Open standards keep documents compatible across different applications and versions.

Keep it Open

The Open Document Format standard is an important reason for the success of LibreOffice. *By Italo Vignoli*

The Open Document Format, or ODF, is an open format for word processing files, spreadsheets, charts, and presentations developed with the aim of providing a standard document format for office applications. ODF was released by a Technical Committee at the Organization for the Advancement of Structured Information Standards (OASIS) consortium in 2005 and approved as an ISO/IEC International Standard in 2006 as ISO/IEC 26300:2006. A document format must be able to accommodate the features of modern applications, such as headers and footers, footnotes, tables of contents, and revision tracking, and also consider issues such as security, document encryption, digital signatures, and problems related to executable code embedded within documents via scripts or macros. In addition, the document format must be resilient, as each file is going to be read and edited multiple times by different applications, including applications that were not involved in creating it. At the end of the process, the file must preserve its integrity and characteristics, and – more important than anything else – guarantee that contents are preserved independently from the application.

THE IMPORTANCE OF DOCUMENT STANDARDS

File compatibility was once based on the assumption that everyone was using the same operating system and the same office suite configured in the same way, with the same fonts. In that case, electronic document exchanges were straightforward. Today, the reality is completely different. Several different operating systems and office suites compete for the user's attention, and documents are shared more than ever before. The best way to

achieve interoperability is to adopt standard formats, which not only allow proper interoperability, but also reduce the impact of "vendor lock-in" and its deleterious consequences on innovation and competition.

A SHORT HISTORY OF ODF

The ODF standard was created as an evolution of OOo XML format contributed by Sun. ODF is maintained by the ODF Technical Committee (TC) within the Organization for the Advancement of Structured Information Standards (OASIS). The first meeting to discuss the standard was held on December 16, 2002. ODF 1.0 specifications were published in May 1, 2005 and then submitted to ISO/IEC Joint Technical Committee 1 (JTC1) on November 16, 2005 under Publicly Available Specification (PAS). After a six-month review period, on May 3, 2006, OpenDocument unanimously passed its six-month DIS (Draft International Standard) ballot in JTC1 (ISO/IEC JTC1/SC34), with broad participation, and was "approved for release as an ISO and IEC International Standard" under the name ISO/IEC 26300:2006. ODF 1.2, approved as ISO/IEC 26300:2015 on June 17, 2015, adds to the specifications additional accessibility features, RDF-based metadata, a spreadsheet formula description based on OpenFormula, support for digital signatures, and some features suggested by the public. While we are writing this article, the TC is drafting ODF 1.3 specifications, which include new interesting features.

INSIDE ODF

An ODF file is typically stored and distributed in a ZIP container, which includes several XML files and the associated binary content, such as images or other media. Each ODF file contains four key embedded XML files: `manifest.xml`, the table of contents for the ZIP container; `meta.xml`, with document-level metadata; `styles.xml`, with style definitions for the document; and `content.xml`, with all the document's structured content. ODF provides a flexible container that can embed almost any existing digital object and allow it to be used in an office document. This is similar to how the HTML standard allows to reference just about anything.

STANDARDS AND PSEUDO-STANDARDS

When software is forced to adapt its format to a dominant product that is not based on open standards, the dominant vendor can use this power position to influence the market. Often the dominant product becomes a de facto, pseudo-standard that has never been vetted or approved but is merely adopted by the market out of necessity. Any flaws in the dominant product are then echoed out to the whole industry. For instance, the 40-year-old "leap year bug," which was inherited from Visicalc and Lotus 123 and passed to Excel, causes the market-leading spreadsheet to incorrectly represent the year 1900 as a leap year, which can cause errors for applications that count days or measure time ranges. Developers of other office suites had to replicate the bug for compatibility. According to the vendor, the faulty state has to be maintained, even across new releases of the product, because fixing it would cause too many problems. So, users are encouraged to keep the faulty spreadsheets, because the vendor protects its commercial interests in spite of the needs of users who have paid a software license.

ODF ADVANTAGES

Software vendors can implement the ODF standard to develop applications that are interoperable by design. Applications compete on the quality of document standards implementation, user interface, ease of use, performance, price, and a host of other features, while keeping customers' data intact and shareable, even if the customer chooses to switch to another competing product. ODF enables software vendors to compete on a level playing field. ODF provides several advantages over non-standard document formats, as it helps to lower ICT costs and improve ICT governance, increase flexibility, make it possible to manage cross-linked documents. In addition, using ODF allows much stricter security checks on incoming and outgoing documents to prevent common cyber-attack scenarios. This helps organizations be much less vulnerable to targeted attacks compared to older binary and pseudo-standard formats, which are among the top 3 most common vulnerabilities. Of course, ODF does not make insecure software safe, but makes it easier to check exactly what is going in and coming out of the document. Because of its advantages, ODF has been selected by several governments as the standard for document exchange within and with the public administration. United Kingdom, the Netherlands, France, Sweden and Taiwan are leading with a firmer stance, but other governments are suggesting ODF as well.

ODF FUTURE

People should be educated about the importance of open document standards. In a world where most documents are exchanged in digital format, an open standard should be a priority in schools and all educational institutions. Of course, this is not an easy task. During the last thirty years, the digital technology has evolved so quickly that it would have been absolutely impossible for any organization to stay up to date, especially with something completely new such as document standards. Governments should take care of the education process of young people but also of the new digital divide based on the lack of digital skills by senior citizens. Free Software organizations are doing as much as they can to fill this educational gap, but the task would be overwhelming even for a large commercial company. To contribute to the task, The Document Foundation is trying to relaunch the ODF Adoption Technical Committee, with the support of other Free Software Foundation and projects, companies, and public administrations. Continued investment in true, open standards will help to protect users and developers from the security, interoperability, and efficiency problems associated with vendor lock-in and closed formats.

This article is based on personal research, which has been partially inspired by the writings of Rob Weir, a former IBM employee who used to be one of the authorities on ODF and has chaired the OASIS ODF TC for quite a long time. In addition, some data and information are sourced from Wikipedia and from documents published by UK Government Cabinet Office. ■■■