

Ηλεκτρονική Δημοσίευση

Σπύρος Βερονίκης

Ηλεκ/γος Μηχ/κός & Μηχ/κός Η/Υ

Δρ. Τμήματος Αρχειονομίας - Βιβλιοθηκονομίας - Μουσειολογίας Ι.Π.

spver@ionio.gr

Οκτώβριος 2020

- 1 Σχετικά με το σημερινό μάθημα (Εβδ.2η)
 - Επισκόπηση θεμάτων
 - Πρότυπα ηλεκτρονικής δημοσίευσης

- 2 Δημοσίευση για τον παγκόσμιο ιστό (Online Publication)
 - Κίνητρα δημοσίευσης στο WWW
 - Πρότυπα γλωσσών επισημείωσης (Markup Language Standards)
 - Το πρότυπο HTML
 - Το πρότυπο XML
 - Μιλώντας για σημασιολογία

Περιεχόμενα

- 1** Σχετικά με το σημερινό μάθημα (Εβδ.2η)
 - Επισκόπηση θεμάτων
 - Πρότυπα ηλεκτρονικής δημοσίευσης

- 2** Δημοσίευση για τον παγκόσμιο ιστό (Online Publication)
 - Κίνητρα δημοσίευσης στο WWW
 - Πρότυπα γλωσσών επισημείωσης (Markup Language Standards)
 - Το πρότυπο HTML
 - Το πρότυπο XML
 - Μιλώντας για σημασιολογία

Επισκόπηση θεμάτων

- E1: Εισαγωγή. Το νέο περιβάλλον δημοσίευσης. Ευκαιρίες & δυνατότητες. Χαρακτηριστικά γνωρίσματα ψηφιακών τεκμηρίων.
- E2: Θέματα δημοσίευσης στο διαδίκτυο. Δόμηση και μορφοποίηση περιεχομένου. Οι γλώσσες επισήμανσης HTML και XML. Μορφοποίηση με CSS και ο μορφότυπος.

Επισκόπηση θεμάτων

- Δημοσιεύοντας στον παγκόσμιο ιστό
- Διαχωρισμός δόμησης και μορφοποίησης
- Οι γλώσσες επισήμανσης HTML και XML.
- Μορφοποίηση με CSS

Περιεχόμενα

- 1** Σχετικά με το σημερινό μάθημα (Εβδ.2η)
 - Επισκόπηση θεμάτων
 - Πρότυπα ηλεκτρονικής δημοσίευσης

- 2** Δημοσίευση για τον παγκόσμιο ιστό (Online Publication)
 - Κίνητρα δημοσίευσης στο WWW
 - Πρότυπα γλωσσών επισημείωσης (Markup Language Standards)
 - Το πρότυπο HTML
 - Το πρότυπο XML
 - Μιλώντας για σημασιολογία

Η ανάγκη για πρότυπα

Προκειμένου ένα δημοσίευμα (έντυπο ή ηλεκτρονικό) να είναι αναγνώσιμο και καταληπτό από τους αναγνώστες του, πρέπει να υπόκειται και να ακολουθεί ορισμένους κανόνες δημοσίευσης. Για παράδειγμα, χρησιμοποιούμε τους χαρακτήρες ενός γνωστού, δημόσιου αλφαβήτου, με τους οποίους σχηματίζουμε λέξεις, προτάσεις, παραγράφους κλπ.

Αυτά τα λεκτικά στοιχεία, υπόκεινται επίσης σε κανόνες γραμματικής & σύνταξης που καθορίζουν την διάρθρωση ενός κειμένου.

Αυτοί οι κανόνες πρέπει να είναι δημοσίως γνωστοί και σαφώς ορισμένοι, δημιουργώντας τα πρότυπα σύνταξης και μορφοποίησης των εγγράφων.

Η ανάγκη για πρότυπα

Όταν τα πρότυπα δημοσίευσης απουσιάζουν, τότε αφήνεται κανείς στις διαθέσεις των εταιρειών παραγωγής λογισμικών ηλεκτρονικής δημοσίευσης, που ήταν πολύ δημοφιλείς στο παρελθόν αλλά πλέον είναι πολύ δύσκολο να αναγνώσει κανείς τα τεκμήρια που δημιουργήθηκαν με τα λογισμικά τους.

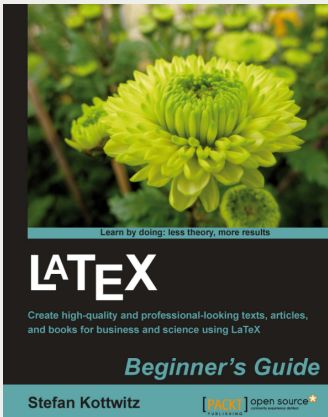
Αντί να προσπαθήσουμε να διατηρούμε διαθέσιμα αυτά τα παλαιά λογισμικά και συνοδά εργαλεία "ε*ς αιεί*", είναι ορθότερο και αποτελεσματικότερο να θεσπίσουμε πρότυπα (standards) δημοσίευσης, τα οποία θα είναι πάντα διαθέσιμα και θα ορίζουν τους κανόνες σύνταξης και προσπέλασης περιεχομένου ηλεκτρονικών τεκμηρίων.

Τα πρότυπα SGML & HTML

Σε κάθε ηλεκτρονικό έγγραφο (ιδίως κείμενο) διακρίνουμε μια δομή ή διάρθρωση (layout). Η δομή αυτή επισημαίνεται (marked-up) με σχόλια από τον συντάκτη ή από τον επιμελητή δημοσίευσης, προκειμένου ο τυπογράφος ή ο γραφίστας να γνωρίζει πώς να διατάξει τα δομικά στοιχεία του κειμένου επάνω στην σελίδα ώστε αυτό να καταστεί ευανάγνωστο.

Τα σχόλια καθορίζουν θέσεις των στοιχείων (π.χ. επικεφαλίδας), τύπο και μέγεθος γραμματοσειράς, μορφή σελιδοποίησης, έκταση περιθωρίων σελίδας, μέγεθος εικόνων, κ.α.

Παράδειγμα επισημειώσεων μορφοποίησης



Font face: Calibri
Size: 11px (standard doc)
Paragraph spacing: 13px

Αριθμός κεφαλαίου

1

Τίτλος

Getting Started with LaTeX

"Quote"

Are you ready to leave those "what you see is what you get" word processors behind and to enter the world of real, reliable, and high-quality typesetting? Then let's go together!

It's great that you decided to learn LaTeX. This book will guide you along the way to help you get the most out of it. Let's speak briefly about LaTeX's benefits and the challenges, and then we shall prepare our tools.

In this chapter, we will:

- Get to know LaTeX and talk about the pros and cons compared to word processors
- Install a complete LaTeX software bundle, including an editor
- Write our first LaTeX document

So, let's get started.

What is LaTeX?

Τίτλος κεφαλίδας

LaTeX is a software for typesetting documents. In other words, it's a document preparation system. LaTeX is not a word processor, but is used as a document markup language.

LaTeX is a free, open source software. It was originally written by Leslie Lamport and is based on the TeX typesetting engine by Donald Knuth. People often refer to it as just TeX, meaning LaTeX. It has a long history; you can read about it at <http://www.tug.org/whatis/latex.html>.

For now, let's continue by looking at how we can make the best use of it.

παραπομπή

Περιεχόμενα

- 1 Σχετικά με το σημερινό μάθημα (Εβδ.2η)
 - Επισκόπηση θεμάτων
 - Πρότυπα ηλεκτρονικής δημοσίευσης
- 2 Δημοσίευση για τον παγκόσμιο ιστό (Online Publication)
 - Κίνητρα δημοσίευσης στο WWW
 - Πρότυπα γλωσσών επισημείωσης (Markup Language Standards)
 - Το πρότυπο HTML
 - Το πρότυπο XML
 - Μιλώντας για σημασιολογία

Δημοσίευση στον Ιστό

Γιατί να δημοσιεύσουμε on-line;

- Χρειάζεται να δημοσιεύουμε γρηγορότερα.
- Χρειάζεται να δημοσιεύουμε με μικρότερο κόστος.
- Χρειάζεται να κάνουμε τα δημοσιεύματά μας πιο εύκολα διαθέσιμα.
- Βιώνουμε ρυθμούς δημοσιεύσεων και πληροφόρησης όπως ποτέ στο παρελθόν.
- Χρειάζεται να μπορούμε να έχουμε και να επεξεργαζόμαστε πληροφορίες ταχύτατα, προκειμένου να παραμείνουμε ανταγωνιστικοί.
- Ο χρόνος κατά τον οποίο διατίθεται μια πληροφορία (timeliness) συχνά καθορίζει και την αξία της.

Δημοσίευση στον Ιστό

- Η ανάπτυξη και ευρεία διάχυση Η/Υ και διαδικτύου διευκολύνει την ηλεκτρονική δημοσίευση και εξυπηρετεί τους προηγούμενους στόχους
- Οι εκδότες εκτιμούν ότι το κόστος δημοσίευσης μπορεί να μειωθεί έως και κατά 30%, παρά το γεγονός ότι εγείρονται νέα κόστη (π.χ. για Η/Υ, servers, λογισμικά, εξειδικευμένο προσωπικό, κ.α.
- Τα ηλεκτρονικά δημοσιεύματα (τεκμήρια) έχουν ορισμένες ελκυστικές αλλά και πολύτιμες ιδιότητες

Ιδιότητες ηλεκτρονικών δημοσιευμάτων

- **Υψηλή διαθεσιμότητα:** ένα ηλεκτρονικό δημοσίευμα μπορεί πλέον να είναι διαθέσιμο οποιαδήποτε χρονική στιγμή και όχι μόνον κατά τις "ώρες γραφείου". Επίσης, μπορεί να διατίθεται οπουδήποτε στον κόσμο υπάρχει σύνδεση στο διαδίκτυο.
- **Timeliness & immediacy:** ένα ηλεκτρονικό τεκμήριο μπορεί να διατίθεται άμεσα, με αποτέλεσμα το περιεχόμενό του να παραμένει επίκαιρο κατά τη στιγμή της ανάγνωσής του. Μετά τη δημοσίευσή του, το τεκμήριο είναι άμεσα διαθέσιμο προς τους τελικούς αποδέκτες.

Ιδιότητες ηλεκτρονικών δημοσιευμάτων

- **Διάθεση, ανεξάρτητη από την τοποθεσία:** ασχέτως με το που βρίσκεται ο τελικός αποδέκτης και με το που δημοσιεύθηκε το ηλεκτρονικό τεκμήριο, αυτό καθίσταται άμεσα διαθέσιμο μέσω του διαδικτύου.
- **Ολοκληρωμένη κάλυψη (comprehensive coverage):** ο τελικός αναγνώστης δεν περιορίζεται μόνον στα δημοσιεύματα ενός συγκεκριμένου τόπου, παρά αποκτά πρόσβαση και σε δημοσιεύματα πέραν του τόπου του, αποκτώντας μια ολοκληρωμένη εικόνα για το θέμα που τον ενδιαφέρει (βλ. περίπτωση λογοκρισίας/προπαγάνδας).

Ιδιότητες ηλεκτρονικών δημοσιευμάτων

- **Ταυτόχρονη διάθεση/προσπέλαση:** ένα ηλεκτρονικό δημοσίευμα μπορεί να χρησιμοποιείται ταυτόχρονα και ανεξάρτητα από πολλαπλούς χρήστες/αναγνώστες, χωρίς να υπόκειται σε περιορισμό πλήθους αντιτύπων.
- **Ασφαλής διατήρηση:** ένα ηλεκτρονικό δημοσίευμα είναι ευκολότερο να διατηρείται μακροχρόνια, ακέραιο, δηλαδή χωρίς φθορά του περιεχομένου του (απώλεια σελίδων, ξεθώριασμα χρωμάτων, κλπ).
- **Ενσωμάτωση πολυμεσικού περιεχομένου:** πέραν του κειμένου που συνοδεύει ένα δημοσίευμα, τα ηλεκτρονικά τεκμήρια μπορούν ταυτόχρονα να φέρουν και αποσπάσματα ήχων και animation/video.

Ιδιότητες ηλεκτρονικών δημοσιευμάτων

- **Δυνατότητα αλληλεπίδρασης:** τα έγγραφα μπορούν να υποστηρίξουν λειτουργίες αλληλεπίδρασης με τον αναγνώστη/χρήστη. Για παράδειγμα, μπορεί να περιλαμβάνουν υπερσυνδέσμους (hyper-links) ή να επιτρέπουν την διαμόρφωση της εμφάνισής τους (αλλαγή μεγέθους/τύπου γραμματοσειράς), κ.α.

“Το νόημα της ηλεκτρονικής διάθεσης πληροφοριών είναι να εκμεταλευθεί κανείς τις νέες χρήσης πληροφορίας οι οποίες δεν θα ήταν διαθέσιμες αν η πληροφορία απλά αποτυπώνονταν σ’ ένα φύλλο χαρτί”, Boyce, Pilachowski and Dalterio (1993), p.133

Περιεχόμενα

- 1 Σχετικά με το σημερινό μάθημα (Εβδ.2η)
 - Επισκόπηση θεμάτων
 - Πρότυπα ηλεκτρονικής δημοσίευσης
- 2 Δημοσίευση για τον παγκόσμιο ιστό (Online Publication)
 - Κίνητρα δημοσίευσης στο WWW
 - Πρότυπα γλωσσών επισημείωσης (Markup Language Standards)
 - Το πρότυπο HTML
 - Το πρότυπο XML
 - Μιλώντας για σημασιολογία

Τα πρότυπα SGML & HTML

Επισημειώσεις εγγράφων (Document Markup)

Χάρη σ' αυτές τις οδηγίες, το έγγραφο αποκτά μια συγκεκριμένη μορφή, η οποία μπορεί να εφαρμόζεται και σε άλλα έγγραφα, προσδίδοντας έτσι μια ομοιομορφία μεταξύ των δημοσιεύσεων (π.χ. άρθρα σε επιστημονικά περιοδικά, πτυχιακές εργασίες σ' ένα ακαδημαϊκό Τμήμα, κ.α.)

Επιπλέον, η μορφοποίηση έχει οριστεί με τρόπο σαφή και μονοσήμαντο και δεν εξαρτάται από τον ίδιο τον τυπογράφο, ο οποίος πλέον ακολουθεί οδηγίες και κοινούς, γνωστούς κανόνες.

(rely on processes, not persons.)

Γλώσσες επισημειώσεων (Markup Languages)

Μια γλώσσα επισημειώσεων (Markup Language) ορίζει τους κανόνες σύμφωνα με τους οποίους θα διαμορφώνονται τα χαρακτηριστικά των δομικών στοιχείων ενός εγγράφου, π.χ. τίτλος, κεφαλίδες, παράγραφοι, λίστες, πίνακες, εικόνες, παραπομπές, κ.α.

Η γλώσσα επισημειώσεων καθορίζει την μορφή και όχι το περιεχόμενο του εγγράφου.

Standard Generalized Markup Language - SGML

Το 1986 δημοσιεύεται από τον ISO το πρότυπο ISO-8879, το οποίο συνιστά το πρώτο δημοσιευμένο πρότυπο περιγραφής της εμφάνισης ενός ηλεκτρονικού εγγράφου, δηλαδή ένα σύνολο κανόνων που περιγράφουν την μορφή του εγγράφου.

Μέρη ενός προτύπου ηλεκτρονικής δημοσίευσης

Κάθε ηλεκτρονικό έγγραφο το οποίο ακολουθεί το πρότυπο SGML αποτελείται από 4 μέρη :

- 1** **μια δήλωση (declaration)** η οποία ορίζει τον τρόπο με τον εφαρμόζονται συγκεκριμένα τμήματα του προτύπου,
- 2** μια περιγραφή προσδιορισμού του τύπου του εγγράφου (**document type definition - DTD**) η οποία περιγράφει τη σχέση μεταξύ των δομικών στοιχείων του εγγράφου,
- 3** το ίδιο το κείμενο (**text instance**) στο οποίο έχει καταγραφεί η πρωτογενής πληροφορία,
- 4** ένα φύλλο μορφοποίησης (**style sheet**) το οποίο καθορίζει πώς θα εμφανίζονται τα δομικά στοιχεία του εγγράφου, στο τελικό, δημοσιευμένο κείμενο.

Μέρη ενός προτύπου ηλεκτρονικής δημοσίευσης

- Εάν ένα από αυτά τα 4 στοιχεία διαφοροποιηθεί, τότε θα αλλάξει και η τελική μορφή του ηλεκτρονικού δημοσιεύματος. Για παράδειγμα, αλλάζοντας την περιγραφή DTD μπορούμε να δημιουργήσουμε μια έκδοση του κειμένου μας σε μορφή Braille.
- Το πρότυπο SGML αναπτύχθηκε πριν την ευρεία διάδοση του παγκόσμιου ιστού γι' αυτό αν και είναι ένα εξαιρετικό πρότυπο περιγραφής εγγράφων για αρχειοθέτηση και δημοσίευση, δεν είναι κατάλληλο για την ηλεκτρονική δημοσίευση.

Ανάγκη διαφοροποίησης εμφάνισης

There's more...

You can change the layout of the book in many ways. Choose your settings at the beginning, or even better—start writing your content without any hesitation—once you have a decent amount of text, you can better see the effects of layout changes. You can do this at any time. Let's take a look at some design ideas.

Changing the page layout

When the book is bound after printing, the binding can cost space, that is, there may be less of the inner margin visible. You can specify a binding correction to compensate and to preserve layout proportions. So, if you would see 5 mm less of the inner margin after binding, add `BCOR=5mm` as class the option at the beginning. A similarly produced book may give you an idea of a good value.

Το κείμενο παραμένει ευανάγνωστο στο μεγάλο μέγεθος της οθόνης του επιτραπέζιου Η/Υ.

Ανάγκη διαφοροποίησης εμφάνισης



Όμως όταν διατίθεται για προβολή σε μικρή οθόνη η αναγκαία σμίκρυνση της γραμματοσειράς το καθιστά ακατάλληλο για ανάγνωση.

Ανάγκη διαφοροποίησης εμφάνισης

Πώς μπορώ από αυτό...

The Progression of Digital Publishing: Innovation and the E-volution of E-books

John W. Warren, RAND Corporation, CA, USA

Abstract: E-books are beginning to emerge from their incunabula stage. While some may think of an e-book as just an electronic image of a paper product, others have used the electronic format to broaden the spectrum of publishing in the digital age. This paper examines three innovative examples that demonstrate the potential and challenges of electronic publications. The first is an online resource providing information on the U.S. health care system, descriptions of policy proposals, and an interactive microsimulation model that estimates the effects of commonly proposed policy changes. The second example is a digital novel utilizing text, sound, images, and gaming in storytelling. The third is a survey of efforts to create digital textbooks with online study resources. Each case study provides insight into the possible future of the e-book.

Keywords: E-books, Information Society, Electronic Texts, Innovation, Digital Tools, Collaboration, Future of E-books

Out of Incunabula

E-BOOKS ARE EMERGING from their incunabula state, although this transition is far from complete. The term incubula (meaning infancy, from the Latin for "cradle" or "swaddling clothes") refers to the earliest books printed with movable type, during the 50-year transitional phase from the 1454 Gutenberg Bible to 1501. These books were initially considered inferior to illuminated manuscripts (Trithemius, 1492, p. 35). They were even considered dangerous, as printed books threatened the livelihood of monks, represented loss of control to elites, and led inexorably to the democratization of books and reading. Today, the e-book is emerging from a similar transition.

Ανάγκη διαφοροποίησης εμφάνισης

να δημιουργήσω αυτό...

The Progression of Digital Publishing: Innovation and the E-volution of E-books

John W. Warren, RAND Corporation, CA, USA

Abstract: E-books are beginning to emerge from their incunabula stage. While some may think of an e-book as just an electronic image of a paper product, others have used the electronic format to broaden the spectrum of publishing in the digital age. This paper examines three innovative examples that demonstrate the potential and challenges of electronic publications. The first is an online resource providing information on the U.S. health care system, descriptions of policy proposals, and an interactive microsimulation model that estimates the effects of commonly proposed policy changes. The second example is a digital novel utilizing text, sound, images, and gaming in storytelling. The third is a survey of efforts to create digital textbooks with online study resources. Each case study provides insight into the possible future of the e-book.

Keywords: E-books, Information Society, Electronic Texts, Innovation, Digital Tools, Collaboration, Future of E-books

Out of Incunabula

E-BOOKS ARE EMERGING from their incunabula state, although this transition is far from complete. The term *incubula* (meaning infancy, from the Latin for “cradle” or “swaddling clothes”) refers to the earliest books printed with movable type, during the 50-year transitional phase from the 1454 Gutenberg Bible to 1501. These books were initially considered inferior to illuminated manuscripts (Trithemius, 1492, p. 35). They were even considered dangerous, as printed books threatened the livelihood of monks, represented loss of control to elites, and led inexorably to the democratization of books and reading. Today, the e-book is emerging from a similar transition.

In an earlier article, I discussed the theme of e-books and innovation (Warren, 2009). The past 12 months have been marked by significant media attention on e-books, numerous e-reader devices have entered the market, and e-book sales have grown. This article discusses

Ανάγκη διαφοροποίησης εμφάνισης

ή αυτό...

The Progression of Digital Publishing: Innovation and the E-volution of E-books

John W. Warren, RAND Corporation, CA, USA

Abstract: E-books are beginning to emerge from their incunabula stage. While some may think of an e-book as just an electronic image of a paper product, others have used the electronic format to broaden the spectrum of publishing in the digital age. This paper examines three innovative examples that demonstrate the potential and challenges of electronic publications. The first is an online resource providing information on the U.S. health care system, descriptions of policy proposals, and an interactive microsimulation model that estimates the effects of commonly proposed policy changes. The second example is a digital novel utilizing text, sound, images, and gaming in storytelling. The third is a survey of efforts to create digital textbooks with online study resources. Each case study provides insight into the possible future of the e-book.

Keywords:

E-books, Information Society, Electronic Texts, Innovation, Digital Tools, Collaboration, Future of E-books

Out of Incunabula

E-BOOKS ARE EMERGING from

their incunabula state, although this transition is far from complete. The term incubula (meaning infancy, from the Latin for “cradle” or “swaddling clothes”) refers to the earliest books printed with movable type, during the

50-year transitional phase from the 1454 Gutenberg Bible to 1501. These books were initially considered inferior to illuminated manuscripts (Trithemius, 1492, p. 35). They were even considered dangerous, as printed books threatened the livelihood of monks, represented loss of control to elites, and led inexorably to the democratization of books and reading. Today, the e-book is emerging from a similar transition.

Περιεχόμενα

- 1 Σχετικά με το σημερινό μάθημα (Εβδ.2η)
 - Επισκόπηση θεμάτων
 - Πρότυπα ηλεκτρονικής δημοσίευσης
- 2 Δημοσίευση για τον παγκόσμιο ιστό (Online Publication)
 - Κίνητρα δημοσίευσης στο WWW
 - Πρότυπα γλωσσών επισημείωσης (Markup Language Standards)
 - Το πρότυπο HTML
 - Το πρότυπο XML
 - Μιλώντας για σημασιολογία

Το πρότυπο HTML



Το 1990 ο Tim Berners-Lee στο CERN δημιούργησε το πρότυπο HyperText Markup Language - HTML, ως μια επέκταση της SGML, κατάλληλης για χρήση στον παγκόσμιο ιστό.

Βασικές απαιτήσεις για έγγραφα HTML

- Η μορφοποίηση του κειμένου θα έπρεπε να είναι δυναμική και όχι σταθερή και αμετάβλητη.
- Το κείμενο θα έπρεπε να είναι διαθέσιμο στον καθένα, μέσω του παγκόσμιου ιστού, και να μην εξαρτάται από το σύστημα που χρησιμοποιεί κανείς για την ανάγνωσή του ούτε από το πρόγραμμα που χρησιμοποιήθηκε για τη συγγραφή του.

Το πρότυπο HTML

Κατά τις αρχές του 1990:

Έγγραφα τα οποία δημιουργούνται κατά το πρότυπο HTML:

- χρησιμοποιούν μια περιγραφή DTD η οποία διαθέτει ένα περιορισμένο πλήθος επισημειώσεων (tags)
- είναι προσπελάσιμα/αναγνώσιμα από το ελεύθερα διαθέσιμο λογισμικό Mosaic
- υποστηρίζουν διασύνδεση (hyperlinks) προς άλλα έγγραφα.

Πλέον βρισκόμαστε στην 5η έκδοση του προτύπου (HTML 5).

<https://html.spec.whatwg.org/>

Το πρότυπο HTML

Κατά το πρότυπο HTML κάθε έγγραφο που συμμορφώνεται με το πρότυπο:

- 1** διαθέτει μια δήλωση τύπου αρχείου (DOCTYPE declaration)
- 2** συνοδεύεται από μια περιγραφή DTD που ορίζει δομικά στοιχεία του κειμένου
- 3** περιλαμβάνει την πρωτογενή πληροφορία (κυρίως κείμενο)
- 4** συνδέεται με μια περιγραφή μορφοποίησης των δομικών του στοιχείων (αυτών που ορίστηκαν στην περιγραφή DTD).

Μέρος 1ο: DOCTYPE declaration

Η δήλωση του τύπου αρχείου του εγγράφου γίνεται στην αρχή και χρησιμοποιείται για να υποδείξει στο πρόγραμμα προβολής και ανάγνωσης του κειμένου **τι είδους** αρχείο πρόκειται να προβάλλει.

Οι συνηθέστερα χρησιμοποιούμενοι τύποι αρχείων για δημοσίευση στο διαδίκτυο είναι οι εξής:

- XHTML
- HTML 4
- HTML 5

Ο προσδιορισμός του τύπου έχει σημασία επειδή συνδέεται με τα είδη των διαφορετικών δομικών στοιχείων που μπορεί να αξιοποιήσει. Δεν είναι όλα τα δομικά στοιχεία διαθέσιμα για όλους τους τύπους εγγράφων.

Μέρος 1ο: DOCTYPE declaration

Παράδειγμα δήλωσης (DOCTYPE declaration)

Η δήλωση τύπου αρχείου συνδέει ένα επισημειωμένο αρχείο (π.χ. SGML/HTML) με ένα σύνολο οδηγιών δόμησης (του εγγράφου) που περιγράφεται σε μια περιγραφή DTD. Περιλαμβάνεται στο ίδιο το πρωτογενές κείμενο και έχει την ακόλουθη μορφή:

```
<!DOCTYPE html PUBLIC
  "-//W3C//DTD XHTML 1.0 Transitional//EN"
  "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
<html lang="ar" dir="ltr" xmlns="http://www.w3.org/1999/xhtml">
```

Στην δήλωση αυτή καθορίζεται εάν η περιγραφή DTD είναι δημόσια (PUBLIC) ή ιδιωτική (SYSTEM).

Αν πρόκειται για δημόσια περιγραφή, ακολουθεί εντός εισαγωγικών το επίσημο αναγνωριστικό δημοσίευσης (*Formal Public Identifier - FPI*). Μετά ακολουθεί το αναγνωριστικό συστήματος (System Identifier), το οποίο εξειδικεύει τον FPI.

Μέρος 1ο: DOCTYPE declaration

Q: Γιατί έχει σημασία η δήλωση τύπου εγγράφου;

A: Επειδή καθορίζει τα είδη των τύπων δεδομένων που μπορούν να χρησιμοποιηθούν.

Tag	HTML5	HTML4	XHTML
< <i>a</i> >	Nαι	Nαι	Nαι
< <i>abbr</i> >	Nαι	Nαι	Nαι
< <i>acronym</i> >	Όχι	Nαι	Nαι
< <i>address</i> >	Nαι	Nαι	Nαι
...

Μέρος 2ο: Περιγραφή ενός DTD

Παράδειγμα Document Type Definition - DTD

- Η περιγραφή DTD καθορίζει τη δομή και τους τύπους δεδομένων (δομικών στοιχείων ή **entities**) που εμφανίζονται στο σώμα του κειμένου.
- Τέτοιοι τύποι μπορεί να είναι: αποσπάσματα κειμένου, ημερομηνίες, αριθμοί, σχήματα, σύνδεσμοι, κ.α.

Έτσι, μπορούμε να ορίσουμε διάφορους τύπους δεδομένων στην περιγραφή DTD ώστε ακολούθως να μπορούμε εύκολα να διαφοροποιούμε τον τρόπο εμφάνισης αυτών των τύπων/τμημάτων του εγγράφου μας.

Μέρος 2ο: Περιγραφή ενός DTD

Για παράδειγμα, ένα κεφάλαιο βιβλίου πρέπει να περιλαμβάνει:

- τον αριθμό κεφαλαίου
- τον τίτλο του κεφαλαίου
- μια εισαγωγή
- ένα υποσέλιδο στο οποίο θα εμφανίζεται ο αριθμός σελίδας
- μια περίληψη του κεφαλαίου, στο τέλος.
- ένα σύνολο βιβλιογραφικών παραπομπών

Η περιγραφή DTD ορίζει τους τύπους των δομικών στοιχείων που περιλαμβάνονται στο έγγραφο και δεν διαμορφώνει την εμφάνισή τους. Τους ορίζει, προκειμένου κατά την φάση της μορφοποίησης αυτοί να αναγνωρισθούν και να μορφοποιηθούν καταλλήλως.

Μέρος 2ο: Περιγραφή ενός DTD

The screenshot shows a news article on the ERT website. Blue boxes highlight various HTML elements: the main title, a sub-header, a photo, a text block, a sidebar section, and a footer. The article text discusses the discovery of the bronze statue of Alexander the Great in Pella, Macedonia.

EPT ΕΙΔΗΣΕΙΣ ▾ ΠΡΟΓΡΑΜΜΑ ΕΡΤFLIX LIVE ▾ ΑΡΧΕΙΟ ΠΕΡΙΦΕΡΕΙΑ ▾ 🔍

Αποκαλύπτεται το ανάκτορο που γεννήθηκε ο Μέγας Αλέξανδρος

ΕΙΔΗΣΕΙΣ

ΤΕΛΕΥΤΑΙΑ ΝΕΑ

Το πρώτο επίσημο Βαρόμετρο των Περιφερειών και Δήμων της ΕΕ παρουσιάζει ο Α. Τσίπρας, παρουσία της προέδρου της Κομισιόν

Θεσσαλονίκη: Σύλληψη δύο αλλοδαπών για βουαιήση ναρκωτικών μεταστατών

Θεσσαλονίκη: Εμπρησμοί σε δύο αυτοκίνητα τα Σπυριδάκια της Τρίτης

Θεσσαλονίκη: Δύο μακερυνέσιες σήμερα στο κέντρο της πόλης

Άλλες θετικές: Μία από τις λίστες η υπερπαιδική κρίση μάσκα παικτών (video)

ροή ειδήσεων

Ο Δήμος Ρόδου εγκαινιάζει τα παλάτια των Φοιδί

Το πρώτο επίσημο Βαρόμετρο των Περιφερειών και Δήμων της ΕΕ παρουσιάζει ο Α. Τσίπρας, παρουσία της προέδρου της Κομισιόν

Μιλύβος σε φερίκι ναρκωτικών στην ασπαστική Μύλη

Το ερώτημα καινούργια εκτιμάται ότι θα ανοίξει τις πύλες του επισκέπτες το ανάκτορο στο οποίο γεννήθηκε ο Μέγας Αλέξανδρος στην παραθαλάσσια τότε Πέλλα. Στο κτίριο αυτό, τμήμα του οποίου είχε ανακαταστεί το 1953 όταν κανείς δεν γνώριζε ότι επρόκειτο για το ανάκτορο, μεγάλου ο βουαίήση των Μακεδόνων ενώ στην παλαιότητα που βρίσκεται σε μικρή απόσταση από το κτίριο, ο βουαίήση σφαιλούνταν με τους γόνους της αριστοκρατίας και στην τεράστια κολυμβητική δεξαμενή γύριζαν το σώμα τους.

«Το ανάκτορο είχε δημόσιο χαρακτήρα και σε αυτό βρισκόταν η αίθουσα όπου γίνονταν τα συμβόσια, η αίθουσα του θρόνου, αν μου επιτρέπεται να χρησιμοποιήσω αυτή την έκφραση και αναλογία» εξήγησε στο ΑΠΕ - ΜΠΕ η προϊσταμένη της Εφορείας Αρχαιοτήτων Πέλλας, Ελισάβετ Τσιγαρίδα. Η ίδια τόνισε ότι οι εργασίες στο κτίριο βρίσκονται αυτή τη στιγμή σε πλήρη εξέλιξη, με χρηματοδότηση από το ΕΣΠΑ, και σημειώνει ότι το σημείο θα μπορεί να είναι επισκεπτικό το καλοκαίρι του 2021, αν όλα εξέλθουν ούλα.

Παράδειγμα ενός διαδικτυακού άρθρου

Αναγνωρίζουμε διάφορα δομικά στοιχεία του άρθρου, όπως:

■ τίτλο, ημερομηνία δημοσίευσης, κυρίως σώμα κειμένου, ημερομηνίες, κλπ.

Μέρος 2ο: Παράδειγμα περιγραφής DTD

Απόσπασμα μιας περιγραφής Document Type Definition

```
<!--===== Imported Names =====>
<!ENTITY % Text "CDATA">
  <!-- used for titles etc. -->
<!ENTITY % Datetime "CDATA">
  <!-- date and time information. ISO date format -->
<!ENTITY % Number "CDATA">
  <!-- one or more digits -->
<!ENTITY % Shape "(rect|circle|poly|default)">
<!ENTITY % URI "CDATA">
  <!-- a Uniform Resource Identifier, see [RFC2396] -->
```

Τα δομικά στοιχεία του εγγράφου μπορεί να είναι προκαθορισμένες οντότητες (ENTITIES) ή κάποια από τα δομικά στοιχεία ενός κειμένου HTML τα οποία υποστηρίζει ο τύπος εγγράφου (DOCTYPE) που χρησιμοποιούμε (π.χ. heading, paragraph, κ.α.).
(*CDATA stands for character data.*)

Μέρος 3ο: Δείγμα κειμένου HTML

Παράδειγμα κειμένου HTML

```
<!DOCTYPE HTML PUBLIC
"-//W3C//DTD HTML 4.01 Transitional//EN"
"http://www.w3.org/TR/html4/loose.dtd">
<html>
  <head>
    <title>A sample of an HTML document</title>
    <link rel="stylesheet" href="styles.css">
  </head>

  <body>
    <h1>This is a heading</h1>
    <p>This is a paragraph which contains <strong>some words
in boldface</strong>.
    </p>
    <p>Here is a MARC exerpt:</p>
    <code>
01041cam 2200265 a 4500 <br/>
001 ###89048230 /AC/r91 <br/>
005 19911106082810.9 <br/>
008 891101d1990 maua j 001 0 eng <br/>
010 ## $a $$89048230 /AC/r91 <br/>
020 ## $a 0316107514 <br/>
    </code>
  </body>
</html>
```


Μέρος 4ο: Παράδειγμα οδηγιών μορφοποίησης (Style sheets)

Παράδειγμα μορφοποίησης κειμένου HTML

```
body {  
  background-color: powderblue;  
}  
h1 {  
  color: blue;  
}  
p {  
  color: red;  
}
```

Οι δηλώσεις μορφοποίησης συχνά περιγράφονται σε συνοδευτικά αρχεία τα οποία λέγονται "*Style sheets*"

Κατά την περιγραφή μορφοποίησης, ορίζονται οι τιμές των ιδιοτήτων των αντικειμένων που συνιστούν τα δομικά στοιχεία του εγγράφου (π.χ. το χρώμα των επικεφαλίδων ή του κειμένου κάθε νέας παραγράφου, το είδος της γραμματοσειράς, κ.α.).

Παράδειγμα μορφοποιημένου εγγράφου κατά HTML

Παράδειγμα τελικής εμφάνισης

This is a heading

This is a paragraph which contains **some words in boldface**.

Here is a MARC excerpt:

```
01041cam 2200265 a 4500
001 ###89048230 /AC/r91
005 19911106082810.9
008 891101d1990 maua j 001 0 eng
010 ## $a $$$89048230 /AC/r91
020 ## $a 0316107514
```

Πλέον, το περιεχόμενο εμφανίζεται σύμφωνα με τους κανόνες μορφοποίησης και αυτό είναι ανεξάρτητο από το πρόγραμμα που χρησιμοποιείται για την εμφάνιση του περιεχομένου.

Επισημειώσεις (tags) ενός εγγράφου HTML

Το πρότυπο HTML υποστηρίζει διάφορα στοιχεία (**tags or elements**) επισημειώσεων για την περιγραφή της δομής ενός εγγράφου. Τα στοιχεία αυτά αφορούν:

- στα μεταδεδομένα του εγγράφου (π.χ. *head, title, meta, style, ...*)
- στις ενότητες ενός εγγράφου (π.χ. *body, article, section, heading, header, footer, ...*)
- σε ομαδοποιήσεις του κειμένου του εγγράφου (π.χ. *paragraph, blockquote, division, lists, figure, table, ...*)
- σε διαμορφώσεις χαρακτήρων (π.χ. *bold, italics, small, abbreviation, time, code, ...*)
- σε παραπομπές (links) (π.χ. *a, area*)

Παραπομπή: <https://html.spec.whatwg.org/>

Επισημειώσεις (tags) ενός εγγράφου HTML

Κάθε στοιχείο επισημείωσης (markup) φέρει ένα opening tag και ένα closing tag. Οι ονομασίες των tags εσωκλείονται μεταξύ των συμβόλων < και >.

Προκειμένου να διακρίνεται ένα closing tag από ένα opening tag, το πρώτο φέρει τον χαρακτήρα / αμέσως μετά το σύμβολο < και πριν την ονομασία του.

Για παράδειγμα:

```
<strong>This is displayed as boldface</strong>
```

Οι επισημειώσεις μπορούν και να συνδυάζονται (nesting). Π.χ.:

```
<strong>The word <i>house</i> appears as both  
boldface and emphasized</strong>
```

Ιδιότητες (attributes) επισημειώσεων

Κάθε στοιχείο (tag or element) που χρησιμοποιείται για την περιγραφή της δομής ενός HTML εγγράφου, φέρει ορισμένες ιδιότητες (**attributes**) οι οποίες καθορίζουν τα χαρακτηριστικά του.

Για παράδειγμα, το στοιχείο παραπομπής **a** (anchor) φέρει την ιδιότητα *href* η οποία χρησιμοποιείται για να προσδιορίσει το URL της παραπομπής.

```
<a href="https://www.w3schools.com">Visit  
W3Schools.com!</a>
```

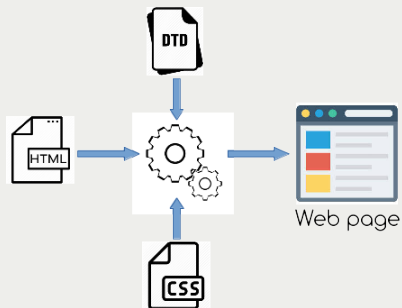
Καθορισμός μορφοποίησης με Style sheets

Η περιγραφή μορφοποίησης μπορεί να γίνεται είτε εντός του ίδιου του αρχείου HTML ή όταν είναι εκτενής και αρκετά λεπτομερής, να γίνεται σε συνδεδεμένο αρχείο (π.χ. styles.css)

```
h1 {  
display : block;  
font-size : 2em;  
margin-top : 0.67em;  
margin-bottom: 0.67em;  
margin-left : 0;  
margin-right : 0;  
font-weight : bold;  
}
```

Η περιγραφή μορφοποίησης μπορεί να ορίζει χρώματα, γραμματοσειρές, αποστάσεις, κ.α.

Δημιουργία εγγράφου ιστοσελίδας



Ανακεφαλαιώνοντας:

Η πρωτογενής πληροφορία περιγράφεται στο HTML αρχείο, μαζί με τις σχετικές επισημειώσεις.

Το αρχείο DTD περιγράφει τις οντότητες και τύπους δεδομένων που διακρίνονται στο έγγραφο. Το αρχείο CSS περιλαμβάνει τις οδηγίες μορφοποίησης των δομικών στοιχείων του εγγράφου. Οι πληροφορίες αυτών των τριών αρχείων (τμημάτων) συνδυάζονται για να δημιουργήσουν τελικώς την ιστοσελίδα που εμφανίζεται στην οθόνη της συσκευής του αναγνώστη.

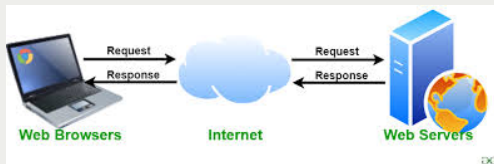
Το πρότυπο HTML

Μια πλήρης λίστα των στοιχείων επισημειώσεων που υποστηρίζουν οι 3 δημοφιλέστεροι τύποι κειμένου για τον παγκόσμιο ιστό (HTML5, HTML4, XHTML Doctypes) παρατίθεται εδώ:

https://www.w3schools.com/tags/ref_html_dtd.asp

Δημοσίευση στον παγκόσμιο Ιστό

- Μπορούμε πλέον να δημιουργούμε έγγραφα για τον παγκόσμιο ιστό
- Τα έγγραφα αυτά μεταφέρονται σε διακομιστές (H/Y) οι οποίοι είναι διαρκώς σε λειτουργία και διασυνδεδεμένοι
- Όποιος επιθυμεί να δει το σχετικό έγγραφο, χρησιμοποιεί τον φυλλομετρητή (browser) της συσκευής του και ζητά από τον διαμομιστή την αποστολή του περιεχομένου του.



Κωδικοποίηση χαρακτήρων κειμένου (HTML encoding)

Όπως γνωρίζουμε, για την αναπαράσταση συμβόλων χαρακτήρων χρησιμοποιείται ένας πίνακας αντιστοίχισης μεταξύ των χαρακτήρων ενός αλφαβήτου και δυαδικών λέξεων που αποτελούνται από 0 και 1.

- Το πρόγραμμα προβολής περιεχομένου αρχείων HTML πρέπει να γνωρίζει τον τρόπο αντιστοίχισης που χρησιμοποιήθηκε κατά την συγγραφή του κειμένου.
- Ένας γνωστός και απλός πίνακας αντιστοίχισης των χαρακτήρων του λατινικού αλφαβήτου είναι ο πίνακας ASCII.
- Ωστόσο, αυτός αναπαριστά μόνον 256 σύμβολα και μεταξύ αυτών δεν περιλαμβάνονται χαρακτήρες πέραν του λατινικού αλφαβήτου.

Προκειμένου να συντάξουμε ένα κείμενο το οποίο είναι πολύγλωσσο (π.χ. ελληνικό/αγγλικό) χρειαζόμαστε έναν άλλο πίνακα αντιστοίχισης.

Ο δημοφιλέστερος και προεπιλεγμένος τρόπος κωδικοποίησης χαρακτήρων είναι ο **UTF-8** ο οποίος περιλαμβάνει τα σύμβολα πάρα πολλών αλφαβήτων (ακόμη και γλώσσας Klingon).

Προκειμένου να γνωρίζει ο browser ποια κωδικοποίηση θα χρησιμοποιήσει για το κείμενό μας, ενσωματώνουμε σ' αυτό μια σχετική δήλωση, εντός της επισημείωσης `<meta>` που εσωκλείεται στην επισημείωση `<head>` του κειμένου μας, ως εξής:

```
<meta http-equiv="Content-Type"  
content="text/html; charset=UTF-8">
```

Περιεχόμενα

- 1 Σχετικά με το σημερινό μάθημα (Εβδ.2η)
 - Επισκόπηση θεμάτων
 - Πρότυπα ηλεκτρονικής δημοσίευσης
- 2 Δημοσίευση για τον παγκόσμιο ιστό (Online Publication)
 - Κίνητρα δημοσίευσης στο WWW
 - Πρότυπα γλωσσών επισημείωσης (Markup Language Standards)
 - Το πρότυπο HTML
 - Το πρότυπο XML
 - Μιλώντας για σημασιολογία

XML definition

Η XML είναι μια ακόμη γλώσσα επισημειώσεων η οποία στοχεύει στην δυνατότητα επέκτασης λειτουργικότητων, εξ' ου και **eXtensible Markup Language (XML)**

Extensible

Χαρακτηρίζεται ως επεκτάσιμη επειδή μας δίνει την δυνατότητα να ορίσουμε τις **δικές μας** επισημειώσεις, τη **σειρά** με την οποία εμφανίζονται και τον **τρόπο** με τον οποίο θα επεξεργαστούν ή θα εμφανιστούν.

XML definition

Markup

Επιτρέπει την χρήση επισημειώσεων (markup elements), όπως και στην γλώσσα HTML. Δεν υπάρχει περιορισμός για το πλήθος και τον τύπο των επισημειώσεων. Το περιεχόμενο των επισημειώσεων καθορίζει τον τρόπο δόμησης της πρωτογενούς πληροφορίας.

Language

Η XML είναι και αυτή μια γλώσσα επισημειώσεων, παρόμοια με την HTML, επιτρέποντας όμως τη χρήση των δικών μας (custom) επισημειώσεων.

Μπορεί επίσης να θεωρηθεί ως υπερκείμενη γλώσσα (meta-language): μας επιτρέπει δηλαδή να δημιουργήσουμε και να ορίσουμε άλλες γλώσσες, όπως RSS, MathML ή ακόμη και εργαλεία όπως XSLT.

XML definition

Η XML περιγράφει την δομή αναπαράστασης πληροφορίας και όχι τον τρόπο εμφάνισής της.

- Σχεδιάστηκε για να διευκολύνει τον τρόπο αποθήκευσης και μεταφοράς/μετάδοσης δεδομένων και πληροφοριών
- Σχεδιάστηκε προκειμένου να είναι μια γλώσσα αυτο-περιγραφική (self-descriptive)
- Αποτελεί μια σύσταση (recommendation) του World Wide Web Consortium (W3C)

Παράδειγμα χρήσης XML

Ένα XML κείμενο μπορεί να αναπαριστά την ακόλουθη δομή:

Note

To: Tove

From: Jani

Reminder

Don't forget me this weekend!

Το κείμενο διαθέτει:

- **τύπο** (Note)
- **παραλήπτη** (Tove)
- **συντάκτη** (Jani)
- **τίτλο** (Reminder)
- **κείμενο** (Don't forget me this weekend)

Γιατί χρειαζόμαστε την XML;

Η XML:

- δημιουργήθηκε για να περιγράψει τη δομή οργάνωσης δεδομένων και πληροφοριών (structured data)
- δεν προορίζεται αποκλειστικά για χρήση στον παγκόσμιο ιστό
- δεν αφορά την μόνον την αλληλεπίδραση ανθρώπων-περιεχομένου. Μπορεί να χρησιμοποιείται και για την ανταλλαγή πληροφοριών/δεδομένων μεταξύ υπολογιστών (π.χ. κατά τη χρήση διαδικτυακών υπηρεσιών).

Σε αντίθεση με την HTML, δεν σχεδιάστηκε μόνον για τον παγκόσμιο ιστό, ούτε για να καθορίσει τον τρόπο εμφάνισης/απεικόνισης πληροφοριών.

Σε τι διαφέρει το πρότυπο XML από το πρότυπο HTML;

Σκοπός δημιουργίας της HTML ήταν να μπορέσει να επιτρέψει στον καθένα να δημιουργεί γρήγορα έγγραφα για τον παγκόσμιο ιστό τα οποία να μπορεί να διαθέτει σε άλλους, απομακρυσμένους χρήστες (π.χ. αρθρογραφία ειδήσεων).

Πώς αλλιώς θα μπορούσατε ως δημοσιογράφοι να διαθέσετε στους αναγνώστες ένα άρθρο το οποίο γράψατε; Σίγουρα, δεν μπορείτε να είστε σε επικοινωνία με τον κάθε ενδιαφερόμενο προκειμένου να του το ταχυδρομήσετε.

Σε τι διαφέρει το πρότυπο XML από το πρότυπο HTML;

Η γλώσσα HTML δεν περιγράφει πώς είναι δομημένο ένα έγγραφο ή τι σημαίνει η πληροφορία που εμπεριέχεται σ' αυτό. Είναι μια γλώσσα παρουσίασης (μορφοποίησης εμφάνισης) ενώ η XML είναι μια γλώσσα περιγραφής δόμησης δεδομένων.

Η HTML περιγράφει τον τρόπο μορφοποίησης ενός εγγράφου και αυτός θα είναι ο ίδιος είτε το έγγραφο απεικονίζεται στην μεγάλη οθόνη ενός επιτραπέζιου Η/Υ είτε στην μικρή οθόνη ενός κινητού τηλεφώνου (καθιστώντας το μάλλον δυσανάγνωστο).

Η XML μπορεί να ορίσει την δομή του κειμένου ώστε στην περίπτωση εμφάνισής του σε μια μικρότερη οθόνη, να διαφοροποιείται η διάρθρωσή του, καθιστώντας το ευανάγνωστο.

Ένα παράδειγμα κειμένου HTML

Περιγραφή μορφοποίησης κατά HTML και αποτέλεσμα παρουσίασης.

```

<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
  "https://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
<html xmlns="https://www.w3.org/1999/xhtml">
<head>
  <meta http-equiv="Content-Type"
    content="text/html; charset=UTF-8">
  <title>Κατάλογος βιβλίων</title>
</head>

<body>
<h1>Κατάλογος βιβλίων</h1>
<h2>Βιβλίο Α</h2>
<p>Το βιβλίο Α πραγματεύεται την ιστορία του ...</p>
<p><b>Τιμή: $19.95</b></p>
<p><b>ISBN: 978-3-16-148410-0</b></p>
<h2>Βιβλίο Β</h2>
...
<h3>Βιβλίο Γ</h3>
<p><i>Τιμή: $24.95</i></p>
<p>Αυτό το βιβλίο είναι ένας πολύτιμος οδηγός επιβίωσης <br/>
στο σύγχρονο περιβάλλον ....
</p>
...
</body>
</html>

```

Κατάλογος Βιβλίων

Βιβλίο Α

Το βιβλίο Α πραγματεύεται την ιστορία του ...

Τιμή: \$19.95

ISBN: 978-3-16-148410-0

Βιβλίο Β

...

Βιβλίο Γ

Τιμή: \$24.95

Αυτό το βιβλίο είναι ένας πολύτιμος οδηγός επιβίωσης
στο σύγχρονο περιβάλλον

...

Περιεχόμενα

- 1 Σχετικά με το σημερινό μάθημα (Εβδ.2η)
 - Επισκόπηση θεμάτων
 - Πρότυπα ηλεκτρονικής δημοσίευσης

- 2 Δημοσίευση για τον παγκόσμιο ιστό (Online Publication)
 - Κίνητρα δημοσίευσης στο WWW
 - Πρότυπα γλωσσών επισημείωσης (Markup Language Standards)
 - Το πρότυπο HTML
 - Το πρότυπο XML
 - Μιλώντας για σημασιολογία

Διαχωρισμός περιεχομένου από την εμφάνιση

Για τον υπολογιστή, η περιγραφή μορφοποίησης ορίζει **ΠΩΣ** πρέπει να απεικονιστεί το κείμενο.

Ωστόσο, δεν του παρέχει καμία πληροφορία για το **ΤΙ ΠΕΡΙΓΡΑΦΕΙ** το κείμενο.

Δεν αποτυπώνεται πληροφορία που θα του επιτρέψει να κάνει τα κατάλληλα ενοιολογικά/σημασιολογικά άλματα για να κατανοήσει το περιεχόμενό του.

Θα μπορούσαμε να αλλάξουμε το "Βιβλίο Α" σε "Barack Obama" και ο Η/Υ δεν θα μπορούσε να αναγνωρίσει/εντοπίσει το σφάλμα καταχώρησης (εγγραφή ανθρώπινου ονόματος εκεί που αναμένεται ονομασία βιβλίου).

Οι άνθρωποι είμαστε ικανότατοι στο να εξάγουμε πληροφορίες από δεδομένα και να τα ερμηνεύουμε. Οι Η/Υ αδυνατούν να το κάνουν, όταν δεν έχουν κατάλληλη καθοδήγηση.

***Tip:** τα τελευταία χρόνια η επιστήμη της Τεχνητής Νοημοσύνης (AI) έχει επιδείξει σημαντική πρόοδο
(βλ. <https://youtu.be/yDI5oVn0RgM>)

Οι άνθρωποι μπορούμε εύκολα να συμπεράνουμε ότι αυτό που περιγράφεται με έντονα (**bold**), μεγάλα γράμματα αντιστοιχεί στην ονομασία του βιβλίου, ότι αυτό που ακολουθεί είναι μια συνοπτική περιγραφή του αντικειμένου, και ότι οι επόμενες δύο γραμμές αφορούν στην εμπορική αξία και τον κωδικό ταυτοποίησης του βιβλίου.

Ακόμη και το γεγονός ότι για το 3ο βιβλίο :

- ο τίτλος του δεν εμφανίζεται με τόσο μεγάλα και έντονα γράμματα όπως των άλλων δύο,
- η τιμή εμφανίζεται με πλάγια γράμματα και όχι με έντονα,
- η περιγραφή ακολουθεί την τιμή και δεν προηγείται αυτής

δεν μας εμποδίζει να κατανοήσουμε τι είναι αυτό που περιγράφεται στο κείμενο. Για ένα υπολογιστικό σύστημα, καμία από αυτές τις

λειτουργίες δεν είναι δυνατή με μόνο εφόδιο την περιγραφική γλώσσα HTML. Η γλώσσα αυτή θα καθορίσει απλώς τον τρόπο εμφάνισης της πληροφορίας στην οθόνη, χωρίς να παρέχει κάποια πληροφορία για το νόημα του περιεχομένου της.

Παράδειγμα ενός XML αρχείου

```
<?xml version="1.0"?>
<bookListing title="ABC Products">
  <book>
    <name>Βιβλίο Α</name>
    <description>Το βιβλίο Α πραγματεύεται την
    ιστορία του ...</description>
    <cost>$19.95</cost>
    <isbn>978-3-16-148410-0</isbn>
  </book>

  <book>
    <name>Βιβλίο Β</name>
    ...
  </book>

  <book>
    <name>Βιβλίο Γ</name>
    <description>Αυτό το βιβλίο είναι ένας
    πολύτιμος οδηγός επιβίωσης στο σύγχρονο
    περιβάλλον ...</description>
    <cost>$24.95</cost>
    <isbn>$0.00</isbn>
  </book>

  ...
</bookListing>
```

Η διπλανή περιγραφή ενός καταλόγου βιβλίων γίνεται σε γλώσσα XML. Παρατηρήστε ότι δεν υπάρχει κάποια περιγραφή που να καθορίζει τον τρόπο εμφάνισης του περιεχομένου (π.χ. μέγεθος γραμματοσειράς, έντονοι ή πλάγιοι χαρακτήρες, κλπ) Αυτό μπορεί να καθορισθεί σε επόμενο στάδιο, κατά τη χρήση τεχνολογιών CSS και XSLT. Η XML διαχωρίζει την περιγραφή δομής (structure) από την εμφάνιση (appearance).

Εξετάζοντας το περιεχόμενο του παραπάνω αρχείου XML μπορούμε να αντιληφθούμε το περιεχόμενο και να διακρίνουμε μια **δομή** σ' αυτό.

Τι ακριβώς είναι μια δομή σ' ένα έγγραφο;

Η δομή ενός εγγράφου είναι ένα μοτίβο (pattern) το οποίο διαχωρίζει τα "συστατικά" ή δομικά στοιχεία του εγγράφου, όπως ενότητες, παράγραφοι, προτάσεις, απαριθμήσεις, σχήματα, κ.α. Το μοτίβο αυτό μπορεί να εφαρμόζεται ακόμα και σε επίπεδο πρότασης, αναδεικνύοντας σ' αυτήν λέξεις (π.χ. έντονες, πλάγιες, ιδιαίτερης σημασίας, κλπ).

Εξετάζοντας το περιεχόμενο του παραπάνω αρχείου XML μπορούμε εύκολα να αναγνωρίσουμε την δομή του και να καταλάβουμε ότι αφορά μια λίστα βιβλίων για κάθε ένα εκ των οποίων καταγράφεται ο τίτλος του, μια συνοπτική περιγραφή, η αξία του, ο αριθμός ταυτοποίησης ISBN, κ.α.

Θα μπορούσαμε να πούμε πως η ίδια η XML είναι αυτο-περιγραφική (self-describing) και εύκολα αναγνώσιμη από ανθρώπους αλλά και Η/Υ.

Οι Η/Υ "λατρεύουν" τα μοτίβα και είναι σε θέση να τα αναγνωρίσουν εύκολα και γρήγορα όταν ξέρουν πώς μοιάζουν. Είναι "πρωτάρηδες" όσον αφορά στα λογικά άλματα (inference).

Ας υποθέσουμε ότι ζητάμε από 2 άτομα να καταρτίσουν μια λίστα βιβλίων που διαθέτουν, ακολουθώντας το παραπάνω σχήμα περιγραφής, δηλαδή:

- η λίστα θα περιλαμβάνει μια σειρά καταγραφών (βιβλίων)
- για κάθε καταγραφή θα αναφέρονται τα εξής στοιχεία: ονομασία βιβλίου, συνοπτική περιγραφή, αξία και αριθμός ISBN.

Ακόμη κι αν ένας από τους δύο άλλαζε την σειρά καταγραφής των στοιχείων (π.χ. αποτυπώνεται η αξία πριν την συνοπτική περιγραφή) εμείς (αλλά και ένας Η/Υ) θα είμασταν σε θέση να εντοπίσουμε τα στοιχεία που χρειαζόμαστε ώστε να καταρτίσουμε μια ενιαία λίστα.

Πηγές / Παραπομπές



John Kane, Joseph R. Makuch, —1998— *"An overview of Electronic Publishing and Extensible Markup Language"*, *Journal of Applied Communications: Vol.82, issue 2*,
<https://doi.org/10.4148/1051-0834.21335>.



Thomas Myers, —2006— *"A Really, Really, Really Good Introduction to XML"*, book chapter in *"No Nonsense XML Web Development With PHP"*, ISBN 978-097-524-0205, SitePoint.