

Εισαγωγή στην Επιστήμη του Ιστού

Ενότητα 7 : Data analytics / Crowdsourcing

Διδάσκων: Νίκος Τσέλιος

Τμήμα Επιστημών της Εκπαίδευσης και της Αγωγής στην
Προσχολική Ηλικία



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης

ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Σημείωμα Αναφοράς

Copyright Πανεπιστήμιο Πατρών, Σχολή Κοινωνικών και Ανθρωπιστικών Επιστημών, Τμήμα Επιστημών της Εκπαίδευσης και Αγωγής στην Προσχολική Ηλικία, Νικόλαος Τσέλιος, «Εισαγωγή στην Επιστήμη του Ιστού». Έκδοση: 1.0.

Πάτρα 2014. Διαθέσιμο από τη δικτυακή διεύθυνση:

<https://eclass.upatras.gr/courses/PN1427/>

Σημείωμα Αδειοδότησης

Το παρόν υλικό διατίθεται με τους όρους της άδειας χρήσης Creative Commons Αναφορά, Μη Εμπορική Χρήση Παρόμοια Διανομή 4.0 [1] ή μεταγενέστερη, Διεθνής Έκδοση. Εξαιρούνται τα αυτοτελή έργα τρίτων π.χ. φωτογραφίες, διαγράμματα κ.λ.π., τα οποία εμπεριέχονται σε αυτό και τα οποία αναφέρονται μαζί με τους όρους χρήσης τους στο «Σημείωμα Χρήσης Έργων Τρίτων».



[1] <http://creativecommons.org/licenses/by-nc-sa/4.0/>

Ως **Μη Εμπορική** ορίζεται η χρήση:

- που δεν περιλαμβάνει άμεσο ή έμμεσο οικονομικό όφελος από την χρήση του έργου, για το διανομέα του έργου και αδειοδόχο
- που δεν περιλαμβάνει οικονομική συναλλαγή ως προϋπόθεση για τη χρήση ή πρόσβαση στο έργο
- που δεν προσπορίζει στο διανομέα του έργου και αδειοδόχο έμμεσο οικονομικό όφελος (π.χ. διαφημίσεις) από την προβολή του έργου σε διαδικτυακό τόπο

Ο δικαιούχος μπορεί να παρέχει στον αδειοδόχο ξεχωριστή άδεια να χρησιμοποιεί το έργο για εμπορική χρήση, εφόσον αυτό του ζητηθεί.

Χρηματοδότηση

- ☉ Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα εκτός κι αν αναφέρεται διαφορετικά
- ☉ Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Πατρών**» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού
- ☉ Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους



Σκοποί ενότητας

- ⦿ Να εξοικειωθεί με την έννοια του Crowdsourcing, δηλαδή την ανάθεση έργου σε χρήστες του Διαδικτύου
- ⦿ Να εξοικειωθεί με αντιπροσωπευτικές υπηρεσίες του είδους όπως το Mechanical Turk
- ⦿ Να αναγνωρίσει διαφορετικά είδη crowdsourcing

Περιεχόμενα ενότητας

- ⦿ Περιγραφή: Γίνεται αναφορά στο Crowdsourcing. Δίνεται ο ορισμός και γίνεται επεξήγηση του όρου
- ⦿ Λέξεις Κλειδιά: Crowdsourcing

Crowdsourcing (1/2)

- ⦿ Διαδικασία ανάθεσης ενεργειών σε ένα κατανεμημένο σύνολο
- ⦿ Παραλλαγή της λέξης outsourcing (από τον Jeff Howe "The Rise of Crowdsourcing"
<http://www.wired.com/wired/archive/14.06/crowds.html>)
- ⦿ Σχετίζεται με διαδικασίες **ανθρώπινου** υπολογισμού, δηλαδή με επίλυση προβλημάτων από ανθρώπους

Crowdsourcing (2/2)

- ◎ Η αρχική ιδέα: Ανάθεση προβλήματος προς επίλυση
 - Ανακοίνωση του προβλήματος
 - Εμπλοκή συμμετεχόντων (αρχάριων χρηστών ή έμπειρων που εργάζονται σε άλλες εταιρίες)
 - Επιλογή λύσης (με 'ηθική' ή χρονική επιβράβευση)

Ορισμός

- Estellés and González (2012) μετά από μελέτη 40 μορφών crowdsourcing
- «Μια συμμετοχική διαδικτυακή δραστηριότητα όπου κάποιος ιδιώτης ή οργανισμός, κερδοσκοπικός ή μη προτείνει σε πλήθος ιδιωτών με ποικιλία γνώσεων, χαρακτηριστικών να συμμετάσχουν με τις γνώσεις, την εργασία και την εμπειρία τους σε μια διαδικασία»
- Η διαδικασία αυτή σχεδόν πάντα χαρακτηρίζεται από αμοιβαίο όφελος
- Estellés A.E., & González L. F. (2012). Towards an integrated crowdsourcing definition. *Journal of Information Science*, 38(2), 189-200.

Αρχική ιδέα

- ⦿ Sir Francis Galton «Σοφία του Πλήθους» Εργασία στο Nature το 1907
- ⦿ Εμπνευσμένη από μια επίσκεψή του το 1906 σε έκθεση πουλερικών και ζώων κρεατοπαραγωγής
 - Κλήθηκαν 800 συμμετέχοντες να μαντέψουν πόσα κιλά κρέας θα προέκυπταν από τη σφαγή ενός βοδιού
 - Κανείς από τους 800 δεν μάντεψε σωστά, ωστόσο ο μ.ο. της εκτίμησης ήταν εκπληκτικά κοντά
- ⦿ 2011: BBC 'The code'. Εμφάνισαν ένα βάζο με 4510 καραμέλες και ρώτησαν το πλήθος πόσες καραμέλες έχει
 - 160 ερωτηθέντες, απαντήσεις από 400 ως 50000. Μόνο 4 έπεσαν πολύ κοντά.
 - Ο Μ.Ο. ήταν 4514,89!

Διαδίκτυο και crowdsourcing

- ◎ Σήμερα το CS διεξάγεται αποκλειστικά σχεδόν με τη χρήση του Διαδικτύου
 - Εύκολη ανακοίνωση, διανομή και χειρισμός της διαδικασίας
 - Οι χρήστες συγκεντρώνονται στην εργασία χωρίς παρενόχληση και αρνητική κριτική
- ◎ Βασικές κατηγορίες σήμερα
 - Πληθοπορισμός για web & graphics design
 - Crowdfunding
 - Microtasking

Αίτια συμμετοχής στη διαδικασία

⦿ 'Εργοδότης'

- Ποικιλία λύσεων σε σχετικά μικρό χρόνο
- Χαμηλό κόστος

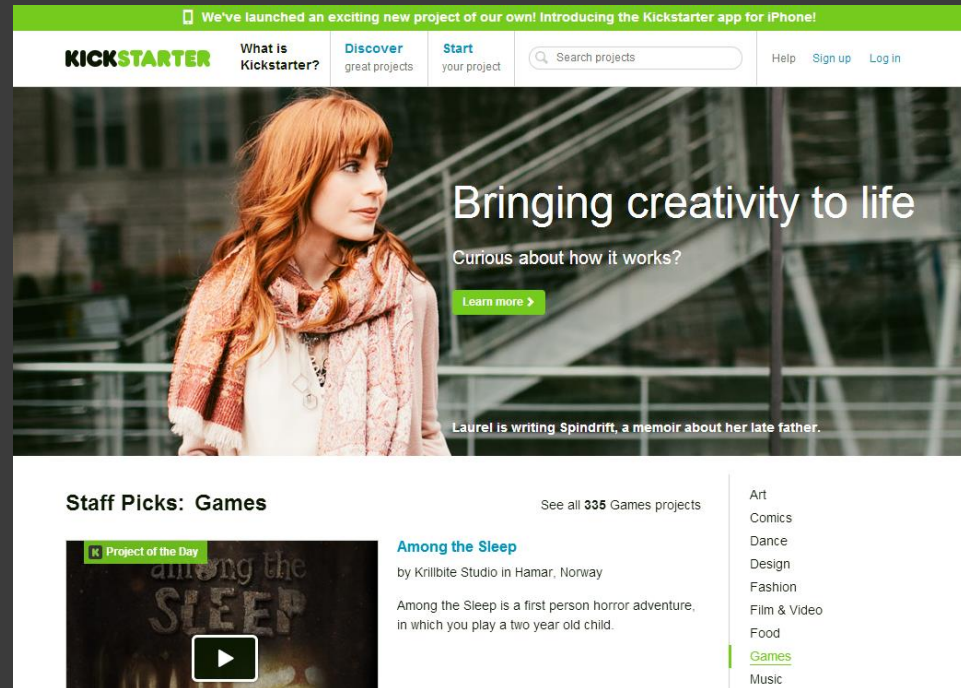
⦿ 'Εργαζόμενος'

- Εσωτερικά κίνητρα
 - Όφελος από τη συμμετοχή στο κοινωνικό δίκτυο
 - Ικανοποίηση
 - Δημιουργία
 - Αίσθηση επάρκειας
- Εξωτερικά κίνητρα
 - Συνήθως χρηματικά

Παραδείγματα

- **Oxford English Dictionary:** Ανοιχτή πρόσκληση για κατάθεση αγγλικών λέξεων και καταλογογράφησή τους σε μια περίοδο 70 ετών (στα μέσα του 19^{ου} αιώνα)
- Δημιουργία γενεαλογικών δένδρων
- Γαλλική κυβέρνηση: Διαγωνισμός για την επίλυση προβλημάτων έναντι χρηματικού επάθλου
- Linux! Δημιουργία λειτουργικού συστήματος από κοινότητα χρηστών
- Wikipedia
- Μελέτη φωτογραφιών από δορυφόρο για εντοπισμό ναυαγών

Crowdfunding: Kickstarter



- ◎ Χρηματοδότηση καινοτόμων ιδεών
 - Στην τέχνη, στις επιστήμες...
 - Πιο απλή από τραπεζικό δανεισμό
 - Τα δικαιώματα παραμένουν στο δημιουργό
 - 4 εκατ. Άνθρωποι έχουν επενδύσει 611 εκ. δολάρια από το 2009 σε 41.000 projects
 - Taxibeat, Uber

Mechanical Turk

(<https://www.mturk.com/mturk/welcome>)



Artificial Intelligence

[Your Account](#) [HITs](#) [Qualifications](#)

[Introduction](#) | [Dashboard](#) | [Status](#) | [Account Settings](#)

Mechanical Turk is a marketplace for work.
We give businesses and developers access to an on-demand, scalable workforce.
Workers select from thousands of tasks and work whenever it's convenient.
49,563 HITs available. [View them now.](#)

Make Money

by working on HITs

HITs - Human Intelligence Tasks - are individual tasks that you work on. [Find HITs now.](#)

As a Mechanical Turk Worker you:

- Can work from home
- Choose your own work hours
- Get paid for doing good work

Find an interesting task



Work



Earn money



[Find HITs Now](#)

or [learn more about being a Worker](#)

Already have an account?
[Sign in as a Worker](#) | [Requester](#)

Get Results

from Mechanical Turk Workers

Ask workers to complete HITs - Human Intelligence Tasks - and get results using Mechanical Turk. [Register Now](#)

As a Mechanical Turk Requester you:

- Have access to a global, on-demand, 24 x 7 workforce
- Get thousands of HITs completed in minutes
- Pay only when you're satisfied with the results

Fund your account



Load your tasks



Get results



[Get Started](#)

or [learn more about being a Requester](#)

Mechanical Turk

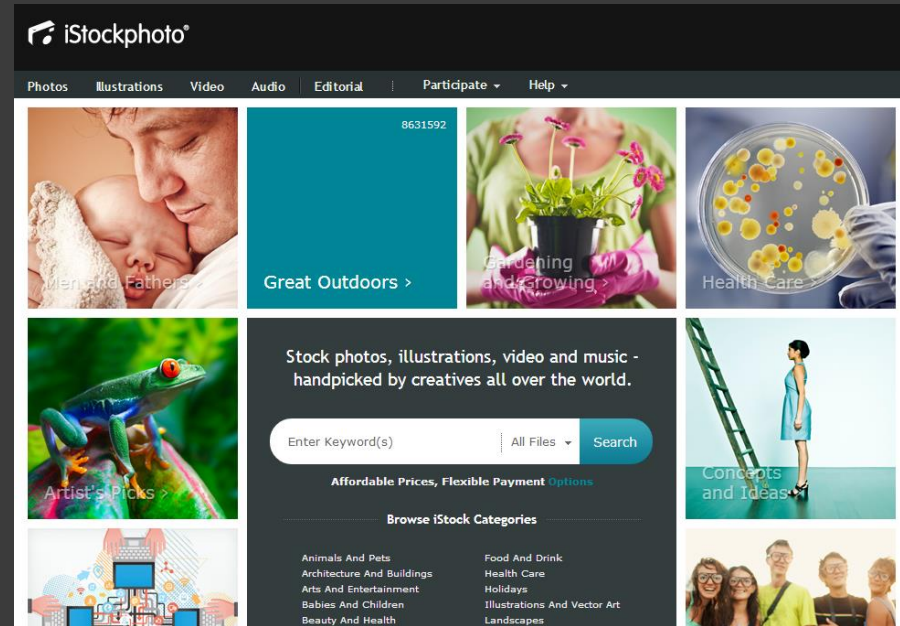
- ◉ Το Mechanical Turk αποτελεί μια ηλεκτρονική αγορά εργασίας για μικροεργασίες
- ◉ Οι αξιολογητές εγγράφονται στην υπηρεσία και επιλέγουν να ολοκληρώσουν εργασίες (πχ tagging φωτογραφιών, λέξεων κλειδιά σε κείμενο, συμπλήρωση ερωτηματολογίων) έναντι μικρής αμοιβής
- ◉ Ένας ερευνητής για να κάνει πχ αξιολόγηση ευχρηστίας μπορεί
 - να περιγράψει κάποιες μικρές εργασίες που θα πρέπει να ολοκληρωθούν στο δικτυακό τόπο και ακολούθως λαμβάνουν ανάδραση από τους αξιολογητές σχετικά με το πόσο εύκολη ή δύσκολη ήταν η ολοκλήρωσή του
 - Δεν υπάρχει η δυνατότητα συμπληρωματικών ερωτήσεων
- ◉ Άλλες υπηρεσίες χρησιμοποιούν έμμεσα το Mechanical Turk διεκπεραιώνοντας τη διαδικασία. Απευθείας χρήση του είναι συνεπώς φθηνότερη αλλά απαιτεί μεγαλύτερη προσπάθεια από την πλευρά του ερευνητή
- ◉ **Pros** – Ευρύ και χαμηλού κόστους δίκτυο για τη συλλογή ανάδρασης
- ◉ **Cons** – Απαιτεί κάποια προσπάθεια για την οργάνωση της εργασίας. Επίσης απαιτεί λογαριασμό στο Amazon.com για την πληρωμή. Η ποιότητα της ανάδρασης δεν είναι εγγυημένη
- ◉ **Pricing** – Πολύ χαμηλό κόστος, που αυξάνεται με τον αριθμό των ζητούμενων εργασιών και το συνολικό χρόνο που θα απασχοληθεί ο χρήστης

Επισημείωση προβλημάτων προσβασιμότητας



- Hara, K., Le, V., & Froehlich, J. (2013). Combining Crowdsourcing and Google Street View to Identify Street-level Accessibility Problems. Proceedings of CHI 2013, Paris, France, Apr 24-May 2, 2013.
- Προσβάσιμο από:
http://www.cs.umd.edu/~jonf/publications/Hara_CombiningCrowdsourcingAndGoogleStreetViewToIdentifyStreet-levelAccessibilityProblems_CHI2013.pdf

istockphoto



- Επαγγελματίες (ή μη) φωτογράφοι μεταφορτώνουν τις φωτογραφίες τους
- Οι χρήστες αγοράζουν τα δικαιώματα χρήσης έναντι χαμηλού τιμήματος (από 1 USD).

Είδη κινήτρων

- ⦿ Διαφέρουν ανάλογα με το χρήστη: πχ
 - οι γυναίκες εκτιμούν περισσότερο την κοινωνική επαφή
 - Εργαζόμενοι μερικής απασχόλησης έχουν κύριο κίνητρο την αύξηση των δεξιοτήτων τους (όχι οι πλήρους απασχόλησης ή οι φοιτητές)
- ⦿ Εξωτερικά:
 - Άμεσα: Πληρωμή
 - Έμμεσα: Αύξηση δεξιοτήτων
 - Prestige, social status
- ⦿ Εσωτερικά: Ικανοποίηση, συμμετοχή σε κοινότητα

Δημογραφικά

- ⦿ Ross et al. (2009). Μελέτη 400.000 χρηστών **Mechanical Turk**:
 - 36% Ινδοί (εκ των οποίων 67% άνδρες, 66% είχαν BSc, 67% <10.000 εισόδημα, για το 27% ήταν κύρια πηγή εσόδων)
- ⦿ US user: 2.3\$ per task, 1.58\$ indian worker
- ⦿ Η πλειοψηφία δούλευε <5 ώρες την εβδομάδα. 18% δούλευε >15ώρες/εβδομάδα
- ⦿ Συνήθως μικρότερο από τον ελάχιστο μισθό
- ⦿ **Microworkers.com**, χρήστες από 197 χώρες, 25% Ινδία, Ινδονησία 18%, Μπαγκλαντές 17%
- ⦿ 28% εργοδοτών US
- ⦿ Σε γενικές γραμμές οι χρήστες έχουν καλό επίπεδο εκπαίδευσης

Είδη

- ⦿ Explicit: ανάθεση συγκεκριμένης εργασίας (πχ αξιολόγηση, υλοποίηση κλπ)
- ⦿ Implicit:
 - Standalone: Επίλυση προβλήματος ως παράπλευρο αποτέλεσμα εμπλοκής σε μια διαδικασία
 - Πχ Αναγνωρίστε μια φωτογραφία! Τελικά χρησιμοποιείται ως διαδικασία tagging
 - piggyback : Πληροφορίες χρηστών από 3^ο δικτυακό τόπο (πχ Google Adsense)

Doan, A., Ramakrishnan, R., & Halevy, A. (2011). ["Crowdsourcing Systems on the World Wide Web"](#), *Communications of the ACM* 54(4): 86–96

Ερευνητικά

- ⦿ Για δημιουργία βάσεων δεδομένων ομιλίας ή κειμένου
- ⦿ Για συμμετοχή σε πειράματα (πρόσβαση σε χρήστες με 'ευρύτερα' δημογραφικά χαρακτηριστικά)
- ⦿ Φαίνεται ότι το framing της εργασίας (δηλαδή η παρουσίαση ότι αφορά ερευνητικό σκοπό) αυξάνει τη συμμετοχή των χρηστών
 - Προς το παρόν όμως δεν διαφαίνεται επίδραση στην ποιότητα του αποτελέσματος

Kittur, A., Chi, E.H., & Sun, B (2008). ["Crowdsourcing user studies with Mechanical Turk"](#), CHI 2008.

Komarov, S., Reinecke, K., & Gajos, K. Z. (2013). Crowdsourcing Performance Evaluations of User Interfaces.

Ερωτήματα

- ◎ Πως ωθούμε τους χρήστες να συνεισφέρουν;
 - Υποχρεωτικά. Διευθυντής ζητά από κάθε υπάλληλο να κάνει κάτι
 - Με πληρωμή
 - Εθελοντικά
 - Του παρέχουμε μια υπηρεσία Α αφού κάνει κάτι Β
 - Έμμεσα (από τις αναζητήσεις πχ συμπεραίνουμε για τη σωστή ορθογραφία)

Κριτική

⦿ Ποιότητα δουλειάς

- Συχνά υπάρχει έμφαση στην ταχύτητα όχι στην ποιότητα
- Ανάγκη για συμμετοχή αρκετών χρηστών ώστε να 'διορθώνονται' πιθανά λάθη
- Η έλλειψη αλληλεπίδρασης οδηγεί συχνά σε παρανοήσεις

⦿ Χαμηλές αμοιβές/ Έλλειψη συμβολαίων

- Εγείρει ερωτήματα ηθικής φύσεως

Concept feedback (1/2)

Concept feedback

Concepts Tags Members How it works

Free Feedback for Marketers and Designers.

Getting ready to launch a brand new concept? Would you like to know what other professionals think before you release it to the world? Concept Feedback is designed specifically for you!

- Upload and share your website, advertisement, logo and more
- Receive quick, actionable feedback from a professional community
- Improve your concept and get better results!

Redesign website, PowerStorm

Review Concept

Category	Review	Rating	Like
Design	★★★★★	5.00	1
Concept	★★★★★	5.00	1
Originality	★★★★★	5.00	1
Implementation	★★★★★	5.00	1

Submit Review

Concept feedback (2/2)

- **Concept Feedback** σχεδιάστηκε για τη συλλογή ανάδρασης από ειδικούς για νέα σχέδια κυρίως στο marketing και στη διαφήμιση
- Μπορεί όμως να χρησιμοποιηθεί για τον ίδιο σκοπό για σχέδια δικτυακών τόπων ή διεπιφανειών χρήσης
- Πως δουλεύει: Μεταφορτώνετε το σχέδιό σας στην κοινότητα και οι αξιολογητές σχολιάζουν το σχέδιό σας και παρέχουν σχόλια για το σχεδιασμό
- Κάθε 'ειδικός' βαθμολογείται διαχρονικά για την ποιότητα των σχολίων του συνεπώς η γνώμη των ειδικών με τη μεγαλύτερη βαθμολογία πιθανά έχει τη μεγαλύτερη βαρύτητα και αξία
- **Pros** Δωρεάν και ταχύτατη συμβουλευτική. Επιπρόσθετα, μπορείτε να παρακολουθείτε την κοινότητα και να δικτυωθείτε με αυτή
- **Cons** – Ο σχολιασμός γίνεται στη βάση μιας στατικής εικόνας, συνεπώς δεν μπορούν να αξιολογηθούν χαρακτηριστικά αλληλεπίδρασης αλλά μόνο φυσικού σχεδιασμού
- **Pricing** – Δωρεάν

ACM Interactions, άρθρο για crowdsourcing (1/2)

Ένα δείγμα από βασικούς τύπους συστημάτων CS στον Παγκόσμιο Ιστό:

ΦΥΣΗ ΣΥΝΕΡΓΑΣΙΑΣ: ΑΡΡΗΤΗ (έμμεση)

Αρχιτεκτονική	Απαραίτητη πρόσληψη χρηστών?	Τι κάνουν οι χρήστες	Παραδείγματα	Στοχοθετημένα προβλήματα	Σχόλια
Αυτόνομη	Ναι	-Παίζουν παιχνίδια με σκοπό -ποντάρουν στις αγορές πρόβλεψης -χρησιμοποιούν ιδιωτικούς λογαριασμούς -επιλύουν captcha -αγοράζουν/ πουλούν/ οργανώνουν δημοπρασίες, παίζουν μαζικά παιχνίδια με πολλούς παίκτες	-ESP -intrade.com, Iowa Electronic Markets -IMDB privet accounts -recaptcha.net -eBay, <i>World of Warcraft</i>	-ονομασία εικόνων -πρόβλεψη γεγονότων -αξιολόγηση ταινιών -ψηφιοποίηση γραπτών κειμένων -δημιουργία μιας κοινότητας χρηστών (για σκοπούς όπως διαφήμιση ή αμοιβές)	Οι άνθρωποι σε όλους τους ρόλους. Οι συνδυασμοί των εισροών μπορεί να είναι είτε χαλαροί είτε σφιχτοί.
Στηρίζονται σε άλλα συστήματα	Όχι	-αναζητούν με λέξεις- κλειδιά -αγοράζουν προϊόντα -περιηγούνται σε ιστότοπους	-Google, Microsoft, Yahoo -Recommendation feature of Amazon, -Adaptive Web sites (e.g., κεντρική σελίδα Yahoo!)	- Ορθογραφικός έλεγχος, πρόβλεψη επιδημίας -σύσταση προϊόντων -αναγνώριση ενός δικτυακού τόπου για καλύτερη πρόσβαση	Οι άνθρωποι σε όλους τους ρόλους. Οι συνδυασμοί των εισροών μπορεί να είναι είτε χαλαροί είτε σφιχτοί.

ACM Interactions, άρθρο για crowdsourcing (2/2)

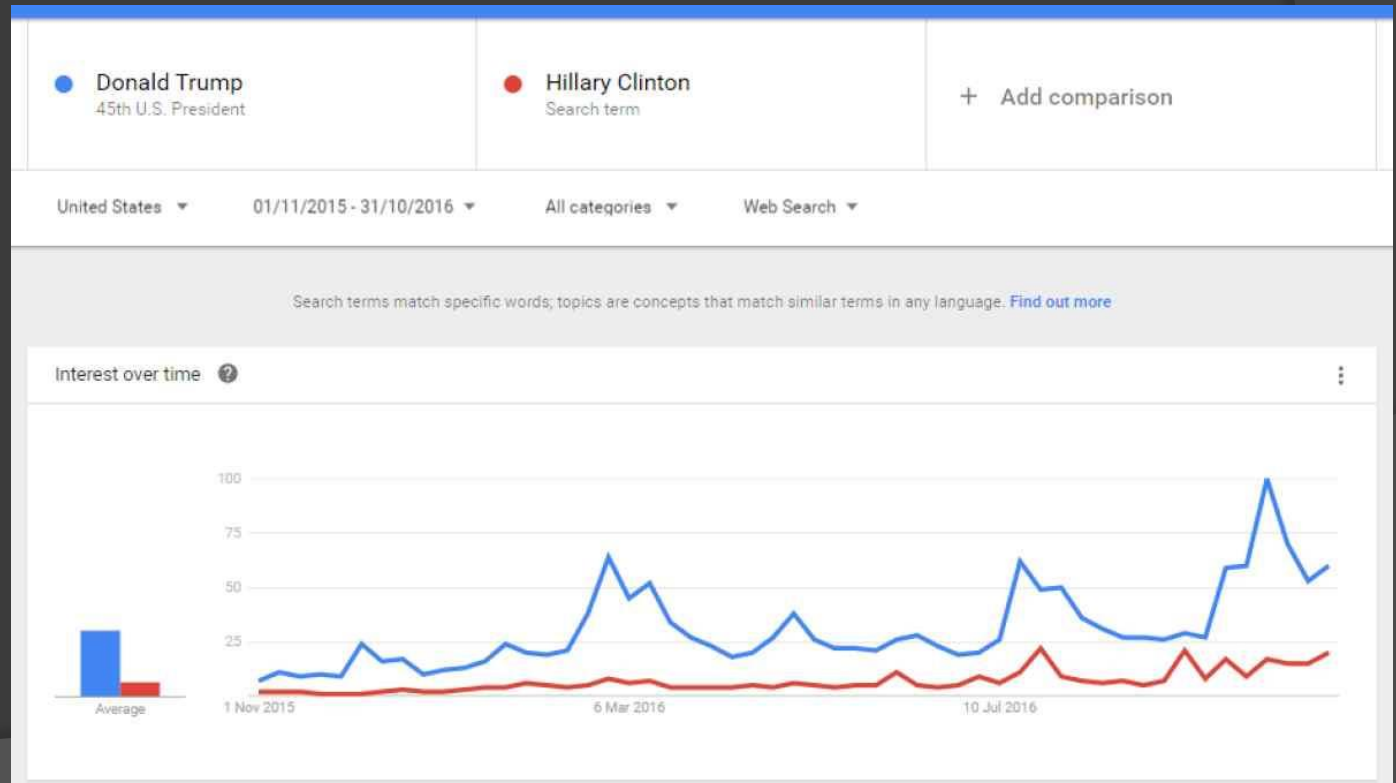
Ένα δείγμα από βασικούς τύπους συστημάτων CS στον Παγκόσμιο Ιστό:

ΦΥΣΗ ΣΥΝΕΡΓΑΣΙΑΣ: ΡΗΤΗ, ΑΡΧΙΤΕΚΤΟΝΙΚΗ:ΑΥΤΟΝΟΜΗ, ΑΠΑΡΑΙΤΗΤΗ ΠΡΟΣΛΗΨΗ ΧΡΗΣΤΩΝ: ΝΑΙ

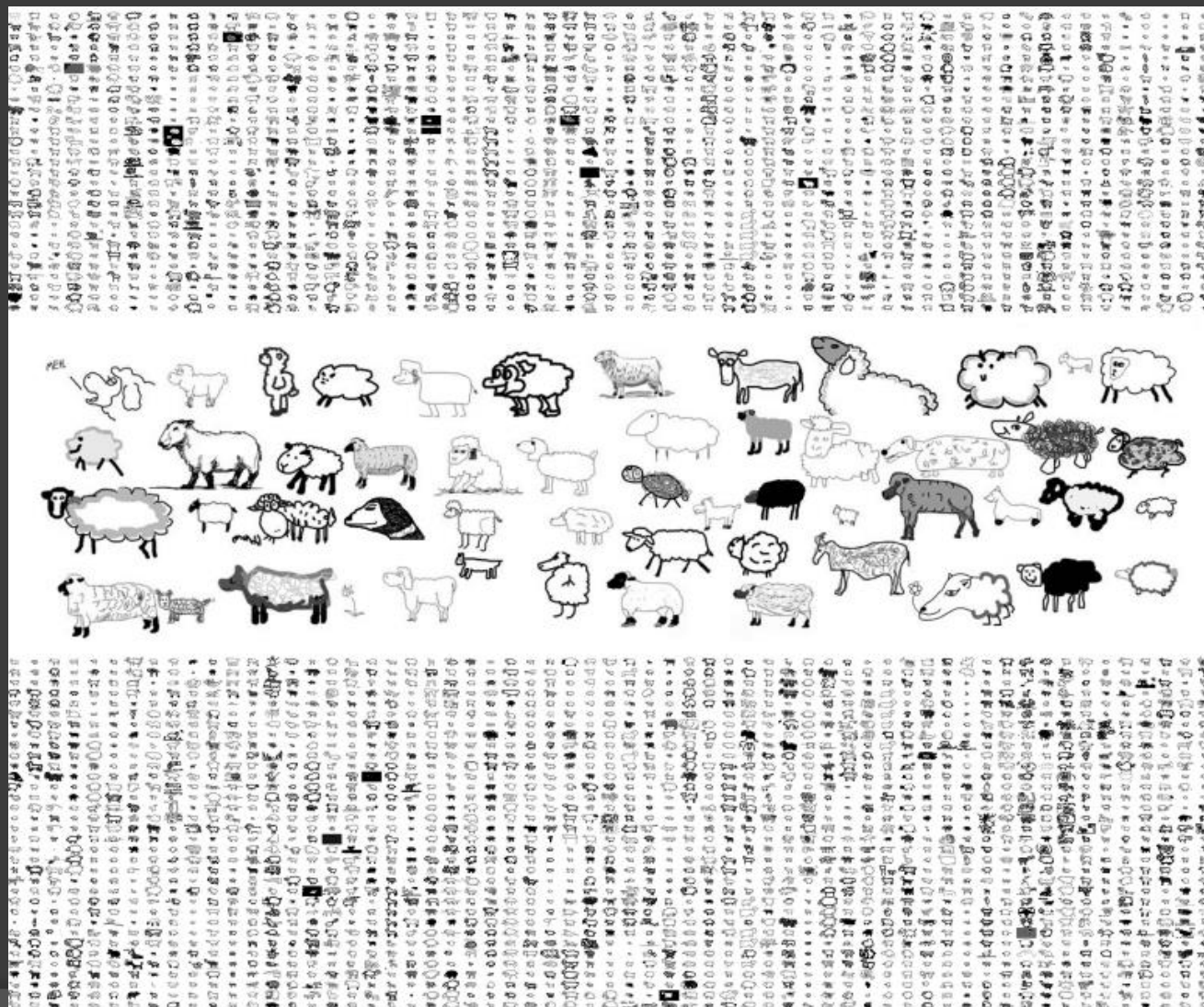
Τι κάνουν οι χρήστες	Παραδείγματα	Στοχοθετημένα προβλήματα	Σχόλια
Αξιολογούν -κριτικάρουν, ψηφίζουν, σημειώνουν	Επανεξέταση και ψήφισμα στο Amazon, σημείωση δικτυακών τόπων στο del.isi.ous.com και Google Co-op	Αξιολόγηση μιας συλλογής στοιχείων (π.χ. προϊόντα, χρήστες)	Οι άνθρωποι ως πάροχοι απόψεων (perspective). Χαλαρός ή καθόλου συνδυασμός εισροών.
Μοιράζονται -αντικείμενα/ στοιχεία -γραφτή γνώση -δομημένη γνώση	-Napster, YouTube, Flickr, CPAN, programmableweb.com, -λίστες ηλεκτρονικού ταχυδρομείου, Yahoo!, Answers, QUIQ, ehow.com, Quora -Swivel, Many eyes, Google Fusion, Tables, Google base, bmrw.wisc.edu, Galaxyzo, Piazza, Orchestra	Οικοδόμηση μιας (διανεμημένης ή κεντρικής) συλλογής στοιχείων που μπορεί να μοιραστεί μεταξύ των χρηστών.	Οι άνθρωποι ως πάροχοι περιεχομένου. Χαλαρός ή καθόλου συνδυασμός εισροών
Δικτυώνονται	-LinkedIn, MySpace, Facebook	Οικοδόμηση κοινωνικών δικτύων	Οι άνθρωποι ως πάροχοι στοιχείων. Χαλαρός συνδυασμός εισροών.
Οικοδομούν τεχνουργήματα -λογισμικά -βάσεις γραπτής γνώσης -βάσεις δομημένης γνώσης -συστήματα -άλλα	-Linux, Apache, Hadoop -Wikipedia, Openmind, Intellipedia, ecolcommunity -Wikipedia, Infoboxes/BDpedia, IWP, Google Fusion Tables, YAGO- NAGA, Cimple/ BDLife -Wikia Search, mahalo, Freebase, Eureka -εφημερίδες στο Digg.com, Second life	Οικοδόμηση φυσικών τεχνουργημάτων	Οι άνθρωποι σε όλους τους ρόλους. Τυπικός, σφιχτός συνδυασμός εισροών. Κάποια συστήματα ζητούν από ανθρώπους και μηχανές να συνεισφέρουν.
Εκτελούν εργασίες	-Finding extraterrestrials, εκλογές, εξεύρεση ανθρώπων, δημιουργία περιεχομένου (πχ. Μέσα ζήτησης	Ενδεχομένως οποιοδήποτε πρόβλημα	

Google trends

- ⦿ Αποτελεσματικότερο εργαλείο εκτίμησης εκλογικής επιρροής σε σχέση με τις παραδοσιακές δημοσκοπήσεις
- ⦿ Κυρίως σε
 - Ελληνικό δημοψήφισμα
 - Βρετανικό δημοψήφισμα
 - Εκλογές Προέδρου ΗΠΑ



Πληρωμή 0.02\$ για κάθε σχέδιο προβάτου



The Sheep Market by Aaron Koblin is a collection of 10,000 sheep made by workers on Amazon's Mechanical Turk. Workers were paid 0.02 (USD) to "draw a sheep facing to the left." Animations of each sheep's creation may be viewed at thesheepmarket.com. Credits to: Aaron Koblin

Πως δημιουργούμε κοινότητα

- ⦿ Μπορούμε να το απαιτήσουμε (πχ υποχρεωτικό στα πλαίσια ενός μαθήματος)
- ⦿ Πληρωμή
- ⦿ Εθελοντισμός
- ⦿ Έμμεση πληρωμή (πχ να απαντήσει κάποιος σε ένα captcha πριν κάνει ανάρτηση. Το captcha όμως θα αφορά σε αναγνώριση γραφής)
- ⦿ Piggyback (χρησιμοποιώ πχ search queries για έλεγχο ορθογραφίας)

Crowdsourcing

- ◎ Πως μπορούμε να αξιοποιήσουμε το crowdsourcing στην Εκπαίδευση;
- ◎ Σχετικό course στο Amherst
<http://people.cs.umass.edu/~dganesan/courses/fall10/syllabus.html>
- ◎ Λίστα με crowdsourcing projects
http://en.wikipedia.org/wiki/List_of_crowdsourcing_projects

Ο Παγκόσμιος Ιστός ως ένα κοινωνικό εργαστήριο

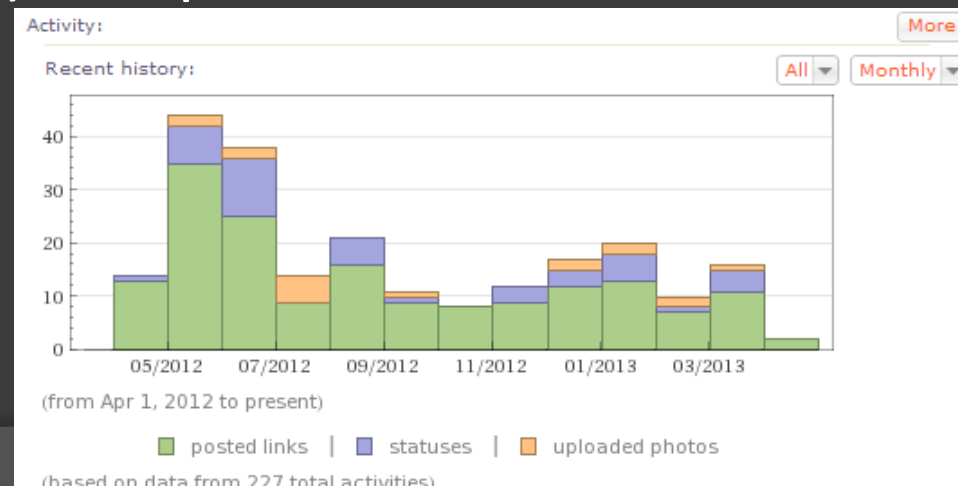
⦿ Εύκολη συλλογή δεδομένων

- Από κοινωνικά μέσα/δίκτυα (πληκτρολογήσεις, συσχετίσεις δικτύων)
- Από αποστολές δεδομένων μεταξύ μηχανών (αισθητήρες, ροή πληροφοριών)
- Από ψηφιακές καρτέλες με στοιχεία όπως πχ κλήσεις τηλεφώνου ή δημοσιεύσεις μελών ΔΕΠ κλπ
- Από δεδομένα χρηστών (chat, forum, ιστολόγια κλπ)

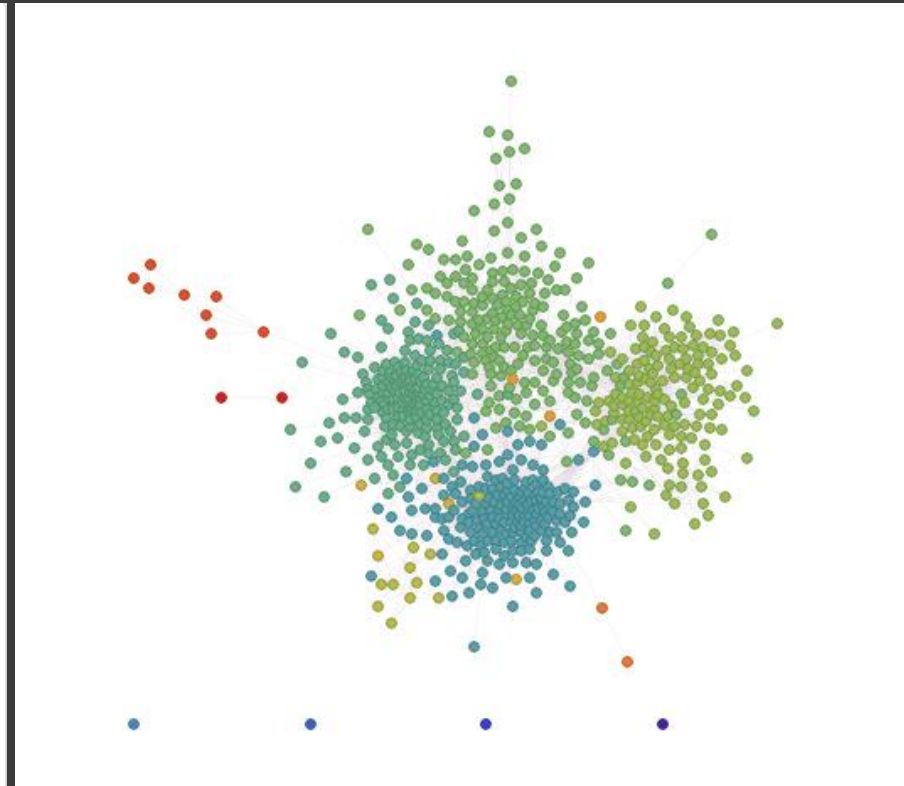
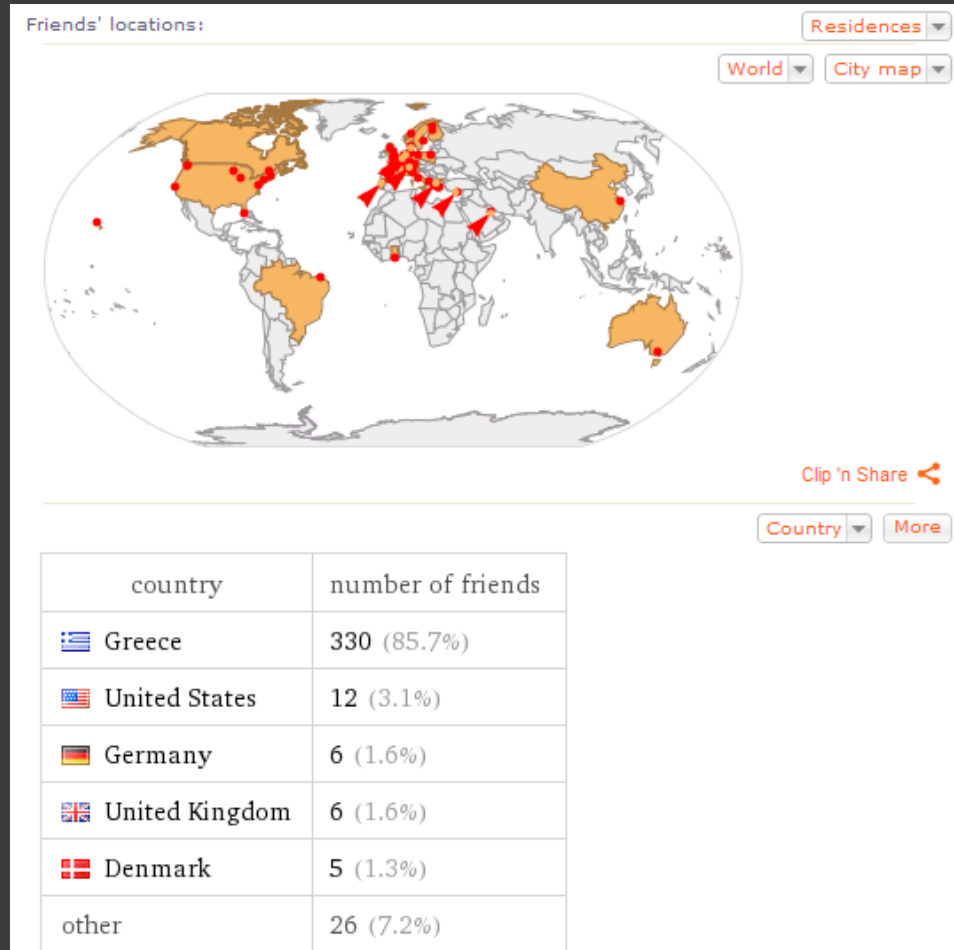
Δημιουργήστε το χάρτη του κοινωνικού σας δικτύου

◎ <http://www.wolframalpha.com/>

- Γράψτε 'Facebook Report'
- Μηνιαία δραστηριότητα (σύνδεσμοι, κείμενο, μεταφορτώσεις φωτογραφιών)
- Φίλοι
- Μέσος αριθμός Like/posts



Facebook report



<http://www.wolframalpha.com/facebook/>

Facebook degrees of separation

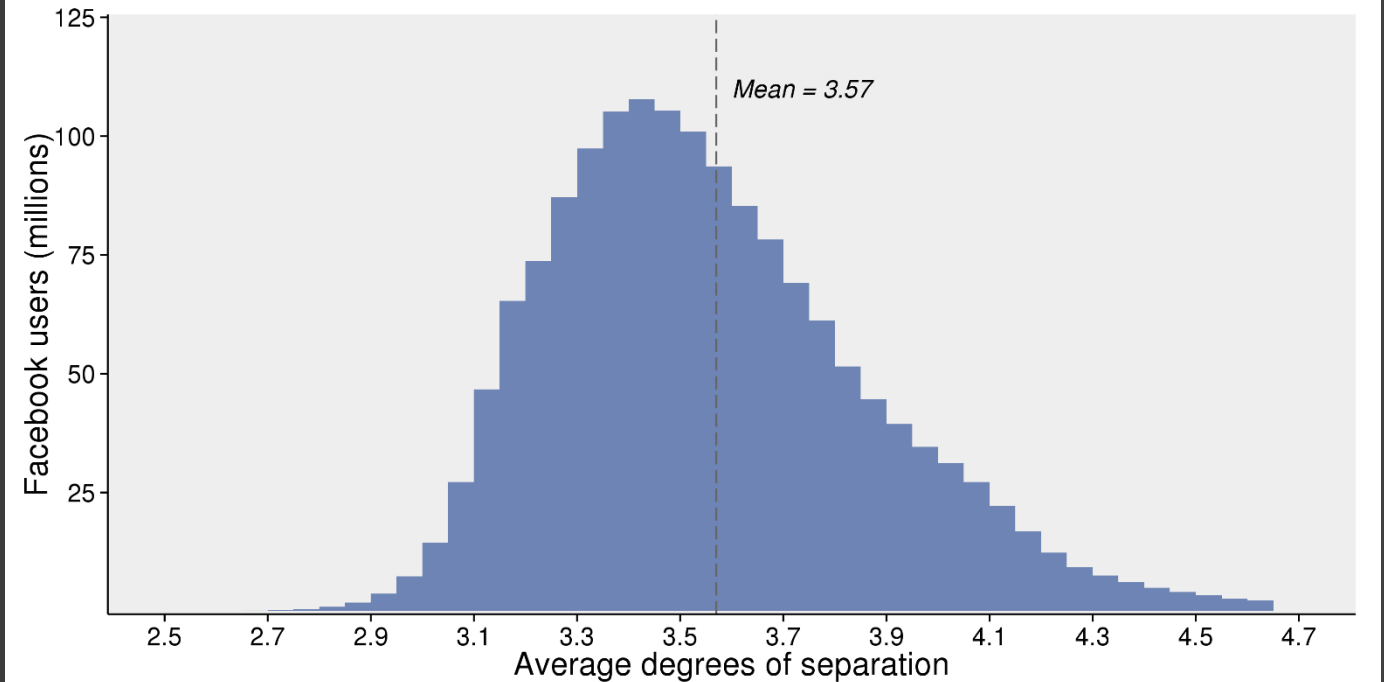
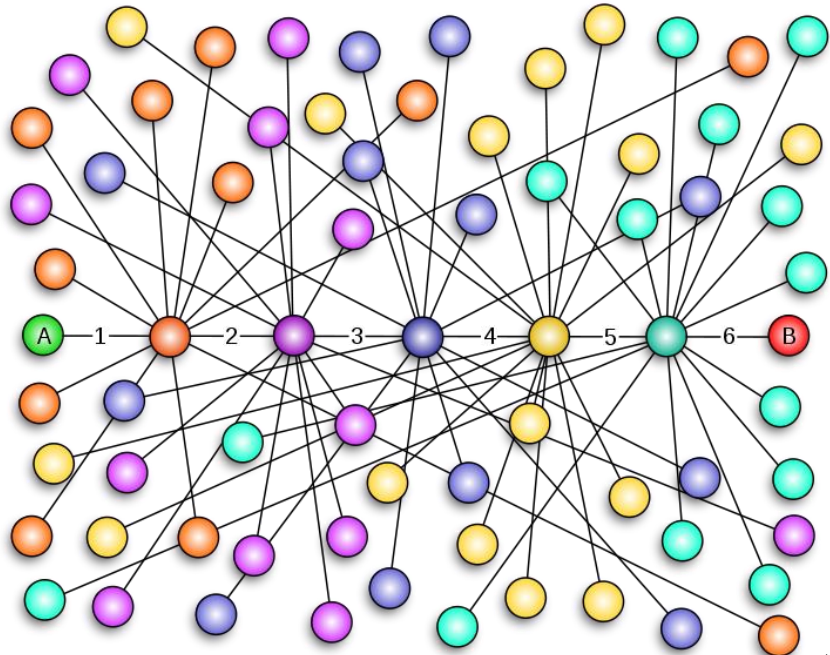
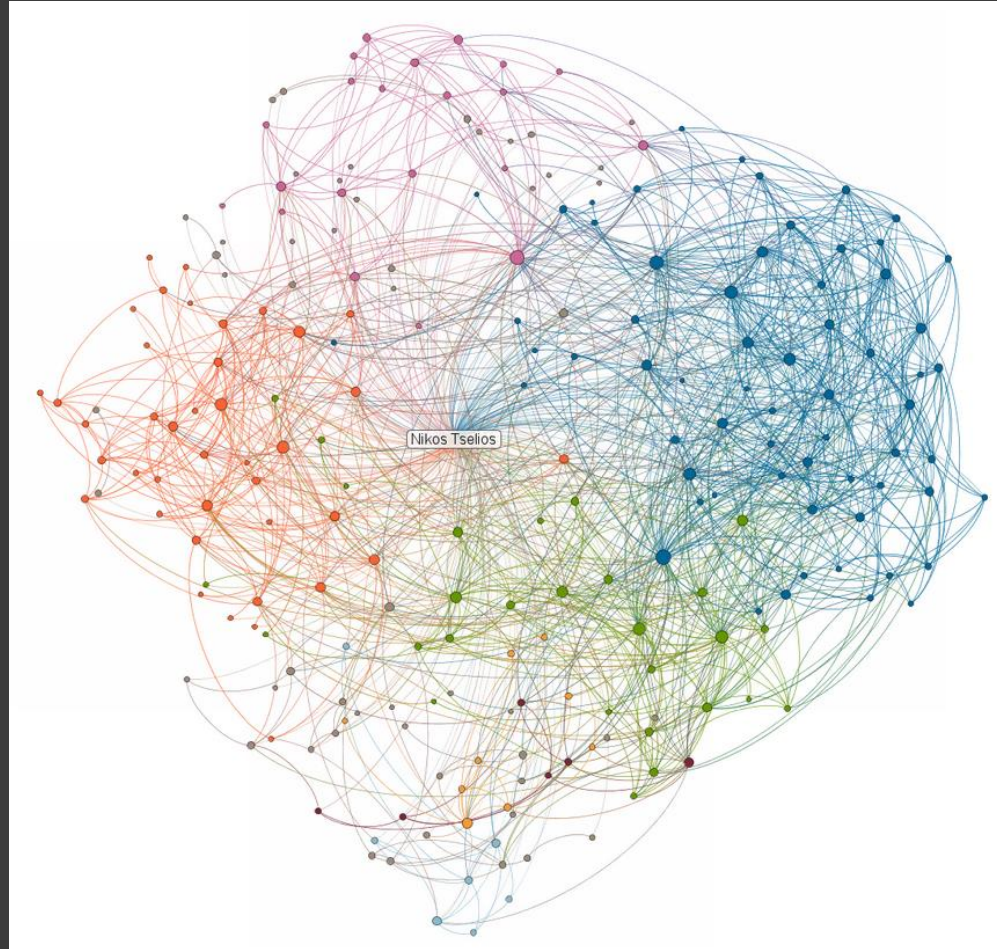


Figure 1. Estimated average degrees of separation between all people on Facebook. The average person is connected to every other person by an average of 3.57 steps. The majority of people have an average between 3 and 4 steps. <https://research.facebook.com/blog/three-and-a-half-degrees-of-separation/>

Linkedin Labs

<http://inmaps.linkedinlabs.com/>

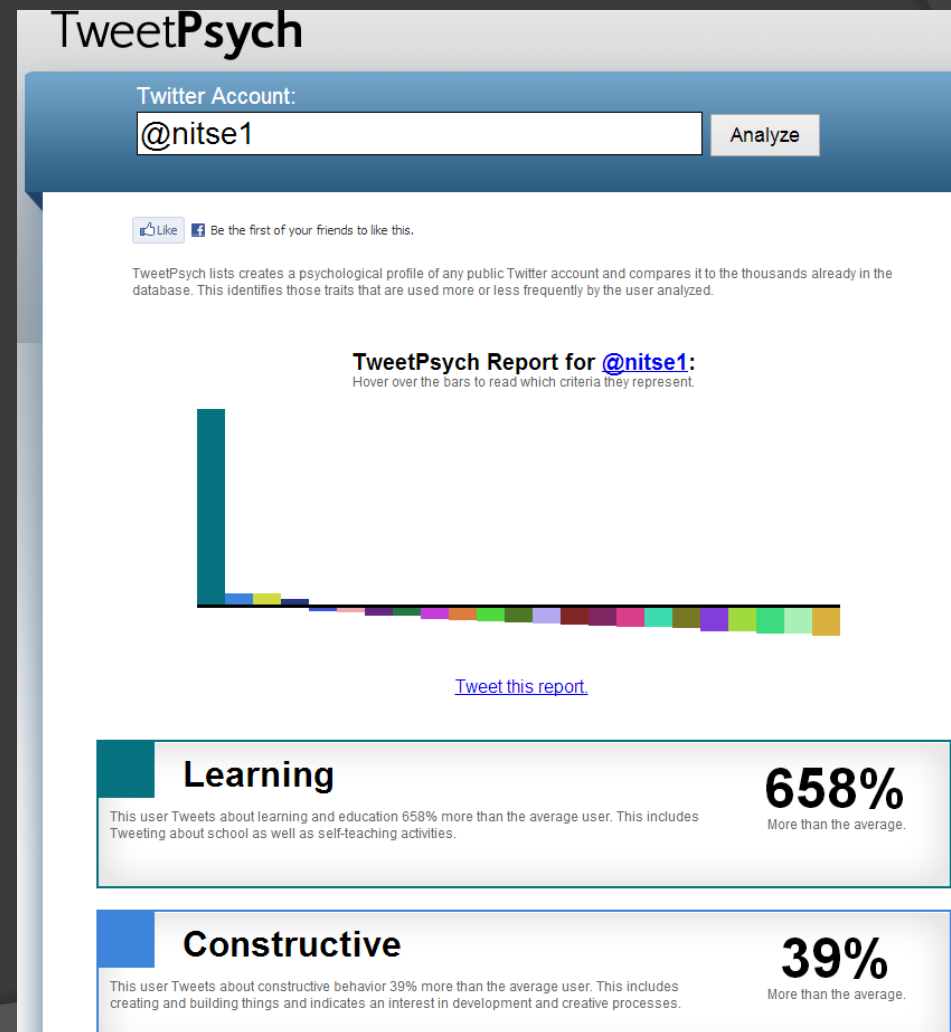


Αυτόματη ανάλυση των Like

- Cambridge Psychometrics Centre/ Microsoft Research Cambridge
- 58.000 US Facebook Users
- Τα facebook likes συνδυάστηκαν με πληροφορίες από τα προφίλ και τεστ προσωπικότητας για να 'εκπαιδεύσουν τους αλγόριθμους'
- Η προβλεπτική ικανότητα ήταν
 - 95% επιτυχία μεταξύ Καυκάσιων και Αφρο-Αμερικάνων
 - Θρήσκευμα 82% επιτυχία
 - Πολιτικές πεποιθήσεις 82%
 - Σχέσεις 65-73%
 - Διαζύγιο γονέων <21, 60%
- Αποδείχτηκε το ίδιο αποτελεσματικό με το τεστ προσωπικότητας
- "Private traits and attributes are predictable from digital records of human behavior," Michal Kosinski, David Stillwell, and Thore Graepel, PNAS, 2013

Twitter analytics

- Κυβερνοψυχογράφημα ☺
- Αναλύει τα tweets ενός χρήστη και εμφανίζει πόσο διαφέρει σε μια σειρά από διαστάσεις από το 'μέσο χρήστη'
- Χρησιμοποιεί αλγόριθμους γλωσσολογικής ανάλυσης
- <http://tweetpsych.com>
- Άλλα εργαλεία ανάλυσης
- <http://www.honeytechblog.com/top-50-twitter-tracking-and-analytics-tools/>



Sentiment analysis

- ⦿ Ανάλυση συναισθημάτων
- ⦿ Με χρήση αλγορίθμων τεχνητής νοημοσύνης γίνεται η ανάλυση κειμένων στα social media (πχ Facebook, Twitter, blogs κ.α.)
- ⦿ Ποσοστό επιτυχίας γύρω στο 70% θεωρείται καλό
- ⦿ Και αυτό γιατί ακόμα και άνθρωποι αξιολογητές έχουν ποσοστό συμφωνίας 79%
- ⦿ Χρησιμοποιείται ως εργαλείο καταγραφής της κοινής γνώμης (πχ marketing, πολιτική, οικονομία)

Google Analytics (1/2)



Google Analytics (2/2)

- Google Analytics: Παρέχει δωρεάν τη δυνατότητα καταγραφής μετρικών επισκεψιμότητας και ανάλυσής τους. Μέσα από τη μελέτη των δεδομένων επισκεψιμότητας μπορούν να αναγνωριστούν πιθανά προβλήματα. Τα δεδομένα συνίστανται σε
 - συνεδρίες χρήσης
 - επισκέψεις
 - θέαση σελίδων
 - στοιχεία επισκέπτη (φυλλομετρητής) κλπ
- Εξαιρετικά καλή λύση μέτρησης επισκεψιμότητας. Ειδικά για eCommerce web sites επιτρέπεται η παρατήρηση και συσχέτιση δεδομένων επισκεψιμότητας με αγορές (conversion events) που επιτρέπει τη διερεύνηση βέλτιστων πρακτικών που μπορούν να αυξήσουν τις πωλήσεις
- Δεν παρέχει εργαλεία διερεύνησης και ερμηνείας της συμπεριφοράς του χρήστη. Θα πρέπει να συνδυάζεται με στοχευμένες συνεδρίες ευχρηστίας με αντιπροσωπευτικούς χρήστες
- **Pros** – Δωρεάν εργαλείο καταγραφής web metrics
- **Cons** – Τα δεδομένα που συλλέγονται δεν μπορούν να ερμηνεύσουν τη συμπεριφορά του χρήστη
- **Pricing** – Δωρεάν

Παράδειγμα: Ανάλυση δεδομένων επιστημονικής έρευνας με το SCImago

SJR

SCImago
Journal & Country
Rank

EST MODUS IN REBUS
Horatio (Satire 1,1,106)

[Home](#)
[Journal Rankings](#)
[Journal Search](#)
[Country Rankings](#)
[Country Search](#)
[Compare](#)
[Map Generator](#)
[Help](#)
[About Us](#)

Science Analysis

The SCImago Journal & Country Rank is a portal that includes the journals and country scientific indicators developed from the information contained in the Scopus® database (Elsevier B.V.). These indicators can be used to assess and analyze scientific domains.

This platform takes its name from the SCImago Journal Rank (SJR) indicator, developed by SCImago from the widely known algorithm Google PageRank™. This indicator shows the visibility of the journals contained in the Scopus® database from 1996.

[Read more about us >](#)

Title	SJR	H index	Total Cites	Quotient
Journal of the American Medical Association	22,438	147	25	84
Journal of the American Academy of Pediatrics	16,100	133	30	86
Journal of the American College of Surgeons	15,224	204	552	1.2
Journal of the American Society of Nephrology	11,193	96	28	
Journal of the American Association of Nurse Practitioners	10,484	133	234	
Journal of the American Association of Nurse Practitioners	12,249	114	175	

Rank

Journals or Countries

Analyze

Journals or Countries

Compare

Journals or Countries

Visualize

Subject areas and categories

What's New?

March 23, 2011
[Scimago Lab's blog on Institutional Collaboration in Global Science](#)

February 21, 2011
[Scimago Lab starts up a new blog](#)

January 17, 2011
[SJR January 2011 Update Release Note](#)

November 03, 2010
[Ranking of Research Institutions SIR World Report 2010 released](#)

December 10, 2009
[SJR 2008 Update. New features](#)

[More news >](#)

e-mail news subscription:

Delivered by FeedBurner

Follow us:

[Twitter](#) [Facebook](#)

SCIMAGO
LAB

Powered by
Scopus

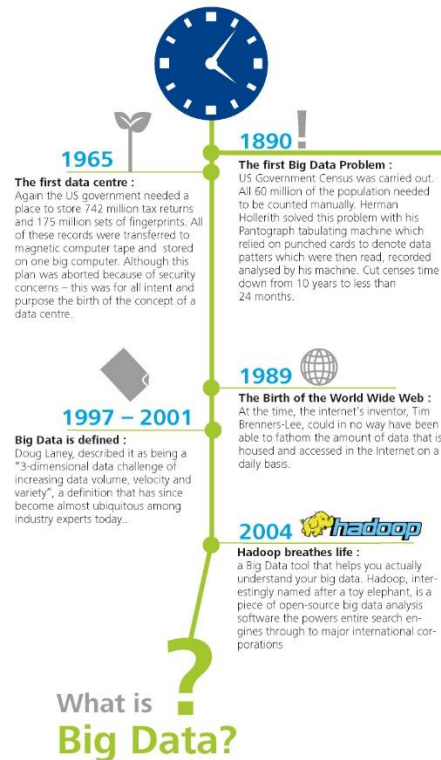
SCImago Lab, Copyright 2007-2013. Data Source: Scopus®

Κατάταξη χωρών με βάση τις αναφορές

	Country	Documents	Citable documents	Citations	Self-Citations	Citations per Document	H index
1	 Switzerland	350.253	329.198	6.873.551	966.536	22,46	537
2	 Denmark	183.880	173.771	3.444.509	514.632	21,17	399
3	 Netherlands	487.784	457.933	8.928.850	1.524.755	20,82	545
4	 United States	6.149.455	5.738.593	114.546.415	54.226.872	20,51	1.305
5	 Sweden	337.135	321.725	6.111.804	1.005.775	19,78	484
6	 Finland	170.476	165.195	2.771.982	462.377	18,28	352
7	 Canada	885.197	836.836	13.928.114	2.727.913	18,19	621
8	 United Kingdom	1.711.878	1.550.373	27.919.060	6.703.673	18,03	802
9	 Belgium	265.913	251.632	4.161.308	630.041	17,81	428
10	 Israel	204.262	194.752	3.283.119	483.335	17,35	393
11	 Norway	141.143	133.311	2.021.938	339.172	17,19	308
12	 Australia	592.533	551.667	8.180.664	1.770.774	16,65	481
13	 Austria	188.440	177.324	2.688.324	387.884	16,51	355
14	 Germany	1.581.429	1.490.140	23.229.085	6.171.727	16,19	704
15	 Ireland	91.125	85.341	1.149.729	141.683	16,18	254
16	 France	1.141.005	1.073.718	16.068.688	3.749.874	15,58	646
17	 New Zealand	114.495	107.441	1.504.946	248.529	15,43	264
18	 Italy	851.692	803.004	11.279.167	2.639.721	15,00	550
19	 Spain	665.977	623.236	7.640.544	1.958.835	13,66	448
20	 Hong Kong	144.935	139.331	1.722.546	262.368	13,52	268
21	 Chile	58.768	56.934	592.148	115.648	13,13	181
22	 Portugal	117.469	113.411	1.150.280	234.405	12,77	218
23	 Singapore	126.881	122.436	1.330.684	191.033	12,51	240
24	 Japan	1.604.017	1.563.732	18.441.796	5.520.032	12,09	602
25	 Greece	160.760	152.000	1.589.963	289.460	11,93	247
26	 Hungary	100.137	96.842	1.058.391	182.169	11,57	239
27	 South Africa	107.976	101.434	1.013.102	225.507	11,11	216
28	 Argentina	105.216	101.695	1.027.553	230.483	11,11	206
29	 Thailand	69.637	67.274	528.969	93.304	10,66	156
30	 South Korea	497.681	487.459	3.988.716	917.147	10,32	309

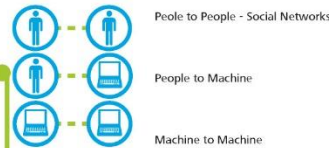
http://www.scimagojr.com/countryrank.php?area=0&category=0®ion=all&year=all&order=cd&min=30000&min_type=it

The Quick Reference Guide To Big Data and Data Analytics



The 3 V's of big data

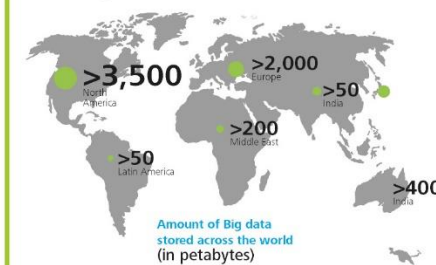
Variety (mixture of structured and unstructured data)



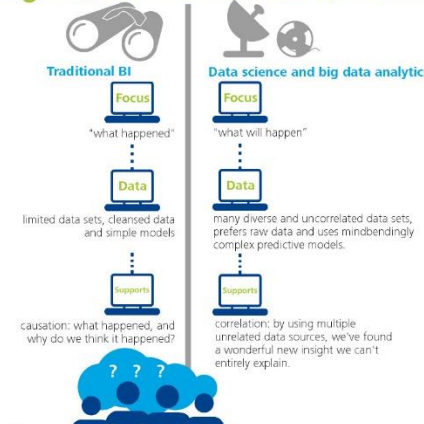
Variety (mixture of structured and unstructured data)

- 48 hours of video uploaded to YouTube every minute
- 2 000 000 queries on Google every minute
- 47 000 app downloads on iTunes every minute

Volume (large amounts of data)



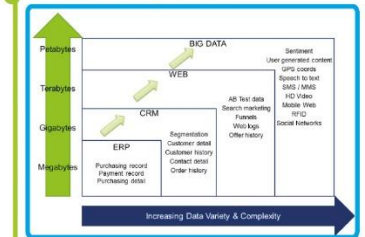
Big Data vs Data Analysis



Leading Question...
How can companies capitalize on insights from "big data"?

- Monitor the flow, rather than a fixed supply, of data.
- Work with data scientists rather than data analysts.
- Integrate analytics into core business and operational functions.

Data vs Big Data



Types of Big Data

- Web and social media**
data includes clickstream and interaction data from social media such as Facebook, Twitter, LinkedIn, and blogs.
- Machine-to-machine**
data includes readings from sensors, meters, and other devices as part of the so-called "Internet of things."
- Big transaction data**
includes healthcare claims, telecommunications call detail records (CDRs), and utility billing records that are increasingly available in semi-structured and unstructured formats.
- Biometric**
data includes fingerprints, genetics, handwriting, retinal scans, and similar types of data.
- Human-generated**
data includes vast quantities of unstructured and semi-structured data such as call center agents' notes, voice recordings, email, paper documents, surveys, and electronic medical records.

Areas where Big Data could be utilized

- Healthcare**
potentially improve the level of healthcare whilst still increasing operating margins.
- Government**
potentially reduce government operating costs
- Retail**
potentially increase retailers operating margins
- Manufacturing**
potentially cut operating costs

Follow @DeloitteTechSA on Twitter for more Big Data information and solutions

Sources: Hadoop.apache.org, wipro.com, capgemini.com

© 2012 Deloitte & Touche. All rights reserved. Member of Deloitte Touche Tohmatsu Limited.

Designed and produced by Creative Solutions at Deloitte, Johannesburg. (0000000/den)

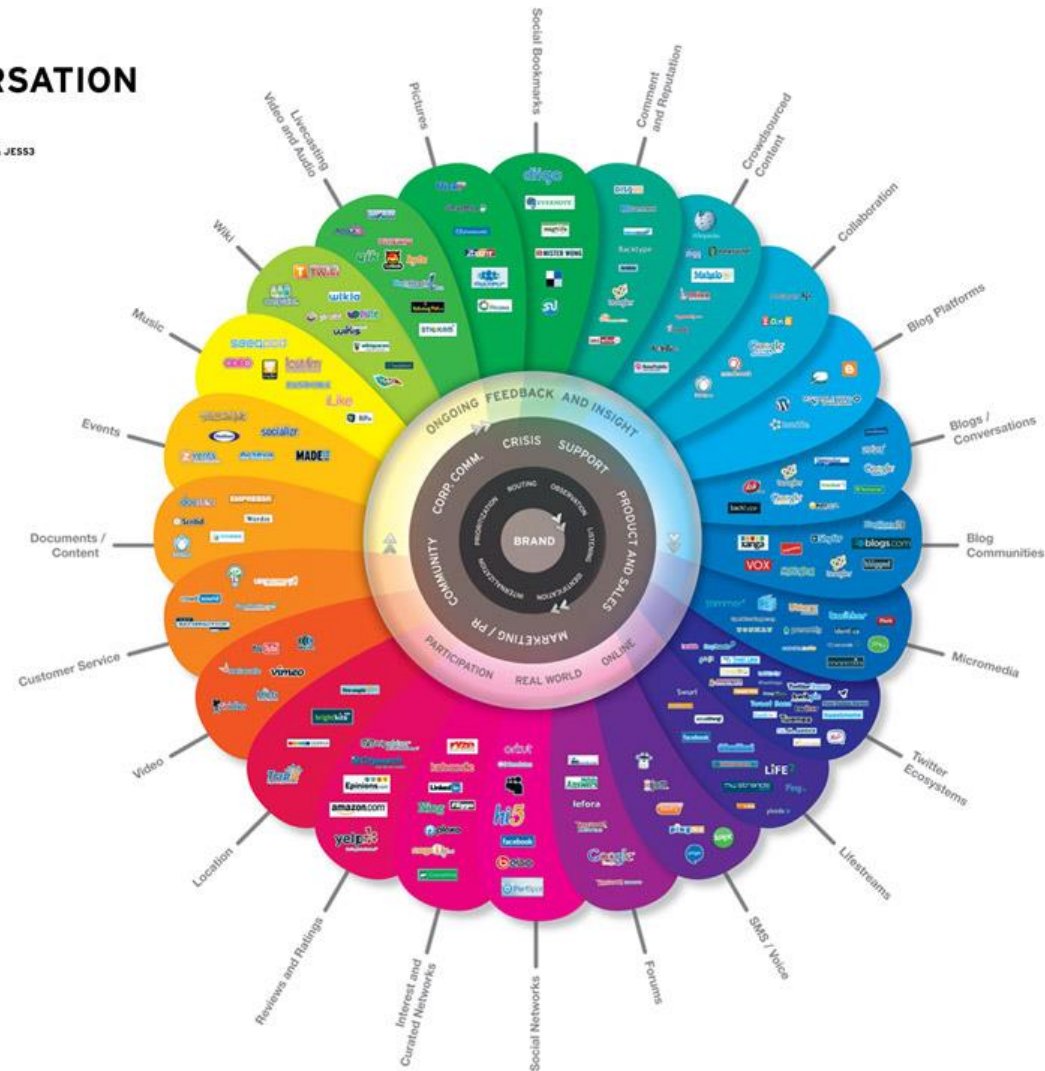
Deloitte Contact

"Contact Deloitte for more information on Big Data and the effects it can have on your business and keep watching for more in depth cloud infographics."

Deloitte.co.za

THE CONVERSATION PRISM

Brought to you by Brian Solis & JESS3



theconversationprism.com

Βήματα για τη δημιουργία της δικής σας παρουσίας στα social media

- ⦿ Φωτογραφία
- ⦿ Σύντομο βιογραφικό (1 παράγραφο)
- ⦿ Εκτενές βιογραφικό
- ⦿ Συλλογή συνδέσμων της 'ψηφιακής παρουσίας σας'
 - Twitter/Facebook/Pinterest
 - LinkedIn, About.me, Academia.edu, Researchgate.net
 - Flickr/Instagram
 - Youtube/Vimeo
 - Blog/ Εταιρικό προφίλ

Troll

- Σκόπιμα 'επιθετική' συμπεριφορά για να προκαλέσει την αντίδραση των χρηστών
- Αποθαρρύνει το διάλογο και την ανταλλαγή απόψεων
- Συχνά δημιουργεί προβλήματα σε χρήστες (συκοφαντία / προσβολή κλπ)
- Ποια είναι τα όρια; Πως προστατευόμαστε χωρίς να περιορίζουμε την ελευθερία της έκφρασης;
 - Αποφεύγουμε να μπαίνουμε σε τέτοιου είδους αντιπαράθεσεις
 - Don't feed the trolls
<http://www.youtube.com/watch?v=ulNSIES1Fds>
- <http://buzzmachine.com/2012/09/13/advice-to-muslims-media-dont-feed-the-trolls/>



Χρήσιμες πηγές

- ◎ <http://socialmediatoday.com/> Άρθρα για τάσεις/ιδέες αξιοποίησης/επίδραση κοινωνικών δικτύων
- ◎ Art of community, Jono Bacon (δωρεάν βιβλίο <http://ubuntuone.com/0n352YwUjlcFR8PjIELH67>)
- ◎ The Wealth of Networks: How Social Production Transforms Markets and Freedom, Yochai Benkler
[http://cyber.law.harvard.edu/wealth_of_networks/Download PDFs of the book](http://cyber.law.harvard.edu/wealth_of_networks/Download_PDFs_of_the_book)
- ◎ Social capital primer. Putnam
http://bowlingalone.com/?page_id=13
- ◎ Social Network Sites: Definition, History and Scholarship
<http://onlinelibrary.wiley.com/doi/10.1111/j.1083-6101.2007.00393.x/pdf>

Τέλος Ενότητας



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης

ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

