What on Earth is...
OPENDOCUMENT?

Marco Fioretti introduces the new format for office files that's making waves in Brussels and Boston.

What is OpenDocument?

The full name for it is OASIS Open Document Format for Office Applications. It's a standard based on XML (eXtensible Markup Language) for rich text documents, spreadsheets and presentations..

What's OASIS? Who are they to be dishing out standards that we're supposed to adhere to?

OASIS is the not-for-profit Organization for the Advancement of Structured Information Standards (**www.oasis-open.org**). Among other things, they are also the authors of UBL, the Universal Business Language, which we covered in *LXF65*. OASIS members include companies as big as IBM, Boeing, Sun, Corel and, yes, Microsoft. From the end user's point of view, this guarantees that OpenDocument is not a one-man, undocumented hack that may disappear overnight.

You mean it's not new software – just another format?

Just another format? Wash your mouth out! Formats

the Cyrillic alphabet, or Arabic script. The words would still be your words, your ideas, in your language, but they would be represented differently on the page. And you'd have a pretty hard time using your work now that everyone else had moved on to the new alphabet standard.

Does OpenDocument have any competition? I'm not going to use a standard if there are loads of other 'standards' out there.

As a format, OpenDocument is not tied to any single software application. It isn't even limited, for that matter, only to FOSS programs, or only to proprietary ones. Having said that, OpenDocument is here to replace for good the .doc, .xls, and .ppt files used by *Microsoft Word*, *Excel* and *PowerPoint*.

Er... Why? I've never had any problem reading and writing Microsoft formats with *OpenOffice.org*.

If you have never had any problem at all exchanging files with Microsoft users in those formats, you are either very lucky, or, frankly, your files have been

The specifications for Microsoft formats are not completely accessible to FOSS developers – besides the fact that they belong to only one profit-making company. You may release the 100% perfect FOSS filter for .doc files, if you work night and day for six months. But Microsoft could change the specification and release the corresponding service pack in the next six days, negating all that glorious work without breaking a sweat. That war cannot be won.

Well, why not .rtf then? That's only text, no funny binary stuff – isn't it?

It is indeed. But it still belongs to Microsoft, and is not XML. This latter fact is important too, because the more the content of a file is structured using standard technologies, the easier it is to automate its generation and processing with standard (GPL of course) software tools.

You mean that OpenDocument is not going to provide any extra sophisticated functionality never heard of before?

Stop obsessing about new features already! The real reason we need universal support for OpenDocument is to guarantee that your data will always be completely readable, editable, indexable and searchable. Today, tomorrow and for the rest of your life – whatever application you use.

"WE FOSS USERS WILL NO LONGER BE DISCRIMINATED AGAINST BECAUSE OUR FILES LACK PERFECT FILTERS."

can be much more important than single software programs or even, in some cases, software licences. Intrinsically, formats have nothing to do with this or that program. They are like alphabets: if they are secret or expensive or change too fast nobody could communicate any more! Also, the information stored in such ephemeral formats would never really belong to whoever wrote it.

Why not? Surely once I've written something, it can't be changed.

Imagine that you wrote something, in English, using our 26 characters. Now imagine that someone abolished the Roman alphabet and replaced it with

relatively simple. Just ask around for more than ten minutes and you're sure to find other users who had problems opening complex tables or equations – to name just two of the compatibility headaches. These limitations may be OK when you're writing a shopping list, but they're unacceptable for corporations and public administration.

Sure, but why add another format to the confusion? Why not improve the current filters for what is the de facto standard?

There is nothing wrong with doing that, but in the long term it cannot accomplish anything relevant.

But none of my files is that important. I won't need my shopping list in 50 years' time.

Wrong. Who ever said that your only files are those in your hard disk? Think about your pension or mortgage payments. What if, sometime between now and 2040 when you will need them, the only paper copies get lost?

No problem – my bank and pension department keep digital copies in Office XP format... Oh, wait...

See? This is what OpenDocument is really about, and why you should ask that it is supported everywhere. A format with wide industry support and

no restrictions on use is much more likely to remain completely usable than a proprietary format that's only supported by one or two companies. More on this subject at Groklaw, **www.groklaw.net/article. php?story=20050130002908154**.

>> Is anyone actually using it yet?

Of course! OpenDocument is already the default file format for *OpenOffice.org*. And if you aren't ready yet for the jump to version 2, upgrade to *OO.o1.5*: it's just like the version you had before, but it can manage OpenDocument files.

>> What about other programs?

Support is present elsewhere, even if it's not as complete as that offered in *OO.o*. *KOffice* (**http:// koffice.kde.org**) is moving towards native support of OpenDocument, and you can already open and edit the corresponding files (with some limitations) with the current version of this suite. And the cross-platform word processor *AbiWord* (**www.abisource. com**) can read OpenDocument text files (.odt). The best known representative in the proprietary software field is *IBM WorkPlace*. Other server and desktop products compatible with OpenDocument are listed at **http://opendocumentfellowship.org/ Applications/HomePage**.

>> Is OpenDocument complete, or still in development?

OpenDocument version 1.0 has been available for several months. That said, nothing is perfect. If, by magic, every office suite in the world became 100% OpenDocument-compatible tonight, there would still be two non-portable things that you could still find in an OpenDocument file. The first is macros, the second formulas.

>> Why are macros so difficult?

Macros are often embedded, or distributed, inside office files. The OpenDocument specification allows this, but doesn't impose any language or other characteristic for macros. The main reason for this is that most macros extend the functionality of a particular application, not of a file.

>> And formulas?

Formulas are a much more serious concern, for two reasons. As we've said, several suites are already supporting (or are working

towards supporting) OpenDocument – but none of them is consistently implementing formulas. The base 10 logarithm of a number, for example, is expressed as LOG(X;10) in *OO.o* and as LOG(X) in *KOffice*. Luckily, developers are already actively working on this problem.

>> Is OpenDocument encumbered by patents and similar?

No. OASIS has committed to keeping OpenDocument freely usable by anyone, without enforcing any patent or similar restrictions. The complete policy is at **www.oasis-open.org/ committees/office/ipr.php**. Other OASIS standards do not, unfortunately, enjoy the same status, but OpenDocument can be safely adopted by free software users.

>> If anyone can use it, does that mean that Microsoft could too?

Why, yes! As a matter of fact, there is already an online petition (**http://opendocumentfellowship. org/petition**) to show Microsoft that people want it to support OpenDocument in its products.

>> Wait a second. If this is true, OpenDocument will hurt free software, won't it?

Yes, when *Microsoft Office* uses OpenDocument, several users will have one less reason to switch to *OO.o* and Gnu/Linux systems. At the same time, however, it will be much easier for them to migrate without losing data. Above all, those who already run free software will no longer be discriminated against because our applications don't have perfect filters for proprietary formats.

>> Is anybody else besides OASIS promoting OpenDocument?

Yes – the OpenDocument Fellowship

(**http://opendocumentfellowship.org**), a volunteer organisation with members around the world. Its goal is to promote the universal adoption of the OpenDocument format and support the development of software designed to operate on it.

>> Are there any resources for developers?

The official OpenDocument specification is available in the document section of the project's homepage (**www.oasis-open.org/committees/tc_home. php?wg_abbrev=office**). Python, Ruby and Perl modules to create and process these files already exist. To learn how to use them, start with the online book *OASIS OpenDocument Essentials* (**http://books. evc-cit.info**).

>> So what's the next step for OpenDocument?

It needs to be ratified by the International Organization for Standardization (ISO). Given ISO's authority, this procedure (which has already started), will definitely make OpenDocument suitable for widespread adoption in large organisations.

>> Large scale as in...

In October the Commonwealth of Massachusetts declared that it will only use OpenDocument for archiving its files. OpenDocument has also been declared by the European Union to be the only XML format that satisfies its criteria for openness. This is no fly-by-night standard, you know... **LXF**