Spotlight Overview

Carbon > Design Guidelines

2007-05-27

Apple Inc.
© 2004, 2007 Apple Inc.
All rights reserved.

No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, mechanical, electronic, photocopying, recording, or otherwise, without prior written permission of Apple Inc., with the following exceptions: Any person is hereby authorized to store documentation on a single computer for personal use only and to print copies of documentation for personal use provided that the documentation contains Apple's copyright notice.

The Apple logo is a trademark of Apple Inc.

Use of the "keyboard" Apple logo (Option-Shift-K) for commercial purposes without the prior written consent of Apple may constitute trademark infringement and unfair competition in violation of federal and state laws.

No licenses, express or implied, are granted with respect to any of the technology described in this document. Apple retains all intellectual property rights associated with the technology described in this document. This document is intended to assist application developers to develop applications only for Apple-labeled computers.

Every effort has been made to ensure that the information in this document is accurate. Apple is not responsible for typographical errors.

Apple Inc.
1 Infinite Loop
Cupertino, CA 95014
408-996-1010

Apple, the Apple logo, Carbon, Cocoa, Mac, Mac OS, Objective-C, QuickTime, and Rosetta are trademarks of Apple Inc., registered in the United States and other countries.

Finder and Spotlight are trademarks of Apple Inc.

Simultaneously published in the United States and Canada.

Even though Apple has reviewed this document, APPLE MAKES NO WARRANTY OR REPRESENTATION, EITHER EXPRESS OR IMPLIED, WITH RESPECT TO THIS DOCUMENT, ITS QUALITY, ACCURACY, MERCHANTABILITY, OR FITNESS FOR A PARTICULAR PURPOSE. AS A RESULT, THIS DOCUMENT IS PROVIDED "AS IS," AND YOU, THE READER, ARE ASSUMING THE ENTIRE RISK AS TO ITS QUALITY AND ACCURACY.

IN NO EVENT WILL APPLE BE LIABLE FOR DIRECT, INDIRECT, SPECIAL, INCIDENTAL, OR CONSEQUENTIAL DAMAGES RESULTING FROM ANY DEFECT OR INACCURACY IN THIS DOCUMENT, even if advised of the possibility of such damages.

THE WARRANTY AND REMEDIES SET FORTH ABOVE ARE EXCLUSIVE AND IN LIEU OF ALL OTHERS, ORAL OR WRITTEN, EXPRESS OR IMPLIED. No Apple dealer, agent, or employee is authorized to make any modification, extension, or addition to this warranty.

Some states do not allow the exclusion or limitation of implied warranties or liability for incidental or consequential damages, so the above limitation or exclusion may not apply to you. This warranty gives you specific legal rights, and you may also have other rights which vary from state to state.

Contents

Introduction to Spotlight 7

Who Should Read This Document 7
Organization of This Document 7
See Also 7

What is Spotlight? 9

What does Spotlight provide? 9
Where Does Metadata Come From? 9
Searching Metadata in your Applications 10
Not Only About Search 10

How Does Spotlight Work? 11

Extracting Metadata from Files 11
Querying Spotlight for Metadata 11

Spotlight Metadata Attributes 13

Security and Privacy Considerations 15

Protecting Sensitive Data 15
Untrusted Applications and Bundled Importers 15

Spotlight and Document Bundles 17

Create an Invisible Document Bundle 17
Encapsulate Your Document with NSFileWrapper 17

Document Revision History 19

2007-05-27 | © 2004, 2007 Apple Inc. All Rights Reserved.

2007-05-27 | © 2004, 2007 Apple Inc. All Rights Reserved.

Figures

How Does Spotlight Work? 11

Figure 1 Extracting Metadata 11
Figure 2 Querying Spotlight 12

2007-05-27 | © 2004, 2007 Apple Inc. All Rights Reserved.

2007-05-27 | © 2004, 2007 Apple Inc. All Rights Reserved.

Introduction to Spotlight

Spotlight is a fast desktop search technology that allows users to organize and search for files based on metadata. Spotlight is extensible, allowing developers to provide metadata importers for their application's documents.

Who Should Read This Document

Spotlight is a fundamental feature of Mac OS X, and all developers should be familiar with its capabilities. Developers of applications that save documents to disk should consider providing Spotlight support by implementing a metadata importer.

Important: Spotlight metadata is available to applications running Mac OS X version 10.4 and later.

Organization of This Document

The following articles cover key concepts in understanding how Spotlight works:

- "What is Spotlight?" (page 9) provides an overview of what Spotlight provides.
- "How Does Spotlight Work?" (page 11) describes how Spotlight creates metadata indexes, and how they are queried.
- "Spotlight Metadata Attributes" (page 13) provides an overview of metadata attributes.
- "Security and Privacy Considerations" (page 15) describes how Spotlight addresses security and privacy issues.
- "Spotlight and Document Bundles" (page 17) describes how an application should create document bundles with Spotlight in mind.

See Also

There are other technologies, not fully covered in this document, that are fundamental to integrating Spotlight into your applications. Refer to these documents for more details:

- *Spotlight Query Programming Guide* describe how to query Spotlight from your applications.
- *Spotlight Importer Programming Guide* describe the plug-ins that extract metadata from document files.
- *Spotlight Metadata Attributes Reference* describe the metadata attributes provided by Apple.

2007-05-27 | © 2004, 2007 Apple Inc. All Rights Reserved.

2007-05-27 | © 2004, 2007 Apple Inc. All Rights Reserved.

What is Spotlight?

Organizing data on a computer is a difficult job, and in most cases is the sole responsibility of the user. However, even the most organized user may find it nearly impossible to arrange their files in a way that makes it easy to find information. Because the underlying file systems offer only one way of organizing information, users must resort to special tools to search for what they want. The problem is that most search tools can be slow and limited in how they do their search. Also, users may want to search more than files. Users may want to search their mail archives, address book contacts, or other digital assets embedded inside a file.

What does Spotlight provide?

Spotlight provides a new way of organizing and accessing information in on your computer by using metadata. Metadata is data about a file, rather than the actual content stored in the file.

Metadata can include familiar information such as an asset's author and modification date but it can also be keywords or other information that is custom to a particular asset. For example, an image file might have metadata describing the image's dimensions and color model. Spotlight can use this information to allow a user to find all their high-resolution images using a CMYK colorspace.

Where Does Metadata Come From?

For Spotlight searching to work, it has to have metadata. While some bits of metadata (modification dates, file type, path name) are easy to gather for a given file, most of the interesting data is embedded inside the file. To gather this embedded information you must provide a Spotlight importer.

A Spotlight importer is a small plug-in bundle that you create to extract information from files created by your application. Spotlight importers are used by the Spotlight server to gather information about new and existing files.

Apple provides importers for many standard file types that the system uses, including RTF, JPEG, Mail, PDF and MP3. However, if you define a custom document format, you must create a metadata importer for your own content.

2007-05-27 | © 2004, 2007 Apple Inc. All Rights Reserved.

Important: It is imperative that developers provide Spotlight importers for their own custom document formats. Spotlight importers improve the user experience greatly by making sure your documents can be found during searches.

Searching Metadata in your Applications

The simplest way to provide Spotlight searching in your application is to display the standard Spotlight search window to users directly from your application using the `HISearchWindow` function, defined in the Carbon framework.

The Spotlight metadata framework provides a query API, `MDQuery`, that you can use in your application to search for files based on metadata values. Using this API, you initiate a query using the query expression syntax. You can sort the results based on different attributes and use those sorting criteria to organize the data before you present it to the user.

The Cocoa framework provides the `NSMetadataQuery` class that is an Objective-C interface to the `MDQuery` API. `NSMetadataQuery` uses the `NSPredicate` classes to construct queries rather than the Spotlight query expression syntax.

Not Only About Search

Spotlight is not only a means of searching for documents. Spotlight importers define metadata that Finder can display in its Get Info panel. This information provides more context and details about documents.

For example:

- Image files provide their dimensions, pixel depth and other color related information.
- QuickTime movies provide their duration.
- PDF files provide information on the authors, security, dimensions, encoding, and where they originated.

It is the richness of the metadata that you provide that will improve the user experience when using your custom files. See Spotlight Importer Schema Format for more information on displaying your attributes in Finder.

2007-05-27 | © 2004, 2007 Apple Inc. All Rights Reserved.

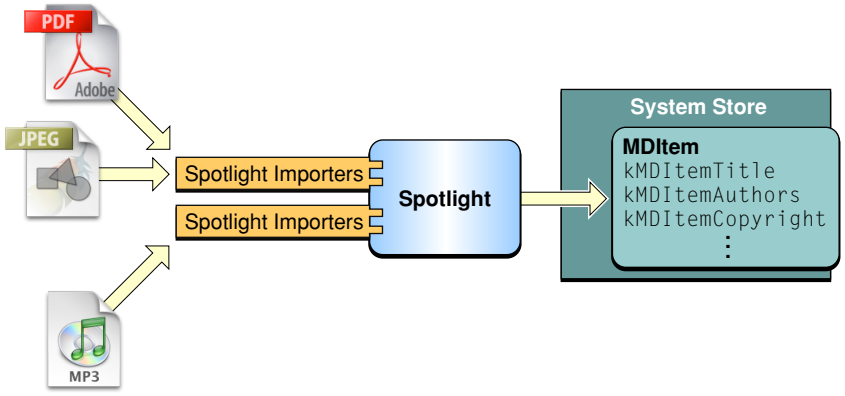
How Does Spotlight Work?

Spotlight provides fast desktop searching by extracting metadata in the background and storing the indexed metadata for future searches. When a query is made, the indexed metadata is searched for matching files.

Extracting Metadata from Files

Every time a file is created, modified or deleted, the kernel notifies the Spotlight engine that it needs to update the system store for changed file. Using Launch Services, Spotlight determines the uniform type identifier of the file and attempts to find an appropriate importer plug-in. If an importer exists and is authorized, it is loaded and passed the path to the file.

Figure 1 Extracting Metadata



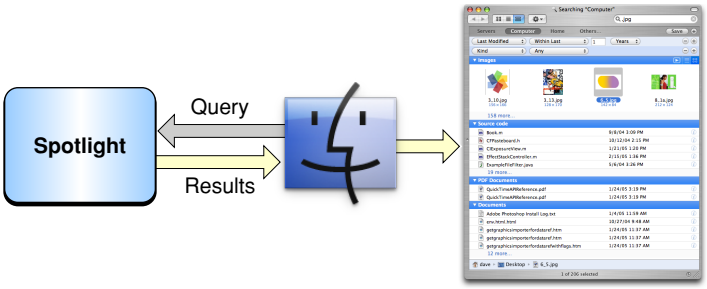
It is the importer's responsibility to then read the data file and construct a dictionary that contains the appropriate metadata. When finished extracting the metadata, the dictionary is returned to the Spotlight engine, which then updates the system store.

Querying Spotlight for Metadata

Spotlight queries are made by client applications, such as the Finder. The application constructs the appropriate query expression for the search, specifies the scope of the search, how the data is to be grouped when it is returned, and then executes the query. The query is passed to the Spotlight engine, which begins the initial result-gathering phase of the search. During this phase the system store is searched for metadata that matches the query, and it returns the search results to the application.

2007-05-27 | © 2004, 2007 Apple Inc. All Rights Reserved.

Figure 2 Querying Spotlight



If the query is configured to return live-update results, Spotlight notifies the client application when a change to the system store is made that causes the search results to change. Changes to the system store can cause additional files to match the query or cause files that initially matched to no longer match the query. Spotlight notifies the client application of the type of change, and the client application can update its results as appropriate.

2007-05-27 | © 2004, 2007 Apple Inc. All Rights Reserved.

Spotlight Metadata Attributes

The first step in providing support for Spotlight is to decide what information a user will want to look for in documents they create using your application. Once you determine the data that should be extracted and provided to users, you need to assign those values to metadata attributes.

Spotlight defines standard metadata attributes that provide a wide range of options for storing your document's metadata. Users can restrict searching to specific attributes, and so it is important that you should use Spotlight's standard metadata attributes whenever possible.

For example, many documents contain a company name that should be extracted as metadata. Spotlight doesn't provide an explicit company name attribute. It does, however, define an attribute that stores an "organization"—`kMDItemOrganizations`—which is appropriate for a company name. The Spotlight provided metadata attributes are documented in *Spotlight Metadata Attributes Reference*.

If an appropriate metadata attribute isn't provided by Spotlight, you should look to third-party developers to see if they have already defined an appropriate attribute key. If no appropriate Spotlight or third-party metadata attribute has been defined, then you can create your own custom metadata attributes.

2007-05-27 | © 2004, 2007 Apple Inc. All Rights Reserved.

2007-05-27 | © 2004, 2007 Apple Inc. All Rights Reserved.

Security and Privacy Considerations

On systems with separate user accounts, Spotlight respects the ownership of users' files—even though the system store is shared. Spotlight filters all query results, removing any files that the user isn't permitted to see.

However, developers should still be aware of users' data privacy and security when extracting metadata from documents.

Protecting Sensitive Data

When determining the data to extract from your application's documents you should consider the implications of making that data searchable. For example, could making the author of the document available be a liability? Should the full content of a document be available for indexing, or only selected fields?

Developers should provide users with information about which data is indexed by their Spotlight importers. Developers should also consider providing preferences that allow users control over what data is extracted as metadata.

Untrusted Applications and Bundled Importers

Spotlight importers typically reside within an application wrapper. When first copied to the user's computer an application is considered untrusted. The first time that the application is launched, Mac OS X presents a warning to the user. If the user accepts, the application is launched and considered "trusted."

Spotlight does not load and use a bundled importer if the associated application is not trusted.

2007-05-27 | © 2004, 2007 Apple Inc. All Rights Reserved.

2007-05-27 | © 2004, 2007 Apple Inc. All Rights Reserved.

Spotlight and Document Bundles

Spotlight is notified when an application creates or updates a document, and imports the metadata from the file. However, if your application saves documents as a document bundle Spotlight is often notified of the creation of the document bundle directory before an application has the opportunity to write the contents of the document to disk. There are several strategies available to work around this issue.

Create an Invisible Document Bundle

One technique for saving a document bundle is to take advantage of the fact that Spotlight does not import metadata from invisible files or directories.

Create the document bundle directory as an invisible directory by prefacing the filename with the "." character. Your application should create the directory using a temporary and unique filename with the appropriate filename extension. To ensure file-system compatibility, the temporary name should not exceed the length of the user specified filename. The application then writes the contents of the document to the invisible document bundle as normal. When completed the application then renames the invisible document bundle to the original document name.

Encapsulate Your Document with NSFileWrapper

Cocoa applications can use the NSFileWrapper class to create the document bundle as a directory wrapper. When the NSFileWrapper instance is written to disk the document bundle is created in such a way that Spotlight is notified when the entire document bundle has been saved.

2007-05-27 | © 2004, 2007 Apple Inc. All Rights Reserved.

2007-05-27 | © 2004, 2007 Apple Inc. All Rights Reserved.

Document Revision History

This table describes the changes to *Spotlight Overview*.

Date	Notes
2007-05-27	Added information about displaying file information in Finder.
2006-04-04	Clarified filename length issues in "Spotlight and Document Bundles."
2006-03-08	Updated to reflect Spotlight compatibility with Rosetta.
2006-02-07	Added note about recompiling as a universal binary.
2005-04-29	Added guidelines that describe how to create document bundles with Spotlight in mind.
	Corrected typos. Changed title from "Metadata."
2004-06-28	New document that introduces the main components and services of Apple's metadata technology.

2007-05-27 | © 2004, 2007 Apple Inc. All Rights Reserved.

2007-05-27 | © 2004, 2007 Apple Inc. All Rights Reserved.