In proceedings of 3rd Workshop on Knowledge Graphs and Big Data in Conjunction with IEEE BigData 2023

IoT-Reg: A Comprehensive Knowledge Graph for Real-Time IoT Data Privacy Compliance

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

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

***Abstract*—The proliferation of the Internet of Things (IoT) has led to an exponential increase in data generation, especially from wearable IoT devices. While this data influx offers unparalleled insights and connectivity, it also brings significant privacy and security challenges. Existing regulatory frameworks like the United States (US) National Institute of Standards and Technology Interagency or Internal Report (NISTIR) 8228, the US Health Insurance Portability and Accountability Act (HIPAA), and the European Union (EU) General Data Protection Regulation (GDPR) aim to address these challenges but often operate in isolation, making their compliance in the vast IoT ecosystem inconsistent. This paper presents the IoT-Reg ontology, a holistic semantic framework that amalgamates these regulations, offering a stratified approach based on the IoT data lifecycle stages and providing a comprehensive yet granular approach to IoT data handling practices. The IoT-Reg ontology aims to transform the IoT domain into a realm where regulatory controls are seamlessly integrated system components by emphasizing risk management, compliance, and the pivotal role of manufacturers' privacy policies, ensuring consistent adherence, enhancing user trust, and promoting a privacy-centric IoT environment. We include the results of validating this framework against risk mitigation for Wearable IoT devices.**

***Index Terms*—privacy, IoT, semantic interoperability, wearables, regulations**

I. Introduction

The Internet of Things (IoT) has ushered in a new connectivity and data generation era. With billions of devices, ranging from simple home sensors to advanced wearable health monitors, the IoT ecosystem is vast and continuously evolving. This unparalleled connectivity has the potential to revolutionize sectors like healthcare, transportation, and manufacturing by harnessing the power of big data analytics. However, the rapid proliferation of IoT devices also leads to an exponential growth in data, posing challenges to data privacy, security, and compliance.

In the face of these challenges, global regulatory bodies have been proactive, instituting robust guidelines and standards such as the United States (US) National Institute of Standards and Technology Internal Report (NISTIR) 8228 [1], the US Health Insurance Portability and Accountability Act (HIPAA) [2], and European Union's (EU) General Data Protection Regulation (GDPR) [3]. These regulations aim to protect user data and ensure its ethical handling in the vast landscape of IoT and big data. However, the diverse and dynamic nature of the IoT domain, combined with the complexities of big data, poses challenges in achieving consistent regulatory compliance. While these regulations are comprehensive, they often operate in silos, leading to potential gaps in their application. This incongruence highlights the need for a unified framework that integrates these regulations, especially when contextualized within the IoT data lifecycle stages.

We have developed a comprehensive, semantically rich framework, IoT-Reg (Internet of Things - Regulations) to address this gap. The IoT-Reg ontology seeks to amalgamate the various data privacy and security regulations governing the IoT data space. IoT-Reg emphasizes stratification according to the IoT data lifecycle, covering stages like data collection, retention, processing, sharing, and deletion. This stratified approach ensures that the framework is holistic and detailed, providing stakeholders with a clear roadmap for data handling.

Risk management and compliance are pivotal in this endeavor. The IoT-Reg ontology promotes an environment where IoT devices, especially wearables, can operate securely within the big data ecosystem. By understanding and proactively addressing potential risks, the ontology serves as:
(i) a guide for manufacturers in designing privacy-aware devices tailored for data-intensive applications and
(ii) a resource for users, empowering them to make informed decisions in a data-driven world.
Privacy policies play a crucial role in the broader narrative of data privacy in the IoT-big data landscape. Manufacturers are responsible for developing clear and transparent policies, ensuring that user data is handled with integrity. With its comprehensive design, the IoT-Reg ontology can assist manufacturers in formulating privacy policies that are consistent with regulatory requirements.

As the digital fabric of our world becomes increasingly interwoven with big data, the importance of data privacy and security assumes a position of preeminence. The IoT-Reg ontology emerges as a beacon in this context, emphasizing

unification, stratification, and risk management. It reinforces the belief that the IoT and big data ecosystem can flourish without compromising data privacy and security principles.

The remaining sections of the paper are structured as follows: Section II examines the related work and background. Section III describes the key regulations that impact IoT privacy and the privacy requirements that we have incorporated into our design for each phase of the IoT data lifecycle. Section IV details our IoT-Reg methodology, and in Section V, using a use case scenario, we outline our risk management ontological model. In section VI, we define rules that assess the validity of our ontology. Section VII provides a summary of our conclusions and future work.

II. RELATED WORK AND BACKGROUND

A. Internet of Things (IoT)

IoT has attracted increased research interest, particularly in the field of ontologies. As the IoT ecosystem grows, a unified ontology framework becomes essential. [4] comprehensively reviews existing ontologies in the Internet of Things (IoT) domain, emphasizing the significance of fundamental IoT application concepts. These basic concepts, including Augmented Entity, User, Device, Resource, and Service, serve as the basis for comprehending and depicting the vast and diverse IoT landscape. The IoT ecosystem relies heavily on location-based ontologies. The WGS84 ontology, for example, provides abstract concepts for defining SpatialThings and TemporalThings. [4] highlight initiatives such as IoT-Lite, which extends the SSN ontology with SAO, and the VITAL ontology, which combines concepts from multiple ontologies to define sensors, measurements, time, and location. [5] highlight the limitations of current solutions in their quest for IoT interoperability, which focus mainly on communication protocols and message formats and often neglect the semantic essence of data. To address this, they developed the Comprehensive Ontology for the Internet of Things (COIoT), which reuses core concepts from well-known ontologies and introduces key concepts such as policy, context, and monitoring. Through COIoT, their work focuses on transforming raw sensor data into actionable knowledge. While their approach offers a holistic perspective on the Internet of Things by addressing interoperability and context awareness, it does not delve deeply into the regulatory and privacy dimensions. This void highlights the niche that our IoT-Reg ontology intends to fill.

In the realm of IoT security, the paper by Mozzaquatro et al. [6] introduces the IoT Security (IoTSec) ontology, a structured knowledge representation that explores the intricate relationships between the traditional components of risk analysis. IoTSec identifies different technologies as assets, emphasizing their requisite security properties such as confidentiality and integrity. The ontology categorizes vulnerabilities associated with M2M (Machine-to-Machine) technologies and potential threats that intelligently exploit these vulnerabilities. IoTSec provides a comprehensive view of the technical facets of IoT security, but it predominantly sidesteps the regulatory aspects. Also, the DS4IoT ontology provides a structured representation of concepts central to data security in IoT ecosystems. The DS4IoT Ontology [7] is centered on the SecureData class, which represents data annotated with security information, and its specialized subclasses, such as SecretData. Other essential concepts in SecureData are Regulation, Certificate, and ProvenanceProvider, which are classes linked to URLs of corresponding regulations, certification authorities, and provenance providers. The ontology's emphasis on AccessControl mechanisms, such as Role-Based and Attribute-Based Access Control, is consistent with the theme of ensuring robust data security in IoT.

Mayke and Renato [8] offer a novel perspective on privacy protection in the IoT domain by introducing the IoT-Priv ontology tailored to IoT privacy needs. This ontology includes, among other concepts, Data Providers, Recipients, Consumption Requests, and Access Purposes. Notably, the ontology highlights the significance of data consumers' obligations and discusses various anonymization techniques, such as obfuscation, to protect personally identifiable information. Although the IoT-Priv ontology provides a structured method for ensuring privacy protection, it focuses primarily on the technical aspects of privacy. Our ontology, in contrast, aims to bridge both the technical and regulatory aspects of privacy, ensuring a more holistic approach to data privacy in the IoT ecosystem. However, the insights from the IoT-Priv ontology serve as foundational knowledge, particularly when considering the technical mechanisms for enforcing privacy.

Another study by [9] investigates privacy-aware Internet of Things (IoT) experimentation, highlighting the need to address privacy in data collection and sharing. The authors developed an ontology that permits dynamic policy changes and real-time consent management using established methodologies and alignment with ontologies such as SSN. However, it focuses primarily on data collection and sharing, leaving some GDPR-mandated aspects of the IoT data cycle untouched. With 520 classes, they designed their ontology for interoperability with frameworks such as SSN, IoT-lite, and oneM2M. The paper identifies areas for enhancement, especially around consent management and user-centric capabilities. While the paper introduces a privacy graph tailored to GDPR requirements, the IoT-Reg ontology seeks to offer a broader, more integrated framework encompassing other pertinent IoT data lifecycle stages and regulations.

[10] discusses the difficulties of securing personal information throughout the lifecycle of IoT devices. They highlight the vulnerabilities of IoT devices, particularly when they are not physically protected, making them susceptible to unauthorized data extraction. The paper emphasizes the stages of the personal information lifecycle, including collection, storage, use, distribution, and destruction. The second-level classes of our ontology, such as Data Collection, Data Storage, Data Processing, Data Sharing, and Data Deletion, resonate with these stages.

B. Semantic Web

To automate IoT regulations, it is crucial to define the domain knowledge to agree on a common meaning, not only for the data but also for the data protection regulations. One possible approach is to employ Semantic Web techniques for modeling and reasoning about IoT data protection policies. We have used this approach to develop our knowledge graph. The Semantic Web primarily deals with the data instead of documents. It allows data to be annotated with machine-understandable meta-data, allowing the automation of their retrieval and their usage in incorrect contexts. Semantic Web technologies include languages such as Resource Description Framework (RDF) and Web Ontology Language (OWL) for defining knowledge graphs or ontologies and illustrating meta-data using these ontologies as well as tools for reasoning over these descriptions. These technologies can be used to support common semantics of regulatory policies, enabling all agents who understand basic Semantic Web technologies to transmit and use each other's Services and data efficiently. In our prior works, we developed integrated Knowledge graphs to capture various data protection regulations that apply to Big Data [11], [12], [13], [14], [15]. We extracted the rules based on the keywords listed in the glossary or appendix of any regulation. Also, we have developed ontologies to represent legal documents about cloud data like Service Level Agreements [16] [17] and Data Privacy policies [18].

III. IoT PRIVACY ELEMENTS

A. IoT Privacy Regulatory Integration

The IoT domain's expansive and diverse ecosystem has necessitated the establishment of numerous regulations to ensure data privacy, security, and ethical handling. As the IoT landscape continues to evolve, integrating these regulations to provide a coherent and consistent framework for stakeholders is crucial. Several regulations pertain to personal data privacy in the IoT space. For simplicity and with the extensibility of our ontology in mind, we only captured the regulatory interplay of the NISTIR 8228, HIPAA, and GDPR (See Fig.1).

1) NISTIR 8228: This regulation was published to assist organizations in understanding and managing the cybersecurity and privacy risks associated with IoT devices throughout their lifecycles. It describes potential challenges to achieving specific goals, such as protecting device security, data security, and individual privacy. It emphasizes the need for up-to-date organizational policies, procedures, and risk mitigation practices in its recommendations for addressing these challenges.
2) HIPAA: This regulation focuses on healthcare data. Integrating HIPAA regulations with the proliferation of wearable health monitors and other health-related IoT devices ensures that personal health information (PHI) is protected and handled ethically. As a cornerstone of health-related data privacy, HIPAA's regulations were instrumental in shaping the ontology's approach to health data generated by the IoT devices.
3) GDPR: This is a data privacy and security law enforced by the European Union (EU). It imposes obligations on organizations everywhere if they target or collect data related to people in the EU. Its principles, including data minimization, the right to erasure, and data portability, are essential for IoT devices that process the personal data of EU citizens.

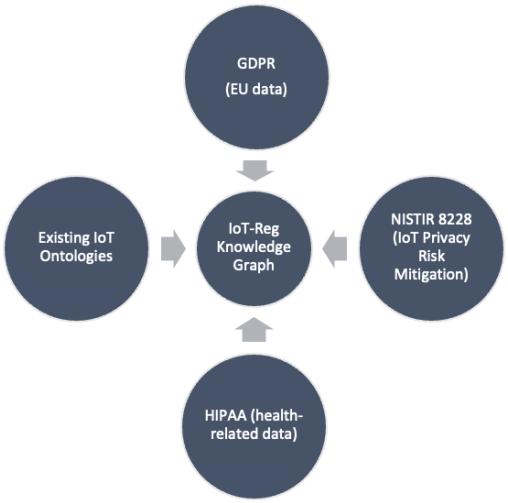

Fig. 1: The IoT-Reg framework, built using policies from three major data protection regulations that regulate IoT privacy

In previous works, we have addressed these regulations in detail and handled the rules in isolation. In [19] [20], for instance, the HIPAA component of the ontology was expanded to include mappings to its respective Privacy and Security Rules, detailing how IoT devices must handle encryption of PHI, ensure data integrity, and maintain data availability. Also, we referenced GDPR's Articles in [21], addressing requirements such as data subject rights, consent mechanisms, and the legal basis for data processing. In [22], we referenced the NISTIR 8228, which provides guidance on risk management and defines goals for identifying, assessing, and mitigating cybersecurity and privacy risks across the IoT landscape. Within IoT-Reg, we capture and operationalize these foundational components and technicalities mandated by the regulations to ensure privacy and security in wearable devices.

We examined these three regulations to extract clauses and guidelines pertinent to the IoT domain, especially wearable devices. Integrating these regulations into the IoT-Reg ontology requires not only amalgamation but also an understanding of the subtleties of each regulation and their alignment with the stages of the IoT data lifecycle. Currently, this review process is manual, and we are working towards automating this knowledge extraction from IoT regulatory documents.

B. IoT Data Lifecycle and Privacy

IoT-Reg adopts a stratified methodology, illustrated in Fig.2, to effectively address privacy concerns throughout the entire

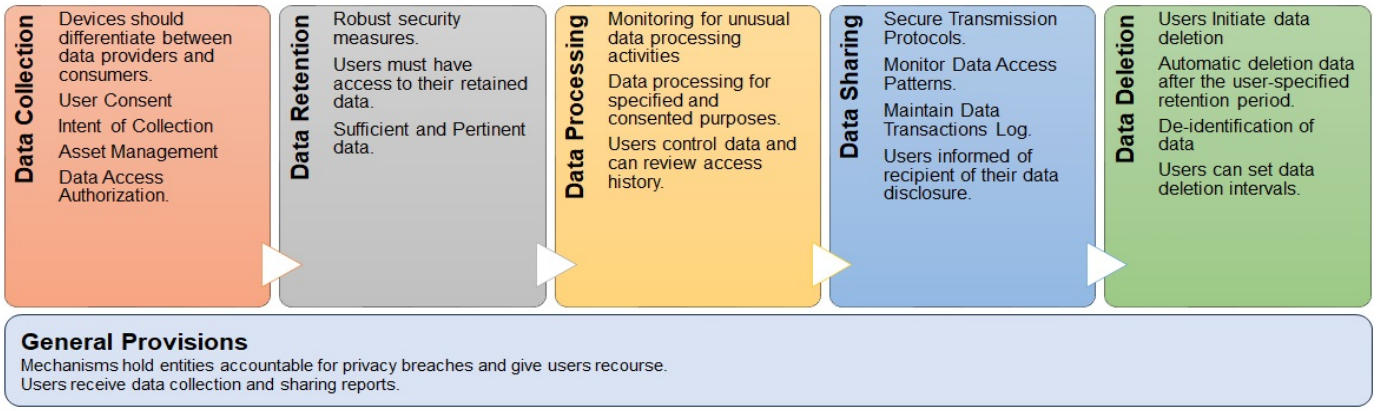

Fig. 2: The comprehensive IoT-Reg ontology incorporates privacy concerns in each phase of the IoT data lifecycle

lifecycle of IoT data, encompassing data collection, retention, sharing, processing, and deletion stages. They constitute the second-level classes of the IoT-Reg ontology, each with distinct privacy requirements for safeguarding personally identifiable information (PII). The differentiation in risk mitigation is of major significance, which we will expound upon in a subsequent section. We derived the ontological privacy requirements presented in this study from analyzing relevant regulations and the scholarly contributions of [8] and [9].

1) **Data Collection**: Data collection devices should differentiate between data providers and consumers. Before data collection, explicit user consent is essential, and users can deny or withdraw consent. The intent of data collection must be transparent, ensuring that users are well-informed. Asset management is essential, with devices keeping track of components and upgrades. Regular software updates should patch known vulnerabilities, and access to specific data should be restricted to authorized parties via appropriate mechanisms. Users should determine which entities can collect their PII by setting permissions.
2) **Data Retention**: Robust security measures, including encryption, must be implemented to protect data retained by the Controller, Covered Entities, and Business Associates. Users must have access to, be able to edit, and be able to delete their retained data. The retained data should be sufficient, pertinent, and not excessive, focusing on minimization.
3) **Data Processing**: Monitoring for unusual data processing activities is essential, as it alerts users to security or privacy breaches. Data processing should adhere to specified purposes, with transparent post-processing and per-user consent. Users must have control over their data and be aware of its access history.
4) **Data Sharing**: Data transmission protocols should prioritize security, preventing data interception and tampering en route. Monitoring data access patterns is crucial for detecting suspicious data access or transfer. Maintaining a comprehensive log of all data transactions is imperative to uphold transparency in data sharing. Also, users should be duly informed of the recipient of their data disclosure.
5) **Data Deletion**: Users should initiate data deletion on their own accord. There should be mechanisms for automatically deleting data after the user-specified retention period. Features that de-identify data ensuring that it cannot be traced back to users, are essential. Users should be able to set specific data deletion intervals.
6) **General Provisions**: Mechanisms should hold entities accountable for privacy breaches and give users recourse. Users should receive transparent data collection and sharing information, allowing informed decision-making. As our ontology focuses on unification, stratification, and risk management, it should be adapted to changing industry standards and user needs. It emphasizes that the IoT can thrive without compromising data privacy and security.

IV. IoT-REG DEVELOPMENT METHODOLOGY

Noy et al [23] have outlined a set of established methodologies for ontology development, which we have adhered to in our approach. Several studies have devised numerous ontologies to address privacy and security issues within the IoT domain. We aim to incorporate many classes from pre-existing IoT ontologies, focusing on meticulously capturing regulatory interactions. In this analysis, we examine Internet of Things (IoT) devices without making distinctions based on their transducing capabilities.

Our IoT-Reg ontology is rooted in a comprehensive understanding of the IoT domain, especially in the context of wearable devices. The ontology encapsulates the regulations and guidelines provided by NISTIR 8228, GDPR, and HIPAA. Considering the IoT data lifecycle, it offers stratified coverage of compliance and risk mitigation areas affecting data privacy in various devices and services. These devices, such as smartwatches, fitness trackers, and health monitors, collect, store, process, share, and eventually delete user data *in no particular order*.

A. Competency Questions

Competency questions are crucial for ontology development because they aid in defining the scope and ensuring the ontology meets its intended purpose. By posing the following deterministic competency questions for the IoT-Reg ontology, we model the stratified interaction between regulations, privacy, and the IoT device, particularly wearables.

- What kinds of information does a particular wearable device collect?
- Which regulations *primarily* govern wearables' collection of health-related data?
- Which regulations mandate user consent before data sharing?
- With which parties are the collected data shared?
- What IoT data lifecycle phase is most affected by the stipulated risk mitigation areas?
- How is user consent for wearable data collection obtained?
- How long does a specific wearable device retain data?
- From what existing ontology does a particular class or property derive?
- What data privacy responsibilities do manufacturers have, as defined by the ontology?

B. IoT-Reg Ontology Definitions

In the following discourse, we expound upon the pivotal classes and properties of the IoT-Reg ontology. Previous studies, reviewed in Section II, frequently examine data privacy regulations independently, without incorporating risk mitigation strategies. To bridge this void, we meticulously designed the IoT-Reg ontology to accommodate these intricate demands while focusing on the multifaceted dimensions of the wearable IoT domain. Fig.3 features some of the top-level classes. The alignment of particular provisions from GDPR, HIPAA, and NISTIR 8228 with ontological components of IoT-Reg is detailed in Table I, showcasing the ontology's ability to not only abstract the specific provisions of these regulations but also substantively embody them, thereby improving its expressiveness and regulatory adherence. IoT-Reg also introduces classes that do not have direct regulatory mappings but are instrumental in the practical enforcement of real-time data privacy adherence.

The ontology further integrates classes from several established ontologies, each contributing to the comprehensive representation of the IoT domain and its regulatory nuances. In addition to the Consent ontology [24] (prefixed with *con:*), we incorporate elements from the Semantic Sensor Network (SSN[1]) ontology, such as *ssn:System* and *ssn:Device*, which provides a foundation for describing the physical layout and capabilities of IoT devices. The *iot-lite:Service* class, derived from the IoT-Lite ontology [25], specifies the functional aspects of IoT services, essential for service discovery and interaction. Additionally, the ontology harnesses classes from the Sensor, Observation, Sample, and Actuator (SOSA[2]) ontology [26], enriching the semantic model with detailed descriptions of sensing. The *iot-priv:* prefix identifies classes/properties from the IoT-Priv ontology [8] that underpins the modeling of access control and privacy policies, while the *priv:* prefixed elements derive from the GDPR-inspired ontology [9] encompassing IoT semantic interoperability and privacy-preserving aspects. For geospatial data integral to IoT device context, we draw upon W3C's Basic Geo[3] (WGS84 lat/long) Vocabulary denoted by the *geo:* prefix, providing precise locational contextualization. We utilize the GDPRtEXT ontology [27] for its pointed GDPR-related classes, further ensuring our ontology's alignment with the stringent privacy regulations of the GDPR. We provide class and property definitions below.

1) **ssn:System** serves as an aggregate of interconnected IoT devices engineered to function as a unified computational unit within an IoT domain. It includes a subclass, **ssn:Device**, which acts as a hardware element executing particular functions within the System, such as wearables. Each Device possesses a **geo:location** property (cardinality: exactly 1) that ranges to **geo:Point**, a geographical coordinate defined by spatial data properties **geo:lat** and **geo:long**, specifying the Device's geographical location. Additionally, **ssn:Device** is tasked with data acquisition via sensors. For the sake of simplicity, our approach consolidates sensors and devices into a single component. Instances of the Device class have the **sosa:madeObservation** property (cardinality: min of 1), representing the act of data collection (**sosa:Observation**). The **sosa:FeatureOfInterest** refers to the environmental or situational aspect under observation or measurement, while **sosa:ObservableProperty** is an attribute or quality subject to measurement or observation, denoted by **sosa:hasObservableProperty** role. The property **sosa:isObservedBy** identifies the Device responsible for the observation, and **priv:ownedBy** indicates the entity holding ownership over the generated data.

iot-lite:Service is a software element offering specific capabilities via IoT Devices. It has an optional **iot-priv:hasServiceDescription** property that describes the Service. We assume all data the Device collects is sensitive and personally identifiable, as in wearables. The Service makes the Device accessible through the **iot-lite:exposes** property. Each Service maintains an access list (**iot-priv:hasAccessList**, cardinality: min of 1), an enumerated set of entities granted permission to utilize the Service. We represent this list by the **iot-priv:AccessList** class. Authorized recipients who can engage with particular Services or data are associated with the **iot-priv:Recipient** class (representing a specific entity given access to resources) via the **iot-priv:hasRecipient** property (cardinality: min of 1), and equivalent to **con:AllowedParty**.

2) **GDPRtEXT:DataSubject** denotes an individual whose personal data is engaged at any point in the IoT data lifecycle, per GDPR. This class is equivalent to **iot-priv:DataProvider** and **con:ConsentingParty** and is characterized by the

[1]https://www.w3.org/ns/ssn/

[2]https://www.w3.org/ns/sosa/

[3]https://www.w3.org/2003/01/geo/wgs84_pos#

TABLE I: Regulatory Mappings of the IoT-Reg Ontology

Ontological Element	Regulatory Reference	Description/Notes
iot-reg:BusinessAssociate	HIPAA 45 CFR § 160.103	Represents entities that perform functions involving the use or disclosure of Protected Health Information.
iot-reg:DataLifecycle	GDPR Articles 5, 17, 30; HIPAA Privacy Rule 45 CFR Part 160 and 164	Describes the phases of data from collection to deletion, ensuring compliance with data handling and retention policies.
iot-reg:ProtectedHealthInformation	HIPAA 45 CFR § 160.103	Defines the type of information covered under HIPAA's privacy and security rules.
iot-reg:SecurityRule	HIPAA Security Rule 45 CFR Part 160 and 164	Outlines the standards for the protection of electronic health information.
iot-reg:DataCollection	GDPR Article 5(1)(b)	Pertains to the lawful collection of personal data with a specific purpose.
iot-reg:DataSharing	GDPR Article 20	Covers the right to data portability and the sharing of personal data between systems.
iot-reg:DataRetention	GDPR Article 5(e)	Details the requirements for the limited retention of personal data.
iot-reg:DataProcessing	GDPR Article 28	Outlines the obligations and roles of data processors.
iot-reg:DataDeletion	GDPR Article 17	Specifies the right to the erasure of personal data ("right to be forgotten").
iot-reg:hasRegulatoryControls	GDPR Article 5(2)	Relates to the demonstration of compliance with the principles of data processing.
iot-reg:RiskMitigationGoal	NISTIR 8228	Defines objectives for reducing risks in IoT data management and operations.
iot-reg:IndividualIoTDeviceChallenge	NISTIR 8228	Describes specific risks and challenges associated with individual IoT devices.
iot-reg:ImplicationsForTheUser	NISTIR 8228	Discusses the implications for users in terms of privacy and security in IoT.
iot-reg:hasDeletionTime	GDPR Article 5(1)(e)	Specifies the conditions and time frame for data storage and deletion.

property **iot-reg:hasPersonalData** (cardinality: min of 1). **GDPRtEXT:PersonalData**, which is the actual data attributed to an individual, is represented by this class which identifies a **GDPRtEXT:DataSubject** and is processed under the supervision of a **GDPRtEXT:Controller**. The data may be disclosed to a designated recipient entity via the **iot-priv:hasRecipient** property (cardinality: min 0, max n, where n is a positive integer). The Controller is the entity responsible for defining the objectives and methods for processing personal data and managing **GDPRtEXT:DataBreach** incidents. The DataBreach class symbolizes a security incident that results in accidental or unlawful access to personal data. The **iot-reg:hasBreachNotificationTime** data property constructs the notification window of 72 hours for GDPR and 60 days for HIPAA. The **iot-reg:notifiesBreach** specifies the subsequent notifications property, which has the following sub-properties: **iot-reg:notifiesController**, with **GDPR:tEXT:Processor** and Controller as its domain and range respectively, and **iot-reg:notifiesDataSubject**, with Controller and DataSubject as its domain and range respectively - representing the act of informing the Data Subject about the usage of their data. The Controller also regulates the observability of a property through the **priv:controls** property. The property under observation and control belongs to the **sosa:ObservableProperty** class. **iot-reg:involvesPersonalData** is a property encompassing all inferences involving personal data, either in processing or during a data breach.

3) **iot-reg:BusinessAssociate** identifies an entity that performs functions or activities for, or provides services to, an **iot-reg:CoveredEntity** that involves using or disclosing **iot-reg:ProtectedHealthInformation**. The **iot-reg:SecurityRule** describes the specific regulations for electronic individually identifiable health data. The **iot-reg:hasHipaaRule** property governs the Covered Entity, a Health Plan, Healthcare Clearinghouse, or Provider responsible for disseminating Protected Health Information.

4) **con:Action** is a human or automated operation on data authorized by **con:Consent**. The **con:activityHasPurpose** property describes the rationale behind the Action. **con:Purpose** defines the objective for processing Personal

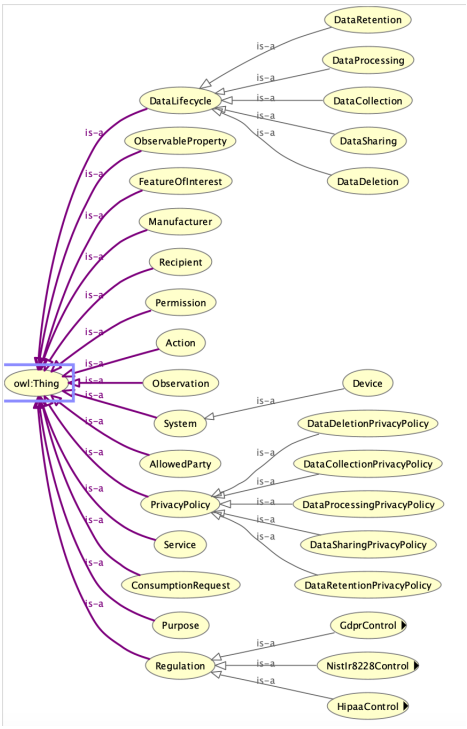

Fig. 3: IoT-Reg Top-Level Classes

Data. **con:Permission** is the formal authorization granted by the Consent to execute the Action, as indicated by the **con:permissionGivenForActivity** property, which specifies the data or attribute (the Observable Property) for which the Permission is granted.

5) **iot-reg:DataLifecycle** describes the various phases data goes through in the IoT ecosystem, from collection to deletion. The class **iot-reg:Regulations** encapsulates the actual regulations that the personal data must adhere to, such as GDPR, HIPAA, and NISTIR_8228. The **iot-reg:isRegulatedBy** property has a domain of **iot-reg:DataLifecycle** and a range of **iot-reg:Regulations**, which specify the regulations that govern the IoT Data Lifecycle. The sub-properties of the **iot-reg:hasDataLifecycleAction** property include collects, shares, deletes, processes, and retains. We represent them by the following five triples, each of which represents a distinct

lifecycle stage:
(iot-reg:DataCollection, iot-reg:collects, Personal Data).
(iot-reg:DataSharing, iot-reg:shares, Personal Data).
(iot-reg:DataRetention, iot-reg:retains, Personal Data).
(iot-reg:DataProcessing, iot-reg:processes, Personal Data).
(iot-reg:DataDeletion, iot-reg:retains, Personal Data).
The property **iot-reg:hasDataLifecycleMethod** defines the method employed for data handling throughout its lifecycle. It has a domain of Action and a range of **iot:reg:DataLifecycleMethod**. For instance, in the Data Collection phase (also applicable to the other four stages), **iot-reg:DataCollection** involves *collecting* personal data, characterized by the **iot-reg:collects** property. **iot-reg:hasDeletionTime** specifies the deletion time for retained data. The property **iot-reg:hasRegulatoryControls** specifies the regulatory controls applicable to any of the **iot-reg:DataLifecycle** subclasses (**iot-reg:GdprControls**, **iot-reg:HipaaControls**, or **iot-reg:NistIr8228Controls**).

6) **iot-reg:RiskMitigationGoal** signifies the objectives to reduce risk in the personal data lifecycle. The three goals (*Goal1*, *Goal2*, and *Goal3*) are outlined in the NISTIR_8228 publication and are addressed by the IoT-Reg ontology through the **iot-reg:addressesRiskMitigationArea** property, targeting specific risk mitigation areas (**iot-reg:RiskMitigationArea**). The property **iot-reg:aimsForGoals** focuses on particular risk mitigation goals based on the current IoT data lifecycle stage. We represent the challenges unique to individual IoT devices by the **iot-reg:IndividualIoTDeviceChallenge** class. This class describes the challenges IoT devices may pose to **iot-reg:Expectation**, as specified by the **iot-reg:hasExpectation** property. **iot-reg:Expectation** represents built-in IoT device functionalities that could mitigate risk. The **iot-reg:hasChallenge** property specifies the challenges based on the affected **iot-reg:NISTSP800-53Revision5Controls** and the corresponding user implications (**iot-reg:ImplicationsForTheUser**), particularly in wearables.

7) The IoT-Reg ontology acknowledges that manufacturers' privacy policies play a crucial role. They are essential to ensuring these IoT devices are inherently compliant with regulations. Manufacturers can demonstrate in writing that they know the possibility of data breaches and have taken precautions to prevent such incidents. By integrating regulatory requirements, manufacturers can create robust privacy policies aligning with global privacy and security standards, fostering trust and ensuring compliance. Privacy Policies should supplement the privacy-by-design campaign [28]. We consolidate this by leveraging our previous work [22] to ensure that wearable IoT devices have privacy policies that address the eleven risk mitigation areas stipulated in the NISTIR 8228 publication.

iot-priv:PrivacyPolicy class represents a formal, legally enforceable set of rules and practices that govern the management and protection of personally identifiable information. Through the **iot-priv:hasOptIn** property, a data recipient conforms to a specific privacy policy. Each instance of **iot-priv:DataProvider** (an entity providing data or a service generating data), as specified by the **iot-priv:hasPrivacyPolicy** property (max cardinality: 1), must be governed by one Privacy Policy. The same entity also possesses the **iot-priv:providesConsentTo** property, which represents granting consent for data usage in response to a formal data consumption request (**con:ConsumptionRequest**). Given that multiple consents may be granted, its cardinality ranges from 1 to a finite positive integer, n. The **iot-priv:isProtectedBy** property identifies the Privacy Policy associated with and protecting the Service. **iot-priv:isCoveredBy** (cardinality: min 1, max n) identifies the policy governing a particular **iot-reg:RiskMitigationArea**. For enhanced user documentation and explainability, the PrivacyPolicy class is linked to the following data properties:
iot-reg:hasManufacturerDetail: Ranges to ^^*xsd:string* and provides details about the IoT device's manufacturer, such as location or contact information.
iot-reg:hasPolicyUrl: Ranges to ^^*xsd:string* and specifies a URL where the complete privacy policy can be accessed.
iot-reg:hasEffectiveDate (cardinality: exactly 1): Ranges to ^^*xsd:dateTime* and marks the date when the Privacy Policy goes into effect.
iot-reg:hasUpdateDate (cardinality: exactly 1): Ranges to ^^*xsd:dateTime* and marks the date when the Privacy Policy was last revised.

V. RISK MANAGEMENT ONTOLOGICAL MODELING - A USE CASE SCENARIO

Using a use case and a backward approach, we demonstrate our ontological risk management model, which enables actionable guidelines. These guidelines ensure regulatory compliance and allow end-users to make informed decisions about their data while benefiting from the advancements in healthcare made possible by wearable IoT devices. The individual risk mitigation areas are not exclusive to their corresponding IoT data lifecycle class mapping indicated below. The connected classes are the risk mitigation area's most affected IoT data lifecycle stage. Fig.4 illustrates this further.

A. Scenario: Privacy in Health Monitoring SmartWatch

John, a 45-year-old health-conscious individual, has recently decided to take real-time control of his health. He acquires a Fitbit smartwatch to track his daily activities, sleep patterns, and heart rate. John's smartwatch continuously collects and transmits big data to the Fitbit app on his smartphone. The app processes and sends the collected data to a cloud server, where it is further processed and analyzed to provide John with insights into his health and fitness. Depending on the data lifecycle phases:
The Fitbit watch *collects* John's Personal Data (Name, Age, Gender, Weight, and Height), Health Data (Heart Rate, Step Count, Sleep Pattern, and Calories Burned), and Location Data (GPS coordinates of his daily routes). The Fitbit app *processes* the collected data to provide insights into John's daily activities, including Step Count, Distance Covered, and

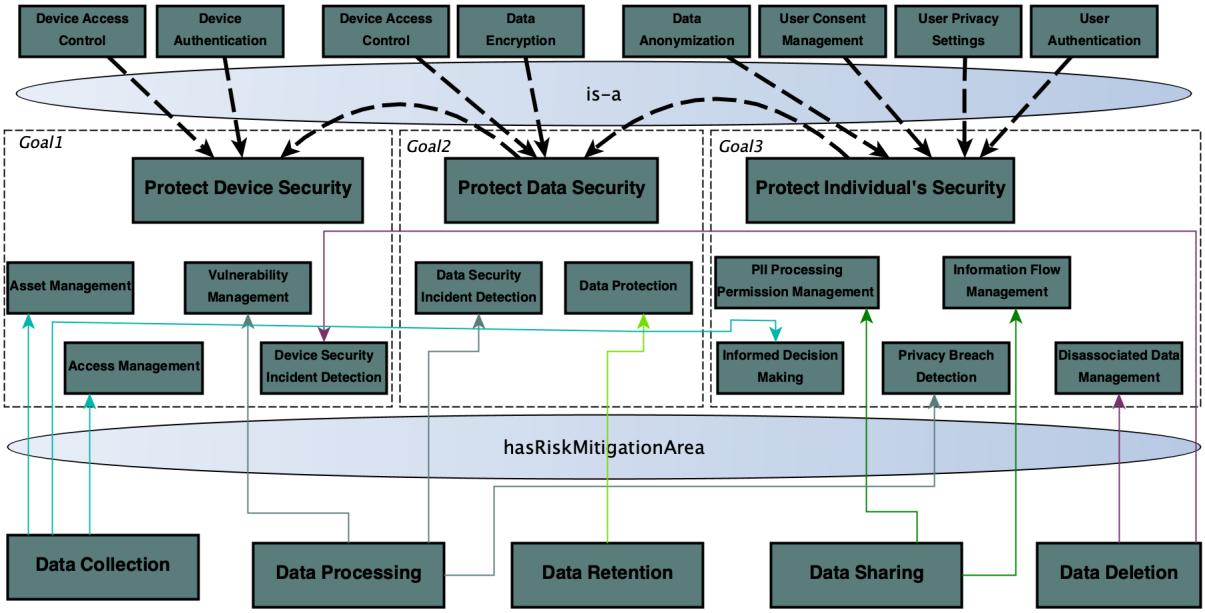

Fig. 4: Wearable IoT data privacy risk mitigation ontological model

Calories Burned, Sleep Analysis (including Sleep Duration, Sleep Quality, and Sleep Stages), and Heart Rate Analysis (including Resting Heart Rate and Heart Rate Zones during workouts). John decides to *share* his heart health data with his physician and a health research organization for a study. The data is securely *stored* on the cloud servers of Fitbit. John has the right to request the *deletion* of his personal data at any time. If he decides to stop using the Fitbit watch, he can request that all his data be deleted.

B. IoT-Reg Framework Application

Applying the HIPAA and NISTIR_8228 policies in IoT-Reg privacy risk mitigation areas to this use case of John's health monitoring smartwatch provides insights into how we can manage these areas effectively. We addressed them regarding their most pertinent IoT data lifecycle stages, followed by an associated triple in IoT-Reg.

Data Collection Phase:

1) Asset Management: John's smartwatch accurately logs each heartbeat, step, and sleep cycle by recording all sensor activities. It keeps track of all its sensors, applications, and other components. The integrity of sensors and applications directly impacts the quality and dependability of health data, making Asset Management essential during Data Collection. *(:JohnSmartwatch, iot-reg:hasAsset, :SensorRecord)*

2) Access Management: John's smartwatch has a built-in locking mechanism to restrict access to his heart rate, sleep patterns, and other health metrics to only authorized apps. Access Management becomes critical during the Data Collection phase for wearables, where unauthorized data collection could directly breach personal and health data. *(:JohnSmartwatch, iot-reg:hasAccessControl, :AuthorizedApps)*

3) Informed Decision Making: When John sets up his smartwatch, it provides clear, understandable information about the types of data that will be collected, allowing him to make informed decisions - thoroughly informing wearable device users of the data collected by their devices. Consequently, the stage of Data Collection is essential for informed consent and decision-making. *(:JohnSmartwatch, iot-reg:providesInformation, :DataCollectionInfo)*

Data Retention Phase:

4) Data Protection: John's smartwatch stores his health information in an encrypted format, preventing unauthorized access to the data. Due to the highly sensitive data that wearables store, the DataProtection class becomes exceedingly critical during the Data Retention stage to ensure encryption and security measures are in place. *(:JohnHealthData, iot-reg:isEncrypted, True)*

Data Sharing Phase:

5) Information Flow Management: John's smartwatch records all instances in which data was shared, with whom, and for what purpose, providing him with a clear understanding and control over his shared data. This class is essential during the data-sharing phase of wearables, as health data may be shared with multiple entities, such as healthcare providers and fitness apps. *(:JohnSmartwatch, iot-reg:hasTransactionLog, :DataTransactionRecord)*

6) PII Processing Permissions Management: John can choose which apps or services can process his PII, such as his location data or heart rate statistics. The Personally Identifiable Information (PII) collected by wearables is particularly sensitive. Therefore, permissions must be

explicitly set, especially during the Data Sharing phase. *(:John, iot-reg:hasPIIPermission, :AllowedApps)*

Data Processing Phase:

7) Privacy Breach Detection: John's smartwatch has built-in mechanisms to alert him in real time if it detects a potential privacy breach during data processing, such as unauthorized data access or manipulation. Since data is most susceptible to malicious manipulation during Data Processing, algorithms to detect potential privacy breaches become imperative. *(:JohnSmartwatch, iot-reg:hasBreachDetection, :PrivacyBreachAlert)*

8) Data Security Incident Detection: The smartwatch employs machine learning algorithms to flag suspicious data processing or access patterns, immediately alerting John's smartphone. During Data Processing, monitoring abnormal patterns that may indicate a security incident is vital. *(:JohnSmartwatch, iot-reg:hasDataMonitoring, :SuspiciousAccessAlert)*

9) Vulnerability Management: There are mechanisms on John's smartwatch that alert him in real time if it detects a potential privacy breach during data processing, such as unauthorized access or unusual data manipulation, as mentioned earlier. As the Data Processing stage frequently involves software-level manipulations, it is essential to patch vulnerabilities during this phase actively. *(:JohnSmartwatchSoftware, iot-reg:hasPatch, :SecurityPatch)*

Data Deletion Phase:

10) Disassociated Data Management: Before John deletes his historical data from the smartwatch, it allows him to anonymize specific datasets so they cannot be traced back to him. Before data deletion, wearables should provide anonymization features. During the Data Deletion phase, this class becomes particularly significant. *(:JohnHealthData, iot-reg:isAnonymized, True)*

11) Device Security Incident Detection: When John initiates data deletion, his smartwatch performs a security check to ensure that the data is being deleted securely and notifies him if any security breaches are detected during the process. During Data Deletion, security measures to detect any unusual activities or security breaches are of the utmost importance, ensuring that data is securely and permanently removed from the device. *(:JohnSmartwatch, iot-reg:hasIncidentDetection, :UnusualActivityAlert)*

VI. IoT-Reg Ontology Rule Definitions

To assess the validity of our ontology, we have formulated a set of rules to improve our ontology's robustness, expressivity, and inferential capabilities, with a particular emphasis on the wearable IoT domain. These rules ensure that the IoT-Reg ontology can capture and enforce critical elements of data privacy, cybersecurity, and regulatory compliance. We use Description Logic (DL), a formalism that enables the specification of intricate relationships, constraints, and reasoning processes, to codify the rules that govern our ontology. Using Semantic Web Rule Language (SWRL) and the Pellet reasoning engine, we implemented the following rules in Protégé. See Fig. 5.

A. User Consent

In data privacy, user consent is a pillar. Our ontology stipulates that any Device class' device that collects personal information must obtain explicit user consent. We modeled this requirement using the Consent class, ensuring data collection with consent.

DL: $\forall d \in \text{Device}, \exists o \in \text{Observation}, \exists pd \in \text{PersonalData}, \exists c \in \text{Consent}.(d \,\text{madeObservation}\, o \wedge o \,\text{involvesPersonalData}\, pd) \Rightarrow pd \,\text{hasConsent}\, c$

B. Location Privacy

Numerous wearable Internet of Things (IoT) devices have geolocation capabilities. Our ontology necessitates that devices with these capabilities that run location-dependent services adhere to a privacy policy encapsulated by the PrivacyPolicy class, thereby protecting location privacy.

DL: $\forall d \in \text{Device}, \exists o \in \text{Observation}, \exists pd \in \text{PersonalData}, \exists c \in \text{Consent}.(d \,\text{madeObservation}\, o \wedge o \,\text{involvesPersonalData}\, pd) \Rightarrow pd \,\text{hasConsent}\, c$

C. Data Deletion After Use

The principle of data minimization is foundational to data protection regulations. Once the primary purpose of data collection has been met, our ontology dictates that the data should be marked for deletion appropriately. The range of the hasDeletionTime property specifies the deletion time.

DL: $\forall o \in \text{Observation}, \exists p \in \text{Purpose}, \exists pd \in \text{PersonalData}, \exists dt \in \text{xsd:dateTime}.(o \,\text{activityHasPurpose}\, p \wedge o \,\text{involvesPersonalData}\, pd) \Rightarrow (o \,\text{activityHasPurpose}\, \text{deletionpurpose} \wedge o \,\text{hasDeletionTime}\, dt)$

D. Access Control

It is essential to restrict data access to prevent unauthorized data dissemination. Our ontology specifies that if a service, denoted by the class Service, maintains an AccessList class, only the entities on the list, designated by the class Recipient (equivalent to the AllowedParty class), may access the data.

DL: $\forall s \in \text{Service}, \exists al \in \text{AccessList}, \exists r \in \text{Recipient}.(s \,\text{hasAccessList}\, al \wedge al \,\text{hasRecipient}\, r) \Rightarrow r \,\text{permissionGivenTo}\, s$

VII. Conclusion and Future Work

We have developed a semantically rich IoT-Reg framework to capture the policies and rules for IoT data privacy as articulated in data protection regulations like NISTIR 8228, GDPR, and HIPAA. Currently, due to the textual nature of regulatory standards, it is challenging for wearable IoT vendors to comply with the rules outlined in these documents in real time. This framework enables both IoT providers and users to query and reason over the integrated IoT-Reg knowledge graph to identify all rules that apply to their IoT device or application, thereby facilitating real-time IoT data compliance.

This study is a part of our goal of developing a semantically rich framework to automate IoT data privacy compliance.

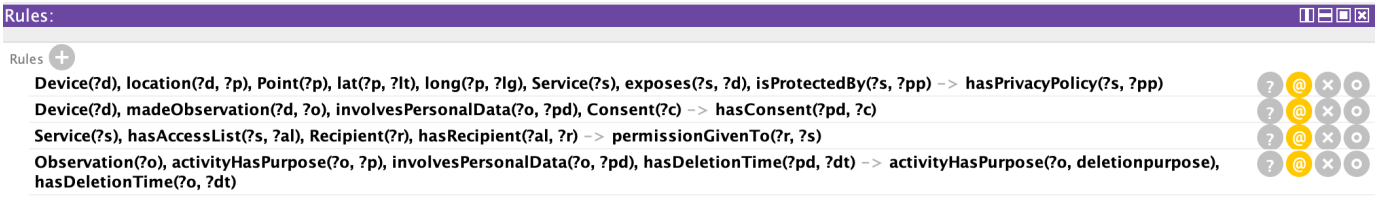

Fig. 5: SWRL rules

As part of our future work, we are developing sophisticated text extraction approaches to automatically populate the IoT-Reg knowledge graph with future policy changes in IoT privacy regulations. The IoT-Reg ontology is not just a static framework but a dynamic, evolving entity designed to meet the ever-changing landscape of IoT privacy regulations.

ACKNOWLEDGMENT

This research was partially supported by the NSF award 1747724, Phase I IUCRC UMBC: Center for Accelerated Real time Analytics (CARTA).

REFERENCES

[1] 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.

[2] Centers for Medicare & Medicaid Services, "The Health Insurance Portability and Accountability Act of 1996 (HIPAA)," Online at http://www.cms.hhs.gov/hipaa/, 1996.

[3] European Parliament and Council of the European Union. Regulation (EU) 2016/679 of the European Parliament and of the Council. [Online]. Available: https://data.europa.eu/eli/reg/2016/679/oj

[4] G. Bajaj, R. Agarwal, P. Singh, N. Georgantas, and V. Issarny, "A study of existing ontologies in the iot-domain," *arXiv preprint arXiv:1707.00112*, 2017.

[5] V. M. Tayur and R. Suchithra, "A comprehensive ontology for internet of things (coiot)," in *2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP)*. IEEE, 2019, pp. 1–6.

[6] B. A. Mozzaquatro, R. Jardim-Goncalves, and C. Agostinho, "Towards a reference ontology for security in the internet of things," in *2015 IEEE International Workshop on Measurements & Networking (M&N)*. IEEE, 2015, pp. 1–6.

[7] P. Gonzalez-Gil, J. A. Martinez, and A. F. Skarmeta, "Lightweight data-security ontology for iot," *Sensors*, vol. 20, no. 3, p. 801, 2020.

[8] M. F. Arruda and R. F. Bulcão-Neto, "Toward a lightweight ontology for privacy protection in iot," in *Proceedings of the 34th ACM/SIGAPP symposium on applied computing*, 2019, pp. 880–888.

[9] R. Agarwal, T. Elsaleh, and E. Tragos, "Gdpr-inspired iot ontology enabling semantic interoperability, federation of deployments and privacy-preserving applications," in *2022 IEEE 8th World Forum on Internet of Things (WF-IoT)*. IEEE, 2022, pp. 1–8.

[10] Y. Lee and G. Y. Lee, "Security management suitable for lifecycle of personal information in multi-user iot environment," *Sensors*, vol. 21, no. 22, p. 7592, 2021.

[11] L. Elluri and K. P. Joshi, "A Knowledge Representation of Cloud Data controls for EU GDPR Compliance," in *11th IEEE International Conference on Cloud Computing (CLOUD)*, July 2018.

[12] A. Nagar and K. P. Joshi, "A Semantically Rich Knowledge Representation of PCI DSS for Cloud Services," in *6th International IBM Cloud Academy Conference ICACON 2018, Japan*. IBM, May 2018.

[13] L. E. Ankur Nagar and K. P. Joshi, "Automated Compliance of Mobile Wallet Payments for Cloud Services," in *7th IEEE International Conference on Big Data Security on Cloud (BigDataSecurity 2021)*, IEEE. IEEE, May 2021.

[14] K. P. Joshi and S. Saha, "A Semantically Rich Framework for Knowledge Representation of Code of Federal Regulations (CFR)," *Digital Government: Research and Practice*, December 2020.

[15] L. E. Karuna Pande Joshi and A. Nagar, "An Integrated Knowledge Graph to Automate Cloud Data Compliance," *IEEE Access*, vol. 8, pp. 148 541–148 555, November 2020.

[16] C. P. Sudip Mittal, Karuna Pande Joshi and A. Joshi, "Automatic Extraction of Metrics from SLAs for Cloud Service Management," in *2016 IEEE International Conference on Cloud Engineering (IC2E 2016)*, April 2016.

[17] K. P. Joshi and C. Pearce, "Automating Cloud Service Level Agreements using Semantic Technologies," in *CLaw Workshop, IEEE International Conference on Cloud Engineering (IC2E)*, IEEE. Tempe, AZ, USA: IEEE Computer Society, March 2015.

[18] S. M. C. P. A. J. Karuna Pande Joshi, Aditi Gupta and T. Finin, "Semantic Approach to Automating Management of Big Data Privacy Policies," in *IEEE BigData 2016*, December 2016, p. 10.

[19] K. P. Joshi, Y. Yesha, T. Finin *et al.*, "An ontology for a hipaa compliant cloud service," in *4th International IBM Cloud Academy Conference ICACON 2016*, 2016.

[20] D.-y. Kim and K. P. Joshi, "A semantically rich knowledge graph to automate hipaa regulations for cloud health it services," in *2021 7th IEEE Intl Conference on Big Data Security on Cloud (BigDataSecurity), IEEE Intl Conference on High Performance and Smart Computing,(HPSC) and IEEE Intl Conference on Intelligent Data and Security (IDS)*. IEEE, 2021, pp. 7–12.

[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] K. Echenim, L. Elluri, K. P. Joshi *et al.*, "Ensuring privacy policy compliance of wearables with iot regulations," in *IEEE International Conference on Trust, Privacy and Security in Intelligent Systems, and Applications (IEEE TPS 2023)*, 2023.

[23] N. F. Noy, D. L. McGuinness *et al.*, "Ontology development 101: A guide to creating your first ontology," 2001.

[24] K. Fatema, E. Hadziselimovic, H. J. Pandit, C. Debruyne, D. Lewis, and D. O'Sullivan, "Compliance through informed consent: Semantic based consent permission and data management model." *PrivOn@ ISWC*, vol. 1951, pp. 1–16, 2017.

[25] M. Bermudez-Edo, T. Elsaleh, P. Barnaghi, and K. Taylor, "Iot-lite: a lightweight semantic model for the internet of things," in *2016 INTL IEEE conferences on ubiquitous intelligence & computing, advanced and trusted computing, scalable computing and communications, cloud and big data computing, internet of people, and smart world congress (uic/atc/scalcom/cbdcom/iop/smartworld)*. IEEE, 2016, pp. 90–97.

[26] K. Janowicz, A. Haller, S. J. Cox, D. Le Phuoc, and M. Lefrançois, "Sosa: A lightweight ontology for sensors, observations, samples, and actuators," *Journal of Web Semantics*, vol. 56, pp. 1–10, 2019.

[27] H. J. Pandit, K. Fatema, D. O'Sullivan, and D. Lewis, "Gdprtext-gdpr as a linked data resource," in *European semantic web conference*. Springer, 2018, pp. 481–495.

[28] A. Aljeraisy, M. Barati, O. Rana, and C. Perera, "Privacy laws and privacy by design schemes for the internet of things: A developer's perspective," *ACM Computing Surveys (Csur)*, vol. 54, no. 5, pp. 1–38, 2021.