Librevenge is suite

What is new in the world of import filters and what is coming soon

Fridrich Štrba

LibreOffice Contributor

LibreOffice

Agenda

- New developments in the importer world

- librevenge and libodfgen

- What is next?

LibreOffice

New developments in the importer world

New babies in the family

- Libetonyek
    - Support first for *Keynote* documents
    - Extending support to *Numbers* and *Pages*
- Libe-book
    - Supports a host of e-book file-formats
- Libfreehand
    - Started the implementing of Freehand import filter
- Libabw
    - Now we can load documents of our "cousin"
- … and more still to come

LibreOffice

New document types

- Previously only text documents and graphics
  - Text documents based on **libwpd API**
    - Libwpd, libwps, libmwaw
  - Graphics based on **libwpg API**
    - Libwpg, libvisio, libcdr, libmspub, libfreehand

- New presentation support
  - Presentations based on **libetonyek API**
    - Libwpg's API was too limited for presentations

- Need to extend to spreadsheets too
  - Libmwaw
  - Libwps

libodfgen and **librevenge**

libodfgen

- ODF Generation was duplicated in several places
  - LibreOffice *writerperfect* module
  - Standalone *writerperfect*
  - Calligra sources
  - It makes sense to collect all bugs in the same place

- **OdtGenerator** class
  - Implementations of *WPXDocumentInterface*
- **OdgGenerator** class
  - Implementation of *WPGPaintInterface*
- **OdpGenerator** class added later
  - Implementation of *KEYPresentationInterface*
- **OdfDocumentHandler** interface
  - SAX-like interface to output XML in a generic way

Librevenge

- **Interface of each document type in different library**
    - Libwpd, libwpg, libetonyek
- **The common types in libwpd**
    - Libwpd is a text-related library
    - All others had to link to it
- **Consolidating the types and interfaces**
    - **Interfaces**
        - RVNGTextInterface, RVNGDrawingInterface,
        - RVNGPresentationInterface, RVNGSpreadsheetInterface
    - **Types**
        - RVNGProperty, RVNGPropertyList, RVNGPropertyListVector
            - Extended the capacities
        - RVNGBinaryData, RVNGString, RVNGStringVector

Librevenge-stream

- RVNGInputStream interface
  - Extended to handle structured documents a bit more efficiently
- Several implementations:
  - RVNGFileStream
    - Implementation using file name
  - RVNGStringStream
    - Implementation using a buffer of data
  - RVNGDirectoryStream
    - Accesses a directory structure as if it was a structured document
- OLE2 and ZIP documents handled transparently
  - No need to know what is the container type
  - Gives the responsibility to the implementers!

Librevenge-generators

- Useful implementations of the different interfaces
- Raw Generators
  - Implementations of the different RVNG interfaces
    - printing callbacks called and properties passed
  - Used for regression testing
- CSV generator for spreadsheets, HTML, Text generators
- SVG generators
  - Exception: SVG generator for drawings
    - included in librevenge core library
    - Historical reasons
- ODF generators in libodfgen
  - More complicated
  - Historical reasons

Advantage of the design

- Parser libraries independent and self-contained
  - Much easier life of filter writers
    - Enough to focus on the structure of document to parse
    - Call the interface callbacks that one needs
  - Avoid sucking in unrelated libraries
    - Librevenge itself and libodfgen have boost as build-time dependency
    - No need to link text-related libraries in drawing application

- Considerable reduction of code duplication
  - Less risk to have bugs fixed in one place and hanging around in another
  - Faster to start a library skeleton

What is next?

Future file-formats to import?

- Google Summer of Code
  - The possibility for a student to work with outstanding mentors
    - Valentin Filippov
    - David Tardon
    - Your faithful

- Several formats ready for straight engineering
  - Apple Numbers, Pages
  - Adobe PageMaker
  - Adobe InDesign
  - … and any other you are interested in

LibreOffice

Impact within LibreOffice and the known universe

◤ **Happy users will reward you**

  ◤ You will be the hero of the people who can now read their documents...

  ◤ … and they will get on your nerves listing features that are not converted.

◤ **Users outside LibreOffice**

  ◤ Inkscape reuses `libvisio` and `libcdr` in 0.49

  ◤ Calligra reuses `libvisio` and (possibly) `libcdr` since 2.5

◤ **Great fun within the community**

QA and Stoning session