

"Εφαρμογές Διαδικτύου"

Ο Παγκόσμιος Ιστός

WWW – World Wide Web

Το πρωτόκολλο HTTP

Εφαρμογές WWW

Ο Παγκόσμιος Ιστός

- γενικευμένος χώρος πληροφοριών
- υλοποιείται στο Διαδίκτυο
- δυνατότητα πρόσβασης από ένα μέρος του σε ένα άλλο

Ο Παγκόσμιος Ιστός

βασικές έννοιες

συνδυάζει τέσσερις βασικές έννοιες:

1. το μοντέλο Πελάτη-Εξυπηρετητή (Client-Server)

Ο Παγκόσμιος Ιστός

βασικές έννοιες

2. το υπερ-κείμενο (hypertext)

- είδος μορφοποίησης της πληροφορίας
- επιτρέπει τη μετακίνηση
 - από ένα μέρος ενός αρχείου σε ένα άλλο
 - από ένα αρχείο σε ένα άλλομέσω των υπερ-συνδέσμων (hyperlinks)

Ο Παγκόσμιος Ιστός

βασικές έννοιες

3. τα Αναγνωριστικά Πόρων (Resource Identifiers)

- μοναδικά αναγνωριστικά
- χρησιμοποιούνται για τον εντοπισμό ενός συγκεκριμένου πόρου στο δίκτυο

Ο Παγκόσμιος Ιστός

βασικές έννοιες

4. τη γλώσσα σήμανσης (markup language)

- πρόκειται για χαρακτήρες και κωδικούς
- παρεμβάλλονται μέσα σε ένα κείμενο
- σηματοδοτούν τη δομή, τη σημασία ή τον τρόπο παρουσίασης του κειμένου

Ο Παγκόσμιος Ιστός

βασικές έννοιες

Ένα πρόγραμμα-πελάτης (φυλλομετρητής ιστού) ανακτά πόρους από εξυπηρετητές ιστού χρησιμοποιώντας ενιαίους εντοπιστές πόρων (URL) και τους εμφανίζει.

Ακολουθώντας υπερ-συνδέσμους μιας ιστοσελίδας, μπορεί κανείς να περιηγηθεί στον Παγκόσμιο Ιστό, να αποστείλει πληροφορίες πίσω στον εξυπηρετητή.

- σερφάρισμα στον Ιστό

Ο Παγκόσμιος Ιστός Λειτουργία

Ο χρήστης δίνει στον φυλλομετρητή του τη **διεύθυνση URL** της σελίδας ή απλά ακολουθεί έναν **υπερ-σύνδεσμο** που δείχνει στη σελίδα αυτή.

αναζήτηση σε έναν εξυπηρετητή DNS της διεύθυνσης
IP του εξυπηρετητή που αναφέρεται στη διεύθυνση URL.

αποστολή ενός αιτήματος HTTP προς τον εξυπηρετητή ιστού
στη διεύθυνση IP που βρέθηκε προηγουμένως

ζητούνται το κείμενο HTML, οι εικόνες και άλλα αρχεία της ιστοσελίδας

ο εξυπηρετητής τα αποστέλλει στον φυλλομετρητή του χρήστη.

ο φυλλομετρητής αρχίζει να απεικονίζει την ιστοσελίδα
σύμφωνα με τις οδηγίες που περιγράφονται στα αρχεία HTML, CSS, ...

Ο Παγκόσμιος Ιστός Λειτουργία

Οι περισσότερες ιστοσελίδες εμπεριέχουν
υπερ-συνδέσμους

- προς άλλες σχετικές ιστοσελίδες
- προς αρχεία για κατέβασμα, αρχεία δεδομένων κ.ά.

Μία τέτοια **συλλογή** σχετικών μεταξύ τους **πόρων**
που διασυνδέονται μέσω υπερ-συνδέσμων,
σχηματίζει έναν **ιστό πληροφοριών**.

Καθιστώντας τον ιστό αυτό διαθέσιμο στο **Διαδίκτυο**,
δημιουργήθηκε αυτό που ο *Tim Berners-Lee* αποκάλεσε
Παγκόσμιο Ιστό στις αρχές της δεκαετίας του 1990.

Ο Παγκόσμιος Ιστός Λειτουργία

Ο **πυρήνας του Ιστού** αποτελείται από

- το Ενιαίο Αναγνωριστικό Πόρων
(Uniform Resource Identifier - **URI**)
παγκόσμιο σύστημα αναφοράς σε πόρους του Ιστού
- το πρωτόκολλο **HTTP**
(HyperText Transfer Protocol)
καθορίζει την επικοινωνία
μεταξύ φυλλομετρητή και εξυπηρετητή ιστού
- τη γλώσσα σήμανσης **HTML**
(HyperText Markup Language)
χρησιμοποιείται για τον καθορισμό της δομής,
και του περιεχομένου των αρχείων υπερ-κειμένου

Ο Παγκόσμιος Ιστός

τεχνολογίες

Java

- αντικειμενοστραφής γλώσσα προγραμματισμού της Sun Microsystems
- ενσωμάτωση applets στις ιστοσελίδες
 - εκτελούνται στον πελάτη
 - δεν επιβαρύνουν τον εξυπηρετητή
- τελικά επικράτησε για ανάπτυξη εφαρμογών στους εξυπηρετητές...

Ο Παγκόσμιος Ιστός τεχνολογίες

JavaScript

- καμία σχέση με τη Java...
- αντικειμενοστραφής γλώσσα σεναρίων της Netscape
- σχεδιάστηκε για χρήση σε ιστοσελίδες
- client-side
- είναι ιδιαίτερα αποτελεσματική

Ο Παγκόσμιος Ιστός τεχνολογίες

Ajax

- **A**synchronous **J**avaScript **A**nd **X**ML
- βασίζεται στη JavaScript
- επιτρέπει την πολύ πιο γρήγορη και αποτελεσματικότερη ανανέωση των ιστοσελίδων
- Gmail
- Google Maps

Ο Παγκόσμιος Ιστός

τεχνολογίες

PHP

- Personal Home Page/Form Interpreter
- PHP: Hypertext Preprocessor
- αντικειμενοστραφής γλώσσα σεναρίων
- σχεδιασμένη για χρήση σε ιστοσελίδες
- server-side
- ιδιαίτερα δημοφιλής

Ο Παγκόσμιος Ιστός

URI και URL

- Ενιαίο Αναγνωριστικό Πόρων
Uniform Resource Identifier – **URI**
 - μία ακολουθία χαρακτήρων για την αναγνώριση ενός πόρου
- Ενιαίος Εντοπιστής Πόρων
Uniform Resource Locator – **URL**
 - ένα είδος URI που επιπλέον καθορίζει και τον τρόπο εντοπισμού.

Ο Παγκόσμιος Ιστός

URL

- αναφέρεται συνήθως ως διεύθυνση URL
- χαρακτηρίζεται από τη μέθοδό της (**scheme**)
 - **http** πόροι HTTP
 - **https** πόροι HTTP με χρήση TLS/SSL
 - **file** πόροι διαθέσιμοι στο σύστημα αρχείων του τοπικού Η/Υ
 - **ftp** File Transfer Protocol
 - **mailto** διεύθυνση e-mail
 - ...

π.χ.

http://alpha.physics.uoi.gr/internet_applications

Ο Παγκόσμιος Ιστός

URL

■ γενική σύνταξη

scheme://authority/path?query#fragment

- **authority**
όνομα ή διεύθυνση IP ενός εξυπηρετητή
 - μπορεί να περιλαμβάνει
 - θύρα TCP
 - username, password
- **path**
καθορίζει τη θέση του πόρου με μία ιεραρχημένη δομή
- **query**
χρησιμοποιείται για το πέρασμα παραμέτρων
- **fragment**
σηματοδοτεί ένα συγκεκριμένο μέρος ενός πόρου

Ο Παγκόσμιος Ιστός

URL

■ απόλυτες διευθύνσεις URL

<http://en.wikipedia.org/w/wiki.phtml?title=Train&action=history>

http://en.wikipedia.org/wiki/Train#Freight_trains

http://alpha.physics.uoi.gr/internet_applications

■ σχετικές διευθύνσεις URL

<//nl.wikipedia.org/wiki/Train>

</wiki/Train>

Train\#Passenger_trains

Ο Παγκόσμιος Ιστός

φυλλομετρητές Ιστού

- πρόκειται για εφαρμογές πελάτες
- εντοπίζουν ιστοσελίδες με βάση τις διευθύνσεις URL
- εμφανίζουν τις ιστοσελίδες
 - μορφοποιημένο κείμενο
 - εικόνες
 - σύνδεσμοι
- είναι διαδραστικά
 - επιτρέπουν στο χρήστη να περιηγηθεί από σύνδεσμο σε σύνδεσμο

Ο Παγκόσμιος Ιστός

φυλλομετρητές Ιστού

- Οι πιο δημοφιλείς είναι
 - Mozilla Firefox
 - Opera
 - Netscape
 - Konqueror
 - Internet Explorer

Ο Παγκόσμιος Ιστός φυλλομετρητές Ιστού

- Υποστήριξη πρωτοκόλλων/τεχνολογιών
 - HTTP, HTTPS και FTP
 - HTML, XML, XHTML
 - εικόνες GIF, PNG, JPEG, SVG
 - Φύλλα Επάλληλων Στυλ
Cascading Style Sheets - CSS
 - JavaScript (Dynamic HTML) και XMLHttpRequest
 - Cookies
 - Ψηφιακά πιστοποιητικά
 - Favicons
 - RSS, Atom

Ο Παγκόσμιος Ιστός φυλλομετρητές Ιστού

■ Χαρακτηριστικά ευκολίας χρήσης

- Αυτόματη συμπλήρωση σε διευθύνσεις URL και φόρμες
- Χωρική πλοήγηση με το πλήκτρο Tab (Tabbed browsing)
- Χωρική πλοήγηση με τα πλήκτρα κατεύθυνσης (Spatial, caret navigation)
- Ανάγνωση περιεχομένων οθόνης, φωνητικός έλεγχος

Ο Παγκόσμιος Ιστός φυλλομετρητές Ιστού

- Αποφυγή ενοχλήσεων
 - Μπλοκάρισμα διαφημιστικών παραθύρων pop-up
 - Φιλτράρισμα διαφημιστικών
 - Προστασία από το "Phishing"

Ο Παγκόσμιος Ιστός

εξυπηρετητές Ιστού

- δέχονται τις αιτήσεις HTTP (εξ ορισμού στη θύρα TCP 80)
- αποστέλλουν τα κατάλληλα δεδομένα
- πιο δημοφιλής είναι ο **Apache** (www.apache.org)
 - διατίθεται σε όλα τα σημερινά λειτουργικά συστήματα
 - συνεργάζεται και με άλλα πακέτα λογισμικού
 - διαχειστής βάσεων δεδομένων MySQL
 - γλώσσα σεναρίων PHP
- τα παραπάνω πακέτα λογισμικού διατίθενται δωρεάν
 - άδεια χρήσης ελεύθερου κώδικα **GNU GPL** (www.gnu.org)
- μπορεί κανείς εύκολα να εγκαταστήσει το δικό του εξυπηρετητή Παγκόσμιου Ιστού!

πρωτόκολλο HTTP

HTTP – HyperText Transfer Protocol

- μεταφορά των πληροφοριών στον Παγκόσμιο Ιστό
- αναπτύσσεται υπό την επίβλεψη του World Wide Web Consortium (W3C)
- είναι ένα πρωτόκολλο αίτησης/απάντησης μεταξύ πελατών και εξυπηρετητών
 - είναι δυνατό να παρεμβάλλονται διάφοροι ενδιάμεσοι όπως proxies, gateways και tunnels.

πρωτόκολλο HTTP

- ένας πελάτης HTTP ξεκινά μία αίτηση, αποκαθιστώντας μία σύνδεση TCP σε έναν απομακρυσμένο υπολογιστή
- ένας εξυπηρετητής HTTP που *ακούει* στη θύρα αυτή περιμένει από κάποιον πελάτη ένα μήνυμα αιτήσεως
- Λαμβάνοντας το αίτημα, ο εξυπηρετητής επιστρέφει στον πελάτη μία γραμμή κατάστασης όπως π.χ. "HTTP/1.1 200 OK" και ένα δικό του μήνυμα που μπορεί να είναι το αρχείο που ζητήθηκε, κάποιο μήνυμα λάθους ή κάποια άλλη πληροφορία.

πρωτόκολλο HTTP

Μήνυμα αίτησης HTTP

- αποτελείται από τα ακόλουθα μέρη
 - Τη γραμμή αίτησης, όπως π.χ.
"GET /images/logo.gif HTTP/1.1"
η οποία ζητά το αρχείο με όνομα **logo.gif**
από τον κατάλογο **/images**
 - Κεφαλίδες, όπως π.χ. "Accept-Language: en"
 - Μία κενή γραμμή
 - Προαιρετικά, το σώμα του μηνύματος

πρωτόκολλο HTTP

Μέθοδοι αίτησης

- το HTTP ορίζει οκτώ μεθόδους αίτησης
 - σηματοδοτούν την επιθυμητή δράση που θα πρέπει να εκτελεστεί σε κάποιον πόρο
1. **HEAD**
Ζητά την ίδια απάντηση με μία αντίστοιχη αίτηση GET, χωρίς όμως το κυρίως σώμα.
 2. **GET**
Ζητά την αποστολή ενός συγκεκριμένου πόρου.
 3. **POST**
Υποβάλλει δεδομένα που παρέχει ο χρήστης προς κάποιον αναγνωρισμένο πόρο.
Τα δεδομένα συμπεριλαμβάνονται στο σώμα του αιτήματος.

πρωτόκολλο HTTP

4. **PUT**
Ανεβάζει στον εξυπηρετητή κάποιον συγκεκριμένο πόρο.
5. **DELETE**
Διαγράφει κάποιον συγκεκριμένο πόρο.
Η μέθοδος αυτή σπανίως υλοποιείται.
6. **TRACE**
Αναμεταδίδει πίσω την αίτηση, ώστε ο πελάτης να δει ποιοι ενδιάμεσοι εξυπηρετητές προσθέτουν ή μεταβάλλουν πληροφορίες της αίτησης.
7. **OPTIONS**
Επιστρέφει τις μεθόδους HTTP τις οποίες υποστηρίζει ο εξυπηρετητής.
Χρησιμοποιείται για τον έλεγχο λειτουργικότητας του εξυπηρετητή.
8. **CONNECT**
Η μέθοδος αυτή χρησιμοποιείται σε συνδυασμό με κάποιο proxy το οποίο μπορεί να μεταβληθεί σε κάποιο τούνελ SSL

πρωτόκολλο HTTP

- Οι μέθοδοι **GET** και **HEAD** χαρακτηρίζονται ασφαλείς, αφού έχουν σκοπό μόνο την ανάκτηση πληροφοριών
- Μη ασφαλείς μέθοδοι (όπως οι **POST**, **PUT** και **DELETE**) πρέπει να παρουσιάζονται στο χρήστη με ειδικό τρόπο, ώστε να προειδοποιείται για ζητήματα ασφάλειας
- Οι εξυπηρετητές HTTP οφείλουν να υλοποιούν τουλάχιστον τις μεθόδους GET και HEAD

πρωτόκολλο HTTPS

- HTTP + TLS/SSL
- αναπτύχθηκε από τη Netscape
- παρέχει έλεγχο αυθεντικότητας και κρυπτογραφημένη επικοινωνία
- είναι ευρέως διαδεδομένο στον Παγκόσμιο Ιστό για την ανταλλαγή ευαίσθητων δεδομένων (π.χ. σε τραπεζικές συναλλαγές).
- χρησιμοποιεί εξ ορισμού τη θύρα TCP 443

εφαρμογές WWW

προσωπικές ιστοσελίδες - ιστοχώροι

- ιστοσελίδα
 - πηγή πληροφοριών που βρίσκεται στον Παγκόσμιο Ιστό
 - μπορεί να εμφανιστεί μέσω ενός φυλλομετρητή
 - η πληροφορία κωδικοποιείται μέσω της γλώσσας HTML
- 2005: ~10 δις ιστοσελίδες
- ιστοχώροι (websites)
 - σύνολα ιστοσελίδων με σχετικό περιεχόμενο
π.χ. ο ιστοχώρος του Τμήματος Φυσικής
<http://www.physics.uoi.gr>

εφαρμογές WWW

αναζήτηση πληροφοριών

- μηχανές αναζήτησης
τεράστιες βάσεις δεδομένων και εξειδικευμένο λογισμικό
αναζήτησης λέξεων-κλειδίων σε αυτές
- ανανεώνονται συνεχώς
ειδικά προγράμματα (spiders, crawlers, robots κ.ά.)
διατρέχουν συνεχώς τον Ιστό συλλέγοντας
πληροφορίες για όλες τις ιστοσελίδες
- Τα αποτελέσματά τους ταξινομούνται
με βάση το βαθμό συσχετισμού τους με τις λέξεις κλειδιά
και το πόσο δημοφιλείς είναι οι ιστοσελίδες-αποτελέσματα

εφαρμογές WWW

αναζήτηση πληροφοριών

- Η μεγαλύτερη και δημοφιλέστερη μηχανή αναζήτησης είναι της εταιρίας **Google** (www.google.com)
- Για την ταξινόμηση των αποτελεσμάτων του, το Google χρησιμοποιεί πάνω από **150** κριτήρια!
- μετα-μηχανές αναζήτησης (π.χ. www.metacrawler.com)
 - δεν διατηρούν δικές τους βάσεις δεδομένων
 - αναζητούν τις λέξεις-κλειδιά σε πολλές μηχανές αναζήτησης και επιστρέφουν όλα τα αποτελέσματα

εφαρμογές WWW

ηλεκτρονική διακυβέρνηση

- Δημόσιοι φορείς εξυπηρετούν τους πολίτες μέσω εξειδικευμένων ιστοχώρων
- Μετά από ταυτοποίηση του χρήστη από τον εξυπηρετητή, είναι δυνατή η πραγματοποίηση περίπλοκων διαδικασιών, π.χ. υποβολή φορολογικής δήλωσης μέσω του www.gsis.gr
- Με την ανάπτυξη τέτοιου είδους ιστοχώρων, επιτυγχάνεται η λεγόμενη ηλεκτρονική διακυβέρνηση.
- Τα πλεονεκτήματα είναι πολλά και άμεσα. Μπορεί κανείς να διεκπεραιώσει τις υποθέσεις του με το Δημόσιο εξοικονομώντας κόπο και χρόνο

εφαρμογές WWW

ηλεκτρονικό εμπόριο

- σχεδόν αμέσως με την εξάπλωση του Διαδικτύου, αναπτύχθηκε το ηλεκτρονικό εμπόριο (e-commerce)
- Σήμερα, ο αριθμός ηλεκτρονικών καταστημάτων είναι τεράστιος. Τα καταστήματα αυτά εξυπηρετούν πάρα πολλές ανάγκες (βιβλία, μουσική, ταινίες, τρόφιμα, ηλεκτρικά είδη, ...)
- Στην κατηγορία αυτή θα μπορούσε κανείς να εντάξει και τις τραπεζικές συναλλαγές μέσω Διαδικτύου (internet banking).
- Σε όλες τις παραπάνω περιπτώσεις, είναι ιδιαίτερα ευαίσθητο το ζήτημα της ασφάλειας των συναλλαγών
- το πρωτόκολλο που χρησιμοποιείται είναι το HTTPS το οποίο παρέχει την απαιτούμενη ασφάλεια

εφαρμογές WWW

πύλες (portals)

- Πρόκειται για ιστοσελίδες ποικίλου περιεχομένου.
- Συνήθως κατασκευάζονται από δημοσιογραφικούς οργανισμούς.
- Παραδείγματα ελληνικών πυλών
 - www.ert.gr
 - www.in.gr
 - www.flash.gr
- συνήθως παρέχουν ειδήσεις της επικαιρότητας, ενώ γύρω υπάρχουν διάφοροι σύνδεσμοι για ποικίλα θέματα (καιρός, διασκέδαση, σπίτι, αυτοκίνητο, χρηματιστήριο κ.ά.).
- Οι πύλες περιλαμβάνουν περιεχόμενο που ανανεώνεται συνεχώς, όπως άλλωστε και οι ειδήσεις

εφαρμογές WWW

τηλεδιασκέψεις - VoIP

- Η διάδοση του ευρυζωνικού διαδικτύου επιτρέπει τη χρήση του για τη μετάδοση εικόνας και ήχου που ψηφιοποιούνται
- Εφαρμογές: VRVS (www.vrvs.org), EVO (evo.caltech.edu)
- Πέρα από τα πρωτόκολλα HTTP και HTTPS, χρησιμοποιούν και άλλα πρωτόκολλα (π.χ. το UDP) και τεχνολογίες (π.χ. τη Java) για το χειρισμό των απαιτούμενων συσκευών
- συνεργασία ομάδων που δεν βρίσκονται στον ίδιο χώρο.
- Ιδιαίτερη είναι η εφαρμογή τους στο χώρο της Φυσικής Υψηλών Ενεργειών,
 - πειράματα τεραστίων διαστάσεων
 - κατασκευάζονται από ομάδες που εργάζονται παράλληλα σε όλο τον κόσμο

εφαρμογές WWW πολυμέσα

- χρησιμοποιώντας βοηθητικά προγράμματα (plugins) των φυλλομετρητών είναι δυνατή η παρακολούθηση στον υπολογιστή τηλεοπτικών και ραδιοφωνικών προγραμμάτων που αναμεταδίδονται στο Διαδίκτυο.
- πολλοί σταθμοί που παρέχουν τη δυνατότητα αυτή π.χ. www.ert.gr, www.redfm.gr, ...
- αναπτύσσονται νέες υπηρεσίες που παρέχουν video π.χ. video.google.com