Eunomia: A Real-time Privacy Compliance Firewall for Alexa Skills

Javaria Ahmad*, Fengjun Li†, Razvan Beuran‡, and Bo Luo†
* *CISA, University of Central Missouri, Warrensburg, MO, USA. Email: ahmad@ucmo.edu*
† *EECS and I2S, The University of Kansas, Lawrence, KS, USA. Email: fli@ku.edu; bluo@ku.edu*
‡ *Japan Advanced Institute of Science and Technology, Nomi, Ishikawa, Japan*

***Abstract*—Voice assistants (VAs), such as Amazon Alexa, are integrated with numerous smart home devices to process user requests using apps called *skills*. With their growing popularity, VAs also pose serious privacy concerns. Sensitive user data captured by VAs may be transmitted to third-party skills without users' consent or knowledge about how their data is handled. Privacy policies are a standard medium to inform the users of the skills' data practices. However, privacy policy compliance verification of such skills is challenging, since the source code is controlled by the skill developers, who can make arbitrary changes to the behaviors of the skill without being audited; hence, conventional defense mechanisms using static/dynamic code analysis can be easily evaded. In this paper, we present Eunomia, the first real-time privacy compliance *firewall* for Alexa skills. As the skills interact with the users, Eunomia hijacks and examines their communications from the skills to the users, and validates them against the published privacy policies that are parsed using a BERT-based policy analysis module. When non-compliant skill behaviors are detected, Eunomia stops the interaction and warns the user about the non-compliance. We evaluate Eunomia with 55,898 skills on Amazon skills store to demonstrate its effectiveness and to provide a privacy compliance landscape of Alexa skills.**

***Index Terms*—Alexa Skills, Privacy Compliance, Privacy Firewall**

1. Introduction

Voice assistants (VAs), such as Google Assistant and Amazon Alexa, are integrated into household smart devices such as speakers, home hubs, and smartphones. They are activated using wake words, e.g., "Alexa," and then follow the voice commands to perform certain actions or services, e.g., playing media, making calls, reading emails, and interacting with the Internet of Things (IoT) devices. With the growing popularity of VA devices and services, security and privacy concerns arise. For example, the audio being sent to the cloud may contain sensitive content, such as users' birthdays or health-related information. VAs may also misinterpret the commands resulting in unintended purchases. Moreover, some VAs make use of third-party apps that pose further threats as documented in the literature [1]–[7], such as, generating malicious instructions with a frequency that is detectable and perceived by VAs but not by people, a security flaw in an app that allows calling contacts through voice instructions even when there is lock on the phone, malicious users accessing and instructing VAs to disclose VA owners' stored account information, use of synthesized speech that mimics a VA owner's voice to deliver instructions to the VA, and use of exploiter apps that swap out real information with false information before delivery to the users to cause miscommunications. Some defenses that VA manufacturers and app developers can employ against these threats are also proposed by this literature, for example, informing the VA users on receipt of any instructions through beeping or using screen notifications, integrating advanced speech recognition techniques so malicious users giving voice commands would not be successful and ensuring that the locked devices needs to be unlocked before accepting any commands, developing and installing apps on VAs to screen and purge user information so it does not transmit to cloud, and regulatory bodies checking for apps and taking actions if the apps have logic to swap out legitimate information with the fake data.

Amazon Alexa's apps, in particular, are called *Skills*. They are used to enhance VA capabilities, but malicious skills and data leaks may risk the privacy of the users [8], [9]. Skills run outside the VA hardware on *third-party servers* selected by the skill developers. Sensitive user data may be transmitted to these third parties without users' knowledge. Privacy policies are a standard way to inform the users of the skills' practices. However, Amazon only requires developers to provide a privacy policy if the skill collects sensitive user information, and Amazon is also known to enforce such requirements loosely [10]–[12]. In addition, when a skill is enabled using a voice command, the user may not have a chance to review the privacy policy. Moreover, Alexa may implicitly enable specific skills, meaning that the user is not asked to review and approve the privacy policies before user's requests are processed. As soon as the user answers the skill's data collection questions, the provided information is transmitted to the cloud that hosts the skill. Hence, such an information collection attempt, if not properly disclosed to the user, is considered a violation of Alexa's policies and a threat to user's privacy. The violations involving sensitive user data could even result in legal consequences on a global level, for example, if non-compliant skills from the US market are made available to EU consumers, EU's General Data Protection Regulation

(GDPR) [13], applies to the US organization that handles the data of EU consumers. Better enforcement of the privacy requirements in the US market would protect the users as well as the US organizations against foreign laws.

Amazon enforces a skill verification process and there are also research efforts for compliance validation of skills [14]–[16]. However, they may be easily avoided, since the codebase of the skills and the delivery of the services are fully controlled by the developers: (1) They may provide a compliant version (consistent with the privacy practice declarations) at validation time but later change to a non-compliant version without going through re-verification. (2) Existing research validates privacy compliance by interacting with the skills to trigger all their behaviors, which is not only time-consuming but also less effective against logic bombs or time bombs, that is, malicious behavior that only manifests at a certain time after the validation process or after certain logic takes place. (3) Other compliance-check approaches, e.g., Liao et al. in [17], compare the privacy policy with the app descriptions. However, skills are not required to fully depict their functions in the descriptions.

The limitations of the privacy compliance studies of VAs motivate us to develop a new solution called Eunomia, which is a real-time privacy-compliance firewall for Alexa skills. We argue that *real-time* firewall is the only adequate solution for effective compliance checks of the skills. Thus, Eunomia is designed as a firewall that protects the users from revealing their sensitive information to non-compliant skills. **Eunomia is unique from the existing approaches because the real-time firewall defends the users against any non-compliant data collection practices** that it encounters. This is in contrast to the existing studies that function as fuzzing tools for detecting violations. The effectiveness of fuzzing tools relies on their coverage and they must test all possible interactions which has its challenges as mentioned earlier in the section. In practice, Eunomia captures the queries from the skills on the fly as the user interacts with Alexa. The monitoring component of the Eunomia firewall validates the queries for compliance against the practices declared in the privacy policies. If a skill's action is determined to be non-compliant at any time, Eunomia firewall plays an active role by issuing a command to pause the interaction before the user responds with potentially sensitive information and warning the user about the non-compliance. If the skill's behavior is compliant, the conversation continues as usual. In our prototype implementation, we categorize sensitive information into higher-risk sensitive personal information and lower-risk non-personal information. Sensitive personal information includes users' identifiable attributes and health/medical information [18], [19] such as name, email, address, and gender, and generally-known health information such as heart rate and blood group. Sensitive non-personal information includes personal habits, family, lifestyle, coarse location, etc. Some data types are less risky, for example, coarse location, but their use needs to be disclosed in the privacy policies as well to be considered compliant as required by Amazon[1] which states that developers "must disclose" these. The coarse location is controlled by the permission `ACCESS_COARSE_LOCATION`. In the deployment of Eunomia, the user could specify her own privacy expectations by adjusting (adding to or removing from) the private information ontology.

Eunomia consists of four main components (to be articulated in Section 3): (1) an ontology definition to identify the relationships between data objects in privacy policies to aid in the compliance validation; (2) a Bidirectional Encoder Representations from Transformers (BERT) based model to extract objects (entities and data) in the privacy policies and then identify the declared privacy practices; (3) a module to capture and analyze the audio/textual output from skills to identify the privacy practices of skills; and (4) a compliance validation module to detect the inconsistencies between the actual skill practices with the privacy disclosures on-the-fly.

We evaluate Eunomia through extensive experiments, which also map out the current privacy compliance landscape of Alexa skills. For testing the Eunomia firewall and providing the measurement dataset, we adopt a chatbot developed by Young et al. in [16], which achieves a test coverage of 92.1%. This test coverage is a measure of how well the chatbot navigates through all possible pathways of skill interaction. We interact with the skills and invoke Eunomia to provide real-time compliance validation. To capture the outputs from all the 55,898 skills from the Amazon skills store, we created 15 Amazon developer accounts to interact with the skills simultaneously using 4 Windows 10 machines over 4 months. Among 55,898 skills, we found that 1,405 skill outputs contain sensitive data practices. Experimental results show that Eunomia achieves a precision of 96-100% and a recall of 96.3%. This measurement study answers (in Section 5) four important questions regarding the development and use of Amazon skills: (1) *How effective is Eunomia in protecting the users?* (2) *What is the overall compliance status of Alexa skills?* (3) *Which particular compliance gaps are there in Alexa skills?* (4) *Which type of compliance issues are more common than others?*

The technical contributions of this paper are as follows:

1. We present Eunomia, the first real-time privacy-compliance firewall for VAs. It extracts the skills' privacy practices on the fly and validates against the disclosures in the policies. Eunomia also functions as a firewall to defend users from malicious skills, as it pauses the skills when non-compliant behavior is detected. We make our code publicly available[2].

2. With the adoption of state-of-the-art NLP algorithms, and automated ontology and synonym extraction, Eunomia identifies fine-grained disclosure types, and achieves outstanding performance in both precision and recall.

3. We provide a thorough analysis of the privacy policy landscape and the state of compliance of Alexa skills through a

1. https://developer.amazon.com/docs/policy-center/user-data-privacy.html

2. https://github.com/Eunomia-skills

manual evaluation and an automated large-scale assessment using Eunomia.

The rest of the paper is organized as follows: We introduce the problem, the related work, the challenges, and the defense model in Section 2, followed by the technical details of the Eunomia solution in Section 3. We then present the experiment results with the skills' privacy compliance landscape in Section 4, answer the research questions in Section 5, and finally conclude the paper in Section 6.

2. Motivation and Related Work

Alexa skills may collect sensitive data from users. For example, "Rental Retreat" collects personal information as follows:

```
Skill:   Welcome to Rental Retreat!
         You can view the Alexa app to
         give permission to use your
         name and email...
```

To be compliant, the skill must declare the sensitive data collection practices in its privacy policy. The privacy policy of Rental Retreat declares the following:

```
This app requires permission to use
your name, email, and phone for some
features.
```

On the contrary, if a skill fails to declare the practices in its policy, the skill is non-compliant. In Eunomia, we aim to detect such non-compliant behaviors, and, more importantly, we protect the users by providing a firewall that stops the interaction with any skill that attempts non-compliant behaviors. For instance, the skill "Obesity Checker" asks for the user's height and weight. However, the skill did not provide a valid URL to its privacy policy. As shown below, Eunomia detects and terminates the non-compliant behaviors and displays a warning message for the user. Note that we demonstrate the communication in text, while the actual Eunomia does not need a text/graphical interface.

```
Skill:   Welcome to Obesity Checker.
         Saying your height and body
         weight, you will see your BMI
         and suitable weight. Say your
         height and weight.
-- Eunomia: stop
Skill:   Good Bye
-- Eunomia: exit

Eunomia: Alert, alert, the skill
         attempts to collect your
         height, weight without
         disclosure.
```

Malicious Skills and Certification. Cheng et al. test Amazon's certification process and find that all the non-compliant skills passed the certification [10]. They also identify other threats such as fake reviews, lack of re-verification for changes in skill code, and lack of automated compliance validation tools. Talebi et al. screen and remove sensitive user information going to the cloud [5]. Skill squatting attacks exploit the ambiguity in the pronunciation of skills. The attackers develop skills that sound similar to legit skills to trick users into enabling the malicious skills [9], [24].

Skill Compliance Check. Liao et al. [17] check the privacy policy of the voice apps against the app descriptions to find inconsistencies. This study does not provide conclusive compliance analysis because it does not validate the skill actions, while the skill descriptions do not always accurately reflect its actions. Lentzsch et al. [25] analyze skill privacy policies against the requested permissions. Le et al. [26] compare skill responses with the permissions but not with the privacy policy declarations. Guo et al. in SkillExplorer [15] use an interaction model to capture the skills' collection of user data and validate the practices against their policy declarations. Young et al. [16] also develop an interaction model to detect policy violations. Xie et al. [23] use the existing interaction model proposed by Guo et al. [15] for compliance validation. Edu et al. [14] conduct a 3-year measurement study of permissions and privacy disclosures of the Alexa ecosystem. It also invokes SkillExplorer [15] in its interactions with the skills.

Existing Studies on Privacy Compliance & Challenges. In Table 1, we summarize the SOTA privacy protection solutions for Alexa skills and compare them with Eunomia. Among these approaches, SkillExplorer [15] develops the first-ever interaction model with the skills while providing an analysis of privacy compliance. It employs PolicyLint [21] to recognize the disclosed privacy practices from the policies. PolicyLint employs limited ontologies, privacy term synonyms, and a simple NLP model (spaCy [27]), which provides limited performance. SkillDetective [16] adapts the data ontologies from PoliCheck [20], which utilizes keyword-based matching on a small set of 21 privacy terms from NIST [18]. Skipper [23] does not consider ontologies in policy analysis, hence, it cannot handle cases when the policy uses generic terms (e.g., "health data") instead of the specific terms (e.g., "heart rate"). Aligning with this limitation, other disclosure categories, such as unclear disclosures, are not considered besides either exact matching or non-matching. Skipper also uses a simple spaCy NLP model to extract privacy terms and their synonyms.

Earlier studies on privacy policy compliance mostly focus on mobile apps [20]–[22]. They are unsuitable for off-the-shelf skill compliance checks without domain adaptation for specific data types. Mallojula et al. [28] study the security issues and privacy compliance in automotive Android mobile apps. Zhao et al. [29] discover malicious iOS apps by employing code analysis and consistency checks between app descriptions and reviews. Studies on IoT devices/apps [30] focus on data types such as device-specific and sensor-based data which is different from the data captured through user interaction in VA studies. Similarly, studies for specific domains, for example, websites [31] and Oculus Virtual Reality (OVR) [32], cannot be directly applied to VAs. Previous studies are usually based on

TABLE 1. COMPARISON OF EUNOMIA WITH EXISTING APPROACHES.

Study Name	Study Type	Policy Analysis NLP Tech				RT F.Wall	Issue Warn	Disc Cat	Int Cvr	Prec (%)	Rec (%)	# Skill Issues	# Skills Tested
		Ont	Syn	Neg	A. NLP								
PoliCheck [20]	Mobile	~	✓	✓	~	×	×	✓	~	90.8	N/A	N/A	N/A
PolicyLint [21]		~	✓	✓	~	×	×	×	~	97.3	81.7	N/A	N/A
Maps [22]		×	×	✓	×	×	×	×	~	82	100	N/A	N/A
SkillExplorer [15]	Skill	~	~	✓	~	×	×	×	~	N/A	67.2†	815	30,801
SkillDetective [16]		~	×	✓	×	×	×	✓	~	90.5	88.5	623	53,859
Skipper [23]		×	~	✓	~	×	×	×	~	71-100‡	100‡	1,012	61,505
Eunomia	**Skill**	✓	✓	✓	✓	✓	✓	✓	✓	**96-100**	**96.3**	**1,041**	**55,898**

✓/×/~: covered/not covered/partially covered; Ont: domain-specific ontology; Syn: automated synonyms extraction; Neg: negations handling; A. NLP: *advanced NLP models*; RT F. Wall: real-time firewall; Issue Warn: notification of non-compliance to users; Disc Cat: fine-grained disclosure categories; Int Cvr: complete interaction testing coverage; Prec/Rec: precision and recall of policy violation identification; †SkillExplorer did not report its recall. SkillDetective evaluated SkillExplorer and reported its recall in a comparative study. ‡Skipper reported 100% recall and 71% to 100% precision.

a small hand-annotated corpus [22], [33]–[35], which is limited and error-prone. The NLP task of analyzing the privacy policies is challenging due to the difficulty of identifying contradictions in the text [21], confusing language, and statements spanning multiple sentences [20]. Moreover, detecting negative statements in privacy policy with high accuracy is also challenging using bigrams and regex [21], [22], [36]. More recent studies [37] utilizing advanced NLP techniques for privacy policy analysis are out of the scope of our study as they focus on capturing the purpose behind data usage by various entities that capture the data. Last, spaCy's CNN-based NER engine is the common approach adopted by many studies for NLP tasks, but it provides limited performance as compared to the transformer-based BERT [38] model which also allows for fine-tuning of the pre-trained model [39].

Eunomia tackles these challenges by considering the text semantics using automated ontology and synonym extraction by employing the SOTA BERT model that is fine-tuned on a large corpus of 2,000 privacy policies prepared using a combination of manual and automated techniques. The model yields 52 data ontology pairs and 7,592 synonyms. As a result of the enhancements, we can detect the skill statements containing sensitive data that other approaches in Table 1 would not be able to. For instance, Eunomia was able to identify skills that prompt for coarse location information or exercise/health training information, while all other solutions miss such privacy practices. Meanwhile, Eunomia provides a unique solution as compared to previous works by implementing a real-time firewall for skill compliance validation, which is capable of catching newly introduced violations and time/logic bombs.

3. Eunomia: A Real-Time Compliance Validation Firewall

3.1. The Defense Model and Assumptions

In our defense model, Eunomia does not have any access to the source code of the skills. Instead, it is deployed on the user's side to monitor all the skill outputs during the interaction. As shown in Figure 1 (I) and (II), Eunomia could be deployed as a standalone device or embedded in an existing VA device. There are three major differences between these two implementations: (1) the computing platform that hosts Eunomia, (2) the mechanism to capture the skill's output, and (3) the mechanism to terminate the skill and warn the user. More details of the implementation of Eunomia will be articulated in Section 3.7. In either implementation, Eunomia needs access to the following:

(1) User-skill interaction. Eunomia needs to intercept the (verbal) communication from the skill to the user. This could be achieved by invoking system functions in VA devices in the embedded deployment, or capturing the voice communication and employing a speech-to-text conversion tool in the standalone model.

(2) Privacy policies. Amazon requires skills that collect personal information to "provide a URL that sends users *directly* to a legally adequate privacy policy" [40], hence, the policies are expected to be publicly accessible. Eunomia directly retrieves the URL from Amazon's skills store.

(3) Communication channels to skill/user. To effectively terminate the non-compliant skills and to notify the user, Eunomia needs access to the VA's API to communicate to or terminate the skill, and/or access to an audio device to send warnings to the user.

Eunomia and User Privacy. Eunomia only processes one-way communication from the skills to the users. Whether the communication is captured in text or audio depends on the design choice of Eunomia (Section 3.7). If Eunomia captures any surrounding audio that is not a skill's output, it drops it. Eunomia is a *stateless* system that does *not* collect, memorize, store, or send out any user information/activities that could be used to infer the user's behavior. Eunomia is installed locally, with *fully open-sourced code*, which does *not* transmit/share information with any external party. All policy caches and non-compliance logs are local. Therefore, Eunomia does *not* introduce any privacy threat.

3.2. The Architecture of Eunomia

The detailed architecture of Eunomia is shown in Fig. 1 (III). It consists of three main components:

• **[Section 3.4] Policy Analysis Ⓕ**. It extracts the privacy policy Ⓔ from the skills store Ⓒ and performs an NLP-based policy analysis. It identifies the declared privacy practices Ⓖ as the output.

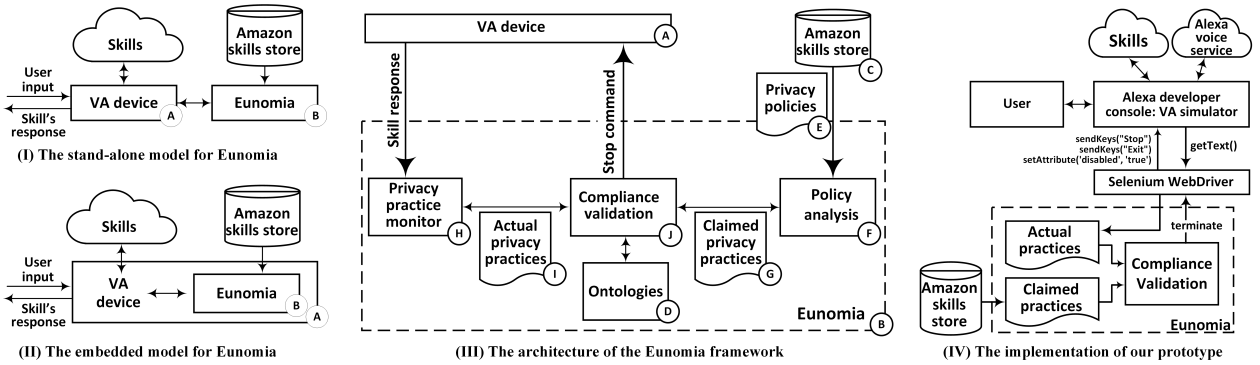

Figure 1. The deployment models, the system architecture, and the prototype implementation of Eunomia.

• **[Section 3.5] Privacy Practice Monitor Ⓗ**. It intercepts the communication from the skill (running in the VA) to the user and analyzes the communication data to extract queries about sensitive information.

• **[Section 3.6] Compliance Validation Ⓙ**. It evaluates the actual privacy practices Ⓘ against the claimed privacy practices Ⓖ using a predefined ontology Ⓓ [Section 3.3] to match data objects.

3.3. Ontology Definition and Domain Adaptation

Eunomia develops an ontology to capture the "is-a" relationships between the terms in privacy policies. They are necessary for understanding the semantics of policies. In particular, the policies may use generic or specific terms, e.g., collecting `personal information` or more specifically, collecting `blood pressure`. Ontologies capture the subsumptive relationship between the data terms, that is, `blood pressure` is `health-related data`, which is `personal information`. In compliance validation, if a skill's privacy policy claims that it does not collect personal information, but the skill prompts for blood pressure, we conclude that there is non-compliance between the declared and the actual practice.

Inspired by [20], [39], we design the following steps to generate more comprehensive ontologies and more VA-specific data types, such as fitness data and health-related data. (1) We first identify 150 privacy policies from random apps. One of the authors with a privacy research background and two other annotators unrelated to this project annotate all the selected policies with *Data* objects and *is-a* relationships among them. To perform object and relationship annotations, the three annotators receive separate copies of the selected privacy policies and they work independently. (2) We compare the results of the three annotators to reach a consensus on each annotation. We select the results with a majority vote followed by a discussion of any disagreement. We observe that the first annotator dissents from the second and third annotators in 8 annotations of data objects and 20 relationship annotations. The second annotator dissents from the first and third annotators in 2 data objects and 6 relationship annotations. The third annotator dissents in 1 data object and 3 relationships. The final annotation results contain 1,314 data objects and 2,348 relationships. (3) This corpus is used to train a Tok2Vec relation classifier. We also fine-tune a BERT [38] model on 2000 privacy policies. We then apply the BERT model and the Tok2Vec classifier to extract data objects and their relationships from privacy policies. (4) From the ontologies, we create the *data ontology graph*. We also extract the synonyms from the ontologies to recognize similarities in the objects, for example, `account information` and `account detail` are identified as synonyms.

Dictionary of Data Types. We identify the data types automatically by applying the fine-tuned BERT model on privacy policies. We also include the data types from literature [16], [20] as well as from the NIST report [18]. Some sensitive data types in our study are shown in Table 6. Finally, the dictionary of the data types from the generated ontologies and synonyms are used for compliance validation.

3.4. Privacy Policy Analysis

In this component, Eunomia conducts an automated analysis of skills' privacy policies to extract the privacy practices. We conduct sentence-level NLP to identify the data sharing and collection statements. We employ the following steps during the policy analysis:

• **Tuning a SOTA BERT Model for Domain Adaptation.** To identify the data objects and entities in the sentences, we adopt and fine-tune a pre-trained transformer-based BERT model [38]. In the same process as described in Section 3.3, we manually annotate 500 privacy policies with data objects and entities but not relationships. We then obtain the dictionary of these annotations and apply it to 2,000 privacy policies. We use 1,600 of the policies to fine-tune the BERT model while we use the remaining 400 for validation. The tuned model demonstrates a significant performance improvement (precision of 87.9%; recall of 88.09% for data object detection) as compared to CNN-based approaches, such as PolicyLint [21] (precision of 82.2%; recall of 79.8%). Besides notable enhancements of accuracy, transformers also allow for an economical training process (time and resources) through fine-tuning the pre-trained models [41]. The training took only 2 hours on an NVIDIA Tesla P100 GPU (PCI-E 16GB).

• **Creating Dependency-Based Parse Trees.** Based on the literature [21], we create dependency-based parse trees by

processing each sentence of the privacy policies to understand and extract the association between the parts of the sentence, e.g., *entities* taking *actions* on the *data types*. For each sentence, we first apply our fine-tuned BERT model to identify the data objects and entities. We then identify the verbs associated with sharing or collection practices by the entities on data types as mentioned in each sentence. We also identify the pronouns in the sentences as entities, e.g., "I" and "We". The trees we create manifest the relationships between the data objects, entities, and privacy practice verbs. For example, an entity "we" that collects personal information is semantically represented using a dependency parse tree, where the nodes are labeled as "personal information", "we", and "collect". As we are interested only in relationships related to data collection and sharing, we discard the nodes that are unrelated to such practices.

• **Recognizing Negative Practices.** We also recognize the negative practices when certain data is *not* collected or shared. The negative verbs, e.g., *do not* collect/share, are identified in the trees, as well as the exceptions (except, besides, aside, etc.) to recognize the practices where certain actions are not performed. An example of such a practice is: *"We do not collect your personal information"*.

• **Representing Practices in Tuples.** Eventually, the trees are converted to tuples of privacy practices. The tuples contain information on actors performing actions on data objects, e.g., '`(we, collect, email address)`'.

Besides the BERT-based privacy practice extraction, we complement it with keyword-based term matching with a dictionary of data types from Section 3.3 for additional coverage. We pass the analysis results to compliance validation.

3.5. Analysis of Skill Actions

When a skill is invoked, it often engages in a series of questions with the user. We examine such information collection attempts using the following steps.

• **Capturing Skill Outputs.** The Alexa voice cloud transmits audio outputs to VA devices [5]. The communication is intercepted with one of the following: (1) For VA devices with built-in speech-to-text converters, e.g. devices that show subtitles on a screen, we directly capture the text. (2) For devices missing this feature, we implement a customized module to (internally or externally) capture audio, invoke a speech-to-text converter, and send the text to Eunomia. The communication between Alexa cloud and VA devices is encrypted [42], but the encryption has no impact on Eunomia, as we capture the communication from VA devices when it is transmitted to the users in decrypted form. Please see Section 3.7 for prototype implementation.

• **Extracting Privacy Practices from Skill Actions.** For the skill responses, we iterate over the sentences and extract the practices by detecting the data objects using a combination of fine-tuned BERT model and keyword-based term matching with a dictionary of data types from Section 3.3. We check the context of the output to identify data collection attempts, i.e., if the skill directly asks for user information or gives instructions to provide such data. For example, for both the outputs, "What is your name?" and "Tell me your name.", we identify them as prompting for user data. Informational outputs, e.g., "I can tell you animal names", are not marked as data collection attempts.

Eunomia offers enhanced coverage in identifying the privacy practices even when the skill responses are evasive and do not state the practices in clear well-known key terms, e.g., "name". For example, for the skill responses, "what may I call you?", "which station do you board?" and "what is the region you currently occupy?", Eunomia correctly identifies them as name and location collection attempts.

3.6. Compliance Validation

The final component is to validate the captured privacy practice against the declared practice. We first feed the policy analysis results (Section 3.4) and the real-time skill action analysis data (Section 3.5) along with the skill ID to the *compliance validation module*. In this module, the skill practice is compared with each privacy policy tuple. With the help of the ontology and synonyms, we identify if the data object in the practice: (1) is *irrelevant* to the data object in the policy tuple, (2) exactly matches with the tuple and the collection action is consistent with the tuple, (3) matches with the tuple but the policy is negated, or (4) is a subset of the data object of the policy tuple. Eventually, we identify four types of disclosures:

(i) Compliant and Clear (Clear). The data collection practice of the skill is clearly disclosed in the privacy policy, i.e., one exact match is identified.

(ii) Compliant but Unclear (Unclear). A privacy policy uses umbrella text to disclose the skill's action, i.e., the data object of the action does not match any policy, but it is a subset of a policy. For example, if a privacy policy declares that it "collects personal information" without specifically stating the information type, while the skill collects the user's name and address, we consider this an unclear disclosure.

(iii) Undisclosed and Non-compliant (Undisclosed). The policy does not disclose the skill's practice, i.e., all the policy tuples are irrelevant to the data collection practice.

(iv) Inaccurate and Non-compliant (Inaccurate). A privacy policy incorrectly states that it does not collect user information while the skill collects user information, i.e., the action matches with a negated policy.

For example, if a skill prompts the user: "I need to know your address" and the skill's privacy practice states "We collect the address and email from you", the practice is classified as "Clear" because the practice is precisely declared in the policy. If a skill asks the user: "What is your age?" but only declares in generic terms that "We collect personal information" without specifying "age", it is classified as "Unclear". When the skill asks: "Tell me your phone number." but does not disclose the practice either in direct or generic terms, the practice is "Undisclosed". If a skill says, "Tell me your age." and the privacy policy states that "We do not collect any personal data from you.", we mark the disclosure as "Inaccurate".

We consider the skills with precise and unclear disclosures in the privacy policies as *compliant*, while skills with undisclosed and inaccurate disclosures are *non-compliant*. Our definitions of compliance and non-compliance align with Amazon's regulations [40] and literature, e.g., PoliCheck [20] and IoTPrivComp [39]. Amazon requires that a skill's policy should be complete, transparent, written in a language that matches the skill's language, specific to the skill, and has an accessible URL that directly takes users to the policy [40]. Amazon does not specify that a skill's practice declarations need to be in clear terms, so as long as the declarations are present in clear or vague terms, the practices are considered compliant. Meanwhile, if the practices are either not declared at all in the policy or inaccurately declared, the skill is considered non-compliant by Amazon and the privacy research literature [20], [39]. When non-compliant actions are detected, we log the information about the non-compliance and issue the "stop" command to stop the skill (See Section 3.7 for details). We also notify the user of the privacy issue and allow them to add an exception to allow the data collection practice.

3.7. Prototype and Implementation Details

We build a prototype with a baseline system that enables the communication between the user and skills (a VA device), while Eunomia "watches" the communication, analyzes the skill's responses on the fly, and stops the skill if non-compliance is detected.

The Baseline System. As described in Section 3.1, Eunomia could be deployed in a standalone model or an embedded model. In our prototype, a simulated VA device built with Amazon's Alexa developer console (Fig. 1 (IV)) loads the skills from AWS or other cloud servers, invokes the Alexa voice service to communicate with the user, and transmits the user's responses back to the Skill. Meanwhile, Eunomia runs on the same physical device as the VA simulator and directly communicates with the VA through its APIs, hence, it is considered an embedded deployment.

Intercepting the Skill-User Communication. Eunomia uses the Selenium WebDriver [43] to interact with the VA simulator. The WebDriver accesses the simulator using its "get()" method. From the text eavesdropped with the "`getText()`" method, Eunomia identifies the current skill using the skill name and matches it with the skill ID from the skills store. Even though the VA may misunderstand the voice command and invoke a wrong skill, Eunomia always examines the skill invoked by the VA.

Compliance Validation. Eunomia implements three main methods in compliance validation: (1) The `readPolicy()` method loads and parses the privacy policies to extract the declared privacy practices. It implements the policy analysis algorithms in Section 3.4 and stores the declared practices in subject-action-object tuples. (2) The `readOutput()` method monitors and parses the skills' output obtained from the WebDriver. It implements the BERT-based method (Section 3.5) to extract the actual privacy practices. (3) Finally, `complianceCheck()` validates the consistency between skill's practices and privacy policy's disclosures (Section 3.6). If non-compliance is detected, it invokes the WebDriver to terminate the skill and warn the user. In the embedded deployment, the warning message can display as text as well as play to the user as audio through device's speakers. The sample warning message from Eunomia is shown in Section 2.

Eunomia can also be used to prevent the collection of certain data types regardless of the existence of the privacy policy, for example, in case of parental control for protecting children's data even if the privacy policies declare the data collection practices. This can be achieved by configuring a "master policy" that disallows certain data collection practices and overrides the privacy policies provided by skills. For usability considerations, our prototype also supports an allowlist, i.e., when the user wishes to accept interaction with non-compliant skills, they can explicitly add them to the list. This feature of Eunomia is particularly useful when interacting with emergency skills. As a user's interaction with emergency skills should not be interrupted, the user may add the skills they intend to use in case of emergencies to the allowlist beforehand. Alternatively, Eunomia may also generate a list of emergency skills from Alexa skills store and add it to the allowlist. For emergencies, however, Amazon provides a paid service called "Alexa Emergency Assist"[3] that does not use any skills.

Efficiently Loading Privacy Policies. Downloading the privacy policies may be slow since they are stored on 3rd-party servers. Eunomia caches a SHA256 hash of the recently used policies and the corresponding tuple representation of the declared practices. When a cached skill is invoked, Eunomia first utilizes the cached policy for compliance check while fetching the current policy. It compares the hash of the newly downloaded policy with the cached hash to determine if the policy has changed and updates are needed. For unknown skills, any private data collection is considered non-compliant before the policy is loaded. The downloading and parsing of a new policy takes $\sim$5 seconds. Skills usually do not ask for private data in the first dialogue and allow longer than 5 seconds before they start collecting data.

Reliably Terminating Non-compliant Communication. In the prototype, Eunomia invokes `sendKeys()` of WebDriver to send a "stop" and an "exit" command, and display a warning message to the user. In our experiments, this approach worked well for all non-compliant skills. We measure the time between a non-compliant inquiry (e.g., `tell me your age.`) to the successful termination of the skill (the skill says `Goodbye` and quits). The prototype effectively terminates the skill within 0.5 to 2 seconds (average=1.14 seconds). Eunomia's response time of 1.14 seconds is typically faster than a user's response time, so Eunomia successfully terminates the non-compliant skills before the user responds to the skills with their sensitive information. Eunomia protects its users against non-compliance effectively and timely.

3. https://www.amazon.com/dp/B0BZSZBK3T

As the skills are fully managed by their owners, neither Amazon nor the VA has control over their code or behaviors. Malicious skills may ignore the "stop" command and keep the session open [15] to maliciously capture private information from the user. In defense, we could programmatically disable the element on the developer console by sending `setAttribute('disabled', 'true')`. The element encloses the fields for the user's inputs, therefore, disabling it effectively prevents any input from being delivered to the skill. In our prototype, if a skill does not quit in 2 seconds, Eunomia disables the element.

Other Design Choices. Another deployment option is the standalone model (Fig. 1 (I)), in which Eunomia sits externally on a small computing device, such as a Raspberry Pi or a PC, which is placed near the VA device. Eunomia captures the VA's output using microphone, and converts the audio to text using a speech-to-text transcriber, e.g., DeepSpeech [44]. In case of non-compliance, Eunomia generates a "stop/exit" command, which is played to the VA device using text-to-speech (TTS) conversion. It also generates a warning message as shown in Section 2 that is played to the user using PC's speakers in an audio format. This model does not rely on any support from Amazon or the VA device, and the hardware cost could be easily managed at under $50. However, its performance counts on the accuracy of the speech recognition module. In particular, works in [9], [24] show that it could be challenging to identify skills from the voice command. Such issues could be mitigated with recent advances in speech recognition, e.g., [45], [46].

Local Deployment and Trade-offs. Whether Eunomia is embedded or standalone, it is deployed entirely locally on a user's device. Unless Eunomia is integrated into VAs and distributed by the VA manufacturers, the local deployment involves certain overhead for the users, for example, the users are responsible for setting up the environment, installing Eunomia, managing any future updates, and providing the needed space on the device. However, Eunomia is lightweight and not resource intensive as the pre-trained BERT model is shipped with Eunomia. The advantages of the local implementation outweigh the disadvantages. The local implementation ensures the privacy of data as any logs generated are locally stored and no information is transmitted outside the local devices. Users have full control of the installation and any customization they desire.

Eunomia's Adoption by VAs. In the embedded model (Fig. 1 (II)), Eunomia is implemented inside the VA device and thus has direct access to the communication from the skill to the user. This model requires support from Amazon and/or the device. For instance, Amazon has commercialized the Alexa Voice Service (AVS) Device SDK, so that any embedded implementation employing this SDK will incur a fee. This is the main reason why we developed our prototype using the Alexa simulator. However, VA manufacturers could embed Eunomia at no additional cost as they already have access to the AVS Device SDK and they have incentives to propose privacy-compliant solutions to their users to gain market advantage. Besides potentially embedding Eunomia in the VAs for the users, the VA manufacturers may also use Eunomia externally for themselves to ensure compliance of the skills they recommend with their devices to provide high-quality products to their users.

4. Privacy Compliance Landscape

4.1. Policy Collection and Initial Analysis

Data Collection. We collect all the skills available on the Amazon Alexa website as of November 2023. We develop a Scrapy script to crawl the skills' data, including name, link, developer, description, invocation name, link to privacy policy, languages, etc. We crawl the data for 55,898 skills across all 23 categories. We then follow the privacy policy links to download the available policies.

Missing Privacy Policy Links. 71% (39,628 out of 55,898) of the crawled skills have missing privacy policy links. A missing policy link is acceptable *only if* the skill does not collect any user information. Eunomia pertains to and checks *all 55,898 skills* for compliance including the ones with missing policy links. We further invoke Eunomia to evaluate whether these skills attempt to collect user information (Section 4.4). If the skill attempts to collect any sensitive data from the user, it is non-compliant.

Broken Policy Links. Out of the 16,270 policy links, 2,481 privacy policies fail to download. Among these, 1,634 sites cannot be reached (DNS error or connection timed out), 841 return the page-not-found (404) error, and 6 have non-text policies. It appears that either the developers provided the links as placeholders to get the skills certified by Amazon or their policy pages have moved. In either case, there is a lack of continued policy enforcement from Amazon after the certification process.

Non-English Policies. 356 successfully downloaded policies are non-English, while 151 (42.4%) of them have English as the only supported language. The number of skills with English descriptions but non-English privacy policies is non-trivial and concerning.

Policies with Insufficient Information. We scrape the 13,433 skills with privacy policies in English. However, the presence of a privacy policy does not guarantee that privacy practices are disclosed. For example, the policy of `Gary Cantrell Podcast` only gives an introduction to the skill. The policy page of `Fuller Elementary` only has the placeholder text "Privacy Policy".

Repeated Privacy Policy Links. There is a non-trivial number of privacy policies that repeat for numerous skills. The skills sharing the same policy link do not always have the same owner. It is our speculation that they may be developed by the same contractor or with the same developing tool. We show the 10 most repeated privacy policy links in Table 7 in Appendix B. Some of these links are even broken, e.g., the first link, https://getstoryline.com/public/privacy.html, is used in 669 skills but the site is unreachable (attempted multiple times in 2023 and 2024). getstoryline.com was a platform for the development of Alexa skills which ceased

service. Meanwhile, http://corp.patch.com/privacy, appearing in 597 skills, leads to a page without a policy. This link is used in the news skills created by "Patch.com". Their privacy policy seems to be moved to https://patch.com/privacy, but the skills have not updated the policy link. The link https://www.lottostrategies.com/script/showpage/1001029/b/privacy_policy.html also leads to a 404 error page, while the hosting site is denied by some ISPs. The skills using this link are all lottery information providers created by "Tinbu LLC".

In summary, while Amazon requires the skill developers to provide a privacy policy if the skill collects user information, we find that skills fall short in declaring their practices. Moreover, merely having a privacy policy link does not guarantee full disclosure of the practices. For a large number of skills that provide the policy links, the links appear to be just placeholders to meet Amazon's criteria as they either lead to error pages or dummy policies. Similar observations about privacy policies on mobile apps [22] and IoT apps [39] have been reported. Given the implementation and enforcement of regulations in recent years, the policy landscape of skills seems worse than the IoT/mobile apps. The discrepancy could be the result of how regulations are imposed. Amazon only requires privacy policies for skills that collect personal information. The flexibility could be misused by skill creators.

4.2. Skills' Compliance Using Manual Evaluation

Before we move to conduct a comprehensive, autonomous validation using Eunomia, we select 100 skills and manually investigate their data collection practice and privacy policy disclosures. This analysis provides some preliminary insights about the state of (non)compliance and it also provides examples for understanding violations.

App Selection and Private Information Collection. We pick 100 skills whose names imply that they are likely to need certain types of user information, e.g., the skill name "Metabolic Calculator" implies that it collects users' measurements to calculate the metabolic rate. We manually evaluate each skill in the following steps: (1) we first read the skills' descriptions to gain a basic understanding of the function of the skill. (2) One of the authors and another person independent of this work manually evaluate the skills' interactions and privacy policies. They verbally interact with each skill to trigger as many different interactions as possible and record all the possible information collection attempts. They then share the collected skill interactions and independently compare each interaction with the privacy policy. (3) Each evaluator marks each skill interaction as collecting or not collecting sensitive information. (4) For each interaction that collects sensitive information, the annotator reads the policies for the corresponding skills to determine if the policies declare the data collection practice. The annotator marks the practice as clear, unclear, undisclosed, or inaccurate. (5) The results were compared and discussed. Between the annotators, we observe 3 discrepancies for marking the skills as either collecting or not collecting the sensitive information and 5 discrepancies for determining the disclosure types. The annotators discuss the results and finalize them through mutual consensus.

TABLE 2. SKILLS THAT COLLECT PERSONAL INFORMATION.

Info	App Count	Has Policy	Disclosed	Unclear	Undisclosed	Inaccurate	Rating
Name	16	8	2	1	13	0	3.6
Birthday	6	4	0	0	6	0	4.5
Email	6	4	2	1	3	0	3.2
Health	9	4	0	1	8	0	3.0
Location	13	3	3	0	10	0	3.2
Fitness	13	4	0	4	9	0	3.6
Phone	5	3	0	1	4	0	2.2

Compliance Analysis and Conclusions. 39 out of 100 skills do not collect any sensitive user information. The other 61 skills collect one or more of the seven types of sensitive information: user's names, birthdays, email addresses, phone numbers, health & wellness information, location, and fitness information. Out of the 61 skills, 36 do not have an available privacy policy. The details of the skills that collect user information are shown in Table 2. The sum of the total number of skills against each information collection category in Table 2 is more than 61 because some skills collect multiple types of data. The average user rating of the skills in each category is also provided. In Table 3, we further provide examples of skill descriptions and information collection interactions from two non-compliant skills.

Through manual investigation of these skills, we have the following observations: (1) We observe more frequent compliance violations for some categories, such as `Name`, `Location`, and `Fitness`. For example, 16 skills attempt to collect names, but only 8 skills have privacy policies, while only 2 skills properly disclose this practice. (2) The non-compliant skills are from different manufacturers, hence, compliance oversight appears to be a widespread problem. (3) While one would think that the skills working with more sensitive functions, e.g., healthcare data, are more likely to clearly declare the practices, such speculation is not true according to our observations. (4) The highest-rated non-compliant skills do not even provide privacy policies, which implies that the users generally do not consider privacy disclosures in their ratings. (5) The skill descriptions sometimes imply the information collection practice. However, they do not always disclose all the practices and this is not the legitimate method to declare the practices.

The manual analysis establishes the groundwork for us to further examine the compliance violations using an automated approach. For this purpose, we focus only on the sensitive data types collected by the skills. Examples of the data types are shown in Table 6 of Appendix A.

4.3. Deployment and Performance Evaluation

We deploy the Eunomia prototype on 4 Windows 10 machines to conduct experiments to evaluate all the skills on Amazon skills store. The experiments took approximately 4 months from April 2023 to July 2023. For conversations

TABLE 3. DESCRIPTIONS AND SAMPLE INTERACTIONS FROM NON-COMPLIANT SKILLS.

Info	Example Description (DES) and Skill Action (SKA)	
Birthday	DES	"Open the skill and ask Alexa for one of the pre-defined countdown timers. Alexa will read the into message and the countdown (10, 9, 8, ... 3, 2, 1), and then play a fun sound effect."
	SKA	"Hello! What is your birthday?"
Email	DES	"This Alexa skill is made to assist and notify you during your shopping journey. You can enable notifications by saying 'Turn on notifications' or turn them off by saying 'turn off notifications'."
	SKA	"Please enable permission to access your name and email so that we can get your order details."

with the skills, we simulate user activities with a chatbot [16] to interact with the simulator. Each skill is invoked using the utterances from Amazon skills store and the conversation continues while being monitored by Eunomia. Note that the chatbot was developed by Young et al. in [16]. It is not a contribution of this paper.

Efficiency, Precision, and Recall. Eunomia's *efficiency* is measured by the time lag between the receipt of a non-compliant query and the confirmation of the terminated skill. In the example from Section 2, the time lag is measured between "Say your height and weight." is delivered and "Good Bye" is received. Eunomia's reaction records to be between 0.5 and 2 seconds with an average of 1.14 seconds (depending on the size of interactions) which makes it a real-time defense tool.

The *precision* is defined as the proportion of correctly identified disclosures (clear, unclear, undisclosed, and inaccurate) out of all reported disclosures. In the evaluation, we randomly select 100 reported disclosures from each of the clear, unclear, and undisclosed categories. We also select 30 inaccurate disclosures. The selected cases are manually validated by interacting with the skills and examining the privacy policies. 100 clear and 100 undisclosed practices are correctly identified (precision of 100%), 96 unclear disclosures and 29 inaccurate disclosures are correctly identified (precisions of 96% and 96.7%), respectively. In summary, Eunomia achieves precision of 96% to 100% in identifying different types of compliant and non-compliant disclosures.

Finally, the *recall* is defined as the proportion of correctly detected information collection actions out of all information collections (both compliant and non-compliant). We measure Eunomia's recall and compare with two SOTA privacy compliance detectors, SkillExplorer [15] and SkillDetective [16]. SkillExplorer assembled a list of 100 skills that collected private information in 2020, while SkillDetective adopted the list and found that 61 skills still collected private information in 2022. The authors of SkillExplorer and SkillDetective kindly shared their lists with us. With manual inspection, we identified that only 27 skills were still collecting private data in May 2023. As reported in Table 1, Eunomia correctly detects 36 out of 37 private information collection actions (recall with respect to data collection practices = 97.3%) in 26 out of the 27 skills (recall with respect to number of skills = 96.3%). We report the recall with respect to the number of skills to be consistent with other comparative studies in Table 1. We also measure Eunomia's recall on our dataset from Section 4.2. Out of 61 skills that collect sensitive data, Eunomia detects 59 skills correctly. We find that the actions which Eunomia does not capture lack proper sentence structure or are unclear about what they ask for (see *False Negatives* below).

False Positives and Negatives. A *false negative* is a non-compliant sensitive data collection query from the skill that is *not* detected/blocked by Eunomia. We observe that when the skill responses are ambiguous and not properly structured, Eunomia may misidentify such practice. For example, without the context, Eunomia did not correctly detect the skill's question "where to?" or "what city?"

A *false positive* is a complaint action (usually an action that does not collect sensitive data) that is misidentified by Eunomia as non-compliant. For example, a skill question, "If you could be any age what age would you want to be?" was identified as collecting sensitive information (and undisclosed). Eunomia was unable to detect that the skill is not asking for the actual age.

4.4. Skills' Compliance Landscape

To measure the effectiveness of Eunomia in a large-scale experiment, we interact automatically with all the 55,898 Alexa skills on the Amazon skills store. The privacy disclosures of the skills identified by Eunomia are shown in Table 4, as grouped by sensitive data categories, e.g., the "health" category contains weight, height, blood group, blood pressure, medication, and body weight. The miscellaneous category has calendar, account information, device information, race, income, etc. A skill may have multiple disclosure types, e.g., a skill may ask for "age" and "gender" but only disclose "age" in the policy. In this case, we count it in both clear and undisclosed categories. Through automated evaluation, we have the following observations.

(1) *Only a small portion of private data collections are clear.* Out of all the data collection practices that request private data from the users, only 25.9% are correctly disclosed.

(2) *7.2% of the data collection practices are compliant but unclear.* They ask the users for specific private data, such as name and age, but they are not specific in their policies about the data type. Most of them mention in generic terms that they collect 'personal information' from the users.

(3) *Majority of the data collection practices are non-compliant.* 64.3% of the data collection practices are completely undisclosed in the privacy policies. We also find 37 non-compliant practices to be inaccurate, that is, the privacy policy claims that the skill does not collect the specific data but the skill asks the user for that data.

(4) *Non-compliances are mostly caused by missing privacy policies.* The 903 undisclosed data collection practices belong to 622 skills. Only 129 out of 622 skills have a privacy policy link. The majority of the non-compliance sources are from missing privacy policies.

TABLE 4. NUMBER OF COMPLIANT AND NON-COMPLIANT SKILLS AND DISCLOSURES IDENTIFIED BY EUNOMIA.

Data Types	Compliant		Non-Compliant	
	Clear	Unclear	Undisc.	Inaccurate
Name	97 (110)	5 (9)	118 (129)	4 (4)
Address	19 (21)	2 (2)	37 (38)	2 (2)
Age	5 (5)	9 (11)	21 (25)	2 (2)
Birthday	4 (4)	12 (13)	41 (51)	3 (3)
Email	63 (67)	4 (5)	24 (28)	2 (2)
Gender	0 (0)	0 (0)	1 (2)	0 (0)
Payment	4 (4)	2 (2)	4 (4)	0 (0)
Phone	36 (37)	6 (6)	29 (30)	4 (5)
Health	4 (10)	2 (2)	19 (28)	0 (0)
Location	69 (93)	22 (27)	196 (306)	10 (16)
Fitness activity	6 (9)	14 (21)	190 (248)	1 (2)
Weather	3 (3)	1 (1)	7 (10)	0 (0)
Miscellaneous	1 (1)	2 (2)	4 (4)	1 (1)
Total Disclosure #	225 (364)	76 (101)	622 (903)	25 (37)
Disclosure Rate (%)	24.2 (25.9)	8.2 (7.2)	66.8 (64.3)	2.7 (2.6)

Total Skills Tested by Eunomia	55,898
Skills w. Sensitive Practices Identified by Eunomia	931
Total Sensitive Practices Captured by Eunomia	1405

X (Y): number of skills (number of data collection practices); Total Disclosure #: number of skills and data collection practices in each disclosure category; Disclosure Rate: fraction of skills and data collection practices in each disclosure category out of all skills (931) and practices (1405), respectively.

4.5. Case Studies and Discussion

A. Compliance Issues of Types of Sensitive Data. Table 5 shows case studies of different types of sensitive data collection practices and their disclosure status.

Birthdays. Only 4 out of 71 birthday collections (5.6%) are clearly disclosed in the privacy policies. The majority of the birthday-related data collections are either undisclosed or disclosed in generic terms. Many skills collecting birthdays are designed to deliver birthday wishes. They usually have a missing policy but their descriptions imply the collection of birthdays. The authors did not hide the data collection practice, rather, they might have been unaware of the disclosure requirements. To address this, the authors need to realize that collecting personal data even for entertainment purposes should be disclosed in privacy policies. Declarations in descriptions are not the legitimate means for declaration.

Health Data. Health-related data is particularly sensitive, however, Eunomia finds that 70% (28/40) of the health-related data collections are non-compliant. These skills are intended to help with the users' health issues such as obesity, metabolism, and growth issues, yet the privacy policy declarations are missing. We notice that the skill descriptions often imply the collection of health data, but declarations in descriptions alone do not make these skills compliant.

Fitness Activities. A vast majority, over 89% (i.e., 250/280), of all the fitness activity disclosures are non-compliant. It is worth noting that about half of these skills fall under the Health & Fitness category on Amazon Skills Store and one would expect the policy declarations to be taken seriously. Again, the skill descriptions are detailed. The developers put effort into writing the descriptions and they solely rely on that even when the regulations require privacy policies to be the standard for the declaration of practices.

Location, Address, Email, and Phone. Among all the data types, location data has the highest number of non-compliant data collection practices (322). Other notable non-compliance issues are email and phone numbers with about 30% (30/102) and 45% (35/78) non-compliant practices, respectively. Skills collect location information for various reasons, e.g., to find local businesses or report weather. Location is the most commonly ignored practice in privacy policies. The reason could be that the developers may have neglected the sensitive nature of the location data.

Name Disclosures. Another personal data type that skills frequently fail to disclose is the name, with 53% (133/252) non-compliant data collections. The name-collecting skills mostly belong to the gaming category and authors may not find it necessary to declare practices in this case, since many of them are just expecting nicknames from the gamers.

From analyzing the non-compliant disclosures, we conclude that the most important solution to addressing the missing policies/declarations is to increase awareness and provide education to the authors so they better understand the need for declaring the practices in the privacy policies.

B. Examples of Compliance Issues. For a skill to be compliant, it must declare its practices in the privacy policy that is accessible to its users as required by Amazon (Section 3.1). This compliance requirement aligns with compliance regulations such as the General Data Protection Regulation (GDPR) [13] and California Privacy Rights Act (CPRA) [47]. Though we focus on the skills in the US market in the Eunomia study, besides Amazon skills store regulations, we draw references from GDPR because it is recognized as a comprehensive compliance regulation in many countries and US regulatory bodies are now adapting from GDPR [48]. The examples discussed below are either non-compliant or compliant but unclear.

- **Blood Donation Helper** [49] skill collects health-related data, specifically, the blood group information from its users to match blood donors with recipients.
- **Workout Coach** [50] skill provides guidance to users with their exercises and it personalizes the workouts for them.
- **Oil Analysis** [51] asks the users about their names to verify the Oil Analysis account.
- **Boston Bike** [52] inquires the users about their address to find the bike stations close to them.
- **Body Mass Index** [53] asks the users their health-related information to calculate the body mass index.

Eunomia identifies Blood Donation Helper, Workout Coach, and Oil Analysis as compliant but unclear. Their policies disclose the practices in generic terms such as "personal data" instead of "birthday" or "weight", etc. Unclear policies are often umbrella policies that cover a number of skills from the same owner or contracted developer. Unclear declarations do not give the user a clear understanding of their data collection. They are concerning especially when the practice is about highly sensitive data.

The privacy policy of Boston Bike states that it does not collect any data. Eunomia finds the policy inaccurate because the skill collects the address. The skill Body Mass

TABLE 5. CASE STUDIES OF DIFFERENT TYPES OF SENSITIVE DATA COLLECTION.

Case 1: Birthday	Privacy Policy	"Can access usage information ..., which geographical regions they are located in, ... users' language, device type, and length and frequency of use."
Undisclosed	Skill Action	"What is your birthday?"
Case 2: Health	Privacy Policy	The privacy policy page does not have privacy policy content.
Undisclosed	Skill Action	"Please, tell me your weight in kilograms."
Case 3: Address	Privacy Policy	"We don't store your data, period. Our iPhone and Android mobile apps, our Amazon Echo Skills, our Facebook, Twitter and other Bots do not store any personal data."
Inaccurate	Skill Action	"This skill requires your address in order to find the closest bike stations."

Index has practices that are undisclosed in the privacy policy. Eunomia marks these skills as non-compliant.

Responsible Disclosure. The identified non-compliant issues were disclosed to the skill developers who provided contact information. We notified 25 manufacturers via email and 3 via web forms in January 2024. Within a week, 3 manufacturers acknowledged the receipt of the disclosure, while one confirmed the creation of a ticket to fix the issue. We visited the skill pages and privacy policies in February and May 2024. We noticed only one change. The manufacturer who created a ticket to update the incomplete policy now has a broken policy link on the skill page.

5. Research Questions and Discussion

The development and evaluation of Eunomia helps us understand the effectiveness of Eunomia and the landscape of Alexa skill compliance. Here we answer the questions presented in Section 1.

1. *How effective is the Eunomia firewall in protecting the users?*

Answer: Eunomia successfully provides defense in real-time by stopping a skill's interaction in an average of 1.14 seconds a non-compliant output is presented to the user. With a recall of 96.3%, Eunomia is very effective in identifying the actions that collect private information. It also accurately identifies the disclosure types for these actions with a precision of 96-100%.

2. *What is the overall compliance status of the Alexa skills?*

Answer: The majority of the data collection practices containing sensitive data are non-compliant (940 out of 1405), where the skill actions are not disclosed in the privacy policies or there are contradictions in the disclosure. While these skills collect sensitive user data, many of them do not have an available privacy policy. This relatively poor overall compliance status of Alexa skills shows a need for Amazon to enhance the certification process and a need for stricter enforcement of user data protection regulations.

3. *Which particular compliance gaps are there in Alexa skills?*

Answer: Most of the compliance gaps are caused by undisclosed practices, which mainly result from missing policies. Inaccurate disclosures are rare, with only 37 occurrences in 1405 data collections. There also exists a non-trivial proportion of unclear disclosures. Out of all the data collection practices, 7.2% are compliant but unclear. A probable reason for commonly occurring unclear disclosures could be that Amazon does not specifically require the policy disclosures to be in clear terms. According to Amazon, as long as the privacy policy has practice disclosures for a skill, the skill's practices are considered compliant.

4. *Which type of compliance issues are more common than others?*

Answer: Certain data types are observed to be more commonly neglected in the privacy policies; in particular, location-related and fitness activity violations occur most frequently. Names are also more likely to be omitted.

Takeaways and Eunomia's Role. In summary, our measurement study shows that there exists a significant amount of compliance violations in Alexa skills that need to be addressed by the platform owner and by the regulations. Amazon has taken the preliminary steps by carefully outlining the requirements for privacy policy declarations in case of sensitive data collection from the users. Yet, we notice that a non-trivial number of skills do not abide by the regulations set forth by Amazon. Amazon needs to enforce its compliance requirements in a stricter manner. In particular, Amazon should have more checks in place that prevent the bypassing of its compliance requirements. Skill developers also need to responsibly declare all the specific data types their skills collect from the users. The developers should carefully follow the privacy requirements of Amazon during publishing and maintenance of the skills. Eunomia can help the regulators and the developers through this process by validating the skills in real-time. Real-time validation is the only effective solution to protect users against non-compliant skills. This approach always works even in the cases where the other existing code analysis and fuzzing compliance-validation techniques fall short, such as, when developers change a previously compliant skill version to a non-compliant version or when fuzzing techniques try to trigger all interactions to provide defense but this approach is resource intensive, time-consuming and not exhaustive.

Limitations. Eunomia is only a compliance validator, *not* a private information shield, i.e., it only validates private information collections against privacy policies, but it does not validate the rationale of the private information collection. That is, a skill may collect an excessive amount of personal information that is not necessary for its functions (e.g., a clock skill asking for birthday), as long as the practices are disclosed, Eunomia (as well as other compliance validation tools [15], [16]) considers the skill compliant. It is in our future plan to identify whether a specific permission or a piece of information is indeed necessary for a skill to perform its service. We study the skills from the US Amazon skills store in Eunomia but we refer to foreign regulations

such as GDPR besides the regulations from Amazon as they both align in their compliance requirements. Moreover, GDPR influences the regulatory bodies in the US.

6. Conclusion

We present Eunomia, a real-time privacy compliance validation tool for Alexa skills. We first analyze the privacy policies using a fine-tuned BERT model, dependency-based parse trees, synonyms, and ontologies. We capture and analyze the data collection queries from the skills, and validate the actions against the privacy policies. Eunomia provides a real-time defense to the users against non-compliant skills by stopping the interaction as soon as non-compliance is detected. We evaluate Eunomia using all the Alexa skills from Amazon skills store and find that Eunomia detects and stops non-compliant actions in an average of 1.14 seconds. The evaluation also leads to a compliance landscape of Alexa skills. We find that a vast majority of all the private data collection practices are non-compliant.

Acknowledgements

Bo Luo and Fengjun Li were supported in part by NSF IIS-2014552, DGE-1565570, and the Ripple University Blockchain Research Initiative. The authors would like to thank the anonymous reviewers and the shepherd for their valuable comments and suggestions. We would also like to thank the authors of SkillExplorer [15] and SkillDetective [16] for generously sharing their data with us.

References

[1] N. Carlini, P. Mishra, T. Vaidya, Y. Zhang, M. Sherr, C. Shields, D. Wagner, and W. Zhou, "Hidden voice commands," in *25th USENIX security symposium (USENIX security 16)*, 2016, pp. 513–530.

[2] W. Diao, X. Liu, Z. Zhou, and K. Zhang, "Your voice assistant is mine: How to abuse speakers to steal information and control your phone," in *Proceedings of the 4th ACM Workshop on Security and Privacy in Smartphones & Mobile Devices*, 2014, pp. 63–74.

[3] M. B. Hoy, "Alexa, siri, cortana, and more: an introduction to voice assistants," *Medical reference services quarterly*, vol. 37, no. 1, pp. 81–88, 2018.

[4] D. Mukhopadhyay, M. Shirvanian, and N. Saxena, "All your voices are belong to us: Stealing voices to fool humans and machines," in *European Symposium on Research in Computer Security*. Springer, 2015, pp. 599–621.

[5] S. M. S. Talebi, A. A. Sani, S. Saroiu, and A. Wolman, "Megamind: a platform for security & privacy extensions for voice assistants," in *Proceedings of the 19th Annual International Conference on Mobile Systems, Applications, and Services*, 2021, pp. 109–121.

[6] T. Vaidya, Y. Zhang, M. Sherr, and C. Shields, "Cocaine noodles: exploiting the gap between human and machine speech recognition," in *9th USENIX Workshop on Offensive Technologies (WOOT 15)*, 2015.

[7] F. Sharevski, P. Jachim, P. Treebridge, A. Li, A. Babin, and C. Adadevoh, "Meet malexa, alexa's malicious twin: Malware-induced misperception through intelligent voice assistants," *International Journal of Human-Computer Studies*, vol. 149, p. 102604, 2021.

[8] D. J. Dubois, R. Kolcun, A. M. Mandalari, M. T. Paracha, D. Choffnes, and H. Haddadi, "When speakers are all ears: Characterizing misactivations of iot smart speakers," *Proceedings on Privacy Enhancing Technologies*, vol. 2020, no. 4, 2020.

[9] D. Kumar, R. Paccagnella, P. Murley, E. Hennenfent, J. Mason, A. Bates, and M. Bailey, "Skill squatting attacks on amazon alexa," in *27th USENIX security symposium (USENIX Security 18)*, 2018, pp. 33–47.

[10] L. Cheng, C. Wilson, S. Liao, J. Young, D. Dong, and H. Hu, "Dangerous skills got certified: Measuring the trustworthiness of skill certification in voice personal assistant platforms," in *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security*, 2020, pp. 1699–1716.

[11] J. S. Edu, X. Ferrer-Aran, J. Such, and G. Suarez-Tangil, "Skillvet: automated traceability analysis of amazon alexa skills," *IEEE Transactions on Dependable and Secure Computing*, vol. 20, no. 1, pp. 161–175, 2021.

[12] W. Seymour, M. Cote, and J. Such, "When it's not worth the paper it's written on: A provocation on the certification of skills in the alexa and google assistant ecosystems," in *Proceedings of the 4th Conference on Conversational User Interfaces*, 2022, pp. 1–5.

[13] P. Voigt and A. Von dem Bussche, "The eu general data protection regulation (gdpr)," *A Practical Guide, 1st Ed., Cham: Springer International Publishing*, vol. 10, p. 3152676, 2017.

[14] J. Edu, X. Ferrer-Aran, J. Such, and G. Suarez-Tangil, "Measuring alexa skill privacy practices across three years," in *Proceedings of the ACM Web Conference 2022*, 2022, pp. 670–680.

[15] Z. Guo, Z. Lin, P. Li, and K. Chen, "Skillexplorer: Understanding the behavior of skills in large scale," in *29th USENIX Security Symposium (USENIX Security 20)*, 2020, pp. 2649–2666.

[16] J. Young, S. Liao, L. Cheng, H. Hu, and H. Deng, "Skilldetective: Automated policy-violation detection of voice assistant applications in the wild," in *USENIX Security Symposium*, 2022.

[17] S. Liao, C. Wilson, L. Cheng, H. Hu, and H. Deng, "Measuring the effectiveness of privacy policies for voice assistant applications," in *Annual Computer Security Applications Conference*, 2020, pp. 856–869.

[18] E. McCallister, *Guide to protecting the confidentiality of personally identifiable information*. Diane Publishing, 2010, vol. 800.

[19] T. H. Journal, "What is considered phi under hipaa?" https://www.hipaajournal.com/considered-phi-hipaa/, 2023, accessed: 2023-10.

[20] B. Andow, S. Y. Mahmud, J. Whitaker, W. Enck, B. Reaves, K. Singh, and S. Egelman, "Actions speak louder than words: Entity-sensitive privacy policy and data flow analysis with policheck," in *Proceedings of the 29th USENIX Security Symposium (USENIX Security'20)*, 2020.

[21] B. Andow, S. Y. Mahmud, W. Wang, J. Whitaker, W. Enck, B. Reaves, K. Singh, and T. Xie, "Policylint: Investigating internal privacy policy contradictions on google play." in *USENIX Security Symposium*, 2019, pp. 585–602.

[22] S. Zimmeck, P. Story, D. Smullen, A. Ravichander, Z. Wang, J. R. Reidenberg, N. C. Russell, and N. Sadeh, "Maps: Scaling privacy compliance analysis to a million apps," *Proc. Priv. Enhancing Tech.*, vol. 2019, p. 66, 2019.

[23] F. Xie, Y. Zhang, C. Yan, S. Li, L. Bu, K. Chen, Z. Huang, and G. Bai, "Scrutinizing privacy policy compliance of virtual personal assistant apps," in *37th IEEE/ACM International Conference on Automated Software Engineering*, 2022, pp. 1–13.

[24] N. Zhang, X. Mi, X. Feng, X. Wang, Y. Tian, and F. Qian, "Dangerous skills: Understanding and mitigating security risks of voice-controlled third-party functions on virtual personal assistant systems," in *2019 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2019, pp. 1381–1396.

[25] C. Lentzsch, S. J. Shah, B. Andow, M. Degeling, A. Das, and W. Enck, "Hey alexa, is this skill safe?: Taking a closer look at the alexa skill ecosystem," *Network and Distributed Systems Security (NDSS) Symposium2021*, 2021.

[26] T. Le, D. Zhao, Z. Wang, X. Wang, and Y. Tian, "Alexa, is the skill always safe? uncover lenient skill vetting process and protect user privacy at run time," in *Proceedings of the 46th International Conference on Software Engineering: Software Engineering in Society*, 2024, pp. 34–45.

[27] M. Honnibal and I. Montani, "spacy 2: Natural language understanding with bloom embeddings, convolutional neural networks and incremental parsing," *To appear*, vol. 7, no. 1, pp. 411–420, 2017.

[28] P. Mallojula, F. Li, X. Du, and B. Luo, "Companion apps or backdoors? on the security of automotive companion apps," in *European Symposium on Research in Computer Security*. Springer, 2024, pp. 24–44.

[29] Y. Zhao, L. Yu, Y. Sun, Q. Liu, and B. Luo, "No source code? no problem! demystifying and detecting mask apps in ios," in *Proceedings of the 32nd IEEE/ACM International Conference on Program Comprehension*, 2024, pp. 358–369.

[30] Y. Nan, X. Wang, L. Xing, X. Liao, R. Wu, J. Wu, Y. Zhang, and X. Wang, "Are you spying on me?{Large-Scale} analysis on {IoT} data exposure through companion apps," in *32nd USENIX Security Symposium (USENIX Security 23)*, 2023, pp. 6665–6682.

[31] A. Oltramari, D. Piraviperumal, F. Schaub, S. Wilson, S. Cherivirala, T. B. Norton, N. C. Russell, P. Story, J. Reidenberg, and N. Sadeh, "Privonto: A semantic framework for the analysis of privacy policies," *Semantic Web*, vol. 9, no. 2, pp. 185–203, 2018.

[32] R. Trimananda, H. Le, H. Cui, J. T. Ho, A. Shuba, and A. Markopoulou, "{OVRseen}: Auditing network traffic and privacy policies in oculus {VR}," in *31st USENIX security symposium (USENIX security 22)*, 2022, pp. 3789–3806.

[33] W. B. Tesfay, P. Hofmann, T. Nakamura, S. Kiyomoto, and J. Serna, "Privacyguide: towards an implementation of the eu gdpr on internet privacy policy evaluation," in *ACM Workshop on Security and Privacy Analytics*, 2018, pp. 15–21.

[34] S. Zimmeck, Z. Wang, L. Zou, R. Iyengar, B. Liu, F. Schaub, S. Wilson, N. Sadeh, S. Bellovin, and J. Reidenberg, "Automated analysis of privacy requirements for mobile apps," in *AAAI Fall Symposium*, 2016.

[35] H. Harkous, K. Fawaz, R. Lebret, F. Schaub, K. G. Shin, and K. Aberer, "Polisis: Automated analysis and presentation of privacy policies using deep learning," in *27th USENIX Security Symposium (USENIX Security 18)*, 2018, pp. 531–548.

[36] E. Okoyomon, N. Samarin, P. Wijesekera, A. Elazari Bar On, N. Vallina-Rodriguez, I. Reyes, Á. Feal, S. Egelman *et al.*, "On the ridiculousness of notice and consent: Contradictions in app privacy policies," in *Workshop on Technology and Consumer Protection (ConPro 2019), in conjunction with the 39th IEEE Symposium on Security and Privacy*, 2019.

[37] H. Cui, R. Trimananda, A. Markopoulou, and S. Jordan, "{PoliGraph}: Automated privacy policy analysis using knowledge graphs," in *32nd USENIX Security Symposium (USENIX Security 23)*, 2023, pp. 1037–1054.

[38] J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: pre-training of deep bidirectional transformers for language understanding," *CoRR*, vol. abs/1810.04805, 2018. [Online]. Available: http://arxiv.org/abs/1810.04805

[39] J. Ahmad, F. Li, and B. Luo, "IoTPrivComp: A measurement study of privacy compliance in iot apps," in *Computer Security–ESORICS 2022: 27th European Symposium on Research in Computer Security, Copenhagen, Denmark, September 26–30, 2022, Proceedings, Part II*. Springer, 2022, pp. 589–609.

[40] R. Ungureanu, "Certification requirements for privacy policy URLs," https://developer.amazon.com/en-US/blogs/alexa/alexa-skills-kit/2023/02/certification-requirements-privacy-urls-feb-2023, 2023, online; accessed 12 September 2024.

[41] T. Wolf, L. Debut, V. Sanh, J. Chaumond, C. Delangue, A. Moi, P. Cistac, T. Rault, R. Louf, M. Funtowicz *et al.*, "Transformers: State-of-the-art natural language processing," in *Proceedings of the 2020 conference on empirical methods in natural language processing: system demonstrations*, 2020, pp. 38–45.

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

[43] Selenium, "Selenium WebDriver," https://www.selenium.dev, 2023, accessed: 2023-02.

[44] R. Morais, "DeepSpeech 0.6: Mozilla's speech-to-text engine gets fast, lean, and ubiquitous," https://hacks.mozilla.org/2019/12/deepspeech-0-6-mozillas-speech-to-text-engine/, 2023, accessed: 2023-02.

[45] T. Kawase, M. Okamoto, T. Fukutomi, and Y. Takahashi, "Speech enhancement parameter adjustment to maximize accuracy of automatic speech recognition," *IEEE Transactions on Consumer Electronics*, vol. 66, no. 2, pp. 125–133, 2020.

[46] Y. Shi, Q. Huang, and T. Hain, "Robust speaker recognition using speech enhancement and attention model," *arXiv preprint arXiv:2001.05031*, 2020.

[47] California Legislature, "California Privacy Rights Act of 2020," https://thecpra.org, 2020.

[48] W. G. Voss and K. A. Houser, "Personal data and the gdpr: providing a competitive advantage for us companies," *American Business Law Journal*, vol. 56, no. 2, pp. 287–344, 2019.

[49] Leslie Correa, "Blood donation helper," https://www.amazon.com/Leslie-Correa-Blood-donation-helper/dp/B07N626993, 2024, online; accessed 10 September 2024.

[50] Voice Craft LLC, "Workout Coach," https://www.amazon.com/Voice-Craft-LLC-Workout-Coach/dp/B07FKZ8GN5, 2024, online; accessed 10 September 2024.

[51] 1stWave, "Oil Analysis," https://www.amazon.com/1stWave-Oil-Analysis/dp/B082J4DQZN, 2024, online; accessed 10 September 2024.

[52] lbrenman, "Boston Bike," https://www.amazon.com/lbrenman-Boston-Bike/dp/B076HHMLL8, 2024, online; accessed 10 September 2024.

[53] Luis TM, "Body Mass Index," https://www.amazon.com/Luis-TM-Body-Mass-Index/dp/B07ZKXJ4V8, 2024, online; accessed 10 September 2024.

Appendix A. Sensitive Data Types

Table 6 shows some examples of sensitive data types that are considered in the Eunomia approach. The mechanism used to identify such data and build the ontology is presented in Section 3.3.

Appendix B. Additional Examples

Table 7 presents the most repeated privacy policies found in the Amazon skills store. Many of them appear to be umbrella policies that are not specific to any skill. The

TABLE 6. EXAMPLES OF USER DATA TYPES.

Data Types
'name', 'phone number', 'address', 'SSN', 'email address', 'age', 'gender', 'birthday', 'medical record number', 'health plan beneficiary number', 'driver license number', 'ethnicity', 'zip code', 'bank account number', 'health and wellness', 'social media information', 'geographical location', 'payment', 'vehicle information', 'salary', 'vehicle identification number', 'fitness activity information', 'employment information', 'body weight', 'credit history', 'blood glucose level', 'heart rate', 'body mass index', 'financial account information', 'blood group', 'blood pressure', 'payment'

skills sharing the same policy link do not always appear to have the same owner. It is our speculation that they may be developed by the same contractor or with the same developing tool or framework.

- The first link, https://getstoryline.com/public/privacy.html, is used in 669 skills but the site is unreachable (attempted multiple times in 2023 and 2024). getstoryline.com was a platform for the development of Alexa skills. It seems that when the development platform ceased service, the skills were left with a broken privacy policy link.
- The second policy link, http://corp.patch.com/privacy, appearing in 597 skills, leads to a page without a policy. This link is used in the news skills created by "Patch.com". The url https://patch.com is still functional and their privacy policy seems to be moved to https://patch.com/privacy, but the skills have not updated the previous privacy policy link.
- The third link, https://www.lottostrategies.com/script/showpage/1001029/b/privacy_policy.html, also leads to a 404 error page, while the hosting site is denied by some ISPs. The skills using this link are all created by "Tinbu LLC". These skills provide lottery information. The privacy policy for https://www.lottostrategies.com has the link https://www.lottostrategies.com/cgi-bin/showpage/1001029/b/privacy_policy.html, but the skills have not updated to the functional URL.
- The link https://www.advicelocal.com/privacy-policy/ appears in education-related skills created by either "Voice Advice Application" or "Voice App Developer".
- https://radio.co/terms/alexa appears in music and radio related skills created by "Online Radio".
- All the skills using the policy link, https://www.govocal.ai/privacypolicy.html, are created by "GoVocal.AI". The site is unreachable (attempted in May and August 2023 and April 2024).

TABLE 7. TOP 10 MOST REPEATED PRIVACY POLICIES.

Privacy Policy Links	Background	Count
https://getstoryline.com/public/privacy.html	Site unreachable, various skill types and creators, development platform seems updated without updating the policy link	669
http://corp.patch.com/privacy	news skills created by "Patch.com"	597
http://www.lottostrategies.com/script/showpage/1001029/b/privacy_policy.html	404 page, lottery skills created by "Tinbu LLC"	263
https://www.advicelocal.com/privacy-policy/	Education-related skills created by either "Voice Advice Application" or "Voice App Developer"	239
https://radio.co/terms/alexa	Music and radio skills created by "Online Radio"	185
https://www.govocal.ai/privacypolicy.html	Site unreachable, skills created by "GoVocal.AI"	122
http://www.newsbreakapp.com/privacy	News-related skills from the creator "Particle Media Inc"	109
https://skilexa.com/doneforyou/privacy/	Platform to help create flash briefings and increase traffic, various skill categories and creators	100
https://cir.st/privacy-policy	The entity provides streaming services, streaming-related skills by various creators	90
https://creator.voiceflow.com/creator/terms?name=Nicola&skill=Random%20Sleep%20Sounds&children=false	Music/sound skills created by either "FeRUE" or "Smart Skills"	84