

# Γλώσσες Σήμανσης (Markup Languages)



Τεχνολογία Διαδικτύου και Ηλεκτρονικό Εμπόριο

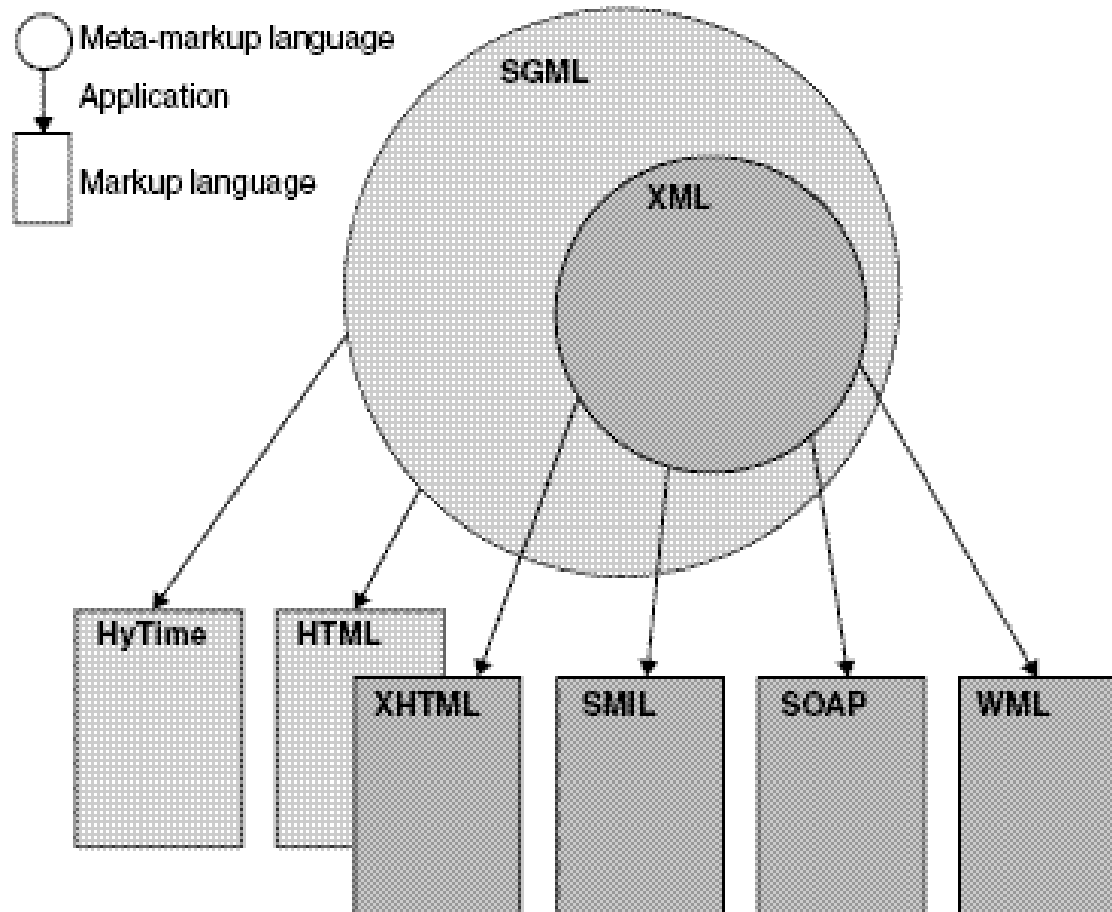
# Γλώσσες Σήμανσης

- Γλώσσες σήμανσης:
  - Αρχικά για τον καθορισμό εμφάνισης σελίδων, γραμματοσειρών.
  - Στη συνέχεια επεκτάθηκαν στην κάλυψη διάφορων ακολουθιών ελέγχου (μορφοποίηση, εκτύπωση, κ.α.)
  - Χρήση στην ψηφιακή τυπογραφία
  - Παραδείγματα: LaTeX, troff, SGML (Standard Generalized Markup Language)

# Γλώσσες και μετα-γλώσσες

- Η SGML είναι γλώσσα για τον καθορισμό συγκεκριμένων γλωσσών σήμανσης (μετα-γλώσσα).
- Η HTML (HyperText Markup Language) είναι εφαρμογή της SGML.
- Η XML (eXtensible Markup Language) είναι υποσύνολο της SGML, δηλαδή μετα-γλώσσα.
- Εφαρμογές της XML είναι π.χ. XHTML, SOAP, WML, κ.α.

# SGML, XML και εφαρμογές



# Τμήματα SGML εφαρμογής

- SGML declaration: καθορίζει νόμιμους χαρακτήρες και διαχωριστές
- Document Type Definition (DTD): καθορίζει νόμιμες κατασκευές σήμανσης
- Προδιαγραφή ανάθεσης σημασιολογίας στη σήμανση.
- Στιγμιότυπα αντικειμένων με περιεχόμενο και σήμανση.

# SGML declaration

- Ετοιμάζει την υποδομή για την κατανόηση των DTD.
- Καθορίζει σύνολο χαρακτήρων, διαχωριστές, περιορισμούς για το τι μπορεί να περιέχεται στο DTD.

# Document Type Definition

- Καθορίζει το συντακτικό των κατασκευών σήμανσης
  - Ορισμός εννοιών (entities)
  - Ορισμός στοιχείων (elements)
  - Προσδιορισμός ιδιοτήτων (attributes)

# HTML

- Ευέλικτη προδιαγραφή
  - π.χ. πολλά στοιχεία έχουν προαιρετικές ετικέτες κλεισίματος
  - πολλά έγγραφα HTML παραβιάζουν ακόμη και αυτούς τους εύκαμπτους περιορισμούς
- Η σημασία των στοιχείων καθορίζεται μέσα από κείμενο
- Τελευταία έκδοση: 4.01 Δεκέμβριος '99
- Ουσιαστικό πρόβλημα: συμβατότητα με υπάρχοντα HTML έγγραφα.



# HTML: Δομή και συντακτικό

- Αναφορά στην έκδοση HTML που υλοποιεί.
- Επικεφαλίδα HTML
- Σώμα HTML

# Παράδειγμα HTML

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN"
  "http://www.w3.org/TR/html4/strict.dtd">
<HTML>
  <HEAD>
    <TITLE>Sample HTML Document</TITLE>
  </HEAD>
  <BODY>
    <P>I don't have to close the <P> tag.
  </BODY>
</HTML>
```

# Επικεφαλίδα HTML

- Περιλαμβάνει δηλώσεις για ολόκληρο το έγγραφο.
- Παραδείγματα από στοιχεία
  - <HEAD>
    - <TITLE>
    - <META>
    - <STYLE>
    - <SCRIPT>

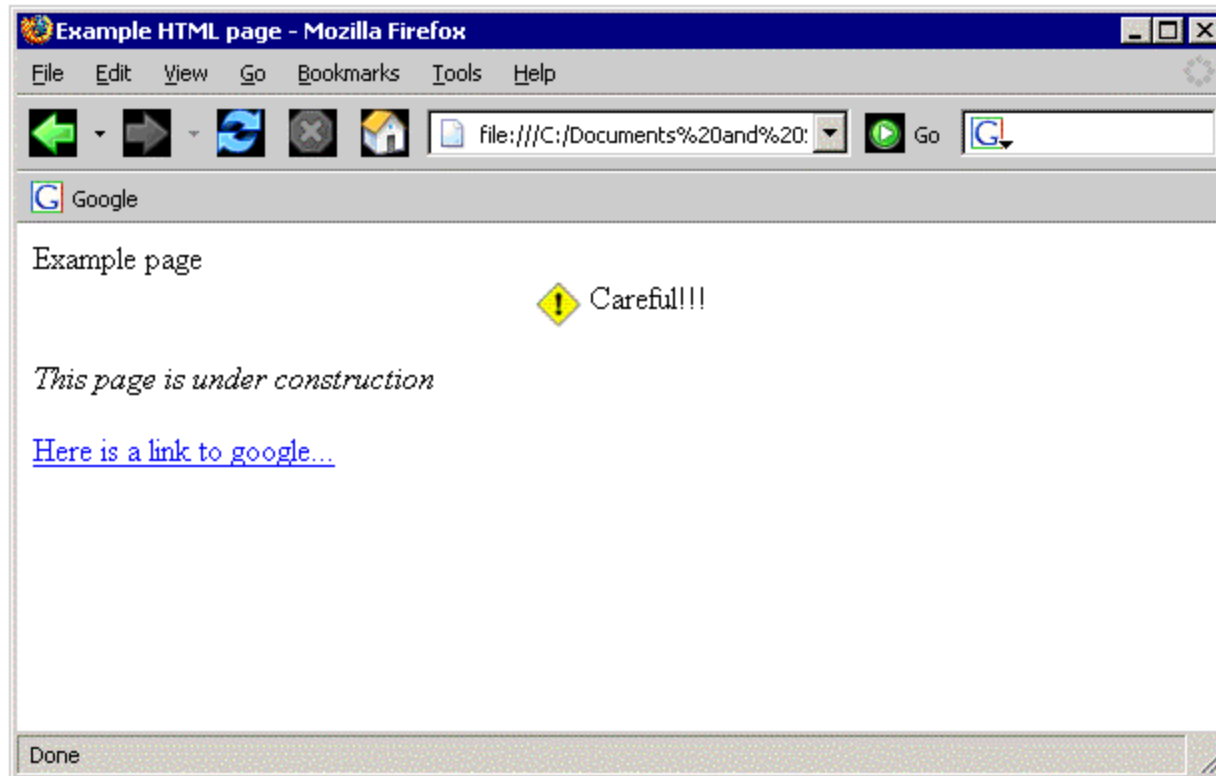
# Σώμα HTML

- Περιλαμβάνεται στο στοιχείο <BODY>
- Παραδείγματα από στοιχεία:
  - Επικεφαλίδες (<H1>, <H2>, κτλ)
  - <TABLE>
  - <OL>, <UL>
    - <LI>
  - <a>, <img>

# Παράδειγμα HTML

```
view-source: - Source of: file:///C:/Documents%20and%20Settings/sarantis/Desktop/odt/example/ex...
File Edit View
<html>
  <title>Example HTML page</title>
  <body bgcolor="#ffffff">
    Example page
    <center>
      
        Careful!!!
      </img>
      <p></p>
    </center>
    <p>
      <em>This page is under construction</em>
    <p>
      <a href="http://www.google.com/">Here is a link to google...</a>
  </body>
</html>
```

# Εμφάνιση στο browser



# Τροποποίηση εμφάνισης

- Cascading Style Sheets (CSS)
- Javascript
- DHTML

# Cascading Style Sheets

- Μηχανισμός ελέγχου εμφάνισης HTML.
- Αποτελείται από κανόνες που εφαρμόζονται σε ένα στοιχείο HTML και προδιαγράφουν κάποια ιδιότητα εμφάνισης.
- Ιδιότητες που καθορίζονται από κανόνες CSS:
  - γραμματοσειρών, χρώματος, κειμένου



# Χρήση CSS

- Εσωτερικά στο HTML έγγραφο:
  - Με χρήση του στοιχείου <STYLE>
- Με ανεξάρτητο αρχείο
  - Με χρήση του στοιχείου <LINK>

# Javascript

- Οι γλώσσες σήμανσης είναι ικανοποιητικές όταν το περιεχόμενο είναι στατικό (δεν αλλάζει).
- Ένα πρόβλημα είναι ότι η παρουσίαση δεν αλλάζει όταν έχει τελειώσει η επεξεργασία.
- Χρειαζόταν κάτι να παρέχει προγραμματιστική λειτουργικότητα στον browser.
- Λύση: Εισαγωγή γλώσσας προγραμματισμού με πρόσβαση στα αντικείμενα του browser.

# Javascript

- Αρχικά αναπτύχθηκε για Netscape Navigator, στη συνέχεια υποστηρίζεται από όλους τους browsers.
- Υπάρχουν κάποιες διαφορές στην υλοποίηση.
- Δεν επεξεργάζονται όλες οι εντολές Javascript την ίδια ώρα.
  - Πριν ή κατά τη διάρκεια επεξεργασίας εμφάνισης.
  - Κατά την ενεργοποίηση κάποιου συμβάντος.

# Dynamic HTML (DHTML)

- Συνδυασμός HTML + Javascript + CSS.
- Προσφέρει ένα βαθμό διαδραστικού ελέγχου στην παρουσίαση της σελίδας.
- Η παρουσίαση αλλάζει μέσω ανάδρασης με το χρήστη.
- Διαφορά με δυναμικές web εφαρμογές:
  - τοπική λειτουργία αντί για δημιουργία νέας σελίδας.

# Παραδείγματα DHTML

- Συμπεριφορές mouse-over
- Επικύρωση πεδίων φόρμας
- Τεχνικές layering

# XML: Παρελθόν και μέλλον

- XML: ένα απλούστερο υποσύνολο της SGML.
  - XML DTDs υποσύνολα των SGML DTDs.
- Αναπτύχθηκαν και άλλοι μηχανισμοί.
- XML Schema: ορισμός περιορισμών εφαρμογής
  - XPath: για χειρισμό XML εγγράφων
  - XSLT: για μετασχηματισμό εγγράφων

# Διαφορές XML-SGML

- Υπάρχει προκαθορισμένο σύνολο χαρακτήρων και διαχωριστών – δεν απαιτείται δήλωση τους.
- Τα απαιτούμενα DTDs είναι απλούστερα.
- Το XML Schema αντικαθιστά το DTD (ορίζεται το ίδιο ως XML εφαρμογή)
- Εισαγωγή του χώρου ονομάτων (namespace), για διαχωρισμό σημασίας μεταξύ διαφορετικών εφαρμογών.

# Έγγραφα XML

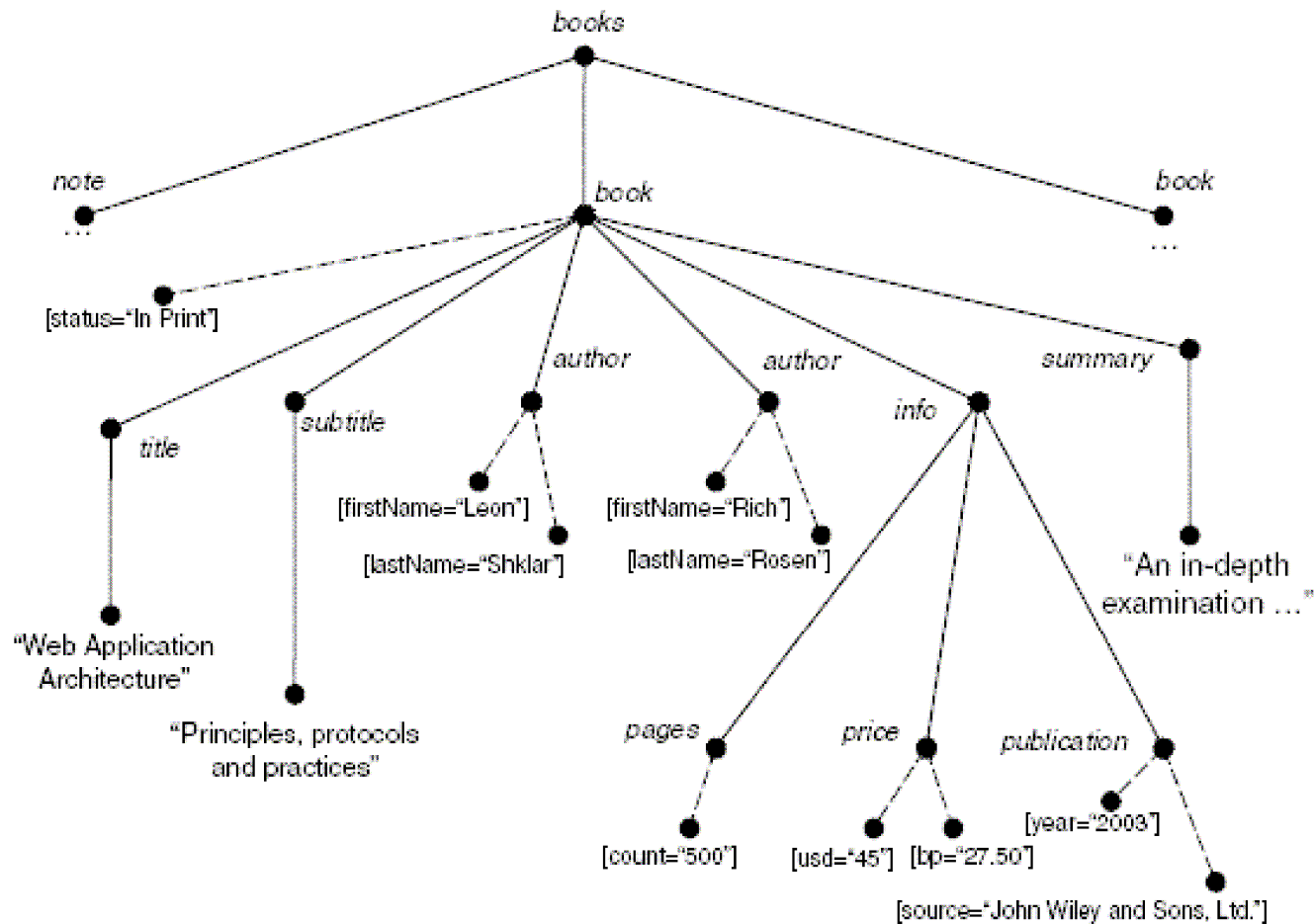
- Αποτελούνται από δηλώσεις, στοιχεία, ετικέτες, σχόλια, αναφορές χαρακτήρων, κτλ.
- Υπάρχει ένα μοναδικό στοιχείο ρίζας.
- Όλα τα άλλα στοιχεία πρέπει να είναι σωστά εμφωλευμένα το ένα μέσα στο άλλο.
- “well-formed” αν περιλαμβάνει δήλωση εγγράφου, ικανοποιεί την εμφώλευση στοιχείων και δεν περιλαμβάνει αδιευκρίνηστες οντότητες.
- “valid” αν ικανοποιεί τους περιορισμούς από DTD ή Schema.



# Παράδειγμα XML

```
<?xml version="1.0" standalone="yes"?>
<!-- XML example for the Web Architecture book -->
<!DOCTYPE books SYSTEM "books.dtd">
<books>
  <book status="In Print">
    <title>Web Application Architecture</title>
    <subtitle>Principles, protocols and practices</subtitle>
    <author firstName="Leon" lastName="Shklar"/>
    <author firstName="Rich" lastName="Rosen"/>
    <info>
      <pages count="500"/>
      <price usd="45" bp="27.50"/>
      <publication year="2003" source="ajw"/>
    </info>
    <summary>An in-depth examination of the basic concepts and general
    principles associated with Web application development.
    </summary>
  </book>
  <!--
  <book><TED>
  </book>
  -->
  <notes><![CDATA[
  &lt;!-- This is our CDATA example that hides syntactically incorrect
  comments and XML fragments -->
  <books><book>
  ]]>
  </notes>
</books>
```

# XML δένδρο στοιχείων



# XML DTDs

- Η σημασία των XML στοιχείων δεν είναι προκαθορισμένη.
  - π.χ. τι τιμές μπορεί να πάρει το στοιχείο <price>;
  - Πώς ανατίθεται νόημα στα στοιχεία XML για συγκεκριμένες εφαρμογές;
- Τα DTDs καθορίζουν τους περιορισμούς κατασκευής σήμανσης.
- Περιορισμένη χρησιμότητα για πολυπλοκότερους περιορισμούς.

# XML Schemas

- Εναλλακτικό του XML DTD.
- Βασίζεται σε XML.
- Περιλαμβάνει επιπλέον κατασκευές για πολυπλοκότερους περιορισμούς.
- Υποστηρίζει “complex” και “simple” τύπους.
  - Οι απλοί τύποι προέρχονται από τους ενσωματωμένους με επιπλέον περιορισμούς.
  - Οι πολύπλοκοι τύποι είναι συνδυασμοί άλλων απλών και πολύπλοκων τύπων.

# Προβλήματα HTML

- Η δομή HTML εγγράφων είναι χωρίς πολλούς περιορισμούς
  - π.χ. οι ετικέτες κλεισίματος είναι προαιρετικές
  - στην πραγματικότητα ακόμη και οι χαλαροί κανόνες καταστρατηγούνται από έγγραφα HTML.

# XHTML

- XHTML είναι η επαναδιατύπωση της HTML 4.01 ως XML εφαρμογής.
- Η μετάβαση στην XHTML επιβάλλει τήρηση αυστηρών κανόνων δομής, αλλά και εγκατάλειψη των προβλημάτων υποστήριξης κακών εγγράφων.
- Διαφορές με HTML
  - όλες οι ετικέτες πρέπει να κλείνουν.
  - οι ετικέτες είναι πεζές (όχι αυθαίρετα πεζά/κεφαλαία).
- Μερικές φορές έγκυρο HTML  $\neq$  έγκυρο XHTML

# Διαφορές HTML-ΧHTML

```
...  
<p>I have to close the &lt;p> tag.</p> <!-- HTML + XHTML -->  
<p>I have to close<br> the &lt;p> tag.</p> <!-- HTML -->  
<p>I have to close<br/> the &lt;p> tag.</p> <!-- XHTML -->  
<p>I have to close<br></br> the &lt;p> tag.</p> <!-- HTML + XHTML -->  
...
```

- Η HTML είναι αποσυρόμενη προδιαγραφή.
- Σταδιακά οι διαφορές μεταξύ HTML και XHTML θα διευρύνονται.

# WML – Wireless Markup Language

- WML = XML εφαρμογή
- Στοχεύει στη χρήση από συσκευές με μικρή επεξεργαστική ισχύ, μικρές δυνατότητες οθόνης, μικρό εύρος ζώνης (π.χ. κινητά τηλέφωνα).
- DTD ορίζεται από το WAP forum.
- Ένα έγγραφο = πολλές διαδοχικές μικρότερες οθόνες.



# XSLT

- eXtensible Stylesheet Language Transformation
- Βασίζεται στη γλώσσα XSL.
- Επιλέγονται τμήματα σε ένα XML έγγραφο και μετατρέπονται σε διαφορετικά XML τμήματα.