

Gestió i edició de continguts

Antoni Oliver

PID_00159194

Temps de lectura i comprensió: **4 hores**



Índex

Introducció	5
Objectius	6
1. Editors i processadors de textos	7
2. Organització del document i format final	8
2.1. Processadors de textos: estils i plantilles	8
2.1.1. Estils	9
2.1.2. Plantilles	10
2.2. El format LaTeX	11
2.2.1. Requeriments per a fer servir LaTeX	12
2.2.2. Les principals instruccions de LaTeX	12
2.2.3. Processament de documents en LaTeX	14
2.2.4. Conclusions	15
2.3. El format DocBook	16
2.3.1. Requisits per a fer servir DocBook	16
2.3.2. Les principals etiquetes de DocBook	16
2.3.3. Processament de documents en DocBook	18
2.3.4. Conclusions	19
3. Edició col·laborativa	21
3.1. Edició col·laborativa amb Open Office	21
3.1.1. Control de canvis	21
3.1.2. Comparació de documents	24
3.1.3. Control de versions	24
3.2. Edició col·laborativa amb Google Docs	25
3.2.1. Creació i edició de documents	25
3.2.2. Compartir documents i edició col·laborativa	26
3.2.3. Publicar i imprimir documents	27
3.2.4. Altres eines interessants	28
4. Gestió de la bibliografia	29
4.1. Bibliografies amb OpenOffice Writer	29
4.2. BibTeX: bibliografia per a LaTeX	33
4.3. Bibliografia en DocBook	38
4.4. Zotero	40
5. Bases de dades documentals i Google Desktop	44
6. Sistemes de gestió de continguts	47

6.1. Gestors de contingut web	47
6.2. Blogs	47
6.3. Wikis	48
6.4. Knols	49
Resum	50
Bibliografia	51

Introducció

En aquest mòdul estudiarem tècniques i eines relacionades amb l'edició i la gestió de continguts. El que pretén aquest mòdul és presentar alguns aspectes relacionats amb la generació, organització, publicació, recuperació i reutilització de continguts generats per nosaltres mateixos, ja sigui en solitari o de manera col·laborativa. Les noves tecnologies, i algunes que no són tan noves, ens ofereixen un gran ventall d'aplicacions per agilitzar la generació de documents i noves formes de publicar-los i compartir-los.

Molts dels aspectes que estudiarem en aquest mòdul seran ja coneguts, almenys de manera parcial, però els principis metodològics i les eines que presentarem seran sense dubte d'utilitat i aplicació per a la generació i publicació dels nostres treballs.

En tot el mòdul presentarem eines de programari lliure, per facilitar-hi l'accés, tot i que hi ha altres eines comercials que proporcionen les mateixes utilitats.

Des del moment de la redacció d'aquest mòdul fins que arribi a les mans del lector de ben segur que hauran aparegut noves eines i utilitats, però els principis bàsics que presentem romandran pràcticament inalterats.

La intenció de l'autor és que la informació d'aquest mòdul sigui d'utilitat i d'aplicació tant en les tasques acadèmiques com en la vida personal i professional.

Objectius

Els objectius que assolireu amb els materials que componen aquest mòdul són els següents:

- 1.** Conèixer les principals eines que ens permeten editar de manera àgil un document i obtenir sortides de gran qualitat.
- 2.** Conèixer les principals eines que ens permeten editar un document de manera col·laborativa.
- 3.** Saber gestionar de manera eficient bases de dades bibliogràfiques i generar de manera automàtica la bibliografia d'un document.
- 4.** Saber què són i per a què serveixen les bases de dades documentals.
- 5.** Conèixer les principals característiques dels sistemes de gestió de continguts.
- 6.** Valorar l'ús d'aplicacions de programari lliure.

1. Editors i processadors de textos

Actualment la majoria de documents que redactem els fem amb l'ús d'un ordinador. Tots nosaltres estem ja bastant habituats a treballar amb eines informàtiques que ens permeten escriure i emmagatzemar documents. Una primera distinció que cal tenir present pel que fa a aquest tipus d'eines és la diferència que hi ha entre un editor de textos i un processador de textos.

Els **editors de textos** són programes que permeten editar fitxers de text, és a dir, fitxers que contenen només text, sense cap tipus de format (tipus i mides i colors de lletra, per exemple). Com veurem més endavant, hi ha uns formats específics de documents que expressen el format mitjançant text, i que per tant sí que podran editar-se mitjançant editors de textos).

Els **processadors de textos** són programes que permeten editar documents afegint format (tipus de lletra, mides, colors, afegir imatges, inserir taules, etc.). Aquests documents es guarden en uns arxius d'algun tipus determinat que contenen tant la informació textual, com la informació de format.

Sovint es confonen aquests dos tipus d'aplicacions. De fet, un processador de textos pot editar documents de text i, per tant, funcionar com si fos un editor de text. Tanmateix, un editor de textos pot editar alguns formats de documents (aquells que tenen les marques de format de tipus textual).

El processador de textos al qual farem referència durant el mòdul és Open Office Writer. Open Office és un conjunt d'aplicacions ofimàtiques (processador de textos, full de càlcul, programa per a fer presentacions, base de dades, etc.) de programari lliure. Està disponible per a Windows, Linux i Mac. En tot el mòdul farem referència a la versió en anglès d'aquesta eina, tot i que també està disponible en català.

2. Organització del document i format final

Quan escrivim els nostres documents convé oblidar-nos de les qüestions de format i presentació i centrar-nos en l'estructura i contingut. Això vol dir que no importa si els títols han d'estar en font Arial 16 i negreta o en Times New Roman 17 i cursiva. Tampoc ens ha d'importar en aquest moment si les pàgines han d'estar numerades o no i si el número de pàgina ha d'estar centrat o en negreta.

Quan escrivim un treball ens hem de preocupar de la seva estructura i del seu contingut. És a dir, hem de tenir clar com organitzem el nostre treball (apartats, subapartats, etc.) i també hem de tenir clar què escrivim en cada un dels apartats.

Un cas pràctic

En una de les assignatures que curseu hem fet un treball molt bo. El consultor ens proposa que publiquem aquest treball en una revista de la Universitat. El problema està en què el termini per a les propostes d'articles acaba avui mateix. A la revista exigeixen la presentació en un format molt determinat (defineixen exactament la font per als títols de cada nivell, els marges per a les pàgines, com ha de ser l'encapçalament i el peu de pàgina, etc.). El treball té més de 30 pàgines, però com que l'hem redactat amb el nostre processador de textos fent un bon ús dels estils, no triguem més de 10 minuts a acomplir els requeriments de la revista i podem presentar la proposta d'article a temps.

2.1. Processadors de textos: estils i plantilles

Tots els processadors de textos actuals permeten treballar amb estils. Els estils ens permeten oblidar-nos de pensar en els atributs físics del document que estem escrivint (formats com tipus de lletra, marges, numeració de pàgines, etc.) i centrar-nos en els atributs lògics. Per exemple, no hem de pensar que el títol ha d'estar en Arial 14 negreta i centrat, sinó que simplement hem d'especificar que el que estem escrivint és el títol de document. Això té enormes avantatges si hem de fer canvis d'última hora en els nostres documents o bé si volem reaprofitar parts d'un document en altres documents.

Les plantilles permeten crear tipus de documents a partir dels quals es poden crear nous documents. En una plantilla es poden definir totes les característiques que es poden definir en qualsevol document (des dels estils, encapçalaments i peus de pàgina, la presència d'un determinat logo en els encapçalaments, etc.). Quan creem un document a partir de la plantilla, aquest nou document tindrà incorporades totes aquestes característiques.

2.1.1. Estils

En l'Open Office Writer hi ha cinc tipus d'estils: de paràgraf, de caràcter, de pàgina, de marc i de llista.

1) Els **estils de paràgraf** afecten el paràgraf sencer marcat amb l'estil. Això vol dir que no podem marcar un fragment d'un paràgraf amb un estil de paràgraf diferent del de la resta del paràgraf. Els estils de paràgraf serveixen per a definir propietats del text (font, mida, indentació, justificació). Alguns dels estils de paràgraf per defecte que presenta l'Open Office Writer són els següents:

- *Title*: per a marcar el títol de document.
- *Subtitle*: per a marcar el subtítol d'un document.
- *Heading1*: per a marcar el títol de secció de nivell 1.
- *Heading2..10*: per a marcar el títol de secció de nivell 2 fins al nivell 10. Normalment en els documents serà suficient amb dos o tres nivells.
- *Text*: per a marcar el text normal dels documents.

Hi ha molts més estils de paràgraf i també és possible definir els nostres propis estils.

2) Els **estils de caràcter** serveixen també per a definir propietats del text (font, mida, color del text, etc.) però es poden aplicar a unes paraules determinades del paràgraf, sense necessitat d'aplicar-lo a tot un paràgraf. A continuació presentem alguns dels estils de caràcter que presenta per defecte l'Open Office Writer, però també en aquest cas es poden crear estils de caràcter nous:

- *Default*: és l'estil de caràcter per defecte, el que es fa servir en el text normal.
- *Definition*: es fa servir per a marcar una definició.
- *Emphasis*: es fa servir per a marcar text ressaltat.
- *Example*: es fa servir per a marcar un exemple.
- *Internet Link*: es fa servir per a marcar un enllaç d'Internet.
- *Strong Emphasis*: es fa servir per a marcar un text amb un ressaltat més fort.
- *Visited Internet Link*: es fa servir per a marcar un enllaç d'Internet que ja ha estat visitat.

3) Els **estils de pàgina** permeten definir propietats de la pàgina (marges, mides de la pàgina, capçalera i peu de pàgina, etc.). Alguns dels estils de pàgina que presenta l'Open Office Writer per defecte són els següents:

- *Default*: és l'estil de pàgina per defecte.
- *Envelope*: defineix les mides d'un sobre.
- *First page*: defineix les propietats de la primera pàgina d'un document.
- *Index*: defineix les propietats de la pàgina o pàgines que contenen l'índex.
- *Left Page*: defineix les propietats de les pàgines de l'esquerra.
- *Right Page*: defineix les propietats de les pàgines de la dreta.

4) Els **estils de marc** permeten definir les propietats dels marcs que contenen certs objectes de text o de gràfics. Els estils de marc per defecte que presenta l'Open Office Writer són els següents:

- *Formula*: és l'estil de marc que es fa servir per a les fórmules, inclou per defecte la propietat *AutoSize*, que ajusta la mida del marc a la mida de la fórmula.
- *Frame*: és l'estil per defecte per als marcs.
- *Graphics*: és l'estil per defectes per als gràfics, que també inclou *AutoSize*.
- *Marginalia*: és un estil per a incloure un marc més enllà del marge esquerre i s'acostuma a fer servir per a incloure comentaris al marge.
- *OLE*: és l'estil de marc per defecte per a fer servir amb objectes OLE (*Object Linking and Embedding*).
- *Watermark*: és l'estil per defecte per a les marques d'aigua, és a dir, per a incloure un gràfic com a fons d'un text.

5) Els **estils de llista** (que també reben el nom d'*estils d'enumeració*) defineixen les característiques de les enumeracions (identació, alineació, estil de número o de marca d'enumeració, etc.).

Web recomanat

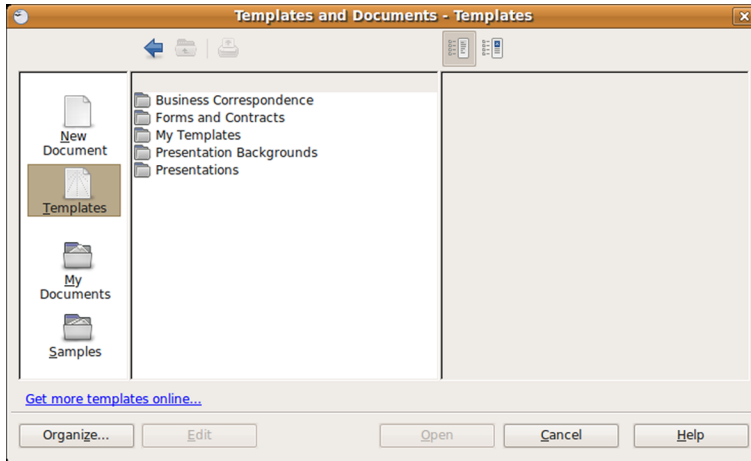
Per a completar la informació sobre l'ús d'estils amb l'Open Office consulteu el manual de l'aplicació, que podeu descarregar a <http://documentation.openoffice.org/>. Concretament als capítols 6 ("Introducció als estils") i 7 ("Treballar amb estils"), trobareu tota la informació complementària.

2.1.2. Plantilles

Una plantilla és una mena de model que es fa servir per a crear altres documents. Una plantilla pot contenir exactament el mateix que pot contenir un document (text, gràfics, un conjunt d'estils, un idioma determinat, etc.). Tots els documents que es creen amb l'Open Office Writer es basen en plantilles. Si quan creem un nou document no especifiquem cap plantilla, el document es basarà en la plantilla per defecte per a documents de text.

Si fem *File > New > Templates and Documents* s'obre una finestra com la que es mostra a la figura 1.

Figura 1. Pantalla de plantilles i documents de l'Open Office Writer



En aquesta pantalla (figura 1), en la carpeta *Templates* tenim un seguit de plantilles i podem guardar-hi les que creem nosaltres mateixos. Fixeu-vos també en l'enllaç *Get more templates online...* que ens permet descarregar noves plantilles.

2.2. El format LaTeX

LaTeX és un llenguatge de marcatge i un sistema de preparació de documents per al sistema tipogràfic TeX. LaTeX es fa servir molt en l'àmbit científic i acadèmic, ja que permet generar uns documents amb una qualitat visual molt gran.

LaTeX es basa totalment en la idea que l'autor d'un document s'ha de centrar en el contingut sense preocupar-se en absolut per l'aspecte visual final del document. Quan s'escriuen documents en LaTeX l'autor especifica l'estructura lògica del document fent servir conceptes com capítol, secció, taula, figura, etc.

A la figura 2 podem veure un exemple de document en LaTeX (extret de la Viquipèdia). D'una banda observem el codi LaTeX, és a dir, el que escriu l'autor, i de l'altra observem l'aspecte final un cop processat el document.

Figura 2. Document en LaTeX (esquerra) i resultat final un cop processat (dreta)

```

\documentclass[12pt]{article}
\usepackage{amsmath}
\title{\LaTeX}
\date{}
\begin{document}
\maketitle
\LaTeX{} is a document preparation system for the \TeX{}
typesetting program. It offers programmable desktop publishing
features and extensive facilities for automating most aspects of
typesetting and desktop publishing, including numbering and
cross-referencing, tables and figures, page layout, bibliographies,
and much more. \LaTeX{} was originally written in 1984 by Leslie
Lamport and has become the dominant method for using \TeX; few
people write in plain \TeX{} anymore. The current version is
\LaTeXe.

% This is a comment; it is not shown in the final output.
% The following shows a little of the typesetting power of LaTeX
\begin{align}
E &= mc^2 \\
m &= \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}
\end{align}
\end{document}

```

LaTeX

LaTeX is a document preparation system for the TeX typesetting program. It offers programmable desktop publishing features and extensive facilities for automating most aspects of typesetting and desktop publishing, including numbering and cross-referencing, tables and figures, page layout, bibliographies, and much more. LaTeX was originally written in 1984 by Leslie Lamport and has become the dominant method for using TeX; few people write in plain TeX anymore. The current version is LaTeX 2_ε.

$$E = mc^2 \quad (1)$$

$$m = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}} \quad (2)$$

Lectura complementària

Podeu aprendre a fer servir plantilles amb l'Open Office Writer llegint el capítol 10 del manual, titulat "Treballar amb plantilles".

LaTeX normalment es distribueix juntament amb TeX i té una llicència de programari lliure. LaTeX està disponible per a molts sistemes operatius, entre ells Unix, Linux, Mac i Windows.

2.2.1. Requeriments per a fer servir LaTeX

Per a escriure documents en LaTeX únicament és necessari disposar d'un editor de textos. Tot i això, com veurem ara, hi ha una sèrie d'editors especials per a escriure i processar documents en LaTeX. Ara bé, per a processar aquests documents i obtenir una sortida en algun format imprimible (per exemple PDF) necessitarem disposar del paquet LaTeX i TeX. A continuació expliquem com obtenir i instal·lar aquests paquets per a Linux, Mac i Windows.

1) **Instal·lació de LaTeX per a Linux:** la majoria de distribucions de Linux proporcionen tant TeX com LaTeX i estan disponibles en el sistema de gestió de paquets. Per a editar els documents podem fer servir qualsevol editor de textos, o bé algun editor específic, com per exemple Kile o Lyx (que trobareu també fàcilment en el vostre sistema de gestió de paquets).

2) **Instal·lació de LaTeX per a Mac:** la distribució MacTeX conté tot el que es necessita per a escriure i processar documents en LaTeX amb Mac, des d'un sistema TeX complet fins a editors per a escriure els documents.

3) **Instal·lació de LaTeX per a Windows:** la distribució ProTeXt permet instal·lar una distribució de TeX i l'editor TeXnicCenter que permet confeccionar fàcilment documents en LaTeX sota Windows.

2.2.2. Les principals instruccions de LaTeX

En aquest subapartat explicarem les principals instruccions de LaTeX mitjançant el següent document d'exemple:

```
\documentclass[12pt]{article}
\usepackage[utf8]{inputenc}
\usepackage[catalan]{babel}

\title{Un document d'exemple en \LaTeX}
\author{Antoni Oliver}
%\date{23 de maig de 2010}
\begin{document}
  \maketitle

  \begin{abstract}
    Aquest article és una molt breu introducció a \LaTeX. No pretén mostrar tota la
    seva potència, sinó simplement les comandes més útils per a filòlegs i
    lingüistes. Per aquest motiu no poso cap exemple de fórmules
  \end{abstract}
```



```
\section{Introducció}
```

Aquest exemple de document serveix per a explicar alguns aspectes bàsics de `\LaTeX`.

```
\section{Organització del document en seccions}
```

El document es pot organitzar en seccions.

```
\subsection{Organització del document en subseccions}
```

El document també es pot organitzar en subseccions.

```
\subsubsection{Organització del document en subsubseccions}
```

I també en secció de tercer nivell, és a dir, subsubseccions.

```
\section{Conclusions}
```

Ara ja podem escriure el nostre propi document en `\LaTeX`.

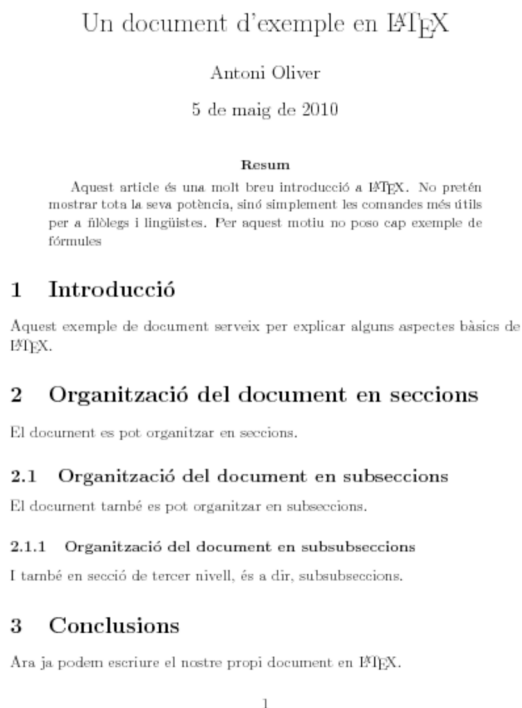
```
\end{document}
```

Si processem aquest document obtenim la sortida que mostrem a la figura 3.

Vegeu també

En el subapartat següent expliquem com es processen els documents en `\LaTeX`.

Figura 3. Resultat de processar el document d'exemple en LaTeX



Fixeu-vos que la data que apareix en el document final no és la mateixa que l'especificada en el document LaTeX, ja que la instrucció *date* es troba comentada (precedida pel caràcter %). La data que apareix en el document final és la corresponent al dia en què es va processar el document.

2.2.3. Processament de documents en LaTeX

Normalment es converteixen els documents en LaTeX en arxius PDF, que es poden visualitzar i imprimir sense problemes en qualsevol sistema operatiu. Si feu servir algun editor específic de LaTeX (TeXnicCenter, Kile, Lyx, etc.) tindreu una opció del menú o de la barra d'eines que us permetrà fer aquesta conversió.

Si treballeu amb Linux o Mac podreu processar els documents manualment des d'un terminal, mitjançant el conjunt d'ordres següent. El primer que cal fer és (considerem que el nostre document en LaTeX es diu "exemple.tex"):

```
latex exemple.tex
```

Amb això obtenim un arxiu en format DVI. Si voleu veure com ha quedat (ja que serà exactament igual que el PDF final) podeu fer:

```
xdvi exemple.dvi
```

Ara convertirem l'arxiu DVI en PS:

```
dvips exemple.dvi
```

Per a convertir el PS en PDF hem de fer:

```
ps2pdf exemple.ps
```

Amb això ja tenim creat l'arxiu "exemple.pdf". Aquest procés es pot simplificar una mica fent:

```
latex exemple.tex  
dvi2pdf exemple.dvi
```

I encara més, fent-t'ho tot en una sola instrucció:

```
pdflatex exemple.tex
```

2.2.4. Conclusions

Amb aquesta petita introducció al format LaTeX ja podem crear els nostres propis documents en aquest format. Aquest format es fa servir molt en àmbits acadèmics i per a realitzar treballs de tesis doctorals. Tot i que sembla un format antiquat i incòmode, les grans funcionalitats que proporciona el fan molt aconsellable. Entre les grans qualitats d'aquest format en podem destacar:

- 1) La gran qualitat visual dels documents que es generen.
- 2) La gran facilitat per escriure fórmules matemàtiques.
- 3) La gran facilitat (com veurem més endavant) per a gestionar bibliografia, fer les corresponents cites i generar l'apartat de bibliografia.
- 4) Quan escrivim un document ens podem oblidar totalment de qüestions d'aspecte final i centrar-nos en el contingut. Al final podrem generar el document final seleccionant l'estil que més ens agradi i modificant els estils existents a les nostres necessitats o preferències. Això inclou també la bibliografia, de manera que no ens hem de preocupar de pensar si l'any va al final o darrera del nom, si el títol va en negreta o no. Tot això ho generarà automàticament el sistema a partir d'uns estils bibliogràfics.

Aquesta és, sense cap mena de dubtes, una de les millors opcions actuals per a fer treballs acadèmics.

2.3. El format DocBook

DocBook és un format per a representar documents basat en XML. XML és un llenguatge de marcatge estàndard que permet la representació i intercanvi d'informació estructurada. El llenguatge de marcatge XML està prenent cada dia més importància en tots els àmbits (de fet, el format intern de l'Open Office és també un XML).

DocBook segueix una idea similar a la de LaTeX, on l'autor no s'ha de preocupar per la presentació, sinó per l'organització del document i pel seu contingut. Com que DocBook es basa en XML i XML és un llenguatge de marcatge estàndard i molt estès, hi ha moltes eines que ens permetran tractar el format DocBook.

2.3.1. Requisits per a fer servir DocBook

El format XML és un format de text i per tant DocBook també ho és. En aquest sentit podem fer servir qualsevol editor de textos per a escriure documents en DocBook. També podem fer servir algun editor específic per a fitxers XML i disposar d'alguna funcionalitat extra. Una altra bona opció és fer servir l'Open Office Writer amb les seves extensions per a escriure documents en DocBook. Un editor d'XML senzill que proporciona diverses funcionalitats relacionades amb DocBook és l'XML Copy Editor.

Descarregar l'XML Copy Editor

L'XML Copy Editor es pot descarregar des de <http://xml-copy-editor.sourceforge.net/>.

Per a processar documents escrits en DocBook i generar una sortida, per exemple, en PDF, tenim diverses opcions:

- 1) Fer servir fulls d'estil (*stylesheets*) i processar el document amb aquests fulls. Molts dels editors d'XML permeten fer aquesta operació.
- 2) Fer servir el paquet *dblatex*.

2.3.2. Les principals etiquetes de DocBook

En aquesta secció presentarem les principals etiquetes de DocBook mitjançant un parell d'exemples. En primer lloc presentarem un exemple d'article molt similar al de l'exemple corresponent al format LaTeX.

```
<?xml version="1.0" encoding="UTF-8"?>
<article lang="ca">
  <title>Un document d'exemple en DocBook</title>
  <articleinfo>
```

```

<author>
  <firstname>Antoni</firstname>
  <surname>Oliver</surname>
</author>
<date>23 de maig de 2010</date>
</articleinfo>
<abstract>
<para>Aquest article és una molt breu introducció a DocBook. No pretén mostrar
tota la seva potència, sinó simplement les comandes més útils per a filòlegs i
lingüistes.</para>
</abstract>
<sect1>
  <title>Introducció</title>
  <para>Aquest exemple de document serveix per a explicar alguns aspectes
bàsics de DocBook. </para>
</sect1>
<sect1>
<title>Organització del document en seccions</title>
<para>El document es pot organitzar en seccions. </para>
<sect2>
  <title>Organització del document en subsubseccions</title>
  <para>El document també es pot organitzar en subseccions. </para>
<sect3>
  <title>Organització del document en subsubseccions</title>
  <para>I també en secció de tercer nivell, és a dir, subsubseccions. </para>
  <para>Fixem-nos que a diferència de LaTeX es marca cada paràgraf
del document.</para>
</sect3>
</sect2>
</sect1>
<sect1>
  <title>Conclusions</title>
  <para>Ara ja podem escriure el nostre propi document en DocBook. </para>
</sect1>
</article>

```

Amb l'etiqueta `<article lang="ca">` estem definint que el nostre document és un article i que està escrit en català. Fixeu-vos que aquesta etiqueta engloba tot el document, ja que no es tanca amb `</article>` fins al final de tot. De seguida definim el títol del document amb `<title>` i posteriorment la informació associada al document amb l'etiqueta `<articleinfo>`. Un dels grans avantatges de DocBook és que es molt autoexplicatiu, ja que mirant el nom de les etiquetes tenim una informació força acurada del que contenen.

Cal tenir en compte que els diferents nivells de les seccions s'especifiquen amb `<sect1>`, `<sect2>`, `<sect3>`, etc. Fixeu-vos que les etiquetes dels diferents nivells estan encavalcades, és a dir, una secció de nivell 2 està dins d'una de nivell 1 i no es pot tancar la de nivell 1 si no es tanca abans la de nivell 2. Fixeu-vos també que els diferents paràgrafs es marquen amb l'etiqueta `<para>`.

Vegeu també

En el subapartat "Processament de documents en DocBook" podeu veure com queda aquest article.

A continuació observem un exemple de llibre en format DocBook:

```
<?xml version="1.0" encoding="UTF-8"?>
<book lang="ca">
  <title>Un exemple de llibre senzill escrit en DocBook</title>
  <bookinfo>
    <author>
      <firstname>Antoni</firstname>
      <lastname>Oliver</lastname>
    </author>
  </bookinfo>
  <chapter>
    <title>Què és DocBook</title>
    <para>DocBook és un llenguatge basat en XML per a escriure documentació. Es fa servir molt per a escriure documentació tècnica de programes, tot i que es pot fer servir per a escriure qualsevol tipus de document</para>
    <para>DocBook, igual que LaTeX, permet a l'autor concentrar-se en l'organització del document i en el seu contingut.</para>
  </chapter>
  <chapter>
    <title>Quins avantatges té DocBook?</title>
    <para>Un avantatge és que DocBook és XML i per tant pot fer servir totes les eines que hi ha per a processar XML.</para>
  </chapter>
</book>
```

Web recomanat

Podeu consultar una referència ràpida sobre les diferents etiquetes que permet DocBook a: <http://www.docbook.org/tdg/en/html/docbook.html>

2.3.3. Processament de documents en DocBook

A continuació presentarem els resultats de processar els documents d'exemple del subapartat anterior fent servir el paquet *dblatex*. Per a processar el document amb els paràmetres per defecte simplement hem de fer:

```
dblatex exemple-article.xml
```

Amb aquestes instruccions es generarà un fitxer "exemple-article.pdf" que conté diverses pàgines. Amb *dblatex* podem especificar l'estil amb què volem generar l'article. Es fan servir els mateixos estils que en la generació a partir de documents LaTeX. Per exemple, si fem servir l'estil simple:

```
dblatex -T simple exemple-article.xml
```

Ens apareix una sortida com la que es mostra a la figura 4.

Figura 4. Resultat de processar el document DocBook d'exemple amb l'opció simple

Un document d'exemple en DocBook

Oliver, Antoni
10 de maig de 2010

Índex

1	Introducció	1
2	Organització del document en seccions	1
2.1	Organització del document en subseccions	1
2.1.1	Organització del document en subseccions	1
3	Conclusions	1

Resum

Aquest article és una molt breu introducció a DocBook. No pretén mostrar tota la seva potència, sinó simplement les comandes més útils per a filòlegs i lingüistes.

1 Introducció

Aquest exemple de document serveix per explicar alguns aspectes bàsics de DocBook.

2 Organització del document en seccions

El document es pot organitzar en seccions.

2.1 Organització del document en subseccions

El document també es pot organitzar en subseccions.

2.1.1 Organització del document en subseccions

I també en secció de tercer nivell, és a dir, subseccions.

Fixem-nos que a diferència de LaTeX es marca cada paràgraf del document.

3 Conclusions

Ara ja podem escriure el nostre propi document en DocBook.

1

Web recomanat

Per a usuaris més avançats, podeu configurar tot el que fa referència a estils i crear estils nous consultant l'enllaç següent: <http://dblatex.sourceforge.net/doc/manual/sec-style.html>

2.3.4. Conclusions

Com hem pogut observar en aquest apartat, DocBook és un format molt interessant ja que és XML i es pot processar amb qualsevol de les eines disponibles per a XML. És possiblement el format de document més interessant per a experts en formats XML, tot i que potser és una mica complicat per a usuaris amb coneixements mitjans.

És molt possible que a poc a poc vagin apareixent eines senzilles per a treballar amb aquest format. També és d'esperar que el suport per a DocBook d'Open Office millori en les properes versions.

3. Edició col·laborativa

En aquest apartat veurem un seguit d'eines que ens permeten editar documents d'una manera col·laborativa, és a dir, amb la intervenció de diversos autors. En aquest sentit, veurem les funcionalitats que ofereix Open Office Writer. Quan en un document intervé més d'un autor, és molt important tenir un control de què ha afegit, canviat i eliminat cada un dels autors. També és molt important tenir un control de versions força exhaustiu, que ens permeti en qualsevol moment recuperar el document en un estat anterior.

D'altra banda, veurem l'eina **Google Docs**, que ens permet editar documents de diferents tipus de manera col·laborativa. Aquests documents els podran editar els usuaris escollit pel seu creador i també es mantindrà un historial de canvis.

3.1. Edició col·laborativa amb Open Office

Open Office ofereix tres eines bàsiques per a controlar els canvis en un document:

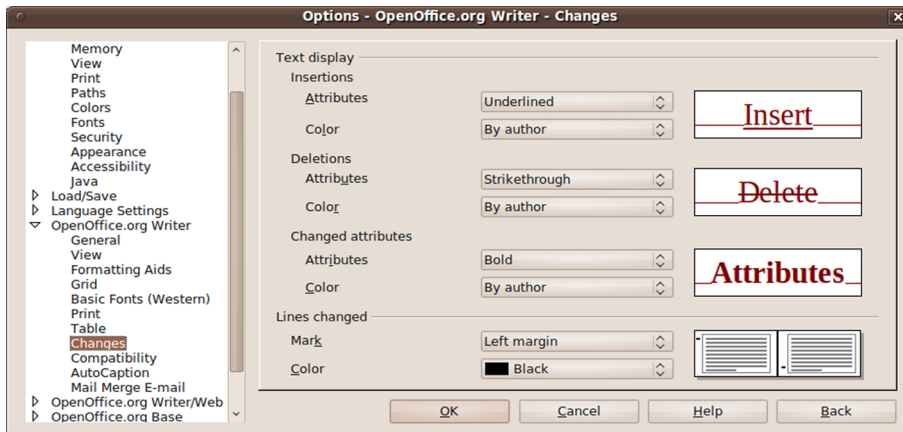
- 1) Control de canvis
- 2) Comparació de documents
- 3) Control de versions

Aquestes funcionalitats poden ser de gran ajuda quan un document ha estat editat per diversos autors i revisors.

3.1.1. Control de canvis

Per activar el control de canvis cal fer *Tools > Options > OpenOffice.org Writer > Changes* i apareixerà una pantalla com la que es mostra a la figura 5.

Figura 5. Pantalla de control de canvis de l'Open Office Writer



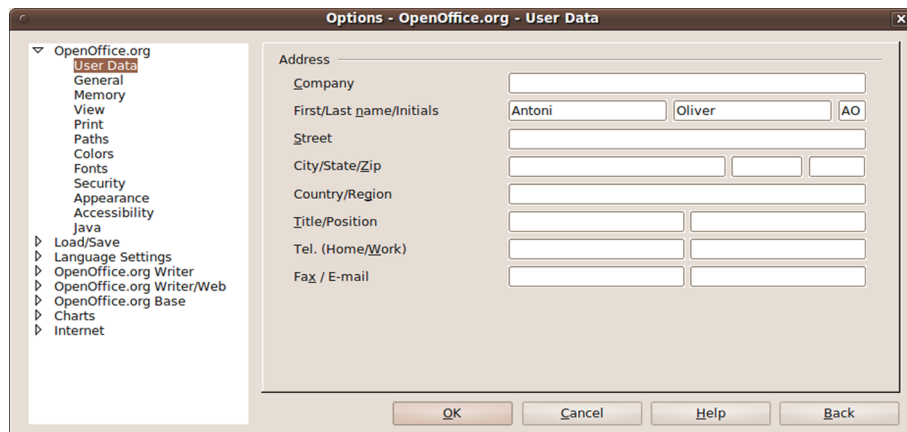
Aquesta pantalla permet especificar com es visualitzaran els diferents tipus de canvis en el document:

- 1) *Insertions*: contingut que està en la segona versió del document i no en la primera.
- 2) *Deletions*: contingut que està en la primera versió del document i no en la segona.
- 3) *Changed attributes*: qualsevol canvi de format.
- 4) *Lines changed*: indiquem on volem que aparegui la barra que indica un canvi a la línia.

Per a *insertions*, *deletions* i *changed attributes* es pot assignar el color que marcarà els canvis fets per l'usuari actual, però aquesta assignació de colors la pot fer OpenOffice automàticament.

Per tal que el control de canvis sigui efectiu és molt important que Open Office Writer disposi de la informació d'usuari que es pot afegir des de *Tools > Options > OpenOffice.org > User Data*. Com a mínim cal introduir el nom, el cognom i les inicials, com a la figura 6.

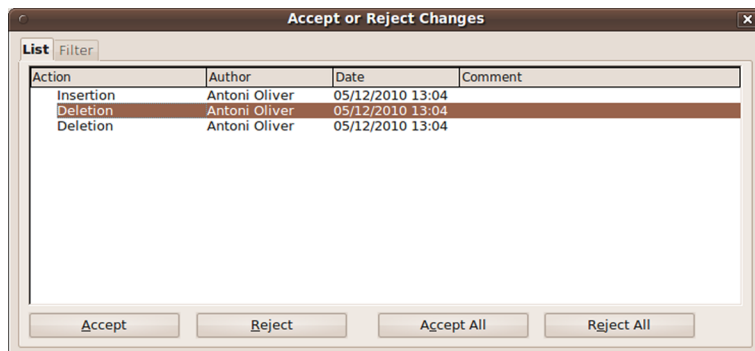
Figura 6. Pantalla de dades de l'usuari d'Open Office



Si Open Office Writer no té aquesta informació els canvis es llistaran com si estiguessin fets per un usuari anomenat *Unknown* i serà molt més complicat fer un bon seguiment dels canvis.

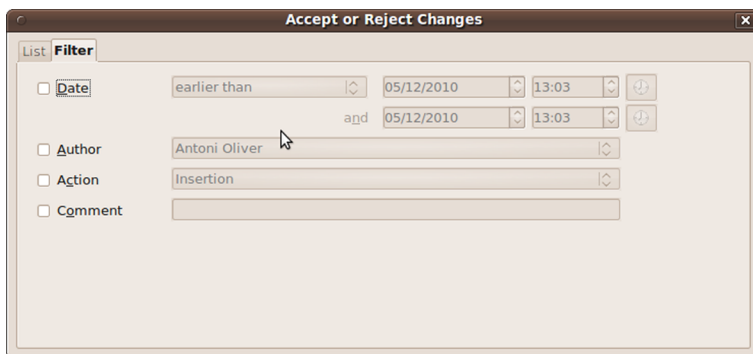
Per a poder treballar amb el control de canvis el primer que hem de fer és activar-lo amb *Edit > Changes > Record*. A partir d'aquest moment es registraran tots els canvis.

Si volem observar els canvis que s'han realitzat en el document podem activar la pantalla *Accept or Reject Changes* amb *Edit > Changes > Accept or reject*. Aquesta pantalla ofereix dues pestanyes, l'anomenada *List* (figura 7).

Figura 7. Pantalla d'acceptació o rebuig de canvis de l'Open Office Writer (pestanya *List*)

En aquesta pestanya *List* podem observar tots els canvis i acceptar-los o rebutjar-los. Pel que fa a la pestanya *Filter* (figura 8), és útil per a documents on s'han fet molts canvis i on han intervingut molts autors, ja que ens permet filtrar els canvis segons diversos paràmetres.

Figura 8. Pantalla d'acceptació o rebuig de canvis de l'Open Office Writer (pestanya *Filter*)



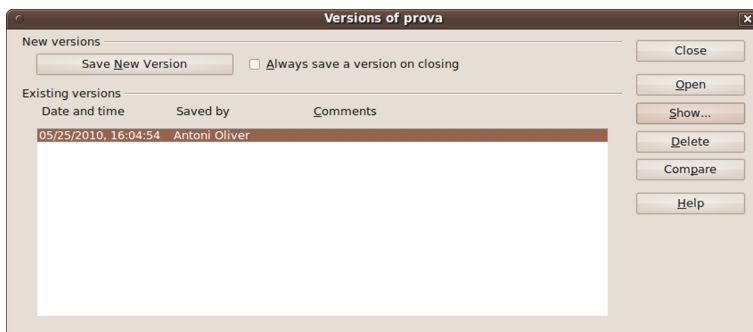
3.1.2. Comparació de documents

Open Office ofereix una eina de comparació de documents que pot ser útil en alguns casos (per exemple per a comparació de diverses versions d'un mateix document). L'eina s'activa fent *Edit > Compare Documents*. En aquest moment podem seleccionar un altre document i s'activarà el control de canvis com si el document inicial fos el resultat de canviar el document que acabem de seleccionar. D'aquesta manera, si volem mantenir el contingut que està únicament en el primer document hem d'acceptar els canvis, i per a mantenir el contingut que està en el segon document hem de rebutjar-lo.

3.1.3. Control de versions

Open Office permet també l'ús de versions del document. Si volem mantenir un control de versions d'un mateix document no és necessari gravar diversos fitxers corresponents al document en diferents versions. En comptes d'això, podem crear una nova versió que es guardarà en el mateix fitxer. Això permet recuperar una versió anterior en qualsevol moment. Per a fer això anem a *File > Versions* i hi apareix una pantalla com la que es mostra a la figura 9.

Figura 9. Pantalla de control de versions de l'Open Office Writer



Aquesta pantalla ens permet generar una nova versió del document amb *Save New Version*. Per a cada versió es guardarà la data i l'hora, l'usuari que l'ha guardat i si es vol, un petit comentari explicatiu. També podem activar la casella *Always save a version on closing* per a fer que cada cop que tanquem el document es generi una nova versió.

Amb aquest control de versions podrem accedir a qualsevol de les versions guardades. Podem obrir-les en una nova finestra, comparar-la amb la versió actual amb la finestra de control de canvis i també esborrar-la. Cal tenir en compte que cada cop que guardem una versió la mida del fitxer augmenta. Si el document és llarg i guardem massa versions podem tenir problemes amb la mida de l'arxiu. Per aquest motiu cal anar en compte a no guardar massa versions i a esborrar les versions que ja no farem servir.

3.2. Edició col·laborativa amb Google Docs

Google Docs permet crear, compartir i editar documents en línia d'una manera molt fàcil. Per a poder accedir a aquest servei gratuït simplement cal tenir un compte de Google. Algunes de les coses que es pot fer amb Google Docs són:

- Carregar documents de Word, OpenOffice, RTF, HTML o fitxers de text.
- Crear documents des de zero i descarregar-los posteriorment en algun d'aquests formats.
- Editar documents simultàniament amb altres usuaris. El creador del document pot assignar quins usuaris poden visualitzar el document i quins usuaris el poden editar.
- Controlar els canvis del document, saber quin usuari ha fet un determinat canvi i restaurar versions anteriors del document.
- Publicar documents com a webs o com a entrades d'un blog.
- Enviar els documents com a adjunts a un missatge de correu electrònic.

Un compte a Google

Per a tenir un compte a Google, us podeu registrar a www.google.com

3.2.1. Creació i edició de documents

Amb Google Docs es poden pujar documents existents o bé crear-los de nou o a partir d'una plantilla. Per a crear un document simplement hem d'anar a la nostra llista de documents i fer *New > Document*. Podem crear documents, presentacions, fulls de càlcul i formularis per a fer enquestes. Aquest darrer tipus és molt interessant, ja que després podrem enviar l'enllaç d'aquestes enquestes i veure els resultats en forma de full de càlcul.

Per a guardar el document simplement hem de fer *Save* i es guardarà en el servidor. Si volem tenir una còpia local del document hem de fer *Download As* i triar un dels formats que ens proposa.

També es poden crear documents a partir d'una plantilla. Google Docs ofereix una gran varietat de plantilles.

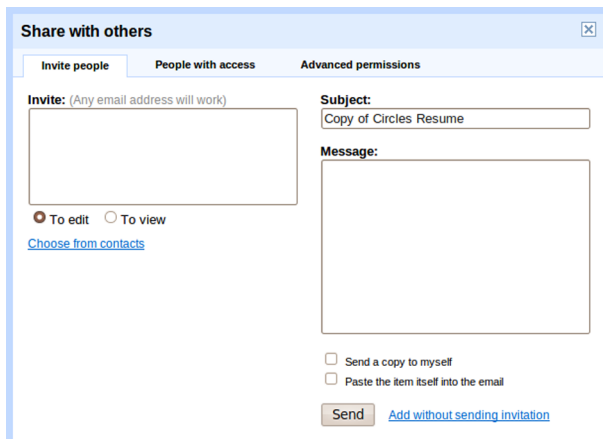
Recordem que també és possible pujar un document que ja estigui creat. Un cop pujat el podem editar des de la interfície de Google Docs.

Una característica molt interessant és la possibilitat d'afegir comentaris als documents. Això es pot fer amb *Insert > Comment*. Aquests comentaris no apareixeran quan es publiqui o s'imprimeixi el document.

3.2.2. Compartir documents i edició col·laborativa

Una de les característiques més interessants de Google Docs és que permet l'edició col·laborativa simultània d'un mateix document. Per a poder editar simultàniament un document primer cal compartir-lo. Això es pot fer tant des del propi document com des de la llista de documents. Simplement cal seleccionar *Share > Invite People*. Apareixerà una pantalla com la de la figura 10.

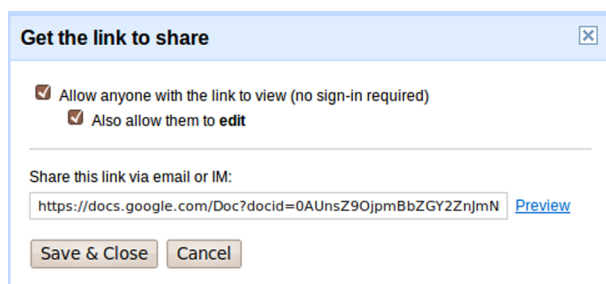
Figura 10. Pantalla per a compartir el document de Google Docs



En aquesta pantalla podem posar tantes adreces de correu electrònic com vulguem i enviar un missatge amb la invitació. Podem restringir la invitació a visualització o a edició.

Una altra possibilitat per a compartir documents és fer *Share > Get the link to share*. En aquest cas apareix una pantalla com la de la figura 11.

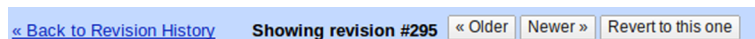
Figura 11. Pantalla per a obtenir l'enllaç i compartir el document de Google Docs



En aquest cas podem escollir entre permetre només la visualització o permetre'n també l'edició. L'enllaç que es genera permet accedir al document. Podem enviar aquest enllaç per correu electrònic.

Google Docs també permet fer un control dels canvis, veure quin usuari els ha fet i recuperar una versió anterior del document des de *File > See revision history*. Per a recuperar una versió anterior simplement hem de seleccionar una de les versions mitjançant l'enllaç. Llavors es mostrarà aquesta versió del document i a dalt de tot els botons que podem observar a la figura 12.

Figura 12. Botons de navegacions per a les versions del document de Google Docs

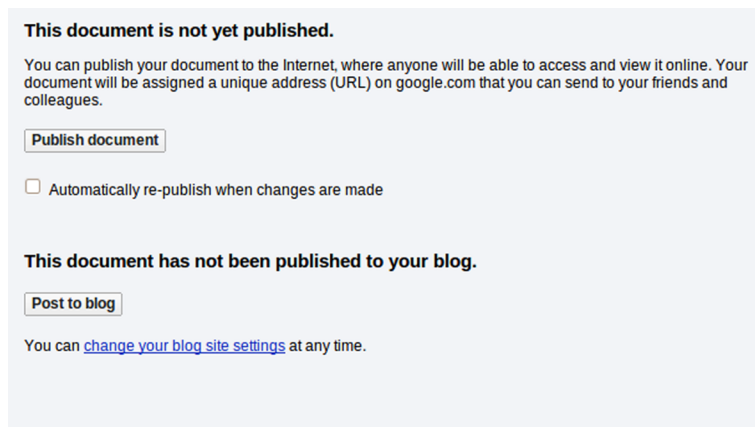


Podem tornar a la pantalla de revisió de versions, o bé veure una versió més antiga o més nova. Finalment podem fer que la versió actual sigui la corresponent a la que estem visualitzant fent *Revert to this one*.

3.2.3. Publicar i imprimir documents

Des de Google Docs podem publicar el nostre document com una pàgina web o com una entrada d'un blog. Per a aconseguir això, des de la llista de documents seleccionem el document que volem publicar i fem *More Actions > Publish*. Apareixerà una pantalla com la de la figura 13.

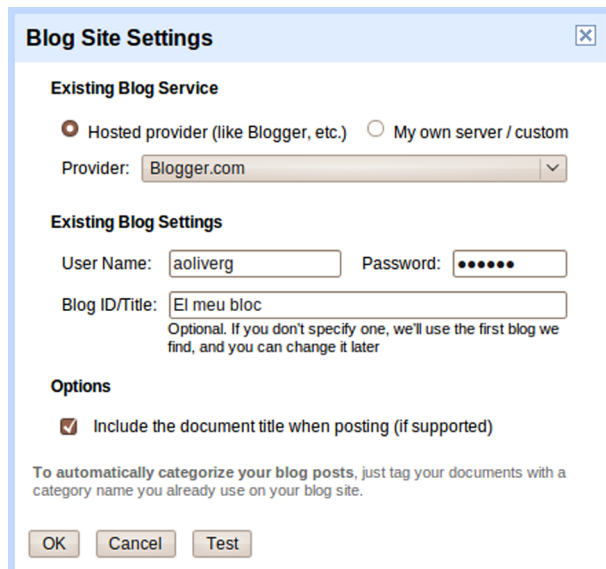
Figura 13. Pantalla per a publicar documents de Google Docs



Si seleccionem *Publish document* se li assignarà un URL i el document serà accessible des de qualsevol navegador d'Internet. Podem activar la opció de que es torni a publicar de manera automàtica si fem canvis al document.

També podem posar aquest document en el nostre blog. Abans haurem de seleccionar la configuració del blog des de l'enllaç *Change your blog site settings*. Apareixerà una pantalla com la de la figura 14.

Figura 14. Configuració per a publicar el document de Google Docs en un blog



Si fem *Print* des de Google Docs es generarà un PDF. Aquest PDF sí que el podem imprimir a la nostra impressora.

3.2.4. Altres eines interessants

Google Docs presenta altres eines que poden ser interessants (s'hi pot accedir des de *Tools*):

- *Translate document*: permet traduir el document amb el motor Google Translate.
- *Check spelling*: per a fer la correcció ortogràfica.
- *Look up word*: permet buscar definicions, sinònims o entrades d'enciclopèdia a partir d'una paraula (només funciona per a algunes llengües).
- *Search the web for word*: permet cercar la paraula o paraules seleccionades fent servir Google.

4. Gestió de la bibliografia

La correcta gestió de la bibliografia pot estalviar volums de treball molt importants, especialment si estem fent un treball d'una mida considerable amb força cites bibliogràfiques. Hi ha diverses maneres de fer una referència bibliogràfica i de presentar les entrades de la bibliografia. Imaginem-nos que tot just hem acabat un treball de 300 pàgines que conté 50 o 60 cites bibliogràfiques. Hem decidit fer les cites d'una determinada manera i presentar un únic recull de bibliografia al final del treball. Tot just abans de lliurar el treball ens assabentem que les cites s'han de fer d'una altra manera, que s'ha de presentar el recull de bibliografia per a cada capítol i que a més la manera de presentar les referències és diferent. Si hem gestionat de manera correcta la bibliografia i l'eina triada és l'adequada, aquest canvi no ens portarà més de 10 minuts. Si no, això es pot convertir en un veritable malson.

Un avantatge addicional de gestionar la bibliografia de manera adequada és que podem aprofitar tota la informació d'una determinada referència per a més d'un treball i podem acabar tenint una important base de dades bibliogràfica del nostre domini d'interès.

En aquest apartat presentarem algunes eines per a gestionar la bibliografia. Presentarem la manera de gestionar la bibliografia si fem servir OpenOffice Writer, LaTeX i DocBook. La darrera eina que presentarem és Zotero, un connector de Mozilla que ens permet gestionar la bibliografia i que es pot fer servir amb diferents processadors de textos.

4.1. Bibliografies amb OpenOffice Writer

OpenOffice permet crear fàcilment una base de dades de bibliografia. La creació d'aquesta base de dades és molt útil, ja que permet reaprofitar les entrades bibliogràfiques entre més d'un treball. Si fem *Tools > Bibliography Database* apareix una pantalla com la de la figura 15.

Vegeu també

OpenOffice Writer, LaTeX i DocBook estan relacionades amb les eines que hem presentat a l'apartat 2, "Organització del document i format final". Consulteu aquest apartat.

Figura 15. Base de dades bibliogràfica de l'Open Office

Identifier	Type	Address	Annote	Author	Booktitle	Chapter	Edition	Editor	Howpublish	Institut	Journal	Month	Note	Number	Organizat	Pages	Publisher	School	Series	Title	RepType
KRA01	1			Kraus, Klat												300	Kraus Verlag				StarOffic
REN00	1			Martin, Rer												304	Hanser Fachb				Berechnu
SCHO00	1			Schoblick,												509	Franzis				OpenOffi
HOR00	1			Horst, Ben												108	Hentzenwerki				The Tiny
PIT00	1			Pitonyak, A												460	Hentzenwerki				OpenOffi
SOPH00	1			Gautier, So												330	Eyrolles				OpenOffi
KRU00	1			Krumbein,												525	Galileo Press				OpenOffi
RAH00	1			Rahempon,												577	Software & Su				OpenOffi
SUN00	1			Sun Micros												456	Universe, Inc				Staroffic
SUN01	1			Sun Micros												67	Universe, Inc				Staroffic
SUN02	1			Sun Micros												204	Universe, Inc				Staroffic
SUN03	1			Sun Micros												219	Universe, Inc				Staroffic
GRU00	1			Grund, Gfr												155	IKON				OpenOffi
BRU00	1			Brunsteine												150	IKON				OpenOffi
MAN00	1			Manzano, J												192	Erica				OpenOffi
BR400	1			Braga, Will												200	Alta Books				Open Off
EST00	1			Esteves, Vi												160	Ciència Model				Dominar
LAM00	1			Lamas, Mu												376	Beto Brito				OpenOffi
WEB00	1			Weber, Jea												172	Weberwomar				Taming C

Aquesta pantalla ens permet editar els continguts de la base de dades bibliogràfica, eliminar registres, crear-ne de nous, etc. També podem seleccionar una altra base de dades fent clic al botó *Data Source*.

Per a introduir una referència a una entrada de la base de dades dins d'un text hem de posar el cursor en el punt on volem introduir la referència i fer *Insert > Indexes and Tables > Bibliographic entry*. Apareixerà una pantalla com la de la figura 16.

Figura 16. Pantalla per a inserir una entrada bibliogràfica amb l'Open Office

En aquesta pantalla podem seleccionar *From bibliography database* i directament el *Short Name* que volem de la llista desplegable corresponent. Això introduiria una referència de l'estil següent, fent servir el *Short Name*:

Bla bla bla bla bla [BOR04] bla bla bla...

En aquest sentit és interessant crear un *Short Name* que estigui d'acord amb com volem fer les cites, per exemple podríem haver triat com a *Short Name* "Borges, 2004" per a obtenir una referència de l'estil:

Bla bla bla bla bla [Borges, 2004] bla bla bla...

Podem canviar la manera de fer les referències i triar que es numerin:

Bla bla bla bla bla [1] bla bla bla...

Per a generar la bibliografia del document hem de situar-nos en el punt del document on volem inserir-la i fer *Insert > Indexes and Tables > Indexes and Tables* i canviar el tipus a *Bibliography*. Apareixerà una pantalla com la de la figura 17.

Figura 17. Pantalla per a inserir Índex/Taula de l'Open Office



En la pestanya *Index/Table* podem portar a terme les accions següents:

1) Donar un títol específic a la secció de bibliografia, canviant el valor del camp *Title*.

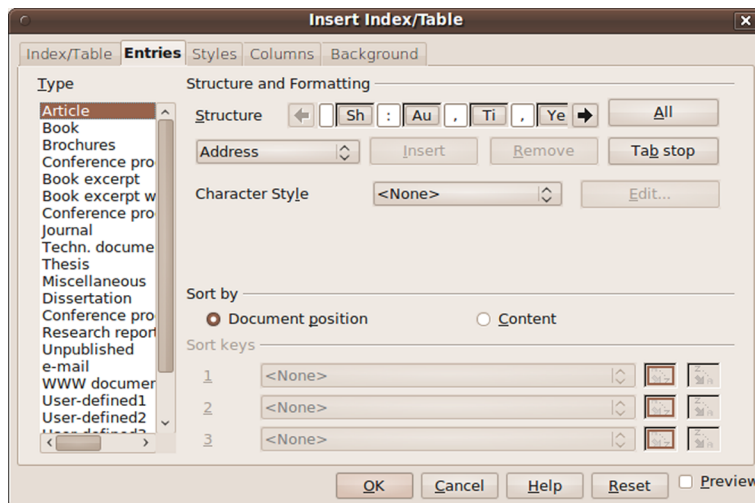
2) Si seleccionem la casella *Protect against manual changes* no serà possible canviar la bibliografia manualment. Si desactivem aquesta opció sí que serà possible fer canvis com si fos text normal, però els canvis que fem no es veuran reflectits en la base de dades de bibliografia.

3) Podem fer que les entrades estiguin numerades o bé que aparegui l'identificador de l'entrada.

4) Podem definir la manera d'ordenar les entrades. Ara per ara només es pot endreçar de manera alfanumèrica. En la pestanya *Entries* podrem especificar que s'endrecin en el mateix ordre com apareixen en el text.

A la figura 18 mostrem la pestanya *Entries*:

Figura 18. Pestanya *Entries* de la pantalla per a inserir Índex/Taula de l'Open Office



En aquesta pestanya podem definir com serà l'entrada, és a dir, quins camps hi apareixeran, en quin ordre i en quin format. Aquesta selecció és individual per a cada tipus d'entrada. Per exemple, si seleccionem *Book* i veiem que té l'estructura:

Sh : Au , Ti , Ye

(és a dir, *Short Name* dos punts *Autor* coma *Títol* coma *Any*)

es generarà una bibliografia com la següent:

Bibliografia

Borges, 2004: Borges, Malte; Schumacher, Jorg; Redeker, Torsten, StarOffice 7, 2003

Si volem, podem canviar aquesta estructura segons les nostres preferències o necessitats. A més, en cada un dels camps li podem donar un format. Ara farem que l'entrada tingui l'autor en negreta, l'any entre parèntesis i el títol en cursiva. Forçarem que l'estructura sigui:

Au (Ye) Ti

i donarem l'estil de caràcter que volem en cada camp. Es generarà una bibliografia com la següent:

Bibliografia

Borges, Malte; Schumacher, Jorg; Redeker, Torsten (2003). *StarOffice 7*.

En aquesta mateixa pestanya també podem especificar que les entrades estiguin endreçades per l'ordre en què apareixen al document (marcant l'opció *Document position*) o bé endreçades alfabèticament (seleccionant *Content*). Amb *Sort Keys* podem afinar molt els criteris d'ordenació, ja que ens permet seleccionar fins a tres camps.

4.2. BibTeX: bibliografia per a LaTeX

LaTeX disposa d'un sistema de gestió de la bibliografia molt eficient. D'una banda, les entrades bibliogràfiques es guarden en uns fitxers de text amb extensió *.bib*. Les entrades d'aquests fitxers tenen l'aspecte següent:

```
@Article{ Mikheev-2002,
  title = "Periods, capitalized words, etc",
  author = "A Mikheev",
  journal = "Computational Linguistics",
  volume = "28",
  number = "3",
  year = "2002",
  pages = "239--318"
}

@TechReport{ Grefenstette-1994,
  author = "Gregory Grefenstette and Papso Tapanainen",
  title = "What is a word, What is a sentence? Problems of Tokenization",
  year = "1994",
  institution = "Xerox"
}

@Book{ Zipf-1949,
  title = "Human Behaviour and the Principle of Least Effort",
  author = "G.K. Zipf",
  year = "1949",
  address = "Cambridge",
  publisher = "Addison-Wesley"
}

@InProceedings{ Benitez-1998,
  author = "Laura Benitez and Sergi Cervell and Gerard Escudero and Monica Lopez and German Rigau and Mariona Taule",
  title = "Methods and Tools for Building the Catalan WordNet",
  url = "http://www.citebase.org/abstract?id=oai:arXiv.org:cmp-lg/9806009",
  year = "1998",
  booktitle = "In Proceedings of ELRA Workshop on Language Resources for European Minority Languages"
}
```

Per a introduir les referències bibliogràfiques al nostre document només hem de fer el següent (adjuntem un exemple de document mínim):

```
\documentclass{article}
\bibliographystyle{plain}
\begin{document}

Per a ampliar informació consulteu \cite{Mikheev-2002}, \cite{Grefenstette-1994}, \cite{Zipf-1949} i \cite{Benitez-1998}.

\bibliography{bibliografia}
\end{document}
```

Després de definir el tipus de document definim l'estil de bibliografia, en aquest cas *plain*. Per a introduir les referències bibliogràfiques podem fer servir `\cite{}` i el nom de la referència. En el punt on volem inserir la bibliografia fem servir la instrucció `\bibliography{}` i el nom del fitxer que conté les dades de la bibliografia en Bibtex. Per a processar aquest fitxer farem:

```
latex document.tex
bibtex document
```

Sovint s'ha de repetir aquest conjunt d'instruccions més d'una vegada. El document resultant es mostra a la figura 19.

Figura 19. Exemple de resultat de bibliografia amb LaTeX amb l'estil *plain*

Per ampliar informació consulteu [3], [2], [4] i [1].

Referències

- [1] Laura Benitez, Sergi Cervell, Gerard Escudero, Monica Lopez, German Rigau, and Mariona Taule. Methods and tools for building the catalan wordnet. In *In Proceedings of ELRA Workshop on Language Resources for European Minority Languages*, 1998.
- [2] Gregory Grefenstette and Papso Tapanainen. What is a word, what is a sentence? problems of tokenization. Technical report, Xerox, 1994.
- [3] A Mikheev. Periods, capitalized words, etc. *Computational Linguistics*, 28(3):239–318, 2002.
- [4] G.K. Zipf. *Human Behaviour and the Principle of Least Effort*. Addison-Wesley, Cambridge, 1949.

Si ara canviem simplement la línia on definim l'estil de la bibliografia per:

```
\bibliographystyle{alpha}
```

obtenim el següent, que es mostra a la figura 20.

Figura 20. Exemple de resultat de bibliografia amb LaTeX amb l'estil *alpha*

Per ampliar informació consulteu [Mik02], [GT94], [Zip49] i [BCE⁺98].

Referències

- [BCE⁺98] Laura Benitez, Sergi Cervell, Gerard Escudero, Monica Lopez, German Rigau, and Mariona Taule. Methods and tools for building the catalan wordnet. In *In Proceedings of ELRA Workshop on Language Resources for European Minority Languages*, 1998.
- [GT94] Gregory Grefenstette and Pappo Tapanainen. What is a word, what is a sentence? problems of tokenization. Technical report, Xerox, 1994.
- [Mik02] A Mikheev. Periods, capitalized words, etc. *Computational Linguistics*, 28(3):239–318, 2002.
- [Zip49] G.K. Zipf. *Human Behaviour and the Principle of Least Effort*. Addison-Wesley, Cambridge, 1949.

I si ara ho canviem per:

```
\bibliographystyle{abstract}
```

obtenim el resultat que es mostra a la figura 21.

Figura 21. Exemple de resultat de bibliografia amb LaTeX amb l'estil *abstract*

Per ampliar informació consulteu [Mikheev-2002], [Grefenstette-1994], [Zipf-1949] i [Benitez-1998].

Referències

- [Benitez-1998] Laura Benitez, Sergi Cervell, Gerard Escudero, Monica Lopez, German Rigau, and Mariona Taule. Methods and tools for building the catalan wordnet. In *In Proceedings of ELRA Workshop on Language Resources for European Minority Languages*, 1998.
- [Grefenstette-1994] Gregory Grefenstette and Pappo Tapanainen. What is a word, what is a sentence? problems of tokenization. Technical report, Xerox, 1994.
- [Mikheev-2002] A Mikheev. Periods, capitalized words, etc. *Computational Linguistics*, 28(3):239–318, 2002.
- [Zipf-1949] G.K. Zipf. *Human Behaviour and the Principle of Least Effort*. Addison-Wesley, Cambridge, 1949.

Hi ha molts estils disponibles i a més és relativament senzill crear-ne de nous. Hi ha alguns paquets que amplien encara més les possibilitats d'ajustar-se més a les necessitats dels usuaris. A continuació mostrem l'ús d'un d'aquests paquets:

```
\documentclass{article}
\bibliographystyle{chicago}
\usepackage{natbib}
\usepackage[utf8]{inputenc}
```

```
\usepackage[catalan]{babel}

\begin{document}

Per a ampliar informació consulteu el llibre de \citeauthor{Zipf-1949} de l'any
\citeyear{Zipf-1949}, i també \cite{Grefenstette-1994}, \citep{Mikheev-2002} i
\citet{Benitez-1998}.

\bibliography{bibliografia}

\end{document}
```

I el resultat que s'obté el podem observar a la figura 22.

Figura 22. Exemple de resultat de bibliografia amb LaTeX fent servir el paquet *natbib*

Per ampliar informació consulteu el llibre de Zipf de l'any 1949, i també Grefenstette and Tapanainen (1994), (Mikheev, 2002) i Benitez et al. (1998).

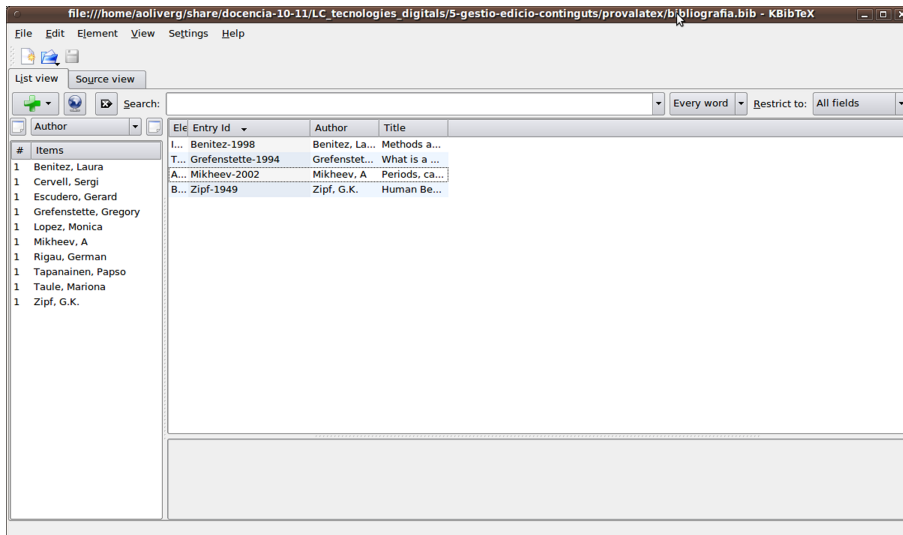
Referències

- Benitez, L., S. Cervell, G. Escudero, M. Lopez, G. Rigau, and M. Taule (1998). Methods and tools for building the catalan wordnet. In *In Proceedings of ELRA Workshop on Language Resources for European Minority Languages*.
- Grefenstette, G. and P. Tapanainen (1994). What is a word, what is a sentence? problems of tokenization. Technical report, Xerox.
- Mikheev, A. (2002). Periods, capitalized words, etc. *Computational Linguistics* 28(3), 239–318.
- Zipf, G. (1949). *Human Behaviour and the Principle of Least Effort*. Cambridge: Addison-Wesley.

Com veiem BibTeX és un sistema de gestió de bibliografia molt potent i fàcil de fer servir. Pot semblar una mica complicat mantenir els fitxers .bib de bases de dades bibliogràfiques, però per sort hi ha una sèrie d'eines que ens poden ajudar.

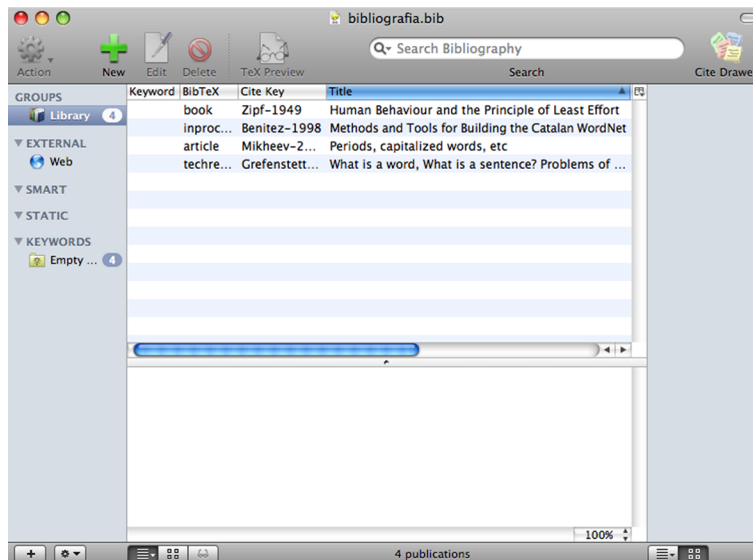
Una de molt eficient i que funciona sota Linux és KbibTex, que poden observar a la figura 23.

Figura 23. Gestor de bibliografia KbibTex per a Linux



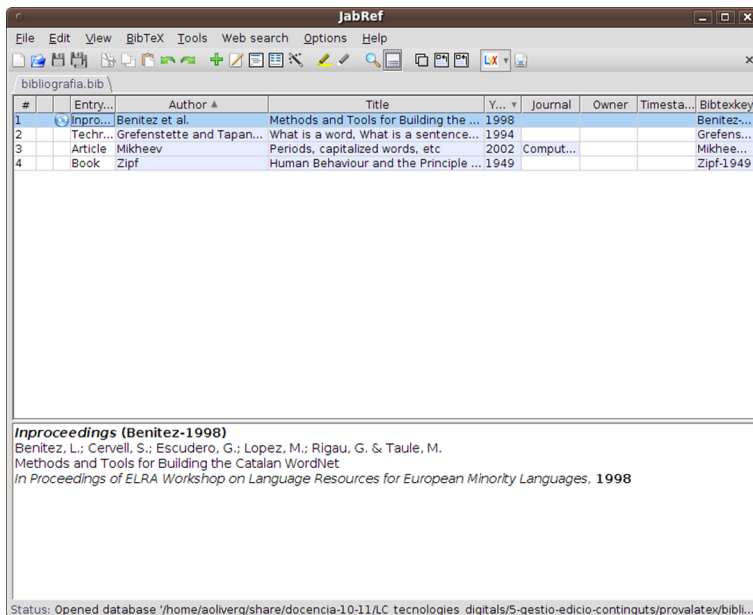
Per a Mac disposem de Bibdesk, que és una aplicació que s'instal·la per defecte quan instal·lem el LaTeX per a Mac (figura 24).

Figura 24. Gestor de bibliografia Bibdesk per a Mac



També podem optar per JabRef. Com que és una aplicació feta en Java pot funcionar sense problemes a Linux, Mac i Windows (figura 25).

Figura 25. Gestor de bibliografia JabRef per a Windows, Linux i Mac



4.3. Bibliografia en DocBook

DocBook disposa d'un seguit d'etiquetes pròpies que ens permeten definir una bibliografia d'un document. A continuació oferim un exemple d'article amb bibliografia en DocBook:

```
<?xml version="1.0" encoding="UTF-8"?>
<article lang="ca">
  <title>Un document d'exemple d'ús de bibliografia amb DocBook</title>
  <articleinfo>
    <author>
      <firstname>Antoni</firstname>
      <surname>Oliver</surname>
    </author>
    <date>23 de maig de 2010</date>
  </articleinfo>
  <abstract>
<para>Aquest article és una breu exemple a l'ús de la bibliografia amb DocBook</para>
</abstract>
  <sect1>
    <title>Introducció</title>
    <para>Aquest exemple de document serveix per a explicar alguns aspectes
      bàsics de DocBook <citation>TDG1999</citation> i les transformacions XSL
      <citation>REC-XSLT</citation>. </para>
  </sect1>

  <bibliography>
  <biblioentry>
  <abbrev>TDG1999</abbrev>
  <authorgroup>
```

```
<author><firstname>Norman</firstname><surname>Walsh</surname></author>
<author><firstname>Leonard</firstname><surname>Muellner</surname></author>
</authorgroup>
<title>DocBook</title>
<subtitle>The Definitive Guide</subtitle>
<pubdate>1999</pubdate>
<edition>1</edition>
<isbn>ISBN: 156592-580-7</isbn>
<pagenums>648</pagenums>
<releaseinfo>
<ulink url="http://www.oreilly.com/catalog/docbook/index.html">Oreilly's catalog
entry</ulink>. An online version <ulink url="http://www.oasis-
open.org/docbook/documentation/reference/html/docbook.html">is available</ulink>
</releaseinfo>
</biblioentry>

<biblioentry>
<abbrev>REC-XSLT</abbrev>
<editor><firstname>James</firstname><surname>Clark</surname></editor>
<title><ulink url="http://www.w3.org/TR/xslt">XSL Transformations (XSLT) Version 1.0</ulink>
</title>
<publishername>W3C Recommendation</publishername>
<pubdate>16 November 1999</pubdate>
</biblioentry>

</bibliography>

</article>
```

Si processem aquest document obtindrem una sortida com la de la figura 26.

Figura 26. Exemple de bibliografia en un document DocBook

Un document d'exemple d'ús de bibliografia amb DocBook

Oliver, Antoni

20 de maig de 2010

Índex

1	Introducció	1
2	Referències	1

Resum

Aquest article és una breu exemple a l'ús de la bibliografia amb DokBook

1 Introducció

Aquest exemple de document serveix per explicar alguns aspectes bàsics de DocBook [TDG1999] i les transformacions XSL [REC-XSLT].

2 Referències

[REC-XSLT] *XSL Transformations (XSLT) Version 1.0*, 16 November 1999.

[TDG1999] Walsh, Norman i Muellner, Leonard, *DocBook*, The Definitive Guide, 1999, 1, ISBN: 156592-580-7, 648, [Oreilly's catalog entry](#). An online version [is available](#) .

4.4. Zotero

Zotero és una extensió del navegador Firefox que permet fer reculls de fonts bibliogràfiques i fer cites bibliogràfiques. Es pot integrar fàcilment en diferents processadors de textos. En aquesta secció farem una breu introducció a aquesta eina.

Com que Zotero és una extensió del navegador Firefox el primer pas necessari serà instal·lar aquest navegador si encara no el tenim. Un cop instal·lat Firefox a l'ordinador podrem instal·lar Zotero.

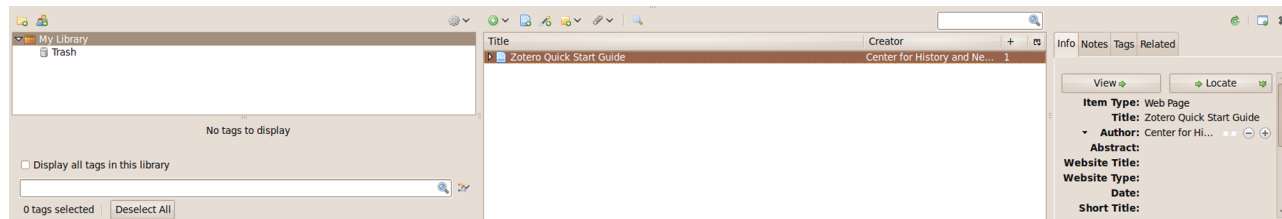
Instal·lació de Firefox i Zotero

Per a instal·lar el navegador Firefox, segueix les instruccions de <http://www.mozilla.com>.

Per instal·lar Zotero, aneu a <http://zotero.org> fent servir el Firefox i feu clic a *Download*. La instal·lació és molt senzilla i només cal anar seguint les instruccions que es presenten per pantalla.

Un cop instal·lat, apareixerà una icona que posa Zotero a la part baixa dreta del Firefox. Si fem clic en aquesta icona apareix una pantalla com la de la figura 27.

Figura 27. Pantalla principal del sistema de gestió de bibliografia Zotero



En aquesta pantalla (per ara buida) tindrem accés a totes les nostres referències bibliogràfiques. També podem afegir noves referències de manera manual fent clic a la icona verda amb una signe "+" a dins. Podrem escollir entre diversos tipus d'obres (llibres, articles, etc.) i posar les dades manualment.

La veritable potència de Zotero és que pot recollir automàticament les referències bibliogràfiques de la web. Si accedim a pàgines web de reculls bibliogràfics podrem recollir la informació de manera automàtica.

Webs recomanats

Algunes pàgines web que permeten recollir la informació de manera automàtica (i n'hi ha d'altres) són:

<http://scholar.google.com>

<http://citeseerx.ist.psu.edu/>

<http://socrates.stanford.edu>

<http://search.ebscohost.com>

També la biblioteca de la UOC, <http://biblioteca.uoc.edu>

Quan visitem alguna d'aquestes pàgines web i tenim oberta la pàgina d'una obra concreta, si a la dreta de la barra d'adreça del navegador apareix una icona com aquesta:



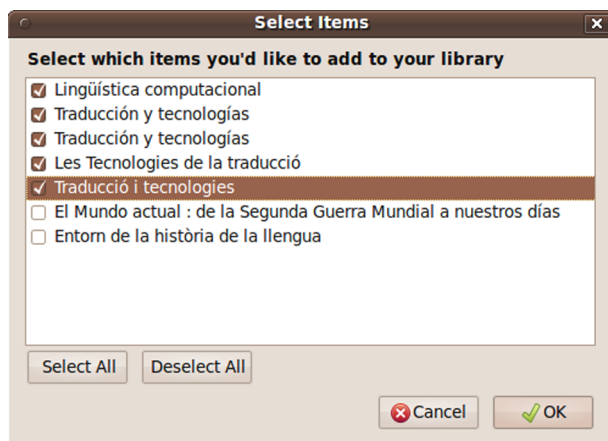
Fent clic sobre aquesta icona podrem introduir la referència a aquesta obra a la nostra base de dades de bibliografia.

Si quan visitem una d'aquestes pàgines web a la pantalla tenim una llista d'obres i a la dreta de la barra d'adreça del navegador apareix una icona com aquesta:



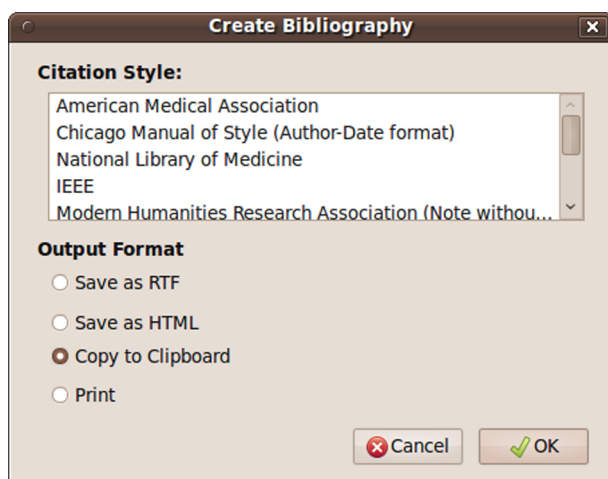
Fent clic sobre aquesta icona podrem importar totes les referències de les obres que hi apareixen, o bé seleccionar quines volem importar. Apareixerà una pantalla com la que es mostra a la figura 28, on podrem seleccionar quines entrades volem importar a la nostra base de dades de referències bibliogràfiques.

Figura 28. Selecció dels elements que es volen afegir a la biblioteca de referències bibliogràfiques de Zotero



Quan la nostra base de dades ja té algunes entrades podrem crear una bibliografia amb les entrades que seleccionem. Per a fer això fem clic a la icona Zotero del navegador. En la llista d'article seleccionem els que volem fer servir i fent clic sobre el botó dret del ratolí escollim l'opció *Create Bibliography from Selected Items*. Apareixerà una pantalla com la de la figura 29 que ens permetrà seleccionar l'estil de la bibliografia que volem generar.

Figura 29. Selecció de l'estil de bibliografia i el format de sortida de Zotero



En aquesta pantalla podem seleccionar el *Citation Style* i l'acció que volem fer. Si seleccionem, per exemple, *Copy to Clipboard*, podrem enganxar la bibliografia en el document.

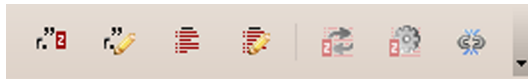
El resultat de fer un *Copy to Clipboard* podria ser aquest:

Gómez Guinovart, Javier. 2010. *Lingüística Computacional*. Barcelona: UOC Universitat Oberta de Catalunya.

Zotero permet una integració amb diversos processadors de textos, entre ells l'OpenOffice.

Un cop instal·lat el connector cal reiniciar també l'Open Office i llavors apareixerà en la barra d'eines les icones que es mostren a la figura 30.

Figura 30. Icones del connector de Zotero per a Open Office



Mitjançant aquestes icones podem (ordre d'esquerra a dreta):

- 1) Inserir una cita.
- 2) Editar una cita.
- 3) Inserir la bibliografia.
- 4) Editar la bibliografia.
- 5) Regenerar les cites i la bibliografia.
- 6) Establir les preferències del document.
- 7) Eliminar els codis de camp.

Zotero i Open Office

Podeu consultar com fer la integració de Zotero amb Open Office i instal·lar els requeriments a http://www.zotero.org/support/word_processor_plugin_installation

5. Bases de dades documentals i Google Desktop

Una base de dades documental és un recull de documents organitzats i indexats de manera que sigui fàcil i ràpid fer cerques sobre aquests documents. Els documents poden ser de diferents tipus i estar en diferents formats. El sistema ha de ser capaç de trobar documents que siguin rellevants per a la cerca que fa l'usuari.

Aquest tipus de bases de dades es fan servir en molts àmbits. Un típic és l'àmbit legislatiu, on els documents són lleis i disposicions del Govern. A la figura 31 es pot observar l'aspecte de la interfície de cerca de la base de dades documental del Diari Oficial de la Generalitat de Catalunya:

Figura 31. Interfície de cerca de la base de dades documental del Diari Oficial de la Generalitat de Catalunya

The screenshot shows the search interface for the DOGC (Diari Oficial de la Generalitat de Catalunya) database. At the top left is the DOGC logo and the text 'Diari Oficial de la Generalitat de Catalunya'. Below it is a link 'RECCERCA A LA BASE DE DADES'. On the right, there are buttons for 'Buscar', 'Restaurar', and 'Aplicar valors'. A message states: 'Els documents continguts a la Base de dades estan en format html i pdf. En el primer cas, porten les correccions d'errada introduïdes en el text; no en el segon cas, ja que reproduïxen la imatge impresa del DOGC. Obviament, les correccions d'errada apareixen en el pdf del DOGC on s'hagin publicat. Les modificacions produïdes per altres normes no estan introduïdes en els textos, sinó que cal consultar la taula de vigències de cada document.' The search form includes fields for 'Titul document', 'Paraules del text', 'Número DOGC', 'Data DOGC (dd.mm.aaaa)' with a dropdown set to 'Igual', 'Organisme', 'Tipus document', 'Data document (dd.mm.aaaa)' with a dropdown set to 'Igual', 'Número disposició', 'Número de control', 'Descriptors temàtics', and 'Descriptors geogràfics'. There are also checkboxes for 'Secció' (Disposicions, Concursos i anuncis, Administració local) and 'Càrrecs i personal' (Anuncis diversos, Administració de Justícia). On the right, there are sorting options for 'ORDENACIÓ' (Asc./Desc.) and a 'Borrar criterios' button. At the bottom, there is an 'Operador' dropdown and a note: '*Informació disponible a partir del DOGC núm. 785 de 02/01/1987'.

Accés a la base de dades documental

Podeu accedir a la base de dades documental del Diari Oficial de la Generalitat de Catalunya a <http://www.gencat.cat/dogc/>.

El nostre ordinador segurament també és un repositori important de documents. Pot resultar interessant disposar de funcionalitats de cerca sobre aquests documents. Els sistemes operatius disposen de funcionalitats bàsiques de cerca d'arxius, en les quals es pot cercar per contingut. Aquestes funcionalitats acostumen a ser insuficients. Normalment no es duu a terme una indexació prèvia dels documents i per aquest motiu la cerca acostuma a ser molt lenta.

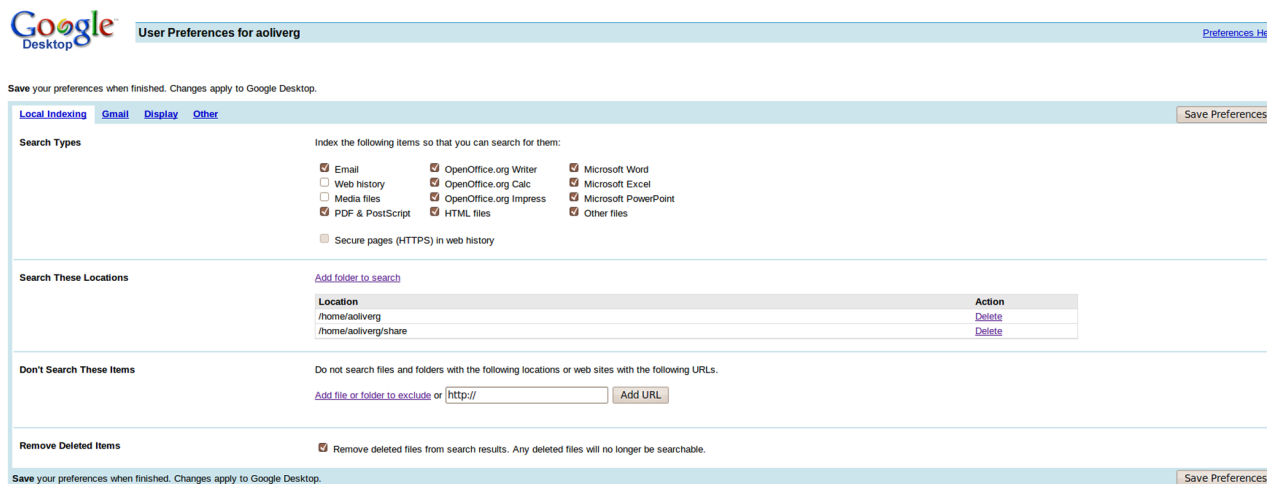
Podem simular un sistema de base de dades documental al nostre ordinador mitjançant Google Desktop. Aquesta aplicació realitza una indexació del contingut del nostre ordinador i permet fer cerques al nostre ordinador com si féssim cerques a Internet.

Google Desktop

Podeu descarregar Google Desktop a <http://desktop.google.com>. Hi ha versions per a Windows, Linux i Mac.

Si decidiu provar Google Desktop, el que heu de fer immediatament després d'instal·lar-lo és modificar les preferències i ajustar-les a les vostres necessitats. Si no ho feu, indexarà absolutament tot el contingut del vostre ordinador. Per a ajustar les preferències heu de fer clic al botó dret del ratolí mentre el teniu situat sobre la icona de Google Desktop i seleccionar *Preferences*. Apareixerà al vostre navegador una pantalla com la que es mostra a la figura 32.

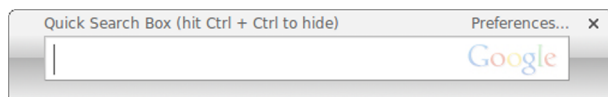
Figura 32. Pantalla de selecció de preferències de Google Desktop



En aquesta pantalla podeu seleccionar el tipus de documents i els directoris que s'indexaran. Si teniu un compte de Gmail podeu fer que indexi també els missatges del vostre compte de correu i que apareguin en la llista de resultats. Quan instal·leu l'aplicació i ajusteu les preferències comença el procés d'indexació. L'aplicació està dissenyada perquè aquest procés d'indexació no afecti massa al rendiment de la màquina i no arribi a bloquejar-se. Per aquest motiu, el procés pot trigar força estona. Les cerques que feu justament després d'instal·lar l'aplicació donaran resultats molt parcials.

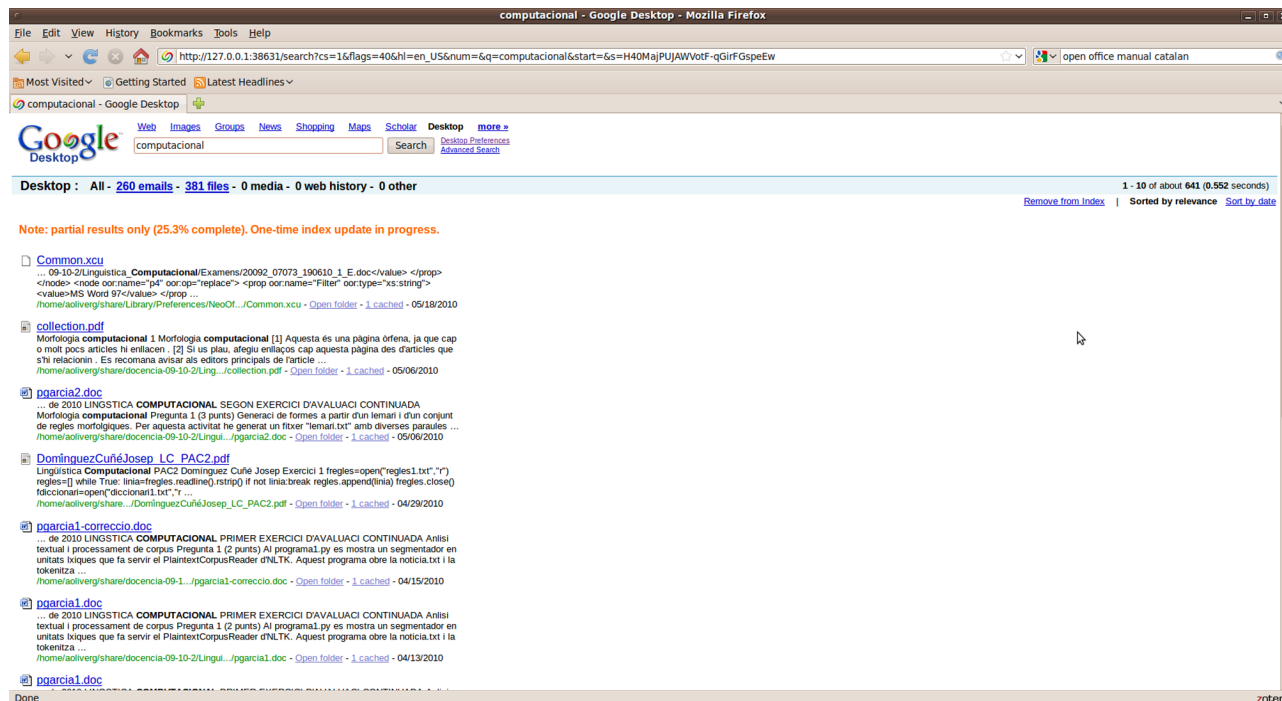
Per a fer cerques podeu fer clic a la Icona de Google Desktop i posar els criteris de cerca en la pantalla que hi apareix (figura 33).

Figura 33. Pantalla de cerca de Google Desktop



Si introduïm un terme de cerca a la pantalla anterior obtenim un seguit de resultats (figura 34).

Figura 34. Resultat d'una cerca amb Google Desktop



També és possible dur a terme cerques combinades al nostre ordinador i a Internet.

Aquesta eina pot ser molt interessant, però és extremadament important configurar-la segons les nostres necessitats. Cal evitar que indexi tot el contingut del nostre ordinador, ja que els fitxers d'índexs poden arribar a ser força grans.

6. Sistemes de gestió de continguts

El concepte de sistema de gestió de continguts agrupa a una sèrie d'aplicacions de programari que permeten generar continguts de manera col·laborativa, fer la gestió de versions i modificacions i fer la publicació d'aquests documents. Molts sistemes d'aquest tipus estan pensats per a gestionar continguts web, tot i que hi ha sistemes adaptats a altres formats.

6.1. Gestors de contingut web

Els gestors de contingut web (*Web Content Management Systems* – WCM, WCMS o Web CMS) són sistemes de gestió de continguts implementats com a aplicacions web que serveixen per a crear i gestionar contingut web. Aquests sistemes faciliten la creació de continguts, el control sobre aquests continguts, l'edició i les funcions de manteniment del lloc web.

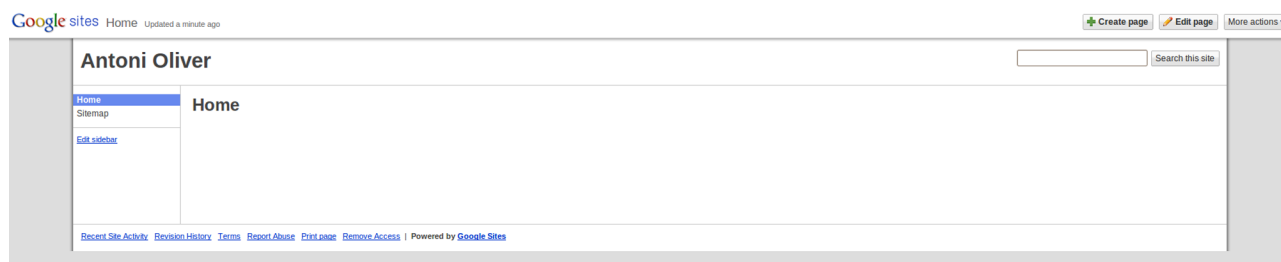
Aquests tipus de sistemes permeten que usuaris amb molt pocs coneixements de llenguatges de programació i llenguatges de marcatge siguin capaços de gestionar el contingut de llocs web amb facilitat.

La instal·lació i la configuració d'aquests sistemes requereixen força coneixements tècnics, però en canvi el seu ús acostuma a ser molt senzill.

Web recomanat

Per a tenir una idea de com funciona un sistema de gestió de contingut web podeu accedir a <http://sites.google.com>. Si disposeu d'un compte de Google podeu fer servir aquest servei per a crear d'una manera molt fàcil la vostra pròpia pàgina web. En la mateixa adreça podeu trobar totes les instruccions necessàries per a poder desenvolupar el vostre lloc web.

Figura 35. Exemple de creació d'un lloc web amb Google Sites



6.2. Blogs

Un blog és una mena de diari interactiu que es publica a Internet i està dissenyat de manera que cada apunt tingui una data de publicació i que es pugui seguir d'una manera temporal tot el que s'ha publicat i editat. Normalment les entrades d'un blog es mostren endreçades per data de publicació i generalment

en ordre cronològic invers, és a dir, les més noves es mostren primer. L'autor del blog pot donar permisos perquè els lectors deixin comentaris que l'autor mateix pot contestar.

És una manera molt senzilla de publicar informació a la Web i hi ha disponibles molts serveis gratuïts que permeten crear blogs de manera molt senzilla.

Crear un bloc

Podeu crear el vostre blog amb <http://www.blogger.com>.

Figura 36. Exemple de blog creat amb Blogger



6.3. Wikis

Un wiki és un lloc web col·laboratiu que pot ser editat de manera molt senzilla, fàcil i ràpida. Un dels wikis més coneguts és la Wikipedia. Aquesta eina es pot fer servir amb èxit per a l'escriptura col·laborativa de qualsevol tipus de document.

La instal·lació i configuració d'un wiki pot ser una tasca relativament complicada. Hi ha una sèrie de serveis que permeten crear wikis, entre els quals es pot destacar el Wikispaces.

Figura 37. Fragment d'un wiki creat amb Wikispaces



6.4. Knols

Knol és un servei de Google que permet escriure unitats de coneixement i publicar-les fàcilment. Per a poder escriure un nou Knol s'ha de disposar d'un compte de Google. La resta d'usuaris poden fer revisions i comentaris sobre el Knol.

Escriure un Knol

Podeu accedir al servei de Knols de Google per a escriure el vostre Knol a <http://knol.google.com>.

Figura 38. Exemple de Knol

The screenshot shows a Knol page titled "Languages of the World" by the Knol Foundation. The page content includes:

- Title:** Languages of the World
- Text:** "There are 6,909 living languages throughout the world. Following is a list in the order of the largest numbers of native speakers in the world. 1. Mandarin Chinese 2. Spanish 3. English 4. Hindu/Urdu 5. Arabic 6. Bengali 7. Portuguese 8. Russian 9. Japanese"
- Actions:** Link, Citation, Email, Print, Favorite, Collect this page
- Related Content:** English Language Translator (with a row of flags)
- Rating:** Article rating: 5 stars (21 Ratings); Your rating: No rating
- Footer:** Moderated collaboration, All Rights Reserved, Top Pick Knol Award, Version: 87, Last edited: May 5, 2010 1:54 AM.

The right sidebar features a "Write a knol" button and profiles for Knol Foundation, Randall Kleinet, and Christine Kleinet, all marked as "Verified".

Resum

En aquest mòdul hem repassat diverses eines i tècniques per a la redacció, gestió i publicació de contingut. L'ús d'aquestes eines pot simplificar molt la redacció de documents, especialment si volem fer treballs col·laboratius entre diversos autors. Per sort hi ha una gran quantitat d'eines per triar i moltes d'elles són de programari lliure o d'ús gratuït.

Hem dedicat una especial atenció a la gestió i creació de bibliografies per la gran importància que té aquest aspecte en els treballs acadèmics.

Bibliografia

Anon (2002). *Content Management Systems*. Arden House: Glasshaus.

Boiko, Bob (2005). *Content Management Bible* (2a. ed.). Indianapolis, Ind.: Wiley.

Brockmeier, Joe (2001). *Docbook Publishing*. Roseville: Prima Tech.

Delgado Cabrera, José María (2009). *OpenOffice.org 3.0*. Madrid: Anaya Multimedia.

Meloni, Julie C. (2005). *Plone Content Management Essentials*. Indianapolis, Ind.: Sams.

Mittelbach, Frank (2004). *The LaTeX Companion* (2a. ed.). Boston: Addison-Wesley.

Walsh, Norman i Hamilton, Richard L. (2010). *DocBook 5: The Definitive Guide*. O'Reilly Media, Inc., April 15.

White, Martin S. (2005). *The Content Management Handbook*. Londres: Facet Publishing.

