

PHONY INFRASTRUCTURE

Αυτόματη δημιουργία Feature-Agnostic συνόλων δεδομένων για ανίχνευση ψευδών ειδήσεων στα μέσα κοινωνικής δικτύωσης

Danae Pla Karidi

Research assistant

IMSI



AOHNA

Έρευνα & Καινοτομία
Τεχνολογίες Πληροφορίας

Τα προαναφερθέντα έργα υλοποιήθηκαν στο πλαίσιο της Πρόσκλησης με κωδικό 031 με τίτλο «ΔΡΑΣΗ ΣΤΡΑΤΗΓΙΚΗΣ ΑΝΑΠΤΥΞΗΣ ΕΡΕΥΝΗΤΙΚΩΝ ΚΑΙ ΤΕΧΝΟΛΟΓΙΚΩΝ ΦΟΡΕΩΝ», που χρηματοδοτείται από το Επιχειρησιακό Πρόγραμμα «Ανταγωνιστικότητα, Επιχειρηματικότητα και Καινοτομία», ΕΣΠΑ 2014-2020 ΕΣΠΑ 2014-2020, με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης (Ευρωπαϊκό Ταμείο Περιφερειακής Ανάπτυξης). Αναλυτικά: "Υπολογιστικές Επιστήμες και Τεχνολογίες Δεδομένων, Περιεχομένου και Αλληλεπίδρασης" (MIS 5002437) με 3 υποέργα: Υποέργο 1 «Γλωσσικές Τεχνολογίες Ανάλυσης Περιεχομένου και Αλληλεπίδρασης», Υποέργο 2 «Τεχνολογίες Ανάλυσης Περιεχομένου στον Πολιτισμό», Υποέργο 3 «ΜΕΔΑ: Μεγάλα Δεδομένα – Προκλήσεις, Μέθοδοι και Τεχνικές Αποδοτικής Διαχείρισης» και "Καινοτόμος εφαρμογή του βιομηχανικού διαδικτύου των πραγμάτων σε ευφυή περιβάλλοντα" (MIS 5002434) με Υποέργο 1: "Καινοτόμος εφαρμογή του βιομηχανικού διαδικτύου των πραγμάτων σε ευφυή περιβάλλοντα".



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Ταμείο
Περιφερειακής Ανάπτυξης



ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
ΥΠΟΥΡΓΕΙΟ
ΑΝΑΠΤΥΞΗΣ ΚΑΙ ΕΠΙΧΕΙΡΗΣΕΩΝ
ΕΙΔΙΚΗ ΓΡΑΜΜΑΤΕΙΑ ΔΙΑΦΟΡΟΤΙΚΩΝ ΠΡΟΓΡΑΜΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ ΕΠΙΧΕΙΡΗΣΕΩΝ



ΕΠΑΝΕΚ 2014-2020
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΑΝΤΑΓΩΝΙΣΤΙΚΟΤΗΤΑ
ΕΠΙΧΕΙΡΗΜΑΤΙΚΟΤΗΤΑ
ΚΑΙΝΟΤΟΜΙΑ

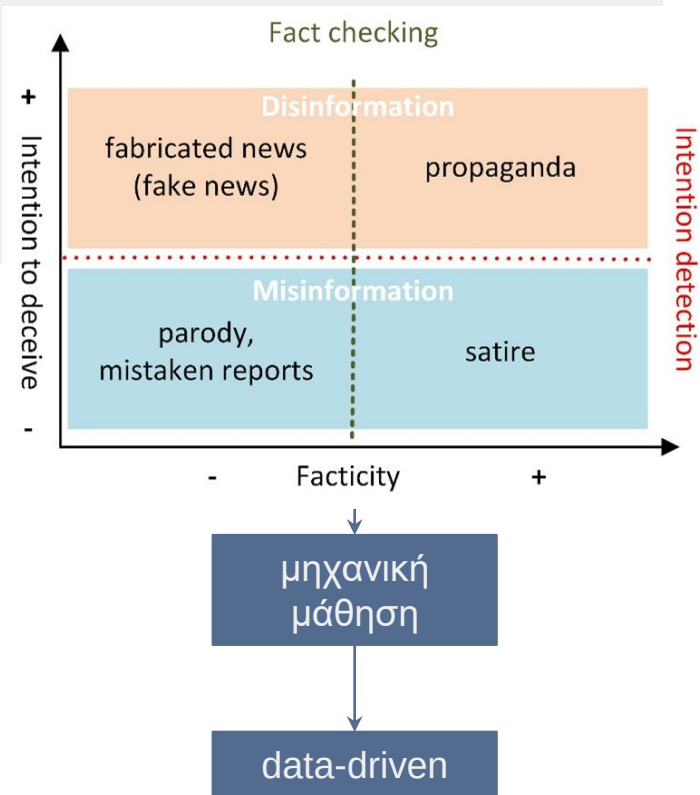


ΕΣΠΑ
2014-2020
ανάπτυξη - εργασία - αλληλεγγύη

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Αυτόματη ανίχνευση ψευδών ειδήσεων

- Fake ή Real ειδήσεις στα social? Ή όχι μόνο?
- Πόσα μοντέλα? Για ποια fakes?
- Datasets?





Υποδομή RHONY

- Αυτοματοποιεί όσο γίνεται τη διαδικασία δημιουργίας των συνόλων δεδομένων
- Παράγει silver-standard σύνολα: ground truth annotations με θόρυβο
- Παράγει σύνολα δεδομένων που επιτρέπουν την εξαγωγή υποσυνόλων ανάλογα με τις ανάγκες και καλύπτουν όλα τα features της βιβλιογραφίας!



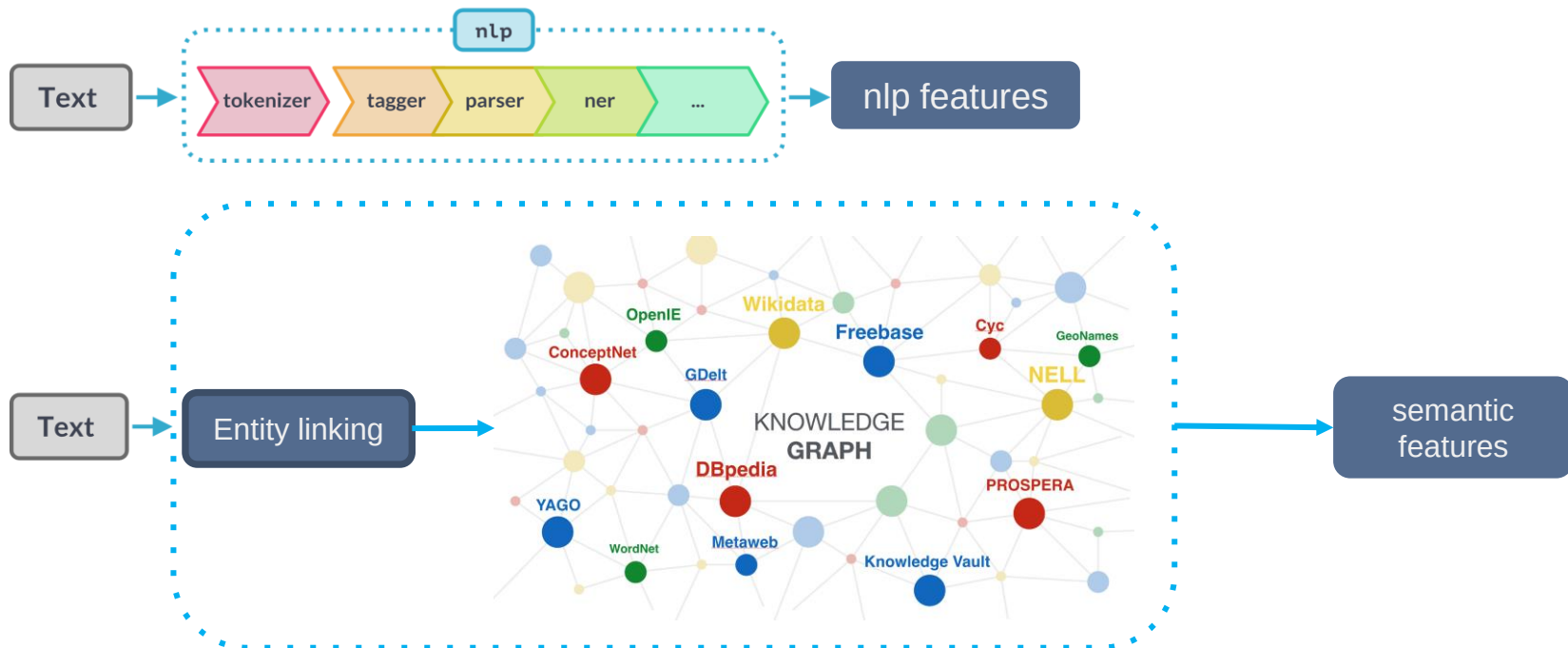


Features Ανίχνευσης

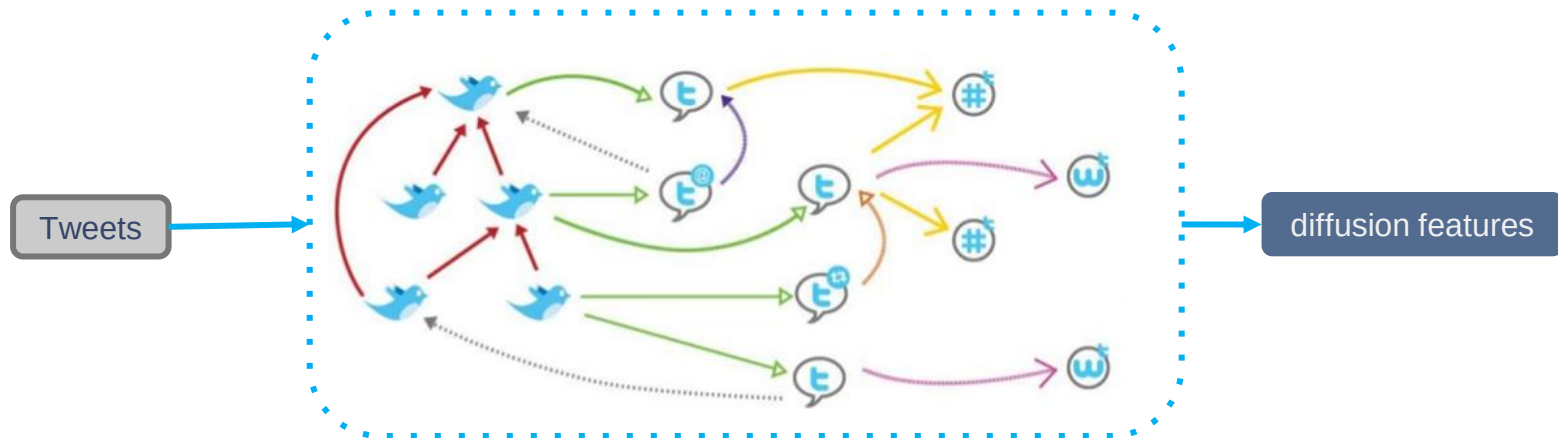
- Πλήθος και ποικιλία features
- Κάποια σύνολα δεδομένων περιέχουν τμήμα, ένα είδος ή ελλιπή features



Γλωσσικά και σημασιολογικά features

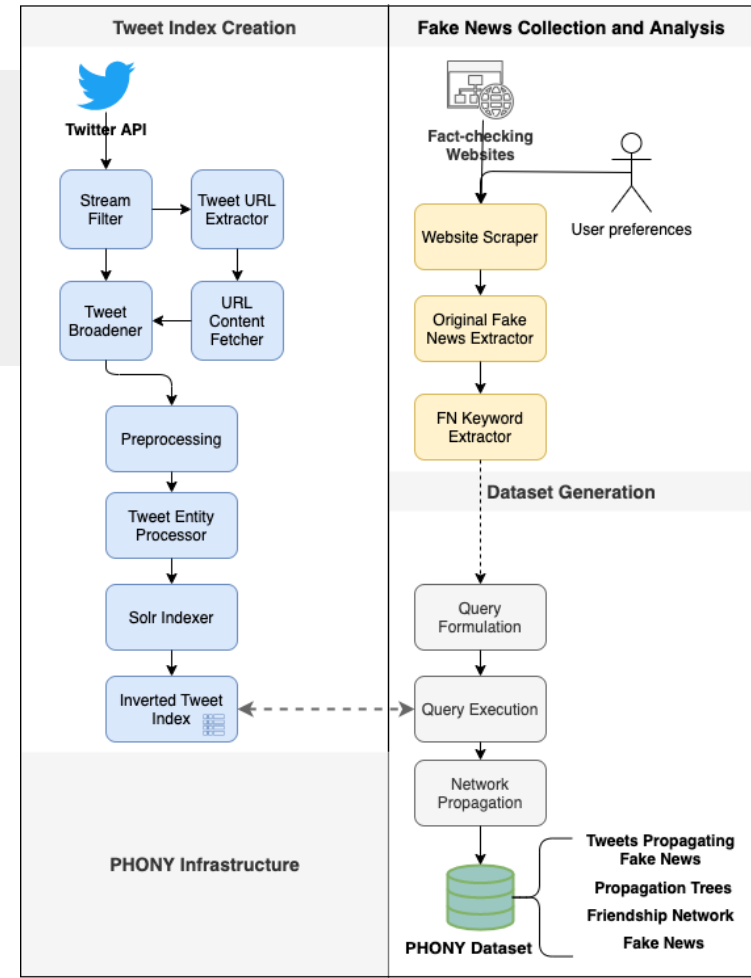


Features διάδοσης στο δίκτυο



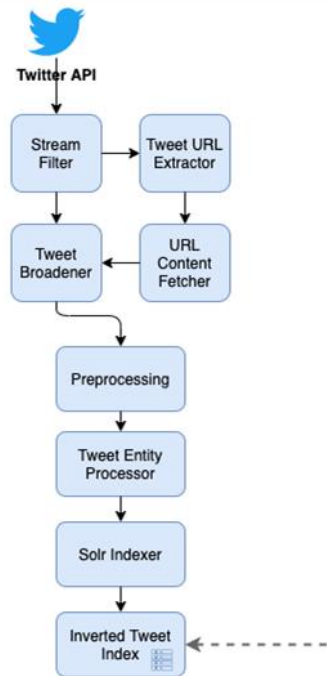
Αρχιτεκτονική

- Συνδυάζει δεδομένα από το Twitter και fact-checking υπηρεσίες
- Δημιουργία ευρετηρίου Tweets
 - incremental inverted index
- Fake News Collection and Analysis
 - με βάση κριτήρια χρήστη (τύπος είδησης, χρονική περίοδος, πηγή ειδήσεων)
- Δημιουργία συνόλου δεδομένων
 - queries στο index και εξαγωγή ίχνους διάδοσης για τις ψευδείς ειδήσεις



Tweet Index Creation

- Tweet Index: word ↔ tweet
- Tweet Broadening:
 - Κείμενο tweet + μεταδεδομένα + web δεδομένα που περιλαμβάνονται στο tweet
 - Tweet URL Extractor + URL Content Fetcher
 - Τα broadened tweets είναι αντικείμενα που περιλαμβάνουν
 - τα original πεδία του tweet: (created at, entities, extended v , favorite count, favorited, filter level, id, id str, is quote status, lang, possibly sensitive, retweet count, retweeted, source, text, timestamp ms)
 - ένα επιπλέον πεδίο: (broadened text), που περιλαμβάνει το κείμενο των συνδέσμων σε άλλες ιστοσελίδες οι οποίες περικλείονται στο tweet



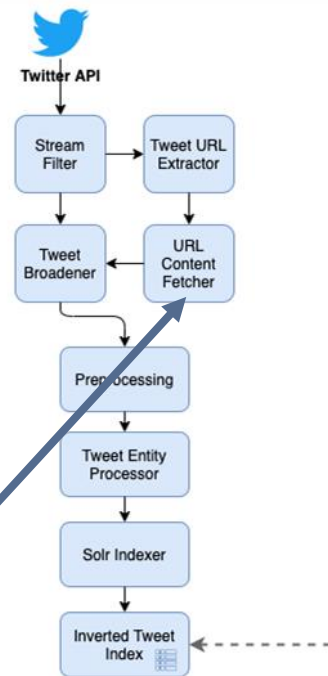
Example Tweet Broadening

- Tweet: “Impressive anticancer action on two cannabis substances
<https://t.co/YGLx59U1m5>
<https://t.co/19advJR7Hb> “
- Μήκος του tweet object πριν το broadening 3299 chars
- Μήκος του tweet object μετά το broadening 4708 chars

Δύο δραστικές ουσίες που περιέχονται στο φυτό της κάνναβης (από όπου, μεταξύ άλλων, προέρχεται η μαρικουάνα και το χασίς), έφεραν εντυπωσιακά αντικαρκινικά αποτελέσματα σε πειραματόζωα, όταν συνδυάστηκαν με παραδοσιακή ακτινοθεραπεία. Είναι η πρώτη φορά που διαπιστώνεται κάτι τέτοιο.

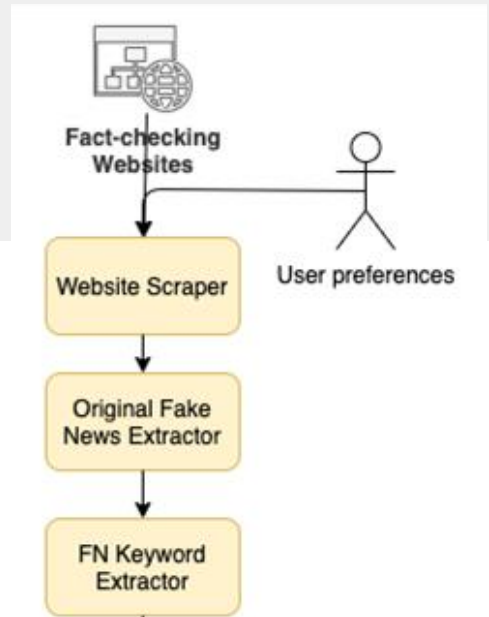
Επιστήμονες του Πανεπιστημίου Σεν Τζορτζ του Λονδίνου, με επικεφαλής τον δρ Γουάι Λιου, που έκαναν τη σχετική δημοσίευση στην επιστημονική επιθεώρηση Molecular Cancer Therapeutics, έκαναν πειράματα με ποντίκια που είχαν καρκινικούς όγκους στον εγκέφαλο.

Οι ερευνητές δοκίμασαν στο εργαστήριο τις χημικές ουσίες της κάνναβης τετραϋδροκανναβιδόλη και κανναβιδιόλη είτε μόνες τους, είτε σε διάφορους συνδυασμούς με συμβατικές θεραπείες. Τα πειράματα έδειξαν ότι οι δύο αυτές ουσίες, μαζί με ακτινοθεραπεία, πέτυχαν να «φρενάρουν» ή και ουσιαστικά να εξαφανίσουν τους όγκους στα πειραματόζωα.



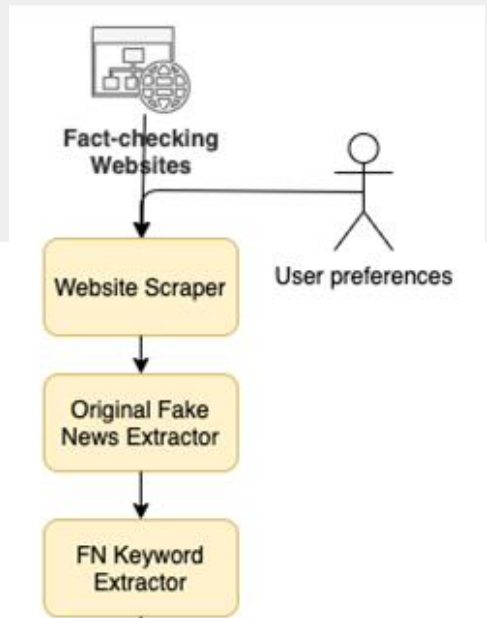
Example Fake News Collection and Analysis

- Επιλέγουμε τη χρονική περίοδο 2017
- Η πηγή μας είναι τα EllinikaHoaxes
- Επιλέγουμε τον τύπο “Fabrication”
 - Ο Website Scraper κάνει το scraping των άρθρων από τα EllinikaHoaxes και αποθηκεύει τον τίτλο και το κείμενο σε μια MongoDB.



Example fake news id=215

- Τίτλος: "Can cannabis cure cancer?"
- Ο Original Fake News Extractor αναλύει το κείμενο (NLP) και βρίσκει τις πηγές της είδησης (instances) και αποθηκεύει το HTML περιεχόμενό τους σε μια MongoDB:
 - 1. "A cannabis protocol for the treatment of cancer": <https://archive.fo/8KdYc>
 - 2. "How cannabis cures cancer": <https://archive.fo/XpqEc>
 - 3. "Medical Cannabis: How It Cures Cancer, Children's Epilepsy, And Not Only": <https://archive.fo/vQ1ub>
 - 4. "Cannabis oil and cancer": <https://archive.fo/YtWO2>
 - 5. "Ambrosia eliquids electronic cigarette replenishment fluid with CBD - Cannabis vapor has health benefits": <https://archive.fo/W8yJZ>
- Ο FN Keyword Extractor το περιεχόμενο των instances και εξαγάγει αντιπροσωπευτικές λέξεις-κλειδιά χρησιμοποιώντας το TF-IDF: cannabinoid, cancer, cannab, marijuana, treatment, thc, cell, receptor, hemp, simpson, rick, oil, oleum, plant, solvent, ambrosia, cannabidiol, liquid, extract.





TF*IDF Keyword extraction

- TF:
 - μετράει τη συχνότητα την εμφάνισης μιας λέξης σε ένα κείμενο
 - ανάλογα με το μέγεθος του κειμένου είναι πιθανό μία λέξη να εμφανίζεται πιο πολύ στα μεγάλα και λιγότερο στα μικρά κείμενα
- IDF:
 - μετράει πόσο σημαντική είναι μια λέξη
 - λαμβάνει υπόψιν το πόσο σπάνια είναι αυτή η λέξη



TF*IDF Keyword extraction

- Για κάθε λέξη w : $W_w = TF * IDF$
- Το IDF είναι υψηλότερο στις λέξεις που είναι πιο σπάνιες (mobilegeddon), ενώ όσο μικρότερη η συχνότητα ενός όρου τόσο πιο μικρό είναι και το IDF.

$$TF_{(w,j)} = \frac{(\text{Number of times term } w \text{ appears in a document})}{(\text{Total number of terms } w \text{ in the document})}$$

$$IDF_{(w)} = \log \frac{\text{Total number of documents}}{\text{Number of documents with term } w \text{ in it}}$$

Term of Interest	Number of Documents Containing that Term	Total Docs / Docs Containing Terms	IDF Values for These Terms
a	100,000,000	1	-
boat	1,000,000	100	2
mobile	100,000	1,000	3
mobilegeddon	1,000	100,000	5





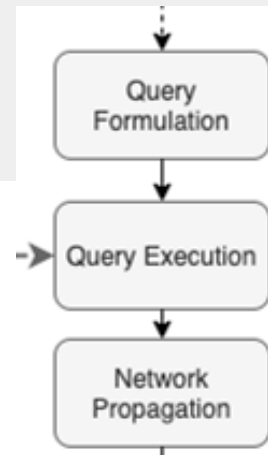
Παράδειγμα $TF*IDF$

- Θέλω να εξάγω τα top-5 keywords από μία είδηση.
- Συνολικά διαθέτω 10,000,000 κείμενα ειδήσεων στη συλλογή μου.
- Για κάθε όρο του κειμένου που εξετάζω, υπολογίζω το $(TF*IDF)$ του όρου.
- Το κείμενο της είδησης είναι 100 λέξεων και περιέχει τον όρο «γάτα» 12 φορές, επομένως
 - $TF_{cat} = 12/100 = 0.12$
- Εάν από τα 10,000,000 κείμενα ειδήσεων, τα 300,000 περιλαμβάνουν τον όρο «γάτα», τότε
 - $IDF_{cat} = \log(10,000,000/300,000) = 1.52$
- $W_{cat} = (TF_{cat} * IDF_{cat}) = 0.12 * 1.52 = 0.182$
- Τα top-5 keywords είναι αυτά με τα πέντε μεγαλύτερα W scores.



Example Fake News Collection and Analysis

- Η μονάδα Query Formulation
 - υπολογίζει τον αριθμό των όρων που πρέπει να περιέχει το query $\geq n$
 - συνδυάζει τις top-n λέξεις-κλειδιά με τον τελεστή AND
 - συνδυάζει τα queries του κάθε instance με τον τελεστή OR
- Η μονάδα Query Execution τρέχει το query στο Solr inverted index των broadened tweets και αποθηκεύει τα tweets του αποτελέσματος σε MongoDB.
- Το παραπάνω query επέστρεψε 652 broadened tweets





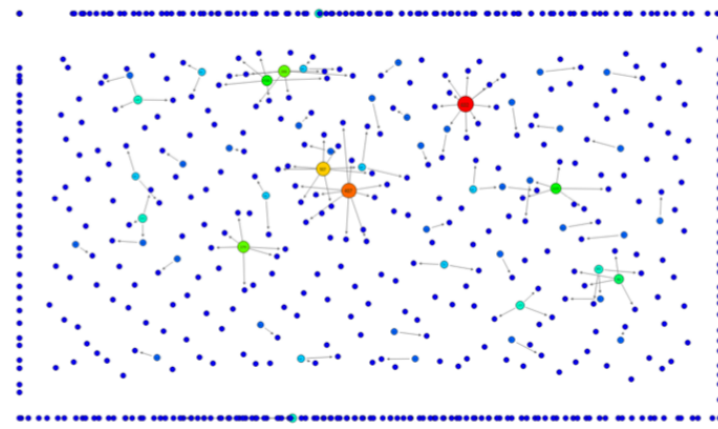
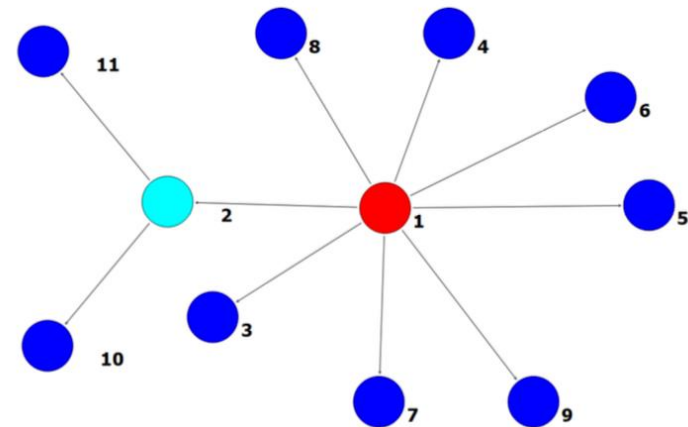
Δείγμα από τα 652 tweets

- Cannabis: A miraculous banned anti-cancer drug <https://t.co/Fwv4jsO8x2>
- Impressive anticancer action on two cannabis substances
<https://t.co/YGLx59U1m5><https://t.co/19advJR7Hb>”
- “The healing value of cannabis - A testimony <https://t.co/vbWbnq6TNp>”
- “Wake Up: The Healing Value of Cannabis - A Testimony: The Other Side ... Read More. Wake Up...
<https://t.co/uztSU8dAqC>”
- “VIDEO Greek #patients talk publicly about how they beat death with #oil #hemp. By #Andreas-Roumeliotis ... <https://t.co/eLUN7cjlQ>”
- “Medical Cannabis: How It Cures Cancer, Children’s Epilepsy, and not only <https://t.co/bZFXnicBeJ>”
- “It’s the most effective herbal blend for cancer. <https://t.co/k6vLpws20A> EMEDI’s Herbal Cancer Blend
<https://t.co/Qdtb0C9JN2>”



Example network propagation

- Ένας χρήστης κάνει retweet περιεχόμενο που έχουν ποστάρει ένας ή περισσότεροι χρήστες που ακολουθεί.
- Για κάθε retweet:
 - Βρίσκουμε ποιος followee έχει δημοσιεύσει το tweet νωρίτερα και αντιστοιχίζουμε το retweet σε αυτό.
 - Εάν αυτοί οι followees είναι περισσότεροι τότε αντιστοιχίζουμε το retweet:
 - Σε όλες αυτές τις δημοσιεύσεις ή
 - Στην πιο πρόσφατη από αυτές ή
 - Στην πιο δημοφιλή από αυτές





Dataset structure

- **fakenews_stories:** περιλαμβάνει τις ψευδείς ειδήσεις.
- **fakenews_texts:** περιλαμβάνει τα κείμενα των ειδήσεων σε χωριστά αρχεία κειμένου. Κάθε αρχείο έχει ένα instance.
- **fakenews_search_results:** περιλαμβάνει πληροφορίες για τη δημιουργία του dataset.
- **fakenews_tweets:** περιλαμβάνει τα broadened tweets είδηση σε χωριστά JSON Lines αρχεία.
- **fakenews_user_network:** περιλαμβάνει το δίκτυο των χρηστών, δηλαδή τις σχέσεις follower-followee, για όσους συμμετείχαν στη διάδοση μιας είδησης.
- **fakenews_propagation_graphs:** περιλαμβάνει τη διάδοση της κάθε είδησης ανάλογα με τη στρατηγική αντιστοίχισης. Περιλαμβάνει ένα γράφο διάδοσης ανά γραμμή.



SAMPLE GREEK Dataset

- Ελληνικά tweets από Ιανουάριο 2016 έως Απρίλιο 2018
- Ψευδείς ειδήσεις από EllinikaHoaxes με τύπο “fake news” από Ιανουάριο 2016 έως Απρίλιο 2018

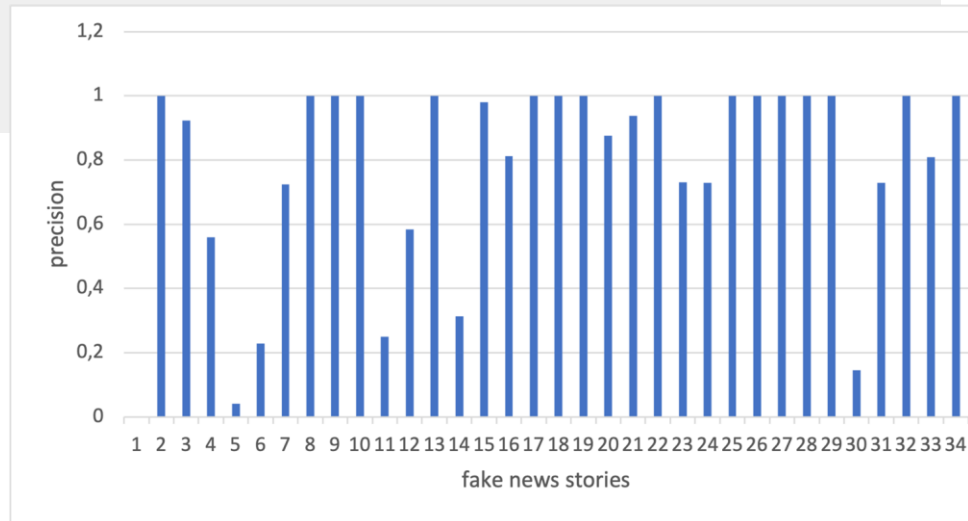
Number of tweets in Tweet Index	238,685,450
Number of unique users in Tweet Index	1,381,799
Number of external links in Tweet Index	57,264,673
Number of terms in Tweet Index	9,237,157
Size of Tweet Index	126.6 GB

Fake News Stories	677
URLs of Fake News	2,835
Tweets propagating Fake News	381,350
Unique Users propagating Fake News	21,223
Propagation Trees of Fake News	322,265
Fake News Stories that did not propagate	68
Avg. tweets per Fake News	562.58
Max. tweets per Fake News	27,072
Min. tweets per Fake News	0
Median tweets per Fake News	38



Αξιολόγηση sample Greek dataset

- silver-standard datasets
- average precision 77,56%
- 19 fake news (54%) πέτυχαν precision >90%





Ευχαριστώ!