

Modélisation et interopérabilité :

Semaine 41, cours 4

Benoît Valiron <benoit.valiron@monoidal.net>

<http://inf356.monoidal.net/>

Partiel en salle : mardi 10 novembre
12h – 15h

Projets

- Vous avez tous dû recevoir une copie
- Sinon, c'est sur le site du cours
- Instructions :
 - Tout seul ou en binôme
 - À rendre jeudi 3 décembre
 - Deux questions sur Xpath et XSLT seront rajoutés à la fin du mois
 - Mardi prochain je vous demanderai les noms des binômes (ou des uni-ôme), ainsi que la ligne de train choisie.
 - Indépendant des TPs, mais nous sommes là pour répondre à vos questions !
 - COMMENCEZ, ET POSEZ VOS QUESTIONS SI VOUS EN AVEZ !

Validation avec les DTDs

Appel général :

- Pour les fils
- Caractère des fils (obligatoire...) dans le format du père

`<!ELEMENT elt format>`

Appel général :

- Pour l'attribut lui-même
- Caractère explicite

`<!ATTLIST elt att format caractère>`

Fils de type texte :

`<!ELEMENT elt (#PCDATA)>`

Attribut de type texte :

`<!ATTRIBUT ... .. CDATA ...>`

CDATA = “character data” ← ici, < et > n'ont pas de sens spécial...

PCDATA = “parsed character data” ← ici, < et > ont un sens spécial

Section CDATA en XML : `<![CDATA[ ... ]]>` ← pas de sens spécial pour < et >

Aujourd'hui, fin du tour du
propriétaire.

Les formats XML...

- Peuvent définir une arborescence avec une racine unique :
 - XHTML → Racine `<html> ... </html>`
 - SVG → Racine `<svg> ... </svg>`
- Peuvent juste définir un ensemble de mots dissociés :
 - Dublin Core, définissant des balises `<creator>`, `<date>`, ...
- Mais peuvent aussi juste définir des attributs.

XLink

- Espace de nom : <http://www.w3.org/1999/xlink>
- But : attacher des liens aux documents XML.
- Méthode :
 - Attacher les attributs href et type du langage Xlink à un élément.
 - href → Lien vers une resource
 - type → “simple” (mais il y en a d'autres)
 - Il faut définir un préfixe puisque **les attributs sans préfixe sont toujours sans espaces de noms.**

```
<roman xmlns:xlink="http://www.w3.org/1999/xlink"
      xlink:type="simple"
      xlink:href="http://archive.org/pub/etext/etext93/wizoz10.txt">
  <titre>The Wonderful Wizard of Oz</titre>
  <auteur>L. Frank Baum</auteur>
  <année>1900</année>
</roman>
```

Le préfixe spécial xml

- Un préfixe spécial ne nécessite pas de définition, le préfixe `xml`
- Deux attributs spéciaux :
 - `xml:lang` → Langue (humaine) utilisée
 - `xml:space` → Les espaces sont-ils significatifs ?
- Se doivent d'être définis dans la DTD, le cas échéant.

Formats de documents texte

- XHTML est spécialisé pour l'affichage pour le web.
- D'autres formats existent :
 - Polyvalent : Docbook
 - Orientés impression papier :
 - XSL-FO
 - Suites bureautiques :
 - ODF : OpenOffice
 - OOXML : Microsoft Office

Docbook

- Origine : format pour de la documentation technique.
- Sa définition comme standard par OASIS (*Organization for the Advancement of Structured Information Standard*) a généralisé son utilisation.
- Documentation du format sur le site <http://www.docbook.org>

Document Docbook

```
<!DOCTYPE xxx PUBLIC "-//OASIS//DTD DocBook XML V4.5//EN"  
    "http://www.oasis-open.org/docbook/xml/4.5/docbookx.dtd">
```

- Livres, avec la racine <book> ... </book>
 - Découpé en chapitre
 - Page de titre
 - Préface
- Articles, avec la racine <article> .. </article>
 - En un seul tenant, pas de sauts de page
 - Pas de chapitres
 - Exemple : le polycopié du TD 2 écrit en format article de DocBook

Example

```
<?xml version="1.0" encoding="utf-8"?>
<article>
  <articleinfo>
    <title>TD-2 : Espaces de noms</title>
    <author>
      <firstname>Benoît</firstname>
      <surname>Valiron</surname>
    </author>
  </articleinfo>
  <sect1>
    <title>Déclaration d'un espace de noms</title>
    <sect2>
      <title>Un titre de sous-section</title>
      <para>
        Texte de <emphasis>l'exercice 1</emphasis>.
      </para>
    </sect2>
  </sect1>
  <sect1>
    <title>Appartenance à un espace de noms</title>
    <para>
      <orderedlist>
        <listitem><para>Question 2.1</para></listitem>
        <listitem><para>Question 2.2</para></listitem>
        <listitem><para>Question 2.3</para></listitem>
      </orderedlist>
    </para>
  </sect1>
</article>
```

Principaux éléments pour un article

- Cinq niveaux hiérarchique <sect1> à <sect5>
- Paragraphes encadrés par <para> ... </para>
- Listes
 - Ordonnées : orderedlist
 - Non-ordonnées : itemizedlist
 - Items des listes : listitem
- Citations : blockquote
- Morceau de code (DTD, XML, programme...) : programlisting
- Dans le texte :
 - Emphase : emphasis
 - Liens hypertextes : ulink
 - Code : code

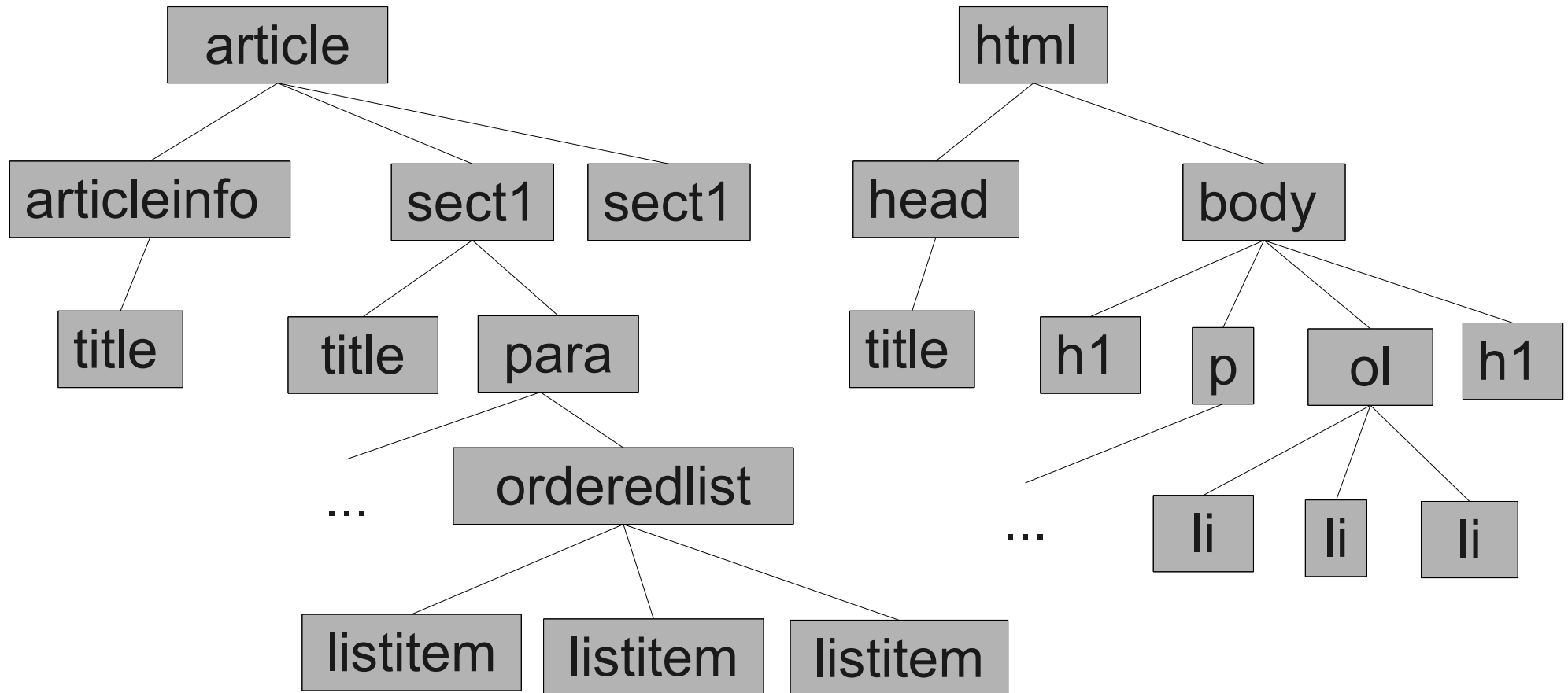
Exemple : programlisting

```
<?xml version="1.0" encoding="utf-8"?>
<article>
  <articleinfo>
    <title>TD-2 : Espaces de noms</title>
  </articleinfo>
  <sect1>
    <title>Une DTD pour RSS</title>
    <para>
      Un flot de données peut être encodé en format RSS.
      Une DTD simplifiée pour ce format est la suivante.
      <programlisting><![CDATA[<!ELEMENT rss (channel+)>
<!ELEMENT channel (title, description, item*)>
<!ELEMENT title (#PCDATA)>
<!ELEMENT description (#PCDATA)>
<!ELEMENT item (title|pubDate|description)*>
<!ELEMENT pubDate (#PCDATA)>]]></programlisting>
      Chaque <code>channel</code> correspond à une
      liste de données, avec un titre, une
      description et les données, une
      par <code>item</code>.
    </para>
  </sect1>
</article>
```

Comparaison avec XHTML

```
<?xml version="1.0" encoding="utf-8"?>    <?xml version="1.0" encoding="utf-8"?>
<article>                                <html>
  <articleinfo>                          <header>
    <title>TD-2 : Espaces de noms</title>    <title>TD-2 : Espaces de noms</title>
    <author>                                </header>
      <firstname>Benoît</firstname>          <body>
      <surname>Valiron</surname>            <h1>Déclaration d'un espace de noms</h1>
    </author>                                <p>
  </articleinfo>                            Du texte <em>important</em>.
  <sect1>                                    </p>
    <title>Un titre</title>                <ol>
    <para>                                    <li>Question 2.1</li>
    Du texte <emphasis>important</emphasis>    <li>Question 2.2</li>
    <orderedlist>                            <li>Question 2.3</li>
      <listitem>                                </ol>
        <para>Q 2.1</para>                    <h1>Autre titre</h1>
      </listitem>                                ...
      <listitem>                                </body>
        <para>Q 2.2</para>                    </html>
      </listitem>
      <listitem>
        <para>Q 2.3</para>
      </listitem>
    </orderedlist>
  </para>
</sect1>
<sect1> Autre section... </sect1>
</article>c
```

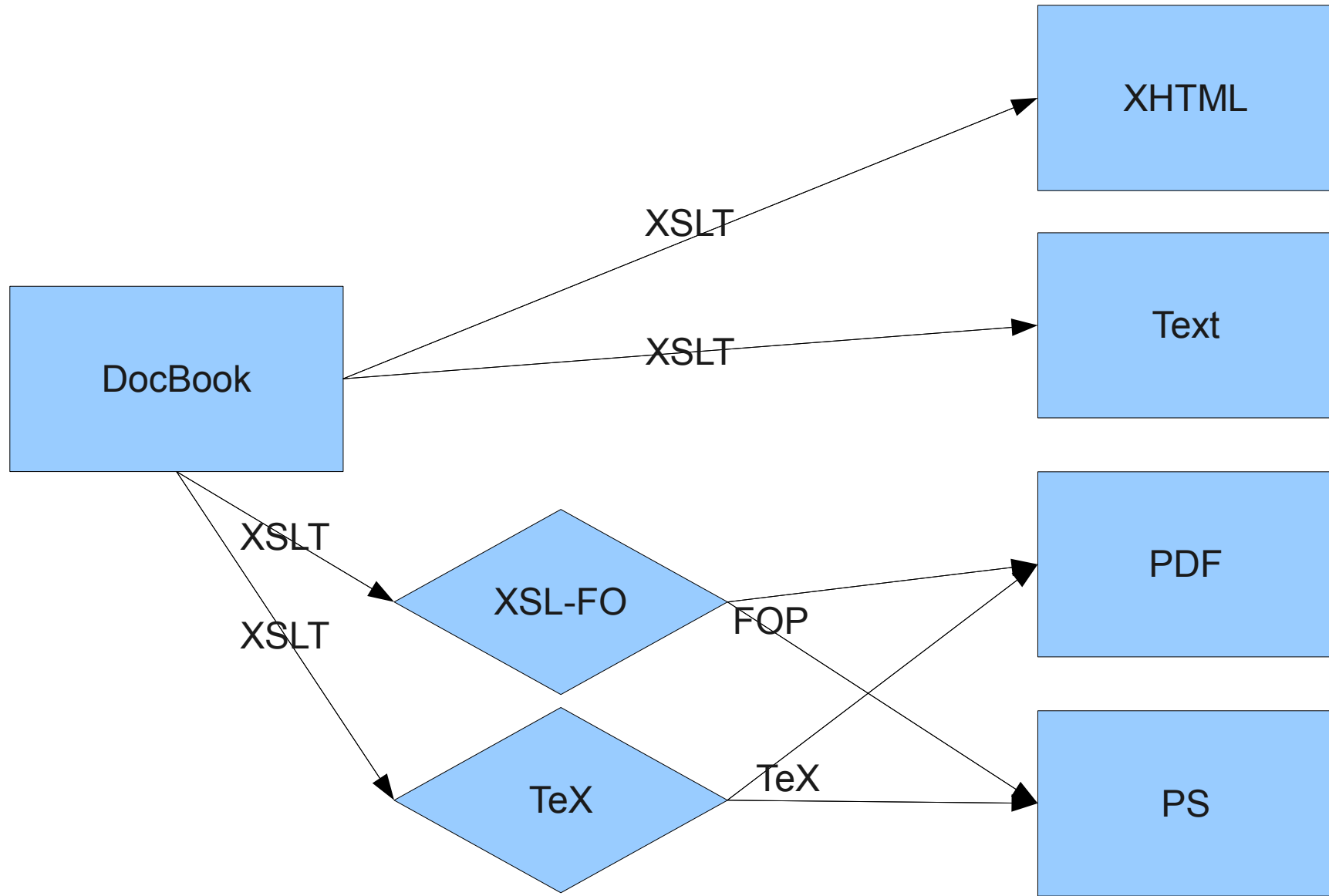
Comparaison avec XHTML



Extensions

- Docbook vient avec de nombreuses possibilités:
 - Table des matières
 - Index
 - Format pour de la bibliographie
 - ...
- Le format est modulaire : nul besoin de tout connaître pour produire un document.

Conversion



Résumé

- Docbook est orienté utilisateur humain
- Il est possible sans difficulté de concevoir un document docbook en utilisant un éditeur de texte de type Notepad+, emacs ou gedit (le TD 2 a été créé comme ça).
- Docbook est modulaire : de nombreuses fonctionnalités peuvent être ajoutée mais n'entrave pas l'écriture de document simple.
- Un fichier docbook peut être transformé en de nombreux formats.

XSL-FO

- Recommandation du W3C depuis 2001.
- Format XML pour décrire un document imprimé : Taille du papier, marges, ...
- Espace de noms :
<http://www.w3.org/1999/XSL/Format>

XSL-FO : Format

<root>

<layout-master-set>

<simple-page-master master-name="couverture" ...>

<!-- page unique : description de la couverture -->

</simple-page-master>

<page-master-sequence master-name="contenu" ...>

<!-- description physique des pages internes -->

</page-master-sequence>

</layout-master-set>

<page-sequence master-reference="couverture">

<!-- Texte de la couverture -->

</page-sequence>

<page-sequence master-reference="contenu">

<!-- Texte du livre -->

</page-sequence>

</root>

XSL-FO : Mise en forme

- Simple-page-master

```
<simple-page-master master-name="contenu"  
  page-height="11cm" page-width="8cm"  
  margin-top="1cm" margin-bottom="1cm"  
  margin-left="1cm" margin-right="1cm">  
  <region-body region-name="contenu-region" />  
</simple-page-master>
```

- Page-sequence

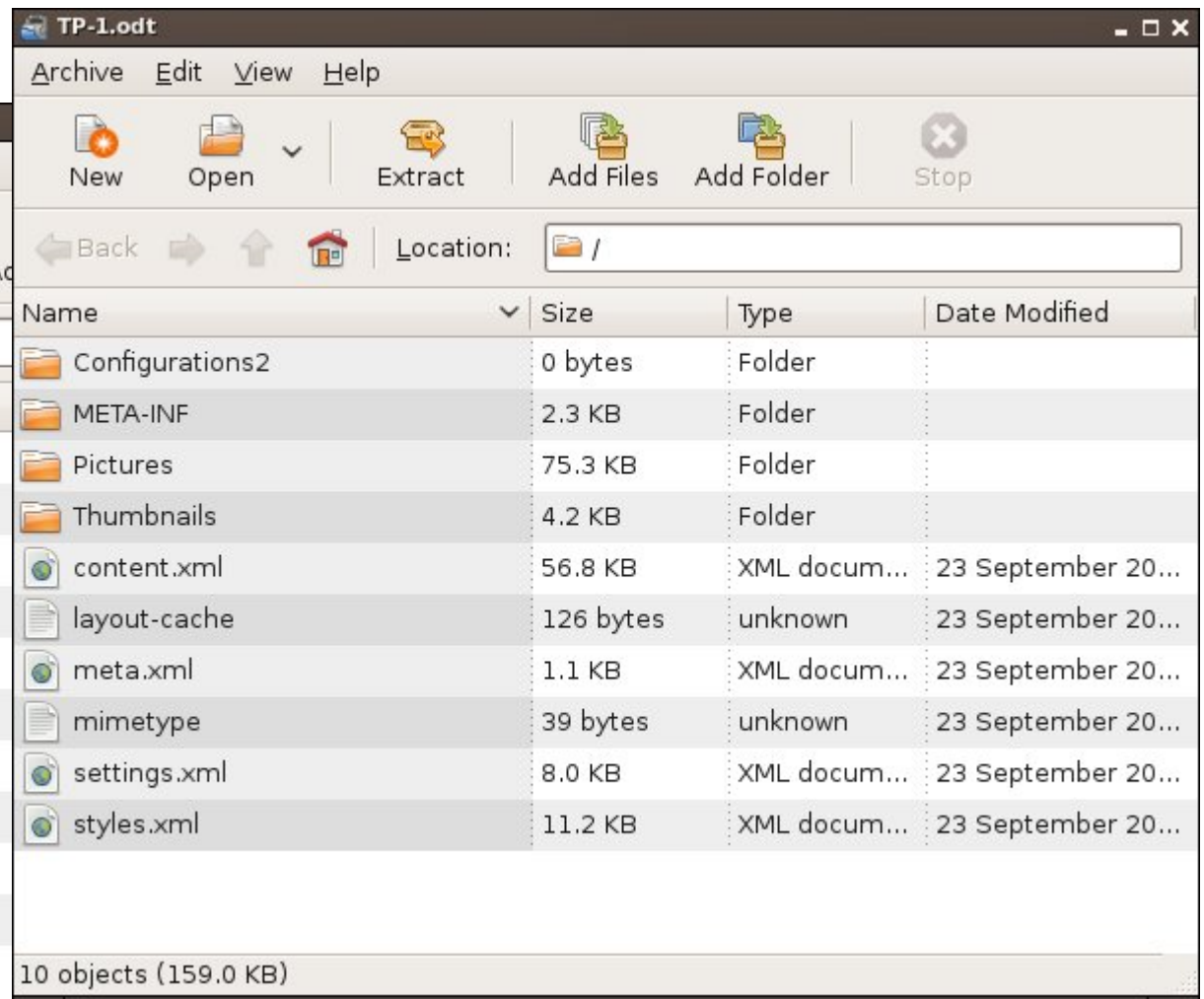
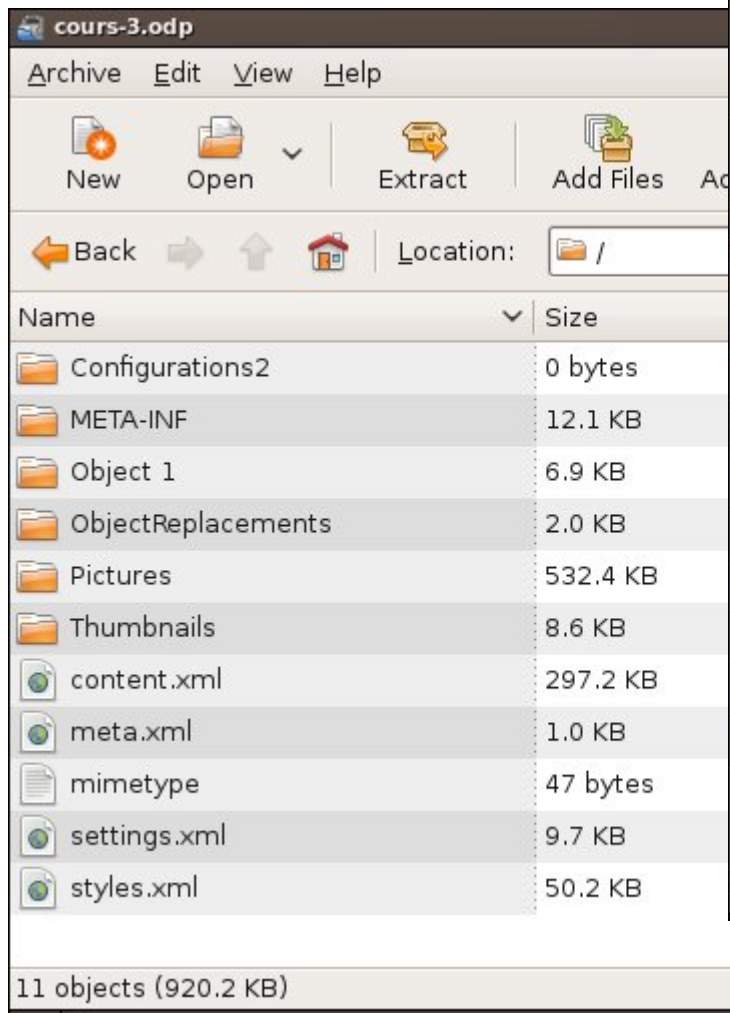
```
<page-sequence master-reference="contenu">  
  <flow flow-name="contenu-region">  
    <block font-family="Helvetica" font-size="20pt"  
      font-weight="bold" text-align="center">  
      Un titre  
    </block>  
    <block font-family="NewRoman" font-size="12pt"  
      font-weight="normal" text-align="left">  
      Du texte  
    </block>  
  </flow>  
</page-sequence>
```

Commentaire

- XSL-FO est particulièrement verbeux. Ceci dit, la structure du document est suffisamment simple pour pouvoir écrire un canevas et s'en servir pour formater du texte.
- Il est spécialement utile dans le cas de transformations automatiques, par exemple en partant de DocBook.
- Son usage principal est la production de PDF au niveau des serveurs de contenu.

ODF et OpenOffice

- Un fichier OpenOffice est un fichier compressé au format zip.



Dedans...

- Du XML !
- En particulier :
 - content.xml → le contenu du document
 - styles.xml → la description des styles utilisés
 - meta.xml → méta-données : titre, auteur, ...
 - settings.xml → les réglages de la dernière sauvegarde (niveau de zoom, position du curseur...)

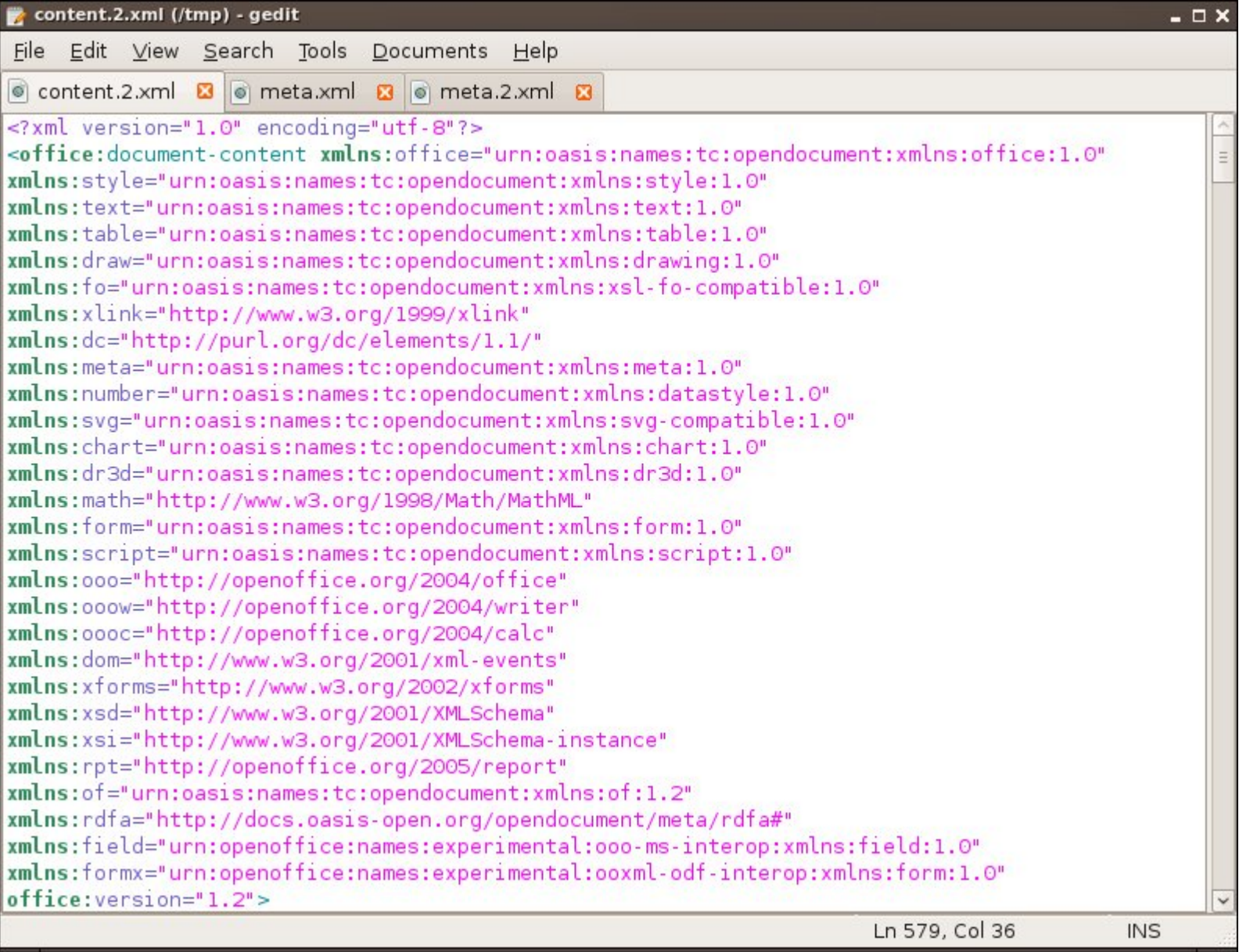
content.xml

- Contenu différent en fonction du type de document : texte, tableur, présentation...
- Élément racine toujours le même :
`office:document-content`
- Espaces de noms:
 - `urn:oasis:names:tc:opendocument:xmlns:text:1.0`
 - `urn:oasis:names:tc:opendocument:xmlns:presentation:1.0`

content.xml : Structure

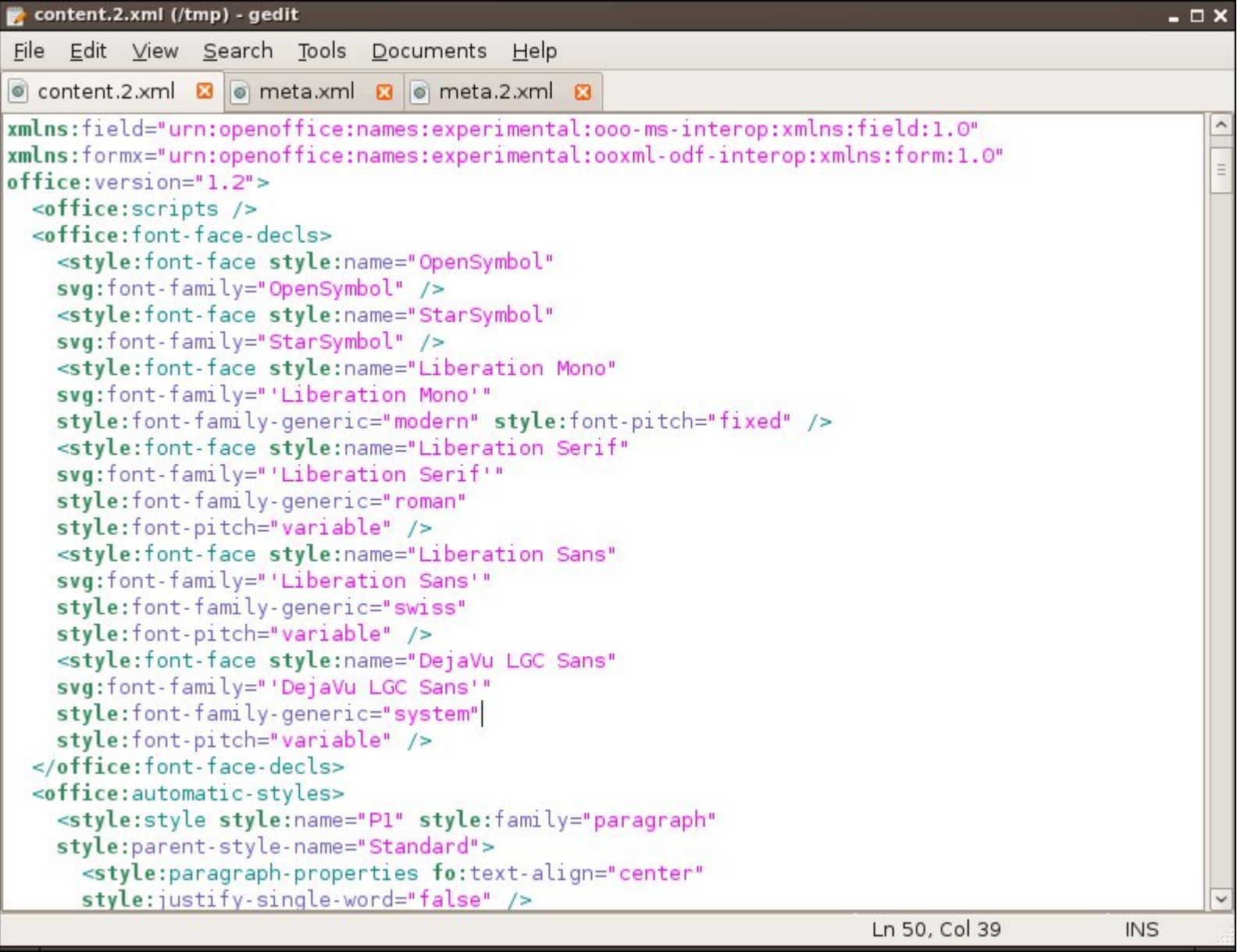
```
<office:document-content
  xmlns:office="urn:oasis:names:tc:opendocument:xmlns:office:1.0"
  xmlns:style="urn:oasis:names:tc:opendocument:xmlns:style:1.0"
  xmlns:text="urn:oasis:names:tc:opendocument:xmlns:text:1.0"
  xmlns:table="urn:oasis:names:tc:opendocument:xmlns:table:1.0"
  xmlns:draw="urn:oasis:names:tc:opendocument:xmlns:drawing:1.0"
  xmlns:fo="urn:oasis:names:tc:opendocument:xmlns:xsl-fo-compatible:1.0"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  ...>
<office:scripts>
  <!-- Scripts définis -->
</office:scripts>
<office:font-face-decls>
  <style:font-face ...>
    <!-- Polices utilisées -->
</office:font-face-decls>
<office:automatic-styles>
  <style:style ...>
    <!-- Styles internes -->
</office:automatic-style>
<office:body>
  <!-- Corps du document -->
</office:body>
```

content.xml : exemple



```
<?xml version="1.0" encoding="utf-8"?>
<office:document-content xmlns:office="urn:oasis:names:tc:opendocument:xmlns:office:1.0"
xmlns:style="urn:oasis:names:tc:opendocument:xmlns:style:1.0"
xmlns:text="urn:oasis:names:tc:opendocument:xmlns:text:1.0"
xmlns:table="urn:oasis:names:tc:opendocument:xmlns:table:1.0"
xmlns:draw="urn:oasis:names:tc:opendocument:xmlns:drawing:1.0"
xmlns:fo="urn:oasis:names:tc:opendocument:xmlns:xsl-fo-compatible:1.0"
xmlns:xlink="http://www.w3.org/1999/xlink"
xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:meta="urn:oasis:names:tc:opendocument:xmlns:meta:1.0"
xmlns:number="urn:oasis:names:tc:opendocument:xmlns:datastyle:1.0"
xmlns:svg="urn:oasis:names:tc:opendocument:xmlns:svg-compatible:1.0"
xmlns:chart="urn:oasis:names:tc:opendocument:xmlns:chart:1.0"
xmlns:dr3d="urn:oasis:names:tc:opendocument:xmlns:dr3d:1.0"
xmlns:math="http://www.w3.org/1998/Math/MathML"
xmlns:form="urn:oasis:names:tc:opendocument:xmlns:form:1.0"
xmlns:script="urn:oasis:names:tc:opendocument:xmlns:script:1.0"
xmlns:ooo="http://openoffice.org/2004/office"
xmlns:ooow="http://openoffice.org/2004/writer"
xmlns:oooc="http://openoffice.org/2004/calc"
xmlns:dom="http://www.w3.org/2001/xml-events"
xmlns:xforms="http://www.w3.org/2002/xforms"
xmlns:xsd="http://www.w3.org/2001/XMLSchema"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:rpt="http://openoffice.org/2005/report"
xmlns:of="urn:oasis:names:tc:opendocument:xmlns:of:1.2"
xmlns:rdfa="http://docs.oasis-open.org/opendocument/meta/rdfa#"
xmlns:field="urn:openoffice:names:experimental:ooo-ms-interop:xmlns:field:1.0"
xmlns:formx="urn:openoffice:names:experimental:ooxml-odf-interop:xmlns:form:1.0"
office:version="1.2">
```

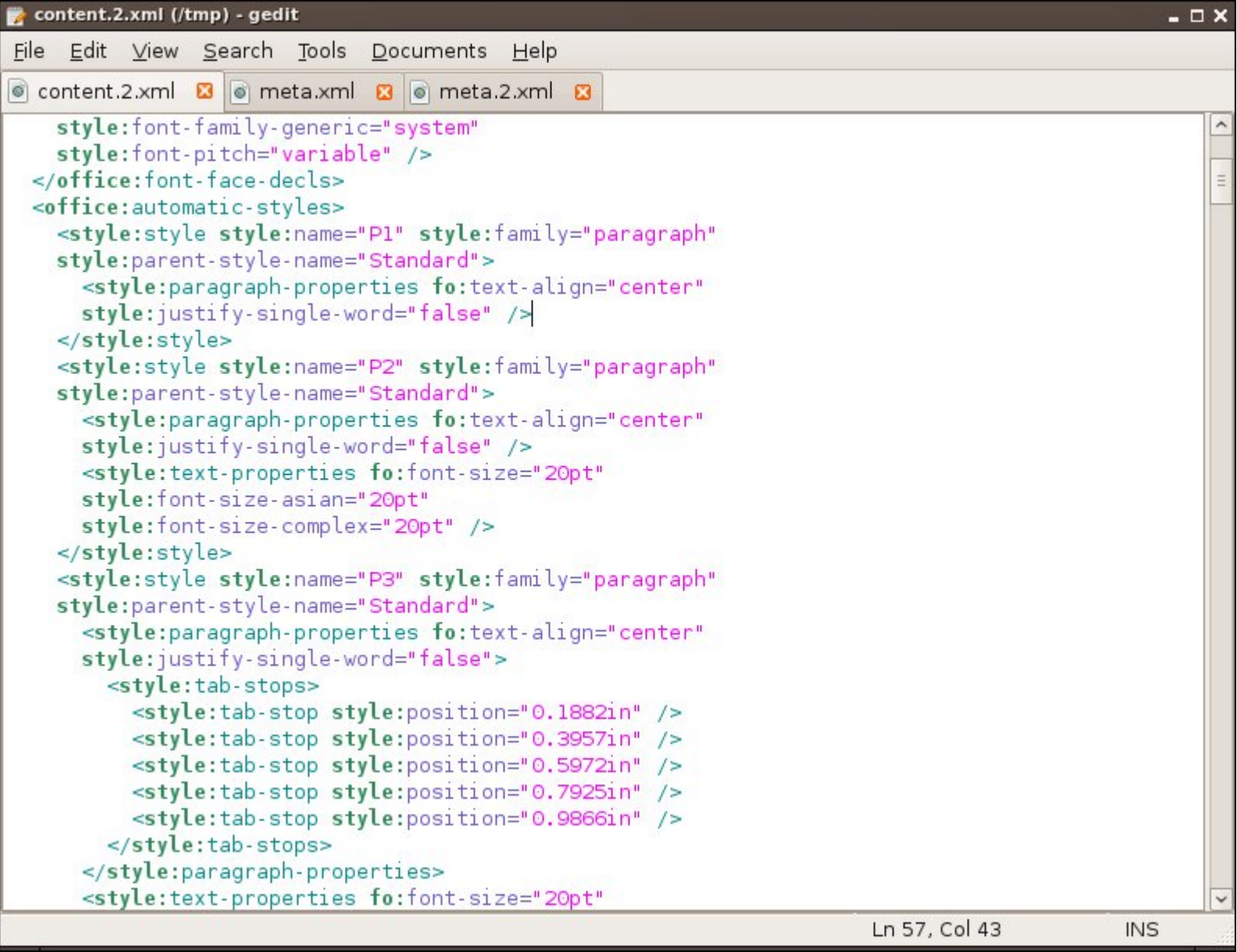
content.xml : exemple



```
content.2.xml (/tmp) - gedit
File Edit View Search Tools Documents Help
content.2.xml meta.xml meta.2.xml
xmlns:field="urn:openoffice:names:experimental:ooo-ms-interop:xmlns:field:1.0"
xmlns:formx="urn:openoffice:names:experimental:ooxml-odf-interop:xmlns:form:1.0"
office:version="1.2">
  <office:scripts />
  <office:font-face-decls>
    <style:font-face style:name="OpenSymbol"
    svg:font-family="OpenSymbol" />
    <style:font-face style:name="StarSymbol"
    svg:font-family="StarSymbol" />
    <style:font-face style:name="Liberation Mono"
    svg:font-family="'Liberation Mono'"
    style:font-family-generic="modern" style:font-pitch="fixed" />
    <style:font-face style:name="Liberation Serif"
    svg:font-family="'Liberation Serif'"
    style:font-family-generic="roman"
    style:font-pitch="variable" />
    <style:font-face style:name="Liberation Sans"
    svg:font-family="'Liberation Sans'"
    style:font-family-generic="swiss"
    style:font-pitch="variable" />
    <style:font-face style:name="DejaVu LGC Sans"
    svg:font-family="'DejaVu LGC Sans'"
    style:font-family-generic="system"
    style:font-pitch="variable" />
  </office:font-face-decls>
  <office:automatic-styles>
    <style:style style:name="P1" style:family="paragraph"
    style:parent-style-name="Standard">
      <style:paragraph-properties fo:text-align="center"
      style:justify-single-word="false" />
    </style:style>
  </office:automatic-styles>
</office:content>
```

Ln 50, Col 39 INS

content.xml : exemple

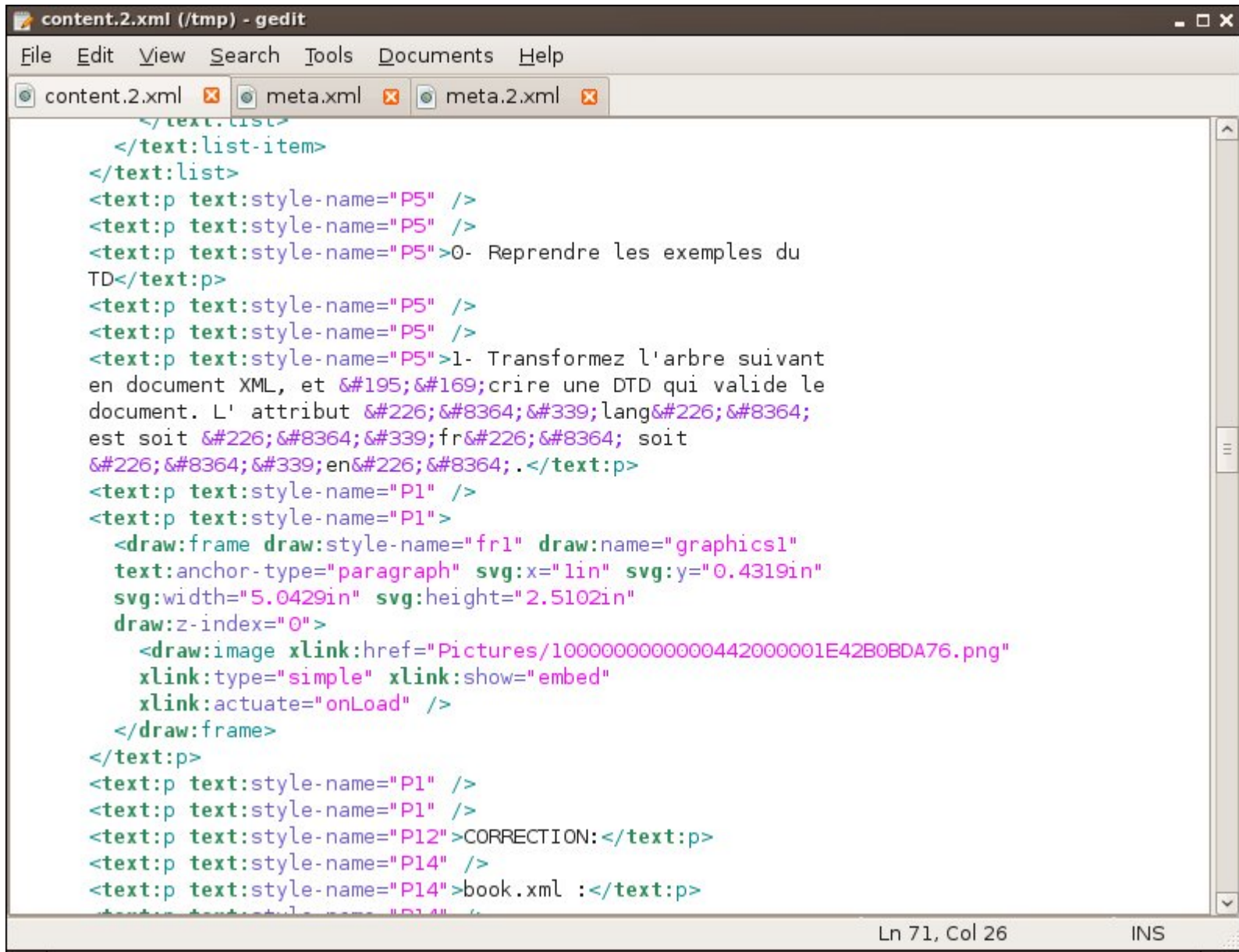


```
content.2.xml (/tmp) - gedit
File Edit View Search Tools Documents Help
content.2.xml meta.xml meta.2.xml

<style:font-family-generic="system"
  style:font-pitch="variable" />
</office:font-face-decls>
<office:automatic-styles>
  <style:style style:name="P1" style:family="paragraph"
    style:parent-style-name="Standard">
    <style:paragraph-properties fo:text-align="center"
      style:justify-single-word="false" />
  </style:style>
  <style:style style:name="P2" style:family="paragraph"
    style:parent-style-name="Standard">
    <style:paragraph-properties fo:text-align="center"
      style:justify-single-word="false" />
    <style:text-properties fo:font-size="20pt"
      style:font-size-asian="20pt"
      style:font-size-complex="20pt" />
  </style:style>
  <style:style style:name="P3" style:family="paragraph"
    style:parent-style-name="Standard">
    <style:paragraph-properties fo:text-align="center"
      style:justify-single-word="false">
      <style:tab-stops>
        <style:tab-stop style:position="0.1882in" />
        <style:tab-stop style:position="0.3957in" />
        <style:tab-stop style:position="0.5972in" />
        <style:tab-stop style:position="0.7925in" />
        <style:tab-stop style:position="0.9866in" />
      </style:tab-stops>
    </style:paragraph-properties>
    <style:text-properties fo:font-size="20pt"
```

Ln 57, Col 43 INS

content.xml : exemple

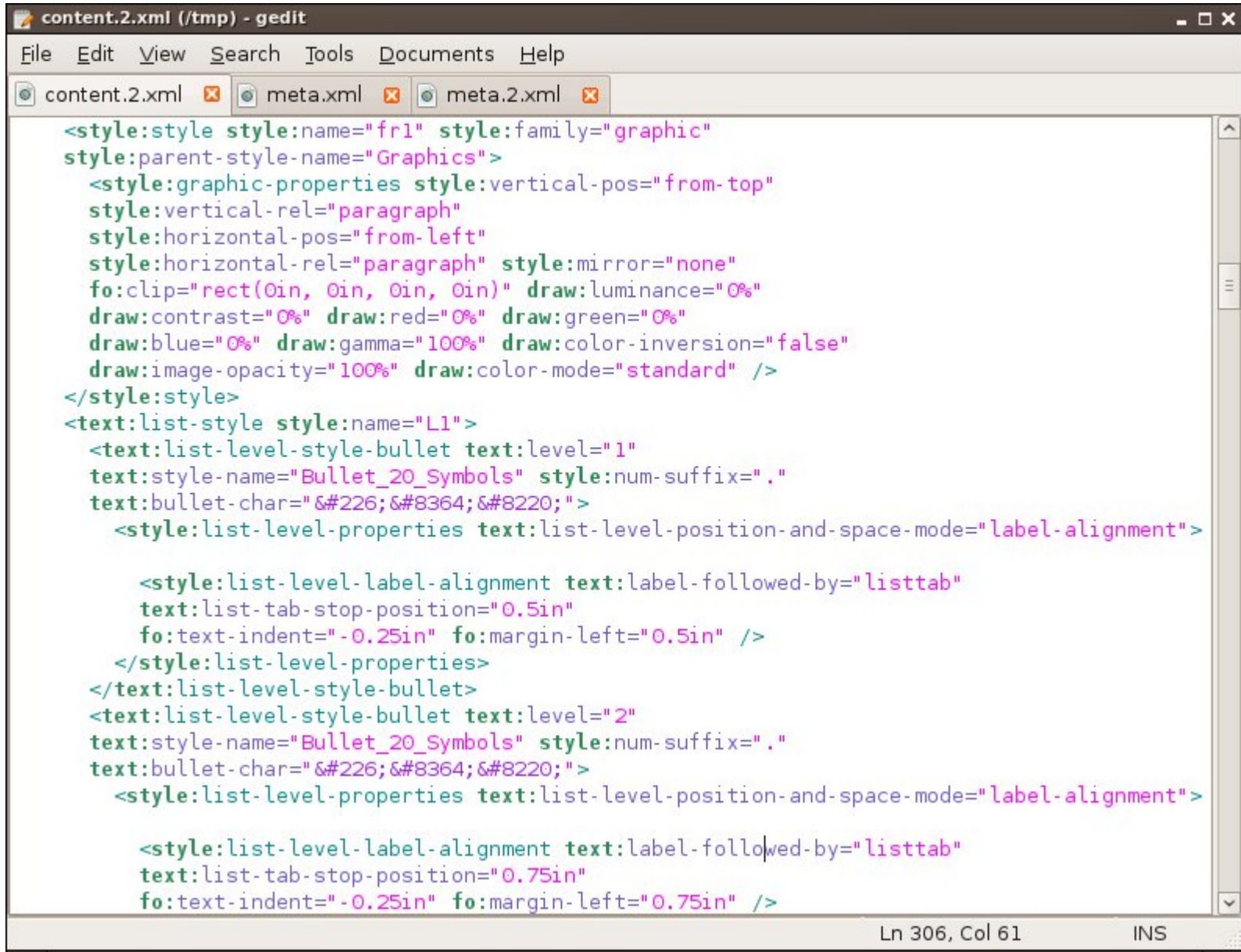


```
content.2.xml (/tmp) - gedit
File Edit View Search Tools Documents Help
content.2.xml meta.xml meta.2.xml

</text:list-item>
</text:list>
<text:p text:style-name="P5" />
<text:p text:style-name="P5" />
<text:p text:style-name="P5">0- Reprendre les exemples du
TD</text:p>
<text:p text:style-name="P5" />
<text:p text:style-name="P5" />
<text:p text:style-name="P5">1- Transformez l'arbre suivant
en document XML, et &#195;&#169;crire une DTD qui valide le
document. L' attribut &#226;&#8364;&#339;lang&#226;&#8364;
est soit &#226;&#8364;&#339;fr&#226;&#8364; soit
&#226;&#8364;&#339;en&#226;&#8364;.</text:p>
<text:p text:style-name="P1" />
<text:p text:style-name="P1">
  <draw:frame draw:style-name="fr1" draw:name="graphics1"
    text:anchor-type="paragraph" svg:x="1in" svg:y="0.4319in"
    svg:width="5.0429in" svg:height="2.5102in"
    draw:z-index="0">
    <draw:image xlink:href="Pictures/10000000000000442000001E42B0BDA76.png"
      xlink:type="simple" xlink:show="embed"
      xlink:actuate="onLoad" />
    </draw:frame>
  </text:p>
<text:p text:style-name="P1" />
<text:p text:style-name="P1" />
<text:p text:style-name="P12">CORRECTION:</text:p>
<text:p text:style-name="P14" />
<text:p text:style-name="P14">book.xml :</text:p>
<text:p text:style-name="P14" />

Ln 71, Col 26 INS
```

content.xml : exemple

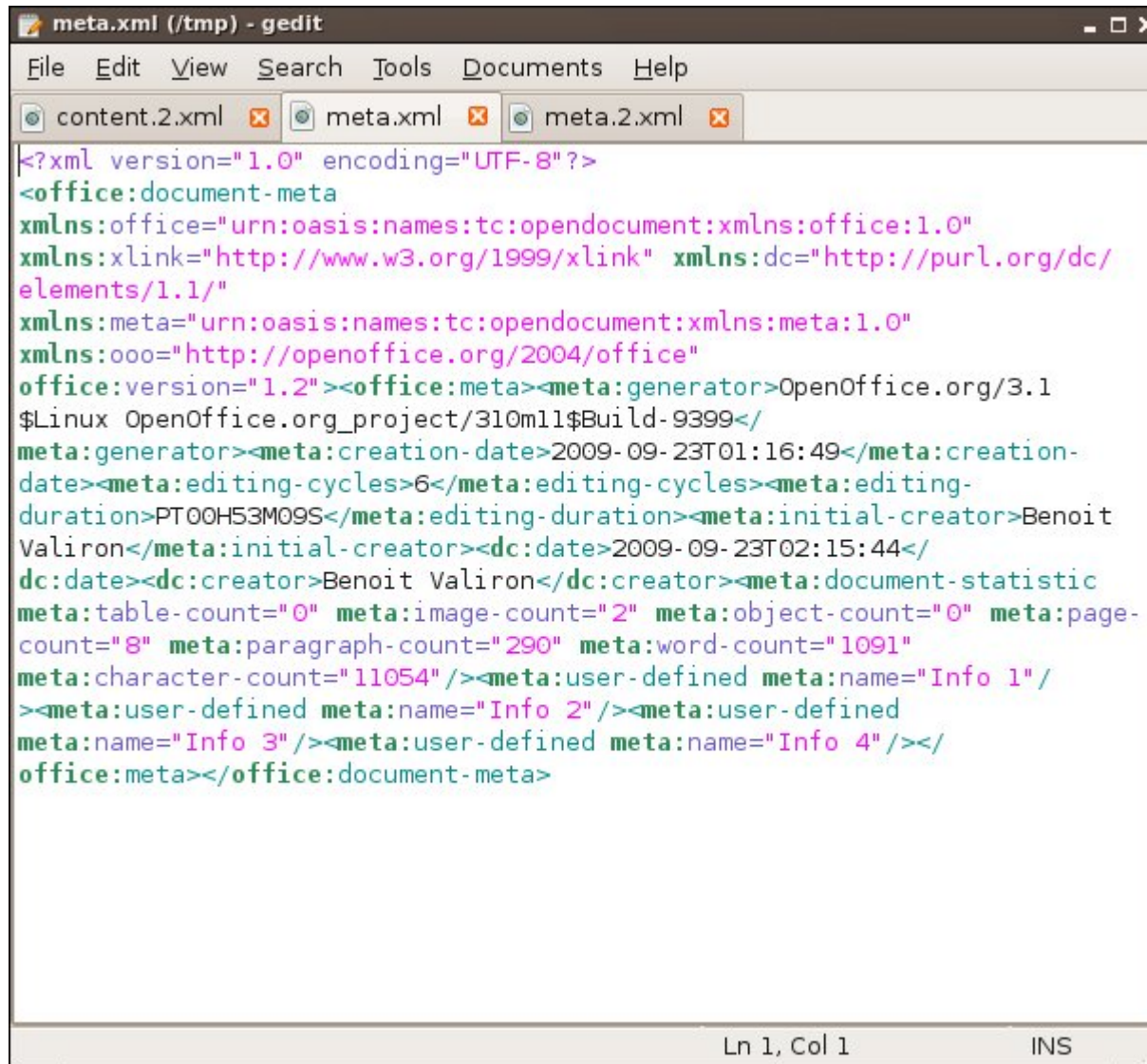
A screenshot of a gedit editor window titled 'content.2.xml (/tmp) - gedit'. The window has a menu bar with 'File', 'Edit', 'View', 'Search', 'Tools', 'Documents', and 'Help'. Below the menu bar is a tab bar showing three open files: 'content.2.xml', 'meta.xml', and 'meta.2.xml'. The main editing area displays XML code for a content.xml file. The code defines a style for a graphic, a list-style, and two list-level styles. The list-level styles specify bullet characters and indentation. The status bar at the bottom indicates 'Ln 306, Col 61' and 'INS' mode.

```
<style:style style:name="fr1" style:family="graphic"
style:parent-style-name="Graphics">
  <style:graphic-properties style:vertical-pos="from-top"
style:vertical-rel="paragraph"
style:horizontal-pos="from-left"
style:horizontal-rel="paragraph" style:mirror="none"
fo:clip="rect(0in, 0in, 0in, 0in)" draw:luminance="0%"
draw:contrast="0%" draw:red="0%" draw:green="0%"
draw:blue="0%" draw:gamma="100%" draw:color-inversion="false"
draw:image-opacity="100%" draw:color-mode="standard" />
</style:style>
<text:list-style style:name="L1">
  <text:list-level-style-bullet text:level="1"
text:style-name="Bullet_20_Symbols" style:num-suffix="."
text:bullet-char="¶;¶;¶">
    <style:list-level-properties text:list-level-position-and-space-mode="label-alignment">

      <style:list-level-label-alignment text:label-followed-by="listtab"
text:list-tab-stop-position="0.5in"
fo:text-indent="-0.25in" fo:margin-left="0.5in" />
    </style:list-level-properties>
  </text:list-level-style-bullet>
  <text:list-level-style-bullet text:level="2"
text:style-name="Bullet_20_Symbols" style:num-suffix="."
text:bullet-char="¶;¶;¶">
    <style:list-level-properties text:list-level-position-and-space-mode="label-alignment">

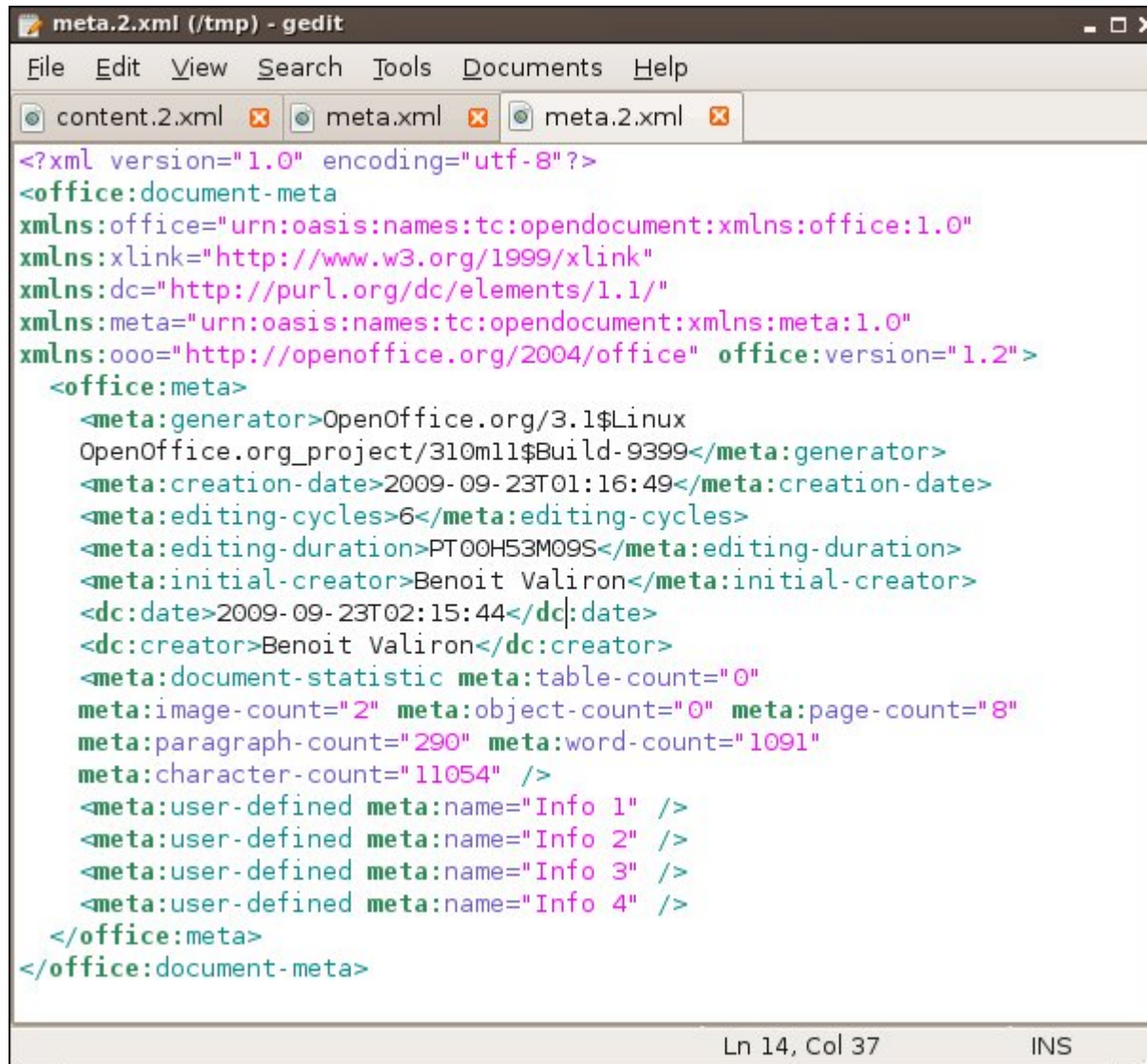
      <style:list-level-label-alignment text:label-followed-by="listtab"
text:list-tab-stop-position="0.75in"
fo:text-indent="-0.25in" fo:margin-left="0.75in" />
    </style:list-level-properties>
  </text:list-level-style-bullet>
</text:list-style>
```


meta.xml



```
<?xml version="1.0" encoding="UTF-8"?>
<office:document-meta
xmlns:office="urn:oasis:names:tc:opendocument:xmlns:office:1.0"
xmlns:xlink="http://www.w3.org/1999/xlink" xmlns:dc="http://purl.org/dc/
elements/1.1/"
xmlns:meta="urn:oasis:names:tc:opendocument:xmlns:meta:1.0"
xmlns:ooo="http://openoffice.org/2004/office"
office:version="1.2"><office:meta><meta:generator>OpenOffice.org/3.1
$Linux OpenOffice.org_project/310m11$Build-9399</
meta:generator><meta:creation-date>2009-09-23T01:16:49</meta:creation-
date><meta:editing-cycles>6</meta:editing-cycles><meta:editing-
duration>PT00H53M09S</meta:editing-duration><meta:initial-creator>Benoit
Valiron</meta:initial-creator><dc:date>2009-09-23T02:15:44</
dc:date><dc:creator>Benoit Valiron</dc:creator><meta:document-statistic
meta:table-count="0" meta:image-count="2" meta:object-count="0" meta:page-
count="8" meta:paragraph-count="290" meta:word-count="1091"
meta:character-count="11054"/><meta:user-defined meta:name="Info 1"/
><meta:user-defined meta:name="Info 2"/><meta:user-defined
meta:name="Info 3"/><meta:user-defined meta:name="Info 4"/></
office:meta></office:document-meta>
```

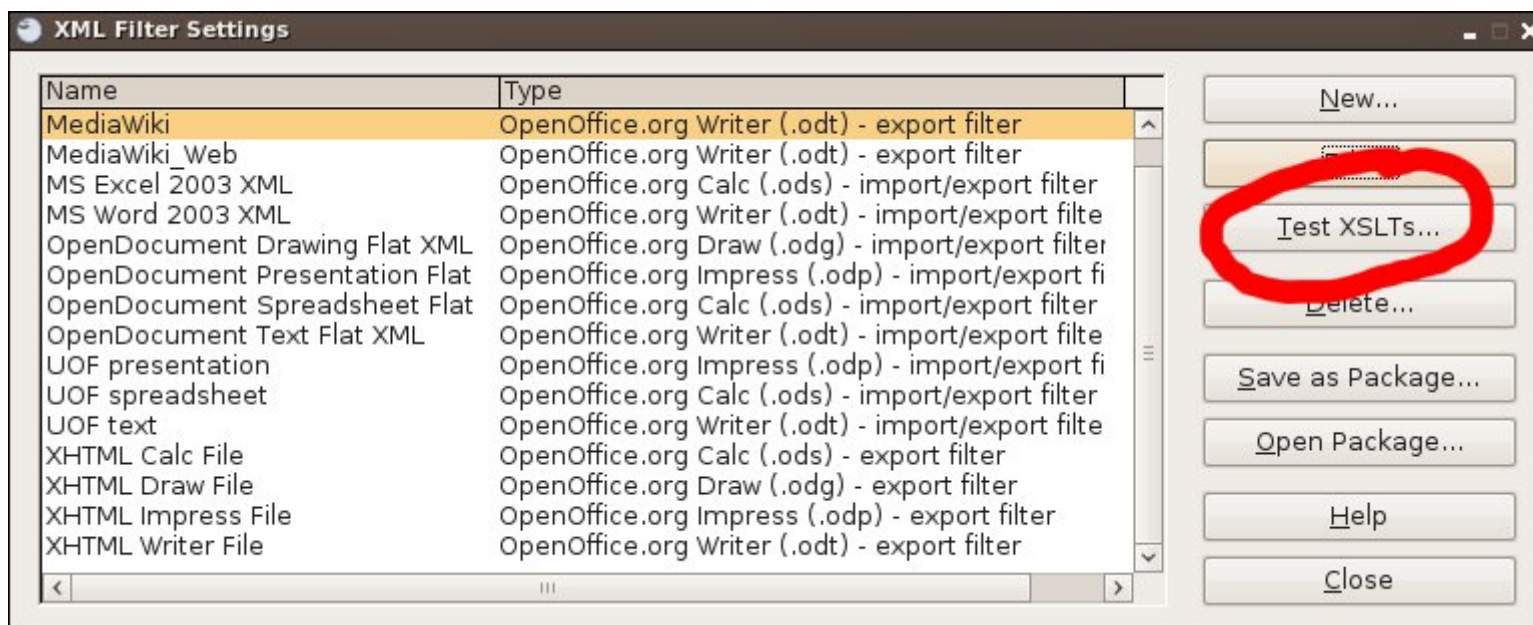
meta.xml



```
<?xml version="1.0" encoding="utf-8"?>
<office:document-meta
xmlns:office="urn:oasis:names:tc:opendocument:xmlns:office:1.0"
xmlns:xlink="http://www.w3.org/1999/xlink"
xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:meta="urn:oasis:names:tc:opendocument:xmlns:meta:1.0"
xmlns:ooo="http://openoffice.org/2004/office" office:version="1.2">
  <office:meta>
    <meta:generator>OpenOffice.org/3.1$Linux
OpenOffice.org_project/310m11$Build-9399</meta:generator>
    <meta:creation-date>2009-09-23T01:16:49</meta:creation-date>
    <meta:editing-cycles>6</meta:editing-cycles>
    <meta:editing-duration>PT00H53M09S</meta:editing-duration>
    <meta:initial-creator>Benoit Valiron</meta:initial-creator>
    <dc:date>2009-09-23T02:15:44</dc:date>
    <dc:creator>Benoit Valiron</dc:creator>
    <meta:document-statistic meta:table-count="0"
meta:image-count="2" meta:object-count="0" meta:page-count="8"
meta:paragraph-count="290" meta:word-count="1091"
meta:character-count="11054" />
    <meta:user-defined meta:name="Info 1" />
    <meta:user-defined meta:name="Info 2" />
    <meta:user-defined meta:name="Info 3" />
    <meta:user-defined meta:name="Info 4" />
  </office:meta>
</office:document-meta>
```

Ln 14, Col 37 INS

XML implique outils standards



Commentaires

- Le format ODF d'OpenOffice a beau être XML, il est néanmoins plutôt orienté machine (même si il reste lisible, suivant la philosophie XML qui demande des noms de balises explicites).
- À la différence de Docbook, ici on est plus concerné par la mise en page que par la structure logique du texte.
- Le formatage est strictement séparé dans des styles.
- Ceci dit, si on le souhaite on peut récupérer le texte et une partie du formatage.
- D'ailleurs, une série de macros XSLT permettent de le faire.

OOXML : Microsoft Office

- Vrai nom : ECMA-376 : Office Open XML file format.
- Maintenant, après de nombreuses péripéties, c'est un standard ISO.
- Descendant du format binaire : .xls, .doc, .ppt
- Néanmoins, le corps du document est en XML.

Un fichier .docx

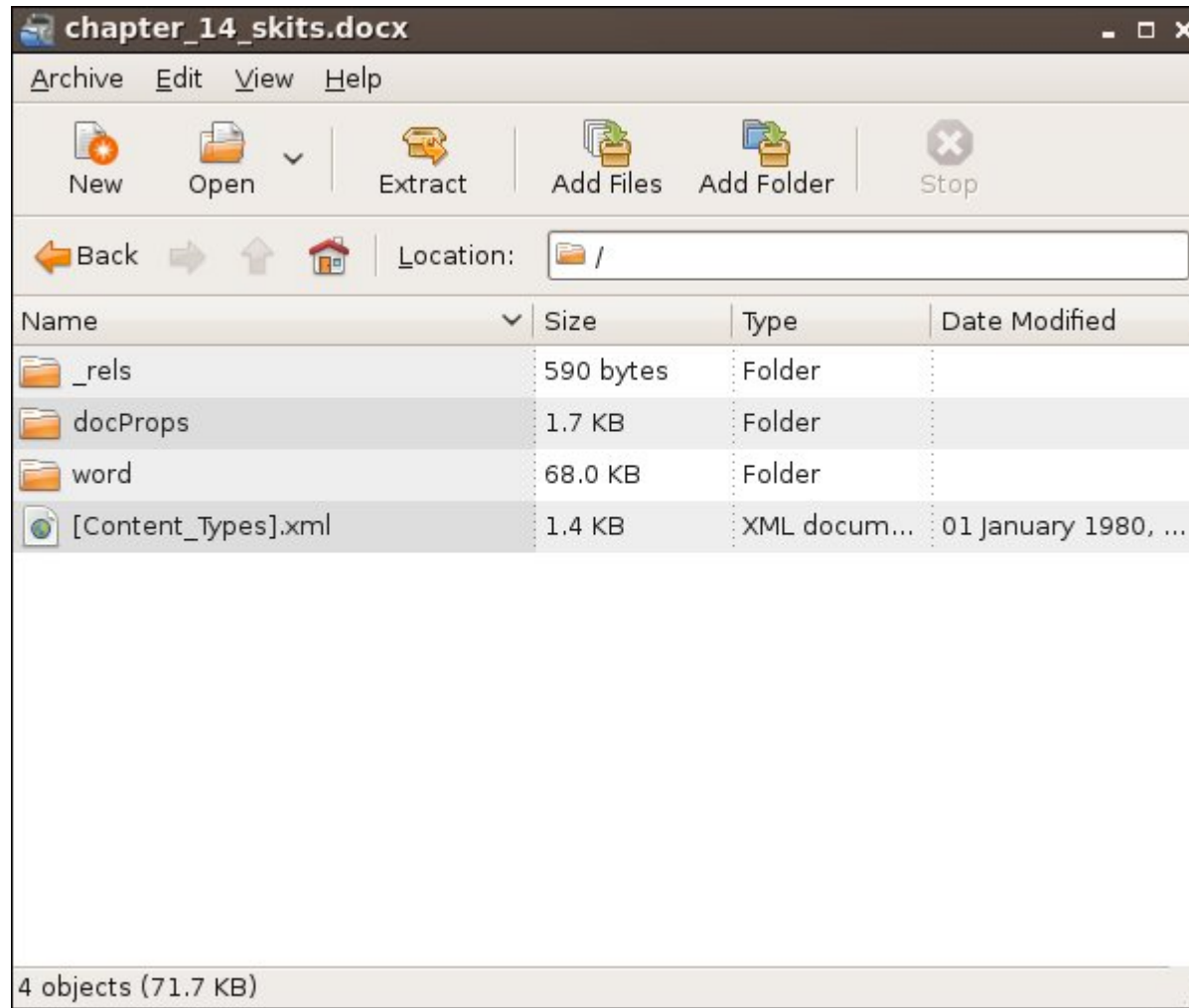
- Décompressé :
 - [Content_Types].xml → décrit le contenu de l'archive
 - _rels → dossier contenant des fichiers XML décrivant les relations entre les composantes de l'archive.
 - DocProps → dossier avec les propriétés
 - De l'application : app.xml
 - Du document : core.xml
 - Définies par l'utilisateur : custom.xml
 - Un répertoire pour le type du document (ex : word) contenant document.xml

Exemple de document.xml

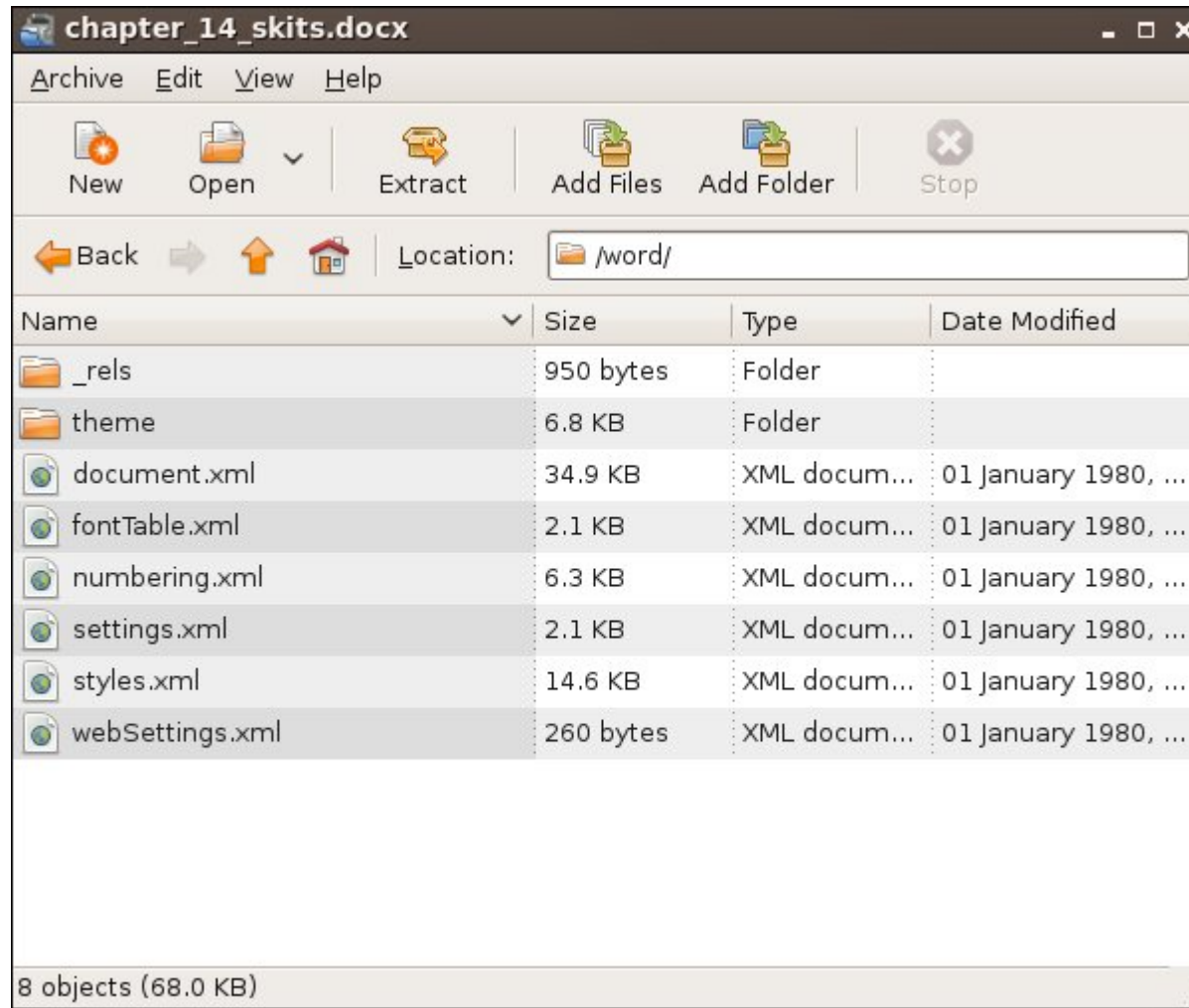
Un paragraphe contenant les mots “gras” et “italique” en gras et en italique :

```
<w:p>
  <w:r>
    <w:rPr><w:b/></w:rPr>
    <w:t>gras</w:t>
  </w:r>
  <w:r>
    <w:rPr><w:i/></w:rPr>
    <w:t>italique</w:t>
  </w:r>
</w:p>
```

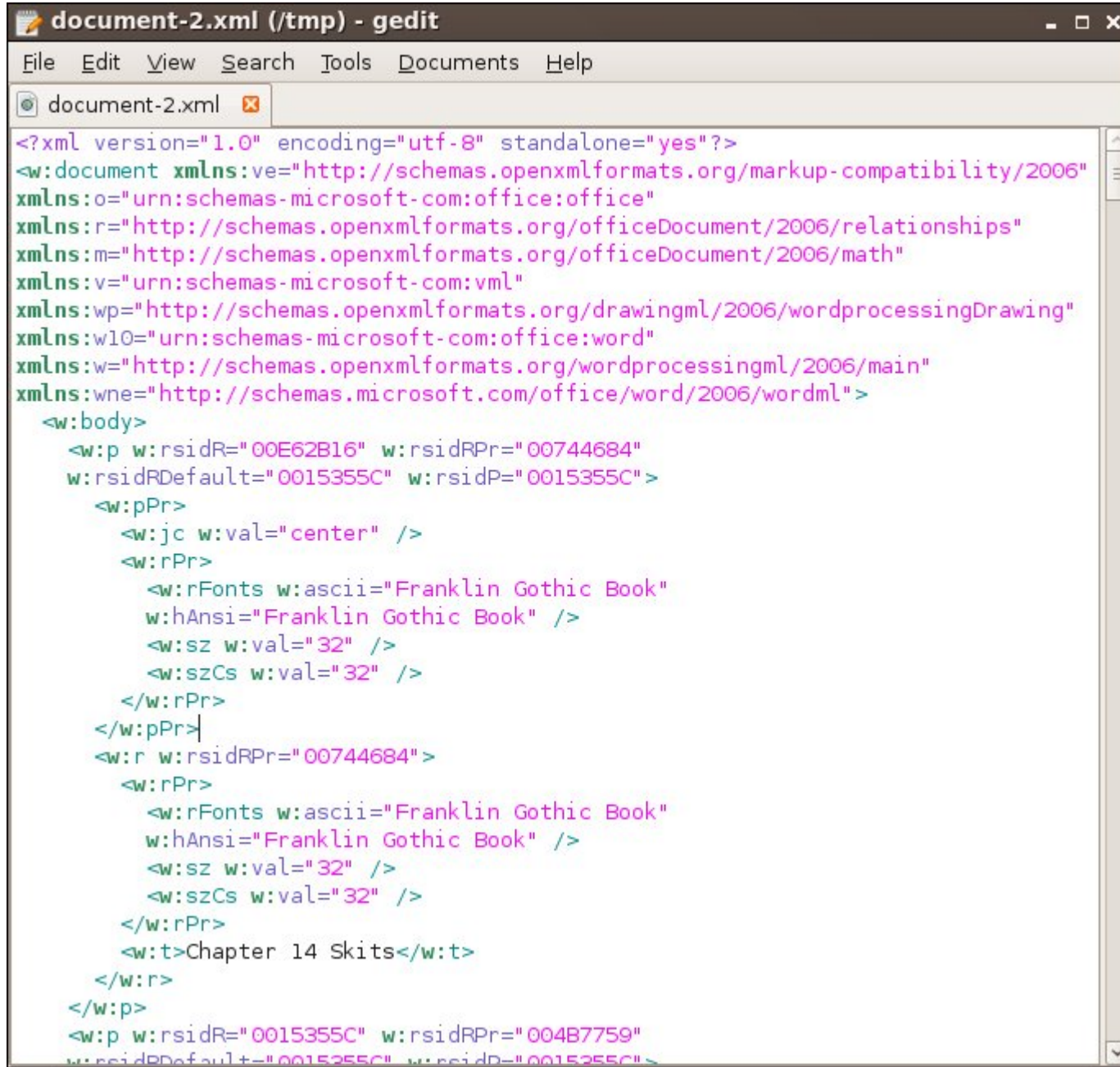
Un exemple tiré du ouaibe (je n'ai pas Office 2007...)



Un exemple tiré du ouaibe (je n'ai pas Office 2007...)

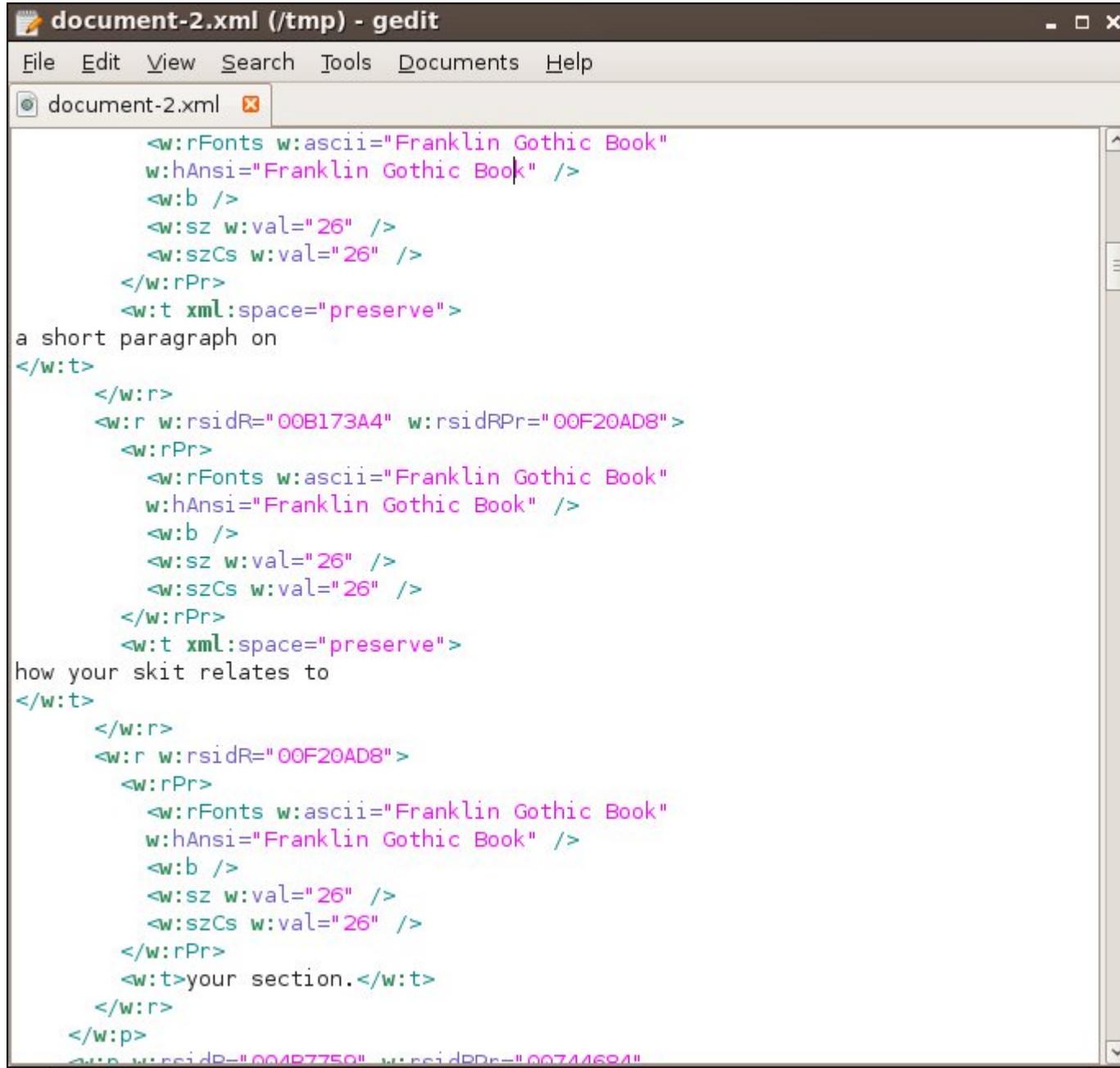


document.xml



```
<?xml version="1.0" encoding="utf-8" standalone="yes"?>
<w:document xmlns:ve="http://schemas.openxmlformats.org/markup-compatibility/2006"
xmlns:o="urn:schemas-microsoft-com:office:office"
xmlns:r="http://schemas.openxmlformats.org/officeDocument/2006/relationships"
xmlns:m="http://schemas.openxmlformats.org/officeDocument/2006/math"
xmlns:v="urn:schemas-microsoft-com:vml"
xmlns:wp="http://schemas.openxmlformats.org/drawingml/2006/wordprocessingDrawing"
xmlns:w10="urn:schemas-microsoft-com:office:word"
xmlns:w="http://schemas.openxmlformats.org/wordprocessingml/2006/main"
xmlns:wne="http://schemas.microsoft.com/office/word/2006/wordml">
  <w:body>
    <w:p w:rsidR="00E62B16" w:rsidRPr="00744684"
w:rsidRDefault="0015355C" w:rsidP="0015355C">
      <w:pPr>
        <w:jc w:val="center" />
        <w:rPr>
          <w:rFonts w:ascii="Franklin Gothic Book"
w:hAnsi="Franklin Gothic Book" />
          <w:sz w:val="32" />
          <w:szCs w:val="32" />
        </w:rPr>
      </w:pPr>
      <w:r w:rsidRPr="00744684">
        <w:rPr>
          <w:rFonts w:ascii="Franklin Gothic Book"
w:hAnsi="Franklin Gothic Book" />
          <w:sz w:val="32" />
          <w:szCs w:val="32" />
        </w:rPr>
        <w:t>Chapter 14 Skits</w:t>
      </w:r>
    </w:p>
    <w:p w:rsidR="0015355C" w:rsidRPr="004B7759"
w:rsidRDefault="0015355C" w:rsidP="0015355C">
```

document.xml



```
document-2.xml (/tmp) - gedit
File Edit View Search Tools Documents Help
document-2.xml
<w:rFonts w:ascii="Franklin Gothic Book"
w:hAnsi="Franklin Gothic Book" />
<w:b />
<w:sz w:val="26" />
<w:szCs w:val="26" />
</w:rPr>
<w:t xml:space="preserve">
a short paragraph on
</w:t>
</w:r>
<w:r w:rsidR="00B173A4" w:rsidRPr="00F20AD8">
<w:rPr>
<w:rFonts w:ascii="Franklin Gothic Book"
w:hAnsi="Franklin Gothic Book" />
<w:b />
<w:sz w:val="26" />
<w:szCs w:val="26" />
</w:rPr>
<w:t xml:space="preserve">
how your skit relates to
</w:t>
</w:r>
<w:r w:rsidR="00F20AD8">
<w:rPr>
<w:rFonts w:ascii="Franklin Gothic Book"
w:hAnsi="Franklin Gothic Book" />
<w:b />
<w:sz w:val="26" />
<w:szCs w:val="26" />
</w:rPr>
<w:t>your section.</w:t>
</w:r>
</w:p>
<w:r w:rsidR="00A97750" w:rsidRPr="00744694">
```

Commentaires

- Comme pour ODF, on a un format orienté machine.
- À la différence d'ODF, les styles sont inclus dans le contenu.
- Il est particulièrement peu lisible, avec des balises et des espaces de noms peu explicites, ce qui ne l'empêche pas d'être verbeux.
- Il porte le lourd héritage du format binaire

Conclusion

- Il y a mille façons d'utiliser le format XML pour encoder du texte :
 - Format humain
 - Format machine
 - Orienté ouaibe ou orienté impression papier