

# Installing Retrieval and Classification (TREX) in a Non- Portal Environment



**Release TREX 6.0 SP1**



## Copyright

© Copyright 2003 SAP AG. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or for any purpose without the express permission of SAP AG. The information contained herein may be changed without prior notice.

Some software products marketed by SAP AG and its distributors contain proprietary software components of other software vendors.

Microsoft®, WINDOWS®, NT®, EXCEL®, Word®, PowerPoint® and SQL Server® are registered trademarks of Microsoft Corporation.

IBM®, DB2®, DB2 Universal Database, OS/2®, Parallel Sysplex®, MVS/ESA, AIX®, S/390®, AS/400®, OS/390®, OS/400®, iSeries, pSeries, xSeries, zSeries, z/OS, AFP, Intelligent Miner, WebSphere®, Netfinity®, Tivoli®, Informix and Informix® Dynamic Server™ are trademarks of IBM Corporation in USA and/or other countries.

ORACLE® is a registered trademark of ORACLE Corporation.

UNIX®, X/Open®, OSF/1®, and Motif® are registered trademarks of the Open Group.

Citrix®, the Citrix logo, ICA®, Program Neighborhood®, MetaFrame®, WinFrame®, VideoFrame®, MultiWin® and other Citrix product names referenced herein are trademarks of Citrix Systems, Inc.

HTML, DHTML, XML, XHTML are trademarks or registered trademarks of W3C®, World Wide Web Consortium, Massachusetts Institute of Technology.






JAVA® is a registered trademark of Sun Microsystems, Inc.

JAVASCRIPT® is a registered trademark of Sun Microsystems, Inc., used under license for technology invented and implemented by Netscape.

MarketSet and Enterprise Buyer are jointly owned trademarks of SAP AG and Commerce One. SAP, SAP Logo, R/2, R/3, mySAP, mySAP.com and other SAP products and services mentioned herein as well as their respective logos are trademarks or registered trademarks of SAP AG in Germany and in several other countries all over the world. All other product and service names mentioned are trademarks of their respective companies.

Regular expression support is provided by the PCRE library package, which is open source software, written by Philip Hazel, and copyright by the University of Cambridge, England. FTP site for the source: <ftp://ftp.csx.cam.ac.uk/pub/software/programming/pcre/>

## Icons

Icon	Meaning
	Caution
	Example
	Note
	Recommendation
	Syntax

## Typographic Conventions

Type Style	Description
<i>Example text</i>	Words or characters that appear on the screen. These include field names, screen titles, pushbuttons as well as menu names, paths and options.  Cross-references to other documentation.
<b>Example text</b>	Emphasized words or phrases in body text, titles of graphics and tables.
EXAMPLE TEXT	Names of elements in the system. These include report names, program names, transaction codes, table names, and individual key words of a programming language, when surrounded by body text, for example, SELECT and INCLUDE.
Example text	Screen output. This includes file and directory names and their paths, messages, source code, names of variables and parameters as well as names of installation, upgrade and database tools.
EXAMPLE TEXT	Keys on the keyboard, for example, function keys (such as F2) or the ENTER key.
<b>Example text</b>	Exact user entry. These are words or characters that you enter in the system exactly as they appear in the documentation.
<Example text>	Variable user entry. Pointed brackets indicate that you replace these words and characters with appropriate entries.

## Contents

Installing Retrieval and Classification (Trex) .....	6
Installing Trex .....	8
Installation Process .....	10
Planning .....	10
Checking Prerequisites .....	11
Preparation .....	12
Installing SAP Gateway .....	12
Installation .....	14
Input Parameters for the Trex Installation .....	14
Installing Trex .....	17
Post Installation .....	18
Checking Performance Settings for the Operating System .....	18
Configuring SAP Gateway .....	19
Copying vpd.properties to the Trex Directory .....	19
Configuring the RFC Server .....	20
Checking the Configuration of the Trex Daemon .....	21
Optional: Activating Python Extensions .....	22
Checking Access to Documents .....	23
Creating an RFC Destination .....	24
Creating an RFC Destination with Activation Type 'Start' .....	24
Creating an RFC Destination with Activation Type 'Registration' .....	26
Creating a Search Server Relation .....	28
Testing a Search Server Relation .....	29
Scheduling Asynchronous Indexing and Deindexing .....	30
Installation Check .....	31
Checking Trex Processes .....	31
Executing an Installation Test Script .....	32
Additional Information .....	33
Troubleshooting .....	33
Checking SAP Gateway .....	34
Checking Trex .....	35
Checking for Problems in IMS Monitoring .....	36
Starting and Stopping Trex .....	37
Starting and Stopping the Trex Service .....	37

Starting and Stopping Individual TREX Servers.....	38
Starting and Stopping the TREX Daemon in Debug Mode .....	39
Deinstalling TREX .....	40



## Installing Retrieval and Classification (TREX)

### Purpose

This guide describes the installation of Retrieval and Classification (TREX) 6.0 SP1 in a non-portal environment.

The guide describes how to plan and carry out the installation of TREX, and explains the steps that are necessary after the installation in order to put TREX into operation. The target group of the guide consists of system administrators and consultants.

### Implementation Considerations

#### Operating System Platforms

Only the Windows version of TREX 6.0 SP1 has been released for non-portal environments.

#### Minimal TREX System

TREX offers a flexible architecture that can be modified to different requirements. A minimal TREX system consists of a single host that provides all TREX functions. You can use a minimal system as a demo and test system, or as a productive system.



For a productive system, we recommend that you install TREX on a single host that is used exclusively for TREX.



If you are installing TREX together with other components, you have to make sure that you have enough main memory space for all components. All TREX processes together can take up a **maximum** of 6 GB main memory. Make sure that TREX has exclusive use of the required main memory space.

#### SAP Notes on Installation



Read the SAP Notes on installation before you begin. The SAP Notes contain current installation information and corrections to the installation documentation.

Make sure that you use the current version of the SAP Notes. You can find the SAP Notes in the SAP Service Marketplace (<http://service.sap.com/notes>) or in SAPNet - R/3 Frontend.

#### Relevant SAP Notes

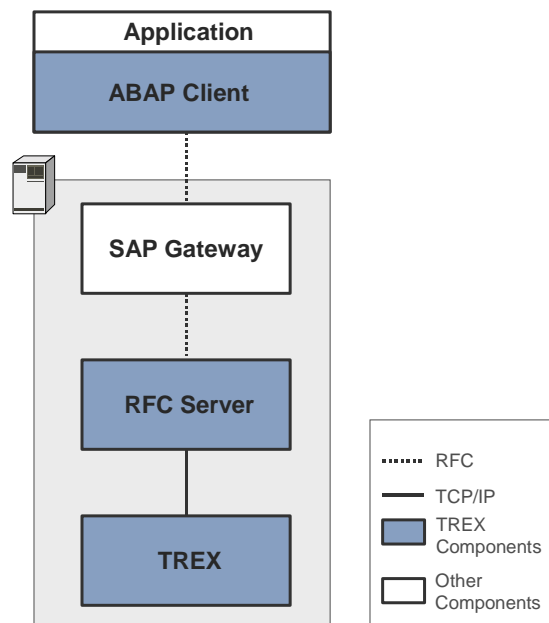
SAP Note Number	Title
565347	TREX 6.0: Central Note
631390	TREX 6.0: Additional Languages
179175	SAP R/3 Gateway (KPro Server Infrastructure Components)-Inst
0137478	Inst. Dialog/Gateway Instance -6.10, 4.6x, 4.5B, 3.11 SR1

SAP Note Number	Title
492305	INST: SAP J2EE Engine/Dialog Inst. /Gateway Inst. 6.20
532989	TREX 5.0 or 6.0: Installation Gateway 6.20
612272	TREX 6.0 Installation of SAP Gateway in non Portal scenarios
0080474	Automatic start of a gateway instance under NT

## Integration

SAP systems can use TREX over an ABAP interface. The TREX ABAP client is integrated into the application and allows the application to access the TREX functions.

The application using the TREX functions firstly sends searching and indexing requests to the TREX RFC server using an SAP gateway. The RFC server forwards the requests to the TREX components that are responsible for the actual processing. The TREX components then send their responses to the RFC server. The RFC server forwards the responses to the application using the gateway.





# Installing TREX

## Use

Use the tables below as a checklist for the installation of the system and to navigate through the installation procedure.

All necessary installation phases are listed in these tables.

Use the links to the general descriptions of actions and to additional information that will help you when executing the actions. By doing this, you ensure that you do not overlook any important information.



This guide also describes the configuration that takes place after the installation as long as this is relevant to general settings on the TREX host and for TREX itself.

Additional configuration steps might be necessary for each application that uses the TREX functions. For more information, see the documentation on the application or TREX implementation scenario in question.

## Prerequisites

We recommend that you proceed as follows when installing TREX.

1. Print out the tables below.
2. Carry out the installation in the order specified in the tables.
  - When carrying out an obligatory step during the installation, follow the link to the relevant section.
  - Then carry out the work steps described there.
  - When the installation step has been successfully completed, place a ✓ next to the relevant entry in the table in order to record your progress.
  - Then continue with the next step listed in the table.

## Planning

✓	Action
	<a href="#">Check prerequisites [Page 11]</a>

## Preparation


✓	Action
	<a href="#">Install SAP Gateway [Page 12]</a>
	Restart the host
	<a href="#">Check whether the gateway has been started [Page 34]</a> (process gwrd.exe)



## Installation

✓	<b>Action</b>
	<a href="#">Note input parameters for the TREX installation [Page 14]</a>
	<a href="#">Install TREX [Page 17]</a>

## Post Installation

✓	<b>Action</b>
	<b>Operating system</b>
	<a href="#">Check performance settings for the operating system [Page 18]</a>
	<b>SAP Gateway</b>
	<a href="#">Configure the gateway [Page 19]</a>
	<b>TREX</b>
	<a href="#">Copy the vpd.properties file to the TREX Directory [Page 19]</a>
	<a href="#">Configure the RFC server [Page 20]</a>
	Only if you want to create an RFC destination with the activation type <i>Registration</i> : <a href="#">Checking the Configuration of the TREX Daemon [Page 21]</a> .
	Optional: <a href="#">Activate the Python extensions [Page 22]</a>
	<p>We recommend that you restart the host. The central TREX service (the TREX daemon) starts automatically when the host is restarted.</p> <p>If you do not want to restart the host, you can start the TREX daemon using a prompt. Open a <b>new prompt</b>, go to the installation directory, and enter <b>TrexDaemon.exe</b>.</p> <div style="text-align: center;">  </div> <p>You have to use a new prompt (one opened after the setup). Otherwise, the required environment variables are not set, and the TREX daemon cannot start.</p> <p>The TREX daemon starts without notifying you. You can check whether it is running by displaying the processes in the Task Manager.</p> <p>You can minimize the prompt but not close it. If you close the prompt, the TREX daemon stops.</p>
	<a href="#">Check access to documents [Page 23]</a>
	<b>SAP system</b>
	Depending on the requirements or your application: <a href="#">Create an RFC Destination with Activation Type 'Start' [Page 24]</a> or <a href="#">Create an RFC Destination with Activation Type 'Registration' [Page 26]</a> .
	<a href="#">Create a search server relation [Page 28]</a>
	<a href="#">Test a search server relation [Page 29]</a>
	Only with KPro: <a href="#">Schedule asynchronous indexing and deindexing [Page 30]</a>



## Installation Process

### Purpose

The TREX installation process consists of four phases: Planning, preparation, installation, and post installation.



## Planning

### Purpose


During the planning phase, you check the hardware and software requirements, and prepare an installation host.



## Checking Prerequisites

### Procedure

Make sure that the installation host fulfills the following prerequisites:

Prerequisite Type	Prerequisite
Hardware Prerequisites	<ul style="list-style-type: none"> <li>• Hard disk capacity: <ul style="list-style-type: none"> <li>○ Installation directory: At least 600 MB for the TREX software</li> <li>○ Index directory: For the indexes, a minimum of 80 GB, depending on the number and type of documents to be indexed.  If documents exist in different formats (Microsoft Word, PDF, and so on), the index needs approximately half as much disk space as the documents. For pure HTML documents, the index needs about 1.5 times as much disk space as the documents.</li> <li>○ Queue directory: Approximately three quarters of the disk space required by the indexes. The documents to be indexed are kept temporarily in the queue directory before being forwarded to actually be indexed.</li> <li>○ Backup directory: Approximately 1.5 times as much disk space as required by the indexes. The backup directory is only relevant if you want to implement index replication or a backup/restore procedure. If this is the case, the index backups are stored in the backup directory.</li> </ul> </li> <li>• RAM: <ul style="list-style-type: none"> <li>○ A minimum of 1 GB (demo and test system)</li> <li>○ Recommended: 2 - 6 GB (productive system)</li> </ul> <div style="text-align: center;">  </div> <p>The index server, queue server, and preprocessor can each need up to 2 GB main memory. This gives a total value of 6 GB. If TREX runs on one host together with other components, make sure that TREX has exclusive use of the necessary memory space.</p> </li> <li>• Processor: <ul style="list-style-type: none"> <li>○ At least Pentium III with 2 processors, each with a tact frequency of at least 1 GHz.</li> <li>○ We recommend Pentium IV with 2 processors, each with a tact frequency of 2 GHz.</li> </ul> </li> </ul>

Prerequisite Type	Prerequisite
Software Prerequisites	<ul style="list-style-type: none"> <li>Operating system: <ul style="list-style-type: none"> <li>Microsoft Windows 2000 Advanced Server (English version) with service pack 3</li> <li>Recommended: Microsoft Windows 2000 Advanced Server (English version) with service pack 3</li> </ul> </li> <li>Python: Version 2.1.3. A Python version by ActiveState is part of the delivery, and is installed by TREX setup if no suitable Python version is already installed</li> <li>SAP Gateway (Standard) 6.10 or 6.20. (Alternatively, version 4.6D from compilation 3)</li> </ul>



## Preparation

### Purpose

Before you install TREX, install SAP Gateway.



## Installing SAP Gateway

### Use

Before you install TREX, install SAP Gateway on the TREX host. For information on the appropriate gateway version, see [Checking Prerequisites \[Page 11\]](#).

### Prerequisites

If you had previously installed an older gateway version on the TREX host, you have now deinstalled it. Refer to SAP Note 179175 for further information.

### Procedure

1. Read SAP Note 0137478 first.
2. Insert the *Presentation CD* into your CD-ROM disk drive.
3. Follow the instructions of the SAP Gateway installation guide. You can find this on the CD in the subdirectory *Docu*.



When using the gateway in combination with TREX, you have to enter the following parameters when you install the gateway:

SAPSYSTEMNAME:      **GWS**

SAPSYSNR:              **47** (gateway instance)

If you want to use an SAP system number other than 47, you have to make sure that you enter the same SAP system number for the following settings:

- During the TREX installation  
RFC Connection Parameters dialog box (See: [Input Parameters for the TREX Installation \[Page 14\]](#))
  - Parameter of RFC destination (transaction SM59).  
Submenu *Gateway Options*, field *Gateway Service*, value  
`sapgw<SAPSYSNR>`.
4. Restart the host once the installation has finished.



If you are using version 4.6D of gateway, you have to set up your system so that the gateway starts automatically when you restart your computer. The procedure is described in SAP Note 0080474.

If you are using version 6.10 or 6.20 of the gateway, you do not need to do this. The required system settings are entered automatically during the installation of the gateway.



## Installation

### Purpose

The following sections describe the installation of TREX.






### Input Parameters for the TREX Installation



#### Procedure

To prepare for the input phase of the TREX installation, we recommend that you note down the necessary entries **before** starting the setup program. If you prepare the necessary information in advance, you can avoid unnecessary delays and errors.

The table below shows the names of the windows that are displayed during the TREX installation process, and the entries that you need to make.

Window	Entry
Welcome	Choose <i>Next</i> .
TREX Directory	<p>Select the target directory in which TREX is to be installed. For optimum performance, this directory needs to be on a local hard disk.</p>  <p>Make sure that there is enough disk space available for the TREX software, indexed, queues, and any necessary index backups (see <a href="#">Checking Prerequisites [Page 11]</a>).</p>
Required Third-Party Software, Python Directory	<p>The setup program checks whether an applicable Python version has been installed. If a suitable version of Python is not installed, you can choose the target directory for the Python installation.</p> <p>Note the following:</p> <ul style="list-style-type: none"> <li>• The directory name must contain the character string <code>Python</code> with the exact capitalization used here.</li> <li>• The directory name cannot contain spaces.</li> </ul> <p>Examples: C:\Python21, D:\Programs\Python21, D:\Programs\TREX_Python21</p>
Setup Type	Choose <i>Custom</i> .
Components	Select <i>RFC SAP Retrieval Service</i> and remove the selection of <i>TREX HTTP Server</i> .

Window	Entry
Proxy Settings	<p data-bbox="444 247 1343 342">If TREX can only reach documents to be indexed using a proxy server, the proxy server needs to be registered. Enter the following data and check the default settings:</p> <ul data-bbox="477 359 1326 768" style="list-style-type: none"><li>• <i>Proxy Server Name</i> – host name and domain of the proxy server Example: <code>proxy.mylocation.mycompany.com</code></li><li>• <i>Proxy Server Port</i> Example 8080</li><li>• <i>Proxy User Name</i> – user name needed to access the proxy server</li><li>• <i>Proxy User Password</i> – password defined for the user</li><li>• <i>Proxy Exclusions</i> – addresses for which the proxy server is not to be used. As a rule, you enter one or more character strings in which the addresses in your intranet end. Separate multiple addresses using a semicolon.</li></ul> <div data-bbox="599 785 651 835"></div> <p data-bbox="599 852 812 877"><code>mycompany.com</code></p> <p data-bbox="599 900 626 926">or</p> <p data-bbox="599 945 985 970"><code>mylocation.mycompany.com</code></p> <p data-bbox="599 993 626 1018">or</p> <p data-bbox="599 1037 1356 1092"><code>onelocation.mycompany.com;anotherlocation.mycompany.com</code></p> <div data-bbox="599 1108 651 1159"></div> <p data-bbox="599 1176 1042 1201">Do not use asterisks as placeholders.</p>

Window	Entry
Index Server, Queue Server, Preprocessor, Name Server, HTTP Server	<p>Check the host names and ports of the individual TREX servers.</p> <ul style="list-style-type: none"> <li>If you are not using a DNS (domain name system) server, replace the host name of the host on which you are carrying out the installation with its IP address.</li> </ul> <p></p> <p>Do not use the name <code>localhost</code>. Communication using <code>localhost</code> is much slower than using the host name or IP address.</p> <ul style="list-style-type: none"> <li>The TREX servers use the following ports by default: <ul style="list-style-type: none"> <li>Index server 8351</li> <li>Queue server 8352</li> <li>Preprocessor 8357</li> <li>Name server 8355</li> </ul> </li> </ul> <p>If these ports are already used, change the port numbers.</p> <p></p> <p>In a non-portal environment, the HTTP server is not relevant.</p>
RFC Connection Parameters	<p>Enter the following data so that TREX can communicate with the SAP system using the SAP Gateway:</p> <ul style="list-style-type: none"> <li>The host name of the host on which the Gateway is installed.</li> <li>The SAP system number that you entered during the Gateway installation (see <a href="#">Installing SAP Gateway [Page 12]</a>).</li> </ul>
Document Languages	<p>These settings are only valid for the automatic language recognition process.</p> <p>If you are installing TREX in a non-Portal environment, the document language settings are not currently relevant. This is because the application transmits the document language to TREX, meaning that language recognition is not necessary.</p>
Settings	<p>The setup program displays the chosen settings. To start the installation, choose <i>Next</i>.</p>





## Installing TREX

### Use

This section describes how to install TREX on released Windows platforms.

### Prerequisites

If you deinstalled TREX previously, you restarted the host after the deinstallation.

### Procedure

1. Log on to the installation host **with administrator permissions**.
2. Insert the *TREX CD* into your CD drive.
3. Start the TREX setup program:  
`<TREX CD>:\setupW2k.exe`
4. Follow the instructions of the setup program. You can find the required input parameters under [Input Parameters for the TREX Installation \[Page 14\]](#).
5. The installation starts when you have entered all necessary data. A message appears at the end of the installation telling you that it has been successful. Choose *Finish*.



## Post Installation

### Purpose

The following sections describe all steps to be carried out after the installation.



## Checking Performance Settings for the Operating System

### Use

To optimize the performance of TREX when using the released Windows platform, you need to check your Windows configuration and make changes if necessary.

### Optimizing Data Throughput For Network Applications

The Windows installation normally makes caching settings that are optimized for file servers. The operating system then reserves a large part of the main memory for the caching of files. Since this file-system cache impairs performance when indexing, you ought to change these settings.

1. Use the secondary mouse button to choose *My Network Places* from the Windows desktop, and choose *Properties*.
2. Use the secondary mouse button to click on *Local Area Connection*, and then choose *Properties*.
3. Under *Components checked are used by this connection*, choose *File and Printer Sharing for Microsoft Networks*.
4. Choose *Properties*, and select *Maximize data throughput for network applications*.
5. Choose *OK* twice.

### Optimizing Performance for Background Processes



Programs such as Microsoft SQL Server and Microsoft Exchange make the setting described below automatically when they are installed. If you have installed one of these programs, you do not need to make any changes.

The setting is only relevant if TREX is running as a service.

1. Use the secondary mouse button on *My Computer*, and choose *Properties*.
2. Choose the *Advanced* tab, and then choose *Performance Options*.
3. Under *Application Response*, choose the *Background Services* field.
4. Choose *OK* twice.



## Configuring SAP Gateway

### Use

If you are using version 4.6D of SAP Gateway, you have to set up your system so that the gateway starts automatically when you restart your computer.



If you are using version 6.10 or 6.20 of SAP Gateway, the system is automatically configured, and you do not have to make any changes.

### Procedure

The procedure for version 4.6D is described in note 0080474.



## Copying vpd.properties to the TREX Directory

### Use

The TREX setup program enters the TREX components installed into the file `vpd.properties`. Whenever you upgrade TREX, the TREX setup program accesses the information in this file.

For security reasons, we recommend that you copy the file `vpd.properties` to the TREX directory **every time** you install or upgrade TREX.

### Procedure

You find the file `vpd.properties` in the directory `C:\WINNT`.

Copy the file to the TREX directory.

### Result

If problems occur during the installation or upgrade of TREX, TREX support can use this copy for troubleshooting.



Other applications also enter information into the file `vpd.properties`. Therefore, never copy the file from the TREX directory to `C:\WINNT`. If you do so, the functionality of other applications can be impaired. Only TREX support may use the copy in the TREX directory for troubleshooting.



## Configuring the RFC Server

### Use

Settings for the RFC server are defined in the <TREX\_Directory>\imsconfig.cfg configuration file. Amongst other things, this file is used to configure whether the RFC server forwards documents to be indexed to the TREX queue server. The queue server collects the documents and transmits them to the index server according to the conditions defined in the queue parameters. The actual indexing then takes place on the index server.

In the standard configuration, the RFC server does not use the queue server. The ideal configuration depends on your application.



SAP Note 658052 contains recommendations for individual applications.

### Procedure

1. If necessary, edit the configuration file <TREX\_Directory>\imsconfig.cfg.
2. In the [CONNECTION] section, set the USE\_QUEUESERVER parameter to YES.

```
[CONNECTION]
```

```
USE_QUEUESERVER=YES
```

```
...
```

### Result

The changes take effect when you next start the RFC server. The RFC server is automatically started by the TREX daemon and/or by SAP Gateway.

Check the queue parameters and set them in line with your requirements. Make sure that you configure the intervals at which the queue server is to transmit documents to the index server. The settings that are recommended depend on how often documents are to be indexed, and how quickly you want them to be available for the search.

In a non-Portal environment, you use the queue client to configure the queue parameters. For more information, see the guide *Retrieval and Classification (TREX) – Administrating the Queue Server with the Queue Client*.



## Checking the Configuration of the TREX Daemon

### Use

If you want to create an RFC destination with the activation type 'Registration' (see [Creating an RFC Destination with Activation Type 'Registration' \[Page 26\]](#)), make sure that the TREX RFC server is entered into the configuration file `TREXDaemon.ini`. The TREX RFC server is then started automatically by the TREX daemon.

### Procedure

1. Edit the `<TREX_Directory>\TREXDaemon.ini` configuration file.
2. Check that the configuration file contains a section for the TREX RFC server. If not, add to the section:

```
[rfc_sapretrieval]
executable=rfc_sapretrieval.exe
arguments=imsconfig.cfg
startdir=<your_TREX_directory>
instances=<number_of_instances>
```

Take care with the parameter `instances`: The number of instances cannot be higher than the number of CPUs on the host. If your host has 4 CPUs, you can start up to 4 instances of the TREX RFC server. The default value is 1.

3. Check that the `programs` parameter in the `[daemon]` section contains an entry for the RFC server. If this is not the case, add the name of the RFC server section. The order of the entries does not matter.

```
[daemon]
programs=indexserver, queueserver, preprocessor, rfc_sapretrieval
```

### Result

Changes to the configuration file take effect when you next start the TREX daemon.



## Optional: Activating Python Extensions

### Use

Depending on the scenario, you might need to activate the TREX Python extensions. Firstly, you have to activate the Python extension handler, and then you register the necessary Python extensions.

### Procedure

#### Activating the Python extension handler

1. Edit the <TREX\_Directory>\TREXExtensions.ini configuration file.
2. Check that the [activate] section has the structure below, and modify the section if necessary.

```
[activate]
imsapi=search, thesaurus, admin
preprocessor
```

3. In the [extensionhandlers] section, add the line **trexxpy** and/or remove the comment sign (#).

```
[extensionhandlers]
trexxpy
```

#### Registering Python extensions

1. Edit the <TREX\_Directory>\extensions.py configuration file.
2. In the relevant section, change the entry `if 0:` to `if 1:`. You identify the extensions by the class name.

Extension	Class
XML attribute extension	XmlExtractor
Expansion of linguistic search queries	LinguistFix
Metadata extraction	AttributeExtractor
Topic maps	XtmExpander
Index replication – extension for the master	ReplicationMaster
Index replication – extension for the slave	ReplicationSlave



Activate XML attribute extension:

```
# XML attribute extractor extension
# -----
if 1:

    sys.path.append(os.path.join(os.getenv('SAP_RETRIEVAL_PATH'),
                                     'extensions', 'attribute-extractor'))

    from xml extractor import XmlExtractor
    trexx.registerExtension(trexx.EXTCLASS_INDEXING,
                           XmlExtractor(debug=0, mimetypes=['text/xml']))
```

## Result

The changes take effect when you next start the TREX daemon. If errors occur during routine operation and the required functions are not available, check the trace file (<TREX\_Directory>\trace\PythonExtension.log). This contains information on the incorrect entries in the TREX configuration files. If you cannot solve the problem, contact TREX support.



## Checking Access to Documents

### Use

Check that you can access the documents to be indexed from the TREX host.

### Procedure

1. Open a Web browser on the TREX host.
2. Enter the URL of one of the documents to be indexed.

If the document cannot be displayed, you might have to modify the proxy settings for the browser. Display the proxy settings for the browser, and enter the proxy server between TREX and the documents to be indexed.



## Creating an RFC Destination

### Purpose

You need to create an RFC destination so that the SAP system can send search and indexing requests to the TREX RFC server. You can choose from the following two activation types:

- `Start` (standard)
- `Registration` (only if your application requires this)

Refer to the documentation on the application in question to see which activation type is needed.



## Creating an RFC Destination with Activation Type 'Start'

### Prerequisites

- TREX and SAP Gateway are installed.
- Your application needs an RFC destination with the activation type *Start*.

### Procedure



If you have previously used an older TREX version, the RFC destination is already available. However, you should check the parameters as a precaution.

1. Choose transaction *SM59* in the SAP system, and then choose *Create*.
2. In the *RFC Destination* field, enter **IMSDEFAULT**.
3. Choose connection type *T*.
4. Enter a description, for example, **RFC Destination for Indexing and Searching**.
5. Save your entries.

The rest of the procedure is dependent on the TREX version that you are using.

### Release 4.6D

In the dialog box that now appears, define the settings for the **IMSDEFAULT** RFC destination.

1. Choose the activation type *Start*.
2. In the *Start On* area, choose the *Explicit Host* option.
3. In the *Program* field, enter **ims\_server\_admin.exe**.
4. For *Save As*, choose the required option (host or IP address).
5. In the *Target Host* field, enter the host name (with domain name if necessary) or the IP address of the TREX host.





Program: `ims_server_admin.exe`

Target Host: `<hostname.domainname>`

Save As: Hostname (selected option)

6. Save your entries.

Now define the settings for SAP Gateway.

7. Choose the *Gateway Options* submenu from the *Destination* menu.
8. In the *Gateway Host* field, enter the host name (with domain name if necessary) or the IP address of the host on which SAP Gateway is installed.
9. In the *Gateway Service* field, enter the name of the Gateway. This value should normally be `sapgw47`.



RFC Destination: `IMSDEFAULT`

Gateway Host: `<hostname.domainname>`

Gateway Service: `sapgw47`

10. Choose *OK*.

You return to the previous dialog box.

11. Choose *Save*, and then *Test Connection*.

The RFC connection is established.



If you used the host name of the target host, and the connection test fails, enter the IP address for the host instead, and test the connection again. If the RFC connection can still not be established, see [Troubleshooting \[Page 33\]](#).

### Releases 6.10 and 6.20

1. On the *Technical Settings* tab page, choose the activation type *Start on Explicit Host*.
2. In the *Program* field, enter `ims_server_admin.exe`.
3. For *Save As*, choose the required option (host name or IP address)
4. In the *Target Host* field, enter the host name (with domain name if necessary) or the IP address of the TREX host.



Program: `ims_server_admin.exe`

Target Host: `<hostname.domainname>`

Save As: Hostname (selected option)

Now define the settings for SAP Gateway.

5. In the *Gateway Host* field, enter the host name (with domain name if necessary) or the IP address of the host on which SAP Gateway is installed.

6. In the *Gateway Service* field, enter the name of the Gateway. This value should normally be `sapgw47`.



Gateway Host: `<hostname.domainname>`

Gateway Service: `sapgw47`

7. Choose *Save*, and then *Test Connection*.

The RFC connection is established.



If you used the host name of the target host, and the connection test fails, enter the IP address for the host instead, and test the connection again. If the RFC connection can still not be established, see [Troubleshooting \[Page 33\]](#).



## Creating an RFC Destination with Activation Type 'Registration'

### Prerequisites

- TREX and SAP Gateway are installed.
- Your application needs an RFC destination with the activation type *Registration*.
- The TREX RFC server `rfc_sapretrieval.exe` is entered in the configuration file `TREXDaemon.ini` (see [Checking the Configuration of the TREX Daemon \[Page 21\]](#)).

### Procedure

1. Choose transaction *SM59* in the SAP system, and then choose *Create*.
2. In the *RFC Destination* field, enter `IMSDEFAULT_REG`.
3. Choose connection type *T*.
4. Enter a description, for example, **RFC Destination for Indexing and Searching**.
5. Save your entries.

The rest of the procedure is dependent on the TREX version that you are using.

#### Release 4.6D

In the dialog box that now appears, define the settings of the RFC destination `IMSDEFAULT_REG`.

1. Choose the activation type *Registration*.
2. In the *Registration* area, enter the host name of the TREX host and the registration program `rfc_sapretrieval` in the *Program ID* field.

Example: *Program ID*    `MYHOST.rfc_sapretrieval`



Enter the host name in upper case, and the registration program `rfc_sapretrieval` in lower case. If you do not do this, the connection test will fail.

3. Save your entries.

Now define the settings for SAP Gateway.

4. Choose the *Gateway Options* submenu from the *Destination* menu.
5. In the *Gateway Host* field, enter the host name (with domain name if necessary) or the IP address of the host on which SAP Gateway is installed.
6. In the *Gateway Service* field, enter the name of the Gateway. This value should normally be `sapgw47`.



*RFC Destination:* IMSDEFAULT\_REG

*Gateway Host:* <hostname.domainname>

*Gateway Service:* sapgw47

7. Choose *OK*.

You return to the previous dialog box.

8. Choose *Save*, and then *Test Connection*.

The RFC connection is established.



If you used the host name of the target host, and the connection test fails, enter the IP address for the host instead, and test the connection again. If the RFC connection can still not be established, see [Troubleshooting \[Page 33\]](#).

### Releases 6.10 and 6.20

1. On the tab page *Technical Settings*, choose the activation type *Registered Server Program*.
2. Enter the host name of the TREX host and the registration program `rfc_sapretrieval` in the *Program ID* field.

Example: *Program ID*    `MYHOST.rfc_sapretrieval`



Enter the host name in upper case, and the registration program `rfc_sapretrieval` in lower case. If you do not do this, the connection test will fail.

3. In the *Gateway Host* field, enter the host name (with domain name if necessary) or the IP address of the host on which SAP Gateway is installed.

4. In the *Gateway Service* field, enter the name of the Gateway. This value should normally be `sapgw47`.



Gateway Host: `<hostname.domainname>`

Gateway Service: `sapgw47`

5. Choose *Save*, and then *Test Connection*.

The RFC connection is established.



If you used the host name of the target host, and the connection test fails, enter the IP address for the host instead, and test the connection again. If the RFC connection can still not be established, see [Troubleshooting \[Page 33\]](#).



## Creating a Search Server Relation

### Procedure



If you have previously used an older TREX version, the search server relation is already available. However, you should check the parameters as a precaution.

1. Choose transaction *SRMO* in the SAP system.
2. Choose *Create SSR*.
3. In the dialog box that now appears, enter a key or name into the *Search Server Relation ID* name.

Example: *Search Server Relation ID* `SSR_NAME`

4. Choose *Create SSR*.
5. In the *Search Engine* field, enter the name of the TREX search engine. The name is **DRFUZZY**.



Make sure that you enter the name **DRFUZZY** in uppercase and in the format specified.

6. In the field *RFC Destination (TCP/IP)*, enter the name of the RFC destination that you created. This entry must tally with the name that you assigned when you created the RFC destination (see [Creating an RFC Destination with Activation Type 'Start' \[Page 24\]](#) or [Creating an RFC Destination with Activation Type 'Registration' \[Page 26\]](#)).

Example: *RFC Destination (TCP/IP)* `IMSDEFAULT` or `IMSDEFAULT_REG`

7. Enter a description of the search server relation into the *Description* field.

Example: *Description* **Search Server Relation for Retrieval Service**

8. Save your entries.

You return to the previous dialog box.

9. Select the newly created search server relation in the table.
10. Choose *Set SSR as Default*.
11. In the confirmation prompt that appears, choose *Yes*.

The search server relation is then shown as default in the table.

## Result

You can now test your search server relation (see [Testing a Search Server Relation \[Page 29\]](#)).



## Testing a Search Server Relation

### Prerequisites

The search server relation exists (see [Creating a Search Server Relation \[Page 28\]](#)).

### Procedure

1. Choose transaction *SRMO* in the SAP system.
2. Choose the *RFC Destinations* tab page.  
You receive two RFC destinations for the search server relation that you created. One is for indexing (I), and the other is for searching (S).
3. Select one of the entries, for example, the entry with Action = S.
4. In the *Search Engine Settings* sub area, choose *Connection Test*.

The connection with the TREX RFC server and the TREX search engine is established. You can see this in the version information that is shown for the TREX components.



If the connection cannot be established, refer to [Troubleshooting \[Page 33\]](#).



## Scheduling Asynchronous Indexing and Deindexing

### Use



This step is only needed if you use the Knowledge Provider (KPRO) component.

Asynchronous indexing and deindexing is triggered using the report *RSTIRIDX*. Schedule this report as a daily background process.



The report *RSTIRIDX* starts the indexing and deindexing of scheduled documents, and logs indexing errors in the productive system. The chosen recipient can view the contents of the report using the transaction SO01. If an error occurs, check in IMS Monitoring (see [Checking for Problems in IMS Monitoring \[Page 36\]](#)).

### Prerequisites

To use a report, you need *Batch Administrator* authorization for the authorization object *Batch Processing*.

### Procedure

1. In the SAP system, choose *System* → *Services* → *Jobs* → *Job Definition* (transaction SM36).
2. In the field *Jobname*, enter `INDEXING`.
3. Choose the Job Class *B*.
4. In the field *Target Server*, enter the name of the host on which the background process is to be carried out.
5. Choose *Spool List Recipient*.
6. In the field *Recipient*, enter the name of the desired recipient, select the required general attributes, and choose *Copy*.
7. Choose *Start Date*.
8. Choose *Date/Time*.
9. Enter the required start date.
10. Select the option *Execute Job Periodically*.
11. Choose *Period Values*.
12. Choose *Daily*.
13. Choose *Save*.
14. Choose *Save again*.
15. Choose *Steps*.
16. Choose *ABAP Program*.
17. In the *ABAP Program* area, enter the name `RSTIRIDX` in the *Name* field.

18. Choose *Print Specifications*.
19. In the area *Spool Control*, choose the option *Print Immediately*.
20. Choose *Save*.



## Installation Check

### Purpose

If you have carried out all the steps described, TREX should be ready for operation. You can carry out checks to ascertain whether the installation was successful.



## Checking TREX Processes

### Procedure

Open the Task Manager and check whether the following processes are running:

- `rfc_sapretrieval.exe`
- `TREXDaemon.exe`
- `TREXIndexServer.exe`
- `TREXQueueServer.exe`
- `TREXPreprocessor.exe`



## Executing an Installation Test Script

### Use

TREX delivers a Python script that you can use to test the basic functions of TREX. If the Python script is executed successfully, you know that TREX has been installed properly, the configuration files contain the necessary entries, and the TREX servers are running.

### Procedure

1. Open a prompt and go to the directory in which the Python scripts are located:

```
cd <TREX_Directory>\python_support\test_tools\lib
```

2. Execute the script `runInstallationTest.py`.

- If you are using the default ports of the TREX servers, execute the script without arguments:

```
runInstallationTest.py
```

- If you changed the default ports during the installation of TREX, execute the script with the following arguments:

```
runInstallationTest.py -- indexServer=<hostname>:<port> --  
queueServer=<hostname>:<port>
```

### Result

The script carries out the following tests:

- Deleting any test indexes that were generated for a previous script run
- Creating a test index
- Indexing documents
- Testing search functions
  - Exact, error-tolerant (fuzzy), and linguistic searches
  - Search with the Boolean operators AND and OR

After the script run the Web browser is started and a table is displayed with the results of the script run. You see the tested calls and their statuses (OK or Failed).

When you run the script for the first time, the call 'Delete Index' has the status `Failed`. This is because there was no existing text index to be deleted. If this is the only cell with the status `Failed`, the test was successful.



If the Web browser is not started automatically, you can display the results of the script run by opening the following file with the Web browser:

```
<TREX_Directory>\python_support\test_tools\lib\script_results\  
log_first_test_index\<current_date>\start.html
```





## Additional Information

### Purpose

The sections below contain information on the following topics:

- Troubleshooting
- Starting and stopping the TREX demons and individual TREX servers
- Deinstalling TREX



## Troubleshooting

### Purpose

The following sections contain information on troubleshooting.



Make sure you observe the correct format for input parameters. In some cases, the system distinguishes between uppercase and lowercase.

The parameters must therefore be entered exactly as defined, paying attention to uppercase and lowercase letters.



## Checking SAP Gateway

### After the installation of the gateway and the restart of the host

Open the Task Manager and check whether the `gwrd.exe` process has started.



Make sure that the gateway starts automatically when the host is restarted. The procedure is described in note 0080474.

### Settings for the gateway service

1. Choose *Start* → *Settings* → *Control Panel* → *Administrative Tools* → *Services*.
2. Check the start category of `SAPGWS_47`. It should start automatically. Change it if necessary. The service needs to be started.

### Has the gateway started?

If you are using version 4.6D of the gateway, you have to make sure that the gateway itself has started.



If you are using version 6.10 or 6.20 of the gateway, the system is automatically configured, and you do not have to make any changes.

You can use the SAP system management console to check whether the gateway has started, and to start it manually if necessary.

1. Choose *Start* → *Programs* → *SAP System Management Console*.
2. Check whether the relevant gateway instance has started.
3. To start the gateway instance, select it and choose *Action* → *Start*.



## Checking TREX

### Access rights for the installation directory

1. Navigate to the TREX directory, and select it.
2. In the *File* menu, choose the *Properties* entry.
3. Choose the *Security* tab page, and then choose *Permissions*.
4. Check that the following permission is set: *Everyone Full Control (All)/All*.
5. Select the *Replace Permissions on Subdirectories* option.
6. Choose *OK*.

### Connection test that fails when setting up RFC destinations and the search server relation for TREX

1. Check your entries (see [Creating an RFC Destination with Activation Type 'Start' \[Page 24\]](#) or [Creating an RFC Destination with Activation Type 'Registration' \[Page 26\]](#)). Pay attention to lowercase and uppercase letters in the input parameters.
2. Check that the host is connected to the network.
3. Check that the process `gwrld.exe` (SAP Gateway) has started (see [Checking SAP Gateway \[Page 34\]](#)).
4. In the `<TREX_Directory>\imsconfig.cfg` file, check the host name and gateway instance:

```
[CONNECTION]
HOST=<hostname>
INSTANCE=sapgw47
```



If you have updated TREX, the standard setting of the gateway service changed from `sapge00` in release 4.6A, to `sapgw47` in release 4.6B.

Make sure that the parameter `INSTANCE=sapgw<SAPSYSNR>` shows the number of the SAP system that you specified when you installed the gateway.

The entries must be made in lowercase.



## Checking for Problems in IMS Monitoring

### Use



This step is only needed if you use the Knowledge Provider (KPRO) component.

### Procedure

1. Choose transaction SKPR07 in the SAP system.
2. Choose *Scheduled Documents* in the *Extras* area.

The system displays a list of documents that are due to be indexed or deindexed. The following information is available:

- Client
- Name of index category (32-place GUID)
- Document class
- Document language
- Processing type (I = indexing, D = deindexing)
- The number of scheduled documents for this index category

3. To see details on the scheduled documents:

- Select the required list entry
- Choose *View*

All scheduled documents are listed. To see the number of attempts to index or deindex individual documents:

- Select the required documents
- Choose *View*

If an error has occurred whilst a document was being indexed or deindexed, it is scheduled for the process again. If the number of retries is large, there is probably an error that is preventing the indexing or deindexing of the document in general. Select such documents in the list of scheduled documents, and delete them.

Documents that could not be indexed or deindexed at the first attempt (including those documents that you have deleted from the list of scheduled documents manually) are treated by the system as problem cases.

4. To see a list of problem cases:

- In transaction SKPR07, choose *Problems*

You can filter this list by document class and by the number of attempts to index or deindex a document.

You can delete the listed documents, or you can mark them to be indexed or deindexed again.



## Starting and Stopping TREX

### Purpose

The following sections explain how to start and stop TREX.

The TREX daemon is registered as a service when TREX is installed. The service is configured so that it starts automatically when the host is started up, and stops automatically when the host is shut down. If the TREX service is running, TREX is ready to use.

You can start and stop the TREX service manually if necessary. For test purposes or troubleshooting, you can also:

- Stop and start the TREX servers individually
- Stop and start the TREX daemon in debug mode



## Starting and Stopping the TREX Service

### Use

The TREX daemon is registered as a service when TREX is installed. The service is configured so that it starts automatically when the host is started up, and stops automatically when the host is shut down.

You can start and stop the TREX service manually if necessary.

### Prerequisites

You restarted the host after you installed TREX.

### Starting the TREX Service Manually

Choose *Start* → *Programs* → *SAP TREX* → *TREX Service* → *Start*.

### Stopping the TREX Service Manually

Choose *Start* → *Programs* → *SAP TREX* → *TREX Service* → *Stop*.

Certain processing steps, for example, writing an index, cannot be interrupted. Such steps are completed before the TREX service is stopped. Therefore, it can take a certain amount of time to stop the service.

With large indexes, it can take up to a few hours to stop the service if lots of documents are currently being indexed.



## Starting and Stopping Individual TREX Servers

### Use

You can start individual TREX servers for test purposes and for troubleshooting. You can then track the program output on the screen.

### Starting the TREX Servers

1. Stop the TREX service (see [Starting and Stopping the TREX Service \[Page 37\]](#)).
2. Open a separate prompt for each TREX server.
3. Go to the TREX installation directory and start the server.

TREX Server	Command
Index server	<code>TREXIndexServer.exe</code>
Preprocessor	<code>TREXPreprocessor.exe</code>
Queue server	<code>TREXQueueServer.exe</code>
RFC server	<code>rfc_sapretrieval.exe - imsconfig.cfg</code>



In the properties of the prompt, deactivate the *QuickEdit Mode* option.

Leave the prompt open. If you want, you can minimize the window so that it is shown as a pushbutton in the Windows task bar.

### Stopping the TREX Servers

1. Display the window in which you started the TREX server.
2. Use CTRL+C or close the window.

Certain processing steps, for example, writing an index, cannot be interrupted. Such steps are completed before the TREX servers are stopped. Therefore, it can take a certain amount of time to stop the service.

With large indexes, it can take up to a few hours to stop the service if lots of documents are currently being indexed.



Do not use the Task Manager to stop the TREX servers. Doing so can lead to the loss of data. Affected indexes can be irreparably damaged.



## Starting and Stopping the TREX Daemon in Debug Mode

### Use

The TREX daemon is the program that is registered as the service. If you start the TREX daemon in debug mode, you can track the program output on the screen. You do this for test purposes and troubleshooting.

### Starting the TREX Daemon

1. Stop the TREX service (see [Starting and Stopping the TREX Service \[Page 37\]](#)).
2. Open a prompt and go to the TREX installation directory.
3. Execute the following command:

```
TREXDaemon.exe -d
```

### Stopping the TREX Daemon

1. Display the window in which you started the TREX daemon.
2. Use CTRL+C or close the window.

Certain processing steps, for example, writing an index, cannot be interrupted. Such steps are completed before the TREX daemon is stopped. Therefore, it can take a certain amount of time to stop the daemon.

With large indexes, it can take up to a few hours to stop the daemon if a large number of documents are currently being indexed.



Do not use the Task Manager to stop the TREX daemon. Doing so can lead to the loss of data. Affected indexes can be irreparably damaged.



## Deinstalling TREX

### Procedure

1. Stop the TREX service (see [Starting and Stopping the TREX Service \[Page 37\]](#)).
2. Choose *Start* → *Settings* → *Control Panel* → *Add/Remove Programs*.  
The dialog box that now appears shows a list of programs that can be removed automatically from the host.
3. Select the entry *KM Platform - Retrieval & Classification (TREX)*.
4. Choose *Change/Remove*.  
This starts the TREX deinstallation program.
5. Choose *Next* three times.
6. Check whether you want to keep or delete the configuration files `bartho.ini`, `imsconfig.cfg`, `TREXConfigMgr.ini`, `TREXIndexServer.ini`, `TREXQueueServer.ini`, and `TREXTcpipClient.ini`
  - If you do not want to use the existing indexes after the deinstallation, you can delete these configuration files.
  - If you want to reinstall TREX after the deinstallation, and continue to use the existing indexes, you must keep these configuration files.



Rename the configuration files after the deinstallation, for example, `old_TREXIndexServer.ini`.

The configuration files contain index information that you have to transfer to the new configuration files after installing TREX. Only carry out this step in consultation with TREX Support.

7. Choose *Finish*.
8. Restart the host.

### Result

For security reasons, not all files are deleted when the deinstallation is carried out. No indexes, queues, or trace files are deleted. If you want to completely deinstall TREX, delete the rest of the content in the TREX directory manually.