

pdftotext-Fehler bei der Archivierung

Einzelne Belege konnten bei einem Kunden nicht archiviert werden. In den Documents-Logs war keine Fehlermeldung zu sehen. Es wurde eine Meldung in den Status der Mappe geschrieben. Die wurde jedoch abgeschnitten. Im EDA werden Archivfehler in dieses Log geschrieben:

...\Documents5\eas\http\logs\error.log

Die komplette Fehlermeldung sieht wie im folgenden Beispiel aus. Entscheidend in dem Text ist der Aufruf der pdftotext.exe und die Fehlermeldung "**Copying of text from this document is not allowed**".

```
[ Wed Aug 31 12:12:28.794794 2022] [ eas: error] [ pid 5760: tid 1160] [ client ::1:58829]
[ mod_eas] D: \\GitLab-
Runner\\builds\\ohPMYjtT\\0\\eas\\eas\\Core\\src\\otris\\eas\\transactions\\Transaction.cpp(68):
Throw in function void __cdecl otris::transactions::Transaction::commit(void) \nDynamic
exception type: struct boost::wrapexcept
<struct otris::transactions::TransactionAbortedException>\nstd::exception::what:
D: \\GitLab-
Runner\\builds\\ohPMYjtT\\0\\eas\\eas\\Core\\src\\otris\\eas\\parser\\ParseCommand.cpp(255):
Throw in function void __cdecl otris::parser::ParseCommand::read(const class std::basic_string
<char,struct std::char_traits<char>,class std::allocator<char> > &)\n
Dynamic exception type: struct boost::wrapexcept<struct
otris::InvalidOperationException>\nstd::exception::what: Unknown exception\n
[struct otris::tag_message * __ptr64] = Unable to filter file!\n
[struct otris::tag_string_value * __ptr64] =
D: /Applications/Documents5/eas/http/../../config/stores/../../filter/pdftotext -enc
UTF-8 E: \\DATA\\DOCUMENTS_DATA\\Archive_RW\\store1\\SP00L\\7fdb951-0cf7-4f7c-b5fa-
196c78945323.tmp -\nPermission
Error: Copying of text from this document is not allowed.
\r\n\r\n[struct otris::tag_inner_exception * __ptr64] =
D: \\GitLab-
Runner\\builds\\ohPMYjtT\\0\\eas\\eas\\Core\\src\\otris\\eas\\parser\\ParseCommand.cpp(255):
Throw in function void __cdecl otris::parser::ParseCommand::read(const class std::basic_string
<char,struct std::char_traits<char>,class std::allocator<char> > &)\n
Dynamic exception type: struct boost::wrapexcept
```

```
<struct otris::InvalidOperationException>\nstd::exception::what:  
Unknown exception\n  
[struct otris::tag_message * __ptr64] = Unable to filter file!\n  
[struct otris::tag_string_value * __ptr64] =  
D: /Applications/Documents5/eas/http/../../config/stores/../../filter/pdftotext -enc  
UTF-8 E: \\DATA\\DOCUMENTS_DATA\\Archive_RW\\store1\\SP00L\\7fdb951-0cf7-4f7c-b5fa-  
196c78945323.tmp -\n  
Permission Error: Copying of text from this document is not allowed.  
\\r\\n\\n\\n[struct otris::tag_message * __ptr64] = Transaction aborted!\n
```

Die **pdftotext.exe** schafft es nicht den PDF-Volltext zu generieren. Die einfachste Lösung ist es bei diesen Belegen auf den PDF-Volltext zu verzichten und die Verarbeitung auch bei Fehlern der **pdftotext.exe** einfach fortzusetzen. Hierzu muss in der ini-Datei zum Store ein passender Filter gesetzt werden. Die ini-Dateien befinden sich hier:

...\\Documents5\\eas\\config\\stores

Standardmäßig ist in der ini bereits ein pdf-Eintrag gesetzt, der wie folgt geändert werden muss.

```
[filter]  
;pdf=utf-8;../../filter/pdftotext -enc UTF-8 %d - ;Standard Windows  
; Lösung Windows:  
pdf=utf-8;suppress-error;../../filter/pdftotext.exe -q -enc UTF-8 %d -  
  
; Lösung Linux:  
;pdf=utf-8;suppress-error;pdftotext -q -enc UTF-8 '%d' - ``
```

Im Anschluss muss der Archiv-Dienst neu gestartet werden.

Revision #1

Created 31 August 2022 12:57:11 by Markus Meisner

Updated 26 September 2022 11:20:56 by Markus Meisner