Association for Computing Machinery
Advancing Computing as a Science & Profession

Statewide Databases of Registered Voters:

Study Of Accuracy, Privacy, Usability, Security, and Reliability Issues commissioned by the U.S. Public Policy Committee of the Association for Computing Machinery

February 2006

Study Committee Members:
Paula Hawthorn, Co-chair of Study
Barbara Simons, Co-chair of Study
Chris Clifton
David Wagner
Steven M. Bellovin
Rebecca N. Wright
Arnon Rosenthal
Ralph Spencer Poore
Lillie Coney
Robert Gellman
Harry Hochheiser

Preface

The Association for Computing Machinery (ACM) is an educational and scientific society uniting the world's computing educators, researchers and professionals to inspire dialogue, share resources and address the field's challenges. ACM strengthens the profession's collective voice through strong leadership, promotion of the highest standards, and recognition of technical excellence. As such, ACM cares deeply about the dependability and reliability of computing technology. Voter registration systems encompass not only the databases that house voter information, but also an entire information technology infrastructure that must be carefully managed by election officials. The U.S. Public Policy Committee of the ACM (USACM) commissioned this study to provide objective technical information and expert recommendations to state and local election officials, policy makers, and the public about these systems.

The USACM serves as the focal point for ACM's interaction with U.S. government organizations, the computing community, and the U.S. public in all matters of U.S. public policy related to information technology.

Supported by ACM's Washington, D.C., Office of Public Policy, USACM responds to requests for information and technical expertise from U.S. government agencies and departments, seeks to influence relevant U.S. government policies on behalf of the computing community and the public, and provides information to ACM on relevant U.S. government activities. USACM also identifies potentially significant technical and public policy issues and brings them to the attention of ACM and the public.

More information about ACM may be found on the World Wide Web at http://www.acm.org, and information on USACM may be found at http://www.acm.org/usacm.

Table of Contents

"An adequate and effective registration will go far toward assuring honesty and fairness in the conduct of elections. Upon the honest and faithful maintenance of the registration books depends the purity of the ballot box. And upon the purity of the ballot box depends the success or failure of our democratic form of government."

-- *Registration of Voters in Louisiana*, Alden L. Powell and Emmett Asseff, Bureau of Government Research, Louisiana State University, 1951

Executive Summary

The voter registration process may seem simple to most voters. They give their names, addresses, birth date, and in some cases party affiliations to election officials with the expectation that they will be able to vote on Election Day. In reality, election officials must oversee a complex system managing this process. They must ensure that the voters' information is accurately recorded and maintained, that the system is transparent while voter information is kept private and secure from unauthorized access, and that poll workers can access this information on Election Day to determine whether or not any given voter is eligible. A well-managed voter registration system is vital for ensuring public confidence in elections.

State and local governments have managed voter registration using different approaches among different jurisdictions. In 2002, Congress sought to make these disparate efforts more uniform by passing the Help America Vote Act, which required that each state have a computerized statewide voter registration database. In implementing this mandate, state and local governments still have differing approaches, but it is clear that information technology underpins each of their efforts. While technology will help election officials manage this complex system, it also creates new risks that must be addressed.

This study focuses on five areas that election officials should address when creating statewide voter registration databases (VRDs): accuracy, privacy, usability, security, and reliability. Each chapter contains detailed discussions and recommendations. The following are some of the overarching goals for VRDs and selected recommendations for achieving them.

1. The policies and practices of entire voting registration systems, including those that govern VRDs, should be transparent both internally and externally.

VRDs control access to voting; therefore, they have a direct impact on the fairness of elections, as well as the public's perception of fairness. It must be possible to convince voters, political parties, politicians, academics, the press, and others that VRDs are correct and are operating appropriately. Internal procedures and interfaces also must be clear to election workers in order to minimize errors. Transparency can be provided by allowing voters to verify their voter registration status and data; publicly disclosing outside data sources that officials use for verification; indefinitely keeping a secure write-

once VRD archive in electronic form to allow audits of previous elections; and using independent experts to audit and review VRD security policies. Other goals such as accountability, audits, and notification also support transparency and are discussed below.

2. Accountability should be apparent throughout each VRD.

It should be clear who is proposing, making, or approving changes to the data, the system, or its policies. Security policies are an important tool for ensuring accountability. For example, access control policies can be structured to restrict actions of certain groups or individual users of the system. Further, users' actions can be logged using audit trails (discussed below). Accountability also should extend to external uses of VRD data. For example, state and local officials should require recipients of data from VRDs to sign use agreements consistent with the government's official policies and procedures.

3. Audit trails should be employed throughout the VRD.

VRDs that can be independently verified, checked, and proven to be fair will increase voter confidence and help avoid litigation. Audit trails are important for independent verification, which, in turn, makes the system more transparent and provides a mechanism for accountability. They should include records of data changes, configuration changes, security policy changes, and database design changes. The trails may be independent records for each part of the VRD, but they should include both who made the change and who approved the change.

4. Privacy values should be a fundamental part of the VRD, not an afterthought.

Privacy policies for voter registration activities should be based on Fair Information Practices (FIPs), which are a set of principles for addressing concerns about information privacy. FIPs typically address collection limitation, data quality, purpose specification, use limitation, security safeguards, openness, individual participation, and accountability. There are many ways to implement good privacy policies. For example, we recommend that government both limit collection to only the data required for proper registration and explain why each piece of personal information is necessary. Further, privacy policies should be published and widely distributed, and the public should be given an opportunity to comment on any changes.

5. Registration systems should have strong notification policies.

Voters should be informed about their status, election information, privacy policies of the government, and security issues. As with audit trails, notification procedures can improve transparency; however, they are not always widely embraced. A recent survey found that approximately two-thirds of surveyed states do not notify voters who have been purged from election rolls. Voters should be notified by mail about their polling places, any changes that may affect their ability to vote, or any security breaches that expose private data.

6. Election officials should rigorously test the usability, security and reliability of VRDs while they are being designed and while they are in use.

Testing is a critical tool that can reveal that "real-world" poll workers find interfaces confusing and unusable, expose security flaws in the system, or that the system is likely to fail under the stress of Election Day. All of these issues, if caught before they are problems through testing will reduce voter fraud and the disenfranchisement of legitimate voters. We recommend many different ways to test various aspects of VRDs throughout the report. Examples include, evaluation of VRD interfaces by laypersons and experts for consistency, feedback, and error handling; testing interfaces with real-world users and conditions, including extreme or sub-optimal conditions such as high processor load or network congestion; and allowing thorough, independent evaluations of the security and reliability of the VRD.

7. Election officials should develop strategies for coping with potential Election Day failures of electronic registration databases.

VRDs are complex systems. It is likely that one or more aspects of the technology will fail at some point. Different strategies can be employed to adjust for various failures. For example, Election Day verifications can be done via any of the following: paper systems, personal computers or hand-held devices with DVD-ROMs or other methods of holding static copies of the voter list, or via personal computers or hand-held devices connected by electronic communication links to central VRDs. Regardless of the method used, a fallback process should be devised to deal with a VRD failure. When appropriate, these processes should operate in tandem with provisional balloting and other measures designed to protect the voters' right to vote.

8. Election officials should develop special procedures and protections to handle large-scale merges with and purges of the VRD.

One of HAVA's main requirements is that VRDs be coordinated with other state databases (such as motor vehicle records). Ensuring that voter records reflect up-to-date information from other databases can improve the accuracy of VRD, but coordination can introduce errors from the same databases, thereby undermining accuracy. Because large-scale merges and purges can render voters ineligible, the action should only be performed by a senior election official with procedures that force some sort of manual review of the changes. Further, if large-scale purges occur, they should be done well in advance of any election, and anyone purged from the database should receive notification so that any errors can be corrected.

Conclusion. State and local election officials face an ongoing and challenging task in creating and implementing statewide voter registration databases. We hope that the discussion and recommendations in this report will help inform officials and the public on how to meet these challenges.

In issuing this report, we recognize that many states have been working diligently

toward meeting the federal requirement to have an operational statewide VRD. Both because many states will not meet this deadline, and because there will be ongoing maintenance and changes to any such system, state and local governments will also face the issues identified in this report well beyond the federal deadline. For this reason, we offer our continued guidance to officials who may wish to discuss any of the topics raised in this report.

Chapter Overviews and Recommendations

Accuracy

Databases are only as good as the data they contain. Quality assurance is a challenge for any database because data entry and necessary merges and purges of data within the system can create errors. Maintaining accurate VRDs is even more difficult considering the mobility of the U.S. population[1] and the wide variety of information sources voting officials must use to verify registration records. Further, voting officials must balance between competing concerns of ensuring that each legally registered voter can cast his or her vote and preventing ineligible voters from casting votes. Accuracy concerns often lie at the center of these debates. An additional complication is that voter eligibility rules are determined state-by-state, and VRD design and implementation are likely to differ state-by-state.

Accuracy Recommendations

Voter Verification

- Voters should easily be able to determine if they are registered.
- Voters should be able to verify that they are registered through the use of a computer or handheld device located at any of the polling places in that state. Responses should not include personally identifiable information about the potential voter.
- Voters should be able to view the relevant contents of their voter registration records to check for accuracy and should be provided with easy-to-use mechanisms and contacts for correcting errors.
- Electronic Election Day updates to registration records are risky and should be implemented only after careful testing, if at all. Paper forms are a well-understood alternative.

Notice

- Whenever a voter or potential voter is determined to be ineligible to vote, the reason and source of information for the determination of ineligibility should be noted in the VRD for the potential voter to review and contest, if appropriate.
- Voters should be notified when their records change in any way that affects their eligibility to vote.
- Public notice of polling places should be provided well in advance of an election (e.g., signs in neighborhoods, prominent notices on local web sites).
- Each registered voter in the VRD should be mailed a postcard with his or her assigned polling place and registration status in advance of the election.

[1] A recent report of the Commission on Federal Election Reform found that "during the last decade, on average, 41.5 million Americans moved each year."

Polling Place Lists

- Polling place lists (whether paper or electronic) of all registered voters associated with a particular polling place should be generated automatically by the VRD well before Election Day.
- Automatically generated lists should be carefully checked by at least two local officials and far enough in advance of elections to allow time for corrections.

Archiving

- Ineligibility records should be retained in the VRD for at least twenty-two months and possibly longer.
- If for any reason it is determined that an individual is ineligible to vote, that individual's record should be marked accordingly, not deleted.
- When information is sufficiently old (we recommend at least 22 months), it should be moved from the VRD into an offline archival database that is never purged and is protected against unauthorized disclosure or access.

Other Databases

- When other databases, such as driver registration databases, are used to check for eligibility, those databases should be used for screening and not to automatically enroll or de-enroll voters.
- An automated check can be used to flag some voters for further scrutiny, but the final determination of eligibility should be performed only by an appropriate election official.

Merges, Purges, and Batch Updates

- Large-scale automated database merges are error-prone and should be avoided if possible.
- If purges are performed, they should be done well in advance of any election. People whose names are purged from the VRD should receive notification in sufficient time for them to be able to correct any errors.
- A greater level of authority should be required to perform a batch update than is required to make smaller changes.

Accountability

- There must be well-defined accountability for all changes to the VRD including to source code, database schemas, database contents, and system configuration.
- Changes should require approval or sign-off by an authorized individual.
- It should be possible to identify a clear chain of responsibility for each change, and the VRD should be designed to facilitate tracking of this information.

Audits

- A complete audit trail should log all modifications to the VRD.

Privacy

The public is increasingly aware that personal information in electronic form can pose new risks, such as identity theft, to personal privacy. As state and local governments digitize, centralize, and share this data, the stakes are raised still higher. While VRDs may pose threats to privacy, technology also opens up new opportunities to protect privacy. As governments design and implement these systems, privacy values must be considered a fundamental part of the design process, not simply applied as an afterthought.

Privacy policies for voter registration activities should be based on Fair Information Practices (FIPs), which are a set of principles for addressing concerns about information privacy. FIPs typically address collection limitation, data quality, purpose specification, use limitation, security safeguards, openness, individual participation, and accountability.

Privacy Recommendations

Openness (Transparency)

- Publish on the main election board website a complete notice of policies and practices describing the collection, maintenance, use, and disclosure of voter registration data. The notice should include contact information for the office or the officials responsible for voter registration data.
- Publish a readable summary notice in other places, such as voter registration forms, at polling places, on sample ballots, and elsewhere as appropriate.
- Provide a copy of the complete notice to any person who requests it.
- Publish any changes to the notice before the changes become effective, and accept and consider public comments.
- Place a date and version number on notices as they are published. Maintain, and make publicly available, copies of all previous notices, including the periods of time during which they were effective.

Data Collection Limitation

- All data should be collected by lawful and fair means.
- Data should be collected, where appropriate, with the knowledge or consent of the subject.
- Registrants and the public should be informed through the published notice of policies and practices of the sources of all data obtained for voter registration purposes.
- The types of data elements to be collected should be subject to public scrutiny.
- Data collection should be limited to sources and procedures authorized by law and properly described in the published notice.
- Only the minimum information necessary for, and relevant to, voter registration

purposes should be collected and maintained. The reason for collecting each type of personal information should be explained, and the specific data elements collected should be subject to public scrutiny.

Use and Disclosure Limits

- Limit use and disclosure of voter registration data to activities directly related to the election process or to other activities expressly authorized by law.
- Describe all uses and disclosures in the published notice of information practices. Identify publicly all recipients of voter registration data.
- Provide public notice of and, if possible, a chance for public comment on all disclosures of identifiable voter registration data for any activity not directly required for voter registration purposes.
- Restrict access to specific records, specific data elements, and specific classes of voters (e.g., by location) to those election officials who have a need to use those records, data elements, and classes in the performance of their duties.
- For some or all uses by election officials or disclosures to external parties, maintain a record of the date, nature, and recipient of all personal information and make the record accessible to the data subject upon request.
- Restrict disclosures to specific data elements permitted by law and necessary to accomplish the purpose of the disclosure. Withhold data elements that are not essential to accomplish the purpose of the disclosure or that would place data subjects in excessive jeopardy to identity theft or other improper activities.
- Prevent recipients of data from using or redisclosing the data in ways not specifically authorized by law. Asking recipients to sign data use agreements is one way to accomplish this purpose.
- Allow some non-essential uses and disclosures only with the affirmative consent (opt-in) or negative consent (opt-out) of the data subject.
- For some data subjects at risk (e.g., victims of spousal abuse, jurors, some public officials), it may be appropriate to further limit disclosures.
- Even the best use and disclosure policies may be violated by people and software within the election process. Therefore, limit access by each person and each system component.
- Provide access for every voter to a personalized list of those third parties who have been given or purchased access to his or her voter registration data.

Usability

VRDs will be used in many ways by a wide variety of people. Ensuring that well-trained election officials, minimally trained volunteer poll workers, and voters with little to no technical skills can all use different and appropriate aspects of VRDs is a key challenge for designers of these systems. Poorly designed user interfaces might confuse users or, worse yet, disenfranchise voters. This can create the reality or the perception of an unreliable system, thereby undermining the entire process.

Usability Recommendations

General Usability

- Consider the various types of users, tasks, and environments in which the voter registration database will be used. Design user interfaces that address all of these factors, providing different interfaces for different combinations as necessary.
- Use accepted user interface design techniques to build data entry forms and data retrieval components that are clear, usable, and interpretable.

Design and Features

- Involve a wide range of test users of different backgrounds, skills, literacy levels, ages, and roles (county official, election volunteer, voters, etc.) in all stages of user interface design, including gathering of usability requirements, design of user interfaces, and testing and evaluation.
- Treat user interface design as an iterative process: use evaluations of user interface designs to guide revisions that themselves can be evaluated in turn.
- Provide informative feedback (i.e., provide users with detail sufficient for understanding the impact of their actions, results of queries, and characteristics of the current operating environment).
- Eliminate unnecessary functionality and data output in favor of simple, minimal user interfaces.
- Provide online tutorials and help systems for all voter registration database user interfaces. For critical applications such as voter verification on Election Day, appropriate experts should be available to help address any concerns.
- Ensure that public-facing interfaces (e.g., World Wide Web based services) are vendor-neutral and are designed to meet widely accepted technical standards.

Evaluation and Testing

- Use a variety of user interface evaluation techniques, including heuristic evaluation by usability experts, "think-aloud" sessions, and user studies.
- Test interfaces thoroughly with representative users performing tasks under situations that approximate those likely to be found in real use.
- Test user interfaces under extreme or suboptimal conditions, including high processor load, network congestion, and noisy or extreme environments.
- Test web-based user interfaces for use by the public on as wide a range of browsers as possible, including multiple older (and pre-release) versions of popular browsers and screen-reader systems for people with visual impairments.
- Evaluate user interfaces, particularly web-based interfaces, to determine their impact on other system goals such as reliability, security, accuracy, and privacy.

Security

Security underpins each of the issues discussed in this report. Maintaining accurate and

private information is impossible if a VRD is vulnerable to malicious attack. Further, the validity of data within the VRD may be called into question if the system is easily compromised or lax security policies are established. Ultimately, an unsecured VRD could undermine elections. Good security policies address many different factors. Election officials should establish detailed access controls for each user accessing the VRD, procedures to harden VRDs from attack, and mechanisms to deal with and recover from security failures.

Security Recommendations

Designing & Implementing Access Control Policies

- Federal, state, and local election officials should work together to establish a common framework for access control policies, such as common roles and responsibilities of users and their levels of access, as well as who would be responsible for ultimately implementing and enforcing access control policies.
- Access control policies should not grant the same privileges to all users; rather the policies should group people by established roles and geographic areas. For example, the security policy might give the same level of privileges to all data entry officials for a particular county, but privileges should be different for VRD administrators.
- Access control policies should minimize the number of people who receive privileges both to access each piece of information and to grant access to others.
- Access control policies should ensure that each person is granted only the minimal set of privileges needed to do his or her job.
- Access control policy should cover all records stored in the VRD including records on both voters and non-voters.
- VRDs should use access control mechanisms provided in the database management systems provided; trying to implement access control entirely at the application level leaves greater opportunity for security mechanisms to be bypassed or compromised.
- VRDs should create public logs of all changes to the list of authorized users and their access rights, and any changes to either of these should require authorization from two different persons.
- Authorized users of the system should receive security training, including how to protect passwords and how to resist social engineering attacks (attempts to deceive someone into performing certain actions), and the importance of never sharing passwords.
- Older versions of access control policies should be retained, along with their dates of applicability, and possibly made available to the public to increase the transparency of the system.

Administrative Privileges and Emergencies

- The number of people with administrative privileges for the VRD should be limited; very few users should have the ability to grant access to others.
- People with administrative access should not be allowed to grant themselves new access privileges unilaterally; rather, such a change should require the consent of another administrator.

- Officials should create rules that allow trusted election officials to increase temporarily the privileges available to others during emergencies in a controlled and fully audited manner.
- Emergency overrides should require two-person authorization and generation of detailed audit logs.

Security Metrics

- Those responsible for managing VRDs should measure how effectively they have limited VRD users' privileges by determining how many people have access to how much data and by tracking effectiveness over time using these metrics.
- The EAC or some other appropriate organization should help develop and identify appropriate metrics.

Protecting Against Attack

- All communication channels used by the system should be secured. Anything transmitted over open communication networks, such as any wireless connection, the Internet, or the phone system, should be protected using end-to-end cryptography.
- Firewalls should be used to severely limit connectivity between internal and external networks.
- Mechanisms should be deployed to detect any penetration of system defenses or any insider misuse.

Dealing with Security Failure

- It must be possible to recover from security failures (e.g., retaining historical copies as well as the latest, regular backups with offsite storage, etc.)
- Officials should obtain independent security reviews of the VRD before system deployment and periodically thereafter.
- Individuals should be notified if an inappropriate person may have obtained their data.

Reliability

Because VRDs control access to voting, they must meet a very high standard for reliability. If the system fails, it may disenfranchise voters and undermine public confidence in elections. VRDs should be designed to be reliable both during the non-peak times before and after an election, and for high-activity times such as Election Day. While reliability issues are often considered in terms of "always on" electronic systems, registration systems may be economically designed to employ both online VRD and offline solutions, such as distributing DVD-ROMs of registration data to polling places for use on Election Day. State and local governments should assess the entire scope of reliability issues and design systems that have built in redundancy, replication, and distribution, but also incorporate mechanisms that allow the voting process to proceed should the VRD fail. States may choose to implement the VRD by centralizing the

database at the state level or decentralizing it and spreading responsibility among the different local jurisdictions; officials must recognize that reliability issues differ depending on the chosen implementation.

Reliability Recommendations

Redundancy

- Use redundancy to alleviate failures affecting time-critical operations.
- Ensure that redundancy actually increases reliability by conducting system failure tests.

Replicated Data

- There should be multiple copies of the database.
- Copies should be physically separated to protect against physical damage.
- Copies should be logically separated (i.e., in different forms/types of systems) to protect against software failure and attacks.
- The data on physically separate copies (such as DVD-ROMs) should be encrypted. Encryption and decryption mechanisms should be tested.
- Different channels, including alternate network providers and routes, physical media, and printed copies to access different replicas should be provided.

Distribution

- Evaluate the ability of individual databases to function when other parts of the system fail.
- Evaluate distributed database solutions with respect to their ability to meet the HAVA-mandated goal of a single, uniform, official, centralized, interactive computerized statewide voter registration list.

Centralization

- Evaluate the ability of the system as a whole to respond to the unavailability of one or more copies of the centralized database.

Archives

- All changes to the database that affect the ability of an individual to vote must be logged and archived.
- Archival media, including audit logs and backups, must be write-once or otherwise protected to ensure that accurate records of changes to the VRD have been maintained.

Election-Day Fallback Processes

- Develop fallback processes for registration verification so that elections can proceed

even in the face of system failures.

- Ensure that fallback processes will withstand any failure that would not otherwise prevent voting. If a power failure at a polling place does not prevent use of voting machines, then it should not prevent voter registration checks to be performed.

Provisions for Delayed Entry of Registration Information

- Develop processes supporting delayed entry of registrations.
- Analyze the impact of near-deadline registration and early/absentee ballots on the system.

Testing

- A defined and empowered quality assurance group should be in place from the beginning of the project. The group should develop functionality, usability, and reliability tests.
- Periods of peak stress (e.g., immediately before registration deadlines, during elections, and registration verification) should be identified for reliability testing, as should the activity mix during periods of peak stress. Consider questions such as how many simultaneous users or operations are expected, and identify all potential component failures. Testing should check whether system performance will be adequate even when some system components have failed.
- Tests for security against likely attacks (e.g., denial-of-service attacks) should be conducted.

1. Introduction

The Help America Vote Act of 2002[2] (HAVA) mandates that each state create a single, uniform, official, centralized, and interactive computerized statewide voter registration list by 2006. The requirement that the list be both interactive and computerized implies that the only compliant implementation will be as a database. While the goal of mandating the use of databases is to improve and streamline aspects of voter registration, inappropriately designed or implemented databases will have serious negative impacts on the accuracy of elections and on public perception.

In this report, we describe the characteristics that centralized voter registration databases should possess. While some recommendations might not be relevant to some systems, most of our recommendations should be implemented if systems are to be accurate, usable, secure, reliable, and appropriately protective of voters' privacy. In those cases in which systems have already been designed or built, election officials should consider modifications if our recommendations have not yet been included.

We start with an overview of voter registration databases and the Help America Vote Act and then provide technical recommendations.

Voter Registration Databases (VRDs). VRDs are statewide databases of registered voters. With the exception of North Dakota, which is the only state that does not have voter registration, voter registration rules are created at the state level.[3] Prior to the Help America Vote Act, local jurisdictions maintained lists of voters, with list formats and uses varying from jurisdiction to jurisdiction. In general, the lists can consist of the following:

- full legal name,
- date of birth,
- last four digits of the social security number,
- driver's license number,
- address of residence (to assign the precinct),
- mailing address,
- phone number,
- place of birth,
- party affiliation (so the correct election materials can be sent before primaries, and correct ballots can be given at primaries), and
- validity status, noting whether the record is for a valid voter, or if the registrant is not currently allowed to vote.

Some jurisdictions may request the full social security number and a digital image of the individual's signature for visual verification of mail-in ballots and initiatives. Jurisdictions may also retain voting history of registered voters and remove invalid

[2] Public Law No. 107-252, 116 Stat. 1666 (codified at 42 U.S.C. §§ 15301-15545), available online at http://www.fec.gov/hava/law_ext.txt.

[3] For more information about state voter registration deadlines, see http://www.eac.gov/register_vote_deadlines.asp.

registrations from the voting rolls. Invalid voter registrations can occur if a voter has not voted in several elections, has died, or has moved outside of the jurisdiction. If a record indicates that someone is not currently a valid voter, that individual must reregister. Some jurisdictions also include an indicator on a voter's record as to whether or not the address and phone number are to be given to outside organizations.

Election officials mail election materials, such as mail-in ballots and polling place addresses, to the voters listed in the VRD. *Polling books or voter rolls* derived from the VRD enable local officials to verify that a voter is registered in the precinct served by a particular polling place and that the voter has not previously voted in the election via a mail-in ballot or early voting. Polling books can be printed on paper or they can be digitized and put on personal computers or electronic handheld devices, often called electronic polling books. While these devices may differ in design, in general they connect either by phone lines or a wireless link to a master location that has the polling information, or they are stand-alone and contain copies of polling information. VRDs also may be used to produce lists of voters, including names, addresses, and party affiliations. Such lists frequently are used by outside groups to send voters election-related materials and to call voters in get-out-the-vote campaigns. VRDs typically are the basis for Internet-based voter information applications that enable people to determine if they are registered and where their polling places are located.

Standards. In light of recent events and legislation that have underscored the core importance of voting and of public confidence in our electoral system, one might conclude that all VRDs should be built and operated to the highest possible standards. While the highest standards of reliability, privacy, accountability, usability, and security are desirable, they may at times be impractical because of resulting expense or system response. Nonetheless, where practical and reasonable, the highest standards should be applied.

Standards for reliability, privacy, accountability, usability, and security allow for a wide range of applications and choices. Conventional commercial products and normal practices, which may be suitable for business or governmental applications, might not satisfy the difficult political and operational demands of voter registration systems. The cost of failure for a VRD, which may include a major loss of confidence in our political system and institutions, must be considered in the standards-setting process along with the other traditional costs that are the normal subject of evaluation.

This report discusses some standards that exceed the average commercial application for database software. While a higher standard may be recommended or included in a list of options for consideration, the ultimate decisions about standards obviously are not ours to make. We hope that those decisions will be made with an awareness of and sensitivity to the requirements essential to maintaining a high degree of public confidence in our electoral system.

The Administration of HAVA. The U.S. Election Assistance Commission (EAC), created by HAVA, is charged with, among other things, assisting states in the administration of Federal elections and establishing minimum election administration standards. It also provides states federal grants to replace punch card voting systems and

to establish and maintain statewide voter registration lists.[4] The cost of developing and maintaining voter registration lists could be more than half the overall cost of administering elections.[5]

Prior to HAVA, voter registration records often were maintained on a county or other local level; these records frequently were not coordinated across counties. What is new with HAVA is the aggregation of all records statewide under a central administration and in electronic form, thereby creating new challenges, risks, and opportunities.

We address a variety of issues in this report with the understanding that many states are nearing the completion of the HAVA-mandated implementation.[6] As computer systems are rarely finished, it is likely that the VRD implementations will continue to be developed and enhanced and that our recommendations will be relevant well beyond the initial implementations.

Other Studies. This report focuses on the technology aspects of VRDs. There are several other studies that discuss different aspects of VRDs. For example, "Balancing Access and Integrity, The Report of the Century Foundation Working Group on State Implementation of Election Reform"[7] has an excellent chapter on VRDs. This study, while not as detailed as ours, includes more policy-related issues.

The California Voter Foundation has an outstanding study, "Voter Privacy in the Digital Age,"[8] that details how information on voter registration lists is gathered and used. "Assorted Rolls, Statewide Voter Registration Databases Under HAVA" by Electionline.org, is a complete snapshot of the States' different plans and implementations of HAVA-mandated statewide VRDs. The Appleseed Foundation, in a joint effort with Latham & Watkins and the Brennan Center for Justice, released a best practices guide in 2005 offering guidance to states in developing their VRDs.[9]

Scope. We make some assumptions to narrow the scope of our report to the kinds of VRDs that are actually being used by the states.[10] For example, we assume that the VRD is implemented as an application using a commercial off-the-shelf database system. Commercial *database management systems* (DBMSs) are reliable, affordable, and have many features that are needed for the VRD application. However, the use of a commercial DBMS is only part of the implementation. Applications built on top of a

---

[4] 42 U.S.C. § 15322.

[5] Ace Project, Voter Registration Overview web page, http://www.aceproject.org/main/english/vr/vr10.htm.

[6] Electionline.org, 2005, "Assorted Rolls: Statewide Voter Registration Databases Under HAVA," Election Reform Briefing 11, June, available online at http://www.electionline.org/Portals/1/Assorted%20Rolls.pdf.

[7] Century Foundation Working Group on State Implementation of Election Reform, 2005, "Balancing Access and Integrity," available online at http://www.reformelections.org/publications.asp?pubid=542.

[8] California Voter Foundation, 2004, "Voter Privacy in the Digital Age," available online at http://www.calvoter.org/issues/votprivacy/pub/voterprivacy/index.html.

[9] Appleseed, 2005, "The Database Dilemma: Implementation of HAVA's Statewide Voter Registration Database Requirement," available online at http://www.appleseeds.net/download/Appleseed_Brennan_HAVA_Users_Manual.pdf.

[10] Electionline.org, op. cit.

DBMS include user interfaces, system design, and the implementation of various security and reliability policies.

Commercial DBMSs have features that are necessary for the VRD application such as *transaction logs* and *audit logs* that maintain records of changes to the data and database design. The systems also provide mechanisms to *backup* the database. A backup is a complete copy of the database at a known point in time. Transaction logs are used together with backups to rebuild the system if there is a problem, restoring the data to its state at the time of the backup. Audit logs are used to determine if suspicious updates have occurred. Commercial DBMSs also provide *access protections*, so that only users with the correct authorizations can access given data.

VRDs may be *top-down*, *bottom-up*, or some combination of the two.[11] In a top-down approach, state officials administer a single master computer server; all voter records are stored on that central server, and all requests to view or modify voter records are executed on the central server. In a bottom-up approach, each county may keep its own database of records for voters within the county, and the county's records may be reconciled with a database run by the state on a periodic basis.[12] Although these two approaches have some different properties, most of the issues that we discuss apply independent of whether the VRD is top-down or bottom-up. Therefore, when we refer to the VRD, it is worth keeping in mind that this database may in fact be implemented by a collection of computer systems working cooperatively to store and maintain voter registration records.

Software Development. Sound principles of project management must be followed when developing software. The knowledge of the people currently working in the local offices, who may be experts in voter registration, should be assessed. A single person should manage the software development project and also bear responsibility for its success.

Those working on the development of the VRD must be trained professionals who have implemented database systems, preferably with the development tools of the chosen vendor. In addition, from the beginning there must be a trained quality assurance group that is continuously testing the design and ultimately the implementation to make sure that the application is reliable and accurate.

[11] The Electionline.org briefing cited above contains an excellent discussion of the distinction between the two and why both can be considered HAVA-compliant.

[12] 42 U.S.C. § 15483(a)(1)(A)(vi) ("All voter registration information obtained by any local election official in the State shall be electronically entered into the computerized list on an expedited basis at the time the information is provided to the local official.") The EAC Voluntary Guidance has interpreted "expedited" as meaning "at least every 24 hours."

2. Accuracy

Maintaining the accuracy of VRDs requires balancing two opposing concerns. The first concern is that a VRD needs to be inclusive to avoid disenfranchising legitimate voters. The names of all people who have registered and are duly eligible to vote must be included in the VRD; any omissions will exclude eligible voters from voting. The second, somewhat contrary concern is that the VRD must not be overly inclusive. To prevent fraud, only legally registered persons should be listed in the VRD as eligible to vote. We will address both of these concerns.

Not only must VRDs be accurate, the public must also believe that they are accurate. Because VRDs control access to voting, transparency is critical. It must be possible to convince those with interests in elections—including voters, political parties, politicians, academics, and the press—of the correctness of the VRDs. To provide transparency, policies should minimize the possibility of error and facilitate the correction of errors. Election officials must also take responsibility for ensuring adherence to these policies.

Data Entry and Errors. Most errors in individual database records occur during data entry. Errors include misspelling of names and addresses, incorrect recording of unique IDs, misidentification of people to whom access to the system should be allowed or denied, and misdirecting voters to the wrong polling place.

Data is entered into the VRD using one of two methods: manual entry or via automatic scanning devices. An automatic scanning device is a machine that looks like a copier and is used to scan a document into a computer system. Once the document is scanned in, software that can recognize characters transfers the data from the printed form into the VRD, while providing a clerk with the opportunity to correct mistakes. For either manual entry or automatic scanning, a well-designed user interface for the clerk will reduce errors. (Chapter 4 on usability contains further discussion of user interfaces.)

While quality control systems and appropriate supervision of data entry may reduce data entry errors, some errors will inevitably occur. Problems can arise because of variations of name spellings (Stevens or Stephens), first and last names that use accent marks or more than one capital letter (McMullen), and names that have no vowels (Ng). Incorrect or incomplete spellings of street names are additional potential sources of errors. Changes that are primarily entered in other state databases—such as changes in marital status and court approved name changes—also compound the challenge to accuracy.

Voter Verification and Notice. To minimize the impact of errors in the VRD, voters should be provided with (1) opportunities and methods to view and verify their data, and (2) notices about changes to their records. For example, the system might provide an Internet website or automated telephone service where voters can examine parts of their records, check their registration status, and determine their assigned polling places.

Whenever a voter or potential voter is determined to be ineligible to vote, the reason and source of information for the determination of ineligibility should be included in the VRD. This information should be retained so that someone who has been inappropriately labeled as ineligible can easily challenge the decision and demonstrate that an error has occurred.

Finally, election officials should mail each registered voter in the VRD a postcard with his or her registration information and information necessary for voting, such as polling place location or instruction for voting by mail. Voters also should be notified when their registration status changes. A voter removed from the rolls or reassigned to a new polling place should be notified by mail of the change and be provided an opportunity to seek correction if the change is an error. A voter recorded as having moved should be notified by mail sent to both the new address and the old address (similar to the method the United States Postal Service uses with respect to change of address forms).

To help correct errors in voting records, contact information for the person or office responsible for complaints and questions should be provided to voters. Further, voters and system administrators should understand how complaints and errors are addressed, and voters should receive feedback explaining the reasons for a final determination.

One recent survey found that approximately two-thirds of surveyed states do not notify voters who have been purged from the election rolls.[13] Advance notice, which can be facilitated by the VRD, would provide voters with an opportunity to identify mistakes prior to an election. Care must be taken in designing such systems so that violations of privacy and security do not occur.

Notification processes are not always foolproof. For example, in 2004, 8,800 Maricopa County, Arizona, residents received election notification cards listing the wrong polling places in the wrong cities.[14]

To help minimize the impact of incorrect notification, we recommend that public notice be provided well in advance of an election. That notice should include the polling place's geographic location and official name (school, church, library name), a description of the exterior of the polling place to assist voters in locating the entrance, times of poll operation, residential boundary lines, and corresponding zip codes.

Some states allow voters to verify that they are registered through an Internet web site or by phone. For states that use Internet verification the user interface should protect voters' privacy by requiring the voter to provide his or her name and address and limiting the response to "yes, you are registered to vote and here is where you go" or "no, you are not registered to vote." The response should not include personally identifiable information about the potential voter.

Some provision needs to be made to deal with corrections on Election Day because not all errors can be corrected in advance. Poll workers are likely to be preoccupied with running an election and should not be allowed to make changes to the VRD. Under the right circumstances, after extensive testing for accuracy and usability, it might be possible to allow poll workers to send electronic reports of needed changes to election workers. If such a system is implemented, the updates would need to satisfy the auditing and authorization requirements discussed elsewhere in this report.

A simple alternative is to provide paper forms that are filled out at the polling place and submitted to election workers after the close of the election.

Generating the List of Registered Voters. A printed voter registration list for those precincts served by a polling place is typically used to verify registered voters. While

---

[13] Electionline.org, op. cit.

[14] Dennis Wagner, 2004, "8,800 Voting Cards Have Wrong Poll Address," *The Arizona Republic*, October 27, p. B5.

these printed lists are convenient and easy to control, sometimes the wrong list is provided to a polling place. To minimize the chance of the delivery of an incorrect list, we recommend that automated generation of polling place lists be used as much as possible and that the lists be carefully checked by at least two people. Local officials can conduct these checks, but they need to be made far enough in advance of elections to allow time for corrections.

Incorrect voter lists could be delivered to polling places independent of whether the data are provided on paper, DVD-ROMs, in a PC, or in a handheld device. In all of these cases, a computer operator might provide incorrect directions to the computer, resulting in the wrong electronic list going to the polling place. As with paper printouts, we recommend that electronic versions of voter lists be checked by at least two people well in advance of elections to allow time for corrections.

Information Deletion and Retention. In addition to being a list of currently registered voters, a VRD is a comprehensive set of records reflecting voter registration activity and administration. Consequently, we recommend that after records appear to be no longer relevant, they be retained in the VRD at least for the next two Federal elections or for the statutorily-mandated minimum of twenty-two months.[15] The retained record should include a dated annotation stating that the voter is not eligible to vote, along with the reason for ineligibility. Thus, a VRD might contain information about those who have applied, been approved, been questioned, died, moved, or been denied the right to vote, as well as those who currently are eligible to vote.

When records were stored on paper, retaining old records imposed a non-trivial administrative burden. Electronic databases have made the cost of retention negligible, so old information can be retained relatively easily and inexpensively. When information is sufficiently old, it should be moved from the VRD into an offline archival database that is never purged. Retention of such information will enhance transparency and facilitate the correction of errors such as those that can occur when voters are thought to have died, moved, been convicted of a felony, or otherwise determined not to be eligible to participate in a public election.

Other Databases. HAVA requires that states authenticate each potential voter by cross-checking with other state databases—in particular, databases of driver's licenses.[16] If a potential voter does not have a state driver's license, then the last four digits of the voter's Social Security number must be used for authentication.

Because other databases can be inaccurate as a result of ambiguous or incorrectly entered data or computer-related problems, wholly automated procedures are risky. Consequently, we recommend that other databases not be used to enroll or de-enroll voters automatically. External databases could be used for initial screening, but an appropriate election official should perform any final determination of voter eligibility or

[15] The Civil Rights Act of 1960 requires that every officer of elections retain for 22 months registration and other voting records and papers for federal elections. 42 U.S.C. § 1974.

[16] HAVA provides for coordination of voters lists with other state agency databases (42 U.S.C. § 15483(a)(1)(A)(iv)) and requires that registration applications include either a current and valid driver's license or the last 4 digits of the applicant's Social Security number (42 U.S.C. § 15483(a)(5)).

ineligibility. We suggest that every change, addition, or deletion to the VRD require explicit approval by an individual authorized to make that change. We discuss how this might be done in Chapter 5 on security.

Errors can arise because of court-approved changes in legal name that conflict with existing birth records, motor vehicle records, or other state records. Name similarities also can create problems. For example, a death record database may show that Mr. John Smith who lives at 254 Vine St. has died. There may be a Mr. John Smith, Jr. living at the same address who is eligible to vote. If the death record database is applied with no cross checking, John Smith Jr. may learn on Election Day that he has been denied his right to vote.

Databases also can be inaccurate or unreliable because of computer viruses, programming errors, and system failures. For example, in 2003 the Maryland Motor Vehicle Administration (MVA) offices were attacked by a computer worm.[17] The worm shut down the MVA's computers and telecommunication systems, cutting them off from all forms of remote communication and disrupting operations in all 23 MVA offices located throughout the state. A second event occurred on January 20, 2004, when the MVA could not process work on the mainframe computer for about an hour after opening. The problem was characterized as a computer glitch.[18]

A further risk to the accuracy of databases is insider fraud, involving either the VRD itself or external databases, such as driver's license databases, that are used to authenticate voters.[19] Therefore, election officials should carefully consider if the accuracy and security of external databases is sufficient to meet voter registration needs. Risks associated with insider fraud are discussed further in Chapter 5 on security.

Avoid Large-Scale Merges and Purges. Computers make it easy to automate sweeping batch updates to a VRD; at the same time, errors can be magnified by the use of automation. In the context of VRDs, a batch update is a group of updates received from what is believed to be an authorized source (e.g., a local county). Because many voter records could be affected by a single batch transaction, a greater level of authority should be required to perform a batch update than is required to make individual changes. As is the case with all updates, election officials should develop policies and procedures to ensure the accuracy of large batch updates to the VRD. For example, a policy might prohibit batch updates affecting more than a maximum number of voters or jurisdictions (essentially requiring that large changes be broken down into multiple smaller batches that can be reviewed more effectively), or a policy might require individualized review and approval of each voter record that is affected. A policy might specify that batch updates be reviewed by several people or mandate that audits of a statistically-significant

---

[17] Christian Davenport and Hamil R. Harris, 2003, "MD's MVA Offices Forced to Shut Down," *Washington Post*, August 13, p. A09.

[18] "Glitch at MVA Branch Offices Delays Some Transactions for an Hour," 2004, *The Baltimore Sun*, January 21, p. B6.

[19] For example, a Maryland MVA employee was charged with conspiring with others to sell more than 150 state identification cards. See Eric Rich, 2005, "MD, MVA Employee Charged in ID Card Sales," *Washington Post*, April 23, p. B03. For a collection of stories of security problems of motor vehicle records, see Center for Democracy and Technology, *Tracking Security at State Motor Vehicle Offices*, available online at http://www.cdt.org/privacy/030131motorvehicle.shtml.

random sample of records in the batch be performed before approving the batch update.

Given the inaccuracies that exist in many governmental databases, large-scale automated merges between databases increase the risk of errors in a VRD.[20] Consequences of inaccuracies in other databases could result in the widespread disenfranchisement of eligible voters, the inclusion of ineligible voters in a VRD, or both.

We recommend special caution in deploying large-scale purges of VRDs. The move to a statewide VRD may make it tempting to attempt to automatically eliminate duplicates by comparing lists of eligible voters across counties, something that previously could not be done. However, automatic purges of duplicate entries could disenfranchise large numbers of legitimate voters. If large-scale purges occur, they should be done well in advance of any election, and all people whose names are purged from the VRD should receive notification in sufficient time for them to be able to correct any errors arising from the purge.

Accountability. Clearly defined accountability for all changes to the database is a fundamental requirement for helping instill voter confidence in VRDs. Voters, politicians, election officials, the press, and others should be able to determine who is responsible for changes to the VRD.

These changes include, changes to the data such as adding new voters, purging voter records, changing addresses, names, etc.; changes to the software configuration such as incorporating new software releases into the VRD; changes to the security policy and access rights; or changes to the database design. Any of these changes can adversely affect the data, so in order to provide the desired accountability there must be a record of each change, when it occurred, and who approved the change.

Audit Trail. The record of the changes to the VRD is called an *audit trail*. In order to ensure accuracy and transparency, VRDs must be auditable. VRDs that can be independently verified, checked, and proven to be fair will increase voter confidence and help avoid litigation.

The audit trail should include the record of all possible changes mentioned, namely, data changes, configuration changes, security policy changes, and database design changes. Although we call this an audit trail, it is not a single entity. The records of configuration, policy and design changes, including who approved them, can be kept in computer files or on paper as long as they are auditable by a third party. The record of changes to the data, because there will be many of them, must be kept in computer files to facilitate auditing.

In DBMS applications, there are typically two files generated because of a change to the database. The *transaction log* records in a file the data values before and after the change occurred, as well as the time of the change. The *audit log* records information about the user ID of the person who made the change. The transaction log is used to provide backup should a system failure occur.

The content of audit logs varies among DBMSs. In some, it is possible to configure the system so that the audit log tracks changes to the security of the system (the

[20] In 1988, Congress enacted the Computer Matching and Privacy Protection Act to address some of the unfairness and inaccuracies arising from federal government use of computer matching techniques. See Public Law 100-503, 102 Stat. 2507 (codified at 5 U.S.C. §552a).

permissions given to particular users), changes to the data, and changes to the database design. For the purposes of the VRD auditing requirements, this is not sufficient. The VRD should record not only which user made the change, but also the identification of the person who authorized the change. Therefore, it may not be possible to rely on the commercial DBMS's auditing capabilities alone for the audit trail that a VRD requires. VRD implementers will need to augment the application code of the commercial database audit log to provide a complete audit trail.

Well-maintained audit trails are critical because they may allow reconstruction of the circumstances of a system failure, thereby facilitating future improvements to access policies and possibly to the database itself.

Approval Mechanism. Given that there is an audit trail that records whose approval was given for each change, state or local officials must set policies on who is actually authorized to make changes. Access control polices are discussed in more detail in Chapter 5 on security. We assume that the person with ultimate authority to make the changes is an election official, and we recommend that the responsibilities and authorities of such election officials be clearly defined and publicly available.

For system changes, we recommend that there be a formal change control process that states how changes to the system configuration, security policy, and database design are reviewed, approved, and recorded.

Summary reports or excerpts from audit trails should be provided to supervisors and made available to external auditors. These reports should be inspected frequently for unusual or suspicious activities such as access from unexpected Internet Protocol (commonly referred to as "IP") addresses or at unusual times of day, surges in the number of accesses by a single user, and other anomalous activity.

Conclusion. Well-designed accuracy features must be accompanied by appropriate training and resources. Even the best designed VRD will be of little value if officials do not monitor and verify that only authorized changes are made to the VRD. Log files that are never read and system quality control processes that are not supervised will not ensure database accuracy. Since accuracy should be viewed as an ongoing responsibility, election officials should assign specific staff to oversee these continuing activities.

3. Privacy

Policies for voter registration activities should include appropriate protections for the privacy of identifiable data about individuals. A privacy policy should be based on Fair Information Practices (FIPs), a set of principles for addressing concerns about information privacy. FIPs typically address issues such as how data is collected, secured and used, and how policies regarding data practices are disseminated. Specific implementation recommendations are included in the discussion.

The increased computerization and sharing of voter registration records raises the stakes for privacy. While paper records also affect the privacy of data subjects, the risks are greater with electronic records, which may be more vulnerable to improper disclosures by more people. Furthermore, the scope of the disclosures can be much greater. A thief can carry only so many paper records, but an entire electronic database can fit unnoticed in someone's pocket.

Technology also brings opportunities for privacy improvements, making it easier to obtain and enforce the preferences of each voter for the use and disclosure of the voter's personal data. Technological tools also facilitate the tracking of data. To minimize the threats and maximize the benefits of technology for privacy, it is necessary to build the proper capabilities into VRDs.

Fair Information Practices, which form the basis of many privacy laws in the United States and around the world, help to assure that any system of personal information addresses all appropriate privacy elements. The Privacy Act of 1974,[21] a law that applies to federal agencies, was the first statutory implementation of FIPs anywhere in the world, and federal agencies have been operating under that law for more than 30 years.[22] Although there have been numerous restatements and versions of FIPs,[23] core principles address collection limitation, data quality, purpose specification, use limitation, security safeguards, openness, individual participation, and accountability.

While FIPs provide a useful framework for information privacy, the principles are not self-implementing. How they are implemented depends on the type of data, the record keeper, the purpose of processing, the manner in which data is to be used and disclosed, the costs, the technology, and the traditions of the jurisdiction or record keeper. There are often several strategies for implementing the same principle. What is most important is that any privacy policy should consider and address in an appropriate way all elements of FIPs. Some FIPs principles also reflect good record management policies.

The prevalence of identity theft illustrates why any sharing of personal information can be a threat to an individual. There is already some evidence that concerns about privacy affect voter behavior: one survey found that 23 percent of California non-voters

---

[21] 5 U.S.C. § 552a (2002).

[22] Fair Information Practices were invented in America. See Secretary's Advisory Committee on Automated Personal Data Systems (Department of Health, Education & Welfare), 1973, *Records, Computers, and the Rights of Citizens*, available online at http://aspe.os.dhhs.gov/datacncl/1973privacy/tocprefacemembers.htm.

[23] The leading international statement of FIPs is by the Organisation for Economic Cooperation and Development. See *Council Recommendations Concerning Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data* (1980), available online at http://www.oecd.org/document/18/0,2340,en_2649_34255_1815186_1_1_1_1,00.html.

say they haven't registered to vote because they want their information to remain private.[24] If voter records are perceived to be a source of data that contributes to the widespread trafficking in personal information and to identity theft, some potential voters may be discouraged from registering and voting. Larger or centralized databases may exacerbate these concerns. Further, any inadvertent or malicious release of data can affect millions of people and will attract considerable publicity. The move to statewide VRDs raises the privacy stakes considerably.

Privacy values, which too often are an afterthought for collections of personal information, are fundamental for voter registration. For this reason, some privacy issues are intertwined with basic design standards and do not need to be addressed separately. This chapter addresses only those privacy policy matters of openness, data collection limitation, and use and disclosure limits, which are not otherwise considered in this report.

Openness (Transparency). Policies and practices for the collection, maintenance, use, and disclosure of voter registration databases should be transparent, published, and available to all upon request.

Implementation Strategies

- Publish on the main election board website a complete notice of policies and practices describing the collection, maintenance, use, and disclosure of voter registration data. The notice should include contact information for the office or the officials responsible for the voter registration data.
- Publish a readable summary notice of policies and practices in other places such as on voter registration forms, at polling places, on sample ballots, and elsewhere as appropriate.
- Provide a copy of the complete notice to any person who requests it.
- Publish any changes to the notice before the changes become effective, and accept and consider public comments.
- Place a date and version number on notices as they are published. Maintain, and make publicly available, copies of all previous notices, including the periods of time during which they were effective.

Discussion. A notice of policies and practices for the collection, maintenance, use, and disclosure of personal information informs registrants, the public, and interested parties of the relevant policies. It also informs the staff of the election agency about the policies and the need to conform to those policies. Finally, clear notice imposes a discipline on agencies helping prevent them from making ad hoc choices about their data processing activities. By requiring that these activities be properly disclosed in advance, privacy policies prevent agencies from undertaking new data gathering or disclosures without going through a formal process, thereby helping agencies resist pressures to use personal information in new ways without sufficient oversight.

[24] California Voter Foundation, 2005, "California Voter Participation Survey," available online at http://www.calvoter.org.

Formal privacy notices, like other legal notices, are often necessarily long and detailed – likely longer and more complex than an average voter will care to read. Consequently, we recommend that a summary notice that is more accessible to the average voter and brief enough to fit on commonly distributed forms be made available.

Collection Limitation. The following principles should apply to the collection of personal data.

- All data should be collected by lawful and fair means.
- Data should be collected, where appropriate, with the knowledge or consent of the subject.
- Registrants and the public should be informed through the published notice of the policies and practices of all the sources of data obtained for voter registration purposes.
- Data collection should be limited to sources and procedures authorized by law and properly described in the published notice.
- Only the minimum information necessary for and relevant to voter registration purposes should be collected and maintained. The reason for collecting each type of personal information should be explained, and the specific data elements collected should be subject to public scrutiny.

Discussion. There are several reasons why the public should be informed in advance of the collection and the source of the data, when such data about voters is obtained from third party sources for purposes of updating, correcting, verifying, or amending the database. First, the public should know what data sources are being used so that it can assess the validity and utility of the data sources. Second, public disclosure may uncover errors (e.g., use of inappropriate or outdated sources). Third, the election agency will be required to justify its choices, thereby reducing the chance that unnecessary data will be collected. For example, there is no reason for voter registration records to reflect religious or sexual preference.

The collection of data with the knowledge or consent of the individual will be accomplished in most instances through the published notice of policies and practices. The public identification of data sources normally will be sufficient to meet the knowledge standard. However, if data on a specific individual is being collected as part of an examination of that individual's eligibility, it may be appropriate to inform the individual, seek consent for the collection, ask for cooperation in the examination, and provide due process rights before taking any action that affects the individual. It will not always be possible to satisfy the consent standard. For example, if voter registration records are examined in an investigation of voting irregularities, notice or consent to the subject of the investigation may be inappropriate. However, in other circumstances, the data subject may be the best source of information.

Use and Disclosure Limits (Purpose Specification and Use Limitation). There should be limits to the uses and disclosures of voter registration data by the agencies that collect and maintain the data. Personal data should not be shared with anyone outside the election process without legal authority or the consent of the data subject. Within the

election process, use of personal information should be limited to those officials who have a need for the information in the performance of their duties. All uses and disclosures should be specified at the time of collection.

Implementation Strategies

- Limit use and disclosure of voter registration data to those activities directly related to the election process or expressly authorized by law.
- Describe all uses and disclosures in the published notice of information practices.
- Identify publicly all recipients of voter registration data.
- Provide public notice of and, if possible, a chance for public comment on all disclosures of identifiable voter registration data for any activity not directly required for voter registration purposes.
- Restrict access to specific records, specific data elements, and specific classes of voters (e.g., by location) to those election officials who have a need to use those records, data elements, and classes in the performance of their duties.
- For some or all uses or disclosures, maintain a record of the date, nature, and recipient of all personal information and make the record accessible to the data subject upon request.
- Restrict disclosures to specific data elements permitted by law and necessary to accomplish the purpose of the disclosure. Withhold data elements that are not essential to accomplish the purpose of the disclosure or that would place data subjects in excessive jeopardy to identity theft or other improper activities.
- Prevent all recipients of data from using or disclosing the data in ways not specifically authorized by law. Asking recipients to sign data use agreements is one way to accomplish this.
- Allow some non-essential uses and disclosures only with the affirmative (opt-in) or negative consent (opt-out) of the data subject.
- Limit disclosures to the greatest extent possible for data subjects at risk (e.g., victims of spousal abuse, jurors, some public officials).
- Even the best use and disclosure policies may be violated by people and software within the election process. Therefore, limit access by each person and each system component. Chapter 5 on security provides further discussion on access policies.
- Provide access for every voter to a personalized list of those third parties who have been given or purchased access to his or her voter registration data.

Discussion. *Use* refers to the utilization of data internal to the operation of the election agency. *Disclosure* refers to any sharing of data with an external party. Controlling both use and disclosure is essential to maintain proper control over data and to prevent the data from being used in inappropriate ways. Controlling use and disclosure through formal procedures and public notice will help limit function creep, which is the use of data for a purpose unrelated to the purpose for which it was originally collected.

Whenever possible, use or disclosure should include only those data elements that are necessary for the required purpose. Limits on the disclosure of some data elements are *constitutionally* required. The principle is illustrated by a successful challenge in 1993 to the disclosure of Social Security numbers from Virginia voter registration records

(*Greidinger v. Davis*[25]). The plaintiff challenged the public disclosure requirement of Social Security numbers as an unconstitutional burden on the right to vote. The plaintiff argued that the privacy interest in the number is sufficiently strong that the right to vote cannot be predicated on disclosure of the number to the public or to political entities. The Fourth Circuit Court of Appeals agreed. Following the decision, Virginia changed its law. The importance of restricting disclosure of some or all data elements has only been highlighted by the epidemic of identity theft in recent years. The Greidinger decision was issued in 1993, well before identity theft had become a common crime and concern.

The Greidinger decision also highlights the sensitivity of the use of any identification number as part of the voter registration process. While Congress mandated that a registration application include a driver's license number or the last four digits of the Social Security number, excessive reliance on numbers for identity verification in voter registration may not be successful, may create new risks to data subjects, and may expand pressures for the use of identification numbers or identification cards in other contexts.

While some secondary uses and disclosures of voter data are authorized by law, the political process may impose limitations on what secondary activities are to be permitted. Some states make voter registration records public, while other states strictly restrict secondary activities. There are no clear right or wrong choices, but privacy standards argue for limiting secondary data sharing to the greatest extent possible.

Middle ground may be helpful at times. While registration records will appropriately be used for voting purposes, it is possible to offer each individual a choice with respect to some secondary uses or disclosures. The federal Driver's Privacy Protection Act[26] provides a model. It describes a series of activities for which use and disclosure is permitted without the consent of the data subject. For other activities, the affirmative consent of the data subject is a requirement.

Many methods can be used to give voters a choice about how their data will be handled. Under an affirmative consent (opt-in) model, personal information can be used or disclosed for particular purposes only if the data subject agrees. Consent can be obtained orally or in writing. Under a negative consent (opt-out) model, a use or disclosure is permissible unless the data subject has stated an objection. Individuals can be offered choices through check boxes on applications or websites or in other ways. Sometimes it may be possible to ask each individual to make a choice about a use or disclosure without establishing a default option. For example, on a website, an individual can be required to make a selection before moving on to the next screen.

The value of individual choice is that it gives the individual a voice in how his or her records may be used for purposes that are not directly related to the purpose for which information was originally obtained. It is a way to resolve conflicts between data subjects who desire privacy and officials and others who seek to use information in new ways. It is a middle ground between saying that records are never available and that records are freely available. The preference of the individual is a reasonable and significant factor to consider when making decisions. With computerized information

[25] 988 F.2d 1344 (4th Cir. 1993).
[26] 18 U.S.C. § 2721 et seq. (2002).

systems, it is easier as well as practical to keep track of individual choices and to abide by those choices.

4. Usability

VRDs will be used by voters, election workers, and authorized officials from state and local governments to perform crucial tasks. As problems in data entry and interpretation can easily disenfranchise voters, user interfaces for these systems must be designed to minimize opportunities for error.

User interfaces that provide users with inadequate and unclear feedback can lead to the entry of inaccurate data. Displays that fail to provide indicators of system state (e.g., the name of the currently authorized user) can introduce opportunities for malicious users. Data displays that include identifying information beyond the minimum level needed to complete a task might compromise voter privacy. Poorly designed voter registration database user interfaces might confuse users and reduce confidence in the system, thus creating a perception, if not a reality, of reduced reliability.

General Usability. User interfaces for voter registration database systems should be designed to help all users complete their tasks confidently and correctly. The design, development, and testing process should explicitly account for wide ranges in user training, backgrounds, and physical abilities, as well as the physical environments in which these user interfaces will be used.

VRDs will be used by voters, election workers, and other authorized officials to accomplish numerous tasks, including (but not limited to) registering voters, updating registration information, verifying eligibility for a given election, and extracting summary reports. Each of these tasks involves one or more user interfaces that bridge the gap between user tasks and the underlying database.

The range of possible users and uses make user interface design particularly challenging. Although some people—for example, county and state elections staff—are likely to be frequent users who receive detailed training, many others—namely polling place volunteers and voters—will use these systems infrequently, possibly without any training at all. Large variations in background, literacy, computer experience, and physical capabilities (including disabilities) throughout the general voting population complicate matters further. User interfaces should be designed to be easily usable by a wide variety of users in a variety of challenging environments employing strategies such as providing text in multiple languages and providing alternative input and output methods for people with disabilities.

The environments in which these systems will be used present additional challenges. Unlike systems that are only used in one well defined work context such as an office, VRDs might be used in many places, including municipal offices, polling places, and in homes or libraries via the Internet. These differing use contexts require different user interfaces.

The computing environment may also influence usability. Computing platforms for VRDs may have relatively minimal requirements for processor performance, network bandwidth, memory, and display capabilities. However, user interfaces that seem to be functional when a system is not stressed can encounter usability difficulties when there is a high system load, network congestion, or other demanding situations. These issues are discussed in more detail in Chapter 6 on reliability.

Human-computer interaction professionals know that simply adding an interface to an

already-designed system does not work well. Interface design development, documentation, and training materials should be addressed at the beginning of a project and throughout its course of development and implementation. While an early focus on interface design and testing allows more time for refinement, user interface evaluation can provide useful information at almost any stage in the software development process.

The needs of the wide range of likely users should be evaluated during the interface design process. Although it is clear that there will be many different types of users for VRDs, not all types of users initially can be defined or identified completely. Input from classes of many potential users including voters, public officials, poll workers, and others can help clarify user needs. Serious consideration for user concerns also can have the added benefit of building good will toward the project.

Before any user interfaces are designed, techniques such as interviews, group discussions, and observations of users completing typical tasks with existing systems (computerized or paper-based) can be employed to gather usability requirements and help developers understand the contexts of use. Such activities also will help developers understand the difference between classes of users and how those differences will impact user interface design.

Usability requirements can act as a starting point for an iterative cycle of design and feedback. Initially, simple mockups of proposed layouts will stimulate more input from users and further clarify usability requirements. In addition to being inexpensive to produce, paper prototypes and other informal presentations of design proposals can make some users feel freer to make critical comments than if they are presented with an almost-finished version. Feedback can be used to inform subsequent, more fully-realized designs, with further iterations eventually leading to convergence on acceptable designs.

Structured evaluations can be useful for identifying specific usability issues that may not arise in discussions with users. Usability experts can examine user interfaces for consistency, proper feedback, error handling, and other criteria. Known as heuristic evaluation, this technique is often very effective after just a few evaluations. Direct observation of potential users attempting typical tasks with proposed designs can also be very helpful. In so-called "think aloud" sessions, users are asked to tell observers what they are doing and why. This feedback helps developers identify potentially confusing or disorienting aspects of a proposed design. If multiple alternative designs are being considered, a user study involving measurement of user performance (in terms of task completion time, accuracy, or other objective measures) on meaningful tasks can help clarify the strengths and weaknesses of the alternatives.

These measures may seem excessive to some, but frequent, early evaluations increase the chances of finding problems with interface designs and other system features before fixes become prohibitively costly.

The process of evaluating and refining user interfaces should continue after the systems have been deployed. As various users—including many who were not involved in the design discussions and evaluations—work with the system, usability difficulties and challenges will likely be identified. Developers should assume that ongoing feedback will lead to further user interface revisions.

Although specific user interfaces will vary from state to state, all of the VRDs will face similar usability problems. Mechanisms for sharing insights gained during user interface design and evaluation (while still respecting proprietary designs) can help

improve overall usability.

Usability in the Service of Accuracy, Security, Reliability, and Privacy. All user interfaces should be explicitly designed to support the goal of building VRDs that are accurate, secure, reliable, and sensitive to voter privacy concerns.

To be successful, user interface specification and design must be an integral part of the software development process. As mechanisms for ensuring accuracy, reliability, security, and privacy sensitivity are developed, their impact on user interactions should be carefully considered and user interfaces designed accordingly.

Clear and useful feedback regarding the state of the system and the impact of user actions is a crucial component of successful user interface design. Such feedback can play a role in guaranteeing system security and privacy sensitivity. For example, user interfaces used by polling workers or county officials might display a photo of the currently logged in user at all times, allowing onlookers to verify that a task is being performed by the appropriate person. Prominent displays of system date and time can show both users and (when appropriate) voters that the systems are configured correctly. Status alerts listing active network connections, along with indications of any that involve unknown hosts, can be used to identify possible intrusion attempts. Dialog boxes and other alerts that warn users of the potentially undesirable outcomes of their action should be displayed. Well-designed displays of summaries regarding accesses to the system and changes to voter records can help managers ensure that the system is functioning reliably and securely.

User interfaces for VRDs should be minimal, containing only displays and functionality that are necessary for the completion of specific tasks. Because displays of personal information create risks for invasion of voter privacy, these displays should only contain information that is necessary for the task at hand. For example, if the last four digits of the Social Security number are used to verify identity, displays should contain only those four digits, not the full number.

When extraneous functionality is removed from an interface, opportunities for malicious hacking, data theft, or entry of inaccurate data are also removed. For example, hardcopy printouts of voter registration data might contain unnecessary information that violates voter privacy. Proper privacy protection would mandate protecting and destroying the printouts. As modern printers generally receive data over a network connection, hardcopy print facilities also have the potential to introduce security vulnerabilities. Limiting print functionality to cases where it is absolutely necessary can reduce these privacy and security risks.

Eliminating extraneous user interface components can have other benefits as well. Simple user interfaces are often less cluttered and therefore easier to use, particularly for novice users. Decreasing the complexity of the interface also can simplify the underlying implementation, potentially reducing development costs.

Usability considerations must factor in tradeoffs as well. Supporting privacy, accuracy, security, and reliability can sometimes reduce usability as can happen with security measures that are explicitly designed to make systems unusable by unauthorized users. For example, systems that are used in public places might have forced logouts after very short idle times to prevent unattended workstations from becoming inviting targets, even though this will, in some cases, result in annoyance for the authorized users.

Appropriate evaluations and user tests might identify aspects of interface design that could negatively impact other design goals. Each display and control can be evaluated to determine if it might introduce potential problems or if it simply can be removed.

Testing user interfaces under extreme or suboptimal conditions can provide insight into the interplay between user interfaces and reliability. Systems that simply freeze or lock-up under extreme operating conditions are neither usable nor reliable. Wherever possible, systems should respond gracefully to stressful conditions, provide users with appropriate feedback, and degrade to reduced functionality if some services are unavailable.

Usability for Election Staff and Government Workers. Because errors in data entry, retrieval, and interpretation by election workers and government officials can lead to voter disenfranchisement, the VRDs should be designed to maximize the usability for election officials while reducing these common problems. The challenge of constructing user interfaces to minimize these errors is complicated by the nature of the user population. County election officials and other municipal employees regularly use the voter registration system. These users can be provided with training that would enable them to effectively use a reasonably complex system. Volunteer election officials, on the other hand, might use the system infrequently (perhaps one day per year) with minimal training. These users might also be relatively unfamiliar with some election jargon.

Polling places are often crowded, busy, and noisy on Election Day. Noise, interruptions, and other distractions can increase cognitive load on users, potentially leading to an increased error rate. Any election technology user interfaces that will be used during polling should account for Election Day stresses.

Known user interface design techniques can reduce the frequency of errors in data input, retrieval, and interpretation. Data input forms should be designed with layouts that clearly indicate the meaning of each field. When possible, data provided on these forms should be immediately validated for accuracy and consistency. Error messages should be as clear as possible, providing information that can help users respond appropriately, for example, by correcting the input or by accessing external resources, such as documentation or personnel, to clarify any confusion. However, as mentioned in the previous section, messages should avoid disclosing unnecessary information.

Modifications to voter record fields such as address or party affiliation can change a voter's precinct or render the voter ineligible to vote in some primaries or for certain offices. Functionality that might change the ability of one or more citizens to vote should be available only to authorized users, but access controls are only the first step in preventing harmful changes. Exactly as desktop operating systems require users to confirm potentially damaging actions before they are executed ("Are you sure you want to delete this file?"), user interfaces for VRDs should require explicit confirmation from the user before making any changes that would restrict or modify an individual's ability to vote. This confirmation might come in the form of a dialog box, or by requiring that a certain check box be selected. For changes that have wider impact, particularly batch updates, displays should indicate the number of affected records. Confirmation for these changes should make users think twice before making significant changes. Possible approaches include multiple, sequential requests for confirmation, request for reauthentication via retyping of the user name and password, or requiring that users type

a word embedded in an image (a so-called CAPTCHA™ test, commonly used for registration on web sites). Larger batch updates should require confirmation by the current user and a colleague who confirms the action separately. Where possible, undo facilities should be provided.

User interfaces for specifying data retrieval parameters are similar to data entry forms: users must specify one or more values for each of several fields. Once data has been retrieved, it should be presented clearly on screens that indicate both the values of the specified parameters and the fields that match those parameters. Such a presentation will help users distinguish between input errors and result interpretation errors. Important fields such as registration status should be highlighted. Detailed feedback, including appropriate contextual information and links to relevant rules and policies, should be provided especially on problems and policies that might disenfranchise voters. To minimize the risk of infringing upon voter privacy, all displays of personally identifiable information should be limited to include only details that are necessary for the task at hand.

Different users might require different user interfaces and training materials. An interface for election officials might provide information that is more detailed and use specialized language that would be inappropriate for election volunteers. Infrequent, less well-trained users might benefit from training sessions, online tutorials, and online help. The context of use is also an important factor in interface design. While audio indication of input errors may be fine for office workers, noisy conditions in polling places might render such output useless.

Interface designs should be tested thoroughly, with representative users performing typical tasks under situations that simulate as closely as possible those of real use. These challenging tests may identify usability problems that might not have been found during testing under idealized conditions.

Usability for Voters. Usable interfaces for individual voters have the potential to educate voters, provide necessary information, and build confidence in the election process. Voters who are unable to perform voting tasks effectively might require help from election officials. If assistance is not available, a voter might simply walk away, effectively disenfranchised by bad design.

The deployment platform for voter user interfaces is an important concern. Systems for use by election officials and workers are likely to be dedicated, stand-alone packages with completely custom user interfaces. As the deployment of custom software to individual voters is not practical, voters are likely to use web browsers to access registration information. Although the use of standard browsers offers many advantages, including the ease of linking to relevant contextual data, browsers can be somewhat limited in the types of feedback that they can display.

The use of web browsers for general public-access user interfaces also presents testing challenges. These systems need to work with many different hardware and software configurations. Such systems need to have their performance verified on many web browsers, including multiple older (and pre-release) versions of popular browsers and screen-reader systems for people with visual impairments. Designing to generic, vendor-neutral standards is one way of achieving maximum portability; conversely, using one vendor's proprietary extensions is an almost certain way to restrict portability and access

by the full public.

Web-based user interfaces should be designed to maximize privacy and security. Retrieval of information about polling places and election policy should be based on a minimal specification: if the street address is sufficient for identifying a polling place, the voter's name should not be requested.

Conclusion. The importance of ease of use with VRDs cannot be overemphasized. User-friendly interfaces are essential if the systems are to be effective and credible.

5. Security

This chapter examines the security mechanisms that enforce the decisions made about who may read or update VRDs. It also addresses ways of protecting against malicious actions by both insiders and outsiders.

VRDs need to control who may access different kinds of information stored in the VRD and under what circumstances they are authorized to do so. Accordingly, the first part of this chapter discusses access controls. Careful control over who is allowed to read or update the VRD reduces the possibility of intentional abuses and unintentional mistakes by authorized users of the system.

The right to view or modify some portion of the VRD is called an access privilege. The list of rules specifying who has which access privileges is called an access control policy. We examine the following aspects of data access:

- deciding who should specify which parts of the access control policy;
- determining who should have which access privileges;
- enforcing access control policies; and
- authenticating that people are who they say they are so the system can identify what access privileges each user should receive when the system is in use.

Generally speaking, four broad classes of access privileges are commonly found in any database system:

- **Read privileges.** The authority to view, inspect, read, print, or otherwise access certain records without modifying them in any way.
- **Write privileges.** The authority to modify, update, add, or delete certain records.
- **Administrative privileges.** The authority to specify what privileges are made available to other users. This includes the ability to create new user accounts, to assign user accounts to specific employees, to specify or change the privileges available to users, and to delete users. In some systems, this category might also include related privileges such as the authority to modify or patch software, the database schema, and other administrative functions.
- **Execution privileges.** Operations that the user is allowed to perform. Execution privileges are often enforced by another system component called the application server.

Access control policies should minimize the number of people who receive privileges either to access each piece of information or to grant access to others. They should also ensure that each person is granted only the minimal set of privileges needed to do his or her job. Following these guidelines can provide significant protection.

The second part of this chapter discusses how to harden a VRD against attack. If a VRD is not secured adequately, technical attacks by insiders who have access privileges or by outsiders via hacking may undermine the VRD—for example, by inserting the names of ineligible voters into the VRD or by removing names of eligible voters from the VRD. Since there are many ways that an attacker might try to subvert the system, one needs processes that encourage secure system design and detect and close significant

vulnerabilities in the deployed system. We discuss the following:

- providing security against technical attacks and other attempts to subvert the system (system security); and
- dealing with security failures should they occur.

Dividing the Responsibility of Choosing an Access Control Policy. Access control policies provide an automated way for state and local officials to implement the accuracy and privacy policies discussed in Chapters 2 and 3, respectively. Access control can help ensure that only authorized users are allowed to make authorized transactions. Establishing access control policies will likely require the cooperation of state election officials and election officials from each local jurisdiction. For example, state officials might not have detailed knowledge of the staff and their responsibilities in each jurisdiction; county officials are more likely know which county employees should receive which kind of access. However, county officials are unlikely to be able to set statewide policy. Therefore, we believe it will be productive if all relevant offices work together in setting VRD access control policy. We discuss some of the options for structuring this process.

One possibility is a partially centralized model. State officials might identify certain common job roles, suggest a reasonable set of access privileges for each role, and perhaps even require that local registrars adopt these roles and privilege sets. For example, roles might include (1) data entry clerk (who receives access privileges that permit the creation of new records and editing of existing records subject to approval by other officials), (2) election judge (who approves modifications to voter records for all voters within the judge's jurisdiction), or (3) registrar of voters (who receives access privileges that allow him or her to create accounts for new users, assign these users to roles, and change the role assignments for existing users). To allow for local autonomy, localities might be allowed to modify the roles and their privileges.

Alternatively, the partitioning of access privileges could be decentralized and left up to county election officials, leaving the state officials with only tasks such as the following:

- Specifying the access rights that officials in one jurisdiction have to data belonging to others. This policy could be rigid, or subject to revision by the jurisdictions involved.
- Managing a list of job roles and purposes, so that people in different jurisdictions all use the same terminology. In other words, in situations where practices are the same, make the vocabulary the same.
- Specifying (or recommending) maximum privileges that can be granted to each job role and purpose. A jurisdiction would be free to specify narrower privileges, if the jurisdiction's officials felt this was appropriate for their setting.

It is likely that even a centralized scheme will require some aspects of authorization to be decentralized. For example, the roles of authorized users are more suitably managed locally, such as by a county registrar, than from afar by, say, the Secretary of State. In many cases a local registrar knows who local users are and thus is much less likely to be deceived by an impersonator.

There is an opportunity for the EAC, or some other nationwide organization, to

provide sample roles and levels of privilege as suggestions to states and local jurisdictions, leading to a more uniform vocabulary and starting point for states.

Some composite actions might require privileges from more than one of the four categories of access privileges (read, write, administrative, and execution). For example, moving a voter from one jurisdiction's voter rolls to another's might require both write privilege (to delete the voter from the former jurisdiction's voter rolls) and read privilege (to obtain the information needed to add the voter to the new jurisdiction's rolls). Normally, a user should be permitted to take a composite action only if the user has all relevant access privileges. Alternatively, such situations can be handled by access rules that state who may execute the action. The rules can be specified by an authorized user who administers all the necessary underlying privileges and enforced in either a DBMS or an application server.

The process used to assign categories of access privileges need not be the same. For example, it would be possible to assign administrative privileges via a semi-centralized model, yet assign read and write privileges in a decentralized fashion. One could also separate administration of felony status from administration of addresses.

Determination of the access control policy does not need to be tied to details of how data is physically distributed. Access control policy might be determined in a centralized or decentralized fashion regardless of whether the VRD data is stored at a centralized location or is physically distributed.

Assigning Access Privileges. The access control policy's scope should include all types of access to the VRD including records on both voters and non-voters, database schema, and so forth, and the VRD should be designed so that such a policy can be enforced. To reduce the overhead of administering privileges, we recommend the approach of grouping people by their roles. Most DBMSs and application servers support this approach. One might define groups of people, groups of data, groups of actions, and specify rules for whole groups. Election officials should specify very detailed rules on who can access what.

It is advisable not to grant all users the same access privileges. Instead of thinking in terms of access to whole databases (e.g., the list of eligible voters or the database from which the eligible voter list was derived), officials should determine specific access rights for each user or group, limiting each user to appropriate data fields, subsets of voter records, and purposes, as well as appropriate access modes (e.g., read, modify, delete, create). One can specify privileges for individual fields of all voters' records (e.g., authority to modify party affiliation and preferred contact method but not the mailing address). One also can specify access privileges for sets of voter records (e.g., authority to modify any part of the voter record for voters in Boston). Separately, one can specify access privileges in terms of groups of people (e.g., all data entry clerks receive the same set of access privileges) or in terms of individual employees (e.g., a privilege granted only to Alice Jones).

The basic principle underlying a sound access control policy is to minimize the number of people who have routine access (read or write) to each data item, and to minimize the amount of data that each person has access to. The rule of thumb is to give each user of the system the minimum amount of access privileges he or she legitimately needs to get the job done and nothing more. This is often known as the *Principle of Least*

Privilege.

A related guideline is that users' tasks should be structured to minimize the amount of access they need and to minimize the number of people allowed to access information. For example, processes should be organized so that poll workers do not require routine access to voters' Social Security numbers or criminal conviction information.

The Principle of Least Privilege helps reduce the likely impact of security failures and abuse should they occur. For example, if some user's password is discovered by a hacker, then the hacker might gain access to everything to which the user has access. In this case, the damage will be far less if the user has only a limited degree of access to the system. By comparison, if every user receives full privileges to read and write every voter record within the state, then penetration of a single user account could lead to almost unlimited harm to the VRD. The Principle of Least Privilege also helps reduce the likelihood of insider abuse of privileges.

A user's access rights should usually depend on his or her role, location, current purpose, and so forth:

- User access privileges should be limited by jurisdiction. Election officials normally should not be granted privilege to read or modify records for voters registered outside of their jurisdiction. For example, San Diego election officials would normally not need to read or modify the records of a San Francisco voter, so they should not be given access privileges that would let them do so. As a special case for voters who move, a San Diego election official might be permitted to read the record of a San Francisco voter when performing a transfer transaction that moves the voter to San Diego County. Initiating such a transfer also might require approval by a San Francisco election official.
- An employee who processes registration forms might be allowed only to change a voter's driver's license or phone number, while an official responsible for determining eligibility might be allowed only to update whether or not a voter is eligible.
- Access might also be limited by field. For example, on Election Day poll workers need read access to some information from the voter rolls (including voter names, addresses, and party affiliations for some elections) to check voter eligibility at the polls. However, poll workers normally would not be granted any access to other fields of the voter record because such access is not needed to perform their jobs and because poll workers are not vetted as carefully as other users of the system. The access control policy should codify such privileges and restrictions.

Administrative privileges should be particularly restrictive; very few users should have the ability to grant access to others. Privileges also might be limited to account for organizational relationships. In certain circumstances, preventing municipal employees from increasing the access levels of their supervisors might remove the possibility of conflicts between database access policies and manager-employee relationships. Similarly, users with administrative privileges should never be allowed to grant themselves new access privileges; requiring the consent of another administrative user increases accountability.

Use of software that extracts and prints voter information, including the creation of

DVD-ROMs for political parties or poll workers, should be governed by the privileges of the ultimate recipient. In other words, documents or DVD-ROMs should contain only data that all of the recipients are allowed to view, even if the creators of the documents have additional privileges.

It is likely that access control policies will need to be updated periodically. As with privacy policies, older versions of access control policies should be retained, along with their dates of applicability. Furthermore, officials may wish to consider making their access control policies public in some form in the interests of transparency and to make the chain of responsibility clear.

We recommend that those responsible for managing VRDs attempt to measure how effectively they have limited privilege by characterizing how many people have access to how much data and by tracking progress over time using these metrics. For example, one might count for each voter record how many people have some kind of access privilege to at least part of this record and compute the average of this across all voter records. More refined metrics might reflect access to only some of the fields (e.g., affiliation but not full SSN). One might perform separate analyses for read access (to assess privacy risk) and write access (as a risk to accuracy). We stress that we mention these metrics only as examples of what is possible.

The EAC, or other nationwide voting administration organizations, could play a helpful role in coordinating an effort to develop suitable metrics. Ideally, such metrics would be published by each state, enabling independent analysts to evaluate each state's effectiveness at setting access control policies and facilitating comparisons of practices among states in a meaningful way.

Adding election workers to the system in an appropriate fashion is a crucial step in the operation of a VRD. It does no good to have restricted access rights if a corrupt official can add new personnel with arbitrary access privileges. There are two complementary solutions: public logs of all changes to the list of authorized parties including their access rights, and a dual signature requirement for any changes to the list (also known as two-person control). Both should be adopted for most users of the system. An exception might be made in the case of poll workers with very limited read access to the system (e.g., ability to view redacted records of only voters within their precinct) and no write access. In this case, approval by a single full-time election official might replace the dual-signature requirement.

Access During Emergencies. Provisions also need to be in place for handling emergencies. Officials should create rules that allow trusted election officials to temporarily increase privileges available to others. This might be achieved by creating rules that enable additional privileges under emergency conditions, together with a separate mechanism to declare to the system that an emergency exists. Emergency overrides should be tightly controlled, for example by two-person authorization, generation of detailed audit logs regarding such events, notification of the person whose privileges are delegated, and periodic proactive inspection of such audit logs. No single user should be permitted to declare an emergency and elevate his or her access privileges during the declared emergency; instead, exercise of an emergency override should require the active cooperation of at least two people.

Recognizing that people will occasionally be absent or overloaded with work, it will

sometimes be necessary to grant one employee some privileges belonging to another. If access control policies are based on roles, this can be done by temporarily assigning a new role to the appropriate individual. In any case, this should be done without revealing either employee's password to the other employee. Emergency or unanticipated delegation of access privileges should be temporary, preferably with automated procedures to remove the extra authority.

Enforcing Access Policies. DMBS and application server security provide several mechanisms for specifying and enforcing policies with the goal of keeping administration manageable. First, DBMSs provide mechanisms for describing the set of users. One can

- Define groups and assign users to them. Groups, rather than individuals, then become the basic unit of authorization. Similarly, one may define a role to represent a specific set of privileges (e.g., those associated with a job description).
- Give users additional descriptive properties that may be used for decisions. For example, officials might be associated with a list of zip codes for which they are responsible.

Second, DBMSs provide the means for assigning privileges to users and enforcing the access control policy. One can

- Grant a privilege for a group to access a field or specified fields of the database (e.g., encoding a policy that states that this user is permitted to view the voter's address but not the voter's full SSN).
- Grant access to a view that filters or summarizes the data but hides many details. Some views might filter by locality, while others might provide statistical summaries that are widely releasable.
- Grant access in which some items in a database are automatically filtered out based on the current user or task.

Application servers offer some of these capabilities, together with privileges to execute programs that implement business functions larger than a single DBMS request.

The VRDs should use access control mechanisms provided in the DBMS; trying to implement access control entirely at the application level leaves greater opportunity for security mechanisms to be bypassed or compromised. There should be no way for users to bypass the access control mechanisms. For each user request, either the application server policy must approve the entire operation or the DBMS must enforce access controls on each data access or both. This requires examining the user's individual credential and the privileges associated with his or her job. Implementing an access control rule in the DBMS guarantees that the rule applies to *all* operations that developers create.

Authentication: Verifying Identity. In any system with restricted access rights, authentication is crucial. The system needs a way for people to prove who they are; from this, their access rights must be determined and enforced.

Authentication can be done in many different ways. The most common form of

authentication is by user name and password. While superficially attractive, password authentication is subject to many failure modes including password guessing, inappropriate sharing of passwords, and inadvertent or deliberate password leakage.

Authentication schemes based on physical devices can be considerably more secure. Systems based on smart cards or timer-based tokens require the presentation of an appropriately encoded electronic device (possibly within a defined time period) for authentication. Biometrics such as fingerprints or eye scans may also provide greater security than simple textual passwords.

The potential advantages of these alternative authentication techniques may be offset by increases in cost and complexity. Lost smart cards are likely to be more expensive to replace than lost passwords. Biometrics systems may have difficulties in enrollment: difficulties in the initial capture of the finger, eye, or voiceprint may cause later problems with authentication.

Security breaches in authentication mechanisms might be exploited to achieve unfettered access to the underlying systems. To avoid this scenario, authentication mechanisms should be carefully designed and tested. Authentication servers must be highly secured, both physically and technically, and appropriate cryptographic techniques should be used. VRDs should not utilize any authentication techniques that have not been validated by extensive use in production environments.

Biometric systems are especially tricky, because many current deployments have been implemented improperly. The use of fingerprints, retinal scans, facial features, and other biometrics all rely on the conversion of these characteristics into strings of bits that can be stored and processed by computers. If these digitized versions of the biometrics are transmitted across networks or stored on multiple computers, security weaknesses in the networks or remote computers might be exploited to capture the biometrics. A malicious attacker who captures digitized biometrics might be able to use them to gain access to the system. In addition to reducing the security of the VRD, such attacks might compromise the use of the specific biometric by the affected users in any other domain. As a result, biometric data should be stored as close to the user as possible, perhaps used only to unlock a smart card. In this scenario, the user's fingerprint, for example, might be used to verify that she is the authorized user of a smart card that would then be used to access the VRD. As the biometric data would be stored only on the smart card (which is generally under the physical control of the authorized user), there are no network connections or remote hosts to tempt malicious intruders.

Biometrics also should be used only in a supervised setting to foil various forms of spoofing attack. There have been many reports of successful attacks on unsupervised biometric authentication. For example, with some facial recognition systems, holding up a glossy photograph of an authorized user to the camera is sufficient to fool the system. There also have been published reports stating it is possible fool a fingerprint recognition system by lifting the fingerprints of an authorized user off of a surface touched by that person and creating a fake "gummy finger" made out of gelatin that bears the authorized users' fingerprint.[27]

Different authentication schemes might be appropriate for different users or different

[27] Tsutomu Matsumoto, 2002, "Gummy and Conductive Silicone Rubber Fingers: Importance of Vulnerability Analysis," pp. 574-575 in *Advances in Cryptology - ASIACRYPT 2002,* Lecture Notes in Computer Science, Vol. 2501.

tasks. The type of authentication being used should be determined by the type of task that the user is performing, the expense and complexity of the authentication scheme, and the potential harm that may be caused if the authentication system is breached. Advanced authentication schemes are more appropriate for election workers and government officials with access to greater privileges over a wide range of voter records. In these cases, multi-factor authentication (such as requiring both a biometric and password) may be warranted, despite its higher costs or inconvenience.

Another style of authentication relies on a technology known as *certificates*. Apart from authenticating the user, certificates allow for operation in the absence of access to a permission database. A certificate can contain a user's access rights in a form that is mathematically protected from change. When a certificate is presented to a system, that system can enforce the user's access rights using only the data presented. Because certificates are too long to be memorized or typed, they frequently are stored on smart cards.

Once a user has been authenticated to the system, each operation on the database should check that the person's privileges allow him or her to perform that operation. Similarly, the database should create an audit trail for all requests that modify the database. Both of these goals are straightforward to achieve. Logging read operations may be feasible and useful though careful engineering is needed to ensure that the logging system can handle the data volume. As previously discussed, to guarantee that these access controls cannot be bypassed, access control restrictions should be implemented in the database itself, where possible.

The importance of security training cannot be overstated. Authorized users of the system must be taught about protection of passwords, how to resist social engineering attacks—attempts to deceive someone into performing certain actions—and the importance of never sharing their passwords, even with their colleagues and other authorized users. Training should include how to cope with failure scenarios such as how to proceed when normal authentication mechanisms are for some reason not functioning. Because procedures that seem arbitrary are often ignored, users should also learn how and why failure to follow procedures could lead to security breaches. Knowing why a rule is in place is the best motivation for following it.

Operational Security. If a partial or whole database is transferred from a central site to another location, protection becomes more difficult, especially if the data is transferred to system with different security controls. *Digital signature* techniques can protect the integrity of database dumps; thus, a county system that receives a copy of its database on a DVD-ROM could verify that the copy was properly created by the statewide system. Further, a combination of encrypted media and procedural controls (i.e., the presence of two people to decrypt the data) can help.

Security Against Technical Attacks. VRD systems must be secured against technical attacks, including attacks both by outside "hackers" and by insiders. When any system is connected to an open communication network, including the Internet, a wireless network, or the phone system, the risk from hackers becomes substantial. Any network-connected VRD will be exposed to attacks from anyone anywhere in the world who cares to attack it; therefore, system security needs to be sufficiently robust to survive the inevitable

onslaught of attacks. It is imperative that security be considered starting very early in the software development lifecycle so that design decisions can be made in ways that maximize security. Trying to add security as an afterthought to a completed system often leads to catastrophic security failures.

First, all communication channels should be secured. Anything transmitted over open communication networks such as the Internet, wireless network, or the phone system should be protected using end-to-end cryptography (such as a VPN or an encrypted network tunnel). This cryptography requirement applies to all channels of communication including those between local election officials and the central database. It may also be prudent to cryptographically protect all data sent over internal networks to limit the damage if a hacker is able to break into the internal network or if an insider seeks to attack the system. Cryptography is especially important if wireless networks are employed, because otherwise anyone within radio range can effectively gain insider access to the wireless network.

Second, defenses should be applied to prevent outsiders from penetrating internal systems. Firewalls should be used to severely limit connectivity between internal and external networks. One simple strategy might be to completely disconnect voting registration systems from all open networks. For example, county officials might communicate with central servers by sending authenticated DVD-ROMs through the mail. Alternatively, if network connectivity is necessary, firewalls should be used to minimize the set of communication protocols, network services, and destination addresses allowed to cross the firewall communicate from the internal network to the external network or vice versa.

Mission-critical machines should be hardened as much as possible, and they should be professionally administered. All relevant security patches should be applied, and virus scanners should be used where appropriate. Unnecessary network services should be disabled. These machines and networks should be used only for voter registration.

Third, mechanisms should be deployed to detect any penetration of system defenses, as well as any insider misuse. For example, application-specific intrusion detection systems could be used to monitor the number of updates to the VRD. Any large spike in activity, whether by an authorized user or in the aggregate, might warrant human attention. In addition, officials could consider contracting with a third-party network security monitoring service to detect network intrusions and attempted attacks on the system.

Fourth, care should be taken to ensure that it is possible to recover from security failures. Regular backups are a simple and effective method for recovering from known failures. All modifications to the database should be logged to write-once media to provide a trustworthy audit trail and enable after-the-fact forensic investigations. Offsite storage of backups can reduce the risk of catastrophic loss of voter registration data. However, backups themselves must be secured, possibly including encryption, so that their loss does not compromise voter privacy or reveal information.

Denial-of-service attacks are particularly vexing. Such an attack could render the VRD unreachable or non-functional when it is most needed. Election officials should be aware that systems connected to open networks are almost invariably subject to malicious denial-of-service attacks that render the system unavailable or unreachable. Because it is beyond the state-of-the-art to completely prevent denial-of-service attacks, either officials

should have a plan prepared for how to ride out and survive such attacks, or they should avoid the use of open networks. For example, one might arrange to use DVD-ROMs if the network has been rendered unusable. Because of the threat of Election Day denial-of-service attacks, officials should ensure that it is possible to function without any network connectivity on this day. Options might include downloading all critical data to polling places several days in advance or distributing copies of the registration list printed on paper. These issues are also discussed in Chapter 6 on reliability.

Fifth, officials should obtain an independent review of their system before deployment. We recommend hiring a group of skilled experts to evaluate VRD security. These experts will conduct a thorough risk analysis of system requirements, architecture, security processes, and all other aspects of the system. These reviews should check for flaws that would allow attackers to obtain privacy-sensitive information, to compromise the integrity of the database by modifying information without authorization, or to mount denial-of-service attacks that would render the VRD inoperable. The use of technical, physical, and human procedural measures to attempt to penetrate a system can also identify security problems that might otherwise have been overlooked.

Officials should consider including an independent security review and publication of the software as part of the acceptance testing for the system. Claims that the security of the system will be endangered by such a review should be treated with extreme skepticism or rejected outright.

Sixth, the technical security of the system needs to be viewed as an ongoing responsibility, with resources devoted to it accordingly. Election officials may find it useful to perform periodic security audits of their system to ensure that system security is kept up to date as technology and attacks change. As the system will evolve over time, and as the threats will change with time, it is important that the system be tested for security issues on a periodic basis. In particular, the system should be fully evaluated after any major upgrades and after recovery from any significant incident

Dealing with Security Failures. In spite of good security measures, there is always some possibility that an attacker will carry out a successful attack. When successful attacks do occur, the system should protect the ability of users (including both election officials and voters) to carry out their activities with as little disruption as possible. Additionally, because prosecution of attackers can act as a deterrent to future attackers, it is important that systems be designed to support potential identification and prosecution of attackers, for example, by keeping audit logs and maintaining a proper chain of custody for relevant records.

Electronic registration databases heighten the need for well-designed recovery mechanisms, because a statewide electronic database potentially introduces opportunities for more, and more significant, failures. To the extent possible, existing policies and laws should be applied.

We discuss three categories of security failures:

- unauthorized disclosure of data in which some data is seen by someone who is not authorized to see the data;
- breaches of integrity, in which ineligible voters are wrongly registered and/or actually vote or in which eligible voters are disenfranchised or wrongly prevented from

voting; and
- breaches of reliability, possibly occurring on Election Day, in which legitimate users of the database are unable to get necessary results.

Unauthorized Disclosure of Data. Disclosure can occur by accident or on purpose. Unauthorized disclosure can happen when an authorized user of the system exceeds his privileges, or when an outsider gains unauthorized access. VRDs should incorporate audit logs (discussed in Chapter 2 on accuracy) that record all attempts to read registration data. With appropriate scrutiny of these audit logs, it may be possible to detect many cases of unauthorized disclosure.

To the extent possible, individuals should be notified if it is determined that data about them has been or may have been obtained inappropriately. Security breach notification laws in California and other states are already having a beneficial effect in this regard.

Breaches of Integrity. The intentional corruption of official records is both a federal and state crime falling under many different statutes, giving prosecutors a number of options. However, unless appropriate audit trails, procedures, and detective controls are in place, security breaches are unlikely to be noticed and identified as potential criminal acts.

Because of the high legal and public relations cost of disenfranchising legitimate voters or allowing ineligible people to vote on Election Day, states should have procedures for auditing and quantifying the accuracy of registration data before an election. For example, election officials could perform an audit of a statistically-significant random sample of all changes to the voter registration database since the last election to look for anomalies, followed by a more thorough audit if anomalies are found. Such an audit should be performed sufficiently in advance that corrective actions can be taken before Election Day if errors are discovered.

To avoid disenfranchisement of legitimate voters on Election Day, it is also important to avoid creating a culture among poll workers that assumes that the computer is always right. In particular, it should be possible for someone who thinks she is a registered voter but is not in the database to cast a provisional ballot that can be counted later, if it is subsequently determined that she is an eligible voter.

Breaches of Reliability. Unlike breaches of data and integrity, which can go undetected, breaches of reliability are easily detectable. Audit logs, including firewall logs, are crucial for tracing and perhaps prosecuting malicious attackers. To limit the impact of reliability breaches on Election Day, we suggest that each polling place be given a backup copy of the data that will be needed to validate eligible voters within that precinct. This list should contain only the information needed for validating voters. For example, Social Security numbers might be redacted from the backup list. Existing policies allow the polls to be kept open beyond the scheduled closing time if failures occur; we recommend these policies be followed. Reliability issues and fallback procedures are discussed in more detail in the next chapter.

6. Reliability

Reliability is often thought of as system availability (i.e., whether the system is up and running 24×7). However, to better understand reliability, we need to understand the *goals* of reliability in a statewide VRD.

While 24×7 operation may achieve these goals, solutions that are more economical are possible because reliability can often be achieved without continuous online access to the database. For example, if regulations impose a deadline for registration or registration changes sufficiently in advance of an election, static snapshot copies of the database may be adequate for supporting Election Day verification of voter registration. Static copies may well prove more reliable than attempting to guarantee reliable network access from each polling place.

We assume that VRDs will have more intense usage in the months immediately prior to the election, with a very large spike in usage during and immediately following the election itself. Activity before an election includes absentee voting, in-person absentee voting, and early voting. Absentee and in-person absentee voting occurs anywhere from 10 to 45 days before Election Day, and early voting usually occurs 10 to 14 days before Election Day.

With this in mind, we divide the recommendations for design of a VRD into two classes—namely, *technical* and *operational* mechanisms for ensuring reliability. We also give recommendations for ensuring reliability during the development of the database.

Technical Mechanisms for Reliability. The hardware/software combination used to access the VRD needs to provide good response and reliable service. It should be designed to work well both in non-election times, when the major activity is voter registration, and in the high-activity times, immediately prior to and during the election itself.

We list several design choices that can be used to improve reliability and discuss recommendations and caveats to be considered when evaluating choices.

Redundancy. While redundant communications systems (e.g., multiple network connections from different providers) have been used successfully,[28] care must be taken to ensure that the systems are truly redundant. For example, a modem and ADSL connection[29] over the same phone line provides little redundancy; two ADSL lines from different providers probably provides less still, as they likely utilize the same central

[28] The Federal Aviation Administration, for example, makes frequent use of redundant systems for air-traffic control. This includes both alternate communication links such as redundant fiber links for Airport Lighting Control and Monitoring Systems (see AC 150/5345-56) and independent approaches such as using Flight Service Station communications as a backup for relaying Air Route Traffic Control Center instructions if direct ARTCC contact is lost (see Aeronautical Information Manual).

[29] "ADSL, which stands for Asymmetric Digital Subscriber Line, is a broadband communication technology designed for use on regular phone lines. It has the ability to move data over the phone lines at speeds up to 140 times speedier than the fastest analog modems available today." From http://www.dsllife.com/tutorial/faq.htm.

switch leased from the local telephone company.

A more robust form of redundancy is to support independent approaches to accomplish the same task. For example, using online access to a centralized statewide database as the primary means of entering voter registrations allows immediate verification of registration. However, a power failure affecting a central database immediately before a registration deadline could prevent registration workers from entering registrations in time. Allowing local entry, followed by later online validation/verification of the entered values, could provide operational reliability similar to redundant power sources for the central database, but at less cost.

Replication. Replicating data in multiple places has value, but the impact of likely or anticipated types of failures must be evaluated to ensure that replication significantly increases reliability. Replicating the database may not protect against software failures that cause errors to spread to all copies, and keeping the replicated databases in different physical locations has the added cost of space for the replicated system and communication lines between the locations for updating the replicas. Additionally, replicated data may not be useful if communications problems at polling places make network access unavailable. Careful archival procedures combined with adequate fallback procedures may be more cost effective and provide as effective reliability as replication.

For example, sending DVD-ROM copies of the relevant part of the database to polling places shortly before an election would provide both a high degree of replication and a fallback procedure for access to the data if either centralized database or communication failures occur. However, the use of DVD-ROM copies must be tempered by the increased risk of disclosure of information. The information stored on such copies should include only the data that would otherwise be available to the polling place and no more. As discussed in Chapter 5 on security, encryption and digital signatures, along with appropriate policies for their use, should be used to protect these copies.

Building and including redundancy is not sufficient. The system must also be tested under realistic situations as discussed in the testing section of this chapter.

Distribution. As was discussed in the introduction, statewide VRDs are being implemented as top-down, in which the master copy of the database is stored in a centralized location, or bottom-up, in which the master copy of the database is actually distributed among many databases. A properly designed distributed database can provide a centralized list of voters as HAVA mandates. The design of a distributed VRD must be evaluated to ensure both that no single failure (hardware or software) can bring down all the connected databases and that fallback procedures are adequate in each county to protect against localized failure.

Distributed databases can serve as a good backup and contain damage caused by failures, including software failures and actions of malicious insiders. However, designers must be aware that distributed database systems can be vulnerable to mass propagation of errors if processes are designed to apply to all the databases at once. Another potential problem is to design a distributed database system so that individual parts cannot act independently. For example, one can design a system that requires that a county database coordinate with a central database for every transaction. It is important

to design a distributed database so that these possibilities are minimized.

Database distribution also increases the difficulty of ensuring the accuracy of the data, unless the system is designed to coordinate the data in the individual databases. One of HAVA's main requirements is that data should be coordinated between VRDs and other databases; therefore, this element should be part of every distributed VRD's design.

Centralization. Centralized databases face a different set of reliability challenges. If the entire database is stored in a central location, this location becomes a single point of failure. Power difficulties, network problems, or other reliability problems with the central location might bring voter registration activities to a halt throughout the state. Although replication and redundancy can help reduce such risks, additional costs may be involved. The use of alternative methods to access and input data, including DVD-ROMs, printed voter lists, and paper forms, may be particularly important when centralized databases are used.

Archives. When data is backed up, the backup files can be recycled or can be retained as long-term archives. Archives safeguard against loss from software failures, intrusion, or malicious insiders who could damage less resilient kinds of backup. Consequently, an archive must be protected from modification through write-once media such as DVD-ROM to ensure against both accidental and intentional erasure or modification.

A second use of archival material is for forensics—that is, identifying what went wrong when a failure occurs, correcting the problem, and preventing new failures (this includes both human- and system-caused failures). To ensure detection of malicious action, it is important to log and archive all changes to the database. With the decreasing cost and increasing density of backup media, long-term maintenance of such logs, which we recommend, can be achieved at reasonable cost.

Operational Mechanisms for Reliability. Reliability will not be achieved solely through technical means. Provisions must exist to ensure the integrity of the election process in spite of possible Election Day failure of the registration database. Since it should always be assumed that something could go wrong, a system must include operational procedures, or fallback processes, that ensure reliability in spite of technical failures. While these procedures are often tied to the technical design decisions, it is necessary to *document* the operational procedures to be followed in the event of database failures.

We recommend that for each process there be at least one specified alternate process to follow in case of failure. In particular, there should be a fallback procedure for each process that could affect the ability of people to vote on Election Day. For example, suppose the process requires that voters physically sign a voter registration list. In case the correct list is not sent to the polling place, we recommend as a fallback that there be a back-up computer system and communication line available at the polling place, so that people's names can be looked up online. If the process requires that voters' names be looked up online, then a fallback would be to provide paper copies of the list in case the computers or their connections go down. Further, Election Day verifications can be done (1) via paper systems, (2) via personal computers or handheld devices with DVD-ROMs, or other methods of holding static copies of the voter list, or (3) via personal computers

or handheld devices connected by electronic communication links to central VRDs. Regardless of the method used, a fallback process should be devised to deal with its failure. When appropriate, these processes should operate in tandem with provisional balloting and other measures designed to protect a voter's right to vote.

Provisions for Delayed Entry of Registration Information. While direct entry of voter registration information into the database may be desirable, allowing immediate confirmation of registration that requires direct entry could undermine the registration process in case of system failure. As discussed previously, fallback procedures must be developed to support alternate means of registration.

Testing Issues. A VRD must be tested to ensure that it will function reliably when placed into service. The problem is that it is impossible to do a true stress test on a VRD because there is no way to completely replicate the stress of an Election Day except to have an election. However, through effective modeling of the system and its capabilities, it is possible to design tests that effectively simulate the stress of actual use. This imposes the requirements on the contracting agency of ensuring that sufficient information is available for vendors/developers and quality assurance groups to adequately model the system.

Technical measures designed to increase greater reliability also should be tested. When used, replication and redundancy facilities might be tested by trying to operate the system when all or part of the system is unexpectedly taken offline. In accordance with EAC recommendations, archival backups should be tested regularly.[30]

The system also should be secured against external network-based attacks (see Chapter 5 on Security). Tests that simulate denial-of-service and related attacks can be used to evaluate the robustness of the VRD and possibly identify weaknesses that should be addressed.

[30] U.S. Election Assistance Commission, 2005, *Voluntary Guidance on Statewide Voter Registration Lists*, available online at http://www.eac.gov/docs/Statewide_Registration_Guidelines_072605.html.

Appendix A: Glossary

Following is a relatively non-technical glossary of terms referred to in the report. Our intent is for the report to be readable by as many people as possible; for that reason, many of the definitions below are not as technically detailed as they might be. For more exhaustive technical definitions or explanations of these and other related terms, please refer to one of the two documents noted at the end of the glossary.

Access control policy – A list of rules assigning access privileges to system users.

Access privilege – The right to read or update a particular kind of data, or to execute a particular operation.

Application – One or more computer programs developed to provide specific functionality. Examples include such things as word processing applications, web browsers, database software, and so on.

Authentication – The process of verifying that a person is who he or she claims to be – for example, specific knowledge of a personal identification number (or PIN) is often used to authenticate ATM card users.

Backups – Copies made for the purpose of safeguarding information; making regular backups of important data is a widely recognized best-practice.

Batch update – A group of additions, modifications, or deletions to a database received from what is believed to be an authorized source (e.g., a local county).

Biometrics – Authentication techniques that rely on an individual's physical attributes (for example, fingerprints, iris scans, facial recognition, and so on).

Bottom-up – Approach to managing voter registration data whereby each county or municipality may keep its own database of records for voters within the county, and the county's records may be reconciled with a database run by the state on a periodic basis. *See also "top-down."*

CAPTCHA™ (Completely Automated Public Turing Test to Tell Computers and Humans Apart) – is a mechanism used to verify that a human user is completing a form, as opposed to a computer program. Generally, CAPTCHA tests consist of an image that contains distorted text that is easy to read for humans, but very difficult for computer software to interpret.

Certificate – A cryptographic tool used to verify such things as the identity of a computer, the source of a program, the integrity of a message, or the identity of the source of a message.

Ciphertext – Information rendered unintelligible except to those who can decrypt it; (encrypted plaintext).

Data element – A basic data structure in a database (for example, "last name," "address," "city," and so on).

DBMS – A database management system is a computer program (or a suite of programs)

that enables users to store, modify, and retrieve information from a database.

Decryption – The process of turning ciphertext back into plaintext. *See Encryption.*

Denial of service attack – An attack on a system where the objective is to prevent the normal use of that system, often by overwhelming the system with a large volume of seemingly normal transactions or requests for data.

Digital Signature – An electronic signature based upon cryptographic methods of originator authentication, computed by using a set of rules and a set of parameters such that the identity of the signer and the integrity of the data can be verified.

DVD-ROM – Digital versatile disk (originally "digital video disc") is an optical storage disk similar to compact disks (CDs). However, DVDs are capable of storing much more data. ROM, or read-only memory, refers to disks that are capable of being written to only once.

Encryption – The process for turning plaintext (e.g., a person's name and address) into ciphertext, where the meaning of the encrypted plaintext is obfuscated. *See also Decryption.*

FIPs – Fair Information Practices, a widely accepted set of principles (e.g., notice, security, minimization, and so on) for addressing concerns about information privacy.

Firewall – A means for preventing unauthorized access to a given system. Firewalls (both hardware and software firewalls) allow administrators to regulate the kind of traffic and data that flow into and out of a system.

HAVA – The Help America Vote Act of 2002 (P.L. 107-252). Election reform legislation that mandated statewide VRDs.

Heuristic evaluation – A strategy for evaluating user interface designs. In heuristic evaluations, usability experts examine user interfaces for consistency, proper feedback and error handling, and other criteria. Heuristic evaluation can often be a cost-effective alternative to more rigorous evaluation via controlled user studies.

Internet Protocol (commonly referred to as "IP") – is a connectionless, best-effort packet-switching protocol and makes up part of the TCP/IP suite of protocols that enable machines to communicate with each other on the Internet.

Intrusion detection system – An application designed to detect attacks on a network or computer system.

Logs – Records of actions within a system, often contained in specific files (for example, audit log files, error log files, and so on). Information found in logs generally includes a description of what was done, when it was done, who did it, and other details necessary to construct and accurate and complete record of what happened.

Merges/purges – Batch updates that involve the integration, alteration, or removal of large amounts of data in an automated fashion (for example, updating voter records in a database by comparing data with a driver's license database, or removing records in a voter database based on records added to a death record database).

Plaintext – Intelligible information; generally in a form readable by a person (decrypted ciphertext).

SSN – Social Security number.

Top-down – An approach to managing voter registration data whereby state officials administer a single master computer server; all voter records are stored on that central server, and all requests to view or modify voter records are executed on the central server. *See also "bottom-up."*

Truncation – The practice of displaying only part of an identifying number (e.g., a Social Security number) for the purposes of identity verification.

VPN – Virtual private network.

VRD – Voter registration database.

Web-based – Applications that are accessed via the Internet (or an intranet), generally using a web browser (e.g., web-based email services like Google's Gmail or Yahoo! Mail).

Note: Other relevant resources include the glossary associated with Volume One of the U.S. Election Assistance Commission's *Voluntary Voting System Guidelines*, which is available online at http://eac.gov/vvsg_intro.htm, and the Consolidated Security Glossary by the NIST IEEE POSIX P1003.6 Security Working Group, which is available at http://www-08.nist.gov/posix/framework_wg/glossary.asc.

Appendix B: Biographies of Committee Members

Paula Hawthorn, Ph.D., *Co-Chair*

Dr. Hawthorn received her Ph.D. in Electrical Engineering and Computer Science from the University of California in 1979. Her thesis topic was on the performance of database systems. She has spent much of her career as a manager of database development, including Vice-President of Software Development for start-ups such as Britton Lee and Illustra, and both management and individual contributor positions at Hewlett-Packard (working on database performance issues) and Lawrence Berkeley National Laboratory. She is now mostly retired, with occasional consulting and continuing involvement with U.C. Berkeley.

Barbara Simons, Ph.D., *Co-Chair*

Dr. Simons earned her Ph.D. from U.C. Berkeley and was a computer science researcher at IBM Research, where she worked on compiler optimization, algorithm analysis, and scheduling theory. A former President of the Association for Computing Machinery (ACM), Dr. Simons founded ACM's U.S. Public Policy Committee (USACM) and served for many years as chair or co-chair of USACM. She was a member of the National Science Foundation panel on Internet Voting, the security peer review group for the DoD's Internet voting project (SERVE), and the President's Export Council's Subcommittee on Encryption. She is on several boards of directors, including the U.C. Berkeley Engineering Fund and the Electronic Privacy Information Center, as well as the Advisory Board of the Oxford Internet Institute and the Public Interest Registry's .ORG Advisory Council. She has testified before both the U.S. and the California legislatures. Dr. Simons is currently co-authoring a book on voting machines and related issues.

Steven M. Bellovin, Ph.D.

Dr. Bellovin is a Professor of Computer Science at Columbia University. He recently joined the faculty after many years at Bell Labs and AT&T Labs Research. He is an AT&T Fellow and a member of the National Academy of Engineering. Dr. Bellovin is the coauthor of *Firewalls and Internet Security: Repelling the Wily Hacker* (2d ed. 2003) and holds several patents on cryptographic and network protocols. He has served on many National Research Council (NRC) study committees and is a member of the Department of Homeland Security's Science and Technology Advisory Committee. He has been a member of the Internet Architecture Board and co-director of the Security Area of the Internet Engineering Task Force.

Chris Clifton, Ph.D.

Professor Clifton has a Ph.D. in Computer Science from Princeton University, and Bachelor's and Master's degrees from the Massachusetts Institute of Technology. He first worked on reliability and availability of database systems at IBM Research in the 1980s. He also worked on data mining and database security issues while at the MITRE

Corporation and, more recently, has been leading research on privacy-preserving data mining since joining the faculty of Purdue University.

Lillie Coney

Ms. Coney is Associate Director with the Electronic Privacy Information Center (EPIC). Her issue areas include nanotechnology, surveillance, children's privacy, civil rights and privacy, coalition development, spectrum, census, and electronic voting. Ms. Coney also serves as Coordinator of the recently established National Committee on Voting Integrity (NCVI). NCVI was created in 2003 in response to growing concerns about the reliability of electronic voting systems.

Robert Gellman

Robert Gellman is a privacy and information-policy consultant in Washington, D.C. He advises companies, government agencies, and other institutions on how to address privacy concerns on the Internet, implement the federal medical-privacy rules, and integrate privacy law and policy in their national and international operations. A graduate of Yale Law School, Gellman has worked on information-policy issues for more than 25 years. He spent 17 years as chief counsel to a subcommittee in the U.S. House of Representatives responsible for privacy, freedom of information, government information dissemination, health-record confidentiality, and other information-policy matters. He also served as a member of the U.S. Department of Health and Human Service's National Committee on Vital and Health Statistics (1996-2000), a federal advisory committee with responsibilities for health-information infrastructure matters, including the Health Insurance Portability and Accountability Act.

Harry Hochheiser, Ph.D.

Dr. Hochheiser received his Ph.D. in Computer Science from the University of Maryland, and bachelor's and master's degrees from the Massachusetts Institute of Technology. His research interests include information visualization, bioinformatics, human-computer interaction, universal usability, and privacy. A former member of the board of directors of the Computer Professionals for Social Responsibility (CSPR), Dr. Hochheiser wrote CPSR's FAQ on Internet filtering systems. He has also written about the policy implications of Internet privacy protocols. He is a founding member of the ACM SIGCHI Committee on U.S. Public Policy.

Ralph Spencer Poore

Ralph Spencer Poore (Principal Consultant at Inovè LLC and Senior Partner at Pi "R" Squared Consulting LLP) has over thirty years of information technology experience with emphasis on privacy, security, audit and control in electronic commerce, enterprise systems, and enabling technologies. His involvement in national and international standards for electronic commerce includes participation on two Internet Engineering Task Force (IETF) working groups and chairmanship of an ad hoc working group of the

Accredited Standards Committee X9, Financial Services, subcommittee X9F Data and Information Security. He founded and chaired the Standards Review Committee of the Information Systems Security Association (ISSA) and participates on the Global Executive Committee of the Generally Accepted Information Security Principles (GAISP) Committee. Mr. Poore has developed and patented security and privacy products, taught cryptographic security courses, and provided assurance services across a broad range of private sector and governmental organizations. He is an inventor, author, and frequent speaker on topics ranging from privacy to transnational data flows. Mr. Poore is a Certified Fraud Examiner (CFE), Certified Information Systems Auditor (CISA), Certified Information Systems Security Professional (CISSP), and Certified in Homeland Security-Level III (CHS-III).

Arnold Rosenthal, Ph.D.

Dr. Rosenthal is a Principal Scientist at The MITRE Corporation, doing consulting and research on databases and distributed systems. In recent years, his research and consulting has focused on data sharing, privacy, and security. He received a Ph.D. in 1974 from U.C. Berkeley. He was on the faculty of the University of Michigan and worked at Sperry Research and Computer Corporation of America. He has held visiting positions at the Swiss Federal Polytechnic (ETH Zurich) and IBM Research. He has served on numerous conference program committees and is an Associate Editor of the ACM Transactions on Database Systems.

David Wagner, Ph.D.

Professor Wagner is an Assistant Professor in the Computer Science Division at the University of California at Berkeley with extensive experience in computer security and cryptography. Dr. Wagner is an Alfred P. Sloan Research Fellow and a CRA Digital Government Fellow. Dr. Wagner was a co-designer of one of the Advanced Encryption Standard finalists, and he remains active in the areas of computer security, cryptography, and e-voting. In the past, Dr. Wagner has served as a member of the Security Peer Review Group for the SERVE Internet voting project and as a technical advisor to the ACLU Ad-Hoc Committee on Touchscreen Voting. Currently, Dr. Wagner is a member of the California Secretary of State's Voting Systems Technical Assessment Advisory Board.

Rebecca N. Wright, Ph.D.

Dr. Wright is an Associate Professor in the Computer Science Department at Stevens Institute of Technology in Hoboken, New Jersey. Her research spans the area of information security, including cryptography, privacy, foundations of computer security, and fault-tolerant distributed computing. Dr. Wright serves as an editor of the Journal of Computer Security (IOS Press) and the International Journal of Information and Computer Security (Inderscience), and she is a former member of the board of directors of the International Association for Cryptologic Research. She received a Ph.D. in

Computer Science from Yale University in 1994 and a B.A. from Columbia University in 1988.