

Otvoreni formati: OpenDocument i PDF

Vedran Miletić

Document Freedom Day 2011

Zašto otvoreni formati?

- Kryder's Law: cijena pohrane jedinice informacije raste eksponencijalno
- Digitalne arhive dokumenata zamjenjuju klasične arhive (e-izdvadak, e-zdravstvo)
- Zastarijevanje softvera i problem čitanja binarnih formata

OpenDocument is...

- **Quiz:** Complete the sentence, "OpenDocument is ..."
 - (a) An open, XML-based file format.
 - (b) An open standard, supported by the OASIS and ISO standards groups.
 - (c) The default file format in OpenOffice.org since 2.0.
 - (d) Official format for the European Commission.
 - (e) Our best chance to fight vendor lock-in associated with proprietary formats.
 - (f) All of the above.
- The correct answer is (f) All of the above.

OpenDocument

- XML-based
- Promjene imena
 - OpenOffice.org 1.0 format (ali trebao bi biti application-neutral...)
 - Open Office XML (!= MS Office XML)
 - **OpenDocument format**
- OASIS standard 1. svibnja 2005.
- ISO/IEC standard 3. svibnja 2006.

OpenDocument ekstenzije

- Nastale prema mogućnostima tadašnjih OpenOffice.org alata:
 - * .odt for word processing (text) documents
 - * .ods for spreadsheets
 - * .odp for presentations
 - * .odb for databases
 - * .odg for graphics
 - * .odf for formulae, mathematical equations

OpenDocument softver

- LibreOffice (i OpenOffice.org, Oracle Open Office i svi ostali derivati)
- Drugi office alati: KOffice, AbiWord
- Alati za crtanje: Inkscape, Scribus
- Komercijalne aplikacije: IBM Lotus Symphony, Corel WordPerfect Office X4
- Web aplikacije: Google Docs, Adobe Buzzword, Zoho Office Suite
- Preglednici: Evince, Okular

Prihvaćanje OpenDocumenta

- Dugotrajan proces
- 2005. Massachusetts
- Od 2006. nadalje članice EU
- NATO pakt i njegove članice imaju obavezu korištenja OpenDocumenta
- Ne znači nužno prelazak na LibreOffice

Primjer (Venezuela)

- *"All organizations of the Federal Government of Venezuela must:*
 - *Use ODF 1.0 (ISO/IEC 26300) for editable documents*
 - *Use PDF for non-editable documents"*

PDF

- Nasljednik PostScripta
- Vremenom "istisnuo" i PostScript i DjVu iz široke upotrebe
- 1993. Adobe, prvotno nazvan "Camelot"
- 2008. otvoren standard, ISO 32000-1:2008
- Može uključivati tekst, fontove, rastersku i vektorsku grafiku i razne dodatne multimedijске sadržaje (video, SWF, U3D/PRC, ...)

PDF revizije

- 1.2 – interaktivnost (forme, gumbi), Unicode
- 1.3 – ICC boje, JS, attachmenti, poboljšano numeriranje stranica, CID fontovi
- 1.4 – XMP, Tagged PDF, umetanje PDF-a u PDF (omogućuje PDF slike), OCR text layer
- 1.5 – sigurnost, opcionalni sadržaj (slojevi)
- 1.6 – U3D/PRC, OpenType
- **pdfinfo** iz poppler-utilisa

PDF fontovi

- Embedded vs Embedded Subset
- PDF podržava Type 1, TrueType i OpenType:
 - Type 1 => **Type 1** ili Type 1C u PDF-u
 - TrueType => **TrueType** ili CID TrueType u PDF-u
 - OpenType => CID Type 0 ili **CID Type 0C** u PDF-u (nekonzistentna terminologija; Type 1C (CID) znači isto što i CID Type 0C)
- CID-keyed font omogućuje kodiranje većeg broja znakova (16-bit umjesto 8-bit)
- **pdffonts** iz poppler-utilisa

PDF fontovi i kodiranje

- 1990. T1 (Cork, EC) TeX kodiranje
- PDF CMap
- OpenType: CFF (Type 1) vs glyf (TrueType)

PDF i grafika

- Odlična podrška za vektorsku grafiku
 - PDF slike
 - PS->PDF->PS (GhostScript)
 - SVG->PDF->SVG (Inkscape/Cairo)
 - MetaPost->PDF (luamplib)
- Podrška za 3D grafiku
 - PRC
 - U3D
 - samo Adobe Reader pod Windowsima :-(

PDF aplikacije

- Gotovo sve spomenute aplikacije + Apache FOP (XML->PDF)
- Kvaliteta proizvedenog PDF-a jako varira
 - Ovisi o uključivanju fontova
 - Ovisi o kodiranju fontova
- KOffice i LibreOffice mogu *uređivati* PDF

Hvala na pažnji!