Electoral Data in the U.S. -- Uses, Operations, Requirements, and Vulnerabilities

Micah Altman, Center for Research on Open and Equitable Scholarship
Michael P. McDonald, Department of Political Science, University of Florida

Abstract. Systematic management of and access to electoral data are critical for establishing that democratic processes are upheld; for evaluating government institutions and their effectiveness; and for understanding citizens' changing relationships with government. In this chapter, we systematically review the information used to plan, administer, and evaluate elections -- and identify gaps in public access to data, durable long-term access, boundary consistency, and voter registration database versioning. We discuss how these gaps expand vulnerabilities for insiders who aim to suppress votes or undermine public confidence in local election organizations and officials.

Cite as*:* Micah Altman and Michael P. McDonald, "Electoral Data in the U.S. -- Uses, Operations, Requirements, and Vulnerabilities." Forthcoming in Michael Ritter (eds.), *Rising Above Conspiracy: Understanding Elections, Election Administration, and Democracy in America and Abroad in the 2020s*. De Gruyter

Acknowledgments*:* The authors thank the organizers of ADRCON25, Amy O'Hara and Ngan MacDonald, and its participants, for comments on earlier versions of this work.

Funding: A grant from the Houston Endowment partially supported Dr. McDonald in conducting data collection related to this publication.

Author Contributions*:* Authors are listed in alphabetical order. All authors take equal responsibility for revisions, the final publication version, and conclusions therein.

Specific contributions are listed using the standard contributor taxonomy (Allen 2014): Conceptualization: MA, MM; Data curation: MM; Funding acquisition: MM; Methodology: MA, MM; Writing -- origina draft: MM; Writing -- review & editing: MA, MM

1. Introduction

Elections are the heart of democracy. Systematic access to election data is crucial for verifying that democratic processes are functioning properly; evaluating the effectiveness of government institutions and programs, and understanding citizens' relationships with the government. Such data shows how citizens participate, the effects of policies and changing electoral conditions on political behavior, and the operational costs of electoral institutions. Election data are critical to a wide range of policy decisions, including voting identification requirements, early voting options, selection of voting machines, polling place locations and hours, and district boundaries. (See Saltman 2006, Alvarez and Grofman 2014, Hale *et* al. 2015, for discussion of the general roles of election administration, data, and technology, and Alvarez 2023 for a bibliography of related research.)

Despite its importance, members of the public and researchers face significant challenges in accessing, analyzing, and auditing electoral data. The decentralized structure of federal and state election administration yields wide variation in management practices across states and localities: complete data are often unavailable from a single source, are difficult and costly to obtain, and are provided in non-standard formats. Records management and data quality processes are inconsistent: Processes may be opaque to the public, or even resistant to audit -- and the data itself may not be retained, versioned, or authenticated. Process weaknesses and a lack of transparency foster an environment where controversies and conspiracies arise from actors seeking to undermine the legitimacy of democratic governance.

2. Election Data Overview

The United States is exceptional in its federal form of election administration structure. There is no national electoral management board responsible for administering federal elections. Instead, authority is devolved to the states, with nearly all states further devolving important responsibilities to local election officials (LEOs). Election officials across all levels of government are responsible for maintaining election data that include, but are not limited to, individuals' voter registration records, election results, electoral boundaries, and summary statistics on electoral performance. Varying federal and state policies sculpt the data and technology that election officials must manage, and states vary in how they devolve data management to LEOs.

Catalyzed by the federal 2002 Help America Vote Act (HAVA), which required states to compile electronic statewide voter registration databases, election officials shifted their data from paper to electronic systems. As this shift progressed, stakeholders became increasingly concerned that election outcomes could be subverted through unintentional errors or intentional intrusions into election infrastructure. Election data and technology concerns first entered the national stage

following the 2004 presidential election, when activists alleged Ohio voting machines were programmed to give President George W. Bush a victory (Bliefuss and Friedman 2006). The United States intelligence agencies raised the stakes higher when they detected Russian-backed intrusions during the 2016 presidential election into state and local election management systems (U.S. Senate Select Committee on Intelligence 2018).

To distinguish states' approaches to implementing voter registration databases under HAVA, stakeholders generally classify states' election management ecosystems along a single dimension of centrality. In *top-down* states, the state provides election technology to localities. In *bottom-up* states, localities manage their infrastructure, but must regularly transmit voter registration records to the state office to comply with HAVA mandates. In *hybrid* states, a state office provides infrastructure that localities may opt to use.

Election management has expanded to include additional functions, and thus, this notion of centrality may also apply to other election duties, such as mapping tools for election boundaries and ballot tracking. The nature of bottom-up has evolved, with universally dominant statewide vendors emerging in some states that employ the bottom-up approach. Furthermore, states' databases have become increasingly dependent on external databases, such as the federal Systematic Alien Verification and Entitlement (SAVE) and states' Department of Motor Vehicles databases.

3. The State of Electoral Data Availability and Curation

Election officials produce election data through highly localized and diverse processes, resulting in substantial public opacity. These processes vary in their ability to produce accurate and reliable data through their election management systems. These process limitations create opportunities for unintentional or intentional data alteration -- vulnerabilities which can negatively impact democratic performance.

To address these issues, we characterize electoral management practices using four complementary frameworks: public availability, records management best practices, data quality criteria, and administrative process requirements. We then discuss potential vulnerabilities that arise from these gaps.

3.1 Frameworks of Analysis

Records Management:
Authenticity, Preservation, Fixity, Versioning, and Provenance

Election management systems are an aspect of government records management. Government records management systems exist to support their internal purposes and to ensure external

regulatory and legal compliance (Guercio and Thibodeaux 2001: 252). We thus analyze electoral data management by reference to government records management principles and best practices developed by archivists -- much of which are codified in law, regulation, and legal rules of evidence (see Seymour 2017, and Jacobs and Jacobs 2025 for reviews).

Government records management encompasses the production, acquisition, organization, and retrieval of quality data produced during administrative functions. Records management encompasses reliable short-term (*retention*) and long-term data persistence (*preservation*). Records management practices must maintain the *integrity*, *provenance, authenticity* properties of and *version control* over records in order for for access and preservation to reliably meet legal standards of evidence, provide administrative accountability, and support public trust in data (see Guercio and Thibodeaux 2001, ISO/TC 46 2016, Jacobs and Jacobs 2025, and also CNS 2025 with respect to relationships to data quality):

- *Integrity* ensures that data remains unchanged, except through authorized and documented processes. The integrity of records' content is established through fixity. For physical records, assessment of fixity involves inspection of the media (e.g., paper and ink) for alteration. To establish fixity of electronic contents requires distinctive metadata, such as digital cryptographic hashes, to be collected at the stages of data collection, data receipt, and revision.
- *Provenance* refers to the information about the origin, custody (or chain of control), and ownership of data or records.
- *Authenticity* identifies the actors and processes that created the data. Establishing authenticity is generally dependent on establishing provenance and fixity. Legal authenticity may require that the actors and processes be authorized, official, and explicitly attested.
- *Version control* involves the systematic tracking of modifications to data. Best practices generally include recording version identifiers, the history of changes, the authorizing party, the reason for the change, and sufficient information about changes to accurately reconstruct prior versions.

Data Quality: Completeness, Accuracy, Comparability, and Accessibility

Management and Information Sciences commonly refer to "data quality" as its fitness for use (Madnick et al. 2009). Other fields use analogous assessments, such as "value of information", "information content", "data reliability", and "data validity." Data quality depends upon, but is not determined by, record management. Stakeholders cannot trust that data provides reliable evidence of actions and phenomena without its authenticity, provenance, version control, and fixity. Stakeholders can trust using data as reliable evidence for description, decision-making, or analysis if it is of sufficient quality. General frameworks of data quality vary depending on the field of study. At a broad level, when applied to quantified information such as official counts

and statistics, data quality includes, at a minimum, accessibility, timeliness, completeness, accuracy, and comparability with respect to the intended use (see Groves and Lyberg 2010, and Biemer 2010, CNS 2025).

- *Accessibility* refers to a qualitative measure of ease and barriers to obtaining and using data by the designated users.
- *Timely* data provides information to support effective administration and decision-making, particularly for administrative functions with legal deadlines.
- *Completeness* refers to the situation where data contain the necessary measures and provide a sufficient representation of the population of interest for a given purpose.
- *Accuracy* refers to how well data are electronically captured, allowing systematic differences to be reliably distinguished from random error.[1]
- *Comparability* refers to the similarity between different databases. Stakeholders combining different databases must determine how various samples intersect in terms of time, space, and the population of interest; however, the degree of comparability required depends on the intended use.

Administrative Process: Reliability, Compliance, and Auditability

For election institutions to be accountable, transparent, and trustworthy depends upon the properties of election data, records, and administrative processes. (see for discussion of this, and proposed additional principles Urahn and Caudell-Feegan 2008, Altman and McDonald 2012, Carolan and Wolf 2017, USAEE 2025) And, at minimum, records management and data quality (as discussed above) generally require processes that are reliable, compliant, and auditable.

- *Reliable* processes consistently produce results over repetitions, achieving the designated goals for the process. Reliable processes are especially important for generating high-quality data and obtaining accurate and trustworthy results.
- *Compliance* refers to the process by which administrators manage data while ensuring adherence to relevant laws and regulations. In the context of election administration, numerous laws and regulations govern voter privacy, information security, records management, and public transparency.
- *Auditable* processes can be reviewed by an authorized stakeholder. Sufficient internal controls and record-keeping enable a review to verify whether a process is compliant, reliable, timely, and meets its designated goals.

3.2 Election Data Quality, Management, and Process

[1] The total survey error approach (Groves & Lyberg, 2010) is commonly used in surveys and official statistics to provide formal measures of uncertainty, calculated by managing and measuring errors at various stages of the data collection and aggregation estimation process, including errors related to measurement, linkage, coverage, and sampling.

In practice, election data administration may not fully satisfy the properties of data quality, records management, and administrative process. The United States' hyper-decentralization of election administration means that state and county election officials manage election data in different ways, making overarching identification of gaps in administrative processes, records management, and data quality complicated. Some government officials manage their election data in a manner that attempts to achieve best practices, while others are wanting. Generally, gaps are larger among smaller governments that are more resource-constrained (Zohr et al. 2024).

In its operations, election data administration encompasses multiple integrated areas of responsibility. In **Figure 1**, we broadly classify the stages at which election administrators generate these data into four categories to help illustrate how these data relate to administrative functions.

As illustrated, relevant data is collected during four administrative phases:

- **Pre-existing Data Period** refers to data collected before election administration begins
- **Pre-Election Period** refers to the time during which preparation for an election takes place
- **Election Period** refers to the time during which voting activities take place
- **Post-Election Period** refers to the time following voting activities

We intend for these generalizations to capture important dynamics, while definitions may be blurred in practice. We limit our review to major election administration functions at a high level, as we do not have space to delve deeply into all aspects.

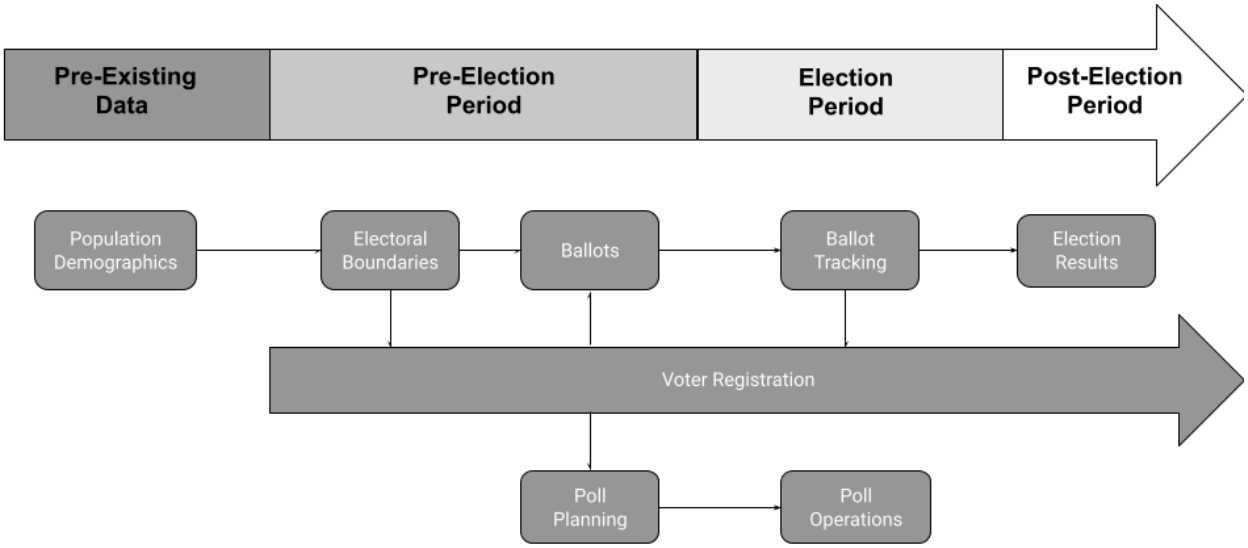

Figure 1: Key election administration areas of responsibility.

Category	Description	Public Access	Preservation status *
Population Counts	*population demographics*	available	active
Electoral Boundaries	*boundaries for districts, local governments, and precincts*	available	active
Voter Registration	*voter registration databases*	mixed	mixed
Ballots	*ballot designs*	mixed	short-term retention
Poll Planning	*precinct boundaries*	mixed	mixed
	voting machine operations polling location performance	poor	poor
Ballot Tracking	*ballot disposition*	poor	short-term retention
Election Results	*totals by office*	available	active
	by geography or voting method	mixed	mixed

Table 1: Summary of data availability and preservation for state and federal elections, by data type. [2]

To provide a view of the overall landscape of election data, we summarize the state of election data access and preservation for federal and state contests in **Table 1**. We focus on these two characteristics since they are necessary for external actors to assess other attributes of good data stewardship. Some election data are deficient with respect to external access and preservation, making it difficult to externally assess their other characteristics.[3]

3.2.1 Pre-Existing Data

Population counts figure prominently in pre-election data. The U.S Census Bureau produces population counts, reported within defined geographic units to protect individuals' confidentiality. State and local governments use population counts to draw legislative districts.

[2] Preservation status is labeled "active" when a preserving institution acts regularly and systematically to obtain authentic records; as well as to perform all actions necessary to ensure long-term access to those records; is labeled "passive" when a preserving institution accepts records of this type and performs all preservation actions with respect to records that are provided to it; and is labeled "retention" in the case where only a non-preservation offers access (possibly for an indefinite period) without making commitments beyond short-term retention.

[3] We do not review local government and special district election data, since compiliting this information remains an open challenge, and discussion of the details are beyond the intended scope of this volume.

The federal government uses these data to identify which localities must provide multilingual ballots under Section 203 of the federal Voting Rights Act.

The U.S. Census Bureau is a highly professionalized organization that operates under federal laws. Its well-documented aggregate data meet high standards of data authenticity, quality, and accessibility. The Census Bureau makes its current and past data publicly available via web portals. The National Archives and Records Administration, research universities, and other interested stakeholders systematically mirror census archives. However, individual data are embargoed for seventy-two years, which prevents timely external auditing.

3.2.2 Pre-Election Period

Electoral boundaries are produced by governments preceding elections. We review two types of electoral boundaries important to election administration: districts and precincts.

Redistricting authorities use census data to draw legislative districts. Current and past federal and state adopted districts are public information. State governments typically disseminate and archive high-quality representations of district boundaries. States transmit district boundaries to the Census Bureau for inclusion in their geographic hierarchy, adding an additional layer of availability and preservation. District drawers may internally track versions and provenance, but they may only provide their final product to the public. Governments may not make accessible or archive officially proposed districts considered, but rejected, during redistricting (see for a discussion, Altman Mac Donald, McDonald 2005; Altman and McDonald 2012; McDonald and Altman 2018).

Election officials must account for offices elected from local government boundaries, such as counties and cities. Local government boundaries are well-defined in state law and thus generally meet high standards of persistent access and preservation.

Election officials draw precincts to identify the polling locations at which voters are eligible to cast in-person votes, and use these as a basis for reporting aggregated election results statistics, thereby protecting each individual's secret ballot. LEOs prefer precinct boundaries to conform with district and local government boundaries; however, this is not always possible when small slices of overlapping districts and local government boundaries exist, or when a polling place location is unavailable within a slice. In part because of this, precinct and election boundaries are frequently difficult to accurately match and align with other official data boundaries. A potential consequence is that voters may be assigned to the wrong districts (Amos and McDonald 2020). Precinct boundaries do not always meet high standards of data accessibility and preservation (Amos, Gerontakis, and McDonald 2024). Rural LEOs, especially, may not have the capacity to generate, post, and archive accessible precinct maps.

V**oter registration databases** (VRDs) are the backbone of election administration, as they manage the approved list of individuals eligible to vote. All states except North Dakota require voter registration. Among their additional important functions VRDs identify registrants' assigned precincts and the offices for which they are eligible to vote. These systems are "live" in that they are updated in real-time across all election periods.

The primary purpose of VRDs is to assist election officials in performing their duties. LEOs first place a new registration application into a suspended status as they review a person's eligibility. If election officials deem an applicant eligible, they move the registrant into an *active* status. Registrants maintain an active status as long as they have contact with LEOs, most typically by voting. If contact lapses, election officials move the registrant into an *inactive* status, which initiates procedures to cancel a voter's registration. During these processes election officials match voter registration records with external databases to manage registrants' status, such as the Systematic Alien Verification and Entitlement to verify citizenship, states' department of motor vehicle databases to manage identification requirements, and the U.S. Post Offices National Change of Address database to identify registrants who may have moved and be placed into inactive status.[4]

VRDs generally meet high standards of accessibility and preservation, although some gaps remain. Election management systems are generally understood to internally track versioning, fixity, provenance, and authenticity to assist election officials in maintaining legal compliance while administering elections. However, LEOs do not publicly release these internal change data, provide public documentation of their internal version processes, nor provide source code for systems that are necessary to externally detect programming errors. Instead, election officials sometimes provide external stakeholders with weekly or monthly VRD snapshots or a database extract at the time of a request. They may not preserve prior snapshots for the public. VRDs are limited in other ways. Election officials do not publicly release protected private data, such as driver's license numbers, and records of at-risk individuals, such as law enforcement and domestic violence victims. States may limit VRD accessibility in other ways. External VRDs may be freely accessible, or may be available only to qualifying individuals and organizations, or exorbitantly costly.

Ballots are the heart of elections. Election officials create *ballot styles* that list the offices and ballot questions for which registrants are eligible to vote. To generate ballot styles, LEOs associate address ranges with the districts and local governments. For example, Main Street addresses with odd numbers from 1 to 99 are in District 1, and so on; these ranges are cross-referenced with registrants' residential addresses.

[4] The innovations of online and automatic registration require external databases to interface with VRDs so that LEOs may review and approve applications from these sources. We do not review the accessibility and preservation of these external databases, which are increasingly important to election administration.

During the pre-election period, LEOs create distinct electronic ballot styles to print paper mail ballots and envelopes, for paper ballots to be distributed at polling locations, and for sample ballots. For large jurisdictions, hundreds of ballot styles may exist. For the impending election, these data generally meet high standards of accessibility since these are the ballots transmitted to voters. After being cast by voters, federal law requires federal election material to be retained for two years or longer, under state law. After the retention period, however, these records are not designated for preservation and may be destroyed.

We include in the definition of ballots the programming and testing of electronic voting machines for ballot format and correctness through a process known as logic and accuracy tests. These tests are performed on-site, publicly. Following an election, there is no preservation, as election officials typically reset electronic voting machines for the next election. Similarly, election officials may program ballot tabulators for a specific election and then simply reset the tabulators after the election is over.

Poll Planning encompasses various functions needed to run an election and generate data. LEOs identify polling places. They recruit, train, and deploy election staff and volunteer poll workers. Campaigns and political parties provide lists of election observers, who are often required to be registered voters. LEOs generate check-in lists of registrants assigned to each precinct, known as poll books. Poll workers further utilize poll books to manage registrants who may need to update their registration records or process registration applications in states that allow same-day registration (SDR), which enables individuals to register and vote at a polling place on the same day. Election officials may provide poll planning data to the public upon request, and they are subject to records retention laws, but they may not always respond to public records requests (Baringer et al. 2021). After the retention period, these records are not designated for preservation and may be legally destroyed.

3.2.3 Election Period

Ballot Tracking refers to how LEOs provide voters with ballots and ensure that each voter casts at most a single valid ballot. The election period formally begins when LEOs provide the first ballots via mail or in person, which may occur before Election Day. To manage mail balloting, LEOs track when ballots are sent, when and how they are returned, and their disposition. Some registrants may request a mail ballot in the pre-election period, and valid requests may continue into the election period.[5] If LEOs determine a mail ballot request or returned ballot is deficient, they may notify the applicant for correction.[6]

[5] LEOs in vote-by-mail states deem all active registered voters as requesting a ballot, while inactive registrants must request a ballot as do those absent from their residential address.
[6] If the Post Office returns any piece of election mail as undeliverable, including a mail ballot, LEOs may designate a registrant as inactive or cancel their registration.

LEOs track ballots cast in-person at early voting locations or on Election Day. If a voter's eligibility is questioned at a polling place, poll workers must provide the voter with a provisional ballot that will be counted if election officials later determine the person is eligible. Persons who cast provisional ballots tend to be members of younger, older, and poorer vulnerable communities, and thus, provisional ballot tracking is critical for compliance.

Internally, ballot tracking data generally meets high standards of data accessibility, as this is a critical election administration function. There are rare exceptions where LEOs manage ballot tracking using paper, as is the case among rural LEOs in states with restrictive absentee ballot laws and low mail ballot usage. LEOs may fall short in administrative functions when they fail to enter ballot tracking information in a timely manner or provide clear reasons for rejections. Externally, ballot tracking data accessibility may, similar to VRDs, involve data censoring of at-risk voters, may be limited to qualifying entities, and may be costly. Archived ballot tracking data typically takes the form of "vote histories," which identify the elections a voter participated in, possibly including their method of voting, but exclude information on the mail ballot request process or disposition of rejected mail ballot requests or ballots. LEOs generally retain these data according to records retention laws, but may not preserve these data for longer periods.

Poll Operations refer to voting activities other than ballot tracking. Operations may include tracking voters' experiences by monitoring the time when voters stand in line, check in, and cast their ballot. Election officials use this information and poll worker communications to perform crisis management, and may post this information to inform the public. For the largest jurisdictions, poll operations involve directing hundreds of people through a service ticket system. Some LEOs provide poll operations information online to improve voters' experiences. However, these records are not designated for retention or preservation and may be discarded at any time

3.2.4 Post-Election Period

Election Results are initially reported and finalized by LEOs during the post-election period. Election officials typically disseminate election results through internet-accessible Election Night Reporting (ENR) systems. Election results reporting involves several steps. Some states permit LEOs to prepare mail and in-person early ballots for counting at their election offices before polls close.[7] When polling places close, LEOs begin reporting initial election results, usually as counting is completed within precincts. During the following canvass period, LEOs check election results for errors, and they count accepted provisional ballots and any valid mail ballots (depending on state law) that arrive after Election Day. At the end of the canvass, LEOs typically

[7] LEOs feed early vote paper ballots through pre-programmed tabulation machines and collect electronic voting machine storage devices. To ensure pre-processed early votes do not influence people who yet to vote, these results are stored on electronic media to be fed into ENR systems when polls close.

present the election results to an electoral board for approval and certification. The election results are then final pending any further legal action. Election officials may update ENR systems with certified results, and official documents may be generated for archiving. LEOs may conduct post-election internal auditing, known as risk-limiting audits, to detect systemic counting errors.

Summary election results generally meet high standards of accessibility and preservation, as these data are a necessary output for democratic elections. Federal and state governments typically preserve certified results for perpetuity. Where election results often fall short in accessibility and preservation is detailed data beyond summary results, such as write-in candidate tallies, registration counts, results and statistics by precinct, and results and statistics by voting method.

3.2.5 Summary

Overall, there are four areas in which substantial gaps in data quality and records management sometimes occur:

1. First, systematic public *access* to many categories of data -- summarized in Table 1 -- is limited. As a result, it is difficult or impossible for independent individuals or organizations outside the local governments to systematically obtain comprehensive data for research and auditing.
2. Second, much of this data is not systematically *preserved* (see Table 1) -- although it may be retained indefinitely at the discretion of each LEO. Moreover, even when preservation is required, it is often neither actively monitored nor enforced.
3. Third, election officials provide data in varied formats and measured over differing geographical units, making it difficult to accurately *compare* data across states and localities.
4. Fourth, few states provide publicly documented, systematic version management of voter registration data. As a consequence, external stakeholders may find it difficult or impossible to retrospectively determine the state of voter registration at the time when relevant decisions are made, to systematically analyze how data changes over time, and how those decisions were affected by interactions with external databases.

4. Election Vulnerabilities Related to Data Management

Durable access to accurate and authentic election data is necessary if we wish to document and understand how well or poorly democracy functions overall. More systematic, complete, and durable access to electoral data is crucial for determining that democratic processes are legally compliant; evaluating the effectiveness of government institutions and programs; and understanding citizens' changing relationships with the government.

Gaps in election records management, data quality, and administrative processes may increase threats to free and fair elections. From the perspective of information security,[8] gaps in record management create vulnerabilities that permit election information to be improperly added, removed, or changed under certain conditions; gaps in data quality weaken the ability to detect these events; and gaps in administrative processes further obstruct detection and mitigation. In combination, these gaps both increase the likelihood of system failures and the scope of exposure to certain threats. The very existence of these gaps may erode public trust in democratic governance.

In the context of election systems, the most significant threats may originate from two types of actors: *insiders* who have access to the election management system through their positions in government, and *malicious outsiders* who aim to disrupt the election process. Insiders may cause systems to fail through *intentional actions* or through *unintentional action*s. Like all humans, election insiders have been known to make data entry and programming errors, and to commit unintentional mistakes across the entire gamut of election processes, from voter registration to printing mail ballot return envelopes. Both insiders and outsiders may aim to accomplish three malicious objectives: *electoral manipulation, vote suppression*, and *weakening electoral institutions*. In the remainder of this section, we discuss plausible scenarios in the current gaps in election data management -- *lack of public access to data, lack of durable long-term access, boundary consistency, and voter registration database versioning* -- that create or expand vulnerabilities that have the potential to result in substantial electoral impacts in the presence of these types of insider and outsider threats.

Electoral manipulation refers to attempts to intentionally interfere in the election process to affect election outcomes by modifying an election data management system's data or outputs. Examples of internal electoral manipulation are exceedingly rare. In a prominent example, the U.S. Senate found no evidence that Russian intrusions during the 2016 November election had penetrated deeply enough to alter voter registration records or election results (U.S. Senate Select Committee on Intelligence, 2018). Presumably, good data practices permitted the U.S. Senate to reach this conclusion. The Heritage Foundation compiled the most comprehensive list of 1,400 allegations of vote fraud spanning decades. The database identifies external actors as the culprits of vote fraud, not election officials.[9] Their examples include persons voting twice, submitting a ballot for a dead spouse, felons mistakenly voting, or unscrupulous campaign operatives

[8] As a framework for analysis we use the conceptual framework of risk modeling common in information security (see JFTI 20212)] which considers risks a function of enumerating threats (an event that potentially negatively affect system function) and their sources; vulnerabilities (characteristics of the system that lead to negative outcomes in the presence of threats), impact (the harm from the threat being realized and resulting system behavior), likelihood (probability that threat will manifest, and probability of impact), and predisposing conditions.

[9] See: https://www.heritage.org/election-integrity/commentary/election-fraud-database-tops-1400-cases (accessed August 20, 2015).

violating their state's laws by sending registrants pre-filled mail ballot request forms or collecting mail ballots. None are intrusions into voting systems.

More often, "manipulation" occurs through unintentional error, such as a Virginia election official who was charged with neglect of duty for data entry and administrative errors that netted Donald Trump 3,975 votes in the 2020 election (Barakat 2024). Among the errors detectable by the public through good data stewardship was the double-entry of a precinct's election results. As an example of the benefits of accessible, high-quality data, scholars worked with election officials to correct the assignment of tens of thousands of voters to the wrong district and precinct, and to explain how they can implement their own validation checks (Amos and McDonald 2020).

A reason why it is difficult for bad actors to directly manipulate elections at scale is that U.S. election administration is highly decentralized. A hypothetically successful attack on one locality is unlikely to affect a statewide election, but if it did, it would likely be detected through comparable data available elsewhere. LEOs are vulnerable, as the U.S. Senate reports, in that LEOs may not have election management systems that meet high standards, and they do not always have effective communication capabilities to warn others if they detect an intrusion. The tradeoff for state governments to support LEOs through top-down election management systems is that while states have more resources to build robust systems, they expand the threat surface to cover all LEOs that use it.

Vote Suppression is the denial of a person's effective vote. The most straightforward denial method is to cancel a person's voter registration or to prevent them from casting a ballot, unless they re-register. Election officials are responsible for canceling voter registrations in accordance with federal and state rules, and for acting in a manner that balances voters' rights with responsible data stewardship.

Voter cancellations or purges can be controversial, with voting rights groups sometimes alleging vote suppression when a purge is overly broad or targets a disadvantaged community. A prominent example preceded the 2000 Florida presidential recount involving a mass purge of ineligible felons (Hasen 2012: 29-30). Florida's Secretary of State provided to LEOs a partially accurate list of felons ineligible to vote under Florida law. Registrants misidentified as felons tended to be from communities of color. LEOs responded to the lists haphazardly, with some accepting all of them and others ignoring the list. Florida LEOs are elected to partisan offices, with Republican LEOs tending to accept the lists while Democrats rejected them. This scenario has repeated numerous times in Florida and elsewhere regarding external databases of felons, noncitizens, and driver's licenses. Some states empower outside groups and individuals to provide their own lists of challenged voters, with no guarantees of data quality.

Local election officials have surprising discretion in how they administer elections within their localities, from managing voter registration to determining polling place locations. LEOs are responsible for identifying polling locations. LEOs can affect polling place accessibility within their localities, as one Florida LEO stated in advance of changing precinct boundaries: "I want them to go down there, and have to walk across town to go over and vote" (Amos, Smith, and Ste. Claire 2017: 137).

Accessible election data plays an important role in both internal and external monitoring of potential vote suppression and enforcing compliance with constitutional and statutory requirements. A publicly accessible and high-quality voter registration database enables monitoring of changes to it, provides evidence of mass cancellations, and identifies affected communities. LEOs providing public notice of polling place locations allow for the assessment of the impact of changes.

Where data are unavailable or difficult to obtain, monitoring and compliance enforcement are challenging. This is of particular concern for VRDs because few states provide demonstrably reliable version control, data in VRDs are often centrally stored (if not centrally managed), and the process of VRD maintenance often depends on external databases. As a consequence, it is often difficult or impossible to determine, after the fact, the state of voter registration at the time relevant decisions were made, or to systematically analyze in what ways the database changes over time, and which specific actions led to those changes.

Institutional Reputation and Confidence. Insiders and outsiders may seek to undermine public trust in election administration with the intention of influencing policies and public opinion. Outsider attacks include willful misinterpretation of election data in their possession, a frequent target of which are voter registration records with missing data values for birthdates, which give the appearance of voters living beyond human life spans. Outsiders may also make inflammatory and unjustified allegations regarding insider election manipulation or incompetence, which can lead to threats against election officials. These attacks may provide legal rationales for election changes ostensibly aimed at addressing such unfounded concerns, but may be motivated by an actor's objective to suppress voting or manipulate elections. A prominent example is President Donald Trump's claims of mail ballot fraud in the 2020 election that led some Republican-controlled states to restrict mail ballot access, a voting mode more frequently preferred by Democrats during the pandemic (McDonald 2022).

Public trust in electoral administration may be undermined by a lack of data transparency, as supposedly hidden secrets spawn conspiracy theories. Among prominent conspiracy theories is the notion that malicious actors can manipulate election outcomes by submitting fake paper ballots, hacking electronic voting machines, and hacking ballot tabulation devices. Hacking concerns were spurred by states using HAVA-supplied funds to replace error-prone voting

devices, such as punch cards with new Direct Recording Electronic (DRE) voting machines that electronically recorded and stored cast ballots using closed-source proprietary software. LEOs verified ballot styles were correctly loaded into DREs through public logic and accuracy tests, and then sealed access to the DRE ports. Activists raised concerns that software bugs or malicious code could pass logic and accuracy tests undetected, or that code could be injected following the tests. These concerns were heightened during the November 2004 election, when the president of DRE manufacturer Diebold stated that he would "deliver" Ohio to George W. Bush (Hasen 2012: 173).

The DRE debate impacted citizens' confidence in elections and, ultimately, public policy. Although no evidence exists of DRE hacking, all states – save Louisiana as of this writing – decertified DREs in favor of paper ballots. Yet, despite election officials' best efforts to demonstrate the security of paper ballots, particularly those sent by mail, and the adoption of policies like risk-limiting audits to assure the public that ballot tabulators perform correctly, conspiracy theories persist. An Antrim County tabulation machine programming error during the 2020 November election spawned a conspiracy theory that Dominion machines were programmed to elect Joe Biden. LEOs quickly identified and fixed the tabulation error, and recounted their paper ballots by hand to verify candidate vote totals. Yet, conspiracy theories persisted, with Fox News Corporation held liable for $787 million in defamation for promoting the false claims (Bauder and Mulvihill 2023). Despite clear evidence that these claims were without merit, these events have had lasting effects. Conservative activists demand the scrapping of machine tabulators in favor of hand-counting ballots, a more error-prone and labor-intensive process (Parks 2022). Republican-controlled local electoral boards have unlawfully refused to certify election results unless LEOs hand-count paper ballots (Cohen 2024).

5. Conclusion

Free and fair elections are the foundations of democracy. Such elections are a culminating event of extensive democratic processes, conducted through political and administrative institutions. Ensuring that election participation and outcomes are free and fair requires that administrative institutions be politically independent and publicly accountable; that administrative processes be reliable, transparent, and judiciable; and that election data be complete, timely, authentic, accurate, comparable, and contestable.

Scale is a fundamental challenge to election administration. State governments have the resources and capacity to create robust election management systems with data characteristics that generally support democratic institutions. Where states limit public access to election data through explicit laws or excessively high costs, thereby limiting public monitoring, courts may grant stakeholders access to opaque data when allegations of legal noncompliance arise.

Data transparency and quality shortfalls are alarming for democratic governance, but these are balanced against the increasing and real threat that malicious actors can subvert elections by hacking election systems. Decentralization creates resistance when an external intrusion into a locality does not expose a neighbor. A bad-faith internal actor can manipulate their locality, but this subversion cannot extend beyond it. Thus, manipulation may be revealed by comparing election data with that of a locality's neighbors.

Decentralization may be desirable to protect against malicious electronic attacks on democratic elections. Devolving administration to state and local governments, which have varying resources and capacities, creates data transparency and quality gaps that malicious actors may exploit. If we set aside the question of whether election administration should be centralized into a national election commission, as exists in many democratic countries, raising data transparency and quality among laggard governments is desirable. Indeed, this is the federal government's approach. For example, HAVA included grants to states to modernize voting machines. Raising election data standards above the bar in terms of data transparency and quality requirements will ensure effective monitoring and compliance, thereby protecting against internal and external threats to election subversion.

References

Allen, Liz, Amy Brand, Jo Scott, Micah Altman, and Marjorie Hlava. "Credit where credit is due." *Nature* 508 (2014): 312-313. https://doi.org/10.1038/508312a

Altman, Micah, Karin MacDonald, and Michael McDonald. 2005. "From Crayons to Computers: The Evolution of Computer Use in Redistricting." *Social Science Computer Review* 23 (3): 334–46. https://doi.org/10.1177/0894439305275855.

Altman, Micah, and Michael P McDonald. 2012. "Redistricting Principles for the Twenty-First Century." Case Western Law Review 62.

Alvarez, R. Michael, and Bernard Grofman, eds. 2014. El*ection Administration in the United States: The State of Reform after Bush v. Gore*. Cambridge University Press.

Alvarez, R. Michael. 2023. "Voting Technology and Election Administration in the United States." In *Bibliographies Political Science,*. Oxford University Press. https://doi.org/10.1093/obo/9780199756223-0358.

Amos, Brian, Gerontakis, Stephen, and McDonald, Michael. 2024. "United States Precinct Boundaries and Statewide Partisan Election Results." *Scientific Data* 11, 1173. https://doi.org/10.1038/s41597-024-04024-2

Amos, Brian and Michael P. McDonald. 2020. "A Method to Audit the Assignment of Voters to Districts." *Political Analysis* 28(3): 356-71. https://www.jstor.org/stable/27116015.

Amos, Brian, Daniel A. Smith, and Casey Ste. Claire. 2017. "Reprecincting and Voting Behavior." *Political Behavior* 39(1): 133–56. https://www.jstor.org/stable/48693872.

Anderson Christopher J, and LoTempio, Andrew J. 2002. "Winning, Losing and Political Trust in America." *British Journal of Political Science* 32(2): 335–51. doi: 10.1017/S0007123402000133.

Anderson Christopher J., Blais André, Bowler Shaun, Donovan Todd, Listhaug Ola. 2005. *Losers' Consent*. Oxford, UK: Oxford University Press.

Barakat, Matthew. 2024. "Prince William County admits election tally in 2020 shorted Joe Biden." NBC Washington, January 12, 2024. https://www.nbcwashington.com/news/local/northern-virginia/prince-william-county-admits-election-tally-in-2020-shorted-joe-biden/3514995/

Baringer, Anna., Eichermuller, Justin, Zelin, William, Shino, Enrijeta, and Smith, Danial A. 2021. "Election Administration and Public Records Responsiveness." *Public Integrity* 24(3): 280–291. https://doi.org/10.1080/10999922.2021.1932330

Bauder, David, and Mulvihill, Geoff. 2023. "Fox, Dominion reach $787M settlement over election claims." Associated Press. April 18, 2023. https://apnews.com/article/fox-news-dominion-lawsuit-trial-trump-2020-0ac71f75acfacc52ea80b3e747fb0afe

Biemer, P. P. 2010. "Total Survey Error: Design, Implementation, and Evaluation." Public Opinion Quarterly 74 (5): 817–48. https://doi.org/10.1093/poq/nfq058.

Carolan, Liz, and Peter Wolf. 2017. Open Data in Electoral Administration. International Institute for Democracy and Electoral Assistance.

Cohen, Matt. 2024. "Virginia Judge Orders Waynesboro Officials to Certify Election." Democracy Docket. November 5, 2024. https://www.democracydocket.com/news-alerts/virginia-county-officials-declare-they-will-refuse-to-certify-november-election/

Committee on National Statistics (CNS), Division of Behavioral and Social Sciences and Education, and National Academies of Sciences, Engineering, and Medicine. 2025. Principles and Practices for a Federal Statistical Agency: Eighth Edition. Edited by Melissa Chiu and Jennifer Park. National Academies Press. https://doi.org/10.17226/27934.

Groves, R. M., and L. Lyberg. 2010. "Total Survey Error: Past, Present, and Future." Public Opinion Quarterly 74 (5): 849–79. https://doi.org/10.1093/poq/nfq065.

Guercio, Maria. 2001. "Principles, Methods, and Instruments for the Creation, Preservation, and Use of Archival Records in the Digital Environment." The American Archivist 64 (2): 238–69. https://doi.org/10.17723/aarc.64.2.n88455np210p8j5v.

Hale, Kathleen, Robert Montjoy, and Brown, Mitchell. 2015. Administering Elections: How American Elections Work. Palgrave Macmillan.

Hasen, Richard L. 2012. *The Voting Wars: From Florida 2000 to the Next Election Meltdown*. New Haven: Yale University Press.

ISO/TC 46. 2016. Information and Documentation — Records Management. ISO 15489.

Jacobs, James Asbury, and James R. Jacobs. 2025. Preserving Government Information: Past, Present, and Future. FreeGovInfo Press.

Joint Task Force Transformation Initiative (JTFTI). 2012. *Guide for Conducting Risk Assessments*. NIST SP 800-30r1. 0 ed. National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-30r1.

Madnick, Stuart E., Richard Y. Wang, Yang W. Lee, and Hongwei Zhu. 2009. "Overview and Framework for Data and Information Quality Research." *Journal of Data and Information Quality* 1 (1): 1–22. https://doi.org/10.1145/1515693.1516680.

McDonald, Michael, and Micah Altman. 2018. *The Public Mapping Project: How Public Participation Can Revolutionize Redistricting*. Brown Democracy Medal. Cornell University Press.

McDonald, Michael P. 2022. *From Pandemic to Insurrection: Voting in the 2020 Presidential Election*. Berlin, Germany: DeGruyter.

Parks, Miles. 2022. “Hand-counting ballots may sound nice. It's actually less accurate and more expensive.” NPR. October 7, 2022. https://www.npr.org/2022/10/07/1126796538/voting-explainer-hand-counting-ballots-accuracy-cost

Saltman, R. G. 1975. Effective use of computing technology in vote-tallying. Technical Report NBSIR 75-687, National Bureau of Standards, Washington, DC.

Saltman, Roy G. 2006. The History and Politics of Voting Technology. Palgrave Macmillan US. https://doi.org/10.1057/9781403977212.

Seymour, Jennifer. 2017. “The Modern Records Management Program: An Overview of Electronic Records Management Standards.” Bulletin of the Association for Information Science and Technology 43 (2): 35–39. https://doi.org/10.1002/bul2.2017.1720430212.

Urahn, Susan, and Michael Caudell-Feagan. 2008. Being Online Is Not Enough. Pew Center for the States.

U.S. Alliance for Election Excellence (USAEE). 2025. Standards for Excellence. https://electionexcellence.org/standards-for-excellence/.

U.S. Senate Select Committee on Intelligence. “Russian Targeting of Election Infrastructure During the 2016 Election: Summary of Initial Findings and Recommendations.” May 29, 2018. Available at: https://www.intelligence.senate.gov/2018/05/29/publications-russian-targeting-election-infrastructure-during-2016-election-summary-initial-findings/.