



Tilannekatsaus: XML ja (La)TEX

Timo Hellgren
VTT Tietopalvelu
timo.hellgren@vtt.fi



Mikä on TEX?

- TEX = Donald E. Knuthin tekemä ilmainen ladontajärjestelmä
- sopii hyvin akateemisiin tutkimuksiin
- soveltuu huonommin esim. mainoslehtisten tekemiseen



Mikä on LaTeX?

- LaTeX = alk. Leslie Lamportin kehittämä makropakkaus loppukäyttäjille
- Sisältää valmiita makroja yleisimmille julkaisutyypeille
- Dokumentti koodataan osittain rakenteisesti
- SGML vastineita esim. TEI



TEX ja XML

- TEXiä voidaan käyttää muuntamaan dokumentti XML:stä Postscriptiin tai PDF:ksi
 - XSLT:llä voidaan määritellä muunnos XML:stä TEX:n makrokielelle
 - Käytetään TEXiä XSL-parserina
- Entä muunnos (La)TEX:sta XML:ksi?



JadeTEX

- SGML+DSSSL → T_EX
- Toimii Jaden kanssa
- Sebastian Rahtzin tekemä L^AT_EX-makropakkaus
- CTAN:/macros/jadetex/



TEI2L^AT_EX

- TEI-Lite → L^AT_EX2_ε
- Jean Daniel Feketen Perl 5 -ohjelma
- <ftp://ftp.lri.fr/LRI/soft/ihtm/tei2latex-0.1c.tar.gz>
- Sopii niille, jotka osaavat perliä paremmin kuin XSLT-kieltä



SGMLS.pm

- David Megginsonin tekemä Perl 5 luokkakirjasto
- osaava voi kirjoittaa näillä Perl-rutiineilla skriptin, joka muuntaa James Clarkin SGMLS-parserin ulosannin vaikka LaTeX:ksi
- tarkoitettu SGML → TeX -muunnoksiin



TeXML

- IBM:n javasovellus, joka konvertoi TeXML.dtd-mukaiset dokumentit TeX:ksi
- <http://www.alphaworks.ibm.com/aw.nsf/techreqs/texml>
- .xml+TeXML+.xsl → [TeXML] → .tex
- vaatii IBM:n XML-parserin ja Java-ympäristön



XML_TEX

- T_EX ei-validoivana XML-parserina
- David Carlislen tekemä makropakkaus
- T_EXin voidaan antaa tehdä ulkoasun muotoilut
- Voidaan käyttää Unicode-merkistöä!



XML_TEX:n käyttö 1

- XML_TEX-makrot ladataan lennosta
- tarvitaan .tex-tiedosto, jossa viitataan XML-dokumenttiin

```
\def\xmlfile{document.xml}
\input xmltex.tex
```
- käynnistys: latex document



XMLTEX:n käyttö 2 ja 3

- formaattina LaTeXin päällä
- omana formaattinaan ilman LaTeXia
- käynnistys: `xmltex document.xml`
- jälkimmäinen sopii guruille, jotka haluavat virittää oman systeeminsä



XMLTEX:n käyttö 4

- tehdään formaatti pdfinitexillä
- nyt lopputulos on suoraan pdf-tiedosto
- käyttö esim. `pdfxmltex document.xml`



XMLTEX:n toiminta (xmt-tiedostot)

- XML-tunnisteet määritellään T_EXin makrokielellä .xmt-tiedostoihin
- `\XMLelement{emph}{}{\itshape}{}{}`
- `\XMLelement{table}{}{
 {\begin{tabular}{*{20}{1}}}
 {\end{tabular}}}`
- `\XMLentity{TeX}{\TeX}`



XMLTEX:n toiminta (catalogue)

- `xmltex.cfg` tiedosto sisältää viittaukset xmt-tiedostoihin, esim.
`\SYSTEM {http://www.oucs.ox.ac.uk/
dtds/tei-oucs.dtd} {tei.xmt} tai
\NAME{html} {html.xmt}`
- XML-dokumentissa on esim. määritelty
`<!DOCTYPE TEI.2 SYSTEM
"http://www.oucs.ox.ac.uk/
dtds/tei-oucs.dtd" [`



XML_{TE}X

- Tuki TEI ja MathML määritysten mukaisille XML-tiedostoille
- CTAN:/macros/xmltex/base/
- .xmt-määritysten teko vaatii LaTeXin makrojen hyvää tuntemusta
- Sopii niille, jotka tuntevat paremmin LaTeXia kuin XSLT:tä



Passive_{TE}X

- Makroja XML_{TE}Xille, jotta TEI-DTD:n mukaisista XML-tiedostoista saadaan PDF- tai PostScript-tiedostoja
- Tarvitsee lisäksi James Clarkin xt:n
- .xml+.xsl → [xt] → .fo
- .fo → [XML_{TE}X] → .dvi → [dvips] → .ps
- .fo → [pdfXML_{TE}X] → .pdf



PassiveTEX

- Toistaiseksi vasta prototyyppiasteella
- <http://users.ox.ac.uk/~rahtz/passivetex/>
- käyttö ei vaadi TEXin tuntemusta
- TEI sopinee useimmille (entä Docbook?)
- XSL-fo:lla helpompi tehdä omia ulkoasumäärittäyksiä kuin LaTeX:lla?



TEX4ht

- Eitan Gurarin muunnossysteemi
- LaTeX → (X)HTML/XML
- Matematiikka voidaan esittää grafiikkana tai MathML:nä
- <http://www.cis.ohio-state.edu/~gurari/TeX4ht/mn.html>



TEX4ht:n toiminta (XML)

- `\usepackage[file ,html]{tex4ht}`
dokumentissa ja käynnistys:
`htlatex file.tex`
- Lukee XML-tagien määrytykset *file.cfg*-tiedostosta
- Lopputulos *file.html*-tiedosto
- Lisäparametrit *.cfg*-tiedossa:
`\Preamble{0.0,edit,hooks}`



TEX4ht ja XML-määrytykset

- Määrittely esim:
`\Configure{section} {\Tg<section>}`
`{\Tg</section>} {\Tg<title>}`
`{\Tg</title>}`
- LaTeX: `\section{otsikko}`
Tekstiä tähän.
- Lopputulos: `<section><title>Otsikko`
`</title>Tekstiä tähän.</section>`



Yhteenveto

- Työkaluja on (osa vähän keskeneräisiä)
- Käyttäjän tehtävä itse määrittäykset omille XML-dokumenttityypeilleen

"Go forth now and create masterpieces of XML publishing!"

Donald E. Knuthia mukailleen