Site Content and Metadata

Metadata is data about data, and is essentially the categorization of the content within a content management system. A good example of metadata is in the classification of documents in a Content Management System (CMS) or Document Management System (DMS)—a proposal document might have some essential attributes associated, aside from the actual content in the proposal document, such as Proposal Number, Client, Date, and so on. These attributes are the metadata associated with the document.

Why is metadata important? Simply, metadata allows rapid search of content in a system and grouping of content when browsing. Sophisticated search engines—such as the one included in SharePoint 2010—allow searching inside the content of documents, but sometimes a user wants to navigate quickly to a document without dealing with the search result noise typically associated with keyword searching. Metadata and categorization assist users to retrieve content without a lot of hunting and pecking into document or file content.

Metadata also assists a search engine. Search Engine Optimization (SEO) algorithms assign more relevance to data stored in metadata, because content owners typically provide metadata explicitly.

In the previous chapter, we looked at social networking, and you saw how taxonomy and folksonomy play an important part in categorization of content. Continuing from this theme, in this chapter we explore metadata in SharePoint 2010, the Content Type Model, and the Managed Metadata Service—the central hub that manages metadata in SharePoint 2010.

The SharePoint 2010 Content Type Model

If metadata is data about data, then a metadata model is a system that allows the creation, editing, and management of metadata. SharePoint 2010 has a metadata model—called the Content Type Model.

The SharePoint Content Type Model consists of metadata types—content types, which themselves consist of metadata fields—Columns, and the management of these types in the core of the SharePoint platform. I cover SharePoint metadata types in the next section of this chapter, but first I’ll present the SharePoint 2010 Content Type model at a high level.

What Are Content Types and Site Columns?

A SharePoint content type is a grouping of fields that describe a data entity in the SharePoint site collection or site—such as a document, calendar event, task, or any other list item. Each content type consists of one or multiple fields, which are known as columns.

Content types exist at either the root site collection or sub-site level. Lower sites in the hierarchy may use content types from parent sites to define data elements. In both cases, content types reside in the Content Type Gallery of the site or site collection.
The fields of content types exist as centrally managed *site columns*, at the same level as the containing content type or at parent levels to it. For example, a content type defined at a sub-site level may contain site columns that reside in the root site collection or parent site in the hierarchy. Sites may not leverage content types and site columns at lower sub-site levels to the current site—just up the hierarchy chain.

There is a difference between centrally managed site columns and list columns. When creating a custom list, the user may define columns specifically for the new list, meaning the columns bind only to that list. By contrast, site columns reside in the Site Column Gallery, and users may reuse these columns in lists (Add an Existing Column) and content types.

Content types may inherit from other content types at the same level in the site hierarchy or a parent level. Even the most basic content types, at least, inherit from the stock Item content type, which contains the Title site column. Through inheritance, users may customize content types by defining new types that have the same columns as the parent, with additional columns to complete the specialization. For example, if you want to customize task list items with new columns, rather than creating a new content type, you can inherit the existing task content type and add the new custom columns. Inheritance provides the added benefit that any changes to a parent content type apply to content subtypes.

---

**Note:** Never change stock content types and site columns; always inherit from content types, and then specialize.

---

**New Content Type Model Functionality**

SharePoint 2007 included lists, libraries, content types, site columns, and list columns. Just as you would expect, SharePoint 2010 includes them, too. This version brings new functionality, either as a direct enhancement to the Content Type Model or as an associated metadata or document management feature. Following are some of the more noteworthy additions to SharePoint 2010:

- Content types and lists that use lookup site columns may now specify additional columns of the lookup list to include, in addition to the primary lookup column.

- Referential integrity between lists for lookup columns and list items that reference the lookup parent list items. For example, when you delete a list item from a lookup list, SharePoint 2010 will either delete list items that look up the parent list item to the one just deleted, prevent the deletion of the parent, or do nothing, depending on a referential integrity setting.

- Document sets, a new content type and feature that allow collection of documents in one list item. They are ideal for collecting multiple documents that combines to a single finished entity, such as a proposal.

- A new Managed Metadata column type that maps to term sets in the Managed Metadata Service term store.

- A new Enterprise Keywords column, which allows users to add their own tags.

- Metadata extracted from the EXIF of images.