How a Feel-Good AI Story Went Wrong in Flint

A machine-learning model showed promising results, but city officials and their engineering contractor abandoned it.

By [Alexis C. Madrigal](#)   January 3, 2019

Workers in Flint, Michigan, replace a lead water-service pipe. (Bill Pugliano / Getty)

More than a thousand days after the water problems in Flint, Michigan, became national news, thousands of homes in the city still have lead pipes, from which the toxic metal can leach into the water supply. To remedy the problem, the lead pipes need to be replaced with safer, copper ones. That sounds straightforward, but it is a challenge to figure out which homes have lead pipes in the first place. The City's records are incomplete and inaccurate. And digging up all the pipes would be costly and time-

consuming.

That's just the kind of problem that automation is supposed to help solve. So volunteer computer scientists, with some funding from Google, designed a machine-learning model to help predict which homes were likely to have lead pipes. The artificial intelligence was supposed to help the City dig only where pipes were likely to need replacement. Through 2017, the plan was working. Workers inspected 8,833 homes, and of those, 6,228 homes had their pipes replaced—a 70 percent rate of accuracy. Heading into 2018, the City signed a big, national engineering firm, AECOM, to a $5 million contract to "accelerate" the program, holding a buoyant community meeting to herald the arrival of the cavalry in Flint.

Few cities have embarked on a pipe-replacement program nearly as ambitious, let alone those that have to deal with the effects of segregation, environmental racism, and the collapse of industry in the upper Midwest. In total, 18,786 families in Flint now know that their pipes are safe, because the City has either dug them up and confirmed that they're copper or replaced them if they were made of lead or galvanized steel. "I think things have gone extremely well," Flint Mayor Karen Weaver told me. "We're a year ahead of schedule and under budget."

But something strange happened over the course of 2018: As more and more people had their pipes evaluated in 2018, fewer and fewer inspections were finding lead pipes. In November 2017, according to meeting notes obtained by local news outlet *MLive*'s Zahra Ahmad, the city's head of public works estimated that about 10,000 of Flint's homes still had lead pipes, roughly in line with the number other experts have floated. The new contractor hasn't been efficiently locating those pipes: As of mid-December 2018, 10,531 properties had been explored and only 1,567 of those digs found lead pipes to replace. That's a lead-pipe hit rate of just 15 percent, far below the 2017 mark.

Read: The 'horrifying' consequence of lead poisoning

There are reasons for the slowdown. AECOM discarded the machine-learning model's predictions, which had guided excavations. And facing political pressure from some residents, Weaver demanded that the firm dig across the city's wards and in every house on selected blocks, rather than picking out the homes likely to have lead because of age, property type, or other characteristics that could be correlated with the pipes.

After a multimillion-dollar investment in project management, thousands of people in Flint still have homes with lead pipes, when the previous program would likely have already found and replaced them.

The declining success of the pipe-replacement program has caused critics of the City to raise the alarm. The Natural Resources Defense Council (NRDC), which represents a community group called the Concerned Pastors for Social Action, has argued in court that the City has [abrogated its court-ordered mandate](#) to get the lead pipes out as quickly as possible. If there are still thousands of homes with lead pipes and the City is doing thousands of excavations, how hasn't it found more of them? "It's the number of lead pipes removed that matters, not the number of holes dug," said Pastor Allen C. Overton, a member of Concerned Pastors for Social Action, [in an NRDC statement](#).

Before things got ugly, the effort to pull the lead pipes out of the ground was shaping up to be a high-tech feel-good story. At Google's AI for Good conference in October, the Georgia Tech computer scientist Jacob Abernethy described how a team of volunteers built the system to predict which homes were most likely to have lead pipes.

The computer scientists saw that an information problem was sitting atop the lead issue in the city. No one knew, exactly, who had lead pipes and who did not. The City had a variety of records: thousands of old cards describing parcels' hookups, and also maps and small updates that had been filed into the system over the years. But a cataloging system is only as good as its maintenance, and the City of Flint had been starved of resources for

decades.

Flint, you probably know, was a key chamber of the heart of the American automobile industry. Through the middle of the 20th century, General Motors had a variety of facilities in the area, employing some 80,000 people. As Flint's position within the automotive industry declined, most white residents took the money they'd earned and [moved to the suburbs, taking their tax dollars and capital out of the city's core](). They created their own regional services in the wealthier Genesee County, while Flint's residents suffered the repercussions of an economy that had moved on: budget cuts, failing schools, and, of course, post-industrial environmental problems. It is not a surprise, then, that before the crisis began, auditing and correcting water-department records from the early-20th century were not top of mind for city officials.

When Flint's money woes got bad enough in the wake of the housing collapse, Michigan Governor Rick Snyder sent in an "emergency" manager to enact cost-cutting measures. Half of Michigan's black residents have lived under an emergency manager, [according to a Michigan Civil Rights Commission report about Flint](). It was Flint's emergency manager who made the call to switch the water supply from the Detroit water system to the Flint River in April 2014 without putting in the right corrosion controls. That's what started the problem.

Many cities share the lead-pipe problem and the informational obstacles layered atop it. The decay of infrastructure built decades ago is not only in the metal, but in the data cataloging that lets the city's government and residents understand the state of the water system. For all the talk of "smart" cities, the real state of play in many older places is that no one even thinks of these things until there's a disaster. People have been saying "[America is 1,000 Flints]()" since the city was booming, and it is still true. Just as there are thousands of lead service lines in Flint, there are something like 6 million lead service lines in America.

When Weaver launched the program to replace Flint's lead service lines, Fast Start, in March 2016, suddenly the city's maintenance debt came back up to the surface. General Michael McDaniel was picked to lead the program, with less than a handful of people working under him.

Some basic things were known about the lead-pipe distribution: The pipes were most likely to be found in postwar homes, built when Flint experienced major expansions, and least likely to be found in newer homes. In February 2016, Martin Kaufman at the University of Michigan at Flint built [some maps of nominal lead pipe placements in the city using City records](). McDaniel's team used them to prioritize initial excavations based on the age of homes and the Department of Environmental Quality's rough sense of where the worst water problems were. Then they asked themselves who would be the most affected by lead in the water. "The very young, the very old, and those with compromised immune systems," McDaniel told me. They determined which homes had kids under 5 years old and adults over 70.

Combining these sources gave them a rough sense of where to start. McDaniel set out to replace 600 lead pipes each in 10 small zones. "It was a matter of what was efficient and what was equitable across the city," he said.

When Abernethy and his collaborator, the University of Michigan's Eric Schwartz, got involved over the summer of 2016, they saw a familiar type of prediction problem: sequential decision making under uncertain conditions. The crews didn't have perfect information, but they still needed the best possible answer to the question Where do we dig next? The results of each new dig could be fed back into the model, improving its accuracy.

Initially, they had little data. In March 2016, only 36 homes had had their pipes excavated. And even as the crews began to do hundreds of digs, they were looking for lead pipes, which meant that they were creating a decidedly unrepresentative sample of the city. Using just that data, the model was likely to overpredict how much lead existed elsewhere in Flint. So the University of Michigan team asked Fast Start to check lines across the city

using a cheaper system called "hydrovacing," which uses jets of water, instead of a backhoe, to expose pipes. The data from those cheaper excavations went back into the model, allowing the researchers to predict different zones of the city more accurately.

Read: How to prevent the next Flint

As they refined their work, they found that the three most significant determinants of the likelihood of having lead pipes were the age, value, and location of a home. More important, their model became highly accurate at predicting where lead was most likely to be found, and through 2017, the contractors' hit rate in finding lead pipes increased. "We ended up considerably above an 80 percent [accuracy] for the last few months of 2017," McDaniel told me.

In late 2017, Weaver announced that the City was awarding a $5 million contract to AECOM, the major national contractor, to run the project. In February 2018, the City held a community forum to "really introduce you to the company that's going to accelerate Fast Start," as Weaver put it. Robert Bincsik, Flint's director of public works, noted at the forum that the City was doing something nearly unprecedented. "There is not anybody else doing this as aggressively as we are," Bincsik said. "Overall, I think we've done a wonderful job."

AECOM's published plans said it intended to "efficiently identify and replace 6,000 [lead service lines] per year." This goal made sense, as the small ragtag and mostly volunteer management team in 2017 had identified and replaced more than 6,000 service lines.

The contractor's process, as laid out at that community meeting, would consist of two steps. First, it would hydrovac in 10 zones laid out by the contractor. Then, after the nature of the pipes was determined, it would go out and replace the lead and galvanized-steel pipes. Bincsik extolled the virtues of hydrovacing: It was cheaper and faster, less intrusive, and created

a lower risk of damaging pipes. Hydrovacing cost $300 or less. Digging up the pipes in a traditional way cost several times more, according to contractor invoices from the 2017 phase of the project—at least $2,500, and as much as $5,000 depending on the type of pipes dug up and replaced.

AECOM's team, however, struggled before it even started. In late October 2018, the project manager, Alan Wong, told me that the problems started during the transition between McDaniel's team and AECOM. Wong's crew was supposed to begin work in October 2017, when McDaniel's contract ended. But AECOM's deal was not actually signed until December 28, 2017. There was no overlap between the teams. "We would have had October, November, and all of December," Wong told me. "We would have been able to mesh, to have a reasonable transition. It didn't work out."

Furthermore, AECOM does not appear to have considered the predictive model central to the project. According to a court declaration, after seemingly positive initial discussions, Schwartz, from the University of Michigan, sent five emails to Wong from January through May 2018, none of which was answered. Wong told me that all his company had was a "heat map" of the city—like an image—but Schwartz said his own team had offered its database, which consisted of individual lead-probability scores for every single address in the city.

AECOM basically approached the problem new, as if other people had not been successfully hammering away at it since June 2016. It discovered, as others had before, that the data the City possessed were neither wholly digitized nor wholly accurate. Wong says the company doing the digitization work pro bono, Captricity, was supposed to be done in January but did not finish until May.

At the same time, Weaver asked AECOM to explore all over the city, in each of the city-council wards. The city administration "did not want to have to explain to a councilperson why there was no work in their district," Wong said. So AECOM created 10 zones spread across the whole city, initially

assigning 600 addresses in each area to contractors.

The problem is that lead pipes are not evenly distributed across the city. When evaluated by any available tool—the actual amount of lead pipes that had been found, the predictions from the University of Michigan model, what the city records said, historical knowledge of construction practices—it was clear that the lead was concentrated in a few areas, mostly in the older places in the core of the city, such as the Fifth Ward, and not in the outer regions, such as the Second, Fourth, or Seventh Wards.
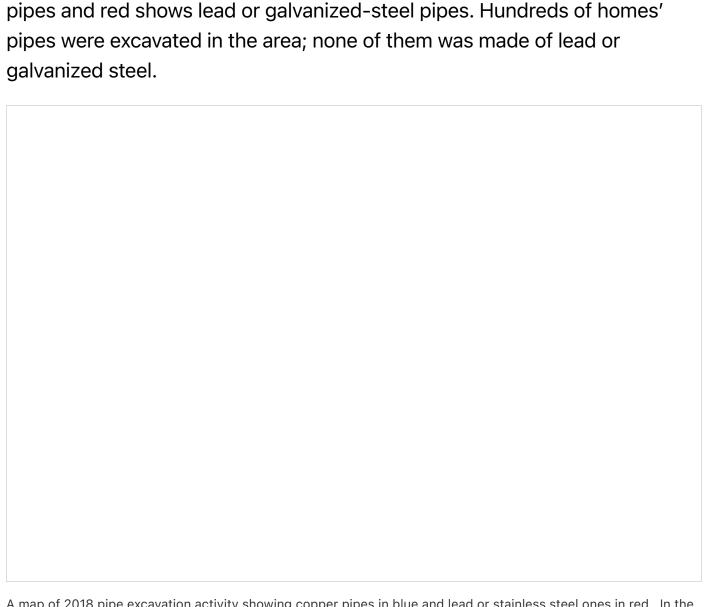
Read: How do you regulate a self-improving algorithm?

Then, in the middle of 2018, some lead was found in pipes that had otherwise seemed to be made of copper. Hydrovacing generally makes a smaller hole than when a backhoe is involved, which had allowed some lead bits to go unnoticed. The mayor made a decision to abandon hydrovacing, opting instead for the gold-standard traditional method. "You get a 100 percent guarantee and that's what we're worth," Weaver told me. Given that AECOM had planned to hydrovac all over the city as a means of identifying lead, that change threw a kink into the company's plans.

Other changes were also afoot. The mayor made a decision to excavate every house in areas where program officials thought they might find lead, rather than skipping over homes that the model indicated probably didn't have lead pipes. "When we started this, people would say, 'You did my neighbor's house and you didn't do mine,'" Weaver said.

"The City did not want to leave anybody behind," Wong told me.

That makes political sense, but it has serious implications for not just the cost of the remediation project, but the speed at which the project could extract the remaining lead service lines in the city. In the outer regions of Flint, block after block of homes were excavated and no lead was found, as in the eastern block of Zone 10, seen below, where blue represents copper

pipes and red shows lead or galvanized-steel pipes. Hundreds of homes' pipes were excavated in the area; none of them was made of lead or galvanized steel.
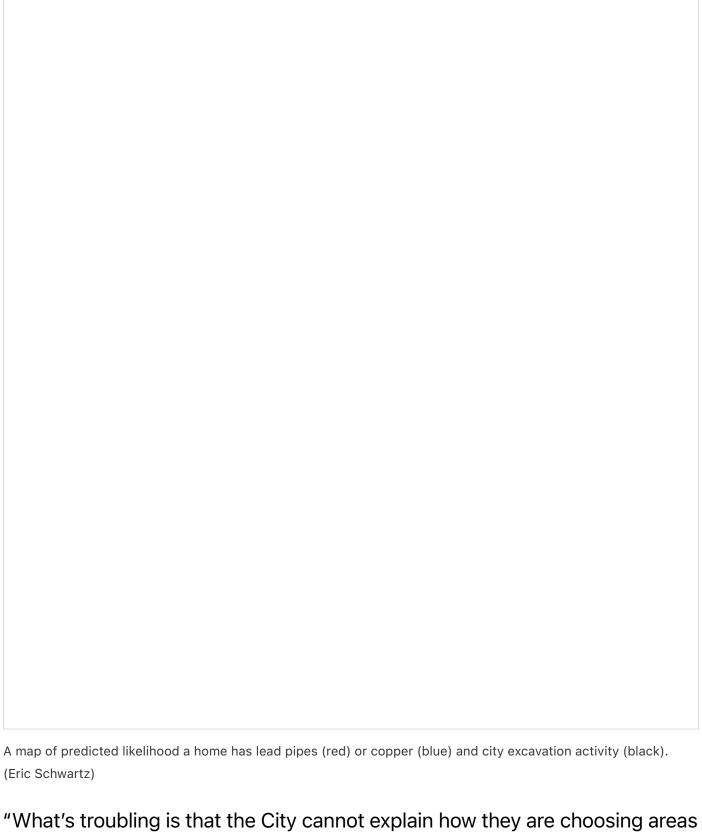

A map of 2018 pipe excavation activity showing copper pipes in blue and lead or stainless steel ones in red. In the three highlighted areas, contractors excavated large numbers of homes and found little or no lead service lines. (City of Flint)

A new directive had begun to guide the program: to excavate, by the most intensive means, every single active water account in the city. Otherwise, citizens could always wonder if they had lead pipes and didn't know it. The program managers would have to tell people, "You'll have to trust a computer model," Wong told me. "The citizens are just not going to trust that."

There are reasonable explanations for why AECOM's hit rate would be lower than the 2017 team's. McDaniel worked in the areas of city with the highest concentrations of lead, and his team generally followed the model's

predictions. AECOM and the City went to work across Flint and did every house along certain blocks. Furthermore, there are fewer lead service lines in the city than originally estimated. Early approximations assumed that 20,000 to 30,000 city pipes were made of lead or galvanized steel. That figure proved too high.

However, the NRDC, which has been suing the City over the way it has conducted the program, still argues that the core priority of its settlement agreement—lead removal—was abandoned. Even given the factors above, the rate at which contractors are finding lead has fallen too precipitously to be explained by reasonable logistical changes to the program. This has had the effect of keeping lead pipes attached to people's homes for longer than is absolutely necessary.

In a [court filing](#), Schwartz estimates that between 4,964 and 6,119 homes with hazardous lines remain in the city. The map below shows, in red, where the AI researchers predict a greater than 90 percent likelihood that hazardous pipes are installed. Blue indicates areas highly unlikely to have lead or steel pipes. The little black dots are where AECOM's team has done work in 2018, as of November. If the model is even generally correct, casual inspection suggests that the work isn't being targeted at the areas most likely to have water lines in need of replacement.

A map of predicted likelihood a home has lead pipes (red) or copper (blue) and city excavation activity (black). (Eric Schwartz)

"What's troubling is that the City cannot explain how they are choosing areas to dig," says Dimple Chaudhary, an NRDC attorney. "You do have this model that is doing a pretty good job of describing 'Here there is lead.' And that model says they are excavating in the wrong places."

To take the most prominent example, the Fifth Ward is expected to have the most remaining lead. The University of Michigan model estimates that crews would find lead 80 percent of the time in that area. Yet from January to August 2018, AECOM contractors did the fewest excavations there, carrying out 163 excavations in the ward out of 3,774 total in the city. They found lead pipes in 156 of those digs—96 percent of them. Meanwhile, over the same time period in the Second Ward, 1,220 homes were investigated and lead was found in 46 of them, just a four percent hit rate. AECOM did the most digging in the two wards that Schwartz and Abernethy's model predicted had the smallest percentage of lead pipes, and the results bore out the predictions of the model.

Looking at this data, the State, which reimburses contractors for their work, has said it is going to suspend payments to the City because of how the program has been managed. "The City made a policy decision to stop prioritizing excavations at homes where lead or galvanized steel service lines were expected to be found," the Department of the Attorney General alleged. Now, the City, the NRDC, the State, and AECOM are negotiating to return to the machine-learning model that was used in 2017. AECOM's contract has been renewed, and appears to include a return to the model. An additional $1.1 million has been allocated to the firm for future work.

City officials have made a good-faith attempt at implementing an ambitious, difficult program. Weaver made important decisions that she saw as protecting the health and safety of all her city's residents. AECOM claims it has done the best it could. But good faith notwithstanding, a heartbreaking fact can't be ignored: Simply continuing the 2017 program's method might have pulled nearly all the remaining lead out of the city during 2018. Instead, thousands of people got the peace of mind that comes with knowing they have copper lines. But others who are more likely to have lead lines that could leach poison into their drinking water will have to wait for digging to commence again to learn for sure.

And that's assuming that the battle between the City and the State about reimbursements doesn't get settled in the State's favor, depriving residents of the support necessary to complete the pipe-replacement project. This tragedy already has more acts than anyone wants to recount, and the stage is now set for yet another one to begin.