In proceedings of IEEE International Conference on Trust, Privacy and Security in Intelligent Systems, and Applications (IEEE TPS 2023)

Ensuring Privacy Policy Compliance of Wearables with IoT Regulations

Kelvin Uzoma Echenim
Information Systems
University of Maryland Baltimore County
Baltimore, MD, USA
kelvine1@umbc.edu

Lavanya Elluri
Computer Information Systems
Texas A&M University - Central Texas
Killeen, TX, USA
elluri@tamuct.edu

Karuna Pande Joshi
Information Systems
University of Maryland Baltimore County
Baltimore, MD, USA
karuna.joshi@umbc.edu

***Abstract*—In an era where wearables, particularly those in non-hospital settings, collect and transmit sensitive personal data, it is imperative to implement stringent privacy safeguards. The National Institute of Standards and Technology (NIST) Internal Report 8228 provides regulations for securing Internet of Things (IoT) devices, data, and the privacy of individuals. We have developed a novel framework for examining the privacy policies governing the data and information utilized by wearable devices to ensure that these IoT devices work in adherence to the NIST controls. Our approach entails constructing an ontology of the pertinent NIST regulations, extracting key regulation terms, establishing clear annotation guidelines, and reasoning over the developed ontology. Our primary contribution is developing a novel method to accurately retrieve the expectations, privacy risk mitigation areas, and the associated regulations using Natural Language Processing and Semantic Web concepts. Ultimately, vendors and users can use our publicly available ontology to semi-automate the privacy compliance process for wearables, ensuring that the data collected and transmitted through the devices are secure, thereby protecting both the devices and the individuals who use them.**

***Index Terms*—Wearables, Wearable Devices, IoT, Semantic Web, Ontology, Privacy, Compliance, Microservices**

I. Introduction

The widespread usage and data transmission, as well as the heightened privacy risks associated with microservices of wearable devices, indicate an apparent need for compliance with existing regulations. For example, an investigation by [1] revealed that many partner applications compatible with the Fitbit fitness tracker transmit personal information, such as geolocation and email addresses, to unintended third parties. This group comprises social networks, analytic and advertising services, and weather-related Application Programming Interfaces. This finding suggests that wearable devices may incorporate microservices that are primarily focused on Personally Identifiable Information (PII), with a real-time collection of sensitive PII, including precise location coordinates and health-specific data [2] [3] [4]. The health and fitness metrics tracked by wearable devices are interconnected with some microservices. Among these services are monitoring of sleep patterns, frequency of standing, movement or walking speed as a measure of physical fitness, heart rate, stress levels, nutritional intake, medication adherence, respiratory rate, glucose levels, menstrual cycle, and mental health [2] [5] [6] [7]. Predominantly, health data is derived from wearables operated outside hospital settings, evading regulations such as the Health Insurance Portability and Accountability Act (HIPAA). According to [8], wearable device manufacturers circumvent HIPAA's "business associate" provisions by not advocating that device users share their personal health information with medical professionals. While our study primarily focuses on privacy policies, the reference to microservices highlights the complexity and granularity of wearable device data management. The intricate data management involved with wearable devices can have implications for privacy policies, particularly concerning the handling and protecting of PII [9] [10].

Wearable devices are widely available for purchase and are used by individuals daily. The fundamental components that drive the business cases for wearable microservices are the PII. Therefore, for users to benefit from or utilize these microservices optimally, they must divulge their PII, posing privacy risks. Such information is frequently transmitted across cloud service meshes, dispersed across numerous networks, many of which are public and where cybersecurity and privacy breaches are prevalent. Unfortunately, users are often unconcerned about the amount of data and information collected by these devices and how the wearable vendors have utilized them without their notice or knowledge. To check this problem, the National Institute of Standards and Technology (NIST) Internal Report (IR) 8228 [11] provides standards to ensure IoT device, data, and individuals' privacy. Often, wearable device manufacturers overlook these guidelines that protect the IoT devices, the data they collect, and the individuals that use them. Therefore, many individuals report privacy concerns after using wearable devices.

We have developed a novel framework for examining the privacy policies that control the data and information utilized on wearables to determine if they conform to the rules outlined in NIST regulation. Because the datasets are available to the public, this study concentrates on wearable devices in non-hospital settings. Our authority document for verification is NISTIR 8228. It has stipulated considerations and mitigation for protecting the data that pass through the IoT devices. Our essential contribution is the novel approach to semi-automatically extract the expectations, mitigation rules, and associated regulations using relevant strategies from Natural Language Processing and Semantic Web technologies. Our framework is useful for both the vendors, who manufacture these wearables, and the users, who wear these devices, to become aware of the binding IoT regulations. Ultimately, it gives them confidence that the personal data passed through the devices are secure.

II. Background

A. WEARABLES - THE MICROSERVICES AND PRIVACY RISKS

IoT platforms originally used monolithic approaches for system design and implementation and for the delivery of IoT services [12]. Microservices are small, independent software components that provide specific functionality to wearable IoT devices in the context of wearables. They are designed to be lightweight and modular and can be combined to create complex applications for wearable devices. Microservices can collect data from sensors on the device, process data, provide user interfaces, and communicate with other devices or services, among other tasks. Microservices for wearables may be deployed on the wearable device or on a cloud-based platform. They can improve wearable applications' performance, flexibility, and scalability, allowing developers to create more customized and individualized user experiences [12]. In recent years, there has been a rapid increase in the number of intelligent wearable products adapted for various applications. Smartwatches, wristbands, eyewear, headsets, earbuds, body straps, foot and hand-worn devices, and smart jewelry are the wearables developed for various applications [13]. This has resulted in a significant rise in the adoption of wearable technology, particularly in the healthcare industry, where incorporating intelligent personal assistants for patients offers ever-increasing potential benefits as a solution for healthcare services [14]. Other application areas include the entertainment industry, industrial logistics, and sports. Nevertheless, the healthcare industry is considered a leader in the widespread adoption of wearable technology due to its ability to improve patient health and healthcare procedures [14]. Microservices have enabled wearables to offer a diverse range of functionalities and features, leading to their widespread acceptance and integration into everyday life.

Due to its practicability, microservices architecture has emerged as the preferred approach for distributed computing in the context of the IoT. IoT systems face numerous obstacles, including interoperability, scalability, data volume, mobility, security and privacy, implementation, and competence. Among these obstacles, the security and privacy of the captured data require considerable attention, necessitating the establishment of suitable standards and practices [15]. With the integration of IoT platforms and the resolution of various issues such as data ownership, data sharing policies, privacy and safety concerns, the full potential of wearable IoT devices can be realized [13]. Due to the uncontrolled dissemination and sharing of data, relying solely on technological solutions to ensure data security is no longer sufficient. This scenario yields two concerns. First, the data collection entity should be responsible for its security. Second, there has been an increase in the use of wearable devices in the healthcare industry, which may contribute to the development of health-related products subject to stringent industry regulations and associated legal and regulatory challenges [16]. The NIST has published a report addressing these privacy risks.

B. NISTIR 8228

The National Institute of Standards and Technology in the United States publishes several special publications and interagency reports to assist the government and private sector in meeting legal and regulatory requirements. Notably, two NIST publications, NIST Special Publication (SP) 800-53 Revision 5 [17] and NISTIR 8228, are significant for two reasons. First, NIST 800-53 provides an extensive array of privacy and security controls, but addresses cybersecurity risks in a manner that is insufficiently specific to addressing privacy risks of wearables. Second, NISTIR 8228 is specifically designed for the IoT, a rapidly expanding and evolving network of technologies that interact with the physical world and is characterized by rapid growth and change [18]. In this paper, our reference regulations document is NISTIR 8228, which directly applies to wearables, and the ontology we have constructed has a representative NIST 800-53 class that captures the Revision 5 controls associated with the specific challenges faced by IoT devices. See Fig.3 for the high-level classes of our NIST ontology. The heterogeneous nature of these devices is reflected in the variability of their generated data, as well as their interoperability (or lack thereof), which presents peculiar efficiency and utilization challenges that can be addressed using the Semantic Web, which includes knowledge graphs as one of the potential solutions to obtain and communicate domain knowledge.

C. SEMANTIC WEB

The semantic web represents the World Wide Web that provides standards for expressing relationships between web information and focuses primarily on data rather than documents. It allows data to be annotated with machine-understandable meta-data, automating their retrieval and use in inappropriate contexts [19] [20]. Semantic Web technologies include languages such as Resource Description Framework (RDF) and Web Ontology Language (OWL) for defining ontologies and describing meta-data using these ontologies, as well as reasoning tools for analyzing these descriptions. These technologies can be used to provide standard semantics of privacy information and policies, allowing all agents who understand fundamental Semantic Web technologies to effectively communicate and utilize each other's data and services

[21] [22] [23]. For reasoning over our OWL ontology, we will use SPARQL [24] [25] queries.

D. ONTOLOGY

An ontology is a formal definition of concepts and their relationships within a particular domain, ranging from simple vocabularies to complex logic-based formalisms. The term is closely related to the schema of the knowledge graph, which can be described using an ontology in the context of knowledge graphs. Ontologies can be created by defining concepts and relationships through domain analysis or by analyzing available data. Ontology, schema, and knowledge graphs can be used interchangeably in a controlled setting [26] [27]. Ontologies provide an automated process (known as reasoning) for retrieving axioms not explicitly included in the knowledge graph. Ontologies can be categorized as general or core ontology and domain ontology based on their conceptualizations [28] [29]. Domain ontology refers to the knowledge of a particular domain, and we will develop an IoT domain ontology for the NISTIR 8228 publication.

III. Related Work

Data privacy enforcement in wearable IoT devices has garnered considerable interest in recent IoT research. [30] proposed a clustering-based k-anonymity method to preserve the privacy of wearable IoT device data while ensuring the usability of collected data. This method offers a practical solution for maintaining user privacy in data-sharing procedures. Privacy and security have been major concerns in this field; [31] proposed using blockchain to securely manage and analyze healthcare big data to address the IoT's privacy concerns. [32] conducted a systematic literature review to examine the security of IoT devices and provided mobile computing countermeasures. [33] proposed a lightweight model, IoT-Stream, to semantically annotate streams, offering a simple yet effective approach for knowledge sharing and inference in IoT. [34] also presented a security context ontology for analyzing the security vulnerabilities of a power system in an IoT-Cloud environment. Their work highlighted the semantic web's and ontology's significance in IoT even further. Compliance with pertinent regulations and security standards is another essential aspect of Internet of Things adoption. [35] comprehensively analyzed the requirements of the smart grid's applicable security standards and guidelines. They proposed a model to map requirements from arbitrary standards, policies, and regulations, accelerating the readiness for cross-standard compliance. This paper builds upon these works by developing a semantically rich knowledge graph to represent NISTIR 8228 regulations, thereby aiding wearable IoT vendors in real-time compliance with these rules.

IV. METHODOLOGY

We advance our analysis using a multifaceted methodology based on the preceding theoretical background. On the one hand, we developed a novel ontology of the regulations document that guided our analysis of the existing privacy risk variables that wearables may encounter; on the other, we empirically analyzed the respective privacy policy documents. Fig. 1 shows the architecture flow of our work. In this paper, we randomly selected and analyzed the privacy policies of six prominent manufacturers of wearable devices. Based on their product market positions, we anticipate their privacy policies will be more comprehensive, covering the most anticipated Risk Mitigation Areas. They are Apple [36] [37], Samsung [38] [39], Fitbit [40], Garmin [41], Withings [42], and Hexoskin [43]. This list is not exhaustive, but the following methods span the entire wearable IoT device industry. While initially based on a select number of devices, the approach is scalable and adaptable to the more extensive wearable IoT industry by providing standard semantics of privacy information and policies. As new devices, technologies, and standards emerge, we can expand and update the ontology to accommodate recent changes, ensuring its continued relevance and applicability. The pertinent IoT privacy risks regulations document, for which we have created an integrated knowledge graph, is the National Institute of Standards and Technology publication, NISTIR 8228 Considerations for Managing Internet of Things Cybersecurity and Privacy Risks [11].

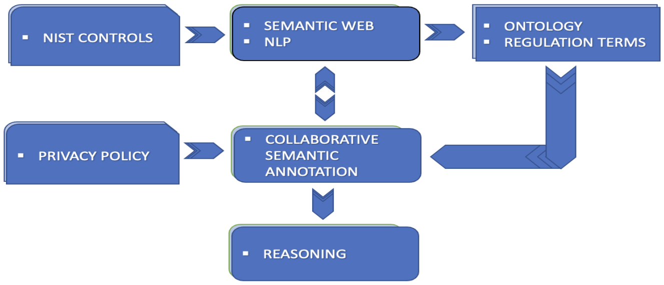


Fig. 1: Architecture Flow

A. ONTOLOGY DEVELOPMENT

Protégé, an open-source framework for building intelligent systems and ontologies, is the primary tool used for our ontology development [44]. We selected it due to its comprehensive support for OWL and its wide adoption in the semantic web research community. In consideration of privacy risks for wearables, NISTIR 8228 is our reference regulations document. This regulation emphasizes that IoT devices (to which wearables belong) affect privacy risks differently than traditional information technology devices. Refer to Fig. 2 to view our unified ontology of this NIST publication as a Unified Modeling Language class diagram. Finally, as a component of our ongoing work, we are consolidating this integrated knowledge graph with existing publicly available technologies.

Privacy and security are interdependent yet distinct concepts within IoT and wearable technologies. These concepts are intertwined in protecting personal information and preventing unauthorized data access, where security measures are the bulwark of privacy [45] [46]. Privacy concerns typically arise when security measures fail, or data is collected without the

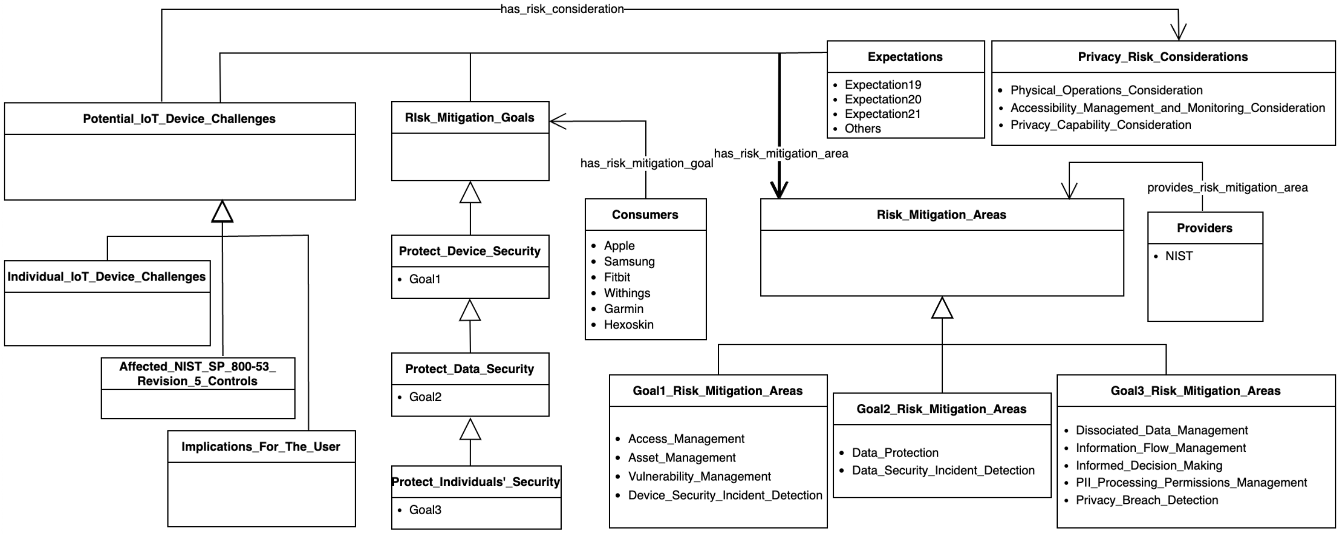

has_risk_consideration
Potential_IoT_Device_Challenges
RIsk_Mitigation_Goals
Expectations
• Expectation19
• Expectation20
• Expectation21
• Others
Privacy_Risk_Considerations
• Physical_Operations_Consideration
• Accessibility_Management_and_Monitoring_Consideration
• Privacy_Capability_Consideration
has_risk_mitigation_goal
has_risk_mitigation_area
provides_risk_mitigation_area
Individual_IoT_Device_Challenges
Protect_Device_Security
• Goal1
Consumers
• Apple
• Samsung
• Fitbit
• Withings
• Garmin
• Hexoskin
Risk_Mitigation_Areas
Providers
• NIST
Affected_NIST_SP_800-53_ Revision_5_Controls
Protect_Data_Security
• Goal2
Implications_For_The_User
Protect_Individuals'_Security
• Goal3
Goal1_Risk_Mitigation_Areas
• Access_Management
• Asset_Management
• Vulnerability_Management
• Device_Security_Incident_Detection
Goal2_Risk_Mitigation_Areas
• Data_Protection
• Data_Security_Incident_Detection
Goal3_Risk_Mitigation_Areas
• Dissociated_Data_Management
• Information_Flow_Management
• Informed_Decision_Making
• PII_Processing_Permissions_Management
• Privacy_Breach_Detection

Fig. 2: High-level Ontology describing the regulations for managing privacy risks of wearables

user's consent. At the same time, security revolves around safeguarding this data against theft, loss, or corruption [47].

Proactively, the NIST controls engage in privacy and security preservation by clearly defining Risk Mitigation Goals. They are the specific objectives set to reduce or eliminate risks in the Risk Mitigation Areas where vendors must take measures to minimize the impact of risks. They can achieve these goals by addressing the potential challenges identified in the corresponding Risk Mitigation Areas. There are three goals, namely:
Goal1: This goal focuses on ensuring the security of the IoT device itself (from threats and vulnerabilities).
Goal2: This goal emphasizes the protection of data that the IoT device stores, processes, and transmits. It involves addressing potential challenges that could lead to unauthorized access, tampering, or loss of data.
Goal3: This goal aims to protect individuals' privacy by mitigating privacy risks arising from authorized PII processing. The goals (*Goal1*, *Goal2*, and *Goal3*) outlined for managing IoT cybersecurity and privacy risks correspond to *Protect Device Security*, *Protect Data Security*, and *Protect Individuals' Security* respectively. It is essential to recognize that absolute privacy is unattainable without comprehensive security measures. Meeting each goal involves addressing a set of Risk Mitigation Areas, which concisely describe the significant cybersecurity and privacy risk mitigation aspects associated with IoT devices most affected by the Risk Considerations. Regarding wearables, *Goal2* and *Goal3* primarily address the most critical privacy Risk Mitigation Areas for services utilizing Personally Identifiable Information (PII). Risk Mitigation Areas "Asset Management" and "Vulnerability Management" straddle the privacy and security domains, as they are concerned with preserving device-related information and mitigating risks. The Risk Considerations provide insight into how IoT devices affect the management of privacy risks.

Similarly, achieving each goal may present difficulties (referring to the Potential Challenges). For example, these wearables may challenge how organizations expect the devices to help mitigate privacy risks for Risk Mitigation Areas and Goals (otherwise referred to as the Expectations). Each Potential Challenge is characterized by the individual device challenges, the affected draft NIST SP 800-53 Revision 5 Controls, and the implications for the user (in this case).

B. ANNOTATION

There were challenges in selecting the appropriate manual annotation tool for our collaborative annotation task. An effective annotation tool should be available, web-based, installable, workable, and schematic [48]. We chose the Doccano Annotation tool [49] for the efficient and accurate annotation of the privacy policies text data based on these requirements, as well as its user-friendly interface and scalability in handling large-scale annotation tasks.. Doccano is an open-source, web-based collaborative platform that allows users to upload, annotate, and visualize text data for various annotation tasks such as text and image classification, intent detection and slot filling, sequence labeling, image captioning, relation extraction, object detection, and more. Other manual annotation tools include Prodigy, Brat, LightTag, Labelbox, and many others.

The definitive objectives articulated in NIST IR 8228 guided the extraction of the 11 key regulation terms against which we performed annotations. Adherence to these fundamental regulatory objectives is critical in ensuring maximum privacy protection for IoT devices, the data they collect, and the people who use them. Our analysis of the privacy policy text corpora used a semi-automated approach centered on the regulation terms extracted via collaborative semantic annotation. These terms represent the key areas that vendors of wearable devices are expected to address in their privacy policies. The manual review phase of the extraction process considers the diverse nature of the information that may be

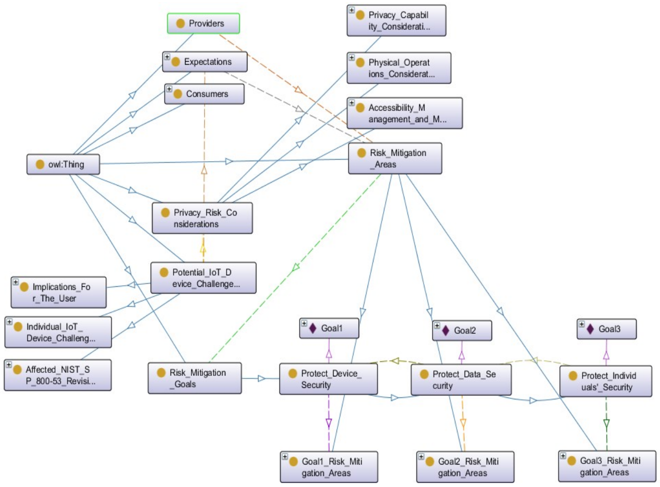

Providers
Privacy_Capabil ity_Considerati...
Expectations
Physical_Operat ions_Considerat...
Consumers
Accessibility_M anagement_and_M...
owl:Thing
Risk_Mitigation _Areas
Privacy_Risk_Co nsiderations
Potential_IoT_D evice_Challenge...
Implications_Fo r_The_User
Individual_IoT_ Device_Challeng...
Affected_NIST_S P_800-53_Revisi...
Risk_Mitigation _Goals
Goal1
Goal2
Goal3
Protect_Device_ Security
Protect_Data_Se curity
Protect_Individ uals'_Security
Goal1_Risk_Miti gation_Areas
Goal2_Risk_Miti gation_Areas
Goal3_Risk_Miti gation_Areas

Fig. 3: NIST 8228 ontology main classes

present in policy documents, accounting for specialized text or characters that may prove difficult for NLP techniques. As a result, this method ensures comprehensive and precise term extraction. The following term definitions, which are excerpts from the NIST regulations document, served as our annotation guidelines to foster consistency:

"*Asset Management*: Maintain a current, accurate inventory of all IoT devices and their relevant characteristics throughout their lifecycles to use that information for cybersecurity and privacy risk management purposes.
Vulnerability Management: Identify and eliminate known vulnerabilities in IoT device software and firmware in order to reduce the likelihood and ease of exploitation and compromise.
Access Management: Prevent unauthorized and improper physical and logical access to, usage of, and administration of IoT devices by people, processes, and other computing devices.
Device Security Incident Detection: Monitor and analyze IoT device activity for signs of incidents involving device security.
Data Protection: Prevent access to and tampering with data at rest or in transit that might expose sensitive information or allow manipulation or disruption of IoT device operations.
Data Security Incident Detection: Monitor and analyze IoT device activity for signs of incidents involving data security.
Information Flow Management: Maintain a current, accurate mapping of the information lifecycle of PII, including the type of data action, the elements of PII being processed by the data action, the party doing the processing, and any additional relevant contextual factors about the processing to use for privacy risk management purposes.
PII Processing Permissions Management: Maintain permissions for PII processing to prevent unpermitted PII processing.
Informed Decision Making: Enable individuals to understand the effects of PII processing and interactions with the device, participate in decision-making about the PII processing or interactions, and resolve problems.
Disassociated Data Management: Identify authorized PII processing and determine how PII may be minimized or disassociated from individuals and IoT devices.
Privacy Breach Detection: Monitor and analyze IoT device activity for signs of breaches involving individuals' privacy." [11]

We used the predefined key regulation terms as labels to meticulously annotate the entities in the privacy policy corpora, which are wearables' privacy Risk Mitigation Areas. Our semantic annotation was performed at the phrase/sentence linguistic component level, identifying phrases and sentences that express the same concept as the key term. For instance, we assigned the entity label "*Asset_Management*" to the concept represented by the key term "Asset Management." We refer to the identified entities as relevant text chunks. This process was repeated for each of the six consumer class privacy policy documents, annotating relevant text chunks correspondingly.

To provide a more transparent illustration of the annotation process, we highlight some specific text chunks extracted and annotated from the privacy policies of different vendors. In the Fitbit privacy policy, examples include "We may share non-personal information that is aggregated or de-identified so that it cannot reasonably be used to identify an individual," annotated as *Disassociated_Data_Management*, and "We use your information when needed to send you Service notifications and to respond to you when you contact us" annotated as *Information_Flow_ Management.* Another example is from Garmin, where the text chunk "If you enable two-factor authentication, Garmin processes your phone number or email address to send the security code via SMS or email" was annotated with the *Data_Security_Incident_Detection* label. Furthermore, in Samsung's privacy policy, a *Data_Protection* annotation was made for the statement, "We maintain safeguards to protect personal information we obtain through the Services." Lastly, an annotation for *Privacy_Breach_Detection* in Apple's privacy policy was made for the statement, "If we learn that a child's personal data was collected without appropriate authorization, it will be deleted as soon as possible." These examples illustrate the application of annotation labels to specific text chunks in accordance with the outlined key regulation terms.

The process of manual annotation is subjective. Therefore, to minimize subjectivity, a team of three individuals carried out the manual annotation process—the authors of this paper—each with a Master's degree or higher. More so, we used the inter-annotator agreement metric to ensure consistency and precision in the annotation process. This method ensured that we addressed any discrepancies in interpretations, boosting the reliability of our annotations even further. Fig. 4 depicts the average span distribution count of entity labels for Fitbit. Fig. 5 depicts the same information for the entire corpus of the privacy policy text. On the average, a total of 476 relevant text chunks were annotated.

For reliability evaluation purposes, we used Krippendorff's alpha score [50] to empirically determine the level of agreement between the annotators regarding the classification of the relevant text chunks in each vendor's privacy policy text corpus. We seek to evaluate compliance; consequently, a high level of agreement was the desired outcome.

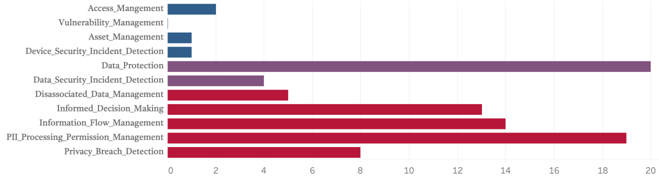


Fig. 4: Average frequency distribution of annotations: Fitbit privacy policy

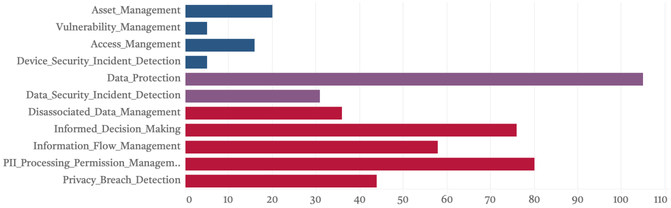


Fig. 5: Average frequency distribution of annotations: 6-vendor-combined privacy policy

TABLE I: Annotation count

Key Terms	**Person1**	**Person2**	**Person3**
Access Management	24	17	19
Vulnerability Management	6	3	6
Asset Management	23	15	10
Device Security Incident Management	4	7	4
Data Protection	128	90	97
Data Security Incident Detection	35	34	24
Information Flow Management	37	75	62
PII Processing Permissions Management	103	65	72
Informed Decision Making	78	80	70
Dissociated Data Management	39	39	30
Privacy Breach Detection	64	30	38
Total	541	455	432

The computation was performed using an open-source Python implementation of Krippendorff [51], and we utilized the same code environment to perform other NLP tasks. We utilized the ordinal Krippendorff's Alpha because it considers the class hierarchy, improving the accuracy of the reliability estimate. We expand this Goal class hierarchical knowledge conceptualization in the Knowledge Extraction section. We achieved a Krippendorff's Alpha score of 0.9123. This score is acceptable, as it demonstrates a high level of agreement between the three annotators, attributed to their cognitive abilities and the well-defined and precise NIST annotation guidelines [52]. If the task required some sentiment analysis, annotation scores could have been as low as 0.5261 on average [53]. This preliminary work is exploratory; as such, we did not perform any modeling.

C. *FEATURE ENGINEERING*

We used several feature engineering and Natural Language Processing (NLP) techniques to preprocess and analyze privacy policies in the proposed framework. We initially tokenized the text into individual words or tokens. Then, using the NLTK [54] library, we eliminated common English words or NLP stopwords, punctuations, and stopwords that were unique to our research context. For instance, we added the situationally generic term "data" and the vendor names (such as "Apple," "Withings," etc.) to our custom list of stopwords. Also, we converted the text to lowercase for consistency.

TABLE II: Top five annotated words for *Dissociated Data Management* - Hexoskin and Fitbit

Word	**Frequency**	**Word**	**Frequency**
Information	5	Information	5
Keep	2	Delete	3
Products	2	Days	2
De-identified	2	Privacy	2
Personal	2	Shield	2
Hexoskin		Fitbit	

TABLE III: Top five annotated words for *Informed Decision Making* - Apple and Garmin

Word	**Frequency**	**Word**	**Frequency**
Consent	14	Consent	11
Personal	10	Information	10
Privacy	9	Provide	5
Informed	8	Marketing	5
Use	6	Process	4
Apple		Garmin	

TABLE IV: Top five annotated words for PII *Processing Permissions Management* - Withings and Samsung

Word	**Frequency**	**Word**	**Frequency**
Personal	16	Information	18
Health	8	Personal	9
Account	6	Services	6
Information	5	Use	5
Privacy	5	Contact	4
Withings		Samsung	

Subsequently, we determined the frequency distribution of the words in the text and extracted the most frequent words to understand the privacy policies' main topics. Unlike many NLP pipelines, we chose not to perform lemmatization to preserve the original context of the words in the privacy policies. This decision was based on our observation that lemmatization could potentially alter the meaning and interpretation of the policy statements. These steps guided us in reducing the data's dimensionality, focusing on key terms, and preparing the data for a more nuanced analysis - creating a function that returns the five highest-occurring words from the collection of annotated text chunks for each extracted key regulation term.

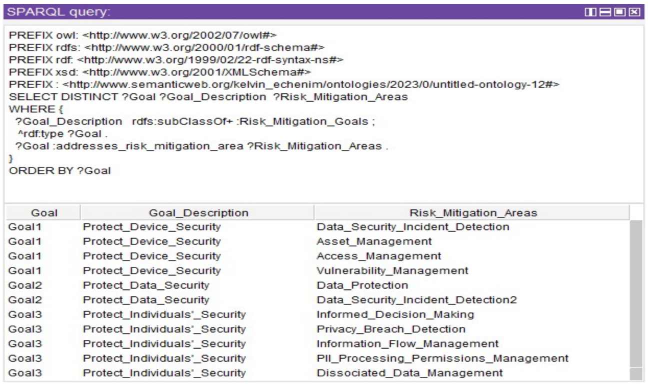
V. RESULT AND DISCUSSION

A. *KNOWLEDGE EXTRACTION*

Our primary contribution is the knowledge architecture derived from the essential NIST publication for wearable IoT devices. It was constructed primarily with RDF triples, and interested parties can access the information using SPARQL queries as it is available in the public domain. For reasoning over our OWL ontology, we used the HermiT reasoner. Also, the featured wearable vendors have been added to the knowledge base as consumers. Using our integrated ontology hosted on the Apache Jena server [55], we investigate some of the binding regulations by examining the following use case listings (assuming a prior definition of the necessary prefixes in each case):

Listing 1: The privacy risk mitigation goals, the goal descriptions, and the associated risk mitigation areas

```
SELECT DISTINCT ?Goal ?Goal_Description  ?
    Risk_Mitigation_Areas
WHERE {
  ?Goal_Description   rdfs:subClassOf+ :
      Risk_Mitigation_Goals ;
    ^rdf:type ?Goal .
  ?Goal :addresses_risk_mitigation_area ?
      Risk_Mitigation_Areas .
}
ORDER BY ?Goal
```



```
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX : <http://www.semanticweb.org/kelvin_echenim/ontologies/2023/0/untitled-ontology-12#>
SELECT DISTINCT ?Goal ?Goal_Description  ?Risk_Mitigation_Areas
WHERE {
 ?Goal_Description   rdfs:subClassOf+ :Risk_Mitigation_Goals ;
  ^rdf:type ?Goal .
 ?Goal :addresses_risk_mitigation_area ?Risk_Mitigation_Areas .
}
ORDER BY ?Goal
```

Goal	Goal_Description	Risk_Mitigation_Areas
Goal1	Protect_Device_Security	Data_Security_Incident_Detection
Goal1	Protect_Device_Security	Asset_Management
Goal1	Protect_Device_Security	Access_Management
Goal1	Protect_Device_Security	Vulnerability_Management
Goal2	Protect_Data_Security	Data_Protection
Goal2	Protect_Data_Security	Data_Security_Incident_Detection2
Goal3	Protect_Individuals'_Security	Informed_Decision_Making
Goal3	Protect_Individuals'_Security	Privacy_Breach_Detection
Goal3	Protect_Individuals'_Security	Information_Flow_Management
Goal3	Protect_Individuals'_Security	PII_Processing_Permissions_Management
Goal3	Protect_Individuals'_Security	Dissociated_Data_Management

Fig. 6: SPARQL Query result for Listing 1

We extracted the three regulation Goals using SPARQL. Refer to Listing 1. We used the recursive path query because each goal builds on the previous goal and does not supplant or negate its requirement. As a result, addressing a set of risk mitigation areas is required to achieve each risk mitigation goal. This concept explains the hierarchical structure of the Goal classes discussed in the prior section. The Goal classes form a hierarchy in which each class subsumes the one before. Fig. 6 shows the results of the preceding query.

Each risk mitigation area specifies an aspect of privacy risk mitigation that is believed to be most significantly or surprisingly impacted by privacy risk considerations for IoT.

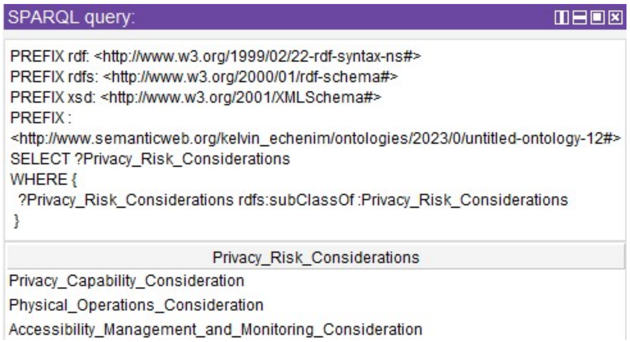
Listing 2: The privacy risk considerations

```
SELECT ?Privacy_Risk_Considerations
WHERE {
  ?Privacy_Risk_Considerations rdfs:subClassOf
        :Privacy_Risk_Considerations
 }
```

These three privacy and cybersecurity risk variables extracted through the query in Listing 2, potentially influence the management of privacy and security risks for wearable/IoT devices. Therefore, vendors of wearables should ensure that they consider the risk factors listed in Fig. 7 throughout the lifecycle of their IoT devices.

Listing 3: Expectations that may be met with one or more challenges posed by IoT devies

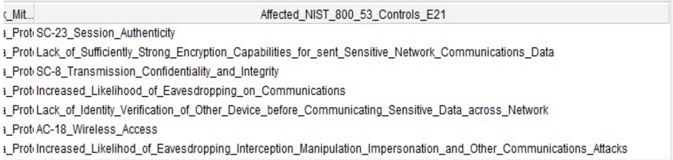
```
SELECT DISTINCT ?Goal ?Goal_Description ?
    Risk_Mitigation_Area ?
    Affected_NIST_800_53_Controls
WHERE {
        ?Expectation a :Expectation21 ;
```



```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX :
<http://www.semanticweb.org/kelvin_echenim/ontologies/2023/0/untitled-ontology-12#>
SELECT ?Privacy_Risk_Considerations
WHERE {
 ?Privacy_Risk_Considerations rdfs:subClassOf :Privacy_Risk_Considerations
}
```

Privacy_Risk_Considerations
Privacy_Capability_Consideration
Physical_Operations_Consideration
Accessibility_Management_and_Monitoring_Consideration

Fig. 7: SPARQL Query result for Listing 2



_Mit...	Affected_NIST_800_53_Controls_E21
_Prot	SC-23_Session_Authenticity
_Prot	Lack_of_Sufficiently_Strong_Encryption_Capabilities_for_sent_Sensitive_Network_Communications_Data
_Prot	SC-8_Transmission_Confidentiality_and_Integrity
_Prot	Increased_Likelihood_of_Eavesdropping_on_Communications
_Prot	Lack_of_Identity_Verification_of_Other_Device_before_Communicating_Sensitive_Data_across_Network
_Prot	AC-18_Wireless_Access
_Prot	Increased_Likelihod_of_Eavesdropping_Interception_Manipulation_Impersonation_and_Other_Communications_Attacks

Fig. 8: SPARQL Query result for Listing 3

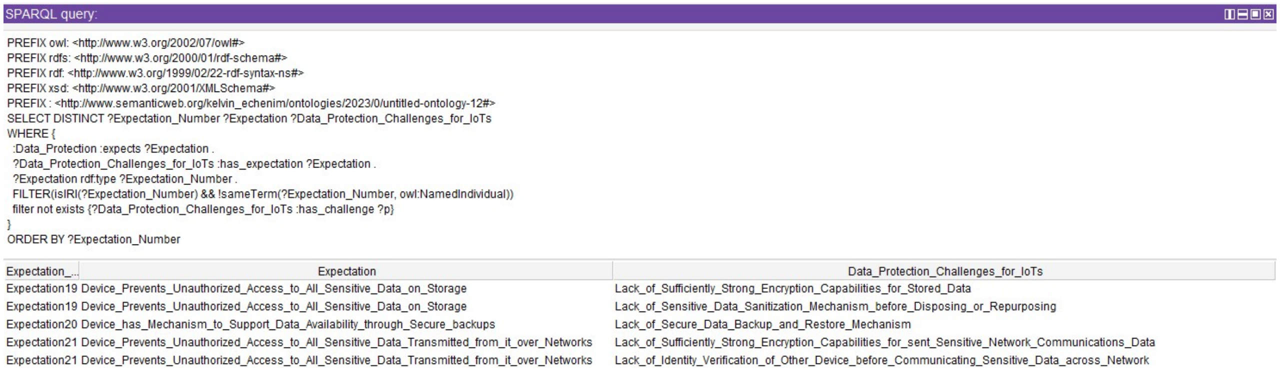
```
        :has_risk_mitigation_area ?
            Risk_Mitigation_Area .
      ?Goal :addresses_risk_mitigation_area
          ?Risk_Mitigation_Area .
      ?Goal rdf:type ?Goal_Description .
      ?Affected_NIST_800_53_Controls :
          has_expectation ?Expectation.
}
```

Using the query in Listing 3, we examined a specific expectation (such as Expectation 21: "Device Prevents Unauthorized Access to Any Sensitive Data Sent from it Across Networks"), and then we identified the corresponding Risk Mitigation Goal and its description, the corresponding NIST 800-53 Revision 5 Control, and the corresponding Risk Mitigation Area. Fig. 8 displays the result showing the Affected NIST 800-53 Revision 5 Controls.

There are recommended approaches for organizations to solve the challenges of privacy risk mitigation. Risk mitigation may be impacted by privacy risk considerations. "Data Protection" is the first area of risk mitigation for attaining *Goal2*: *Protect Data Security*, and the fifth overall. We detailed all its linked Expectations using the SPARQL query in Listing 4. The result is shown in Fig. 9.

Listing 4: Expectations that may be met with one or more challenges posed by IoT devies

```
SELECT ?Expectation_Number ?Expectation ?
    Data_Protection_Challenges_for_IoTs
WHERE {
  :Data_Protection :expects ?Expectation .
  ?Data_Protection_Challenges_for_IoTs :
      has_expectation ?Expectation .
  ?Expectation rdf:type ?Expectation_Number
  filter not exists {?
      Data_Protection_Challenges_for_IoTs :
      has_challenge ?p}
}
ORDER BY ?Expectation_Number
```



```
SPARQL query:

PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX : <http://www.semanticweb.org/kelvin_echenim/ontologies/2023/0/untitled-ontology-12#>
SELECT DISTINCT ?Expectation_Number ?Expectation ?Data_Protection_Challenges_for_IoTs
WHERE {
 :Data_Protection :expects ?Expectation .
 ?Data_Protection_Challenges_for_IoTs :has_expectation ?Expectation .
 ?Expectation rdf:type ?Expectation_Number .
 FILTER(isIRI(?Expectation_Number) && !sameTerm(?Expectation_Number, owl:NamedIndividual))
 filter not exists {?Data_Protection_Challenges_for_IoTs :has_challenge ?p}
}
ORDER BY ?Expectation_Number
```

```
Expectation_...  Expectation                                                                  Data_Protection_Challenges_for_IoTs
Expectation19 Device_Prevents_Unauthorized_Access_to_All_Sensitive_Data_on_Storage           Lack_of_Sufficiently_Strong_Encryption_Capabilities_for_Stored_Data
Expectation19 Device_Prevents_Unauthorized_Access_to_All_Sensitive_Data_on_Storage           Lack_of_Sensitive_Data_Sanitization_Mechanism_before_Disposing_or_Repurposing
Expectation20 Device_has_Mechanism_to_Support_Data_Availability_through_Secure_backups       Lack_of_Secure_Data_Backup_and_Restore_Mechanism
Expectation21 Device_Prevents_Unauthorized_Access_to_All_Sensitive_Data_Transmitted_from_it_over_Networks  Lack_of_Sufficiently_Strong_Encryption_Capabilities_for_sent_Sensitive_Network_Communications_Data
Expectation21 Device_Prevents_Unauthorized_Access_to_All_Sensitive_Data_Transmitted_from_it_over_Networks  Lack_of_Identity_Verification_of_Other_Device_before_Communicating_Sensitive_Data_across_Network
```

Fig. 9: SPARQL Query result for Listing 4

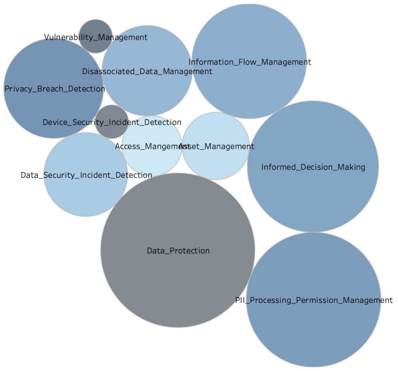


Fig. 10: Average annotation frequency for regulation terms

B. DISCUSSION AND COMPARATIVE ANALYSIS

In a broad sense, the *Data_Protection* label had more annotated entities (as shown in Fig. 10), even though IoT privacy encompasses other aspects besides preventing access to and tampering with data at rest or in transit that might expose sensitive information or allow manipulation or disruption of IoT device operations (as defined in the NIST regulation). This observation implies that Data Protection has been the primary privacy focus of vendors. However, the concept of IoT privacy compliance goes beyond this, as outlined by NIST's Risk Mitigation Areas.

TABLE V: Average annotation count by vendor

Key Regulation Terms	Apple	Samsung	Fitbit	Withings	Garmin	Hexoskin
Access Management	2	1	1	9	4	3
Vulnerability Management	1	0	0	2	2	0
Asset Management	2	0	2	6	5	1
Device Security Incident Management	0	0	1	2	1	1
Data Protection	20	14	20	30	16	5
Data Security Incident Detection	9	3	4	10	4	1
Information Flow Management	9	13	14	10	7	5
PII Processing Permissions Management	15	15	19	14	9	8
Informed Decision Making	17	15	13	7	18	6
Dissociated Data Management	7	3	5	10	5	6
Privacy Breach Detection	11	6	8	8	6	5

A comparative analysis of the privacy policies of the sampled wearable devices shows a significant variation in the number of annotations for the key regulation terms across the different companies. Apple, for instance, had the highest number of annotations for Data Protection, Information Flow Management, and Informed Decision Making, indicating that its privacy policies comprehensively cover these aspects. On the other hand, Withings had the most annotations for Asset Management, Access Management, and Data Security Incident Detection, suggesting a strong emphasis on meeting *Goal1*, *Protect Device Security*. Samsung and Fitbit had a balanced distribution of annotations across the key regulation terms, with a notable emphasis on PII Processing Permission Management. However, Hexoskin had the fewest annotations for most of the regulation terms, indicating a potential improvement area in their privacy policy.

Moving forward, we looked closely at the returned result containing the most frequent words from the annotated relevant text chunks. The frequency of specific words in the annotated text chunks for the key regulation terms offered additional insight into the privacy policies of each vendor. Again, there are parallels and distinctions in the top words associated with each regulation term among the vendors, providing insights into how each vendor approaches compliance with applicable NIST regulations. For Asset Management, the top word for Withings and Fitbit policies is "device," whereas Garmin includes "customer" and "service," and Samsung includes "profile." These disparities suggest that each vendor may prioritize compliance with the Asset Management aspect of the regulations differently. Garmin and Withings emphasized "account" and "research" in the context of Disassociated Data Management, whereas Fitbit and Hexoskin emphasized "information," indicating that diverse approaches to managing disassociated data exist. Fitbit used "information" and "account" most frequently for PII Processing Permissions Management. In contrast, Samsung and Withings emphasized "information" and "personal," demonstrating a shared emphasis on managing personal information and user accounts. In the area of Informed Decision Making, Samsung, Apple, and Garmin emphasized "information," "consent," and "provide," respectively, proposing distinct strategies to ensure that users make informed decisions. Fitbit and Apple prioritized "information" and "personal" for Privacy Breach Detection, whereas Hexoskin prioritizes "users," suggesting various methods for detecting and managing privacy violations.

Exploring the overlapping words, we observed a pattern where "information" is a top word for multiple vendors across the regulation terms; this indicates that privacy risks are a prominent concern for all vendors, signaling a shared emphasis on data and information management. Considered together, the terms "personal" and "information" were prevalent in the vendor privacy policies for PII Processing Permissions Management and Informed Decision Making. The featured wearable vendors placed a minor emphasis on Device Security Incident Detection and Vulnerability Management (See Fig.10).

While there were commonalities among the privacy policies of wearable devices, each company had areas of emphasis that reflected its approach to privacy and data management.

C. METHOD VALIDATION

For validation purposes, we merged the five most frequent annotated words (as partly shown earlier) for each vendor's key regulation terms into a set. We used NLP techniques to determine the total frequency of these words in each vendor privacy policy corpora. Then, we plotted a bar graph featuring each of the six vendor privacy policies, with the height of the bars representing the magnitude of the frequency of the combined word set. Refer to Fig. 11. We eventually populated the our ontology with these terms as instances of their respective classes. Using the semi-automated privacy policy reasoning mechanisms in conjunction with our publicly available ontology, ensuring privacy policy agreement compliance of wearables with IoT regulations can be expedited. Finally, we recommend that vendors of wearables include specific language or sections in their privacy policy agreements for each Risk Mitigation Area, which would further help mitigate the rigor of compliance checking and address the issues of user explainability.

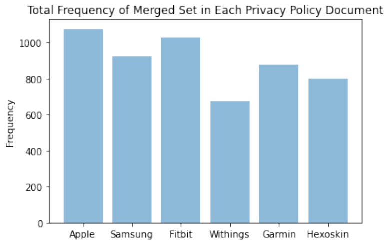


Fig. 11: Frequency distribution of top annotated words

VI. LIMITATIONS

We recognize that our methodology has certain limitations. For instance, while our approach is effective at extracting and analyzing privacy policies, it may only capture some of the subtleties and complexities of legal language, primarily due to the subjectivity of the annotation procedure. In addition, our methodology relies on the precision and exhaustiveness of the privacy policies provided by manufacturers of wearable devices, which may vary in quality and extensive detail. Although the manufacturing of wearable devices is beyond the scope of this paper, the proposed framework can serve as the basis for future work that could include a mechanism to verify the implementation of privacy policies in wearable devices. This implementation could be accomplished by creating a set of compliance tests or checks based on the extracted rules from the regulation ontology. Regarding anticipated revisions to the featured authority document, we propose a mechanism to update the knowledge graph following the modifications to the regulations, necessitating a rerun of the extraction and graph creation process whenever standards are updated.

VII. CONCLUSION

We developed a semantically rich knowledge graph to depict NISTIR 8228 regulations by extracting key terms, establishing clear annotation guidelines, and reasoning over our developed ontology. Due to the textual nature of these regulatory standards, it is challenging for wearable IoT vendors to comply with the rules outlined in these documents in real time. Using Natural Language Processing, Semantic Web concepts, and textual document conversion into a graph-based ontology, our novel framework can accurately extract the expectations, privacy risk mitigation areas, and associated regulations. This approach also helped us map and populate our ontology instances with rules occurring in multiple IoT privacy policy agreements for wearable vendors. Our knowledge graph, available in the public domain, is helpful for reasoning and notifying wearable device vendors of rule violations. Our ontology is semi-automated based on annotations extracted from the privacy policies of numerous wearable IoT device vendors. We intend to automate and populate our ontology in future work fully. We are also developing a custom predictive modeling approach for an easy-to-use graphical interface that will enable IoT vendors to query and reason over the knowledge graph and take swift action.

ACKNOWLEDGMENT

This work was supported in part by NSF award 1747724, Phase I IUCRC UMBC: Center for Accelerated Real-Time Analytics (CARTA).

REFERENCES

[1] A. Kazlouski, T. Marchioro, H. Manifavas, and E. Markatos, "Do partner apps offer the same level of privacy protection? the case of wearable applications," in *2021 IEEE International Conference on pervasive computing and communications Workshops and other affiliated events (PerCom Workshops)*. IEEE, 2021, pp. 648–653.

[2] F. M. Garcia-Moreno, M. Bermudez-Edo, J. L. Garrido, E. Rodríguez-García, J. M. Pérez-Mármol, and M. J. Rodríguez-Fórtiz, "A microservices e-health system for ecological frailty assessment using wearables," *Sensors*, vol. 20, no. 12, p. 3427, 2020.

[3] N. Surantha, O. K. Utomo, E. M. Lionel, I. D. Gozali, and S. M. Isa, "Intelligent sleep monitoring system based on microservices and event-driven architecture," *IEEE Access*, vol. 10, pp. 42 069–42 080, 2022.

[4] V.-A. Stefanescu and I. E. Radoi, "Stress level prediction using data from wearables," in *2019 18th RoEduNet Conference: Networking in Education and Research (RoEduNet)*. IEEE, 2019, pp. 1–6.

[5] A. Keogh, K. Taraldsen, B. Caulfield, and B. Vereijken, "It's not about the capture, it's about what we can learn": a qualitative study of experts' opinions and experiences regarding the use of wearable sensors to measure gait and physical activity," *Journal of NeuroEngineering and Rehabilitation*, vol. 18, no. 1, p. 78, 2021.

[6] N. Elmagboul, B. W. Coburn, J. Foster, A. Mudano, J. Melnick, D. Bergman, S. Yang, L. Chen, C. Filby, T. R. Mikuls *et al.*, "Physical activity measured using wearable activity tracking devices associated with gout flares," *Arthritis Research & Therapy*, vol. 22, pp. 1–9, 2020.

[7] A. Henriksen, M. Haugen Mikalsen, A. Z. Woldaregay, M. Muzny, G. Hartvigsen, L. A. Hopstock, and S. Grimsgaard, "Using fitness trackers and smartwatches to measure physical activity in research: analysis of consumer wrist-worn wearables," *Journal of medical Internet research*, vol. 20, no. 3, p. e110, 2018.

[8] G. Arnow, "Apple watch-ing you: Why wearable technology should be federally regulated," *Loy. LAL Rev.*, vol. 49, p. 629, 2016.

[9] M. Alhajri, A. Salehi Shahraki, and C. Rudolph, "Privacy of fitness applications and consent management in blockchain," *Australasian Computer Science Week 2022*, pp. 65–73, 2022.

[10] E. Bostancı, "Medical wearable technologies: applications, problems and solutions," in *2015 Medical Technologies National Conference (TIPTEKNO)*. IEEE, 2015, pp. 1–4.

[11] K. Boeckl, K. Boeckl, M. Fagan, W. Fisher, N. Lefkovitz, K. N. Megas, E. Nadeau, D. G. O'Rourke, B. Piccarreta, and K. Scarfone, *Considerations for managing Internet of Things (IoT) cybersecurity and privacy risks*. US Department of Commerce, National Institute of Standards and Technology . . . , 2019.

[12] A. J. Muhammad, G. K. Muhammad, A. Sajjad, and C. Ilyoung, "Sensors: Microservices in Web Objects Enabled IoT Environment for Enhancing Reusability," 2018. [Online]. Available: https://www.mdpi.com/1424-8220/18/2/352

[13] F. John Dian, R. Vahidnia, and A. Rahmati, "Wearables and the Internet of Things (IoT), Applications, Opportunities, and Challenges: A Survey," *IEEE Access*, vol. 8, pp. 69 200–69 211, 2020.

[14] W. Salehi, G. Gupta, S. Bhatia, D. Koundal, A. Mashat, and A. Belay, "IoT-Based Wearable Devices for Patients Suffering from Alzheimer Disease," *Contrast Media & Molecular Imaging*, Apr. 2022, publisher: Hindawi.

[15] A. Razzaq, "A Systematic Review on Software Architectures for IoT Systems and Future Direction to the Adoption of Microservices Architecture," *SN Computer Science*, vol. 1, no. 6, Oct. 2020.

[16] J. Gao, "The data privacy regulations for the health data in wearable industry in the United States," 2022, publisher: Malmo University.

[17] S. NIST, "800-53 rev. 5: Security and privacy controls for information systems and organizations," 2020.

[18] R. V. Rose, "New nist revisions–what do they mean for regulatory compliance?" *EDPACS*, vol. 59, no. 6, pp. 5–13, 2019.

[19] K. P. Joshi, Y. Yesha, and T. Finin, "Automating cloud services life cycle through semantic technologies," *IEEE Transactions on Services Computing*, vol. 7, no. 1, pp. 109–122, 2012.

[20] Wikipedia. (12 February, 2023) Semantic Web. Accessed: March 23, 2023. [Online]. Available: https://en.wikipedia.org/wiki/Semantic_Web

[21] K. P. Joshi, L. Elluri, and A. Nagar, "An integrated knowledge graph to automate cloud data compliance," *IEEE Access*, vol. 8, pp. 148 541–148 555, 2020.

[22] W3C. (15 March, 2014) Resource Description Framework. Accessed: March 23, 2023. [Online]. Available: https://www.w3.org/RDF/

[23] W3C. (10 February, 2004) Web Ontology Language. Accessed: March 23, 2023. [Online]. Available: https://www.w3.org/TR/owl-features/

[24] W3C. (21 March, 2013) SPARQL 1.1 Overview. Accessed: March 23, 2023. [Online]. Available: https://www.w3.org/TR/sparql11-overview/

[25] L. Elluri, A. Nagar, and K. P. Joshi, "An integrated knowledge graph to automate gdpr and pci dss compliance," in *2018 IEEE International Conference on Big Data (Big Data)*. IEEE, 2018, pp. 1266–1271.

[26] V. Ryen, A. Soylu, and D. Roman, "Building semantic knowledge graphs from (semi-) structured data: a review," *Future Internet*, vol. 14, no. 5, p. 129, 2022.

[27] K. P. Joshi, L. Elluri, and A. Nagar, "An integrated knowledge graph to automate cloud data compliance," *IEEE Access*, vol. 8, pp. 148 541–148 555, 2020.

[28] M. Yahya, J. G. Breslin, and M. I. Ali, "Semantic web and knowledge graphs for industry 4.0," *Applied Sciences*, vol. 11, no. 11, p. 5110, 2021.

[29] L. Elluri, S. S. L. Chukkapalli, K. P. Joshi, T. Finin, and A. Joshi, "A bert based approach to measure web services policies compliance with gdpr," *IEEE Access*, vol. 9, pp. 148 004–148 016, 2021.

[30] F. Liu and T. Li, "A clustering k-anonymity privacy-preserving method for wearable iot devices," *Security and Communication Networks*, vol. 2018, no. 3, p. 4945152, 2018.

[31] A. D. Dwivedi, G. Srivastava, S. Dhar, and R. Singh, "A decentralized privacy-preserving healthcare blockchain for iot," *Sensors*, vol. 19, no. 2, p. 326, 2019.

[32] B. Liao, Y. Ali, S. Nazir, L. He, and H. U. Khan, "Security analysis of iot devices by using mobile computing: A systematic literature review," *IEEE Access*, vol. 8, pp. 120 331–120 350, 2020.

[33] T. Elsaleh, S. Enshaeifar, R. Rezvani, S. T. Acton, V. Janeiko, and M. Bermudez-Edo, "Iot-stream: A lightweight ontology for internet of things data streams and its use with data analytics and event detection services," *Sensors*, vol. 20, no. 4, p. 953, 2020.

[34] C. Choi and J. Choi, "Ontology-based security context reasoning for power iot-cloud security service," *IEEE Access*, vol. 7, pp. 110 510–110 517, 2019.

[35] M. Stojkov, N. Dalčeković, B. Markoski, B. Milosavljević, and G. Sladić, "Towards cross-standard compliance readiness: Security requirements model for smart grid," *Energies*, vol. 14, no. 21, p. 6862, 2021. [Online]. Available: http://dx.doi.org/10.3390/en14216862

[36] Apple. (2022) Apple privacy policy. Accessed: March 21, 2023. [Online]. Available: https://www.apple.com/legal/privacy/en-ww/

[37] Apple Health. (2020) Apple Health Study Apps privacy policy. Accessed: March 21, 2023. [Online]. Available: https://www.apple.com/legal/privacy/apple-health-studies/en-ww/

[38] Samsung. (December 30, 2022) Samsung privacy policy for the u.s. Accessed: March 21, 2023. [Online]. Available: https://account.samsung.com/membership/terms/privacypolicy

[39] Samsung Health. (April 5, 2021) Samsung Health privacy policy for the U.S. Accessed: March 21, 2023. [Online]. Available: https://www.samsunghealth.com/privacy

[40] Fitbit. (September 16, 2022) Fitbit privacy policy. Accessed: March 21, 2023. [Online]. Available: https://www.apple.com/legal/privacy/en-ww/

[41] Garmin. (May 10, 2022) Garmin privacy policy. Accessed: March 21, 2023. [Online]. Available: https://www.garmin.com/en-US/privacy/global/policy/

[42] Withings. (January 26, 2023) Withings privacy policy. Accessed: March 21, 2023. [Online]. Available: https://www.withings.com/au/en/legal/privacy-policy

[43] Hexoskin. (December 27, 2019) Hexoskin privacy policy. Accessed: March 21, 2023. [Online]. Available: https://www.hexoskin.com/pages/privacy-policy

[44] M. A. Musen, "The protégé project: a look back and a look forward," *AI Matters*, vol. 1, no. 4, pp. 4–12, 2015, accessed: March 18, 2023. [Online]. Available: https://doi.org/10.1145/2757001.2757003

[45] O. Arias, J. Wurm, K. Hoang, and Y. Jin, "Privacy and security in internet of things and wearable devices," *IEEE Transactions on Multi-Scale Computing Systems*, vol. 1, no. 2, pp. 99–109, 2015.

[46] O. D'Mello, M. Gelin, F. B. Khelil, R. E. Surek, and H. Chi, "Wearable iot security and privacy: a review from technology and policy perspective," in *Future Network Systems and Security: 4th International Conference, FNSS 2018, Paris, France, July 9–11, 2018, Proceedings 4*. Springer, 2018, pp. 162–177.

[47] K. W. Ching and M. M. Singh, "Wearable technology devices security and privacy vulnerability analysis," *International Journal of Network Security & Its Applications*, vol. 8, no. 3, pp. 19–30, 2016.

[48] M. Neves and J. Ševa, "An extensive review of tools for manual annotation of documents," *Briefings in Bioinformatics*, vol. 22, no. 1, pp. 146–163, 12 2019. [Online]. Available: https://doi.org/10.1093/bib/bbz130

[49] H. Nakayama, T. Kubo, J. Kamura, Y. Taniguchi, and X. Liang. (2018) doccano: Text annotation tool for human. [Online]. Available: https://github.com/doccano/doccano

[50] K. Krippendorff, "Estimating the Reliability, Systematic Error and Random Error of Interval Data," *Educational and Psychological Measurement*, vol. 30, no. 1, pp. 61–70, Apr. 1970, publisher: SAGE Publications Inc. [Online]. Available: https://doi.org/10.1177/001316447003000105

[51] T. Grill and S. Castro, "Python implementation of krippendorff's alpha—inter-rater reliability," 2017.

[52] K. A. Neuendorf, *The content analysis guidebook*. sage, 2017.

[53] B. R. Chakravarthi, N. Jose, S. Suryawanshi, E. Sherly, and J. P. McCrae, "A sentiment analysis dataset for code-mixed malayalam-english," 2020. [Online]. Available: https://arxiv.org/abs/2006.00210

[54] S. Bird, E. Klein, and E. Loper, *Natural language processing with Python: analyzing text with the natural language toolkit*, 2009.

[55] Apache Jena. (2023) Apache Software Foundation. Accessed: March 21, 2023. [Online]. Available: https://jena.apache.org/