

## **Strategies for achieving anonymity of ballots (paper, pictorial images, and separately, cast vote records) for Colorado. Harvie Branscomb - harvie@electionquality.com**

Colorado is at the forefront of development and implementation of a very beneficial election tabulation audit. The words “risk limiting audit” are now being uttered in the august halls of the U.S. government. This “RLA” offers a scientific method that seeks evidence of sufficient tabulation accuracy to support a reported outcome. The basis for this confidence is the evidence on voter-marked and verified paper records that must be subject to audit verification by persons separate from those who manage and operate the elections. For the full benefits of this audit to be obtained, access to the election evidence must be offered to interested parties and the public at large. The principled desire to provide audit transparency can teach us about sources of failed anonymity so we can make statutory and regulatory remedies such that all auditable records are treated as the public records that they must eventually be.

Colorado law provides for simultaneous and identical voter privacy protection of not just voted paper ballots, but for pictorial images of ballots, and digital records of the voter intent. The concerns about risk for voter privacy ought to differ as the format of the media and means of access to the media differ. These three items need to be treated separately but paper and image are closely related.

### **Substantive self-identification on paper and pictorial images of paper**

Paper ballots and pictorial images are subject to identifiable stray marks that are the responsibility of the voter - some of which may be considered substantively identifiable. State regulations need to exist to define what marks are to be considered substantive. In Colorado the legislature has removed the requirement for a secrecy sleeve around the ballot to protect voter privacy as it is extracted from the ballot envelope. This recent legislative agenda has refocused attention on the issue of identifiability of voter intent as the ballot is accessed. The newly revised statute requires election rulemaking, not yet promulgated, to ensure ballot secrecy.

A first step in that direction would create a regulatory specification of “substantive self-identification” by a voter so that officials know what they must be concerned about when election judges are opening ballots and thereafter. The time of first flattening of paper ballots is the best time to locate and remedy these problems for voter privacy before ballots are scanned and the problematic evidence is spread around and seen by numerous officials such as election judges and watchers. It is unreasonable to expect each county official to determine how much risk to take that a type of identifying mark will be considered a problem by a court. There must be statewide standard set, and best practices put in place to cause problematic ballot self-identification to be uniformly remedied by duplication or mechanical redaction of the potentially harmful mark. Since duplication adds to risk of recording voter intent erroneously, and because CO’s audit regulations require (as they should) the original marked by the voter to be audited, a means of physical redaction reversible under court order is a superior method.

Note that self-identification is largely not a risk with the CVR. So-called “pattern voting” represents the only risk of this type within the CVR and there is no way for a voter to arrange for any pattern to be unique in an election and hence presumably identifiable. However, it is important to recognize that fewer styles with larger numbers of voters associated with each does make this risk even more negligible.

### **Risk to voter privacy within the CVR**

Cast vote records are discussed at some length in a separate document at the SOS website:

<https://www.sos.state.co.us/pubs/elections/VotingSystems/riskAuditFiles/2018/20180309PreservingAnonymityOfCVR.pdf>

Sources of systematic risks to voter privacy (and the most typical associated cause - failed anonymity) relate both to coordination of contests that create rare district styles, the overlap of precinct reporting on top of precinct-ignorant district contests such as special districts and school districts. A separate risk originates from mistakenly coherent batching of uncontrolled, unsorted distributions of ballot styles. Tabulation batches when they correlate to envelope batches lead to some rare style ballots that will be identifiable to a voter as they are stored. These are symptoms of lack of attention to voter privacy that was a norm in a system where the public had little chance to access the ballot records. With pictorial digital images and cast vote records, that technical boundary has been removed and newly designed audits are now ready to take advantage of the accessibility of ballot records to provide for excellent election verifiability with public involvement.

Lastly there is also a potential for delayed, limited release of ballot records to subdivide the election in time such that individual styles when separated by delay become rare. This risk, like others, has a systematic solution. Before systematic solutions to these risks are implemented at the source, we are currently in a position of need for interim measures to provide cleansing of the election records such that the public may access without substantially invading voter privacy.

The Colorado Open Records Act defines a threshold of ten or more for anonymity in numbers of ballots or portions of ballots of same form or style. Counties are individually attempting to amend or withhold their public records to conform to this requirement. It turns out there are many alternative solutions to the redaction requirement. In a nutshell, at least three specific methods have presented themselves in recent efforts:

- 1) Redaction of specific rows, representing specific ballot sheets, for any row that presents an identical "style" to only 8 or fewer other rows (9 or fewer total). All such rows are redacted out of the record and this affects the usefulness for verification of any contest for which voter intent has been redacted away- it will likely affect most contests. This type of redaction is almost certain to remove voter intent for audited contests, thus eliminating the potential for the CVR to be verified or used as part of a verification of an audit of a contest. This method ought to be avoided.
- 2) Redaction of specific rows by numerical aggregation of each grouped by district style. This is a new method that has recently been attempted by Boulder County. This does preserve the potential for summing the columns of each contest in the CVR file for audit verification but it fails to teach the reason why the redaction is necessary- namely the inclusion of specific precinct-ignorant contests that created problematic precinct splits. It also fails to reflect the extent of redaction and requires trust in the redaction if redacted data must be used to verify an audit. A better method, but not best.
- 3) Redaction of specific columns that represent choices for contests that once removed produce a form of ballot that is identical in 10 instances or more throughout the remaining CVR. This method demonstrates the sources of rare styles in terms of particular contests included in coordination. This method points to solutions closer to the source- namely: 1) the decision to coordinate a problem-causing contest on the same style with disparate-border districts, and 2) the design of the district such that it chaotically overlaps with precinct boundaries for standard statutory legislative and county commissioner districts that represent the bulk of a general or primary election ballot. Non-precinct-based and precinct-based districts do not share a ballot sheet well- they produce precinct splits, extra styles and problems for accountability and anonymity. They could be separated onto separate sheets that are not coupled together- particularly when the election already requires two sheets.

Selective redaction of contests – the third approach here - to achieve ballot anonymity is I think the best mechanism for learning the source of failure to provide an anonymous ballot, as required by the Colorado Constitution: “no ballot may be marked in a manner by which it can be determined who voted it.”

Neither row redaction of CVRs nor treating CVRs as if they are ballots pave the best route to long term voter privacy. The eventual goal should be: all tabulated ballots are eventually to be tabulated as anonymous and any paper that is substantially self-identified will upon removal from envelope be reversibly redacted prior to tabulation. That way, any subsequent image made by a scanner will also become anonymous. CCCA presidents Pam Bacon and Lori Mitchell were heard recommending redaction prior to tabulation at a recent Colorado County Clerks Association conference.

I think a first step to cleanup systematic risks to anonymity is to write and share a report about anonymity from the time of election design. Such a report would include how many electors there are for each district style and each precinct-split style. This calculation can be made as the decisions to coordinate are made by local officials. Such a report can be used to estimate the extent of anonymity problems as defined by CORA and potentially to design remedies that will avoid rare styles of ballot sheets once we can accommodate individual styles per sheet in a multi-sheet election. Even now, with Dominion and Clear Ballot, each separate sheet of a ballot produces a separate CVR entry.

Then after election day the CVR file can easily be modified by neutralizing the voter intent to show only the equivalent design data but updated for numbers of actually returned ballot styles. The state could publish this for all counties using as a source the full CVRs uploaded for audit. The resulting statewide file is what could be used by officials and public to determine what paper, what images and what CVR columns are clear of voter privacy issues. Of course, as previously mentioned, paper and image deserve removal of substantive voter self-identification too and that is a separate procedure.

The redacted CVR with zero voter intent included will eventually, when we have removed sources of systematic identification, serve as public proof that we audit paper and CVRs with no remaining anonymity issues. That can be treated as a goal for the future.

Meanwhile we need to treat identifiable ballots as exempt from the audit with a Philip Stark "zombie" approach or similar. When identifiable ballots are encountered in the audit, they can be substituted for by additional auditing by assuming that the inaccessible ballot contains voter intent with the worst prognosis for effect - meaning in the direction of reversal of presumed outcome.

While pursuing CVR anonymity protection officials can learn how to coordinate and paginate elections to avoid creating rare styles, knowing that styles need not span cards in an election and cards need not remain associated together.

## **APPENDIX**

### **Here is a short term strategy to accommodate RLA transparency:**

- 1) When adding contests to coordination, create an “eligible voter record” for the election:  
Eligible Voter Record
  - Model the election by determining district styles needed
  - Create the list of active voters for each district style, by precinct if applicable

- Report the number of voters per precinct-split or district style
  - Estimate the turnout per style and precinct. Compare estimated turnout to the CORA threshold for anonymity- 10 or more
  - Consider methods to protect privacy such as dividing contests into separate styles and separate sheets
- 2) Make sure batching before tabulation creates independence from envelope batches, not only in name but in ballot content. Collections large enough to contain many instances of each style can remain constant throughout the election- these can be enclosed in large tubs while batches within are jumbled.
  - 3) Hold back enough ballots of each style for delayed tabulation- at least 10 per style to ensure that late added ballots are not identifiable by style according to CORA criteria. This in effect means shuffle batches so that late batches are not exclusively late arriving ballots.
  - 4) Do not isolate in-person ballots or UOCAVA or provisional into specific batches or specific tabulators but mix them with mail-in batches. Yes, this requires accounting adjustments just as occasional ballots are removed from batches, some may be added. Also remove the statutory requirement to mark any ballots as “provisional”. Provisional ballots once tabulated should not be recognizable.
  - 5) Have election judges examine ballot cards upon opening for substantive self-identifying marks and redact with pressure sensitive labels (already in use for other purposes) or duplicate only when necessary.
  - 6) After election day, modify the Cast Vote Record file into a “Tabulated Sheet Record” by replacing the record of voter intent (0,1) with a neutral character such as “X.” Remove characterization columns that are row specific such as CVRNumber, Sequence. Then sort the rows with a random row number.
  - 7) Use the above Tabulated Sheet Record to determine which contests can be redacted from the CVR to produce a remaining CVR with new, virtual styles that require no redaction according to CORA. Plan to include contests targeted for audit in this redacted CVR if possible.
  - 8) Produce and publish a revised column-redacted CVR with all columns present except BallotType and the select redacted problematic contest columns.
  - 9) Test the produced CVR by making a new “Tabulated Sheet Record” for the redacted CVR, counting the number of identical form voter intent rows (neutralized voter intent) to check if all are 10 or more. Examples of these techniques are posted here- see the last several comments for March 23, 2018):  
<https://www.sos.state.co.us/pubs/elections/VotingSystems/riskAuditGroup.html>
  - 10) As long as columnar redaction is still required (meaning longer term strategies have not been accomplished) some complete ballot pictorial images will not be accessible to the public (at least until CORA is modified, or Title 1 addresses its own requirement for beneficial public access to ballots for election auditing purposes.) However, under current CORA, incomplete pictorial images of ballot cards could be provided to the public. For example, only front side, or only some geographical portion of the image, depending on style. In future, we could have automated methods for affording access to images by contest or groups of contests. Ideally, the longer term strategy will produce sheets and also CVRs that are entirely free of need for redaction or withholding.

## **Long term strategy**

Plan for a two-sheet (or more) election. Format all legislative districts plus county and state districts onto one sheet (both sides if needed). This will result in a standard minimal number of styles coded by precinct- likely one style per precinct. State Representative and State Senate, Congressional District and County Commissioner Districts will define the sheet 1 styles. There will be no precinct splits. On sheet two, place all coordinated contests and do not report sheet 2 by precinct and do not associate the sheet 2 style with the sheet 1 style. There will be no precinct splits of this second sheet's styles because precincts are not reported for these non-precinct contests. The two sheets have no need to be associated with each other once removed from the envelope and will generate independent CVRs. The accounting for votes counted will be independent of the sheets they arrive on- one count for each sheet. The number of ballots cast will be produced by eligibility data from SCORE, not from the vote tabulation system. The number of votes counted will ideally be reported by contest, not by "ballot" or by sheet.

The incremental cost of the additional ballot sheet can be borne by the districts that request to coordinate. Legislation and rulemaking are likely needed to conform existing statutes to this proposed practice- particularly the Form of Ballot that specifies existing ballot order and ignores pagination.

### **Longest term strategy**

Work with legislators to obtain remedies for redistricting and re-precinct-ing to maximize coordination of district boundaries with precinct boundaries especially for school districts and municipal districts. Avoid any split of a small county into more than one Congressional District. Avoid special, school and municipal district or Congressional District crossover into other counties for small numbers of electors. If any remain, require a separate ballot card for all voters in contests that when coordinated, result in only 30 or fewer electors in any of the coordinated counties in the contest. Those electors should be sent separate ballots for the rare contest to be returned to the county that has the lead DEO for the contest.

Harvie Branscomb 3/22/2018 updated 11/27/2018 [www.electionquality.com](http://www.electionquality.com) harvie@electionquality.com