CASE STUDY #2

WISCONSIN

Greg Grube, GIS Elections Specialist, along with Sarah Whitt, WisVote Functional Lead, both from the Wisconsin Elections Commission, shared the following information on how they geo-enabled their elections a decade ago in the state of Wisconsin.

THE WHY IN GEO-ENABLING ELECTIONS

Before using GIS in our voter system, we districted voters using street ranges. When the statewide voter registration system was initially set up in 2005, we worked with local jurisdictions to determine what precincts, school districts, and special districts they had within their jurisdictions. We used this information to determine what precinct-splits (we call them "District Combos") that jurisdiction had. We then reviewed local maps to determine all the streets that were within those splits, what the house numbers were and what side of the street they were on, placing that information into the voter registration system. When voter registrations were entered, the system would compare the address on the voter registration to the street range tables to determine the voter's district combo and therefore ballot. If the user entered a house number that was out of the range or entered a street name that did not match the ranges, they manually needed to assign the voter to a district combo, or else adjust the ranges to accommodate that address.

In 2011, new precinct boundaries were drawn in response to the 2010 Census. This work required rebuilding all the district combos from the ground up. Due to the significant effort involved in this process, we determined it would be more efficient to use GIS to build the new district combos and use it going forward to assign voters to districts. The Information Technology (IT) team proposed collecting shapefiles for electoral districts and using the Census-based precinct files generated through redistricting to create the new district combos. However, it was determined early on that the census-provided precinct data was not spatially accurate enough for election administration and that accurate shapefiles were not available for all electoral districts such as school districts, sanitary districts, aldermanic districts, and technical colleges. Staff reached out to county Land Information Officers for help, and most counties were able to provide school district boundaries that were more accurate than the Census districts and they helped to line up census data to parcel boundaries. For other districts that were not available as shapefiles, staff read legal and other documentation to determine the district definitions and manually assigned those until maps could be found or created.

All existing addresses were run through a commercial geocoder to determine their latitude and longitude coordinates. These coordinates were then compared to the district combo maps. The system flagged certain types of exceptions for clerks to manually review, such as addresses that appeared in a different district based on the geocode, addresses that were within 50 meters of a boundary line, or other problematic locations. The review of the addresses created a large volume of work for local election officials that was very difficult for them to complete. In some cases, not enough validation was done ahead of time, and

many districting issues were not identified until voters appeared at the polling place to vote. We also did not initially have proper ways for the system to catch issues like duplicates or addresses in the middle of the road.

We have since created better, more accurate shapefiles, and significantly improved our geocoding process. We are fully implementing GIS and have geo-enabled elections, though we still have many improvements to make. We would like to have better source data (especially with addresses), better auditing methods, and improved updating processes.

Our current system identifies errors in district placement by comparing underlying district data with the placement of an address point, which in our system displays with a pin. Three main voting data types determine ballots in Wisconsin: wards, which (often called precincts), school districts, and sanitary districts. Wards determine election types like statewide representatives, congressional representatives, as well as local races like mayor and town board. These are always coincident with or within municipal boundaries. School districts and sanitary districts, on the other hand, are their own entities and have geography that doesn't coincide with municipal boundaries.

To create the smallest possible voting unit, we combine those three into the District Combo. Every active address in the system should have a district combo associated with it, and every district combo should be assigned using the underlying district data. Users of the system can move address pins around if the location is wrong (or if it is a new address and our geocoding system can't find a location). If an address is assigned a district combo that does not match the underlying geography, they get automatically marked as mis-matches, and either state staff or local clerks are supposed to fix it.

In theory, if every address in the system is placed somewhere in Wisconsin and if the district boundaries are all correct, then every address should have its district combo assigned correctly based on location. When a new address gets added to WisVote, we have the user press a geocoding button and it is placed automatically. Otherwise, they should manually place those locations by moving a pin. However, if a clerk would manually place an address incorrectly (for example in a lake instead of on top of a house), then our system assumes that the clerk did that with a high level of authority. When clerks are not precise with their locations, voters may be districted improperly. We can't currently identify those that are manually misplaced without a manual review.

GIS is the best tool for this work because if users are working with both good district and address data, then GIS can use spatial logic to verify that voters are in the correct place, instead of relying on spreadsheets or only local knowledge. It also can be much more intuitive to visualize issues on maps instead of in a table. For example, if a district line is going through the middle of a house, that will hopefully at least cause a user to question the data or process.

THE WHO IN GEO-ENABLING ELECTIONS

When we first embarked on the geo-enabling process, we did not have accurate enough data to make the system work properly. We relied on users to provide address locations, and we had to reach out to all 72 counties in the state to get their district data. The state has since made progress in getting better data. Ward data is now required to be submitted every six months, so we have a better picture of that data. In the past year, school district data has become available in a digital form from the Department of Public Instruction, and

we can cross check that with parcel data that is also required to be collected at the state level every year. Sanitary voting districts are not currently maintained at a statewide level; however, there are fewer than 50 in the entire state that hold elections. We maintain our own GIS files of those to use in WisVote.

Addresses are not as readily available at a statewide level. Most counties in Wisconsin have address points, though they vary in quality from rooftop accuracy to parcel centroid and state centerline points. As a result, we are not currently able to use these comprehensively. Right now, we use commercial geocodes provided by ESRI. We chose to use them based on research of different companies and determined that their pay-per-geocode method was the most cost effective for us. The downside of their pay model is that it doesn't allow us to do re-comparisons or audits easily. We hope to eventually use a hybrid model, where we utilize county address points in highly accurate places, and then rely on commercial geocodes and local clerk knowledge as a backup.

IMPROVING THE PROCESS

As talked about above, we would like to use better data to improve the quality of our address data. We are hoping to get that data directly from counties; however, data from USPS or some other source may be an option as well. We would also like to do a better job of auditing addresses already in the system. Re-comparing addresses with authoritative data every six months or annually would allow us to correct poor manual pin placement, as well as add in new addresses that clerks haven't had the chance to enter.

We would also like to see ward, school, and sanitary district data that better fits parcel and municipal boundaries. Right now, we have to make manual corrections to those lines, but in the future, we are hoping that stakeholders like the US Census Bureau, WI Department of Public Instruction, and county GIS staff can provide data that we can merely consume. Having to manipulate the district data to make it fit election needs is hugely time-consuming and is a duplication of effort.

CHALLENGES

Again, as talked about above, the quality of the data is the primary constraint. Having consistent and accurate data is a critical step to make sure that no voters are districted incorrectly. We must ensure local and county clerks understand the importance of placing addresses in the correct location. When we first started using geo-enabled election processes, clerks were advised to not worry about exact locations. Just place the "pins" in the right district area. Because of this, when the lines change due to redistricting, we may end up with a lot of data that needs to be fixed. With better quality data, we should only have to change the underlying maps during redistricting periods, and if the addresses are accurate, then they should automatically get re-assigned correctly.

Another challenge is explaining how the process works to various stakeholders and getting them to buy into a change to their existing processes. Many of the clerks in Wisconsin that have been around a long time like to do things the way they've always done them. Because Wisconsin is extremely decentralized (1,850 different local election officials), a clerk that must deal with 100 voters during an election has very different problems than large cities that have tens of thousands of voters. Getting everyone to agree and understand how vital

addressing is for correct balloting is difficult, especially when clerks and other state election staff are not geospatial professionals.

We've attempted to solve these problems by explaining how the system works, writing up tutorials, and sharing how this will save time in the future. For the data piece, we want to try to make all our data providers aware of how we are consuming this data and the importance of making it accurate and complete.

ADVICE, SUCCESSES, AND BENEFITS

My advice to other states would be in three parts. First, ensure you have the data to support geo-enabled elections or else a method to get around not having it. In our case, we didn't have reliable address data or district data when we decided to do this process. Both district and address data have improved over time, but much of our initial spatial data was not high quality, and some of our address records in WisVote reflect that fact. Here are some questions to ask yourself: How do you get this data? Is it high enough quality to use or is there a way to supplement with other sources? How do you deal with data that isn't good enough quality? Establish a workflow with the stakeholders and ensure they understand the quality (and consequences) of what they are sending.

Second, ensure you have a plan on how to audit or improve the data over time. Many addresses don't change much, but depending on the area, there may be a lot of new construction or even house numbers or street names changing. Ensure you know how to handle these situations because they will happen. For example, we just had most of a county change their entire addressing scheme, so we had to go into our system and match up old addresses with the new to ensure voter records were moved properly. Because many of those addresses were added manually, it was not always easy to match them. Spelling and syntax errors, duplicate addresses, and other types of user error can make data quality very difficult if these ideas are not well planned and thought out.

Finally, think about how to deal with unique situations because the system can't always accommodate them well. For example, we have a few colleges and universities in Wisconsin that use centralized mailbox locations on their campuses. So, all the mail gets delivered to one location and either it is then delivered by campus employees manually or students pick it up there. Situations like this cause significant problems when trying to locate these voters, especially if there are multiple districts that divide up these areas. Do you have separate living locations and mailing addresses? Where are the actual address points placed on the map? What happens if these students move around from building to building while still maintaining the same mailing box?

Apartment complexes and nursing homes also are challenging. The process is relatively straightforward when dealing with single family homes, but what if a voter registers with Apt 301 and another with Apt 301B? Are these separate or considered the same place? These are all items to consider when designing your system.

We didn't consider the non-election benefits of our data when we started, but the benefits were apparent because many state agencies are using their own proprietary data for things like addresses and roads. Here are some items to consider. Determine how you can tap into that data? Is there a statewide or local effort to bring together large databases? How can you create a feedback loop to audit other agencies' data and make it better?