Capturing the Provenance of Internet of Things Deployments*

David Corsar, Milan Markovic, and Peter Edwards

Computing Science, University of Aberdeen, Aberdeen, UK.
{dcorsar, m.markovic, p.edwards}@abdn.ac.uk

Abstract. This paper introduces the System Deployment Provenance Ontology and an associated set of provenance templates. These can be used to describe Internet of Things deployments.

Keywords: IoT deployments · Semantic Sensor Networks · Provenance

1 Introduction

There is growing recognition that increasing the transparency of Internet of Things (IoT) devices is key to fostering trust between citizens and the IoT [1, 2]. Within the TrustLens project[1] we are working with members of the public to identify what they want to know about IoT deployments. These end-user requirements are influencing the design of an ontological framework that is being used to represent this information, and make it available for use by tools that enable citizens to pose transparency questions of future IoT deployments.

During our initial discussions, users have highlighted a desire to know what IoT devices are doing, the types of sensors that are part of a device, how accurate the sensors are, and what data are being generated. Information such as this can be described by the Semantic Sensor Network Ontology (SSNO) W3C recommendation [4], which provides formalisms to describe sensors and related concepts in domains such as the Internet of Things. SSNO describes *ssn:Systems*[2], as pieces of infrastructure which may be composed of subsystems; three types (subclasses) of *ssn:System* are defined: *sosa:Sensors*[3], *sosa:Actuator*, and *sosa:Sampler*. Systems implement *sosa:Procedures* that can be used to describe the system's intended operations (e.g. how a sensor will make an observation). SSNO also models system capabilities (e.g. accuracy, expected battery life), acts (e.g. making an observation), results of those acts (e.g. an observation value), features of interest (the subject of an act), and properties of a feature of interest that can be observed, sampled, or changed.

* The research described here is supported by the RCUK Digital Economy programme award made to the University of Aberdeen; award reference: EP/N028074/1.

[1] http://www.trustlens.org

[2] Defined by the SSN namespace http://www.w3.org/ns/ssn/, abbreviated to "ssn".

[3] Defined by the SOSA namespace http://www.w3.org/ns/sosa/, abbreviated to "sosa".

Our interactions with end-users also highlighted their interest in the activities that may have occurred before and during IoT deployments. For example, whether any community consultation was conducted, who designed the deployment, and if any assessment was made of potential privacy risks. While SSNO includes a *ssn:Deployment* class, it only describes the deployed system and the platform (such as a wall, shelf, etc.) that hosts it, along with any further details about the platform, such as its location. We argue that having additional information about the activities that influenced the deployment would greatly increase its transparency (as desired by users), and assist with interpreting any data generated by the deployed system. For example, knowledge of maintenance activities performed on a sensor may influence an individual's view of the quality of data it generates. This may involve considerations such as when it was last (re)calibrated, its specified accuracy and drift values [3], or whether the surfaces of an air quality sensor have been recently cleaned [5].

This paper presents the Semantic Sensor Network System Deployment Provenance Ontology (SDPO)[4]. SDPO extends PROV-O[5] with a vocabulary for describing deployments of IoT systems as a collection of PROV activities conducted before or during a deployment, the associated agents and entities (e.g. systems, sensors), that have shaped the deployment in some way.

2 Describing IoT System Deployments

SDPO defines the *sdpo:DeploymentRelatedActivity* class, and an initial hierarchy of subclasses representing various types of such activities. SDPO also asserts that: *ssn:Deployment* is a subclass of *prov:Activity*[6]; *ssn:System* is a subclass of *prov:Collection* and *prov:Agent*; and *ssn:hasSubSystem* is a subproperty of *prov:hadMember*. These subsumptions are based on earlier work [3] aligning the non-normative SSNO published by the W3C SSN-XG with PROV-O.

The types of activities defined by SDPO include those that may have been conducted before deployment, such as *sdpo:SystemSelection*, *sdpo:SiteInspection*, *sdpo:DeploymentDesign*, and *sdpo:Installation*. Various types of maintenance operations are also defined, based on [5]. These include *sdpo:Calibration*, *sdpo:Cleaning*, and *sdpo:Replacement* of a system or subsystem. To support developers use SDPO to describe IoT deployments, a set of PROV-TEMPLATES[7] are available for the ProvToolbox library[8] that provide suggested provenance patterns for several types of SDPO deployment related activities[9].

Fig. 1 illustrates an instantiation of the PROV-TEMPLATE designed to capture the replacement of a subsystem. In this example, the sensor *:electricSensor1*

[4] Namespace `http://www.w3id.org/sdpo/`, abbreviated to "sdpo".
[5] Namespace `http://www.w3.org/ns/prov#`, abbreviated to "prov".
[6] SDPO views a *ssn:Deployment* as an activity during which, for example, a sensor performs the act(s) of making one (or more) observations.
[7] `https://provenance.ecs.soton.ac.uk/prov-template/`
[8] `https://lucmoreau.github.io/ProvToolbox/`
[9] The templates are available at `http://www.github.com/TrustLens/sdpo`.

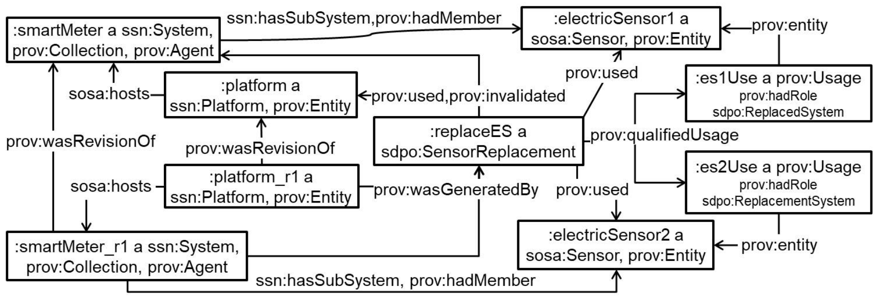

Fig. 1. Provenance record describing the replacement of a sensor in a smart meter.

that monitors electricity consumption as part of a smart meter (*:smartMeter*) is replaced during the *:replaceES* activity. The roles *sdpo:ReplacementSystem* and *sdpo:ReplacedSystem* are used to differentiate the function of the two sensors in *:replaceES*. As the sensors are part of the *:smartMeter* system (described in PROV terms as a collection of system entities), a revision of the system is created to reflect the change in collection membership. Consequently, a revision of the platform (*:platform*) hosting the system must be created to reflect that it now hosts *:smartMeter_r1*. As *:smartMeter* and *:platform* cease to exist following the replacement, they are invalidated by *:replaceES*[10]. While not illustrated in Fig. 1, it is expected that *:replaceES* would link to the activity (e.g., a *sdpo:DataReview*) which identified that *:electricSensor1* should be replaced, or to the entity that triggered the replacement (e.g, a fault report).

While SSNO can link an *ssn:System* with the *sosa:Procedure*s (plans) it implements, SSNO does not define how those procedures should be described. Fig. 2 illustrates the use of P-PLAN[11] to describe a *sosa:Procedure* as a *p-plan:Plan* that a system will enact during a deployment. P-PLAN describes plans as a series of *p-plan:Step*s that can be linked by *p-plan:Variable*s. In Fig. 2 the first step of the plan (*:plan*) is an *sdpo:Observe* step (a subclass of *p-plan:Step*[12]), which has an output variable *:electricityReading*. This variable is input to the *:uploadReading* step, which will send the reading to the energy supplier. An enactment of this plan is captured in the *:obs1-enactment* provenance bundle, which also illustrates the correspondences between the P-PLAN and SSNO concepts. This information can be used to improve the transparency of the expected behaviour of a device (as described by the implemented plan(s)), and contextualise the device's actual behaviour (as described in the retrospective provenance describing plan enactments). Note, as plans are specific to deployments of individual systems, we do not currently define a set of PROV-TEMPLATES for plans.

[10] Note, *:electricSensor1* is not invalidated, as it may subsequently be used by a repair or recycling activity.
[11] Namespace `http://purl.org/net/p-plan#`, abbreviated to "p-plan".
[12] SDPO also defines the steps *sdpo:Sample* and *sdpo:Actuate* corresponding to the *sosa:Sampling* and *sosa:Actuation* acts defined by SSNO.

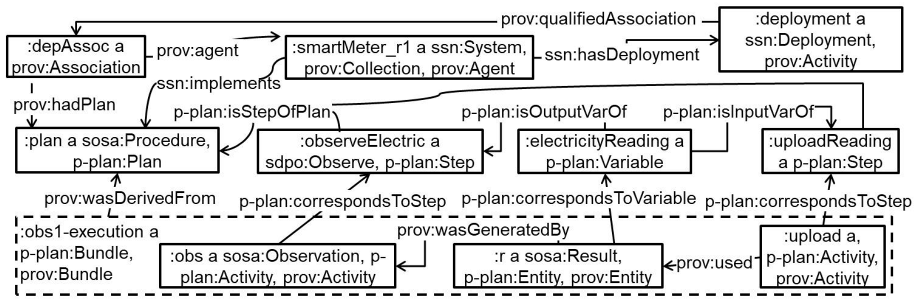

Fig. 2. Example plan for a smart meter to observe and upload the quantity of electricity consumed, and associated retrospective provenance generated during an enactment.

3 Future Work

The plan illustrated in Fig. 2 provides only partial transparency of *:smartMeter_r1*'s expected behaviour. For example, the plan as shown does not record that the smart meter will observe and upload energy usage every 30 minutes, as P-PLAN does not presently include the constructs necessary to model repeat processes. We are defining extensions to P-PLAN that will allow us to provide a more representative view of a device's expected behaviour. In addition to repeat processes, our intended extensions include associating constraints with variables. This will, for example, allow a plan to specify that the inputs to an upload step will be all of the readings made in the past 24 hours, rather than a single reading.

We are also continuing our user engagement activities, with plans to deploy several IoT devices in public spaces. We are currently developing a software framework that will capture ontological descriptions of these deployments and data generated by the devices. These will be used during co-design sessions involving members of the public, the outcomes of which will guide the development of software tools that allow citizens to explore details about IoT deployments, assist them in understanding the risks and benefits associated with IoT devices, and to assess the quality of the data produced.

References

1. Alliance for Internet of Things Innovation: AIOTI Strategy 2017-2021 (2017)
2. Almeida, V.A.F., Doneda, D., Monteiro, M.: Governance challenges for the internet of things. IEEE Internet Computing **19**(4), 56–59 (2015)
3. Compton, M., Corsar, D., Taylor, K.: Sensor data provenance: SSNO & PROV-O together at last. In: Proc. 7th Int. Workshop on Semantic Sensor Networks. vol. 1401. CEUR (2014)
4. Haller, A., Janowicz, K., Cox, S., Le Phyoc, D., Taylor, K., Lefrancois, M.: Semantic sensor network ontology. W3C Recommendation (March 2018)
5. Williams, R., V., K., Snyder, E., Kaufman, A., Dye, T., Rutter, A., Russell, A. abd Hafner, H.: Air Sensor Guidebook. U.S. EPA (2014), ePA/600/R-14/159