Making Smart Cities Explainable: What XAI Can Learn from the "Ghost Map"

Shubhangi Gupta
shubhangi@gatech.edu
Georgia Institute of Technology
Atlanta, Georgia, USA

Yanni Alexander Loukissas
yanni.loukissas@lmc.gatech.edu
Georgia Institute of Technology
Atlanta, Georgia, USA

ABSTRACT

How can we visualize civic algorithms in ways that illuminate both their positive and negative spatial impacts? Civic algorithms guide everyday decisions that cumulatively create city life. Yet, their broader effects remain invisible to their creators and city inhabitants. Recent scholarship on "algorithmic harms" presents an urgent need to make smart cities explainable. We argue that existing Explainable AI (XAI) approaches are limited across four important dimensions: *accessibility, cultural reflexivity, situatedness, and visibility into internal representations*. Our research explores the potential of conventional *maps* in addressing these limits and providing what we call "grounded explanations". As a salient example, we harness the historical case of the "Ghost Map", designed by John Snow to visualize and resolve the 1854 London Cholera epidemic. We believe that such examples can help the XAI community learn from the cultural history of city representations, as they seek to establish public processes for explaining and evaluating "smart cities".

CCS CONCEPTS

• **Human-centered computing** → *Visualization application domains*; **Information visualization**.

KEYWORDS

Explainable AI, Algorithmic audits, Transparency, Smart city, Maps, Visualization

ACM Reference Format:
Shubhangi Gupta and Yanni Alexander Loukissas. 2023. Making Smart Cities Explainable: What XAI Can Learn from the "Ghost Map". In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems (CHI EA '23), April 23–28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3544549.3585847

1 INTRODUCTION

In the past decade, several new algorithms meant to give rise to "smart cities" [52] have been deployed around the world, with a variety of effects that are still not well understood. These algorithms draw on spatial and temporal data to construct models of cities and shape the experiences of their inhabitants [26]. Generally speaking, such algorithms can work to strengthen, maintain, and challenge preexisting power relations within cities [12].

During the same time period, researchers in Human-Computer Interaction (HCI) and related fields have sought to identify instances where algorithms uphold social relations that lead to discrimination and injustice. Sociologists such as Zukin et al. have demonstrated how geographically coded Yelp reviews reinforce prejudice against neighborhoods with people of color thereby influencing civic investments and contributing to processes of urban change such as gentrification. The algorithmic moderation of reviews is not public, and users are unaware of their role in affecting capital flows [65]. Filter bubbles created by Zillow reinforce existing spatial power structures along the axes of race and class. Even as the goal of the site is to give users more control over their home-buying experience, it leaves users unaware of the implications of their own filtering decisions [29]. Safe walking apps, such as Safetipin, risk segregating neighborhoods by attempting to advance the safety of one social group while marginalizing another. Once again, knowledge about the failings of social structures and policy normalized by the app is inaccessible to citizens, as well as the app creators [19]. There is an urgent need to explain the discriminatory logics underlying data creation, aggregation, and consumption and invite public debate and evaluation.

To address these concerns, scholars from fields such as critical data and algorithmic studies are calling for advancements in algorithmic transparency, explainability, and auditing. Their goal is to improve visibility into the workings of algorithms, opening them up for critique by both experts and everyday users. HCI and Machine Learning (ML) researchers have proposed numerous techniques for auditing algorithms for algorithmic justice. However, these audits are generally designed for technical internal use. Meanwhile, there is a growing community of activists and end users who do have adequate tools for understanding, comparing, and representing these algorithms for themselves. This leaves little room for evaluating how the complex world we live in is simplified to be represented in the design of algorithms— who is included, who is excluded, and how cities are quantified and aggregated for computation.

In this paper, we ask: how can the spatial impacts of "smart city" algorithms be explained effectively and ethically for the benefit of their inhabitants? To answer this question, we first analyze existing approaches to explain algorithms and lay out four primary limitations to these approaches. We argue that these limitations are well addressed by one popular visualization method for cities—*maps*. The limits of existing explainable AI (XAI) approaches and the opportunities presented by maps to address those are listed below:

CHI EA '23, April 23–28, 2023, Hamburg, Germany

ACM ISBN 978-1-4503-9422-2/23/04.
https://doi.org/10.1145/3544549.3585847

(1) Accessibility: XAI methods can be incomprehensible for everyday users. On the other hand, maps are well understood—even treated with affection—by a broad spectrum of audiences.
(2) Cultural reflexivity: XAI methods are commonly not representative of the social and political factors that shape algorithms and their designers whereas maps signify their own context of production through their visual languages that are culturally rooted.
(3) Situatedness: XAI methods tend to be removed from situated real-world contexts and experiences of city inhabitants. Maps, particularly when they are created in a participatory fashion, can draw on local knowledge of people and places in their making.
(4) Visibility into internal representations: XAI methods focus on explaining the abstract relationship between input and output variables while overlooking the impact of their internal representations of cities. Maps reveal how their makers conceptualize cities and the spatial components that comprise them.

While these limits of XAI and affordances offered by maps to address them are not exhaustive, they are a good starting point to explore if and how maps can serve as useful tools to explain geospatial civic algorithms. We explore the opportunities presented by maps as stated above by drawing on the work of Dr. John Snow, a physician-geographer, who traced the spread of Cholera in the 1854 London epidemic through a map [50], popularly termed the 'Ghost Map' [24]. This map is amongst the most noteworthy examples of how mapping can serve as a technique to effectively and accessibly explain social phenomena.

We present the related work in visual approaches to XAI and a description of the Ghost Map in the following section. Next, we lay out existing XAI approaches more broadly, their limits, and the opportunities offered by maps. Lastly, we discuss the work of other visual designers and cartographers that aligns with our goals and conclude by presenting future work and limits. Ultimately, the goal of this paper is to encourage the XAI community to take lessons from the cultural history of city representations as they establish public processes for explaining and evaluating algorithms.

2 ALGORITHMIC EXPLAINABILITY

2.1 Visual explanations for AI

Research in the field of XAI is quickly growing. More recently, researchers have presented the importance of the form of algorithmic explanations and its effect on the understanding of AI systems [56]. Information design experts and critical studies scholars have started employing visualizations as tools to understand and explain algorithms. In this section, we summarize works that present visual techniques to explain algorithms.

Integrating XAI and visual design can make explanation interfaces more usable and readable [7]. The effects of static and interactive visualization techniques of white box (where a model's inner working is shown) and black box (where input and output variables' relationships are shown) explanations on user comprehension have been investigated by Cheng et al. [7]. They found that white-box interactive explanations were most effective in increasing user understanding but were worse than black-box explanations in increasing user confidence in their understanding. This may be a result of the complexity and cognitive overload of white-box explanations. An understanding of human cognition and decision-making capabilities is now being used to develop frameworks for explaining algorithms [59]. Researchers have also attempted to visualize ethical frameworks in order to make them more accessible to users who may not be familiar with ethical terminologies [49]. However, more work is needed to design accessible white-box explanations of algorithms.

Share Lab has used data visualizations to represent several aspects of algorithms including algorithmic labor, invisible infrastructures that surround algorithms, and the social and political relations that inform the workings of tech companies [15]. Kate Crawford's work titled 'Anatomy of AI' which displays several invisible aspects of labor, data, and environmental resources in relation to algorithms is another well-known example of visualizing algorithms [10]. These works draw great attention to the socio-political structures that surround the design of algorithms. However, they do not attempt to explain how the complexity of the world we live in is reduced to conform it to data practices.

Interactive visual analytics are being used to help data scientists better understand their systems through the design of tools such as Prospector [27], Gamut [22], Visual Auditor [36], and more. These tools visualize algorithms for controlled assessment and evaluation. More work is needed to incorporate real-life situated perspectives into algorithmic explanations.

Data Comics have been presented as a means to better report HCI and statistical analysis research studies and are being explored as visualization techniques to communicate research processes and practices in accessible and engaging manners [60]. Economist Julia Schneider and Artist Lena Kadriye Ziyal designed a comic series that explains what Artificial Intelligence is, its core properties, and the risks associated with its widespread deployment [46]. Explaining specific design decisions underlying the functioning of algorithms and their impact on the lives of its users still remains underexplored.

These visual XAI approaches are highly innovative and serve as great starting points for furthering research at the intersection of information design and XAI. Opportunities to improve lie in designing explanations that: (1) are easy to comprehend, (2) account for socio-political factors surrounding algorithms, (3) incorporate local knowledges, and (4) reveal how algorithms internally represent cities.

This paper furthers the work of the scholars above and calls for the use of *mapping* to produce grounded explanations of the inner life of cities. We define grounded explanations as representations of phenomena that are accessible, culturally reflexive, situated, and provide visibility into their internal mechanisms. We discuss how maps can be effective in furthering the development of grounded explanations for geospatial algorithms and addressing the gaps in existing XAI research by drawing on the work of Dr. John Snow. We describe his work briefly in the section below.

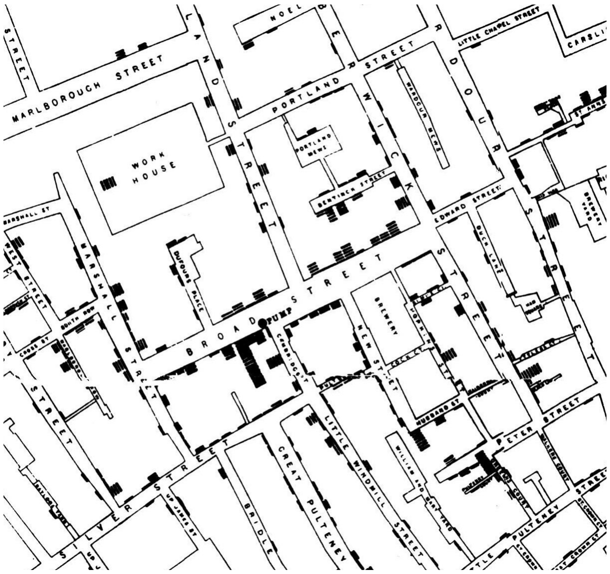

Figure 1: John Snow's map that traces the spread of Cholera deaths during London's 1854 Cholera epidemic

2.2 The Ghost by John Snow: a brief description

Dr. John Snow, a pioneer in social mapping, had been studying the periodic Cholera epidemics in London since the 1830s. He hypothesized that Cholera is water-borne and is caused by the ingestion of contaminated water in contrast to the then-popular belief that Cholera is caused due to poisonous air. In the 1854 Cholera breakout, Snow set out to identify and prove the cause of the disease. To do so, he mapped the fatalities and their proximity to various water pumps in the form of bars and dots. He noticed that the majority of deaths were taking place around the Broad Street water pump. However, there were some anomalies. A major brewery and a workhouse were largely unaffected by the disease. Upon talking to the people who lived and worked there, Snow found out that they both had their own individual wells for water consumption and the brewery workers drank mostly beer, which was unaffected by contamination. He also observed that there were cases of outbreaks in areas distant from the Broad Street pump. Later, Snow discovered that those cases were, in fact, connected to the Broad Street pump; all of them had some personal or professional relationships in the area. Snow presented his map to both residents and the authorities, upon which the city agreed to remove the handle of the Broad Street pump. This prevented people from consuming its water and eventually, the epidemic ended. London has not seen a Cholera breakout since [50]. What can XAI researchers learn from this historical example, as they seek to explain the algorithms that govern "smart cities?"

3 ALGORITHMIC EXPLAINABILITY: LIMITS AND POSSIBILITIES

What can XAI researchers learn from the historical example of the "Ghost Map", as they seek to explain the algorithms that govern "smart cities"? With the rising need to explain algorithms that heavily influence the lives of city inhabitants, XAI researchers have made noteworthy progress in developing methods, tools, and frameworks for algorithmic explainability as described above. However, despite the progress, there remain several limits. In this paper, we argue that current XAI approaches could be substantially improved by incorporating four design strategies introduced below, each of which is found in the Ghost Map and indeed many other conventional maps. In the section that follows, we reflect on these four strategies, which we call—accessibility, cultural reflexivity, situatedness, and visibility into internal representations. In order to do so, we will make use of John Snow's Ghost Map, because it is such a salient example of the explanatory power of mapmaking.

3.1 Accessibility

Much research has focused on developing documentation methods for ML scientists to evaluate the effectiveness of models and data. One such popular framework, 'Model Cards', provides essential information about pre-trained models such as what is the training data, how was the data processed, etc. [35]. More prescriptive approaches to designing responsible algorithms have also been proposed such as 'Method Cards'. Method cards consist of probes that encourage model creators to provide instructions to other researchers who may read the documentation on how to deploy the model responsibly [1]. Alongside the creation of model documentation frameworks, XAI researchers have also designed systems to specifically document data creation processes such as "Data-centric Explanations" [2], "Datasheets" [16], "Dataset Nutrition Labels" [23], and "Data Statements" [4]. Currently, popular XAI research is catered more toward data scientists and engineers to debug these systems rather than end users [5].

Algorithmic explanations, such as the ones listed above, are not designed to reach diverse audiences, particularly those with limited technical knowledge including everyday users who bear the impact of their algorithmic harms. There is even little research that focuses on user-centered XAI in the Global South context [39]. There have been calls to focus on end-users when designing algorithmic explanations such as GDPR's 'right to explainability' [17]. Recent work in user-centered XAI, even though influential, is limited in scope and tends to be disorganized [11]. It may also either still be too technical for certain end users or too simple to be semantically meaningful [64]. Further, users can only audit systems that they interact with. This leaves out systems such as hiring algorithms, recidivism algorithms, or loan estimation algorithms that may be out of reach of end users. Other times, even if users directly interact with these systems, they may be unaware of how the system works affecting their ability to probe and investigate [11]. Users may need guidance from experts or algorithms to better audit algorithmic systems [48]. More research is needed to create organized, accessible, and meaningful approaches that explain and test algorithms with end users.

3.1.1 What makes an explanation accessible? Before John Snow's work on the Ghost Map revealed that Cholera is spread through contaminated drinking water, the London water distribution board was convinced that the epidemic was the result of something in the air: a miasma. They refused to believe the research presented by John Snow until he supplemented his research with a visual, easily understandable, and detailed representation tracing the epidemic's effects. Snow gave spatial form to the General Registrar's data about Cholera deaths. He converted one-dimensional data organized by date of death to two-dimensional data organized by proximity to water sources. This conveyed the cause-effect relationship between Cholera infections and proximity to the Broad Street pump efficiently [53] and clearly. The accessible nature of the map allowed many others to not only understand the map but to reproduce it [24]. It neatly captured the intricacies of the complex urban phenomenon that spread the disease and convinces the general public and authorities of its claim [54, 57]. According to Johnson Stevenson, the map helped make sense of a phenomenon [microorganisms invisible to the naked eye] that previously "defied human understanding" [24]. Ultimately, the map serves as an excellent example of a user-friendly visual explanation.

3.2 Reflexivity

The primary components of XAI addressed by current approaches include the following: local instance explanations (weight of features resulting in a specific prediction), instance explanation comparisons (how do features change to result in diverse predictions), counterfactuals (how do predictions change as you modify features), nearest neighbors (what predictions have similar value), regions of error (when can predictions be inaccurate), and feature importance (what features are most influential for a prediction) [22]. The majority of this work takes a technocentric approach to explain algorithms focusing majorly on what technical features, in what capacity, contribute to a result presented by an algorithm.

The design of algorithms, however, is grounded in contextual factors such as project goals and values, budget, team members, etc., that inform the decisions of computer scientists. A focus on exclusively explaining the technicalities gives little insight into the socio-political factors influencing and being shaped by the design of algorithms. Some scholars have begun explaining the societal influence and impact of algorithms. 'Social Transparency' i.e., making known the social and organizational factors surrounding the design and use of algorithms, has been deemed essential to allow users to make informed decisions based on AI's predictions [13]. Polack investigates another way to proactively account for the limits of algorithms by demonstrating how "design problems, objectives, and needs" presuppose the consequences and impact of algorithms [42]. However, more work is needed to explain not just the algorithmic design decisions made, but also the personal, social, economic, political, and technical motivations and impacts of those decisions.

3.2.1 What makes an explanation reflexive? Visualizations bear the burden of being interpreted as a single objective reality of the world. However, this can be overcome with thoughtful consideration. John Snow, in his mapping and supplementary description, did not remove himself from his map. Rather, he carefully documented the possible errors that may have occurred in the data collection and presentation processes alongside the reasons for the possible errors. First, he clearly showed the contrast between deaths in houses near the Broad Street pump and the brewery and workhouse nearby that remained unaffected. This left room to question the cause-effect relationship that Snow hypothesized. Second, in accompanying texts he laid out how some data he received from the registrar's office was missing house numbers and so could not be visualized [54]. While this map itself is a reflexive artifact in the sense that it captured Snow's acknowledgment of his partial perspective through his mapping and descriptions, it also is an artifact that promotes reflexivity in the reader by challenging their stable perspectives. In its form, it communicates the time period it was created in and for what purposes, the region and scope of analysis, and what London looked like at the time.

3.3 Situated in local knowledges

Several techniques for auditing algorithms have been introduced in recent years. These include code audit (assessing the code of the algorithm), non-invasive user audit (a research survey with the users), direct scrape or scarping audit (querying the system and observing the results), sock puppet audit (impersonating users and evaluating results), carrier puppet (impersonating users such that it affects the real world), collaborative or crowdsourced audit (having end users test the algorithm) [3, 45]. Much of the work in XAI today focuses on providing explanations or conducting audits in a controlled environment, externally or internally, often influenced majorly by researchers' intuition [33].

Even though some of these techniques involve users in their testing processes, most of them are scripted and distant from the situated life of end users. Empirical XAI research is especially limited in the Global South [39]. However, algorithms are used in the wild and thus demand that explanations and audits be contextual and grounded in the everyday reality of users. Given this lack of real-world context and diverse perspectives in existing approaches, there have been rising calls to engage end users in the assessment of algorithmic systems [8, 48, 51, 55, 61]. The few individual and communal user-driven audits that have been conducted, successfully surfaced algorithmic harms in times when controlled auditing failed [11]. Contextual XAI work is needed to highlight specific explanation needs or algorithmic harms that may not be visible when experts in the field audit these systems with a 'view from nowhere' [20].

3.3.1 What makes an explanation situated? The Ghost Map is informed by the everyday lives of the residents of Broad Street and the rest of London. To be able to design the map, Snow needed a fine-grained knowledge of the people of London. He conducted extensive interviews to understand people's movement patterns, sanitary conditions, water consumption practices, etc. He also collaborated with local people such as those who attended the sick. Even though the Ghost Map presented a bird's eye view, it allowed readers to observe patterns built by situated knowledges of local experiences [24]. For example, Snow investigated why a brewery close to Broad Street was unaffected by the disease. He found that it was primarily because the brewery workers mostly drank beer instead of water which saved them from the disease [54]. Alongside considering the spatial distribution of the water sources, Snow also considered the time it would take to walk along the turns of the city for a house member and reach various water sources. This further exemplifies the incorporation of situated elements, including temporality, that inform the design and analysis of the Ghost Map.

3.4 Visibility into internal representations

XAI commonly explains algorithms by (a) the use of ML models that are interpretable by design such as "white-box" models (e.g., logistic regression), or (b) the use of post-hoc explanations (e.g., LIME [44] or SHAP [30]). Most XAI research attempts to explain the mapping of inputs with outputs. However, Rader et al. have argued that alongside the why and how of algorithmic functioning, the what (identifying the presence of algorithms), and unbiased process descriptions (explanations of how the algorithmic system is being continually improved and modified) should also be given importance [43].

Current XAI approaches obscure how algorithms divide and categorize the world in order to make standard data analytical practices applicable to it [40]. To make data work with algorithms and to make predictions, computer scientists process and manipulate the data while constructing simplified representations of the world that guide the working of algorithms. Policy makers have long been doing the same and attempting to simplify cities, disregarding practical social orders and arrangements, for better management and functioning. However, such simplification has often failed in the face of un-codable practical complexities and improvisations [47]. The use of algorithms is highly entangled with the interactions between cities, citizens, and policy makers. This calls for understanding how infrastructures, both physical and social, are constructed through and reinforced by the design of algorithms. Such visibility into what elements and people are included or excluded in the design of civic systems dictates citizens' ability to engage democratically with cities and their governance [38].

3.4.1 How can an explanation make internal representations visible?

The Ghost Map follows and represents a clear hypothesis: Cholera spreads with the consumption of contaminated water. To analyze this hypothesis, Snow clearly displayed only two major components on the map— the water pumps in London and Cholera-related deaths. He had a clear causal relationship in mind that he hoped to investigate and that is made clear to the reader [54]. The clarity of the map also provided space for critical questions such as how the selection of time periods (aggregating Cholera deaths over a week or a day) and boundaries drawn (deaths in a house or a block or another abstract segregation) affected our understanding of the spread of the disease [54]. Tufte critiqued John Snow's dot maps by arguing that visualizing deaths as dots gives little information about the population density of different areas. That is, deaths would likely be more in a densely populated area whether or not a water pump was the source of the disease. Without the map of Snow's analysis, Tufte would not have been able to critique and investigate the basis of the claim made by Snow. Mapping how algorithms represent cities provides similar opportunities for critique. For example, when calculating the safety of a location, does the algorithm aggregate the crime rate without normalizing it by the population density of an area? If yes, what are the impacts of such design decisions?

4 DISCUSSION

With the popularization of the "smart city" as a model for the future of urban life, geo-spatial algorithms have become deeply embedded in our everyday lives. The centralized and authoritative nature of these algorithms gives them the power to create or reinforce unjust (and often invisible) societal structures and inequalities [12]. These algorithms redefine cities using the structures that computers are most adept at processing, for example, a hierarchical tree [32]. However, cities are highly complex artifacts that cannot simply fit into the neat structures demanded by computation. The overly simplistic mathematical representations that algorithms rely on are bound to amplify some features of cities and flatten others. These representational distortions can create the kinds of invisible inequalities and incremental injustices that produce large-scale societal effects, such as gentrification and segregation [29]. In order to critique, improve, or resist such simplifications, it is essential to provide visibility into the design of geo-spatial civic algorithms.

Unfortunately, due to the increasing complexity of algorithms, there is a widening gap between how computational systems work and how they are broadly understood [21]. While current XAI approaches (such as [22, 35, 44]) have made noteworthy strides in promoting algorithmic transparency, they are still limited in a variety of dimensions described in this paper: accessibility, cultural reflexivity, situatedness, and providing visibility into internal representations. These limits have received little attention from algorithmic transparency, explainability, and auditing scholars. In this paper, we have demonstrated the usefulness of *mapping* as a technique to provide grounded explanations of how algorithms represent cities, as well as the effects of those representations on the lives of city inhabitants. The goal of our work is to encourage XAI researchers to take inspiration from the cultural history of visual design. We have identified four strategies that conventional maps use to address the limits of current XAI techniques and uncover the complex inner lives of cities. These strategies, we argue, can prove helpful when explaining geo-spatial algorithmic systems.

Comprehending our urban environment by "making public data public" has been considered of utmost importance by many architects and city planners [62]. While we have focused on the case of the influential Ghost Map by John Snow, many other cartographers, visual and information designers, and urban planners present similar noteworthy works that demonstrate the useful qualities of maps in explaining cities. Here are a few other references that XAI researchers might yet explore:

Accessibility: Popular XAI approaches are inaccessible to everyday users. Maps present an opportunity to explain algorithms in a comprehensible manner. In 'The Image of the City' [31], the proto-"city designer" Kevin Lynch describes how he and his students at MIT identified five constitutive elements that city dwellers use to understand the places they live in: paths, edges, districts, nodes, and landmarks. The implication is that these five elements can form the core of "legible" city design and mapping. With the longest print run of any book by MIT Press (more than eighty years), Lynch's work still serves as a useful tool for thinking about how the complex form of cities can be accessible.

Reflexivity: XAI methods do not explain the social-political contexts that shape algorithmic design. This limitation can be addressed by calls for reflexivity in fields such as 'Critical Cartography' [25]. Such representations can question the status quo in a number of ways. Challenging western positivist cartographic techniques, some mapmakers are creating alternative forms of mapping that foreground indigenous forms of spatial knowledge [41]. Another similarly subversive map designed by Laura Kurgan and her collaborators called "Million Dollar Blocks", draws critical attention to unequal practices of incarceration that specifically target Black neighborhoods in New York City (USA). The map draws attention to a few low-income city blocks where millions of dollars are being invested—not for public health or education—but to remove the inhabitants and put them behind bars in the name of creating a safer city [14].

Situatedness: XAI methods tend to be distant from situated real-world contexts and experiences of city inhabitants. On the other hand, several artists and scholars have experimented with participatory mapping projects, that allow participants to make meaningful connections between authoritative data, such as census or censor readings, and their own local knowledge of the places they live in. One such example is the Map Spot Project [63], an expansion of the St. Louis Map Room [37], that brings together local community members to explore and represent the relationships between civic data and lived experiences. This grassroots mapping effort problematizes the "objective" and "stable" nature of big data and grounds data and its limits in contextual experiences of space.

Figure 2: Atlanta Map Spot Project

Visibility into internal representations: XAI approaches limit their scope to explaining relations between input and output variables while disregarding how algorithms construct cities in their design. Decisions about what one chooses to represent on a map have a significant impact on who is affected by city design positively or negatively [6]. Who is on the map or off the map affects who is included and who is not [58]. While maps construct reality through their representations, they also, through their very form, explain to us their constructed reality, leaving room for critiquing and improving said reality.

In our future work, we will use mapping to explain public safety algorithms to city inhabitants. Specifically, we will explain how training datasets and their aggregations are grounded in the historical, economic, political, and social contexts of cities. Currently, we are creating representations of how algorithms partition cities for computation. Gupta et al. have shown how changes in spatial partitioning of cities for aggregation of data can have major effects on algorithmic results. For example, they demonstrate how the Gini index (a measure of spatial inequality) changes as one modifies the scale of calculation [18]. Given the impact of spatial partitioning on algorithmic outcomes and the need for algorithmic transparency in such cases, we are mapping algorithmic partitioning and superimposing it with other data layers that represent historical and contemporary segregations, such as red-lining maps [34], to ground the partitioning in spatial politics. This could help us evaluate if and how algorithms may reinforce or challenge existing city segregations.

Even though we plan to use mapping to study and question spatial injustices, historically mapping has also been used to propagate discriminatory agendas [9]. Maps may present distant and stable representations of reality which may favor one social group over another. However, with advancements in critical cartography, there are now new ways to question the ideologies and assumptions embedded in "objective" maps. Scholars are also exploring participatory mapping as a technique to problematize the positivist representations of cities and employ mapping for pluralistic exploration and critique [28] We present mapping as an exploratory tool for advancing grounded XAI research while acknowledging its limits.

5 CONCLUSION

We believe that the Ghost Map—along with other similarly "grounded" explanations work listed above—can inspire new forms of public engagement with the algorithmic logics of the "smart city". Following the model we have put forward here, we believe explanations of algorithms can be made more accessible, reflexive, situated, and visible in the internal representations of cities. In future work, we plan to build on these insights, by developing a toolkit that a broad range of designers and researchers can use to visualize the algorithms that shape smart cities. Using such tools, we believe, can enhance widespread efforts to understand, improve, or resist the next generation of algorithms designed to remake cities—for better or worse—around the capacities of AI.

REFERENCES

[1] David Adkins, Bilal Alsallakh, Adeel Cheema, Narine Kokhlikyan, Emily McReynolds, Pushkar Mishra, Chavez Procope, Jeremy Sawruk, Erin Wang, and Polina Zvyagina. 2022. Method cards for prescriptive machine-learning transparency. In *Proceedings of the 1st International Conference on AI Engineering: Software Engineering for AI*. 90–100.
[2] Ariful Islam Anik and Andrea Bunt. 2021. Data-centric explanations: explaining training data of machine learning systems to promote transparency. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–13.
[3] Jack Bandy. 2021. Problematic machine behavior: A systematic literature review of algorithm audits. *Proceedings of the acm on human-computer interaction* 5, CSCW1 (2021), 1–34.
[4] Emily M Bender and Batya Friedman. 2018. Data statements for natural language processing: Toward mitigating system bias and enabling better science. *Transactions of the Association for Computational Linguistics* 6 (2018), 587–604.
[5] Umang Bhatt, Alice Xiang, Shubham Sharma, Adrian Weller, Ankur Taly, Yunhan Jia, Joydeep Ghosh, Ruchir Puri, José MF Moura, and Peter Eckersley. 2020. Explainable machine learning in deployment. In *Proceedings of the 2020 conference on fairness, accountability, and transparency*. 648–657.
[6] Peter Bosselmann. 1998. *Representation of places: reality and realism in city design*. Univ of California Press.
[7] Hao-Fei Cheng, Ruotong Wang, Zheng Zhang, Fiona O'Connell, Terrance Gray, F Maxwell Harper, and Haiyi Zhu. 2019. Explaining decision-making algorithms through UI: Strategies to help non-expert stakeholders. In *Proceedings of the 2019 chi conference on human factors in computing systems*. 1–12.
[8] Michael Correll. 2019. Ethical dimensions of visualization research. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–13.
[9] Jeremy W Crampton. 2011. *Mapping: A critical introduction to cartography and GIS*. Vol. 11. John Wiley & Sons.

[10] Kate Crawford and Vladan Joler. 2023. Anatomy of an AI System. https://anatomyof.ai/

[11] Alicia DeVos, Aditi Dhabalia, Hong Shen, Kenneth Holstein, and Motahhare Eslami. 2022. Toward User-Driven Algorithm Auditing: Investigating users' strategies for uncovering harmful algorithmic behavior. In *CHI Conference on Human Factors in Computing Systems*. 1–19.

[12] Catherine D'ignazio and Lauren F Klein. 2020. *Data feminism*. MIT press.

[13] Upol Ehsan, Q Vera Liao, Michael Muller, Mark O Riedl, and Justin D Weisz. 2021. Expanding explainability: Towards social transparency in ai systems. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–19.

[14] Center for Spatial Research Columbia University. 2023. Million Dollar Blocks. https://c4sr.columbia.edu/projects/million-dollar-blocks

[15] Share Foundation. 2023. SHARE LAB – Research & Data Investigation Lab. https://labs.rs/en/

[16] Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé Iii, and Kate Crawford. 2021. Datasheets for datasets. *Commun. ACM* 64, 12 (2021), 86–92.

[17] Bryce Goodman and Seth Flaxman. 2017. European Union regulations on algorithmic decision-making and a "right to explanation". *AI magazine* 38, 3 (2017), 50–57.

[18] Jayant Gupta, Alexander Long, Corey Kewei Xu, Tian Tang, and Shashi Shekhar. 2021. Spatial Dimensions of Algorithmic Transparency: A Summary. In *17th International Symposium on Spatial and Temporal Databases*. 116–125.

[19] Shubhangi Gupta, Sylvia Janicki, Pooja Casula, and Nassim Parvin. 2022. Rethinking Safe Mobility: The Case of Safetipin in India. *International Conference on Information & Communication Technologies and Development* (2022). In Press.

[20] Donna Haraway. 2020. Situated knowledges: The science question in feminism and the privilege of partial perspective. In *Feminist theory reader*. Routledge, 303–310.

[21] Sabrina Hauser, Johan Redström, and Heather Wiltse. 2021. The widening rift between aesthetics and ethics in the design of computational things. *AI & SOCIETY* (2021), 1–17.

[22] Fred Hohman, Andrew Head, Rich Caruana, Robert DeLine, and Steven M Drucker. 2019. Gamut: A design probe to understand how data scientists understand machine learning models. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–13.

[23] Sarah Holland, Ahmed Hosny, Sarah Newman, Joshua Joseph, and Kasia Chmielinski. 2020. The dataset nutrition label. *Data Protection and Privacy, Volume 12: Data Protection and Democracy* 12 (2020), 1.

[24] Steven Johnson. 2006. *The ghost map: The story of London's most terrifying epidemic–and how it changed science, cities, and the modern world*. Penguin.

[25] Annette M Kim. 2015. Critical cartography 2.0: From "participatory mapping" to authored visualizations of power and people. *Landscape and Urban Planning* 142 (2015), 215–225.

[26] Rob Kitchin. 2014. *The data revolution: Big data, open data, data infrastructures and their consequences*. Sage.

[27] Josua Krause, Adam Perer, and Kenney Ng. 2016. Interacting with predictions: Visual inspection of black-box machine learning models. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. 5686–5697.

[28] Hee Rin Lee, Selma Šabanović, and Sonya S Kwak. 2017. Collaborative map making: A reflexive method for understanding matters of concern in design research. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 5678–5689.

[29] Yanni Alexander Loukissas. 2022. Who wants to live in a filter bubble? from'zillow surfing'to data-driven segregation. *Interactions* 29, 3 (2022), 36–41.

[30] Scott M Lundberg and Su-In Lee. 2017. A unified approach to interpreting model predictions. *Advances in neural information processing systems* 30 (2017).

[31] Kevin Lynch. 1964. *The image of the city*. MIT press.

[32] Shannon Mattern. 2020. A city is not a computer. In *The Routledge Companion to Smart Cities*. Routledge, 17–28.

[33] Tim Miller. 2019. Explanation in artificial intelligence: Insights from the social sciences. *Artificial intelligence* 267 (2019), 1–38.

[34] Bruce Mitchell and Juan Franco. 2018. HOLC "redlining" maps: The persistent structure of segregation and economic inequality. (2018).

[35] Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru. 2019. Model cards for model reporting. In *Proceedings of the conference on fairness, accountability, and transparency*. 220–229.

[36] David Munechika, Zijie J Wang, Jack Reidy, Josh Rubin, Krishna Gade, Krishnaram Kenthapadi, and Duen Horng Chau. 2022. Visual Auditor: Interactive Visualization for Detection and Summarization of Model Biases. In *2022 IEEE Visualization and Visual Analytics (VIS)*. IEEE, 45–49.

[37] The Office of Creative Research in partnership with COCA. 2018. St.Louis Map Room. http://stlmaproom.org/

[38] Dietmar Offenhuber. 2015. Infrastructure legibility—a comparative analysis of open311-based citizen feedback systems. *Cambridge Journal of Regions, Economy and Society* 8, 1 (2015), 93–112.

[39] Chinasa T Okolo, Nicola Dell, and Aditya Vashistha. 2022. Making AI Explainable in the Global South: A Systematic Review. (2022).

[40] Samir Passi and Steven Jackson. 2017. Data vision: Learning to see through algorithmic abstraction. In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*. 2436–2447.

[41] Margaret Wickens Pearce and Renee Pualani Louis. 2008. Mapping indigenous depth of place. (2008).

[42] Peter Polack. 2020. Beyond algorithmic reformism: Forward engineering the designs of algorithmic systems. *Big Data & Society* 7, 1 (2020), 2053951720913064.

[43] Emilee Rader, Kelley Cotter, and Janghee Cho. 2018. Explanations as mechanisms for supporting algorithmic transparency. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–13.

[44] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. 2016. " Why should i trust you?" Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*. 1135–1144.

[45] Christian Sandvig, Kevin Hamilton, Karrie Karahalios, and Cedric Langbort. 2014. Auditing algorithms: Research methods for detecting discrimination on internet platforms. *Data and discrimination: converting critical concerns into productive inquiry* 22 (2014), 4349–4357.

[46] Julia Schneider and Lena Ziyal. 2023. We Need to Talk, AI – A Comic Essay on Artificial Intelligence. https://weneedtotalk.ai/

[47] James C Scott. 2008. Seeing like a state. In *Seeing Like a State*. yale university Press.

[48] Hong Shen, Alicia DeVos, Motahhare Eslami, and Kenneth Holstein. 2021. Everyday algorithm auditing: Understanding the power of everyday users in surfacing harmful algorithmic behaviors. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–29.

[49] Joanna Sleigh, Manuel Schneider, Julia Amann, Effy Vayena, et al. 2020. Visualizing an Ethics Framework: A Method to Create Interactive Knowledge Visualizations From Health Policy Documents. *Journal of medical Internet research* 22, 1 (2020), e16249.

[50] John Snow. 1856. On the mode of communication of cholera. *Edinburgh medical journal* 1, 7 (1856), 668.

[51] Megha Srivastava, Hoda Heidari, and Andreas Krause. 2019. Mathematical notions vs. human perception of fairness: A descriptive approach to fairness for machine learning. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*. 2459–2468.

[52] Anthony M Townsend. 2013. *Smart cities: Big data, civic hackers, and the quest for a new utopia*. WW Norton & Company.

[53] Edward R Tufte. 1985. The visual display of quantitative information. *The Journal for Healthcare Quality (JHQ)* 7, 3 (1985), 15.

[54] Edward R Tufte, Susan R McKay, Wolfgang Christian, and James R Matey. 1998. Visual explanations: Images and quantities, evidence and narrative.

[55] Niels Van Berkel, Jorge Goncalves, Danula Hettiachchi, Senuri Wijenayake, Ryan M Kelly, and Vassilis Kostakos. 2019. Crowdsourcing perceptions of fair predictors for machine learning: A recidivism case study. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–21.

[56] Niels Van Berkel, Jorge Goncalves, Daniel Russo, Simo Hosio, and Mikael B Skov. 2021. Effect of information presentation on fairness perceptions of machine learning predictors. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–13.

[57] Laura Vaughan. 2018. *Mapping society: The spatial dimensions of social cartography*. UCL Press.

[58] Janet Vertesi. 2008. Mind the gap: The London underground map and users' representations of urban space. *Social Studies of Science* 38, 1 (2008), 7–33.

[59] Danding Wang, Qian Yang, Ashraf Abdul, and Brian Y Lim. 2019. Designing theory-driven user-centric explainable AI. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–15.

[60] Zezhong Wang, Jacob Ritchie, Jingtao Zhou, Fanny Chevalier, and Benjamin Bach. 2020. Data comics for reporting controlled user studies in human-computer interaction. *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (2020), 967–977.

[61] Allison Woodruff, Sarah E Fox, Steven Rousso-Schindler, and Jeffrey Warshaw. 2018. A qualitative exploration of perceptions of algorithmic fairness. In *Proceedings of the 2018 chi conference on human factors in computing systems*. 1–14.

[62] Richard Saul Wurman. 1971. Making the City Observable. (1971).

[63] Yanni Youkissas. 2018. Atlanta Map Room Project. https://loukissas.lmc.gatech.edu/uncategorized/atlanta-map-room/

[64] Wencan Zhang and Brian Y Lim. 2022. Towards Relatable Explainable AI with the Perceptual Process. In *CHI Conference on Human Factors in Computing Systems*. 1–24.

[65] Sharon Zukin, Scarlett Lindeman, and Laurie Hurson. 2017. The omnivore's neighborhood? Online restaurant reviews, race, and gentrification. *Journal of Consumer Culture* 17, 3 (2017), 459–479.

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009