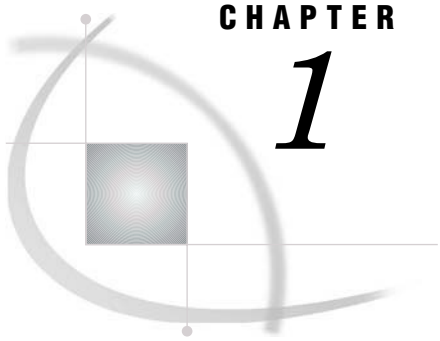


CHAPTER

1

Understanding Metadata Concepts

About This Guide

This guide provides a brief description of metadata, metadata management, and the benefits of using the SAS Open Metadata Architecture and the SAS Open Metadata Interface. It also shows you how to use the SAS Open Metadata Interface by walking you through the steps of identifying what information you need to capture, selecting the metadata types best suited for your information, and setting up a SAS Metadata Server and a metadata repository on Windows NT. In addition, it provides sample SAS Open Metadata Interface method calls for creating, updating, querying, and deleting metadata.

To get the most out of this guide, acquaint yourself with metadata concepts before reading Chapter 2, “Working with Metadata,” on page 5. Then follow the steps in Chapter 2, “Working with Metadata,” on page 5 to create your own sample repository.

Note: This guide does not describe how to control user access to metadata. For information about the SAS Open Metadata Architecture authorization facility, see the *SAS Metadata Server: Setup Guide and Administration Guide*. This guide is available at support.sas.com/rnd/eai/openmeta/. △

The setup guide also contains instructions for setting up a SAS Metadata Server in other host environments.

What Is Metadata?

Metadata is information about the data resources in an organization, or in simpler terms, data about data. Typically in the IT industry, we talk about "inventory data," "personnel data," "budget data," and "payroll data." The first word, a modifier, describes the data and classifies it as belonging to a certain business function. This is metadata. It tells us what the data *is*.

Metadata is also information about how the data is used. To understand this definition, consider the following example: "$100.00" is a piece of data. It could be payroll data, personnel data, inventory data, or budget data. Under the expanded definition,

- metadata is information that provides meaning and context to the piece of data. It tells us that "$100.00" is a monetary amount in U.S. dollars, expressed in terms of dollars and cents.
- metadata also tells us how to understand the way the data is expressed or represented. Metadata helps us to *understand* the data.

In any organization, there are two types of metadata:

Technical metadata describes the physical nature of the data, how the data was created, and how it is managed. This type of metadata is often machine-readable. Borrowing from the previous example, the fact that $100.00 is a monetary value and how it is expressed is physical data. Other examples of physical metadata might answer questions such as

- What is the origin of the data? Does it come from an external source, or is it generated internally?
- Where does the data reside? Is it in a SAS table or some other structure?
- On which server is the structure stored?

Informational metadata describes business rules and definitions on which the data is based. This type of metadata is often intended for people rather than machines. It is informational metadata that tells us whether the $100.00 value is payroll data, personnel data, inventory data, or budget data. Informational metadata would also answer questions such as

- Who is responsible for the accuracy of the data? How can I contact him or her?
- What business process produced this data? How do I execute the business process?
- Which applications should (and do) have access to this data?

What Is Metadata Management?

If metadata helps us to understand data, metadata management enables us to *use* the metadata. Metadata creation is time-consuming and expensive. To be truly useful, once stored, metadata must be centrally available and easy to maintain.

The primary goals of metadata management are

- to promote metadata conformity to enable sharing of metadata by an organization's applications. Metadata that is defined for one application can be copied and easily adapted for use by another application.
- to provide a common, centralized method of searching and managing distinct collections of metadata.

Both goals lower the costs of metadata development and maintenance by promoting standardization and reducing redundancy. Furthermore, when these goals are achieved, metadata can provide meaningful and valuable information, for example,

- impact analysis of technical changes within an organization
- comprehensive technical reporting about the organization's application systems.

Impact analysis gauges the effect of a single technical change on all of the applications in an organization. For example, if an organization stores metadata about its computer systems, it can use that metadata to easily determine which applications will be affected by taking a specific server offline. Or, if all applications store client

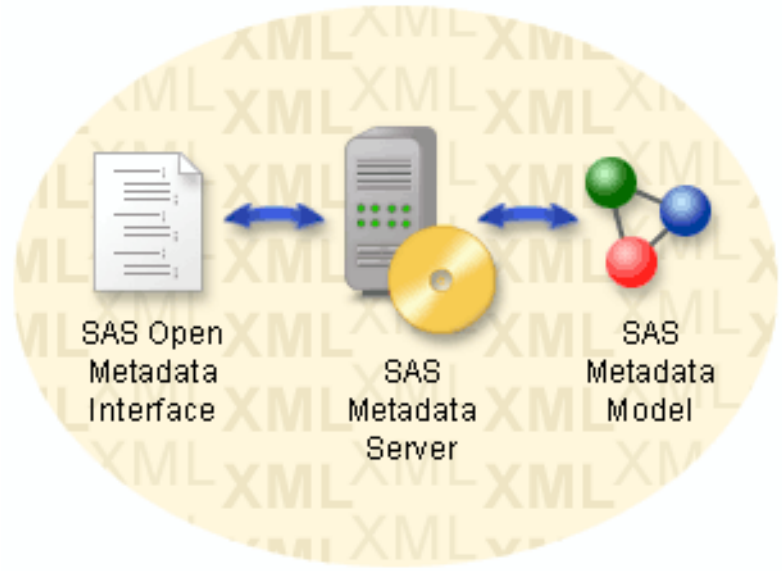
address information in an address object and a change is needed in the way this information is stored – for example, to surface street address, city/state, and country as three separate fields instead of one – the change is easy to identify, make, and propagate.

As electronic data transfers and e-commerce increase, the soundness of the metadata supporting these transactions will become as important as the data itself. Support for industry metadata models and data interchange standards enables organizations to respond quickly and economically to rapidly evolving, external reporting obligations.

What Is the SAS Open Metadata Architecture?

The SAS Open Metadata Architecture is a general-purpose metadata management facility that provides common metadata services to SAS applications. The metadata architecture provides

- *the SAS Metadata Server,* a central, shared location for storing metadata
- *the SAS Open Metadata Interface,* an application programming interface (API) that provides access to the server from a variety of programming environments, including Java, COM/DCOM, and SAS
- *the SAS Metadata Model,* a set of metadata types that are used for saving metadata on the server
- *an XML transport format and XML representation of metadata,* which makes it easy to transform the metadata to HTML and other standard XML representations, like the Object Management Group's Common Warehouse Metamodel (CWM).



Benefits of the SAS Open Metadata Architecture

- The SAS Metadata Model defines metadata types for the most commonly used objects and provides a mechanism for extending the metadata types with application-specific attributes and associations. This enables SAS applications to use a common model for most metadata objects while retaining the custom features that make them unique.
- Metadata objects are stored in application-specific repositories, which are managed by a repository manager. This tiered management approach enables metadata to be maintained separately yet accessed centrally through the repository manager, and guards the integrity of application-specific metadata while enabling global searching.
- A single tool set can be used to create, access, and manage metadata for all of your SAS applications.
- Support for the XML transport format and industry standard metadata models increases the likelihood of compatibility between SAS applications and other software applications.