SUSTAIN Deliverable

D8.3 Data Management Plan

Grant Agreement number	101071179
Action Acronym	SUSTAIN
Action Title	Smart Building Sensitive to Daily Sentiment
Type of action:	HORIZON EIC Grants
Version date of the Annex I against which the assessment will be made	28th March 2022
Start date of the project	1st October 2022
Due date of the deliverable	M6
Actual date of submission	09.05.2023
Lead beneficiary for the deliverable	AALTO
Dissemination level of the deliverable	Public

Action coordinator's scientific representative

Prof. Stephan Sigg
AALTO –KORKEAKOULUSÄÄTIÖ,
Aalto University School of Electrical Engineering, Department of Communications and Networking
stephan.sigg@aalto.fi

Funded by the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Innovation Council and SMEs Executive Agency (EISMEA). Neither the European Union nor the granting authority can be held responsible for them.

European Innovation Council

Authors in alphabetical order		
Name	**Beneficiary**	**e-mail**
Name (from A's to Z's)	SHORT name of organization	name.name@org.de
Stephan Sigg	AALTO	stephan.sigg@aalto.fi

Abstract
This Data Management Plan is a living document which presents the SUSTAIN project handling of data. The objective of creating this document using tools such as OpenAIRE ARGOS, is to make data findable, accessible, interoperable and reusable (FAIR). This Deliverable links to D8.7 Update on Data Management Plan and D8.8 Final update on Data Management Plan, which are updated versions of this document.

Contents

1 Data Summary

Existing data will be used for initial development of the RF-sensing algorithms in WP2 for embedded probabilistic learning, in WP3 for distributed intelligence, in WP4 for Communication and Sensing, as well as for WP6 on the development of the encryption algorithms. In particular, we will exploit the binary I/Q data and point cloud data from https://zenodo.org/record/4459969 as well as from https://version.aalto.fi/gitlab/salamid1/sidelinkchanneldataset.

The project will further conduct new RF measurements containing signal strength fluctuation at receive antennas while activities (movement, gestures) are conducted in the proximity of the receiver. The experiments are carried out at two locations, an anechoic chamber at Aalto University as well as within a pilot building (Aalto Collider building, Otakaari 3b).

Specifically, the new data generated during the project will comprise new CSI (binary I/Q samples, as well as in form of x-y-z-time point cloud format after processing (e.g. range and angle Doppler)) data from radio channel measurements in T4.2, T4.3, T4.4, T7.3. Data will be mostly stored and processed in ascii dat/txt or python pkl files and might contain labels, raw-data and feature values. Data will not contain any real names, gender, or age information. Experiments will be planned beforehand, avoiding unnecessary redundancy. The new data is collected for the purpose of reaching the project objectives:

O2: developing a self-organizing distributed intelligence by auto-selecting the most suitable sensor inputs and deep learning overlay.

O3: Unprecedented accuracy short and long-range RF-based localization, tracking, gesture and sentiment recognition. Furthermore, the data will be exploited as open data-sets and for the design of demonstrator prototypes.

As such, the data will be useful for the project partners to reach the project objectives. In addition, it will be useful for fellow researchers that exploit the open data-sets for benchmarking and testing alternative RF sensing methodologies. We expect that 10GB-10TB will be stored during the respective project phases.

2 FAIR Data

The following sub-chapters explain how the FAIR principles will be implemented in SUSTAIN.

2.1 Making Data Findable, Including provisions for Metadata

As naming convention, the data will be distinguishable via <MeasurementSite>_<timestamp> _<Data/FeatureType>_<ClassLabel>_<DeviceID>_<AntennaNo> We will include search keywords in the documentation files stored together with the data.

The documentation of data files will be provided as separate text files together with the respective data files. Metadata standards will be used (e.g., Dublin Core [DC], Resource Description Format [RDF], the Data Documentation Initiative [DDI].

2.2 Making Data Accessible

During data collection, data will be written to storage media and backed up securely at Aalto University (i.e. aalto.version.fi) with access restricted to the project researcher conducting the measurements and the PI Stephan Sigg. To guarantee sufficient anonymization, the data collected in the field, for instance in WP7, will be stored using storage media of project partners (e.g. version.aalto.fi) that implements regular backup service and data encryption. Most data collected during the experiments will not be shared with project partners or the general public and will be deleted after few weeks or days as much of the data collected is uninformative with respect to RF-perception and mainly used for maintaining the connection.

All published data will be made available via publicly available data storages https://version.aalto.fi, https://zenodo.org, and via the European Open Science Cloud: eosc-portal.eu), will be rigorously anonymized (e.g. using tools like OpenAIRE AMNESIA https://amnesia.openaire.eu/). The zenodo repository ensures that the data is assigned an identifier and resolves the identifier to a digital object.

The developed source code of the algorithms will be stored in publicly accessible databases, such as https://version.aalto.fi, https://zenodo.org, and the European Open Science Cloud eosc-portal.eu.

The data and source code will be owned by the project partner generating the data. The aim is to publish the data and code as openly as possible after possible exploitation activities (e.g. patenting) have been thoroughly analysed. The open source code will be extensively commented to ensure that the code is as intelligible as possible and can be utilized by others after it is published.

The source code and data will be also linked to scientific publications through which the potential users will find out about it. These publications will also describe the tools and instruments necessary for validating the results based on the data.

As specified in the Grant Agreement, the data and metadata will be licensed under the GNU General Public License 3 (GPL3) no later than three months after the project end and possibly sooner, as much as the delay through rigorous anonymization permits. Metadata will contain information to enable the user to access the data. The PI Stephan Sigg will ensure that data will be available and findable for further use at least 5 years after the project end through publicly accessible databases, such as https://version.aalto.fi, https://zenodo.org, and through he European Open Science Cloud eosc-portal.eu. Should all services become

unavailable within this timeframe, the PI will ensure that the data is transferred to an alternative appropriate location that ensures assigning an identifier and to resolve it to a digital object, as well as storing metadata alongside the data.

The scientific articles produced in the project will be stored as pre-publication versions in ArXiv (https://arxiv.org) and also other open access publication archives. In particular, All project dissemination will be green open Access, making them available via University archives, such as Aalto's Aaltodoc Open Access database Acris (research.aalto.fi).

We don't foresee a need for a data access committee (e.g. to evaluate/approve access requests to personal/sensitive data) and no need to ascertain the identity of the persons accessing the data.

Some of the data will require a text editor capable to interpret ascii documents, while other data will require the python object serialization library pickle (https://docs.python.org/3/library/pickle.html). Neither an ascii-capable text editor nor the python library will be included in the repository.

2.3 Making Data Interoperable

Metadata standards will be used (e.g., Dublin Core [DC], Resource Description Format [RDF], the Data Documentation Initiative [DDI].

New CSI data will comprise binary I/Q samples, as well as x-y-z-time point cloud format after processing (e.g. range and angle Doppler)

As naming convention, the data will be distinguishable via <MeasurementSite>_<timestamp>_<Data/FeatureType>_<ClassLabel>_<DeviceID>_<AntennaNo>

We will include search keywords in the documentation files stored together with the data.

Publications in the frame of the project may include qualified references to other data if such data is used in the respective research study.

2.4 Increase Data Re-use

The open-source code will be extensively commented to ensure that the code is as intelligible as possible and can be utilized by others after it is published.

The documentation of data files will be provided as separate text files together with the respective data files, including a description on the provenance of the data (Data Documentation Initiative [DDI]).

The data and software will be licensed under the GNU General Public License 3 (GPL3) no later than three months after the project end and possibly sooner, as much as the delay through rigorous anonymization permits. Data will be available for further use at least 5 years after the project end.

Third parties can use the data under GPL3.

To ensure high data re-use, the documentation quality will be reviewed by external experts active in the field. In particular, we intend to collaborate in this regard with Professor Moustafa Youssef (Alexandria University,

Egypt), Professor Matthias Pätzold (University of Agder, Norway), and Professor Yasamin Mostofi (UC Santa Barbara, US).

Data quality assurance will be assured through annual screening of the data produced and stored. This includes detection of anomalies, identifying and removing obsolete information, as well as data cleaning. In particular, Aalto will in the frame of Task 8.4 be responsible for this data management in the project.

3 Other Research Outputs

The data and software will be licensed under the GNU General Public License 3 (GPL3) no later than three months after the project end and possibly sooner, as much as the delay through rigorous anonymization permits. Data will be available for further use at least 5 years after the project end.

The developed source code of the algorithms will be stored in publicly accessible databases, such as https://version.aalto.fi, https://zenodo.org, and the European Open Science Cloud eosc-portal.eu.

The data and source code will be owned by the project partner generating the data. The aim is to publish the data and code as openly as possible after possible exploitation activities (e.g. patenting) have been thoroughly analysed.

The open-source code will be extensively commented to ensure that the code is as intelligible as possible and can be utilized by others after it is published.

The source code and data will be also linked to scientific publications through which the potential users will find out about it. These publications will also describe the tools and instruments necessary for validating the results based on the data.

Other research outputs will be managed according to the FAIR data principles, in the same way as it is implemented for the data generated throughout the project (see description above).

4 Allocation of Resources

Data management is conducted by Aalto in the frame of T8.4 Data management. The open databases and tools used in the project do not have a cost (e.g. https://arxiv.org, https://zenodo.org, eosc-portal.eu, https://amnesia.openaire.eu). The services provided by Aalto University (https://version.aalto.fi, research.aalto.fi) are offered to the project free of charge.

The PI Stephan Sigg will ensure that data will be available and findable for further use at least 5 years after the project end through publicly accessible databases, such as https://version.aalto.fi, https://zenodo.org, and through the European Open Science Cloud eosc-portal.eu. Should all services become unavailable within this timeframe, the PI will ensure that the data is transferred to an alternative appropriate location that ensures assigning an identifier and to resolve it to a digital object, as well as storing metadata alongside the data.

5 Data Security

During data collection, data will be written to storage media and backed up securely at Aalto University (i.e. aalto.version.fi) with access restricted to the project researcher conducting the measurements and the PI Stephan Sigg. To guarantee sufficient anonymization, the data collected in the field, for instance in WP7, will be stored using storage media of project partners (e.g. version.aalto.fi) that implements regular backup service and data encryption.

Most data collected during the experiments will not be shared with project partners or the general public and will be deleted after a few weeks or days as much of the data collected is uninformative with respect to RF-perception and mainly used for maintaining the connection.

All published data will be made available via publicly available data storages (https://version.aalto.fi, https://zenodo.org, and via the European Open Science Cloud: eosc-portal.eu), will be rigorously anonymized (e.g. using tools like OpenAIRE AMNESIA https://amnesia.openaire.eu/). The Zenodo repository ensures that the data is assigned an identifier and resolves the identifier to a digital object.

The developed source code of the algorithms will be stored in publicly accessible databases, such as https://version.aalto.fi, https://zenodo.org, and the European Open Science Cloud eosc-portal.eu.

The data and source code will be owned by the project partner generating the data. The aim is to publish the data and code as openly as possible after possible exploitation activities (e.g. patenting) have been thoroughly analysed.

The PI Stephan Sigg will ensure that data will be available and findable for further use at least 5 years after the project end through publicly accessible databases, such as https://version.aalto.fi, https://zenodo.org, and through the European Open Science Cloud eosc-portal.eu. Should all services become unavailable within this timeframe, the PI will ensure that the data is transferred to an alternative appropriate location that ensures assigning an identifier and to resolve it to a digital object, as well as storing metadata alongside the data.

6 Ethics

The PI Stephan Sigg has initiated the Request for Statement Regarding Ethicality of Research Studies at Aalto University. The process is lengthy and is conducted by experienced Research Ethics experts, the Vice Rector for Research and a Legal Counsel with special expertise in Personal Data, Research Ethics and GDPR. The study has been described in detail and the information sheets, privacy notices (one for each of the three studies) and consent forms have been submitted for review. All comments have been addressed and the process is near final. The Decision was that the request can be accepted by the Chair after addressing the identified issues and their verification. When the issues can be accepted by the Chair, it means that only minor modifications have been requested.

The project will conduct new RF measurements containing signal strength fluctuation at receive antennas while activities (movement, gestures) are conducted in the proximity of the receiver. The experiments are carried out at two locations, an anechoic chamber at Aalto University as well as within a pilot building (Aalto Collider building, Otakaari 3b).

Specifically, the new data generated during the project will comprise new CSI (binary I/Q samples, as well as in form of x-y-z-time point cloud format after processing (e.g. range and angle Doppler)) data from radio channel measurements in T4.2, T4.3, T4.4, T7.3. Data will be mostly stored and processed in ascii dat/txt or python pkl files and might contain labels, raw-data and feature values. Data will not contain any real names, gender, or age information. Experiments will be planned beforehand, avoiding unnecessary redundancy.

As summarized above, WP4 and WP7 may raise ethical issues due to the data recorded for the development, training, implementation, verification and benchmarking of the algorithms and then stored in the backend. In particular, issues of informed consent and data privacy apply to this project since the case studies involve human participants who should only participate with informed consent, and the experiments will collect personal data related to activities and gestures. The study situation will be videoed, but this lifelike video will exist for less than a week. Within this time frame the will be technically manipulated to eliminate recognizable faces and bodies. The manipulation will be permanent so from that point in time personal data ceases to exist. Only the (irreversibly) anonymized video will be shared with partners.

Informed consent is addressed by designing the case studies accordingly. We take a responsible approach that complies with European (and national) legislation and fundamental ethics principles such as those reflected in the charter of fundamental rights of the European Union and the European Convention on Human Rights and its supplementary protocols. In all case studies, respect for people and for human dignity, fair distribution of research benefits as well as protection of the values, rights and interests of the research participants are ensured. No psychological, social, legal, economic or environmental harm is expected for the participants.

Participation in the case studies will be voluntarily. We will query for potential participants via public mailing lists and message boards of involved partners or for example in local markets. The announcement will be neutral and not in favour of, for instance, gender, age, social group, occupation. The sole requirement for participation will be that participants have to prove that they are adult in advance of the start of the case study.

Vulnerable groups and minors will not participate in the studies.

The experiments will be conducted for an exactly defined period of time, spanning not more than six hours continuously. The areas in which the experiments will be conducted will be clearly, visibly labelled and the area will be monitored by the researchers conducting the study at all time, to avoid accidental access by non-study participants. The experimental areas will be selected such that non-participating subjects may avoid the areas without inconvenience, e.g. it will not be necessary to pass the areas in order to reach any location in the building.

Participation will be rewarded with an insignificant gift card to the Aalto Shop. All participants are informed that they will at any time during the experiment have the opportunity to withdraw from the case study and/or request that their data be deleted. In advance of the studies, participants will be provided an informed consent form and detailed information sheets. These signed consent forms will be digitalised and are kept separate from the recorded data and statistical information. The informed consent forms and the detailed information sheets will be provided to prospective participants at least three months prior to the actual case studies. The information sheets, privacy notices and consent form have already been reviewed in the Ethical Board at Aalto University.

PERSONAL DATA:

Video footage will also be gathered during the three studies. This data is personal data, and it will be processed only by PI Stephan Sigg and a dedicated researcher, who will transform the video into a form that leaves people unrecognizable. This transformation is permanent and as soon as the transformation is complete, the real-life-like video will be permanently destroyed. After this the project will have no video form personal data in its possession. Only after video has been anonymized it will be shared with other researchers in the project, i.e. personal data will not be shared.

Since data is recorded for the participants (RF-fluctuation, CSI and I/Q data, holograms data structures), any possible privacy issues rising in connection to this will be investigated and handled appropriately.

We address these by anonymising all data and by separating personal, identifiable information from the actual recorded data (storage on physically distinct hard disks). Only that data will be recorded and stored which is absolutely necessary to conduct the described research. In particular, when dealing with human subject, will keep summary information on the participant count, age and gender distribution (but not the age and gender information for individual subjects) for statistical reasons and documentation of the case studies. Other than this, no personal information is collected from the participants. We will randomly permutate the order in which participant data is stored with respect to the order in which the statistical information about the participants is stored. To be precise: it will not be possible to associate a particular recorded dataset to the gender or age of the actual participant from the data stored. In addition, signedinformed consent forms will, short of the signature and name, contain no other information that might be used to associate them with the corresponding data recorded during the case study.

All participants are informed of this procedure in advance and will have the possibility to withdraw from the case study if they do not agree to these terms. In this case, the corresponding signed informed consent forms will be destroyed.

Personal data is also gathered in a minimal way outside the context of the actual research by way of the use of the SUSTAIN website. The personal data is processed in order to enable you to use the website; maintain and develop the website; enable communication; and monitor the number of visitors as part of the evaluation of dissemination impact, as the funder requires monitoring the impact of the dissemination work. Furthermore, the personal data is processed for the purposes of data security and in order to prevent and resolve possible misconduct.

The personal data collected and stored in the logs of the web server (nginx) currently are the IP address(es); the browser(s) used; and the page(s) opened (page request). The Privacy Notice can be seen in full as Annex 1 to this Data Management Plan.

7 Other Issues

Other than the procedures described in this document, we will not make use of other national, funder, sectorial, or departmental procedures for data management.

8 Conclusions

The Data Management Plan is a living document and the current version of it will evolve as the project implementation continues. The currently existing data management documents are attached to the Data Management Plan as annexes.

9 Annex 1 – Privacy Notice for SUSTAIN website

Privacy Notice for SUSTAIN website

SUSTAIN website serves as a medium through which the SUSTAIN project showcases the SUSTAIN activities and is part of the communications activities required to achieve the impact of the SUSTAIN project. In order to achieve these purposes, the website collects personal data that relates to the visitors of the website.

Aalto University ("Aalto"), also the coordinator of the SUSTAIN consortium, is responsible for the execution and upkeep of the SUSTAIN website and the designated contact point for data subjects concerning the use of personal data.

Other members of the SUSTAIN consortium are University of Trento, Institut Mines-Telecom, Yildiz Technical University and Universitat Politecnica de Catalunya. The SUSTAIN project receives funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement 101071179. For sake of clarity, Aalto does not share personal data collected from the SUSTAIN website between the other members of the consortium.

Aalto is committed to complying with the requirements of the General Data Protection Regulation ("GDPR") and applicable national legislation in the processing of your personal data. The means and purposes of processing your personal data are described in further detail in this privacy notice (the "Privacy Notice").

This Privacy Notice is subject to change and it will be updated when changes are implemented. You may always find the up-to-date version of this Privacy Notice on this website.

This Privacy Notice covers the following areas:

-Why does SUSTAIN website process personal data?
-What personal data is processed?
-Sources of information
-Legal basis for processing
-Sharing of personal data
-International transfers of personal data
-Retention period
-Your rights as SUSTAIN website visitor
-Contact information

1. Why does SUSTAIN website process personal data?

Your personal data is processed in order to
-enable you to use the website;
-maintain and develop the website;
-enable communication; and
-monitor the number of visitors as part of the evaluation of dissemination impact, as the funder requires monitoring the impact of the dissemination work.

-Further, your personal data is processed for the purposes of data security and in order to prevent and resolve possible misconduct.

2. What personal data is processed?

Following personal data is collected and stored in the logs of the web server (nginx):

- the IP address(es);
- the browser(s) used; and
- the page(s) opened (page request).

3. Sources of information

Personal data is collected from the visitors themselves during their visit(s) on the SUSTAIN website.

4. Legal basis for processing

The legal basis for processing your personal data is the legitimate interest to maintain and develop the website and to communicate the activities of the SUSTAIN project as required for monitoring the impact of the SUSTAIN project.

Legitimate interest is the legal basis for:

-personal data processed for communications purposes;
-information collected through observing the use of the SUSTAIN website; and
-processing of personal data for the purposes of data security and to prevent and resolve possible misconduct.

5. Sharing of personal data

Aalto shares personal data only to the extent necessary for the purposes personal data is processed:

I) Service providers

Aalto may, in the maintenance of the SUSTAIN website and in its service provision, use service providers to process data for the purposes detailed in this Privacy Notice.

II) Statutory reasons

Aalto may disclose your personal data to third parties if access to personal data or other processing of personal data is required to

i) fulfill statutory responsibilities or a court order; and/or
ii) detect, prevent or handle misuse, security risks or technical issues.

6. International transfers of personal data

The server on which the SUSTAIN website is operated is located in the EU. Aalto strives to carry out all services related to the SUSTAIN website by using operators and services located within the EU or the EEA. In some cases, however, services related to the use of the website may also be carried out by operators and on servers

located in third countries. In such cases, your personal data may also be transferred outside the EU or EEA in accordance with applicable legislation.

7. Retention period

Personal data will be retained for the period of validity of the legal basis for processing and for as long as necessary for the processing purposes mentioned in this Privacy Notice. Aalto determines the validity of such in accordance with, e.g., your use of the website as well as the communication between the visitor(s) and Aalto.

8. Your rights as SUSTAIN website visitor

You have rights related to the processing of your personal data. The extent of your rights is tied to the legal basis for processing your data and applicable legislation.

The right to access data: You have the right to receive confirmation on whether we process personal data relating to you and the right to access such personal data. We may ask you to specify your request where necessary, for example with regard to the details of the provision of information.

The right to rectification: You have the right to rectify erroneous or incomplete data.

The right to be forgotten: You have the right to request the removal of your personal data related to your website use. We will comply with your request, provided that there is no legitimate reason to retain the data, e.g. a statutory obligation to continue processing it. Personal data may not be deleted instantly from backup copies and other such data systems, but will be deleted through regular database retention practices.

The right to object to data processing: You have the right to object to the processing of your personal data related to your website use, if your personal data is processed for other purposes than the fulfillment of legal responsibilities or the provision of services.

The right to restrict processing: If you contest the correctness of the data registered about you or the lawfulness of the processing, or if you have objected to the processing of the data in accordance with your right to object, you may request us to restrict the processing of such data to storage only. The processing will be restricted to storage until the correctness of the data can be verified, or until it is assured that our legitimate interests override your right to restrict our processing activities.

If you do not have the right to request the removal of your personal data, you may instead request the restriction of data processing to storage only. If the processing of the data which we have registered about you is solely necessary to assert a legal claim, you may also demand that other processing of such data is restricted to storage. We may process your data for other purposes if this is necessary to assert a legal claim.

9. Contact information

You may ask questions concerning the use of your personal data on the SUSTAIN website and use your rights referred to above by contacting the SUSTAIN consortium's designated contact point for the use of personal data on this website.

Contact person

Tuija Heikura
Aalto University Foundation sr
Otakaari 1, Espoo 02150, Finland

Regarding eventual questions or demands relating to the processing of personal data, please contact Aalto Data Protection Officer, Ms. Anni Tuomela.

Data Protection Officer
Anni Tuomela
anni.tuomela@aalto.fi
Phone number: +358 (9) 47001

10 Annex 2 – Participant Information Sheet, Study 1

Case Study 1: Smart Building Sensitive to Daily Sentiment

Horizon Europe,
EIC Pathfinder Challenge project
SUSTAIN (101071179)

May 2, 2023

Contact: Stephan Sigg
Phone: 0504666941
Email: stephan.sigg@aalto.fi

General description of the study method

Thank you for your participation in the case study. The purpose of the study is the detection of location as well as the tracking of movement of human subjects in an indoor environment. The study will be conducted in the Aalto Collider building, Otakaari 2 (cf. Figure 1a).

For the purpose of this study, sensing devices will be installed in the area where the case study is conducted. Together with other participants, you will be instructed to move around in the area according to a specific pattern. The exact movement details will be discussed with you at the measurement site.
The sensors utilized in the study are standard WiFi routers as well as purpose-build RFID devices, similar to the ones depicted in figure 1b. Particularly, the signal strength, phase and Frequency patterns observed from electromagnetic waves (2.4GHz and 5GHz standard WiFi transmission, as well as 433 MHz RFID) measured in the measurement area are analyzed to estimate the location and movement of human subjects within the area. The signals utilized in the study are according to international standards and are harmless.
Video footage will also be gathered during the study. In a matter of days, the video will be transformed in a way that makes people unrecognizable. The real-life-like video will be permanently deleted immediately after the transformation has been done. The video will not be processed outside Aalto. Only the transformed video will be shared to other researchers in the project.

(a) The Otakaari 2 building (b) Backscatter prototypes

Figure 1: Location of the case study and prototype sensor devices

1

The research group conducting the study has been conducting similar research studies utilizing electromagnetic signals for the detection of environmental infromation since 2009.

Purpose and nature of the research

The study conducted shall demonstrate distributed learning of WiFi and RFID-based passive localization and tracking within the Aalto Collider building. In particular, we will localize non-transceiver-equipped subjects via RFID installations and count the number of people present at various locations in the building. A prototype application will be developed to display the occupancy of rooms and corridors of the building in real-time. An additional functionality of the instrumentation, demonstrating the tracking capability will be the visualization of real-time crowd flows across the various areas of the building.

The data generated during the study will comprise CSI (binary I/Q samples, as well as in form of x-y-z-time point cloud format after processing (e.g. range and angle Doppler)) data from radio channel measurements. Data will be mostly stored and processed in ascii dat/txt or python pkl files and might contain labels, raw-data and feature values. Data will not contain any real names, gender, or age information. The video footage, which will be transformed into a form where people cannot be recognized, is necessary to support the analysis of the data gathered.

Compensation

Participation in the study is rewarded with a gift card to Aalto Shop with a value of 30 EUR. Some coffee, tea and light refreshments will be made available to support the participation.

Voluntary nature of the study

Participation in the study is voluntary. All participants may at all time during the experiment withdraw from the case study and/or request that their data be deleted without obligation to disclose any specific reasons for discontinuing. Participation is conditioned on the signing of an informed Consent Form. These signed forms will not be published and are kept separate from the recorded data and statistical information.

Communication with the research staff

At all time during the case study, research staff will be present to help and assist. Please feel free to reach out to the research staff at any time and with any question you might have.

Study duration

Please be aware that your study session may last for up to three hours.

Additional information

Measures to be taken in cases of unexpected incidental findings:
The tests have not been designed to provide any clinical information. Nonetheless, in the event that the tests happen to reveal a markedly nontypical unexpected finding, which the data subject is not aware of, a specialist doctor of the appropriate field will be consulted about the finding. Based on his/her evaluation, the data subject may be recommended, if necessary, to seek appropriate further examinations.

Insurance coverage:
Study participants (data subjects) are covered by an insurance policy, purchased by Aalto University, to cover accidents and damages occurring during the test (also internationally). The Information Sheet shall be signed by the researcher in charge (with his/her contact information) on behalf of the research group. No unnatural physical activity will be requred and we do not forsee any risks of injury that exceed the risk of normal leasure walk or shopping stroll. Please wear shoes when you participate in the study.

11 Annex 3 – Research Study Privacy Notice, Study 1

Aalto University

EIC pathfinder challenge project SUSTAIN Privacy Notice

This privacy notice describes how your personal data will be used in the research study as an individual participating in the research study. You have also been provided with document called "Participant Information Sheet", which explains in more detail how the study is carried out.

1. What is being studied in this research study and the purpose of processing personal data

The study conducted shall demonstrate distributed learning of WiFi and RFID-based passive localization and tracking within the Aalto Collider building. In particular, we will localize non-transceiver-equipped subjects via RFID installations and count the number of people present at various locations in the building. A prototype application will be developed to display the occupancy of rooms and corridors of the building in real-time. An additional functionality of the instrumentation, demonstrating the tracking capability will be the visualization of real-time crowd flows across the various areas of the building.

This research is funded by the European Commission.

2. What personal data is processed in the research study

The data collected comprises categories of gesture, location and posture.

From gesture and posture, we intend to infer patterns over time, which are interpreted as sentiment. (ground truth measurements on sentiment is not needed since these patterns are instructed to the participants and they will indeed 'play' the respective sentiment)

The data generated will comprise timestamps, channel-state-information (CSI) in the form of binary I/Q samples, as well as in form of x-y-z-time point cloud format after processing (e.g. range and angle Doppler)) data from radio channel measurements. Data will be stored and processed in ascii dat/txt or python pkl files and might contain labels, raw-data and feature values (e.g. variance, mean, max, min, cepstrum, etc). Data will not contain any real names, gender, or age information.

The data is gathered from readings of the radio channel on the low-power backscatter sensor node developed in the project. Ground truth data will be generated using the TI AWR 1445 class of mmWave radar sensors. The TI radar has a point-cloud and a I/Q modus. We will extract data exclusively using the point-cloud modus, which has limited accuracy and e.g. does not allow the recognition of respiration or heart rate of subjects.

Video footage will also be gathered during the study. This data is personal data and it will be processed only by PI Stephan Sigg and a dedicated researcher, who will transform the video into a form that leaves people unrecognizable. This transformation is permanent and as soon as the transformation is complete, the real-life-like video will be permanently destroyed. After this the project will have no personal data in its possession. Only after video has been

anonymized it will be shared with other researchers in the project, i.e. personal data will not be shared.

Furthermore, we will collect signed Consent Forms containing a signature and name of the participant, but no address, gender or age. The Consent Forms will be digitalized and stored in a secure Network Storage System folder with access granted to only PI Stephan Sigg and the Project Manager of the project. The Consent Forms will be deleted three years after the end of the project. Consent Forms are required to be saved for situations where the EC may conduct an audit on the project.

Special categories of personal data (sensitive personal data)

Data belonging to special categories of personal data or other specially protected personal data **will not be processed** in the research study.

3. Processing of necessary personal data

The research study only processes personal data that is necessary for the purpose of the study. The video footage, which will be transformed into a form where people cannot be recognized, is necessary to support the analysis of the data gathered in other ways. This personal data will exist for only less than a week after it has been captured. Within this timeframe it will be transformed into a form where people are not recognisable, which is a form of irreversible anonymization, as a result of which the footage is no longer personal data. Only the PI Stephan Sigg and a researcher conducting this transformation will have access to and modify the data. The research data is processed during the research in a manner, in which a participant will not be identifiable. The participant name will not be associated with research data as only a random identifier will be used for participants.

4. Legal basis for the processing of personal data

The processing of personal data is based on the data subject's consent or explicit consent.

Participation in the study is voluntary. All participants may at any time during the experiment withdraw from the study and/or request that their data be deleted without obligation to disclose any specific reasons for discontinuing. The participant is given an Information Sheet explaining the study and based on being informed, his/her participation is pending on the signing of a Consent Form. These forms will be digitalised on a Network Storage System folder with access granted to only to the PI Stephan Sigg and the Project Manager. They will be kept separate from the recorded data and statistical information.

5. Sharing personal data

No personal data is shared with project partners.

Research data containing a random identifier is shared with the following parties:

Rigorously anonymized data may be shared via publicly available data storages (https://version.aalto.fi, https://zenodo.org, and via the European Open Science Cloud: eosc-portal.eu), e.g. using tools like OpenAIRE AMNESIA https://amnesia.openaire.eu/.

6. International data transfers

No personal data will be transferred internationally. Permanently anonymized research data will be transferred to non-EU / EEA countries or international organizations.

7. Storage and protection of personal data

Your personal data is processed and preserved in secure IT systems, which are approved by Aalto University and suitable for personal data. Access to all computers and IT systems are protected by username and strong personal password. Your personal data will be stored on a password-protected Network Storage System folder at Aalto University with access granted to only to the PI Stephan Sigg and the Project Manager (Consent Forms) or PI Stephan Sigg and a researcher conducting the anonymization (video footage). The Network Storage System has been evaluated to be suitable for storage of personal data.

Information processed in information systems: anonymized data is stored on a university network drive, restricted access, user IDs and firewall

Processing of direct identifiers:
Direct identifiers are not used during the study

The research data is archived: Data management is conducted by Aalto in the frame of T8.4 Data management. The open databases and tools used in the project do not have a cost (e.g. https://arxiv.org, https://zenodo.org, eosc-portal.eu, https://amnesia.openaire.eu). The services provided by Aalto University (https://version.aalto.fi, research.aalto.fi) are offered to the project free of charge.

The PI Stephan Sigg will ensure that data will be available and findable for further use at least 5 years after the project end through publicly accessible databases, such as https://version.aalto.fi, https://zenodo.org, and through the European Open Science Cloud eosc-portal.eu. Should all services become unavailable within this timeframe, the PI will ensure that the data is transferred to an alternative appropriate location that ensures assigning an identifier and to resolve it to a digital object, as well as storing metadata alongside the data.

8. Retention and deletion of personal data

Deletion during and after the study

The video data is transformed into anonymous format within a week of data the recording. After that there is no personal data derived from the research study. We avoid having any other personal data by anonymising all other collected data and by separating personal, identifiable information from the actual recorded data. Only that data will be recorded and stored which is necessary to conduct the described research. In particular, we will keep summary information on the participant count, age and gender distribution (but not the age and gender information for individual subjects) for statistical reasons and documentation of the case studies. Other than this, no personal information is collected from the participants. We will randomly permutate the order in which participant data is stored with respect to the order in which the statistical information about the participants is stored. To be precise: it will not be possible to associate a particular recorded dataset to the gender or age of the actual participant from the data stored. In addition, signed informed consent forms will, short of the signature, contain no information that might be used to associate them with the corresponding data recorded during the case study.

All participants are informed of this procedure in advance and will have the possibility to withdraw from the case study if they do not agree to these terms. In this case, the corresponding signed informed consent forms will be destroyed.

The Consent Forms with signature and name will be digitalized and deleted three years after the end of the project.

We will use and maintain a generation procedure to (re-)generate an anonymous identifier key also after the completion of the study. In particular, the identifier may be of a form like <2nd letter of the first name of the mother [a..z]> <last digit of the day of birth of the grandmother [0..9]> <length (number of letters) of the name of birth-town [1..?]>, <color of favorite ice cream [white, green, blue, …]> etc. (for instance, 'f4' or 'dk' or 'purpleF') . In the case that a pair of study subjects has the same identifier, we will swap the order of digits/letters, use capitalized/non-capitalized letters or use a neighbouring digit or letter instead. Subjects may decide to provide a digit different from the one requested (for instance, when the 2nd letter of the first name of the mother is 'e', provide instead 'j'. Hence, the subjects are able to further anonymize their identifier without informing about it. The identifier will be hand-written on a piece of paper and handed to the participants to help them to remember the identifier for possible later withdrawal of their consent, should they wish to do so.
The identifiers (for instance 'f4') will not be published but kept with a linked generic descriptor (e.g. subject-XX, order of XX permuted from the actual order of participation in the study, for instance 'subject-08') on encrypted and password protected storage space in Aalto University. Data will be published using the generic descriptors.

The described procedure allows us to verify the participation of a subject in the study and to delete the data of this particular subject at any later point in time via a plausible proof but without the need for us to maintain identifiers that can be uniquely linked to a subject.

Anonymized research data is retained in order for research data to be used for further scientific research in the same scientific discipline or in other disciplines that support this research study. Research data may also be transferred to other universities or research organisations for further research projects.

9. Rights of the research participant

According to the General Data Protection Regulation (GDPR), a data subject has the right to:

- receive information on the processing of their personal data
- right to access the personal data collected and processed
- right to rectification of inaccurate personal data
- request that the processing of personal data be restricted
- object the processing of personal data
- right to erasure of personal data if the conditions of Article 17(1) of the Data Protection Regulation are met and processing is no longer necessary for archiving purposes in the public interest or for scientific research or statistical purposes in accordance with Article 89(1)

If the research purpose does not require, or no longer requires the identification of the data subject (note: this is not necessary at any point during the study), the controller shall not be obliged to obtain further information so that the data or the data subject may be identified only for purposes to able the data subject to exercise his/her rights. If the controller is unable to link

the data to a particular data subject, the data subject does not have the right to access or correct the personal data, object the processing, or delete the personal data. However, if the data subject provides additional information that allows their identification from the research data, the rights will not be restricted.

10. Contact details of the controller

The controller of this research study is Aalto University Foundation sr., operating as Aalto University.

Person in charge of the research study

Questions regarding the conduct of the research study may be addressed to the person in charge of the study: Stephan Sigg, 0504666941, Stephan.sigg@aalto.fi

Data Protection Officer

If the research participant has questions or requests related to data protection or the processing of personal data, the research participant should contact the Data Protection Officer of Aalto University: tel. +358 9 47001 (exchange), dpo@aalto.fi .

In this data request service, you can request the exercise of your rights under GDPR from Aalto University as the controller Home - Consent Management (aalto.fi)

If a participant of the research study feels that his or her personal data has been processed in violation of data protection legislation, the participant has the right to lodge a complaint with the supervisory authority, the Data Protection Ombudsman's office (read more: http://www.tietosuoja.fi).

12 Annex 4 – Consent Form

Participation confirmation Smart Building Sensitive to Daily Sentiment,

SUSTAIN Project

I have understood that participation is voluntary and at any point during the research study, I am at liberty to notify that I no longer wish to participate, and all information gathered up until that point will be deleted. Participation in the study means that the data gathered will be used as described in the Privacy Notice of the research study.

I have received sufficient information about the research study, I have had the possibility to have my questions answered, I have understood the information and I wish to participate in the research study.

Signature _____

Name of research participant _____

Contact details:

Principal Investigator: Stephan Sigg
Phone: 050 4666941
Email: stephan.sigg@aalto.fi
Aalto-University