On the Use of Artificial Intelligence to Deal with Privacy in IoT Systems: A Systematic Literature Review

Giammaria Giordano, Fabio Palomba, Filomena Ferrucci

SeSa Lab - Department of Computer Science, University of Salerno, Italy
giagiordano@unisa.it, fpalomba@unisa.it, fferrucci@unisa.it

Abstract

The Internet of Things (IoT) refers to a network of Internet-enabled devices that can make different operations, like sensing, communicating, and reacting to changes arising in the surrounding environment. Nowadays, the number of IoT devices is already higher than the world population. These devices operate by exchanging data between them, sometimes through an intermediate cloud infrastructure, and may be used to enable a wide variety of novel services that can potentially improve the quality of life of billions of people. Nonetheless, all that glitters is not gold: the increasing adoption of IoT comes with several privacy concerns due to the lack or loss of control over the sensitive data exchanged by these devices. This represents a key challenge for software engineering researchers attempting to address those privacy concerns by proposing (semi-)automated solutions to identify sources of privacy leaks. In this respect, a notable trend is represented by the adoption of smart solutions, that is, the definition of techniques based on artificial intelligence (AI) algorithms. This paper proposes a systematic literature review of the research in smart detection of privacy concerns in IoT devices. Following well-established guidelines, we identify 152 primary studies that we analyze under three main perspectives: (1) What are the privacy concerns addressed with AI-enabled techniques; (2) What are the algorithms employed and how they have been configured/validated; and (3) Which are the domains targeted by these techniques. The key results of the study identified six main tasks targeted through the use of artificial intelligence, like *Malware Detection* or *Network Analysis*. *Support Vector Machine* is the technique most frequently used in literature, however in many cases researchers do not explicitly indicate the domain where to use artificial intelligence algorithms. We conclude the paper by distilling several lessons learned and implications for software engineering researchers.

Keywords: Data privacy, Artificial Intelligence, Internet-of-Things, Software Engineering for IoT.

1. Introduction

We live in a world that is more and more virtualized and where people can do anything, anytime, from anywhere [35].This is enabled by the availability of devices and sensors that can capture the surrounding environment and/or the user requests in order to distribute them toward other devices and sensors and produce data, knowledge, actions, communications, entertainment, and others: this is what we call *Internet-of-Things* (a.k.a. IoT) [5]. It is not easy to give an all-encompassing definition of IoT because the field of applicability of these devices is so vast to risk not including some possible domain or sub-domain. However, Strous *et al.* [34] tried to give a general definition that can be summarized as *"IoT is the inter-networking of physical devices such as vehicles, home appliances, medical devices and so on that can collect and exchange data and interact with other devices using the Internet to monitor or control something."* The key objective of IoT is indeed that of providing people with an infrastructure that allows ubiquitous access to devices and service providers [20]. The last decades have seen an ever-growing interest in IoT, and the vast majority of services and communications are currently offered through IoT devices like smartphones and other smart objects [11, 15]. Recent statistics report that, in 2020, the number of IoT devices connected to the Internet is about 8.74 billion, and this number will increase by 25 times in 2030.[1] However, the growth of IoT is not exempt from serious security threats and privacy [42]. In particular, IoT devices typically have low memory and produce large amounts of sensitive data that are sent and elaborated by servers, which then return the outcome of the elaboration to those devices [12]. Such an intensive exchange of data naturally allows external attackers to steal information and use them for malicious reasons [4, 29], as we witness too often in the news. Some of the most recent, resounding examples are connected to the malicious use of smart assistants[2] or even the influence that IoT data leaks might have had on the 2016 US elections.[3] The problem of privacy is so spread in practice that Meneghello *et al.* [25] provocatively defined IoT as the

[1] Source STATISTA.COM: https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide/.

[2] The ALEXA case: https://www.theguardian.com/technology/2019/oct/09/alexa-are-you-invading-my-privacy-the-dark-side-of-our-voice-assistants

[3] The CAMBRIDGE ANALYTICA case: https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election.

Internet-of-Threats, synthesizing the current body of knowledge on security weaknesses of commercial IoT solutions and highlighting the urgent need for automated mechanisms that may support the detection of privacy concerns in IoT systems. Researchers have actively embraced this call through the definition of techniques based on blockchain [17, 18], gateway instrumentation [24], privacy-preserving data aggregation schemas [19, 22], to name a few.

Besides the techniques discussed above, a recent trend is represented by the adoption of artificial intelligence (AI) algorithms and models. These approaches concern the design of supervised and unsupervised methods, meta-heuristics, or reasoning approaches to detect potential privacy leaks or to preserve privacy in IoT systems [27]. For example, Majumder and Izaguirre [85] developed an AI-based security system that, employing motion detection and facial recognition, might prevent the malicious intrusion of externals into IoT systems. Similarly, Liu *et al.* [77] developed a fully encrypted *Convolutional Neural Network* (CNN) [2] to monitor the vital signs of patients: the encryption mechanism allowed to hide personal data during the training phase of the artificial intelligence model, preserving privacy.

Most of this research has been conducted by researchers in the fields of algorithms, cybersecurity [9], and networks. We advocate that it is time for software engineering to come into play by conducting empirical investigations into the matter and proposing novel instruments to support developers of IoT systems. In this respect, we notice a lack of comprehensive knowledge on what are the privacy issues tackled by artificial intelligence approaches, which design and validation choices were applied when assessing those techniques, which are the current limitations induced by these choices, and in which domains the application of AI-based methods seem more promising. An improved understanding of these aspects is crucial for (1) assessing the current capabilities of these methods; (2) pointing out potential limitations of the techniques employed so far; and (3) identifying additional domains and/or methods that may be used to detect possible privacy issues and preserve data privacy in IoT systems. All these angles might be of interest to our research community to identify the areas where to focus our collective effort.

Hence, in this paper, we conduct a Systematic Literature Review (SLR) on the usage of artificial intelligence techniques for detection privacy issues or to preserve privacy in IoT systems. We employ well-established guidelines [14, 38] to systematically search the literature on the matter: from an initial set of 2,202 papers, we identify 152 primary studies that we then analyze to address the three perspectives of interest.

The analysis of the research literature highlights an increasing interest in artificial intelligence methods for the privacy of IoT systems, and, indeed, we find that a large portion of the papers was published in 2020. In addition, we identify two key use cases: artificial intelligence is used to spot privacy issues or prevent their emergence, yet there exist several subfields where AI-based techniques might be applied. While the most widely used approach is *Support Vector Machine*, we discover that only a few papers elaborated on the rationale behind the selection of the AI technique and, perhaps more importantly, that most of the approaches have been assessed through a limited and potentially biased evaluation metrics. Lastly, we find that the vast majority of the published papers do not include explicit indications on the domains where the proposed techniques can be applied, hence threatening their actionability and reproducibility.

Based on our findings, we identify several future research directions and implications for the research community that encompass the adoption, definition, configuration, and validation of artificial intelligence methods for IoT privacy.

Structure of the paper. Section 2 provides background on IoT devices, other than elaborating on the existing literature reviews on IoT privacy and how our work differs from them. In Section 3 we report the research questions and the systematic methodology employed to search and synthesize the literature. Section 4 discusses the results achieved, while Section 5 further analyzes the implications of our findings. The potential threats to validity and our mitigation strategies are discussed in Section 6. Finally, Section 7 concludes the paper and outlines our future research agenda.

2. Background and Related Work

This section reports on the basic information required to understand our work, namely on the logical architecture of an IoT device and the literature reviews previously proposed in the context of privacy of IoT systems.

2.1. Anatomy of an IoT device

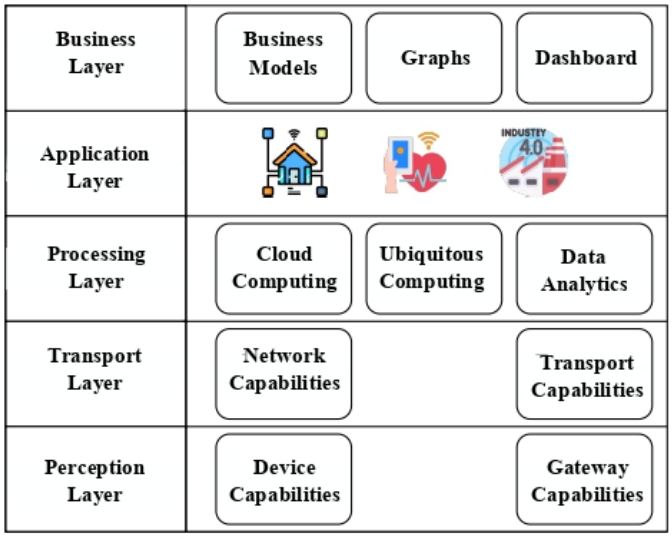

Figure 1: Five Layer Architecture.

Figure 1 shows the current theoretical architecture of a generic IoT device. The architecture is componsed by: *perception layer* that include objects and sensors that compose the device, like GPS sensors, bar-code scanners, and RFID sensors. The *"Transport"* layer that receives pre-elaborated

information and analyzes it through two sub-layers: The *Network Capabilities* and the *Transport Capabilities*.The former allows the device to connect to the network and proceed with the authentication and the access control mechanism; the latter implements the mechanisms needed to transfer the data to the upper levels (e.g., through the definition of wired or wireless protocols like Wi-Fi, RFID sensors, and Bluetooth. The *"Processing"* layer offers a support mechanism to store the data received by the lower layer; this layer usually uses cloud infrastructures, ubiquitous computing, and, finally, it is used to perform data analysis tasks and generate actions that could influence the environment. From the user's perspective, the *"Application"* layer offers an interface between the IoT devices and applications that could be built. It is used to develop and deploy IoT applications like a specific application to govern a smart home or monitor a patient with a healthcare app. Finally, the *"Business"* layer is used to manage and control applications using flow charts, graphs, and dashboards. This is the layer employed for the decision-making process, where one can decide which actions or operations should be done with the information received from the previous layers. This layer is directly involved in protecting the end user's privacy.

In any case, it is worth remarking that, in a real-world scenario, the architectures described are often subject to changes or customization to meet specific requirements of the application to build or because of the heterogeneity of the IoT devices. As such, the architectures should be considered as a starting point for building IoT applications.

2.2. State of Art

In the recent past, some literature reviews targeted the privacy of IoT systems. Most of them treated the problem by investigating the major privacy threats in IoT environments, focusing on the root causes of privacy concerns rather than how artificial intelligence methods have been exploited to identify privacy issues or preserving privacy.

Aleisa and Renaud [3] surveyed the literature from 2009 and 2016 to investigate (1) the geographic distribution of privacy issues, finding this typically concerns Europe and North America; (2) the data collection methods, which were found to be diverse and scattered, other than mainly focusing on quantitative perspective; (3) the hardware technologies, that in about 35% of the cases refer to RFID sensors; (4) the major issue about privacy, namely the lack of privacy-preserving mechanisms; and (5) the topics treated, with authentication and authorization mechanisms being the most popular ones. In doing so, the authors did not only collect published research papers but also news stories and privacy reports to analyze a larger variety of privacy violation perspectives.

Ziegeldorf *et al.* [42] elaborated on a list of IoT environments' privacy threats, reporting the following ones as the most harmful:

1. **Identification**: This relates to the possibility of identifying a device through its IP address or machine name;
2. **Localization and tracking**: This threat refers to the possibility of detecting user traffic in multiple ways, e.g., using a GPS sensor or smartphone localization;
3. **Profiling**: The profiling threat implies the possibility of tracking the user information in order to identify possible relevant information from the target;
4. **Interaction and presentation**: This aspect refers to Machine-Machine Interaction. Indeed, a threat to privacy could arise when these devices share information with other devices;
5. **Lifecycle transitions**: This threat occurs when the devices assume that the information previously shared with other devices has been deleted. However, the devices that receive that information could be storing those data for unclear reasons;
6. **Inventory attacks**: This aspect refers to the possible unauthorized access inside the device. Indeed, the malicious user could detect possible sensible data and use it for multiple illegal actions;
7. **Linkage**: This threat refers to privacy issues arising when multiple devices are connected and share information; in these cases, the devices could be used for unauthorized access inside the system.

The work by Ziegeldorf *et al.* [42] is not meant to be a systematic investigation but rather a viewpoint on the key concerns threatening the privacy of IoT systems.

Two literature reviews have been recently published by Hussain *et al.* [56] and Waheed *et al.* [37]. Similarly to us, both of them investigated the role of artificial intelligence in the context of IoT privacy.

Hussain *et al.* [56] conducted a non-systematic literature review to delineate the current solutions and the future challenges of the use of machine learning in IoT environments, with a particular focus on privacy issues. From a technical standpoint, the authors conducted a meta-analysis of the previous surveys on software security and IoT systems in order to investigate two aspects. First, they synthesized the motivations for using machine learning techniques in the context of IoT. Secondly, they summarized which are the machine learning algorithms employed. In this respect, their focus was mainly on the analysis of the *efficiency* and *complexity* of the machine learning solutions proposed so far. Hence, with respect to the work by Hussain *et al.* [56], ours can be seen as a systematic complementary analysis where we focus on the *design* of the machine learning pipelines, namely the strategies employed to train, build, and validate the models. Furthermore, our systematic literature review (1) does not limit itself to machine learning but explores the broader application of artificial intelligence methods and (2) considers additional dimensions such as the domains, the privacy issues, and the techniques employed for dealing with privacy.

Waheed *et al.* [37] conducted a systematic literature review of the research papers published from 2008 to 2019 that

focused on understanding the role of machine learning and blockchain to deal with security and privacy in IoT systems. Waheed *et al.* focused on threats and countermeasures for security and privacy concerns, reporting the lack of survey efforts in the context of machine learning and privacy.

Based on the analysis of the state of the art, we can claim that our work presents the largest up-to-date systematic investigation into the role of artificial intelligence to deal with privacy in IoT systems.

3. Research Methodology

The *goal* of the study is to survey the research literature that applied artificial intelligence methods for detecting privacy concerns and preserving privacy in IoT systems, with the *purpose* of providing software engineering researchers with actionable items and insights that they can exploit to investigate the matter further and improve the automated support made available to developers and managers to deal with privacy concerns. The *perspective* is that of researchers who are interested in assessing the currently existing methodologies and how to improve them.

To address our goal, we developed and conducted a Systematic Literature Review (SLR), which is a synthesis process through which the existing research papers on a subject of interest are systematically identified, selected, and critically appraised to address one or more research questions [14]. In the context of our work, we followed the well-established guidelines originally proposed by Kitchenham and Charters [14]. To provide additional rigor to the analysis, we also integrated the standard procedure with the so-called *snowballing* procedure [39], i.e., a methodology used to scan the incoming and outcoming references of the primary studies identified by the systematic search for identifying additional sources. We followed the snowballing guidelines provided by Wohlin [38]. In terms of reporting, we followed the *ACM/SIGSOFT Empirical Standards*.[4] and, in particular, the *"General Standard"* and *"Systematic Reviews"* guidelines.

3.1. Research Objectives and Questions

The specific research objectives of the systematic literature review are reported in the following:

Objective 1. Understanding the IoT privacy tasks targeted with artificial intelligence techniques;

Objective 2. Understanding the IoT domains where artificial intelligence techniques have been employed.

Objective 3. Understanding the design, configuration, and evaluation of the artificial intelligence techniques used to deal with privacy in IoT systems.

[4] Available at: shorturl.at/cBDH6.

While the literature on privacy of IoT systems [36] has established a number of static and dynamic instruments that help developers detecting the presence of privacy threats, our objectives are motivated by recent research efforts in the field of privacy and security showing an increasing trend in the adoption of artificial intelligence methods to deal privacy in IoT systems. [16, 26]. For instance, Kuzlu *et al.* [16] advocated the exponential growth in the development of complex artificial intelligence-enabled algorithms to protect IoT systems. This was confirmed by a number of additional studies in the field of privacy and security (e.g., [10, 30, 40]). These observations posed the foundations of our research objectives. We argue the need for a comprehensive understanding of how artificial intelligence methods have been engineered for securing the privacy of IoT systems. This is crucial to assess the software engineering angle of the matter, possibly letting emerge problems and challenges that our research community might help addressing.

Our objectives have driven the definition of our research questions (**RQ**s):

RQ$_1$. *What are the IoT privacy tasks that can be tackled with the use of artificial intelligence techniques?*

RQ$_1$ aimed at addressing the first objective and investigating the most common privacy tasks addressed through the adoption of any form of artificial intelligence method. The research question is motivated by our willingness to provide a comprehensive overview of the state of art regarding privacy tasks treatable with AI-based methods: this may reveal tasks that have been only partially considered by the state of the art, hence suggesting potential future work in the field.

RQ$_2$. *Which are the IoT domains where artificial intelligence techniques have been applied to deal with privacy?*

This research question addressed the second overall objective of the study and was motivated by the willingness to assess the typical domains where artificial intelligence techniques have been applied to the problem of privacy, which may naturally highlight additional domains where the application of these techniques might be worthy.

RQ$_3$. *Which families of artificial intelligence algorithms were used to deal with privacy in IoT systems?*

RQ$_4$. *Which datasets were used to validate the artificial intelligence methods employed to deal with privacy in IoT systems?*

RQ$_5$. *Which strategies were used to validate the artificial intelligence methods employed to deal with privacy in IoT systems?*

RQ$_6$. *What are the evaluation metrics employed to assess the quality of the artificial intelligence methods employed to deal with privacy in IoT systems?*

With the set of research questions from **RQ$_3$** to **RQ$_6$**, we aimed at addressing the last objective and investigating the inner-working of the artificial intelligence techniques employed in the literature. It is important to note that such an analysis was motivated by a key consideration in the field of artificial intelligence and machine learning: the design, configuration, and validation of those techniques might heavily influence the interpretation of the performance [23, 28, 32]. Hence, our research questions shed light on how researchers have defined these techniques, possibly revealing common patterns and limitations to address. In addition, these research angles allowed us to complement previous work on the use of artificial intelligence methods for IoT privacy [37, 56], by providing a deeper understanding of the methodology used to define artificial intelligence pipelines to be used when detecting privacy issues or preserving privacy.

3.2. Research Query definition

One of the key methodological steps of a systematic literature review is identifying appropriate search terms that may help retrieve a comprehensive set of sources. In this respect, we adopted the following strategy:

- For each research question, we first extracted the most relevant keywords—these represented the base to conduct our search;
- For all relevant terms, we identified possible synonymous or alternative spelling;
- We used boolean operators (*AND* and *OR*) to compose the research query.

The outcome was the following research query:

Search Query

("privacy" *OR* "anonymization" *OR* "sensitive information" *OR* "sensitive words") *AND* ("iot" *OR* "Internet-of-Things") *AND* ("machine learning" *OR* "artificial intelligence")

As shown, we put in *OR* all the synonyms of the same concept, while multiple concepts were combined using the *AND* operator. It is important to remark that, before continuing our systematic literature review, we checked the presence of the search query terms related to privacy concerns on the systematic literature reviews described in Section 2. The basic idea behind this step was to verify the consistency and completeness of the selected terms against papers that reported a systematic investigation of the literature. Therefore, they are supposed to contain a complete mapping of the terms used in literature to indicate privacy concerns. This step allowed us to include alternative words in the search query in case these were not included initially. This did not eventually happen since we did not identify any additional terms to include.

As for the artificial intelligence-related terms, the existing literature reviews targeted just part of the problem, i.e., the use of machine learning. Therefore, it was impossible to check the selected terms' completeness against them. In any case, we preferred to include both *"machine learning"* and *"artificial intelligence"* to (1) not miss any of the resources considered by previous systematic literature review and (2) identify sources that did not employ machine learning but other forms of artificial intelligence.

3.3. Search Databases

Once we defined the search query, we selected the databases to use when performing our search. Correct identification of those databases is fundamental to have a successful literature review [14]. For this reason, we selected the top-three research databases,[5] namely:

- IEEEXplore (http://ieeexplore.ieee.org);
- Scopus (www.scopus.com);
- ACM Digital Library (https://dl.acm.org).

These databases are typically used to conduct systematic literature reviews [8, 13] and, perhaps more importantly, guarantee complete coverage of the published research, hence allowing us to access the entire set of papers.

3.4. Exclusion and Inclusion criteria

Exclusion and inclusion criteria allow the selection of resources that address the research questions of a systematic literature review [13]. In the context of our study, we identified and applied the following *"Inclusion/Exclusion"* criteria.

A) **Exclusion criteria**: The resources that met the following constraints were filtered out from our study:

- Papers not written in English;
- Short papers, namely papers with a number of pages lower than seven;
- Workshop papers;
- Duplicated papers;
- Papers whose full text read was not available;
- Conference papers later extended to journal;
- Master Theses.

Using these filters, we could exclude all preliminary research results, e.g., workshop or short papers, but also avoid considering a similar paper multiple times, e.g., in case of an archived journal paper that extends a conference publication or in case of duplicates.

[5] source: https://paperpile.com/g/research-databases-computer-science/.

B) **Inclusion criteria**: Papers that applied artificial intelligence methods to the problem of privacy of IoT systems were *included* in our study.

3.5. Snowballing

The snowballing technique refers to the use of the reference list of an paper or its citations to identify additional papers that might have been missed by the search process [39]. This is typically used *after* the application of the exclusion/inclusion criteria, so that the reference analysis is only performed on the relevant papers that address the research questions of the literature review. As the reader might see, the snowballing technique requires an extensive amount of time and effort: for this reason, we limited ourselves to the application of the so-called *backward* snowballing, that is, the scanning of the reference list of the papers selected.

3.6. Quality assessment

Before proceeding with the extraction of the data required to address our research questions, we assessed the quality and thoroughness of the retrieved resources to discard the papers that did not provide enough details to be used in our study. Particularly, we defined a checklist that included the following questions:

Q1. *Are the artificial intelligence techniques clearly defined?*

Q2. *Are the privacy topics treated in the paper clearly defined?*

Each question could be answered as *"Yes"*, *"Partially"*, *"No"*. We associated a numeric value for each label to better assess the quality and thoroughness of each source: the label *"Yes"* was associated to the value *'1'*, *"Partially"* to *'0.5'*, *"No"* to *'0'*. The overall quality score was computed by summing up the score of the answers to the two questions, and the articles with a quality score of at least 1 were accepted.

3.7. Data extraction

Once we had identified the final set of sources to consider, we extracted the information relevant to address our research questions. In particular, we defined the data extraction form reported in Table 1. Besides the basic information on the privacy topics treated by the paper or on the design/validation of the artificial intelligence techniques, we also sought to extract data on the dataset exploited and the programming language used to develop the technique: these pieces of information could provide additional insights into the characteristics of the considered papers. Also, the data extraction form included a *"Limitation(s)"* field, through which we took note of the possible limitations of the techniques assessed. It is important to remark that the *"Validation Techniques"* field of the data extraction form was voluntarily left optional, as not all the primary studies might have proposed validations of the artificial intelligence techniques proposed.

Dimension	Attribute: Description
Privacy	What kind of privacy concerns have been highlighted in the use of IoT devices?
Machine Learning Algorithms	What kind of algorithms were used to tackle the problem?
Topics of Interest	What are the main topics treated?
Programming Language	What programming languages have been used to address this issue?
Training Strategy	What is the strategy adopted to train the model?
Validation Techniques	What kind of techniques were used to validate the model (if any)?
Dataset	What dataset has been selected to train the Machine Learning model?
Evaluation Metrics	What evaluation metrics has been used to evaluate the model? (e.g., F-score, Accuracy, Precision, Recall).
Limitation	What are the limitations of current techniques?

Table 1: Data Extraction Form

3.8. Search Process Execution

Once we had defined the basic blocks of our systematic literature review, we then proceeded with its execution. An overview of the execution is presented in Figure 2, where we show how the number of primary studies considered varied when applying the various filters we defined. In particular, the execution process worked as follows:

A We run the search query against the three selected databases. In this respect, it is worth remarking that each database requires its parameters to conduct the search process, e.g., in terms of the document types to consider. For the sake of replicability, Table 2 summarizes the parameters defined for each database. The search query output a total amount of 2,202 hits: 1,273 for the ACM Digital Library, 486 for IEEEXplore, and 443 for Scopus. The higher number of hits obtained when querying the ACM Digital Library is motivated by the internal mechanisms that the platform employs to match a query against the content it makes available [7]: in particular, it does not limit the search of each term of a query to the full content of an paper, but also considers the metadata, hence providing a larger amount of candidate relevant papers. In any case, we completed the first step by downloading all the candidate papers and storing them in a local environment for a quicker investigation.

B Each of the candidate papers entered the next phase, which consisted of applying the exclusion criteria. The process was conducted by the paper's first author, who scanned each source and applied the filters. The author first considered each paper's title, abstract, and

Database	Year	Document Type	Publication Stage	Language	Media Format	Subject Area	Source Type	Results
IEEE	2011 - 2021	Conference Journal						486
Scopus	2011 - 2021	Conferences paper	Final	EN		Computer Science Engineering	Conference Proceeding Journal	443
ACM	2011 - 2021	Research Paper			PDF			1,273
Total								2,202

Table 2: Filters applied in the research queries.

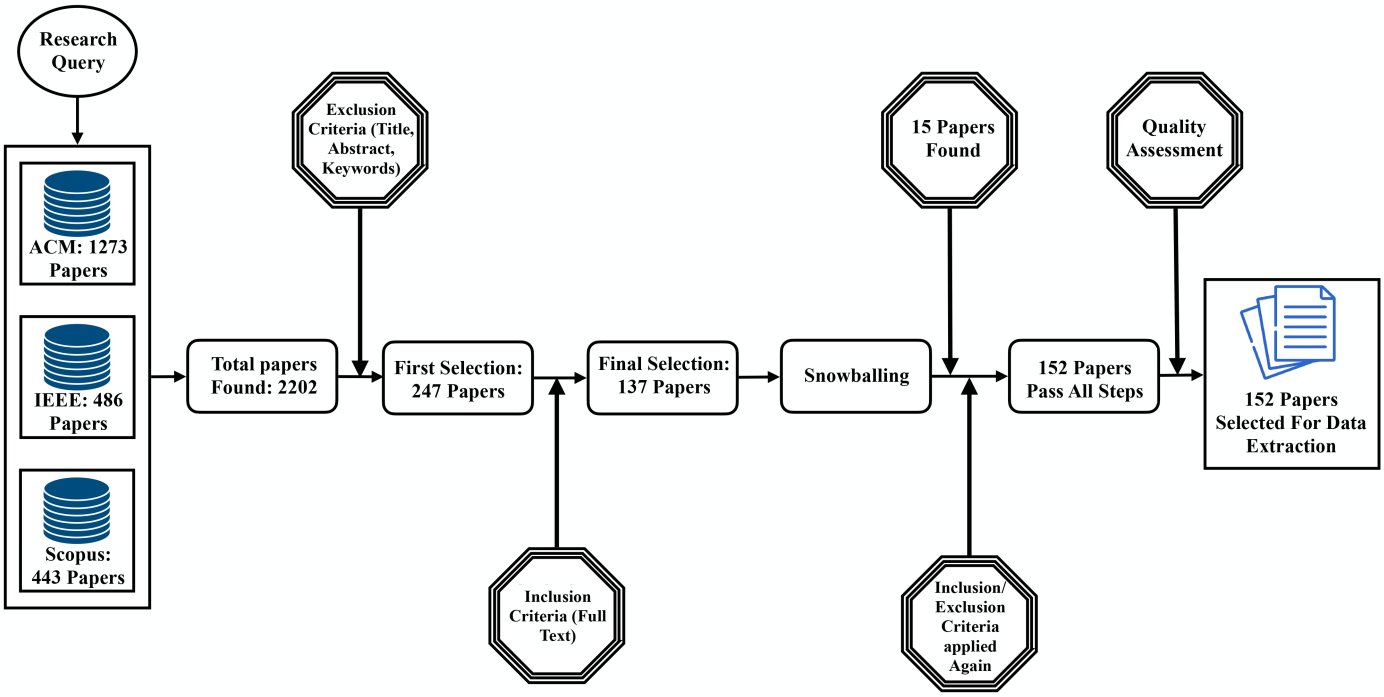

Figure 2: Overview of the papers selection process.

keywords to decide on whether it should have been discarded. If this was enough, the inspector read the content of the paper. Overall, 1,955 papers were excluded, and, therefore, 247 passed to the next step.

C The inclusion criteria were considered. Also, in this case, the paper's first author acted as the inspector and started applying the criteria defined against the 247 papers. Unlike the previous step, the inclusion was assessed by considering the full paper and not only the title, abstract, and keywords. In case of indecision, the inspector brought the case to the attention of the other authors, who could provide feedback and open a discussion that led to a final agreement. As a result of this procedure, we discarded additional 110 sources, leading to a final number of 137 papers included in our systematic review.

D After collecting the relevant papers, the inspector applied the backward snowballing procedure and identified potentially relevant candidates missed by the original search. Then, the inspector let the additional sources pass through the exclusion/inclusion criteria. Similar to what was previously done, the inspector requested the feedback of the other authors whenever needed. The snowballing procedure included 15 new sources, leading to a total of 152 papers.

E The next step is concerned with the application of the quality assessment. This was a particularly critical phase, since we had to rate the papers based on their clarity or the availability of enough information to address our research questions. The process was initially started by the first author of the paper, who provided a first score to each paper. Then, the second author independently repeated the analysis to have a second look and a more stable evaluation of the papers. The scores assigned by the two authors were later compared, and cases of disagreement were discussed. As a result, no paper was excluded, and, therefore, all 152

papers passed the quality assessment.

F Finally, we proceeded with the data extraction. Most of the information required to address our research questions (e.g., the artificial intelligence technique employed) was rather easy to collect and, therefore, the first author independently extracted them. More problematic was instead the analysis of the potential limitations. This required a more careful and focused discussion, which was jointly conducted by the two first authors of the paper. In particular, they analyzed (1) the sections of the papers where potential limitations and threats to validity were discussed; (2) the characteristics and properties of each technique employed, trying to identify additional limitations. In this respect, it is worth remarking that the inspectors have years of expertise in artificial intelligence and are also involved in academic courses on the matter.

The data extracted from the selected papers were then used to provide an answer to our research questions. The following section overviews the main findings of our analysis.

4. Analysis of the Results

Before diving into the results addressing our **RQ**s, it is worth reporting some meta-information on the primary studies accepted for our systematic literature review.

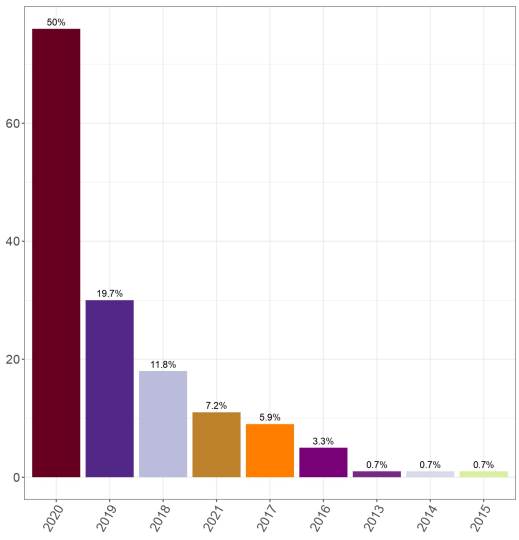

Figure 3: Publication trend by year.

In the first place, Figure 3 depicts bar plots highlighting the number of papers published by year. Looking at the figure, two elements might be noticed. On the one hand, the publication trend recalls an exponential function, which means that the interest in facing privacy in IoT systems using artificial intelligence techniques is rapidly and massively increasing. This may indicate that several other papers will be published in the near future. The publication trend further motivates our work, namely the need for a systematic literature review that analyzes how artificial intelligence techniques have been applied and validated in the field, other than which are the key limitations that future research is called to address. On the other hand, we can also notice that 50% of the papers (75) have been published in 2020 [A1–A4, A7, A10, A15–A17, A20–A23, A25–A27, A29, A31–A33, A36, A38, A40, A42, A44, A45, A49, A54–A56, A61, A62, A65, A66, A68, A71, A75, A77, A78, A80, A83–A85, A87, A88, A90, A91, A93, A95, A96, A98, A104–A106, A109–A112, A117, A120, A125–A127, A129, A130, A135, A136, A138, A140, A142–A144, A147, A148, A150]. While the astonishing number of published material can be connected to the general exponential publication trend, it also indicates how privacy is becoming more and more pressing for researchers. A possible influencing factor is the significant increase in terms of IoT devices acquired by users during the pandemic years [33], which has naturally further increased the need for privacy-preventing mechanisms.

An additional preliminary view on the characteristics of the primary studies is concerned with the programming languages employed to devise the artificial intelligence techniques. We noticed that not all the articles explicitly mentioned the programming languages used. In some cases, this information has been obtained by analyzing side information, i.e., references of third-party libraries, code snippets commented in the articles, or manual analysis of replication packages. Unfortunately, in 91 cases we could not find any information to elicit the programming language adopted.

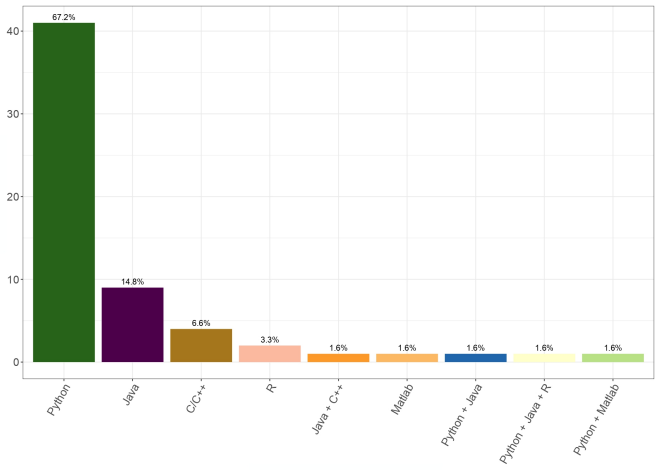

Figure 4: Programming languages used to devise the artificial intelligence techniques proposed in the primary studies.

Figure 4 provides the results of this analysis. We identified PYTHON as the key means enabling the definition of artificial intelligence techniques: this was indeed used in 67% of the articles [A2, A4, A7, A14, A16, A19, A26, A40, A49, A50, A54, A59, A60, A62, A64, A66–A68, A71, A72, A74, A76, A83, A89, A94, A97, A100–A103, A106, A110, A119, A120, A123, A129, A135, A138, A142, A147, A148] This result was somehow ex-

pected, as PYTHON is widely considered as the main programming language for data science and machine learning engineering, as it offers a large amount of data science libraries that make the development of artificial intelligence techniques easier.[6] Other programming languages are less used. In 11% of the cases researchers preferred a combination of multiple programming languages. In these cases, different programming languages were used to implement different steps of the artificial intelligence pipelines: as an example, [A47] employed the R toolkit to perform data cleaning operations and then relied on the *Weka* library[7]—written in JAVA—to devise a machine learning solution.

Hence, such an analysis allows us to recommend the usage of PYTHON for building novel solutions based on artificial intelligence to treat privacy concerns in IoT: this solution would indeed offer an easier chance to build techniques that can extend the existing ones, e.g., by applying specific, tailored mechanisms on top of the techniques proposed in literature, or even compare the performance of the newly proposed techniques with the existing ones.

Summary.

The problem of privacy detection and preservation in IoT using artificial intelligence is now, more than ever, relevant and massively explored by researchers. The publication trend is indeed exponential and about 50% of the primary studies has been released in 2020. PYTHON is the top programming language employed to build the artificial intelligence techniques proposed in literature.

4.1. RQ_1 - *On the privacy tasks tackled with the use of artificial intelligence techniques.*

To address $\mathbf{RQ_1}$, we elicited the privacy task(s) performed in the primary studies. We labeled each paper with the set of tasks considered to enable the analysis. Figure 5 reports the top-6 tasks performed in the primary studies. For the sake of clarity, we focus the following discussion on these tasks since these are the ones considered by at least 10% of the primary studies. Nonetheless, we report in Figure 6 a word cloud that summarizes the whole set of tasks considered.

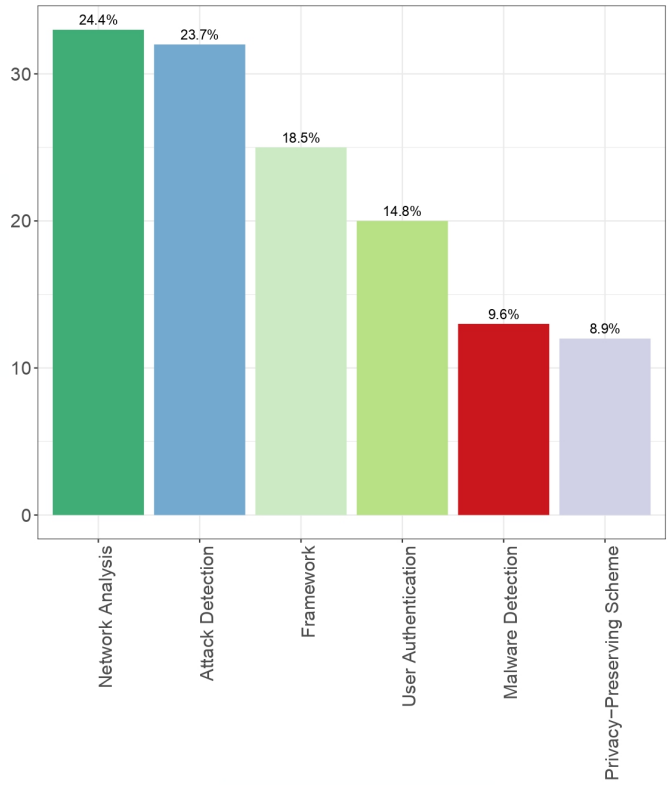

Figure 5: Topics frequency.

"Network Analysis". The most prominent task is the one of *Network Analysis*—24.4% of the primary studies (33) explicitly focused on that. The authors of these primary studies highlighted a critical threat to privacy due to the fact that IoT devices typically share information without secure protocols. For this reason, data might be easily stolen [A64, A114, A115, A139, A148]. More specifically, the task aims at investigating the presence of malicious traffic and/or activities on networks, e.g., the exchange of vulnerable packages. Artificial intelligence models are typically used to classify network traffic and identify sensitive or personal information transmitted by IoT devices [A1, A8, A9, A22, A23, A26, A29, A32, A38, A45, A59, A60, A63, A64, A67, A73, A76, A79, A94, A96, A102, A107, A109, A114, A115, A117, A120, A128, A129, A131, A139, A148, A151]. For instance, Fei *et al.* [A38] collected traffic data from the network environment to feed a *Random Forest* algorithm able to classify an abnormal traffic potentially leading to a denial of service. In a very similar fashion, the other approaches proposed in literature collect information from various sources to train machine learners able to classify malicious inputs to a network.

"Attack Detection'. This task has been taken into account by 23.7% of the primary studies (32). It refers to the possible detection of malicious actions. More particularly, we identified two main use cases: *"Intrusion Detection"* and *"Anomaly Detection"*. The former consists of the definition of a hardware or software component to detect possible attacks on an IoT device. The intrusion detector analyzes the network traffic and pinpoints possible suspicious activities, like phishing and ransomware. To perform this action, the authors typically used artificial intelligence to analyze this traffic, searching for anomalous patterns that can indicate an intrusion on the system [A6, A12,

[6] Top programming languages for data science and machine learning engineering: https://towardsdatascience.com/top-programming-languages-for-data-science-in-2020-3425d756e2a7.

[7] The *Weka* toolkit: https://www.cs.waikato.ac.nz/~ml/weka/.

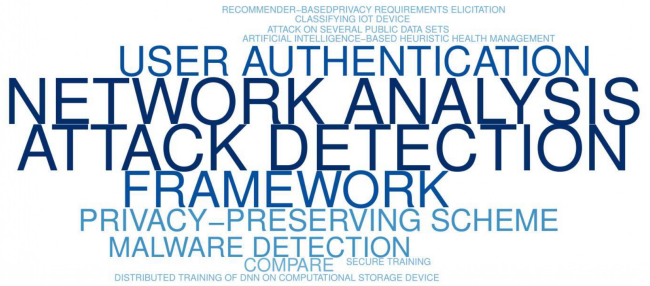

Figure 6: N-Gram topics treated.

A21, A40, A95, A103, A108, A110, A124, A125, A145]. The latter use case, i.e., *"Anomaly Detection"*, may be seen as a sub-category of the *"Intrusion Detection"* one: the purpose, indeed, is exactly the same but with a fundamental difference due to the methodology applied to identify malicious actions. While the *"Intrusion Detection"* analyzes the signatures of known attacks or possible deviations from normal traffic, *"Anomaly Detection"* relies on statistical models to verify the incoming or outgoing traffic [A13, A14, A35, A47, A50, A65, A82, A99, A122, A134]. It is worth noting that, in some cases, the authors of the primary studies did not explicitly indicate the specific use case considered, i.e., they simply refer to *"Attack Detection"* (e.g., [A3, A4, A7, A44, A57, A66, A68, A93, A104, A143, A152].). For this reason, we grouped *"Intrusion Detection"* and *"Anomaly Detection"* under the *"Attack Detection"* task.

"Framework Building". Building a framework to characterize privacy concerns is the focus of 25 studies [A16, A25, A33, A42, A58, A69, A78, A81, A83, A84, A86, A88, A90, A91, A100, A105, A106, A111, A119, A130, A132, A136, A140, A141, A149]. A typical use case is the creation of frameworks that can be then used to experiment new mechanisms to train machine learning models in a distributed environment [A16, A33, A42, A81, A83, A84, A90, A91, A100, A105, A106, A111, A130, A136, A140, A149]. The framework building consists of the design and implementation of usable tools or pipeline that combine multiple artificial intelligence algorithms to detect privacy issues or preserve privacy. As an example, Meurish *et al.* [A90] devised a decentralized and privacy-by-design platform that loads confidential artificial intelligence models into a trusted execution environment, hence protecting users from possible privacy concerns. On a similar note, Wang *et al.* [A130] defined a federated machine learning approach that enables users to deploy complex clustering problems into the cloud.

"User Authentication". 20 primary studies defined new secure authentication mechanisms [A20, A24, A27, A28, A34, A37, A39, A41, A43, A48, A52, A55, A70, A74, A85, A87, A89, A118, A121, A142]. As an example, this category refers to the definition of person authentication tools that exploit biometric sensors: this is especially true in the healthcare field, where biometric sensors are used to monitor patients through the measurement of the blood pressure, heart rate, and others; afterward, the collected parameters are used to generate a unique identifier that can be used to access a system or in a reserved area [A20, A24, A37].

"Malware Detection". This task was the subject of 13 primary studies and refers to the creation of agents that analyze the processes that execute on a host machine to identify possible malware. The most common task consisted of the identification and/or classification [A2, A5, A11, A18, A30, A31, A46, A61, A71, A72, A101, A113, A126] of the various types of malware. More particularly, we recognized two different trends. First, the use of pattern mining to detect malicious applications. For instance, Darabian *et al.* [30] used sequential pattern mining to detect the most frequent opcode sequences of malicious IoT applications; these sequences were then used to distinguish malicious from benign IoT applications. Second, the use of supervised machine learning approaches to classify malware. As an example [5] trained a *Random Forest* algorithm with malware data of ANDROID applications in order to identify malicious mobile apps.

"Privacy-Preserving Scheme". This task was subject of 12 primary studies and refers to the definition of new protocols and schemes to improve privacy. The authors of the primary studies typically include blockchain or similar mechanisms to keep data safe [A51, A60, A62, A77, A80, A92, A98, A112, A137, A138, A144, A146]. An example is represented by the work of Zhao *et al.* [A146], who devised a blockchain-based federated learning approach for IoT devices, where the data collected from multiple sensors are stored within a privacy-preserving blockchain before being consumed by machine learning models.

Other tasks are way less considered, perhaps because they represent emerging topics or because of the lack of datasets that may be used to perform them. We further analyze this in the context of the next research questions.

Summary.

The results of $\mathbf{RQ}_1$ indicate six tasks that are often considered for the application of artificial intelligence techniques: (1) *"Network Analysis"*; (2) *"Attack Detection"*; (3) *"Framework Building"*; (4) *"User Authentication"*; (5) *"Malware Detection"*, and (6) *"Privacy-Preserving Scheme"*. A common approach is that of using artificial intelligence techniques on networks in order to detect possible reserved information exchanged or even the vulnerable IoT devices in a certain environment.

4.2. RQ_2 - *On the IoT domains where artificial intelligence techniques have been applied.*

When addressing $\mathbf{RQ}_2$, we needed to elicit the domain from each of the considered primary studies. In this respect, we labeled each paper with one or more domains: in cases where the domain was not explicitly reported by the authors, we used the label *"Smart Environment"*: this indicates that a certain approach is generic enough to be used in more domains. The results of this analysis are depicted in Figure 7. As reported, the vast majority of the primary studies do not

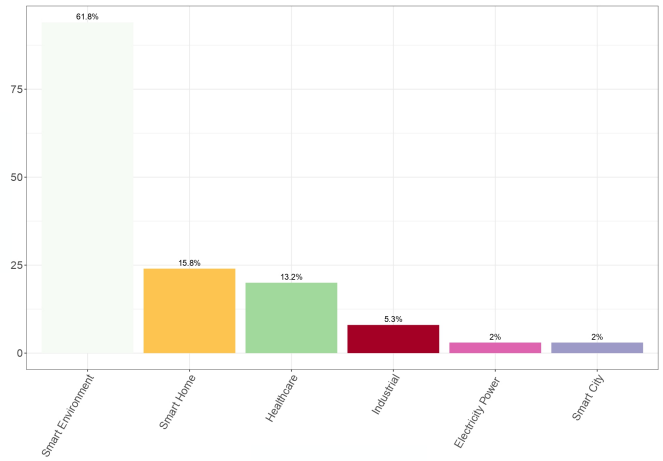

Figure 7: IoT domains where the artificial intelligence methods have been experimented.

explicitly indicate a use case domain for the artificial intelligence approach proposed or experimented. This was the case for 93 papers (61.8%). In most of these studies, the authors limit themselves to generic discussions of IoT environments where their approach might work. This indicates that most techniques are agnostic and can be applied for a variety of purposes. Typically, these have to do with domains like smart factories, military fields, smart home, healthcare, and more [A6, A13, A22, A27, A49, A74, A90, A101].

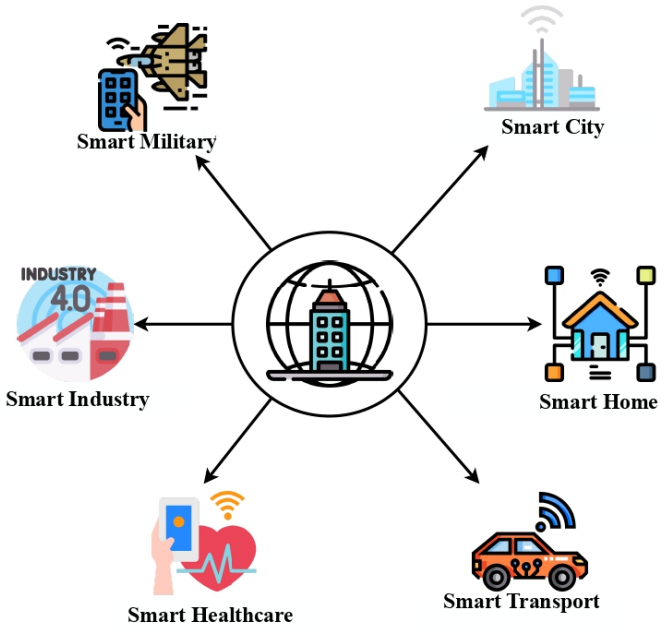

Figure 8: Definition of *"Smart Environment"*, according to the analysis of the papers that do not explicitly report the domains.

Figure 8 reports a taxonomy of the *"Smart Environment"* domain, which was built after analyzing the papers that attempted to devise agnostic techniques. Besides the generic smart environment domain, 24 studies (15,8%) discussed techniques for smart home, while other 20 (13.2%) proposed approaches to manage healthcare-related issues. There are two likely reasons behind this result. On the one hand, the research interest in smart home might be driven by the increase of IoT devices that can be used in such domain. As an example, devices like AMAZON ALEXA or GOOGLE HOME are becoming affordable and popular. As a consequence, the privacy of smart home devices represents a critical challenges to face. On the other hand, the healthcare domain has often caught the attention of researchers, given that IoT techniques can be used to monitor patients and exchange personal data to speed up diagnoses and communications.

Other domains are, instead, less considered so far and represent emerging topics. The application of artificial intelligence to smart industry, for energy considerations or industry [A8, A13, A16, A18, A25, A34, A38, A58, A66, A68, A76, A79, A80, A87, A111, A115, A143, A144, A150] was indeed the object of recent papers published in 2020. This suggests that the research community is trying to approach domains that were not typically targeted.

Artificial Intelligence Technique	Smart Environment	Smart Home	Healthcare	Industrial	Smart Cities	Sum
SVM	21	10	6	1	2	50
Random Forest	17	10	5	2	1	38
K-Nearest Neighbours (k-NN)	10	7	8	2		27
Decision Tree	8	8	5			23
Convolutional Neural Network (CNN)	11	2	4	1		19
Naive Bayes	6	4	5	1	1	18
Multilayer Perceptron (MLP)	9	3	2			14
Logistic Regression	8	2	1			13
Neural Network	6	1	3			13
K-Means	3	2	2	1		7

Table 3: Frequency of artificial intelligence techniques used in literature to deal with privacy concerns in IoT systems.

Summary.

So far, most of the proposed techniques target multiple smart environments and were designed to be generic enough to work in various domains. At the same time, smart home and healthcare are established contexts where privacy concerns are always challenging. Our literature review also identified some emerging domains for artificial intelligence, like, for instance, the application of smart techniques for electricity power reduction.

4.3. $\mathbf{RQ}_3$ - *On the families of artificial intelligence algorithms used to deal with privacy in IoT systems.*

With $\mathbf{RQ}_3$ we analyzed the primary studies in order to identify the artificial intelligence techniques that were used and label them according to their characteristics. Figures 9 and 10 show the results of our analysis, while Table 3 overviews the number of primary studies that adopted each technique, also indicating the domain where these have been experimented.

Shallow machine learning approaches, i.e., approaches that learn from data described by predefined features [41], are more frequently devised. Supervised techniques have been used in various forms: these are connected to the definition of prediction models that can distinguish the characteristics of an unseen instance based on a training base. According to our analysis, a number of algorithms have been proposed, like *Random Forest* [A1–A8, A12, A13, A20, A26, A27, A38, A40, A50, A52, A57, A60, A65, A71, A73, A74, A79, A88, A93–A95, A97, A102, A107, A109, A119, A120, A122, A133, A134], *Support Vector Machines* [A1, A2, A4, A6, A7, A11–A14, A24, A26–A28, A30, A35, A40, A43, A46–A48, A50, A52, A53, A60, A63, A65, A67, A71, A74, A75, A79, A88, A93, A94, A101–A104, A109, A110, A115, A121, A125–A127, A131, A133, A138–A150]. None of the surveyed papers provided motivations leading to the selection of these algorithms. Nonetheless, the higher amount of primary studies proposing supervised learning techniques is likely due to the characteristics of the problems considered: as a matter of fact, most researchers have been working on the definition of classification and/or regression approaches to identify privacy concerns, which calls for the adoption of supervised machine learning techniques. An overview of the privacy tasks considered with each of the machine learning techniques is provided in Figure 11. As shown, typical use cases are the authentication problem and the network traffic analysis. The authors that considered this authentication problem typically applied supervised learning algorithms to classify authorized or unauthorized accesses. Network analysis was instead approached by collecting previous network data and features in order to devise prediction models that could discriminate the likelihood that the current traffic is anomalous and may therefore lead to security threats for an IoT device.

A lower amount of studies focused on unsupervised learning. Specifically, the use of clustering, and the *k-Means* algorithm in particular, allowed researchers to devise approaches able to group together the common properties that may characterize the privacy concerns treated. Clustering algorithms were typically used to cover two macro-areas: data classification and devices aggregation. The former refers to clustering to classify devices or network traffic. The latter refers to the definition of common features or parameters related to IoT devices [A6].

During our analysis, we found that the clustering algorithms were either used as an alternative to supervised learning algorithms [A6] (e.g., to classify or aggregate devices based on some criterion for instance defined commons features or parameters related to IoT devices) or in combination with them [A13, A45, A47, A48, A89, A90, A130, A133, A150]). As an example, the studies performed by Anton et al. [A13] and Hamza et al. [A47] employed clustering to classify abnormal network traffic, hence defining unsupervised approaches that could identify possible anomalies on a network. At the same time, an example of combination was presented in the paper by Hag et al. [A48], who focused on the problem of occupancy detection, i.e., the classification of whether a room is occupied by a person. In this case the authors used time-stamped images of environmental variables like temperature, humidity, light, CO2, to assess the accuracy of a user authentication approach. When gathering the images, the authors applied a k-means clustering algorithm to define a first grouping of normal and malicious room occupancy. These clusters were used to obtain labels that were later exploited to train an *SVM* algorithm.

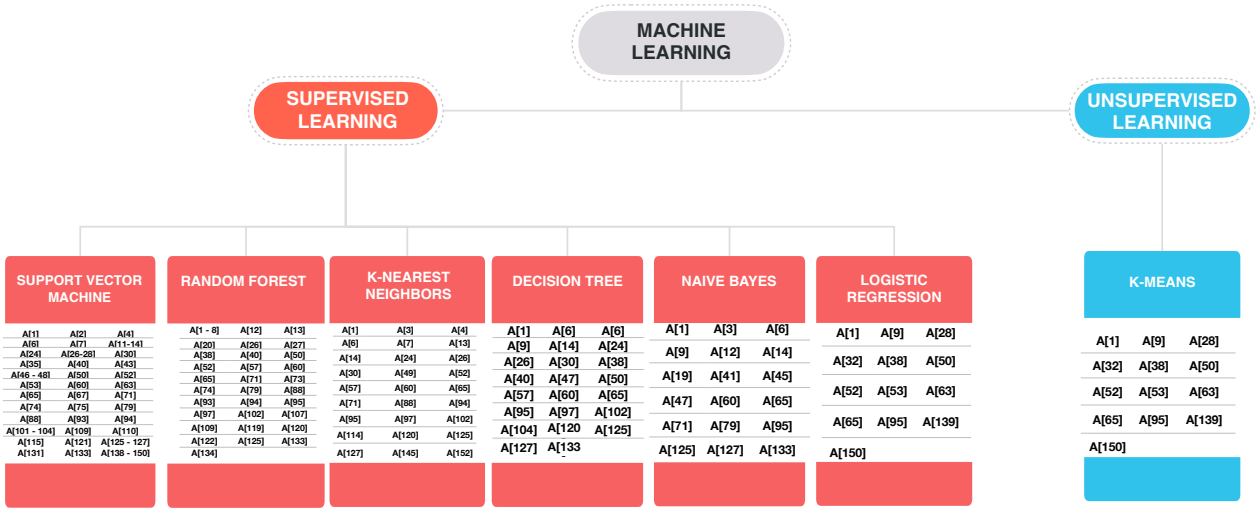

Figure 9: Taxonomy of the machine learning techniques used in literature to deal with privacy concerns.

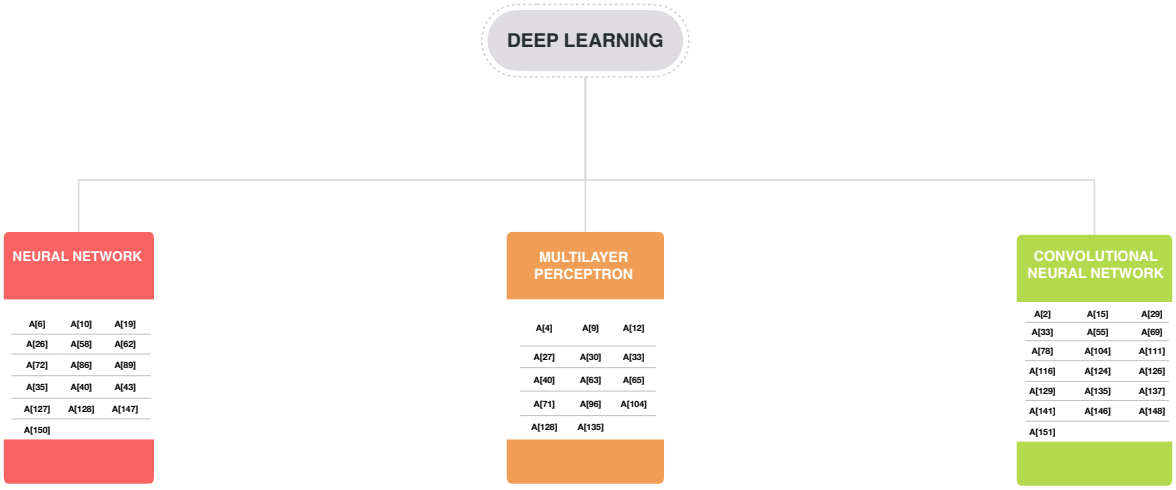

Figure 10: Taxonomy of the deep learning techniques used in literature to deal with privacy concerns.

A more recent trend is the adoption of *deep* learning. We observed that the primary studies that used this type of learning were all published in 2020, indicating a growing interest.

Figure 12 provides a conclusive overview on the artificial intelligence techniques used in literature. In particular, the figure connects the top-6 tasks coming from the results of $\mathbf{RQ}_1$ to the artificial intelligence techniques adopted to solve them. Each task is depicted with a different color; this color characterizes the edges that connect each task to the techniques used in literature. The edges are weighted based on the amount of primary studies using a technique to address a certain task. For instance, the *"Network Analysis"* task is reported in red. The red edges indicate that the task has been addressed in 13 papers with the use of *SVM*, in 11 papers with *Random Forest*, in 7 papers with *KNN*, and so on. From the figure, we can confirm that *Support Vector Machine* has been the most used artificial intelligence algorithm (48 times) to address the majority of the privacy tasks investigated by researchers so far.

Figure 13 overviews the tasks faced by researchers through the use of deep learning. To provide an example of a common task for which deep learning has been used, let consider the OCCLUMENCY framework developed by Lee et al. [69]. This is a cloud-driven solution designed to protect user privacy without reducing the benefits of cloud resources. It is common for IoT applications to collect and share sensitive information through a cloud platform. The OCCLUMENCY framework uses deep learning to encrypt that information without increasing the latency of the cloud platform's response. More in general, we noticed that deep learning had been experimented for tasks previously treated with shallow machine learning techniques also to verify how deep learning approaches can improve the prediction performance of traditional shallow learning algorithms.

As an outcome of our analysis, there are two main observations to make. First, most researches focused on the adop-

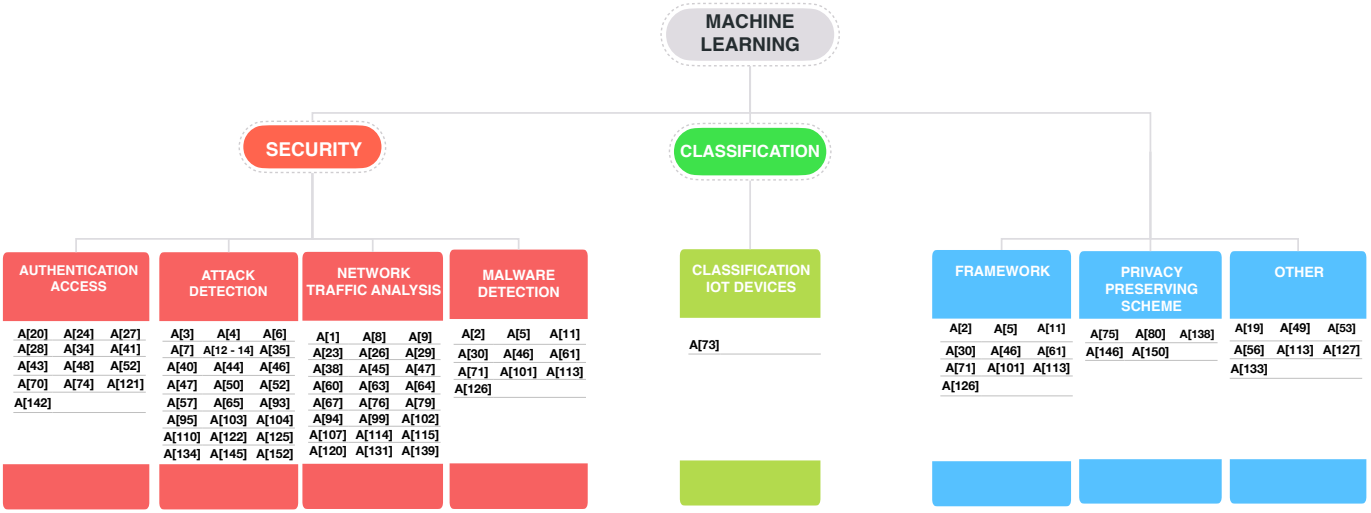

Figure 11: Use cases where the supervised and unsupervised learning techniques have been used.

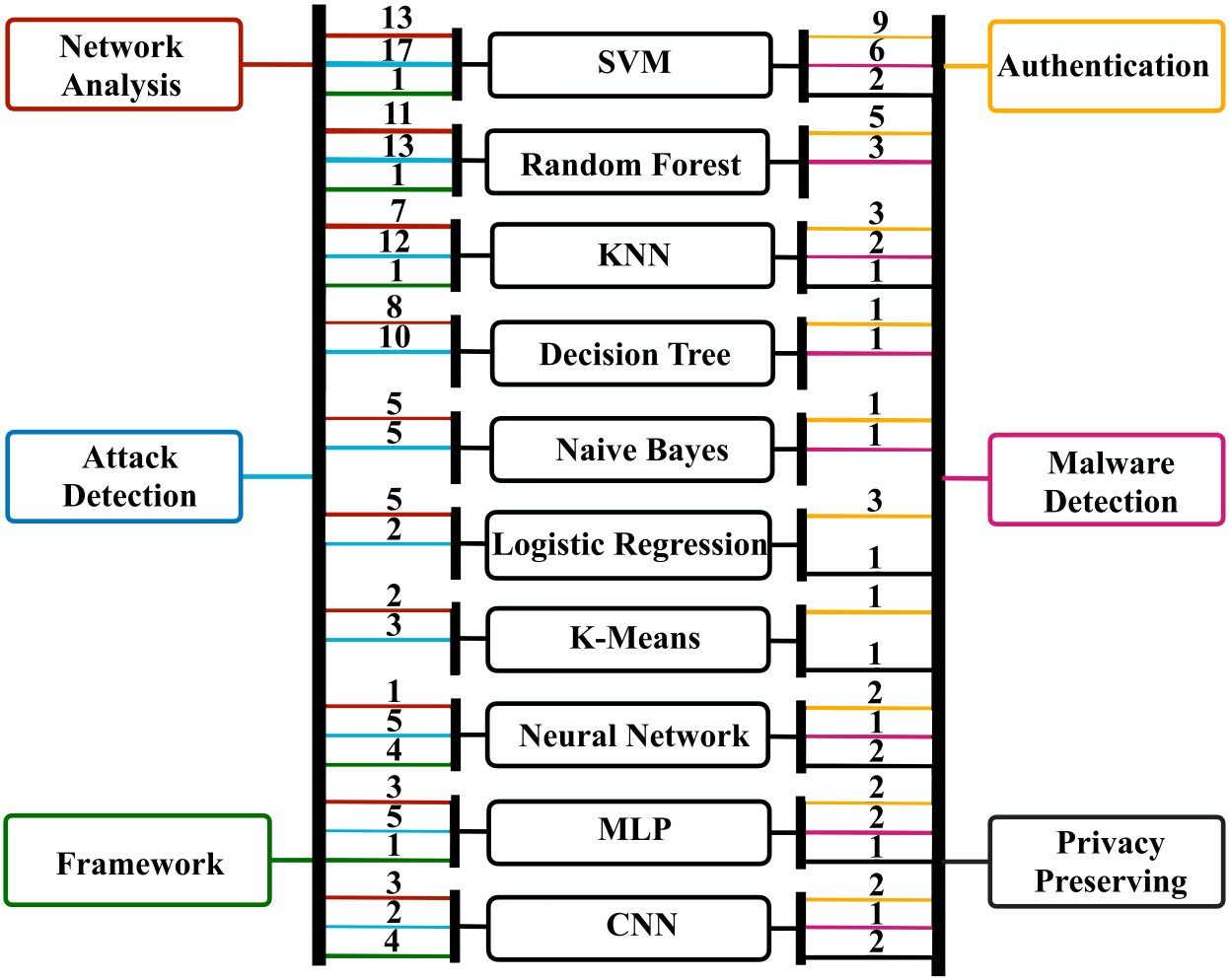

Figure 12: Frequencies of artificial intelligence tasks with the most six tasks considered.

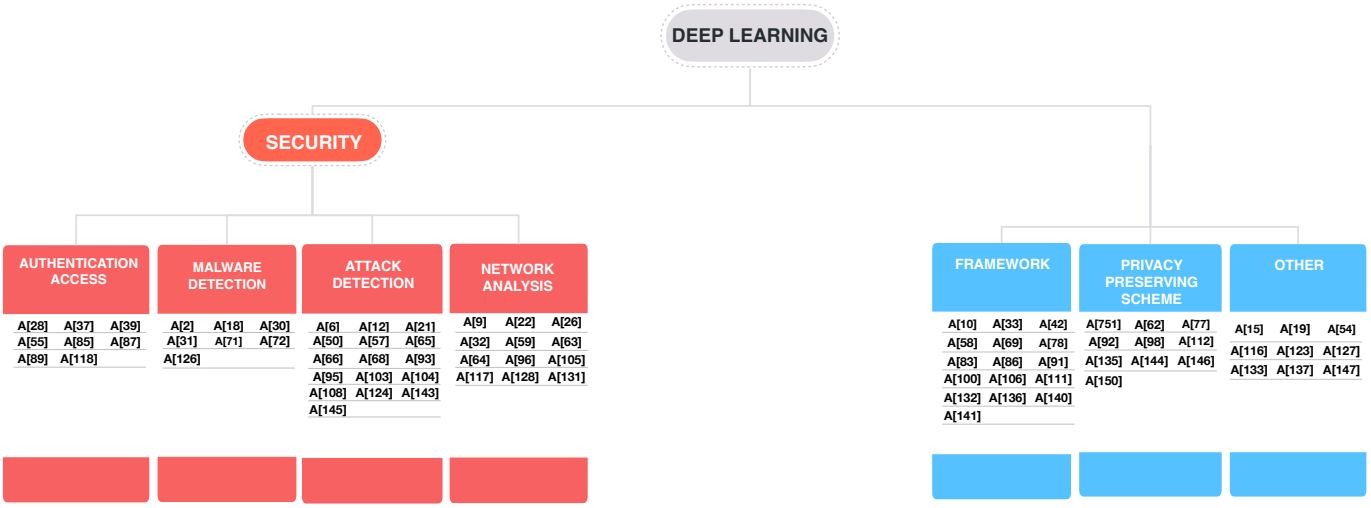

Figure 13: Use cases where the deep learning techniques have been used.

tion of supervised learning, while other types of artificial intelligence techniques seem to have been neglected. As such, our systematic literature review suggests that additional analyses might focus on unsupervised learning and orthogonal techniques, like evolutionary algorithms or pattern recognition. Secondly, we identified only three empirical investigations aimed at comparing the various forms of artificial intelligence techniques employed [A19, A133, A147]. In this sense, we highlight further possibilities for the empirical software engineering community, which might exploit our literature survey's outcome to design and execute empirical investigations into the matter.

Summary.

The results obtained from **RQ**$_3$ indicate that the large majority of primary studies focused on supervised learning techniques to deal with privacy concerns. Yet, we highlight the lack of analyses on other types of artificial intelligence approaches, other than the lack of empirical studies to compare the existing techniques.

4.4. ***RQ***$_4$ *- On the datasets employed by the artificial intelligence methods.*

After providing an overview of the tasks, domains, and families of techniques employed to deal with privacy concerns in IoT systems, we started our fine-grained analysis on the design and evaluation of these techniques. With **RQ**$_4$, we collected data and characteristics of the datasets used by the primary studies. Table 4 reports the list of datasets, along with information on their category, the tasks for which they were employed, and where to find them.

Figure 14 shows top-10 frequency of use of each dataset.

Among the available datasets, we observed that most of the primary studies (42%, 64 papers) relied on *"MINIST"*, while the others have been exploited to a lower extent. More specifically, let us comment on those datasets:

"MNIST". This is a dataset of handwritten digits created for the specific purpose of experimenting machine learning techniques. It indeed contains about 60,000 examples and a test set of 10,000 samples. It was used in 42% of the papers [A15, A33, A53, A59, A77, A78, A81, A83, A91, A100, A105, A132, A135–A137, A142, A144, A146, A148, A151], and is particularly indicated to test authentication techniques that rely on biometric data. For instance, Jiang *et al.* [59] experimented with biometric data obfuscation techniques to verify how the digits classification performance of a deep neural network vary with respect to the case where the network is trained with the original, cleaned digits.

"CIFAR-10". It consists of 60,000 images categorized in 10 classes. The dataset has been created for the specific case of machine learning, as it contains around 6,000 images for each class and is released so that a researcher can use 50,000 images for training and 10,000 images for testing machine learning techniques. We have found ten papers that used this dataset [A15, A29, A77, A91, A100, A105, A132, A136, A137, A151]. Typically, it is used to experiment classification algorithms aiming at addressing the privacy concerns of images and videos, like the problem of understanding whether sensitive data can be derived from fragments of images captured by sensors [A91, A136, A151].

"KDD Cup 99". This dataset contains raw signals obtained in nine weeks from the *TCP dump*. The dataset includes 24 training attacks and 14 types of test data. This dataset is used for the *intrusion detection* learning task and to build supervised algorithms that could learn from these examples to predict the emergence of intrusion attacks [A82, A122, A124, A145].

Dataset	Category	Task	Paper	Link
MNIST	Handwritten	Comparison	[A15, A53, A137]	yann.lecun.com/exdb/mnist/
		Network Analysis	[A59, A148, A151]	
		Framework	[A33, A78, A81, A83, A91, A100, A105, A116, A132, A136]	
		Privacy Preserving Scheme	[A77, A135, A144, A146]	
		User Authentication	[A142]	
CIFAR-10	Image Classification & Object Detection	Comparison	[A15, A137]	www.cs.toronto.edu/~kriz/cifar.html
		Network Analysis	[A29, A151]	
		Framework	[A91, A100, A105, A132, A136]	
		Privacy Preseving Scheme	[A77]	
KDD Cup 99	Cybersecurity	Attack Detection	[A82, A122, A124, A145]	www.kdd.org/kdd-cup/view/kdd-cup-1999/Data
DS2OS	Cybersecurity	Attack Detection	[A50, A65, A68]	www.kaggle.com/francoisxa/ds2ostraffictraces
Adult	Personal Information	Comparison	[A147]	www.kaggle.com/wenruliu/adult-income-dataset
		Network Analysis	[A45]	
Heart Disease	Healthcare	Secure Training	[A49]	archive.ics.uci.edu/ml/datasets/heart+disease
		Network Analysis	[A115]	
CASIA-WebFace	Face Recognition	Network Analysis	[A131]	paperswithcode.com/dataset/casia-webface
		Framework	[A116]	
CTU-13	Cybersecurity	Comparison	[A19]	www.stratosphereips.org/datasets-ctu13
		Malware Detection	[A113]	
Fashion MNIST	Object Detection & Image Classification	Framework	[A81, A100]	www.kaggle.com/zalando-research/fashionmnist
GeoLife	Tracking GPS	Attack Detection	[A134]	www.microsoft.com/en-us/download/details.aspx?id=52367
		Framework	[A90]	

Table 4: List of datasets used to device or experiment the artificial intelligence techniques proposed in literature.

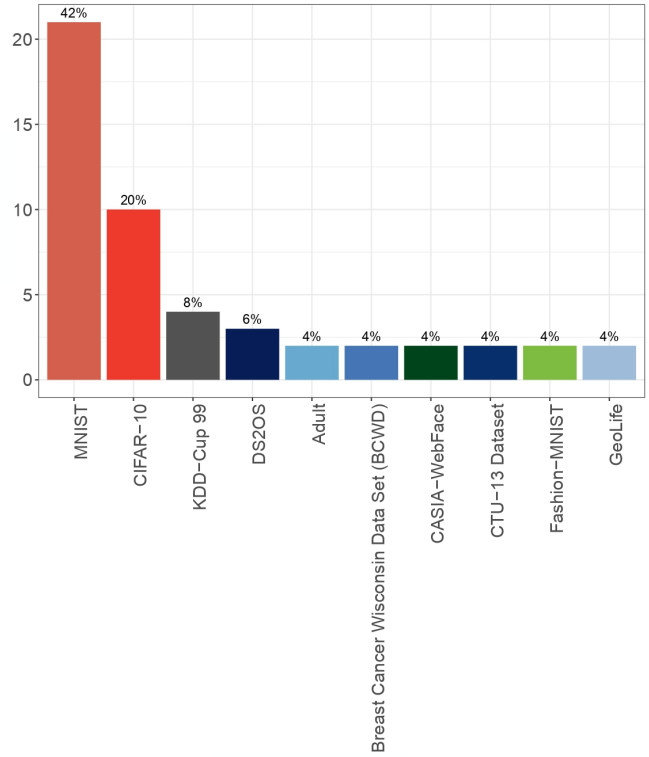

Figure 14: Frequency of use of the datasets exploited by the primary studies.

"DS2OS". The dataset contains traffic traces obtained in IoT environments and is typically used to verify *anomaly detection* algorithms [A50, A65, A68]. For instance, Hasan et al. [50] employed anomaly detection on this dataset to identify possible attacks in IoT sites.

"Adult Dataset". It contains information about people, including the annual income. This is typically exploited by researchers interested in building and/or assessing techniques to detect personal data losses. Similar to the cases above, the dataset seems to be particularly useful for classification algorithms, given that it reports labeled data that can be used for training purposes [A45, A147].

"Breast Cancer Wisconsin DataSet". This dataset is computed from a breast mass digitization image of fine needle aspirate (FNA). The dataset contains 569 instances and 32 attributes (including symmetry, concave points, area). Researchers have been using the dataset to experiment with supervised *classification* techniques that aim at identifying potential personal data losses [A49, A115].

"CASIA-WebFace". The dataset contains 494,414 face images of 10,575 real identities. This is typically used for face verification and face identification tasks [A116, A131]. For instance, Wang et.al. [A131] used this dataset to experiment with a combination of *Deep Neural Network* and *Support Vector Machines* able to analyze video streams and identify and blur faces based on privacy policies, obtaining an accuracy rate close to 92%.

"CTU-13". The dataset contains botnet traffic captured in regular traffic and background traffic. Researchers used the dataset for *malware detection* tasks [A19, A113]. As an example, Bansal et.al. [A19] employed multiple machine learning algorithms, including *Naive Bayes* and *Neural Networks*, to detect botnets, obtaining F-1 scores up to 88%.

"Fashion MNIST". This dataset includes ZALANDO's article images and contains about 60,000 training set images and 10,000 test set examples. It is divided into 10 classes (includes T-shirt, pullover, and coat), and each category contains 10,000 examples [A81, A100]. The primary studies that exploited this dataset were interested in building techniques that might prevent privacy leaks due to the identification of people from their clothes.

"GeoLife". This dataset contains GPS trajectories collected in over three years. The dataset includes information about the time-stamped and information about the latitude, longitude, and altitude [A90, A134]. It has been used to train and test techniques that could prevent the localization of people based on their coordinates.

To broaden the scope of the discussion, it is worth focusing on the tasks for which each of the above datasets has been used—Table 4 reports the details of our analysis. We could first notice that the *"MINIST"* dataset has been used

by multiple authors to pursue several tasks: authors opted for it when performing comparisons among artificial intelligence techniques [A15, A53, A137], building frameworks [A33, A78, A81, A83, A91, A100, A105, A116, A132, A136] or experimenting with network analysis approaches and authentication mechanisms, other than verifying the accuracy of privacy preserving schemas. Similar conclusions can be drawn for the *"CIFAR-10"* dataset. These observations highlight the flexibility of the datasets with respect to multiple tasks. On another note, the other datasets have been used in a more restrictive manner and mostly for dealing with the task of *"Attack Detection"*. This is even the only task considered when employing the *"DS20S"* and *"KDD Cup 99"* datasets.

While we already discussed about the availability of only a few datasets for experimenting with artificial intelligence techniques, some additional observations should be made. The authors of the *"Fashion MNIST"* dataset openly criticized the original *"MNIST"* dataset. Indeed, its structure allows both traditional and deep learning algorithms to achieve very high accuracy without providing enough insights into the actual validity of the predictions performed. In other words, the dataset is built in a biased way that impacts the analysis of the real capabilities of the experimented techniques. Other data scientists and practitioners also remarked this. In April 2017, a GOOGLE BRAIN research scientist and deep learning expert, Ian Goodfellow, advised migrating to other datasets. Later on, another deep learning expert, François Chollet, explained that the *"MNIST"* dataset is not good at representing everyday tasks. These considerations, along with the consideration that a large number of primary studies employed this dataset, allow us to claim that the research in privacy of IoT devices might require a critical re-assessment. This is further confirmed by the fact that 19 primary studies evaluated the proposed approaches only in terms of accuracy [A15, A33, A53, A59, A77, A78, A81, A83, A100, A105, A116, A132, A135–A137, A142, A146, A148, A151], hence possibly biasing the conclusions drawn—more details are reported when addressing **RQ**$_6$.

Another discussion point concerns with the intrinsic characteristics of the datasets. When addressing **RQ**$_4$, we analyzed whether and to what extent the available datasets are balanced. Data balancing is a crucial data quality aspect to take into account while selecting a suitable dataset to create and/or validate privacy approaches [A6]. The availability of balanced datasets, namely of datasets for which there are a similar amount of data for each class, might notably affect the learning capabilities of artificial intelligence techniques [6], other than implying the definition of methodological steps that aim at facing this potential learning bias.

Table 5 summarizes our analysis on the data balancing of the considered datasets. The three most used datasets are balanced, while *"DS20S"* and *"Breast Cancer Wisconsin DataSet"* are not. Researchers can use this information to take appropriate data balancing considerations during the design of their studies, other than exploit it to analyze deeper the validation of the proposed techniques (**RQ**$_6$).

Dataset	Balanced
MNIST	Yes
CIFAR-10	Yes
KDD Cup 99	Yes
DS2OS	No
Adult	No
Breast Cancer Wisconsin DataSet (BCWD)	No
CASIA-WebFace	No
CTU-13	No
Fashion MNIST	Yes
GeoLife	No

Table 5: Classify balanced or unbalanced datasets

Summary.

We point out the need for further open datasets that may cover a larger variety of privacy concerns. About 40% of the papers conducted experiments on the *"MNIST"* dataset (an Handwritten dataset). Nonetheless, it has been criticized, as it may lead to biased interpretations of the results obtained by artificial intelligence techniques. As a consequence, the conclusions drawn by most of the papers published so far might need to be re-assessed.

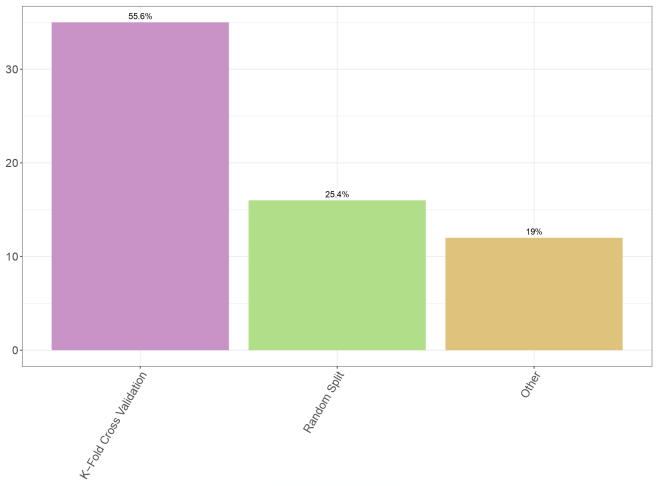

Figure 15: Validation techniques used by the primary studies.

4.5. **RQ**$_5$ - *On the validation strategies employed to assess the artificial intelligence methods.*

In the context of **RQ**$_5$ we analyzed the validation strategies adopted when assessing the capabilities of artificial intelligence methods devised to deal with privacy concerns in IoT systems. As clarified in Section 3.7, not all the primary studies validated the proposed techniques. This was the case for 21 papers [A6, A23, A34, A39, A41, A51, A54, A56, A58, A76, A86, A87, A89, A91, A106, A111, A123, A130, A138, A139, A150]: hence, this research question takes the validation procedures

of 131 primary studies into account. Figure 15 shows the results of our analysis.

55.6% of the studies (35) that explicitly indicate the validation technique used the so-called *k*-fold cross validation [A1, A2, A5, A8, A9, A15, A18, A20, A26, A28, A40, A43, A46, A48, A50, A52, A70, A71, A74, A79, A90, A93, A94, A97, A104, A107, A109, A112, A115, A120, A122, A126–A128, A133], with a value of *k* equals to 5 or 10. This is a method that can be used to estimate the performance of machine learning algorithms: it randomly splits a dataset in *k* groups called folds and (1) takes one fold as test and k-1 folds as training, (2) fits a machine learning model and executes it on the current test fold, and (3) iterates the procedure until all unique folds have been considered exactly once as test set. Upon completion of the validation procedure, the results obtained are summarized by means of statistical indicators like, for instance, the mean number of true positive instances identified over the various runs of the validation.

When analyzing the primary studies, we could not identify the specific reasons leading researchers to use this validation procedure. We cannot provide insights on whether its adoption was the most suitable one or whether other validation procedures would have better fit the specific problems treated in the studies. The only exception to this general discussion concerns the work by Meurisch et al. [A90]. The study proposed an AI-based privacy-preserving mechanism to overcome the need to continuously share user data streams in the cloud, which was later validated through 10-fold cross-validation. The proposed approach employs temporal data, namely data collected over a given time frame and that, for this reason, follow a temporal order. The application of cross-validation in this scenario risks bias the interpretation of the results [8]. Indeed, the cross-validation might accidentally lead future data to be used for training past data, causing a form of data leakage that interprets results biased. Unfortunately, we could not understand if and how the authors have mitigated the risks connected to the adoption of cross-validation. Yet, we can argue that more details on the rationale and the methodology adopted to validate the artificial intelligence methods would be required to properly assess the validity of the insights provided.

Besides the cross validation, another popular strategy is the so-called random split or percentage split. This randomly splits the dataset into training and test sets, e.g., retaining 80% for training and 20% for testing [A4, A29, A33, A37, A49, A55, A68, A73, A78, A86, A103, A113, A117, A124, A135, A138, A147]. Similarly to the discussion above, the primary studies that employed this validation did not explicitly mention the rationale behind its use nor the possible threats that this validation might cause.

The last 29 primary studies used different strategies, which we grouped as *"Other"* in Figure 15. These studies employed various validation methods, like the Monte Carlo cross validation [A88] or time-sensitive analyses [A67].

[8] https://medium.com/@soumyachess1496/cross-validation-in-time-series-566ae4981ce4

Dataset	Training/Validation Strategy
MNIST	K-Fold Cross-Validation, Random Split
CIFAR-10	K-Fold Cross-Validation, Random Split
KDD Cup 99	K-Fold Cross-Validation
DS2OS	K-Fold Cross Validation, Random Split
Adult	Random Split
Breast Cancer Wisconsin DataSet (BCWD)	K-Fold Cross-Validation
CASIA-WebFace	Other
CTU-13	Other, Random Split
Fashion MNIST	Not specified
GeoLife	K-Fold Cross Validation

Table 6: Datasets used and corresponding training/validation strategies employed.

Summary.

Cross-validation and random split are the two most common validation procedures employed in the literature. Nonetheless, the methodological choices made when selecting the validation strategies are often unclear or not specified. In some cases, the validation strategies adopted risk bias the interpretation of the performance of artificial intelligence methods. As such, we argue that more details should be provided for rigorousness, reproducibility, and replicability of the research.

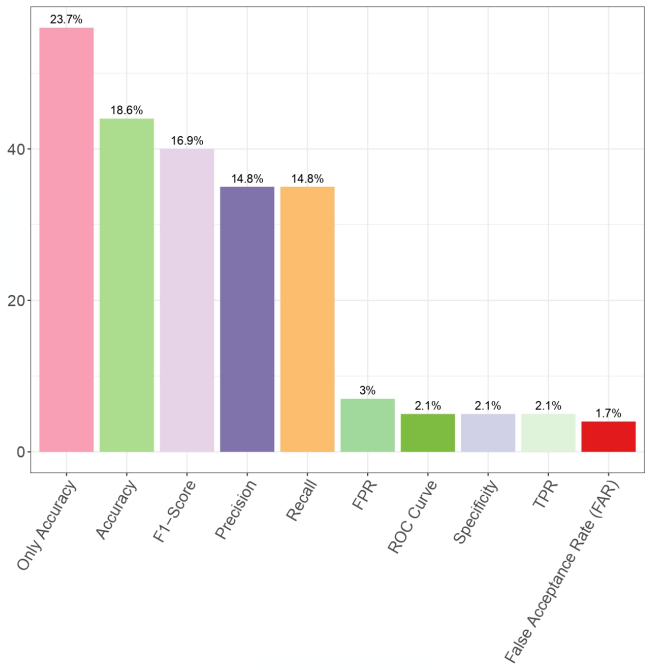

Figure 16: Evaluation Metrics Techniques.

4.6. RQ_6 - *On the evaluation metrics employed to assess the artificial intelligence methods.*

The last perspective of our study was concerned with the evaluation metrics employed to measure the performance of the artificial intelligence techniques proposed in literature. The results for $\mathbf{RQ}_6$ are plotted in Figure 16.

Paper	Dataset	Evalation Metrics
Attack and anomaly detection in IoT sensors in IoT sites using machine learning approaches	DS20S	Accuracy, Precision, Recall, F1-Score, Roc Curves
A Novel Attack Detection Scheme for the Industrial Internet of Things Using a Lightweight Random Neural Network	DS20S	Accuracy, Precision, Recall, F1-score
Ensemble Learning for Detecting Attacks and Anomalies in IoT Smart Home	DS20S	Accuracy, Precision, Recall, F1-Score
Preserving User Privacy for Machine Learning: Local Differential Privacy or Federated Machine Learning?	Adult	Accuracy
A Differentially Private Big Data Nonparametric Bayesian Clustering Algorithm in Smart Grid	Adult	Accuracy
Privacy-Preserving Support Vector Machine Training Over Blockchain-Based Encrypted IoT Data in Smart Cities	Breast Cancer Wisconsin DataSet (BCWD)	Accuracy
Privacy-preserving k-nearest neighbors training over blockchain-based encrypted health data	Breast Cancer Wisconsin DataSet (BCWD)	Accuracy, Precision, Recall

Table 7: Evaluation metrics used when working on unbalanced datasets.

We found that a large quantity of papers only relied on the accuracy metric[A4, A5, A8, A9, A15, A16, A20, A21, A26, A27, A29, A31–A33, A37, A42, A45, A48, A53, A59, A60, A64, A66, A67, A69, A70, A73, A75, A77–A79, A81, A83, A90, A93, A97, A98, A100, A105, A108, A112, A115, A116, A118, A121, A125, A128, A129, A131, A132, A135–A137, A140, A142, A146–A148, A151], that is, the total amount of correct predictions with respect to all the predictions output. In other cases, the primary studies used a combined approach, for instance by computing accuracy and precision, recall and precision, and so on. In any case, there are two observations to make on the choice of the evaluation metrics. In the first place, and similarly to the discussion done in $\mathbf{RQ}_5$, most of the surveyed studies did not report on the rationale for using these metrics nor on their suitability for the considered problem. As an example, let consider the case of accuracy. By definition, the value of the metric increases as the number of both true positives and negatives increases. Some of the datasets currently available are strongly unbalanced and contain only a few elements characterizing privacy issues—this is, for example, the case of the *"Adult"* and *"Breast Cancer Wisconsin DataSet"* discussed in $\mathbf{RQ}_4$. For these datasets, one is reasonably interested in assessing the performance of artificial intelligence techniques with respect to their capabilities in correctly predicting the privacy issues appearing in the minority class of the dataset. Nonetheless, training an AI-based solution with only a few instances of the class of interest might lead the approach not to properly learn how to classify them. On the contrary, the approach might be biased toward the classification of the majority class, namely the set of instances that do not present any privacy concern. As a consequence, measuring the accuracy metric might provide a wrong view of the performance, since the metric tends to reward the approach independently from which class it is able to correctly predict. High accuracy values can therefore indicate that the artificial intelligence approach is able to correctly predict the majority class, which is the least interesting. We could not find considerations of this type in the primary studies considered and, unfortunately, this might have had an impact on the conclusions drawn by various studies. More specifically, Table 7 reports the primary studies that have worked on unbalanced datasets; the last column of the table also reports the evaluation metrics considered. As it is possible to observe, all of them relied on the accuracy—for some of them, this was the only metric considered—but, perhaps more importantly, only a few assessed the performance of the techniques in a more comprehensive manner through other metrics. A second point of discussion is still related to the relation between the datasets exploited and the metrics used for evaluation. Even the primary studies that worked on balanced datasets typically relied on the accuracy. While in these cases the use of accuracy was more reasonable, some peculiarities of the datasets might have influenced the choice of the evaluation metric. For instance, as discussed in $\mathbf{RQ}_4$, most of the primary studies relied on the *"MNIST"* dataset, which turned to be somehow biased toward accuracy, i.e., as already explained, the major criticism made was concerned to the fact that any AI-based technique can easily reach high accuracy levels on this dataset. As such, it is likely to believe that a re-evaluation of the techniques proposed in literature might be beneficial for the research community in order to more appropriately understand the actual value of those techniques.

Summary.

23% of the primary studies only used accuracy to evaluate the quality of the artificial intelligence techniques. However, the characteristics of the datasets might make them biased toward accuracy, implying a biased interpretation of the real capabilities of the proposed techniques.

5. Discussion and Implications

The results of our study pointed out several observations that are worth further discussion. These observations also define key implications and open challenges for researchers working or who are willing to work on the definition of software engineering methods and practices for detecting and preserving privacy concerns in IoT systems.

"AI & IoT privacy" is a hot topic. First and foremost, our systematic literature review provides clear evidence on the relevance of using artificial intelligence methods to deal with privacy concerns in IoT systems. Not only the number of papers is rapidly increasing, but the trend seems to be exponential. Based on this observation, we can provide two key implications that target the software engineering research community as a whole, other than the educational aspects of the matter:

☞ The interest toward the topic is rising fast, and, as indicated by the results of our study, the research community would need additional resources to further elaborate on the most appropriate methods for adopting artificial intelligence approaches in the context of privacy detection and preservation. We see this as an opportunity for the software engineering community, which might come into play and support the research efforts done by researchers of other areas. Our results indeed indicate the need for a joint research effort between various communities, like the ones of machine learning, software engineering, and software security. These represent precious opportunities for fresh Ph.D. students and researchers in general. The former might consider working toward this subject to contribute to a growing field. The latter might consider embracing the current challenges to create joint workgroups that may take advantage of complementary expertise to improve the current support provided to practitioners. In this sense, it is our hope that the findings obtained in our systematic literature review might be inspiring for new researchers.

☞ The definition of artificial intelligence techniques to deal with privacy concerns in IoT systems might not only be of the interest of researchers, but also of the educators, called to provide (under-)graduate students with elements that can be used to improve the state of the practice further. As a consequence, the results of our study might be used to stimulate the creation of ad-hoc study plans and courses that encourage a practical approach to the use of artificial intelligence and configuration of AI pipelines, other than its application to novel contexts like IoT and privacy of IoT systems. This would also likely stimulate an increased collaboration with practitioners and security experts, who might be interested in sharing their own experience to form the new generation of experts in AI and IoT privacy.

On the programming language support. According to our preliminary analysis of the primary studies, PYTHON seems to be the most mature programming language to support the design and development of artificial intelligence techniques for privacy concerns detection and preservation. While this was somehow expected, given the amount of data science libraries and frameworks available for this programming language, we argue that our observations provide two implications for programming language designers and software engineering researchers:

☞ The designers of other programming languages might want to take our results as a motivation to propose novel instruments and tools to let researchers and practitioners work with other programming languages. Also, the maintainers of known libraries, e.g., WEKA, along with software engineering researchers, might further understand the reasons behind their low adoption and design methodologies that increase the overall usability of their APIs.

☞ The popularity of the PYTHON language for data science and artificial intelligence also has clear implications on the educational side. While many institutions have already courses teaching this programming language, others might exploit the observations of this systematic literature review to motivate the introduction of specific data science programming courses where students are exposed to the use of well-known libraries that facilitate the definition of artificial intelligence and/or machine learning techniques.

On verifiability and replicability. As noticed throughout our systematic literature review and analyses, a large number of primary studies do not report granular information to enable neither verification nor replication. In addition, the papers are rarely accompanied by replication packages that make data and scripts publicly available for other researchers. In our humble opinion, this represents a key threat to dissemination and verification of the published research papers, which leads to two implications concerned with the way research papers are disseminated:

☞ Researchers should consider including more methodological details to enable an improved understanding of the design and definition of the proposed techniques. This would be beneficial in terms of dissemination, as practitioners might better understand how to put the defined techniques into practice, increasing the overall impact of the research on the matter. At the same time, this would support research, since additional investigations might be made on top of the findings achieved by previous researchers, further increasing the impact of research.

☞ Researchers should complement their own work with online, publicly available appendix reporting data and scripts used to experiment with artificial intelligence techniques, other than to mine the data required to execute them. Other than providing priceless support for researchers, this would be in line with the most recent guidelines and regulations on open science, made available by public institutions like, for instance, the European Com-

mission.[9] The software engineering research community is, in this sense, pioneering the rise of open data science and data, e.g., the Journal of Systems and Software has recently introduced new open science review processes. By contributing more to the understanding of IoT systems, we believe that software engineering might become the driving wheel of a change in terms of transparency.

AI & IoT privacy: The road ahead for SE4AI research. The results to our research questions clearly indicate that there is still a long and winding road to making artificial intelligence suitable for the problem of IoT privacy. According to our analyses, this pertains to several aspects that, in turn, call for several implications for software engineering for the artificial intelligence research community:

☞ Researchers have been mainly focusing on six tasks that have to do with user authentication, network analysis, and others. At the same time, our systematic search identified several other tasks that have received less attention and that further research might consider. Perhaps more importantly, we highlight the lack of insights from the trenches, namely the lack of empirical investigations that target the practitioners' and IoT users' perspectives and might reveal other relevant tasks that the current body of knowledge has neglected. As a consequence, we claim that the first relevant aspect for future research in the field is represented by a large-scale analysis of IoT privacy in practice.

☞ Our analysis showed that researchers have mainly relied on supervised learning algorithms. While the assessment reported in the primary studies has shown promising results, adopting alternative artificial intelligence approaches might provide an additional boost to current solutions. Future research might devote effort in understanding how and how well-unsupervised learning approaches, rather than evolutionary search-based algorithms, which have been successfully applied in other software engineering domains [21, 31], might be exploited for various privacy tasks. Still, in this respect, we point out the lack of empirical research: this is, in our opinion, a critical threat that precludes an improved assessment of how artificial intelligence can be exploited for privacy detection and preservation. We envision and wish a radical change in this sense, with an ever-increasing amount of empirical studies aiming at comparing and experimenting with various artificial intelligence methods, other than assessing how specific configurations of the techniques (e.g., the hyper-parameter tuning of the machine learners or the role of data balancing) might impact the performance of the existing techniques.

☞ A third, critical issue unveiled by our work relates to the public datasets currently available. Besides having only a few datasets to experiment with, the major criticism is concerned with the level of realism and actual suitability of these datasets. As commented in $\mathbf{RQ}_4$, some datasets are unbalanced, being potentially unsuitable for training artificial intelligence techniques. Moreover, the most widely used datasets are biased toward certain accuracy indicators, hence biasing the interpretation of the results. This is, likely, the most important issue encountered by our systematic review, as it impacts most research conducted so far. Therefore, we argue that concrete steps should be conducted to devise novel, more reliable datasets to re-assess the experiments performed so far. We hope that the indications provided by our work, in terms of limitations and challenges of the methodology employed by current papers, might help design better the empirical investigations into the performance of artificial intelligence techniques.

☞ As a follow-up discussion, we believe there is a need for a joint community effort to establish guidelines and/or standard templates to devise and validate artificial intelligence methods. These guidelines might make researchers aware of common pitfalls to avoid other than best practices to follow. While our study calls for additional work in the area, we can already distill some insights. In the first place, researchers should consider clarifying the rationale behind each methodological decision made when building and validating artificial intelligence approaches to increase the understandability of the proposed techniques and case studies. Second, we noticed that in some cases, the primary studies did not consider some potential threats to validity, e.g., the use of cross-validation in the presence of time sensitive data, which might limit the realism of the conclusions drawn or even provide wrong outcomes. The definition of best and bad practices might mitigate risks connected to biased interpretation of the results. Third, researchers should more carefully select the evaluation metrics employed to assess the proposed techniques to avoid possible interpretation errors. All these challenges are the core of software engineering for artificial intelligence, which is called to define instruments to build and validate artificial intelligence methods for IoT privacy properly.

6. Threats to Validity

As any other systematic literature review, ours has some limitations that may have threatened the validity of the reported findings. This section discusses them along with the mitigation strategies employed to address them.

Literature selection. A critical challenge for a systematic literature review consists of identifying a complete set of papers to enable a comprehensive overview of state of art. In this respect, we have first defined a search query having the goal of retrieving as many papers related to the use of artificial intelligence for dealing with privacy of IoT systems as possible without any temporal limitations: while this choice has implied a higher effort in terms of manual analyses, we preferred it for the sake of completeness. Furthermore, we

[9] The EU regulations on Open Science: https://ec.europa.eu/info/research-and-innovation/strategy/strategy-2020-2024/our-digital-future/open-science_en.

identified synonyms or alternative spellings of terms typically used in literature when defining the search query. In addition, we checked the presence of the search terms within existing systematic literature reviews on IoT privacy to find possible additional terms. To further increase the completeness of our study, we also conducted a backward snowballing session on the papers that passed the exclusion/inclusion criteria. Perhaps more importantly, all the steps leading to the selection of primary studies were always double-checked by at least one of the paper's authors. The combination of these actions makes us confident of the completeness of the literature selection. Nevertheless, for the sake of verifiability and replicability, we have provided as additional contribution an online appendix reporting all steps and intermediate results of our analyses [1].

Literature analysis and synthesis. Upon completion of the selection process, we have applied specific exclusion criteria intending to filter out papers that could not contribute or could provide a limited contribution toward the summarization of state of the art related to the defined research questions. Moreover, we did not limit the selection of primary studies to those that successfully passed the inclusion criteria, but also conducted an additional quality assessment to verify their actual suitability to our purposes. Such a manual assessment has further limited the risk of including resources that did not fit our purposes.

More generally, the literature synthesis has been conducted based on manual analyses, which are subject to human error by nature. In this respect, there are two observations to be done. First, the two first authors have continuously worked together, hence limiting the risk of subjectiveness and/or errors. Second, the paper's third author has been constantly involved in the process and, whenever needed, provided insights on how to conduct the various steps of the systematic literature review.

7. Conclusion

This paper reports on a systematic literature review on the application of artificial intelligence techniques to the problem of privacy detection and preservation in IoT systems. We tackled a number of research angles, which aimed at assessing the current state of the art with respect to the types of privacy concerns treated, the characteristics of the artificial intelligence techniques defined, their limitations, and the domains where these have been applied to. The results of our systematic analysis let emerge that there is a substantial lack of software engineering research in the field of IoT privacy and its management through artificial intelligence approaches. As such, our findings represent a call to actions and provided a number of key implications for future software engineering efforts: as an example, the need for improved validation mechanisms that might more reliably assess the capabilities of the techniques exploited. It is our hope that further researchers and fresh Ph.D. Students interested in working along the lines of privacy of IoT devices and artificial intelligence might be inspired by our work to contribute to the improvement of a research field that may benefit of additional empirical analyses, techniques, and large-scale investigations into the effectiveness of artificial intelligence for IoT privacy. All in all, our paper provides the contributions listed in the following:

- A systematic literature review that summarizes the current knowledge on the use of artificial intelligence techniques for dealing with privacy in IoT systems;
- The identification of limitations, open issues, and challenges of the state of the art, which researchers might use to define the next research steps to improve the support given to developers;
- An online appendix [1] providing data to replicate and verify the systematic work done to conduct our study.

Our future research agenda is driven by the considerations and implications of this systematic literature review. We indeed aim at performing empirical analyses to compare the effectiveness of the currently available solutions, other than proposing novel methodologies and instruments to help developers deal with privacy concerns in practice.

Acknowledgment

Fabio is partially supported by the Swiss National Science Foundation - SNF Project No. PZ00P2_186090 (TED). This work has been partially supported by the EMELIOT national research project, which has been funded by the MUR under the PRIN 2020 program (Contract 2020W3A5FY).

Appendix

[A1] Acar, A., Fereidooni, H., Abera, T., Sikder, A.K., Miettinen, M., Aksu, H., Conti, M., Sadeghi, A.R., Uluagac, S., 2020. Peek-a-boo: I see your smart home activities, even encrypted!, in: Proceedings of the 13th ACM Conference on Security and Privacy in Wireless and Mobile Networks, pp. 207–218.

[A2] Ahmed, F.S., Mustapha, N., Mustapha, A., Kakavand, M., Foozy, C.F.M., 2020. Preliminary analysis of malware detection in opcode sequences within iot environment. Journal of Computer Science 16, 1306–1318.

[A3] AL-Akhras, M., Alawairdhi, M., Alkoudari, A., Atawneh, S., 2020. Using machine learning to build a classification model for iot networks to detect attack signatures. International journal of Computer Networks & Communications 12, 99–116. doi:10.5121/ijcnc.2020.12607.

[A4] Al Mtawa, Y., Singh, H., Haque, A., Refaey, A., 2020. Smart home networks: Security perspective and ml-based ddos detection, in: 2020 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), IEEE. pp. 1–8.

[A5] Alam, M.S., Vuong, S.T., 2013. Random forest classification for detecting android malware, in: 2013 IEEE international conference on green computing and communications and IEEE Internet of Things and IEEE cyber, physical and social computing, IEEE. pp. 663–669.

[A6] Albalawi, U., 2020. A comprehensive analysis on intrusion detection in iot based smart environments using machine learning approaches. International Journal of Scientific & Technology Research 9, 1646–1652.

[A7] Alshehri, A., Granley, J., Yue, C., 2020. Attacking and protecting tunneled traffic of smart home devices, in: Proceedings of the Tenth ACM Conference on Data and Application Security and Privacy, pp. 259–270.

[A8] Alsoliman, A., Rigoni, G., Levorato, M., Pinotti, C., Tippenhauer, N.O., Conti, M., 2021. Cots drone detection using video streaming characteristics, in: International Conference on Distributed Computing and Networking 2021, pp. 166–175.

[A9] Alturki, B., Reiff-Marganiec, S., Perera, C., 2017. A hybrid approach for data analytics for internet of things, in: Proceedings of the Seventh International Conference on the Internet of Things, pp. 1–8.

[A10] Amoon, M., Altameem, T., Altameem, A., 2020. Internet of things sensor assisted security and quality analysis for health care data sets using artificial intelligent based heuristic health management system. Measurement 161, 107861.

[A11] An, N., Duff, A., Naik, G., Faloutsos, M., Weber, S., Mancoridis, S., 2017. Behavioral anomaly detection of malware on home routers, in: 2017 12th International Conference on Malicious and Unwanted Software (MALWARE), IEEE. pp. 47–54.

[A12] Anthi, E., Williams, L., Słowińska, M., Theodorakopoulos, G., Burnap, P., 2019. A supervised intrusion detection system for smart home iot devices. IEEE Internet of Things Journal 6, 9042–9053.

[A13] Anton, S.D., Kanoor, S., Fraunholz, D., Schotten, H.D., 2018. Evaluation of machine learning-based anomaly detection algorithms on an industrial modbus/tcp data set, in: Proceedings of the 13th international conference on availability, reliability and security, pp. 1–9.

[A14] Antonini, M., Vecchio, M., Antonelli, F., Ducange, P., Perera, C., 2018. Smart audio sensors in the internet of things edge for anomaly detection. IEEE Access 6, 67594–67610.

[A15] Arachchige, P.C.M., Bertok, P., Khalil, I., Liu, D., Camtepe, S., Atiquzzaman, M., 2019. Local differential privacy for deep learning. IEEE Internet of Things Journal 7, 5827–5842.

[A16] Arachchige, P.C.M., Bertok, P., Khalil, I., Liu, D., Camtepe, S., Atiquzzaman, M., 2020. A trustworthy privacy preserving framework for machine learning in industrial iot systems. IEEE Transactions on Industrial Informatics 16, 6092–6102.

[A17] Arca, S., Hewett, R., 2020. Privacy protection in smart health, in: Proceedings of the 11th International Conference on Advances in Information Technology, pp. 1–8.

[A18] Azmoodeh, A., Dehghantanha, A., Choo, K.K.R., 2018. Robust malware detection for internet of (battlefield) things devices using deep eigenspace learning. IEEE transactions on sustainable computing 4, 88–95.

[A19] Bansal, A., Mahapatra, S., 2017. A comparative analysis of machine learning techniques for botnet detection, in: Proceedings of the 10th International Conference on Security of Information and Networks, pp. 91–98.

[A20] Batool, S., Hassan, A., Saqib, N.A., Khattak, M.A.K., 2020. Authentication of remote iot users based on deeper gait analysis of sensor data. IEEE Access 8, 101784–101796.

[A21] Bendiab, G., Grammatikakis, K.P., Koufos, I., Kolokotronis, N., Shiaeles, S., 2020. Advanced metering infrastructures: Security risks and mitigation, in: Proceedings of the 15th International Conference on Availability, Reliability and Security, pp. 1–8.

[A22] Bhattacharya, S., Manousakas, D., Ramos, A.G.C., Venieris, S.I., Lane, N.D., Mascolo, C., 2020. Countering acoustic adversarial attacks in microphone-equipped smart home devices. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 4, 1–24.

[A23] Bugeja, J., Jacobsson, A., Davidsson, P., 2020. Is your home becoming a spy? a data-centered analysis and classification of smart connected home systems, in: Proceedings of the 10th International Conference on the Internet of Things, pp. 1–8.

[A24] Cabra, J.L., Mendez, D., Trujillo, L.C., 2018. Wide machine learning algorithms evaluation applied to ecg authentication and gender recognition, in: Proceedings of the 2018 2nd International Conference on Biometric Engineering and Applications, pp. 58–64.

[A25] Cao, H., Liu, S., Zhao, R., Xiong, X., 2020. Ifed: A novel federated learning framework for local differential privacy in power internet of things. International Journal of Distributed Sensor Networks 16, 1550147720919698.

[A26] Caputo, D., Verderame, L., Ranieri, A., Merlo, A., Caviglione, L., 2020. Fine-hearing google home: why silence will not protect your privacy. J. Wirel. Mob. Networks Ubiquitous Comput. Dependable Appl. 11, 35–53.

[A27] Chatterjee, U., Chatterjee, S., Mukhopadhyay, D., Chakraborty, R.S., 2020. Machine learning assisted puf calibration for trustworthy proof of sensor data in iot. ACM Transactions on Design Automation of Electronic Systems (TODAES) 25, 1–21.

[A28] Chauhan, J., Rajasegaran, J., Seneviratne, S., Misra, A., Seneviratne, A., Lee, Y., 2018. Performance characterization of deep learning models for breathing-based authentication on resource-constrained devices. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 2, 1–24.

[A29] Chiu, T.C., Shih, Y.Y., Pang, A.C., Wang, C.S., Weng, W., Chou, C.T., 2020. Semisupervised distributed learning with non-iid data for aiot service platform. IEEE Internet of Things Journal 7, 9266–9277.

[A30] Darabian, H., Dehghantanha, A., Hashemi, S., Homayoun, S., Choo, K.K.R., 2020. An opcode-based technique for polymorphic internet of things malware detection. Concurrency and Computation: Practice and Experience 32, e5173.

[A31] Ding, F., Li, H., Luo, F., Hu, H., Cheng, L., Xiao, H., Ge, R., 2020. Deeppower: Non-intrusive and deep learning-based detection of iot malware using power side channels, in: Proceedings of the 15th ACM Asia Conference on Computer and Communications Security, pp. 33–46.

[A32] Dong, S., Li, Z., Tang, D., Chen, J., Sun, M., Zhang, K., 2020. Your smart home can't keep a secret: Towards automated fingerprinting of iot traffic, in: Proceedings of the 15th ACM Asia Conference on Computer and Communications Security, pp. 47–59.

[A33] Du, W., Li, A., Zhou, P., Xu, Z., Wang, X., Jiang, H., Wu, D., 2020. Approximate to be great: Communication efficient and privacy-preserving large-scale distributed deep learning in internet of things. IEEE Internet of Things Journal 7, 11678–11692.

[A34] El Kalam, A.A., Outchakoucht, A., Es-Samaali, H., 2018. Emergence-based access control: New approach to secure the internet of things, in: Proceedings of the 1st International Conference on Digital Tools & Uses Congress, pp. 1–11.

[A35] Elmisery, A.M., Sertovic, M., Gupta, B.B., 2017. Cognitive privacy middleware for deep learning mashup in environmental iot. IEEE Access 6, 8029–8041.

[A36] Elrawy, M.F., Awad, A.I., Hamed, H.F.A., 2018. Intrusion detection systems for iot-based smart environments: a survey. Journal of Cloud Computing 7, 1–20.

[A37] Farid, F., Elkhodr, M., Sabrina, F., Ahamed, F., Gide, E., 2021. A smart biometric identity management framework for personalised iot and cloud computing-based healthcare services. Sensors 21, 552.

[A38] Fei, J., Yao, Q., Chen, M., Wang, X., Fan, J., 2020. The abnormal detection for network traffic of power iot based on device portrait. Scientific Programming 2020.

[A39] Ferdowsi, A., Saad, W., 2018. Deep learning for signal authentication and security in massive internet-of-things systems. IEEE Transactions on Communications 67, 1371–1387.

[A40] Gassais, R., Ezzati-Jivan, N., Fernandez, J.M., Aloise, D., Dagenais, M.R., 2020. Multi-level host-based intrusion detection system for internet of things. Journal of Cloud Computing 9, 1–16.

[A41] Gebrie, M.T., Abie, H., 2017. Risk-based adaptive authentication for internet of things in smart home ehealth, in: Proceedings of the 11th European Conference on Software Architecture: Companion Proceedings, pp. 102–108.

[A42] Gong, C., Lin, F., Gong, X., Lu, Y., 2020. Intelligent cooperative edge computing in internet of things. IEEE Internet of Things Journal 7, 9372–9382.

[A43] Gong, N.Z., Payer, M., Moazzezi, R., Frank, M., 2016. Forgery-resistant touch-based authentication on mobile devices, in: Proceedings of the 11th ACM on Asia Conference on Computer and Communications Security, pp. 499–510.

[A44] Gu, T., Abhishek, A., Fu, H., Zhang, H., Basu, D., Mohapatra, P., 2020. Towards learning-automation iot attack detection through reinforcement learning, in: 2020 IEEE 21st International Symposium on" A World of Wireless, Mobile and Multimedia Networks"(WoWMoM), IEEE. pp. 88–97.

[A45] Guan, Z., Lv, Z., Sun, X., Wu, L., Wu, J., Du, X., Guizani, M., 2020. A differentially private big data nonparametric bayesian clustering algorithm in smart grid. IEEE Transactions on Network Science and Engineering 7, 2631–2641.

[A46] Ham, H.S., Kim, H.H., Kim, M.S., Choi, M.J., 2014. Linear svm-based android malware detection for reliable iot services. Journal of Applied Mathematics 2014.

[A47] Hamza, A., Gharakheili, H.H., Benson, T.A., Sivaraman, V., 2019. Detecting volumetric attacks on Iot devices via sdn-based monitoring of mud activity, in: Proceedings of the 2019 ACM Symposium on SDN Research, pp. 36–48.

[A48] Han, J., Pan, S., Sinha, M.K., Noh, H.Y., Zhang, P., Tague, P., 2017. Sensetribute: smart home occupant identification via fusion across on-object sensing devices, in: Proceedings of the 4th ACM International Conference on Systems for Energy-Efficient Built Environments, pp. 1–10.

[A49] Haque, R.U., Hasan, A., Jiang, Q., Qu, Q., 2020. Privacy-preserving k-nearest neighbors training over blockchain-based encrypted health data. Electronics 9, 2096.

[A50] Hasan, M., Islam, M.M., Zarif, M.I.I., Hashem, M., 2019. Attack and anomaly detection in iot sensors in iot sites using machine learning approaches. Internet of Things 7, 100059.

[A51] He, X., Jin, R., Dai, H., 2018. Deep pds-learning for privacy-aware offloading in mec-enabled iot. IEEE Internet of Things Journal 6, 4547–4555.

[A52] Heather, K., Shah, K.K., Venkatasubramanian, K.K., Cai, H., Hoyme, K., Seeberger, M., Wiechman, G., 2018. A novel authentication biometric for pacemakers, in: Proceedings of the 2018 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies, pp. 81–87.

[A53] Hegedus, I., Jelasity, M., 2016. Distributed differentially private stochastic gradient descent: An empirical study, in: 2016 24th Euromicro international conference on parallel, distributed, and network-based processing (PDP), IEEE. pp. 566–573.

[A54] HeydariGorji, A., Rezaei, S., Torabzadehkashi, M., Bobarshad, H., Alves, V., Chou, P.H., 2020. Hypertune: Dynamic hyperparameter tuning for efficient distribution of dnn training over heterogeneous systems, in: 2020 IEEE/ACM International Conference On Computer Aided Design (ICCAD), IEEE. pp. 1–8.

[A55] Huang, Y., Wang, W., Wang, H., Jiang, T., Zhang, Q., 2020. Authenticating on-body iot devices: An adversarial learning approach. IEEE Transactions on Wireless Communications 19, 5234–5245.

[A56] Hussain, F., Hussain, R., Hassan, S.A., Hossain, E., 2020. Machine learning in iot security: Current solutions and future challenges. IEEE Communications Surveys & Tutorials 22, 1686–1721.

[A57] Iqtidar Newaz, A., Sikder, A.K., Ashiqur Rahman, M., Selcuk Uluagac, A., 2019. Healthguard: A machine learning-based security framework for smart healthcare systems. arXiv e-prints , arXiv–1909.

[A58] Islam, M.S., Verma, H., Khan, L., Kantarcioglu, M., 2019. Secure real-time heterogeneous iot data management system, in: 2019 First IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications (TPS-ISA), IEEE. pp. 228–235.

[A59] Jiang, L., Tan, R., Lou, X., Lin, G., 2019. On lightweight privacy-preserving collaborative learning for internet-of-things objects, in: Proceedings of the International Conference on Internet of Things Design and Implementation, pp. 70–81.

[A60] Jourdan, T., Boutet, A., Bahi, A., Frindel, C., 2020. Privacy-preserving iot framework for activity recognition in personal healthcare monitoring. ACM Transactions on Computing for Healthcare 2, 1–22.

[A61] Kadiyala, S.P., Alam, M., Shrivastava, Y., Patranabis, S., Abbas, M.F.B., Biswas, A.K., Mukhopadhyay, D., Srikanthan, T., 2020. Lambda: Lightweight assessment of malware for embedded architectures. ACM Transactions on Embedded Computing Systems (TECS) 19, 1–31.

[A62] Kanagavelu, R., Li, Z., Samsudin, J., Yang, Y., Yang, F., Goh, R.S.M., Cheah, M., Wiwatphonthana, P., Akkarajitsakul, K., Wang, S., 2020. Two-phase multi-party computation enabled privacy-preserving federated learning, in: 2020 20th IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing (CCGRID), IEEE. pp. 410–419.

[A63] Kayode, O., Tosun, A.S., 2019. Analysis of iot traffic using http proxy, in: ICC 2019-2019 IEEE International Conference on Communications (ICC), IEEE. pp. 1–7.

[A64] Kennedy, S., Li, H., Wang, C., Liu, H., Wang, B., Sun, W., 2019. I can hear your alexa: Voice command fingerprinting on smart home speakers, in: 2019 IEEE Conference on Communications and Network Security (CNS), IEEE. pp. 232–240.

[A65] Khare, S., Totaro, M., 2020. Ensemble learning for detecting attacks and anomalies in iot smart home, in: 2020 3rd International Conference on Data Intelligence and Security (ICDIS), IEEE. pp. 56–63.

[A66] Krundyshev, V., 2020. Neural network approach to assessing cybersecurity risks in large-scale dynamic networks, in: 13th International Conference on Security of Information and Networks, pp. 1–8.

[A67] Laput, G., Zhang, Y., Harrison, C., 2017. Synthetic sensors: Towards general-purpose sensing, in: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, pp. 3986–3999.

[A68] Latif, S., Zou, Z., Idrees, Z., Ahmad, J., 2020. A novel attack detection scheme for the industrial internet of things using a lightweight random neural network. IEEE Access 8, 89337–89350.

[A69] Lee, T., Lin, Z., Pushp, S., Li, C., Liu, Y., Lee, Y., Xu, F., Xu, C., Zhang, L., Song, J., 2019. Occlumency: Privacy-preserving remote deep-learning inference using sgx, in: The 25th Annual International Conference on Mobile Computing and Networking, pp. 1–17.

[A70] Lee, W.H., Lee, R., 2016. Implicit sensor-based authentication of smartphone users with smartwatch, in: Proceedings of the Hardware and Architectural Support for Security and Privacy 2016, pp. 1–8.

[A71] Lee, Y.T., Ban, T., Wan, T.L., Cheng, S.M., Isawa, R., Takahashi, T., Inoue, D., 2020. Cross platform iot-malware family classification based on printable strings, in: 2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom), IEEE. pp. 775–784.

[A72] Lei, T., Qin, Z., Wang, Z., Li, Q., Ye, D., 2019. Evedroid: Event-aware android malware detection against model degrading for iot devices. IEEE Internet of Things Journal 6, 6668–6680.

[A73] Li, H., Xu, Z., Zhu, H., Ma, D., Li, S., Xing, K., 2016. Demographics inference through wi-fi network traffic analysis, in: IEEE INFOCOM 2016-The 35th Annual IEEE International Conference on Computer Communications, IEEE. pp. 1–9.

[A74] Li, X., Yan, F., Zuo, F., Zeng, Q., Luo, L., 2019. Touch well before use: Intuitive and secure authentication for iot devices, in: The 25th annual international conference on mobile computing and networking, pp. 1–17.

[A75] Lin, H., Garg, S., Hu, J., Wang, X., Piran, M.J., Hossain, M.S., 2020. Privacy-enhanced data fusion for covid-19 applications in intelligent internet of medical things. IEEE Internet of Things Journal .

[A76] Liu, C.H., Lin, Q., Wen, S., 2018. Blockchain-enabled data collection and sharing for industrial iot with deep reinforcement learning. IEEE Transactions on Industrial Informatics 15, 3516–3526.

[A77] Liu, X., Li, H., Xu, G., Liu, S., Liu, Z., Lu, R., 2020a. Padl: Privacy-aware and asynchronous deep learning for iot applications. IEEE Internet of Things Journal 7, 6955–6969.

[A78] Liu, Y., Yang, Y., Ma, Z., Liu, X., Wang, Z., Ma, S., 2020b. Pe-health: Enabling fully encrypted cnn for health monitor with optimized communication, in: 2020 IEEE/ACM 28th International Symposium on Quality of Service (IWQoS), IEEE. pp. 1–10.

[A79] Longo, S., Cheng, B., 2015. Privacy preserving crowd estimation for safer cities, in: Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers, pp. 1543–1550.

[A80] Lu, Y., Huang, X., Dai, Y., Maharjan, S., Zhang, Y., 2020a. Blockchain and federated learning for privacy-preserved data sharing in industrial iot. IEEE Transactions on Industrial Informatics 16, 4177–4186. doi:10.1109/TII.2019.2942190.

[A81] Lu, Y., Huang, X., Zhang, K., Maharjan, S., Zhang, Y., 2020b. Communication-efficient federated learning and permissioned blockchain for digital twin edge networks. IEEE Internet of Things Journal 8, 2276–2288.

[A82] Luo, C., Shrivastava, A., 2018. Arrays of (locality-sensitive) count estimators (ace) anomaly detection on the edge, in: Proceedings of the 2018 World Wide Web Conference, pp. 1439–1448.

[A83] Lyu, L., Bezdek, J.C., Jin, J., Yang, Y., 2020. Foreseen: Towards differentially private deep inference for intelligent internet of things. IEEE Journal on Selected Areas in Communications 38, 2418–2429.

[A84] Ma, X., Zhang, J., Ma, J., Jiang, Q., Gao, S., Xie, K., 2020. Do not perturb me: A secure byzantine-robust mechanism for machine learning in iot, in: 2020 International Conference on Networking and Network Applications (NaNA), IEEE. pp. 348–354.

[A85] Majumder, A.J., Izaguirre, J.A., 2020. A smart iot security system for smart-home using motion detection and facial recognition, in: 2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC), IEEE. pp. 1065–1071.

[A86] Malekzadeh, M., Clegg, R.G., Haddadi, H., 2018. Replacement autoencoder: A privacy-preserving algorithm for sensory data analysis, in: 2018 IEEE/acm third international conference on internet-of-things design and implementation (iotdi), IEEE. pp. 165–176.

[A87] Mao, B., Kawamoto, Y., Kato, N., 2020. Ai-based joint optimization of qos and security for 6g energy harvesting internet of things. IEEE Internet of Things Journal 7, 7032–7042.

[A88] Martins, P., Reis, A.B., Salvador, P., Sargento, S., 2020. Physical layer anomaly detection mechanisms in iot networks, in: NOMS 2020-2020 IEEE/IFIP Network Operations and Management Symposium, IEEE. pp. 1–9.

[A89] McGinthy, J.M., Wong, L.J., Michaels, A.J., 2019. Groundwork for neural network-based specific emitter identification authentication for iot. IEEE Internet of Things Journal 6, 6429–6440.

[A90] Meurisch, C., Bayrak, B., Mühlhäuser, M., 2020. Privacy-preserving ai services through data decentralization, in: Proceedings of The Web Conference 2020, pp. 190–200.

[A91] Mills, J., Hu, J., Min, G., 2019. Communication-efficient federated learning for wireless edge intelligence in iot. IEEE Internet of Things Journal 7, 5986–5994.

[A92] Min, M., Wan, X., Xiao, L., Chen, Y., Xia, M., Wu, D., Dai, H., 2018. Learning-based privacy-aware offloading for healthcare iot with energy harvesting. IEEE Internet of Things Journal 6, 4307–4316.

[A93] Mohammed, H., Hasan, S.R., Awwad, F., 2020. Fusion-on-field security and privacy preservation for iot edge devices: Concurrent defense against multiple types of hardware trojan attacks. IEEE Access 8, 36847–36862.

[A94] Msadek, N., Soua, R., Engel, T., 2019. Iot device fingerprinting: Machine learning based encrypted traffic analysis, in: 2019 IEEE wireless communications and networking conference (WCNC), IEEE. pp. 1–8.

[A95] Mudgerikar, A., Sharma, P., Bertino, E., 2020. Edge-based intrusion detection for iot devices. ACM Transactions on Management Information Systems (TMIS) 11, 1–21.

[A96] Neff, C., Mendieta, M., Mohan, S., Baharani, M., Rogers, S., Tabkhi, H., 2020. Revamp2t: Real-time edge video analytics for multicamera privacy-aware pedestrian tracking. IEEE Internet of Things Journal 7, 2591–2602. doi:10.1109/JIOT.2019.2954804.

[A97] OConnor, T., Mohamed, R., Miettinen, M., Enck, W., Reaves, B., Sadeghi, A.R., 2019. Homesnitch: behavior transparency and control for smart home iot devices, in: Proceedings of the 12th conference on security and privacy in wireless and mobile networks, pp. 128–138.

[A98] Osia, S.A., Shamsabadi, A.S., Sajadmanesh, S., Taheri, A., Katevas, K., Rabiee, H.R., Lane, N.D., Haddadi, H., 2020. A hybrid deep learning architecture for privacy-preserving mobile analytics. IEEE Internet of Things Journal 7, 4505–4518.

[A99] Pahl, M.O., Aubet, F.X., 2018. All eyes on you: Distributed multi-dimensional iot microservice anomaly detection, in: 2018 14th International Conference on Network and Service Management (CNSM), IEEE. pp. 72–80.

[A100] Pang, J., Huang, Y., Xie, Z., Han, Q., Cai, Z., 2020. Realizing the heterogeneity: A self-organized federated learning framework for iot. IEEE Internet of Things Journal 8, 3088–3098.

[A101] Phu, T.N., Hoang, L.H., Toan, N.N., Tho, N.D., Binh, N.N., 2019. Cfdvex: A novel feature extraction method for detecting cross-architecture iot malware, in: Proceedings of the Tenth International Symposium on Information and Communication Technology, pp. 248–254.

[A102] Pinheiro, A.J., de Araujo-Filho, P.F., Bezerra, J.d.M., Campelo, D.R., 2020. Adaptive packet padding approach for smart home networks: A tradeoff between privacy and performance. IEEE Internet of Things Journal 8, 3930–3938.

[A103] Qaddoura, R., Al-Zoubi, A., Almomani, I., Faris, H., 2021. A multi-stage classification approach for iot intrusion detection based on clustering with oversampling. Applied Sciences 11, 3022.

[A104] Qian, X., Chen, H., Jiang, H., Green, J., Cheng, H., Huang, M.C., 2020. Wearable computing with distributed deep learning hierarchy: a study of fall detection. IEEE Sensors Journal 20, 9408–9416.

[A105] Qu, Y., Gao, L., Luan, T.H., Xiang, Y., Yu, S., Li, B., Zheng, G., 2020. Decentralized privacy using blockchain-enabled federated learning in fog computing. IEEE Internet of Things Journal 7, 5171–5183.

[A106] Rahman, M.A., Hossain, M.S., Islam, M.S., Alrajeh, N.A., Muhammad, G., 2020. Secure and provenance enhanced internet of health things framework: A blockchain managed federated learning approach. Ieee Access 8, 205071–205087.

[A107] Ren, J., Dubois, D.J., Choffnes, D., Mandalari, A.M., Kolcun, R., Haddadi, H., 2019. Information exposure from consumer iot devices: A multidimensional, network-informed measurement approach, in: Proceedings of the Internet Measurement Conference, pp. 267–279.

[A108] Saeed, A., Ahmadinia, A., Javed, A., Larijani, H., 2016. Intelligent intrusion detection in low-power iots. ACM Transactions on Internet Technology (TOIT) 16, 1–25.

[A109] Sahoo, K.S., Puthal, D., 2020. Sdn-assisted ddos defense framework for the internet of multimedia things. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM) 16, 1–18.

[A110] Sarumi, O.A., Adetunmbi, A.O., Adetoye, F.A., 2020. Discovering computer networks intrusion using data analytics and machine intelligence. Scientific African 9, e00500.

[A111] Savazzi, S., Nicoli, M., Rampa, V., 2020. Federated learning with cooperating devices: A consensus approach for massive iot networks. IEEE Internet of Things Journal 7, 4641–4654.

[A112] Schiliro, F., Moustafa, N., Beheshti, A., 2020. Cognitive privacy: Ai-enabled privacy using eeg signals in the internet of things, in: 2020 IEEE 6th International Conference on Dependability in Sensor, Cloud and Big Data Systems and Application (DependSys), IEEE. pp. 73–79.

[A113] Shamshirband, S., Chronopoulos, A.T., 2019. A new malware detection system using a high performance-elm method, in: Proceedings of the 23rd international database applications & engineering symposium, pp. 1–10.

[A114] Shen, M., Ma, B., Zhu, L., Du, X., Xu, K., 2018. Secure phrase search for intelligent processing of encrypted data in cloud-based iot. IEEE Internet of Things Journal 6, 1998–2008.

[A115] Shen, M., Tang, X., Zhu, L., Du, X., Guizani, M., 2019. Privacy-preserving support vector machine training over blockchain-based encrypted iot data in smart cities. IEEE Internet of Things Journal 6, 7702–7712.

[A116] Shen, M., Wang, H., Zhang, B., Zhu, L., Xu, K., Li, Q., Du, X., 2020. Exploiting unintended property leakage in blockchain-assisted federated learning for intelligent edge computing. IEEE Internet of Things Journal 8, 2265–2275.

[A117] Shezan, F.H., Hu, H., Wang, J., Wang, G., Tian, Y., 2020. Read between the lines: An empirical measurement of sensitive applications of voice personal assistant systems, in: Proceedings of The Web Conference 2020, pp. 1006–1017.

[A118] Shi, C., Liu, J., Liu, H., Chen, Y., 2017. Smart user authentication through actuation of daily activities leveraging wifi-enabled iot, in: Proceedings of the 18th ACM International Symposium on Mobile Ad Hoc Networking and Computing, pp. 1–10.

[A119] Siboni, S., Sachidananda, V., Meidan, Y., Bohadana, M., Mathov, Y., Bhairav, S., Shabtai, A., Elovici, Y., 2019. Security testbed for internet-of-things devices. IEEE transactions on reliability 68, 23–44.

[A120] Skowron, M., Janicki, A., Mazurczyk, W., 2020. Traffic fingerprinting attacks on internet of things using machine learning. IEEE Access 8, 20386–20400.

[A121] Song, Y., Huang, Q., Yang, J., Fan, M., Hu, A., Jiang, Y., 2019. Iot device fingerprinting for relieving pressure in the access control, in: Proceedings of the ACM Turing Celebration Conference-China, pp. 1–8.

[A122] Sridharan, R., Maiti, R.R., Tippenhauer, N.O., 2018. Wadac: Privacy-preserving anomaly detection and attack classification on wireless traffic, in: Proceedings of the 11th ACM Conference on Security & Privacy in Wireless and Mobile Networks, pp. 51–62.

[A123] Stach, C., Steimle, F., 2019. Recommender-based privacy requirements elicitation-epicurean: an approach to simplify privacy settings in iot applications with respect to the gdpr, in: Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing, pp. 1500–1507.

[A124] Tabassum, A., Erbad, A., Mohamed, A., Guizani, M., 2021. Privacy-preserving distributed ids using incremental learning for iot health systems. IEEE Access 9, 14271–14283.

[A125] Thamilarasu, G., Odesile, A., Hoang, A., 2020. An intrusion detection system for internet of medical things. IEEE Access 8, 181560–181576.

[A126] Tien, C.W., Chen, S.W., Ban, T., Kuo, S.Y., 2020. Machine learning framework to analyze iot malware using elf and opcode features. Digital Threats: Research and Practice 1, 1–19.

[A127] Trevizan, B., Chamby-Diaz, J., Bazzan, A.L., Recamonde-Mendoza, M., 2020. A comparative evaluation of aggregation methods for machine learning over vertically partitioned data. Expert systems with applications 152, 113406.

[A128] Veličković, P., Lane, N.D., Bhattacharya, S., Chieh, A., Bellahsen, O., Vegreville, M., 2017. Scaling health analytics to millions without compromising privacy using deep distributed behavior models, in: Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare, pp. 92–100.

[A129] Wang, C., Kennedy, S., Li, H., Hudson, K., Atluri, G., Wei, X., Sun, W., Wang, B., 2020a. Fingerprinting encrypted voice traffic on smart speakers with deep learning, in: Proceedings of the 13th ACM Conference on Security and Privacy in Wireless and Mobile Networks, pp. 254–265.

[A130] Wang, H., Li, A., Shen, B., Sun, Y., Wang, H., 2020b. Federated multiview spectral clustering. IEEE Access 8, 202249–202259.

[A131] Wang, J., Amos, B., Das, A., Pillai, P., Sadeh, N., Satyanarayanan, M., 2018a. Enabling live video analytics with a scalable and privacy-aware framework. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM) 14, 1–24.

[A132] Wang, J., Zhang, J., Bao, W., Zhu, X., Cao, B., Yu, P.S., 2018b. Not just privacy: Improving performance of private deep learning in mobile cloud, in: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 2407–2416.

[A133] Wang, W., Seraj, F., Meratnia, N., Havinga, P.J., 2019. Privacy-aware environmental sound classification for indoor human activity recognition, in: Proceedings of the 12th ACM International Conference on PErvasive Technologies Related to Assistive Environments, pp. 36–44.

[A134] Xiong, J., Zhao, M., Bhuiyan, M.Z.A., Chen, L., Tian, Y., 2019. An ai-enabled three-party game framework for guaranteed data privacy in mobile edge crowdsensing of iot. IEEE Transactions on Industrial Informatics 17, 922–933.

[A135] Xu, D., Zheng, M., Jiang, L., Gu, C., Tan, R., Cheng, P., 2020a. Lightweight and unobtrusive data obfuscation at iot edge for remote inference. IEEE Internet of Things Journal 7, 9540–9551.

[A136] Xu, H., Li, J., Xiong, H., Lu, H., 2020b. Fedmax: Enabling a highly-efficient federated learning framework, in: 2020 IEEE 13th International Conference on Cloud Computing (CLOUD), IEEE. pp. 426–434.

[A137] Yan, Y., Pei, Q., Li, H., 2019. Privacy-preserving compressive model for enhanced deep-learning-based service provision system in edge computing. IEEE Access 7, 92921–92937.

[A138] Yang, L., Deng, H., Dang, X., 2020. Preference preserved privacy protection scheme for smart home network system based on information hiding. IEEE Access 8, 40767–40776.

[A139] Yang, L., Li, F., 2018. Cloud-assisted privacy-preserving classification for iot applications, in: 2018 IEEE Conference on Communications and Network Security (CNS), IEEE. pp. 1–9.

[A140] Yin, B., Yin, H., Wu, Y., Jiang, Z., 2020. Fdc: A secure federated deep learning mechanism for data collaborations in the internet of things. IEEE Internet of Things Journal 7, 6348–6359.

[A141] Yu, J., Fu, B., Cao, A., He, Z., Wu, D., 2018. Edgecnn: A hybrid architecture for agile learning of healthcare data from iot devices, in: 2018 IEEE 24th International Conference on Parallel and Distributed Systems (ICPADS), IEEE. pp. 852–859.

[A142] Zhang, K., Yiu, S.M., Hui, L.C.K., 2020. A light-weight crowdsourcing aggregation in privacy-preserving federated learning system, in: 2020 International Joint Conference on Neural Networks (IJCNN), IEEE. pp. 1–8.

[A143] Zhang, M., Chen, J., He, S., Yang, L., Gong, X., Zhang, J., 2019a. Privacy-preserving database assisted spectrum access for industrial internet of things: A distributed learning approach. IEEE Transactions on Industrial Electronics 67, 7094–7103.

[A144] Zhang, X., Chen, X., Liu, J.K., Xiang, Y., 2019b. Deeppar and deepdpa: privacy preserving and asynchronous deep learning for industrial iot. IEEE Transactions on Industrial Informatics 16, 2081–2090.

[A145] Zhao, S., Li, W., Zia, T., Zomaya, A.Y., 2017. A dimension reduction model and classifier for anomaly-based intrusion detection in internet of things, in: 2017 IEEE 15th Intl Conf on Dependable, Autonomic and Secure Computing, 15th Intl Conf on Pervasive Intelligence and Computing, 3rd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech), IEEE. pp. 836–843.

[A146] Zhao, Y., Zhao, J., Jiang, L., Tan, R., Niyato, D., Li, Z., Lyu, L., Liu, Y., 2020. Privacy-preserving blockchain-based federated learning for iot devices. IEEE Internet of Things Journal 8, 1817–1829.

[A147] Zheng, H., Hu, H., Han, Z., 2020. Preserving user privacy for machine learning: local differential privacy or federated machine learning? IEEE Intelligent Systems 35, 5–14.

[A148] Zhou, C., Fu, A., Yu, S., Yang, W., Wang, H., Zhang, Y., 2020a. Privacy-preserving federated learning in fog computing. IEEE Internet of Things Journal 7, 10782–10793.

[A149] Zhou, P., Zhong, G., Hu, M., Li, R., Yan, Q., Wang, K., Ji, S., Wu, D., 2019. Privacy-preserving and residential context-aware online learning for iot-enabled energy saving with big data support in smart home environment. IEEE Internet of Things Journal 6, 7450–7468.

[A150] Zhou, T., Shen, J., Ji, S., Ren, Y., Yan, L., 2020b. Secure and intelligent energy data management scheme for smart iot devices. Wireless Communications and Mobile Computing 2020.

[A151] Zhou, W., Li, Y., Chen, S., Ding, B., 2018. Real-time data processing architecture for multi-robots based on differential federated learning, in: 2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), IEEE. pp. 462–471.

[A152] Zhu, L., Tang, X., Shen, M., Du, X., Guizani, M., 2018. Privacy-preserving ddos attack detection using cross-domain traffic in software defined networks. IEEE Journal on Selected Areas in Communications 36, 628–643.

Bibliography

[1] , . Web appendix of the paper. https://drive.google.com/drive/folders/18UkAkO4DeM5WXLyG3syo4qbL-_4__3X9?usp=sharing.

[2] Albawi, S., Mohammed, T.A., Al-Zawi, S., 2017. Understanding of a convolutional neural network, in: 2017 International Conference on Engineering and Technology (ICET), Ieee. pp. 1–6.

[3] Aleisa, N., Renaud, K., 2016. Privacy of the internet of things: a systematic literature review (extended discussion). arXiv preprint arXiv:1611.03340 .

[4] Almagrabi, A.O., Bashir, A., 2021. A classification-based privacy-preserving decision-making for secure data sharing in internet of things assisted applications. Digital Communications and Networks .

[5] Ashton, K., et al., 2009. That 'internet of things' thing. RFID journal 22, 97–114.

[6] Batista, G.E., Prati, R.C., Monard, M.C., 2004. A study of the behavior of several methods for balancing machine learning training data. ACM SIGKDD explorations newsletter 6, 20–29.

[7] Bergmark, D., Phempoonpanich, P., Zhao, S., 2001. Scraping the acm digital library, in: ACM SIGIR Forum, ACM New York, NY, USA. pp. 1–7.

[8] Brereton, P., Kitchenham, B.A., Budgen, D., Turner, M., Khalil, M., 2007. Lessons from applying the systematic literature review process within the software engineering domain. Journal of systems and software 80, 571–583.

[9] Choi, W., Kim, J., Lee, S., Park, E., 2021. Smart home and internet of things: A bibliometric study. Journal of Cleaner Production 301, 126908.

[10] Ghosh, A., Chakraborty, D., Law, A., 2018. Artificial intelligence in internet of things. CAAI Transactions on Intelligence Technology 3, 208–218.

[11] Haller, S., 2010. The things in the internet of things. Poster at the (IoT 2010). Tokyo, Japan, November 5, 26–30.

[12] Khan, R., Khan, S.U., Zaheer, R., Khan, S., 2012. Future internet: the internet of things architecture, possible applications and key challenges, in: 2012 10th international conference on frontiers of information technology, IEEE. pp. 257–260.

[13] Kitchenham, B., Brereton, O.P., Budgen, D., Turner, M., Bailey, J., Linkman, S., 2009a. Systematic literature reviews in software engineering–a systematic literature review. Information and software technology 51, 7–15.

[14] Kitchenham, B., Pearl Brereton, O., Budgen, D., Turner, M., Bailey, J., Linkman, S., 2009b. Systematic literature reviews in software engineering – a systematic literature review. Information and Software Technology 51, 7–15. URL: https://www.sciencedirect.com/science/article/pii/S0950584908001390, doi:https://doi.org/10.1016/j.infsof.2008.09.009. special Section - Most Cited Articles in 2002 and Regular Research Papers.

[15] Kruger, C.P., Hancke, G.P., 2014. Benchmarking internet of things devices, in: 2014 12th IEEE International Conference on Industrial Informatics (INDIN), IEEE. pp. 611–616.

[16] Kuzlu, M., Fair, C., Guler, O., 2021. Role of artificial intelligence in the internet of things (iot) cybersecurity. Discover Internet of things 1, 1–14.

[17] Le, D.P., Meng, H., Su, L., Yeo, S.L., Thing, V., 2018. Biff: A blockchain-based iot forensics framework with identity privacy, in: TENCON 2018-2018 IEEE Region 10 Conference, IEEE. pp. 2372–2377.

[18] Le, T., Mutka, M.W., 2018. Capchain: A privacy preserving access control framework based on blockchain for pervasive environments, in: 2018 IEEE International Conference on Smart Computing (SMARTCOMP), IEEE. pp. 57–64.

[19] Li, F., Luo, B., Liu, P., 2011. Secure and privacy-preserving information aggregation for smart grids. International Journal of Security and Networks 6, 28–39.

[20] Lin, J., Yu, W., Zhang, N., Yang, X., Zhang, H., Zhao, W., 2017. A survey on internet of things: Architecture, enabling technologies, security and privacy, and applications. IEEE internet of things journal 4, 1125–1142.

[21] Lopez-Herrejon, R.E., Linsbauer, L., Egyed, A., 2015. A systematic mapping study of search-based software engineering for software product lines. Information and software technology 61, 33–51.

[22] Lu, R., Liang, X., Li, X., Lin, X., Shen, X., 2012. Eppa: An efficient and privacy-preserving aggregation scheme for secure smart grid communications. IEEE Transactions on Parallel and Distributed Systems 23, 1621–1631.

[23] Mair, C., Kadoda, G., Lefley, M., Phalp, K., Schofield, C., Shepperd, M., Webster, S., 2000. An investigation of machine learning based prediction systems. Journal of systems and software 53, 23–29.

[24] Mazhar, M.H., Shafiq, Z., 2020. Characterizing smart home iot traffic in the wild, in: 2020 IEEE/ACM Fifth International Conference on Internet-of-Things Design and Implementation (IoTDI), IEEE. pp. 203–215.

[25] Meneghello, F., Calore, M., Zucchetto, D., Polese, M., Zanella, A., 2019. Iot: Internet of threats? a survey of practical security vulnerabilities in real iot devices. IEEE Internet of Things Journal 6, 8182–8201.

[26] Mohamed, E., 2020. The relation of artificial intelligence with internet of things: A survey. Journal of Cybersecurity and Information Management 1, 30–24.

[27] Mohanta, B.K., Jena, D., Satapathy, U., Patnaik, S., 2020. Survey on iot security: Challenges and solution using machine learning, artificial intelligence and blockchain technology. Internet of Things 11, 100227.

[28] Nargesian, F., Samulowitz, H., Khurana, U., Khalil, E.B., Turaga, D.S., 2017. Learning feature engineering for classification., in: Ijcai, pp. 2529–2535.

[29] Ogonji, M.M., Okeyo, G., Wafula, J.M., 2020. A survey on privacy and security of internet of things. Computer Science Review 38, 100312.

[30] Osuwa, A.A., Ekhoragbon, E.B., Fat, L.T., 2017. Application of artificial intelligence in internet of things, in: 2017 9th international conference on computational intelligence and communication networks (CICN), IEEE. pp. 169–173.

[31] Ramirez, A., Romero, J.R., Simons, C.L., 2018. A systematic review of interaction in search-based software engineering. IEEE Transactions on Software Engineering 45, 760–781.

[32] Sculley, D., Holt, G., Golovin, D., Davydov, E., Phillips, T., Ebner, D., Chaudhary, V., Young, M., 2014. Machine learning: The high interest credit card of technical debt .

[33] Singh, R.P., Javaid, M., Haleem, A., Suman, R., 2020. Internet of things (iot) applications to fight against covid-19 pandemic. Diabetes & Metabolic Syndrome: Clinical Research & Reviews 14, 521–524.

[34] Strous, L., von Solms, S., Zúquete, A., 2021. Security and privacy of the internet of things. Computers & Security 102, 102148.

[35] Thierer, A., Castillo, A., 2015. Projecting the growth and economic impact of the internet of things. George Mason University, Mercatus Center, June 15.

[36] Thilakarathne, N.N., 2020. Security and privacy issues in iot environment. International Journal of Engineering and Management Research 10.

[37] Waheed, N., He, X., Ikram, M., Usman, M., Hashmi, S.S., Usman, M., 2020. Security and privacy in iot using machine learning and blockchain: Threats and countermeasures. ACM Computing Surveys (CSUR) 53, 1–37.

[38] Wohlin, C., 2014. Guidelines for snowballing in systematic literature studies and a replication in software engineering, in: Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering, Association for Computing Machinery, New York, NY, USA. URL: https://doi.org/10.1145/2601248.2601268, doi:10.1145/2601248.2601268.

[39] Wohlin, C., 2016. Second-generation systematic literature studies using snowballing, in: Proceedings of the 20th International Conference on Evaluation and Assessment in Software Engineering, pp. 1–6.

[40] Wu, H., Han, H., Wang, X., Sun, S., 2020. Research on artificial intelligence enhancing internet of things security: A survey. Ieee Access 8, 153826–153848.

[41] Xu, Y., Zhou, Y., Sekula, P., Ding, L., 2021. Machine learning in construction: From shallow to deep learning. Developments in the Built Environment 6, 100045.

[42] Ziegeldorf, J.H., Morchon, O.G., Wehrle, K., 2014. Privacy in the internet of things: threats and challenges. Security and Communication Networks 7, 2728–2742.