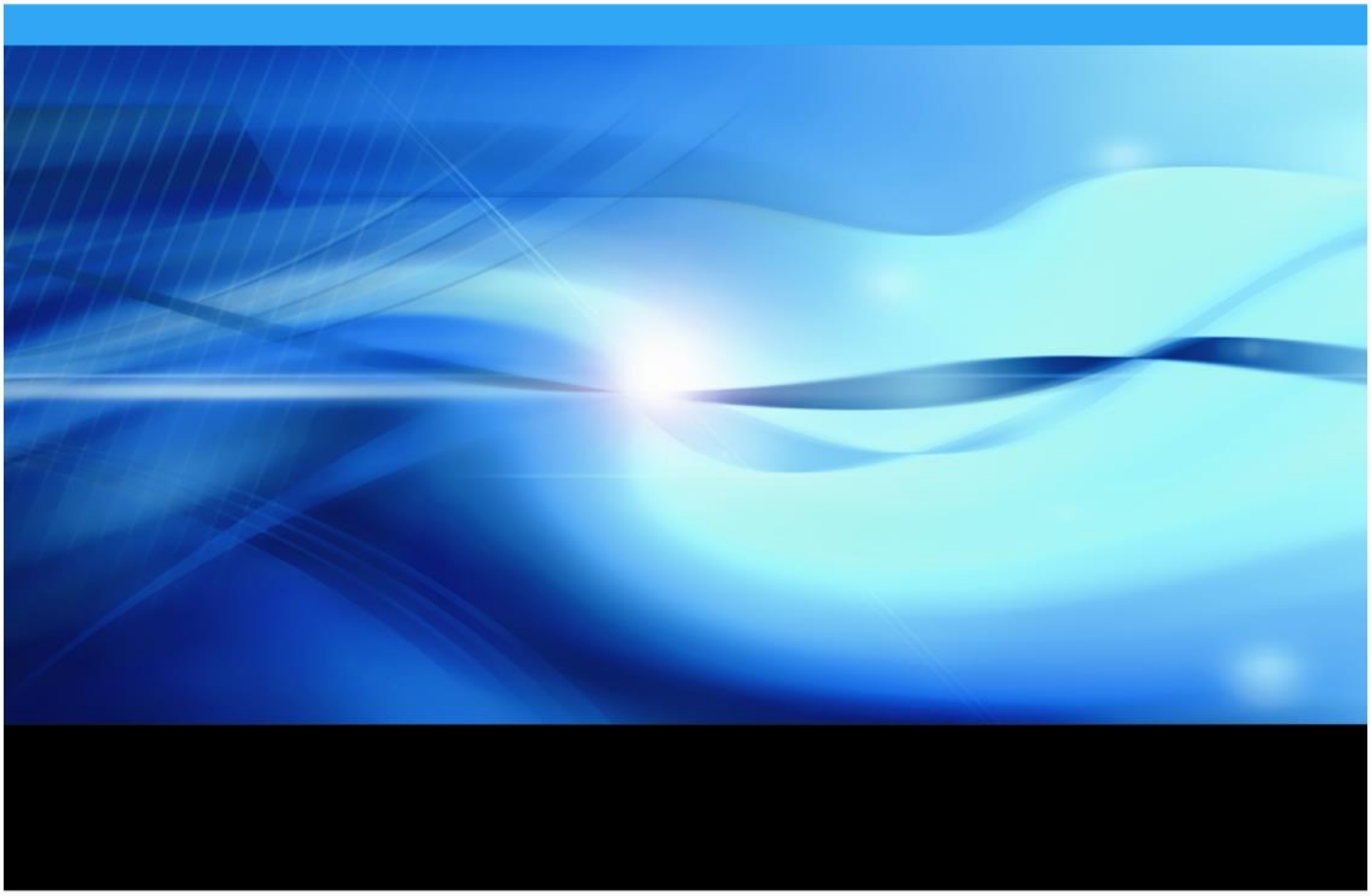
Integrating Search Interface to SAS® Content 3.2



Copyright Notice

The correct bibliographic citation for this manual is as follows: SAS Institute Inc., *Integrating Search Interface to SAS® Content 3.2*, Cary, NC: SAS Institute Inc., 2012.

Integrating Search Interface to SAS® Content 3.2
Copyright © 2012, SAS Institute Inc., Cary, NC, USA.

All rights reserved. Printed in the United States of America. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, by any form or by any means, electronic, mechanical, photocopying, or otherwise, without the prior written permission of the publisher, SAS Institute Inc. Limited permission is granted to store the copyrighted material in your system and display it on terminals, print only the number of copies required for use by those persons responsible for installing and supporting the SAS programming and licensed programs for which this material has been provided, and to modify the material to meet specific installation requirements. The SAS Institute copyright notice must appear on all printed versions of this material or extracts thereof and on the display medium when the material is displayed. Permission is not granted to reproduce or distribute the material except as stated above.

U.S. Government Restricted Rights Notice. Use, duplication, or disclosure of the software by the government is subject to restrictions as set forth in FAR 52.227-19 Commercial Computer Software-Restricted Rights (June 1987).

SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513.

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries.

® indicates USA registration.

Other brand and product names are trademarks of their respective companies.

Table of Contents

Chapter 1 — Initial Integration 3

Integrating with Search Engines 3

Integrating Search Interface to SAS Content with SAS Information Retrieval Studio 3
Integrating Search Interface to SAS Content with Google Search Appliance 3
Integrating Search Interface to SAS Content with Microsoft SharePoint Search Server 3

Chapter 2 — Configuration Information for SAS Information Retrieval Studio and Feeding SAS Contents 4

Step 1 - Verify the Registration Information in Metadata 4

Step 2 - Verify Configuration of SAS Information Retrieval Studio to Support Feeding SAS Contents into the Index 6

Step 3 - Verify Scheduling of the loadindex Script with Required Parameters 7

Advanced Configuration Options 7

Chapter 3 — Registering Search Interface to SAS Content as Google OneBox Module 8

Define the OneBox Module to the GSA 8

Trigger 9
Examples 9
Provider 10
Authentication 10
OneBox Stylesheet Template 11
Additional Authentication Steps 14
Add the OneBox Module Definition to a Front End 14

Chapter 4 — Registering Search Interface to SAS Content as a Federated Location in Microsoft SharePoint Search Server 15

Importing and Modifying the Federated Location 15

Adding the Federated Location WebPart with the Search Results 21

Chapter 5 - Logging for Search Interface to SAS Content and loadindex 23

Logging for Search Interface to SAS Content 23

Logging for loadindex 23

Appendix — Troubleshooting 24

Technical Support 29

Chapter 1 — Initial Integration

Integrating with Search Engines

Integrating Search Interface to SAS Content with SAS Information Retrieval Studio

Refer to "Chapter 2 — Configuring and Feeding SAS Contents to the Index of SAS Information Retrieval Studio" for the procedure to feed the SAS contents to the index server of SAS Information Retrieval Studio.

Integrating Search Interface to SAS Content with Google Search Appliance

Search Interface to SAS Content can be integrated with the Google Search Appliance by integrating it as an OneBox module. Refer to "Chapter 3 — Registering Search Interface to SAS Content as Google OneBox Module."

Integrating Search Interface to SAS Content with Microsoft SharePoint Search Server

Refer to "Chapter 4 — Registering Search Interface to SAS Content as a Federated Location in Microsoft SharePoint Search" for the procedure to integrate Search Interface to SAS Content with a Microsoft SharePoint Search Server.

Chapter 2 — Configuration Information for SAS Information Retrieval Studio and Feeding SAS Contents

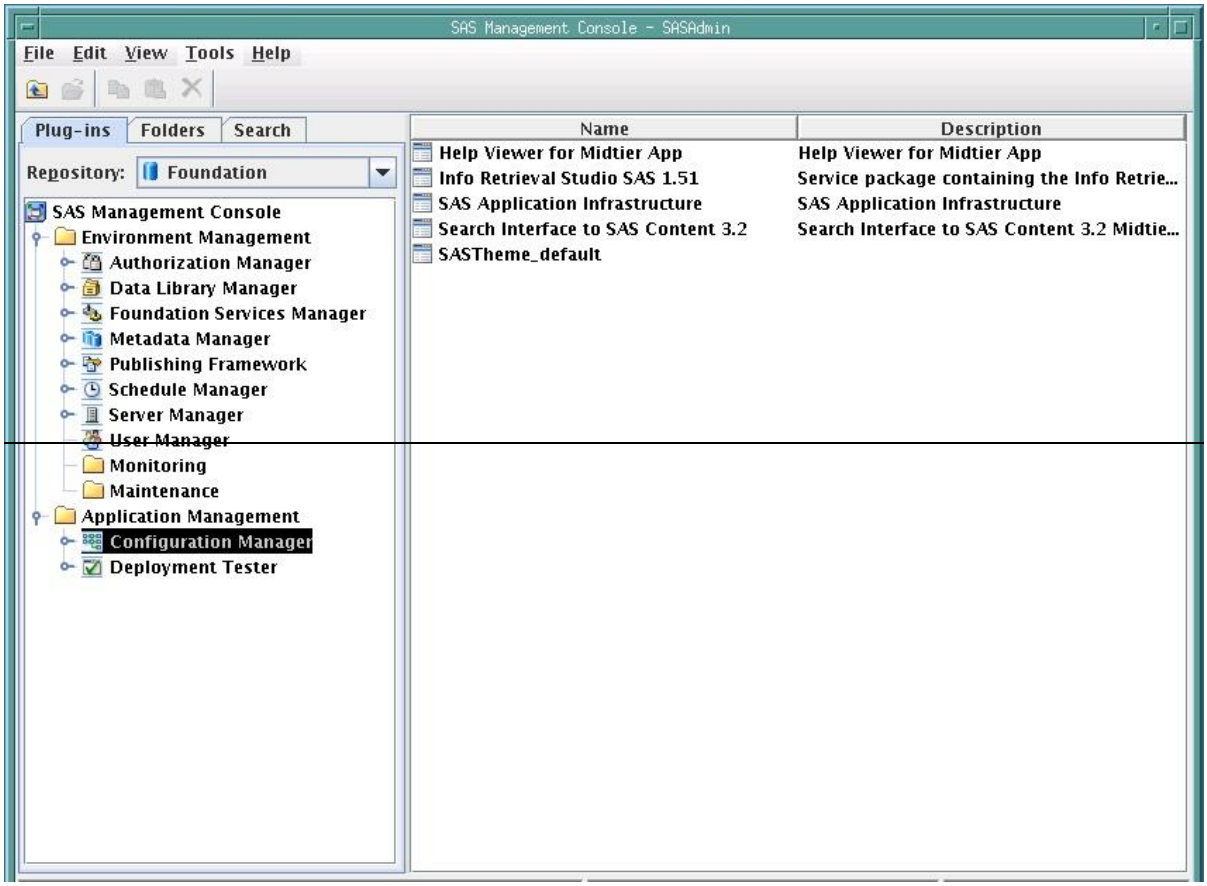
Search Interface to SAS Content 3.2 now supports feeding SAS contents to the index of SAS Information Retrieval Studio 1.51 and searching from it with authorization. Users will be able to view results based on their roles. Configuration is fully automated through the SAS Deployment Wizard (SDW). If any part is not working, follow the steps below to verify the configuration.

Step 1 - Verify the Registration Information in Metadata

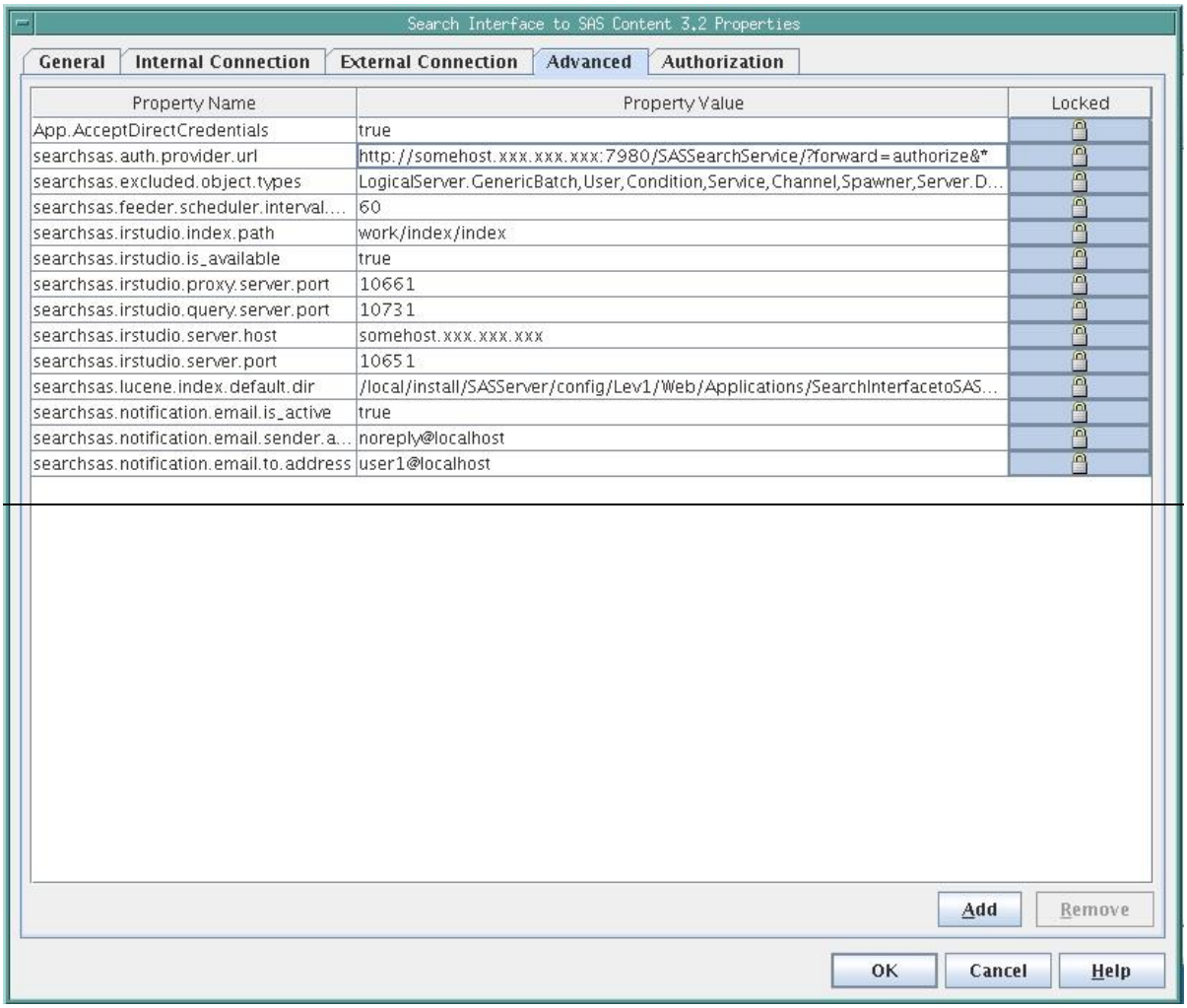
Make sure Search Interface to SAS Content and SAS Information Retrieval Studio are installed and running. Also verify that the correct information is used by Search Interface to SAS Content 3.2 to connect to SAS Information Retrieval Studio 1.51.

Use the following steps for verification:

1. Launch SAS Management Console and login with administrator credentials.
2. Go to the **Plug-ins** tab.
3. Navigate to **SAS Management Console→Application Management** and select **Configuration Manager**.



4. Right-click on **Search Interface to SAS Content** 3.2 and select Properties and then click on the **Advanced** tab in the dialog.



5. Ensure that the following properties have correct information, which should have been used while installing and configuring SAS Information Retrieval Studio:

   - **searchsas.irstudio.server.host:** the host name of the machine on which SAS Information Retrieval Studio servers are running.
   - **searchsas.irstudio.server.port**: the port number of the machine on which SAS Information Retrieval Studio is running.
   - **searchsas.irstudio.query.server.port**: the port number of the machine on which the query server of SAS Information Retrieval Studio is running.
   - **searchsas.irstudio.proxy.server.port**: the port number of the machine on which the proxy server of SAS Information Retrieval Studio is listening.

   Note: *If you are unsure of the correct values, see the configuration document of SAS Information Retrieval Studio.*

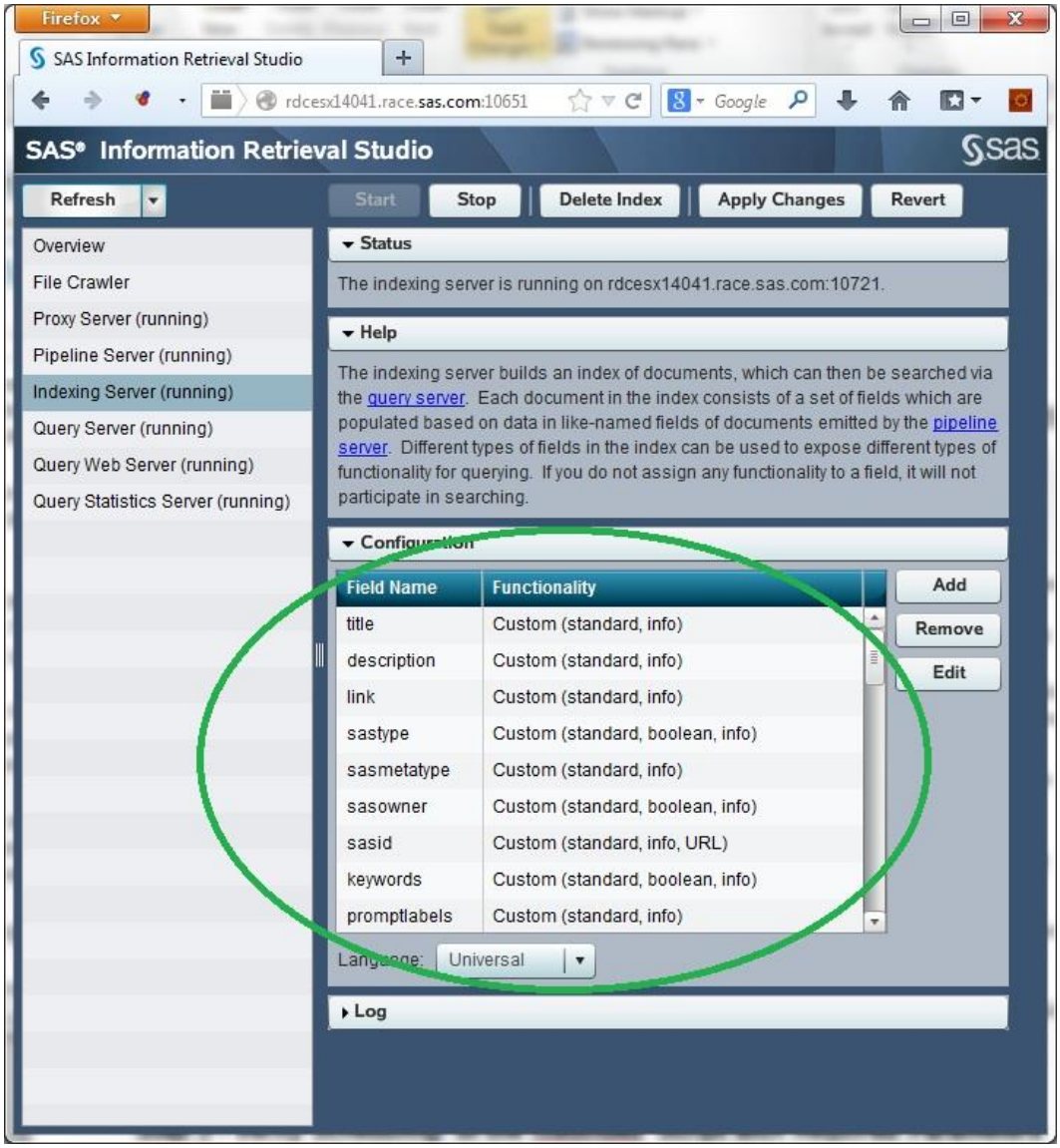
The following are internal properties and should never be changed:

- **searchsas.auth.provider.url**: Search Interface to SAS Content uses this URL for authorization.
- **searchsas.irstudio.index.path**: Search Interface to SAS Content uses this path to initialize SAS Information Retrieval Studio index path.

Step 2 - Verify Configuration of SAS Information Retrieval Studio to Support Feeding SAS Contents into the Index

Search Interface to SAS Content 3.2 needs to configure SAS Information Retrieval Studio to support authorization and feed SAS contents into the index. The configuration will be done automatically when the loadindex program runs for the first time.

To verify the configuration, open the admin console of Information Retrieval Studio and go to Indexing Server. The configuration section should show the fields populated for the information to be indexed.



Step 3 - Verify Scheduling of the loadindex Script with Required Parameters

The configuration process creates the scheduled task for loading index automatically. It can be verified if a cron job is created for the operating system scheduler. If the SAS Deployment Wizard configuration fails to schedule the loadindex task, Instructions.html shows a warning message to indicate that it should be scheduled manually.

For LAX systems, as soon as the cron job is set, there will be changes in the cron settings for the user who performed the configuration.

To modify the cron settings, a user can use the scheduler.sh script which resides in the following directory *with slight modifications:*

SAS-Configuration-Directory`/Lev`*n*`/Web/Applications/SearchInterfacetoSASContent`

If the user wants the alert emails to be sent to a different userid, then the property '`searchsas.notification.email.to.address`' under Advanced tab of the Software Component of 'Search Interface to SAS Content' needs to be modified with appropriate value.

`loadIndex` script (loadindex.sh for LAX and loadindex.bat for win) is available in the configuration directory of Search Interface to SAS Content (for example, `Config/Lev1/Web/Applications/SearchInterfacetoSASContent`).

Advanced Configuration Options

- When the loadindex application runs, you may need to increase the java heap size to avoid OutOfMemoryError. To increase the java heap size, open the `loadindex.ini` file located in the Search Interface to SAS Content installation home directory and add the following commands:

  ```
  JavaArgs_7=-Xmx512m
  JavaArgs_8=-Xms512m
  ```

  Note: *To provide more java heap size, you can change the heap size value from 512 to a more appropriate value.*

- If Search Interface to SAS Content is deployed in a Web app server configured with https, then the SSL certificate has to be imported in the JRE on which the loadindex application is running.

- If you are using a clustered environment, you may need to update the `url_list.txt` file to change the host and port information of the load balancing HTTP server. You can find the `url_list.txt` file in the following directory:

  <*config-dir*>/<*levn-dir*>/Web/Applications/SearchInterfacetoSASContent

Chapter 3 — Registering Search Interface to SAS Content as Google OneBox Module

The Search Interface to SAS Content searches all generic metadata types excluding those mentioned in the excluded list property of software component.

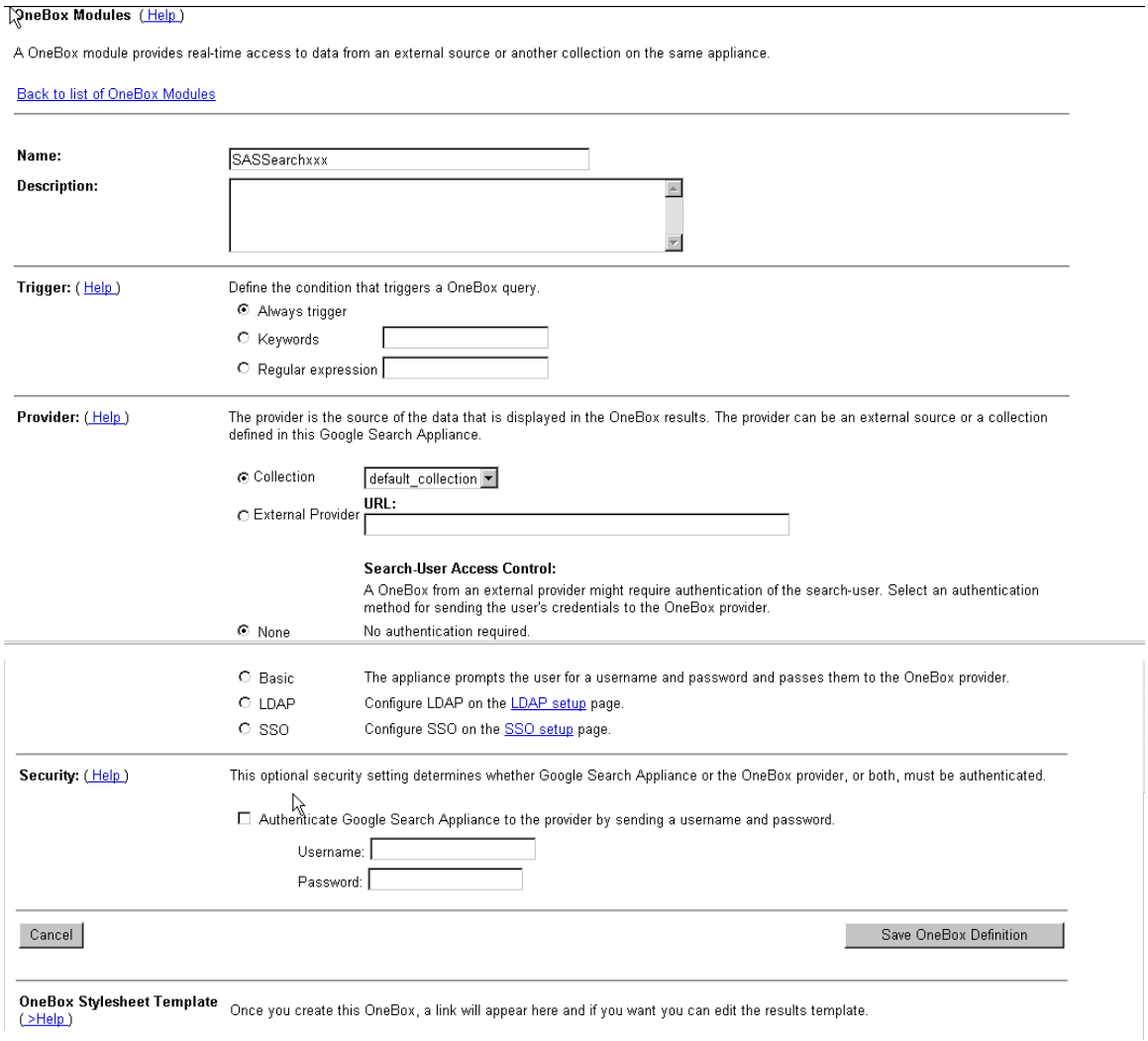
To search both types of content together, Search Interface to SAS Content must be registered as OneBox module. The sections below describe the steps required to use the Google Search Appliance (GSA) Administrative User Interface to perform the registration.

Define the OneBox Module to the GSA

Go to the OneBox Module definition screen in the Serving section of the interface.

At the bottom of this page there is an entry form to create a new OneBox module definition. While you can enter any name, it is recommended that you select a name that is generic enough to cover the triggers and items to be searched, but unique enough so that this module can be easily identified later from a list.

After you create the OneBox module definition, the next window asks you to enter details about the module.



Fill in the following fields.

Trigger

The only option currently supported is Regular expression. In the **Regular expression** field, a regular expression should be entered. When the input search matches the regular expression, this module will be called. The regular expression must contain a "parenthesized" field which will be passed as the search term to the module.

Note: *Regular expression processing can be expensive, so try to keep the regular expression simple.*

Examples

`(.*) report` – would match any search phrase that has the word report in it, and it would pass as the search string the phrase before that word. The phrase "sales report" would search for reports that have the word "sales" in it. "new sales report" would search for reports that have the phrase "new sales" in it.

`^report (.*)` – would match any search phrase where the first word in the search is "report" and it would pass as the search string the phrase after that word. The phrase "report sales"

would search for reports that have the word "sales" in it. The phrase "report new sales" would search for reports that have the phrase "new sales" in it. Note that this definition is equivalent to using the keyword trigger option; however, the keyword trigger option does not pass the appropriate phrase to the module, and thus cannot be used as a replacement for this syntax.

`^(.*sales.*)` – would match any search phrase that has the word "sales" in it. In addition, it will pass the entire entered search string to the module. Thus, if the user enters "new sales in US" that entire phrase will be passed, and only results that match the entire phrase will be returned.

`^promotions|promotion(.*)` – would match any search phrase where the first word in the search is either "promotion" or "promotions" and it would pass the search string the phrase after that word.

Provider

In the provider section, select **External Provider** and enter a URL that points to the location where you deployed the `sas.searchsas.ear` file.

For example, if you have deployed the EAR file to a Web server on *`yyy.mycompany.com`* that is listening on port 8080, you would enter

```
http://yyy.mycompany.com:8080/SASSearchService/?forward=Search&se
archType=reports
```

The HTTP protocol should only be used if you are using none for the Authentication value. If you are using Basic Authentication, then it is preferable to use HTTPS (HTTP would also work but the credentials would be visible in the URL), such as:

```
https://yyy.mycompany.com:8443/SASSearchService/?forward=Search&searchT
ype=data
```

Authentication

The GSA supports multiple types of authentication for OneBox providers. At this time, only None and Basic are supported for the Search Interface to SAS OneBox provider.

None - No Authentication is required to use this module. Thus, when a user does a search with only "Public Content" selected, this module will be called. If you choose this option, then Search Interface to SAS Content provides search results for user SAS Web Anonymous user. If you want to change the user to some other user, then you will have to append the username and password in the parameter as below:

```
http://yyy.mycompany.com:8080/SASSearchService/?forward=Search&searc
hType=reports&userName=sasguest@saspw&password=xxxxxxx
```

Note: *If you provide the user name and password in the URL, all users will have the access to see all the results accessible for this user.*

Basic – The GSA will prompt the user for a user ID and password and will pass these values to the provider. When a user is doing a search, the user must have selected "Public and Secure content" to be searched for this module to be invoked. The authentication scheme allows for the results to be limited to just what that individual user is authorized to access.

After you complete the Security settings, click **Save OneBox Definition**. The main page then appears, and you can find your newly added module in the list. If you want to change/edit the 'OneBox

Stylesheet Template', click on the **Edit** link that corresponds to the module listing. At the bottom of the configuration page, you will see the **Edit XSL** link.

Note: *When using Basic Authentication, Search Interface to SAS Content should use the HTTPS protocol to protect the user ID and password that is being passed to the provider.*

OneBox Stylesheet Template

When the results are returned from the module to the GSA, the GSA can apply an XSL style sheet to the results to format them for display. You should provide a style sheet to make the user experience more robust than is provided by the GSA's default XSL. To edit the XSL, use the **Edit XSL** link at the bottom of this area of the administrative page. The following is a sample XSL:

```
<xsl:template name="results">
 <xsl:variable name="searchLink"><xsl:value-of
select="title/urlLink"/></xsl:variable>
    <table border="0">
      <tr>
      <td>
      <xsl:if test="//IMAGE_SOURCE">
      <xsl:variable name="imageSource"><xsl:value-of
select="//IMAGE_SOURCE[1]"/></xsl:variable>
      <xsl:variable name="imageData">
      <xsl:for-each select="MODULE_RESULT">
            <xsl:choose>
            <xsl:when test="Field[@name='URI'] = $imageSource">
            <xsl:variable name="link"><xsl:value-of
select="U"/></xsl:variable>
            <xsl:choose>
                  <xsl:when test="$link=''">
                  ||||||||<xsl:value-of select="Field[@name='Name']"/>
                  </xsl:when>
                  <xsl:otherwise>
                  ||||<xsl:value-of select="$link"/>||||<xsl:value-of
select="Field[@name='Name']"/>
                  </xsl:otherwise>
            </xsl:choose>
            </xsl:when>
            <xsl:otherwise></xsl:otherwise>
            </xsl:choose>
      </xsl:for-each>
      </xsl:variable>
      <xsl:choose>
      <xsl:when test="$imageData!=''">
            <xsl:choose>
                  <xsl:when test="starts-with(normalize-
space($imageData),'||||||||')">
                        <xsl:variable name="title">
                              <xsl:choose>
                                    <xsl:when test="contains(substring-
after($imageData,'||||||||'),'||||')">
                                     <xsl:value-of select="substring-
before(substring-after($imageData,'||||||||'),'||||')"/>
                                    </xsl:when>
                                    <xsl:otherwise>
```

```
                                    <xsl:value-of select="substring-
after($imageData,'||||||||')"/>
                                    </xsl:otherwise>
                              </xsl:choose>
                        </xsl:variable>
                  <xsl:element name="img">
                              <xsl:attribute name="src"><xsl:value-of
select="$imageSource"/></xsl:attribute>
                              <xsl:attribute name="alt"><xsl:value-of
select="$title"/></xsl:attribute>
                              <xsl:attribute name="title"><xsl:value-of
select="$title"/></xsl:attribute>
                        </xsl:element>
                  </xsl:when>
                  <xsl:otherwise>
                        <xsl:variable name="link">
                        <xsl:value-of select="substring-before(substring-
after($imageData,'||||'),'||||')"/>
                        </xsl:variable>
                        <xsl:variable name="title">
                        <xsl:choose>
                              <xsl:when test="contains(substring-
after(substring-after($imageData,$link),'||||'),'||||')">
                                    <xsl:value-of select="substring-
before(substring-after(substring-
after($imageData,$link),'||||'),'||||')"/>
                              </xsl:when>
                              <xsl:otherwise>
                                    <xsl:value-of select="substring-
after(substring-after($imageData,$link),'||||')"/>
                              </xsl:otherwise>
                        </xsl:choose>
                        </xsl:variable>
                        <a><xsl:attribute name="href">
                              <xsl:value-of select="$link"/>
                              </xsl:attribute><xsl:attribute
name="target">_parent</xsl:attribute>
                              <xsl:element name="img">
                                          <xsl:attribute
name="src"><xsl:value-of select="$imageSource"/></xsl:attribute>
                                          <xsl:attribute
name="alt"><xsl:value-of select="$title"/></xsl:attribute>
                                          <xsl:attribute
name="title"><xsl:value-of select="$title"/></xsl:attribute>
                                          </xsl:element>
                        </a>
                  </xsl:otherwise>
            </xsl:choose>
            </xsl:when>
      <xsl:otherwise>
            <xsl:element name="img">
                  <xsl:attribute name="src"><xsl:value-of
select="$imageSource"/></xsl:attribute>
                  <xsl:attribute name="alt">Powered by SAS</xsl:attribute>
                  <xsl:attribute name="title">Powered by
SAS</xsl:attribute>
            </xsl:element>
```

```
      </xsl:otherwise>
      </xsl:choose>
      </xsl:if>
      </td>
                  <td>
               <!-- The oneBox style guide says to only show 3 responses
at most -->
             <xsl:for-each select="MODULE_RESULT[position() &lt;=3]">
             <xsl:variable name="Type"><xsl:value-of
select="Field[@name='Type']"/></xsl:variable>
             <xsl:choose>
                  <xsl:when test="$Type!='Document'">
                  <a><xsl:attribute name="href"><xsl:value-of
select="U"/></xsl:attribute><xsl:attribute
name="target">_parent</xsl:attribute><b> <xsl:value-of
select="Field[@name='Name']"/></b></a><br/>
                  <!-- If there is a RELATED field, then add the Similar
Reports hyperlink -->

                  <font size="-1"><xsl:value-of
select="Field[@name='Desc']"/>
                  <xsl:if test="Field[@name='RELATED']">
                  <a class="fl"><xsl:attribute name="href"><xsl:value-of
select="Field[@name='RELATED']"/></xsl:attribute><xsl:attribute
name="target">_parent</xsl:attribute> <xsl:value-of
select="Field[@name='SimilarLabel']"/></a>
                  </xsl:if>

                  </font>

                  <br/>
            </xsl:when>
            <xsl:otherwise>
            </xsl:otherwise>
            </xsl:choose>
            </xsl:for-each>
            <xsl:if test="count(MODULE_RESULT) > 3">
                        <!-- Since there are more results than can fit on
one OneBox result, issue the same query and show all the results -->
                <!-- The search criteria is the same in all
MODULE_RESULTS, so just get it from the first one -->
                  <xsl:if test="$searchLink != ''">
                  <font size="-1"><a><xsl:attribute name="href">
                  <xsl:value-of
select="$searchLink"/></xsl:attribute><xsl:attribute
name="target">_parent</xsl:attribute><xsl:value-of
select="MODULE_RESULT[1]/Field[@name='MoreLabel']"/></a></font>
                  <br/>
                  </xsl:if>
        </xsl:if>
        </td>
      </tr>
</table>
</xsl:template>
```

Note: *There are limitations imposed by the GSA on the XSL used here. For more information, see the Google Enterprise Web site,* www.google.com/Enterprise.

Additional Authentication Steps

If you used a value of Basic Authentication for the Authentication option, you will need to perform an additional step in setting up a Google One-Box to use Basic Authentication.

In the Google Admin UI, go to **Crawl and Index→Crawler Access**.

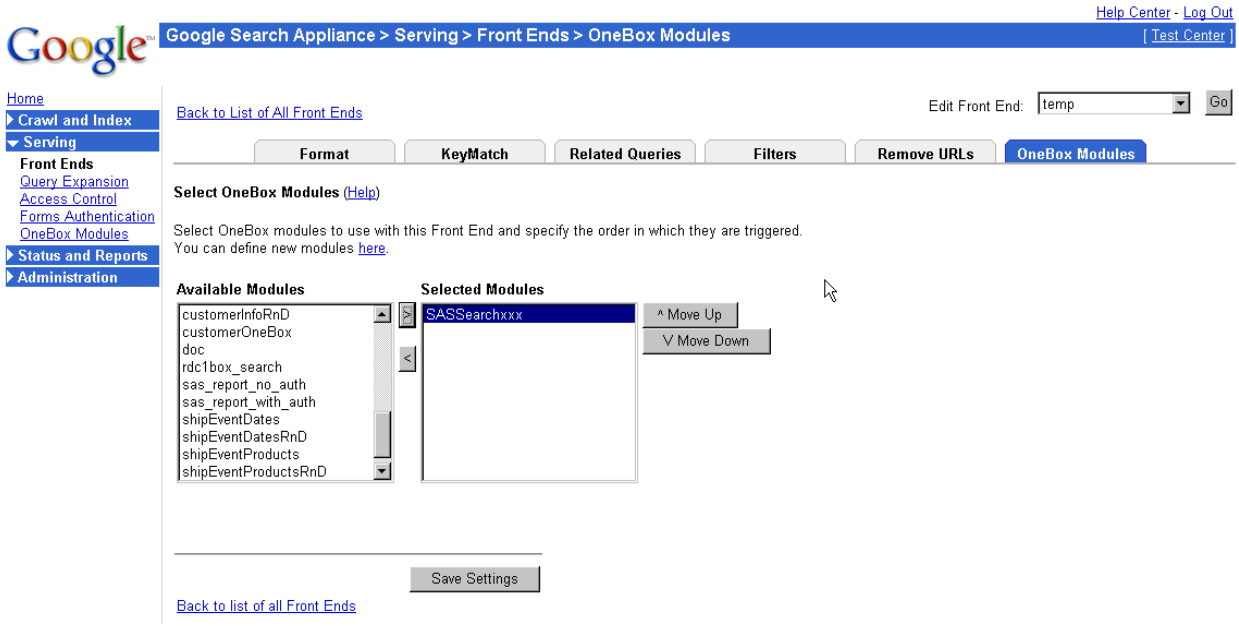
Make sure that the URL pattern of the One Box provider resides there, with a user ID and the "public" checkbox NOT checked. For example, if the web application is deployed to `http://yyy.mycompany.com:8080/SASSearchService`, then this same string needs to be used in the URL pattern line.

Add the OneBox Module Definition to a Front End

For the OneBox module to be called when a user does a query, the OneBox module must be attached to a Front End (for more information on what a Front End is, see `www.google.com/Enterprise`).

To add a OneBox module to a Front End, follow these steps:

1. Go to Google Admin UI→**Serving→Front Ends**.
2. Select the front end by clicking the edit link in front of it.
3. Go to the **OneBox Modules** tab and select your OneBox Module from the list.
4. If required, the sequence of OneBox modules can be changed by using the **Move Up** or **Move Down** buttons.
5. Click the **Save Settings** button to implement the changes.



Chapter 4 — Registering Search Interface to SAS Content as a Federated Location in Microsoft SharePoint Search Server

Search Interface to SAS Content can be integrated with Microsoft SharePoint Search Server as a federated search location. All SharePoint Search products that support integrating federated search locations can be integrated with Search Interface to SAS Content.

The sections below describe the steps required to configure Search Interface to SAS Content as a federated location. There are two steps to configure Search Interface to SAS Content with Microsoft SharePoint Search Server:

1. Import and modify the federated location.
2. Add the federated location WebPart with the search results.

Importing and Modifying the Federated Location

As a first step to integrate Search Interface to SAS Content with Microsoft SharePoint Search Server, Search Interface to SAS Content has to be registered as a federated location with Microsoft SharePoint Search Server. To ease the registration process, a Federated Location Definition (FLD) file is included with the installer, and you can import this FLD file and modify the necessary configuration.

To import and modify the federated location, follow the steps below.

Note: *Currently, importing the same FLD file provided with the installer to add more than one location is not supported. The workaround is to import once and create copies of the same, and make necessary changes to add more than one location.*

1. Open the Search Server administration page.
2. At the left side of the page under the **Queries and Results** section, click on the **Federated Locations** link.
3. On the **Manage Federated Location** page, click on the **Import Location** link.
4. On the **Import Federated Location** page, click the **Browse** button to import the Federated Location Definition File.
5. Point to the `SearchInterfaceToSASContent.FLD` file that resides under the `MICROSOFT` subdirectory of your Search Interface to SAS Content home directory. You can locate the `SearchInterfaceToSASContent.FLD` file in the following default locations:

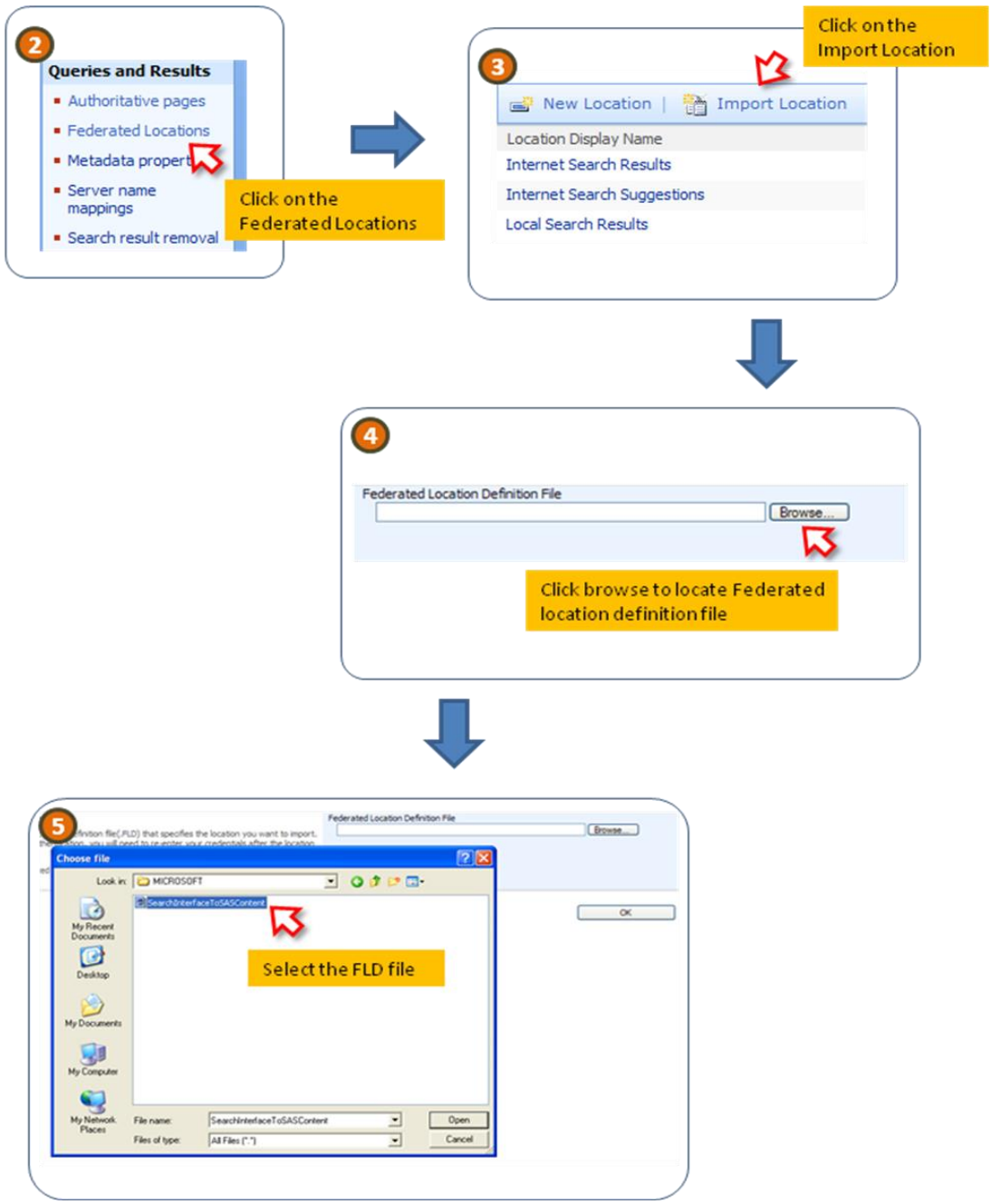
   - For Windows:

     ```
     C:\Program Files\SAS\SearchInterfaceToSASContent\3.2\MICROSOFT\
     SearchInterfaceToSASContent.FLD
     ```

   - For UNIX:

     ```
     /usr/local/SAS/SearchInterfaceToSASContent/3.2/MICROSOFT/
     SearchInterfaceToSASContent.FLD
     ```

Steps 2 to 5 are illustrated in the following diagram.



6. After the FLD file is imported successfully, you receive a success message:

   The location "Search Interface to SAS 3.2" was successfully imported.

   Click **OK** to proceed.

7. After the FLD file is imported successfully, click the **Edit Location** button.

8. In the **Edit Federated Location** page, expand the **Location Information** Section.

   By default the Federated Location will be registered for Report Search. For report search, the Query template will be as follows:

```
http://<host>:<port>/SASSearchService/?forward=Search&query={sear
chTerms}&searchType=reports&authType=none
```

If you want to register for Data search (for searching information maps), modify the searchType parameter value to 'data'.

After the change, the URL will appear this way:

```
http://<host>:<port>//SASSearchService/?forward=Search&query={sea
rchTerms}&searchType=data&authType=none
```

9. Change the host name to the server name where Search Interface to SAS Content is deployed. For example if you deploy the Search Interface to SAS Content in a server having the host name *yyy.mycompany.com* and port *7001*, then the Query template will be as follows:

   For reports:

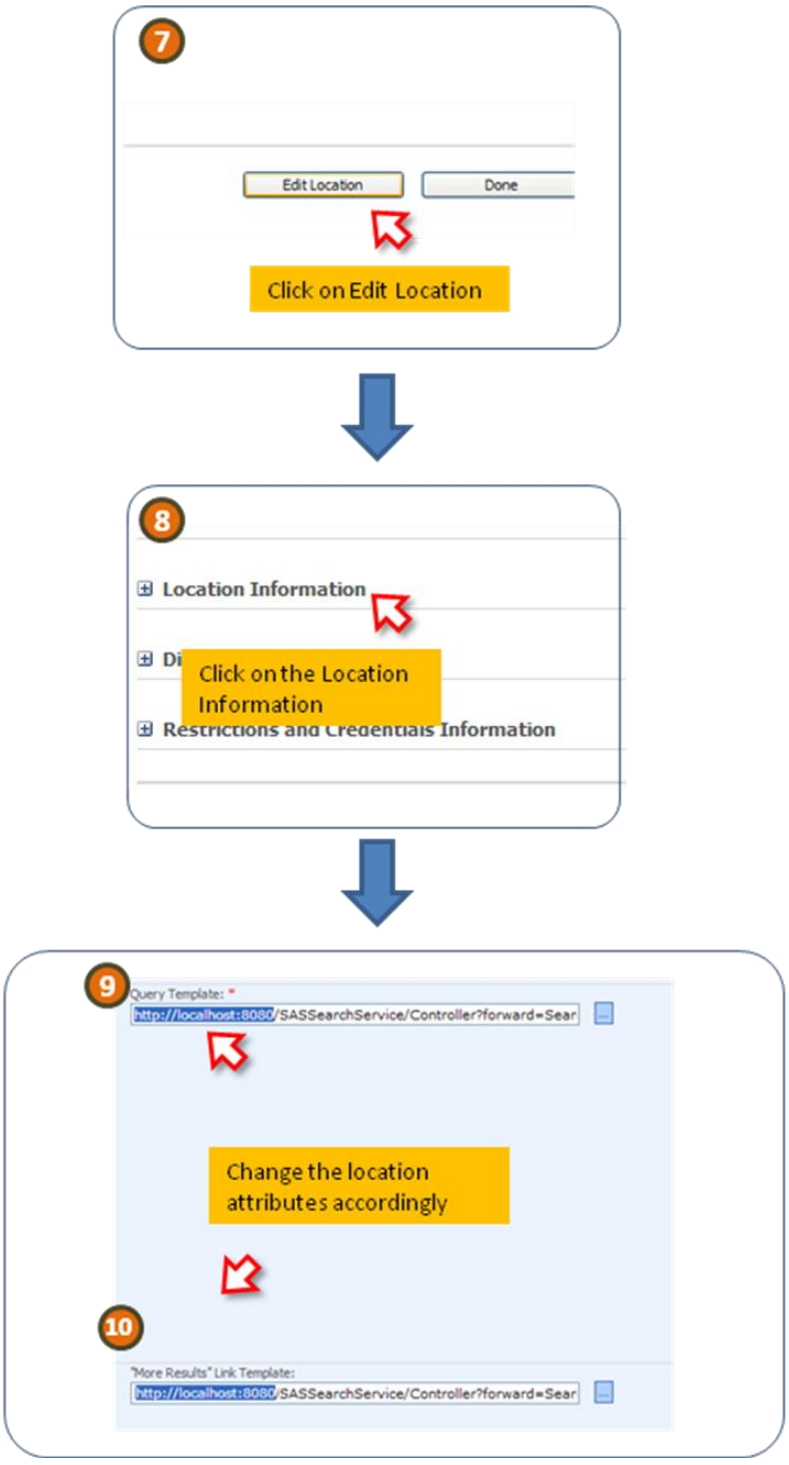
   ```
   http://yyy.mycompany.com:7001/SASSearchService/?forward=Search&qu
   ery={searchTerms}&searchType=reports&authType=none
   ```

   For data:

   ```
   http://xxx.mycompany.com:7001/SASSearchService/?forward=Search&qu
   ery={searchTerms}&searchType=data&authType=none
   ```

10. Change the host name in the "More Results" Link Template also as above.

Steps 7 to 10 are illustrated in the following diagram.



11. Expand the **Restrictions and Credentials Information** section.
12. Modify the access credential.

    By default, the access credentials for this location under **Specify Credentials** section will be **anonymous**. If the access credential is specified as anonymous, then Search Interface to SAS

Content provides search results for user SAS Web Anonymous user. If you want to change the user to some other user, then you will have to modify the query template by removing the **authType=none** parameter and append the username and password.

For example, change the Query Template in the **Location Information** as below:

```
http://yyy.mycompany.com:7001/SASSearchService/?forward=Search&query={searchTerms}&searchType=reports&userName=sasguest@saspw&password=xxxxxxx
```

If you specify the **Basic Authentication – Specify a user name and password** under the **Common** section, you will have to change the **authType** parameter in the Query Template in the Location information as below:

```
http://<host>:<port>//SASSearchService/?forward=Search&query={searchTerms}&searchType=reports&authType=basic
```

and for the account and password section for Basic Authentication under Common section specify valid SAS credentials which will be used to during search.

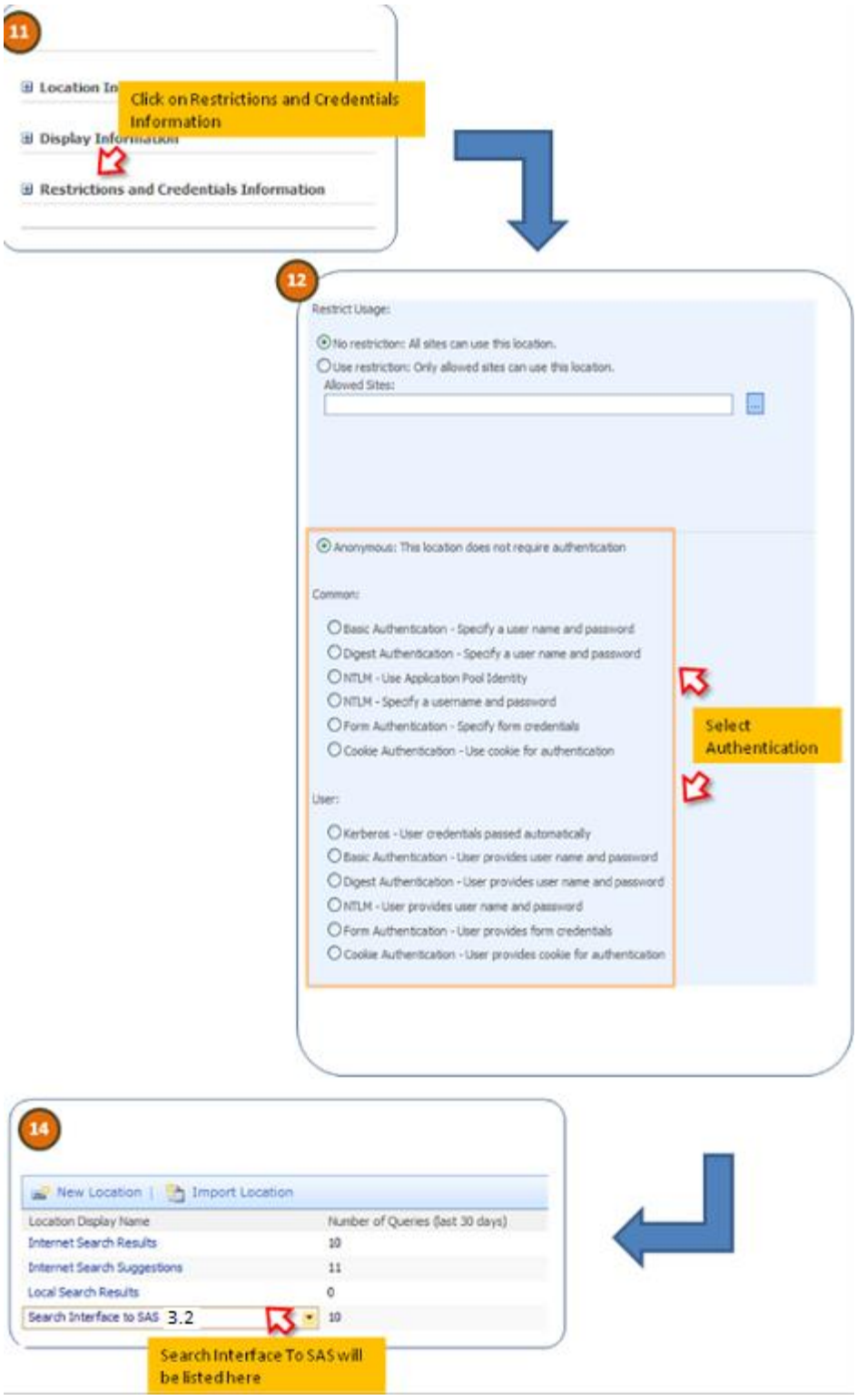
If you specify the **Basic Authentication – User provides user name and password** under the **User** section, you will have to create a **Custom Federated Search Web Part with Credentials UI.**

For more information please visit the MDSN website.

Note: *Search Interface to SAS Content is not tested with access credentials options other than Anonymous access and Basic Authentication – Specify a user name and password.*

13. Click **OK** to save the changes. If the configuration is successful, the "Manage Federated Locations" page will be displayed. If not successful, modify the necessary changes and continue.
14. In the list of Federated Locations, "Search Interface to SAS 3.2" also will be listed.

Steps 11, 12, and 14 are illustrated in the following diagram.

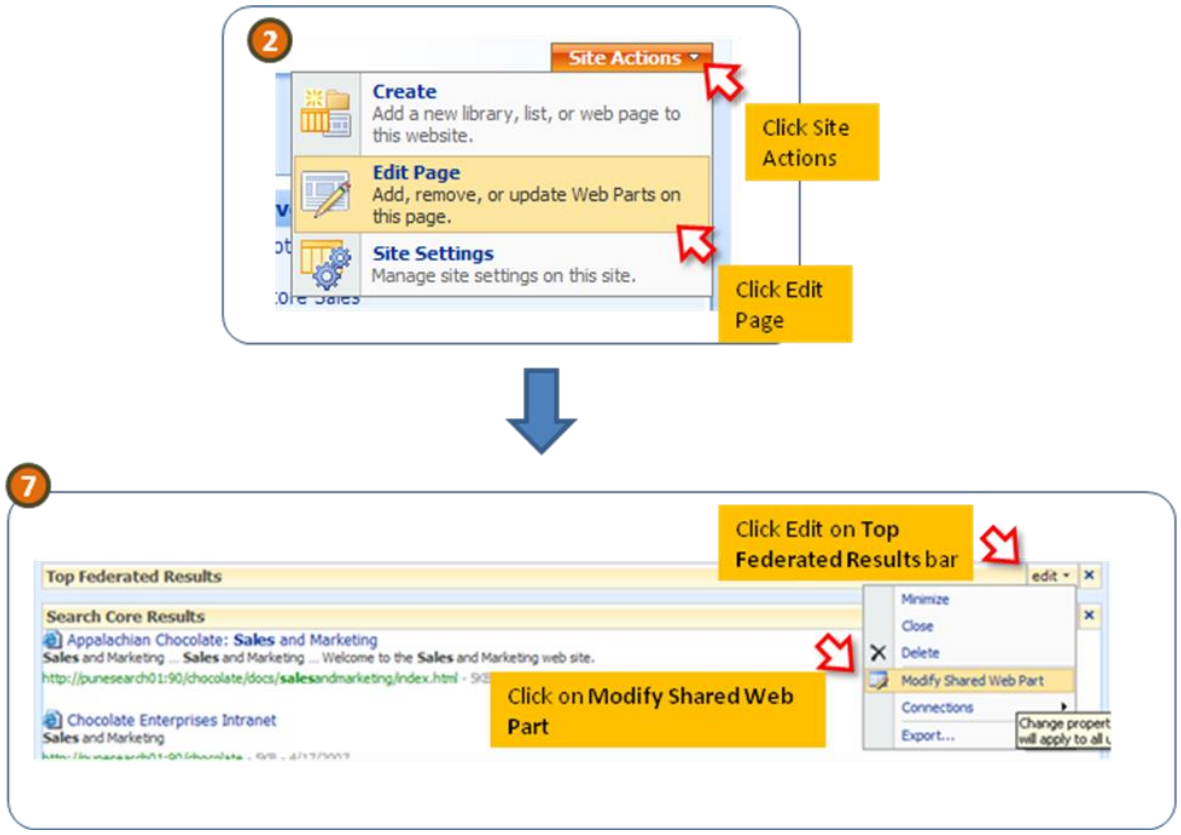

11
Location In
Click on Restrictions and Credentials Information
Display Information
Restrictions and Credentials Information
12
Restrict Usage:
No restriction: All sites can use this location.
Use restriction: Only allowed sites can use this location.
Allowed Sites:
Anonymous: This location does not require authentication
Common:
Basic Authentication - Specify a user name and password
Digest Authentication - Specify a user name and password
NTLM - Use Application Pool Identity
NTLM - Specify a username and password
Form Authentication - Specify form credentials
Cookie Authentication - Use cookie for authentication
Select Authentication
User:
Kerberos - User credentials passed automatically
Basic Authentication - User provides user name and password
Digest Authentication - User provides user name and password
NTLM - User provides user name and password
Form Authentication - User provides form credentials
Cookie Authentication - User provides cookie for authentication
14
New Location | Import Location
Location Display Name
Number of Queries (last 30 days)
Internet Search Results 10
Internet Search Suggestions 11
Local Search Results 0
Search Interface to SAS 3.2 10
Search Interface To SAS will be listed here

Adding the Federated Location WebPart with the Search Results

To include the Search Result from Search Interface to Content with your default search result, you will have to add the Search Interface to SAS Federated Location in to a Top Federated Results WebPart. Follow the steps below to perform this task:

1. Go to the Search Results page by first opening the search home page, and then searching for any string. This will take you to the result page.
2. Click on the **Site Actions** link in the top right corner of the search page. Click on the **Edit page** menu item from the expanded menu.
3. Check whether "Top Federated Results" WebPart is already available on the page. If it is available, proceed to step 7. If it is not available continue with the next step.
4. Click on the **Add a Web Part** link on the Bottom Zone.
5. A popup window containing the list of WebParts opens.
6. Expand all WebPart sections and select the **Top Federated Results** WebPart under the **Search** sub-section. Click **Add** to continue.
7. Click on the **Edit** link provided in the **Top Federated Results** WebPart. Click on the **Modify Shared WebPart** from the expanded menu.

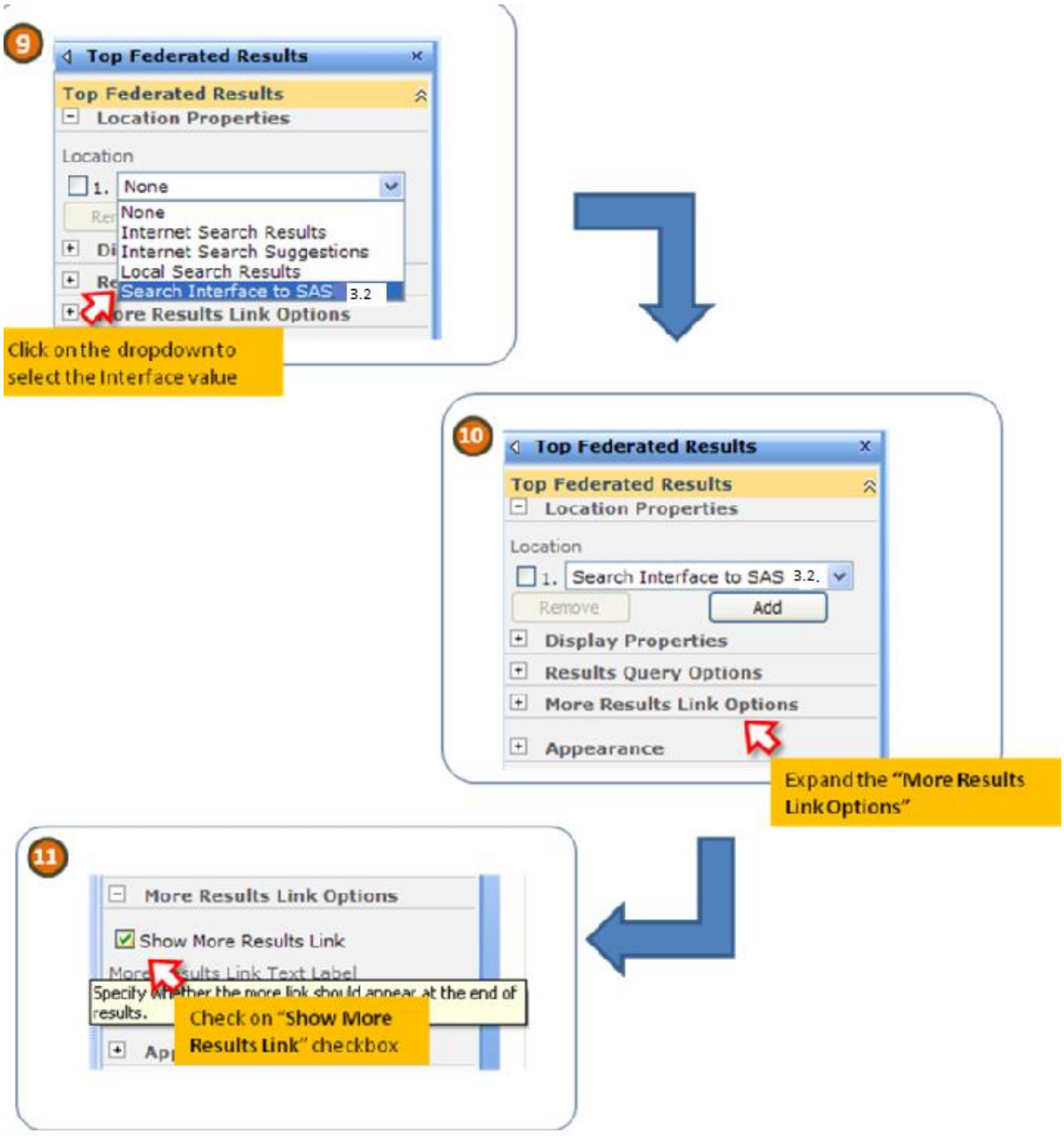
Steps 2 and 7 for adding Federated Locations in the WebPart are illustrated in the following diagram.



8. The property of Shared WebPart will be displayed in the right-hand side of the page.
9. From the Location List, Select **Search Interface to SAS 3.2**.

10. Expand the More Results Link Options.
11. Determine if the **Show More Results Link** option is selected. If not, select it.

Steps 9 to 11 for adding Federated Locations in the WebParts are illustrated in the following diagram.



12. Modify the Display Properties, if required.
13. Click **Apply** and then **OK** to continue.
14. Click **Exit Edit Mode** on the top right side of the page.

Chapter 5 - Logging for Search Interface to SAS Content and loadindex

Logging for Search Interface to SAS Content

Logging has been enabled for Search Interface to SAS Content. The log4j root level is set at *error* which can be changed to *warn, info* as required. The configuration file for Search Interface to SAS Content can be found at this location:

SAS-Configuration-Directory/**Lev***n*/**Web/Common**/LogConfig/SASSearchService-log4j.xml

Logs are generated at this location:

SAS-Configuration-Directory/**Lev***n*/**Web/Logs**

Logging for loadindex

Logging has been enabled for loadindex program. The log4j root level is set as *error* which can be changed to *warn, info* as required. The configuration file for loadindex program can be found at location:

SAS-Configuration-Directory/**Lev***n*/**Web/Applications/SearchInterfacetoSASContent/SASSearchServiceClient-log4j.xml.**

Logs are generated at location:

SAS-Configuration-Directory/**Lev***n*/**Web/Applications/SearchInterfacetoSASContent/Logs**

Appendix — Troubleshooting

How do I validate that Search Interface to SAS Content has been installed properly?

When Search Interface to SAS Content has been installed and configured, the REST-based "Live" search URL can be used to search the query string in the metadata directly. This query does not require any other software component to be installed. You will need to provide appropriate host, port, and credential information in the URL:

```
http://<host name>:<port>/SASSearchService/rest/content/liveSearch/?query=
Sample&userName=<user name>&password=<password>
```

If you are able to get some response, it means that Search Interface to SAS Content has been installed correctly and its REST interface is working.

How do I validate that the integration between Search Interface to SAS Content and SAS Information Retrieval Studio is working?

Two conditions must be met in order for the integration between Search Interface to SAS Content and SAS Information Retrieval Studio to provide results in response to a search query:

- SAS Information Retrieval Studio is configured at runtime by Search Interface to SAS Content when its first URL is hit either for search or for feed extract. The configuration may be validated at the admin page of SAS Information Retrieval Studio as described below.
- The index of SAS Information Retrieval Studio must be populated with the Feed Extract from the Search Interface to SAS Content. The loadIndex script must be executed before searching any content in the index in order to ensure that the data has been fed and that SAS Information Retrieval Studio can return a response from them.

You can test the integration with the following URL, including the appropriate host, port, and credential information:

```
http://<hostname>:<port>/SASSearchService/rest/search/content/?q=Sample
&userName=<UserName>&password=<password>&sourceIndex=teragramIndex
```

I do not see any result when I'm trying to search a query string in the HUB. What can I do?

First you need to validate that Search Interface to SAS Content and SAS Information Retrieval Studio have been appropriately installed and configured.

When you have validated that both are installed and configured, the only remaining issue could be the query string, which may not have any matching record for the user (keeping in mind the authorization constraints). Try using `'*'` as the query string, which should return some results.

How can I validate that SAS Information Retrieval Studio Server is installed correctly?

To verify that the SAS Information Retrieval Studio is installed correctly, go to the Web admin client of SAS Information Retrieval Studio and check the server status. Use the following test with the appropriate host and port:

```
http://<host name>:<port>/
```

If the default values of the installation have not been altered, the default port for the admin client of SAS Information Retrieval Studio is 10651.

The following servers must be listed as up and running on the admin page that opens:

- Proxy Server
- Pipeline Server
- Indexing Server
- Query Server

The servers listed above are required for the Search applications to work. If any of them are missing from the admin page, this might be due to a license issue. Please contact SAS Technical Support.

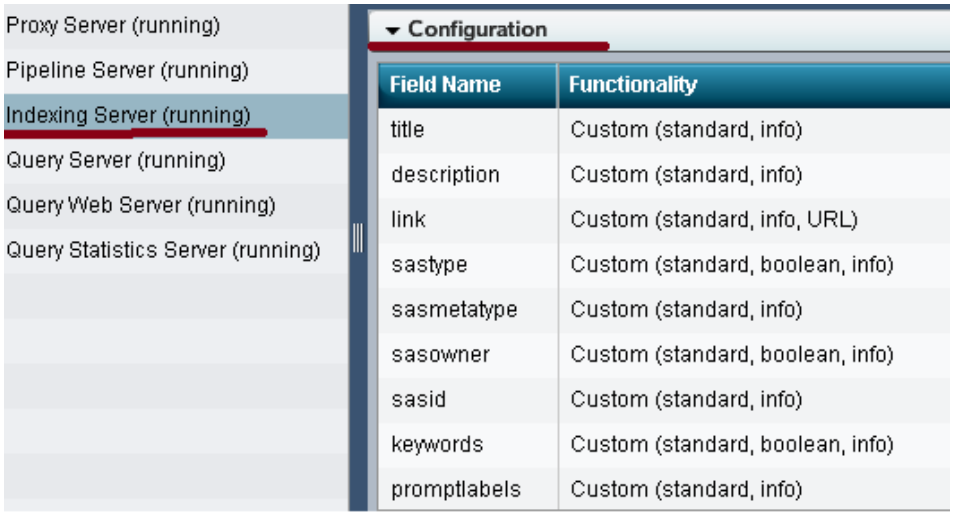
If a server is not listed as running, start the server by clicking on the **Start** button on the top right side of that server's details page. Each of these servers must be in running mode.

If each of these servers is listed and in running mode, SAS Information Retrieval Studio has been appropriately installed.

How can I validate that SAS Information Retrieval Studio has been configured correctly?

When you have validated that SAS Information Retrieval Studio is installed correctly, and that the search application has been installed correctly (done by using the REST URL for search, which also silently configures the SAS Information Retrieval Studio for Search via BI5.1 client), verify the SAS Information Retrieval Studio configuration by checking the following elements in the **Configuration** section of the Indexing Server:

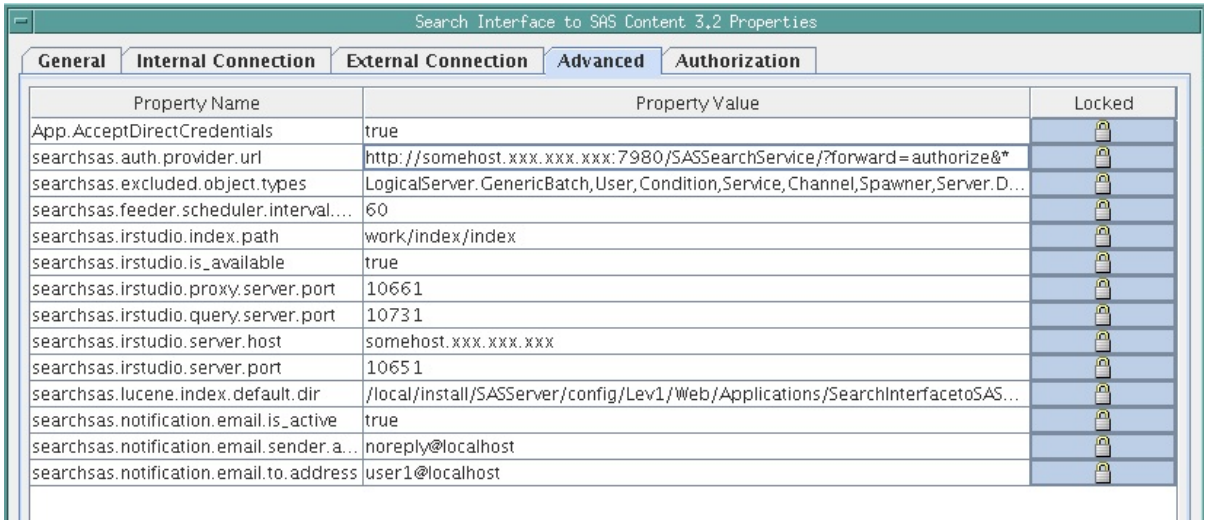
- title
- description
- link
- sastype
- sasmetatype
- sasowner
- sasid
- keywords
- promptlabels



The elements shown in the **Configuration** section of the image must be present in the **Configuration** section of the target set-up. This ensures that the configuration of SAS Information Retrieval Studio was invoked by Search Interface to SAS Content.

How can I verify the host and port of SAS Information Retrieval Studio? What port value should I use for LoadIndex script?

The host and port of SAS Information Retrieval Studio for the admin page and for the IndexLoader script can be taken from the properties of the metadata component of Search Interface to SAS Content.



searchsas.irstudio.server.host: The host name of the machine on which SAS Information Retrieval Studio servers are running.

searchsas.irstudio.server.port: The port number of the machine on which SAS Information Retrieval Studio's Web admin client is running.

searchsas.irstudio.proxy.server.port: The port number on which we push the feed in SAS Information Retrieval Studio. We use this port for the LoadIndex script.

searchsas.irstudio.query.server.port: The port number on which query to Index are sent.

While running the LoadIndex script to push feed, I'm getting an OutOfMemoryError. How should I proceed?

The heap size of the JVM running LoadIndex program may be increased by increasing the value to 1 GB.

To increase the heap size, edit the `loadindex.ini` (which resides in the same folder as the script) in order to add the minimum and maximum memory setting.

Change `JavaArgs_1=-Xmx512m` to `JavaArgs_1=-Xmx1024m` and add `JavaArgs_8=-Xms1024m`

If the issue is not resolved with 1 GB, you can use 2 GB. If the issue persists, check the available free memory on the machine. It is possible that the machine does not have enough memory available. In that case, restarting JBOSS or the machine should help.

What should I do if while running the LoadIndex script, I get a message similar to the one below?

```
Please check that the host/port are correct and the Server is running
Message = Connection refused : status = SOCKET_PROTOCOL_ERROR
```

As the message itself says, either the host and port provided in the LoadIndex command are not correct or the proxy server of SAS Information Retrieval Studio is down. Check the server status and start it if necessary. If you provided an incorrect port, correct it by verifying the port against the metadata. The following is an example command to make this verification:

```
./loadindex.sh -sassearchserverhost <host name> –sassearchserverport
<port>
```

Also, in the same location where this executable resides, you will find a file `url_list.txt`. Edit this file to include the following feed URL including the appropriate host, port, and credential information:

```
http://<host name>:<port>/SASSearchService/Controller?forward=search
&userName=<user name>&password=<password>&searchClient=feeder
```

I am getting automated emails alerting "Search Interface to SAS Content encountered error while feeding the index to index server." What does this mean?

The error message says-
Error Message: com.sas.svcs.search.client.ConcurrentIndexerRunningException: An indexing process is already running. Cannot start a concurrent process. Check the logs for more details.

You are getting this email because the scheduled job 'loadindex' is having a problem during execution. As the message says, a flow for 'loadindex' is already in progress so it will not be able to execute another request. It should be noted that we support 'Incremental feeds' and at any point in time, only a single feeding flow can execute. This is done to ensure that due to multiple running processes, indexing flow does not consumes too much of CPU & RAM. The incremental flow guarantees that no record would be left out even if some records were updated while the flow was running.

I am getting Error message on running 'loadIndex' script like this:
Fetching data from 'Search Interface to SAS Content' failed. Invalid User Credentials provided for 'Search Interface to SAS Content'. An error occurred while pushing feed.

There are two possible reasons for this error:

- As the message itself says, the user credentials are not correct. Please check if the Metadata user account is not locked. If it is locked, please unlock it.
- If the machine was initially deployed on one HOST (ABC.sas.com) and then externalized to another (XYZ.sas.com) via updating it in Metadata, then you also need to update the HOST in the file '`url_list.txt`' which exists at the same location as 'loadindex' script '`\Config\Lev1\Web\Applications\SearchInterfacetoSASContent\`'.

One needs to check both of these cases.

What changes are required in the configuration for reverse proxy environment?

The external host name should be used at all the appropriate places in configuration artifacts. Please make sure that the external host is stored at following places:

- In SAS Management Console, go to **Application Management > Configuration Manager** and open the software component with the name "Search Interface to SAS Content (Version)" and check the property `searchsas.auth.provider.url` in the **Advanced** tab. Ensure that the HOST in the URL is that of 'External Host'.
- Modify the `url_list.txt` present at '`\Config\Lev1\Web\Applications\SearchInterfacetoSASContent\`' so that it points to the external host.
- Stop the IR Studio. Once it is stopped, modify the `query-server.db` file present at `config/Lev1/Applications/SASInformationRetrievalStudioforSAS1.51/work` so that the URL for 'auth provider' points to the external host.

***Note**: Due to limitation of 'IR Studio' supporting only 'HTTP Protocol' for searching, the Authorization Provider URL must have HTTP-based protocol and not HTTPS. If the environment supports only HTTPS (hard redirect from http to https. http port closed.) then Search would not work in those cases.*

How do I recover lost data for the Indexing Server after an Indexing Failure?

It is possible for the 'Load Index' process to throw an exception. In earlier releases, if something failed, we completely aborted the feeding process but now we would continue with the feeding process for rest of the data.

Solution: After the issue has been identified and resolved, one needs to push the lost data to index as well (the data that could not be indexed due to the exception and is therefore not searchable). In that case, the 'IR Studio' index would have to be deleted and the 'Web App Server' restarted. This would initiate the 'FULL-FEED' to capture any data that was lost before due to failure.

Retrieved from
http://sww.sas.com/saspediawiki/index.php?title=SAS_Search_Service_FAQ&oldid=823560

Technical Support

If you need assistance with the software, we ask that only SAS support personnel call our Technical Support Division.

- For U.S. and Canadian customers, support is provided from our corporate headquarters in Cary, North Carolina. You may call (919) 677-8008, Monday through Friday.
- Customers outside of the U.S. can obtain local-language technical support through the local office in their countries. Customers in these locations should contact their local office for specific support hours. See http://support.sas.com/techsup/contact/index.html for contact information for local offices.

Before you call, you may want to explore the SAS Support Web site at http://support.sas.com/techsup/

This site offers access to the SAS Knowledge Base, as well as discussion forums, Technical Support contact options, and other support materials that may answer your questions.

SAS is the leader in business analytics software and services, and the largest independent vendor in the business intelligence market. Through innovative solutions delivered within an integrated framework, SAS helps customers at more than 50,000 sites improve performance and deliver value by making better decisions faster. Since 1976, SAS has been giving customers around the world THE POWER TO KNOW®.