NEW YORK STATE CHIEF INFORMATION OFFICER

CIO OFT

NEW YORK STATE OFFICE FOR TECHNOLOGY

A Strategy for Openness

Enhancing E-Records Access in New York State
Part II: Supporting Documentation

Submitted to:

The Honorable David A. Paterson, Governor
The Honorable Joseph L. Bruno, Temporary President of the Senate
The Honorable Sheldon Silver, Speaker of the Assembly

Executive Co-Sponsors

Dr. Melodie Mayberry-Stewart
New York State Chief Information Officer and Director,
New York State Office for Technology and Office of the
New York State Chief Information Officer

Christine Ward
New York State Department of Education
Assistant Commissioner for Archives & Records
and New York State Archivist

May 2008

Part I – Executive Summary

(Separate Document)

Part II – Supporting Documentation

Table of Contents

Appendix I 4
GLOSSARY 4

Appendix II 11
THE WORKGROUP'S DETAILED FINDINGS: 11
A. Background 11
B. Specific Issues Raised In The Request For Public Comments 18
C. Some Approaches That Were Suggested For The State 34
D. Implementation Issues 44
E. Sustained Operationalization Of Openness 61
F. Conclusions 69

Appendix III 72
COMPARATIVE CHART: 72
National Governments requiring use of open formats 72

Appendix IV 78
COMPARATIVE CHART: 78
Provincial/State/Regional Governments requiring use of open formats 78

Appendix V 81
COMPARATIVE CHART: 81
U.S. State Government Collaborative Processes for Addressing e-Record Policies / Standards 81

Appendix VI 84
PROPOSED NEW YORK STATE ELECTRONIC RECORDS COMMITTEE 84

Appendix VII 87
SPECIFIC NEW YORK STATE ARCHIVES E-RECORDS NEEDS 87
A. State Archives E-Records Resource Needs 87
B. State Archives E-Records Transfer Needs 89
C. State Archives E-Records Testimony 93
D. State Archives Attachment to Testimony 100

Appendix VIII 106
ANALYSIS OF OPEN SOURCE AND NYS GOVERNMENT 106

Appendix IX 123
ENDNOTES 123

Part III - Results of Request for Public Comments

(Separate Document)

Appendix I

GLOSSARY[1]

These are starting definitions at their most basic. Distinct sections of this Part II of the report discuss in detail, for many of these terms, the definition of the term which best suits the purposes of New York State.

- **Access:** Multiple definitions of this term are possible dependent upon context and various applicable laws. In most instances unless otherwise expressly stated, the fine-tuned distinctions do not affect this report. In general, access is the right, opportunity, means of finding, using or retrieving information, usually subject to rules and conditions.

- **Active Record:** A record that has not been closed and which is required for the day-to-day functioning of an agency or person.

- **ASCII/Unicode:** A text file format. ASCII (American Standard Code for Information Interchange) is the most common format for text files in computers and on the Internet. Unix and DOS-based operating systems use ASCII for text files. Windows NT and 2000 uses a newer code, Unicode. Contrasted with binary files, such as executable software (machine language programs), most word processing files and database, spreadsheet and multimedia files. Text and source program files as well as HTML and XML files are ASCII text files, not binary.

- **Assistive Technology:** Hardware and software that help people who are physically impaired. Often called "accessibility options" when referring to enhancements for using the computer, the entire field of assistive technology is quite vast and even includes ramp and doorway construction in buildings to support wheelchairs. Enhancements for using the computer include alternative keyboard and mouse devices, replacing beeps with light signals for the deaf, screen magnifiers and text enlargers and systems that form tactile Braille letters from on-screen text.

- **Born Electronic:** Electronic materials that are not intended to have an analogue equivalent, either as the originating source or as a result of conversion to analogue form. Used to differentiate materials from those that have been created as a result of converting analogue originals, and material that may have originated from a digital source but have been printed to paper, e.g. some electronic records.

- **Chief Information Officer (CIO):** The CIO is the executive officer in charge of information processing in an organization. All systems design, development and datacenter operations fall under CIO jurisdiction. In this report, unless stated otherwise use of this term refers to the New York State CIO. At the time of the report the New York State CIO is also Director of the New York State Office for Technology.

- **CIO Council:** The CIO Council was developed as a means of advancing the Governor's agenda for New York State and facilitating communication between the Chief Information Officer/Office for Technology and other State agencies and authorities. It comprises senior IT

leadership of state agencies, authorities, public benefit corporations and local governments. Exclusively for CIOs and chaired by the NYS Deputy CIO, its current membership includes representatives from more than 85 separate organizations, including ten local governments.

- **CIO Council Action Teams:** Sub-groups of the CIO Council, these teams address specific issues related to agency and interagency information technology concerns.

- **Conversion:** The process of changing records from one medium to another or from one format to another. Conversion involves a change of the format of the record but ensures the record retains the identical primary information (content). Examples include microfilming and digital imaging of paper records.

- **Data Exchange:** The transmission and receipt of information (data, audio and visual) via a computer-linked network, or from disk to disk.

- **European Computer Manufacturers Association (Ecma):** Ecma is an ISO member organization that establishes standards for the information technology and telecommunications industries. Ecma submitted the OOXML format for ISO standardization.

- **Electronic Data/Electronic Documents/Electronic Records:** Multiple definitions of these terms are possible dependent upon context and various applicable laws. In most instances unless otherwise expressly stated, the fine-tuned distinctions do not affect this report. In general, these are data, documents, or records that are created, transmitted, received, or stored in digital format.

- **Electronic discovery:** Also called e-discovery or ediscovery. Refers to any process in which electronic data is sought, located, secured, and searched with the intent of using it as evidence in a civil or criminal legal case.

- **Enterprise:** The entire organization, including all of its subsidiaries. It implies a large corporation or government agency, but it may also refer to a company of any size with many systems and users to manage. It depends on context. A corner candy store is "someone's enterprise." The terms "enterprise," "company," "corporation" and "organization" are used synonymously.

- **Enterprise Architecture/Information Architecture:** The interrelationships of systems in place in an organization. It is used to assist in creating systems that are interoperable rather than duplicating.

- **Enterprise Framework:** A complete environment for developing and implementing a comprehensive information system. Enterprise frameworks provide pre-built applications, development tools for customizing and integrating those applications to existing ones as well as developing new applications.

- **File format:** File formats are the structure of program and data files. Each has its own headers, codes and rules for laying out the content. There are many different file structures for each kind of file, including executable programs, word processing documents, graphics files and databases.

- **Gartner:** Gartner is the largest existing information technology consulting firm specializing in research and analysis. It

has predicted significant adoption of the ODF format.

- **HTML (HyperText Markup Language):** The document format used on the Web. Web pages are built with HTML tags (codes) embedded in the text. HTML defines the page layout, fonts and graphic elements as well as the hypertext links to other documents on the Web. E ach link contains the URL, or address, of a Web page residing on the same server or any server worldwide, hence "World Wide" Web.

- **Inactive Record:** A record that is not required to be readily available for the business purposes of a department or agency and may therefore be transferred to intermediate storage, archival custody or be destroyed subject to applicable laws.

- **Infrastructure:** The fundamental structure of a system or organization. The basic, fundamental architecture of any system (electronic, mechanical, social, political, etc.) determines how it functions and how flexible it is to meet future requirements.

- **International Standards Organization (ISO):** A non-governmental organization that is a network of the national standards bodies of 157 nations.

- **Interoperable/Interoperability:** The ability for one system to communicate or work with another. The capability of two or more hardware devices or two or more software routines to work harmoniously together. For example, in an Ethernet network, display adapters, hubs, switches and routers from different vendors must conform to the Ethernet standard and interoperate with each other.

- **Lock-in:** Vendor lock-in, or just lock-in, is the situation in which customers are dependent on a single manufacturer or supplier for some product (i.e., a good or service), or products, and cannot move to another vendor without substantial costs and/or inconvenience. This dependency is typically a result of standards that are controlled by the vendor (i.e., manufacturer or supplier). It can grant the vendor some extent of monopoly power and can thus be much more profitable than would be the absence of such dependency.

  The term is commonly used in the computer industry to refer to the situation that can occur due to a lack of compatibility between different hardware, operating systems or file formats. Such incompatibility can be intentional or unintentional. A specific way in which lock-in can be created is by a dominant company developing file formats that make it difficult for its users to convert their data to other formats.

  The costs of lock-in can be severe. They can include (1) a substantial inconvenience and expense of converting data to other formats and converting to more efficient, secure and inexpensive application programs and operating systems. They also include (2) a lack of bargaining ability to reduce prices and improve service, (3) vulnerability to forced upgrades and (4) the corruption, or even loss, of critical data while attempting to convert it.

- **Metadata:** Structured information that describes and/or allows users to find, manage, control, understand or preserve other information over time. Metadata is attached to records when they are created and added to as a result of different processes such as sentencing and disposal.

- **Native Format:** The format in which the record was created or in which the originating application stores records.

- **OASIS (Organization for the Advancement of Structured Information Standards):** OASIS is a nonprofit, international consortium whose goal is to promote the adoption of product-independent standards for information formats. It promotes the adoption of interoperability standards, and recommends ways members can provide better interoperability for their users. OASIS submitted the ODF format for ISO standardization.

- **Office Open Extended Markup Language (OOXML):** OOXML is an XML-based file format for saving and exchanging text, spreadsheets, charts, and presentations. OOXML was developed by Microsoft. It was submitted to the International Standards Organization (ISO) for standardization by the European Computer Manufacturers Association (Ecma). The standard is published by ISO as ISO/IEC DIS 29500, *"Information technology – Office Open XML file formats."*

- **Open/Openness:** The simplest definition concerns the extent to which technology is made to operate with other products. Discussion of a proposed detailed definition forms a distinct section of this Part II of the report.

  Open Document Format (ODF): The Open Document Format (ODF) is an XML-based file format for saving and exchanging text, spreadsheets, charts, and presentations. ODF was developed by a committee formed under the OASIS (Organization for the Advancement of Structured Information Standards) consortium. It was submitted to the International Standards Organization (ISO) for standardization by OASIS. The standard is published by ISO as ISO/IEC 26300, *"Open Document Format for Office Applications (OpenDocument)."*

- **Open Standards:** Specifications for hardware and software that are developed by a standards organization or a consortium involved in supporting a standard. Available to the public for developing compliant products, open standards imply "open systems"; that an existing component in a system can be replaced with that of another vendor. Open formats are a subset of open standards.

- **Open Source:** Software that is distributed with its source code so end user organizations and vendors can modify it for their own purposes. Most open source licenses allow the software to be redistributed without restriction under the same terms of the license.

- **Operating System:** The essential program that enables all other programs to be run on a computer, and which establishes an interface between a user and the hardware of the computer.

- **PDF (Portable Document Format):** PDF is a file format that has captured all the elements of a printed document as an electronic image one can view, navigate, print, or forward to someone else. PDF files are created using Adobe Acrobat, Acrobat Capture, or similar products.

- **Plan to Procure/Annual Technology Plan/Intent to Purchase:** A process consistent with the legal authority of the CIO/Office for Technology to help manage the State's Information Technology (IT) investments to fully leverage the State's buying power and create value for delivering better government services across the State enterprise. It is intended to establish a close strategic alignment with the

Administration's priorities and enterprise technology priorities and enable better identification and coordination of IT procurement opportunities among agencies in an effort to reduce duplication and redundant spending and increase the level of joint collaboration for successful IT solutions.

- **Port**: To convert software to run in a different computer environment. For example, the phrase "to port the application to Unix," means to make the necessary changes in the program to enable it to run under Unix.

- **Preservation**: The processes and operations involved in ensuring the technical and intellectual survival of authentic records through time. Preservation encompasses environmental control, security, creation, storage, handling, and disaster planning for records in all formats, including digital records.

- **Procurement**: Synonymous with "purchasing." The procurement department within an organization manages all the major purchases.

- **Productivity Suite**: A suite of business applications that usually includes a word processing program, a spreadsheet, a database program, a communications program, and a presentation graphics program. Also referred to as office suite software.

- **Proprietary Standards**: Specifications for hardware and software that are developed and controlled by one company. Proprietary standards are technically de facto standards such as Microsoft's Windows and Intel's x86 chip family.

- **Records Management**: The field of management responsible for the efficient and systematic control of the creation, receipt, maintenance, use and disposal of records, including processes for capturing and maintaining evidence of and information about business activities and transactions in the form of records.

- **Render**: To convert any coded content to the required format for display or printing. Although the term is typically used to refer to images, it may refer to any data. For example, an HTML page, which contains text and graphics, is said to be "rendered" when it is displayed.

- **Reference Design/Reference Schema**: A technical blueprint of a system that is intended for others to copy. It contains the essential elements of the system; however, third parties may enhance or modify the design as required. Extensible Markup Language (XML) schema definition language for defining the structure, contents and semantics of XML documents.

- **Request for Proposals**: A document that invites a vendor to submit a bid for hardware, software and/or services. It may provide a general or very detailed specification of the system.

- **Request for Public Comments (RFPC)**: A document issued by the State CIO/Office for Technology seeking comments from members of the public and identifying issues which this report proposed to address. An RFPC related to this report was released on December 12, 2007 with an initial comment response date of December 28, 2007. In response to requests from the public, the deadline for responding was extended to January 18, 2008.

- **Rip and Replace:** The term means basically starting from scratch, for example requiring an entity to replace its applications with new applications running on different platforms. Rip and replace strategies are widely considered expensive, complicated and highest-risk.

- **Software Obsolescence:** A situation where software is rendered obsolete because newer versions are not 'backwardly compatible' (able to read older versions of that software) or it is no longer used and has been superseded by other software or it cannot function with newer equipment or software.

- **Source Code:** Programming statements and instructions that are written by a programmer. Source code is what a programmer writes, but it is not directly executable by the computer. It must be converted into machine language by compilers, assemblers or interpreters.

- **Source Code Escrow:** Source code escrow means deposit of the source code of the software into an account held by a third party escrow agent. Escrow is typically requested by a party licensing software (the licensee), to ensure maintenance of the software. The software source code is released to the licensee if the licensor files for bankruptcy or otherwise fails to maintain and update the software as promised in the software license agreement.

- **State CIO/OFT Strategic Plan:** A plan that guides the State's IT decision-making, aligned with the overarching strategic direction for the state, with goals such as improving government operations through consolidating duplicative services, achieving operational efficiencies through automation, or reaching a greater number of citizens and businesses. The 2008 plan includes input from a variety of stakeholders concerned with the procurement of technology, the implementation of IT systems, and the delivery of e-government services, including state agencies and their commissioners, local governments, the Executive Chamber and technology companies.

- **Total Cost of Ownership (TCO):** TCO (total cost of ownership) is a type of calculation designed to help consumers and enterprise managers assess both direct and indirect costs and benefits related to the purchase of any IT component. The intention is to arrive at a final figure that will reflect the effective cost of purchase, all things considered. TCO analysis performs calculations on extended costs for any purchase - these are called fully burdened costs. Fully burdened cost may include costs of purchase, repairs, upgrades, service and support, networking, security, user training, and software licensing. The TCO is compared to the total benefits of ownership (TBO) to determine the viability of the purchase.

- **Use Case:** A use case is a methodology used in system analysis to identify, clarify, and organize system requirements. The use case is made up of a set of possible sequences of interactions between systems and users in a particular environment and related to a particular goal. It consists of a group of elements (for example, classes and interfaces) that can be used together in a way that will have an effect larger than the sum of the separate elements combined. The use case should contain all system activities that have significance to the users. A use case can be thought of as a collection of possible scenarios related to a particular goal, indeed, the use case and goal are sometimes considered to be synonymous.

- **World Wide Web Consortium (W3C):** An international industry consortium founded in 1994 by Tim Berners-Lee to develop standards for the Web. The W3C has standardized many of the fundamental technologies of the Web, including HTML and XML, URLs and URIs, the SOAP protocol and the P3P privacy description.

- **Extensible Markup Language (XML):** An open standard for describing data from the W3C. It is used for defining data elements on a Web page and business-to-business documents. By providing a common method for identifying data, XML supports business-to-business transactions and has become "the" format for electronic data interchange and Web services. The human-readable XML tags provide a simple data format, but the intelligent defining of these tags to serve business needs properly and everyone's adherence to using the same tags determines the real value of XML.

Appendix II

THE WORKGROUP'S DETAILED FINDINGS:

A. Background

In light of the workgroup's findings, the background behind the statute and after evaluating the extensive public comment, the workgroup identified the report's central focus:

> Having recognized that the use of open standards and formats can be a critical feature enabling State government to enhance its control over and public access to electronic records, the State needs to further explore means of securing that feature in its information technology infrastructure.

This focus strikes an appropriate balance between:

> (a) the benefits cited by proponents of greater IT openness, and
> (b) the concerns expressed by others that mandating specific document creation and preservation technology solutions might lead to failure to meet other equally compelling State needs or technology which could easily become outdated.[2]

Given all of the above, the primary recommendation of this report is that the State needs to take measured steps to integrate the desired features of open standards and formats into existing State technology procurement processes ... but not at the cost of other needed functionality.

What the Report is Addressing

The bill that authorized this study and report focused our efforts on determining what New York State should do to make sure its electronic records and electronic record systems remain open to all citizenry over time. It is no coincidence that this focus was captured in the main themes underlying most of the public comments received by the State, including the State's need to identify "ways to maintain effective access to State documents and records," the need to address "interoperability between competing products, so that users have the chance to substitute one solution with another without incurring major risks or costs," and the State's need to make format choices which "foster a market with multiple competing products" in order to reduce costs and avoid vendor lock-in. Overwhelmingly, commenters recommended that a solution to the problems identified would specifically contain the adoption of the Open Document Format (ODF) for word processing documents, spreadsheets, and slideshow-style presentations.

As one commenter noted, the State's failure in the past to avoid the "status quo" of substantial lock-in to the predominant office suite vendor and its products had acted to "preserve the technological disadvantages and difficulties that led to [this] study in the first place."

Proprietary File Formats and Applications

Many of the most commonly used file formats are the intellectual property of software companies that either develop software packages that can create and access files encoded in these formats or permit a small number of third parties to do so. In most instances, these software packages are commercial products subject to various use restrictions. Although proprietary formats and applications are, in many instances, feature-rich, they can impede the effective management and business use of electronic data in the following fashions:

- *Limited choice of software applications.* Users managing significant quantities of data saved in a proprietary format are in effect forced to use the small number of applications that can read data encoded in that format; they cannot readily adopt non-compatible applications that better meet their business needs. Some file formats can be accessed by only one application, altogether eliminating choice.
- *Inability to exchange data with others.* Users wishing to convey data encoded in proprietary formats must verify that the person or organization receiving the data has the software needed in order to access it or convert the data to a format the recipient can access.
- *Difficulty ensuring the long-term accessibility of data.* Data saved in a proprietary format may remain accessible only as long as the format's owner finds it profitable to support the format or the supporting applications. Even if the format continues to be supported and improved, it might not be possible to access older data. The owner of a format will sometimes update the format and the software needed to make use of it but newer versions of the software do not always allow one to access data saved in older format versions.

Moreover, these realities often have fiscal implications:

- *Cost of ownership.* Individuals and organizations forced to limit their software choices to applications that are compatible with the file formats of their existing data cannot adopt non-compatible applications that cost less to purchase or maintain.
- *Cost of exchanging data.* Users seeking to transmit data to people or organizations lacking software that can access the data in its original format must devote time and effort to converting the data into a format the recipient's software can access.
- *Cost of maintaining accessibility.* People and organizations with an ongoing need to access data saved in a proprietary format must regularly convert their data to newer versions of the format and obtain newer versions of the supporting software. If the file formats of older versions of documents become obsolete, users who have valuable data encoded in the format may be forced to recreate the data or secure the services of a specialized data recovery firm. This cost is also borne by the State for records with permanent value that have been, or will be, transferred to the custody of the State Archives. In order to continue to preserve and make the information in these records accessible, the Archives and the State of New York will have ongoing expenses associated with continual transfer to new versions of the format or recovery of unreadable data.

Open File Formats and Applications

In the 1990s, corporations, governments, and individuals concerned about the cost of using proprietary formats and applications and the risk of losing data as a result of technological obsolescence began advocating use of open formats and software. Open formats and applications are free (or largely free) of legal restrictions on use and modification. Also, all of the technical documentation that enables programmers to develop software that can create and access files encoded in the format is publicly available. Use of open formats and software has several key advantages:

- *Greater choice.* Users of open formats can freely move from one software application to another and, in most instances, can choose from a larger number of compatible software applications than users of proprietary formats. Those unhappy with all of the existing applications can use the format's technical documentation to create their own software.
- *Increased ability to exchange data.* Many open formats can be accessed by popular word processing, spreadsheet, presentation, and graphics software packages, thus reducing the need to convert data into a different format prior to sending it to other people or organizations.
- *Support for long-term preservation.* People and organizations storing data in open formats do not have to be as concerned about migrating their data to newer versions of the format. Many applications that can access data stored in open format are themselves open and freely available from multiple non-commercial sources. Even if the applications that can read older versions of the format disappear, the availability of the technical documentation for the standard enables users to write their own programs.

Use of open file formats and applications may also have fiscal benefits:

- *Cost of ownership.* Given that users of open formats are generally able to choose from a larger number of compatible applications, they may be able to select software packages having lower purchase and maintenance costs than proprietary applications that use proprietary formats. Moreover, some open applications are available without charge to anyone who wishes to install and maintain them.
- *Cost of exchanging data.* In some instances people and organizations seeking to share their data with others may not be forced to allocate resources to converting their data into formats the recipients can access.
- *Cost of maintaining accessibility.* In many instances, users who store their data in open formats can devote fewer resources to migrating their data to newer format versions. If they do need to convert their data to a different format, they can consult the technical documentation for the current file format and develop strategies for identifying and correcting the errors that often occur when during such migrations. Moreover, the ready availability of open applications that can access data saved in many open formats reduces the possibility data will have to be recreated or rescued by a data recovery firm.

However, in some instances, there may be practical and fiscal disadvantages to using open formats and applications:

- *Lack of technical support.* Most popular open source applications have active online user forums. But users who do not have the financial or technological resources to solve complex technical problems or to customize open-source software to meet their particular needs often require more assistance than a user forum can provide. However, users seeking to create files in open formats can often use proprietary applications that create files in the desired formats and include specified levels of technical support in the licensing agreement.

- *Security vulnerabilities*. Some industry observers assert that making the full technical documentation for a given application freely available makes it easier for malicious hackers to find and exploit vulnerabilities. However, others argue that allowing many people to view the full documentation increases the chance security problems will be identified and fixed quickly.[3]

- *Small market share*. In many fields, open formats and applications are used by a small number of individuals and organizations. As a result, users of open formats and applications may need to convert data to other formats prior to disseminating it to others.

- *Cost, unavailability, or insufficient other functionality*. For highly specialized software applications, oftentimes there simply is no alternative solution based on open formats and applications. Also, the costs of solutions (taking into account all costs including costs of implementation *plus* total cost of ownership) may possibly be higher for some open formats and applications. Also, oftentimes an entire infrastructure of ancillary features and applications has been built up around proprietary solutions which can be difficult to duplicate with existing open formats and applications.

Productivity Suites, ODF, and OOXML

Many software companies produce office suites (also known as "productivity suites") which allow organizations and individuals to create word processing documents, spreadsheets, and slideshow-style presentations. Some of these applications also allow users to create databases, graphics, or other types of files. Determining the market share of a given suite cannot be done with precision. Some computer manufacturers bundle productivity suites and other software with the hardware that they sell, and some individuals use more than one suite. However, industry observers agree that at present Microsoft's Office suite accounts for ninety to ninety-five percent of the productivity suite market.[4]

Recently, the development of the office suite format known as Open Document Format or "ODF" has attracted great interest. ODF was created by the Organization for the Advancement of Structured Information Standards (OASIS), a non-profit group of software companies, industry groups, universities, governments, and end users.[5] It is based upon Extensible Markup Language (XML), an open standard developed to facilitate the exchange of structured data across different information systems. ODF itself is also open: OASIS has released the format's full technical documentation, and the format became an ISO standard (ISO/IEC 26300) in September 2006.[6]

ODF's creators and supporters anticipated that the format would spur the development of multiple ODF-compatible applications that would allow end users to:

- *Ensure their files remain accessible over time*. Although the format and its supporting applications were expected to evolve over time, the ongoing availability of technical documentation for all versions of the format would enable future programmers to create new applications that could access old ODF files.

- *Choose the software package that best meets their needs*. Some ODF-compatible applications (e.g., OpenOffice) are available free of charge but have relatively small feature sets and minimal technical support. Others are available for a

fee (e.g., Star Office) but offer more features or full-fledged technical support.

- *Benefit from increased innovation in software development.* Because of the open nature of the ODF format, software companies that charged for use of their ODF-compatible software would remain in business by offering features and technical support users desired; they would not be able to build or maintain market share by controlling the ODF file format.[7]

ODF has also attracted criticism on the grounds that it:

- *Does not include features desired by users.* At present the ODF specification approved by ISO does not allow users to place tables in slideshow-style presentations and makes no mention of digital signatures.[8] Although these deficiencies are scheduled to be addressed in the next version of the ODF standard, they might pose a significant short-term problem for some users.

- *May have technical shortcomings.* Some mathemeticians and developers who have evaluated the ODF standard have found fault with its handling of mathematical formulas, Java applets, and macro/scripting capabilities.[9]

- *Is not fully accessible to people with visual disabilities.* Some of the text-reading applications most commonly used by people with visual disabilities are not fully compatible with ODF. [10]

- *May complicate data exchange.* ODF users seeking to share files with people who use applications that are not ODF-compatible must convert their files to a compatible format prior to doing so. They also have to convert files they receive from users of these applications.

- *May be subject to patent restrictions.* Sun Microsystems was actively involved in the development of ODF and holds the patent for the ODF standard. This allows Sun to enforce its U.S. and foreign rights to the standard in the event it is not actively involved in development of future versions of the standard. As a result, Sun could choose to cease its involvement with ODF and to prevent others from refining and expanding the standard.[11]

Responding to the development of ODF and to its customers' desire to use XML to exchange and reuse data, Microsoft has incorporated XML functionality into its Office productivity suite products:

- In 2003, it released the Microsoft Office XML Reference Schemas, which allowed users of Office 2003 to save word processing documents and spreadsheets in XML.

- In 2005-06, it developed Office Open XML (OOXML). [12] This format, which is the default format for files created using Microsoft Office 2007, was developed to replace older versions of the Microsoft Office formats (Word, Excel, and PowerPoint) and to make Microsoft Office files accessible across a wide array of software and hardware platforms.[13]

Microsoft addressed observers' concerns about interoperability by licensing the XML Reference Schemas on a royalty-free basis, thus enabling any interested third party to

access the schemas' complete technical documentation.[14] After critics charged that the terms of the license emphasized that Microsoft retained certain patent rights governing third-party usage and modification of the schemas, the firm responded by formally promising not to assert its patent rights over the Office Reference Schemas.[15] As of December 2007, its promise extended to anyone "making, using, selling, offering for sale, importing or distributing any" application that made use of the schemas or the OOXML format.[16] However, like ODF, OOXML's patent promise has been criticized as potentially retaining the possible subjection of its users to patent restrictions.

Microsoft has also sought to have OOXML formally recognized as an open standard. In 2005, it submitted OOXML's technical documentation to Ecma International, an ISO member organization that establishes standards for the information technology and telecommunications industries. In December 2006, Ecma International formally approved OOXML as an Ecma standard and agreed to assume responsibility for maintaining its technical documentation and facilitating its further development.[17] In March 2007, ISO agreed to expedite its review of OOXML in accordance with its "fast-track" review procedure for existing technology standards.

During the ISO review process, in response to critics who questioned its decision to develop a new XML-based format instead of making the Office suite ODF-compatible, the firm asserted that OOXML would enable users to:

- *Continue accessing files created with older versions of Microsoft Office.* Unlike the ODF standard, OOXML was asserted to be fully compatible with older versions of the Office formats.

- *Choose the file formats and software applications that best met their needs.* Noting that ISO had approved multiple standards for digital image formats and textual documents, Microsoft asserted that the ODF format and ODF-compatible applications lacked Microsoft Office's full range of features and the development of multiple standards for productivity suite files would enable users to select from a wider array of formats and applications.[18]

Ecma International also requested that ISO, which has developed a "fast-track" review process for established technology standards, expedite the balloting process for OOXML. In March 2007, ISO agreed to do so.[19] The following months proved extremely contentious:

- Opponents alleged Microsoft alone would have the right to develop future versions of OOXML, the format might not be fully compatible with application and operating system software developed by others and that Microsoft was actually seeking to supplant ODF, an existing ISO standard.[20]

- Supporters and opponents accused one another of improperly seeking to influence the outcome of the balloting process. In several instances, opponents were able to produce e-mails indicating Microsoft had encouraged its business partners to join ISO member organizations and vote in favor of making OOXML an ISO standard.[21]

In September 2007 ISO's members voted against adopting Office Open XML as an international standard. A subcommittee charged with determining how to revise the OOXML technical documentation so that it met with members' approval convened in February 2008. At the end of the following

month, ISO voted to make OOXML an international standard for office suite file formats. Provided none of ISO's member organizations file a formal appeal of the vote within the two months following the balloting, the technical documentation needed to develop fully OOXML-compatible software will be published as an ISO standard (ISO/IEC 29500).[22]

Challenges of Managing Electronic Records

As a result of its heavy use of information technology to support diverse business functions, of the need to exchange data with citizens and a wide array of external organizations, and the obligation to ensure the long-term accessibility of some data, government watched the development of ODF and OOXML with particular interest. Governments that create and hold electronic records face a number of preservation challenges:

- *Electronic records are inherently unstable.* Technological obsolescence launches a constant four-pronged attack against the possibility of preservation: hardware, software, file formats and media formats are all subject to change and technology advances so quickly and so often that change is a constant. Government organizations need to keep their hardware and software current – upgrading to new versions regularly -- and they must ensure electronic records remain usable in each new environment.

- *Maintaining electronic records over time is not a high priority for government entities.* State agencies and local governments have successfully maintained usable electronic data over time for information vital to their operations. However, there is no evidence they are doing this for all valuable but less critically important records or for permanent records that will eventually be transferred to the State Archives. When government entities have no compelling reason to expend resources to preserve electronic records, it is likely the records will be neglected and rendered unusable.

- *Cost effective techniques for preserving electronic records in useable formats are not yet available.* Archivists and records management practitioners as a whole have not developed a standard solution for the preservation of electronic records.

- *Permanent electronic media does not exist.* Many organizations incorrectly believe that a CD or a computer tape will always last until the end of its greatest possible life expectancy. That is not the case. Disc technologies are often very susceptible to environmental conditions (light, humidity, air-borne particulates, and even human fingerprints) and may have a much shorter expected life span under these less than optimal conditions. Even if stored under optimal conditions, the media will likely be obsolete long before the life expectancy of the information it holds has expired.

B. Specific Issues Raised In The Request For Public Comments

The Study Approach

The decision to use office suite formats to illuminate issues relating to access to and control of State electronic records was supported by commenters in two ways:

- Many commenters stated that this approach was "reasonable" and "a good start"; and

- Most of the commenters emphasized that office suite file formats were of particular concern to them.

However, some commenters believed a better starting point would be an overall State vision and definition of its business needs for electronic data management. Other commenters qualified their support for the study's focus by noting that while it was a good starting point, office suite formats are also "the tip of the iceberg" consisting of parallel situations elsewhere in the State's IT infrastructure.

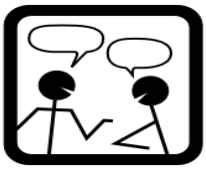
One commenter went further, identifying embedding enterprise Records and Information Management (RIM) into the NYS enterprise architecture and populating a Records and Information Architecture Analytical Framework as key initial steps. The commenter noted that the problem is not limited to the preservation or interoperability of office documents issues alone and lasting solutions to format and digital preservation issues would reveal themselves best when considered within an overall enterprise architecture framework. This comment is reasonable. New York State does not have a formal entity that unites the technology, policy and records administration concerns that this study and commenters highlighted. A foundation recommendation of this report is to establish such a formal entity.

Recommendations

- **Recommendation:** When developing future Statewide strategic plans, CIO Council Strategic Planning Workshops should directly address planning for open standards and including open formats wherever feasible within those plans.

- **Recommendation:** The State should create an ongoing cross-agency electronic records committee as other states have done. A proposed model for the committee is specifically described in this report. Another example of a committee structure that could work here is the New York State Council for Universal Broadband. Virtually all of the reasoning concerning increasing public access which was cited in support of the development of the New York State Council for Universal Broadband could apply equally to universal e-record formats and standards, which also would increase public access. This recommendation for the creation of an ERC is a foundation for the other recommendations contained in this report describing an action agenda for the State going forward.

Definitions: Data, Documents and Records



An "electronic record" is defined in ESRA as "information, evidencing any act, transaction, occurrence, event, or other activity, produced or stored by electronic means and capable of being accurately reproduced in forms perceptible by human sensory capabilities." This definition is consistent with the definition of "records" in the laws governing the admissibility of records in legal proceedings (including Civil Practice Law and Rules sec. 4518), the retention and disposition of government records (Arts and Cultural Affairs Law Art. 57, sections 57.05 and 57.17), and the Freedom of Information Law (Public Officers Law Art. 6, sec. 86).

The RFPC examined the statutory definitions of terms such as electronic "records," "data," and "documents." Many commenters found the RFPC's suggested definitions useful, although some of the more technical public comments suggested using definitions that are more precise. These varied perspectives likely result from the mixture of technical and business questions inherent in any study of this nature. As one commenter noted, the respective definitions can become so precise *"It is probably just as useful to lump everything under the heading of electronic information and let it go at that."* This report has adopted that approach, albeit while recognizing the primary distinction drawn by this commenter: namely, that government's electronic records are an abstraction which have very precise technical mechanisms underlying them.

To the extent the State creates an Electronic Records Committee as recommended in this report, some of the finer distinctions suggested by commenters will be very helpful as that Committee works to recommend more system specific analyses and changes.

Definitions: Access

Consistent with the statute authorizing this study and report, the RFPC also focused on questions concerning access to government records. The commenters generally agreed with this focus.

Recognizing electronic records meet differing needs during a life-cycle over time, the RFPC had suggested conceptualizing records "accessibility" as divided among (a) accessibility for day-to-day utility versus (b) the records' accessibility during their active business use for ancillary purposes (e.g. pursuant to Freedom of Information Law (FOIL) requests, or pursuant to discovery requests in litigation ("e-discovery")) versus (c) accessibility for their historical and research value after having been preserved as official State agency records.

Commenters generally agreed with a trifurcated approach, although there was disagreement as to how to break down these categories. One consistent theme was to maximize accessibility by maximizing openness in all of a record's manifestations. As one commenter flatly stated, "*I define a format's 'accessibility' to include openness -- namely, the format must be based on open standards, and be guaranteed to stay that way in the future. This means that those standards are completely documented and specified, and available to anyone, and will remain so.*" The accessibility of government records is both a public policy and legal obligation of New York State. Lack of format openness inhibits accessibility. Therefore, format openness is one of the desired features and functional requirements

which should be measured for technology used by New York State.

Another commenter noted that while within an organization it is possible to control the use of applications and standards, as soon as data needs to be shared with any external parties, the State has no control over what applications and operating systems the external party uses. As such, "vendor, platform and application neutral standards with broad application and platform support should be chosen for external collaboration." Other commenters noted that full attention by the State to accessibility and preservation needs is less likely to be afforded if differing standards are used which require conversion. Again, many commenters suggested finer distinctions which will be useful to examine in more system-specific analyses. All of these recommendations will be useful for an electronic records committee to explore in more detail in specific contexts. One starting point might be for the ERC to examine exactly what needs particular requestors have for access.

Finally, members of the workgroup noted the State already has policies in place concerning public access to documents made available through the Internet, and while these generally pertain to accommodating assistive technology needs they could be refined to include access to electronic records created and saved in a variety of file formats.[23]

Recommendations

- **Recommendation**: The ERC should explore e-records access issues in relation to the specific needs of specific requestors and the mechanisms by which such records are best made available to requestors. For example:
    - What types of access to records are needed for day-to-day State staff using e-records operationally? What functionality is needed? What functionality is required by law?
    - What types of access would an interested member of the public looking for records need?
    - What types of access would a requestor need in the context of litigation?
    - What types of access would a vendor need?
    - What types of access would a historical researcher need?
    - What types of access would an auditor need?
    - What types of access would an archivist need?
- **Recommendation**: The ERC should also determine what steps would be required for State government agencies to begin accepting office suite documents in open formats such as ODF, as several other state and local government agencies already do in the United States.
- **Recommendation**: CIO/OFT should modify its "Best Practice Guideline G06-001: Accessibility of State Agency Webbased Intranet and Internet Information and Applications" at §14.1 to require that when a State agency posts editable documents to the Internet, open formatted versions of the documents simultaneously be posted.

Definitions: Interoperability

The memo in support of the bill which required this study also focused on new conditions and requirements for State government and the increased usage of electronic documents. The memo expressed serious concerns about the interoperability and preservation of those e-documents.

In its 2004 New York State CIO Council Technology Committee's *"Principles Governing The New York State Information Technology Enterprise Architecture,"* CIO/OFT adopted Gartner's definition of an information technology Enterprise Architecture as a framework that includes helping to *"ensure that IT systems are flexible enough to adapt quickly to new business conditions and requirements."*[24]

So to the degree to which document format choices led to those serious concerns about the interoperability and preservation of e-documents, what are the "new business conditions and requirements" facing IT in terms of formats? They consist at the very least of the following items needing to be addressed by State government IT leaders:

- e-Discovery requirements;
- intellectual property issues;
- availability of multiple formats for doing similar things;
- the increasing feature completeness and capabilities of software applications which use open formats;
- the increasing adoption of open formats by individuals and entities with whom the State does business;
- recognition by courts in the FOIL context that providing the maximum access contemplated by FOIL statutes includes providing access to records digitally;
- the increased movement of more and more State records to e-records instead of paper;
- the increased movement towards XML-based office documents;
- the increased dynamicism of documents rendering more important the need to preserve original functionality; and
- always increasing needs for fiscal restraint and cost savings.

The choice of e-records formats rises to an Enterprise level question in that it affects more than one Enterprise entity (e.g. more than one state agency) or external entity (e.g. multiple county agencies). Its essence is facilitation of the ability to communicate information between disparate entities and individuals.

As stated within the State's Enterprise Architecture document:

> *"There is a compelling case that can be made for setting and enforcing standards. Some operational functions in state agencies are ubiquitous. Understandably, these functions would benefit from having a similar feel and look to promote information sharing, increase usability and reduce IT maintenance costs.*
>
> *Without careful consideration of standardization, the otherwise positive moves toward open architecture and web based e-government solutions will create problems and excessive costs. Without standards, common civil service titles, combined with current Principles Governing the New York State Information Technology Enterprise Architecture hiring practices, promotions and transfers across agencies will result in increased training costs and lower efficiency levels.*

> *Our common goal must be - and will be - to further [an] agenda to improve government service to New York State's citizens through technology. Standards will help us achieve greater efficiencies internally which can then be passed directly on to the taxpayers."*[25]

The e-records study statute required CIO/OFT to make recommendations concerning interoperability. The RFPC published by CIO/OFT recommended a definition of "interoperable" as *"products and systems from multiple vendors that can be used together without modification or development of custom interfaces and tools."*[26]

Many commenters found this definition "useful" or "well-stated." But other commenters offered a variety of alternative definitions with no real commonality among them, focusing on disparate aspects which each commenter valued, such as who controls the format, whether there are costs associated with its use, or the degree to which interoperability is effective without conversion errors. The fact that there were such differences of opinion bolsters the observation of another commenter. That commenter averred that interoperability should be viewed as a continuum among business processes and the technology, applications, data, information and records that support them. The commenter further stated that the term's definition "should be applied to specific business process interoperability objectives and the desired level of the interoperability specified."

Individual RFPC commenters have overwhelmingly voiced support for more openness in formats. While a few comments reflected embittered characterizations, most of these comments focused on practical considerations. These commenters did not seem to buy into the competition between companies and expressed little concern with the State remaining with one commercial solution or another. Rather, their primary concern was the openness of the formats used in documents available to them, irrespective of the software applications used.

Given users' simultaneous demands for more openness in office suite software along with all of the functionality that has built up around the currently predominant office suite, an ideal short-term solution might be for vendors of popular, feature-rich office suite software to directly support ODF in their products. This specific request has internationally been made repeatedly, most recently by the U.K. national government IT and education agency BECTA.[27]

Why proprietary vendors do not do this is unclear. The commonly stated reason is that ODF lacks certain features desired by their customers,[28] but that could equally be said of other formats which are directly supported within their products. The most popular proprietary office suite software solutions offer so many features, both direct and ancillary, that adding direct support for ODF as an additional feature would presumably enhance the value of such products and make those solutions even more compelling, from an end user's perspective, particularly in a complex environment such as the State's.

Supporting ODF in popular proprietary products also will address the needs of the State's constituents. Overwhelmingly, RFPC commenters asked for more document openness and fewer application and cost barriers when they access State records. Choice of applications using open formats was a resounding refrain among individual commenters to the RFPC. These users' needs for openness could be seen as magnified because, while complex entities such as the State may need fuller functionality besides just format openness, for many of the State's end users the functionality accompanying openness offered by low cost or even free

software applications is enough. Thus, with fewer functional needs the State's citizens may very well be more ready than the State to use a variety of office suite software applications.

The State has undertaken several initiatives to end the digital divide such as its broadband program. Efforts to remove barriers to citizen access to government and its records should include removing the barrier of closed formats.

Recommendations

- **Recommendation:** Where it is possible to do so, the State should incorporate open formats into its e-records systems as a means to enhance interoperability and to remove barriers to citizen access to government and its records.

Definitions: Openness and Open Standards

The e-records study statute used very broad terminology setting forth features favorable to publicly accessible electronic records. As the RFPC noted drawing from the terms used in the statute, electronic records should be:

- creatable;
- maintainable;
- exchangeable;
- interoperable;
- accessible;
- readable;
- preservable;
- storable;
- appropriately controllable;
- end-user technology choice capable;
- vendor neutral; and
- cost effectively implementable

There are a multitude of definitions of "open standards" and the sub-set of open standards that are "open formats.[29] As the e-record format discussion continues, the most likely flashpoint in the political realm and elsewhere will be debates about the degree of openness of particular formats, and the minimum set of qualifications which a format will need to possess to be considered "open." This is already where some of the most contentious arguments have been raised in the market.

The statute requiring this report stated certain values and goals which CIO/OFT was asked to study, as listed above. An effective manner for the State to support these values and goals would be to define open standards and formats using the broadest, most State-protective definitions and for the State's IT purchases to aim towards achieving as close to that definition as possible.[30] So long as State agencies retain the flexibility to choose other options should an evaluation of best value demonstrate the weight of competing desired functionality, then broad definitions most strongly benefit end-users such as the State and its citizens. In the RFPC, CIO/OFT proposed to adopt the type of broad definition used by the South African Department of Technology. The comments in response to the RFPC reinforced the wisdom of doing so.[31]

The South African government defines formats as "open" if:

- they are maintained by a non-commercial organization;

- participation in the ongoing development work is based on decision-making processes that are open to all interested parties;
- anyone may access committee documents, drafts and completed standards free of cost or for a negligible fee;
- anyone may copy, distribute and use the standard free of cost;
- the intellectual rights required to implement the standard (e.g. essential patent claims) are irrevocably available, without any royalties attached;
- there are no reservations regarding reuse of the standard; and
- there are multiple implementations of the standard.

Commenters offered additional refinements, including that:

- software using the standard should not extend the standard so that the format in which the documents are saved are not actually in the standardized format;
- there should be full and independent implementations on multiple platforms;
- its development and management process must be collaborative and democratic:
    - participation must be accessible to all those who wish to participate and can meet fair and reasonable criteria imposed by the organization under which it is developed and managed;
    - the processes must be documented and, through a known method, can be changed through input from all participants;
    - the process must be based on formal and binding commitments for the disclosure and licensing of intellectual property rights;
    - development and management should strive for consensus, and an appeals process must be clearly outlined; and
    - the standard specification must be open to extensive public review at least once in its lifecycle, with comments duly discussed and acted upon, if required.
- the standard must describe an interface, not an implementation, and the industry must be capable of creating multiple, competing implementations to the interface described in the standard without undue or restrictive constraints. Interfaces include APIs, protocols, schemas, data formats and their encoding;
- the standard must not contain any proprietary "hooks" creating technical or economic barriers;
- faithful implementations of the standard must interoperate. Interoperability includes the ability to use, convert, or exchange file formats, protocols, schemas, interface information or conventions, so as to permit the computer program to work with other computer programs and users in all the ways in which they are intended to function;
- it must be permissible for anyone to copy, distribute and read the standard for a nominal fee, or even no fee. If there is a fee, it must be low enough to not preclude widespread use;
- it must be possible for anyone to obtain free (no royalties or fees; also known as "royalty free"), worldwide,

non-exclusive and perpetual licenses to all essential patent claims to make, use and sell products based on the standard. The only exceptions are terminations per the reciprocity and defensive suspension terms outlined below. Essential patent claims include pending, unpublished patents, published patents, and patent applications. The license is only for the exact scope of the standard in question:

- may be conditioned only on reciprocal licenses to any of licensees' patent claims essential to practice that standard (also known as a reciprocity clause);
- may be terminated as to any licensee who sues the licensor or any other licensee for infringement of patent claims essential to practice that standard (also known as a "defensive suspension" clause); and
- the same licensing terms are available to every potential licensor; and

- the licensing terms of an open standard must not preclude implementations of that standard under open source licensing terms or restricted licensing terms.[32]

Commenters were generally supportive of the State using a broad definition of open standards and formats, offering such comments as "this is a very good definition" and that it was "the best definition of the many I've seen." As noted a few commenters suggested refinements, and one noted the definition should not be used to mandate open standards usage where such usage needed to yield to practical realities.

The State Legislature could undertake several measures which would be helpful in ensuring openness is integrated into the State's use of technology. For example, France is one jurisdiction where not only its executive branch of government, but also its legislature, is mandating use of ODF. Because the State Legislature likely encounters similar issues which led to the French legislature's decision, the NYS Legislature should study France's implementation and consider the advisability for the NYS Legislature itself adopting open formats.

Also, it is understood that the State Assembly already uses the open Mozilla Firefox web browser for its own uses. When entities migrate to different IT solutions, this often illuminates the absence of features which had been built up around the previously used technology.[33] If staff from the State legislature participate in the recommended State's Electronic Records Committee, then they may be able to bring valuable insight into why and how the Legislature effected its own migration.

In response to questions concerning the definition of "openness," one commenter noted:

> *"We also urge OFT to refrain from making policy or technology recommendations based on preconceptions of what is more (or "fully") or less (or not) open. In the debates of the day, the word "open" is used widely but means different things to different people. For instance, combining our answer on IP questions (number 48) above with the Burton Group's observation that Sun's IP position could stymie ODF (see question 7), one could conclude that ODF's "full" openness is suspect. ODF advocates of course would vehemently disagree. Given that the marketplace is responding to customer demands, it*

would seem unnecessary and unwise for OFT to wade into this swamp."

Most of this comment appears to be a fair statement. Anyone -- any vendor, any format promoter -- proffering a solution that is denominated as "open" needs to demonstrate to the State's satisfaction that the proposed solution truly meets the definition of "openness" that best meets the State's needs.

The proponents of each of the competing new XML-based office suite formats presumably must consider the concept of "openness" important. Otherwise, they would not each have prefaced their standards with that term. After all, the two standards are referred to by their proponents as Office *OPEN* XML and *OPEN* Document Format. (emphases added).

The statement that the marketplace is responding to customer demands is less supportable. As a large customer of office suite products, State government's open standards and open format needs for those products are still largely unmet.

Defining openness correctly, in a manner which benefits the State to the greatest extent, does not in and of itself mandate any disruptive, immediate, rip and replace implementation of products adhering to that definition. Instead, it makes plain the State's intention to render non-openness as a legacy to be responsibly migrated away from.

Open formats are published specifications for storing digital data, "usually maintained by a non-proprietary standards organization, and free of legal restrictions on use."[34]

For electronic records, New York State's day-to-day usage business need is for the maximum combination of features and open standards, including open formats. This openness has the potential to reduce the State's costs, increase the State's choice of software vendors, and further the State's responsiveness to ancillary needs, such as accessibility of its records for responding to reasonable FOIL or e-discovery requests.

Some State agencies such as State Archives have explicitly expressed a user need for this additional functionality of openness. This is because State Archives perceives maximum openness of formats as most easily allowing the conversion of textual records into a pure ASCII or Unicode format that will be relatively simple to preserve over time. Thus, in addition to all of the typical office suite functionality which it has historically sought it makes sense for the State to include a preference for the greatest possible degree of openness in the formats which it procures.

Recommendations

- **Recommendation:** The State should build upon work by other governmental jurisdictions and create a New York State Interoperability Architecture. There are many existing examples by other governments which the State could build upon to develop its own Interoperability Architecture which can be applied to specific State business objectives.

- **Recommendation:** CIO/OFT should modify its Enterprise Architecture Principles to add a new Principle #33 concerning Open Formats to follow the existing Principle #32 concerning Open Standards.

- **Recommendation:** The State should join the British Educational Communications and Technology Agency (BECTA), the pan-European e-government services committee (PEGSCO) and many others in expressing its broad-based request for all of its office suite software suppliers to fully support both ODF and OOXML.

Trends: Predicted Use of Open Formats

As noted in the 2004 ESRA report, "most State agencies accept[ed] electronic records from citizens, businesses and other government entities." At that time the percentage of State agencies accepting electronic records, including the New York State Unified Court System, was seventy-one (71%) percent. Presumably this has grown, and can only be expected to grow further. If citizens, businesses and other government entities increasingly wish to submit e-records to the State in open formats, the State will need a strategy to be able to work with those documents. This was an issue during the preparation of this report. Several RFPC commenters initially submitted their comments in Open Document Format. The workgroup regrettably needed to ask these commenters to choose another format for their submissions that could be opened by workgroup members' currently approved office suite applications.

Commenters generally agreed the use of ODF is growing substantially and Gartner's prediction referred to in the RFPC concerning enormous future usage by governments was probably correct.[35]

It was suggested by some vendor commenters that in general customers were not requesting direct support of ODF in productivity suite products.[36] To some extent this may be true. For example, consumers who have moved on to software applications which support open formats may be using those software applications exclusively and no longer using proprietary products. However, the individuals who have made those software choices who reside in New York State are the State government's "customers." And the State has been hearing about this issue loudly and clearly from our customers, with a distinct and growing

demand for open document formats. If all that proprietary vendors are waiting for before they directly support ODF is a "broad based customer request" then they should be aware that such a demand already exists in New York State.

As such, New York State may be added to the list of those who are asking proprietary vendors to directly support document formats such as ODF or other formats that maintain the same level of ODF's openness. It would be timely for the National Association of Chief Information Officers (NASCIO) to address this demand on behalf of all of the United State's state governments. It would certainly serve the interests of their "customers" -- their citizens -- if state governments were to amplify their voices on the topic of openness in electronic record formats.

In a recent and prominent report, the Burton Group, a noted technical research and consulting services organization, stated that it believes that while governments will be satisfied with applications using the ODF format, the OOXML document format will lead the larger market for three reasons:

> *"First, many enterprises are not that caught up in the standards debate; they just want to use what works for their needs. Microsoft Office 2007 defaults to storing documents in OOXML format, so, by migrating to Office 2007, many companies will let Microsoft make the decision for them. Second, OOXML is an extensible standard. It allows vendors and enterprises to extend the standard within an OOXML-defined framework. For example, the .XLSM file format, used to support a Microsoft Office 2007 Excel macro-enabled workbook, is not part of the base OOXML standard, but rather a Microsoft-created extension. This built-in ability to augment the OOXML standard is a safety valve for future innovation, allowing new features to be added without forcing vendors to invent yet another separate file format or wait for standards bodies to give their approval. While such extensions initially decrease interoperability, it's Burton Group's belief that this issue will resolve itself over time, as popular extensions are adopted by other vendors or eventually move into the baseline specification. Third, OOXML supports "overlay" custom schemas (not in ODF 1.0, promised in ODF 1.2), which can be used as views into the business information stored in documents. This separation of document and views allows enterprises to more easily perform tasks such as programmatically updating a "Stock Price" element or corporate logo within a document, compared to ODF's method of serially inspecting and updating the document itself. In short, because OOXML is more ecosystem- and application-oriented than ODF, most vendors and enterprises will see it as more useful than ODF."*[37]

In response, Sun Microsystems posted a review of the Burton Group's analysis that refuted several "myths" about ODF, including the following:

> *"[Myth:] Customer[s] care about features, not formats*
>
> *That customers increasingly care about formats and actually consider open standards support a key feature and requirement becomes for example evident in the Valoris study conducted by the European Commission in 2003 (http://ec.europa.eu/idabc/en/document/3439#IBM) as well as the workshop about document exchange formats as part of the German EU presidency earlier this year (http://ec.europa.eu/idabc/servlets/Doc?id=27875). At the very least, government agencies around the world consider open standards support to be a key product*

feature just like spell checking or printing."[38]

Both of these analyses appear to be reconcilable and correct. OOXML, which has just been deemed an international standard for office suite files, likely will gain traction in the marketplace, and likely will be favored by some because of its features. However, open standards and open formats support are in fact increasingly being recognized as a *"key product feature just like spell checking or printing."* Optimally, all of the State's desired features will be present in technology solutions which vendors provide to the State.

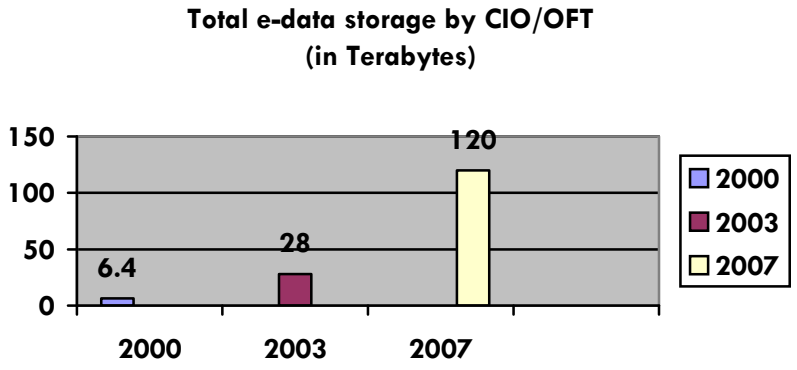
As one commenter's general manager of corporate interoperability and standards group recently stated:

"Interoperability is a necessity [and] is going to be a feature in a product, which customers demand. It is going to be a standard functionality, just like security."[39]

We note that the State of Massachusetts was the first state in the United States to explicitly adopt open formats just as other governmental administrations all over the world have done. It appears that questions about Massachusetts' proposals arose, however, not because of what was done, but because of how it was achieved. As a consequence, the State of New York should address the desired feature of open standards support by following the more measured path identified in the recommendations which follow.

Recommendations

- **Recommendation**: Clear trends describe growing use of document formats which are open, such as the ODF format, in office suite documents. To address these trends, as an IT customer the State should:
    - define open standards and open formats with the definitions meeting the needs of the State;
    - identify open standards and open formats, under those definitions, as a technology feature specifically desired by the State;
    - integrate the acquisition of this feature of openness into the State's technology planning and procurement processes, and in doing so specify openness in procurement requests as one desired feature among many;
    - recognize that for some technology purchases openness may become the distinguishing factor in their acquisition; and
    - ensure collaboration among State agencies to assess the mix of their needs for openness versus other features, and to test and pilot the suitability and cost whenever it appears that technologies have become available where the sufficiency of openness and other feature sets have converged.



Trends: Current State Practices

Between 2000 and 2007, the amount of e-data stored by CIO/OFT for State agencies increased exponentially, by 1,775%.[40] CIO/OFT provides the platforms upon which State agency software applications and data reside. Accordingly, CIO/OFT does not always have particular awareness of what those agencies' materials are.

Only four State agencies provided responsive comments concerning the current adoption of various formats within State agencies. One State agency noted its paper records and documents are scanned through a copier/scanner and then transferred to a Lotus Notes database. Another agency observed merely that it currently holds seventeen (17) gigabytes of digital office documents in office suite software within electronic folders and another sixteen gigabytes of e-records in Lotus/Domino and Websphere applications. As such, the workgroup used different methodologies to develop a basic sense of office suite application usage in sister State agencies.

The reality is that the State has not fully consolidated usage of office suite formats at the State agency level and possibly never will. Just as web browser users vary in their browser preference selections which can cause webpages to render differently (such as disagreeing on whether to allow cookies, or whether or not to use Flash or Javascript), office suite users vary in their office suite application preferences.

It would be incorrect to assert that Microsoft Office is a de facto standard in NYS agencies, as there are pockets of alternatives. Reasons for the use of alternatives vary. As one state agency recently noted concerning an item of proposed legislation, *"the needs of individual agencies for support, reliability, and security vary greatly and their disparate needs may not be met adequately by a consolidated [technology]."*

Wordstar was the most popular word-processing program used on PCs in the 1980s, and still has an active community.[41] WordPerfect surpassed it in popularity in the late 1980s through the early 1990s, and still has active usage. A Jupiter Research survey in 2006 found that "*WordPerfect remains the No. 2 office suite behind Microsoft Office in the consumer, SMB [Server Message Block] and enterprise markets with roughly a 15 percent share in each market."*[42]

Various State agencies still use WordPerfect, including the Office of the Attorney General and the Department of Environmental Conservation. A review of purchasing documents reveals that during the 2007-08 budget year two other agencies purchased,

respectively, Lotus Notes and an EMC ancillary application which integrates with and manages Lotus Notes.

The State Office of General Services (OGS) in 2007 updated a contract through 2008 for Quick Copy / Duplicating and Color Transparencies (All State Agencies - Albany Area) which accommodates the printing of Adobe, Microsoft Word, and Corel WordPerfect formatted documents.[43]

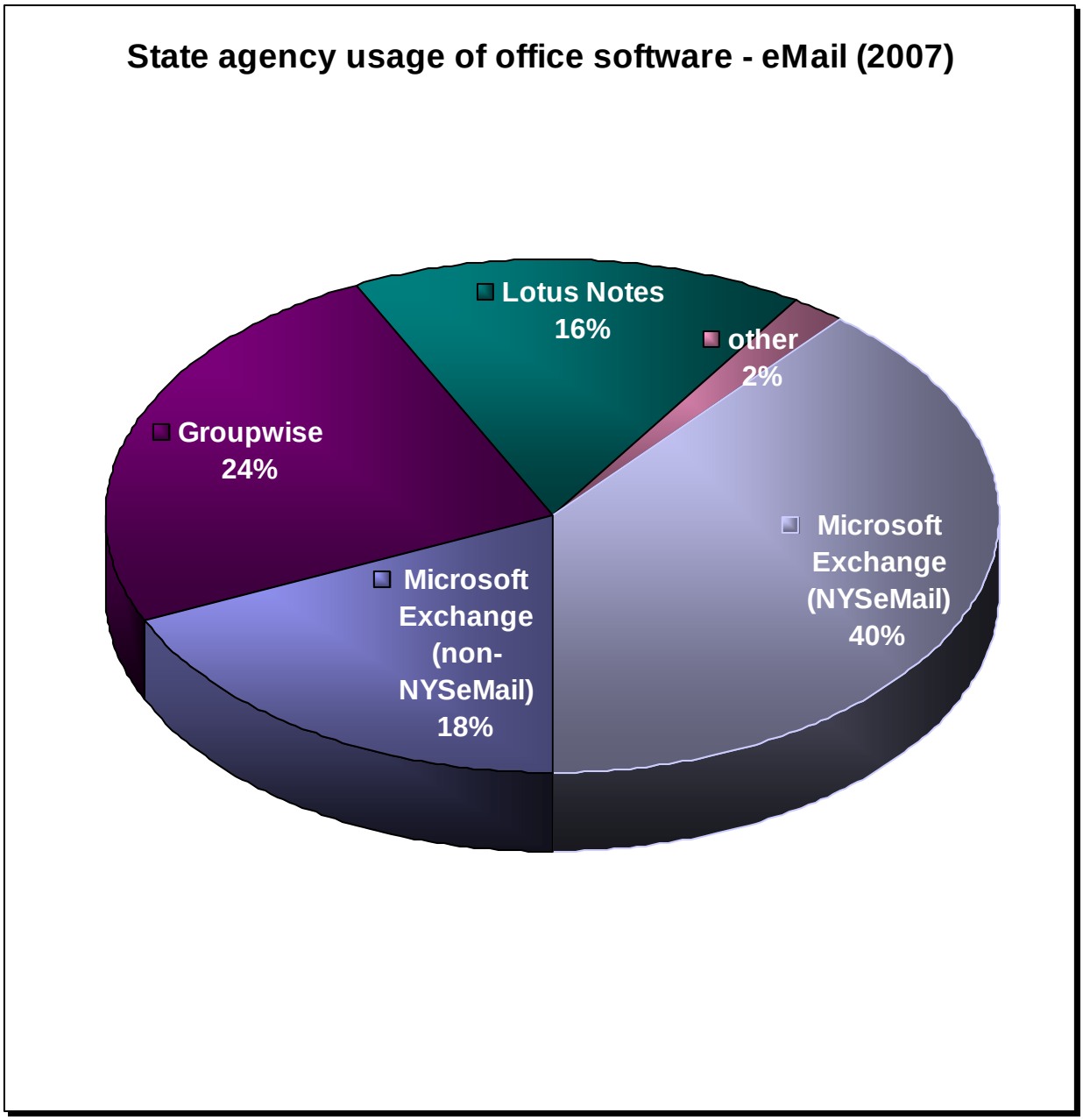
OGS also maintains an "*Assistive Technology for Persons with Disabilities*" contract through 2009, which provides tutorials on the use of Jaws disability software for either Microsoft Word or WordPerfect.[44]

The OGS "Electronic Database Products And Services Service Agreement" with West Group allows downloading documents into formatting for Microsoft Word, Corel WordPerfect, HTML, PDF, and ASCII.[45]

State agencies can buy WordPerfect through the OGS ASAP contract,[46] and OGS's Microsystems Office Automation Systems contract with SUN makes Star Office available to State agencies.[47]

The diversity of e-mail applications used by State agencies also highlights the ways in which agency work practices, security needs, and other concerns sometimes militate against consolidation of technology. (See chart below).

International Data Group's IDC subsidiary studied usage of ODF in the United States in 2007 and found usage at approximately 1% of organizations polled.[48] Using the most recent available New York State census figures, assuming accuracy in the study and assuming New York State's organizational usage is consistent with this national figure, this means that more than five-thousand (5,000) non-farm businesses in New York State and their employees are using ODF formats.[49] This is a significant figure and does not even include the State's households which are using the format. Thus, it is reasonably argued there is already a significant enough degree of ODF usage within New York State to warrant the State trying to accommodate the ODF needs of its citizens.


State agency usage of office software - eMail (2007)
Lotus Notes 16%
other 2%
Groupwise 24%
Microsoft Exchange (NYSeMail) 40%
Microsoft Exchange (non-NYSeMail) 18%

Trends: Other Jurisdictions

The most consistent user need described in the public comments received in response to the RFPC was a demand for State government to publish its documents in ODF format. There are many examples of governmental bodies in the United States already successfully regularly posting documents in multiple formats, including ODF format.[50] Moreover, many such governmental entities are not merely *posting* ODF documents, but also accepting their use for official forms and other submissions to the agency. For example, the Indiana Department of Education accepts application forms for e-Rate funding submitted in only PDF or ODF formats.[51] Similarly, Washington State in a recent broadband solicitation has asked for RFP responses (bids) to be submitted to state government in either Microsoft Word or in ODF format.[52]

The federal Government Printing Office (GPO) currently accepts files from other federal agencies in any format in which the agency produces them. The agency is developing a "Future Digital System" for retaining federal government records. Its procurement requirements have mandated the ability to *"have the capability to deliver OASIS Open Document Format for Office Applications (OpenDocument) v1.0."* [53]

A more detailed analysis of the open format policies of numerous governmental jurisdictions may be found in Appendices III through V within this section of the report. If this report's recommendation for the creation of an Electronic Records Committee (ERC) to address implementing similar requirements in New York State is implemented, then the ERC will have a wealth of sample approaches which it may evaluate.

C. Some Approaches That Were Suggested For The State

Approach: Retaining the Status Quo by Using Converters

The RFPC did not ask detailed questions about document converters, or "translators." However, CIO/OFT received public comments from certain commercial entities stressing two points:

> (a) that the State should make no changes to its existing procurement practices; and
>
> (b) that to the extent the State needs to preserve documents in ODF format, it can use translators to accomplish this.

Many commenters, including State Archives, observed that for various reasons electronic records are best preserved in the same form in which they were originally created.

State Archives also stressed the importance of using open formats. In many instances records created in open formats can be preserved in their native format for a lengthy period of time. If and when the records need to be converted to a different format, archivists can reduce the incidence of conversion-related errors by studying the technical documentation of the native format and developing protocols minimizing the impact of conversion upon the records' appearance, behavior, and informational content.[54]

Archivists recognize that in many instances open formats simply do not offer the functionality that records creators require. However whenever possible, archivists will convert records in proprietary formats to open formats. This approach is driven by two concerns:

a) *The need to balance essential characteristics and accessibility.* Although converting records from one format to another may result in the loss of important characteristics (e.g., appearance, functionality), many records will not remain accessible over time unless they are converted to a different format. If converting the files to an open format would result in unacceptable losses, archivists will take one of two actions until a better preservation strategy eventually emerges.

They will either migrate files to a newer version of their native format or preserve the files in their native format if migration is not an option, as is the case with some CAD and other specialized systems.

b) *The desire to minimize the number of conversions performed upon a given file.* Each conversion increases the risk the record's essential characteristics will be altered or lost. In most instances, records encoded in open formats require fewer conversions than records encoded in proprietary formats. Another reason for preserving electronic documents in their original formats is that document converters do not work very well. For example, Microsoft working together with Novell has created an ODF Add-in for Word, Excel, and PowerPoint, but these document converters are far from seamless to use. Microsoft's ODF Add-in for Word itself in dialog boxes warns about the types of functionality which can be lost through using the translator. (See graphic below.)

> These include header dimensions, table properties, table alignment and header repetition, image cropping, text box positioning, and top and bottom wrapping. [55]

These types of functions which the warning says are subject to alteration during conversion are commonly used in creating electronic documents. In fact, *each of these features was used in the creation of this report.*

Document converters have been widely disparaged over the years. The existing OOXML-ODF translators are no exception. In fact, despite recommending the use of translators as a complete solution, in its response to the RFPC Microsoft has recently itself acknowledged translators are an imperfect solution.

> *"With regard to whether it is possible to create a perfect translator between different formats, Robertson said there is no such thing. The community needs to be involved in the process and translation could be optimized and continually improved, but the fact remains that the two formats [OOXML and ODF] are different, he said."*[56]

Microsoft's partner in creation of that translator, Novell, also has acknowledged the recommended translator does not work optimally:

> *"But clearly, seamless interoperability doesn't occur instantaneously. Novell first shipped a translation tool for ODF and OOXML last year, but it is not yet fully baked. The tool will have 'full supportability across all the components of the Office suite in the first half of 2009,' Dyroff said. In the meantime, the company will continue to ship beta versions, he added. 'It just takes time, he said of the roadmap.'"*[57]

In general, incompatibilities render use of translators in any situation sub-optimal. There is a list of dozens of incompatibilities associated with usage of Microsoft's ODF Add-ins for Word, Excel, and PowerPoint at the Microsoft Add-in software's website.[58]

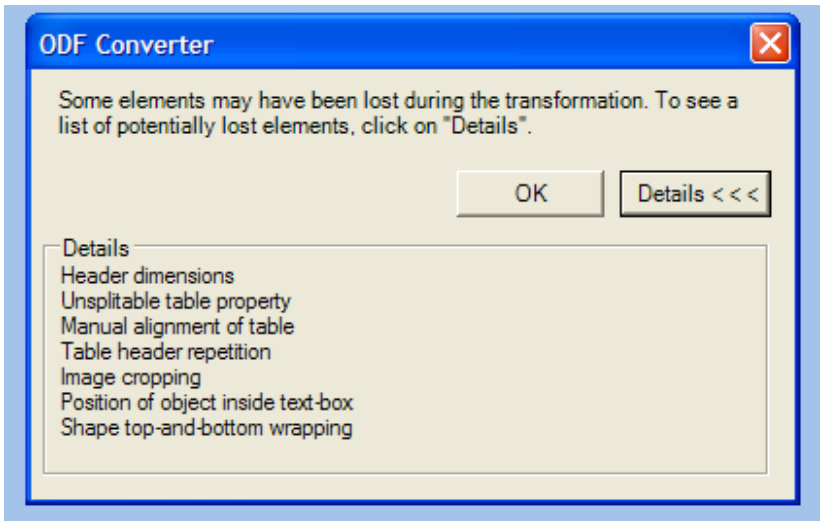
Moreover, requiring use of a translator has been deemed anti-competitive by other governments. As the U.K.'s national government technology and education agency BECTA recently observed:

> *"We identified ten steps that users would need to take in order to locate and install the converter that gives Office 2007 the ability to access ODF files and note that the arrangements for opening and saving ODF files in Microsoft Office 2007 are not intuitive in that they deviate from the normal approach familiar to users. We believe that these arrangements present sufficient technical difficulties for the majority of users to make them disinclined to use competitor products and this may weaken competition."*[59]

Microsoft recently published "Interoperability Principles" designed to "increase the openness of its products and drive greater interoperability."[60] When this pledge for greater future openness is realized perhaps ODF-OOXML converters will perform sufficiently for the State's needs. But as of the date of this report the existing converter options do not appear to be sufficient.

Some commenters have suggested that certain proprietary solutions such as OpenOffice Novell Edition or Sun's StarOffice or Corel WordPerfect 2007 may better be able to achieve acceptable conversions. Under the auspices of the proposed State Electronic Records Committee all existing converters could be vetted professionally in a lab environment with vendor input to determine whether any

might sufficiently meet the State's recommendation for being able to accept and send documents in ODF format as needed.



Recommendations

- **Recommendation:** The State cannot afford to "do nothing" while waiting for converters to improve. Instead, the State should take the types of efforts described in this report to operationalize openness to the extent possible throughout all of its IT systems.
- **Recommendation:** Despite assertions by their proponents that converters may never work properly, the State should remain receptive to their use in the event that they are improved. The proposed Electronic Records Committee should further evaluate converter utility, reaching out to the vendor community to determine definitively their sufficiency.
- **Recommendation:** The State should press for harmonization of OOXML and ODF into one office suite format.

Approach: Provisional Format Approval

The RFPC asked about the possibility of provisionally certifying formats which met the State's own standards (such as the State's needs for various types of functionality including the functionality of openness) with the flexibility to certify additional standards which came into compliance with the State's needs.

RFPC commenters generally disliked this suggestion. One commenter pointed to the complexity and costs that this approach could increase. Another commenter noted that even if the State were to select an initial "preferred format," the State still would have a need to interoperate with others who had chosen their preferred differing formats. One State agency noted that while CIO/OFT can recommend a particular format,

agencies have a need for and should be permitted flexibility.

This request for flexibility appears reasonable. As noted previously the State has a simultaneous need for maximum functionality along with maximum openness, and for many of the State's business purposes there does not appear to be an IT solution which combines both of these. The State also has a need to meet the expectations of particular constituents, such as those served by the State Archives. The Archives must ensure future generations of New Yorkers can access archival electronic records and prefers to receive records created in open formats, and the many businesses and individuals who which prefers truly open formats under recognized archival principles. The State's constituents also include the many businesses and individuals within the State who want to interact with the State using truly open formats. The conundrum then is that for the foreseeable future any State requirement imposing one format or another is likely to cause difficulties for State agencies. Faced with this, the State should move measuredly towards adopting software when it does meet both the State's openness and other functional needs, and work through committees such as the ERC to evaluate interim approaches which further short-term interoperability.

Recommendations

- **Recommendation:** The State should not currently require the use of any particular format within State agencies, either provisionally or as a final recommendation.
- **Recommendation:** The Electronic Records Committee should recommend particular mechanisms by which to meet the document open format requests of State agencies (e.g. State Archives) and the State's citizens.

Approach: Not Raised in the RFPC

The State also asked commenters to provide any alternative approaches which the State may not have considered and described in the RFPC. The comments received varied greatly. One commenter suggested the State should do a better job in describing the exact problem it faces, and then write a business case to address it. For particular use cases this is likely advisable. Several commercial commenters used the opportunity to describe offerings which they could sell to the State to address some of the issues raised. These responses were helpful and should provide food for thought as CIO Council committees and/or the proposed ERC undertake more detailed analyses of specific use cases.

Several State commenters described a need for consistent electronic management systems across all State agencies. State Archives in testimony before the State Legislature also described similar approaches in use elsewhere. While the issues which the RFPC addressed have cross agency implications, and consistency of State-wide approaches is recommended, cost and competing State initiatives will obviously need to be considered for any individual proposal.

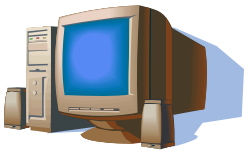
A few commenters recommended the State retain flexibility as formats and IT systems

are still continually evolving. Finally, several commenters advised the State to make a concerted effort to ensure its e-records management practices fit within its records management thinking in general, and seek to leverage its information assets to the greatest degree possible. While the State needs to be cognizant of practical limitations, these are valuable insights.

Recommendations

- **Recommendation:** During their deliberations concerning how to operationalize openness within the State's e-records systems, the proposed ERC working with the appropriate CIO Council action teams should review these suggested alternative approaches.

Approach: Technology Archiving



The RFPC asked about hardware archiving only because there are various commentators within archiving literature who write about long-term preservation and have suggested it. The RFPC commentators roundly criticized the approach. The reasons for rejecting it included that it is not a viable long-term solution given hardware breakage and obsolescence, and suggestions that it would need to be combined with the use of open formats in any event.

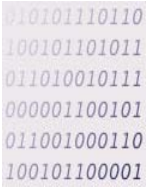
A few commenters mentioned the possibility of using virtualization technology in order to keep outdated software solutions available, but still rejected the idea as non-viable. The workgroup was also aware of perspectives offered within the literature that entities choosing this approach could run into difficulties such as End User License Agreements (EULAs) that prohibit using software on virtualized platforms. Another concern is older proprietary software that was subject to product activation associated with particular hardware. That software might not be capable of being migrated and re-activated because of the shut-down by those proprietary interests of either their companies or of their activation capabilities for that software. Thus, reliance on trying to maintain outdated software could cause the State to lose all access to documents in prior document formats.

For the reasons stated, hardware or software archiving does not appear to be a viable approach to the issues raised by the legislation requiring this study.

Recommendations

- **Recommendation:** Depending on the archiving of older versions of hardware and software would not be a viable preservation strategy for the State, and the State should not pursue it.

Approach: Open Source and Source Code Escrow



Source code escrow means deposit of the source code of the software into an account held by a third party escrow agent. Escrow is typically requested by a party licensing software (the licensee), to ensure maintenance of the software. The software source code is released to the licensee if the licensor files for bankruptcy or otherwise fails to maintain and update the software as promised in the software license agreement.[61]

Concerning both the open source and the source code escrow questions many commenters, particularly those from State government, expressed support for the State acquiring source code when it procures software applications to increase access by the State to its own data. Yet many other commenters expressed skepticism vendors would agree to this.

Source code escrow is important to State Archives. State Archives provides assistance to local governments in managing their records, including grants made through the Local Government Records Management Improvement Fund. One of the requirements for applicants applying for technology improvement grants is that they "include a clause in any software development contract that requires the software code for customized software to be placed in escrow" and requires the vendor "deliver software documentation that meets industry standards."[62]

On both questions the majority of commenters noted that while the suggestions might create other benefits for the State, the question is not pertinent to the format question.

Recommendations

- **Recommendation:** The State should seek to increase the open source and open standards expertise in the State's IT workforce.
- **Recommendation:** The State should include a non-mandatory but desirable requirement that is evaluated in centralized procurement contracts favoring that the source code for the procured software is shared with the State.

Approach: Emergency Preparedness

Failure to expend monies to render data more "readily usable" can lead to unexpected costs in emergency management situations. After the tsunami and Katrina disasters in the past few years where governments were unable to effectively communicate because e-records were held in proprietarily locked, non-interoperable agency-specific silos, it has become clearer that:

(a) any solutions seeking to address e-records preservation and access must take into account emergency preparedness; and

(b) emergency preparedness would best be served by the use of standards and formats sufficiently open so that e-records can readily be accessed by anyone who appropriately has a need to do so.

Three examples of disaster preparedness projects in which State Archives is involved include:

- SEMO/SED partnership for Long Island schools: The State Archives recently collaborated with the State Education Department's Facilities Planning Bureau and the State Emergency Management Office in the development of a new LI School Facility Hurricane Preparedness & Recovery Guide. This is a guidance document developed for the Long Island education community to raise awareness in relation to Long Island's hurricane history; provide suggested preparedness and mitigation measures to protect schools facilities and equipment, vital records, and other facility contents; and to provide color-coded maps illustrating the exact locations of facilities within potential storm surge zones. The guide also offers suggestions for funding sources to support preparedness, mitigation, and recovery efforts.

- IMLS grant project to assess emergency preparedness: The State Archives, with the State Library and State Museum, is leading an Institute of Museum and Library Services (IMLS) Connecting to Collections. Entitled "Partnership for New York's Cultural Heritage," the project will survey the state's cultural repositories to learn the risks to their collections, preservation challenges, and their educational needs. A focus of the assessment will be emergency preparedness and recovery needs. Statewide partners include the New York State Council on the Arts, New York Library Association, New York Archives Conference, Museum Association of New York, Lower Hudson Conference of Historical Agencies & Museums, and Upstate History Alliance. The survey will result in a statewide preservation plan intended to set the course for years to come, thus ensuring sustainability and consistency of action and even greater investment in New York's cultural heritage.

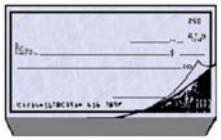
State Archives is also a partner in a nationwide Intergovernmental Preparedness for Essential Records project (IPER). This project uses lessons learned by archivists from Hurricane Katrina and the New York floods of 2006 to increase awareness of the importance of intergovernmental and inter-professional collaboration in protecting essential records. The project also seeks to develop training for emergency managers and records managers at every level of government.

The IPER project received $2.5 million dollars in funding from the Federal Emergency Management Agency (FEMA) and is coordinated by the Council of State Archivists.

Recommendations

- **Recommendation:** The Electronic Records Committee working in conjunction with CIO Council action teams to address e-records preservation and access needs should take into account emergency preparedness and the degree and manner in which open standards and formats can increase the interoperability of information in an emergency.

Approach: State Incentives



The RFPC asked whether the State should provide incentives for either proprietary software vendors to include open formats within their products, or for software which more directly supports open formats to have its other functionality improved. Almost all of the commenters agreed the State should do so. A few commenters suggested the State should itself actively participate in the creation of open standards and the open source projects which adopt those standards. Many commenters recommended that the State should afford greater weight to openness in the procurement process as the State seeks to meet its various IT functional needs. One commenter noted that the questions implied the market is failing to meet the State's IT needs.

The market *is* failing to meet the State's IT needs. The State has an obligation to address the e-records needs of its fellow government agencies and its citizens. But the State has not been able to secure office suite software which simultaneously meets the State's concurrent functional needs of full featuredness and full openness.

An effective approach would include the State continuing to monitor standards and work through organizations such as the National Association of Chief Information Officers (NASCIO) to promote the State's needs in the creation of open standards.

One succinct RFPC response reinforces this point: *"Use the approach of giving state preferences for the use of more open formatted software whenever the functionality of the software exceeds the user base's needs."*

Recommendations

- **Recommendation:** After creating a formal definition of the types of openness which meet the State's IT needs, the State should ensure that the degree to which proposed IT solutions meet that definition is evaluated in State procurements.

- **Recommendation:** Until the market matures and begins better providing for the State's openness needs, meeting the State's definition of openness should be evaluated as a non-mandatory but desirable specification in IT procurements for IT systems which produce electronic data.

Approach: Using Alternative formats

There are many governmental jurisdictions both publishing and receiving office suite documents in open formats such as ODF.

But governments also have taken less formal and various approaches to maximize public access to e-records. For example, some government agencies recognize a need to provide alternative formats in general for e-records, usually accommodating requests for paper copies. The City of Bloomington, Indiana has a specific page on its website where members of the public can request e-records in alternatives formats.[63] And the New York State 2007 ESRA Guidelines recommend, in accordance with law, providing access to e-records in the form the user prefers.[64]

To maximize access, some other governmental agencies automatically post "unofficial" versions of their e-records on their websites for easier public consumption. For example, the New York State Department of State posts versions of the New York Code of Rules and Regulations on its website with the caveat that the posted version is not the *official* format but is being posted merely to increase availability.[65]

Other government agencies have essentially deprecated usage of proprietary formats for Internet publication. For example, the United States Department of Commerce Office of the CIO has published a federal web-publishing guideline entitled "Universal Web Pages: Guidelines for Exceptions" wherein they note:

> *"Use appropriate formats. Where alternative formats are needed, open formats are best because they are available to a broad base of applications on multiple platforms, thereby making the data available to a broad spectrum of users without additional cost. This is not always the case with closed, proprietary formats that are processed by a limited number of programs from specific vendors. ...Use a proprietary format (for example, SAS, SPSS, SQL, MS Excel, MS Word, etc.) only if the format provides functionality not otherwise available. If you must use a proprietary format, ensure that the intended audience is known to have ready access to the appropriate software. Always provide a link to download the appropriate viewer, plug-in, or related software. Be aware that not all viewers or plug-ins for proprietary formats are available for multiple platforms. Plug-ins for open formats have the widest availability."*[66]

Recommendations

- **Recommendation**: In the short-term, it may serve the needs of many of the State's constituents to receive access to unofficial versions of the State's e-records in the formats they desire, even if these versions do not faithfully adhere completely to the official version. The proposed ERC should examine this type of informal approach as well as more formalized formatting requirements.

D. Implementation Issues

The State is clearly encountering problems associated with past failures to fully implement open standards and formats. The idea of imposing in response some immediate, drastic remedies is tempting. But "rip and replace" changes tend to be costly and error-prone. These are some of the issues which need to be examined in greater depth by the proposed Electronic Records Committee to ensure the State imposes sustainable implementation.

Issue: Standards Bodies

Most open formats and applications are created collaboratively. In many instances the technical documentation needed to develop software that can access and create files saved in a given open format is approved and maintained by a national or international standards organization. The most significant of these organizations is the International Organization for Standardization (ISO), a network of the national standards bodies of 157 nations. ISO is a non-governmental organization, but many ISO standards have acquired the force of law through treaties or national standards established by the governments of member nations.

ISO standards for open file formats are developed by the national delegations of experts of technical committees in accordance with established procedures. There are clear advantages to adopting file formats that have received the approval of a national or international standards organization:

- *Reliable documentation.* Organizations such as ISO (and its American member body, the American National Standards Institute) are established entities with considerable resources. Allowing a standards organization to assume responsibility for maintaining and disseminating a file format's technical information increases the chance complete and accurate documentation will be readily accessible for a long time.

- *Increased openness.* Format creators seeking to have their formats recognized as a standard relinquish at least some legal control over third-party use of the format. As noted above, by submitting the technical documentation for the current version of the PDF format to a standards organization, Adobe Systems is ceding its legal right to prevent others from creating software that can create and access files encoded in this version of the format.

However, it must be emphasized that many file formats commonly regarded as open or as standards have not received the approval of a standards organization:

- *De facto open formats.* Some formats are created by corporations that retain their legal rights to the format but freely distribute the technical documentation needed to create compatible software. For example, even prior to its vote in standards bodies for certification as a formalized standard, PDF has long been viewed as a de facto open format because of its widespread use and its creator's decision to release the format's technical documentation to anyone who seeks it.

- *De facto standards.* Some proprietary formats are regarded as de facto standards because of their omnipresence.

Additionally, the process of developing standards is not removed from political or economic considerations. For example, during the debate as to whether the Office Open XML format for word processing documents, spreadsheets and presentations

should be recognized as an international standard, supporters and opponents of the format's standardization repeatedly alleged that various players improperly sought to influence the process.

The RFPC described media reports suggesting the process for standardizing formats is corruptible.[67] It then asked what weight the State should give to whether a particular format has been accepted as a standard by a recognized standards body.

Commenters generally agreed that acceptance by a recognized standards body is an *indication* "that a standard meets the quality criteria for an open standard." But commenters also noted it is most important that a standard possesses actual openness and neutrality best demonstrated by "the availability of solid implementations by different vendors for different platforms as well as the level of substitutability of different implementations."

File formats, optimally, are accompanied by published specification documents and reference implementations describing "exactly how the data is to be encoded" and used for determining "whether or not a particular program treats a particular file format correctly." Some file format developers neglect to develop or deny public access to their specification documents, interfering with the ability of other developers from developing software applications which correctly render those formats.[68] The State's need for openness cannot be met if a standard cannot be implemented by multiple vendors.

Commenters had a wide range of thoughtful opinions about formal standardization. Many of the comments are insightful. Readers of this report are encouraged to review them in Appendix III of this report. It is difficult from the outside for a customer such as the State to discern the absolute truth behind the lack of perfect conversion or interoperability between the various formats. For example, there has been criticism of software applications using either standard (ODF and OOXML) for failure of those software applications purporting to use those formats to use format versions which actually conform to the official standards.[69]

Some have attributed self-interested design of certain office suite software mechanisms for the lack of fidelity during conversion between OOXML and ODF.[70]

What is clearest is that the market has not done a sufficient job of addressing its governmental customers' interests. As countless governments including the State of Minnesota in its recent report have agreed, neither standard, nor their implementations, are currently complete for the needs of government. State governments need harmonizing between the two standards and complete conformance of the software applications which use those standards with the actual specifications. This perspective is common amongst governments who have studied these issues. The same position was taken at the Advancing eGovernment Conference in February 2007.[71] As the conference attendees, twenty-one (21) European Community governments concluded:

"For all parties involved, the exchange of documents and data between authorities, businesses and citizens must be possible without technical barriers. The public administration must not exclude anyone from participating in an electronic procedure owing to the use of a specific product. The Member States are agreed that in the future electronic documents should be exchanged fully on the basis of open document exchange formats."[72]

A demand for harmonization between the two formats was also repeated by multiple nations in their official comments to OOXML's ISO standardization vote in September 2007.[73] Harmonization between the two

standards was called for on the eve of the March 2008 re-vote on ISO approval of the OOXML standard by Microsoft's CEO in France, Eric Boustouller. Mr. Boustouller has proposed the creation of a working group which *"will be in charge of harmonization so as to create a better interoperability between the two distinct open standards, Open XML and ODF."*[74] And as results of the second ISO vote were being announced on April 1, 2008, harmonization of the two standards was also being endorsed by IBM and others.[75]

According to the ISO, the decision to develop standards is made when anybody who needs a particular standard communicates that need to their nation's ISO member organization. In the United States, the ISO member organization is the American National Standards Institute (ANSI).[76]

According to ANSI's document, *"Essential Requirements: Due process requirements for American National Standards,"* if *"any person"* (such as a government agency) has *"a direct and material interest"* in standards development, the agency may express its position to ANSI and have its position considered. Also, ANSI promises to use good faith efforts to resolve potential conflicts between and among existing standards when those are brought to its attention.[77]

In 2007 Massachusetts reversed its decision to use only ODF format in its office suite documents and agreed also to use OOXML, stating a hope that each standard would improve.[78] Now two other states -- Minnesota and New York -- have also said clearly they would prefer one document standard, and that the two standards that are competing within ISO are not sufficient for state government needs. As Minnesota stated in its report, *"Moving in the direction of a fully documented functional document standard that can do all one wants is desirable. But neither of the competing standards [ODF and OOXML] addresses all the government goals and purposes in [Minnesota's study e-records study statute]."*

It is important to note that anyone can denominate their standard as "open." But without a specific definition of what openness means, that is just a label. Moreover, the fact of standardization by a formal standards body may be admirable. But formal standardization is no guarantee that the formalized standard will be faithfully adopted in software made available to the State and as such meets the functional needs of the State and its citizens.

It remains to be seen the extent to which the two competing ISO-approved office suite standards (OOXML and ODF) will be subject to transparent governance by a broad community of stakeholders. But it seems apparent the openness of a format should extend to its governance.

Recommendations

- **Recommendation:** Working with the National Association of Chief Information Officers (NASCIO), New York should explore with other state governments making a formal request to ANSI for the harmonization of ODF and OOXML into a single document standard that meets the needs of government.

- **Recommendation:** Once the State creates its own State-favorable definitions of the terms "open standards" and "open formats" and defines its functional needs to include meeting those definitions, then whenever possible the State needs to use standards which adhere to those definitions.

Issue: Appropriate Government Control

The study is required to show how electronic documents are created and preserved in a manner that encourages "appropriate government control." The term "appropriate government control" was not defined in the statute nor further explained in its bill memorandum. The phrase is susceptible to multiple meanings. One interpretation concerns the application of appropriate security and privacy controls to electronic records. For example, if State agencies were to publish more of their documents on public websites, what would be the best mechanisms to ensure individual privacy rights for the subjects of certain of those records are protected as required by law?

Another interpretation concerns the interoperability and longevity of government electronic records for various purposes of government -- for example, for archiving, researching or creating merged databases of documents for useful comparative analysis -- without any roadblocks placed in government's path by artificial proprietary lock-in barriers. Some obvious and oft-cited examples are the Katrina and tsunami emergencies where governments were unable to control the useful emergency gathering of information because e-records were essentially kept in proprietarily locked, non-interoperable agency-specific silos.

and in turn received a set of responses addressing both issues. Concerning the security and privacy afforded by the use of ODF or OOXML for electronic records, several comments extolled the greater security provided by ODF not just for the reasons usually cited (openness of software code or development processes), but also for such reasons as that:

- It can be encrypted so as to be totally impossible to decipher without a password, unlike OOXML which presents a security risk as it must be converted by an external program to be read on most systems;
- It offers better portability of data, which offers the State a better choice of enciphering tools, which eventually gives better security; and
- ODF is less complex due to the strong reuse of standards and concepts which makes it easier to scan ODF-using documents for malicious or sensitive content.

Throughout the RFPC responses governmental commenters frequently raised their differing needs when it came to e-record formats and standards. This reinforces the recommendation for development of a long-term, cross-agency collaborative entity (an Electronic Records Committee) to address issues such as these.

Recommendations

- **Recommendation:** The State should assess the privacy and security impacts of any new software which it procures including any risks arising from the formats used by that software.

- **Recommendation:** Through the proposed ERC working with the appropriate CIO Council action teams the State should examine the differing needs of different State agencies and the degree to which their distinct IT choices affect government control of its own e-data.

Issue: Assistive Technologies

The RFPC received few comments from entities or individuals identifying themselves as users of assistive technologies.
One State agency described a situation when Microsoft Office 2003 was provided to the agency where "people using screen reading software, like JAWS, could not read Word 2003 documents" because of the way in which the default view had been set by the application.

Both the OOXML and ODF formats have been criticized for their support for assistive technologies. For example, in 2006 Massachusetts noted that ODF-supporting office suites were *"unlikely to be fully supported by assistive technology vendors, or alternatively to include fully functional adaptations in the packaged product,"* by the time frame slated for its migration to ODF formats.[79] By 2008, that apparently still hadn't changed. According to the most recent version of the Massachusetts Enterprise Technical Reference Model (ETRM v. 4.1, dated 2.6.08):

> *"As of the date of publication of the ETRM v. 4.1, there are no office applications that natively support ODF that also provide sufficient accessibility for persons that use assistive technology devices. While work is ongoing in this area, at this time, the only implementation option available to agencies is the use [of] ODF through the use of translator software."*[80]

Similarly, the Adaptive Technology Resource Centre in Canada noted in January 2008 concerning OOXML:

> *"Conclusions*
>
> *There are grave issues with respect to the accessibility of Office Open XML as a format and potential standard that should preclude its adoption at present. It may be the case that OOXML can be improved to ameliorate some of the more specific technical concerns, but it is most likely too late for the higher-level issues, especially those inherent in the process by which OOXML was developed. We suggest that energy would be better spent in the ongoing effort to improve the existing ISO ODF standard (with which OOXML would overlap and compete if it is adopted). In any event, decisions with respect to standardized document formats should be made in consultation with members of disability communities, disabilities experts and developers of assistive technologies, with universal accessibility as a core requirement as opposed to an ad hoc afterthought.*
>
> *6. Addendum [added in Version 1.3]*
>
> *After briefly investigating Microsoft's revision (1.01) to their report, "Accessibility of Ecma Office Open XML File Formats", we are more disappointed than ever. The superficial change in their claims about what is and is not supported seems to be (1) a deliberate attempt to circumvent the spirit of the accessibility guidelines and checkpoints to 'whitewash' OOXML and/or (2) a demonstration of incompetence and lack of understanding with respect to accessibility. Further, we would like to re-emphasize the need for consistent, clear, and interoperable means of providing information to Assistive Technology, requirements that are sorely lacking in the accessibility "solutions" offered by Microsoft and in the proposed OOXML standard itself. We conclude by reiterating, with renewed urgency, that there can be no substitute for a thorough accessibility review by experts, including developers of assistive technologies and members of disability communities, for standards that are as fundamental and important as this."*[81]

Comments like these reinforced the need for the State to fully vet *any* new formats to ensure that moving to those formats will not

decrease the functionality needed by the State's workforce and its other citizens. In relation to assistive technologies, it appears neither the OOXML nor ODF formats currently fully meet the needs of the State.

Several of the commenters noted that assistive technology questions implicate more than just file formats but also application programming interfaces (API). One State agency commenter suggested the State CIO should establish an assistive technology advisory committee. Whenever the State's technology choices risk decreasing access for those using assistive technologies, it would be prudent for the State to consult with assistive technology specialists.[82]

Recommendations

- **Recommendation:** Even for those State agencies which are planning migrations to Microsoft Office 2007, the State should consider not adopting either the ODF or OOXML formats until these assistive technology issues are fully resolved. Earlier formats such as .doc or .pdf could continue to be used.

- **Recommendation:** The Electronic Records Committee should reach out to assistive technology specialists whenever the State risks, because of its technology choices, decreasing access to those who need assistive technologies.

Issue: e-Discovery

The RFPC's questions about e-Discovery engendered some of the widest-ranging opinions by commenters, perhaps because this is still such a dynamic and developing issue. The workgroup noted some basic facts and assumptions:

Discovery is the name of the process used in lawsuits to gather relevant information to support or defend parties' positions. In December 2006, the rules governing discovery in federal civil cases were amended to address how parties preserve, collect, and produce electronically stored information. Although these rules relate specifically to federal civil cases, litigants involved in lawsuits in nearly all jurisdictions must increasingly address issues relating to the discovery of electronically stored information.

The format in which electronically stored information is produced in a lawsuit may itself be the subject of dispute. Depending on the circumstances, parties may produce documents in their original format or in "reasonably usable" formats which facilitate indexing and searchability. In certain situations documents may be produced to reveal embedded data such as mathematical formulas associated with spreadsheets.

Litigation support software applications help render the parsing of huge volumes of electronic data easier during lawsuits, and the State uses some of that software. Examples include, but are not limited to, applications by Concordance (http://www.concordancesoftware.co.uk/) and Summation (http://www.summation.com/).

Other tools which can assist with e-discovery include those applications which pre-categorize data based on search terms, help weed out privileged documents, integrate and work with litigation support software and do Bates numbering.

Depending on when a lawsuit arises, the electronic records may be available and maintained by an agency in the normal course of business, or they may have already been preserved for archival purposes. Therefore, electronic discovery issues should be addressed in both contexts: agencies' maintenance of electronically stored information in the ordinary course of business, and standards and formats adopted for purposes of archival preservation.

Ready preservation of metadata, and assurances of e-record authenticity, can equally serve preservation needs as well as litigation needs. Formats meeting the combined needs of preservation and e-discovery also ensure that appropriate documentation or "metadata" (information about information) is captured and maintained. This metadata is needed to facilitate access to records and support statements about the records authenticity and authority.

It is noted the State Archives' two most prolific types of electronic records requestors are genealogists and then lawyers, in that order.

Several commenters said that formats didn't matter. Others expressed concern that for State Archives records maintained in proprietary formats, e-Discovery demands would require costly efforts by the State to ensure continuing access. Some commenters noted more litigation support software is currently written to work with predominant proprietary office suite software, but others noted that as open formats increase in popularity this will change. Many commenters noted that ideally records would be archived in the same format as they were created, because otherwise disputes concerning reasonable access to original or archived versions of e-records could raise great complexities and costs in litigation.

Commenters were equally divided about the use of alternative formats such as PDF/A. Several State agency respondents expressed strong concern that such formats are susceptible to tampering and lack needed functionality such as the ability to use encryption or to embed audio or video within documents, and may raise intellectual property issues.

A number of commenters noted that e-Discovery issues highlighted the need to develop comprehensive records systems and planning.

Finally, several State commenters recommended in general deferring to the perspective of the State Office of the Attorney General on this type of issue.

Recommendations

- **Recommendation:** The State Office of the Attorney General, State Archives and the Division of the Budget need to be part of any collaborative discussions on integrating openness into the State's technology procurements, including having representation on the recommended Electronic Records Committee. The types of data access needed by the OAG must be addressed, and any solutions must interoperate with the OAG's preferred litigation support software.

Issue: Intellectual Property

The RFPC asked about intellectual property challenges and the means for the State to protect itself from those. It also asked whether formalized patent promises such as the Open Specification Promise and the OpenDocument Patent Statement afforded sufficient protection.

Several commenters noted that the State typically procures software through formalized bidding procedures and vendor contracts, and as such can continue to protect itself from intellectual property claims through protective contractual terms and conditions. Currently, software is usually licensed by the State accompanied by ongoing service agreements for vendor-provided maintenance. Some commenters suggested the State could save money by using open source software which can be obtained at no cost. One commenter interestingly recommended the State use the open-source licensing model to share its internally developed software with other entities.

Any Intellectual Property issues concerning formats can likely be managed through contractual terms with the vendors who either supply or service the office suite software used by the State. This reinforces, however, concerns about the risks associated with treating e-data producing software such as office suite software as mere commodity purchases whose terms and conditions need not be negotiated by the State.

Recommendations

- **Recommendation:** The State should protect itself from IP threats using contractual indemnification for any software procurements and should not rely on vendors' covenants not to sue.

- **Recommendation:** Collaboration amongst State agencies on sharing software developed by the State is desirable. But sharing internally-developed State software outside of the confines of the State raises complex questions with licensing and security concerns which should be evaluated by the proposed Electronic Records Committee working with CIO Council action teams on a case-by-case basis.

Issue: Vendor Neutrality and Promoting Competition

The 2007 Electronic Signature and Records Act (ESRA) Guidelines offer several recommendations also considered crucial by the State Archives. One of those recommendations is that for their e-records it is optimal for State agencies to: *"Maintain the e-record's original functionality to the degree necessary: Many e-records lose their meaning and usefulness if they cannot be used or function as they did when they were in their original environment (e.g., ability to be processed or searched). Determine if it is necessary to retain an e-record's functionality. If so, the record should be retained in a format that can be processed or used by available technology."*

The other recommendation is that State agencies should establish standards for file formats: *"Policy should designate approved data file formats for each record "type." All information stored on a computer system requires software for retrieval and display. This software is subject to change, either by the implementation of new releases, or by changes to operating systems or hardware. A policy of approved media formats for records storage will facilitate data migration to ensure long-term retrieval of e-records."*

The problem is that operationalizing this optimal condition within State agencies presents real difficulties, no matter which "standard file formats" are chosen. As noted, converters do not work very well. And as the State government commenters in particular noted in response to the RFPC's vendor neutrality question, a rip-and-replace strategy will be crushingly complex and costly.

It appears from this study that as the varieties of software on the market that directly support truly open formats gain feature richness, those varieties are more likely, over time, to be adopted by all end users.

One of the most common responses to the RFPC stated in sum or substance that open formats had been designed with vendor and application neutrality in mind, and thus choosing software that directly supported those formats was itself a neutral choice. Some commenters noted the plethora of support for ODF in the marketplace.

Many other commenters noted that the State's procurement laws require vendor neutrality. As such, these commenters believed that merely defining the State's needs and including those needs in procurement documents would ensure neutrality. Finally, one commenter recommended not choosing a particular format, but rather for any given acquisition defining the business case most relevant to that procurement.

This study found that the State's existing procurement processes have sufficient controls to ensure neutrality. This is the best mechanism for the State to use as it measuredly seeks to operationalize openness within purchases of IT which create e-data. Open standards and formats compliance is a feature that will need to be weighed by the State in comparison to its other compelling needs. There are several existing guidance documents which describe how governments can accomplish this task. [83]

That this balancing is needed can easily be illustrated by just one example: specialized software and documents. The RFPC posed a question concerning highly specialized software such as Computer Assisted Drawing (CAD). Most commenters agreed that with this type of software openness principles might need to yield to other considerations for utility's sake. State Archives has agreed with that "exception" as well in the past. And as Minnesota recently recommended in its own e-records study, state governments

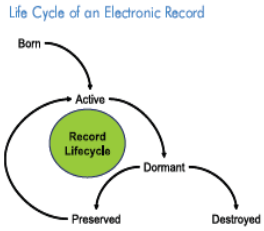
should: *"Respect the effort and objectives of document creators in storing them for future reference. In particular, documents which integrate graphical elements, tables and other components should not be forced into overly simplified formats unable to accurately represent the document purpose and the meaning and intent of the content."*

Recommendations

- **Recommendation:** The State should rely on existing procurement law and practices to ensure vendor neutrality in its IT purchases.
- **Recommendation:** The State should identify the use of open standards and formats as a desirable feature in those procurements to be weighed against other desired specifications.
- **Recommendation:** As individual IT procurements are considered and evaluated for the manner in which they fit within the State's enterprise IT plans, before the specific procurement of openness becomes sought as a mandatory specification the State should develop a full business case explaining its costs and benefits.

Issue: Obsolescence and e-Records Lifecycles

Records are not static.[84] Over time, they may no longer require regular access and most lose their value to organizations. For example, a contract file will be actively used until the terms of the contract have been met. At that point the file has little value and often may be destroyed. However, a small percentage of records, considered to be archival, may retain their value and may warrant preservation.



It is critical for organizations to manage electronic records well at each stage of the life cycle. Each stage is dependent on preceding stages - bad decisions made during the creation stage can render the record inaccessible or make the cost of making the records accessible prohibitive at the preservation stage. Unlike paper records, each stage requires active intervention on the part of the stakeholder involved. For example, if the creator of the record does not migrate the record to a newer software version access to the record will be lost when the original version is no longer supported.

There are usually different stakeholders involved in each stage. Both IT and program unit staff may be involved in the creation stage. Users, staff and the public may be stakeholders in the active use stage. Users, IT staff, and records storage staff may be involved in the inactive use stage. Finally, state and local governments and archival staff may be involved in the preservation or destruction stage.

There are four general stages of the records lifecycle:

1. *Creation.* Records are created in the normal course of business to satisfy legal

and fiscal requirements and to meet other administrative needs. Electronic records may be created:

a. Originally in electronic format, also called "born digital"

b. From paper or microfilm records that are scanned or digitized to an electronic format

Increasingly, electronic records are born digital and no paper or other format equivalents exist. Records creators need to carefully consider the design of the electronic records system and the file formats utilized at this stage of the life cycle. It is much easier (and cheaper) to build in components, such as records retention controls, at this early stage than to wait until a later stage.

2. *Active Use*. During this stage, organizations distribute records to other organizations, frequently refer to the records, and store the records close by for ready access. To facilitate efficient retrieval, records are usually indexed. Organizations should carefully consider how users search for information within records (e.g., by author, date, subject) and ensure adequate and accurate metadata[85] exists for those search categories. Security and access controls should be placed on the records to prevent unauthorized access to records in order to ensure their integrity and authenticity.

3. *Inactive Use*. At some point, records are no longer frequently referred to and they are removed from active storage on computers and active servers. For electronic records, this means storing the records offline on more cost efficient servers or electronic media (e.g., CDs or computer tape).

4. *Preservation or Destruction*. Only a small percentage of electronic records have enduring legal, fiscal, or historical value and warrant long-term or archival preservation. Most have a limited retention period and will be destroyed at the end of their life cycle. Archival records should be transferred by State agencies to the State Archives.

Focusing on the lifecycle of electronic records, the National Association of Chief Information Officers (NASCIO) has issued a wide-ranging set of recommendations for State CIOs to adopt.[86] The recommendations pertain primarily to integrating electronic records management into the respective states' enterprise architectures (EAs) but they equally could pertain to operationalizing within the EAs implementation of the subset of open standards known as open formats. NASCIO described its primary recommendation as follows:

> *"The key recommendation for the state CIO is to collaborate and team with the state experts: records managers archivists and state librarians. Enlist their help in establishing policy, reviewing options, assessing technology solutions along with procurement, project management and enterprise architecture ... The creators of records, records managers, and state archivists must partner with the CIO and the state enterprise architect to ensure that electronic records management and digital preservation issues and discipline are integrated with the states investment process, project and program management, and the overall enterprise architecture strategies for managing knowledge assets of the state.... [Governments] must recognize the knowledge assets of the enterprise and manage them as enterprise assets."*

The federal government similarly has recommended U.S. federal agencies identify and embed e-records management requirements within their enterprise architectures. In doing so, the National

Archives and Records Administration states that federal agencies *"will realize benefits such as compliance with relevant laws and regulations, consistent records management practices across the agency, improved customer service, and real cost savings."*[87]

RFPC comments reflected several common themes. Several commenters suggested use of open formats. Other commenters, particularly those from State government, noted the need within State government for specific guidelines to assist with records characterization and preservation. One commenter pointed the State to several sources of best practice guidelines for electronic records management. The most consistent theme amongst the comments expressed in various ways was a need for the State to integrate its enterprise architecture with its classifications of electronic records to effect business process change. This was expressed succinctly by one commenter as follows:

"We believe the better approach is to have records managers and archivists proactively work with system administrators to define what types of electronic objects are in fact "records" and how they should be managed throughout their lifecycle."

One item which this report has not addressed concerns the best methods for State agencies to transmit e-records in accordance with their agencies' records retention schedules to State Archives for archiving. The reason is there is no one-size-fits-all solution to this issue. It must be addressed on an ongoing, case-by-case basis.

Few State documents require preservation as *official* State records by State Archives (on average five (5%) percent for any state government, including New York). Under existing laws and standard records retention principles, agencies need to make retention decisions. State Archives cannot take all records from an agency simply because they are encoded in a preservation-friendly format.

A State agency may have a need to maintain in or convert to an open format for meeting others of its needs (such as making its documents more accessible during their active lives to the general public). But conversion to open formats prior to an agency's transmission to State Archives needs to be done only for that small percentage of an agency's e-records being transmitted to State Archives for preservation.

Timing of transmission to State Archives is another consideration. The active lives for some agencies' e-records can extend to many dozens of years. Thus, specific standards dictated by State Archives today for transmission to State Archives of records which may not be transmitted until dozens of years from now is impractical. Any transmission standards mandated today are likely to have been superseded by new technology realities dozens of years from now. Rather, agencies will need to develop working structures for retention scheduling of electronic records. Such structures, as recommended by the National Association of State Chief Information Officers and others, should be as automated as possible. At the time of transmission to State Archives, State Archives will need to engage in in-depth analyses which will be highly dependent upon the manner in which respective State agencies ultimately integrated e-records management and open standards into their existing technical frameworks.

State Archives has identified a need for sufficient resources in order to develop strategies for creating ongoing methods for integrating its archiving processes with agencies' e-records management frameworks.

Recommendations

- **Recommendation:** As NASCIO has recommended, State IT personnel need to consistently work with State archivists *"to ensure that electronic records management and digital preservation issues and discipline are integrated with the states investment process, project and program management, and the overall enterprise architecture strategies for managing knowledge assets of the state."*

- **Recommendation:** The Electronic Records Committee should be charged with ensuring this collaborative work.

- **Recommendation:** As more records transferred to the State Archives by State agencies are in electronic format, ensuring that those records meet long-term preservation needs will require additional technical and professional skills and resources. The Electronic Records Committee should explore options for recruiting and retaining qualified State Archives staff and leveraging resources to assist State Archives in preserving archival electronic records.

Issue: Technical Issues

The RFPC asked a series of questions about technical differences between office suite software formats. As already noted, concerning highly specialized software and the e-records it produces, commenters generally recommended using open formats where possible but tolerating closed formats if necessary. Most commenters stressed that the use of closed formatted e-records should be kept to the minimal amount necessary. Several commenters suggested the State learn more about converting these documents from various recommended specialists across the world.

Concerning the State's Freedom of Information Law and the manner in which the State's format choices might impact the State's ability to comply with its legal requirements, several commenters suggested the use of open formats could streamline the process or even create new XML-based ways for the State to provide documents to its citizens on-line. But for the most part, commenters did not see the State's FOIL obligations as significantly affected by format choices. A few commenters noted that if the State needs to convert e-records before providing them to the public pursuant to FOIL, the conversion process could slow down the State's responsiveness.

The RFPC asked about whether migration to other office suite formats would be impacted by the fact some government agencies may have written custom software applications using certain technologies (e.g. ActiveX; scripting; Microsoft Access) which may need to be re-written. Several commenters pointed out that new iterations of formats and software by the same vendor often are so different from each other that this is a concern no matter which path of format the State follows.

This study considered whether the office suite format question particularly matters in terms of the State's interactions with the public. It might be possible in certain contexts (e.g. web publication) for the State to consistently provide documents in formats such as .txt, HTML, or PDF. However, those formats, too, raise their own functionality questions. For example, concerns have been raised that the HTML code produced by Microsoft Word *"is semi-proprietary, and it is prone to include information which cannot be displayed on all platforms."*[88] Thus, creating in or converting to HTML code for publication could be problematic. As for PDF, it is appropriate as a presentation format but it doesn't retain the original nature needed by State Archives nor the editability needed in many cases. And, as one State agency commenter noted, PDF/A lacks certain needed functionality -- it lacks the ability to embed within it audio or video, a capacity for encryption, and may be subject to intellectual property concerns.

Finally, the RFPC asked about the technical shortcomings and benefits of the prominent competing office suite formats, ODF and OOXML. Comments traversed a range of issues. Most of the comments focused on the perceived lack of features possessed by each of the formats. One commenter suggested the most objective manner in which to test these assertions would be for the State to perform a "side-by-side comparison ... in a controlled lab environment." When assessing the functionality available from each format and the software applications which implement them, this is a recommended approach.

This commenter and others further compellingly noted that "standards" change rapidly, and it can be unwise to lock into a particular standard.

There are two ways to circumvent these problems. First, rather than immediately selecting a particular office suite format, the State should initially concentrate on defining the desirable features which it needs and can obtain from the use of particular software formats. Then, specific formats and software can be chosen based on the question of whether they meet the State's specifications. This is why this report has settled on the cleanest and simplest of solutions: ensuring openness gets a seat at the table whenever the State evaluates its IT features needs.

Second, to the extent any particular formats or software are identified as preferred standards which come closest to meeting the State's needs, some type of oversight body needs to ensure those formats and software actually address the State's e-records needs in practice. Such policies or preferences must also remain up-to-date and not become outmoded. The ongoing cross-agency e-records committee could stay on top of issues like these.

Recommendations

- **Recommendation:** The State will need to evaluate a host of issues as it adopts new office suite software and formats. State agencies should take into account the types of technical issues raised by commenters in response to this study's RFPC.
- **Recommendation:** Concerning the State's procurement processes, an ERC can help evaluate these types of issues from a cross-agency perspective.
- **Recommendation:** Rather than immediately selecting a particular office suite format, the State should initially concentrate on defining the desirable features which it needs which the State can obtain from the use of particular formats or software which implements those formats.
- **Recommendation:** The ERC, in conjunction with agency CIOs, should advise and set parameters for formal evaluations of various technologies to determine whether those technologies are meeting the State's needs for openness and interoperability.

Issue: Costs

As one author has noted in an economic analysis of the value of standards, *"interoperable standards allow natural monopolies of **technologies** (standards) while providing for competition among vendors," and that "standard[s] provide a platform **above which** innovation can take place freely and collaboratively."* (emphases added).[89] Accepting this view, in expressing a preference for openness in formats the State would not be rejecting choice but rather aligning its requirements and the needs of its users at the level at which choice from a user perspective is maximized.

The RFPC asked about the costs which would be incurred because of the State's standards and formats choices for citizens, vendors, and the State itself. Concerning costs for citizens, as several commenters noted software directly supporting open formats can be acquired and continually updated to new versions at no cost. Thus, that software can help to bridge the digital divide and reduce barriers to citizen access to government.

Several commenters asked for web-published documents to be provided in PDF format for "presentation" documents which do not need changes and for ODF formatted documents to be made available for forms needing completion and return to the State. Some commenters specified that PDF format should not be used for forms needing completion because this would require purchase of specific software which goes beyond the mere reading of PDF documents. However, the last two versions of Adobe Reader (the Adobe freeware product) allow users to fill in PDF forms. This software is exactly the same software that would be used for "the mere reading of PDF documents."

One persistent issue is that the implementation of comprehensive, well-considered technology architectures can be inhibited by implementation costs and competing State needs. This is one reason that rather than creating comprehensive new rules or employing a sudden "rip and replace" strategy, the State instead should *measuredly* begin adopting greater technology openness. This can be done by integrating the need for openness into existing State practices and guidelines, given the State's recognition of such openness as a desired State technology feature.

To the extent they can, State agencies should design data systems which preserve in open formats before their e-records are transmitted to State Archives for preservation. However, competing goals or realities such as lack of open alternatives or an agency's need for functionality not present in open alternatives may prohibit this. If so, then alternatively to the extent possible those agencies should develop means to convert their e-records to open formats prior to transmitting those records to State Archives.

The e-records study bill asked CIO/OFT to study the "costs of implementation." This presumably concerns the initial costs that would be incurred in implementing open format-deploying solutions within State government. A valid concern is that this implementation cost language, if construed too narrowly, would omit consideration of longer term costs commonly referred to as "Total Cost of Ownership" (TCO). State agencies should consider all prospective costs for any technology solutions which they procure, including not just initial implementation costs but also any costs for retraining, consulting, and service fees.

This report does take note, however, that: (a) a multitudes of studies have found significant cost savings after governments have implemented open format solutions, even taking into account all of the costs besides implementation costs; and (b) despite requesting such information in its RFPC, no studies taking into account all costs which found cost savings from remaining with or migrating to closed format solutions have been brought to CIO/OFT's attention. This latter point is telling. Despite expressly being asked, not one of the one hundred fourteen (114) commenters to the RFPC cited a single study finding HIGHER costs after migration to open format solutions.

New York State citizens should not encounter financial barriers or software application preference barriers when accessing information which they are legally entitled to receive from the State. In order to facilitate transparency in government, the State needs to be able to accommodate choice. The State also needs to remain flexible given that file formats evolve over time. Thus, to the degree to which it is reasonably possible to do so when responding to citizen requests for records, the State should provide e-records to its citizens in a variety of formats, including ODF when requested.

Recommendations

- **Recommendation:** Those State agencies ready to implement more immediate transitions to more open architectures should issue RFQs to discover and share the best methodologies available for meeting the State's desired feature of preserving documents in open formats.

- **Recommendation:** State agencies should develop cost models for integrating openness into their e-records retention or conversions which can be applied within those agencies and shared with and used by other State agencies on a case-by-case basis.

- **Recommendation:** State agencies must perform comprehensive business case analyses before requiring any implementation of particular document formats. These business cases need to include risk analysis, stakeholder impact analysis, and Total Cost of Ownership (TCO) cost and benefits analysis, including costs for conversion, training, maintenance, and support costs.

- **Recommendation:** State agencies should work with the Division of Budget and the Office of the State Comptroller to assist with performing comprehensive cost audits as needed when evaluating e-data producing open format and software choices.

E. Sustained Operationalization Of Openness

Recommended State Procurement Changes

State Technology Law §103(4) delegates to the New York State Office for Technology (OFT) the power to review and coordinate technology purchases. OFT is empowered to review agency IT purchases for consistency with agency technology plans. Those technology plans must be in conformance with the statewide strategic technology plan. The State Chief Information Officer (CIO), who sets IT policy for the State and oversees OFT, is delegated oversight and coordination responsibilities by Executive Order 117. This oversight over State IT purchases is further reinforced by State Division of Budget requirements such as "H-300A – Technology and Information Resource Management Planning Process."[90]

In this oversight role, CIO/OFT assesses the consistency of the purchase request with the statewide strategic technology plan and the requesting agency's technology plan, implementation of statewide technology standards including privacy, confidentiality and data security safeguards, and the proper dissemination of public information. In fulfilling this role, the CIO/OFT could assess conformance with the State's declared adherence to open standards.

Historically, the State has purchased office suite software through the use of State Office of General Services (OGS) centralized commodity contracts. According to the State Procurement Council's procurement Guidelines, a State agency is permitted to purchase from centralized commodity contracts if the item the State agency needs is available in the form, function and utility consistent with an agency's need.[91]

These centralized office suite software contracts typically have been based on single source justifications,[92] negotiated with the strongest focus by OGS on obtaining the best price.

According to the New York State Procurement Guidelines, State agencies have dual requirements when procuring IT. First, "It is the State agency's responsibility to determine that a need exists for a particular Service, Technology or Commodity." The State through CIO/OFT has previously identified that use of open standards is a State need. This report is defining a further State need -- to sustainably operationalize over time the adoption of truly open standards in all State procurements.

The second obligation of State agencies identified by the Procurement Guidelines is for each State agency "to select and document the appropriate procurement technique for meeting [its] need [for a particular Service, Technology or Commodity]." One commonly used procurement technique is to use centralized contracts, based on single source justifications, using template terms and conditions.[93]

But that technique does not readily accommodate the approach this report recommends for ensuring sustained progress towards integrating openness into procured software - evaluating various features of the software offered by competing vendors and ensuring the evaluation addresses a non-mandatory feature of openness.

There are a few possible methods for integrating the State's need for openness into the procurement process. First, the template terms and conditions for contracts such as the centralized office suite software

contracts could be modified to include specific language identifying the State's need. And, either alternatively or as a companion mechanism, the State's Plan to Procure process could be modified to ensure that when State agency procurements are evaluated for adherence to Statewide budgetary and technology strategic approaches, a portion of that evaluation could weigh the extent to which the procuring agency is appropriately integrating open standards and formats into its IT purchase. This was also a recommendation made by Minnesota in its recent study report on these same issues. The Minnesota study recommended revising that State's equivalent to New York State's Intent to Purchase form so it will begin requiring Minnesota state agencies to justify the use of non-open formats and to explain the methods Minnesota state agencies will use to convert e-data to open formats.

Openness has not typically been evaluated as a feature needed by the State. Why is it being so identified now? As this report has demonstrated, there is:

- a growing recognition of e-records interoperability problems. This was illustrated, for example, by events which took place during Hurricane Katrina;
- a greater State sensitivity to cost;
- apparent consensus unrefuted by any countervailing studies that proprietary lock-in generally raises costs;
- greater penetration of Internet access, leading to increased global interactivity;
- an exponential increase in the international usage of open formats;
- a growing recognition of the need for governments to interoperate with citizens using open formats. This has been illustrated by the acceptance and regular use of open formats by, for example, the governments of the State of Washington and in Warren County, New York; and finally
- increased demand for sustainable operationalization of open standards and formats by the State's constituents. Amongst those asking the State to use open standards and formats are other State agencies, the State Archives, and several businesses and individuals in New York State.

Most commenters supported the RFPC's proposal to integrate openness evaluations into the State procurement processes. In fact, several commenters expressly supported this direction as the best possible option for the State in response to the e-records study bill and laid out detailed and sophisticated recommendations for doing so systemically. One State agency commenter expressed concern the CIO/OFT Plan-to-Procure process governing State agency IT purchases is already cumbersome. However, as this report is being issued that process is being streamlined and made more efficient, so this concern will likely soon be resolved.

Defining with specificity its need for openness and carefully integrating the acquisition of that feature in a measured fashion into existing State technology procurement processes will minimize any sudden disruptions of a vibrantly competitive technology marketplace. Additionally, this will afford vendors a level playing field to compete in creating rich software applications meeting all of the State's needs, including its need for format openness. As with any procurements, State agencies will continue to specify the mix of features they need (including openness amongst many, many features), and balance the mix of features offered by bidders with the cost of those solutions. There are no barriers prohibiting any vendors from adding openness to the mix of features their software offers.

In response to the State's RFPC, the Business Software Alliance (BSA) provided comments including the following:

"We understand that government leaders may want to define what product features are needed across the government space, such as security, accessibility, interoperability, reliability, and value. But then such leaders should allow individual government customers to choose the solutions that best meet their specific objectives. To reiterate, governments should avoid a one-size-fits-all approach that could prevent procurement of the best product at the best price. ***The focus should be on how well the product meets the desired features****, not on preferences for specific standards, processes or technologies. Procurement rules should leave room for competing solutions to develop and for agencies to shift to new solutions as technologies advance and needs change. Governments should allow market forces to select the best solutions in individual cases and not mandate a specific approach (such as a standard), except when the industry as a whole has backed a single approach."* (emphasis added).

This perspective is supportable. It needs to be understood, however, that in office suite software, as in any e-data creating software, the user need for openness is increasingly being recognized as essential. Thus, openness will become one of the State's "most desired features." This report's recommendations are in alignment not only with the comments of sophisticated entities such as NASCIO and Lockheed-Martin, but also fully consistent with BSA's quoted statement.

Staffing

One question which the RFPC did not address directly is training needs. One of Minnesota's express recommendations was for that state to *"[p]rovide adequate resources for thorough training of employees who manage government information (data, records, and documents) at every stage in their life cycle. Restore important records management functions within agencies and at a coordinated state level in order to manage government information throughout the entire life cycle of creation, access and final disposition."*

The 2007 State ESRA Guidelines reinforce a similar point: *"Training, which ensures that staff adequately maintains systems used to create and retain e-records, is critically important in preserving the authenticity, integrity and legal admissibility of e-records. In addition, it is important to ensure awareness of the unique management issues associated with e-records, such as the fragile media on which e-records are stored, the technology platform needed to access and use e-records, and the responsibilities to manage e-records diligently to ensure their admissibility in legal proceedings and their accessibility throughout their legal retention periods."*

CIO/OFT is currently enhancing the curriculum of the New York State Technology Academy, which provides training to State agency and local government IT personnel. Expanding the curriculum to include development of open standards skills could accomplish two useful things:

- increase the pool of employees knowledgeable about open standards and open formats, which would facilitate integration of these standards into the State's IT architecture; and
- further demonstrate the State's commitment to operationalizing open standards and make State IT positions more attractive, particularly to new graduates.

Many IT professionals currently view State IT as tied to the moribund technologies of the past, not as a

forward-looking, challenging environment in which to work. [94]

State Archives is also seeking to ensure its staff members have sufficient expertise. In the electronic environment the traditional responsibilities of records managers and archivists -- working with records creators to identify records that warrant permanent preservation, protecting records from tampering and physical harm, and helping researchers identify and use records -- will remain largely unchanged. However, archivists and records managers will need new skills in order to manage, preserve and provide access to electronic records. State Archives staff have sought to acquire these new skills in a variety of ways:

• attending numerous workshops and presentations concerning electronic records preservation;
• conducting pilot inventories of electronic records systems maintained by local governments and State agencies; and
• participating in multi-state grant projects that address specific electronic records preservation and access issues.

Incorporating open standards and formats into the State's IT infrastructure would allow State Archives to amend its training program to:

• devote less attention to learning how to overcome problems associated with the use of proprietary formats; and
• focus upon learning to manage and preserve records created and saved in open formats and to enhance public access to State Archives holdings.

Moreover, State Archives provides training and guidance to State agency and local government records managers. Adoption of open standards and formats would enable State Archives field staff to:

• spend less time helping customers struggling with problems caused by usage of proprietary formats and recordkeeping systems; and
• focus upon proactively helping customers create recordkeeping systems that meet current business and future archival needs.

Electronic Records Committee

This report has made frequent mention of the proposed Electronic Records Committee (ERC). Improving electronic records management is one of NASCIO's and the State's top four identified CIO Priorities for Strategies, Management Processes and Solutions. Integrating openness into State IT can serve many of the State's information technology strategic goals including:

- increasing agility and enterprise-wide leverage in IT systems;
- improving technology capabilities;
- increasing cost efficiencies and greater economies of scale;
- improving service to our State agency customers and to segments of the general public;
- creating conditions which would facilitate transformational changes; and even addressing impending baby boomer retirement issues by making State IT a more attractive place to work.

Meeting these goals will entail not just technology changes but policy changes as well. These types of issues need to be addressed in both CIO Action Team meetings and also through the more policy-oriented ERC.

These procurement recommendations address the long-term openness needs of the State,

but not its shorter-term needs. As such, it is recommended that some of the first issues which the ERC should review include ways to address short-term issues. For example, the ERC could:

- determine whether multiple office suite applications should be installed on State desktop computers so State employees can open files submitted in diverse formats by the public and save files intended for dissemination in multiple formats;
- help agencies identify formats meeting particular needs (for example, ODF might be suitable for simple letters, and OOXML might be preferable for complex spreadsheets with functionality missing from ODF or complex documents tied to databases); and
- reach out to vendors to quickly ascertain whether there are converters good enough to produce converted documents meeting State business and legal requirements.

Once these immediate concerns are resolved, the ERC can then help to identify best practices supporting the State's ability to:

- create and appropriately manage electronic records;
- exchange information across jurisdictional boundaries;
- respond to legal discovery requests;
- prevent inadvertent or unauthorized destruction or alteration of records; and
- preserve records that are of long-term value to their creators or have enduring legal or historical value.

Establishment of an ERC was discussed briefly in the State's Electronic Signature and Records Act report in 2004: *"New York State should consider adopting an interagency collaborative approach in addressing long-term electronic record preservation solutions, thereby focusing the resources and expertise of multiple agencies on an issue of continuing importance to government entities."*[95] This report provides detail as to the proposed structure of the ERC.

Recommendations

- **Recommendation:** An evaluation of the degree to which any given IT procurement adheres to the State's need for open standards and formats should be part of the Annual Technology Plan and Intent to Procure processes or any such successor programs.

- **Recommendation:** The State should consider using RFIs that describe the State's full office software feature needs to help determine if a mix of applications is available in the market that meets the State's needs.

- **Recommendation:** As CIO/OFT works collaboratively with the Division of the Budget and the Office of General Services to jointly promote the use of statewide Technology Aggregate commodity contracts and Enterprise or Universal Licensing Contracts, those agencies should include terms and conditions which address the use of open standards and formats and evaluate the extent to which bids supply those features.

- **Recommendation:** Agencies should reflect in their procurement records consideration of the long-term, ancillary and TCO costs of their office suite software purchases, including the costs which can be incurred from deploying software that does not natively support formats adhering to the State's definition of open formats and open standards.

- **Recommendation:** The NYS CIO Council through its Procurement, Sourcing and Vendor Relationships Action Team should develop strategies for integrating e-records considerations into the Plan-to-Procure and Annual Technology Plan (ATP) evaluative processes.

Recommended Implementation and Migration

One commenter noted: *"As we have discussed throughout, it is important to recognize that data formats are just one feature of an application ... And, new formats can be accessed by pre-existing applications via application updates, translators, plug-ins and the like. Therefore, there is no logic to the argument (often espoused by ODF-only advocates) that adoption of a new format necessitates adoption of a new application. In this regard, the WordPerfect-to-Microsoft transition actually is irrelevant to the ODF/OXML line of questions that permeates the RFPC."*

Under examination, this assertion is not sustainable. The assertion holds true only if the new format is in conformance with published standards and fully supported by the software application vendor. All formats used in software purchased by the State need to be adopted faithfully by software vendors without extensions and deviations from the published formats. This report describes the limitations of those "*application updates, translators, plug-ins and the like.*"

Migration can be necessary between older and newer versions of proprietary office suite software as well, causing additional, ancillary costs and raising that software's Total Cost of Ownership (TCO). For

example one company, ConverterTechnology, specializes in assisting large entities with such Office-to-Office migrations. ConverterTechnology has estimated there are approximately two hundred (200) incompatibilities between Microsoft Office 97 and 2003 which need to be remediated, and another two hundred (200) between Office 2003 and Office 2007.[96]

There are additional "costs" to migration as well, in terms of loss of data and context. In a report published by the Council on Library and Information Resources, the authors concluded, "with migration, as with all translations, some information is lost, no matter how skilled the interpreter. In migration, it is usually the context, rather than the data, that drops out or is improperly reconstructed in the new code. This can be crippling in dynamic formats, in relational databases, and even in simple spreadsheets."[97]

Gartner has recommended not focusing exclusively on the use of open standards, but instead weighing such factors as which selection in a given instance furthers the deployment of public services, and an assessment of cost versus public value. Gartner observes that accepting the use of multiple standards in the short term can allow a migration towards open standards where appropriate in the long run.[98]

Several government entities[99] are allowing a mixture of formats, but that mixture consists of Open Document Format or Microsoft's pre-OOXML formats such as .doc. These entities have defined OOXML as unacceptable, and pre-OOXML proprietary formats as becoming deprecated, but still useful.

Another common migration method is to create Interoperability, Access or Web Publishing Architecture documents such as the Utah Department of Technology Services "Technical Architecture Review"[100] as companions to a government's Enterprise Architecture documents. The State should explore doing so.

One commenter recommended the following: *"Any successful strategy will begin by identifying the stakeholders and working with them to identify processes, both manual and automated, that are currently tied to a particular document format. Then it will set a date for agencies to achieve a stated capability, such as "All agencies must be able to accept documents in ODF format from the public and from other agencies." Capabilities can be rolled out in a natural progression: the ability to accept documents in a given format, the ability to produce documents in a particular format, etc."*

Such progressive approaches make sense, as does the perspective of several commenters stressing that for any such approach, a comprehensive business case should first be developed.

CIO committees and the ERC, working in tandem, can define the best migration path in any given instance.

Recommendations

- **Recommendation:** Working with the appropriate CIO action teams, the ERC should help the State to develop web publishing formats Best Practices and other guidelines after reviewing similar government guidelines such as those that are posted on the US Department of Commerce and other government sites.

- **Recommendation:** Working in tandem, CIO action teams and the ERC can also define the best migration path towards implementing openness in e-data producing IT in any given instance.

- **Recommendation:** During the development of those migration paths, those same entities can develop business cases to fully vet the issues of all stakeholders and ensure that migration decisions are fully supportable.

- **Recommendation:** The financial costs and ancillary impacts must be addressed in any given State technological procurement to ensure a seamless migration.

- **Recommendation:** If an effort is undertaken by NASCIO or others to develop nationwide government standards concerning openness, State agencies should observe and participate in such efforts when reasonable to do so, but not be beholden to their decisions if they do not address the State's specific needs.

Recommended Enforcement

OFT has the statutory authority to promulgate rules and regulations pertaining to the use and acceptance of electronic signatures and records in accordance with the Electronic Signatures and Records Act [State Technology Law § 303(2)(a)]. State Archives has statutory authority to promulgate rules and regulations pertaining to management of all, and long-term preservation of archival, records of state and local government (including those in electronic formats).

More importantly in relation to the subject of this study and report, the State Archives is responsible for making available those State official records, including electronic records, that have been deemed sufficiently valuable to warrant continued preservation by the State.

Commenters did not speak directly to the question concerning whether to provide State Archives with enhanced authority or responsibilities in relation to the issues addressed in this report. Instead, they spoke about State Archives having the ability to reject non-compliant content which is transmitted to it by State agencies. They recommended the State create a collaborative committee which could help to ensure compliance with certain guidelines and audit for failures to comply. This report suggests a number of recommendations at its outset in this regard which the State Legislature might wish to consider.

F. Conclusions

From Minnesota's e-records report:

> *"This report does not recommend the adoption of a particular format standard. The dynamic nature of technology innovation and change make adoption of a single standard problematic. Moving in the direction of a fully documented functional document standard that can do all one wants is desirable. But neither of the competing standards proposed addresses all the government goals and purposes in the law. In any case, the choice or use of a standard must not be to adopt a standard for the sake of adopting a standard. Any choice must be in the context of what value such a decision adds to government."*

This New York study and report have come to very similar conclusions. A single open standard would be optimal for each document type produced by any state government. But sadly because of deficiencies in the market, in terms of office suite documents neither of the competing standards (OOXML and ODF) and the software which implements them fully comprises all of the features needed by New York State. One solution is a more feature-rich but less truly open productivity suite. The other solution is more truly open but less feature-rich. The solution, then, is not to impose one incomplete standard or another. Rather, the solution for the State lies in ensuring the openness feature needed by New York State is recognized and has a full seat at the table so the degree of openness is evaluated as comprehensively as other features typically are by State agencies procuring such software.[101]

There is always a tension between mandating the use of single standards for cost and efficiency's sake, versus accommodating the diverse needs of users. The State has existing processes for evaluating technology purchases financially and technically. Those processes try to ensure a balance between competing needs and continual movement in the direction of best value. Purchases are tested against various State technology principles. Within its Enterprise Architecture principles, the State has already expressly recognized that open standards are one of the State's desired features. However, the State has yet to fully operationalize its processes for evaluating the type of open standards known as open formats.

Certain principles revealed by this study should be re-emphasized:

- The optimal is not currently possible. The optimal solution is in the hands of vendors, not government. The only thing government can do is incentivize vendors to do the right thing, doing so by defining the State's functional needs and including openness within those needs. Rip and replace is not a practical strategy because of cost reasons.

- The IT field is dynamic. The State could issue this report, and shortly thereafter find there have been major technological developments rendering the report outdated. This report is current right up to its publication date, and focuses on global issues which transcend the immediate.

- The State has a diverse constituency of both internal and external customers. Some of its constituency recognizes the value of and are embracing open standards software. Others do not value it as highly. The State needs to try to reasonably meet the needs of as many of its constituents as possible.

- Government analyses like this report consistently have concluded the optimal situation would be for all parties to agree on a single truly open office suite format standard. That way, IT users could choose whatever software they wished to based not on formats but on their requirements for other features such as cost or integration with ancillary software.

This report focused largely on electronic document formats, for five reasons:

- First, that is the context in which the many bills introduced in state legislatures in 2007 were originally created.
- Second, the document format question works very well in serving as an illustrative example of open access to State e-records generally.
- Third, because of the debate over formal standardization by the ISO, there is a wealth of commentary available about different approaches to the formats issue.
- Fourth, members of the public who provided the State with written comments primarily addressed the document format issue.
- And finally, there are many ways in which access to the State's electronic records may be enhanced. But enhancing access through the use of appropriate electronic document formats in itself would be an enormous undertaking which, if accomplished, would facilitate other types of enhancement in the future.

Open Document Format (ODF) and Office Open XML (OOXML) are competing electronic document formats. Some urge that large end-users of software such as New York State government should permit the use of multiple types of document formats because this maximizes "choice." On the surface, that argument seems untenable. It is axiomatic in technology administration that setting a level playing field using *single* standards permits choice elsewhere such as choice of vendors or choice of software applications. After the NYS Office of the Chief Information Officer/Office for Technology (CIO/OFT) solicited responses from the public in a Request for Public Comments published on the agency's website, one commenter flatly noted when referring to ODF and OOXML: *"When it comes to interoperability standards, choice between essentially equivalent standards is NOT desirable."*

But the question is larger than this. Also axiomatic in technology administration is that different users have different needs. For example, worldwide there is significant competition between the most prominent web browsers (Firefox; Internet Explorer; and Safari) because of users' preferences. Even Lynx, a text-only browser that is sixteen (16) years old, still has its adherents, as do other variants. Similarly in State government, while the Microsoft Office software suite is predominant as it is in much of the United States, several State agencies still prefer and continue using Corel Word Perfect because it is perceived as better meeting their functional needs. Differing technologies for e-mail such as Microsoft Outlook, Novell GroupWise, or Lotus Notes each have their State agency adherents.

Software users' preferences are diverse because of the varying types of functionality those users need. It is New York State public policy, expressed in the Freedom of Information Law and other sources of law, for the State to be as accessible to our citizens as it can. The State must accommodate the public's information needs. There is strong evidence of growing demand for Open Document Format within the State, the nation, the world, and in the responses which

CIO/OFT received to its survey. Many members of the public perceive the openness of electronic document formats as an important one of their functional technology needs.

Because of this, *the openness of electronic document formats also becomes a functional technology requirement for State government itself*. The State must balance many different types of functionality in the software which it uses. Software openness is one of the many features the State needs. The State needs it to meet citizen requests. The State needs it to meet its need for long-term accessibility and usability of its most valuable, archival records. The State needs it to maximize the number of and competition between vendors and to limit State costs.

It has been said the Internet works so well because it was designed to perceive interference with its mission as unacceptable and to route around it. The State, like all governments, is caught in a circumstance of market inability to fully deliver the State's functional needs. Like the legislation that spawned it, this report perceives this interference with the State's mission, and offers concrete suggestions for routing around it.

Appendix III

COMPARATIVE CHART:
National Governments requiring use of open formats

Area	Year	Levels of Government Subject to Policy			Decision: What?	Action plan: How?
		Nat.	State	Local		
Belgium	2007	yes	no	no	Government agencies are <u>required</u> to: - be able to access documents in ODF format (September 2007) - use ODF for intra- and interagency exchange of word processing documents, slideshow-style presentations, and spreadsheets (September 2008) Government agencies are <u>permitted</u> to: - use other formats when creating and distributing documents internally	<u>Enforcement</u> is "bottom up," driven by "buy in" and "consensus," not compulsion <u>Ongoing control</u> by FEDICT (Federal Public Service Information and Communication Technology) <u>Next</u> steps - evaluation of ODF and other proposed formats (PDF/A, PDF/X, and SMIL) and inclusion of approved formats into the Belgian Government Interoperability Framework
Brazil	2007	yes	no	no	Government agencies are <u>required</u> to: - incorporate "recommended" file formats into new information systems and upgrades to existing systems - ODF - RTF - PDF/A - UNICODE -HTML -XML	<u>Enforcement</u> through regulation <u>Ongoing control by</u> Ministry of Planning, Budget, and Administration, Secretariat for Logistics and Information Technology <u>Created</u> -Reference document (*e-PING Standards of Interoperability for Electronic Government*) <u>Changed</u> - IT development and procurement requirements <u>Next steps</u> - development of agency-specific implementation plans
Croatia	2007	yes	no	no	Government agencies are <u>required</u> to: - make each government form posted on a public Web site "accessible in a way that makes it legible in accessible freeware applications" - Options for meeting this requirement: - ODF	<u>Enforcement</u> through regulation <u>Ongoing control</u> by Central State Administrative Office for e-Croatia Created

Area	Year	Levels of Government Subject to Policy			Decision: What?	Action plan: How?
		Nat.	State	Local		
					- PDF - HTML	-Reference document (*Operational Plan for the Implementation of e-Croatia*)
Denmark	2008	yes	yes	yes	Government agencies are <u>required</u> to: - accept ODF and OOXML "data-processing documents" created by the public, businesses, and other governmental units Government entities are <u>permitted</u> to: - refrain from implementing the new OOXML and ODF standard if doing so would incur "additional costs or inconveniences" or raise "IT security concerns"	<u>Enforcement</u> through regulation <u>Ongoing control by</u> Ministry of Science Technology and Innovation <u>Created</u> - mandatory standards for document exchange (OOXML and ODF), data exchange (OIOXML), electronic case and document management (FESD), electronic purchasing (OIOUBL), digital signatures (OCES), public Web sites and accessibility, and IT security (DS484) <u>Changed</u> - government procurement practices: "IT solutions purchased after 1 January 2008 must support" ODF, OOXML, or both, "and be capable of accepting data-processing documents in both formats, where necessary through the use of plug-ins (supplementary programs)." <u>Next Steps</u> - third-party evaluation of whether OOXML and ODF standards will remain mandatory after July 1, 2009 -development of implementation plans for other standards
France	2007	yes	no	no	Government agencies are <u>required</u> to: - accept documents submitted in ODF format - use PDF/A to preserve text documents Government agencies are <u>encouraged</u> to: - install OpenOffice.org (an open-source, ODF-based productivity suite) - use ODF to create text documents, spreadsheets, and slideshow-style presentations Government agencies are <u>prohibited</u> from: -migrating to any productivity suite formats other than ODF	<u>Enforcement</u> through regulation <u>Ongoing control by</u> General Directorate for the State's Modernization <u>Created</u> - Interoperability framework (*Référentiel Général d'Interopérabilité*)

Area	Year	Levels of Government Subject to Policy			Decision: What?	Action plan: How?
		Nat.	State	Local		
	2007	yes	no	no	The National Assembly has required that: - open source software, including the ODF-based OpenOffice productivity suite, be installed on all computers used by all members and staff of the Assembly	Enforcement through proclamation Ongoing control by President of the Assembly
Netherlands	2008	yes	yes	yes	Government agencies are required to: - migrate new systems to Open Standards and Open Source - "support ODF alongside existing file formats for reading, writing and exchange of documents" - create open source implementation strategies; - share with each other knowledge gained from agency pilots, migrations, and test cases; - work collaboratively with stakeholders on openness plans	Enforcement through "soft approach" with "high trust" but with "monitoring and ranking" of efforts and award of an annual prize for the "Most Open Public Organization" Ongoing control by Ministries of Economic Affairs and Internal Affairs Created: - broad definitions of the terms "open standards" and "open specifications" - list of open standards - interoperability framework; - advisory process allowing proposed technology purchases to be submitted to advisory board for assistance Changed government procurement practices: - "Comply or explain, and commit" principle: For new systems, modifications, or contract extensions: (a) government purchases must "comply" with open standards, or government purchaser must "explain" why not; (b) purchases must be timed, and good business cases developed, to avoid unnecessary migration costs; and (c) once an open standards compliant purchase has been made, government agencies must commit on an ongoing basis to giving "preference to the application of open standards so that an exception criterion is no longer applicable."
Norway	2009	yes	yes	yes	Government agencies are required to: - make government documents posted on the Web accessible to the public in three (3) mandatory formats: - HTML (presentation) - PDF (presentation) - ODF (manipulable).	Enforcement through regulation Ongoing control by Ministry of Government Administration and Reform Created: - Reference document (*Reference Catalogue of IT Standards in the*

Area	Year	Levels of Government Subject to Policy			Decision: What?	Action plan: How?
		Nat.	State	Local		
					- develop ability to RECEIVE documents from the public in these formats. Government agencies are <u>permitted</u> to: - use other formats, but only if they also produce the documents in PDF or ODF format.	*Public Sector*) <u>Next step</u>: - evaluating additional formats for document exchange with public sector and exchange of documents within the public sector
South Africa	2008	yes	yes	yes	Government agencies are <u>required</u> to: - comply with MIOS standards ("*Minimum Interoperability Standards (MIOS) for Information Systems in Government*") - migrate new systems and major upgrades to MIOS standards - make their IS systems Internet-centered, with access to public services and documents preferably via web browser - consider open source based solutions before proprietary ones - use ONLY the following formats: - for working office suite documents (word-processing, spreadsheet, presentation): ASCII, CSV, or ODF - for office suite documents provided in presentation view: XHTML or PDF - work collaboratively with stakeholders on openness plans	<u>Enforcement</u> through law, audit, and budget: - compliance with MIOS standards falls under South Africa's Public Finance Management Act (PFMA) - funding and purchasing government IS/ICT projects is dependent on compliance with MIOS and is subject to auditing by accounting officers and by the South African Auditor-general - cost of non-compliance with MIOS expressly rests with the non-complying entity, system or organization. <u>Ongoing control by</u> SITA (the South African State Information Technology Agency) <u>Created</u>: - broad definitions of the terms "open standards" and "open specifications" - MIOS standards within S. Africa's Enterprise Architecture, to be reviewed and updated annually after discussions with stakeholders - interoperability framework covering three key areas of technical policy: (1) Interconnectivity; (2) Data Interoperability; and (3) Information Access. <u>Changed</u>: - government procurement practices
Switzerland	2007	yes	yes	yes	Government agencies are <u>required</u> to: - use the following formats when exchanging documents with citizens or other agencies: - PDF/A ("urgently recommended") - PDF/X ("recommended") - ODF ("recommended under observation") - OOXML ("recommended under observation")	<u>Enforcement</u> through regulation <u>Ongoing control by</u> eCH (federal e-government standards agency) <u>Created</u> reference document (*SAGA.ch*)

Area	Year	Levels of Government Subject to Policy			Decision: What?	Action plan: How?
		Nat.	State	Local		
United Kingdom - British Educational Communications and Technology Agency (BECTA)	2008-2009	yes	yes	yes	BECTA is a UK government agency which oversees the procurement of all information and communications technologies (ICT) for all levels of UK schools. Funding for UK educational ICT purchases depends upon buying from an approved list of companies which comply with BECTA's "framework agreements." BECTA urgently recommends that Microsoft provide native support for the OpenDocument format (ODF), and recommends: (a) against any UK educational institution using OOXML; (b) that in the short-term UK educational agencies continue using older Microsoft binary formats (such as .doc); and requires (c) that in the future these agencies purchase office productivity suites that are *"capable of opening, editing and saving documents in the ODF format and setting ODF as the default file format."* BECTA is creating a framework for interoperability compliant procurement of UK educational purchases.	Enforcement through creation of approved purchasing frameworks and through funding Ongoing control by BECTA (federal UK agency which oversees the procurement of all information and communications technologies (ICT) for all levels of UK schools) Created a final interoperability report (January 2008), and is creating an interoperability procurement framework agreement within "the next twelve months."

Sources:

1. Belgium: *"Open Standards: Belgium's Federal Council of Ministers Approves ODF (Open Document Format)"* (Ministry of Employment and Computerization, June 23, 2006), http://presscenter.org/archive/20060623/432d0130470a88df1105dda38d1282b0/?lang=nl&prLang=en; *"Belgian Government Chooses Open Document"* (CNET News, June 23, 2006), http://www.news.com/2100-7344_3-6087275.html; *"Belgian Government Adopts Open Document"* (*Techworld*, 29 June 2006), http://www.techworld.com/applications/news/index.cfm?newsID=6335&pagtype=all; *"File Types and Document Formats"* (Belgian Government Interoperability Framework, September 18, 2007), http://www.belgif.be/index.php/File_type_and_document_formats; *"Open Standards and ODF in Belgium"* (Presentation of Chief Information Officer, FEDICT, at the 1st International ODF User Workshop, October 29, 2007), http://www.odfworkshop.org/3.pdf.

2. Brazil: *"e-PING Standards of Interoperability for Electronic Government, Version 2.0.1"* (Executive Committee on Electronic Government, December 5, 2006), https://www.governoeletronico.gov.br/acoes-e-projetos/anexos/E15_677e-PING_v2.0.1_05_12_06_english.pdf; *"e-PING Padrões de Interoperabilidade de Governo Eletrônico*, Versão 3.0" (Executive Committee on Electronic Government, December 14, 2007), https://www.governoeletronico.gov.br/anexos/e-ping-versao-3.0

3. Croatia: *"Interoperability"* (Central State Administrative Office for e-Croatia, 2006), http://www.e-hrvatska.hr/sdu/en/ProgramEHrvatska/Provedba/Interoperabilnost.html.

4. Denmark: *"Agreement on the Use of Open Standards for Software in the Public Sector"* (Ministry of Science Technology and Information, October 2007), http://videnskabsministeriet.dk/site/frontpage/information-and-communication-technology/agreement-on-the-use-of-open-standards-for-software-in-the-p

5. France: "*Référentiel Général d'Interopérabilité Volet Technique,*" version 0.90 (General Directorate for the State's Modernization, November 14, 2006), https://www.ateliers.modernisation.gouv.fr/ministeres/domaines_d_expertise/architecture_fonctio/public/rgi/referentiel_general1617/downloadFile/file/Referentiel_General_Interoperabilite_Volet_Technique_V0.90.pdf?nocache=1163526872.97; *"Les Postes Micro-informatiques des Députés Seront Dotés de Logiciels Libres à Compter de la Prochaine Legislature"* (National Assembly, November 22, 2006), http://www.assemblee-nationale.fr/presse/divisionpresse/m01.asp; *"Mise á Disposition d'un Kit d'Installation Open Office.Org" (General Directorate for the State's Modernization, June 20, 2007), http://www.thematiques.modernisation.gouv.fr/chantiers/411_75.html; "Rationalisations des Moyens Mutualisations des Méthodes"* (General Directorate for the State's Modernization, August 13,. 2007); *http://www.thematiques.modernisation.gouv.fr/chantiers/409_75.html*

6. Norway: *" Reference Catalogue of IT Standards in the Public Sector"* (Ministry of Government Administration and Reform, December 21, 2007), http://www.regjeringen.no/en/dep/fad/pressesenter/pressemeldinger/2007/Open-document-standards-to-be-obligatory.html?id=494810

7. Netherlands: *"The Netherlands in Open Connection: An Action Plan for the Use of Open Standards and Open Source Software in the Public and Semi-public"* (Netherlands Ministry of Economic Affairs, November 2007) http://appz.ez.nl/publicaties/pdfs/07ET15.pdf

8. South Africa: *"Minimum Interoperability Standards (MIOS) for Information Systems in Government (v.4.1)"* (South Africa Department of Public Service and Administration, October 2007) http://www.oss.gov.za/MIOS_V4.1_final.pdf

9. Switzerland" *"SAGA.ch, standard no. eCH-0014, version 4.01"* (eCH, June 22, 2007), http://ech.ch/index.php?option=com_docman&task=doc_download&gid=1160&Itemid=78&lang=fr; *"CH: New eGovernment Standards"* (ePractice.eu, July 24. 2007), http://www.epractice.eu/document/3724

10. United Kingdom (BECTA): *"Microsoft Vista and Office 2007: Final report with recommendations on adoption, deployment and interoperability" (BECTA Report January 2008), http://news.becta.org.uk/display.cfm?cfid=1989752&cftoken=8d5d40a46f6be0ce-6F4CC738-F9EA-10F9-A47DDB7F8B7EFF9E&resID=35287&page=1658&catID=1633; http://publications.becta.org.uk/download.cfm?resID=35275*

Appendix IV

COMPARATIVE CHART:

Provincial/State/Regional Governments requiring use of open formats

Area	Year	Levels of Government Subject to Policy		Decision: What?	Action plan: How?
		State	Local		
Andalucía (Spain)	2008	yes	no	Government agencies are required to: - use open standards to interchange documents with citizens, non-governmental organizations, and other government bodies" - ODF - PDF - PDF/A - HTML - XHTML - ASCII - UNICODE	Enforcement through regulation Ongoing control by Secretariat of Telecommunications and Information Society Created - interoperability framework
Assam (India)	2007	yes	[unknown]	Government agencies are required to: -use ODF -transition to open-source operating systems	Enforcement through regulation Ongoing control by information technology department
Extremadura (Spain)	2007	yes	no	Government agencies are required to: -create and save in ODF (manipulable) or PDF/A (static) format all documents that will be exchanged with other government entities or with citizens	Enforcement through regulation Ongoing control by Ministry of Education, Science and Technology
Hong Kong (China)	2006	yes	no	Government agencies are required to: - Exchange text documents in ODF, HTML, RTF or a mutually agreed-upon legacy format - exchange spreadsheets in ODF, or a mutually agreed-upon legacy format -exchange slideshow-type presentations in ODF, or a mutually-agreed-upon legacy format	Enforcement through "self-regulation" of bureaus and departments; OOXML not approved as its interoperability is under examination Created - Interoperability framework (*The HKSARG Interoperability Framework*)
Massachusetts (United States)	2007	yes	no	Government agencies are required to: - create and save "official records" in one of six (6) "open" or "acceptable" formats: - ODF (open) - OOXML (open) - HTML (open)	Enforcement through regulation Ongoing control by Information Technology Department Created - Reference document (*Enterprise Technical Reference Manual 4.0*)

Area	Year	Levels of Government Subject to Policy		Decision: What?	Action plan: How?
		State	Local		
				- ASCII (open) - RTF (acceptable) - PDF (acceptable)	Changed - procurement policies
Misiones (Argentina)	2006	yes	no	Government agencies are required to: -create and save in ODF (manipulable) or PDF/A (static) format all documents that will be exchanged with other government entities or with citizens	Enforcement through regulation Ongoing control by Computer Center of the Province of Misiones and the Main Directorate of Communication Networks and Computer Science
Paraná (Brazil)	2007	yes	no	Government agencies and state-owned companies are required to: -use ODF for the creation, storage, and display of all electronic documents	Enforcement through law.

Sources:

1. Andalucía: *"Junta de Andalucía's Digital Interoperability Framework"* (Junta de Andalucía, Council of Innovation, Science, and Business, October 2007), http://www.odfworkshop.org/2pdf

2. Assam: *"Achieving Impeccable Openness in Translation with Open Document Format"* (Assam Electronics Development Corporation, October 2007), http://www.odfworkshop.org/9.pdf; "German Foreign Office Comes Out in Favor of Open Document Format" (*Heise Online*, October 30, 2007), http://www.heise.de/english/newsticker/news/98208

3. Extremadura: *"Agreement for the Implementation of Free Software in the Personal Computers of the Junta de Extremadura," "Extremadura Switches Exclusively to Linux and OpenDocument"* (*Heise Online*, August 1, 2006), http://www.heise.de/english/newsticker/news/76219; *"Standard Open Formats and Libre Software in the Extremadura Public Administration"* (*Upgrade*, December 2006), http://www.upgrade-cepis.org/issues/2006/6/up7-6Millan.pdf

4. Hong Kong: *"The HKSARG Interoperability Framework,"* version 6.0 (Government Chief Information Officer, December 2007), http://www.ogcio.gov.hk/eng/infra/download/s18.pdf

5. Massachusetts: *"Enterprise Technical Reference Manual—Version4.0—Domain: Information"* (Information Technology Department, August 1, 2007), http://www.mass.gov/Aitd/docs/policies_standards/etrmv4dot0/etrmv4dot0information.rtf

6. Misiones: *"Resolución DGCC No. 175/06—Boletín Oficial No. 11928"* (Computer Center of the Province of Misiones, 21 December, 2006), http://www.misiones.gov.ar/egov/index.php?option=com_content&task=view&id=133&Itemid=26

7. Paraná: *"Diáro Oficial Paraná,"* digital edition no. 7621(December 18, 2007), https://www.documentos.dioe.pr.gov.br/dioe/publico/localizar.jsp?dtinicial=&dtfinal=&numero=7621&search=&diario_codigo=3&submit=Localizar [click "download do

diário official"]; "*Lei No. 15742*" (December 18, 2007), http://celepar7cta.pr.gov.br/SEEG/sumulas.nsf/319b106715f69a4b03256efc00601826/a4d30af5cd3749bc832573be00431df8?OpenDocument

Appendix V

COMPARATIVE CHART:

U.S. State Government Collaborative Processes for Addressing e-Record Policies / Standards

STATE:	Iowa	Kansas	Michigan	Nevada	Ohio	Texas	Florida (proposed)	New York (proposed)
COLLABORATIVE'S NAME:	Iowa Electronic Records Committee (ERC)	Kansas Electronic Records Committee (ERC)	Michigan Electronic Records Committee (ERC)	Nevada Electronic Records Committee (ERC)	Ohio Electronic Records Committee (ERC)	Texas Records Management Interagency Coordinating Council (RMIC)	Florida Office of Open Government	New York Electronic Records Committee (ERC)
HOUSED IN:	State Records Commission	Part of the Kansas Department of Administration's Information Technology Advisory Board	Joint effort of the Michigan CIO and Department of Management and Budget (DMB)	Subcommittee of the State Records Committee in the Nevada Secretary of State's office.	The State Archives in conjunction with Ohio's Department of Administrative Services' Office of Policy and Planning.	Created by statute, Texas Gov't Code § 441.203. Role of presiding officer rotates between permanent members	Recommended by Florida State Senate to be an office created by state statute.	Recommended by NYS CIO/OFT January 2008 e-records study report to be jointly sponsored by CIO/OFT and NYS Archives
RESPONSIBLE FOR:	Developing statewide government information policies and guidelines addressing e-record issues common to most state agencies.	Recommending and reviewing policies, guidelines and best practices for the creation, maintenance, preservation of and access to e-records created by Kansas state government.	An advisory group recommending enterprise-wide standards and guidance to state government agencies for the creation, management, accessibility, retention and preservation of information in electronic	Assisting all branches of Nevada State government in the management of electronic resources, especially e-records, through the establishment of statewide record keeping policies and practices.	Drafting policies and guidelines for electronic records.	Studying records management issues, reporting to the Governor and Legislature biennially; adopting records policies applicable to all member agencies.	Implementing e-records standards.	• Identifying best practices supporting the State's ability to: - create and manage e-records; -exchange e-records across jurisdictional boundaries; - respond to legal e-discovery requests; - prevent destruction or

STATE:	Iowa	Kansas	Michigan	Nevada	Ohio	Texas	Florida (proposed)	New York (proposed)
			formats.					alteration of e-records; - preserve e-records of long-term value to their creators or enduring legal or evidentiary value; • Informing the State CIO Council and the State CIO/Office for Technology how these best practices should inform the State's technology procurement process. • Informing the State Legislature of legislative or regulatory changes needed to facilitate good electronic recordkeeping.
MEMBERSHIP:	Iowa Information and Technology Dep't.; State Archivist; State Librarian; various state agencies (executive and judicial branches, State AG), universities and local governments	Representatives from all branches of State government, including Kansas Information Technology Office, Attorney General's Office, Office of Judicial Administration,	Includes records managers, archivists, information technology professionals, purchasing officers, attorneys and end-users from	Includes records managers, archivists, information technology personnel, librarians, other information managers, and policy makers representing the	Includes policy makers, records managers, IT personnel, archivists and librarians from various state agencies, universities, libraries and historical	Permanent: the secretary of state; state auditor; comptroller of public accounts; attorney general; state director and librarian; executive	Collaborate with the Florida Agency for Enterprise Information Technology, Office of Open Government, and Division of Library and Information	New York State: • Office of the Attorney General • Office of General Services • Director of State Operations • Office of the State Comptroller • Division of the

STATE:	Iowa	Kansas	Michigan	Nevada	Ohio	Texas	Florida (proposed)	New York (proposed)
		Legislative Database Manager, State Archivist, Department of Revenue, and University of Kansas.	state agencies.	various state agencies, local governments, and universities of Nevada.	societies.	director of Texas Building and Procurement Commission; the executive director of the Department of Information Resources. Auxiliary appointed members: one faculty member of a public senior college or university with knowledge of records and information management; and two individuals who serve as information resources managers for executive branch state agencies.	Services.	Budget • CIO/Office for Technology • Office of Cyber Security and Critical Infrastructure Coordination • State Archives • State Library • Speaker of the Assembly • Majority Leader of the Senate plus 23 constitutional agencies and select Executive Department agencies and non-profit organizations

Appendix VI

PROPOSED NEW YORK STATE ELECTRONIC RECORDS COMMITTEE

Mission/Charge

New York State government agencies and the Legislature are increasingly reliant upon records existing only in electronic format. These records must be created and managed appropriately in order for the State to meet its current and future legal obligations, make prudent use of its resources, and respond appropriately to citizens' requests for current or archival government records.

The State Electronic Records Committee will guide the State in developing and sustaining its capacity to manage its electronic records and ensure the preservation of electronic records having permanent legal or historical value. It will do so by:

- Identifying best practices supporting the State's ability to:
    - create and manage electronic records;
    - exchange information across jurisdictional boundaries;
    - respond to legal discovery requests;
    - prevent inadvertent or unauthorized destruction or alteration of records;
    - preserve records that are of long-term value to their creators or have enduring legal or evidential value;
- Informing the State CIO Council and the State CIO/Office for Technology as to how these best practices should inform the State's technology procurement process.
- Informing the State Legislature of legislative or regulatory changes needed to facilitate good electronic recordkeeping.

Immediate Action Items

The State Electronic Records Committee will focus first upon:

- Identifying state agencies and nongovernmental organizations that will serve on the committee in an observatory/advisory capacity;
- Developing and issuing recommendations to the State CIO Council regarding the incorporation of recordkeeping concerns into the State Information Technology Strategic Plan;
- Providing advice and support to the CIO/Office for Technology regarding:
    - identification of digital information formats and standards that support preservation and accessibility of government information;

  - mechanisms for encouraging the adoption of those standards and formats;

- Furnishing advice and supporting the existing authority of:
  - State CIO/Office for Technology—purchase of information technology used to create State government records
  - State Archives—disposition of State government records.

Proposed Composition of the Committee

Co-chairs—agency head or designee

- NYS Chief Information Officer
- NYS Commissioner of Education

Electronic Records Committee Steering Committee members—agency head or designee

- Office of the Attorney General
- Office of General Services
- Governor's Director of State Operations
- Office of the State Comptroller
- Division of the Budget
- Chief Information Officer/Office for Technology
- Office of Cyber Security and Critical Infrastructure Coordination
- State Archives
- State Library
- Speaker of the Assembly
- Majority Leader of the Senate

Full Committee members—agency head or designee

The other 23 Constitutional Agencies as well as the following agencies in the Executive Department:

- State Emergency Management Office
- Office of Real Property Services
- Committee on Open Government
- State Librarian
- Office of Court Administration

Interested Non-profit Organizations

Non-NYS government organizations have demonstrated interest in the effective management of electronic government records. The State Electronic Records Committee should be given the discretion to determine which non-profit organizations may take part in its deliberations. Organizations that may be asked to do so include, but are not limited to, the following:

- New York State Forum for Information Resource Management

- New York State Local Government Information Technology Directors Association
- New York Association of Local Government Records Officers
- New York State Historical Records Advisory Board
- New York State Local Government Records Advisory Council
- Center for Technology in Government

Secretariat

Responsible for all administrative matters—will likely range from working collaboratively on simple matters such as setting up meetings to more resource-intensive matters such as creation and maintenance of a State ERC Web site and Webcasting of committee meetings:

- Chief Information Officer/Office for Technology
- State Archives

Appendix VII

SPECIFIC NEW YORK STATE ARCHIVES E-RECORDS NEEDS

A. State Archives E-Records Resource Needs

One of the values of using open formats is that they improve public access to government records. But, that access will decline over time if New York does not develop an adequate program to manage and preserve its electronic records. Even records in open formats are difficult to preserve. Electronic records created by state government are also seriously endangered because they increasingly have no paper analog. Electronic records require regular and continued attention to ensure their functionality is maintained as technology continues to progress and change. Ensuring the ongoing preservation of and access to these records is a complicated task, requiring sophisticated technical and professional skills. Failure to manage state government electronic records well will result in their irrevocable loss.

To begin to address these challenges, the Archives must perform significant additional research into best practices nationwide and develop a long-term strategy for preserving New York's permanent electronic records. The state will require significant resources to address this serious need. The federal government, for instance, has budgeted $58 million in this fiscal year alone for the development of a full-scale solution to this problem. The State of New York's costs will be much lower but likely just as costly in relative terms to its budget.

Needs

- Digital records that document state government activities must be captured, preserved, and made available publicly.
- Access to government records is limited or incomplete, and vital records and historical information will be lost if they are not preserved.
- The State Archives does not have the staff to review the large volume of state electronic records, identify those worth preserving, and manage the long-term preservation of these records.

The Archives needs funding to

- Assess the electronic records situation in New York's state government.
- Research and identify best practices to share with state and local governments.
- Develop a strategy for managing state government records in electronic form to ensure their long-term preservation and accessibility.

Benefits

- The state will be able to use the electronic records it needs access to long term.
- The public will continue to have access to essential state and local government records in electronic form.
- Research on the best ways to preserve records created in electronic formats will be shared among state agencies and local governments.

B. State Archives E-Records Transfer Needs

The State Archives seeks to preserve "all books, papers, maps, photographs, or other documentary materials, **regardless of physical form or characteristics**," that were "made or received by any agency of the state or by the legislature or the judiciary in pursuance of law or in connection with the transaction of public business" and that contain information of enduring legal or historical value [New York State Arts and Cultural Affairs Law, Section 57.05; emphasis added].

Records in digital formats present special challenges for long-term preservation:

- Hardware and software quickly become obsolete.
- Electronic media are fragile and have unpredictable life spans.

Electronic records having long-term value thus require careful stewardship. In order to ensure they remain uncorrupted and accessible over time, they must be copied onto new media at regular intervals. In many instances, they must be migrated from their native file format to a newer version of the same format or converted to an entirely new format.

Archivists who work with electronic records perform these actions with great care. They:

- Test various migration and conversion options.
- Select the approach that has the least impact upon the intellectual content and appearance of a given type of records.
- Whenever possible, reduce or eliminate the need for future migrations or conversions (e.g., by converting records in proprietary formats—those owned and controlled by a single entity—to fully documented open formats that can be accessed by existing software or software that the archivists themselves can create).

Records commonly created in office environments—word-processing documents, slideshow-style presentations, and spreadsheets—pose particular challenges:

- The most widely used applications to create these types of documents are owned by private corporations that are not obligated to ensure interoperability or long-term readability of the files that their software creates.
- The files often do not fall under good organizational control and are often stored on local hard disk drives or on network drives with broad access and limited security controls, making it easy to modify, delete or otherwise corrupt the files.
- Staff working collaboratively may repeatedly modify and exchange copies of these files and save the files to multiple locations on shared or individual drives.

- Unless explicitly discouraged from doing so, some employees may save important files on the hard drives of their desktop computers, thus rendering those records inaccessible to others.

The State Archives prefers that records be transferred to the Archives in their native formats whenever possible. For example, if the records are in Microsoft Word 95 format, they should be transferred to the Archives in that format, unless the agency needs to migrate the records to a more current software format for active use. By limiting the number of migrations and maintaining records in their native formats, agencies can help ensure that the Archives can preserve the most complete and useable version of the records.

The State Archives preserves documents, regardless of physical form or format in a way that ensures their integrity and authenticity over time and that also ensures that documents remain readable and useable despite changes in technologies. The Archives holds the protection and maintenance of these valuable government records as its primary mission.

In addition to these general suggestions, there are a number of steps that records creators can take to ensure that their important records are can be preserved and made accessible:

Rely on records management practices and methods to identify records that should be transferred to the State Archives.

Agencies should develop filing systems, naming conventions, document format standards, and transfer plans based on existing Records Disposition Authorizations (RDAs) or the General Records Retention and Disposition Schedule. Doing so will enable agencies to identify records that will be transferred to the State Archives soon after the records are created and ensure that those records are easily moved to the Archives or disposed of appropriately.

Options for meeting this criterion include:

- Working with the State Archives to develop new RDAs as needed.
- Ensuring that employees are familiar with agency records management policies and that they create and save files in accordance with agency and program office protocols.
- Using non-proprietary software to create important records.
- Establishing a filing structure or metadata schema that facilitates periodic transfer of records to the Archives.
- Regularly disposing of records that have satisfied their retention periods.

Maintain records in a secure environment in order to protect their integrity.

Office documents that are to be transferred to the State Archives should not be modified or updated prior to transfer. These documents should be sequestered to ensure that they are protected from inappropriate deletion or modification.

Options for meeting this criterion include:

- Developing a discrete filing structure on an agency network and establishing folder- and file-level access controls.
- Writing files to CD or DVD and physically securing the discs.
- Having in place and monitoring audit trails.

To the extent that it is possible and practical to do so, use non-proprietary file formats and structures to create and maintain records that have long-term value.

Proprietary formats often make it difficult, if not impossible, to access and use records outside of their native software environments. While not all records creators can avoid the use of highly specialized proprietary formats, agencies should consider the long-term usability of certain proprietary formatting features when creating records that have permanent or long-term retention periods. Some features (e.g., tables) do not migrate or translate from one file format to another or from an older version of a given format to a newer version of the same format.

Options for meeting this criterion include:

- Avoiding the use of highly proprietary formatting features in important word-processing documents that have long-term value.
- Removing document passwords or other encryption methods prior to transfer.
- Avoiding the use of compression techniques to reduce the amount of space that files require (e.g. zip).

Maintain the structure, integrity and organization of files and their content.

Records often rely on an organizational structure in order to have meaning and context. Removing files from their organizational structure often makes it difficult to identify versions of a specific file, understand the relationship among files or even identify the creator of a file.

Options for meeting this criterion include:

- Including the directory structure when transferring files to the Archives.
- Transferring files whose contents have not been redacted. The Archives will work with agencies to ensure that sensitive information is protected in an appropriate manner.

- Migrating data to newer software and platforms as necessary. Data loss may occur during migration so it is important that agencies conduct tests that ensure that the changes made do not corrupt the content or structure of office documents prior to committing to a final migration of important records.

C. State Archives E-Records Testimony

Testimony Of Christine W. Ward, State Archivist And Assistant Commissioner for Archives And Records, New York State Education Department, before the New York State Senate Committee On Investigations And Government Operations February 12, 2008

Good afternoon, Chairman Winner and members of the committee. I am Christine Ward, the state archivist and assistant commissioner for archives and records at the New York State Education Department.

Thank you for the opportunity to provide our assessment of issues and needs for electronic records in New York State government. I am neither a technical specialist nor an expert in legal matters, but can offer you a perspective on electronic records and systems from an operations point of view: the management and preservation of records as part of the day-to-day business of government.

The New York State Archives' Responsibility for Records of New York State Government

The New York State Archives was established in law in 1971. We have a broad range of responsibilities for the records of New York's state and local governments and non-profit archives. Today I will focus on our role with respect to state government records, including the records of the Executive and Legislative branches.

The State Archives is responsible for providing guidance and advice on managing the current records of state government. For the last two decades we have been helping agencies manage records in rapidly expanding electronic systems. We operate the State Records Center, where inactive records of state government, including electronic records, can be stored until ready for destruction. We operate the State Archives facility, where state government records of long-term value in a variety of formats, including electronic, are permanently preserved. And, we provide advice, assistance and grants to local governments for managing their current, inactive and archival records, including electronic records, which are a critical challenge for them as well.

"Official records" are defined in the State Archives statute, Arts and Cultural Affairs Law, section 57.05 as those that "include all books, papers, maps, photographs, or other documentary materials, *regardless of physical form or characteristics*, (emphasis added)made or received by any agency of the state or by the legislature or the judiciary in pursuance of law or in connection with the transaction of public business and preserved or appropriate for preservation by that agency or its legitimate successor as evidence of the organization, functions, policies, decisions, procedures, operations, or other activities, or because of the information contained therein." It is the content and context of the record that is important, not the format or media within which the information is contained.

One of the State Archives' key missions is to ensure that New York's government records of permanent value are preserved and made available for future research. This is accomplished chiefly through systematic records management, identifying records with archival value and helping agencies develop schedules for their records that guide their decisions on retention and

legal disposition. Records with long-term value are transferred to the State Archives when they are no longer needed for business purposes. Less than 5 percent of all records created by New York's state government are determined to be archival. Still, this represents a huge amount of material. The State Archives holds more than 200 million paper documents and many more in electronic and machine readable formats.

The Challenge of Preserving Electronic Records

A crisis has been developing over the last couple of decades, generated by the proliferation of records in electronic formats without statewide policy for their creation, management and preservation. Without proper management it is unlikely that important records will be identified and transferred to the State Archives.

Think of all the paper records that we have from past administrations and governmental programs. Before the age of electronic communications everything was written down and therefore much more easily managed. It was simpler to identify what was important to save and what had little value. Today, with cell phones, e-mail, text messaging and other electronic systems, a new generation of technology-savvy users define archiving as saving something for a period measured in months or maybe a few years but certainly not permanently. We are faced with the very real possibility that much of our state's modern history is in danger of being lost. Consider what we know of the actions of former governors and legislators and how we are able to analyze and evaluate governmental programs and policies of the past. This is because we have a written record, much of it preserved in the State Archives. Without full documentation, the complete picture so to speak, historians and policy makers of the future will be working with historical gaps, unable to interpret and learn from the past to create a better future.

The problem is not just e-mail and text messages, but records in all electronic and digital systems: office applications, databases, geographic information systems, old mainframe applications, digital images and websites.

Electronic records suffer from a perception problem. They are too often regarded as tools to get business done efficiently and quickly and not as records that are subject to existing laws and regulations.

Records created in electronic formats are extremely vulnerable. Hardware, software and media can become obsolete with stunning speed. One example is the old 5 ¼" floppy disks used for many years for a variety of office applications. Today, it is difficult to find the equipment to run these disks. Moreover, the software used just for word processing has progressed through many versions in recent years, rendering documents created in the older versions obsolete and sometimes unreadable. Even if we could find hardware with 5 ¼" drives and the right software to run the application, the data stored on these disks may well be corrupted because of inappropriate storage.

Electronic data is prone to loss through human error in design, entry, backup, and storage procedures as well as through deliberate acts of electronic sabotage such as hacking and computer viruses.

Electronic records are especially vulnerable to natural disasters because all too often the organizations that create and maintain them store vast amounts of vital data on single servers

with no backup or backup that is not located a sufficient distance from the original. Our most current, and most used, records are now likely to be in electronic systems that have no paper counterpart. Any situation that impedes access to these records, such as a natural disaster or system malfunction, can seriously impair or prevent the continuing and efficient operation of government.

Electronic records are costly to maintain, even over a short time. They require a commitment of resources that few governments and agencies are able to bear during times of fiscal austerity. We estimate that it would cost about $4 million to design and construct a centralized digital archive to hold the permanent records of state agencies and possibly local governments and at least $1.5 million per year to maintain it. These figures do not include the cost of a building to house it. A digital archive for New York would provide the technological and intellectual infrastructure needed to manage the terabytes of permanent records the state creates in electronic form. (A terabyte is a thousand billion bytes or a thousand gigabytes.) The Archives is currently examining various other models for managing the state's electronic records, including working with other states to build a collaborative repository or contracting with an established digital archives on a fee-for-service basis. We are looking for solutions that both address the needs of the state and are as cost-effective as possible.

Electronic Records Issues Facing New York

The proliferation of state government's electronic records systems without coordination, standardization, oversight or guidance has generated a number of issues that must be addressed.

Citizen access: All citizens have the right to access information created by their government. There is a growing public expectation that records and information will be available on-line with 24/7 access. Many state agencies put a wealth of material on their websites but are unable to ensure that key records are preserved so they will continue to be accessible and usable long-term. For example, when the gubernatorial administration changed last year agency websites, including the departing Governor's, were partly or totally taken down. The State Archives was able to carry out a "web crawl" to capture electronically pages from key sites before they were removed. Had we not done this a great deal of information on policy and programs would have been lost forever. Unfortunately, we hadn't the resources to do more and undoubtedly some important information was lost.

Ability to share information: State government needs to be able to share information across agencies, with the federal and local governments and with citizens. New York lacks unified electronic records policies and initiatives, although many in government recognize the need to combine efforts and share information, knowledge, equipment and other resources. The State Archives is currently working with the State Chief Information Officer and the Office for Technology on several fronts, such as open document formats, to begin to deal with electronic records issues, including making it possible to share information and eliminate redundancies.

Ensuring legal compliance: Each year an increasing number of agencies are required to produce electronic records during the discovery phase of a lawsuit. They are often unprepared because they have not instituted records management systems for their electronic records. It can be extremely expensive to find and produce the data required. If they cannot comply legal sanctions can result.

Managing e-mail: Literally trillions of e-mails are sent and received annually. Because so much business is transacted this way, records in these systems are especially vulnerable to e-discovery. The organization and preservation of e-mail has become a major problem for state and local governments. The State Archives and those in many other states have developed rudimentary guidelines intended to help agencies identify e-mail messages that contain policies, directives and other information that warrants preservation. At present, however, most agency e-mail practices are driven by IT departments struggling to keep their e-mail servers from being overwhelmed, not by the need to preserve information of enduring value to the agencies or to the people of New York. Some agencies place the responsibility for identifying e-mails that warrant preservation on individual users, an approach that almost always proves unworkable. Others may be preserving backup tapes that contain every message that has passed through their e-mail servers, a practice that may lead to staggering expenses associated with e-discovery, not to mention a drain on storage space. Some federal government agencies and agencies in other states are implementing tools that facilitate individual users' efforts to identify and preserve important e-mails or automate at least some aspects of the retention process and then move e-mails slated for preservation from the e-mail server to a separate, secure storage system. New York needs to promulgate policies and practices that follow their lead.

Current Actions

The New York State Archives has been working on electronic records issues for almost two decades, but our resources have been limited. In an attempt to bring greater visibility to the problem and the need, we have identified electronic records as one of four major priority areas for the next five years. We will be taking advantage of every opportunity to inform both the Executive and the Legislature about this issue. The State Education Department and the Regents have asked for $500,000 in 2008/2009 budget to begin an evaluation of e-records issues and status in New York's state and local governments and develop a strategy to deal with the problems.

Pursuant to legislation passed last year, the State Chief Information Officer and Office for Technology, with assistance from the State Archives, will submit a report on electronic records policy in New York to the Executive and the Legislature on April 30, 2008. This document examines productivity software—the word processing, spreadsheet and slideshow-style presentation files found throughout state government. And, it describes the manner in which the file formats used to create and save these materials can limit citizen access to government information, increase the state's information technology costs and impede the State Archives' efforts to preserve electronic archival records. Many of the public's comments about this report recommend that the state weigh interoperability, functionality, accessibility, preservation needs and other concerns when evaluating software options and take measured steps to integrate file format openness into existing state technology procurement processes. File format openness includes the free availability of a format's technical specifications, which allows multiple vendors to develop the means to produce and use files in that format.

The State Archives has developed e-mail management guidelines for state agencies and local governments. These describe how to manage e-mail centrally, identify e-mail that are records and must be preserved for a set amount of time, file e-mail and preserve email to ensure its long-term usability. These guidelines are available on our website. We are also just completing comprehensive email policy which will be issued in the spring.

We are working with the rest of the State Education Department on a pilot project to develop strategies for transferring archival data in its electronic systems to the State Archives. The results of this project should be transferable to electronic records systems created by other agencies.

The importance of partnerships to deal with the problem of electronic records cannot be overstated. The problem is so encompassing and the costs so great that no one part of state government can do it all alone. The State Archives is working closely with other agencies on important electronic records issues. I mentioned our partnership with the State Chief Information Officer and Office for Technology, which will be key as we move forward. Other important partnerships include:

- The Office of Cyber Security and Critical Information Coordination (CSCIC) with whom we collaborate on local government records security issues. We have worked with CSCIC to develop easy-to-use, understandable local government security guidelines that are being distributed nationally.
- The New York State Forum, with which we work closely to develop and deliver electronic records training for state agencies and local governments. They are also partners in the development of statewide e-mail guidelines.

Other States and the Federal Government

Partnerships extend beyond the boundaries of New York. The problem is national and international. We need to partner with other states and the federal government as we move ahead, sharing information and knowledge and working together to develop solutions to mutual problems.

Approximately half the states, including New York, have issued guidelines relating to aspects of e-records management. These vary widely in their focus and detail. Some provide only general guidance on managing electronic records, while others provide specific detail about preservation. The State Archives has developed a number of guidance products, all with a practical focus. We are completing guidelines on acceptable file formats and transfer methods for the electronic records state agencies send to the Archives for preservation.

Five states, not including New York, have intergovernmental electronic records committees that consist of representatives from the state archives, the chief information officer's office, the audit authority, the budget agency, and other key stakeholders. These committees identify and promote electronic records policy, standards and best practices. The workgroup from the State Chief Information Officer's office that is developing the report to the Legislature on electronic records is examining whether New York should follow this approach.

Washington is the only state actively preserving state and local government electronic records of permanent value. It has built a digital archives facility at a cost of $14.5 million, provided $2.5 million in start-up costs and committed to support an annual operating budget of between $800,000 and $1.7 million, supplemented by fees paid by local governments storing e-records there.

The National Archives and Records Administration has awarded a $308 million contract to Lockheed Martin to create an Electronic Records Archive system that will facilitate the transfer of

electronic records from agencies, record information on records retention and preserve and provide access to archival e-records. Expected completion is 2011. The National Archives anticipates that by 2022 it will have preserved 350,000 terabytes of archival electronic records. Since 2001, a portion of the National Archives' budget has been set aside for the Electronic Records Archive system; $58 million is budgeted for the current fiscal year.

The Risk from Doing Nothing

New York could enter a Digital Dark Age. We lost a large piece of our state's history as the result of the 1911 Capitol fire. That loss will seem minor in comparison to the volume of electronic records that could be lost because of technological obsolescence, failure to invest adequate resources to deal with the problem or failure to act at all.

Vital information will be lost. Governments are in danger of losing information and data critical to operations and day-to-day functioning. Planning for the future requires the capacity to learn from the past. If decades of data are lost governments will lose the ability to develop comprehensive and consistent plans for the future.

Citizens' rights will be endangered. Electronic records currently maintained by state and local governments document individual residents' right to vote, property ownership, eligibility for and participation in government services and programs, educational progress and attainment, possession of professional and occupational licenses, and much more.

The cost of government operations will increase. Taxpayers have the right to expect that their tax dollars are being used wisely. Loss of records or the inability to find records can result in needless expense. There is a high cost associated with recovering or recreating data that is corrupted or lost and with duplicating effort because of an inability to share information across agencies and among state, federal and local governments. For instance, in complicated electronic systems such as geographic information systems, known as GIS, and computer-assisted design applications, the software and file formats are proprietary and owned and controlled by individual companies. It is often difficult or impossible to access files created with an older version of an application using a newer version of the same program. As a result, agencies that rely heavily on GIS systems to track data and analyze changes in zoning and land use over time will have to devote substantial time and effort to making sure that older data remains accessible and useful. In some instances, agencies may have to reconstruct older data simply when changing from one version of the software to another.

What New York Needs

In order to deal with the problems that I have identified, we need to focus our efforts in several specific areas, including carrying out a full evaluation of the status of electronic records in state and local government and developing a set of strategies to meet the growing needs. We also need the following:

1. Stronger electronic records laws. Expand the definition of records to encompass those created in electronic format. Following the federal model, define the records of the state's chief executive as belonging to the people of New York. Ensure the transfer of governors' records to the State Archives at the end of each administration. Give the State Archives the authority to intervene directly when state government electronic records are at risk. Provide the resources

necessary to manage and ensure the preservation of the electronic records of New York's state government.

The State Archives supports adding more precise clarification, greater consistency and specificity relating to modern record formats in the various statutory definitions of records and official records. The law also should clarify what constitutes public records and private records. The law should contain more consistency of language across all statutes that deal with records. The State Archives recommends that the law clarify public ownership of the official records of the governor's office and require that records that document the official actions of the governor and staff and are of long-term value to the people of New York be transferred to the State Archives for permanent preservation.

Build strong partnerships. The challenge is so large that no one agency or entity can do it alone. We need partnerships among state agencies whose missions relate to records and information and among implementers of new technology in state agencies and local governments to reduce costs, complexity and redundancy.

Develop best practices and models. Implement statewide policies for managing electronic records, including e-mail. Encourage state contracts for electronic records solutions, such as open software products, and data recovery and electronic vaulting services. Develop model strategies for evaluating electronic records systems in state government and transferring those with long-term value to the State Archives.

Conclusion

The challenge is great and the stakes are high. It is imperative that we act now to ensure that essential information for the state's fiscal, legal and administrative continuity survives and that New York's historical record and legacy are preserved for posterity.

D. State Archives Attachment to Testimony

States Electronic Records Programs: Attachment to Testimony of Christine W. Ward, State Archivist and Assistant Commissioner for Archives and Records, New York State Education Department, before the Senate Standing Committee on Investigations and Government Operations February 12, 2008

States with Electronic Records Staff

Although 31 state archives and records management programs reported having staff in place to manage electronic records in FY 2006,[1] the level of programming varies widely from state to state. Some states (such as Washington) have the technological and administrative infrastructure needed to acquire, preserve, and provide access to electronic records in a variety of formats. (Washington State has constructed an entire separate facility to manage the preservation of electronic records of state and local government.) Other states are in the process of developing this infrastructure. These include Arizona, Florida, Kentucky, Wisconsin, and even New York. Finally, the remaining states (including Connecticut, Indiana, and Ohio) simply take in records from state agencies in the hopes that they will eventually develop the capacity to care for these records appropriately. The State of Wisconsin did not report to the Council of State Archivist in 2006 that it was employing staff to work on electronic records, but they currently have such staff.

Alabama
Arizona
California
Delaware
Florida
Georgia
Hawaii
Kansas
Kentucky
Louisiana
Maine
Maryland
Michigan
Minnesota
Mississippi
Missouri
Nebraska
New Hampshire
New Jersey
New Mexico
New York
North Carolina
North Dakota
Pennsylvania
South Carolina
South Dakota
Tennessee
Utah
Virginia
Washington
Wyoming

Electronic Records Held by State Archives

As of FY 2006, the state archives below reported to the Council of State Archivists that they had the following records in their collections. The quantity and type of electronic records held by these state archives ranges dramatically. Note that this list includes three states (Connecticut, Indiana, and Ohio) that reported having no staff to manage electronic records and that it does not include eight of the states reporting that they did have such staff.

Alabama

12 gold compact discs of state agency websites and 4 full sets of websites for 120 state agency, college, and university websites gathered and stored by the Internet Archive.

[1] Council of State Archivists (CoSA), *The State of State Records: A Status Report on State Archives and Records Management Programs in the United States*, Iowa City, Ia., 2007. http://statearchivists.org/reports/2007-ARMreport/StateARMs-2006rpt-final.pdf

Arizona
Currently, harvesting webpages in concert with the Law and Research Library Division that are stored on DAT tapes with limited access, 900 GBs to date. Also share custody with Library of 1,736 born-digital state reports hosted on and stored in CONTENTdm. Archives holds several hundred born-digital photographs and several electronic collections as well.

California
Less than 25 cubic feet.

Connecticut
310 VHS videotapes, 55 audiotapes, 8 microcassettes, 2 DVDs

Delaware
Audio, video, phonograph, and images; total of 736 cubic feet stored in archival containers, maintained and remastered as needed. Less than 1 cubic foot of electronic data (unknown MB); primarily state publications plus a smattering of agency records.

Florida
6,155 megabytes of Department of State executive offices, including email, Word documents, etc., plus miscellaneous disks in a few other collections.

Georgia
912 compact discs. First records now being added to Digital Archives servers.

Indiana
Miscellaneous records on compact discs, i.e., governor's email, state website backups, etc.

Kansas
The KHS has ingested 87 items into the Kansas State Publications Archival Collection (kspace.org) pilot digital repository. Most of these items are annual reports and special reports to the legislature that, in a Web-based environment, represent a gray area between state government publications and state government records.

Kentucky
70,000 files, or approximately 48 gigabytes.

Louisiana
Less than 50 kilobytes.

Maine
10 geographic information systems data layers.

Maryland
Amount of electronic data imaged: 65,000 gigabytes. Number of database records managed: 6,191,225

Michigan
Under 200 compact discs.

Minnesota
Approximately 2 terabytes.

Mississippi
404 gigabytes.

New York
Seven series of archival datasets and 15 gigabytes of electronic records commingled with paper records.

North Carolina
Governor's Office correspondence (6.38 gigabytes).

North Dakota
Minor amount of data; volume is in video and sound.

Ohio
441 floppy disks; 64 zip disks; 99 compact discs.

Pennsylvania
Approximately 1,000 compact discs/DVDs which represent use copies of scanned archival records.

South Carolina
10,000 audio tapes; 50 video tapes and compact discs; 3 databases.

South Dakota
2 cubic feet.

Virginia
154 megabytes; 706 compact discs; 32 websites.

Washington
7 terabytes.

Wyoming
110 compact discs; 214 optical disks; 55 IBM cartridges.

State Archives that Held No Electronic Records (as of FY 2006)

Arkansas
Colorado
District of Columbia
Hawaii
Idaho
Illinois
Iowa
Missouri
Montana
Nebraska (rec'd records 2007)
Nevada
New Hampshire
New Jersey
New Mexico
Oklahoma
Tennessee
Texas
Vermont
West Virginia

State Archives that Did Not Indicate Whether They Held Any Electronic Records (as of FY 2006)

Alaska
Massachusetts
Oregon
Rhode Island
Wisconsin

State Archives that Had Issued E-mail Management Guidelines (as of FY 2006)

Forty state archives have released email guidelines for use by state and local government. Each of these policies provides three basic types of guidance: They differentiate between e-mails that are records (those messages sent and received by government employees using government resources in the course of performing government business) and those e-mails that are not records (including listserv messages and announcements of staff events). These policies state that record e-mail must be managed in accordance with existing records management statutes and regulations. Finally, these emphasize individual employees' responsibility for managing record e-mail and the importance of educating staff about records management, existing records retention schedules, and proper management of record e-mail. Other states either have no formal e-mail policies or their policies focus entirely on the narrow issue of the acceptable use of e-mail by government employees.

Alabama
Alaska
Arizona
Arkansas
California
Colorado
Connecticut
Delaware
Florida
Georgia
Hawaii
Idaho
Indiana
Iowa
Kansas
Kentucky
Louisiana
Maine
Massachusetts
Michigan
Minnesota
Mississippi
Missouri
Montana
Nebraska
Nevada
New Jersey
New York
North Carolina
Ohio
Oregon
Pennsylvania
South Carolina
Texas
Vermont
Virginia
Washington
West Virginia
Wyoming

E-mail Policy at the National Archives and Records Administration

In January 1989, on the last day of the Reagan administration, the National Security Archive sued the Executive Office of the President (EOP) to block the destruction of several years' worth of Reagan White House emails. The case became known as *Armstrong v. EOP* (or just "*Armstrong*"). *Armstrong* dragged out in the courts through the entire Bush I administration and was finally decided in 1993 in favor of the plaintiffs. This landmark case led to sweeping changes in federal regulations regarding retention of email messages by federal agencies.

The most important rule change that followed *Armstrong* requires that the recordkeeping copy of electronic mail messages deemed to be Federal records be moved to a true archival system unless the electronic mail system itself meets several minimum criteria. The system must be able to

- allow for "grouping of related records into classifications according to the nature of the business purposes the records serve"
- permit "easy and timely retrieval of both individual records and files or other groupings or related records"
- Retain "the records in a usable format for their required retention period." In other words, the system must truly preserve the records, not just warehouse them. The corollary implication is that records managers have to budget for the long-term cost of preservation.

- make electronic mail messages "accessible by individuals who have a business need for information in the system." This means that "secret" e-mail records files are not permitted.
- preserve "the transmission and receipt data," meaning that basic metadata (such as date, sender, and recipient) must accompany each message
- "agencies must permit transfer of permanent records" to the National Archives and Records Administration

The National Archives has yet to issue a set of guidelines to federal agencies to help them comply with law and regulations, and recent evidence suggests that compliance with regulations is being resisted.

Appendix VIII

ANALYSIS OF OPEN SOURCE AND NYS GOVERNMENT

[Before New York State Technology Law § 305(4) (requiring this study) was enacted, business units at the New York State Office of the Chief Information Officer/Office for Technology (CIO/OFT) had already started an analysis of open source software and its place within the State's information technology (IT) enterprise systems. That nascent study was adapted for inclusion in this report, as follows]

The widespread implementation of open source software presents New York State with an unprecedented opportunity to think about the way it manages information technology. This may present opportunities on transformation allowing the State to utilize technology in ways that are more flexible, responsive to business needs and more cost effective.

According to Gartner, by 2008 open source solutions will directly compete with intellectual property (closed-source) products in all infrastructure markets. By 2010, 10-20% of all application software used by government agencies will either be reused from other agencies or open source.[102]

This is a high level overview that will:

- Define and provide examples of what open source software is and how other organizations use it;
- Identify the benefits and risks that open source adoption would bring to New York State; and
- Recommend measures the State could take to begin adopting open source.

Open Source Defined

Open source software (OSS) is software which unlike traditional proprietary software allows anyone to use, read, redistribute and modify the source code without a royalty or other fees. The Open Source Initiative defines ten features which distinguish OSS from other licensing models:[103]

1. Free Redistribution: The software can be transferred to other entities without charge.
2. Source Code: All code needed to compile or build the software must be included or made available to any interested party at a reasonable cost.
3. Derived Works: Code may be altered or embedded within another software package under the same license conditions as the original work.
4. Integrity of the author's code: Derived works may not interfere with or distort the original author's work. Author may require derived works carry a different name from the original software.
5. No discrimination against persons or groups.
6. No discrimination against fields of endeavor: Distributed software cannot place restrictions based on user's intent (i.e. "academic use only" or "non-commercial use only" clauses are not acceptable).
7. Distribution of license: The rights of the program must apply to all whom the program is re-distributed to without need for an additional license.
8. License must not be specific to a product: The openness of a software package may not be tied to using another specific software package.

9. License must not contaminate other software.
10. License must be technology neutral.

OSS has become increasingly popular over the years within the technical and academic communities that created the Internet as we know it today. Critical technologies we use every day like web browsers, DNS, email and TCP/IP are built upon or popularized by OSS.

In today's enterprise IT environment, open source software has moved into the software mainstream. The most popular platform for the Oracle Database Server is Linux and the Tomcat Application Server is the reference implementation of a J2EE application server. Open source has made a big impact in software-dependent industries like financial services and telecommunications.

In the personal computing arena, desktop applications like OpenOffice and Mozilla Firefox are becoming increasingly popular due to their low cost, advanced functionality and perceived level of security. Apple Computer has made hundreds of open source products an integral part of its OS/X operating system, which is built upon the open source Mach microkernel.

Established Internet firms like Yahoo, Google and Amazon take advantage of OSS throughout their businesses. Yahoo replaced internal applications based on C++ and in-house languages with the open source PHP language[104]. Google runs their business on a customized Linux and the open source Python programming language. Amazon is a big user of Linux, Perl, the Mason template engine and MySQL.[105] These firms see open source as a way to accelerate development, cut costs and compete.

The Federal Government is an active open source user and collaborates on dozens of products as well. According to Forbes, "Open Source is critical to DOD's central nervous system … if open source were banned from the DOD, costs would spike as capability and security dropped."[106] Some Federal agencies actively collaborate on open source projects as well. The Lustre file system, a secure system designed for use with computing clusters, was started as a joint venture between Hewlett Packard and the National Nuclear Security Administration.

Benefits of OSS for NYS

Strategic Factors

According to an IDC Study[107], companies with high rates of open-source adoption tend to be in "industries [i.e. telcos, Internet firms, financial and business services] that perceive software as the most important to their ability to compete." According to the study, U.S. firms tend to report cost as the most important factor for using open source, while European firms focus on flexibility and software quality.

These industries share a desire to cut operating costs by replacing expensive proprietary solutions with less expensive alternatives. Google is an extreme case, operating clusters of hundreds of thousands of servers in dozens of data centers. A more typical case includes financial services firms like Euronext NV, which say: "In a nutshell, performance, cost and scalability were the three most important factors … We're getting 30 times more bang for our buck with moving over to Linux."[108]

Cost is only one factor for choosing open source. IDG survey respondents and Google indicate quality and flexibility are the most important motivators for adopting OSS. Some firms enjoy the flexibility cost-free licensing provides for scalability or disaster-recovery scenarios, or use that flexibility to keep the details of their computer infrastructure secret and gain a competitive edge over rivals. Chris DiBona, Open Source Program Manager at Google, explains:

> *"If we had to go and buy software licenses ... people would absolutely know what the Google infrastructure looks like ... the use of open source software, that's one more way we can control our destiny."*[109]

The ability to customize is another major factor for OSS adopters. Instead of independently developing applications, companies can take an existing, proven application and add features to it. For example, IBM[110] and the NSA[111] have added major features to Linux to meet their own needs, which were subsequently incorporated into the next version of the Linux kernel.

Exploit the Global Knowledge Base

Students in technical majors today use open source software every day. Many universities base their core Computer Science curriculum around languages like C and Java, and teach using open source development environments. Source code, Google and the open source community are the primary reference tools for today's Computer Science student.

Purchasing proprietary software represents an onerous cost for educational institutions and students. Computer Science departments discovered open source years ago. Now mainstream educators see open source as a way to get more technology in front of more students at a lower cost.

The State of Indiana's ACCESS (Affordable Classroom Computers for Every Secondary Student) program has provided over 80,000 students in 80 high schools with Linux workstations. The mission of the ACCESS program is to provide Indiana's students with access to technology to enhance the learning experience at a cost school districts and the State can afford to bear.[112] The program is considered a success and will result in a generation of children familiar with Linux and using Internet tools to collaborate and learn.

The reaction of students to the new Linux environment is evidence of the strides Linux Desktop operating systems have made:

> Huffman [an Indiana technology coordinator] is eager to get a read on student acceptance of Linux. In surveying one classroom last year, he asked a student what he thought of using a Linux desktop vs. a Windows desktop, and the student responded, *"Who cares?"*[113]

The success of Indiana's program has attracted the attention of other education departments as well. In New York, the Madison-Oneida Regional Information Center is considering the use of Linux desktops to provide computing services to over 80,000 students in the Rome City School District and other districts supported by the center.[114]

The widespread adoption of the open source philosophy in education is evident in the technology choices made by cutting edge technology and startup firms, which tend to be founded by students and recent graduates. Successful startups like Flickr (acquired by Yahoo), Splunk and established

Internet companies like Amazon and Google all build their core service offerings around an open source foundation.

Cost Containment

Depending on how widely and deeply open source solutions are deployed, cost savings can be significant.

In the case of Indiana's ACCESS program, by adopting open source operating systems like Novell's SLED[115], Sun's OpenOffice and other open source packages Indiana reduced the total cost of software to five ($5.00) dollars.[116]

Total cost for a workstation and monitor range from $250 - $300. To put that in perspective, Microsoft licensing for one NYS agency includes over $325 in license subscription costs over the 3-year life of a workstation.

For New York State, the potential for savings could be substantial:

- On the desktop, proprietary office suites deployed on Human Services Enterprise Network workstations represent an annual subscription cost of approximately $2.1 million.
- On the server-side, departmental-scale databases currently depending on $10,000+ proprietary SQL Server licenses could be deployed with free, open source databases like MySQL or PostgreSQL.

The real cost savings derive from strategically shifting how New York State manages technology. Currently, server and network infrastructure is purchased and deployed to meet the needs of a particular project. Decoupling the procurement of hardware, operating system and application software encourages the centralization and sharing of services that benefit from the economies of scale, which aggregated procurement and administration bring.

For example:

- Virtualization and standardized server configurations allow the State Data Center to control overhead costs. Staff can plan and think in the long term instead of reacting to immediate needs.
- A standardized environment means fewer system administrators can manage more servers and respond to customers better.
- Application owners benefit from faster server provisioning and lower overall costs.

Open Source is more than just Linux!

Linux is one of the best know open source projects – but there are thousands of other OSS projects. From Java Application Servers to Development environments to core Internet infrastructure like DNS and email, there is an open source project to address most needs. Prominent open source projects are listed below.

What measures might New York State take to Implement Open Source Software?

Focus on Staff Development and Education

Adopting open source software as a replacement for existing proprietary platforms represents a major change for employees who have spent years becoming expert at operating and designing solutions using proprietary software.

A two-phased approach should be taken to address the training of end-users and education of IT staff:

1 - Educate and Train Staff in Open Source Technologies

For end users, reference material in the form of websites, quick reference cards and training classes need to be delivered to ease the transition from proprietary applications. Open source desktop transitions will likely introduce the most need for end-user training, while changes to infrastructure will often be invisible to the end user.

For IT staff, a more comprehensive program of training is required. Easy access to reference materials like books, online courses, and access to formal courses at colleges and universities would facilitate the transition of experts in proprietary applications to a similar level of expertise using new, open source applications. Additionally, "seeding" IT departments with subject matter experts in appropriate open source applications and technology will increase the effectiveness of training and reduce the resistance to change.

Open source is more than an application migration or upgrade for IT professionals. It represents a sea change in thought process and a new, unfamiliar landscape that may be interpreted as a threat. It is critical any migration be carefully managed and be presented as an opportunity for career development and not a threat.

2 - Identify Potential Partnerships with Higher Education

Students and recent graduates are often some of the most qualified technical resources for implementing and developing open source technologies. Many of the individual contributors to open source projects are students, and most Computer Science programs make extensive use of open source projects as teaching and learning tools.

New York State should work to establish partnerships with the State University system to gain access to the enormous potential of the students. Using skilled student interns is a cost-effective way to seed State IT departments with open source experts and recruit exceptional students for employment after graduation. This will allow students to work with experienced IT workers and gain "real world" knowledge through these interactions.

The State should also explore integrating complex State IT problems into the university curriculum. MIS and Management students would value the opportunity to observe and assist program and project managers; Computer Science and IT students would benefit from programming and other IT roles.

3 – Support Partnerships

State agencies must establish partnerships with vendors who have established open source consulting, training and support practices. Companies such as IBM, Hewlett Packard, Sun, Red Hat, Novell, MySQL and Canonical all have OSS support offerings that can bridge the skill gap and allow employees to succeed with unfamiliar software.

Identify Potential "Quick Win" Opportunities and Establish Pilots

Agencies can launch efforts to identify systems running on proprietary software platforms where potential "drop-in" solutions exist. Replacing proprietary office suites with open source office suites would be a perfect example of a "quick win" in many circumstances.

Once candidates are identified, proof of concept and pilot projects should be established to fully evaluate open source solutions in an operational setting. CIO/OFT Customer Networking Solutions (CNS) has been conducting an open source desktop pilot in FY07-08.

Identify Current Software Implementations

All State agencies use a wide variety of software to accomplish their missions, all of which have associated costs. Agencies should review all software coming up for renewal and evaluate open source alternatives.
Common functions across agencies, such as case management and customer relations management should be identified as candidates to be re-usable application components, where appropriate.

Determine If Procurements Can Be Open Source Friendly

Open source solutions may not fit well with formal procurement processes, as there usually is no vendor (nor partnership with hardware or consulting vendors) marketing the product.

Agencies should consider if open source solutions are a good fit for the program or service area under consideration. This analysis would occur early in the evaluation process, based upon project requirements, and prior to starting the procurement process. The resulting design and procurement record could lay-out open source options that address the agencies needs. This would then allow for a subsequent evaluation of the responses.

Use of Open Standards for Enterprise Solutions

In some situations, open source may not be appropriate to the task. When adopting open source solutions is not feasible, proprietary packages should adhere to open standards that permit full interoperability. Organizations should avoid solutions that "lock-in" other proprietary products and methods.

Shared Services and Open Source

For decades organizations have sought to reuse software components but have been stymied by three major stumbling blocks, which Gartner defines as:[117]

1. Interconnecting and adapting disparate components: Web service standards like SOAP and XML-RPC have made this hurdle less of a problem;

2. Maintaining and licensing components to ensure reliability and sustainability over time: open source development and licensing offers a viable approach; and
3. Overcoming issues between potential users of reusable application components in competitive industry sectors: This problem remains and tends to limit reuse to lower levels of the software stack (i.e. Operating systems, J2EE Servers, Databases, etc.), although reuse is possible where competition is not an issue such as in the public sector.

Gartner defines Shared Services as "…*the aggregated provision of services between multiple, largely autonomous entities. Shared Services aim to achieve benefits by using a single group to provide a service to multiple agencies or units, rather than each agency requiring its own capacity to provide that service.*"[118]

Many believe Shared Services are an idea whose time has come. On the public Internet, Amazon.com has been a shared service pioneer, offering web service interfaces access to Amazon's vast catalog via Real Simple Syndication (RSS), Representative State Transfer (REST) and Simple Object Access Protocol (SOAP) interfaces for several years. Today, Amazon has branched out into on-demand storage, queuing services and virtual servers that can be provisioned in minutes[119] and are billed on a "pay as you use" basis (similar to an electric or gas utility).

Open source software makes the economics of shared services work. Removing the marginal cost of licensing additional instances of database servers, application servers and operating systems creates an environment where IT infrastructure has the flexibility to add and redeploy resources as demand dictates. Legacy infrastructure can be integrated with shared service platforms by using Web Services technologies like SOAP and XML-RPC, which are freely available for most environments via open source packages.

Develop a Centralized State Code Library

Most state government entities create code for applications, websites and to assist with managing IT resources. In the State of New York, this code is the property of the State and can and should be shared across organizational boundaries.

Adopting the spirit of open source development in addition to making use of the benefits of open source could further reduce overall costs and encourage the delivery of higher quality code. As part of its top to bottom analysis of State government operations, the California Performance Review[120] (CPR) discovered a significant amount of programmer time was being spent "reinventing the wheel" because agencies and even divisions within agencies do not have any mechanism to collaborate. Date calculation routines are a perfect example of a relatively simple problem that is re-implemented hundreds or thousands of times at great cost.

The CPR believes that in California's case, the use of a central code library has the potential to significantly reduce contract programmer expenses. This centralized, easily accessible code library would also ease open source migration difficulties by increasing collaboration between agencies and divisions within agencies. Programmers and IT staff making the transition from proprietary to open source would have a place to start to locate code to solve problems and create new applications.

Most importantly, a well used code library has the potential to recreate the peer-review process that encourages the high quality of open source software within State government.

Looking Ahead: Choosing Open Source for New York State

Goals

A systematic adoption of open source software within New York State government should be centered on three key goals, to:

- Save taxpayer money;
- Improve the operational efficiency of State IT operations; and
- Increase the operational agility of State IT initiatives.

The Decision Matrix: Putting Open Source on the Table

Open source software does not have marketing and sales forces actively promoting products to business and technology leaders. The "open" nature of open source can be disconcerting to government IT managers, many of whom equate open source applications to "shareware."

These concerns are not without merit -- some OSS projects are less mature than others or don't have a large active user base to grow from. To address these concerns, consultancies such as Cap Gemini[121] and Navica[122] have developed formal methodologies for measuring the maturity of OSS.

For example, the Cap Gemini Open Source Maturity Model provides a systematic way to measure four key product indicator groups:[123]

- Product Internals;
- Integration with other products or infrastructure;
- Use - How a user is supported in the day to day operation of the product; and
- Acceptance - How the user and developer community has received the product.

Using established models like OSMM to pre-screen open source software will ensure solutions evaluated by the State will be of the highest quality available.

The Evaluation and Procurement Process

To encourage the adoption of OSS for enterprise software procurements, the information technology evaluation and procurement processes should encourage the evaluation of open source solutions or components. This encouragement can be achieved both through mandates and building infrastructure reflecting the low cost and agile deployment capability that OSS provides.

Suggested modifications to technology planning and procurement processes might include:

- Providing incentives or preferences in the bidding process to vendors who propose the implementation of open source solutions or components;
- Mandating that agencies considering technology deployments over a specific dollar threshold evaluate open source solutions. (For example, if an agency evaluates three proprietary products, at least one open source product must be evaluated as well);

- If no open source solution is available or viable, evaluating open source components for a proprietary system (Example: If a proprietary business system is under consideration, evaluate an open source database or operating system); and
- Offering a "fast track" server provisioning service. The unique economics of open source may make it feasible to offer pre-provisioned, rapidly deployed virtual or physical servers in State datacenters.

The purpose of these modifications is to get open source on the table and evaluated equally beside proprietary solutions. Functionality while reducing costs must continue to drive IT procurements.

Sample Decision Matrix

Table 1: Sample Open Source vs. Proprietary Decision Matrix

Attribute	Weight	Solution 1: Proprietary Solution	Solution 2: Proprietary Solution	Solution 3: Open Source Solution
Solution meets or exceeds technical requirements				
Support Availability from Vendor				
Support Availability from internal resources				
Solution interoperates with existing infrastructure				
Solution adheres to relevant open standards				
Solution utilizes open source technology				
Initial costs for initial software licensing & support				
Initial costs for hardware and/or physical infrastructure				
Initial costs for training				
Ongoing support costs				
Ongoing license renewal or other costs				

Attribute	Weight	Solution 1: Proprietary Solution	Solution 2: Proprietary Solution	Solution 3: Open Source Solution
Maturity/market position of solution				
Anticipated longevity of solution				
TOTALS				

Examples of Open Source Software

This table describes some of the major open source packages on the market and identifies the category of organization governing the product. Three high level categories are defined:

- Foundations: Typically not for profit corporations organized to hold intellectual property associated with the OSS project;
- Communities: Ad hoc organizations of contributors, including corporate contributors and individual actors; and
- Commercial: Some corporations directly sponsor open source products or release the source of proprietary products. Some projects, such as Mozilla, OpenSolaris and Eclipse are later "spun off" into independent foundations.

Table 2: Organizations producing prominent OSS packages[124]

Project / Project Sponsor	Category	Role
Apache Software Foundation	Foundation	Sponsors a broad range of open source projects such as Tomcat, Velocity and the Apache Web Server.
BSD	Community	Produces various versions of the Berkeley System Distribution, an open source Unix operating system.
Drools	Community	An object-oriented rules engine for Java, widely used in the healthcare and financial community.

Project / Project Sponsor	Category	Role
Eclipse Foundation	Foundation	An open platform for software development; one of the leading Java IDEs (Integrated Development Environments).
GNU Project	Community	Develops tools and utilities that are essential parts of any Unix or Linux system. One of the pioneers of the modern open source movement.
Groovy	Community	A dynamic scripting language for the JVM (Java Virtual Machine).
Jabber Software Foundation	Foundation	Jabber is an open source instant messaging platform. It is the foundation of IM (Instant Messaging) offerings from Google and Sun.
JBoss	Commercial	A for-profit subsidiary of Red Hat that develops the JBoss Java application server and associated software.
Mozilla Foundation	Foundation	Mozilla is an open source browser technology distributed as Firefox. Other projects include Thunderbird, which is a POP and IMAP email client and Sunbird, a calendaring package.
MySQL	Commercial Open Source	A for-profit company that develops the MySQL database.
Novell	Commercial Open Source	A for-profit company that distributes Suse Linux.
ObjectWeb	Consortia	An open source community created to foster development of open source distributed middleware.
Open Source Development Labs (OSDL)	Consortia	The home of Linux kernel development led by Linus Torvalds.
Perl Foundation	Foundation	Develops the Perl programming language.
Python Foundation	Foundation	Develops the development of the Python programming language.
Ruby on Rails	Community	An accessible, easy to use rapid web development framework.
Spring	Community	A Java application framework for Java/J2EE applications.
SugarCRM	Commercial Open Source	A for-profit company that distributes SugarCRM, an open source customer relationship management package.

Project / Project Sponsor	Category	Role
Sun Microsystems	Commercial Open Source	A for-profit company that distributes a number of open source products like OpenSolaris, OpenOffice, NetBeans and the ZFS filesystem.
Yahoo	Commercial Open Source	A for-profit company that distributes several open source resources for web developers such as the Yahoo UI Library and Yahoo Design Pattern Library.

Cost Illustration

Figure 1: Current 3-year Cost Breakdown for a Human Services Enterprise Network (HSEN) NYS agency PC

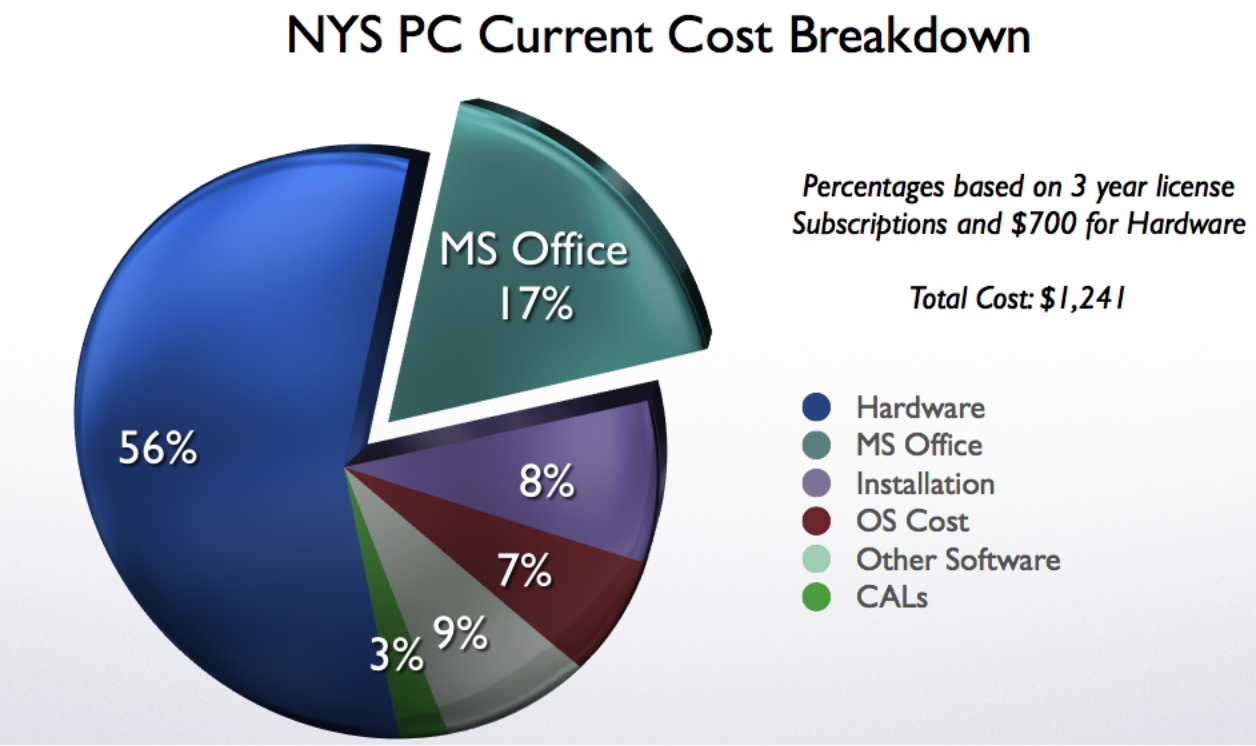
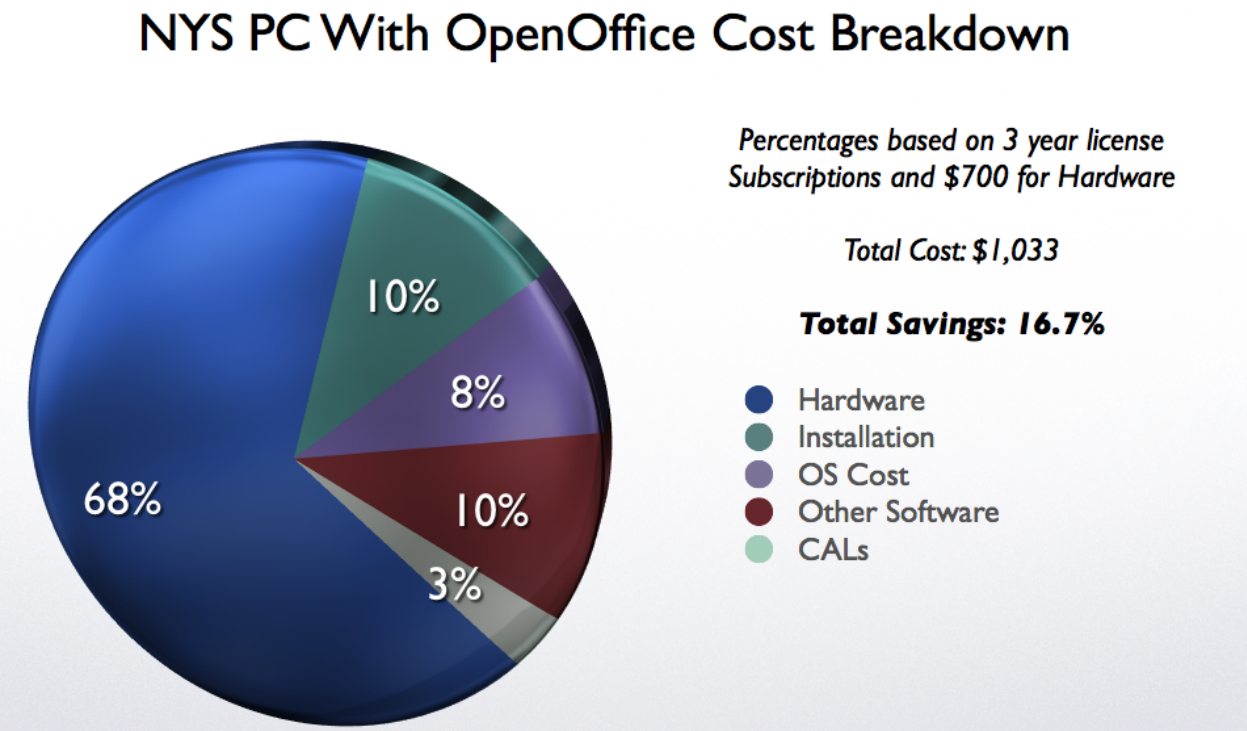


Figure 2:
3-year Cost Breakdown for an HSEN agency PC using OpenOffice



State CIO Survey

A survey conducted during the 2005 National Association of State CIOs Conference provides some great insight into how other state governments see the role of open source in the enterprise.

The results of this survey indicate that:

- State governments adopt open source because it is cheaper and works better;
- Most states are in the early phases of open source adoption;
- CIOs are not yet comfortable with the open source support model or the availability of qualified staff; and
- State CIOs are overwhelmingly happy with the open source deployments they have conducted.

Figure 3: Open Source employment[125]

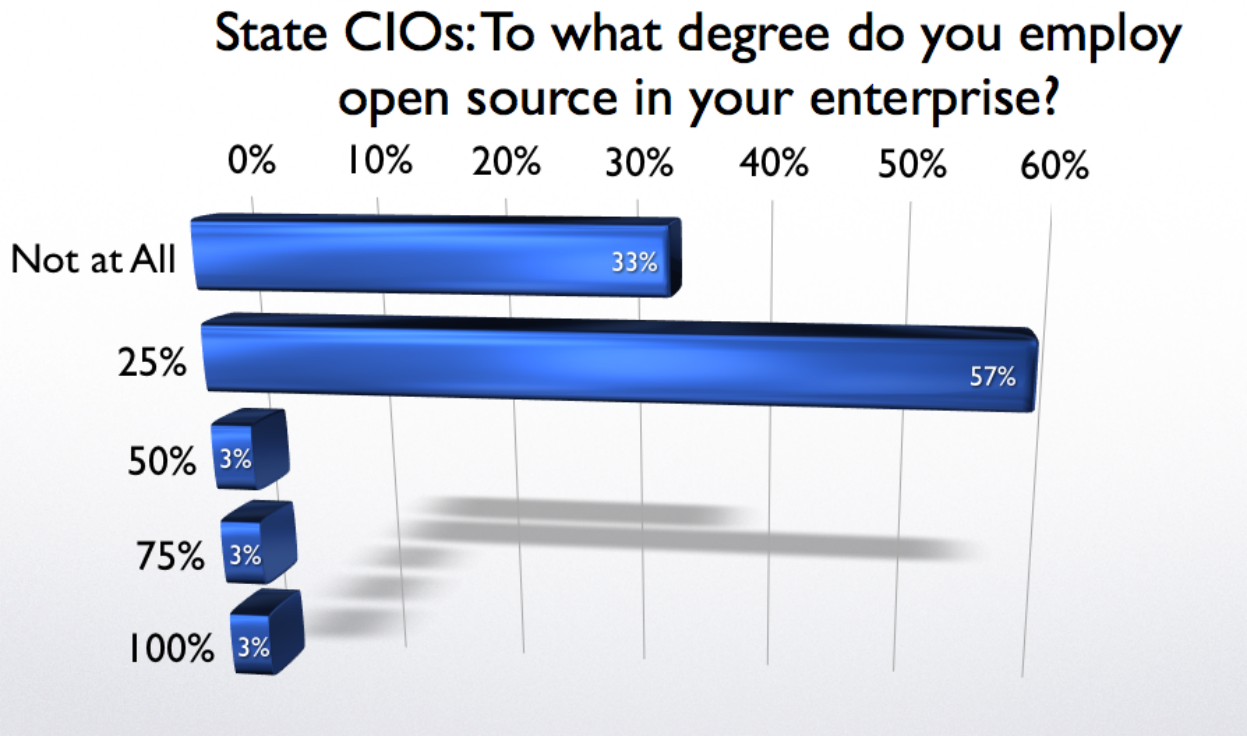


Figure 4: Why did you choose OSS?[126]

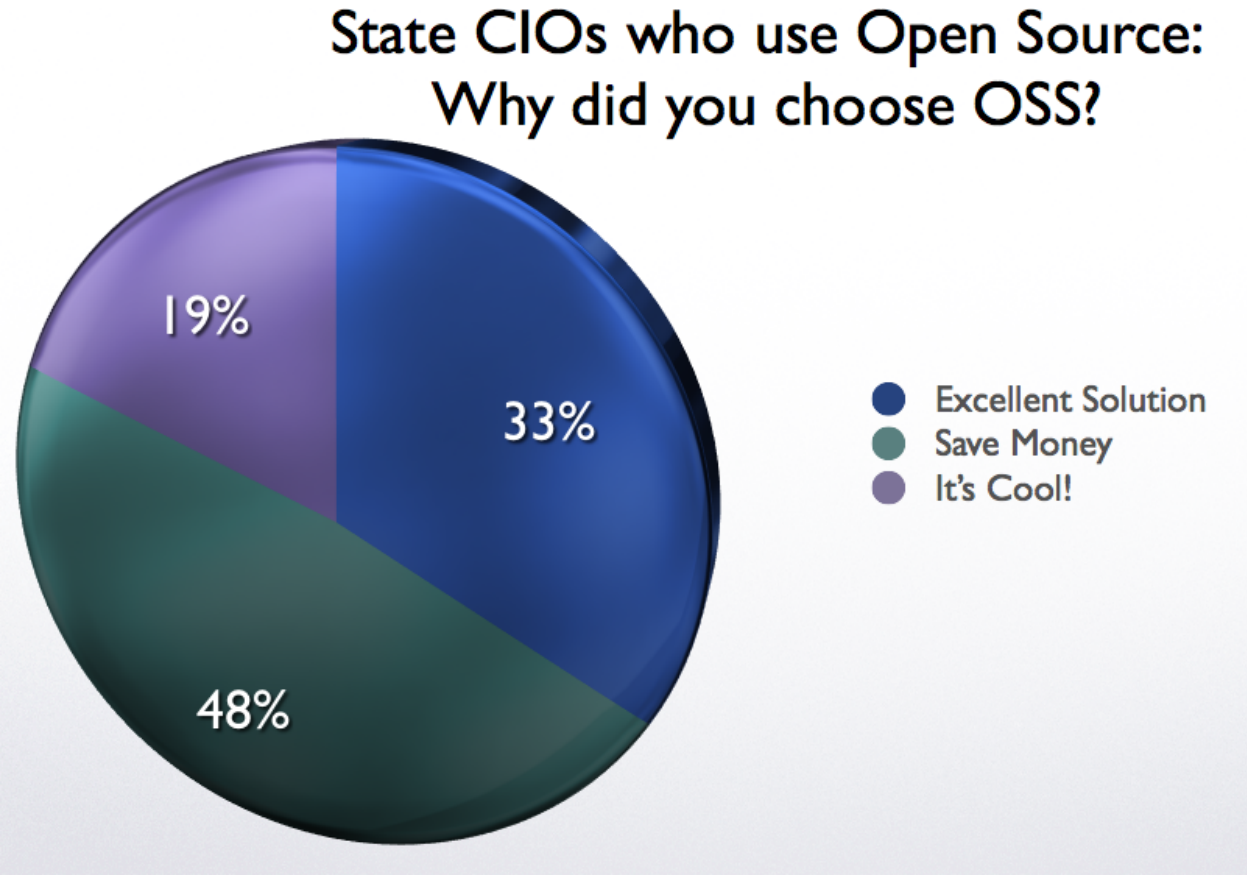


Figure 5: OSS Concerns[127]

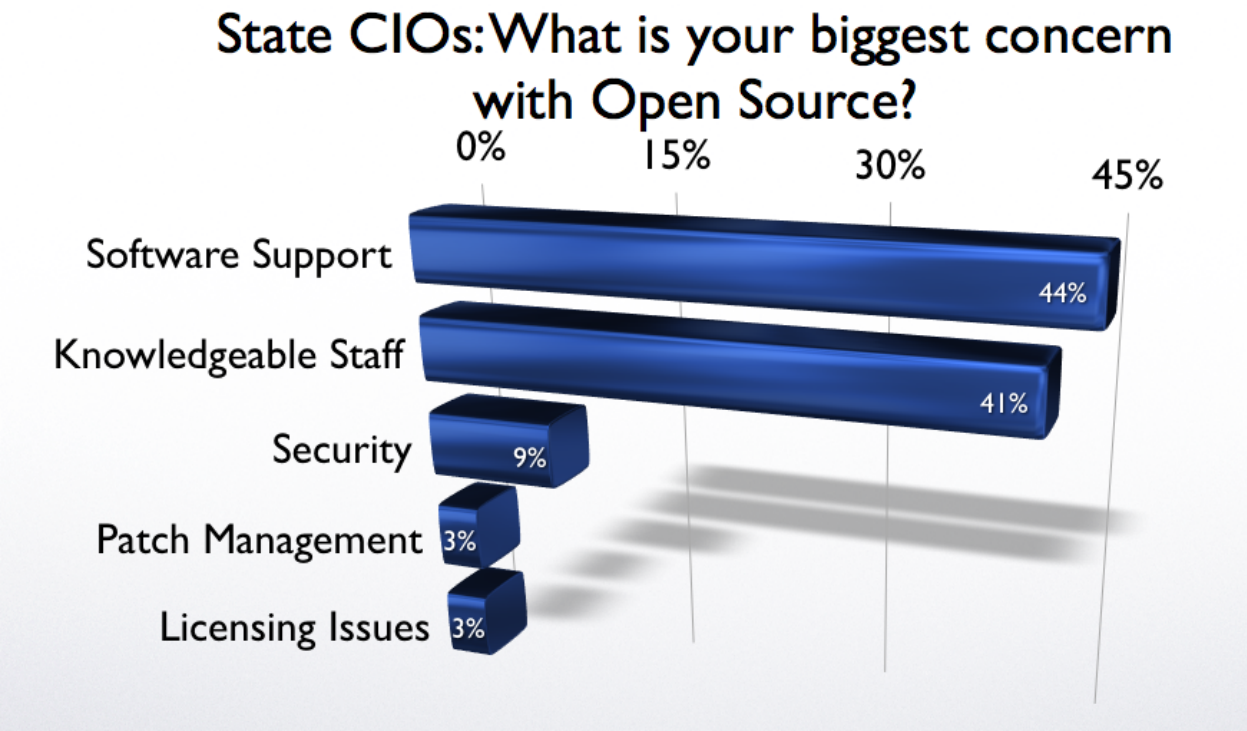
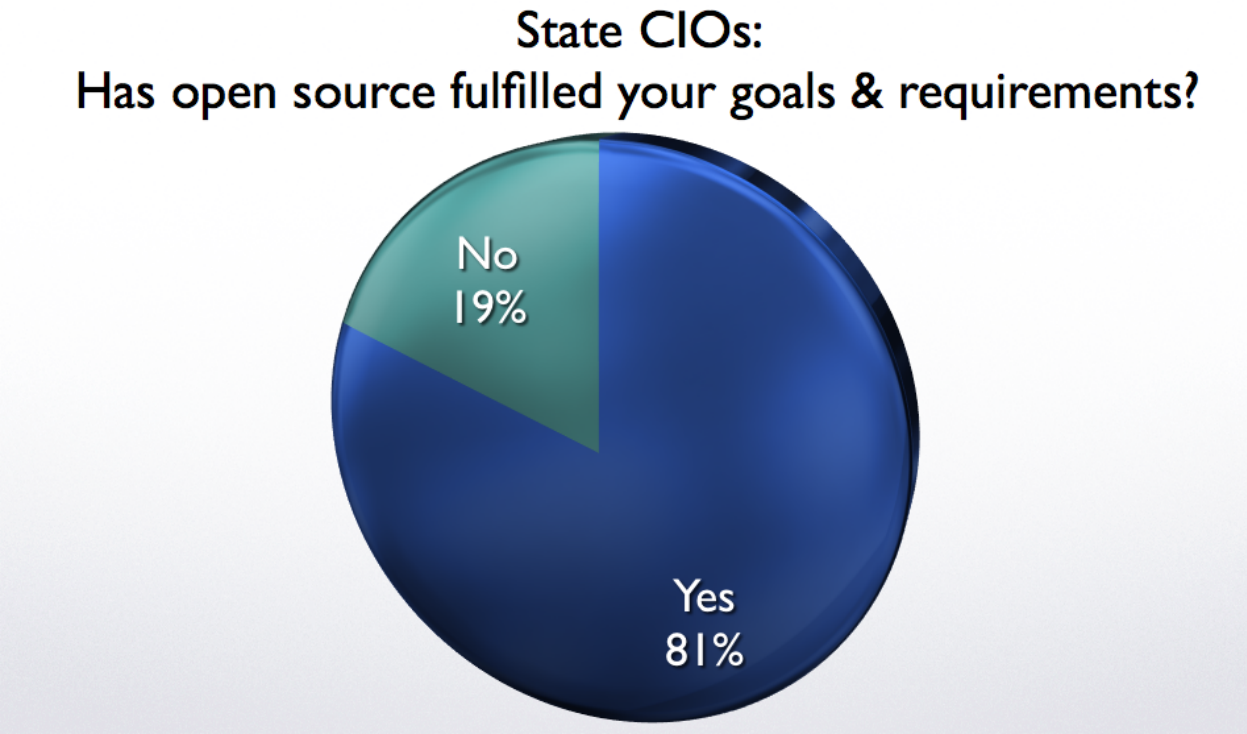


Figure 6: Satisfaction[128]



Appendix IX

ENDNOTES

[1] Definitions were derived from those found at various websites such as http://www.techweb.com/encyclopedia, http://searchwinit.techtarget.com, http://www.allbusiness.com/company-activities-management/operations-customer/5204610-1.html, www.nildram.net/jargon_buster/, http://www.naa.gov.au/records-management/publications/glossary.aspx and http://en.wikipedia.org/, and from experience.

[2] As some have stated, closed formats under certain circumstances may offer specific benefits not provided by open formats. See for example: http://www.infograph.com/opentext/CSF.htm

> *"Open versus Proprietary Formats*
>
> *Open formats, like TIFF and PDF, make it easy to share data. There are many applications that can create and view these files and they are ideal for many uses; however, because they are publicly shared formats, there are also many software programs that can edit these files. There are even applications that can open password-protected files.*
>
> *As a proprietary format, CSF is not shared with third party software companies. Only IGC products can create and view CSF files. Proprietary formats are better-suited for short-term, secure content sharing (as opposed to long-term archival, for example). Because CSF is encrypted and has security options is more secure than an open format."*

[3] Natalie Whitlock, "The Security Implications of Open Source Software," IBM (March 1, 2001), http://www.ibm.com/developerworks/linux/library/l-oss.html; John Viega, "Open Source Security: Still a Myth," *O'Reilly.com* (September 16, 2004), http://www.oreillynet.com/pub/a/security/2004/09/16/open_source_security_myths.html.

[4] "More to Life than the Office," *Business Week* (July 3, 2006), http://www.businessweek.com/magazine/content/06_27/b3991412.htm; "Rivals Set Their Sights on Microsoft Office: Can They Topple the Giant?" *Knowledge@Wharton* (August 22, 2007), http://knowledge.wharton.upenn.edu/article.cfm?articleid=1795

[5] "OASIS Frequently Asked Questions," Organization for the Advancement of Structured Information Standards (n.d.), http://www.oasis-open.org/who/faqs.php

[6] International Organization for Standardization, *ISO/ICE 26300:2006* [2006], http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=43485; Peter Galli, "Open Document Format Gets ISO Approval," *eWeek* (May 3, 2006), http://www.eweek.com/article2/0,1895,1957321,00.asp.

[7] "About Open Document Format: An Introduction," Open Document Alliance (n.d.), http://www.odfalliance.org/resources/AboutODF.pdf. The Open Document Alliance is not affiliated with OASIS, but the two groups have many members in common and both promote the adoption and ongoing development of the ODF format; see "Frequently Asked Questions (FAQs)," Open Document Alliance (n.d.), http://www.odfalliance.org/resources/FAQs.pdf.

[8] Brian Jones, "Quick Question for ODF Experts," Brian Jones: Open XML Formats, July 20, 2006, http://blogs.msdn.com/brian_jones/archive/2006/07/20/673323.aspx; *Wikipedia*, s.v., "Open Document," http://en.wikipedia.org/wiki/OpenDocument (accessed January 2, 2007).

[9] *Wikipedia*, s.v., "Open Document," http://en.wikipedia.org/wiki/OpenDocument (accessed January 2, 2007).

[10] However, users of text-reading software have also had difficulty accessing files using proprietary formats owned by Microsoft and other companies. Linda Tucci, "*Advocates for Disabled Wary of Open Source,*" *SearchCIO.com* (January 3, 2006), http://searchcio.techtarget.com/originalContent/0,289142,sid182_gci1154992,00.html.

[11] Organization for the Advancement of Structured Information Standards, "Sun OpenDocument Patent Statement, submitted by Sun Microsystems, Inc., September 29, 2005" (September 29, 2005), http://www.oasis-open.org/committees/office/ipr.php.

[12] In its public comments, Microsoft asked that the ECMA-376 standard be referred to by New York State in this report as "Open XML" or "OXML" instead of "Office Open XML" ("OOXML"). OOXML is the manner in which NYS's RFPC referred to it. Other commenters asked NYS to refer to the format as MOOXML (for "Microsoft Office Open XML").

The formal name of the document standard submitted to standards organizations ECMA and the ISO is "Office Open XML" (see, e.g., ECMA International, *Standard ECMA-376: Office Open XML File Formats* (December 2006), http://www.ecma-international.org/publications/standards/Ecma-376.htm; International Organization for Standardization, ISO/IEC DIS 29500: Information Technology -- Office Open XML File Formats (April 10, 2008), http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=45515).

Office Open XML also appears to be the most commonly used name for the format amongst the public in general including amongst those members of the public who responded to the RFPC. In fact, a Google search revealed that Microsoft itself uses the Office Open XML name more than 6,000 times on its own www.microsoft.com website, including as the formal name in the specification which Microsoft lists as legally protected by its Microsoft Open Specification Promise (see http://www.microsoft.com/interop/osp/default.mspx).

Therefore, in the absence of the format's name having been formally changed by ISO or Microsoft, the report refers to the format's legal name of "Office Open XML" or "OOXML."

[13] Frank Rice, "Introducing the Office (2007) Open XML File Formats," *Microsoft Developer Network* (May 2006), http://msdn2.microsoft.com/en-us/library/aa338205.aspx#office2007aboutnewfileformat_introduction.

[14] Microsoft, "Microsoft Announces Availability of Open and Royalty-Free License For Office 2003 XML Reference Schemas" (November 17, 2003), http://www.microsoft.com/presspass/press/2003/nov03/11-17XMLRefSchemaEMEAPR.mspx.

[15] Mike Gunderloy, "Thoughts on the Office XML Schemas," *Application Development Trends* (November 20, 2003), http://www.adtmag.com/article.aspx?id=8550&page=; Microsoft Office Online, "New Covenant vs. Old License for Office 2003 XML Schemas" (n.d.), http://office.microsoft.com/en-us/products/HA102136561033.aspx?pid=CL100796341033.

[16] Microsoft, "Microsoft Open Specification Promise" (December 18, 2007), http://www.microsoft.com/interop/osp/default.mspx.

[17] Ecma International, "Ecma International Approves Office Open XML Standard" (December 7, 2006), http://www.ecma-international.org/news/PressReleases/PR_TC45_Dec2006.htm.

[18] Microsoft, "Interoperability, Choice, and Open XML" (February 14, 2007), http://www.microsoft.com/interop/letters/choice.mspx.

[19] "Microsoft Guns Open XML onto ISO Fast Track, *Computerworld* (March 12, 2007), http://www.computerworld.com/action/article.do?command=viewArticleBasic&articleId=9012860.

[20] Arif Mohamed, "Fears on Standards as Microsoft Pushes Office Open XML Format," *Computer Weekly* (December 7, 2005), http://www.computerweekly.com/Articles/2005/12/07/213312/fears-on-standards-as-microsoft-pushes-office-open-xml.htm; Charles Babcock, "Linux Foundation Urges 'No' Vote on Microsoft's Open XML Format," *Information Week* (August 27, 2007), http://www.informationweek.com/news/showArticle.jhtml?articleID=201803012.

[21] Wolfgang Gruener, "IBM Abuses File Format Standardization Process to Limit Choice—Microsoft,: *TG Daily* (February 15, 2007), http://www.tgdaily.com/content/view/30978/118/; Tom Espiner, "Microsoft Accused of Rigging OOXML Votes," *ZD Net UK* (August 30, 2007), http://news.zdnet.co.uk/software/0,1000000121,39288959,00.htm; Stephen J. Vaughn-Nichols, "Expert: Open XML Loses Standards Battle," *eWeek* (September 4, 2007), http://www.eweek.com/article2/0,1895,2178836,00.asp

[22] Peter Galli, "Office Open XML Down but Not Out," *eWeek* (September 7, 2007), http://www.eweek.com/article2/0,1759,2180707,00.asp

[23] From: New York State Office for Technology, *Best Practice Guideline G06-001: Accessibility of State Agency Webbased Intranet and Internet Information and Applications - Issue Date: October 25, 2006; Publication Date: October 25, 2006*: http://www.oft.state.ny.us/policy/G06-001/index.htm

"Standard 14: Downloadable/Embedded Objects

14.1 — When downloadable documents [e.g., word processing documents, spreadsheets, Portable Document Format (PDF), java applets] are used, a link to accessible HTML or text version(s) will be made available.

If documents cannot be converted from their original format to an accessible format, post a notice to that effect and include contact information for users who need the information in another format. The agency still has a responsibility to make that information available in some other format (e.g., plain text, audio, etc.). WCAG 1.0: 11.3 (Pri. 3); Section 508: 1194.22(m)

14.2 — When hardware, software and assistive technology devices are controlled by a state agency, downloadable documents [e.g., word processing documents, spreadsheets, Portable Document Format (PDF)] available through an intranet or extranet will be allowed.

This standard was adopted to allow the use of technologies and content that would be considered questionable on a public site, where the visitor's hardware/software is unknown. This does not eliminate the agency's obligation to provide reasonable accommodation (e.g., screen reader, screen magnifier, adaptive hardware). WCAG 1.0: N/A Section 508: N/A"

[24] New York State Office for Technology, *Principles Governing the New York State Information Technology Enterprise Infrastructure* (March 5, 2004), http://www.oft.state.ny.us/Policy/p04-001/NYSTechPolicyP04-001.pdf

[25] Ibid.

[26] *Metaglossary*, s.v., "Interoperable," http://www.metaglossary.com/terms/interoperable/

[27] BECTA, *Microsoft Vista and Office 2007: Final report with recommendations on adoption, deployment and interoperability* (January 2008), Becta, http://publications.becta.org.uk/download.cfm?resID=35275 See also: Becta, "Becta Releases Major Report on Microsoft Vista and Office 2007 for Schools and Colleges" (n.d.), http://news.becta.org.uk/display.cfm?cfid=1989752&cftoken=8d5d40a46f6be0ce-6F4CC738-F9EA-10F9-A47DDB7F8B7EFF9E&resID=35287&page=1658&catID=1633

[28]

"Microsoft said that there is nothing wrong with having multiple file formats. The company cannot adopt ODF in its own Office suite, it said, because it cannot migrate the legacy of billions of documents in older Microsoft formats onto it. But it does allow users to export their file in ODF format.

"Any investment we make in the future of information work has to take into account what has been done in the past," said Microsoft Office project manager Gray Knowlton. "It's very important when migrating to open file formats that we take older documents into account."

"ODF was designed to omit the functionality of existing documents," Knowlton said. "We, on the other hand, cannot start from scratch. Our customers would never accept that.""

From: Brett Winterford, "Microsoft makes last-gasp OOXML push," *ZDNet* (January 29, 2008): http://news.zdnet.com/2100-3513_22-6228247.html

[29] See, e.g., Wikipedia, s.v. "Open Standard," http://en.wikipedia.org/wiki/Open_standard

[30] CIO/OFT already requires that State agencies use various methodologies to *"Maintain[ ] Authentic and Reliable E-records that are Accessible Over Time,"* including establishing standards for file formats. See *"NYS BEST PRACTICE GUIDELINE G04-001: Electronic Signatures and Records Act (ESRA) Guidelines"* (revised September 2007) http://www.oft.state.ny.us/Policy/G04-001/index.htm

[31] Eric Lai, "South Africa, Netherlands, and Korea Striding Toward ODF," *Computer World* (November 21, 2007), http://www.computerworld.com/action/article.do?command=viewArticleBasic&articleId=9048521&intsrc=hm_list

[32] Some of the specifics of these suggested definitions raise complex business and legal issues. See for example this article: Pamela Jones, *"What is Wrong with RAND,?"* Groklaw (April 17, 2008), http://www.groklaw.net/article.php?story=20080417104016186. The proposed electronic records committee will need to vet these issues fully before settling on a final recommended definition.

[33] For example, while the Firefox web browser has been gaining significant market share from usage by both individuals and entities, the absence of certain features accessible to the longer and more widely used Internet Explorer (IE) web browser but missing from Firefox has been decried. These include: (a) concerns that Firefox updates cannot be patched from a central console, as IE can be through Microsoft's Windows Server Update Services; (b) concerns that Firefox cannot be managed and secured using Active Directory, a popular tool for setting group policies; and (c) concerns that the Mozilla Foundation does not provide paid technical support services to risk-averse institutional users.

Eric Lai, "IE Still Top Dog Over Firefox in Corporate Browser Kennel," *Computer World* (January 11, 2008), http://www.computerworld.com/action/article.do?command=printArticleBasic&articleId=9056780

[34] *Wikipedia*, s.v., "Open Format," http://en.wikipedia.org/wiki/Open_format

[35] OpenOffice.org software, which is one office suite application which provides direct support for ODF within its software, already has significant market share. See: OpenOffice.org Wiki, "Market Share Analysis," http://wiki.services.openoffice.org/wiki/Market_Share_Analysis

[36] For example, in response to the question, "Who really cares about this ODF vs. OOXML battle?" Microsoft's Tom Robertson was quoted as follows:

"Robertson: Governments are the most concerned with the issue of ISO-standardized document formats. Not only do some governments have requirements to accept communications if they are presented in ISO-approved formats, but there are other factors as to why XML-based formats matter (such as long-term archival) as well. The concepts of interoperability, greater choice of solutions, and the ability to translate between formats are all important to governments. In general, we are not hearing about this issue from our enterprise or consumer customers – it is localized to governments today." Interview: Mary Jo Foley, "Microsoft: Why the ODF vs. OOXML battle matters" (February 28th, 2007), http://blogs.zdnet.com/microsoft/?p=290

[37] Burton Group, *What's Up, .DOC? ODF, OOXML, and the Revolutionary Implications of XML in Productivity Applications* (January 11, 2008): http://www.burtongroup.com/Guest/Ccs/WhatsUpDoc.aspx

[38] "Dispelling Myths Around ODF" (January 11, 2008), http://blogs.sun.com/dancer/entry/dispelling_myths_around_odf

[39] From: Erwin Oliva, "Microsoft readies openXML for global standardization," *Inquirer* [Merkati City, Philippines] (January 17, 2008),: http://newsinfo.inquirer.net/breakingnews/infotech/view/20080117-112973/Microsoft-readies-openXML-for-global-standardization

[40] Source: CIO/OFT Budget Hearing Testimony (February 7, 2008), http://www.oft.state.ny.us/News/SFY2008-09BudgetTestimony.pdf

[41] See e.g. one of the Wordstar message boards at: http://wordstar2.com/pipermail/wordstar_wordstar2.com/

[42] http://www.appleinsider.com/articles/06/01/26/iwork_has_no_game_against_office_or_wordperfect.html

[43] http://www.ogs.state.ny.us/purchase/spg/awards/5002220020Can.htm (This website may have been archived by the date of this report's issuance. It may still be accessed by using an Internet search engine's cache function).

[44] http://www.ogs.state.ny.us/purchase/snt/awardnotes/7570219465can.pdf

[45] http://www.ogs.state.ny.us/purchase/snt/awardnotes/791241427spec.pdf

[46] http://www.ogs.state.ny.us/purchase/snt/awardnotes/7951818503can.htm

[47] http://www.ogs.state.ny.us/purchase/snt/awardnotes/75016T940052can.htm (This website may have been archived by the date of this report's issuance. It may still be accessed by using an Internet search engine's cache function).

[48] http://www.openxmlcommunity.org/documents/IDC%20Document%20Adoptions%20White%20Paper.pdf

[49] See: http://quickfacts.census.gov/qfd/states/36000.html

[50] See for example:

- *Warren County, New York*: http://www.co.warren.ny.us/rp/rolls.php
- *Chatham County Schools, North Carolina*: http://www.chatham.k12.nc.us/vce/onlinereader.odt
- *New Mexico, Federal Court system*: http://www.nmcourt.fed.us/web/DCDOCS/dcindex.html
- *Western Oklahoma State College*: http://www.western.cc.ok.us/~njhood/password%20assist.odt
- *Indiana Department of Education*: http://www.doe.state.in.us/technology/docs/08techplanupdate.odt
- *Massachusetts Information Technology Division*: http://www.mass.gov/Aitd/docs/policies_standards/etrm3dot5/etrmv3dot5intro.odt
- *Argonne National Laboratory*: http://www.aps.anl.gov/epics/docs/USPAS2007/labs/DBexercise.odt
- *US Department of Energy Chemical Sciences Division*: http://actinide.lbl.gov/gtsc/seminars/2006/Soderholm_Flyer_061011.odt
- *National Aeronautics and Space Administration (NASA)*: http://fits.gsfc.nasa.gov/registry/foreign/foreign_fits_extension.odt
- *National Security Administration (NSA)*: http://www.nsa.gov/selinux/papers/gconf07-paper.odt
- *National Park Service*: http://www.nps.gov/acad/forteachers/upload/sea_app_07.odt
- *Brookhaven National Laboratory*: http://www.phenix.bnl.gov/phenix/WWW/publish/frawley/white_paper/draft/rhic2_wp_draft_dec30.odt
- *Lawrence Berkeley National Laboratory*: http://hahn.lbl.gov/svn/rsxap/tags/r1/doc/RSXAP_documentation.odt

[51] http://www.doe.state.in.us/technology/erate_documents.html

[52] See: http://www.wutc.wa.gov/webimage.nsf/0/0c107f2aecec013a8825733800684fcf/$FILE/Broadband%20RFP.doc

> *"2.4 SUBMISSION OF PROPOSALS*
>
> *Consultants must submit proposals as follows:*
>
> - *Two (2) hard-copy originals with original signatures.*
> - *One (1) electronic copy on CD-ROM in a commonly-available electronic file format such as Microsoft Office 2003 or Open Document Format that allows easy re-use, re-formatting and analysis of any underlying calculations."*

[53] See: Jason Miller, "Future Digital System Similar to NARA's Record System Only on the Surface," Government Computing News (August 14, 2006), http://www.gcn.com/print/25_24/41657-1.html; U.S. Government Printing Office, *Requirements Document (RD v2.1) for the Future Digital System (FDsys)* (April 18, 2006), http://www.gpo.gov/projects/pdfs/FDsys_RD_v2.1.pdf

[54] See Gregory W. Lawrence et al, *Risk Management of Digital Information: A File Format Investigation* (Council on Library and Information Resources, 2000), http://www.clir.org/pubs/reports/pub93/contents.html

[55] Rob Weir, "Interoperability by Design" (May 22, 2007), http://www.robweir.com/blog/2007/05/interoperability-by-design.html

[56] Tom Robertson, Microsoft General Manager for Corporate Responsibility and Standards group, quoted in "Microsoft: No File Format or Standard Is Perfect," *eWeek*, (January 16, 2008): http://www.eweek.com/c/a/Web-Services-and-SOA/Microsoft-No-File-Format-or-Standard-Is-Perfect/

This question was recently asked on a Microsoft software developers' blog:

> *"Hasn't it struck you that the need for an Open XML/ODF translator is evidence that ODF is widely used and accepted by your customer base? Why force customers to use a translator when you could have just used the existing standard and saved them the hassle?"*

The answer stated that the reason the company developed an OOXML Translator was not because it had received broad-based customer demand for full support for ODF, but because a few key government agencies raised legitimate concerns about their legal requirements to be able to accept ISO standard formats if sent to them by a constituent. The response continued, *"Over time, it may be that customers broadly ask for ODF - who am I to guess on that front. If they do, and we get broad-based requests for that to be in the product, then we should listen...."* http://blogs.msdn.com/jasonmatusow/archive/2007/02/06/open-xml-iso-iec-standardization.aspx

Microsoft has further been quoted to the effect that full compatibility may never be possible between the two formats:

"The Open XML formats are unique in their compatibility and fidelity to billions of Office documents, helping protect customers' intellectual investments. Open XML formats are also distinguished by their approach to accessibility support for disabled workers, file performance and flexibility to empower organizations to access and integrate their own XML data with the documents they use every day. In contrast, ODF focuses on more limited requirements, is architected very differently and is now under review in OASIS subcommittees to fill key gaps such as spreadsheet formulas, macro support and support for accessibility options. As a result, certain compromises and customer disclosures will be a necessary part of translating between the two formats." http://www.microsoft.com/presspass/press/2006/jul06/07-06OpenSourceProjectPR.mspx

[57] *"Novell execs discuss Microsoft's interop pledge,"* http://www.infoworld.com/article/08/03/05/Novell-execs-discuss-Microsoft-interop-pledge_1.html

[58] See the "Known Issues" section at: http://odf-converter.sourceforge.net/features.html

[59] *"Microsoft Vista and Office 2007: Final report with recommendations on adoption, deployment and interoperability"* (BECTA Report January 2008), http://news.becta.org.uk/display.cfm?cfid=1989752&cftoken=8d5d40a46f6be0ce-6F4CC738-F9EA-10F9-A47DDB7F8B7EFF9E&resID=35287&page=1658&catID=1633; http://publications.becta.org.uk/download.cfm?resID=35275

The British Educational Communications and Technology Agency (BECTA) recently called for all office suite software suppliers to fully support each format in their software:

"While we are not commenting on the technical merits of either the existing international document standard (ODF) or the proposed second international document standard (OOXML), we remain convinced that multiple incompatible international standards that address essentially the same area of functionality are not in the interests of educational users. It will introduce confusion, complexity and unnecessary costs; and it will constitute a lost opportunity of considerable proportions which will damage the marketplace, the educational community and indeed the concept of international standards per se."

"Becta supports the recommendation of the recent pan-European e-government services committee (PEGSCO) that 'suppliers should develop applications that can handle all relevant international standards, leaving the choice to their customers as to what format will be used 'by default'."

[60] *"Microsoft Makes Strategic Changes in Technology and Business Practices to Expand Interoperability,"* http://www.microsoft.com/presspass/press/2008/feb08/02-21ExpandInteroperabilityPR.mspx

[61] http://en.wikipedia.org/wiki/Source_code_escrow

[62] State Archives' Local Government Records Management Improvement Fund (LGRMIF), Grant Application and Reference Materials 2008-2009. See http://www.archives.nysed.gov/a/grants/grants_lgrmif_appl.shtml

[63] Most of the requests the City receives are for paper or computer tape versions. See: http://www.bloomington.in.gov/egov/scripts/docs.php?path=doc&id=25247&id2=23764&linked=0&fDD=407-1950

"Request Alternative Formats: Materials on the City of Bloomington website, as well as other materials published by the City, may be available in alternative formats upon reasonable request. To request materials contact the Citizen Services Coordinator at 812-349-3589 (TDD 812-349-3458) or email fultonc@bloomington.in.gov for more information."

[64] See: http://www.oft.state.ny.us/Policy/G04-001/G04-001.pdf

"Provide access to e-records in the form the user prefers: Some people do not have access to the technology needed to use e-records or prefer records in paper form. ESRA, and the ESRA regulation (see Part 540.5(b)(1)) require governmental entities to provide access to e-records as permitted by statute and in paper form if requested. This does not mean that governmental entities must maintain paper copies of e-records, only that they have the technical capability to generate copies of e-records that are accessible under the law in both paper and electronic form. This will likely require appropriate output devices, such as a high-quality printer capable of producing legible or useable copies of records."

[65] See: http://www.dos.state.ny.us/info/nycrr.htm:

"The complete unofficial NYCRR is now available online. This initial release of the online NYCRR does not yet include a table of contents or all graphic images. To find desired text, visitors simply enter a search term or NYCRR citation. The table of contents and graphic images will be available in the near future. " and it links to:

http://government.westlaw.com/linkedslice/default.asp?SP=nycrr-1000:

"Welcome to the online New York Codes, Rules and Regulations. This website is maintained by Thomson West under contract with the New York Department of State to provide free public access to the full text of the New York Codes, Rules and Regulations. You may access the online New York Codes, Rules and Regulations through the following link:

... This site from Thomson West provides free access to an unannotated version of the New York Codes, Rules and Regulations.

The information contained on this site is not the official version of the Official Compilation of the Codes, Rules and Regulations of the State of New York (NYCRR). No representation is made as to its accuracy, nor may it be read into evidence in New York State courts. To ensure accuracy and for evidentiary purposes, reference should be made to the official NYCRR. The official NYCRR is available from Thomson West...." (emphasis added)

66 http://www.osec.doc.gov/webresources/bestpractices/universal_Access_guidance2.html

67 See for example this article, and the links cited therein:
http://www.groklaw.net/article.php?story=20070827111019189

68 http://en.wikipedia.org/wiki/File_format

69 *"...the Office Open XML format being assessed by the ISO 'is not what Microsoft implements in the Office suite,' Vinje said, adding that 'If you implement OOXML, you don't get interoperability with Office.'" "Microsoft's ISO win may worsen its antitrust woes,"* April 1, 2008: http://www.infoworld.com/article/08/04/01/Microsofts-ISO-win-may-worsen-its-antitrust-woes_1.html

See also: *"Documents that conform to the OpenDocument specification may contain elements and attributes [extensions] not specified within the OpenDocument schema." "Game Over for Open Document?,"* July 23, 2007:
http://www.linuxworld.com/news/2007/072307-opendocuments-grounded.html?page=1

70 See for example: "Weak welcome for Microsoft's interoperability promise," February 25, 2008:
http://www.macworld.co.uk/business/news/index.cfm?newsid=20527 and "*Game Over for Open Document?,"* July 23, 2007: http://www.linuxworld.com/news/2007/072307-opendocuments-grounded.html?page=1

See also:

"Microsoft has been roundly and deservedly criticized for putting application-specific document settings markup in the OOXML specification. Is ODF somehow superior because it allows the creation and use of such custom markup without any specification whatsoever? Here is what it says in the ODF conformance section:

"Documents that conform to the OpenDocument specification may contain elements and attributes [extensions] not specified within the OpenDocument schema. Such elements and attributes must not be part of a namespace that is defined within this specification and are called foreign elements and attributes.

...

"The various <style:-properties> elements may have arbitrary attributes attached ...*

...

"There are no rules regarding the elements and attributes that actually have to be supported by conforming applications, except that applications should not use foreign elements and attributes for features by the OpenDocument schema.

ISO/IEC:26300-2006 OpenDocument section 1.5.
http://standards.iso.org/ittf/PubliclyAvailableStandards/index.html

*Short story: ODF "interoperability" is a complete and utter myth and the OpenDocument TC isn't exactly excited about making the myth come true. "Open" and "interoperable" are *not* synonyms, whether that bit of disinformation comes from Microsoft or anyone else. When ODF advocates criticize OOXML on grounds of non-interoperability, it's no more than the proverbial pot calling the kettle black."*

71 http://ec.europa.eu/idabc/en/document/6474

72 http://ec.europa.eu/idabc/servlets/Doc?id=27956

73 *"The Case for Harmonization,"* January 31, 2008: http://www.robweir.com/blog/2008/01/case-for-harmonization.html . See also: *"ODF-OOXML Harmonization: Yes, We Can!,"*
http://www.odfalliance.org/resources/IssueBriefHarmonization.pdf

[74] http://www.lesechos.fr/medias/2008/0331//300253699.pdf See also: *"The France Shift From No to Abstain -- HP helped Microsoft France do it - Updated,"* March 31 2008: http://www.groklaw.net/article.php?story=20080331212042460

[75] *"The need for truly open standards and processes is demonstrably more important than ever. IBM will continue to be an active supporter of ODF. We look forward to being part of the community that works to harmonize ODF and OOXML for the sake of consumers, companies and governments, when OOXML control and maintenance is fully transferred to JTC1 (ISO/IEC)."* http://www.groklaw.net/article.php?story=2008040212120873

[76] http://www.ansi.org/

"How does ISO decide what standards to develop? Working through the ISO community, it is the people who need the standards that decide. What happens is that the need for a standard is felt by an industry or business sector, which communicates the requirement to one of ISO's national members. The latter then proposes the new work item to ISO as a whole. If accepted, the work item is assigned to an existing technical committee. Proposals may also be made to set up technical committees to cover new scopes of technological activity. In order to use resources most efficiently, ISO only launches the development of new standards for which there is clearly a market requirement."

http://www.iso.org/iso/support/faqs/faqs_standards.htm

[77]
http://publicaa.ansi.org/sites/apdl/Documents/Standards%20Activities/American%20National%20Standards/Procedures,%20Guides,%20and%20Forms/2008%20ANSI%20Essential%20Requirements/2008%20ANSI%20Essential%20Requirements%20031108.pdf

[78] See: http://www.eweek.com/c/a/Linux-and-Open-Source/Mass-Gives-Open-XML-the-Official-Green-Light/

"The Commonwealth continues on its path toward open, XML-based document formats without reflecting a vendor or commercial bias in ETRM v4.0. Many of the comments we received identify concerns regarding the Open XML specification. We believe that these concerns, as with those regarding ODF, are appropriately handled through the standards-setting process, and we expect both standards to evolve and improve."

[79] http://www.mass.gov/Aitd/docs/legal/odf_accessibility_midyear_ltr.pdf

[80]
http://www.mass.gov/?pageID=itdterminal&L=4&L0=Home&L1=Policies%2c+Standards+%26+Guidance&L2=Enterprise+Architecture&L3=Enterprise+Technical+Reference+Model+-+Service+Oriented+Architecture+(ETRM+v.+4.1)&sid=Aitd&b=terminalcontent&f=policies_standards_etrm4dot1_ETRM4dot1Final_ETRM4dot1InformationFD&csid=Aitd#_ftn1

[81] From: *"Accessibility Issues with Office Open XML,"* January 4, 2008, http://atrc.utoronto.ca/index.php?option=com_content&sectionid=14&task=view&hidemainmenu=1&id=371

[82] There have also been calls for the use of single, open formats by various assistive technology and religious groups. See e.g.: http://lxer.com/module/newswire/view/47810/index.html

[83] Two persuasive examples include: (a) the findings of the Massachusetts State Auditor in the report it issued concerning the Commonwealth's open format plans: *"Office of the State Auditor's Report on the Examination of the Information Technology Division's Policy for Implementing the Open Document Standard"* (September 2007): http://www.mass.gov/sao/200608844t.pdf; and (b) a document published in 2005 by the Berkman Center for Internet and Society at Harvard Law School which reflected an approach recommended by individuals from numerous national governments entitled the *"Roadmap For Open ICT Ecosystems"* (2005): http://cyber.law.harvard.edu/epolicy/roadmap.pdf

[84] *"Electronic Records Management and Digital Preservation: Protecting the Knowledge Assets of the State Government Enterprise PART I: Background, Principles and Action for State CIOs"* (May 2007): http://www.nascio.org/committees/EA/download.cfm?id=82

[85] *"Metadata describes other data. It provides information about a certain item's content. For example, an image may include metadata that describes how large the picture is, the color depth, the image resolution, when the image was created, and other data. A text document's metadata may contain information about how long the document is, who the author is, when the document was written, and a short summary of the document."* http://www.techterms.com/definition/metadata

[86] *"Electronic Records Management and Digital Preservation: Protecting the Knowledge Assets of the State Government Enterprise,"* Parts I, II, and III, available at: http://www.nascio.org/publications/

- Partner actively with your electronic records management and digital preservation function to develop strategies for proactively managing records and digital archives.
- Support an enterprise approach to electronic records management and preservation.
- Require attention to electronic records management and preservation in capital investment proposals, and project plans.
- Create an electronic records management and digital preservation domain under the Enterprise Architecture program to foster collaboration, shared decisions and common enterprise solutions.
- Embrace the principles of electronic records management.
- Understand electronic records management and digital preservation as disciplines for managing knowledge assets of the enterprise.
- Become familiar with the economic, organizational, and technological issues related to electronic records management and digital preservation. Bring this thinking into the culture of the CIO office and IT operations.
- Lead the establishment of the necessary relationships and project planning delivery processes to ensure electronic record retention rules are automated as much as possible. Avoid reliance on administrative controls to implement records management retention rules.
- Create an electronic records management and digital preservation domain under the Enterprise Architecture program to foster collaboration, shared decisions and common enterprise solutions.
- Prepare a baseline for your state to more fully understand the legal framework, institutional roles, responsibilities, authorities and existing services for managing electronic records. Leverage expertise to expose the gaps and identify at-risk state government digital information.
- Champion the promotion of digital capabilities for managing enterprise knowledge assets and the inherent capabilities of digital preservation related to disaster planning.
- Lead the development of collaborative relationships with and among the functions of records management, archiving, library services and digital preservation. Promote the development of a consistent operating discipline across all branches of state government.
- Establish the CIO office as the lead for the operating discipline for managing knowledge assets as part of the state Enterprise Architecture Program. Include electronic records management and digital preservation as a domain within the state Enterprise Architecture Program.
- Lead the establishment of standards for project and capital investment proposals to include the total cost of ownership including the long term cost of managing the enterprise knowledge assets that are created and referenced by these investments. Ensure the state project management training includes material on this topic.
- Partner with state expertise centers for records management and digital preservation to establish goals, objectives and strategies for managing knowledge assets. Leverage national initiatives and vendor solutions related to digital preservation technology. Because they are in the early stages of development, maintain a healthy skepticism toward these initiatives.
- Build awareness and lead the development of a global perspective across the enterprise relative to global sourcing, and offshoring of digital assets. Be a communicator of the risks and long term effects of moving digital assets offshore. Lead the development or enhancement of project management delivery processes that include proper attention to viability analysis and risk analysis. These processes must include evaluation of economic and political factors and appropriate attention to national security and defense when evaluating proposals. Be the conscience of the enterprise.
- Most, if not all, states and the federal government have laws that provide a definition of what constitutes a "record." These laws may establish regulations for records retention and disposition. Records should not be destroyed arbitrarily or capriciously by a government employee. Some process must exist to authorize the destruction of records to ensure that it is done systematically to protect the state from charges of spoliation.
- All records have value. That value may be short term (until received and read) or long term (forever). Records managers and archivists evaluate administrative, fiscal, legal and historical value of records to determine how long records need to be kept and when destruction should be authorized.
- Records managers and archivists must maintain an enterprise view, and compare one record to all other records created by an enterprise when appraising relative value.

- Record keeping has become distributed. This has a huge impact on the volume, variety and maintenance of electronic records. IT Operations and Data Management need to collaborate with records managers, and archivists regarding how to address these challenges.
- Desktop recordkeeping issues are different than database/system issues. There is no one-size-fits-all solution.
- If an electronic record needs to be retained for more than 10 years, the original technology that was used to create/store it will most likely become obsolete during its lifecycle. IT Operations will be responsible for migrating/maintaining the electronic records during this time. This will require an investment in resources and skills. IT Operations needs to work with records managers and archivists to understand, assess and address this situation. State archivists will eventually take custody of some of these electronic records for permanent preservation. They need help from IT Operations and new funding for this task. Otherwise, essential records that protect the rights and interests of the state and individuals, as well as the history and culture of the people will disappear.
- Electronic Records Management must be an integral part of any project or investment proposal. Project managers must work with records managers and state archivists to determine appropriate plans and associated investments required to maintain the records generated and/or referenced by any business process, or system. These plans must be part of presenting the Total Cost of Ownership associated with any project plans to deliver business processes, or systems.
- Electronic Records Management discipline must be viewed as an integral part of the state enterprise architecture. Records Management (and Knowledge Management) touch every aspect of Enterprise Architecture – Business, Information, Process, Organization, Technical, Program & Project Management, Security. NASCIO has made a similar case regarding the integral nature of enterprise architecture with project/program management, security, and procurement.
- Records Management has inherent risk management issues. What to keep and what to destroy will always constitute a balancing act. No one can fully anticipate what knowledge and information will be sought now or at some point in the future. State government cannot keep everything. Records Management policy will have to be established to drive the decision making process for managing records.
- State CIOs and Enterprise Architects must partner with Records Managers and State Archivists, and State Librarians in order to establish necessary elements for managing digital assets. These elements include policy, responsibility, capacity, and, understanding and awareness among state employees. The desired outcome is to ensure the state has the information it needs today – and tomorrow.
- State CIOs are focusing more on business strategy as their roles, demands and expectations are expanding. Governors rely on them to interpret, organize and effectively harness information technology to serve the state. In the area of electronic records and digital preservation, NASCIO stresses the need for cooperative, collaborative relationships – particularly with state offices that have statutory responsibility for these lines of business. State CIOs and Enterprise Architects must partner with the appropriate officials in order to ensure Records Management Policy, Principles and Best Practices are implemented effectively.

[87] *"Federal Enterprise Architecture Records Management Profile"* (December 2005): http://www.archives.gov/records-mgmt/pdf/rm-profile.pdf

[88] http://www.openformats.org/en61

[89] *"Open standards, economics, and innovation"* (April 2007): http://www.thebolingroup.com/collaborativeadvantage/downloads/Rishab%20Ghosh%20Beijing%20Standards%20Edge.pdf

[90] http://www.budget.state.ny.us/guide/bprm/h/h-300a.pdf

[91] From the NYS Procurement Council's Procurement Guidelines:

http://www.ogs.state.ny.us/procurecounc/PDFdoc/Guidelines.pdf

"Section 5-C. Utilization of Centralized Contracts

The following outlines the general procedures and agency responsibilities for using centralized contracts.

1. OGS will disseminate information about centralized contract offerings to agencies through bulletins, contract award notices, electronic access and user groups. A listing of agency representatives has been established to receive such information and will be updated periodically. Purchases under the centralized services contracts will be

generally processed through a purchase order which should be identified by "PT" or "PS" contract numbers. Centralized commodity contract purchases are identified by "P" contract numbers. Alternatively, when the contract has a "CMS" designation, it must be accessed by a contract encumbrance. The "PT" contract identifier provides for purchases of technologies. The "PS" contract identifier provides for purchases of services. The "CMS" contract identifier primarily provides for purchases of services and technology which utilize payment schedules which qualify for the automatic payment process and contracts requiring a mini-bid.

2. OGS may provide more than one contract that could be used to address the needs of an agency. The agency determines the most appropriate centralized contract that addresses their needs and provides the most cost effective solution. A large volume purchase requirement may enable an agency to solicit best and final offers from potential vendors under the centralized contract pricing. Such purchases will still be made under the centralized contract, but at the special pricing offered by the vendor.

3. Agencies are to purchase from centralized commodity contracts if the item is available in the form, function and utility consistent with an agency's need. Agencies have the option of using centralized services contracts, unless otherwise specified by the State Procurement Council, or establishing their own contracts. A filed requirement approach may also be used by PSG. With such an approach, agencies will be asked to define their need and commit to use of the centralized contract. This information may be utilized in the bid solicitation to assist in ensuring the most cost effective contract.

4. OGS will, as necessary, establish contracts through a sole source or single source procurement. To support these procurement methods, agencies may need to provide documentation which details the special circumstances and factors that justify a sole or single source procurement.

5. The benefits associated with centralized contracts generally exist from the merging of multiple agency needs; however, OGS may undertake a contract for a single agency for a pilot project or a prototype acquisition. On a limited basis, PSG will also establish a contract for a specific agency, upon request.

6. Vendor lists established by PSG are also available to State agencies when undertaking independent competitive procurements.

The PT contracts recently available for State agency purchases of office suite software include:

1. Microsoft software products such as Microsoft Office have been available to State agencies through a centralized agreement with Hewlett-Packard corporation as a reseller, under OGS Group # 76304 Award # 18766 Contract # PT 61408:

http://www.ogs.state.ny.us/purchase/snt/awardnotes/7630418766spec.pdf

2. Sun software products such as StarOffice have been available to State agencies through a centralized agreement directly with Sun Microsystems, under Group # 75016 Award # T940052 Contract # PT00086:

http://www2.ogs.state.ny.us/cs.html?url=http%3A//www.ogs.state.ny.us/purchase/prices/75016T940052prices.pdf&charset=iso-8859-1&qt=url%3Apurchase+%7C%7C+pt00086&col=mergall&n=1&la=en

3. In the past Corel products such as WordPerfect have been available through similar centralized contracts (Centralized Microcomputer Software contract Group # 76314) and now are available through centralized reseller contracts such as the ASAP Software Miscellaneous Software contract Group # 79518 Award # 18503 Contract # PT60291.

These contracts all include contractual terms and conditions similar to the following:

"PROCUREMENT PROCEDURES

The following guidelines will assist State Agencies in procuring large purchases resulting from the use of this contract. State Agencies should carefully consider all alternatives to determine which provides the best products and cost.

Agencies are reminded that procurements resulting from this contract of all software purchases greater

than $20,000 or hardware greater than $50,000 in value require prior approval from the NYS Office for Technology. Refer to OFT Technical Policy 96-2 and 96-2A, available at OFT's web site at http://www.oft.state.ny.us/policy/index.html

Agencies are required to have a procurement record for purchases for software above $50,000. This record should detail the procurement procedure used and why the product purchased has been determined to be the most cost-effective for current and future needs, and will be subject to a post audit by OSC.

In accordance with the Guidelines for use of the technology contracts, it is recommended that the agency aggregate its requirements and make as few purchases a year as possible. Each purchase of product should be acquired in accordance with the agency's established policy.

No approval by the Office of General Services is required.

Although not under OGS purchasing authority, political subdivisions should exercise similar judgment when making large purchases from OGS PSG contracts.

These procedures may be modified in the future as needed to reflect changes in procurement law."

[92] In New York State a "sole source" contract means a procurement in which only one Contractor is capable of supplying the required product or service. This is contrasted with a "single source" procurement in which although two or more Contractors (or participating resellers) can supply the required product or service, an Authorized User, upon written findings setting forth the material and substantial reasons therefore, may award the order to one Contractor (or participating reseller) over the other(s). The Authorized User is required to document in the procurement record the circumstances leading to the selection of the vendor, including the alternatives considered, the rationale for selecting the specific vendor and the basis upon which it determined the cost was reasonable. See the State Procurement Council Guidelines for more information: http://www.ogs.state.ny.us/procurecounc/pdfdoc/guidelines.pdf

[93] The template terms and conditions can be found here: http://www.ogs.state.ny.us/purchase/BidTemplate/SoftwareTemplate.doc

[94] Source: CIO/OFT Budget Hearing Testimony (February 7, 2008): http://www.oft.state.ny.us/News/SFY2008-09BudgetTestimony.pdf

[95] http://www.oft.state.ny.us/esra/ESRA_Report_2004/ESRA_Report_2004.pdf Many of the recommendations from that report retain full validity, including that:

- State agencies should keep their hardware and software current -- upgrading to new versions regularly -- and they must ensure that electronic records remain usable in each new environment, to prevent technological obsolescence resulting from changes in hardware, software, file formats and media formats;
- State agencies should work with the archival community to embrace cost-effective, standard solutions for the preservation of electronic records in usable formats;
- The Legislature should consider providing the State Archives sufficient resources to preserve large quantities of electronic records in a large number of formats, and to redevelop its capacity for managing archival electronic records, including for its Electronic Records Team to complete a plan and methodology for managing e-records, including:
    - developing staff training plans in electronic records;
    - implementing specific electronic records projects;
    - purchasing hardware to be used to copy, error-check, and manage digital storage media; and
    - hiring professional staff members responsible for the entire range of activities regarding the management of archival electronic records.
- For e-signed records, State agencies need to work with State Archives to preserve the context and links between components of electronic records, such that additional evidence is provided to support the reliability and authenticity of the signed electronic record and/or to actually constitute the electronic signature itself, by:
    - determining what information needs to be retained to maintain a valid, authentic, and reliable signed electronic record; and
    - preserving the link or association between the various components of a signed record over time; and

• For the State Archives and CIO/OFT to collaborate more concerning the systematic collection of data to better identify electronic record preservation challenges faced by public entities and ascertain opportunities for solutions.

[96] *"How to minimize the pain of an Office 2007 upgrade: Ignoring dead documents, wrestling with templates, and other changeover joys,"* (December 2007): http://www.computerworld.com/action/article.do?command=viewArticleBasic&articleId=9050439&intsrc=hm_list

[97] "Risk Management of Digital Information: A File Format Investigation," by Gregory W. Lawrence, William R. Kehoe, Oya Y. Rieger, William H. Walters, and Anne R. Kenney. Published by the Council on Library and Information Resources, June 2000. See http://www.clir.org/pubs/reports/pub93/contents.html

[98] See discussion of the European Commission's *"European Interoperability Framework (EIF)"* and *"Architecture Guidelines (AG)"* at: http://ec.europa.eu/idabc/en/document/6227 and *"Preparation for Update European Interoperability Framework 2.0 - FINAL REPORT"* (Gartner, April 2007): http://ec.europa.eu/idabc/servlets/Doc?id=29727 and

[99] These include BECTA in the United Kingdom.

[100] From: http://dts.utah.gov/main/etechresearch/docs/tareviews/TA%20Review%20Office%20Suites%2012.11.2007R2fs.pdf

Utah Department of Technology Services: "TECHNICAL ARCHITECTURE REVIEW"

> *"Recommendations: The assumptions in group A seem to align most closely with what the State has done and what it can likely afford to do in the future. Based on Group A as a recommendation platform the State should:*
>
> *· Establish a common document exchange standard for word processing, spreadsheet, presentation, visualization, and light-weight database files. This standard should incorporate existing Microsoft file formats and be interoperable with Office 2007 applications.*
>
> *· Default installations of office suites in agencies to the common document exchange format standards.*
>
> *· Allow agencies to use Office 2007 Pro Standard subject to agency preferences and budget availability.*
>
> *· Support OpenOffice.org as a standard office suite with all document formats defaulted to the approved document exchange standard. Agencies with older versions of the Microsoft Office Suite, but with a limited upgrade budget, should be encouraged to migrate to OpenOffice under the Novell MLA.*
>
> *· Use the OpenOffice license11 provided under the MLA with Novell for implementation.*
>
> *o Novell licenses fonts from AGFA to match the kerning and spacing of the fonts available in Microsoft Office. This means that pagination and layout will remain the same across office suites, and an organization can have a mix of OpenOffice.org and MS Office users that can easily collaborate and share documents.*

The United States Department of Commerce has also posted several webpages describing in detail the formats it considers appropriate to use when communicating with the public. See e.g.: http://www.osec.doc.gov/webresources/bestpractices/bp6_universal_webpage.htm and http://www.osec.doc.gov/webresources/bestpractices/universal_Access_guidance2.html

[101] The risks in remaining with a format adopted by a single company's software application were recently amply illustrated by the decision of Microsoft to issue a service pack to its Office 2003 office suite application which unilaterally disabled users' access to document formats which the software previously had opened. Microsoft's stated reason was that the portion of its office suite software code which enabled opening and saving in these various file types was insecure. Commentators warned that entities which had saved large numbers of documents in those formats risked losing access to their saved documents if they upgraded their Microsoft office suite software.

See: http://www.microsoft-watch.com/content/business_applications/sp3_downgrades_office_2003.html; http://www.tgdaily.com/content/view/35504/118/

[102] DiMaio, Andrea; "Open Source in Government and Software Vendors: Size Matters, Up to a Point"; Gartner Research, March 29, 2006

[103] Bruce Perens, "The Open Source Definition", http://www.opensource.org
[104] Festa, Paul; "Yahoo Shifts to Open Source Scripting"; CNet News; http://news.com.com/2100-1023-963937.html

[105] O'Reilly, Timothy; http://www.oreilly.com/pub/a/oreilly/ask_tim/2004/amazon_0204.html

[106] Brown, Ericka; "The Feds Love Linux"; Forbes; July 20, 2003; http://www.forbes.com/2003/06/20/cz_eb_0620linux.html?partner=yahooandreferrer

[107] "Western European End-User Survey: 2005 Spending Priorities, Outsourcing, Open Source and Impact of Compliance", IDC Research

[108] "Red Hat: Born on Wall Street and Here to Stay"; EWeek; April 24, 2006

[109] "Google Revealed: The IT Strategy That Makes it Work"; Information Week; August 28, 2006

[110] See "JFS for Linux"; http://jfs.sourceforge.net

[111] See "Security Enhanced Linux"; http://www.nsa.gov/selinux/

[112] InAccess; http://www.doe.state.in.us/inaccess/

[113] "Hoosier Daddy? In Indiana Schools, It's Linux"; CRN Magazine; August 6, 2006; http://www.crn.com/sections/breakingnews/breakingnews.jhtml?articleId=192201386

[114] "New York School Districts Select Linux Desktops from Novell to Improve Student Success"; Novell Corp. Press Release courtesy of Reuters; http://today.reuters.com/news/articleinvesting.aspx?view=PR&symbol=NOVL.O&storyID=218458+15-Aug-2006+PRN

[115] Suse Enterprise Linux Desktop

[116] "Hoosier Daddy? In Indiana Schools, It's Linux"; CRN Magazine; August 6, 2006; http://www.crn.com/sections/breakingnews/breakingnews.jhtml?articleId=192201386
[117] Di Maio, Andrea; "Local Governments in France Move to Open Source Applications", Garter Research; July 8, 2005

[118] Harris, Richard G.; "Shared Services Offer Promise for Governments", Gartner Research Report, 22 January 2004

[119] See Amazon.com for details; http://www.amazon.com/AWS-home-page-Money/b?ie=UTF8andnode=3435361

[120] http://cpr.ca.gov

[121] Seriouslyopen.org; http://www.seriouslyopen.org/nuke/html/index.php

[122] Open Source Maturity Model; http://www.navicasoft.com/pages/osmm.htm

[123] Cap Gemini Open Source Maturity Model v1.5.3; http://www.seriouslyopen.org/nuke/html/modules/Downloads/osmm/GB_Expert_Letter_Open_Source_Maturity_Model_1.5.3.pdf

[124] Based on data provided from the report "How Firms Should Work With the Open Source Ecosystem"; Forrester Research; October 4, 2005

[125] Data source: 2005 NASCIO Conference; http://www.nascio.org

[126] Ibid.

[127] Ibid.

[128] Ibid.