Search Indexing tab

Overview

Use this tab to configure specific MIME types in order to search within files (such as PDF files) that have been uploaded to the Tiki File Gallery.

To Access

From the File Gallery Admin page, click the **Search Indexing** tab.

Bosnian (bos...

Note

In order to search within uploaded files, your server may require additional applications, such as **strings** or **pdftotext**. It also relies on utilities on the server, more information on searching within files at Search-within-files. To check support on your server, use the admin tool. Tiki Check.

Related Topics

Search within files

Option	Description	Default
Automatic indexing of file content	Uses command line tools to extract the information from the files based on their MIME types.	Disabled
Automatic indexing of emails stored as files	Parses message/rfc822 types of files (aka eml files) and stores individual email headers and content in search index.	Disabled
Asynchronous indexing		Enabled
OCR Files	Extract and index text from supported file types.	Disabled
DCR Every File	Attempt to OCR every supported file.	Disabled
Allow file level OCR languages	Allow users to change the default languages that will be used to OCR a file.	Enabled
OCR limit languages	Limit the number of languages one can select from this list. ⇒ Auto detect languages Afrikaans (Afrikaans) Albanian (Shqip) Amharic (አলেের) Arabic Arabic (العربية) Armenian Armenian (Հայերեն) Assamese (অসমীয়া) Azerbaijani (azərbaycan dili) Azerbaijani	None

(azərbaycan dili) (cyrl) | Basque (euskara, euskera) | Belarusian (беларуская мова) | Bengali | Bengali (বাংলা) |

Bif blank, the \$PATH will be used, but will likely fail with scheduler. Path to the location of the binary. Defaults to the \$PATH location. If blank, the \$PATH will be used, but will likely fail with scheduler. Deform Description Description Automatic indexing of file content Automatic indexing of emails stored is files Automatic indexing of emails stored content in search index. Description Default Disabled Parses message/rfc822 types of files (aka eml files) and stores individual email headers and content in search index. Disabled Disabled Disabled DOCR Files Extract and index text from supported file types. DOCR Every File Attempt to OCR every supported file. Allow dile level OCR languages Allow users to change the default languages that will be used to OCR a file. Limit the number of languages one can select from this list. Extract detect languages Afrikaans (Afrikaans) Albanian (Shqip) Amharic (************************************	Description	Default
Description Descr		/usr/bin/tesseract
Automatic indexing of file content Automatic indexing of file content Automatic indexing of emails stored as files Automatic indexing of emails stored as files Automatic indexing of emails stored as files Asynchronous indexing Disabled Disabled Disabled DOCR Files Extract and index text from supported file types. Disabled Disabled Disabled Disabled Disabled DOCR Every File Attempt to OCR every supported file. Disabled Allow users to change the default languages that will be used to OCR a file. Enabled DISABLED DISABL		/usr/bin/pdfimages
Parses message/rfc822 types of files (aka eml files) and stores individual email headers and content in search index. Disabled content in search index. Enabled DCR Files Extract and index text from supported file types. Disabled DCR Every File Attempt to OCR every supported file. Allow users to change the default languages that will be used to OCR a file. Disabled DCR limit languages Limit the number of languages one can select from this list. Auto detect languages Afrikaans (Afrikaans) Albanian (Shqip) Amharic (사ヤでが) Arabic Arabic (ルナルリ) Armenian Armenian (と山) Armenian (と山) Assamese (匈和利利) Azerbaijani (azərbaycan dili) Azerbaijani (azərbaycan dili) (cyrl) Basque (euskara, euskera) Belarusian (беларуская мова) Bengali Bengali (বাংলা) Bosnian (bos Path to the location of the binary. Defaults to the \$PATH location. In fi blank, the \$PATH will be used, but will likely fail with scheduler. Disabled D	Description	Default
Asynchronous indexing CCR Files Extract and index text from supported file types. DCR Every File Attempt to OCR every supported file. Allow file level OCR languages Allow users to change the default languages that will be used to OCR a file. Enabled CCR limit languages Limit the number of languages one can select from this list. Enabled CCR limit languages Limit the number of languages one can select from this list. Auto detect languages Afrikaans (Afrikaans) Albanian (Shqip) Amharic (A***C***) Arabic Ara	Uses command line tools to extract the information from the files based on their MIME types.	Disabled
DCR Files Extract and index text from supported file types. Disabled DCR Every File Attempt to OCR every supported file. Allow file level OCR languages Allow users to change the default languages that will be used to OCR a file. Enabled DCR limit languages Limit the number of languages one can select from this list. Enabled None Allow users to change the default languages that will be used to OCR a file. Image: Auto detect languages Afrikaans Albanian (Shqip) Amharic (♣आट१) Arabic Arabic (◄आ८१) Armenian Belarusian (беларуская мова) Bengali Bengali (चाजा) Bosnian (bos Path to the location of the binary. Defaults to the \$PATH location. In It blank, the \$PATH will be used, but will likely fail with scheduler. Path to the location of the binary. Defaults to the \$PATH location. Just/bin/pdfimages path		Disabled
Attempt to OCR every supported file. Allow file level OCR languages Allow users to change the default languages that will be used to OCR a file. Enabled Disabled Disabled Enabled DCR limit languages Limit the number of languages one can select from this list. Enabled None Auto detect languages Afrikaans Albanian (Shqip) Amharic (አጥርኛ) Arabic Arabic (العربية) Armenian Armenian (Հայերեն) Assamese (আসমীয়া Azerbaijani (azərbaycan dili) Azerbaijani (azərbaycan dili) Azerbaijani (azərbaycan dili) Bosnian (bos Path to the location of the binary. Defaults to the \$PATH location. If blank, the \$PATH will be used, but will likely fail with scheduler. Path to the location of the binary. Defaults to the \$PATH location. /usr/bin/pdfimages path		Enabled
Allow file level OCR languages Allow users to change the default languages that will be used to OCR a file. DCR limit languages Limit the number of languages one can select from this list. Image: Auto detect languages Afrikaans (Afrikaans) Albanian (Shqip) Amharic (क्रण्टाङ) Arabic Arabic (العربية) Armenian Armenian (Հայերեն) Assamese (अल्लीका) Azerbaijani (azərbaycan dili) Azerbaijani (azərbaycan dili) (cyrl) Basque (euskara, euskera) Belarusian (беларуская мова) Bengali Bengali (बाला) Bosnian (bos Path to the location of the binary. Defaults to the \$PATH location. If blank, the \$PATH will be used, but will likely fail with scheduler. Path to the location of the binary. Defaults to the \$PATH location. /usr/bin/pdfimages path	Extract and index text from supported file types.	Disabled
DCR limit languages Limit the number of languages one can select from this list. Auto detect languages Afrikaans (Afrikaans) Albanian (Shqip) Amharic (ਮਿਆਟਨਾ) Arabic Arabic (العربية) Armenian Armenian (حسابات المعالية) Assamese (অসমীয়া) Azerbaijani (azərbaycan dili) Azerbaijani (azərbaycan dili) (cyrl) Basque (euskara, euskera) Belarusian (беларуская мова) Bengali Bengali (বাংলা) Bosnian (bos Path to the location of the binary. Defaults to the \$PATH location. If blank, the \$PATH will be used, but will likely fail with scheduler. pdfimages path Path to the location of the binary. Defaults to the \$PATH location. /usr/bin/pdfimages	Attempt to OCR every supported file.	Disabled
चित्रपार detect languages Afrikaans (Afrikaans) Albanian (Shqip) Amharic (१ পেরেই) Arabic Arabic (।	Allow users to change the default languages that will be used to OCR a file.	Enabled
♣ If blank, the \$PATH will be used, but will likely fail with scheduler. pdfimages path Path to the location of the binary. Defaults to the \$PATH location. /usr/bin/pdfimages	三 Auto detect languages Āfrikaans (Afrikaans) Albanian (Shqip) Amharic (አო৫४) Arabic Arabic (العربية) Armenian Armenian (২шյերեն) Assamese (অসমীয়া) Azerbaijani (аzərbaycan dili) Azerbaijani (аzərbaycan dili) (cyrl) Basque (euskara, euskera) Belarusian (беларуская мова) Bengali Bengali (বাংলা)	None
		/usr/bin/tesseract
an blank, the girthi will be doed, but him men, tan man benedater.	Path to the location of the binary. Defaults to the \$PATH location. If blank, the \$PATH will be used, but will likely fail with scheduler.	/usr/bin/pdfimages
		Path to the location of the binary. Defaults to the \$PATH location. If blank, the \$PATH will be used, but will likely fail with scheduler. Path to the location of the binary. Defaults to the \$PATH location. If blank, the \$PATH will be used, but will likely fail with scheduler. Description Uses command line tools to extract the information from the files based on their MIME types. Parses message/rfc822 types of files (aka eml files) and stores individual email headers and content in search index. Extract and index text from supported file types. Attempt to OCR every supported file. Allow users to change the default languages that will be used to OCR a file. Limit the number of languages one can select from this list. E Auto detect languages Afrikaans (Afrikaans) Albanian (Shqip) Amharic (トマルマン) Arabic (コルリン) Armenian Armenian (Հայերեկ) Assamese (অসবীয়া) Azerbaijani (azərbaycan dili) Azerbaijani (azərbaycan dili) (cyrl) Basque (euskara, euskera) Belarusian (беларуская мова) Bengali Bengali (বारला) Bosnian (bos Path to the location of the binary. Defaults to the \$PATH location. If blank, the \$PATH will be used, but will likely fail with scheduler. Path to the location of the binary. Defaults to the \$PATH location.

Default

Option

Description