STANDARDS UPDATE

David Q. McDowell, Editor

This issue of Standards Update focuses on a variety of document format standards being done by several different industry and standards groups. It is strictly coincidental that the reports of these various activities all seemed to arrive on my desk within a short period of time. So it seemed like a logical topic for this article.

It begins with a piece by Betsy A. Fanning, the Director of Standards & Content of AIIM. We will also look at some of the other document-related file format standards work that is going on in ISO and Ecma.

Betsy's article, which first appeared in the AIIM E-DOC Magazine, does such an excellent job of describing several standards efforts in the document and storage format area that I felt that it should be shared here.

AIIM was founded in 1943 as the National Microfilm Association and later became the Association for Information and Image Management. AIIM is now known as AIIM—The Enterprise Content Management Association. The AIIM E-DOC Magazine can be found at www.edocmagazine.com and the AIIM home page is www.aiim.org.

AIIM is the Secretariat for ISO TC 171, the Secretariat for ISO 171 SC 2, and Secretariat for the USTAG for ISO 171. Betsy is responsible for all those activities.

Document and Storage Formats–ODF, OOXML, XPS, and XAM

by Betsy A. Fanning

With this column we are taking a brief step back from our normal focus on AIIM standards work to take a look at four other standards efforts that bear examination due to their relationship to AIIM's document and content management area. The new document formats being introduced to the international standards process include Open Document Format; Office Open XML, and XML Paper Specification. There is also a storage standard, XAM, that we will be taking a brief look at. The common thread across all four of these standards is that they are all based in XML, eXtensible Markup Language.

These standards efforts are originating in different standards development organizations prior to being submitted to ISO, the International Organization for Standardization for ratification. OASIS, the Organization for the Advancement of Structured Information Standards (www.oasis-open.org) is responsible for the Open Document Format and Ecma International (www.ecma-international.org) is responsible for the Open XML and XML Paper Specification (XPS) standards. SNIA, the Storage Network Industry Association (www.snia.org) is responsible for XAM.

Before discussing these standards, it is important to take a step back and review a few basics so that we can better understand the benefits each will bring to the industry. Filename extensions, those letters after the dot in a filename actually tell us about the format of the file. Not only do those letters tell us the format but they also tell us what type of data is contained in the file. In many cases the filename extension or file type will initiate the appropriate viewer for the file so that the file recipient will be able to view the file as its creator had intended it to be viewed. For example, a .pdf file will cause a PDF viewer to display the document; a .doc will have a word processing application open the file and a .tif will display the document image using an image viewer.

In early 2006, IBM, Novell, Oracle, and Sun Microsystems joined with end users to develop an XML-based Open Document Format for office documents. Their initial goal was to raise awareness of governments to the benefits offered by an open file format and the assurances of easy access, retrieval and use that such a file format would provide. By standardizing the file format governments would have more control over their documents and realize benefits like making their processes more efficient and flexible which result in better customer service to the constituencies that they serve. Open Document Format (ODF) is viewed as the only open standard for office documents that is completely vendor neutral.

The Open Document Format standard was approved as an OASIS standard in May 2005 and submitted to ISO/IEC JTC1 (International Organization for Standardization International Electrotechnical Commission's Joint Technical Committee) for approval as an ISO standard. This standard defines an XML schema for office applications including the associated metadata. The schema covers all types of office documents, including text documents, spreadsheets, charts and graphical documents like drawings or presentations, but is not restricted to just these documents. It defines XML structures for office documents that will easily translate to other applications by using XSLT or similar XML-based tools. The OASIS developed OpenDocument v1.1 provides an introduction to the format, describes the metadata that is contained in the documents, and describes the text and paragraph content, tables, text fields, etc. OpenDocument makes use of existing standards such as HTML, SVG, XSL, SMIL, XLink, XForms, MathML and Dublin Core wherever possible to promote interoperability.

The Open Document Format (ODF) Alliance (www.odfalliance.org) was formed in 2006 to promote and advance the use of the Open Document Format (ODF) as the primary format for governments. Through the use of this open standard, the ODF Alliance hopes to enable governments and their constituents to manage critical documents in a truly open environment that is independent of applications or enterprise platforms used to create or access the documents.

In response to numerous requests for a way to ensure that their organization's in-

vestment in the Microsoft Office Suite is protected, Office Open XML or OOXML was developed. Office Open XML or OOXML is an XML based file format for all types of electronic office documents that will promote interoperability and further enable workflow or business process management in the office. It is intended to simplify the exchange of information between Microsoft Office products and enterprise applications. Open XML is based on the industry accepted XML standard and ZIP technologies. It was designed for the unstructured content created on products developed by the Microsoft Office products. It is a container format with specialize XML-based markup languages being used that correspond to the individual applications within the Microsoft Office product line. It defines multiple vocabularies. Basically, OpenXML a modular, componentized representation of the document storing a chart, table or executable, for instance, in a separate component that is properly displayed when the document is viewed. Using ZIP technology with the file format, means that the files created using OOXML may be smaller. OOXML includes Office MathML, a mathematical markup language, which while different from the MathML defined by the W3C (Worldwide Web Consortium) is compatible.

XML Paper Specification or XPS is a document storage and viewing specification developed by Microsoft that describes electronic paper in a way that it can be read by hardware, software and the human eye. XPS provides a page view of the way the document will print. It describes the appearance of fixed format documents by using an XML based format so that the layout won't change. XPS is viewed as a potential competitor to PDF (Portable Document Format) but will not replace PDF in the instances where dynamic content capabilities are required which XPS cannot handle. By dynamic content, we mean the type of content that may be contained in a drop-down on a form. Microsoft released XPS with a royalty-free patent license to encourage wide adoption by the industry. The XPS document format is included with Windows Vista and 2007 Microsoft Office System products. The XPS viewer is included with Windows Vista and is available for Windows XP and Windows Server 2003. The viewer allows users to open, read and apply digital signatures to XPS documents without needing the full XPS generating software. XPS allows electronic documents to print better, be shared easier, archived and better maintain the security of the information in the document.

XPS relies on two additional specifications- the Open Package Conventions which describe the method for creating packages of content, resources, and metadata in a compressed ZIP file and Open XML Markup Compatibility Specification that describes the mechanism to support versioning and extensibility of XML. Ecma International's TC46 chaired by Martin Bailey, Global Graphics, will continue the work on the XPS standard and facilitate its fast track processing at ISO to become an ISO standard. XPS is platform independent.

XAM, the eXtensible Access Method Specification developed by SNIA, the Storage Networking Industry Association defines a standard access method between application software and storage systems to manage fixed content reference information storage services. The intent of this specification is to provide an interface for the longevity and mobility of reference information that is independent from any specific storage system technology, storage solution or the location of a storage system or data. Possible implementations of XAM include interfaces between applications and storage systems that coordinate metadata to achieve interoperability, storage transparency, and automation for ILM (information lifecycle management) practices, long term records retention and information security. The XAM specification will enable storage products to seamlessly integrate, giving end users more flexibility when selecting storage technology.

The work on this specification began in October 2004 as a project between IBM and EMC Corporations. There were later joined by HP, Hitachi Data Systems and Sun Microsystems.

A new working group was formed in April 2007, to create a protocol based on the XAM specification for storing the metadata that describes fixed content, such as email, medical records and financial data. The working group will create a software development kit to help write applications to the XAM specification. This latest effort is expected to be completed by early 2008 and will focus on email archiving.

This effort is being facilitated by SNIA, the Storage Networking Industry Association (SNIA) which is a not-for-profit global organization, made up of more than 460 member companies and close to 7,000 active individuals spanning virtually the entire storage industry. For more information on XAM, please visit SNIA at www.snia.org.

The amount of adoption that these standards is really up to you, the end

IS&T REPORTER

The *IS&T* Reporter is published bimonthly by IS&T—The Society for Imaging Science and Technology.

Editorial

Executive Editor	*Vivian Walworth*
Managing Editor	*Donna K. Smith*
Production Editor	*Itzel Morales*
Standards Editor	*David McDowell*

Articles and letters to the editor in this newsletter do not necessarily constitute endorsement or reflect the opinions of the editors or IS&T. Advertising and copy are subject to acceptance by the editors.

IS&T is an international, non-profit society whose goal is to keep members aware of the latest developments in fields related to imaging science through conferences, journals, and other publications. IS&T focuses on imaging in all its aspects, with particular emphasis on silver halide, digital printing, electronic imaging, color science, photofinishing, image preservation, pre-press technology, and hybrid imaging systems.

IS&T publishes the *Journal of Imaging Science & Technology* and, in conjunction with SPIE, the *Journal of Electronic Imaging*.

users. You and your organization will need to carefully evaluate each of them to determine the appropriateness of them for your organization in light of the technology that you have installed and make the best possible decision that you can. The decision process won't be easy as you will most likely encounter support for each of them in your organization. Watch the AIIM Standards Watch Blog at http://aiimstandardswatch.typepad.com for developments on these standards. Also, visit the AIIM Standards web pages at www.aiim.org/standards to see how you can become involved in AIIM's standards work.

Thank you Betsy.

From the ISO Central Office

On September 4, 2007 ISO issued a press release that included the following:

Vote closes on draft ISO/IEC DIS 29500 standard

A ballot on whether to publish the draft standard ISO/IEC DIS 29500, Information technology–Office Open XML file formats, as an International Standard by ISO (International Organization for Standardization) and IEC (International Electrotechnical Commission) has not achieved the required number of votes for approval.

The five-month ballot process ended on 2 September and was open to the IEC and ISO national member bodies from 104 countries, including 41 that are participating members of the joint ISO/IEC technical committee, JTC 1, Information technology.

Approval requires at least 2/3 (*i.e.*, 66.66 %) of the votes cast by national bodies participating in ISO/IEC JTC 1 to be positive; and no more than 1/4 (*i.e.* 25 %) of the total number of national body votes cast negative. Neither of these criteria were achieved, with 53 % of votes cast by national bodies participating in ISO/IEC JTC 1 being positive and 26 % of national votes cast being negative.

ISO/IEC DIS 29500 is a proposed standard for word-processing documents, presentations and spreadsheets that is intended to be implemented by multiple applications on multiple platforms. According to the submitters, one of its objectives is to ensure the long-term preservation of documents created over the last two decades using programmes that are becoming incompatible with continuing advances in the IT field.

ISO/IEC DIS 29500 was originally developed as the Office Open XML Specification by Microsoft Corporation which submitted it to Ecma International for transposing into an ECMA standard. Following a process in which other IT industry players participated, Ecma International subsequently published the document as ECMA standard 376.

Ecma International then submitted the standard in December 2006 to ISO/IEC JTC 1, with whom it has category A liaison status, for adoption as an International Standard under the JTC 1 "fast track" procedure. This allows a standard developed within the IT industry to be presented to JTC 1 as a Draft International Standard (DIS) that can be adopted after a process consisting of a one-month review by the national bodies of JTC 1 and then a five-month ballot open to all voting national bodies of ISO and IEC.

Other Activities

XML Paper Specification (XPS)

Ecma International has also recently created a technical committee, TC46 - XML Paper Specification (XPS), to standardize XPS. The published scope of ECMA TC46 is as follows.

ECMA TC46 Scope: The goal of the Technical Committee is to produce a formal standard for an XML-based electronic paper format and XML-based page description language which is consistent with existing implementations of the format called the XML Paper Specification (XPS).

The Technical Committee will use the format called the XML Paper Specification (XPS) as a starting point with the aim to provide a standard, secure, and highly trustworthy format that enables a wide set of applications, devices, tools and platforms to implement compatible paginated-document workflows. An additional goal will be to enable the interoperability of independently created software and hardware systems that produce, consume or otherwise process XPS content. The Technical Committee will be responsible for the ongoing maintenance and evolution of the standard.

XPS is describes as a document storage and viewing specification developed by Microsoft that is intended to describe electronic paper in a way that it can be read by hardware, software and the human eye. XPS is intended to provide a page view of the way the document will print. It will describe the appearance of fixed format documents by using an XML based format so that the layout won't change. It is a key component of the Microsoft VISTA Office Suite

ISO 32000, Document management —Portable document format—PDF 1.7

ISO TC 171/SC2 (Document management applications/Application issues) has recently introduced ISO 32000, D*ocument management—Portable document format —PDF 1.7* for fast-track ballot. If approved, it is the intent of TC171/SC2 to form a Working Group within SC2 to carry this standard to completion and to begin the development of the next ISO version of PDF.

In discussions within the US shadow committee supporting this work it was felt to be important to provide a statement to the technical experts of other national bodies that described the intent of this standard. That statement said in part:

"The U.S. National Body committee that proposed and is supporting the work of TC171/SC2 with respect to the work item ISO 32000 wishes to communicate some background information relevant to this work and urges you to support this work in your own national body deliberations.

In preparing ISO/DIS 32000, Document management—Portable document format (PDF 1.7), the source material, Adobe's PDF Reference 1.7 (including associated errata), was edited to create a document that is both easier to read and understand, and more in line with the

style requirements of the international standards community.

The key goals during that process were to ensure that the resultant document was technically consistent with the PDF Reference 1.7 published by Adobe in November 2006 and to incorporate the interpretations of that specification provided by PDF application developers. We supported this action.

The members of the U.S. technical committee who have reviewed the ISO draft and PDF Reference 1.7 agree that these documents are technically consistent."

The ballot on ISO 32000 closes on December 2, 2007.

PDF/X

ISO TC130 recently completed the final balloting of PDF/X-4 (ISO 15930-7, *Graphic technology—Prepress digital data exchange using PDF—Part 7: Complete exchange of printing data (PDF/X-4) and partial exchange of printing data with external profile reference (PDF/X-4p) using PDF 1.6) and PDF/X-5 (15930-8, Graphic technology -- Prepress digital data exchange using PDF -- Part 8: Partial exchange of printing data using PDF 1.6 (PDF/X-5)).*

These documents were approved and are in publication at ISO Central Secretariat. Publication is expected within the next 2 months.

vPDF

TC130 recently approved a new work item (NWI) for the development of ISO 16612-2 which has a tentative title of *Graphic technology—Variable printing data exchange —Part 2: Using PDF/X-4 and PDF/X-5 (VPDF/X-200X).*

The proposed scope of this standard is:"This part of ISO 16612 specifies the use of the Portable Document Format (PDF) Version 1.6, as restricted by PDF/X-4 and PDF/X-5, for the exchange of all content elements and metadata necessary to produce a variable data printing job as intended by the sender. Externally referenced content and metadata includes provision for unambiguous identification.

This part of ISO 16612 specifies document layout, content elements, and interaction of content elements in a transparency-based graphics model. To make use of this content in a print product and corresponding production process requires the use of JDF, or similar job description protocol.

Colour-managed, CMYK, Gray, RGB or spot colour data are supported in any combination; as are PDF transparency and optional content (layers). Files may be prepared for use with Gray, RGB, CMYK and n-colorant printing characterizations."

This work is being done in TC130/WG2/TF3 under the leadership of Tim Donahue of Kodak and Dov Isaacs of Adobe.

For suggestions for (or input to) future updates, or standards questions in general, please contact the author at mcdowell@npes.org or mcdowell@kodak.com