



Data Analytics and the Fight against Housing Blight: A Guide for Local Leaders

Citation

Pough, Bradley and Qian Wan. 2007. Data Analytics and the Fight Against Housing Blight: A Guide for Local Leaders. Responsive Communities (March 2017).

Published Version

<https://cyber.harvard.edu/publications/2017/03/DataAnalytics>

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:31852257>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

DATA ANALYTICS AND THE FIGHT AGAINST HOUSING BLIGHT: A GUIDE FOR LOCAL LEADERS

Bradley Pough and Qian Wan

As cities across the country publicly grapple with the challenges associated with a recent rise in urban in-migration,¹ countless others silently struggle through the seemingly intractable problems of urban disinvestment, population loss, and, most relevant to this paper, housing blight. From mid-western Rust-Belt cities working through a painful economic transition,² to southern Sun-Belt locales still pulling themselves out of a deep foreclosure crisis,³ numerous American municipalities now find themselves asking the same hard question: How do we prevent our housing stock from falling into further disrepair? This question cuts to the very core of what it means to be a city, as housing dictates so much of our communities' financial, demographic, and economic fortunes.

For many cities, digital technology and the promise of big data offer potential solutions to the challenges of housing blight. As the cost of computing continues to decrease, local governments – sorely strapped for cash – are turning in masse to the world of analytics to address the problems that traditional human capital cannot. Cities like New Orleans and Detroit have illustrated for the rest of the country that, when used correctly, data analytics can be a powerful tool in the fight to preemptively target, and ultimately ameliorate blight. Now, other cities are looking to replicate their success, hoping to find in their own data the answers to their housing woes.

This paper will serve as a guide for cities inter-

ested in employing data analytics as a way to combat housing blight. Relying on multiple case studies as well as interviews with experts in the field, this report provides local decisionmakers with a set of recommendations for how they can develop a data solution tailored to their local needs. Part I introduces the reader to the issue of blight and discusses why it has proven so vexing for so many cities. Part II describes the role data can play in helping to target housing blight, and introduces the various case studies from which this paper pulls its recommendations. Part III closes with a set of five action steps that cities interested in developing a blight-targeting data solution can follow to get their project off the ground.

I. Background – What is Housing Blight and why is it a Problem?

One of the most challenging aspects in developing an effective solution for addressing housing blight is first settling on a definition of what housing blight is. Unfortunately, cities and states across the country define housing blight differently⁴ – driven in part by varying legal structures (at all levels of government) that incentivize the over, or under, categorization of housing as blight.⁵ While this paper stresses the need for local leaders to think critically about how they define their blight, it is crucial for readers to start with a baseline understanding of where blight definitions come from and what forces shape those categorizations.

The Evolution of the Concept of Blight

Although the idea of using housing degradation as a justification for urban redevelopment seems natural today, for much of our cities' histories the issue of blight went unmentioned.⁶ However,

⁴ See Martin E. Gould & Lynne B. Sagalyn, *The Use and Abuse of Blight in Eminent Domain*, 38 FORDHAM URB. L.J. 1119, 1124 (2011) (discussing how “[f]ederalism . . . prevailed in the definition and determination of blight[.]”).

⁵ See, e.g., Christopher S. Brown, *Blinded by the Blight: A Search for a Workable Definition of “Blight” in Ohio*, 73 U. CIN. L. REV. 207, 210 (2005) (arguing that Ohio’s blight definitions were incentivized, in part, by Urban Renewal programs and the desire to exercise eminent domain).

⁶ Gould & Sagalyn, *supra* note 4, at 1121.

¹ See Lucy Westcott, *More Americans Moving to Cities, Reversing the Suburban Exodus*, THE ATLANTIC (March 27, 2014) <http://www.theatlantic.com/national/archive/2014/03/more-americans-moving-to-cities-reversing-the-suburban-exodus/359714/>.

² See Brian Louis, *Rust Belt Cities Demolish Homes as Defaults Blight Neighborhoods*, THE SEATTLE TIMES (November 19, 2010) <http://www.seattletimes.com/business/real-estate/rust-belt-cities-demolish-homes-as-defaults-blight-neighborhoods/>.

³ See Alison Fitzgerald & Jared Bennet, “Zombie” Homes Haunt Florida Neighborhoods, THE CENTER FOR PUBLIC INTEGRITY (September 15, 2014) <https://www.publicintegrity.org/2014/09/15/15519/zombie-homes-haunt-florida-neighborhoods>.

Suggested Citation: Pough, Bradley and Qian Wan. *Data Analytics and the Fight Against Housing Blight: A Guide for Local Leaders*. (March 2017). Responsive Communities. Available at: <http://cyber.harvard.edu/publications/2017/03/DataAnalytics>

starting in the 1920's, several pioneering localities searching for a legal precondition for slum clearance began tying their states' eminent domain powers to the existence of unlivable housing conditions in certain neighborhoods.⁷ The federal government followed soon after, and in 1949 passed the Housing Act which, in part, authorized and funded the razing of slums and the construction of middle and upper-income dwellings in their place.⁸ Interestingly, although the federal government assumed as much as two-thirds of the cost of redeveloping these neighborhoods, Congress refrained from defining which neighborhoods would qualify for this kind of large-scale intervention.⁹ Instead, they left that decision to the states and localities, allowing federalisms laboratories to produce widely divergent definitions of what qualified as a "blighted area[]."¹⁰

Attorney Hudson Hayes Luce performed a 50 state canvass of the various blight statutes, breaking the laws into 12 overarching categories of common blight characteristics.¹¹ According to his research, all states, to one degree or another, define blight as the existence of structural defects and health hazards in a house or dwelling.¹² Beyond that, however, the definitions varied greatly. For example, approximately two-thirds of states defined a blighted home as one that lacked sufficient condition of title (e.g., ownership was vested in too many people at once).¹³ One quarter of states defined blight by the character of the neighborhood surrounding a dwelling.¹⁴ And just under a fifth of states defined blight by housing occupancy (i.e., whether or not a dwelling was vacant).¹⁵

This tapestry of blight definitions is driven by varying legal regimes that incentivize localities to categorize homes as blighted as a way of facilitating the transfer of property from a private party to either the government or to a developer. For a city like New York, whose laws allow it

to acquire title for tax-delinquent properties,¹⁶ characterizing tax-delinquency as blight might be in their political best interest. However, in a city like New Orleans, which auctions off debt from tax-delinquent properties to private investors,¹⁷ targeting tax-delinquency through blight efforts may have the unintended consequence of increasing vacancy rates in homes that are now owned by private banks or investment funds. Although this paper passes no judgments on the normative value of either of these legal regimes, they help illustrate how different legal motivations can drive the way cities define and ultimately target blight.

These widely divergent definitions illustrate that as a legal concept, blight has no universally agreed upon definition, prompting historian Colin Gordon to say that "blight has lost any substantive meaning as a description of urban conditions or a target for public policy."¹⁸ Although numerous states have attempted to reign in these subjective categorizations through a variety of methods, few are convinced that our state and local governments are approaching consensus on this topic.¹⁹

Why Blight Is a Problem

Regardless of the specific definition, blight has proven to be a challenging foe for many cities struggling with urban disinvestment and population loss. Often characterized a "disease," housing blight seems to spread through a city like cancer, starting out too small to be perceptible, and slowly expanding until it has grown into something too large to be controlled. This growth brings with it a whole host of negative externalities impacting cities' financial and physical well-being. Coupled with the fact that for many places blight is notoriously difficult to track, it is no wonder that cities are eager to explore new and innovative solutions to addressing this almost century-old problem.

7 *Id.*

8 *Id.*

9 *Id.* at 1123-24.

10 *Id.* at 1124.

11 Hudson Hayes Luce, *The Meaning of Blight: A Survey of Statutory and Case Law*, 35 *REAL PROPERTY, PROBATE, AND TRUST J.* 389, 394-96 (2000).

12 *Id.* at 397.

13 *Id.* at 399.

14 *Id.*

15 *Id.* at 402.

16 See, Student Note, *In Rem Foreclosure as a Cure to Tax Delinquency*, 26 *ST. JOHN'S L. REV.* 283, 288-91 (2013) (describing New York's process for in rem foreclosures on delinquent properties).

17 Interview with Oliver Wise, Director of Office of Performance and Accountability, City of New Orleans, (October 31, 2016) (arguing that New Orleans's process of transferring liens to private investors restricts the City's ability to adequately enforce its building code).

18 Colin Gordon, *Blighting the Way: Urban Renewal, Economic Development, and the Elusive Definition of Blight*, 31 *FORDHAM URB. L. J.* 305, 307 (2004).

19 See Gould & Sagalyn, *supra* note 4, at 1128.

One of the primary reasons why blight is so problematic is that it can potentially depress the home values and marketability of nonblighted neighboring properties. One study performed in Cuyohoga County, Ohio found that each additional blighted house within 500 feet of a nonblighted house would reduce the nonblighted house's selling price by between 1 and 2%.²⁰ This makes intuitive sense. If blighted homes are often magnets for crime and other dangerous activity, potential buyers would likely be less interested in purchasing near blight without a significant discount.

An additional reason why blight weighs so heavily on some cities is the way it impacts municipal revenues.²¹ Cities are funded predominantly by property taxes, which are assessed based on the market value of homes.²² If blight does in fact depress home values, then it would also have the second order effect of depressing a city's tax revenues. Furthermore, blight often co-presents (and is sometimes coterminous) with vacancy.²³ If no one is living in a home, then it is likely that no one is paying taxes on that home, further eroding the city's revenue base. Studies support this theory,²⁴ illustrating how blight impacts not only its immediate neighborhood, but its entire city's ability to operate and adequately provide services.

Finally, blight can create a host of additional social challenges for cities that are often already struggling to address the needs of their low-income populations. Blighted neighborhoods are frequently associated with higher crime rates,²⁵

more juvenile delinquency,²⁶ increased rates of disease,²⁷ more frequent fires,²⁸ and decreased social cohesion.²⁹ While one could argue that many of these problems may have actually preceded the blight (and potentially caused it), few would contend that allowing blight to fester would do anything but exacerbate these social issues.

A Taxonomy of the Blight Remediation Landscape

Although we typically speak about blight as a government issue, blight problems implicate entire cities, and therefore require combatants from all sectors of public and private life to address it. The following list attempts to outline the various actors involved in the world of blight remediation in urban centers, although it is in no way exhaustive:

City Governments: City governments are the undeniable leaders in the fight against blight. Because the problem impacts their cities, their finances, and their built environments so acutely, they have a clear stake in addressing this problem as best they can. Municipal efforts take a variety of forms, many of which are outlined in greater detail later in this paper. However, some efforts include blight tracking projects, coalition building, punitive code enforcement measures, rehabilitation, resale, and when all other avenues have been exhausted, demolition.³⁰

Research Non-profits and Universities: Research non-profits and universities serve instrumental roles in the fight against urban blight. While they may not have the same kind of financial resources or physical presence of a government actor, they add value in two other ways

20 Stephen Whitaker & Thomas J. Fitzpatrick IV, *Deconstructing distressed-property spillovers: The effects of vacant, tax-delinquent, and foreclosed properties in housing*, 22 J. OF HOUSING ECON. 79, 91 (2013). submarkets

21 See U.S. Housing and Urban Development, *Vacant and Abandoned Properties: Turning Liabilities Into Assets*, EVIDENCE MATTERS 1, 2 (2014) (Research links foreclosed, vacant, and abandoned properties with reduced property values . . . and increased costs for municipal governments.”).

22 See, e.g., *Where do Texas Cities get Their Money*, TEXAS MUNICIPAL LEAGUE 1 <http://www.tml.org/HCW/WhereCitiesGetMoney.pdf>.

23 See, e.g., Luce, *supra* note 11, at 402 (referencing cities that defined housing blight by vacancy).

24 See, e.g., Tri-COG Collaborative, *Financial Impact Analysis on Blight 4* (2013) http://housingalliancepa.org/wp-content/uploads/2014/01/FinancialImpactofBlightonTriCG_FullReport.pdf (stating that blight caused an over \$8 million loss in tax revenue for TRI-COG cities).

25 See Erin Graves & Elizabeth Shuey, *The Social Impact of Home Rehabilitation in Low-Income Neighborhoods*, FEDERAL RESERVE BANK OF BOSTON 1, 4-5 (2013) (referencing the research on the relationship between housing vacancy and crime.).

26 See generally, Philip Lloyd Jackson, *A SPATIAL ANALYSIS OF URBAN BLIGHT AND JUVENILE DELINQUENCY IN PHOENIX, ARIZONA*, 1968 (1970).

27 See Wendell E. Pritchett, *The “Public Menace” of Blight: Urban Renewal and the Private Uses of Eminent Domain*, 21. YALE L. & POL’Y REV. 1, 45 (2003).

28 See, e.g., Jeremy Brilliant, *Vacant Houses Attract Crime, Add to Blight*, WTHR (September 14, 2015) <http://www.wthr.com/article/vacant-houses-attract-crime-add-to-blight> (attributing a string of potential arson outbreaks to housing blight).

29 See Graves & Shuey, *supra* note 25 (arguing that although there has been little empirical research on the matter, there is strong theoretical evidence that home abandonment strains neighborhood social relations.).

30 For a detailed example of a city government exploring multiple ways to address blight, see generally, *Every Neighborhood has a Future and it Doesn’t Include Blight*, Detroit Blight Removal Task Force, (May 2014).

– longevity and expertise.³¹ An unfortunate reality about government-led initiatives is that they have a tendency to end after an administration leaves office. Research institutions do not have that problem, and can therefore see blight remediation efforts through various mayoral (or gubernatorial) administrations. Similarly, government actors often struggle to attract the kind of human capital with the skill set necessary for advanced analytics on blight. Research institutions by their very nature do not have that problem, and have therefore proven invaluable in providing governments with the requisite empirical data to focus their blight-fighting efforts.

Community Development Organizations (CDC's): CDC's are non-profit organizations that serve a wide range of functions, all predominantly aimed at revitalizing underserved communities.³² CDC's typically act as lightning rods for government and private sector investment, and subsequently redistribute those dollars in the form of investment in affordable housing, commercial property, health centers, and educational facilities. CDC's have been instrumental in the fight against blight for many cities, as they have helped catalyze neighborhoods that have fallen into, or were teetering on the edge of, a blighted state.

Residents: In a lot of ways residents exist on the front lines of a city's fight against blight. More than perhaps any other actor, they confront the blight every day during their normal routines. Many cities rely on resident's 311 or (unfortunately) 911 calls as their primary mechanism for tracking blight within their borders.³³ Now that cities are turning to digital technology to address blight, many governments are deputizing their residents to "crowdsource" what they see regarding blight using smartphone apps.³⁴ Finally, residents are critical sources of

political capital for motivating city governments to act on blight. According to New Orleans's Oliver Wise, part of the reason why the City opens all of its blight strategy sessions to the public is to hold itself accountable for the problem.³⁵ As long as citizens are allowed to access these discussions, they can press their elected officials for more resources or more action when they see little movement on the blight problems in their communities.

While each of these actors plays a slightly different role in the fight against urban blight, the mechanics of that fight have similar themes from one city to the next. Regardless of the locality, cities typically have only one of three general options for how to deal with their blight:

Rehabilitation: For most cities this is by far the most attractive option. During the early stages of blight, many houses have not fallen so far into disrepair that they cannot be fixed. Homes with minor code violations or other visible, structural maladies can typically be repaired with limited financial expenditures. Often, rehabilitation is simply a matter of "nudging" homeowners to take some restorative action, either through some kind of punitive measure, or simply by way of reminder.³⁶

Resale: This can take numerous forms but, at a high-level, resale typically entails the government acquiring an abandoned property and selling it to a new owner (often after significant rehabilitation). This process is often complicated for municipal or state governments, especially when the laws around acquiring title to blighted property are not on their sides. Additionally, resale does not guarantee that the property will not fall back into disrepair, so governments typically try to go to great lengths to ensure that the property is going into good hands. Programs like New Orleans Lot Next Door target homeowners of houses that share a common boundary with a blighted property, in hopes that someone with a stake in the community, and a proven track record of responsibility, will turn out to be a good

31 But see Karen Chapple and Miram Zuk, *Forewarned: The Use of Neighborhood Early Warning Systems for Gentrification and Displacement*, CITYSCAPE 1, 6 (arguing that part of the reason why University-led early warning programs that tracked blight were discontinued was that they lacked administrative longevity.).

32 See Overview: Community Development Organizations (CDCs), Community Wealth.Org <http://community-wealth.org/strategies/panel/cdcs/index.html>.

33 See, e.g., Jason Sheuh, *Detroit Revitalizes City with 311 App*, GOVTECH (May 7, 2015) <http://www.govtech.com/dc/articles/Detroit-Revitalizes-City-with-311-App.html>.

34 See, e.g., Anna Clark, *Will "Blexting" Help with Detroit Blight?*, NEXT CITY (July 28, 2014) <https://nextcity.org/daily/entry/detroit-blight-blexting-houses-motor-city-mapping>.

35 Interview with Oliver Wise, Director of Office of Performance and Accountability, City of New Orleans, (October 31, 2016)

36 For an example of this "nudging" in action, see Katherine Hillenbrand, *New Orleans Brings Data-Driven Tools to Blight-Remediation*, DATA SMART CITY SOLUTIONS (October 12, 2016) (discussing New Orleans pilot effort to use letters to "nudge" property owners into compliance with housing code).

steward of the property in question.³⁷

Demolition: Typically the last recourse for addressing blight, demolition occurs when a property has fallen into such a state of disrepair that there is no hope for rehabilitation. Cities often try to avoid this avenue for multiple reasons. First, demolition is expensive. According to Professor Margaret Dewar, in Detroit, the demolition of one property can cost up to \$15,000.³⁸ Second, demolition permanently takes a property off a city's tax rolls. Finally, while the demolition of one blighted building can often be beneficial for a community, too many demolitions in the same area can drain a neighborhood of a lot of architectural vibrancy and character.³⁹ For these reasons cities typically do everything in their power to avoid demolishing a property, but are often left with no other choice.

II. Using Data to Fight Blight

Why Should Cities use Data in the Fight Against Blight?

Outside of the socioeconomic externalities, blight also poses a logistical challenge for cities. For most cities, blight is reported on an ad hoc basis, either through complaints from neighbors, intermittent code inspections, or the random sighting by a city official. The reality is that for many of these cash-strapped cities, cataloging the tens of thousands of homes within their borders that have fallen into disrepair would be a Herculean effort. For them, data-driven solutions, which drastically reduce the time, effort, and resources necessary for targeting blight, are potentially their only avenues for redress against this pernicious problem.

Data-driven solutions can help cities fight blight in two primary ways. First, good data can expand a city's geographic reach as they attempt

to catalogue the blight within their borders. As was stated previously, many cities currently rely on word of mouth (311 calls, etc.) and eyes on the street (deployed inspectors, vigilant elected officials, etc.) as their primary mechanism for discovering when a house has fallen (or is about to fall) into blight.⁴⁰ For small cities with few blight problems, this might be sufficient. But for larger cities struggling with blight across a dispersed geography, simply waiting for a blighted property to come across their desks will not be enough to get a handle on the scope of the problem. For these larger cities (Detroit, New Orleans, etc.) big data analytics that use blight proxies – like utility usage and mail receipt – have proven instrumental in helping shed light on blight wherever it occurs.⁴¹

Second, data analytics can help a city begin remediation tactics earlier in a property's "blight cycle." Currently, too many cities are alerted to a blighted property at the point when demolition is the only real solution. The sooner government officials know about a building's degradation, the sooner they can utilize other tactics in their tool kits (rehabilitation, resale, etc.) to effect change. Data analytics are making this kind of early detection a real possibility.⁴² Now, instead of waiting to hear about blight only after a building has fallen into an irreparable state, analysts can look to past data for early warning signals of blight. Indicators like minor code violations and small abnormalities in utility payments may not signal the kind of blight that requires a major intervention, but they might presage a level of inattentiveness that, if caught early, could prevent more intractable problems from taking root.

Data is also being used in less intuitive ways. For example, in New Orleans government officials currently employ a lengthy review process to determine if a blighted property should be rehabilitated, resold, or demolished. Recognizing that there might be discernable commonalities

37 See Greg LaRose, *Find out if NOLA Will Sell the 'Lot Next Door' to you*, NOLA (January 25, 2016) http://www.nola.com/business/index.ssf/2016/01/find_out_if_nora_is_selling_th.html.

38 Interview with Margaret Dewar, Professor of Urban and Regional Planning, University of Michigan (November 3, 2016).

39 See Steve Lockwood, *Stop the Demolitions, or If All You Have is a Bulldozer, Every House Looks Like a Pile of Sticks*, SMART CITY MEMPHIS (March 10, 2015) <http://www.smartcitymemphis.com/2015/03/stop-the-demolitions-or-if-all-you-have-is-a-bulldozer-every-house-looks-like-a-pile-of-sticks/> ("Scraping a house clean costs about \$4,000, but in low-value communities will generally result in an overgrown vacant lot for decades to come. No one will be coming to build a new house there anytime soon.").

40 See Sheuh, *supra* note 34.

41 See Interview with Chad Kenney, Director of Development System Performance, City & County of Denver (October 28, 2016) (discussing how water usage was a great indicator of housing blight); Interview with Oliver Wise, Director of Office of Performance and Accountability, City of New Orleans, (October 31, 2016) (discussing how in the past, the city relied on USPS data to quantify New Orleans' overall blight).

42 See e.g., Reyes, Helsby, et. al., *Early detection of properties at risk of Blight Using Spatiotemporal data*, DATA SCIENCE FOR SOCIAL GOOD (2015).

between the kinds of buildings that received each treatment, data scientists developed a tool that could “predict” the outcome of the City’s review process before it began. Although the City still relies on the human-driven review process for its final determination, government officials can turn to this tool to help validate their choices. As cities get a better handle on what their data can be used for, we will likely see a proliferation of these types of cases.

How Data Is Used Today

Although using computer-driven algorithms to target blight is a relatively new tactic, cities and other researchers have been employing empirical analyses and statistics to measure neighborhood quality for decades. Starting in the 1980’s, Neighborhood Early Warning Systems (NEWS) pooled newly available neighborhood and parcel level data to track and predict neighborhood change in cities like Chicago, Los Angeles, Philadelphia, and Minneapolis.⁴³ These initiatives looked for correlative relationships by gathering a set of indicator variables (e.g., crime, tax delinquencies, code violations, utility shut-offs, etc.) and assessing their value in predicting outcomes like housing degradation or neighborhood decline.⁴⁴ Unfortunately, NEWS were typically run out of universities or non-profits, and often had little interaction with the governments of the cities they sought to better understand.⁴⁵ Perhaps due to this disconnection from city government, the NEWS initiatives were gradually discontinued,⁴⁶ but still clearly serve as the intellectual precursors to the municipally driven data initiatives proliferating today.

Recent advances in technology have prompted cities (as well as other institutions) to pick up where the universities’ NEWS initiatives left off in using data analytics to track and predict housing blight. These advances include:

Geographic Information Systems (GIS):

One of the earliest innovations that spurred the reinvestment in blight-focused data analytics was the proliferation of GIS – a system designed to capture, store, analyze, and depict geo-spa-

tial information in a digital format.⁴⁷ By allowing users to code data geo-spatially, GIS added a new dimension to potential blight analyses. Now, instead of having to focus solely on indicator variables that mapped on to a particular parcel or census tract, analysts could better explore the way neighborhood variables influenced the likelihood a house would fall into blight. How, for example, did a house’s proximity to crime influence its future blight potential? Did houses with high densities of crime within a 500-foot radius experience blight at significantly higher rates than those with that were, instead, 1000-feet removed from serious crime? Although these kinds of questions may have been answerable prior to GIS’ proliferation, the advent of GIS certainly made answering them (and depicting that answer) much easier.

Crowdsourcing and the Proliferation of Smartphones:

An additional influential innovation has been the recent expansion of smartphone ownership and crowdsourcing apps. As previously stated, one of blight’s major complications is how difficult it is for a single city government to track. It is a geographically dispersed problem that moves in unpredictable ways. For a cash-strapped city, canvassing all of the homes within its borders with enough frequency to keep an accurate account of its blight problems would be impossible. Instead, many cities are now able to delegate that canvassing job to the people most acutely impacted by blight – their citizens. Inspired by private apps like See-Click-Fix, multiple cities have created tools that their residents can download to send photographs and impressions of housing level blight instantly.⁴⁸ This added data source has proven invaluable to cities, not only in that it augments their knowledge of where blight is occurring, but also in how it sheds light on which blight occurrences weigh most heavily on city residents.

Open Data Initiatives: The push for municipalities to open up their data to the public has helped the fight against blight in two significant ways. First, by opening up their data, cities are now granting countless independent actors the

⁴³ See Chapple & Zuk, *supra* note 31, at 3-4.

⁴⁴ For a list of indicator variables see *id.* at 5-6.

⁴⁵ See *id.* at 6.

⁴⁶ See *id.*

⁴⁷ See *What is GIS*, ESRI <http://www.esri.com/what-is-gis>.

⁴⁸ See, e.g. Clark, *supra* note 32; Michelle M. Thompson, *The City of New Orleans Blight Fight: Using GIS Technology to Integrate Local Knowledge*, 22 HOUSING POL’Y DEBATE 101 (2011).

chance to perform the kind of rigorous analytics projects that governments often cannot. This “data democratization” has effectively allowed cities to expand their pools of human capital without having to hire new employees or contract with expensive third parties.⁴⁹ Second, often by going through the exercise of publishing open data sets, cities are finding they are better equipped to deal with data questions in-house. Open data initiatives force cities to take stock of what data they have available and push them to take on the often arduous task of cleaning the data and coding it in such a way that it is ready for analytic research.⁵⁰ By the end of that process, if done right, data sets from the cities’ police departments, fire departments, and code enforcement agencies should be able to “talk” to each other, making a data-driven blight initiative much easier to undertake.

Machine learning: As the amount of data collected increase, the relationship between each dataset also becomes complex and difficult to discern by eye or intuition. Machine learning methods usually accompany data-driven analytics, as they can extract complex relationships between datasets, and make highly informed decisions on classification and predictions of blight status. Most machine learning algorithms used today are for identifying houses that are already blighted but not yet reported. However, they could potentially be used to predict blight before it occurs. The theory for prediction is simple. Instead of using particular variables as proxies for current blight (i.e. water being shut off), the algorithm culls past datasets to see which factors historically predicted blight at some future state. Researchers have taken various approaches to achieving this end and are beginning to produce results that, if replicable, could help cities stop blight before it ever occurs.⁵¹

⁴⁹ See Alex Howard, *President Obama’s New Open Data Initiative Could Help Cities Help Themselves*, TECH REPUBLIC (March 8, 2016) <http://www.techrepublic.com/article/president-obamas-new-open-data-initiative-could-help-cities-help-themselves/> (“To paraphrase President John F. Kennedy, they’ve challenged the nation to ask not what your country can code for you, but ask what you can code for your country.”).

⁵⁰ See Victoria Lemieux, *Why We are Failing to get the Most out of Open Data*, WORLD ECONOMIC FORUM (September 29, 2014) <https://www.weforum.org/agenda/2014/09/open-data-information-governance-quality/> (arguing that open data initiatives require serious data curation before they can be successful).

⁵¹ See e.g., Reyes, Helsby, et. al., *supra* note 46.

Case Studies

Today, several cities, universities, and non-profits around the country are partnering to employ these technological advances in new and exciting ways. From New Orleans us of a form of machine learning to inform the way it determines blight remediation strategies, to Detroit’s use of crowdsourced data to track its growing blight problem, cities are turning to data en masse as one of the primary solutions to their blight struggles. Now, as many of this movement’s pioneers are beginning to see results from these initiatives, it’s important that researchers take a step back and ask: what has worked, what has not, and where do we go from here?

Below are four case studies from which this paper pulls its recommendations. It’s important to note that these case studies weren’t chosen because they paint a cohesive picture of what a data solution for blight should look like. Indeed, each of these initiatives takes a different approach to the fight against blight, employing data in sometimes drastically different ways (some predictive, others not). Instead, these initiatives were chosen because of their diversity, with hopes that they’ll serve as an examples of what is possible for cities interested in using data analytics to get their blight under control.

New Orleans: Although New Orleans blight-fighting efforts are colloquially known as Blight Stat, according to Director of Performance and Accountability Oliver Wise, Blight Stat technically only refers to a “series of management meetings” where city managers and public citizens share insights regarding the success of those blight-fighting efforts.⁵² But New Orleans blight programs extend far beyond Blight Stat, and include the City’s efforts to develop an information technology system for tracking code enforcement data (LAMA),⁵³ develop a public website for tracking blighted properties (Blight Status),⁵⁴ promulgate punitive responses to code violations (Sheriff Sales, demolition, etc.),⁵⁵ employ predictive analytics to determine how best to respond to a blighted parcel,⁵⁶ and several others.

⁵² Interview with Oliver Wise, Director of Office of Performance and Accountability, City of New Orleans, (October 31, 2016)

⁵³ *Blight Reduction Report*, City of New Orleans 13 (January 2014).

⁵⁴ *Id.*

⁵⁵ *Id.* at 14.

⁵⁶ See Katherine Hillenbrand, *New Orleans Brings Data-Driven Tools*

By themselves, these efforts may have not been enough to significantly reduce the City's blight numbers. However, in concert they were able to improve the condition of New Orleans's housing stock considerably. Just between September 2010 and April of 2013, the City was able to reduce the number of blighted properties within its borders by over 10,000 homes – including over 4,000 units that had been demolished and over 2,000 that were brought into compliance with the City's code enforcement regulations through the administrative hearing process.⁵⁷ These numbers helped Mayor Landrieu achieve (and surpass) his goal of reducing the number of blighted properties in the city by 10,000 – a goal that was initially described as “aggressive” when it was first made.⁵⁸

According to Director Wise, getting a handle on the City's data was instrumental for New Orleans in achieving its 10,000-home target.⁵⁹ Unlike other cities, New Orleans's primary difficulty was not quantifying how much blight they already had, but figuring out how what to do with that blight once they found it. To achieve this, the City (across multiple departments) worked to organize its code enforcement data, align the efforts of departmental managers, and develop a predictive analytics tool that “learned” which cases were best suited for demolition as opposed to rehabilitation.⁶⁰ Additionally, New Orleans relied on data to make the blight-fighting process more transparent, helping to pull the public into an exercise that excluded them for years. This effort included partnering with Esri to develop a crowdsourcing app for blight tracking, as well as the development of Blight Status – a website created to visually depict all of the blight in an area and provide citizens insight about what was being done to address it.⁶¹

Detroit: Similar to New Orleans, Detroit has

historically been characterized as one of the “most extreme” examples of a blighted municipality in the United States.⁶² According to Professor Dewar, much of Detroit's problems have stemmed from the lack of “attention” and “investment” from the City's government.⁶³ As the City declined, so too did the capacity for its government to address difficult problems like housing degradation.⁶⁴ Today, the City is forced to address blight across a highly dispersed geographic area with significantly fewer resources than other cities.

Unlike New Orleans, Detroit's difficulties with blight stem both from their inability to accurately track their blight problem, as well as their difficulty determining how to address the blight once they've found it. For this reason, the City of Detroit (in partnership with non-profit Data Driven Detroit and start-up Loveland Technology)⁶⁵ has developed a two-pronged, data driven approach to fighting their blight, first focusing on accurately tracking the issue, and then turning to how best to address it:

Detroit's Tracking Initiative: Known as the Motor City Mapping project, Detroit's data driven tracking efforts first aim to empower local residents with the development of a crowdsourcing app.⁶⁶ Through this application, citizens are encouraged to “Blext”⁶⁷ images and impressions of neighboring housing degradation, helping city officials get a handle on their blight problem without having to send inspectors into the field.⁶⁸ Next, data analysts employed by the city work to clean and standardize the information sourced from the Blexting app.⁶⁹ This data is finally merged with other in-house sources of data (such as fire department data, or data on code violations) to create a fuller picture of the City's blight landscape.⁷⁰

Detroit's Efforts to Address Blight: Once the City

to Blight-Remediation, DATA SMART CITY SOLUTIONS (October 12, 2016) (describing New Orleans's program for using data from past code enforcement hearings to “predict” the outcomes of future hearings. This helps adjudicators know which remediation strategy has most often been used in situations similar to the one on which they have to decide.”).

57 *Blight Reduction Report*, City of New Orleans 21-22 (January 2014).

58 Frank Donze, *Mayor Mitch Landrieu sets Goal of Cleaning 10,000 Blighted Eyesores*, NOLA (October 1, 2010) http://www.nola.com/politics/index.ssf/2010/10/mayor_mitch_landrieu_sets_goal.html.

59 Interview with Oliver Wise, Director of Office of Performance and Accountability, City of New Orleans, (October 31, 2016)

60 *Id.*

61 *Id.*

62 Interview with Margaret Dewar, Professor of Urban and Regional Planning, University of Michigan (November 3, 2016).

63 *Id.*

64 *Id.*

65 For a complete list of the groups partnering with the city in their blight-fighting endeavors, see *Every Neighborhood has a Future and it Doesn't Include Blight*, Detroit Blight Removal Task Force 3 (May 2014).

66 *Id.* at 41.

67 A portmanteau of “blight” and “text.” *Id.* at 39.

68 *Id.* at 41.

69 *Id.* at 43.

70 *Id.*

has determined where their blight is located, they must then determine how best to fight it. One component of this effort is the Maximizing Community Impact (MCI) tool, an algorithm that helps Detroit prioritize which neighborhoods are at the greatest need for blight remediation.⁷¹ MCI pulls together two risk indicators – one illustrating how blighted a neighborhood is (Neighborhood Dynamics Indicator), the next illustrating how at risk that neighborhood is to “tipping” further into disrepair (Neighborhood Threat Index).⁷² This tool recognizes that not all blighted neighborhoods are created equally. Some neighborhoods, though horribly blighted, are not at much risk of getting worse. Others, though less blighted in real terms, may deserve more attention because they have much further to fall.

As of 2014, 30% of Detroit’s buildings (almost 80,000 of them) were dilapidated and over 100,000 were vacant.⁷³ Given these numbers, the Detroit Blight Removal Task Force recommended that the city spend nearly \$850 million to tear down or repair 40,000 homes across the city.⁷⁴ MCM, MCI, and numerous other efforts have been employed to help the city reach their 40,000-home goal. Although it is still too early to judge the efficacy of those programs, the City’s early commitment to data-driven solutions illustrates how integral these tools have become in fighting blight around the country.

Cincinnati: Last year, Ed Cunningham (the head of Cincinnati’s Building Code Enforcement Office), chanced upon a young data scientist named Chad Kenney during his time at the organization Data Science for Social Good (DSSG).⁷⁵ At the time, Kenney was discussing the possibility that cities could use legacy data on blight to develop an algorithm that could predict blight moving forward.⁷⁶ The organization was prepared to put together a team to test this

hypothesis, and was looking for a city to partner with as they rolled out this pilot.⁷⁷ Cunningham bit, and soon after, the partnership between the City of Cincinnati and DSSG was born.

The team’s idea was relatively simple. Using almost 50 variables from across three years of data (inspections data, crime data, housing value data, etc.) the team “taught” an algorithm to find the trends that best predicted whether a home would fall into blight.⁷⁸ Then, using code enforcement data from the fourth year, the team tested if the algorithm could accurately predict whether a code inspector would inspect a home that was actually blighted, or one that was not.⁷⁹ The algorithm correctly predicted 78% of the serious code violations that occurred during that fourth year of inspections.⁸⁰ Compared with the code inspector’s 53% success rate, this number represented a significant improvement over Cincinnati’s original blight-tracking methodology.⁸¹

Given how recently this pilot program occurred, Cincinnati has not yet been able to operationalize what they have learned.⁸² With that said, according to Chad Kenney, this project offers cities several key insights into how they too can construct a predictive model for finding blight. Kenney stressed how important it is for cities to be able to gather the requisite data inputs in one place in a machine-readable format.⁸³ For many cities, this will represent the most serious obstacle to constructing this kind of predictive algorithm, as cleaning, curating, and sharing data across departments can be an immensely expensive and time-consuming process.

City Nexus: City Nexus represents a slightly different model for using data to track and address blight. Not driven by a particular city, City Nexus instead is a municipally agnostic tool that provides any city government the chance to

71 See *How does this Information Guide us? TIME TO END BLIGHT* <http://report.timetoendblight.org/guide/>.

72 *Id.*

73 Monica Davey, *Detroit Urged to Tear Down 40,000 Buildings*, N.Y. TIMES (May 27, 2014) https://www.nytimes.com/2014/05/28/us/detroit-task-force-says-blight-cleanup-will-cost-850-million.html?_r=0.

74 *Id.*

75 See Susan Crawford, *Bluegrass, Blight, and the Future of Cities: How a Fiddler and an Astrophysicist Introduced Predictive Analytics to Cincinnati*, BACKCHANNEL (September 30, 2015) <https://backchannel.com/how-a-fiddler-and-an-astrophysicist-brought-predictive-analytics-to-cincinnati-c5583468f4cc#.3ph7tev17>.

76 *Id.*

77 *Id.* See also Reyes, Helsby, et. al., *Early detection of properties at risk of Blight Using Spatiotemporal data*, DATA SCIENCE FOR SOCIAL GOOD (2015).

78 See Crawford, *supra* note 80.

79 See *id.*

80 *Id.*

81 *Id.*

82 As of a year ago, the City was looking for a full-time employee to move this project forward. See *Data Scientist: City of Cincinnati Office of Performance and Data Analytics*, City of Cincinnati <https://www.kaggle.com/jobs/16868/city-of-cincinnati-office-of-performance-and-data-analytics-data-scientist>.

83 Interview with Chad Kenney, Director of Development System Performance, City & County of Denver (October 28, 2016).

glean insights about blight from whatever data they have at their disposal.⁸⁴ Started by a team of computer programmers affiliated with the Harvard Kennedy School, the tool pulls in whichever data sets a city chooses, cleans them, then gives city officials the opportunity to “weight” those variables by however much they feel is necessary.⁸⁵ After weighting their variables, the city officials can run the algorithm, which would produce a list of the most at-risk homes within the city’s borders.⁸⁶

The key insight of City Nexus’s creator Sean Alaback is that cities and data scientists are still pretty far away from knowing what the best variables are for predicting blight.⁸⁷ Coupled with the knowledge that those “best” variables will likely change from one city to the next, Alaback understood that the most powerful blight-prediction tool had to be flexible enough to work in any city context, with any collection of data sets.⁸⁸ Additionally, Alaback recognized that this tool should put city officials in the driver’s seat, and allow them to iterate and adjust inputs based on a changing blight landscape, or changing political needs.⁸⁹ So far, this model has proven attractive for several cities: Salem, Fitchburg, Chelsea, Lawrence, Winthrop have all partnered with City Nexus to begin tackling their blight problems.⁹⁰

III. Recommendations for Cities

Big-data driven analytics have the potential to improve municipal blight remediation significantly. The following section gives a series of recommendations for local governments that plans to use data analytics to streamline their blight amelioration efforts.

Aggregate Municipal Data

The first step to any analysis is having quality data. Unfortunately, information that is historically used to find blight, such as water data, postal inactivity, fire code and health code vio-

lations,⁹¹ are all kept by different departments, under different formats. This means a significant amount of an analyst’s effort will be spent gathering and cleaning data instead of performing analyses. Additionally, when databases are built on project-by-project bases, they are usually only organized and updated in an ad-hoc manner, which means they can be difficult to share and distribute. They also become outdated once the project expires, forcing future programmers to redeploy the same efforts when other projects call for similar data.

Therefore, it is imperative that the first steps any local government take is to centralize, digitize, and standardize existing information collected on home properties, as well as allocate permanent staff positions to maintain this database. The data can include 311 reports, 911 calls, municipal utilities, postal activity, inspection results from fire and health departments, property titles, tax records, and house market values. Over 40 cities and counties in the U.S. are already putting together some version of an open data portal for public access.⁹² Many are specific to housing, including New Orleans’s BlightSTAT, Ohio’s NEO CANDO, the private company Loveland Technologies, etc.

The goal of this data center is to make information easily accessible and searchable, and to guarantee the sources and quality of information in a format that is both machine and human readable. Once the data is collected in one place, even simple analytics can reveal relationships that are too complex to notice by eye.

This by no means is an easy step. Understandably, not all departments are ready to share their information. Encouraging and prepping them to do so will take political prowess. Standardizing and organizing data will need someone who knows not only the data science, but also the meaning behind it. To avoid being stuck in the investment phase without seeing any returns, there are intermediate steps that can help move the development process forward. For example, a city might consider collecting data from a few highly relevant departments, or

84 Interview with Sean Alaback, Founder of City Nexus (November 10). For more information on City Nexus, see Sandra Larson, *Silo-busting Data-analytics Help Mass. Cities Tackle Vacant Properties*, NEXT CITY (October 14, 2016) <https://nextcity.org/daily/entry/blight-data-tools-mass-cities-vacant-properties>.

85 Interview with Sean Alaback, Founder of City Nexus (November 10).

86 *Id.*

87 *Id.*

88 *Id.*

89 *Id.*

90 *Id.*

91 Interview with Chad Kenney, Director of Development System Performance, City & County of Denver (October 28, 2016)

92 See *Open Data in the United State*, DATA.GOV <https://www.data.gov/open-gov/>.

beginning with a database that is only useable internally.

Although few cities are likely ready to remove humans from the blight-fighting equation entirely, data-driven analysis is still advantageous in that it can rank targets by risks and allocate resources accordingly. Additionally, as cities collect new data, algorithms can be updated easily, greatly expanding the reach of these analytic programs. Finally, it is important to remember that data collection reaps dividends outside the world of blight policy. Maintaining a robust data bank can provide the foundation for other cost-effective municipal projects in the future.

Fight Blight Through Data

Define Blight

As discussed earlier, there is no universal definition of blight, and it would be unproductive to insist on one – the existing definition for each city or state reflects its own unique history, politics, and legal priorities. However, for the purposes of data analytics, a common flaw of existing definitions is that they are often open-ended, vague, and subjective.⁹³ This makes it difficult to evaluate blight severity, prioritize interventions, measure progress, and weigh the contribution of different factors.

A definition that is useful for data analysis should be concrete, quantitative, and measurable. This will help enforce targeted, objective, and consistent data collection, so that all the properties can be evaluated by the same criteria and ranked against each other. When desired criteria are not easily measurable, inferences should be made through data that is available or easily measurable. For example, the structural integrity of the property is considered in all U.S. states in blight definition.⁹⁴ But private homes, regardless of their blight level, rarely have frequent structural inspection records. However, records such as fire-inspection results, home-appraisal-inspection results, or perhaps even frequency and result of termite inspections, can be directly or highly correlated with structural integrity, and therefore can be used as a proxy for de-

sired structural integrity data.

Additionally, cities should consider their remediation strategies when crafting a definition. For example, if, for whatever reason, demolition were the only operation available to a city government, there would be little point to defining blight as anything other than the most severe cases. Therefore, a city's operating definition of blight should strike a balance between the traditional policy considerations, the available data, and the potential strategies for remediation.

Find and Rank Blight

The most common method for finding blight is to physically inspect the outside of the houses either by driving through a neighborhood or on foot. This kind of ground-level inspection can be very comprehensive and consistent. But it requires a significant expenditure of human capital, which translates to battles for funding and time. Many cities, such as Memphis and Detroit, have responded to this reality by constructing a digital blight map, but that too requires a significant investment of time and resources to develop successfully. A less labor-intensive way of collecting physically inspected data is to mobilize citizens and let them send in reports of blighted properties, such as through 311 calls, an app, or a website. However, while crowdsourcing strategies can reduce the need for government-paid staffers to inspect all the houses, the data received from average citizens can be biased (depending on the caller) or inconsistently evaluations.

Hence, while crowdsourcing can be very valuable for gathering objective information such as addresses, in order to leverage the crowd for more serious evaluations, there needs to be quality control for standards. New Orleans devised a crowdsourcing property-evaluation strategy that both saved man power and guaranteed a high level of data consistency. First they photographed all of the houses through a Google-map style drive through the city, distributed the photographs on an app, then asked the citizens to evaluate the properties. They made sure the evaluation questions were as objective as possible, which included “count the numbers of broken windows” or “is there litter on the

⁹³ See Gould & Sagalyn, *supra* note 4, at 1124.

⁹⁴ See Luce, *supra* not 11.

premise?”⁹⁵ This way, the City was able to guarantee a level of objectivity and consistency to the data it received.

A more sophisticated way of utilizing large data is to be able to identify blight properties even if they have not yet been reported or inspected. This is where machine learning algorithms shine. They use real data to make educated guesses about the likelihood a house will fall into blight at some point in the future. Building such an algorithm starts with using properties that have already been identified as blighted (most likely through ground-level inspection), and label these houses as “training data.” Then using a series of optimization methods, parameters (i.e. code violations, tax records, etc.) that are most consistently associated with these blighted houses can be identified. Properties that are not already labeled as blighted but exhibit similar values for those parameters will then be considered “at-risk” for future blight. The same algorithms can usually also output a confidence level regarding their conclusion, so that inspectors can then be sent out to target the properties that are most confidently described as at risk of blight. If it turns out the guess is incorrect, then that property can be marked as a negative example, which is just as valuable as positive examples, and analysts can run a new round of analysis to update the algorithm.⁹⁶

Regardless of which type of analytics a city uses, there is still need to collect data from ground-level inspections from time to time. This helps to establish some baseline for objective truth that can be used for training and validating algorithms. As mentioned earlier, the burden in this step can be significantly reduced if residents are engaged to help with reporting, evaluating, and inspecting their own neighborhoods. Therefore, outreach, information sharing, and educating citizens on the importance of blight amelioration efforts are all still indispensable.

Address and Preempt Blight

As expressed previously, remediation methods

can be as simple as nudging the property owners through a notification letter,⁹⁷ or some level of rehabilitation, reselling, and eventually demolition if all else fails. Additionally, depending on the legal structures of the cities, there might be other intermediate steps before operating on the house, such as transferring titles from private owner to the city.⁹⁸

Regardless the steps, an attractive feature in data-driven analysis is that with very little overhead, emphasis can be placed on different parameters to influence the outcome. For example, in Detroit where lack of title bottlenecks the blight-amelioration process,⁹⁹ a blight-targeting algorithm could select for houses by ownership, producing a list of houses that are both at risk of blight and carry the requisite title. Cities with other priorities can use different weights, even for the same set of parameters, but the outcome will be customized to each city’s needs.

This is an example where the definition of blight and the available intervention methods are complementary --- a city’s definition can be influenced by the steps they can take to change things, and how they define blight may also govern what intervention methods they choose to use. Hence the definition combines what is intuitively important, what is measurable, and what can be done reliably.

One step further than identifying existing blight is to predict which houses are heading toward it. Catching blight early can minimize cost associated with heavy restoration, demolition, and declining of housing values. Not many cities are there yet. It’s still unclear whether it is realistic to make predictions on house-by-house basis. Nevertheless, neighborhood level predictions could still prove valuable given that some solutions for fighting blight involve investment in public infrastructure,¹⁰⁰ which are beneficial on the neighborhood level.

⁹⁵ Interview with Oliver Wise, Director of Office of Performance and Accountability, City of New Orleans, (October 31, 2016)

⁹⁶ See, e.g., Ben Green, Alejandra Cara, et. al., Mining Administrative Data to Spur Urban Revitalization, presented at 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (August 10, 2015) <http://scholar.harvard.edu/files/bggreen/files/kdd2015.pdf>.

⁹⁷ See Hillenbrand, *supra* note 40.

⁹⁸ Interview with Margaret Dewar, Professor of Urban and Regional Planning, University of Michigan (November 3, 2016).

⁹⁹ Interview with Margaret Dewar, Professor of Urban and Regional Planning, University of Michigan (November 3, 2016).

¹⁰⁰ See, e.g., Alva James Johnson, *The Fight Against Blight: City seeks investors to redevelop former Wade Street crime ‘hole’*, LEDGER-ENQUIRER (September 7, 2013) <http://www.ledger-enquirer.com/news/local/article29304352.html> [discussing how infrastructure investment was part of the strategy for revitalizing this neighborhood].

Evaluate Progress

It is important to have an unbiased evaluation method to measure progress and to check the efficacy of the algorithms. It is easy to get caught up with improving the numbers and forgetting that the original definition is only an approximation or a reflection of the underlying problems. If not careful, the results can become a self-fulfilling cycle that doesn't produce any progress. For example, if an algorithm found that blight is highly correlated with houses that don't have their front doors painted, then as part of the amelioration effort, the city can paint all the doors. The result will pass evaluation with flying colors because now there are no houses with unpainted doors. Mission accomplished? Clearly not. While perhaps the paint is a reliable indication of a bigger problem and therefore picked up by the algorithm as an important parameter, painting just a door as a solution will not truly alleviate any real blight condition (unless, of course, the house has no other problem except the unpainted door).

Therefore, it is important to keep the definitions flexible, expect many iterations of fine tuning, make sure qualitative judgements to check the soundness of the quantitative measurements are used, and periodically collect real world examples to validate the accuracy of the algorithm.

Data Analytics Pitfalls

Big data analysis may be more powerful than almost any tools we have ever had, but it is not without faults. It is crucial to understand the potential pitfalls of using big data analytics, because using a powerful tool incorrectly can cause devastating mistakes:

Biased or incomplete sampling: Your analysis can only be as good as the data you have, and biased or incomplete data can severely skew results. For example, a weather forecast will never predict rain if it has never seen rain in its data, or the amount of overgrowth that would qualify a house to be blighted in the inner city may not be applicable for houses in rural, wooded areas. Hence, a good practice for data collection is making sure the sample taken spans a range that is representative of the city at question, or at the very least, addresses the

type of blight that city is targeting. This means taking a small but targeted sample of the neighborhoods, types of properties, levels of severity, and any other parameter that may be important but has a diverse range in the city.

Confusing correlation with causation: Some data are measuring parameters that are directly causing blight, such as a property's structural code violation --- the definition of blight almost always involves some version of structural instability.¹⁰¹ But many parameters will be a proxy, such as the frequency with which crime is associated with a property --- crime does not cause blight, but its occurrence is highly correlated with blighted houses. In fact, one can argue blight is a cause for attracting crimes. This is a small but important distinction: reducing blight may reduce the crime rate associated with the property, but reducing crime surrounding that property may not directly solve blight. Understanding the difference between correlation and causation is the key to understand which of the many problems to attack to reduce blight. Too narrow a focus on the proxy data can potentially distract the amelioration effort or create a self-confirmation loop that does not relieve actual blight.

Missing the big picture: Even with best practices for data analytics, the best solutions to ameliorate blight can be a moving target, rendering the analyses ineffective over time. When New Orleans first started its attack on blight, it focused on reducing the number of overall blight and did so by addressing each property one by one. It worked well for a while, but as studies later showed, in some cases, it is potentially more profitable to "bind the parcels together to make land available for a larger impact."¹⁰² Hence, data analysis should not be done in a vacuum, and its results should be regularly measured not just by definition of blight, but by its ability to build safe and welcoming communities for the residents.

Unintentional discrimination: Since machine learning is essentially looking for patterns in existing data, it is not unlikely that race or

¹⁰¹ See Gould & Sagalyn, *supra* note 4, at 1125.

¹⁰² See Richard A. Webster, *New Orleans Should Rethink its Blight Strategy, Study Says*, NOLA (March 18, 2016) http://www.nola.com/politics/index.ssf/2016/03/new_orleans_needs_to_rethink_b.html.

class could be found to be highly correlated with blighted houses. Whether this is intentional discrimination is still under debate,¹⁰³ but blight amelioration efforts, especially data-based blight identification, are relatively low risk when it comes to systemically putting a certain class or race of citizens at a greater disadvantage, because a city could validate whatever unintended racialized proxy data they use by physically verifying that the house is blighted. Blight prediction can be a bit more controversial, for example if the prediction leads to a certain population to be inspected more frequently. In these cases, governments are encouraged to use as much “nudging” as possible when the property is not yet exhibiting obvious signs of blight, and only interfere when the evidence of blight becomes clear.

Conclusion

As the case studies and recommendations illustrate, how our cities should best use data to fight the scourge of blight is far from settled. City leaders across the country are experimenting with various tactics – from crowdsourcing to predictive analytics – in hopes that they will find an inexpensive and effective solution to their particular housing woes. For that reason, recommending one strategy or type of technology would be irresponsible. Data-driven solutions should be multi-faceted, flexible, and above all, grounded in the neighborhoods they aim to serve. There is little doubt that the data analytics landscape will change drastically over the next few years. What today feels like cutting edge technology will soon be eclipsed by something unforeseen and unimagined. But, by adhering to the overarching principles of good data management, flexibility, and appreciation for local context, city leaders can rest comfortably knowing that they are prepared for the future’s curve balls.

¹⁰³ For an example of how this debate is playing out in another field, see Julia Angwin, et. al., *Machine Bias*, PROPUBLICA (May 23, 2016) <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.