Auditing Elections Using Ballot Images and AuditEngine -- General Background

Raymond Lutz
Executive Director and Lead Developer,
Citizens' Oversight Projects

V2 2022-10-04

Citizens' Oversight Projects
is a 501(c)3 nonprofit and
nonpartisan organization
http://citizensoversight.org

Welcome to AuditEngine

Auditing elections is an essential element in thwarting attacks on our democracy, providing solid evidence that the results are complete, honest, and trustworthy.

AuditEngine, a software platform developed by Citizens' Oversight Projects and led by Raymond Lutz, is an important and powerful tool that can prove or disprove consistency between ballot images created early in the tabulation process and the final tabulation. This consistency is important evidence that our election outcomes are accurate and trustworthy.

Because audits of ballot images are relatively new and the public and officials may be unacquainted with this approach, the following introduction provides general information about ballot image auditing in general and the specific approach provided by AuditEngine, without providing details or findings from any specific audit. Maintaining the background information in this separate document allows us to more briefly describe the details from any specific election while relying on this document for background information and terminology.

AuditEngine is a cloud application that uses the horsepower of up to 10,000 computers in parallel to completely process all the ballot images, taking only minutes instead of weeks.

Because AuditEngine is relatively new, the following sections provide a generic introduction to the process and will defer treatment of specific issues in separate audit reports.

A note on writing style

Throughout this document, we will use "programmer" style quotes, which always frame the terms and do not include punctuation. Numbers are always shown in numerical form and commas will always be included in conjunctive lists.

TABLE OF CONTENTS

1 The Problem

1.1 Our Elections are not Trusted

Far too often, the public finds that election evidence is processed behind closed doors, inside computers controlled by just a handful of for-profit corporations. Paper ballot evidence, if it exists, is often sealed and sequestered and therefore impossible to review by members of the public. This can lead to a loss of confidence in the results, and may lead to unnecessary court cases and unfounded claims of fraud.

Voting System

Before we proceed, it will be important to define a voting system. Such a system consists of typically many voting machines in many individual polling locations and also the central "Election Management System" which handles many aspects of the election, including formatting the ballots and tabulating the results. Polling locations may have nothing more than paper ballots and pens, and a single voter-facing scanner that can give feedback to the voter in terms of overvotes or if the ballot appears to be completely blank. Sometimes, they may have no scanner there and just return all the ballots to the central location for scanning. The voter-facing scanners typically also tabulate the ballots they process and then those results are provided to the EMS by hand-carrying flash memory devices and sometimes additionally cell modem transmission.

Newer systems also may have Ballot Marking Devices (BMDs) which allow the voter to vote using a touch-screen and then print out a ballot summary card which has their votes listed typically also with a QR Code or Barcodes encoding the votes. These devices do not tabulate votes, and instead, the ballot summary cards are scanned by the same scanners that process hand-marked ballots.

Old-school Direct Response Electronic (DRE) machines where there is no paper record are extremely hazardous because there is no paper record and so these are being phased out. For vote-by-mail or absentee voters, the voting system may use a central scanning operation where high-capacity high-speed scanners are used.

High resolution ballot images

In recent years, voting machines for processing paper ballots have evolved into full-page document scanners which create digital images of both sides of every ballot sheet, typically at the resolution of 200 pixels per inch. Voting machines use the ballot image rather than the paper ballot itself to evaluate voter intent by recognizing and interpreting the marks in the image as votes. Thus, because the images are a critical step in the operation of these machines, they must be preserved[1], and are (or should be) considered "public records" and frequently can be provided under freedom of information laws, while still preserving the original paper records.

Since the ballot images are created very early in the process, they can be used to thwart all later fraudulent modifications to the tabulation, and can be a check on the quality of the evaluation of the vote by election processing systems. The images can be secured using digital security mechanisms to ensure that any changes to the ballot images themselves can be detected[2].

There is a razor thin chance that a sophisticated hacker might alter ballot images as they are made. The risk that this remaining hazard exists can be minimized through the use of cybersecurity techniques as well as inspections of ballot images as they are produced on a sampling basis. Indeed, as the sophistication of these controls improves, the ballot image audit can provide the best transparency of the election results to reassure the public while simultaneously cutting off false claims of impropriety.

The tabulation can be audited also using other methods, such as a "full hand-count audit", where all ballots are counted by hand, or a "risk-limiting audit" where batches or individual ballots are randomly sampled to create a check on the result within a specified risk that the sampling may not detect an altered election. These audits can work in concert with ballot image audits. However, there is no audit that has the precision of a ballot image audit because (if the recommended data is provided) a ballot image audit can compare every ballot to the official result, for all contests, regardless of how close the results are. This is very difficult in the case of sampling audits and hand-counts because of the induction of human error in the process and the cost of handling paper.

---

[1] Federal law, 52 USC 20701: "Retention and preservation of records and papers by officers of elections; deposit with custodian; penalty for violation" requires that "Every officer of election shall retain and preserve, for a period of twenty-two months from the date of any general, special, or primary election... all records and papers which come into his possession relating to any application, registration, payment of poll tax, or other act requisite to voting in such election." The U.S. Dept of Justice published further clarification "Jurisdictions must therefore also retain and preserve records created in digital or electronic form." in the document "Federal Law Constraints on Post-Election “Audits”", on July 28, 2021, https://www.justice.gov/opa/press-release/file/1417796/download

[2] Although this will be possible in the future, these security mechanisms have not been standardized nor adopted by the voting machine makers. However, there are standards initiatives in process today that will provide the building blocks to ensure that election data artifacts are fully secured using the most advanced security protocols available.

1.2 Paper Ballots are Essential

Paper ballots are a key component in the improved security of our voting system, and hand-marked paper ballots are the best user interface. They can be easily marked by the voter as that voter also verifies that their vote was recorded properly. This single enhancement in recent years has significantly improved the security of the system over the pure touch-screen interface with no paper record ("DRE" machines) where it is impossible verify that the vote was recorded, and the vote could be modified at will by any insider or hacker with no way to check it at all.

Yet despite the superiority of the hand-marked paper ballot, the paper record is also difficult to fully trust since they can also be easily modified, added, or subtracted, damaged or destroyed. As a result, paper ballots require utmost security of the chain of custody. This security largely depends on election officials' honesty and competence. If ballot images are produced, they can be an important check on the paper, and the paper ballots can be a check on the images. The review of both should be included in any complete auditing program.

1.3 Some aspects are not covered by a tabulation audit

A ballot image audit is the most thorough and intensive check of the tabulation from when the voter first casts their ballot, to the final result, but it does <u>not</u> cover aspects prior to when the ballots are cast, such as the question of eligibility, whether those ballots are all appropriately cast, and whether they represent the complete set of all the ballots cast and no more. These aspects can be checked by other types of audits, including chain of custody, eligibility and procedural compliance audits.

AuditEngine addresses the tabulation portion and does not at this time directly audit voter eligibility or signature validation, although the tabulation results can be compared with the number of voters that are recorded as casting a ballot to verify that the number of ballots matches the number of voters that cast a ballot. This was a specific design decision to avoid any accusation that we might be linking the voter identity to their vote. We don't have the voter's identity, and so it is not possible to link anything in this respect using the data processed by AuditEngine.

1.4 Ballot Image Audits provide extreme scrutiny

Review of the ballot images to create an independent tabulation of the election can detect many types of errors or subsequent modification of the totals, either by honest mistakes, a compromised insider, or intrusion by an outside hacker.

Audit Engine provides a very detailed review of the inventory of ballots in the election which can expose processing mistakes and help election officials and election software vendors to improve their procedures so these mistakes can be avoided in the future. Even when election winners are not in doubt, ballot image audits help to improve the quality of election processing.

2. The Landscape

Today, elections equipment and voting machines are primarily supplied by three top vendors[3]:

1. Election Systems and Software (ES&S) -- This company holds approximately 47% of the market. ES&S serves about 1,668 county-level jurisdictions across the U.S.[4]
2. Dominion Voting Systems -- This company holds about 37% of the market, serving more than 1,300 jurisdictions, including 9 of the largest 20 counties in the USA.[5]
3. Hart Intercivic[6] -- This company has about 10% of the market.

There are also a number of other vendors that have less of a market share, including Clear Ballot, Voting Works, and others.

At this time, AuditEngine is compatible with data files generated by ES&S and Dominion, and we are in the process of adding Hart and other vendors, which should be largely compatible.

We are concerned mainly with hand-marked paper ballots and ballot summary cards produced by Ballot Marking Devices.

Hand-marked paper ballots have ovals or rectangles ("targets") that are darkened with a pen. As the voter completes their ballot, they also simultaneously verify that the marks they are marking are next to the appropriate option. Typically, these ballots are inserted into a voting machine scanner which can perform a quick check for over-votes or other

---

[3] Verified Voting "Verifier" provides 21 pages of results for Dominion and 64 pages of results for ES&S. But it would be more accurate to consider not the number of jurisdictions, but the number of voters using each type of voting system. Unfortunately, that information is not directly available.

[4] https://www.essvote.com/faqs/

[5] https://www.dominionvoting.com/about/

[6] https://www.hartintercivic.com/

mistakes, and alert the voter to provide the opportunity to correct those issues before casting the ballot. These ballots are also commonly used for vote-by-mail voting.

Ballot Marking Devices (BMDs) are touch-screen machines that produce ballot summary cards that list the votes cast by the voter for each contest, and which can be verified by the voter. The BMD ballot summary typically also includes linear or 2-D barcodes identifying the vote for machine interpretation. Voting systems generally only read the barcodes and not the human-readable summary, and the barcodes are difficult or impossible to be checked by the voter. The BMD machines do not tabulate the ballots they produce.

Unlike hand-marked paper ballots, BMDs do not provide simultaneous marking and verification, and the resulting ballot summary cards are rarely adequately verified by voters, missing altered ballot selections 93% of the time in university studies[7]. And since voting tabulators typically only read the barcodes, there is a chance that the BMD device could be altered or designed so that the barcodes could specify one candidate while the written text could indicate another.

Alex Halderman, a computer science professor at the University of Michigan exposed this flaw in the Dominion, and frankly all BMD voting systems:

> Halderman found that malicious software could be installed on voting touchscreens so that votes are changed in QR codes printed on paper ballots, which are then scanned to record votes, according to court documents. QR codes aren't readable by the human eye, and voters have no way to know whether they match the printed text of their choices[8].

Audit Engine uses optical character recognition (OCR) to "read" the human readable text and generate an independent cast vote record of each and every BMD ballot, and then compare with the official result, rather than depend on the barcodes, as the voting system does. AuditEngine will therefore catch the exploit described by Halderman, as long as the text is what the voter intended.

If the text is also changed from the selections of the voter, then it can't be detected except by the voter, and as mentioned, voters usually don't review their ballot summary

---

7 https://news.engin.umich.edu/2020/01/new-study-finds-voters-not-detecting-ballot-errors-potential-hacks/ Matt Bernhard and J. Alex Halderman found that "voters missed over 93% of errors on printed ballots that they filled out using electronic ballot marking devices."

8 "Secret report finds flaw in Georgia voting system, but state in the dark" https://www.ajc.com/politics/secret-report-on-georgia-voting-system-finds-flaws-but-state-shows-no-interest/YKFEET2WE5BBPJ7TYVOYMBTIKQ/

cards to verify that the text matches their intent. In contrast, hand-marked paper ballots don't separate the verification from the marking steps and they don't have this problem.

The education of voters to review their ballot summary cards, or BMD devices adding a "directed review"[9] of the readable text, would improve the confidence of auditors that the ballot summary cards reflect voter intent.

After refinement of our BMD parsing algorithm during our audit of Bartow County, GA, AuditEngine parsed 100% of human-readable text on the 39,967 BMD ballots, with very few cases (1) when the text could not be read and compared with the official records. The very clean and complete data set provided by Bartow officials also facilitated refining AuditEngine to accommodate prospective audits in other Dominion jurisdictions.

3. The Data

3.1 Ballot Images

Modern voting equipment used in most districts in the U.S. create ballot images[10] as they process the ballots. Of the ten battleground states in the 2020 General Election, our review indicates that 93% of the most populous jurisdictions (accounting for 80% of the electorate in each state) use voting equipment that creates ballot images as the ballots are scanned. Thus, the use of ballot images is now widespread and growing.

3.1.1 Ballot Images are Essential Data

Ballot Image data is the first digital representation of a hand-marked paper ballot or BMD ballot summary card. It includes the maximum amount of data, including whether the ovals were completely marked, checked, circled, or any number of other possibilities. Once the vote is extracted, the data is "compressed" to a maximum extent, and the compression is considered "lossy", because it is not possible to reconstruct the ballot image from the summary of the vote.

It is essential, therefore, that the election equipment be configured, and procedures established, so ballot images are not deleted. Election departments should preserve

---

[9] Such a directed review would interact with the voter via the touch screen to have them check the printed ballot summary card and compare it with what they chose during their private voting session. This would eliminate the risk that the human-readable text does not match what the voter intended, but obviously would slow down the voting process.

[10] The term "ballot image" was initially used to refer to the digital summary of a DRE (Direct Recording Electronic) AKA "touch-screen" voting machine, and was not an actual picture of a hand-marked paper ballot, which is how the term is now used. This has been reflected in the latest Voluntary Voting Systems Guidelines (VVSG) 2.0 now being processed by the Election Assistance Commission.

and routinely publish ballot images for all elections. This issue is particularly relevant to ES&S equipment which can be set to delete the ballot images. Sometimes running a subsequent election may overwrite the data from a prior election.

It is federal law that election data be preserved for 22 months after every election. The U.S. Department of Justice clarified this to include digital and electronic data[11]. This certainly includes ballot images, which are extremely important because they are the first instance of digital data produced by voting systems.

3.1.2 Privacy and Ballot Anonymity

Ballot images are anonymous[12]. There is no way to determine the identity of any voter by looking at the ballot images alone, and any names or personal information that may be seen (and improperly added to the ballot) cannot be proven to be of the person who actually voted that ballot. However, there are sometimes cases when voters use special ballots, such as federal-contests-only ballots, when there are so few of these that by triangulating with other data, someone could match up votes with voters in these rare cases.

The privacy of other types of data, such as medical data, can be anonymized by adding data to the set that does not change the overall statistics, while making it impossible to pair up names with medical history. However, in the case of election data, that approach cannot be taken because adding ballots to the set creates another opportunity for manipulation, although some jurisdictions say they might do so[13].

Although jurisdictions might be tempted to modify the election data, we must insist that the exact results are published.

---

[11] The U.S. Dept. of Justice published further clarification "Jurisdictions must therefore also retain and preserve records created in digital or electronic form." in the document "Federal Law Constraints on Post-Election “Audits”", on July 28, 2021, https://www.justice.gov/opa/press-release/file/1417796/download

[12] Some states still may add a unique serial number to ballots, supposedly so they can retrieve any ballot should a voter die prior to election day but after they have already cast an early ballot. This practice should cease because it means those ballots are not anonymous and any unscrupulous election official may pair up the exact votes of any voter.

[13] NC says: "Why do the vote counts in the Precinct Sorted Data files differ slightly from those in the Election Results files? To protect the secrecy of voters’ ballots, statistical noise has been added to the precinct sorted results files. The process of adding statistical noise is a common practice among statisticians, demographers, and other organizations working with datasets containing sensitive and/or confidential information. The addition of statistical noise to the precinct sorted data ensures that, when joined to other datasets, the precinct sorted data cannot reveal the ballot choices of any voter. Due to the addition of statistical noise, the vote counts in the precinct data will differ slightly from those in the Election Results. As such, the precinct sorted data should not be considered a source for reporting official election results." https://www.ncsbe.gov/results-data/election-results/historical-election-results-data

It is therefore our recommendation that rare ballot styles are minimized, and when there are ballots that fall into rare styles, that these ballots can be omitted from the ballot image set without seriously impacting the results of the audit. Also, we recommend that it be considered a crime to match up and use or publish the vote of any voter with their identity.

AuditEngine specifically has been designed to exclude voter identification information so it is not possible to link votes with voters, and as a consequence, it does not provide audits of voter eligibility. The information required to link any ballot to a voter is not utilized by our application and thus cannot link them through any mechanisms such as rare styles. However, we do a rare style analysis to determine if there is any widespread risk of voter privacy intrusion. Thus far, it is quite infrequent to find cases when a voter might be matched up to their voted ballot, and certainly, there is a very low risk that an arbitrary voter could be linked to their ballot.

Another fact that preserves privacy in these cases, AuditEngine does not rely on human review of individual ballot images for the vast majority of ballots. We have conducted a number of audits in various counties and since we have no access to voter identification information, there is no risk of exposure by using AuditEngine alone. The actual ballots required to be reviewed due to a disagreement between the evaluation of voter intent by AuditEngine and the voting system are usually less than 1% of the ballots.

However, we do allow any ballot to be fully researched by our system, so as to provide every bit of data we process, and thus you can check that AuditEngine is producing honest results. The fact is that it will find a number of discrepancies in every election, and with this level of detailed results, it certainly provides a level of confidence. Yet, like in science, the only way to trust results is to have others replicate the analysis.

3.1.3 Ballot Image Security

Fundamental to any ballot image audit is the assumption that the ballot images are a faithful representation of the ballots themselves. There exists a very small chance that ballot images might be modified as they are being produced[14]. Although this hazard cannot be eliminated completely, the risk that it can occur can be minimized to the point that it can effectively be neglected.

[14] See "UnclearBallot: Automated Ballot Image Manipulation" -- Matthew Bernhard, Kartikeya Kandula, Jeremy Wink, and J. Alex Halderman, Department of Electrical Engineering and Computer Science, University of Michigan https://mbernhard.com/papers/unclearballot.pdf and our response: https://copswiki.org/Common/M1976

We have proposed a method where the scanning equipment will cryptographically secure the ballot images to lock-out any changes as soon as they are created by calculating the SHA[15] code of each image and signing it using the embedded private key in the ballot image scanner, so that it is infeasible to change any ballot image without detection. Publishing SHA codes of the ballot images enables recipients of the ballot images to recalculate the supplied hash code and therefore confirm that no alteration occurred enroute, while not actually revealing the ballot images themselves until after the election.[16]

Ballot images are transferred to the "Election Management System" (EMS), typically during the processing of the election by officials. The EMS is a self-contained, air-gapped system that performs many aspects of election processing and is provided by the voting machine vendor.

Scanners that are located in polling places scan the paper ballots to create images, capture votes from the ballot images, and then the result is transferred to the election office. The transfer sometimes uses a communication link (such as a cellular modem) to transmit the totals of each contest from that location to the central office so as to allow rapid reporting to the media on election night[17]. Flash drives are used to transfer the details of the results, including ballot images, by transporting these with the paper ballots back to the central office. The flash drives are paired with each scanner and include configuration information to allow the scanners to properly interpret the vote.

Vote-by-mail (VBM, AKA "absentee") ballots and sometimes some or all in-person ballots are scanned in centralized scanning operations. For such central-count scanning, the images are typically stored in large (maybe 1TB or larger) internal hard-disk drives in the scanner, and may be transferred to the EMS using a local-area

---

[15] SHA means secure hash algorithm, which creates a distinctive bit sequence for any given ballot image which will likely differ from any other ballot image. There are different varieties, such as SHA-256, SHA-512, etc.

[16] Work is proceeding at the Internet Engineering Task Force (IETF) "Supply Chain Integrity, Transparency, and Trust" working group (SCITT), for general software security and other applications. It will support the secure submission of certificates to transparency services so that no alteration of ballot images (or other data) will be possible without the possibility of detection, and the software in voting systems can be similarly tracked. We will be following and contributing to this effort.

[17] Cellular modems are generally not used to transmit ballot images. Typical cellular data transmission rates range from 100 KB/s to 1MB/s. A typical precinct voting machine with perhaps 1,000 ballots would take about 5 minutes to transmit typical ballot images of 300KB each. So it is not unreasonable, particularly as transmission rates continue to increase, that the ballot images could be transmitted by cellular modem while also still transporting the flash drive media to double check that the images were correctly transmitted. Cellular modem connections are generally feared but can be secured with robust handshakes and encryption.

network which does not connect to the Internet or using "sneaker net", or hand-carrying flash drives[18].

Once the election is complete, or perhaps as each phase is completed prior to certification, the EMS can export the ballot images for use by AuditEngine or other reviewers. These should be placed into ZIP archives[19] with about 10,000 to 50,000 ballots per archive, for ease of handling. AuditEngine accepts image formats as provided by the leading vendors, including PDF[20], TIFF[21], PNG[22], PBM[23], and JPG[24].

In their most recent voting systems deployed, Dominion uses the TIFF image format and ES&S uses PDF files. These files consume about 200KB to 600KB to store both front and back of a single ballot sheet. BMD ballots may take substantially less space. We have found that avoiding gray backgrounds on the ballot can substantially improve legibility while also reducing the size of the resulting image files. They are easier to record to smaller flash media, being likely half or one-third the size of those with gray backgrounds.

3.2 Cast Vote Record Files

3.2.1 CVR Files in General

The purpose of an audit is to check the official outcome of the election. The official outcome can be provided at different levels of granularity. AuditEngine can provide the most detailed reports if the official result from every ballot is available, and it can be matched up to the ballot image. The result of the interpretation of the votes on one ballot sheet is called the "Cast Vote Record" (CVR) for that ballot sheet, and the set of all CVRs are combined into a set of files (CVR files).

---

[18] https://www.wsj.com/articles/sneakernet-helps-election-officials-process-results-11604440573 "'Sneakernet' Helps Election Officials Process Results -- With the internet representing a cyber risk, hand-carrying voting data is often the norm" The Wall Street Journal, 2020-11-03

[19] https://en.wikipedia.org/wiki/ZIP_(file_format)

[20] https://en.wikipedia.org/wiki/PDF

[21] https://en.wikipedia.org/wiki/TIFF

[22] https://en.wikipedia.org/wiki/Portable_Network_Graphics

[23] https://en.wikipedia.org/wiki/Netpbm#File_formats -- used only by legacy ES&S machines

[24] https://en.wikipedia.org/wiki/JPEG

The CVR files must be machine readable to avoid errors that might otherwise be introduced[25]. Adding up the votes for each contest on each of these records should produce counts that are (nearly) identical to the officially reported outcomes.

If CVR files that allow a ballot-by-ballot comparison are not available, AuditEngine can still produce an independent high-level or precinct level result which can be compared to the official outcome at the same level of granularity.[26] This will unfortunately provide almost no diagnostic information that is truly valuable for improving election department procedures and voting systems.

Frequently, CVR files are provided in spreadsheet format, where each row is a ballot sheet, front and back, and each column is the contests and options (this has been typical for ES&S). In other cases, they may be provided as JSON[27] or XML[28] format or CSV[29] (character separated values) files. Dominion uses the JSON format which includes both the 'original' interpretation and also the 'modified' adjudication, if the ballot was adjudicated and a change occurred to that ballot by the adjudication.

The National Institute of Standards and Technology (NIST) has recently promoted a "Common Data Format"[30] that may provide some standardization in the future, but at present, we accept the proprietary formats and will strive to read any reasonable file format. The JSON format as used by Dominion is similar but not exactly the NIST standardized format. At this point, it is the only vendor offering something similar to the NIST standard.

---

[25] In older releases of the Dominion EMS software, it provides the ballot-level official interpretation only in a graphical image called the "AuditMark" as the third page in the ballot image file which is NOT machine readable without performing an error-prone OCR (optical character recognition, i.e "computer reading") conversion, and is therefore not very useful due to the errors that are induced with that conversion. The newer releases now provide the CVR files that do not have this source of error, but they still provide the AuditMark image.

[26] AuditEngine is the only ballot image auditing system that provides ballot-level comparison with the cast vote records. Other systems, like Clear Ballot as used in Florida, and many counties in NY and other places, rescans the ballots and processes those without using the voting system data. This does have the benefit of not relying on the scans by the voting system but means that the new scans cannot be compared with the voting system result on a ballot-by-ballot basis, and can only be compared on an aggregated basis (like by batch or precinct) which drastically reduces the diagnostic power of the audit, because if the aggregation differs, it is hard to find the exact ballot sheet(s) that caused the difference.

[27] https://en.wikipedia.org/wiki/JSON

[28] https://en.wikipedia.org/wiki/XML

[29] https://en.wikipedia.org/wiki/Comma-separated_values

[30] https://www.nist.gov/publications/cast-vote-records-common-data-format-specification-version-10

Note: The term "Cast Vote Record" can be the source of some confusion because of historical ambiguity. ES&S uses the term to refer to both spreadsheets format, which includes all ballot sheets, and also PDF files which can be produced with the ballot images. One ES&S CVR PDF file summarizes in readable format the voting-system interpretation of one ballot sheet. Dominion has a similar summary they call the "AuditMark"[31] which is included as the third page in the ballot image file itself. AuditEngine does not use these as the primary CVR source because using files which consolidate all the records is more secure. The reason is that it is much harder to modify, add, or subtract one or a few ballots using consolidated records.

To disambiguate this term, we will always say "CVR Files" to indicate the "spreadsheet" or database files (typically XLSX[32] or CSV, or structured JSON) that provide the interpretation of the voter intent on the ballots in a machine readable format.

There are a number of essential fields used in ballot image auditing and may be included in the CVR, perhaps using vendor-specific nomenclature:

1. **ballotid** -- a unique identifier of ballot image, corresponding to one ballot sheet, including two images for the front and back. Sometimes, the ballotid is imprinted on the corresponding paper ballot and it also appears in the ballot image. Generally, there are no identifying numbers preprinted on a ballot. If the identifier is imprinted on the paper ballot when it is scanned, the identifier will appear in the image and can be used to marry the ballot image to its physical paper ballot. Anyone who wants to later confirm that the ballot image is a faithful representation of the physical ballot can, in theory, find the physical ballot and inspect it.
2. **ballot_style** -- a numeric identifier for the style of ballot, which indicates the contests, language, and option order on that ballot sheet. It is helpful to have the style of the ballot also listed in the cast vote record for each ballot. The *ballot_style* value should have a direct relationship to the code encoded on the ballot using a graphical barcode. The style of the ballot will be covered in more detail in a later section of this document.

[31] https://www.votescount.us/Portals/16/New%20voting%20system/AuditMark%20Brochure%20-%20final.pdf AuditMark is a registered trademark of Dominion Voting Systems.

[32] https://en.wikipedia.org/wiki/Office_Open_XML

3. **precinct** -- the precinct to which ballots refer, and typically represents the geographic area where the voter resides and includes the set of contests appropriate for all the voters in that region. This is necessary to allow comparison with other reports that group the results by precinct, particularly if the CVR is not provided at ballot-level detail. Sometimes this is provided in the ballot image archives as a component of the filename. To enhance voter privacy, precincts should be as large as possible.
4. **batch** -- Election-day ballots are sometimes automatically sorted into precincts due to the fact that voters go to their precinct to vote. But some areas use voting-centers where they will service any voter, and when ballots are received as vote-by-mail or absentee ballots, the ballots may be grouped into mixed-precinct batches. It is best to also denote the batch so the paper ballots can be located and reviewed if desired. Unfortunately, the batch report is frequently not provided. Dominion, however, does provide this information and it is a big advantage over ES&S CVR data.

3.3 Other Configuration Files

Other files are required for configuration of AuditEngine depending on the strategy for mapping, and the extent that the files, such as CVR Files, are available and have consistent contest and option names.

3.3.1 Election Information File (EIF)

The EIF provides contest information and the options, such as candidate names or yes/no, as would appear a) in the CVR, b) in other reports, and c) on ballots; the number of votes allowed in the contest; how many write-in lines are there, etc. We also need to know the text that is actually used on BMD ballots to facilitate optical character recognition (OCR) reading.

This file is required regardless of whether we have CVR Files or not. When we do have CVR Files, the EIF can be largely automatically generated, and then amended by hand in the configuration process.

3.3.2 Target Map

For the purposes of evaluating the voter's marks on a hand-marked paper ballot, it is necessary to know the location of each "target" i.e. the oval or rectangle that corresponds to each option of a given contest. To vote and select the option, the voter darkens the target with a pen. Establishing the correspondence of target location and the contest option is the mapping process.

If we can get the target map from the "Election Management System" (EMS), this can reduce the work involved in generating the map using the interactive tools provided by AuditEngine. However, there are no established standards among voting equipment vendors for providing this data. Even if we could get this information, we would still need to review and verify it, which amounts to almost the same amount of work in small elections as mapping it using our tools.

AuditEngine has substantial functionality to assist in the creation of the Target Map, which will be discussed in more detail later. In fully independent mapping, the ballot images are used as the source data along with the TargetMapper app, to assist in the process of getting all the styles accurately mapped. Another approach uses the Ballot Style Masters, which are PDF files that were generated by the voting system to be sent to the printing subcontractor to print hand-marked paper ballots, and must be in "searchable" PDF format. These also are processed using the TargetMapper app, and if available, the work is substantially reduced to just the number of contests that have different variations (language or order on the ballot) rather than the (number of styles) x (the number of contests) on each style.

3.3.4 Job Settings File

Each election also has a number of specific settings that include characteristics of the district (population, registered voters, etc mostly for reports) and settings to allow the audit to complete, such as the names of ballot image archives and CVR file names, special case overrides and the like. This file is largely consistent for each vendor and generation of their software.

AuditEngine has a user-friendly "frontend" that runs in a browser and allows the audit to be processed in a series of "stages" in a "audit pipeline".

4. The Process

This section provides a brief overview of the AuditEngine process. The entire process can be reduced to the following major phases:

SETUP and CHECK DATA: Gather the input data, create job control files, capture metadata, and create initial reports. This step may include an initial full automated review of ballot images to extract style information and other metadata.

MAP STYLES: The mapping process determines the location of all the targets on each ballot style. AuditEngine uses browser-based utilities to assist the audit team to create an accurate map.

EXTRACT: This is the main phase where all ballots are reviewed and the vote is evaluated on each ballot without reference to the official results and without relying on any software or hardware component of the voting system.

COMPARE and REPORT: Finally, the result of the evaluation by AuditEngine is compared with the official result, preferably ballot-by-ballot, and a report is created which details any discrepancies located.

4.1 Step-by-step Overview

Here, we will briefly describe the steps in slightly more detail.

Setup and Check Data

Step	Description
Create Election	The first step is to create the election in AuditEngine. We now offer the ability to create a single uploading link for any jurisdiction that will remain the same over many elections. After the election is created, then data can be uploaded to the cloud as described below.
Gather Input Data	AuditEngine needs ballot image data and CVR files. Ballot image files and CVR files can be uploaded directly to AuditEngine's datacenter in the cloud. If the CVR files are not available, then the complete list of contests and options as well as the official results must be included in configuration files. Ballot Style Masters files are desired to allow for automated target mapping. Information about how to generate these files, exporting and uploading can be found here: https://auditengine.org/user-guide/exporting_guide/
Audit Job and Job Settings	After data has been uploaded, an Audit Job can be created that will use the data uploaded. From within the election screen, click "Audits" and then [New Audit Job]. If the Ballot Images, Cast Vote Records, and Ballot Style Masters are already uploaded, the job settings file can be created by clicking the "Create Job File" button. If no Job File exists, then it will be created based on the vendor specified, with the appropriate settings initialized to refer to the files uploaded to the election. A series of settings will require input, such as total recent population of the district, registered voters, etc. (This process is still being optimized and may require

	hand-holding at first).
Precheck	The very first active stage in the audit process is to precheck the files as referenced in the Job Settings file. This function is very quick as it does not process these files beyond listing them and providing the hash value for each file. This hash value can be compared with the value generated independently for each file to ensure that the file is identical with the official file. (We use hash values provided by the cloud storage system so the files need not be downloaded to check them).
Capture Metadata	The first few steps of the processing pipeline extracts the metadata from various sources and combines it into a uniform record for each ballot. In each case, the metadata listed for each ballot is called the "bif" (ballot information file). **biabif** -- this is the metadata that is extracted from the ballot image archives. The biabif report also provides an analysis of the file sizes. ES&S frequently provides metadata such as precinct, group (early vote vs. election day) and the like in the structure of the files in the zip archive rather than providing in a more extensive CVR as does Dominion. **cvrbif** -- this is the metadata extracted from the cast-vote record files. **bltbif** -- this is the metadata extracted from the ballot images in an automated review, extracting the ballot style (card_code) which is typically encoded into proprietary barcodes printed around the edges of the ballot. This pass does not attempt to read the vote, but it may read other metadata, such as the county name, precinct designations, and the like, based on the settings established for the job. The next stage simply joins the separately generated metadata. It is generally our policy to use the card_code style indication directly from the ballot rather than depending on the style indication that may be in the CVR, because those may be incorrect. The one exception for this is when we process BMD ballots, because the style indication is encoded in the QR Code, which we generally avoid reading.
BIF Report	The Metadata report (BIF Report) provides a set of reports on aspects of the ballot image and CVR set. We have found this is an important step because of frequent differences in the data provided from the various sources. It reports on missing, repeated, or extra ballot images or CVR records. It also provides reports to help analyze the ballot style scheme used.
Derive EIF	If the CVR is available, it can be used to largely generate the Election

	Information File (EIF) which provides the list of contests and the options related to each contest. In addition to what can be automatically derived from the CVR, it will be usually necessary to add information to the EIF which are not found in the normal reports but are needed for AuditEngine to run the audit.

Map Styles

Step	**Description**
Derive style templates	Templates are created for each style identified. We create the style templates from the ballot images themselves.
Map Targets	It is necessary to determine, for each ballot style, the location of each contest option target that could be darkened if the voter selects that option. AuditEngine provides computer-assisted mapping where a graphical user interface (GUI) is used by a human operator to associate the targets on the ballot images with the contest options in the election. There are two options: 1. **Use Ballot-Style Masters**, which would normally be sent to the printing subcontractor and must be "searchable" PDFs of each ballot style. Our mapping software requires that a human operator match up the contests as discovered on the ballots with the official contest and option names, which must be done once for every contest. 2. **Independent (Computer Assisted) Mapping** does not require the Ballot Style Masters and instead uses only the ballot templates created from the ballot image data. This is more work, requiring a separate step for each style instead of only for each contest, but is fully independent and will always complete, but may be significantly more difficult approximately by the factor of the number of styles.
Create Redline Proofs	Redline proofs of the mapping are prepared that show the targets outlined and annotated so they can be easily reviewed. Mapping is a critical step because an error in the map can significantly alter the result. However, if there is an error in the map, it will normally be obvious later in the process.

Extract

Step	**Description**

Vote Extraction	With the mapping information approved, the complete ballot image set can be processed. Each ballot image is read from the ZIP archives, is aligned and trimmed, the darkness of each target is evaluated, and the vote is determined. AuditEngine uses an adaptive thresholding algorithm to adjust to the habits of each voter and the density of the image. BMD ballots are evaluated by "reading" the text on the ballot rather than reading the QR/barcodes. Because AuditEngine runs in the cloud, up to 10,000 virtual computers can be used in parallel to reduce the overall time for extraction to less than 15 minutes for up to 2 million images. The same process would take weeks or months using a single computer.
Extraction Report	This report documents the extraction process itself, the number of marks detected, and the distribution of those marks.
Summary Results Report	The report of the extraction provides the overall statistics for each contest. This can be compared with the overall outcome of the election.

Compare and Report

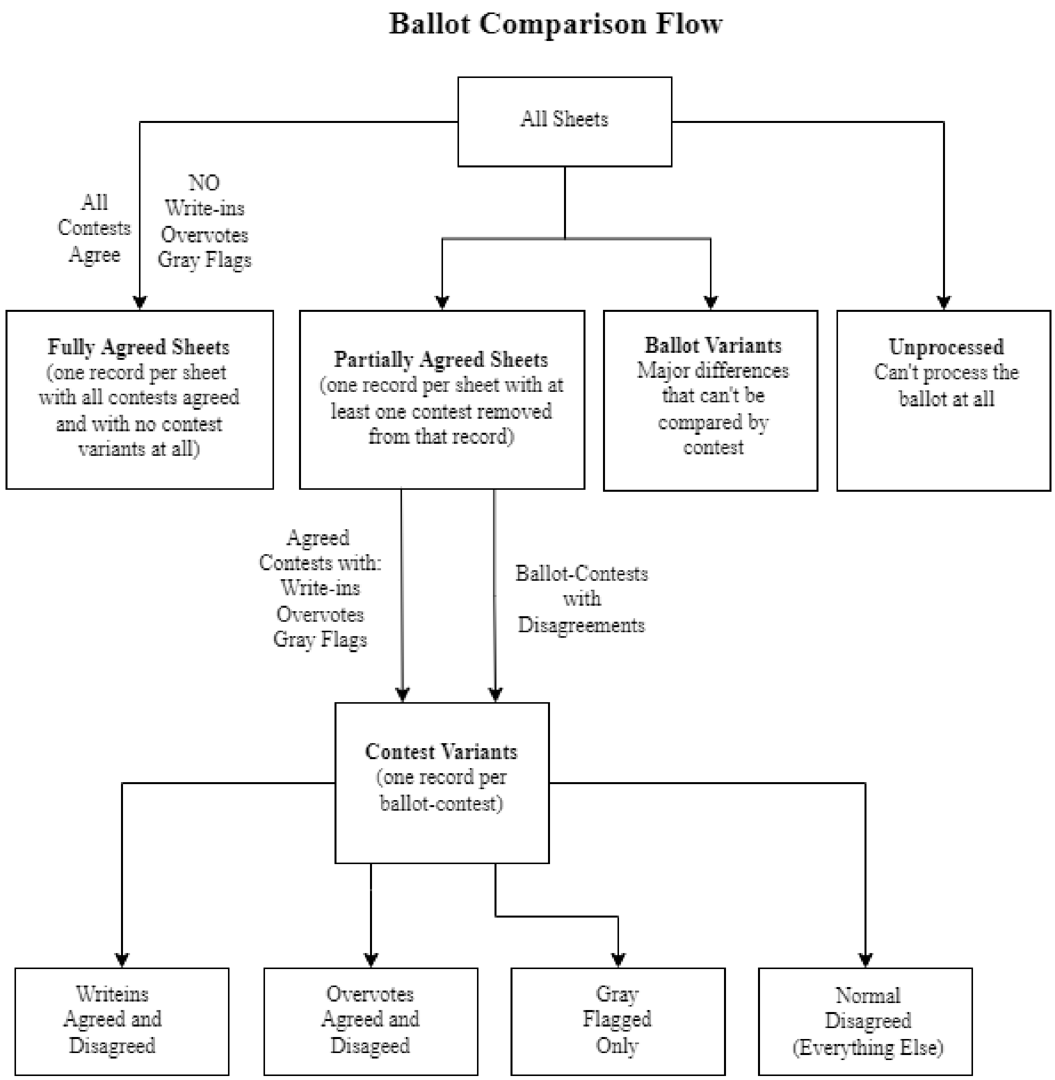
Step	**Description**
Compare with CVR	If the cast-vote record (CVR) files are available with a record for each ballot, then the result of the vote extraction can be compared with the official result on a ballot-by-ballot basis. This is also performed in parallel by up to 10,000 computers in the cloud.
Discrepancy Report	The result of this step is a Discrepancy Report, where the number of disagreements between the audit result and the official result is detailed, and the image of the ballots with discrepancies are provided for review. If there is no CVR, this report still details the ballots that were considered marginal or "gray-flagged" and can help to direct further image review to produce a very precise result. This is a very long report which includes information on each case of disagreement or ballots that may need to be reviewed by campaigns to decide if further review is required. The report may have very many pages (perhaps 1000s) and thus it is not appropriate to be printed. If there is specific interest in a particular contest, an in-depth report can be included for that contest. Default settings include details on the first 10 contests, the 5 closest contests (in terms of %margin of victory), and the 5

	most variant and most disagreed contests.
Final Report	A final report enumerates all the other reports. It is also appropriate to produce a narrative report including an executive summary pointing out any issues found in the audit that go beyond the automated reports.
Adjudication	An adjudication application is available to allow users to review any ballot images that deserve further review so as to refine the evaluation of the vote. Overvotes, writeins, undervotes, and gray-flags can be reviewed to make sure the results of AuditEngine are as accurate as possible.

5. Discrepancy Report

The most important report is the Discrepancy Report as it is the ultimate result of a ballot image audit.

One of the key differences between ballot image audits, most specifically with the analysis provided by AuditEngine,and other types of auditing is that the comparison between the results from AuditEngine and the official results occurs for all ballots at the ballot level.



To be able to describe the differences, we have a number of terms that have been developed, as follows:

1. **Sheets** -- Each ballot image comprises a single paper sheet, including both sides. This is the case even if multiple sheets are included in the physical ballot for a single voter. (The number of sheets in tables is not necessarily a direct sum because a single sheet may include multiple contests in different categories.)
2. **Contest or Ballot-Contest** -- On each sheet, there are a number of contests. A single contest on a single ballot is a "ballot-contest" and sometimes just "contest" in this context. This is not the entire contest with all votes from all ballots summed, but simply the evaluation of that single contest on that sheet. Frequently, a ballot contest is called a "vote" (except that sometimes a single ballot-contest will allow multiple votes, when the voter can vote for more than 1).
3. **Gray-Flags --** If there were any ambiguous marks or if AuditEngine used heuristics or machine-learning to evaluate a contest, these ballot-contests are "gray-flagged" to allow AuditEngine's best guess to be later evaluated by human-eye in close contests. Gray Flags are not reliant on the comparison with the voting system results but are independently determined by AuditEngine.

4. **Write-in** -- If the write-in oval is marked or if the write-in area is filled in, then it is considered that the voter intent is to write-in the name for some other candidate.

   How this is handled is specific to different states, but most often, the written-in name must also be from a list of qualified write-in candidates.

   Write-ins on BMD ballots do not need human-eye adjudication at this stage because the names are keyed-in.

   The name written-in should be examined in close contests even if there were no candidates that were qualified for this contest, because the name written-in may be of a listed candidate, and if so, then the vote is accepted for that listed candidate. At this stage, most jurisdictions do not fully filter whether the name written-in is a qualified write-in candidate, but may do so during adjudication.

   Write-ins normally account for at least 80% of Contest Variants.
5. **Overvotes** -- One overvote occurs when more than the "vote-for" number are selected. Overvotes do not happen on BMD ballots. Very frequently, overvotes are misinterpreted by the voting system and should be fully reviewed in close contests. If a contest is classified as an overvote, it will not also be marked as a write-in, even if the write-in is selected in the CVR. This is a minor shortcoming of the CVR formats being used.

6. **Undervotes** -- Undervotes occur very commonly when voters choose to leave a contest blank or select fewer options than the "vote-for" number. AuditEngine does not treat every undervote as a "variant contest" or there would be far more false positives than necessary. Therefore, it creates a contest-variant record only when the evaluation by the voting system (the official cast-vote record) and the evaluation by AuditEngine disagree, or if they are "gray-flagged".

   AuditEngine will correctly interpret many marks that would be considered undervotes by voting systems, such as when ovals are circled or if a checkmark does not actually get inside the oval.

   Undervotes on nonBMD ballots in critical contests may be of interest for additional review. These are not included in contest variants in this version of the Discrepancy Report but will be added as an option in the future for contests that are marked as "critical".

   It is sometimes a jurisdiction's operating procedure to mark confirmed overvotes as undervotes. If this is the case, all these cases will already be included as contest variants, but will obviously mis-report the number of overvotes in statistics.

7. **Adjudicated** -- For some voting systems, the CVR may also provide not only the 'original' evaluation of the ballot sheet, but also a 'modified' record which includes a new evaluation for all contests on that sheet. The 'modified' record may modify zero or many contests after human-eye review by election staff. If the modified record does not show a change in a given contest, it is not possible to know from that record if the contest was inspected and confirmed or not. The Dominion Voting System supports modified records (if the module is purchased for the software) but ES&S does not have any integrated adjudication solution, and their CVR does not have modified records.

8. **Disagreed** -- a contest is considered "disagreed" unless the voting system and AuditEngine fully agree on the outcome, including whether it was overvoted, undervoted, or had write-ins.

9. **Fully Agreed Sheets** -- A ballot sheet is classified as Fully Agreed if it has no write-ins, overvotes, and gray-flags, and it also every contest fully agrees between the evaluation of the voting system and AuditEngine. Fully Agreed sheets are maintained as a single record but AuditEngine keeps track of the contests on each ballot for summations.

10. **Partially Agreed Sheets** -- If a ballot sheet has at least one contest-variant detected, then that contest variant is logically pulled from this comparison record, and what remains are the rest of the non-variant contests on that sheet. Note that if all contests are "pulled" and contest-variant records are created for each one, then a Partially Agreed Sheet record will still exist but it will have no contests left.

11. **Contest Variant** -- A contest variant is a ballot-contest which has write-ins, overvotes, is gray-flagged, or if it is 'disagreed', i.e. if there is any disagreement between the evaluation of the vote by AuditEngine and the official result. Undervotes are generally not considered variants unless they are 'disagreed' or in critical contests. Contest Variants in tables can normally be summed by contest, because each record concerns only a single contest. (In contrast, sheets cannot always be summed because one sheet may have multiple contest variants, and the same sheet may appear in multiple categories.)

12. **Ballot Variant** -- If a complete sheet is blank, if the image was detected as corrupted, or if the contests included differ from what was expected for that style, then this is classified as a ballot-variant.

13. **Unprocessed** -- Unprocessed sheets are those that were NOT successfully processed by AuditEngine, but may have been processed successfully by the voting system. These are generally due to images that are damaged due to 'stretching' of the image when the page is not evenly fed through the scanner, barcodes that were corrupted and unreliable, or BMD ballots that were not perfectly read using OCR. Cleaning the scanner rollers can help to avoid corrupted images. AuditEngine does not track contests on unprocessed sheets.

    Note, if possible, unprocessed are tracked as Ballot Variants and this category may be empty.

14. **Images Missing** -- Sometimes we can't access all the images and so these are accounted for as images missing. AuditEngine does not attempt to sum the contests on ballot images that are missing.

Thus, we have four major types of records. The first three have one record per ballot sheet, and the final one is one record per ballot-contest.

- **Fully Agreed:** One record per sheet which has all the contests on that sheet, when they are Fully Agreed and there are no contest variants at all.

- **Partially Agreed:** One record per sheet that includes the contests that have not been pulled and "moved" to the contest-variant list. Those that are "pulled" are moved because they are classified as "disagreed", or have overvotes, write-ins, or gray-flags.

- **Ballot-Variants:** One record per sheet that is completely blank or is corrupted. These may be partially included in the Fully Agreed set (when they are agreed and fully blank). It would be better to call these "sheet variants".

- **Contest-Variants:** A separate record for each ballot-contest that is disagreed, or has overvotes, write-ins, or gray-flags.

The Discrepancy Report includes columns for both Sheets and Contests. Please note that a given sheet may include several variant contests, and one sheet may be included in different categories. Thus, the total number of sheets is not always a direct sum of the sheets in each category, and in some cases it is listed as a direct sum and that is only an estimate rather than an exact value.

These comparison records can then be used to produce the discrepancy report.

5.1 Contest Discrepancy Report

The most effective report is the Contest Discrepancy Report because the disagreements can be related to the margin of victory in a specific contest.

For any particular contest, we can focus on the "Disagreed% of Margin" or the "Variant% of Margin". The margin of victory, in votes, for the contest is between the last-winning candidate and the first-losing candidate. This is not the "pairwise" margin, but the actual margin including all other candidates. For ease of reading, the closest 5 contests are highlighted in terms of the Disagreed% of Margin and Variant% of Margin, and also contests with margins of victory below 10% are highlighted.

The following is an example of the first few lines of the Contest Discrepancy Report Table. You can click on the contest name (in the actual discrepancy report, not in this narrative report) to review the details of the contest. "NonVariant" cases are fully agreed without overvotes, write-ins or gray-flags. If you add the highlighted columns, the sum is shown in "All Variants" column. "Disagreed% of Margin" and "Variant% of Margin" can be used to see which contests had the most disagreements and contest variants when compared with the vote margin in the contest.

Contest	Total	NonVariant	Agreed Overvotes	Agreed Writeins	Disagreed	No CVR	Gray Only	All Variants	Disagreed% of Margin	Variant% of Margin	Vote Margin	Margin%
President of the United States	50,677	50,517	11	136	10	0	3	160	0.04%	0.63%	25,582	50.69%
US Senate (Perdue)	50,677	50,594	7	62	13	0	1	83	0.13%	0.81%	10,303	20.57%
US Senate (Loeffler) - Special	50,677	50,517	67	70	17	0	6	160	0.32%	3.06%	5,236	10.55%
Public Service Commission District 1	50,649	50,599	5	30	13	0	2	50	0.05%	0.19%	25,914	52.52%

5.2 Precinct Discrepancy Report

The precinct report can be helpful if there are style-specific or machine-specific issues where disagreements exist and can be located in larger numbers in specific precincts but may not be obvious in the contest report. The Disagreed% and Variant% in this report is a percentage of all ballots in the precinct, and across all contests, and these counts are in ballot sheets (not contests), so it you will not be able to add up the figures in each row, because a given sheet may have multiple issues.

The highest 5 Disagreed% and Variant% values are highlighted.

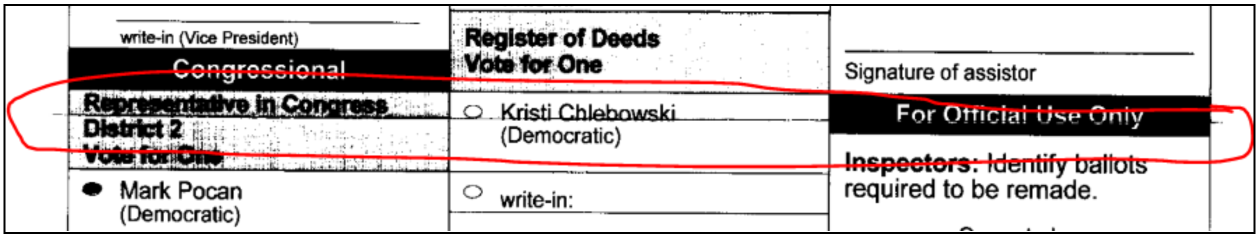
Precinct Report

Precinct	Style	Total	Agreed & NonVariant	Agreed Overvotes	Agreed Write-ins	Agreed Blank	No CVR	Gray Only	Disagreed	All Variants	Disagreed% of Total	Variant% of Total	Disagreed Examples
C Fitchburg Wd 14	[9, 29]	1,512	1,468	1	39	0	0	3	3	44	0.20%	2.91%	[260975, 261481, 262337]
C Fitchburg Wds 1-4	[7]	2,716	2,639	10	56	0	0	15	8	77	0.29%	2.84%	[249167, 248461, 249539, 249712, 250036, 250505, 250924, 251014]
C Fitchburg Wds 10-13	[6, 7]	2,633	2,559	5	62	0	0	10	3	74	0.11%	2.81%	[258914, 259900, 260067]
C Fitchburg Wds 15, 18-19	[7, 6]	3,824	3,720	0	95	0	0	7	7	104	0.18%	2.72%	[262650, 263733, 264708, 265800, 266200, 266478, 266718]
C Fitchburg Wds 16-17	[9, 29]	1,316	1,280	0	34	0	0	2	1	36	0.08%	2.74%	[266924]
C Fitchburg Wds 5-9	[6, 7]	4,560	4,440	0	106	0	0	15	5	120	0.11%	2.63%	[illegible]

By way of example, we can note in this report, one precinct has an unusually high number of Disagreed% and Variant% values in this report:

V Waunakee Wds 6-12	[24]	4,778	4,619	1	155	0	0	3	1	159	0.02%	3.33%	[284175]
V Windsor Wds 1-2	[19]	453	420	0	10	0	0	22	22	33	4.86%	7.28%	[320050, 320310, 320353, 320407, 320442, 320444, 320536, 320542, 320605, 320625]
V Windsor Wds 3-5	[10]	1,802	1,728	3	67	0	0	6	2	74	0.11%	4.11%	[320011, 321786]
V Windsor Wds 6-10	[3]	3,138	3,009	4	113	0	0	11	11	129	0.35%	4.11%	[318586, 319519, 319854, 317034, 317332, 317621, 317684, 317715, 318078, 319249]

You can click on the Precinct name (in the actual discrepancy report, not in this document or in the narrative reports) to jump to the details for this precinct. For example, we can discover in the report that a fold-line crease caused improper evaluations of the "Register of Deeds" contest.


write-in (Vice President)
Congressional
Representative in Congress
District 2
Vote for One
Mark Pocan
(Democratic)
Register of Deeds
Vote for One
Kristi Chlebowski
(Democratic)
write-in:
Signature of assistor
For Official Use Only
Inspectors: Identify ballots
required to be remade.

In this case, the contest was uncompetitive, and Kristi Chlebowski had no opposition. In a few cases (22), AuditEngine improperly awarded a vote for Chlebowski when the contest was actually blank. AuditEngine uses a slightly larger evaluation area to allow for marks that do not go into the oval but are clearly the voter intent to vote for that ballot option. Such is the case when voters circle the oval. This strategy may have this negative result in some cases. This mistake was not made by the voting system. We will be continuing to improve the ability of AuditEngine to avoid this mistake as we progress in AuditEngine development. As it is, when we disagree with the voting system, AuditEngine does a much better job of detecting voter intent correctly without any adjudication, between 70% and 90% better than the voting system in recent audits. We will be continuing to enhance and refine our ability to classify voter intent using machine learning and heuristics.

5.3 Variant Notation

In our classification methodology, we use a multi-component notation:

[bmd_](1)_(2)_(3)[_gry]

bmd_

bmd_ is optionally prefixed if the variant applies only to BMD ballots. If not shown, then it is only nonBMD ballots or may apply to all ballots. BMDs have fewer variant categorizations because overvotes do not occur, and write-ins are typed in, so it is not necessary to adjudicate BMDs very often.

AuditEngine does not rely on the QR/barcodes, but rather uses OCR to "read" the human readable summary.

If there is a difference between the text on the ballot summary sheets and the CVR, then it is either a major voting system failure and very unlikely, or a deficiency in AuditEngine's parsing. Recently, AuditEngine has had few, if any, failures to read this text.

Write-ins from BMD ballots are not in doubt, but they may be excluded later if the written-in name is not from the qualified write-in candidate list, depending on state law.

(1)

The next first component (1) is AuditEngine's classification, and can be:

'ov' - overvoted

'uv' - undervoted and possibly blank

'wi' - a write-in mark is detected and it is not an overvote.

'x' - (one or more) votes, neither overvoted nor undervoted

If there is no CVR, then this is the only classification. (2) and (3) are not used, but "_gry" is used to indicate that the contest should be reviewed with human-eye in close elections.

(2)

The second component is the CVR interpretation. If no CVR is available, then classifications with this component are not used. It uses the same notation as above, but adds adding also:

'x' - votes which match the votes in component (1)

'y' - votes which differ from the votes in component (1)

(3)

The third component is the voting system adjudication, which is currently available only if the CVR is available, and only at this time if it is from the newest Dominion software suite with adjudication records, i.e." 'modified' snapshots". The abbreviations are the same as mentioned above. It is important to stress here that just because an 'modified' adjudication record is available, it does not mean this specific contest has been actually reviewed by election staff if the modified record is the same, i.e. component (2) is the same as component (3).

So for example, the pattern 'wi_ov_x' means that AuditEngine evaluated the contest as a write-in, while the voting system CVR specified that it was an overvote, and after adjudication, it was resolved to a (non-writein) vote.

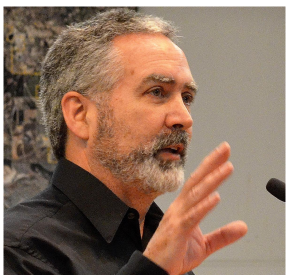
6. Narrative Report

In addition to the automatically generated reports, an audit generally includes a narrative report which provides specific findings and any curious features of the election audited. The Narrative Report will refer to this overview document as background information.

7. Conclusion

Since Ballot Image Auditing, and particularly the approach used by AuditEngine is relatively new, we hope this introductory document will be helpful in your reading of any specific audit reports. We will be striving to continually improve this information so check back each season to see our newest work.

Primary Author: Raymond Lutz



Raymond Lutz is the founder and executive director of Citizens' Oversight Projects, a 501(c)3 nonpartisan nonprofit organization that has been involved in providing oversight to elections for over 15 years. Lutz has a Masters degree in electronics and software engineering, with experience in the document management and printer/scanner/fax/copier industry, and medical device industry. He is the lead developer of AuditEngine.