

IBM SPSS Collaboration and Deployment Services 5 web services overview

What are web services?

At a high level, a **web service** is a set of functionality that is distributed across a network (LAN or the Internet) using a common communication protocol. The web service serves as an intermediary between an application and its clients, providing both a standardized information structure and a standardized communication protocol for interaction between the two. Where other methods of distributed application architecture rely on a single programming language being used on both the application and its clients, a web service allows the use of loosely coupled services between non-homogenous platforms and languages. This provides a non-architecture-specific approach allowing, for example, Java services to communicate with C# clients, or vice-versa.

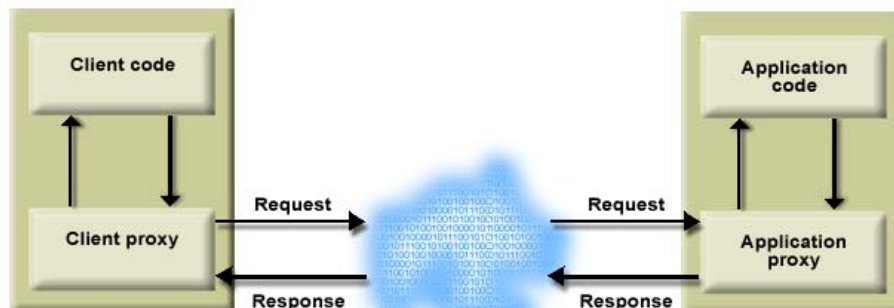
There are several advantages to implementing application functionality as web services, such as the following:

- Software written in different languages (Java or C#) running on different platforms (UNIX or Windows) can exchange services and data.
- Application functionality can be accessed by a variety of clients. For example, both a thin-client interface and a rich-client interface can take advantage of the web service operations.
- Updates to the service are immediately available to all service clients.

Web service system architecture

Web services are deployed and made publicly available using an application server, such as JBoss®, WebSphere®, or WebLogic®. The published web services are hosted by this application server to handle application requests, access permissions, and process load. A high-level architecture of how web services are implemented is displayed in the following diagram.

Figure 1-1
Web service architecture



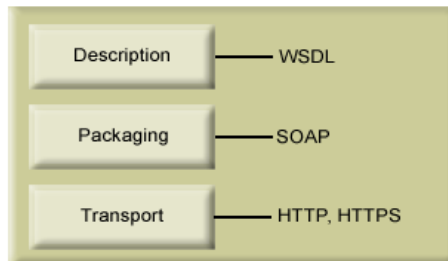
The client code supplies input to an operation offered by a proxy class. The proxy class generates a request containing a standardized representation of the input and sends it across the network to the application. A proxy class on the server receives the request and unmarshals the contents into objects for processing by the application. Upon completing the operation, the application supplies a proxy with the output. The proxy creates a standardized representation of that output and sends the response back to the client. The client proxy unmarshals the response into native objects for subsequent processing by the client code.

Standardizing the format of the information passing between the client and the application allows a client written in one programming language to communicate with an application written in another. The proxy classes, which are automatically generated from a web service description by a variety of toolkits, handle the translation between native programming objects and the standardized representation.

Web service protocol stack

A web service implementation depends on technologies often organized in a layered stack. The implementation itself defines a standard protocol for each technology layer, with each layer depending on the layers appearing below it in the stack.

Figure 1-2
Web service protocol stack



Beginning at the bottom of the stack, the Transport layer defines the technology standards for communication, allowing information to move across the network. HTTP or HTTPS are often used as the standard for the transport layer.

The Packaging layer rests on top of Transport and defines the standard for structuring information for transport across the network. The SOAP format is commonly used, which offers an XML structure for packaging the data. For more information, see the topic [SOAP: Simple Object Access Protocol](#) on p. 3 .

The topmost layer is Description and identifies the standards used by the layers below it in the stack, and provides the definition of the interface available for client use. The most common means of conveying this information is through the use of a WSDL file. For more information, see the topic [WSDL: Web Service Description Language](#) on p. 3 .

SOAP: Simple Object Access Protocol

The Simple Object Access Protocol (SOAP) is a way to pass information between applications in XML format. SOAP messages are transmitted from the sending application to the receiving application, typically over an HTTP session. The actual SOAP message is made up of the Envelope element, which contains a Body element and an optional Header element.

WSDL: Web Service Description Language

A Web Service Description Language (WSDL) file provides an XML-based map of what functionality the published web service allows, separating the implementation in the service from the interface. The WSDL defines the following:

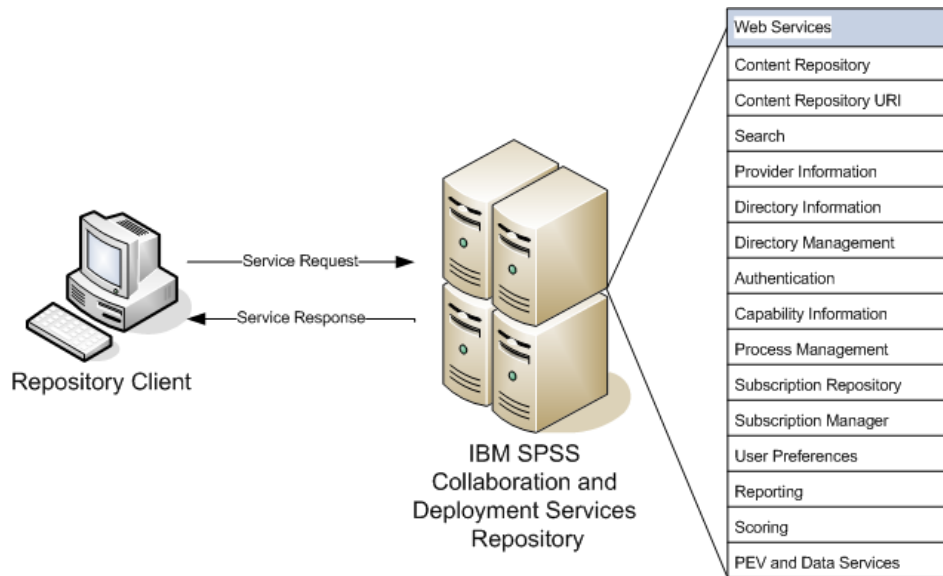
- The access location of the web service
- Operations the web service exposes
- Parameters the exposed operations accept
- Any request or response messages associated with the operations

The WSDL provides the information necessary to generate a client-side proxy in the desired programming language.

IBM SPSS Collaboration and Deployment Services web services

In IBM® SPSS® Collaboration and Deployment Services, the IBM® SPSS® Collaboration and Deployment Services Repository serves as a centralized location for storing analytical assets, such as models and data. Client applications access those assets using web services available on the repository server. The client sends a functional request as a SOAP message to a specific service, and receives a SOAP message response from the service in return.

Figure 1-3
IBM SPSS Collaboration and Deployment Services web services



IBM SPSS Collaboration and Deployment Services includes web services for item storage and retrieval, search, security, scheduled asset execution, and notification.

Content Repository Service

The IBM® SPSS® Collaboration and Deployment Services Repository provides the storage facilities for IBM® SPSS® Collaboration and Deployment Services. The repository stores objects in a hierarchical system similar to the folder/file structure used in operating systems. The objects themselves may exist in several different versions to accommodate and track changes to the object over time.

The Content Repository Service provides remote access to the repository for general storage and retrieval of content and meta-data. For example, a new object can be stored in the repository with particular name and keyword values. To retrieve the object, a query can be defined to search for objects having the specified name and keywords. Objects returned by the query can be retrieved for subsequent use. Retrieval of the object also returns the meta-data associated with it.

Content Repository URI Service

The Content Repository URI Service provides remote access to the repository for general storage and retrieval of content and meta-data. For example, a new object can be stored in the repository with particular name and keyword values. Retrieval of the object also returns the meta-data associated with it. The Content Repository URI Service offers functionality similar to that found in the Content Repository Service. The primary difference is the use of **uniform resource identifiers** to reference objects stored in the repository instead of identification specifiers.

Search Service

The Search Service provides a query mechanism for locating content in the repository that meets specified criteria. The query may be a global search for a specified string or a more structured search for information in specific fields. The information contained in the search result set can be customized to be as broad or focused as desired. In addition, large result sets can be returned as individual pages containing a specified number of hits to optimize client performance.

Provider Information Service

A user's login/password combination must be verified before access to IBM® SPSS® Collaboration and Deployment Services can be granted. The Provider Information Service provides information about the enabled security providers against which the user credentials can be authenticated. The data returned by the service is typically used to allow users to select the provider to use for authentication of their login information.

Directory Information Service

The Directory Information Service returns information about principals available to IBM® SPSS® Collaboration and Deployment Services.

Principals fall into one of three categories:

- A **user** is an individual who needs access to the system.
- A **group** is a set of users who need access to the system.
- A **role** is a set of one or more privileges, or actions. Roles are assigned to users or groups to manage system security.

Directory Management Service

The Directory Management Service allows control over user and group access to IBM® SPSS® Collaboration and Deployment Services. Specifically, the service offers the ability to perform the following tasks:

- Configure security providers for user authentication
- Create users and groups for the system, and edit their properties
- Assign users and groups to security roles that control access to system functionality

Authentication Service

The Authentication Service provides methods for users of client applications to connect to and disconnect from the IBM® SPSS® Collaboration and Deployment Services server. When a user attempts to log in to the system, the service provides access if the user credentials are valid. For a valid user, the service reports information about the user, such as actions the user can perform, as controlled by the authorization mechanism. In addition, the Authentication Service offers validated users the ability to modify their passwords.

Capability Information Service

The Capability Information Service allows a client to obtain detailed information about the IBM® SPSS® Collaboration and Deployment Services server. The information returned in the capabilities message includes the following:

- User details
- The service version and instance
- Actions the user can perform, as controlled by the authorization mechanism
- Web services available on the server
- Settings for the web services that can be configured by the user

The Capability Information Service is primarily designed to allow clients to collect a variety of information when first connecting to a IBM SPSS Collaboration and Deployment Services server. The capabilities information has ramifications for how a particular user interacts with the server. Consequently, a request for capabilities data should be included as part of the action of logging in to the server.

Process Management Service

The Process Management Service allows a client to manage the artifacts associated with running jobs. The client can submit an existing job for processing, and retrieve the execution meta-data and results. In addition, schedules can be created and assigned to jobs, allowing execution at a future date or on a recurring basis.

Subscription Repository Service

The Subscription Repository Service provides a general facility for sending event notifications to subscribers. Components supplying events that can trigger notifications use the service to register the events and their properties in IBM® SPSS® Collaboration and Deployment Services. Clients can query the IBM® SPSS® Collaboration and Deployment Services Repository to discover the properties that they can expect to be contained in events of a certain type, and can write well-formed subscription expressions to match events in which they are interested. Subscribers subscribe to the events of interest, having the notification messages sent to any defined destinations.

Subscription Manager Service

The Subscription Manager Service allows a client to manage notification plug-ins, which augment the standard services with additional functionality. For instance, plug-ins can generate e-mail distribution lists from database queries. The service also includes operations for message template management.

User Preferences Service

The User Preferences Service allows users of a client application to store and retrieve individual values for preference items defined in the system, permitting a customized experience for each user. For example, a user can specify his or her e-mail address and have it persist across sessions. In addition, the service includes administrative functionality for managing preference items, such as identifying which users have specified preference values.

Coordinator of Processes Service

The Coordinator of Processes Service provides remote interaction with the COP server. Servers use the web service to store configuration files and status information in the IBM® SPSS® Collaboration and Deployment Services Repository. Clients can retrieve a list of available servers and server clusters for subsequent connections.

Reporting Service

The Reporting Service allows a client to submit a predefined report for processing, such as a report created with BIRT Report Designer for IBM® SPSS®. Using information contained within the report, the client can validate input sources and create prompts for user input to direct report processing. Report output is available in a variety of formats for optimal display in any type of client.

Scoring Service

The Scoring Service is a document literal web service allowing client applications to employ real-time scores derived from predictive models developed in IBM® SPSS® Modeler, IBM® SPSS® Statistics, or third party tools. The service fetches a specified model, loads it, invokes the correct scoring implementation, and returns the result to the client. Supported models include regression (linear and logistic), decision trees, decision lists, neural networks, and naïve Bayes defined in SPSS Modeler streams or in PMML from SPSS Statistics.

Scoring can be either synchronous or asynchronous, depending on whether the client needs to wait for a score before proceeding or not. The service can load multiple models simultaneously for scoring and can be virtualized across multiple servers in a cluster configuration to handle large processing loads. The service logs all scoring activity for regulatory audit purposes. Configuring models for scoring and monitoring the service performance can be done using the IBM® SPSS® Collaboration and Deployment Services Deployment Manager.

The Scoring Service is often used in conjunction with the Data Services Service to access the data for scoring configurations. For more information, see Data Services Service documentation.

PevServices Service

The PevServices Service provides functionality used when working with the input data sources for analytic and scoring tasks. In general, the service provides the ability to perform the following tasks:

- Manage Enterprise View and Application View resource versions, allowing two versions of a view to be merged into a new composite version
- Retrieve metadata about the variables available in data sources
- Search for Application View and Data Provider Definition - Real Time resources that are compatible with specified criteria

The PevServices Service is often used in conjunction with the Scoring Service. Use the PevServices Service to discover data sources that are compatible with a particular scoring configuration. The metadata for variables in those data sources can be used to construct user interfaces that prompt for scoring input and to validate any supplied data values.

Data Services Service

The Data Services Service provides functionality used when working with the data sources defined in the IBM® SPSS® Collaboration and Deployment Services Repository for analytic and scoring tasks. In general, the service provides the ability to perform the following tasks:

- Retrieve metadata about the tables available in data sources
- Retrieve information about table columns and links

The Data Services Service is often used in conjunction with the Scoring Service. Use the Data Services Service to access information about the data used for a particular scoring configuration.

Single sign-on services

Single sign-on functionality for IBM® SPSS® Collaboration and Deployment Services clients is enabled by the following web services:

- **SSO Authentication Service.** Enables single sign-on access.
- **SSO Directory Management Service.** Enables management and configuration of IBM SPSS Collaboration and Deployment Services single sign-on.

The SSO Authentication Service provides methods for users of client applications to connect to a single sign-on-enabled IBM SPSS Collaboration and Deployment Services server by supplying the client single sign-on provider information and distributing Kerberos tokens.

The SSO Directory Management Service allows control over single sign-on configuration and user and group access to IBM SPSS Collaboration and Deployment Services. Specifically, the service offers the ability to perform the following tasks:

- Configure the single sign-on provider

- Create users and groups for the system, and edit their properties
- Assign users and groups to security roles that control access to system functionality