

# DSSSL και T<sub>E</sub>X

---

Απόστολος Συρόπουλος

28ης Οκτωβρίου 366

671 00 Ξάνθη

E-mail: [apostolo@obelix.ee.duth.gr](mailto:apostolo@obelix.ee.duth.gr)

## Abstract

We briefly describe the SGML and XML markup languages. Next, we briefly describe the DSSSL language. Then we demonstrate how one can typeset SGML/XML documents with T<sub>E</sub>X by transforming them to a file format that can be processed with the jadetex package. Moreover, we briefly discuss the necessary changes to jadetex in order to be able to work with Ω

## 1. Εισαγωγή

Η γλώσσα DSSSL (Document Style Semantics and Specification Language) αποτελεί ένα εργαλείο για την επεξεργασία κειμένων γραμμένα είτε σε SGML είτε σε XML.

Η SGML (Standard Generalized Markup Language) είναι μια γενική γλώσσα επισήμανσης η οποία σχεδιάστηκε ώστε να επιτρέπει την ανταλλαγή ηλεκτρονικών εγγράφων. Ως γλώσσα επισήμανσης η SGML επιτρέπει να ορίσουμε το συντακτικό και την σημασιολογία (semantics) μιας νέας γλώσσας επισήμανσης. Για παράδειγμα, στο κείμενο που περιγράφει το πρότυπο της γλώσσας HTML οι ορισμοί των διαφόρων ετικετών δίνονται σε SGML. Φυσικά, η SGML μπορεί να χρησιμοποιηθεί και για την προετοιμασία ηλεκτρονικών εντύπων (αυτός ήταν άλλωστε και ο λόγος της δημιουργίας της). Το σημαντικό είναι να κατανοήσει κάποιος ότι ένα αρχείο SGML δεν περιέχει πληροφορίες που αφορούν τη μορφοποίηση του κειμένου αλλά πληροφορίες καθαρά δομικού περιεχομένου. Συνήθως κάθε αρχείο SGML συνοδεύεται από ένα αρχείο DTD (Document Type Definition) στο οποίο ορίζουμε την σύνταξη που θα χρησιμοποιηθεί στο ηλεκτρονικό έγγραφο, δηλαδή ποιες ετικέτες θα μπορούν να χρησιμοποιηθούν. Σε περιπτώσεις απλών εγγράφων το αρχείο DTD μπορεί να είναι τμήμα του αρχείου SGML, δηλαδή ο

κώδικας του να εμφανίζεται στο αρχείο SGML. Δυστυχώς η SGML είναι φοβερά πολύπλοκη και έτσι δε μπορεί να χρησιμοποιηθεί για την ανταλλαγή εγγράφων μέσω του λεγομένου pagk'osmiou isto'u. Για τον σκοπό αυτό σχεδιάστηκε η XML. Για μια μη τεχνική εισαγωγή στην SGML μπορείτε να επισκεφθείτε τον δικτυακό τόπο: [http://www.arbortext.com/Think\\_Tank/SGML\\_Resources/Getting\\_Started\\_with\\_SGML/getting\\_started\\_with\\_sgml.html](http://www.arbortext.com/Think_Tank/SGML_Resources/Getting_Started_with_SGML/getting_started_with_sgml.html).

Η XML (eXtensible Markup Language) έχει σχεδιασθεί με σκοπό την ικανοποίηση των αναγκών εκείνων που ετοιμάζουν έγγραφα για διανομή μέσω του Δικτύου. Έτσι η XML παρέχει πολύ μεγαλύτερες δυνατότητες από αυτές που παρέχει η HTML. Η XML είναι αποτέλεσμα μιας προσπάθειας να γίνει εφικτή η διανομή εγγράφων σε SGML μέσω του Η XML αποτελεί ένα πολύ μικρό υποσύνολο της SGML και ικανοποιεί τους σχεδιαστικούς της στόχους, ενώ προσφέρει σε αυτούς που ετοιμάζουν σελίδες HTML και δεν ενδιαφέρονται για τα πλεονεκτήματα της SGML μια ευκαιρία να την εξατομικεύσουν και να προσθέσουν αποκλειστικά στοιχεία στην HTML. Συνήθως υπάρχουν 3 αρχεία τα οποία επεξεργάζεται μια εφαρμογή που μπορεί να προβάλει δεδομένα XML:

- Το αρχείο XML το οποίο περιέχει τα δεδομένα του εγγράφου το κείμενο και την λογική διάρθρωσή του, η οποία σημειώνεται με ετικέτες.
- Ένα αρχείο στυλ γραμμένο στην γλώσσα XSL (eXtensible Stylesheet Language). Το αρχείο στυλ υπαγορεύει το πως θα πρέπει να γίνει η μορφοποίηση των διαφόρων στοιχείων όταν αυτά προβάλλονται μέσω ενός προγράμματος φυλλομέτρησης ή ένα επεξεργαστή κειμένου. Φυσικά μπορούμε να εφαρμόσουμε διαφορετικά αρχεία στυλ στο ίδιο αρχείο XML επιτυγχάνοντας έτσι την αλλαγή της εμφάνισης με τον απλούστερο δυνατό τρόπο.
- Ένα αρχείο ορισμού του τύπου του εγγράφου (DTD) με λειτουργικότητα ανάλογη του αντίστοιχου αρχείου της SGML.

Για μια εισαγωγή στη γλώσσα XML μπορείτε να επισκεφθείτε το δικτυακό τόπο [http://pdbeam.uwaterloo.ca/~rlander/XML\\_Tutorial](http://pdbeam.uwaterloo.ca/~rlander/XML_Tutorial). Ειδικά προγράμματα επεξεργασίας αρχείων XSL και DSSSL χρησιμοποιούνται για τη μετατροπή αρχείων SGML και XML σε εκτυπώσιμη μορφή αλλά και μορφή κατανοητή από τα προγράμματα φυλλομέτρησης με ένα τελείως ελεγχόμενο και προβλέψιμο τρόπο.

Στο κείμενο αυτό παρουσιάζουμε τα βασικά χαρακτηριστικά της DSSSL, το πρόγραμμα μετατροπής OpenJade και το πακέτο jadetex για τη στοιχειοθεσία της εξόδου του προγράμματος OpenJade.

<pre> &lt;!DOCTYPE COOKBOOK [ &lt;!ELEMENT COOKBOOK    - - (ctitle, recipe*)&gt; &lt;!ELEMENT RECIPE      - - (title, ingredient*)&gt;  &lt;!ELEMENT CTITLE      - o (#PCDATA)&gt; &lt;!ELEMENT TITLE       - o (#PCDATA)&gt; &lt;!ELEMENT INGREDIENT  - o (#PCDATA)&gt; ]&gt;  &lt;COOKBOOK&gt;   &lt;CTITLE&gt;The Cookbook   &lt;RECIPE&gt;     &lt;TITLE&gt;Cake     &lt;INGREDIENT&gt;500g Flour     &lt;INGREDIENT&gt;200g Sugar     &lt;INGREDIENT&gt;300g Butter   &lt;/RECIPE&gt; &lt;/COOKBOOK&gt; </pre>	<pre> &lt;!doctype style-sheet PUBLIC "-//James Clark//DTD DSSSL Style Sheet//EN"&gt;  (element COOKBOOK   (make simple-page-sequence))  (element CTITLE   (make paragraph    font-size: 30pt    quadding: 'center')) ;alignment  (element TITLE   (make paragraph    font-size: 24pt))  (element INGREDIENT   (make paragraph    font-size: 12pt)) </pre>
Αρχείο SGML	Αρχείο DSSSL

Σχήμα 1: Ένα αρχείο SGML και ένα απλό αρχείο DSSSL για την επεξεργασία του.

## 2. Τι είναι η DSSSL;

Η DSSSL χρησιμοποιείται για τη μορφοποίηση είτε αρχείων SGML είτε αρχείων XML. Η διαδικασία μορφοποίησης επιτυγχάνεται με την εφαρμογή *κατασκευαστικών κανόνων* στα στοιχεία του αρχείου SGML/XML. Στο σχήμα 1 φαίνεται ο κώδικας ενός αρχείου SGML, το οποίο περιέχει και τις πληροφορίες του αρχείου DTD, και ο κώδικας ενός αρχείου DSSSL. Βλέπουμε ότι στο αρχείο SGML ορίζουμε την ετικέτα CTITLE και την χρησιμοποιούμε για σημειώσουμε τη λέξη The Cookbook. Αντίστοιχα στο αρχείο DSSSL ορίζουμε ότι τα περιεχόμενα αυτής της ετικέτας θα πρέπει να στοιχειοθετηθούν στο κέντρο με γραμματοσειρά μεγέθους 30 στιγμών. Ανάλογες 'ρυθμίσεις' υπάρχουν και για τις άλλες ετικέτες που ορίζονται στο αρχείο SGML. Το στοιχειοθετημένο αποτέλεσμα του αρχείου SGML με βάση τους συγκεκριμένους κανόνες του αρχείου DSSSL φαίνεται στο σχήμα 2. Στην αρχή της ενότητας αναφερθήκαμε στους κατασκευαστικούς κανόνες οι οποίοι επιτρέπουν την στοιχειοθεσία ενός αρχείου SGML/XML. Αλλά τι ακριβώς είναι ένας κατασκευαστικός κανόνας; Ένας κατασκευαστικός κανόνας μπορεί να αναφέρεται είτε σε όλο το έγγραφο είτε σε ένα μόνο στοιχείο. Στο αρχείο DSSSL του παραδείγματός μας έχουμε μόνο κανόνες της δεύτερης μορφής. Η γενική μορφή ενός κατασκευαστικού κανόνα είναι η εξής:

```

(όνομα-κανόνα (έκφραση-"ταϊριάσματος")
  (παράσταση-δράσης)
)
```

# The Cookbook

## Cake

500g Flour  
200g Sugar  
300g Butter

Σχήμα 2: Το στοιχειοθετημένο αποτέλεσμα του αρχείου SGML του σχήματος 1.

Η έκφραση-"ταιριάσματος" καθορίζει σε ποια ετικέτα θα εφαρμοσθεί ο συγκεκριμένος κανόνας, ενώ η παράσταση-δράσης τι ακριβώς θα πρέπει να συμβεί όταν συναντήσουμε μια ετικέτα αυτού του είδους. Για παράδειγμα, αν έχουμε το παρακάτω απόσπασμα αρχείου XML

```
<?xml version="1.0" ?>
<book>
  <section>Construction rules
    <header> Introduction</header>
  </section>
  <section>Construction languages
    <header>Introduction</header>
  </section>
</book>
```

Ο παρακάτω κανόνας θα εφαρμοσθεί στο στοιχείο header το οποίο μπορεί να βρίσκεται σε ένα στοιχείο section

```
(element (section header)
  (make paragraph
    font-family-name: "Helvetica"
    font-weight: 'bold
    font-posture: 'oblique
    (process-children)
  )
)
```

Στο συγκεκριμένο παράδειγμα η δράση που καθορίζεται σχετίζεται με το είδος της γραμματοσειράς που θα χρησιμοποιηθεί για την στοιχειοθεσία του

κειμένου. Πιο συγκεκριμένα το περιεχόμενο της ετικέτας header θα εμφανισθεί χρησιμοποιώντας την γραμματοσειρά Helvetica σε έντονη πλάγια μορφή. Περισσότερες πληροφορίες σχετικά με την DSSSL μπορείτε να βρείτε στους δικτυακούς τόπους <http://www.netfolder.com/DSSSL/MarkupRule.htm> και <http://www.jclark.com/dsssl/>. Ειδικά από το δεύτερο μπορείτε να 'κατεβάσετε' και του προτύπου ISO/IEC 10179:1996 που ορίζει την DSSSL.

### 3. Το προγράμματα Jade και OpenJade

Τα προγράμματα Jade και OpenJade αποτελούν υλοποιήσεις του προτύπου ISO/IEC 10179:1996 που ορίζει την γλώσσα DSSSL. Το πρώτο πρόγραμμα αποτελεί προσωπική δουλειά του James Clark, ενώ το δεύτερο βασίζεται στο πρώτο. Το πρόγραμμα OpenJade αναπτύχθηκε και συντηρείται από την ομάδα OpenJade. Το πρόγραμμα Jade μπορείτε να το προμηθευτείτε από τον δικτυακό τόπο <http://www.jclark.com/jade/>, ενώ το πρόγραμμα OpenJade μπορείτε να το προμηθευτείτε από το δικτυακό τόπο <http://openjade.sourceforge.net/>. Τέλος, οι λάτρεις... των λειτουργικών συστημάτων του Bill Gates μπορούν να προμηθευτούν υλοποιήσεις των προγραμμάτων αυτών από το δικτυακό τόπο <http://www.sscd.de/openjade/>. Τα προγράμματα Jade και OpenJade έχουν ταυτόσημη συμπεριφορά και εμείς εδώ θα αναφερθούμε μόνο στο OpenJade. Το πρόγραμμα OpenJade λειτουργεί από την γραμμή εντολών ακριβώς όπως και το T<sub>E</sub>X. Το πρόγραμμα δέχεται ως είσοδο είτε ένα αρχείο SGML είτε ένα αρχείο XML και το μετατρέπει σε μορφές αρχείων που συμπεριλαμβάνουν:

- Αναπαράσταση της ροής του αντικειμενικού δένδρου σε XML.
- Τη μορφή RTF για εκτύπωση μέσω των προγραμμάτων επεξεργασίας κειμένου της Microsoft, όπως το Word<sup>1</sup>.
- Τη μορφή T<sub>E</sub>X.
- Τη μορφή MIF για παραπέρα επεξεργασία με τον επεξεργαστή κειμένου Framemaker.
- Σε μορφές SGML και XML πράγμα που επιτρέπει τη μετατροπή αρχείων SGML/XML.

Το πρόγραμμα είναι εντελώς δωρεάν και το μόνο που ζητούν οι σχεδιαστές του συστήματος OpenJade σε περίπτωση εμπορικής χρήσης είναι να αναγράφεται μια σχετική ευχαριστία!!!

---

<sup>1</sup> Χωρίς να είμαι απόλυτα σίγουρος, πιστεύω ότι τα αρχεία RTF μπορεί να τα επεξεργαστεί και το σύστημα StarOffice που προσφέρει δωρεάν η Sun Microsystems.

Η χρήση του προγράμματος είναι πολύ απλή: πηγαίνουμε σε ένα κατάλογο που περιέχει τα απαραίτητα αρχεία DSSSL και δίνουμε την εντολή:

```
openjade demo.sgml
```

Το αποτέλεσμα είναι να δημιουργηθεί ένα αρχείο XML με όνομα `demo.fot`. Αν θέλουμε να δημιουργήσουμε ένα αρχείο  $\TeX$ , θα πρέπει να δώσουμε την εντολή

```
openjade -t tex demo.sgml
```

Η εντολή δημιουργεί το αρχείο `demo.tex`. Το ερώτημα που εύλογα τίθεται είναι: *Πως μπορούμε να διαχειριστούμε το παραγόμενο αρχείο  $\TeX$* ; Την απάντηση δίνει η επόμενη ενότητα.

#### 4. Η φόρμα jadetex

Η φόρμα jadetex σχεδιάστηκε από τον Sebastian Rahtz και επιτρέπει την στοιχειοθεσία αρχείων SGML/XML με το  $\TeX$ /Ωμέγα μέσω του  $\TeX$ /Ωμέγα. Ειδικά η δυνατότητα στοιχειοθεσίας με την χρήση του Ωμέγα προτάθηκε στον Sebastian Rahtz από τον γράφοντα ο οποίος μάλιστα υπέδειξε και τον τρόπο που αυτό θα έπρεπε να γίνει. Την φόρμα jadetex μπορείτε να το προμηθευτείτε από το δικτυακό τόπο <http://www.tug.org/applications/jadetex/>. Η φόρμα απαιτεί τη χρήση του πακέτου hyperref το οποίο μπορείτε να προμηθευτείτε από το δικτυακό τόπο <http://www.tug.org/applications/hyperref/>. Επειδή για άγνωστους λόγους το Ωμέγα δεν συνεργάζεται με το πακέτο hyperref, ο Rahtz δημιούργησε το αρχείο nohyperref, το οποίο αποτελεί τμήμα του πακέτου hyperref, που απενεργοποιεί όλες τις αναφορές στο πακέτο hyperref που υπάρχουν στο αρχείο jadetex.tex. Ένα άλλο πρόβλημα είναι ότι πρέπει να δημιουργηθεί το αρχείο ut1ptm.fd από το αρχείο ot1omlgs.fd αλλάζοντας όλα τα κέρματα OT1 σε UT1 και όλα τα κέρματα omlgs σε ptm. Αν έχετε εγκατεστημένη τη τελευταία έκδοση του Ωμέγα στον υπολογιστή σας, μπορείτε να προμηθευτείτε το αρχείο jadeomega.zip από τους δικτυακούς τόπους <ftp://ocean1.ee.duth.gr/pub/TeX/macros> και <ftp://obelix.ee.duth.gr/pub/TeX> και με αυτό να εγκαταστήσετε την έκδοση του jadetex για το Ωμέγα. Οδηγίες εγκατάστασης περιέχονται στο αρχείο. Στο σχήμα 3 φαίνεται το αρχείο SGML του σχήματος 1 και με ελληνικό κείμενο αλλά και το στοιχειοθετημένο αποτέλεσμα όπως δημιουργήθηκε από το πρόγραμμα OpenJade και τη φόρμα jadetex. Σημειώστε ότι σε περίπτωση που αλλάξετε το όνομα του αρχείου που περιέχει τον κώδικα SGML, θα πρέπει να αλλάξετε και το όνομα του αρχείου DSSSL.

<pre> &lt;!DOCTYPE COOKBOOK [ &lt;!ELEMENT COOKBOOK      - - (ctitle, recipe*)&gt; &lt;!ELEMENT RECIPE        - - (title, ingredient*)&gt; &lt;!ELEMENT CTITLE        - o (#PCDATA)&gt; &lt;!ELEMENT TITLE         - o (#PCDATA)&gt; &lt;!ELEMENT INGREDIENT    - o (#PCDATA)&gt; ] &gt;  &lt;COOKBOOK&gt;   &lt;CTITLE&gt;Ο Τσελεμεντές   &lt;RECIPE&gt;     &lt;TITLE&gt;Cake     &lt;INGREDIENT&gt;500g Αλεύρι     &lt;INGREDIENT&gt;200g Ζάχαρη     &lt;INGREDIENT&gt;300g Βούτυρο   &lt;/RECIPE&gt; &lt;/COOKBOOK&gt; </pre>	<h1>Ο Τσελεμεντές</h1>
<p><b>Cake</b>  500g Αλεύρι  200g Ζάχαρη  300g Βούτυρο</p>	

Σχήμα 3: Αρχείο SGML με ελληνικά και το στοιχειοθετημένο αποτέλεσμα που δημιουργήθηκε με τη νέα έκδοση του jadetex.

## 5. Συμπεράσματα

Σε τούτο το άρθρο παρουσιάσαμε τον τρόπο με τον οποίο μπορεί κανείς να στοιχειοθετήσει αρχεία SGML/XML που περιέχουν ελληνικό κείμενο χρησιμοποιώντας το Ωμέγα. Αυτό έγινε δυνατό με την τροποποίηση της φόρμας jadetex ώστε αυτή να μπορεί να δουλεύει και με το Ωμέγα. Ίσως αυτή να είναι η πρώτη εφαρμογή του Ωμέγα που δημιουργήθηκε από κάποιον εκτός των δύο σχεδιαστών του Ωμέγα, δηλαδή του Γιάννη Χαραλάμπους και του Τζων Πλές. Δυστυχώς η ελλιπής τεκμηρίωση (ελληνικά... documentation) του συστήματος να είναι και ένας βασικός λόγος που, δυστυχώς, ακόμη και μεταξύ ελλήνων δεν είναι ακόμη διαδεδομένο.