Mapping the Smart City: Participatory approaches to XAI

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

ABSTRACT

How can we explain the broad and uneven spatial effects of Machine Learning (ML) algorithms that mediate the everyday lives of smart city residents? The discriminatory impacts of civic algorithms remain opaque to city inhabitants and experts alike. Current Explainable AI (XAI) approaches, while influential, are limited in their ability to explain the inequitable algorithmic spatial effects in an accessible, critical, and grounded manner. My thesis explores the potential of *participatory mapping* as a critical and collaborative technique to address these limits. My work draws on (1) scholarship on critical data and algorithmic studies, (2) qualitative research with domain experts from history and criminology, and (3) participatory mapping sessions with city residents and ML practitioners. Ultimately, my research will inform the design of a toolkit to help people in classrooms and community centers collaboratively reflect on how city residents may unevenly experience the impact of artificially intelligent systems guiding contemporary urban life.

CCS CONCEPTS

• **Human-centered computing → Information visualization**.

KEYWORDS

Explainable AI, Participatory Methods, Transparency, Smart city, Mapping, Visualization

ACM Reference Format:
Shubhangi Gupta. 2023. Mapping the Smart City: Participatory approaches to XAI. In *Designing Interactive Systems Conference (DIS Companion '23), July 10–14, 2023, Pittsburgh, PA, USA.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3563703.3593063

1 INTRODUCTION

Civic algorithms, i.e., Machine Learning (ML) algorithms used for civic purposes, have grown to profoundly guide the everyday lives of smart city residents. However, without careful consideration of their effects, they risk reproducing or even amplifying historical systems of discrimination. Several smart city systems have been criticized by activists and scholars alike for their inequitable effects on space. Planning algorithms, such as the Market Value Analysis, present governing bodies with a "data-driven objective" means of distributing public resources by classifying and segregating communities through the construction of color-coded boundaries. Such organization of space standardizes unjust classification systems and restructures the public sphere in ways that favor some neighborhoods over others. These systems are opaque, inaccessible, and are rarely if ever, developed through a public and participatory process [53]. Online websites such as Zillow and Yelp affect civic investments and neighborhood reputations in ways that further marginalize racialized and poor neighborhoods. Yet, their effects remain invisible to the general public [39, 72]. Black-boxed surveillance tools, situated within racial and neoliberal histories, disregard the perspectives of those who are most affected by these systems and work to promote biopolitics while encroaching on peoples' liberties [61]. Even as these systems continue to cause harm, their impacts remain invisible to both the users and makers of these tools.

Critical geographers have demonstrated how social discrimination and marginalization are inherently entangled with space [56]. In my thesis, I focus on explaining the emergent spatial effects of public safety algorithms. I define algorithmic *spatial effects* as the localized impacts of smart city systems, as reported by the people who experience them. Oftentimes, *spatial effects* are echoes of existing spatial *information infrastructures* such as zip codes, GPS coordinates, or neighborhood boundaries that ML algorithms rely on to make cities legible, predictable, and ultimately programmable [42, 47]. This way of reading cities has an unfortunate tendency to incorporate and even amplify historical biases. Akpinar et al. argue, for example, that a public safety algorithm called PredPol reproduces existing spatial disparities and bias in geocoded crime data, resulting in over-policing of predominately Black neighborhoods [1].

Even though there has been growing research in the realm of explaining algorithms, most current research provides little support in explaining the spatial effects of civic algorithms in a manner that is grounded in the everyday experiences of people most affected by these systems [28]. In my Ph.D. work, I build upon current XAI research, while addressing their limits, to design a participatory mapping toolkit that can help people in classrooms and community centers reflect on how city residents may be impacted unevenly by the spatial effects of artificially intelligent civic algorithms.

To understand data processes and their impacts, we need to understand the historical, social, and political systems they are embedded in [11]. Therefore, my participatory mapping toolkit will help ground the effects of safety algorithms in the lives of city residents and provide a means to investigate their underlying information infrastructures in a pluralistic manner. Ultimately, my goals when designing for algorithmic spatial explainability are (1) to make visible the spatial information infrastructures guiding the design of safety algorithms, (2) to provide a collaborative and grounded approach to reflect on the spatial effects of civic algorithms, and (3) to problematize the unfounded trust in the 'objectiveness' and

DIS Companion '23, July 10–14, 2023, Pittsburgh, PA, USA

ACM ISBN 978-1-4503-9898-5/23/07.
https://doi.org/10.1145/3563703.3593063

'neutrality' of AI. Ultimately, my work will support the public examination and scrutiny of smart cities and provide a means to hold tech makers accountable for the impact of their work.

2 BACKGROUND AND RELATED WORK

2.1 Public Safety Algorithms and Harms

A variety of ML algorithms have been developed with the hope to mitigate crime and advance citizen safety in smart cities. Popular examples include— COMPAS, which predicts recidivism risk for an individual; Predpol, which predicts geographic areas where crime is most likely to happen; Arnold Public Safety Assessment, which provides judges with sentencing recommendations. These algorithms, even as they aim to promote public safety in cities, tend to reinforce discrimination along the axes of race and class. Jefferson demonstrates how Predpol legitimizes the bias embedded in official crime datasets and has resulted in the over-policing of already heavily surveilled neighborhoods [30]. Risk assessment tools build upon and reinforce the racist policies and infrastructures underlying carceral systems in the US. Additionally, they define 'risk' at the level of an individual, disregarding how 'risk' is a reflection of societal prejudice against various social groups [23]. In India, centralized systems and norms along with the subjectivities of individual police officers lead to historical, representational, and measurement bias in recorded crime data for the CMAPS predictive policing tool. Further, the opaque design of CMAPS allows for discrimination against immigrant colonies and minority settlements by promoting the belief that crime rises in specific neighborhoods by virtue of the above-mentioned communities living there [41].

Transparency is a much-needed feature for the effective assessment and development of public safety algorithms [51]. Given the limited potential of techniques designed to "de-bias" public safety algorithms, there is an urgent need to make the data assemblages [33] surrounding public safety algorithms transparent and accessible to city residents and governmental bodies [41].

2.2 Existing XAI Methods and Tools

Several methods and tools, such as model [44] and data documentation [4, 7, 22, 29] frameworks, have been designed to explain AI to experts [17]. However, more recently, the XAI community has presented the need to make AI processes visible to the general public to build trust in the artificially intelligent systems that guide their lives [34]. As such, there is a growing focus on developing methods and frameworks that provide user-centered algorithmic explanations [43, 59, 65]. Human cognitive abilities [66], users' explanatory needs [37, 62], users' situated real-world experiences [16], users' ability to collaborate and form counter publics [57] have been some of the guiding factors in advancing user-centered XAI research. Another strategy to make explanations accessible, which has also been previously explored at DIS [55], is the integration of visual design and XAI efforts [13]. Artists and scholars are using visual techniques to explain the core properties and risks of AI [67] and the sociotechnical infrastructures surrounding AI [15, 21].

There is also growing work in designing open-source toolkits to identify and assess algorithmic harms and biases [6, 8, 68]. AI Fairness 360 (AIF360) [6] and Fairlearn [8] are tools that aim to help practitioners understand 'bias' metrics and allow them to detect algorithmic biases. These tools also help mitigate said biases by providing a variety of mitigation algorithms. Another tool called "What-If" uses visualizations to help users and practitioners investigate how a model will perform in hypothetical scenarios created by changes in data points [68]. The AIX360 toolkit considers the varied explanation needs of different stakeholders and compiles a variety of explanation algorithms with supporting case studies to help users understand algorithms [5]. These tools attempt to advance efforts in algorithmic transparency, auditing, and risk mitigation to reach practitioners outside of an organization's internal technology teams as well as users of the systems.

2.3 Participatory Approaches to XAI

Beyond merely catering the explanations to users, there have been few but rising calls for centering local communities in the design and development of ML algorithms [12, 31, 35, 36, 58, 70]. Shen et al. have presented a model cards toolkit that can be used by community members to deliberate on which model, amongst a variety of models, aligns with their values and interests [58]. Lee et al. have developed a framework that supports community members in building policy to govern algorithms in a participatory manner [36].

Participatory methods have also been previously explored for analyzing the impacts of artificially intelligent systems at DIS [9] and related venues [2, 35, 49]. Blair et al. propose to explore the potential of participatory art installations to perform a public assessment of predictive algorithms [9]. The AI Now Institute proposed the Algorithmic Impact Assessment (AIA) framework and noted the need for affected communities and governmental bodies to be aware of how black-boxed automated decision-making systems work [49]. Alvarado et al. propose Algorithmic Experience (AX), an analytical tool that can be employed in a participatory manner to understand users' experiences of AI-driven tools such as the Facebook News Feed [2]. The Algorithmic Equity Toolkit (the AEKit) offers a collection of methods to increase the participation of the public in algorithmic advocacy [35].

I build on the work described above to develop a participatory mapping toolkit that can help local community members and ML researchers reflect on the spatial effects of public safety algorithms. At the same time, my work differs from the approaches described in Sections 2.2 and 2.3 in a multitude of ways– (1) Subject: Unlike current approaches that focus on explaining input-output relationships, ML processes, or general AI impacts, my work focuses specifically on explaining the spatial effects of civic algorithms; (2) Methodology: Instead of relying on AI models or tech experts to develop AI explanations in a manner that is removed from the local experiences of community members, I use the technique of *participatory mapping*, as described below, for providing grounded explanations; (3) Goals: In contrast to existing approaches to XAI that aim to build trust in users [19, 38], the goal of my thesis is to problematize said trust and empower city inhabitants and ML practitioners to effectively evaluate ways in which AI systems organize cities.

2.4 Participatory Mapping as a Critical Practice

Historically, mapmaking as a practice has been used to support the ideologies of its makers and has served as a tool of persuasion and power to advance colonization and imperialism [71]. However, with the growth of participatory mapping, local communities have repurposed this historical practice to fulfill their own purposes and demand social and political justice [60]. Techniques such as counter-mapping [46], collaborative cartography [10], and participatory GIS [20] aim to engage local communities in the process of map-making to represent and visualize existing systemic injustices and propose better futures [10]. There exist several influential examples. In 1971, the Detroit Geographic Expedition and Institute (DGEI) released a map called "Where Commuters Run Over Black Children on the Pointes-Downtown Track". The map was the result of a collaboration between young black adults from local neighborhoods and academics. Through the collaboration, the youth learned cutting-edge mapping techniques to transform their local knowledges into tools to demand justice [18]. Another provocative example is the Anti Eviction Mapping effort, a participatory oral history project, which is a result of collaboration with local partners and people being evicted with the goal of resisting urban gentrification [40].

Participatory methods aim to develop social and technical systems directly in collaboration with end users [45]. *Critical making* allows participants to focus on the shared processes of construction as a site to develop a conceptual understanding of critical sociotechnical issues [48]. I plan to design a *participatory mapping toolkit as a form of collaborative critical making* to help the makers and users of civic algorithms reflect on both the positive and negative spatial effects of public safety algorithms.

Despite the opportunities presented by participatory mapping to investigate social justice issues, there remain limits. Participatory maps, even as they strive to be pluralistic, may still silence the voices of the least privileged while enhancing the perspective of a chosen elite. The activity may also become a benign way of involving community members while reserving the power and decision-making capabilities for the more technically proficient people [24]. For example, in the participatory mapping sessions I plan to host, people who are less acquainted with AI may feel uncomfortable participating in a collaborative setting. I acknowledge the limits of my methodology and will actively work to create a safe and respectful environment for all participants.

3 RESEARCH OBJECTIVES AND METHODS

Through my research, I aim to develop methods that can support the XAI community in their efforts to explain the spatial effects of civic algorithms. My thesis will result in the design of a participatory mapping toolkit that helps represent the spatial information infrastructures underlying the design of public safety algorithms and their impacts. While my focus is on risk assessment algorithms used to calculate safety scores of various geographic locations in a city, the mapping toolkit will also be able to examine other civic algorithms. In line with these goals, I have divided my thesis into three major parts: (1) Problem Statement and Methodology (2) Research and Design (3) Testing and Results

3.1 Part 1: Problem Statement and Methodology

This part of my thesis focuses on understanding ways in which public safety technologies may contribute to the harmful and discriminatory segregation of cities. I demonstrate the need for the XAI community to step in and make the spatial logics of civic algorithms and their possible social impacts accessible and visible to ML researchers as well as city inhabitants. To address this need, I present participatory mapping as a technique that can support pluralistic exploration and explanation of the spatial boundaries embedded in and reinforced by algorithms. This part of my thesis motivates my problem statement and methodology.

This part began with a critical analysis of a renowned safe walking app primarily deployed in India, called 'Safetipin' [52]. Safetipin recommends 'safe' paths to users from an origin to a destination by calculating 'safety scores' for various paths. These safety scores are calculated by aggregating crowdsourced 'safety data' such as the amount of lighting, or presence of security officers, in various locations in a city. In my past work [27], I demonstrate Safetipin's ability to (1) restrict women's movement to computationally calculated 'safe' neighborhoods and (2) reinforce caste and religion based segregations in India. By disregarding the prejudice about vulnerable neighborhoods that governs the 'feeling of safety' of its users who contribute to the crowdsourced data, it fails to situate itself in the broader historical politics of safety in the city [32] that continue to marginalize people of lower socioeconomic status and minority religions. Nonetheless, the app and its underlying information infrastructures that promote segregation, have been enthusiastically accepted and celebrated [26, 69]. This presents a dire need to explain the impact of spatially distributed data inputs and aggregations on the city and its people. This work motivates the need to understand and explain how emerging public safety technologies organize cities and their impact on spatial segregation and discrimination.

Having established the need to explain the spatial effects of civic algorithms, my past work presents mapping as a useful technique to ground public safety algorithms in historical, social, and political contexts [28]. I argue that mapping can provide explanations that are accessible, culturally reflexive, situated, and provide visibility into how cities are structured by AI processes. However, mapping risks portraying space as stable and objective, representing the realities of the dominant and the powerful [14]. To address this shortcoming, I am now exploring the potential of *participatory mapping* to support the pluralistic understanding and exploration of spatial politics of smart city algorithms. Currently, I am performing an analysis of historical and contemporary participatory mapping cases, to understand how various features of participatory mapping can help explain ML algorithms and their potential limits. Feedback from the Doctoral Consortium (DC) can help me reflect on the opportunities and challenges presented by participatory mapping as an XAI approach.

3.2 Part 2: Research and Design

The goal of this part of my thesis is to (1) identify what aspects of the data construction and computation processes need to be considered in order to explain the impact of algorithms on space, and (2) design a participatory mapping toolkit that incorporates this information onto a map for accessible and collaborative reflection on the spatial politics of civic algorithms. This part focuses on research and design.

To progress towards Goal 1 listed above, I am performing a close reading of critical scholarship that discusses how public safety algorithms result in harm in order to deduce the role of space in the same. As a starting point, I have identified five primary components:

Data Construction and Consumption sites: The spatial disparities embedded in data creation and consumption affect location-based decision making [1].

Data Properties: Bias in who is represented in the training data and variables, including proxy variables, associated with them affect the calculation of equitable predictions [3].

Spatial Partitioning of Data: Ways in which cities are partitioned to aggregate spatial data affect algorithmic calculations [25].

Relationship to Other Data: Spatial components of training data may be correlated with other civic data that may reinforce the impacts of one another [50].

Temporal boundaries: Temporal dimensions of automation affect systems of oppression in space [63].

I plan on conducting a semi-structured qualitative study with domain experts from criminology, history, sociology, and tech ethics, framed around the components identified above, in order to refine and revise these components and identify related relevant information that can help ground civic AI systems. Drawing on the literature review and expert evaluation, I will develop a framework that will be incorporated into the design of a participatory mapping toolkit as stated in Goal 2. I will appreciate feedback from the DC on the mapping elements described above as I develop my framework and toolkit. I also would benefit from advice on my methodology—including recruitment and data analysis strategies.

3.3 Part 3: Testing and Revisions

The goal of this part of the thesis is to use the mapping toolkit mentioned above and explore how it can be used (1) by machine learning students and scientists to learn about the spatial politics of their work, and (2) by city residents to understand smart city algorithms and problematize the unfounded trust citizens may place in algorithmic technologies. This part focuses on testing, results, and revisions.

I will perform two participatory sessions with the groups mentioned above to test the mapping toolkit and identify the opportunities offered by and limits of participatory mapping as an XAI approach. I will partner with local organizations, including my own university, to recruit participants for the participatory sessions. I am in the process of identifying community partners by collaborating with initiatives such as 'Serve-Learn-Sustain' at Georgia Tech [64]. The work can be extended to involve stakeholders not mentioned above such as policymakers and urban planners, etc. I would benefit from advice on how to develop relationships with local community partners in meaningful ways. Additionally, given the limits of participatory mapping as mentioned in Section 2.4, I would appreciate feedback on best practices for creating a safe collaborative environment when conducting participatory research.

4 DISSERTATION STATUS AND LONG-TERM GOALS

I have completed Part 1 as described in Section 3 and am currently working towards Part 2, with a goal of completion by August 2023, followed by progress towards Part 3 between August 2023 and May 2024. I successfully passed my qualifier exams in Spring 2022 and plan to propose my dissertation in May 2023. My next steps include developing the participatory mapping toolkit in a manner that supports the meaningful exploration of the spatial effects of civic algorithms and finalizing the logistics of collaborating with local partners. I would greatly appreciate feedback from the Doctoral Consortium about any shortcomings or gaps in my Ph.D. work that need to be urgently addressed.

Long term, I plan to contribute to the fields of critical data and algorithmic studies by designing tools and methods that center local communities in the design of smart cities. Eventually, I plan to evolve the focus of my work to the context of my home country—India, and join the efforts to support the responsible design of technology in the Global South [54].

5 PROGRAM INFORMATION

I am a third-year Ph.D. student majoring in Digital Media at the Georgia Institute of Technology. I am advised by Dr. Yanni Loukissas. My expected graduation date is May 2024. I have never attended a Doctoral Consortium and my work could greatly benefit from constructive feedback on my research questions, arguments, and methods.

REFERENCES

[1] Nil-Jana Akpinar, Maria De-Arteaga, and Alexandra Chouldechova. 2021. The effect of differential victim crime reporting on predictive policing systems. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency.* 838–849.

[2] Oscar Alvarado and Annika Waern. 2018. Towards algorithmic experience: Initial efforts for social media contexts. In *Proceedings of the 2018 chi conference on human factors in computing systems.* 1–12.

[3] Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner. 2016. Machine bias. In *Ethics of data and analytics.* Auerbach Publications, 254–264.

[4] 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.

[5] Vijay Arya, Rachel KE Bellamy, Pin-Yu Chen, Amit Dhurandhar, Michael Hind, Samuel C Hoffman, Stephanie Houde, Q Vera Liao, Ronny Luss, Aleksandra Mojsilović, et al. 2019. One explanation does not fit all: A toolkit and taxonomy of ai explainability techniques. *arXiv preprint arXiv:1909.03012* (2019).

[6] Rachel KE Bellamy, Kuntal Dey, Michael Hind, Samuel C Hoffman, Stephanie Houde, Kalapriya Kannan, Pranay Lohia, Jacquelyn Martino, Sameep Mehta, Aleksandra Mojsilovic, et al. 2018. AI Fairness 360: An extensible toolkit for detecting, understanding, and mitigating unwanted algorithmic bias. *arXiv preprint arXiv:1810.01943* (2018).

[7] 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.

[8] Sarah Bird, Miro Dudík, Richard Edgar, Brandon Horn, Roman Lutz, Vanessa Milan, Mehrnoosh Sameki, Hanna Wallach, and Kathleen Walker. 2020. Fairlearn: A toolkit for assessing and improving fairness in AI. *Microsoft, Tech. Rep. MSR-TR-2020-32* (2020).

[9] Kathryn Blair, Pil Hansen, and Lora Oehlberg. 2021. Participatory Art for Public Exploration of Algorithmic Decision-Making. In *Companion Publication of the 2021 ACM Designing Interactive Systems Conference.* 23–26.

[10] A. Bosse. 2021. Collaborative Cartography. *The Geographic Information Science & Technology Body of Knowledge (3rd Quarter 2021 Edition)* (2021).

[11] danah boyd. 2020. Questioning the legitimacy of data. *Information Services & Use* 40, 3 (2020), 259–272.

[12] Anna Brown, Alexandra Chouldechova, Emily Putnam-Hornstein, Andrew Tobin, and Rhema Vaithianathan. 2019. Toward algorithmic accountability in public services: A qualitative study of affected community perspectives on algorithmic decision-making in child welfare services. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems.* 1–12.

[13] 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.
- [14] Jeremy W Crampton. 2011. *Mapping: A critical introduction to cartography and GIS*. Vol. 11. John Wiley & Sons.
- [15] Kate Crawford and Vladan Joler. 2023. Anatomy of an AI System. https://anatomyof.ai/
- [16] 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 *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–19.
- [17] Shipi Dhanorkar, Christine T Wolf, Kun Qian, Anbang Xu, Lucian Popa, and Yunyao Li. 2021. Who needs to know what, when?: Broadening the Explainable AI (XAI) Design Space by Looking at Explanations Across the AI Lifecycle. In *Designing Interactive Systems Conference 2021*. 1591–1602.
- [18] Catherine D'Ignazio and Lauren Klein. 2021. Who Collects the Data? A Tale of Three Maps. (2021).
- [19] Upol Ehsan, Pradyumna Tambwekar, Larry Chan, Brent Harrison, and Mark O Riedl. 2019. Automated rationale generation: a technique for explainable AI and its effects on human perceptions. In *Proceedings of the 24th International Conference on Intelligent User Interfaces*. 263–274.
- [20] Sarah Elwood. 2006. Critical issues in participatory GIS: Deconstructions, reconstructions, and new research directions. *Transactions in GIS* 10, 5 (2006), 693–708.
- [21] Share Foundation. 2023. SHARE LAB – Research & Data Investigation Lab. https://labs.rs/en/
- [22] 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.
- [23] Ben Green. 2020. The false promise of risk assessments: epistemic reform and the limits of fairness. In *Proceedings of the 2020 conference on fairness, accountability, and transparency*. 594–606.
- [24] Jo Guldi. 2017. A history of the participatory map. *Public Culture* (2017).
- [25] 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.
- [26] Poorvi Gupta. 2019. India's Safetipin App Wins Asia Foundation's Lotus Leadership Award. https://www.shethepeople.tv/news/safetipin-app-asia-foundation-award/
- [27] 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.
- [28] Shubhangi Gupta and Yanni 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*. In Press.
- [29] Sarah Holland, Ahmed Hosny, Sarah Newman, Joshua Joseph, and Kasia Chmielinski. 2020. The dataset nutrition label. *Data Protection and Privacy* 12, 12 (2020), 1.
- [30] Brian Jordan Jefferson. 2018. Predictable policing: Predictive crime mapping and geographies of policing and race. *Annals of the American Association of Geographers* 108, 1 (2018), 1–16.
- [31] Michael Katell, Meg Young, Dharma Dailey, Bernease Herman, Vivian Guetler, Aaron Tam, Corinne Bintz, Daniella Raz, and PM Krafft. 2020. Toward situated interventions for algorithmic equity: lessons from the field. In *Proceedings of the 2020 conference on fairness, accountability, and transparency*. 45–55.
- [32] Leslie Kern. 2021. *Feminist city: Claiming space in a man-made world*. Verso Books.
- [33] Rob Kitchin and Tracey Lauriault. 2014. Towards critical data studies: Charting and unpacking data assemblages and their work. (2014).
- [34] Bran Knowles and John T Richards. 2021. The sanction of authority: Promoting public trust in ai. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. 262–271.
- [35] PM Krafft, Meg Young, Michael Katell, Jennifer E Lee, Shankar Narayan, Micah Epstein, Dharma Dailey, Bernease Herman, Aaron Tam, Vivian Guetler, et al. 2021. An action-oriented AI policy toolkit for technology audits by community advocates and activists. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*. 772–781.
- [36] Min Kyung Lee, Daniel Kusbit, Anson Kahng, Ji Tae Kim, Xinran Yuan, Allissa Chan, Daniel See, Ritesh Noothigattu, Siheon Lee, Alexandros Psomas, et al. 2019. WeBuildAI: Participatory framework for algorithmic governance. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–35.
- [37] Q Vera Liao, Daniel Gruen, and Sarah Miller. 2020. Questioning the AI: informing design practices for explainable AI user experiences. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–15.
- [38] Zachary C Lipton. 2018. The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery. *Queue* 16, 3 (2018), 31–57.
- [39] 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.
- [40] Manissa M Maharawal and Erin McElroy. 2018. The anti-eviction mapping project: Counter mapping and oral history toward bay area housing justice. *Annals of the American Association of Geographers* 108, 2 (2018), 380–389.
- [41] Vidushi Marda and Shivangi Narayan. 2020. Data in New Delhi's predictive policing system. In *Proceedings of the 2020 conference on fairness, accountability, and transparency*. 317–324.
- [42] Shannon Mattern. 2020. A city is not a computer. In *The Routledge Companion to Smart Cities*. Routledge, 17–28.
- [43] Tim Miller. 2019. Explanation in artificial intelligence: Insights from the social sciences. *Artificial intelligence* 267 (2019), 1–38.
- [44] 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.
- [45] Michael J Muller and Sarah Kuhn. 1993. Participatory design. *Commun. ACM* 36, 6 (1993), 24–28.
- [46] Kollektiv Orangotango+. 2018. *This Is Not an Atlas: A Global Collection of Counter-Cartographies*. transcript Verlag.
- [47] 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.
- [48] Matt Ratto. 2011. Critical making: Conceptual and material studies in technology and social life. *The information society* 27, 4 (2011), 252–260.
- [49] Dillon Reisman, Jason Schultz, Kate Crawford, and Meredith Whittaker. 2018. Algorithmic Impact Assessments: A Practical Framework for Public Agency. *AI Now* (2018).
- [50] Rashida Richardson, Jason M Schultz, and Kate Crawford. 2019. Dirty data, bad predictions: How civil rights violations impact police data, predictive policing systems, and justice. *NYUL Rev. Online* 94 (2019), 15.
- [51] Cynthia Rudin, Caroline Wang, and Beau Coker. 2020. The age of secrecy and unfairness in recidivism prediction. *Harvard Data Science Review* 2, 1 (2020), 1.
- [52] Safetipin. 2023. Safetipin, Creating Safe Public Spaces for Women. https://safetipin.com/
- [53] Sara Safransky. 2020. Geographies of algorithmic violence: Redlining the smart city. *International Journal of Urban and Regional Research* 44, 2 (2020), 200–218.
- [54] Nithya Sambasivan. 2021. Seeing like a dataset from the global south. *Interactions* 28, 4 (2021), 76–78.
- [55] Bianca G Schor, Jatinder Singh, and Alan Blackwell. 2022. Making Algorithmic Systems More Meaningfully Transparent for Users Through Visualisation. In *Designing Interactive Systems Conference*. 18–20.
- [56] Tim Schwanenand and MP Kwan. 2012. Critical space–time geographies thinking the spatiotemporal. guest editorial. *Environment and Planning A* 44 (2012), 2043–2048.
- [57] 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.
- [58] Hong Shen, Leijie Wang, Wesley H Deng, Ciell Brusse, Ronald Velgersdijk, and Haiyi Zhu. 2022. The model card authoring toolkit: Toward community-centered, deliberation-driven AI design. In *2022 ACM Conference on Fairness, Accountability, and Transparency*. 440–451.
- [59] Ben Shneiderman. 2020. Human-centered artificial intelligence: Reliable, safe & trustworthy. *International Journal of Human–Computer Interaction* 36, 6 (2020), 495–504.
- [60] Bjørn Sletto, Alfredo Wagner, Joe Bryan, and Charles Hale. 2020. *Radical cartographies: participatory mapmaking from Latin America*. University of Texas Press.
- [61] Gavin JD Smith. 2020. The politics of algorithmic governance in the black box city. *Big Data & Society* 7, 2 (2020), 2053951720933989.
- [62] Aaron Springer and Steve Whittaker. 2020. Progressive disclosure: When, why, and how do users want algorithmic transparency information? *ACM Transactions on Interactive Intelligent Systems (TiiS)* 10, 4 (2020), 1–32.
- [63] Daniel Susser. 2022. Decision Time: Normative Dimensions of Algorithmic Speed. In *2022 ACM Conference on Fairness, Accountability, and Transparency*. 1410–1420.
- [64] Georgia Tech. 2023. Serve-Learn-Sustain. https://serve-learn-sustain.gatech.edu/
- [65] Jennifer Wortman Vaughan and Hanna Wallach. 2020. A human-centered agenda for intelligible machine learning. *Machines We Trust: Getting Along with Artificial Intelligence* (2020).
- [66] 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.
- [67] 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.
- [68] James Wexler, Mahima Pushkarna, Tolga Bolukbasi, Martin Wattenberg, Fernanda Viégas, and Jimbo Wilson. 2019. The what-if tool: Interactive probing of machine learning models. *IEEE transactions on visualization and computer graphics* 26, 1

(2019), 56–65.
[69] Womanity. 2023. Womanity Award. https://womanity.org/womanity-award/
[70] Pak-Hang Wong. 2020. Democratizing algorithmic fairness. *Philosophy & Technology* 33 (2020), 225–244.
[71] Denis Wood and John Fels. 1992. *The power of maps.* Guilford Press.
[72] 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.