουσιάζεται σε [10]. Το “κλασικά” μεταδεδομένα του Twitter, όπως χρήστες, οντότητες και συσχετίσεις, εμπλουτίζονται με πληροφορίες από το περιεχόμενο των εικόνων με χρήση του Google Cloud Vision API.

Οι συγγραφείς στο [11] χρησιμοποιούν περιεχόμενο ΚΔ για την αξιολόγηση της αποτελεσματικότητας στην υποστήριξη συστημάτων διαχείρισης της κυκλοφορίας. Για το σκοπό αυτό, ένας συνδυασμός μεθόδων εφαρμόζεται σε tweets σχετικά με την κυκλοφορία, συμπεριλαμβανομένης της επεξεργασίας φυσικής γλώσσας (Natural Language Processing - NLP) στο κείμενο και της αυτόματης αναγνώρισης οντοτήτων στις εικόνες χρησιμοποιώντας το Google Cloud Vision API. Η μελέτη καταλήγει στο συμπέρασμα ότι οι χρήστες ΚΔ σπάνια αναφέρουν λεπτομέρειες κίνησης, όπως ροή και μήκος ουράς, οι οποίες είναι σημαντικές παράμετροι για τη διαχείριση της κυκλοφορίας σε πραγματικό χρόνο.

Τέλος, στο [12] πραγματοποιήθηκε ποιοτική ανάλυση στα labels εικόνων συγκεκριμένης κατηγορίας, που προέρχεται από ανθρώπους και από το Google Cloud Vision API. Η αξιολόγηση έδειξε ότι το δεύτερο είναι πιο αποτελεσματικό. Συγκεκριμένα, τα labels προερχόμενα από ανθρώπους ήταν λιγότερα, συχνά εσφαλμένα και επηρεασμένα από τον τομέα ανάλυσης. Επιπλέον, μοντέλα που εκπαιδεύτηκαν στα δεδομένα προερχόμενα από

<sup>2</sup> <https://azure.microsoft.com/en-us/services/cognitive-services/computer-vision/>

<sup>3</sup> <https://cloud.google.com/vision/>

ανθρώπους είχαν μικρότερη ακρίβεια σε σύγκριση με μοντέλα που εκπαιδεύτηκαν σε δεδομένα από το Google Cloud Vision API.

## 2. Η προτεινόμενη εφαρμογή

Σε αυτή την ενότητα παρουσιάζεται η αρχιτεκτονική και η υποδομή της προτεινόμενης εφαρμογής για τη αυτόματη εξαγωγή labels και κειμένου από φωτογραφίες ΚΔ. Επιπλέον, παρουσιάζουμε τη χρήση του Google Cloud Vision API για την εξαγωγή labels και κειμένου με τεχνικές OCR.

### 2.1 Επισκόπηση

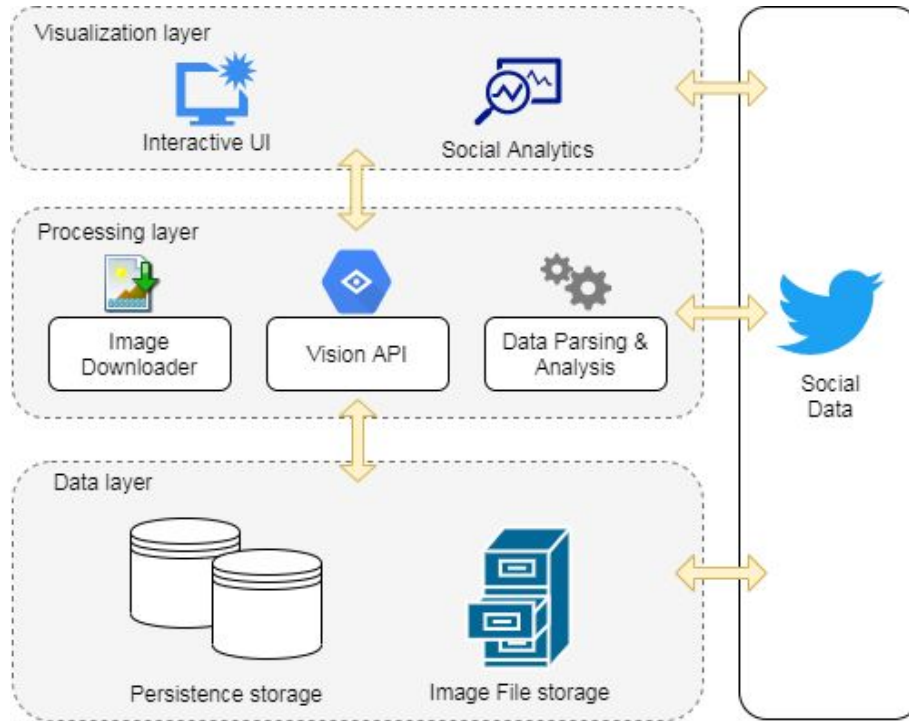
Η εφαρμογή αποτελείται από μια αρχιτεκτονική τριών επιπέδων, όπως παρουσιάζεται και στην Εικ. 1. Το επίπεδο “δεδομένων” (data layer) είναι υπεύθυνο για την αποθήκευση και την ανάκτηση δεδομένων στη βάση δεδομένων (H2 Database - Ενότητα 2.2.2.2). Επίσης, είναι υπεύθυνο για την αποθήκευση των εικόνων που μεταφορτώνονται από τα ΚΔ.

Το επίπεδο “επεξεργασίας” (processing layer) είναι υπεύθυνο για τη διασύνδεση και ανάκτηση δεδομένων από εξωτερικά συστήματα, καθώς και μετατροπή των δεδομένων αυτών σε μορφή κατάλληλη για αποθήκευση και περαιτέρω επεξεργασία. Αναλυτικότερα αποτελείται από 3 υποσυστήματα:

1. Το “Google Cloud Vision API” για την ανάλυση, εξαγωγή και εμπλουτισμό με labels,
2. Το “Ανάλυση και Μετατροπή δεδομένων” για την μετατροπή των δεδομένων που λαμβάνονται από το API, ώστε να είναι κατάλληλα για επεξεργασία και αποθήκευση,
3. Το “Μεταφόρτωση Εικόνων” για την ανάκτηση και αποθήκευση αρχείων εικόνας.

Στις περιπτώσεις που ένα tweet συσχετίζεται με μια εικόνα ή ένα βίντεο, τότε συνοδεύεται από έναν σύνδεσμο URL του πραγματικού αρχείου ή ενός thumbnail. Προφανώς αν μια εικόνα έχει δημοσιευθεί από πολλούς λογαριασμούς, τότε μεταφορτώνεται και αποθηκεύεται μια φορά στο σύστημά μας.

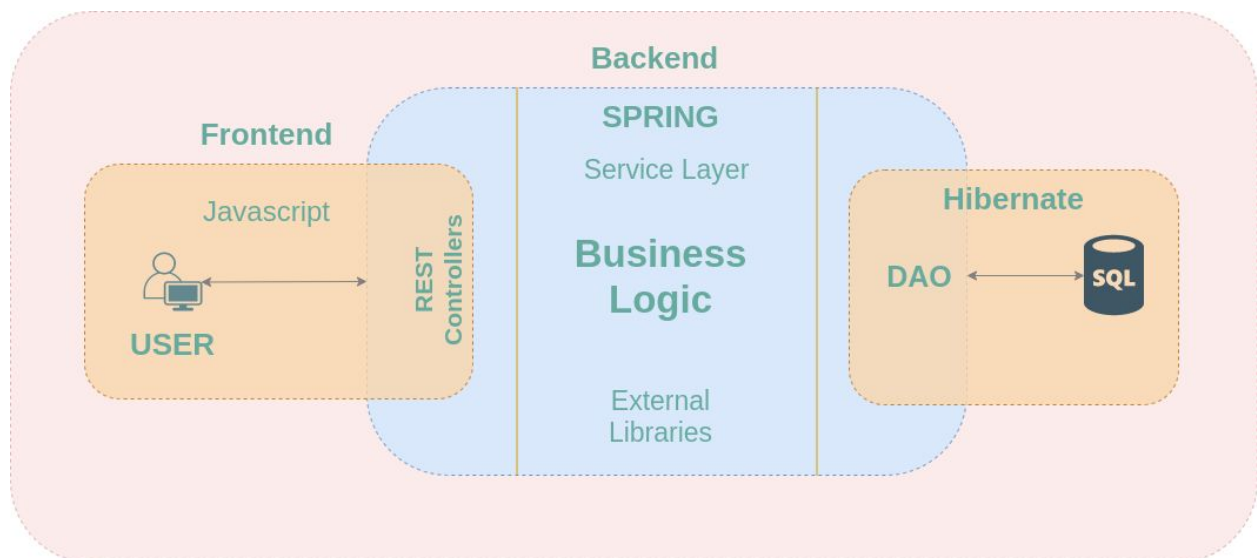
Τέλος, υπάρχει και το επίπεδο “παρουσίασης” (visualization layer), που είναι υπεύθυνο για την παρουσίαση αναλυτικών στοιχείων πάνω στο δημοσιευμένο υλικό των χρηστών.



Εικ.1 Η Αρχιτεκτονική της προτεινόμενης εφαρμογής

## 2.2 Τεχνική Ανάλυση

Από πλευρά προγραμματιστικής υλοποίησης η εφαρμογή αποτελείται από δύο διακριτά μέρη το Frontend και το Backend.

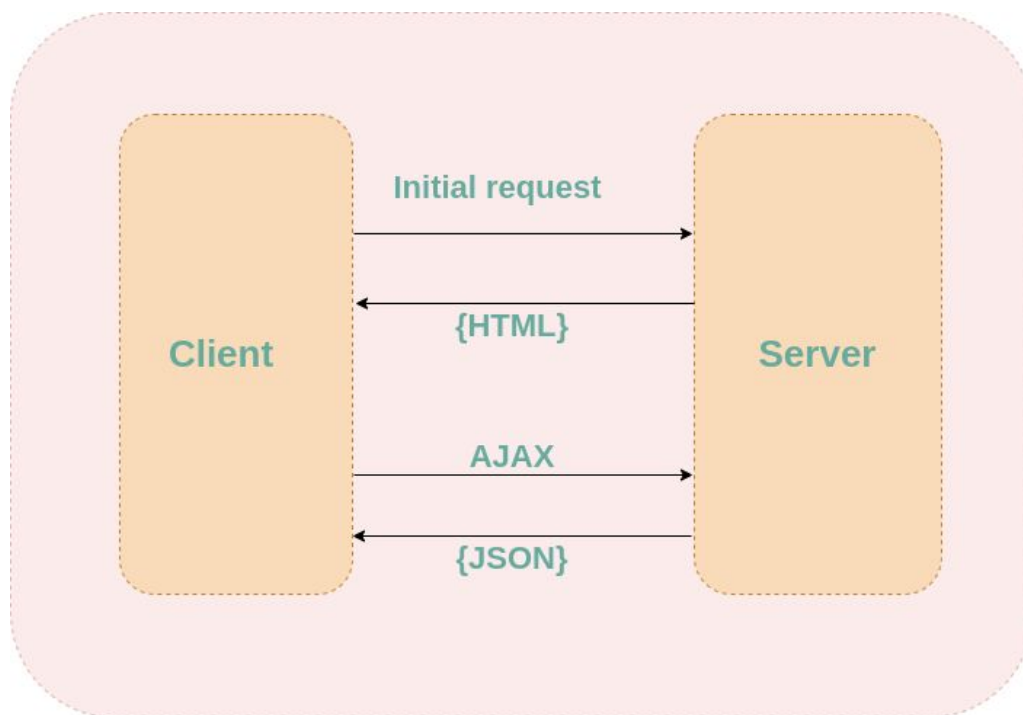


Εικ.2 Τεχνική ανάλυση αρχιτεκτονικής

Στις παρακάτω ενότητες παρουσιάζονται τα δομικά στοιχεία από τα οποία αποτελείται η εφαρμογή και πώς αυτά επικοινωνούν μεταξύ τους, τα οποία απεικονίζονται στην Εικ.2.

### 2.2.1 Frontend

Το frontend περιλαμβάνει τις γραφικές διεπαφές (GUI) που είναι απαραίτητες για την αλληλεπίδραση του χρήστη με την εφαρμογή. Είναι δομημένο σύμφωνα με την αρχιτεκτονική του **Single-page application** (SPA) κατά την οποία ο περιηγητής (browser) του χρήστη φορτώνει μια σελίδα μαζί με όλα τα απαραίτητα αρχεία που την αποτελούν, μία φορά. Σε κάθε αλληλεπίδραση που έχει ο χρήστης με την εφαρμογή δεν χρειάζεται να κατεβάσει μια καινούρια σελίδα, παρά μόνο τα δεδομένα και η εφαρμογή αναλαμβάνει να ενημερώσει την τρέχουσα σελίδα. Με αυτό τον τρόπο βελτιώνουμε την απόδοση της εφαρμογής και μειώνουμε τον φόρτο του web-server στο ελάχιστο δυνατό.



Εικ.3 Αρχιτεκτονική Single-page application

#### 2.2.1.1 Γλώσσα Προγραμματισμού & Frameworks

Οι τεχνολογίες που χρησιμοποιούνται είναι vanilla **Javascript** και το framework **jQuery**. Η σελίδα αποτελείται από διάφορα μέρη υπό την μορφή tabs. Κάθε tab της σελίδας αποτελείται από ένα Javascript αρχείο, το οποίο εκτελείται άμεσα κατά την παραλαβή του από τον browser

μέσω **Self-Invoking Functions**, δηλαδή μεθόδων που εκτελούνται αυτόματα και παρέχουν την λειτουργικότητα της εφαρμογής.

### 2.2.1.2 AJAX

Τα δεδομένα που ζητούνται από το frontend τμήμα της εφαρμογής γίνονται μέσω κλήσεων **AJAX** (Asynchronous JavaScript and XML) με χρήση του **jQuery**. Κατά την διάρκεια AJAX κλήσεων η σελίδα δε χρειάζεται να περιμένει (blocked) τα δεδομένα, αλλά γίνεται ενημέρωσή της όταν αυτά γίνουν διαθέσιμα στον browser.

### 2.2.1.3 Functionality

Η εφαρμογή αποτελείται από δύο tabs. Το πρώτο, **Upload** (Ενότητα 2.4.1) tab, είναι υπεύθυνο για όλες τις αλληλεπιδράσεις του χρήστη που αφορούν τη μεταφόρτωση εικόνων, εισαγωγή δεδομένων και αποστολή εικόνων στο Google Vision API. Το δεύτερο tab, **Statistics** (Ενότητα 2.4.2) είναι υπεύθυνο για την εξαγωγή αναλυτικών στοιχείων μέσω στατιστικών γραφημάτων.

## 2.2.2 Backend

Το backend τμήμα της εφαρμογής είναι υπεύθυνο για την επεξεργασία δεδομένων, την αποθήκευση τους στη βάση και την σύνδεση με εξωτερικά συστήματα μέσω APIs (π.χ. Google Vision). Είναι δομημένο σε αρχιτεκτονική 3 επιπέδων:

#### 1. *Controller*

Ο Controller λαμβάνει το request από το frontend, ελέγχει ότι είναι έγκυρο και καλεί όλα τα services από το επόμενο επίπεδο που είναι απαραίτητα για την εξυπηρέτηση του αιτήματος.

#### 2. *Service*

Το επίπεδο service εμπεριέχει όλο το business logic της εφαρμογής και χρησιμοποιεί το επίπεδο DAO για την επικοινωνία με τη βάση.

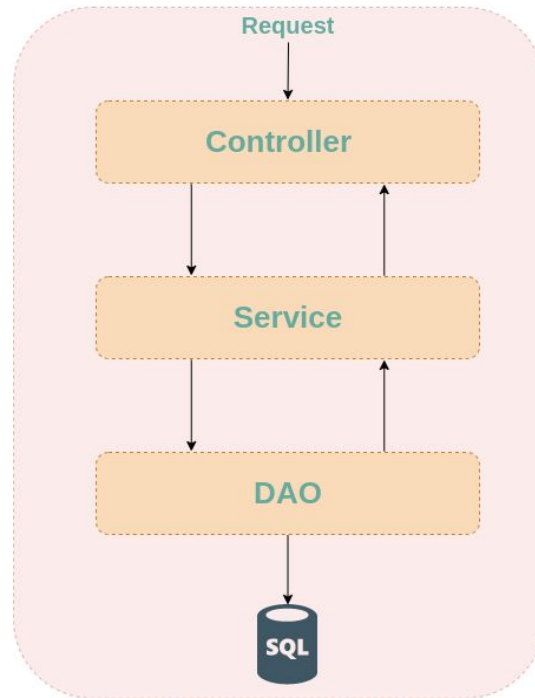
#### 3. *DAO (Data Access Layer)*

Το επίπεδο DAO είναι αποκλειστικά υπεύθυνο για την ανάκτηση και αποθήκευση των δεδομένων στη βάση

### 2.2.2.1 Γλώσσα Προγραμματισμού & Frameworks

Η προτεινόμενη εφαρμογή είναι υλοποιημένη στην γλώσσα προγραμματισμού Java 11 και χρησιμοποιεί το framework Spring.

Το Spring είναι το πιο διαδεδομένο framework για τη δημιουργία enterprise εφαρμογών σε Java.



Εικ.4 Αρχιτεκτονική τριών επιπέδων

Μερικά από τα πλεονεκτήματα που παρέχει το Spring και χρησιμοποιήθηκαν στην εφαρμογή είναι:

1. Δημιουργία loose coupled κώδικα
2. Διασύνδεση με βάσεις δεδομένων μέσω
3. Σύνδεση με third-party υπηρεσίες (π.χ. Google Vision API μέσω Spring Cloud)
4. Ευκολία στη δημιουργία test σεναρίων.

#### 2.2.2.2 Βάση Δεδομένων

Η βάση δεδομένων που χρησιμοποιήθηκε είναι η σχεσιακή embedded H2 Database και η επικοινωνία γίνεται με χρήση του ORM εργαλείου Hibernate. Το Hibernate παρέχει ένα abstraction layer ανάμεσα στην εφαρμογή και τη βάση, που μας δίνει την δυνατότητα να την αντικαταστήσουμε με κάποια άλλη της επιλογής μας αν αυτό κριθεί απαραίτητο.

Η H2 Database μπορεί να λειτουργήσει με δυο τρόπους, είτε in-memory αποθηκεύοντας τα δεδομένα προσωρινά στην RAM, είτε in-file για αποθήκευση σε αρχεία. Ο πρώτος τρόπος χρησιμοποιείται συνήθως κατά τη διαδικασία ανάπτυξης της εφαρμογής. Στη δική μας περίπτωση επιλέχθηκε η αποθήκευση σε αρχεία.

Στους παρακάτω πίνακες περιγράφονται αναλυτικά οι πίνακες της Βάσης Δεδομένων μας, καθώς και οι συσχετίσεις που υπάρχουν μεταξύ τους. Τέλος, στην Εικ.5 παρουσιάζεται το Entity-Relationship διάγραμμα τη βάσης.

#### USER

Ο πίνακας USER μοντελοποιεί έναν χρήστη με στοιχεία όπως το όνομα χρήστη στο Twitter.

Στήλη	Περιγραφή
ID	Το primary key της εγγραφής
NAME	Το όνομα του χρήστη
PARTY	Το πολιτικό κόμμα που ανήκει ο χρήστης (αν υπάρχει)
CATEGORY	Η κατηγορία που ανήκει ο χρήστης (αν υπάρχει)

#### FILE

Ο πίνακας FILE μοντελοποιεί ένα αρχείο με στοιχεία το όνομα αρχείου και αν έχει σταλεί στο GC Vision API.

Στήλη	Περιγραφή
ID	Το primary key της εγγραφής
NAME	Το όνομα του αρχείου
SENT	Κατάσταση του αρχείου, αν έχει αποσταλεί στο GC Vision API

#### GOOGLE\_RESPONSE

Ο πίνακας GOOGLE\_RESPONSE αποθηκεύει τις πληροφορίες που λαμβάνουμε από το GC Vision API.

Στήλη	Περιγραφή
ID	Το primary key της εγγραφής
DESCRIPTION	Το label ή το OCR κείμενο του GC Vision API
FILE_ID	Το primary key του αρχείου
SCORE	Το Score του label σύμφωνα με το GC Vision API
TYPE	Τύπος εγγραφής (0:label, 1:OCR text)
COSINE_SIM	Το cosine similarity των labels με το OCR text
FINAL_DESCRIPTION	Επεξεργασμένο DESCRIPTION



## FILE\_USER\_ASSOC

Ο πίνακας FILE\_USER\_ASSOC υλοποιεί μια many-to-many συσχέτιση ανάμεσα στους πίνακες USER και FILE. Αυτό μας επιτρέπει να συσχετίζουμε έναν χρήστη με πολλά αρχεία, αλλά και ένα αρχείο με πολλούς χρήστες. Επομένως αν ένας χρήστης ανεβάσει αρχείο που ήδη υπάρχει, δεν αποθηκεύεται δεύτερη φορά ούτε αποστέλλεται στο GC Vision API, παρά μόνο προστίθεται μια νέα εγγραφή στον πίνακα FILE\_USER\_ASSOC.

Στήλη	Περιγραφή
FILE_ID	Το primary key του αρχείου
USER_ID	Το primary key του χρήστη

## NLP

Ο πίνακας NLP αποθηκεύει τα αποτελέσματα από την ανάλυση φυσικής γλώσσας, όπως τη λέξη, το μέρος του λόγου και στατιστικά όπως το πλήθος εμφάνισης.

Στήλη	Περιγραφή
ID	Το primary key της εγγραφής
FILE_ID	Το primary key του αρχείου
POS	Το μέρος του λόγου που ανήκει η λέξη
RTF	Το πλήθος εμφανίσεων της λέξης στο έγγραφο
TF	Η τιμή της διαίρεσης RTF/TWD
TWD	Ο συνολικός αριθμός των λέξεων στο έγγραφο
WORD	Η λέξη καθεαυτή

## STOP\_WORDS

Ο πίνακας STOP\_WORDS αποθηκεύει λέξεις οι οποίες φιλτράρονται κατά την παρουσίαση των αποτελεσμάτων.

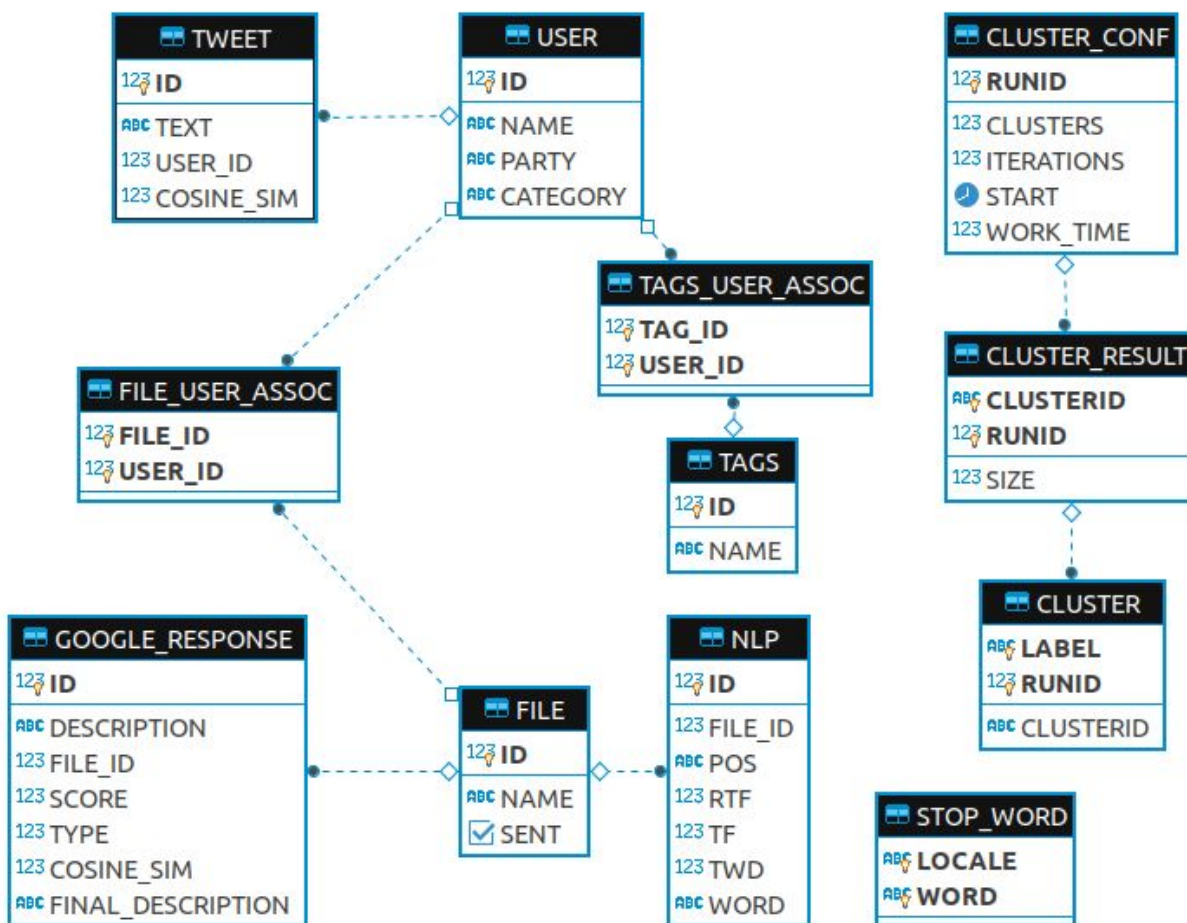
Στήλη	Περιγραφή
LOCALE	Η γλώσσα στην οποία ανήκει η λέξη
WORD	Η λέξη καθεαυτή

TAG, TAG\_USER\_ASSOC

Οι πίνακες TAG και TAG\_USER\_ASSOC αποθηκεύουν τις κατηγορίες στις οποίες ανήκει ένας χρήστης, μέσω μιας συσχέτισης many-to-many.

Στήλη	Περιγραφή
ID	To primary key της εγγραφής
NAME	To tag

Στήλη	Περιγραφή
TAG_ID	To primary key του tag
USER_ID	To primary key του χρήστη



Εικ.5 ER-Diagram

### 2.2.2.3 REST API

Η επικοινωνία με το frontend γίνεται με το παρεχόμενο REST API μέσω της υποδομής του Spring. Οι υπηρεσίες που παρέχονται είναι:

- Προσθήκη νέων χρηστών,
- Ανέβασμα αρχείων φωτογραφίας,
- Αποστολή φωτογραφίας στο Google Vision API,
- Ανάκτηση αποτελεσμάτων φωτογραφίας,
- Μαζική φόρτωση και αποστολή φωτογραφιών με την καταχώρηση URLs,
- NLP επεξεργασία μέσω της βιβλιοθήκης [Stanford CoreNLP](#).

### 2.2.3 Google Vision API

Το Google Vision είναι ένα API με δυνατότητες computer vision. Ο χρήστης μπορεί να ανεβάζει φωτογραφίες και μέσω τεχνικών μηχανικής μάθησης παρέχει λειτουργίες όπως:

- Αναγνώριση προσώπων, αντικειμένων,
- Αναγνώριση τοποθεσιών, logos ,
- Εξαγωγή labels,
- Αναγνώριση κειμένου (OCR).



Εικ.6 Εικόνα από τον λογαριασμό @CNN, όπως αναλύθηκε από το GC Vision API. Με πράσινο πλαίσιο είναι η επισήμανση του αναγνωρισμένου κειμένου.

Για τη διασύνδεση της εφαρμογής με το GC Vision API χρησιμοποιήθηκε το module Spring Cloud. Το Spring Cloud προσφέρει έναν εύχρηστο client για την αποστολή των αρχείων

φωτογραφίας στο GC Vision API και τη λήψη των αποτελεσμάτων. Η πρόσβαση στο REST API απαιτεί χρήση ενός API key με το οποίο γίνεται ταυτοποίηση του χρήστη.

Το GC Vision API είναι βασισμένο σε τελευταίας τεχνολογίας Νευρωνικά Δίκτυα (Deep Neural Network) και αλγόριθμους μηχανικής μάθησης. Στη διπλωματική εργασία η χρήση του GC Vision API είναι διττή:

- Κατηγοριοποίηση των εικόνων με βάση labels από ένα μεγάλο προκαθορισμένο σύνολο, τα οποία συνοδεύονται και από αντίστοιχους βαθμούς εμπιστοσύνης (βαθμός  $\in [0, 1]$ ),.
- Αυτόματη ανίχνευση κειμένου χρησιμοποιώντας προηγμένες τεχνικές OCR.

Κάθε μία από τις προαναφερθείσες υπηρεσίες του GC Vision API αποτελεί ξεχωριστό request, επομένως για κάθε φωτογραφία που ανεβάζει ο χρήστης, η εφαρμογή καλεί το API δυο φορές. Στις Εικ.7 και Εικ.8 φαίνονται παραδείγματα από τα responses του GC Vision API για την Εικ.6, που δημοσιεύτηκε στο Twitter από τον λογαριασμό του CNN, κατά τη διάρκεια συνέντευξης του Anthony Fauci, διευθυντή του Εθνικού Ινστιτούτου Αλλεργίας και Λοιμωδών Νοσημάτων (National Institute of Allergy and Infectious Diseases, NIAID) των ΗΠΑ, για τον κορωνοϊό.

Όπως φαίνεται στην Εικ.7, κάθε label (“description”) σχετίζεται έναν βαθμό εμπιστοσύνης (“score”) μαζί με ένα αναγνωριστικό (“mid”), που συσχετίζει το label με μια εξωτερική πηγή σημασιολογικών δεδομένων, όπως τη [Wikidata](#). Τέλος, στην Εικ.8 παρουσιάζεται το κείμενο (“description”) της Εικ.6 μαζί με τη γλώσσα (“locale”), όπως αναγνωρίζεται στις OCR περιοχές (πράσινα πλαίσια). Ο ειδικός χαρακτήρας “\n” υποδηλώνει ότι κείμενο ανήκει σε διαφορετικό πλαίσιο. Σύμφωνα με τα παραπάνω, είναι προφανές ότι ακόμη και μία εικόνα εμπεριέχει μια πληθώρα λανθανουσών πληροφοριών, οι οποίες με τη σειρά τους μπορούν να χρησιμοποιηθούν για τον εμπλουτισμό της ανάλυσης κοινωνικών δικτύων.

```

1  {
2      "labelAnnotations": [{
3          "description": "News",
4          "mid": "/m/05jhg",
5          "score": 0.96244496
6      },
7      {
8          "description": "Newscaster",
9          "mid": "/m/025m070",
10         "score": 0.95149094
11     },
12     {
13         "description": "Spokesperson",
14         "mid": "/m/01xr66",
15         "score": 0.89320624
16     },
17     {
18         "description": "Television program",
19         "mid": "/m/0f2f9",
20         "score": 0.8240909
21     },
22     {
23         "description": "Speech",
24         "mid": "/m/09x0r",
25         "score": 0.63054645
26     }
27 ]
28 }

```

Εικ.7 Δείγμα από response του GC Vision API για labels

```

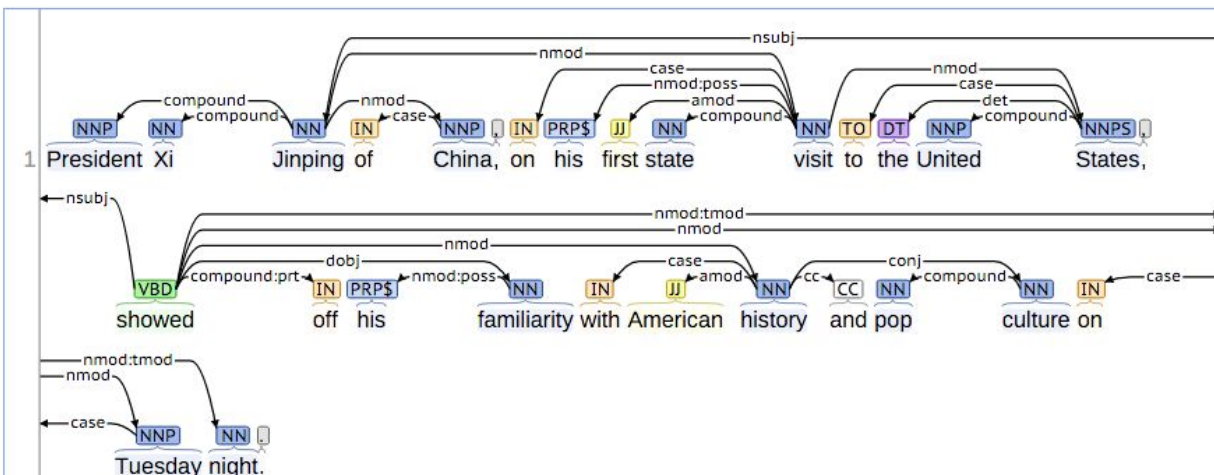
1  {
2      "textAnnotations": [{
3          "boundingPoly": {
4              "vertices": [
5                  {
6                      "x": 100,
7                      "y": 100,
8                      "x2": 100,
9                      "y2": 100,
10                 },
11                 {
12                     "x": 100,
13                     "y": 100,
14                     "x2": 100,
15                     "y2": 100,
16                 },
17                 {
18                     "x": 100,
19                     "y": 100,
20                     "x2": 100,
21                     "y2": 100,
22                 }
23             ]
24         },
25         "description": "DE\nUVE ON CNN TONIGHT BP
26         ET\nBIDEN\nSANDERS\n#CNNSOTU\nLIVE\nCORONAVIRUS
27         SPREADS FEAR AND TRANSFORMS LIFE ACROSS
28         AMERICA\nCM\n9:13 AM ET\n",
29         "locale": "en"
30     }
31 ]
32 }

```

Εικ.8 Δείγμα από response του GC Vision API για labels

## 2.2.4 Stanford CoreNLP

Το Stanford CoreNLP είναι μια βιβλιοθήκη που παρέχει δυνατότητα για επεξεργασία NLP. Σε συνέχεια του προηγούμενου OCR κειμένου που εξήχθη από το Google Vision API, έγινε χρήση του Stanford CoreNLP για περαιτέρω επεξεργασία και ανάλυση του κειμένου σε προτάσεις και λέξεις όπως και αναγνώριση του μέρους του λόγου που ανήκουν.



Εικ.9 Διάγραμμα λειτουργίας Stanford CoreNLP διάγραμμα

Με τα παραπάνω αποτελέσματα υπολογίστηκαν οι μετρικές TF, TFIDF ανά χρήστη, ανά ομάδα χρηστών και για όλους τους χρήστες.

- TF (Term Frequency)  
Είναι ο αριθμός των εμφανίσεων ενός όρου  $t$  στο έγγραφο  $d$ .
- TF-IDF (Term frequency–Inverse document frequency)  
Είναι το γινόμενο των  $TF * IDF$ , όπου το IDF υπολογίζεται ως εξής:

$$idf(t, D) = \log \frac{N}{|\{d \in D : t \in d\}|}$$

όπου

$N$ : ο συνολικός αριθμός των εγγράφων

$|\{d \in D : t \in d\}|$ : Ο αριθμός των εγγράφων στα οποία εμφανίζεται ο όρος  $t$ .

Μια υψηλή τιμή TF-IDF δείχνει μεγάλη συχνότητα ενός όρου, σε μικρό αριθμό εγγράφων. Με αυτό τον τρόπο φιλτράρονται “κοινοί όροι” που εμφανίζονται συχνά σε όλα τα έγγραφα και έχουν χαμηλή σημασιολογική αξία.

## 2.3 Δεδομένα Twitter και Συλλογή Twitter Images

Σε αυτήν την ενότητα, παρέχουμε πολύτιμες πληροφορίες σχετικά με τη διαδικασία συλλογής των δεδομένων από το Twitter, που χρησιμοποιούνται στη μελέτη μας. Συγκεκριμένα, περιγράφουμε τη μεθοδολογία για τον προσδιορισμό λογαριασμών που σχετίζονται με τον τομέα της πολιτικής, καθώς και τη διαδικασία συλλογής, ανάλυσης και αποθήκευσης του περιεχομένου και των μηνυμάτων τους.

### 2.3.1 Προσδιορισμός των λογαριασμών Twitter

Όπως περιγράφεται και στο [3], προσδιορίστηκαν 44 θεματικές κατηγορίες για την επισήμανση των λογαριασμών Twitter. Οι 40 από αυτές προέρχονται από το DBpedia, ενώ οι υπόλοιπες τέσσερις προστεθηκαν από τους δημιουργούς της υπηρεσίας InfluenceTracker για λόγους πληρότητας. Από τις 40 κατηγορίες οι 36 έχουν βασιστεί στις τιμές της σημασιολογικής ιδιότητας "rdf:type" της [DBpedia](#), υποδεικνύοντας ότι είναι στιγμιότυπο μιας οντολογικής κλάσης, ενώ οι υπόλοιπες τέσσερις από τις τιμές της ιδιότητας "dbr:position", που αντιπροσωπεύει την πολιτική θέση. Αυτή η ιδιότητα χρησιμοποιήθηκε για την δημιουργία νέας κατηγορίας, όταν δεν μπορούσαν να βρεθούν οι κατάλληλες στην DBpedia.

Είναι συνηθισμένο το γεγονός πως διαφορετικά σημασιολογικά σχήματα χρησιμοποιούν διαφορετικούς σημασιολογικούς ορισμούς, για ίδιες ή παρόμοιες έννοιες. Για παράδειγμα οι κλάσεις `dbo:Politician`, `yago:Politician110450303` και `umbel-rc:Politician` αναφέρονται στην ίδια έννοια, αυτή του πολιτικού. Ως εκ τούτου, για να βρεθούν οι κυρίαρχες θεματικές ενότητες, χρησιμοποιήθηκαν οι κλάσεις (οι τιμές της ιδιότητας "rdf: type" δηλαδή) που μπορούν να ομαδοποιηθούν στην ίδια θεματική ενότητα, αλλά και τη συχνότητα εμφάνισής τους. Επιπλέον, γενικές κατηγορίες όπως: `owl:Thing`, `dbo:Agent`, και `foaf:Person` δεν λήφθηκαν υπόψιν.

Μια σημαντική προϋπόθεση για τον προσδιορισμό αυτών των θεματικών κατηγοριών είναι η εύρεση της κατάλληλης οντότητας από την DBpedia (URI resource), που περιγράφει καλύτερα τον εν λόγω λογαριασμό Twitter. Για αυτό τον σκοπό στο [3] 2 σχετικές μεθοδολογίες αναπτύχθηκαν.

### 2.3.2 Συλλογή Δεδομένων

Στο [13] περιγράφεται το οντολογικό σχήμα "InfluenceTracker" και όλες οι σχετικές τεχνολογίες για μετατροπή των οντοτήτων του Twitter σε διασυνδεδεμένα δεδομένα (Linked Data) όπως λογαριασμούς, χρήστες, αναφορές (mentions), απαντήσεις (replies), hashtags, φωτογραφίες, διευθύνσεις URL, καθώς και κοινωνικές συσχετίσεις και χαρακτηριστικά.

Αυτό το σχήμα είναι η βάση της πλατφόρμας του InfluenceTracker. Τα σημασιολογικά δεδομένα αποθηκεύονται και ανακτώνται μέσω ερωτημάτων SPARQL, δίνοντας τη δυνατότητα για προηγμένες αναζητήσεις. Η πλατφόρμα συνδυάζει χρήση σχεσιακής βάσης (MySQL Server), RDF store (Open Link Virtuoso Server) και βάση NoSQL (CouchDB).

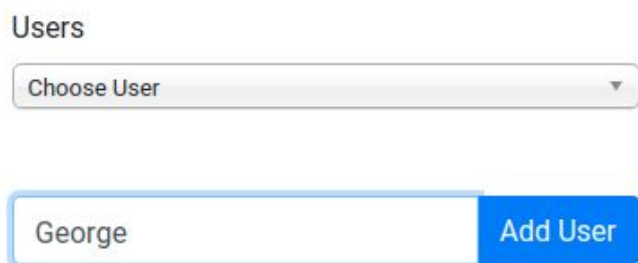
## 2.4 Ανάλυση χρήσης εφαρμογής

Όπως έχει προαναφερθεί (Ενότητα 2.2.1) η εφαρμογή προσφέρει δύο βασικές λειτουργίες. Παρακάτω γίνεται παρουσίαση χρήσης της εφαρμογής.

### 2.4.1 Upload Tab

Το upload tab περιέχει όλα τα UI components που επιτρέπουν στον χρήστη να ανεβάζει αρχεία εικόνας στην εφαρμογή, αλλά και να βλέπει όσα ήδη υπάρχουν. Το tab αποτελείται από 4 μέρη όπως παρουσιάζονται παρακάτω.

Εισαγωγή - Επιλογή Χρήστη

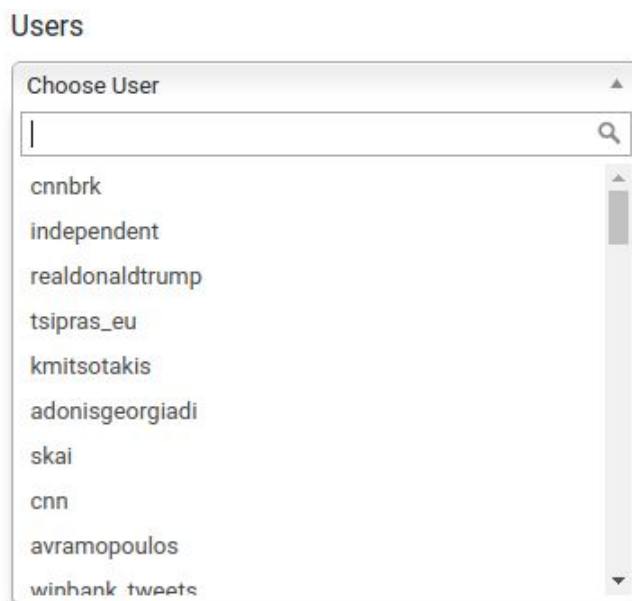


Users

Choose User

George Add User

α)



Users

Choose User

cnnbrk  
independent  
realdonaldtrump  
tsipras\_eu  
kmitsotakis  
adonisgeorgiadi  
skai  
cnn  
avramopoulos  
winhank tweets

β)

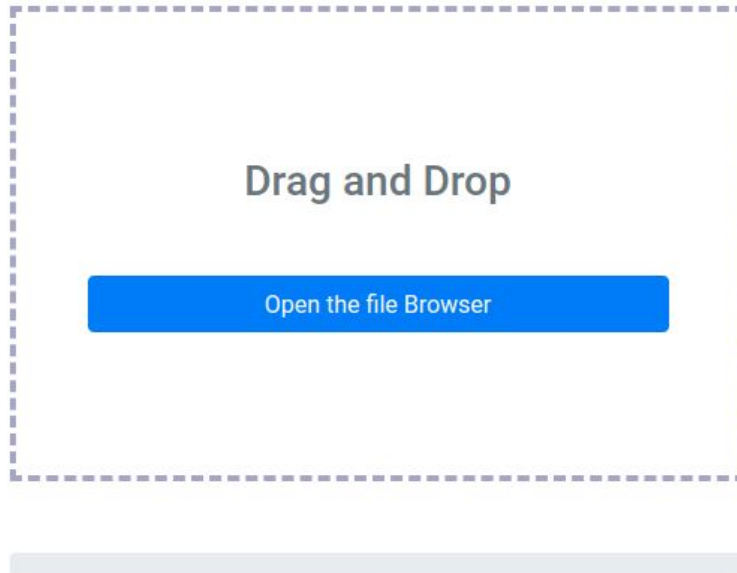
Εικ.10 Διεπαφή επιλογής χρήστη



Ο χρήστης επιλέγει με ποιον λογαριασμό θέλει να συσχετίσει το αρχείο που ανεβάζει (Εικ. 10.α), είτε επιλέγοντας έναν από τους διαθέσιμους από την dropdown λίστα (Εικ. 10.β), είτε προσθέτοντας έναν καινούριο (Εικ. 10.α).

## Μεταφόρτωση Αρχείου

Για να ανεβάσουμε ένα αρχείο στην εφαρμογή έχουμε δύο τρόπους:



Εικ.11 Διεπαφή μεταφόρτωσης αρχείων

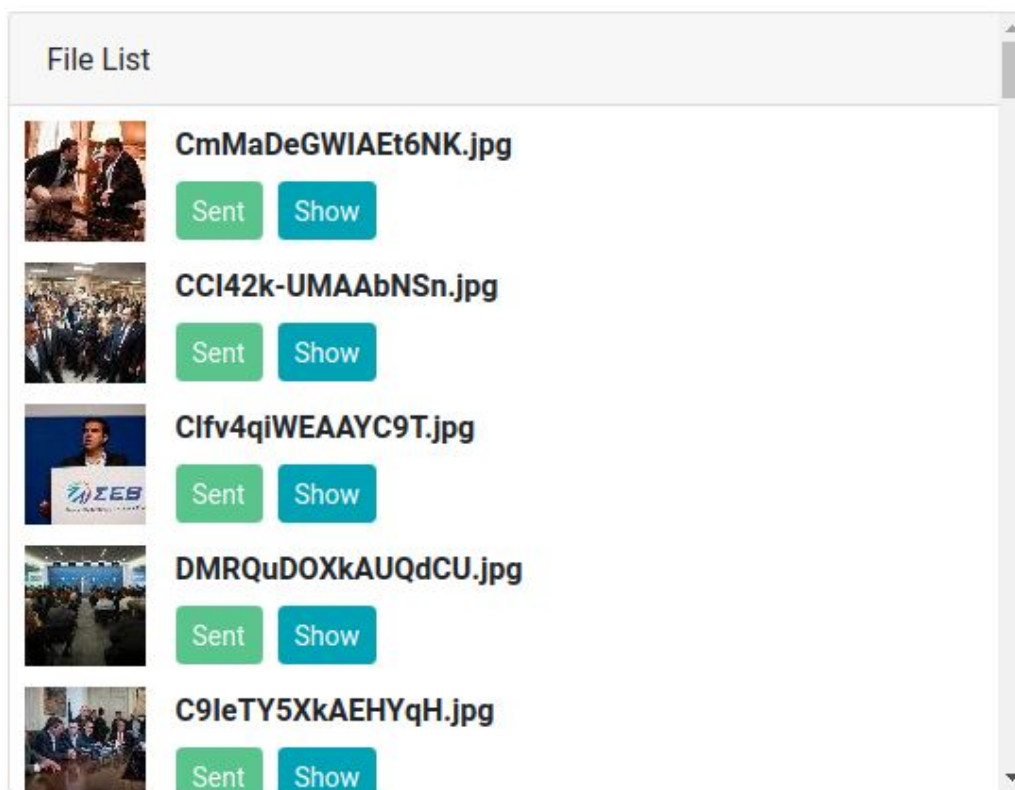
1. Ο χρήστης πατώντας το κουμπί “Open the file Browser” μπορεί να επιλέξει ένα τοπικό αρχείο από τον υπολογιστή του. Επίσης παρέχεται και η δυνατότητα για “Drag and Drop” όπου ο χρήστης ανεβάζει ένα αρχείο απλά μεταφέροντας το πάνω στην συγκεκριμένη περιοχή.
2. Με την παρακάτω φόρμα ο χρήστης έχει τη δυνατότητα να κάνει “copy paste” τα URLs αρχείων που θέλει να ανεβάσει και η εφαρμογή αναλαμβάνει να κάνει τη διαδικασία.

### Add Urls

Send

Εικ.12 Διεπαφή μαζικής υποβολής

### Λίστα Αρχείων




Εικ.13 Διεπαφή λίστας αρχείων

Η λίστα αρχείων περιέχει τα αρχεία που είναι ήδη ανεβασμένα στην εφαρμογή. Επίσης παρέχει τις παρακάτω λειτουργίες:

1. Send/Sent για την αποστολή του αρχείου στο Google Vision API ή ένδειξη ότι έχει ήδη σταλεί.

2. Show που παρουσιάζει σε pop-up παράθυρο τα αποτελέσματα του Google Vision API, όπως φαίνεται παρακάτω.

**Google Response**
✕



Description	Score
Soil	0.95561236
Field	0.9467015
Agricultural machinery	0.9086913
Vehicle	0.8975245
Agriculture	0.87346137
Plough	0.8439052
Transport	0.83513474
Farm	0.8062798
Tractor	0.76767075
Crop	0.74094105

Showing 1 to 10 of 10 entries      Previous    1    Next

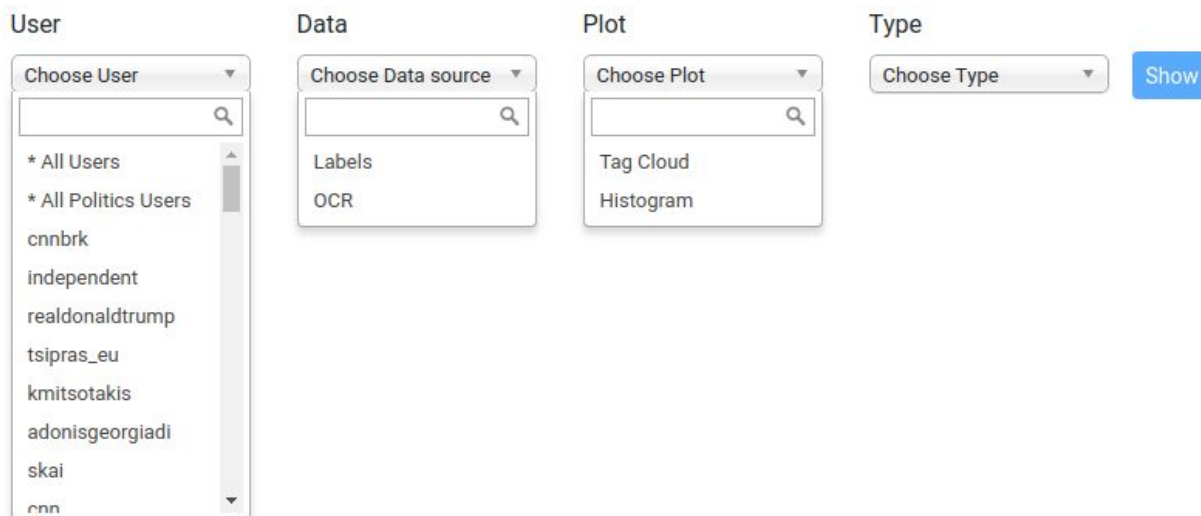
**OCR Text**

ευρώ  
 για το 2017  
 50 Περιφερειακό Συνέδριο  
 για την Παραγωγική Ανασυγκρότηση.

Εικ. 14 Διεπαφή επισκόπησης δεδομένων

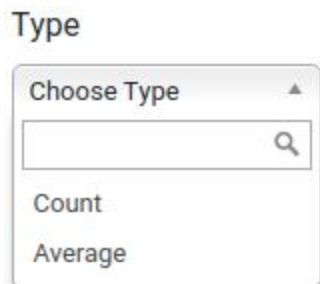
## 2.4.2 Statistics Tab

Το statistics tab δίνει τη δυνατότητα στον χρήστη να εξάγει μια σειρά από χρήσιμα στατιστικά στοιχεία. Ο χρήστης επιλέγει από τις dropdown λίστες το είδος των στατιστικών που θέλει.

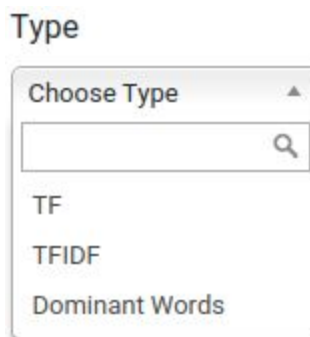


Το Type εξαρτάται από το Data source:

1. labels



2. OCR



Εικ.15 Διεπαφή επιλογής γραφημάτων

1. Κατά την επιλογή χρήστη μπορούμε να επιλέξουμε είτε έναν μεμονωμένο, είτε κάποια από τις προϋπολογισμένες κατηγορίες χρηστών π.χ. All Users, All Politics Users.
2. Από τις επιλογές Data επιλέγουμε είτε labels, είτε OCR.
3. Το plot αναφέρεται στο είδος της γραφική απεικόνιση των δεδομένων.
4. Το type αναφέρεται στη μετρική που θα χρησιμοποιηθεί για την οπτικοποίηση των δεδομένων:

a. Count

Υπολογίζει με πόσα αρχεία του χρήστη συσχετίζεται το κάθε label

b. Average

Υπολογίζει τον μέσο όρο των Scores (των βαθμών εμπιστοσύνης που μας επιστρέφει το GC Vision API).

c. TF

Υπολογίζει τις λέξεις με το μεγαλύτερο term frequency (TF).

d. TFIDF

Όπως το (c), υπολογίζοντας το TFIDF.

e. Dominant Words

Υπολογίζει τις λέξεις που μεγιστοποιούν τη μετρική (Ενότητα 3.1) που έχει οριστεί.

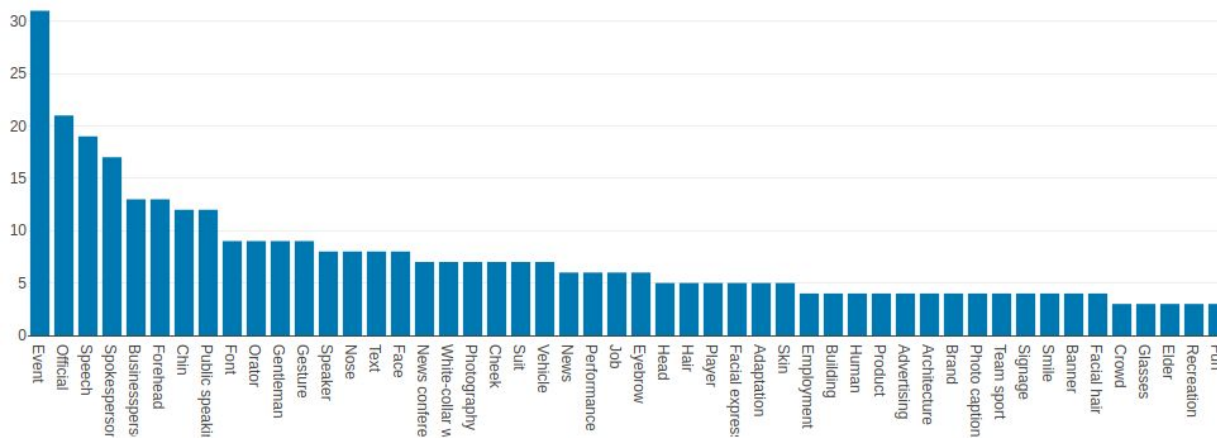
Όπως φαίνεται και στην Εικ.15, ο χρήστης μπορεί να επιλέξει δύο τύπους γραφικής απεικόνισης των δεδομένων, το Tag Cloud και το Histogram. Το Tag Cloud είναι μια οπτική αναπαράσταση δεδομένων κειμένου, στη δική μας περίπτωση είναι είτε labels είτε λέξεις από το OCR κείμενο που αναγνωρίστηκε, και η σημασία τους υποδηλώνεται από το μέγεθος και το χρώμα της γραμματοσειράς. Επομένως οι όροι που έχουν τη μεγαλύτερη τιμή στην υπό εξέταση μετρική, εμφανίζονται με μεγαλύτερη γραμματοσειρά και διαφορετικό χρώμα. Το Histogram είναι μια γραφική αναπαράσταση της κατανομής των δεδομένων. Στον κάθετο άξονα απεικονίζεται η τιμή της υπό εξέταση μετρικής και στον οριζόντιο άξονα ο εκάστοτε όρος. Τέλος, παρατηρείται ότι τα ιστογράμματα έχουν την τάση να ακολουθούν μια Power-law κατανομή, που δείχνει ότι λίγοι όροι είναι αυτοί που έχουν πολύ μεγάλη σημασία ή την περισσότερη χρήση, ενώ οι περισσότεροι όροι το αντίστροφο. Μπορούν δηλαδή να θεωρηθούν ως “κοινοί όροι”, καθώς έχουν τιμές που δείχνουν την μικρή συνεισφορά τους.

Ακολουθούν μερικά παραδείγματα αυτών των γραφικών απεικονίσεων.



Εικ.16 Tag Cloud των labels του csnbrk με μετρική count

Το Tag Cloud για το CNN, δείχνει ότι το περιεχόμενο που δημοσιεύεται εμπειριέχει λανθάνουσες πληροφορίες σχετικές με όρους όπως Event, Spokesperson, Speech και Public Speaking, που είναι αναμενόμενο για ένα ειδησεογραφικό μέσο.



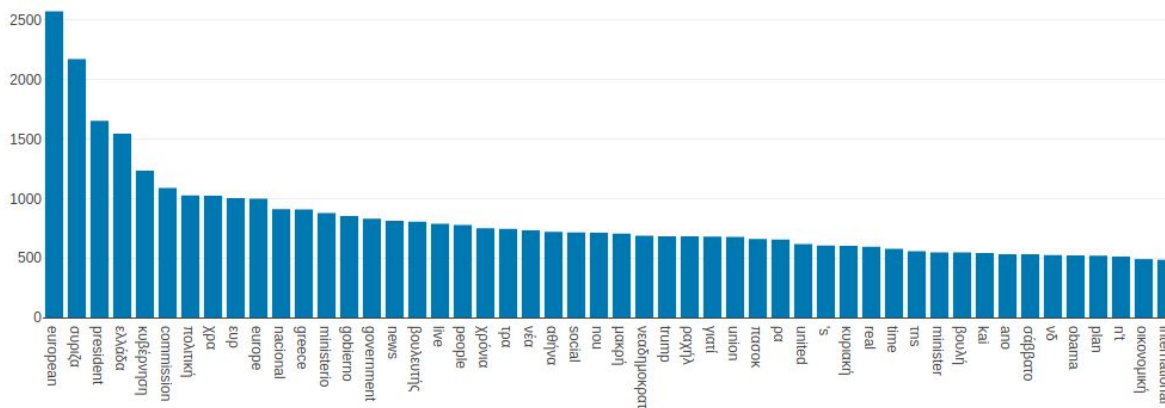
Εικ.17 Histogram των labels του csnbrk με μετρική count

Η ίδια λογική παρατηρείται και στο ιστόγραμμα του CNN για τα labels που εμφανίζονται πιο συχνά.





Εικ.20 Tag Cloud των OCR όλων των πολιτικών με μετρική Dominant Words



Εικ.21 Histogram των OCR όλων των πολιτικών με μετρική Dominant Words

Κάνοντας χρήση της μετρικής Dominant Words για το OCR κείμενο που αναγνωρίστηκε στις δημοσιευμένες εικόνες παρατηρούμε συσχέτιση με όρους όπως european, συριζα, president, κυβέρνηση, που σχετίζονται άμεσα με τον χώρο της πολιτικής.



### 3. Αποτελέσματα και Αξιολόγηση

Σε αυτή την ενότητα θα παρουσιάσουμε αναλυτικά τα αποτελέσματα από την προτεινόμενη μεθοδολογία για την αυτόματη εξαγωγή labels και κειμένου, όπως περιγράφηκε παραπάνω, η οποία εφαρμόστηκε σε ένα εστιασμένο υποσύνολο των δεδομένων.

#### 3.1 Πειράματα - Μεθοδολογία

Στη δική μας περίπτωση μελέτης θα επικεντρωθούμε σε 130 λογαριασμούς Twitter που ανήκουν στον τομέα της πολιτικής. Οι λογαριασμοί κατηγοριοποιήθηκαν περαιτέρω σε 8 υποκατηγορίες ανάλογα με το που ήταν πολιτικά προσκείμενοι, όπως φαίνεται και στον παρακάτω πίνακα. Για καθένα από αυτούς κατεβάσαμε και αποθηκεύσαμε έως 500 φωτογραφίες από τα tweets τους και τις υποβάλαμε στο Google Vision API.

Political Party	Center
Politics	Center Right
Far Right	Center Left
Right Wing	Left Wing
Place	Economics
Journalist	

Πιν.1 Κατηγορίες των λογαριασμών

Με το πέρας της διαδικασίας δημιουργήθηκε ένα σύνολο δεδομένων από 39.499 φωτογραφίες από τις οποίες έγινε εξαγωγή 4.059 μοναδικών labels. Επίσης έγινε εξαγωγή κειμένου σε 26.953 από το σύνολο των φωτογραφιών, που δείχνει ότι η πλειονότητα περιέχει και κείμενο σε ποσοστό 68,2%. Το σύνολο των μοναδικών όρων που αναγνωρίστηκαν ήταν 188.074. Συμπερασματικά σε κάθε φωτογραφία αναλογούσαν 7,4 labels και 16,5 όροι, όπως φαίνεται και στον παρακάτω πίνακα.

Οντότητα	Τιμή
Εικόνες (Σύνολο)	39.499 (40.585)
GC Vision API Labels	4.059
Labels ανά εικόνα	7,4
Εικόνες με κείμενο	26.953 (68,2%)
OCR όροι	188.074

OCR όροι ανά εικόνα	16,5
Λογαριασμοί Twitter	130
Labels λογαριασμών Twitter	11

Πιν.2 Επισκόπηση δεδομένων

Όπως ήδη αναφέρθηκε, και οι 130 λογαριασμοί σχετίζονται με το πολιτικό τομέα, συμπεριλαμβανοντας Έλληνες και ξένους πολιτικούς ή κόμματα. Στην Ελλάδα, υπάρχουν τρεις βασικές πολιτικές θέσεις «Δεξιά», «Κέντρο» και «Αριστερά», όπου το κάθε μία από αυτές μπορεί να αναλυθεί περαιτέρω. Για να πραγματοποιήσουμε την ανάλυσή μας και να αξιολογήσουμε τα περιεχόμενα πολυμέσων, ταξινομήσαμε τους λογαριασμούς στις ακόλουθες τέσσερις ομάδες:

1. “Πολιτικοί” : όπου περιλαμβάνονται και οι 130 λογαριασμοί,
2. “Αριστερά” : που περιλαμβάνει τις κατηγορίες “Left Wing” and “Center Left”,
3. “Κέντρο” : που περιλαμβάνει τις κατηγορίες “Center”, “Center Right” and “Center Left”,
4. “Δεξιά” : που περιλαμβάνει τις κατηγορίες “Far Right”, “Right Wing” and “Center Right”.

Η ανάλυσή μας επικεντρώνεται στις πληροφορίες που περιέχονται στις δημοσιευμένες φωτογραφίες αυτών των λογαριασμών. Για τον σκοπό αυτό, πραγματοποιήσαμε δύο τύπους αναλύσεων, η πρώτη ανάλυση στα labels που παρείχε το GC Vision API και η δεύτερη στο κείμενο που αναγνωρίστηκε με τεχνικές OCR.

Όπως αναφέρθηκε, κάθε label συνοδεύεται από έναν βαθμό εμπιστοσύνης που αποτελεί ένδειξη του πόσο καλά περιγράφει μια φωτογραφία. Επίσης μια επιπλέον ένδειξη της σημαντικότητας του label είναι το πόσο συχνά εμφανίζεται στις φωτογραφίες του χρήστη. Για να λάβουμε υπόψιν και τους δύο βαθμούς σημαντικότητας ορίζουμε στην Εξ.1 την Τιμή Σημαντικότητας (Dominance Metric) το γινόμενο του αθροίσματος των βαθμών εμπιστοσύνης των labels επί τον αριθμό εμφανίσεων του εκάστοτε label. Το δεύτερο υπολογίζεται στη λογαριθμική κλίμακα, με τρόπο κατάλληλο για την αποφυγή των μηδενικών τιμών.

$$Dominance Metric = \sum_{i=1}^n score \times \log_{10}(1 + n)$$

$$where n > 0 \text{ and } n \in Z$$

Εξ.1: Υπολογισμός της Τιμής Σημαντικότητας (Dominance Metric)

### 3.2 Αξιολόγηση Αποτελεσμάτων

Κάνοντας χρήση της παραπάνω μετρικής “Dominance Metric”, παρουσιάζονται τα πρώτα 15 labels για κάθε ομάδα πολιτικών.

Category	Label	Dominance	Category	Label	Dominance
P o l i t i c s	Event	53,566	R i g h t	Event	10,592
	Text	41,875		Text	9,230
	Businessperson	16,569		Businessperson	4,304
	Official	10,663		Official	2,547
	Crowd	10,479		Spokesperson	2,512
	Spokesperson	8,870		White-collar	2,356
	Public speaking	8,811		Public speaking	2,117
	White-collar worker	8,302		Photo caption	2,100
	Suit	8,256		Speech	2,051
	Speech	8,152		Suit	2,043
	Community	7,876		Crowd	1,877
	People	7,444		News+	1,736
	Photo caption	6,926		Photography	1,733
	Brand	6,816		Sky	1,558
	Photography	6,550		Brand	1,438
a)			b)		
Category	Label	Dominance	Category	Label	Dominance
C e n t e r	Text	6,554	L e f t	Event	10,673
	Event	5,438		Text	7,425
	Photo caption	1,295		Businessperson	2,935
	Businessperson	1,252		Crowd	2,566
	Crowd	1,232		Public speaking	2,134
	Public speaking	1,155		Spokesperson	1,997
	Photography	960		People	1,731
	Sky	850		Speech	1,668
	People	835		Official	1,560
	Spokesperson	787		Photo caption	1,465
	Brand	783		Audience	1,391
	Convention	771		Community	1,389
	Community	743		White-collar	1,276
	Audience	725		Photography	1,256
	Speech	717		Convention	1,224
c)			d)		

Εικ.22 Τα πρώτα 15 labels ανά κατηγορία

Όπως φαίνεται από τους παραπάνω πίνακες όλες οι εικόνες από όλες τις ομάδες πολιτικών έχουν υψηλή συσχέτιση με τα labels “Text” και “Photo caption”, λόγω της υψηλής περιεκτικότητας τους σε λεκτικό περιεχόμενο. Επίσης, επειδή οι πολιτικοί είναι μια κατηγορία ανθρώπων που δίνουν πολλές συνεντεύξεις και δημόσιες ομιλίες, συσχετίζονται με labels όπως

“Event”, “Speech”, “Crowd” και “Spokesperson”. Τέλος, λόγω των κανόνων ένδυσης που συνηθίζονται στο χώρο της πολιτικής παρατηρούμε και labels όπως “Businessperson” και “White-collar worker”.

Μια πιο εμπειριστατωμένη ανάλυση δείχνει ότι 9 από τα 15 (60%) πιο συχνά labels εμφανίζονται και στις 4 κατηγορίες πολιτικών, όπως επίσης ότι 14 από τις 15 είναι κοινές σε τουλάχιστον 3 κατηγορίες. Μόνο το label “News” εμφανίζεται μία φορά (επισημαίνεται με πράσινο Εικ.22), ως αποτέλεσμα συνεντεύξεων σε ειδησεογραφικούς σταθμούς (βλ. Εικ.6). Επίσης για τους πολιτικούς στις ομάδες “Right” και “Center” υπάρχει συσχέτιση με το label “Sky”, ενώ για αυτούς της ομάδας “Left” με το label “Convention”. Η έρευνα μας έδειξε ότι οι μεν πρώτοι έκαναν τις ομιλίες τους σε ανοιχτούς χώρους, ενώ οι δεύτεροι σε συνεδριακούς χώρους. Η παραπάνω ανάλυση επιβεβαιώνει την ύπαρξη ενός μοτίβου στο περιεχόμενο που μοιράζονται οι χρήστες ανάλογα με την ομάδα στην οποία ανήκουν.

Στον επόμενο πίνακα παρουσιάζονται οι 20 πιο συχνόι όροι που εμφανίζονται τα κείμενα που εξήχθησαν με τεχνικές OCR, έχοντας αφαιρέσει τις “κοινές λέξεις” (stopwords) (π.χ. άρθρα, προθέσεις, κλπ).

Category	Term	Dominance	Category	Term	Dominance
P o l i t i c s	europaean	2789	R i g h t	μακρή	716
	συριζα	2238		ραχήλ	701
	president	1742		ελλάδα	633
	ελλάδα	1597		νεαδημοκρατια	603
	eu	1439		συριζα	568
	κυβέρνηση	1288		κυβέρνηση	539
	commission	1160		βουλευτής	379
	χώρα	1052		ελληνες	375
	europa	1045		@kmitsotakis	362
	πολιτική	1044		europaean	340
	greece	954		νέα	307
	nacional	947		δημοκρατια	299
	ministerio	920		πολιτική	296
	world	902		νδ	289
	government	883		χρόνια	286
	gobierno	877		αθήνα	257
news	845	hrs	225		
βουλευτής	828	greece	221		
people	812	τσίπρα	218		
trump	741	πρόεδρος	215		
a)			b)		
Category	Term	Dominance	Category	Term	Dominance
C e n t e r	πασοκ	435	L e f t	συριζα	1110
	κυβέρνηση	334		πασοκ	466
	ελλάδα	322		europaean	438
	europaean	297		greece	420
	συριζα	258		ελλάδα	419
	χώρα	248		europa	333
	αθήνα	248		κυβέρνηση	331
	ποταμι	246		αθήνα	262
	greece	239		χώρα	251
	βουλευτής	225		eu	245
	χρόνια	201		βουλευτής	245
	europa	186		athens	237
	eu	171		νέα	231
	βουλή	171		χρόνια	227
	πολιτική	156		αριστερας	203
	χώρας	154		πολιτική	203
δημιουργία	151	enoitha	192		
athens	146	υπουργός	190		
νέα	144	συζήτηση	171		
θέμα	141	γιανης	167		
c)			d)		

Εικ.23 Οι 20 πιο συχνοί όροι με βάση τη μετρική Dominance Metric ανά κατηγορία

Η μετρική που χρησιμοποιήθηκε είναι μια παραλλαγή της ήδη υπάρχουσας “Dominance Metric”, όπου το άθροισμα των βαθμών εμπιστοσύνης αντικαταστάθηκε με τον αριθμό των φωτογραφιών που εμφανίζεται ο όρος. Πριν τον υπολογισμό της μετρικής, αφαιρέσαμε τις “κοινές λέξεις” στα Ελληνικά, Αγγλικά, Ισπανικά, Γερμανικά και Γαλλικά. Η ανάλυση έδειξε ότι 50 από τους 80 (62,5%) επικρατέστερους όρους εμφανίζονται σε πολλαπλές κατηγορίες, ενώ οι μοναδικοί όροι εμφανίζονται με πράσινο. Καθώς οι περισσότεροι λογαριασμοί των κοινωνικών δικτύων ανήκουν σε Έλληνες πολιτικούς, η πλειοψηφία των κυρίαρχων όρων εμφανίζονται στα

Ελληνικά. Επίσης εμφανίζονται και όροι στα αγγλικά (trump, commission) και στα ισπανικά (gobierno, ministerio), ωστόσο σχετίζονται άμεσα με τον χώρο της πολιτικής.

Περαιτέρω ανάλυση δείχνει μια ισχυρή σχέση ανάμεσα στους κυρίαρχους όρους και τις πολιτικές θέσεις σε κάθε μια κατηγορία. Δυο ενδεικτικά παραδείγματα είναι η εμφάνιση των όρων “@kmitsotakis”, “νεαδημοκρατια”, “νεα”, και “δημοκρατια” για το κόμμα της Νέας Δημοκρατίας στο χώρο της δεξιάς και οι όροι “συριζα”, “αριστερας” και “γιαννης” για το κόμμα του Σύριζα στο χώρο της αριστεράς.

Όπως και στην περίπτωση των labels, η ανάλυση του κειμένου των εικόνων έδειξε την ύπαρξη ομοιοτήτων στην λανθάνουσα πληροφορία που εμπεριέχουν λογαριασμοί που ανήκουν στον ίδιο τομέα

### 3.3 Συμπεράσματα και Μελλοντικές Επεκτάσεις Εφαρμογής

Σε αυτή τη διπλωματική εργασία προτείναμε μια μεθοδολογία που εξάγει πληροφορίες που περιέχονται σε εικόνες στο ΚΔ Twitter, κάνοντας χρήση του GC Vision API, με σκοπό τον εμπλουτισμό της ανάλυσης κοινωνικών δικτύων με σημασιολογικό περιεχόμενο και πληροφορίες κειμένου. Επιπλέον, με χρήση τεχνικών αναγνώρισης κειμένου OCR, έγινε εξαγωγή του κειμένου που απεικονίζεται στις εικόνες. Ανάλυση στην εξαγόμενη πληροφορία, έδειξε την ύπαρξη κοινών χαρακτηριστικών (patterns) στις εικόνες που δημοσιεύονται από λογαριασμούς που ανήκουν στις ίδιες ή επικαλυπτόμενες κοινότητες (communities). Επίσης η μελέτη δείχνει ότι το περιεχόμενο που δημοσιεύουν οι χρήστες, οι πληροφορίες που εμπεριέχονται στις εικόνες, οι πληροφορίες κειμένου και τα διασυνδεδεμένα δεδομένα (linked data) αποτελούν πολύτιμη πηγή πληροφοριών, οι οποίες μέχρι τώρα αγνοούνταν. Τέλος δημοσιεύουμε το σύνολο των δεδομένων που δημιουργήθηκαν κατά την εκπόνηση της διπλωματικής εργασίας και τον πηγαίο κώδικα της εφαρμογής στο δημόσιο αποθετήριο <https://github.com/gtheofilou/gvapi> για περαιτέρω χρήση και αξιολόγηση από την ερευνητική κοινότητα.

Η διπλωματική μας εργασία διαφοροποιείται σε σχέση με τις μελέτες που αναφέρονται στην βιβλιογραφία, με δύο τρόπους. Πρώτον, λαμβάνονται υπόψιν οι πληροφορίες που εξάγονται από τις εικόνες, με αποτέλεσμα να αποκαλύπτονται κοινά χαρακτηριστικά σε λογαριασμούς που ανήκουν σε ίδιους ή συγγενείς τομείς. Δεύτερον, η ύπαρξη σημασιολογικών δεδομένων τόσο σε επίπεδο λογαριασμών όσο και σε επίπεδο περιεχομένου, μπορεί να αξιοποιηθεί στον προσδιορισμό παρόμοιων λογαριασμών, στην ανάλυση κοινωνικών προτύπων και στη δημιουργία προφίλ χρηστών.

Στο μέλλον, σκοπεύουμε να επεκτείνουμε τα στατιστικά που προσφέρει η εφαρμογή, προσθέτοντας διαδικασία ομαδοποίησης (clustering) των labels με χρήση του αλγορίθμου k-means. Τέλος, θα θέλαμε να αξιολογήσουμε τον βαθμό συσχέτισης των labels των εικόνων

του GC Vision API, του κειμένου των tweets που τις περιέχουν, καθώς και του εξαγόμενου κειμένου με τεχνικές OCR.

## Βιβλιογραφία

- [1] K. Lee, “How to Tweet - New research on what drives engagement”, Buffer, 2016, <https://buffer.com/resources/twitter-data-1-million-tweets>. Accessed 17 March 2020
- [2] A. Vempala, and D. Preotiuc-Pietro, “Categorizing and Inferring the Relationship between the Text and Image of Twitter Posts”, Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 2019, pp. 2830-2840. DOI: <http://dx.doi.org/10.18653/v1/P19-1272>
- [3] G. Razis, I. Anagnostopoulos, and P. Saloun, “Thematic labeling of Twitter accounts using DBpedia properties”, 11th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP), Thessaloniki, Greece, 2016, pp. 106-111. DOI: <https://doi.org/10.1109/SMAP.2016.7753393>
- [4] T. Chen, D. Lu, M.-Y. Kan, and P. Cui, “Understanding and classifying image tweets”, In Proceedings of the 21st ACM international conference on Multimedia (MM '13), Association for Computing Machinery, New York, NY, USA, 2013, pp. 781-784. DOI: <https://doi.org/10.1145/2502081.2502203>
- [5] T. Sakai, K. Tamura, H. Kitakami and T. Takezawa, “Photo image classification using pre-trained deep network for density-based spatiotemporal analysis system”, IEEE 10th International Workshop on Computational Intelligence and Applications (IWCIA), 2017, pp. 207-212. DOI: <https://doi.org/10.1109/IWCIA.2017.8203586>
- [6] S.Sakaki, Y. Miura, X. Ma, K. Hattori, and T. Ohkuma, “Twitter User Gender Inference Using Combined Analysis of Text and Image Processing”. Proceedings of the Third Workshop on Vision and Language, 2014, pp. 54-61. DOI: <https://doi.org/10.3115/v1/W14-5408>
- [7] B. Ferwerda, and M. Tkalcic, “Predicting Users’ Personality from Instagram Pictures: Using Visual and/or Content Features?”, In Proceedings of the 26th Conference on User Modeling, Adaptation and Personalization (UMAP '18), Association for Computing Machinery, 2018, pp. 157-161. DOI: <https://doi.org/10.1145/3209219.3209248>
- [8] A. Alfarrarjeh, S. Agrawal, S. H. Kim and C. Shahabi, “Geo-Spatial Multimedia Sentiment Analysis in Disasters,” 2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA), Tokyo, Japan, 2017, pp. 193-202. DOI: <https://doi.org/10.1109/DSAA.2017.77>

- [9] S. Zannettou, B. Bradlyn, E. D. Cristofaro, G. Stringhini, and J. Blackburn, “Characterizing the Use of Images by State-Sponsored Troll Accounts on Twitter”, 14th International AAAI Conference on Web and Social Media (ICWSM 2020), 2020, In Press.
- [10] R. Boselli, M. Cesarini, F. Mercurio, M. Mezzanzanica, and A. Vaccarino, “A Pipeline for Multimedia Twitter Analysis through Graph Databases: Preliminary Results”, In Proceedings of the 6th International Conference on Data Science, Technology and Applications (DATA), 2017, pp. 343-349. DOI: <https://doi.org/10.5220/0006490703430349>
- [11] M. Noaen, and B. H. Far, “Social media analysis for traffic management”, In Proceedings of the 14th International Conference on Global Software Engineering (ICGSE '19), IEEE Press, 2019, pp. 72-73. DOI: <https://doi.org/10.1109/ICGSE.2019.00012>
- [12] A. O’Neal, B. Rodgers, J. Segler, D. Murthy, N. Lakuduva, M. Johnson, and K. K. Stephens, “Training an Emergency-Response Image Classifier on Signal Data”, 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL, USA, 2018, pp. 751-756. DOI: <https://doi.org/10.1109/ICMLA.2018.00119>
- [13] G. Razis, and I. Anagnostopoulos, “Semantifying Twitter: The Influence Tracker Ontology”, 9th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP), Corfu, Greece, 2014, pp. 98-103. DOI: <https://doi.org/10.1109/SMAP.2014.23>
- [14] G. Razis, and I. Anagnostopoulos, “Discovering similar Twitter accounts using semantics”, Engineering Applications of Artificial Intelligence, vol. 51, 2016, pp. 37-49. DOI: <https://doi.org/10.1016/j.engappai.2016.01.015>