

Oracle® Enterprise Repository
Software File Identification (SFID) Guide
10g Release 3 (10.3)

July 2009

Oracle Enterprise Repository Software File Identification (SFID) Guide, 10g Release 3 (10.3)

Copyright © 2008, 2009, Oracle. All rights reserved.

Primary Author: Vimika Dinesh

Contributing Author: Jeff Schieli, Sharon Fay, Casey Stella, Scott Spieker, Phil Reed, Adam Wallace, David Difranco

The Programs (which include both the software and documentation) contain proprietary information; they are provided under a license agreement containing restrictions on use and disclosure and are also protected by copyright, patent, and other intellectual and industrial property laws. Reverse engineering, disassembly, or decompilation of the Programs, except to the extent required to obtain interoperability with other independently created software or as specified by law, is prohibited.

The information contained in this document is subject to change without notice. If you find any problems in the documentation, please report them to us in writing. This document is not warranted to be error-free. Except as may be expressly permitted in your license agreement for these Programs, no part of these Programs may be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose.

If the Programs are delivered to the United States Government or anyone licensing or using the Programs on behalf of the United States Government, the following notice is applicable:

U.S. GOVERNMENT RIGHTS Programs, software, databases, and related documentation and technical data delivered to U.S. Government customers are "commercial computer software" or "commercial technical data" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, use, duplication, disclosure, modification, and adaptation of the Programs, including documentation and technical data, shall be subject to the licensing restrictions set forth in the applicable Oracle license agreement, and, to the extent applicable, the additional rights set forth in FAR 52.227-19, Commercial Computer Software--Restricted Rights (June 1987). Oracle USA, Inc., 500 Oracle Parkway, Redwood City, CA 94065.

The Programs are not intended for use in any nuclear, aviation, mass transit, medical, or other inherently dangerous applications. It shall be the licensee's responsibility to take all appropriate fail-safe, backup, redundancy and other measures to ensure the safe use of such applications if the Programs are used for such purposes, and we disclaim liability for any damages caused by such use of the Programs.

Oracle, JD Edwards, PeopleSoft, and Siebel are registered trademarks of Oracle Corporation and/or its affiliates. Other names may be trademarks of their respective owners.

The Programs may provide links to Web sites and access to content, products, and services from third parties. Oracle is not responsible for the availability of, or any content provided on, third-party Web sites. You bear all risks associated with the use of such content. If you choose to purchase any products or services from a third party, the relationship is directly between you and the third party. Oracle is not responsible for: (a) the quality of third-party products or services; or (b) fulfilling any of the terms of the agreement with the third party, including delivery of products or services and warranty obligations related to purchased products or services. Oracle is not responsible for any loss or damage of any sort that you may incur from dealing with any third party.

Oracle Enterprise Repository

Software File Identification (SFID) Guide

Table of Contents

- [Overview](#)
- [Enabling SFID](#)
- [Adding the SFID Element to a Type](#)
- [Adding an SFID to an Asset](#)
- [Limitations](#)
- [SFID Toolkit](#)
- [SFID Manager](#)
- [Calculating the SFID Programmatically](#)

Overview

Software File Identification (SFID) provides the ability to determine asset usage independent of the manual asset **Use - Download** process within Oracle Enterprise Repository. The SFID process tags selected files within an asset with a unique SFID fingerprint. This tag is then used to detect when and where an asset is used, even if the asset was acquired through means other than the **Use - Download** button in Oracle Enterprise Repository. An instance of usage is recorded by Oracle Enterprise Repository when tagged files within the asset are opened in a developer's IDE.

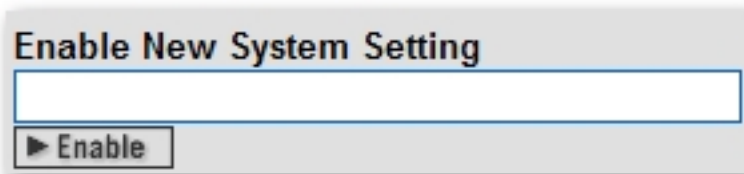
Important! Depending on your IDE environment, SFID requires the installation of the Oracle Enterprise Repository Plug-in for WorkSpace Studio, which is an Eclipse-based IDE, or the Oracle Enterprise Repository plug-in for Visual Studio .NET. For more information about repository access capabilities from a supported IDE, refer to the following documentation:

- *Repository Access in Eclipse Using WorkSpace Studio*
at http://download.oracle.com/docs/cd/E13164_01/oer/docs10134/eclipse/index.html
- *Oracle Enterprise Repository Integration with Visual Studio .NET*
at http://download.oracle.com/docs/cd/E13164_01/oer/docs10134/pdf/OER103-TP_VS_NET_Guide.pdf

Enabling SFID

This procedure is performed on the Oracle Enterprise Repository **Admin** screen.

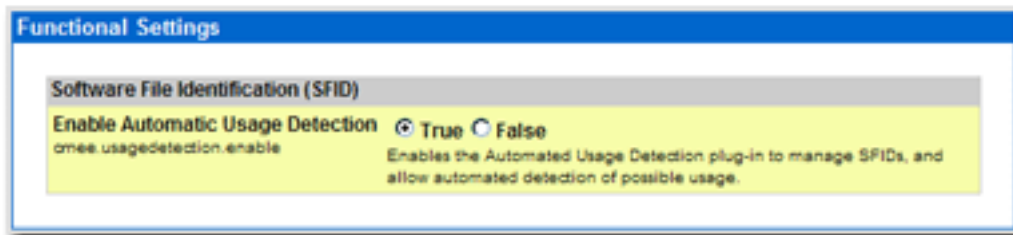
1. Click **System Settings** in the left pane.
2. Enter `cmee.usagedetection.enable` in the **Enable New System Settings** text box.



The image shows a dialog box titled "Enable New System Setting". It contains a text input field and a button labeled "Enable" with a right-pointing arrow.

3. Click **Enable**.

Enable Automated Usage Detection appears in the list of properties in the main pane.



The image shows a "Functional Settings" pane. Under the "Software File Identification (SFID)" section, the "Enable Automated Usage Detection" property is set to "True". The property name is `cmee.usagedetection.enable`. A description below reads: "Enables the Automated Usage Detection plug-in to manage SFIDs, and allow automated detection of possible usage."

4. Make sure the property is set to **True**.
5. Click **Save**.

SFID is now enabled.

- The **SFID** element can now be added to types in the **Type Manager**.
- When the **SFID** element is added to an asset type, an **SFID** section will

appear on assets of that type when viewed in the **Asset Editor**. It is within that section that files can be selected for SFID fingerprinting.

Automated Usage Detection

Please select the files that uniquely identify this asset. This information will be use...

Name	Description	Generated Date
------	-------------	----------------

Add

Remove

Update

- **SFID** functions are now available in the Oracle Enterprise Repository REX.

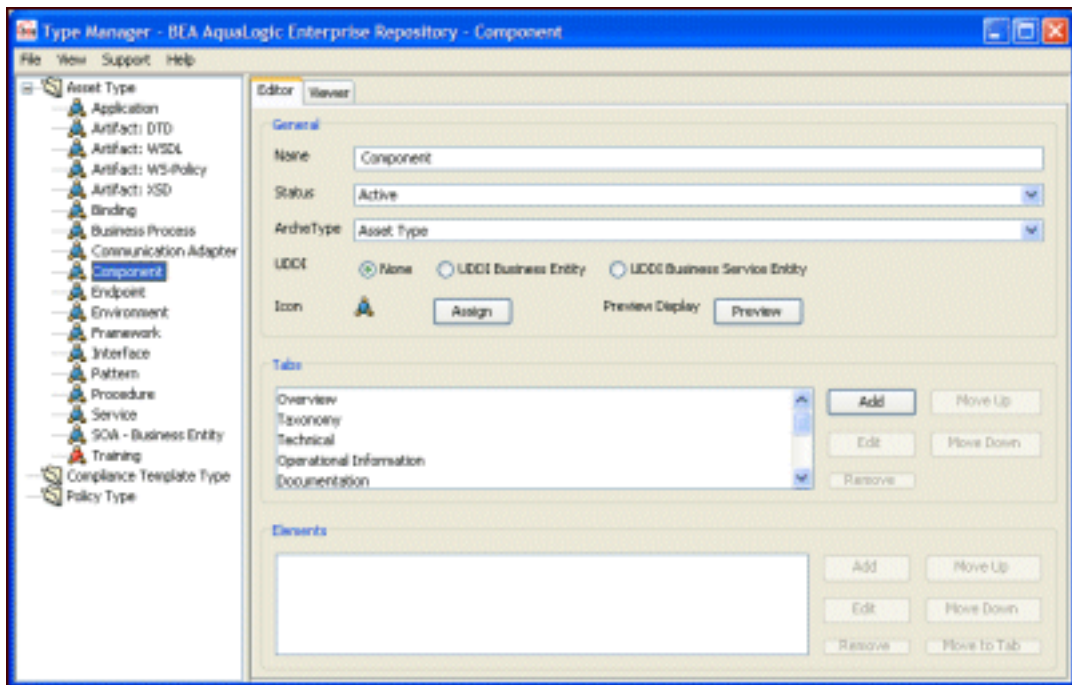
Adding the SFID Element to a Type

Adding the SFID element to a type enables assets of that type to be automatically detected.

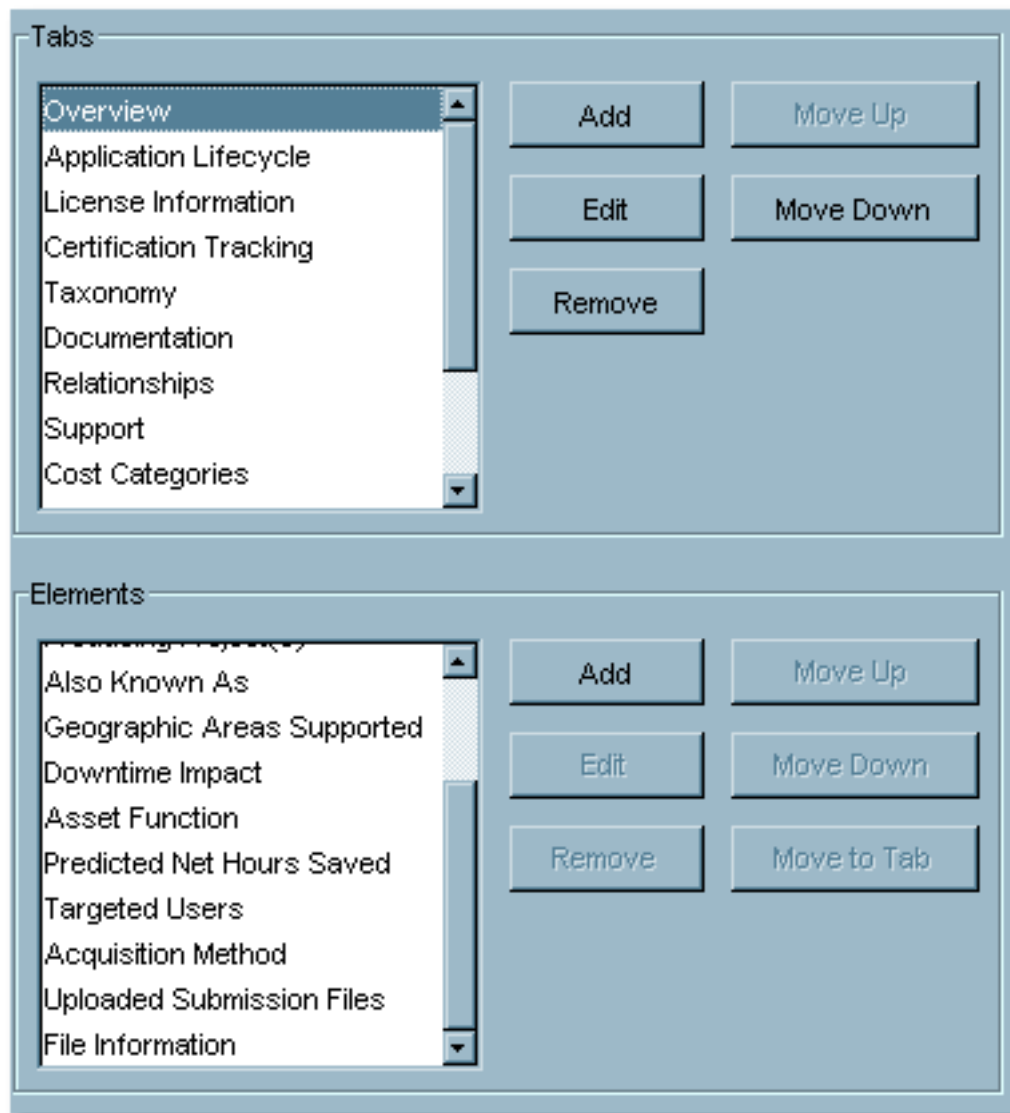
Note: You will also need to add an SFID to the individual assets to be tracked.(See Adding an SFID to an Asset).

This procedure is performed in the **Type Manager**.

1. Select the type to which the **SFID** element is to be added.



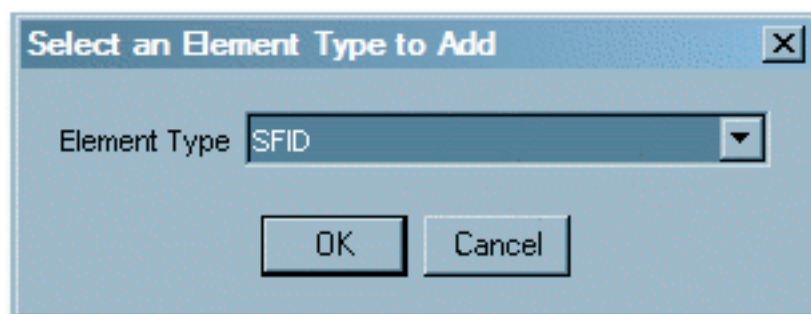
2. Click the **Editor** tab.
3. From the list in the **Tabs** section, select the tab to which to add the **SFID** element.



A list of the elements on that tab appears in the **Elements** section.

4. Click the **Add** button in the **Elements** section.

The **Select an Element Type to Add** pop-up opens.

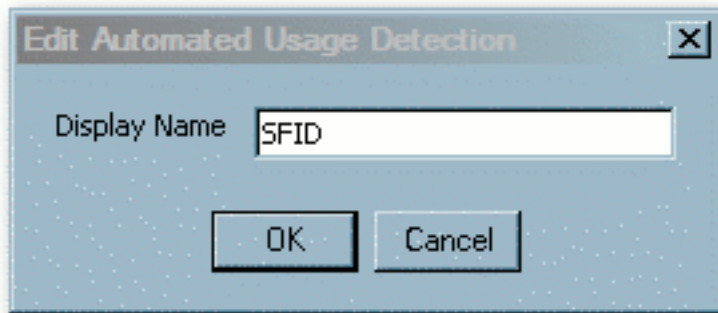


5. Select **SFID** from the drop-down menu.

6. Click **OK**.

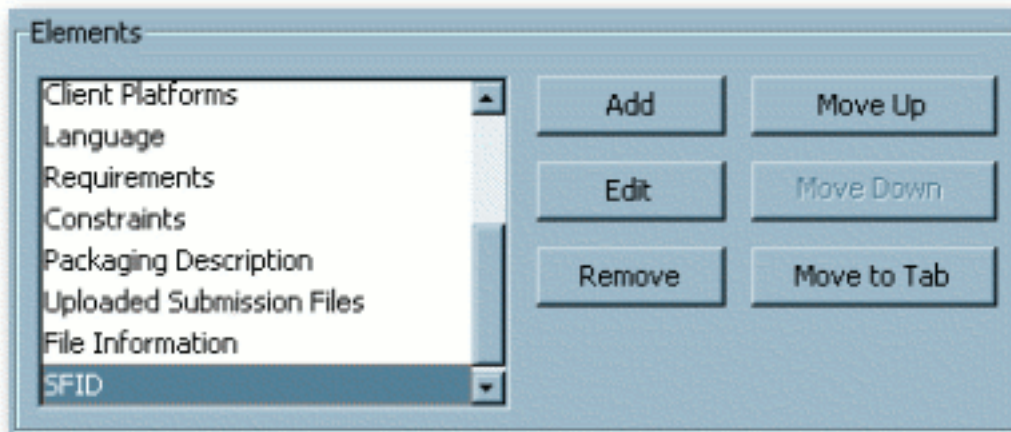
The **Select an Element Type to Add** pop-up closes, and the **Edit Automated Usage Detection** pop-up appears.

7. Enter a name for the **SFID** element.



8. Click **OK**.

The new element is listed in the **Elements** section.



9. Use the **File -> Save** menu to save the modified asset type.
10. Exit the **Type Manager**.
11. In the **Asset Editor**, open an asset of the type to which the **SFID** element was added.

The **SFID** section now appears on the tab to which it was assigned.

Automated Usage Detection

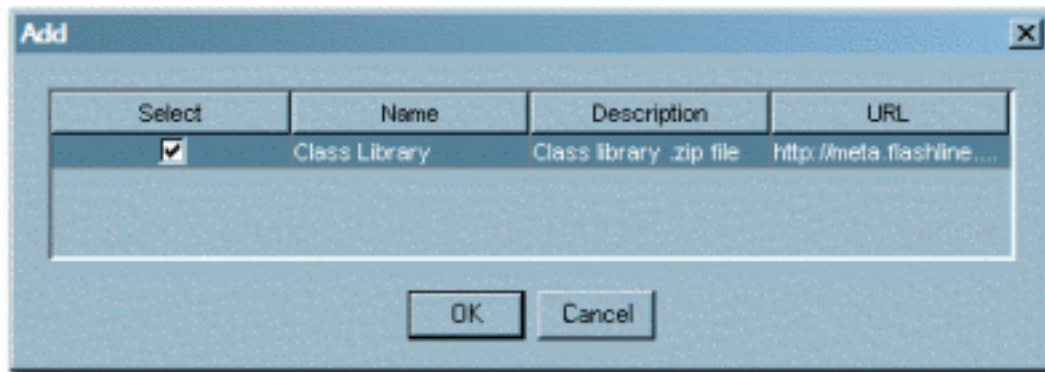
Please select the files that uniquely identify this asset. This information will be use...

Name	Description	Generated Date	Add
			Remove
			Update

Adding an SFID to an Asset

To track assets through an IDE, select the unique artifact file representing the asset. The file will be assigned an SFID, and usage will be automatically incremented every time the file is found in an IDE project. This procedure is performed in the **Asset Editor**.

1. Select the asset to which an SFID is to be added.
2. Click the **Add** button in the **Automated Usage Detection** section.
3. Select the files to be fingerprinted from the list in the **Add** dialog box.



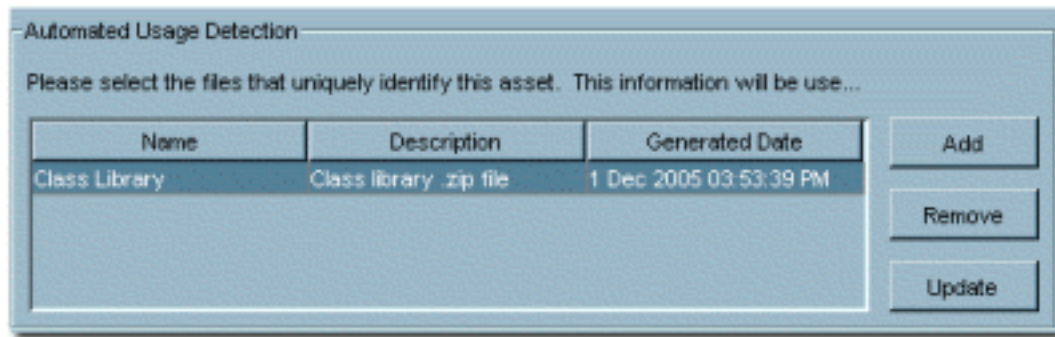
4. Click **OK**.

The system downloads the selected files, calculates the SFID, and then deletes the local (downloaded) copies of the files. A warning message is displayed if fingerprinting for any of the selected files is unsuccessful.

Update an SFID file

This procedure is performed in the **Asset Editor**.

1. Select the SFID file to be updated from the list in the **Automated Usage Detection** section.



2. Click the **Update** button.

The **Perform Operation** window opens, indicating the progress of the update.

Remove an SFID File

This procedure is performed in the **Asset Editor**.

1. Select the SFID file to be removed from the list in the **Automated Usage Detection** section.
2. Click the **Remove** button.

The file is removed from the list.

Limitations

SFID System Limitations

- The file must be accessible to generate the SFID.
 - The client machine/user must be provided with a path (URI) to the file. The URI must be accessible from the client system, without the need for user intervention, such as requests for authentication information).
- SFIDs are associated with files (1:1) which are associated with assets (n:n)
- SFIDs can be added to files through the **Asset Editor** or through **REX**.

Limitations of the Asset Editor (SFID generation)

- An SFID can be associated only with extractable files.
- An SFID can be generated only for files that are accessible to the client.

Limitations of the ANT Integration (SFID comparison)

- SFIDs can be generated for ALL accessible files (for comparison).

Oracle Enterprise Repository

Software File Identification (SFID) Tool Kit

Overview

The **SFID Toolkit** provides a framework that allows third party developers to develop applications that integrate with Oracle Enterprise Repository's REX subsystem for the purpose of submitting SFID Candidates. The SFID toolkit provides an application that will create an SFID Candidate for each file in a specified directory and then submit those SFID Candidates to the Oracle Enterprise Repository system for storage and analysis.

Notes on the Calculation of SFIDs

- A file is considered a text file if it contains no characters above ASCII code 128.
- The content of a text file is normalized before the SFID hash is computed.
 - All CR, LF, and Tab chars are stripped.
 - All text is converted to lower case.
- XML Files are canonicalized as per the W3C canonicalization standard (<http://www.w3.org/TR/xml-c14n>) before the SFID hash is computed.

Installation / Configuration

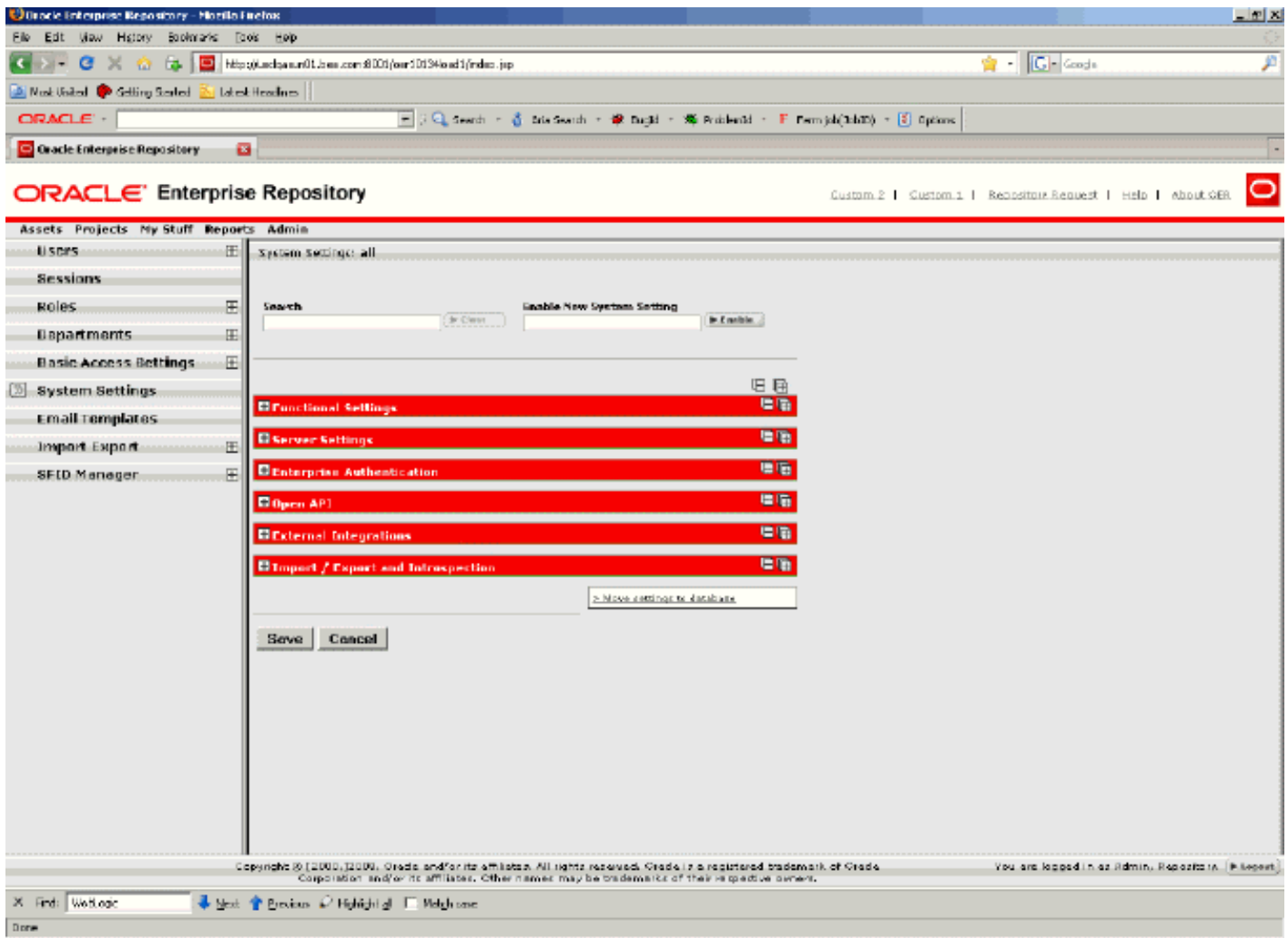
Enable Harvesting of Unrecognized SFIDs

Some organizations mandate that every SOA asset deployed should be in the Oracle Enterprise Repository. Companies can monitor compliance to this mandate by harvesting unrecognized SFIDs. For example, companies may want to monitor the build scripts for unrecognized SFIDs. Additionally, Service monitoring tools may discover services that are not in OER. These "rogue services" can be assigned an SFID and harvested into OER as asset "{candidates. Once in OER, they can be further evaluated and processed.

This procedure is performed on the Oracle Enterprise Repository **Admin** screen.

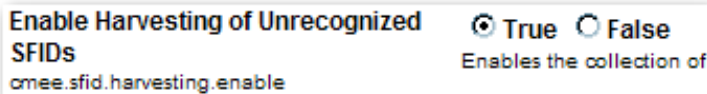
1. Click **System Settings** in the left pane.

The **System Settings** section opens in the main pane.



2. Enter `cmee.sfid.harvesting.enable` in the **Enable New System Setting** text box.
3. Click the **Enable** button.

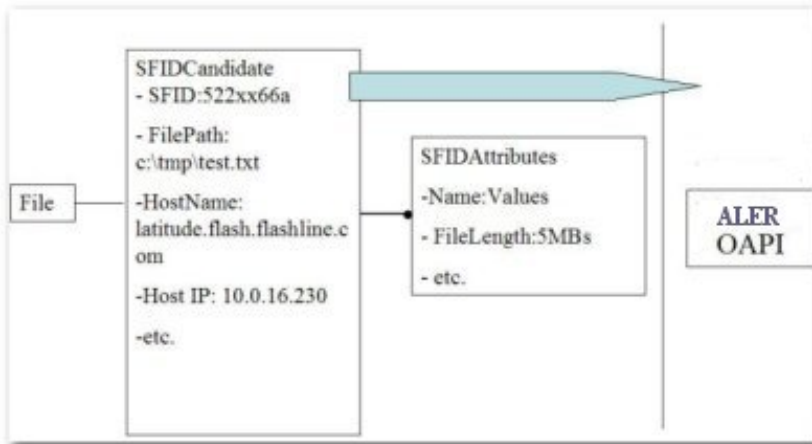
The **Enable Harvesting of Unrecognized SFIDs** property appears in the list of system settings.



4. Make sure property is set to **True**.
5. Click **Save**.

The property is now enabled.

Object Model



Description of Sample Directory Crawler

- Iterate through all the files and folders of a specified folder.
 - Instantiate an `SFIDCandidate` for each candidate
 - Populate general metadata such as path and sfid.
 - For each `SFIDCandidate`, create `SFIDAttributes` that describe arbitrary data to be captured, such as Eclipse Workspace, etc.
- Invoke the Oracle Enterprise Repository REX method `sfidCandidateSubmit`, passing an array of `SFIDCandidate` objects.

• AlerUtil Object

- Instantiate a `AlerUtil` object to connect to Oracle Enterprise Repository.

```
String IUrl = "http://localhost:9080/oer/services/FlashlineRegistry";
String IUserName = "username";
String IPassword = "password";
IAlerUtil = new AlerUtil(IUrl, IUserName, IPassword);
```

- Use the `AlerUtil` object to invoke a Oracle Enterprise Repository REX method.

```
IAlerUtil.getFlashlineRegistry().sfidCandidateSubmit(IAlerUtil.getAuthToken(), ISfidCandidates);
```

- The `AlerUtil` object provides a series of convenience methods.

- **getAuthToken**

Obtain a valid `authToken` for the for current username and password. If the `authToken` is expired, obtain a new `authToken`. Otherwise, a `UserValidationException` is thrown.

- **calculateSfidValue**

Calculate the SFID for a file or path and return the hash as a string.

- **sfidCandidateListToArray**

Convert a list of `SFIDCandidates` to an array of `SFIDCandidates`.

- **`sfidAttributeListToArray`**

Convert a list of `SFIDAttributes` to an array of `SFIDAttributes`.

- **`sfidAttributeValueListToArray`**

Convert a list of `SFIDAttributeValues` to an array of `SFIDAttributeValues`.

- **`createCandidate(File)`**

Create a standard `SFIDCandidate` for a file. `SFIDCandidate` fields `filepath`, `hostip`, `hostname`, and `sfid` will be pre-populated. An `SFIDAttribute` named *Filelength* will be attached to the `SFIDCandidate` with the value of length of the file.

- **`createCandidate(File, Map)`**

Create a standard `SFIDCandidate` for a file as specified by the method `createCandidate(File)`. Each `(String)key/(String)value` in the map will attach an additional `SFIDAttribute` to the `SFIDCandidate` whose name is `key` and whose value is `value`.

- **`createCandidate(File, List)`**

Create a standard `SFIDCandidate` for a file as specified by the method `createCandidate(File)`. Attach each `SFIDAttribute` passed in the list to the `SFIDCandidate`.

- **`sendCandidate`**

Send a single `SFIDCandidate` for harvesting.

- **`sendCandidates`**

Send a list of `SFIDCandidates` for harvesting.

SFID Manager

Overview

The **SFID Manager** provides an easy way to view SFID candidates harvested by Oracle Enterprise Repository.

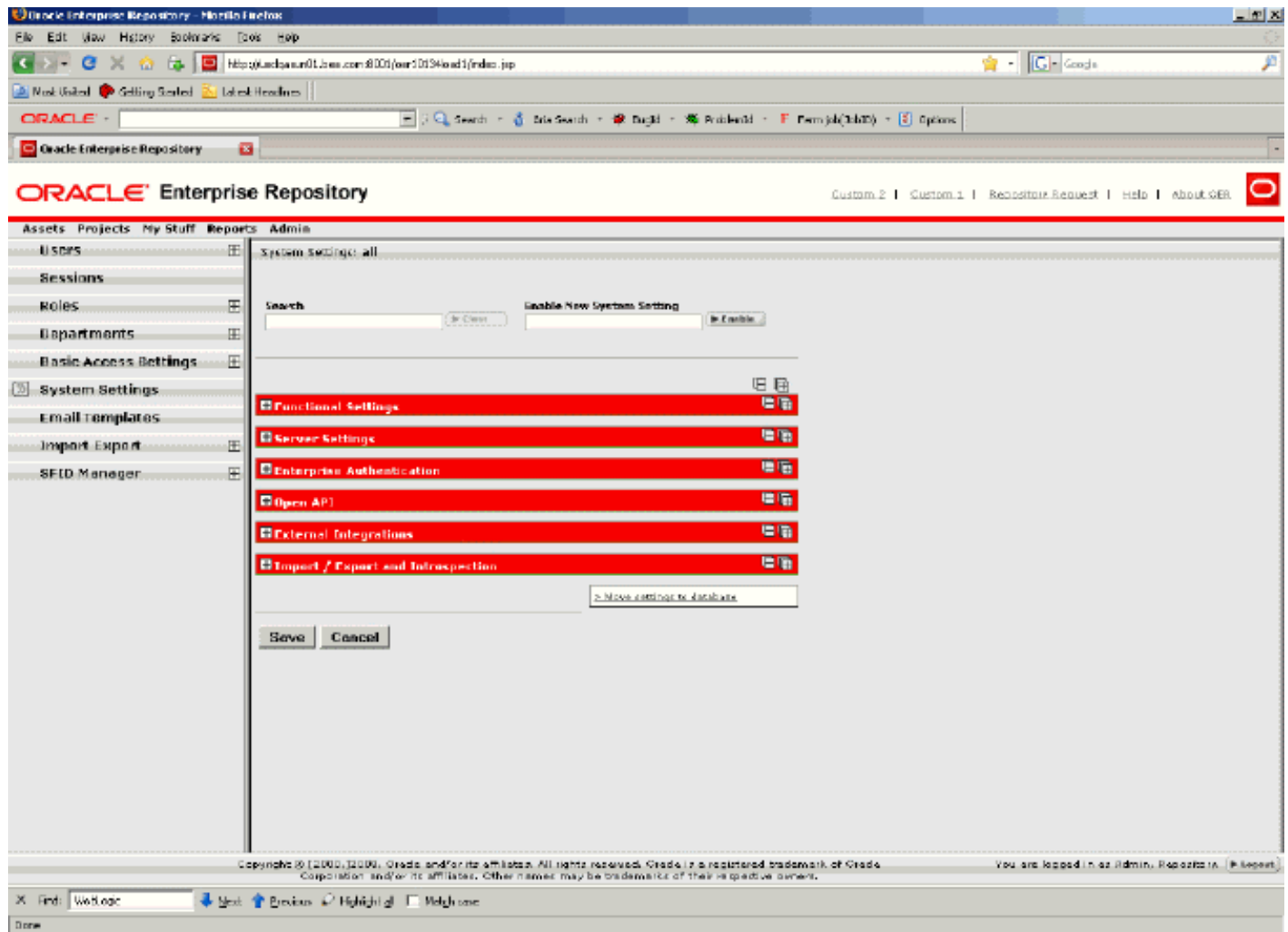
Use of the **SFID Manager** requires System Administrator privileges.

Enabling the SFID Manager

This procedure is performed on the **Admin** screen in Oracle Enterprise Repository.

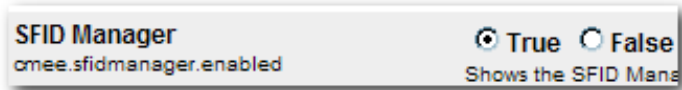
1. Click **System Settings** in the left pane.

The **System Settings** section opens in the main pane.



2. Enter `cmee.sfid.harvesting.enable` in the **Enable New System Setting** text box.
3. Click the **Enable** button.

SFID Manager is displayed in the list of system settings.



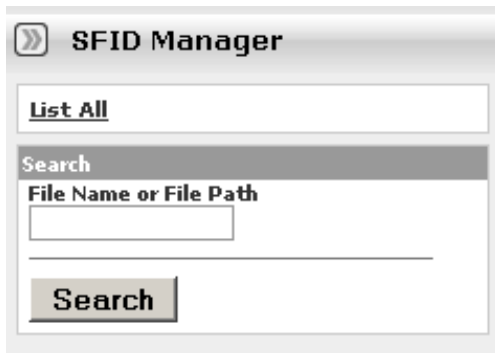
4. Make sure the **True** radio button is selected.
5. Click **Save**.
6. Click **Admin** in the Oracle Enterprise Repository menu bar to refresh the **Admin** screen.

The **SFID Manager** section now appears in the left pane.

Using the SFID Manager

This procedure is performed on the **Admin** screen in Oracle Enterprise Repository.

1. Click **SFID Manager** in the left pane.
2. Click **List All** in the **SFID Manager** section to display a list of available SFID candidates.



Be advised that the number of available candidates affects the time it takes for this operation to complete.

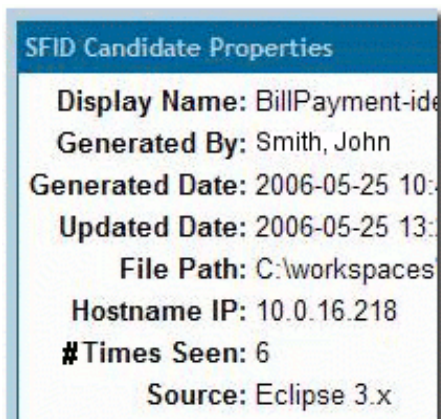
SFID candidates can also be located using the **SFID Manager** search function.

Identified candidates will be listed in the upper section of the main pane.

Name	#	SFID
BillPayment-ideal.wsdl	6	ef89ff1d1e09dffe4529f0ebdd082fc0
Shell.java	4	aad1e63a213a770cddd906ee875ef3d8
Taco.java	6	f0f95effad81248d2f53fe9491a2c5da

The number of time the listed SFID has been detected (**Times Seen**) is indicated in the **#** column.

3. Click any listed candidate to display extended information in the bottom section of the main pane.



The displayed information consists of a core set of attributes common to all candidates:

- **Display Name**
 - The name of the candidate. Generally, the *simple* file name.
- **Generated By**
 - The username of the user that caused the candidate to be generated.
- **Generated Date**
 - The date on which the candidate was created.
- **Updated Date**
 - The last time that the candidate record was modified in the database.
- **File Path**
 - The absolute path to the file on the machine from which the candidate was harvested.
- **Hostname/IP**
 - The hostname or IP address of the machine from which the candidate was harvested
- **# Times Seen**
 - The number of times this candidate has been viewed in this context. This metric gives an indication of the frequency of file use, which provides some indication of the value of the file in this context.
- **Source**
 - A description of the context (e.g., Eclipse, Ant, etc.)

The display also includes an extended set of attributes that will vary based on the context from which the candidate was harvested.

Calculating the SFID Programmatically

Overview

The "[ArtifactFingerprinter](#)" java library provides functionality for calculating SFID fingerprints programmatically in the client-side code.

Installation

Ensure that you have the `components-flashtrax.jar` file and it is in your java classpath for installation.

Usage

The `com.oracle.artifact.ArtifactAlgorithm` class contains several static instances, each of which encapsulates the fingerprinting algorithm for a particular type of file.

The supported algorithms are

- DEFAULT: fingerprints based on the file contents, without any transformations
- XML: normalizes the XML contents to remove whitespace, and normalize namespace abbreviations
- XSD: same as XML, plus inlines XSD imports and includes
- WSDL: same as XSD, plus inlines WSDL imports and includes, removes documentation, and orders WSDL elements by name
- WSDL_ABSTRACTING: same as WSDL, plus removes endpoint and binding information

Example

An example that describes calculating the SFID programmatically is as follows:

```
ArtifactFingerprinter fp = ArtifactAlgorithm.WSDL.newFingerprinter();
```

```
String sfid = fp.getFingerprint(new File("mytest.wsdl"));
```