Zentrum für sichere Informationstechnologie – Austria
Secure Information Technology Center – Austria

A-1030 Wien, Seidlgasse 22 / 9
Tel.: (+43 1) 503 19 63–0
Fax: (+43 1) 503 19 63–66

A-8010 Graz, Inffeldgasse 16a
Tel.: (+43 316) 873-5514
Fax: (+43 316) 873-5520

http://www.a-sit.at
E-Mail: office@a-sit.at

DVR: 1035461
ZVR: 948166612
UID: ATU60778947

LOD AND LOV FOR AUTHORIZATION CONCEPTS

Project Report

Version 1.0, 03.7.2017

Bojan Suzic – bojan.suzic@a-sit.at

Abstract:

Linked Open Data (LOD) and Linked Open Vocabularies (LOV) deal with the definition of reusable concepts, models, and architectures that facilitate the integration of data and services on a web scale. With the ever-growing heterogeneity of available standards and implementation approaches, such integration faces various barriers that make this process costly and less effective in practice. Open and reusable vocabularies aim to lower these barriers by providing the foundations for conceptual annotations that allow the abstraction and bridging of concepts among different entities and domains in reusable, scalable and machine-readable manner. LOV, as one of the initiatives supporting the underlying paradigm, has emerged from the DataLift project and supported by Open Knowledge Foundation. LOV today represents the largest dataset that systematically gathers, analyzes and presents data about semantic vocabularies from different domains.

In this project, we examine the current state of vocabularies that support the annotation and integration of security-related concepts on the web. Based on our initial analysis, we establish a range of contributions that provide the underlying technical base for cross-domain application and exchange of data related to security processes. Our primary goal is to support authorization functionality, which allows the efficient permission management across different domains. We have structured and submitted our contribution to LOV portal for further indexing and expert scrutiny.

Contents

Figures

1. *Introduction*

The overall cloudification of services and an increasing amount of data that is stored in, delivered from, or exchanged between different systems deployed in the cloud requires rethinking existing paradigms of security and privacy management. The use-cases enclosing the traditional environment and architectural models assumed the user's or organizational data to be primarily stored and processed at a single entity, inside the premises of a particular organization. In contrast to that, current business models in a broader application, as well as their emerging derivations, assume the dynamic processing and reuse of data in multi-sectoral transactions. This expectation also implies the sharing, traversal and processing of data across different organizational entities, for various purposes and beyond its original or primary intention. Considering the complexity of such interactions, the unified management of data processing and its application in inter-organizational levels becomes a rising concern and challenge.

In this project, we approach the challenge of security management from the perspective that considers interoperability of controls on the semantic level. For this purpose, we rely on initiatives such as Linked Data Platform (LDP) [1] and Linked Open Vocabularies (LOV) [2], aiming to extend their reach and contribute with structured meta-models that are applicable in the field of information security.

In the following section of this report, we provide a brief overview of existing vocabularies that deal with security management and authorization. In the third section, we introduce our contribution, which is a framework that consists of three descriptive layers, whereas each of them is aimed to provide modular and reusable vocabularies that deal with the specific part of the problem. By following the minimalistic and modular approach, we aim to support reuse of these building blocks and their integration with the existing vocabularies in the way that is suitable for domain-specific applications, inducing a lower adoption and integration overhead. We then discuss our contribution and provide a brief conclusion.

The annex of this document provides the draft specifications which resulted this project.

2. Existing Vocabularies

This section presents the overview of prior work on defining open vocabularies for use in the areas of security management and authorization. The presented vocabularies have been provided at different levels of completeness and applied in limited use-cases.

2.1. W3C ACL System

World Wide Web Consortium (W3C) has implemented and described an overall approach [3] applied in their system for the purpose of access control. The goal of this work was to express access rules in logical, unambiguous and machine-accessible format, whereas the accessing client should be able to prove its access privileges, based on the use of different authentication mechanisms. The schema applied in this project is published on their web site as ACL Schema [4]. It consists of four classes and five properties. The initial concepts provided in the schema relate to an access rule `(ResourceAccessrule)`, which is constrained in the time interval and refers to a particular identity that is allowed an access of specific type to the resource. In addition to that, the ACL Schema allows the management of group memberships.

This system has been considered as an experimental setup applied to solve a single problem in the particular organization using a new approach. Apart from that, its adoption and application beyond this use-case remains undocumented.

2.2. Web Access Control

Web Access Control (WAC) is another approach that originated at W3C [5]. It considers users and groups as subjects that can be identified with HTTP URIs. This concept, which relies on dereferencing of URIs, allows the decentralized integration of various entities that reside outside of the originating system and employ diverse means for authentication. The vocabulary that serves as a basis for WAC is published as Basic Access Control Ontology [6].

WAC relies on WebID specifications [7] drafted by WebID Incubator Group at W3C. The overall work is further motivated by Tim Berners-Lee vision of socially aware cloud storage [8]. WebID provides a distributed identification mechanism that is extensible, builds on practices of web architecture and aims to seamlessly integrate with other protocols and standards including HTML and RDF. It furthermore relates to Linked Data Platform (LDP) [1], which is another initiative hosted at W3C with the aim to deliver the architecture for read-write linked data on the web. These two initiatives aim at synergic reuse of other protocols and their building blocks in order to achieve a higher degree of data and service integration at the web scale.

The notable motivations behind WebID include:

1) establishing the ability of the agents to control their identity
2) linking identities and relationships across web sites
3) agent-based view and establishment of trust relationships
4) enabling global authentication supporting strong privacy mechanisms
5) transforming and integrating identity management lifecycle to HTTP-based interaction

Similarly to ACL System, WAC is based on the concept of access control lists, which specify a range of operations that are applicable to a single resource by the referenced user. In the case of WAC, the predefined operations include `read`, `write` and `append`, complemented with `control`, which allows delegating the management of the access list to a third party.

Performed in the scope of different working groups that are backed by well-established standardization organization, WAC initiative is active since several years. The provided model, however, is too simplistic and restricted to the access control model that is historically proven but offers limited capacities to address the complexity and scale of internetworked world. The ACL-backed approach also does not provide support for dynamic adaptation and rich expressivity of controls, which are of significance for complex interactions and changing security requirements in modern web.

2.3. ReLL-S Ontology

A broader approach to semantic definition of authorization concepts has been proposed by Sepulveda et al. [9]. In their work, Sepulveda et al. first defined ReLL ontology for resource linking for machine-based clients of RESTful services. ReLL provides a declarative meta-model of RESTful services, including the mechanisms of URI generation, extraction, parsing and dynamic late binding. ReLL descriptions primarily allow descriptions of RESTful services using the concepts of `Resource`, `Representation`, `Links` that follow hypermedia constraints and `Requests` and `Responses` that are used to interact with services.

In the second iteration of their work, Sepulveda et al. enhanced their initial ontology with ReLL-S, an extension that conceptualizes security requirements and allows machine-based clients to interact with resources secured by the services using dynamically changing constraints. These constraints can describe any categories of confidentiality, integrity, authentication, and authorization. Based on these categories, ReLL-S further specializes by providing the descriptions for various concepts that implement each of these categories. In this sense, ReLL establishes a framework for specifying access constrains such as API keys, various HTTP-based authentication types and OpenID and OAuth protocols.

Although it follows a holistic approach that relates service descriptions with security-relevant specifications, the primary motivation behind ReLL and ReLL-S is the implementation of service composition based on predefined goal. From this perspective, the provided descriptions are read-only for external actors and serve only to describe what is expected or requested from the service which is necessary to automate service compositions. Hence, the users do not have the ability to manage or alter security requirements, nor to position them in the context that is broader than the single service description.

2.4. OWL Model for ABAC

In a more recent work, Sharma and Joshi [10] proposed an ontology for representing Attribute-based Access Control Policies (ABAC) in OWL [11]. Their ontology relies on the general ABAC$_\alpha$ framework proposed by Jin et al. [12] that is based on generic attributes that allow capturing concepts present in broadly deployed DAC, MAC, and RBAC access control models.

The concepts introduced in this work propose separate classes for each access control model covered under ABAC$_\alpha$. These include `RequestedAction`, `User`, `Object` and `Permission` for general model, and then `Clearance`, `Classification`, `uRole` for other subsumed models.

Their initial findings demonstrate the potential of OWL to represent ABAC policies. Among that, they have employed EYE reasoner [13] to derive authorization decisions based on provided policies. However, their model can be considered as incomplete and experimental, as it does not provide dynamic separation of duties, exhibits a low level of expressivity and provides only a basic description of subjects that restricts the practical application of the model. Moreover, the described model considers involved entities from the perspective of the closed, single system, which does not grant the application of the model beyond a single organization.

3. Proposed Vocabularies

To provide a solution that addresses a broader range of scenarios, we propose a general framework that consist of several domain-specific and use-case-specific vocabularies.

We establish this framework in three layers, as shown on Figure 1, and describe them as follows:

Service layer grounds a service description for the purpose of authorization management. It can potentially build on or relate to other service description models that may serve different goals. The service layer specifies the information model and the behavioral model of a service. The information model includes the description of data structures and their interrelationships, such as the hierarchical organization of resources present in RESTful API, or their interdependencies. The behavioral model specifies the applicable intents (actions), their expected parameters and the way they affect a service or its exposed entities.

Interaction layer provides concepts for the description of the contextual model and interactions that occur both in the internal and external environment. These interactions include different types of requests and responses that are used to retrieve resources or, for instance, check the authorization status of some transaction. The contextual model establishes an abstract level, aiming to cover both extrinsic and intrinsic parameters. The former describe particular states of the environment, system or external conditions, while the later represent the states that relate to resources or interactions in the reference system.

Authorization layer defines controls that are employed to manage authorization capabilities in a particular system or across the different systems. This layer currently includes representations for two authorization frameworks with different purposes. LEAR[1] is the access control model designed for RESTful APIs, while OAuth 2 is broadly used web authorization framework to manage data sharing across the services on the basis of resource owner consents.

Following the principle of modularity, each of presented layers can accommodate an arbitrary number of vocabularies that deliver meta-models for domain-specific or use-cases-specific applications. For instance, while DASP-Core intends to specify of RESTful APIs by considering intent-based and resource-based perspectives, some other vocabularies may be applied to describe interfaces based on WebSockets, WebRTC or messaging queues. Similarly, while DASP-interaction focuses on providing overall context descriptions and structure requests and responses for the purpose of authorization management and security enforcement, some other vocabularies may be introduced to extend existing or provide new concepts that go beyond these requirements.

[1] **L**ightweight and **E**xpressive **A**ccess Control Model for **R**ESTful Services

Finally, the authorization layer from the overall framework may be extended to provide specifications for other access control models, such as RBAC, ABAC or EBAC.

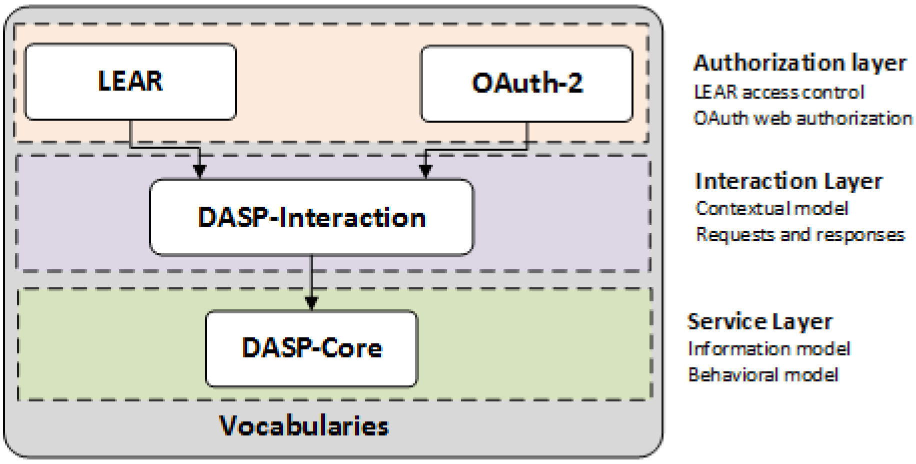

Figure 1: Three descriptive layers and their interdependences

3.1. LEAR Authorization Vocabulary

To evaluate the feasibility of our approach, we have developed a policy vocabulary that exposes concepts from LEAR, an access control model for Web APIs. The detailed overview of LEAR is available in [14]. The specification of this vocabulary is attached as an annex to this report. This section briefly introduces basic concepts of LEAR vocabulary.

The primary entity in LEAR vocabulary is the *security policy*, a minimal element of security evaluation expressed as a set of *security rules* that formulate a particular security goal. A rule establishes a relationship between *intended activity* and/or the exposed resource, positioned in the context that describes an arbitrary transactional, environmental or basic properties of the resource. We perceive a *context* as the state that further specializes in *intrinsic* and *extrinsic* conditions. Extrinsic conditions relate to contextual parameters that are external to the system, such as accessing client's properties, credentials, system time, or evaluation of an external function. *Intrinsic* conditions relate to the request, action or the product of the interaction. Using intrinsic conditions, we can dynamically reduce the scope of the context to a particular value present in the request parameter or the value of a retrievable element provided in the response.

Each security rule specifies an *effect* resolved in the case of successful rule match. The concept of an effect particularly encompasses different classes which can be categorized as *plain effects*, such as the ones that imply straightforward decisions. *Complex effects* refer to the subclasses that extend plain effects with additional functionality. This includes a class of *transformational effects*, which extend the permit class with operations that have to be executed in the course of the interaction. Their functionality enables achieving an acceptable balance between security and utility by applying the operations such as parameter redaction over incoming requests or data masking over outgoing responses in Web API interactions.

Considering that security rules may provide different or contradictory effects, a security policy includes a reference to *decision function*, which is applied over a set of rule effects with the purpose to resolve conflicts and derive a final effect of the policy.

As shown in Figure 1*, LEAR vocabulary* directly references the concepts from *DASP-Interaction,* and indirectly depends on *DASP-Core.* These relationships are relevant for four phases of policy management: In the *policy specification* phase*,* it is used to fetch available concepts from the target service, such as actions or elements, which can be reused to specify policies. *Policy validation* compares both policies and service descriptions to find inconsistencies. In *policy evaluation,* the descriptions of targeted services are retrieved and traversed to dynamically generate *e*xtraction paths and operations necessary for retrieval and comparison of values referenced in the policies. Finally, policy enforcement relies on both descriptions to execute policy decisions and apply optional transformations over resources.

3.2. DASP-Interaction and OAuth

In the scope of this work, we additionally specified *DASP-Interaction*, *OAuth-2* and enhanced the specification of *DASP-Core* vocabulary, which was defined in the course of previous work. The detailed specifications of these vocabularies are provided in the annex.

DASP-Interaction specifies the concepts from three main categories: *Context*, *Request* and *Response*. The context is further specialized in *Extrinsic Context* and *Intrinsic Context*, with each of them defining further classes that describe entities and states depending on their position relative to the system or a transaction. Hence, intrinsic context specifies the states or conditions that occur in requests, such as their parameters, or the properties of resources and their parts[2] that are subject to interaction. In addition to that, intrinsic contexts are used to define restrictions that apply to the property. This capability is suitable to restrict the matching of conditions, which can then be reused by other entities for evaluation of security policies or other purposes.

OAuth-2 is a domain specific vocabulary that represents the concepts defined in this authorization framework. The current release of this vocabulary contains the classes and properties that allow describing entities such as *Authorization Server* or *Tokens* that are used in interaction to authorize the requests. Similarly as in the previous cases, these entities are reused for various purposes. In our validation cases we use OAuth-2 vocabulary to reference client tokens in security policies, effectively providing the integration of existing token-bearer authorization framework with a broader LEAR access control model.

3.3. Discussion

Following the emergence of semantic technologies, many ideas materialized that relied on the application of semantics in existing use cases or in the generation of new business models. Although the use of these technologies in many cases can still be considered as experimental or immature, many advanced solutions rely on or in some way integrate semantic technologies. The initiatives such as Core Vocabularies [15] or Schema.org[3] demonstrate the perception of the relevance of these technologies in establishing cross-organizational interoperability on the semantic level both across public and commercial organizations.

One of the barriers to the adoption of semantic technologies is the diversity in approaches and vocabularies, as well as their non-existence (in some cases) or a lack of a broader use (following the networking effect). LOV initiative [2] partially aims to address the challenge of automated cataloging, classification, and retrieval of vocabularies, which are of significant relevance for their further reuse and interconnection of both vocabularies and available data sets.

In this work, we considered semantic technologies and particularly, the application of open vocabularies to establish inter-organizational semantic interoperability for the purpose of security management. In the first part, we provided the brief overview of existing approaches in definition and application of vocabularies (ontologies) that relate to the area of authorization and access control. Although the experimental setups were initiated and tested in different scenarios, the solutions such as W3C ACL System or Web Access Control [6, 3] provided only the restricted descriptive power and limited capabilities for their practical application. Their primary drawbacks are the focus on a simple access control model (access control lists) as well as simple use cases that consider specific setup at the particular organization. These approaches can be rather described as the application of semantic technologies to particular use case than the use of semantic technologies to address the broader problem of interoperability and management and provide a novel and innovative approach. The similar can be stated for OWL-ABAC [10], as well.

[2] Relates to previously exposed service information and behavioral models using DASP-Core

[3] http://www.schema.org

On the other hand, the other proposals, such as RELL-S [9], provided a richer level of expressivity and broader support for diverse security-related mechanisms, but again, considered in restricted application scenario, RELL-S does not address the management of security functions as it provides only read-only descriptions of exposed service capabilities and requirements.

In this work, we introduced a range of vocabularies that aim to address the gaps in existing approaches. In the first line, we acknowledged the lack[4] of public vocabularies that deal with the topics of security, access control and authorization, and specified and published the vocabularies that allow describing RESTful API services and security policies in a holistic manner. The proposed framework is built as a minimalistic and modular solution, promoting simple reuse and extensibility to support other authorization models and application scenarios. In current scenario, we allow for integration of information and behavioral models of Web APIs and their integration with LEAR access control policies and OAuth 2 framework, which are already broadly adopted on the internet.

4. Conclusion

In this project, we examined the current state of vocabularies that support the annotation and integration of security-related concepts on the web. Following the initial analysis, we specified a range of vocabularies that provide the underlying technical base for cross-domain application and exchange of data related to security processes. Our primary goal is to support authorization functionality, which allows the efficient permission management across different domains. We address this challenge from the holistic perspective by providing an abstract model for the unified description of informational and behavioral models of diverse Web APIs, which we then integrate with security policy management and existing web authorization framework to achieve a broader applicability within existing and emerging scenarios. We have structured and submitted our contribution to LOV portal for further indexing and expert scrutiny.

5. Annex

The following documents are attached in the annex of this report:

1. DASP-Core Vocabulary, Release 30 Juni 2017
2. DASP-Interaction Vocabulary, Release 30 Juni 2017
3. LEAR Authorization Vocabulary, Release 30 Juni 2017
4. OAuth 2 Vocabulary, Release 30 Juni 2017

References

[1] S. Speicher, J. Arwe and A. Malhotra, "Linked Data Platform 1.0," W3C, 2014.

[2] V. Pierre-Yves and G. A. Atemezing, "Linked Open Vocabularies (LOV): a gateway to reusable semantic vocabularies on the Web," *Semantic Web,* pp. 437--452, 2017.

[3] W3C, "W3C ACL System," 2004. [Online]. Available: https://www.w3.org/2001/04/20-ACLs.

[4] W3C, "A namespace for describing Access Control Lists," 2001. [Online]. Available:

[4] To the best of our knowledge

https://www.w3.org/2001/04/ACLS/Schema.rdf.

[5] W3C, "WebAccessControl," 2016. [Online]. Available: https://www.w3.org/wiki/WebAccessControl.

[6] W3C, "Basic Access Control Ontology," 2012. [Online]. Available: https://www.w3.org/ns/auth/acl.

[7] H. Story, "WebID Specifications," W3C, 2013.

[8] T. Berners-Lee, "Socially Aware Cloud Storage," 2011.

[9] C. Sepulveda, R. Alarcon and J. Bellido, "QoS aware descriptions for RESTful service composition: security domain," in *World Wide Web*, 2015.

[10] N. K. Sharma and A. Joshi, "Representing attribute based access control policies in OWL," in *2016 IEEE Tenth International Conference on Semantic Computing (ICSC)*, 2016.

[11] W3C Owl Working Group, *OWL 2 Web Ontology Language Document Overview,* W3C, 2009.

[12] J. Xin, R. Krishnan and R. Sandhu, "A Unified Attribute-Based Access Control Model Covering DAC, MAC and RBAC," in *IFIP WG 11.3 Working Conference on Data and Applications Security and Privacy*, 2012.

[13] R. Verborgh and J. De Roo, "Drawing Conclusions from Linked Data on the Web - The EYE Reasoner," in *IEEE Software*, 2015.

[14] B. Suzic, "Integrated Authorization Management for RESTful Services," in *(in submission)*, 2017.

[15] PwC EU Services, e-Government Core Vocabularies Handbook, 2015.

[16] D. Hardt, *The OAuth 2.0 authorization framework.,* 2012.

[17] B. Suzic, *Multidimensional Security Policies,* Graz: Zentrum für sichere Informationstechnologie - Austria (A-SIT), 2016.

[18] B. Suzic, "User-centered Security Management of API-based Data Integration Workflows," *NOMS 2016 - 2016 IEEE/IFIP Network Operations and Management Symposium,* pp. 1233-1238, 2016.

[19] B. Suzic, "Securing integration of cloud services in cross-domain distributed environments," in *Proceedings of the 31st Annual ACM Symposium on Applied Computing*, 2016.

[20] R. Cyganiak, D. Wood and M. Lanthaler, "RDF 1.1 concepts and abstract syntax," W3C, 2014.

[21] I. Horrocks, P. Patel-Schneider and H. Boley, "SWRL: A semantic web rule language combining OWL and RuleML," W3C Member submission, 2004.