





Ανάλυση Πολυμέσων για Εξαγωγή Πληροφορίας από Κοινωνικά Δίκτυα και Αναζήτηση

Δρ. Γιάννης Κομπατσιάρης
Ερευνητής Β' - ΙΠΤΗΛ



Μερικά νούμερα

- 7,000,000 εικόνες στο  (online images marketplace)
- 5,000,000,000 εικόνες στο  (Σεπτ. 2010)
- Μόνο το Σαββατοκύριακο της Πρωτοχρονιάς ανέβηκαν 750,000,000 φωτογ 
- Η εφαρμογή  ξεπέρασε τον αριθμό των 2,000,000 χρηστών μέσα σε δύο μήνες.



Το Πρόβλημα

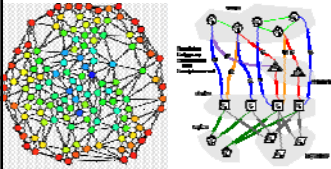
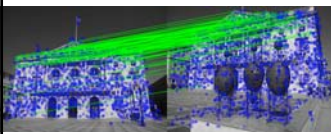


- Υπέρογκα ποσά πληροφορίας
- Ελλιπείς περιγραφές
- Κείμενο όχι πάντοτε επαρκές
- Μη σχετικό περιεχόμενο (θόρυβος)
- Χρήσιμη πληροφορία που δεν προκύπτει από απλή αναζήτηση

Ο κρυμμένος πλούτος

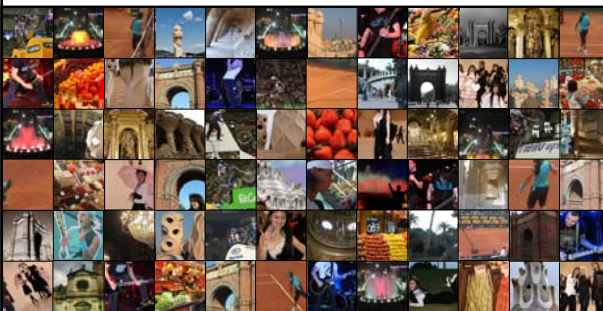


Η τεχνολογία



- Περιγραφή πολυμέσων βάσει τις οπτικής τους πληροφορίας
- Αυτόματη εξαγωγή λεκτικών περιγραφών
- Υπολογισμός ομοιότητας εικόνων
- Ομαδοποίηση και κατηγοριοποίηση εικόνων

Ομαδοποίηση (Clustering)









Η Θεσσαλονίκη μέσα από το ClustTour

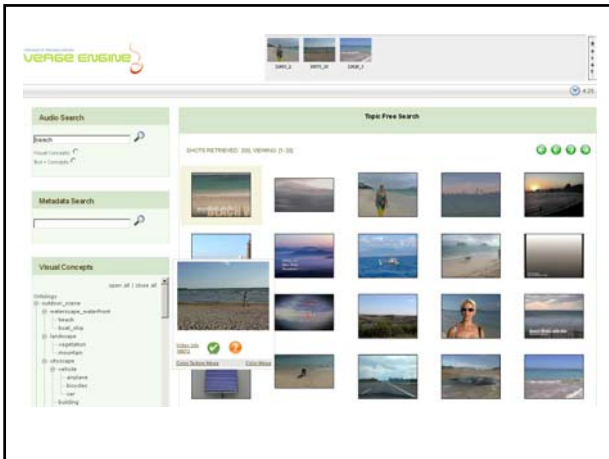
"Everything is automatic and when we say everything we mean everything"



CT: Εφαρμογές, παρόν, μέλλον

- Εφαρμογές
 - Τουριστικοί οδηγοί
 - Υπεύθυνοι προβολής - διαφήμισης
 - Κινητές συσκευές
- Μελλοντικές επεκτάσεις:
 - Εμπλουτισμός με εξωτερικές πηγές περιεχομένου (π.χ. Wikipedia, travel blogs)
 - Εφαρμογή σε real-time δεδομένα
- Εφαρμογή σε άλλους τομείς:
 - Διαχείριση κρίσεων: ομαδοποίηση tweets σε σχέση με γεγονότα έκτακτης ανάγκης (πυρκαγιές, πλημμύρες, κλπ.)
 - Ηλεκτρονική συμμετοχή σε διακυβέρνηση: ομαδοποίηση από γνώμες πολιτών, στατιστική ανάλυση/οπτικοποίηση αποτελεσμάτων
 - Περιβάλλον/Πόλη: Ομαδοποίηση και κατάταξη προβλημάτων (π.χ. από παράπονα πολιτών)
 - Πολιτιστική Κληρονομιά





Αναζήτηση: Εφαρμογές, παρόν, μέλλον

- Εφαρμογές
 - Δελτία ειδήσεων
 - Πατέντες
- Μελλοντικές επεκτάσεις:
 - Implicit user feedback (eye tracking, emotions)
- Εφαρμογή σε άλλους τομείς:
 - Διαδικτυακοί κόμβοι με πολυμεσικό υλικό, stock images/video
 - Ειδησεογραφικά πρακτορεία και εκδοτικοί οίκοι
 - Μουσεία, ιστορικά, πολιτιστικά αρχεία, τηλεοπτικοί σταθμοί
 - Γραφεία ευρεσιτεχνιών και μεγάλες εταιρίες που διαθέτουν τμήματα ανάπτυξης παραγωγής ευρεσιτεχνιών
 - Εταιρείες ανάπτυξης λογισμικού portals, μηχανοργάνωσης, workflows, προβολής, οργάνωσης εταιρικού περιεχομένου

Εκμετάλλευση – Συνεργασίες

- Licensing τεχνολογίας
 - Ολοκλήρωση με υπάρχουσες εφαρμογές για υπηρεσίες προστιθέμενης αξίας
 - Παροχή APIs
- Δημιουργία νέας εταιρείας (Spin-off)
 - Υπηρεσίες για τουρισμό
 - Νέα – Ενημέρωση
 - Brand monitoring
 - Ηλεκτρονική Διακυβέρνηση
 - Διασκέδαση – Πολιτισμός
 - Πατέντες
 - Stock photos

- www.clusttour.gr
- <http://mklab-services.iti.gr/patmedia/>
- <http://mklab.iti.gr/verge/>

Δρ. Γιάννης Κομπατσιάρης
ikom@iti.gr, <http://mklab.iti.gr>

Η έρευνα για το ClustTour γίνεται σε συνεργασία με το ΑΠΘ,
Τμήμα Πληροφορικής, Καθ. Αθηνά Βακάλη