Library Service-Oriented Architecture to Enhance Access to Science

Richard Akerman*

*National Research Council Canada (NRC), Canada Institute for Scientific and Technical Information (CISTI)
richard.akerman@nrc-cnrc.gc.ca

Abstract

In the digital world, library services need to be transformed to recognize that automation and machine-to-machine communication of information enables many advanced features. NRC CISTI, Canada's National Science Library, has addressed this challenge using Service-Oriented Architecture (SOA). Using SOA based on Enterprise Architecture, we have modeled services provided by the catalogue, services related to document delivery, and advanced services to provide more open data access. We will present a completed model of some library business functions, discuss the challenges of implementing an SOA within the library technology environment, and present our vision of how library SOA is an important enabler for wider, easier, and more powerful access to science.

Keywords: service-oriented architecture, enterprise architecture, software development.

1 Introduction

The library community has recognized the importance of identifying, standardizing, and communicating shared models of library services [1, 2]. CISTI has been pursuing an internal Enterprise Architecture initiative for over three years. The models created and resulting services identified have been used to guide system development, and are beginning to deliver on the promise of reduced duplication of development effort, as additional service offerings can be built on the foundations of existing services.

2 Enterprise Architecture and Service-Oriented Architecture

Enterprise Architecture is a methodology for describing an enterprise's IT systems and processes, and for planning for changes in these systems. Unfortunately, it often meets with limited success, due to the overwhelming complexity of IT systems [3].

In order to address this challenge, CISTI chose a methodology from InfoMajic, which focuses on Enterprise Architecture as the process of transforming business desires into defined implementation projects [4].

The importance of using a practical methodology and the challenges of developing and maintaining an Enterprise Architecture should not be underestimated. There may be many organizations that simply do not have the scale of software development needed to support this

effort. At each stage, the organization must concentrate on practical steps forward, rather than being distracted by complex technologies and approaches.

The use of this focused methodology has enabled CISTI to develop models and guide development successfully. However, it was recognized that the addition of Service-Oriented Architecture would further enhance the ability of the organization to implement an agile, sustainable architecture.

Service-Oriented Architecture is a methodology for identifying business functions that have certain desirable characteristics:

- autonomous
- loosely-coupled
- coarse-grained
- accessed through well-defined interfaces
- able to support reuse
- able to suppose composition

We will briefly explain each of these essential elements in the following sub-sections.

2.1 Autonomous and Loosely-Coupled Services

Services should be autonomous, that is, they should be able to operate independently of other services. This loose-coupling of services is an essential part of the service-oriented design methodology. It helps to counteract the tendency of systems to be developed in a siloed manner, in which a closed system has many internal interdependencies and its internal components cannot be reused outside of the system.

2.2 Coarse-Grained Services

Services should encapsulate significant amounts of functionality, otherwise they will not deliver adequate capabilities in return for the effort required to reuse them.

2.3 Services with Well-Defined Interfaces

Essential to maintaining and sustaining the use of services is guaranteeing a stable, well-defined interface. By careful specifying and maintaining this interface, the scope of the service is controlled. As well, the inner workings (the implementation) of the service may change, while the interface remains the same.

2.4 Reuse and Composition

The ultimate goal of service development is to provide the capability to reuse the services, and to create combinations or compositions of different services in order to deliver novel functionality. Maintaining the previously-outlined service development elements will help to ensure that services can be reused and composed together.

3 Example: CISTI's Pay Per Article Offering

CISTI Pay Per Article (PPA) [5] is the organization's first public offering to have been developed entirely using a service-oriented approach. Enterprise Architecture, followed by service-oriented analysis, helped to identify the necessary business functions and data flows.

Functions such as "Find Document" and "Get Document", which often have quite a generic meaning in a library context, were for the purposes of Pay Per Article defined in terms of specific data flows, and support particular data definitions (MODS). Well-defined interfaces allow these services to be reused in future development, beyond the PPA context.

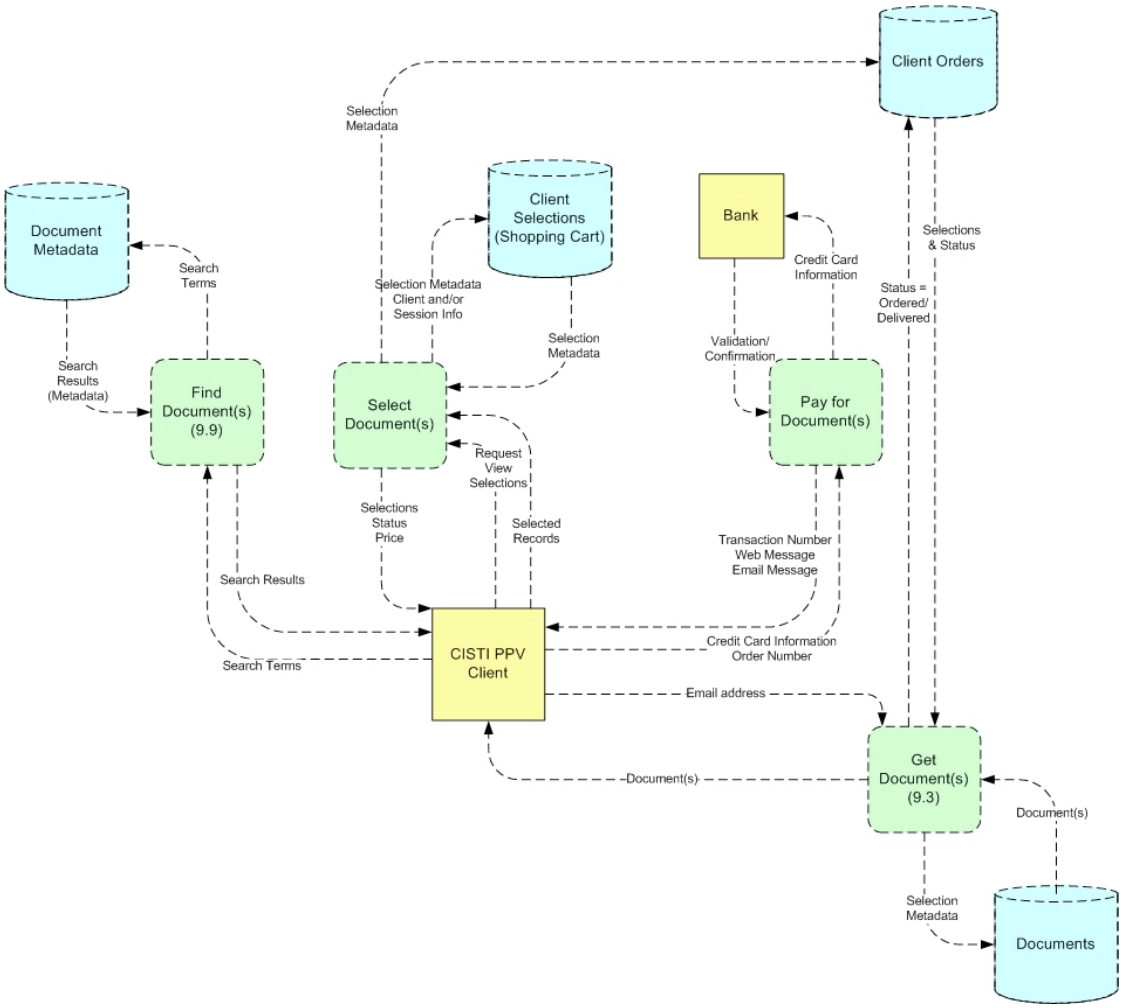

Figure 1. Simplified Pay Per Article Architecture.

This analysis flowed from the existing Enterprise Architecture and the identified business need for a credit card payment system. The simplified diagram of the PPA architecture (Figure 1) is one piece within the much larger context of the overall CISTI architecture.

This figure does not represent the program flow; it is not a sequence diagram. Instead, it portrays the data relationships between data stores (cylinders), business functions (rounded squares), and external interfaces (squares). These data relationships help to substantially define and constrain the service interfaces that are developed.

4 Library Integration with Scientific Workflows

One may wonder how library Web Services can help to enhance access to science. By providing standard machine-to-machine services for querying library holdings, and providing data and content over Web Services, the library can participate more fully in scientific workflows, particularly automated ones involving automated queries and analysis.

Currently tools such as Kepler [6] are enabling scientists to construct data and computation workflows. Libraries are often unable to expose information services that can be added to such compositions.

5 Discussion

This work demonstrates the feasibility of Enterprise Architecture and Service-Oriented Architecture in a library technology context. CISTI has therefore moved beyond basic frameworks and enterprise information technology theory, to demonstrated implementation.

Many large academic and research libraries struggle with legacy technology, including catalogues that were not originally designed to integrate with other systems (such as authentication and course management). There is increasing demand from patrons for more integration capabilities, for example, exposing data services to enable "mashups" and other Web 2.0 types of unified data views.

If this is done in an ad-hoc manner, there is a high risk that such implementations may not be sustainable, whether from an interface or a systems perspective. Additionally, if many different interfaces are developed at different libraries that provide similar capabilities, it will introduce a substantial development overhead when one attempts to develop cross-library applications.

Therefore, we recommend that libraries with substantial technical capability investigate and coordinate the use of architecture modeling and service standardization.

6 Future Work

CISTI is exploring the use of Web Services from copyright.com, and is also open to the collaborative development and use of library and scientific Web Services.

Substantial work is needed by the entire academic and research library community to come to agreement on a common framework for services and to begin implementing such services. Efforts such as the e-Framework for Education and Research [7] and the Digital Library Federation Service Framework for Digital Libraries [8] are stepping-stones on the path to true library integration with scientific workflows.

References

[1] B. Lavoie, G. Henry, L. Dempsey. "A Service Framework for Libraries", *D-Lib Magazine*, **12**, No. 7/8, (2006).

[2] "National Library of Australia: IT Architecture Report", http://www.nla.gov.au/dsp/documents/itag.pdf, (2007, accessed April 22, 2007).

[3] R. Sessions. “A Better Path to Enterprise Architectures”, Microsoft Technical Articles, http://msdn2.microsoft.com/en-us/library/aa479371.aspx, (accessed April 22, 2007).
[4] J. Carbone. *IT Architecture Toolkit*. Prentice-Hall PTR, (2004).
[5] http://ppa-ac.cisti-icist.nrc-cnrc.gc.ca/payperview/jsp/find.jsp?lang=en (accessed April 22, 2007).
[6] http://kepler-project.org/ (accessed April 22, 2007).
[7] http://www.e-framework.org/ (accessed April 22, 2007).
[8] http://www.diglib.org/architectures/serviceframe/ (April 22, 2007).