

# ΜΥΕ003: Ανάκτηση Πληροφορίας

*Διδάσκουσα: Ευαγγελία Πιτουρά*

Κεφάλαια 8, 11: Περιλήψεις αποτελεσμάτων, Πιθανοτική ανάκτηση πληροφορίας.

## Τι θα δούμε σήμερα;

- Πιθανοτική ανάκτηση
- Περιλήψεις αποτελεσμάτων
  - Κάνοντας τα καλά αποτελέσματα χρήσιμα

# ΠΙΘΑΝΟΤΙΚΗ ΑΝΑΚΤΗΣΗ ΠΛΗΡΟΦΟΡΙΑΣ

# Πιθανοτική Ανάκτηση Πληροφορίας

- Βασική ιδέα: Διάταξη εγγράφων με βάση την *πιθανότητα* να είναι συναφή με το ερώτημα
- Τυχαία μεταβλητή  $R$
- Διάταξη με βάση το
$$P(R=1 | q, d)$$

## Αρχή Πιθανοτικής Κατάταξης

«Αν η απόκριση ενός συστήματος ανάκτησης πληροφορίας σε κάθε ερώτημα είναι η *κατάταξη των εγγράφων κατά φθίνουσα πιθανότητα συνάφειας* προς τον χρήστη που υπέβαλε το ερώτημα, όπου οι πιθανότητες εκτιμώνται με τον ακριβέστερο δυνατό τρόπο βάσει των όποιων δεδομένων είναι διαθέσιμα στο σύστημα για τον σκοπό αυτό, η συλλογική *αποτελεσματικότητα είναι η βέλτιστη εφικτή* με βάση αυτά τα δεδομένα» [van Rijsbergen 1979]

Probability Ranking Principle (RPR):

decreasing probability of relevance => best overall effectiveness

## Δυαδικό Μοντέλο Ανεξαρτησίας (Binary Independence Model (BIP))

Θα αναπαραστήσουμε κάθε έγγραφο  $d$  (και αντίστοιχα, το ερώτημα) ως ένα δυαδικό (Boolean) διάνυσμα

$$\vec{x} = (x_1, \dots, x_M)$$

$$x_i = 1 \quad \text{ανν το έγγραφο περιέχει τον όρο } i$$

**Υπόθεση ανεξαρτησίας όρων** (independence assumption):  
Οι όροι εμφανίζονται στα έγγραφα ανεξάρτητα ο ένας από τον άλλο

Επίσης, η συνάφεια ενός εγγράφου είναι ανεξάρτητη από τη συνάφεια των άλλων εγγράφων (**ανεξαρτησία εγγράφων**)

## Δυαδικό Μοντέλο Ανεξαρτησίας

Επειδή μας ενδιαφέρει *η διάταξη* θα χρησιμοποιήσουμε τη σχετική πιθανότητα - **odds**

$$O(R | q, \vec{x}) = \frac{p(R = 1 | q, \vec{x})}{p(R = 0 | q, \vec{x})}$$

Από τον κανόνα του Bayes

$$O(R | q, \vec{x}) = \frac{p(R = 1 | q, \vec{x})}{p(R = 0 | q, \vec{x})} = \frac{\frac{p(R = 1 | q) p(\vec{x} | R = 1, q)}{p(\vec{x} | q)}}{\frac{p(R = 0 | q) p(\vec{x} | R = 0, q)}{p(\vec{x} | q)}}$$

## Δυαδικό Μοντέλο Ανεξαρτησίας

$$O(R | q, \vec{x}) = \frac{p(R = 1 | q, \vec{x})}{p(R = 0 | q, \vec{x})} = \frac{p(R = 1 | q)}{p(R = 0 | q)} \frac{p(\vec{x} | R = 1, q)}{p(\vec{x} | R = 0, q)}$$

Σταθερά για  
κάθε ερώτημα

Εκ των προτέρων πιθανότητα (prior probability) να ανακτήσουμε συναφές (μη συναφές) ερώτημα

Πρέπει να  
εκτιμήσουμε αυτήν  
την ποσότητα

Αν έχουμε ανακτήσει ένα συναφές (μη συναφές) έγγραφο, αυτό να έχει αναπαράσταση  $\vec{x}$



## Δυαδικό Μοντέλο Ανεξαρτησίας

Θα χρησιμοποιήσουμε την *υπόθεση ανεξαρτησίας των όρων* (η παρουσία ή απουσία ενός όρου σε ένα έγγραφο είναι ανεξάρτητη από την παρουσία ή απουσία οποιουδήποτε άλλου όρου – Naïve Bayes conditional independent assumption)

$$\frac{p(\vec{x} | R = 1, q)}{p(\vec{x} | R = 0, q)} = \prod_{i=1}^M \frac{p(x_i | R = 1, q)}{p(x_i | R = 0, q)}$$

Άρα:

$$O(R | q, \vec{x}) = O(R | q) \prod_{i=1}^n \frac{p(x_i | R = 1, q)}{p(x_i | R = 0, q)}$$

## Δυαδικό Μοντέλο Ανεξαρτησίας

$$O(R | q, \vec{x}) = O(R | q) \prod_{i=1}^n \frac{p(x_i | R = 1, q)}{p(x_i | R = 0, q)}$$

Αφού τα  $x_i$  είναι είτε 1 είτε 0, σπάμε το γινόμενο

$$O(R | q, \vec{x}) = O(R | q) \prod_{x_i=1} \frac{p(x_i = 1 | R = 1, q)}{p(x_i = 1 | R = 0, q)} \prod_{x_i=0} \frac{p(x_i = 0 | R = 1, q)}{p(x_i = 0 | R = 0, q)}$$

## Δυαδικό Μοντέλο Ανεξαρτησίας

Για τον όρο  $i$ , ορίζουμε ως  $p_i$  την πιθανότητα να εμφανίζεται σε **συναφές** και ως  $r_i$  την πιθανότητα να εμφανίζεται σε **μη συναφές**

$$p_i = p(x_i = 1 | R = 1, q)$$

$$r_i = p(x_i = 1 | R = 0, q)$$

	Έγγραφο	Συναφές (R=1)	Μη συναφές (R=0)
Όρος $i$ εμφανίζεται	$x_i = 1$	$p_i$	$r_i$
Όρος $i$ δεν εμφανίζεται	$x_i = 0$	$(1 - p_i)$	$(1 - r_i)$

## Δυαδικό Μοντέλο Ανεξαρτησίας

Αν υποθέσουμε ότι οι όροι που δεν εμφανίζονται στο ερώτημα είναι το ίδιο πιθανό να εμφανίζονται σε συναφή και σε μη συναφή

$$p_i = r_i$$

$$O(R | q, \vec{x}) = O(R | q) \prod_{\substack{x_i=1 \\ q_i=1}} \frac{p_i}{r_i} \prod_{\substack{x_i=0 \\ q_i=1}} \frac{(1-p_i)}{(1-r_i)}$$

## Δυαδικό Μοντέλο Ανεξαρτησίας

$$O(R | q, \vec{x}) = O(R | q) \prod_{x_i=q_i=1} \frac{p_i}{r_i} \prod_{\substack{x_i=0 \\ q_i=1}} \frac{1-p_i}{1-r_i}$$

Οι όροι του ερωτήματος που εμφανίζονται

Οι όροι του ερωτήματος που δεν εμφανίζονται

Μοιράζουμε το μεσαίο όρο (=1) στα αριστερά και δεξιά (είναι οι όροι του ερωτήματος που εμφανίζονται στο έγγραφο)

$$O(R | q, \vec{x}) = O(R | q) \prod_{\substack{x_i=1 \\ q_i=1}} \frac{p_i}{r_i} \prod_{\substack{x_i=1 \\ q_i=1}} \left( \frac{1-r_i}{1-p_i} \cdot \frac{1-p_i}{1-r_i} \right) \prod_{\substack{x_i=0 \\ q_i=1}} \frac{1-p_i}{1-r_i}$$

$$O(R | q, \vec{x}) = O(R | q) \prod_{x_i=q_i=1} \frac{p_i(1-r_i)}{r_i(1-p_i)} \prod_{q_i=1} \frac{1-p_i}{1-r_i}$$

Οι όροι του ερωτήματος που εμφανίζονται

Όλοι οι όροι

## Δυαδικό Μοντέλο Ανεξαρτησίας

$$O(R | q, \vec{x}) = O(R | q) \cdot \prod_{x_i=q_i=1} \frac{p_i(1-r_i)}{r_i(1-p_i)} \cdot \prod_{q_i=1} \frac{1-p_i}{1-r_i}$$

Σταθερά για κάθε ερώτημα

Η μοναδική ποσότητα που πρέπει να εκτιμηθεί

Θα χρησιμοποιήσουμε το  $\log$  αυτής της ποσότητας  
**Retrieval Status Value (RSV)** (τιμή κατάστασης ανάκτησης)

$$RSV = \log \prod_{x_i=q_i=1} \frac{p_i(1-r_i)}{r_i(1-p_i)} = \sum_{x_i=q_i=1} \log \frac{p_i(1-r_i)}{r_i(1-p_i)}$$

## Τιμή Κατάστασης Ανάκτησης (RSV)

$$RSV = \log \prod_{x_i=q_i=1} \frac{p_i(1-r_i)}{r_i(1-p_i)} = \sum_{x_i=q_i=1} \log \frac{p_i(1-r_i)}{r_i(1-p_i)}$$

$$RSV = \sum_{x_i=q_i=1} c_i \quad c_i = \log \frac{p_i(1-r_i)}{r_i(1-p_i)}$$

Τα  $c_i$  (log odd ratios) έχουν το ρόλο των βαρών σε αυτό το μοντέλο

	Έγγραφο	Συναφές (R=1)	Μη συναφές (R=0)
Όρος $i$ εμφανίζεται	$x_i = 1$	$p_i$	$r_i$
Όρος $i$ δεν εμφανίζεται	$x_i = 0$	$(1 - p_i)$	$(1 - r_i)$

Πως θα εκτιμήσουμε αυτά τα  $c_i$ ;

## Θεωρητική Εκτίμηση των RSV συντελεστών

Για κάθε όρο  $i$ , κατασκευάζουμε έναν πίνακα μετρητών

Έστω  $N$  έγγραφα,  $S$  συναφή, και ο όρος εμφανίζεται σε συνολικά σε  $n$  έγγραφα από τα οποία τα συναφή είναι  $s$

Έγγραφα	Συναφή (R=1)	Μη συναφή (R=0)	Συνολικά
$x_i = 1$	$s$	$n - s$	$n$
$x_i = 0$	$S - s$	$N - n - (S - s)$	$N - n$
	$S$	$N - S$	$N$

Εκτιμήσεις:

$$p_i \approx \frac{s}{S} \quad r_i \approx \frac{(n-s)}{(N-S)}$$

$$c_i \approx K(N, n, S, s) = \log \frac{s/(S-s)}{(n-s)/(N-n-S+s)}$$

Αγνοεί τους  
όρους που δεν  
εμφανίζονται



## Εκτίμηση των RSV συντελεστών

Αν τώρα υποθέσουμε ότι τα μη συναφή είναι περίπου όλη η συλλογή, τότε το  $r_i$  (η πιθανότητα εμφάνισης του όρου σε μη συναφές έγγραφο) είναι  $n/N$

$$\log \frac{1 - r_i}{r_i} = \log \frac{N - n - S + s}{n - s} \approx \log \frac{N - n}{n} \approx \log \frac{N}{n} = \text{IDF!}$$

# Εκτίμηση των RSV συντελεστών

- Το  $p_i$  (πιθανότητα εμφάνισης σε συναφή έγγραφα) είναι πιο δύσκολο να εκτιμηθεί
- Πιθανοί τρόποι εκτίμησης του  $p_i$ :
  - Από συναφή έγγραφα, αν κάποια από αυτά είναι γνωστά
  - Με μία σταθερά – τότε απλώς χρησιμοποιούμε το idf (with  $p_i=0.5$ )

$$RSV = \sum_{x_i=q_i=1} \log \frac{N}{n_i}$$

- Ανάλογο των εμφανίσεων του όρου στη συλλογή  
έχει προταθεί:  $1/3 + 2/3 \text{ df}_i/N$

# Πιθανοτική Ανάδραση Συνάφειας (Relevance Feedback)

1. Υποθέτουμε κάποιες αρχικές τιμές για τα  $p_i$  και  $r_i$  - τις οποίες χρησιμοποιούμε για να ανακτήσουμε ένα αρχικό σύνολο συναφών εγγράφων (εγγράφων με  $R=1$ )
2. Αλληλοεπιδρούμε με το χρήστη για να βελτιώσουμε αυτές τις τιμές: οι χρήστες χαρακτηρίζουν ένα σύνολο έγγραφων  $V$  ως συναφή ( $R = 1$ ) και μη συναφή ( $R = 0$ )
3. Επανα-υπολογίζουμε τα  $p_i$  και  $r_i$ 
  - Ή τα συνδυάζουμε με τις αρχικές μας εκτιμήσεις χρήση Bayesian prior):

$$p_i^{(2)} = \frac{|V_i| + \kappa p_i^{(1)}}{|V| + \kappa}$$

$\kappa$ : prior weight

4. Επαναλαμβάνουμε, έως σύγκλιση

# PRP και BIM

(+) Λογικές προσεγγίσεις για τις πιθανότητες

(-) Περιοριστικές υποθέσεις:

- Ανεξαρτησία όρων
- Ανεξαρτησία συνάφειας εγγράφων
- Οι όροι που δεν εμφανίζονται στο ερώτημα δεν επηρεάζουν το αποτέλεσμα
- Δυαδική αναπαράσταση (αγνοούμε tf, μήκος εγγράφου, κλπ)

# Στάθμιση Οκαρι BM25

- BM25 “Best Match 25”
- Αναπτύχθηκε στα πλαίσια του συστήματος Οκαρι
- Πιθανοτικό μη δυαδικό μοντέλο που λαμβάνει υπόψη συχνότητες όρων και μήκη εγγράφων

# Αρχικές Εκδοχές του BM25

Εκδοχή 1:

$$c_i^{BM25v1}(tf_i) = c_i^{BIM} \frac{tf_i}{k_1 + tf_i}$$

Εκδοχή 2 (απλοποίηση με IDF):

$$c_i^{BM25v2}(tf_i) = \log \frac{N}{df_i} \times \frac{(k_1 + 1)tf_i}{k_1 + tf_i}$$

# Στάθμιση BM25 με Κανονικοποίηση με το Μήκος Εγγράφου

Μεγαλύτερα έγγραφα πιο πιθανόν να έχουν μεγάλες τιμές  $tf_i$

Μεγάλα έγγραφα

- πλεονασμός (verbosity) το  $tf_i$  που παρατηρούμε είναι πού μεγάλο
- γενικότερου σκοπού – το  $tf_i$  που παρατηρούμε μπορεί να είναι ακριβές

Πρέπει να σταθμίσουμε αυτά τα δύο

# Στάθμιση BM25 με Κανονικοποίηση με το Μήκος Εγγράφου

- Μήκος εγγράφου

$$dl = \sum_{i \in V} tf_i \quad B = \left( (1-b) + b \frac{dl}{avdl} \right)$$

- $avdl$ : μέσο μήκος εγγράφου στη συλλογή
- $b$  - Συντελεστής κανονικοποίησης μήκους

Κανονικοποίηση

$$tf_i' = \frac{tf_i}{B}$$



# Οκρί BM25

$$RSV^{BM25} = \sum_{i \in q} \log \frac{N}{df_i} \cdot \frac{(k_1 + 1)tf_i}{k_1((1-b) + b \frac{dl}{avdl}) + tf_i}$$

- $k_1$  ελέγχει τη διαβάθμιση της συχνότητα όρου (term frequency scaling)
  - $k_1 = 0$  δυαδικό μοντέλο;  $k_1 =$  μεγάλο – μετράει η καθαρή συχνότητα
- $b$  ελέγχει την κανονικοποίηση του μήκους εγγράφου
  - $b = 0$  μη κανονικοποίηση;  $b = 1$  είναι η σχετική συχνότητα (πλήρης κανονικοποίηση)
- Συνήθως,  $k_1$  στο 1.2–2 και το  $b$  γύρω στο 0.75
- Υπάρχουν και εκδοχές του BM25 που συμπεριλαμβάνουν και βάρη στους όρους του ερωτήματος καθώς και ανάδραση συνάφειας.

## Τι θα δούμε σήμερα;

- Πιθανοτική ανάκτηση
- Περιλήψεις αποτελεσμάτων
  - Κάνοντας τα καλά αποτελέσματα χρήσιμα

# Παρένθεση 1: Ασκήσεις

## Ακολουθούμε τις οδηγίες

- Για τον τρόπο υποβολής (pdf, email, κλπ)
- Για τις περιεχόμενο (10000 όχι 1000, κλπ)
- κλπ

Θα αφαιρεθούν βαθμοί

# Παρένθεση 2: emails

1. **username**
  - Στέλνουμε από το account του Τμήματος
  - Σε κάθε περίπτωση: Το username πρέπει να είναι «σοβαρό»
2. Δε ξεχνάμε το **subject** (το οποίο πρέπει να είναι ενδεικτικό του περιεχομένου)
3. Προσφώνηση
  - ανάλογα με τον αποδέκτη
  - αν θέλετε να αναφερθείτε προσωπικά: Τίτλος Επώνυμο
    - Πχ Αγαπητή κ. Πιτουρά ή Dear Prof. Pitoura, **όχι κυρία ή κυρία Μαρία**
4. Υπογράφουμε (Όνομα Επώνυμο)
5. Να αποφεύγονται:
  - ορθογραφικά λάθη, ιδιωματισμοί, πολλές συντομογραφίες, έλλειψη τόνων, ...
6. Να είστε όσο σύντομοι χρειάζεται

ΠΩΣ ΠΑΡΟΥΣΙΑΖΟΥΜΕ ΤΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΣΤΟ  
ΧΡΗΣΤΗ;

## Περιλήψεις αποτελεσμάτων

- Αφού έχουμε διατάξει τα έγγραφα που ταιριάζουν με το ερώτημα, θέλουμε να τα παρουσιάσουμε στο χρήστη
- Πιο συχνά ως μια λίστα από **τίτλους** εγγράφων, URL, μαζί με μια μικρή **περίληψη (result snippet)**, aka “10 blue links”

### [John McCain](#)

**John McCain** 2008 - The Official Website of **John McCain's** 2008 Campaign for President ... African American Coalition; Americans of Faith; American Indians for **McCain**; Americans with ...  
[www.johnmccain.com](http://www.johnmccain.com) · [Cached page](#)

### [JohnMcCain.com - McCain-Palin 2008](#)

**John McCain** 2008 - The Official Website of **John McCain's** 2008 Campaign for President ... African American Coalition; Americans of Faith; American Indians for **McCain**; Americans with ...  
[www.johnmccain.com/Informing/Issues](http://www.johnmccain.com/Informing/Issues) · [Cached page](#)

### [John McCain News- msnbc.com](#)

Complete political coverage of **John McCain**. ... Republican leaders said Saturday that they were worried that Sen. **John McCain** was heading for defeat unless he brought stability to ...  
[www.msnbc.msn.com/id/16438320](http://www.msnbc.msn.com/id/16438320) · [Cached page](#)

### [John McCain | Facebook](#)

Welcome to the official Facebook Page of **John McCain**. Get exclusive content and interact with **John McCain** right from Facebook. Join Facebook to create your own Page or to start ...  
[www.facebook.com/johnmccain](http://www.facebook.com/johnmccain) · [Cached page](#)

## Περιλήψεις αποτελεσμάτων

- Η περιγραφή του εγγράφου είναι *κρίσιμη* γιατί συχνά οι χρήστες βασίζονται σε αυτήν για να αποφασίσουν αν το έγγραφο είναι σχετικό
  - Δε χρειάζεται να διαλέξουν ένα-ένα τα έγγραφα με τη σειρά

*Ο τίτλος αυτόματα από μεταδεδομένα, αλλά πώς να υπολογίσουμε τις περιλήψεις;*

# Περίληψεις αποτελεσμάτων

## Δύο βασικά είδη περιλήψεων

- Μια **στατική περίληψη** (static summary) ενός εγγράφου είναι πάντα η ίδια ανεξάρτητα από το ερώτημα που έθεσε ο χρήστης
- Μια **δυναμική περίληψη** (dynamic summary) εξαρτάται από το ερώτημα (**query-dependent**). Προσπαθεί να εξηγήσει γιατί το έγγραφο ανακτήθηκε για το συγκεκριμένο κάθε φορά ερώτημα



## Στατικές Περιλήψεις

- Σε ένα τυπικό σύστημα η στατική περίληψη είναι ένα υποσύνολο του εγγράφου
- Απλός ευριστικός: οι πρώτες περίπου 50 λέξεις του εγγράφου
  - cached κατά τη δημιουργία του ευρετηρίου
- Πιο εξελιγμένες μέθοδοι (text summarization) – βρες από κάθε έγγραφο κάποιες **σημαντικές προτάσεις**
  - Απλή γλωσσολογική επεξεργασία (NLP) με ευριστικά για να βαθμολογηθεί κάθε πρόταση (πληροφορία θέσης: πρώτη και τελευταία παράγραφος, πρώτη και τελευταία πρόταση στην παράγραφο, και περιεχομένου: σημαντικές λέξεις)
  - Η περίληψη αποτελείται από τις προτάσεις με το μεγαλύτερο βαθμό
- Ή και πιο περίπλοκη γλωσσολογική επεξεργασία για τη σύνθεση/δημιουργία περίληψης

## Δυναμικές Περιλήψεις

- Παρουσίασε ένα ή περισσότερα «*παράθυρα*» (windows, snippets) μέσα στο έγγραφο που να περιέχουν αρκετούς από τους όρους του ερωτήματος
  - “*KWIC*” snippets: αναπαράσταση Keyword-in-Context

The image displays three search engine results for the query "christopher manning". Each result shows a search engine logo, a search bar with the query, and a snippet of text from the search results.

**Google**™ christopher manning  
**Christopher Manning, Stanford NLP**  
 Christopher Manning, Associate Professor of Computer Science and Linguistics, Stanford University.  
[nlp.stanford.edu/~manning/](http://nlp.stanford.edu/~manning/) - 12k - [Cached](#) - [Similar pages](#)

**Google**™ christopher manning machine translation  
**Christopher Manning, Stanford NLP**  
 Christopher Manning, Associate Professor of Computer Science and Linguistics, ... computational semantics, **machine translation**, grammar induction, ...  
[nlp.stanford.edu/~manning/](http://nlp.stanford.edu/~manning/) - 12k - [Cached](#) - [Similar pages](#)

**YAHOO!**™ christopher manning  
**Christopher Manning, Stanford NLP**  
 Christopher Manning, Associate Professor of Computer Science and Linguistics, Stanford University ... **Chris Manning** works on systems and formalisms that can ...  
[nlp.stanford.edu/~manning](http://nlp.stanford.edu/~manning/) - [Cached](#)

## Δυναμικές Περιλήψεις

- Για τον υπολογισμό τους χρειαζόμαστε τα ίδια τα έγγραφα (δεν αρκεί το ευρετήριο)
- Cache εγγράφων – που πρέπει να ανανεώνεται
- Συχνά όχι όλο το έγγραφο αν είναι πολύ μεγάλο, αλλά κάποιο πρόθεμα του
- Βρες μικρά παράθυρα στα έγγραφα που περιέχουν όρους του ερωτήματος
  - Απαιτεί γρήγορη αναζήτηση παράθυρου στην cache των εγγράφων

## Δυναμικές Περιλήψεις

- Βαθμολόγησε κάθε παράθυρο ως προς το ερώτημα
  - Με βάση διάφορα χαρακτηριστικά: το πλάτος του παραθύρου, τη θέση του στο έγγραφο, κλπ
  - Συνδύασε τα χαρακτηριστικά
- Δύσκολο να εκτιμηθεί η ποιότητα
- Positional indexes (words vs bytes)
- Ο χώρος που διατίθεται για τα παράθυρα είναι μικρός

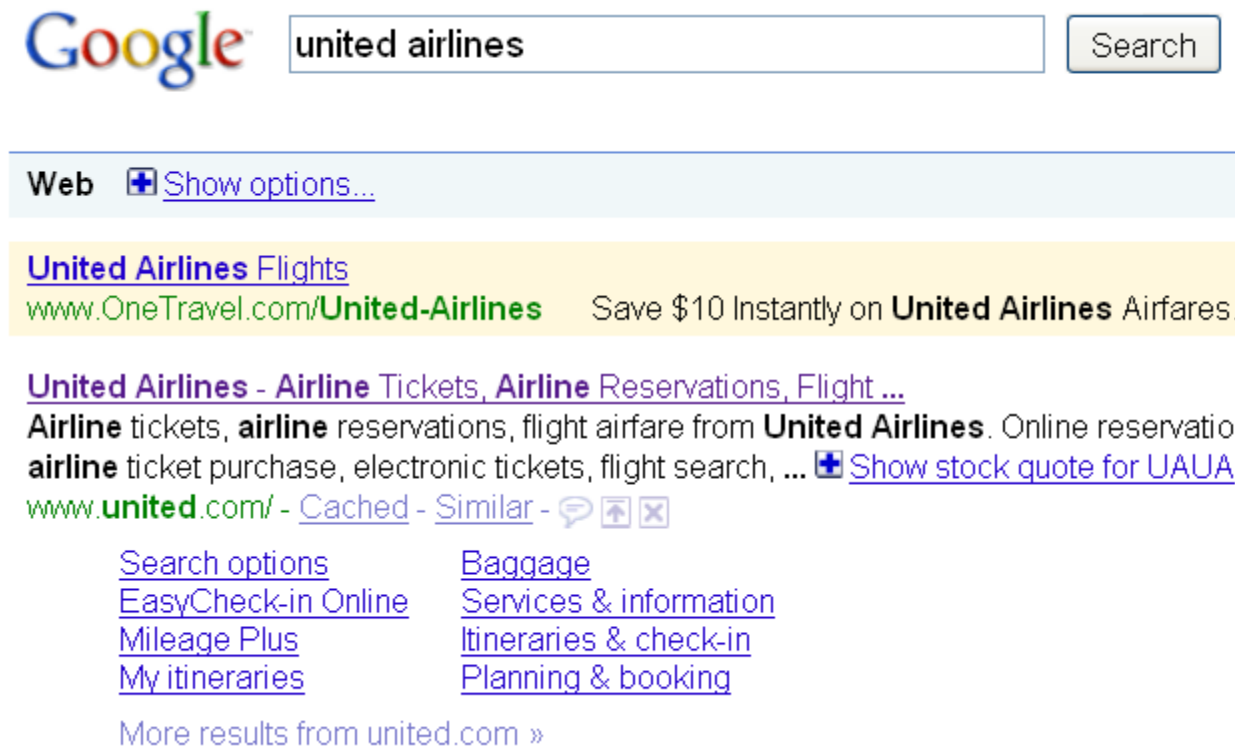
## Δυναμικές Περιλήψεις

Query: “new guinea economic development” Snippets (in bold) that were extracted from a document: . . .

**In recent years, Papua New Guinea has faced severe economic difficulties and** economic growth has slowed, partly as a result of weak governance and civil war, and partly as a result of external factors such as the Bougainville civil war which led to the closure in 1989 of the Panguna mine (at that time the most important foreign exchange earner and contributor to Government finances), the Asian financial crisis, a decline in the prices of gold and copper, and a fall in the production of oil. **PNG’s economic development record over the past few years is evidence that** governance issues underly many of the country’s problems. Good governance, which may be defined as the transparent and accountable management of human, natural, economic and financial resources for the purposes of equitable and sustainable development, flows from proper public sector management, efficient fiscal and accounting mechanisms, and a willingness to make service delivery a priority in practice. . . .

# Quicklinks

- Για *navigational query* (όταν ψάχνουμε μια συγκεκριμένη σελίδα) όπως **united airlines** οι χρήστες πιθανόν να ικανοποιούνται από τη σελίδα [www.united.com](http://www.united.com)
- Quicklinks παρέχουν navigational cues σε αυτή τη σελίδα



The image shows a Google search interface. The search bar contains the text "united airlines" and a "Search" button. Below the search bar, there is a "Web" section with a plus icon and a link to "Show options...". The first search result is highlighted in yellow and is for "United Airlines Flights" with the URL "www.OneTravel.com/United-Airlines" and a snippet "Save \$10 Instantly on United Airlines Airfares.". Below this, there is another search result for "United Airlines - Airline Tickets, Airline Reservations, Flight ..." with a snippet "Airline tickets, airline reservations, flight airfare from United Airlines. Online reservation airline ticket purchase, electronic tickets, flight search, ..." and a link to "Show stock quote for UUA". At the bottom of the search results, there are several quicklinks: "Search options", "EasyCheck-in Online", "Mileage Plus", "My itineraries", "Baggage", "Services & information", "Itineraries & check-in", and "Planning & booking". At the very bottom, there is a link to "More results from united.com »".

Google united airlines Search

Web [+ Show options...](#)

[United Airlines Flights](#)  
[www.OneTravel.com/United-Airlines](http://www.OneTravel.com/United-Airlines) Save \$10 Instantly on **United Airlines** Airfares.

[United Airlines - Airline Tickets, Airline Reservations, Flight ...](#)  
Airline tickets, airline reservations, flight airfare from **United Airlines**. Online reservation  
airline ticket purchase, electronic tickets, flight search, ... [+ Show stock quote for UUA](#)  
[www.united.com/](http://www.united.com/) - [Cached](#) - [Similar](#) - [🗨](#) [📄](#) [✕](#)

[Search options](#) [Baggage](#)  
[EasyCheck-in Online](#) [Services & information](#)  
[Mileage Plus](#) [Itineraries & check-in](#)  
[My itineraries](#) [Planning & booking](#)

[More results from united.com »](#)

united airlines

Search Pad

SearchScan - On

102,000,000 results for united airlines:

Show All

United Air Lines

Wikipedia

Also try: [united airlines reservations](#), [united airlines flight](#), [More...](#)

[United Airlines - Airline Tickets, Airline Reservations ...](#) (Nasdaq: [UAUA](#))

Official site for **United Airlines**, commercial air carrier transporting people, property, and mail across the U.S. and worldwide.

[www.united.com](#) - 65k - [Cached](#)

[Planning & Booking](#)

[Shop for Flights](#)

[Itineraries & Check-in](#)

[Special Deals](#)

[Mileage Plus](#)

[Flight Status](#)

[Services & Information](#)

[Customer Service](#)

[more results from united.com »](#)

united airlines



## UNITED AIRLINES

[United Airline Fleet](#)

[United Airline Schedule](#)

[United Airlines Reservations](#)

[United Airline Jobs](#)

[Reference](#)

ALL RESULTS

[Cheap Flight Tickets](#) · [www.CheapOair.com](#)

CheapOair - The Only Way to Go!! Find Over 18 Million Exclusive Fares.

[Fly United Airlines](#) · [www.OneTravel.com/United-Airline](#)

Save \$10 Instantly on **United Airlines** Flights. Book Now, Hurry!

Best match

[United Airlines - Airline Tickets, Airline Reservations, Flight ...](#)

[www.united.com](#) · Official site

**Airline** tickets, **airline** reservations, flight airfare from **United Airlines**. Online reservations, **airline** ticket purchase, electronic tickets, flight search, fares and availability ...

[Flights](#)

[Redeem miles](#)

[Check In Online](#)

[Children, pets, & assistance](#)

[My itineraries](#)

[Change your travel plans](#)

[Baggage](#)

[Special deals](#)

Customer service 800-864-8331

## RELATED SEARCHES

[United Airlines Flight Status](#)

[US Airways](#)


[Continental Airlines](#)

# Εναλλακτικές αναπαραστάσεις;

YAHOO!®

Web Images Video Local Shopping News more ▾

uni Search

- united airlines** ▶ **UNITED AIRLINES - AIRLINE TICKETS,...**  
Airline tickets, airline reservations, flight airfare from United Airlines.  
Online reservations, ...  
 [www.united.com](http://www.united.com)
- univision
- university of phoenix
- asian unicorn
- universal studios
- united states postal service
- united healthcare

**MORE INFO**

<a href="#">Flights</a>	<a href="#">Check In Online</a>
<a href="#">Mileage Plus</a>	<a href="#">My Itineraries</a>
<a href="#">Baggage</a>	<a href="#">Redeem Miles</a>



ΤΕΛΟΣ (τμήματος) Κεφαλαίων 8, 11

Ερωτήσεις?

*Χρησιμοποιήθηκε κάποιο υλικό των:*

- ✓ *Pandu Nayak and Prabhakar Raghavan, CS276:Information Retrieval and Web Search (Stanford)*
- ✓ *Hinrich Schütze and Christina Lioma, Stuttgart IIR class*