The Use of XML in Election Results Reporting,

A historical review of Virginia's use of XML in 2008

Paul E. Stenbjorn, M.S.
Information Services Manager, Virginia State Board of Elections

September 2009

Background

Each Election Day candidates, voters, public interest groups, and the media anxiously await publication of official results from the various election administration entities. Many states and municipalities have developed online reporting systems to which election results are posted online. Media outlets race to collect results by calling local and state officials, scouring the internet for results, and through any means available to be first in line with the freshest results.

The 2008 Democratic primary season saw unprecedented interest in election results reporting in a significant number of states. Virginia found itself in the crosshairs of the political and media spotlight during the so-called Potomac Primary of February 19, 2008.

The Virginia State Board of Election had historically presented its election results on its website through the use of a .NET web service which presented data as compiled from localities. These data were collected through its statewide election and registration information system and delivered to the internet via a web launched application developed in concert with Virginia Interactive, the private entity contracted to provide Virginia.gov and other web development technologies.

The Associated Press and other media outlets approached the Virginia State Board of Elections with a request to provide election results data in a real-time data feed. The Virginia State Board of Elections technology staff researched available standards for election results data distribution and discovered that there was no nationally accepted standard for the collection and distribution of election results data. In response, the technology staff developed an ad hoc XML schema loosely based on the internationally accepted EML (election markup language) standard.

The Virginia State Board of Elections generated XML data files and an XSL rendering file to media outlets for the February primary based on several jurisdictional boundaries and developed ETL packages that regenerated these files regularly.

Findings

Inconsistency of Data Presentation

Media outlets and other reporting entities have yet to develop a standard presentation logic for formatting and delivering election results. Television media requested data feeds for their scrawls yet had not functional method for mapping this technology with an XML feed. Many smaller media outlets and interest groups lacked the technical resources to develop applications. In response to the inconsistency of data presentation methodologies, the Virginia State Board of Elections developed associated XSL renderings of election results provided from the feed. Each XSL provided a distinct geographic slice of data (e.g. statewide, locality, precinct).

Taxonomy

With few examples to draw from, the Virginia State Board of Elections developed its own terminology to refer to election results data objects. While there are implicit data types in this development (a name of a candidate will be a text field and number of votes should be an integer), the relationships between and normalization of data became problematic. When examining the prospect of producing XML results for statewide elections, the relationship of election districts to voting precincts added layers of complexity. The Virginia State Board of Elections settled on providing data streams that were either as inclusive as possible (precinct results for each candidate) or as limited as possible (statewide tabulation).

Methodology

The Election Results Markup Language Schema

The Virginia State Board of Elections developed the following structure and schema for election results reporting.

Element : ELECTION

Attributes:
ELECTION_NAME type: string
ELECTION_DATE type: date
PRECINCT_COUNT type: integer
PRECINCTS_REPORTING type: integer

Child Element: LOCALITY

Attributes:
LOCALITY_NAME type: string
LOCALITY_CODE type: string
PRECINCT_COUNT type: integer
PRECINCTS_REPORTING type: integer
CONGRESSIONAL_DISTRICT type: string

Child Element: DISTRICT

Attributes:
DISTRICT_NAME: type: string
PRECINCT_COUNT: type: integer
PRECINCTS_REPORTING: type: integer

Child Element: PRECINCT

Attributes:
PRECINCT_NAME type: string
CONGRESSIONAL_DISTRICT type: string

Child Element: OFFICE

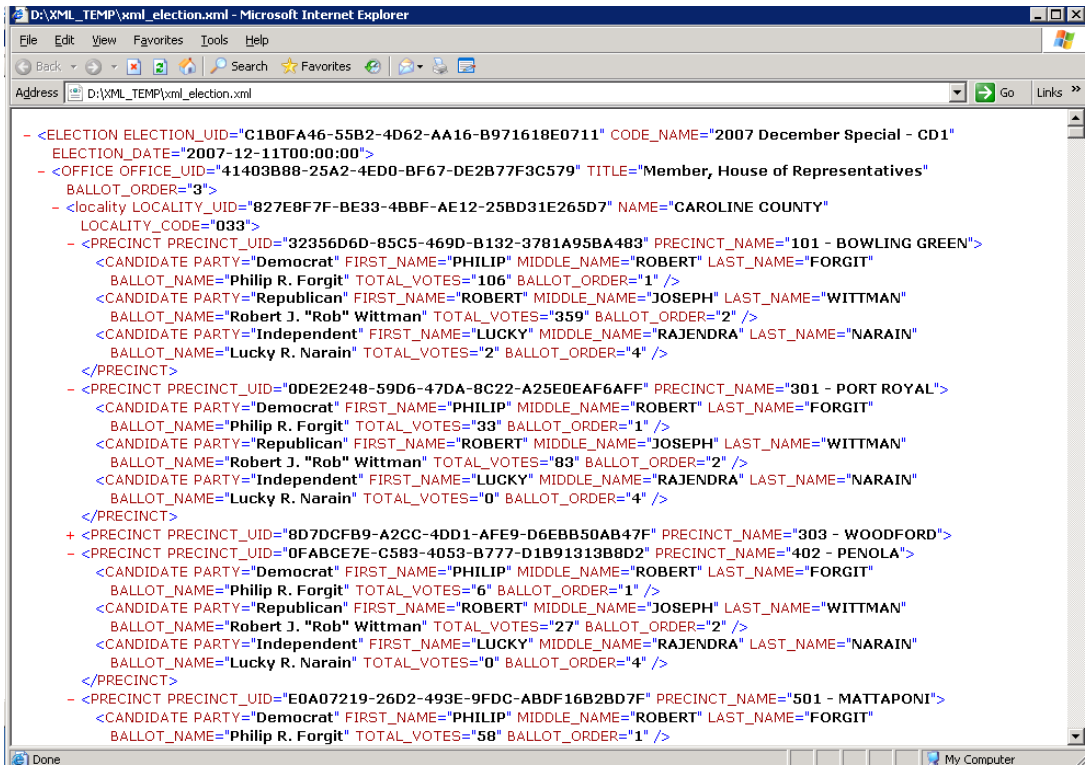
Attributes:
TITLE type: string
BALLOT_ORDER type:integer

Child Element: CANDIDATE

Attributes:
PARTY type: string
FIRST_NAME type: string
MIDDLE_NAME type: string
LAST_NAME type: string
BALLOT_NAME type: string
TOTAL_VOTES type: integer
BALLOT_ORDDER type: integer

Note: All data elements also contained a unique identifier field that allowed for relationships with other data elements within the SBE database and between data elements.

Figure 1-1 Election Schema



Deficiencies of current Schema

As previously mentioned the creation of district based election results adds complexity to the data model and will require the development of a district schema and a relationship model of districts to election results.

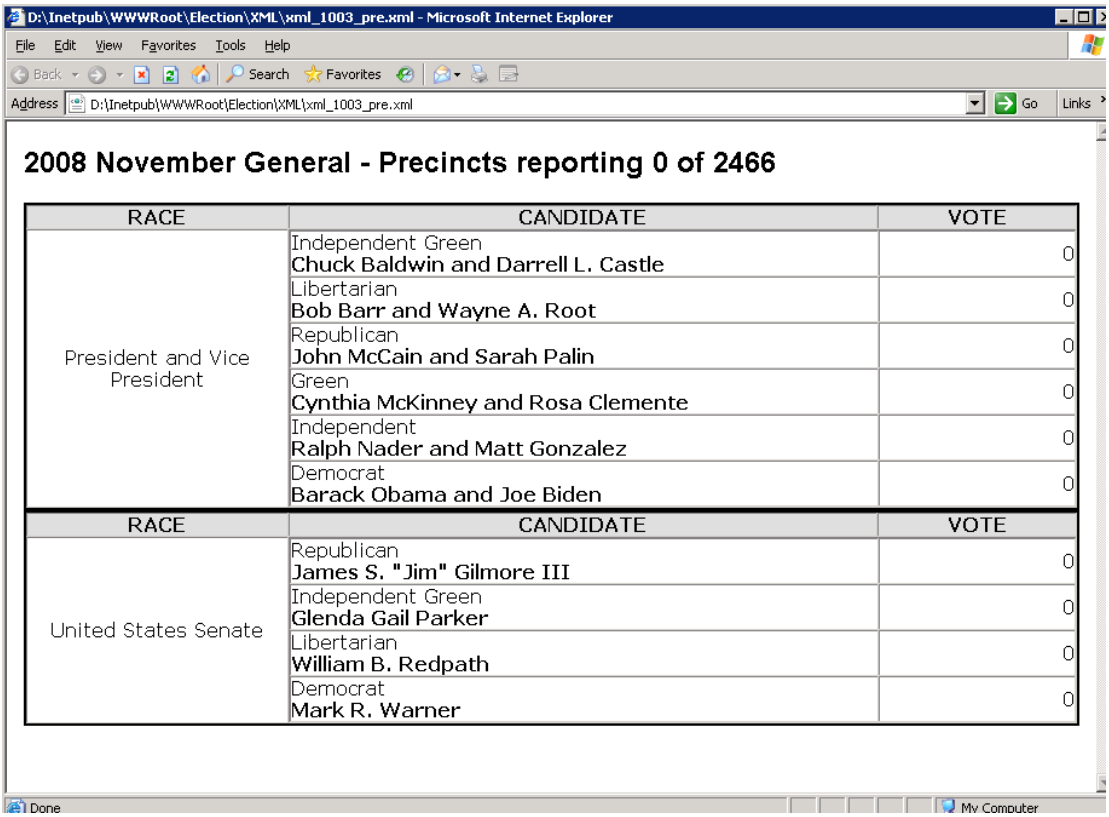
Data Driven XML File Development

The Virginia State Board of Elections technology staff leveraged the state-wide voter registration and elections management system's SQL 2005 database to generate the XML files used in the presentation of data. Using the XML data types available in SQL 2005 and a series of SQL Server Integration Services (SSIS) packages, transact SQL (T-SQL) stored procedures and SQL agent scheduling tasks, we generated XML files published to a webserver every 3 minutes. Initially, files were generated with and posted with placeholder (zeroes) as results to allow news agencies and developers to attach to the files. Problems arose from the frequency of the data refresh and generating redistributable files from the data set. Improving the process to generate virtual files (files that are created from XML data calls) will improve performance.

XSL Presentation Logic

Largely due to the lack of data standards, the Virginia State Board of Elections created XSL sheets to aid in the presentation of XML by reporting entities. These XSL sheets were inherently self limiting as they were designed to present data by specific jurisdictional boundaries. Future development will require the use of more adaptive technologies and the development of online reusable components (e.g. widgets) for the presentation of data uniformly.

Figure 1-2 – presentation logic using XSL



Conclusions

Despite the fairly ad hoc methodology followed by the Virginia State Board of Elections for the presentation of election results data in 2008 via an XML stream, the consumers of the data appreciated and used the results for online presentation. In fact, subsequent elections have led to calls for more election results data to be delivered in this format.

However, until data standards are adopted, each state/reporting entity will need to coordinate its efforts with media outlets independently. Additionally, the scarcity of state funded technology resources familiar with the development and implementation of XML schemas presents a challenge for the maintenance of these efforts.