Data Privacy Ontology for Ubiquitous Computing

Narmeen Zakaria Bawany
FAST-National University of Computer and Emerging Sciences, Karachi, Pakistan
Jinnah University for Women, Karachi, Pakistan

Zubair A. Shaikh
President
Mohammad Ali Jinnah University,
Karachi, Pakistan

***Abstract*—Privacy is an ability to understand, choose, and regulate what personal data one shares, with whom, for how long and under what context. Data owners must not lose the rights of ownership, once the data is shared. Privacy decisions have many components that include identity, access granularity, time and context. We propose an ontology based model for data privacy configuration in terms of producer and consumer. Producer is an IP entity who owns data, that is Data owner. Consumer is an IP entity with whom data is shared. We differentiate between consumer and data holder, also and IP entity, which may not have similar access rights as consumer. As we rely on Semantic Web technologies to enable these privacy preferences, our proposed vocabulary is platform independent and can thus be used by any system relying on these technologies. Ideally, producers can specify a set of attributes which consumers must satisfy in order to be granted access to the requested information. Privacy can be configured not only in terms of typical read and edit, but novel attributes like location and time are also included in the proposed ontology.**

Keywords—Privacy Ontology; Data Privacy; Location based privacy; Time based privacy; Ubiquitous Computing

I. INTRODUCTION

Security and privacy are two growing concerns in developing and deploying ubiquitous computing systems. The problem of privacy and protection of personal data has been addressed in literature since long[1], [2]. However this issue has been aggravated with new computing paradigms such as ubiquitous computing.

Unauthorized use of personal data has become significant threat to persons' privacy[3]. Although, privacy leakages may lead to untoward incidents[4], the mere advantages that information and communication systems provide in terms of usability and user comfort will surely outweigh privacy concerns for most users. Realizing the escalating concerns legislative acts such as Health Insurance Portability and Accountability Act (HIPAA) [5] for healthcare and Gramm Leach Bliley Act (GLBA) [6] for financial institutions has been formed. Various strategies have been adopted to protect customers' privacy such as P3P[7], TRUSTe[8], ESRB, BBBOnline, and CPAWebTrust. However these policies fail to provide any systematic mechanism that can put privacy protection into the place, providing no assurance at grassroots level. In all such systems, data owner does not know how personal data is actually handled after it is collected.

Preventing users from sharing data is not a viable strategy for privacy protection. Thus, we present a better stratagem that treats data as an asset of its owner. The Data Policy Ontology defines vocabulary for representing privacy policies on data. Policy is a set of rules that is specified by a producer that is data owner, to restrict data access. Data is an asset of its owner hence producer must be able to set terms and conditions on the usage of data. Producer may use policies to configure who is allowed to read, edit and share data. Data usage may also be protected in terms of time and location. Producer maintains the rights on data even after it is shared to variety of users. We call those users as Consumers. The basic idea is, data can only be consumed by those consumers that are allowed by producer after satisfying the privacy requirements setup by the producer. We present an approach that focuses on maintaining the privacy of data through-out its life cycle. That is data can be shared without losing data ownership and its access rights. We also present a novel idea of protecting data privacy by applying policies with respect to time, sharing medium and location.

II. RELATED WORK

The Policy is a technique for controlling and adjusting the low-level system behaviors by specifying high-level rules. Current implementations have been limited to Role based access and policies are defined in terms of read and write. The Context Broker Architecture (CoBrA) is a broker-centric, agent-based architecture for supporting context-aware computing in intelligent spaces[9]. Standard Ontology for Ubiquitous and Pervasive Applications (SOUPA) is designed to model and support pervasive computing applications includes modular component vocabularies to represent intelligent agents with associated beliefs, desires, and intentions, time, space, events, user profiles, actions, and policies for security and privacy [10]. This ontology typically revolves around specifying policies to restrict the type of personal information that can be shared by the public services. Gaia is an infrastructure for smart spaces, which are pervasive computing environments that encompass physical spaces. The main characteristic of Gaia is that it brings the functionality of an operating system to physical spaces. It employs common operation system functions including events, signals, file systems, security, and processes), and extends them with context, location awareness, mobile computing devices, and actuators. Using this functionality, Gaia integrates devices and physical spaces, and allows the physical and virtual entities to seamlessly interact [11]. Policy languages such as P3P enables Websites to express their privacy practices in a standard format. These languages were defined to automatically enforce privacy specifications but those languages usually lack a formal semantics [7]. Spiekermann and Cranor use a three-layer model of user privacy concerns to relate them to system operations (data transfer, storage, and processing) and examine their effects on user behavior. They

also presents two approaches "privacy-by-policy" and "privacy-by-architecture." The privacy-by policy approach focuses on the implementation of the notice and choice principles of fair information practices, while the privacy-by architecture approach minimizes the collection of identifiable personal data and emphasizes anonymization and client-side data storage and processing. [12]

III. Data Privacy Ontology

We present hypothetical scenarios to illustrate the usage of proposed model.

A. Smart Office Scenario

Imagine that Mr. Ahmed, marketing representative of company ABC is invited by Mr. Salim, Regional Sales manager company XYZ at latter's office to discuss a potential business deal. Mr. Salim shares an official document with Mr. Ahmed for meeting discussion only. The document must not be accessible outside Mr. Salim's office.

The figure 1 represents the scenario in terms of proposed ontology. Mr. Salim is an instance of class Producer and Mr. Ahmed is an instance of class Consumer. The official document, "Official_Doc" is an instance of class Data, on which Location Privacy Policy is configured by creating an instance pp1122. Mr. Salim enables the location privacy policy on his document and sets the accessible location to his office. Mr. Ahmed, consumer of data, will not be allowed to access this document, the data, outside the location set by Mr. Salim, the producer of data.

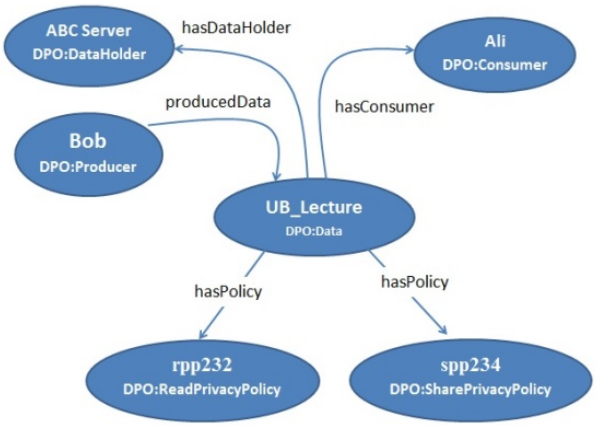

Fig. 1. Location Privacy Policy Scenario

B. Virtual Classroom Scenario

Consider another scenario. Professor Bob has hosted his video lectures on ABC server. Professor allows his registered students to view the lecture once only. Also he does not want his lectures to be shared via email or on social networks. The figure 2 illustrates this scenario in terms of proposed ontology. Professor Bob is an instance of Producer class and Ali is the instance of Consumer class. ABC server is an instance of DataHolder class where video lecture titled UB_Lecture, an instance of Data class, is hosted. ReadPrivacyPolicy class instance rpp232 is configured such that ReadPrivacyPolicy property ViewLimit is set to one and ReadOnly is set true. SharePrivacyPolicy instance, spp234 is configured such that canShare property is set to false.

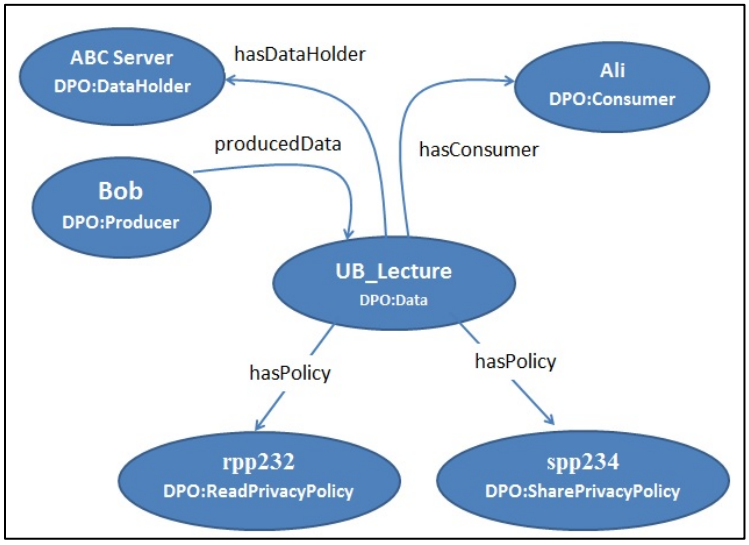

Fig. 2. Share Privacy Policy Scenario

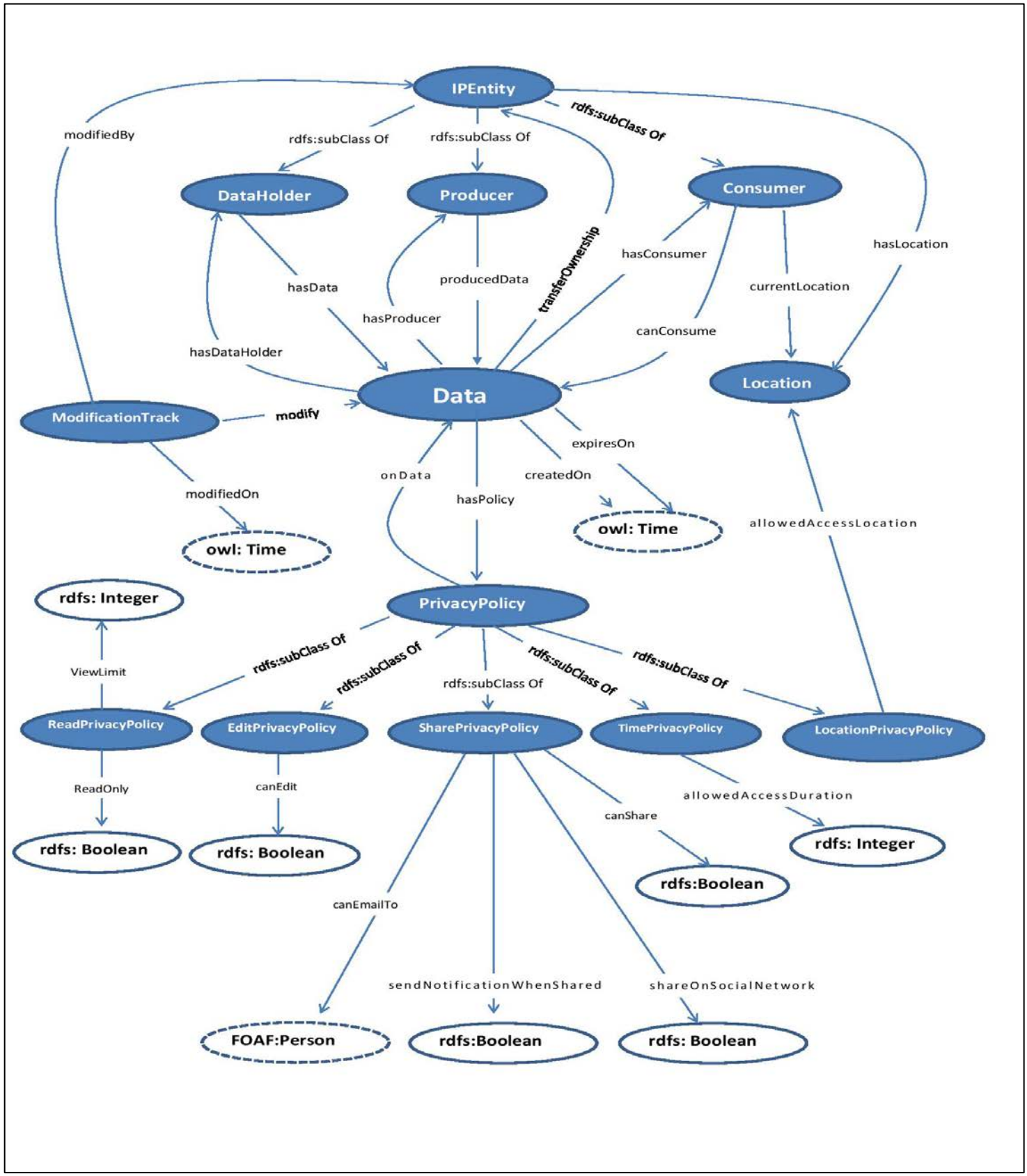

Fig. 3. Data Privacy Ontology (DPO)

C. Virtual Recruitment System Scenario

Ms. Sarah applies for job in Star Security Organization. The organization states in the advertisement that job application process will be completed within a week. She shares her resume with the Human Resource Manager of the organization. However, Sarah does not want her resume to be available to organization after the job application process has been completed.

Figure 4 shows how TimePolicy can be used to set the duration of access. Sarah configures the TimePolicy and sets the duration for access to one week. Note, duration time will be set in minutes. Once the configured duration has passed, Sarah's resume will not be accessible to organization.

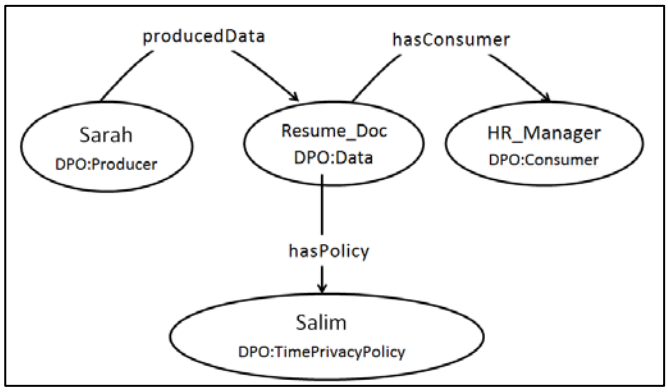

Fig. 4. Time Privacy Policy Scenario

IV. The Privacy Manager

Figure 4 illustrates work flow of the system. Producer defines the data privacy policy using privacy configuration manager. An instance of policy is created and transmitted to the policy enforcer. Consumer request for data is passed through the policy enforcer. The Policy enforcer will permit the access only if it is allowed in policy.

For example a consumer requests to share the data via email. The readOnly property is true and canEmail property of the particular data is set to false. Policy enforcer will not allow to email this data.

V. Description of Ontology

This section presents a brief description of data privacy ontology. The Figure 3 shows complete layout of the proposed data privacy ontology. Producer, Consumer and DataHolder are subclasses of IP entity.

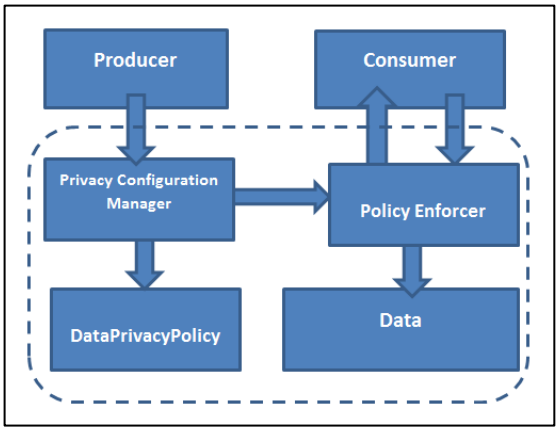

Fig. 5. Privacy Manager

The ontology representation of privacy policy is defined by PrivacyPolicy class. This class has five subclasses namely, ReadPrivacyPolicy, EditPrivacyPolicy,SharePrivacyPolicy, LocationPrivacyPolicy and TimePrivacyPolicy. ReadPrivacyPolicy has data properties that can restrict the data access to read only and set a limit on number times a document can be viewed.

SharePrivacyPolicy class has object properties to restrict sharing of data via email or social networks. The sendNotificationWhenShared property is used to enable notifications to producer whenever data is shared.

TimePrivatePolicy is used to define duration after which data will not be accessible.

LocationPrivacyPolicy is used to define location where data will remain accessible. Location can be both physical or virtual.

Location class is used for describing sensed location context of a consumer or an object. The location context is information that describes the whereabouts of a consumer or an object, which includes both temporal and spatial properties.

VI. Conclusion and Future Work

We presented an ontology based solution for data privacy in ubiquitous computing environment. In contrast to role based security models, our model presents a novel idea of protecting the data by embedding the security model within. We argue that data remains the property of its owner and its privacy and security must be maintained throughout its lifecycle. Data must be accessible to its legitimate users and

the terms of usage shall be dictated by its producer. We intend to classify data with respect to its type in future. We plan to build an open source systems based on this ontology to prove the effectiveness of this research. As the concept of smart cities is now beginning to be implemented, we believe this research will open new directions in protecting users' privacy.

ACKNOWLEDGEMENTS

The authors are grateful to the National University of Computer and Emerging Sciences, Karachi for providing support in carrying out this research.

REFERENCES

[1] Pedar, A. "Privacy, Security, and Protection in Distributed Computing Systems." Offene Multifunktionale Büroarbeitsplätze und Bildschirmtext. Springer Berlin Heidelberg, 1985. 230-246..

[2] S. Akl and P. Taylor, "Cryptographic solution to a problem of access control in a hierarchy," ACM Trans. Comput. Syst., 1983.

[3] "Privacy- Unauthorized access report of June 2012," 2012. [Online]. Available: http://www.bcit.ca/privacy/faq.shtml.

[4] "Security Fix - Payment Processor Breach May Be Largest Ever," 2009. [Online]. Available: http://voices.washingtonpost.com/securityfix/2009/01/payment_processor_breach_may_b.html. [Accessed: 22-Jan-2014].

[5] "Health Insurance Portability and Accountability Act of 1996 (HIPPA)."

[6] X. Zhang, L. T. Yang, C. Liu, and J. Chen, "A Scalable Two-Phase Top-Down Specialization Approach for Data Anonymization Using MapReduce on Cloud," IEEE Trans. Parallel Distrib. Syst., p. 1, 2013.

[7] W. W. W. Consortium, "Platform for privacy preferences (P3P) project," June. Retrieved November., 2000.

[8] "TRUSTe.org. An independent, nonprofit enabling trust based on privacy for personal information on the internet." [Online]. Available: http://www.truste.org/.

[9] Chen, Harry, Tim Finin, and Anupam Joshi. "An ontology for context-aware pervasive computing environments." The Knowledge Engineering Review 18.03 (2003): 197-207.

[10] Chen, H., Perich, F., Finin, T., & Joshi, A. " Soupa: Standard ontology for ubiquitous and pervasive applications", In IEEE Mobile and Ubiquitous Systems: Networking and Services, (pp. 258-267), 2004

[11] Román, M., Hess, C., Cerqueira, R., Ranganathan, A., Campbell, R. H., & Nahrstedt, K. (2002). A middleware infrastructure for active spaces. IEEE pervasive computing, 1(4), 74-83.

[12] Spiekermann, S., & Cranor, L. F. (2009). Engineering privacy. Software Engineering, IEEE Transactions on, 35(1), 67-82.

[13] Modeling privacy control in context-aware systems. Jiang, X., & Landay, J. A. 2002, Pervasive Computing, IEEE, pp. 1(3), 59-63.

[14] On the anonymity of home/work location pairs. Golle, P., & Partridge, K. 2009, Pervasive Computing, Springer Berlin Heidelberg., pp. 390-397.

[15] An approach for privacy protection based-on ontology. Gao, F., He, J., Peng, S., Wu, X., & Liu, X. 2010. Second International Conference on Networks Security Wireless Communications and Trusted Computing(NSWCTC). pp. Vol. 2, pp. 397-400.

[16] Conformance verification of privacy policies . Fu, X. s.l. : Springer, 2011. Web Services and Formal Methods. pp. pp. 86-100.

[17] Mandatory enforcement of privacy policies using trusted computing principles. Kargl, Frank, Florian Schaub, and Stefan Dietzel. 2010. AAAI Spring Symposium: Intelligent Information Privacy Management.

[18] Niu, Chun Cheng, et al. "Security and Privacy Issues of the Internet of Things."Applied Mechanics and Materials 416 (2013): 1429-1433.

[19] Bawany, Narmeen Shawoo, and Nazish Nouman. "A Step towards Better Understanding and Development of University Ontology in Education Domain." Research Journal of Recent Sciences, Volume 2, Issue (10), 57-60(2013) ISSN 2277: 2502.

[20] Bawany, Narmeen Zakaria, and Jawwad A. Shamsi. "Smart City Architecture: Vision and Challenges." International Journal of Advanced Computer Science & Applications 1.6: 246-255.