HANCOM
한글과컴퓨터

DocsConverter SDK Guide

2017. 02

㈜한글과컴퓨터

Table of Contents

1. DocsConverter SDK

1.1. Introduction

DocsConverter SDK provides CLI (Command Line Interface) that handles conversion requests. By calling with a parameter combination, it requests a conversion and receives a result.

1.2. Supported Formats

Write (word processor)
- Input: doc, docx, dot, dotx, rtf
- Output: html, pdf, jpg/png, text

Calc (spreadsheet)
- Input: xls, xlsx, xlt, xltx, xlsxm, cell (Hancom spreadsheet)
- Output: html, pdf, jpg/png, text

Show (presentation)
- Input: ppt, pptx, pps, ppsx, pot, potx, show (Hancom presentation)
- Output: html, pdf, jpg/png, text

Hwp (Hangul word processor)
- Input: hwp, hwpx, hwt, hwdt, hml
- Output: html, pdf, jpg/png, text, hwp

2. Conversion

2.1. Commands

Scripts to execute on Linux and Windows are as follows.

Name	Description
convert.sh	Executive script for Linux OS
convert.bat	Executive script for Windows OS

The execution syntax is as follows.

```
$ convert.sh <input file> <output directory> <output type> [-D<key=value> [-D<key=value>...]]
```

If you execute the following, sample.doc under the ‘samples’ directory is converted to html and

saved under the 'out-dir' directory.

```
$ convert.sh samples/sample.doc out-dir/ html
```

2.2. Parameters

<Required Parameters>

Name	Type	Description
input file	string	Directory path of input files
output directory	string	Directory path of output files
output type	string	Format of output files [html, pdf, image, text] If it is not specified, files are converted to PDF.

2.3. Optional Parameters

You can enter them as -Dkey=value. If it is omitted, the default value will be used.

<Common conversion options>

Name	Type	Description
firstPage	integer	Set the first page of a conversion document. If it is omitted, 1 is used as a default value which indicates the first page. ex) -DfirstPage=3
lastPage	integer	Set the last page of a conversion document. If it is omitted, the index value of the last page is used as a default value. ex) -DlastPage=5
maxPage	integer	Set the maximum number of pages to convert. If it is used with lastPage, maxPage is preferred. If it is omitted, it complies with the setting of firstPage and lastPage.
imgType	string	Convert a file to an image. Optional: jpg (default), png
pageSplit	boolean	Specify whether to generate each page as a separate file when converting to 'html' or 'pdf'. If it is 'true', as many files as the number of pages are generated. Optional: true, false (default)

<Hwp-only options>

Name	Type	Description
requestXmlPath	string	When editing a Hwp file or getting a value, specify an absolute path of XML file including actions to execute. (Refer to 3. Hwp Action Functions)
pdfCompatible	boolean	When converting 'hwp' to 'html', specify to be compatible with phantomJS. phatomJS does not fully support SVG Fill Pattern, so that it allows using pre-render image instead of an image for SVG Fill Pattern while generating 'html'.

2.4. Conversion Examples

<Basic HTML conversion>

```
$ convert.sh sample1.hwp output/ html

- input: sample1.hwp
- output: output/document.html
```

<Generating as many html files as the number of pages specified on the 'pageSplit' option>

```
$ convert.sh sample1.hwp output/ html -DpageSplit=true

- input: sample1.hwp
- output: output/document_0001.html, ... , output/document_nnnn.html
```

<Basic PDF conversion>

```
$ convert.sh sample3.docx output/ pdf

- input: sample3.docx
- output: output/document.pdf
```

<Basic image conversion (generating jpg image)>

```
$ convert.sh sample4.ppt output/ image

- input: sample4.ppt
- output: output/document_0001.jpg, ... , output/document_nnnn.jpg
```

<Generating png image based on the imgType option>

```
$ convert.sh sample4.ppt output/ image -DimgType=png

- input: sample4.ppt
- output: output/document_0001.png, ... , output/document_nnnn.png
```

3. Hwp Action Functions

3.1. hwp-hwp Conversion

hwp-hwp filter of DocsConverter can insert an image or text in a particular cell of a Hwp document, or get cell information of the document.

You can input or output an seal image and information for electronic approval with using this filter.

3.2. Action Functions

Action functions of hwp-hwp filter are as follows. They are created as a xml file and then specified as parameters.

Function	Description
insertImage	Insert an image into a particular cell of a Hwp document.
insertText	Insert text into a particular cell of a Hwp document.
getTextValue	Get text from a particular cell of Hwp document.
getFieldList	Get a list of cell information of Hwp document.

Conversion examples using the action.xml file are as follows.

```
$ convert.sh sample6.hwp output/ hwp -DrequestXmlPath=/path/to/action.xml

- input: sample6.hwp
- xml input: /path/to/action.xml
- output: document.hwp
```

3.3. Action XML Creation

Action: insertImage

```
<action id="insertImage">
  <param id="fieldName">SignImageCell1</param>
  <param id="imagePath">/path/to/image.jpg</param>
  <param id="insertType">0</param>
</action>
```

- fieldName: name of cell where an image is inserted
- imagePath: absolute path of an image

Action: insertText

```
<action id="insertText">
  <param id="fieldName">SignTextCell1</param>
  <param id="inputValue">2010-05-10 10:25</param>
</action>
```

- fieldName: name of a cell where text is inserted
- inputValue: text string to be inserted

Action: getTextValue

```
<action id="getTextValue">
  <param id="fieldName">DocumentNoCell</param>
</action>
```

- fieldName: name of a cell which text is get from

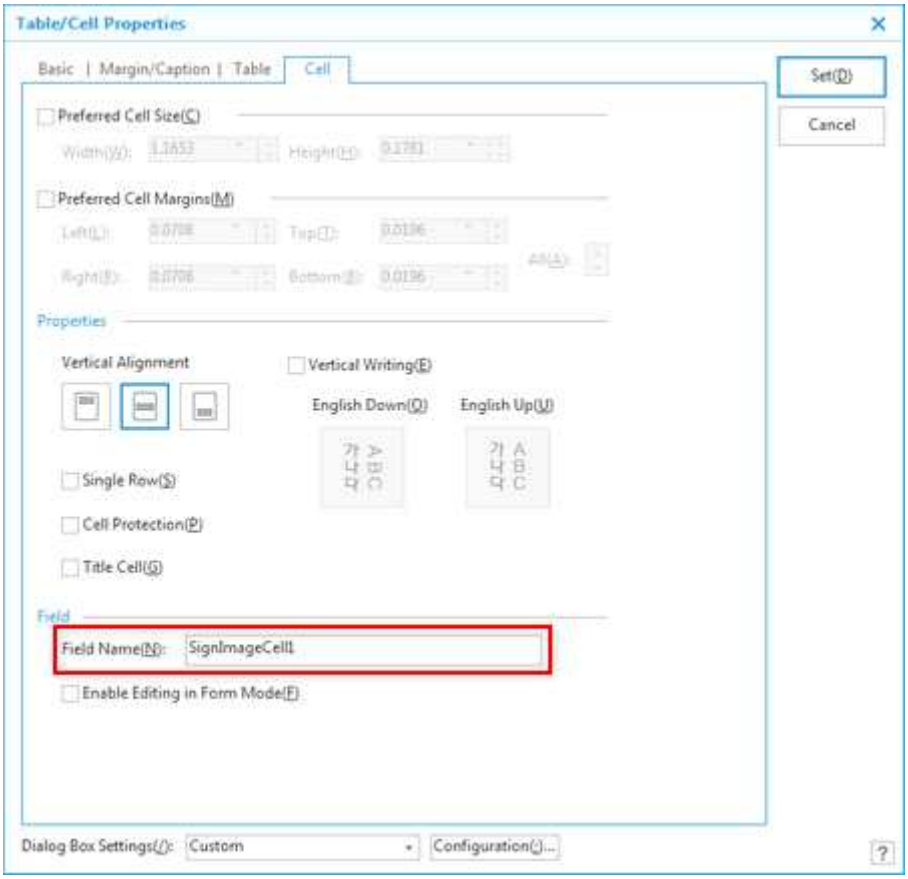
Action: getFieldList

```
<action id="getFieldList">
</action>
```

3.4. Hwp Document Settings

You can set a filed name on the original Hwp document as follows.

3.4.1. Select a cell as a block (F5), and right click and select Table/Cell Properties.



3.4.2. Enter “SignImageCell1” in the Field Name box and click Set.

3.4.3. Once the Hwp document is saved, it is used as a template document in the electronic approval system where you can add an seal image or information.

4. Preparing for installation

4.1. Requirements

DocsConverter SDK can be installed on both Linux and Windows, and your system should meet the following requirements or environment to install it.

4.2. Hardware

- CPU: Intel Xeon 2.6 GHz or faster, 4 or 8 cores or faster (recommended)
- RAM: 8 GB (minimum), 16 GB or more (recommended)

- HDD: 1 GB or more for installation, free space for saving conversion deliverables (recommended)

4.3. Operating System

Linux and Windows are supported, and the following list shows the details.

Linux 64bit
- CentOS 6, CentOS 7
- Ubuntu 12.04, Ubuntu 14.04
- Debian SID

Windows 64bit
- Windows 7, Windows 8, Windows 10
- Windows Server 2008, Windows Server 2012

4.4. Additional Components based on OS

Additional components are required to use the Hwp document conversion filter. They are automatically installed while you are setting up on Linux, but they should be manually installed on Windows.

Linux
- ImageMagick
- Cairo

Windows
- Windows: Microsoft .NET Framework 4.x or higher
- Visual Studio 2013 Redistributable Packages (32bit and 64bit)

4.5. Java

DocsConverter requires JDK, and Open JDK or Oracle JDK is supported here. While installing DocsConverter, a control server 'Tomcat' and a conversion processor 'Processing' are registered as your system services.

- Java: Open JDK 1.7.x or Oracle JDK 1.7.x
- JAVA_HOME environment variable should be set

4.6. Install Files

Two different install files are provided for Linux and Windows as follows. Once you register a

licence after completing installation, you can perform the document conversion.

Install files for Linux

- CentOS 6&7: hermes-sdk-version-centos6-7.0.1.x.zip
- Ubuntu 12: hermes-sdk-version-ubuntu12-7.0.1.x.zip
- Ubuntu 14: hermes-sdk-version-ubuntu14-7.0.1.x.zip

Install file for Windows

- hermes-sdk-version-win64-7.0.1.x.zip

License file

- license.dat

5. Installation

Two types of DocsConverter SDK installers are provided for Linux and Windows. JDK should be installed in advance before trying to install DocsConverter SDK.

5.1. Installation on Linux

This guide assumes that the recent version of CentOS 6.X is installed and package could be updated using yum command.

Install yum update and unzip package as follows.

```
[root@docsconverter]# yum update

[root@docsconverter]# yum install unzip

[root@docsconverter]# cat /etc/redhat-release
CentOS release 6.5 (Final)
```

Install Java 7 and check the version and $JAVA_HOME directory as follows.

```
[root@docsconverter]# java -version
java version "1.7.0_85"
OpenJDK Runtime Environment (rhel-2.6.1.3.el6_7-x86_64 u85-b01)
OpenJDK 64-Bit Server VM (build 24.85-b03, mixed mode)

[root@docsconverter]# echo $JAVA_HOME
/etc/alternatives/jre
```

Upload and unzip the installer, and then locate its directory on the server. Start the installer by

typing "install-hnc.sh" and make sure to have the root authorization. Otherwise, it may cause problems during installation.

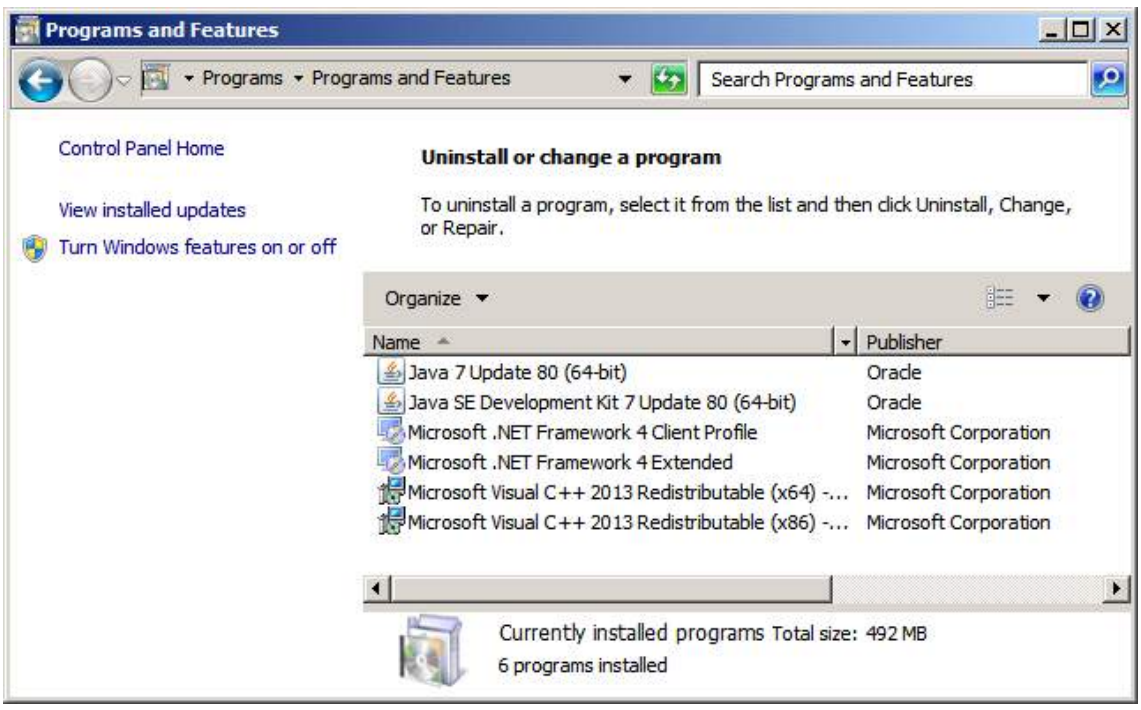
DocsConverter SDK uses open source projects such as ImageMagick, xvfb to support the image-image and pdf-image conversion filter. These are installed together during DocsConverter SDK installation.

5.2. Installation on Windows

Microsoft .NET Framework 4.0 or higher is required. If you currently have a lower version, go to https://www.microsoft.com/en-us/download/details.aspx?id=17851 to download and install the recent version.

Visual Studio 2013 Redistributable Packages are also required. If you want to download and install it, go to https://www.microsoft.com/en-US/download/details.aspx?id=40784. The 64bit version of Windows needs both 32/64bit versions of packages.

The installed programs or packages are shown in the control panel as follows.



Install Java 7 and check the version and $JAVA_HOME directory as follows.

```
C:₩>java -version
java version "1.7.0_67"
Java(TM) SE Runtime Environment (build 1.7.0_67-b01)
Java HotSpot(TM) 64-Bit Server VM (build 24.65-b04, mixed mode)
```

```
C:₩>echo %JAVA_HOME%
C:₩Program Files₩Java₩jdk1.7.0_67
```

Upload and unzip the installer, and then locate its directory on the server. Start the installer by typing "install-hnc.bat" and make sure to run a command prompt as an administrator. Otherwise, it may cause problems during installation.

5.3. Register License

You can register a license file by copying 'license.dat' in the 'conf' directory underneath the DocsConverter SDK install directory.