Proceedings of the 53rd Hawaii International Conference on System Sciences | 2020

A Literature Survey on Smart Toy-related Children's Privacy Risks

Marcelo Fantinato,
Otávio de Paula Albuquerque
University of São Paulo, Brazil
m.fantinato@usp.br,
otavioalbuquerque@usp.br

Anna Priscilla de Albuquerque,
Judith Kelner
Federal Univ. of Pernambuco, Brazil
apa@cin.ufpe.br,
jk@cin.ufpe.br

Benjamin Yankson
OntarioTech University, Canada
benjamin.yankson@uoit.ca

Abstract

Smart toys have become popular as technological solutions offer a better experience for children. However, the technology used increases the risks to children's privacy, which does not seem to have become a real concern for toy makers. Most researchers in this domain are vague in defining their motivations due to lack of an expert survey to support them. We conducted a literature survey to find papers on smart toy-related children's privacy risks and mitigation solutions. We analyzed 26 papers using a taxonomy for privacy principles and preserving techniques adapted from the IoT context. Our analysis shows that some types of risks received more attention, especially (a) confidentiality, (b) use, retention and disclosure limitation, (c) authorization, (d) consent and choice, (e) openness, transparency and notice and (f) authentication. As for solutions, few were effectively presented; the vast majority related to data restriction – (a) access control and (b) cryptographic.

1. Introduction

Smart toys can collect and process data at real time favored by miniaturization and lower costs of processing circuits. They can connect with other toys or gadgets such as smartphones and access to mobile and cloud services hence permeating the domains of the Internet of Things (IoT). Smart toys use short- or long-range wireless communications protocols for the acquisition, computing and transfer of child user's personal and non-personal information. The National Institute of Science and Technology (NIST) defines Personally Identifiable Information (PII) as any information about an individual maintained by an organization that can be used to distinguish or trace an individual's identity, including any information linkable to an identifiable individual.

Smart toys use PII for several reasons. For example, pervasive location-based applications, such as Niantic's Pokemon GO, collect Global Positioning System (GPS) data for outdoor playing. Smart toys such as Cognitoys' Dino and Mattel's Hello Barbie as well as companion robots such as Asus' Zenbo offers conversation functions that use speech recognition and Artificial Intelligence (AI) to establish and maintain a reasoning-based dialogue with the child. Thus, they collect, share and store PII (i.e., the child's voice recordings).

Similar to other IoT products, ensuring the privacy of the information collected is a critical challenge to smart toys [1]. The concern is even greater in this case as PPI is collected from underage users. Smart toy solutions can be opportunities for many risks to children's privacy, with PII becoming a target for unauthorized data collection, storage or disclosure [2]. Smart toy makers should consider threats and vulnerabilities to mitigate privacy risks before marketing new products that could compromise children's privacy. Smart toy solutions should at least ensure compliance with the information privacy and security regulations related to child users' PII.

As this is a emergent research area, most researchers in this domain are vague in defining their motivations due to lack of an expert survey to support them. To contribute with this research area, this paper presents an analysis of the scientific papers that address some aspect related to smart toy-related children's privacy. We conducted a systematic search[1][2][3][4] and found 26 papers published from 2009 to 2019. We divide our analysis into two parts: first, we present a content analysis of these papers that allows us to understand the evolution of the topic over the years; then, we present an axial analysis of how these papers deal with different aspects of privacy risks. Since authors from interdisciplinary areas do not always follow a common and well-defined nomen-

[1] **String**: "(privacy OR private OR security OR secret OR confidential OR safety OR vulnerable OR protect OR preserve OR reliability OR sensitive) AND ((smart OR connected OR clever OR intelligent OR anthropomorphic OR humanoid OR humanlike OR interactive OR cognitive OR social OR educational OR internet OR IoT OR online) W/5 (toys OR playthings OR dolls OR barbie))". Stems were used to help find relevant papers (e.g., 'vulnerab*' was used for 'vulnerable').

[2] **Databases**: Scopus and Web of Science.

[3] **Inclusion criterion**: the paper discusses children's privacy issues when using smart toys and possibly solutions to such privacy issues.

[4] **Exclusion criteria**: the paper is not fully published in English, published as an abstract, or not a peer reviewed scientific work.

URI: https://hdl.handle.net/10125/63921
978-0-9981331-3-3

clature for terms related to privacy and security risks, threats and vulnerabilities, we use *risk* here as a general term to refer to all these related terms.

This paper introduces children's privacy concepts, compares related work, presents our content analysis followed by the axial analysis and ends with conclusion.

2. Children's Privacy

IoT-related products offer risks to the user's privacy as they can collect, store and manage PII, including sharing users' PII with third parties. Governmental entities are worldwide taking efforts to regulate data privacy protection rules to be employed in companies and organizations. For example, the General Data Protection Regulation (GDPR) [3] gives control to individuals over their personal data and simplifies the regulatory environment for international business by unifying the regulation within the Europe Union. GDPR defines *Personal Data* (PD) as anything containing directly or indirectly compromising information that can expose user privacy and allow the singling out of individual behavior. Despite small differences between the concepts of NIST's PII and GDPR's PD, both refer to similar contexts distinguished only by the regulations they govern. PII is the most widely used term in North America; thus, in this study, we use the term PII.

Privacy protection aims to enable parents to control the privacy of their children by specifying their toy-related privacy preferences, assuming there is an accurate privacy policy with which the toy complies through a privacy protection mechanism attached to the toy [4]. Other examples of privacy regulations are the Canadian Personal Information Protection and Electronic Documents Act (PIPEDA) [5] and Brazilian General Law of Protection of Personal Data (GLPPD) [6]. Further, the American Federal Trade Commission (FTC) offers specific PII regulations to children's privacy, governed by the Children's Online Privacy Protection Act (COPPA) [7]. COPPA protects the online privacy of children under 13 and states that a child's PII cannot be collected without parental consent. In 2010, an amendment to COPPA further elaborated that PII includes geolocation information, photographs and videos [8].

Those privacy regulations became quality references for protecting the smart toy-related children's privacy since consumers, particularly parents, are aware of these privacy risks. Thus, addressing concrete solutions to mitigate those risks is essential to support the promising future of smart toy markers. Privacy risks can compromise the future for smart toys in the consumer's market, such as by affecting purchase intent and the distribution of these products to a worldwide audience [9, 10].

3. Related Work

With a significant number of works addressing privacy and security concerns regarding smart toys, no systematic study has been published to date that analyzes. Some literature reviews cover smart toy-related topics without covering any privacy or security concerns. Two of them address the application of smart toys in a specific context: clinical treatment [11, 12]. Another review identifies the different types of interactions between users and smart toys, classifying them into different categories, genres, and setups based on their play and interface features [13]. Secondary studies on privacy and security in broader research areas such as IoT [14], Bring Your Own Device (BYOD) [15] and big data [16] fail to address smart toys as this is a very specific subject and they refer to it in more general aspects.

As a contribution, we conducted a systematic search [17, 18] of scientific papers addressing *smart toy-related children's privacy risks* and, potentially, *mitigation solutions*. As a result, we found 26 papers. To analyze the coverage of these works on smart toy-related children's privacy risks, we followed two steps: content analysis and axial analysis. Content analysis presents a chronological and contextualized analysis of each paper. Through this analysis, it is possible to follow and understand the evolution of the research that has been carried out in this area. Second, axial analysis classifies the 26 papers based on the taxonomy proposed by Loukil *et al.* [14], considering smart toys as a type of IoT.

4. Content analysis

We present here a chronological and contextualized analysis of each of the 26 papers. Through this analysis, it is possible to follow and understand the evolution of the research that has been carried out in this area.

The first work, robot toys. Denning *et al.* [19] found issues related to children's privacy risks associated with the use of smart robot toys that present characteristics such as mobility, dexterity, sensing and output capability, and wireless communication. The authors conducted non-extensive experiments with two robot toys to identify privacy and security vulnerabilities and classified the findings into: *remote identification*, *eavesdropping*, *unauthorized operational notifications*, *unauthorized remote control* and *compromised network security*. They found several potential risks that included *spying*, *vandalism* and *psychological attacks*. Spying is the most serious and refers to the possibility of an attacker compromising one of the robot toys and leveraging its built-in video camera to spy on a child in their bedroom. Vandalism refers to damage of fragile objects in the surrounding environment whereas psychological attacks may include for example hijacking audio capabilities to cause distress to a child or remotely controlling the robot to arrange objects on the ground in a threatening symbol. The authors proposed 14 design questions for the development of household robots, including five related to security and privacy questions.

A new type of threat: online pedophiles. Yong *et al.* [20] were the only ones to address the threat of online pedophiles in the smart toy context. They were concerned with small children being exploited online when using these toys as they are technically vulnerable to virtual *eavesdropping* if not secure in terms of remote access, either by close proximity, by the internet connection or by VoIP. The authors provide four scenarios of potential threats in which smart toys can help pedophiles. For risk mitigation, they propose technological solutions for toy companies and consumers that can be categorized into *VoIP protection strategies*, *remote control connection*, *wireless connection*, *parental control strategies*, *camera protection strategies*, *shared risk perspective* and *general protection strategies*. As in the previous work [19], these are very high-level guidelines to help the work of smart toy developers.

Stealing data... and the toy. Pleban, Band and Creutzburg [21] found two major security-related technical vulnerabilities: *(1)* full access control to a toy drone can be gained with root shell over an unencrypted Wi-Fi network and *(2)* data on a USB device connected to a toy drone can be accessed remotely by people connected to the FTP server. Possible threats are, respectively, theft of the toy itself or of private data. As a solution, they presented technical steps to perform the cross-compilation of the programs involved, which require technical knowledge. Although these are concrete steps and not just general guidelines, they represent not a generic solution for any similar toy with similar issues.

Hello Barbie, not keeping secrets. Jones and Meurer [22] were concerned with *data disclosure* and conducted experimental works to evaluate the toy's abilities in keeping data privacy, considering also parental supervision. They did not identify technical issues; but they concluded that the toy is not designed for keeping secrets, neither child's from parents nor parents' from third parties. Simulating for children that the toy keeps their secrets from their parents can be a *psychological* risk. While this is not exactly a problem as it depends on how parents and society view this issue, *design decisions supported by children's specialists* could be made to improve this point. The authors propose *greater transparency by the toy maker*, offering *choice opportunities for the parents*. Later [23], they highlighted the need to adapt privacy policies to the IoT world.

Hacking Hello Barbie. Holloway and Green [24] analyzed how the news and public discourses describe risks of smart toys and studied consumer concerns. They were motivated by general risks such as: *children's dataveillance* (i.e., corporate and government surveillance of children's activities and encroachments upon their data privacy and security), *illegitimate access to children's information* through *hacking of internet-connected toys*, *geolocational tracking of children* and *unauthorized remote control of toys*' recording and speaking devices. They evaluated VTech's Learning Lodge and Hello Barbie. They analyzed several media reports, including the hacking of Hello Barbie by a security expert. Concerns raised include the lack of regulatory direction relating to children's data privacy and security. They raised the need for more research to encourage the creation of *new legislation* with *effective policies and possible sanctions* for toy makers. However, as in most previous works, they focus on the discussion of privacy risks, without proposing effective solutions.

Hello Barbie, indiscreet subjects. Hung *et al.* [1] analyzed the Hello Barbie's *privacy policies* and features of its *parental control* tool. They identified a *poor parental control* over the collection and retention of sensitive user data. They raised privacy concerns from the predetermined phrases the toy uses to talk to children, which might encourage them to *reveal private information indirectly*. Possible leaks of sensitive information to third parties can cause a *physical safety* risk to the child. They suggest as future work to address special *privacy requirements for child* users because of their vulnerability and propose a *framework that can provide adequate parental control* for the child users' parents.

Privacy requirements for smart toys. Hung, Fantinato and Rafferty [8] discussed *privacy requirements* targeted at child using smart toys. The motivation is the risk of *physical safety* for child users caused, for example, by child predators. The requirements are grouped into: *end-user requirements for children*, *parental control*, *privacy laws and regulations*, *health Canada's safety requirements for children's toys and related products*, *industry guidelines and best practices*. The authors again pointed out the need for a standardized children's protection *framework for parental control* as future work. While they are high-level requirements, this was the first paper to contribute more concretely to the context of smart toy-related children's privacy.

Conceptual model for privacy rules. Rafferty *et al.* [25] proposed a *conceptual model* to enable parents (or legal guardians) to provide consent through *access rules* that define that their child's data to be shared according to their *privacy preferences*. The proposed model involves concepts of smart toy, mobile services, device, location and guidance with related privacy entities: purpose, recipient, obligation and retention for smart toys. The model allows parents to create privacy rules and receive acknowledgments regarding their children's privacy-sensitive data. Rules created by the parents can include, for example, the rights to request the restriction of the data usage, to access and inspect the stored data, to request data deletion and to be notified about data sharing. The conceptual model includes the access control model (core and extended), a policy rule creation process and an access control decision process. The proposal in this work is a concrete contribution that could be taken as the basis for the toy industry – either naturally or through legislation – to support the development of appropriate parental control tools.

Toys that listen... and children unaware. Semi-structured interviews were conducted by MyReynolds *et al*. [26] with parent-child pairs interacting with two smart toys. They were motivated by the news reports on parent's and children's *data storage*, the *hacking of data* that allowed *children identification and location* and issues with toy's *privacy policies*. The main findings were: *children were often unaware* that others might be able to hear what was said to the toy whereas parents were concerned with what *data would be retained* by the company. Based on their findings, the authors proposed a *set of privacy-related recommendations*: *(1)* for toy makers – producing toys that make *children aware of recording*, using less remote processing to *avoid the need for recording*, *deleting recording* after use and considering *ethical aspects* when offering resources; and *(2)* for lawmakers – helping children, parents and toy makers become more *aware of privacy issues*, strengthen the *existing children's privacy protections* such as COPPA and addressing the same issues with *other similar internet-connected devices* also used by children such as Siri, Amazon Echo and Alexa.

Security requirements for smart toys. Carvalho and Eler [27, 28, 29] addressed *smart toy-related security* rather than privacy, assuming security needs should be guaranteed first so privacy can also be ensured. First, they used a well-defined security requirements engineering approach for guidance. They identified 12 security issues, 15 security threats and 20 security requirements associated with smart toys [27]. The security issues were raised from COPPA and PIPEDA. The security threats are grouped into *spoofing*, *tampering*, *repudiation*, *information disclosure*, *denial of service* and *elevation of privilege*. The security requirements address security concepts mainly related to *confidentiality*, *integrity*, *availability*, *privacy*, *non-repudiation* and *authenticity*. Later [28, 29], they related GDPR with extra security issues and security requirements. They also proposed a *testing strategy* to evaluate security issues and requirements, which envisions the execution of different testing types at different stages of development – implementation, verification and release – of smart toys. They evaluated flaws reported in the news for five smart toys. They concluded that many of the reported breaches would have been avoided if these security issues and requirements had been addressed by toy makers.

Hello Barbie... the duty to report! While some researchers consider recording and listening to children's conversations with Hello Barbie and similar smart toys an ethical issue as it can pose a privacy and trust violation, Moini [30] considers that the toy (in fact, the company responsible for it) has an obligation to do so. For the author, computer technicians, service providers and toy companies have a *duty to monitor and track suspicious recordings* on critical matters such as sexual abuse. However, according to her studies, existing children's *privacy regulations do not guarantee such an obligation* to companies, not even COPPA. Thus, she proposed an *amendment to COPPA* to include this obligation. She presented a full and detailed draft proposal of this amendment for COPPA.

Dino asking to open the house door? Valente and Cardenas [31] conducted experiments with Dino, and two problems were found. On privacy, they showed it was possible to *eavesdrop on the communication* between the toy and the cloud by decryption. As for security, they showed it was possible to *inject audio into the toy* so it could be played for the child. Audio injection causes risks, including *physical safety* (e.g., the toy can ask the child to open the house door) and *psychological safety* (e.g., the toy can bully the child). Both the eavesdropping and audio injection attacks are performed by injecting RTP (Real-time Transport Protocol) traffic directly into the network traffic communication to a Dino toy, which cannot be avoided due to the protocol nature. A solution is to change RTP by SRTP (Secure Real-time Transport Protocol), which provides message authentication and integrity checking, and protection against replay attacks on the RTP traffic. The authors reported collaborating with the toy maker to correct the issue.

Which smart toy is less vulnerable? Mahmoud *et al*. [32] proposed an analytical framework to evaluate smart toy privacy practices with 17 privacy-sensitive criteria split into five categories: application authenticity and permissions; privacy policy documentation; terms of user documentation; information collection; and information storage, sharing and protection. With this framework, they analyzed 11 smart toys. The evaluation was based on publicly available privacy policies and terms of use documentation as well as static analysis of toys' companion apps. They reported: *(1)* excessive collection of children's private information, *(2)* incomplete/lack of information on data storage location and legal compliance and *(3)* over-privilege with the potential to leak child's private information in most apps. Some of the 17 features evaluated were not found for any of the toys, such as *reasonable-permissions*, *reasonable-PII-collection* and *data-storage-location*. Only two of the toys offer more than half of the evaluated features: Hello Barbie (11 full) and Anki Cozmo (8 full and 3 partial). The scenario may be worse considering a dynamic evaluation since some reported features may not have been implemented or correctly implemented.

Is any smart toy getting a badge A? Haynes *et al.* [33] proposed an *evaluation framework* that presents a set of standards that should be fulfilled as *privacy and security prerequisites* for a toy to be marketed. It is based on the NIST documentation and includes six criteria related to *product packaging notification*, *secure communication and data storage*, *ongoing vulnerability scanning*, *independent security audits*, *no default passwords* and *secure remote access and updates*. These authors innovate by suggesting a consumer-oriented badge for these toys, which would be classified as A+, A, B+, B or C, depending on the number of met criteria. The authors considered issues in smart toys reported in the media news. They assessed seven smart toys. None of them received A+ or A, and only the last two were the best ranked with B+. In addition, no smart toy met the *independent security audits* criterion.

Fictitious toy... real risks. Demetzou, Böck and Hantter [34] followed the *privacy by design* approach to identify smart toy-related children's privacy risks. The authors used a generic model that identified risks associated with: *data misuse*; *data leakage*; *lack of control* over own data; algorithmic classification of children in *incorrect profiles*; *surveillance/eavesdropping* for blackmail, kidnapping, pedophilia etc.; and *parents checking on* their children. They advocated following the *data minimization* principle in line with the GDPR's *purpose limitation* principle to mitigate those risks and discussed technical solutions to be adopted to achieve these two principles. Even with data minimization, there is still a need to reduce the risk of attacks with security measures, such as: *(1)* using a *Trusted Third Party* (TTP) to prevent unauthorized *targeted advertisements*, *(2) anonymizing communication endpoints* with anonymizing services to prevent *tracking the location* of the toy and hence the child; *(3)* using robust *encryption* protocols to prevent external parties from *accessing the transmitted data*; *(4)* using robust *authentication* to prevent even those with the correct application from gaining *unauthorized access*; and *(5) encrypting all stored data* in the toy to prevent data from being *accessed in case of theft*.

Would you buy a Hello Barbie for your child? Fantinato *et al.* [10] studied how consumers in emerging countries, whose markets still do not have access to most smart toys, see privacy risks related to Hello Barbie. They conducted a survey with 118 potential consumers in Brazil and Argentina. The authors assessed the *perceptions of consumers on the innovativeness, risks and benefits* of the *toy's conversation function* as well as the participants' purchase intents. The results showed participants from both countries rated smart toys as equally *innovative and risky*, which shows that these potential consumers are already concerned on data privacy, although they were not informed of such problems through the news. The results also showed that Brazilians have a more positive perception of the toy and hence higher purchase intent than the Argentinians. Such a difference may be explained by small cultural differences between the two countries, such as higher risk tolerance, although both are neighbors and have a similar economy. The authors concluded their paper with suggestions for toy makers to reduce privacy issues by *improving parental control tools using data mining*.

Should parents be the only responsible for protecting the child's PII? Kshetri and Voas [35] presented a brief analysis of children's privacy risks in this scenario. First, they discussed several reasons that may make it more interesting to *steal children's data and identity* than adults' ones. They then examined *regulatory initiatives* to protect children from the harmful effects attributed to smart toys. Finally, they analyzed initiatives by government agencies and private watchdog groups to *alert the public* about the privacy and security risks of smart toys. The authors concluded that: *(1)* makers do not seem to have the capability, resources and motivation to strengthen the security level of their smart toys; *(2)* regulatory efforts to address this issue are incipient, a gap that various government agencies and consumer watchdog groups are trying to fill; and *(3)* for now, *parents are solely responsible* for monitoring the use of smart toys and protecting children's data.

Does everything need to go to the cloud? Espinosa-Aranda *et al.* [36] were concerned with the *emotional recognition* functionality of smart toys. This functionality typically needs to send the child's captured data for processing in the cloud, leading to different types of privacy risks. They proposed a technique using deep learning so that the *processing is carried out locally* in the toy. The authors argue that the emotional state of a child or infant is a piece of private information that must be preserved. In addition to not being sent to the cloud for processing, the information can be *stored encrypted in the toy* and accessed through *secure password access mechanisms*. Besides, they conducted a pilot experiment to show the feasibility of the proposed solution and were satisfied with the results.

How good are the privacy policies of toy companies? Chowdhury [37] investigated the *privacy policies* of 15 connected or smart toys. The work aimed to answer 16 questions on privacy and security related to *security level*, *password*, *encryption*, *user control*, *user data removal*, *user awareness*, *non-essential data collection*, *third-party access protection*, the existence of *privacy policy document*, *notification of a change in the privacy policy*, among others. Most toy makers *do not mention or clarify* in their privacy policy documents how their toys protect users' security and privacy. Thus, they recommended *more explicit privacy policies* so parents can have transparency to decide whether to buy smart toys for their children. A Frequently Asked Questions *(FAQ) with privacy items* should accompany privacy policy documents to make it easier to see what and how information is collected, used and disclosed. In addition, toy makers should reconsider the need to share user data with remote, potentially hacked databases.

Privacy vs. personalization vs. advertising culture. Smith and Shade [38] analyzed *data surveillance and commercialization* practices in digital playgrounds that trouble *children's and parent's privacy* rights. The analysis was based on the *privacy sweep* method, used to recreate the consumer experience and assess the transparency of PII practices against some indicators. Three indicators were adopted: *(i)* What *political economy* elements surround the implementation of the digital playground and its privacy policy? *(ii)* Do *privacy communications* adequately explain how PII is collected, used and disclosed? *(iii)* Are *users fully informed* about how PII collected by the device is stored and safeguarded? The authors presented five reflections: *(1)* parents are tasked to act as data proxies and supervisors for children, *(2)* personalization is proposed as the benefit for disclosing data, *(3)* advertising and promotional culture is present in children's digital playgrounds, *(4)* the algorithms that power digital playgrounds remain opaque for parents and *(5)* there are patterns in data stewardship for digital playgrounds that parents may wish to examine.

Which smart toy is the most vulnerable? Shasha *et al.* [2] proposed a *vulnerability taxonomy* for potential security and privacy flaws that can lead to private information leakage or allow toy remote control for malicious use. The taxonomy classifies the attacks into 15 types, split into three levels of proximity – *physical*, *nearby* and *remote*. They analyzed 11 smart toys, w.r.t. network traffic analysis, reverse engineering and code analysis of companion apps. They detected an *excessive collection of unique IDs* that facilitate *tracking of children* on different services or platforms and *sending children's PII to unauthorized entities*. Most toys expose children to threats that can be exploited by physical, nearby or remote attackers. Only two smart toys showed no vulnerability to any type of attack. *Insecure bluetooth practice* is the type of attack for which smart toys are most vulnerable, identified for six toys; followed by *unauthorized configuration physical*, *unencrypted communication channels* and *URL redirect*. The authors present recommendations to smart toy makers: *fixing vulnerabilities*, *limiting data collection* and *improving storage and communication security*, *securing Bluetooth and WiFi connections*, granting parents with *fine-grained access control*, preparing the toy *end-of-life* and being cautious with *advertisement and tracking*.

Formalizing the data flow and its privacy risks. Yankson *et al.* [39] are interested in *privacy-sensitive points* in the smart toys *data flow*. They argue that current data flow modeling techniques do not have adequate elements to address privacy in smart toys transactions. This limitation of modeling techniques impairs the risk analysis during the smart toy development and can increase the chance of threats to the children's privacy and physical safety. As a solution, the authors propose to use *colored Petri nets* to support it. Thus, they modeled simulations to show that it is possible to *include elements on privacy* in the data flows. The authors concluded that using this modeling technique allows for flow analysis to minimize the risk breaching the privacy of the child's data with the use of the toy being developed. Smart toy developers can use this technique to *analyze the privacy-sensitive data flow* for the toy being developed.

Not even basic best practices for security and privacy are followed. Chu, Apthorpe and Feamster [40] conducted three anonymous case studies on security and privacy risks on three smart toys. They analyzed *network and application vulnerabilities* through static analysis (with publicly available documentation analysis and application binary decompilation) and dynamic analysis (with network monitoring through experimentation). They uncovered a number of *vulnerabilities undisclosed to the public* that *violate COPPA* and the *toys' individual privacy policies*, which included *security flaws* in network communications with first-party servers. These were: *lack of data encryption*, *lack of authentication for accessing sensitive user information*, *POST token reuse*, *sensitive user information in crash reports to third parties* and *secret keys in source code constant files*. Both COPPA and individual privacy policies of the toys are violated with such vulnerabilities. Moreover, they had an indicative that *toy makers are not following best security and privacy practices*, not even basic ones. Finally, they point to the need for automated analysis to identify privacy and security vulnerabilities for smart toys.

5. Axial analysis

An axial analysis allowed us to evaluate both the context addressed by current studies and a view of possible future research topics, as presented in this section.

5.1. Current scenario

Axial coding was performed for the 26 papers to complement the content analysis. We adapted the taxonomy proposed by Loukil *et al.* [14], using four facets related to privacy and security from the original seven as our study has a more specific aim. In addition, we grouped *security properties* and *security requirements* as a single sub-facet of *information security* to make the taxonomy more suitable for our purposes. Table 1 shows the classification of the findings, based on the adapted taxonomy, that considers the concerns expressed by their authors according to the terms used by them. The classification used in our study has two facets. The former (*ISO privacy principles*) refers to privacy risks involved in using smart toys, including threats and vulnerabilities. The latter (*privacy-preserving techniques*) refers to solutions to mitigate such privacy risks.

As relied by Loukil *et al.* [14], we assumed smart toy makers should adhere to *ISO privacy principles* so as not to pose risks to children. There are 11 privacy principles, with one of them – *information security* – split into six other principles. This second classification level is important as information security is one of the most important aspects to ensure information privacy. As for *privacy-preserving techniques*, five main techniques are considered: *noise addition* and *anonymization* (related to *data perturbation*) and *access control*, *cryptographic* and *blockchain*-based (related to *data restriction*).

Axial analysis consisted of identifying in the papers which were presenting concerns on which of the principles in the adopted taxonomy. Hence, we searched for keywords that could show the concern of the authors regarding each subject of the taxonomy. Each term was accounted for in the data extraction for a given paper only when a clear and explicit concern manifestation was found in the paper on the corresponding term. The data shown in Table 1 reflect the researchers' primary concerns on children's privacy risks in using smart toys as well as on the types of solutions to mitigate those risks. Some of these risks and solutions are only briefly mentioned in some papers, with shallow coverage, while others have deep coverage. Thus, for a better characterization, our analysis considered two possible levels of coverage for each topic in each paper, considering whether the paper addresses a particular risk or solution in a *Deep (**D**)* or *Shallow (**S**)* way.

The analysis of this data is split into two parts: *(i)* by principles or techniques addressed in the different papers and *(ii)* by papers addressing different principles or techniques. For both cases, the numbers of D and S are summed but with different weights: D was considered as 1.0 point and S was considered as 0.5 point.

For the analysis by the 16 *ISO privacy principles*, one can verify that some principles have raised more concern of the authors than other principles. The six types of risk most commonly addressed in the papers published so far, with more than 10.0 points, are related to the following principles: *(i) confidentiality* (20.5 points), referring to child's data protection from unauthorized accesses, disclosures and processes – one of the *information security* sub-principles; *(ii) use, retention and disclosure limitation* (18.0 points), referring to the need to limit to the maximum extent possible the use, retention and disclosure of the child's data; *(iii) consent and choice* (17.0 points), referring to the need for consent of the child's parent for any type of operation to be performed with the child's data, including offering different options for them; *(iv) authorization* (16.5 points), referring to provide permissions towards the child's data – one of the *information security* sub-principles; *(v) openness, transparency and notice* (15.5 points), referring to the need for broad transparency by the smart toy maker so that the child' parent has extensive knowledge of what is involved in the operations with the child's data; and *(vi) authentication* (13.0 points), referring to ensure that a claimed characteristic of an entity is correct; one of the *information security* sub-principles.

Through this axial analysis of the 26 studies, one can easily notice a pattern of concern by the authors, who express a clear interest in basic privacy risks related to smart toys. The top six risks are related to the basic principles to ensure children's privacy while using smart toys. *Confidentiality* summarizes and represents the whole concern on privacy as the ultimate intention is that confidentiality is protected, which naturally leads also to *authorization* and *authentication*. The other three principles are correlated and linked to the need for toy makers and companies to be transparent, responsible and careful with their data usage policies. While these basic issues are still the subject of concern by researchers, others are still timidly treated, although also important to the ultimate goal.

Regarding the other ten principles, *collection limitation* stands out (9.0 points). A reasonable number of authors expressed concerns on the amount of data collected by toy companies larger than necessary. On the other hand, *data minimization*, the extreme of the principle of collection limitation, was the least mentioned (1.5 points). Minimizing data means that there should be a

Paper reference	Year	ISO privacy principles																Privacy-preserving techniques					TOTAL
											Security properties and requirements							Data perturbation		Data restriction			
		Consent and choice	Purpose legitimacy and specification	Collection limitation	Data minimization	Use, retention and disclosure limitation	Accuracy and quality	Openness, transparency and notice	Individual participation and access	Accountability	Confidentiality	Integrity	Availability	Authentication	Authorization	Accountability	Privacy compliance	Noise addition	Anonymization	Access control	Cryptographic	Block chain	
[19]	2009	-	-	-	-	-	-	-	-	-	D	-	-	S	S	S	S	-	-	-	S	-	3,5
[20]	2011	-	-	-	-	S	-	-	-	-	D	-	-	S	D	-	-	-	-	S	D	-	4,5
[21]	2014	-	-	-	-	-	-	-	-	-	D	-	S	S	S	-	-	-	-	-	S	-	3,0
[22]	2016	D	-	-	-	S	-	D	-	-	D	-	-	-	S	-	-	-	-	D	-	-	5,0
[23]	2016	D	-	-	-	-	-	D	-	S	-	-	-	-	S	S	-	-	-	-	-	-	3,5
[24]	2016	D	S	S	S	D	-	S	S	S	S	-	-	S	S	S	-	-	-	-	-	-	7,0
[1]	2016	D	D	D	-	D	S	S	-	S	D	-	-	-	S	S	-	-	-	D	S	-	9,0
[8]	2016	D	-	D	-	D	-	D	S	-	S	-	S	-	D	-	-	-	S	D	S	-	8,5
[25]	2017	D	D	D	-	D	-	S	D	S	S	-	-	-	D	S	S	S	S	D	S	-	11,0
[26]	2017	D	S	-	-	S	-	S	-	S	S	-	-	-	D	S	-	-	-	D	S	-	6,5
[27]	2017	D	D	-	-	D	D	D	-	-	D	D	D	D	D	-	-	-	-	D	D	-	12,0
[30]	2017	D	-	S	-	D	-	D	-	-	D	-	-	D	-	S	-	-	-	-	S	-	6,5
[31]	2017	-	-	-	-	S	-	-	-	-	D	S	-	D	S	-	-	-	-	-	D	-	4,5
[32]	2017	D	-	D	-	D	-	S	-	D	D	-	-	S	D	S	-	-	-	D	S	-	9,0
[33]	2017	S	-	-	-	-	-	S	-	-	D	-	-	D	D	D	D	-	-	-	D	-	7,0
[28]	2018	D	D	-	-	D	D	D	-	-	D	D	D	D	D	-	S	-	-	D	D	-	12,5
[29]	2018	D	D	-	-	D	D	D	-	-	D	D	D	D	D	-	-	-	-	D	D	-	12,0
[34]	2018	S	D	D	D	D	-	D	-	-	D	-	S	D	D	-	-	-	S	S	D	-	11,0
[10]	2018	S	-	D	-	D	-	S	-	-	S	-	-	-	S	-	-	-	-	D	-	-	5,0
[35]	2018	S	-	-	-	S	-	S	-	-	S	-	-	-	-	S	-	-	-	-	-	-	2,5
[36]	2018	-	-	-	-	-	D	-	-	-	S	-	S	D	-	-	-	-	-	-	S	-	3,5
[37]	2018	S	S	D	-	S	-	D	S	-	S	-	S	-	S	S	-	-	-	S	S	-	7,0
[38]	2018	D	S	-	-	D	-	D	-	-	S	-	-	S	-	S	D	-	-	S	-	-	6,5
[2]	2018	D	-	D	-	D	-	D	S	-	D	S	-	S	D	-	-	-	-	D	D	-	9,5
[39]	2019	-	-	-	-	D	-	-	S	S	D	S	-	S	S	S	D	-	-	D	S	-	7,5
[40]	2019	S	S	-	-	D	-	S	-	S	D	S	-	D	S	S	D	-	-	-	D	-	8,5
Total		17,0	8,5	9,0	1,5	18,0	4,5	15,5	3,5	4,5	20,5	5,0	5,5	13,0	16,5	7,5	5,5	0,5	1,5	14,0	14,5	0,0	

Tab. 1. Classification of the analyzed papers based on our adapted taxonomy et al. [14] (S-Shallow; D-Deep).

limitation on data processing, not just data collection. In general, more specific privacy concerns, which are usually not obvious, have not yet been addressed. For example, few authors have addressed *individual participation and access* (3.5 points), which refers to giving users the right to review, change, amend and remove their data in a simple manner whenever desired.

As for *privacy-preserving techniques*, only two *data restriction*-related techniques received a reasonable number of mentions: *cryptographic* (14.5 points) and *access control* (14.0 points). The high number of cryptographic-related solutions reflects one of the researchers' larges concerns – the risk of a confidentiality breach. Several researchers addressed the need to encrypt children's PII, and possibly related data, when using smart toys. Researchers are also interested in proposing ways to control access to toys only to those authorized. Access control can be treated with specific levels and goals. One form is through parental control tools, addressed in some of these papers as a proposed solution to reduce the risks to children's privacy.

Anonymization and *noise addition* were merely mentioned as solutions. It remains to be seen whether these *data perturbation*-related techniques are not appropriate for smart toys or whether they have not yet been considered by researchers in this context. At least for the general IoT context, according to Loukil *et al.* [14], the number of solutions found related to anonymization was about 60% greater than to access control. This scenario is very different from the one found for smart toys. Finally, solutions related to blockchain were not mentioned in any paper. However, this could be expected as blockchain is a very recent technique. Even for the more general IoT context, very few papers mentioned blockchain as a possible solution to privacy risks [14].

Finally, one can observe that each of the 26 papers addresses different perspectives of smart toy-related children's privacy risks and solutions. A few of them do not address any type of solution, but only existing privacy risks [23, 24, 35]. Considering 21 items (16 *ISO privacy principles* and five *privacy preservation techniques*), five papers can be considered quite comprehensive, with more than 10.0 points counted [25, 27, 28, 29, 34]. This result does not mean that these five works are necessarily better than the others, but only that they address more concerns in a single paper. This view provides only a reference for researchers interested in aspects related to smart toy-related children's privacy.

5.2. Future research topics

Ensuring privacy and security is hard for researchers and industry, especially for involving child users. Table 1 shows the magnitude of this complexity, as more than 15 privacy principles should be addressed to ensure an appropriate environment for the use of smart toys. Several of these principles were only shyly addressed, which may show the need for future research. This does not necessarily mean that there is no need for further research for others, especially when it turns out that few solutions have been proposed for them.

From a risk standpoint, one can see that few authors have addressed the need to *limit data collection* and even less the need to *minimize data use*. Apparently, toy makers simply collect and use data in an *ad hoc* way without careful analyzing the scenario at hand. Industry should realize that personal data is valuable and should not be treated as a disposable and worthless asset. This is a relevant research line yet to be dealt with in more depth from technical, social and policy point of views.

Other principles less addressed so far may have been overlooked because they are more generic and can be applied directly to the field of smart toys without any specificity. However, specific studies should be conducted to refine this hypothesis. For example, *accuracy and quality* are indeed common to any application domain, but it is still not clear whether this principle has specific characteristics regarding smart toys. This also applies to *individual participation and access*, as the need for users to be able to review their data and change, if necessary, is also a common requirement, but specific requirements may exist for the smart toys domain.

Accountability and *privacy compliance* have also been only marginally addressed so far and should be the focus of future work. Once privacy policies have been set, they need to be enforced, which should be done through accountability and privacy compliance.

Specifically regarding security properties and requirements, *integrity* and *availability* seem to be being neglected for this domain. Similar to the points outlined above, these two principles may be subject to future research, as all properties and security requirements must be guaranteed to ensure privacy as well.

Finally, from the solution point of view, the summary presented in Table 1 shows that almost no author sought to apply data perturbation techniques, such as *noise addition*. Experts in this type of technique could investigate whether this application domain could benefit from this type of solution beyond data restriction techniques. In addition, privacy preservation solutions in general represent, in fact, the great potential for contribution in this area. Even studies that provide data restriction-related solutions involving access control and cryptograpic still do so to a very limited extent. For now, the vast majority of studies are much more focused on presenting problems than solutions.

6. Conclusion

The advent and use of smart toys have exacerbated the children's privacy risks due to the amount of PII that they can collect and how ease toy makers can retain user data. Parents usually expect privacy when their children are using smart toys, which is not confirmed by many of the 26 papers analyzed in this study. Ensuring children's privacy while using smart toys is vital because of the risk associated with disclosure of sensitive information that can lead to different types of harm to children. Significant issues are data *confidentiality*, data *use, retention and disclosure limitation*, *authorization*, *consent and choice* to collect smart toy-related data, *openness, transparency and notice* about collected data and *authentication* mechanisms ensuring accuracy of data.

Some authors propose mitigation solutions to address the raised risks at different stages of the development and use life-cycle of smart toys. However, privacy basics are the most questioned by researchers, while other more complex principles have been overlooked. This may show that toy makers have not yet successfully addressed the minimum necessary to ensure privacy in this area of interest. Overall, most analyzed papers addressed the implementation of security controls to protect the privacy of PII. The proposed solutions seek technical, administrative and operational security or privacy controls preventing unauthorized access and disclosure of information. This is shown through the results of our taxonomy-based analysis.

As future work, a new evaluation framework can be proposed, with a similar aim to those found and shown herein, but more systematic and comprehensive, involving the whole classification adopted in this study.

References

[1] P. C. K. Hung, F. Iqbal, S.-C. Huang, M. Melaisi, and K. Pang, "A glance of child's play privacy in smart toys," in *2nd Int. Conf. on Cloud Comp. and Sec.*, pp. 217–231, 2016.

[2] S. Shasha, M. Mahmoud, M. Mannan, and A. Youssef, "Playing with danger: A taxonomy and evaluation of threats to smart toys," *IEEE Int. of Things J.*, vol. 6, no. 2, pp. 2986–3002, 2018.

[3] EU, "GDPR – General Data Protection Regulation," 2016. Eur. Parliament and Council of the Eur. Union.

[4] P. C. K. Hung, M. Fantinato, and J. Roa, "Children privacy protection," in *Encyclopedia of Computer Graphics and Games* (N. Lee, ed.), 2019.

[5] Canada, "PIPEDA – Personal Information Protection and Electronic Documents Act," 2018. Min. of Just., Canada.

[6] Brazil, "GLPPD – General Law of Protection of Personal Data," 2018. Brazil's National Congress and Presidency of the Republic.

[7] USA, "COPPA – Children's Online Privacy Protection Act," 2002. Federal Trade Commis., US Congress, USA.

[8] P. C. K. Hung, M. Fantinato, and L. Rafferty, "A study of privacy requirements for smart toys," in *20th Pac. Asia Conf. on Inf. Sys.*, pp. 1–7, 2016.

[9] M. Fantinato, P. C. K. Hung, Y. Jiang, J. Roa, P. Villarreal, M. Melaisi, and F. Amancio, "A survey on purchase intention of hello barbie in brazil and argentina," in *Computing in Smart Toys* (J. K. Tang and P. C. K. Hung, eds.), pp. 21–34, 2017.

[10] M. Fantinato, P. C. K. Hung, Y. Jiang, J. Roa, P. Villarreal, M. Melaisi, and F. Amancio, "A preliminary study of hello barbie in Brazil and Argentina," *Sust. Cities and Soc.*, vol. 40, pp. 83–90, 2018.

[11] E. P. d. S. Nunes, E. M. Lemos, C. Maciel, and C. Nunes, "Human factors and interaction strategies in three-dimensional virtual environments to support the development of digital interactive therapeutic toy: A systematic review," in *7th Int. Conf. on Vir., Augm. and Mixed Real.*, pp. 368–378, 2015.

[12] E. P. d. S. Nunes, V. A. d. Conceição Jr, L. V. G. Santos, M. F. L. Pereira, and L. C. L. D. F. Borges, "Inclusive toys for rehabilitation of children with disability: A systematic review," in *11th Int. Conf. on Univ. Access in Human-Comp. Inter.*, pp. 503–514, 2017.

[13] A. P. d. Albuquerque and J. Kelner, "Toy user interfaces: Systematic and industrial mapping," *J. of Sys. Arch.*, vol. 97, no. Aug, pp. 77–106, 2018.

[14] F. Loukil, C. Ghedira-Guegan, A. N. Benharkat, K. Boukadi, and Z. Maamar, "Privacy-aware in the IoT applications: A systematic literature review," in *25th Int. Conf. on Coop. Inf. Sys.*, pp. 552–569, 2017.

[15] T. Oktavia, Y. Tjong, H. Prabowo, and Meyliana, "Security and privacy challenge in bring your own device environment: A systematic literature review," in *2016 Int. Conf. on Inf. Man. and Tech.*, pp. 194–199, 2017.

[16] B. Nelson and T. Olovsson, "Security and privacy for big data: A systematic literature review," in *2016 IEEE Int. Conf. on Big Data*, pp. 3693–3702, 2016.

[17] B. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering," Tech. Rep. EBSE-2007-01, Keele Univ. and Univ. of Durham, UK, 2007. Version 2.3.

[18] J. vom Brocke, A. Simons, B. Niehaves, K. Riemer, R. Plattfaut, and A. Cleven, "Reconstructing the giant: On the importance of rigour in documenting the literature search process," in *17th Eur. Conf. on Inf. Sys.*, pp. 2206–2217, 2009.

[19] T. Denning, C. Matuszek, K. Koscher, J. R. Smith, and T. Kohno, "A spotlight on security and privacy risks with future household robots: attacks and lessons," in *11th Int. Conf. on Ubiq. Comp.*, pp. 105–114, 2009.

[20] S. Yong, D. Lindskog, R. Ruhl, and P. Zavarsky, "Risk mitigation strategies for mobile Wi-Fi robot toys from online pedophiles," in *2011 Int. Conf. on Priv., Secur., Risk and Trust*, pp. 1220–1223, 2011.

[21] J.-S. Pleban, R. Band, and R. Creutzburg, "Hacking and securing the AR.Drone 2.0 quadcopter: Investigations for improving the security of a toy," in *Mobile Devices and Multimedia: Enabling Technologies, Algorithms, and Applications*, pp. 1–12, 2014.

[22] M. L. Jones and K. Meurer, "Can (and should) Hello Barbie keep a secret?," in *2016 IEEE Int. Symp. on Ethics in Engin., Sci. and Tech.*, pp. 1–7, 2016.

[23] M. L. Jones, "Your new best frenemy: Hello barbie and privacy without screens," *Engaging Sci., Tech., and Soc.*, vol. 2, pp. 242–246, 2016.

[24] D. Holloway and L. Green, "The internet of toys," *Comm. Res. and Prac.*, pp. 1–14, 2016.

[25] L. Rafferty, P. Hung, M. Fantinato, S. M. Peres, F. Iqbal, S. Kuo, and S. Huang, "Towards a privacy rule conceptual model for smart toys," in *50th Hawaii Int. Conf. on Sys. Sci.*, pp. 1–10, 2017.

[26] E. McReynolds, S. Hubbard, T. Lau, A. Saraf, M. Cakmak, and F. Roesner, "Toys that listen: A study of parents, children, and internet-connected toys," in *2017 ACM SIGCHI Conf. on Human Factors in Comp. Sys.*, pp. 5197–5207, 2017.

[27] L. G. Carvalho and M. M. Eler, "Security requirements for smart toys," in *19th Int. Conf. on Ent. Inf. Sys.*, pp. 144–154, 2017.

[28] L. G. Carvalho and M. M. Eler, "Security requirements and tests for smart toys," in *19th Int. Conf. on Ent. Inf. Sys.*, pp. 291–312, 2018.

[29] L. G. Carvalho and M. M. Eler, "Security tests for smart toys," in *20th Int. Conf. on Ent. Inf. Sys.*, pp. 111–120, 2018.

[30] C. Moini, "Protecting privacy in the era of smart toys: Does hello barbie have a duty to report," *Cath. Univ. J. of Law and Tech.*, vol. 25, no. 2, pp. 281–318, 2017.

[31] J. Valente and A. A. Cardenas, "Security & privacy in smart toys," in *1st Work. on Int. of Things Sec. and Priv.*, pp. 19–24, 2017.

[32] M. Mahmoud, M. Z. Hossen, H. Barakat, M. Mannan, and A. Youssef, "Towards a comprehensive analytical framework for smart toy privacy practices," in *7th Work. on Socio-Tech. Asp. in Sec. and Trust*, pp. 64–75, 2017.

[33] J. Haynes, M. Ramirez, T. Hayajneh, and M. Z. A. Bhuiyan, "A framework for preventing the exploitation of IoT smart toys for reconnaissance and exfiltration," in *10th Int. Conf. on Sec., Priv. and Anony. in Comp., Comm. and Storage*, pp. 581–592, 2017.

[34] K. Demetzou, L. Böck, and O. Hanteer, "Smart bears don't talk to strangers: Analysing privacy concerns and technical solutions in smart toys for children," in *Living in the Int. of Things: Cybersec. of the IoT*, pp. 1–7, 2018.

[35] N. Kshetri and J. Voas, "Cyberthreats under the bed," *Computer*, vol. 51, no. 5, pp. 92–95, 2018.

[36] J. L. Espinosa-Aranda, N. Vallez, J. M. Rico-Saavedra, J. Parra-Patino, G. Bueno, M. Sorci, D. Moloney, D. Pena, and O. Deniz, "Smart doll: Emotion recognition using embedded deep learning," *Symmetry*, vol. 10, no. 9, pp. 1–18, 2018.

[37] W. Chowdhury, "Toys that talk to strangers: A look at the privacy policies of connected toys," in *Conf. of Fut. Tech. Conf.*, pp. 152–158, 2018.

[38] K. L. Smith and L. R. Shade, "Childrens digital playgrounds as data assemblages: Problematics of privacy, personalization, and promotional culture," *Big Data & Soc.*, vol. 5, no. 2, pp. 1–12, 2018.

[39] B. Yankson, F. Iqbal, Z. Lu, X. Wang, and P. C. K. Hung, "Modeling privacy preservation in smart connected toys by petri-nets," in *52nd Hawaii Int. Conf. on Sys. Sci.*, pp. 1696–1705, 2019.

[40] G. Chu, N. Apthorpe, and N. Feamster, "Security and privacy analyses of internet of things children's toys," *IEEE Int. of Things J.*, vol. 6, no. 1, pp. 978–985, 2019.