

L^AT_EX2HTML ervaringen

Van Handleiding in L^AT_EX tot Hulp Module op het Internet

Arno Kemperman

Laboratorium voor Analytische Chemie (LAC), Katholieke Universiteit Nijmegen (KUN)
gpstud1@sci.kun.nl

Abstract

Welkom op de digitale snelweg! Steeds meer informatie wordt er aangeboden op het Internet, en met het uitgroeien van het Internet groeit de onoverzichtelijkheid. In dit artikel wordt een toepassing van L^AT_EX2HTML beschreven waarbij een L^AT_EX document wordt omgezet naar een reeks goed gestructureerde pagina's op het World-Wide Web. Deze toepassing betreft een document die een handleiding is voor Internet-gebruikers (voornamelijk voor medewerkers van Laboratorium voor Analytische Chemie, KUN, Nijmegen). Er worden wat voordelen en voorbeelden van het L^AT_EX2HTML gebruik gegeven en hoe je het programma kunt aanpassen naar je eigen wensen. Bekijk de L^AT_EX documenten ook eens op deze moderne manier en blader eens met de muis door je tekst, getransformeerd tot een digitaal wonder op je scherm. Oftewel: L^AT_EX2HTML2!

1 Inleiding

De mens heeft altijd behoefte aan communicatie gehad: het uitwisselen van informatie. Een wereld zonder communicatie zou eenvoudigweg niet mogelijk zijn. Steeds weer zoeken we dan ook naar nieuwe en snellere vormen van communicatie om het welzijn van de mens te verhogen. Neem bijvoorbeeld de ontwikkeling van de televisie. De televisie heeft toch wel een cultuurschok teweeg gebracht. Mensen worden graag geïnformeerd over alles wat ze interesseert. De televisie geeft echter niet altijd de gewenste informatie. Dit komt met name doordat de televisie niet interactief is. Je moet het doen met wat je wordt voorgeschoteld.

2 Het Internet en World-Wide Web

Het Internet is zo'n medium waarmee het interactief uitwisselen van informatie mogelijk is. Het Internet is een wereldomspannend netwerk gebruik makend van allerlei standaarden en protocollen om informatie te transporteren tussen enorm veel computersystemen. World-Wide Web (WWW) maakt gebruik van het Internet. Een officiële beschrijving van het WWW project is: 'wereldwijde hypermedia informatie-uitwisseling met het doel een universele toegang te verlenen tot een groot aantal documenten'. De informatie wordt verstuurd via het HyperText Transfer Protocol (HTTP). De documenten zijn geschreven in HyperText Markup Language (HTML).

Het HTTP protocol zorgt voor de snelle communicatie tussen de verschillende computersystemen; m.a.w. tussen vraag (client) en aanbod (de WWW server). Als client gebruik je een zogenaamde Web browser (Mosaic, Netscape) die de informatie toont op het beeldscherm. Het protocol zorgt voor een vrije interpretatie van de HTML-code voor

de verschillende browsers, zoals Netscape, voor de MAC, PC en UNIX omgevingen. De exacte opbouw voor de verschillende systemen zal dus verschillen.

Het bijzondere gebruik van (HTML) hypertext door WWW is met name de mogelijkheid die wordt geboden om een verwijzing te maken naar een document welke zich op een andere computer bevindt. Via de browser is zo'n verwijzing meestal zichtbaar doordat het onderstreept is. Aanklikken met de muis brengt je dan naar het gewenste document ergens op een computer in de wereld. Elk document heeft zijn eigen specifieke URL-naam. Een URL zorgt dus eigenlijk voor een unieke naam voor een WWW-pagina waardoor de browser dit document kan vinden op het Internet. Zo'n URL is opgebouwd uit een aantal componenten:

```
<protocol>://<server>/<locatie>
```

<protocol> staat voor het soort protocol voor het verzenden van de informatie (HTTP kan ook bijvoorbeeld FTP aan)

<server> staat voor een 'knooppunt' op het net: een server-naam waarmee het computersysteem, waarop het gevraagde document zich bevindt, kan worden gevonden.

<locatie> geeft de vindplaats van het document aan op de computer van de server.

Er kan verwezen worden naar andere documenten, ASCII-teksten, gopher- en FTP-adressen, nieuwsgroepen, geluid, filmpjes en plaatjes. Naast deze 'links' biedt WWW nog meer mogelijkheden zoals het opnemen van plaatjes in een pagina, invulformulieren, opsommingen, een nette lay-out met verschillende letterstijlen en sinds kort formules en vergelijkingen, tabellen, figuren met tekst er omheen, achtergrond maken en het veranderen van de kleuren van de tekst. Dit maakt het WWW systeem een interessant en aantrekkelijk medium om informatie te zoeken waardoor het

in korte tijd uitgegroeid is tot de meest gebruikte standaard op de digitale snelweg.

3 L^AT_EX2HTML

Er zijn een aantal programma's beschikbaar die documenten, gemaakt met je favoriete tekstverwerker, kunnen omzetten naar HTML. Hoewel de HTML-code zeer eenvoudig is (er kan gemakkelijk met behulp van een simpele editor een WWW-pagina worden gemaakt), is het soms toch de moeite waard zo'n programma te gebruiken. L^AT_EX2HTML is zo'n programma wat L^AT_EX documenten rechtstreeks kan omzetten naar HTML-documenten die gelezen kunnen worden door de verschillende Web browsers. Het programma is geschreven door Nikos Drakos, Computer Based Learning Unit, University of Leeds.¹

Laat ik wat redenen geven om L^AT_EX2HTML te gebruiken:

- Een document is direct beschikbaar te maken op het Internet via World-Wide Web.
- Hierdoor kan het tijdswinst opleveren.
- Je hoeft de HTML-codes niet te kennen om toch een HTML pagina te maken.
- Het brengt structuren aan in een document zodat bijvoorbeeld stukken tekst snel te vinden zijn ('links' naar hoofdstukken bijvoorbeeld, via een inhoudsopgave of de woorden in de index zijn aanklikbaar).
- Het zet zo veel mogelijk automatisch om naar het HTML formaat: de stijl (bijvoorbeeld grootte), inhoudsopgave, index, formules, voetnoten en speciale tekens.
- Het is mogelijk (hyper)links te maken naar andere adressen/documenten. Zo wordt een document wel erg dynamisch: je referenties zijn bijvoorbeeld direct bereikbaar door simpel klikken met de muis.
- Je kunt een document nog steeds uitprinten in je ouwe vertrouwde L^AT_EX lay-out.

4 Een Toepassing, Inleiding en Aanleiding

Begin 1994 ging ik stage lopen voor mijn scheikunde studie op de afdeling Laboratorium voor Analytische Chemie (LAC), Katholieke Universiteit Nijmegen. In die tijd was Mosaic de enige browser die werd gebruikt op de afdeling. Tegenwoordig wordt voornamelijk Netscape gebruikt. In het begin van het gebruik van Mosaic werd het eigenlijk vaker gebruikt voor minder nuttige doeleinden (zoals Interzappen of Web-surfing; Interzappen is voor mij 'het rondzwerven op het Internet op zoek naar nuttige of minder nuttige informatie en soms niet meer weten hoe je er bent gekomen'). Toen was nog niet duidelijk dat het Internet ook zeer nuttige informatie voor voor het wetenschappelijk onderzoek kon opleveren. Denk bijvoorbeeld aan artikelen en informatie over mailing lijsten, Frequently Asked Questions (FAQ's), congressen, nieuwsgroepen en FTP-adressen.

Langzamerhand werden de verschillende toepassingsmogelijkheden van het World-Wide Web ook door de afdeling ontdekt en ontstond de vraag of het LAC zich ook niet moest kunnen presenteren op het Internet. Het LAC kon zo bereikbaar zijn voor onderzoekers in hetzelfde vakgebied en informatie beschikbaar stellen in de vorm van informatie over medewerkers, hun onderzoek en eerder gepubliceerde artikelen. Begin 1995 werden er een aantal WWW-pagina's geschreven die in april '95 officieel op het Internet zijn gepresenteerd.²

Mijn bijdrage was voornamelijk het inventariseren en structureren van allerlei verwijzingen naar nuttige adressen op het Internet, met name naar informatie over onderwerpen op de afdeling. Deze onderwerpen zijn voornamelijk chemometrie, kunstmatige intelligentie, neurale netwerken en genetische algoritmen. Ook wordt in een pagina vele verwijzingen gegeven naar informatie over het Internet en WWW zelf, zoals allerlei handleidingen en interessante startpunten. Het opzetten van deze pagina's vereist natuurlijk wel wat kennis over het Internet en hoe en waar bepaalde informatie te vinden is. De opgedane kennis wilde ik in een verslag zetten zodat dit als een handleiding kan gaan dienen voor medewerkers van de afdeling om snel informatie te kunnen vinden over allerlei onderwerpen, en verwijzingen naar WWW-pagina's waar eventueel nog meer informatie over het gewenste onderwerp te vinden is.

De handleiding wilde ik schrijven in L^AT_EX omdat het een programma was waardoor ik een simpele editor (xedit) in UNIX X-Windows kon gebruiken om mijn teksten te schrijven (handig als je veel 'verft' met je muis, het kopiëren van lappen tekst, en met name URL-adressen die toch vaak erg lang en ingewikkeld zijn). Bovendien wilde ik er gewoon wel eens mee leren werken. Met de juiste L^AT_EX commando's is een fraaie lay-out met een mooi lettertype zo gemaakt. Tijdens een van mijn ritten op de digitale snelweg kwam ik in een document een verwijzing naar het L^AT_EX2HTML programma tegen.

Dit programma converteert documenten naar het formaat dat geschikt is voor WWW en dit leek me wel geschikt om mijn scriptie om te zetten naar HTML. Wat is mooier dan een handleiding over het Internet op het Internet zelf! Het leek me ideaal om alle verwijzingen naar het Internet aanklikbaar te kunnen maken.

Na het overhalen van het programma via ftp bleek tijdens het installeren dat het programma wel wat andere programma's nodig had om te kunnen werken zoals `perl`, `dvips` en `gs`. Tijdens het zoeken naar deze programma's op het netwerk van de universiteit bleek L^AT_EX2HTML er zelf al op te staan! Een eerste poging om een L^AT_EX tekst om te zetten was toen snel gedaan en het resultaat was al verbluffend. Ook een aantal medewerkers op de afdeling zijn nu enthousiast over dit programma. Ik heb wel wat mogelijkheden gebruikt om het programma aan te passen. Deze veranderingen waren simpel uit te voeren door de confi-

¹Zie ook: <http://cbl.leeds.ac.uk/nikos/tex2html/doc/latex2html/latex2html.html>

²<http://www.sci.kun.nl/cac/www-home.html>

guratiefile `.latex2html-init` aan te passen die met het programma wordt meegeleverd. Nikos Drakos heeft bij het programma een prima handleiding geschreven waarmee alle mogelijkheden duidelijk worden uitgelegd. Een aantal toepassingen en aanpassingen aan het programma heeft er uiteindelijk voor gezorgd dat mijn L^AT_EX documenten worden omgezet naar HTML, aangepast op mijn wensen, terwijl ik het ook gewoon netjes kan uitprinten in L^AT_EX.

5 Gebruik en Aanpassing L^AT_EX2HTML Versie 95

Met het programma L^AT_EX2HTML wordt een voorbeeld configuratiefile `.latex2html-init` (`dot.latex2html-init`) meegeleverd om je eigen instellingen te veranderen. Hieronder volgen de veranderingen en waarom ik die heb gemaakt.

dot.latex2html-init

1. `$MAX_SPLIT_DEPTH = 8;`
`# Stop making separate files at this`
`#` `depth`
 Veranderd van 8 naar 2 omdat ik overzicht wilde behouden in mijn HTML versie van mijn L^AT_EX document. Een maximale diepte van 2 zorgt ervoor dat alle `\subsection` delen in 1 HTML pagina kwamen. De hoeveelheid tekst in een subsection vond ik niet groot genoeg om deze in een apart bestand te zetten. Het geheel wordt leesbaarder. In mijn geval wordt een hele `\section` in één keer leesbaar.
2. `$MAX_LINK_DEPTH = 4;`
`# Stop showing child nodes at this depth`
 Veranderd naar 2. Ook dit om niet onnodig te laten verwijzen naar allerlei sub-paragraafjes, dat maakt het niet overzichtelijker.
3. `$TITLE = "Het Internet";`
 Titel van het document en de startpagina in HTML.
4. `$AUTO_NAVIGATION = 1;`
 Niet veranderd, de navigation gebruik ik om te 'bladeren' in het HTML document naar een volgende en vorige bladzijde etc. Ik heb wel eigen icoontjes gebruikt en de volgorde veranderd.
5. `$CHILDLINE = "
 <HR>\n";`
 veranderd naar:
`$CHILDLINE = "
 <IMG SRC=`
`blauwe_lijn.gif>
";`
 Geeft dus een lijn (gif-plaatje) in het HTML document in plaats van een gewone lijn (`<HR>`).
6. `$WORDS_IN_PAGE = 1000;`
 Naar 300. Als er meer woorden zijn dan komt onderaan het document ook navigatie. Wel handig omdat je dan niet hoeft terug te scrollen om op het icoontje 'volgende bladzijde' te klikken.
7. `$default_language = 'english';`
 naar:
`$default_language = 'dutch';`
 Het verslag is in het Nederlands.
8. `$WORDS_IN_NAVIGATION_PANEL_TITLES = 4;`
 Veranderd naar 10. Oorspronkelijk staat de tekstge-

stuurde navigatie bovenaan de pagina naast elkaar. Hij kapt dan bijvoorbeeld een titel van een hoofdstuk af na 4 woorden, maar aangezien ik het onder elkaar wilde hebben is het afkappen niet echt nodig.

9. `$EXTERNAL_UP_LINK =`
`"http://www.sci.kun.nl/cac/www-home.html";`
`$EXTERNAL_UP_TITLE =`
`"Terug naar de LAC-page";`
 Deze gewoon ingevuld. Geeft een link op de eerste HTML pagina zodat je terug kunt verwijzen naar het document waarop verwezen werd naar het L^AT_EX2HTML document (waar je dus vandaan kwam).
10. `$NETSCAPE_HTML = 1;`
 Op 1 gezet, blijkt iets mooier te zijn indien je met Netscape werkt.
11. `$TITLES_LANGUAGE = "dutch";`
 Naar dutch gezet dus.
12. `sub dutch_titles {`
`$toc_title = "Inhoudsopgave";`
`$lof_title = "Lijst Figuren";`
`$lot_title = "Lijst Tabellen";`
`$idx_title = "Index";`
`$bib_title = "Referenties";`
`$info_title = "Over dit dokument...";`
`}`
 Dit toegevoegd. In de handleiding bij L^AT_EX2HTML wordt een voorbeeld gegeven hoe je dit kon aanpassen naar de gewenste taal. Dit voorbeeld overgenomen en veranderd naar het Nederlands.
13. **Navigation Panel**
`# Start with a horizontal rule (3-d`
`#` `dividing line)`
`"
 <HR>".`
 Naar:
`# Start with a horizontal rule (3-d`
`#` `dividing line)`
`"

".`
 Mooiere blauwe lijn ingevoegd ipv een gewone lijn.
14. `# Now add a few buttons with a space`
`#` `between them`
`"$NEXT $UP $PREVIOUS $CONTENTS $INDEX`
`$CUSTOM_BUTTONS" .`
 Naar:
`# Now add a few buttons with a space`
`#` `between them`
`"$PREVIOUS $UP $NEXT $CONTENTS $INDEX`
`$CUSTOM_BUTTONS" .`
 De volgorde veranderd! Een wat nettere lay-out omdat ik pijlen gebruik in het navigatiepaneel. Door de volgorde `$PREVIOUS $UP $NEXT` te nemen krijg je de pijlvolgorde: links – boven – rechts, deze verwijzen naar respectievelijk: vorige pagina — 'omhoog in document structuur' — volgende pagina. Er werden geen `$CUSTOM_BUTTONS` gebruikt.
15. `# If ``next`` section exists, add its title`
`#` `to the navigation panel`
`($NEXT_TITLE ? " Next: $NEXT_TITLE\n"`
`: undef) .`
`# Similarly with the ``up`` title ...`
`($UP_TITLE ? "Up: $UP_TITLE\n"`
`: undef) .`
`#... and the ``previous`` title`

```
($PREVIOUS_TITLE ? "<B> Previous:</B>
$PREVIOUS_TITLE\n" : undef) .
Veranderd naar:
# ... and the 'previous' title
($PREVIOUS_TITLE ? "<BR> <B> vorige blz:</B>
$PREVIOUS_TITLE <BR>" : undef) .

# Similarly with the 'up' title ...
($UP_TITLE ? "<B> op:</B> $UP_TITLE <BR> "
: undef) .
```

```
# If 'next' section exists, add its
# title to the navigation panel
($NEXT_TITLE ? "<B> volgende blz:</B>
$NEXT_TITLE <BR>" : undef) .
```

Hier de taal aangepast voor de tekstgestuurde navigatie. Door de \n te vervangen door
 komen de verwijzingen onder elkaar te staan wat ik mooier vond.

```
16. # Line Break, horizontal rule (3-d
# dividing line) and new paragraph
"<BR> <HR> <P>\n"
```

```
# Line Break, horizontal rule (3-d
# dividing line) and new paragraph
"<BR> <IMG SRC=blauwe_lijn.gif> <P>\n"
Ook hier een mooie scheidinglijn aangebracht.
```

```
17. %icons =
(
'cross_ref_visible_mark' , 'd_hand.gif' ,
'anchor_mark' , '&#160;' ,
'anchor_invisible_mark' , '&#160;' ,
'up_visible_mark' , 'cyuarrw.gif' ,
'next_visible_mark' , 'cyrarrw.gif' ,
'previous_visible_mark' , 'cylarrw.gif' ,
'next_page_visible_mark' , 'cyrarrw.gif' ,
'previous_page_visible_mark' , 'cylarrw.gif' ,
'contents_visible_mark' , 'type_wr.gif' ,
'index_visible_mark' , 'note01.gif' ,
'footnote_mark' , 'd_hand.gif' ,
'up_inactive_visible_mark' , 'cyuarrw.gif' ,
'next_inactive_visible_mark' , 'cyrarrw.gif' ,
'previous_inactive_visible_mark' , 'cylarrw.gif' ,
'next_page_inactive_visible_mark' ,
'cyrarrw.gif' ,
'previous_page_inactive_visible_mark' ,
'cylarrw.gif' ,
'change_begin_visible_mark' , 'yl_ball.gif' ,
'change_end_visible_mark' , 'yl_ball.gif' ,
'change_delete_visible_mark' , 'yl_ball.gif'
);
```

Aangemaakt in de configuratiefile aangezien ik andere icoontjes wilde gebruiken. Er zijn ook inactieve icoontjes. Deze worden gebruikt als er bijvoorbeeld geen pagina 'terug' is. Voor de index en inhoudsopgave zijn ook aparte icoontjes gekozen.

```
18. Om de icoontjes te laten zien, moet worden veranderd:
$ICONSERVENR = 'http://ac7.sci.kun.nl:
6666/latex/gif';
```

6 Gebruik L^AT_EX2HTML opties

Algemeen:

```
latex2html [optie] latex.tex
```

Veel van de L^AT_EX2HTML opties kunnen ook in de configuratiefile .latex2html-init worden geregeld zoals:

```
-t NAAM          naam van het document
-no_navigation 0  geen navigatie (of 1: wel
navigatie)
```

De enige optie die ik gebruik heb (dus buiten de configuratie-file) is:

```
-dir directory  directory waar de HTML
bestanden moeten worden
geïnstalleerd
```

Voor meer van deze opties verwijst ik weer naar de handleiding van Nikos Drakos.

7 Gebruik Extra Commando's in L^AT_EX

Voor L^AT_EX is een extra .sty bestand beschikbaar: html.sty. Deze bevat informatie over extra toe te voegen commando's met betrekking tot het toevoegen van extra HTML mogelijkheden. Een aantal daarvan heb ik gebruikt:

1. Toevoegen html.sty:

```
\documentstyle[dutch,epsf,html]{article}
```
2. De hyperlink 'Tekst' wordt onderstreept en aanklikbaar in HTML en als een voetnoot wordt

```
http://lycos.cs.cmu.edu/ afgedrukt in LATEX:
```

```
\htmladdnormallinkfoot{Tekst}{http://}
lycos.cs.cmu.edu/
```

3. Hetzelfde als hierboven, alleen wordt nu geen voetnoot gegeven:

```
\htmladdnormallink{Tekst}{http:}
```

4. Plaatje invoegen in HTML door verwijzing naar een locatie (URL), wel onhandig als je het plaatje ook wil gebruiken voor L^AT_EX. Dan moet je echt de normale L^AT_EX commando's gebruiken en bijvoorbeeld het plaatje als PostScript in je directory zetten (L^AT_EX2HTML zet deze netjes in je WWW pagina met onderschrift etc.):

```
\htmladdimg{URL}
```

5. Begin en einde 'platte' html-code, handig als je hele stukken HTML-code hebt. Deze worden gewoon in de WWW pagina tussengevoegd. Heb ik gebruikt om hier en daar wat plaatjes toe te voegen voor een leuke lay-out. L^AT_EX zelf doet niets met deze code:

```
\begin{rawhtml}
\end{rawhtml}
```

6. Begin 'platte' L^AT_EX-code. Deze wordt niet omgezet naar HTML, wel naar L^AT_EX uiteraart:

```
\begin{latexonly}
\end{latexonly}
```

7. Hyperlink naar een referentie:

```
\htmlref{tekst}{fig:example}
```

8. Eigen button naar een eigen locatie, dus invullen:

```
\htmladdtonavigation
{\htmladdnormallink
{\htmladdimg{http:?.gif}
}
{http:}}
```



Figuur 1: Voorbeeld van een L^AT_EX2HTML pagina layout. Bovenaan het navigatiepaneel. De pijlen verwijzen naar een volgende of vorige bladzijde of naar een niveau hoger in het document. Daaronder de aanklikbare titels.

2.1.1 Er was eens...

Er was eens..., een voorstel. In maart 1989 werd het eerste [voorstel](#) [3] voor het WWW project geschreven door [Tim Berners-Lee](#), werkzaam bij [CERN](#) [2] (zie ook [History to date](#), [CERN](#) [5]). Dit leidde in november 1990 tot een voorstel voor een universeel hypertext- systeem met het doel effectief research- onderwerpen en ideeën te transporteren door de organisatie. Effectieve communicatie was jaren lang een doel voor CERN aangezien veel medewerkers in veel verschillende landen zaten.

Figuur 2: Voorbeeld van een stuk tekst in HTML waarbij de onderstreepte woorden links zijn naar referentiedocumenten. Deze links zijn in L^AT_EX aangemaakt door de speciale commando's in `html.sty`.

8 Slot

L^AT_EX2HTML gebruik ik nu dagelijks.³⁴ Zo kan ik zien hoe mijn document er in hypertext-vorm uit ziet en probeer ik het zo gebruikersvriendelijk mogelijk te laten zijn. Doordat een aantal hoofdstukken een eigen URL hebben is het nu mogelijk te verwijzen naar dit hoofdstuk als een hulp-optie bij de verschillende onderwerpen op de LAC pagina op het Internet.

Ik wil ook nog wijzen op de mogelijkheden die L^AT_EX2HTML biedt voor niet-Internetters. Met de verschillende browsers kunnen ook bestanden beke-

ken worden via de zogenaamde localhost (via een `file://localhost/..` of `file://..` URL). Dit geeft de mogelijkheid om je document zelf door te lezen. Als een soort luxe alternatief om je L^AT_EX document door te nemen. Met name de mogelijkheid dat je hoofdstukken direkt aan kunt klikken, werkt naar mijn mening zeer prettig en geeft je een prima overzicht over je document (vooral als dat zeer groot is). Ik gebruik het zelf nu om mijn teksten makkelijk na te kunnen lezen zonder ik met mijn ogen een weg moet kunnen banen tussen al die L^AT_EX commando's (het is daarnaast ook mooier dan bijvoorbeeld Ghostview). Al met al heb ik hierboven veel veranderingen aangegeven, maar het beste is gewoon om het zelf eens uit te proberen: trial and error!

³L^AT_EX2HTML pagina: <http://www.sci.kun.nl/cac/latex/scriptie/scriptie.html>

⁴De LAC Home Page: <http://www.sci.kun.nl/cac/www-home.html>