General Description Framework as a Key Enabler to Resource Integration and Use

Ken Laskey[1] klaskey@mitre.org
The MITRE Corporation

Integration of resources that enable integration and use of heterogeneous data and services requires description adequate to understand

- what the resource is, i.e. what functionality it provides or supports,
- what conditions of use govern the resource, e.g. policies governing access and use, necessary operating agreements,
- what metrics describe the nonfunctional aspects, e.g. speed, quality, provenance, and
- how to interact with the resource, both in terms of vocabularies supporting information exchange and accessibility to enable the exchange.

An "impedance mismatch" in any of these aspects of description, not just a mismatch at the data level, could make a resource unusable. For example, if a word processing service does not support bulleted lists, it would not be sufficient for creating this position paper.

This position paper will expand upon the requirements for effective description and introduces the General Description Framework (GDF) as a basis for creating adequate description of any class of resources.

Past description efforts

Description of resources has traditionally focused on the description and control of documentation to specify programming interfaces. Web Services continued this initial focus with service description that concentrated on the programming aspects needed to invoke the services. Version 1.1 of the Web Service Description Language (WSDL) [1] specifies "an XML format for describing network services as a set of endpoints operating on messages containing either document-oriented or procedure-oriented information." Version 2.0 of WSDL [2] expands the thinking on what needs to be described by noting it "enables one to separate the description of the abstract functionality offered by a service from concrete details of a service description such as 'how' and 'where' that functionality is offered." Still, the degree to which functionality can be elaborated is minimal, and while WSDL provides necessary support to developers, it is lacking from the perspective of consumers finding the components of their problem solutions.

The idea of using description to support discovery was addressed in the Universal Description Discovery & Integration (UDDI) [3] specification. A registry built on UDDI was to support "description and discovery of (1) businesses, organizations, and other Web services providers, (2) the Web services they make available, and (3) the technical interfaces which may be used to access those services." However, the focus for UDDI was not on the functionality being provided or the conditions of use but on the providing entity. Furthermore, an effective description using UDDI requires building technical models (tModels) to represent the particular concepts and constructs of description, and the UDDI registry owners rarely followed through on the necessary effort.

[1] The author's affiliation with The MITRE Corporation is provided for identification purposes only, and is not intended to convey or imply MITRE's concurrence with, or support for, the positions, opinions or viewpoints expressed by the author.

The basic concepts involved for description in a service oriented architecture (SOA) context have been explored by the OASIS Service Oriented Architecture Reference Model Technical Committee [4]. In October 2006, the committee completed work on the SOA Reference Model (SOA-RM) [5] in which description was identified as a key concept. Work has proceeded on a SOA Reference Architecture Foundation (SOA-RAF) [6] that includes modeling of the elements of description in general and service description in particular. The SOA-RAF makes use of experience gained in the development and application of numerous description approaches as a starting point for its modeling, and concentrates more attention on all the aspects of description listed above.

As work proceeded on the implementation of SOA-RAF models for service description, it became apparent that the artifacts associated with services were many and varied, and descriptions are also applicable to many classes of resources in addition to services. It was during this time that the idea of a generalized description framework became apparent, where the general framework would capture the ''essence'' of what is common and necessary for all descriptions (i.e., descriptions of all classes of resources), and that the description of each kind of resource would then be a specialization of the general framework. The description of SOA services became the first specialization of the General Description Framework and subsequent specializations under development or in use include description for content collections, software, organizations, standards, widgets, and projects.

Underlying assumptions regarding description

The General Description Framework has been developed to recognize and respond to challenges that impact description, especially as relates to identification and use of resources in a SOA ecosystem. The ''SOA ecosystem'' is composed both of the participants interacting with resources to "do their business", and the resources themselves, all of which must be adequately described for the ecosystem to function effectively.

''Resource'', in the current context, is broadly defined as any entity of interest that can be uniquely identified and can be associated with a party responsible for its creation, maintenance, visibility, and accessibility. Resources include SOA services, content collections, widgets, and any other relevant class of items. It is less important to decide whether something should be considered a resource and more important to describe whatever may provide value and facilitate its use.

The following are a set of assumptions about description that motivate the form, fit, and function of the GDF:

- There is no one “right” description for all time. Any template for describing any resource is a snapshot in time.
- There is no single “right” description to satisfy all stakeholders. There will be an overlap in what various stakeholders find important but each stakeholder group is likely to have unique interests. Segregating portions of description and allocating specific segments to specific stakeholder groups will result in duplication and unreasonable constraints on the use of available description.
- Different stakeholders may associate different values to a descriptive property, where the values are concurrently valid. For example, someone concerned with an OMB (Office of Management and Budget) form A300 will choose values for business functionality from the Federal Enterprise Architecture (FEA) Reference Models [7] whereas someone with expertise in the discipline associated with the functionality will find a precise technical term more useful.

- There is likely no unique use for any element of description. Different stakeholder groups will use descriptive information in different ways for purposes they find relevant. For example, the concept of zip codes was created by the Post Office to streamline mail delivery but is now a key for store locators for businesses, locale indicators for weather, “chip” identifiers for map segments.
- Providers and consumers of description must use the same vocabularies or have ways to mediate among vocabularies if relevant matches are to be accurately identified. Thus, it is necessary to unambiguously declare the semantics being used. This applies to both the properties used for description and the values assigned to the properties. If the values are numeric, it is necessary to unambiguously indicate the units of measure.
- Description elements and values should be reusable across classes of resources. For example, color can be an element of description for any resource where the idea of color applies, and the value red should be equally applicable across resource classes. Value definitions may be prescribed through well-documented taxonomies or other collections of terms and relationships.
- Description should be concise. Long text passages are difficult to check for exact matches and are difficult to keep synchronized across multiple descriptions if a change is introduced as part of some use. A URI used to indicate retrievable text is easy to compare and multiple uses of the same URI remain synchronized if the retrievable text changes.
- There will always be a legacy collection of descriptions based on previous versions of descriptive property sets and previous versions of description values. The more success we have in getting descriptions created today means more legacy to accommodate future change.

<u>The Framework</u>

The General Description Framework has been formulated as a set of well-defined semantics that should be applicable for description of any class of resources in a SOA ecosystem. Configuration of the GDF for a particular class is done through defining a set of properties (referred to as a Property Set) and any applicable constraints in assigning values to the constituent properties. In particular, the Service Description Framework is an initial configuration of the GDF in support of description and discovery of SOA services.

The key element of the description is the Property, and a description is comprised of a set of Properties that is considered sufficient to adequately describe instances of the resource class and distinguish one instance of the class from another. The framework applies to any Property Set, thus enabling a consistent description framework that can address the needs of multiple communities and should remain flexible as the description itself evolves.

The Property is composed of the following three parts:

1. *Property Identifier*: The Property Identifier is the commonly used name of the Property. It must have an associated definition and the definition may include guidance on assigning values to the Property. The Property Identifier should be chosen to be applicable across communities where a description with this Property may be used, but the associated guidance may contain or reference aspects that are community specific.
2. *Property Value*: The Property Value is the value or set of values assigned to a Property. For a given instance of a resource class, the set of Property Identifiers and associated Property Values constitutes the description of the instance. The Property Value should be concise and make use

of the Value Definition Source (the third Property part) or other referenced sources if additional detail is required. Considerations for assigning a Property Value are discussed below.

3. *Value Definition Source (VDS)*: The Value Definition Source identifies one or more references that define the semantics or other source information of the values assigned as Property Values. The VDS entry should be a URI that dereferences to a Web-accessible resource. The resource may range from a text file that elaborates any semantics inherent in the Property Value to an entry in a taxonomy of terms to a concept and related axioms in a formal ontology. It may also point to other rationale for assigning this value.

An example of a Property is shown in Figure 1.

Property Identifier: version Property Value: 3.7.4 (20100704) VDS: http://www.myscheme.org/definitions#version_number

Figure 1 GDF Property example

The Property Identifier ''version'' is assigned the single value 3.7.4 (20100704). The VDS identifies http://www.myscheme.org/definitions#version_number as a source of information that enables an unambiguous interpretation of the Property Value. The definition of the Property Identifier ''version'' is assumed to have been identified when ''version'' was chosen as part of a Property Set.

Conclusions

The full details of the specialization and use of the General Description Framework is beyond the scope of a position paper. However, the important points are that past experience has led to a wider understanding of the requirements for description, especially to support the use of distributed resources that may be governed by different domains and need to be accessed by communications that cross different ownership boundaries. The General Description Framework is current work that has shown success in addressing these requirements.

References

[1] Web Services Description Language (WSDL) 1.1, http://www.w3.org/TR/wsdl , W3C Note 15 March 2001.

[2] Web Services Description Working Group Published W3C Recommendations, http://www.w3.org/2002/ws/desc/ , Recommendations dated 26 June 2007.

[3] Universal Description Discovery & Integration (UDDI) Version 3.0.2, http://www.oasis-open.org/committees/uddi-spec/doc/spec/v3/uddi-v3.0.2-20041019.htm , 19 October 2004.

[4] OASIS SOA Reference Model Technical Committee, http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=soa-rm .

[5] Reference Model for Service Oriented Architecture v1.0, links to HTML, PDF, and Microsoft Word formats at http://www.oasis-open.org/specs/index.php#soa-rmv1.0 , 12 October 2006.

[6] OASIS Reference Architecture Foundation for Service Oriented Architecture 1.0, Committee Draft 2, http://docs.oasis-open.org/soa-rm/soa-ra/v1.0/soa-ra-cd-02.pdf , 14 October 2009.

[7] Federal Enterprise Architecture (FEA) Consolidated Reference Model Document Version 2.3, http://www.whitehouse.gov/sites/default/files/omb/assets/fea_docs/FEA_CRM_v23_Final_Oct_2007_Revised.pdf , October 2007.