

In-home sensing, air quality, annotation, thematic analysis, longitudinal deployment

Managing In-home Environments through Sensing, Annotating, and Visualizing Air Quality Data

JIMMY MOORE*, University of Utah
 PASCAL GOFFIN, University of Utah
 MIRIAH MEYER, University of Utah
 PHILIP LUNDRIGAN, University of Utah
 NEAL PATWARI, University of Utah
 KATHERINE SWARD, University of Utah
 JASON WIESE, University of Utah

Air quality is important, varies across time and space, and is largely invisible. Pioneering past work deploying air quality monitors in residential environments found that study participants improved their awareness of and engagement with air quality. However, these systems fielded a single monitor and did not support user-specified annotations, inhibiting their utility. We developed MAAV— a system to **M**eaure **A**ir quality, **A**nnotate data streams, and **V**isualize real-time $PM_{2.5}$ levels – to explore how participants engage with an air quality system addressing these challenges. MAAV supports collecting data from multiple air quality monitors, annotating that data through multiple modalities, sending text message prompts when it detects a $PM_{2.5}$ spike; it also features an interactive tablet interface for displaying measurement data and annotations. Through six long-term field deployments (20-47 weeks, mean 37.7 weeks), participants found these system features important for understanding the air quality in and around their homes. Participants gained new insights from between-monitor comparisons, reflected on past $PM_{2.5}$ spikes with the help of their annotations, and adapted their system usage as they familiarized themselves with their air quality data and MAAV. These results yield important insights for designing residential sensing systems that integrate into users' everyday lives.

CCS Concepts: • **Human-centered computing** → **Empirical studies in HCI**; **Empirical studies in ubiquitous and mobile computing**;

Additional Key Words and Phrases: Air quality, longitudinal deployment, annotation, thematic analysis, microclimate, microenvironment, case study, user engagement, personal informatics, text messaging, push notifications, asthma

ACM Reference format:

Jimmy Moore, Pascal Goffin, Miriah Meyer, Philip Lundrigan, Neal Patwari, Katherine Sward, and Jason Wiese. 2018. Managing In-home Environments through Sensing, Annotating, and Visualizing Air Quality Data. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 3, Article 39 (September 2018), 28 pages.
<https://doi.org/0000001.0000001>

*This is the corresponding author

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2018 Association for Computing Machinery.
 2474-9567/2018/9-ART39 \$15.00
<https://doi.org/0000001.0000001>

Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, Vol. 2, No. 3, Article 39. Publication date: September 2018.

1 INTRODUCTION

The World Health Organization estimates approximately 3 million people die annually as a result of ambient air pollution [44]. Exposure to fine-particulate matter ($PM_{2.5}$) – particles with diameters smaller than 2.5 microns – has the greatest adverse health effects and is linked with an increased incidence of cardiac arrhythmia, lung cancer, heart disease, and mortality [4, 36, 45, 46, 52]. Recent advances in sensor technology enable motivated citizens, grass-roots organizations, and researchers to bring new, low-cost, real-time sensors online [7, 29, 48], addressing the measurement gaps left by the sparse distribution of expensive, government-run sensors that fail to capture known microenvironments of $PM_{2.5}$ [5, 12, 53]. New, smaller sensors have improved our understanding of air quality conditions and sources of pollution [13, 21, 23, 35]. Although outdoor air quality is a growing concern for urban areas around the world, studies have found poor correlations between outdoor $PM_{2.5}$ levels and personal exposure measurements due to the large percentage of time that people spend inside [39]. For example, most Americans are estimated to spend upwards of 90 percent of their time indoors, with about 70 percent of their day spent at home [32].

Recent work has deployed low-cost air quality monitors and a basic visualization of the data to empower residents to visualize the air quality in their homes [16, 26, 29–31]. This foundational work has demonstrated the value of deploying an air quality monitor in a residential environment: study participants report being more aware of the air quality in their homes and more engaged in its management. However, these same studies also identify some limitations of single-monitor residential deployments. One such limitation is a single monitor’s inability to identify the existence of indoor and outdoor microenvironments. These micro-environments require multiple monitors in order to reliably detect and characterize $PM_{2.5}$ variability [40, 50], as well as to understand the effect of outdoor conditions on indoor air quality [3, 33, 34, 41]. Another challenge in these recent studies has been in reviewing and interpreting air quality sensor data, which can rapidly become unintelligible without annotations capturing potential sources of poor air quality. However, user-driven labeling of air quality events can be challenging because changes in $PM_{2.5}$ levels are often not detectable by humans at the time, and it is difficult to remember what was going on when revisiting data to annotate after the fact. These issues have limited the utility of residential air quality sensing deployments.

The goal of this work is to capture the additional value that users find when they have access to an air quality monitoring system that collects data from multiple monitors, supports proactive and in situ annotation of that data, and presents real-time air quality data and annotations in a interactive visualization. To accomplish this, we developed MAAV, a system to **M**eaure Air quality, **A**nnotate data streams, and **V**isualize real-time $PM_{2.5}$ levels. MAAV includes multiple $PM_{2.5}$ monitors placed both inside and outside a home to capture $PM_{2.5}$ microenvironments; three different annotation modalities, including a system-initiated prompt, to enable residents to contextualize data streams; and a tablet-based interactive visualization for exploring measured $PM_{2.5}$ levels and annotations.

We deployed MAAV to 6 families over a period of 20-47 weeks (mean 37.7 weeks). Over this time, we conducted 34 interviews with participant families to understand their experience with MAAV. Results showed that **1.** MAAV’s multiple monitors helped participants make meaningful between-monitor comparisons; **2.** the availability of multiple annotation modalities led participants to generate many annotations throughout the deployment, sustaining long-term engagement; **3.** annotations and the interactive visualization supported participants to draw insights from their data; and **4.** participants remained engaged with MAAV over a long period of time, but that initial active engagement gave way to a “maintenance mode”. These results yield insights about new types of non-temporal comparisons, the value of system-initiated annotation prompts, and the potential for the system interface to “evolve” over time.

2 RELATED WORK

A significant amount of work explores the technical aspects of air quality sensing from various perspectives and application areas: commercial monitoring and HVAC control [7, 25, 27], system infrastructure and platform development [8, 26, 28, 47], mobile sensing [13, 21, 26], personal exposure monitoring [26, 42], and source detection and classification [16, 49]. These studies primarily focus on a prototype's technical contributions or proof-of-concept systems architecture [7, 8, 21, 24, 25, 28, 47, 49]. Early work by Postolache et al. [47] proposes a multi-monitor air quality sensing system, with significant work on modeling, calibrating, and processing sensor data to ensure an accurate system. Jiang et al.'s *MAQS* system presents CO₂ exchange rate and n-gram augmented Bayesian room localization models to estimate indoor air quality and personal exposure, respectively [26]. Both projects develop purpose-built systems for addressing research questions, but they do not explore the needs or questions of the end-users.

The increasing availability of low-cost, commercially available PM_{2.5} sensors enables recent work to focus on human-centered aspects of air quality sensing in the home. Early work by Kim et al. [29–31] deploy the *inAir* system to support residents in understanding and managing potential air quality health threats. Using a single PM_{2.5} monitor and ambient display, *inAir* succeeds in increasing participants' awareness of and reflection on indoor air quality, with an improved ability to make connections between in-home activities and air quality levels. Although visualizing indoor air quality for the user, the system's ambient display did not support user-interactivity or direct data annotation, making it difficult for residents to contextualize events and reason about their causality. *inAir* also elicits feelings of powerlessness and frustration because residents can not determine the sources of PM_{2.5} spikes in their homes [29]. Furthermore, single-monitor deployments prevent participants from knowing the air quality within multiple indoor areas or the relationship between indoor and outdoor air quality in their homes [30].

Subsequent work demonstrates the ability to automatically classify PM_{2.5} sources within a small set of detection categories. Fang et al. [16] address the issue of PM_{2.5} source classification and targeted feedback with the use of a machine learning model. Their *AirSense* system is able to automatically detect and identify three sources of indoor pollution: cooking, smoking, and spraying pesticide. Their system also estimates personal exposure and provides actionable suggestions to help people improve their air quality [16]. However, similar to *inAir*, this system also fields only a single monitor and does not address issues related to source localization, outdoor air quality, air quality in other parts of the home, or contextualization of collected data.

For those collecting and consuming air quality sensor data, its lack of context can complicate its interpretation. Work by Tolmie et al. [54] identifies *why* annotation is important by demonstrating the large amount of articulation work that goes into interpreting raw sensor data. They conclude that "data generated through [a] networked sensing system is opaque when considered in isolation". These findings highlight the importance of supporting annotation in sensing systems not just to improve sensemaking, but also to reduce the amount of misplaced effort and resources related to incorrect data interpretation. Yet, no systems support this low-barrier creation, overlay, and review of annotation to contextualize the data. We incorporate these findings into MAAV and support multiple integrated annotation mechanisms to allow participants to label and contextualize their indoor air quality data in situ as it is collected.

Very limited work has been done to explore the nature of participant engagement with air quality sensing systems over longer term deployments. Kim et al. finds participants' degree of engagement remained constant in separate 4-week [30] and 4-month [31] studies, although these results were based on self-report, and the broader work primarily focuses on the quantifiable air quality improvements via participants' behavior change – not the evolution of their interactions with the system over time. Fang et al. similarly finds that hosting an air quality sensing system improves participants' sense of awareness and competence over 6- and 10-week deployments, but their study did not explore the nature of participants' engagement or use of the system beyond this period.



Fig. 1. Elements of a MAAV deployment. From left to right: wireless Dylos air quality monitor, Tablet display, Google Home and Raspberry Pi gateway computer (front). Wireless connectivity allows distributed instrumentation throughout participants' homes. Each deployment received three air quality monitors but is capable of instrumenting many more.

[16]. Our work outlines shifts in participant engagement over long-term deployments ranging between 20 - 47 weeks (average 37.7 weeks) relating to changes in curiosity, familiarity, goals, and seasonality.

Our focus on end-users and data legibility is in line with a variety of research conducted broadly in the space of eco-feedback, not just for air quality. For example, past work discusses the importance of *disaggregated* measurements [19] to support a finer granularity of resource usage akin to multi-monitor deployments for measuring air quality in different parts of the home. Past work has also explored the importance of technology probes and display designs in the area of water conservation [18], specifically emphasizing views that support different levels of data and time granularity. This same work prioritizes data comparison for highlighting relative differences and deemphasizing absolute values and units. MAAV similarly supports participants to make relative comparisons between monitors and over different granularities of time. Finally, research has shown that systems providing clear, specific, and frequent feedback via computerized and interactive tools successfully evoke lasting engagement [17]. Each of these pieces – interactivity, annotation, comparative views, and active feedback – come together in MAAV.

Building on this breadth of work, we deploy a flexible and open-ended technology probe utilizing multiple wireless $PM_{2.5}$ monitors, a host of annotation mechanisms, and an interactive tablet interface to focus on participants' goals and use cases of such a system in the home. This work complements the previous *inAir* [29–31] and *AirSense* [16] systems by exploring the ways participants use a multi-monitor annotation-enabled system, the questions they develop, and how these interaction mechanisms change over time. To the best of our knowledge, this is the first study combining these system features, as well as the longest running indoor air quality study with persistent deployments.

3 MAAV: A MULTI-MONITOR AIR QUALITY SENSING AND FEEDBACK SYSTEM

Drawing insights from past work, we developed a system to **M**easure Air quality, **A**nnotate data streams, and **V**isualize real-time $PM_{2.5}$ levels, which we call MAAV. MAAV consists of multiple low-cost air quality monitors, a gateway device to upload air quality sensor data to our back end server, mechanisms for users to annotate collected sensor data, and a tablet-based visualization that shows collected sensor data and annotations (Figure 1). This project is part of a larger multi-institutional effort to develop tools for epidemiologists to collect data and

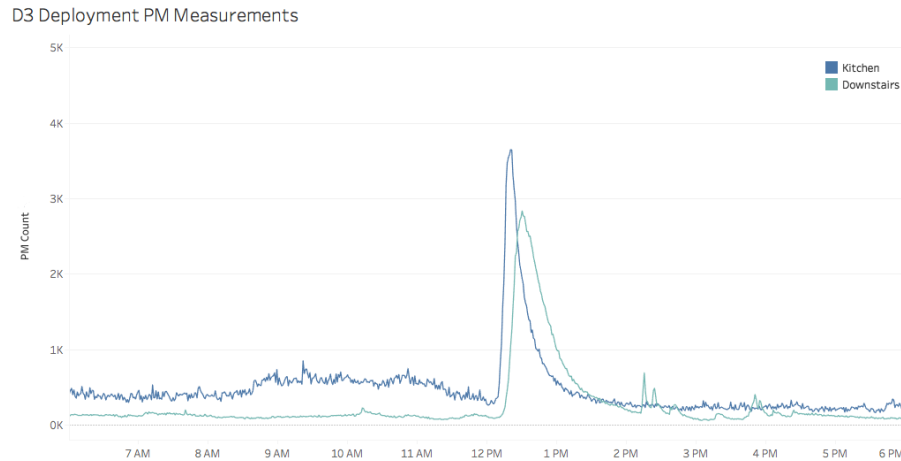


Fig. 2. $PM_{2.5}$ cooking spike in P3's kitchen. Note its diffusion and registration in the downstairs monitor a short time after. Both rooms return to baseline levels after approximately two hours.

conduct research on the relationship between air quality and asthma. The following subsections detail each of MAAV's components.

3.1 Low-Cost Monitors and Gateway Computer

MAAV brings together air quality data from multiple $PM_{2.5}$ monitors. We use Dylos air quality monitors¹ for their sensitivity at low $PM_{2.5}$ levels and ability to detect the full range of various household spikes (Figure 2). These detectors operate via optical light scattering: ambient air is drawn in through the device and across a laser, where an optical sensor measures the amount of light scattered by suspended particles in the airflow. This detected level is output as a $PM_{2.5}$ count, averaged over a 60-second window and logged each minute.

To integrate the stock air quality monitors within MAAV, we modified each with a BeagleBone Black embedded computer² for data formatting and wireless network connectivity. We also installed an improved RGB LCD for communicating system status and measurements to participants and technicians. Once running, the monitors log their measurements to internal memory and transmit them to a Raspberry Pi running Home Assistant³, which acts as the local gateway computer.

Dylos monitors come precalibrated from the factory and have been found to closely track laboratory-grade air quality instruments [55]. We validated this finding with a week-long pilot study involving 10 Dylos monitors colocated with several calibrated, laboratory-grade air quality monitors^{4,5,6} in the first author's home. All monitors were subjected to a host of domestic activities: lighting and extinguishing candles, vacuuming, changing bed sheets, and doing laundry, the results of which verify the monitors' measurement accuracy against calibrated air quality monitoring hardware [22]. During the study, a trained research assistant periodically cleaned deployed air

¹Dylos DC1100 Pro - <http://www.dylosproducts.com/dcproairqumo.html>

²<https://beagleboard.org/black>

³<http://hass.io>

⁴GRIMM 11-B: <http://wiki.grimm-aerosol.de/index.php?title=IAQ-11-B>

⁵Dustrack: <http://www.tsi.com/DUSTTRAK-DRX-Aerosol-Monitor-8533/>

⁶minivol: <http://www.airmetrics.com/products/minivol/index.html>

quality monitors following the manufacturer’s guidelines for maintaining the board-level air quality sensor but did not recalibrate. To ensure Dylos monitors remained within a satisfactory operational envelope throughout their deployment, we logged a variety of diagnostic and monitor health data, including case temperature, fan speed, and network statistics. Although the potential for sensor drift exists, any such long-term changes were overshadowed by $PM_{2.5}$ spike dynamics and other events examined in this work. Furthermore, the qualitative and human-centered approach of this study minimizes the need for quantitative metrics, instead focusing on the impact of significant AQ changes in the home.

All air quality measurements and participant annotations recorded with this system are saved on a HIPAA-compliant server. Back end services further monitor, analyze, and alert participants based on their indoor air quality. More information on data aggregation and processing can be found in Lundrigan et al. [38], and our Dylos testing procedures are outlined in Hegde et al. [22].

3.2 Contextualizing Air Quality Monitor Data with Annotation

Past work on data legibility has shown that “personal data generated through networked sensing systems are *opaque* when considered in isolation ... [and] that fine grain understandings of interaction cannot be ‘read off’ the data alone.” [54]. In the context of indoor air quality systems, Kim and Paulos [30] find that such a lack of contextual awareness can produce feelings of frustration and powerlessness in end-users and, ultimately, distrust in the system in severe cases. Annotation is therefore essential not only to improve data legibility but also to empower users and help legitimize deployed systems. Finally, the act of annotating can encourage users to reflect [10] on the connection between their daily activities and their indoor $PM_{2.5}$ levels.

The open-ended nature of daily activities that can affect air quality requires manual, user-provided annotation. However, manual annotation can also present a major burden for users. In an attempt to increase the convenience and reduce the burden of manually annotating, MAAV supports multiple annotation modalities: a tablet-based mechanism incorporated into the visualization interface; a text messaging prompt based on automatic detection of $PM_{2.5}$ spikes; and voice transcription via a smart speaker. We also offered participants the ability to manually record their activities in Google sheets as was done in previous studies [16], but all participants declined. Annotations are stored in their own database, along with annotation modality origin, logs of $PM_{2.5}$ spike activity, deployment spike notifications, and all relevant timestamps. Each annotation modality offers different trade-offs to support different situations where users may want to annotate their data.

3.2.1 Tablet. Users can record tablet annotations with a long-press at the desired location within the interface’s main view area (Section 3.3). A pop-up window displays the selected date and time, along with a textbox for typing the annotation. When submitted, descriptions appear as interactive glyphs on the visualization (Figure 3e).

Tablet annotation is effective in situations where users want to annotate an event when reviewing their air quality data in the tablet interface. This modality is not always convenient, however, such as when the tablet is not close at hand.

3.2.2 Text Messaging. MAAV incorporates text messaging⁷ to prompt users for an annotation when a $PM_{2.5}$ spike occurs. An online peak detection algorithm [1] evaluates incoming $PM_{2.5}$ measurements, and those surpassing independent rolling average and standard deviation thresholds are written to the database and generate a text message to the participant, including the spike’s location and time: “We detected a PM spike in the bedroom at 10:56 AM. Any idea what caused it?”. Users’ replies are logged as annotations and associated with the time the $PM_{2.5}$ spike was detected. Participants can also text the same phone number to record an unprompted annotation. Text prompts are disabled during participant-specified quiet hours or within a defined proximity to other detected peaks. For example, if a participant requests quiet hours from 9:00 PM - 9:00 AM, then MAAV will not send

⁷<http://www.twilio.com>

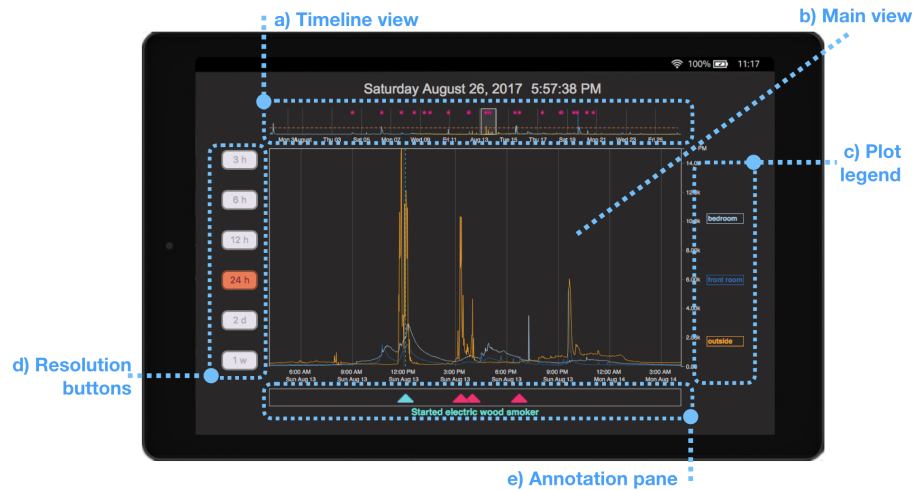


Fig. 3. The tablet visualization interface. This image shows a number of indoor (blue) and outdoor (orange) spikes in the main view related to $PM_{2.5}$ measurements over a 24-hour time span. The annotation pane shows that the earliest and largest spike is annotated with the participant message: “Started electric wood smoker”. These annotations lend context and legibility to the air quality dynamics shown on the interface. The timeline view shows that this is one of many such contextualizing annotations provided by the participant over their deployment.

text notifications for $PM_{2.5}$ spikes within this time range. Similarly, participants requesting no more than three notifications per hour will have messages rate-limited to a 20-minute wait time. The system uses detection time to determine whether to actively prompt a participant to annotate.

Text messages are an obvious choice for delivering a prompt to annotate air quality data. Daily activities that may have caused $PM_{2.5}$ to spike can be easily forgotten. The immediacy of receiving a text message and the convenience of responding to that message with a few short words offer users a high-value, low-burden solution. Prompts can be responded to in the moment or at a later, more convenient time. On the other hand, it can be useful to view $PM_{2.5}$ data while annotating, which is not easily supported by text messages. Additionally, past work has shown that people do not necessarily have their phones close at hand while at home [11].

3.2.3 Voice Annotation. As a complement to tablet and text modalities, MAAV also supports voice annotation. MAAV leverages Google Home and DialogFlow to support users in dictating their annotations with a custom command: “OK, Google ... Tell MAAV [annotation text]”. Their annotation text is then transcribed and logged to the database.

Voice annotations are especially useful when users are unable to type – such as when preparing dinner or cleaning. Participants can dictate their annotations in these conditions, either in response to a text notification or as stand-alone annotations. Similar to tablet annotations, physical proximity is also necessary to engage with this modality, which can limit its utility and use-cases. Also, similar to text message annotation, it is not possible to review $PM_{2.5}$ data when providing voice annotations.

3.3 Tablet Visualization

MAAV uses an interactive tablet interface to allow users to engage with their $PM_{2.5}$ data and annotations (Figure 3). We designed the interface to default to a glanceable, always-on visualization that shows the most

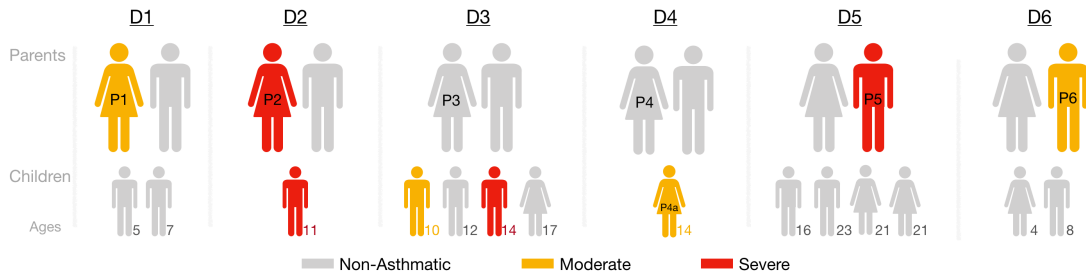


Fig. 4. Participant family structures for deployments D1 - D6. Yellow and red figures represent moderate and severe asthmatics, respectively. Remaining figures are nonasthmatic.

recent 24 hours of data in a line chart (Figure 3b). However, the visualization is also interactive and designed to support users in reviewing, exploring, and annotating the data. Participants can change the main view (Figure 3b) by:

- zooming in or out on the data at predefined intervals by tapping a button on the left (Figure 3d)
- scrubbing across the top timeline view to look at data from the previous 30 days (Figure 3a)
- toggling the line for a monitor on or off by tapping that monitor's name in the legend (Figure 3c)

The visualization also shows annotations the user has provided for the data, represented by pink triangles underneath the main view (Figure 3e). The user can tap on a triangle to view the text of that annotation. The timeline view (Figure 3a) also shows pink asterisks at the top where there are annotations. As described in Section 3.2.1, the user can also long-press on the main view to provide a new annotation. Figure 3 illustrates the visualization's interactive components.

This interface was implemented on an Amazon Fire HD8 tablet for its low cost and customizability [2], using JavaScript and D3.js [6], and run in a kiosk browser⁸ to provide a fixed, full-screen environment. We also incorporated Google Analytics to capture participants' interactions with the interface throughout their deployments.

4 FIELD DEPLOYMENTS

Prior work shows that single-monitor indoor air quality systems with basic visualization capabilities improve participants' air quality awareness and increase their engagement [29–31]. Building on this, we sought to better understand how users would engage with MAAV's interactive visualization and rich annotation capabilities to support them in characterizing their home spaces. We deployed MAAV to six families for a longitudinal study and conducted a series of interviews along the way to understand their experience. By recruiting entire households, our goal was to observe ways in which the entire family's engagement in and awareness of their indoor air quality changed during the deployments. These field deployments were staggered across Summer 2017 into Winter 2018 as participants were available and able to host MAAV.

4.1 Study Participants

We recruited eight households from a concurrent university-run medical study for families in which at least one family member has moderate or severe asthma [43]. To be eligible for our study, families needed a high-speed Internet connection and a wireless home network. They also needed to be willing to host a year-long system deployment. From these 8 families, 6 persisted throughout the study. Two families dropped out before we began

⁸<http://www.ozerov.de/fully-kiosk-browser/>



Fig. 5. The MAAV study timeline. Each phase corresponds to participants' level of interaction with the system: Phase I - no data provided (Passive), Phase II - interactive feedback (Active), Phase III - routine interaction (Long-term deployment)

our deployments, citing measurement fatigue from the parent study. We labeled these remaining deployments D1 - D6. Although we encouraged the entire family to participate throughout the study and interview process, we did not require their complete involvement. Each of the six deployments settled into a pattern with a singularly motivated participant who assumed the primary communication and feedback role. We denote these participants as *primary participants*, labeled P1 - P6 according to their deployment. Across all deployments, the primary participant was either the enrolled participant from the parent study (P1, P5, P6) or the primary care giver to the child who was enrolled in that study (P2, P3, P4). We also included P4's teenage daughter, P4a, as she was sufficiently involved in the deployment and interview process. None of the other children participated in the Phase I & II interviews.

Figure 4 illustrates the family make-up, distribution of asthmatic participants, and ages of the children in the long-term deployments. In deployments D1, D5, and D6, the primary participants (P1, P5, and P6) are the only asthmatics. Deployment D2 is the only deployment with joint parent, child asthmatics. In deployments D3 and D4, P3 and P4 are each primary caregivers to asthmatic children.

Participants' asthma triggers were individually diverse, but all asthmatics reported experiencing their worst symptoms during the winter months, especially because Salt Lake City often experiences periods of sustained moderate to severe air pollution during these times. No participants had prior experience with commercial air quality monitoring systems.

4.2 Phased Deployments

Participants received MAAV in a phased roll out to habituate them to hosting the system (Figure 5). These phases are delineated by participants' interaction with MAAV over the deployment: Phase I - passive deployment with no interaction (6-13 weeks), Phase II - active deployment with high interaction (2-6 weeks), and Phase III - long-term deployment with routine interaction (4 - 38 weeks), for a total operational time of 20-47 weeks (mean 37.7 weeks, Table 1). Each of the three phases began with user interviews to understand participants' air quality awareness (Phase I, predeployment interview), expectations for MAAV (Phase II, deployment interview), and experience with MAAV (Phase III, postdeployment interview). Although we encouraged all family members to participate, the majority of deployment families seemed disengaged and did not participate. These interviews consequently focused on the primary participants. To better understand the (lack of) motivation for the remaining family members, we conducted a fourth engagement interview with each parent of each family in an interview, as well as two additional interviews each with children from D3 and D5 to better understand reasons for their lack of engagement. In all, we conducted 34 interviews with over 20 hours of interview audio (Table 1). All interview guides are available in this manuscript's supplemental material.

4.2.1 Phase I - passive deployment (no interaction). Phase I began with installing air quality monitors in participants' homes to habituate them to hosting the hardware. Each deployment received three air quality monitors and a small gateway computer for gathering, formatting, and packaging air quality measurements. A trained research assistant placed the monitors following the parent study's established protocol to capture the

Table 1. Top: Phase durations for deployments D1-D6 (in weeks). Bottom: Number and length of participant interviews for each deployment. Counts include predeployment, deployment, and postdeployment interviews with the primary participant, plus separate engagement interviews for both the primary and non-primary participant(s), netting five interview sessions per deployment. Deployments D3 and D5 had an additional two interviews each from separate non-primary engagement interviews with their teenage children.

Deployment Duration	D1	D2	D3	D4	D5	D6	Avg.
Phase I	6.3	7.3	7.1	9.7	12.7	5.9	8.2
Phase II	3.0	4.1	2.0	2.6	4.1	6.3	3.7
Phase III	37.9	35.7	34.9	34.7	4.0	7.9	25.9
Total	47.1	47.1	44.0	47.0	20.9	20.0	37.7
Participant Interviews							Total
Number	5	5	7	5	7	5	34
Total Length (mins.)	139	192	160	245	230	258	1224

majority of possible indoor activity: 1 outside, 1 in a common living space (kitchen or living room), and 1 in the asthmatic's bedroom. To measure air quality within the breathing zone, she placed indoor monitors anywhere between tabletop and head height, approximately 30 - 60 inches, and away from vents, obstructions, or other environmental obstacles for accurate measurements. Outdoor monitors were mounted shoulder-height on the home exterior by the front door and placed inside a vented, non-temperature-controlled enclosure to permit sufficient airflow while keeping precipitation out. These monitors gave participants insights into the air quality around their home as well as providing a comparative data stream when understanding and assessing indoor air quality.

Once instrumented, we began collecting deployment-specific baseline data for use in the deployment interview. We disabled the monitor's on-board LCD display to not influence participants during this initial phase. Next we conducted the predeployment interview, with questions relating to the participants' awareness, understanding, and perception of air quality, both indoors and outdoors.

Phase I lasted 6-13 weeks (mean 8.2 weeks) and ended when we deployed the visualization and annotation modalities during the deployment interview. This deployment interview included an in-home demonstration of MAAV's interface, interactivity methods, and various annotation options. Participants saw the month of data that had been collected for the first time during the demonstration, which motivated a personalized explanation of their indoor air quality and exploration of the interactive tablet interface. We also conducted a separate data literacy exercise to gauge participants' ability to interpret time series information. Each participant understood the data abstraction, and none had any difficulty using the tablet interface to answer questions about the data.

We conducted the interview in tandem with the system demonstration in order to capture the participants' reaction to seeing their data for the first time, along with their expectations of how they would use MAAV. When demonstrating the annotation modalities, we stressed that annotation was completely optional and not a requirement for participating in the study. The Google Home and text message annotation modalities were optional; however, all participants chose to receive both modalities. We also purposefully did not provide any information on ways participants could improve their air quality to increase the likelihood that changes observed in subsequent interviews could be attributed to their use of the system.

4.2.2 Phase II - active deployment (high interaction). Phase II began after the deployment interview when participants were able to annotate and explore their air quality data. Phase II ran for a period of 2 - 6.3 weeks (average = 3.7 weeks) subject to participants' availability. During this phase, participants were free to use and interact with MAAV at their convenience. This deployment phase ended when we returned to conduct our postdeployment interview.

During the postdeployment interview, we focused on how participants used the system and on any insights or outcomes as a result of exploring their air quality data and making annotations. We asked additional questions for comparison against previous expectations relating to indoor air quality variability, system sensitivity, air quality awareness and engagement, and overall reflection on system utility. At the end of this interview, we gave each participant the option to return some or all of MAAV. Only P4 opted to return anything: Google Home.

4.2.3 Phase III - long-term deployment. Phase III began after the postdeployment interview, with the goal of leaving MAAV in place to capture the evolution of the participants' opinions and interactions after several months of familiarity. Phase III is ongoing and has run for a range of 4 - 38 weeks (average = 25.9 weeks) at the time of this writing.

We conducted a follow-up engagement interview 4 - 35 weeks (average = 23.3 weeks) after the participants' postdeployment interview to investigate whether or not they were still using MAAV and how their use had evolved since they first received it. We also questioned the remaining non-primary participants to gain insights into their perceptions of the system, its relevance to them, and mechanisms for their disengagement.

4.3 Qualitative Data Analysis

We transcribed and analyzed all interviews via case study analysis [20] and thematic analysis [9] to distill themes across the entire data set.

This joint approach allowed us to identify and analyze within-interview-round themes between individual deployment phases (thematic analysis) and within-participant themes across all deployment phases (case study analysis).

We conducted the first round of qualitative analysis with data from participants P1-P4 and proceeded by deployment phases. One researcher open-coded initial interviews to characterize participant replies, and two additional researchers then joined to collaboratively refine these codes, developing a set of closed codes. The original researcher then applied these closed codes to the remainder of the interviews in that phase and repeated this process for each phase. We met again at the end of coding each interview phase to iterate on codes and group them into higher level themes.

After completing our analyses with P1-P4, we enrolled P5 and P6 to validate our initial findings and verify thematic saturation. These additional participants received the same system through the same phases with identical interviews. This second round of qualitative analysis proceeded by coding P5 and P6 interviews using the closed codes identified in the first round, but with an eye to identifying anything new. No new codes were found, but the feedback and insights provided by these extra participants helped refine our findings. All resulting themes from the analysis are available in the supplemental materials.

5 FINDINGS

Results from deploying MAAV highlight the benefits of a flexible multi-monitor deployment and the diversity of tasks and questions it enabled participants to address. Findings in this section are deeply rooted in the particular context of each deployment. Drawing on interview data, we characterize each participant (see Table 2), which is useful for understanding and interpreting participants' perspectives throughout this section.

Table 2. Participant characterizations.

P1	Stay-at-home mom and adult asthmatic. Self-identifies as non-technical, passive, and reactionary with regard to asthma management. Admitted she was not overly engaged with checking air quality conditions, despite sensitivities. Primarily interested in using MAAV to inform her family's outdoor activities.
P2	Medically disabled from asthmatic symptoms. Non-technical, but engaged due to her and her child's more severe asthma symptoms. Actively seeks air quality forecast information through a combination of phone app and local forecasts. Curious what she could learn through using MAAV.
P3	Work-from-home mom, non-asthmatic. Two of four children are asthmatic: one suffers from Common Variable Immune Deficiency and asymptomatic asthma. Interested in using MAAV to experiment to find the cause of spikes and to monitor indoor air quality to guard against potential triggers.
P4, P4a	Nurse at a university hospital and caregiver to a teenage asthmatic daughter (P4a, student). Both were initially disengaged from air quality monitoring, but became more aware over the course of the deployment. Interested in drawing health correlations from visualizations.
P5	Public school administrator and severe asthmatic. Already aware of personal triggers, P5 was eager to use MAAV to explore spike sources and understand room activity levels.
P6	Public health employee and adult asthmatic. Motivated to use MAAV to characterize living space and improve personal sense of air quality control.

5.1 Multi-Monitor Deployments Enable Comparisons in and around the Home

In addition to showing the benefits from deploying a single monitor [16, 26, 29–31], MAAV supported participants in making comparisons between monitors. With three monitors, participants could observe and perform between-room and indoor/outdoor comparisons. These comparisons led to insights on home characteristics from observing $PM_{2.5}$ spike diffusivity (i.e., comparing two indoor monitors) and the ability to draw comparisons between indoor and outdoor monitors (Figure 6). Some participants found these comparisons so compelling that they asked if they could receive monitors for every room in their home. Similarly, when asked about their opinions on the adequacy of 3 monitors versus 1, participants unanimously preferred the flexibility of seeing multiple measurements at once.

P5: It was nice to see there are differences, like what's going on in different rooms. If there's just one sensor for an entire house, it doesn't really tell you anything.

5.1.1 Between-Room Comparisons. When air quality changes within a home, it does not happen uniformly. Instead, it is subject to characteristics of the home itself, such as air exchange rates and airflow between rooms. During the deployment, participants used MAAV to reason about ways that the home itself might influence air quality. For example, P2 noticed differences in how $PM_{2.5}$ spikes diffused from one room to another:

P2: It has surprised me to see that when [the living room] spikes, [the bedroom] tends to spike too. But when [the bedroom] has any activity [the living room] doesn't necessarily. I don't know if that's just the way the air flows in the house, or what.

Having multiple air quality monitors also enabled participants to detect air quality spikes that would otherwise be missed had there only been a single monitor.

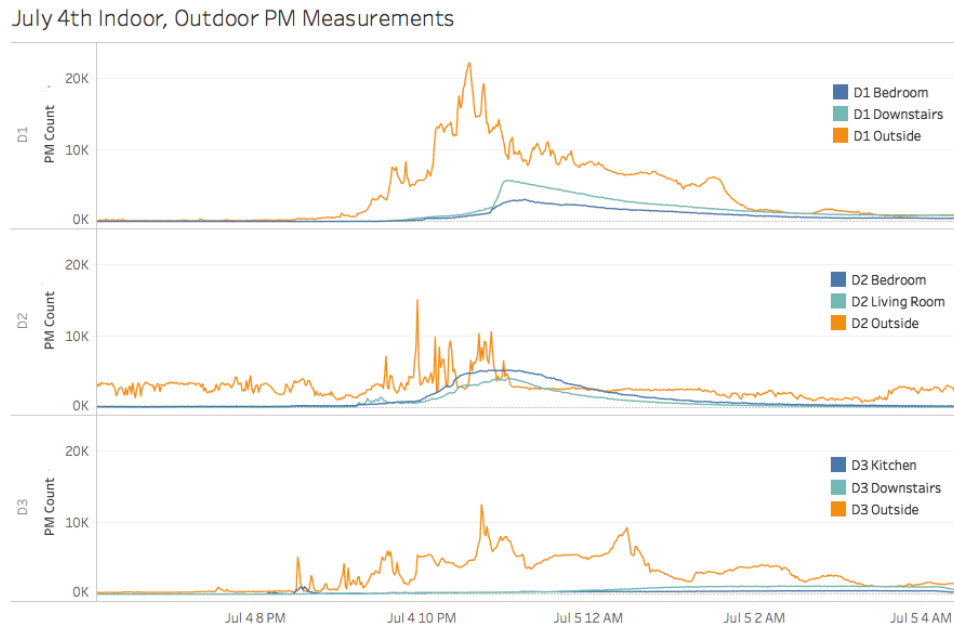


Fig. 6. Impact of July 4th fireworks on outdoor and indoor air quality. Indoor $PM_{2.5}$ readings noticeably increase for P1 and P2 whereas P3 shifts only slightly. In each case, it takes several hours for $PM_{2.5}$ readings to return to baseline levels. The readings from the outdoor monitors are different between deployments as well, highlighting the value of deploying monitors outdoors.

P1: I find it humorous that I can tell the days – when it’s a Wednesday, my husband works from home and lights incense in the office downstairs, and there’s a spike downstairs.

P3: I was definitely looking at the kitchen a lot because I think things change a lot more in the kitchen than the bedroom. Everything that happens in the house, happens in the kitchen. So it was always interesting when I saw a bedroom spike because ... what did they do? Seems really strange.

These situations where air quality events occur in localized spaces demonstrate microenvironments within the home, a phenomenon that can be observed only when there are at least two indoor monitors. Participants recognized these dynamics in their own living spaces, making observations and building a more nuanced understanding of how indoor air quality changes in the home.

P4a: It’s really interesting that we can compare [$PM_{2.5}$ spikes in different rooms] and see how some things that happen in different parts of the house can still affect other parts ... I think it’s interesting that if there’s big spikes outside, if you open the door – just a little bit – it affects the inside of the house. And it doesn’t affect it just while the door’s open, it kind of stays around for a while, because it’s been let in. I think that’s pretty interesting.

P4a’s experience highlights the importance of multiple monitors. With MAAV, she was able to reflect on the source of an indoor spike by relating it to the air quality outside, something that would not have been apparent from a single-monitor deployment.

5.1.2 Indoor/Outdoor Comparisons. Past work indicates participants in single-monitor deployments have wanted to relate their indoor and outdoor air quality [29], and the primary method for doing so has been to pull data from third parties, such as local air quality agencies [7, 8, 31]. Unfortunately, these external sources are typically government-run sensors operating in fundamentally different ways, and providing PM_{2.5} measurements at different time resolutions, spatial scales, and measurement units. It is therefore difficult to directly compare these third party measurements with a participant's own monitor. Moreover, many local activities (e.g., vehicular traffic, neighbors mowing their lawns, or barbecuing) do not show up at the municipal level, but they can directly influence air quality around a home. For example, Figure 6 shows very different outdoor readings across three different deployments during 4th of July fireworks.

MAAV's outdoor monitor lets participants directly compare indoor and outdoor measurements. They reported they were glad to have access to this information, and, in some cases, came to rely on this data in lieu of less representative public sources.

P6: I wanna see what's happening with all three sensors. The outdoor is usually the one that stands out. Historically, that's where the higher [readings are].

Participants in this study all had some concerns about indoor air quality because someone in their home – either themselves or their child – had asthma. Before receiving MAAV, their only options for managing indoor air quality were to react quickly when symptoms manifest, guess about sources of poor air quality and manage them, or follow generalized medical advice. For asthmatics, one common guideline recommends to stay indoors and keep windows closed, carrying the implicit assumption that the air quality indoors is generally better than outdoors. Despite this recommendation, past work has found indoor air quality can oftentimes be worse than outdoor [56].

Prior to receiving MAAV, participants had no way of verifying that indoor air quality was actually better at a given moment and whether the best course of action was to keep windows closed or air out their homes:

P2: Sometimes you feel that people are like, oh, you should stay in because the [outdoor] air's bad. But, is it really that much worse than what's in my house?

After receiving MAAV, participants made use of the indoor and outdoor monitors to track fluctuations and trends between the two environments. By facilitating comparisons between indoor and outdoor PM_{2.5} measures, MAAV empowered participants to evaluate the current situation, rather than depending on the generalized advice:

P3: I have learned to pay a lot more attention to the difference between the air quality inside and the air quality outside. I think previously, I just kinda assumed that the air quality inside was always really great, because our windows are shut and our doors are shut and we don't open our windows ... and that's what I was told: just don't open your windows and use air conditioning and you're good. But I realized that's not really always the case. And that's been a big eye-opener for me.

Whereas P2 and P3 were relieved to find their indoor air quality was often better than the outdoors, P4 was relieved to see that their outdoor air quality was oftentimes just as good.

P4: We open the windows all the time if the weather is nice, and I haven't noticed many spikes. That surprised me that there weren't that many.

P4's observation allowed her to relax from strictly following the generalized advice to always keep the windows closed. These findings indicate that having monitors both inside and outside of the home empowered participants to make informed, personalized decisions based on the unique environments in and around their homes. For many participants, the most compelling aspect of this comparison was the peace of mind in knowing that their indoor air quality was better than outdoors.

P3: I loved that it showed that the inside didn't get really bad. That's one thing I was looking for – how was the outdoor air quality affecting indoor air quality.

P6: The [greatest] value is knowing, “Hey, we're better than outside ... the air's better in here, guys”, and checking that every day and having that reassurance. I don't know if I expected that that would be my best value, or my most favorite value.

Section 5.1 Takeaway: *Participants made use of multiple monitors in their deployment to make comparisons between monitors, which supported new insights into the environment in and around their homes.*

5.2 Annotation Improves User-Experience, Engagement

The ability to annotate in MAAV enabled participants to reflect and reason about the ways their behaviors impacted indoor air quality. Figure 7 illustrates participants' various use patterns and annotation modality preferences along with the evolution of those behaviors longitudinally over the deployments. All participants noted that annotation was essential for helping them recall past events, make connections, and find patterns.

P2: I'm really glad those annotations are in there. It makes it easier to understand and put value on the data I'm seeing.

P3: I don't think this [visualization] would do me much good without annotations ... You have to have the combination of [data and annotation]; otherwise [the visualization] is kind of useless.

P4a: I look at annotations a lot. I like being able to see it. What I look at them for, I pay attention for if there's repeating things [patterns]. Today we played with the dog, and last week we played with the dog, but the spike is bigger.

Tolmie and Crabtree [54] found that raw sensor data is difficult to interpret without contextualizing annotations. By tightly integrating tablet, text, and voice annotation modalities into our system, participants could contextualize their data with personally meaningful annotations in whichever way suited their interaction preferences. Early versions of *inAir* [29, 30] provided a website for reviewing past data and logging activities via a calendar interface, although this manual form of journaling was not overlaid on the PM_{2.5} data. Despite the effort to contextualize raw sensor data, participants reported feelings of powerlessness and frustration over a lack of context for the current data they were shown [29]. By tightly integrating annotation methods and interactive visualization, participants were able to use MAAV to alleviate this shortcoming and make annotations available for review as they were captured.

Annotations were facilitated through *user-initiated* and *system-initiated* annotations. User-initiated annotations relied on action from the user to engage, such as with the tablet interface or voice annotation, whereas system-initiated annotations were directly generated by MAAV to prompt participants for action.

5.2.1 User-Initiated Annotations: Engaging through Tablet and Voice. Participants recorded *user-initiated annotations* through the tablet interface and voice modalities, often reviewing their data on the interface to explore the detected event, verify their annotation, and explore previous activities captured by past annotations. The combination of increased engagement and improved data legibility led participants to reason on the cause-effect relationships between their behaviors and the air quality in their homes. During the deployment interview, P3 speculated about the potential for MAAV to support self-experimentation:

P3: It would be nice to see if there's changes depending on when I have to use our giant oven versus the little convection oven and what difference that makes. I think I would definitely annotate those things as I start to make dinners.

Afterwards, P3 used MAAV to methodically test cooking behaviors and deduce sources of cooking spikes (Figure 2). When we returned for the postdeployment interview, P3 shared her process for annotating her cooking activities to determine which decisions influenced spikes – the combination of real-time sensing, longitudinal



Fig. 7. Long-term system interaction over phases II and III for each deployment. This plot illustrates the daily distribution of system-initiated notifications (black), received text annotations (light blue), tablet annotations (blue green), and voice annotations (blue), along with participants' interface interactions (light purple). Interaction patterns for P2, P3, and P4 show varied cycles of engagement, and patterns for P3 and P5 exhibit re-engagement with the system immediately after the engagement interview. P3 also shows patterns of making voice annotations when experimentally determining spike sources. All participants came to rely on text messaging for their primary annotation mechanism as the study progressed.

visualization, and annotation across multiple monitors enabled her to perform classic A/B experimental design and keep track of her test cases:

P3: If my theory is that it went really high that day because I used the stove top and used something that made a lot of smoke, then the next time I'm gonna do that, I'd start that and [annotate] "I'm doing this again," and then see what happened.

I figured out I get kitchen spikes when I cook with olive oil instead of avocado oil ... I don't use olive oil any more. I haven't had nearly as many kitchen spikes since then. Kinda cool.

Table 3. Deployment annotations per modality over phase II and III. The Annotation column is the sum of all annotation modalities. Participant annotations include those made in response to MAAV notifications as well as user-initiated annotations, exceeding 100% compliance rates for many deployments.

	<i>Phase II – Compliance</i>						<i>Phase III – Compliance</i>					
	Text	Tablet	Voice	Anno.	Notif.	(%)	Text	Tablet	Voice	Anno.	Notif.	(%)
D1	4	4	1	9	8	112.5%	47	9	0	56	110	50.9%
D2	16	5	3	24	18	133.3%	81	7	3	91	103	88.3%
D3	5	4	12	21	5	420.0%	81	8	20	109	98	111.2%
D4	7	5	2	14	14	100.0%	61	2	1	64	382	16.8%
D5	19	9	4	32	19	168.4%	8	11	0	19	7	271.4%
D6	31	1	1	33	36	91.7%	24	0	0	24	33	72.7%

P3’s annotation led her to conclude that cooking with olive oil contributed to kitchen spikes and that using avocado oil eliminated these spikes. This is another case where participants performed comparisons and reached conclusions that would not have been possible without the use of integrated annotation and visualization capabilities with the sensing system. Of further note, P3 used voice annotations for self-experimentation, which allowed her to continue cooking without taking the time to type an annotation.

5.2.2 System-Initiated Annotations: Push Notifications Engage Participants and Improve Awareness in Real-Time. One important feature for facilitating annotation in MAAV was the inclusion of *system-initiated notifications* via text alerts. These alerts actively engaged the users, prompting them to annotate the spike with a returned text message. Motivated by work on *just in time interventions* [51], the prompt system notifies participants as spikes occur in order to maximize the likelihood that they will be able to accurately annotate the data. Whereas all participants used more than one annotation modality during their deployment, text message annotation in response to notifications was the most common and consistent mode across participants (Figure 7).

P2: The thing I found most helpful were the texts. I love that it’s “hey, we had a spike” and I’m just like “this is what it is”. This is what I’m doing right now, or in the last 10 minutes. That’s my favorite feature. I love being able to see how much and what is causing spikes.

Beyond presenting an opportunity to contextualize air quality data, text alerts also became proxies for other indicators. In postdeployment and engagement interviews, P4 and P4a talked about using MAAV’s sensitivity to infer others’ daily schedules, such as leaving for work or getting home from school.

P4a: I’ll wake up to a text: “at this time in the morning there was a spike ... ” Oh, that’s when my parents went to work! Or when I get home from school, the rooms will spike.

P4: I have noticed a small pattern ... It’s a little more common to get spikes around 3:30pm.

P4a: That’s when I come home, drop all my stuff, and get food. I actually dance more until parents get home. I don’t dance when my parents are home.

5.2.3 Annotation Modality Influenced by Review Routine. Participants’ personal schedules, asthmatic symptoms, and personality helped influence the way they interacted with and annotated their air quality data (Table 2). Figure 7 and Table 3 show how participants used MAAV over time. P1’s outdoor sensitivities motivated her to regularly check the tablet interface to keep informed of her immediate outdoor air quality which, combined with her habit of reviewing her indoor air quality data at night, found her logging half of her annotations through the

tablet interface over phase II. P2 adopted a weekly interface review regimen, leading her to be more reliant on text notifications to gauge indoor air quality. Therefore, P2 relied on the receipt of text notifications to know when her air quality became worse. Consequently, she annotated much more frequently using the texting modality. P3 engaged in self-experimentation to understand the impact of alternative behaviors on her in-home air quality, causing her to rely on voice annotation for annotating her self-experiments.

Deployment number D4 was slightly unique from others in the study in that the primary participant (P4) was considerably less engaged with MAAV than other primaries in their deployments, which stemmed from two main factors. First, P4's job outside the home meant she was frequently away from the house when receiving text notifications. Her solution to this was to text P4a to determine what was going on at home. Second, the tablet interface was set up in P4a's room, meaning P4 had comparatively less access and opportunity to review the data than the primaries in the other deployments. D4 experienced an equal number of text and tablet annotations within phase II, but as the deployment progressed, they came to rely on text messaging as the primary mode of annotation.

Both P5 and P6 adopted an after-work routine to review their data and annotations out of curiosity and were equally engaged with MAAV, primarily relying on text messaging to maintain high compliance rates with system notifications.

Section 5.2 Takeaway: *Participants used different annotation modalities to annotate their data in different situations that seem to be tied to the affordances of that modality, indicating that the availability of multiple modalities led to more annotations.*

5.3 The Interactive Tablet Interface and Annotations Enabled Data Exploration

All participants frequently engaged with their tablet interface within phase II for reviewing and exploring their air quality data (Figure 7). Addressing prior findings that users wanted a larger snapshot of time when reviewing their air quality information [30], the tablet interface provides a one-month backlog of data to explore. Participants appreciated the ability to scroll back to review their annotations and prior $PM_{2.5}$ spikes, although most primarily used the interface to explore the previous 1-2 days.

P2: I'm more interested in what is happening right now, the last couple of hours, how it's affected (or might affect) what's going on today or tomorrow. I liked the idea of being able to go back a month or more, at first. I'm just finding I'm not using that.

P5: I was more just [looking at] current data, like the last day or two.

The ability to look back farther also motivated some to explore the full extent of their data.

P4a: I'd go back as far as it would let me. Most of the time I probably go 1-2 weeks back. Depends if I've looked at the spikes before ... Most of the time I'll just look over the past week. Especially the day that I'm looking at it and the day before, 'cause those ones are the most ... newer spikes that I haven't seen yet. Sometimes I'll go back and "what was I doing". Oh yeah, and I'll add an annotation.

Several participants reported not being interested in the value of the readings themselves and that seeing the relative dynamics was enough, echoing findings from design in eco-feedback [18]. Instead, they primarily focused their attention on the presence and distribution of spikes, along with their previous annotations.

P1: I'm definitely more of a visual [person], so seeing the spikes, to me, is enough of an [indicator].

P4a: Most of the time I'll just look at new spikes, but then I'll look at the little bar at top and "hey! there's a whole bunch of [annotations] here. I don't remember what this was" and I'll go back and look through it.

When reasoning about $PM_{2.5}$ spikes, some participants expressed interest in discovering and comparing patterns that MAAV's visualization does not make salient.

P1: It would be kind of cool to see “You’ve done these last 5 things cooking dinner. Here’s how they overlap and compare.” That would be an interesting thing.

This desire to view and compare groups of activities over time suggests that realizing the value of a longer data history might require alternative visualization and processing techniques beyond the pervasive time series view.

Section 5.3 Takeaway: *Participants used the tablet interface to review their annotations and air quality data, drawing insights that would not have been possible without annotated data.*

5.4 Participants Maintained Long-Term Engagement, but Changed That Use over Time

Many facets of user engagement and system integration cannot be fully observed or validated over a short-term deployment. For air quality, long-term effects of the changing seasons can have a significant impact on the nature of data shown to participants and, correspondingly, can affect sense of engagement. In this context, our long-term deployment provides a rigorous probe into the notion of disengagement by avoiding short-term effects or confounds. Figure 8 and Table 4 provide evidence for deployments’ continued involvement on the days where they received at least one text prompt, showing an 80% or greater compliance rate for 5 out of 6 deployments. We did not track non-notification days as we would not necessarily expect participants to interact with the system on days without notifications, which indicate few if any significant changes to the air quality inside their homes. By focusing on notification days, we can track the methods of participants’ responses, which serve as a reasonable proxy for user engagement. Although participants were primarily driven by curiosity to explore their data when first receiving the interface, their motivations evolved as deployments wore on. Mechanisms for this shift included an increase in familiarity with the data and a corresponding lack of surprises, but also evolving goals and seasonal impacts, as well as participants changing the questions they sought to answer.

P2: I used it more in the winter when we have such bad air quality [...] I’ve not been using it as much [recently], but mostly because as the seasons have changed we’re not having that terrible air anymore.

As their experiences grew, participants also began to engage the data with more specific questions, which our technology probe was not well suited to answer.

P2: If it were to be something that I could have a more personalized overview that I could take to my doctor, I would probably use it more and be more interested in keeping track of a daily “how I’m feeling” [...] To be able to connect health with the data is important.

It is important to note that these changes were not always a result of disinterest. P3’s interest in self-experimentation lent itself to periodic engagement (Figure 7).

P3: Once I figured some things out ... the kitchen thing ... I had a lot of fun trying to figure out why it would go off at various times in the kitchen. So I did a lot of experimenting early on, but once I had my questions answered it was good.

Regardless of these variations, participants’ text response rate was largely unchanged throughout the deployment and had become the preferred annotation method for most participants by the time of the engagement interview (Figure 7).

P1: Honestly ... I didn’t think the text messages would be as easy to respond to ... It’s a matter of, if it texts me right away, “hey this happened at this.” I can either go “Oh yeah, I know what that is”, annotate, done. That ease of [answering] even if I’m outside or [away from home] ... honestly, that’s what’s changed how much I’ve annotated, is just that ease of the text message.

P5: Phone texting was much more immediate and much more responsive. [texts] would draw out a response immediately that we would look into.

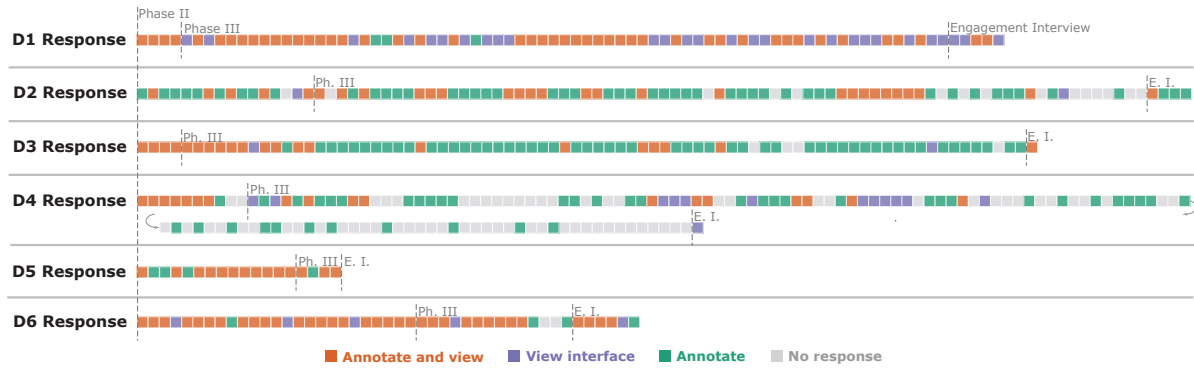


Fig. 8. Engagement chart for system deployments D1 - D6. Boxes represent individual days when a deployment received at least one notification from MAAV, indicating significant air quality changes in the home. We focus on these days as we do not expect that participants would necessarily interact with MAAV on days without notifications. Each box is color-coded according to the deployment's response to the notification on that day: viewing the interface (purple), making an annotation (green), both viewing the interface and annotating (orange), or not responding that day (gray). The sparsity of gray boxes illustrates participants' willingness to engage when we should expect it, serving as a proxy for their continued engagement. Outside of D4, which received significantly more notifications than other deployments, the remaining homes all had engagement rates above 80% (Table 4).

Not everyone remained engaged with text messaging, however. Both P4 and P4a reduced their annotations, but for different reasons.

P4: It just got tedious. I figured out that everything in the house makes this room spike. It doesn't matter what we're doing. But I just got tired of repeating the same thing over and over again.

P4a: When I first started [annotating], it was because I could annotate everything. I would still probably be doing that. I thought of it as "this is cool – this will help me realize what causes me to

Table 4. Engagement data from Figure 8. 'Notification Days' lists the number of boxes for each deployment and its corresponding percentage of deployment time. Participant response is further binned by 'Days Ignored' (gray boxes), and 'Days Engaged' (non-gray boxes). The final column is the percentage of notification days garnering a deployment response. Total engagement with MAAV was above 80% for 5 out of 6 total deployments for the entire study.

	Phase II & III Length (days)	Notification Days	% Days Notified	Days Ignored	Days Engaged	% Notification Days Engaged
D1	265	78	29.4%	0	78	100.0%
D2	250	95	36.0%	15	80	84.2%
D3	244	81	33.2%	4	77	95.1%
D4	243	144	59.3%	70	74	51.4%
D5	28	18	64.3%	0	18	100.0%
D6	55	45	81.8%	2	43	95.6%

stop breathing as much”. That was really exciting. I stopped annotating as much because [the tablet died].

Due to an uncommunicated hardware error, P4a was left without a functioning tablet at the end of their long-term deployment phase. P4a’s feedback on how the ability to review their data motivated annotation points to a synergy between annotating and reviewing data not found in extant literature.

For those who continued annotating, their messages became shorter as they defaulted to a subset of personally relevant keywords to describe tasks that set off text alerts.

P1: I would say it’s much more succinct. “Doing laundry”, “cooked dinner”, instead of being “I cooked dinner and there was burnt pizza on the bottom and there was smoke and blah blah blah”. They’re kind of more “oh yeah” reminders than “oh ‘this is this ...’ ” and I need to explain this whole situation. It’s definitely condensed down to keywords.

Fang et al. [16] have developed an air quality sensing platform for detecting and classifying three common sources of indoor pollution: cooking, smoking, and spraying pesticide. Although we found a moderate amount of overlap in participants’ keywords, we did not have instances of smoking or pesticide use. The remaining annotated activities, and participants’ convergence to a set of abbreviated keywords, point to a more diverse set of in-home activities capable of generating PM_{2.5} spikes. Moreover, arriving at these personalized, ecologically valid keywords takes time and would not necessarily be achieved within the shorter term deployments in past work.

5.4.1 Lack of Family Engagement. Maintaining deployments over many months allowed us to observe usage patterns and characteristics that would be challenging to discern over shorter term installations. One observation was how each deployment contained at most one engaged participant from the onset of the study.

In discussing their own senses of engagement, primary participants highlighted a division of labor or a sense of responsibility very early in the deployment and postdeployment interviews.

P3: Anyone in my family? Noooooooo. The reason is because I’m the one that takes care of the stuff. If they don’t feel responsibility to take care of it, then they don’t care.

This sentiment was shared by spouses:

P6 Spouse: I don’t really get involved – it’s his thing. He shares everything anyway, so I hear about it that way [...] my schedule is already pretty busy and this is just another thing.

P4 Spouse: Yeah, [poor air quality] sucks for some people, but it doesn’t affect me.

We were able to rule out insufficient communication or interview scheduling as engagement barriers over time, and conducted a dedicated fourth round engagement interview with the goal of understanding why certain participants were engaged and others were not.

When discussing engagement with non-asthmatic participants, we asked whether they could imagine circumstances where their level of engagement may have been different. Each cited their lack of asthma as a barrier. We also observed that age influenced a participant’s willingness to engage; other than P4a, no other teenage or younger participant was aware of, or concerned with, air quality issues.

Engagement interviews also identified inconvenience and unintentional gate-keeping as other factors affecting people’s ability to engage with the interface. By coordinating interviews with a single family contact, in some deployments it came to be seen as “*their thing*”, or “*the study’s*”, which limited others’ sense of responsibility.

P5: Putting it over there so that my family didn’t break it probably didn’t help them to be more interactive with it.

Family dynamics are complex and vary between households. It is possible that with targeted effort we could have seen more engagement by multiple family members, but their reactions here suggest that, left to their own

devices, they may also naturally settle into a routine where only one person is primarily concerned with their air quality monitoring system.

5.4.2 Participants Converged to a Single Interaction, Annotation Modality over Time. As tablet use decreased, participants relied more on text alerts to annotate and gauge their indoor air quality:

P6: If it sucks enough for you to message me, that's going to indicate that something is up.

When asked what motivated them to continue annotating, primary participants cited text messaging's low barrier to responding and the importance of contextualizing their data. Some also mentioned a hope of receiving more accurate medical recommendations or individually tailored air quality advice as a future study benefit. Mobile interaction proved popular with participants, and although feedback during postdeployment interviews indicated that participants enjoyed having a dedicated tablet for the visualization, by the time of the engagement interview, each primary participant – and even some non-primary participants – identified that a mobile app would be more useful and would improve their engagement.

P1 Spouse: If it had been an app rather than in one place in the house, I could see being a lot more involved. Have the sensors running here [home], then I can pull it up at work, and see how things are at the house. It would become a thing like checking weather, I think. Before work, I check the weather on my phone. I check the air quality around the house just the same.

P6: If I had an app that would show me the exact same thing, I would use that app a ton.

Participants' convergence toward text notifications as their primary annotation and interaction modality lends strong support for the power of system-initiated prompts delivered with a mobile device for long-term engagement (Table 3).

Section 5.4 Takeaway: *Participants changed how they engaged with MAAV over time, transitioning from an initial phase of deeper regular engagement to a maintenance mode, while their family members remained mostly disengaged throughout.*

6 DISCUSSION

In our long-term deployment of MAAV, participants exhibited several behaviors not observed (or possible) in previous work. These behaviors included drawing insights from multi-monitor comparisons, leveraging annotations to facilitate both awareness and sensemaking, and changing (while still sustaining) their use of the system over time. These results have several important implications. First, annotations have the potential to facilitate new types of comparisons in users' data. Second, the surprising long-term adherence to our mobile phone push notifications offers a promising direction for data labeling and collection in future studies, and perhaps in different contexts. Third, participants' shifting usage of MAAV over time suggests opportunities for a user interface that supports different stages of system use. We discuss each of these implications in detail.

6.1 Beyond Direct Temporal Comparison

Participants in this study were able to view synchronized data from multiple monitors along with annotations, and they were able to use that representation of the data to draw insights that would not have been possible without multiple monitors or without annotations. Integrating annotation into MAAV supported participants in making the data legible [54]. When reviewing their data, they were less interested in scales, magnitudes, and baseline fluctuations communicated with the time series representation, and instead they focused on the locations and distributions of $PM_{2.5}$ spikes. Seeing their data annotated with their activities and routines allows participants to go beyond the numerical representation and begin to reason about these impacts at a higher level than would be possible without the context afforded by annotated data.

Over time, participants wanted to go beyond the comparison capabilities supported by MAAV. Through our interviews, we observed participants thinking across time and wanting to compare relationships between temporally disjoint events, e.g. compare the five most recent times I vacuumed. The fact that participants requested this type of feature demonstrates the value of annotations and of collecting long-term data logs. Supporting such user-driven *cuts* [14] requires re-thinking annotation logs as the current annotation mechanism tags single points in time, though the tracked activities may describe events taking place over a duration. At a conceptual level this type of comparison interface is straightforward to describe, but the implementation details it are less straightforward. For example, developing methods on how best to define, capture, and align a collection of PM_{2.5} events, and the corresponding semantic rules or computational metrics for comparing event similarity are much more involved. Although neither MAAV nor other interfaces in the air quality literature support these temporal cuts, future systems should consider event tracking and processing techniques for supporting comparisons across time, place, and activity.

6.2 System-Initiated Notifications Sustained Engagement

From the perspective of sustained participant engagement, much of the success of MAAV seems to come from the system-initiated text message prompts. Figure 8 shows the majority of participants remained engaged with MAAV throughout their deployment, with engagement rates over 80% on days they received notifications, for 5 of the participants out of 6 total. During interviews, participants indicated that the text messages were useful as a proxy for gauging indoor air quality and that the messages were lightweight and easy to respond to. Yet for P4, who received comparably more notifications than other deployments, these messages led to frustration over their repetition (Table 3). As an outlier, P4's experience points to a possible upper bound for user notifications, in other words, participant engagement may decline or drop off entirely over some amount of sent notifications (Figure 8). Based on P4's notification-fatigue, it may be preferable to allow participants to selectively mute specific indoor monitors or have control over their system's thresholding and rate-limiting parameters. Another possibility is to analyze spikes in the context of past spikes and user engagement to dynamically adjust which notifications to deliver. For example, multiple spikes that seem to be similar might be bundled together, or notifications might be reserved for spikes that appear to be sufficiently different from past spikes.

System notifications make people aware of changes in their environment in a way that is easy to miss with an ambient display. The next-longest longitudinal indoor air quality study [31] utilized ambient displays, requiring those participants to be in front of the monitor to witness spike activity, which even then showed only the most recent four hours of data. Thus, these results are new and novel for the way that they directly elicit and probe user engagement.

As an annotation prompt technique, text messaging was successful at keeping participants engaged with MAAV and thinking about air quality without requiring them to proactively and regularly check the monitor. The fact that our participants sustained their engagement, measured by response rate to text message prompts, over a long period of time suggests this level of engagement may be a reasonable steady state for an air quality monitoring system. Eventually, with enough labeled data, it might become possible to train a machine learning model to identify the highest frequency data labels, which could support labeling household-specific events, rather than focusing on a closed label set, e.g. Fang, et al.'s classifier for cooking, smoking, and pesticide events [16].

6.3 Evolving System Usage During Long-Term Deployments

Whereas prior work focuses on the benefits of deploying residential air quality monitor systems and its impact on air quality [16, 29–31], this work focuses on how users engage with an air quality sensing system and how this engagement changes over time. It is notable that MAAV's tablet interface seems to have served participants

well through their trajectory of system usage, starting at a high engagement level during an initial period of discovery and sensemaking, and then transitioning into a steady-state maintenance mode of sustained lower activity. Although the current interface supported this transition in some ways, future systems could design for these stages more directly. Toward the beginning, the system could take advantage of high engagement to learn more about the user's interests and relevant lifestyle factors. As data begins to accumulate, the system could begin to suggest comparisons, correlations, or views on the data. As the user transitions into a maintenance mode, the system could similarly scale back to a lighter weight visualization. It could even streamline annotation, such as asking the participant to confirm predicted annotations, rather than typing them out from scratch. This notion of an evolving interface could be more generally useful in other human-in-the-loop data collection domains, including eco-feedback and personal informatics.

Interpreting air quality monitoring within a personal informatics framework is compelling because we did not initially consider it within this domain. Air quality, even indoors, is often perceived as a characteristic of an area. Yet, air quality's influence is clearly relevant to personal informatics when relating it to personal health impacts. Findings in this work also show $PM_{2.5}$ to be an unexpected proxy for human activity. Reflecting further on participants' evolving engagement, we began to consider this work in the context of personal informatics models [15, 37], with the initial high engagement similar to the *discovery* phase in Li's stage-based model for personal informatics systems [37]. Participants also exhibited changes in their engagement related to increased familiarity, seasonality, and new questions, leading to periodic interaction with the tablet (D2, D3, D4, D6) and voice annotation modality (D3). Participants' periodic engagement with MAAV maps neatly to the lapsing and resuming cycles described in Epstein's lived informatics model [15]. Shorter term deployments D5 and D6 exhibited similar behavior, notably with occasional tablet use (D5) and re-engagement with the tablet interface (D6) immediately after the engagement interview.

The diversity in spike activity within deployments offers opportunities to reflect on the impact a residence has on indoor air quality and engagement. The combination of D4's low participant engagement and burden of high text notifications may account for their comparably low notification compliance, but we have no insight into mechanisms behind their increased $PM_{2.5}$ activity. Home permeability, furnace efficiency, and behavioral patterns can contribute to better understanding these causes, but these factors were outside the scope of this work. Extrapolating or incorporating direct measurement of any of these factors could help structure what appeared to D4 to be an overly sensitive system. Conversely, the highest compliant deployments (D3, D5) were those most curious to find out what was going on in their homes. This innate curiosity, along with co-enrollment in an ongoing clinical study, helped sustain participants' engagement. Moving forward, the findings of this work will help us to design systems and interfaces to foster engagement and curiosity, enabling users to pose and answer their own questions through more sophisticated exploration, comparison, and hypothesis generation techniques. Combined with an adaptive system, users could receive more relevant and actionable information.

7 LIMITATIONS

The findings presented in this paper should be interpreted in the context of specific limitations and study design decisions. Recognizing that multiple factors affect and influence indoor air quality, it is important to acknowledge that any self-labeled annotations reasoning about $PM_{2.5}$ spikes represent a participant's best guess and are not necessarily indicative of the reported source.

The air quality monitor's design – in particular the monitor's internal air quality sensor – was a source of some initial confusion. These sensors are counting airborne particles, and therefore they report only the observed concentration but say nothing about chemical composition or its health impact. Therefore, an increase in $PM_{2.5}$ particle counts does not necessarily correlate to an unhealthy source. For example, humidity can create significant

spikes. This distinction became clear as participants recognized the monitor's responsiveness to a range of activities.

Our participants were concurrently enrolled in an ongoing medical study, which has both benefits and drawbacks. Advantageously, we engaged a highly motivated population already familiar with data collection and interested in supporting the project. However, their involvement in a clinical study came with a number of caveats regarding the information we could disclose to avoid introducing confounds. Clinicians are interested in participants self-labeling and contextualizing data. However, as our findings show, participants who engage with and annotate their own air quality data are also inclined to connect air quality to their own behaviors, reflect on that relationship, and take action to improve their $PM_{2.5}$ exposure. To mitigate these broader concerns over how our interviews may influence participants and potentially confound the parent study's clinical findings, our interviews were vetted by practitioners in the parent project to ensure we were not encouraging or influencing participants to change their behavior.

Further, our field deployments consisted of six households, limiting the generalizability of our findings to a larger population, as does the asthmatic status of the participating families. This status also influences baseline levels of engagement, as asthma severity has been shown to be positively correlated with engagement of a technology probe [57].

We also encountered some technology limitations. Early in phase II, our voice transcription skill went offline from firmware updates outside our control, which prevented some participants from using this feature and consequently caused them to lose faith in this annotation modality. Furthermore, two participants experienced difficulty with Google Home recognizing and parsing their voice commands, further reducing its appeal. Lastly, staggered deployment scheduling resulted in deployments D5 and D6 starting later than the other participants, leading to a shorter overall deployment length.

8 CONCLUSIONS AND FUTURE WORK

Air quality and $PM_{2.5}$ exposure are important and often invisible aspects of personal health. We deployed MAAV, a multi-monitor air quality measurement, annotation, and visualization system, in a longitudinal field study with six families to understand how this combination of features could support participants in managing their air quality at home. The results of the deployment indicate that MAAV supported participants in gaining insights by comparing multiple monitor streams, recording and exploring annotations for contextualizing their air quality data, and remaining engaged over their long-term deployments by relying on low-burden text message alerts. Participants appropriated the tool in various ways that suited their needs and produced individual, personalized insights into the air quality in and around their homes.

Building on the work presented here, we plan to focus on redesigning the interface as an event viewer. Participants were less interested in seeing precise values and instead evaluated their data based on the presence and frequency of detected spikes. This event viewer could propose multiple comparative views to better externalize differences between events and across time, while evaluating air quality measurements between rooms and inside and outside the home. We also plan to re-engineer the annotation system to facilitate sorting, searching, highlighting, and filtering across a variety of dimensions, including time, specific keywords, and classes of activities. These additional features would enable users to perform more complex filtering queries such as reviewing 'All cooking events on Monday evenings.'

Another possibility would be to investigate the potential of a mobile application to integrate notifications and data review, which would provide greater control over interaction mechanisms and a more personalized air quality exposure. Such a mobile application could integrate a variety of external data sources by leveraging output from mobile health sensors such as smartwatches or other wearables, going beyond the standard wearable health metrics to incorporate next-generation air quality hardware, capable of measuring air pollutants (lead, ozone,

SO₂, NO₂, fine and coarse particulate matter, CO₂) in a mobile platform. Accessing new data streams provides greater opportunities to not only integrate additional health metrics, but also provide supplemental context via a visual overlay for viewing PM_{2.5} events in relation to other data such as health events. This collection of modular sensors could integrate within our system and provide external triggers to collect further feedback on user health and well-being.

ACKNOWLEDGMENTS

Research reported in this publication was supported by the National Institute of Biomedical Imaging and Bioengineering of the National Institutes of Health under Award Number U54EB021973. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

REFERENCES

- [1] 2014. Smoothed Z Score Algorithm, Stack Overflow. <https://stackoverflow.com/questions/22583391/peak-signal-detection-in-realtime-timeseries-data/22640362#22640362>. (March 25, 2014). Accessed: 2017-05-15.
- [2] 2016. How To Install Google Play Store on FireOS(5.x). <https://forum.xda-developers.com/amazon-fire/general/how-to-install-google-play-store-fire-t3486603>. (October 24, 2016). Accessed: 2017-06-01.
- [3] Eileen Abt, Helen H. Suh, Paul Catalano, and Petros Koutrakis. 2000. Relative contribution of outdoor and indoor particle sources to indoor concentrations. *Environmental Science & Technology* 34, 17 (2000), 3579–3587.
- [4] American Lung Association. 2015. *State of the Air 2015*. Technical Report. American Lung Association.
- [5] Michelle L. Bell, Keita Ebisu, and Roger D. Peng. 2011. Community-level spatial heterogeneity of chemical constituent levels of fine particulates and implications for epidemiological research. *Journal of Exposure Science and Environmental Epidemiology* 21, 4 (07 2011), 372–384. <http://dx.doi.org/10.1038/jes.2010.24>
- [6] Michael Bostock, Vadim Ogievetsky, and Jeffrey Heer. 2011. D³ data-driven documents. *IEEE Transactions on Visualization and Computer Graphics* 17, 12 (2011), 2301–2309.
- [7] Xuxu Chen, Yu Zheng, Yubiao Chen, Qiwei Jin, Weiwei Sun, Eric Chang, and Wei-Ying Ma. 2014. Indoor air quality monitoring system for smart buildings. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, New York, NY, USA, 471–475. <https://doi.org/10.1145/2632048.2632103>
- [8] Yun Cheng, Xiucheng Li, Zhijun Li, Shouxu Jiang, Yilong Li, Ji Jia, and Xiaofan Jiang. 2014. AirCloud: a cloud-based air-quality monitoring system for everyone. In *Proceedings of the 12th ACM Conference on Embedded Network Sensor Systems*. ACM, 251–265.
- [9] Victoria Clarke and Virginia Braun. 2014. *Thematic Analysis*. Springer New York, New York, NY, 1947–1952. https://doi.org/10.1007/978-1-4614-5583-7_311
- [10] Enrico Costanza, Sarvapali D. Ramchurn, and Nicholas R. Jennings. 2012. Understanding domestic energy consumption through interactive visualisation: a field study. In *Proceedings of the ACM Conference on Ubiquitous Computing*. ACM, 216–225.
- [11] Anind K. Dey, Katarzyna Wac, Denzil Ferreira, Kevin Tassini, Jin-Hyuk Hong, and Julian Ramos. 2011. Getting closer: an empirical investigation of the proximity of user to their smart phones. In *Proceedings of the 13th International Conference on Ubiquitous Computing*. ACM, New York, NY, USA, 163–172. <https://doi.org/10.1145/2030112.2030135>
- [12] Naihua Duan. 1982. Models for human exposure to air pollution. *Environment International* 8, 1-6 (1982), 305–309.
- [13] Prabal Dutta, Paul M. Aoki, Neil Kumar, Alan Mainwaring, Chris Myers, Wesley Willett, and Allison Woodruff. 2009. Common sense: participatory urban sensing using a network of handheld air quality monitors. In *Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems*. ACM, 349–350.
- [14] Daniel Epstein, Felicia Cordeiro, Elizabeth Bales, James Fogarty, and Sean Munson. 2014. Taming data complexity in lifelogs. *Proceedings of the 2014 conference on Designing interactive systems - DIS '14* (2014), 667–676. <https://doi.org/10.1145/2598510.2598558>
- [15] Daniel A. Epstein, An Ping, James Fogarty, and Sean A. Munson. 2015. A lived informatics model of personal informatics. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 731–742. <https://doi.org/10.1145/2750858.2804250>
- [16] Biyi Fang, Qiumin Xu, Taiwoo Park, and Mi Zhang. 2016. AirSense: an intelligent home-based sensing system for indoor air quality analytics. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 109–119.
- [17] Corinna Fischer. 2008. Feedback on household electricity consumption: A tool for saving energy? *Energy Efficiency* 1, 1 (2008), 79–104. <https://doi.org/10.1007/s12053-008-9009-7>
- [18] Jon Froehlich, Leah Findlater, Marilyn Ostergren, Solai Ramanathan, Josh Peterson, Inness Wragg, Eric Larson, Fabia Fu, Mazhengmin Bai, Shwetak Patel, and James A. Landay. 2012. The design and evaluation of prototype eco-feedback displays for fixture-level water usage data. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 2367–2376.

- <https://doi.org/10.1145/2207676.2208397>
- [19] Jon Froehlich, Eric Larson, Sidhant Gupta, Gabe Cohn, Matthew Reynolds, and Shwetak Patel. 2011. Disaggregated end-use energy sensing for the smart grid. *IEEE Pervasive Computing* 10, 1 (2011), 28–39. <https://doi.org/10.1109/MPRV.2010.74>
 - [20] Dawson R. Hancock and Bob Algozzine. 2016. *Doing case study research: A practical guide for beginning researchers*. Teachers College Press.
 - [21] David Hasenfratz, Olga Saukh, Silvan Sturzenegger, and Lothar Thiele. 2012. Participatory air pollution monitoring using smartphones. *Mobile Sensing* 1 (2012), 1–5.
 - [22] Shruti Hegde, Kyeong Min, Jimmy Moore, Kerry Kelly, Scott Collingwood, Neal Patwari, and Philip Lundrigan. in submission. Household indoor particulate matter measurement using a network of low-cost sensors. In *Environmental Pollution*.
 - [23] Yen-Chia Hsu, Paul Dille, Jennifer Cross, Beatrice Dias, Randy Sargent, and Illah Nourbakhsh. 2017. Community-empowered air quality monitoring system. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, 1607–1619.
 - [24] Yang Huang, Liang Hu, Disheng Yang, and Hengchang Liu. 2017. Air-Sense: indoor environment monitoring evaluation system based on ZigBee network. In *IOP Conference Series: Earth and Environmental Science*, Vol. 81. IOP Publishing, 012208.
 - [25] Karol Jabłoński and Tomasz Grychowski. 2017. Fuzzy inference system for the assessment of indoor environmental quality in a room. *Indoor and Built Environment* (2017). <https://doi.org/10.1177/1420326X17728097>
 - [26] Yifei Jiang, Kun Li, Lei Tian, Ricardo Piedrahita, Xiang Yun, Omkar Mansata, Qin Lv, Robert P. Dick, Michael Hannigan, and Li Shang. 2011. MAQS: A personalized mobile sensing system for indoor air quality monitoring. In *Proceedings of the 13th International Conference on Ubiquitous Computing*. ACM, New York, NY, USA, 271–280. <https://doi.org/10.1145/2030112.2030150>
 - [27] Ming Jin, Nikolaos Bekiaris-Liberis, Kevin Weekly, Costas Spanos, and Alexandre Bayen. 2015. Sensing by proxy: Occupancy detection based on indoor CO2 concentration. *UBICOMM 2015* 14 (2015).
 - [28] Jung-Yoon. Kim, Chao-Hsien Chu, and Sang-Moon Shin. 2014. ISSAQ: An integrated sensing systems for real-time indoor air quality Monitoring. *IEEE Sensors Journal* 14, 12 (Dec 2014), 4230–4244. <https://doi.org/10.1109/JSEN.2014.2359832>
 - [29] Sunyoung Kim and Eric Paulos. 2009. inAir : Measuring and Visualizing Indoor Air Quality. *Proceedings of The International Conference on Ubiquitous Computing* (2009), 81–84. <https://doi.org/10.1145/1620545.1620557>
 - [30] Sunyoung Kim and Eric Paulos. 2010. InAir: sharing indoor air quality measurements and visualizations. *Proceedings of the Conference on Human Factors in Computing (CHI)* (2010), 1861–1870. <https://doi.org/10.1145/1753326.1753605>
 - [31] Sunyoung Kim, Eric Paulos, and Jennifer Mankoff. 2013. inAir: a longitudinal study of indoor air quality measurements and visualizations. *Proceedings of the ACM Conference on Human Factors in Computing Systems* (2013), 2745. <https://doi.org/10.1145/2470654.2481380>
 - [32] Neil E Klepeis, William C Nelson, Wayne R Ott, John P Robinson, Andy M Tsang, Paul Switzer, Joseph V Behar, Stephen C Hern, and William H Engelmann. 2001. The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants. *Journal of Exposure Science and Environmental Epidemiology* 11, 3 (2001), 231.
 - [33] Ismo K Koponen, Ari Asmi, Petri Keronen, Katri Puhto, and Markku Kulmala. 2001. Indoor air measurement campaign in Helsinki, Finland 1999–the effect of outdoor air pollution on indoor air. *Atmospheric Environment* 35, 8 (2001), 1465–1477.
 - [34] Pramod Kulkarni, Paul A Baron, and Klaus Willeke. 2011. *Aerosol measurement: principles, techniques, and applications*. John Wiley & Sons.
 - [35] Stacey Kuznetsov, Scott E Hudson, and Eric Paulos. 2014. A low-tech sensing system for particulate pollution. In *Proceedings of the 8th International Conference on Tangible, Embedded and Embodied Interaction*. ACM, 259–266.
 - [36] Johanna Lepeule, Francine Laden, Douglas Dockery, and Joel Schwartz. 2012. Chronic exposure to fine particles and mortality: An extended follow-up of the Harvard six cities study from 1974 to 2009. *Environmental Health Perspectives* 120, 7 (2012), 965–970. <https://doi.org/10.1289/ehp.1104660>
 - [37] Ian Li, Anind Dey, and Jodi Forlizzi. 2010. A stage-based model of personal informatics systems. In *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10*. 557. <https://doi.org/10.1145/1753326.1753409>
 - [38] P. Lundrigan, K. Min, N. Patwari, S. Kasper, K. Kelly, J. Moore, M. Meyer, S. C. Collingwood, F. Nkoy, B. Stone, and K. Sward. 2017. EpiFi: An In-Home Sensor Network Architecture for Epidemiological Studies. *ArXiv e-prints* (Sept. 2017). [arXiv:cs.NI/1709.02233](https://arxiv.org/abs/1709.02233)
 - [39] Qing Yu Meng, Barbara J Turpin, Leo Korn, Clifford P Weisel, Maria Morandi, Steven Colome, Junfeng Zhang, Thomas Stock, Dalia Spektor, Arthur Winer, et al. 2005. Influence of ambient (outdoor) sources on residential indoor and personal PM2.5 concentrations: analyses of RIOPA data. *Journal of Exposure Science and Environmental Epidemiology* 15, 1 (2005), 17–28.
 - [40] Christian Monn. 2001. Exposure assessment of air pollutants: a review on spatial heterogeneity and indoor/outdoor/personal exposure to suspended particulate matter, nitrogen dioxide and ozone. *Atmospheric environment* 35, 1 (2001), 1–32.
 - [41] Engineering National Academies of Sciences, Medicine, et al. 2016. *Health Risks of Indoor Exposure to Particulate Matter: Workshop Summary*. National Academies Press.
 - [42] Nima Nikzad, Nakul Verma, Celal Ziftci, Elizabeth Bales, Nichole Quick, Piero Zappi, Kevin Patrick, Sanjoy Dasgupta, Ingolf Krueger, Tajana Šimunić Rosing, and William G. Griswold. 2012. CitiSense: Improving Geospatial Environmental Assessment of Air Quality Using a Wireless Personal Exposure Monitoring System. In *Proceedings of the Conference on Wireless Health (WH '12)*. ACM, New York, NY, USA, Article 11, 8 pages. <https://doi.org/10.1145/2448096.2448107>

- [43] National Institute of Biomedical Imaging and Bioengineering. 2015. Pediatric Research Using Integrated Sensor Monitoring Systems. (2015). <https://www.nibib.nih.gov/research-funding/prisms>
- [44] World Health Organisation. 2016. *Ambient Air Pollution: A Global Assessment of Exposure and Burden of Disease*. Working Papers. eSocialSciences. <https://EconPapers.repec.org/RePEc:ess:wpaper:id:11368>
- [45] Annette Peters, Emerson Liu, Richard L Verrier, Joel Schwartz, Diane R Gold, Murray Mittleman, Jeff Baliff, J Annie Oh, George Allen, Kevin Monahan, and Douglas W Dockery. 2000. Air pollution and incidence of cardiac arrhythmia. *Epidemiology* 11, 1 (2000). <http://journals.lww.com/epidem/Fulltext/2000/01000/Air>
- [46] A.C. Pope, Richard T. Burnett, Michelle C. Turner, Aaron Cohen, Daniel Krewski, Michael Jerrett, Susan M. Gapstur, and Michael J. Thun. 2011. Lung cancer and cardiovascular disease mortality associated with ambient air pollution and cigarette smoke: Shape of the exposure-response relationships. *Environmental Health Perspectives* 119, 11 (2011), 1616–1621. <https://doi.org/10.1289/ehp.1103639>
- [47] O. A. Postolache, J. M. Dias Pereira, and P. M. B. Silva Girao. 2009. Smart Sensors Network for Air Quality Monitoring Applications. *IEEE Transactions on Instrumentation and Measurement* 58, 9 (2009), 3253–3262. <https://doi.org/10.1109/TIM.2009.2022372>
- [48] PurpleAir. 2017. PurpleAir Air Quality Monitoring: An air quality monitoring network built on a new generation of "Internet of Things" sensors (<http://map.purpleair.org/>) Accessed 2017-11-13. (2017). <https://www.purpleair.com/>
- [49] Shaharil Saad, Allan Andrew, Ali Shakaff, Abdul Saad, Azman Kamarudin, and Ammar Zakaria. 2015. Classifying Sources Influencing Indoor Air Quality (IAQ) Using Artificial Neural Network (ANN). *Sensors* 15, 5 (May 2015), 11665–11684. <https://doi.org/10.3390/s150511665>
- [50] I Salma, K Dosztály, T Borsós, B Söveges, T Weidinger, G Kristóf, N Péter, and Zs Kertész. 2013. Physical properties, chemical composition, sources, spatial distribution and sinks of indoor aerosol particles in a university lecture hall. *Atmospheric environment* 64 (2013), 219–228.
- [51] Hillol Sarker, Moushumi Sharmin, Amin Ahsan Ali, Md Mahbubur Rahman, Rummana Bari, Syed Monowar Hossain, and Santosh Kumar. 2014. Assessing the availability of users to engage in just-in-time intervention in the natural environment. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 909–920.
- [52] Lihua Shi, Antonella Zanobetti, Itai Kloog, Brent A Coull, Petros Koutrakis, Steven J Melly, and Joel D Schwartz. 2016. Low-Concentration PM(2.5) and Mortality: Estimating Acute and Chronic Effects in a Population-Based Study. *Environmental Health Perspectives* 124, 1 (01 2016), 46–52. <https://doi.org/10.1289/ehp.1409111>
- [53] Susanne Steinle, Stefan Reis, and Clive Eric Sabel. 2013. Quantifying human exposure to air pollution—moving from static monitoring to spatio-temporally resolved personal exposure assessment. *The Science of the Total Environment* 443 (jan 2013), 184–93. <https://doi.org/10.1016/j.scitotenv.2012.10.098>
- [54] Peter Tolmie, Andy Crabtree, Tom Rodden, James Colley, and Ewa Luger. 2016. “This Has to Be the Cats”: Personal Data Legibility in Networked Sensing Systems. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. ACM, New York, NY, USA, 491–502. <https://doi.org/10.1145/2818048.2819992>
- [55] Robert J. Vercellino, Darrah K. Sleeth, Rodney G. Handy, Kyeong T. Min, and Scott C. Collingwood. 2018. Laboratory evaluation of a low-cost, real-time, aerosol multi-sensor. *Journal of Occupational and Environmental Hygiene* 15, 7 (2018), 559–567. <https://doi.org/10.1080/15459624.2018.1468565> arXiv:<https://doi.org/10.1080/15459624.2018.1468565> PMID: 29683781.
- [56] LA Wallace, ED Pellizzari, TD Hartwell, C Sparacino, R Whitmore, L Sheldon, H Zelon, and R Perritt. 1987. The TEAM (Total Exposure Assessment Methodology) Study: personal exposures to toxic substances in air, drinking water, and breath of 400 residents of New Jersey, North Carolina, and North Dakota. *Environmental research* 43, 2 (August 1987), 290–307. [https://doi.org/10.1016/s0013-9351\(87\)80030-0](https://doi.org/10.1016/s0013-9351(87)80030-0)
- [57] Tae-Jung Yun Tae-Jung Yun, Hee Young Jeong Hee Young Jeong, Hee Rin Lee Hee Rin Lee, Rosa I. Arriaga, and Gregory D. Abowd. 2010. Assessing asthma management practices through in-home technology probes. *Proceedings of the 4th International ICST Conference on Pervasive Computing Technologies for Healthcare* (2010), 1–9. <https://doi.org/10.4108/ICST.PERVASIVEHEALTH2010.8839>

Received November 2017; revised July 2018; accepted September 2018