



ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ

ΤΕΧΝΟΛΟΓΙΕΣ ΔΙΑΧΕΙΡΙΣΗΣ ΠΟΛΙΤΙΣΜΙΚΗΣ ΠΛΗΡΟΦΟΡΙΑΣ

Εργαστήριο 2: XML

Γιώργος Καρυδάκης
Τμήμα Πολιτισμικής Τεχνολογίας και
Επικοινωνίας



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Άδειες Χρήσης

- Το παρόν εκπαιδευτικό υλικό υπόκειται σε άδειες χρήσης Creative Commons.
- Για εκπαιδευτικό υλικό, όπως εικόνες, που υπόκειται σε άλλου τύπου άδειας χρήσης, η άδεια χρήσης αναφέρεται ρητώς.



Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Αιγαίου**» 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.



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



Εργαστήριο 2: XML

Τεχνολογίες Διαχείρισης Πολιτισμικής
Πληροφορίας

Γ. Καρυδάκης

SGML (Standard Generalized Markup Language)

- Διεθνές πρότυπο ορισμού μεθόδων
 - αναπαράστασης πληροφοριών ανεξάρτητα από
 - συσκευή/σύστημα και χρήση/ανάγνωση
- Εφαρμογές SGML
 - HTML (όχι 5)
 - XML

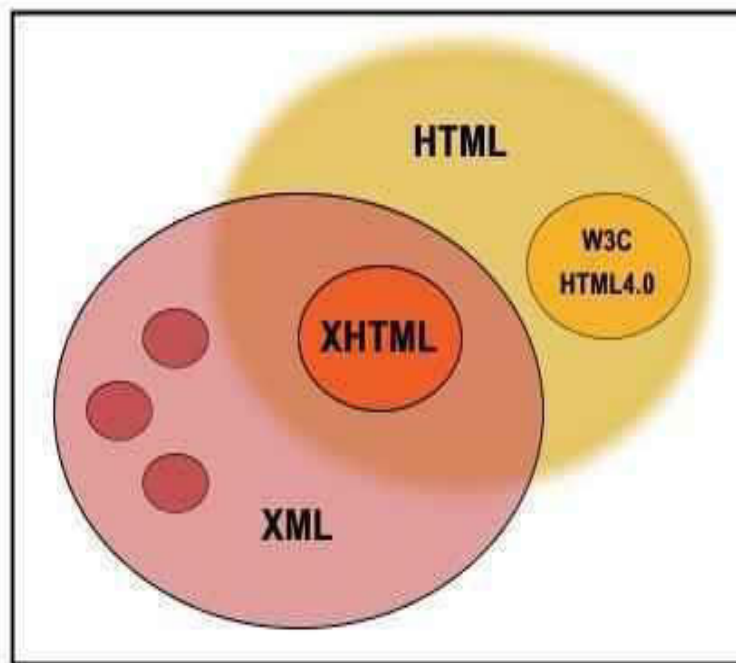
HTML

```
<html>
<head><title>Books</title></head>
<body>
<h2>Books</h2>
<hr>
<em>Sense and Sensibility</em>, <b>Jane Austen</b>,
1811<br>
<em>Pride and Prejudice</em>, <b>Jane Austen</b>,
1813<br>
<em>Alice in Wonderland</em>, <b>Lewis Carroll</b>,
1866<br>
<em>Through the Looking Glass</em>, <b>Lewis
Carroll</b>, 1872<br>
</body>
</html>
```

XML

```
<books>
  <book>
    <title>Sense and Sensibility</title>
    <author>Jane Austen</author>
    <year>1811</year>
  </book>
  <book>
    <title>Pride and Prejudice</title>
    <author>Jane Austen</author>
    <year>1813</year>
  </book>
  ...
</books>
```

HTML & XML



Ομοιότητες

- Είναι γλώσσες σήμανσης (markup)
- Χρησιμοποιούν ετικέτες (tags)
- Εύκολα αναγνώσιμες από άνθρωπο

Διαφορές

- Όλες οι ετικέτες στην XML πρέπει να κλείνουν
- Η HTML είναι δύσκολα αναγνώσιμη από μηχανές
- Η HTML δεν περιέχει δομική πληροφορία (οντότητες- σχέσεις)
- Η XML επιτρέπει ορισμό περιορισμών σε τιμές (π.χ. έτος)
- Οι ετικέτες στην HTML είναι προκαθορισμένες
- Οι ετικέτες στην HTML αφορούν μορφοποίηση
- Η XML είναι μια μέτα-γλώσσα σήμανσης
 - οι ετικέτες δεν είναι προκαθορισμένες
 - ορίζονται από τους γρήστες

Δήλωση XML (XML declaration)

- `<?xml version="1.0" [encoding="UFT-16"?]>`
- `<?xml version="1.0" [encoding="UFT-16"? standalone = "no"]>`
- Αναφορά σε εξωτερικά έγγραφα
 - `<!DOCTYPE book SYSTEM "book.dtd">`
 - οι δομικές πληροφορίες βρίσκονται στο τοπικό αρχείο `book.dtd`

Στοιχεία - Elements

- Αναφέρονται στις κύριες έννοιες ενός εγγράφου XML
- Στοιχείο `<tag-name>content</tag-name>`
 - π.χ. `<title>Big Lebowski</title>`
- Τα ονόματα ετικετών (tag-names) ορίζονται από τον χρήστη
- Το περιεχόμενο (content) ενός στοιχείου μπορεί να είναι
 - κείμενο
 - αριθμός

Χαρακτηριστικά - Attributes

- Ένα κενό στοιχείο μπορεί να περιέχει ορισμένες ιδιότητες με τη μορφή «χαρακτηριστικών»
- Χαρακτηριστικό
 - <όνομα>=<τιμή>
 - π.χ. <dancer hairstyle="funky">
- Elements vs Attributes
 - πολλαπλές τιμές
 - δενδρική δομή
 - επεκτάσιμα

Σχόλια

- <!-- Αυτό είναι ένα σχόλιο -->

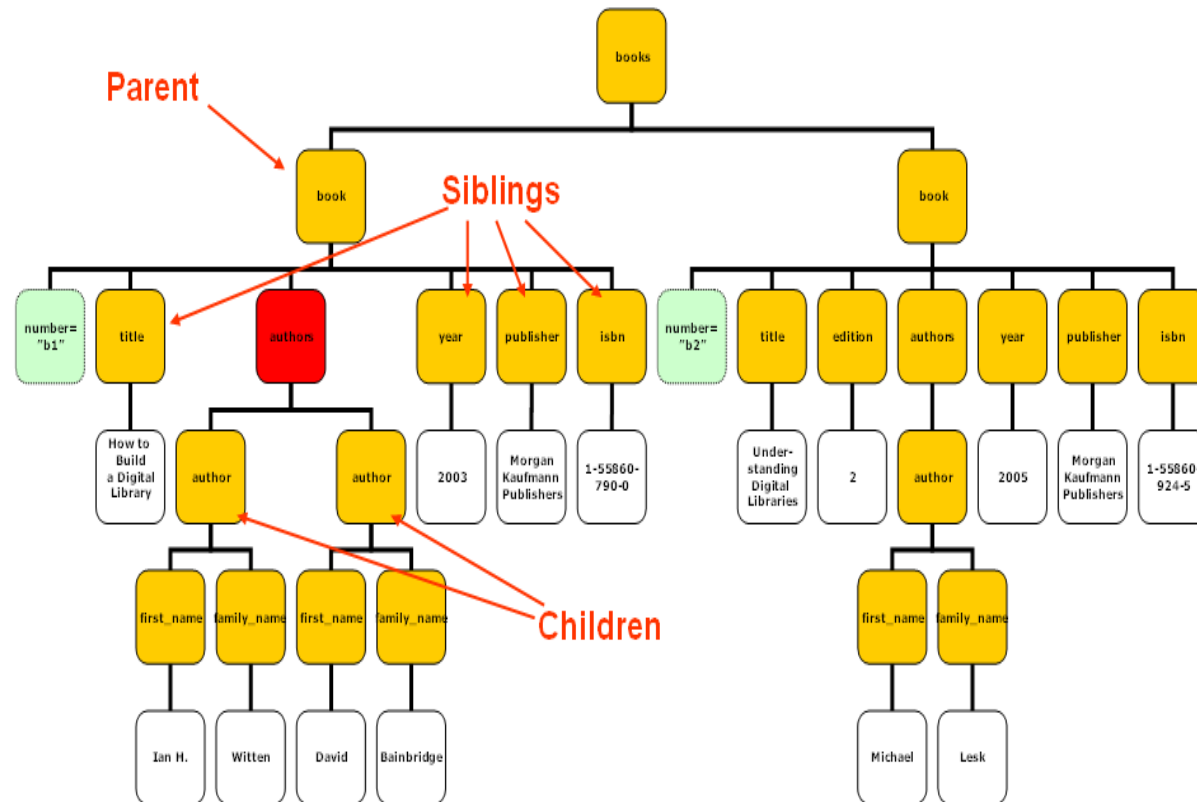
Κανόνες Δόμησης Εγγράφου XML

- Μόνο ένα εξωτερικό στοιχείο
 - στοιχείο ρίζα
- Κάθε στοιχείο περιέχει μια ετικέτα ανοίγματος και μια αντίστοιχη κλεισίματος
- Δεν υπάρχουν επικαλύψεις ετικετών
 - ~~<author><name>Lee Hong</author></name>~~
- Τα χαρακτηριστικά σ' ένα στοιχείο έχουν μοναδικά ονόματα
- Τα ονόματα των στοιχείων και των ετικετών πρέπει να είναι επιτρεπτά

Δενδρικό Μοντέλο

- Είναι εφικτό να αναπαριστούμε σωστά δομημένα έγγραφα XML σαν δένδρα
- Κανόνες:
 - Υπάρχει ακριβώς μια ρίζα
 - Δεν υπάρχουν κύκλοι
 - Κάθε κόμβος έχει ακριβώς ένα γονέα (πλην της ρίζας)
 - Κάθε κόμβος έχει μια ετικέτα
 - Η σειρά των στοιχείων είναι σημαντική

Δενδρικό Μοντέλο Δόμησης



Άσκηση

- Επιλέξτε ένα θέμα που σας ενδιαφέρει
- Σχεδιάστε στο χαρτί την δενδρική δομή της πληροφορίας
- Μετατρέψτε την δομή αυτή σε XML
- Δημιουργείστε πολλά στιγμιότυπα