

# Configuration

---

## Configuration

Together Teamlösungen EDV-Dienstleistungen GmbH

Elmargasse 2-4

A-1190

Vienna

Austria

+43 (0) 5 04 04 - 122

+43 (0) 5 04 04 - 11 122

<office@together.at>

<http://www.together.at/together/index.html>

Copyright © 2006 Together Teamlösungen EDV-Dienstleistungen GmbH

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of the Together Teamlösungen EDV-Dienstleistungen GmbH.

Together Teamlösungen EDV-Dienstleistungen GmbH DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

---

---

# Table of Contents

.....	1
Configuration .....	1

---

# List of Tables

- 1. 'Together Search Server Admin' application configuration ..... 1
- 2. 'Together Search Server' (search) application configuration ..... 5
- 3. 'Together Document Viewer'application configuration .....6

---

# Configuration

'Together Document Viewer' and 'Together Search Server' are without database dependency. Configuration process uses configuration files placed in 'tss-resources' directory to configure definition of 'Sites'. List of files placed in 'tss-resources' directory :

1. 'configuration\_schema.xsd', represent the schema of 'siteConf.xml' configuration file.
2. 'siteConf.xml', represent the shared configuration file that will configure applications. (Path to the configuration file is defined in 'web.xml' file of the applications).
3. 'documentGroupFile.xml', represent the default document group file list, that will be used by configuration process in case that 'DOCUMENTGROUPLIST' is not defined in configuration file('siteConf.xml') for all sites.

This table contains the configuration options for the 'Together Search Server Admin' application. Values of 'Together Search Server Admin' configuration parameters are contained in XML documents (not in \*.conf files).

**Table 1. 'Together Search Server Admin' application configuration**

Attribute	Description
IndexDir	Folder where the indexes are stored ("C:\index" by default)
SearcherFactoryClassName	Wrapper class used for searching
IndexerFactoryClassName	Wrapper class used for indexing
LoggingManagerClassName	Logging implementation class
role-name	By default, this value is set to 'admin'. It is defined in Tomcat's tomcat-users.xml file ('%together_application_server_root%/conf'). In order to change this role(username/password), it has to be changed in tomcat-users.xml. Please refer to Tomcat's documentation to for more info.
MaxFieldLength	Maximal field length (in bytes) of the document contents field (being indexed) (10000 by default)
MaxIndexLength	Maximum number of files that will be processed in single index/update procedure. If 'Indexer.ReIndexMode' is set to 'true' next index/update procedure will start from the last file from previous procedure. Default value is 500000.
LogicalNameFromDatabase	0/1 – If 1 the value from a metadata(if available) table will be used for document title
DocumentLogicalName	Metadata table's key column's value describing document title
DocumentUpdate	Metadata table's key column's value describing document modification date.

DocumentUpdatePattern	This parameter represents the pattern of the modification date in column value.
DBFetchSize	Fetch size to be used during calls to metadata, include list and exclude list table queries
TempDir	Used in 'Together Search Server Admin' application. Temporary directory where zip files will be unpacked into temporary files, parsed and deleted. Default value is user home directory +/.snapperTemp directory.
RelativeIndexPath	If you want to use one index file on different server locations and/or operating systems (e.g. Windows/Linux) you should use index file which contains relative file paths for each Document (relative to index root). If you set this parameter to true, you can move your index files if you change index root path in database. Possible values: true/false. Default value: false
FileSeparator	Possible values for this parameter is system-dependent (default value), and unix. If unix value is set, 'Together Search Server Admin' and 'Together Search Server' will use / as a File separator for all paths in database and in index files.
Indexer.MountTitleInContent	Possible values for this parameter is 'false'(default value) and 'true'. If value is 'true', 'Together Search Server Admin' will mount title of the document on the beginning of the persed text (content).
Indexer.MountFilePathInContent	Possible values for this parameter is 'false'(default value) and 'true'. If value is 'true', 'Together Search Server Admin' will mount File path of the document on the beginning of the persed text (content).
Indexer.MountPropertiesInContent	Possible values for this parameter is 'false'(default value) and 'true'. If value is 'true', 'Together Search Server Admin' will mount properties of the document on the beginning of the persed text (content).
Indexer.MountMetaDataInContent	Possible values for this parameter is 'false'(default value) and 'true'. If value is 'true', 'Together Search Server Admin' will mount meta data of the document on the beginning of the persed text (content).
ForceIndexUnlock	Possible values for this parameter is 'false'(default value) and 'true'. If value is 'true', 'Together Search Server Admin' will unlock index files before updating index. It is dangerous option if more users update index at the same time.
MaxPropertiesLength	Maximal properties length of the document that will be indexed in propertie field (5000 by default)
Indexer.ReIndexMode	Possible values for this parameter is 'true'(default value) and 'false'. If value is 'true', 'Together Search Server Admin' will continue with index process from last indexed file, or start from beginning if index process is finished before . If value is 'false' every click on index button will start index process from beginning(previously behavior).
DocumentGroupConfFile	This parameter represents the path to the default definition of Document group list file . Default value of this parameter is '/tss-resources/documentGroupConf.xml', relative path to the < catalina.base> (<together_application_server_root>/multiserver). Note

	: DocumentGroupConfFile can be defined as absolut file path.
SiteConfFile	This parameter represent the path to the default definition of Sites . Default value of this parameter is '/tss-resource/siteConf.xml', relative path to the < catalina.base> ('<Enhydra-7.0-dir>/multiserver'). Note : SiteConfFile can be defined as absolute file path.
StatisticActive	Possible values for this parameter is 'false'(default value) and 'true'. If value is 'true', 'Together Search Server Admin' will try to acces the StatisticDirectory to read and show statistc information.
StatisticDirectory	This parameter represnt the path to the StatisticDirectory where statistic files are placed. Note : this parameter must have the same value as in 'Together Search Server' application coniguration file.
IndexOSSpecific	Possible values for this parameter is 'false'(default value) and 'true'. If value is 'true', 'Together Search Server Admin' will access system specific resources through 'Java Native Interface' to collect additional information about the file (creation and last access time) . Note : Win32File.dll or libLinuxFile.so must be placed on the 'java.library.path'. By default they are placed in '<tas-dir>/multiserver/bin' directory.
Parser.Excel.ConverterClassName	Possible value for this parameter is 'org.enhydra.snapper.parsers.fileparsers.parsers.excelparser.ExcelToText' (default value) . 'Together Search Server Admin' will store pure text in index.
Parser.PowerPoint.ConverterClassName	Possible value for this parameter is 'org.enhydra.snapper.parsers.fileparsers.powerparser.PowerToText'. (default value) . 'Together Search Server Admin' will store pure text in index.
Parser.Word.ConverterClassName	Possible value for this parameter is 'org.enhydra.snapper.parsers.fileparsers.wordparser.WordToText'. (default value) . 'Together Search Server Admin' will store pure text in index.
Parser.Excel2007.ConverterClassName	Possible value for this parameter is 'org.enhydra.snapper.parsers.fileparsers.parsers.excel2007parser.Excel2007ToText' (default value) . 'Together Search Server Admin' will store pure text in index.
Parser.PowerPoint2007.ConverterClassName	Possible value for this parameter is 'org.enhydra.snapper.parsers.fileparsers.powerpoint2007parser.PowerPoint2007ToText'. (default value) . 'Together Search Server Admin' will store pure text in index.
Parser.Word2007.ConverterClassName	Possible value for this parameter is 'org.enhydra.snapper.parsers.fileparsers.wordparser2007parser.Word2007ToText'. (default value) . 'Together Search Server Admin' will store pure text in index.
SaveConvertedFile	Possible values for this parameter is 'true'(default value) and 'false'. If value is 'true', 'Together Search Server Admin' will if Converter has ability to convert file to html form be saved to file sistem.
PathOfConvertedFiles	Folder where the converted 'html' files will be stored.
Parser.Excel2007.ConverterClassName	Name of converter for Excel 2007 files.Alternative is

	org.enhydra.snapper.parsers.fileparsers.parsers.excel2007parser.Excel2007ToText.In combination with 'SaveConvertedFile' or 'SaveConvertedExcel2007' set to 'true',converted HTML files will be saved on file system.
Parser.PowerPoint2007.ConverterClassName	Name of converter for PowerPoint 2007 files.Alternative is org.enhydra.snapper.parsers.fileparsers.parsers.powerpoint2007parser.PowerPoint2007ToText.In combination with 'SaveConvertedFile' or 'SaveConvertedPowerPoint2007' set to 'true',converted HTML files will be saved on file system.
Parser.Word2007.ConverterClassName	Name of converter for Word 2007 files.Alternative is org.enhydra.snapper.parsers.fileparsers.parsers.wordparser2007parser.Word2007ToText.In combination with 'SaveConvertedFile' or 'SaveConvertedWord2007' set to 'true',converted HTML files will be saved on file system.
Word2007TransformationPath	Path to custom XSL transformation path.If set,it will be used instead of default XSL transformation for Word 2007.
Excel2007TransformationPath	Path to custom XSL transformation path.If set,it will be used instead of default XSL transformation for Excel 2007.
PowerPoint2007TransformationPath	Path to custom XSL transformation path.If set,it will be used instead of default XSL transformation for PowerPoint 2007.
SaveConvertedWord	Should converted Word files be saved in HTML format on file system.If 'true',converted files can be used for preview option.Default value is 'false'.
SaveConvertedExcel	Should converted Excel files be saved in HTML format on file system.If 'true',converted files can be used for preview option.Default value is 'false'.
SaveConvertedPowerPoint	Should converted PowerPoint files be saved in HTML format on file system.If 'true',converted files can be used for preview option.Default value is 'false'.
SaveConvertedWord2007	Should converted Word 2007 files be saved in HTML format on file system.If 'true',converted files can be used for preview option.Default value is 'false'.
SaveConvertedExcel2007	Should converted Excel 2007 files be saved in HTML format on file system.If 'true',converted files can be used for preview option.Default value is 'false'.
SaveConvertedPowerPoint2007	Should converted PowerPoint 2007 files be saved in HTML format on file system.If 'true',converted files can be used for preview option.Default value is 'false'.
CheckConnection	Should “dummy” sql query be used before executing other queries in order to check if DB connection is established.Default value is 'true'.
ParserCharacterLimit	This parameter represent the limit of characters in the document that will be parsed. Default value is '0'. If value of this parameter is '0' no limit will be performed (complete document will be parsed).
CharacterLimitForParser	This parameter represent the list of parsers that will use the parameter 'ParserCharacterLimit'. Parsers that support parameter 'ParserPageLimit' are : Microsoft Excel



	parser ( value : 'xls' ) , Message parser ( value : 'msg' ), Open office parser ( value : 'oo' ) , Microsoft Power Point parser ( value : 'ppt' ), Rich text documents parser ( value : 'rtf' ) , Microsoft Word parser ( value : 'doc' ) and Text parser ( value : 'txt' ).
ParserPageLimit	This parameter represent the number of pages in the document that will be parsed. Default value is '0'. If value of this parameter is '0' no limit will be performed (complete document will be parsed).
PageLimitForParser	This parameter represent the list of parsers that will use the parameter 'ParserPageLimit'. Default value is 'ppt,pdf,docx,xlsx,pptx'.
MaxFileSize	If file size exceeds this limit,file will not be parsed.If 0, no limit will be applied.Limit is set in bytes (for example for limit of 1MB set 1024).Default value is 0-no limit.
TimeLimit	If file parsing time exceeds this limit,parsing procedure will end.If 0, no limit will be applied. Limit is set in milliseconds (for example for limit of 1sec set 1000).Default value is 0-no limit.
OptimizeOnIndex	Should index be optimized after indexing procedure ends.
OptimizeOnUpdate	Should index be optimized after updating procedure ends.

**Table 2. 'Together Search Server' (search) application configuration**

Attribute	Description
IndexDir	Folder where the indexes are stored (“C:\index” by default)
SearcherFactoryClassName	Wrapper class used for searching
LoggingManagerClassName	Logging implementation class
Xsl.Directory	Folder the xsl filed used for search result representation are stored
MaxRange	Restricts the maximum number of documents that can be displayed per results page. This parameter is used as assurance, in case of large number of search result per one page.
RelativeIndexPath	If you want to use one index file on different server locations and/or operating systems (Windows/Linux) you should use index file which contains relative file paths for each Document (relative to index root). If you set this parameter to true, you can move your index files if you change index root path in database. Possible values:true/false. Default value: false. Use the same value as in admin application.
FileSeparator	Possible values for this parameter is system-dependent (default value), and unix. If unix value is set, 'Together Search Server Admin' and 'Together Search Server' will use / as a File separator for all paths in database and in index files. Use the same value as in admin application.

PreviewURL	URL where 'Together Document Viewer' application where installed and where 'Together Document Viewer' request should be sent. Default is http://localhost:8080/tdv/ Preview.po. It is URL used for building preview link (previewURL + preview parameters)
GoogleSearcherURL	This parameter represent the url of the google web service.
GoogleSearcherKey	This parameter represent the key that is used by the google web service to recognize the user. To access the Google Web APIs service, you must have a Google Account [https://www.google.com/accounts/NewAccount?continue=http://api.google.com/createkey&followup=http://api.google.com/createkey&gtcreate] and obtain a license key. Your Google Account and license key entitle you to 1,000 automated queries per day.
GoogleResultLimit	This parameter represents the number of google search result displayed in result page. The maximum value is 10 result per page (Restriction by google web service).
SearchDatePattern	This parameter represents the pattern of 'Modified' input field (document attribute) in advanced search. Possible values are : 'yyyyMMdd', 'dd.MM.yyyy', 'yyyy.MM.dd', 'yyyy-MM-dd' and 'dd-MM-yyyy'. Default value is 'yyyyMMdd'!
ResultDatePattern	This parameter represent the pattern of 'Modified' date in search result. Possible values are : 'dd.MM.yyyy', 'yyyy.MM.dd', 'yyyy-MM-dd' and 'dd-MM-yyyy'. Default value is 'yyyy.MM.dd',
DocumentGroupConfFile	This parameter represent the path to the default definition of Document group list file . Default value of this parameter is '/resource/documentGroupConf.xml', relative path to the < catalina.base> ('<Enhydra-7.0-dir>/multiserver'). Note : Document-GroupConfFile can be defined as absolute file path.
SiteConfFile	This parameter represent the path to the default definition of Sites . Default value of this parameter is '/resource/siteConf.xml', relative path to the < catalina.base> ('<Enhydra-7.0-dir>/multiserver'). Note : SiteConfFile can be defined as absolute file path.
StatisticActive	Possible values for this parameter is 'false'(default value) and 'true'. If value is 'true', 'Together Search Server' will log statistic information.
StatisticDirectory	This parameter represnt the path to the StatisticDirectory where statistic files are placed. Note : this parameter must have the same value as in 'Together Search Server Admin' application coniguration file.
ReReadConfigFilesEveryMinutes	Configuration files are rereaded by default every 5 minutes.
ParseOnFly	Should documents be parsed “on the fly” for preview. If false, HTML preview will be created from files stored on file system. If file does not exist on file system,preview will not be created. Default value is 'true'.

**Table 3. 'Together Document Viewer' application configuration**

Attribute	Description
SearcherFactoryClassName	Wrapper class used for searching
LoggingManagerClassName	Logging implementation class
Xsl.Directory	Folder the xsl filed used for search result representation are stored
ParserPageLimit	This parameter represent the number of pages in the document that will be parsed. Default value is '0'. If value of this parameter is '0' no limit will be performed (complete document will be parsed).
PageLimitForParser	This parameter represent the list of parsers that will use the parameter 'ParserPageLimit'. Default value is 'ppt,pdf,docx,xlsx,pptx'.
ParserCharacterLimit	This parameter represent the limit of characters in the document that will be parsed. Default value is '0'. If value of this parameter is '0' no limit will be performed (complete document will be parsed).
CharacterLimitForParser	This parameter represent the list of parsers that will use the parameter 'ParserCharacterLimit'. Parsers that support parameter 'ParserPageLimit' are : Microsoft Excel parser ( value : 'xls' ) , Message parser ( value : 'msg' ), Open office parser ( value : 'oo' ) , Microsoft Power Point parser ( value : 'ppt' ), Rich text documents parser ( value : 'rtf' ) , Microsoft Word parser (value : 'doc') and Text parser ( value : 'txt' ).
DocumentGroupConfFile	This parameter represent the path to the default definition of Document group list file . Default value of this parameter is 'resource/documentGroupConf.xml', relative path to the < catalina.base > ( ' < Enhydra-7.0-dir > /multiserver ). Note : Document-GroupConfFile can be defined as absolute file path.
SiteConfFile	This parameter represent the path to the default definition of Sites . Default value of this parameter is ' / resource/siteConf.xml', relative path to the < catalina.base > ( ' < Enhydra-7.0-dir > /multiserver ). Note : SiteConfFile can be defined as absolute file path.
ReReadConfigFilesEveryMinutes	Configuration files are rereaded by default every 5 minutes.
PathOfConvertedFiles	Folder where the converted 'html' files are stored.
ReadConvertedFile	Preview of 'Excel' files can be done from index (pure text) (parameter value 'false'), or if the Excel file is converted during index proces then converted (HTML file from file system) will be displayed.
Parser.Excel.ConverterClassName	Possible value for this parameter is 'org.enhydra.snapper.parsers.fileparsers.parsers.excelparser.ExcelToText' (default value) .
Parser.PowerPoint.ConverterClassName	Possible value for this parameter is 'org.enhydra.snapper.parsers.fileparsers.parsers.powerparser.PowerToText' . (default value) .
Parser.Word.ConverterClassName	Possible value for this parameter is

	'org.enhydra.snapper.parsers.fileparsers.parsers.wordparser.WordToText'. (default value) .
Parser.Excel2007.ConverterClassName	Possible value for this parameter is 'org.enhydra.snapper.parsers.fileparsers.parsers.excel2007parser.Excel2007ToText' (default value) .
Parser.PowerPoint2007.ConverterClassName	Possible value for this parameter is 'org.enhydra.snapper.parsers.fileparsers.parsers.powerpoint2007parser.PowerPoint2007ToText'. (default value) .
Parser.Word2007.ConverterClassName	Possible value for this parameter is 'org.enhydra.snapper.parsers.fileparsers.parsers.wordparser.Word2007ToText'. (default value) .
TimeToSleep	Time (in minutes;default 1440=24h) after which a thread for cleaning temporary files and folders will be activated. Process repeats every "TimeToSleep" minutes.
JBIGConverterPath	Relative path to JBIG converter from "webapps\tdv\WEB-INF\classes".This converter will be used for image files compressed with JBIG2 standard.
TimeLimit	If file parsing time exceeds this limit,parsing procedure will end.If 0, no limit will be applied. Limit is set in milliseconds (for example for limit of 1sec set 1000).Default value is 10000 - 10 seconds.
ContainerFilesPreviewLimit	Number of files that will be shown inside container preview option.If set to 0,all container files will be shown.Default value is 3.
Word2007TransformationPath	Path to custom XSL transformation path.If set,it will be used instead of default XSL transformation for Word 2007.
Excel2007TransformationPath	Path to custom XSL transformation path.If set,it will be used instead of default XSL transformation for Excel 2007.
PowerPoint2007TransformationPath	Path to custom XSL transformation path.If set,it will be used instead of default XSL transformation for PowerPoint 2007.