The Election Markup Language

e-democracy is one of the hottest political topics worldwide. Whilst e-democracy incorporates activities such as using technology to inform the public and collect opinions, the highest profile aspect is the contribution that technology can make to modernising the voting process. Although there had been a gradual increase in the contribution that technology has made to voting over several years, the whole subject hit the headlines following the US Presidential election in 2000. Since then, electronic voting has remained in the headlines, often in connection with "should we?" "shouldn't we?" discussions based on risks both real and imagined.

The Election Markup Language (EML) is being brought into this environment to provide "a standard for the structured interchange of data among hardware, software, and service providers who engage in any aspect of providing election or voter services to public or private organizations". In this article, we will look at what EML is, how it can contribute to elections (whether there is electronic voting or not), where it is being used, and how it addresses some of the security questions.

What is EML?

EML is both more and less than electronic voting. More, in that it defines messages to flow between different parts of the election process all the way from registering electors to counting votes. So an election could be based entirely on postal votes (something that has been tried successfully in the UK), but still use EML in many parts of the process. But EML is less than electronic voting in that its role is only to define messages flowing between systems. It does not define security mechanisms, although it has features that support them; it does not define the transport mechanism or encryption to be used. These must be done separately as part of organizing the election, and will vary for different election types.

EML was developed, and is being maintained, by the Election and Voter Services Technical Committee of OASIS, the Organization for the Advancement of Structured Information Standards. OASIS is a not-for-profit, international consortium that drives the development, convergence and adoption of e-business standards, and is recognized within the e-GIF as a source of standards for use in the UK public sector. Although an international effort by many governments and suppliers of election services and systems, the UK is leading the work through the Office of the e-Envoy, and we were one of the first to pilot use of EML in the 2003 local elections.

At a technical level, EML defines a set of messages to flow within the election process, for example, for an electoral roll system to provide a list of eligible voters to an election system. These messages are based on XML, and are defined in two complementary ways. The first uses the XML Schema language to provide a set of message specifications that are common internationally. These specify, for example that a vote must include the voter's selection of candidate, and may include the voter's identity. A further, and much smaller, set of definitions is used to tailor the specifications for a particular type of election, effectively forming a set of EML profiles for different elections. Whilst a vote in a company AGM will have to identify the voter, a vote in a general election must not. These additional specifications are defined using the Schematron language. Both XML Schema and Schematron are

widely used aspects of XML, and both allow EML messages to be checked automatically to ensure that they meet the specifications.

Where is EML Being Used?

Although EML is still relatively new, an early version was used in the 2002 local elections, and a more complete version in the 2003 local elections. It also forms part of the data interchange standard for the CORE electoral roll project in the UK. Many European countries have expressed an interest and are evaluating its suitability for their elections, and it is likely to be recommended for adoption throughout the countries of the Council of Europe. The Australian government has also been a contributor to the standards, and once fully ratified within OASIS, EML is expected to be mandated for all federal US elections.

What is it about EML that is causing such widespread adoption? After all, most elections are local affairs. One reason for the wide adoption is that EML benefits both suppliers and customers of election systems. For the suppliers, the move to common standards cuts development time and costs. Whilst different elections will continue to need tailoring of systems as long as election systems vary, the amount of tailoring can be reduced as common standards are adopted. For the election organizers, this can mean lower costs. However, more importantly for them, they can now source different parts of the election system from different suppliers. If they have an electoral roll system in place, they can buy an election system from a different supplier without having to define and implement a new interface. Furthermore, although elections may be locally organized, some aspects of the overall election process have a wider base. A European parliamentary election requires results from several areas to be combined into a single European region. EML allows the count from each area to be sent in the same format, making declaring the result a simpler and faster process. Boundary changes become simple as electoral roll systems exchange data in a common format. Overall, the use of standards such as EML benefits everyone as common interfaces between systems speed both development and use.

EML and Election Security

Technology has increased the number of ways in which people can vote, although they fall into two main categories: supervised and unsupervised. Casting a paper vote or using an electronic kiosk at a polling station are both supervised; using an Internet connection from home or a postal ballot are unsupervised. Each provides different security concerns, as does the use of electronic or paper voting methods.

In UK public elections, the main security and audit requirements are are:

- the vote must be secret and anonymous, but traceable under court order
- each voter must only be able to vote once, and only for candidates for offices for which he or she is entitled to vote
- the vote must not be read or altered between the voter and the count
- the voter must not be subject to coercion in the casting of their vote
- it must be possible to demonstrate that all votes cast have been counted

Some of these concerns have a technical solution, others do not. EML provides many security and audit facilities. For example, it allows votes to be "sealed" so that any tampering is evident. It also allows checks to be made that all votes cast have been counted. However, there are some security aspects that cannot easily be countered through technology, although technology can help. For example, for any form of remote voting, it is hard to eliminate coercion by a family member. This applies as much to postal voting as Internet voting or voting by text message. However, systems such as allowing people multiple votes but only counting either the first or last can help. Another threat is the use of Trojan Horses on PCs. Many home computers have been compromised by software that allows keystrokes to be monitored remotely. The risk of these and other threats must be evaluated and managed, but EML is not the complete solution.

A Hypothetical Use of EML

Having gained an understanding of EML and what it does and does not do within the election process, we can look at a hypothetical example of its use. We have already mentioned the CORE project where EML is used to provide common interfaces between electoral roll systems and between them and other systems. We will now look at how EML can be used within an election with multiple means of voting.

Voters might have used EML to register for the electoral roll and to indicate how they would like to vote. So postal voters will get postal ballots and Internet and SMS voters might get instructions relating to their chosen method of voting as part of the information on their poll card (which will have sent to the printer as EML). These voters will also receive something that is termed a "voting token" (or VToken) in EML. This is a credential that indicates the right to vote, but not necessarily the identity of the voter. This could be as simple as a PIN and password. A link between the voter and credential might be stored securely in case it is needed later for audit purposes.

In many cases of electronic voting, the link from the voter, such as a mobile phone for SMS voting, will not support EML, or even XML. In these cases a gateway is used to manage the conversation with the voter using suitable protocols, using EML to communicate with the election system. The gateway will seal the vote and keep a record of the number of votes it has sealed. EML is used in messages to authenticate the voter, present a ballot, collect the votes and send them to a counting machine. The output of this count is then sent, with the output of other machines such as one with an optical scanner used for electronic counting of paper ballots, for collation. Meanwhile, the number of votes sealed can be compared to the number of sealed votes counted as part of the audit process.

That was a very rapid run through a hypothetical election. EML is designed to cope with many more types of election and referendum, including the use of various voting methods such as first past the post, party lists and the additional member system.

More information can be found on the OASIS web site at http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=election.

Paul Spencer is a Director of Boynings Consulting Ltd and a Consultant on the use of XML in Government. He led the development of the XML definition of the Election Markup Language for OASIS on behalf of the Office of the e-Envoy.

First draft for comment 23/3/2004
paul.spencer@boynings.co.uk