Virginia NIST Whitepaper

January 10th 2018

I. Introduction

As part of the Help America Vote Act, the Election Assistance Commission (EAC) has been tasked with creating common data formats (CDFs) to support various aspects of the election process. These include, but are not limited to: voter registration records, pre-election information, ballot definition, election results, machine event logs, and ballot auditing information. The EAC went to the National Institute of Standards and Technology (NIST), and had them create working groups to identify and involve stakeholders in this process, including state and local election officials, vendors, media organizations, technology companies, and other interested third parties.

The first data standard to come out of this effort was the NIST 1500-100 Election Results Reporting Common Data Format, published in February 2016. The 1500-100 specification -- referred to here as "1500-100" -- defined common elements, including contests, parties, candidates, political geography, and ballot counts. At this time we were in the process of re-implementing our end-to-end data management system. While we did not implement 1500-100 exports at the time, we left open the possibility that we would in the future.

During the winter of 2017 we agreed to work with Google to pilot the publication of data surrounding our November 2017 general election. We planned to publish information in the 1500-100 format, and Google would display results on their search page. This process would give us an opportunity to understand the costs and benefits of implementing 1500-100 compatibility, and provide lessons to other states making decisions around implementing such CDFs.

This paper provides details of our experience, metrics from the Virginia Department of Elections and Google, and concluding observations.

II. Experience

Google approached us in June of 2016, with an interest in collaborating on a pilot project based on the NIST 1500-100 data standard. Between July and September we attempted to create a test feed, but by mid-September we agreed that it was too late to create a working feed in time for the November election, and deferred until 2017. Virginia conducts state-level elections in odd-numbered years, and in 2017 there would be a gubernatorial election, and all seats in the House of Delegates would be up for election; the state senate is on a four year cycle and all state Senate seats will be contested in 2019.

Virginia met with Google again in February of 2017. We agreed upon a goal of publishing a 1500-100 feed in November, and Google would display results of the Gubernatorial election directly on the Search page when people used query terms like "Virginia election results" or "Virginia governor election". We agreed that the primary metric of interest would be **the amount of web traffic seen by the Google search result**. If more people saw authoritative and up-to-date election results directly in Google, there would be fewer hits on the web infrastructure of the Department of Elections. Less traffic to the state would mean we would spend less money on web hosting and decrease the risk of our website failing during peak traffic on election night, all while reaching a larger audience.

At this time the Department of Elections was in the process of re-implementing our backend data storage infrastructure. This infrastructure includes our voter registration database and voter file, list of candidates, list of polling places, and election results. One unexpected benefit of the 1500-100 project was our ability to leverage the data model on which the standard is built. The NIST working groups put considerable effort into creating a sensible data model which strikes a balance between simplicity and comprehensiveness. The NIST data model helped inform our internal data model; that, in turn, lowered the amount of work necessary to export results in the 1500-100 format. We discovered that the 1500-100 format was able to represent all of our pre-election and election night information in both the partisan primaries in June and the general election in November.

The Virginia Gubernatorial election ended up being a high-profile contest, with a close race between Lieutenant Governor Ralph Northam and former RNC Chairman Ed Gillespie. On election night the Virginia Department of Elections website served approximately 1.78 million pages with results for the Gubernatorial election, with about two-thirds of those coming from mobile devices. Google, however, reported 4.4 million page views from 2.3 million users, also with two-thirds coming from mobile devices. Google reported that 95% of the traffic came from within the United States, 2% from Canada, and the rest from other countries.

While we were not able to determine how many of those 4.4 million page views would have instead gone to our website, even half of them would have more than doubled the traffic to our site. Furthermore, Google's election results product supported approximately ten languages, furthering the reach of our data. While we hope to see more clients of this data than just Google, this is an example of how common data formats enable third parties to provide authoritative information to communities that a state or local government may not otherwise be able to reach due to budgetary realities. This has convinced us that a common data format which has widespread adoption has the potential to increase the reach of accurate and authentic election data at a reduced cost and risk to state and local governments.

III. Conclusions

Our experience learning about and implementing the NIST 1500-100 common data format was a positive experience for the Virginia Department of Elections. It helped inform the development of our backend data storage infrastructure, addressed our public-facing use cases, and enabled a partnership with a third party that reduced the cost and risk of running our election results website. Additionally, we learned that the 1500-100 standard is compatible with version 5.1 of the Voting Information Project (VIP) specification. This compatibility allowed us to implement exports of VIP 5.1 feeds in a matter of weeks instead of months, and will enable us to provide richer, more detailed information to our citizens in 2018 and onward. We would encourage other states to engage with the Election Assistance Commission and NIST to learn more about their common data format projects, and assess how these projects could provide tangible benefits to states and counties in the form of more efficient, lower-cost, reduced-risk operations while increasing the reach and benefits to citizens of election administration departments.