

**Ballot manifest discussion- Colorado Risk Limiting Vote Tabulation Audit** Harvie Branscomb  
3/12/2017

**Background:**

This description is limited to counties using the Dominion Democracy Suite that produce a CVR.

The risk limiting audit is an evidence based audit. The evidence for the audit is the collection of voter-marked, or machine-marked but voter-verified paper ballots.

Intermediate records that are used are election night sub-tallies, cast vote records from individual ballots, and a ballot manifest that ideally reflects every cast ballot.

While during precinct polling and precinct tabulation the scanner that tabulates the vote also has first-hand knowledge of the casting of the ballot, in Colorado this is not the case. Colorado has no in-precinct voting and almost vanishingly little in-voting-place scanning or tabulation besides the remaining DREs..

Ballots in Colorado arrive almost entirely in envelopes, either dropped in election-specific drop boxes or mailed in. Less than 10% are voted in person, either voter marked on the same format as the mail or machine-marked as selections-only QR-encoded voter intent.

**Sources for the ballot manifest**

There are no ballots cast into scanners in the Democracy Suite system. Hence, scanners are not a suitable place to capture a ballot manifest - at least not for the long term.. In Colorado, ballots arrive within the election system long before they reach a scanner. In this paper I will present several scenarios for creation of ballot manifest that would allow the RLVTA to sample from a list closer to all ballots cast than just all ballots scanned. For simplicity and in the startup year of the RLVTA it might make sense to have the ballot manifest originate at the scanner but this is not a viable long term solution.

In Colorado ballot styles are almost always chaotically distributed by ballot style in collections such as batches. Batches that come from VSPCs (Voter Service and Polling Centers) may have predominance of a few local styles, but can contain any style in a county election. For the predominant "mail/drop ballots" some counties may actually sort their ballots within return envelopes upon arrival, in which case the batch may contain a uniform style or desired selection of styles. It is possible to regulate the number of styles in a batch by sorting either in the envelope or once the ballot is out of the envelope and rendered anonymous. Large counties have machines that can automatically sort ballots in the envelope but these machines have limited hardware "pockets" in which to sort the envelopes.

**Contents of the ballot manifest**

The ballot manifest ideally will contain or index into a database containing the style of each ballot and its location in the batch, the identity of the batch, the box, the room and the county where the ballot resides. The ballot manifest does not and ought not contain a reference to the voter. If all this style and location information is available to the manifest then very efficient sampling can be performed, targeting the contests remaining to be audited. Sampling is done by matching the ballot manifest to

the list of integers produced by a PRNG (pseudo-random number generator).. At a minimum the ballot manifest contains the quantity of ballots in a batch, the identifier of the batch, and indirect pointers to the identifier of the box, of the room and of the county. The ballot manifest should be able to be used on a statewide basis without any confusion.

### **Benefits of knowledge of style in the ballot manifest**

If the sequence position of each style in a batch is not available in the ballot manifest, then efficient sampling can still be done for countywide contests without additional information - all ballots contain the contests of interest-, but for intra-county contests (including multi county districts that represent part of a county) it will be necessary for best efficiency to find the style of each sampled ballot either in the ballot manifest or alternatively in the CVR associated with the selected sample. One may determine from the CVR if the sampled ballot contains the desired contest or contests. The CVR may not correctly characterize the ballot but if so, an audit discrepancy will occur. If sorting by style has been done, then this lookup problem is probably solved. Alternatively the paper ballot may be physically obtained and the style determined from direct observation. That is the inefficient way, wasting time if the ballot sampled is not a matching style but may still be practical for some purposes.

### **Choice of election step to begin making the ballot manifest**

There are three general categories of intake process for mail/drop ballot envelopes in Colorado counties:

- 1) Large county intake with envelope scanner/sorter machines;
- 2) mid and small county systematic "binking" of envelopes at ballot intake to form batches before sigver (signature verification); and
- 3) mid and small county binking individual envelopes at the time of signature verification.

Binking is the colloquial term for using a wand to read the bar code on the front of a ballot envelope.

A system could be developed to build the ballot manifest either from the envelope sorter, or from SCORE (statewide voter registration database) or both. Initial batches of envelopes would require sorting or adequate mixing into subsequent batches for which voter identity does not carry over by sequence or by style. I do not suggest that this be done in 2017 other than by pilot experiment to understand the potential and pitfalls, if any.

There are three other places later in the processing where the ballot manifest might originate:

- A) At the time of opening/flattening the ballot. Sometimes the style is checked at this point. Also there is a check for no-ballot-present or two or more ballots in the envelope (or mismatching cards).
- B) At some time prior to scanning when an identifying number might be imprinted
- C) At the time the scanner reads the ballot and a CVR is created.

### **Benefits of creating ballot manifest early and privacy concerns**

Ideally the RLVTA will sample from literally all incoming paper ballots (cast ballots) and will be able to test for accuracy in determination of eligibility, or at least test that an eligibility process was undertaken. The caveat always remains that the identity of the envelope must not carry onto an

identity of the ballot after the separation of ballot from envelope takes place. This will require some careful planning and batch management- likely more time than we have before the elections of November 2017.

Colorado open records law presumes that when 10 or more ballots of the same style are present in a collection, it is not possible to identify which was marked by a voter even if there is corresponding list of voters. This means that if any batch for which there is a list of voters has less than 10 ballots of same style, that privacy might be violated under current Colorado law.

If batches are of only 25 ballots, unless there are two or fewer styles, this will require effort to accomplish. It can be done by sorting by style and batching by style. Privacy through confidentiality of records has been the prevailing mode of operation- and anonymity of ballots has been provided for just before records are made available to the public. In 2017 Colorado could choose to maintain this same less effective style of privacy protection as an interim measure.

### **Compromise suggestion for 2017**

A less than perfect source of the ballot manifest is the voting system itself. Since the voting system is also being audited, its documentary artifacts such as the CVR, including the style represented by the CVR, may be suspect. In initial implementations of the RLVTA it could make sense to simply start with the CVR as a stand-in for the ballot manifest with an intention to implement a more effective ballot manifest based on other techniques that may take time to implement for 2018 when the challenges for the audit are greater because of multi-county legislative district elections. This simplification will not apply to the non-CVR counties, but their scanners may provide the information adequate for such an interim manifest.

### **Intra-county contest audits prefer knowledge of style for efficiency**

For audit of only county-wide contests, no knowledge of ballot style is needed. For audit of intra-county contests within a single county, it is helpful but not essential to be able to determine which sampled ballots contain the intra-county contests.

If the CVR does not directly contain the ballot style, at a minimum the style can be extracted from the CVR by noting the contests that have been voted or undervoted. A numerical coding system could be created that turns the actual contest pattern on the CVR into a style code that can be used for purposes of sampling intra-county contests. The election night reporting system could produce the number of ballots counted in intra-county contests, and the CVRs can be aggregated to determine how many ballots per contest are in each collection. This, along with ENR-based margins are what is needed for an initial implementation of RLVTA for county-wide and entirely intra-county contests.

### **Compromise solution for quick implementation**

A simplification that is available in the 2017 election is to ignore multi-county intra-county contests completely and simply plan to audit county-wide contests to risk limits and allow single county intra-county contests to be subject only to risk-measurement. Multi-county intra-county contests

could be ignored for purposes of the initial implementation of the 2017 RLVTA. Not until 2018 will these include legislative districts. By then this simplification will be inappropriate.

Eventually we will want to create the ballot manifest as soon as paper is available to the county in order to be sure to audit every eligible ballot even if it did not due to an error reach the scanner, but in 2017 this step could be avoided. If we wait for the CVR, or a voting system report from a non-CVR county, to make the ballot manifest, we can get at least overall style information from the system. With CVRs we know styles by collection.

Eventually, and presumably in 2018 we will want to audit legislative districts that are both multi-county, intra-county and both multi and intra county and of course, state-wide. For these purposes a bit more planning will be needed, that ought also to be prepared for during this year and with the human resources now at hand.

#### **Other data needed as soon as it becomes available**

Once we see the exact format of the unified Cast Vote Record format to be produced by Dominion, we will know better what is needed to create the ballot manifest. I hope that as soon as the SOS does obtain a copy of such a CSV-formatted CVR that it will be published on the SOS website. Likewise, those of us advising on the RLVTA are also waiting to know which counties will be auditing with CVRs in 2017. As this information becomes available, please post it immediately. Lastly, the CO RLVTA depends on ENR (election night reporting) data that may undergo a change of vendor this year. As soon as there is information about the format of the ENR data, we would like to know what it is. We are hoping that residual votes, that is under and overvotes, or both combined, are accurately reflected in ENR results. We also need "vote-for" information - that is how many winners there are in each contest, and a way to determine which counties are reporting votes in each contest.

Harvie Branscomb

Harvie at [electionquality.com](http://electionquality.com) [this email and the link below need not be redacted]

<http://electionquality.com>