
From: Harvie Branscomb
Sent: Monday, March 5, 2018 7:34 PM
To: Dwight Shellman; Matthew Crane
Cc: Neal McBurnett
Subject: test of CORA 205.5 compliance, Arapahoe County
Attachments: first2K_Arapahoe_022218_CORA_Request_CVR_Export.csv; testG_first2K_Arapahoe_022218_CORA_Request_CVR_Export.xlsx; first2K_Redacted_CVR_Export_20171116151057.csv

Dwight, Matt and Jennifer and Neal

Gradually a process is becoming clear to me for optimally satisfying transparency needs of CVR for verification of audit and also meeting CORA 205.5 standards. It is not necessary or desirable to publish a separate CVR for every contest. That will lead to confusion. I suggest columnar redaction and eventual removal of the redacted contests onto a possible separate CVR (and more eventually, a separate card). Forget the second CVR or card for the purpose of this email.

Counties and requestors and possibly SOS need access to a file containing a map of each election per county showing the potential for risk of voter privacy. I am going to call that for now the Cast Card Record CCR. (A Cast Ballot Record would be what the eligibility system knows about- one row per voter rather than a row per card). The CCR is an edited and sorted version of the CVR file. Arapahoe County supplied me with it and using it I then requested from Arapahoe only contests that meet the CORA standard and Matt and Jennifer supplied me with a redacted file, attached.

What I am calling a CCR is what I asked Arapahoe for at first (slightly modified as I learned how to do this):

1) Create a Cast Card Record CCR that is a modified CVR file but with no voter intent, only evidence that voter intent was present. (Ideally Dominion will produce this - of course each county can massage the CVR to CCR but that is a waste of effort if everyone is doing it.)

replace all 0 in voter intent with X; replace all 1 in voter intent with X; leave blanks as blanks

remove data (not the headers) in columns that identify individual rows- CVRNumber, RecordID, ImprintedID,

leave in CountingGroup, PrecinctPortion, BallotType and also TabulatorNum and BatchID;

sort by BallotID or otherwise destroy the original order of rows so they cannot be matched later with fulfilled voter intent.

Arapahoe provided a file similar to this attached (I removed most of the rows to simplify):
first2K_Redacted_CVR_Export_20171116151057.csv

2) Post this CCR for everyone- it is harmless but gives contest names, how many ballots per contest, enough data to determine how many rows will have fewer than 10, etc. Also separately informs about anonymity of

the images and paper ballots without revealing any voter intent. (SOS instead of Dominion could massage data into CCRs - by redaction, replacement and sorting-starting from the uploaded CVRs counties supply to SOS prior to dice rolling and SOS could then post centrally- that makes sense.)

3) Counties hash the contest columns (all choices per contest together) in the CVR separately by contest prior to upload so that whole groups of columns can be redacted and the hashes still compared later.

4) Before CVR release someone (at the moment it is me as a requestor- possibly could be automated for use by counties or SOS) takes the CCR and figures out which contests to redact from CVR to obtain releasable data (all new rows having more than 10 instances). Proposed redacted releasable file must redact BallotType and probably CountingGroup and PrecinctPortion (if splits are reported and causing identity problems). Ideally, audited contests are not redacted in the proposed releasable file.

5) test the release candidate CVR by calculating a new ballot style for each row, sorting by style and counting rows.

Attached is Arapahoe's redacted CVR as provided to me- but only the first 2000 rows to save email attachment space:

first2K_Arapahoe_022218_CORA_Request_CVR_Export.csv.

Note three columns F,G,H redacted plus several contest columns.

The version converted by me to test the redaction is:

testG_first2K_Arapahoe_022218_CORA_Request_CVR_Export.xlsx.

Arapahoe's redaction passed the test as the lowest count of identical format rows was 12 (the list is shown on the second sheet in the file).

The procedure for the test is a bit similar to creation of the CBR:

- replace all 0 in voter intent with 1; replace all blanks in voter intent with 0; leave 1 as 1

- use empty columns to create a concatenation of all voter intent rows- producing a binary number that characterizes the new style of the row
- convert the binary to decimal in another column

- sort data rows by new decimal style

- count these rows (various methods available), check if any show count of less than 10.

(this process should be implemented as a tool for any county or SOS to use- to test conformance with CORA 205.5)

6) If the test passes, release the redacted CVR.

As you can see I did some reformatting to make it readable for my purposes, such as the redaction test. CVRs can be formatted much more conveniently than they are from Dominion once they are out of the .csv format.

Thanks to Arapahoe for participating in this experiment. It does show Arapahoe which contests are the ones that impinge on voter privacy according to CORA- If these redacted contests were on a second card, voter privacy might not be a problem because problematic contests would appear in a separate card and hence CVR. Likewise voter privacy of the image and ballot go away as the rare styles caused by cramming all contests on one card go away when using two cards. Once I receive the CBR from all counties I can report the voter privacy issue per county without ever seeing actual voter intent.

El Paso County did not require any redaction of its CVR- having no rare styles.

Adams County had many contests that needed redaction.

Harvie