Cost Modeling Data Standards for Model Integration

Jeff McDowell
jmcdowell@tecolote.com
Tecolote Research, Inc.
4950 Corporate Drive; Suite 140-O
Huntsville, AL 35805-6227

1. Introduction

Service-Oriented Architectures (SOA) promise enterprises endless advantages: increased code reuse, reduced integration expense, better security, and - the big payoff - greater business agility. To achieve these goals, US DoD is migrating to web-based SOA. SOA is independent of, and removes concerns about, the underlying hardware or operating system of the participating systems in a data exchange. It requires adherence to a set of standards for the description, invocation, and exchange of information between systems.

The increasing complexity of modeling and simulation tools, combined with the imperative to control acquisition costs motivates interest in SOA to automate cumbersome and often repetitive processes such as the exchange of data between the technical and cost estimating communities. As the modeling and simulation community moves towards automated discovery and composability of models, the clarity of the data descriptions will be of utmost importance, as will adhering to common standards.[1] Standards are integral to web services, and web services are key to a service-oriented architecture.[2] This need for common data standards motivated the Army's Office of Deputy Assistant Secretary of the Army - Cost & Economic (ODASA-CE) to sponsor the effort described in this paper.

2. Background

In most organizations, it makes sense to begin SOA policy-making efforts with standards[3]. By adapting cost estimating data structures to XML, this effort lays the groundwork to support SOA integration with other models and simulations as well as other acquisition processes that provide information to, or require information from, cost estimators.

The Department of Defense Net-Centric Data Strategy requires that all Information Technology systems operating within the Global Information Grid (GIG)[4] be "advertised" to enable their discovery and subsequent use by the widest possible audience. As expressed in the DOD Net-Centric Data Strategy, "advertising" means tagging information resources with metadata that illuminates their identity, nature, and content; key persons and organizations responsible for them; a variety of associated dates and product formats; and other information. It also means posting or publishing these advertisements in globally accessible registries, directories and catalogs[5].

The objective of this initial standards effort is to tag cost model content so it can be used and reused in different contexts. Tags may be thought of as simply descriptive labels. Collectively, the labels form a vocabulary. If a cost analysis tagging vocabulary exists which is well-understood, it is better to use this vocabulary for data exchange rather

---

[1] McDowell, Jeff and Schaefer, Tom; Integrating Cost Estimating and Design Using SOA; Simulation Interoperability Standards Organization's (SISO) Simulation Interoperability Workshop Fall 2006.
[2] Service Oriented Architecture Reference Guide; U.S. Army Enterprise Solutions Competency Center; PEO-EIS; Jan07.
[3] InfoWorld: Governing SOA; By Phillip J. Windley; January 19, 2006.
[4] The Global Information Grid (GIG) is the collective of all of DoD's personnel who are on-line at any given time, DOD communications and other enterprise infrastructure, and all warrior, intelligence, and business applications. DoD Data Asset Visibility; Defense Information Systems Agency by Net-Centric Enterprise Services Program Management Office; January 26, 2005.
[5] Department of Defense Net-Centric Data Strategy memorandum; DoD CIO; October 23, 2003.

than re-invent model-to-model exchange formats for each future application. In this context, the goal of the effort described in this paper is to tag the cost analysis Community of Interest (COI)[6] content.

3. Use of Data Standards in SOA

Figure 1 depicts a setting with a COI-implemented data standard. Near the top of the figure, the COI has defined its vocabularies and taxonomies; Its data assets have been made visible; The metadata is published; and its models and databases have been made available as services on the GIG. Consider now the construct from the viewpoint of a user, who in an SOA setting is a data consumer. The user does not go to each model developer to express their data needs who in turn would "push" the specified data to them. Rather the user takes on the burden of "pulling" their needed data from the shared space. The benefit to each modeler is they do not concern themselves with multiple user needs and they especially avoid the n-factorial problem of model data exchange.

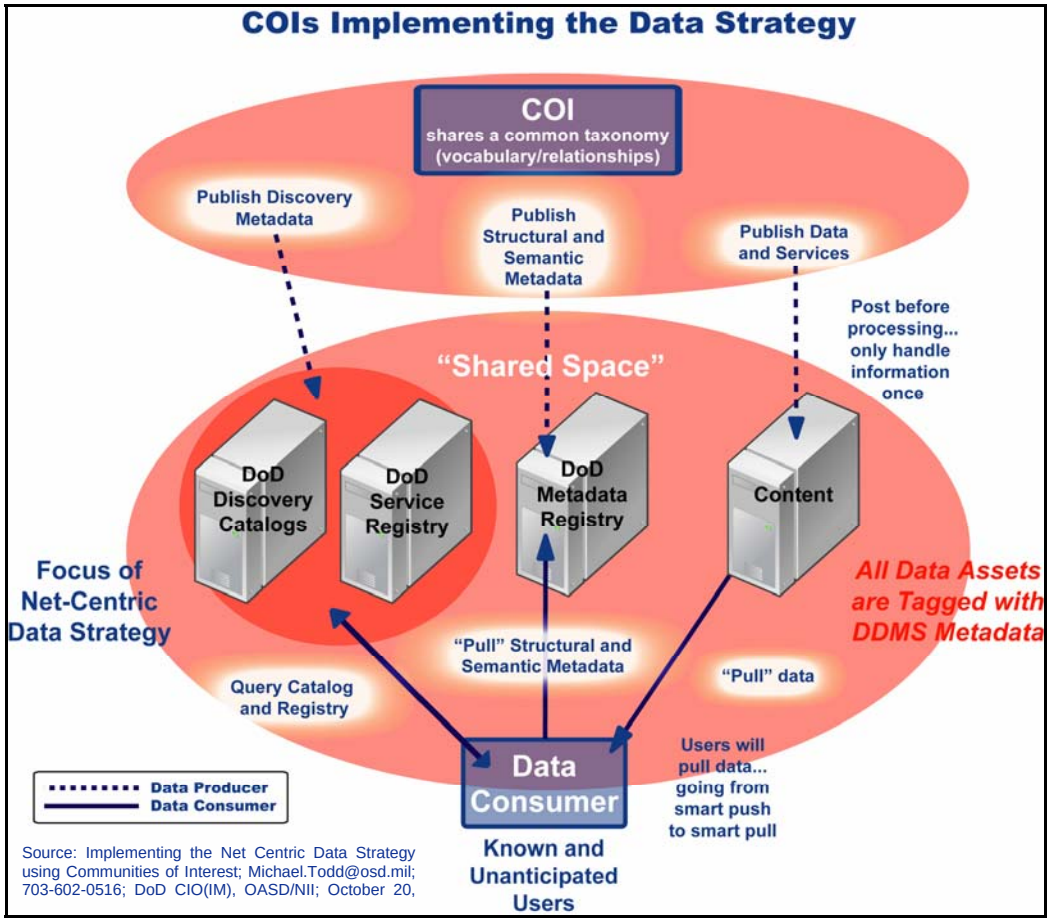

Figure 1. COI Data Strategy in an SOA Setting

[6] Communities of Interest (COI) are collaboration groups of users, who must exchange information in pursuit of their shared goals, interests, missions, or business processes, and who, therefore must have shared vocabulary for the information they exchange. DoD Data Asset Visibility; Defense Information Systems Agency by Net-Centric Enterprise Services Program Management Office; January 26, 2005.

Figure 2 illustrates several applications adapting to the data standard via adaptation layers. The adaptation layer is intended to be a modest effort sized in hours or days. The emphasis is that each application's internal data structure need not change.

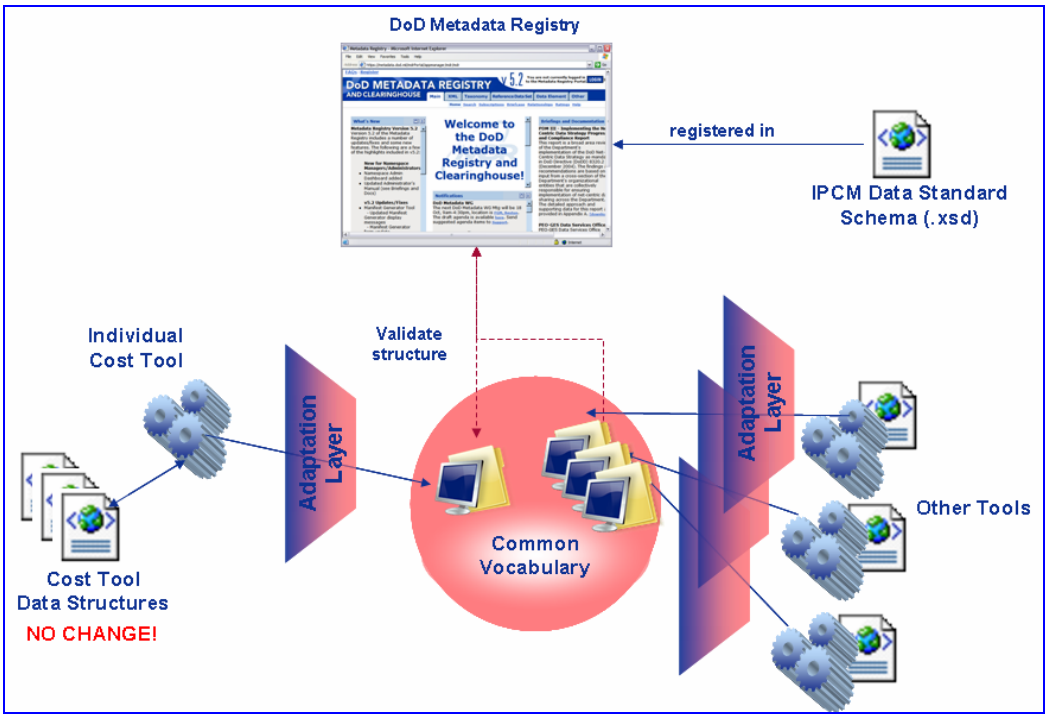

Figure 2. Models Using a Shared Vocabulary

Figure 3 illustrates the SOA from the standpoint of a data user conducting as analysis of alternatives. In this figure two COIs, Cost Analysis and Modeling & Simulation, are shown. In this construct, the user has not integrated each model with one another but has simply pulled and combined data from all the source models within the shared space.

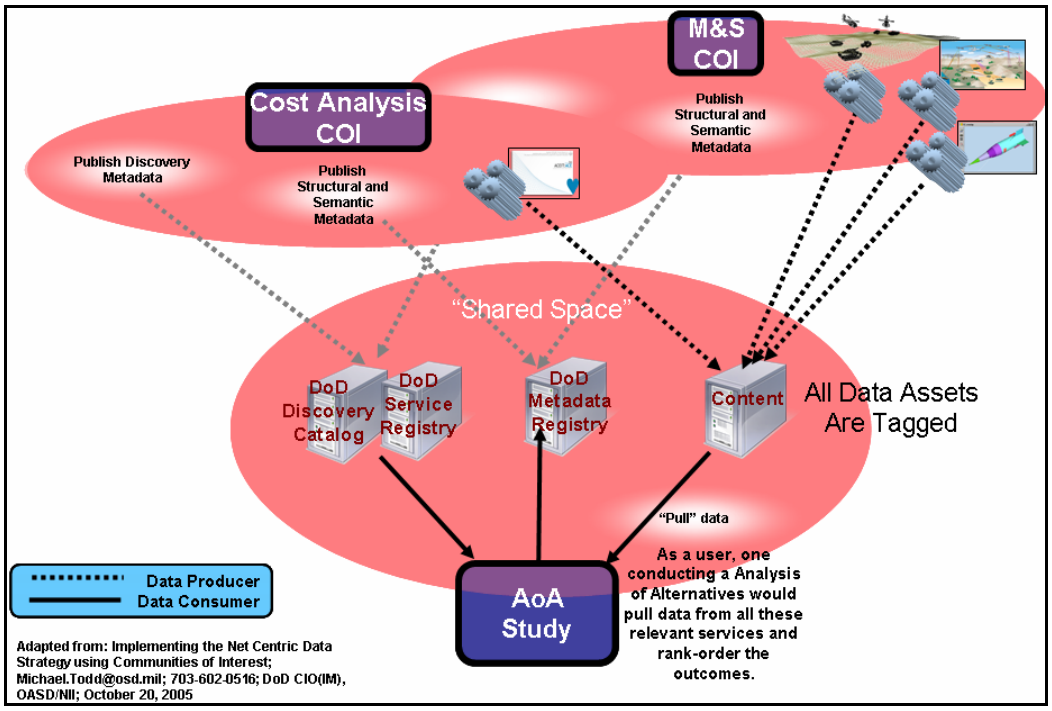

Figure 3. An AoA Study in an SOA

4. Standards Development

The sequence of analysis to develop the initial cost modeling standards is described in this section. The process began with a review of the DoD Metadata Gallery to identify existing schema from which, it was hoped, repeatable segments could be lifted or, at a minimum, learn from existing registered XML products. Second was to review the body of variables and data elements contained in existing Army cost analysis models, with special emphasis on the commodity cost performance models. The body of knowledge examined included provided several IPCM cost performance models, access to Force and Organization Cost Estimating System (FORCES), Army Military-Civilian Cost System (AMCOS), Operating and Support Management Information System (OSMIS), Logistics Cost Estimating Tool (LCET), and numerous Cost Analysis Requirements Descriptions (CARDs). All the parameters that were deemed cost drivers and cost model results were examined and tabulated.

These parameters were then categorized into a first-level set of four bins: Technical, Programmatic, Logistics, and Cost. Next, a taxonomy was devised that took the categories downward one additional level. Finally, the original list of parameters was mapped against the taxonomy to test the completeness of the taxonomy. These steps were iterated several times until a workable taxonomy was achieved.

The taxonomy was then used to layout the top level set of schema elements in XML. The development suite XMLSpy from Altova was used as the development platform. The Federal XML Naming and Design Rules[7] were followed in developing the .xsd schema files.

The terminal end of each branch of the schema tree specifies how the data will reside in future XML applications utilizing the schema. Continuous consideration was given into whether the values should be elements or attributes. The XML language is ambivalent on the matter (it permits either) and there is no community consensus on when to use

[7] Federal XML Naming And Design Rules 9 June 2005 Draft**;** XML Schema Interoperability Working Group**.**

attributes or elements. Attributes have certain virtues relating to lexical typing and convenience and elements have the primary virtue of being able to contain sub-structure.

For the purposes of this standard the following were used:

- Attributes for metadata about the parent element.
- Attributes for data that is semantically tied to the enclosing element.
- Elements for data items that have a meaning separate from the parent and sibling elements.
- Attributes when the value will be frequently present in order to improve the human readable form of an XML instance document and to make them easier to parse.

At the time of this writing, seventeen (17) separate schemas have been developed. Figure 4 shows the key ones with their primary elements. The CostObject specifies the information necessary to fully describe a value of cost (results) and the CostModelingObject specifies the information necessary to model an item's cost (computation). The CostStructureObject specifies an hierarchical structure (WBS) in which CostObjects and CostModelingObjects would reside. The ProgrammaticObject, ManpowerObject, and the LogisticsObject specify cost-driving information in the categories each of their names implies. The DefenseSystemObject is a generic object for describing systems which is tailored by system type such that all the cost drivers are specified. Examples of such schemas are MissileSystemObject, VehicleSystemObject, etc. The center of Figure 4 presents a set of attributes applicable to all of the objects. These are the metadata necessary to manage the data in future applications.

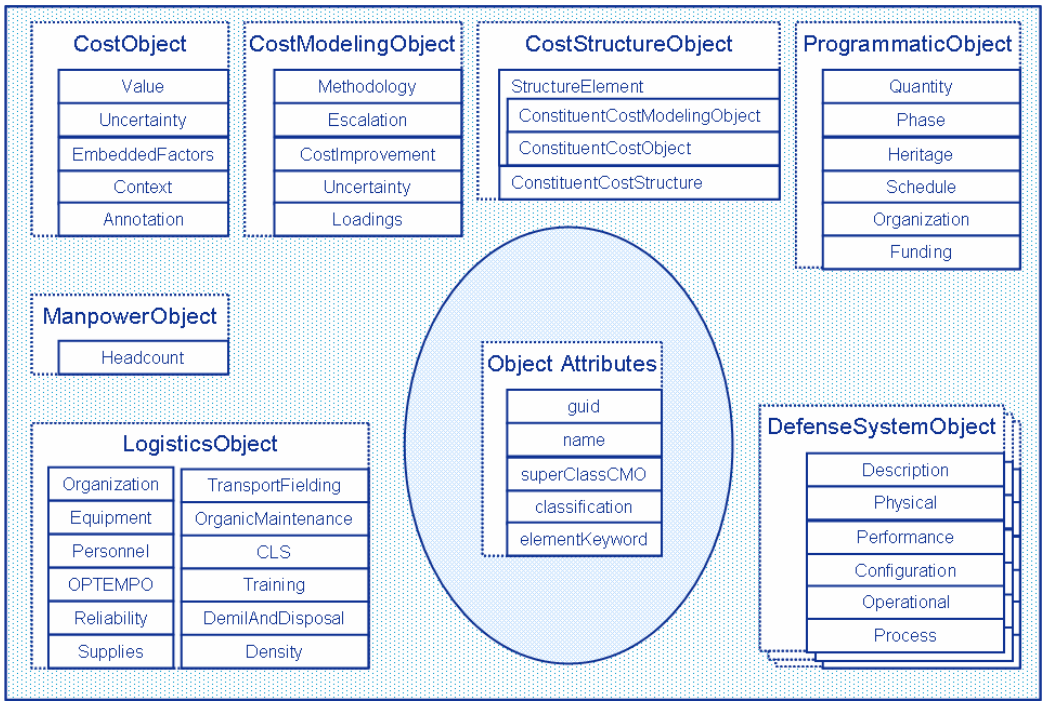

Figure 4. Data Schemas

Figures 5, 6, 7, and 8 present the schema tree for the CostObject, CostModelingObject, and the CostStructureObject. Each branch terminates with an element which, in turn, contains attributes. The attributes have the additional role of specifying data type such are integer, decimal, text, etc. For example the "Value" element in Figure 5 has five attributes. So the properly specified cost would comprised not only as its value (e.g. ten), but also its units (e.g. US Dollars), its scale (e.g. millions), its inflation basis (then-year), and its inflation year (e.g. 2007).

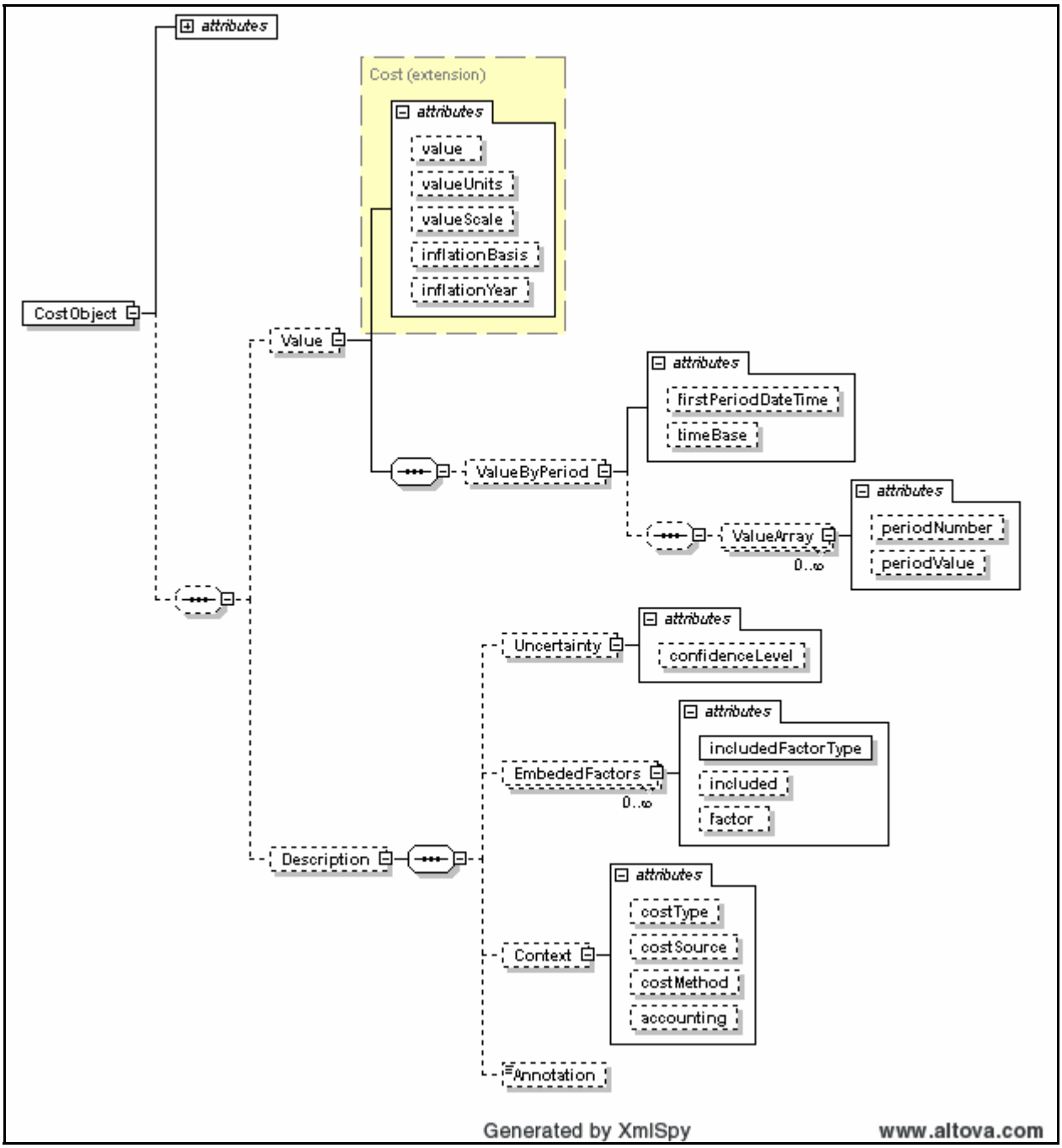

Figure 5. Cost Object

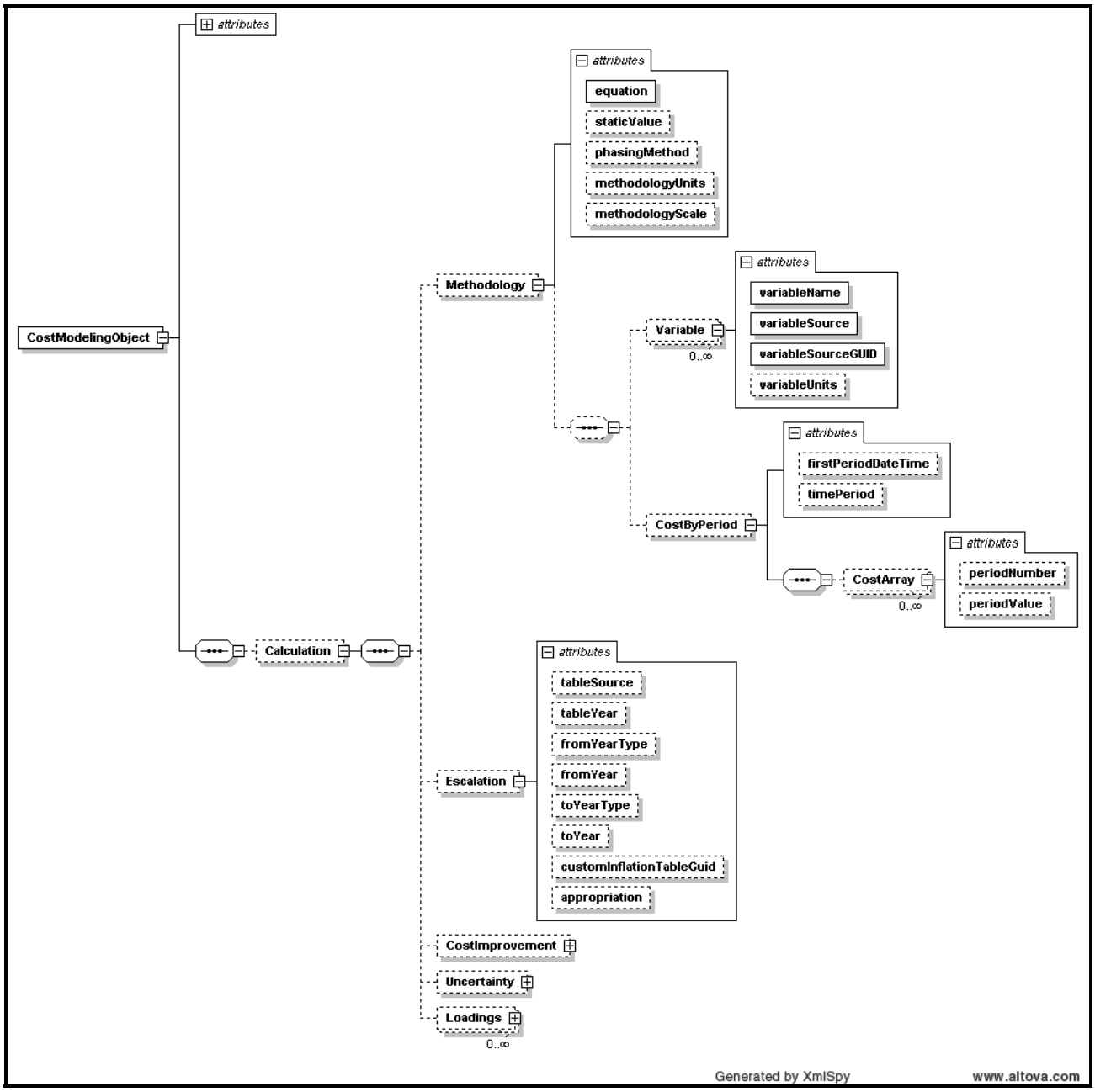

Figure 6. Cost Modeling Object (1 of 2)

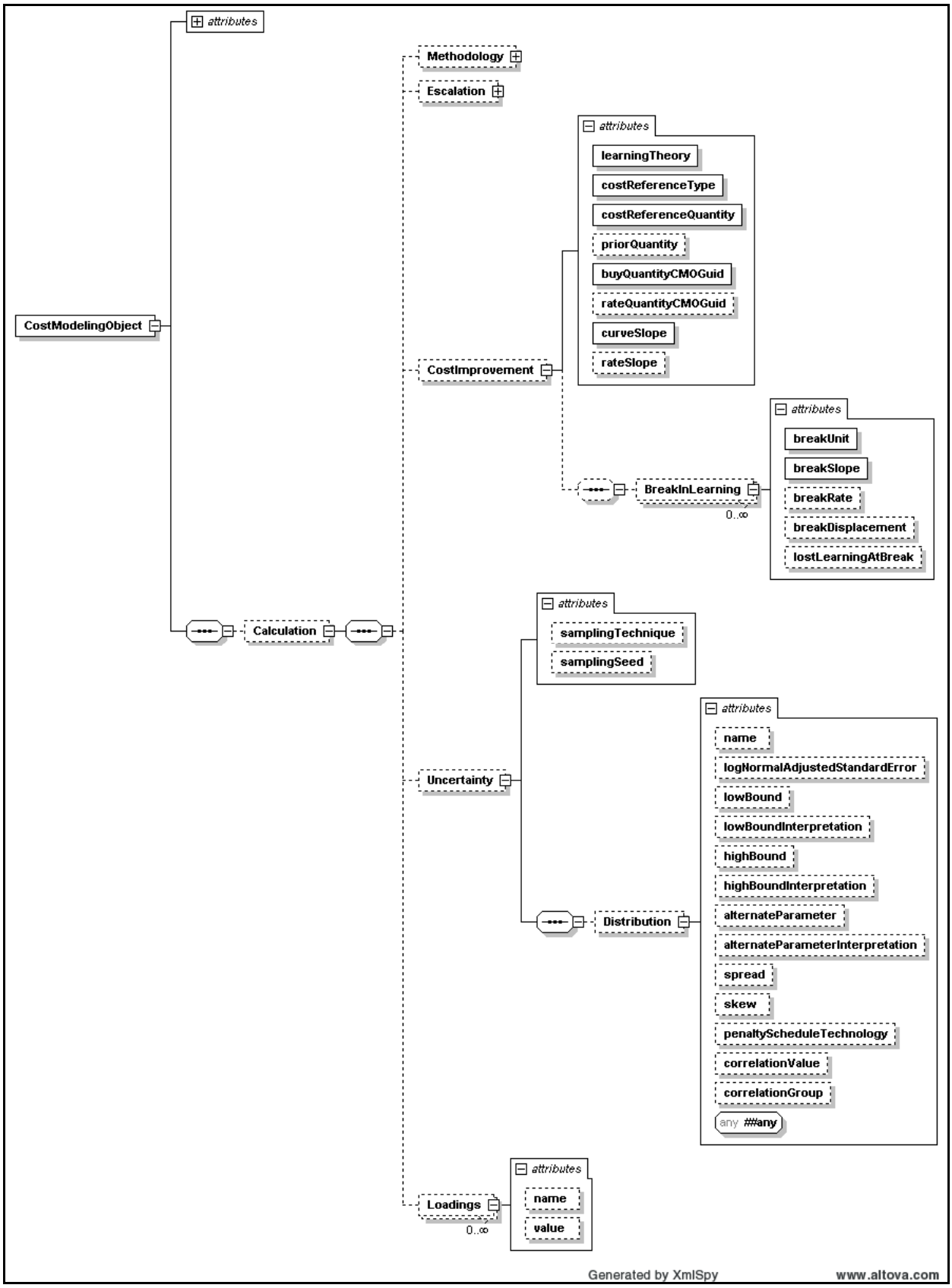

Figure 7. Cost Modeling Object (2 of 2)

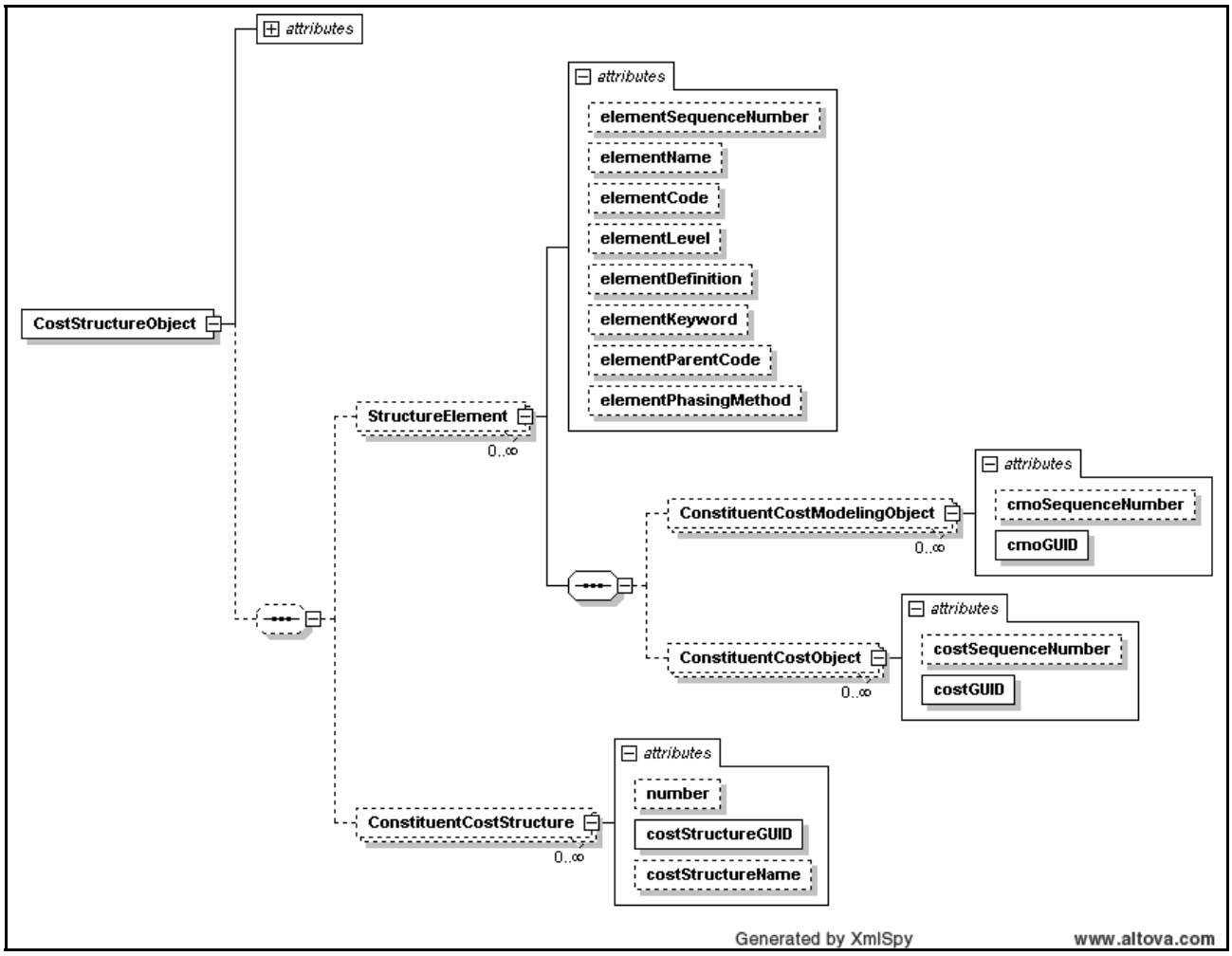

Figure 8. Cost Structure Object

5. Summary

The schema packages have been run through the DoD Metadata Registry validator and have been submitted into its Modeling and Simulation Namespace. By adapting the cost estimating data structures to XML, this effort has laid a strong foundation for enabling cost model utility in an SOA setting. The author encourages others in the cost and technical modeling and simulation communities to participate in a future formation of a Cost Analysis COI to further refine and promulgate cost analysis data standards.