OPEN DATA’S IMPACT

OHIO, USA: KENNEDY V. CITY OF ZANESVILLE

Open Data as evidence

By Christina Rogawski, Stefaan Verhulst and Andrew Young

January 2016
OPEN DATA’S IMPACT

OHIO, USA: KENNEDY V. CITY OF ZANESVILLE

Open Data as evidence

By Christina Rogawski, Stefaan Verhulst and Andrew Young

January 2016

Special thanks to Akash Kapur who provided crucial editorial support for this case study, and to the peer reviewers [odimpact.org/about] who provided input on a pre-published draft.

www.odimpact.org

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License
Summary

For over 50 years, while access to clean water from the City of Zanesville water line spread throughout the rest of Muskingum County, residents of the predominantly African-American area of Zanesville, Ohio were only able to use contaminated rainwater or drive to the nearest water tower to truck water back to their homes. After years of legal battles, one of the key pieces of evidence used during Kennedy v. the City of Zanesville was a map derived from open data from the water company displaying houses connected to the water line and data showing town demographics. The insights from the map showed significant correlation between the houses occupied by the white residents of Zanesville and the houses hooked up to the city water line, and the case went in favor of the African-American plaintiffs, awarding them a $10.9 million settlement.

Key Takeaways

- Access to open data, combined with other forms of data, can lead to important insights and evidence of conditions on the ground and how they are impacting different communities – in this case, highlighting systemic inequalities.

- Awareness of open data is an important first step that may be overlooked, particularly if stakeholders are not particularly data- or tech-savvy.

- Owners of data sets may add new (or more rigorously enforce existing) barriers to access if the data can negatively impact them.

- The usefulness and relevance of open data can be amplified when data sets are supplemented with data collected through other means – like crowdsourcing or surveys.

Dimension of Impact

- Solving Public Problems
- Data-Driven Assessment
I. CONTEXT AND BACKGROUND

For decades, residents of the Coal Run neighborhood in Zanesville Ohio, a predominately African-American neighborhood, were denied public water service despite living within one mile of public water lines.\(^1\) The situation went back to 1956, when a now-defunct water board refused to extend service to parts of Coal Run. As some residents described in a 2008 *New York Times* article, the water stopped “where the black folks started.”\(^2\)

![Figure 1: Map of Water Line Extensions in Muskingum County\(^3\)](image)

Many residents were forced to rely on extreme measures to source water. For example, they had to use electric pumps to retrieve water from a cistern that was fouled with animals and residue from old coal deposits. Due to contamination, many residents couldn’t even use the water and spent time and money instead trucking water in.\(^4\) Others collected rain via buckets and gathered snow in the winter.\(^5\)

---

Not only did the situation impose a daily burden on residents, it was also demeaning and humiliating. One lawyer cited the following example, which reveals the racial disparities in water distribution:

“One man ... spent the whole morning trying to get water or deal with the water shortage. Meanwhile, he could see his white neighbor caddy corner to him sprinkling his lawn. It became clear if you were white and living outside Zanesville you would get water, but if you were black, you wouldn’t.”

In 2002, some two dozen black residents of Coal Run filed a complaint with the Ohio Civil Rights Commission, saying they had been denied service because of race. The next year, the commission found “probable cause” of discrimination and a month after that, Muskingum County officials announced they would extend water to Coal Run, to be completed in 2004.

The decision to extend the water lines did not mark the end of the battle, however. In 2005, after construction on the new water lines was completed, 67 residents of the Coal Run neighborhood filed a lawsuit, alleging that the City of Zanesville and the East Muskingum Water Authority had refused to provide them public water service for over 50 years simply because they lived in the one predominately African-American neighborhood in a virtually all-white county – in the 2000 Census, Muskingum County was found to have a 93.9 percent white population, with the black community making up only 4 percent of the county.

The case was eventually taken up by the civil rights law firm Relman, Dane & Colfax, based in Washington, D.C. In 2008, after a three-year trial, a federal jury returned verdicts totaling nearly $11 million against the City of Zanesville. This case study examines the innovative use of public data that went into building the successful case and, in the process, addressing a decades-old civil rights violation.

---

9 http://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk
II. CASE DESCRIPTION AND INCEPTION

To determine whether there was an association between race and access to public water services in the Coal Run neighborhood, the plaintiff’s lead attorneys, John Relman and Reed Colfax, obtained the demography services of Dr. Allan Parnell of the Cedar Grove Institute for Sustainable Communities. The Cedar Grove Institute is a nonprofit in Mebane, North Carolina that provides technical assistance, analysis and training to help community groups promote equitable community development.\(^{10}\) It grew out of the for-profit company, McMillan and Moss Research, Inc., whose “research and analyses were being called for in cases involving civil rights, predatory lending, segregation in schools, institutionalized discrimination and community economic development.”\(^{11}\) Mr. Colfax reached out to Dr. Parnell based on the recommendation of Jennifer Klar, now a partner at Relman, Dane & Colfax who met Dr. Parnell during a conference, and based on the Cedar Grove Institute’s reputation for work in civil rights cases and community development. Ms. Klar, as well as pro bono attorneys from the law firm Jones Day and other organizations, also worked with Mr. Relman and Mr. Colfax on the case.\(^{12}\)

As civil rights attorneys, Relman and Colfax were well aware of how public data can help provide critical evidence in cases, but lacked the technical know-how to analyze the data themselves. Parnell, a well-known public data specialist, regularly serves as an expert witness in civil rights cases using open data. For example, he was one of the plaintiff experts in *Texas Department of Housing and Community Affairs v. Inclusive Communities Project*, the 2015 Supreme Court decision that affirmed the validity of disparate impact cases where public data are key.\(^{13}\) Parnell therefore led the data research and analysis for Relman and Colfax, eventually settling on a strategy of combining data from multiple sources (public GIS data, water billing data and demographic data) to create maps that established a clear pattern of racial discrimination.\(^{14,15}\)

Relman and Colfax agreed with Parnell’s strategy in using public data. Once the decision was taken, however, Parnell soon realized that census data would not be effective due to the small size of the neighborhood in question and the distribution of residents within the blocks. For example, within each block, the northern part was typically predominantly white and the southern predominantly non-white.

Instead of using census data, Parnell proposed using publicly available geographic information systems (GIS) data from Muskingum County to perform a house-by-house analysis within the

---

10 http://www.cedargroveinst.org/mission.php  
11 http://www.cedargroveinst.org/mission.php  
12 GovLab interview with Allan Parnell, Ph.D., Vice President, Cedar Grove Institute for Sustainable Communities, October 2, 2015.  
13 GovLab interview with Allan Parnell, Ph.D., Vice President, Cedar Grove Institute for Sustainable Communities, July 21, 2015.  
14 GovLab interview with Allan Parnell, Ph.D., Vice President, Cedar Grove Institute for Sustainable Communities, July 21, 2015.  
neighborhood.\textsuperscript{16} Although not available via an open data portal, GIS data is typically available by request, and Parnell and Reed successfully obtained the needed GIS data through a direct request to Muskingum County officials, and it was provided in a standard, machine-readable format. GIS data uses spatio-temporal location as the key index variable. Parnell explained that for most municipalities, one needs to first fill out a form requesting access to such GIS data, but that there is a “wealth of data available if you know how or where to ask for it.”\textsuperscript{17} Essentially, the nature of GIS data allows users to analyze and interpret data in ways that make it easier to identify, manipulate and understand relationships, patterns and trends, and then visualize that data in forms that are accessible for anyone (data expert or not) to understand and share (e.g., maps, globes, reports and charts).\textsuperscript{18} Parnell, being experienced with GIS data, recognizes the opportunities GIS data presents more readily than those with less experience, such as attorneys. However, should the open data movement continue to grow, more people of all backgrounds could recognize and use GIS data.

Parcel data – which identified all occupied houses in the study areas, the location of water lines with dates of construction, Zanesville’s city limits and the street locations – provided the backbone of the case. Additionally, Relman, Dane & Colfax obtained water-billing data, which provided the addresses of all houses with public water service. With this data in hand, Parnell’s team of paralegals undertook a door-to-door effort to: a) confirm that property identified in the parcel data was an occupied house; and b) determine racial composition and how long each resident had lived in that location in order to determine that “there was no difference between the people with and without water other than race.”\textsuperscript{19}

Using the public GIS data, the household survey information, the plaintiff information and the addresses of houses with billed water service, Parnell’s colleague, Ben Marsh, Ph.D., Professor of Geography and Environmental Studies at Bucknell University, built the GIS layers for the maps showing a clear pattern of racial discrimination. Parnell wrote the expert report used

\begin{quote}
“The easier it is to access data, and the more people that can access data without having to pay for it, the more egalitarian society we will have.”

– Tara Ramchandani, Relman, Dane, & Colfax PLLC
\end{quote}


\textsuperscript{17} GovLab interview with Allan Parnell, Ph.D., Vice President, Cedar Grove Institute for Sustainable Communities, July 21, 2015.

\textsuperscript{18} “What is GIS?” ESRI. http://www.esri.com/what-is-gis/howgisworks

\textsuperscript{19} GovLab interview with Allan Parnell, Ph.D., Vice President, Cedar Grove Institute for Sustainable Communities, July 21, 2015.
in the case based on the maps, survey information and additional information taken from the plaintiffs. During the trial, Relman and Colfax walked the jury through the information contained in the maps by rebuilding Parnell’s maps, “layer by layer,” while explaining how each piece of information was obtained and what it uncovered about water access discrimination.20

Meanwhile, the expert testifying on behalf of the defense attorneys representing the City of Zanesville, Muskingum County and the East Muskingum Water Authority also attempted to use data-driven maps to support the opposition’s case. The defense’s expert, however, did not effectively manipulate the data or maps, leading to a mismatch between the claims made by the attorneys and the information being displayed. The defense used the GIS and census data to try and argue that race did not affect who had water service, claiming that all residents of a certain census block had water if a water line intersected any part of that census block. This is demonstrably false, and the defense was unable to prove the claim. Parnell countered the assertion using the same census data, showing that in 2000, the water lines for the block in question only served 34 African-

---


American residents – all of whom lived in a nursing home whose population was 88 percent white.\textsuperscript{22}

The use of open data, in some cases drawn from the same source, to make contradictory points from two sides of the same court case demonstrates the danger of selective, perhaps manipulative use of data to lead people to make incorrect conclusions. The defense was unable to use the data to create a compelling case, however, and, as a result, appeared “clueless.”\textsuperscript{23}

\section*{III. IMPACT}

The impact of using open data in the Kennedy v. Zanesville case can broadly be split into two categories: immediate, and long term.

The immediate impact was clear, tangible and in many respects highly positive. The use of open data (along with maps) was at the heart of a legal strategy that identified and remedied a long-standing civil rights violation. As Parnell put it: Some black residents were unaware that “for 50 years their [white] neighbor had a hot tub [while] they couldn’t turn the tap on.”\textsuperscript{24} Open data helped visualize and irrefutably identify a systematic form of discrimination that had long been woven into the fabric of daily life in Zanesville.

Significant monetary damages were also assessed:

The federal jury awarded $11 million against the City of Zanesville and the East Muskingum Water Authority for illegally denying water on the basis of race.

- The jury also awarded $80,000 in damages to Fair Housing Advocates Association, the agency that initially assisted the plaintiffs with their administrative complaints before the Ohio Civil Rights Commission.\textsuperscript{25}

All told, the plaintiffs were eligible for payments of between $15,000 and $300,000.\textsuperscript{26}

The medium and longer-term impact, while in many ways positive, is somewhat less clear. On the one hand, the use of data and maps quantified and solidified the case made by the plaintiffs that their neighborhood had long suffered discrimination. Although the city of Zanesville had completed laying water pipes before the jury reached its verdict, the data-driven case established by the legal team validated the original complaint, and potentially made it harder for municipal authorities to scale back their expanded water distribution or deny water services to other neighborhoods, as communities or attorneys could reference this case in the future.

\textsuperscript{23} GovLab interview with Allan Parnell, Ph.D., Vice President, Cedar Grove Institute for Sustainable Communities, July 21, 2015.
\textsuperscript{24} GovLab interview with Allan Parnell, Ph.D., Vice President, Cedar Grove Institute for Sustainable Communities, July 21, 2015.
\textsuperscript{25} “Kennedy v. City of Zanesville.” Relman, Dane & Colfax PLLC. http://www.relmanlaw.com/civil-rights-litigation/cases/zanesville.php
\textsuperscript{26} “Kennedy v. City of Zanesville.” Relman, Dane & Colfax PLLC. http://www.relmanlaw.com/civil-rights-litigation/cases/zanesville.php
In addition, in July 2015, The U.S. Department of Housing and Urban Development (HUD) announced new rules to the Fair Housing Act of 1968, which barred outright racial discrimination, then routine, and required active desegregation in housing.\textsuperscript{27} The new rules require “cities and towns all over the country to scrutinize their housing patterns for racial bias and to publicly report, every three to five years, the results. Communities will also have to set goals, which will be tracked over time, for how they will further reduce segregation.”\textsuperscript{28} While not a direct result of this case, the new HUD rules will create even more open data that can be applied to civil rights and fair housing cases. Law firms engaging in civil rights litigation often rely on GIS or similar open data sets for evidence in a variety of cases, including fair housing, school segregation, redevelopment and relocation, code enforcement abuse and unequal service provision.\textsuperscript{29} Data helps illuminate the local political geography, providing “a link to a diversity of policy decision information, such as the relationship of race to the proximity of superfund sites.”\textsuperscript{30} In illegal discriminatory lending cases, open data enables attorneys to determine where loans from the bank are going, and whether loans are given to minority neighborhoods or not.\textsuperscript{31} Regarding the new HUD rules, civil rights attorneys representing fair housing cases will have even more data to identify and prove patterns of discrimination, with civil-rights groups lauding the decision as “an important advancement on what’s been one of the most fraught frontiers of racial progress.”\textsuperscript{32}

Yet despite such advances, research conducted by Parnell (and others) has, somewhat paradoxically, concluded that the overall cause of open data has been little advanced in the community and among Muskingum County officials. In fact, perhaps because the case cost the county $11 million and significant negative publicity, its aftermath has actually been marked in many cases to a tightening of data availability and supply. For example, Parnell found that public GIS data from Muskingum County, which although not online, had previously been available through a simple phone or form request, became more challenging to locate and access, requiring more phone calls and forms, and, often, the skills of an attorney or someone familiar with navigating the bureaucracy.\textsuperscript{33} Parnell has also experienced similar difficulties in obtaining water and sewer information in certain California and North Carolina municipalities, with officials mandating that only necessary third parties, such as engineering companies, may access such data. The process for obtaining GIS data greatly varies across governments – by location and by level – although Parnell observed that this increase in steps to access GIS data became even more apparent after the attacks on September 11, 2001, as data custodians had more leeway in citing security concerns to delay access.\textsuperscript{34} Unfortunately, Parnell believes that

\begin{itemize}
  \item \textsuperscript{27} http://www.huduser.org/portal/affht_pt.html#final-rule
  \item \textsuperscript{29} “Expert Witness Work.” Cedar Grove Institute for Sustainable Communities. http://www.cedargroveinst.org/partners.php
  \item \textsuperscript{30} GovLab interview with Allan Parnell, Ph.D., Vice President, Cedar Grove Institute for Sustainable Communities, July 21, 2015.
  \item \textsuperscript{31} GovLab interview with Tara Ramchandani, Attorney, Relman, Dane, & Coffax PLLC, August 3, 2015.
  \item \textsuperscript{33} GovLab interview with Allan Parnell, Ph.D., Vice President, Cedar Grove Institute for Sustainable Communities, October 2, 2015.
  \item \textsuperscript{34} GovLab interview with Allan Parnell, Ph.D., Vice President, Cedar Grove Institute for Sustainable Communities, October 2, 2015.
\end{itemize}
uncovering the true motives behind these actions, if they are indeed to discourage lawsuits and/or hide potential wrong-doing, would prove extremely time- and resource-intensive.\textsuperscript{35}

In several cases, these restrictions on data were justified in the name of security. For example, county officials said that making infrastructure data more fully open, for example by publishing it online in a downloadable format, posed a potential security threat that would allow terrorists to locate targets such as water or energy plants. However, Parnell argues that since that type of location information is not particularly difficult to find without access to GIS data, should one be motivated, these policy changes may need to be re-evaluated if the result is less transparency and decreased access to public data.\textsuperscript{36}

\section*{IV. CHALLENGES}

These less-than-optimal outcomes of the Kennedy v. Zanesville case point to some recurring obstacles faced by open data advocates. In particular, the deployment of security as a justification for restricting access to information is fairly common, as several other case studies in this series indicate.

Broadly, the Zanesville case suggests three challenges to the wider release and dissemination of open data:

\textbf{Security Concerns}

After the incidents of September 11, 2001, concerns over security are frequently raised by custodians of data. These can express themselves as anxieties over national security, data security or other forms of security. These concerns are, of course, often masks for other reasons (for example, a desire to restrict negative publicity or avoid lawsuits). As Parnell, who has dealt with officials across the spectrum of transparency, in counties across the country, explains, “You’re either going to hide your data or you’re going to fix things.”

Nonetheless, putting aside the validity of security as a justification for restricting access, there are steps open data advocates can take to mitigate such concerns.

\textbf{Awareness and Usability}

Open data offers a powerful way to combat discrimination (and various other injustices), but like all technologies, it is only a tool, its potential defined by the extent to which it is usable and actually used. Repeatedly, we see instances where data is made available but lies under-utilized due to a lack of awareness or barriers to usability. A similar pattern was very much

\textsuperscript{35} GovLab interview with Allan Parnell, Ph.D., Vice President, Cedar Grove Institute for Sustainable Communities, July 21, 2015.

\textsuperscript{36} GovLab interview with Allan Parnell, Ph.D., Vice President, Cedar Grove Institute for Sustainable Communities, July 21, 2015.
evident in Zanesville, where the data finally deployed to such great success in the lawsuit had in fact been available for years but residents were unable to use it.

As Tara Ramchandani, an attorney at Relman, Dane & Colfax, explains:

How do you know if you’re being denied water service if you have to know to get the data, and then hire a lawyer to actually get and use the data? That makes it unfair. The more information that’s easily accessible [in simplified formats], the more easily you can understand what’s happening to you, and you can put your experience in context of the population around you.37

She points out that even attorneys, who may use this type of data more regularly than the average citizen, often have to rely on experts in order to identify which data sets are useful, and how to access them. In addition, open data is often most useful when combined with other forms of closed or proprietary data (in this case, for example, the free and open public GIS data was combined with a Zanesville-specific, door-to-door survey conducted by Parnell’s team to confirm the race of residents in the neighborhood). The sophisticated technical and other skills required to access and combine data are quite often out of reach for ordinary citizens. As Ramchandani puts it: “The easier it is to access data, and the more people that can access data without having to pay for it, the more egalitarian society we will have.”38

Advocates of open data therefore need to first increase awareness of open data, as “it never crosses most people’s minds that this information is out there.”39 Civil rights groups and law firms in particular should be targeted in outreach campaigns, and provided resources to learn how to access, use or find experts in open data to help support their cases.

Data – the Human Factor

Open data is a tool.40 Its true potential derives from the way it is used by humans. Data unquestionably played a key role in winning the case, but one of the plaintiffs’ attorneys from Relman, Dane & Colfax noted that testimony given at trial was an equally, if not more, important factor in winning the case than the data and mapping evidence. For example, plaintiffs describing in their own words the experiences of discrimination, the hardships of not having access to water and the disappointment in repeatedly asking for water and being denied, painted a very moving, emotional story for the jury to consider. Meanwhile, the testimony of the defendants and all the people involved in deciding to deny people water, and the investigative process of discovering discrimination patterns in decision-makers’ behavior, also proved to be a powerful narrative in the courtroom, helping the jury to more fully understand and relate to the experiences of the community.41

Therefore, while the data and maps proved to be critical, hard evidence in the plaintiffs’ case, the more visceral aspects of the case help describe the real-world impact.42 For example, many Coal

---

37 GovLab interview with Tara Ramchandani, Attorney, Relman, Dane, & Colfax PLLC, August 3, 2015.
38 GovLab interview with Tara Ramchandani, Attorney, Relman, Dane, & Colfax PLLC, August 3, 2015.
39 GovLab interview with Allan Parnell, Ph.D., Vice President, Cedar Grove Institute for Sustainable Communities, July 21, 2015.
40 GovLab interview with Allan Parnell, Ph.D., Vice President, Cedar Grove Institute for Sustainable Communities, July 21, 2015.
41 GovLab interview with Tara Ramchandani, Attorney, Relman, Dane, & Colfax PLLC, August 3, 2015.
42 GovLab interview with Tara Ramchandani, Attorney, Relman, Dane, & Colfax PLLC, August 3, 2015.
Run residents, including Doretta Hale, 74, wept the first time clean water came through her pipes, describing, “I could wash clothes whenever I wanted ... I could go out and water the flowers.”

While not a challenge experienced by the plaintiffs in this case, it is clear that, especially in situations involving advocacy or persuasion, cold, hard data can only go so far. Personal experiences can help exemplify and ground the insights uncovered through open data, perhaps making takeaways easier to understand and relate to.

V. LOOKING FORWARD

Although Muskingum County seems to have pushed back against open data in some manner at the local level as a result of this case, water services are now supplied to residents equally. More broadly, the open data movement in Ohio on the state level does seem to be growing. The Office of the Ohio Treasurer, for example, has launched OhioCheckbook.com, an interactive online tool allowing users to search and access state spending data, as part of its transparency program. Meanwhile, the state House of Representatives is currently working on a bill to launch DataOhio, an initiative that promotes open data standards and transparency in a number of ways. First, state and local agencies in Ohio will be required to adhere to an open data standard. Similar to the federal open data portal and portals established in other states, DataOhio would establish an online catalog (data.ohio.gov) to provide descriptions of data sets, tutorials and tools. To help provide the financial backing for spurring open data activity, DataOhio also calls for the disbursement of $10,000 in grants to local governments as an incentive to provide budgetary, staffing and compensation information online in an open data format using uniform accounting. The city of Cincinnati, meanwhile, has its own open data portal to “provide access to government data, improve services, increase accountability and stimulate economic activity.”

As more state-wide and large-city open data initiatives are implemented across Ohio, there is potential for trickle-down effects. Government officials in smaller localities like Zanesville could be encouraged to embrace open data and help boost awareness among residents of the broad potential impacts of making government data more publicly accessible.

45 https://data.cincinnati-oh.gov/