ACCURATE
A CENTER FOR CORRECT, USABLE, RELIABLE,
AUDITABLE, AND TRANSPARENT ELECTIONS

PUBLIC COMMENT ON THE VOLUNTARY VOTING SYSTEM GUIDELINES, VERSION II (FIRST ROUND)*

Submitted to
The United States Election Assistance Commission

May 5, 2008

*This material is based upon work supported by the National Science Foundation under A Center for Correct, Usable, Reliable, Auditable and Transparent Elections (ACCURATE), Grant Number CNS-0524745. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. This public comment narrative was prepared by Aaron Burstein and Joseph Lorenzo Hall of the Samuelson Law, Technology and Public Policy Clinic along with comments from the Principal Investigators and Advisory Board Members of the NSF ACCURATE Center.

ACCURATE Principal Investigators

Aviel D. Rubin
ACCURATE Director
Department of Computer Science
Johns Hopkins University
rubin@cs.jhu.edu
http://www.cs.jhu.edu/~rubin/

Dan S. Wallach
ACCURATE Associate Director
Department of Computer Science
Rice University
dwallach@cs.rice.edu
http://www.cs.rice.edu/~dwallach/

Dan Boneh
Department of Computer Science
Stanford University
dabo@cs.stanford.edu
http://crypto.stanford.edu/~dabo/

Michael D. Byrne
Department of Psychology
Rice University
byrne@rice.edu
http://chil.rice.edu/byrne/

David L. Dill
Department of Computer Science
Stanford University
dill@cs.stanford.edu
http://verify.stanford.edu/dill/

Douglas W. Jones
Department of Computer Science
University of Iowa
jones@cs.uiowa.edu
http://www.cs.uiowa.edu/~jones/

Deirdre K. Mulligan
School of Law
University of California, Berkeley
dmulligan@law.berkeley.edu
http://www.law.berkeley.edu/faculty/profiles/facultyProfile.php?facID=1018

Peter G. Neumann
Computer Science Laboratory
SRI International
neumann@csl.sri.com
http://www.csl.sri.com/users/neumann/

David A. Wagner
Department of Computer Science
University of California, Berkeley
daw@cs.berkeley.edu
http://www.cs.berkeley.edu/~daw/

Brent Waters
Computer Science Laboratory
SRI International
bwaters@csl.sri.com
http://www.csl.sri.com/users/bwaters/

Preface

A Center for Correct, Usable, Reliable, Auditable and Transparent Elections (ACCURATE),[1] a multi-institution, multidisciplinary, academic research project funded by the National Science Foundation's (NSF) "CyberTrust Program,"[2] is pleased to provide these comments on the Voluntary Voting System Guidelines to the Election Assistance Commission (EAC). ACCURATE was established to improve election technology. ACCURATE conducts research investigating software architecture, tamper-resistant hardware, cryptographic protocols and verification systems as applied to electronic voting systems. Additionally, ACCURATE is evaluating voting system usability and how public policy, in combination with technology, can better facilitate voting nationwide.

Since receiving NSF funding in 2005, ACCURATE has made a number of important contributions to the science and policy of electronic voting.[3] The ACCURATE Center has published groundbreaking results in cryptography, usability, and verification of voting systems. ACCURATE has also been actively contributing to the policy discussion through regulatory filings, through testimony and advising decisionmakers as well as conducting policy research.[4] ACCURATE researchers have participated in running elections and assisting election officials in activities such as unprecedented technical evaluation of voting systems and redesigning election procedures.[5] Finally, the education and outreach mission of ACCURATE has flourished through the development of numerous undergraduate and graduate classes and the creation of the premier venue for research involving voting systems.[6]

With experts in computer science, systems, security, usability, and technology policy, and knowledge of election technology, procedure, law and practice, ACCURATE is uniquely positioned to provide helpful guidance to the EAC as it attempts to strengthen the specifications and requirements that ensure the functionality, accessibility, security, privacy and equality of our voting technology.

We welcome this opportunity to further assist the EAC and hope this process continues the collaboration between the EAC and independent, academic experts in order to sustain improvements in election systems and procedures.

---

[1] *See:* http://www.accurate-voting.org/

[2] National Science Foundation Directorate for Computer & Information Science & Engineering, Cyber Trust, at http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=13451&org=CISE.

[3] *2006 Annual Report*. A Center for Correct, Usable, Reliable, Auditable and Transparent Elections, January 2007 ⟨URL: http://accurate-voting.org/wp-content/uploads/2007/02/AR.2007.pdf⟩; *2007 Annual Report*. A Center for Correct, Usable, Reliable, Auditable and Transparent Elections, January 2008 ⟨URL: http://accurate-voting.org/wp-content/uploads/2008/01/2007.annual.report.pdf⟩

[4] *List of ACCURATE Testimony*. ACCURATE Website ⟨URL: http://accurate-voting.org/pubs/testimony/⟩; Aaron J. Burstein, Joseph Lorenzo Hall and Deirdre K. Mulligan, *Public Comment on the Manual for Voting System Testing & Certification Program (submitted on behalf of ACCURATE to the U.S. Election Assistance Commission)*. October 2006 ⟨URL: http://accurate-voting.org/wp-content/uploads/2006/11/ACCURATE_VSTCP_comment.pdf⟩.

[5] ACCURATE researchers have participated in the comprehensive voting system evaluations sponsored by the States of California and Ohio. We reference these in Section 2.

[6] For more on our educational output, please see those sections of our Annual Reports (see note 3). The joint USENIX/ACCURATE Workshop on Electronic Voting Technology (EVT), colocated with the USENIX Security Symposium, was started in 2006 and continues to attract high caliber voting technology research. See: http://www.usenix.org/event/evt08/.

Contents

1 Introduction and Background

The current draft of the Voluntary Voting System Guidelines (VVSG) aptly identifies the properties that a voting system should embody: fairness, accuracy, transparency, security, accessibility, verifiability and timeliness. Experience with electronic voting systems has demonstrated that the requirements and testing in previous standards and guidelines were unable to produce systems that exhibit all of these properties. As ACCURATE pointed out in its comments on the 2005 VVSG, only requirements written with an understanding of how they will affect design, testing, and implementation are likely to lead to real systems that embody these properties.[1] Two of the main recommendations in those comments were (1) that the EAC adopt guidelines that create requirements reflecting the state of the art in specific disciplines, rather than relying on functional testing; and (2) that the guidelines provide mechanisms to incorporate experience with fielded systems into revisions of the requirements.

We are pleased to find that the current draft of the VVSG takes significant steps toward adopting these approaches. The result is a set of guidelines that present detailed, coherent requirements for voting system security, usability, accessibility, and auditability. Moreover, the current draft would help make data available to conduct ongoing examinations of several important facets of voting systems. Put together with the EAC's Voting System Test Laboratory Accreditation Program, Voting System Certification Program, and the EAC's development as a clearinghouse for research and reports on many aspects of voting systems, the draft guidelines will form part of a system that will help create and maintain the trustworthiness of voting systems in the United States.

A fundamental insight that underlies the draft is that voting technologies are so complex that it is not realistic to definitively establish that a given device or system conforms to a certain high-level property, such as security or usability. As we discuss throughout these comments, the VVSG draft contains a number of innovative ways of handling this complexity. With respect to security, the concept of software independence provides a groundbreaking framework for requirements that should prevent undetected changes in voting technology from affecting the outcome of an election. In Section 2 we discuss how this framework ties together requirements for security, auditing, accessibility, and documentation. Sections 3 and 4 explain how the VVSG draft significantly improves upon previous guidelines in terms of taking voting system complexity into account when setting requirements for security, reliability, and usability testing. Nevertheless, further improvements are needed. Section 5 highlights how changes in documentation requirements will lead to voting systems submissions that test labs can more easily evaluate and documentation that election officials and pollworkers can more easily use. Finally, Section 6 outlines ways in which the EAC can lend ongoing institutional support to ensure that the VVSG incorporates feedback from the field as well as changes in the several disciplines that must inform voting system design.

2 Software Independence is Critical for the Future of Voting System Certification

Software independence is one of the major conceptual advances in the current draft of the VVSG.[2] As the definition in Part 1:2.7 states, "software independence means that an undetected error or fault in the

[1] *Public Comment on the 2005 Voluntary Voting System Guidelines*. A Center for Correct, Usable, Reliable, Auditable and Transparent Elections (ACCURATE), September 2005 ⟨URL: `http://accurate-voting.org/accurate/docs/2005_vvsg_comment.pdf`⟩.

[2] The initial formulation of software independence was given by Rivest and Wack: Ronald L. Rivest and John Wack, *On the Notion of "Software Independence" in Voting Systems*. National Institute of Standards and Technology HAVA Technical Guidelines Development Committee, July 2006 ⟨URL: `http://vote.nist.gov/SI-in-voting.pdf`⟩.

voting system's software is not capable of causing an undetectable change in election results."[3] Though this definition may appear to be rather abstract, it addresses a broad array of practical problems facing electronic voting systems. To see why this is so, we discuss in this section what software independence does and does not do.

Software independence represents a general, flexible requirement to counter a problem that any electronic voting system is likely to encounter: The software, hardware, and other technologies[4] necessary to support election activities are typically so complex that it is effectively impossible to verify their correct operation by either formal proofs or testing.[5] Moreover, even if the logic in voting devices could be fully verified in a test lab, it would still be necessary to ensure that the hardware, software and firmware used in actual elections are identical to the systems that were tested. While the VVSG draft sets forth important improvements in testing requirements that support this assurance, improved testing alone will never be able to replace software independence as a security measure.

The underlying premise of the software independence approach is that, no matter how hard one looks for errors or faults in voting system software, there is no way to guarantee that one has found them all. Even if no errors or faults are found, there is no way to guarantee that none exist.

Software independence provides a conceptual framework to ensure that accidental programming errors do not affect the outcome of an election, as well as to detect intentionally introduced malicious software. Examples of accidental programming errors in voting systems are legion. For example, in November 2004, 4,400 votes were permanently lost after DREs in Carteret County, North Carolina exceeded their vote storage capacity without alerting voters or pollworkers.[6] Far more subtle issues arising from programming errors have also been found. During a volume test of DREs in California, for example, testers found that voters with long fingernails who used a dragging motion on the touch screen could cause the device to crash.[7] Both incidents illustrate the risks of recording votes on a single electronic device.

Of course, numerous studies have shown that currently deployed voting systems are susceptible to undetectable malicious attacks. The voting systems produced by all four manufacturers with significant market share in the United States have been subjected to thorough batteries of adversarial testing, source code review, accessibility testing and documentation review.[8] All of these systems have vulnerabilities

[3] To clarify that software independence applies to any number of errors or faults, the Commission might consider changing the definition to read: "software independence means that undetected errors or faults in the voting system's software are not capable of causing an undetectable change in election results."

[4] As a November 2006 NIST staff discussion draft on software independence noted, the phrase " '[s]oftware independence' should be interpreted to really mean *complex technology independence*" to include software implemented in hardware, such as programmable read-only memory and circuit boards. *See:* National Institute of Standards and Technology, *Requiring Software Independence in VVSG 2007: STS Recommendations for the TGDC*. November 2006 ⟨URL: `http://vote.nist.gov/DraftWhitePaperOnSIinVVSG2007-20061120.pdf`⟩

[5] Note, however, that recent research has shown that it is possible to starkly reduce the scope of what one must trust in a voting system. See, for example, Ka-Ping Yee, *Building Reliable Voting Machine Software*. Ph. D thesis, University of California, Berkeley, 2007, ⟨URL: `http://zesty.ca/pubs/yee-phd.pdf`⟩; Ronald L. Rivest and Warren D. Smith, Three-VotingProtocols: ThreeBallot, VAV, and Twin. In Proceedings of the Second Electronic Voting Technology Workshop (EVT). August 2007 ⟨URL: `http://www.usenix.org/events/evt07/tech/full_papers/rivest/rivest.pdf`⟩.

[6] *More than 4,500 North Carolina Votes Lost Because of Mistake in Voting Machine Capacity*. USA Today (Associated Press), November 2004 ⟨URL: `http://www.usatoday.com/news/politicselections/vote2004/2004-11-04-votes-lost_x.htm`⟩.

[7] David Jefferson et al., *Lessons from the Diebold TSx "sliding finger" bug (unpublished)*. Oct 2005.

[8] *Software Reviews and Security Analyses of Florida Voting Systems*. Florida State University's Security and Assurance in Information Technology Laboratory, February 2008 ⟨URL: `http://www.sait.fsu.edu/research/evoting/index.shtml`⟩; Patrick McDaniel et al., *EVEREST: Evaluation and Validation of Election-Related Equipment, Standards and Testing (Academic Final Report)*. December 2007 ⟨URL: `http://www.sos.state.oh.us/sos/info/EVEREST/14-AcademicFinalEVERESTReport.pdf`⟩; *Top-To-Bottom Review of California's Voting Systems*. California Secretary of State, March 2007 ⟨URL: `http://www.sos.ca.gov/elections/elections_vsr.htm`⟩; Ariel J. Feldman, J. Alex Hal-

that could relatively easily be exploited to alter the results of an election. These studies demonstrate that individual vote-capture devices as well as central-count systems are susceptible to attacks that could lead to undetected changes in election results.

The usefulness of software independence is also evident in situations in which the presence of a voting system fault is a matter of dispute. In the November 2006 election for a representative from Florida's 13th Congressional District, an unusually high proportion of votes cast on paperless DREs in Sarasota County recorded no vote for this race. Subsequent litigation, academic and government studies, and public debate explored whether ballot design, miscalibration, software errors, or some other cause (e.g., voters choosing not to vote) was responsible for the undervotes in this race. Though the official study of this election by the Government Accountability Office "did not identify any problems that would indicate that the machines were responsible for the undervote,"[9] others have pointed out that the scope of this study was too narrow to rule out miscalibration and other hypotheses.[10] In any event, this investigation, which took more than half of the Congressional term to bring to a conclusion, was likely prolonged, and the controversy intensified, by the fact that the voting devices in question did not produce a record of votes that was independent of those recorded by the DREs.

It is against this background—unreliability in the field; the prospect of undetectable, malicious attacks; and the inconclusiveness of post-election analysis in purely electronic systems—that the EAC should view the software independence requirement. Software independence is flexible enough to realistic assumptions about voter behavior. Some voters might neglect to inspect the independent records that some software-independent voting systems (e.g., DREs with a voter-verifiable paper audit trail [VVPAT] and optical scan systems) produce.[11] Others might be unable to do so because of visual impairment or other disabilities. In both cases, however, software independence is still achievable. The point of software independence is not that each voter must be able to verify that his or her selections are captured accurately by two independent channels. Instead, software independence requires that any change in the vote record that is counted is detectable at some point.

For example, in an optical scan system (perhaps used in conjunction with an electronic ballot marking device), software independence would not require that each voter be able to verify that the scanner correctly interprets and records the marks on his or her ballot. Instead, properly designed and executed post-election recounts of optically scanned paper ballots can expose errors in the machine tally. This independent check on election results supports the software independence of optical scan systems.

A larger scheme of routine post-election audits and technical requirements for records to support such audits are integral to achieving software independence. Many other sections of the VVSG draft provide these supporting technical requirements.[12] In particular, the current draft's requirements for

derman and Edward W. Felten, Security Analysis of the Diebold AccuVote-TS Voting Machine. In Proceedings of USENIX/ACCURATE Electronic Voting Technology Workshop. August 2007 ⟨URL: `http://www.usenix.org/events/evt07/tech/full_papers/feldman/feldman.pdf`⟩

[9] U.S. Government Accountability Office, *Results of GAO's Testing of Voting Systems Used in Sarasota County in Florida's 13th Congressional District (Statement Before the Task Force for the Contested Election in the 13th Congressional District of Florida, Committee on House Administration, House of Representatives)*. February 2008 ⟨URL: `http://www.gao.gov/new.items/d08425t.pdf`⟩.

[10] Verified Voting Foundation, *GAO Report Not a Clean Bill of Health for Voting Machines: Limited Scope Investigation Not Conclusive*. February 2008 ⟨URL: `http://www.verifiedvotingfoundation.org/downloads/VVF-Statement-GAO.pdf`⟩.

[11] Sarah P. Everett, *The Usability of Electronic Voting Machines and How Votes Can Be Changed Without Detection*. Rice University PhD Thesis, May 2007 ⟨URL: `http://chil.rice.edu/alumni/petersos/EverettDissertation.pdf`⟩.

[12] Specifying audit procedures, on the other hand, would be outside the scope of the VVSG. Still, given the increasing number of states that require routine post-election audits and the prospect of a federal audit requirement, audits are a crucial piece of election administration that the VVSG should address. For a review of state audit laws and recent scholarly work on post-election audits, see Lawrence Norden et al., *Post-Election Audits: Restoring Trust in Elections*. Brennan Center for Justice at The New York University School of Law and The Samuelson Law, Technology and Public Policy Clinic at the University

an audit architecture (Part 1:4.2), vote and report data exchange (Part 1:6.6-B), and for independent voter-verifiable records (IVVR) (Part 1:4.4) would help ensure voting systems produce records that support audits designed to detect discrepancies between two independent sources of an election tally. (See Section 2.1 for more extensive comments on the VVSG draft's treatment of voting system auditing architecture.) An example of such an architecture is the combination of electronic records and VVPAT records from a DRE-VVPAT system. The current draft also leaves room for new technologies to improve upon or replace current systems; though the draft specifies that providing an IVVR is one way that a voting system may achieve software independence, it does not require this approach. The innovation class (Part 1:2.7) would allow other approaches to be recognized as software independent.[13]

Still, though a requirement of software independence is necessary to guard against changes in the outcome of an election, it is not, by itself, sufficient to guard against all instances in which a voter's intended ballot selections differ from those that are actually cast. In particular, a software independence requirement does not supplant the need for broader software reliability requirements and testing. For example, software that occasionally causes a DRE system to skip a page of the ballot could cause undervotes in the contests on that page, but the two records of the vote would not show a discrepancy.[14] Or voting system software might run more slowly, or crash more frequently, once a specific candidate is chosen.[15] These types of errors are not readily addressed within the software independence framework; the reliability, usability, and accessibility testing requirements that we address later in these comments are necessary complements to software independence.

To summarize, the software independence requirements are integral to the overall structure of the current VVSG draft. Software independence represents a well-defined objective for the trustworthiness of elections conducted using highly complex, electronic voting devices. It provides a framework to greatly increase the likelihood of detecting changes in election results caused by software errors, relative to formal testing and analysis of these systems. Finally, many other requirements in the VVSG draft support software independence, and their full utility is achieved when they are tied to the overarching requirement of software independence.

We would like to reiterate that testing and analysis alone will never be able to confirm correct operation of voting systems, and therefore cannot replace software independence as an accuracy, integrity and security measure.

2.1 The Requirements for Software Independence and Auditing Architecture Are Intimately Related

The VVSG draft emphasizes and articulates the importance of post-election audits. A core requirement of software independence obliges voting systems to recover from software failures. The methods for recovery currently contemplated by the draft VVSG involve auditing; that is, checking or counting, often by hand, audit records via a means independent of the voting system. ACCURATE researchers have long recognized the importance of auditing elections.[16] Fortunately, most states require or have

of California, Berkeley School of Law (Boalt Hall), 2007 ⟨URL: `http://www.brennancenter.org/dynamic/subpages/download_file_50227.pdf`⟩

[13] We comment in detail on the innovation class in section 6.1.

[14] Yee (as in n. 5), pages 181-185 discusses this and other examples in greater depth.

[15] *See id.*

[16] Peter G. Neumann, Risks in Computerized Elections. Communications of the ACM, 33 November 1990. For more recent commentary and research from ACCURATE on audits, see: David Dill and Joseph Lorenzo Hall, *Testimony: Post-Election Audits of California's Voting Systems*. The California Secretary of State's Post-Election Audit Standards (PEAS) Working Group, July 2007; Arel Cordero, David Wagner and David Dill, The Role of Dice in Election Audits—Extended Abstract. IAVoSS Workshop on Trustworthy Elections 2006 (WOTE 2006), June 2006 ⟨URL: `http://www.cs.berkeley.edu/~daw/papers/dice-wote06.pdf`⟩; Rebekah Gordon, Elections Office Gets Tips from Experts. San Mateo County Times, November 2006

procured voting systems that produce audit trails.[17] In this section, we highlight how the VVSG draft establishes requirements to help ensure that audit records support the goal of auditability.

It is essential that national-level requirements specify a basis for auditing that all voting systems must support. Well-specified audit support requirements applied at the national level will ensure that voting systems can support a wide variety of auditing schemes. This will help to guarantee that voting systems will have the capacity to support new methods of conducting audits in the future as new laws are adopted and new audit methods are vetted by the scientific community. In terms of forensic capability, the draft VVSG audit requirements appropriately require voting systems to capture and keep evidence of error or possible fraud, at an appropriate level of granularity.

First, we comment generally on the term "post-election audit". In general, election auditing encompasses checking for agreement and consistency amongst records used with or created by the voting system. There are other types of audits and audit-related activities other than those specified in the VVSG that election systems should be designed to support. For example, auditing event logs—logs that record the times and descriptions of voting system events—allow detecting anomalous events such as machines being opened before polls were open, machines being reset or rebooted, or even unusual patterns of ballot casting. In the election audit community, the term "post-election audit" has come to refer to the more narrow practice of conducting a manual tally of physical audit records and comparing the result to the electronic result stored by the EMS (the third type of audit in the list below). Even within post-election audits, the Carter Center has introduced the idea of "hot" and "cold" audits, where the former can impact the certified result and the latter are used as part of a continual quality monitoring program and do not affect the outcome of the certified result.[18]

That being said, the VVSG draft refers to three types of audits:

- The phrase "pollbook audit" (Part 1:4.2.1) refers to counting pollbook signatures and comparing that count to the vote data reported by the tabulator.
- The phrase "hand audits of IVVR records" (Part 1:4.2.2) refers to manually counting audit records and comparing to the vote totals reported by the tabulator.
- The phrase "ballot count and vote total audit" (Part 1:4.2.3) refers to manually counting audit records and comparing to the vote totals reported by the EMS. We will call this a "manual tally" audit.

For election officials, the pollbook audit is typically only one part of a larger process, often called "ballot reconciliation", that starts immediately after election day and involves the mentioned pollbook audit but also includes activities such as balancing the voted, spoiled and unused ballot stock with the number of ballots sent to each precinct. To our knowledge, few if any jurisdictions employ the second notion of auditing above, comparing a hand audit of audit records to totals produced by a tabulator, regardless of what the EMS reports.[19]

⟨URL: `http://www.shapethefuture.org/press/2006/insidebayareacom113006.asp`⟩

[17] Norden et al. (as in n. 12).

[18] *Summary of Proceedings, Automated Voting and Election Observation*. The Carter Center, March 2005 ⟨URL: `http://www.ciaonet.org/wps/car071/car071.pdf`⟩.

[19] Note: The last two types of audits may seem equivalent at first blush; however, the difference is that the manual count in each case is compared to two different sets of electronic records: those from the precinct tabulator device and, in the other case, from the central Election Management System software.

2.1.1 The VVSG Draft's Auditing Requirements Will Significantly Enhance Voting System Auditability

The VVSG's chapter on Audit Architecture requirements (Part 1:4) will greatly enhance the auditability of voting systems certified to the guidelines. These requirements cover much of the ground towards achieving auditability of IVVR voting systems; they include requirements by type of audit being performed (pollbook audits, tabulator audits and manual tallies), requirements for electronic audit records, and requirements for physical audit records. With one exception, discussed in the next section, the VVSG draft addresses each area of auditing from a systems perspective.

The VVSG draft is also appropriately forward-thinking with respect to support for auditability. For example, none of the major manufacturers currently support digital signatures for audit data.[20] This is problematic, as auditors need to be able to compare results of a manual audit to digitally-signed electronic results. Without verification support using tools such as digital signatures, parties with an interest in corrupting the audit or hiding evidence of error could fabricate the audit records or render them unusable through denial-of-service attacks. The VVSG draft, however, requires digital signatures be used with electronic audit data so that the content can be verified as produced by a specific device at a specific time.

The draft further addresses problematic features of currently deployed voting technologies. For example, Part 1:4.4.2.2-B requires that voting systems with VVPAT capability be able to detect problems that might affect the printing, recording, or storage of the VVPAT record and, upon such a detection, prohibit the voter's ballot from being cast. Currently, only one manufacturer's VVPAT subsystem (Hart InterCivic's eSlate DRE with VBO VVPAT) has this capability. Missing, destroyed or unreadable VVPAT records have become increasingly prevalent, affecting the quality and, in some cases, the *possibility* of conducting post-election manual tallies of VVPAT records.

Finally, the VVSG draft supports some future directions of voting system auditing models that are now only nascent. For example, the requirements in Part 1:4.4.3.1-A–A.1 allow precinct-count optical scan (PCOS) systems to make optional marks on ballots during the casting and scanning process while restricting these optional marks to specific areas of the ballot face for security reasons. Researchers are now working on methods to increase the effectiveness and efficiency of manual tally audits using machine-assisted auditing that would require optional marks to be written on a ballot at the time of casting.[21]

2.1.2 Further Enhancements of the VVSG Draft Are Needed to Better Support Auditing

Currently deployed systems exhibit a number of shortcomings with respect to supporting audit activities. For example, manual tally procedures often specify that the vote totals—the quantities being audited—must be made available to the public before the random selection and manual tally.[22] However, some manufacturers' EMSs do not report totals in a way that would be useful for an auditor or public observer. For example, vote totals for ballots cast on PCOS systems in the precinct are often automatically mixed with totals for DRE+VVPAT votes cast in the same precinct. Mixing two or more sets of vote totals for devices that require different auditing methods frustrates auditing and observing efforts; hand counting PCOS ballots is a different process from hand counting VVPAT records.

[20] Even in the places that manufacturers do use digital signatures, they often misuse them. California Top-To-Bottom Review (as in n. 8); McDaniel et al. (as in n. 8)

[21] Joseph A. Calandrino, J. Alex Halderman and Edward W. Felten, Machine-Assisted Election Auditing. USENIX/ACCURATE Electronic Voting Technology Workshop 2007, August 2007 ⟨URL: `http://www.usenix.org/events/evt07/tech/full_papers/calandrino/calandrino.pdf`⟩.

[22] This ensures that the public can verify that the tally agrees with results to which the election official has previously committed.

In some cases, the manufacturers' EMSs will not report machine-specific results within a precinct. Unfortunately, this often means that a manual tally of, say, four to five VVPAT rolls for a given precinct can be compared only with aggregate precinct totals, instead of on a machine-by-machine basis. Considering that it might take one tally team of four people over 4 hours to tally the VVPAT rolls for one precinct, finding a discrepancy after all that effort is ineffective; if there is a discrepancy, the EMS report contains no information that would be helpful in locating on which VVPAT roll the discrepancy might be contained.[23] This can result, due to blind counting rules,[24] in the tally team having to redo the tally for that precinct's VVPAT rolls. If the EMS had reported vote totals for each machine in the precinct, the tally team would have had to retally a small number of VVPAT rolls.

State-of-the-art auditing methodologies can also place distinct requirements on voting systems. For example, statistically conservative audit schemes[25] start with a flat percentage audit, then require the auditor to calculate a statistical confidence value and, if needed, increase the sample size of the audit. However, some manufacturers' EMSs will produce meaningful results only in PDF format, a format useful for presentation of information but not useful for computation. To calculate a statistical quantity with data from hundreds of precincts in such an unusable format would require an army of transcribers. If EMSs had the capability to output vote totals in an open, machine-readable and machine-processable format, they would better support more sophisticated forms of election audits.

It is clear that adequate support for auditing manual tallies requires two important features:

1. The vote data stored by the EMS should be kept at the level of ballot and device granularity appropriate for the manual tally; and,

2. The EMS must be able to output this information in a form useful for all parties involved in the manual tally procedure.

These guidelines illustrate a few important points about the EMS's storage of vote data. First, different ballot types need be kept separate in the EMS database according to the type of casting methods as well as ballot status (e.g., provisional, regular, vote-by-mail, and early voting). Data is meaningful for audit purposes only if the EMS can output reports that include this level of detail.

Second, this data should be kept at a level of granularity that corresponds to the audit unit. For lower-capacity voting devices, the device level is probably the best level of granularity here as opposed to the level of individual VVPAT-rolls, which might be difficult for the machine to keep track of. For high-capacity devices such as central-count optical scanners, storing data on the batch level makes more sense. Some of these requirements might be covered by Part 1:4.2.2-A.1 of the VVSG, but only at a high-level; it would seem wise to attempt to specify these elements in more detail.

A related recommendation is that the system should support locating types of ballots to support the auditing context. For example, if a jurisdiction is performing a precinct-level audit, it will need to locate all the ballots for that precinct. For vote-by-mail (VBM) ballots, which are often scanned centrally in batches rather than sorted into precincts, it makes sense for the EMS to provide reports that list in which batch a precinct's VBM ballots are located and how many are in each batch.[26]

[23] To take this example to an extreme, even if a precinct uses 40 DREs with VVPAT printers, all the votes might be combined into one quantity by the EMS. The obvious problem with this design is that after counting 40 machines-worth of VVPAT rolls by hand, if there is a discrepancy, the system gives the auditor no information about which machine(s) might hold the discrepancies.

[24] A *blind count* is where the tally team manually counts the ballots without knowing the result they should achieve. Blind counting ensures that no conscious or unconscious incentives exist for artificially making the tally and electronic count match.

[25] Philip B. Stark, Conservative Statistical Post-Election Audits (in press). The Annals of Applied Statistics, 2008 ⟨URL: `http://www.stat.berkeley.edu/~stark/Preprints/conservativeElectionAudits07.pdf`⟩.

[26] To support including valid provisional ballots cast on DRE+VVPAT machines, the EMS should be able to tell the auditor

The VVSG draft also recognizes the need for audit record output to be in a public format that supports auditing; however, there is under-specification as to what constitutes an "open" format. Specifically, Part 1:4.3.1-A requires that audit records be available in a "fully specified, public format". However, "fully specified" and "public" do not necessary correspond to "open". The requirement in Part 1:4.4.1-A.8 is more specific about the requirements for an open format being "non-proprietary" and "requiring no special knowledge of confidential or proprietary or trade secret information". These elements of this definition should, at a minimum, be copied to the aforementioned section. A more general solution would specify in one part of the VVSG what constitutes an "open" format and then simply incorporate that definition by reference.

In addition to openness, the VVSG should specify that EMSs provide machine-readable and machine-processable output to support auditing. Auditors and members of the public will need access to vote data and audit data to conduct or oversee complex auditing methods or to verify the vote count in elections using alternative voting algorithms, such as Instant-Runoff Voting (IRV). The VVSG draft approaches this subject only narrowly; the discussion for 4.3.2-A specifies that the "[tabulator] record must be output in a human-readable format," but it says nothing about machine-readability or -processability. In the most basic sense of machine-processability, EMSs should output vote and audit data in a spreadsheet format such as Comma-Separated Value (CSV) or Open Document Spreadsheet (ODS) format. However, it is also important for the VVSG to require manufacturer support for output in a standardized data-rich XML format. Such support would require an effort to define minimum reporting requirements for relevant data and define a single standard format to be used, eventually, in all jurisdictions and by all manufacturers. The only current candidate for such an XML format is the OASIS standard Election Markup Language (EML).[27]

3 The New Security and Reliability Testing Will Greatly Improve Voting System Security and Reliability

Past versions of national voting system standards were severely lacking in terms of requirements and evaluation of system security and reliability. The current draft goes a long way towards correcting these deficiencies. We believe that each of the sets of requirements for security and reliability should be retained in future drafts.

Security and reliability are closely linked concepts. Secure systems are engineered to function dependably in an environment that includes malicious activity, error and plain bad luck. A system is reliable to the extent that it is resistant to error and malfunction. Naturally, these two qualities interact and complement each other: less reliable systems are often less secure; and less secure systems are not likely to remain reliable in the presence of misuse or other security anomalies.

Two promising methodologies for discovering and correcting security flaws and reliability issues are, respectively, adversarial vulnerability testing and volume testing of voting systems. *Adversarial vulnerability testing* of currently deployed voting systems has proven to be effective at finding flaws in voting systems that should have been addressed during certification testing. *Volume testing* provides a more realistic method of measuring a voting system's reliability than methods in previous standards.[28]

on which machine a valid provisional vote was cast for a given audited precinct. This avoids having to, for example, unroll a whole precinct's-worth of VVPAT rolls to find a single provisional ballot that was cast in the wrong precinct.

[27] John McCarthy, *Strengthening 2007 Voluntary Voting Systems Guidelines for Inter-operability, Data Publication and Election Auditing: The Case for Requiring EML 5.0 (or higher).* Verified Voting Foundation, May 2008.

[28] The mean-time-between-failure (MTBF) metric of 163 hours used in the past has been problematic, virtually ensuring that reliability problems would surface during the 12+ hours of election day. Note that an MTBF of 163 hours results in a probability of machine failure of $1/163$ in a given hour. For a 13-hour election day, this corresponds to a failure rate of 8.0% ($13/163$). See: Howard Stanislevic, *DRE Reliability: Failure by Design?* Vote Trust USA E-Voter Education Project, March

Coupled with the requirement that voting systems be software independent, these added security and reliability tests have the potential to be the biggest contributions to voting system integrity in the next VVSG.

In this section, we discuss the track record of these methods applied to voting systems and why we feel they are an essential part of the next generation of voting system guidelines.

3.1 Adversarial Vulnerability Testing Will Increase the Security of Certified Voting Systems

State-level deployments of voting systems have substantially benefited from the use of vulnerability testing. Studies mentioned previously[29] show that many deficiencies have fallen through the cracks of national testing. To find these flaws, it took adversarial vulnerability testing including source code review and penetration testing.[30]

Adversarial vulnerability testing involves source code review and penetration testing—methodologies from the domain of computer security analysis. In source code review, reviewers analyze a system's source code—statically or dynamically—using a variety of tools—from automated flaw-finding tools to step-by-step debugging tools to human inspection of the code. The flaws that reviewers find are then evaluated in terms of impact, in terms of possible interactions with other features of the system and in terms of the system's response to exploitation.

Penetration testing involves developing attacks that could conceivably be used in the voting system's development and operational environment to modify, add, subtract or destroy voting data. Penetration testers typically use a model of threats against the system along with knowledge of the system to hypothesize possible attacks and attempt to demonstrate them on the voting system configured for operation. When combined together, using both source code review and penetration testing can expose complex sets of vulnerabilities that can then be prioritized in terms of difficulty, required resources and possible impact.

Some draft VVSG commentors refer to vulnerability testing as an impossible test that has no requirements. However, the draft VVSG clearly identify the requirements for adversarial vulnerability testing as well as the pass/fail criteria. Part 3:5.4 lists requirements for vulnerability testing including team composition, the scope of testing, resources made available, level of effort to be expended and the rules of engagement for evaluation of the system. The team make-up and qualifications requirements are designed such that the testers possess a high level of expertise. Part 3:5.4.4 specifies the fail criteria for vulnerability tests: a system can fail if (1) the manufacturers system in conjunction with use procedures and security controls do not adequately mitigate significant threats (Part 3:5.4.4-B); or (2) if found vulnerabilities could be used to: "change the outcome of an election, interfere with voters' ability to cast ballots or have their votes counted during an election, or compromise the secrecy of vote [...]" (Part 3:5.4.4-C)

Related to this point, some draft VVSG commentors have rallied around what amounts to a semantic

2006 ⟨URL: `http://www.wheresthepaper.org/StanislevicDRE_ReliabilityMTBF.pdf`⟩

[29] See note 8.

[30] It is interesting to note that many of the vulnerabilities found in the existing set of DREs could have been found by a method called *fuzz testing*. Fuzz tests involve identifying every point where data is input into a computer system and providing random inputs to that point. Fore example, in the case of file systems, this requires identifying all file names the system might try to open and providing it with files by those names holding random content. Where the system checks file authenticity by, for example, checksum or cryptographic means, the fuzz tester will equip their fuzz data with correct authenticators to force the system to digest garbage. A system is considered to fail a particular fuzz test if it hangs or throws an unhandled exception. See: Barton P. Miller et al., *Fuzz Revisited: A Re-examination of the Reliability of UNIX Utilities and Services*. University of Wisconsin-Madison, Computer Sciences Dept, 1995 ⟨URL: `http://reactos.ccp14.ac.uk/advocacy/fuzz-revisited.pdf`⟩

issue in opposition to this requirement. The draft VVSG uses the term "Open-Ended Vulnerability Testing" (OEVT) to describe adversarial vulnerability testing. Some argue—for example, in the EAC Standards Board Resolution 2007-07—that *open-ended* testing is an invitation to fail systems because it allows a form of testing "without restriction" and in a manner "not conducive to a conformance assessment".[31] As outlined above, the draft VVSG contains specific requirements, restrictions and fail criteria relating to adversarial vulnerability testing. Statements such as the resolution above demonstrate a lack of understanding of adversarial vulnerability testing as well as what the VVSG says.

3.2 Volume Testing Will Increase the Reliability of Certified Voting Systems

The draft VVSG proposes volume testing as a testing method that will "contribute substantially to the evaluations of reliability, accuracy, and misfeed rate" (Part 3:5.2.3-D). Volume testing is a testing procedure that requires casting a large amount of votes simultaneously on many voting devices. Testers keep an error log in order to record all anomalies experienced during the volume test. (The draft VVSG protocol for volume testing is essentially an enhanced version of the volume testing protocol used by the State of California since 2006.[32])

Volume testing is a vital element of future certification with respect to voting system reliability. It simulates the load that a typical machine might encounter during its peak use period and does so on many devices at once. This allows testers to observe the ways voting systems fail and measure the consistency of certain types of failures across an inventory of exemplar machines. Volume testing has become an essential part of certification in California, where it has exposed subtle, hard-to-diagnose flaws with voting systems that would not have been detected otherwise.[33]

However, state-level volume testing can never be as instrumentally effective as volume testing performed during national certification. Flaws found during national certification can be fixed immediately and the system re-certified during the ongoing certification process instead of having to re-submit a delta change under a new certification attempt. Of course, as we have stated previously, conducting this kind of intense testing and doing it right is much more efficient at the national level; otherwise, states and jurisdictions have to do it on their own and the cost multiplies accordingly. Because it permits an effective and significant measurement of reliability and because of its high expense, volume testing should be conducted at the national level instead of piecemeal at the state or jurisdictional level.

4 The Expanded Usability and Accessibility Requirements and Testing Will Profoundly Improve the User Experience for All Voters

A welcome feature to the current VVSG draft is the significantly expanded set of usability and accessibility tests and requirements. As ACCURATE discussed in its comments to the 2005 VVSG (as well as in our own usability research), the previous federal qualification regimes lacked comprehensive usability and accessibility standards and evaluation.[34] We argued then that this type of standard

[31] See Resolution 2007-07 of: *EAC Standards Board Resolutions: December 2007*. U.S. Election Assistance Commission, December 2007 ⟨URL: `http://www.eac.gov/about/docs/sb-resolutions-and-cover-letter-dec-2007.doc/attachment_download/file`⟩. Contrast this with the more positive language of the EAC Board of Advisors Resolution 2007-D7: *EAC Board of Advisors Resolutions: December 2007*. U.S. Election Assistance Commission, December 2007 ⟨URL: `http://www.eac.gov/about/committees/advisors/docs/2007-resolutions.pdf/attachment_download/file`⟩.

[32] *Protocol for Volume Testing of Voting Systems*. California Secretary of State, 2006 ⟨URL: `http://www.sos.ca.gov/elections/voting_systems/volume_test_protocol_final.pdf`⟩.

[33] Jefferson et al. (as in n. 7).

[34] ACCURATE's Comments on the 2005 VVSG (as in n. 1); Michael D. Byrne, Kristen K. Greene and Sarah P. Everett, Usability of Voting Systems: Baseline data for paper, punch cards, and lever machines. Association for Computing Machinery,

should include expert review of systems, usability tests with actual voters and usability testing of voting systems during the certification process. In this respect, we are encouraged by the requirements for manufacturer-conducted usability testing and the development of specific performance benchmarks for VSTL-conducted usability tests. The new requirements and required testing will undoubtedly improve the user experience for all voters, with or without disabilities.

The VVSG requires manufacturers to conduct summative usability tests. Summative usability testing is testing performed on a finished product to demonstrate that the developed system is effective, efficient, and satisfactory to its intended users. The requirements that the manufacturer perform summative usability tests appear in six crucial places: usability tests of the system for the general population (Part 1:3.2.1.2-A), for alternative language requirements (Part 1:3.2.7-A.4), for poll worker usability (Part 1:3.2.8.1-B), for voters with low vision (Part 1:3.3.2-A), for blind voters (Part 1:3.3.3-A) and for voters with dexterity difficulties (Part 1:3.3.4-A). These cases are vital for demonstrating that the system is usable by a wide variety of users, from the general population of voters to voters with specific disabilities.

The VVSG separates these usability *design requirements*, binding on the manufacturers, from usability *performance requirements* for which the Voting System Testing Laboratories (VSTLs) will perform testing and report results.[35] This distinction between manufacturer-conducted and VSTL-conducted usability testing needs to be made more clear. That is, nowhere in the performance testing section (Part 1:3.2.1.1) does it mention that these benchmarks will be evaluated by a VSTL. The sole mention of who will conduct testing in this section of the VVSG is one mention of the manufacturer in Part 1:3.2.1.2. In addition, there is a mention of summative testing (Part 1:3.2.1.2) which appears *after* the discussion of performance testing benchmarks (Part 1:3.2.1.1), further obscuring the temporal order of these events. It should be perfectly clear that the manufacturer submits evidence of performed summative usability testing to the VSTL and that the VSTL then (1) evaluates this evidence and (2) conducts usability performance testing.

The five usability benchmarks used in performance testing in Part 1:3.2.1.1—Total Completion Score, Perfect Ballot Index, Voter Inclusion Index, Average Voting Session Time and Average Voter Confidence—are also welcome and encouraging. The first three of these benchmarks will provide an absolute target that voting systems must meet to satisfy the recognized goal of effectiveness.[36] While there are no specific targets for efficiency and satisfaction, requiring manufacturers to report data about those measures may also lead to improvements in those areas. NIST has since published the Voting Performance Protocol (VPP) which establishes a standard testing methodology for testing each of these benchmarks.[37] While perhaps not all of the details of the VPP represent our preferred choices, the fact that such usability benchmarks exist at all is a critical and welcome step. It is worth noting that the usability measures collected using the VPP are likely to overstate the true usability of voting systems because the VPP instructs voters on who to vote for in the test, which tends to decrease error rates.

2007 ⟨URL: http://doi.acm.org/10.1145/1240624.1240653⟩.

[35] "Voting System Testing Laboratory" is the technical term for a testing laboratory authorized by the EAC under NIST's National Voluntary Accreditation Laboratory Program (NVLAP). These are the equivalent of the Independent Testing Authorities (ITAs) in the previous NASED voting system qualification regime.

[36] Note that state-level rules about what constitutes a valid vote will continue to impact vote counting consistency as measured in the field. For example, in optical scan systems, the nexus between what the machine reads as a valid mark, what mark the voter makes in response to the ballot's marking instructions and state laws that govern the rules for valid marks will result in different vote totals in different jurisdictions. Unfortunately, this inconsistency, caused by HAVA's requirements that each state define a "valid vote", has moved the question of "what is a valid mark" outside the realm of what can be tested at the national level.

[37] *Usability Performance Benchmarks For the Voluntary Voting System Guidelines*. National Institute of Standards and Technology HAVA Technical Guidelines Development Committee, August 2007 ⟨URL: http://vote.nist.gov/meeting-08172007/Usability-Benchmarks-081707.pdf⟩.

Besides specifying benchmarks, the draft VVSG also requires certain specific user interface features for DRE-style systems. Many currently-deployed DREs do not meet the feature requirements as laid out in the draft VVSG, and it is our belief that few, if any, are likely to meet the benchmark targets. Thus, adoption of these standards will have a substantial impact on future voting systems and their usability.

While there does seem to be some standardization effort being expended on the first three of the five benchmarks (performance requirements), which cover effectiveness, we would like to see some more attention paid to the last two reporting requirements that cover the time taken to vote and voter confidence in the system. For example, the System Usability Scale (SUS) was developed to facilitate comparison of subjective usability assessments across similar, but slightly different, systems and its use is supported by recent psychometric research.[38] The SUS score (or a subset of the SUS questionnaire) would likely make more sense as a reporting requirement for satisfaction than the undefined "Average Voter Confidence" measurement in Part 1:3.2.1.1-D.3.[39]

We suspect that the benchmark targets are high enough that this will encourage manufacturers to move towards a a User-Centered Design (UCD) model[40] in the design and manufacturing of their voting systems, including both formative and summative usability testing, which would require in-house usability testing and user experience expertise. This implicit recognition of the benefits of UCD is welcome; however, the use of UCD should be explicitly encouraged in the VVSG. Manufacturers that eschew formative usability testing, prescribed by the UCD model, during research and development will find their internal testing might not pass muster with the required VSTL evaluation of manufacturer-performed summative usability tests.

Furthermore, the relationship between the usability and accessibility standards laid out in Section 3 of the VVSG and the standards for voter verification laid out in Section 4 are unclear. Will VVPAT systems be required to meet the same usability benchmarks? We believe, in principle, that they should, but there is so little research on the usability of voter verification systems that it is hard to know how difficult it would be to meet such standards.

Finally, we are further encouraged by the explicit mention of usability by poll workers in the VVSG. We hope that future versions of the VVSG will set the same kinds of usability benchmarks for poll worker activities such as set-up and configuration of voting systems.

5 The Draft Recognizes the Importance of Adequate Documentation

The current draft of the VVSG makes three important advances in the treatment of voting system documentation. First, the draft adopts a significant change in perspective by viewing documentation as a distinct part of the voting system, rather than an element of voting system's functionality or performance (Part 2:1.1.1). An important consequence of this change is that voting system documentation becomes part of a structure that supports the high-level goals of trustworthy elections through software independence. Second, a set of national guidelines pertaining to structure and content will likely make voting systems more comprehensive and usable, thus helping to relieve states of some of the need to require additional documentation from manufacturers or develop their own documentation. Third, the draft

[38] John Brooke; Patrick W. Jordan et al., editors, SUS: A 'Quick and Dirty' Usability Scale. London: Taylor and Francis, 1996, Usability Evaluation in Industry; Aaron Bangor, Philip T. Kortum and James T. Miller, An Empirical Evaluation of the System Usability Scale (SUS) (in press). International Journal of Human-Computer Interaction 2008.

[39] Note that the VPP goes into more detail in defining a Likert scale-based question to measure voter confidence. (See: NIST's Voting Performance Protocol (as in n. 37) at 21-22) However, we would recommend a known metric, like SUS, which is comparable across disparate types of technology, and to which usability research on voting systems can be compared.

[40] Sharon Laskowski et al., *Improving the Usability and Accessibility of Voting Systems and Products*. April 2004 ⟨URL: `http://www.vote.nist.gov/Final%20Human%20Factors%20Report%20%205-04.pdf`⟩.

adopts a more balanced view of confidentiality and intellectual property rights than previous guidelines. Still, some requirements need clarification.

5.1 The Documentation Requirements Generally Lend Support to the Draft's Stated Objectives for Voting System Properties

The shift to defining documentation as part of the voting system also brings the promise of promoting documentation that supports the VVSG's high-level goals: "fair, accurate, transparent, secure, accessible, timely, and verifiable elections" (Part 1:2.7.2). To draw attention to a few particularly important areas, we first note that requirements 2:4.4.5-B and C will help create more effective documentation for detecting and recovering from VVPAT errors while protecting voter privacy.[41] Given the importance of VVPATs to currently available, software-independent DRE systems, detailed guidance from manufacturers about how to detect and correct faults in VVPAT printers is critical.

Similarly, Part 2:4.3 outlines requirements for documenting elements of election administration that are critical to maintaining the security of voting systems. These include documenting the system's access controls (Part 2:4.3.1), system event logs (Part 2:4.3.2), physical security (Part 2:4.3.4), and producing records that are necessary to audit the system (Part 2:4.3.6). The requirements for high-level descriptions of a voting system's security (requirement 2:3.5.1-B and Table 3-1), for example, will allow test labs to understand more easily the threats that manufacturers consider during development, as well as which design choices were made to address these threats. These descriptions, in turn, will likely provide useful guidance to test labs as they develop plans to test conformance to the VVSG's security requirements. To the extent that states or other election jurisdictions obtain Technical Data Packages (TDPs), these security-related documents will be useful to the development of state-level security plans and any future voting system evaluations.

Still, the document requirements could go further toward supporting software independence. Specifically, the VVSG should require manufacturers to provide an audit plan as part of their documentation.[42] The current requirements are likely to lead to detailed documentation of specific steps in an audit. A useful complement to this detail would be a high-level plan that outlines the steps involved in an audit and links them to physical security, chain of custody, and other issues that must be considered to make an audit secure.[43] States, of course, have different audit laws (if they have them at all); and it would be difficult to capture all of these variations in a single audit plan. Still, the basic elements remain the same; and we simply recommend placing them in the broader context of setting up, running, and closing an election. A high-level, schematic audit plan would be the best way to provide this context.

---

[41] These requirements read, respectively:

> Manufacturers of VVPATs SHALL provide documentation for procedures to recover from VVPAT printer errors and faults including procedures for how to cancel a vote suspended during an error.

and

> Manufacturers of paper-roll VVPATs SHALL provide documentation describing necessary procedures for handling the paper roll in a way that preserves voter privacy.

[42] The audit documentation requirements in Part 2:4.3.6 should also contain references to the descriptions of a pollbook audit (Part 1:4.2.1), hand audit (Part 1:4.2.2), ballot count and vote total audit (Part 1:4.2.3), and observational testing (Part 1:4.2.4).

[43] Evidence of how the detail currently required may be seen in the discussions of requirements 2:4.3.6-A–E. For example, the discussion for requirement 2:4.3.6-A states that conforming documentation for a pollbook audit "includes explaining how to generate all needed reports, how to check the reports against one another for agreement, and how to deal with errors and other unusual problems that come up during the audit step."

This information is undoubtedly critical to conducting a secure pollbook audit. It would be helpful, however, to situate these details within the larger context of pre- and post-election events, as discussed in the main text.

5.2 The VVSG's Documentation Requirements Would Aid State and Local Election Administration

The need for better national-level guidance for voting system documentation has been made clear by recent voting system reviews in California and Ohio. Document review reports from the California "Top-to-Bottom Review," for example, found that user documentation—intended for use by election officials and pollworkers—lacked detail and clarity sufficient to allow jurisdictions to run elections without either external information or assistance.[44] These reports also found that the lack of national-level guidance for voting system documentation leaves gaps that states seek to fill by requiring additional documentation from manufacturers. At best, this puts manufacturers to needless expense; it would be more efficient to create documentation that meets jurisdictions' needs at the outset. At worst, writing state-by-state, ad hoc supplements creates the potential that a state (or other election jurisdiction) will receive internally contradictory documentation. The confusion that this sows makes it more likely that voting equipment will be set up or used improperly.

5.3 The Confidentiality and Intellectual Property Requirements Need Clarification

Finally, the draft's treatment of voting system documentation confidentiality is sensible but in need of some refinement. As the draft points out, manufacturers should be allowed to mark certain materials as confidential (Part 2:3.1.3-A), with the caveats that "[a]n accredited test lab may reject a TDP if it is so encumbered by intellectual property claims as to obstruct the lab's delivery of the Test Plan. . . or Test Report. . . " and that "[a]n overuse of trade secret and patent protection may prevent certification by a certification authority." These limitations mark a change, relative to VSS II.2.1.3, to bring the TDP requirements into line with the EAC's *Voting System Testing and Certification Manual*.

We recommend, however, that the mention of patent protection in the final paragraph of requirement 2:3.1.3-A be removed. A patent gives the patentholder a right to restrict others from practicing the invention described in the patent. A patent does not restrict disclosures of information about a system containing a patented invention. To the contrary, the grant of a patent is conditional upon publication of how to practice the invention covered by the patent. Thus, a patent will not protect information submitted by the manufacturer from public disclosure.[45] A more appropriate focus for requirement 2:3.1.3-A is any material designated as confidential or as a trade secret. Requirements 2:3.1.3-A and B would be clearer if they eliminated references to "intellectual property" and "proprietary information" and replaced them with "confidential or trade secret information."

6 Some Features of the Draft Necessitate Increased Institutional Support from the EAC

There are a few things in the new VVSG that will need increased institutional support.

[44] Candice Hoke and Dave Kettyle, *Documentation Assessment of the Diebold Voting Systems*. July 2007 ⟨URL: `http://www.sos.ca.gov/elections/voting_systems/ttbr/diebold_doc_final.pdf`⟩; Joseph Lorenzo Hall and Laura Quilter, *Documentation Review of The Hart Intercivic System 6.2.1 Voting System*. July 2007 ⟨URL: `http://www.sos.ca.gov/elections/voting_systems/ttbr/hart_doc_final.pdf`⟩; Aaron J. Burstein, Nathan S. Good and Deirdre K. Mulligan, *Review of the Documentation of the Sequoia Voting System*. July 2007 ⟨URL: `http://www.sos.ca.gov/elections/voting_systems/ttbr/sequoia_doc_final.pdf`⟩.

[45] In any event, determining the scope and validity of a patent claim may be extremely time-consuming and expensive. Jurisdictions may wish to consider patent issues when purchasing voting systems, but considering them during testing would add little value to that process.

6.1 The Innovation Class Will Require a Support Structure Outside the Scope of the VVSG

As we stated in Section 2, the VVSG draft mitigates the risk of causing voting technology to ossify by specifying all permissible classes of voting systems while maintaining a requirement of software independence. The mechanism that the VVSG provides for covering new technologies is the "innovation class" (Part 1:2.7.2). The requirements for the innovation class are basically sound, but making the innovation class work in practice will require additional and ongoing institutional support from the EAC. We outline requirements for such support and present some possibilities for encoding it into the VVSG.

We wish to make clear that software independence is crucial to the innovation class. The innovation class simply is not acceptable unless devices within this class are required to be software independent.

Given this constraint, the innovation class meets a significant need created by the overall formal structure of the VVSG. This structure, set forth in Part 1:2.5 of the VVSG, creates a tractable way of identifying where a given piece of voting equipment falls into the voting process. This makes clear which requirements apply to a given voting system or device, which, in turn, will facilitate the preparation of clear and thorough test plans. Without the innovation class, however, the VVSG might exclude technologies that do not fit within the classes specified in the guidelines.

In the broad contours presented in Part 1:2.7.2, the innovation class represents a sensible solution to this problem. The innovation class balances the stability and coherence of the VVSG's class structure with the ability to accommodate technologies that are currently in early stages of development or even entirely unforeseen.

The current innovation class requirements, however, would benefit from some amendments; and making the innovation class work in practice will require a clear plan for ongoing institutional support from the EAC. We first address changes to the currently proposed requirements. It is unclear why an innovation class submission must separately justify a device's innovativeness (requirement 1:2.7.2-B). It would seem to be a sufficient demonstration of innovativeness that a device performs or supports one or more recognized voting activities, and that it does so in a manner that is not contemplated by other specific classes. As the draft's discussion of requirement 1:2.7.2-B states, the threshold consideration for an innovation class submission is "whether the creation of a new class is justified" based on a description of the device's "functionality, boundaries, and interactions with other devices." This description actually provides a more concrete guide to the identification of innovativeness than does the text proposed for requirement 1:2.7.2-B. Accordingly, we recommend that the requirement incorporate the concrete language of the discussion.

A further suggestion for improving the VVSG draft's innovation class structure is to make the class a focal point (both during the EAC's decision-making about the final VVSG, as well as in the use of the final guidelines to design, test, and certify voting systems) for further consideration of interpreted code. The current VVSG draft requires any interpreted code to "run under a specific, identified version of a COTS runtime interpreter" (Part 1:6.4.1.7-A.4). At least one recently developed voting system prototype makes extensive use of interpreted code that runs under a custom interpreter.[46] This system offers the prospect of dramatically simplifying the code necessary to support voting; but, under plausible readings of the VVSG draft's interpreted code restrictions, this system would not conform with the VVSG. Though we recognize the difficulties that interpreted code poses with respect assuring the

[46] See: Yee (as in n. 5). Also note that Premier runs a custom interpreter that has been the focus of some scrutiny in the past. David Wagner et al., *Security Analysis of the Diebold AccuBasic Interpreter*. Voting Systems Technology Assessment Advisory Board, February 2006 ⟨URL: `http://www.sos.ca.gov/elections/voting_systems/security_analysis_of_the_diebold_accubasic_interpreter.pdf`⟩; *Voting System Memory Card Issues*. National Association of State Election Directors, March 2006 ⟨URL: `http://www.nased.org/ITA%20Information/NASED%20Memory%20Card%20Report.pdf`⟩

integrity of voting system software, such promising systems warrant a closer look at the interpreted code restriction. Perhaps the innovation class will provide the practical framework for evaluating voting systems containing interpreted code. Before the EAC decides that issue, however, we recommend that the Commission reconsider the basic approach to interpreted code, perhaps with input from NIST or an ancillary process for establishing related guidelines.[47]

The decision of whether a device warrants the creation of a new device class raises a deeper question that the draft VVSG does not address at all: *Who* will make this decision? The EAC, of course, has the authority under the Help America Vote Act of 2002 (HAVA) to adopt and modify the VVSG through a notice-and-comment process which includes at least one public hearing.[48] Aside from providing this authority, however, HAVA does not articulate a structure for evaluating VVSG modifications in general or the adoption of new device classes in particular.

This void is troublesome, as it leaves the implementation of the innovation class—and a source of the VVSG's continuing vitality—open to doubt. Though we recognize that a full specification of an institutional support structure for the innovation class may be beyond the scope of the VVSG, we view this support as critical to managing the development of the innovation class; and we urge the EAC to consider this issue in tandem with the VVSG as a whole. We do not propose in these comments a full structure for institutional support of the innovation class. Instead, we offer two broad considerations to guide the development of such a structure.

First, the EAC should seek to maintain the TGDC/NIST relationship or a similar body as a source of technical, elections and scientific expertise on voting systems in order to evaluate innovation class submissions. Under the requirements proposed in Part 1:2.7.2, evaluating an innovation class submission will require an ability not only to understand the details of the submitted device but also to compare this device to the relevant technologies that are currently in use. Making these assessments will call for a breadth and depth of technical knowledge that the EAC may not have the resources to maintain on its own and which may be difficult to reassemble if the TGDC is disbanded.

Second, given the likely complexity of considerations that will go into innovation class submissions, the EAC should consider either creating a set of procedural guidelines to accompany the substantive requirements in the VVSG, or modifying the VVSG to include such procedures.[49] Innovation class submissions will likely raise issues similar to those addressed in the *Voting System Testing and Certification Manual*. For example, there is a need to balance claims of confidentiality in submissions with the need for the release of information—and, in the case of innovation class submissions, public comment—about submitted devices. It will also be important to publish the basis for the EAC's decision to accept or reject an innovation class submission, in order to give guidance to election officials and voters. As the experience with the *Voting System Testing and Certification Program* demonstrated, a comment period allowed considerable illumination of the complex issues present in submitting systems for evaluation,

[47] Some relevant considerations for more nuanced interpreted code requirement include:

- Does the introduction of the interpreter cause a significant net (i.e., taking into account extra code from adding the interpreter itself) reduction in the size and complexity of the software overall?
- Is the source code of the custom interpreter included in the code base submitted for review and testing?
- Does the interpreter itself run on a specific, identified version of a COTS compiler or interpreter?
- Does the interpreter have limited access for only a limited purpose (as compared to the voting system as a whole), and is there documentation provided to make a solid case that it cannot exceed that limited access?

[48] *Help America Vote Act of 2002 (HAVA) § 222.*

[49] HAVA provides the EAC with the general authority to modify the VVSG and to seek outside technical expertise when doing so. See, for example, HAVA § 221(e)(1), which authorizes the TGDC to seek technical support from NIST "to carry out its duties under this subtitle," which include "assist[ing] the Executive Director of the Commission in the development of the voluntary voting system guidelines" (HAVA § 221(b)(1).

and the comments that were submitted resulted in a stronger standard. A similarly complex set of issues will likely go along with innovation class submissions, and we urge the Commission to devote a similar level of attention to this context.

6.2 The VVSG Should Incorporate Incident Reporting and Feedback into the Certification Process

As part of the Voting System Testing and Certification Program, the EAC has committed to developing a program to monitor field incidents with election technologies. The EAC has adopted a set of policies for "Field Anomaly Reporting" in § 8.7 of the *Voting System Testing and Certification Program Manual*.[50] However, there have been no field anomalies reported publicly by the EAC despite a number of recent high-profile cases of election incidents.[51] As we, and others, have argued in the past, robust incident reporting and feedback into the certification process would ensure that known problems would not continue to affect fielded voting systems.[52]

Non-governmental, non-partisan organizations have been collecting incident reports since 2004. VotersUnite has been tracking incidents reported in the press since 2004.[53] Organizations affiliated with the Election Protection Coalition (EPC) have written software and manned call centers to track election incidents for each election since August 2004.[54] Other groups and researchers have also collected incident data at the national and local level.[55] Instead of serving as a substitute, these groups' efforts would be better focused to enhance official incident reporting by the EAC.[56]

Recently, the EAC's own Board of Advisers passed a resolution calling for improved incident reporting. Their resolution stated that "many incidents and irregularities [...] have not been collected and made usable by election officials, vendors and the public", that the current system is "highly restrictive in terms of how input is provided and what types of incidents are reported" and calling for the EAC to

[50] *Voting System Testing and Certification Program Manual*. U.S. Election Assistance Commission, December 2006 ⟨URL: http://www.eac.gov/voting%20systems/docs/testingandcertmanual.pdf/attachment_download/file⟩.

[51] For example, see: California Nonpartisan Voters Report Trouble at Polls. cbs2.com (Los Angeles, CA), February 2008 ⟨URL: http://cbs2.com/politics/Ballot.Double.Bubble.2.646580.html⟩; Diane C. Walsh, Election-Machine Problems Spur Call for Study. New Jersey Star-Ledger, March 2008 ⟨URL: http://www.nj.com/news/ledger/jersey/index.ssf?/base/news-9/1205300247279401.xml&coll=1⟩

[52] ACCURATE's Comments on the 2005 VVSG (as in n. 1); Burstein, Hall and Mulligan (as in n. 4); Wendy Weiser, *Written Testimony of Wendy R. Weiser, Deputy Director, Democracy Program, Brennan Center for Justice at NYU School of Law before the Subcommittee on Financial Services and General Government of the House Appropriations Committee*. U.S. House of Representatives, February 2008 ⟨URL: http://www.brennancenter.org/content/resource/testimony_before_congress_regarding_the_eac/⟩.

[53] John Gideon and Ellen Theisen, *Election Problem Log: 2004 to Date*. ⟨URL: http://www.votersunite.org/electionproblems.asp⟩.

[54] In the 2004 and 2006 elections, the Election Protection Coalition ran the Election Incident Reporting System (EIRS). In the 2008 election cycle, the software has been rewritten and renamed to Total Election Awareness (TEA). See: *Election Incident Reporting Systems*. Election Incident Reporting System, 2006 ⟨URL: http://verifiedvotingfoundation.org/article.php?list=type&type=85⟩; *Total Election Awareness*. February 2008 ⟨URL: http://www.eff.org/deeplinks/2008/02/total-election-awareness⟩

[55] Another national effort collected audio-based incident data in 2006; see: Christopher Patusky, Allison Brummel and Timothy Schmidt, *MyVote1 National Election Report: Voice of the Electorate 2006*. Fels Institute of Government, Unniversity of Pennsylvania, August 2007 ⟨URL: http://www.fels.upenn.edu/Projects/myvote1_report_8_20_07.pdf⟩. Researchers Michael Alvarez, Thad Hall, D. Roderick Kiewiet and Jonathan N. Katz collected a rich set of incident reports from a single jurisdiction in 2006; see pages 48-70 of: *DRE Analysis for May 2006 Primary, Cuyahoga County, Ohio*. Election Sciences Institute, August 2006 ⟨URL: http://bocc.cuyahogacounty.us/GSC/pdf/esi_cuyahoga_final.pdf⟩.

[56] The EAC administers a survey research program to collect information from local jurisdictions each presidential election year. In 2004, the first instance of this data collection resulted in a report rich in aggregate data about election administration. However, this 2004 survey contained very little information on the frequency, extent and causes of machine failures. *2004 Election Day Survey*. U.S. Election Assistance Commission, September 2005 ⟨URL: http://www.eac.gov/clearinghouse/clearinghouse/2004-election-day-survey/⟩

create "an effective compilation of voting system incident reports that have been reported by local or State officials, keyed to different voting system vendors and models".[57]

The EAC should invest more effort into incident reporting. However, incident reporting should be one part of a feedback cycle where incidents reported from the field are included as data in the certification process. Systems should not pass federal certification if they previously displayed problems in the field that would otherwise violate requirements in the VVSG. The necessary feedback loop is not closed unless there are consequences for maufacturers that might submit voting systems to certification with known flaws. The EAC can complete this feedback mechanism by adding a requirement to the VVSG that instructs the VSTLs to submit voting systems to tests meant to replicate problems reported in the field.

7 Conclusion

In our comments, we have focused on a few crucial themes. First, software independence is an inextricable feature of the draft VVSG. Software independent voting systems will be more secure and will better support auditability. Second, the new standards for security and reliability testing—especially adversarial vulnerability testing and volume testing—will help ensure that voting system flaws and vulnerabilities do not slip through national certification as they have in the past. Third, the requirements for new usability and accessibility testing encourage a User-Centered Design model of research and development, which tends to improve usability, while requiring VSTLs to measure voting system usability performance against a set of meaningful benchmarks. Although it needs some clarification, a new Part of the draft VVSG properly recognizes the importance of voting system documentation and the crucial role that documentary information plays in voting system evaluation and administration. Finally, certain features of the draft VVSG—such as the innovation class and incident reporting—need further institutional support from the EAC to maximize the responsiveness of the certification regime.

The VVSG draft is a significant and positive step forward that we expect will contribute to and support improvements to the nation's voting systems. We laud the efforts of NIST and the TGDC in this ambitious overhaul of a complex standard. The ultimate measure of success for this certification regime can be measured by looking at the quality of our voting systems and the extent to which state and local jurisdictions find national certification useful and effective. Unfortunately, recent history has been unkind along both of these dimensions; one state is even considering undertaking its own certification process and passing legislation that would forgo any federal certification requirement.[58] The new VVSG shows promise of drastic improvements; when they go into effect, we will have higher-quality voting technology that better serves the voters and election officials of our nation.

[57] See Resolution 2007–[D14] of: EAC Board of Advisors 2007 December Resolutions (as in n. 31)

[58] Mark Niquette, Brunner wants to use voting devices that feds haven't yet OK'd. The Columbus Dispatch, May 2008 ⟨URL: http://www.dispatchpolitics.com/live/content/local_news/stories/2008/05/02/voting_machines.html?sid=101⟩.