DAMA – A transparent meta-assistant for data self-determination in smart environments

Christopher Ruff [1], Andrea Horch[2], Benedict Benthien[3], Wulf Loh[4] and Alexander Orlowski[5]

Abstract: Global sales of AI-based smart voice assistants and other smart devices are increasing every year. Smart devices are becoming ubiquitous, including living and workspaces. These spaces often have very high privacy requirements, like living rooms, bedrooms or meeting rooms in office environments. Users of smart devices have security and privacy concerns regarding personal data collection, data storage and the use of such data by the devices and the providers. These concerns are aggravated by a lack of transparency by the device manufacturers. As a result, users have limited possibilities to make an informed decision due to missing information or interfaces. While this leads to limited trust regarding the security and privacy of smart devices, for most users, the practical benefit dominates. The project DAMA wants to address user's security and privacy concerns by creating transparency and regulating the smart devices in connection with the respective context (e.g. when users are alone at home or when they have visitors). For this purpose, the project is developing a "meta-assistant", an assistant that regulates other AI-based assistants and other smart devices. It uses artificial intelligence (AI) for context detection and device regulation. The regulation processes are based on established ethical guidelines, which are adjusted to the project context.

Keywords: Smart Home, Smart Office, Smart Speaker, Privacy.

1 Introduction

The use of smart speakers like Amazon Echo or Google Home for services like music streams or search engines or to control other smart devices within a Smart (Home) environment via voice command are becoming more and more popular every year. According to [Br20a], global sales of Smart Speakers increased from 32.8 Million in 2017 to 86.5 Million, and were estimated to reach 91.3 Million by 2019. In a survey, introduced in [Br20b], 1.000 Germans indicated where they use their smart speakers in their Smart Homes. The result shows that the top four rooms equipped with smart speakers are the following: 67% living room, 53% kitchen, 44% office room, 43% bedroom. The survey of [Xu2020] shows the most popular functions of smart speakers. The top three functions are (1) asking general questions (55%), (2) getting information like weather, travel or

[1] Fraunhofer IAO, Identity Management, Nobelstr. 12, Stuttgart, 70569, christopher.ruff@iao.fraunhofer.de

[2] Fraunhofer IAO, Identity Management, Nobelstr. 12, Stuttgart, 70569, andrea.horch@iao.fraunhofer.de

[3] Fraunhofer IAO, Identity Management, Nobelstr. 12, Stuttgart, 70569, benedict.benthien@iao.fraunhofer.de

[4] Internationales Zentrum für Ethik in den Wissenschaften (IZEW), Wilhelmstr. 19, Tübingen, 72074, wulf.loh@izew.uni-tuebingen.de

[5] Internationales Zentrum für Ethik in den Wissenschaften (IZEW), Wilhelmstr. 19, Tübingen, 72074, alexander.orlowski@uni-tuebingen.de

sports updates (45%) and (3) the use of in-built music streaming services (35%). Home automation ranked on place seven (18%). The survey also states that at the end of 2019 an average U.S. home owned 9.2 connected devices. The most popular devices in the survey were phones, desktop and tablet computers, streaming boxes and sticks, connected TVs, gaming consoles, smart speakers, watches and thermostats. These surveys show that the use of smart speakers and other smart devices becomes more and more commonplace and the number of devices in use increases year by year.

Through the use of and interaction with Internet of Things (IoT) devices, a large amount of data is generated. Since personalized services are a central component here, the identification of the respective users creates profiles that contain a variety of personal data, from daily schedules, personal preference up to medical issues and biometric data used for identification [Wa2018]. As applications of smart devices in one's own home concern a specially protected private space, the trade-offs between the need for privacy protection on the one hand, and access to modern IoT technology on the other, are difficult not only from a data protection standpoint, but also from an ethical one [Hu2018]. Risks arise not only from consciously shared data, but also from reconfiguration with other existing data sets. These so-called *derivable data* [Gh2016] enable conclusions to be drawn about individuals far beyond their intended use. These possibilities are likely to undermine informational self-determination, i.e. in this case the possibility of users to control their shared data.

Due to the multitude of devices, the intransparency of data collection and processing, as well as the inconsistent and insufficient regulatory schemes, an understanding of the impact on privacy is often difficult to comprehend, even for experts. Therefore, DAMA seeks to empower users to regain their informational self-determination by giving back control over the devices and thus over the collected data. Regulatory efforts such as the General Data Protection Regulation (GDPR) address some of these issues. However, there are still gaps here [Wa2018], that prevent users from making informed privacy decisions as well as gaining adequate control over the privacy settings on their IoT devices. DAMA addresses this gap in the current legislation in order to create a possibility for users to use IoT devices in a self-determined way.

These issues not only concern legal and ethical challenges of IoT and Smart Home applications, but also user trust. A consumer study published in [HFA20] investigated the security and privacy concerns of smart home users concerning the smart devices in their homes. The results show a lack of transparency of the smart devices regarding data collection, data storage and the use of data by the devices and their providers. According to the study, the consumers are afraid of e.g. that their households are profiled, that the providers could sell their data, or that the government or someone from outside could get access to the data. The perceived lack of transparency may lead to a loss of confidence regarding the technology. The study also mentions that some of the users take simple and non-technical mitigation actions regarding their concerns, which do not have much effect, due to limited technical options or the lack of technical expertise. The lack of technical options can be traced back to the fact that most multi-agent systems in Smart Homes are closed systems, which use miscellaneous protocols, interfaces and standards as stated in [Fi2020]. Additionally, devices like smart speakers exhibit false detections regarding the wake word, thus being triggered without the user uttering the wake word [Sc2020]. Especially in situations subjectively considered as intimate, these false triggers confuse and unsettle the users and may lead to a loss of trust in the technology.

The research project DAMA (Data Autonomy Meta-Assistant) is addressing the users' needs as described above. The project aims to ensure privacy and informational self-determination - and thereby create trust - by developing a meta assistant called DAMA. This assistant will regulate the devices within the Smart Home environment in accordance with the respective context (e.g. resident is alone, resident has visitors) as well as depending on the sensors and functions of the devices. In order to secure privacy and increase trust, the meta-assistant will also be able to regulate itself, e.g. by turning off its own speech recognition. In order to maximize transparency, the meta-assistant informs the users about the privacy context DAMA has determined (alone, visitors, etc.), and which devices/sensors are currently active. The users also have the possibility to access this information at any time by asking DAMA. The respective context will be identified by using machine learning on sensor data and user input. The regulation will consider ethical aspects in order to enhance privacy, informational self-determination, and transparency for the users of Smart Home environments.

2 Related Work

The High-Level Expert Group on AI of the European Commission presented ethics guidelines for a trustworthy artificial intelligence (AI), which can serve as a starting point for the design of trustworthy AI systems [EU2019] like smart speakers or other smart devices. However, the report provides little guidance on how to implement the explicability of AI systems. In addition, DAMA makes use of the guidelines of the German Ethics Council on Big Data and AI [DE2017]. They go a little bit into more detail on how to establish control, transparency, and traceability for users, as well as conditions for explainability.

Large companies often have guidelines and policies to improve the trustworthiness of their AI products. Examples are the guidelines for AI of the Telekom [Fu2018], the baselines for AI of SAP [Ma2018] or the ethical AI principles of Microsoft [Bo2019].

In addition, there are different national and international initiatives for AI ethics standards, such as the IEEE P70xx series of ethical technology standards [Ci2019].

A practical approach to create trust for consumers of IoT devices is given in [Ca2021]. The authors describe a privacy and security label for IoT devices to enable consumers to make well-informed choices when they select an IoT device for their applications. The label informs consumers e.g. about included sensors, the data storage on the devices and in the cloud and the groups, the data collected by the device is shared with. Another initiative is [AI2020], in which a group of scientists and standardization experts made a first effort to move AI ethics “from principles to practice” and create an AI ethics label as well as introduce the idea of criticality levels for AI applications.

The goal of the EU project uTRUSTit (2010-2013) was to provide understandable security and trustworthiness feedback to the users of IoT devices in order to create transparency and trust. As described in [Ho2012] the project developed an overview of data sent by the smart devices, like device or network information and personal data. The overview is displayed on a mobile device like a mobile phone or tablet in a comprehensible and user-friendly manner.

The project PAPAYA (PlAtform for PrivAcY preserving data Analytics) [Pa2021] develops dedicated privacy preserving data analytics modules in order to enable users of devices and applications, which use third-party processors to analyse the user’s data, to get valuable information about the processed data.

Several research projects created a good set of Transparency Enhancing Technologies (TETs) which provide users with all necessary information about their data being stored, processed, exchanged and used by applications and devices. This data also includes private information of the users like voice or sensor data collected by smart devices in Smart Home environments. An overview of such technologies and tools is given in [Ja2013].

The BlockIT project introduced in [Lo2018] describes a blockchain-based solution for privacy preserving translucent data sharing on a smart IoT-aided grid by using smart contracts to regulate the interactions between the nodes in the grid. The blockchain provides transparency concerning the transactions and a wallet of pseudonyms offers a solution to preserve data privacy.

3 Objectives

The main objective of the DAMA project is to create more transparency and privacy for users of Smart Home applications and thereby increase their trust in the technology. To reach this goal, different context levels with different privacy settings are introduced. They describe typical situations in smart environments (use cases), for example: the resident of a Smart Home is alone or the resident has visitors. The devices within the smart environment are then configured and regulated automatically according to the respective context level that the users explicitly or implicitly set. The regulation can entail the activation or deactivation of specific devices or sensors in order to assure the user’s privacy needs and informational self-determination within the given situation. The users are notified about the deactivation and activation of devices or single sensors, which also helps to increase the transparency of the whole system. Additionally, users can retrieve the information about the status of all smart sensor devices at any time and check whether they are active or not, by using the assistant’s intuitive user interface. The algorithms and AIs in DAMA will detect the context of a situation and control the devices and configuration. In addition to the regulation of the other devices within the smart environment, DAMA also regulates itself, including a trustable, complete shutdown, to assure the adequate privacy level in situations of very high privacy needs.

4 Context Levels and Use Cases

The context levels are a convenient and practical way to regulate the Smart Home and its different applications. By changing context levels, it is easy for the users to react to a situational change without having to interact with all the different devices. At the same time, the user is given full control over the various devices while DAMA also displays the collected data by the devices. To make this an even more seamless user experience, DAMA will be able to automatically detect privacy patterns in user's behavior and sets context levels accordingly through a ML algorithm.

Three basic modes can be distinguished for the context levels:

1) Maximum functionality: all devices of the Smart Home / Smart Office are activated (context level 1).

2) Intermittent customizable modes - only certain sensors, displays and devices are activated; this depends on the respective situation and existing devices (context level 2-9).

3) Maximum privacy: all Smart Home / Smart Office devices are switched off. In addition, DAMA's microphones and speech recognition are deactivated (context level 10).

The second category is a compromise between functionality and privacy, depending on the user's preferences for a given situation. It is therefore not a single mode, but various customizable context levels are possible. By giving DAMA control over the other applications and their sensors, and letting the user self-define the context levels and when they are applied, the project maximizes the user's informational self-determination. The user can define for herself what constitutes an "appropriate flow of information" [Ni2010: 114] within her home. At the same time, the user receives status information about the active applications / sensors, which increases her informational self-determination [Lo18a]. In addition, the information about the data flows and uses provides her with the opportunity to make an informed decision whether and what data it opts to share in return for functionality [Lo2021].

For this, it is important to distinguish between sensors (for example, Alexa's microphone) and displays (for example, calendar entries on a smart mirror). The different context levels allow users to either hide their private data from visitors, e.g. the personal calendar displayed on a screen or a smart mirror. In addition, it allows them to protect their own and their visitor's privacy from datafication by the smart devices.

In the project, we identified different use cases, which represent different situations in which ideal-typical demands for privacy and functionality are raised. This allows us to elaborate and define the different context levels. The aim of these use cases is to create a solid basis that shows the usefulness of the system to enhance privacy and informational self-determination (proof of concept). In addition, a demonstrator will be developed that covers typical settings, but also is limited to certain idealized assumptions about intended use and usage environment. Based on this foundation created in the DAMA project, further use cases and requirements can be realized in the future, addressing a larger variety of usage environments as well as unintended usage. The development of DAMA also includes analyses of the range of various sensors. Specifically, the extent to which devices in adjacent rooms or on other floors can also collect data is examined.

The use cases identified by the project are:

- Use Case 1 - Resident comes home
- Use Case 2 - Resident gets a visitor in the evening
- Use Case 3 - Resident has a visitor with a child
- Use Case 4 - Craftsman or service provider must enter the flat

We will use a short example to explain the function of the context levels: After work, the resident of the Smart home comes into the house with a colleague. DAMA recognizes that the resident is not alone and therefore does not automatically switch to the mode that is usually selected for the end of the day: by playing a relaxation playlist and DAMA reading out private messages. Instead, DAMA switches off the private displays. In course of the evening, the conversation turns to important developments in the company that should not be made public, so the occupant switches to context level 10 "maximum privacy" - all devices and DAMA itself are switched off.

5 Technical Environment

The technical implementation of the AI prototype (meta-assistant DAMA) is work in progress. DAMA will be integrated into an existing Smart Home / Smart Office laboratory in the office building of the Fraunhofer IAO in Stuttgart. The laboratory is spatially divided into a *Smart Home section* and a *Smart Office section* as shown in Figure 1.

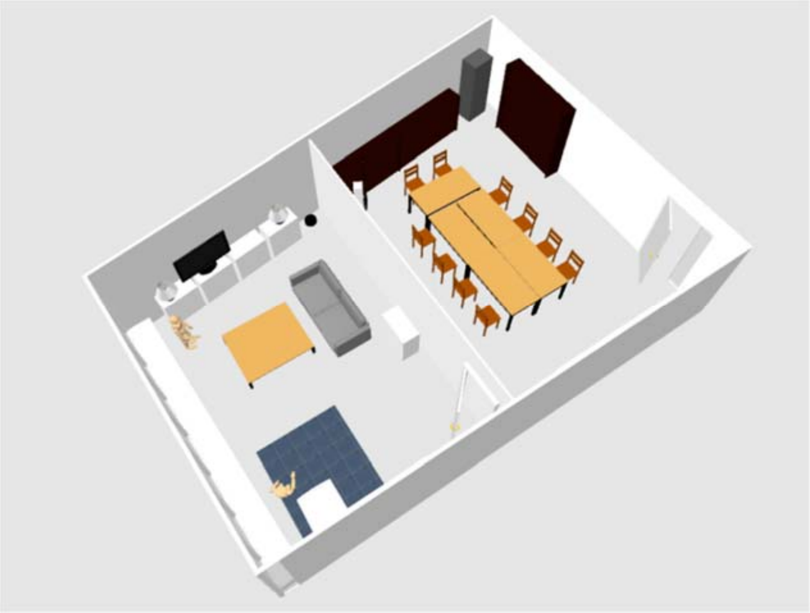

Fig. 1 - Smart Home / Smart Office laboratory

The Smart Home section includes devices like a smart fridge, a smart couch, a smart couch table, a smart mirror and a smart TV. Fig. 1 shows the Smart Home laboratory and some of its devices. The Smart Office section includes devices like a Mircosoft Surface Hub coffee table, a projector, a smart coffee machine and various environment-sensors, measuring things as e.g. the air quality, humidity or the air temperature in the room.

6 The Meta-Assistant »DAMA«

In the following, the technical concepts and architecture of the DAMA system components are detailed. The overall technical concept of the DAMA system can be viewed as a layered architecture, as seen in Fig. 3.

1. Context Detection Layer

The context detection layer includes all the components that gather information via sensory systems, like microphones in voice assistants, environmental sensors but also AI powered algorithms coupled with computer vision.

The basis for context detection is to derive information about who and how many people are present in the current smart home (or office) environment that the DAMA assistant has access to. As DAMA is designed to maximize privacy, the system will use minimally privacy-invasive detection technology and avoid biometrical recognition, such as face, voice, iris recognition etc. Currently, the system uses computer vision, based on OpenCV and machine learning to detect people entering the smart home. Cameras are located in less private areas like the entrance focusing only a small area and also only from above. The AI only tries to recognize people coming or leaving and does not record any biometrical information. Data processing is done locally.

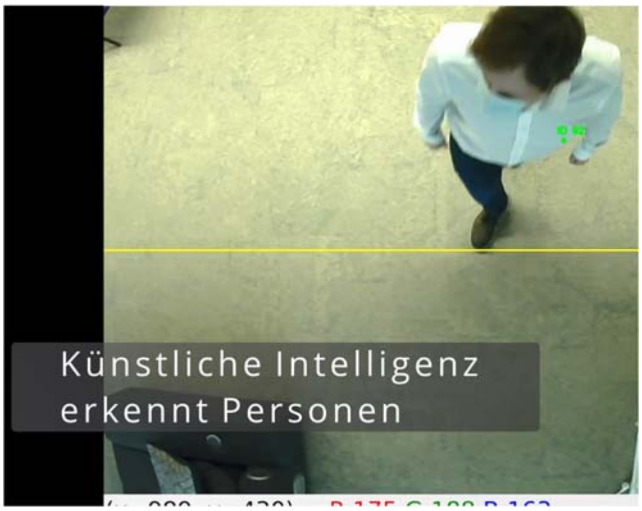

Fig. 2 - Detecting people entering the home using computer vision

This will be coupled with other systems such as infrared cameras and light barriers. The sensor systems are chosen to only collect the minimal amount of data necessary to tell whether someone is present in the smart. To identify the residents and their closer friends or partners, we are using MAC-scanners to identify smart phones of registered users to allow more fine tuned privacy settings. All devices can be switched off.

In order to solve the dilemma of introducing more sensors into the smart home to enhance privacy and informational self-determination, the DAMA system will focus on transparency. In general, it will only collect the information needed to identify the context (privacy by design) and avoid re-identification. The system is designed to run locally without sharing information with external servers. Furthermore, the system itself will shut down if requested by the user.

2. Controller

In the controller layer, the data and information collected by the various sensors and devices are collected and processed. Based on the user configuration, including the setting of different context levels, the DAMA Assistant's business logic is responsible for the processing and control of the devices in the Smart-Home through the Actor-Layer.

This layer will have a web-based user interface than can be shown on smart-TVs as well as other devices. To streamline and simplify connection to a variety of smart devices, smart home controllers are connected to the DAMA Assistant as well and are used where the devices support the required protocols. In the final version, data connection between devices and the controller will use TLS encryption to safeguard against tempering of the information sent.

3. Actors

The actor layer consists of all the smart devices the DAMA Assistant can control and regulate. The modular structure of the system will allow various devices and setups to work with minimal integration effort.

4. Transport layer

The DAMA Assistant needs to interoperate with a lot of devices and information sources, so an efficient bi-directional communication infrastructure is required. The MQTT [HTS08] protocol is lightweight, robust and as such capable to handle our requirements. Implementations of MQTT Clients are available for all major platforms and programming languages. While it can handle different communication scenarios, the publisher/subscriber using a MQTT-Broker as an intermediary is used in our context.

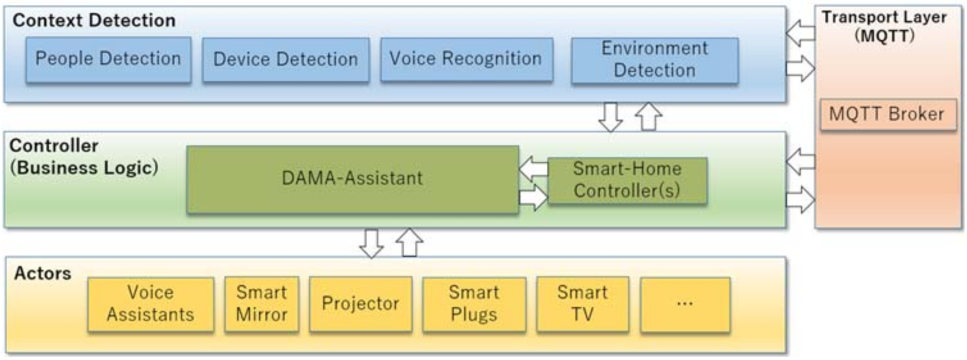

Fig. 3 - DAMA System Architecture

5. User Interface

The user interface is designed to address the needs of ordinary residents without a lot of technical knowledge. The system will be controlled via an intuitive user interface that allows registering new users, their respective smart device environment and their desired privacy context levels.

To facilitate the setup for less technically inclined users, default settings will be provided, based on a general classification of devices in terms of their possible impact on privacy. For the classification of devices, a privacy impact assessment is necessary. As common standards for classification are still in development, the classification will be based on a survey of existing approaches and criteria as of the “IT-Gütesiegel” certificate initiative of the German state government and the privacy and security label [GI14]. This and other sources are taken into account to create a project internal privacy matrix to structure this classification.

The DAMA backend logic is contained in several python scripts. Users can either call a web application to authenticate and change the context level directly, or it can be changed by sensors located inside the trusted smart environment. The MQTT Protocol is again used for the communication between DAMA, actors and sensors. Changes made to the context level are stored in a database with a timestamp and the used method for later reference.

7 Conclusion and Future Work

In this paper, we presented the underlying technical and ethical concept of a meta-assistant that is able to mitigate the privacy challenges that the use of smart assistants inevitably introduce, such as voice assistants and other smart devices that gather, display and share personal and potentially sensitive information in our living and work spaces. We tackle this problem by defining privacy related context levels, based on ethical values and assessments, which are subsequently implemented in a technical system that regulates said devices in terms of activity and data handling according to the current context.

By using these concepts, many of the challenges of dealing with a multitude of individual smart devices and their settings are hidden from the users, which in turn can assess on a much more abstract level, which privacy requirements they have for a given situation.

One of the challenges is the regulation of devices that do not allow users to have fine-grained control over said devices and do not provide interfaces to regulate them. While the DAMA assistant tries to use provided application interfaces when possible, lesser accessible devices are controlled by means of cutting power supply, shutters for visual devices or blocking network traffic. If interfaces allow, more fine-grained settings are possible, i.e. to activate or de-activate only part of the full functionality of the devices - for user-controlled periods.

The here presented system allows the user to exercise control and transparency over the multitude of smart devices while also enabling users to respect and cater to the needs of guests and visitors in terms of their privacy concerns, as well as their own. The work on the DAMA assistant is ongoing. In future work, privacy contexts and privacy settings are to be increasingly set by using machine learning to identify potentially privacy sensitive situations and automatically suggest the best system settings accordingly in more complex user scenarios.

8 Acknowledgement

We thank and acknowledge the Baden-Württemberg Stiftung for financing the DAMA

project. For more information, please visit: https://www.bwstiftung.de/.

Bibliography

[AI2020] AI Ethics Impact Group: From Principles to Practice – An interdisciplinary framework to operationalise AI ethics, VDE/Bertelsmann, https://www.ai-ethics-impact.org/, accessed 12/02/2021.

[Bo2019] Böhm T.: Künstliche Intelligenz und Ethik: Warum KI ethische Prinzipien braucht, um ein Erfolg zu werden, https://news.microsoft.com/de-de/ethik-prinzipien-kuenstliche-intelligenz/, accessed 04/02/2021.

[Br20a] Brandt, M.: Smart Speaker-Absatz - Amazon - Nummer 1 mit knappem Vorsprung, https://de.statista.com/infografik/20675/geschaetzter-weltweiter-smart-speaker-absatz/, accessed 26/01/2021.

[Br20b] Brandt, M.: Smart Speaker - Wo Alexa und Co. im Einsatz sind, https://de.statista.com/infografik/20414/orte-an-denen-smart-speaker-genutzt-werden/, accessed: 26/01/2021.

[Ca2021] Carnegie Mellon University: IoT Security & Privacy Label, https://www.iotsecurityprivacy.org/, accessed: 28/01/2021.

[Ci2019] Cihon P.: Standards for AI Governance: International Standards to Enable Global Coordination in AI Research & Development, https://www.fhi.ox.ac.uk/wp-content/uploads/Standards_-FHI-Technical-Report.pdf, accessed 04/02/2021.

[DE2017] Deutscher Ethikrat: Big Data and Health – Data Sovereignty as the Shaping of Informational Freedom. 2017.

[EU2019] European Commission, High-Level Expert Group on AI: Ethics Guidelines for Trustworthy Artificial Intelligence, https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai, accessed 04/02/2021.

[Fi2020] Firouzi R., Rahmani R., Kanter T.: An Autonomic IoT Gateway for Smart Home Using Fuzzy Logic Reasoner, Procedia Computer Science, Vol. 177, pp. 102-111, 2020.

[Fu2018] Fulde V.: Guidelines for Artificial Intelligence, https://www.telekom.com/en/company/digital-responsibility/details/artificial-intelligence-ai-guideline-524366, accessed 04/02/2021.

[Gh2016] Ghiglieri, M., Hansen, M., Nebel, M., Pärschke, J.V., Fhom, H.S.: Smart-TV und Privatheit. Bedrohungspotenziale und Handlungsmöglichkeiten. Hg. v. Forum Privatheit. 2016.

[GI14] IoT, Security & Privacy Label, Carnegie Mellon University, https://www.iotsecurityprivacy.org/, accessed: 13/04/2021.

[HFA20] Haney J. M., Furman S. M., Acar Y.: Smart Home Security and Privacy Mitigations: Consumer Perceptions, Practices, and Challenges, HCI for Cybersecurity, Privacy and Trust, pp. 393-411, 2020.

[Ho2012] Hochleitner C., Graf C., Unger D., Tscheligi M.: Making Devices Trustworthy: Security and Trust Feedback in the Internet of Things, Fourth International Workshop on Security and Privacy in Spontaneous Interaction and Mobile Phone Use (IWSSI/SPMU), Value4, 2012.

[HTS08] Hunkeler, U.; Truong, H. L.; Stanford-Clark.: "MQTT-S — A publish/subscribe protocol for Wireless Sensor Networks," 2008 3rd International Conference on Communication Systems Software and Middleware and Workshops (COMSWARE '08), Bangalore, India, 2008, pp. 791-798, doi: 10.1109/COMSWA.2008.4554519.

[Hu2018] Hummel, P.; Braun, M.; Augsberg, S.; Dabrock, P.: Sovereignty and Data Sharing. ITU Journal: ICT Discoveries, Special Issue No. 2, 23 Nov. 2018.

[Ja2013] Janic M., Wijbenga J. P., Veugen T.: Transparency Enhancing Tools (TETs): An Overview, 2013 Third Workshop on Socio-Technical Aspects in Security and Trust, New Orleans, LA, USA, 2013, pp. 18-25, doi: 10.1109/STAST.2013.11.

[Lo18a] Loh, W.: A Practice-Theoretical Account of Privacy, in: Ethics & Information Technology 20/4, pp. 233-247, 2018.

[Lo18b] Lombardi, F.: An IoT data sharing dilemma: Transparency or Translucency?, https://medium.com/cybersoton/an-iot-data-sharing-dilemma-transparency-or-translucency-9cd85ca278c3, accessed 06/04/2021.

[Lo2021] Loh, W.: Social Pathologies of Informational Privacy, in: Journal of Social Philosophy, forthcoming.

[Ma2018] Machmaier C.: Die Grundsätze für Künstliche Intelligenz von SAP, https://news.sap.com/germany/2018/09/ethische-grundsaetze-kuenstliche-intelligenz/, accessed 04/02/2021.

[Ni2010] Nissenbaum, H.: Privacy in context: Technology, policy, and the integrity of social life, Stanford Law Books, Stanford, 2010.

[Pa2021] PAPAYA Project: PAPAYA project website, http://www.papaya-project.eu/, accessed 06/04/2021.

[Sc2020] Schönherr L., Golla M., Eisenhofer T., Wiele J., Kolossa D., Holz T.: Unacceptable, where is my privacy? Exploring Accidental Triggers of Smart Speakers, arXiv.org, 2020, https://arxiv.org/abs/2008.00508, accessed: 04/02/2021.

[Wa2018] Wachter, S.: Normative Challenges of Identification in the Internet of Things: Privacy, Profiling, Discrimination, and the GDPR, in: Computer Law & Security Review, 34 (3), 436-449, 2018, DOI: 10.2139/ssrn.3083554.

[Xu2020] Xu, C.: The Smart Speaker Frenzy and Why It's Happening, https://www.comscore.com/ger/Insights/Blog/The-Smart-Speaker-Frenzy-and-Why-Its-Happening, accessed 26/01/2021.