IEEE P1622 Standards Development Strategy

This document lays out a strategy of developing a series of narrowly-focused use case/standards to comprise a comprehensive standard for a common data format for election systems, under the auspices of IEEE Project P1622, Voting Systems Electronic Data Interchange.

Overview

The Election Assistance Commission (EAC) and its Technical Guidelines Development Committee (TGDC) have requested that NIST identify a common data format (CDF) for election system data for inclusion in a VVSG. The EAC desires that CDF facilitate interoperability and component certification of voting devices and believes that an interoperable format will facilitate automated testing of voting devices. As part of its voting program, NIST is undertaking a project whose goal is to work with IEEE to create an IEEE standard for a CDF for election systems.

The scope of the data covered by the common data format work includes:

- Exports of voter registration information from state voter registration databases
- Voter registration data and precinct information processed by electronic pollbooks
- Candidate and ballot definition information
- Voted ballot information and tabulations
- Device logs
- Other information that may be produced by election management systems

NIST is working with the IEEE Project P1622 working group, in conjunction with the Organization for the Advancement of Structured Information Standards (OASIS), to arrive at a CDF standard by calendar years 2012-13. NIST is also consulting with the TGDC working group for comments on CDF-related requirements and to assist in prioritizing development of CDF use case standards for scopes of election activities. IEEE P1622 is using OASIS EML (Election Markup Language) as the basis for the development of its CDF standard.

Background

In June 2001 the IEEE Standards Association inaugurated a new standards project tasked with the development of standards for voting equipment, known as subcommittee P1583. A closely related subcommittee known as P1622 was begun for definition of a data interchange standard for voting systems. P1622 was approved on July 30, 2002 with the following scope and purpose:

This project will develop electronic data interchange formats to be used by components of the voting system for exchange of electronic data. Data interchange specifications for voting systems may include information used to define and create ballots, as well as reporting of voting activity and results.

Jurisdictional and geo-political information is included in that information, as are contest definitions, voter registration, and voting area (a.k.a., precinct) information. All such data is included in the scope of this project. The project scope does not include implementation/vendor specific information.

P1622 produced a draft standard in 2007 that did not advance beyond the subcommittee, and the subcommittee was inactive for the subsequent next two years. In October 2009, the P1622 reformed with new membership under the same scope and purpose. NIST agreed to become an active member and to work towards a CDF standard.

NIST, in 2007, produced the draft TGDC VVSG Recommendations, now known colloquially as "VVSG 2.0." It contained several requirements specifying that manufacturers must document the data formats they use and that manufacturers should use a common format. In the December 2009 TGDC meeting, the EAC tasked the TGDC with completing work on a CDF for inclusion in the VVSG 2.0.

At a NIST and EAC-sponsored P1622 meeting in February, 2011, OASIS agreed to work within the IEEE framework for identifying a CDF standard and agreed to make changes to EML as necessary to enable it work well with U.S. manufacturers and election requirements. EML is an international standard that needs to be tailored for local environments, e.g., the US election environment. P1622 subsequently decided to base further CDF development on OASIS EML, with the intention that OASIS and IEEE remain consistent and do not diverge.

Use case standard strategy for CDF development

The prior incarnation of P1622 attempted to produce a comprehensive standard but was unable to achieve working group consensus and to subsequently proceed further with balloting. NIST and P1622 ~~have~~ decided upon a different strategy, therefore, of identifying and agreeing upon use case~~s~~ standards for certain scopes/applications for election data formats, with the goal being that the sum of use case standards will constitute a comprehensive common data format for election systems in general.

Briefly, a P1622 use case standard includes the following sorts of information:

- Summary information about the application and Its associated stakeholders and system actors,
- An informative model of the data, possibly expressed in UML,
- Normative schemas in XML,
- Identification of requirements for the election data fields involved in the application, and
- Any associated issues and constraints.

Each IEEE standard requires its specific Project Authorization Request (PAR), which identifies aspects of the standard to be developed including the scope and purpose of the standard. For an "umbrella" of P1622 standards, a series of PARs could be developed, one for each use case standard, e.g.,

1. P1622.1 – UOCAVA Electronic Blank Ballot Distribution
2. P1622.2 – Voter Registration Database Export Use Case Standard
3. P1622.3 – Election Reporting Use Case Standard
4. P1622.4 – Election Auditing Use Case Standard

5. …

Ultimately, a final comprehensive standard could be developed to tie the other standards together.

The use case standard strategy should allow for more rapid deployment of a CDF. Device-to-device interoperability, using a CDF in part as the means, will be difficult and require the involvement and cooperation of many parties. Given this, the strategy for use case standard development involves addressing the "easier" aspects of a CDF that involve the least amount of device-to-device interoperability and have the highest odds of early success, and then subsequently working towards those aspects that are more difficult but achieve more device-to-device interoperability. Addressing these standards in parallel as opposed to serially allows more flexibility and capability to take advantage of external assistance or collaboration with other interested parties or coalitions.

Note: At the February 2011 meeting, the Federal Voting Assistance Program (FVAP) and the EAC, because of their need to respond quickly to the Military and Overseas Voter Empowerment (MOVE) Act, requested that P1622's first priority be to develop a CDF standard for the scope of election data to support electronic distribution of blank ballots to overseas military and citizens. Thus, the first P1622 CDF standard addresses only those data elements needed for electronic distribution of blank ballots. It is proposed that this standard be numbered as the first P1622.x standard.

Four stages of use case standard development

Accordingly, four stages of use case standard development have been defined, focusing on the endpoints of the voting system at level one and working towards greater aspects of device interoperability at levels 2, 3, and 4[1]. The following sections identify use case standards for each of the four levels, with a cross reference to the relevant EML schemas. In terms of development priority, level one use case standards would come first, followed in general by the other levels.

Level one use cases standards:

Level one focuses on inputs into the voting system and exports from the voting system and could be considered as to address the "low hanging fruit," that is, data that does not necessarily require device interoperability and that is relatively straightforward to place in a CDF. The devices addressed are voter registration databases (VRDB) and election management systems (EMS), both of which generally permit imports/exports in a comma-separated value (CSV) format. Translators could thus be built from CSV to EML.

1. **VRDB export** - Registered voter data from VRDBs is currently exported into EMSs; VRDBs are also sometimes updated via exports from electronic poll books and other devices within the voting system. This use case standard would focus specifically on the EML 300 and 200 series

---

[1] The 4 stages of use case standard development are based directly on analysis performed by Ian Piper of Dominion Voting, available at http://www.nist.gov/itl/vote/ieee-p1622-bbd-use-case.cfm.

schemas needed for registered voter information and other related information that may be stored, e.g., status information associated with a voter's ballot.

2. **Election results data (state roll-ups)** - The EMS currently exports election results data that is reported to central offices and local jurisdictions and generally reported upward within a state (known as a state roll-up). From there, this information can be used nationally and by other organizations to analyze results of contests. This use case standard would focus specifically on the EML 510, 520, 530, and 110, 210, and 470 schemas needed for election reporting.

3. **Election auditing** – An EMS can provide election results data, ballot cast records and audit event logs to facilitate the auditing, analysis and possible forensic investigation of an election. The information would be available through data exports provided in the EMS. Note: Current EMS already provide this audit information through either data exports and/or reports (electronic or printed), but the formats vary from system to system. This use case standard would focus on the EML 500 series schemas.

Level two use cases standards:

Level two addresses data that is more complicated to export in some cases, e.g., event log data or cast vote record data, as well as data that may allow limited device interoperability, e.g., blank ballot information, minus state-specific formatting, that can be used by ballot distribution systems (BDS).

1. **VRDB and EMS export of information to build electronic blank ballots minus formatting details[2]** – An EMS can export data which includes the following information:

   - election header,
   - contest list,
   - contest selections,
   - contest 'vote for' quantity, and
   - ballot/precinct IDs.

   This ballot information would not provide the contest order, rotation, layout or formatting information. Those items are part of the Level 3 Interoperability. At Level 2, if these items were legally required by a voter's jurisdiction, they would have to be manually loaded into the BBDS for ballot printing or ballot presentation to the voter. With the information in the above list, ballots can be laid out on a display screen, marked electronically, and printed or printed on paper and marked manually. In either case, for the cast ballot selections to be entered into the EMS election results, they would have to be manually transferred to an actual scannable ballot (then scanned, tabulated and uploaded), or the selections could be manually entered into the EMS. This use case standard focuses on the EML 505, 110, 230, 330, 410, and 470 schemas.

2. **EMS and voting device export of event log and other auditing information** – A possible route to placing event logs in a CDF would be to utilize the NIST Event Management Automation Protocol

---

[2] This use case standard is in final review, but included here for completeness.

(EMAP)[3] XML schemas and add them to EML. It would also be beneficial to have a common lexicon for audit log entries through which an analysis of an audit log can be done more efficiently and effectively. This may require invention of new EML schemas as well as use of the EML 480.

3. **EMS export of cast vote records** - An EMS can provide election results data (from Level One), ballot cast records and audit event logs to facilitate the auditing, analysis and possible forensic investigation of an election. The information would be available through data exports provided in the EMS. This use case standard focuses on the EML 400 series schemas.

Level three use cases standards:

Level three focuses specifically on EMS export of information to dynamically build blank ballots with state-specific formatting, known as ballot definition data. This would permit voting devices to accept this information and display ballots according to state requirements. Practically speaking, the state-specific information might be contained in templates or configuration files that may work in tandem with a CDF, but if made part of the CDF specification, this would permit interoperability among EMSs and voting devices. Level three can be broken out into three segments.

1. The first segment would be the information required for generating an electronic ballot and would also be a portion of the data required for printing a physical paper ballot (note that some of this information is already covered in the level two use case standard for electronic blank ballot distribution). The information would include:

    - election header,
    - contest list/order,
    - contest selections/rotations,
    - contest 'vote for' quantity,
    - endorsements/cross-endorsements,
    - straight party/recall contest associations, and
    - ballot/precinct IDs.

    With this information, ballots can be laid out on a display screen and marked electronically or printed and marked manually. In either case, for the cast ballot selections to be entered into the EMS election results, they would have to be manually transferred to an actual scannable ballot (then scanned, tabulated and uploaded), or the selections could be manually entered into the EMS.

2. The second segment would be the information required for printing a scannable ballot. This information would include the first segment and also the:

    - voting target shape/size/color/position,
    - timing/control/special marks and their shape/size/color/position, and
    - ballot/precinct ID coding and their shape/size/color/position.

---

[3] http://scap.nist.gov/emap/

3. The third segment is information that characterizes the appearance of the ballot. This information would include items such as:

- font types/sizes/color,
- background color/watermarks/color striping, and
- instructional text.

Some jurisdictions have state statutes and/or regulations that mandate some if not all of these characteristics. For those jurisdictions, the ballots presented to their voters (and returned to them for counting) would have to include this information, whether it be presented on paper and manually marked or presented on a display and electronically marked.

Level four use cases standards:

Level four involves EMS export of other voting device configuration data (memory card configuration, other device initialization information) to achieve interoperability among voting devices in general. At the base level within the EMS, the machine configuration data that is downloaded to the vote capture devices (e.g., DREs) could be provided to vote capture devices from other voting systems. This information would provide all the detail required for the machine's operation, including:

- Memory media ID
- Election header information (title, date, version, poll center ID, copy number)
- Security and verification data (signatures, check counter values, serial numbers)
- Election Data (counter groups, voter groups, base precincts, district categories, languages, contests, candidates, rotations, relationships, endorsements, 'vote for' number, ballot style IDs, voting box type/size/position/justification, sequencing index, report format)
- Rejection criteria for overvotes, undervotes, blank voted contests, and blank voted ballots
- Ballot sorting options
- Cast vote records (data and/or images), tally results
- Modem upload phone number
- Audit logs

For DRE display interface devices, the information would also include:

- Header/footer sizes, number of columns, scaling factors
- Button type/size/position/justification
- Flags for voting, rendering, text wrapping
- Background colors for page/labels/contests/candidates
- Instructional text, write-ins, audio, and default volumes

Although the downloaded memory devices would be interchangeable between VCDs, only those memory devices created by one EMS for an election would be allowed to upload to that EMS.

Next Steps

The next step is to formalize this strategy within P1622 and create from it a "unifying" draft standard whose purpose is to provide structure to the use case standards. The unifying standard would include the terminology used in the use case standards and would be updated as each use case standard is developed. It may be useful to locate the terminology on the P1622 site and make it available to others in the community so that we can achieve some consistency in usage.

Following that, individual PARs for each use case standard need to be developed, with the following four use case standards as the first priority:

1. Logging system and election events
2. Election reporting
3. VRDB export
4. Audit information export

Summary table

Following is a summary of use case standards, linked to the corresponding EML schemas.

Level	Use Case Standard Name	Associated EML Schemas
1	VRDB export	200, 300 series
1	Election results export	110, 210, 470, 510, 520, 530
1	Election audit export	500 series
2	Blank ballot export minus formatting details	505, 110, 230, 330, 410, 470
2	Event log export	480, TBD new
2	Cast vote record export	400 series
3	Ballot definition file export	TBD new
4	Voting device configuration data	TBD new