



Without You it's
just a Tube, with
you it becomes a
profiling database

Dimitris Gritzalis

October 2014



Online Social Networks: Αναμνήσεις από το ...μέλλον

Καθηγητής Δημήτρης Γκρίτζαλης

Εργαστήριο Ασφάλειας Πληροφοριών & Προστασίας Κρίσιμων Υποδομών
Τμήμα Πληροφορικής, Οικονομικό Πανεπιστήμιο Αθηνών

Δεδομένα στα Online Social Networks



Broadcast Yourself™
Profile

- Η δομή των Online Social Networks επιτρέπει **προσωποποιημένη χρήση**
- Η πλειονότητα των δεδομένων στο διαδίκτυο **παράγονται από τους χρήστες**
- Οι χρήστες αποκαλύπτουν, αβίαστα, **στοιχεία της προσωπικότητάς τους**
- Υπάρχουν ποικίλα **κίνητρα** χρήσης τους (επαγγελματικά, διασκέδαση, επικοινωνία κά)
- Οι χρήστες **αναπαράγουν τη συμπεριφορά τους** (και) στο Διαδίκτυο^{1,2}

Έτσι συγκεντρώνεται τεράστιος όγκος δημόσια διαθέσιμων δεδομένων τύπου **Open Source Intelligence (OSINT)**, δηλαδή πληροφορίες προερχόμενες από δημόσια διαθέσιμα δεδομένα, που αξιοποιούνται μεταδιδόμενες έγκαιρα για την αντιμετώπιση ειδικών αναγκών πληροφόρησης

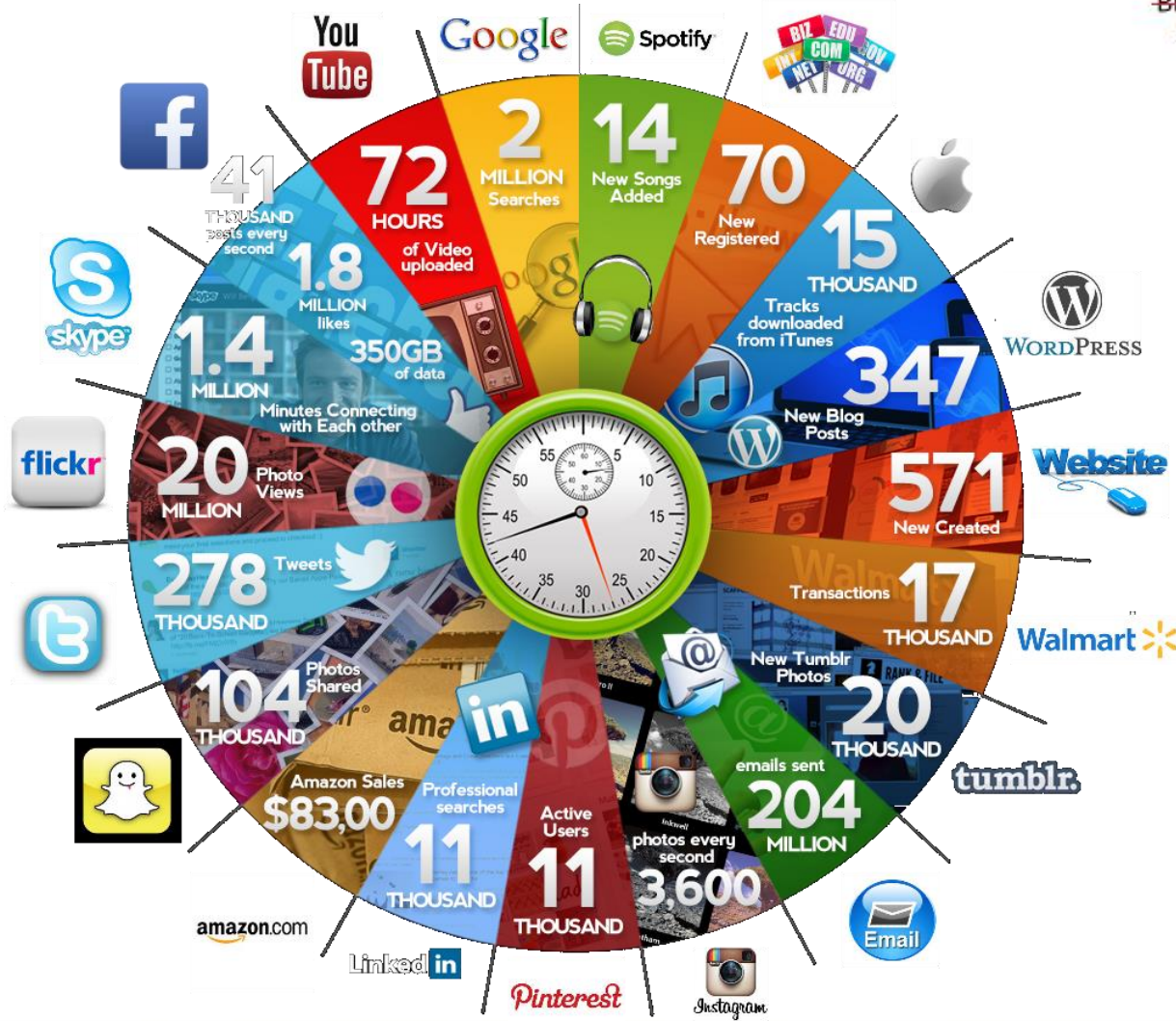
[1]. Ross C., Orr E, Sisis M., Arseneault J., Simmering M., Orr R., "Personality and motivations associated with Facebook use", *Computers in Human Behavior*, Vol. 25, pp. 578-586, 2009.

[2]. Amichai-Hamburger Y., Vinitzky G., "Social network use and personality", *Computers in Human Behavior*, Vol. 26, pp. 1289-1295, 2010.

Τι συμβαίνει online μέσα σε 1 λεπτό!



Broadcast Yourself™
Profile



- Δημιουργία μορφότυπων (profiles) *χρήσης και χρηστών για:*
 - Προσωποποίηση περιεχομένου
 - Στοχευμένη προώθηση (marketing)
 - Βελτίωση εμπειρίας χρήσης
- Αποκάλυψη *ευαίσθητων* προσωπικών δεδομένων μέσω **άμεσων** συσχετίσεων³:
 - Πολιτικές πεποιθήσεις
 - Θρησκευτικές επιλογές
 - Σεξουαλικές προτιμήσεις
- Εξαγωγή **πολιτικών πεποιθήσεων** πολιτών από το δεδομένα που αναρτούν στο YouTube (μέσω **έμμεσων** συσχετίσεων)



Πανοπτικόν: Αποκάλυψη ευαίσθητων προσωπικών δεδομένων μέσω YouTube

Πανοπτικόν: Η «ιδανική φυλακή» του J. Bentham. Ένα φανταστικό οικοδόμημα, σχεδιασμένο με τρόπο ώστε οι κατάδικοι να παρακολουθούνται διαρκώς από ένα κεντρικό σημείο.

Αποκαλύπτοντας πολιτικές πεποιθήσεις μέσω ...data mining και ...YouTube!



Broadcast Yourself™
Profile

- **Online Social Network**

YouTube

- **Δεδομένα** (Νοε 2005 - Δεκ 2012)

12.964 χρήστες, 207.377 βίντεο, 2.034.362 σχόλια

Ανωνυμοποίηση όλων των δεδομένων, για να μην υπάρχει συσχέτιση ονόματος χρήστη-πεποιθήσεων (MD5 hashes αντί ονομάτων χρηστών), με βάση το N. 2472/1997.

- **Πρόσφορο πεδίο εφαρμογής**

Πολιτικό περιεχόμενο, οπτικοακουστικά ερεθίσματα, συναισθηματική φόρτιση, ευρεία συμμετοχή χρηστών.

- **Μέθοδοι ανάλυσης**

Γραφοθεωρητική ανάλυση (Small World Phenomenon, Indegree/Outdegree Distribution, Node Loneliness)

Ανάλυση περιεχομένου (εξαγωγή συμπερασμάτων από ανάλυση σχολίων χρηστών, λεκτική ανάλυση - Opinion Mining, Machine Learning)

Ανάλυση νέφους ετικετών (συγκέντρωση ετικετών βίντεο και απεικόνιση σε νέφος ώστε να μελετηθεί το περιεχόμενο - Tag Cloud)

Εξελιγμένοι μηχανισμοί
διευκόλυνσης της συλλογής
δεδομένων χρηστών

Μια πιθανή (ή πραγματική;) μέθοδος



Broadcast Yourself™
Profile

- Ορίζονται τρεις (ενδεικτικές) **κατηγορίες ταξινόμησης**:
 - «**Ριζοσπαστικοί**», «**Ουδέτεροι**», «**Συντηρητικοί**»
 - Οι παραδοχές εξαρτώνται από το κοινωνικό πλαίσιο
 - Το αντικείμενο του πειράματος είναι μια πραγματική κοινότητα χρηστών
 - Ανακλάται ένα (θεωρητικό) ιστορικοπολιτικό πλαίσιο της Ελλάδας
- Ορίζεται (ενδεικτικά) η ακόλουθη **αντιστοίχιση**:
 - **Ριζοσπαστικοί** χρήστες:
 - Όσοι δείχνουν κεντροαριστερές ή αριστερές πολιτικές απόψεις
 - **Ουδέτεροι** χρήστες:
 - Όσοι δεν επιδεικνύουν κάποια πολιτική ταύτιση
 - **Συντηρητικοί** χρήστες:
 - Όσοι δείχνουν κεντροδεξιές ή δεξιές πολιτικές απόψεις

Μια πιθανή (ή πραγματική) μέθοδος



Broadcast Yourself™
Profile

- Τα σχόλια του YouTube εντάσσονται και ταξινομούνται στις καθορισμένες κατηγορίες πολιτικών απόψεων
 - Η ταξινόμηση των σχολίων έγινε ως ταξινόμηση κειμένου⁴
 - Εκπαίδευση λογισμικού με βάση το περιεχόμενο των σχολίων και την κατηγορία όπου εντάχθηκαν
 - Ένταξη των σχολίων σε κατηγορίες μέσω ειδικών (Κοινωνιολόγος, Πολιτικός Επιστήμονας)
- Χρήση κατάλληλου δείγματος για επιβεβαίωση εγκυρότητας μεθόδου
 - Εισάγονται προσημασμένα δεδομένα στο σύστημα
 - Ελέγχεται αν η αρχική σήμανση (από τον ειδικό) συμπίπτει με το συμπέρασμα του συστήματος
- Η πλειονότητα των δεδομένων είναι στο ελληνικό αλφάβητο
- Παράλληλη χρήση λατινικού αλφαβήτου από τους χρήστες (“greeklish”)
 - Αντιμετωπίστηκαν ως δύο διαφορετικές «γλώσσες»
 - Τα δεδομένα εκπαίδευσης των δύο «γλωσσών» ενώθηκαν σε έναν ταξινομητή

Γραφοθεωρητική ανάλυση χρηστών



Broadcast Yourself™

Small World Phenomenon

- Κάθε χρήστης απέχει ≤ 6 βήματα από κάθε άλλον χρήστη μέσα στο γράφο

Indegree Distribution

- Πολλοί χρήστες με λίγες επαφές και λίγοι με πολλές. Λίγοι χρήστες είναι δημοφιλείς (οι άλλοι ασχολούνται μαζί τους)

Outdegree Distribution

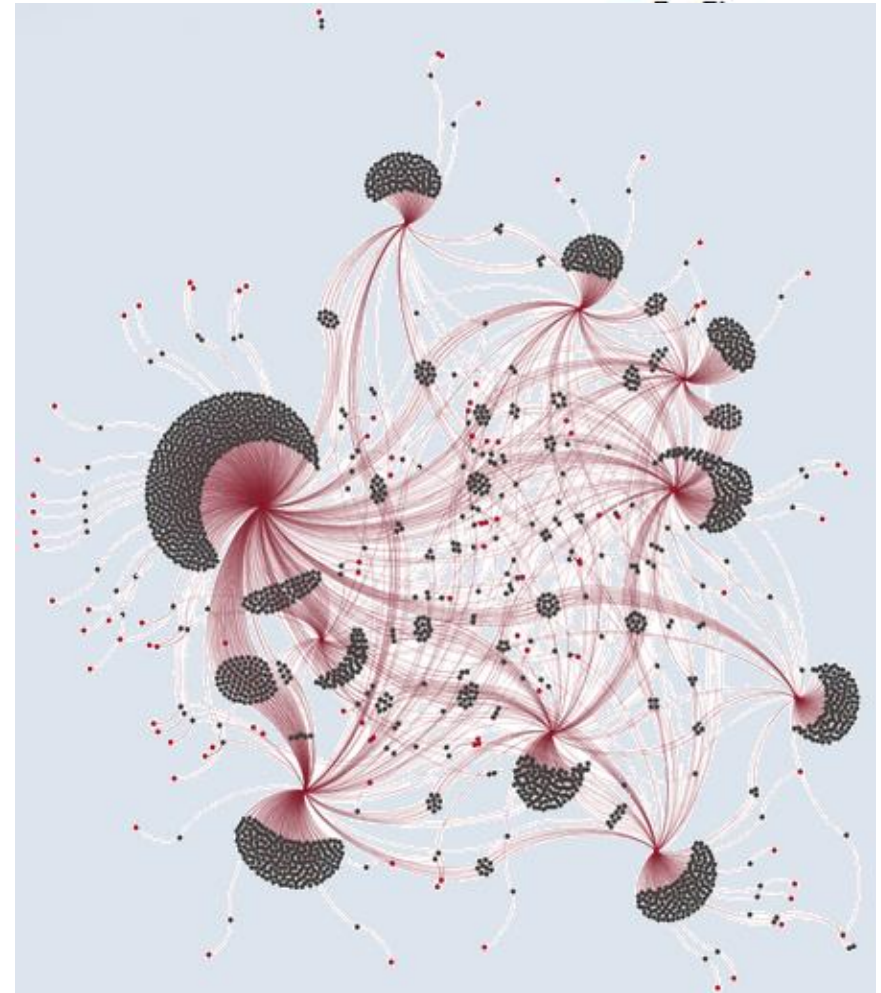
- Όλοι παράγουν περιεχόμενο. Λίγοι παράγουν το μεγάλο ποσοστό του περιεχομένου (κανόνας Pareto: Το 80% του υλικού παράγεται από το 20% των χρηστών)

Tag cloud

- Εντοπίζεται πολιτική φρασεολογία

Γράφημα χρηστών

- Οπτικοποίηση γράφου χρηστών



Αποκαλύπτοντας πολιτικές πεποιθήσεις μέσω ...Data Mining και ...YouTube!



Broadcast Yourself™
Profile

Κατηγορίες Χρηστών (ενδεικτικές):

«Κεντρώοι-Κεντροαριστεροί», «Ουδέτεροι», «Κεντρώοι-Κεντροδεξιοί»

Συνδρομή ειδικών (εντοπισμός μορφοτύπων):

Κοινωνιολόγος και Πολιτικός Επιστήμονας

Αλγόριθμος: Multinomial Logistic Regression (MLR)

ΚΑΤΗΓΟΡΙΕΣ ΜΕΤΡΙΚΕΣ	Κεντρώοι - Κεντροαριστεροί	Ουδέτεροι	Κεντρώοι - Κεντροδεξιοί
Precision	83%	91%	77%
Recall	77%	93%	78%
F-Score	80%	92%	77%
Accuracy	87%		

Precision: Ο αριθμός χρηστών που έχουν κατηγοριοποιηθεί ορθά, διαιρούμενος με το πλήθος των χρηστών που εντάχθηκαν σε αυτή την κατηγορία.

Recall: Ο αριθμός χρηστών που έχουν κατηγοριοποιηθεί ορθά, διαιρούμενος με το συνολικό αριθμό των χρηστών της κατηγορίας.

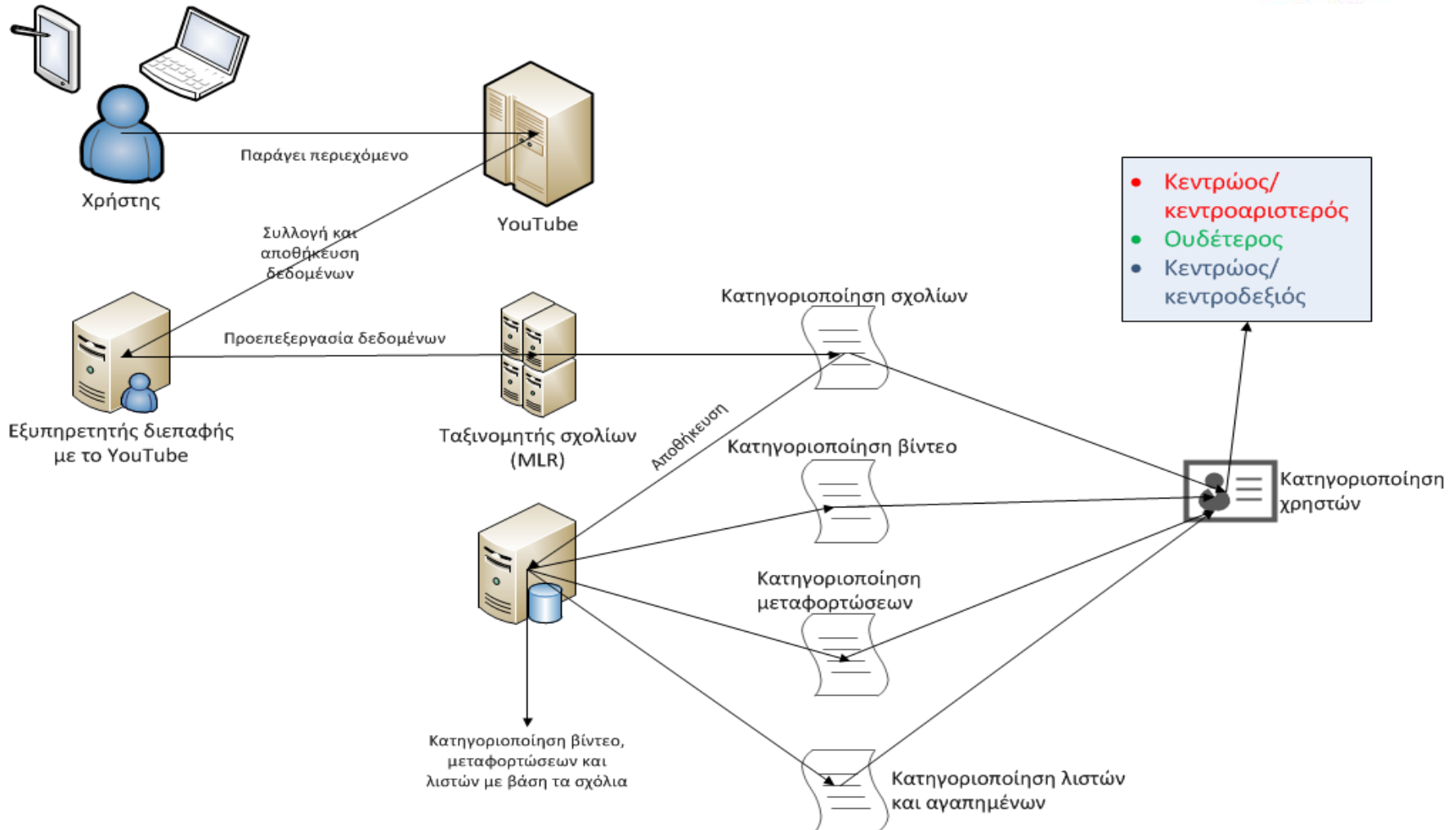
F-Score: Σταθμισμένος αρμονικός μέσος των **Precision** και **Recall**.

Accuracy: Ο αριθμός των ορθών κατηγοριοποιήσεων (ισούται με το πηλίκο των ορθών κατηγοριοποιήσεων δια του συνόλου των δεδομένων).

Διαδικασία εξαγωγής συμπερασμάτων (horror story!!!)



Broadcast Yourself™



Παρατηρήσεις (1/2)



Broadcast Yourself™
Profile

2% των **σχολίων** αφορούν πολιτικά ζητήματα (0.7% Ριζοσπαστικά, 1.3% Συντηρητικά)

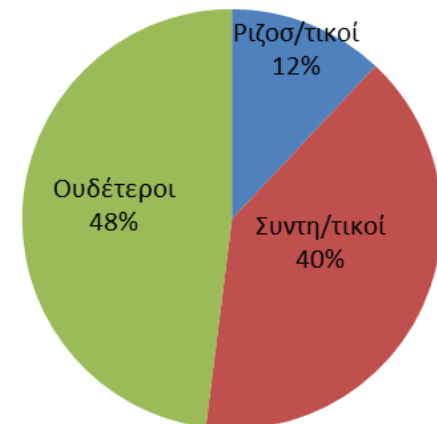
- Δηλαδή, σχεδόν **41.000 σχόλια** (από τα 2.000.000) έχουν ρητό πολιτικό περιεχόμενο

7% των **βίντεο** κατηγοριοποιήθηκαν ως πολιτικά (2% Ριζοσπαστικά, 5% Συντηρητικά)

- Δηλαδή, σχεδόν **14.000 βίντεο** (από τα 200.000) έχουν ρητό πολιτικό περιεχόμενο

12% των **χρηστών** εκφράζουν **Ριζοσπαστική** πολιτική τοποθέτηση και 40% **Συντηρητική**

- Αυτό συνεπάγεται ότι **6.760 χρήστες** «αποκαλύπτουν» την πολιτική τους τοποθέτηση



Παρατηρήσεις (2/2)



Broadcast Yourself™
Profile

Ριζοσπαστικοί:

- 20% των σχολίων τους περιλαμβάνει πολιτική τοποθέτηση
- Προτιμούν το ελληνικό αλφάβητο (54% σχολίων ελληνικά, 33% greeklish, 13% και τα δύο).
- Σχολιάζουν μαζικά κάποια συγκεκριμένα βίντεο
- Προτιμούν βίντεο με πολιτικό περιεχόμενο (γεγονότα, μουσική κλπ.)
- Προσθέτουν στα αγαπημένα βίντεο ντοκιμαντέρ και μουσικά βίντεο

Συντηρητικοί:

- Προτιμούν να γράφουν σε greeklish (55% greeklish, 35% ελληνικά, 10% και τα δύο).
- Προτιμούν «συνωμοσιολογία» βίντεο και βίντεο με εθνικιστικό περιεχόμενο.

Τα σχόλια σε Greeklish είναι **πιο σύντομα** και **πιο επιθετικά**

Τα σχόλια με ελληνικό αλφάβητο είναι **πιο εκτενή** και **επεξηγηματικά**

Όσο **πιο επιθετικό** είναι ένα σχόλιο, τόσο **πιο ανορθόγραφο**

7% των βίντεο δημοσιεύονται **με Creative Commons** άδεια χρήσης

- 55% από Ριζοσπαστικούς, 10% από Συντηρητικούς, 35% από Ουδέτερους

Ιστορίες από το παρελθόν ή διηγήσεις για το παρόν και το μέλλον;



Broadcast Yourself™
Profile

- Αυτόματη κατηγοριοποίηση χρηστών, ανάλογα με τις πολιτικές τους πεποιθήσεις!
- Στατιστικώς αξιόπιστα αποτελέσματα, χωρίς να αποκλείονται false positives/negatives
- Αποκάλυψη ευαίσθητων προσωπικών δεδομένων, για μεγάλο πλήθος χρηστών, εύκολα και γρήγορα
- Επιβεβαιώνονται δημογραφικά και δημοσκοπικά δεδομένα της χρονικής συγκυρίας διενέργειας της έρευνας

Η αιτία είναι ο εχθρός, όχι το αποτέλεσμα



Broadcast Yourself™
Profile



References

1. Amichai-Hamburger Y., Vinitzky G., "Social network use and personality", *Computers in Human Behavior*, Vol. 26, pp. 1289-1295, 2010.
2. Gritzalis D., Stavrou V., Kandias M., Stergiopoulos G., "Insider Threat: Enhancing BPM through Social Media", in *Proc. of the 6th IFIP International Conference on New Technologies, Mobility and Security*, Springer, UAE, 2014.
3. Gritzalis D., Kandias M., Stavrou V., Mitrou L., "History of Information: The case of Privacy and Security in Social Media", in *Proc. of the History of Information Conference*, Law Library Publications, Athens, 2014.
4. Kandias M., Mitrou L., Stavrou V., Gritzalis D., "Which side are you on? A new Panopticon vs. privacy", in *Proc. of the 10th International Conference on Security and Cryptography*, pp. 98-110, Iceland, 2013.
5. Kandias M., Galbogini K., Mitrou L., Gritzalis D., "Insiders trapped in the mirror reveal themselves in social media", in *Proc. of the 7th International Conference on Network and System Security*, pp. 220-235, Springer (LNCS 7873), Spain, June 2013.
6. Kandias M., Stavrou V., Bozovic N., Mitrou L., Gritzalis D., "Can we trust this user? Predicting insider's attitude via YouTube usage profiling", in *Proc. of 10th IEEE International Conference on Autonomic and Trusted Computing*, pp. 347-354, IEEE Press, Italy, 2013.
7. Kandias M., Stavrou V., Bosovic N., Mitrou L., Gritzalis D., "Proactive insider threat detection through social media: The YouTube case", in *Proc. of the 12th Workshop on Privacy in the Electronic Society*, pp. 261-266, ACM Press, Germany, 2013.
8. Kosinski M., Stillwell D., Graepel T., "Private traits and attributes are predictable from digital records of human behavior", *Proc. of the National Academy of Sciences*, 2013.
9. Mitrou L., Kandias M., Stavrou V., Gritzalis D., "Social media profiling: A Panopticon or Omnipticon tool?", in *Proc. of the 6th Conference of the Surveillance Studies Network*, Spain, 2014.
10. Mylonas A., Meletiadiis V., Mitrou L., Gritzalis D., "Smartphone sensor data as digital evidence", *Computers & Security*, Vol. 38, pp. 51-75, October 2013.
11. Ross C., Orr E., Sisic M., Arseneault J., Simmering M., Orr R., "Personality and motivations associated with Facebook use", *Computers in Human Behavior*, Vol. 25, pp. 578-586, 2009.
12. Sebastiani, F., Machine learning in automated text categorization, *ACM Computing Surveys*, vol. 34, no. 1, pp. 1-47, 2002.
13. Stavrou V., Kandias M., Karoulas G., Gritzalis D., "Business Process Modeling for Insider threat monitoring and handling", in *Proc. of the 11th International Conference on Trust, Privacy & Security in Digital Business*, Springer, Germany, 2014.