



How Open Can Europe Get?

A Position Paper by OpenForum Europe on Open Document Formats for Public Administrations across Europe

November 2004

OpenForum Europe was set up to accelerate, broaden and strengthen the use of OSS including Linux, within business - importantly within the context of open verifiable standards. 'Not for profit' and independent it draws its membership from both the supply and user communities. OpenForum Europe acknowledges all the input received from its members and partners in the compilation of this document. However, it does not seek to represent the OSS developer community nor present its opinions as being unanimously supported by its full membership.

www.openforumeurope.org

Winkworth House, 83 St. Judes Road, Englefield Green. Surrey. TW20 0DF, UK
Tel: +44 1784 473005 Email: info@openforumeurope.org

1. Introduction

In April 2004 OpenForum Europe published a white paper on the contribution Open Source Software can make to Interoperability in Europe. This paper argued that an essential goal of the EU is to improve efficiency for both citizens and business, maximising the opportunity of operating in a cohesive environment. It recognised the particular integration issues of Europe dwarf the needs of the US, and that we talk of common processes, eGovernment, and a variety of programmes that are fundamentally there to promote interoperability.

Interoperability is a phrase that the ICT sector has seized on for years to promote independence, avoidance of proprietary lock in, and maximisation of market competitiveness. But it is only recently that the potential pitfalls have been fully recognised. These have largely been as a result of limited competition, and dominance in the market by a single supplier. So what can Europe do? What is the reality of interoperability in the real world?

The last White Paper presented the issue of Interoperability in the context of Open Source Software, against a background of the needs of the individual and community and not from a technologist viewpoint.

This White Paper is the product resulting from a briefing /discussion held by OFE in early September which explored the issues surrounding Open Document Formats. Open Document Formats have been identified as a key area of interoperability, and for this reason have been identified for specific focus in this series of white papers – ‘How Open Can Europe Get’.

2. A User-Led Definition

The IDA programme of the EC (See note 1) published earlier this year a timely Working Paper, “Linking up Europe: the Importance of Interoperability for eGovernment Services”, which has drawn some conclusions and provided pointers to success.

We selectively quote, “Interoperability is not simply a technical issue... it goes beyond this to include the sharing of information... and the reorganisation of administrative processes to support the seamless delivery...”. It does not stop at national or administrative boundaries, linking together organisations, administrations, enterprises or citizens. To be effective it has to cover the three aspects of,

Technical interoperability– definition of open interfaces, protocols etc

Semantic interoperability– meaning of exchanged information

Organisational interoperability– aligning business processes, information architectures

This is a concise analysis that recognises the role of the individual as well as the process and infrastructure in guaranteeing openness.

Technical aspects of interoperability are necessary but not uniquely sufficient to achieve working interoperability. To achieve the levels of interoperability and ease of use which users are demanding requires a concerted effort. Given the rapid advance of technology and users’ expectations, it is essential that policy makers avoid defining interoperability solely as a technical issue.

Ultimately, interoperability should be defined from the users' perspective.

“Interoperability” is the fulfillment of users' need to exchange and use information among various devices and software products from multiple vendors or service providers. Technical barriers to interoperability should only be those resulting from limitations in technology. They should not be intentionally introduced or sustained by vendors or service providers except where required to meet security or other external requirements. In these cases, the rationale for the existence of barriers to access should be evident to the user.

3. The Need for “Openness”

There is general consensus in global government around the world that “Openness” is both pivotal and a positive. However, many people debate the meaning, implications and limitations of “Openness”. The definition “Open” is also a potentially critical dynamic in the competitive landscape of the ICT industry. As a result there are various points of views, and conflicting definitions. So how do we come to terms with what is open? How can openness be leveraged to support dynamic, responsive, and cost effective government?

Governments must be “open” to their primary clients - citizens and businesses. Ensuring wider and equitable access to government services and enhanced responsiveness is the objectives of eGovernment. “Open” in this context implies that public administrations allow access to government applications via a choice of platforms based on a variety of technologies which do not to impose a single platform, operating system or HW configuration on the general public.

Information systems are essential to permit governments to deal with the complexity, globalization, security, mobility, and fiscal demands and constraints. Government has the right to define the configuration and flexibility of their information systems. ICT software and hardware are now as essential to national infrastructure as rail, road and utility systems.

Those systems should communicate seamlessly with other required systems. They need to be easily reconfigurable. They need to have the flexibility to source technology from a variety of vendors and leverage innovative emerging technology from any source originating under any development model. Enabling this flexibility is at the essence of “Openness”.

Openness is not a political statement, development method, or assessment of competing economic models. Openness is simply a means to an end. It is essential that we do not lose sight of the objective. The objectives of "Openness" include:

- Ensuring flexibility
- Ensuring interoperability
- Avoiding vendor lock-in
- Avoiding imposing technology decisions on the community
- Driving cost effectiveness
- Ensuring future access to information
- Ensuring a level playing field

- Maximizing freedom of action
- Supporting social inclusion

Technical and organisational interoperability in particular create an innovation-friendly environment for effective business model innovation and functional/technical innovation. This creates real economic value and an entry point for SMEs. An open IT ecosystem is a core part of an innovation friendly environment for the evolution of IT systems and new business models.

Document formats have a powerful networking effect. The requirement to exchange, access and archive documents with fidelity re-enforces the usage of a limited number of file formats. If this format is proprietary and controlled by one vendor, then the principles of openness are violated. Government's flexibility is reduced, public access is limited. There are significant technological and intellectual property constraints that compound the difficulty of other vendors to participate in delivering government services and spur innovation.

This lack of alternatives implies inefficiency in the market and low value for money. Public administrations are locked in to the technology, strategy and pricing of a single supplier. There is often a defacto absence of alternatives in practice, if not in theory. The use of a proprietary document file format implies that a particular technology decision is imposed on administrations, businesses and citizens.

The public authorities may not have the technical capability or the rights to view and manipulate archived documents at some future date – of which they are the intellectual originators and owners. As a result of concerns and the failure to address the confusion in this area, many governments feel a compelling impetus to move to a common “Open” document format.

4. Open, Verifiable Standards

“*Open, verifiable standards*” form the first part of the opportunity from Openness. However, it is only too easy to be drawn into the world of the academic definition rather than the pragmatism that is necessary for organisations to draw full benefit. The characteristics of openness when applied to standards are:

- Published without undue or unreasonable restriction, and
- Controlled by an independent organization with an open, well-defined process for evolution of the standard, not by a vendor.
- Freely available for adoption by the industry without constraint,
- Implemented by offerings that are available in the market.

Lack of restriction allows any vendor, or a public administration itself, to view, understand and design implementations that can create and manipulate documents stored in the format.

Control by an organisation allows for stability and the assurance of open and fair access. Requirements can be openly submitted, prioritised, and incorporated in the specification. The industry organisation can also facilitate the proper evolution of a specification in the broader ecosystem of related and complementary open standards.

Free availability removes artificial barriers to entry to the market. It promotes healthy competition, cost effectiveness and the delivery of innovative alternatives. It also allows for very specialised and niche solutions that may need to use a document format. Such access must as a minimum be on a reasonable, non-discriminatory basis, and preference without royalty.

Implementation by various vendors in the market implies choice, alternatives, and longevity.

In this way proprietary “standards” may satisfy such a pragmatic definition of openness provided they are then endorsed by a consortium or standards setting body. Standards evolve and mature driven by pragmatism, lifecycle, take-up rates and efficiency. Examples that many people are familiar with include HTTP, HTML, WAP, TCP/IP, VoiceXML, XML, and SQL. They are typically built by software engineers from various IT/software companies who collaborate under the auspices of organizations such as W3C, OASIS, OMA, ISO, and IETF.

Such a pragmatic approach is illustrated firstly by OpenOffice which originated in the proprietary world and secondly PDF, which despite being introduced by Adobe, is widely accepted by the industry as open. It has been used for the publication of this white paper and is commonly found on EU websites.

Insistence on openness does need to be tempered with such pragmatism. In reality, openness is a continuum. There are varying degrees of openness. There may be situations where public authorities may have requirements that can be addressed best by new technologies that have not fully exploited the standardisation process. However, intent to make full use of the process should always be a pre-condition of selection.

In these situations, and especially when considering technologies that impact the access, exchange and interaction of information, it is wise to select technologies that are further along the open continuum. In general, the best procurement policies are those which meet the technical requirements and adhere to the principles of openness, which incorporate user requirements and cost effectiveness as primary drivers.

In contrast, “Proprietary” describes interfaces that are developed by and controlled by one company and which have not been made freely available for adoption by industry. Proprietary software uses non-public interfaces or formats. When an interface is non-public, the owner of the proprietary interface controls the interface, including when and how the interface changes, and whether, how and who can adopt it and the terms and conditions of use.

Most public authorities and most software vendors are very supportive of the principle of **Open Document Formats**. There are a number of alternative approaches being advocated and each has merit. The two significant alternatives are, as the Valoris report commissioned by the European Commission concluded, the Office 2003 XML and the OpenOffice.org XML formats.

The Office 2003 XML format is being supported by Microsoft Inc. and OpenOffice by the Organization for the Advancement of Structured Information Standards (OASIS) and will subsequently be submitted to ISO.

5. OpenOffice.org XML Format

The OpenOffice.org XML file format currently stands out from the alternatives in terms of openness, completeness, and wide implementation. It has been under development since 1999, and has been in the hands of an independent Technical Committee at OASIS since 2002.

The OpenOffice.org XML file format has advantages over alternatives because:-

- it is not controlled by any single vendor,
- its definition is in the hands of a Working Group,
- the working group has open membership
- there is a formal change-control process,
- it is not hampered by any legal or licensing conditions.

The Working Group published its consensus Committee Draft in early 2004. It is in the process of finalizing the file format as an OASIS standard. Following this, it has been proposed, and agreed to, that the format should be submitted to ISO to be considered for ratification as a true International Standard.

The OpenOffice.org XML format, therefore appears to comply with the defined conditions of an open solution . It also covers all major types of office documents: presentations, word processing documents, and spreadsheets. Furthermore, it uses XML Namespaces properly to ensure that these types of documents can be combined in an efficient, modular fashion.

Due to the availability of the OpenOffice.org software the OpenOffice.org XML format can be used on all major desktop platforms including Microsoft Windows, Linux, Mac OSX and the Solaris operating system. In addition, OpenOffice.org and derivative products are available for 44 languages with many more under development. Since the XML file format is the default in the OpenOffice.org suite most OpenOffice.org users are already creating open XML document files.

Finally, the OpenOffice.org format has been implemented by multiple vendors, including IBM, Sun, Novell, KingSoft, and Red Hat. It is proven, reliable, and interoperable, and would seem a prudent choice for those who are concerned with the openness, availability, and longevity of their office documents.

6. The Microsoft Office 2003 XML Reference Schema

The Office 2003 Office Reference Schema is an XML based file format which has been published and will be licensed royalty free with minimum restrictions. This bodes well for interoperability; nevertheless, there remain a number of issues:

The most significant concern is that WordML is not really the document format of Microsoft products. The underlying file formats remain .doc; .xls, and .ppt . These are the formats in which files are stored when an end user clicks “save”. User intervention is required if the user wants to avoid saving a file in .doc, .xls or .ppt and save the file in WordML. The WordML format will never be used in significant volume as long as .doc and related formats are the defaults.

Word itself has long had the ability to save documents in various formats including 3 web formats, RTF, 3 variations of plain text, 6 versions of word, 5 versions of WordPerfect, and 2 versions of works. The reality is that the default setting and the network effect of the requirements for file interchange forces end users towards adoption of the default binary file formats. Should government department applications be allowed to bias users to one particular format by specific auto save functions?

In these circumstances, the network effects risk driving an increasing dependency on unique file formats and by extension a dependency on specific software products. To date, the Microsoft Office 2003 XML reference schemas are incomplete and do not include definitions for presentation material, such as a XML representation for PowerPoint. Therefore, this proposed solution as an open file format does not alter the status quo.

A second issue is that a mixture of proprietary data with the XML file format could hinder interoperability. In some situations, tags such as like <bindata>, where proprietary, platform specific, non portable information which is unreadable, is mixed with the OfficeXML render the document non-portable. The net effect of these extensions is that vendors can restrict other applications from properly viewing, accessing and manipulating a file created with their application. The result is a potential extension of the network effect based control that a vendor may have, introduce compatibility problems, and thus reduce a government's choice and flexibility.

Thirdly, the format is licensed and not "controlled by an open industry organisation with a well-defined inclusive process for evolution of the standard". There is no open process for requirements definition and prioritisation, which allows Microsoft the right to modify the specification without consultation. This seems to introduce considerable leverage and a competitive advantage as a result of that control. This means that the playing field is not yet level and that maximum value, choice and flexibility are not being delivered to public administrations.

Finally, there are legal concerns regarding the terms under which Office XML file formats can be used by other developers. The right unilaterally to change the licensing terms is the most disconcerting. There is room for real concern that competitive software development efforts could be impeded by legal means. Legal advice concerning this issue is recommend as per any large scale implementation.

7. Other Complementary Formats?

The essence of Open standards is the provision of choice and continuing competition and innovation in the market. One specific area relates to the wide spread use of PDF in relation to publishing formats.

The document life-cycle includes not only a need for open standards among office file formats, but also for open delivery and publishing formats. The requirements of delivery documents are different from those of office documents. Delivery documents need to preserve the look and feel of their originals without being impacted by user hardware and software environments. They need to ensure document authenticity, security, and in some cases provide document rights management and usage restriction features. Additionally, there is a growing requirement for the accurate archival and long-term preservation of electronic documents.

PDF has become a published specification and a de facto standard for delivery documents. It is, however, a commercial application; nevertheless, it has a high level of openness with a large and active community of OSS and commercial developers of PDF creation-manipulation tools. PDF-reading software is freely available and widely distributed across every major hardware platform and operating system. Additionally, PDF creation tools are tightly bound to both main office applications Microsoft Office and OpenOffice.

The capability of PDF to combine and preserve content from any application, including specialized ones like CAD and desktop publishing, can provide access to documents that otherwise may be unavailable to many users. Access can be provided to documents that might otherwise be unavailable to most users. The ability to comment and markup, implement metadata with open standard RDF XML, and for user-defined XML electronic forms are positive innovation examples. Digital signatures and content protection with document rights management have not been developed it to date in any other proposed solution.

PDF-based specifications are emerging as open ISO standards for specific usages. For example ISO standards are being developed around PDF for long-term preservation and archiving, engineering documentation, accessible content for the disabled, and ensuring pre-press document accuracy. As and when these PDF variants are ratified as ISO standards, the PDF specification will take on an even greater level of openness and acceptance

8. Recommendations

We commend the European Commission for taking the initiative to encourage vendors to support open file formats and engage in the OASIS technical committee discussions. OFE strongly endorses this initiative.

OFE recommends:

- That the European Commission and member states move towards the use of open document formats as a matter of principle.
- Support should be required in the tendering process for such open formats in software procured by public administrations.
- Software vendors should define the open format as a user default in software used by public administrations.
- Administrations should ensure that vendors implement ODF as defaults in their applications sold to government;

Which open document format is selected is not a political decision but a usability, accessibility decision. We strongly recommend that European public authorities select open document formats that comply with the principles of openness. OFE believes that the OASIS OpenOffice XML formats have considerable technical merit, relatively broad support with great potential. They will evolve in an appropriate manner with broad industry participation.

OFE would thus further recommend that:

- Public administrations across Europe evaluate the OASIS OpenOffice XML formats as a standard
- Open delivery and ISO-standardized archival file formats are considered to be important by administrations, because of increasing needs of document control and preserving documents for the long term, independent of any hardware platform or operating system.
- Microsoft are encouraged to consider the donation of schemas to OASIS.
- The Commission should encourage all vendors to participate in the technical committee at OASIS.

- The Commission and other Public Administrations should contribute requirements to the OASIS ODF technical committee.

Notes

1. IDA Interchange of Data between administrations www.europa.eu.int/ISPO/ida