Mites FAQs

Please note, we have prepared this document to answer many of the questions that have been asked about the Mites infrastructure in TCS Hall by its occupants and our community members in a transparent manner. It also includes many of our research ideas and our long-term vision and goals. We have divided this set of FAQs into several categories. However, we encourage you to read it in its entirety.

Executive Summary (TLDR;): TCS Hall has a building-wide sensor network based on the "Mites", a ubiquitous sensing platform with 9 different environmental sensors. Our vision is to create compelling smart building apps that are useful to our community and study how users interact with the Mites sensors and these apps. We have been working on the Mites and planning the deployment of these sensors in TCS Hall for more than two years. We have talked to, and worked, with a number of people throughout this process and held town halls for our community. We have also solicited questions that we can answer to help make everyone comfortable. We have worked with CMU's IRB over a 4 month period to follow all processes, have them vet and approve our research protocol, answer any questions and make changes as needed.

While modern buildings already have numerous sensors for their management, we also recognize that any additional sensors in our spaces can lead to natural privacy and security concerns. We take privacy and security seriously, and our entire system has several privacy and security safeguards built-in, both technical but also social, and in terms of processes we follow (see section Privacy and Security). The sensors on the Mites sense the environment around them. The Mites do not have a camera to sense images or video. The data from the microphone is processed and featurized on the sensor itself, such that the source sound cannot be reconstructed. All other sensor data is also featurized similarly such that it is not Personally Identifiable Information (PII) (see section General Questions). In public spaces, the sensors are turned on as the data contains no personally identifiable information. In individual/shared offices sensor data may be indirectly linked to the occupants based on the sensor location but this association risk is further mitigated by using an obfuscated ID mapping of the device location with the Mite data and thus this risk is significantly lower than those from existing sensors in buildings (see FAQ#D2, FAQ#D3, and FAQ#P1). However, sensors in offices may still lead to concerns and we provide extensive privacy controls for occupants in these offices to disable any (or even all) of the sensors using the Mites Mobile App or just sending us an email (see FAQ#UC1). To enable the microphone in an office all occupants will have to provide informed consent in the MitesApp and turn it on. Using the MitesApp, and using its functions such as installing applets based on Mites sensor data, is completely voluntary and informed consent will be required from users (see section User Interaction, Controls, and Data Collection). Our custom Mites

hardware is secured using industry-standard cryptographic protocols, is network isolated, and only communicates with an on-campus server. Our backend is access controlled, protected, and hardened using industry-standard best practices. All applets will be vetted carefully, and informed consent will be obtained as required (see sections Privacy and Security).

We are grateful for the trust you have placed in us. We also realize that this is an ambitious project which will hopefully help us study responsibly how future smart buildings infrastructure using IoT devices should be designed, deployed, and used, with privacy and security as first-class objectives. We want to learn with you as partners. We will provide multiple mechanisms to hear your concerns and get your feedback. In addition to providing feedback directly in the MitesApp, you can also send us an email at questions@mites.io or just come and talk to us in person! We also plan to have annual town halls where we can answer questions and also learn from our community what works, what did not, and what you would love to see in a future smart building application!

Basic Questions:

This section consists of basic definitions of Human Subject Research (HSR). While some of the members of our community likely already know this, we would like to discuss these important concepts and how they pertain to the Mites Project and the deployment at TCS in general and hopefully alleviate several privacy and security concerns.

FAQ#0 What is the definition of Human Subject Research?

The Institutional Review Board (IRB) and the Federal regulations define Human subject research based on 45 CFR 46.102(e) as:

> *Human subject research means a living individual about whom an investigator (whether professional or student) conducting research:*
>
> - *Obtains information or biospecimens through intervention or interaction with the individual, and uses, studies, or analyzes the information or biospecimens; OR*
> - *Obtains, uses, studies, analyzes or generates identifiable private information or identifiable biospecimens.*

45 CFR 46.102(2) defines *identifiable private information* as “private information for which the identity of the subject is or may readily be ascertained by the investigator or associated with the information.”

FAQ#1 What aspects of the Mites sensor deployment in TCS are considered human subjects research?

The presence of the Mites in the TCS building and the collection of environmental data from them does not constitute research involving human subjects per the definition in FAQ#0 for the reasons outlined below:

- The data collected from the Mites devices is not obtained by the intervention (i.e., an attempt to modify behavior or outcomes) or interaction with individuals.
- The data collected from the Mites devices is not identifiable private information because there would not be any way for the researchers to specifically identify an individual who was the source of any of the data from the Mites and the data elements being collected are environmental in nature (e.g., temperature, humidity, sound).

However, a user who chooses to claim a Mite in an office or install a Mite in their home and then interact with that Mite through the app is considered a human subject because data is being collected about an individual (see FAQ#D2). The research involving human subjects was reviewed and approved by the IRB. This research will examine how users interact with the Mites and use the app. These individuals will sign the informed consent prior to claiming the Mites and using the app, and prior to any research data collection on the individual’s use of Mites. Any part of the Mites system or any application built using the Mites system that collects data that is attributed to a specific individual or requires the interaction or intervention of a Human subject requires us to obtain Informed consent. As with any research, there are often risks involved and we do our best to address them in the sections Privacy and Security).

Category: General Questions

FAQ#G1: What are the goals of the Mites research project?

Our overarching aim is to build an end-to-end, secure, and privacy-aware, IoT infrastructure using a combination of sensor technologies (our Mites device) that is scalable, extensible, and usable. . One domain we are exploring is Smart Buildings, which already have many sensor systems in them and which are often not designed with security and privacy in mind. As part of our overall vision, we would like to develop

this infrastructure first and then study different applications and use cases on top of it, as well as how users interact with the infrastructure, including the security and privacy controls we develop and provide.

FAQ#G2: What are Mites?

Mites were designed at CMU as a collaborative project (SynergyLabs and FigLab) to be all-in-one sensing devices for smart spaces. We integrated nine distinct sensors, which translate to twelve distinct sensor streams.

FAQ#G3: What can Mites capture and what can they not?

The Mites devices have 9 different sensors that can sense several physical quantities in the ambient environment: ambient temperature, relative atmospheric humidity, pressure, magnetic fields, vibrations, motion, light intensity, light color, WiFi signal strength, thermal temperature, electromagnetic noise, Bluetooth devices nearby, and sound. The raw data obtained from the sensors on each Mites device is processed **on the device** in a series of steps that essentially convert it into a non-reconstructable featurized representation that consists of basic statistical features (min, max, range, average, sum, standard deviation, and centroid) and aggregated frequency representation values (using a Fast Fourier Transform (FFT)). This featurization and denaturing of data is done primarily to mitigate any privacy concerns such that the essence of the signals can be extracted while preventing the reconstruction of the original signals. Notably, all this processing and denaturing happens on the Mites device itself in its secure firmware; thus, the raw sensor data never leaves the Mites device. So, in the case of the audio sensor, no raw audio data is ever sent from the Mites devices and the original audio data cannot be reconstructed from the featurized data.

FAQ#G4: Where do the Mites send data and how is that kept secure?

The Mites devices use end-to-end encrypted communication to connect to a pre-programmed server on campus (details in FAQ#S1). The server is a standard Linux machine that will be kept up to date with regular security updates and patches. In addition, it has a firewall and is secured using standard Linux tools. Only the researchers have access to the server and all communication is authenticated using industry-standard protocols (HTTPS, OAuth2.0, etc.).

FAQ#G5: What is the actual capability of the microphone on the Mites? Are you recording a conversation? Are you violating PA Wiretap law?

The microphone on the Mites device is a low-cost, MEMS-based embedded microphone.

It can capture audio, but as articulated in FAQ#G3 this audio data is denatured and featurized on the sensor board itself to prevent reconstruction. This level of non-reconstructable featurization does not allow us to record any conversation.

No, we are not violating the PA wiretap law. We have had extensive discussions with the CMU - Office of the General Counsel and they have verified that we are not violating the PA wiretap law

FAQ#G6: What data are the sensors currently collecting and what data will they collect in the near future?

Currently, no sensor data is being collected. Only the Mites sensor's health information (device status, device heartbeat, packet rate, etc.) is being sent to the server. As per the approved IRB protocol, the team will start data collection from all the sensors on the Mites in public spaces (notices will be going up as per the protocol). In individual and shared office spaces, once we have the Mobile App released we will notify everyone via email that they can use it to interact with the Mites in their offices and we will start sensor data collection from the Mites in the individual/shared spaces (except the audio sensor, which requires explicit consent by all occupants of the office). FAQ#U1 mentions the controls users will have for their individual or shared offices.

Category: User Interaction Questions:

FAQ#U1: How can occupants of the TCS Building interact with the Mites?

We worked on building a Mobile App that allows the occupants to provide consent for research, control the sensors on the Mites device and install applications called “Applet” that provide useful insights from the Mites devices.

This Mites Mobile App, which we are currently testing and will be releasing soon, (iOS and Android) will be the primary mechanism for the occupant to interact with the infrastructure and the Mites sensors. . All users in TCS who decide to participate will have to create an account to authenticate (username, password, using OAuth) with the system. Using these credentials they can then authenticate themselves in the MitesApp. While we will be adding functionality to the Mites App and updating it over time, from the beginning, the app will at least have the ability for users to provide consent to the research and opt-in, then “claim” Mites in their own offices, and turn ON/OFF any of the individual sensors on their office. In addition, users will be able to view historical featurized data for their Mites for any of the enabled sensors (e.g., average temperature or humidity). In the future, we will add functionality to the MitesApp to also allow users to request access to a Mites sensor (or a subset of the sensors on a Mite, e.g. the temperature or the microphone) from other authenticated users of the system. Similarly, users will be able to grant or revoke access to other users who have requested sensor access. Over time we will add functionality to allow users to install “Applets” that use the data from Mites in their individual or shared office if they wish, or install “Applets” that use data from public sensors from within the MitesApp. More information on the applets is provided in FAQ#U6.

FAQ#U2 How can users benefit from the sensor data and what are example “applets”?

Users can benefit from the sensor data from their own offices, for example, to look at the trend of the average light level in their office, temperature, and humidity, or even how the ambient noise varies throughout the day which may affect their productivity. We use the term “applets” to denote applications that users can install from within the MitesApp (FAQ#U1) to use the sensor data from individual or shared offices or public

spaces. Examples of “public applets” are: “Find an available conference room,” and “Find a quiet space in the building to sit.” In individual or shared offices, users might install an applet that would allow them to train a virtual sensor to detect when the train passes on the track below TCS so they can ask the applet “How many trains have passed by today?”. An applet that computed aggregate statistics from both individual or shared offices and public spaces can also be used for several safety and security applications such as notifying emergency staff in case of fire or other safety issues. An applet used for the facilities manager might list rooms that are over 80 degrees or below 68 degrees to notify a maintenance crew to check if the HVAC is working correctly. Our longer-term vision is to also explore advanced techniques like differential privacy to develop applets in the future. Importantly, users will install these applets themselves and be able to know what data they are requesting for what purpose, and all these applets will be vetted before even being available for users to install (See FAQ#U7). If an IRB modification is needed to enable any applets, the researchers will get it prior to making it available.

FAQ#U3: Where can I download the Mites App? Who maintains it? Does the Mites App collect any data?

The Mites Application can be downloaded from the iOS and Android Playstore and will be maintained by researchers at SynergyLabs. Apart from the usual data required for the functionality of the Mites App, we also collect feedback from the user about any concerns or issues that they are facing with the Mites infrastructure.

FAQ#U4: How are the users authenticated via the App?

All the users require their Andrew ID to login into their system. The Mites App uses standard **OAuth2.0 protocol** to communicate with our middleware OS, BuildingDepot, to interact with the Mites devices. Additionally, we use a unique access token (based on OAuth2.0) for every login/access to the system that is refreshed every 24 hours (or) on logout to prevent any access to unauthorized data.

FAQ#U5: How are the users and the data from the Mites devices associated?

The users and the data from the Mites devices are never directly associated and will never be. All the sensor data from the Mites devices are stored to a Unique Identifier (UUID) and these identifiers are tagged with an obfuscated location ID of the Mite in the Building. When this data is used for research (or) by other Applets, the data associated

with the obfuscated location of the Mite is used and thus, the user information is never associated (see FAQ#P2 and FAQ#P8 for more details).

FAQ#U6: How does the installation of "Applets" work? Are there Applets that collect PII information?

When the registered users select an Applet to be installed in the system, they are spawned in an isolated environment specific to the Applet. Notably, for security and privacy considerations these applets do not run on any arbitrary machine but only in our sandboxed, tightly controlled, and isolated environment on a server that the researchers maintain at CMU. These isolated environments have a unique identifier (UUID) that is then mapped with sensor UUID, which the user specifies. Similar to **FAQ#U4**, the Applets then use the industry-standard **OAuth2.0 protocol** to communicate with our middleware OS, BuildingDepot, to get access to the data. Note, the user's Andrew ID is not involved in this process.
Yes, there may be Applets that collect PII information, and such Applets during installation would require your consent to use your information and associate it with the sensor data.

FAQ#U7: How are the Applets vetted? What process do the Researchers use for vetting Applets?

The researchers will carefully vet every Applet stored in the Mites system to ensure that it does not violate the users' privacy and security of the Mites system. Such Applets will also be carefully vetted for the purpose of data access from the Mites and their functionality to ensure that they are not doing anything malicious. Moreover, Applets which associates the user and the sensor data will be required to ask for consent from the user. An idea that we are also exploring is for applets to have something akin to a manifest that declares its functionality, what data it needs, why it needs it (purpose), etc for manual, and eventually automatic, verification. To start, these Applets will likely be only developed by our research team but we imagine over time other members of our community may propose new applets which we will carefully vet.

Category: Controls Provided and Related Questions:

FAQ#UC1:What controls do I have as an occupant of my individual or shared office space?

FAQ#U1 lists how occupants will be able to interact with a Mite in their own office using the mobile MitesApp. Once the occupants provide their consent through the MitesApp, they can enable or disable any (even all) of the sensors on the Mites in their individual or shared office. Note, occupants, do not have to provide consent in the MitesApp to be able to turn off any (or all) of the sensors on the Mites in their office and can just email at questions@mites.io to do so (see FAQ#D7). These requests can be handled by an ombudsman person selected by the department head from the department IT support group. Only the ombudsperson will receive a list of occupants who decided not to participate or who turned off sensors in their offices.

Again, we provide these controls despite the environmental data collected in individual or shared offices not being personally identifiable (and thus not associated with Human Subjects Research). (See #FAQ1).

FAQ#UC2: What controls do I have in Public spaces?

Occupants in public spaces do not have any controls. The sensors will remain on in these spaces.

FAQ#UC3: How does Mites Device control work in a shared office with multiple occupants?

In a shared office space with multiple occupants, when an occupant wants to enable or disable any sensor in their office they would need to discuss it with the other occupants before doing so. If there is any disagreement on control of the sensor, please send us an email at questions@mites.io. We will always use the most conservative settings, meaning only mutually agreed-upon sensors will be turned on for a shared office.

Category: Data Collection Related Questions:

FAQ#D1 What are the types of data that we are collecting?

We are collecting multiple types of data from the Mites system at the TCS deployment. Below are the types of data that we collect and the rationale behind collecting this information:

- ***System metrics data for sensor health:*** To monitor the health of the Mites devices, we will collect information such as the status of the device (online/offline/damaged), device memory usage, network packet rate over time, variation in the packet rate, and sensor data, round trip time for transmission of a packet from the devices, uptime of the device, number of over-the-air updates to the device, WiFi signal strength, and the number of reboots for each device. We also log the IP address of each device deployed.
- ***Mites Sensor Data:*** We collect non-reconstructable featurized data from the sensors on the Mites device (mentioned in FAQ#G3) which consists of basic statistical features (min, max, mean, and std. dev) and aggregated frequency representation values (using a Fast Fourier Transform (FFT)) that are used for various Applets (mentioned in FAQ#U2). **Note:** For the microphone sensor on the Mite we will not collect the data in an individual or shared office until we obtain the Informed Consent from all of the occupants (see FAQ#C2).
- ***Applet data:*** To assess different applications around the system and evaluate the application functions we plan to collect the data from the "Applets"(mentioned in FAQ#U2). **Note:** We will ask for consent for an Applet that collects PII data (see FAQ#C2, FAQ#P5)
- ***Participant metrics:*** To provide control and claiming of each Mites device after the participant consents, the Name, Andrew IDs of the participants will be used to register to the Mites system. The participant's Andrew IDs registered with the Mites system will be associated with the sensor data. The Mites system never uses Andrew ID for any of the internal functions (See FAQ#U4, FAQ#P2)
- ***Mites Application Feedback Data:*** To assess different concerns the users may have with the app or system or any suggestions in general we plan to obtain feedback from the user. **Note:** We will ask for consent when surveying users (see FAQ#C2).

FAQ#D2 What is the default data collection from the Mites Devices in individual or shared offices?

In individual/shared offices the default case as per our IRB protocol is to (a) collect telemetry data for sensor health and (b) collect data from all other sensors except the audio sensor. The rationale for making the default case to collect data from these other sensors except audio is that they do not collect any data that is personally identifiable, should have a very low-risk perception with users, and most of them already exist in building systems by default for building management (e.g. PIR based

movement/presence sensors, temperature, humidity, are all part of HVAC control sensors in almost every building). In fact, modern WiFi networks such as those at CMU already have information on each and every WiFi device authenticated and connected to the network for reasons like network security and health. Since audio sensing using the microphone can be perceived as sensitive despite the featurized data not being personally identifiable or reconstructable, all occupants of an office have to explicitly give their consent (i.e. opt-in) to enable the audio sensor data collection. We intend for this default to make residents of offices feel more comfortable with the presence of the Mites if they do not wish to participate in the human subjects research (using the app to interact with the Mites). To be clear, however, the audio sensor will not record voices or conversations. (See FAQ#D7 and FAQ#D9 on how to change this default data collection behavior in individual or shared offices.)

FAQ#D3 What is default data collection from the Mites Devices in Public Spaces?

By default, in public spaces, all sensor data will be collected. Note, as mentioned earlier, the processed data from the Mites is not personally identifiable (not PII).

FAQ#D4 Why are the researchers collecting data from an individual or shared office? What are the types of applications that are possible with such data?

The data collected from both the individual/shared offices and public spaces can be used to build several smart building applications. For example, applications to compute aggregate statistics at the building level such as obtaining coarse-grained occupancy by a hallway, floor, or at the building level for authenticated members allow several use cases to understand space utilization efficiency, energy management, and occupant comfort. Additionally, these computed aggregate statistics can also be used for several safety and security applications such as notifying emergency staff in case of fire or other safety issues. Another example is to identify HVAC faults (abnormally high temp on multiple offices on a floor) to the Building Manager. We also want to explore mechanisms such as differential privacy to balance privacy vs utility/accuracy tradeoffs. The data collected from both the individual/shared offices and public spaces, thus, allow us to analyze the data for several such smart building applications that can function even with this level of noisy data. Again, our goal is not to use applications (without consent) that attempt to link sensor data to individual users and we are not allowed to do so. When such applets exist that link sensor data to individual users, we will require the applet to ask for consent from the individuals assigned to that individual or a shared office (See FAQ#P5).

FAQ#D5: Can I unplug/remove the Mites devices if I want to opt out from the data collection process in a public space or in an individual/shared office?

No, please do not unplug the Mites devices from the walls or ceilings. The Mites devices are CMU property and are subject to the university policies for damage to CMU property (see Student Damage to University Property - University Policies). Additionally, these devices were installed as part of the building infrastructure with consultation from the Campus Facilities Design and Construction (CDFD) team and the Architects of the TCS building to ensure that this device adheres to Workplace and Construction Safety (see: Safety Consulting - Environmental Health & Safety). Removing these devices creates a safety hazard not just for you but all the occupants in the building, and such actions will be taken very seriously. The Mites in your own office can easily be turned off if you wish. Please look at question **FAQ#D9** for more details.

FAQ#D6: Is there any hardware switch on the Mites to disable the device or a particular sensor on the devices?

No, there is no hardware switch on the Mites to disable any sensor or the device itself. However, for individual or shared offices we have several options for users to be able to disable any sensors, including an option for users to be able to request disconnection of the Mites in their office if they want. (Please refer to **FAQ#D9**).

FAQ#D7: What is the process of opting out of having data collected in an individual or shared office space?

The occupants of the office can send a request to questions@mites.io with their office location and copy all the other occupants of the office to disable a sensor on the device or turn off all sensors.

These requests will be handled by an ombudsperson person selected by the department head from the department IT support group. The ombudsperson can also verify that these requests have been carried out.

The ombudsperson will not share the list of people who have consented to participate, declined participation, or those who have requested their sensor(s) to be turned off.

FAQ#D8: What is the process of opting out of having data collected in public spaces, such as conference rooms or common areas?

Because it is not possible to identify an individual from the sensor data, there is no way for an individual to change the default data collection behavior (opt-out) of coming into contact with the Mites in public spaces. The Mites do not collect identifiable information. As per our IRB-approved protocol, we will have notices in public spaces to let occupants

and visitors know about the data collection and a way for them to share their concerns with us (over email) and ask us questions.

FAQ#D9: What if I am still not comfortable and would still like to completely opt-out of any data collection in my individual/shared office?

We understand that some colleagues may still be uncomfortable with any sensor data collection in their individual/shared offices and the controls we provide. In this case, we request the occupants of that office to send us a request to questions@mites.io with their office location and copy all the other occupants of the office, to disable the Mites in their office. We will have the campus networking team disable the sensor connection from the network closet so that it is not provided power anymore. Note, the actual sensor will still be present in the wall/ceiling but will not be powered on. We request that occupants use this as a last resort since we would not even get network heartbeats from our sensors. Note, under no circumstances should occupants take it upon themselves to physically unscrew the sensors themselves (Reasons FAQ#D5). The next occupant of the office may want to enable the Mites. We hope the fact that the sensor can be powered off addresses any remaining concerns.

If the option to have the Mites powered off by the campus networking in your office does not suffice then occupants can request the physical disconnection of the Ethernet cable from the Mites in their office and even the removal of the Mites device as a last resort. We would like to note that physical disconnection/removal (and later re-connection/re-installation if the next occupant of the office desires it) comes at a significant logistical and financial cost thus we would sincerely request you to only request this in extreme cases. If any concerns still remain in your office please reach out to the ombudsman to see how we can further help resolve them.

The ombudsman will not share the list of people who have consented to participate, declined participation, or those who have requested their sensor(s) to be turned off.

Category: Consent-related questions:

FAQ#C1: What is Informed Consent? Why does the IRB not require the Mites Project to ask for consent from the occupants for collecting the Mites sensor data in individual/shared offices or public spaces?

Informed consent is the method of telling potential research participants about the key elements of a research study and what their participation entails to facilitate a prospective subject's decision about whether or not to participate in research. Participation in human subjects research is voluntary.

Collecting data from individual or shared offices and public spaces are not considered to be Human Subject Research and thus does not require informed consent from occupants or visitors to the building. Please see #FAQ0 and #FAQ1. Additionally, the Mites device provides strong privacy and security guarantees of the Mites sensor data and the underlying system overall. Please look at FAQ#U4, FAQ#U5, FAQ#P1, FAQ#P2, FAQ#P2, FAQ#D2, FAQ#D3 for more details.

Any part of the Mites system or any application built using the Mites system that collects data attributable to a specific individual (or) requiring the interaction or intervention of a Human subject requires us to obtain Informed consent. Because the use of the app is considered human subjects research, informed consent will be obtained from any individual using the app prior to their data being collected.

FAQ#C2: When do I have to consent?

Informed consent must be obtained from any subject participating in human subject research prior to participation or data collection. We are required to ask an individual to provide Informed consent for the below information:

- Using the app to interact with the Mites and claiming it (including turning on the microphone).
- Associate the actual location of the Mites device in their office with the data collected from the Mites device. (If the office is shared by multiple occupants, we would then require consent from all the occupants of the office before we can associate the Mites data with its actual location.)
- Associate the email ID of the participant in the office with the data collected from the Mites device.
- Collecting In-App Feedback from the Mites Application about any concerns or issues that they are facing with the Mites infrastructure.
- Any Applet that collects PII information requires your consent to use your information and associate it with the sensor data.

Since audio sensing using the microphone may be perceived as sensitive by some, despite the featurized data not being personally identifiable or reconstructable, we will ask for users to provide informed consent to collect the audio data. Please see #FAQ0 and #FAQ1 for more details.

FAQ#C3: How is consent obtained?

The consent is obtained through the MitesApp. Once the occupant of the individual/shared office or any public space has downloaded the MitesApp, they will be directed to a set of screens that onboard them, and on one of the screens, the occupants will be asked for their consent (see FAQ#C2 more details on when to

consent). For future deployed Applets, when it is being installed, and if the Applet requires the user to consent, the specific Applet will ask for the user's consent before it can be installed.

FAQ#C4: What happens if you don't want to provide Informed Consent?

Participation in human subjects research is voluntary. The occupant of the individual/shared spaces can decide not to provide informed consent to claim the sensor in the office space, for the collection of audio data from the Microphone and their feedback that will be used for research. Additionally, the occupants also have the choice to decline participation in data collection for research in Applets.

If you decide not to provide informed consent, you will be guided to a screen on the Mites Application to send us an email at questions@mites.io with requests such as to enable/disable sensors on Mites, or the Mites device itself. You will not be able to use the App functionalities to claim the sensor, no data will be collected from the microphone sensor, and the feedback will never be used for research.

Category: IRB Related Questions:

FAQ#I1: What was the category of IRB review that this project received? What was the process followed for your IRB protocol?

The research protocol was reviewed by the IRB under the Expedited category (non-exempt). It underwent several rounds of reviews over an almost 4 month period. We received many insightful comments and questions from the IRB reviewers and we submitted our responses and made modifications to our protocol. Notably, through several discussions, our IRB reviewers determined the following: (a) The processed information collected from any of the sensors on the Mites devices is not personally identifiable information. (b) Any information that may be perceived as PII (e.g. microphone sensor in an office) will be consented for.

FAQ#I2: Can I see the approved IRB protocol?

We currently do not plan to share our approved IRB protocol as it is meant to be confidential and impacts some of our research goals around user perception of privacy and security. However, we have summarized the key points of the IRB protocol in a more detailed manner in this FAQ document (which has also been vetted by the CMU IRB). However, once we release the MitesApp, users can view the user-consent form, which we are happy to share.

FAQ#I3: On what basis has the IRB categorized this project in the Expedited category?

The Federal Regulations governing Human Subjects Research create certain categories of research and corresponding requirements. The IRB categorizes the Mites project as part of the Expedited category as this research involves no more than minimal risk, and meets one or more of the OHRP Expedited Review Categories. Minimal risk is determined when the probability and magnitude of harm from participation in the research is not greater than what an individual would encounter in their normal daily activities or routine physical/psychological tests. The IRB has determined that the Mites project meets the below categories under Expedited review:

> Category *6. Collection of data from voice, video, digital, or image recordings made for research purposes.*
>
> *Category 7. Research on individual or group characteristics or behavior (including, but not limited to, research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies.*

Privacy Risks and Security:

Category: Privacy Risks from the Data from Mites:

FAQ#P1: Is the Data Personally Identifiable?

No, The sensor data obtained from the devices is not Personally Identifiable. This is based on how the information obtained from the sensors on the Mites is processed and how it is communicated so that it is anonymized, with strong privacy and security guarantees. To reiterate how we do this, the raw data obtained from the sensors on each Mites device is converted into a non-reconstructable featurized format that consists of basic statistical features (min, max, range, average, sum, standard deviation, and centroid) and aggregated phaseless frequency representation values (using a Fast Fourier Transform (FFT)). This fundamental denaturing of data happening on the Mites device does not permit the reconstruction of the original signals. Additionally, this featurized information is affected by the ambient environmental variation and noise, the presence of multiple people, sensing “range” from placement of the device, etc.

Given how the information is obtained from any of the sensors on the device and how they are affected by the environmental factors, this information cannot be used to identify an individual. See FAQ#G3 and FAQ#G4 for more details.

FAQ#P2: Is there a risk that an adversary (including a researcher) can associate the data obtained with the individuals indirectly?

To obtain the data from the Mites devices an adversary needs to get access to the system, understand how information is stored in the system, find the mapping of the obfuscated location information to the actual room location and identify the sensor data related to the individual room’s location, and finally, understand what data values mean. Several security practices of the Mites System system mentioned in Device and Server security (see FAQ#S1 and FAQ#G4) keep the devices in TCS secure from external hosts on the public internet.

Hypothetically, even if an adversary gains access to the system and obtains all this information, they still cannot use this sensor data to indicate directly if it is related to a particular individual. The reason is due to the way data is featurized on the Mites and the effect that different environmental noise has on the featurized data. Thus, the obtained data does not provide enough assurance to accurately indicate that it belongs to a particular individual even through indirect means.

FAQ#P3: Can the firmware of the device be manually or remotely updated by an adversary?

The possibility of an adversary taking over a Mites device remotely, or manually, is extremely low due to the number of security mechanisms we have in place. The Mites devices are pre-programmed with asymmetric cryptographic keys in their firmware such that they mutually authenticate themselves with, and only send data to, our backend server (specific hostname) that resides on the CMU Campus. Our backend server is itself secured using standard Linux security tools and best practices and will be kept up to date with regular patches and updates. Additionally, all the Mites devices deployed in the TCS Hall connect to the encrypted CMU-Device WiFi network, which provides a private IP address (172.X.X.X) such that devices at TCS are unreachable from anywhere outside of the campus (i.e. inaccessible from the Internet).
In order to manually program the firmware someone would need to know the hardware specifics of our sensor design and our proprietary communication protocols and have physical access to the sensors on the walls and ceiling in TCS.

FAQ#P4: Can the device be replaced with an identical/similar-looking device?

This problem with replacing the device with a similar-looking device is somewhat mitigated as only authorized CMU users can access the Mite devices on campus, and TCS Hall is access controlled. Additionally, the devices are located in places that are not easily accessible, such as the ceiling tiles in public spaces or in the wall or the ceiling in personal offices (which require physical keys to access). Furthermore, if a device goes missing, the researchers will be made aware of it since the Mites backend keeps track of the Mites devices that should be online. Most importantly, replacing the Mites device with another similar-looking device or maliciously programming the firmware to do high-fidelity recordings surreptitiously would be illegal and a gross violation of numerous laws. It is akin to a bad actor installing a covert surveillance device or installing malicious software on someone's laptop or phone. Such actions would be taken very seriously and will be reported by the researchers to the IRB and to the Campus Police if needed.

FAQ#P5: Can the Applet Collect PII Data? Will I be asked for consent for such applets?

Not all the Applets that use the Mites Sensor data can collect PII data. When an Applet is being installed that associates the sensor data with the users then users will be asked for their consent before installing that Applet. All the Applets that are being built using the Mites infrastructure will be carefully vetted by the Researchers.

FAQ#U6 provides examples of certain Applets. For example, the Applet "Count of the number of trains that have passed by today" may ask the user to provide examples of what the train passing by sounds like and obtain the data from the Mites devices. The Applet then may use this user to provide training data to predict the trains that have

passed by TCS Hall. For such Applets, where it associates the user and the mites sensor data, the users will be asked for consent before installing them. Other examples of Applets that do not associate the sensor data with the users, such as "Find an available conference room" (or) "Find a quiet space in the building to sit," do not require the user to ask for consent.

FAQ#P6: Can the "Applets" being developed on the Mites be used for "algorithmic management" of certain individuals or a group of individuals?

The researchers during the vetting process will not approve such an applet that can be used for any malicious purpose. The goal for "Applets" is to provide insightful information that can be obtained from the sensor data which promotes occupant comfort, encourages awareness about their space, explores newer insights, and improves building management efficiency. The Applets are never designed to surveil building occupants or use some sort of algorithmic management of individuals or any other malicious purposes. Our key research goal is to make it much harder for bad actors to invade our privacy or conduct surveillance and we are exploring multiple mechanisms to create these "Applets" to be more transparent and adhere to their stated purpose.

FAQ#P7: Is non-consensual data collection for research unethical?

All data collection for Human Subjects Research will require consent to participate. Anyone using the Mites app will provide consent prior to any data collection. This is the only current aspect of the Mites project that meets the definition of Human Subjects Research. Data collection for research that does NOT qualify as Human Subjects Research per the Federal definition does not require consent because the data collected is not private identifiable data and it is not collected via interaction or intervention with an individual. The Federal Regulations, which dictate and direct research procedures, are built on ethical considerations.

FAQ#P8: Is there a risk that an adversary can use the location information to associate the data obtained with the individuals indirectly?

As mentioned before (see **FAQ#P2**), to obtain the data from the Mites devices an adversary needs to get access to the system, understand how information is stored in the system, find the mapping of the obfuscated location information to the actual room location, identify the sensor data related to the individual room's location, and finally understand what data values mean. While the individuals assigned to an office can be

associated with an office location and thus a Mite device, the risk that an individual's behavior will be exposed due to Mite data collection is extremely low due to the featurization of the data and our protocol for protecting this data. The risk is lower than the risk of exposing an individual's behavior based on data collected by other sensors in the building such as sensors in the existing Building Management Systems and the WiFi infrastructure for example.

Category: Security Best Practices for the Mites Infrastructure:

FAQ#S1: What are the security practices that are followed by the Mites infrastructure? How are the Mites devices themselves secured?

The Mites devices are custom designed by our research team and use an industrial-standard compute module with WiFi from a well-known IoT vendor (Particle). We largely use Particle's Device OS and build on top of it with our own application firmware and backend. Our device firmware is pre-programmed with asymmetric cryptographic keys to mutually authenticate themselves with our backend and use that to derive symmetric cryptographic keys using industry-standard protocols, for secure end-to-end communication. Notably, the Mites devices are pre-programmed to only communicate with our backend server and no other host using our custom and proprietary packet format. Finally, the Mites are on a non-internet routable campus encrypted WiFi network that is unreachable from outside the campus for additional security. They are installed using screws in the walls and ceiling by the facilities personnel in TCS Hall to prevent easy removal. We also get network heartbeats from the Mites to be able to account for all the sensors and their network status.

FAQ#S2: What happens when there is a Security Vulnerability?

The researchers will constantly evaluate the performance and security of the Mites. If a concern arises regarding security, we will ensure that all events are reported to the IRB in a timely manner and immediate necessary actions are taken to mitigate any risk.

FAQ#S3: Is the source code audited by a third party to prove that the implementation is secure?

No, the source code is not audited by a third party. However, we use a highly customized version of a mature middleware building OS, called BuildingDepot, which is open source and available for everybody.

FAQ#S4: What is the life cycle plan for the Mites devices and the data collected from them?

We are planning for this project to continue for several years and we shall keep the IRB protocol in place for the entire period. Once the project ends we plan to delete the data collected and have a decommissioning plan for the Mites.