An OASIS White Paper

OASIS EML – USA Case Study

Experiences with California Secretary of State Election Night Reporting

Submitted for NIST Workshop on a Common Data Format for Electronic Voting Systems

Abstract

The OASIS Election and Voter Services Technical Committee was approached to assist the California Secretary of State office (CA SoS) in the use of XML for reporting election results. Initially the goal was to show that the previous election for 2006 for state governor (the "Schwarzenegger" Election) results could be exported into XML successfully from the existing SQL database content.

The project went from proof-of-concept to actual live field use for the 2007 and 2008 elections.

We discuss here the goals, resources, challenges, outcomes and lessons learned. Also compare and relate these topics to the NIST workshop initial assertions for requirements, scope and testing.

The Challenge and Resources

California needed to ensure real time results reporting service to media sources during Election Day coverage. That format had to allow rendering into a variety of human readable formats such as web, broadcast and print media. Their existing software was using a SQL database and written in Perl code with one fixed layout for internal reporting purposes. It accepts and created feeds in limited CSV text file formats.

The initial challenge was verifying that the OASIS EML structure and layout for information could handle the existing data contained in the SQL database. Combined with this was staff knowledge and familiarity as the internal CA SoS staff had not used XML or XML tools before. Management was also skeptical of the approach and did not see the need for a technology refresh (using Microsoft Visual Developer and .NET).

The OASIS EML committee members were able to provide an initial straw man XML example and rendering stylesheet (XSLT) that all worked in Microsoft IE web browser for the CA SoS staff to review. Then materials to talk to the technology pro's and con's for management decision makers, and also find a local XML class for staff training in Sacramento. The training proved vital in allowing CA SoS staff to assimilate and work with XML and provide the necessary implementation from the SQL database.

Outcomes and Lessons Learned

Over the ensuing months CA SoS was able to implement first a pilot that converted the 2006 election results from the SQL database into EML 510. This took about 3 months from inception to completion and management approval to continue further. At this point management decided to poll 10 major news media providers as to the applicability of providing direct XML feeds of election results. To their surprise the news media technical staff overwhelming responded with positive "Yes please, when can we have that, now?" This convinced management of the need and the approach.

Over the next three months work proceeded on adapting the EML 510 layout to a new EML 530 statistics reporting layout that was co-developed between CA SoS and the OASIS Election TC members and added to the overall specifications from OASIS. The EML 530 includes special support for USA district elections style reporting and an extensible statistics mechanism so that any kind of statistical values can be inserted into the reporting structure. An example stylesheet (XSLT) was also developed to render the information as a web page.

Once this work was complete CA SoS then implemented the remaining piece which was to provide a secure interface via ftp remote access to allow media sources to retrieve results files directly from the CA SoS server. Information is provided below in the resources section on all these details.

The lessons learned here include that when public election authorities work with the OASIS EML TC to develop standards the outcomes are extremely positive, are developed rapidly and without imposing prohibitive demands on internal staff resources or skills. Average developers with typical software development skills are fully capable of successfully implementing solutions.

Also the OASIS EML approach of providing Schema structures that define the information content needed is able to smoothly adapt to USA voting scenarios. Furthermore in the new OASIS V6.0 specifications the core structure definitions have been enhanced in line with the CA SoS experience to ensure that the complete suite of EML information structures for election and voter services now support those capabilities to handle USA district and precinct style elections. This further demonstrates the power of the modular approach engineered into the OASIS EML standard.

Compare and Relate to NIST Topics

In the call for papers the following topics are given; our notes relating these to the CA SoS use of EML are given in *blue italic* text inline:

- Requirements
    - Human-readable versus machine readable-human-readability matters for transparency and auditing, whereas integration of equipment indicates use of an existing data exchange syntax that is optimized for machine-reading rather than human-reading.

    In this case CA SoS was able to show how the OASIS machine readable formats enhanced the ability to deliver human readable content to new media sources that could then be directly rendered from that content using stylesheets. Also XML is intrinsically more human accessible as the CA SoS technical resource guide illustrates (see Resource link below).

    - Flexible and extensible versus uniform and complete-interoperability is helped by a standard data format only if the format is used in the same way by everyone; other goals may accommodate a more flexible and extensible format.

    OASIS EML is not provided constrained by this dogma and thinking. It provides extensible mechanisms that adapt and leverage a consistent dictionary of common components. The whole point is to provide a standard framework that implementers can then tailor to their exact needs and then ensure alignment through use of the OASIS EML dictionary of components. The same approach is true of the US NIEM.gov approach exploiting the power and flexibility of XML driven mechanisms (see the paper provided on OASIS EML and information integration tools including NIEM approach).

    - Secure versus convenient-securing the records to make them valid for auditing could also hamper intermediate processing that might be needed to integrate disparate pieces of equipment.

OASIS EML is able to fully support the CA SoS needs in this area and provide media sources with trusted information feeds. OASIS EML provides additional use of digital certificates and more security mechanisms that were not defined as needed for CA SoS purposes. Integration across equipment is not an issue since the digital signature component does not prevent processing by systems that are not XML digital signature aware.

- Format versus content-to "get the data out" it is more important to specify the data to be provided, the points in the process at which it is to be provided, and the mechanism for export than the format of the data.

*Using XML based standards provides flexibility. It is **not** true to make the generalization made here in regard to election systems and getting any old data out. In certain exchanges strict adherence is vital to ensure verifiable information and that all security and other procedures are met. If someone is sending you information about concert tickets or the types of flower arrangements they can provide you, then you may be laissez faire about it. But if someone is reporting actual legal election results for the most important free democracy on the planet, then they had better do that in a precise and exact manner that is verifiable. OASIS EML specification provides detailed discussions about the points in the process and the election workflows.*

- Scope
  - Publishing results involves only the reporting phase of the election and should not include such information as individual ballot images.

That is self-evident and OASIS EML fully supports voter privacy and voter rights. In fact a two year study was commissioned across 13 European Union member states into all these aspects and OASIS EML fully enshrined those rights for use in the European Union elections. These also coincidently fully support and exceed US needs (the OASIS EML specification provides links to these EU Commission reports).

  - Interoperability might only be needed at certain points in the system architecture.

Again self-evidently, internal information exchanges can use local formats; however any external intra-system or intra-device information exchanges must use compatible standards to ensure correct operation.

  - Registration may or may not be within scope.

OASIS EML supports voter services, not just Election Day voting. For example the new EML 150 geo-districting, co-developed with Pew Foundation and Google staff and OASIS Election TC now supports online web mapping of polling places and district boundaries. This project is described in a separate paper.

  - Low-level event log data may only be useful for auditing.

OASIS EML provides three types of applicable audit logging and event tracking mechanisms and associated XML structures. These were not applicable to the CA SoS use however.

- Testing
    - If the data format is to be used for auditing, usability testing with auditors is indicated.
    - If the goal is interoperability, interoperability testing is different from conformity assessment

*OASIS EML provides template formats in addition to schemas to verify precise information content and structure use in data exchanges. These mechanisms are used in industry standard best practices to test and verify XML information exchanges. In the CA SoS instance the schema verification method was used along with visual inspection of rendered election report information. In addition a test site was established to allow media systems to test prior to election night (*see http://www.sos.ca.gov/media/).

Information Transparency, Access and Standards

The CA SoS initiative represents a first step toward better transparency and full availability of election information in formats that provide easy access for citizens and election interest groups and organizations. Standards fulfil a key role in transparency and ease of access. When information is in predictable and prescribed formats then software tools can be created to access the data and provide a variety of tools. For example the EML 530 statistics may be rendered as Microsoft Excel spreadsheet XML so that anyone can then review those details on a PC equipped with software capable of reading in that format. That software conversion tool can then be made available and shared. When that information is predictable over time then this permits long term analysis and the intrinsic value increases. The CA SoS now has data available for 2006, 2007, and 2008 elections all in the same EML format. Similarly for news media organizations whom were the original target for CA SoS information sharing, their investment in software to process CA SoS data becomes more leveraged as more States provide their information in the same formats and data quality.

A gauge of the demand for the information is shown by the fact that the CA SoS public site running on a T3 line was initially overwhelmed by demand at 8pm on the November, 2008 election night by people wanting to view online the results. This demonstrates the need to distribute results widely to multiple reputable news sources and avoiding a single point of failure.

Conclusions

The CA SoS has now successfully used their interface for results reporting and statistics for the 2007 and 2008 elections to news media sources with live information. The OASIS EML 510 and EML 530 formats used are capable of meeting the needs of one of the largest and most populous States in America. The CA SoS November 2009 election download site is already open for developers to access in test mode.

Open standards are the base on which to build elections that will be trustworthy, open and creditable. Using consistent data and exchanging that at recognised interface points is essential for trusted elections. EML meets all known election requirements and is the only available public open standard that can meet the needs of elections officials, solution providers, and deliver the services that voters demand.

Resources

CA SoS site to portal interface and example:

http://www.sos.ca.gov/elections/ca_elect_results/result_example.htm

Test site and retrieval with data examples:

http://www.sos.ca.gov/media/

Instructions on EML 510 and EML 530 night of election data feeds:

http://www.sos.ca.gov/media/09ss/election-night-data-feed-user-guide.pdf

CA SoS election results public site: http://vote.sos.ca.gov/

OASIS EML V6.0 specification documents:

OASIS EML Top Ten Reasons for Use brochure:
http://www.oasis-open.org/committees/download.php/30366/EML-Top-Reasons.pdf

OASIS Election & Voter Services Technical Committee

22 August 2009