Designing Interactive Systems for Community Citizen Science

Yen-Chia Hsu

CMU-RI-TR-18-29
June 2018

The Robotics Institute
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

Thesis Committee:
Illah Nourbakhsh, Chair, CMU RI
Aaron Steinfeld, CMU RI
Jeffrey Bigham, CMU HCII
Eric Paulos, UC Berkeley EECS

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Keywords: Community citizen science, visualization, crowdsourcing, computer vision, machine learning, artificial intelligence, interaction design, human-computer interaction, sustainable HCI, adversarial design, participatory design, community engagement, community empowerment, civic engagement, air quality, interactive storytelling, environmental health, public health, ubiquitous computing, mobile computing

For the current and future generations of humanity.

Abstract

Citizen science forges partnerships between experts and citizens through collaboration and has become a trend in public participation in scientific research over the past decade. Besides this trend, public participation can also contribute to participatory democracy, which empowers citizens to advocate for their local problems. This strategy supports citizens to form a community, increase environmental monitoring, gather evidence, and tell convincing stories. Researchers believe that this "community citizen science" strategy can contribute to the well-being of communities by giving them the power to influence the general public and decision makers.

Community citizen science requires collecting, curating, visualizing, analyzing, and interpreting multiple types of data over a large spacetime scale. This is highly dependent on community engagement (i.e., the involvement of citizens in local neighborhoods). Such large-scale tasks require the assistance of innovative computational tools to give technology affordance to communities. However, existing tools often focus on only one type of data, and thus researchers need to develop tools from scratch. Moreover, there is a lack of design patterns for researchers to reference when developing such tools. Furthermore, existing tools are typically treated as products rather than ongoing infrastructures that sustain community engagement.

This research studies the methodology of developing computational tools by using visualization, crowdsourcing, and artificial intelligence techniques to support the entire community engagement lifecycle, from initiation, maintenance, to evaluation. This research will make methodological and empirical contributions to community citizen science and sustainable human-computer interaction. Methodological contributions include detailed case studies with applied techniques from information technology systems that are deployed in real-world contexts. Empirical contributions include generalizable empirical insights for developing interactive systems that integrate multiple types of scientific data.

In this dissertation, I first define "community citizen science" and explain corresponding design challenges. Then, I review existing computational tools and techniques that are related to this research. Next, I present four interactive systems centered around the research scope: (1) a timelapse editor that supports building evidence-based narratives, (2) an air quality monitoring system that integrates heterogeneous data and computer vision to support the formation of scientific knowledge, (3) a visualization tool that reveals the impact of oil and gas development, and (4) a mobile crowdsourced application for reporting and visualizing pollution odors. Finally, I synthesize findings from all four works into generalizable design implications for future researchers and developers.

Acknowledgments

I greatly thank my thesis advisor, Illah Nourbakhsh, for the support and guidance of this research. I first met Illah on his Fall 2012 human-robot interaction course and had become interested in engaging users with robotics technology. After talking to Illah about my interest, I am very grateful that he gave me an opportunity to work with the CREATE Lab. Illah is a great advisor and role model for me. He encouraged me to explore and implement concepts that I was interested. The timelapse viewer project that I first participated in the lab is a significant milestone for me to realize that interactive systems can engage people on a vast scale. This project motivated me to apply for the Ph.D. program and further inspired me to study the topic of empowering communities with interactive systems. During my Ph.D. career, Illah not only provided academic feedback but also encouraged me when I felt frustrated about research results. I also deeply appreciate his patience and faith, which is very considerable for me to overcome my stress when writing this dissertation.

I am very grateful to other thesis committee members, Aaron Steinfeld, Jeffrey Bigham, and Eric Paulos, for advising and encouraging this research. I thank Srinivasa Narasimhan for serving on my Ph.D. speaking qualifier committee and the guidance for computer vision techniques. I am thankful to Randy Sargent for the technical guidance and support. I thank Jennifer Cross for serving on my Ph.D. speaking qualifier committee, advising all my qualitative studies, and collaborating with me in writing research papers. I acknowledge Ryan Roderick in the Global Communication Center for proofreading this thesis and tutoring my writing skills. I greatly thank Paul Dille for offering extensive system development support and proofreading this thesis. I am thankful to Beatrice Dias and Emily Hamner for the advice of obtaining IRB approvals. Thank you, Beatrice Dias, Jessica Pachuta, and Jennifer Cross for talking to me when I was stressful during thesis writing. I acknowledge the technical suggestions and assistance from Michael Tasota, Christopher Bartley, and Michael Taylor. I am grateful to Yen-Chi Chen (陳彥吉) for the advice on statistical analysis. I appreciate the support from other CREATE Lab members, the Heinz Endowments, and all other communities, especially the Allegheny County Clean Air Now.

I would like to thank my college advisor, Tay-Sheng Jeng (鄭泰昇), when I studied architecture design in Taiwan. Tay-Sheng enlightened me to conduct research in the field of human-computer interaction and motivated me to apply for my master program at Carnegie Mellon University. I would like to also thank my master program advisor, Mark Gross, for strengthening my skills in implementing systems. I also extend my thanks to all my friends who kept me encouraged and hopeful, especially Cheng-An Pan, Chia-Hsun Lee, Shih-En Wei, Hui-Chun Chen, Chun-Liang Li, Olive Ho, Ting-Yao Hu, Carol Cheng, Jen-Hao Chang, Po-Wei Wang, Chieh Lin, Ting-Hao Huang, Chen-Hsuan Lin, Chieh Lo, Guan-Horng Liu, Sz-Rung Shiang, Ching-Hang Chen, and Meng Chen.

Finally, my deepest thanks are for my family. I thank my single parent, Chiu-Liang Chou (周秋良), for the love and assistance that she has given me all the time for pursuing my dreams. I thank my grandparent, Te-Shun Chou (周得順), for the support that he provided when I pursued advanced academic degrees. I thank my spouse, Shao-Wen Chiu (邱劭雯), for her patience and love every day. I am very grateful that she kindly responded to my numerous complaints and provided considerable mental and emotional support for me to overcome my stress during dissertation research.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In a democratic process, citizens must advocate for regulatory changes related to social, political, or environmental issues, such as inequality, urban renewal, or air pollution. To influence and convince stakeholders, such as regulators, businesses, and the general public, citizens often need to forge and present reliable scientific evidence. Scientific evidence can be formed from multiple types of data, which typically need to be collected and curated over a large geographic area and an extended period. However, citizens' efforts to collect and curate data are often limited since they lack sufficient technological fluency and need to seek assistance from experts in governmental agencies, academic institutions, business companies, or non-governmental organizations. Such large-scale collaborative tasks require community engagement and the intervention of modern information technology, which includes visualization, crowdsourcing, and artificial intelligence techniques. These three techniques play different critical roles when initiating, maintaining, and evaluating community engagement. Applying these techniques poses new challenges related to data quality, science communication, and evaluation metrics. This research studies the methodology of developing and using information technology systems to democratize scientific knowledge over the entire course of community engagement. The primary research question is:

- **How can we design interactive systems with visualization, crowdsourcing, and artificial intelligence to support the engagement lifecycle in community citizen science?**

In this chapter, I begin with framing the research scope, **community citizen science**, which is a subfield of citizen science. Then, I explain the design principle and challenges of this research. Next, to be successful in the intervention of information technology, I discuss the importance of considering the lifecycle of community engagement, which forms the research question. Finally, I present the contributions of this research.

1.1 Research Scope

I frame the scope of this research under the field of designing interactive systems to support **citizen science**, especially when used for addressing issues related to the public good in civil society. In general, citizen science refers to empowering amateurs and professionals to form partnerships and produce **scientific knowledge** through actual participation or collaboration [27, 28, 80, 159, 196]. Gaining scientific knowledge is one of the most significant differences between

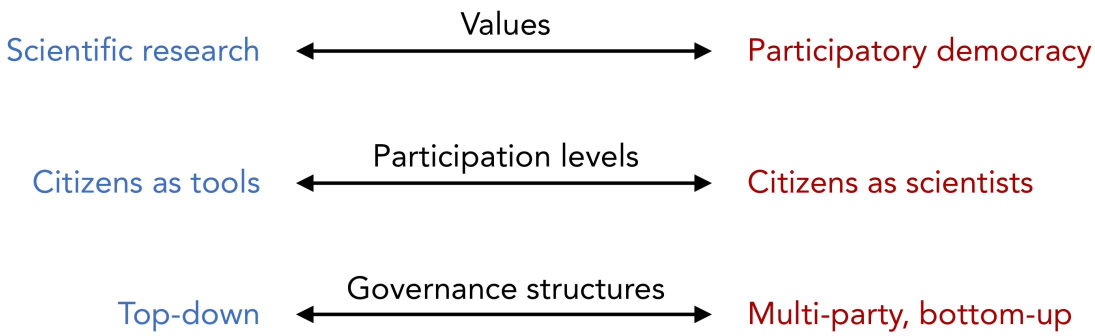


Figure 1.1: Citizen science has two main strands: research-oriented (the left part in this figure) and community-oriented (the right part in this figure). This research focuses on the latter.

citizen science and conventional public communication methods, such as newsletters or public hearings. Researchers and policy-makers often rely on scientific knowledge to provide answers. Besides treating science as a pure problem-solving activity, scientific knowledge and endeavor can also contribute to revealing the condition of a social, political, or environmental problem and encourage follow-up critical discussions.

There are three dimensions of citizen science research projects: values, participation levels, and governance structures (Figure 1.1). One can consider each citizen science project as a data point in a space that is spanned by these dimensions. It is essential to consider these dimensions since projects which fall in distinct subspaces have various goals, which in turn require different design principles and pose different challenges. In the following subsections, I discuss these dimensions and define a subspace, **community citizen science**, as the research scope.

1.1.1 Values

Citizen science has different scientific, educational, societal, and policy-making values [196]. These values determine the type of impacts and contributions that a citizen science project makes to the civil society. Historically, there are two main approaches [52]:

- **Scientific research**, which raises awareness and increases public understanding of science by spreading knowledge among common people [25, 26, 47, 53, 62, 64, 161, 204].
- **Participatory democracy**, which empowers lay people to represent their needs, address community concerns, and influence policy-making by producing and exchanging scientific knowledge [44, 93, 115, 116, 117, 171, 174, 209, 210, 228].

The **scientific research** approach has the goal of solving large-scale scientific research questions which are infeasible for scientists to tackle alone. Research questions under this approach are often propelled by professional scientists. Researchers applying this approach study how scientists can encourage the public to participate in scientific research, asking "What can common people do for professional scientists?" This scientific research approach was originated from Rick Bonney in the Cornell Lab of Ornithology, which uses citizens science to engage people in bird watching and environmental conservation [25, 62]. In contrast, the **participatory democracy** approach aims to democratize science by equipping citizens with tools to directly target community concerns for advocacy. Research questions under this approach are often driven by community members. Researchers applying this approach explore how scientists can engage in social and ethical issues that are promulgated by citizens, asking "What can professional scientists do for common people?" This participatory democracy approach originates from Alan Irwin and emphasizes scientific citizenship through openness and transparency of scientific governance and public engagement [115, 117].

This thesis adopts the **participatory democracy** approach to explore how information technology can be used to strengthen the link between scientific research and civil society. Besides asking how citizens can be engaged to act as scientists, this research asks how scientists can think and act as citizens. Rather than involving communities to participate in solving scientific research questions, the primary focus is to address community concerns directly. This research seeks to generate scientific evidence from community data to support citizen-driven exploration, understanding, and dissemination of public good concerns. The ultimate goal is to empower citizens to advocate for themselves in solving local specific issues, which may be social, environmental, or political.

1.1.2 Participation Levels

Citizen science research has various participation levels between scientists and citizens [25, 53, 101, 200, 225, 227]. Participation levels determine the level of investment and effort to participate in a citizen science project. They typically include five stages:

- **Problem definition**, which involve form (e.g. site or environment), function (e.g. people, activities, or relationships), economy (e.g. budget, operating costs, lifecycle costs, or available resources), and time (e.g. past, present, or future) [176]
- **Plan development**, which involve data collection protocols, system development and deployment, or steps of community actions
- **Data collection**, which involves collecting or labeling multiple types of data, such as sensor readings, images, or crowdsourced reports
- **Data analysis**, which involves information visualization, exploratory data analysis, hypothesis testing, machine learning, or computer vision
- **Decision making**, which involves data interpretation, formation of scientific knowledge, storytelling, or policy-making

One can consider participation along a spectrum that ranges from **citizens as tools** to **citizens as scientists**. At one end of the spectrum, scientists treat volunteers as powerful tools that provide or label data, which is also called crowdsourcing [79] or citizen cyberscience [95]. One such example is *Galaxy Zoo*, which utilized the knowledge collected from volunteers to classify a large number of galaxies via a web-based platform [152]. The classification result was aggregated from various participants with different weights, which were determined by their tenancy to agree with the majority. Volunteers were not directly involved in the design process of the platform. At the other end of the spectrum, scientists treat volunteers as collaborators over the entire project lifecycle, which involves almost all stages listed above. One such example is the *Neighborhood Networks* project, a participatory design practice which used robotics and sensing technology to engage residents in collecting environmental data and making critical discussions about local environmental concerns [68, 70]. Learning how to use the technology and apply it to real-world contexts happened together during the iterative design process. Participants in this project were involved in both designing and using the computational tools.

This thesis takes the concept of **citizens as scientists**, where volunteers and scientists establish a strong partnership through collaboration and engagement. Participants are treated as co-designers that bring diverse skills and expertise from multiple disciplines to inform the design of computational tools. This thesis studies how researchers take on the role of supporters that facilitate utilizing and disseminating technology, instead of supervisors that oversee and control the entire community engagement procedure. Design decisions are dynamically adjusted in response to the needs of communities. This co-design approach can grant the flexibility to discover, leverage, and adopt new design implications that are necessary for achieving the goal of democratizing science envisioned by Alan Irwin [115, 117].

1.1.3 Governance Structures

Citizen science research has different governance structures, which affects the power relationships between communities, governments, businesses, and non-governmental organizations [51]. A citizen science project can be led by a central organization, multiple stakeholders, or a community. This corresponds to three following governance structures:

- **Top-down**, which means a central organization or government invites local communities to contribute data or participates in the decision-making process
- **Multi-party**, which means multiple stakeholders (e.g. citizens, local communities, academic institutions, non-profit organizations, businesses, or government agencies) collaborate together in running a citizen science project
- **Bottom-up**, which means local communities initiate, organize, and lead grassroots movements about specific local issues and seek assistance from experts

One example of the **top-down** governance structure is *EyeWire*, where trained volunteers collaborate to provide knowledge for a machine learning model that reconstructs 3-dimensional representations of retinal neurons [125]. Errors in the volunteer-contributed data were resolved by majority agreements or inspected by domain experts. *EyeWire* was led by researchers from academic institutions with pre-designed engagement procedures. An example of the **multi-party** governance structure is the *Creek Watch*, a mobile and web application for users to report images and texts about local waterway conditions to assist water management policy-making [127]. The design process of *Creek Watch* integrated user needs and feedback from multiple stakeholders, such as government organizations, consulting companies, and local communities. An example of the **bottom-up** governance structure is the *Bucket Brigade*, which was a low-cost device for citizens to collect air samples [163]. Affected residents organized several communities to use the device for monitoring local air quality and measuring the impact of local industrial activities. This project was started in 1995 by Attorney Edward Masry, who sued a refinery for air pollution on behalf of a local community, and it was later maintained by advocacy groups.

This thesis focuses on empowering citizens to advocate for themselves using the **bottom-up** and **multi-party** governance structures. Local communities play a significant role in providing organizational network and disseminating critical findings to influence policy-making. Thus, I believe that these two structures are more suitable for linking technology to local concerns. The top-down structure is typically used for engaging the public in scientific research, rather than democratizing scientific knowledge. Moreover, these two structures align well with the concept that the researcher acts as a supporter and treats community members as co-designers. On the other hand, the top-down structure often considers citizens as data-contributors.

1.2 Design Principle

In Section 1.1, I discuss the space of citizen science research projects, which can be defined by three dimensions: values, participation levels, and governance structures. Projects which are located in difference subspaces have specific goals. When developing information technology to achieve particular goals in these projects, researchers accordingly adopt two opposing design principles: **consensus** and **agonism**. The following subsection discusses these design principles, and this thesis focuses on the **agonism** design principle.

1.2.1 Consensus versus Agonism

The **consensus** design principle supports structured deliberation to improve the situation of decision makers. This principle is often applied to citizen science projects which have scientific research values. Research questions and directions under this principle are often defined by professionals in academic institutions, non-governmental organizations, or government agencies. One notable project that uses the consensus principle is *eBird*, which provided computational tools for birdwatchers, scientists, and policy-makers to collect, visualize, and analyze bird data collaboratively [213, 214]. The tool documented bird migration timing, seasonal occurrence, and relative abundance on various spatial and temporal scales. This information can be used for not only promoting bird science education to the public but also prioritizing species-specific conservation actions. Another example is the *Community Resource Messenger*, which applied ubiquitous computing to facilitate and improve the communication between staff and residents at a shelter for homeless mothers [146]. Shelter staff could send residents information about available services and appointment reminders via text messaging, and residents could ask their case managers or therapist for assistance. Another instance, *Tiramisu*, was a transit information system for collecting GPS location data and problem reports from bus commuters [230]. Based on the user-contributed location data and the official bus schedule information, the system provided real-time bus arrival estimation, which was a significant concern for commuters. *Tiramisu* demonstrated that citizens could contribute valuable data to increase public service. In these three examples, user groups can generally be categorized as service providers (e.g., regulators, staffs, academic researchers) and consumers (e.g., citizens, amateurs), where their power relationships tend to be more equal and balanced. Although projects applying the consensus design principle opened up discussions of research results and interpretations to citizens, it rarely enabled those citizens to negotiate or reframe problem definitions, research questions, or research goals.

In contrast, this thesis embraces the **agonism** design principle, which directly targets social, political, or environmental issues that community members need to advocate for themselves [66, 164]. This principle aligns with community citizen science projects which have participatory democracy values. Research questions and directions under this principle are often defined by citizens in communities. One project that uses the agonism approach is *Feral Robot*, which served as low-cost mobile sensor network nodes for grassroots communities to collect, map, and present chemical pollution data in a local park [140, 141]. The goal of *Feral Robot* is to empower affected residents to apply robotics sensing for social activism around local environmental concerns. By measuring invisible environmental factors, such as air quality and noise, residents can make sense of the impact of pollution in the context of their neighborhoods. The air quality monitoring project, *Bucket Brigade*, discussed in subsection 1.1.3 is also an instance of the agonism design principle. In these examples, the power relationships between service providers and consumers tend to be more contradictory and unbalanced. A typical case is that citizens affected by industrial activities may not be satisfied with the actions taken by government regulators. The agonism design principle leverages street science [54] and adversarial design [67]. Street science emphasizes fusing local and professional knowledge to produce scientific evidence which can inform or challenge policy-making. Adversarial design promotes critical political discussions and challenges the current unbalanced power structure between citizens, governments, and businesses. The agonism design principle adopted in this thesis is not to support the mechanism and procedure of governance, but rather to open up debates and improve the condition of society through openness and transparency.

1.3 Design Challenges

The main goal of my research scope, **community citizen science** (see Section 1.1), is to democratize scientific knowledge through the **intervention of information technology** and to shape more equitable power relationships between citizens and stakeholders in a measurable way. To achieve the goal, this research embraces the **agonism** design principle (see Section 1.2). Agonism has a close relationship with public participation in politics, where citizens organize a community (or the public) around certain issues to advocate for themselves [57]. The agonism design principle embraces agonistic plurality, controversy, and variability in communities and thus enables the articulation of community concerns, such as social, political, or environmental problems. However, while researchers intend to enable citizens to generate scientific evidence with interactive systems, they are unable to accurately predict if citizens will contribute sufficient data needed for drawing meaningful insights. It is also difficult to determine the critical amount of human effort required for extracting knowledge [230]. Moreover, there are various methods of collecting, presenting, and using the data. It is not feasible to explore and evaluate all possible methods without deploying the system in the real-world context. These challenges, combined with other conditions, form a "wicked" problem [50, 190]. The concept of a wicked problem is in contrast to a "tame" problem, which can be understood and solved by following existing methodologies. A **wicked problem** has the following properties:

- Each problem is unique and dependent on context.
- There is no alternatively enumerable set of solutions.
- Each problem cannot be understood unless the corresponding solution is framed.
- Each solution is a one-shot operation and has no opportunity for trial and error.

- There is no optimal solution or stopping rule.
- There is no right or wrong, only good or bad.

To reveal and address a wicked problem, it is essential to obtain reliable **scientific knowledge**. However, forming scientific knowledge is not a trivial process. It may require using large-scale data over long periods of time and broad geographical areas. This is where the **intervention of information technology** becomes prominent. The interactive systems are designed to influence citizen participation and reveal local community concerns simultaneously, which is different from observational studies that use existing data, such as one that correlated air quality keywords from social media contents with environmental sensor measurements [85]. Due to the nature of wicked problems in community citizen science, traditional software design principles that tackle clear user needs are not suitable.

This thesis considers interactive systems as an **ongoing infrastructure** to sustain communities over time (as mentioned in [57]), rather than a software product which solves a single well-defined problem. This design principle is similar to how architects and urban designers address wicked problems. When approaching a community or city-scale problem, architects and urban planners first explore problem attributes (form, function, economy, and time [176]) and then design specific solutions (e.g., buildings or urban infrastructures) based on prior empirical experiences. I translate this idea to provide technology affordances, as mentioned in [90, 92, 166, 167], for seeking and revealing the condition of the local air quality problems, past and present. Technology affordance refers to the possibilities that information technology offers the people who may use and interact with it. Such affordances involve collecting, visualizing, analyzing, and interpreting multiple types of data over a large spatial and temporal scale as scientific evidence. This poses three challenges: data quality, science communication, and evaluation metrics.

- **Data quality**: how can we collect and curate large-scale heterogeneous data efficiently?
- **Science communication**: how can we use multiple types of data to form and share scientific knowledge?
- **Evaluation metrics**: how can we evaluate the intervention of information technology in a real-world context?

In the following subsections, I explain the importance of these challenges and their corresponding needs in order.

1.3.1 Data Quality

How can we collect and curate large-scale heterogeneous data efficiently?

Modern citizen science research often requires collecting data over a large temporal and spatial scale. As mobile, sensing, and information technology proliferates, citizens now have tools to gather data, such as sensors, cameras, and mobile devices. However, collecting data manually by using these tools has several problems. First, the process is time-consuming and laborious. For instance, to collect evidence of air pollution over a long period, citizens use cameras to take pictures of smoke emissions. Nevertheless, it is difficult to capture air quality violations in this way since smoke emissions can happen at any time. Having citizens monitor potential pollution sources 24/7 is infeasible. Second, the process can be error-prone and subject to bias. For example, the large amount of collected air quality data may contain missing values, have different formats or units, and be separated from multiple locations. This makes it difficult to retrieve information. Without proper curation, these data are unusable. Third, citizen-generated data often contain personal information, such as residential locations and IP addresses, which raises serious privacy concerns. Therefore, there is a strong need to develop computational tools to automate the data collection process [165], improve data quality [26, 27, 63, 101, 107, 159, 189, 227], make data retrievable [63, 165], and address privacy concerns by de-identifying personal information [159].

1.3.2 Science Communication

How can we use multiple types of data to form and share scientific knowledge?

Besides collecting and curating data, citizens need to present and share convincing stories about local problems to journalists, regulators, and the general public. Convincing stories often integrate personal experiences and scientific evidence, which requires computational tools to visualize, analyze, and make sense of large-scale heterogeneous data. This leads to several new challenges. First, to build a system which enables forming scientific knowledge, there is a lack of reusable computational tools. Existing tools are often designed to handle only one type of data, such as images. Thus, researchers need to develop new tools to integrate multiple types of data from scratch, which requires a considerable number of resources. Second, even with these new computational tools, analyzing data manually with pure human power can take a tremendous amount of time. For instance, it is challenging to ask citizens to find smoke emissions in videos captured from monitoring cameras over a long period. Third, citizen science data often depend on time and contain human errors. Independent variables (predictors) can have high dimensions or be correlated. These ill-conditioned data make it hard to apply traditional statistical methods for pattern recognition. In sum, there is a strong need to develop reusable tools for visualizing and analyzing large-scale heterogeneous data [26, 27, 159, 165] and to use automatic approaches (e.g., machine learning or computer vision) for knowledge discovery and extraction [22, 107].

1.3.3 Evaluation Metrics

How can we evaluate the intervention of information technology in a real-world context?

Measuring information and communication technology (ICT) interventions in community advocacy is generally challenging. Community advocacy has the ultimate goal of policy change, yet it is difficult to causally prove how critical to a successful policy change the communities' actions have been. For example, how can we tell if an air quality monitoring system promotes a sustainable power relationship among communities, governments, and industries in the long term, even when community concerns are addressed, or policy goals are achieved eventually? Such systems succeed in how the behaviors and attitudes of citizens, policy-makers, and businesses change and how the relationships among them evolve. Moreover, because local problems of community concerns are often wicked [50, 190], it is infeasible to follow a prescribed design guideline to reach a solution directly. The solution that is adopted in one context may not be suitable for another similar context. Under various real-world contexts, stakeholders often have dramatically different views of interpreting wicked problems. Both constraints of a wicked problem and the required resources to solve it can also change over time. This means that addressing a wicked problem requires a large number of people to change their mindsets, and thus there is no opportunity for testing by trial and error. Without real-world deployment and evaluation metrics, it is difficult to provide insights that can inform future researchers in developing such systems. In summary, there is a strong need to deploy information technology [32], evaluate the scientific, social, and political impact of such technology in a real-world context [26, 32, 210], and provide generalizable design insights [159].

1.4 Research Question

In Section 1.3, I explain that forming scientific knowledge is essential to understanding and addressing wicked problems and requires the **intervention of information technology**. The success of adopting interactive systems in community citizen science is highly dependent on **community engagement** [208], the involvement of citizens in local neighborhoods. Community engagement, which is also called public engagement, can have a significant impact on democratic governance [210]. There are various definitions of a public (community). This research applies the definition that a public (community) is formed by citizens who are indirectly or directly affected by negative or positive issues in civil society and are dedicated to making sure that these issues get recognized and resolved [61]. These issues can be social, political, or environmental. With the support of modern information technology, community engagement protocols have the potential to empower citizens in addressing these issues, which require collecting, analyzing, and interpreting data over large space-time scales. This process can scale local and professional knowledge up to a level capable of producing scientific evidence for communities to take political action [54].

The intervention of information technology in this research focuses on three techniques: **visualization**, **crowdsourcing**, and **artificial intelligence**. These techniques can contribute to democratizing scientific knowledge and supporting community engagement. First, by using visualization techniques to map data into visual elements, humans can explore and interpret data intuitively with perceptual skills, which may help communicate facts and promote public engagement [104, 105]. Also, visualization can be integrated with other types of media into compelling and memorable stories [135, 155]. Sharing and presenting these stories is a powerful way to raise public awareness about certain issues and to serve as evidence for convincing stakeholders. Second, crowdsourcing techniques can distribute tasks to crowds and aggregate ideas from those crowds to produce scientific knowledge collaboratively to facilitate decision making [29, 79]. The scalability of crowdsourcing techniques can engage diverse community members in the problem-seeking process over large geographical areas and long periods of time. User-generated content from crowdsourcing not only can be used to index sensor and image data but also can be valuable for indicating levels of public participation and engagement. Finally, artificial intelligence techniques, including machine learning and computer vision, can automate repeated processes and significantly reduce the workloads for citizens to organize scientific evidence. The predictive power of machine learning models can provide feedback to users, such as sending push notifications to inform users of the potential presence of events. Machine learning is also a

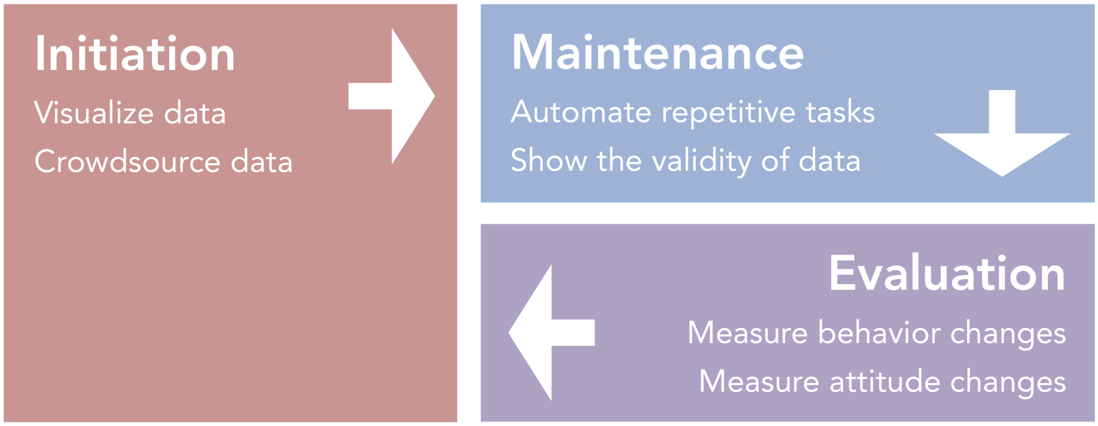


Figure 1.2: This figure shows the concept of the community engagement lifecycle.

promising approach for identifying and interpreting patterns among multiple types of citizen science data, which are typically high-dimensional, correlated, biased, and noisy [17, 22, 107, 124].

Rather than treating information technology as a product, this thesis treats it as ongoing infrastructure to sustain communities rather than products (see Section 1.3). Therefore, when developing interactive systems to support community citizen science, it is important to consider the **engagement lifecycle** as a whole 1.2. The engagement lifecycle refers to the concept of viewing community engagement over the course of its entire life, which involves (1) initiating citizen participation in contributing data, (2) maintaining engagement for a long-term, and (3) evaluating the performance and impact of engaging with the system. This concept is inspired by the design process where architects and urban planners continue refining their methods in response to user behavior changes under various unique contexts at different phases. Based on the challenges and needs described in Section 1.3, this thesis explores methods for developing **information technology** with the **agonism** design principle to support the **engagement lifecycle**. The main research question is formulated as the following:

> **How can we design interactive systems with visualization, crowdsourcing, and artificial intelligence to support the engagement lifecycle in community citizen science?**

In the following subsections, I discuss the three stages of the community engagement lifecycle.

1.4.1 Initiate Community Engagement

How can we motivate citizens to participate in community citizen science?

This stage consists of data quality and science communication challenges that are discussed in Subsection 1.3.1 and 1.3.2. Communities with a high awareness of local issues may already have sufficient motivation to start collecting or providing data. However, when a community has low initial activation, there is a need to empower citizens in the community by using visualization and crowdsourcing techniques. First, by visualizing data, whose raw forms often do not make sense for the general public, community members can gain a better understanding of local issues. For example, visualizing particulate matter in the air helps citizens scientifically track and understand the air quality around their living areas. Second, crowdsourcing enables citizens to actively and efficiently contribute data by dividing a large-scale task into smaller micro-tasks. The act of participating can implant the concept that citizens can change conditions in their community. For instance, reporting industrial odors to local health departments via a mobile application can encourage citizens to pay attention to the air quality in their living environment in a scientific manner.

1.4.2 Maintain Community Engagement

How can we maintain public participation in community citizen science?

This stage consists of data quality and science communication challenges that are discussed in Subsection 1.3.1 and 1.3.2. When using information technology to produce scientific knowledge over large space-time scales, it is vital to sustain a sufficient level of motivation and participation of citizens. This requires adopting several approaches in designing visualization and crowdsourcing systems. First, showing perceived values to citizens and giving community members proper credit can reveal the impacts of their scientific endeavors, which may help improve their confidence in achieving their goals, such as policy changes. Second, simplifying tasks by using artificial intelligence, such as machine learning and computer vision, can significantly reduce workload for citizens, which may increase their willingness to participate. Third, supporting various participation levels for a wide spectrum of active and passive user behaviors allows citizens to contribute their time and efforts efficiently.

1.4.3 Evaluate Community Engagement

How can we evaluate the performance and impact of community engagement?

This stage consists of evaluation metrics challenges that are discussed in Subsection 1.3.3. Section 1.3 mentions that this research treats information technology systems as ongoing infrastructure in communities to sustain public participation in politics [57]. When using this concept to improve deployed systems and provide design insights to future researchers, it is important to evaluate the impact of the system by using qualitative or quantitative metrics. Community engagement has in general two major metrics: changes in external behaviors (variable X) and changes in internal attitudes (variable Y). The direction and strength of the relation between X and Y can be influenced by a moderator [46], the intervention of information technology (variable Z). Figure 1.3 shows the interrelationships among these three variables. In this research, the intervention of information technology refers to features in interactive systems. Behavior changes refer to patterns about how users interact with system features over time, which can be observed by analyzing the participation history from server log files or implementing built-in metrics, such as website traffic trackers. Attitude changes refer to how psychological conditions of citizens in a community evolve, which can be measured explicitly by using self-report surveys or interviews. Evaluating community engagement has several advantages. On a narrower level,

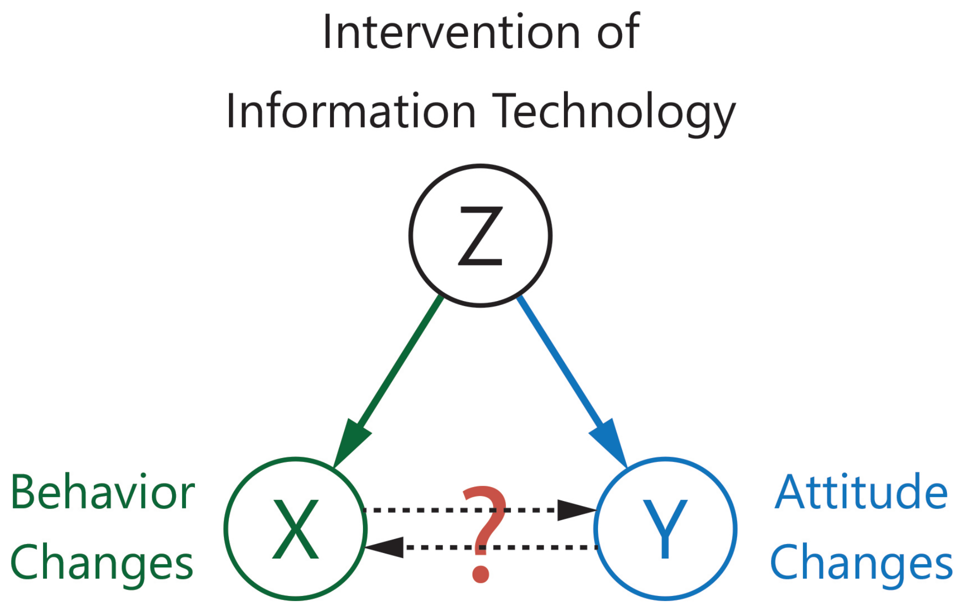


Figure 1.3: This figure shows the interrelationships among the intervention of information technology (Z) and two primary categories of community engagement metrics: observed behavior changes of citizens (X), and measured attitude changes of a community (Y). The intervention of information technology serves as the moderator [46] to influence the strength or direction of the relationship between the other two variables: behavior and attitude changes. It is possible to identify the corresponding behavior and attitude changes after deploying systems. However, the causation links between behavior and attitude changes are unclear.

systems can distribute different types of tasks to suitable groups of users based on their participation level. On a broader level, communities, researchers, organizations, or governments need quantitative or qualitative methodologies, techniques, and metrics to understand the impact of community citizen science projects.

1.5 Contributions

This chapter has defined the scope of this research (community citizen science), described the design principle (agonism), and explained design challenges (data quality, science communication, and evaluation metrics). Also, this chapter has framed the research question around initiating, maintaining, and evaluating community engagement by using information technology with visualization, crowdsourcing, and artificial intelligence techniques. The intervention of information technology has the ultimate goal of empowering communities to advocate for their local issues. To achieve this goal, it is essential to design and deploy such systems in real-world contexts, evaluate the impacts of such systems, and provide generalizable design methodologies and insights for future researchers. This research aligns with large-scale sustainable HCI (Human-Computer Interaction) [24, 37, 69, 71, 75, 158], which studies the intervention of information technology for increasing the awareness of sustainability, changing user behaviors, and influencing attitudes of communities as a whole. This thesis makes the following methodological and empirical contributions to **sustainable HCI** and **community citizen science**:

- **Methodological contribution**, including detailed case studies with applied methodologies of information technology systems that are deployed in real-world contexts to support community citizen science
- **Empirical contribution**, including generalizable empirical implications for developing interactive systems that integrate multiple types of scientific data, such as sensor readings, images, and crowdsourced content

1.6 Outline

This thesis is organized as follows. Chapter 2 reviews previous works related to the history of public participation, the importance of integrating human-generated and machine-generated data, and the approach of using computer vision and machine learning when analyzing data. Chapter 3, 4, 5, and 6 present four deployed interactive systems. Two of these systems focus on initiating community engagement, and the other two systems tackle the entire engagement lifecycle. Chapter 7 concludes all experiences and findings into generalizable design implications for future researchers to develop interactive systems that support community citizen science. The appendix consists of surveys for evaluating these deployed systems.

Chapter 2

Related Work

As mentioned in the introduction chapter, this research focuses on empowering local communities to collaboratively advocate for their local problems using scientific knowledge. Forming and democratizing scientific knowledge at a broad spatial-temporal scale requires community engagement in collecting, curating, visualizing, comparing, and making sense of data from heterogeneous sources. It is the power of modern interactive systems that make this data-driven community engagement approach viable. Developing such systems to improve technological fluency among citizens and foster public communication for participatory democracy is still an ongoing research topic [27, 159]. This chapter reviews techniques that can support the development of interactive systems, which involve crowdsourcing, visualization, and artificial intelligence. First, I start by describing an early approach of engaging communities in environmental epidemiology before the existence of the citizen science concept and modern information technology. Using historical examples, I highlight the significance of collecting and utilizing lay knowledge when addressing community concerns. Next, thanks to the development of modern information technology, citizens can now use environmental sensing devices, mobile computing applications, and web-based tools to crowdsource multiple types of human and machine inputs. These heterogeneous data sources, when being integrated together via visualization techniques, can provide better contexts that reveal local problems. However, these noisy, high-dimensional, and potentially correlated datasets are complex to analyze and interpret using traditional statistical methods. I then discuss approaches that apply artificial intelligence, including machine learning and computer vision, for automating repetitive tasks, predicting future events, and extracting knowledge from citizen science data.

2.1 Community Engagement in Epidemiology

Citizen science strategies are particularly valuable for addressing large-scale public health issues [58]. Historically, before the existence of the citizen science framework, community engagement had been applied in epidemiological research, which studied connections between human health and environmental factors. In a classic study of the cholera epidemic in London in 1854 [205], Dr. John Snow wrote: “Within two hundred and fifty yards of the spot where Cambridge Street joins Broad Street, there were upwards of five hundred fatal attacks of cholera in ten days.”

Contrary to the then-popular belief that cholera was spread through the air, Snow argued that sewage contamination caused the epidemic. To build this argument, Snow carefully documented the locations of affected houses, examined death cases, and verified if those cases accessed the contaminated water pump. Then, by plotting the location of water pumps and the death cases geographically on a map, Snow was able to provide convincing evidence that cholera was spread through contaminated drinking water. After presenting this evidence, the handle of the contaminated water pump was removed, and the outbreak quickly diminished.

Snow's method for using citizen-contributed lay knowledge to inspect how cholera spread in London has been widely adopted in epidemiology to identify the distribution of diseases and understand factors that affect the distribution. Researchers further adapted this approach to **popular epidemiology** and **community-based participatory research**, where citizens directly engage in gathering data and extracting scientific knowledge from these data for advocacy and activism [33, 34, 35, 150]. In the study of childhood leukemia cases that were clustered near contaminated water wells in Woburn, residents recruited epidemiologists to show the relationship between the risk of childhood leukemia and the hazardous chemicals in their drinking water [36]. The Woburn case led to the increase of national funding to clean up toxic waste sites and study the connection between human health and toxic contamination.

These historical examples suggest that gathering local experimental community data, which are often inaccessible to scientists, can link expert and lay knowledge to produce convincing scientific evidence. This data-driven evidence, especially when integrated with community narratives, is essential for citizens to make sense of local environmental issues and take community action [173]. This concept is especially beneficial when lay perspectives contradict professional ones, and thus advocacy or activism is needed to inform policy-makers about the perceptions and views of community concerns [34, 58]. This thesis research draws on popular epidemiology by including lay knowledge to track and interpret the distribution of local environmental concerns, but it is different in its use of computational tools to speed up the process and facilitate communication.

2.2 Computational Tools for Community Engagement

Due to the advancement of modern information technology, citizens can now collect data with computational tools to contextualize and express their concerns. There are typically two types of community data, generated from either **machine** or **human inputs**. Each type of data provides a small fragment of evidence. As discussed in the previous section, lay knowledge is vital to the conversation about community concerns. However, using human-generated data alone is not sufficient for producing convincing evidence in community citizen science. When it comes to resolving and revealing community concerns, human-generated data can show how real-world living experiences of residents are affected by local issues, but it is typically noisy, ambiguous, and hard to quantify at a consistent scale. Machine-generated data can complement human-generated data by providing temporally dense and reliable measurements that reflect the real situations of the surroundings, but it fails to explain how community members perceive and experience the environment. Most of the research effort has been concentrated on only one type of data. Without integrating both types of data, it is difficult to understand the context of

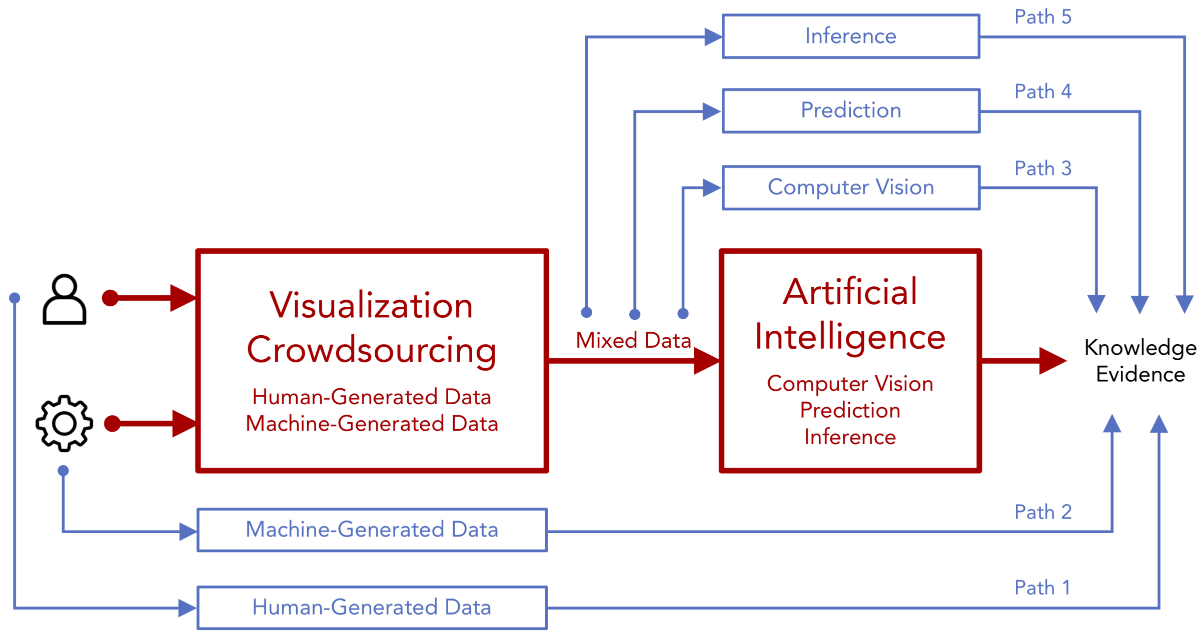


Figure 2.1: The bottom blue paths 1 and 2 depict a gap in integrating human-generated (subsection 2.2.1) and machine-generated data (subsection 2.2.2). The top blue path 3 shows the approach of using computer vision to support extracting patterns from image or video data. The top blue paths 4 and 5 show that there is a need to consider both prediction (subsection 2.3.2) and inference (subsection 2.3.3) when analyzing community citizen science data. The middle bold and red path demonstrates the approach that this thesis suggests for designing interactive systems that support community citizen science.

local concerns. In the following subsections, I discuss a range of computational tools that are designed to collect and visualize either human-generated or machine-generated data, as shown in the bottom two separate blue paths in Figure 2.1.

2.2.1 Human-Generated Data

Human-generated data includes personal observations contributed by community members. Modern computational tools enable citizens to provide and share volunteered geographic information as defined in [101]. Once this information is collected, the tools can aggregate these human-generated data to produce scientific knowledge and collective intelligence [20, 29, 215]. For example, *Ushahidi* is an open-source online platform that crowdsources crisis information via text messages or its website to provide timely transparent information to a broader audience [169]. The platform allowed citizens to share information that reflected real situations, such as incidents of violence after an election or protest, when rumors and doubts were prevailing in the mainstream news media. Another example is *NoiseTude*, a mobile application that empowered citizens to report noise via their mobile phones and map urban noise pollution on a geographical heatmap [76, 156]. The tool could be utilized for not only understanding the context of urban noise pollution but also measuring short-term or long-term personal exposure, which might benefit large-scale epidemiological research.

In addition to aggregating data, computational tools can facilitate community engagement, and community members can provide feedback for developers to refine the tools iteratively. For instance, *Creek Watch* is a monitoring system which enabled citizens to report water flow and trash data in creeks [127]. The system consisted of a mobile application for collecting user-generated content (e.g., images) and a website for visualizing and sharing data by using a map and a table. The iterative design process involved regulators. User studies showed that participants were satisfied with the data quality and believed that the system would promote public engagement and education in watershed health. Another instance is *Sensr*, a tool for creating environmental data collection and management applications on mobile devices without programming skills [129, 131]. Non-profit organizations were involved in the design iterations. Project managers could use the framework to create a campaign website around a specific issue, such as air quality. Community members could report data with geographic information, such as images, via a mobile application. Case studies that created citizen science campaigns with this tool indicated the need to control data quality, support both dense and sparse citizen participation, and create data-driven narratives to facilitate communication.

Besides engaging local communities, computational tools also afford data collection at a vast spacial-temporal scale. For instance, *Encyclopedia of Life* is a platform for curating species information that was crowdsourced from professionals and non-expert volunteers [192]. Curators could comment and make "trust," "untrust," or "hide" decisions on the user-generated content. Content providers could then improve the data based on this collaborative feedback. User studies were conducted by using grounded theory [212] to categorize key design recommendations, which included establishing sub-communities based on user interests, endorsing the scientific value of community-contributed data, distributing tasks to the right people, attributing the effort of collecting data to users, and rewarding contributors. Another example, *eBird*, is a crowdsourcing platform to engage birdwatchers, scientists, and policy-makers to collect and analyze bird data collaboratively [213, 214]. The platform enabled users to submit birdwatching data, such as dates, locations, and species. The data were visualized on a map to show frequency distributions and seasonal migrations of birds. Researchers indicated that the balance between quantity and quality, the accessibility, and the variety of data are key factors to make *eBird* successful.

The advancement of modern computational tools can support community members in contributing and visualizing human-generated data. From the examples shown above, when developing these tools, it is essential to understand what data stakeholders need, how to validate data quality, how to make data useful, and how to deliver data effectively. Researchers in these works considered tools as an integrated system, which supported different levels of participation, rather than individual and separate components. This insight pointed out the importance of considering how a computational tool could support the entire lifecycle of community engagement.

2.2.2 Machine-Generated Data

Machine-generated data include environmental measurements quantified with sensing devices. These sensors are often designed for citizens to run by themselves to monitor their surroundings collaboratively with minimal to no assistance from experts. Several previous works focused on gathering data from deployed sensing equipment. For example, Tian et al. implemented a low-cost and calibrated wearable sensor, *MyPart*, to measure airborne particles [220]. These real-time sensor readings were visualized on a mobile application for users to make sense of their air quality data. In a preliminary user study where participants took a specific route with the device, the result showed that visualization could engage participants in interacting with and interpreting air quality data. Another example is a fine particulate matter sensor, *Speck*, developed by Taylor et al. [217, 218]. The sensor was calibrated for operating indoors. The screen on the device enabled users to view the current estimated particle concentration as well as historical data. A user study revealed that participants had better awareness about their indoor environment after using Speck. The result also showed an increased level of confidence to mitigate the effect of poor indoor air quality.

Traditionally, systems that collected machine-generated data were evaluated based on performance and accuracy, but it is also critical to evaluate attitude and behavior changes in communities during the design process. For instance, Kuznetsov et al. developed a monitoring system which involved low-cost air quality sensors and a map-based visualization [138]. The system was deployed in four different types of communities for supporting public engagement and activism. User studies revealed that community members perceived the system as a tool for expressing and understanding local public health concerns, such as traffic exhaust and industrial air pollution. Another instance is an indoor air quality monitoring system implemented by Kim et al. [130]. The system used Arduino to gather air quality data from commercial sensors. The data were visualized together with other sensor data provided by AirNow [4]. User studies of system deployment found increased awareness of indoor air pollution problems and changes of habitual behaviors to improve indoor air quality, such as turning on a fan while cooking. Another example is a low-tech and low-cost paper sensing system, developed by Kuznetsov et al., to trap particulates in the air [139]. The system was deployed in a local air quality activist group. User studies showed that the system allowed community activists to observe various pollutants and understand how air pollution travels at local regions across different times of the day and week.

Insights from these previous works showed that sensing technology, especially accompanied with visualizations, could provide context and evidence that might raise awareness and engage local communities to participate in political activism. These works emphasized studying the influence of deploying systems in communities by analyzing changes in how users interact with information technology over time. Similar to these works, this thesis evaluates interactive systems not only according to system performance and usability, but also its real-world impact on communities regarding behavior and attitude changes.

2.3 Artificial Intelligence for Community Engagement

Modern computational tools, as discussed in the previous section, have enabled community members to crowdsource and visualize a large quantity of human-generated or machine-generated data. When the amount of data is small, visualization techniques are sufficient for community members to search and document evidence. However, the quantity of data in citizen science is often too large to be manually inspected through visualization. Multiple interrelated patterns may exist in the data, which may not be revealed initially in the visualization. Identifying and documenting all evidence from the visualization can take community members a considerable amount of time and effort. Moreover, there is skepticism about whether lay people can contribute valid and reliable data for scientific research, as mentioned in [27, 47, 58, 170]. Citizen science data are typically high-dimensional, noisy, potentially correlated, and spatially or temporally sparse. These characteristics make data validation and interpretation problematic when using traditional statistical methods. To convince stakeholders, researchers and community members need to not only automate the process of collecting evidence but also validate citizen science data when producing scientific knowledge [226].

To automate the process of searching and validating evidence from "big data", artificial intelligence provides promising techniques, including **computer vision** [86, 102, 216] and **machine learning** [23, 103, 119, 121, 162], to assist community members by identifying and documenting patterns in citizen science data. Computer vision enables extracting patterns from image and video data automatically, which can assist community members in collecting scientific evidence. Machine learning involves **prediction** and **inference** techniques [17, 22, 108, 201]. Prediction focused on increasing the performance of forecasting future events with sophisticated models. However, interpreting these models to identify patterns that are particularly relevant to local concern could be challenging. Inference is specialized for explaining models. However, using inference alone without considering prediction could over-interpret models due to overfitting the data, and the result could be poorly generalizable for other similar contexts. Thus, there is a need to integrate both approaches that complement each other when evaluating the impact of interactive systems, as discussed in [108, 201]. The following subsections discuss how previous works leveraged the power of computer vision, prediction, and inference, as shown in the top blue separate paths in Figure 2.1.

2.3.1 Computer Vision

Computer vision enables computers to make sense of image or video data and also extract patterns from them, which can augment human perception when completing certain tasks. Several previous works used information crowdsourced from participants to complement or enhance the computer vision algorithm. For example, *Glance* is a video coding system which asked crowd workers on Mechanical Turk to label small clips in parallel [144]. The system aggregated the labels from multiple workers based on their quality and levels of agreement. This crowdsourcing approach enabled researchers to analyze behaviors that were hard to detect with existing computer vision algorithms. Without searching the entire video, researchers could use the aggregated crowdsourcing labels to identify events quickly. Another example is *Zensors*, a mobile image recognition application that combined crowdsourcing and computer vision to answer user-defined questions [143]. Initially, answers were provided by workers on Mechanical Turk. The system then used these answers as labels to train a computer vision classifier, which could take the image recognition task when its accuracy reached a threshold for a specified period. Another instance, *VizLens*, is a mobile application that assisted visually impaired people on using various interfaces [98]. A blind person was first instructed by crowd workers to take a picture of the interface. Crowdworkers then labeled elements on the interface collaboratively. Next, a computer vision algorithm, which was trained by using these labels, took over the image recognition task. When the blind person touched an element on the interface, the application generated a spoken response based on the image recognition result.

These works emphasized the idea that using other data sources to index real-time or archived videos could help users in making sense of video content efficiently and reliably. This idea inspired this research to adopt computer vision to automate repetitive tasks, which made it easier for community members to gather visual evidence from the data. Also, this idea strengthened the approach, mentioned in the previous section, to blend multiple types of human-generated and machine-generated data for comparison.

2.3.2 Prediction

Prediction aims to forecast the future based on previous observations of predictors and responses. Several previous works used machine learning to predict air quality. For example, Zheng et al. developed a framework to forecast air quality readings of a monitoring station over the next 48 hours based on meteorological data, weather forecasts, and sensor readings from other nearby monitoring stations [229]. The framework consisted of linear regression and an artificial neural network that modeled temporal and spatial trends respectively. Their predictions were weighted according to weather conditions. Additionally, sudden changes of air quality were modeled by using a separate rule-based model. Another work, conducted by Azid et al., used principal component analysis and an artificial neural network to identify significant pollution sources and predict air pollution respectively [14]. Donnelly et al. combined kernel regression and multiple linear regression to forecast the concentrations of nitrogen dioxide over the next 24 and 48 hours [74]. Hsieh et al. utilized a graphical model to predict the air quality of a given location grid based on data from sparse monitoring stations [111]. Nodes in the graph indicated location grids at different spatial-temporal states, and edge weights represented correlations between these states. The graph was constructed by optimizing a loss function that modeled the distance between labeled and unlabeled distributions of air quality index.

The primary goal of applying machine learning in these studies was to increase the predictive power of the data. These works applied prediction techniques to help citizens plan daily activities and also inform regulators in controlling air pollution sources. This idea inspired this research to not only visualize data in interactive systems but also provide informative predictions as a way of demonstrating the perceived values of citizen-contributed data to encourage and maintain community engagement.

2.3.3 Inference

Inference refers to mining and extracting knowledge about the interrelationships between predictors and responses. Understanding how changes of predictors affect responses is essential in analyzing the impacts of environmental issues in the long-term [34, 58]. Several previous works focused on using machine learning to increase the explanatory power and infer potential patterns in the data. For instance, Gass et al. investigated the joint effects of outdoor air pollutants on emergency department visits for pediatric asthma by applying Decision Tree learning [89]. Predictors (air pollutants) were simplified from continuous values into quartiles. The decision trees, trained with the CART algorithm [149], were used as a supervised version of hierarchical clustering for explaining patterns and generating hypothesis, rather than predicting the future. The authors suggested using Decision Tree learning to hypothesize about potential joint effects of predictors for further investigation. Another work, conducted by Stingone et al., trained decision trees to identify possible interaction patterns between air pollutants and math test scores of kindergarten children [211]. The interaction terms were added to a linear regression model for estimating the effect. With control for confounding factors, the result showed relationships between isophorone exposure and math scores among children who lived in densely populated urban areas. Another study, conducted by Hochachka et al., fused traditional statistical techniques with boosted regression trees to extract species distribution patterns from the data collected via the eBird platform [107]. The results could be used to guide species conservation planning and management.

Applying inference techniques to find associations among variables is not trivial. Typically, selecting representative variables requires domain knowledge from experts or hypotheses generated when inspecting the visualization. These previous works utilized domain knowledge to fit machine learning models with high explanatory powers on filtered citizen science data. Results showed that understanding the structure of machine learning models can inform decision-making. These works inspired this research to extract knowledge from the data and study the relationships between predictors and responses, rather than merely concentrating on evaluating how well these models represent the data. The extracted knowledge can reveal local concerns and serve as convincing evidence for communities in taking action.

2.4 Summary

This thesis shows that the approach of integrating human-generated and machine-generated data can produce scientific evidence that lay people and experts can interpret. The interactive systems can further be enhanced by combining interactive design with artificial intelligence. The middle red path in Figure 2.1 demonstrates the concept of designing interactive systems for community citizen science. In the following chapters, I present four computational tools that demonstrate the value of integrating multiple types of data to provide narratives or evidence. Two of these tools specifically utilize artificial intelligence techniques to reduce the workload of community members, forecast future events to support personal decision-making, and interpret patterns of local community concerns.

Chapter 3

A Web-based Large-scale Timelapse Editor for Creating and Sharing Guided Video Tours and Interactive Slideshows

Scientists, journalists, and photographers have used advanced camera technology to capture extremely high-resolution timelapse and developed information visualization tools for data exploration and analysis. However, it takes a great deal of effort for professionals to form and tell stories after exploring data, since these tools usually provide little aid in creating visual elements. This chapter presents a web-based timelapse editor to support the creation of guided video tours and interactive slideshows from a collection of large-scale spatial and temporal images. Professionals can embed these two visual elements into web pages in conjunction with various forms of digital media to tell multimodal and interactive stories. The editor provides technology affordance for forming convincing scientific narratives that reveal critical landscape changes on the Earth, such as deforestation, coral bleaching, and drying lakes. This work addresses the science communication challenges (as mentioned in section 1.3) to initialize community engagement in community-oriented citizen science. The main contribution of the work is to provide reusable computational tools for forming and sharing scientific knowledge.

3.1 Preface

As camera technology proliferates, the quantity and resolution of digital images have increased exponentially. Researchers have worked on creating tools for generating, exploring, and sharing large-scale timelapses after capturing high quality images. For instance, Sargent et al. [193] have developed an integrated solution to capture and stitch gigapixel timelapses, generate multi-resolution video tiles, and visualize the results in an interactive web-based viewer. Professionals have used the technology to document the entire context of a site, capture extreme details in scenes, and share high resolution timelapses on the Internet. One example is the Google Annual Earth Timelaspe [1] consisting of 29 cloud-free mosaics of the planet from Landsat satellite imagery between 1984 and 2012, with each frame containing nearly 1 trillion explorable pixels.

Such visualization tools enable browsing large-scale images and provide powerful data exploration experiences to professionals. However, most existing tools lack the capability for professionals to create visual elements for data-driven storytelling through space and time, affording the presentation of a sequence of related facts found during exploration. If professionals want to create video tours, they need to import timelapses into video editing software. This process is impractical and takes a huge amount of time as most software cannot handle large datasets that do not fit into memory.

To address this problem, this work presents a web-based timelapse editor [77] operating along large-scale space and time to assist professionals in creating visual elements based on facts found during exploration. The editor allows the creation of **guided video tours** and **interactive slideshows** to enhance different story structures described by Segel and Heer [197]. Guided video tours present changes over time in author-driven stories, having linear visualization paths and limited interactivity. Interactive slideshows store various interesting locations and facilitate follow-up exploration for reader-driven stories, having little prescribed orderings and high interactivity. Each tour or slideshow is a micro-story and can be integrated by professionals with other types of media into a mega-story [118].

3.2 System

When scientists find interesting events while exploring the timelapse in the zoomable and pannable viewer, they can use the editor attached at the bottom of the viewer (Figure 3.1) to create guided video tours defined by sequences of keyframes containing the time, location, and scale of different views. Users can use the functions provided on the main control bar to add keyframes. On the left side of the viewer above the scale bar, there is a small box displaying the satellite image quality relative to a resolution to assist scientists in choosing appropriate scales. If a keyframe is unwanted or misplaced, users can select the keyframe and then delete it or drag it to a desired place in the sequence. Each keyframe in the container has auxiliary functions for users to update the keyframe to the current view, duplicate the keyframe, and add corresponding descriptions.

Users can click on the play button on the main control bar to preview the tour animated by using a linear motion for each pair of keyframes using default transition settings. In the keyframe container, users can specify transition parameters between two consecutive keyframes. There are two main types of transitions: speed and duration. A speed transition uses the user-defined playback rate relative to the original video rate and automatically computes the duration. The editor uses 100% speed as the default transition setting. In contrast, a duration transition calculates the speed accordingly from the user-defined duration. For short timelapses (e.g. less than 100 frames), users can assign a desired looping parameter, the number of times to loop through the entire timelapse video between two keyframes. While looping the entire video, the editor introduces a 0.5 second dwelling time for a better transition effect, meaning that the animation pauses at the very beginning and end of the timelapse for 0.5 second. By using the animation logic described above, users can perform the following five different camera motions:

- Animate zooming, panning, and time simultaneously by adding two keyframes at different locations and dates, and then setting speed or duration to a non-zero value.
- Pause zooming and panning but animate time by adding two keyframes at different dates but the same location, and then setting speed or duration to a non-zero value.

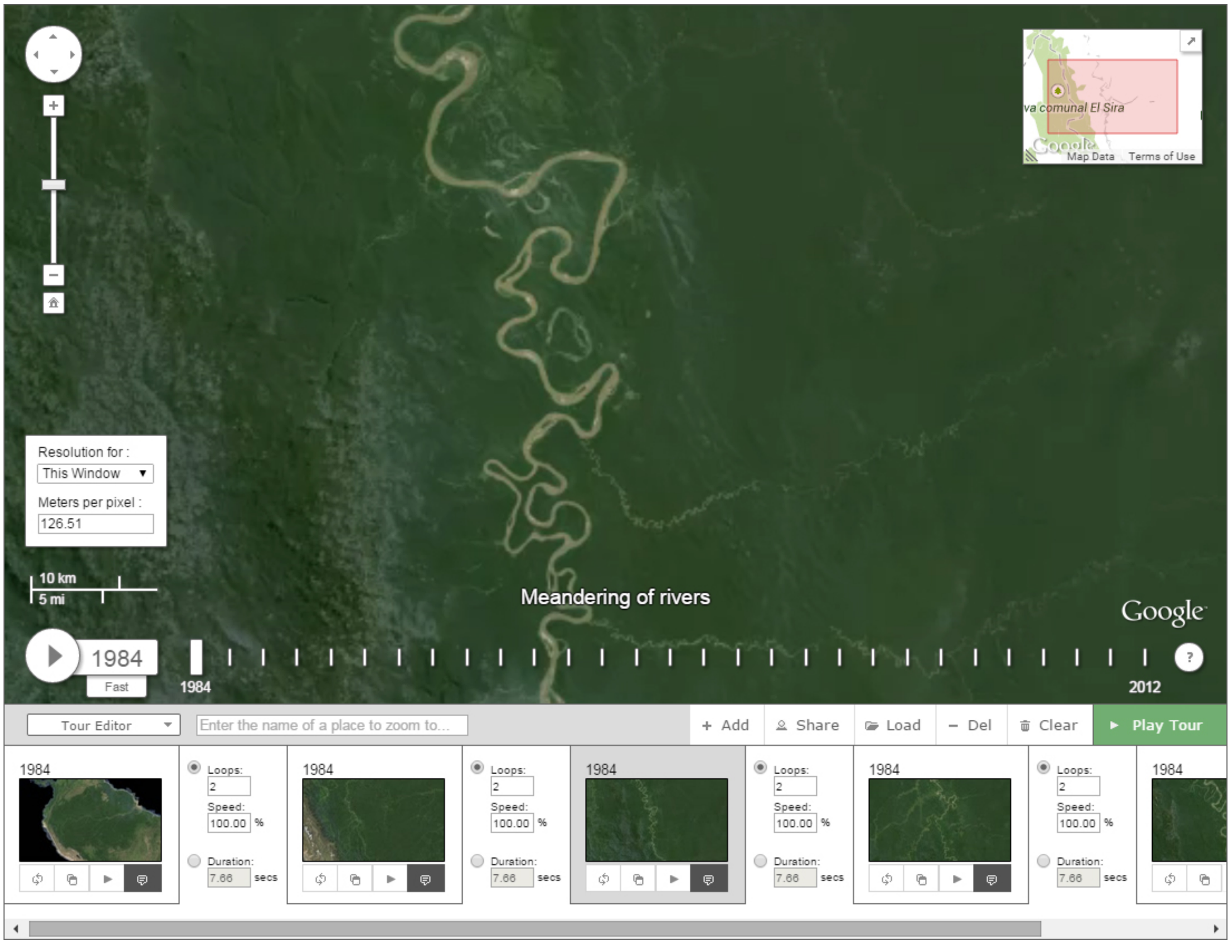


Figure 3.1: This figure shows the user interface of the timelapse editor. The top part is a viewer which shows a timelapse imagery dataset. Users can navigate the dataset by zooming and panning the viewer. There is a toolbar at the bottom of the viewer, which provides functions for editing a video tour or an interactive slideshow. The bottom part of the interface shows a series of keyframes, which compose the narrative, and the transition parameters among these keyframes.

- Pause time but animate zooming and panning by adding two keyframes at different locations but the same date, and then setting duration to a non-zero value.
- Pause zooming, panning, and time simultaneously by adding two keyframes cloned at the same location and date, and then setting duration to a non-zero value.
- Jump immediately from the first keyframe to the second one by forcing duration to be zero.

When finished editing, users can click the share button to disseminate or embed the guided video tour (Figure 3.2) encoded in an URL (uniform resource locator) into a storytelling webpage. Descriptions associated with each keyframe show up as video captions. The tour interface displays a time stamp, a scale bar, and a small Google map to provide contextual information. A button at the top left of the interface allows users to stop the tour.

Professionals can also use the editor to create interactive slideshows (Figure 3.3) containing a collection of locations. The workflow is similar to the one for creating video tours. The ed-

Figure 3.2: This figure shows the user interface of a guided video tour. The tour animates automatically by following a series of keyframes that are created in the editor.

itor turns keyframes into slides and omitting all transition parameters. Audiences first see an overview and then can choose an interesting location to zoom in and to request detailed information. When audiences hover a mouse onto a slide, a message box containing corresponding descriptions fades into the interface. Clicking on a slide animates the viewer to a keyframe representing an interesting location. Professionals can use interactive slideshows for storytelling in a webpage or for visual exhibitions on hyperwalls in museums.

3.3 Discussion and Summary

After releasing the editor in 2012, journalists at TIME magazine used the editor to create video tours for telling stories about extreme natural resources, global climate changes, and urban explosions on Earth. The tool rendered these tours into high-quality videos and then the journalist integrated these videos with the timelapse viewer and other forms of digital media into a compelling story [2]. Audiences first experienced a prescribed author-driven story with rich multimedia containing space-time tours and then were free to explore the timelapse by themselves. In addition, scientists in the Explorables team in Pittsburgh in 2014 created interactive slideshows for telling a reader-driven story about landscape changes along Taiwan's coastline over a two-decade period [84]. In 2014, scientists at the Exploratorium museum in San Francisco installed a visual exhibition showing an interactive slideshow on a hyperwall. These professionals were able to create tours or slideshows by using the editor over approximately half an hour of training and to use these visual elements in forming interesting stories. However, there are several open research questions:

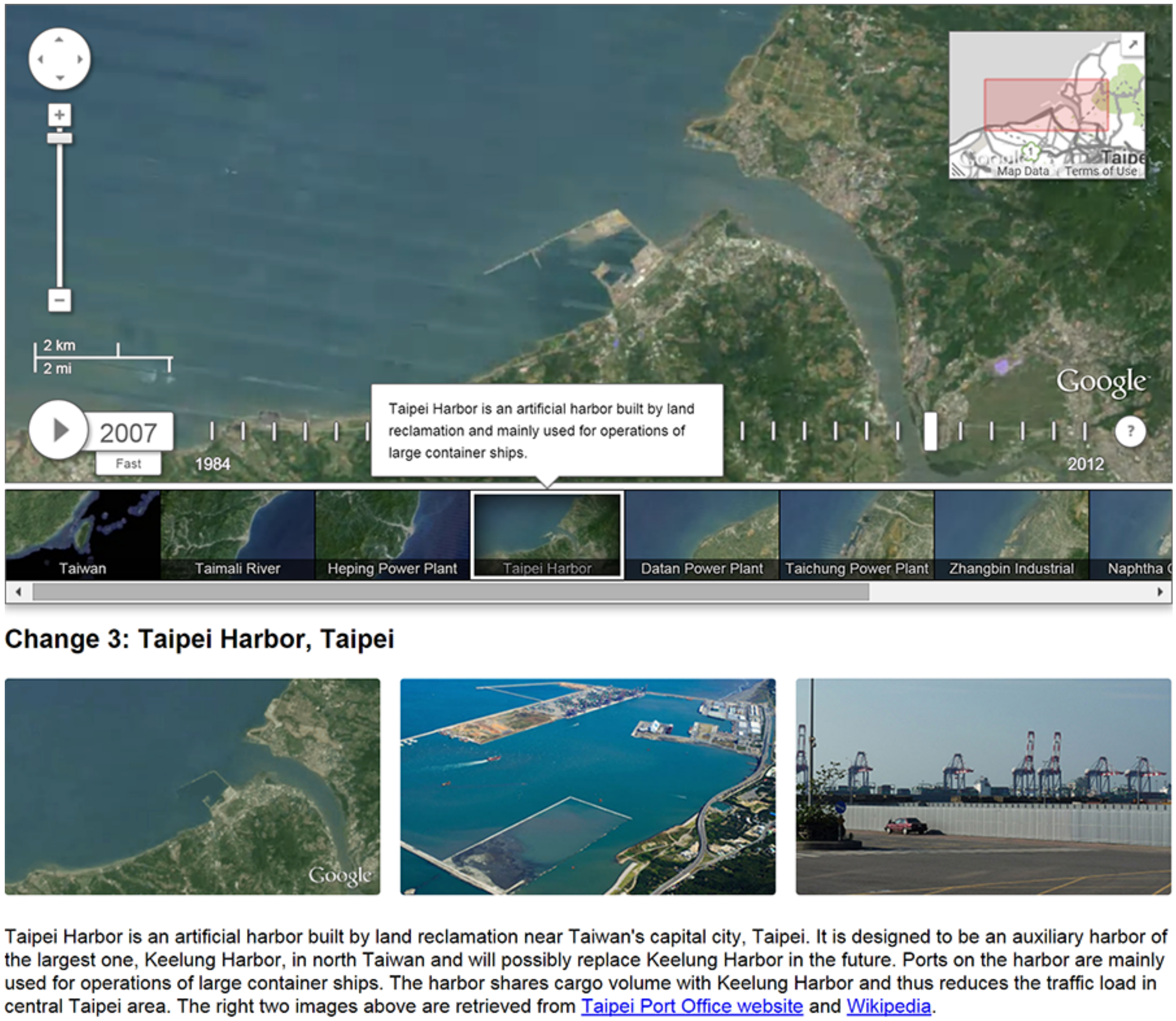


Figure 3.3: This figure shows the user interface of an interactive slideshow with other media.

- What is the efficiency of the editor? Can professionals use it easily without spending too much effort?
- Do stories integrating custom guided tours and interactive slideshows encourage audiences to explore the timelapse?
- Do custom guided tours and interactive slideshows help audiences gain insights from stories formed by professionals?

Developing tools to support the creation of visual elements for storytelling depends heavily on the user needs from professionals. It is vital to collaborate with target users and keep them in the design loop. Future works include conducting a medium to long term user evaluation [202] to answer the research questions and analyzing the locations that users focused on by parsing server log files (i.e. requests of images). The ultimate goal is that the editor can truly help professionals focus more on the content of stories rather than time-consuming and laborious work.

Chapter 4

Community-Empowered Air Quality Monitoring System

Developing information technology to democratize scientific knowledge and support citizen empowerment is a challenging task. In our case, a local community suffered from air pollution caused by industrial activity. The residents lacked the technological fluency to gather and curate diverse scientific data to advocate for regulatory change. We collaborated with the community in developing an air quality monitoring system which integrated heterogeneous data over a large spatial and temporal scale. The system afforded strong scientific evidence by using animated smoke images, air quality data, crowdsourced smell reports, and wind data. In our evaluation, we report patterns of sharing smoke images among stakeholders. Our survey study shows that the scientific knowledge provided by the system encourages agonistic discussions with regulators, empowers the community to support policy making, and rebalances the power relationship between stakeholders. The system critically reveals the local air pollution problem of a community and empowers citizens to advocate for themselves. This work addresses the data quality, science communication, and evaluation metrics challenges (as mentioned in section 1.3) for initializing, maintaining, and evaluating community engagement. The contributions of this work include the methodology of curating and visualizing multiple types of scientific data, the concept of using computer vision to support forming scientific knowledge, and the result of tracking behavior and attitude changes.

4.1 Preface

Air pollution is a critical environmental issue for people who live near industrial sites. To address this problem, it takes communities a great effort to gather scientific evidence at a large spatial and temporal scale, which requires the assistance of information technology in collecting, curating, and visualizing various types of data. In our case, 70,000 residents near Pittsburgh suffer from air pollution caused by a coke (fuel) plant. Under some unusual situations, the coke plant leaks hazardous smoke irregularly, known as fugitive emissions (see Figure 4.1), into the atmosphere. The resulting toxic emissions with fine particulates pose risks to health and have negative impacts to living quality [122, 177].

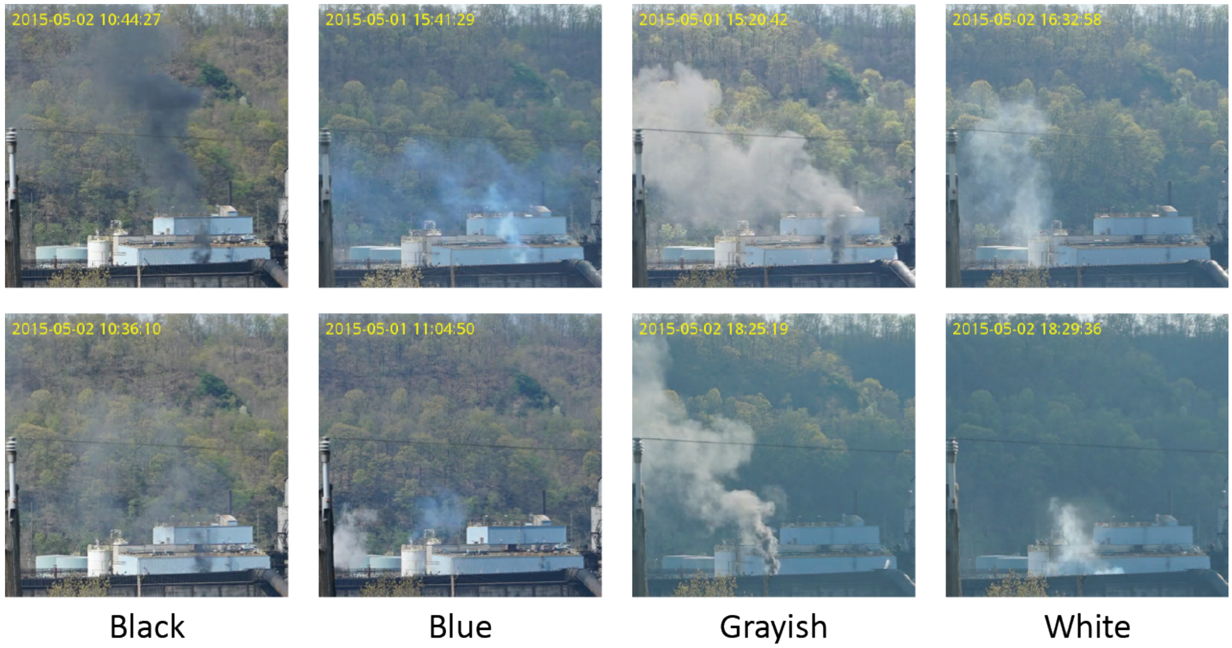


Figure 4.1: This figure shows emissions with various lightings, appearance, and opacities.

To address air pollution, residents formed the ACCAN (Allegheny County Clean Air Now) group. In several community meetings, residents mentioned that adults and children developed respiratory problems because of exposure to coke oven gas. In addition, residents must close windows at night because of irritating burning smells. They also said that the air quality was so poor that they could not exercise outside. To pursue environmental justice, the community took a series of actions, such as gathering evidence of violations and filing petitions to the government. They envisioned that these actions could raise public awareness about air quality issues and pressure the government to deal with air pollution problems.

To advocate for themselves in improving the local air quality, the community needed to gather convincing evidence in communicating with stakeholders. Traditionally, the community collected scientific data manually, which was time-consuming, error-prone, and offered limited scientific validity. The community lacked technological fluency and required the assistance of experts in setting up an automatic system to collect and archive data from various sources. Starting in January 2015, we aided the community to set up outdoor air quality sensors and live cameras pointed at the coke oven where smoke usually occurred. We also created an electronic process for capturing smell reports. To visualize hybrid data (sensor readings, smell reports, real-time high resolution imagery, and wind information), we developed a web-based air quality monitoring system. Community members could use the system to manually search for smoke in timelapse videos and use a thumbnail generator to create animated images. But searching and documenting all smoke emissions required manpower and took an impractical investment of time. Therefore, we implemented a computer vision tool to detect smoke and produce corresponding animated images (see Figure 4.5), which could then be curated in online documents and shared on social media. With the monitoring system, community members could tell stories with concrete scientific evidence about what happened (using animated smoke images) and how these events

Figure 4.2: This figure shows steam, shadow, and the mixture of steam and smoke.

affected the local neighborhood (using sensor readings, smell reports, and wind information).

To evaluate community engagement, we analyzed the server logs, which store HTTP requests of thumbnails from August 2015 to July 2016. In addition, we conducted a survey study with the research question: does interacting with the air quality monitoring system increase community engagement in addressing air pollution concerns? We anticipated that the intervention of the system increases awareness, self-efficacy [15, 41], and sense of community [160], which are the three dependent variables in our survey study. Awareness means participants know a problem exists and has impact on daily lives. Self-efficacy means the strength of participants' belief in their ability to successfully reach the community's goal. Sense of community means participants feel they have influence in the community and a sense of belonging. We form three corresponding hypotheses: interacting with the system improves the ability to perceive air quality problems, strengthens the belief that the ACCAN community can reach its goal of improving air quality, and makes people think that they are influential and fit in the community. The independent variables are involvement, age range, and education level. Involvement is the level of participation, such as exploring, documenting, and sharing data from the system.

In this chapter, we explore the formation and use of scientific knowledge in citizen empowerment via the intervention of information technology. Our design principle is to stimulate critical discussions and confront the current unbalanced power relation between stakeholders. We begin by explaining the research scope and reviewing similar projects. Then, we describe the design process and the implemented web-based air quality monitoring system. In addition, we discuss the results of smoke image usage from server logs and survey study. Finally, we provide insights in developing systems to empower data-driven community action and conclude with limitations. Our contributions are:

- Detailed documentation of a worked example which used scientific data from heteroge-

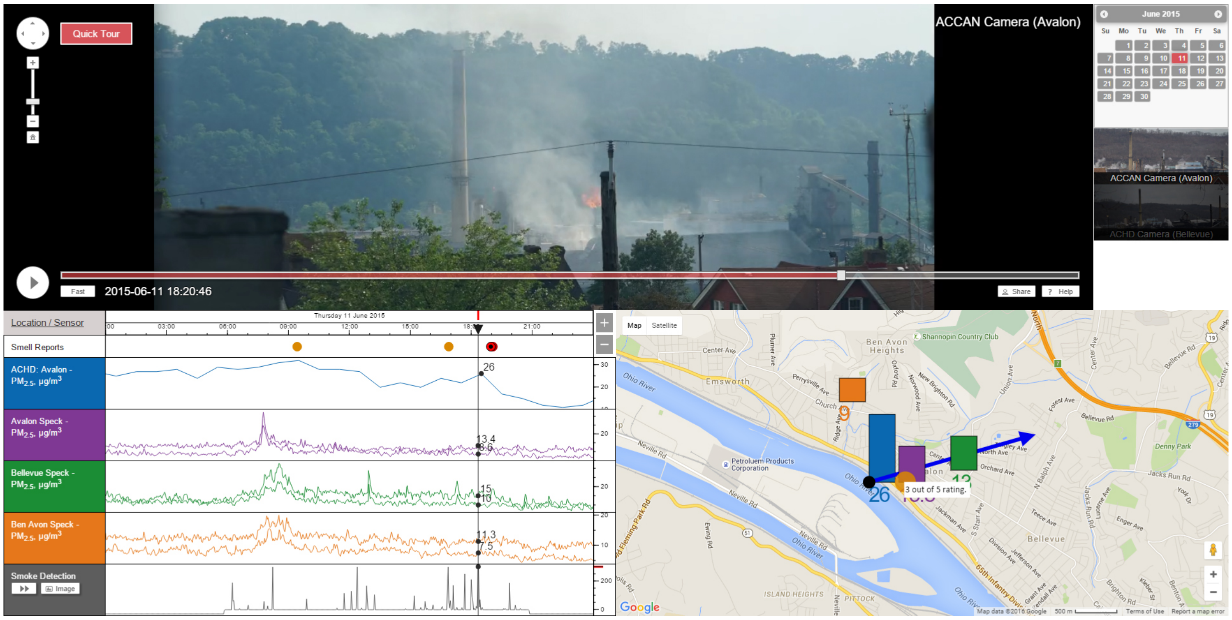


Figure 4.3: The user interface of the web-based air quality monitoring system. The top-left part is a zoomable and pannable viewer which shows the timelapse video. The bottom-left charts visualize crowdsourced smell reports, PM2.5 sensor readings, and automatic smoke detection results. The blue line shows readings from the sensor operated by the local health department. The purple, green, and orange lines shows readings from six sensors that we deployed in the community. The bottom-right map indicates wind speed (length of the blue arrow), wind direction (orientation of the blue arrow), and sensor locations (bar charts). The colors and heights of bar charts on the map correspond to the colors and readings on the line charts respectively.

neous sources to critically reveal, question, and challenge environmental conditions.

- Analysis of community behavior changes after the intervention of information technology and participatory design.
- Analysis of how the community uses smoke images over a long-term participation period (12 months).
- Insights for researchers to develop environmental monitoring systems that combine politics, community, and information technology.

4.2 Design Process and Challenges

We began by participating in monthly community meetings to understand the context of air pollution issues. The community was taking a series of actions, such as reporting industrial smells and filing petitions to the local health department and the EPA (Environmental Protection Agency). Our roles were as supporters, which use information technology to assist the citizen-led grassroots movement around local air quality issues, and as researchers, which study the effect of the technological intervention.

The successfulness of the intervention of information technology is highly dependent on community engagement [208], the involvement of citizens in local neighborhoods. During initial discussions with the community, we found that the most significant gap in community engagement is the lack of scientific evidence. For instance, it was difficult for residents to report the exact time when an air quality violation occurred and its environmental impact to government regulators. Therefore, we proposed building an air quality monitoring system, which could afford exploring, archiving, presenting, and sharing scientific evidence among stakeholders.

The problem that the community dealt with is in nature wicked [50, 190]. One characteristic of a wicked problem is that it cannot be fully observed, which means that solving a subset of a wicked problem reveals new ones. Based on this idea, we argue that our work requires an iterative design approach to handle and solve design challenges step by step. Thus, we adopt the community-based participatory design approach [72]. It is iterative in the sense that citizens and developers explore design options collaboratively.

We collaborated closely with the community and implemented system features based on iterative feedback from community members. There were two major design challenges in setting up the monitoring system. First, the community did not have sufficient technological fluency. Our system had to curate and visualize data in a way that users could easily perceive and document the seriousness of smoke emissions and their impacts to local neighborhoods. Second, this work had a timing issue, where residents had to form and use strong scientific evidence to convince regulators on a planned community meeting with the local health department and the EPA. These challenges served as constraints that affected our design decisions.

4.3 System

We now explain system components together with three design iterations, which naturally emerged during the design process. The number of iterations depends on the complexity of the wicked problem [50, 190] that the community tackles. Each iteration contained system features which were implemented based on the challenges revealed iteratively.

4.3.1 First Iteration:

Interactive Web-based Timelapse Viewer

Starting in January 2015, we installed a live camera which was oriented towards the coke plant from a volunteer home. The live camera takes a high quality image every 5 seconds for a total of 17,000 each day. We streamed the time-series imagery to our servers and used an open source tool developed by Sargent et al. [193] to process these images into multi-resolution video tiles. The tool was implemented in JavaScript/HTML and provided an **interactive web-based timelapse viewer** (top-left part of Figure 4.3) where users could search for fugitive emissions by panning, zooming, and playing the video. The viewer loaded and showed the video tile corresponding to the zoom and pan level. Users could share a particular view online with other people. After we developed the web-based viewer, community members were excited and shared screenshots with each other via emails. At that time, the community pointed out two major challenges. Static images such as screenshots could not represent the dynamics and persistent time quality of smoke emissions. In addition, although smoke images indicated the source of air pollution, they

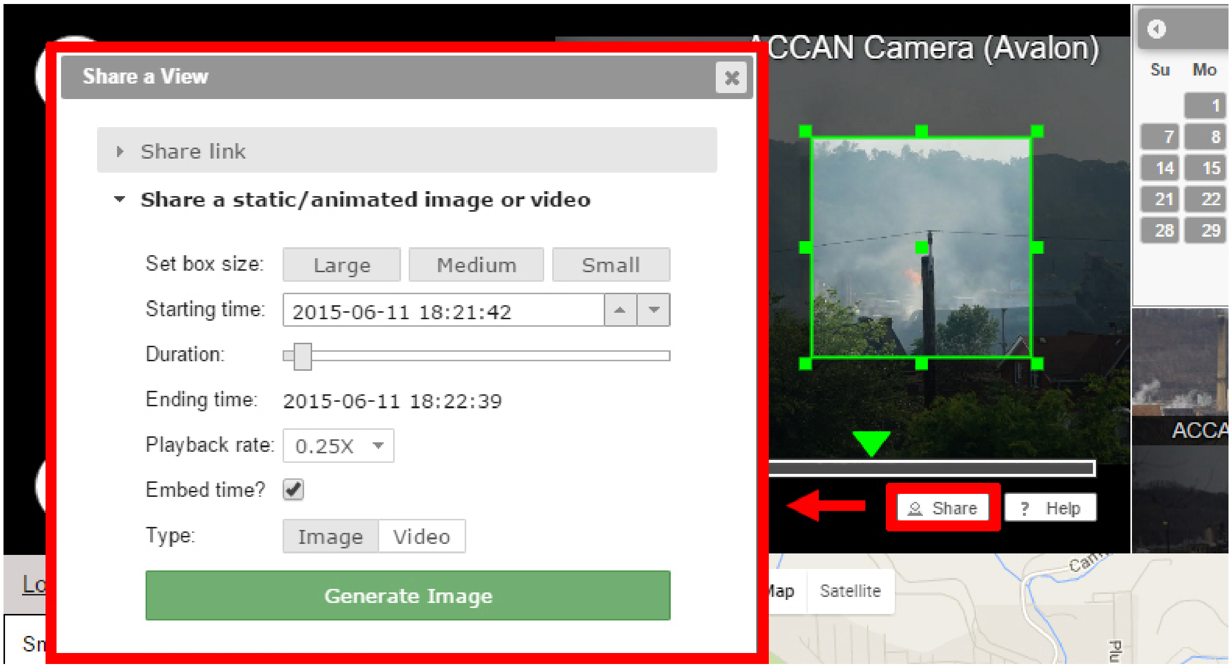


Figure 4.4: Clicking the share button on the timelapse viewer on the main user interface (see Figure 4.3) shows the thumbnail tool, which is used for generating sharable animated images. Users can edit the image size by resizing the green box on the viewer. The dialog window provides adjustable parameters, such as starting time and duration of the animated image.

did not show the impacts to local air quality. These challenges led to the next design iteration.

4.3.2 Second Iteration:

Thumbnail Generator and Sensor Data Visualization

To address the emergent challenges, we implemented a **thumbnail generator**, which allowed community members to create, document, and share animated smoke images as visual evidence (Figure 4.4). We also visualized $\mathrm{PM}_{2.5}$ (particle pollution) data from a sensor station operated by the local health department. In addition, we visualized smell reports which were collected via a Google Form, only available to community members. In the form, we asked community members to rate the severity of the pollution odors from 1 to 5, with 5 being the worst. The form was disseminated to the community via a Google Groups email and phone calls. The **visualization of air quality data and smell reports** showed how smoke emissions affected the living quality of the community. With these new features, residents could compare smoke images together with sensor and crowdsourced data to identify correlations. We recorded a tutorial video and taught residents how to use these features during community meetings. The community was using the tool to find, generate, and share animated smoke images. However, searching smoke emissions manually from a large amount of time-series imagery was laborious and time-consuming. Moreover, the government-operated sensor station reported data only once per hour, which had difficulties in identifying air quality changes over a shorter time period. Furthermore, the lack of visualized wind data and sensor locations hindered the ability to determine how

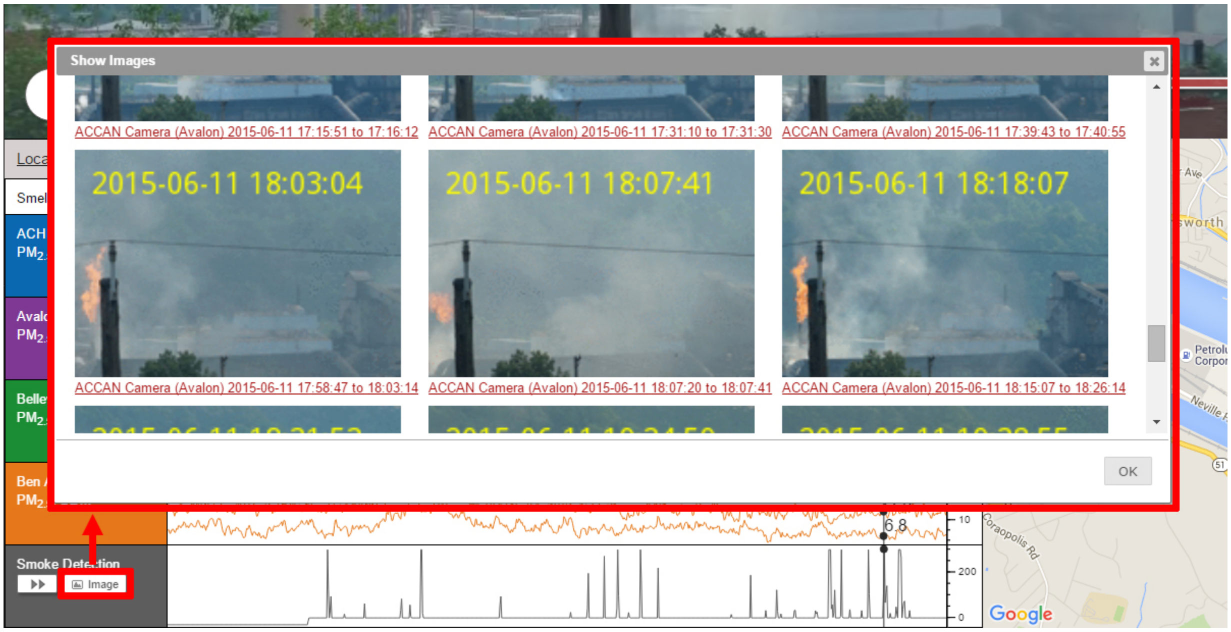


Figure 4.5: Clicking the image button on the line charts on the main user interface (see Figure 4.3) shows web links and animated images produced by the smoke detection algorithm. Users can quickly select representative images and insert them into an online document. Users can also click on a peak of a spike on the line chart to seek to a video frame with fugitive emissions.

pollutants affected the air quality hyperlocally. These challenges again led to another design iteration.

4.3.3 Third Iteration:

Citizen Sensors, Computer Vision Tool, and Map Visualization

To account for the challenges from the previous iteration, we deployed **six commercial air quality sensors** [206, 217] in local areas with finer time resolutions. These sensors reported $\mathrm{PM}_{2.5}$ data to our server via wireless Internet once per minute. The location of sensors and the Internet services were provided by community volunteers. Furthermore, we developed a **computer vision tool** based on a smoke detection algorithm (see the next section) for finding fugitive emissions automatically. The algorithm identified the number of smoke pixels for each video frame at daytime (bottom chart in Figure 4.5) and automatically produced corresponding sharable animated images (see Figure 4.5). We also added a **map visualization** for showing wind direction, wind strength, and sensor locations (bottom-right part of Figure 4.3). All sensor data and smoke detection results were plotted on multiple charts (bottom-left part of Figure 4.3). Users could use the charts as indicators for finding unusual events such as fugitive emissions. Clicking on a smell report or a peak of a spike on the chart jumped the video to the corresponding time. Users could also click on the image button near the smoke detection chart to bring up a dialog box with animated smoke images, which could be shared via social media or archived into a Google Doc.

The final design enabled community members to fully explore and compare data from heterogeneous sources (animated smoke images, finer air quality data, crowdsourced smell reports, and wind information). When residents noticed industrial smells like sulfur, they could use the timelapse viewer to check if the coke plant emitted smoke at a specific time. They could then compare sensor readings, smell reports, and wind data to verify if the emission came from the coke plant and affected the local air quality. With the system, the community could form and share convincing narratives grounded with scientific evidence aggregated from hybrid data.

4.4 Smoke Detection

There are three general approaches appearing in previous research for detecting the presence of smoke emissions in a single image or across multiple frames: (1) color modeling; (2) change detection; and (3) texture analysis.

Color modeling describes the characteristics of image intensity values. For instance, smoke is grayish and has low saturation. Previous research used color models to identify smoke pixels [43] or extract color histogram features [148].

Change detection [185] determines moving objects in an image, which provides candidate regions containing smoke emissions for further analysis. One common technique is background subtraction [45, 49] which estimates an image without moving objects from an image sequence, subtracts the estimated image from the current one to get a residual image, and thresholds the residual image to obtain a binary mask. In addition, there are background modeling approaches [88, 207] which learn a probabilistic model of each pixel using a mixture-of-Gaussians and determine the background pixels according to the probability distribution. Other techniques involve computing the entropy of the optical flow field [134] to identify smoke and checking flickering pixels at the edge of candidate smoke regions [222].

Texture analysis measures texture energy in a single image or texture changes between multiple frames. One approach is to apply texture descriptors, such as a wavelet transform, on small blocks in an image for obtaining feature vectors and train a classifier using these features [40, 96].

Each of these approaches has distinct strengths and weaknesses. Color modeling is straightforward, but suffers from situations where smoke and non-smoke objects have the same chrominance (e.g. white smoke and steam, dark shadow and black smoke) or the background does not contain plentiful color information due to various weather and lighting conditions (e.g. fog, nighttime images). Background subtraction and background modeling do not distinguish smoke from non-smoke regions since they find all moving objects including shadow, steam, and smoke. Optical flow can determine smoke motions, but has high computational cost. It is difficult to extract useful information from texture analysis if the background does not contain sufficient texture information. Several research has integrated these methods into a system for better performance. Toreyin et al. [222] combined background subtraction, edge flickering, and texture analysis into a final result. Lee et al. [148] used change detection to extract candidate regions, computed feature vectors based on color modeling and texture analysis, and trained a support vector machine classifier using these features.

We are aware of other advanced machine learning approaches. For instance, Hohberg [109] trains a convolutional neural network for recognizing wildfire smoke. Tian et al. [219] present a physical based model and use sparse coding to extract reliable features for single image smoke detection. However, a simpler heuristic approach combining color modeling, change detection, and texture analysis is sufficient for our current needs.

Inspired by prior method integration approaches, we have implemented a smoke detection algorithm for detecting fugitive emissions during the daytime from a static camera. The algorithm contains five steps: preprocessing, change detection, texture segmentation, region filtering, and event detection. Change detection identifies moving pixels containing smoke, steam, and shadow. Texture segmentation clusters pixels into several candidate regions based on texture information. Region filtering iteratively evaluates each candidate region based on shape, color, size, and the amount of change to determine if it matches the appearance and behavior of smoke. Event Detection groups video frames with smoke together to identify the starting and ending time of fugitive emissions.

4.4.1 Preprocessing

We apply the algorithm on 9700 daytime frames for each day and ignore nighttime. To reduce the computational cost, we first scale the original image at time t down to one-fourth of the original size to obtain a downsampled image I_t. Then we estimate the background image B_t by taking the median over the previous 60 images as shown in (4.1).

$$B_t(x, y) = \text{median}\big(I_t(x, y), ..., I_{t-59}(x, y)\big) \tag{4.1}$$

where (x, y) indicates the position of a pixel. Finally we convert all RGB images with 8-bit unsigned integer format to double precision ranging from 0 to 1.

4.4.2 Change Detection

Change detection finds moving pixels in video frames by computing changes in high frequency signals (e.g. edges, textures) and image intensity values (e.g. colors).

High Frequency Change Detection

Smoke is semi-transparent with various opacities and occludes parts of the background upon presence, which causes changes of high frequency signals across frames. First we compute the difference of Gaussian (DoG) of I_t and B_t to obtain I_{dog} and B_{dog} as shown in (4.3)

$$G_\sigma(x, y) = \frac{1}{2\pi\sigma^2}\exp\Big(-\frac{x^2 + y^2}{2\sigma^2}\Big) \tag{4.2}$$

$$\begin{aligned} I_{dog}(x, y) &= \big(G_{\sigma_1}(x, y) - G_{\sigma_2}(x, y)\big) * I_t(x, y) \\ B_{dog}(x, y) &= \big(G_{\sigma_1}(x, y) - G_{\sigma_2}(x, y)\big) * B_t(x, y) \end{aligned} \tag{4.3}$$

where the asterisk sign $*$ indicates the convolution operator and $G_\sigma(x, y)$ is a Gaussian kernel with variance σ^2 and mean zero. The DoG image contains high frequency information for the current and the background images.

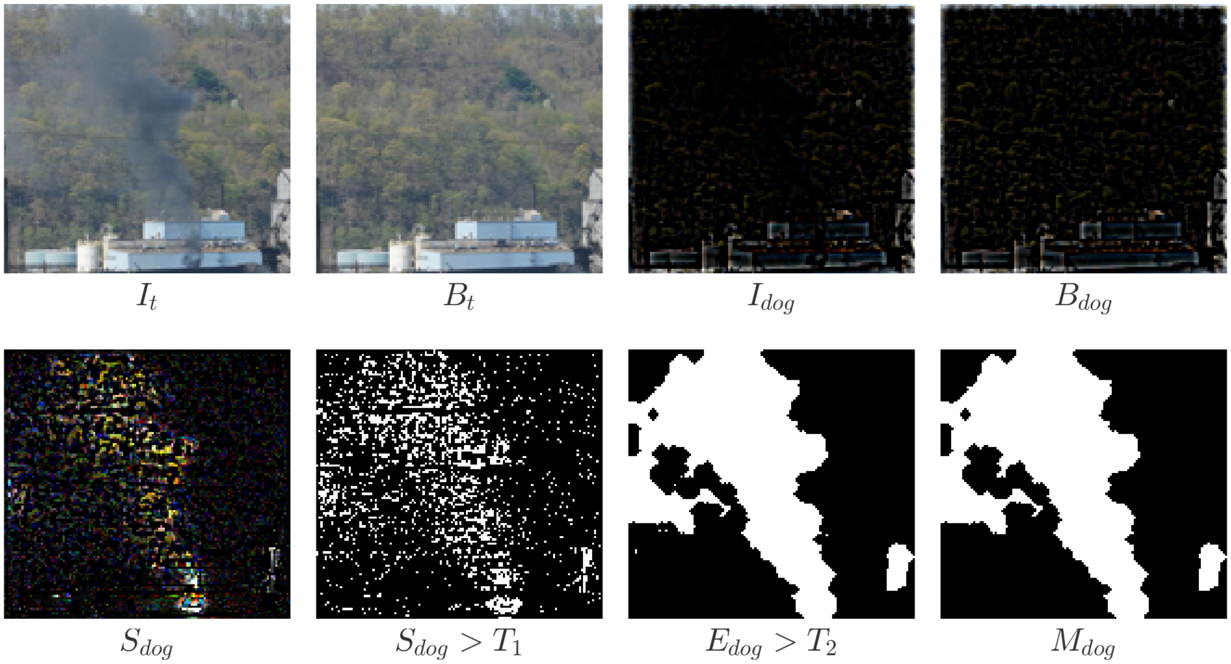


Figure 4.6: This figure visualizes the steps of high frequency change detection. Refer to section 4.4.2 for detailed explanation.

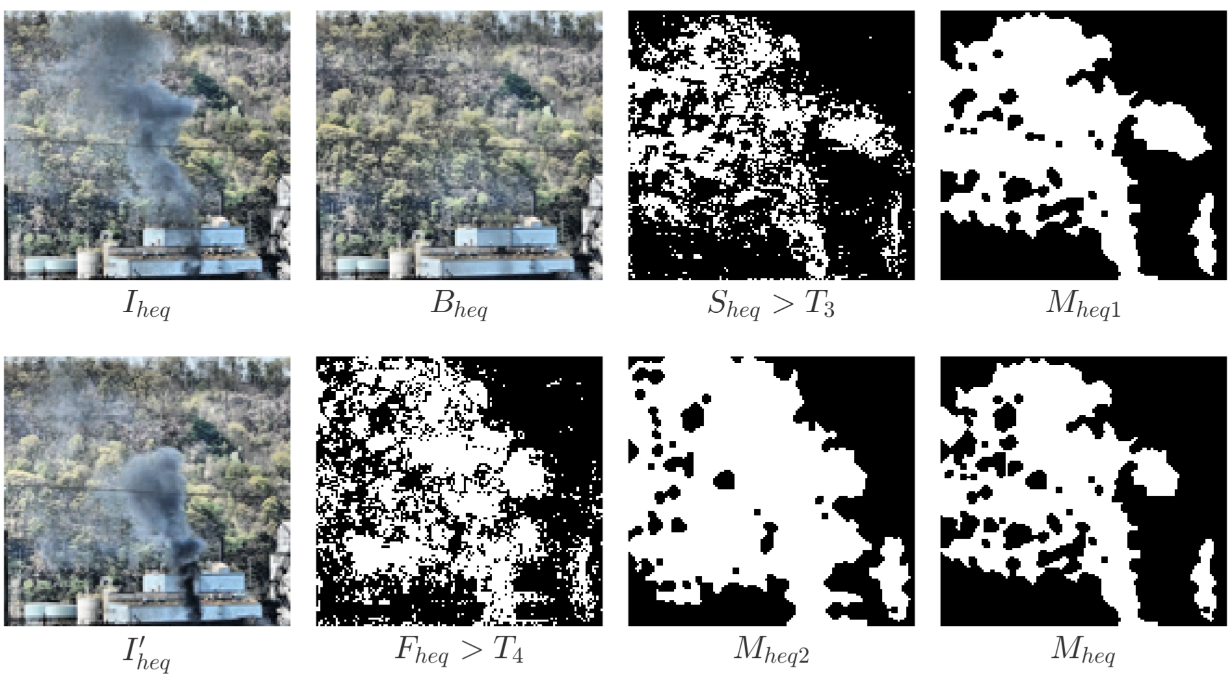


Figure 4.7: This figure visualizes the steps of image intensity change detection. Refer to section 4.4.2 for detailed explanation.

Then we perform background subtraction on I_{dog} and B_{dog} to obtain $S_{dog} = \text{bgSub}(I_{dog}, B_{dog})$ as shown in (4.4).

$$\text{bgSub}(I, B) = \frac{|I - B|}{\max\big(I + B, 0.1\big)} \tag{4.4}$$

Dividing the background subtraction term in the nominator by $\max\big(I + B, 0.1\big)$ alleviates the effect of illumination in images. The $\max$ function in the denominator in (4.4) prevents dividing to an extremely small value or zero. One way to interpret the S_{dog} image is that it measures the change of high frequency signals such as edges and texture between the current and background image. Thresholding channels in S_{dog} yields a binary image. Computing the local entropy of the 9-by-9 neighborhood centered around each pixel in the binary image gives an entropy image E_{dog} as show in (4.5).

$$E_{dog} = \text{entropyFilter}\big(\text{bgSub}(I_{dog}, B_{dog}) > T_1\big) \tag{4.5}$$

Finally we threshold the entropy image E_{dog} to obtain a binary image $E_{dog} > T_2$. Performing morphological closing, removing noise using a median filter, and discarding small regions using connected component algorithm on the binary image yields the smoothed image M_{dog} as shown in (4.6).

$$M_{dog} = \text{smooth}(E_{dog} > T_2) \tag{4.6}$$

Figure 4.6 visualizes the steps of high frequency change detection. If the M_{dog} image contains no regions (i.e. all pixel values are zero), the smoke detection algorithm terminates at this step and outputs zero as the response.

Image Intensity Change Detection

Changes of pixel intensity values across frames indicate candidate regions containing smoke. We first enhance the contrast of image I_t, I_{t-2}, and B_t by using CLAHE (contrast-limited adaptive histogram equalization [231]) to obtain I_{heq}, I'_{heq}, and B_{heq}. CLAHE limits the contrast to avoid over-amplifying noise and operates on small local regions in the image. The desired shape of the histogram in a local region is approximately flat and follows a uniform distribution. One reason for performing contrast enhancement is that the color and saturation of smoke may be similar to the background under some lighting conditions.

Next we perform background subtraction as shown in (4.4) on each channel of the two image pairs (I_{heq}, B_{heq}) and (I_{heq}, I'_{heq}) to obtain $S_{heq} = \text{bgSub}(I_{heq}, B_{heq})$ and $F_{heq} = \text{bgSub}(I_{heq}, I'_{heq})$, which provides information about the change of image intensity values between the current frame, background, and the previous frame. Smoothing the binary images $S_{heq} > T_3$ and $F_{heq} > T_4$ by using the process described in section 4.4.2 yields M_{heq1} and M_{heq2}.

Finally we combine M_{heq1} and M_{heq2} by using an AND operator into the resulting image M_{heq} as shown in (4.7). Figure 4.7 visualizes the steps of image intensity change detection.

$$\begin{aligned} M_{heq1} &= \text{smooth}\big(\text{bgSub}(I_{heq}, B_{heq}) > T_3\big) \\ M_{heq2} &= \text{smooth}\big(\text{bgSub}(I_{heq}, I'_{heq}) > T_4\big) \\ M_{heq} &= M_{heq1} \text{ and } M_{heq2} \end{aligned} \tag{4.7}$$

4.4.3 Texture Segmentation

Texture segmentation partitions images into regions based on their texture information. This step computes filter responses by convolving an image with a filter bank, clusters the responses into a set of textons [157], and partitions the image into separate regions by using these textons. We first combine the results of change detection algorithms by performing an AND operation on M_{dog} and M_{heq} to obtain M_{cd} as shown in Figure 4.8. If all pixel values in image M_{cd} are zero, the smoke detection algorithm stops at this step and outputs zero as the response.

Next we compute the filter bank using a variation of Laws' texture energy measures [145] as shown in (4.8).

$$\begin{array}{lcl} L5 & = & [\ 1 \ \ 2 \ \ 3 \ \ 2 \ \ 1\ ] \ \text{(Level)} \\ E5 & = & [\ \text{-1} \ \ \text{-2} \ \ 0 \ \ 2 \ \ 1\ ] \ \text{(Edge)} \\ S5 & = & [\ \text{-1} \ \ 0 \ \ 2 \ \ 0 \ \ \text{-1}\ ] \ \text{(Spot)} \\ W5 & = & [\ \text{-1} \ \ 2 \ \ 0 \ \ \text{-2} \ \ 1\ ] \ \text{(Wave)} \\ R5 & = & [\ 1 \ \ \text{-4} \ \ 6 \ \ \text{-4} \ \ 1\ ] \ \text{(Ripple)} \end{array} \tag{4.8}$$

The filter bank is a set of 5-by-5 convolution masks obtained by calculating the outer products of pairs of texture vectors in (4.8). The L5, E5, S5, W5, and R5 vectors detects gray level, edges, spots, waves, and ripples in the image respectively.

Then we take the contrast-enhanced image I_{heq}, subtract it with the mean value of I_{heq}, and convolve it with the filter bank for each RGB channel to obtain feature vectors. Each vector represents the corresponding pixel in I_{heq} in the feature space and has 125 dimensions. Then the algorithm performs Principal Component Analysis which preserves 98% of the energy (eigenvalues) on the feature vectors to reduce dimensions. Using the contrast-enhanced image alleviates the problem that some weather circumstances such as fog cause a decrease in background texture information.

Finally we perform an accelerated k-means++ algorithm [13, 81] which chooses better initialized values (seed points) to cluster the feature vectors into textons and divide the current image into various regions as shown in image R_t in Figure 4.8. Smoothing the image R_t by discarding small regions, removing noise by using a median filter, and performing morphological closing yields R_{smooth} in Figure 4.8.

4.4.4 Region Filtering

Region filtering determines if a region matches the appearance and behavior of smoke by evaluating shape, color, size, and the amount of change. We first use the connected component algorithm to find all separated regions and remove the ones which are thin and narrow. Mathematically speaking, for each region, the ratio of width to height of its bounding box exceeds a certain threshold. Or the ratio of the size of the region and its bounding box is smaller than a threshold.

Next we adjust the contrast of each channel in I_t to produce I_{adj} in Figure 4.8 by stretching intensity values so that 1% of the data is saturated at low and high intensities of I_t. We group nearby white regions and black ones based on I_{adj} to reconstruct the shapes of objects. Since the color of smoke is usually grayish or bluish, we can remove regions having non-grayish and

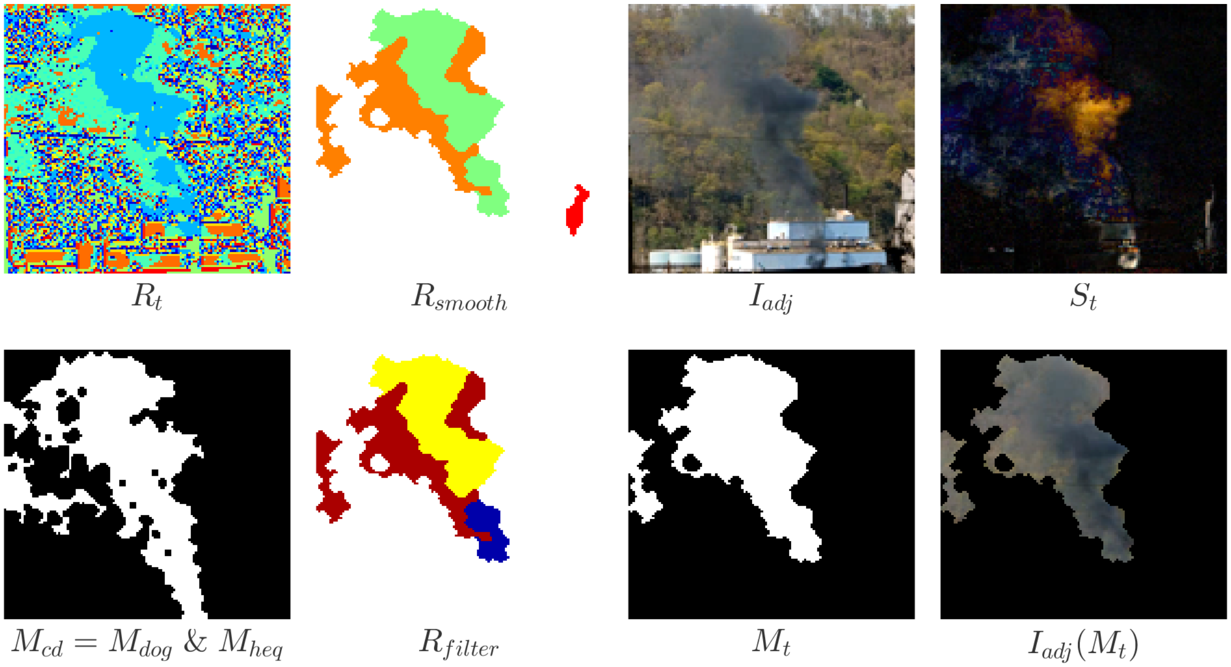


Figure 4.8: This figure demonstrates the steps of texture segmentation and region filtering. See section 4.4.3 and 4.4.4 for detailed explanation.

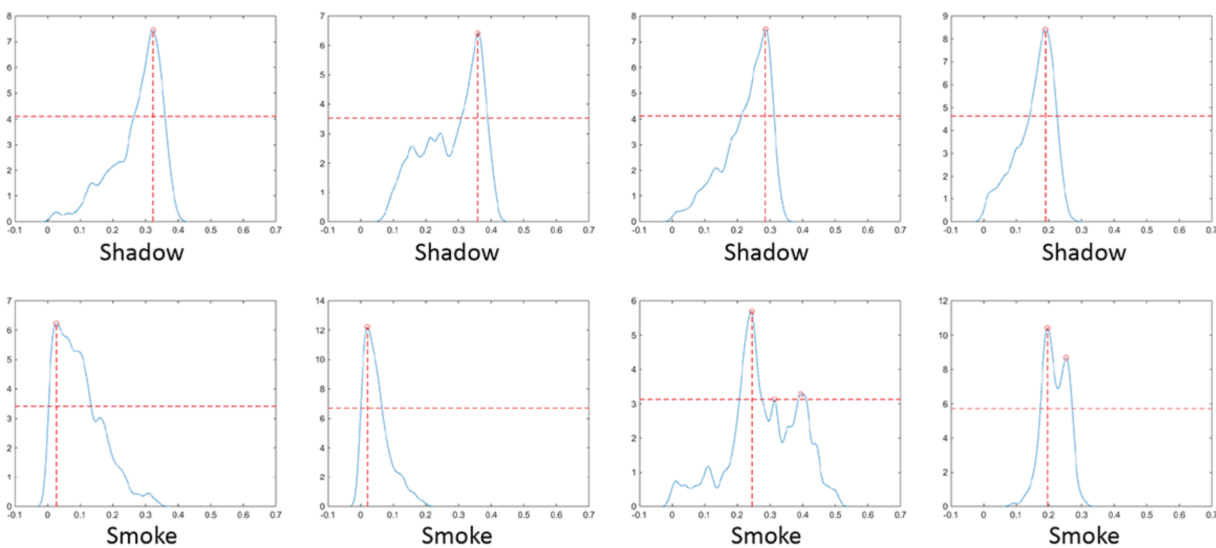


Figure 4.9: Each small graph shows the probability density function of a smoke or shadow region's corresponding pixel values in S_t (see Figure 4.8) using kernel density estimation. The x-axis represents the pixel values in S_t. The horizontal red line is the threshold for computing number of peaks. The vertical red line indicates the pixel value of the highest peak.

non-bluish colors described by (4.9)

$$
\begin{aligned}
&|c_1 - c_2| \geq t_1 \;\text{ or }\; |c_2 - c_3| \geq t_2 \;\text{ or }\; |c_1 - c_3| \geq t_3 \\
&c_j = \text{median}\big(I_{adj}(x, y, j)\big) \;\; \forall (x, y) \in R_i
\end{aligned}
\tag{4.9}
$$

where j indicates different channels in I_{adj}, R_i denotes the i^{th} region, (x, y) means the location of pixels, and $\{c_j : j = 1, 2, 3\}$ are the median of corresponding pixel values in R_i in the RGB channels of I_{adj}. We also remove regions having light colors described by (4.10) because steam is usually white.

$$
c_1 \geq t_4 \;\text{ and }\; c_2 \geq t_5 \;\text{ and }\; c_3 \geq t_6 \tag{4.10}
$$

Then we compute the size of each region and remove large or small ones which may be noise and shadow respectively. Furthermore, we remove the i^{th} region R_i if it does not have sufficient amount of change by summing up the corresponding pixel values in M_{cd} by using (4.11)

$$
\sum_{\forall (x,y) \in R_i} M_{cd}(x, y) \leq t_7 \tag{4.11}
$$

where (x, y) denotes the location of pixels in region R_i.

Finally we remove regions which may contain shadow. The algorithm performs background subtraction using (4.4) on I_t and B_t to obtain $S_t = \text{bgSub}(I_t, B_t)$ in Figure 4.8. Then we compute the probability density function (PDF) of each region's corresponding pixel values in S_t using kernel density estimation [203] with a Gaussian kernel.

$$
\hat{p}(x) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h} K\left(\frac{x - X_i}{h}\right) \;\text{ where }\; X_i \in S_t \tag{4.12}
$$

Because the PDF of shadow and smoke regions have distinct characteristics (see Figure 4.9), we can describe shadow regions by utilizing (4.13)

$$
\underset{x}{\text{argmax}}\; p(x) > t_8 \;\text{ and }\; \sum_{x_i \in X} \mathbf{1}_{\{p(x_i) > t_9\}} < t_{10} \tag{4.13}
$$

where x indicates pixel values, $p(x)$ is the probability density function, $\text{argmax}_x\; p(x)$ means the pixel value of the highest peak, X is a set of pixel values of the corresponding peaks, $\mathbf{1}_A$ is the indicator function of a set A, and $\sum_{x_i \in X} \mathbf{1}_{\{p(x_i) > t_9\}}$ is the number of peaks having their heights exceed a certain threshold.

Applying all the above region filtering steps on R_{smooth} yields R_{filter} (see Figure 4.8). We compute a mask M_t which is a binary image based on R_{filter} and output the response at time t as the sum of all pixel values in mask M_t.

4.4.5 Event Detection

Event Detection identifies the starting and ending time of fugitive emissions. We first select daytime frames for each day and ignore nighttime ones because of lighting issues. Next, we apply

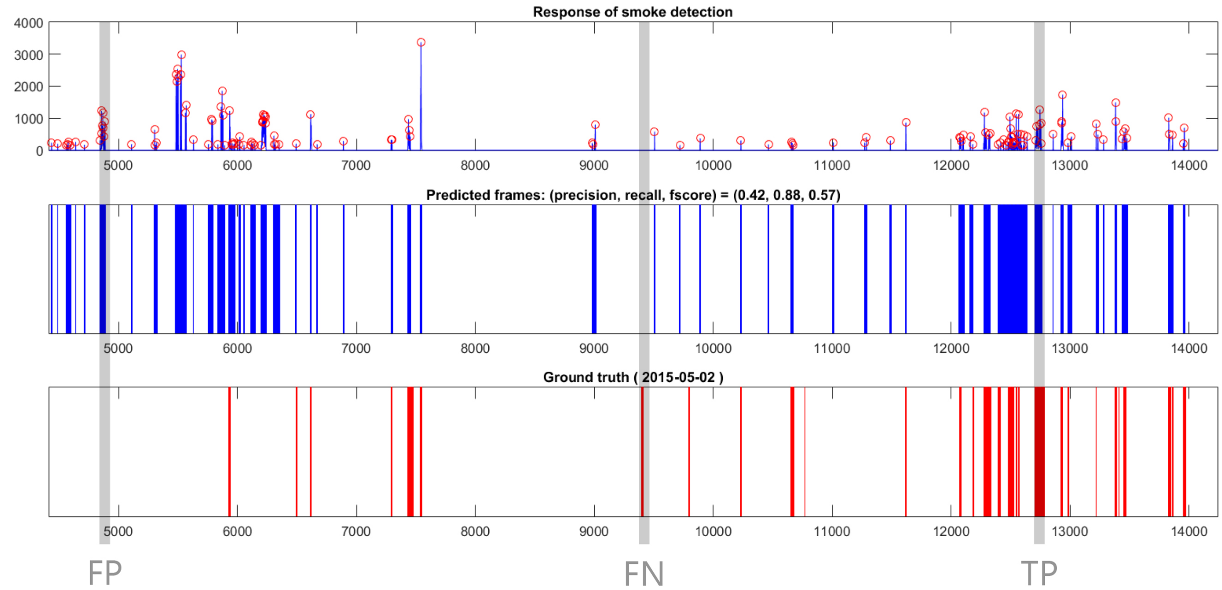


Figure 4.10: This figure shows the result of smoke detection. The x-axis and y-axis indicate the frame number and the amount of pixels identified as smoke. The bottom graph is the ground truth of May 2, 2015. The top and middle graphs show the response and the prediction of all daytime frames. The red circles in the top graph represent the local peaks. The gray bars indicate true positive (TP), false positive (FP), and false negative (FN).

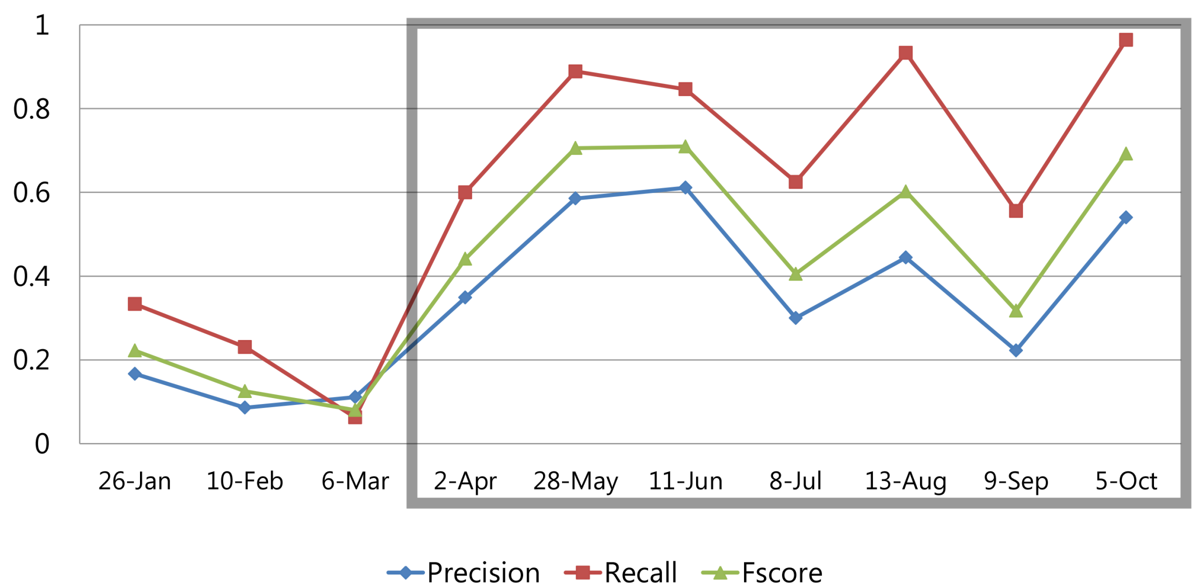


Figure 4.11: Evaluation of the smoke detection algorithm on 12 randomly chosen days for each month in 2015.

change detection, texture segmentation, and region filtering on these frames to obtain a time-series signal (see the top chart in Figure 4.10). Each value in the time-series signal represents the number of smoke pixels in a corresponding video frame. Then we compute segments in the time-series signal by finding peaks and corresponding peak widths. Finally we merge nearby segments into events (see the middle chart in Figure 4.10).

4.4.6 Experiment

We used MATLAB to develop the smoke detection algorithm and VLFeat [223] library to run the accelerated k-means++ algorithm for clustering feature vectors during the texture segmentation step. Each timelapse of a day consisted of 16838 frames. We ran the smoke detection algorithm on a window with 496-by-528 pixels in the timelapse video for 21 days in 2015 during daytime. The processing time was 30 minutes on average for a day by using all cores on a workstation with two hex-core CPU (Intel Xeon X5670).

We manually labeled these 21 days to evaluate the performance of the algorithm. The bottom graph of Figure 4.10 shows the ground truth labels on May 2. The middle and bottom graphs demonstrate the response and the prediction of smoke emissions. Table 4.1 and 4.2 show the results of evaluation from May 1st to 9th with and without the frames having steam. Table 4.3 shows the accuracy of twelve randomly picked days for each month in 2015.

We calculate true positives (TP), false positives (FP), false negatives (FN). Denote the boolean array of ground truth labels G and predictions P which contains only true and false entries. We first group the continuous true entries in G into a series of segments and apply the same process on prediction P. Next, for each segment in P, denote the starting and ending frame indices m_p and n_p. We mark a segment as a true positive if 30% of the entries in the segment contains true ground truth labels, which is described in (4.14). Otherwise, we mark the segment as a false positive.

$$\frac{\sum_{i=m_p}^{n_p} G(i)}{n_p - m_p + 1} > 0.3 \tag{4.14}$$

For each segment in G, denote the starting and ending frame indices m_g and n_g. We mark a segment as a false negative if $\sum_{i=m_g}^{n_g} P(i) = 0$, which means no entries in the segment contains true predictions. Finally, we compute precision (PR), recall (RE), and F-score by using (4.15).

$$\begin{aligned} \text{PR} &= \text{TP}/(\text{TP} + \text{FP}) \\ \text{RE} &= \text{TP}/(\text{TP} + \text{FN}) \\ \text{F-score} &= 2 * \text{PR} * \text{RE}/(\text{PR} + \text{RE}) \end{aligned} \tag{4.15}$$

4.5 Evaluation

Google Analytics evaluation of our website shows that from August 2015 to July 2016 there were 542 unique users, which contributed 1480 sessions. The average session duration was three minutes. We now discuss the image usage study for identifying how community members used animated images. Then we present the results of the survey study.

Table 4.1: The evaluation of all daytime frames for 9 days on May 2015.

Date	TP	FP	FN	Precision	Recall	F-score
May 1	15	36	4	0.2941	0.7895	0.4286
May 2	21	29	3	0.4200	0.8750	0.5676
May 3	24	28	8	0.4615	0.7500	0.5714
May 4	25	25	5	0.5000	0.8333	0.6250
May 5	14	19	4	0.4242	0.7778	0.5490
May 6	17	11	4	0.6071	0.8095	0.6939
May 7	26	16	3	0.6190	0.8966	0.7324
May 8	22	22	4	0.5000	0.8462	0.6286
May 9	16	23	1	0.4103	0.9412	0.5714
Avg				0.4707	0.8355	0.5964

Table 4.2: The evaluation of all daytime frames (exclude frames containing steam) for 9 days on May 2015.

Date	TP	FP	FN	Precision	Recall	F-score
May 1	13	8	4	0.6190	0.7647	0.6842
May 2	18	11	3	0.6207	0.8571	0.7200
May 3	24	19	6	0.5581	0.8000	0.6575
May 4	25	17	4	0.5952	0.8621	0.7042
May 5	13	9	3	0.5909	0.8125	0.6842
May 6	15	4	4	0.7895	0.7895	0.7895
May 7	26	6	3	0.8125	0.8966	0.8525
May 8	22	18	4	0.5500	0.8462	0.6667
May 9	14	17	1	0.4516	0.9333	0.6087
Avg				0.6209	0.8402	0.7075

Table 4.3: Evaluation of the smoke detection algorithm on 12 randomly chosen days for each month in 2015. TP, FP, and FN indicates true positive, false positive, and false negative respectively.

Date	TP	FP	FN	Precision	Recall	F-score
Dec 22	18	21	7	0.4615	0.7200	0.5625
Nov 15	18	6	1	0.7500	0.9474	0.8372
Oct 05	27	23	0	0.5400	0.9643	0.6923
Sep 09	10	35	8	0.2222	0.5556	0.3175
Aug 13	28	35	2	0.4444	0.9333	0.6022
Jul 08	15	35	9	0.3000	0.6250	0.4054
Jun 11	22	14	4	0.6111	0.8462	0.7097
May 28	24	17	3	0.5854	0.8889	0.7059
Apr 02	15	28	10	0.3488	0.6000	0.4412
Mar 06	1	8	15	0.1111	0.0625	0.0800
Feb 10	3	32	10	0.0857	0.2308	0.1250
Jan 26	1	5	2	0.1667	0.3333	0.2222

Figure 4.12: This figure shows a part of the collection of animated images generated by the timelapse viewer according to the results of smoke detection. The local community can select desired images and drag them into a Google Doc for documentation, presentation, and storytelling.

4.5.1 Image Usage Study

We evaluated the usage patterns of animated smoke images by parsing server logs. The logs stored HTTP requests of images from our server over an 11-month period from August 2015 to July 2016. Each request contained the source IP address, requested date, image URL, and browser type. Each image URL indicated its bounding box, size, time, and dataset. We first excluded all IP addresses from our research institute. Then for each HTTP request, we subtracted the requested date from the image taken date to get D, the difference in days, which indicated how far back in time a user viewed an image compared to when the image was taken. Table 4.4 shows summary statistics of animated images and users. The number of views of algorithm-generated images greatly exceeds the ones of human-generated images. Next we discuss two sub-studies which focus on images and users.

Image-based Sub-study

For the image-based sub-study, we separated images into two sets: created by human or created by the computer vision tool. Then for each set, we aggregated the number of images, views, viewed datasets, and users based on three criteria: viewing date (date that the image was viewed), dataset date (date that the image was taken), and D (difference in days). We now present three interesting findings.

First, while human-generated images were suitable for initiating community engagement, algorithm-generated images were useful for maintaining community engagement. In Figure 4.13, we aggregated number of views based on D, difference in days. The top graph in Figure 4.13 showed that a large portion of views of human-generated images had small D, which indicated a short period between when a user viewed an image and when the image was taken. This suggested that our users tended to create animated images manually by using the thumbnail generator after a recent event (e.g. smoke emission), which showed the purpose of initiating community engagement. However, most of the views of algorithm-generated images had high D (see the bottom graph in Figure 4.13). This showed that community members tended to use images generated automatically by the computer vision tool to review events occurring well beforehand, which demonstrated the objective of maintaining community engagement.

Second, the computer vision tool encouraged community members to explore more datasets. In Figure 4.14, we aggregated the number of views based on dataset date, the time that the image was taken. The top and bottom graphs in Figure 4.14 show results for human-generated and algorithm-generated images respectively. By comparing these graphs, the number of views of algorithm-generated images were more distributed across datasets than the ones of human-generated images, which were concentrated on specific days.

Third, the existence of the coke plant was significant in motivating the community to interact with the monitoring system. In Figure 4.15, we aggregated the number of views based on viewing date, the time that image was viewed. The figure shows that community members viewed much less human-generated and algorithm-generated images after Jan 2016, which was the time that the coke plant was closed.

# of unique and viewed HG images	135
# of views of all HG images	477
# of unique and viewed AG images	6745
# of views of all AG images	11043
# of total views	11520
# of users who created HG images	32
# of users who viewed HG images	85
# of users who viewed AG images	75
# of total users	141

Table 4.4: Summary statistics of animated smoke images and users. The "HG" and "AG" abbreviations mean "human-generated" and "algorithm-generated" respectively. The "#" sign means "number of". We can see that the number of views of algorithm-generated images greatly exceeds the ones of human-generated images.

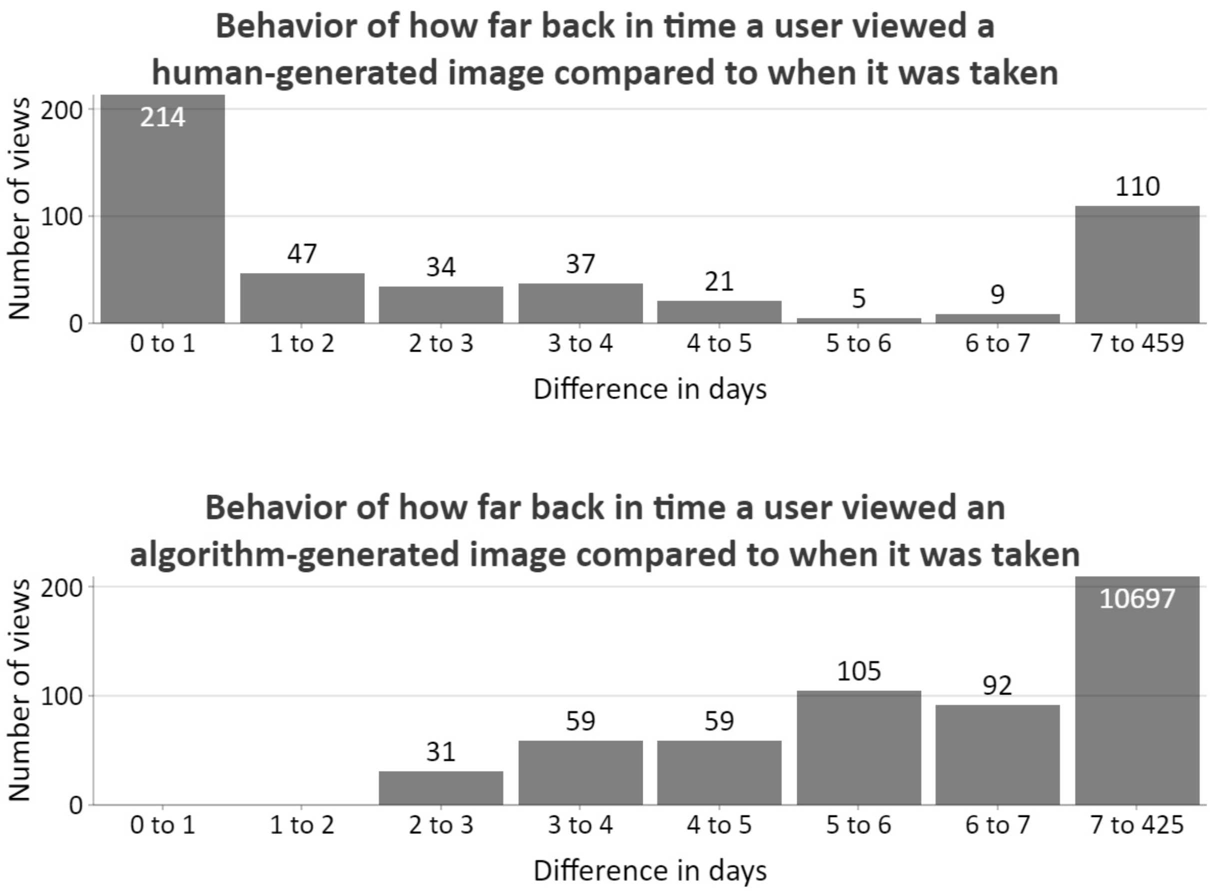
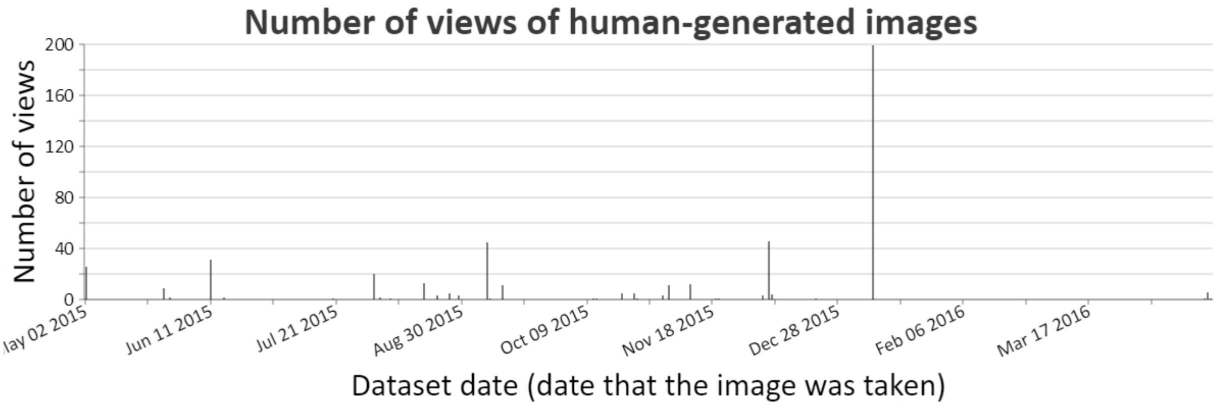


Figure 4.13: Behavior of how far back in time a user viewed a human-generated or algorithm-generated image compared to when it was taken. The x-axis is the difference in days (denote D) between the dates that an image was viewed and taken. Image views with small or large D mean they are used for verifying if an event, such as fugitive emissions happened (e.g. fugitive emission) or reviewing previous events respectively. While human-generated images were often viewed in less than one day after events occur, algorithm-generated images were usually viewed at least a week after the events.

User-based Sub-study

For the user-based sub-study, we aggregated the number of images, views, and viewed datasets based on unique IP addresses to obtain a series of vectors. To find relationships, we computed the correlation matrix of five vectors into the number of: created human-generated images, viewed human-generated images, viewed datasets in human-generated images, viewed algorithm-generated images, and viewed datasets in algorithm-generated images. We now summarize two findings.

First, there were strong correlations within the usage of human-generated images. Community members who created more images by using the thumbnail generator also viewed more human-generated images (Pearson's R Correlation = 0.91) and explored more datasets (Pearson's R Correlation = 0.89). Moreover, community members who viewed more human-generated images also explored more datasets (Pearson's R Correlation = 0.8).



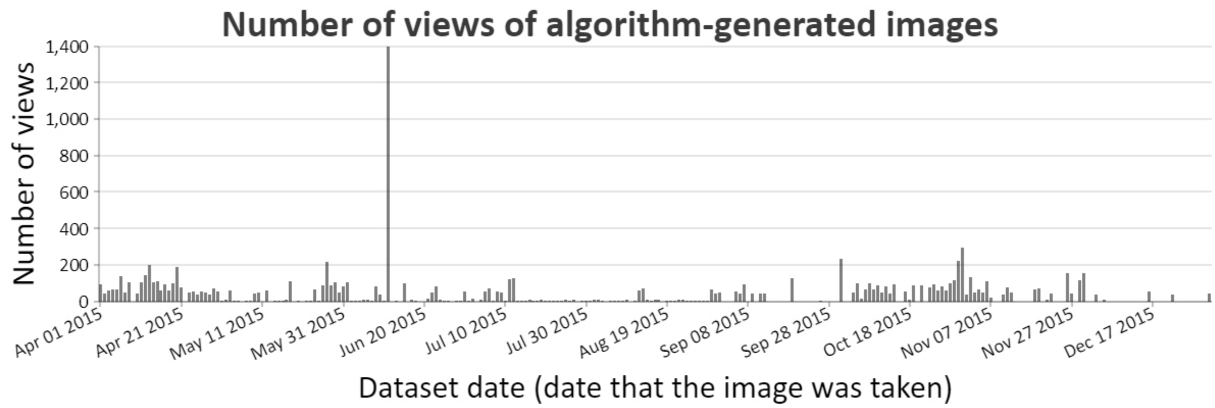


Figure 4.14: Number of views of human-generated or algorithm-generated images which are aggregated by dataset date. From these two graphs, we can see that the views of algorithm-generated images are more distributed across datasets, which means that users tend to use algorithm-generated images to explore events in different dates.

Second, it appeared that there was no obvious relationship between the usage of human-generated and algorithm-generated images. Community members who created or viewed more human-generated images did not necessarily view more algorithm-generated images (Pearson's R Correlation = 0.13 and 0.07 respectively). Furthermore, there were no strong correlations within the usage of algorithm-generated images. Community members who viewed more algorithm-generated images did not necessarily explore more datasets (Pearson's R Correlation = 0.35). The rhetorically compelling power of human-generated data should not be underestimated.

4.5.2 Survey Study

We now discuss the survey study for evaluating changes in the community's attitude after the intervention of our system.

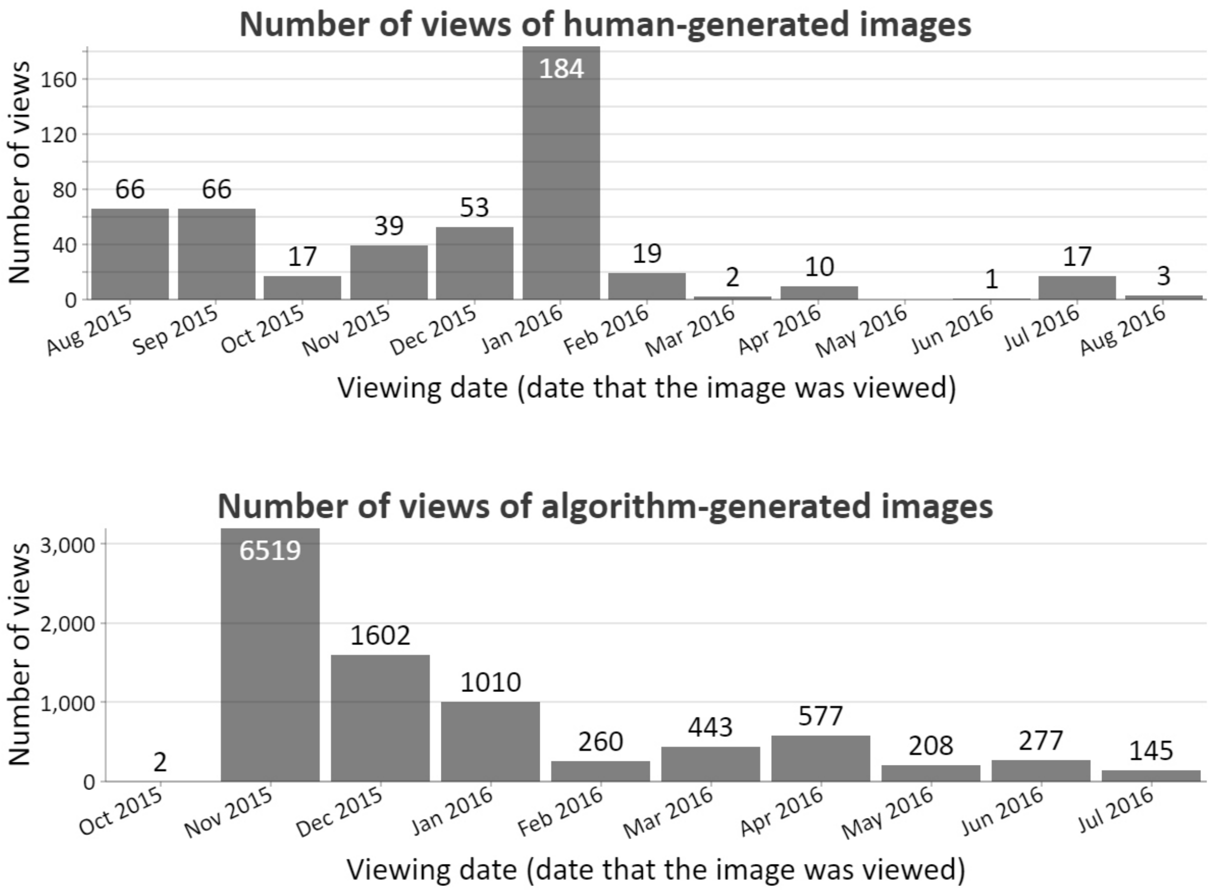


Figure 4.15: Number of views of human-generated or algorithm-generated images which are aggregated by viewing date. There is a significant decrease after January 2016, which was when the coke plant was closed.

Participants

ACCAN members were the primary users of the air quality monitoring system. Adult volunteers (age 18 and older) were recruited from these users through a Google Groups email. The email described the research purpose and included a link to an online survey. Paper surveys were also provided at a community meeting. All responses were kept confidential and there was no compensation. There was a brief consent script to review before taking the survey. We received 24 responses in total from 83 community members on the Google Groups (29% response rate). One invalid response which contained inconsistent answers and five incomplete ones were discarded. Most of the participants had a high education level and were over the age of 35 (see Table 4.5 for demographics).

Procedure and Materials

Participants filled out a survey. The survey was expected to take less than 30 minutes and contained three question types. The first type measured participants' involvement in the community action, such as exploring, documenting, and sharing data on the system. The second type mea-

	18-24	25-34	35-44	45-54	55-64	64-74	75+	Sum
No degree	0	0	0	0	0	1	0	1
Bachelor	1	1	1	0	2	2	0	7
Master	0	0	2	2	2	3	0	9
Doctor	0	0	0	0	0	0	1	1
Sum	1	1	3	2	4	6	1	18

Table 4.5: Age and education level for the participants of 18 valid survey responses. Participants have a high education level in general.

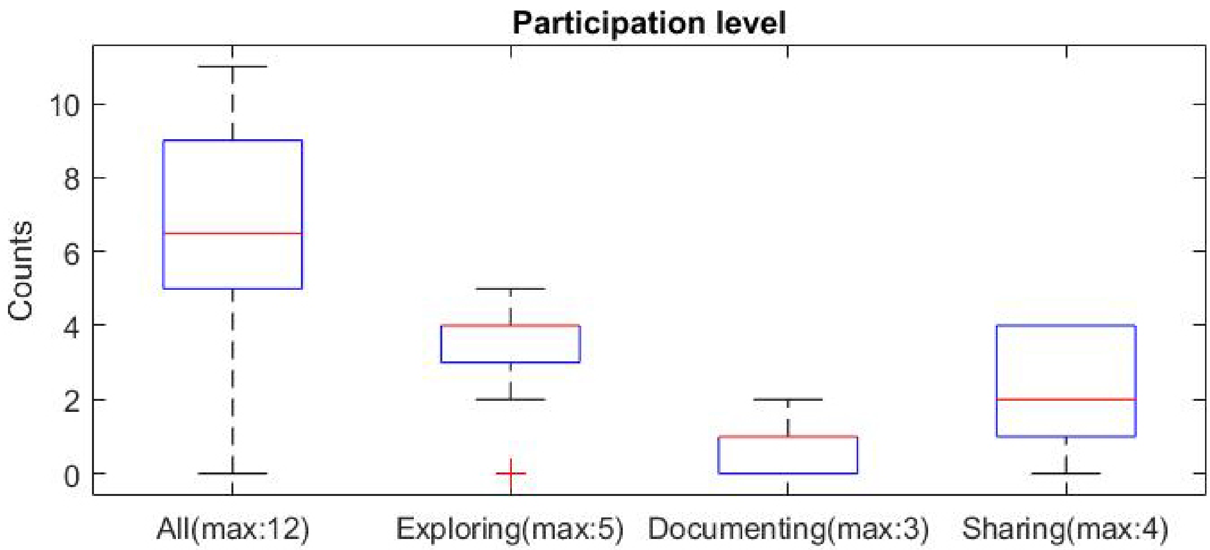


Figure 4.16: The boxplot of the participation level. We asked three multi-choice questions related to how users explore, document, and share the data provided by the system (the x-axis). These three questions had 5, 3, and 4 choices respectively. We summed up the number of choices that were selected by participants in each question to obtain participation levels (the y-axis). In general, the users had high participation levels.

sured community engagement, which included Likert scale questions related to the dependent variables: awareness, self-efficacy [15, 41], and sense of community [160]. The third type asked demographics, such as age range and education level. The range of the Likert scale was from 1 to 5, with 5 being the highest attitude.

Analysis

In the survey, participants answered three questions about how they explored, documented, or shared data by using the system. These three questions contained 5, 3, and 4 choices respectively. We summed up the number of choices that were selected by participants in each question to obtain participation levels (see Figure 4.16). We also asked questions about the frequency (from 1 to 5, with 5 being the highest frequency) of browsing the data in the system after noticing bad smells, number of people that a participant discussed the system with, and number of monthly meetings (from 0 to 12) attended in 2015 (see Table. 4.6).

	Browsing (V_b)	People discussed (V_d)	Meetings (V_m)
$\mu\|\sigma$	2.94|1.35	22.28|21.85	7.83|3.60

Table 4.6: The mean (μ) and standard deviation (σ) of other independent variables. V_b is the frequency (from 1 to 5, with 5 being the highest) of browsing the data in the system after noticing bad smells. V_d is the number of people that a participant discussed the system with. V_m is the number of monthly community meetings (from 0 to 12) attended in 2015. In general, participants were active in the community.

Figure 4.17: The boxplots of the changes of mental states among all participants after interacting with the monitoring system. The x-axis indicates dependent variables. The y-axis is the differences in Likert scale. Positive values mean increases, and vice versa.

For a dependent variable, participants answered a question set twice based on the time before (denote S_i^b) and after (denote S_i^a) they learned about the air quality monitoring system. Each question set had two Likert scale questions. We then averaged the Likert scales in set S_i^b and S_i^a to obtain a pair of scores. Figure 4.17 showed the difference of scores for each dependent variable. Positive values indicated increases, and vice versa.

Our directional null hypotheses were that the community did not have significant increases in awareness, self-efficacy, and sense of community. Since the differences of our paired samples did not follow a normal distribution (see Figure 4.17), we performed a right-tailed Wilcoxon signed-rank test, a nonparametric version of paired t-test. Table 4.7 showed the p-values and confidence interval.

Results

According to the analysis (see Table 4.7), the result favored the alternative hypotheses, which claimed there were significant increases ($p < 0.05$) in self-efficacy and sense of community after interacting with the system. The average increases in these two dependent variables were 0.53 and 0.56 respectively in Likert scale. However, we retained the null hypothesis, which

	Awareness	Self-efficacy	Community sense
p-value	0.2500	0.0042	0.0010
CI	0.08±0.15	0.53±0.40	0.56±0.38

Table 4.7: The p-value of right-tailed Wilcoxon signed-rank test and the confidence interval on the differences of paired samples. CI indicates 95% confidence interval. Gray cells indicate statistical significance ($p < 0.05$) or the confidence interval which does not contain zero.

stated there was no significant increase in awareness, since $p > 0.05$ and the confidence interval contained zero.

Open-ended answers in surveys showed that the monitoring system could encourage agonistic discussion with regulators and empower the community in supporting local policy making. With the system, community members could report concrete scientific evidence of fugitive emissions to the local health department, such as animated smoke images and the exact time of emissions, instead of vague reports.

> *"I **made screenshots** of the [system name] dashboard at different times/days when wind was strong and in the direction of my community. I inserted these screenshots into Powerpoint slides. I **shared printed versions of these slides** with my Township commissioner when asking for assistance in reducing emissions."*
>
> *"I **continually spoke** at regional meetings, City, County, Health Department, Clairton, Lawrenceville, etc. Wrote numerous letters to the editor, most did get published, not all."*
>
> *"I reported specific emissions from [coke plant name] to ACHD. I was able to **provide specific times** so that ACHD could review the exact episodes that I was reporting."*
>
> *"I **shared web links** to the [system name] when I submitted complaints to the health department"*
>
> *"Confronted ACHD staffers repeatedly with **uncomfortable info**."*
>
> *"I **e-mailed images** to others, including regulators."*

Moreover, others mentioned that their confidence in taking action was significantly improved after interacting with the system. One important reason was that integrating heterogeneous data (smoke images, air quality data, smell reports, and wind information) formed strong scientific evidence, which was powerful in communicating with regulators and thus changed the power relationship between citizens and the government.

> *"I felt that the more information/proof that I made available might help justify my concern and spur action. I **felt** that my concerns with what I was experiencing were **grounded in actual imagery, wind data and spatial data**."*
>
> *"I **believe** that the [system name] was very important in helping us **get the attention of regulators** (ACHD and EPA) and get them to take our concerns seriously."*
>
> *"The [system name] was one of the most important tools the community has in holding the plant accountable. I **believe** that images presented at the Nov. 2015 EPA ACHD ACCAN meeting **provided a tipping point for the plant's shutdown**."*

	$\mu|\sigma$
The timelapse video	4.81|0.54
Zooming in and out of the video	4.50|0.73
Sharing a web link of a view and time	4.43|0.85
Smell reports	4.38|0.81
Line charts showing sensor readings	4.31|0.87
The map showing sensor values	4.44|0.73
The thumbnail tool	4.19|0.83
The automatic smoke detection tool	4.31|0.70
Smoke images shown on the meeting with EPA	4.94|0.25

Table 4.8: The mean and standard deviation ($\mu|\sigma$) of the importance rating of features on the air quality monitoring system. In general, participants rated all features important.

> *"I **believe** that the [system name] images shown at the November 2015 community meeting 'tipped the balance' for the EPA and may have resulted directly in the closing of [coke plant name]. In fact, **without those images, it may have taken years to close the plant**."*

In addition, several community members specifically identified the political and educational values of the monitoring system. In addition, they showed a desire of reproducing the monitoring system on other neighborhoods.

> *"Background as a **environmental law paralegal**."*
>
> *"Fantastic **educational** tool."*
>
> *"I would like to see **similar monitoring of other pollution sites** in Pittsburgh, ie. the [other coke plant name] and others mentioned in the Toxic Ten listing."*

4.6 Discussion

The community that we collaborated with has fought for decades to resolve the air pollution problem, which existed since 1999. The monitoring system was launched in Fall 2015. In November 2015, the community held a meeting at their local church with government officials from the ACHD (Allegheny County Health Department) and the EPA. During the meeting, as information technology supporters, we demonstrated the system and the visualization. In addition, the community projected hundreds of animated smoke images generated by the system on a large screen in front of ACHD and EPA regulators. Community members described how their living quality was affected by the air pollution together with animated smoke images, air quality sensors, crowdsourced smell reports, and wind data. The scientific knowlege demonstrated how heavy air pollution flowed into the neighborhood. The community successfully combined personal experiences and scientific evidence into a story to convince regulators. The story showed that the pollution source was the coke plant, and its fugitive emissions acturally affected the local air quality. This forced regulators to respond to the air quality problem publicly. The acting director of the EPA from the Region III Air Protection Division in Philadelphia pointed at the screen and said: *"But what I see in the video, is totally unacceptable."* In addition, the local air quality problem became available for further debate and investigation. The administrator agreed that the EPA would continue to review the coke works' compliance with the 2012 federal consent decree. Furthermore, on December 2015, the parent company of the coke works announced the closure of the plant, which was the ultimate goal that the community had tried to achieve for decades.

4.6.1 Insights

Based on the major community meeting described in the previous paragraph and the results presented in the previous section, we now summarize our findings into three key insights and offer suggestions to future researchers.

Use a Flexible and Iterative Design Process

We encourage using a flexible and iterative procedure instead of a single and prescribed one. This practice is also mentioned by DiSalvo et al. [70] as community co-design [30], a process which involves community members when designing a system that supports citizen empowerment. Often there are attempts to duplicate successful systems in another similar real-world context. However, this is unlikely to succeed because the environmental problem that the community deals with is wicked [50, 190]. Every wicked problem has no clear formulation, is unique, and cannot be fully observed. Therefore, like the experience we describe in the design process and system sections, we recommend scheduling multiple design phases to reveal unique challenges and to apply specific solutions on these challenges iteratively. In the survey study, participants rate the importance of features of the system (see Table 4.8). The rating scale is from 0 to 5, with 5 being the most important. The average ratings are all above 4, which verifies that the iterative design process help develop altogether useful system features to the community.

Initiate and Maintain Community Engagement

It is critical to initiate and maintain community engagement via actual participation in using the system. We recommend combining manual and automatic approaches, which are the thumbnail generator and the computer vision tool respectively in this work, to serve two different purposes in citizen participation. First, a manual approach can initiate citizen participation and lead to follow-up interactions. The image usage study shows that community members use the thumbnail generator to manually create images after they notice unusual events (see Figure 4.13), such as industrial smell or hazardous smoke. Correlation analysis of image usage indicates that users who create more images also view more images and explore more datasets (see the User-based Sub-study subsection). Second, an automatic approach can encourage community members to participate in a long temporal horizon. Smoke images generated automatically by the computer vision tool are used for reviewing fugitive emissions (see Figure 4.13). The computer vision tool encourages community members to explore more datasets (see Figure 4.14). However, it appears that there are no clear correlations between the manual and automatic approach (see the

User-based Sub-study subsection). How to integrate these two approaches seamlessly to open up and maintain citizen participation remains an important research question.

Enable the Formation of Scientific Knowledge via Hybrid Data

Data requires being interpreted into scientific knowledge to be impactful in changing unbalanced power relations between citizens and governments. Besides collecting data, providing affordance for citizens to make sense of the relationship among various types of data is key to generating scientific knowledge. We suggest integrating image, sensor, and crowdsourced data from both human and machines into such a system. Analysis in the survey study is limited by the small sample size of total users, and this should be taken as a caveat in regards to analysis of statistical significance. Nonetheless, Figure 4.17 shows the changes of participants' attitudes and Table 4.7 includes statistical significance findings in self-efficacy and sense of community. Open responses in the survey show that with scientific knowledge, citizens can present data in meaningful ways to regulators who have the power to make policy changes. At the meeting in November 2015, the community successfully influenced the attitude of the government after presenting the evidence. Scientific knowledge gives citizens power to advocate for their living quality and to influence other stakeholders.

4.6.2 Limitation

Measuring information and communication technology (ICT) interventions in community advocacy is generally challenging. Community advocacy has the ultimate goal of policy change, yet it is difficult to causally prove how critical to a successful policy change the communities' actions have been. Such projects succeed not only when policy goals are achieved, but in how the relationship between citizens, policy makers, and businesses evolves. This work shows that making scientific data transparent to stakeholders can foster sustainable relationships among them. It is sustainable in the sense that the system promotes a healthy and balanced power structure for democracy in the long term. We believe patterns of scientific data usage and changes of mental state among community members are useful proxies for evaluating the effectiveness of such projects. To better understand usage patterns, we suggest tracking the usage of data in the system. Future research about how to evaluate ICT interventions is still needed. For instance, qualitative research, like in-depth interviews, will be needed to identify key factors for successful collaboration between stakeholders and to understand changes of power dynamics among citizens, scientists, developers, and regulators. Moreover, forming scientific knowledge about the relationship between the smoke emissions and the severity of the air pollution by using the monitoring system currently relies on human interpretation. Additional future research involves enhancing the knowledge by analyzing the correlations between various types of data. The analysis can explain how these data reinforce or conflict with each other, which provides strong statistical scientific evidence.

Another limitation is that the sample size of participants in the survey study is too small and the statistical analysis conclusion (see subsection Results) is weak. Participants only represent a fraction of the population in the neighborhood near the coke works. They have high education (see Table 4.5) and involvement levels (see Table 4.6 and the left-most boxplot in Figure 4.16), which includes interacting with the system, discussing the system with others, and attending monthly community meetings. Most of them have strong activation before learning about the monitoring system, which causes the failure to reject the null hypothesis related to awareness (see Table 4.7). The strong activation may also result in the high correlation between community members who created and viewed smoke images (see subsection User-based Sub-study). Nevertheless, one alternative explanation of this limitation is that without high awareness, it would be impossible to support community advocacy with ICT interventions. In other words, high awareness may be a necessary condition for successful citizen empowerment. How attitude may change among people with low education or low involvement level after interacting with the air quality monitoring system still remain an open research question.

Furthermore, the smoke detection algorithm used in the system is tuned to operate in our settings. Currently, the algorithm uses a heuristic method and has too many tuning parameters, which is not robust enough for similar contexts for other communities. One approach to generalize the system is to collect crowdsourced labels via mobile or online platforms, which requires deeper citizen participation. These labels can then be used to train a smoke image classifier using machine learning. Moreover, it appears that the existence of the coke plant, which poses personal risk, is the major source of motivation (see Figure 4.15). This crowdsourcing approach may provide extra motivations to the community. Besides collecting labels, organizing the hybrid scientific data collected in the system into a comprehensive dataset can potentially assist future academic research related to environmental problems.

4.7 Summary

This chapter presents a web-based air quality monitoring system which integrates image, sensor, and crowdsourced data. It is an instance of adversarial design [66, 67] which critically reveals, questions, and challenges a real-world environmental problem. The system provides technological affordance for forming strong scientific evidence. We discuss the iterative participatory design process that leads to decisions of system features with the community. We describe our evaluation, which includes an image usage study from server logs and a survey study. The survey study indicates statistically significant increases in self-efficacy and sense of community among users after interacting with the system. Open responses in the study show that the system promotes critical discussions with policy makers and empowers citizens to participate in community actions. Based on the evaluation, we offer three key insights about using an iterative design process, encouraging community engagement, and forming scientific knowledge. Finally, we mention limitations and future research directions related to evaluating the intervention of information technology, studying user behavior of community members with low participation level, and generalizing the smoke detection algorithm by collecting crowdsourced labels. We hope that this work can inspire other researchers to contribute towards developing innovative information technology that supports citizen empowerment.

Chapter 5

Visualization Tool for Environmental Sensing and Public Health Data

To assist residents affected by oil and gas development, public health professionals in a non-profit organization have collected community data, including symptoms, air quality, and personal stories. However, the organization was unable to aggregate and visualize these data computationally. This chapter presents the Environmental Health Channel, an interactive web-based tool for visualizing environmental sensing and public health data. The tool enables discussing and disseminating scientific evidence to reveal local environmental and health impacts of industrial activities. This work addresses the science communication and data quality challenges (as mentioned in section 1.3) to initialize community engagement in community-oriented citizen science. The main contribution of the work is to provide reusable computational tools for forming and sharing evidence related to environmental health.

5.1 Preface

Air quality and its impacts on public health are critical environmental issues for residents who live near oil and gas development sites [48]. A vital step towards addressing these issues is through the collection and dissemination of convincing scientific evidence of these impacts [112, 113]. However, conveying this evidence, especially with multiple types of data at a large temporal and geographic scale, requires the assistance of computational tools. In the pursuit of developing a tool for this purpose, we collaborated with a local non-profit organization that is working to study and assist communities that are potentially affected by oil and gas development. Since 2014, the organization has collected data which includes (1) particulate measurements from air quality sensors, (2) physical and psychosocial symptoms from surveys, and (3) personal stories from interviews. These citizen-contributed data were stored across multiple incompatible systems, which hindered retrieving information, visualizing trends, and disseminating findings. Moreover, the organization lacked the resources to independently develop computational tools for aggregating and visualizing data to facilitate user decision-making. Therefore, we collaborated with health professionals from the non-profit organization to develop the Environmental Health Channel (EHC), an interactive web-based data visualization tool (see Figure 5.1). The

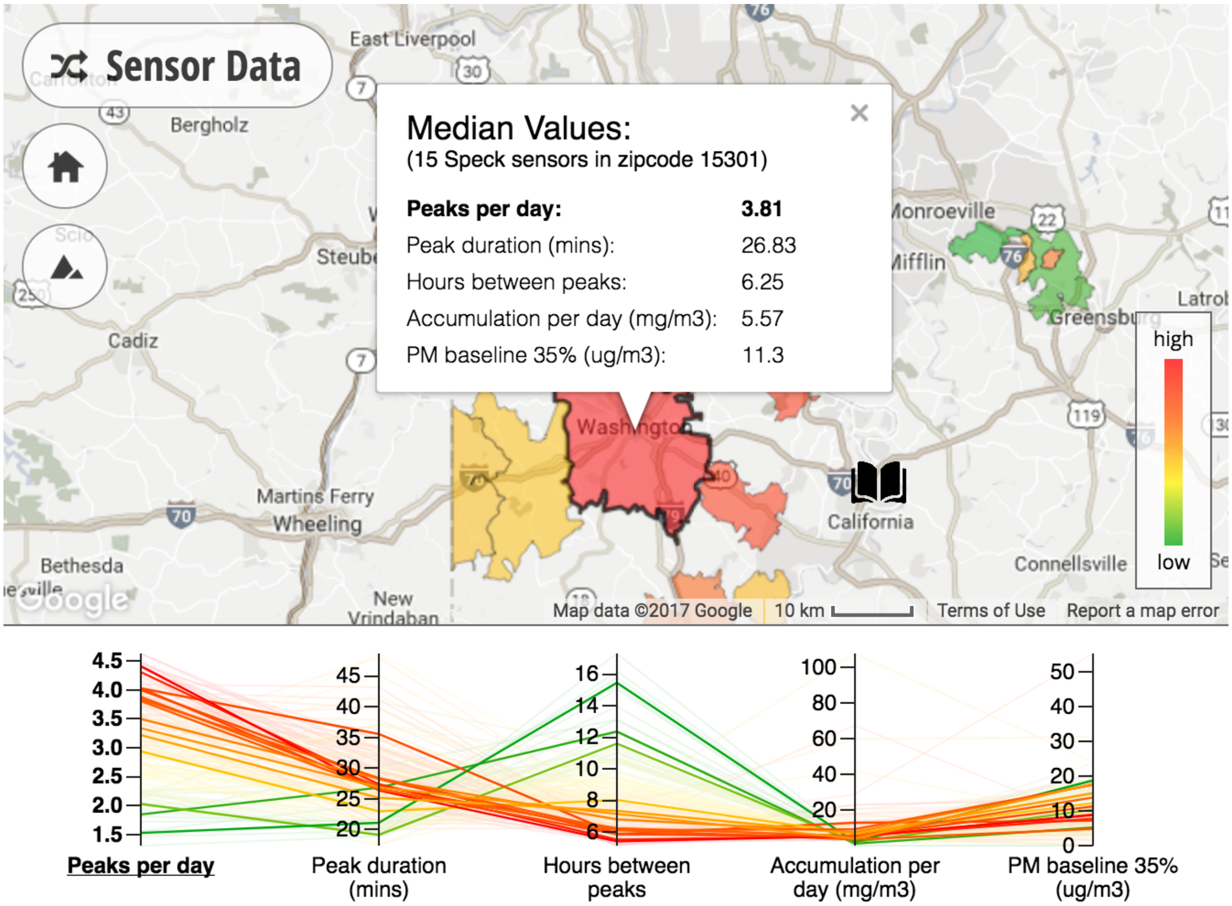


Figure 5.1: The user interface of the Environmental Health Channel, which visualizes the analysis of air quality sensors.

goals were to (1) make citizen-contributed data explorable through visualization, (2) enable users to communicate and share air quality issues with scientific evidence, and (3) empower community members to make evidence-supported decisions.

5.2 System

During system development, we collaborated with health professionals from the non-profit organization in implementing system features. We began the design process by investigating the data types that the non-profit organization gathered from affected residents, as different data types require distinct visualization affordances. There were three data types: air quality metrics, self-reported health symptoms, and personal stories with images. Since 2014, the non-profit organization has provided portable air quality sensors [206, 217] to affected residents. After a month of placing sensors indoors and outdoors, the organization collected the sensors, computed air quality statistics from the raw sensor values, and presented these statistics to affected residents in report form. Also, affected residents filled out a self-reporting health survey to indicate physical and psychosocial symptoms that they experienced during the period when sensors were placed.

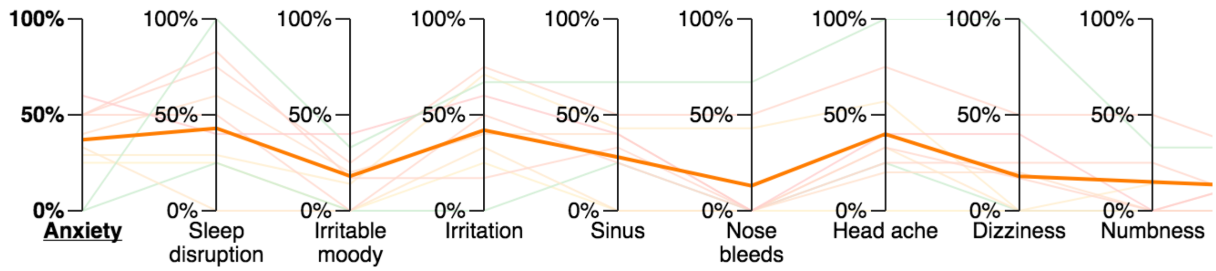


Figure 5.2: When selecting health data by clicking on the top-left button in Figure 5.1, the bottom parallel coordinate plot changes.

The organization interviewed several affected residents about their personal stories of living near oil and gas drilling sites and collected photographs of their home environments. From these interviews, the organization created a series of photos with narrative text. Integrating the sensor, survey, and interview data into EHC posed privacy issues. To protect the privacy of participants, we de-identified and aggregated data based on zip code boundaries. This approach addressed the concern that confidentiality could be compromised by re-identification of data. EHC stored these de-identified data in a Google Sheet, which enabled the stakeholders to work collaboratively on adding more citizen-contributed data in the future with ease without programming skills. To automate the process of updating data, a Python script on the server periodically parsed the Google Sheet data into suitable formats for each visualization.

EHC permits reviewing and comparing aggregated data among different regions simultaneously. To enable interpreting patterns and identifying key policy issues from multiple types of data, we implemented a **heatmap**, a **parallel coordinate plot**, and a **story slider** in HTML and JavaScript. The **heatmap** (see the top part of Figure 5.1) contains colored polygons to indicate zip code regions which contain air quality sensor data. A color legend (see bottom-right of the map part of Figure 5.1) displays the relative color scale from green, yellow, orange, to red, which corresponds to -1, -0.5, 0.5, and 1 standard deviation away from the mean value respectively. When users click on a colored zipcode, an information window shows up to provide summary statistics of air quality data in the corresponding zip code region. The **parallel coordinate plot** [114] (see the bottom part of Figure 5.1) displays the distribution of summary statistics describing air quality or health data. Each axis of the plot represents one statistic, such as the average number of air quality peaks per day. This plot allows users to visually compare relative values of a statistic across different zip code regions. For instance, when the number of peaks per day is selected (see Figure 5.1), red-colored zip code regions on the map have a relatively higher number of peaks per day than all other regions. Users can select a statistic by clicking on the corresponding label on the axis. The **story slider** (see Figure 5.3) shows personal stories and images collected from interviews. This combined visual and narrative presentation offers insight into personal experiences with oil and gas exposures and their involvement with air monitoring. Users can click on open-book icons on the heatmap to explore stories on the slider.

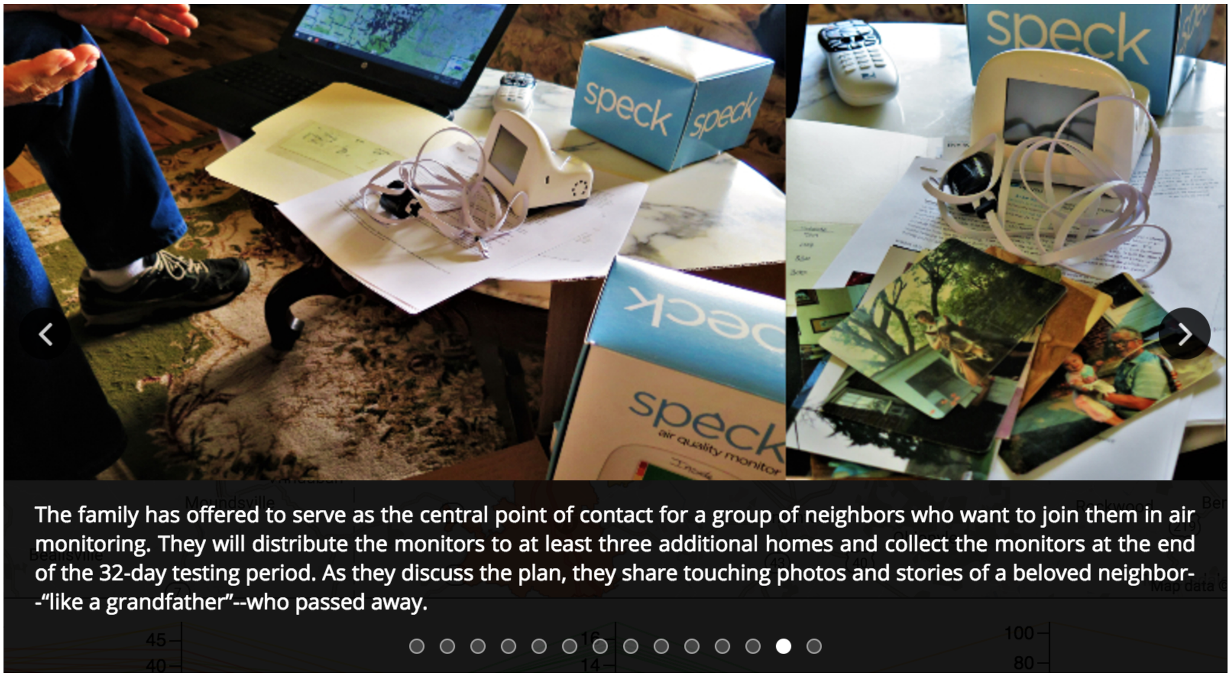


Figure 5.3: The image slider of personal stories from residents.

5.3 Evaluation

We conducted a 2-hour focus group study [136, 142, 187] and applied affinity diagramming [16, 18, 110, 123, 154] to gain insights about: (1) potential issues about system features and (2) affordances that EHC provided or would support in the future. Seven air quality experts were invited to discuss EHC with a software developer and three health professionals. We found that the discussion was centered around three themes found in previous research [68, 126, 127, 128, 137, 138]: exploration, investigation, and advocacy. First, **exploration** refers to supporting the understanding of air quality variables, data sources, and visualizations. For instance, participants mentioned the importance of providing instructions and explanations to users about the provided sensor statistics and the health variables. Participants also suggest that the color red should always indicate a qualitatively worse situation as it relates to potential health impacts, instead of a numerically higher value. Second, **investigation** pertains to recognizing and comparing data patterns, forming hypotheses, and building narratives with evidence. For example, providing methods for simultaneously comparing health and air quality data is critical for allowing users to investigate the hypotheses that interest them. Additionally, participants recommended adding background variables, such as demographics, to provide more context and enhance scientific evidence. Third, **advocacy** refers to validating data, taking actions with scientific evidence, and advocating for social impact and political change. For instance, as stories are compelling in evoking emotions and may leave users with the desire to take action, participants suggested adding resources at the end of the story slider to encourage community engagement. Moreover, participants pointed out that there is a need for abstracting data and visuals into concise and convincing reports that can easily be shared with stakeholders and raise the awareness of air quality issues.

5.4 Discussion and Summary

EHC has been deployed in the local community affected by oil and gas development. Although EHC is being iteratively improved, it enables and encourages health professionals in the non-profit organization to add, visualize, and share incoming data interactively among stakeholders and citizens without the assistance of computer scientists. With the help of air quality experts and health professionals, we have conducted a focus group study to understand issues about system features and determined possible future directions. The result supports the findings in previous research conducted by DiSalvo [68], Kuznetsov [137, 138], and Kim [126, 127, 128]. As participants in this study were limited to experts, the result does not reflect the opinions of users with other levels of participation and expertise, such as residents or the general public. Future work will involve conducting more focus group studies to receive feedback from a broader audience. Moreover, we have not evaluated the impact of EHC on experts nor residents. Future research is needed to understand motivations of participation and evaluate attitude changes after using EHC, such as changes in the awareness of air quality problems, confidence in reaching goals, and sense of belonging in a community. We hope that this work will lay a foundation for researchers who develop information technology that provides scientific evidence from multiple perspectives to empowers citizens.

Chapter 6

Smell Pittsburgh: A Crowdsourced Mobile Application for Reporting and Visualizing Pollution Odors

Urban air pollution can have a negative impact on human health. Citizens who suffer from poor air quality mostly rely on experts to identify pollution sources due to the lack of accessible computational tools. This chapter presents Smell Pittsburgh, a crowdsourced mobile application that equips citizens with the capabilities to report pollution odors and track where these odors are frequently concentrated. The smell reports are sent to the local health department and visualized on a map along with fine particulate matter and wind data from the local federal monitoring stations. The visualization provides a convincing overview of the urban air pollution landscape. Additionally, Smell Pittsburgh applies machine learning methods to periodically generate push notifications that inform citizens about the potential presence of pollution odors. This work also assesses the validity of using citizen-contributed data in drawing meaningful insights to identify air quality problems through statistical prediction and inference. In the evaluation, we conduct qualitative and quantitative studies to measure changes in engagement and understand motivation factors for submitting smell reports. The results reveal generalizable design implications for developing and deploying similar tools in other real-world contexts. This work addresses the data quality, science communication, and evaluation metrics challenges (as mentioned in section 1.3) for initializing, maintaining, and evaluating community engagement. The contributions of this work include a methodology of crowdsourcing and visualizing smell reports on a city-wide scale, a procedure of evaluating the value of citizen science data with machine learning, and a study of identifying attitude changes and motivation factors.

6.1 Preface

Urban air pollution is of great concern due to its negative impact on human health and quality of life [73, 122, 177, 181, 224]. Conventional techniques for addressing air pollution involve negotiations between corporations and regulators, who in general hold power to improve air quality. Although air quality policy significantly affects the health of citizens, they rarely participate in

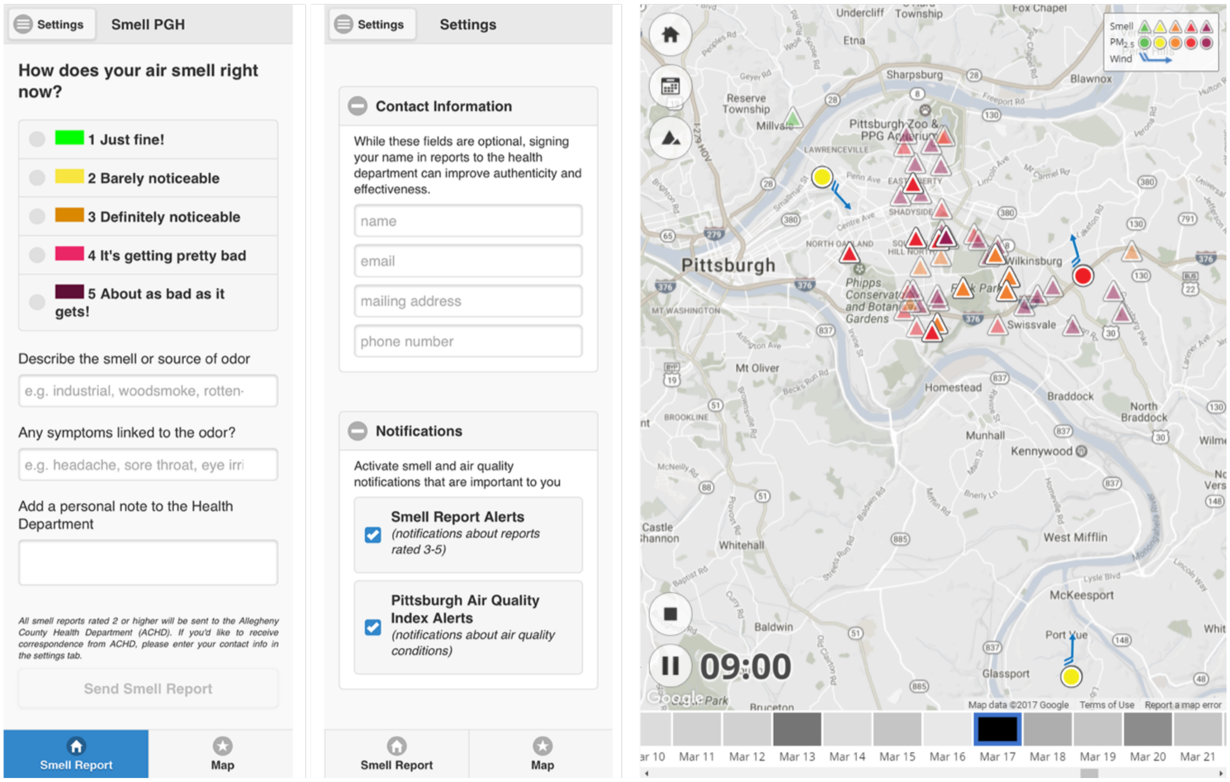


Figure 6.1: The user interface of Smell Pittsburgh. The left image shows the submission console for selecting and describing smell characteristics, explaining symptoms, and providing notes for the local health department. The middle image shows the setting menu for push notifications and personal identifiers when submitting smell reports. The right image shows the visualization of smell reports, sensors, and wind directions.

policy-making directly, and their voices typically fail to persuade decision-makers. To influence policy, citizens often need to present reliable scientific evidence to support their argument [170]. Forming such evidence requires collecting, processing, and interpreting multiple sources of data over a large geographic area and an extended period. This task is challenging due to the requirements of financial resources, organizational networks, and access to technology. As a result, affected residents must rely on experts in governmental agencies, academic institutions, or non-governmental organizations to analyze and track pollution sources.

Citizen science and crowdsourcing is a promising approach for citizens to pool resources and efforts to gather scientific evidence for advocacy. However, crowdsourced data is often held in low regard because the information can be unreliable or include errors during data entry. Additionally, there may be insufficient citizen participation to validate the data. For instance, the city involved in this study, Pittsburgh, is one of the ten most polluted cities in the United States [5]. Currently, Pittsburgh citizens report air quality problems to the local health department via its phone line or a textbox on its website. Nevertheless, the quality of gathered data is doubtful. Citizens may not remember the exact time and location that pollution odors occurred. Asking citizens to submit complaints retrospectively is hard for capturing accurate details and prone to errors. Such errors can result in missing or incomplete data that can affect the outcome of statistical analysis to identify pollution sources [60]. Furthermore, the reporting process is not transparent and does not encourage citizens to contribute data. There is no real-time feedback or ways of sharing experiences to encourage citizen participation and forge a sense of community. Without data that adequately represents the Pittsburgh community, it is difficult to know if an air pollution problem is at a neighborhood or city-wide scale. This approach is inadequate for gathering citizen data and hinders the participation in bringing air quality issues to the attention of regulators and advocating for policy changes.

To improve the crowdsourced data quality and increase citizen engagement, we propose a computational tool, Smell Pittsburgh. Citizens can use this mobile application to report pollution odors to the local health department with accurate time and GPS location data via cellular networks, anywhere in the city. The application also visualizes these odor complaints in real-time, which enables community members to confirm their personal experiences by viewing if others also share similar experiences. Additionally, users receive push notifications whenever a predictive model detects potential presence of pollution odors. Using Smell Pittsburgh data collected over the past year and a half, we can assess the value of citizen-contributed data by applying machine learning methods to identify connections between smell reports and air quality sensor readings. In the evaluation section, we describe qualitative and quantitative studies designed to understand changes in user engagement and motivation over time. Finally, we provide design implications for developing systems to empower data-driven community action and conclude with limitations. Overall, our contributions are:

- Detailed documentation of our approach to collecting and visualizing odor complaints at a city-wide scale
- Evaluation of citizen-contributed data, and the value of crowdsourced information in identifying and revealing air quality concerns
- Analysis of citizen engagement and motivation factors
- Insights for developing information technology to empower citizens for environmental advocacy

6.2 Design Principles and Challenges

The ultimate goal is to develop an interactive system to (1) lower the barriers for citizens to participate in scientific research and (2) democratize scientific knowledge for citizens to advocate for better air quality. While our community is dedicated to identifying pollution sources, this is challenging because air quality is affected by atmospheric conditions and may not always be visible. To understand the impact of urban air pollution, we focus on crowdsourcing one specific type of lay knowledge: smell. The human sense of smell is highly sensitive and can potentially be a useful measurement of urban air pollution events, such as high concentration of volatile organic compounds. The human olfactory system can distinguish more than one trillion odors [39] and outperform sensitive measuring equipment in odor detection tasks [199]. Furthermore, community members frequently use smell to indicate pollution events [170] and support decision making in daily lives [168].

Prior works have applied a smell-walking approach to record and map the landscape of smell experiences by recruiting participants to travel in cities [106, 182, 183]. Although Quercia et al.

correlated air pollutants with odors obtained by using smell-walking, the goal of smell-walking is to construct and validate a generalizable dictionary of smell types instead of revealing air quality concerns [182]. There is a lack of research in understanding the potential of using smell as an indicator of urban air pollution. Also, while data generated from smell-walking has a dense temporal resolution, it poses high workloads for participants to be extremely involved in completing tasks, and thus restricts the scale of the project. In our case, we intend to maximize community engagement by minimizing the efforts for citizens to contribute smell data. This tradeoff makes our data temporally sparse but grants the capability to engage a large number of citizens on a city-wide scale.

Chapter 4 has shown promising results of utilizing cameras, air quality sensors, and smell experiences to form and present scientific evidence about air pollution collaboratively. However, its approach to collect smell reports via an online Google Form is not scalable. The scope of the study in Chapter 4 is limited to a local community affected by one known and nearby pollution source. In contrast, we are interested in revealing air quality concerns on a city-wide scale with more than 300,000 affected residents over several years. In our case, there are multiple undetermined pollution sources within different distances from the impacted city [3]. Conducting community citizen science over such large-scale becomes much more complicated. Smell Pittsburgh integrates both human-generated and machine-generated data, including air quality sensor measurements and citizen-contributed smell reports, to provide a contextualized landscape of air pollution in industrialized urban areas. To the best of our knowledge, Smell Pittsburgh is the first system of its kind that can crowdsource and visualize smell experiences at such large scale to form evidence about urban air quality issues.

To invite citizens to contribute data when launching Smell Pittsburgh, we made use of an existing network of community advocacy groups, including ACCAN [6], GASP [8], Clean Air Council [7], PennFuture [9], and PennEnvironment [3]. These groups were pivotal in shaping the design of Smell Pittsburgh and providing insights into how to engage the broader Pittsburgh community. To sustain participation, we visualized smell report data on a map and also engage residents through push notifications. To add more weight to citizen-contributed pollution odor report, we engineered the application to send smell reports directly to the Allegheny County Health Department (ACHD). This strategy ensured that the local health department could access high resolution citizen-generated pollution data to ascertain better and address potential pollution sources in our region. We met and worked with staff in ACHD to determine how they hoped to utilize smell report data and adjusted elements of the application to better suit their needs, such as sending data directly to their database and using these data as evidence of air pollution. Based on their feedback, the system submitted all smell reports to the health department, regardless of the smell rating. This approach provided ACHD with a more comprehensive picture of the local pollution landscape.

6.3 System

To initiate and sustain citizen participation, we developed Smell Pittsburgh, a mobile application on iOS and Android devices to crowdsource and track pollution odors in industrialized urban areas. We now describe two system features: (1) a mobile interface for submitting and visualizing odor complaints and (2) push notifications for predicting the potential presence of odor events.

6.3.1 Submitting and Visualizing Smell Reports

Users could report odor complaints via Smell Pittsburgh from their mobile devices via the submission console (the left-most image in Figure 6.1). To submit a report, users first selected a smell rating from 1 to 5, with one being "just fine" and five being "about as bad as it gets." These ratings, their color, and the corresponding descriptions were designed by affected local community members to mimic the US EPA Air Quality Index [82]. Also, users could fill out optional text fields where they could describe the smell (e.g., industrial, rotten egg), their symptoms related to the odor (e.g., headache, irritation), and their personal experiences. Once a user submitted a smell report, the system sent it to the local health department and anonymously archived it on our backend database. In the setting panel (the middle image in Figure 6.1), users could decide if they were willing to provide their contact information to the health department. Regardless of the setting, our database did not record the personal information.

Upon receiving smell reports, we visualized them on a map that also depicted fine particulate matter and wind data from government-operated air quality monitoring stations (the right-most image in Figure 6.1). All smell reports were anonymous, and their geographical locations were skewed to preserve privacy. When clicking or tapping on the playback button, the application animated 24 hours of data for the currently selected day, which served as convincing evidence of air quality concerns. Triangular icons indicated crowdsourced smell reports with colors that correspond to smell ratings. Users could click on a triangle to view details of the associated report. Circular icons showed government-operated air quality sensor readings with colors based on the Air Quality Index [82] to indicate the severity of particulate pollution. Blue arrows showed wind directions measured from nearby monitoring stations. The timeline on the bottom of the map represented the concentration of smell reports per day with grayscale squares. Users could view data for a specific date by selecting the corresponding square.

6.3.2 Sending Push Notifications

Smell Pittsburgh sent two different types of smell event notifications to encourage citizen participation: a crowdsourced notification and a predictive notification. When there were a sufficient number of poor odor reports during the previous hour, the system sent a crowdsourced notification: "Many residents are reporting poor odors in Pittsburgh. Were you affected by this smell event? Be sure to submit a smell report!" The intention of sending this notification was to encourage users to check and report if they had similar odor experiences. Second, we applied machine learning [23, 103, 119, 162] to model the relationships between crowdsourced smell reports and air quality measurements from the past to predict the occurrence of abnormal odors in the future. Each day, whenever the model predicted a smell event, the system sent a separate predictive notification: "Local weather and pollution data indicates there may be a Pittsburgh smell event in the next few hours. Keep a nose out and report smells you notice." The goal of making the prediction was to support users in planning daily activities and encourage community members to pay attention to the air quality. To keep the smell prediction system updated, we computed a

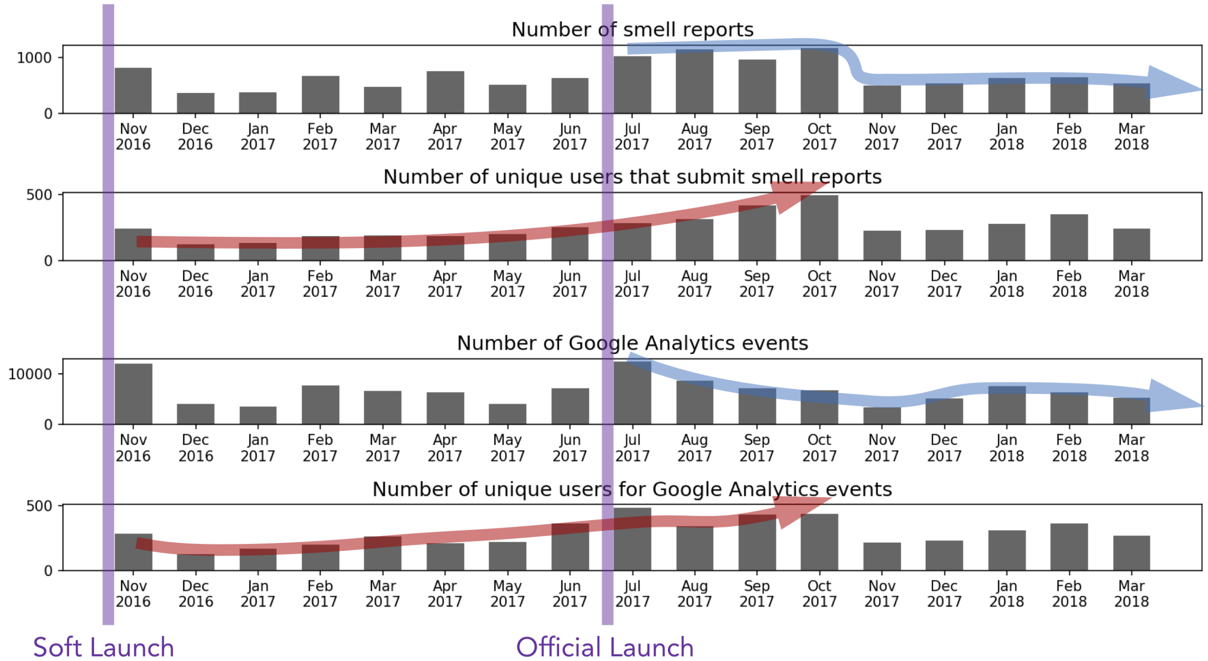


Figure 6.2: The distribution of submitted smell reports, Google Analytics events, and unique users over month. Although our users grew over 11 months (red arrows) after the soft and official launch (purple bars), there was a decrease in engagement recently (blue arrows).

new machine learning model every Sunday night based on the data collected previously. Details of this prediction task will be discussed in section 6.4.2.

6.4 Evaluation

To evaluate citizen participation, we showed that using odor experience was practical and scalable for revealing urban air quality concerns. We now discuss three studies: (1) system usage patterns of citizen-contributed smell reports and interaction events, (2) assessment of crowdsourced data validity by using statistical prediction and inference, and (3) survey of attitude changes and motivation factors.

6.4.1 System Usage Study

In this study, we evaluated the usage patterns on mobile devices by parsing server logs and Google Analytics events. From our initial testing with the community on September 2016 to the end of March 2018, we had 2,064 and 849 installations of Smell Pittsburgh on iOS and Android devices respectively in the United States. We excluded data generated during the system stability testing phase in September and October 2016. From our soft launch in November 2016 to the end of March 2018 over 17 months, there were 2,858 unique anonymous users in the Pittsburgh region. Our users contributed 11,700 smell reports, 383,767 alphanumeric characters in the

Table 6.1: Statistics of different user groups

	% of users	% of smell reports	% of characters in text field	% of events from Google Analytics
Enthusiasts	47%	91%	94%	76%
Contributors	12%	9%	6%	N/A
Observers	41%	N/A	N/A	24%
Total	100% (N=2,858)	100% (N=11,700)	100% (N=383,767)	100% (N=114,899)

Table 6.2: Statistics and Mann-Whitney U test results of variables among user groups

	Submitted reports ∀ user	Interaction events ∀ user	Characters in text fields ∀ report	Hours between hit and data timestamps ∀ event
Enthusiasts	Mdn=3 (n=1,345)	Mdn=20 (n=1,345)	Mdn=14 (n=10,692)	Mdn=11.5 (n=84,150)
Contributors	Mdn=1 (n=343)	N/A	Mdn=10 (n=1,008)	N/A
Observers	N/A	Mdn=9 (n=1,170)	N/A	Mdn=28.5 (n=25,718)
Test Result	p<.001 (U=303,087)	p<.001 (U=1,032,434)	p<.001 (U=5,884,441)	p<.001 (U=851,002,633)

Abbreviations "Mdn" and "n" indicate median and sample size respectively. Symbol ∀ represents "for each".

submitted text fields, and 114,899 events of interacting with the visualization (e.g., clicking on icons on the map). Among all smell reports, 75% of them had ratings larger or equal than three.

Data aggregated by month showed that our user engagement grew after a year following the soft launch and decreased noticeably afterward, as shown in Figure 6.2. There were two spikes of Google Analytics events in November 2016 and July 2017. These spikes corresponded to our soft and official launches of Smell Pittsburgh, which received widespread media coverage. After the soft launch, the number of our users grew four-fold over 11 months, from December 2016 to October 2017. However, there was a decrease in citizen engagement recently, and the number of users declined by more than half by November 2017, which was four months after the official launch.

To investigate the distribution of smell reports and interaction events among our users, we divided all users into three types: enthusiasts, contributors, and observers (Table 6.1). Contributors were those who submitted smell reports but did not interact with the visualization. Observers were those who interacted with the visualization but did not submit reports. Enthusiasts participated in both submitting reports and interacting with the visualization. We were interested in four variables with different distributions among user groups, which represented their characteristics (Figure 6.3). First, for each user, we computed the number of submitted smell reports and interaction events. Then, for each smell report, we calculated the number of alphanumeric characters in the submitted text fields. Finally, for interaction events that involved investigating previous data, we computed the time difference between hit timestamps and data timestamps. These two timestamps represented when users interacted with the system and when the data were archived in the system respectively. All variables differed from a normal distribution (normality test p<.001). Thus, to determine if there were significant differences among groups, we applied two-tailed Mann-Whitney U test (a nonparametric test for two independently sampled

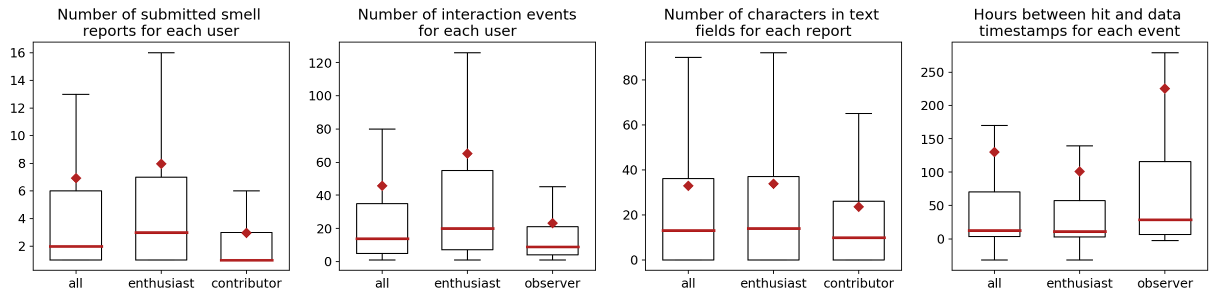


Figure 6.3: The box plots show distributions of different variables among user groups. The red lines in the middle of the box indicate the median ($Q2$). The red-filled diamonds represent the mean. The top and bottom edges of a box indicate 75% ($Q3$) and 25% ($Q1$) quantiles respectively. The boxes represent inter-quantile ranges $IQR = Q3-Q1$. The top and bottom whiskers show $Q3 + 1.5 * IQR$ and $Q1 + 1.5 * IQR$ respectively. This plot excludes outliers that are beyond the range of whiskers.

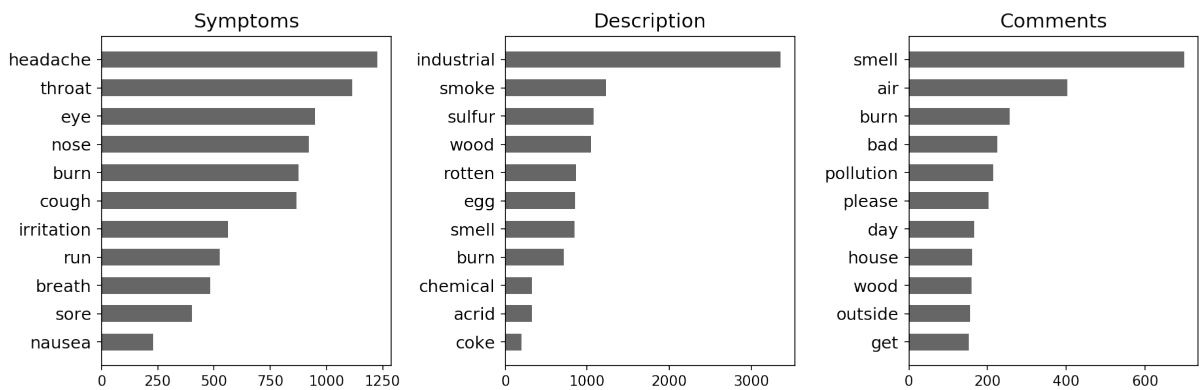


Figure 6.4: The frequency of words in different text fields of all submitted smell reports. Most of the high frequency words describe industrial pollution odors and related symptoms, especially hydrogen sulfide (rotten egg smell).

groups) and reported the results in Table 6.2.

From the user group study, we found that user contributions were highly skewed (Figure 6.3). Approximately 34% of the users submitted only one report, and 50% of the users submitted less than three reports, which aligned with the typical pattern in citizen science projects that many volunteers participated for only a few times [194]. Moreover, these three user groups differed regarding the type and amount of data they contributed. Table 6.1 shows that enthusiasts, corresponding to less than half of the users, contributed 91% smell reports and 76% interaction events. Table 6.2 indicates that all four variables were statistically significant among user groups. Enthusiasts tended to contribute more smell reports, the number of alphanumeric characters of reports, and interaction events. Observers tended to browse data that were far away from the interaction time. Also, by further investigating the enthusiast group, we found a moderate positive association (Pearson correlation coefficient r=.51, n=1,345, p<.001) between the number of submitted smell reports and the number of user interaction events.

To identify critical topics in citizen-contributed smell reports, we analyzed the frequency of words in three text fields. Before counting words, we used python NLTK package [21] to remove

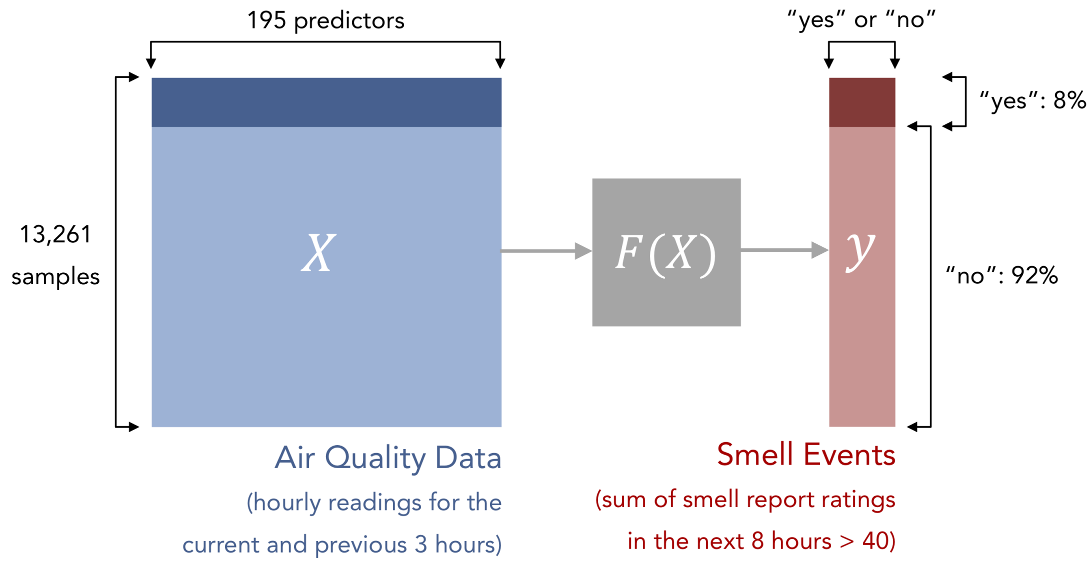


Figure 6.5: To enable odor prediction, we used machine learning techniques to estimate a function that maps air quality data (predictor matrix X) to smell events (response vector y).

stop words and group similar words with different forms (lemmatization). Figure 6.4 shows that high-frequency words mostly described industrial pollution odors and related symptoms, especially hydrogen sulfide that has rotten egg smell and can cause a headache, dizziness, eye irritation, sore throat, cough, nausea, and shortness of breath [55, 97, 151, 186]. This finding inspired us to examine how hydrogen sulfide affected urban odors in the next study.

6.4.2 Data Validity Study

The standardized regulatory procedure to assess air quality, which focuses on generating expert knowledge about average concentrations over long periods, naturally tends to resist lay knowledge that focuses on short-term events of sudden increases in air pollution readings [170]. In this study, we applied machine learning to show that the short-term events identified from crowd-sourced anonymous smell reports, when linked to air quality sensor measurements, could contribute to drawing meaningful insights for local air pollution concerns. Mathematically, machine learning involves a set of methods to approximate a function F such that $y = F(X)$, where y is the response vector and X is the predictor matrix. There are two main reasons for estimating the function F: prediction and inference [119]. We framed the odor prediction and inference tasks as a supervised learning problem to infer the function F that mapped inputs to an output based on pairs of previous observations. For the prediction problem of forecasting future smell events, we treated the model as a black box and focused on increasing its performance. For the inference problem of understanding the relationship between odor reports and environmental measurements, we used a white box model with an exact form to explain its internal decision-making process. The implementation in this study was based on python scikit-learn package [175].

Dataset

The input predictor matrix X had size n by m (Figure 6.5). The notation n meant the number of observations and m meant the number of predictor variables, which was also called "features". The input predictor matrix X consisted of hourly-recorded time-series data collected from government-operated air quality monitoring stations at different locations in Pittsburgh, such as Lawrenceville, Liberty, Flag Plaza, Parkway East, Avalon, North Braddock, and Glassport. Rows of matrix X represented observations, which could be viewed as data points in a high dimensional space formed by predictor variables. Columns of matrix X represented current and previous p hours of sensor readings, such as particulate matters, sulfur dioxide, carbon monoxide, nitrogen oxides, ozone, hydrogen sulfide, and wind information (direction, speed, and standard deviation of direction). Wind directions were decomposed into cosine and sine components. We considered p as a dataset parameter and set it to 3 by using cross-validation that will be described later. To equalize the effect of predictors and stabilize the time-series data, we normalized each column of matrix X to zero mean and unit variance. We also replaced missing values with mean values of the corresponding predictors. Finally, we added days of the week, hours of the day, and days of the month into the predictor matrix, which resulted in 195 features.

The output response vector y had size n by 1 (Figure 6.5), which was also called "labels". Vector y contained observations about whether a smell event would occur in the future q hours or not. The occurrence of events was represented by binary class labels 0 and 1, where 1 meant "yes" and vice versa. The i^{th} observation of vector y corresponded to the i^{th} row of the predictor matrix X. Because the pollution sources were undetermined, it was not feasible to obtain the "ground truth" labels. Thus, smell events were defined by majority agreements of our users. We specifically chose the geographic regions that have sufficient amount of data when computing smell events, as shown in Figure 6.9. To obtain the binary class labels, we first aggregated smell report data from the selected geographic regions of the coming q hours into an odor value, which was a weighted combination of smell reports with ratings greater than or equal to 3. For instance, if there were 10 smell reports with rating 4 in the next q hours, the odor value would be 40. In this study, we assigned a fixed value 8 to q for reducing parameters. Then, if the odor value was larger than a threshold r, its binary class label of the occurrence of a smell event was "yes", and vice versa. We considered r as another dataset parameter and set it to 40 by using cross-validation that will be discussed later. The dataset was highly imbalanced, where only 8% of the labels were positive.

Predicting Smell Events

Figure 6.6 and 6.7 visualized the dataset by using Principal Component Analysis (PCA) [120] and PCA with a radial basis function kernel [195] respectively. To model the relationships between the input predictor matrix X and output response vector y, we implemented two ensemble-based models, Extremely Randomized Trees [91] and Random Forests [31]. These algorithms build a collection of decision trees using the CART algorithm [149], where the leaves represent the binary class label and the branches represent the logical conjunction of predictors. The ensemble method [65] is effective in reducing model variance and sensitivity of overfitting, which are both problematic for individual decision trees [31, 91, 103]. When splitting a tree node during

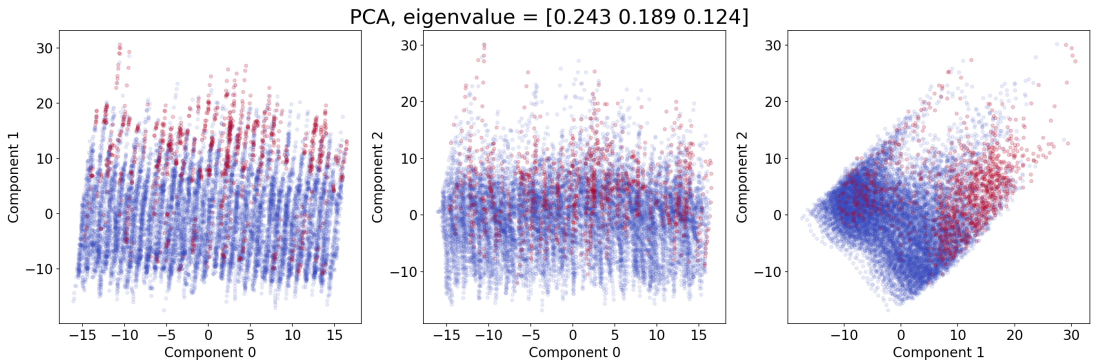


Figure 6.6: Principal Component Analysis. Blue and red dots indicate negative (without smell event) and positive labels (with smell event) respectively.

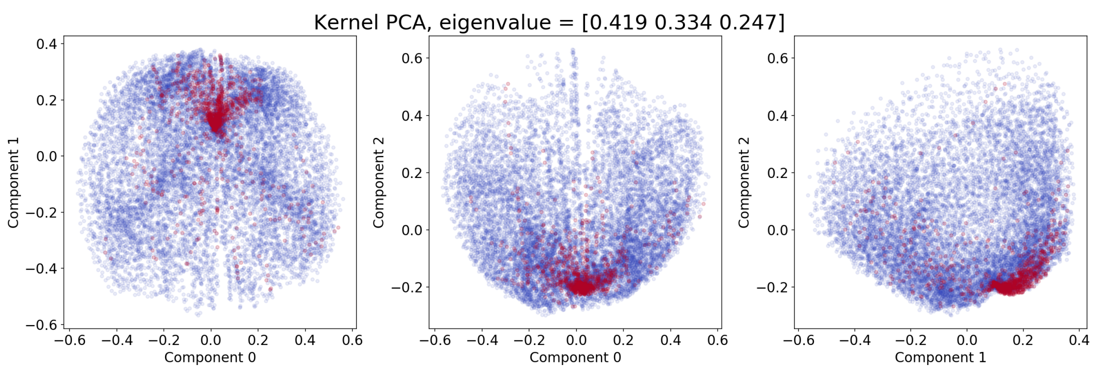


Figure 6.7: Principal Component Analysis with a radial basis function (RBF) kernel. Blue and red dots indicate negative (without smell event) and positive labels (with smell event) respectively.

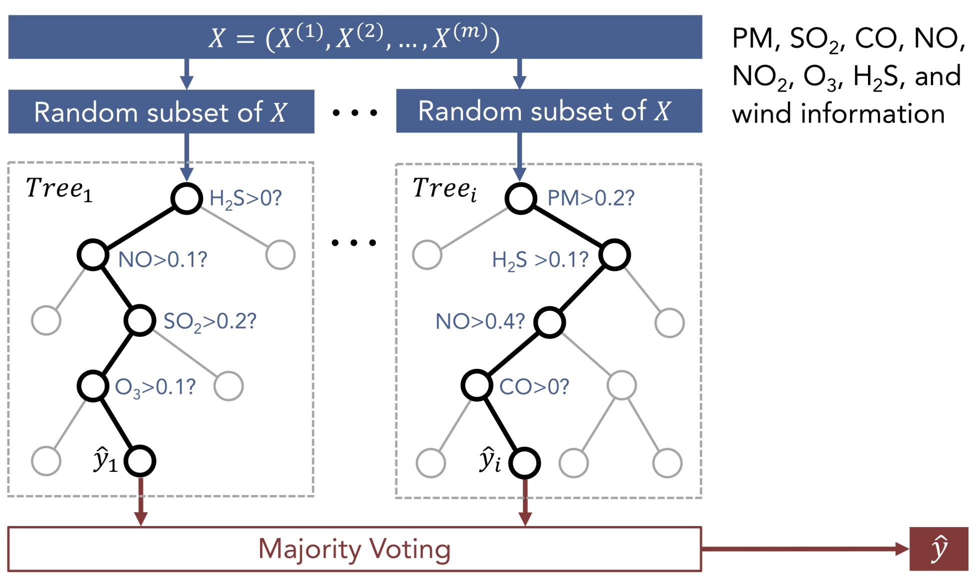


Figure 6.8: We used ensemble-based models, a collection of Decision Trees, to predict smell events ($\hat{y}$) by using air quality data (X).

the training process, Random Forests compute the optimal cut-point values from a subset of randomly chosen predictors based on bootstrap samples [78], while Extremely Randomized Trees use random cut-point values and the whole original samples. Also, our data contained highly correlated features. The randomization process when splitting nodes makes the model robust to correlated time-series predictor variables [87]. Finally, the prediction result is aggregated by a majority vote of all trees. There were three tunable model parameters: the number of trees u in the model, the number of features v to select randomly for splitting a tree node, and the minimum number of samples w required to split a tree node. We reduced these parameters to only v and w by always using 1,000 trees in both models.

Selecting Dataset and Model Parameters

To evaluate model performance, we used F-score [180], with its best value at 1 and worst value at 0. We first merged consecutive positive samples to compute the starting and ending time of smell events. Then, if a predicted event overlapped with a crowdsourced event, we counted this event as a true positive (TP). Otherwise, we counted a non-overlapped predicted event as a false positive (FP). For crowdsourced events that had no overlapping predicted events, we counted them as false negatives (FN). Figure 6.11 shows examples of TP, FP, and FN. Finally, we computed the precision, recall, and F-score by using the following equations:

$$\begin{aligned} \text{Precision} &= \text{TP}/(\text{TP}+\text{FP}) \\ \text{Recall} &= \text{TP}/(\text{TP}+\text{FN}) \\ \text{F-score} &= 2 * \text{Precision} * \text{Recall}/(\text{Precision}+\text{Recall}) \end{aligned} \tag{6.1}$$

When computing these metrics, we considered only daytime events because people rarely submitted smell reports during nighttime (Figure 6.10). We defined daytime from 5 am to 7 pm. Because the model predicted if a smell event would occur in the next 8 hours, we only needed to evaluate the prediction generated from 5 am to 11 am.

The method that we used for choosing parameters was time-series cross-validation [12, 133], where the entire dataset was partitioned and rolled into several pairs of training and testing subsets for evaluation (Figure 6.12). This method was different from the traditional cross-validation for time-independent data. Because our predictors and responses were all time-dependent, we used previous samples to train the models and evaluated them on future data. We first divided all samples (from October 9th 2016 to April 15th 2018) into 79 folds, with each fold approximately representing a week. Then, starting from fold 49, we took the previous 48 folds as training data (about 8,000 samples) and the current fold as testing data (about 168 samples), which resulted in 31 iterations for computing evaluation metrics. This procedure reflected the setting of the deployed system, where we trained a new model from scratch on every Sunday night by using previous 8,000-hour data.

We performed a two-stage grid search from a set of values with cross-validation to select dataset parameters (p, r) and model parameters (v, w). During model selection, there was a trade-off between precision and recall. We preferred the model that made fewer false predictions instead of forecasting all possible events. Therefore, for parameter sets that had similar F-scores, we selected the set that had the least number of false positives (highest precision). At the first stage, we only searched dataset parameters (p, r) with fixed model parameters $(v, w) = (\sqrt{m}, 2)$

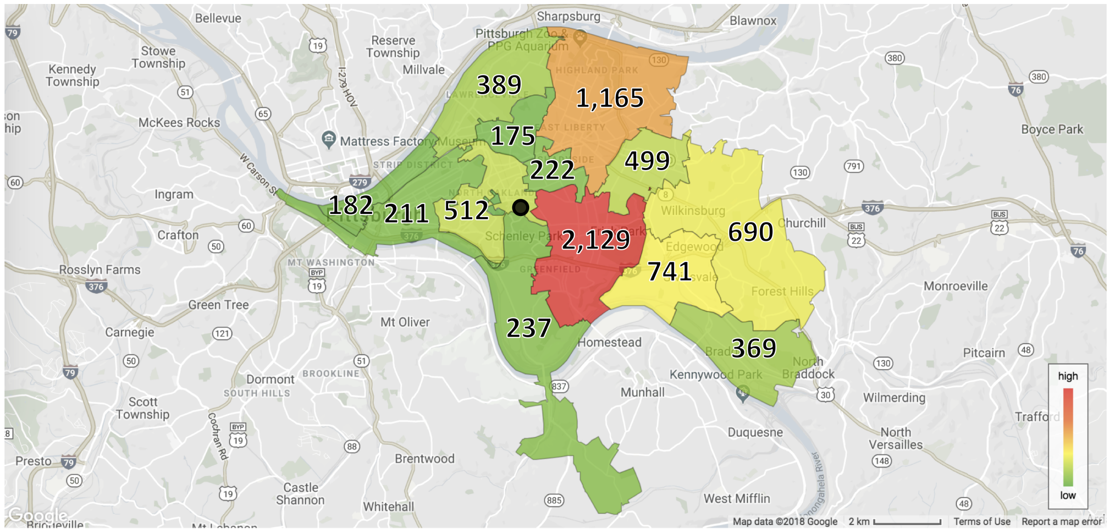


Figure 6.9: The distribution of smell reports geographically on selected zip code regions from October 9th 2016 to April 15th 2018. The integers on each zip code region indicate the number of smell reports. The black dot shows the location of Carnegie Mellon University.

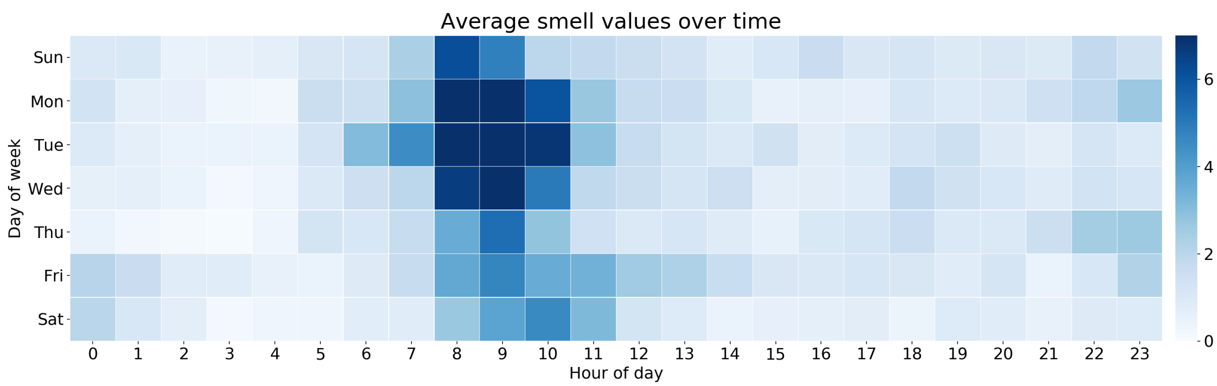


Figure 6.10: The average smell values aggregated by hour of day and day of week. This figure shows that our users rarely submit smell reports at nighttime.

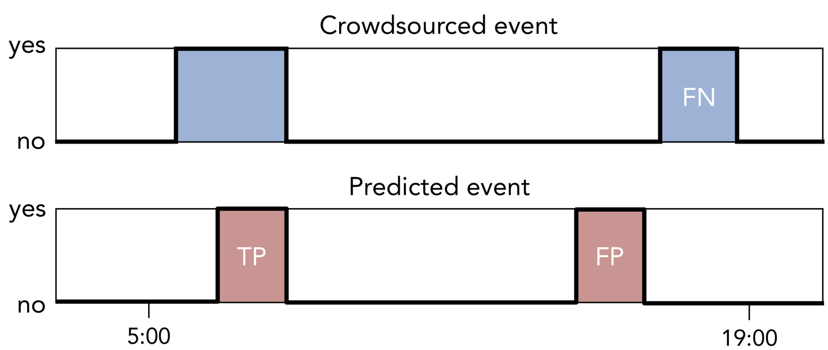


Figure 6.11: This figure shows the original and predicted smell events. The x-axis represents time. The blue and red boxes indicate crowdsourced and predicted smell events respectively. Abbreviations TP, FP, and FN mean true positives, false positives, and false negatives respectively.

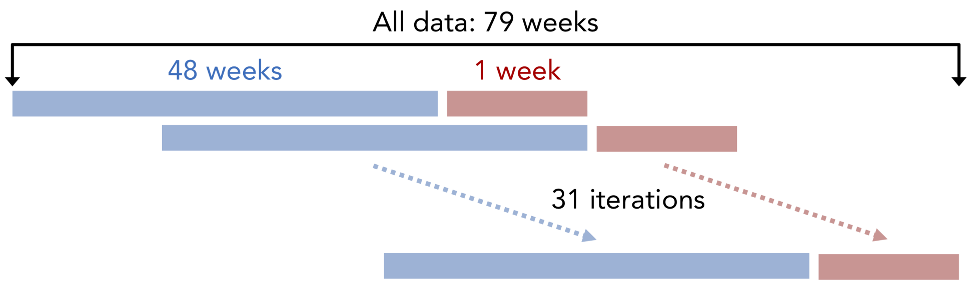


Figure 6.12: The entire dataset was partitioned and rolled into several pairs of training and testing subsets for cross-validation.

Table 6.3: Cross-validation result of models for statistical prediction

	TP	FP	FN	Precision	Recall	F-score
ExtraTrees	28.65 ± 0.48	3.99 ± 0.72	14.35 ± 0.48	0.88 ± 0.02	0.67 ± 0.01	0.76 ± 0.01
Random Forest	28.84 ± 0.60	5.86 ± 0.68	14.16 ± 0.60	0.83 ± 0.02	0.67 ± 0.01	0.74 ± 0.01
Always Yes	43	170	0	0.20	1.00	0.34

The cell format is "mean $\pm$ standard deviation". We run this experiment on 31 weeks of testing data for 100 times. Abbreviations "ExtraTrees", "TP", "FP", and "FN" indicates Extremely Randomized Trees, true positives, false positives, and false negatives respectively. The last model, "Always Yes", indicates the baseline that always makes positive predictions.

Table 6.4: Cross-validation result of models for statistical inference

	TP	FP	FN	Precision	Recall	F-score
Decision Tree	16.81 ± 1.32	6.63 ± 1.55	4.18 ± 1.52	0.72 ± 0.05	0.80 ± 0.07	0.76 ± 0.04

The cell format is "mean $\pm$ standard deviation". We run this experiment on 31 weeks of testing data for 100 times. Abbreviations "TP", "FP", and "FN" indicates true positives, false positives, and false negatives respectively.

suggested by the Extremely Randomized Trees paper [91], where m was the number of features. The dataset parameters with the best cross-validated F-score was $(p, r) = (3, 40)$. To verify if these parameters were reasonable, we applied them to compute the dataset and plotted the time-lagged point-biserial correlation coefficients between continuous predictors and binary responses (Figure 6.13). All coefficients having more than 3-hour time lag were less than 0.3. Next, we fixed the dataset parameters for searching better model parameters (v, w). The best parameters for the Extremely Randomized Trees and Random Forests were $(v, w) = (90, 32)$ and $(v, w) = (30, 2)$ respectively. Table 6.3 reports the evaluation metrics after cross-validating these two models for 100 times with various random seeds on 31-week testing data. The result showed that the performance is better than the baseline, which is a model that always makes positive predictions.

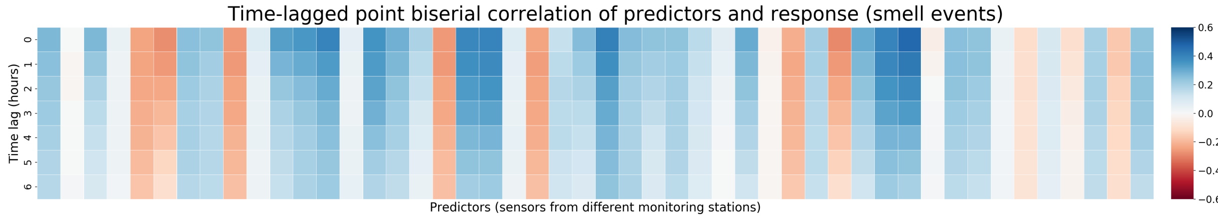


Figure 6.13: The time-lagged point-biserial correlation of continuous predictors (sensor readings from different monitoring stations) and binary response (smell events). Top five predictors with highest correlations are particulate matter at Glassport (r=.47, n=13,264, p<.001) and Liberty (r=.40, n=13,264, p<.001), carbon monoxide at Flag Plaza (r=.41, n=13,264, p<.001) and Lawrenceville (r=.40, n=13,264, p<.001), and hydrogen sulfide at Liberty (r=.36, n=13,264, p<.001). None of the correlation coefficients exceed 0.5.

Interpreting Smell Events

While these two ensemble-based models enabled us to predict future events, they were typically considered as black box models and not suitable for statistical inference. Although these two models provided feature importances, interpreting these weights could be problematic because several predictors in the dataset were highly correlated, which might appear less significant than other uncorrelated counterparts. Inspired by several previous works related to extracting knowledge from data [42, 89, 198], we utilized a white box model, Decision Tree, to explain a representative subset of predictors and samples (Figure 6.14), which were selected by applying feature selection [99] and cluster analysis. First, we used domain knowledge to manually select features. Based on the knowledge obtained from several informal community meetings and the result discovered in the text analysis of smell reports (Figure 6.4), we chose hydrogen sulfide, wind direction, wind speed, and standard deviation of wind direction from all monitoring stations. The current and up to two-hour time lagged readings were all included. Also, we added interaction terms of all predictors, such as hydrogen sulfide multiplied by the sine component of wind direction. This manual feature selection procedure produced 781 features.

Next, we applied a clustering algorithm, DBSCAN [83], to choose a representative subset of samples. The parameters for DBSCAN were minimum 30 samples within a 0.7 Eps-neighborhood distance for a point to be defined as a core sample, which indicated the density of a cluster. The distance matrix D for clustering was derived from a Random Forest fitted on the manually selected features. The Random Forest parameters were $(v, w) = (0.15 * m, 2)$, where m was the number of features. For each pair of samples, we counted the number of times that they appeared in the same leaf of all trees in the model. The results were assembled into a similarity matrix S and normalized to the range between 0 and 1. We converted the similarity matrix into a distance matrix by using $D = 1 - S$. This procedure identified a cluster with about 25% of the total 1,069 positive samples. The cluster represented about 50% of the total 46 events, with each event indicating consecutive positive samples, as shown in Figure 6.11. To focus on interpreting this cluster, we set all positive samples outside the cluster to negative samples.

Furthermore, we performed recursive feature elimination (RFE) by iteratively removing features that had smaller weights [100]. These feature importance weights represented the mean decrease impurity [153] of a Random Forest with parameters $(v, w) = (\sqrt{m}, 2)$, where m was the number of features. Parameters for RFE include eliminating 50 features for each iteration and

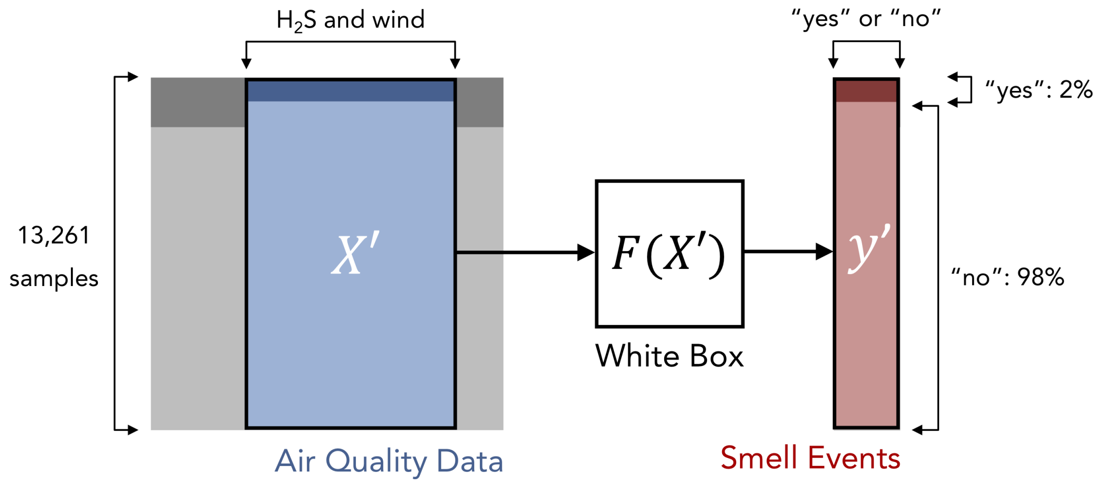


Figure 6.14: We used a Decision Tree (white box model) to explain a subset of predictors and positive samples, which was selected by applying community knowledge and cluster analysis.

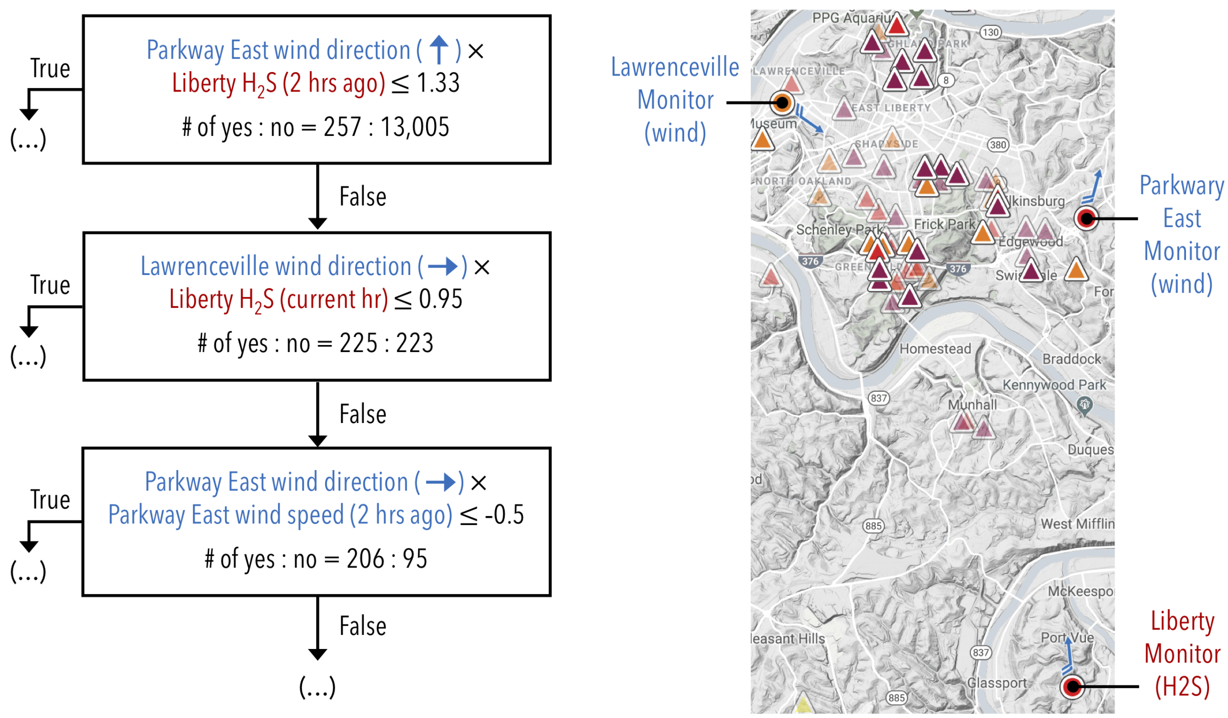


Figure 6.15: The right terrain map shows smell reports and sensor readings at 10:30 am on December 3rd, 2017. Important predictors are marked on the map. The left graph shows a part of the Decision Tree model for interpreting patterns with F-score 0.81. For simplification, only the first three depth levels of the tree are plotted. This model explains the pattern of about 50% smell events, which contain the interactions of hydrogen sulfide and wind information from different monitoring stations. The first two lines of a tree node shows the corresponding feature and its threshold for splitting. The third line of a tree node indicates the ratio of the number of positive samples (with smell event) and negative samples (no smell event). The most important predictor is the interaction between the sine component of wind directions at Parkway East and the previous 2-hour hydrogen sulfide readings at Liberty (r=.62, n=13,262, p<.001). The second most important predictor is the interaction between the cosine component of wind directions at Lawrenceville and the hydrogen sulfide readings at Liberty (r=.45, n=13,262, p<.001). Notation "r" means the point-biserial correlation of the predictor and smell events.

Table 6.5: Demographics of participants (ages and education levels)

	18-24	25-34	35-44	45-54	55-64	65-74	Total
Associate's	0	0	1	0	0	0	1
Bachelor's	2	2	2	0	1	1	8
Master's	0	2	2	0	0	4	8
Doctoral	0	1	1	1	5	0	8
Total	2	5	6	1	6	5	25

Table 6.6: Frequency of system usage (sorted by percentage)

	Count	Percentage
Other (the open-response text field)	9	36%
At least once per month	7	28%
At least once per week	4	16%
At least once per day	3	12%
At least once per year	2	8%

Table 6.7: Choices for measuring participation level (sorted by percentage)

	Count	Percentage
I submitted smell reports.	22	88%
I checked other people's smell reports on the map visualization.	22	88%
I opened Smell Pittsburgh when I noticed unusual smell.	22	88%
I discussed Smell Pittsburgh with other people.	21	84%
I provided my contact information when submitting smell reports.	14	56%
I paid attention to smell event alert notifications provided by Smell Pittsburgh.	13	52%
I shared Smell Pittsburgh publicly online (e.g. email, social media, news blog).	13	52%
I clicked on the playback button to view the animation of smell reports.	9	36%
I took screenshots of Smell Pittsburgh.	9	36%
I mentioned or presented Smell Pittsburgh to regulators.	6	24%
I downloaded smell reports data from the Smell Pittsburgh website.	4	16%

selecting 30 most important features at the final iteration. Finally, we trained a Decision Tree using the CART algorithm [149] to interpret the cluster and the selected features. The parameters for the Decision Tree model were minimum 5 samples for being a leaf node, minimum 20 samples for splitting a node, and maximum depth 8 of the tree. All parameters for data interpretation (DBSCAN, Random Forest, RFE, and Decision Tree) were selected by using cross-validation. Table 6.4 reports the evaluation metrics after cross-validating the model for 100 times with various random seeds on 31-week testing data. The result showed that the model was capable of explaining the underlying pattern of about 50% of the smell events, which was a joint effect of wind information and hydrogen sulfide readings (Figure 6.15).

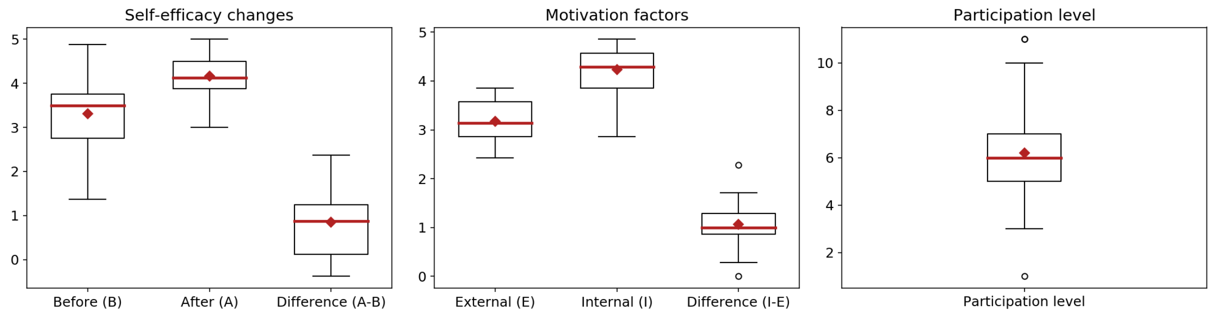


Figure 6.16: The box plots show distributions of self-efficacy changes, internal and external motivations, and participation level for 25 valid survey responses. The red lines in the middle of the box indicate the median ($Q2$). The red-filled diamonds represent the mean. The top and bottom edges of a box indicate 75% ($Q3$) and 25% ($Q1$) quantiles respectively. The boxes represent inter-quantile ranges $IQR = Q3 - Q1$. The top and bottom whiskers show $Q3 + 1.5 * IQR$ and $Q1 + 1.5 * IQR$ respectively. Black hollow circles show outliers that are beyond the range of whiskers.

6.4.3 Survey Study

When developing computational tools to support citizen science, it is essential to evaluate attitude changes or motivations to inform and reflect upon system design [113, 191]. Many crowdsourcing or citizen science projects have found it informative to characterize the motivations of their user bases [10, 11, 56, 184, 221, 230]. We developed a survey (described below) to measure the motivations and self-efficacy of Smell Pittsburgh users. We define self-efficacy as beliefs about how well an individual can achieve desired effects through actions [15].

Participants

We recruited adult participants via snowball sampling, as described by [19]. We delivered an anonymous online survey via email to community advocacy groups and asked them to distribute the survey to potential participants. Paper surveys were also provided. All responses were kept confidential, and there was no compensation. We received 29 responses in total over one month from March 20th to April 20th, 2018. Four responses were excluded due to incomplete questions or no experiences in interacting with the system, which gave 25 valid survey responses. There were 8 males, 16 females, and 1 person with undisclosed gender information. All but one participant had a Bachelor's degree at minimum. The demographics of the sample population (Table 6.5) were not typical for the region.

Procedures and Materials

We administered a single survey to people who had used Smell Pittsburgh since its release. The survey had three sections: (1) Self-Efficacy Changes, (2) Motivation Factors, (3) System Usage Information. These last two sections were also complemented by open-response text fields where participants provided additional comments and messages.

For Self-Efficacy Changes, we measured changes to user confidence mitigating air quality problems. This section was framed as a retrospective pre-post self-assessment. The items were divided between pre-assessment, “BEFORE you knew about or used Smell Pittsburgh,” and post-assessment, “AFTER you knew about or used Smell Pittsburgh.” For both assessments, we used a scale developed by the Cornell Lab of Ornithology to measure self-efficacy in citizen science projects [59, 179]. The scale was customized for air quality to suit our purpose. The scale consisted of eight Likert-type items (from 1 “Strongly Disagree” to 5 “Strongly Agree”).

The Motivation Factors section was based on a scale developed by the Cornell Lab of Ornithology [59, 178] with 14 Likert-type items (from 1 “Strongly Disagree” to 5 “Strongly Agree”). The scale was customized for air quality and measured both internal (7 items) and external factors (7 items). Internal factors involved enjoyment during participation and the desire to achieve the goal of improving air quality. On the other hand, external factors involved obtaining rewards and avoiding negative consequences if not taking actions. A text field with question “Are there other reasons that you use Smell Pittsburgh?” was provided for open responses.

In the System Usage Information section, we collected individual experiences with Smell Pittsburgh. We documented participation level through a multiple-choice and multiple-response question, “How did you use Smell Pittsburgh?” as shown in Figure 6.16 (right). This question allowed participants to select from a list of 11 activities. We identified the frequency of system usage through a multiple-choice question, “How often do you use Smell Pittsburgh?” as shown in Table 6.6. Text fields were provided for both of the above two questions.

Finally, we asked an open-response question “Do you have any other comments, questions, or concerns?” at the end of the survey. Our analysis of these responses is presented below in conjunction with each related question.

Analysis and Results

For Self-Efficacy Changes, we averaged the scale items to produce total self-efficacy pre score (Mdn=3.50) and post score (Mdn=4.13) for each participant. A two-tailed Wilcoxon Signed-Ranks test (a nonparametric version of a paired t-test) indicated a statistically significant difference ($W=13.5$, $Z=-3.79$, $p<.001$), as shown in Figure 6.16 (left). This finding indicated that there were increases in self-efficacy during participation.

For Motivation Factors, we computed the average score of internal (Mdn=4.29) and external (Mdn=3.14) motivation scores for each participant. A two-tailed Wilcoxon Signed-Ranks test indicated a statistically significant difference ($W=0$, $Z=-4.29$, $p<.001$), as shown in Figure 6.16 (middle). This result suggested that internal factors were primary motivations for our participants rather than external factors. Open-ended answers showed that 12 participants (48%) were motivated by the technical affordance provided by the system. Among them, nine participants (36%) mentioned the affordance to contribute data as scientific evidence efficiently and intuitively, as shown in the following selected quotes. Bold emphases were added by researchers to highlight key user sentiments.

> “I used to try to use the phone to call in complaints, but that was highly unsatisfactory. I never knew if my complaints were even registered. With Smell Pittsburgh, **I feel that I’m contributing to taking data**, as well as to **complaining when it’s awful**. [...]”

> “It’s seems to be the most **effective way to report wood burning** that can fill my neighborhood with the smoke and emissions from wood burning.”

> "The Smell app **quantifies observations in real time.** Researchers can use this qualitative information along quantitative data in real time. Added benefit is to have [the health department] receive this information in real time **without having to make a phone call or send separate email**. I have confidence that the recording of Smell app data is **quantified more accurately** than [the health department]'s."

> "It is an **evidence based way** for a citizen to register what is going on with the air where I live and work."

> "I believe in science and data and think **this can help build a case**. [...]"

Additionally, four participants (16%) indicated the affordance to validate personal experiences based on the data provided by others. Selected quotes were shown below.

> "I used to (and sometimes still do) call reports in to [the health department]. I love how the map displays after I post a smell report. **Wow! I'm not alone!**"

> "It **validates my pollution experiences** because others are reporting similar experiences."

> "I like using it for a similar reason that I like checking the weather. It helps me understand my environment and **confirms my sense of what I'm seeing** (or in this case smelling)."

We also found that altruism, the concern about the welfare of others, was another motivation. Six participants (24%) mentioned the desire to address climate changes, activate regulators, raise awareness of others, expand air quality knowledge, influence policy-making, and build a sense of community. Selected quotes were presented:

> "Because **climate change** is one of our largest challenges, [...] Also, the ACHD isn't as active as they should be, and **needs a nudge**."

> "I use [Smell Pittsburgh] to **demonstrate to others how they can raise their own awareness**. I've also pointed out to others that many who have grown up in this area of Western PA have grown up with so much pollution, **to them air pollution has become normalized** and many do not even smell the pollution any more. This is extremely dangerous and disturbing."

> "I want to help **expand the knowledge and education of air quality** in Pittsburgh and believe the visuals Smell Pittsburgh provides is the best way to do that."

> "I believe in the power of crowd-sourced data to **influence policy decisions**. I also believe that the air quality activism community will find more willing participants if there is a **very easy way for non-activists to help support clean air**, and the app provides that mechanism. It is basically a very **easy onramp for potential new activists**. The app also acts as a way for non-activists to see that they are not alone in their concerns about stinky air, which I believe was a major problem for **building momentum in the air quality community** prior to the app's existence."

For System Usage Information, we reported the counts for system usage frequency questions (Table 6.7). The result showed that our users had a wide variety of system usage frequency. Open-responses indicated that instead of using the system regularly, eight participants (32%) only submitted reports whenever they experienced poor odors. To quantify participation levels, we counted the number of selected choices for each participant, as shown in Figure 6.16 (right). We also counted the number of participants who selected each choice in the "How did you use Smell Pittsburgh?" question (Table 6.6). We found that our participation levels were normally distributed. In the open-response text field for this question, two participants (8%) mentioned using personal resources to help promote the system.

> "I **ran a Google Adwords campaign to get people to install Smell Pittsburgh**. It turns out that about $6 of ad spending will induce someone to install the app."

> "I take and share so many screenshots! Those are awesome. [...] I also **made two large posters of the app screen**– one on a very bad day, and one on a very good day. I **bring them around to public meetings** and try to get county officials to look at them."

In the open-ended question to freely provide comments and concerns, two participants (8%) were frustrated about the lack of responses from regulators and unclear values of using the data to take action, as shown in the following quotes.

> "After using this app for over a year, and making many dozens of reports, I **haven't once heard from the [health department]**. That is disappointing, and makes me wonder, why bother? [...] Collecting this data is clever, but towards what end? I sometimes **don't see the point in continuing to report**."

> "**It wasn't clear when using the app that my submission was counted**, so it made me feel like the work I did was useless. I want to be able to see directly that my smell reports are going somewhere and being used for something. [...]"

Also, five (20%) participants suggested augmenting the current system with new features and offering this mobile computing tool to more cities. Such features involved reporting smell at a different location and time, viewing personal submission records, and earlier predictive push notifications about odor events. The followings showed several quotes.

> "I **get around mostly by bike**, so it is difficult to report smells the same moment I smell them. I wish I could report smells in a different location than where I am so that I **could report the smell once I reach my destination**."

> "It would be nice to be able to **add a retroactive report**. We often get the strong sulfur smells in Forest Hills in the middle of the night [...] but I strongly **prefer to not have to log in to my phone at 3 am to log the report** as it makes it harder to get back to sleep."

> "This app should let me **see/download all of my data**: how many times I reported smells, what my symptoms and comments were and how many times the [health department] didn't respond [...]"

> "[...], right now [the predictions] are a little sparse and **often come without enough warning time** for me to plan my exercise around them."

6.5 Discussion

We have shown that Smell Pittsburgh, as a modern mobile computing tool, can equip citizens with the capability of collecting and visualizing a large amount of air quality data. The system collected 8,720 smell reports in 2017. This amount is 10-fold more than the 796 complaints collected by the health department regulators in 2016. Furthermore, all smell reports in our system had location data, while the location information was missing from 45% of the regulator-collected complaints. In the survey study, participants mentioned that the system enables them to contribute and communicate data-driven evidence for advocacy. Although the survey study was limited by the small sample size of total users, the result showed a statistically significant increase in self-efficacy after using the system. Several participants were even willing to use their resources to encourage others to install the system and engage in reporting odor events. These findings from the survey study, combined with the drastic increase in the amount of collected data, suggest that Smell Pittsburgh can lower the barrier and reduce the workload for citizens to participate in large-scale environmental epidemiology research. Although we showed the increase in data quantity, scientists may criticize the reliability of studies involving these crowdsourced data, since lay experiences may be prone to noise and exaggeration. For exaggerated smell reports, we view it as a significant sign of the need to expand environmental health protection about air quality issues, rather than as an argument to exclude these data in scientific research. We have demonstrated that using machine learning techniques to interpret these noisy data, even when crowdsourced anonymously, can provide informative insights to reveal the pattern of local environmental problems.

6.5.1 Implications

Based on our experiences in developing the system and the results presented in our evaluation studies, we now summarize our findings into three design implications.

Use qualitative data and domain knowledge to inform or explain quantitative analysis

Qualitative data, when collected and combined with domain knowledge, can be instrumental in informing or explaining quantitative analysis. In the system usage study, text analysis of the self-reported symptoms and smell descriptions revealed that hydrogen sulfide might be the primary source of odor events. This finding inspired us to choose hydrogen sulfide and wind information from all of the other available predictors, which was critical for the data validity study. As there were many highly correlated features, selecting a subset of features arbitrarily for interpreting patterns in the data was impractical. Moreover, open-ended questions in the survey study pointed out frustrations about lacking perceived values in using the collected data to advocate for policy changes. This finding provided a possible explanation for the decrease of citizen engagement discussed in the system usage study. Additionally, the system usage study identifies a moderate correlation between contributing smell reports and interacting with the visualization. This finding could be explained by the survey study that community members were motivated by the technical affordance of viewing data provided by others to validate personal experiences.

Consider prediction and inference when evaluating values of citizen science data

We encourage applying both statistical prediction and inference techniques to evaluate the value of citizen science data, which may not reveal meaningful information at first glance. Due to the nature of wicked problems in citizen science, the collected data may suffer from many types of bias and error that sometimes can even be unavoidable [22, 38]. Making sense of such noisy data has been a significant concern in citizen science research framework [165, 172], and machine learning offers potential techniques to address this concern [22, 108]. In the data validity study, we not only used ensemble-based black box models to predict the occurrence of odor events but also used a white box model, Decision Tree, to interpret the patterns. This approach enabled us to explain about 50% of the smell events, which was a joint effect of hydrogen sulfide and wind information. Our goal was not to show causality or statistical significance, but to present explainable insights about the impact of urban air quality problems that lay people and professionals would understand. The potential connections between predictors and responses, which were found from the statistical inference, could serve as hypotheses for future epidemiological studies.

Treat interactive systems for community citizen science as an ongoing infrastructure

To support environmental health advocacy, we believe that considering the entire community engagement life cycle is critical in developing information technology infrastructure and estimating how many resources are needed to accomplish goals. For initiating engagement, we developed a mobile application for crowdsourcing odor data and invited citizens through news media and our established network of advocacy groups. For maintaining engagement, we applied standard user interface design to visualize smell reports and added more weights to the system by sending these reports to the health department. Community members have utilized our publicly-released data and the visualization as evidence for taking action. For evaluating engagement, we conducted qualitative and quantitative studies to understand the impact of the system and patterns in the collected data. These three steps form a cycle in the sense that we can iterate through those steps multiple times to refine the system in response to user behavior changes.

6.5.2 Limitation

Maintaining citizen participation is not a trivial task. The push notification feature was designed for this purpose. However, we deployed the predictive model recently, and the sample size was insufficient for determining if push notifications were effective. Moreover, the system usage study indicated that less than 50% of the users contributed most of the data, and there was a decrease in engagement after the official launch of Smell Pittsburgh. Also, the survey study showed that several participants were frustrated about the stagnation of using crowdsourced smell data for advocacy. Future directions involve studying the effect of push notifications to engage community groups, exploring methods to activate regulators in pursuing policy changes, and establishing a feedback loop that shows efforts from both citizens and regulators.

We have explained the design, deployment, and evaluation of a mobile computing tool for Pittsburgh communities to collect and visualize odor events. However, our sample size (n=25) of total users (N=2,858) in the survey study was small, and the conclusion from the statistical analysis was weak. Additionally, our community members might be unique in their characteristics, such as the awareness of the air quality problem, the tenacity of advocacy, and the power relationships with other stakeholders. Involving citizens to address urban air pollution collaboratively is a wicked problem, and thus attempts to replicate our approach may be doubtful. It is possible that interactive systems like Smell Pittsburgh can only be practical for communities with specific characteristics, such as high awareness [113]. Whether this research can be fully replicated in other contexts is an open question. Future research is needed to study the impact of deploying this system in other cities that have similar or distinct community characteristics compared to Pittsburgh.

In the data validity study, there were still many odor events that the model was unable to predict (i.e., false negatives). Furthermore, 50% of the smell events remained unexplained. Recent research has shown that deep neural networks [147] usually outperforms other models. However, deep neural networks require a significant amount of data to make the performance compelling, and the number of crowdsourced smell reports has not reached such level. Future work involves adding more predictors (e.g., weather forecasts, air quality index), training deep neural networks for prediction, and using generalizable data interpretation techniques that can explain any predictive model to identify more patterns [188].

6.6 Summary

This chapter explores the design and impact of the intervention of a mobile computing tool, Smell Pittsburgh, that provides technological affordance to empower citizens in advocating for better air quality. We applied crowdsourcing to gather odor experiences from citizens without the support of professionals. Moreover, we used data visualization to present the context of air quality concerns from multiple perspectives as scientific evidence. In the evaluation, we identified the distribution and trend of smell reports and interaction events among different types of users. By adopting machine learning, we developed the push notification system feature and revealed patterns within the crowdsourced data. Using a survey, we studied motivation factors for submitting smell reports and measured user attitude changes after using the system. Based on the evaluation, we summarized findings into three insights: using qualitative data to support quantitative analysis, applying both statistical prediction and inference when evaluating data validity, and considering the community engagement life cycle when developing similar systems. Finally, we discussed limitations and future directions: studying the effect of push notifications, exploring methods to engage more users, deploying the system in other cities, and using advanced techniques for pattern recognition. We envision that this research can inspire engineers, designers, and researchers to develop computational tools that support advocacy and citizen empowerment.

Chapter 7

Conclusion

Community citizen science aims to representing community voices, addressing community concerns, and influence policy-making by using scientific knowledge. However, it is challenging for community members to collect and form reliable scientific knowledge due to the requirements of financial resources, organizational networks, and access to technology. This thesis has explored methods of using information technology to empower lay people in producing and sharing scientific knowledge. When power relationships among citizens and stakeholders are unbalanced and contradictory, community citizen science (as defined in chapter 1) plays a significant role to mitigate this situation. Modern computational tools grant lay people the autonomy to crowdsource, visualize, and share multiple types of human-generated and machine-generated data collaboratively. Moreover, computer vision and machine learning allow communities to extract scientific knowledge and evidence from the data, which are essential for communities to express their needs and concerns. Although crowdsourcing, visualization, and artificial intelligence techniques have received extensive attention and been widely adopted in commercial products, we are still at the beginning of integrating these techniques for common good. This dissertation believes that democratizing both computational tools and scientific knowledge is vital to empower communities and promote the welfare of human beings. The core research question, "**How can we design interactive systems with visualization, crowdsourcing, and artificial intelligence to support the engagement lifecycle in community citizen science?**" is answered through the design, deployment, and evaluation of four interactive systems. Each system makes methodological and empirical contributions to sustainable HCI (discussed in section 1.5), as briefly reviewed below.

- Chapter 3 introduced a timelapse editor for creating guided tours and interactive slideshows from cloud-free annual mosaics of satellite imagery. The tool was designed to make global imagery data transparent and meaningful through interactive visualization. By democratizing data that were typically accessible to only domain experts, the video tours generated from this tool enabled journalists to communicate critical global issues, such as climate change and urban expansion.
- Chapter 4 described an air quality monitoring system to reveal the local air pollution concern by integrating multiple types of data, which included images, smell reports, air quality measurements, and wind direction. This tool utilized a computer vision algorithm to gen-

erate video clips that displayed smoke emissions, which reduced the workload for community members to collect evidence of poor air quality. A survey study indicated increases in self-efficacy and sense of community after interacting with the system.

- Chapter 5 depicted a web-based tool for health professionals in a non-profit organization to visualize geographically aggregated community data, including air quality measurements, self-reported symptoms, and personal stories. The data were collected from residents who believed that their living quality had been affected by local industrial activities. A focus group study showed that it was essential to provide system features for understanding data, comparing patterns, and advocating for social changes.
- Chapter 6 discussed a mobile computing system for crowdsourcing odor complaints submitted by residents. These complaints were visualized and animated with sensor measurements from air quality monitoring stations. The visualization enabled community members to track how pollution odors travel in the city. The system used a machine learning model to predict the occurrence of poor odor events and send corresponding push notifications to users. A data validation study revealed a prevailing pattern that half of the odor events were related to a joint effect of hydrogen sulfide and wind directions. A survey study showed the prevalence of internal motivations and the increase of self-efficacy after using the system.

7.1 Design Implications

These four computational tools presented in the previous chapters differ in their contexts, users, scales, and goals. Despite their differences, these tools share similar design processes and principles under the general theme of community citizen science. Based on the experiences of developing and deploying these computational tools, this section summarizes all findings into generalizable implications for future researchers to design interactive systems that support community citizen science.

7.1.1 Co-design Interactive Systems with Communities

Community citizen science addresses problems and concerns that are deeply grounded in local regions. Community members have personal experiences and attachments to local concerns, and thus it is essential to treat them as **co-designers** [30], which bring diverse expertise and knowledge that researchers may not have. Designing computational tools to empower communities is highly iterative and reflective. It is also a two-way communication and knowledge exchange process between scientists and communities. This thesis adopts the value that scientists can think and act as citizens to gain and understand experiences of local concerns when designing interactive systems, as discussed in section 1.1.1. Understanding "What does the community actually need?" is fundamentally different from "What does the researcher think the community needs?" Similar to how architects and urban designers tackle wicked problems (discussed in section 1.3), the designers present an artifact (concept or prototype) to a group (peers, seniors, and targeting communities) and describe the rationality of design principles. The design is then open to constructive criticisms and questions to develop a shared understanding of how the artifact is embodied in its context. The entire design process is based on the engagement, discussion, and critiques of communities and stakeholders.

The design process that this thesis embraces is similar to architecture or urban design studios, where participants engage in developing ideas into artifacts or criticize the rationality of a proposed artifact. For instance, chapter 4 described the iterative participatory design process with community members. Researchers attended community meetings multiple times to present the system prototypes and ask community members to test system features. Feedback from residents was applied to refine the system design, and the community members were able to combine evidence provided by the system with personal stories to convince regulators. Another example is the focus group study described in chapter 5. Developers and stakeholders, who were involved in the study, pointed out critical problems and suggested useful future improvements of system features for another design iteration. Additionally, Smell Pittsburgh, mentioned in chapter 6, included community members in designing the scales for measuring the severity of odors. Regulators were also involved in the design process to strengthen the weight of citizen-contributed reports. Feedback from regulators and community members is essential for developing and deploying the mobile computing tool to collect a large quantity of data, which made it possible to study the relationships between pollutants and odor experiences.

7.1.2 Contextualize Scientific Evidence

From an epidemiological point of view, conducting causal inference studies in community citizen science is not practical. For example, it is not ethical to perform randomized experiments on how urban air pollution affects residents by manipulating emissions. Moreover, data generated from human and machine inputs with modern computational tools can be very high-dimensional and noisy, which makes it challenging to select important variables and control confounding factors rigorously. This thesis suggests identifying joint effects among multiple types of data to provide scientific evidence from various perspectives, which forms a **context** that lay people and experts can understand. This context can serve as a clue and hypothesis for further epidemiological studies to investigate the joint effects and adjust confounding factors. This section summarizes general recommendations for contextualizing scientific evidence.

Integrate human-generated and machine-generated data

Human-generated and machine-generated data provide different perspectives of evidence, as discussed in section 2.2. Integrating these data can provide a better context for identifying underlying patterns of community concerns. For instance, the monitoring system in chapter 4 visualized camera inputs, air quality and wind information from sensors, and citizen-contributed smell reports. The visualization enabled community members to tell a convincing story about when the pollution happened, how smoke emissions affected the community, and how people experienced the pollution. This data-driven story successfully changed the attitudes of regulars during a community meeting with the Allegheny County Health Department and the Environmental Protection Agency. Additionally, the visualization provided a clue that wind information was a possible confounding factor when studying the relationships between smoke emissions and odor complaints. Another example is the interactive tool in chapter 5, which discussed the importance of comparing air quality data with self-reported health symptoms to generate hypotheses. The focus group participants specifically mentioned the need to provide more background information, such as demographics, that could further enhance the context.

Apply both prediction and inference in machine learning

Prediction and inference can complement each other by providing predictive and explanatory powers, as mentioned in section 2.3. Applying both approaches can prevent over-interpreting machine learning models and understand if their decision-making processes are reasonable. For example, the mobile computing tool in chapter 6 trained ensemble-based models for predicting the presence of poor odor events. Another tree-based model was constructed based on a subset of filtered predictors and samples to explain a prevailing pattern among the data. Strategies for selecting representative predictors included manual and automatic approaches. The manual approach used knowledge obtained from the text analysis of smell reports. The automatic approach recursively removed variables that had lower importance weights. The pattern provided an explanation about how pollution traveled to urban areas, which is a joint effect of hydrogen sulfide and wind directions, as a reasonable hypothesis for future epidemiological studies.

Use artificial intelligence to support collecting evidence

Visualizations enable citizens and experts to explore patterns in the data interactively. However, a single case may not be sufficient to explain a pattern, and it is difficult to examine all patterns manually in the data to provide convincing evidence. Artificial intelligence techniques, such as computer vision and machine learning, can reduce the workload by expediting and automating the process of collecting data-driven evidence. The intention of applying this automatic approach is not to replace the manual effort, but to augment human capabilities. For instance, the air quality monitoring system in chapter 4 used a smoke detection algorithm to assist users in identifying a large number of hazardous emissions and generating animated smoke images, which were presented in a community meeting to influence the attitude of regulators. Another example is the Decision Tree model for explaining the relationships between air quality data and smell events in chapter 6. The model can be used to automatically extract and visualize all similar patterns about how air pollutants affect odor experiences on different dates in the dataset.

7.1.3 Evaluate the Impact of Interactive Systems

This thesis emphasizes the importance of evaluating the impact of deploying interactive systems on communities. Merely focusing on usability testing, such as measuring the time of completing tasks, may restrict the perspective of system design [94]. To answer the core research question (discussed in section 1.4), this thesis believes that instead of asking “Is the system useful?” it is more appropriate to ask **“Is the system influential?”** However, unlike observational studies, community citizen science applies information technology to produce scientific knowledge and influence community members simultaneously. If researchers frame this question of identifying the causal relationships as an observational study, it is difficult to track and control confounding factors that may influence their behaviors and attitudes, such as the effect of news and social media. One can treat the intervention of information technology as a randomized experiment. But, it is not practical to randomly sample a control group with sufficient size from affected residents, since the information can spread among communities. Even if there is a way to prevent the control group from accessing the information about the deployed system, it is not ethical and contradicts the value of democratizing scientific knowledge. One may consider a more ethical way to compare the changes in the targeting community with another independent one that shares similar concerns but does not have access to the tool at the beginning. Nevertheless, community citizen science is by nature not replicable since it addresses wicked problems (discussed in section 1.3). Each community has distinct characteristics and power relationships, and the results obtained by conducting the randomized experiment on two independent communities can be misleading. In this sense, it is extremely difficult to statistically verify if the computational tool truly empowers communities and causes attitude or behavior changes.

Measure attitude and behavior changes with qualitative and quantitative analysis

Although it is difficult to statistically and rigorously validate the impact of interactive systems for community citizen science, understanding **"How can the system be influential?"** and **"Does the community think that the system is influential?"** can benefit and inform system design, especially at early stages of development [132]. These findings can provide insights about how computational tools are used to support community citizen science. For instance, chapter 4 studied how community members used animated smoke images and found that both manual and automatic approaches for generating images are essential during the engagement lifecycle. Despite the small sample size in the analysis of self-efficacy and sense of community, the survey study explained that the capability of using data-driven evidence from multiple perspectives is an important reason that the communities felt more confident after interacting with the system. Moreover, applying both qualitative and quantitative analysis can further strengthen the evaluation of impact. For instance, chapter 6 found that motivations for community members to use Smell Pittsburgh came mainly from internal factors, including the desire of contributing data-driven evidence, the concern about the welfare of others, and the capability of validating personal experiences using the visualization. This result of the qualitative analysis was reinforced by the quantitative analysis of system usage, which identified a moderate association between contributing data and interacting with the visualization.

7.2 Final Words

Community citizen science aims to empower everyday citizens and scientists to represent their voices, reveal local concerns, and advocate for social changes by using scientific knowledge. Collaboratively producing and exchanging scientific knowledge requires the intervention of interactive systems. These systems provide technological affordance for community members to collect, visualize, and make sense of data at an extensive spacial-temporal scale. When designing these systems, it is essential to apply visualization, crowdsourcing, and artificial intelligence techniques to initiate, maintain, and evaluate community engagement. Through developing and deploying four computational tools during the entire community engagement lifecycle, this thesis offers generalizable implications: treating community members as co-designers, contextualizing scientific evidence with multiple types of data, and measuring attitude and behavior changes after deploying systems. Although these four tools may not be entirely replicable, they serve as concrete examples and case studies. This thesis is the beginning of designing interactive systems that support community citizen science, and I hope that its methodological and empirical contributions can enlighten and inspire future researchers in this field.

Chapter 8

Appendix

The [System Name] Survey

This survey is part of the [system name] research. The purpose is to study how people use, share, and think about the [system name]. The information will help develop similar systems for supporting environmental justice. The survey is completely anonymous and is expected to take less than 30 minutes.

Please answer each question carefully and to the best of your ability. Thank you!

Before you knew the [system name]

1. Please select the answer that shows how you feel about the statement.
   (Answer this question based on the time **before** you knew the [system name])

	Strongly agree	Agree	Neither agree or disagree	Disagree	Strongly disagree
I was concerned about air quality.	○	○	○	○	○
My involvement with ACCAN's actions was active.	○	○	○	○	○
I was confident when I discussed air quality issues with other people.	○	○	○	○	○
I felt my actions were influential in improving the local air quality with ACCAN.	○	○	○	○	○
I was comfortable with local air quality.	○	○	○	○	○
I was confident that I and ACCAN could achieve the goal of improving the local air quality.	○	○	○	○	○

After you knew the [system name]

2. Please select the answer that shows how you feel about the statement.
   (Answer this question based on the time **after** you knew the [system name])

	Strongly agree	Agree	Neither agree or disagree	Disagree	Strongly disagree
I was concerned about air quality.	○	○	○	○	○
My involvement with ACCAN's actions was active.	○	○	○	○	○
I was confident when I discussed air quality issues with other people.	○	○	○	○	○
I felt my actions were influential in improving the local air quality with ACCAN.	○	○	○	○	○
I was comfortable with local air quality.	○	○	○	○	○
I was confident that I and ACCAN could achieve the goal of improving the local air quality.	○	○	○	○	○

3. About how many people did you discuss the [system name] with?

4. About how many times did you go to the monthly ACCAN meetings in 2015?

5. Did you **visit** or **use** the [system name]?
   - ○ Yes
   - ○ No [Skip question 4 and 5]

6. How did you **explore** data on the [system name]?
   (Ignore this question if you did not explore data.)
   [Please select all answers that apply]
   - ○ I clicked on the video playback button to watch the video
   - ○ I zoomed in and out of the video
   - ○ I used the line charts showing air quality data
   - ○ I clicked on the fast-forwarding button to jump to smoke emissions
   - ○ I browsed the images produced by the automatic smoke detection tool

   Other (please specify)

7. How did you **document** data on the [system name]?
   (Ignore this question if you did not document data.)
   [Please select all answers that apply]
   - ○ I found smoke in the video and used the thumbnail tool to generate images. Then I collected these smoke images in a document (e.g. Google Doc, Microsoft Word)
   - ○ I selected the images produced by the automatic smoke detection tool. Then I collected these smoke images in a document (e.g. Google Doc, Microsoft Word)
   - ○ I wrote stories on online platforms (e.g. Facebook, Twitter, and Post-Gazette), and these stories referred to the [system name].

   Other (please specify)

8. How did you **share** data via email or online platforms?
(Data include images, web links, stories, etc. Online platforms include Facebook, Twitter, Post-Gazette, etc. Ignore this question if you did not share data.)
[Please select all answers that apply]
   - ○ I shared smoke images produced on the [system name]
   - ○ I shared screenshots of the [system name]
   - ○ I shared web links to the [system name]
   - ○ I shared stories which refer to the [system name]

   Other (please specify)

9. If you **shared** data, what were your **motivations**?
(Ignore this question if you did not share data.)
[Please select all answers that apply]
   - ○ I thought that the data is valuable. I shared to make people aware of air quality issues.
   - ○ I shared to show that I care about air quality problems
   - ○ I shared to stay close to people who also care about air quality problems
   - ○ I shared to get feedback from others. This made me feel I was a part of the community
   - ○ I shared to spread the word about the [system name]

   Other (please specify)

10. How often did you browse the [system name] after you noted smoke or bad smells? (E.g. look for evidence, check smell reports)

Every time	Almost every time	Sometimes	Almost never	Never
○	○	○	○	○

11. The [system name] shows and records air quality data. How likely do you think that browsing the [system name] makes people care about air quality problems?

Extremely likely	Likely	Neither likely nor unlikely	Unlikely	Extremely unlikely
○	○	○	○	○

12. There are published air quality stories which refer to the [system name]. How likely do you think that reading these stories make people care about air quality problems?

Extremely likely	Likely	Neither likely nor unlikely	Unlikely	Extremely unlikely
○	○	○	○	○

[Images of the system]

13. Please rate the importance of the features on the [system name].

[Skip the ones that you do not know]	Extremely important	Very important	Moderately important	Slightly important	Not at all important
Short stories on the first page	◯	◯	◯	◯	◯
The timelapse video	◯	◯	◯	◯	◯
Zooming in and out of the video	◯	◯	◯	◯	◯
Sharing a web link of a view and time	◯	◯	◯	◯	◯
Smell reports	◯	◯	◯	◯	◯
Line charts showing sensor readings	◯	◯	◯	◯	◯
The line chart showing smoke detection	◯	◯	◯	◯	◯
The map showing sensor values	◯	◯	◯	◯	◯
The thumbnail tool that people can use to manually generate smoke images	◯	◯	◯	◯	◯
The automatic smoke detection tool which produces smoke images	◯	◯	◯	◯	◯
Smoke images which were shown during the community meeting with the EPA	◯	◯	◯	◯	◯

14. Do you have other comments?

15. What is your age?

- ○ 18-24
- ○ 25-34
- ○ 35-44
- ○ 45-54
- ○ 55-64
- ○ 65-74
- ○ 75+

16. What is your education level

- ○ No formal educational credential
- ○ High school diploma or equivalent
- ○ Some college, no degree
- ○ Postsecondary nondegree award
- ○ Associate's degree
- ○ Bachelor's degree
- ○ Master's degree
- ○ Doctoral or professional degree

You have completed the survey. This survey is part of the [system name] research. Thank you very much for your participation.

Smell Pittsburgh Community Empowerment Survey

This survey is part of the Smell Pittsburgh project. The information will help develop similar systems for supporting environmental justice. The survey is completely anonymous and is expected to take less than 10 minutes.

Please answer each question carefully and to the best of your ability. Thank you!

1. Answer this question based on the time **BEFORE** you knew about or used Smell Pittsburgh.

Please indicate how much you DISAGREE or AGREE with each of the following statements about your influence on local air quality. Please respond as you really felt, rather than how you think "most people" would feel.

	Strongly disagree	Disagree	Neutral	Agree	Strongly Agree
I felt confident in my ability to help protect local air quality.	○	○	○	○	○
I was capable of making a positive impact on local air quality.	○	○	○	○	○
I was able to help take care of local air quality.	○	○	○	○	○
I believed I could contribute to solutions to local pollution problems by my actions.	○	○	○	○	○
Compared to other people, I thought I could make a positive impact on local air quality.	○	○	○	○	○
I didn't think I could make any difference in solving local pollution problems.	○	○	○	○	○
I believed that I personally, working with others, could help solve local air issues.	○	○	○	○	○
It was hard for me to imagine myself helping to protect local air quality.	○	○	○	○	○

2. Answer this question based on the time **AFTER** you knew about or used Smell Pittsburgh.

Please indicate how much you DISAGREE or AGREE with each of the following statements about your influence on local air quality. Please respond as you really felt, rather than how you think "most people" would feel.

	Strongly disagree	Disagree	Neutral	Agree	Strongly Agree
I feel confident in my ability to help protect local air quality.	○	○	○	○	○
I am capable of making a positive impact on local air quality.	○	○	○	○	○
I am able to help take care of local air quality.	○	○	○	○	○
I believe I can contribute to solutions to local pollution problems by my actions.	○	○	○	○	○
Compared to other people, I think I can make a positive impact on local air quality.	○	○	○	○	○
I don't think I can make any difference in solving local pollution problems.	○	○	○	○	○
I believe that I personally, working with others, can help solve local air issues.	○	○	○	○	○
It's hard for me to imagine myself helping to protect local air quality.	○	○	○	○	○

3. Please indicate how much you DISAGREE or AGREE with each of the following statements. Please respond as you really feel, rather than how you think "most people" feel.

Think about some of the ways you use Smell Pittsburgh to help solve local air pollution problems. Why do you use Smell Pittsburgh?

	Strongly disagree	Disagree	Neutral	Agree	Strongly Agree
Because I think it's a good idea to do something for local air quality	○	○	○	○	○
Because I enjoy doing it	○	○	○	○	○
For the pleasure I experience while doing it	○	○	○	○	○
Because other people will be disappointed in me if I don't	○	○	○	○	○
Because I'm concerned about what could happen to people I care about if I don't do anything	○	○	○	○	○
Because I would feel guilty if I didn't do anything for local air quality	○	○	○	○	○
Because I'm concerned about what could happen to me if I don't do anything	○	○	○	○	○

4. Continued from the previous page...

Please indicate how much you DISAGREE or AGREE with each of the following statements. Please respond as you really feel, rather than how you think “most people” feel.

Think about some of the ways you use Smell Pittsburgh to help solve local air pollution problems. Why do you use Smell Pittsburgh?

	Strongly disagree	Disagree	Neutral	Agree	Strongly Agree
Because I think it’s a good idea to protect local air quality	○	○	○	○	○
Because it's fun to do it	○	○	○	○	○
Because I think it’s important to take care of local air quality	○	○	○	○	○
Because I'm concerned about what could happen to local air quality if I don't do anything	○	○	○	○	○
Because people I look up to think it's a really good thing to do	○	○	○	○	○
For the recognition I get from others	○	○	○	○	○
Because I want people to see me as a good person	○	○	○	○	○

Are there other reasons that you use Smell Pittsburgh?

5. How often do you use Smell Pittsburgh?

- ○ At least once per day
- ○ At least once per week
- ○ At least once per month
- ○ At least once per year
- ○ Never use Smell Pittsburgh
- ○ Other (please specify)

6. How did you use Smell Pittsburgh? [Please select all answers that apply]

- ○ I submitted smell reports.
- ○ I provided my contact information when submitting smell reports.
- ○ I checked other people's smell reports on the map visualization.
- ○ I clicked on the playback button to view the animation of smell reports.
- ○ I opened Smell Pittsburgh when I noticed unusual smell.
- ○ I paid attention to smell event alert notifications provided by Smell Pittsburgh.
- ○ I downloaded smell reports data from the Smell Pittsburgh website.
- ○ I discussed Smell Pittsburgh with other people.
- ○ I mentioned or presented Smell Pittsburgh to regulators.
- ○ I shared Smell Pittsburgh publicly online (e.g. email, social media, news blog).
- ○ I took screenshots of Smell Pittsburgh.

Other (please specify)

7. Do you have other comments?

8. What is your age?
- ○ 18-24
- ○ 25-34
- ○ 35-44
- ○ 45-54
- ○ 55-64
- ○ 65-74
- ○ 75+
- ○ Prefer not to say

9. What is your education level
- ○ No formal educational credential
- ○ High school diploma or equivalent
- ○ Some college, no degree
- ○ Postsecondary nondegree award
- ○ Associate's degree
- ○ Bachelor's degree
- ○ Master's degree
- ○ Doctoral or professional degree
- ○ Prefer not to say

10. What is your gender?
- ○ Female
- ○ Male
- ○ Prefer not to say

You have completed the survey. This survey is part of the Smell Pittsburgh project. Thank you very much for your participation.

Bibliography

[1] Google earth engine - landsat annual earth timelapse, 2012. https://earthengine.google.org/#intro/Amazon. 3.1

[2] Timelapse: Landsat satellite images of climate change, via google earth engine, 2012. http://world.time.com/timelapse/. 3.3

[3] PennEnvironment, 2015. http://www.pennenvironment.org/reports/pae/toxic-ten. 6.2

[4] AirNow, 2016. https://www.airnow.gov/. 2.2.2

[5] American Lung Association, 2016. http://www.lung.org/our-initiatives/healthy-air/sota/city-rankings/most-polluted-cities.html. 6.1

[6] Allegheny County Clean Air Now, 2018. http://accan.org/. 6.2

[7] Clean Air Council, 2018. https://cleanair.org/. 6.2

[8] Group Against Smog and Pollution, 2018. http://gasp-pgh.org/. 6.2

[9] PennFuture, 2018. https://www.pennfuture.org/. 6.2

[10] Tanja Aitamurto. Motivation factors in crowdsourced journalism: Social impact, social change, and peer learning. *International Journal of Communication*, 9:21, 2015. 6.4.3

[11] Tanja Aitamurto and Jorge Saldivar. Motivating participation in crowdsourced policymaking: The interplay of epistemic and interactive aspects. *Proc. ACM Hum.-Comput. Interact.*, 1(CSCW):18:1–18:22, December 2017. ISSN 2573-0142. doi: 10.1145/3134653. URL http://doi.acm.org/10.1145/3134653. 6.4.3

[12] Sylvain Arlot and Alain Celisse. A survey of cross-validation procedures for model selection. *Statist. Surv.*, 4:40–79, 2010. doi: 10.1214/09-SS054. URL https://doi.org/10.1214/09-SS054. 6.4.2

[13] David Arthur and Sergei Vassilvitskii. K-means++: The advantages of careful seeding. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1027–1035, Philadelphia, PA, USA, 2007. Society for Industrial and Applied Mathematics. 4.4.3

[14] Azman Azid, Hafizan Juahir, Mohd Ekhwan Toriman, Mohd Khairul Amri Kamarudin, Ahmad Shakir Mohd Saudi, Che Noraini Che Hasnam, Nor Azlina Abdul Aziz, Fazureen Azaman, Mohd Talib Latif, Syahrir Farihan Mohamed Zainuddin, Mohamad Romizan Osman, and Mohammad Yamin. Prediction of the level of air pollution using principal component analysis and artificial neural network techniques: a case study in malaysia.

Water, Air, & Soil Pollution, 225(8):2063, Jul 2014. ISSN 1573-2932. doi: 10.1007/s11270-014-2063-1. URL https://doi.org/10.1007/s11270-014-2063-1. 2.3.2

[15] Albert Bandura. Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, 84(2):191, 1977. 4.1, 4.5.2, 6.4.3

[16] Kathy Baxter, Catherine Courage, and Kelly Caine. *Understanding your users: A practical guide to user research methods*. Morgan Kaufmann, 2015. 5.3

[17] Colin Bellinger, Mohomed Shazan Mohomed Jabbar, Osmar Zaïane, and Alvaro Osornio-Vargas. A systematic review of data mining and machine learning for air pollution epidemiology. *BMC public health*, 17(1):907, 2017. 1.4, 2.3

[18] Hugh Beyer and Karen Holtzblatt. *Contextual design: defining customer-centered systems*. Elsevier, 1997. 5.3

[19] Patrick Biernacki and Dan Waldorf. Snowball sampling: Problems and techniques of chain referral sampling. *Sociological methods & research*, 10(2):141–163, 1981. 6.4.3

[20] Jeffrey P Bigham, Michael S Bernstein, and Eytan Adar. Human-computer interaction and collective intelligence. *Handbook of collective intelligence*, 57, 2015. 2.2.1

[21] Steven Bird, Ewan Klein, and Edward Loper. *Natural language processing with Python: analyzing text with the natural language toolkit*. " O'Reilly Media, Inc.", 2009. 6.4.1

[22] Tomas J Bird, Amanda E Bates, Jonathan S Lefcheck, Nicole A Hill, Russell J Thomson, Graham J Edgar, Rick D Stuart-Smith, Simon Wotherspoon, Martin Krkosek, Jemina F Stuart-Smith, et al. Statistical solutions for error and bias in global citizen science datasets. *Biological Conservation*, 173:144–154, 2014. 1.3.2, 1.4, 2.3, 6.5.1

[23] Christopher Bishop. *Pattern Recognition and Machine Learning*. Springer-Verlag New York, 2006. ISBN 978-0-387-31073-2. 2.3, 6.3.2

[24] Eli Blevis. Sustainable interaction design: Invention & disposal, renewal & reuse. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '07, pages 503–512, New York, NY, USA, 2007. ACM. ISBN 978-1-59593-593-9. doi: 10.1145/1240624.1240705. URL http://doi.acm.org/10.1145/1240624.1240705. 1.5

[25] Rick Bonney, Heidi Ballard, Rebecca Jordan, Ellen McCallie, Tina Phillips, Jennifer Shirk, and Candie C Wilderman. Public participation in scientific research: Defining the field and assessing its potential for informal science education. a caise inquiry group report. *Online Submission*, 2009. 1.1.1, 1.1.2

[26] Rick Bonney, Caren B. Cooper, Janis Dickinson, Steve Kelling, Tina Phillips, Kenneth V. Rosenberg, and Jennifer Shirk. Citizen science: A developing tool for expanding science knowledge and scientific literacy. *BioScience*, 59(11):977–984, 2009. doi: 10.1525/bio.2009.59.11.9. URL http://bioscience.oxfordjournals.org/content/59/11/977.abstract. 1.1.1, 1.3.1, 1.3.2, 1.3.3

[27] Rick Bonney, Jennifer L. Shirk, Tina B. Phillips, Andrea Wiggins, Heidi L. Ballard, Abraham J. Miller-Rushing, and Julia K. Parrish. Next steps for citizen science. *Science*,

343(6178):1436–1437, 2014. ISSN 0036-8075. doi: 10.1126/science.1251554. URL http://science.sciencemag.org/content/343/6178/1436. 1.1, 1.3.1, 1.3.2, 2, 2.3

[28] Rick Bonney, Tina B Phillips, Heidi L Ballard, and Jody W Enck. Can citizen science enhance public understanding of science? *Public Understanding of Science*, 25(1):2–16, 2016. 1.1

[29] Daren C. Brabham. Crowdsourcing as a model for problem solving: An introduction and cases. *Convergence: The International Journal of Research into New Media Technologies*, 14(1):75–90, 2008. doi: 10.1177/1354856507084420. URL http://con.sagepub.com/content/14/1/75.abstract. 1.4, 2.2.1

[30] Peter Bradwell and Sarah Marr. *Making the most of collaboration: An international survey of public service co-design*. Demos, 2008. 4.6.1, 7.1.1

[31] Leo Breiman. Random forests. *Machine Learning*, 45(1):5–32, Oct 2001. ISSN 1573-0565. doi: 10.1023/A:1010933404324. URL https://doi.org/10.1023/A:1010933404324. 6.4.2

[32] Dominique Brossard, Bruce Lewenstein, and Rick Bonney. Scientific knowledge and attitude change: The impact of a citizen science project. *International Journal of Science Education*, 27(9):1099–1121, 2005. 1.3.3

[33] Phil Brown. Popular epidemiology: Community response to toxic waste-induced disease in woburn, massachusetts. *Science, Technology, & Human Values*, 12(3/4):78–85, 1987. 2.1

[34] Phil Brown. Popular epidemiology and toxic waste contamination: lay and professional ways of knowing. *Journal of health and social behavior*, pages 267–281, 1992. 2.1, 2.3.3

[35] Phil Brown. Integrating medical and environmental sociology with environmental health: crossing boundaries and building connections through advocacy. *Journal of health and social behavior*, 54(2):145–164, 2013. 2.1

[36] Phil Brown and Edwin J Mikkelsen. *No safe place: Toxic waste, leukemia, and community action*. Univ of California Press, 1997. 2.1

[37] Hronn Brynjarsdottir, Maria Håkansson, James Pierce, Eric Baumer, Carl DiSalvo, and Phoebe Sengers. Sustainably unpersuaded: How persuasion narrows our vision of sustainability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 947–956, New York, NY, USA, 2012. ACM. ISBN 978-1-4503-1015-4. doi: 10.1145/2207676.2208539. URL http://doi.acm.org/10.1145/2207676.2208539. 1.5

[38] Matthias Budde, Andrea Schankin, Julien Hoffmann, Marcel Danz, Till Riedel, and Michael Beigl. Participatory sensing or participatory nonsense?: Mitigating the effect of human error on data quality in citizen science. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(3):39, 2017. 6.5.1

[39] Caroline Bushdid, Marcelo O Magnasco, Leslie B Vosshall, and Andreas Keller. Humans can discriminate more than 1 trillion olfactory stimuli. *Science*, 343(6177):1370–1372,

2014. 6.2

[40] Simone Calderara, Paolo Piccinini, and Rita Cucchiara. Smoke detection in video surveillance: A mog model in the wavelet domain. In *Computer Vision Systems*, volume 5008 of *Lecture Notes in Computer Science*, pages 119–128. Springer Berlin Heidelberg, 2008. 4.4

[41] John M. Carroll, Mary Beth Rosson, and Jingying Zhou. Collective efficacy as a measure of community. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '05, pages 1–10, New York, NY, USA, 2005. ACM. ISBN 1-58113-998-5. doi: 10.1145/1054972.1054974. URL http://doi.acm.org/10.1145/1054972.1054974. 4.1, 4.5.2

[42] Rich Caruana, Mohamed Elhawary, Art Munson, Mirek Riedewald, Daria Sorokina, Daniel Fink, Wesley M Hochachka, and Steve Kelling. Mining citizen science data to predict orevalence of wild bird species. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 909–915. ACM, 2006. 6.4.2

[43] Turgay Çelik, Hüseyin Özkaramanli, and Hasan Demirel. Fire and smoke detection without sensors: Image processing based approach. In *European Signal Processing Conference*, pages 1794–1798, 2007. 4.4

[44] Ramya Chari, Luke J Matthews, Marjory S Blumenthal, Amanda F Edelman, and Therese Jones. The promise of community citizen science. 2017. 1.1.1

[45] Sen-Ching S. Cheung and Chandrika Kamath. Robust background subtraction with foreground validation for urban traffic video. *EURASIP J. Appl. Signal Process.*, 2005:2330–2340, January 2005. 4.4

[46] Jacob Cohen, Patricia Cohen, Stephen G West, and Leona S Aiken. *Applied multiple regression/correlation analysis for the behavioral sciences*. Routledge, 2013. (document), 1.4.3, 1.3

[47] Jeffrey P. Cohn. Citizen science: Can volunteers do real research? *BioScience*, 58(3):192–197, 2008. doi: 10.1641/B580303. URL http://bioscience.oxfordjournals.org/content/58/3/192.abstract. 1.1.1, 2.3

[48] Theo Colborn, Carol Kwiatkowski, Kim Schultz, and Mary Bachran. Natural gas operations from a public health perspective. *Human and Ecological Risk Assessment: An International Journal*, 17(5):1039–1056, 2011. doi: 10.1080/10807039.2011.605662. URL http://dx.doi.org/10.1080/10807039.2011.605662. 5.1

[49] Robert Collins, Alan Lipton, Takeo Kanade, Hironobu Fujiyoshi, David Duggins, Yanghai Tsin, David Tolliver, Nobuyoshi Enomoto, and Osamu Hasegawa. A system for video surveillance and monitoring. Technical report, The Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, May 2000. 4.4

[50] Jeff Conklin. *Dialogue Mapping: Building Shared Understanding of Wicked Problems*. John Wiley & Sons, Inc., New York, NY, USA, 2005. ISBN 0470017686. 1.3, 1.3.3, 4.2, 4.3, 4.6.1

[51] Cathy C. Conrad and Krista G. Hilchey. A review of citizen science and community-based environmental monitoring: issues and opportunities. *Environmental Monitoring and Assessment*, 176(1):273–291, 2011. ISSN 1573-2959. doi: 10.1007/s10661-010-1582-5. URL http://dx.doi.org/10.1007/s10661-010-1582-5. 1.1.3

[52] Caren B. Cooper and Bruce V. Lewenstein. Two meanings of citizen science. In Darlene Cavalier and Eric B. Kennedy, editors, *The Rightful Place of Science: Citizen Science*. Consortium for Science, Policy & Outcomes, Arizona State University, 2016. 1.1.1

[53] Caren B Cooper, Janis Dickinson, Tina Phillips, and Rick Bonney. Citizen science as a tool for conservation in residential ecosystems. *Ecology and Society*, 12(2):11, 2007. 1.1.1, 1.1.2

[54] Jason Corburn. *Street Science: Community Knowledge and Environmental Health Justice (Urban and Industrial Environments)*. The MIT Press, 2005. 1.2.1, 1.4

[55] National Research Council, Committee on Acute Exposure Guideline Levels, et al. *Acute exposure guideline levels for selected airborne chemicals*, volume 9. National Academies Press, 2009. 6.4.1

[56] Vickie Curtis. Motivation to participate in an online citizen science game: A study of foldit. *Science Communication*, 37(6):723–746, 2015. 6.4.3

[57] Christopher A Le Dantec and Carl DiSalvo. Infrastructuring and the formation of publics in participatory design. *Social Studies of Science*, 43(2):241–264, 2013. doi: 10.1177/0306312712471581. 1.3, 1.4.3

[58] Lea Den Broeder, Jeroen Devilee, Hans Van Oers, A Jantine Schuit, and Annemarie Wagemakers. Citizen science for public health. *Health promotion international*, page daw086, 2016. 2.1, 2.3, 2.3.3

[59] Developing, Validating, and Implementing Situated Evaluation Instruments, 2010. http://www.birds.cornell.edu/citscitoolkit/evaluation/instruments. 6.4.3

[60] Rodolphe Devillers and Robert Jeansoulin. *Fundamentals of Spatial Data Quality (Geographical Information Systems Series)*. ISTE, 2006. ISBN 1905209568. 6.1

[61] John Dewey and Melvin L Rogers. *The public and its problems: An essay in political inquiry*. Penn State Press, 2012. 1.4

[62] Janis L. Dickinson and Rick Bonney. *Citizen Science: Public Participation in Environmental Research*. Cornell University Press, 1 edition, 2012. ISBN 0801449111. 1.1.1

[63] Janis L. Dickinson, Benjamin Zuckerberg, and David N. Bonter. Citizen science as an ecological research tool: Challenges and benefits. *Annual Review of Ecology, Evolution, and Systematics*, 41:149 – 172, Jan-12-2010 2010. ISSN 1543-592X. doi: 10.1146/annurev-ecolsys-102209-144636. URL http://www.annualreviews.org/doi/abs/10.1146/annurev-ecolsys-102209-144636. 1.3.1

[64] Janis L Dickinson, Jennifer Shirk, David Bonter, Rick Bonney, Rhiannon L Crain, Jason Martin, Tina Phillips, and Karen Purcell. The current state of citizen science as a tool for ecological research and public engagement. *Frontiers in Ecology and the Environment*, 10

(6):291–297, 2012. ISSN 1540-9309. doi: 10.1890/110236. URL http://dx.doi.org/10.1890/110236. 1.1.1

[65] Thomas G. Dietterich. Ensemble methods in machine learning. In *Multiple Classifier Systems*, pages 1–15, Berlin, Heidelberg, 2000. Springer Berlin Heidelberg. ISBN 978-3-540-45014-6. 6.4.2

[66] Carl DiSalvo. Design, democracy and agonistic pluralism. In *Proceedings of the design research society conference*, pages 366–371, 2010. 1.2.1, 4.7

[67] Carl DiSalvo. *Adversarial Design*. The MIT Press, 2012. ISBN 0262017385, 9780262017381. 1.2.1, 4.7

[68] Carl DiSalvo, Illah Nourbakhsh, David Holstius, Ayça Akin, and Marti Louw. The neighborhood networks project: A case study of critical engagement and creative expression through participatory design. In *Proceedings of the Tenth Anniversary Conference on Participatory Design 2008*, PDC '08, pages 41–50, Indianapolis, IN, USA, 2008. Indiana University. ISBN 978-0-9818561-0-0. URL http://dl.acm.org/citation.cfm?id=1795234.1795241. 1.1.2, 5.3, 5.4

[69] Carl DiSalvo, Kirsten Boehner, Nicholas A. Knouf, and Phoebe Sengers. Nourishing the ground for sustainable hci: Considerations from ecologically engaged art. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 385–394, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-246-7. doi: 10.1145/1518701.1518763. URL http://doi.acm.org/10.1145/1518701.1518763. 1.5

[70] Carl DiSalvo, Marti Louw, Julina Coupland, and MaryAnn Steiner. Local issues, local uses: Tools for robotics and sensing in community contexts. In *Proceedings of the Seventh ACM Conference on Creativity and Cognition*, C&C '09, pages 245–254, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-865-0. doi: 10.1145/1640233.1640271. URL http://doi.acm.org/10.1145/1640233.1640271. 1.1.2, 4.6.1

[71] Carl DiSalvo, Phoebe Sengers, and Hrönn Brynjarsdóttir. Mapping the landscape of sustainable hci. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 1975–1984, New York, NY, USA, 2010. ACM. ISBN 978-1-60558-929-9. doi: 10.1145/1753326.1753625. URL http://doi.acm.org/10.1145/1753326.1753625. 1.5

[72] Carl DiSalvo, Andrew Clement, and Volkmar Pipek. Participatory design for, with, and by communities. In Jesper Simonsen and Toni Robertson, editors, *International Handbook of Participatory Design*, pages 182–209. Routledge, 2013. ISBN 9780415720212. 4.2

[73] Douglas W. Dockery, C. Arden Pope, Xiping Xu, John D. Spengler, James H. Ware, Martha E. Fay, Benjamin G. Jr. Ferris, and Frank E. Speizer. An association between air pollution and mortality in six u.s. cities. *New England Journal of Medicine*, 329(24):1753–1759, 1993. doi: 10.1056/NEJM199312093292401. URL http://dx.doi.org/10.1056/NEJM199312093292401. PMID: 8179653. 6.1

[74] Aoife Donnelly, Bruce Misstear, and Brian Broderick. Real time air quality forecasting using integrated parametric and non-parametric regression techniques. *Atmospheric En-*

vironment, 103:53 – 65, 2015. ISSN 1352-2310. doi: https://doi.org/10.1016/j.atmosenv.2014.12.011. URL http://www.sciencedirect.com/science/article/pii/S1352231014009595. 2.3.2

[75] Paul Dourish. Hci and environmental sustainability: The politics of design and the design of politics. In *Proceedings of the 8th ACM Conference on Designing Interactive Systems*, DIS '10, pages 1–10, New York, NY, USA, 2010. ACM. ISBN 978-1-4503-0103-9. doi: 10.1145/1858171.1858173. URL http://doi.acm.org/10.1145/1858171.1858173. 1.5

[76] Ellie D'Hondt, Matthias Stevens, and An Jacobs. Participatory noise mapping works! an evaluation of participatory sensing as an alternative to standard techniques for environmental monitoring. *Pervasive and Mobile Computing*, 9(5):681–694, 2013. 2.2.1

[77] Editor. Timelapse Editor. http://timemachine.cmucreatelab.org/wiki/EarthEngineTourEditor. 3.1

[78] Bradley Efron and Robert J Tibshirani. *An introduction to the bootstrap*. CRC press, 1994. 6.4.2

[79] Serge Egelman, Ed H. Chi, and Steven Dow. *Crowdsourcing in HCI Research*, pages 267–289. Springer New York, New York, NY, 2014. ISBN 978-1-4939-0378-8. doi: 10.1007/978-1-4939-0378-8_11. URL http://dx.doi.org/10.1007/978-1-4939-0378-8_11. 1.1.2, 1.4

[80] MV Eitzel, Jessica L Cappadonna, Chris Santos-Lang, Ruth Ellen Duerr, Arika Virapongse, Sarah Elizabeth West, Christopher Conrad Maximillian Kyba, Anne Bowser, Caren Beth Cooper, Andrea Sforzi, et al. Citizen science terminology matters: Exploring key terms. *Citizen Science: Theory and Practice*, 2(1), 2017. 1.1

[81] Charles Elkan. Using the Triangle Inequality to Accelerate K-Means. In *Proceedings of the Twentieth International Conference on Machine Learning (ICML-2003)*, 2003. 4.4.3

[82] EPA. Air quality index - a guide to air quality and your health. *U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Outreach and Information Division*, 2014. 6.3.1

[83] Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd*, volume 96, pages 226–231, 1996. 6.4.2

[84] Explorables. Landscape Changes along Taiwan's Coastline over a Two-Decade Period, 2012. http://explorables.cmucreatelab.org/explorables/timelapse-story-telling.html. 3.3

[85] Bonne Ford, Moira Burke, William Lassman, Gabriele Pfister, and Jeffrey R Pierce. Status update: is smoke on your mind? using social media to assess smoke exposure. *Atmospheric Chemistry and Physics*, 17(12):7541–7554, 2017. 1.3

[86] David A Forsyth and Jean Ponce. *Computer vision: a modern approach*. Prentice Hall Professional Technical Reference, 2002. 2.3

[87] Ian Foster, Rayid Ghani, Ron S Jarmin, Frauke Kreuter, and Julia Lane. *Big data and*

social science: A practical guide to methods and tools. CRC Press, 2016. 6.4.2

[88] Nir Friedman and Stuart Russell. Image segmentation in video sequences: A probabilistic approach. In *Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence*, UAI'97, pages 175–181, San Francisco, CA, USA, 1997. Morgan Kaufmann Publishers Inc. 4.4

[89] Katherine Gass, Mitch Klein, Howard H Chang, W Dana Flanders, and Matthew J Strickland. Classification and regression trees for epidemiologic research: an air pollution example. *Environmental Health*, 13(1):17, 2014. 2.3.3, 6.4.2

[90] William W. Gaver. Technology affordances. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '91, pages 79–84, New York, NY, USA, 1991. ACM. ISBN 0-89791-383-3. doi: 10.1145/108844.108856. URL http://doi.acm.org/10.1145/108844.108856. 1.3

[91] Pierre Geurts, Damien Ernst, and Louis Wehenkel. Extremely randomized trees. *Machine Learning*, 63(1):3–42, Apr 2006. ISSN 1573-0565. doi: 10.1007/s10994-006-6226-1. URL https://doi.org/10.1007/s10994-006-6226-1. 6.4.2, 6.4.2

[92] James J Gibson. *The ecological approach to visual perception*. Psychology Press, 2013. 1.3

[93] Bernard Greaves and Gordon Lishman. The theory and practice of community politics. A.L.C. Campaign Booklet No. 12, 1980. http://www.rosenstiel.co.uk/aldc/commpol.htm. 1.1.1

[94] Saul Greenberg and Bill Buxton. Usability evaluation considered harmful (some of the time). In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 111–120, New York, NY, USA, 2008. ACM. ISBN 978-1-60558-011-1. doi: 10.1145/1357054.1357074. URL http://doi.acm.org/10.1145/1357054.1357074. 7.1.3

[95] Francois Grey. The age of citizen cyberscience. CERN Courier, 2009. http://cerncourier.com/cws/article/cern/38718. 1.1.2

[96] Jayavardhana Gubbi, Slaven Marusic, and Marimuthu Palaniswami. Smoke detection in video using wavelets and support vector machines. *Fire Safety Journal*, 44:1110 – 1115, 2009. 4.4

[97] Tee L Guidotti. Hydrogen sulfide: advances in understanding human toxicity. *International journal of toxicology*, 29(6):569–581, 2010. 6.4.1

[98] Anhong Guo, Xi ang 'Anthony' Chen, Haoran Qi, Samuel White, Suman Ghosh, Chieko Asakawa, and Jeffrey P. Bigham. Vizlens: A robust and interactive screen reader for interfaces in the real world. In *Proceedings of the 29th Annual ACM Symposium on User Interface Software and Technology*, UIST '16. ACM, 2016. 2.3.1

[99] Isabelle Guyon and André Elisseeff. An introduction to variable and feature selection. *Journal of machine learning research*, 3(Mar):1157–1182, 2003. 6.4.2

[100] Isabelle Guyon, Jason Weston, Stephen Barnhill, and Vladimir Vapnik. Gene selection for cancer classification using support vector machines. *Machine learning*, 46(1-3):389–422,

2002. 6.4.2

[101] Muki Haklay. Citizen science and volunteered geographic information: Overview and typology of participation. In Daniel Sui, Sarah Elwood, and Michael Goodchild, editors, *Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice*, pages 105–122. Springer Netherlands, Dordrecht, 2013. ISBN 978-94-007-4587-2. doi: 10.1007/978-94-007-4587-2_7. URL http://dx.doi.org/10.1007/978-94-007-4587-2_7. 1.1.2, 1.3.1, 2.2.1

[102] Richard Hartley and Andrew Zisserman. *Multiple view geometry in computer vision*. Cambridge university press, 2003. 2.3

[103] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Second Edition*. Springer-Verlag New York, 2009. ISBN 0387848576. 2.3, 6.3.2, 6.4.2

[104] Jeffrey Heer and Ben Shneiderman. Interactive dynamics for visual analysis. *Commun. ACM*, 55(4):45–54, April 2012. ISSN 0001-0782. doi: 10.1145/2133806.2133821. URL http://doi.acm.org/10.1145/2133806.2133821. 1.4

[105] Jeffrey Heer, Michael Bostock, and Vadim Ogievetsky. A tour through the visualization zoo. *Commun. ACM*, 53(6):59–67, June 2010. ISSN 0001-0782. doi: 10.1145/1743546.1743567. URL http://doi.acm.org/10.1145/1743546.1743567. 1.4

[106] Victoria Henshaw. *Urban smellscapes: Understanding and designing city smell environments*. Routledge, 2013. 6.2

[107] Wesley M. Hochachka, Daniel Fink, Rebecca A. Hutchinson, Daniel Sheldon, Weng-Keen Wong, and Steve Kelling. Data-intensive science applied to broad-scale citizen science. *Trends in Ecology & Evolution*, 27(2):130 – 137, 2012. ISSN 0169-5347. doi: http://dx.doi.org/10.1016/j.tree.2011.11.006. URL http://www.sciencedirect.com/science/article/pii/S0169534711003296. Ecological and evolutionary informatics. 1.3.1, 1.3.2, 1.4, 2.3.3

[108] Jake M Hofman, Amit Sharma, and Duncan J Watts. Prediction and explanation in social systems. *Science*, 355(6324):486–488, 2017. 2.3, 6.5.1

[109] Simon Philipp Hohberg. Wildfire smoke detection using convolutional neural networks. Technical report, Freie Universität Berlin, Berlin, Germany, September 2015. 4.4

[110] Karen Holtzblatt, Jessamyn Burns Wendell, and Shelley Wood. *Rapid contextual design: a how-to guide to key techniques for user-centered design*. Elsevier, 2004. 5.3

[111] Hsun-Ping Hsieh, Shou-De Lin, and Yu Zheng. Inferring air quality for station location recommendation based on urban big data. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '15, pages 437–446, New York, NY, USA, 2015. ACM. ISBN 978-1-4503-3664-2. doi: 10.1145/2783258.2783344. URL http://doi.acm.org/10.1145/2783258.2783344. 2.3.2

[112] Yen-Chia Hsu, Paul S Dille, Randy Sargent, and Illah Nourbakhsh. Industrial smoke detection and visualization. Technical Report CMU-RI-TR-16-55, Robotics Institute, Pitts-

burgh, PA, September 2016. 5.1

[113] Yen-Chia Hsu, Paul Dille, Jennifer Cross, Beatrice Dias, Randy Sargent, and Illah Nourbakhsh. Community-empowered air quality monitoring system. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 1607–1619. ACM, 2017. 5.1, 6.4.3, 6.5.2

[114] A Inselberg and Bernard Dimsdale. Parallel coordinates: a tool for visualizing multidimensional geometry.(1990). *DOI: http://dx. doi. org/10.1109/VISUAL*, 1990. 5.2

[115] Alan Irwin. *Citizen science: A study of people, expertise and sustainable development*. Psychology Press, 1995. 1.1.1, 1.1.2

[116] Alan Irwin. Constructing the scientific citizen: Science and democracy in the biosciences. *Public Understanding of Science*, 10(1):1–18, 2001. doi: 10.1088/0963-6625/10/1/301. URL `http://pus.sagepub.com/content/10/1/1.abstract`. 1.1.1

[117] Alan Irwin. The politics of talk: Coming to terms with the 'new' scientific governance. *Social Studies of Science*, 36(2):299–320, 2006. doi: 10.1177/0306312706053350. URL `http://sss.sagepub.com/content/36/2/299.abstract`. 1.1.1, 1.1.2

[118] R. Jain and M. Slaney. Micro stories and mega stories. *IEEE MultiMedia*, 20(1):86–90, Jan 2013. ISSN 1070-986X. doi: 10.1109/MMUL.2013.6. 3.1

[119] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An introduction to statistical learning*, volume 112. Springer, 2013. 2.3, 6.3.2, 6.4.2

[120] Ian T Jolliffe. Principal component analysis and factor analysis. In *Principal component analysis*, pages 115–128. Springer, 1986. 6.4.2

[121] M. I. Jordan and T. M. Mitchell. Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245):255–260, 2015. ISSN 0036-8075. doi: 10.1126/science.aaa8415. URL `http://science.sciencemag.org/content/349/6245/255`. 2.3

[122] Marilena Kampa and Elias Castanas. Human health effects of air pollution. *Environmental Pollution*, 151(2):362 – 367, 2008. Proceedings of the 4th International Workshop on Biomonitoring of Atmospheric Pollution (With Emphasis on Trace Elements). 4.1, 6.1

[123] Jiro Kawakita. The original kj method. *Tokyo: Kawakita Research Institute*, 1991. 5.3

[124] Muin J. Khoury and John P. A. Ioannidis. Big data meets public health. *Science*, 346 (6213):1054–1055, 2014. ISSN 0036-8075. doi: 10.1126/science.aaa2709. URL `http://science.sciencemag.org/content/346/6213/1054`. 1.4

[125] Jinseop S Kim, Matthew J Greene, Aleksandar Zlateski, Kisuk Lee, Mark Richardson, Srinivas C Turaga, Michael Purcaro, Matthew Balkam, Amy Robinson, Bardia F Behabadi, et al. Space-time wiring specificity supports direction selectivity in the retina. *Nature*, 509(7500):331, 2014. 1.1.3

[126] Sunyoung Kim and Eric Paulos. Inair: Sharing indoor air quality measurements and visualizations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 1861–1870, New York, NY, USA, 2010. ACM. ISBN 978-1-60558-929-9. doi: 10.1145/1753326.1753605. URL `http://doi.acm.org/10.1145/1753326.1753605`. 5.3, 5.4

[127] Sunyoung Kim, Christine Robson, Thomas Zimmerman, Jeffrey Pierce, and Eben M. Haber. Creek watch: Pairing usefulness and usability for successful citizen science. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 2125–2134, New York, NY, USA, 2011. ACM. ISBN 978-1-4503-0228-9. doi: 10.1145/1978942.1979251. URL http://doi.acm.org/10.1145/1978942.1979251. 1.1.3, 2.2.1, 5.3, 5.4

[128] Sunyoung Kim, Jennifer Mankoff, and Eric Paulos. Sensr: evaluating a flexible framework for authoring mobile data-collection tools for citizen science. In *Proceedings of the 2013 conference on Computer supported cooperative work*, pages 1453–1462. ACM, 2013. 5.3, 5.4

[129] Sunyoung Kim, Jennifer Mankoff, and Eric Paulos. Sensr: Evaluating a flexible framework for authoring mobile data-collection tools for citizen science. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, CSCW '13, pages 1453–1462, New York, NY, USA, 2013. ACM. ISBN 978-1-4503-1331-5. doi: 10.1145/2441776.2441940. URL http://doi.acm.org/10.1145/2441776.2441940. 2.2.1

[130] Sunyoung Kim, Eric Paulos, and Jennifer Mankoff. inair: A longitudinal study of indoor air quality measurements and visualizations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 2745–2754, New York, NY, USA, 2013. ACM. ISBN 978-1-4503-1899-0. doi: 10.1145/2470654.2481380. URL http://doi.acm.org/10.1145/2470654.2481380. 2.2.2

[131] Sunyoung Kim, Jennifer Mankoff, and Eric Paulos. Exploring barriers to the adoption of mobile technologies for volunteer data collection campaigns. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 3117–3126, New York, NY, USA, 2015. ACM. ISBN 978-1-4503-3145-6. doi: 10.1145/2702123.2702378. URL http://doi.acm.org/10.1145/2702123.2702378. 2.2.1

[132] Predrag Klasnja, Sunny Consolvo, and Wanda Pratt. How to evaluate technologies for health behavior change in hci research. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 3063–3072, New York, NY, USA, 2011. ACM. ISBN 978-1-4503-0228-9. doi: 10.1145/1978942.1979396. URL http://doi.acm.org/10.1145/1978942.1979396. 7.1.3

[133] Ron Kohavi. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2*, IJCAI'95, pages 1137–1143, San Francisco, CA, USA, 1995. Morgan Kaufmann Publishers Inc. ISBN 1-55860-363-8. URL http://dl.acm.org/citation.cfm?id=1643031.1643047. 6.4.2

[134] I. Kopilovic, B. Vagvolgyi, and T. Sziranyi. Application of panoramic annular lens for motion analysis tasks: surveillance and smoke detection. In *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, volume 4, pages 714–717 vol.4, 2000. 4.4

[135] R. Kosara and J. Mackinlay. Storytelling: The next step for visualization. *Computer*, 46

(5):44–50, May 2013. ISSN 0018-9162. 1.4

[136] Richard A Krueger and Mary Anne Casey. *Focus groups: A practical guide for applied research*. Sage publications, 2014. 5.3

[137] Stacey Kuznetsov and Eric Paulos. Participatory sensing in public spaces: Activating urban surfaces with sensor probes. In *Proceedings of the 8th ACM Conference on Designing Interactive Systems*, DIS '10, pages 21–30, New York, NY, USA, 2010. ACM. ISBN 978-1-4503-0103-9. doi: 10.1145/1858171.1858175. URL http://doi.acm.org/10.1145/1858171.1858175. 5.3, 5.4

[138] Stacey Kuznetsov, George Davis, Jian Cheung, and Eric Paulos. Ceci n'est pas une pipe bombe: Authoring urban landscapes with air quality sensors. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 2375–2384, New York, NY, USA, 2011. ACM. ISBN 978-1-4503-0228-9. doi: 10.1145/1978942.1979290. URL http://doi.acm.org/10.1145/1978942.1979290. 2.2.2, 5.3, 5.4

[139] Stacey Kuznetsov, Scott E. Hudson, and Eric Paulos. A low-tech sensing system for particulate pollution. In *Proceedings of the 8th International Conference on Tangible, Embedded and Embodied Interaction*, TEI '14, pages 259–266, New York, NY, USA, 2013. ACM. ISBN 978-1-4503-2635-3. doi: 10.1145/2540930.2540955. URL http://doi.acm.org/10.1145/2540930.2540955. 2.2.2

[140] G. Lane, C. Brueton, D. Diall, D. Airantzis, N. Jeremijenko, G. Papamarkos, G. Roussos, and K. Martin. Community-based public authoring with mobile chemical sensor networks. In *Intelligent Environments, 2006. IE 06. 2nd IET International Conference on*, volume 2, pages 23–29, July 2006. 1.2.1

[141] Giles Lane, Camilla Brueton, George Roussos, Natalie Jeremijenko, George Papamarkos, Dima Diall, Dimitris Airantzis, and Karen Martin. Public authoring & feral robotics. *Proboscis. Cultural Snapshot Number Eleven*, 2006. 1.2.1

[142] Joe Langford and Deana McDonagh. *Focus groups: supporting effective product development*. CRC press, 2003. 5.3

[143] Gierad Laput, Walter S. Lasecki, Jason Wiese, Robert Xiao, Jeffrey P. Bigham, and Chris Harrison. Zensors: Adaptive, rapidly deployable, human-intelligent sensor feeds. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 1935–1944, New York, NY, USA, 2015. ACM. ISBN 978-1-4503-3145-6. doi: 10.1145/2702123.2702416. URL http://doi.acm.org/10.1145/2702123.2702416. 2.3.1

[144] Walter S. Lasecki, Mitchell Gordon, Danai Koutra, Malte F. Jung, Steven P. Dow, and Jeffrey P. Bigham. Glance: Rapidly coding behavioral video with the crowd. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*, UIST '14, pages 551–562, New York, NY, USA, 2014. ACM. ISBN 978-1-4503-3069-5. doi: 10.1145/2642918.2647367. URL http://doi.acm.org/10.1145/2642918.2647367. 2.3.1

[145] Kenneth I. Laws. *Textured Image Segmentation*. PhD thesis, University of Southern

California, Los Angeles., Jan 1980. 4.4.3

[146] Christopher A. Le Dantec, Robert G. Farrell, Jim E. Christensen, Mark Bailey, Jason B. Ellis, Wendy A. Kellogg, and W. Keith Edwards. Publics in practice: Ubiquitous computing at a shelter for homeless mothers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 1687–1696, New York, NY, USA, 2011. ACM. ISBN 978-1-4503-0228-9. doi: 10.1145/1978942.1979189. URL http://doi.acm.org/10.1145/1978942.1979189. 1.2.1

[147] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *nature*, 521(7553): 436, 2015. 6.5.2

[148] Chen-Yu Lee, Chin-Teng Lin, Chao-Ting Hong, and Miin-Tsair Su. Smoke Detection Using Spatial and Temporal Analysis. *International Journal of Innovative Computing, Information and Control*, 8:4749–4770, 2012. 4.4

[149] Richard A. Olshen Charles J. Stone Leo Breiman, Jerome Friedman. *Classification and regression trees*. Chapman & HallCRC, 1984. 2.3.3, 6.4.2, 6.4.2

[150] Margaret W Leung, Irene H Yen, and Meredith Minkler. Community based participatory research: a promising approach for increasing epidemiology's relevance in the 21st century. *International journal of epidemiology*, 33(3):499–506, 2004. 2.1

[151] Joerg Lindenmann, Veronika Matzi, Nicole Neuboeck, Beatrice Ratzenhofer-Komenda, Alfred Maier, and Freyja-Maria Smolle-Juettner. Severe hydrogen sulphide poisoning treated with 4-dimethylaminophenol and hyperbaric oxygen. 2010. 6.4.1

[152] Chris J. Lintott, Kevin Schawinski, Anže Slosar, Kate Land, Steven Bamford, Daniel Thomas, M. Jordan Raddick, Robert C. Nichol, Alex Szalay, Dan Andreescu, Phil Murray, and Jan Vandenberg. Galaxy zoo: morphologies derived from visual inspection of galaxies from the sloan digital sky survey. *Monthly Notices of the Royal Astronomical Society*, 389 (3):1179–1189, 2008. doi: 10.1111/j.1365-2966.2008.13689.x. URL http://mnras.oxfordjournals.org/content/389/3/1179.abstract. 1.1.2

[153] Gilles Louppe, Louis Wehenkel, Antonio Sutera, and Pierre Geurts. Understanding variable importances in forests of randomized trees. In *Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 1*, NIPS'13, pages 431–439, USA, 2013. Curran Associates Inc. URL http://dl.acm.org/citation.cfm?id=2999611.2999660. 6.4.2

[154] Andrés Lucero. Using affinity diagrams to evaluate interactive prototypes. In *Human-Computer Interaction*, pages 231–248. Springer, 2015. 5.3

[155] K. L. Ma, I. Liao, J. Frazier, H. Hauser, and H. N. Kostis. Scientific storytelling using visualization. *IEEE Computer Graphics and Applications*, 32(1):12–19, Jan 2012. ISSN 0272-1716. doi: 10.1109/MCG.2012.24. 1.4

[156] Nicolas Maisonneuve, Matthias Stevens, Maria E. Niessen, Peter Hanappe, and Luc Steels. Citizen noise pollution monitoring. In *Proceedings of the 10th Annual International Conference on Digital Government Research: Social Networks: Making Connections Between Citizens, Data and Government*, dg.o '09, pages 96–103. Dig-

ital Government Society of North America, 2009. ISBN 978-1-60558-535-2. URL http://dl.acm.org/citation.cfm?id=1556176.1556198. 2.2.1

[157] Jitendra Malik, Serge Belongie, Thomas Leung, and Jianbo Shi. Contour and texture analysis for image segmentation. *International Journal of Computer Vision*, 43:7–27, 2001. 4.4.3

[158] Jennifer C. Mankoff, Eli Blevis, Alan Borning, Batya Friedman, Susan R. Fussell, Jay Hasbrouck, Allison Woodruff, and Phoebe Sengers. Environmental sustainability and interaction. In *CHI '07 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '07, pages 2121–2124, New York, NY, USA, 2007. ACM. ISBN 978-1-59593-642-4. doi: 10.1145/1240866.1240963. URL http://doi.acm.org/10.1145/1240866.1240963. 1.5

[159] Duncan C McKinley, Abraham J Miller-Rushing, Heidi L Ballard, Rick Bonney, Hutch Brown, Daniel M Evans, Rebecca A French, Julia K Parrish, Tina B Phillips, Sean F Ryan, et al. Investing in citizen science can improve natural resource management and environmental protection. *Issues in Ecology*, 19, 2015. 1.1, 1.3.1, 1.3.2, 1.3.3, 2

[160] David W. McMillan and David M. Chavis. Sense of community: A definition and theory. *Journal of Community Psychology*, 14(1):6–23, 1986. ISSN 1520-6629. doi: 10.1002/1520-6629(198601)14:1⟨6::AID-JCOP2290140103⟩3.0.CO;2-I. URL http://dx.doi.org/10.1002/1520-6629(198601)14:1<6::AID-JCOP2290140103>3.0.CO;2-I. 4.1, 4.5.2

[161] Abraham Miller-Rushing, Richard Primack, and Rick Bonney. The history of public participation in ecological research. *Frontiers in Ecology and the Environment*, 10(6):285–290, 2012. ISSN 1540-9309. doi: 10.1890/110278. URL http://dx.doi.org/10.1890/110278. 1.1.1

[162] Tom Mitchell. *Machine Learning*. McGraw Hill, 1997. ISBN 0070428077. 2.3, 6.3.2

[163] Global Community Monitor. Bucket brigade, 2016. http://www.gcmonitor.org/communities/start-a-bucket-brigade/. 1.1.3

[164] Chantal Mouffe. *The Democratic Paradox*. verso, 2000. 1.2.1

[165] Greg Newman, Andrea Wiggins, Alycia Crall, Eric Graham, Sarah Newman, and Kevin Crowston. The future of citizen science: emerging technologies and shifting paradigms. *Frontiers in Ecology and the Environment*, 10(6):298–304, 2012. ISSN 1540-9309. doi: 10.1890/110294. URL http://dx.doi.org/10.1890/110294. 1.3.1, 1.3.2, 6.5.1

[166] Don Norman. *The design of everyday things: Revised and expanded edition*. Basic Books (AZ), 2013. 1.3

[167] Donald A. Norman. Affordance, conventions, and design. *interactions*, 6(3):38–43, May 1999. ISSN 1072-5520. doi: 10.1145/301153.301168. URL http://doi.acm.org/10.1145/301153.301168. 1.3

[168] Marianna Obrist, Alexandre N Tuch, and Kasper Hornbaek. Opportunities for odor: experiences with smell and implications for technology. In *Proceedings of the SIGCHI*

Conference on Human Factors in Computing Systems, pages 2843–2852. ACM, 2014. 6.2

[169] Ory Okolloh. Ushahidi, or 'testimony': Web 2.0 tools for crowdsourcing crisis information. *Participatory learning and action*, 59(1):65–70, 2009. 2.2.1

[170] Gwen Ottinger. Buckets of resistance: Standards and the effectiveness of citizen science. *Science, Technology, & Human Values*, 35(2):244–270, 2010. 2.3, 6.1, 6.2, 6.4.2

[171] Gwen Ottinger. Social movement-based citizen science. In Darlene Cavalier and Eric B. Kennedy, editors, *The Rightful Place of Science: Citizen Science*. Consortium for Science, Policy & Outcomes, Arizona State University, 2016. 1.1.1

[172] Gwen Ottinger. Crowdsourcing undone science. *Engaging Science, Technology, and Society*, 3:560–574, 2017. 6.5.1

[173] Gwen Ottinger. Making sense of citizen science: stories as a hermeneutic resource. *Energy Research & Social Science*, 31:41–49, 2017. 2.1

[174] Eric Paulos, RJ Honicky, and Ben Hooker. Citizen science: Enabling participatory urbanism. *Urban Informatics: Community Integration and Implementation*, 2008. 1.1.1

[175] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. Scikit-learn: Machine learning in python. *Journal of machine learning research*, 12(Oct):2825–2830, 2011. 6.4.2

[176] William M. Pena and Steven A. Parshall. *Problem Seeking: An Architectural Programming Primer*. Wiley; 5 edition (February 28, 2012), 2012. ISBN 9781118084144. 1.1.2, 1.3

[177] C Arden Pope III and Douglas W Dockery. Health effects of fine particulate air pollution: lines that connect. *Journal of the air & waste management association*, 56(6):709–742, 2006. 4.1, 6.1

[178] N. Porticella, T. Phillips, and R. Bonney. Motivation for environmental action (generic). *Technical brief series*, 2017. 6.4.3

[179] N. Porticella, T. Phillips, and R. Bonney. Self-efficacy for environmental action (seea, generic). *Technical brief series*, 2017. 6.4.3

[180] David Martin Powers. Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation. 2011. 6.4.2

[181] Annette Prüss-Üstün and Maria Neira. *Preventing disease through healthy environments: a global assessment of the burden of disease from environmental risks*. World Health Organization, 2016. 6.1

[182] Daniele Quercia, Rossano Schifanella, Luca Maria Aiello, and Kate McLean. Smelly maps: the digital life of urban smellscapes. *arXiv preprint arXiv:1505.06851*, 2015. 6.2

[183] Daniele Quercia, Luca Maria Aiello, Rossano Schifanella, et al. The emotional and chromatic layers of urban smells. In *ICWSM*, pages 309–318, 2016. 6.2

[184] M Jordan Raddick, Georgia Bracey, Pamela L Gay, Chris J Lintott, Carie Cardamone, Phil Murray, Kevin Schawinski, Alexander S Szalay, and Jan Vandenberg. Galaxy zoo:

Motivations of citizen scientists. *arXiv preprint arXiv:1303.6886*, 2013. 6.4.3

[185] R.J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam. Image change detection algorithms: a systematic survey. *Image Processing, IEEE Transactions on*, 14(3):294–307, March 2005. 4.4

[186] RJ Reiffenstein, William C Hulbert, and Sheldon H Roth. Toxicology of hydrogen sulfide. *Annual review of pharmacology and toxicology*, 32(1):109–134, 1992. 6.4.1

[187] Roger A Rennekamp and Martha A Nall. Using focus groups in program development and evaluation. *UK Cooperative service, University of Kentucky, College of Agriculture: A Practical Guide for Applied research*, 2000. 5.3

[188] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. 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*, pages 1135–1144. ACM, 2016. 6.5.2

[189] Hauke Riesch and Clive Potter. Citizen science as seen by scientists: Methodological, epistemological and ethical dimensions. *Public Understanding of Science*, 23(1):107–120, 2014. doi: 10.1177/0963662513497324. URL http://pus.sagepub.com/content/23/1/107.abstract. 1.3.1

[190] Horst W. J. Rittel and Melvin M. Webber. Dilemmas in a general theory of planning. *Policy Sciences*, 4(2):155–169, 1973. ISSN 1573-0891. doi: 10.1007/BF01405730. URL http://dx.doi.org/10.1007/BF01405730. 1.3, 1.3.3, 4.2, 4.3, 4.6.1

[191] Dana Rotman, Jenny Preece, Jen Hammock, Kezee Procita, Derek Hansen, Cynthia Parr, Darcy Lewis, and David Jacobs. Dynamic changes in motivation in collaborative citizen-science projects. In *Proceedings of the ACM 2012 conference on computer supported cooperative work*, pages 217–226. ACM, 2012. 6.4.3

[192] Dana Rotman, Kezia Procita, Derek Hansen, Cynthia Sims Parr, and Jennifer Preece. Supporting content curation communities: The case of the encyclopedia of life. *Journal of the American Society for Information Science and Technology*, 63(6):1092–1107, 2012. ISSN 1532-2890. doi: 10.1002/asi.22633. URL http://dx.doi.org/10.1002/asi.22633. 2.2.1

[193] Randy Sargent, Chris Bartley, Paul Dille, Jeff Keller, and Illah Nourbakhsh. Timelapse GigaPan: Capturing, Sharing, and Exploring Timelapse Gigapixel Imagery. In *Fine International Conference on Gigapixel Imaging for Science*, 2010. 3.1, 4.3.1

[194] Henry Sauermann and Chiara Franzoni. Crowd science user contribution patterns and their implications. *Proceedings of the National Academy of Sciences*, 112(3):679–684, 2015. 6.4.1

[195] Bernhard Schölkopf, Alexander Smola, and Klaus-Robert Müller. Nonlinear component analysis as a kernel eigenvalue problem. *Neural computation*, 10(5):1299–1319, 1998. 6.4.2

[196] Bristol Science Communication Unit, University of the West of England. Science for environment policy indepth report: Environmental citizen science. *Report produced for the*

European Commission DG Environment, December 2013. URL http://ec.europa.eu/science-environment-policy. 1.1, 1.1.1

[197] E. Segel and J. Heer. Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1139–1148, Nov 2010. ISSN 1077-2626. doi: 10.1109/TVCG.2010.179. 3.1

[198] Torgyn Shaikhina, Dave Lowe, Sunil Daga, David Briggs, Robert Higgins, and Natasha Khovanova. Decision tree and random forest models for outcome prediction in antibody incompatible kidney transplantation. *Biomedical Signal Processing and Control*, 2017. 6.4.2

[199] Gordon M Shepherd. The human sense of smell: are we better than we think? *PLoS biology*, 2(5):e146, 2004. 6.2

[200] Jennifer L Shirk, Heidi L Ballard, Candie C Wilderman, Tina Phillips, Andrea Wiggins, Rebecca Jordan, Ellen McCallie, Matthew Minarchek, Bruce V Lewenstein, Marianne E Krasny, et al. Public participation in scientific research: a framework for deliberate design. *Ecology and Society*, 17(2):29, 2012. 1.1.2

[201] Galit Shmueli. To explain or to predict? *Statistical science*, pages 289–310, 2010. 2.3

[202] Ben Shneiderman and Catherine Plaisant. Strategies for evaluating information visualization tools: Multi-dimensional in-depth long-term case studies. In *Proceedings of the 2006 AVI Workshop on BEyond Time and Errors: Novel Evaluation Methods for Information Visualization*, BELIV '06, pages 1–7, New York, NY, USA, 2006. ACM. ISBN 1-59593-562-2. doi: 10.1145/1168149.1168158. URL http://doi.acm.org/10.1145/1168149.1168158. 3.3

[203] Bernard W Silverman. *Density estimation for statistics and data analysis*, volume 26. CRC press, 1986. 4.4.4

[204] Jonathan Silvertown. A new dawn for citizen science. *Trends in Ecology & Evolution*, 24(9):467 – 471, 2009. ISSN 0169-5347. doi: http://dx.doi.org/10.1016/j.tree.2009.03.017. URL http://www.sciencedirect.com/science/article/pii/S016953470900175X. 1.1.1

[205] John Snow. *On the mode of communication of cholera*. John Churchill, 1855. 2.1

[206] Speck Air Quality Sensor, 2015. https://www.specksensor.com/. 4.3.3, 5.2

[207] Chris Stauffer and W.E.L. Grimson. Adaptive background mixture models for real-time tracking. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, 1999. 4.4

[208] Matthias Stevens, Michalis Vitos, Julia Altenbuchner, Gillian Conquest, Jerome Lewis, and Muki Haklay. Taking participatory citizen science to extremes. *IEEE Pervasive Computing*, 13(2):20–29, 2014. ISSN 1536-1268. doi: http://doi.ieeecomputersociety.org/10.1109/MPRV.2014.37. 1.4, 4.2

[209] Jack Stilgoe. *Citizen Scientists: reconnecting science with civil society*. Demos London, 2009. 1.1.1

[210] Jack Stilgoe, Simon J. Lock, and James Wilsdon. Why should we promote public en-

gagement with science? *Public Understanding of Science*, 23(1):4–15, 2014. doi: 10.1177/0963662513518154. URL http://pus.sagepub.com/content/23/1/4.abstract. 1.1.1, 1.3.3, 1.4

[211] Jeanette A Stingone, Om P Pandey, Luz Claudio, and Gaurav Pandey. Using machine learning to identify air pollution exposure profiles associated with early cognitive skills among us children. *Environmental Pollution*, 230:730–740, 2017. 2.3.3

[212] Anselm Strauss, Juliet Corbin, et al. *Basics of qualitative research*, volume 15. Newbury Park, CA: Sage, 1990. 2.2.1

[213] Brian L. Sullivan, Christopher L. Wood, Marshall J. Iliff, Rick E. Bonney, Daniel Fink, and Steve Kelling. ebird: A citizen-based bird observation network in the biological sciences. *Biological Conservation*, 142(10):2282 – 2292, 2009. ISSN 0006-3207. doi: http://dx.doi.org/10.1016/j.biocon.2009.05.006. URL http://www.sciencedirect.com/science/article/pii/S000632070900216X. 1.2.1, 2.2.1

[214] Brian L. Sullivan, Jocelyn L. Aycrigg, Jessie H. Barry, Rick E. Bonney, Nicholas Bruns, Caren B. Cooper, Theo Damoulas, André A. Dhondt, Tom Dietterich, Andrew Farnsworth, Daniel Fink, John W. Fitzpatrick, Thomas Fredericks, Jeff Gerbracht, Carla Gomes, Wesley M. Hochachka, Marshall J. Iliff, Carl Lagoze, Frank A. La Sorte, Matthew Merrifield, Will Morris, Tina B. Phillips, Mark Reynolds, Amanda D. Rodewald, Kenneth V. Rosenberg, Nancy M. Trautmann, Andrea Wiggins, David W. Winkler, Weng-Keen Wong, Christopher L. Wood, Jun Yu, and Steve Kelling. The ebird enterprise: An integrated approach to development and application of citizen science. *Biological Conservation*, 169:31 – 40, 2014. ISSN 0006-3207. doi: http://dx.doi.org/10.1016/j.biocon.2013.11.003. URL http://www.sciencedirect.com/science/article/pii/S0006320713003820. 1.2.1, 2.2.1

[215] J. Surowiecki. *The Wisdom of Crowds: Why the Many are Smarter Than the Few and how Collective Wisdom Shapes Business, Economies, Societies, and Nations*. Doubleday, 2004. ISBN 9780385503860. URL https://books.google.com/books?id=bA0c4aYTD6gC. 2.2.1

[216] Richard Szeliski. *Computer vision: algorithms and applications*. Springer Science & Business Media, 2010. 2.3

[217] MD Taylor and IR Nourbakhsh. A low-cost particle counter and signal processing method for indoor air pollution. *Air Pollution XXIII*, 198:337, 2015. 2.2.2, 4.3.3, 5.2

[218] Michael D. Taylor. *Calibration and Characterization of Low-Cost Fine Particulate Monitors and their Effect on Individual Empowerment*. dissertation, Carnegie Mellon University, 2016. 2.2.2

[219] Hongda Tian, Wanqing Li, Philip Ogunbona, editor="Cremers Daniel Wang, Lei", Ian Reid, Hideo Saito, and Ming-Hsuan Yang. *Single Image Smoke Detection*, chapter Computer Vision – ACCV 2014: 12th Asian Conference on Computer Vision, Singapore, Singapore, November 1-5, 2014, Revised Selected Papers, Part II, pages 87–101. Springer International Publishing, 2015. 4.4

[220] Rundong Tian, Christine Dierk, Christopher Myers, and Eric Paulos. Mypart: Personal,

portable, accurate, airborne particle counting. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 1338–1348, New York, NY, USA, 2016. ACM. ISBN 978-1-4503-3362-7. doi: 10.1145/2858036.2858571. URL http://doi.acm.org/10.1145/2858036.2858571. 2.2.2

[221] Ramine Tinati, Markus Luczak-Roesch, Elena Simperl, and Wendy Hall. An investigation of player motivations in eyewire, a gamified citizen science project. *Computers in Human Behavior*, 73:527 – 540, 2017. ISSN 0747-5632. doi: https://doi.org/10.1016/j.chb.2016.12.074. URL http://www.sciencedirect.com/science/article/pii/S0747563216309037. 6.4.3

[222] B.U. Toreyin, Y. Dedeoglu, and A.E. Cetin. Wavelet based real-time smoke detection in video. In *Signal Processing Conference, 2005 13th European*, pages 1–4, Sept 2005. 4.4

[223] A. Vedaldi and B. Fulkerson. VLFeat: An open and portable library of computer vision algorithms, 2008. http://www.vlfeat.org/. 4.4.6

[224] WHO. Ambient (outdoor) air quality and health, 2016. http://www.who.int/mediacentre/factsheets/fs313/en/. 6.1

[225] A. Wiggins and K. Crowston. From conservation to crowdsourcing: A typology of citizen science. In *2011 44th Hawaii International Conference on System Sciences (HICSS)*, pages 1–10, Jan 2011. doi: 10.1109/HICSS.2011.207. 1.1.2

[226] Andrea Wiggins and Yurong He. Community-based data validation practices in citizen science. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, CSCW '16, pages 1548–1559, New York, NY, USA, 2016. ACM. ISBN 978-1-4503-3592-8. doi: 10.1145/2818048.2820063. URL http://doi.acm.org/10.1145/2818048.2820063. 2.3

[227] Candie C Wilderman. Models of community science: design lessons from the field. In *Citizen Science Toolkit Conference, C. McEver, R. Bonney, J. Dickinson, S. Kelling, K. Rosenberg, and JL Shirk, Eds., Cornell Laboratory of Ornithology, Ithaca, NY*, 2007. 1.1.2, 1.3.1

[228] James Wilsdon, Jack Stilgoe, and Brian Wynne. *The public value of science: or how to ensure that science really matters*. Demos London, 2005. 1.1.1

[229] Yu Zheng, Xiuwen Yi, Ming Li, Ruiyuan Li, Zhangqing Shan, Eric Chang, and Tianrui Li. Forecasting fine-grained air quality based on big data. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '15, pages 2267–2276, New York, NY, USA, 2015. ACM. ISBN 978-1-4503-3664-2. doi: 10.1145/2783258.2788573. URL http://doi.acm.org/10.1145/2783258.2788573. 2.3.2

[230] John Zimmerman, Anthony Tomasic, Charles Garrod, Daisy Yoo, Chaya Hiruncharoenvate, Rafae Aziz, Nikhil Ravi Thiruvengadam, Yun Huang, and Aaron Steinfeld. Field trial of tiramisu: crowd-sourcing bus arrival times to spur co-design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1677–1686. ACM, 2011. 1.2.1, 1.3, 6.4.3

[231] Karel Zuiderveld. Contrast Limited Adaptive Histograph Equalization. pages 474–485. Academic Press Professional, Inc., 1994. 4.4.2